Ann Inst Stat Math (2016) 68:77–104
DOI 10.1007/s10463-014-0486-5

The complex multinormal distribution, quadratic forms in complex random vectors and an omnibus goodness-of-fit test for the complex normal distribution

Gilles R. Ducharme · Pierre Lafaye de Micheaux · Bastien Marchina

Received: 6 March 2014 / Published online: 21 September 2014

Abstract This paper first reviews some basic properties of the (noncircular) complex multinormal distribution and presents a few characterizations of it. The distribution of linear combinations of complex normally distributed random vectors is then obtained, as well as the behavior of quadratic forms in complex multinormal random vectors. We look into the problem of estimating the complex parameters of the complex normal distribution and give their asymptotic distribution. We then propose a virtually omnibus goodness-of-fit test for the complex normal distribution with unknown parameters, based on the empirical characteristic function. Monte Carlo simulation results show that our test behaves well against various alternative distributions. The test is then applied to an fMRI data set and we show how it can be used to "validate" the usual hypothesis of normality of the outside-brain signal. An R package that contains the functions to perform the test is available from the authors.

Keywords Characteristic function · Complex multinormal distribution · Complex normality test · Complex random vectors · fMRI · Goodness-of-fit test · Independence · Quadratic forms

G. R. Ducharme · B. Marchina
Institut de mathématiques et de modélisation de Montpellier (I3M UMR 5149), cc 051, Université Montpellier II, Place Eugène Bataillon, 34095 Montpellier Cedex 5, France
e-mail: ducharme@math.univ-montp2.fr

B. Marchina
e-mail: marchina@math.univ-montp2.fr

P. Lafaye de Micheaux (✉)
Département de mathématiques et de statistique, Université de Montréal, Pavillon André-Aisenstadt, Montreal, QC H3C 3J7, Canada
e-mail: lafaye@dms.umontreal.ca

1 Introduction

Data in the complex domain are increasingly encountered in many fields, in particular those related to signal processing (see references in Adali et al. 2011), medical imaging (Calhoun et al. 2002; Rowe and Logan 2004; Bernard et al. 2006) and pattern recognition (González-Jiménez and Alba-Castro 2007). However, despite much recent developments on probability models adapted to such data, statistical procedures for analyzing complex data are still trailing in number and scope.

Up to the year 1995, most works on complex random vectors (c.r.v.) pertained to so-called circular, or proper, probability models. In the important case of the complex normal distribution (Goodman 1963), this means that the real and imaginary data are independent real normal random variables with the same variance, so that there was not much need for specific statistical methods. However, since the introduction of the noncircular, or improper, complex normal distribution (van den Bos 1995) and some of its generalizations (Novey et al. 2010), it has emerged that complex data present a challenge of their own and require adapted statistical methods, much as the analysis of directional data requires specific approaches that take into account their particularities.

To this day, the literature on statistics and complex data has mainly focused on exploratory methods and simple modeling techniques where the complex normal distribution has played a prominent role. Inferential procedures (estimation and testing) for complex data are still nascent and there is a paucity of results in the field of complex mathematical statistics.

The goal of this paper is to develop some statistical results for c.r.v., which are then applied to the construction of a virtually omnibus goodness-of-fit test for the noncircular complex normal distribution with unknown parameters. As a consequence, these results are presented with a less utilitarian view than we would have taken if our interest resided only in the goodness-of-fit test.

In Sect. 2 we gather for convenience several basic properties of the complex normal and multinormal distribution as well as a few characterizations. These results remind some well-known facts about real random variables (r.v.) and vectors. A detailed presentation of the distribution of linear combinations of complex normally distributed random vectors is given, as this is a building block of our goodness-of-fit procedure. In Sect. 3, we show that the matrix in a quadratic form in complex random vectors must have a particular form to be statistically useful and prove a general theorem giving its asymptotic behavior. In Sect. 4, we estimate the parameters of the complex normal distribution (which are themselves complex quantities) and obtain their asymptotic distribution. In Sect. 5, we build the goodness-of-fit test for the complex normal distribution. The test is based on the empirical characteristic function, itself a complex quantity. The main reason for this choice is that goodness-of-fit tests based on this function have been recognized as very good for testing the real multinormal distribution, both from theoretical arguments, as in Henze (2002), and from Monte Carlo power studies, as in Mecklin and Mundfrom (2005). First, some properties of the complex empirical characteristic process are reviewed and the general behavior of a test statistic is derived from the results of the previous sections. This is then specialized to the case of the complex normal with known parameters. It is then shown how these results can be adapted to the important case where the parameters of the distribution are unknown and must be estimated. In Sect. 6, a simulation study shows that the test behaves rather well. In Sect. 7, we investigate an example where the assumption of complex normality is important for the modeling process and where our test proves useful in assessing the validity of this assumption. Section 8 concludes the paper. All proofs are provided in the Appendix.

2 Complex normal vectors

2.1 Definitions and general properties

A complex random vector (c.r.v.) Z of $\mathbb{C}^d$ is constructed from a pair $\mathcal{X} = (X_1^\top, X_2^\top)^\top$ (where $^\top$ denotes transposition) of real random vectors (r.v.), each in $\mathbb{R}^d$, as

$$Z = X_1 + \jmath X_2,$$

where $\jmath = \sqrt{-1}$. Thus, $\mathrm{Re}(Z) = X_1,\ \mathrm{Im}(Z) = X_2$. We define the $2d \times 2d$ matrices

$$M = \frac{1}{2}\begin{pmatrix} I_d & I_d \\ -\jmath I_d & \jmath I_d \end{pmatrix} \qquad \text{and} \qquad M^{-1} = \begin{pmatrix} I_d & \jmath I_d \\ I_d & -\jmath I_d \end{pmatrix} = 2M^{\mathcal{H}},$$

where I_d is the identity matrix of order d and whose determinant is $|M| = \jmath^d 2^{-d}$. Throughout the paper, for a vector $\nu \in \mathbb{C}^d$, ν^* will denote its conjugate and $\nu^{\mathcal{H}} = (\nu^*)^\top$ its transpose conjugate. This notation extends to matrices, e.g., $V^{\mathcal{H}}$ for the conjugate transpose of matrix V. The notation $\underset{\sim}{\nu}$ will stand for $(\nu^\top, \nu^{\mathcal{H}})^\top$ and will be referred to as the augmented complex vector associated with ν. In particular, $\underset{\sim}{Z} = (Z^\top, Z^{\mathcal{H}})^\top$ is the augmented complex random vector (a.c.r.v.) associated with Z. We have the relationships

$$\mathcal{X} = M\underset{\sim}{Z}, \quad \underset{\sim}{Z} = M^{-1}\mathcal{X} = 2M^{\mathcal{H}}\mathcal{X}. \tag{1}$$

Thus, the basic probabilistic properties of a c.r.v. Z derive from those of the corresponding real $\mathcal{X}$. For example, if $f_{\mathcal{X}}(x)$ is the density of $\mathcal{X}$ (with respect to the Lebesgue measure on $\mathbb{R}^{2d}$), A is a measurable subset of $\mathbb{C}^d$ and $B = \{x = (x_1^\top, x_2^\top)^\top \in \mathbb{R}^{2d};\ z = x_1 + \jmath x_2 \in A\}$, then

$$\mathbb{P}[Z \in A] = \int_B f_{\mathcal{X}}(x)dx.$$

Also, if the required moments of $\mathcal{X}$ exist, then the expectation of Z is

$$\mu = \mathbb{E}(Z) = \mathbb{E}(X_1) + \jmath\mathbb{E}(X_2),$$

and its covariance matrix is defined by

$$\Gamma = \mathbb{C}\mathrm{ov}(Z) = \mathbb{E}\left[(Z-\mu)(Z-\mu)^{\mathcal{H}}\right].$$

The $d \times d$ matrix Γ is complex, hermitian ($\Gamma = \Gamma^{\mathcal{H}}$) and non-negative definite.

Another quantity arises in relation with c.r.v., the relation matrix

$$P = \mathbb{E}\left[(Z-\mu)(Z-\mu)^{\mathsf{T}}\right],$$

which is needed to fully characterize the second-order moments of Z. The matrix P is complex and symmetric. Note that in early statistical analyses of complex data, this latter matrix was omitted. The main reason for this can be traced back to the introduction of c.r.v. in Wooding (1956) where, from the motivating example he considered, this matrix was implicitly null, which corresponds to what Picinbono (1996) calls (second-order) circularity. But in general applications, this is inadequate and a proper modeling of c.r.v. requires taking into consideration this matrix P.

The Γ and P matrices are related to the covariance matrix $\Sigma_{\mathcal{X}}$ of $\mathcal{X}$ through the so-called covariance-relation matrix

$$\Gamma_P = \begin{pmatrix} \Gamma & P \\ P^{\mathcal{H}} & \Gamma^* \end{pmatrix} = \begin{pmatrix} \Gamma & P \\ P^* & \Gamma^* \end{pmatrix}, \tag{2}$$

by the following relations:

$$\Sigma_{\mathcal{X}} = M\Gamma_P M^{\mathcal{H}}\,, \quad \Gamma_P = M^{-1}\Sigma_{\mathcal{X}}(M^{\mathcal{H}})^{-1} = 4M^{\mathcal{H}}\Sigma_{\mathcal{X}}M, \tag{3}$$

and, when they exist, $\Sigma_{\mathcal{X}}^{-1} = (M^{\mathcal{H}})^{-1}\Gamma_P^{-1}M^{-1}$ and $\Gamma_P^{-1} = M^{\mathcal{H}}\Sigma_{\mathcal{X}}^{-1}M$.

2.2 Some basic facts about the complex multinormal distribution

The particular case of the complex normal distribution with $P = 0$ was introduced by Wooding (1956). The case $P \neq 0$ was first explored by van den Bos (1995) with Γ_P positive definite. In this subsection, we regroup some known properties of this more general complex multinormal distribution and state some new results.

Suppose $\mathcal{X} \sim N_{2d}(\mu_{\mathcal{X}}, \Sigma_{\mathcal{X}})$ is the $2d$-dimensional (real) normal distribution. By the standard characterization property (Mardia et al. 1979, p. 60), we define this as meaning that all linear combinations of the components of $\mathcal{X}$ are normally distributed. In the particular case where $\Sigma_{\mathcal{X}}$ is positive definite, $\mathcal{X}$ has density

$$f_{\mathcal{X}}(x) = \frac{1}{(2\pi)^d\,|\Sigma_{\mathcal{X}}|^{1/2}} \exp\left\{-\frac{1}{2}(x-\mu_{\mathcal{X}})^{\mathsf{T}}\Sigma_{\mathcal{X}}^{-1}(x-\mu_{\mathcal{X}})\right\}.$$

In view of this, a constructive first definition of the complex normal is given below.

Definition 1 Let $\mathcal{X} \sim N_{2d}(\mu_{\mathcal{X}}, \Sigma_{\mathcal{X}})$. Then $Z = X_1 + \jmath X_2$ has the d-dimensional complex normal distribution.

We can write down the density of Z (with respect to the Lebesgue mesure on $\mathbb{C}^d$) when $\Sigma_{\mathcal{X}}$ is positive definite. Because of (1), the density of Z, at each point $z \in \mathbb{C}^d$, satisfies $f_Z(z) = f_{\mathcal{X}}(x)$, where $x = M\underset{\sim}{z}$. Using (3), this leads to

$$f_Z(z) = \frac{1}{(2\pi)^d\,|\Sigma_{\mathcal{X}}|^{1/2}} \exp\left\{-\frac{1}{2}\left((z^{\mathcal{H}}, z^{\mathsf{T}}) - (M^{-1}\mu_{\mathcal{X}})^{\mathcal{H}}\right)\Gamma_P^{-1}\left(\begin{pmatrix} z \\ z^* \end{pmatrix} - M^{-1}\mu_{\mathcal{X}}\right)\right\}.$$

In view of $|\Sigma_{\mathcal{X}}| = |M|\,|\Gamma_P|\,\left|M^{\mathcal{H}}\right| = 2^{-2d}\,|\Gamma_P|$, and upon writing $\mu = \mathbb{E}(Z)$ with $\underset{\sim}{\mu} = M^{-1}\mu_{\mathcal{X}} = (\mu^{\mathsf{T}}, \mu^{\mathcal{H}})^{\mathsf{T}}$, we get

$$f_Z(z) = \frac{1}{\pi^d\,|\Gamma_P|^{1/2}} \exp\left\{-\frac{1}{2}(\underset{\sim}{z} - \underset{\sim}{\mu})^{\mathcal{H}}\Gamma_P^{-1}(\underset{\sim}{z} - \underset{\sim}{\mu})\right\}. \tag{4}$$

Hence, we have a second definition of the complex multinormal distribution, which holds when Γ_P is positive definite.

Definition 2 The c.r.v. $Z \in \mathbb{C}^d$ is said to have a d-dimensional complex multinormal distribution if its density function, with respect to the Lebesgue measure on $\mathbb{C}^d$, has the form (4), where $\mu \in \mathbb{C}^d$, Γ is $d \times d$ hermitian, P is $d \times d$ symmetric and Γ_P is positive definite. This will be noted $Z \sim CN_d(\mu, \Gamma, P)$.

We remark that, when $P = 0$, (4) reduces to

$$f_Z(z) = \frac{1}{\pi^d\,|\Gamma|} \exp\left\{-(z - \mu)^{\mathcal{H}}\Gamma^{-1}(z - \mu)\right\}, \tag{5}$$

which is the circular density of Wooding (1956); see also Picinbono (1996).

We now recall that the characteristic function of a d-dimensional c.r.v. Z is (Andersen et al. 1995, Definitions 1.10 and 1.20), for ν in $\mathbb{C}^d$,

$$\varphi_Z(\nu) = \mathbb{E}\left(\exp(\jmath\,\mathrm{Re}(\nu^{\mathcal{H}}Z))\right) = \mathbb{E}\left(\exp\left(\frac{1}{2}\jmath\,\underset{\sim}{\nu}^{\mathcal{H}}\underset{\sim}{Z}\right)\right) = \varphi_{\underset{\sim}{Z}}(\underset{\sim}{\nu}). \tag{6}$$

When $\mathcal{X} \sim N_{2d}(\mu_{\mathcal{X}}, \Sigma_{\mathcal{X}})$, its characteristic function is (Mardia et al. 1979, p. 61)

$$\varphi_{\mathcal{X}}(t) = \exp\left\{\jmath\,t^{\mathsf{T}}\mu_{\mathcal{X}} - \frac{1}{2}t^{\mathsf{T}}\Sigma_{\mathcal{X}}t\right\},$$

where $t = (t_{(1)}^{\mathsf{T}}, t_{(2)}^{\mathsf{T}})^{\mathsf{T}} \in \mathbb{R}^{2d}$.

Because each point $\nu \in \mathbb{C}^d$ corresponds uniquely to a point $t \in \mathbb{R}^{2d}$ via $t = M\underset{\sim}{\nu}$, the characteristic function of Z is given through $\varphi_Z(\nu) = \varphi_{\mathcal{X}}(t)$. Now, from (2.1), (3) and (2)

$$\begin{aligned}\varphi_Z(\nu) &= \exp\left\{\jmath\,\frac{1}{2}\underset{\sim}{\nu}^{\mathcal{H}}M^{-1}\mu_{\mathcal{X}} - \frac{1}{2}\underset{\sim}{\nu}^{\mathcal{H}}M^{\mathcal{H}}\Sigma_{\mathcal{X}}M\underset{\sim}{\nu}\right\} \\ &= \exp\left\{\jmath\,\mathrm{Re}(\nu^{\mathcal{H}}\mu) - \frac{1}{4}\left(\nu^{\mathcal{H}}\Gamma\nu + \mathrm{Re}(\nu^{\mathcal{H}}P\nu^*)\right)\right\},\end{aligned} \tag{7}$$

because Γ is hermitian. This leads to a third definition.

Definition 3 $Z \sim CN_d(\mu, \Gamma, P)$ if its characteristic function $\varphi_Z(\cdot)$ is given by (7). This holds even if Γ_P is positive semidefinite.

2.3 Linear combinations of complex normal random vectors

Now, let A be a $q \times d$ matrix whose components are complex and consider the c.r.v. $Y = AZ$. Then, from (7),

$$\begin{aligned}\varphi_Y(w) &= \varphi_Z(A^{\mathcal{H}}w)\\ &= \exp\left\{\jmath\,\mathrm{Re}(w^{\mathcal{H}}A\mu) - \frac{1}{4}\left(w^{\mathcal{H}}A\Gamma A^{\mathcal{H}}w + \mathrm{Re}(w^{\mathcal{H}}APA^{\mathsf{T}}w^*)\right)\right\},\end{aligned}$$

from which we conclude from Definition 3 that $Y \sim CN_q(A\mu, A\Gamma A^{\mathcal{H}}, APA^{\mathsf{T}})$. From this follows the fact that any complex linear combination of a *CN* is also *CN*.

Inversely, with $Z \in \mathbb{C}^d$, suppose that for any $\ell \in \mathbb{C}^d\backslash\{0\}$, $\ell^{\mathcal{H}}Z \sim CN_1(\ell^{\mathcal{H}}\mu, \ell^{\mathcal{H}}\Gamma\ell, \ell^{\mathcal{H}}P\ell^*)$. Then, the characteristic function of $\ell^{\mathcal{H}}Z$ at $v \in \mathbb{C}$ is

$$\varphi_{\ell^{\mathcal{H}}Z}(v) = \exp\left\{\jmath\,\mathrm{Re}(v^{\mathcal{H}}\ell^{\mathcal{H}}\mu) - \frac{1}{4}\left(v^{\mathcal{H}}\ell^{\mathcal{H}}\Gamma\ell v + \mathrm{Re}(v^{\mathcal{H}}\ell^{\mathcal{H}}P\ell^* v^*)\right)\right\}.$$

Hence,

$$\varphi_{\ell^{\mathcal{H}}Z}(1) = \exp\left\{\jmath\,\mathrm{Re}(\ell^{\mathcal{H}}\mu) - \frac{1}{4}\left(\ell^{\mathcal{H}}\Gamma\ell + \mathrm{Re}(\ell^{\mathcal{H}}P\ell^*)\right)\right\} = \varphi_Z(\ell),$$

showing that $Z \sim CN_d(\mu, \Gamma, P)$. This leads to a fourth definition of the complex multinormal distribution, which also holds when Γ_P is positive semidefinite.

Definition 4 A d-dimensional c.r.v. Z is said to have a complex multinormal distribution if every complex linear combination of the components of Z has a complex normal distribution. We again write this as $Z \sim CN_d(\mu, \Gamma, P)$, letting the context make precise whether the corresponding Γ_P is positive definite or positive semidefinite.

Now, using (6), (3) and (7), write

$$\varphi_{\underset{\sim}{Z}}(\underset{\sim}{v}) = \exp\left\{\jmath\frac{1}{2}\underset{\sim}{v}^{\mathcal{H}}\underset{\sim}{\mu} - \frac{1}{8}\underset{\sim}{v}^{\mathcal{H}}\Gamma_P\underset{\sim}{v}\right\}, \tag{8}$$

and let A and B be two $q \times d$ complex matrices. Consider

$$A_B\underset{\sim}{Z} = \begin{pmatrix} A & B\\ B^* & A^*\end{pmatrix}\underset{\sim}{Z} = \begin{pmatrix} AZ + BZ^*\\ (AZ + BZ^*)^*\end{pmatrix} = \underset{\sim}{Y},$$

thus forming the c.r.v. $Y = AZ + BZ^*$ and the associated a.c.r.v. $\underset{\sim}{Y}$. Then,

$$\varphi_{\underset{\sim}{Y}}(\underset{\sim}{w}) = \varphi_{\underset{\sim}{Z}}(A_B^{\mathcal{H}}\underset{\sim}{w}) = \exp\left\{\jmath\frac{1}{2}\underset{\sim}{w}^{\mathcal{H}}(A_B\underset{\sim}{\mu}) - \frac{1}{8}\underset{\sim}{w}^{\mathcal{H}}A_B\Gamma_P A_B^{\mathcal{H}}\underset{\sim}{w}\right\}, \tag{9}$$

from which we conclude that Y follows a complex multinormal distribution whose covariance-relation matrix is given by $\Gamma_{P,Y} = A_B\Gamma_P A_B^{\mathcal{H}}$. This suggests extending complex normality to the a.c.r.v. $\underset{\sim}{Z}$. From now on the notation $Z \sim CN_d(\mu, \Gamma, P)$ is equivalent to $\underset{\sim}{Z} \sim \underset{\sim}{CN}_d(\underset{\sim}{\mu}, \Gamma_P)$, whose characteristic function is given in (8). This last notation is convenient to establish in a compact form which parallels with results about real random vectors. For example, (9) shows that

$$A_B\underset{\sim}{Z} \sim \underset{\sim}{CN}_q(A_B\underset{\sim}{\mu}, A_B\Gamma_P A_B^{\mathcal{H}}). \tag{10}$$

Also, if $\underset{\sim}{Z} \sim \underset{\sim}{CN}_d(\underset{\sim}{\mu}, \Gamma_P)$ and Γ_P is invertible, setting $A_B = \Gamma_P^{-1/2}$, where $\Gamma_P^{-1/2}$ is a square root of Γ_P^{-1}, gives

$$\Gamma_P^{-1/2}(\underset{\sim}{Z} - \underset{\sim}{\mu}) \sim \underset{\sim}{CN}_d(\underset{\sim}{0}, I_{2d}).$$

We refer to the $CN_d(0, I_d, 0)$ or the $\underset{\sim}{CN}_d(\underset{\sim}{0}, I_{2d})$ as the standard complex multinormal distribution. The above shows how to reduce a general complex multinormal to its standard form.

We close this section by giving necessary and sufficient conditions ensuring that some components of $Z \sim CN_d(\mu, \Gamma, P)$ are independent.

Theorem 1 *Let $Z = (Z_1, \ldots, Z_d)^{\mathsf{T}} \sim CN_d(\mu, \Gamma, P)$. The components $Z_1, \ldots, Z_d$ are independent if and only if Γ and P are diagonal.*

Corollary 1 *Let $Z \sim CN_{d_1+d_2}(\mu, \Gamma, P)$. Partition Z as $(Z_1^{\mathsf{T}}, Z_2^{\mathsf{T}})^{\mathsf{T}}$ where Z_1 is $d_1 \times 1$, and likewise μ into (μ_1, μ_2), Γ in $\begin{pmatrix} \Gamma_1 & \Gamma_{12} \\ \Gamma_{12}^{\mathcal{H}} & \Gamma_2 \end{pmatrix}$ and similarly for P. Then, Z_1 and Z_2 are independent if and only if $\Gamma_{12} = P_{12} = 0$.*

3 Quadratic forms

We give the stochastic behavior of quadratic forms in the augmented vector $\underset{\sim}{Z}$,

$$Q = \underset{\sim}{Z}^{\mathcal{H}}R\underset{\sim}{Z}, \tag{11}$$

where R is a $2d \times 2d$ complex matrix and $Z \sim \underset{\sim}{CN}_d(\underset{\sim}{\mu}, \Gamma_P)$. This behavior is another building block of our test procedure in Sect. 5. Because this behavior has an interest of its own, we present the result in some generality.

For proper use in statistical procedures, we require Q in (11) to be a real non-negative random variable. This forces R to be an hermitian matrix. Indeed, in view of (1),

$$\underset{\sim}{Z}^{\mathcal{H}}R\underset{\sim}{Z} = \mathcal{X}^{\mathsf{T}}(M^{-1})^{\mathcal{H}}RM^{-1}\mathcal{X} = 2\mathcal{X}^{\mathsf{T}}S\mathcal{X} \tag{12}$$

where $S = MRM^{-1}$ is real and must be a symmetric positive semidefinite matrix. Writing $S = \begin{pmatrix} S_{11} & S_{12} \\ S_{12}^{\mathsf{T}} & S_{22} \end{pmatrix}$ using $d \times d$ blocks, we have

$$\begin{aligned} R &= M^{-1}SM \\ &= \frac{1}{2}\begin{pmatrix} S_{11} + S_{22} + \jmath(S_{12}^{\mathsf{T}} - S_{12}) & S_{11} - S_{22} + \jmath(S_{12} + S_{12}^{\mathsf{T}}) \\ S_{11} - S_{22} - \jmath(S_{12} + S_{12}^{\mathsf{T}}) & S_{11} + S_{22} - \jmath(S_{12}^{\mathsf{T}} - S_{12}) \end{pmatrix} = \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix}. \end{aligned}$$

The symmetry of S forces $R_{11} = R_{22}^{*}$ and $R_{21} = R_{12}^{\mathcal{H}}$. Hence, R must be of the form

$$R = \begin{pmatrix} G & K \\ K^{\mathcal{H}} & G^{*} \end{pmatrix}, \tag{13}$$

where G is $d \times d$ hermitian positive semidefinite and K is $d \times d$ symmetric complex (i.e., $K^{\mathcal{H}} = K^{*}$). Therefore, R is hermitian. The form (13) is also encountered in Eriksson (2010).

Note that previous works on hermitian quadratic forms, starting with Turin (1960), consider expressions of the form $Z^{\mathcal{H}}GZ$, where $Z \sim CN_d(0, \Gamma, 0)$ and G is hermitian positive semidefinite. To relate this to (11), note that

$$\underset{\sim}{Z}^{\mathcal{H}} \begin{pmatrix} G & 0 \\ 0 & G^{*} \end{pmatrix} \underset{\sim}{Z} = 2Z^{\mathcal{H}}GZ.$$

Therefore, former works on hermitian quadratic forms can be seen as a particular case of the more general expression (11).

Theorem 2 *Let* $\underset{\sim}{Z} \sim \underset{\sim}{CN}_d(\underset{\sim}{\mu}, \Gamma_P)$ *and* R *be as in* (13). *Then,*

$$\underset{\sim}{Z}^{\mathcal{H}} R \underset{\sim}{Z} \sim \sum_{k=1}^{q} \alpha_k \chi_1^2\left(\delta_k^2\right) + \sum_{k=q+1}^{2d} 2\tau_k^2,$$

where the $\chi_1^2(\delta_k^2)$ *are independent noncentral* χ_1^2 *random variables with noncentrality parameters* $\delta_k^2 = \frac{2\tau_k^2}{\alpha_k}$, $\tau_k = \frac{1}{2}x_k^{\mathcal{H}} R \underset{\sim}{\mu}$, *with* α_k *and* x_k *being respectively the sorted (in decreasing order) nonzero eigenvalues and eigenvectors of the* $2d \times 2d$ *matrix* $\Gamma_P R$ *of rank* q. *When* $\underset{\sim}{\mu} = \underset{\sim}{0}$, $\tau_k = \delta_k = 0$ *for all* k.

Finally, we show how to build a quadratic form whose distribution will be χ^2.

Theorem 3 *Let* $\underset{\sim}{Z} \sim \underset{\sim}{CN}_d(\underset{\sim}{0}, \Gamma_P)$ *and* Γ_P^{+} *be the Moore–Penrose pseudo-inverse of* Γ_P. *Then,* $Q = \underset{\sim}{Z}^{\mathcal{H}} \Gamma_P^{+} \underset{\sim}{Z} \sim \chi_q^2$, *where* $q = \mathrm{rank}(\Gamma_P)$. *When* Γ_P *is of full rank,* $Q = \underset{\sim}{Z}^{\mathcal{H}} \Gamma_P^{-1} \underset{\sim}{Z} \sim \chi_{2d}^2$.

4 Estimation in the complex multinormal distribution

Let $Z_1, \ldots, Z_n$ be independent copies of Z, a c.r.v. in $\mathbb{C}^d$ with expectation μ, covariance matrix Γ and relation matrix P.

First, for completeness, we state a central limit theorem (C.L.T.) for c.r.v.. See also Ollila and Koivunen (2010, p.106). The symbol $\rightsquigarrow$ represents convergence in distribution.

Theorem 4 *Let $\bar{Z}_n = n^{-1}\sum_{i=1}^n Z_i$. Then,*

$$\sqrt{n}(\underset{\sim}{\bar{Z}}_n - \underset{\sim}{\mu}) \rightsquigarrow \underset{\sim}{CN_d}(\underset{\sim}{0}, \Gamma_P).$$

We also give a simple form of Slutsky's Theorem for c.r.v. useful to prove the corollary below.

Proposition 1 *Let X (resp. $\{X_n\}$) be some (resp. a sequence of) c.r.v. in $\mathbb{C}^d$ and A be an $m \times d$ complex matrix. If $X_n \rightsquigarrow X$, then $AX_n \rightsquigarrow AX$. This is also true for a.c.r.v.*

Corollary 2 *Let A and B be $m \times d$ matrices, $A_B = \begin{pmatrix} A & B \\ B^* & A^* \end{pmatrix}$ and $\bar{Z}_n = n^{-1}\sum_{i=1}^n Z_i$. Then,*

$$\sqrt{n}A_B\left(\underset{\sim}{\bar{Z}}_n - \underset{\sim}{\mu}\right) \rightsquigarrow \underset{\sim}{CN_d}(\underset{\sim}{0}, A_B\Gamma_P A_B^{\mathcal{H}}).$$

Now, we consider the estimation of the parameters of a complex multinormal distribution.

Proposition 2 *Let $Z_1, \ldots, Z_n$ be independent copies of $Z \sim CN_d(\mu, \Gamma, P)$. The method of moments and the maximum likelihood estimators (m.l.e.) of the three parameters coincide. They are given respectively by:*

$$\hat{\mu} = \bar{Z}_n,$$
$$\hat{\Gamma} = \frac{1}{n}\sum_{k=1}^n (Z_k - \bar{Z}_n)(Z_k - \bar{Z}_n)^{\mathcal{H}},$$
$$\hat{P} = \frac{1}{n}\sum_{k=1}^n (Z_k - \bar{Z}_n)(Z_k - \bar{Z}_n)^{\mathsf{T}}.$$

We will need the asymptotic behavior of the m.l.e. For brevity, we consider the case $d = 1$. Write these estimators $\hat{\mu}$, $\hat{\gamma}$ and $\hat{p}$ as obtained from a sample $Z_1, \ldots, Z_n$ of independent copies of $Z \sim CN_1(\mu, \gamma, p)$.

Because $\mathcal{X} = M\underset{\sim}{Z} \sim N_2(\mu_{\mathcal{X}}, \Sigma_{\mathcal{X}})$ with

$$\mu_{\mathcal{X}} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \quad \Sigma_{\mathcal{X}} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix},$$

it is immediate that

$$\begin{pmatrix} \mu \\ \gamma \\ p \\ \mu^* \\ p^* \end{pmatrix} = \begin{pmatrix} 1 & \jmath & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & -1 & 2\jmath \\ 1 & -\jmath & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & -2\jmath \end{pmatrix} \begin{pmatrix} \mu_1 \\ \mu_2 \\ \sigma_1^2 \\ \sigma_2^2 \\ \sigma_{12} \end{pmatrix}. \tag{14}$$

Therefore the m.l.e. of μ, γ, p are linear functions of $\hat{\mu}_1, \hat{\mu}_2, \hat{\sigma}_1^2, \hat{\sigma}_2^2, \hat{\sigma}_{12}$, the standard m.l.e. of the $N_2(\mu_{\mathcal{X}}, \Sigma_{\mathcal{X}})$. Applying the results of Bilodeau and Brenner (1999, Section 13.3.1 and Example 6.4 p. 80), we obtain

$$\sqrt{n}\left(\begin{pmatrix} \hat{\mu}_1 \\ \hat{\mu}_2 \\ \hat{\sigma}_1^2 \\ \hat{\sigma}_2^2 \\ \hat{\sigma}_{12} \end{pmatrix} - \begin{pmatrix} \mu_1 \\ \mu_2 \\ \sigma_1^2 \\ \sigma_2^2 \\ \sigma_{12} \end{pmatrix}\right) \rightsquigarrow N_5(0, W), \tag{15}$$

with

$$W = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & 0 & 0 & 0 \\ \sigma_{12} & \sigma_2^2 & 0 & 0 & 0 \\ 0 & 0 & 2\sigma_1^4 & 2\sigma_{12}^2 & 2\sigma_1^2\sigma_{12} \\ 0 & 0 & 2\sigma_{12}^2 & 2\sigma_2^4 & 2\sigma_2^2\sigma_{12} \\ 0 & 0 & 2\sigma_1^2\sigma_{12} & 2\sigma_2^2\sigma_{12} & \sigma_1^2\sigma_2^2 + \sigma_{12}^2 \end{pmatrix}.$$

Now, from (14), (15) and Proposition 1 applied using a 3×5 diagonal (of 1's) matrix,

$$\sqrt{n}\begin{pmatrix} \hat{\mu} - \mu \\ \hat{\gamma} - \gamma \\ \hat{p} - p \end{pmatrix} \rightsquigarrow CN_3(0, \Gamma_\psi, P_\psi),$$

where $\psi = (\mu, \gamma, p)^\top$ and the elements of Γ_ψ, P_ψ are extracted from W and the matrix in (14) to give

$$\Gamma_\psi = \begin{pmatrix} \gamma & 0 & 0 \\ 0 & \gamma^2 + |p|^2 & 2p^*\gamma \\ 0 & 2p\gamma & 2\gamma^2 \end{pmatrix} \text{ and } P_\psi = \begin{pmatrix} p & 0 & 0 \\ 0 & \gamma^2 + |p|^2 & 2p\gamma \\ 0 & 2p\gamma & 2p^2 \end{pmatrix}.$$

Finally, in Sect. 5, we will need the following. Let

$$\begin{aligned} \theta_R &= (\mathrm{Re}(\mu), \mathrm{Im}(\mu), \gamma, \mathrm{Re}(p), \mathrm{Im}(p))^\top, \\ \hat{\theta}_R &= (\mathrm{Re}(\hat{\mu}), \mathrm{Im}(\hat{\mu}), \hat{\gamma}, \mathrm{Re}(\hat{p}), \mathrm{Im}(\hat{p}))^\top. \end{aligned} \tag{16}$$

Through (14),

$$\theta_R = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} \mu_1 \\ \mu_2 \\ \sigma_1^2 \\ \sigma_2^2 \\ \sigma_{12} \end{pmatrix}.$$

From this and (15), $\sqrt{n}(\hat{\theta}_R - \theta_R) \rightsquigarrow N_5(0, \mathcal{G})$, with

$$\mathcal{G} = \begin{pmatrix} \frac{1}{2}(\gamma + \mathrm{Re}(p)) & \frac{1}{2}(\mathrm{Im}(p)) & 0 & 0 & 0 \\ \frac{1}{2}(\mathrm{Im}(p)) & \frac{1}{2}(\gamma - \mathrm{Re}(p)) & 0 & 0 & 0 \\ 0 & 0 & \gamma^2 + |p|^2 & 2\gamma\,\mathrm{Re}(p) & 2\gamma\,\mathrm{Im}(p) \\ 0 & 0 & 2\gamma\,\mathrm{Re}(p) & \gamma^2 + \mathrm{Re}(p^2) & \mathrm{Im}(p^2) \\ 0 & 0 & 2\gamma\,\mathrm{Im}(p) & \mathrm{Im}(p^2) & \gamma^2 - \mathrm{Re}(p^2) \end{pmatrix}. \tag{17}$$

5 Goodness-of-fit tests

In this section, we use the results of the paper to build a goodness-of-fit test for the complex normal distribution. We focus on the important case $d = 1$. As stated in Sect. 1, for real data, tests of normality based on the empirical characteristic function have been recognized as having good power over broad classes of alternatives. Consequently, we will base our test of complex normality on the empirical characteristic function.

5.1 Goodness-of-fit test based on the characteristic function

Here, we sketch the basic framework that will be adapted to our goodness-of-fit problem. Let $X_1, \ldots, X_n$ be a sample of real random variables with probability law $\mathbb{P}$ and characteristic function $\varphi_X(\cdot)$. The problem is to test $H_0 : \mathbb{P} = \mathbb{P}_0$ for some $\mathbb{P}_0$. If $\varphi_0(\cdot)$ is the characteristic function of $\mathbb{P}_0$, this problem is equivalent to testing $H_0 : \varphi_X(\cdot) = \varphi_0(\cdot)$. Consider the empirical characteristic process

$$U_n(\cdot) = \sqrt{n}(\hat{\varphi}_n(\cdot) - \varphi_0(\cdot)),$$

where $\hat{\varphi}_n(\cdot)$ is the empirical characteristic function

$$\hat{\varphi}_n(t) = \frac{1}{n}\sum_{i=1}^{n} e^{\jmath t X_i}.$$

Let $t_1, \ldots, t_m$ be a collection of points in $\mathbb{R}$ and let $\mathbf{U}_n = (U_n(t_1), \ldots, U_n(t_m))^\mathsf{T}$. If $\Gamma_{P,U}$ denotes the covariance-relation matrix of $\mathbf{U}_n$, the test statistic $\xi_n = \underset{\sim}{\mathbf{U}_n}^{\mathcal{H}} \Gamma_{P,U}^{+} \underset{\sim}{\mathbf{U}_n}$

converges in distribution to a $\chi^2_{\text{rank}(\Gamma_P)}$ from Theorems 4, 3 and the continuous mapping theorem.

5.2 Goodness-of-fit test for the complex normal distribution: the simple hypothesis case

We adapt the above reasoning to the case where $\mathbb{P}_0$ is the complex normal distribution. Let $Z_1, \ldots, Z_n$ be a sample of c.r.v. with characteristic function $\varphi_Z(\cdot)$. Under $H_0 : Z_i \sim CN_1(\mu, \gamma, p)$, where $\mu \in \mathbb{C}$, $\gamma \in \mathbb{R}$ and $p \in \mathbb{C}$ with $|p| < \gamma$, are known constants. Here, $\varphi_0(\cdot)$ is the characteristic function of a $CN_1(\mu, \gamma, p)$, whose expression is given in (7).

The empirical characteristic function is [see (6)],

$$\hat{\varphi}_n(\nu) = \frac{1}{n}\sum_{i=1}^{n} \exp(\jmath\, \mathrm{Re}(\nu^{\mathcal{H}} Z_i)). \tag{18}$$

Because $\varphi_{\mathcal{X}}(t_1, t_2) = \varphi_Z(t_1 + \jmath t_2)$, it follows from Csörgő (1981) that, on every bounded subset $S \subset \mathbb{C}^d$, $\sup_{\nu \in S} |\hat{\varphi}_n(\nu) - \varphi_Z(\nu)| \to 0$ a.s.. Thus, a test based on the complex empirical process $U_n(\cdot) = \sqrt{n}(\hat{\varphi}_n(\cdot) - \varphi_0(\cdot))$ can in principle detect any departures from H_0.

From Csörgő (1981), the covariance function of $U_n(\cdot)$ is

$$C(s, \nu) = \mathbb{E}(U_n(s)U_n(\nu)^*) = \varphi_0(s - \nu) - \varphi_0(s)\varphi_0(\nu)^*$$

and its relation function is

$$P(s, \nu) = \mathbb{E}(U_n(s)U_n(\nu)) = \varphi_0(s + \nu) - \varphi_0(s)\varphi_0(\nu).$$

Choose m points $\nu_1, \ldots, \nu_m$ in $\mathbb{C}$ and consider the complex random vectors $\mathbf{U}_n = (U_n(\nu_1), \ldots, U_n(\nu_m))^{\mathsf{T}}$. Because under H_0, $\mathbb{E}(\hat{\varphi}_n(\nu)) = \varphi_0(\nu)$, Theorem 4 yields under H_0, $\mathbf{U}_n \rightsquigarrow \mathbf{U} \sim CN_m(0, \Gamma_U, P_U)$, where

$$\Gamma_U = \mathbb{E}(\mathbf{U}\mathbf{U}^{\mathcal{H}}) = (C(\nu_k, \nu_{k'}))_{k,k'=1,\ldots,m}, \tag{19}$$

and

$$P_U = \mathbb{E}(\mathbf{U}\mathbf{U}^{\mathsf{T}}) = (P(\nu_k, \nu_{k'}))_{k,k'=1,\ldots,m}. \tag{20}$$

Therefore, in the same manner as in Subsect. 5.1,

$$\xi_n = \underset{\sim}{\mathbf{U}_n}^{\mathcal{H}} \Gamma^{+}_{P,U} \underset{\sim}{\mathbf{U}_n} \rightsquigarrow \chi^2_{\text{rank}(\Gamma_{P,U})},$$

where $\Gamma_{P,U} = \begin{pmatrix} \Gamma_U & P_U \\ P_U^{\mathcal{H}} & \Gamma_U^* \end{pmatrix}$.

Note that when H_0 is false and if at the point ν_k, $\varphi_Z(\nu_k) - \varphi_0(\nu_k) = \mu_k \neq 0$, then, from Theorem 4, for some $\gamma_Z(\nu_k)$, $p_Z(\nu_k)$,

$$U_n(\nu_k) = CN_1(0, \gamma_Z(\nu_k), p_Z(\nu_k)) + \sqrt{n}\mu_k + o_p(1) \to \infty,$$

so that the test based on ξ_n is consistent for such a ν_k. When H_0 is false, there exists some points for which $\varphi_Z(\nu_k) \neq \varphi_0(\nu_k)$ and, thus, our test becomes virtually omnibus as m increases. A variant where $\varphi_Z(\nu_k) - \varphi_0(\nu_k) = n^{-1/2}\mu_k$ (contiguous alternatives) could be developed, but because it is of a theoretical interest only, this will not be pursued here.

5.3 Goodness-of-fit test for the complex normal distribution: the composite case

In most practical applications, the parameters of the complex normal distribution are unknown. The problem becomes that of testing the composite hypothesis $H_0 : \mathbb{P} \in \{CN_1(\mu, \gamma, p)\ ;\ \mu \in \mathbb{C}, \gamma \in \mathbb{R}, p \in \mathbb{C} \text{ with } |p| < \gamma\}$.

With $Z_1, \ldots, Z_n$ being independent copies of Z, define $\underset{\sim}{Y}_k = \gamma_p^{-1/2}(\underset{\sim}{Z}_k - \underset{\sim}{\mu})$, where

$$\gamma_p = \begin{pmatrix} \gamma & p \\ p^* & \gamma \end{pmatrix},$$

and $\hat{\underset{\sim}{Y}}_k = \hat{\gamma}_p^{-1/2}(\underset{\sim}{Z}_k - \hat{\underset{\sim}{\mu}})$, using the m.l.e. of Proposition 2. Again, let $\nu_1, \ldots, \nu_m$ be given points in $\mathbb{C}$ and consider $\hat{\mathbf{U}}_n = (\hat{U}_n(\nu_1), \ldots, \hat{U}_n(\nu_m))^\mathsf{T}$, where $\hat{U}_n(\nu) = \sqrt{n}(\varphi_{n,\hat{Y}}(\nu) - \varphi_0(\nu))$, $\varphi_0(\cdot)$ is the characteristic function of a $CN_1(0, 1, 0)$ c.r.v. and $\varphi_{n,\hat{Y}}(\nu)$ is the empirical characteristic function given in (18) computed from the $\hat{Y}_k = (1, 0)\hat{\underset{\sim}{Y}}_k$. Thus, here $\theta_R = (0, 0, 1, 0, 0)^\mathsf{T}$.

The Taylor expansion of $\varphi_{n,\hat{Y}}(\nu)$ as a function of $\hat{\theta}_R$ about θ_R gives

$$\sqrt{n}(\varphi_{n,\hat{Y}}(\nu) - \varphi_{n,Y}(\nu)) = (\nabla_{\theta_R}\varphi_{n,Y}(\nu))^\mathsf{T}\sqrt{n}(\hat{\theta}_R - \theta_R) + o_p(1),$$

where $\nabla_{\theta_R} = \left(\frac{\partial}{\partial\,\mathrm{Re}(\mu)}, \frac{\partial}{\partial\,\mathrm{Im}(\mu)}, \frac{\partial}{\partial\gamma}, \frac{\partial}{\partial\,\mathrm{Re}(p)}, \frac{\partial}{\partial\,\mathrm{Im}(p)}\right)^\mathsf{T}$ is the gradient operator evaluated at θ_R. The law of large numbers implies that $\nabla_{\theta_R}\varphi_{n,Y}(\nu)$ converges in probability to a certain point $J(\nu)$ that will be exhibited below. It follows that $\hat{U}_n(\nu) = U_n(\nu) + J(\nu)^\mathsf{T}\hat{\Theta}_n + o_p(1)$, where $\hat{\Theta}_n = \sqrt{n}(\hat{\theta}_R - \theta_R)$. Thus,

$$\hat{\mathbf{U}}_n = \left(I_m\ ,\ J_{\boldsymbol{\nu}}^\mathsf{T}\right)\begin{pmatrix}\mathbf{U}_n \\ \hat{\Theta}_n\end{pmatrix} + o_p(1),$$

where $J_{\boldsymbol{\nu}} = \left(J(\nu_1), \ldots, J(\nu_m)\right)$. Classical results in m.l.e. theory ensure that $\hat{\Theta}_n = \frac{1}{\sqrt{n}}\sum_{k=1}^{n} l(Z_k) + o_p(1)$, where $l(Z_k) = \mathcal{G}(\nabla_{\theta_R}(\log(f_0(Z_k, \theta_R)))$. Here, $f_0(z, \theta_R)$

denotes the density of the complex normal distribution given in (4) parametrized with θ_R, and $\mathcal{G}$ is given in (17). Straightforward algebra yields

$$l(Z_k)=\begin{pmatrix} \mathrm{Re}(Z_k)-\mathrm{Re}(\mu) \\ \mathrm{Im}(Z_k)-\mathrm{Im}(\mu) \\ -\gamma+(\mathrm{Re}(Z_k)-\mathrm{Re}(\mu))^2+(\mathrm{Im}(Z_k)-\mathrm{Im}(\mu))^2 \\ -\mathrm{Re}(p)+(\mathrm{Re}(Z_k)+\mathrm{Im}(Z_k)-\mathrm{Re}(\mu)-\mathrm{Im}(\mu))(\mathrm{Re}(Z_k)-\mathrm{Im}(Z_k)-\mathrm{Re}(\mu)+\mathrm{Im}(\mu)) \\ -\mathrm{Im}(p)+2(\mathrm{Re}(Z_k)-\mathrm{Re}(\mu))(\mathrm{Im}(Z_k)-\mathrm{Im}(\mu)) \end{pmatrix}.$$

We now obtain the asymptotic distribution of $\hat{\mathbf{U}}_n$ under H_0.

Theorem 5 *Let $Z_1,\ldots,Z_n$ be independent copies of $Z \sim CN_1(\mu,\gamma,p)$. Let $\nu_1,\ldots,\nu_m \in \mathbb{C}$. Then,*

$$\begin{pmatrix}\mathbf{U}_n \\ \hat{\Theta}_n\end{pmatrix} \rightsquigarrow CN_{m+5}\left(0, \begin{pmatrix}\Gamma_U & C_{\boldsymbol{\nu}} \\ C_{\boldsymbol{\nu}}^{\mathcal{H}} & \mathcal{G}\end{pmatrix}, \begin{pmatrix}P_U & C_{\boldsymbol{\nu}} \\ C_{\boldsymbol{\nu}}^{\top} & \mathcal{G}\end{pmatrix}\right),$$

where Γ_U, P_U are given in (19), (20), $\mathcal{G}$ is given in (17) and $C_{\boldsymbol{\nu}}$ is a complex matrix that will be detailed below. As a result,

$$\hat{\mathbf{U}}_n \rightsquigarrow CN_m(0, \Gamma(\boldsymbol{\nu}), P(\boldsymbol{\nu})), \tag{21}$$

with

$$\Gamma(\boldsymbol{\nu}) = \Gamma_U + C_{\boldsymbol{\nu}} J_{\boldsymbol{\nu}}^* + J_{\boldsymbol{\nu}}^{\top} C_{\boldsymbol{\nu}}^{\mathcal{H}} + J_{\boldsymbol{\nu}}^{\top}\mathcal{G} J_{\boldsymbol{\nu}}^*, \tag{22}$$

$$P(\boldsymbol{\nu}) = P_U + C_{\boldsymbol{\nu}} J_{\boldsymbol{\nu}} + J_{\boldsymbol{\nu}}^{\top} C_{\boldsymbol{\nu}}^{\top} + J_{\boldsymbol{\nu}}^{\top}\mathcal{G} J_{\boldsymbol{\nu}}. \tag{23}$$

Proof Consider the real random vector

$$\begin{pmatrix}\mathrm{Re}(\mathbf{U}_n) \\ \hat{\Theta}_n \\ \mathrm{Im}(\mathbf{U}_n)\end{pmatrix} = \sqrt{n}\left(\frac{1}{n}\begin{pmatrix}\mathrm{Re}(\boldsymbol{\varphi}_n)-\mathbb{E}(\mathrm{Re}(\boldsymbol{\varphi}_n)) \\ \sum_{k=1}^n l(Z_k) \\ \mathrm{Im}(\boldsymbol{\varphi}_n)-\mathbb{E}(\mathrm{Im}(\boldsymbol{\varphi}_n))\end{pmatrix}\right) + o_p(1), \tag{24}$$

where $\boldsymbol{\varphi}_n = (\varphi_{n,Y}(\nu_1),\ldots,\varphi_{n,Y}(\nu_m))^{\top}$. The C.L.T. ensures that (24) converges to a normal random vector. Therefore, $\mathbf{U}_{\theta,n} = (\mathrm{Re}(\mathbf{U}_n), \hat{\Theta}_n, \mathrm{Im}(\mathbf{U}_n), \mathbf{0}_{5\times 1})^{\top}$ also converges to a (degenerate) normal random vector. Because

$$M^{-1}\mathbf{U}_{\theta,n} = \underset{\sim}{\begin{pmatrix}\mathbf{U}_n \\ \hat{\Theta}_n\end{pmatrix}},$$

where M^{-1} is the $2(m+5)\times 2(m+5)$ matrix given in (2.1), it follows that

$$\begin{pmatrix}\mathbf{U}_n \\ \hat{\Theta}_n\end{pmatrix} \rightsquigarrow CN_{m+5}\left(0, \begin{pmatrix}\Gamma_U & C_{\boldsymbol{\nu}} \\ C_{\boldsymbol{\nu}}^{\mathcal{H}} & \mathcal{G}\end{pmatrix}, \begin{pmatrix}P_U & C_{\boldsymbol{\nu}} \\ C_{\boldsymbol{\nu}}^{\top} & \mathcal{G}\end{pmatrix}\right),$$

in which $C_{\boldsymbol{\nu}} = \mathbb{E}(\mathbf{U}_n \hat{\Theta}_n^{\top}) = (C(\nu_1), \ldots, C(\nu_m))^{\top}$, with

$$C(\nu) = \mathbb{E}\left(U_n(\nu) \times \frac{1}{\sqrt{n}} \sum_{k=1}^{n} l(Z_k) \right).$$

The asymptotic behavior of $\hat{\mathbf{U}}_n$ follows from (10). □

The test statistic for $H_0 : Z \sim CN_1(\mu, \gamma, p),\ \mu,\ \gamma,\ p$ unknown, is

$$\hat{\xi}_n(\boldsymbol{\nu}) = \hat{\underset{\sim}{\mathbf{U}}}_n^{\mathcal{H}} \Gamma_P^{+}(\boldsymbol{\nu}) \hat{\underset{\sim}{\mathbf{U}}}_n, \tag{25}$$

where $\Gamma_P(\boldsymbol{\nu})$ is the covariance-relation matrix of $\hat{\mathbf{U}}_n$ whose components are given in (22) and (23). Under H_0, by Theorem 3, $\hat{\xi}_n(\boldsymbol{\nu}) \rightsquigarrow \chi^2_{\mathrm{rank}(\Gamma_P(\boldsymbol{\nu}))}$. When H_0 is false and at least one point ν_k is such that $\varphi_Z(\nu_k) \neq \varphi_0(\nu_k)$, then $\hat{\xi}_n(\boldsymbol{\nu}) \to \infty$ and the test is consistent.

It turns out that the components of $\Gamma_P(\boldsymbol{\nu})$ do not depend on unknown parameters that would then need to be estimated. This interesting property follows from the fact that $\Gamma_P(\boldsymbol{\nu})$ is evaluated at parameter space point $\theta_R = (0, 0, 1, 0, 0)^{\top}$. Then,

$$\mathcal{G} = \begin{pmatrix} 1/2 & 0 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

and, for each point ν_k,

$$C(\nu_k) = \frac{1}{4} \begin{pmatrix} 2\jmath \,\mathrm{Re}(\nu_k) \\ 2\jmath \,\mathrm{Im}(\nu_k) \\ -|\nu_k|^2 \\ -\mathrm{Re}(\nu_k)^2 + \mathrm{Im}(\nu_k)^2 \\ -2\,\mathrm{Re}(\nu_k)\,\mathrm{Im}(\nu_k) \end{pmatrix} \times \varphi_0(\nu_k)$$

and

$$J(\nu_k) = \begin{pmatrix} -\jmath \,\mathrm{Re}(\nu_k) \\ -\jmath \,\mathrm{Im}(\nu_k) \\ \frac{1}{4}|\nu_k|^2 \\ \frac{1}{4}(\mathrm{Re}(\nu_k)^2 - \mathrm{Im}(\nu_k)^2) \\ \frac{1}{2}\,\mathrm{Re}(\nu_k)\,\mathrm{Im}(\nu_k) \end{pmatrix} \times \varphi_0(\nu_k).$$

6 Simulations

To investigate the behavior of our test statistic (25), a simulation study is conducted to assess its level and power under various settings of n and $(\nu_1, \ldots, \nu_m)$. The first part of our simulation explores the actual levels of $\hat{\xi}_n(\boldsymbol{\nu})$ to determine whether the asymptotic χ^2 distribution provides a good approximation in moderate samples. The number of points in $\boldsymbol{\nu}$ ranges from $m = 1$ to 6 and we test two different patterns : points located either on a circle or on a line in $\mathbb{C}$. In all cases, we choose points neither too close (because $U_n(0) \equiv 1$) nor too far from the origin (because the convergence of $\varphi_n(\cdot)$ occurs on bounded sets in $\mathbb{C}$). Finally, for ease of computations, the points are taken such that $\Gamma_P^+(\nu) = \Gamma_P^{-1}(\nu)$. This leads to the following sets of points :

- $m = 1 : \nu_{1,1} = 0.50 + 0.50\jmath$
- $m = 1 : \nu_{1,2} = 5\nu_{1,1}$
- $m = 2$, circular pattern:

$$\boldsymbol{\nu}_{2,1} = (0.25 + 0.433\jmath, -0.129 + 0.483\jmath)^{\mathsf{T}}$$

- $m = 2$, circular pattern:

$$\boldsymbol{\nu}_{2,2} = 5\boldsymbol{\nu}_{2,1}$$

- $m = 2$, linear pattern:

$$\boldsymbol{\nu}_{2,3} = (0.79 + 0.13\jmath, 0.02 + 0.33\jmath)^{\mathsf{T}},$$

- $m = 3$, circular pattern:

$$\boldsymbol{\nu}_{3,1} = (0.935 - 1.173\jmath, 0.935 + 1.173\jmath, -1.351 + 0.650\jmath)^{\mathsf{T}},$$

- $m = 3$, linear pattern:

$$\boldsymbol{\nu}_{3,2} = (0.5 + 0.5\jmath, 1 + 1\jmath, 1.5 + 1.5\jmath)^{\mathsf{T}}$$

- $m = 4$, circular pattern:

$$\boldsymbol{\nu}_4 = 1.5(\exp(\jmath), \exp(2\jmath), \exp(3\jmath), \exp(4\jmath))^{\mathsf{T}}$$

- $m = 5$, circular pattern:

$$\boldsymbol{\nu}_5 = 1.5(\exp(\jmath), \exp(2\jmath), \exp(3\jmath), \exp(4\jmath), \exp(5\jmath))^{\mathsf{T}}$$

- $m = 6$, circular pattern:

$$\boldsymbol{\nu}_6 = 1.5(\exp(\jmath), \exp(2\jmath), \exp(3\jmath), \exp(4\jmath), \exp(5\jmath), \exp(6\jmath))^{\mathsf{T}}$$

Table 1 gives the levels of $\hat{\xi}_n(\boldsymbol{\nu})$ for $n = 50, 100, 250$ and 500 and for the above choices of points. It appears that the test using the asymptotic approximation holds its level. In general, for the above sample sizes we may recommend, level-wise, the use of a circular pattern with $m = 3$ or 4.

Table 1 Percentage points for the observed distribution of $\hat{\xi}_n(\boldsymbol{\nu})$ based on 100,000 repetitions for $\nu_{1,1}$, $\nu_{1,2}$, $\boldsymbol{\nu}_{2,1}$, $\boldsymbol{\nu}_{2,2}$ and $\boldsymbol{\nu}_{2,3}$

n	$\nu_{1,1}$	$\nu_{1,2}$	$\boldsymbol{\nu}_{2,1}$	$\boldsymbol{\nu}_{2,2}$	$\boldsymbol{\nu}_{2,3}$	$\boldsymbol{\nu}_{3,1}$	$\boldsymbol{\nu}_{\mathbf{3,2}}$	$\boldsymbol{\nu}_4$	$\boldsymbol{\nu}_5$	$\boldsymbol{\nu}_6$
level	5 %	5 %	5 %	5 %	5 %	5 %	5 %	5 %	5 %	5 %
50	0.039	0.051	0.048	0.046	0.046	0.048	0.032	0.050	0.055	0.062
100	0.042	0.051	0.052	0.047	0.050	0.049	0.038	0.051	0.055	0.062
250	0.045	0.050	0.053	0.049	0.052	0.050	0.046	0.052	0.055	0.060
500	0.047	0.050	0.053	0.050	0.051	0.050	0.050	0.050	0.052	0.056

The second part of our simulation pertains to the power of our test procedure. We apply our test to data generated from a number of alternative distributions, listed below that cover some interesting departures from the CN distribution.

- Khintchine distributions $Kh(a)$: let $W = X + \jmath Y$ where X, Y are drawn from a $U[0, 1]$ distribution. Let $G \sim Gamma(1.5, 1)$. A datum from a complex Khintchine distribution with parameter a is generated through

$$\sqrt{\frac{3\Gamma(1.5)}{\Gamma(1.5+2a)}}G^a(2W - 1 - 1\jmath).$$

  We consider the cases $a = 0.5$ and $a = 0.3989$, for which the distribution has almost normal marginals on its real and imaginary parts, but the level curves of the density are rectangular in $\mathbb{C}$.
- Mixture distributions: let $W \sim CN_1(0, 2, 0)$, $Y \sim CN_1(3 + 3\jmath, 2, 0)$ and $X \sim CN_1(0, 10, 1.8\jmath)$. We consider the following mixture models, $pW + (1 - p)Y$ (location mixture) with $p = 0.5, 0.7$ and 0.9 and $qW + (1 - q)X$ (scale mixture) with $q = 0.5$ and $q = 0.9$.
- Complex Student distribution $Ct(d)$: let $Z \sim CN_1(0, 2, 0)$, $Y \sim \chi^2_d$ and $R = \sqrt{d/Y}$. We say that $X = RZ$ follows a complex Student distribution. We consider the cases $d = 3$, $d = 10$ and $d = 15$.
- Complex B–P logistic CBPL: let $G \sim \text{Gamma}(1, 1)$, $Y_1 \sim \text{Exp}(1)$ and $Y_2 \sim \text{Exp}(1)$ be independent. Let $U_1 = \left(1 + \frac{Y_1}{G}\right)^{-1}$, $U_2 = \left(1 + \frac{Y_2}{G}\right)^{-1}$ and $\Phi^{-1}(\cdot)$ be the quantile function of an $N(0, 1)$. Then, $Z = \Phi^{-1}(U_1) + \jmath\Phi^{-1}(U_2)$ is said to follow a complex B–P logistic distribution, whose density tends to have triangular levels in $\mathbb{C}$.
- Complex Pearson type II distribution $CP2(r)$: let $U \sim U[0, 2\pi]$, $Y = \cos(U) + \jmath\sin(U)$ and $B \sim \beta(1, 1 + r)$; then $W = BY$ is said to follow a complex Pearson type II distribution. The density of this distribution is supported in the unit circle in $\mathbb{C}$ and becomes sharper at the origin as r increases. We consider the cases $r = 5$ and $r = 10$.
- Laplace-type distributions: let $U, W, W \sim \text{Exp}(1)$ be independent. Then $X = V - U$, $Y = W - U$ and $Z = X + \jmath Y$.
- Independent exponential: let $X, Y \sim \text{Exp}(1)$ be independent. Then, $Z = X + \jmath Y$.

Table 2 Power of $\hat{\xi}_n(\boldsymbol{\nu}_{3,1})$, based on 100,000 Monte Carlo repetitions and using the empirical critical points

n	50	100	250	500
significance level	5 %	5 %	5 %	5 %
Kh(0.5)	0.269	0.459	0.829	0.988
Kh(0.3989)	0.270	0.459	0.828	0.988
Location mixture 1	0.163	0.822	1.000	1.000
Location mixture 2	0.356	0.942	1.000	1.000
Location mixture 3	0.860	0.996	1.000	1.000
Scale mixture 1	0.484	0.756	0.986	1.000
Scale mixture 2	0.508	0.742	0.960	0.998
Ct(3)	0.862	0.985	1.000	1.000
Ct(10)	0.273	0.420	0.698	0.915
Ct(15)	0.173	0.255	0.422	0.636
CBPL	0.258	0.536	0.947	1.000
CP2(5)	0.855	0.990	1.000	1.000
CP2(10)	0.923	0.998	1.000	1.000
Lap	0.922	0.998	1.000	1.000
Exp	0.999	1.000	1.000	1.000
Lognorm	0.808	0.991	1.000	1.000
χ^2	0.751	0.980	1.000	1.000

- Complex log-normal (0.5): let $U, V, W \sim \text{LogNormal}(0, 0.25)$ be independent. Then, $X = UV$, $Y = UW$ and $Z = X + \jmath Y$.
- Complex $\chi^2(8)$: let $U, V, W \sim \text{Gamma}(2, 0.5)$ be independent. Then $X = U + V$, $Y = U + W$ and $Z = X + \jmath Y$.

Our simulation study is conducted using the points $\boldsymbol{\nu}_{3,1}$, $\boldsymbol{\nu}_{3,2}$ and $\boldsymbol{\nu}_4$ which yields consistent tests. To give a proper reflection of the power of our test, we use the empirical quantiles obtained from the first part of our simulation. The results are summarized in Tables 2, 3 and 4, associated with $\boldsymbol{\nu}_{3,1}$, $\boldsymbol{\nu}_{3,2}$ and $\boldsymbol{\nu}_4$, respectively. The test has rather good power against all alternatives.

Note that $\boldsymbol{\nu}_{3,1}$, a circular pattern, yields globally the most powerful test and as default values we may recommend using these points. With $\boldsymbol{\nu}_4$, another circular pattern, the power is globally slightly weaker but better than with $\boldsymbol{\nu}_{3,2}$, a linear pattern. Complementary power studies (not shown) have established that the behavior of $\hat{\xi}_n(\boldsymbol{\nu})$ is rather stable among the circular patterns (from $m = 3$ to $m = 6$), and that the power of the tests is lower if $m = 1$ or 2. This is intuitively easy to understand.

7 Example

Functional magnetic resonance imaging (fMRI) is a neuro imaging procedure that measures brain activity in real time. Data are acquired as natively complex (Calhoun

Table 3 Power of $\hat{\xi}_n(\boldsymbol{\nu}_{3,2})$, based on 100,000 Monte Carlo repetitions and using the empirical critical points

n	50	100	250	500
significance level	5 %	5 %	5 %	5 %
Kh(0.5)	0.249	0.381	0.669	0.919
Kh(0.3989)	0.248	0.378	0.671	0.919
Location mixture 1	0.917	1.000	1.000	1.000
Location mixture 2	0.944	1.000	1.000	1.000
Location mixture 3	0.937	0.999	1.000	1.000
Scale mixture 1	0.393	0.627	0.937	0.999
Scale mixture 2	0.411	0.642	0.910	0.991
Ct(3)	0.650	0.888	0.997	1.000
Ct(10)	0.184	0.282	0.484	0.700
Ct(15)	0.127	0.185	0.297	0.439
CBPL	0.107	0.161	0.323	0.616
CP2(5)	0.696	0.930	1.000	1.000
CP2(10)	0.782	0.968	1.000	1.000
Lap	0.538	0.792	0.987	1.000
Exp	0.775	0.989	1.000	1.000
Lognorm	0.430	0.744	0.992	1.000
χ^2	0.454	0.774	0.996	1.000

Table 4 Power of $\hat{\xi}_n(\boldsymbol{\nu}_4)$, based on 100,000 Monte Carlo repetitions and using the empirical critical points

n	50	100	250	500
significance level	5 %	5 %	5 %	5 %
Kh(0.5)	0.221	0.381	0.748	0.969
Kh(0.3989)	0.223	0.381	0.748	0.969
Location mixture 1	0.085	0.637	1.000	1.000
Location mixture 2	0.246	0.873	1.000	1.000
Location mixture 3	0.849	0.996	1.000	1.000
Scale mixture 1	0.522	0.797	0.993	1.000
Scale mixture 2	0.537	0.767	0.968	0.999
Ct(3)	0.881	0.989	1.000	1.000
Ct(10)	0.289	0.444	0.728	0.935
Ct(15)	0.183	0.271	0.445	0.673
CBPL	0.211	0.474	0.928	0.999
CP2(5)	0.887	0.995	1.000	1.000
CP2(10)	0.944	0.999	1.000	1.000
Lap	0.928	0.999	1.000	1.000
Exp	0.999	1.000	1.000	1.000
Lognorm	0.785	0.986	1.000	1.000
χ^2	0.711	0.969	1.000	1.000

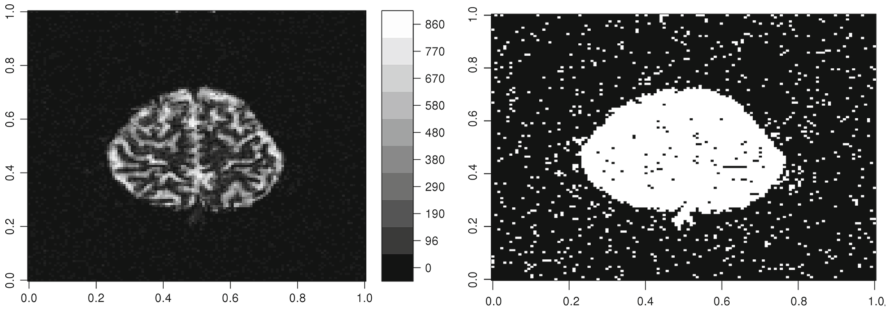

Fig. 1 Heat map of the fMRI dataset, first with respect to the values of $\hat{\xi}_n(\boldsymbol{\nu}_{3,1})$ (*left*), then versus the 95 %-percentile of the χ^2_6 distribution (*right*)

and Adali 2012), but for various reasons only the so-called magnitude (given by ZZ^*) part of the data is typically used in the ensuing analysis. Over the past few years, interest in exploiting the full information has been growing and new methods for modeling and processing the whole data have been developed. In particular, Rowe and Logan (2004) present an activation model for complex fMRI data. Their dataset is structured as a 128 × 128 array of volume elements (called voxels) representing spatial positions in a given slice of the brain. In each of these voxels, 269 complex fMRI observations are collected over time, while the subject performs a finger-tapping task. Each observation can be viewed as a (indirect) measure of the neuronal activity in this small voxel region and at a given time. However, some of these voxels are located outside the brain tissue and, in their model, Rowe and Logan (2004) make the assumption that the corresponding 269 data points follow a complex normal distribution, while data pertaining to voxels lying in brain tissue behave differently. Their model uses this difference to allow discriminating between the different spatial positions of the voxels.

Lacking a formal way to validate this complex normality assumption, they resort to showing histograms of the real and imaginary part of the data which seem to corroborate complex normality. However, note that even if such histograms are consonant with normal distributions for each of the parts, this is not sufficient to get complex normality : there exists complex probability distributions (e.g., the complex Khintchine distribution), whose marginal distributions, in the real and imaginary part, are indeed almost normal while the joint distribution is not.

To assess the complex normality assumption of Rowe and Logan (2004), we have applied $\hat{\xi}_n(\boldsymbol{\nu}_{3,1})$ to the datasets in each voxel and obtained a 128 × 128 array where each component is the value of $\hat{\xi}_n(\boldsymbol{\nu}_{3,1})$. We set aside considerations regarding this multiple testing situation and concentrate on the individual tests.

The left panel of Fig. 1 shows the values of $\hat{\xi}_n(\boldsymbol{\nu}_{3,1})$, represented by a heat map.

It is remarkable to see the image of the brain tissue emerging very clearly and corresponding to high values of our test statistic, while voxels located outside the brain yields small values of the test statistic. Note that this panel is very similar to their Fig. 1c, representing the magnitude image of their raw data. The right panel of Fig. 1 represents the results of applying our test at level 0.05; white pixels correspond to significant tests. Besides the fact that the brain structure is still well apparent from this two-color map, this panel offers a validation of the complex normality assumptions in Rowe and Logan (2004).

Thus, our test gives strong arguments in favor of Rowe and Logan (2004)'s complex normal noise hypothesis which is an important assumption in their modeling of fMRI images. Thus, the present test contributes to the validation of their fMRI activation model and the conclusions they derive.

8 Conclusion

Complex data are becoming common in applications and, in response, statistical models have been proposed to analyze such data while taking into account their structural characteristics. In these models, and in analogy with real data, the assumption of complex normality plays a central role. But there is a lack of procedures to assess this distributional assumption. Extensive work with real data has shown that goodness-of-fit tests of (real) normality based on the empirical characteristic function are very powerful in a variety of situations. Thus, the present work proposes a goodness-of-fit test for the complex normal distribution based on its empirical characteristic function. The results of a simulation study show that the levels of the test, which are based on an asymptotic approximation, are close to nominal for moderate sample sizes. Moreover, the power of the test is rather good, being close to 1 against almost all alternatives we have considered when $n = 250$.

The derivation of our test requires information about (a) the behavior of linear combinations in complex normal vectors, (b) the statistical behavior of quadratic forms in complex normal vectors and (c) the problem of estimating the parameters in the distribution along with their asymptotic distribution. A few results appear in the literature on each of these points, but not at the level we require. Thus, they are developed here in some detail, as they may have an interest of their own, e.g., as building blocks for other statistical procedures for complex data. For example, they may serve in extending our goodness-of-fit test procedure to the case of the complex multinormal, as well as for other distributional models for complex multidimensional data.

Appendix

Proof of Theorem 1. From Mardia et al. (1979, p. 33), $Z_1, \ldots, Z_d$ are independent if and only if

$$\varphi_{(Z_1,\ldots,Z_d)}(v_1, \ldots, v_d) = \prod_{k=1}^{d} \varphi_{Z_k}(v_k). \tag{26}$$

We proceed by induction starting with $d = 2$. Let $Z = (Z_1, Z_2)^\mathsf{T} \sim CN_2(\mu, \Gamma, P)$. If $\mu = (\mu_1, \mu_2)^\mathsf{T}$, $v = (v_1, v_2)^\mathsf{T}$ and $\Gamma = \mathrm{diag}(\gamma_{11}, \gamma_{22})$, $P = \mathrm{diag}(p_{11}, p_{22})$, then, from (7),

$$\varphi_{(Z_1,Z_2)}(v) = \exp\Big\{\jmath \operatorname{Re}\left(v_1^*\mu_1\right) + \jmath \operatorname{Re}\left(v_2^*\mu_2\right) \\ - \frac{1}{4}\left[|v_1|^2\gamma_{11} + |v_2|^2\gamma_{22} + \operatorname{Re}\left((v_1^*)^2 p_{11} + (v_2^*)^2 p_{22}\right)\right]\Big\} \\ = \varphi_{Z_1}(v_1)\varphi_{Z_2}(v_2).$$

Inversely, assume $Z_1,\ Z_2$ are independent so that (26) is verified. Letting $\Gamma = (\gamma_{ij})$ and $P = (p_{ij})$, we have for all $v = (v_1, v_2)^{\mathsf{T}} \in \mathbb{C}^2$

$$-4\ln(\varphi_{(Z_1,Z_2)}(v)) = -4\ln(\varphi_{Z_1}(v_1)\varphi_{Z_2}(v_2)),$$

so that

$$-4\jmath \operatorname{Re}(v^{\mathcal{H}}\mu) + v^{\mathcal{H}}\Gamma v + \operatorname{Re}(v^{\mathcal{H}}Pv^*) = -4\jmath(\operatorname{Re}(v_1^*\mu_1) + \operatorname{Re}(v_2^*\mu_2)) \\ + |v_1|^2\gamma_{11} + |v_2|^2\gamma_{22} + \operatorname{Re}\left((v_1^*)^2 p_{11} + (v_2^*)^2 p_{22}\right),$$

which shows that

$$\operatorname{Re}(v_1^* v_2 \gamma_{12}) + \operatorname{Re}((v_1 v_2)^* p_{12}) = 0.$$

With $v_1^* = a_1 - \jmath b_1$ and $v_2 = a_2 + \jmath b_2$, simple algebra leads to the equivalent equation:

$$(a_1a_2 + b_1b_2)\operatorname{Re}(\gamma_{12}) + (a_1a_2 - b_1b_2)\operatorname{Re}(p_{12}) \\ -(a_1b_2 - b_1a_2)\operatorname{Im}(\gamma_{12}) + (b_1a_2 + a_1b_2)\operatorname{Im}(p_{12}) = 0, \tag{27}$$

which must hold for all a_1, a_2, b_1, b_2. In particular, they must hold for the following cases:

$$a_1 = a_2 = b_1 = b_2 = \lambda \\ a_1 = a_2 = b_1 = -b_2 = \lambda \\ b_1 = b_2 = 0,\ a_1 = a_2 = \lambda \\ a_2 = b_1 = 0,\ a_1 = b_2 = \lambda$$

for some $\lambda \in \mathbb{R}\backslash\{0\}$. This leads to the linear system $\lambda^2 AX = 0$ with $X = (\operatorname{Re}(\gamma_{12}), \operatorname{Re}(p_{12}), \operatorname{Im}(\gamma_{12})), \operatorname{Im}(p_{12})))^{\mathsf{T}}$ and

$$A = \begin{pmatrix} 2 & 0 & 0 & 2 \\ 0 & 2 & 2 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix}.$$

Since $\det(A) \neq 0$, the only solution is $\operatorname{Re}(\gamma_{21}) = \operatorname{Re}(p_{12}) = \operatorname{Im}(\gamma_{21}) = \operatorname{Im}(p_{12}) = 0$, which proves that Γ and P must be diagonal.

Now, for a given $d > 2$, let

$$Z_{(d+1)} = (Z_1, \ldots, Z_d, Z_{d+1})^{\mathsf{T}} \sim CN_{d+1}(\mu_{d+1}, \Gamma_{d+1}, P_{d+1}),$$

where $Z_1, \ldots, Z_d$ are independent, which, by the induction hypothesis, implies $\Gamma_d = \operatorname{diag}(\gamma_{11}, \ldots, \gamma_{dd})$ and $P_d = \operatorname{diag}(p_{11}, \ldots, p_{dd})$. Let

$$\Gamma_{d+1} = \left(\begin{array}{c|c} \Gamma_d & \Gamma_{(\cdot,d+1)} \\ \hline \Gamma_{(d+1,\cdot)} & \gamma_{d+1,d+1} \end{array}\right),$$

with

$$\Gamma_{(\cdot,d+1)} = \begin{pmatrix} \gamma_{1,d+1} \\ \vdots \\ \gamma_{d,d+1} \end{pmatrix} = \Gamma^{\mathcal{H}}_{(d+1,\cdot)},$$

and similarly

$$P_{d+1} = \left(\begin{array}{c|c} P_d & P_{(\cdot,d+1)} \\ \hline P_{(d+1,\cdot)} & p_{d+1,d+1} \end{array}\right),$$

with

$$P_{(\cdot,d+1)} = \begin{pmatrix} p_{1,d+1} \\ \vdots \\ p_{d,d+1} \end{pmatrix} = P^{\mathsf{T}}_{(d+1,\cdot)}.$$

The proof is that

$$\varphi_{(Z_1,\cdots,Z_{d+1})^{\mathsf{T}}}((\nu_1, \cdots, \nu_{d+1})) = \prod_{k=1}^{d+1} \varphi_{Z_k}(\nu_k)$$

when $\Gamma_{(.,d+1)} = 0$ and $P_{(.,d+1)} = 0$ is straightforward. We focus on the converse. Consider the equality

$$-4 \ln \left(\varphi_{(Z_1,\ldots,Z_{d+1})}(v_1, \ldots, v_{d+1})\right) = -4 \ln \left(\prod_{k=1}^{d+1} \varphi_{Z_k}(v_k) \right),$$

which should hold for any values of $v_1, \ldots, v_{d+1}$. It is easy to see that the right hand side of this equation is

$$- 4\jmath \,\mathrm{Re}(v^{\mathcal{H}}_{(d+1)}\mu_{d+1}) + \sum_{k=1}^{d+1} |v_k|^2 \gamma_{kk} + \mathrm{Re}\left(\sum_{k=1}^{d+1} (v_k^*)^2 p_{kk}\right), \tag{28}$$

with $v_{(d+1)} = (v_1, \ldots, v_{d+1})^{\mathsf{T}}$. As for the left hand side, we have from (7)

$$\begin{aligned}
-4\ln(\varphi_{Z_{(d+1)}}(v_{(d+1)})) &= -4\jmath \,\mathrm{Re}(v^{\mathcal{H}}_{(d+1)}\mu_{d+1}) + v^{\mathcal{H}}_{(d+1)}\Gamma_{d+1}v_{(d+1)} \\
&\quad + \mathrm{Re}(v^{\mathcal{H}}_{(d+1)} P_{d+1} v^*_{(d+1)}) \\
&= -4\jmath \,\mathrm{Re}(v^{\mathcal{H}}_{(d+1)}\mu_{d+1}) + \sum_{k=1}^{d+1}\left(|v_k|^2\gamma_{kk}\right) \\
&\quad + \mathrm{Re}\left(\sum_{k=1}^{d+1}(v_k^*)^2 p_{kk}\right) \\
&\quad + 2\,\mathrm{Re}\left(\sum_{k=1}^{d} v_k^* v_{d+1}\gamma_{k,d+1} + (v_{d+1}v_k)^* p_{k,d+1}\right).
\end{aligned}$$

Comparing this with (28) shows that we must have:

$$\mathrm{Re}\left(\sum_{k=1}^{d+1} v_k^* v_{d+1}\gamma_{k,d+1} + (v_{d+1}v_k)^* p_{k,d+1}\right) = 0.$$

Now, if there is only one nonzero v_j among $\{v_1, \ldots, v_d\}$ while $v_{d+1} \neq 0$, then

$$\mathrm{Re}(v_j^* v_{d+1}\gamma_{j,d+1}) + \mathrm{Re}((v_j v_{d+1})^* p_{j,d+1}) = 0.$$

The argument used in the case $d = 2$ shows that $\gamma_{j,d+1} = p_{j,d+1} = 0$. Because j is arbitrary, $\Gamma_{(\cdot,d+1)} = P_{(\cdot,d+1)} = 0$ and, thus, $\Gamma_{(d+1)}$ and $P_{(d+1)}$ are diagonal. This concludes the proof. □

Proof of Corollary 1. Without loss of generality, assume $\mu_1 = \mu_2 = 0$. The proof that Z_1 and Z_2 are independent when $\Gamma_{12} = P_{12} = 0$ follows from noticing that this entails $\varphi_Z(\nu_1, \nu_2) = \varphi_{Z_1}(\nu_1)\varphi_{Z_2}(\nu_2)$.

Inversely, let $Z_{1,k} \sim CN_1(0, \gamma_{k,k}, p_{k,k})$ be the k-th component of Z_1 and $Z_{2,\ell} \sim CN_1(0, \gamma_{\ell,\ell}, p_{\ell,\ell})$ the ℓ-th component of Z_2.

Assuming that Z_1, Z_2 are independent, $Z_{1,k}$ and $Z_{2,\ell}$ are independent and Theorem 1 ensures that the (k, ℓ) coefficients in Γ_{12} and P_{12} are equal to zero.

Because k, ℓ are abitrary, the proof follows. □

Proof of Theorem 2. From (12), $\underline{Z}^{\mathcal{H}} R \underline{Z} = 2\mathcal{X}^{\mathsf{T}} S \mathcal{X}$ where $\mathcal{X} \sim N_{2d}(\mu_{\mathcal{X}}, \Sigma_{\mathcal{X}})$. Let $S^{1/2}$ be a real $2d \times 2d$ matrix such that $(S^{1/2})^{\mathsf{T}} S^{1/2} = S$. Using an eigenvalue decomposition, we have $S^{1/2}\Sigma_{\mathcal{X}}(S^{1/2})^\top = O\Lambda O^\top$ where Λ is the diagonal matrix of the real eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{2d} \geq 0$ and O is orthogonal. Hence, $O^\top S^{1/2}\mathcal{X} \sim N_{2d}(O^\top S^{1/2}\mu_{\mathcal{X}}, \Lambda)$. This entails that

$$\underset{\sim}{Z}^{\mathcal{H}} R \underset{\sim}{Z} = 2(O^\top S^{1/2}\mathcal{X})^\top(O^\top S^{1/2}\mathcal{X}) = 2\sum_{k=1}^{q}\lambda_k\left(\mathcal{N}_k + \frac{\tau_k}{\sqrt{\lambda_k}}\right)^2 + 2\sum_{k=q+1}^{2d}\tau_k^2,$$

where q is the number of nonzero eigenvalues λ_k, the $\mathcal{N}_k$ are independent $N(0,1)$ random variables and τ_k is the k-th component of $O^\top S^{1/2}\mu_{\mathcal{X}}$.

Taking $S^{1/2} = MR^{1/2}M^{-1}$, we get $O^\top S^{1/2}\mu_{\mathcal{X}} = O^\top MR^{1/2}\underset{\sim}{\mu}$, so that $\tau_k = e_k^\top O^\top MR^{1/2}\underset{\sim}{\mu}$ where e_k is the k-th vector in the canonical basis for $\mathbb{R}^{2d}$. We then have

$$\underset{\sim}{Z}^{\mathcal{H}} R \underset{\sim}{Z} \sim 2\sum_{k=1}^{q}\lambda_k\chi_1^2(\delta_k^2) + 2\sum_{k=q+1}^{2d}\tau_k^2,$$

where $\delta_k^2 = \frac{\tau_k^2}{\lambda_k}$ $(1 \leq k \leq q)$, which vanishes if $\mu = 0$.

Now, if x_k is the kth eigenvector of $\Gamma_P R$ corresponding to the eigenvalue α_k, we have $\Gamma_P R x_k = \alpha_k x_k$. This gives

$$\begin{aligned}\Gamma_P M^{-1}SMx_k &= \alpha_k x_k,\\ S^{1/2}M\Gamma_P M^{-1}(S^{1/2})^\top S^{1/2}Mx_k &= \alpha_k S^{1/2}Mx_k.\end{aligned}$$

which, from (3) and after taking the conjugate, gives

$$S^{1/2}\Sigma_{\mathcal{X}}(S^{1/2})^\top(S^{1/2}M^*x_k^*) = \frac{\alpha_k}{2}(S^{1/2}M^*x_k^*).$$

We thus have $S^{1/2}M^*x_k^* = Oe_k$ and $\alpha_k/2 = \lambda_k$. We finally obtain

$$\begin{aligned}\tau_k &= e_k^\top O^\top MR^{1/2}\underset{\sim}{\mu}\\ &= x_k^{\mathcal{H}}M^{\mathcal{H}}(S^{1/2})^\top MM^{-1}S^{1/2}M\underset{\sim}{\mu}\\ &= x_k^{\mathcal{H}}M^{\mathcal{H}}MRM^{-1}M\underset{\sim}{\mu}\\ &= \frac{1}{2}x_k^{\mathcal{H}}R\underset{\sim}{\mu}.\end{aligned}$$

□

Proof of Theorem 3. First, we recall the following property of the Moore–Penrose pseudo-inverse (Penrose 1955). Let A be $m \times n$ complex and B be $n \times p$ complex, if $AA^{\mathcal{H}} = I_m$ or $B^{\mathcal{H}}B = I_p$, then the Moore–Penrose pseudo-inverse $(AB)^+$ of AB satisfies $(AB)^+ = B^+A^+$.

From (1) and equation above (1), $Q = 2\mathcal{X}^{\mathsf{T}} M \Gamma_P^{+} M^{-1} \mathcal{X}$, where $\mathcal{X} \sim N_{2d}(0, \Sigma_{\mathcal{X}})$. From (3) and the above property,

$$\Gamma_P^{+} = \left(M^{-1} \Sigma_X (M^{\mathcal{H}})^{-1}\right)^{+} = \frac{1}{2} M^{-1} \Sigma_{\mathcal{X}}^{+} M.$$

Therefore, $Q = \mathcal{X}^{\mathsf{T}} \Sigma_{\mathcal{X}}^{+} \mathcal{X}$. Now, from (Searle 1971, Corollary 2s.2, p.69), $Q \sim \chi^2_{\mathrm{tr}(\Sigma_{\mathcal{X}}^{+} \Sigma_{\mathcal{X}})}$ and (Rao 1973, Proposition (ii)b p. 25) ensures that $\mathrm{tr}(\Sigma_{\mathcal{X}}^{+} \Sigma_{\mathcal{X}}) = \mathrm{rank}(\Sigma_{\mathcal{X}}) = \mathrm{rank}(\Gamma_P)$. When Γ_P is of full rank, $\Gamma_P^{+} = \Gamma_P^{-1}$ and $\mathrm{rank}(\Gamma_P) = 2d$. □

Proof of Theorem 4. Because we were not able to find a proof of this theorem in the literature, we provide one here.

In view of (1), we have $\sqrt{n}(\underset{\sim}{\bar{Z}_n} - \underset{\sim}{\mu}) = M^{-1} n^{-1/2} \sum_{i=1}^{n} (\mathcal{X}_i - \mu_{\mathcal{X}})$. Hence,

$$\varphi_{\sqrt{n}(\underset{\sim}{\bar{Z}_n} - \underset{\sim}{\mu})}(\underset{\sim}{\nu}) = \varphi_{n^{-1/2} \sum_{i=1}^{n} (\mathcal{X}_i - \mu_{\mathcal{X}})}(t) \underset{n \to \infty}{\longrightarrow} \varphi_{N_{2d}(0, \Sigma_{\mathcal{X}})}(t)$$

using a result in Feldman (1965) and the standard C.L.T. for real r.v. But

$$\varphi_{N_{2d}(0, \Sigma_{\mathcal{X}})}(t) = \exp\left(-\frac{1}{2} t^{\mathsf{T}} \Sigma_{\mathcal{X}} t\right) = \exp\left(-\frac{1}{2} \underset{\sim}{\nu}^{\mathcal{H}} M^{\mathcal{H}} \Sigma_{\mathcal{X}} M \underset{\sim}{\nu}\right) = \varphi_{\underset{\sim}{CN}_d(\underset{\sim}{0}, \Gamma_P)}(\underset{\sim}{\nu}),$$

from (8). □

Proof of Proposition 1. From (6), for all ν, $\varphi_{AX_n}(\nu) = \varphi_{X_n}(A^{\mathcal{H}}\nu) \underset{n \to \infty}{\longrightarrow} \varphi_X(A^{\mathcal{H}}\nu) = \varphi_{AX}(\nu)$. □

Proof of Corollary 2. The result follows from Theorem 4, Proposition 1 and (10). □

Proof of Proposition 2. The result is obvious regarding the method of moments estimators. As for the m.l.e., the likelihood function can we written from (4) as

$$L(\mu, \Gamma_P) = \frac{1}{\pi^{nd} |\Gamma_P|^{n/2}} \exp\left\{-\frac{1}{2} \sum_{k=1}^{n} \left(\underset{\sim}{Z_k} - \underset{\sim}{\mu}\right)^{\mathcal{H}} \Gamma_P^{-1} (\underset{\sim}{Z_k} - \underset{\sim}{\mu})\right\}$$
$$= \frac{1}{\pi^{nd} |\Gamma_P|^{n/2}} \exp\left\{-\frac{1}{2} \mathrm{tr}\left(\Gamma_P^{-1} (\underset{\sim}{\mathbf{Z}} - \underset{\sim}{\mu} e^{\mathsf{T}})(\underset{\sim}{\mathbf{Z}} - \underset{\sim}{\mu} e^{\mathsf{T}})^{\mathcal{H}}\right)\right\}$$

where $\underset{\sim}{\mathbf{Z}} = \left(\underset{\sim}{Z_1}, \ldots, \underset{\sim}{Z_n}\right)$ is a $2d \times n$ matrix of a.c.r.v. and $e = \left(1, \ldots, 1\right)^{\mathsf{T}}$ is n-dimensional. With $\underset{\sim}{\bar{\mathbf{Z}}} = n^{-1} \sum_{k=1}^{n} \underset{\sim}{Z_k}$, we notice that $\underset{\sim}{\mathbf{Z}} e = n \underset{\sim}{\bar{\mathbf{Z}}}$. Moreover,

$$(\underset{\sim}{\mathbf{Z}} - \underset{\sim}{\mu} e^{\mathsf{T}})(\underset{\sim}{\mathbf{Z}} - \underset{\sim}{\mu} e^{\mathsf{T}})^{\mathcal{H}} = n \hat{\Gamma}_P + n(\underset{\sim}{\bar{\mathbf{Z}}} - \underset{\sim}{\mu})(\underset{\sim}{\bar{\mathbf{Z}}} - \underset{\sim}{\mu})^{\mathcal{H}},$$

where $n\hat{\Gamma}_P = \underset{\sim}{\mathbf{Z}}\underset{\sim}{\mathbf{Z}}^{\mathcal{H}} - n\bar{\underset{\sim}{\mathbf{Z}}}\bar{\underset{\sim}{\mathbf{Z}}}^{\mathcal{H}}$. Thus,

$$L(\mu, \Gamma_P) = \frac{1}{\pi^{nd}\,|\Gamma_P|^{n/2}} \exp\left\{-\frac{n}{2}\,\mathrm{tr}\left[\Gamma_P^{-1}\hat{\Gamma}_P\right]\right\}$$
$$\exp\left\{-\frac{n}{2}\,\mathrm{tr}\left[\Gamma_P^{-1}\left((\bar{\underset{\sim}{\mathbf{Z}}} - \underset{\sim}{\mu})(\bar{\underset{\sim}{\mathbf{Z}}} - \underset{\sim}{\mu})^{\mathcal{H}}\right)\right]\right\}.$$

Obviously, $L(\mu, \Gamma_P)$ is maximized in μ when $\underset{\sim}{\mu} = \bar{\underset{\sim}{\mathbf{Z}}}$. Moreover,

$$\begin{aligned} L(\hat{\mu}, \Gamma_P) &= \frac{1}{\pi^{nd}\,|\Gamma_P|^{n/2}} \exp\left\{-\frac{n}{2}\,\mathrm{tr}\left(\Gamma_P^{-1}\hat{\Gamma}_P\right)\right\} \\ &\leq \frac{1}{\pi^{nd}\left|\hat{\Gamma}_P\right|^{n/2}} \exp\{-dn\} \\ &= L(\hat{\mu}, \hat{\Gamma}_P), \end{aligned}$$

where the inequality holds in view of Srivastava and Khatri (1979, Theorem 1.10.4). This gives the m.l.e. for Γ_P. We recover the corresponding estimators for Γ and P by extracting the corresponding terms in $\hat{\Gamma}_P$. □

Acknowledgments The authors would like to thank Daniel Rowe for having provided the fMRI data used in Sect. 7. The research of the second author has been funded by the NSERC of Canada.

References

Adalı, T., Schreier, P. J., Scharf, L. L. (2011). Complex-valued signal processing: The proper way to deal with impropriety. *IEEE Transactions on Signal Processing*, *59*, 5101–5125.

Andersen, H. H., Højbjerre, M., Sørensen, D., Eriksen, P. S. (1995). *Linear and graphical models for the multivariate complex normal distribution*, vol. 101 of Lecture Notes in Statistics. New York: Springer. doi:10.1007/978-1-4612-4240-6.

Bernard, O., D'hooge, J., Friboulet, D. (2006). Statistics of the radio-frequency signal based on k distribution with application to echocardiography. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, *53*, 1689–1694.

Bilodeau, M., Brenner, D. (1999). *Theory of multivariate statistics*. Springer Texts in Statistics. New York: Springer.

Calhoun, V., Adalı, T., Pearlson, G., Van Zijl, P., Pekar, J. (2002). Independent component analysis of fmri data in the complex domain. *Magnetic Resonance in Medicine*, *48*, 180–192.

Calhoun, V. D., Adalı, T. (2012). Analysis of complex-valued functional magnetic resonance imaging data: are we just going through a "phase"? *Bulletin of the Polish Academy of Sciences, Technical Sciences*, *60*, 371–387.

Csörgő, S. (1981). Limit behaviour of the empirical characteristic function. *The Annals of Probability*, *9*, 130–144.

Eriksson, J. (2010). Essential statistics and tools for complex random variables. *IEEE Transactions in Signal Processing*, *58*, 5400–5408.

Feldman, J. (1965). A short proof of the levy continuity theorem in hilbert space. *Israel Journal of Mathematics*, *3*, 99–103.

González-Jiménez, D., Alba-Castro, J. L. (2007). Shape-driven gabor jets for face description and authentication. *IEEE Transactions on Information Forensics and Security*, *2*, 769–780.

Goodman, N. R. (1963). Statistical analysis based on a certain multivariate complex gaussian distribution (an introduction). *The Annals of Mathematical Statistics*, *34*, 152–177.

Henze, N. (2002). Invariant tests for multivariate normality: a critical review. *Statistical Papers*, *43*, 467–506.
Mardia, K. V., Kent, J. T., Bibby, J. M. (1979). *Multivariate analysis*. Probability and Mathematical Statistics: A Series of Monographs and Textbooks. London: Academic Press [Harcourt Brace Jovanovich Publishers].
Mecklin, C. J., Mundfrom, D. J. (2005). A monte carlo comparison of the type i and type ii error rates of tests of multivariate normality. *Journal of Statistical Computation and Simulation*, *75*, 93–107.
Novey, M., Adalı, T., Roy, A. (2010). A complex generalized gaussian distribution: characterization, generation, and estimation. *IEEE Transactions on Signal Processing*, *58*, 1427–1433.
Ollila, E., Koivunen, V. (2010). Robust estimation techniques for complex-valued random vectors. In: T. Adalı, S. Haykin (Eds.), *Adaptive Signal Processing, Next-Generation Solutions* (pp. 87–141). Hoboken: Wiley.
Penrose, R. (1955). A generalized inverse for matrices. *Proceedings of the Cambridge Philosophical Society*, *51*, 406–413.
Picinbono, B. (1996). Second-order complex random vectors and normal distributions. *IEEE Transactions on Signal Processing*, *44*, 2637–2640.
Rao, C. R. (1973). *Linear statistical inference and its applications*, 2nd ed. Wiley Series in Probability and Mathematical Statistics. New York: Wiley.
Rowe, D. B., Logan, B. R. (2004). A complex way to compute fmri activation. *NeuroImage*, *23*, 1078–1092.
Searle, S. R. (1971). *Linear models*. New York: Wiley.
Srivastava, M., Khatri, C. (1979). *An introduction to multivariate statistics*. New York: Elsevier.
Turin, G. L. (1960). The characteristic function of Hermitian quadratic forms in complex normal variables. *Biometrika*, *47*, 199–201.
van den Bos, A. (1995). The multivariate complex normal distribution-a generalization. *IEEE Transactions on Information Theory*, *31*, 537–539.
Wooding, R. (1956). The multivariate distribution of complex normal variables. *Biometrika*, *43*, 212–215.