A ROBUST ESTIMATOR AND DETECTOR OF CIRCULARITY OF COMPLEX SIGNALS

Esa Ollila, Visa Koivunen
Aalto University
Dept. of Signal Processing and Acoustics
P.O.Box 13000, FI-00076 Aalto, Finland

H. Vincent Poor
Princeton University
Department of Electrical Engineering
Princeton, NJ 08544, USA

ABSTRACT
Recent research has revealed that circularity (or, propriety) of complex random signals can be exploited in developing optimal signal processors. In this paper, a robust estimator of circularity is proposed. The estimate is found by solving M-estimation equations and employing a novel weighting scheme. A simple iterative algorithm for its computation is introduced. A robust circularity detector stemming from the large sample properties of the estimator is proposed. It is shown to be valid detector under the broad class of complex elliptically symmetric (CES) distributions. An illustrative example demonstrating the reliable performance of the detector in a practical signal processing application is provided.

Index Terms— Improper, 2nd-order circular complex random variable; circularity coefficient; circularity quotient

1. INTRODUCTION
Recent research has revealed that circularity/non-circularity (or, propriety/impropriety) of complex random signals can be exploited in developing optimal signal processors; e.g. see the recent text-book [1] solely devoted to this topic. Recall that a complex random variable (r.v.a.) $z = x + jy$ is said to be circular or, to have a circularly symmetric distribution (around the origin), if its distribution remains invariant under multiplication by any (complex) number on the unit complex sphere, i.e. if $z$ has the same distribution as $e^{j\theta}z$ $\forall \theta \in \mathbb{R}$. A circular r.v.a. possesses the property that it is statistically uncorrelated with its complex-conjugate, i.e. the pseudo-variance $E[z^2] = 0$, a property that is referred to as 2nd-order circularity or propriety. We note that a complex r.v.a. can be 2nd-order circular (or, proper) but not necessarily circularly symmetric.

Standard techniques derived for the circular case are suboptimal in the noncircular case. Hence to exploit the non-circularity property (or lack of it), a detector of circularity is needed. In [2], a generalized likelihood ratio test (GLRT) of circularity assuming complex Gaussian (normal) samples was derived and further studied in [3, 4, 5, 6]. The adjusted GLRT of circularity derived in [7] has the desirable feature of being robust to departures from Gaussianity within CES distributions [9] with finite 4th-order moments. In the univariate case, [8] considered the GLRT of circularity assuming a complex generalized Gaussian distribution, whereas Wald’s type detectors were constructed using the sample circularity quotient in [9].

Robustness (i.e. insensitivity to measurement errors, outliers, impulsive heavy-tailed noise and model class selection errors) of the developed signal processing techniques has been widely accepted as an important design criterion [10]. In this paper a robust circularity estimator is developed that solves M-estimating equations based on a robust and novel weighting scheme and a simple iterative algorithm is provided for its computation; see Section 2. In Section 3, we derive the asymptotic distribution of the estimator under CES distributions. The asymptotic distribution is then used in devising a robust detector of circularity. In Section 4, the utility of the detector is illustrated with a communications example. In Section 5, a connection to Tyler’s $M$-estimator of shape matrix is established.

2. ROBUST CIRCULARITY MEASURE
A natural measure of circularity of a complex r.v.a. $z = x + jy$ is the circularity quotient [4], defined as the ratio of the pseudo-variance and the variance:

$$\varrho(z) \triangleq E[z^2]/E[|z|^2] \in \Omega,$$

where $\Omega = \{z \in \mathbb{C} : |z| \leq 1\}$ denotes the closed unit disk, $\partial \Omega$ its boundary, the unit circle and $\Omega^c = \Omega \setminus \partial \Omega$ the open unit disk. The modulus of the circularity quotient, $|\varrho| \in [0,1]$, is called the circularity coefficient [11] of $z$ and its argument $\vartheta \triangleq \text{Arg}(\varrho)$ as the circularity angle [4]. The circularity coefficient measures the “degree of circularity” as it equals the squared eccentricity of the ellipse defined by the real covariance matrix of $(x,y)^T$. Thus $|\varrho| = 0$ when $z$ is 2nd-order circular and $|\varrho| = 1$, when $z$ is maximally non-circular (i.e. $x$ or $y$ is equal to zero with probability 1, or, $x$ is a linear function of $y$).

If $z_1, \ldots, z_n$ is an independent and identically distributed (i.i.d.) random sample from a distribution with finite 2nd-order moments, then

$$\hat{\varrho} = \frac{1}{n} \sum_{i=1}^{n} z_i^2 / \sum_{i=1}^{n} |z_i|^2, \tag{1}$$

and $|\hat{\varrho}|$ are natural sample estimates of $\varrho$ and $|\varrho|$. Let

$$u_i = \begin{cases} z_i/|z_i| = \exp(j\theta_i), & \text{if } |z_i| \neq 0 \\ 0, & \text{if } |z_i| = 0 \end{cases}$$

denote the normalizations of the data points onto the unit circle $\partial \Omega$, i.e. their spatial signs, $\theta_i = \text{Arg}(z_i) \in (-\pi, \pi]$. We can now write (1) as a weighted mean of the squares of the spatial signs:

$$\hat{\varrho} = \frac{1}{n} \sum_{i=1}^{n} u_i^2 w_i, \quad \text{where } w_i \triangleq \frac{|z_i|^2}{\sum_{j=1}^{n} |z_j|^2} \tag{2}$$

Esa Ollila performed the work while visiting Princeton University, Department of Electrical Engineering. This paper was prepared in part under the support of U.S. Office of Naval Research under Grant N00014-09-1-0342.
are weights that satisfy
\[ 1 = \frac{1}{n} \sum_{i=1}^{n} w_i. \]  
\[ (3) \]
Note that \[ u_i^2 = \exp(j2\theta_i) \in \partial\Omega. \]

We now assume that \( z_1, \ldots, z_n \neq 0 \) and \( n > 2 \). In the spirit of (2) and (3), we define a robust circularity estimator, denoted by \( \hat{q} \), as the choice of \( q \in \Omega^2 \) solving the pair of M-estimating equations:
\[ q = \frac{1}{n} \sum_{i=1}^{n} u_i^2 \varphi_i \]  
\[ (4) \]
\[ 1 = \frac{1}{n} \sum_{i=1}^{n} \varphi_i \]  
\[ (5) \]
where \( \varphi_i = \varphi(u_i^2; q) \) and \( \varphi(\cdot, \cdot) : \partial\Omega \times \Omega \rightarrow [0, \infty) \) is a weighting function. We consider the following robust weighting scheme:
\[ \varphi(u_i^2; q) = \frac{1 - |q|^2}{1 - |u_i^2|^2} = \frac{1 - |q|^2}{1 - |q| \cos(\alpha_i)}. \]  
\[ (6) \]
where \( \alpha = \angle(u_i^2, q) \in [0, \pi] \) is the angle between \( u_i^2 \in \partial\Omega \) and \( q \in \Omega^2 \) and \( \langle \cdot, \cdot \rangle \) denotes the (Euclidean) inner product of a vector space \( \mathbb{C} \), \( \langle z_1, z_2 \rangle = \text{Re}(\bar{z}_1 z_2) \) for \( z_1, z_2 \in \mathbb{C} \). The pair of equations (4) and (5) form implicit estimating equations whose solution \( \hat{q} \) can be found by the simple iterative algorithm given at the end of this section in equation (7). In Section 5 we show that the estimator is intimately related to Tyler’s M-estimator [12] of the shape matrix.

It is instructive to compare the weights \( w_i \) of \( \hat{q} \) in (2) to those of the robust estimator \( \hat{q} \) solving (4) and (5). Now the proposed robust weighting scheme assigns weights
\[ \varphi_i = \varphi(u_i^2; \hat{q}) = \frac{1 - |\hat{q}|^2}{1 - |\hat{q}| \cos(\alpha_i)}, \]
where \( \alpha_i = \angle(u_i^2, \hat{q}) \). Note that \( \varphi_i \in [1 - |\hat{q}|, 1 + |\hat{q}|] \), where the maximal weight \( 1 + |\hat{q}| \) is obtained if \( u_i^2 \) points in the same direction as \( \hat{q} \) (i.e. \( \hat{q} \parallel u_i^2 \)) whereas the minimum value \( 1 - |\hat{q}| \) is obtained if they point in the opposite direction (i.e. \( \hat{q} \perp u_i^2 \)). If vectors are orthogonal (i.e. \( \hat{q} \perp u_i^2 \)), then the weight is \( \varphi_i = 1 - |\hat{q}|^2 \). Note that the weight attached to each \( u_i^2 \) is solely based on its alignment with \( \hat{q} \) but not on the norm \( |z_i| \) of the observation. The weighting given by (2) is sensitive to outliers since the weights \( w_i \) depend only on the radius of the observations. Clearly, an observation \( z_i \) with large \( |z_i| \) leads to weights such that \( w_i \) is large and simultaneously all \( w_j \), \( j \neq i \), are small due the dominant term \( |z_i|^2 \) in the denominator of each \( w_j \). Thus, in case of a single large outlier \( z_i \), the value of \( \hat{q} \) is in essence determined solely by the term \( u_i^2 w_i \).

We now illustrate the above with an example. We generated a sample of \( n = 10 \) observations from a complex normal (CN) distribution with unit variance and circularity quotient \( \varrho = 0.9 \exp(j\pi/4) \). The left column of Figure 1 shows the squared spatial signs \( u_i^2 \) of the observations \( z_i \), \( i = 1, \ldots, n \), along with \( \varrho \) and its estimates \( \hat{\varrho} \) and \( \hat{\varrho} \). Figure 1 also depicts the corresponding weights, where the dashed vertical lines illustrate the minimum and maximum attainable weight values: \( u_i \in (0, n) \) and \( \varphi \in [1 - |\varrho|, 1 + |\varrho|] \). The robust estimator \( \hat{\varrho} \) is aligned with the majority of \( u_i^2 \)'s; \( u_i^2 \)'s with indices 1, 2, 5, 7, 10 that are largely dispersed on the unit circle are given smaller weights. However, the weights \( w_i \) depend only on the magnitudes of the observations \( z_i \) but not on their phases (orientations). The right hand side of Figure 1 shows what happens if we multiply the first observation \( z_1 \) by 1000, that is, its orientation remains constant but its magnitude is increased drastically. It is seen that \( \hat{\varrho} \) is drawn towards the observation \( u_1^2 \) and that \( w_1 \) is given almost maximal weight \( (w_1 \approx n) \) whereas the remaining weights are negligible \( (w_j \approx 0 \text{ for } j = 2, \ldots, 10) \). The estimator \( \hat{\varrho} \) and the respective weights on the contrary remain unchanged regardless of the highly deviating observation, i.e. outlier. Hence, it is robust.

Fig. 1. Left column: Illustration of \( \varrho \) and \( \hat{\varrho} \) and the corresponding weights for a sample of length \( n = 10 \) from a unit variance CN distribution with \( \varrho = 0.9 \exp(j\pi/4) \). Right column: \( \hat{\varrho} \) and \( \hat{\varrho} \) and the respective weights after the 1st observation \( z_1 \) is scaled by a 1000.

The estimator can be calculated by a simple iterative algorithm given below. It can be proved that for any given sample \( z_1, \ldots, z_n \neq 0 \), the solution \( \hat{\varrho} \) to (4) and (5) exists and that the sequence \( \{q_m, m = 0, 1, 2, \ldots\} \) defined by the iterative scheme
\[ q_{m+1} = \sum_{i=1}^{n} u_i^2 \varphi_{m,i} \sum_{i=1}^{n} \varphi_{m,i} \]  
\[ (7) \]
where \( \varphi_{m,i} = \varphi(u_i^2, q_m) \) converges to the solution \( \hat{\varrho} \) provided that the data stem from a continuous complex distribution and \( n > 2 \). The initial value \( q_0 \) can be any complex number in \( \Omega^2 \); e.g. \( q_0 = \varrho \) (provided that \( |\varrho| \neq 1 \)). The proof follows by establishing the relationship of (7) with the corresponding fixed-point iteration scheme of Tyler’s M-estimator [12] of scatter as in Section 5.
 Recall that in order to be able to apply optimal signal processing procedures, we need to detect circularity/non-circularity of the received data. Recall also that  \( q \) vanishes for a circular r.v.a). Thus, the distance of \( |q|^2 \) from zero is a very natural test statistic for circularity whose rejection region can be based on the asymptotic distribution of \( |q|^2 \). For this purpose, we now derive asymptotics of the estimator under the family of CES distributions \([9]\) which is a natural extension of the circular distributions.

We now assume that \( z_1, \ldots, z_n \) is an i.i.d. random sample from a complex elliptically symmetric (CES) distribution symmetric about the origin \([9] \), the density function of which is of the form

\[
 f(z) = g \left( \frac{|z|^2 - (z^*)^2, q}{\sigma^2 (1 - |q|^2)} \right) \frac{1}{\sigma^2 \sqrt{1 - |q|^2}},
\]

where \( g(\cdot) \) is a non-negative function called the density generator, \( \sigma^2 > 0 \) is the dispersion parameter and \( q \in \Omega^0 \) is a quotient parameter of the distribution. As an example, the complex normal (CN) distribution is obtained when \( g(t) = \pi^{-1} \exp(-t) \). Other examples include the complex-\( t \)-, complex generalized Gaussian and symmetric \( \alpha \)-stable (SoS) distributions. If \( q = 0 \), then the CES distributions reduce to the class of circular distributions, denoted by \( z \sim \text{Cir}(\sigma^2, q) \). We can now investigate the asymptotic distribution of \( |q|^2 \) under the null hypothesis \( H_0 : |q| = 0 \) (circular distribution) against the alternative \( H_1 : |q| \neq 0 \). We note that CES distributions can also be defined more generally without the assumption of a density as in \([9] \) and that \( q \) is assumed to lie inside the unit disk (as \( |q| = 1 \) implies a singular CES distribution).

The parameters \( \sigma^2 \) and \( q \) are inherently related to the variance \( E[|z|^2] \) and the circularity quotient \( q \) when the 2nd-order moment exists, i.e. when \( q \) satisfies condition \( \int_0^\infty t g(t) dt < \infty \). In this case, \( E[|z|^2] = \sigma^2 (\int_0^\infty t g(t) dt) \) and \( g = \varrho \), that is, \( \sigma^2 \) is equal to the variance up to a positive constant and the quotient parameter \( q \) is equal to the circularity quotient \( g \) of the distribution. Hence, if the 2nd-order moments exist, we write \( g \) in place of \( q \). If the 2nd-order moments do not exist (as is the case for complex Cauchy for example), the modulus \( |q| \in [0, 1) \) is still a valid measure of circularity as it equals the squared eccentricity of the elliptical contours whereas \( \varrho^2 \) is a measure of the scale of the contours.

**Theorem 1** Under \( H_1 \), the limiting distribution of \( \sqrt{n} (|q| - |q|) \) is a zero mean normal distribution with (asymptotic) variance \( \text{ASV}(|q|) = 2(|q|^2 - 1)^2 \). Under \( H_0 \), \( \sqrt{n}/2|q| \rightarrow \chi_2 \).

The notation \( \rightarrow_L \) means convergence in distribution and \( \chi_2 \) denotes the chi-distribution with 2 degrees of freedom (d.f.) (i.e. Rayleigh distribution with unit scale). Due to space limitations, the proof of Theorem 1 is omitted. For the limiting normal distribution, \( |q| \) requires finite 4th-order moments \([9]\) whereas \( q \) does not require any finite moment assumptions. Note that the asymptotic distribution under \( H_0 \) or \( H_1 \) does not depend on the reference CES distribution at all. Thus \( |q| \) is asymptotically distribution free under the class of CES distributions. This desirable property is based on the connection of \( q \) with Tyler’s \( M \)-estimator of the shape matrix.

Based on Theorem 1, a robust (Wald type) circularity detector rejects the null hypothesis \( H_0 \) if

\[
 n|q|^2/2 > \chi^2_{\alpha},
\]

where \( \chi^2_{\alpha} \) denotes the upper \( \alpha \)-quantile of the chi-squared distribution with 2 d.f. and \( \chi^2_{\alpha, \alpha} = 2 \chi^2_{\alpha}/n \). This detector is asymptotically valid with probability of false alarm (PFA) \((\text{type I error})\) equal to \( \alpha \) under an arbitrary CES distribution without any finite moment assumptions. For example, with PFA \( \alpha = 0.05 \), \( \chi^2_{0.05} = 5.9915 \).

Based on Theorem 1, the probability of detection is approximately (for large \( n \))

\[
 \gamma(q) = \Pr(\text{Reject } H_0 \text{ when } H_1 : q \neq 0 \text{ holds})
 = \Pr(|q| > c_{\alpha, \alpha} | q) \approx 1 - \Phi \left( \frac{\sqrt{n}(c_{\alpha, \alpha} - |q|)}{\sqrt{2} (1 - |q|^2)} \right),
\]

where \( \Phi \) denotes the cumulative distribution function of the standard normal distribution. The remarkable feature of the robust detector is that the probability of detection remains the same (for large \( n \)) regardless of the reference CES distribution.

We now investigate the power of the test by simulations. The sample \( z_1, \ldots, z_n \) is generated from a unit-dispersion \((\sigma^2 = 1)\) CN distribution and Cauchy distribution, respectively. For each generated sample, \( |q| = |q| \) when 2nd moments exists is kept fixed. Figure 2(a) shows the performance of the optimal GLRT-detector \([4]\) and the robust detector \((9)\) in the CN case by depicting the proportion of correct rejections (observed probability of detection) as a function of \( |q| \). In this simulation, the PFA was \( \alpha = 0.05 \), sample size was \( n = 500 \) and the number of generated samples (for each fixed \(|q|\)) was 1000. Note that at \( |q| = 0 \), the probability of rejection for both detectors are close to the nominal 0.05 level. As expected, the GLRT has better probability of detection (i.e. steeper slope). Results for the adjusted GLRT \([7]\) and the robust detector in the Cauchy case are shown in Figure 2(b). Although the adjusted GLRT is a valid test under the CES distribution, it fails in this case due to finite 4th-order moments assumption which does not hold in case of Cauchy. As expected, the probability of detection of the robust detector remains the same regardless of the reference CES distribution.

**4. SIGNAL PROCESSING EXAMPLE**

We consider the same example as in \([3]\) in which independent equiprobable binary symbols \( b \in \{-1, 1\} \) are transmitted over an additive noise channel so that the received signal is \( z = \exp(bz) + \varepsilon \), where \( \varepsilon \) is the additive circularly symmetric noise with dispersion \( \varrho^2 \) (i.e. \( \varepsilon \sim \text{Cir}(\varrho^2, q) \)) and \( \theta \) is the channel phase. Under \( H_0 \) (resp. \( H_1 \)), we assume that \( \theta \sim \text{Unif}(\pi, \pi) \) (resp. \( \theta \sim \mathcal{N}(\theta_0, \sigma^2) \)) is independent of \( b \). Thus we are interested in classifying the channel as either noncoherent or coherent, that is, is this channel rotationally invariant \((H_0)\), or can some phase information be extracted from the received data \((H_1)\)? Even though \( z \) is not a complex elliptically symmetric r.v.a., the proposed robust test is well suited for this hypothesis test. We consider the cases
in which the additive circularly symmetric noise $\varepsilon$ has a complex normal distribution and a complex symmetric $\alpha$-stable (SoS) distribution with $\alpha = 0.5$ (i.e., highly impulsive additive noise). Since $E[\theta^2] = 1$, we define the (generalized) signal to noise ratio (SNR) as $-10 \log_{10}(\sigma^2)$ (dB). Figure 3 shows the (observed) probability of detection (averaged over 2000 samples) versus SNR in both noise environments when the phase is $\theta_0 = \pi/4$, the phase tracking error variance is $\sigma^2 = 0.1$ and the sample size $n = 500$. In additive Gaussian noise, the adjusted GLRT has better detection performance; however when the noise follows the SoS distribution, the adjusted GLRT behaves practically like a random guess due to the lack of robustness. The robust detector, on the contrary, shows detection performance very similar to that in the Gaussian case.

![Fig. 3. Observed probability of detection versus SNR in the communications example; $n = 500$ and the number of simulated samples (for each SNR) is 2000.](image)

5. CONNECTION TO TYLER’S M-ESTIMATOR

Let $(x_i, y_i)^T$, $i = 1, \ldots, n$, denote the composite real-valued bi-variate data set formed by stacking the real part and imaginary part of $z_i = x_i + jy_i$, $i = 1, \ldots, n$, where $z_i \neq 0$. Let $P$ denote the set of all positive definite symmetric real $2 \times 2$ matrices. The choice of $C \in P$ that solves the $M$-estimation equation:

$$C = \frac{1}{n} \sum_{i=1}^{n} \hat{\phi}_i \left( \begin{array}{c} x_i \\ y_i \end{array} \right) \left( \begin{array}{c} x_i \\ y_i \end{array} \right)^T$$  \hspace{1cm} (10)

where $\hat{\phi}_i$ is the positive weight $\hat{\phi}_i = 2/(|x_i, y_i|C^{-1}(x_i, y_i)^T)$ is called Tyler’s $M$-estimator of scatter and is denoted by $\hat{C}$. Note that $\hat{C}$ is not uniquely defined as the estimating equation is scale invariant, i.e., if $\hat{C}$ is a solution, then so is $a\hat{C}$ for any $a > 0$. To impose uniqueness, we impose a scale constraint on the solution. Scatter matrices with a suitable scale constraint are commonly called as shape matrices. Hence we say that Tyler’s $M$-estimator of the shape matrix is the unique matrix $\hat{C} \in P^0 = \{ C \in P \mid Tr(C) = 1 \}$ that solves (10). First we note that any matrix $C \in P^0$ can be represented as

$$C = C(q) = \frac{1}{2} \left( \begin{array}{cc} 1 + \text{Re}(q) & \text{Im}(q) \\ \text{Im}(q) & 1 - \text{Re}(q) \end{array} \right)$$

where $q \in \Omega^0$. To verify this, note that $C$ is symmetric with $Tr(C) = 1$ and $\text{det}(C) = 1 - \text{Re}(q)^2 - \text{Im}(q)^2 = 1 - |q|^2 > 0$ iff $q \in \Omega^0$, thus $C(q)$ is positive definite. Using the fact that

$$C^{-1} = \frac{2}{1 - |q|^2} \left( \begin{array}{cc} 1 - \text{Re}(q) & -\text{Im}(q) \\ -\text{Im}(q) & 1 + \text{Re}(q) \end{array} \right)$$

and straightforward algebra, we can rewrite $\hat{\phi}_i$ in complex form as

$$\hat{\phi}_i = \frac{1 - |q|^2}{|z_i|^2 - (z_i^T q)} = \frac{\hat{\phi}_i}{|z_i|}$$

where $\hat{\phi}_i$ is defined by (6). Now multiplying (10) from the left by a vector $(1, j)$ and from the right by a vector $(1, j)^T$, we obtain the equation (7). Furthermore, taking matrix traces of both sides of equation (10) yields $1 = \frac{1}{n} \sum_{i=1}^n \hat{\phi}_i$. Thus the real-valued estimating equation (10) can be recast in complex form as a pair of estimating equations (4) and (5). In other words, if $q$ is the solution to (4) and (5), then $C = C(q)$ is the solution to (10), i.e., $C$ is Tyler’s $M$-estimator of shape and vice versa.

6. REFERENCES


