THE DEGREE OF IMPROPRIETY (NONCIRCULARITY) OF COMPLEX RANDOM VECTORS

Peter J. Schreier

School of Electrical Engineering and Computer Science
The University of Newcastle
Callaghan, NSW 2308, Australia
Peter.Schreier@newcastle.edu.au

ABSTRACT

A complex random vector is called improper (noncircular) if it is correlated with its complex conjugate. We consider measures for the degree of impropriety that are invariant under linear transformation. These measures are functions of the canonical correlations between the vector and its complex conjugate, which have been termed the circularity coefficients. However, we show that these circularity coefficients do not tell the whole story: Two random vectors with identical covariance matrix and identical circularity coefficients can still behave differently in second-order estimation and detection.

Index Terms— Improper complex random vector, widely linear, strong uncorrelating transform, circularity coefficients, canonical correlations.

1. INTRODUCTION

Let $\mathbf{s} \in \mathbb{C}^n$ denote a zero-mean complex random vector with covariance matrix $\mathbf{R} = E\,\mathbf{s}\mathbf{s}^H$ and complementary covariance (or pseudo-covariance) matrix $\widetilde{\mathbf{R}} = E\,\mathbf{s}\mathbf{s}^T$. If $\widetilde{\mathbf{R}} = \mathbf{0}$, $\mathbf{s}$ is called *proper*, otherwise *improper*. Improper complex random vectors and processes have received a great deal of attention in the literature lately: Correctly accounting for impropriety can lead to significant performance gains in many communications and signal processing applications.

In this paper, we are investigating numerical measures for the degree of impropriety of complex random vectors. Because propriety is preserved by linear but not widely linear (linear–conjugate linear) transformations, we require that these measures be invariant under linear transformation. As a consequence, the measures must be functions of the canonical correlations between $\mathbf{s}$ and $\mathbf{s}^*$ [1]. These canonical correlations have been called circularity coefficients by [2].

There is, however, an important caveat. Two random vectors $\mathbf{s}_1$ and $\mathbf{s}_2$ with identical covariance matrix $\mathbf{R}$ and identical circularity coefficients, and hence identical degree of impropriety, can still behave differently in second-order estimation and detection. We will illustrate this with a simple estimation problem.

This work was supported by the Australian Research Council (ARC) under Discovery Project Grant DP0664365.

2. CIRCULARITY COEFFICIENTS

2.1. Augmented algebra

In order to describe the second-order characteristics of $\mathbf{s}$, it is convenient to work with an augmented vector $\underline{\mathbf{s}} = [\mathbf{s}^T, \mathbf{s}^H]^T$ whose covariance matrix

$$\underline{\mathbf{R}} = E\,\underline{\mathbf{s}}\,\underline{\mathbf{s}}^H = \begin{bmatrix} \mathbf{R} & \widetilde{\mathbf{R}} \\ \widetilde{\mathbf{R}}^* & \mathbf{R}^* \end{bmatrix} \tag{1}$$

is called the *augmented covariance matrix* of $\mathbf{s}$ [3]. The advantage of working with this augmented algebra is that it allows access to the vast number of results on 2×2 block matrices.

A transformation of the form $\mathbf{s}' = \underline{\mathbf{A}}_1\mathbf{s} + \underline{\mathbf{A}}_2\mathbf{s}^*$ is called widely linear [4]. It may be represented in augmented form as

$$\underline{\mathbf{s}}' = \begin{bmatrix} \mathbf{s}' \\ \mathbf{s}'^* \end{bmatrix} = \underline{\mathbf{A}}\,\underline{\mathbf{s}} = \begin{bmatrix} \underline{\mathbf{A}}_1 & \underline{\mathbf{A}}_2 \\ \underline{\mathbf{A}}_2^* & \underline{\mathbf{A}}_1^* \end{bmatrix} \begin{bmatrix} \mathbf{s} \\ \mathbf{s}^* \end{bmatrix}. \tag{2}$$

Note that in this augmented representation, $\underline{\mathbf{R}}$ and $\underline{\mathbf{A}} \in \mathbb{C}^{2n\times 2n}$ satisfy a particular block structure where the northwest is the conjugate of the southeast block, and the northeast is the conjugate of the southwest block. In factorizations of $\underline{\mathbf{R}}$, all factors must also have this block structure.

2.2. Strong uncorrelating transform

Propriety is preserved under *linear* (but not widely linear) transformation, which includes rotation and scaling of $\mathbf{s}$. A maximal set of invariants for the augmented covariance matrix $\underline{\mathbf{R}}$ under nonsingular linear transformation is the set of *canonical correlations* [5] between $\mathbf{s}$ and $\mathbf{s}^*$ [1]. This means that any function of $\underline{\mathbf{R}}$ that is invariant under linear transformation must be a function of these canonical correlations.

Following [6], we begin with the *coherence matrix*

$$\mathbf{M} = \mathbf{R}^{-1/2}\widetilde{\mathbf{R}}\mathbf{R}^{-T/2}. \tag{3}$$

Since $\mathbf{M}$ is complex symmetric, $\mathbf{M} = \mathbf{M}^T$, there exists a special singular value decomposition (SVD), called *Takagi's factorization*, which is

$$\mathbf{M} = \mathbf{F}\mathbf{K}\mathbf{F}^T. \tag{4}$$

The complex matrix $\mathbf{F}$ is unitary, and $\mathbf{K} = \mathbf{Diag}(k_1, k_2, ..., k_n)$ contains the canonical correlations $1 \geq k_1 \geq k_2 \geq \cdots \geq k_n \geq 0$ on its diagonal. The latent description $\mathbf{s}' = \mathbf{F}^H\mathbf{R}^{-1/2}\mathbf{s}$ is said to be given in *canonical coordinates*. The canonical coordinates are uncorrelated and have unit variance:

$$E\, s_i' s_j'^* = E\, s_i' s_j' = 0 \text{ for } i \neq j \tag{5}$$

$$E\, |s_i'|^2 = 1. \tag{6}$$

However, they are generally improper as

$$E\, s_i'^2 = k_i. \tag{7}$$

In [2], vectors that are uncorrelated with unit variance, but possibly improper, are called *strongly uncorrelated*, and the transformation $\mathbf{F}^H\mathbf{R}^{-1/2}$, which transforms $\mathbf{s}$ into canonical coordinates $\mathbf{s}'$, is called the *strong uncorrelating transform*. The canonical correlations $\{k_i\}_{i=1}^n$ are referred to as the *circularity coefficients* of $\mathbf{s}$ in [2]. A more thorough discussion of properties of $\{k_i\}$ is contained in [2, 1]. The insight that the $\{k_i\}$ are canonical correlations is critical as it enables us to utilize the many results on this topic in the literature.

The canonical coordinates are given by $\mathbf{s}' = \mathbf{F}^H\mathbf{R}^{-1/2}\mathbf{s}$, so that their complementary correlation matrix is

$$E\, \mathbf{s}'\mathbf{s}'^T = \mathbf{F}^H\mathbf{R}^{-1/2}\widetilde{\mathbf{R}}\mathbf{R}^{-T/2}\mathbf{F}^* = \mathbf{F}^H\mathbf{M}\mathbf{F}^*. \tag{8}$$

In order to make this matrix diagonal, Takagi's factorization (4) rather than the "regular" SVD $\mathbf{M} = \mathbf{U}\mathbf{K}\mathbf{V}^H$ (that generally yields $\mathbf{U} \neq \mathbf{V}^*$) must be employed. It is shown in [7, Sec. 4.4] how to compute the Takagi factorization. If all circularity coefficients are distinct, the matrix $\mathbf{F}$ is the product of the matrix of singular vectors $\mathbf{U}$ with a diagonal unitary matrix.

2.3. Entropy

Combining our results so far and proceeding along the lines of [6], we may factor $\underline{\mathbf{R}}$ as

$$\begin{bmatrix} \mathbf{R} & \widetilde{\mathbf{R}} \\ \widetilde{\mathbf{R}}^* & \mathbf{R}^* \end{bmatrix} = \begin{bmatrix} \mathbf{R}^{1/2} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}^{*/2} \end{bmatrix} \begin{bmatrix} \mathbf{F} & \mathbf{0} \\ \mathbf{0} & \mathbf{F}^* \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{K} \\ \mathbf{K} & \mathbf{I} \end{bmatrix} \times \begin{bmatrix} \mathbf{F}^H & \mathbf{0} \\ \mathbf{0} & \mathbf{F}^T \end{bmatrix} \begin{bmatrix} \mathbf{R}^{H/2} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}^{T/2} \end{bmatrix}. \tag{9}$$

Note that each factor is an augmented matrix. The factorization (9) establishes, similarly to [8],

$$\begin{aligned} \det\underline{\mathbf{R}} &= \det^2\mathbf{R}\,\det(\mathbf{I} - \mathbf{K}^2) \\ &= \det^2\mathbf{R}\prod_{i=1}^{n}(1 - k_i^2). \end{aligned} \tag{10}$$

This allows us to derive the connection between the entropy of an improper Gaussian random vector with augmented covariance matrix $\underline{\mathbf{R}}$ and the corresponding proper Gaussian random vector with covariance matrix $\mathbf{R}$ (see [2, Cor. 2]) in the bargain as

$$\begin{aligned} H_{\text{improper}} &= \frac{1}{2}\log[(\pi e)^{2n}\det\underline{\mathbf{R}}] \\ &= \underbrace{\log[(\pi e)^n \det\mathbf{R}]}_{H_{\text{proper}}} + \frac{1}{2}\log\prod_{i=1}^{n}(1 - k_i^2). \end{aligned} \tag{11}$$

This again shows the classic result that $H_{\text{improper}} \leq H_{\text{proper}}$, and the circularity coefficients determine the loss in entropy.

3. DEGREE OF IMPROPRIETY

We would now like to introduce measures for the degree of impropriety. We want these measures to be invariant under linear transformation because propriety itself is preserved by linear transformation. Thus, these measures must be functions of the circularity coefficients $\{k_i\}$, but several functions seem plausible, for instance:

$$d_1 = 1 - \prod_{i=1}^{n}(1 - k_i^2) = 1 - \det\underline{\mathbf{R}}\det^{-2}\mathbf{R} \tag{12}$$

$$d_2 = \prod_{i=1}^{n} k_i^2 = \det(\widetilde{\mathbf{R}}\mathbf{R}^{-*}\widetilde{\mathbf{R}}^*)\det^{-1}\mathbf{R} \tag{13}$$

$$d_3 = \frac{1}{n}\sum_{i=1}^{n} k_i^2 = \frac{1}{n}\operatorname{tr}(\mathbf{R}^{-1}\widetilde{\mathbf{R}}\mathbf{R}^{-*}\widetilde{\mathbf{R}}^*). \tag{14}$$

These functions have been discussed in [9] as measures of multivariate association between an arbitrary pair of real vectors. They satisfy $0 \leq d_i \leq 1$, and can be defined for reduced rank $r < n$, considering only the r largest circularity coefficients in the computation. When the eigenvalues of $\underline{\mathbf{R}}$ are specified, it is possible to derive tight upper and lower bounds on the degree of impropriety [10].

We consider d_1 the most compelling measure mainly for two reasons. Firstly, as shown above in (11), d_1 connects the entropy of the proper and improper cases. Secondly, it is a measure of the linear dependence between $\mathbf{s}$ and $\mathbf{s}^*$ and as such, can be used to design a generalized likelihood ratio test for impropriety [8, 1].

3.1. Scalar case

In the case of a scalar random variable s with covariance $R = E\,|s|^2$ and complementary covariance $\widetilde{R} = E\,s^2$, the measure

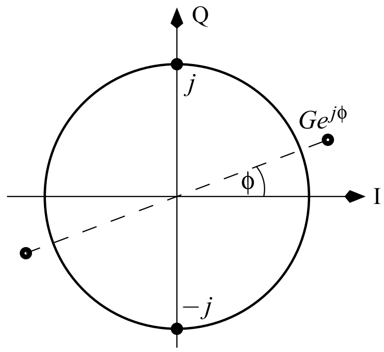

Fig. 1. QPSK with I/Q imbalance.

d_1 becomes particularly simple:

$$d_1 = \frac{|\widetilde{R}|^2}{R^2}. \tag{15}$$

For example, M-PSK, $M \geq 4$, and QAM symbols are proper, $d_1 = 0$, whereas BPSK and PAM symbols are maximally improper, $d_1 = 1$. This reflects the fact that M-PSK and QAM are rotationally symmetric, whereas BPSK and PAM are maximally statistically redundant.

As another example, consider QPSK with I/Q imbalance characterized by gain imbalance (factor) $G > 0$ and quadrature skew ϕ, as depicted in Fig. 1. It is easy to show that

$$d_1 = \frac{|G^2 e^{2j\phi} - 1|^2}{(1+G^2)^2} = \frac{G^4 - 2G^2\cos 2\phi + 1}{(1+G^2)^2}. \tag{16}$$

Clearly, QPSK with perfect I/Q balance has $G = 1$, $\phi = 0$, and thus $d_1 = 0$, whereas the worst possible I/Q imbalance $\phi = \pi/2$ results in $d_1 = 1$.

3.2. Most improper vectors

With specified covariance matrix $\mathbf{R}$, it follows from (3) and (4) that all valid complementary covariance matrices are of the form

$$\widetilde{\mathbf{R}} = \mathbf{R}^{1/2}\mathbf{F}\mathbf{K}\mathbf{F}^T\mathbf{R}^{T/2}, \tag{17}$$

where $\mathbf{F}$ is an arbitrary unitary matrix, and $\mathbf{K}$ is a matrix of arbitrary circularity coefficients.

In the most improper case $\mathbf{K} = \mathbf{I}$, we have $d_1 = d_2 = d_3 = 1$, and (17) becomes

$$\widetilde{\mathbf{R}} = \mathbf{R}^{1/2}\mathbf{F}\mathbf{F}^T\mathbf{R}^{T/2}. \tag{18}$$

Then, the augmented covariance matrix

$$\underline{\mathbf{R}} = \begin{bmatrix} \mathbf{R} & \mathbf{R}^{1/2}\mathbf{F}\mathbf{F}^T\mathbf{R}^{T/2} \\ \mathbf{R}^{*/2}\mathbf{F}^*\mathbf{F}^H\mathbf{R}^{H/2} & \mathbf{R}^* \end{bmatrix} \tag{19}$$

has a zero Schur complement $\mathbf{R} - \widetilde{\mathbf{R}}\mathbf{R}^{-*}\widetilde{\mathbf{R}}^* = \mathbf{0}$, which means that the rank of $\underline{\mathbf{R}}$ equals the rank of $\mathbf{R}$.[1] If $\mathbf{s}$ is most improper, the conjugate $\mathbf{s}^*$ is perfectly *linearly* estimable from $\mathbf{s}$ (obviously, $\mathbf{s}^*$ can always be written as a *widely linear* function of $\mathbf{s}$.)

[1] We assume $\mathbf{R}$ to be invertible, so the rank of $\mathbf{R}$ must be n.

4. THE CIRCULARITY COEFFICIENTS DO NOT TELL THE WHOLE STORY

One might be tempted to assume that two random vectors $\mathbf{s}_1$ and $\mathbf{s}_2$ with the same covariance matrix $\mathbf{R}$ and the same circularity coefficients $\{k_i\}_{i=1}^n$ would behave identically in second-order estimation and detection. This is not the case. We will illustrate this for a simple estimation scenario.

4.1. Widely linear estimation

Consider estimating an improper random message $\mathbf{s}$ with augmented covariance matrix $\underline{\mathbf{R}}$ in complex white (proper) Gaussian noise $\mathbf{n}$ with augmented covariance matrix $N_0\mathbf{I}$. The observations are

$$\mathbf{r} = \mathbf{s} + \mathbf{n}, \tag{20}$$

and $\mathbf{n}$ is assumed to be uncorrelated with $\mathbf{s}$. The optimum widely linear estimator that minimizes the mean squared error can be written in augmented representation as [3]

$$\underline{\hat{\mathbf{s}}} = \underline{\mathbf{R}}(\underline{\mathbf{R}} + N_0\mathbf{I})^{-1}\underline{\mathbf{r}}. \tag{21}$$

The MMSE is [3, 11]

$$\begin{aligned} E\|\mathbf{s} - \hat{\mathbf{s}}\|^2 &= \tfrac{1}{2}\operatorname{tr}\left(\underline{\mathbf{R}} - \underline{\mathbf{R}}(\underline{\mathbf{R}} + N_0\mathbf{I})^{-1}\underline{\mathbf{R}}^H\right) && (22) \\ &= \frac{N_0}{2}\sum_{i=1}^{2n}\frac{\lambda_i}{\lambda_i + N_0}, && (23) \end{aligned}$$

where $\{\lambda_i\}_{i=1}^{2n}$ are the eigenvalues of $\underline{\mathbf{R}}$. For fixed noise level N_0, it follows from [12, 3.C.1] that the MMSE is a Schur-concave function[2] of $\{\lambda_i\}_{i=1}^{2n}$. Hence, if $\mathbf{s}_1$ and $\mathbf{s}_2$ are two random vectors and the eigenvalues of the augmented covariance matrix of $\mathbf{s}_1$ majorize those of $\mathbf{s}_2$, then the MMSE estimating $\mathbf{s}_1$ is less than the MMSE estimating $\mathbf{s}_2$.

Let $\{\lambda_i\}_{i=1}^{2n}$ again denote the eigenvalues of $\underline{\mathbf{R}}$, and $\{\mu_i\}_{i=1}^{n}$ the eigenvalues of $\mathbf{R}$. The following majorization inequality was established by [11]:

$$[\lambda_1, \ldots, \lambda_{2n}]^T \prec [2\mu_1, \ldots, 2\mu_n, 0, \ldots, 0]^T. \tag{24}$$

This says that, for given $\mathbf{R}$, the Schur-concave MMSE (23) is minimized if

$$[\lambda_1, \ldots, \lambda_{2n}]^T = [2\mu_1, \ldots, 2\mu_n, 0, \ldots, 0]^T. \tag{25}$$

[2] see the Appendix for the definitions of majorization and Schur-concave functions

4.2. Most improper vectors

It follows from (25) that in order to minimize the MMSE for a given covariance matrix $\mathbf{R}$, the augmented covariance matrix $\underline{\mathbf{R}}$ must have rank n, i.e., zero Schur complement $\mathbf{R} - \widetilde{\mathbf{R}}\mathbf{R}^{-*}\widetilde{\mathbf{R}}^* = \mathbf{0}$. Thus, $\mathbf{s}$ must be most improper. However, will any most improper vector with $\mathbf{K} = \mathbf{I}$ minimize the MMSE? This is what Schreier et al. [11] claimed but, unfortunately, it is wrong.

The crux of the matter is that the eigenvalues of $\underline{\mathbf{R}}$ given by (19) do not necessarily satisfy (25). This can be shown by simple computation for numerical examples. However, there is indeed *at least one* most improper vector with covariance matrix $\mathbf{R}$ such that the augmented covariance matrix $\underline{\mathbf{R}}$ has the eigenvalues (25). If $\mathbf{R} = \mathbf{U}\mathbf{M}\mathbf{U}^H$ is the eigenvalue decomposition of $\mathbf{R}$, then choosing the complementary covariance as $\widetilde{\mathbf{R}} = \mathbf{U}\mathbf{M}\mathbf{U}^T$ results in $\mathbf{K} = \mathbf{I}$ and $\underline{\mathbf{R}}$ satisfying (25). If $\mathbf{R}$ has repeated eigenvalues, then there will be more than one possible choice for $\widetilde{\mathbf{R}}$.

This shows that there can be two random vectors with the same covariance matrix $\mathbf{R}$ and the same circularity coefficients ($\mathbf{K} = \mathbf{I}$), but different performance in a second-order estimation problem.

5. CONCLUSION

We have investigated measures for the degree of impropriety of a complex random vector. Because propriety is preserved by linear transformation, these measures are functions of the circularity coefficients. Nevertheless, we have shown that the circularity coefficients do not tell the full story. In particular, we have established the following result for most improper random vectors ($\mathbf{K} = \mathbf{I}$): In order to maximize (minimize) a Schur-convex (Schur-concave) function of $\{\lambda_i\}_{i=1}^{2n}$ for fixed $\mathbf{R}$, it is necessary, but not sufficient, that $\mathbf{s}$ be most improper. Examples of Schur-concave and -convex functions are MMSE (as considered above) and deflection (as a performance measure for detection [11]).

6. APPENDIX: MAJORIZATION

This appendix defines majorization and Schur-convex functions. An excellent overview of majorization theory is given in [12].

A vector $\mathbf{x} \in \mathbb{R}^n$ is said to be *majorized* by a vector $\mathbf{y} \in \mathbb{R}^n$, written as $\mathbf{x} \prec \mathbf{y}$, if

$$\sum_{i=1}^{r} x_{[i]} \leq \sum_{i=1}^{r} \mu_{[i]}, \quad r = 1, \ldots, n, \tag{26}$$

with equality for $r = n$. Here, $[\cdot]$ is a permutation such that $x_{[1]} \geq \cdots \geq x_{[n]}$. Intuitively, if $\mathbf{x} \prec \mathbf{y}$, then the components of $\mathbf{x}$ are "less spread out" or "more equal" than the components of $\mathbf{y}$.

The idea of majorization becomes most powerful when it is combined with the concept of Schur-convexity. Functions that are Schur-convex preserve the preordering of majorization. A real-valued function f defined on a set $D \subset \mathbb{R}^n$ is said to be *Schur-convex on D* if

$$\mathbf{x} \prec \mathbf{y} \text{ on } D \Rightarrow f(\mathbf{x}) \leq f(\mathbf{y}), \tag{27}$$

and *Schur-concave on D* if

$$\mathbf{x} \prec \mathbf{y} \text{ on } D \Rightarrow f(\mathbf{x}) \geq f(\mathbf{y}). \tag{28}$$

7. REFERENCES

[1] P. J. Schreier, L. L. Scharf, A. Hanssen, "A generalized likelihood ratio test for impropriety of complex signals," *IEEE Signal Processing Lett.*, vol. 13, no. 7, pp. 433–436, July 2006.

[2] J. Eriksson, V. Koivunen, "Complex random vectors and ICA models: Identifiability, uniqueness, and separability," *IEEE Trans. Inform. Theory*, vol. 52, no. 3, pp. 1017–1029, Mar. 2006.

[3] P. J. Schreier, L. L. Scharf, "Second-order analysis of improper complex random vectors and processes," *IEEE Trans. Signal Processing*, vol. 51, no. 3, pp. 714–725, Mar. 2003.

[4] B. Picinbono, P. Chevalier, "Widely linear estimation with complex data," *IEEE Trans. Signal Processing*, vol. 43, no. 8, pp. 2030–2033, Aug. 1995.

[5] H. Hotelling, "Relations between two sets of variates," *Biometrika*, vol. 28, pp. 321 – 377, 1936.

[6] L. L. Scharf, C. T. Mullis, "Canonical coordinates and the geometry of inference, rate, and capacity," *IEEE Trans. Signal Processing*, vol. 48, no. 3, pp. 824–831, Mar. 2000.

[7] R. A. Horn, C. R. Johnson, *Matrix Analysis*, Cambridge, UK: Cambridge University Press, 1985.

[8] E. Ollila, V. Koivunen, "Generalized complex elliptical distributions," in *Proc. SAM Workshop*, pp. 460–464, July 2004.

[9] E. M. Cramer, W. A. Nicewander, "Some symmetric, invariant measures of multivariate association," *Psychometrika*, vol. 44, no. 1, pp. 43–54, Mar. 1979.

[10] P. J. Schreier, "Bounds on the degree of impropriety," *IEEE Signal Processing Lett.*, submitted.

[11] P. J. Schreier, L. L. Scharf, C. T. Mullis, "Detection and estimation of improper complex random signals," *IEEE Trans. Inform. Theory*, vol. 51, no. 1, pp. 306–312, Jan. 2005.

[12] A. W. Marshall, I. Olkin, *Inequalities: Theory of Majorization and Its Applications*, New York, NY: Academic Press, 1979.