Properness and Widely Linear Processing of Quaternion Random Vectors

Javier Vía, *Member, IEEE*, David Ramírez, *Student Member, IEEE*, and Ignacio Santamaría, *Senior Member, IEEE*

***Abstract*—In this paper, the second-order circularity of quaternion random vectors is analyzed. Unlike the case of complex vectors, there exist three different kinds of quaternion properness, which are based on the vanishing of three different complementary covariance matrices. The different kinds of properness have direct implications on the Cayley–Dickson representation of the quaternion vector, and also on several well-known multivariate statistical analysis methods. In particular, the quaternion extensions of the partial least squares (PLS), multiple linear regression (MLR) and canonical correlation analysis (CCA) techniques are analyzed, showing that, in general, the optimal linear processing is *full-widely linear*. However, in the case of jointly $\mathbb{Q}$-proper or $\mathbb{C}^{\eta}$-proper vectors, the optimal processing reduces, respectively, to the *conventional* or *semi-widely linear* processing. Finally, a measure for the degree of improperness of a quaternion random vector is proposed, which is based on the Kullback–Leibler divergence between two zero-mean Gaussian distributions, one of them with the actual augmented covariance matrix, and the other with its closest proper version. This measure quantifies the entropy loss due to the improperness of the quaternion vector, and it admits an intuitive geometrical interpretation based on Kullback–Leibler projections onto sets of proper augmented covariance matrices.**

***Index Terms*—Canonical correlation analysis (CCA), properness, propriety, quaternions, second-order circularity, widely linear (WL) processing.**

I. Introduction

In recent years, quaternion algebra [1] has been successfully applied to several signal processing and communications problems, such as array processing [2], wave separation [3]–[5], design of orthogonal space-time-polarization block codes [6], and wind forecasting [7]. However, unlike the case of complex vectors [8]–[17], the properness/propriety[1] (or second-order circularity) analysis of quaternion random vectors has received limited attention [4], [5], [18], [19], and a clear definition of quaternion widely linear processing is still lacking [7].

Manuscript received July 27, 2009; revised January 11, 2010. Current version published June 16, 2010. This work was supported by the Spanish Government (MICINN) under projects TEC2007-68020-C04-02/TCM (MultiMIMO) and CONSOLIDER-INGENIO 2010 CSD2008-00010 (COMONSENS), and FPU grant AP2006-2965.

The authors are with the Department of Communications Engineering, University of Cantabria, 39005 Santander, Cantabria, Spain (e-mail: jvia@gtas.dicom.unican.es; ramirezgd@gtas.dicom.unican.es; nacho@gtas.dicom.unican.es).

Communicated by E. Serpedin, Associate Editor for Signal Processing.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2010.2048440

[1]In this paper, we will mainly use the term properness. However, it should be noted that both propriety [14]–[16] and properness [8], [18]–[20] have been used in the literature as synonyms of second-order circularity.

In this paper, we analyze the different kinds of properness for quaternion-valued random vectors, study their implications on optimal linear processing, and provide several measures for the degree of quaternion improperness. In particular, in Section III, we introduce the definition of the complementary covariance matrices, which measure the correlation between the quaternion vector and its involutions over three pure unit quaternions, and show their relationship with the Cayley–Dickson representation of the quaternion vector. Then, we present the definitions of $\mathbb{R}^{\eta}$-properness (cancelation of one complementary covariance matrix), which resembles the properness conditions on the real and imaginary parts of complex vectors; $\mathbb{C}^{\eta}$-properness (cancelation of two complementary covariance matrices), which results in the complex joint-properness of the vectors in the Cayley–Dickson representation; and $\mathbb{Q}$-properness (cancelation of the three complementary covariance matrices), which combines the two previous definitions. The $\mathbb{C}^{\eta}$ and $\mathbb{Q}$ properness definitions in this paper are closely related, but different, to those in [4], [5], [18], and [19]. More precisely, unlike the previous approaches, which are based on the invariance of the second-order statistics (SOS) to left Clifford translations, the definitions in this paper are directly based on the complementary covariance matrices (in analogy with the complex case), and they naturally result in SOS invariance to right Clifford translations. Even more importantly, unlike previous approaches, the proposed kinds of properness are invariant to quaternion linear transformations, i.e., if $\mathbf{x}$ is a proper quaternion vector, then $\mathbf{F}_1^H\mathbf{x}$ (with $\mathbf{F}_1$ a quaternion matrix) is also proper. Analogously to the complex case, the invariance to quaternion linear transformations represents a key property for signal processing applications.

In Section IV, several well-known multivariate statistical analysis methods are generalized to the case of quaternion vectors. Specifically, we show that in the cases of principal component analysis (PCA) [21], partial least squares (PLS) [22], multiple linear regression (MLR) [23] and canonical correlation analysis (CCA) [24], [25], the optimal linear processing is in general *full-widely linear*, which means that we must simultaneously operate on the four real vectors composing the quaternion vector, or equivalently, on the quaternion vector and its three involutions. Interestingly, in the case of jointly $\mathbb{Q}$-proper vectors, the optimal processing is linear, i.e., we do not need to operate on the vector involutions, whereas in the $\mathbb{C}^{\eta}$-proper case, the optimal processing is *semi-widely linear*, which amounts to operate on the quaternion vector and its involution over the pure unit quaternion η. Thus, we can conclude that different kinds of quaternion improperness require different kinds of linear processing.

In Section V, we propose an improperness measure for quaternion random vectors, which is based on the Kullback–Leibler divergence between multivariate quaternion Gaussian distributions. In particular, we consider the divergence between the distribution with the actual augmented covariance matrix, and its Kullback–Leibler projection onto the space of Gaussian proper distributions. Although the different kinds of properness result in different measures, all of them can be obtained from a (generalized) CCA problem [24]–[26], and can be interpreted as the mutual information among the quaternion vector and its involutions. In other words, the proposed measure provides the entropy loss due to the quaternion improperness. Finally, we show that the proposed improperness measure admits a straightforward geometrical interpretation based on projections onto sets of proper augmented covariance matrices. In particular, we illustrate the complementarity of the $\mathbb{R}^\eta$ and $\mathbb{C}^\eta$-properness by showing that the $\mathbb{Q}$-improperness measure can be decomposed as the sum of the $\mathbb{R}^\eta$ and $\mathbb{C}^\eta$ improperness.

II. Preliminaries

A. Notation

Throughout this paper, we will use bold-faced upper case letters to denote matrices, bold-faced lower case letters for column vectors, and light-faced lower case letters for scalar quantities. Superscripts $(\cdot)^*$, $(\cdot)^T$ and $(\cdot)^H$ denote quaternion (or complex) conjugate, transpose and Hermitian (i.e., transpose and quaternion conjugate), respectively. The notation $\mathbf{A} \in \mathbb{R}^{m\times n}$ (respectively $\mathbf{A} \in \mathbb{C}^{m\times n}$ or $\mathbf{A} \in \mathbb{H}^{m\times n}$) means that $\mathbf{A}$ is a real (respectively complex or quaternion) $m \times n$ matrix. $\mathrm{Tr}(\mathbf{A})$ and $|\mathbf{A}|$ denote the trace and determinant of $\mathbf{A}$, $\mathrm{diag}(\mathbf{a})$ is a diagonal matrix with vector $\mathbf{a}$ along its diagonal, $\otimes$ is the Kronecker product, $\mathbf{I}_n$ is the identity matrix of dimension n, and $\mathbf{0}_{m\times n}$ denotes the $m \times n$ zero matrix. Additionally, $\mathbf{A}^{1/2}$ (respectively $\mathbf{A}^{-1/2}$) is the Hermitian square root of the Hermitian matrix $\mathbf{A}$ (respectively $\mathbf{A}^{-1}$). Finally, E is the expectation operator, and in general, $\mathbf{R}_{\mathbf{a},\mathbf{b}}$ is the cross-correlation matrix for vectors $\mathbf{a}$ and $\mathbf{b}$, i.e., $\mathbf{R}_{\mathbf{a},\mathbf{b}} = E\mathbf{a}\mathbf{b}^H$.

B. Properness of Complex Vectors

Let us start by considering a n-dimensional zero-mean[2] complex vector $\mathbf{x} = \mathbf{r}_1 + i\mathbf{r}_i$ with real and imaginary parts $\mathbf{r}_1 \in \mathbb{R}^{n\times 1}$ and $\mathbf{r}_i \in \mathbb{R}^{n\times 1}$, respectively. The second-order statistics (SOS) of $\mathbf{x}$ are given by the covariance $\mathbf{R}_{\mathbf{x},\mathbf{x}} = E\mathbf{x}\mathbf{x}^H$ and complementary covariance $\mathbf{R}_{\mathbf{x},\mathbf{x}^*} = E\mathbf{x}\mathbf{x}^T$ matrices [11], [14], or equivalently by the $2n \times 2n$ augmented covariance matrix [13], [14]

$$\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}} = E\bar{\mathbf{x}}\bar{\mathbf{x}}^H = \begin{bmatrix} \mathbf{R}_{\mathbf{x},\mathbf{x}} & \mathbf{R}_{\mathbf{x},\mathbf{x}^*} \\ \mathbf{R}^*_{\mathbf{x},\mathbf{x}^*} & \mathbf{R}^*_{\mathbf{x},\mathbf{x}} \end{bmatrix}$$

where $\bar{\mathbf{x}} = \left[\mathbf{x}^T, \mathbf{x}^H\right]^T \in \mathbb{C}^{2n\times 1}$ is defined as the augmented complex vector.

[2]Through this paper, we consider zero-mean vectors for notational simplicity. The extension of the results to the nonzero mean case is straightforward.

With the above definitions, the complex vector $\mathbf{x}$ is said to be *proper* (or *second-order circular*) if and only if (iff) [8]

$$\mathbf{R}_{\mathbf{x},\mathbf{x}^*} = \mathbf{0}_{n\times n} \tag{1}$$

i.e., iff $\mathbf{x}$ is uncorrelated with its complex conjugate $\mathbf{x}^*$. Obviously, the definition of a proper complex vector can also be made in terms of the real vectors $\mathbf{r}_1$ and $\mathbf{r}_i$ [16]. In particular, it is easy to check that (1) is equivalent to the two following conditions:

$$\mathbf{R}_{\mathbf{r}_1,\mathbf{r}_1} = \mathbf{R}_{\mathbf{r}_i,\mathbf{r}_i} \tag{2}$$
$$\mathbf{R}_{\mathbf{r}_1,\mathbf{r}_i} = -\mathbf{R}^T_{\mathbf{r}_1,\mathbf{r}_i} \tag{3}$$

which, in the scalar case, reduce to have uncorrelated real and imaginary parts with the same variance. However, in the general vector case, condition (1) provides much more insight than conditions (2) and (3) [14], [17].

The properness definition can be easily extended to the case of two complex random vectors $\mathbf{x} \in \mathbb{C}^{n\times 1}$ and $\mathbf{y} \in \mathbb{C}^{m\times 1}$. In particular, $\mathbf{x}$ and $\mathbf{y}$ are *cross proper* iff the complementary cross-covariance matrix $\mathbf{R}_{\mathbf{x},\mathbf{y}^*} = E\mathbf{x}\mathbf{y}^T$ vanishes. Finally, $\mathbf{x}$ and $\mathbf{y}$ are *jointly proper* iff they are proper and cross proper, or equivalently, iff the composite vector $\left[\mathbf{x}^T, \mathbf{y}^T\right]^T$ is proper [14], [17].

From a practical point of view, the (joint)-properness of random vectors translates into the optimality of *conventional* linear processing. Consider as an example the problem of estimating a vector $\mathbf{y} \in \mathbb{C}^{m\times 1}$ (or its augmented version $\bar{\mathbf{y}}$) from a reduced-rank (with rank r) version of $\mathbf{x} \in \mathbb{C}^{n\times 1}$. In a general case, the optimal linear processing is of the form $\hat{\bar{\mathbf{y}}} = \overline{\mathbf{G}}_{\bar{\mathbf{y}}}\overline{\mathbf{F}}^H_{\bar{\mathbf{x}}}\bar{\mathbf{x}}$, where $\hat{\bar{\mathbf{y}}}$ is the estimate of $\bar{\mathbf{y}}$, $\overline{\mathbf{F}}_{\bar{\mathbf{x}}} \in \mathbb{C}^{2n\times 2r}$ and $\overline{\mathbf{G}}_{\bar{\mathbf{y}}} \in \mathbb{C}^{2m\times 2r}$ are *widely linear operators* given by [14]

$$\overline{\mathbf{F}}_{\bar{\mathbf{x}}} = \begin{bmatrix} \mathbf{F}_1 & \mathbf{F}_i^* \\ \mathbf{F}_i & \mathbf{F}_1^* \end{bmatrix}, \quad \overline{\mathbf{G}}_{\bar{\mathbf{y}}} = \begin{bmatrix} \mathbf{G}_1 & \mathbf{G}_i^* \\ \mathbf{G}_i & \mathbf{G}_1^* \end{bmatrix}$$

$\mathbf{F}_1 \in \mathbb{C}^{n\times r}$ and $\mathbf{F}_i \in \mathbb{C}^{n\times r}$ are the projection matrices, and $\mathbf{G}_1 \in \mathbb{C}^{m\times r}$ and $\mathbf{G}_i \in \mathbb{C}^{m\times r}$ are the reconstruction matrices. The above solution is an example of widely linear processing [10], [14], which is a linear transformation operating on $\bar{\mathbf{x}}$, i.e., both on $\mathbf{x}$ and its conjugate. Obviously, this is a more general processing than that given by the conventional linear transformations. However, if $\mathbf{x}$ and $\mathbf{y}$ are jointly proper, the optimal linear processing takes the form $\hat{\mathbf{y}} = \mathbf{G}_1\mathbf{F}_1^H\mathbf{x}$, i.e., $\mathbf{F}_i = \mathbf{0}_{n\times r}$, $\mathbf{G}_i = \mathbf{0}_{m\times r}$. In other words, the widely linear processing of jointly proper vectors does not provide any advantage over the conventional linear processing [14], [17].

C. Quaternion Algebra

In this subsection, the basic quaternion algebra concepts are briefly reviewed. For an advanced reading on quaternions, we refer to [27], as well as to [3], [28] for several important results on matrices of quaternions.

Quaternions are 4-D hypercomplex numbers invented by Hamilton [1]. A quaternion $x \in \mathbb{H}$ is defined as

$$x = r_1 + ir_i + jr_j + kr_k \tag{4}$$

where r_1, r_i, r_j, r_k are four real numbers, and the imaginary units (i, j, k) satisfy the following properties:

$$\begin{aligned} ij &= k = -ji \\ jk &= i = -kj \\ ki &= j = -ik \\ i^2 &= j^2 = k^2 = ijk = -1. \end{aligned}$$

Quaternions form a noncommutative normed division algebra $\mathbb{H}$, i.e., for $x, y \in \mathbb{H}$, $xy \neq yx$ in general. The conjugate of a quaternion x is $x^* = r_1 - ir_i - jr_j - kr_k$, and the conjugate of the product satisfies $(xy)^* = y^* x^*$. The inner product between two quaternions $x, y \in \mathbb{H}$ is defined as the real part of xy^*, and two quaternions are orthogonal if and only if (iff) their inner product is zero. The quaternion norm is defined as $|x| = \sqrt{xx^*} = \sqrt{r_1^2 + r_i^2 + r_j^2 + r_k^2}$, and it is easy to check that $|xy| = |x||y|$. The inverse of a quaternion $x \neq 0$ is $x^{-1} = x^*/|x|^2$, and we say that η is a pure unit quaternion iff $\eta^2 = -1$ (i.e., iff $|\eta| = 1$ and its real part is zero). Quaternions also admit the Euler representation

$$x = |x|e^{\eta\theta} = |x|\left(\cos\theta + \eta \sin\theta\right)$$

where $\eta = (ir_i + jr_j + kr_k)/\sqrt{r_i^2 + r_j^2 + r_k^2}$ is a pure unit quaternion and $\theta = \arccos(r_1/|x|) \in \mathbb{R}$ is the angle (or argument) of the quaternion. Thus, given an angle θ and a pure unit quaternion η, we can define the left (respectively right) Clifford translation [29] as the product $e^{\eta\theta}x$ (resp. $xe^{\eta\theta}$). Let us now introduce the rotation and involution operations.

Definition 1 (Quaternion Rotation): Consider a quaternion $a = |a|e^{\eta\theta} = |a|\left(\cos\theta + \eta\sin\theta\right)$, then[3]

$$x^{(a)} = axa^{-1}$$

represents a 3-D rotation of the imaginary part of x [27]. In particular, the vector $[r_i, r_j, r_k]^T$ is rotated clockwise an angle 2θ in the pure imaginary plane orthogonal to η.

Definition 2 (Quaternion Involution): The involution of a quaternion x over a pure unit quaternion η is

$$x^{(\eta)} = \eta x \eta^{-1} = \eta x \eta^* = -\eta x \eta$$

and it represents the reflection of x over the plane spanned by $\{1, \eta\}$ [27].

[3]From now on, we will use the notation $\mathbf{A}^{(a)}$ to denote the element-wise rotation of matrix $\mathbf{A}$.

With the above definitions, and given two quaternions $a, b \in \mathbb{H}$, it is easy to check the following properties [4], [5]:

$$\begin{aligned} x^{(a)*} &= (x^*)^{(a)}, & \forall\, x \in \mathbb{H} \\ \left(x^{(a)}\right)^{(b)} &= x^{(ba)}, & \forall\, x \in \mathbb{H} \\ (xy)^{(a)} &= x^{(a)}y^{(a)}, & \forall\, x, y \in \mathbb{H} \\ ab &= b^{(a)}a = ba^{(b^*)}. \end{aligned}$$

Here we must point out that the real representation in (4) can be easily generalized to other orthogonal bases. In particular, we will consider an orthogonal system $\{1, \eta, \eta', \eta''\}$ given by

$$\begin{bmatrix} 1 \\ \eta \\ \eta' \\ \eta'' \end{bmatrix} = \begin{bmatrix} 1 & \mathbf{0}_{1\times 3} \\ \mathbf{0}_{3\times 1} & \mathbf{Q} \end{bmatrix} \begin{bmatrix} 1 \\ i \\ j \\ k \end{bmatrix}$$

where $\mathbf{Q} \in \mathbb{R}^{3\times 3}$ is an orthogonal matrix, i.e., $\mathbf{Q}^T\mathbf{Q} = \mathbf{I}_3$. Furthermore, we will assume that the signs of the rows of $\mathbf{Q}$ are chosen in order to ensure

$$\begin{aligned} \eta\eta' &= \eta'' = -\eta'\eta \\ \eta'\eta'' &= \eta = -\eta''\eta' \\ \eta''\eta &= \eta' = -\eta\eta'' \\ \eta^2 &= \eta'^2 = \eta''^2 = \eta\eta'\eta'' = -1. \end{aligned}$$

Thus, any quaternion can be represented as

$$x = r_1 + \eta r_\eta + \eta' r_{\eta'} + \eta'' r_{\eta''} \tag{5}$$

where $[r_\eta, r_{\eta'}, r_{\eta''}] = [r_i, r_j, r_k]\,\mathbf{Q}^T$. Moreover, we can use the following *modified Cayley–Dickson* representations

$$x = a_1 + \eta'' a_2, \quad x = b_1 + \eta b_2, \quad x = c_1 + \eta' c_2 \tag{6}$$

where

$$\begin{aligned} &a_1 = r_1 + \eta r_\eta, \quad a_2 = r_{\eta''} + \eta r_{\eta'}, \quad b_1 = r_1 + \eta' r_{\eta'} \\ &b_2 = r_\eta + \eta' r_{\eta''}, \quad c_1 = r_1 + \eta'' r_{\eta''}, \quad c_2 = r_{\eta'} + \eta'' r_\eta \end{aligned}$$

can be seen as complex numbers in the planes spanned by $\{1, \eta\}$, $\{1, \eta'\}$ or $\{1, \eta''\}$.

Finally, it is important to note that the Cayley–Dickson representations in (6) differ from those in [4], [5], [18], and [19].[4] Although this is only a notational difference, we will see later that the choice of the formulation in (6) results in a clear relationship between the quaternion properness definitions and the statistical properties of the complex vectors in the Cayley–Dickson representation.

[4]In particular, the Cayley–Dickson representations in the cited papers can be rewritten as $x = a_1 + a_3\eta'' = b_1 + b_3\eta = c_1 + c_3\eta'$, with $a_3 = a_2^*, b_3 = b_2^*$ and $c_3 = c_2^*$. Therefore, the results in this paper can be easily rewritten in terms of these alternative Cayley–Dickson formulas.

TABLE I
CORRESPONDENCE BETWEEN THE QUATERNION COVARIANCE MATRICES AND THE REAL AND COMPLEX (CROSS)-COVARIANCES

	Expression based on the real representation	Expressions based on the Cayley-Dickson representations
Covariance Matrix $\mathbf{R}_{\mathbf{x},\mathbf{x}} = E\mathbf{x}\mathbf{x}^H$	$\mathbf{R}_{\mathbf{x},\mathbf{x}} = \left(\mathbf{R}_{\mathbf{r}_1,\mathbf{r}_1} + \mathbf{R}_{\mathbf{r}_\eta,\mathbf{r}_\eta} + \mathbf{R}_{\mathbf{r}_{\eta'},\mathbf{r}_{\eta'}} + \mathbf{R}_{\mathbf{r}_{\eta''},\mathbf{r}_{\eta''}}\right)$ $+\eta\left(\mathbf{R}^T_{\mathbf{r}_1,\mathbf{r}_\eta} - \mathbf{R}_{\mathbf{r}_1,\mathbf{r}_\eta} - \mathbf{R}_{\mathbf{r}_{\eta'},\mathbf{r}_{\eta''}} + \mathbf{R}^T_{\mathbf{r}_{\eta'},\mathbf{r}_{\eta''}}\right)$ $+\eta'\left(\mathbf{R}^T_{\mathbf{r}_1,\mathbf{r}_{\eta'}} - \mathbf{R}_{\mathbf{r}_1,\mathbf{r}_{\eta'}} + \mathbf{R}_{\mathbf{r}_\eta,\mathbf{r}_{\eta''}} - \mathbf{R}^T_{\mathbf{r}_\eta,\mathbf{r}_{\eta''}}\right)$ $+\eta''\left(\mathbf{R}^T_{\mathbf{r}_1,\mathbf{r}_{\eta''}} - \mathbf{R}_{\mathbf{r}_1,\mathbf{r}_{\eta''}} - \mathbf{R}_{\mathbf{r}_\eta,\mathbf{r}_{\eta'}} + \mathbf{R}^T_{\mathbf{r}_\eta,\mathbf{r}_{\eta'}}\right)$	$\mathbf{R}_{\mathbf{x},\mathbf{x}} = \left(\mathbf{R}_{\mathbf{a}_1,\mathbf{a}_1} + \mathbf{R}^*_{\mathbf{a}_2,\mathbf{a}_2}\right) + \eta''\left(\mathbf{R}^H_{\mathbf{a}_1,\mathbf{a}_2} - \mathbf{R}^*_{\mathbf{a}_1,\mathbf{a}_2}\right)$ $= \left(\mathbf{R}_{\mathbf{b}_1,\mathbf{b}_1} + \mathbf{R}^*_{\mathbf{b}_2,\mathbf{b}_2}\right) + \eta\left(\mathbf{R}^H_{\mathbf{b}_1,\mathbf{b}_2} - \mathbf{R}^*_{\mathbf{b}_1,\mathbf{b}_2}\right)$ $= \left(\mathbf{R}_{\mathbf{c}_1,\mathbf{c}_1} + \mathbf{R}^*_{\mathbf{c}_2,\mathbf{c}_2}\right) + \eta'\left(\mathbf{R}^H_{\mathbf{c}_1,\mathbf{c}_2} - \mathbf{R}^*_{\mathbf{c}_1,\mathbf{c}_2}\right)$
Complementary Covariance Matrix $\mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta)}} = E\mathbf{x}\mathbf{x}^{(\eta)H}$	$\mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta)}} = \left(\mathbf{R}_{\mathbf{r}_1,\mathbf{r}_1} + \mathbf{R}_{\mathbf{r}_\eta,\mathbf{r}_\eta} - \mathbf{R}_{\mathbf{r}_{\eta'},\mathbf{r}_{\eta'}} - \mathbf{R}_{\mathbf{r}_{\eta''},\mathbf{r}_{\eta''}}\right)$ $+\eta\left(\mathbf{R}^T_{\mathbf{r}_1,\mathbf{r}_\eta} - \mathbf{R}_{\mathbf{r}_1,\mathbf{r}_\eta} + \mathbf{R}_{\mathbf{r}_{\eta'},\mathbf{r}_{\eta''}} - \mathbf{R}^T_{\mathbf{r}_{\eta'},\mathbf{r}_{\eta''}}\right)$ $+\eta'\left(\mathbf{R}^T_{\mathbf{r}_1,\mathbf{r}_{\eta'}} + \mathbf{R}_{\mathbf{r}_1,\mathbf{r}_{\eta'}} - \mathbf{R}_{\mathbf{r}_\eta,\mathbf{r}_{\eta''}} - \mathbf{R}^T_{\mathbf{r}_\eta,\mathbf{r}_{\eta''}}\right)$ $+\eta''\left(\mathbf{R}^T_{\mathbf{r}_1,\mathbf{r}_{\eta''}} + \mathbf{R}_{\mathbf{r}_1,\mathbf{r}_{\eta''}} + \mathbf{R}_{\mathbf{r}_\eta,\mathbf{r}_{\eta'}} + \mathbf{R}^T_{\mathbf{r}_\eta,\mathbf{r}_{\eta'}}\right)$	$\mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta)}} = \left(\mathbf{R}_{\mathbf{a}_1,\mathbf{a}_1} - \mathbf{R}^*_{\mathbf{a}_2,\mathbf{a}_2}\right) + \eta''\left(\mathbf{R}^H_{\mathbf{a}_1,\mathbf{a}_2} + \mathbf{R}^*_{\mathbf{a}_1,\mathbf{a}_2}\right)$ $= \left(\mathbf{R}_{\mathbf{b}_1,\mathbf{b}_1^*} + \mathbf{R}^*_{\mathbf{b}_2,\mathbf{b}_2^*}\right) + \eta\left(\mathbf{R}^T_{\mathbf{b}_1,\mathbf{b}_2^*} - \mathbf{R}^*_{\mathbf{b}_1,\mathbf{b}_2^*}\right)$ $= \left(\mathbf{R}_{\mathbf{c}_1,\mathbf{c}_1^*} - \mathbf{R}^*_{\mathbf{c}_2,\mathbf{c}_2^*}\right) + \eta'\left(\mathbf{R}^T_{\mathbf{c}_1,\mathbf{c}_2^*} + \mathbf{R}^*_{\mathbf{c}_1,\mathbf{c}_2^*}\right)$
Complementary Covariance Matrix $\mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta')}} = E\mathbf{x}\mathbf{x}^{(\eta')H}$	$\mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta')}} = \left(\mathbf{R}_{\mathbf{r}_1,\mathbf{r}_1} - \mathbf{R}_{\mathbf{r}_\eta,\mathbf{r}_\eta} + \mathbf{R}_{\mathbf{r}_{\eta'},\mathbf{r}_{\eta'}} - \mathbf{R}_{\mathbf{r}_{\eta''},\mathbf{r}_{\eta''}}\right)$ $+\eta\left(\mathbf{R}^T_{\mathbf{r}_1,\mathbf{r}_\eta} + \mathbf{R}_{\mathbf{r}_1,\mathbf{r}_\eta} + \mathbf{R}_{\mathbf{r}_{\eta'},\mathbf{r}_{\eta''}} + \mathbf{R}^T_{\mathbf{r}_{\eta'},\mathbf{r}_{\eta''}}\right)$ $+\eta'\left(\mathbf{R}^T_{\mathbf{r}_1,\mathbf{r}_{\eta'}} - \mathbf{R}_{\mathbf{r}_1,\mathbf{r}_{\eta'}} + \mathbf{R}_{\mathbf{r}_\eta,\mathbf{r}_{\eta''}} - \mathbf{R}^T_{\mathbf{r}_\eta,\mathbf{r}_{\eta''}}\right)$ $+\eta''\left(\mathbf{R}^T_{\mathbf{r}_1,\mathbf{r}_{\eta''}} + \mathbf{R}_{\mathbf{r}_1,\mathbf{r}_{\eta''}} - \mathbf{R}_{\mathbf{r}_\eta,\mathbf{r}_{\eta'}} - \mathbf{R}^T_{\mathbf{r}_\eta,\mathbf{r}_{\eta'}}\right)$	$\mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta')}} = \left(\mathbf{R}_{\mathbf{a}_1,\mathbf{a}_1^*} - \mathbf{R}^*_{\mathbf{a}_2,\mathbf{a}_2^*}\right) + \eta''\left(\mathbf{R}^T_{\mathbf{a}_1,\mathbf{a}_2^*} + \mathbf{R}^*_{\mathbf{a}_1,\mathbf{a}_2^*}\right)$ $= \left(\mathbf{R}_{\mathbf{b}_1,\mathbf{b}_1} - \mathbf{R}^*_{\mathbf{b}_2,\mathbf{b}_2}\right) + \eta\left(\mathbf{R}^H_{\mathbf{b}_1,\mathbf{b}_2} + \mathbf{R}^*_{\mathbf{b}_1,\mathbf{b}_2}\right)$ $= \left(\mathbf{R}_{\mathbf{c}_1,\mathbf{c}_1^*} + \mathbf{R}^*_{\mathbf{c}_2,\mathbf{c}_2^*}\right) + \eta'\left(\mathbf{R}^T_{\mathbf{c}_1,\mathbf{c}_2^*} - \mathbf{R}^*_{\mathbf{c}_1,\mathbf{c}_2^*}\right)$
Complementary Covariance Matrix $\mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta'')}} = E\mathbf{x}\mathbf{x}^{(\eta'')H}$	$\mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta'')}} = \left(\mathbf{R}_{\mathbf{r}_1,\mathbf{r}_1} - \mathbf{R}_{\mathbf{r}_\eta,\mathbf{r}_\eta} - \mathbf{R}_{\mathbf{r}_{\eta'},\mathbf{r}_{\eta'}} + \mathbf{R}_{\mathbf{r}_{\eta''},\mathbf{r}_{\eta''}}\right)$ $+\eta\left(\mathbf{R}^T_{\mathbf{r}_1,\mathbf{r}_\eta} + \mathbf{R}_{\mathbf{r}_1,\mathbf{r}_\eta} - \mathbf{R}_{\mathbf{r}_{\eta'},\mathbf{r}_{\eta''}} - \mathbf{R}^T_{\mathbf{r}_{\eta'},\mathbf{r}_{\eta''}}\right)$ $+\eta'\left(\mathbf{R}^T_{\mathbf{r}_1,\mathbf{r}_{\eta'}} + \mathbf{R}_{\mathbf{r}_1,\mathbf{r}_{\eta'}} + \mathbf{R}_{\mathbf{r}_\eta,\mathbf{r}_{\eta''}} + \mathbf{R}^T_{\mathbf{r}_\eta,\mathbf{r}_{\eta''}}\right)$ $+\eta''\left(\mathbf{R}^T_{\mathbf{r}_1,\mathbf{r}_{\eta''}} - \mathbf{R}_{\mathbf{r}_1,\mathbf{r}_{\eta''}} + \mathbf{R}_{\mathbf{r}_\eta,\mathbf{r}_{\eta'}} - \mathbf{R}^T_{\mathbf{r}_\eta,\mathbf{r}_{\eta'}}\right)$	$\mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta'')}} = \left(\mathbf{R}_{\mathbf{a}_1,\mathbf{a}_1^*} + \mathbf{R}^*_{\mathbf{a}_2,\mathbf{a}_2^*}\right) + \eta''\left(\mathbf{R}^T_{\mathbf{a}_1,\mathbf{a}_2^*} - \mathbf{R}^*_{\mathbf{a}_1,\mathbf{a}_2^*}\right)$ $= \left(\mathbf{R}_{\mathbf{b}_1,\mathbf{b}_1^*} - \mathbf{R}^*_{\mathbf{b}_2,\mathbf{b}_2^*}\right) + \eta\left(\mathbf{R}^T_{\mathbf{b}_1,\mathbf{b}_2^*} + \mathbf{R}^*_{\mathbf{b}_1,\mathbf{b}_2^*}\right)$ $= \left(\mathbf{R}_{\mathbf{c}_1,\mathbf{c}_1} - \mathbf{R}^*_{\mathbf{c}_2,\mathbf{c}_2}\right) + \eta'\left(\mathbf{R}^H_{\mathbf{c}_1,\mathbf{c}_2} + \mathbf{R}^*_{\mathbf{c}_1,\mathbf{c}_2^*}\right)$

III. PROPERNESS OF QUATERNION VECTORS

A. Augmented Covariance Matrix

Analogously to the case of complex vectors, the circularity analysis of a n-dimensional quaternion random vector $\mathbf{x} = \mathbf{r}_1 + \eta\mathbf{r}_\eta + \eta'\mathbf{r}_{\eta'} + \eta''\mathbf{r}_{\eta''}$ can be based on the real vectors $\mathbf{r}_1$, $\mathbf{r}_\eta$, $\mathbf{r}_{\eta'}$ and $\mathbf{r}_{\eta''}$ [18]. However, here we follow a similar derivation to that in [19] for the case of scalar quaternions. In particular, we define the augmented quaternion vector as $\bar{\mathbf{x}} = \left[\mathbf{x}^T, \mathbf{x}^{(\eta)^T}, \mathbf{x}^{(\eta')^T}, \mathbf{x}^{(\eta'')^T}\right]^T$, whose relationship with the real vectors is given by

$$\bar{\mathbf{x}} = 2\mathbf{T}_n\mathbf{r}_\mathbf{x}$$

where $\mathbf{r}_\mathbf{x} = \left[\mathbf{r}_1^T, \mathbf{r}_\eta^T, \mathbf{r}_{\eta'}^T, \mathbf{r}_{\eta''}^T\right]^T$, and

$$\mathbf{T}_n = \frac{1}{2}\begin{bmatrix} +1 & +\eta & +\eta' & +\eta'' \\ +1 & +\eta & -\eta' & -\eta'' \\ +1 & -\eta & +\eta' & -\eta'' \\ +1 & -\eta & -\eta' & +\eta'' \end{bmatrix} \otimes \mathbf{I}_n \tag{7}$$

is a unitary quaternion operator, i.e., $\mathbf{T}_n^H\mathbf{T}_n = \mathbf{I}_{4n}$.

Based on the above definitions, we can introduce the augmented covariance matrix

$$\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}} = \begin{bmatrix} \mathbf{R}_{\mathbf{x},\mathbf{x}} & \mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta)}} & \mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta')}} & \mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta'')}} \\ \mathbf{R}^{(\eta)}_{\mathbf{x},\mathbf{x}^{(\eta)}} & \mathbf{R}^{(\eta)}_{\mathbf{x},\mathbf{x}} & \mathbf{R}^{(\eta)}_{\mathbf{x},\mathbf{x}^{(\eta'')}} & \mathbf{R}^{(\eta)}_{\mathbf{x},\mathbf{x}^{(\eta')}} \\ \mathbf{R}^{(\eta')}_{\mathbf{x},\mathbf{x}^{(\eta')}} & \mathbf{R}^{(\eta')}_{\mathbf{x},\mathbf{x}^{(\eta'')}} & \mathbf{R}^{(\eta')}_{\mathbf{x},\mathbf{x}} & \mathbf{R}^{(\eta')}_{\mathbf{x},\mathbf{x}^{(\eta)}} \\ \mathbf{R}^{(\eta'')}_{\mathbf{x},\mathbf{x}^{(\eta'')}} & \mathbf{R}^{(\eta'')}_{\mathbf{x},\mathbf{x}^{(\eta')}} & \mathbf{R}^{(\eta'')}_{\mathbf{x},\mathbf{x}^{(\eta)}} & \mathbf{R}^{(\eta'')}_{\mathbf{x},\mathbf{x}} \end{bmatrix}$$

where we can readily identify the covariance matrix $\mathbf{R}_{\mathbf{x},\mathbf{x}} = E\mathbf{x}\mathbf{x}^H$ and three complementary covariance matrices $\mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta)}} = E\mathbf{x}\mathbf{x}^{(\eta)^H}$, $\mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta')}} = E\mathbf{x}\mathbf{x}^{(\eta')^H}$ and $\mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta'')}} = E\mathbf{x}\mathbf{x}^{(\eta'')^H}$. The relationship among these matrices, the real representation in (5), and the Cayley–Dickson representations in (6), can be obtained by means of straightforward but tedious algebra, and are summarized inTable I.

As we have previously pointed out, the different definitions of quaternion properness are based on the cancelation of the complementary covariance matrices. However, before proceeding, we must introduce the following lemmas, which present three key properties of the augmented covariance matrix.

Lemma 1: The structure (location of zero complementary covariance matrices) of $\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}$ is invariant to linear transformations[5] of the form $\mathbf{u} = \mathbf{F}_1^H \mathbf{x}$, with $\mathbf{F}_1 \in \mathbb{H}^{n\times r}$.

Proof: It can be easily checked that $\mathbf{R}_{\mathbf{u},\mathbf{u}} = \mathbf{F}_1^H \mathbf{R}_{\mathbf{x},\mathbf{x}} \mathbf{F}_1$ and $\mathbf{R}_{\mathbf{u},\mathbf{u}^{(a)}} = \mathbf{F}_1^H \mathbf{R}_{\mathbf{x},\mathbf{x}^{(a)}} \mathbf{F}_1^{(a)}$, $\forall\, a \in \mathbb{H}$. The proof concludes particularizing a for η, η', η''. ∎

Lemma 2: A rotation $\mathbf{u} = \mathbf{x}^{(a)}$ results in a simultaneous rotation of the orthogonal basis $\{1, \eta, \eta', \eta''\}$ and the augmented covariance matrix

$$\mathbf{R}_{\bar{\mathbf{u}},\bar{\mathbf{u}}}(\{1,\eta,\eta',\eta''\}) = \mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}^{(a)}(\{1,\eta^{(a^*)},\eta'^{(a^*)},\eta''^{(a^*)}\})$$

where the expressions in parentheses make explicit the bases for the augmented covariance matrices.

Proof: The covariance matrix can be easily obtained as $\mathbf{R}_{\mathbf{u},\mathbf{u}} = Ea\mathbf{x}a^{-1}a\mathbf{x}^H a^{-1} = a\mathbf{R}_{\mathbf{x},\mathbf{x}}a^{-1} = \mathbf{R}_{\mathbf{x},\mathbf{x}}^{(a)}$. On the other hand, $\forall\, b \in \mathbb{H}$, we have

$$\mathbf{R}_{\mathbf{u},\mathbf{u}^{(b)}} = Ea\mathbf{x}a^{-1}ba\mathbf{x}^H a^{-1}b^{-1}$$

and right-multiplying by aa^{-1}, we obtain

$$\mathbf{R}_{\mathbf{u},\mathbf{u}^{(b)}} = a\left(E\mathbf{x}a^{-1}ba\mathbf{x}^H a^{-1}b^{-1}a\right)a^{-1} = \mathbf{R}_{\mathbf{x},\mathbf{x}^{(b^{(a^*)})}}^{(a)}.$$

The proof concludes particularizing b for η, η', and η''. ∎

Lemma 3: The augmented covariance matrices in two different orthogonal bases are related as

$$\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}(\{1,\nu,\nu',\nu''\}) = \mathbf{\Gamma}\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}(\{1,\eta,\eta',\eta''\})\mathbf{\Gamma}^H$$

where

$$\mathbf{\Gamma} = \begin{bmatrix} 1 & \mathbf{0}_{1\times 3} \\ \mathbf{0}_{3\times 1} & \mathbf{\Lambda}_\nu \mathbf{Q}\mathbf{\Lambda}_\eta^H \end{bmatrix} \otimes \mathbf{I}_n$$

$\mathbf{Q} \in \mathbb{R}^{3\times 3}$ is the matrix for the change of basis $[\nu,\nu',\nu''] = [\eta,\eta',\eta'']\mathbf{Q}^T$, $\mathbf{\Lambda}_\nu = \mathrm{diag}([\nu,\nu',\nu''])$, and $\mathbf{\Lambda}_\eta = \mathrm{diag}([\eta,\eta',\eta''])$.

Proof: Let us consider the pure unit quaternion $\nu = \eta\nu_\eta + \eta'\nu_{\eta'} + \eta''\nu_{\eta''}$, where $[\nu_\eta, \nu_{\eta'}, \nu_{\eta''}]$ is the first row of $\mathbf{Q}$. Thus, the involution of $\mathbf{x}$ over ν is

$$\begin{aligned}\mathbf{x}^{(\nu)} &= \nu\mathbf{x}\nu^* = \nu\mathbf{x}(\eta^*\nu_\eta + \eta'^*\nu_{\eta'} + \eta''^*\nu_{\eta''})\\ &= \nu\nu_\eta\eta^*\mathbf{x}^{(\eta)} + \nu\nu_{\eta'}\eta'^*\mathbf{x}^{(\eta')} + \nu\nu_{\eta''}\eta''^*\mathbf{x}^{(\eta'')}.\end{aligned}$$

Repeating this procedure for ν' and ν'', we obtain the mapping between the augmented quaternion vectors in the two different bases

$$\begin{bmatrix} \mathbf{x} \\ \mathbf{x}^{(\nu)} \\ \mathbf{x}^{(\nu')} \\ \mathbf{x}^{(\nu'')} \end{bmatrix} = \mathbf{\Gamma} \begin{bmatrix} \mathbf{x} \\ \mathbf{x}^{(\eta)} \\ \mathbf{x}^{(\eta')} \\ \mathbf{x}^{(\eta'')} \end{bmatrix}.$$

Finally, as a direct consequence of the previous relationship, we have $\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}(\{1,\nu,\nu',\nu''\}) = \mathbf{\Gamma}\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}(\{1,\eta,\eta',\eta''\})\mathbf{\Gamma}^H$. ∎

Lemma 1 ensures the invariance of the structure of $\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}$ to linear transformations, which will translate into the invariance of the properness definitions. On the other hand, Lemma 2 states that, taking into account the rotation of the orthogonal basis $\{1,\eta,\eta',\eta''\}$, the structure of the augmented covariance matrix is also invariant to rotations, which include involutions as a particular case. This property will allow us to easily relate the properness of the original quaternion vector with that of its rotated version. Finally, Lemma 3 shows that the complementary covariance matrices in an arbitrary base $\{1,\nu,\nu',\nu''\}$ can be easily obtained as quaternion linear combinations of $\mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta)}}$, $\mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta')}}$ and $\mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta'')}}$. From our point of view, these nice properties justify the use of the augmented covariance matrix instead of other cross-covariance matrices based on the real or Cayley–Dickson representations [4], [5].

B. $\mathbb{R}^\eta$-Properness

Let us start by the weakest properness definition.

Definition 3 ($\mathbb{R}^\eta$-Properness): A quaternion random vector $\mathbf{x}$ is $\mathbb{R}^\eta$-proper iff the complementary covariance matrix $\mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta)}}$ vanishes.

To our best knowledge, the definition of $\mathbb{R}^\eta$-proper vectors is completely new. Obviously, it translates into the following structure in the augmented covariance matrix

$$\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}} = \begin{bmatrix} \mathbf{R}_{\mathbf{x},\mathbf{x}} & \mathbf{0}_{n\times n} & \mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta')}} & \mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta'')}} \\ \mathbf{0}_{n\times n} & \mathbf{R}_{\mathbf{x},\mathbf{x}}^{(\eta)} & \mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta'')}}^{(\eta)} & \mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta')}}^{(\eta)} \\ \mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta')}}^{(\eta')} & \mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta'')}}^{(\eta')} & \mathbf{R}_{\mathbf{x},\mathbf{x}}^{(\eta')} & \mathbf{0}_{n\times n} \\ \mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta'')}}^{(\eta'')} & \mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta')}}^{(\eta'')} & \mathbf{0}_{n\times n} & \mathbf{R}_{\mathbf{x},\mathbf{x}}^{(\eta'')} \end{bmatrix}$$

and its main implication can be established with the help of the Cayley–Dickson representation summarized in Table I. In particular, we can see that a quaternion vector $\mathbf{x}$ is $\mathbb{R}^\eta$-proper iff

$$\mathbf{R}_{\mathbf{a}_1,\mathbf{a}_1} = \mathbf{R}_{\mathbf{a}_2,\mathbf{a}_2}^T \quad (8)$$

$$\mathbf{R}_{\mathbf{a}_1,\mathbf{a}_2} = -\mathbf{R}_{\mathbf{a}_1,\mathbf{a}_2}^T \quad (9)$$

which can be seen as the complex analogue of the conditions in (2) and (3) for the real and imaginary parts of a complex proper vector. From a practical point of view, the implications of this kind of properness are rather limited. In particular, unlike the $\mathbb{C}^\eta$ and $\mathbb{Q}$ properness, it does not translate into a simplified kind of quaternion linear signal processing, and neither implies the invariance of all the SOS of $\mathbf{x}$ to a right Clifford translation. However, the next lemma proves the equivalence between $\mathbb{R}^\eta$-properness and a relaxed[6] kind of SOS invariance.

Lemma 4: A quaternion random vector $\mathbf{x} = \mathbf{a}_1 + \eta''\mathbf{a}_2$ is $\mathbb{R}^\eta$-proper iff the covariance $\mathbf{R}_{\mathbf{a}_1,\mathbf{a}_1}$, $\mathbf{R}_{\mathbf{a}_2,\mathbf{a}_2}$ and cross covariance $\mathbf{R}_{\mathbf{a}_1,\mathbf{a}_2}$ matrices are invariant to a right multiplication by the pure unit quaternion η''.

Proof: As a result of the right product, we have

$$\mathbf{x}\eta'' = \mathbf{g}_1 + \eta''\mathbf{g}_2 = -\mathbf{a}_2^* + \eta''\mathbf{a}_1^*$$

[5]In this paper, we focus on left multiplications, which agrees with most of the quaternion signal processing literature [2], [3], [7].

[6]Note that Lemma 4 only considers right Clifford translations with angle $\pi/2$ $\left(x\eta'' = xe^{\eta''\frac{\pi}{2}}\right)$, and it does not ensure the invariance of the second order statistics given by $\mathbf{R}_{\mathbf{a}_1,\mathbf{a}_1^*}$, $\mathbf{R}_{\mathbf{a}_2,\mathbf{a}_2^*}$, and $\mathbf{R}_{\mathbf{a}_2,\mathbf{a}_2^*}$.

and the new covariance and cross-covariance matrices are

$$\begin{aligned}\mathbf{R}_{\mathbf{g}_1,\mathbf{g}_1} &= \mathbf{R}_{\mathbf{a}_2^*,\mathbf{a}_2^*} = \mathbf{R}_{\mathbf{a}_2,\mathbf{a}_2}^T\\ \mathbf{R}_{\mathbf{g}_2,\mathbf{g}_2} &= \mathbf{R}_{\mathbf{a}_1^*,\mathbf{a}_1^*} = \mathbf{R}_{\mathbf{a}_1,\mathbf{a}_1}^T\\ \mathbf{R}_{\mathbf{g}_1,\mathbf{g}_2} &= -\mathbf{R}_{\mathbf{a}_2^*,\mathbf{a}_1^*} = -\mathbf{R}_{\mathbf{a}_1,\mathbf{a}_2}^T.\end{aligned}$$

Obviously, the covariance and cross-covariance matrices are invariant to the product $\mathbf{x}\eta''$ iff $\mathbf{R}_{\mathbf{g}_1,\mathbf{g}_1} = \mathbf{R}_{\mathbf{a}_1,\mathbf{a}_1}$, $\mathbf{R}_{\mathbf{g}_2,\mathbf{g}_2} = \mathbf{R}_{\mathbf{a}_2,\mathbf{a}_2}$ and $\mathbf{R}_{\mathbf{g}_1,\mathbf{g}_2} = \mathbf{R}_{\mathbf{a}_1,\mathbf{a}_2}$. Thus, we have

$$\begin{aligned}\mathbf{R}_{\mathbf{g}_1,\mathbf{g}_1} &= \mathbf{R}_{\mathbf{a}_1,\mathbf{a}_1} \Leftrightarrow \mathbf{R}_{\mathbf{a}_1,\mathbf{a}_1} = \mathbf{R}_{\mathbf{a}_2,\mathbf{a}_2}^T\\ \mathbf{R}_{\mathbf{g}_2,\mathbf{g}_2} &= \mathbf{R}_{\mathbf{a}_2,\mathbf{a}_2} \Leftrightarrow \mathbf{R}_{\mathbf{a}_1,\mathbf{a}_1} = \mathbf{R}_{\mathbf{a}_2,\mathbf{a}_2}^T\\ \mathbf{R}_{\mathbf{g}_1,\mathbf{g}_2} &= \mathbf{R}_{\mathbf{a}_1,\mathbf{a}_2} \Leftrightarrow \mathbf{R}_{\mathbf{a}_1,\mathbf{a}_2} = -\mathbf{R}_{\mathbf{a}_1,\mathbf{a}_2}^T\end{aligned}$$

which are the necessary and sufficient conditions for $\mathbb{R}^\eta$-properness given in (8) and (9). ∎

Additionally, we will see later that the $\mathbb{R}^\eta$-properness definition allows us to shed some light on the relationship between the two main kinds of quaternion properness, which are presented in Sections III-C and D. Finally, we must note that the definition of $\mathbb{R}^\eta$-proper vectors obviously depends on the choice of the pure unit quaternion η, but it is independent of the two orthogonal quaternions η' and η''.

C. $\mathbb{C}^\eta$-Properness

In this subsection, we introduce the definition of $\mathbb{C}^\eta$-proper vectors, which is closely related (but different) to those in [4], [5], [18], and [19]. The main difference is due to the fact that the previous approaches were based on the invariance of the SOS to left Clifford translations $e^{\eta\theta}\mathbf{x}$, whereas the definition in this paper naturally results in SOS invariance to right Clifford translations $\mathbf{x}e^{\eta\theta}$. More importantly, as a direct consequence of Lemma 1, the properness definitions in this paper are invariant to linear quaternion transformations of the form $\mathbf{u} = \mathbf{F}_1^H\mathbf{x}$, which is not the case if we impose the invariance of the SOS to left Clifford translations.[7] Obviously, this is a very desirable property from a practical point of view, which has its well-known counterpart in the case of complex vectors. Therefore, we think that the properness definitions in this paper will be more useful for the signal processing community.

Definition 4 ($\mathbb{C}^\eta$-Properness): A quaternion random vector $\mathbf{x}$ is $\mathbb{C}^\eta$-proper iff the complementary covariance matrices $\mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta')}}$ and $\mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta'')}}$ vanish.

At this point, one could be tempted to think that the definition of $\mathbb{C}^\eta$-proper vectors depends on η' and η''. However, the following lemma ensures that it only depends on η.

Lemma 5: The definition of $\mathbb{C}^\eta$-properness for quaternion vectors depends on η, but not on the particular choice of η' and η''. In other words, $\mathbf{x}$ is $\mathbb{C}^\eta$-proper iff it is $\mathbb{R}^{\eta^\perp}$-proper for all pure unit quaternions $\eta^\perp$ orthogonal to η.

Proof: The proof can be seen as a particular case of Lemma 3. It is based on the fact that all pure unit quaternions $\eta^\perp$ orthogonal to η can be written as real linear combinations of η' and η'', which also implies that $\mathbf{x}^{(\eta^\perp)}$ can be written as a quaternion linear combination of $\mathbf{x}^{(\eta')}$ and $\mathbf{x}^{(\eta'')}$. Therefore, if $\mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta')}}$ and $\mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta'')}}$ vanish, so does $\mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta^\perp)}} = E\mathbf{x}\mathbf{x}^{(\eta^\perp)^H}$. ∎

Analogously to the previous case, and from the expressions in Table I, we can conclude that a vector $\mathbf{x}$ is $\mathbb{C}^\eta$-proper iff

$$\mathbf{R}_{\mathbf{a}_1,\mathbf{a}_1^*} = \mathbf{R}_{\mathbf{a}_2,\mathbf{a}_2^*} = \mathbf{R}_{\mathbf{a}_1,\mathbf{a}_2^*} = \mathbf{0}_{n\times n}. \tag{10}$$

In other words, $\mathbf{x}$ is $\mathbb{C}^\eta$-proper iff it can be represented by means of two jointly proper complex vectors ($\mathbf{a}_1 = \mathbf{r}_1 + \eta\mathbf{r}_\eta$ and $\mathbf{a}_2 = \mathbf{r}_{\eta''} + \eta\mathbf{r}_{\eta'}$) in the plane spanned by $\{1,\eta\}$. Here, we must note that a similar conclusion was obtained in [18], [19] for the definition of $\mathbb{C}^\eta$-proper vectors based on the SOS invariance to left Clifford translations.

From a practical point of view, it is clear that the augmented covariance matrix of a $\mathbb{C}^\eta$-proper quaternion vector can be written as

$$\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}} = \begin{bmatrix}\mathbf{R}_{\tilde{\mathbf{x}},\tilde{\mathbf{x}}} & \mathbf{0}_{2n\times 2n}\\ \mathbf{0}_{2n\times 2n} & \mathbf{R}_{\tilde{\mathbf{x}},\tilde{\mathbf{x}}}^{(\eta')}\end{bmatrix}$$

where $\mathbf{R}_{\tilde{\mathbf{x}},\tilde{\mathbf{x}}} = E\tilde{\mathbf{x}}\tilde{\mathbf{x}}^H$ can be defined as a semi-augmented covariance matrix and $\tilde{\mathbf{x}} = \left[\mathbf{x}^T, \mathbf{x}^{(\eta)^T}\right]^T$ is the semi-augmented quaternion vector. Thus, it is easy to prove that the $\mathbb{C}^\eta$-properness is invariant under *semi-widely linear* transformations, i.e., linear transformations of the form

$$\mathbf{u} = \mathbf{F}_{\tilde{\mathbf{x}}}^H\tilde{\mathbf{x}} = \mathbf{F}_1^H\mathbf{x} + \mathbf{F}_\eta^H\mathbf{x}^{(\eta)} \tag{11}$$

where $\mathbf{F}_{\tilde{\mathbf{x}}} = \left[\mathbf{F}_1^T, \mathbf{F}_\eta^T\right]^T$ and $\mathbf{F}_1, \mathbf{F}_\eta \in \mathbb{H}^{n\times r}$. In other words, if $\mathbf{x}$ is $\mathbb{C}^\eta$-proper, all the vectors $\mathbf{u}$ obtained as (11) are $\mathbb{C}^\eta$-proper.

Finally, the following lemma establishes the equivalence between $\mathbb{C}^\eta$-properness and the invariance of the SOS to right Clifford translations in the plane $\{1,\eta\}$.

Lemma 6: A quaternion random vector $\mathbf{x}$ is $\mathbb{C}^\eta$-proper iff its SOS are invariant under right Clifford translations $\mathbf{x}e^{\eta\theta}$, $\forall\,\theta\in\mathbb{R}$.

Proof: As we have seen, the SOS of a quaternion vector are given by the covariance and three complementary covariance matrices. Consider the right product $\mathbf{u} = \mathbf{x}a$, with $a \in \mathbb{H}$, which can be rewritten as $\mathbf{u} = a\mathbf{x}^{(a^*)}$. Thus, from Lemma 2, we obtain

$$\begin{aligned}\mathbf{R}_{\mathbf{u},\mathbf{u}} &= a\mathbf{R}_{\mathbf{x},\mathbf{x}}^{(a^*)}a^* = |a|^2\mathbf{R}_{\mathbf{x},\mathbf{x}}\\ \mathbf{R}_{\mathbf{u},\mathbf{u}^{(\eta)}} &= a\mathbf{R}_{\mathbf{x},\mathbf{x}^{(\nu)}}^{(a^*)}(a^*)^{(\eta)} = -|a|^2\mathbf{R}_{\mathbf{x},\mathbf{x}^{(\nu)}}\nu\eta\\ \mathbf{R}_{\mathbf{u},\mathbf{u}^{(\eta')}} &= a\mathbf{R}_{\mathbf{x},\mathbf{x}^{(\nu')}}^{(a^*)}(a^*)^{(\eta')} = -|a|^2\mathbf{R}_{\mathbf{x},\mathbf{x}^{(\nu')}}\nu'\eta'\\ \mathbf{R}_{\mathbf{u},\mathbf{u}^{(\eta'')}} &= a\mathbf{R}_{\mathbf{x},\mathbf{x}^{(\nu'')}}^{(a^*)}(a^*)^{(\eta'')} = -|a|^2\mathbf{R}_{\mathbf{x},\mathbf{x}^{(\nu'')}}\nu''\eta''\end{aligned}$$

[7]If the SOS of $\mathbf{x}$ are invariant to left Clifford translations of the form $\mathbf{u} = \eta\mathbf{x}$, the covariance matrices of $\mathbf{u}$ and $\mathbf{x}$ should be identical. Thus, we have $\mathbf{R}_{\mathbf{u},\mathbf{u}} = \mathbf{R}_{\mathbf{x},\mathbf{x}}^{(\eta)} = \mathbf{R}_{\mathbf{x},\mathbf{x}}$, which implies that the elements of $\mathbf{R}_{\mathbf{x},\mathbf{x}}$ belong to the plane $\{1,\eta\}$. Now, it is easy to find a linear transformation $\mathbf{v} = \mathbf{A}\mathbf{x}$ (for instance, $\mathbf{A} = (\eta+\eta')\mathbf{I}_n$) such that $\mathbf{R}_{\mathbf{v},\mathbf{v}} = \mathbf{A}\mathbf{R}_{\mathbf{x},\mathbf{x}}\mathbf{A}^H \neq \mathbf{R}_{\mathbf{v},\mathbf{v}}^{(\eta)}$, i.e., the properness of $\mathbf{x}$ can be lost due to a linear transformation (and vice versa).

where $\nu = \eta^{(a)}$, $\nu' = \eta'^{(a)}$, $\nu'' = \eta''^{(a)}$. Now, particularizing for $a = e^{\eta\theta}$, we have $\nu = \eta$, $\nu' = a^2\eta'$, $\nu'' = a^2\eta''$, which yields

$$\begin{aligned}\mathbf{R}_{\mathbf{u},\mathbf{u}} &= \mathbf{R}_{\mathbf{x},\mathbf{x}}\\ \mathbf{R}_{\mathbf{u},\mathbf{u}^{(\eta)}} &= \mathbf{R}_{\mathbf{x},\mathbf{x}^{(\nu)}}\\ \mathbf{R}_{\mathbf{u},\mathbf{u}^{(\eta')}} &= \mathbf{R}_{\mathbf{x},\mathbf{x}^{(\nu')}}a^2 = E\mathbf{x}a^2\mathbf{x}^{(\eta')^H}\\ \mathbf{R}_{\mathbf{u},\mathbf{u}^{(\eta'')}} &= \mathbf{R}_{\mathbf{x},\mathbf{x}^{(\nu'')}}a^2 = E\mathbf{x}a^2\mathbf{x}^{(\eta'')^H}\end{aligned}$$

i.e., the covariance $\mathbf{R}_{\mathbf{x},\mathbf{x}}$ and complementary covariance $\mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta)}}$ are invariant under right Clifford translations $\mathbf{x}e^{\eta\theta}$. On the other hand, writing $a = e^{\eta\theta} = \cos\theta + \eta\sin\theta$ the complementary covariance matrices $\mathbf{R}_{\mathbf{u},\mathbf{u}^{(\eta')}}$ and $\mathbf{R}_{\mathbf{u},\mathbf{u}^{(\eta'')}}$ can be further simplified to

$$\begin{aligned}\mathbf{R}_{\mathbf{u},\mathbf{u}^{(\eta')}} &= \mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta')}}\cos 2\theta + \mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta'')}}\eta\sin 2\theta\\ \mathbf{R}_{\mathbf{u},\mathbf{u}^{(\eta'')}} &= \mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta')}}\eta\sin 2\theta + \mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta'')}}\cos 2\theta.\end{aligned}$$

Thus, it is easy to see that the SOS are invariant under right Clifford translations $\mathbf{u} = \mathbf{x}e^{\eta\theta}$ iff

$$[\,\mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta')}} \quad \mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta'')}}\,]\boldsymbol{\Psi} = \mathbf{0}_{n\times 2n} \tag{12}$$

where

$$\boldsymbol{\Psi} = \left(\begin{bmatrix}\cos 2\theta - 1 & \eta\sin 2\theta\\ \eta\sin 2\theta & \cos 2\theta - 1\end{bmatrix}\otimes \mathbf{I}_n\right)$$

and $\boldsymbol{\Psi}\boldsymbol{\Psi}^H = 2(1-\cos 2\theta)\mathbf{I}_{2n}$. Therefore, excluding the trivial case of $a = \pm 1$, (12) is only satisfied for $\mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta')}} = \mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta'')}} = \mathbf{0}_{n\times n}$, i.e., the quaternion vector $\mathbf{x}$ is invariant to right Clifford translations $\mathbf{x}e^{\eta\theta}$ ($\forall\, \theta \in \mathbb{R}$) iff it is $\mathbb{C}^\eta$-proper. ∎

D. $\mathbb{Q}$-Properness

So far, we have presented two different kinds of properness for quaternion random vectors. The last and strongest kind of properness can be seen as a combination of the $\mathbb{R}^\eta$ and $\mathbb{C}^\eta$ properness and is defined as follows:[8]

Definition 5 ($\mathbb{Q}$-Properness): A quaternion random vector $\mathbf{x}$ is $\mathbb{Q}$-proper iff the three complementary covariance matrices $\mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta)}}$, $\mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta')}}$ and $\mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta'')}}$ vanish.

The following lemmas establish the main properties of $\mathbb{Q}$-proper quaternion vectors.

Lemma 7: A quaternion random vector $\mathbf{x}$ is $\mathbb{Q}$-proper iff all the complementary covariance matrices $\mathbf{R}_{\mathbf{x},\mathbf{x}^{(\nu)}}$ (for all pure unit quaternions ν) vanish. In other words, the definition of $\mathbb{Q}$-proper vectors does not depend on the orthogonal basis $\{1,\eta,\eta',\eta''\}$, and it is equivalent to the $\mathbb{R}^\eta$ and $\mathbb{C}^\eta$ properness of $\mathbf{x}$ for all η.

Proof: This is a direct consequence of Lemma 3 and the $\mathbb{Q}$-properness definition. Note that the complementary covariance matrix $\mathbf{R}_{\mathbf{x},\mathbf{x}^{(\nu)}}$ is given by a quaternion linear combination of $\mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta)}}$, $\mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta')}}$ and $\mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta'')}}$. Thus, if $\mathbf{x}$ is $\mathbb{Q}$-proper we have $\mathbf{R}_{\mathbf{x},\mathbf{x}^{(\nu)}} = \mathbf{0}_{n\times n}$ for all pure unit quaternions ν. Obviously, this also implies that $\mathbf{x}$ is $\mathbb{R}^\eta$-proper and $\mathbb{C}^\eta$-proper for all pure unit quaternions η. ∎

Lemma 8: The covariance matrix of a $\mathbb{Q}$-proper quaternion vector can be written as

$$\begin{aligned}\mathbf{R}_{\mathbf{x},\mathbf{x}} &= E\mathbf{x}\mathbf{x}^H = 2\left(\mathbf{R}_{\mathbf{a}_1,\mathbf{a}_1} + \eta''\mathbf{R}^H_{\mathbf{a}_1,\mathbf{a}_2}\right)\\ &= 2\left(\mathbf{R}_{\mathbf{b}_1,\mathbf{b}_1} + \eta\mathbf{R}^H_{\mathbf{b}_1,\mathbf{b}_2}\right)\\ &= 2\left(\mathbf{R}_{\mathbf{c}_1,\mathbf{c}_1} + \eta'\mathbf{R}^H_{\mathbf{c}_1,\mathbf{c}_2}\right)\end{aligned}$$

regardless of the choice of the orthogonal basis $\{1,\eta,\eta',\eta''\}$. Equivalently, the vectors in the real representation of $\mathbf{x}$ satisfy

$$\begin{aligned}\mathbf{R}_{\mathbf{r}_1,\mathbf{r}_1} &= \mathbf{R}_{\mathbf{r}_\eta,\mathbf{r}_\eta} = \mathbf{R}_{\mathbf{r}_{\eta'},\mathbf{r}_{\eta'}} = \mathbf{R}_{\mathbf{r}_{\eta''},\mathbf{r}_{\eta''}}\\ \mathbf{R}^T_{\mathbf{r}_1,\mathbf{r}_\eta} &= -\mathbf{R}_{\mathbf{r}_1,\mathbf{r}_\eta} = -\mathbf{R}_{\mathbf{r}_{\eta'},\mathbf{r}_{\eta''}} = \mathbf{R}^T_{\mathbf{r}_{\eta'},\mathbf{r}_{\eta''}}\\ \mathbf{R}^T_{\mathbf{r}_1,\mathbf{r}_{\eta'}} &= -\mathbf{R}_{\mathbf{r}_1,\mathbf{r}_{\eta'}} = \mathbf{R}_{\mathbf{r}_\eta,\mathbf{r}_{\eta''}} = -\mathbf{R}^T_{\mathbf{r}_\eta,\mathbf{r}_{\eta''}}\\ \mathbf{R}^T_{\mathbf{r}_1,\mathbf{r}_{\eta''}} &= -\mathbf{R}_{\mathbf{r}_1,\mathbf{r}_{\eta''}} = -\mathbf{R}_{\mathbf{r}_\eta,\mathbf{r}_{\eta'}} = \mathbf{R}^T_{\mathbf{r}_\eta,\mathbf{r}_{\eta'}}.\end{aligned}$$

Proof: This can be seen as a consequence of the simultaneous $\mathbb{R}^\eta$ and $\mathbb{C}^\eta$ properness, and can be easily checked with the help of Table I. ∎

Lemma 9: A quaternion random vector $\mathbf{x}$ is $\mathbb{Q}$-proper iff its SOS are invariant to right Clifford translations $\mathbf{x}e^{\eta\theta}$ for all pure unit quaternions η and $\forall\,\theta\in\mathbb{R}$.

Proof: This is a direct consequence of Lemma 6 and the $\mathbb{C}^\eta$-properness of $\mathbf{x}$ for all η. ∎

To summarize, we can say that $\mathbb{Q}$-properness combines the two previous kinds of properness as follows: First, the $\mathbb{R}^\eta$-properness ensures the equality (up to a complex conjugation) of the covariance matrices, and the skew-symmetry of the cross covariance between $\mathbf{a}_1$ and $\mathbf{a}_2$ [see (8) and (9)], which can be seen as the complex version of (2) and (3) for proper complex vectors. On the other hand, the $\mathbb{C}^\eta$-properness ensures that the complex vectors $\mathbf{a}_1$ and $\mathbf{a}_2$ are jointly proper. Thus, $\mathbb{R}^\eta$-properness and $\mathbb{C}^\eta$-properness can be seen as two complementary kinds of properness for quaternion random vectors, which together result in $\mathbb{Q}$-properness.

E. Extension to Two Random Vectors

In order to conclude this section, we introduce properness definitions for two quaternion random vectors $\mathbf{x}\in\mathbb{H}^{n\times 1}$ and $\mathbf{y}\in\mathbb{H}^{m\times 1}$. Analogously to the complex case, we start by the definition of cross-proper vectors.

Definition 6 (Cross Properness): Two quaternion random vectors $\mathbf{x}$ and $\mathbf{y}$ are:

- cross $\mathbb{R}^\eta$-proper iff the complementary cross-covariance matrix $\mathbf{R}_{\mathbf{x},\mathbf{y}^{(\eta)}} = E\mathbf{x}\mathbf{y}^{(\eta)^H}$ vanishes;

[8]Note again that the $\mathbb{Q}$-properness definition in this paper differs from those based on the invariance of the SOS to left Clifford translations [4], [5], which are not invariant to quaternion linear transformations.

- cross $\mathbb{C}^{\eta}$-proper iff the complementary cross-covariance matrices $\mathbf{R}_{\mathbf{x},\mathbf{y}^{(\eta')}} = E\mathbf{x}\mathbf{y}^{(\eta')^H}$ and $\mathbf{R}_{\mathbf{x},\mathbf{y}^{(\eta'')}} = E\mathbf{x}\mathbf{y}^{(\eta'')^H}$ vanish;
- cross $\mathbb{Q}$-proper iff all the complementary cross-covariance matrices ($\mathbf{R}_{\mathbf{x},\mathbf{y}^{(\eta)}}$, $\mathbf{R}_{\mathbf{x},\mathbf{y}^{(\eta')}}$ and $\mathbf{R}_{\mathbf{x},\mathbf{y}^{(\eta'')}}$) vanish.

Finally, combining the definitions of properness and cross properness, we arrive to the concept of jointly proper vectors.

Definition 7 (Joint-Properness): Two quaternion random vectors $\mathbf{x}$ and $\mathbf{y}$ are jointly $\mathbb{R}^{\eta}$ (respectively $\mathbb{C}^{\eta}$ or $\mathbb{Q}$) proper iff the composite vector $\left[\mathbf{x}^T, \mathbf{y}^T\right]^T$ is $\mathbb{R}^{\eta}$ (resp. $\mathbb{C}^{\eta}$ or $\mathbb{Q}$) proper. Equivalently, $\mathbf{x}$ and $\mathbf{y}$ are jointly proper iff they are proper and cross proper.

IV. Full and Semi-Widely Linear Processing of Quaternion Random Vectors

To our best knowledge, the only work dealing with widely linear processing of quaternion random vectors is [7]. In that work, inspired by the case of complex vectors, the authors propose to simultaneously operate on the quaternion vector $\mathbf{x}$ and its conjugate $\mathbf{x}^*$. Here, we show that, unlike the complex case, there exist different kinds of quaternion widely linear processing. The most general linear transformation, which we refer to as *full-widely linear* processing, consists in the simultaneous operation on the four involutions

$$\mathbf{u} = \mathbf{F}_{\bar{\mathbf{x}}}^H \bar{\mathbf{x}} = \mathbf{F}_1^H \mathbf{x} + \mathbf{F}_{\eta}^H \mathbf{x}^{(\eta)} + \mathbf{F}_{\eta'}^H \mathbf{x}^{(\eta')} + \mathbf{F}_{\eta''}^H \mathbf{x}^{(\eta'')}$$

where $\mathbf{F}_{\bar{\mathbf{x}}} = \left[\mathbf{F}_1^T, \mathbf{F}_{\eta}^T, \mathbf{F}_{\eta'}^T, \mathbf{F}_{\eta''}^T\right]^T \in \mathbb{H}^{4n\times r}$ is a quaternion matrix. In terms of the augmented vectors $\bar{\mathbf{x}}$ and $\bar{\mathbf{u}}$, the above equation can be written as

$$\bar{\mathbf{u}} = \overline{\mathbf{F}}_{\bar{\mathbf{x}}}^H \bar{\mathbf{x}} \tag{13}$$

where

$$\overline{\mathbf{F}}_{\bar{\mathbf{x}}} = \underbrace{\begin{bmatrix} \mathbf{F}_1 & \mathbf{F}_{\eta}^{(\eta)} & \mathbf{F}_{\eta'}^{(\eta')} & \mathbf{F}_{\eta''}^{(\eta'')} \\ \mathbf{F}_{\eta} & \mathbf{F}_1^{(\eta)} & \mathbf{F}_{\eta''}^{(\eta')} & \mathbf{F}_{\eta'}^{(\eta'')} \\ \mathbf{F}_{\eta'} & \mathbf{F}_{\eta''}^{(\eta)} & \mathbf{F}_1^{(\eta')} & \mathbf{F}_{\eta}^{(\eta'')} \\ \mathbf{F}_{\eta''} & \mathbf{F}_{\eta'}^{(\eta)} & \mathbf{F}_{\eta}^{(\eta')} & \mathbf{F}_1^{(\eta'')} \end{bmatrix}}_{4n\times 4r}$$

is a general full-widely linear operator. Equivalently, we can use the real version of (13)

$$\mathbf{r}_{\mathbf{u}} = \mathbf{F}_{\mathbf{r}_{\mathbf{x}}}^T \mathbf{r}_{\mathbf{x}}$$

where $\mathbf{r}_{\mathbf{x}} = \left[\mathbf{r}_1^T, \mathbf{r}_{\eta}^T, \mathbf{r}_{\eta'}^T, \mathbf{r}_{\eta''}^T\right]^T = \frac{1}{2}\mathbf{T}_n^H \bar{\mathbf{x}}$, $\mathbf{r}_{\mathbf{u}} = \frac{1}{2}\mathbf{T}_r^H \bar{\mathbf{u}}$, and $\mathbf{F}_{\mathbf{r}_{\mathbf{x}}} \in \mathbb{R}^{4n\times 4r}$ is given by

$$\mathbf{F}_{\mathbf{r}_{\mathbf{x}}} = \mathbf{T}_n^H \overline{\mathbf{F}}_{\bar{\mathbf{x}}} \mathbf{T}_r \tag{14}$$

with $\mathbf{T}_n$ (and $\mathbf{T}_r$) defined in (7).

In this section, we follow a similar derivation to that in [17] for the case of complex vectors. Our goal is to present a rigorous generalization of several well-known multivariate statistical analysis techniques to the case of quaternion vectors and, more importantly, to show the implications of the $\mathbb{C}^{\eta}$ and $\mathbb{Q}$ properness on the optimal linear processing.

A. Multivariate Statistical Analysis of Quaternion Vectors

Several popular multivariate statistical analysis techniques amount to maximize the correlation (under different constraints or invariances) between projections of two random vectors [17]. In this subsection, we focus on the general problem of maximizing the correlation between the following r-dimensional projections of the quaternion vectors $\mathbf{x} \in \mathbb{H}^{n\times 1}$ and $\mathbf{y} \in \mathbb{H}^{m\times 1}$

$$\mathbf{r}_{\mathbf{u}} = \mathbf{F}_{\mathbf{r}_{\mathbf{x}}}^T \mathbf{r}_x, \quad \mathbf{r}_{\mathbf{v}} = \mathbf{G}_{\mathbf{r}_{\mathbf{y}}}^T \mathbf{r}_{\mathbf{y}}$$

where $\mathbf{F}_{\mathbf{r}_{\mathbf{x}}} \in \mathbb{R}^{4n\times 4r}$, $\mathbf{G}_{\mathbf{r}_{\mathbf{y}}} \in \mathbb{R}^{4m\times 4r}$ are real operators,[9] and $r \le p = \min(m, n)$. Specifically, our problem can be written as

$$\underset{\mathbf{F}_{\mathbf{r}_{\mathbf{x}}}, \mathbf{G}_{\mathbf{r}_{\mathbf{y}}}}{\arg\max} \operatorname{Tr}\left(\mathbf{F}_{\mathbf{r}_{\mathbf{x}}}^T \mathbf{R}_{\mathbf{r}_{\mathbf{x}},\mathbf{r}_{\mathbf{y}}} \mathbf{G}_{\mathbf{r}_{\mathbf{y}}}\right)$$

where $\mathbf{R}_{\mathbf{r}_{\mathbf{x}},\mathbf{r}_{\mathbf{y}}} = E\mathbf{r}_{\mathbf{x}}\mathbf{r}_{\mathbf{y}}^T$. Obviously, in order to avoid trivial solutions, some constraints (or invariances) have to be imposed in the previous problem. In fact, the choice of constraints makes the difference among the following well-known multivariate statistical analysis techniques.

- Partial least squares (PLS) [22]: PLS maximizes the correlations between the projections of two random vectors subject to the unitarity of the projectors, i.e., the constraints are $\mathbf{F}_{\mathbf{r}_{\mathbf{x}}}^T \mathbf{F}_{\mathbf{r}_{\mathbf{x}}} = \mathbf{G}_{\mathbf{r}_{\mathbf{y}}}^T \mathbf{G}_{\mathbf{r}_{\mathbf{y}}} = \mathbf{I}_{4r}$. In the particular case of $\mathbf{y} = \mathbf{x}$, PLS reduces to the principal component analysis (PCA) technique [21].
- Multivariate linear regression (MLR) [23]: For this method, which is also known as the rank-reduced Wiener filter, half canonical correlation analysis [14], or orthogonalized PLS [30], the constraints can be written as $\mathbf{F}_{\mathbf{r}_{\mathbf{x}}}^T \mathbf{R}_{\mathbf{r}_{\mathbf{x}},\mathbf{r}_{\mathbf{x}}} \mathbf{F}_{\mathbf{r}_{\mathbf{x}}} = \mathbf{G}_{\mathbf{r}_{\mathbf{y}}}^T \mathbf{G}_{\mathbf{r}_{\mathbf{y}}} = \mathbf{I}_{4r}$.
- Canonical correlation analysis (CCA) [24], [25]: This technique imposes the energy and orthogonality constraints on the projections $\mathbf{r}_{\mathbf{u}}$ and $\mathbf{r}_{\mathbf{v}}$, i.e., the constraints are $\mathbf{F}_{\mathbf{r}_{\mathbf{x}}}^T \mathbf{R}_{\mathbf{r}_{\mathbf{x}},\mathbf{r}_{\mathbf{x}}} \mathbf{F}_{\mathbf{r}_{\mathbf{x}}} = \mathbf{G}_{\mathbf{r}_{\mathbf{y}}}^T \mathbf{R}_{\mathbf{r}_{\mathbf{y}},\mathbf{r}_{\mathbf{y}}} \mathbf{G}_{\mathbf{r}_{\mathbf{y}}} = \mathbf{I}_{4r}$.

After a straightforward algebraic manipulation, the three previous problems can be rewritten as

$$\begin{aligned} \underset{\mathbf{U}_{\mathbf{r}_{\mathbf{x}}}, \mathbf{V}_{\mathbf{r}_{\mathbf{y}}}}{\arg\max} \quad & \operatorname{Tr}\left(\mathbf{U}_{\mathbf{r}_{\mathbf{x}}}^T \mathbf{C}_{\mathbf{r}_{\mathbf{x}},\mathbf{r}_{\mathbf{y}}} \mathbf{V}_{\mathbf{r}_{\mathbf{y}}}\right) \\ \text{s. t.} \quad & \mathbf{U}_{\mathbf{r}_{\mathbf{x}}}^T \mathbf{U}_{\mathbf{r}_{\mathbf{x}}} = \mathbf{V}_{\mathbf{r}_{\mathbf{y}}}^T \mathbf{V}_{\mathbf{r}_{\mathbf{y}}} = \mathbf{I}_{4r} \end{aligned} \tag{15}$$

where $\mathbf{C}_{\mathbf{r}_{\mathbf{x}},\mathbf{r}_{\mathbf{y}}} = \mathbf{S}_{\mathbf{r}_{\mathbf{x}},\mathbf{r}_{\mathbf{x}}}^{-\frac{1}{2}} \mathbf{R}_{\mathbf{r}_{\mathbf{x}},\mathbf{r}_{\mathbf{y}}} \mathbf{S}_{\mathbf{r}_{\mathbf{y}},\mathbf{r}_{\mathbf{y}}}^{-\frac{1}{2}}$, $\mathbf{U}_{\mathbf{r}_{\mathbf{x}}} = \mathbf{S}_{\mathbf{r}_{\mathbf{x}},\mathbf{r}_{\mathbf{x}}}^{\frac{1}{2}} \mathbf{F}_{\mathbf{r}_{\mathbf{x}}}$, $\mathbf{V}_{\mathbf{r}_{\mathbf{y}}} = \mathbf{S}_{\mathbf{r}_{\mathbf{y}},\mathbf{r}_{\mathbf{y}}}^{\frac{1}{2}} \mathbf{G}_{\mathbf{r}_{\mathbf{y}}}$, and the expressions for $\mathbf{S}_{\mathbf{r}_{\mathbf{x}},\mathbf{r}_{\mathbf{x}}}$ and $\mathbf{S}_{\mathbf{r}_{\mathbf{y}},\mathbf{r}_{\mathbf{y}}}$ in the three studied cases are summarized in Table II. Obviously, the solutions $\mathbf{U}_{\mathbf{r}_{\mathbf{x}}}$, $\mathbf{V}_{\mathbf{r}_{\mathbf{y}}}$ of (15) are given by the singular vectors associated to the $4r$ largest singular values of the matrix $\mathbf{C}_{\mathbf{r}_{\mathbf{x}},\mathbf{r}_{\mathbf{y}}}$, whose singular value decomposition (SVD) can be written as

$$\mathbf{C}_{\mathbf{r}_{\mathbf{x}},\mathbf{r}_{\mathbf{y}}} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{V}^T$$

with $\mathbf{U} \in \mathbb{R}^{4n\times 4p}$, $\mathbf{V} \in \mathbb{R}^{4m\times 4p}$ unitary matrices and $\boldsymbol{\Lambda} \in \mathbb{R}^{4p\times 4p}$ a diagonal matrix containing the singular values. In par-

[9]Note that $4r$-dimensional real projections are equivalent to r-dimensional full-widely linear quaternion projections.

TABLE II
SUMMARY OF THE PRESENTED METHODS AND CONDITIONS FOR OPTIMALITY OF SEMI-WIDELY OR CONVENTIONAL LINEAR PROCESSING

	PLS (also PCA)	MLR	CCA
Matrices in the real formulation	$\mathbf{S}_{\mathbf{r_x},\mathbf{r_x}} = \mathbf{I}_{4n}$ $\mathbf{S}_{\mathbf{r_y},\mathbf{r_y}} = \mathbf{I}_{4m}$ $\mathbf{C}_{\mathbf{r_x},\mathbf{r_y}} = \mathbf{R}_{\mathbf{r_x},\mathbf{r_y}}$	$\mathbf{S}_{\mathbf{r_x},\mathbf{r_x}} = \mathbf{R}_{\mathbf{r_x},\mathbf{r_x}}$ $\mathbf{S}_{\mathbf{r_y},\mathbf{r_y}} = \mathbf{I}_{4m}$ $\mathbf{C}_{\mathbf{r_x},\mathbf{r_y}} = \mathbf{R}_{\mathbf{r_x},\mathbf{r_x}}^{-\frac{1}{2}}\mathbf{R}_{\mathbf{r_x},\mathbf{r_y}}$	$\mathbf{S}_{\mathbf{r_x},\mathbf{r_x}} = \mathbf{R}_{\mathbf{r_x},\mathbf{r_x}}$ $\mathbf{S}_{\mathbf{r_y},\mathbf{r_y}} = \mathbf{R}_{\mathbf{r_y},\mathbf{r_y}}$ $\mathbf{C}_{\mathbf{r_x},\mathbf{r_y}} = \mathbf{R}_{\mathbf{r_x},\mathbf{r_x}}^{-\frac{1}{2}}\mathbf{R}_{\mathbf{r_x},\mathbf{r_y}}\mathbf{R}_{\mathbf{r_y},\mathbf{r_y}}^{-\frac{1}{2}}$
Matrices in the augmented-vector formulation	$\mathbf{S}_{\bar{\mathbf{x}},\bar{\mathbf{x}}} = \mathbf{I}_{4n}$ $\mathbf{S}_{\bar{\mathbf{y}},\bar{\mathbf{y}}} = \mathbf{I}_{4m}$ $\mathbf{C}_{\bar{\mathbf{x}},\bar{\mathbf{y}}} = \mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{y}}}$	$\mathbf{S}_{\bar{\mathbf{x}},\bar{\mathbf{x}}} = \mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}$ $\mathbf{S}_{\bar{\mathbf{y}},\bar{\mathbf{y}}} = \mathbf{I}_{4m}$ $\mathbf{C}_{\bar{\mathbf{x}},\bar{\mathbf{y}}} = \mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}^{-\frac{1}{2}}\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{y}}}$	$\mathbf{S}_{\bar{\mathbf{x}},\bar{\mathbf{x}}} = \mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}$ $\mathbf{S}_{\bar{\mathbf{y}},\bar{\mathbf{y}}} = \mathbf{R}_{\bar{\mathbf{y}},\bar{\mathbf{y}}}$ $\mathbf{C}_{\bar{\mathbf{x}},\bar{\mathbf{y}}} = \mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}^{-\frac{1}{2}}\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{y}}}\mathbf{R}_{\bar{\mathbf{y}},\bar{\mathbf{y}}}^{-\frac{1}{2}}$
Sufficient conditions for optimality of semi-widely linear processing	Cross $\mathbb{C}^\eta$-properness of $\mathbf{x}$ and $\mathbf{y}$ ($\mathbb{C}^\eta$-properness of $\mathbf{x}$ in the particular PCA case)	Cross $\mathbb{C}^\eta$-properness of $\mathbf{x}$ and $\mathbf{y}$ and $\mathbb{C}^\eta$-properness of $\mathbf{x}$	Joint $\mathbb{C}^\eta$-properness of $\mathbf{x}$ and $\mathbf{y}$
Sufficient conditions for optimality of conventional linear processing	Cross $\mathbb{Q}$-properness of $\mathbf{x}$ and $\mathbf{y}$ ($\mathbb{Q}$-properness of $\mathbf{x}$ in the particular PCA case)	Cross $\mathbb{Q}$-properness of $\mathbf{x}$ and $\mathbf{y}$ and $\mathbb{Q}$-properness of $\mathbf{x}$	Joint $\mathbb{Q}$-properness of $\mathbf{x}$ and $\mathbf{y}$

ticular, we will order the singular values $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{4p}$ in $\mathbf{\Lambda}$ as

$$\mathbf{\Lambda} = \begin{bmatrix} \mathbf{\Lambda}_1 & \mathbf{0}_{p\times p} & \mathbf{0}_{p\times p} & \mathbf{0}_{p\times p} \\ \mathbf{0}_{p\times p} & \mathbf{\Lambda}_2 & \mathbf{0}_{p\times p} & \mathbf{0}_{p\times p} \\ \mathbf{0}_{p\times p} & \mathbf{0}_{p\times p} & \mathbf{\Lambda}_3 & \mathbf{0}_{p\times p} \\ \mathbf{0}_{p\times p} & \mathbf{0}_{p\times p} & \mathbf{0}_{p\times p} & \mathbf{\Lambda}_4 \end{bmatrix}$$

with

$$\begin{aligned} \mathbf{\Lambda}_1 &= \mathrm{diag}(\lambda_1, \lambda_5, \ldots, \lambda_{4p-3}) \\ \mathbf{\Lambda}_2 &= \mathrm{diag}(\lambda_2, \lambda_6, \ldots, \lambda_{4p-2}) \\ \mathbf{\Lambda}_3 &= \mathrm{diag}(\lambda_3, \lambda_7, \ldots, \lambda_{4p-1}) \\ \mathbf{\Lambda}_4 &= \mathrm{diag}(\lambda_4, \lambda_8, \ldots, \lambda_{4p}). \end{aligned}$$

At this point, taking (14) into account, the full-widely linear operators $\overline{\mathbf{F}}_{\bar{\mathbf{x}}}$ and $\overline{\mathbf{G}}_{\bar{\mathbf{y}}}$ can be obtained as

$$\begin{aligned} \overline{\mathbf{F}}_{\bar{\mathbf{x}}} &= \mathbf{T}_n \mathbf{F}_{\mathbf{r_x}} \mathbf{T}_r^H = \mathbf{T}_n \mathbf{S}_{\mathbf{r_x},\mathbf{r_x}}^{-\frac{1}{2}} \mathbf{U}_{\mathbf{r_x}} \mathbf{T}_r^H \\ \overline{\mathbf{G}}_{\bar{\mathbf{y}}} &= \mathbf{T}_m \mathbf{G}_{\mathbf{r_y}} \mathbf{T}_r^H = \mathbf{T}_m \mathbf{S}_{\mathbf{r_y},\mathbf{r_y}}^{-\frac{1}{2}} \mathbf{V}_{\mathbf{r_y}} \mathbf{T}_r^H \end{aligned}$$

and due to the unitarity of the operator $\mathbf{T}_n$, we can write

$$\begin{aligned} \overline{\mathbf{F}}_{\bar{\mathbf{x}}} &= \left(\mathbf{T}_n \mathbf{S}_{\mathbf{r_x},\mathbf{r_x}}^{-\frac{1}{2}} \mathbf{T}_n^H\right)\left(\mathbf{T}_n \mathbf{U}_{\mathbf{r_x}} \mathbf{T}_r^H\right) = \mathbf{S}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}^{-\frac{1}{2}} \overline{\mathbf{U}}_{\bar{\mathbf{x}}} \\ \overline{\mathbf{G}}_{\bar{\mathbf{y}}} &= \left(\mathbf{T}_m \mathbf{S}_{\mathbf{r_y},\mathbf{r_y}}^{-\frac{1}{2}} \mathbf{T}_m^H\right)\left(\mathbf{T}_m \mathbf{V}_{\mathbf{r_y}} \mathbf{T}_r^H\right) = \mathbf{S}_{\bar{\mathbf{y}},\bar{\mathbf{y}}}^{-\frac{1}{2}} \overline{\mathbf{V}}_{\bar{\mathbf{y}}} \end{aligned}$$

where

$$\mathbf{S}_{\bar{\mathbf{x}},\bar{\mathbf{x}}} = \mathbf{T}_n \mathbf{S}_{\mathbf{r_x},\mathbf{r_x}} \mathbf{T}_n^H, \quad \mathbf{S}_{\bar{\mathbf{y}},\bar{\mathbf{y}}} = \mathbf{T}_m \mathbf{S}_{\mathbf{r_y},\mathbf{r_y}} \mathbf{T}_m^H$$

are shown in Table II for the three studied cases, and $\overline{\mathbf{U}}_{\bar{\mathbf{x}}} = \mathbf{T}_n \mathbf{U}_{\mathbf{r_x}} \mathbf{T}_r^H$, $\overline{\mathbf{V}}_{\bar{\mathbf{y}}} = \mathbf{T}_m \mathbf{V}_{\mathbf{r_y}} \mathbf{T}_r^H$ are unitary full-widely linear operators. Furthermore, defining the matrix

$$\mathbf{C}_{\bar{\mathbf{x}},\bar{\mathbf{y}}} = \mathbf{T}_n \mathbf{C}_{\mathbf{r_x},\mathbf{r_y}} \mathbf{T}_m^H = \mathbf{S}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}^{-\frac{1}{2}} \mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{y}}} \mathbf{S}_{\bar{\mathbf{y}},\bar{\mathbf{y}}}^{-\frac{1}{2}}$$

the operators $\overline{\mathbf{U}}_{\bar{\mathbf{x}}}$ and $\overline{\mathbf{V}}_{\bar{\mathbf{y}}}$ can be directly obtained from the decomposition

$$\mathbf{C}_{\bar{\mathbf{x}},\bar{\mathbf{y}}} = \underbrace{\left(\mathbf{T}_n \mathbf{U} \mathbf{T}_p^H\right)}_{\overline{\mathbf{U}}} \underbrace{\left(\mathbf{T}_p \mathbf{\Lambda} \mathbf{T}_p^H\right)}_{\overline{\mathbf{\Lambda}}} \underbrace{\left(\mathbf{T}_n \mathbf{V} \mathbf{T}_p^H\right)^H}_{\overline{\mathbf{V}}^H} \tag{16}$$

which can be seen as an extension of the singular value decomposition used in [14] for the second-order circularity analysis of complex vectors. In particular, it is easy to check that $\overline{\mathbf{U}} \in \mathbb{H}^{4n\times 4p}$, $\overline{\mathbf{V}} \in \mathbb{H}^{4m\times 4p}$ are unitary full-widely linear operators, and

$$\overline{\mathbf{\Lambda}} = \begin{bmatrix} \mathbf{\Sigma}_1 & \mathbf{\Sigma}_2 & \mathbf{\Sigma}_3 & \mathbf{\Sigma}_4 \\ \mathbf{\Sigma}_2 & \mathbf{\Sigma}_1 & \mathbf{\Sigma}_4 & \mathbf{\Sigma}_3 \\ \mathbf{\Sigma}_3 & \mathbf{\Sigma}_4 & \mathbf{\Sigma}_1 & \mathbf{\Sigma}_2 \\ \mathbf{\Sigma}_4 & \mathbf{\Sigma}_3 & \mathbf{\Sigma}_2 & \mathbf{\Sigma}_1 \end{bmatrix}$$

with

$$\mathbf{\Sigma}_1 = \frac{1}{4}\left(\mathbf{\Lambda}_1 + \mathbf{\Lambda}_2 + \mathbf{\Lambda}_3 + \mathbf{\Lambda}_4\right) \tag{17}$$

$$\mathbf{\Sigma}_2 = \frac{1}{4}\left(\mathbf{\Lambda}_1 + \mathbf{\Lambda}_2 - \mathbf{\Lambda}_3 - \mathbf{\Lambda}_4\right) \tag{18}$$

$$\mathbf{\Sigma}_3 = \frac{1}{4}\left(\mathbf{\Lambda}_1 - \mathbf{\Lambda}_2 + \mathbf{\Lambda}_3 - \mathbf{\Lambda}_4\right) \tag{19}$$

$$\mathbf{\Sigma}_4 = \frac{1}{4}\left(\mathbf{\Lambda}_1 - \mathbf{\Lambda}_2 - \mathbf{\Lambda}_3 + \mathbf{\Lambda}_4\right). \tag{20}$$

B. Practical Implications of Quaternion Properness

In this subsection, we point out the main implications of $\mathbb{C}^\eta$ and $\mathbb{Q}$ properness in the previous multivariate statistical analysis techniques. We will start by analyzing the case of jointly $\mathbb{C}^\eta$-proper vectors $\mathbf{x}$ and $\mathbf{y}$, which also paves the way for the $\mathbb{Q}$-proper case.

From the joint $\mathbb{C}^\eta$-properness definition it is clear that the matrices $\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{y}}}$, $\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}$ and $\mathbf{R}_{\bar{\mathbf{y}},\bar{\mathbf{y}}}$ (and, therefore, also $\mathbf{C}_{\bar{\mathbf{x}},\bar{\mathbf{y}}}$, $\mathbf{S}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}$ and $\mathbf{S}_{\bar{\mathbf{y}},\bar{\mathbf{y}}}$) take the block-diagonal structure

$$\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{y}}} = \begin{bmatrix} \mathbf{R}_{\tilde{\mathbf{x}},\tilde{\mathbf{y}}} & \mathbf{0}_{2n\times 2m} \\ \mathbf{0}_{2n\times 2m} & \mathbf{R}_{\tilde{\mathbf{x}},\tilde{\mathbf{y}}}^{(\eta')} \end{bmatrix}$$
$$\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}} = \begin{bmatrix} \mathbf{R}_{\tilde{\mathbf{x}},\tilde{\mathbf{x}}} & \mathbf{0}_{2n\times 2n} \\ \mathbf{0}_{2n\times 2n} & \mathbf{R}_{\tilde{\mathbf{x}},\tilde{\mathbf{x}}}^{(\eta')} \end{bmatrix}$$
$$\mathbf{R}_{\bar{\mathbf{y}},\bar{\mathbf{y}}} = \begin{bmatrix} \mathbf{R}_{\tilde{\mathbf{y}},\tilde{\mathbf{y}}} & \mathbf{0}_{2m\times 2m} \\ \mathbf{0}_{2m\times 2m} & \mathbf{R}_{\tilde{\mathbf{y}},\tilde{\mathbf{y}}}^{(\eta')} \end{bmatrix}$$

where $\mathbf{R}_{\tilde{\mathbf{x}},\tilde{\mathbf{y}}} = E\tilde{\mathbf{x}}\tilde{\mathbf{y}}^H$, $\mathbf{R}_{\tilde{\mathbf{x}},\tilde{\mathbf{x}}} = E\tilde{\mathbf{x}}\tilde{\mathbf{x}}^H$, $\mathbf{R}_{\tilde{\mathbf{y}},\tilde{\mathbf{y}}} = E\tilde{\mathbf{y}}\tilde{\mathbf{y}}^H$ are the semi-augmented (cross)-covariance matrices, which are obtained from the semi-augmented vectors $\tilde{\mathbf{x}} = \left[\mathbf{x}^T, \mathbf{x}^{(\eta)T}\right]^T$ and $\tilde{\mathbf{y}} = \left[\mathbf{y}^T, \mathbf{y}^{(\eta)T}\right]^T$. Thus, the block-diagonal structure also appears in the decomposition in (16), which can be written as $\mathbf{C}_{\bar{\mathbf{x}},\bar{\mathbf{y}}} = \overline{\mathbf{U}}\overline{\mathbf{\Lambda}}\overline{\mathbf{V}}^H$, with

$$\overline{\mathbf{U}} = \begin{bmatrix} \tilde{\mathbf{U}} & \mathbf{0}_{2n\times 2p} \\ \mathbf{0}_{2n\times 2p} & \tilde{\mathbf{U}}^{(\eta')} \end{bmatrix}$$
$$\overline{\mathbf{V}} = \begin{bmatrix} \tilde{\mathbf{V}} & \mathbf{0}_{2m\times 2p} \\ \mathbf{0}_{2m\times 2p} & \tilde{\mathbf{V}}^{(\eta')} \end{bmatrix}$$
$$\overline{\mathbf{\Lambda}} = \begin{bmatrix} \tilde{\mathbf{\Lambda}} & \mathbf{0}_{2p\times 2p} \\ \mathbf{0}_{2p\times 2p} & \tilde{\mathbf{\Lambda}} \end{bmatrix}$$

and

$$\tilde{\mathbf{U}} = \begin{bmatrix} \mathbf{U}_1 & \mathbf{U}_\eta \\ \mathbf{U}_\eta^{(\eta)} & \mathbf{U}_1^{(\eta)} \end{bmatrix}$$
$$\tilde{\mathbf{V}} = \begin{bmatrix} \mathbf{V}_1 & \mathbf{V}_\eta \\ \mathbf{V}_\eta^{(\eta)} & \mathbf{V}_1^{(\eta)} \end{bmatrix}$$
$$\tilde{\mathbf{\Lambda}} = \begin{bmatrix} \mathbf{\Sigma}_1 & \mathbf{\Sigma}_2 \\ \mathbf{\Sigma}_2 & \mathbf{\Sigma}_1 \end{bmatrix}.$$

Now, we can state the two following theorems.

Theorem 1: For jointly $\mathbb{C}^\eta$-proper vectors $\mathbf{x}$ and $\mathbf{y}$, the optimal PLS, MLR, and CCA projections reduce to semi-widely linear processing, i.e., they have the form

$$\mathbf{u} = \mathbf{F}_1^H\mathbf{x} + \mathbf{F}_\eta^H\mathbf{x}^{(\eta)}, \quad \mathbf{v} = \mathbf{G}_1^H\mathbf{y} + \mathbf{G}_\eta^H\mathbf{y}^{(\eta)}.$$

Proof: The proof follows directly from the structure of $\overline{\mathbf{U}}$, $\overline{\mathbf{V}}$ and the block-diagonality of $\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}$ and $\mathbf{R}_{\bar{\mathbf{y}},\bar{\mathbf{y}}}$. ■

Theorem 2: Given two jointly $\mathbb{C}^\eta$-proper vectors $\mathbf{x}$ and $\mathbf{y}$, the singular values $\lambda_1,\ldots,\lambda_{4p}$ of $\mathbf{C}_{\mathbf{r}_\mathbf{x},\mathbf{r}_\mathbf{y}}$ (and $\mathbf{C}_{\bar{\mathbf{x}},\bar{\mathbf{y}}}$) have multiplicity greater than or equal to two.

Proof: The block diagonal structure of $\overline{\mathbf{\Lambda}}$ implies $\mathbf{\Sigma}_3 = \mathbf{\Sigma}_4 = \mathbf{0}_{p\times p}$, which from (19) and (20) results in $\mathbf{\Lambda}_1 = \mathbf{\Lambda}_2$ and $\mathbf{\Lambda}_3 = \mathbf{\Lambda}_4$. ■

Theorem 1 constitutes a sufficient condition for the optimality of semi-widely linear processing. In other words, we should not expect any performance advantage from full-widely (instead of semi-widely) linear processing two jointly $\mathbb{C}^\eta$-proper vectors. However, we must note that the joint-properness is not a necessary condition. As a matter of fact, several relaxed sufficient conditions can be easily obtained by taking into account the particular expressions for $\mathbf{C}_{\bar{\mathbf{x}},\bar{\mathbf{y}}}$ (see Table II). On the other hand, Theorem 2 ensures that the augmented covariance matrices of $\mathbb{C}^\eta$-proper vectors have eigenvalues with multiplicity (at least) two, which is also the multiplicity of the singular values of the augmented cross-covariance matrices of cross $\mathbb{C}^\eta$-proper vectors.

In the case of jointly $\mathbb{Q}$-proper vectors $\mathbf{x}$ and $\mathbf{y}$, the analysis can be easily done following the previous lines. The two main results, which are analogous to those in Theorems 1 and 2 are the following.

Theorem 3: For jointly $\mathbb{Q}$-proper vectors $\mathbf{x}$ and $\mathbf{y}$, the optimal PLS, MLR and CCA projections reduce to conventional linear processing, i.e.,

$$\mathbf{u} = \mathbf{F}_1^H\mathbf{x}, \quad \mathbf{v} = \mathbf{G}_1^H\mathbf{y}.$$

Proof: The proof is based on the block-diagonality (four blocks of the same size) of the matrices in the decomposition $\mathbf{C}_{\bar{\mathbf{x}},\bar{\mathbf{y}}} = \overline{\mathbf{U}}\overline{\mathbf{\Lambda}}\overline{\mathbf{V}}^H$. ■

Theorem 4: Given two jointly $\mathbb{Q}$-proper vectors $\mathbf{x}$ and $\mathbf{y}$, the singular values $\lambda_1,\ldots,\lambda_{4p}$ of $\mathbf{C}_{\mathbf{r}_\mathbf{x},\mathbf{r}_\mathbf{y}}$ (and $\mathbf{C}_{\bar{\mathbf{x}},\bar{\mathbf{y}}}$) have multiplicity greater than or equal to four.

Proof: The block-diagonal structure of $\overline{\mathbf{\Lambda}}$ (four blocks of size $p\times p$) implies $\mathbf{\Sigma}_2 = \mathbf{\Sigma}_3 = \mathbf{\Sigma}_4 = \mathbf{0}_{p\times p}$, and combining (17)–(20), we obtain $\mathbf{\Lambda}_1 = \mathbf{\Lambda}_2 = \mathbf{\Lambda}_3 = \mathbf{\Lambda}_4$. ■

As can be seen, Theorem 3 ensures the optimality of conventional linear processing of jointly $\mathbb{Q}$-proper vectors (see Table II for more relaxed sufficient conditions), whereas Theorem 4 shows that the augmented (cross)-covariance matrices of (cross) $\mathbb{Q}$-proper vectors have singular values (or eigenvalues [3], [28]) with multiplicity (at least) four. Thus, Theorems 3 and 4 can be seen as extensions of Theorems 1 and 2. In particular, we already knew that if $\mathbf{x}$ and $\mathbf{y}$ are jointly $\mathbb{Q}$-proper, then they also are jointly $\mathbb{C}^\eta$-proper and Theorems 1 and 2 apply. However, the joint $\mathbb{Q}$-properness also implies joint $\mathbb{R}^\eta$-properness, which finally results in Theorems 3 and 4.

Finally, we must point out that the results in this section can be seen as an extension to quaternion vectors of the results in [14], [17]. Moreover, following the lines in [14], we could also introduce the concepts of generalized $\mathbb{C}^\eta$ and $\mathbb{Q}$-properness, which would be based on the multiplicities of the eigenvalues of the augmented covariance matrices, and would translate into similar results to those in [14] for the case of complex vectors.

V. Improperness Measures for Quaternion Vectors

In the case of complex random vectors, improperness measures have been proposed in [15], [20], [31]. Here, we extend this idea to the case of quaternion vectors. In particular, given a random vector $\mathbf{x} \in \mathbb{H}^{n\times 1}$ with augmented covariance matrix $\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}$, we propose to use the following improperness measure:

$$\mathcal{P} = \min_{\hat{\mathbf{R}}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}\in\mathcal{R}} D\left(\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}} \| \hat{\mathbf{R}}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}\right) \tag{21}$$

where $\mathcal{R}$ denotes the set of proper augmented covariance matrices (with the required kind of quaternion properness), and $D(\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}\|\hat{\mathbf{R}}_{\bar{\mathbf{x}},\bar{\mathbf{x}}})$ is the Kullback–Leibler divergence between two quaternion Gaussian distributions with zero mean and augmented covariance matrices $\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}$ and $\hat{\mathbf{R}}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}$.

TABLE III
PROBABILITY DENSITY FUNCTION, ENTROPY AND KULLBACK-LEIBLER DIVERGENCE OF QUATERNION GAUSSIAN VECTORS

	Probability density function (pdf)	Entropy	Kullback-Leibler Divergence
Expression from the real vector $\mathbf{r}_{\mathbf{x}}$	$p_{\mathbf{x}}(\mathbf{r}_{\mathbf{x}}) = \frac{\exp\left(-\frac{1}{2}\mathbf{r}_{\mathbf{x}}^T \mathbf{R}^{-1}_{\mathbf{r}_{\mathbf{x}},\mathbf{r}_{\mathbf{x}}} \mathbf{r}_{\mathbf{x}}\right)}{(2\pi)^{2n}\lvert\mathbf{R}_{\mathbf{r}_{\mathbf{x}},\mathbf{r}_{\mathbf{x}}}\rvert^{1/2}}$	$H_{\mathbf{x}}(\mathbf{R}_{\mathbf{r}_{\mathbf{x}},\mathbf{r}_{\mathbf{x}}}) = 2n\ln(2\pi e) + \frac{1}{2}\ln\lvert\mathbf{R}_{\mathbf{r}_{\mathbf{x}},\mathbf{r}_{\mathbf{x}}}\rvert$	$D(\mathbf{R}_{\mathbf{r}_{\mathbf{x}},\mathbf{r}_{\mathbf{x}}}\Vert\hat{\mathbf{R}}_{\mathbf{r}_{\mathbf{x}},\mathbf{r}_{\mathbf{x}}}) = \frac{1}{2}\ln\left(\frac{\lvert\hat{\mathbf{R}}_{\mathbf{r}_{\mathbf{x}},\mathbf{r}_{\mathbf{x}}}\rvert}{\lvert\mathbf{R}_{\mathbf{r}_{\mathbf{x}},\mathbf{r}_{\mathbf{x}}}\rvert}\right) + \frac{1}{2}\mathrm{Tr}\left(\hat{\mathbf{R}}^{-1}_{\mathbf{r}_{\mathbf{x}},\mathbf{r}_{\mathbf{x}}}\mathbf{R}_{\mathbf{r}_{\mathbf{x}},\mathbf{r}_{\mathbf{x}}}\right) - 2n$
Expression from the augmented vector $\bar{\mathbf{x}}$	$p_{\mathbf{x}}(\bar{\mathbf{x}}) = \frac{\exp\left(-\frac{1}{2}\bar{\mathbf{x}}^H \mathbf{R}^{-1}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}\bar{\mathbf{x}}\right)}{(\pi/2)^{2n}\lvert\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}\rvert^{1/2}}$	$H_{\mathbf{x}}(\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}) = 2n\ln\left(\frac{\pi e}{2}\right) + \frac{1}{2}\ln\lvert\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}\rvert$	$D(\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}\Vert\hat{\mathbf{R}}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}) = \frac{1}{2}\ln\left(\frac{\lvert\hat{\mathbf{R}}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}\rvert}{\lvert\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}\rvert}\right) + \frac{1}{2}\mathrm{Tr}\left(\hat{\mathbf{R}}^{-\frac{1}{2}}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}\hat{\mathbf{R}}^{-\frac{1}{2}}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}\right) - 2n$
Expression from the semi-augmented vector $\tilde{\mathbf{x}}$ ($\mathbb{C}^{\eta}$-proper case)	$p_{\mathbf{x}}(\tilde{\mathbf{x}}) = \frac{\exp\left(-\tilde{\mathbf{x}}^H \mathbf{R}^{-1}_{\tilde{\mathbf{x}},\tilde{\mathbf{x}}}\tilde{\mathbf{x}}\right)}{(\pi/2)^{2n}\lvert\mathbf{R}_{\tilde{\mathbf{x}},\tilde{\mathbf{x}}}\rvert}$	$H_{\mathbf{x}}(\mathbf{R}_{\tilde{\mathbf{x}},\tilde{\mathbf{x}}}) = 2n\ln\left(\frac{\pi e}{2}\right) + \ln\lvert\mathbf{R}_{\tilde{\mathbf{x}},\tilde{\mathbf{x}}}\rvert$	$D(\mathbf{R}_{\tilde{\mathbf{x}},\tilde{\mathbf{x}}}\Vert\hat{\mathbf{R}}_{\tilde{\mathbf{x}},\tilde{\mathbf{x}}}) = \ln\left(\frac{\lvert\hat{\mathbf{R}}_{\tilde{\mathbf{x}},\tilde{\mathbf{x}}}\rvert}{\lvert\mathbf{R}_{\tilde{\mathbf{x}},\tilde{\mathbf{x}}}\rvert}\right) + \mathrm{Tr}\left(\hat{\mathbf{R}}^{-\frac{1}{2}}_{\tilde{\mathbf{x}},\tilde{\mathbf{x}}}\mathbf{R}_{\tilde{\mathbf{x}},\tilde{\mathbf{x}}}\hat{\mathbf{R}}^{-\frac{1}{2}}_{\tilde{\mathbf{x}},\tilde{\mathbf{x}}}\right) - 2n$
Expression from the vector $\mathbf{x}$ ($\mathbb{Q}$-proper case)	$p_{\mathbf{x}}(\mathbf{x}) = \frac{\exp\left(-2\mathbf{x}^H \mathbf{R}^{-1}_{\mathbf{x},\mathbf{x}}\mathbf{x}\right)}{(\pi/2)^{2n}\lvert\mathbf{R}_{\mathbf{x},\mathbf{x}}\rvert^{2}}$	$H_{\mathbf{x}}(\mathbf{R}_{\mathbf{x},\mathbf{x}}) = 2n\ln\left(\frac{\pi e}{2}\right) + 2\ln\lvert\mathbf{R}_{\mathbf{x},\mathbf{x}}\rvert$	$D(\mathbf{R}_{\mathbf{x},\mathbf{x}}\Vert\hat{\mathbf{R}}_{\mathbf{x},\mathbf{x}}) = 2\ln\left(\frac{\lvert\hat{\mathbf{R}}_{\mathbf{x},\mathbf{x}}\rvert}{\lvert\mathbf{R}_{\mathbf{x},\mathbf{x}}\rvert}\right) + 2\mathrm{Tr}\left(\hat{\mathbf{R}}^{-\frac{1}{2}}_{\mathbf{x},\mathbf{x}}\mathbf{R}_{\mathbf{x},\mathbf{x}}\hat{\mathbf{R}}^{-\frac{1}{2}}_{\mathbf{x},\mathbf{x}}\right) - 2n$

The probability density function (pdf) of quaternion Gaussian vectors can be easily obtained from the pdf of the real vector $\mathbf{r}_{\mathbf{x}}$ (see also [18], [32] for previous works on quaternion Gaussian vectors), and it can be simplified in the case of $\mathbb{C}^{\eta}$-proper or $\mathbb{Q}$-proper vectors. Table III shows the pdf, entropy, and Kullback–Leibler divergence expressions for quaternion Gaussian vectors.[10]

Before proceeding, we must remark the following reasons for the choice of the measure in (21).

- First, the Gaussian assumption is justified by the fact that Gaussian vectors are completely specified by their second-order statistics. Therefore, the improperness measure should also be a noncircularity measure for Gaussian vectors.
- As we have pointed out in Lemma 1, the structure of the augmented covariance matrix $\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}$ is invariant under quaternion linear transformations. As we will see later, the improperness measure in (21) preserves this invariance. Moreover, in the case of $\mathbb{C}^{\eta}$-properness, it is also invariant to semi-widely linear transformations.
- The choice of the Kullback–Leibler divergence is justified by its information-theoretic implications. On one hand, the measure in (21) is closely related to the concepts of entropy and mutual information. On the other hand, $D(\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}\Vert\hat{\mathbf{R}}_{\bar{\mathbf{x}},\bar{\mathbf{x}}})$ provides the error exponent of the Neyman-Pearson detector for the binary hypothesis testing problem of deciding whether a set of i.i.d. vector observations belongs to a zero-mean Gaussian distribution with augmented covariance matrix $\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}$ or $\hat{\mathbf{R}}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}$ [33].[11] Moreover, taking into account the minimization in (21), $\mathcal{P}$ can be interpreted as a *worst-case* error exponent, or equivalently, as the error exponent associated to the problem of deciding between $\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}$ and $\mathcal{R}$, i.e., *all* the augmented covariance matrices with the required properness structure.

A. Measure of $\mathbb{Q}$-Improperness

Let us start our analysis by the strongest kind of quaternion properness. The set $\mathcal{R}_{\mathbb{Q}}$ of $\mathbb{Q}$-proper augmented covariance matrices is

$$\mathcal{R}_{\mathbb{Q}} = \left\{\hat{\mathbf{R}}_{\bar{\mathbf{x}},\bar{\mathbf{x}}} \middle| \hat{\mathbf{R}}_{\mathbf{x},\mathbf{x}^{(\eta)}} = \hat{\mathbf{R}}_{\mathbf{x},\mathbf{x}^{(\eta')}} = \hat{\mathbf{R}}_{\mathbf{x},\mathbf{x}^{(\eta'')}} = \mathbf{0}_{n\times n}\right\}$$

[10]Note that, due to the noncommutativity of the quaternion product, the term $\mathrm{Tr}\left(\hat{\mathbf{R}}^{-1}_{\mathbf{r}_{\mathbf{x}},\mathbf{r}_{\mathbf{x}}}\mathbf{R}_{\mathbf{r}_{\mathbf{x}},\mathbf{r}_{\mathbf{x}}}\right)$ in the Kullback–Leibler expression has to be rewritten as $\mathrm{Tr}\left(\hat{\mathbf{R}}^{-1}_{\mathbf{r}_{\mathbf{x}},\mathbf{r}_{\mathbf{x}}}\mathbf{R}_{\mathbf{r}_{\mathbf{x}},\mathbf{r}_{\mathbf{x}}}\right) = \mathrm{Tr}\left(\hat{\mathbf{R}}^{-1/2}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}\hat{\mathbf{R}}^{-1/2}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}\right)$. Alternatively, we could have written $\mathrm{Tr}\left(\hat{\mathbf{R}}^{-1}_{\mathbf{r}_{\mathbf{x}},\mathbf{r}_{\mathbf{x}}}\mathbf{R}_{\mathbf{r}_{\mathbf{x}},\mathbf{r}_{\mathbf{x}}}\right) = \Re\left[\mathrm{Tr}\left(\hat{\mathbf{R}}^{-1}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}\right)\right]$, where $\Re(a)$ denotes the real part of the quaternion a.

[11]The error exponent is defined as the rate of exponential decay of the miss probability under a constant false alarm probability. Here, the miss probability is the probability of deciding $\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}$ when $\hat{\mathbf{R}}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}$ is true.

and the matrix $\hat{\mathbf{R}}_{\bar{\mathbf{x}},\bar{\mathbf{x}}} \in \mathcal{R}_{\mathbb{Q}}$ minimizing $D(\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}\|\hat{\mathbf{R}}_{\bar{\mathbf{x}},\bar{\mathbf{x}}})$ is

$$\hat{\mathbf{R}}_{\bar{\mathbf{x}},\bar{\mathbf{x}}} = \mathbf{D}_{\mathbb{Q}} = \begin{bmatrix} \mathbf{R}_{\mathbf{x},\mathbf{x}} & \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} \\ \mathbf{0}_{n\times n} & \mathbf{R}^{(\eta)}_{\mathbf{x},\mathbf{x}} & \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} \\ \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} & \mathbf{R}^{(\eta')}_{\mathbf{x},\mathbf{x}} & \mathbf{0}_{n\times n} \\ \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} & \mathbf{R}^{(\eta'')}_{\mathbf{x},\mathbf{x}} \end{bmatrix}.$$

Thus, the $\mathbb{Q}$-improperness measure reduces to

$$\mathcal{P}_{\mathbb{Q}} = -\frac{1}{2}\ln|\mathbf{\Phi}_{\mathbb{Q}}|$$

where we have defined $\mathbf{\Phi}_{\mathbb{Q}} = \mathbf{D}_{\mathbb{Q}}^{-\frac{1}{2}}\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}\mathbf{D}_{\mathbb{Q}}^{-\frac{1}{2}}$ as the $\mathbb{Q}$-coherence matrix. Interestingly, this matrix naturally appears in the quaternion version of the maximum variance (MAXVAR) generalization of canonical correlation analysis (CCA) to four random vectors [26], [34], [35]. Therefore, the $\mathbb{Q}$-improperness measure is obtained from the canonical correlation analysis of the random vectors $\mathbf{x}$, $\mathbf{x}^{(\eta)}$, $\mathbf{x}^{(\eta')}$ and $\mathbf{x}^{(\eta'')}$. Furthermore, we can easily check that $\mathcal{P}_{\mathbb{Q}}$ is invariant under rotations $\mathbf{x}^{(a)}$, basis changes, linear transformations $\mathbf{F}_1^H\mathbf{x}$, and it can also be written as

$$\mathcal{P}_{\mathbb{Q}} = H_{\mathbf{x}}(\mathbf{D}_{\mathbf{Q}}) - H_{\mathbf{x}}(\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}) = H_{\mathbf{x}}(\mathbf{R}_{\mathbf{x},\mathbf{x}}) - H_{\mathbf{x}}(\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}})$$

which represents the entropy loss due to the improperness of $\mathbf{x}$. That is, $\mathcal{P}_{\mathbb{Q}}$ can be seen as a measure of the mutual information among the random vectors $\mathbf{x}$, $\mathbf{x}^{(\eta)}$, $\mathbf{x}^{(\eta')}$ and $\mathbf{x}^{(\eta'')}$ [36].

B. Measure of $\mathbb{C}^{\eta}$-Improperness

In this case, the set of $\mathbb{C}^{\eta}$-proper augmented covariance matrices is $\mathcal{R}_{\mathbb{C}^{\eta}} = \left\{\hat{\mathbf{R}}_{\bar{\mathbf{x}},\bar{\mathbf{x}}} | \hat{\mathbf{R}}_{\mathbf{x},\mathbf{x}^{(\eta')}} = \hat{\mathbf{R}}_{\mathbf{x},\mathbf{x}^{(\eta'')}} = \mathbf{0}_{n\times n}\right\}$, and the matrix $\hat{\mathbf{R}}_{\bar{\mathbf{x}},\bar{\mathbf{x}}} \in \mathcal{R}_{\mathbb{C}^{\eta}}$ minimizing $D(\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}\|\hat{\mathbf{R}}_{\bar{\mathbf{x}},\bar{\mathbf{x}}})$ is

$$\hat{\mathbf{R}}_{\bar{\mathbf{x}},\bar{\mathbf{x}}} = \mathbf{D}_{\mathbb{C}^{\eta}} = \begin{bmatrix} \mathbf{R}_{\tilde{\mathbf{x}},\tilde{\mathbf{x}}} & \mathbf{0}_{2n\times 2n} \\ \mathbf{0}_{2n\times 2n} & \mathbf{R}^{(\eta')}_{\tilde{\mathbf{x}},\tilde{\mathbf{x}}} \end{bmatrix}.$$

Therefore, the $\mathbb{C}^{\eta}$-improperness measure reduces to

$$\mathcal{P}_{\mathbb{C}^{\eta}} = -\frac{1}{2}\ln|\mathbf{\Phi}_{\mathbb{C}^{\eta}}|$$

where $\mathbf{\Phi}_{\mathbb{C}^{\eta}} = \mathbf{D}_{\mathbb{C}^{\eta}}^{-\frac{1}{2}}\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}\mathbf{D}_{\mathbb{C}^{\eta}}^{-\frac{1}{2}}$ is the $\mathbb{C}^{\eta}$-coherence matrix in the quaternion extension of CCA for the random vectors $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{x}}^{(\eta')}$. Furthermore, it is easy to prove that $\mathcal{P}_{\mathbb{C}^{\eta}}$ is invariant to semi-widely linear transformations $\mathbf{F}_1^H\mathbf{x} + \mathbf{F}_{\eta}^H\mathbf{x}^{(\eta)}$, and it also represents the entropy loss due to the $\mathbb{C}^{\eta}$-improperness of $\mathbf{x}$, i.e.,

$$\mathcal{P}_{\mathbb{C}^{\eta}} = H_{\mathbf{x}}(\mathbf{D}_{\mathbb{C}^{\eta}}) - H_{\mathbf{x}}(\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}) = H_{\mathbf{x}}(\mathbf{R}_{\tilde{\mathbf{x}},\tilde{\mathbf{x}}}) - H_{\mathbf{x}}(\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}).$$

Additionally, rewriting the semi-augmented vector $\tilde{\mathbf{x}}$ in terms of the Cayley–Dickson representation

$$\underbrace{\begin{bmatrix} \mathbf{x} \\ \mathbf{x}^{(\eta)} \end{bmatrix}}_{\tilde{\mathbf{x}}} = \underbrace{\left(\begin{bmatrix} 1 & \eta'' \\ 1 & -\eta'' \end{bmatrix} \otimes \mathbf{I}_{n,n}\right)}_{\mathbf{L}} \underbrace{\begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \end{bmatrix}}_{\mathbf{a}}$$

and taking into account the unitarity of the operator $\mathbf{L}$, the $\mathbb{C}^{\eta}$-improperness measure can be rewritten as

$$\mathcal{P}_{\mathbb{C}^{\eta}} = -\frac{1}{2}\ln|\mathbf{\Phi}_{\tilde{\mathbf{a}}}|$$

where $\mathbf{\Phi}_{\tilde{\mathbf{a}}} = \mathbf{D}_{\tilde{\mathbf{a}}}^{-\frac{1}{2}}\mathbf{R}_{\tilde{\mathbf{a}},\tilde{\mathbf{a}}}\mathbf{D}_{\tilde{\mathbf{a}}}^{-\frac{1}{2}}$ is the coherence matrix for the complex vector $\tilde{\mathbf{a}} = \left[\mathbf{a}^T, \mathbf{a}^H\right]^T$, and

$$\mathbf{D}_{\tilde{\mathbf{a}}} = \begin{bmatrix} \mathbf{R}_{\mathbf{a},\mathbf{a}} & \mathbf{0}_{n\times n} \\ \mathbf{0}_{n\times n} & \mathbf{R}^*_{\mathbf{a},\mathbf{a}} \end{bmatrix}.$$

Thus, the $\mathbb{C}^{\eta}$-improperness measure reduces to an improperness measure of the complex vector $\mathbf{a}$ [15], [20], [31], which is also a measure of the degree of joint-improperness of the complex vectors $\mathbf{a}_1, \mathbf{a}_2$. That is, as pointed out in Section III-C, the $\mathbb{C}^{\eta}$-properness of a vector $\mathbf{x}$ can be seen as the joint-properness of the complex vectors in the Cayley–Dickson representation $\mathbf{x} = \mathbf{a}_1 + \eta''\mathbf{a}_2$.

C. Measure of $\mathbb{R}^{\eta}$-Improperness

For the $\mathbb{R}^{\eta}$-improperness measure, the problem is more involved than in the previous cases. This is due to the fact that, given the set $\mathcal{R}_{\mathbb{R}^{\eta}} = \left\{\hat{\mathbf{R}}_{\bar{\mathbf{x}},\bar{\mathbf{x}}} | \hat{\mathbf{R}}_{\mathbf{x},\mathbf{x}^{(\eta)}} = \mathbf{0}_{n\times n}\right\}$, obtaining the matrix $\hat{\mathbf{R}}_{\bar{\mathbf{x}},\bar{\mathbf{x}}} \in \mathcal{R}_{\mathbb{R}^{\eta}}$ minimizing $D(\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}\|\hat{\mathbf{R}}_{\bar{\mathbf{x}},\bar{\mathbf{x}}})$ is far from trivial, and it is closely related to the problem of maximum likelihood estimation of structured covariance matrices [23], [37].

Here we focus on an alternative and more meaningful measure. In particular, we consider the measurement of the $\mathbb{R}^{\eta}$-improperness of $\mathbb{C}^{\eta}$-proper vectors. That is, given an augmented covariance matrix $\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}} \in \mathcal{R}_{\mathbb{C}^{\eta}}$, we look for the closest (in the Kullback–Leibler sense) matrix $\hat{\mathbf{R}}_{\bar{\mathbf{x}},\bar{\mathbf{x}}} \in \mathcal{R}_{\mathbb{R}^{\eta}}$, and with a slight abuse of notation define

$$\mathcal{P}_{\mathbb{R}^{\eta}} = D(\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}\|\hat{\mathbf{R}}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}).$$

Thus, following the lines in the previous subsections, the $\mathbb{R}^{\eta}$-improperness measure reduces to

$$\begin{aligned}\mathcal{P}_{\mathbb{R}^{\eta}} &= -\frac{1}{2}\ln|\mathbf{D}_{\mathbb{Q}}^{-\frac{1}{2}}\mathbf{D}_{\mathbb{C}^{\eta}}\mathbf{D}_{\mathbb{Q}}^{-\frac{1}{2}}| \\ &= -\ln|\mathbf{D}_{\mathbb{R}^{\eta}}^{-\frac{1}{2}}\mathbf{R}_{\tilde{\mathbf{x}},\tilde{\mathbf{x}}}\mathbf{D}_{\mathbb{R}^{\eta}}^{-\frac{1}{2}}| = -\ln|\mathbf{\Phi}_{\mathbb{R}^{\eta}}|\end{aligned}$$

where

$$\mathbf{D}_{\mathbb{R}^{\eta}} = \begin{bmatrix} \mathbf{R}_{\mathbf{x},\mathbf{x}} & \mathbf{0}_{n\times n} \\ \mathbf{0}_{n\times n} & \mathbf{R}^{(\eta)}_{\mathbf{x},\mathbf{x}} \end{bmatrix}$$

and $\mathbf{\Phi}_{\mathbb{R}^{\eta}} = \mathbf{D}_{\mathbb{R}^{\eta}}^{-\frac{1}{2}}\mathbf{R}_{\tilde{\mathbf{x}},\tilde{\mathbf{x}}}\mathbf{D}_{\mathbb{R}^{\eta}}^{-\frac{1}{2}}$ is the $\mathbb{R}^{\eta}$-coherence matrix, which appears in the canonical correlation analysis of the random vectors $\mathbf{x}$ and $\mathbf{x}^{(\eta)}$. Finally, analogously to the previous cases, the measure $\mathcal{P}_{\mathbb{R}^{\eta}}$ is invariant to linear transformations, and it provides the entropy loss due to the $\mathbb{R}^{\eta}$-improperness of the vector $\mathbf{x}$

$$\begin{aligned}\mathcal{P}_{\mathbb{R}^{\eta}} &= H_{\mathbf{x}}(\mathbf{D}_{\mathbb{Q}}) - H_{\mathbf{x}}(\mathbf{D}_{\mathbb{C}^{\eta}}) \\ &= H_{\mathbf{x}}(\mathbf{D}_{\mathbb{R}^{\eta}}) - H_{\mathbf{x}}(\mathbf{R}_{\tilde{\mathbf{x}},\tilde{\mathbf{x}}}) = H_{\mathbf{x}}(\mathbf{R}_{\mathbf{x},\mathbf{x}}) - H_{\mathbf{x}}(\mathbf{R}_{\tilde{\mathbf{x}},\tilde{\mathbf{x}}})\end{aligned}$$

or equivalently, the mutual information between $\mathbf{x}$ and $\mathbf{x}^{(\eta)}$.

D. Further Comments

As we have shown, the three proposed improperness measures are directly related to the canonical correlation analysis technique and its extension to four random vectors. In the case of complex vectors, similar results have been obtained in [15],

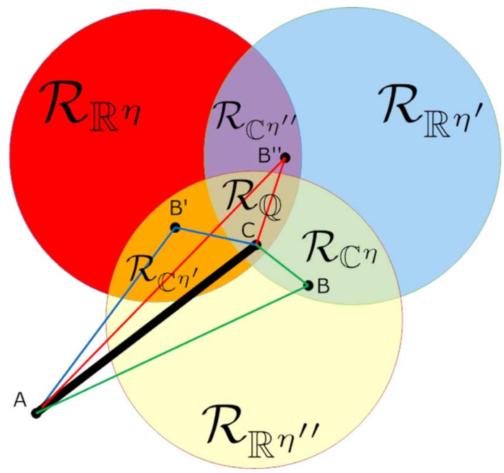

Fig. 1. Illustration of the $\mathbb{Q}$-improperness measure decomposition. The figure shows the sets of $\mathbb{R}$-proper ($\mathcal{R}_{\mathbb{R}^\eta}$, $\mathcal{R}_{\mathbb{R}^{\eta''}}$ and $\mathcal{R}_{\mathbb{R}^{\eta''}}$), $\mathbb{C}$-proper ($\mathcal{R}_{\mathbb{C}^\eta}$, $\mathcal{R}_{\mathbb{C}^{\eta''}}$, and $\mathcal{R}_{\mathbb{C}^{\eta''}}$), and $\mathbb{Q}$-proper ($\mathcal{R}_{\mathbb{Q}}$) augmented covariance matrices. Point A represents a general augmented covariance matrix $\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}$. B is the closest (in the Kullback–Leibler sense) point to A in $\mathcal{R}_{\mathbb{C}^\eta}$ (matrix $\mathbf{D}_{\mathbb{C}^\eta}$). C (matrix $\mathbf{D}_{\mathbb{Q}}$) is the projection of A onto $\mathcal{R}_{\mathbb{Q}}$, which coincides with the projection of B onto $\mathcal{R}_{\mathbb{R}^\eta}$. The length of the segment $\overline{AC}$ represents the measure $\mathcal{P}_{\mathbb{Q}}$, which is equal to the sum of the lengths of the segments $\overline{AB}$ ($\mathcal{P}_{\mathbb{C}^\eta}$) and $\overline{BC}$ ($\mathcal{P}_{\mathbb{R}^\eta}$). The same interpretation can be done in terms of the points B' and B''.

[20], [31], where the authors have shown that the canonical correlations (eigenvalues of the coherence matrix) provide a measure of improperness, entropy loss and mutual information.

Interestingly, the improperness measures proposed in this paper satisfy

$$\mathcal{P}_{\mathbb{Q}} = \mathcal{P}_{\mathbb{C}^\eta} + \mathcal{P}_{\mathbb{R}^\eta} \tag{22}$$

which can be seen as a direct consequence of the Pythagorean theorem for exponential families of pdf's [38], [39], and corroborates our intuition about the complementarity of $\mathbb{C}^\eta$ and $\mathbb{R}^\eta$ properness. Moreover, since the $\mathbb{Q}$-improperness measure does not depend on the orthogonal basis $\{1, \eta, \eta', \eta''\}$, $\mathcal{P}_{\mathbb{Q}}$ can be decomposed as (22) for all pure unit quaternions η. In other words, the Kullback–Leibler "distance" from an augmented covariance matrix $\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}$ to the closest $\mathbb{Q}$-proper matrix $\mathbf{D}_{\mathbb{Q}}$ can be calculated as the divergence from $\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}$ to the closest $\mathbb{C}^\eta$-proper matrix $\mathbf{D}_{\mathbb{C}^\eta}$, plus the divergence from $\mathbf{D}_{\mathbb{C}^\eta}$ to the closest $\mathbb{R}^\eta$-proper matrix $\mathbf{D}_{\mathbb{Q}}$. This fact is illustrated in Fig. 1 for three orthogonal pure unit quaternions η, η' and η''.

VI. Conclusion

The properness of quaternion-valued random vectors has been analyzed, showing its similarities and differences with the complex case. In particular, the second-order statistics of quaternion vectors are captured by the covariance matrix and three complementary covariance matrices, which are obtained as the correlation between the quaternion vector and its involutions over three pure unit quaternions. The existence of three complementary covariance matrices translates into three different kinds of properness, all of them with direct implications on the Cayley–Dickson representations of the quaternion vector. Analogously to the complex case, the optimal linear processing of quaternion vectors is in general full-widely linear, which means that we have to simultaneously operate on the quaternion vector and its involutions. However, in the case of $\mathbb{Q}$-proper and $\mathbb{C}^\eta$-proper vectors, the optimal processing reduces to conventional and semi-widely linear processing, respectively. Finally, the improperness of a quaternion vector can be measured by the Kullback–Leibler divergence between two Gaussian distributions, one of them with the augmented covariance matrix, and the other with its closest proper version. This measure, which is closely related to the canonical correlation analysis technique, provides the entropy loss due to the improperness of the quaternion vector, and it admits a straightforward geometrical interpretation based on Kullback–Leibler projections onto different sets of proper augmented covariance matrices.

Acknowledgment

The authors would like to thank the anonymous reviewers for their valuable suggestions, especially those regarding the invariance of the second-order statistics to Clifford translations.

References

[1] W. R. Hamilton, "On quaternions," in *Proc. Royal Irish Acad.*, Nov. 11, 1844.
[2] S. Miron, N. Le Bihan, and J. Mars, "Quaternion-music for vector-sensor array processing," *IEEE Trans. Signal Process.*, vol. 54, no. 4, pp. 1218–1229, Apr. 2006.
[3] N. Le Bihan and J. Mars, "Singular value decomposition of quaternion matrices: A new tool for vector-sensor signal processing," *Signal Process.*, vol. 84, no. 7, pp. 1177–1199, Jul. 2004.
[4] N. Le Bihan and S. Buchholz, "Optimal separation of polarized signals by quaternionic neural networks," presented at the XIV European Signal Processing Conference (EUSIPCO), Florence, Italy, 2006.
[5] S. Buchholz and N. Le Bihan, "Polarized signal classification by complex and quaternionic multi-layer perceptrons," *Int. J. Neural Syst.*, vol. 18, no. 2, pp. 75–85, 2008.
[6] J. Seberry, K. Finlayson, S. Adams, T. Wysocki, T. Xia, and B. Wysocki, "The theory of quaternion orthogonal designs," *IEEE Trans. Signal Process.*, vol. 56, no. 1, pp. 256–265, Jan. 2008.
[7] C. Took and D. Mandic, "The quaternion LMS algorithm for adaptive filtering of hypercomplex processes," *IEEE Trans. Signal Process.*, vol. 57, no. 4, pp. 1316–1327, Apr. 2009.
[8] F. Neeser and J. Massey, "Proper complex random processes with applications to information theory," *IEEE Trans. Inf. Theory*, vol. 39, no. 7, pp. 1293–1302, Jul. 1993.
[9] B. Picinbono, "On circularity," *IEEE Trans. Signal Processing.*, vol. 42, no. 12, pp. 3473–3482, Dec. 1994.
[10] B. Picinbono and P. Chevalier, "Widely linear estimation with complex data," *IEEE Trans. Signal Process.*, vol. 43, no. 8, pp. 2030–2033, Aug. 1995.
[11] B. Picinbono, "Second-order complex random vectors and normal distributions," *IEEE Trans. Signal Process.*, vol. 44, no. 10, pp. 2637–2640, Oct. 1996.
[12] B. Picinbono and P. Bondon, "Second-order statistics of complex signals," *IEEE Trans. Signal Process.*, vol. 45, no. 2, pp. 411–420, Feb. 1997.
[13] A. van den Bos, "The multivariate complex normal distribution-a generalization," *IEEE Trans. Inf. Theory*, vol. 41, no. 3, pp. 537–539, Mar. 1995.
[14] P. Schreier and L. Scharf, "Second-order analysis of improper complex random vectors and processes," *IEEE Trans. Signal Process.*, vol. 51, no. 3, pp. 714–725, Mar. 2003.
[15] P. Schreier, L. Scharf, and A. Hanssen, "A generalized likelihood ratio test for impropriety of complex signals," *IEEE Signal Process. Lett.*, vol. 13, no. 7, pp. 433–436, Jul. 2006.
[16] A. Walden and P. Rubin-Delanchy, "On testing for impropriety of complex-valued Gaussian vectors," *IEEE Trans. Signal Process.*, vol. 57, no. 3, pp. 825–834, Mar. 2009.
[17] P. Schreier, "A unifying discussion of correlation analysis for complex random vectors," *IEEE Trans. Signal Process.*, vol. 56, no. 4, pp. 1327–1336, Apr. 2008.
[18] N. N. Vakhania, "Random vectors with values in quaternion Hilbert spaces," *Theory Probab. Appl.*, vol. 43, no. 1, pp. 99–115, 1999.

[19] P. Amblard and N. Le Bihan, "On properness of quaternion valued random variables," in *Proc. IMA Conf. Mathematics in Signal Processing*, Cirencester, U.K., 2004, pp. 23–26.
[20] J. Eriksson and V. Koivunen, "Complex random vectors and ICA models: Identifiability, uniqueness, and separability," *IEEE Trans. Inf. Theory*, vol. 52, no. 3, pp. 1017–1029, Mar. 2006.
[21] K. I. Diamantaras and S. Y. Kung, *Principal Component Neural Networks, Theory and Applications*. Hoboken, NJ: Wiley, 1996.
[22] H. Wold, *Encyclopedia of the Statistical Sciences*. Hoboken, NJ: Wiley, 1985, ch. Partial least squares, pp. 581–591.
[23] L. Scharf, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*. Reading, MA: Addison-Wesley, 1990.
[24] H. Hotelling, "Relations between two sets of variates," *Biometrika*, vol. 28, pp. 321–377, 1936.
[25] L. L. Scharf and C. T. Mullis, "Canonical coordinates and the geometry of inference, rate and capacity," *IEEE Trans. Signal Process.*, vol. 48, no. 3, pp. 824–831, Mar. 2000.
[26] J. R. Kettenring, "Canonical analysis of several sets of variables," *Biometrika*, vol. 58, no. 3, pp. 433–451, 1971.
[27] J. P. Ward, *Quaternions and Cayley Numbers: Algebra and Applications*. Dordrecht, The Netherlands: Kluwer, 1997.
[28] F. Zhang, "Quaternions and matrices of quaternions," *Linear Algebra Appl.*, vol. 251, pp. 21–57, 1997.
[29] H. S. M. Coxeter, "Quaternions and reflections," *Amer. Math. Month.*, vol. 53, no. 3, pp. 136–146, 1946.
[30] K. J. Worsley, J.-B. Poline, K. J. Friston, and A. C. Evans, "Characterizing the response of PET and fMRI data using multivariate linear models," *NeuroImage*, vol. 6, no. 4, pp. 305–319, Nov. 1997.
[31] P. Schreier, L. Scharf, and C. Mullis, "Detection and estimation of improper complex random signals," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 306–312, Jan. 2005.
[32] N. Le Bihan and P. O. Amblard, "Detection and estimation of Gaussian proper quaternion valued random processes," presented at the 7th IMA Conf. Mathematics in Signal Processing, Cirencester, U.K., Dec. 2006.
[33] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Hoboken, NJ: Wiley, 1991.
[34] J. Vía, I. Santamaría, and J. Pérez, "A learning algorithm for adaptive canonical correlation analysis of several data sets," *Neural Netw.*, vol. 20, no. 1, pp. 139–152, Jan. 2007.
[35] J. Vía, I. Santamaría, and J. Pérez, "Deterministic CCA-based algorithms for blind equalization of FIR-MIMO channels," *IEEE Trans. Signal Process.*, vol. 55, no. 7, pp. 3867–3878, Jul. 2007.
[36] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*. Hoboken, NJ: Wiley, 2001.
[37] J. Burg, D. Luenberger, and D. Wenger, "Estimation of structured covariance matrices," *Proc. IEEE*, vol. 70, no. 7, pp. 963–974, Sep. 1982.
[38] S.-I. Amari, H. Nagaoka, and A. M. E. Society, *Methods of Information Geometry*. Oxford, U.K.: Oxford Univ. Press, 1993.
[39] J. Cardoso, *Unsupervised Adaptive Filtering, Volume 1, Blind Source Separation*. Hoboken, NJ: Wiley, 2000, vol. 1, ch. Entropic contrasts for source separation: geometry and stability, pp. 139–190.

Javier Vía (M'08) received the Telecommunication Engineer Degree and the Ph.D. degree in electrical engineering from the University of Cantabria, Spain, in 2002 and 2007, respectively.

In 2002, he joined the Department of Communications Engineering, University of Cantabria, Spain, where he is currently an Assistant Professor. He has spent visiting periods at the Smart Antennas Research Group of Stanford University and at the Department of Electronics and Computer Engineering (Hong Kong University of Science and Technology). He has actively participated in several European and Spanish research projects. His current research interests include blind channel estimation and equalization in wireless communication systems, multivariate statistical analysis, quaternion signal processing, and kernel methods.

David Ramírez (S'07) received the Telecommunication Engineer Degree from the University of Cantabria, Spain, in 2006. He is currently pursuing the Ph.D. degree at the Communications Engineering Department, University of Cantabria, Spain, under the supervision of I. Santamaría and J. Vía.

In 2009, he spent a visiting period at the University of Newcastle under the supervision of Prof. P. J. Schreier. His current research interests include Signal Processing for wireless communications, MIMO Systems, MIMO testbeds, and multivariate statistical analysis.

Ignacio Santamaría (M'96–SM'05) received the Telecommunication Engineer Degree and the Ph.D. degree in electrical engineering from the Universidad Politécnica de Madrid (UPM), Spain, in 1991 and 1995, respectively.

In 1992, he joined the Departamento de Ingeniería de Comunicaciones, Universidad de Cantabria, Spain, where he is currently a Full Professor. He has been a visiting researcher at the Computational NeuroEngineering Laboratory (University of Florida) and at the Wireless Networking and Communications Group (University of Texas at Austin). He has authored more than 100 publications in refereed journals and international conference papers. His current research interests include signal processing algorithms for wireless communication systems, MIMO systems, multivariate statistical techniques, and machine learning theories. He has been involved in several national and international research projects on these topics. He is currently serving as a member of the Machine Learning for Signal Processing Technical Committee of the IEEE Signal Processing Society.