WIDELY AND SEMI-WIDELY LINEAR PROCESSING OF QUATERNION VECTORS

Javier Vía, David Ramírez, Ignacio Santamaría and Luis Vielva

Dept. of Communications Engineering. University of Cantabria. Spain
e-mail: {jvia,ramirezgd,nacho,luis}@gtas.dicom.unican.es

ABSTRACT

In this paper the two main definitions of quaternion properness (or second order circularity) are reviewed, showing their connection with the structure of the optimal quaternion linear processing. Specifically, we present a rigorous generalization of the most common multivariate statistical analysis techniques to the case of quaternion vectors, and show that the different kinds of quaternion improperness require different kinds of widely linear processing. In general, the optimal linear processing is *full-widely linear*, which requires the joint processing of the quaternion vector and its involutions over three pure unit quaternions. However, in the case of jointly $\mathbb{Q}$-proper and $\mathbb{C}^\eta$-proper vectors, the optimal processing reduces, respectively, to the *conventional* and *semi-widely linear processing*, with the latter only requiring to operate on the quaternion vector and its involution over the pure unit quaternion η. Finally, a simulation example poses some interesting questions for future research.

***Index Terms*—** Quaternions, properness, second-order circularity, widely linear processing.

1. INTRODUCTION

Although quaternion algebra [1] has been successfully applied to several signal processing and communications problems [2, 3, 4], the properness (or second order circularity) analysis of quaternion random vectors has received limited attention [5, 6, 7], and a clear definition of quaternion widely linear processing is still lacking [4].

In this paper the two main kinds of quaternion properness are reviewed, and their implications on the optimal linear processing are analyzed. In particular, we revisit the concepts of $\mathbb{Q}$-properness and $\mathbb{C}^\eta$-properness [5, 6, 7], showing the relationships between the complementary covariance matrices and the Cayley-Dickson representations of the quaternion vector, which results in new insights on the structure of proper quaternion vectors. Furthermore, we present a unified approach to quaternion multivariate statistical analysis and show that, in general, the optimal linear processing is *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 $\mathbb{Q}$-proper vectors, the optimal processing is linear, i.e., we do not need to operate on the vector involutions, whereas in the case of $\mathbb{C}^\eta$-proper vectors, 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 widely linear processing.

This work was supported by the Spanish Government, Ministerio de Ciencia e Innovación (MICINN), under project MultiMIMO (TEC2007-68020-C04-02), project COMONSENS (CSD2008-00010, CONSOLIDER-INGENIO 2010 Program) and FPU grant AP2006-2965.

2. PRELIMINARIES

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 $(\cdot)^T$ and $(\cdot)^H$ denote transpose and Hermitian (i.e., transpose and quaternion conjugate), respectively. The notation $\mathbf{A} \in \mathbb{R}^{m\times n}$ (respectively $\mathbf{A} \in \mathbb{H}^{m\times n}$) means that $\mathbf{A}$ is a real (respectively quaternion) $m \times n$ matrix. $\text{Tr}(\mathbf{A})$ denotes the trace of matrix $\mathbf{A}$, $\otimes$ 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. Finally, $\text{diag}(\mathbf{a})$ denotes the diagonal matrix with vector $\mathbf{a}$ along its diagonal, 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., $\mathbf{R}_{\mathbf{a},\mathbf{b}} = E\mathbf{a}\mathbf{b}^H$.

2.1. Quaternion Algebra

Quaternions are four-dimensional hypercomplex numbers introduced by Hamilton [1]. A quaternion $x \in \mathbb{H}$ is described by four real numbers $(r_1, r_\eta, r_{\eta'}, r''_\eta)$ and three imaginary units[1] (η, η', η'')

$$x = r_1 + \eta r_\eta + \eta' r_{\eta'} + \eta'' r_{\eta''}, \tag{1}$$

where the orthogonal imaginary units satisfy

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

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 $x^* = r_1 - \eta r_\eta - \eta' r_{\eta'} - \eta'' r_{\eta''}$, its norm is $|x| = \sqrt{xx^*} = \sqrt{r_1^2 + r_\eta^2 + r_{\eta'}^2 + r_{\eta''}^2}$, and we say that x is a pure unit quaternion if and only if (iff) $x^2 = -1$. The involution of a quaternion x over a pure unit quaternion ν is

$$x^{(\nu)} = -\nu x \nu,$$

and it represents the reflection of x over the plane spanned by $\{1, \nu\}$. Finally, we can introduce the *Cayley-Dickson* representations

$$x = a_1 + \eta'' a_2, \qquad x = b_1 + \eta b_2, \qquad x = c_1 + \eta' c_2, \tag{2}$$

where

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

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

[1] A particular choice of the imaginary axes is the canonical basis $\{i, j, k\}$. However, in this paper we use the more general representation $\{\eta, \eta', \eta''\}$.

3. PROPERNESS OF QUATERNION RANDOM VECTORS

In this section we present the main kinds of quaternion properness (or second-order circularity) [5, 6, 7]. In particular, we show that the formulation based on the complementary covariance matrices, as well as the Cayley-Dickson representation in eq. (2), allows us to obtain additional insight into the structure of proper quaternion random vectors. Some mathematical details, which are omitted here due to the lack of space, can be found in [8].

3.1. Complementary Covariance Matrices

Analogously to the case of complex vectors, the second-order 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 $\mathbf{r}_{\eta''}$ [5, 6, 7]. However, clearer results can be obtained by following a similar derivation to that in [6] for the case of scalar quaternions. Thus, we define the $4n \times 1$ 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

$$\mathbf{T}_n = \frac{1}{2}\begin{bmatrix} +1 & +\eta & +\eta' & +\eta'' \\ +1 & +\eta & -\eta' & -\eta'' \\ +1 & -\eta & +\eta' & -\eta'' \\ +1 & -\eta & -\eta' & +\eta'' \end{bmatrix} \otimes \mathbf{I}_n, \tag{3}$$

is a unitary quaternion operator, i.e., $\mathbf{T}_n^H\mathbf{T}_n = \mathbf{I}_{4n}$. Finally, 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 relationships among these matrices, the real representation in (1), and the Cayley-Dickson representations in (2), can be obtained by means of straightforward but tedious algebra [8], and are omitted here due to the lack of space.

3.2. Quaternion Properness

For quaternion random vectors, there exist two main kinds of properness, which are defined as follows:

Definition 1 ($\mathbb{C}^\eta$-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.*

Definition 2 ($\mathbb{Q}$-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.*

Here, we must point out that the definition of $\mathbb{C}^\eta$-proper vectors only depends on the pure unit quaternion η, whereas the definition of $\mathbb{Q}$-properness is independent of the orthogonal imaginary basis $\{\eta, \eta', \eta''\}$, and therefore it implies $\mathbb{C}^\eta$-properness $\forall\eta$. Furthermore, it can be easily checked that the $\mathbb{C}^\eta$-properness is invariant under *semi-widely linear transformations*, which are linear transformations of the form $\mathbf{u} = \mathbf{F}_1^H\mathbf{x} + \mathbf{F}_\eta^H\mathbf{x}^{(\eta)}$, with $\mathbf{F}_1, \mathbf{F}_\eta \in \mathbb{H}^{n\times r}$. On the other hand, the $\mathbb{Q}$-properness is invariant under quaternion linear transformations, i.e., if $\mathbf{x}$ is a $\mathbb{Q}$-proper vector, then $\mathbf{u} = \mathbf{F}_1^H\mathbf{x}$ is also $\mathbb{Q}$-proper.

Using the Cayley-Dickson representations in (2) it is easy to prove some interesting implications of the two kinds of quaternion properness, which are summarized as follows:

- A vector $\mathbf{x}$ is $\mathbb{C}^\eta$-proper iff $\mathbf{R}_{\mathbf{a}_1,\mathbf{a}_1^*} = \mathbf{R}_{\mathbf{a}_2,\mathbf{a}_2^*} = \mathbf{R}_{\mathbf{a}_1,\mathbf{a}_2^*} = \mathbf{0}_{n\times n}$. In other words, $\mathbf{x}$ is $\mathbb{C}^\eta$-proper iff it can be represented by means of two jointly-proper complex vectors[2] ($\mathbf{a}_1 = \mathbf{r}_1 + \eta\mathbf{r}_\eta$ and $\mathbf{a}_2 = \mathbf{r}_{\eta''} + \eta\mathbf{r}_{\eta'}$) in the plane spanned by $\{1, \eta\}$.
- A quaternion vector $\mathbf{x}$ is $\mathbb{Q}$-proper iff the covariance matrix can be written as

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

  or equivalently, iff the vectors in the real representation $\mathbf{x} = \mathbf{r}_1 + \eta\mathbf{r}_\eta + \eta'\mathbf{r}_{\eta'} + \eta''\mathbf{r}_{\eta''}$ satisfy

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

Finally, the extension of the above definitions for two quaternion random vectors is analogous to that in the case of complex vectors [9]. In particular, $\mathbf{x}$ and $\mathbf{y}$ are jointly $\mathbb{C}^\eta$ (respectively $\mathbb{Q}$) proper iff the composite vector $\left[\mathbf{x}^T, \mathbf{y}^T\right]^T$ is $\mathbb{C}^\eta$ (resp. $\mathbb{Q}$) proper.

4. 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 [4]. 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

$$\bar{\mathbf{u}} = \overline{\mathbf{F}}_{\bar{\mathbf{x}}}^H\bar{\mathbf{x}}, \tag{4}$$

where

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

[2] Two complex vectors $\mathbf{a}_1$ and $\mathbf{a}_2$ are jointly proper iff they are proper and cross proper, or equivalently iff the composite vector $[\mathbf{a}_1^T \mathbf{a}_2^T]$ is proper.

is a general *full-widely linear operator*. Equivalently, we can use the real version of (4)

$$\mathbf{r_u} = \mathbf{F}_{\mathbf{r_x}}^T \mathbf{r_x},$$

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

$$\mathbf{F}_{\mathbf{r_x}} = \mathbf{T}_n^H \overline{\mathbf{F}}_{\bar{\mathbf{x}}} \mathbf{T}_r, \tag{5}$$

with $\mathbf{T}_n$ (and $\mathbf{T}_r$) defined in (3).

4.1. Multivariate Statistical Analysis of Quaternion Vectors

Several popular multivariate statistical analysis techniques amount to maximize the correlation between projections of two random vectors [10]. 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 $\mathbf{y} \in \mathbb{H}^{m\times 1}$,

$$\mathbf{r_u} = \mathbf{F}_{\mathbf{r_x}}^T \mathbf{r}_x, \qquad \mathbf{r_v} = \mathbf{G}_{\mathbf{r_y}}^T \mathbf{r_y},$$

where $\mathbf{F}_{\mathbf{r_x}} \in \mathbb{R}^{4n\times 4r}$, $\mathbf{G}_{\mathbf{r_y}} \in \mathbb{R}^{4m\times 4r}$ are real operators,[3] and $r \leq p = \min(m,n)$. Specifically, our problem can be written as

$$\underset{\mathbf{F}_{\mathbf{r_x}}, \mathbf{G}_{\mathbf{r_y}}}{\arg\max}\ \mathrm{Tr}\left(\mathbf{F}_{\mathbf{r_x}}^T \mathbf{R}_{\mathbf{r_x},\mathbf{r_y}} \mathbf{G}_{\mathbf{r_y}}\right),$$

subject to some constraint to avoid trivial solutions. In fact, the choice of the constraints makes the difference among the following well-known multivariate statistical analysis techniques:

- *Partial least squares (PLS)*: PLS maximizes the correlations subject to the unitarity of the projectors, i.e., the constraints are $\mathbf{F}_{\mathbf{r_x}}^T \mathbf{F}_{\mathbf{r_x}} = \mathbf{G}_{\mathbf{r_y}}^T \mathbf{G}_{\mathbf{r_y}} = \mathbf{I}_{4r}$. In the particular case of $\mathbf{y} = \mathbf{x}$, PLS reduces to the principal component analysis (PCA) technique.
- *Multivariate linear regression (MLR)* [11]: For this method, which is also known as the rank-reduced Wiener filter, half canonical correlation analysis, or orthogonalized PLS, the constraints are $\mathbf{F}_{\mathbf{r_x}}^T \mathbf{R}_{\mathbf{r_x},\mathbf{r_x}} \mathbf{F}_{\mathbf{r_x}} = \mathbf{G}_{\mathbf{r_y}}^T \mathbf{G}_{\mathbf{r_y}} = \mathbf{I}_{4r}$.
- *Canonical correlation analysis (CCA)* [12]: This technique imposes the energy and orthogonality constraints on the projections $\mathbf{r_u}$ and $\mathbf{r_v}$, i.e., the constraints are $\mathbf{F}_{\mathbf{r_x}}^T \mathbf{R}_{\mathbf{r_x},\mathbf{r_x}} \mathbf{F}_{\mathbf{r_x}} = \mathbf{G}_{\mathbf{r_y}}^T \mathbf{R}_{\mathbf{r_y},\mathbf{r_y}} \mathbf{G}_{\mathbf{r_y}} = \mathbf{I}_{4r}$.

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

$$\begin{aligned} &\underset{\mathbf{U}_{\mathbf{r_x}}, \mathbf{V}_{\mathbf{r_y}}}{\arg\max}\quad \mathrm{Tr}\left(\mathbf{U}_{\mathbf{r_x}}^T \mathbf{C}_{\mathbf{r_x},\mathbf{r_y}} \mathbf{V}_{\mathbf{r_y}}\right), \\ &\quad\text{s. t.}\quad \mathbf{U}_{\mathbf{r_x}}^T \mathbf{U}_{\mathbf{r_x}} = \mathbf{V}_{\mathbf{r_y}}^T \mathbf{V}_{\mathbf{r_y}} = \mathbf{I}_{4r}, \end{aligned} \tag{6}$$

where $\mathbf{C}_{\mathbf{r_x},\mathbf{r_y}} = \mathbf{S}_{\mathbf{r_x},\mathbf{r_x}}^{-\frac{1}{2}} \mathbf{R}_{\mathbf{r_x},\mathbf{r_y}} \mathbf{S}_{\mathbf{r_y},\mathbf{r_y}}^{-\frac{1}{2}}$, $\mathbf{U}_{\mathbf{r_x}} = \mathbf{S}_{\mathbf{r_x},\mathbf{r_x}}^{\frac{1}{2}} \mathbf{F}_{\mathbf{r_x}}$, $\mathbf{V}_{\mathbf{r_y}} = \mathbf{S}_{\mathbf{r_y},\mathbf{r_y}}^{\frac{1}{2}} \mathbf{G}_{\mathbf{r_y}}$, and the expressions for $\mathbf{S}_{\mathbf{r_x},\mathbf{r_x}}$ and $\mathbf{S}_{\mathbf{r_y},\mathbf{r_y}}$ in the three studied cases are summarized in Table 1. Obviously, the solutions $\mathbf{U}_{\mathbf{r_x}}$, $\mathbf{V}_{\mathbf{r_y}}$ of (6) are given by the singular vectors associated to the $4r$ largest singular values of the matrix $\mathbf{C}_{\mathbf{r_x},\mathbf{r_y}}$, whose singular value decomposition (SVD) can be written as $\mathbf{C}_{\mathbf{r_x},\mathbf{r_y}} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{V}^T$, 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 particular, we will order the singular values $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_{4p}$ in $\boldsymbol{\Lambda}$ as

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

	$\mathbf{S}_{\mathbf{r_x},\mathbf{r_x}}$	$\mathbf{S}_{\mathbf{r_y},\mathbf{r_y}}$	$\mathbf{S}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}$	$\mathbf{S}_{\bar{\mathbf{y}},\bar{\mathbf{y}}}$
PLS (and PCA)	$\mathbf{I}_{4n}$	$\mathbf{I}_{4m}$	$\mathbf{I}_{4n}$	$\mathbf{I}_{4m}$
MLR	$\mathbf{R}_{\mathbf{r_x},\mathbf{r_x}}$	$\mathbf{I}_{4m}$	$\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}$	$\mathbf{I}_{4m}$
CCA	$\mathbf{R}_{\mathbf{r_x},\mathbf{r_x}}$	$\mathbf{R}_{\mathbf{r_y},\mathbf{r_y}}$	$\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}$	$\mathbf{R}_{\bar{\mathbf{y}},\bar{\mathbf{y}}}$

Table 1. Values of the matrices $\mathbf{S}$ in the three studied cases.

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

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

Finally, using the relationships in (4) and (5), we can obtain the solutions of the above problems in terms of the augmented vectors and matrices. In particular, defining the matrix $\mathbf{C}_{\bar{\mathbf{x}},\bar{\mathbf{y}}} = \mathbf{S}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}^{-\frac{1}{2}} \mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{y}}} \mathbf{S}_{\bar{\mathbf{y}},\bar{\mathbf{y}}}^{-\frac{H}{2}}$ (see Table 1 for the particular values of $\mathbf{S}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}$ and $\mathbf{S}_{\bar{\mathbf{y}},\bar{\mathbf{y}}}$), the optimal projectors are given by

$$\overline{\mathbf{F}}_{\bar{\mathbf{x}}} = \mathbf{S}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}^{-\frac{1}{2}} \overline{\mathbf{U}}_{\bar{\mathbf{x}}}, \qquad \overline{\mathbf{G}}_{\bar{\mathbf{y}}} = \mathbf{S}_{\bar{\mathbf{y}},\bar{\mathbf{y}}}^{-\frac{1}{2}} \overline{\mathbf{V}}_{\bar{\mathbf{y}}},$$

where the unitary widely-linear operators $\overline{\mathbf{U}}_{\bar{\mathbf{x}}} = \mathbf{T}_n \mathbf{U}_{\mathbf{r_x}} \mathbf{T}_r^H \in \mathbb{H}^{4n\times r}$ and $\overline{\mathbf{V}}_{\bar{\mathbf{y}}} = \mathbf{T}_m \mathbf{V}_{\mathbf{r_y}} \mathbf{T}_r^H \in \mathbb{H}^{4m\times r}$ can be directly obtained from the decomposition

$$\mathbf{C}_{\bar{\mathbf{x}},\bar{\mathbf{y}}} = \underbrace{\left(\mathbf{T}_n \mathbf{U} \mathbf{T}_p^H\right)}_{\overline{\mathbf{U}}} \underbrace{\left(\mathbf{T}_p \boldsymbol{\Lambda} \mathbf{T}_p^H\right)}_{\overline{\boldsymbol{\Lambda}}} \underbrace{\left(\mathbf{T}_n \mathbf{V} \mathbf{T}_p^H\right)^H}_{\overline{\mathbf{V}}^H}, \tag{7}$$

which can be seen as an extension of the singular value decomposition used in [9] 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

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

with

$$\begin{bmatrix} \boldsymbol{\Sigma}_1 \\ \boldsymbol{\Sigma}_2 \\ \boldsymbol{\Sigma}_3 \\ \boldsymbol{\Sigma}_4 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} +\mathbf{I}_p & +\mathbf{I}_p & +\mathbf{I}_p & +\mathbf{I}_p \\ +\mathbf{I}_p & +\mathbf{I}_p & -\mathbf{I}_p & -\mathbf{I}_p \\ +\mathbf{I}_p & -\mathbf{I}_p & +\mathbf{I}_p & -\mathbf{I}_p \\ +\mathbf{I}_p & -\mathbf{I}_p & -\mathbf{I}_p & +\mathbf{I}_p \end{bmatrix} \begin{bmatrix} \boldsymbol{\Lambda}_1 \\ \boldsymbol{\Lambda}_2 \\ \boldsymbol{\Lambda}_3 \\ \boldsymbol{\Lambda}_4 \end{bmatrix}.$$

4.2. Implications of $\mathbb{C}^\eta$ and $\mathbb{Q}$ Properness

In this subsection we point out the main implications of the two kinds of quaternion properness on the previous multivariate statistical analysis techniques. Due to the lack of space, we skip some mathematical details (which can be found in [8]) and present the main results as four 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*

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

Theorem 2 *Given two jointly $\mathbb{C}^\eta$-proper vectors $\mathbf{x}$ and $\mathbf{y}$, the singular values $\lambda_1, \ldots, \lambda_{4p}$ of $\mathbf{C}_{\mathbf{r_x},\mathbf{r_y}}$ (and $\mathbf{C}_{\bar{\mathbf{x}},\bar{\mathbf{y}}}$) have multiplicity two.*

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.,*

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

Theorem 4 *Given two jointly $\mathbb{Q}$-proper vectors $\mathbf{x}$ and $\mathbf{y}$, the singular values $\lambda_1, \ldots, \lambda_{4p}$ of $\mathbf{C}_{\mathbf{r_x},\mathbf{r_y}}$ (and $\mathbf{C}_{\bar{\mathbf{x}},\bar{\mathbf{y}}}$) have multiplicity four.*

The proofs of the theorems are based on the block diagonal structure of the matrices ($\mathbf{S}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}$, $\mathbf{S}_{\bar{\mathbf{y}},\bar{\mathbf{y}}}$ and $\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{y}}}$) involved in the obtention of $\mathbf{C}_{\bar{\mathbf{x}},\bar{\mathbf{y}}}$. This diagonal structure (two diagonal blocks in the $\mathbb{C}^\eta$-proper case and four diagonal blocks in the $\mathbb{Q}$-proper case) is propagated to the matrices in the decomposition in eq. (7), from which the theorems can be easily proved.

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. On the other hand, Theorem 2 ensures that the augmented covariance matrices of $\mathbb{C}^\eta$-proper vectors have eigenvalues with multiplicity two. Finally, Theorems 3 and 4 can be seen as the counterpart of Theorems 1 and 2 for jointly $\mathbb{Q}$-proper vectors. In particular, Theorem 3 ensures that we can not expect any gain from full or semi-widely linear processing $\mathbb{Q}$-proper vectors.

5. NUMERICAL EXAMPLE AND CONCLUSIONS

The previous ideas are illustrated here by means of a simulation example. In particular, we consider a $\mathbb{Q}$-proper Gaussian vector $\mathbf{y}$ of dimension $m = 2$ with zero mean and covariance $\mathbf{R}_{\mathbf{y},\mathbf{y}} = \mathbf{I}_2$. From $\mathbf{y}$, we form the observation vector $\mathbf{x} \in \mathbb{H}^{4\times 1}$ as

$$\mathbf{x} = [\mathbf{w}^T, \mathbf{z}^T]^T + \mathbf{n}^T,$$

where $\mathbf{w} = \mathbf{y} + 0.8\mathbf{y}^{(\eta)} + \mathbf{0}.6\mathbf{y}^{(\eta')} + 0.5\mathbf{y}^{(\eta'')}$, $\mathbf{z} \in \mathbb{H}^{2\times 1}$ is a $\mathbb{Q}$-proper Gaussian quaternion vector with zero mean and covariance $\mathbf{R}_{\mathbf{z},\mathbf{z}} = \mathbf{I}_2$, and $\mathbf{n} \in \mathbb{H}^{4\times 1}$ is a $\mathbb{Q}$-proper Gaussian quaternion vector with zero mean and covariance $\mathbf{R}_{\mathbf{n},\mathbf{n}} = 10^{-2}\mathbf{I}_4$.

With these definitions, it is clear that $\mathbf{x}$ is improper, and therefore the full-widely linear processing should outperform the semi-widely or conventional linear processing. Specifically, we apply the MLR technique to estimate $\mathbf{y}$ from $\mathbf{x}$, which results in a theoretical (i.e., with perfect knowledge of the second-order statistics) mean square error (MSE) of 0.31 if we apply full-widely linear processing, 0.61 for semi-widely linear processing, and 1.12 for quaternion linear processing.

Additionally, the averaged results of 1000 Monte-Carlo simulations with estimated covariance matrices from a finite set of observations are shown in Fig. 1. As can be seen, for a sufficiently large number of observations N, the obtained results coincide with the theoretical values. However, the figure also suggests that in some situations such as small sample sizes, it can be advisable to use the simpler quaternion linear (or semi-widely linear) model instead of the theoretically optimal full-widely linear processing. In other words, although the data can not be correctly represented by quaternion vectors, the simplicity of quaternion linear processing still provides some practical advantages. Future research lines include the estimation of the optimal pure unit quaternion η for semi-widely linear processing, as well as the theoretical analysis of the gain provided by full or semi-widely linear processing.

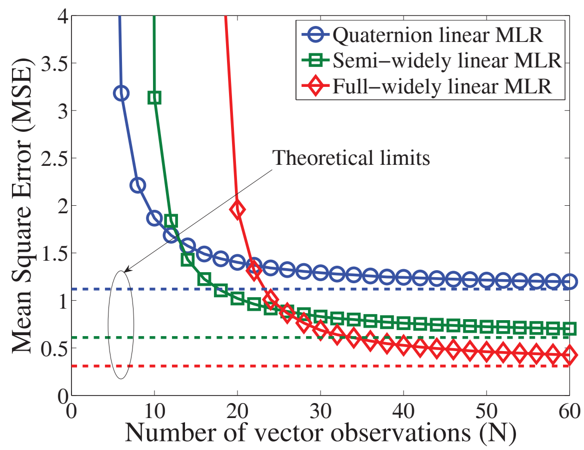

Fig. 1. MSE in the reconstruction of $\mathbf{y}$ from $\mathbf{x}$. Theoretical values and evolution with the number of vector observations.

6. REFERENCES

[1] W. R. Hamilton, "On quaternions," *Proceedings of the Royal Irish Academy*, November 11 1844.

[2] S. Miron, N. Le Bihan, and J.I. Mars, "Quaternion-MUSIC for vector-sensor array processing," *IEEE Transactions on Signal Processing*, vol. 54, no. 4, pp. 1218–1229, Apr. 2006.

[3] J. Seberry, K. Finlayson, S.S. Adams, T.A. Wysocki, Tianbing Xia, and B.J. Wysocki, "The theory of quaternion orthogonal designs," *IEEE Transactions on Signal Processing*, vol. 56, no. 1, pp. 256–265, Jan. 2008.

[4] C.C. Took and D.P. Mandic, "The quaternion LMS algorithm for adaptive filtering of hypercomplex processes," *IEEE Trans. on Sig. Proc.*, vol. 57, no. 4, pp. 1316–1327, Apr. 2009.

[5] N. N. Vakhania, "Random vectors with values in quaternion hilbert spaces," *Theory of Probability and its Applications*, vol. 43, no. 1, pp. 99–115, 1999.

[6] P. Amblard and N. Le Bihan, "On properness of quaternion valued random variables," in *IMA Conference on Mathematics in Signal Processing*, Cirencester (UK), 2004, pp. 23–26.

[7] N. L. Bihan and S. Buchholz, "Optimal separation of polarized signals by quaternionic neural networks," in *XIV European Signal Processing Conference*, Florence, Italy, 2006.

[8] J. Vía, D. Ramírez, and I. Santamaría, "Properness and widely linear processing of quaternion random vectors," submmited to *IEEE Transactions on Information Theory*, 2009.

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

[10] P.J. Schreier, "A unifying discussion of correlation analysis for complex random vectors," *IEEE Transactions on Signal Processing*, vol. 56, no. 4, pp. 1327–1336, April 2008.

[11] L. Scharf, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*, Addison-Wesley, New York, 1990.

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