Generalized Likelihood Ratios for Testing the Properness of Quaternion Gaussian Vectors

Javier Vía, *Member, IEEE*, Daniel P. Palomar, *Senior Member, IEEE*, and Luis Vielva

***Abstract*—In a recent paper, the second-order statistical analysis of quaternion random vectors has shown that there exist two different kinds of quaternion *widely linear* processing, which are associated with the two main types of quaternion *properness*. In this paper, we consider the problem of determining, from a finite number of independent vector observations, whether a quaternion Gaussian vector is proper or not. Specifically, we derive three generalized likelihood ratio tests (GLRTs) for testing the two main kinds of quaternion properness and show that the GLRTs reduce to the estimation of three previously proposed quaternion improperness measures. Interestingly, the three GLRT statistics (improperness measures) can be interpreted as an estimate of the entropy loss due to the quaternion improperness. Additionally, we analyze the case in which the orthogonal basis for the representation of the quaternion vector is unknown, which results in the problem of estimating the principal $\mathbb{C}$-properness direction, i.e., the pure unit quaternion minimizing the $\mathbb{C}$-improperness measure. Although this estimation problem is not convex, we propose a technique based on successive convex approximations, which can be solved in closed form. Finally, some simulation examples illustrate the performance and practical application of the proposed tests.**

***Index Terms*—Generalized likelihood ratio test (GLRT), principal $\mathbb{C}$-properness direction, properness, propriety, quaternions, second-order circularity.**

I. Introduction

APART from its traditional use in aerospace [1], [2] and computer graphics [3] problems, quaternion signal processing has recently encountered interesting applications in image processing [4]–[8], wind modeling [9]–[12], processing of polarized waves [13], [14], and design (and processing) of space–time (and space–time-polarization [15]) block codes [16]–[20]. However, the statistical analysis of quaternion random vectors has received limited attention [9], [14], [21]–[24], and only recently the concept of *widely linear* processing has been extended from complex to quaternion vectors [25]. A complex random vector is said to be proper if it is uncorrelated with its complex conjugate, which results in the optimality of the *conventional* linear processing. However, in the more general case of (possibly) improper complex vectors, the optimal linear processing is widely linear, i.e., we have to simultaneously operate on the data vector and its complex conjugate [26]–[34].

Unlike the complex case, we can define two different types of quaternion *widely linear* processing [25], which are strongly related to the two main kinds of quaternion properness. In particular, the most general quaternion linear processing (which we refer to as *full-widely* linear processing) requires the operation on the quaternion vector $\mathbf{x}$ and its involutions over the three pure unit quaternions in an orthogonal basis $\{\eta, \eta', \eta''\}$. However, for $\mathbb{Q}$-proper data, the optimal linear processing reduces to *conventional* linear processing (we do not need to operate on the vector involutions), whereas for $\mathbb{C}^\eta$-proper vectors, the optimal linear processing (referred to as *semi-widely* linear) only requires the operation on the quaternion vector and its involution over η [25]. In other words, taking into account the isomorphism among quaternion, complex, and real numbers, we can consider three different scenarios. 1) If our quaternion vector $\mathbf{x}$ is $\mathbb{Q}$-proper, we can apply *conventional* quaternion linear processing. 2) If it is $\mathbb{C}^\eta$-proper, it has to be decomposed into two complex vectors, which will be jointly processed. 3) If $\mathbf{x}$ is improper (i.e., if it is not $\mathbb{Q}$- or $\mathbb{C}$-proper), we need to directly operate on the four real vectors composing $\mathbf{x}$. Finally, we must note that in [25], the authors introduced the definition of $\mathbb{R}^\eta$-properness, which allows us to easily relate the two main kinds of quaternion properness. Roughly speaking, we can say that the $\mathbb{R}^\eta$-properness is all what a $\mathbb{C}^\eta$-proper quaternion vector needs to become $\mathbb{Q}$-proper.

Analogously to the complex case [35], algorithms adapted for improper signals can fail or suffer from slow convergence when they are used for proper signals. This is due to the fact that the number of free parameters in a conventional linear algorithm is multiplied by four (respectively, by two) in its full-widely (respectively, semi-widely) linear counterpart. Therefore, since the complexity of the associated parameter estimation problem depends on the selected model (conventional, semi-widely, or full-widely), we should follow the principle of parsimony and choose the simplest model exploiting the statistical properties of the data. As a consequence, it becomes crucial to determine whether a quaternion random vector is $\mathbb{Q}$-proper, $\mathbb{C}^\eta$-proper, or improper. As a practical example, which is illustrated in Section VI, we consider an optical communication system based on dual polarization [36]–[38]. Thus, depending on several system parameters, the signals in the fiber can be represented by $\mathbb{Q}$-proper, $\mathbb{C}$-proper, or improper random quaternions.

Manuscript received May 31, 2010; revised October 02, 2010 and December 09, 2010; accepted December 11, 2010. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Jean-Yves Tourneret. This work was supported by the Spanish Government, Ministerio de Ciencia e Innovación (MICINN), under projects COMONSENS (CSD2008-00010, CONSOLIDER-INGENIO 2010 Program) and COSIMA (TEC2010-19545-C04-03) and by the Hong Kong RGC 618709 research grant.

J. Vía and L. Vielva are with the Department of Communications Engineering, University of Cantabria, 39005 Santander, Spain (e-mail: jvia@gtas.dicom.unican.es; luis@gtas.dicom.unican.es).

D. P. Palomar is with the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Kowloon, Hong Kong (e-mail: palomar@ust.hk).

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/TSP.2010.2101067

In this paper, we consider the three binary hypotheses testing problems associated with the evaluation of two out of the three possible scenarios. In particular, assuming zero-mean quaternion Gaussian data, we derive three generalized likelihood ratio tests (GLRTs), which are also the key ingredient for solving the more general multiple-hypotheses testing problem. Although suboptimal in the Neyman–Pearson sense, this kind of detector is easy to obtain, performs well in practice, and in our case results in simple detection rules, which admit straightforward interpretations. Specifically, we show that the proposed GLRTs reduce to the estimation of three improperness measures, which can be interpreted as the entropy loss due to the different kinds of quaternion improperness.

Two previous works related to our detection problems are [23] and [39]. Specifically, [39] studied the problem of testing whether "a covariance matrix with complex structure has quaternion structure," which can be shown to be equivalent to one of the three GLRTs derived in this paper. However, although the formulation in [39] can be useful for obtaining the moments of the test statistic, our derivation based on the quaternion complementary covariance matrices makes the derivation and interpretation of the GLRTs easier, as well as their generalization to arbitrary orthogonal bases. Finally, in [23] the authors considered the detection of $\mathbb{C}$-proper quaternion random variables in white $\mathbb{Q}$-proper noise.

Additionally, we consider the case in which the principal $\mathbb{C}$-properness direction η is unknown, and therefore it has to be estimated as well. In this case, the estimation problem reduces to the minimization of the $\mathbb{C}^\eta$-improperness measure or, equivalently, the maximization of the $\mathbb{R}^\eta$-improperness. Although the maximum-likelihood (ML) estimation of η results in a nonconvex optimization problem in the general vector case, we propose an algorithm based on successive convex approximations [40]–[42] of the nonconvex cost function, which guarantees the convergence to a solution satisfying the Karush–Kuhn–Tucker (KKT) conditions [43]. Finally, after a brief discussion on the general classification problem and the distribution of the test statistics, several simulation examples illustrate the accuracy and fast convergence of the proposed algorithm, as well as the performance and practical application of the three derived GLRTs.

A. Notation

In this paper, we use boldfaced uppercase letters to denote matrices, boldfaced lowercase letters for column vectors, and lightfaced lowercase letters for scalar quantities. Superscripts $(\cdot)^*$, $(\cdot)^{\mathrm{T}}$ and $(\cdot)^{\mathrm{H}}$ denote quaternion (or complex) conjugate, transpose, and Hermitian (i.e., transpose and quaternion conjugate), respectively. The notation $\mathbf{A} \in \mathbb{F}^{n\times m}$ denotes that $\mathbf{A}$ is an $n \times m$ matrix with entries in $\mathbb{F}$, where $\mathbb{F}$ can be $\mathbb{R}$, the field of real numbers, $\mathbb{C}$, the field of complex numbers, or $\mathbb{H}$, the skew-field of quaternion numbers. $\Re(\mathbf{A})$, $\mathrm{Tr}(\mathbf{A})$, and $|\mathbf{A}|$ denote the real part, trace, and determinant of matrix $\mathbf{A}$. $\mathbf{A}^{1/2}$ (respectively, $\mathbf{A}^{-1/2}$) is the Hermitian square root of the Hermitian matrix $\mathbf{A}$ (resp. $\mathbf{A}^{-1}$). The diagonal matrix with vector $\mathbf{a}$ along its diagonal is denoted by $\mathrm{diag}(\mathbf{a})$, and $\mathrm{vec}(\mathbf{A})$ is the column-wise vectorized version of matrix $\mathbf{A}$. $\mathbf{I}_n$ is the identity matrix of dimension n, and $\mathbf{0}_{n\times m}$ is the $n \times m$ zero matrix. Finally, the Kronecker product is denoted by $\otimes$, 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., $\mathbf{R}_{\mathbf{a},\mathbf{b}} = E\mathbf{a}\mathbf{b}^{\mathrm{H}}$.

II. Preliminaries

A. Quaternion Algebra

Quaternions are four-dimensional hypercomplex numbers invented by Hamilton [44]. A quaternion $x \in \mathbb{H}$ is defined as

$$x = r_1 + ir_i + jr_j + kr_k \tag{1}$$

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

$$i^2 = j^2 = k^2 = ijk = -1 \tag{2}$$

which also implies

$$ij = k = -ji \quad jk = i = -kj \quad ki = j = -ik. \tag{3}$$

Quaternions form a skew field $\mathbb{H}$ [45], which means that they satisfy the axioms of a field except for the commutative law of the product, i.e., for $x, y \in \mathbb{H}$, $xy \neq yx$ in general, although we must note that $\Re(xy) = \Re(yx)$. The conjugate of a quaternion x is defined as $x^* = r_1 - ir_i - jr_j - kr_k$, and the conjugate of the product satisfies $(xy)^* = y^*x^*$. The inner product between two quaternions $x, y \in \mathbb{H}$ is defined[1] as xy^*, and two quaternions are orthogonal if and only if (iff) their scalar product (the real part of the inner product) is zero. The absolute value of a quaternion 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 $\eta \in \mathbb{H}$ is a pure unit quaternion iff $\eta^2 = -1$ (i.e., iff $|\eta| = 1$ and its real part is zero). A particularly important operation is the quaternion involution.

Definition 1 (Quaternion Involution): The involution of a quaternion x over a pure unit quaternion η is

$$x^{(\eta)} = \eta x \eta^{-1} = \eta x \eta^* = -\eta x \eta \tag{4}$$

and it represents a rotation of angle π in the imaginary plane orthogonal to $\{1, \eta\}$ [45].

Some basic properties of the quaternion involution, which can be easily checked, are [14], [24]

$$x^{(\eta)^*} = (x^*)^{(\eta)} \qquad \forall x \in \mathbb{H} \tag{5}$$

$$\left(x^{(\eta)}\right)^{(\eta)} = x \qquad \forall x \in \mathbb{H} \tag{6}$$

$$(xy)^{(\eta)} = x^{(\eta)}y^{(\eta)} \qquad \forall x, y \in \mathbb{H} \tag{7}$$

$$x\eta = \eta x^{(\eta)} \qquad \forall x \in \mathbb{H}. \tag{8}$$

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

$$\begin{bmatrix} 1 \\ \eta \\ \eta' \\ \eta'' \end{bmatrix} = \begin{bmatrix} 1 & \mathbf{0}_{1\times 3} \\ \mathbf{0}_{3\times 1} & \mathbf{Q} \end{bmatrix} \begin{bmatrix} 1 \\ i \\ j \\ k \end{bmatrix} \tag{9}$$

[1]Other definitions of the quaternion inner product are possible; see for instance [45].

where $\mathbf{Q} \in \mathbb{R}^{3\times 3}$ is a rotation matrix, (i.e., $\mathbf{Q}^{\mathrm{T}}\mathbf{Q} = \mathbf{I}_3$ and $|\mathbf{Q}| = 1$), which implies

$$\eta^2 = {\eta'}^2 = {\eta''}^2 = \eta\eta'\eta'' = -1. \tag{10}$$

Thus, any quaternion can be represented as

$$x = r_1 + \eta r_\eta + \eta' r_{\eta'} + \eta'' r_{\eta''} \tag{11}$$

where $[r_\eta, r_{\eta'}, r_{\eta''}] = [r_i, r_j, r_k]\mathbf{Q}^{\mathrm{T}}$.

Finally, a quaternion x can also be represented by means of the *Cayley–Dickson* construction

$$x = a_1 + \eta'' a_2 \quad x = b_1 + \eta b_2 \quad x = c_1 + \eta' c_2 \tag{12}$$

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 $\{1, \eta\}$, $\{1, \eta'\}$, or $\{1, \eta''\}$.

B. Second-Order Statistics of Quaternion Random Vectors

The statistical analysis of a quaternion random vector $\mathbf{x} \in \mathbb{H}^{n\times 1}$ can be based on its real representation $\mathbf{r}_{\mathbf{x}} = [\mathbf{r}_1^{\mathrm{T}}, \mathbf{r}_\eta^{\mathrm{T}}, \mathbf{r}_{\eta'}^{\mathrm{T}}, \mathbf{r}_{\eta''}^{\mathrm{T}}]^{\mathrm{T}}$, which will allow us to obtain the moments of the GLRT statistics [39]. However, working with the augmented quaternion vector $\bar{\mathbf{x}} = [\mathbf{x}^{\mathrm{T}}, {\mathbf{x}^{(\eta)}}^{\mathrm{T}}, {\mathbf{x}^{(\eta')}}^{\mathrm{T}}, {\mathbf{x}^{(\eta'')}}^{\mathrm{T}}]^{\mathrm{T}}$ will make easier the definition of quaternion properness, the derivation of the tests, and the generalization of the results to an arbitrary orthogonal basis $\{\nu, \nu', \nu''\}$. Thus, the second-order statistical information of the quaternion vector is given by the augmented covariance matrix

$$\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}} = \begin{bmatrix} \mathbf{R}_{\mathbf{x},\mathbf{x}} & \mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta)}} & \mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta')}} & \mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta'')}} \\ \mathbf{R}^{(\eta)}_{\mathbf{x},\mathbf{x}^{(\eta)}} & \mathbf{R}^{(\eta)}_{\mathbf{x},\mathbf{x}} & \mathbf{R}^{(\eta)}_{\mathbf{x},\mathbf{x}^{(\eta'')}} & \mathbf{R}^{(\eta)}_{\mathbf{x},\mathbf{x}^{(\eta')}} \\ \mathbf{R}^{(\eta')}_{\mathbf{x},\mathbf{x}^{(\eta')}} & \mathbf{R}^{(\eta')}_{\mathbf{x},\mathbf{x}^{(\eta'')}} & \mathbf{R}^{(\eta')}_{\mathbf{x},\mathbf{x}} & \mathbf{R}^{(\eta')}_{\mathbf{x},\mathbf{x}^{(\eta)}} \\ \mathbf{R}^{(\eta'')}_{\mathbf{x},\mathbf{x}^{(\eta'')}} & \mathbf{R}^{(\eta'')}_{\mathbf{x},\mathbf{x}^{(\eta')}} & \mathbf{R}^{(\eta'')}_{\mathbf{x},\mathbf{x}^{(\eta)}} & \mathbf{R}^{(\eta'')}_{\mathbf{x},\mathbf{x}} \end{bmatrix} \tag{13}$$

which contains the covariance matrix $\mathbf{R}_{\mathbf{x},\mathbf{x}} = E\mathbf{x}\mathbf{x}^{\mathrm{H}}$ and three complementary covariance matrices $\mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta)}} = E\mathbf{x}{\mathbf{x}^{(\eta)}}^{\mathrm{H}}$, $\mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta')}} = E\mathbf{x}{\mathbf{x}^{(\eta')}}^{\mathrm{H}}$, and $\mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta'')}} = E\mathbf{x}{\mathbf{x}^{(\eta'')}}^{\mathrm{H}}$.

Finally, we must point out that, given two different orthogonal bases $\{\eta, \eta', \eta''\}$ and $\{\nu, \nu', \nu''\}$ related by means of a rotation matrix $\mathbf{Q} \in \mathbb{R}^{3\times 3}$ as

$$\begin{bmatrix} \nu \\ \nu' \\ \nu'' \end{bmatrix} = \mathbf{Q} \begin{bmatrix} \eta \\ \eta' \\ \eta'' \end{bmatrix} \tag{14}$$

we can easily relate the augmented quaternion vectors and covariance matrices as stated in the following lemmas.

Lemma 1: Given a quaternion random vector $\mathbf{x} \in \mathbb{H}^{n\times 1}$ and two different orthogonal bases $\{\eta, \eta', \eta''\}$ and $\{\nu, \nu', \nu''\}$, the corresponding augmented quaternion vectors are related as

$$\begin{bmatrix} \mathbf{x} \\ \mathbf{x}^{(\nu)} \\ \mathbf{x}^{(\nu')} \\ \mathbf{x}^{(\nu'')} \end{bmatrix} = \mathbf{\Gamma} \begin{bmatrix} \mathbf{x} \\ \mathbf{x}^{(\eta)} \\ \mathbf{x}^{(\eta')} \\ \mathbf{x}^{(\eta'')} \end{bmatrix} \tag{15}$$

where $\mathbf{\Gamma}$ is a unitary quaternion operator given by

$$\mathbf{\Gamma} = \begin{bmatrix} 1 & \mathbf{0}_{1\times 3} \\ \mathbf{0}_{3\times 1} & \mathbf{\Lambda}_\nu \mathbf{Q} \mathbf{\Lambda}_\eta^{\mathrm{H}} \end{bmatrix} \otimes \mathbf{I}_n \tag{16}$$

$\mathbf{\Lambda}_\nu = \mathrm{diag}([\nu, \nu', \nu'']^{\mathrm{T}})$, and $\mathbf{\Lambda}_\eta = \mathrm{diag}([\eta, \eta', \eta'']^{\mathrm{T}})$.

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

$$\begin{aligned} \mathbf{x}^{(\nu)} &= \nu\mathbf{x}\nu^* = \nu\mathbf{x}(\nu_\eta\eta^* + \nu_{\eta'}\eta'^* + \nu_{\eta''}\eta''^*) \\ &= \nu_\eta\nu\eta^*\mathbf{x}^{(\eta)} + \nu_{\eta'}\nu\eta'^*\mathbf{x}^{(\eta')} + \nu_{\eta''}\nu\eta''^*\mathbf{x}^{(\eta'')}. \end{aligned} \tag{17}$$

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

Lemma 2: The augmented covariance matrices in two different orthogonal bases are related as

$$\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}(\{1, \nu, \nu', \nu''\}) = \mathbf{\Gamma}\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}(\{1, \eta, \eta', \eta''\})\mathbf{\Gamma}^{\mathrm{H}} \tag{18}$$

where the expressions in parentheses make the bases explicit.

Proof: This is a direct consequence of Lemma 1 and the definition of the augmented covariance matrix. ■

C. Properness of Quaternion Vectors

In the complex case, a vector $\mathbf{x} \in \mathbb{C}^{n\times 1}$ is proper iff the complementary covariance matrix $\mathbf{R}_{\mathbf{x},\mathbf{x}^*} = E\mathbf{x}\mathbf{x}^{\mathrm{T}}$ is zero [26], [29], [30], [46]. The quaternion case is a bit more complicated, and we can define different kinds of properness [21], [22], [25], which also have different implications on the optimal linear processing of a quaternion random vector [25]. In this paper, we focus on the two main kinds of quaternion properness.

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.

Definition 3 ($\mathbb{C}^\eta$-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.

These properness definitions satisfy some interesting properties, which include the invariance to linear quaternion transformations and the invariance of the second-order statistics (SOS) to different types of right-Clifford translations [25], [47]—i.e., right products $\mathbf{x}c$, with c a unit quaternion. Here, we summarize the two main properties of the $\mathbb{Q}$- and $\mathbb{C}^\eta$-properness definitions [25].

Property 1 ($\mathbb{Q}$-Proper Vectors): A quaternion random vector $\mathbf{x}$ is $\mathbb{Q}$-proper iff it is $\mathbb{C}^\eta$-proper for all pure unit quaternions η.

Property 2 ($\mathbb{C}^\eta$-Proper Vectors): A quaternion random vector $\mathbf{x}$ is $\mathbb{C}^\eta$-proper iff the vectors $\mathbf{a}_1$, $\mathbf{a}_2$ in its Cayley–Dickson representation $\mathbf{x} = \mathbf{a}_1 + \eta''\mathbf{a}_2$ are jointly complex-proper, i.e., iff the composite vector $\mathbf{a} = [\mathbf{a}_1^{\mathrm{T}}, \mathbf{a}_2^{\mathrm{T}}]^{\mathrm{T}}$ is proper.

From a practical point of view, the two previous definitions have a strong impact on the structure of the optimal linear processing. In general, the optimal linear processing of a quaternion random vector $\mathbf{x} \in \mathbb{H}^{n\times 1}$ takes the form

$$\mathbf{u} = \mathbf{F}_1^{\mathrm{H}}\mathbf{x} + \mathbf{F}_\eta^{\mathrm{H}}\mathbf{x}^{(\eta)} + \mathbf{F}_{\eta'}^{\mathrm{H}}\mathbf{x}^{(\eta')} + \mathbf{F}_{\eta''}^{\mathrm{H}}\mathbf{x}^{(\eta'')} \tag{19}$$

with $\mathbf{u} \in \mathbb{H}^{p\times 1}$ and $\mathbf{F}_1, \mathbf{F}_\eta, \mathbf{F}_{\eta'}, \mathbf{F}_{\eta''} \in \mathbb{H}^{n\times p}$. That is, we have to simultaneously operate on the quaternion vector and its involutions, which is referred to as *full-widely* linear processing. However, in the case of $\mathbb{Q}$- and $\mathbb{C}^\eta$-proper vectors, the optimal linear processing simplifies as follows [25].

Property 3 (Linear Processing of $\mathbb{Q}$-Proper Vectors): The optimal linear processing of a $\mathbb{Q}$-proper vector $\mathbf{x}$ takes the form

$$\mathbf{u} = \mathbf{F}_1^{\mathrm{H}}\mathbf{x} \tag{20}$$

and it is referred to as *conventional* linear processing.

Property 4 (Linear Processing of $\mathbb{C}^\eta$-Proper Vectors): The optimal linear processing of a $\mathbb{C}^\eta$-proper vector $\mathbf{x}$ takes the form

$$\mathbf{u} = \mathbf{F}_1^{\mathrm{H}}\mathbf{x} + \mathbf{F}_\eta^{\mathrm{H}}\mathbf{x}^{(\eta)} \tag{21}$$

and it is referred to as *semi-widely* linear processing.

Finally, in [25] the authors have introduced a third kind of quaternion properness based on the cancellation of only one complementary covariance matrix.

Definition 4 ($\mathbb{R}^\eta$-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.

Unfortunately, this third kind of quaternion properness does not result in a simplification of the optimal linear processing. However, the $\mathbb{R}^\eta$-properness definition clearly relates the two previous kinds of quaternion properness. That is, we can say that the $\mathbb{C}^\eta$- and $\mathbb{R}^\eta$-properness are complementary and, together, result in $\mathbb{Q}$-properness. This relationship will become useful in the derivation of the GLRTs in Section III.

III. Generalized Likelihood Ratio Tests

As we have seen, the $\mathbb{Q}$- and $\mathbb{C}^\eta$-properness have a strong impact on the structure of the optimal linear processing of quaternion random vectors. Therefore, it is crucial to determine whether our quaternion data $\mathbf{x} \in \mathbb{H}^{n\times 1}$ are $\mathbb{Q}$-proper, $\mathbb{C}^\eta$-proper, or improper. Clearly, in its general formulation, this is a multiple-hypotheses testing problem. However, here we focus on the three binary hypotheses testing problems obtained by considering two out of the three different hypotheses, which is justified by the two following facts.

- The binary hypotheses testing problems can arise in practical situations when the problem structure yields some *a priori* information about the properness of the data. For instance, if we consider the problem of detecting the presence of a zero-mean improper Gaussian signal (with unknown augmented covariance matrix) in zero-mean $\mathbb{Q}$-proper Gaussian noise (with unknown covariance matrix), the optimal detector amounts to determining whether the observations are $\mathbb{Q}$-proper or not. That is, in this situation, the problem structure allows us to discard the hypothesis of $\mathbb{C}^\eta$-proper observations (which is still implicit in the improper hypothesis).
- The binary hypotheses testing problems result in simple detection rules, which provide a clear insight about the structure of the overall testing problem. Moreover, we will show that the three binary tests can be seen as the core of a practical multiple-hypotheses test based on the approximation of the *a posteriori* probabilities of each hypothesis.

In this section, we propose three GLRTs for solving the associated binary hypotheses tests. Although suboptimal in the Neyman–Pearson sense, the GLRT provides satisfactory results in practical situations [32], [33], [48]. Furthermore, the derivation of the GLRT is usually simpler than other alternative detectors and, in our particular problem, it permits a straightforward and intuitive interpretation of the detection rules.

A. ML Estimates of the Augmented Covariance Matrix

Let us start by writing the probability density function (pdf) of a quaternion Gaussian vector with zero mean and nonsingular augmented covariance matrix $\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}$ as [25]

$$p(\mathbf{x}; \mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}) = \frac{1}{(\pi/2)^{2n}|\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}|^{1/2}} \exp\left(-\frac{1}{2}\bar{\mathbf{x}}^{\mathrm{H}}\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}^{-1}\bar{\mathbf{x}}\right). \tag{22}$$

Thus, given T independent realizations $\mathbf{x}[t]$ $(t = 0, \ldots, T-1)$ of a quaternion Gaussian vector $\mathbf{x}$, we can take the logarithm of the pdf to obtain the log-likelihood function, which (up to a scaling factor T and constant terms) is given by

$$\mathcal{L}(\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}) = -\frac{1}{2}\ln|\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}| - \frac{1}{2}\Re\left[\mathrm{Tr}\left(\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}^{-1}\hat{\mathbf{R}}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}\right)\right] \tag{23}$$

where

$$\hat{\mathbf{R}}_{\bar{\mathbf{x}},\bar{\mathbf{x}}} = \frac{1}{T}\sum_{t=0}^{T-1}\bar{\mathbf{x}}[t]\bar{\mathbf{x}}^{\mathrm{H}}[t] = \begin{bmatrix} \hat{\mathbf{R}}_{\mathbf{x},\mathbf{x}} & \hat{\mathbf{R}}_{\mathbf{x},\mathbf{x}^{(\eta)}} & \hat{\mathbf{R}}_{\mathbf{x},\mathbf{x}^{(\eta')}} & \hat{\mathbf{R}}_{\mathbf{x},\mathbf{x}^{(\eta'')}} \\ \hat{\mathbf{R}}_{\mathbf{x},\mathbf{x}^{(\eta)}}^{(\eta)} & \hat{\mathbf{R}}_{\mathbf{x},\mathbf{x}}^{(\eta)} & \hat{\mathbf{R}}_{\mathbf{x},\mathbf{x}^{(\eta'')}}^{(\eta)} & \hat{\mathbf{R}}_{\mathbf{x},\mathbf{x}^{(\eta')}}^{(\eta)} \\ \hat{\mathbf{R}}_{\mathbf{x},\mathbf{x}^{(\eta')}}^{(\eta')} & \hat{\mathbf{R}}_{\mathbf{x},\mathbf{x}^{(\eta'')}}^{(\eta')} & \hat{\mathbf{R}}_{\mathbf{x},\mathbf{x}}^{(\eta')} & \hat{\mathbf{R}}_{\mathbf{x},\mathbf{x}^{(\eta)}}^{(\eta')} \\ \hat{\mathbf{R}}_{\mathbf{x},\mathbf{x}^{(\eta'')}}^{(\eta'')} & \hat{\mathbf{R}}_{\mathbf{x},\mathbf{x}^{(\eta')}}^{(\eta'')} & \hat{\mathbf{R}}_{\mathbf{x},\mathbf{x}^{(\eta)}}^{(\eta'')} & \hat{\mathbf{R}}_{\mathbf{x},\mathbf{x}}^{(\eta'')} \end{bmatrix} \tag{24}$$

can be seen as the sample covariance matrix estimator of $\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}$. Here, we must note that in the transition from (22) to (23) we have used the relation $\bar{\mathbf{x}}^{\mathrm{H}}\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}^{-1}\bar{\mathbf{x}} = \Re(\bar{\mathbf{x}}^{\mathrm{H}}\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}^{-1}\bar{\mathbf{x}}) = \Re(\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}^{-1}\bar{\mathbf{x}}\bar{\mathbf{x}}^{\mathrm{H}})$. Alternatively, we could write $\Re[\mathrm{Tr}(\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}^{-1}\hat{\mathbf{R}}_{\bar{\mathbf{x}},\bar{\mathbf{x}}})] = \mathrm{Tr}(\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}^{-1/2}\hat{\mathbf{R}}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}^{-1/2})$.

Finally, assuming for notational simplicity that $\hat{\mathbf{R}}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}$ is nonsingular[2] (which obviously requires $T \geq 4n$), we are ready to obtain the ML estimates of the augmented covariance matrix $\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}$ under the three different hypotheses.

1) $\mathbb{Q}$-Proper Vectors (Hypothesis $\mathcal{H}_{\mathbb{Q}}$): In the case of $\mathbb{Q}$-proper vectors, the ML estimation problem can be written as

$$\underset{\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}\in\mathcal{R}_{\mathbb{Q}}}{\text{maximize}} \quad \mathcal{L}(\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}) \tag{25}$$

[2]We must note that, replacing matrix inverses by Moore–Penrose pseudoinverses, the derived GLRTs can be directly applied in the case of rank-deficient sample covariance matrices $\hat{\mathbf{R}}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}$. See [35] and [49] for the complex counterpart.

where $\mathcal{R}_{\mathbb{Q}}$ denotes the convex set of $\mathbb{Q}$-proper augmented covariance matrices, i.e.,

$$\mathcal{R}_{\mathbb{Q}} = \{\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}} | \mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta)}} = \mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta')}} = \mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta'')}} = \mathbf{0}_{n\times n}\}. \tag{26}$$

Lemma 3: Under the hypothesis $\mathcal{H}_{\mathbb{Q}}$, the ML estimate of the augmented covariance matrix is given by

$$\hat{\mathbf{D}}_{\mathbb{Q}} = \begin{bmatrix} \hat{\mathbf{R}}_{\mathbf{x},\mathbf{x}} & \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} \\ \mathbf{0}_{n\times n} & \hat{\mathbf{R}}_{\mathbf{x},\mathbf{x}}^{(\eta)} & \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} \\ \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} & \hat{\mathbf{R}}_{\mathbf{x},\mathbf{x}}^{(\eta')} & \mathbf{0}_{n\times n} \\ \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} & \hat{\mathbf{R}}_{\mathbf{x},\mathbf{x}}^{(\eta'')} \end{bmatrix} \tag{27}$$

which results in a log-likelihood function

$$\mathcal{L}(\hat{\mathbf{D}}_{\mathbb{Q}}) = -\frac{1}{2}\ln|\hat{\mathbf{D}}_{\mathbb{Q}}| - 2n. \tag{28}$$

Proof: Let us start by noting that, under $\mathcal{H}_{\mathbb{Q}}$, $\mathcal{L}(\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}})$ can be rewritten as

$$\begin{aligned}\mathcal{L}(\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}) &= -\frac{1}{2}\ln|\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}| - \frac{1}{2}\Re\left[\operatorname{Tr}\left(\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}^{-1}\hat{\mathbf{D}}_{\mathbb{Q}}\right)\right] \\ &= -D(\hat{\mathbf{D}}_{\mathbb{Q}} \,\|\, \mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}) - \frac{1}{2}\ln|\hat{\mathbf{D}}_{\mathbb{Q}}| - 2n \end{aligned} \tag{29}$$

where $D(\hat{\mathbf{D}}_{\mathbb{Q}} \,\|\, \mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}})$ is the Kullback–Leibler divergence [25], [50] between two zero-mean quaternion Gaussian distributions with augmented covariance matrices $\hat{\mathbf{D}}_{\mathbb{Q}}$ and $\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}$. Therefore, the ML estimation problem reduces to the minimization of the positive term $D(\hat{\mathbf{D}}_{\mathbb{Q}}\|\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}})$, which vanishes for $\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}} = \hat{\mathbf{D}}_{\mathbb{Q}}$. ■

2) $\mathbb{C}^{\eta}$-Proper Vectors (Hypothesis $\mathcal{H}_{\mathbb{C}^{\eta}}$): In this case, the ML estimation problem is

$$\underset{\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}\in\mathcal{R}_{\mathbb{C}^{\eta}}}{\text{maximize}} \quad \mathcal{L}(\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}), \tag{30}$$

with the convex set

$$\mathcal{R}_{\mathbb{C}^{\eta}} = \left\{\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}} \,|\, \mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta')}} = \mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta'')}} = \mathbf{0}_{n\times n}\right\}. \tag{31}$$

Lemma 4: Under the hypothesis $\mathcal{H}_{\mathbb{C}^{\eta}}$, the ML estimate of the augmented covariance matrix is

$$\hat{\mathbf{D}}_{\mathbb{C}^{\eta}} = \begin{bmatrix} \hat{\mathbf{R}}_{\mathbf{x},\mathbf{x}} & \hat{\mathbf{R}}_{\mathbf{x},\mathbf{x}^{(\eta)}} & \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} \\ \hat{\mathbf{R}}_{\mathbf{x},\mathbf{x}^{(\eta)}}^{(\eta)} & \hat{\mathbf{R}}_{\mathbf{x},\mathbf{x}}^{(\eta)} & \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} \\ \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} & \hat{\mathbf{R}}_{\mathbf{x},\mathbf{x}}^{(\eta')} & \hat{\mathbf{R}}_{\mathbf{x},\mathbf{x}^{(\eta)}}^{(\eta')} \\ \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} & \hat{\mathbf{R}}_{\mathbf{x},\mathbf{x}^{(\eta)}}^{(\eta'')} & \hat{\mathbf{R}}_{\mathbf{x},\mathbf{x}}^{(\eta'')} \end{bmatrix} \tag{32}$$

which yields the log-likelihood

$$\mathcal{L}(\hat{\mathbf{D}}_{\mathbb{C}^{\eta}}) = -\frac{1}{2}\ln|\hat{\mathbf{D}}_{\mathbb{C}^{\eta}}| - 2n. \tag{33}$$

Proof: The proof is identical to that of Lemma 3. ■

3) Possibly Improper Vectors (Hypothesis $\mathcal{H}_{\mathcal{I}}$): Finally, in the case of (possibly) improper vectors, we do not need to impose any particular structure on the augmented covariance matrix, and the ML estimation problem is

$$\underset{\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}}{\text{maximize}} \quad \mathcal{L}(\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}). \tag{34}$$

Therefore, the ML estimate of the augmented covariance matrix is directly given by the sample covariance matrix estimator, which results in a log-likelihood function

$$\mathcal{L}(\hat{\mathbf{R}}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}) = -\frac{1}{2}\ln|\hat{\mathbf{R}}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}| - 2n. \tag{35}$$

B. $\mathbb{Q}$-Properness GLRT

After obtaining the ML estimates of the augmented covariance matrix under the three different hypotheses, the derivation of the GLRTs is straightforward. Let us start by considering the following binary hypothesis test:

$$\begin{aligned} &\text{Null Hypothesis } \mathcal{H}_{\mathbb{Q}}: && \mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}} \in \mathcal{R}_{\mathbb{Q}} \\ &\text{Alternative Hypothesis } \mathcal{H}_{\mathcal{I}}: && \mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}} \notin \mathcal{R}_{\mathbb{Q}}. \end{aligned}$$

That is, we want to decide whether $\mathbf{x}$ is $\mathbb{Q}$-proper or not. Thus, taking the logarithm of the generalized likelihood ratio, we obtain the GLRT statistic

$$\hat{\mathcal{P}}_{\mathbb{Q}} = \mathcal{L}(\hat{\mathbf{R}}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}) - \mathcal{L}(\hat{\mathbf{D}}_{\mathbb{Q}}) = -\frac{1}{2}\ln|\hat{\mathbf{\Phi}}_{\mathbb{Q}}| \tag{36}$$

where $\hat{\mathbf{\Phi}}_{\mathbb{Q}} = \hat{\mathbf{D}}_{\mathbb{Q}}^{-1/2}\hat{\mathbf{R}}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}\hat{\mathbf{D}}_{\mathbb{Q}}^{-1/2}$ is defined as the $\mathbb{Q}$-coherence matrix, which is closely related to the multiset extensions of canonical correlation analysis (CCA) [51]–[53]. Specifically, $\hat{\mathbf{\Phi}}_{\mathbb{Q}}$ appears in the CCA of the quaternion random vectors $\mathbf{x}$, $\mathbf{x}^{(\eta)}$, $\mathbf{x}^{(\eta')}$, and $\mathbf{x}^{(\eta'')}$.

Interestingly, the test statistic $\hat{\mathcal{P}}_{\mathbb{Q}}$ can be seen as an estimate of the $\mathbb{Q}$-improperness measure proposed in [25], which is based on the Kullback–Leibler divergence [50] between two zero-mean quaternion Gaussian distributions and provides the entropy loss due to the $\mathbb{Q}$-improperness of the quaternion random vector $\mathbf{x}$—that is, due to the additional correlation (not contained in $\mathbf{R}_{\mathbf{x},\mathbf{x}}$) among the real components of the quaternion vector. Moreover, we must note that $\hat{\mathcal{P}}_{\mathbb{Q}}$ satisfies the following important properties.

Property 5: $\hat{\mathcal{P}}_{\mathbb{Q}}$ is invariant to invertible linear transformations.

Proof: Consider the linearly transformed data $\mathbf{u}[t] = \mathbf{F}_1^{\mathrm{H}}\mathbf{x}[t]$ $(t = 1, \ldots T)$, with $\mathbf{F}_1 \in \mathbb{H}^{n\times n}$ an invertible matrix. Then, it is easy to see that the associated $\mathbb{Q}$-properness GLRT statistic is

$$-\frac{1}{2}\ln|\bar{\mathbf{F}}^{\mathrm{H}}\hat{\mathbf{R}}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}\bar{\mathbf{F}}| + \frac{1}{2}\ln|\bar{\mathbf{F}}^{\mathrm{H}}\hat{\mathbf{D}}_{\mathbb{Q}}\bar{\mathbf{F}}| = -\frac{1}{2}\ln|\hat{\mathbf{\Phi}}_{\mathbb{Q}}| = \hat{\mathcal{P}}_{\mathbb{Q}} \tag{37}$$

where $\bar{\mathbf{F}}$ is defined as

$$\bar{\mathbf{F}} = \begin{bmatrix} \mathbf{F}_1 & \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} \\ \mathbf{0}_{n\times n} & \mathbf{F}_1^{(\eta)} & \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} \\ \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} & \mathbf{F}_1^{(\eta')} & \mathbf{0}_{n\times n} \\ \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} & \mathbf{F}_1^{(\eta'')} \end{bmatrix}. \tag{38}$$

That is, the GLRT statistics for the original $(\mathbf{x}[t])$ and transformed $(\mathbf{u}[t])$ data are identical. ■

Property 6: $\hat{\mathcal{P}}_{\mathbb{Q}}$ is independent of the orthogonal basis $\{\eta, \eta', \eta''\}$.

Proof: Let us write the GLRT statistic as

$$\hat{\mathcal{P}}_{\mathbb{Q}} = -\frac{1}{2}\ln|\hat{\mathbf{\Phi}}_{\mathbb{Q}}| = -\frac{1}{2}\ln|\hat{\mathbf{R}}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}| + \frac{1}{2}\ln|\hat{\mathbf{D}}_{\mathbb{Q}}|.$$

Now, as a direct consequence of Lemma 2, $|\hat{\mathbf{R}}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}|$ is independent of the orthogonal basis. Moreover, it is easy to see that $\ln|\hat{\mathbf{D}}_{\mathbb{Q}}| = 4\ln|\hat{\mathbf{R}}_{\mathbf{x},\mathbf{x}}|$. Therefore, we can conclude that $\hat{\mathcal{P}}_{\mathbb{Q}}$ does not depend on the particular choice of the orthogonal basis $\{\eta, \eta', \eta''\}$. ■

Summarizing, the $\mathbb{Q}$-properness GLRT reduces to the comparison of $\hat{\mathcal{P}}_{\mathbb{Q}}$ with some fixed threshold $\gamma_{\mathbb{Q}}$

$$\hat{\mathcal{P}}_{\mathbb{Q}} \underset{\mathcal{H}_{\mathbb{Q}}}{\overset{\mathcal{H}_{\mathcal{I}}}{\gtrless}} \gamma_{\mathbb{Q}}. \quad (39)$$

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

Here, we consider the problem of determining whether $\mathbf{x}$ is $\mathbb{C}^{\eta}$-proper or not, i.e., our hypotheses testing problem can be written as

$$\begin{array}{ll} \text{Null Hypothesis } \mathcal{H}_{\mathbb{C}^{\eta}}: & \mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}} \in \mathcal{R}_{\mathbb{C}^{\eta}} \\ \text{Alternative Hypothesis } \mathcal{H}_{\mathcal{I}}: & \mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}} \notin \mathcal{R}_{\mathbb{C}^{\eta}}. \end{array}$$

Following the lines in Section III-B, we easily obtain the GLRT statistic

$$\hat{\mathcal{P}}_{\mathbb{C}^{\eta}} = \mathcal{L}(\hat{\mathbf{R}}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}) - \mathcal{L}(\hat{\mathbf{D}}_{\mathbb{C}^{\eta}}) = -\frac{1}{2}\ln|\hat{\mathbf{\Phi}}_{\mathbb{C}^{\eta}}| \quad (40)$$

where now $\hat{\mathbf{\Phi}}_{\mathbb{C}^{\eta}} = \hat{\mathbf{D}}_{\mathbb{C}^{\eta}}^{-1/2}\hat{\mathbf{R}}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}\hat{\mathbf{D}}_{\mathbb{C}^{\eta}}^{-1/2}$ is the $\mathbb{C}^{\eta}$-coherence matrix, which appears in the CCA of the random vectors $\tilde{\mathbf{x}} = [\mathbf{x}^{\mathrm{T}}, \mathbf{x}^{(\eta)^{\mathrm{T}}}]^{\mathrm{T}}$ and $\tilde{\mathbf{x}}^{(\eta')}$.

Analogously to the previous case, $\hat{\mathcal{P}}_{\mathbb{C}^{\eta}}$ can be seen as an estimate of the $\mathbb{C}^{\eta}$-improperness measure proposed in [25], which provides the entropy loss due to the $\mathbb{C}^{\eta}$-improperness of $\mathbf{x}$, and satisfies the following invariance property.

Property 7: $\hat{\mathcal{P}}_{\mathbb{C}^{\eta}}$ is invariant to invertible *semi-widely* linear transformations.

Proof: Let us define the *semi-widely* linear transformation $\mathbf{u} = \mathbf{F}_1^{\mathrm{H}}\mathbf{x} + \mathbf{F}_{\eta}^{\mathrm{H}}\mathbf{x}^{(\eta)}$, with $\mathbf{F}_1, \mathbf{F}_{\eta} \in \mathbb{H}^{n\times n}$ providing an invertible matrix

$$\bar{\mathbf{F}} = \begin{bmatrix} \mathbf{F}_1 & \mathbf{F}_{\eta} & \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} \\ \mathbf{F}_{\eta}^{(\eta)} & \mathbf{F}_1^{(\eta)} & \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} \\ \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} & \mathbf{F}_1^{(\eta')} & \mathbf{F}_{\eta}^{(\eta')} \\ \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} & \mathbf{F}_{\eta}^{(\eta'')} & \mathbf{F}_1^{(\eta'')} \end{bmatrix}. \quad (41)$$

Then, the associated $\mathbb{C}^{\eta}$-properness GLRT statistic is

$$-\frac{1}{2}\ln|\bar{\mathbf{F}}^{\mathrm{H}}\hat{\mathbf{R}}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}\bar{\mathbf{F}}| + \frac{1}{2}\ln|\bar{\mathbf{F}}^{\mathrm{H}}\hat{\mathbf{D}}_{\mathbb{C}^{\eta}}\bar{\mathbf{F}}| = -\frac{1}{2}\ln|\hat{\mathbf{\Phi}}_{\mathbb{C}^{\eta}}| = \hat{\mathcal{P}}_{\mathbb{C}^{\eta}}. \quad (42)$$

■

Moreover, $\hat{\mathcal{P}}_{\mathbb{C}^{\eta}}$ can be rewritten in terms of the vectors in the Cayley–Dickson representation $\mathbf{x} = \mathbf{a}_1 + \eta''\mathbf{a}_2$ as [25]

$$\hat{\mathcal{P}}_{\mathbb{C}^{\eta}} = -\frac{1}{2}\ln\left|\hat{\mathbf{\Phi}}_{\bar{\mathbf{a}}}\right| \quad (43)$$

where $\hat{\mathbf{\Phi}}_{\bar{\mathbf{a}}}$ is the coherence matrix for the complex random vector $\mathbf{a} = [\mathbf{a}_1^{\mathrm{T}}, \mathbf{a}_2^{\mathrm{T}}]^{\mathrm{T}}$. That is, defining the augmented vector $\bar{\mathbf{a}} = [\mathbf{a}^{\mathrm{T}}, \mathbf{a}^{\mathrm{H}}]^{\mathrm{T}}$, the coherence matrix is obtained as $\hat{\mathbf{\Phi}}_{\bar{\mathbf{a}}} = \hat{\mathbf{D}}_{\bar{\mathbf{a}}}^{-1/2}\hat{\mathbf{R}}_{\bar{\mathbf{a}},\bar{\mathbf{a}}}\hat{\mathbf{D}}_{\bar{\mathbf{a}}}^{-1/2}$, where

$$\hat{\mathbf{R}}_{\bar{\mathbf{a}},\bar{\mathbf{a}}} = \frac{1}{T}\sum_{t=0}^{T-1}\bar{\mathbf{a}}[t]\bar{\mathbf{a}}^{\mathrm{H}}[t] = \begin{bmatrix} \hat{\mathbf{R}}_{\mathbf{a},\mathbf{a}} & \hat{\mathbf{R}}_{\mathbf{a},\mathbf{a}^*} \\ \hat{\mathbf{R}}_{\mathbf{a},\mathbf{a}^*} & \hat{\mathbf{R}}_{\mathbf{a},\mathbf{a}}^* \end{bmatrix} \quad (44)$$

is the sample covariance estimator of the complex augmented covariance matrix, and

$$\hat{\mathbf{D}}_{\bar{\mathbf{a}}} = \begin{bmatrix} \hat{\mathbf{R}}_{\mathbf{a},\mathbf{a}} & \mathbf{0}_{n\times n} \\ \mathbf{0}_{n\times n} & \hat{\mathbf{R}}_{\mathbf{a},\mathbf{a}}^* \end{bmatrix}. \quad (45)$$

Interestingly, $\hat{\mathcal{P}}_{\mathbb{C}^{\eta}} = -(1/2)\ln|\hat{\mathbf{\Phi}}_{\bar{\mathbf{a}}}|$ is also the GLRT statistic for determining whether $\mathbf{a}$ is (complex) proper or not, or equivalently, for determining if $\mathbf{a}_1$ and $\mathbf{a}_2$ are jointly complex-proper [32], [33], [48], [54], [55]. That is, as suggested by Property 2, the $\mathbb{C}^{\eta}$-properness test reduces to the evaluation, and comparison to a previously fixed threshold, of the complex-improperness measure of $\mathbf{a}$

$$\hat{\mathcal{P}}_{\mathbb{C}^{\eta}} \underset{\mathcal{H}_{\mathbb{C}^{\eta}}}{\overset{\mathcal{H}_{\mathcal{I}}}{\gtrless}} \gamma_{\mathbb{C}^{\eta}}. \quad (46)$$

D. $\mathbb{C}^{\eta}$-Properness Versus $\mathbb{Q}$-Properness GLRT

Finally, let us consider the case in which we already know that the quaternion random vector $\mathbf{x}$ is $\mathbb{C}^{\eta}$-proper (see the simulations for a practical example). Then, we should determine whether it is also $\mathbb{Q}$-proper, and our testing problem is

$$\begin{array}{ll} \text{Null Hypothesis } \mathcal{H}_{\mathbb{Q}}: & \mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}} \in \mathcal{R}_{\mathbb{Q}} \\ \text{Alternative Hypothesis } \mathcal{H}_{\mathbb{C}^{\eta}}: & \mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}} \in \mathcal{R}_{\mathbb{C}^{\eta}}. \end{array}$$

Following the lines in Sections III-A and III-B, we obtain the GLRT statistic

$$\hat{\mathcal{P}}_{\mathbb{R}^{\eta}} = \mathcal{L}(\hat{\mathbf{D}}_{\mathbb{C}^{\eta}}) - \mathcal{L}(\hat{\mathbf{D}}_{\mathbb{Q}}) = \hat{\mathcal{P}}_{\mathbb{Q}} - \hat{\mathcal{P}}_{\mathbb{C}^{\eta}} = -\frac{1}{2}\ln|\hat{\mathbf{\Phi}}_{\mathbb{R}^{\eta}}| \quad (47)$$

where $\hat{\mathbf{\Phi}}_{\mathbb{R}^{\eta}} = \hat{\mathbf{D}}_{\mathbb{Q}}^{-1/2}\hat{\mathbf{D}}_{\mathbb{C}^{\eta}}\hat{\mathbf{D}}_{\mathbb{Q}}^{-1/2}$ is defined as the $\mathbb{R}^{\eta}$-coherence matrix. Analogously to the previous cases, $\hat{\mathcal{P}}_{\mathbb{R}^{\eta}}$ can be seen as an estimate of the $\mathbb{R}^{\eta}$-improperness degree in [25], which is a measure of the entropy loss due to the $\mathbb{Q}$-improperness of the $\mathbb{C}^{\eta}$-proper vector $\mathbf{x}$, and satisfies the following invariance property.

Property 8: $\hat{\mathcal{P}}_{\mathbb{R}^{\eta}}$ is invariant to invertible linear transformations.

Proof: This is a direct consequence of the decomposition $\hat{\mathcal{P}}_{\mathbb{R}^{\eta}} = \hat{\mathcal{P}}_{\mathbb{Q}} - \hat{\mathcal{P}}_{\mathbb{C}^{\eta}}$ and the invariances of $\hat{\mathcal{P}}_{\mathbb{Q}}$ and $\hat{\mathcal{P}}_{\mathbb{C}^{\eta}}$. ■

Therefore, as we have pointed out before, the $\mathbb{R}^{\eta}$-properness naturally appears as the *difference* between the two main kinds of quaternion properness, and the $\mathbb{C}^{\eta}$- versus $\mathbb{Q}$-properness GLRT reduces to

$$\hat{\mathcal{P}}_{\mathbb{R}^{\eta}} \underset{\mathcal{H}_{\mathbb{Q}}}{\overset{\mathcal{H}_{\mathbb{C}^{\eta}}}{\gtrless}} \gamma_{\mathbb{R}^{\eta}} \quad (48)$$

where $\gamma_{\mathbb{R}^{\eta}}$ is a previously fixed threshold.

IV. GLRTs for Unknown Bases

In Section III, we have derived the GLRTs for testing the properness of a quaternion random vector under the *a priori* knowledge of the orthogonal basis $\{\eta, \eta', \eta''\}$. That is, we have assumed that the pure unit quaternion η, for the analysis of the $\mathbb{C}^\eta$-properness, is fixed. However, in practice this parameter could be unknown, and we should consider the $\mathbb{C}^\eta$-properness tests for all the possible values of η.

In this section, we generalize the previous results by including the estimation of η in the GLRTs. As we will see, the ML estimation problem amounts to finding the principal $\mathbb{C}$-properness direction η, that is, the pure unit quaternion η minimizing the $\mathbb{C}^\eta$-improperness measure $\hat{\mathcal{P}}_{\mathbb{C}^\eta}$. Interestingly, this problem can be seen as that of finding a decomposition of the quaternion random vector $\mathbf{x}$ into two jointly proper complex vectors. Finally, we should also note that after obtaining the ML estimate of η, we will be ready to apply the optimal *semi-widely* linear processing $\mathbf{u} = \mathbf{F}_1\mathbf{x} + \mathbf{F}_\eta\mathbf{x}^{(\eta)}$.

A. Problem Statement

Following the derivation in Section III, we should start by obtaining the joint ML estimates of η and $\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}$ under the three different hypotheses. However, under $\mathcal{H}_{\mathbb{Q}}$ and $\mathcal{H}_{\mathcal{I}}$, the maximum of the log-likelihood function does not depend on η, which can be seen as a direct consequence of Lemma 2. Therefore, our problem reduces to the joint ML estimation of η and $\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}$ under the hypothesis of $\mathbb{C}$-proper vectors,[3] i.e.,

$$\underset{\eta\in\mathcal{Q},\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}\in\mathcal{R}_{\mathbb{C}^\eta}}{\text{maximize}} \quad \mathcal{L}(\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}) \tag{49}$$

where $\mathcal{Q}$ denotes the set of pure unit quaternions. Now, it is clear that the previous problem can be rewritten as

$$\underset{\eta\in\mathcal{Q}}{\text{maximize}}\ \underset{\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}\in\mathcal{R}_{\mathbb{C}^\eta}}{\text{supremum}} \quad \mathcal{L}(\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}) \tag{50}$$

or, as a direct consequence of Lemma 4

$$\underset{\eta\in\mathcal{Q}}{\text{maximize}} \quad \mathcal{L}(\hat{\mathbf{D}}_{\mathbb{C}^\eta}). \tag{51}$$

Thus, noting that $\hat{\mathcal{P}}_{\mathbb{C}^\eta} = \mathcal{L}(\hat{\mathbf{R}}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}) - \mathcal{L}(\hat{\mathbf{D}}_{\mathbb{C}^\eta})$, and since $\mathcal{L}(\hat{\mathbf{R}}_{\bar{\mathbf{x}},\bar{\mathbf{x}}})$ is independent of η, our ML estimation problem can be written as

$$\underset{\eta\in\mathcal{Q}}{\text{minimize}} \quad \hat{\mathcal{P}}_{\mathbb{C}^\eta} \tag{52}$$

or equivalently

$$\underset{\eta\in\mathcal{Q}}{\text{maximize}} \quad \hat{\mathcal{P}}_{\mathbb{R}^\eta}. \tag{53}$$

That is, as we could expect, we are looking for the pure unit quaternion η minimizing (equivalently maximizing) the estimated $\mathbb{C}^\eta$ (equiv. $\mathbb{R}^\eta$) improperness measure. Finally, taking into account the ML estimates of η and $\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}}$, the overall $\mathbb{C}$-properness test is

$$\underset{\eta\in\mathcal{Q}}{\text{infimum}}(\hat{\mathcal{P}}_{\mathbb{C}^\eta}) \underset{\mathcal{H}_{\mathbb{C}}}{\overset{\mathcal{H}_{\mathcal{I}}}{\gtrless}} \gamma_{\mathbb{C}} \tag{54}$$

[3]We say that a vector is $\mathbb{C}$-proper iff it is $\mathbb{C}^\eta$-proper for some η.

and equivalently, the problem of testing $\mathbb{C}$-properness versus $\mathbb{Q}$-properness results in

$$\underset{\eta\in\mathcal{Q}}{\text{supremum}}(\hat{\mathcal{P}}_{\mathbb{R}}) \underset{\mathcal{H}_{\mathbb{Q}}}{\overset{\mathcal{H}_{\mathbb{C}}}{\gtrless}} \gamma_{\mathbb{R}}. \tag{55}$$

B. Formulation of the Optimization Problem

Let us start by rewriting the $\mathbb{R}^\eta$-coherence matrix as

$$\begin{aligned}\hat{\mathbf{\Phi}}_{\mathbb{R}^\eta} &= \hat{\mathbf{D}}_{\mathbb{Q}}^{-1/2}\hat{\mathbf{D}}_{\mathbb{C}^\eta}\hat{\mathbf{D}}_{\mathbb{Q}}^{-1/2} \\ &= \begin{bmatrix} \mathbf{I}_n & \hat{\mathbf{R}}_{\mathbf{y},\mathbf{y}^{(\eta)}} & \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} \\ \hat{\mathbf{R}}_{\mathbf{y},\mathbf{y}^{(\eta)}}^{(\eta)} & \mathbf{I}_n & \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} \\ \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} & \mathbf{I}_n & \hat{\mathbf{R}}_{\mathbf{y},\mathbf{y}^{(\eta)}}^{(\eta')} \\ \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} & \hat{\mathbf{R}}_{\mathbf{y},\mathbf{y}^{(\eta)}}^{(\eta'')} & \mathbf{I}_n \end{bmatrix}\end{aligned} \tag{56}$$

where $\hat{\mathbf{R}}_{\mathbf{y},\mathbf{y}^{(\eta)}} = \hat{\mathbf{R}}_{\mathbf{x},\mathbf{x}}^{-1/2}\hat{\mathbf{R}}_{\mathbf{x},\mathbf{x}^{(\eta)}}(\hat{\mathbf{R}}_{\mathbf{x},\mathbf{x}}^{-1/2})^{(\eta)}$ can be seen as an estimate of the complementary covariance matrix of the prewhitened vector $\mathbf{y} = \mathbf{R}_{\mathbf{x},\mathbf{x}}^{-1/2}\mathbf{x}$. This reduces the $\mathbb{R}^\eta$-improperness measure to $\hat{\mathcal{P}}_{\mathbb{R}^\eta} = -\ln|\mathbf{I}_n - \hat{\mathbf{R}}_{\mathbf{y},\mathbf{y}^{(\eta)}}\hat{\mathbf{R}}_{\mathbf{y},\mathbf{y}^{(\eta)}}^{\mathrm{H}}|$, which results in the following ML estimation problem:

$$\underset{\eta\in\mathcal{Q}}{\text{maximize}} \quad -\ln\left|\mathbf{I}_n - \hat{\mathbf{R}}_{\mathbf{y},\mathbf{y}^{(\eta)}}\hat{\mathbf{R}}_{\mathbf{y},\mathbf{y}^{(\eta)}}^{\mathrm{H}}\right|. \tag{57}$$

Finally, defining the matrices $\hat{\mathbf{\Psi}}_\nu = \hat{\mathbf{R}}_{\mathbf{y},\mathbf{y}^{(\nu)}}\nu$ for all pure unit quaternions ν, the objective function in (57) can be rewritten as $\hat{\mathcal{P}}_{\mathbb{R}^\eta} = -\ln|\mathbf{I}_n - \hat{\mathbf{\Psi}}_\eta\hat{\mathbf{\Psi}}_\eta^{\mathrm{H}}|$, and applying Lemma 2 we have

$$\hat{\mathbf{\Psi}}_\eta = \eta_\nu\hat{\mathbf{\Psi}}_\nu + \eta_{\nu'}\hat{\mathbf{\Psi}}_{\nu'} + \eta_{\nu''}\hat{\mathbf{\Psi}}_{\nu''} \tag{58}$$

where $\eta_\nu, \eta_{\nu'}, \eta_{\nu''} \in \mathbb{R}$ are the coordinates of the pure unit quaternion η in the arbitrary orthogonal basis $\{\nu, \nu', \nu''\}$, i.e., $\eta = \eta_\nu\nu + \eta_{\nu'}\nu' + \eta_{\nu''}\nu''$. Thus, (57) can be rewritten as

$$\begin{aligned} \underset{\hat{\mathbf{\Psi}}_\eta,\boldsymbol{\eta}}{\text{minimize}} \quad & \ln\left|\mathbf{I}_n - \hat{\mathbf{\Psi}}_\eta\hat{\mathbf{\Psi}}_\eta^{\mathrm{H}}\right| \\ \text{subject to} \quad & \hat{\mathbf{\Psi}}_\eta = \eta_\nu\hat{\mathbf{\Psi}}_\nu + \eta_{\nu'}\hat{\mathbf{\Psi}}_{\nu'} + \eta_{\nu''}\hat{\mathbf{\Psi}}_{\nu''} \\ & \|\boldsymbol{\eta}\| \le 1 \end{aligned} \tag{59}$$

where $\boldsymbol{\eta} = [\eta_\nu, \eta_{\nu'}, \eta_{\nu''}]^{\mathrm{T}}$, and the last constraint, which forces $\|\boldsymbol{\eta}\| = 1$ (equivalently $|\eta| = 1$), has been relaxed to an inequality because the cost function is monotonically decreasing with $\|\boldsymbol{\eta}\|$.

C. Proposed Algorithm: Successive Convex Approximations

Unfortunately, the above optimization problem is not convex due to the cost function $\ln|\mathbf{I}_n - \hat{\mathbf{\Psi}}_\eta\hat{\mathbf{\Psi}}_\eta^{\mathrm{H}}|$ in (59), which precludes its solution by means of standard convex optimization tools [43]. Here, in order to find reliable approximated solutions, we propose to apply the successive convex approximations method [40]–[42]. This technique relies on solving a series of convex problems, in which the nonconvex cost function $f(\boldsymbol{\eta}) = \ln|\mathbf{I}_n - \hat{\mathbf{\Psi}}_\eta\hat{\mathbf{\Psi}}_\eta^{\mathrm{H}}|$ is replaced by a convex approximation $\hat{f}(\boldsymbol{\eta})$. The following lemmas provide sufficient conditions for the convergence of the successive convex approximations method, as well as a particular approximation satisfying the convergence conditions.

Lemma 5: Consider the optimization problem

$$\underset{\boldsymbol{\eta}\in\mathcal{S}}{\text{minimize}}\, f(\boldsymbol{\eta}) \tag{60}$$

where $\mathcal{S}$ is a convex set and $f(\boldsymbol{\eta})$ is a nonconvex function. If the convex approximations $\hat{f}(\boldsymbol{\eta})$ of the cost function $f(\boldsymbol{\eta})$ satisfy:

- $\hat{f}(\boldsymbol{\eta}) \geq f(\boldsymbol{\eta})$ for all $\boldsymbol{\eta} \in \mathcal{S}$;
- $\hat{f}(\boldsymbol{\eta}_0) = f(\boldsymbol{\eta}_0)$, where $\boldsymbol{\eta}_0$ is the optimal solution of the approximated problem in the previous iteration;
- $\nabla \hat{f}(\boldsymbol{\eta}_0) = \nabla f(\boldsymbol{\eta}_0)$, where ∇ is the gradient operator;

then the successive convex approximations method, based on the solutions of the convex problems

$$\underset{\boldsymbol{\eta}\in\mathcal{S}}{\text{minimize}}\, \hat{f}(\boldsymbol{\eta}) \tag{61}$$

guarantees the convergence of $f(\boldsymbol{\eta})$, and after the convergence $\boldsymbol{\eta}$ satisfies the KKT conditions of the original problem.

Proof: See [40] and [41] for the proof and some minor technical details. ■

Lemma 6: Consider the cost function $f(\boldsymbol{\eta}) = \ln|\mathbf{I}_n - \hat{\boldsymbol{\Psi}}_\eta \hat{\boldsymbol{\Psi}}_\eta^{\mathrm{H}}|$ and the matrix $\hat{\boldsymbol{\Xi}}_0 = \hat{\boldsymbol{\Psi}}_0 \hat{\boldsymbol{\Psi}}_0^{\mathrm{H}}$, where $\hat{\boldsymbol{\Psi}}_0$ denotes the value of $\hat{\boldsymbol{\Psi}}_\eta$ in the previous iteration. Then, the approximation

$$\begin{aligned}\hat{f}(\boldsymbol{\eta}) &= \ln|\mathbf{I}_n - \hat{\boldsymbol{\Xi}}_0| + \mathrm{Tr}((\mathbf{I}_n - \hat{\boldsymbol{\Xi}}_0)^{-1}\hat{\boldsymbol{\Xi}}_0) \\ &\quad - \mathrm{Tr}\left(\hat{\boldsymbol{\Psi}}_\eta^H (\mathbf{I}_n - \hat{\boldsymbol{\Xi}}_0)^{-1} \hat{\boldsymbol{\Psi}}_\eta\right)\end{aligned} \tag{62}$$

satisfies the convergence conditions in Lemma 5.

Proof: Defining $\hat{\boldsymbol{\Xi}}_\eta = \hat{\boldsymbol{\Psi}}_\eta \hat{\boldsymbol{\Psi}}_\eta^{\mathrm{H}}$, the cost function is $f(\boldsymbol{\eta}) = \ln|\mathbf{I}_n - \hat{\boldsymbol{\Xi}}_\eta|$, and $\hat{f}(\boldsymbol{\eta})$ is its first-order Taylor's series approximation (with respect to $\hat{\boldsymbol{\Xi}}_\eta$) around $\hat{\boldsymbol{\Xi}}_0$. Now, it is easy to check that the approximation satisfies the second and third convergence conditions in Lemma 5. Finally, since the cost function is concave in $\hat{\boldsymbol{\Xi}}_\eta$, the approximation $\hat{f}(\boldsymbol{\eta})$ also satisfies the first convergence condition. ■

Using the proposed approximation, the convex problem to be solved in each iteration is

$$\begin{aligned}\underset{\hat{\boldsymbol{\Theta}}_\eta, \boldsymbol{\eta}}{\text{maximize}} \quad & \mathrm{Tr}\left(\hat{\boldsymbol{\Theta}}_\eta^{\mathrm{H}} \hat{\boldsymbol{\Theta}}_\eta\right) \\ \text{subject to} \quad & \hat{\boldsymbol{\Theta}}_\eta = \eta_\nu \hat{\boldsymbol{\Theta}}_\nu + \eta_{\nu'} \hat{\boldsymbol{\Theta}}_{\nu'} + \eta_{\nu''} \hat{\boldsymbol{\Theta}}_{\nu''} \\ & ||\boldsymbol{\eta}|| \leq 1\end{aligned} \tag{63}$$

where

$$\hat{\boldsymbol{\Theta}}_\nu = (\mathbf{I}_n - \hat{\boldsymbol{\Xi}}_0)^{-1/2} \hat{\boldsymbol{\Psi}}_\nu \tag{64}$$

$$\hat{\boldsymbol{\Theta}}_{\nu'} = (\mathbf{I}_n - \hat{\boldsymbol{\Xi}}_0)^{-1/2} \hat{\boldsymbol{\Psi}}_{\nu'} \tag{65}$$

$$\hat{\boldsymbol{\Theta}}_{\nu''} = (\mathbf{I}_n - \hat{\boldsymbol{\Xi}}_0)^{-1/2} \hat{\boldsymbol{\Psi}}_{\nu''}. \tag{66}$$

Thus, defining the matrix

$$\hat{\boldsymbol{\Theta}} = \left[\mathrm{vec}\left(\hat{\boldsymbol{\Theta}}_\nu\right) \quad \mathrm{vec}\left(\hat{\boldsymbol{\Theta}}_{\nu'}\right) \quad \mathrm{vec}\left(\hat{\boldsymbol{\Theta}}_{\nu''}\right)\right] \tag{67}$$

the previous problem can be rewritten as

$$\begin{aligned}\underset{\boldsymbol{\eta}}{\text{maximize}} \quad & \boldsymbol{\eta}^{\mathrm{T}} \boldsymbol{\Omega} \boldsymbol{\eta} \\ \text{subject to} \quad & ||\boldsymbol{\eta}||^2 \leq 1\end{aligned} \tag{68}$$

Algorithm 1: Principal $\mathbb{C}$-Properness Direction

Input: Estimates $\hat{\boldsymbol{\Psi}}_\nu$, $\hat{\boldsymbol{\Psi}}_{\nu'}$, $\hat{\boldsymbol{\Psi}}_{\nu''}$ in some basis $\{\nu, \nu', \nu''\}$.
Output: Principal $\mathbb{C}$-Properness Direction η.
Initialize: η at some arbitrary value.
repeat
  Obtain $\hat{\boldsymbol{\Psi}}_\eta = \eta_\nu \hat{\boldsymbol{\Psi}}_\nu + \eta_{\nu'} \hat{\boldsymbol{\Psi}}_{\nu'} + \eta_{\nu''} \hat{\boldsymbol{\Psi}}_{\nu''}$.
  Compute $\hat{\boldsymbol{\Xi}}_0 = \hat{\boldsymbol{\Psi}}_\eta \hat{\boldsymbol{\Psi}}_\eta^{\mathrm{H}}$.
  Obtain $\hat{\boldsymbol{\Theta}}_\nu$, $\hat{\boldsymbol{\Theta}}_{\nu'}$, and $\hat{\boldsymbol{\Theta}}_{\nu''}$ from (64)–(66).
  Obtain $\hat{\boldsymbol{\Theta}} = [\mathrm{vec}(\hat{\boldsymbol{\Theta}}_\nu) \quad \mathrm{vec}(\hat{\boldsymbol{\Theta}}_{\nu'}) \quad \mathrm{vec}(\hat{\boldsymbol{\Theta}}_{\nu''})]$.
  Extract $\boldsymbol{\eta}$ as the principal eigenvector of $\boldsymbol{\Omega} = \Re(\hat{\boldsymbol{\Theta}}^{\mathrm{H}} \hat{\boldsymbol{\Theta}})$.
until Convergence.

with $\boldsymbol{\Omega} = \Re(\hat{\boldsymbol{\Theta}}^{\mathrm{H}} \hat{\boldsymbol{\Theta}})$. Finally, the solution $\boldsymbol{\eta}$ is given by the principal eigenvector of the matrix $\boldsymbol{\Omega}$, and the overall algorithm for the estimation of the pure unit quaternion η is summarized in Algorithm 1.

V. Further Discussion

In this section, we provide some additional details about the distribution of the test statistics, the particularization of the obtained results to the scalar case, and the general multiple-hypotheses testing problem.

A. Distribution of the GLRT Statistics

As we have seen, the three proposed GLRTs reduce to the comparison of the estimated improperness measure with a threshold. Typically, the selection of the threshold is based on some performance criterion, such as a constant false alarm probability, which in this paper is defined as the probability of accepting the alternative (improper) hypothesis when the null (proper) hypothesis is true. Therefore, the selection of the threshold requires the knowledge of the GLRT statistic distribution under the null hypothesis.

Although the theoretical derivation of the distributions is in general a very difficult problem, the properties of the test statistics allow us to draw some interesting conclusions. First, the invariance of the statistics $\hat{\mathcal{P}}_{\mathbb{Q}}$, $\hat{\mathcal{P}}_{\mathbb{R}^\eta}$ (resp. $\hat{\mathcal{P}}_{\mathbb{C}^\eta}$) under linear (resp. *semi-widely* linear) transformations can be easily exploited to numerically determine the thresholds for a fixed false alarm probability in the case of known η. That is, following the lines in [33] for the case of complex vectors, the distributions of the statistics under the null hypothesis can be obtained by simulation using $\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}} = \mathbf{I}_{4n}$. Thus, as illustrated in Section VI-B, we *only* need to tabulate the simulation results for the different values of n (vector dimensionality) and T (number of available observations). Second, we must note that any complementary covariance matrix $\mathbf{R}_{\mathbf{x},\mathbf{x}^{(\nu)}}$ satisfies $\mathbf{R}_{\mathbf{x},\mathbf{x}^{(\nu)}}^{\mathrm{H}} = \mathbf{R}_{\mathbf{x},\mathbf{x}^{(\nu)}}^{(\nu)}$, which implies that the diagonal elements are orthogonal to ν. Therefore, a complementary covariance matrix $\mathbf{R}_{\mathbf{x},\mathbf{x}^{(\nu)}} \in \mathbb{H}^{n\times n}$ is completely specified by $n(2n+1)$ real numbers ($3n$ in the diagonal and $2n(n-1)$ above the diagonal). Thus, as a direct consequence of the Wilks' theorem [33], [56], we have that under the null hypothesis

$$2T\hat{\mathcal{P}}_{\mathbb{R}^\eta} \simeq \chi^2_{d_{\mathbb{R}^\eta}} \quad 2T\hat{\mathcal{P}}_{\mathbb{C}^\eta} \simeq \chi^2_{d_{\mathbb{C}^\eta}} \quad 2T\hat{\mathcal{P}}_{\mathbb{Q}} \simeq \chi^2_{d_{\mathbb{Q}}} \tag{69}$$

where $\simeq$ indicates convergence (for $T \to \infty$) in distribution, χ_d^2 denotes a chi-square random variable with d degrees of freedom, and

$$d_{\mathbb{R}^\eta} = \frac{1}{2} d_{\mathbb{C}^\eta} = \frac{1}{3} d_{\mathbb{Q}} = n(2n+1). \tag{70}$$

The case of estimated η is more complicated. Obviously, $\hat{P}_{\mathbb{Q}}$ is independent of the orthogonal basis, and the previous results apply. However, the $\mathbb{C}$-properness test statistic is not invariant under *semi-widely* linear transformations, which means that its distribution under the null hypothesis cannot be obtained from simulations using $\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}} = \mathbf{I}_{4n}$. In this case, if $\mathbf{x}$ is not $\mathbb{Q}$-proper, the direct application of the Wilks' theorem provides an asymptotic distribution

$$2T \operatorname*{infinum}_{\eta \in \mathcal{Q}} \left(\hat{\mathcal{P}}_{\mathbb{C}^\eta} \right) \simeq \chi^2_{d_{\mathbb{C}^\eta}-2}$$

Finally, the distribution of $\hat{\mathcal{P}}_{\mathbb{R}^\eta}$ under the null hypothesis can be obtained by means of simulations with $\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}} = \mathbf{I}_{4n}$, but the Wilks' theorem does not apply due to the inconsistency[4] of the η estimates [56].

The derivation of the small sample distributions is much more complicated. However, focusing on the case of known η, and taking into account the equivalence between the $\mathbb{C}^\eta$-properness test and the GLRT for testing the improperness of complex random vectors [32], [48], [54], [55], we can directly apply the results for the complex case [33], [39]. Furthermore, it can be proved that the $\mathbb{C}^\eta$-properness versus $\mathbb{Q}$-properness GLRT is equivalent to the problem of "testing the hypothesis that a covariance matrix with complex structure has quaternion structure," which was studied by Anderson *et al.* in [39]. Thus, using the results in [39] for the moments (of order r) of the test statistics under the null hypothesis, we have

$$E|\hat{\mathbf{\Phi}}_{\mathbb{R}^\eta}|^r = \prod_{j=1}^{n} \frac{\Gamma(T-j+3/2)\Gamma(2r+T-n-j+1)}{\Gamma(T-n-j+1)\Gamma(2r+T-j+3/2)} \tag{71}$$

for $T \geq 2n$, and

$$E|\hat{\mathbf{\Phi}}_{\mathbb{C}^\eta}|^r = \prod_{j=1}^{2n} \frac{\Gamma\left(\frac{T-j}{2}+1\right)\Gamma\left(r+\frac{T-j+1}{2}-n\right)}{\Gamma\left(\frac{T-j+1}{2}-n\right)\Gamma\left(r+\frac{T-j}{2}+1\right)} \tag{72}$$

for $T \geq 4n$, where $\Gamma(\cdot)$ is the gamma function. Now, following the lines in [33] for the application of the Box's approximation method [57], we conclude that under the null hypothesis

$$2(T-n+1/4)\hat{\mathcal{P}}_{\mathbb{R}^\eta} \approx \chi^2_{d_{\mathbb{R}^\eta}} \quad 2(T-2n)\hat{\mathcal{P}}_{\mathbb{C}^\eta} \approx \chi^2_{d_{\mathbb{C}^\eta}} \tag{73}$$

where $\approx$ denotes approximated distribution. Moreover, taking into account that $\hat{\mathbf{D}}_{\mathbb{C}^\eta}$ and $|\hat{\mathbf{\Phi}}_{\mathbb{C}^\eta}|$ are independent under the null hypothesis [39], it is clear that $|\hat{\mathbf{\Phi}}_{\mathbb{R}^\eta}|$ and $|\hat{\mathbf{\Phi}}_{\mathbb{C}^\eta}|$ are also independent, and therefore the rth moments of $|\hat{\mathbf{\Phi}}_{\mathbb{Q}}|$ are given (for $T \geq 4n$) by the product of (71) and (72), which can be exploited to approximate the null distribution of $\hat{\mathcal{P}}_{\mathbb{Q}}$. Here, we must note that, for $T \to \infty$, these approximations become equivalent to the Wilks' approach in (69).

[4]Note that under the null ($\mathbb{Q}$-proper) hypothesis, we cannot define a *true* value of the principal $\mathbb{C}$-properness direction η.

Finally, the distributions of the test statistics under the alternative (improper) hypothesis are not easy to obtain. There are some available results in the complex case [35], [49], but in general we can only say that, for $T \gg 1$ and *true* improperness measures $(\mathcal{P}_{\mathbb{R}^\eta}, \mathcal{P}_{\mathbb{C}^\eta}, \mathcal{P}_{\mathbb{Q}})$ close to zero, the test statistics are approximately distributed as [58]–[60]

$$2T\hat{\mathcal{P}}_{\mathbb{R}^\eta} \approx \chi^2_{d_{\mathbb{R}^\eta}}(\mu_{\mathbb{R}^\eta}) \tag{74}$$

$$2T\hat{\mathcal{P}}_{\mathbb{C}^\eta} \approx \chi^2_{d_{\mathbb{C}^\eta}}(\mu_{\mathbb{C}^\eta}) \tag{75}$$

$$2T\hat{\mathcal{P}}_{\mathbb{Q}} \approx \chi^2_{d_{\mathbb{Q}}}(\mu_{\mathbb{Q}}) \tag{76}$$

where now $\chi_d^2(\mu)$ denotes the noncentral chi-square distribution with d degrees of freedom and noncentral parameter μ, and $\mu_{\mathbb{R}^\eta}, \mu_{\mathbb{C}^\eta}, \mu_{\mathbb{Q}}$ depend on the actual distribution [58]–[60].

B. Particularization to the Scalar Case

The GLRT statistics provide additional insights in the scalar case $x = r_1 + \eta r_\eta + \eta' r_{\eta'} + \eta'' r_{\eta''}$. Specifically, defining the real vector $\mathbf{r} = [r_1, r_\eta, r_{\eta'}, r_{\eta''}]^T$ and the unitary matrix

$$\mathbf{T} = \frac{1}{2} \begin{bmatrix} +1 & +\eta & +\eta' & +\eta'' \\ +1 & +\eta & -\eta' & -\eta'' \\ +1 & -\eta & +\eta' & -\eta'' \\ +1 & -\eta & -\eta' & +\eta'' \end{bmatrix} \tag{77}$$

we can write $\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}} = 4\mathbf{T}\hat{\mathbf{R}}_{\mathbf{r},\mathbf{r}}\mathbf{T}^{\mathrm{H}}$, $\hat{\mathbf{D}}_{\mathbb{Q}} = \mathrm{Tr}(\hat{\mathbf{R}}_{\mathbf{r},\mathbf{r}})\mathbf{I}_4$, and

$$\hat{\mathcal{P}}_{\mathbb{Q}} = -2 \ln \frac{|\hat{\mathbf{R}}_{\mathbf{r},\mathbf{r}}|^{1/4}}{\mathrm{Tr}(\hat{\mathbf{R}}_{\mathbf{r},\mathbf{r}})/4} \tag{78}$$

which is the GLRT statistic for the well-known sphericity test of the real vector $\mathbf{r}$ [61]. Furthermore, using the Cayley–Dickson representation $x = a_1 + \eta'' a_2$, and defining the vector $\mathbf{a} = [a_1, a_2]^{\mathrm{T}}$, we obtain [25]

$$\hat{\mathcal{P}}_{\mathbb{C}^\eta} = -\frac{1}{2} \ln |\hat{\mathbf{\Phi}}_{\bar{\mathbf{a}}}| \quad \hat{\mathcal{P}}_{\mathbb{R}^\eta} = -2 \ln \frac{|\hat{\mathbf{R}}_{\mathbf{a},\mathbf{a}}|^{1/2}}{\mathrm{Tr}(\hat{\mathbf{R}}_{\mathbf{a},\mathbf{a}})/2}. \tag{79}$$

That is, as previously pointed out, the $\mathbb{C}^\eta$-properness GLRT reduces to testing the complex properness of $\mathbf{a}$, and the $\mathbb{R}^\eta$-properness GLRT is the sphericity test for the complex vector $\mathbf{a}$.

Finally, we must note that in the scalar case, the cost function in (59) reduces to a quadratic function, and the optimization problem can be solved in closed form. Thus, the principal $\mathbb{C}$-properness direction is obtained after the first iteration of the proposed successive convex approximations algorithm.

C. Multiple-Hypotheses Testing Problem

As we have previously shown, the $\mathbb{R}^\eta$-properness is all what a $\mathbb{C}^\eta$-proper vector needs to become $\mathbb{Q}$-proper, which is confirmed by the relationship $\hat{\mathcal{P}}_{\mathbb{Q}} = \hat{\mathcal{P}}_{\mathbb{C}^\eta} + \hat{\mathcal{P}}_{\mathbb{R}^\eta}$. Interestingly, this fact can also be used to easily relate the three proposed GLRTs. In particular, consider the classification problem with hypotheses $\mathcal{H}_{\mathbb{Q}}, \mathcal{H}_{\mathbb{C}^\eta}$, and $\mathcal{H}_{\mathcal{I}}$, assume as true the ML estimates of the augmented covariance matrices, and assign some *a priori* probabilities $P_{\mathcal{H}_{\mathbb{Q}}}, P_{\mathcal{H}_{\mathbb{C}^\eta}}, P_{\mathcal{H}_{\mathcal{I}}}$ satisfying

$$P_{\mathcal{H}_{\mathbb{Q}}} + P_{\mathcal{H}_{\mathbb{C}^\eta}} + P_{\mathcal{H}_{\mathcal{I}}} = 1, \qquad P_{\mathcal{H}_{\mathbb{Q}}} > P_{\mathcal{H}_{\mathbb{C}^\eta}} > P_{\mathcal{H}_{\mathcal{I}}}. \tag{80}$$

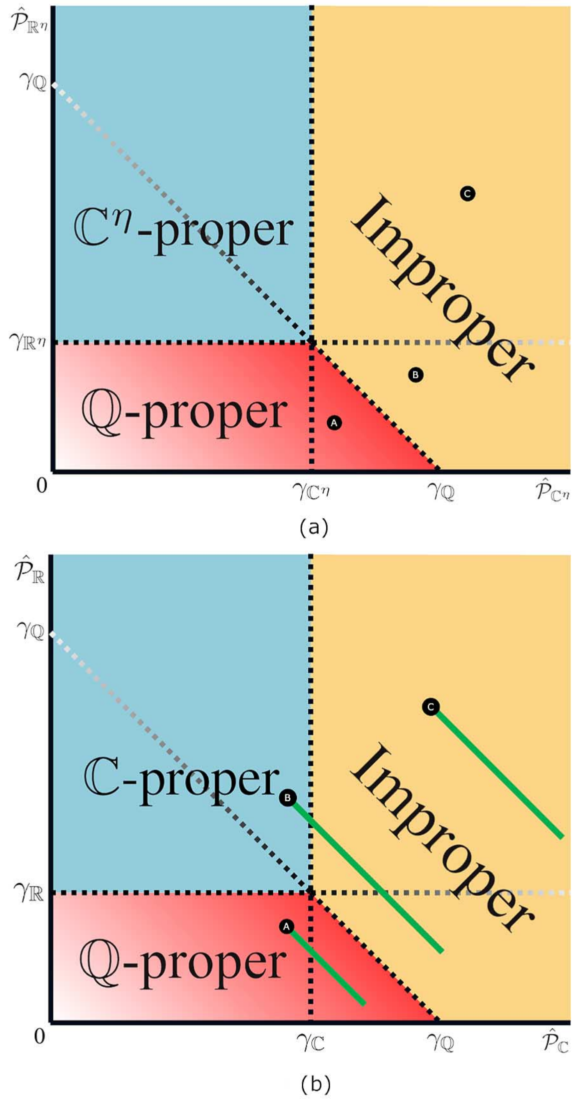

Fig. 1. Space of values for the two improperness measures (GLRT statistics) $\hat{\mathcal{P}}_{\mathbb{C}^\eta}$ and $\hat{\mathcal{P}}_{\mathbb{R}^\eta}$ in the case $\gamma_{\mathbb{Q}} = \gamma_{\mathbb{C}^\eta} + \gamma_{\mathbb{R}^\eta}$, which divides the space into three regions corresponding to the different properness hypotheses. (a) Scenario with *a priori* knowledge of the $\mathbb{C}$-properness direction η. (b) Scenario with unknown η. The solid lines represent the achievable pairs of values for different values of η.

Then, it is easy to prove that the maximum *a posteriori*[5] (MAP) classification rule would be based on the thresholds

$$\gamma_{\mathbb{C}^\eta} = \frac{1}{T}\ln\left(\frac{P_{\mathcal{H}_{\mathbb{C}^\eta}}}{P_{\mathcal{H}_{\mathcal{I}}}}\right) \quad \gamma_{\mathbb{R}^\eta} = \frac{1}{T}\ln\left(\frac{P_{\mathcal{H}_{\mathbb{Q}}}}{P_{\mathcal{H}_{\mathbb{C}^\eta}}}\right) \tag{81}$$

and more importantly

$$\gamma_{\mathbb{Q}} = \frac{1}{T}\ln\left(\frac{P_{\mathcal{H}_{\mathbb{Q}}}}{P_{\mathcal{H}_{\mathcal{I}}}}\right) = \gamma_{\mathbb{C}^\eta} + \gamma_{\mathbb{R}^\eta}. \tag{82}$$

As an example, Fig. 1(a) shows the space of values of the pair $(\hat{\mathcal{P}}_{\mathbb{C}^\eta}, \hat{\mathcal{P}}_{\mathbb{R}^\eta})$, as well as the regions associated with a particular choice of the thresholds $\gamma_{\mathbb{C}^\eta}$, $\gamma_{\mathbb{R}^\eta}$, and $\gamma_{\mathbb{Q}} = \gamma_{\mathbb{C}^\eta} + \gamma_{\mathbb{R}^\eta}$. Additionally, Fig. 1(b) represents the case without previous knowledge of the principal $\mathbb{C}$-properness direction. As can be seen, the optimization in η moves the points toward the upper left corner, and in particular, point B moves from the improperness region to the $\mathbb{C}^\nu$-properness region (for some pure unit quaternion ν).

[5]Note that this is not a true MAP classification technique because we are directly plugging the ML estimates of the augmented covariance matrices.

TABLE I
SECOND-ORDER STATISTICS FOR THE SIMULATION EXPERIMENTS

	$\mathbf{R}_{\mathbf{x},\mathbf{x}}$	$\mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta)}}$	$\mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta')}}$	$\mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta'')}}$
$\mathcal{H}_{\mathcal{I}}$	$\mathbf{I}_4$	$\mathbf{\Lambda}_\eta$	$\mathbf{\Lambda}_{\eta'}$	$\mathbf{0}_{4\times 4}$
$\mathcal{H}_{\mathbb{C}^\eta}$	$\mathbf{I}_4$	$\mathbf{\Lambda}_\eta$	$\mathbf{0}_{4\times 4}$	$\mathbf{0}_{4\times 4}$
$\mathcal{H}_{\mathbb{Q}}$	$\mathbf{I}_4$	$\mathbf{0}_{4\times 4}$	$\mathbf{0}_{4\times 4}$	$\mathbf{0}_{4\times 4}$

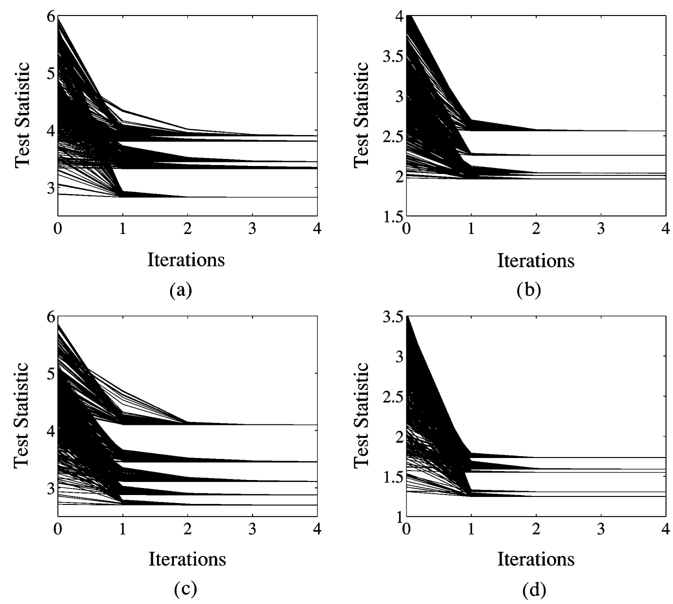

Fig. 2. Convergence example. Evolution of the test statistic $\hat{\mathcal{P}}_{\mathbb{C}}$ for five independent experiments, each one with 100 different initialization points. (a) $\mathcal{H}_{\mathcal{I}}$ and $T = 20$. (b) $\mathcal{H}_{\mathcal{I}}$ and $T = 30$. (c) $\mathcal{H}_{\mathbb{C}^\eta}$ and $T = 20$. (d) $\mathcal{H}_{\mathbb{C}^\eta}$ and $T = 30$.

VI. SIMULATION RESULTS

In this section, the performance of the proposed GLRTs is illustrated by means of some simulation results, which have been obtained using the MATLAB quaternion Toolbox [62]. Unless otherwise stated, the experiments are based on T i.i.d realizations of a four-dimensional quaternion Gaussian vector (i.e., $n = 4$) with zero mean and second-order statistics as illustrated in Table I, where the diagonal matrices $\mathbf{\Lambda}_\eta$ and $\mathbf{\Lambda}_{\eta'}$ are

$$\mathbf{\Lambda}_\eta = \frac{1}{4}\begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad \mathbf{\Lambda}_{\eta'} = \frac{1}{8}\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}.$$

Thus, with the specified SOS, the principal $\mathbb{C}$-properness direction (under $\mathcal{H}_{\mathbb{C}^\eta}$ and $\mathcal{H}_{\mathcal{I}}$) is η.

A. Convergence of the Successive Convex Approximations Method

The first set of examples illustrate the convergence of the successive convex approximations method. Specifically, Fig. 2

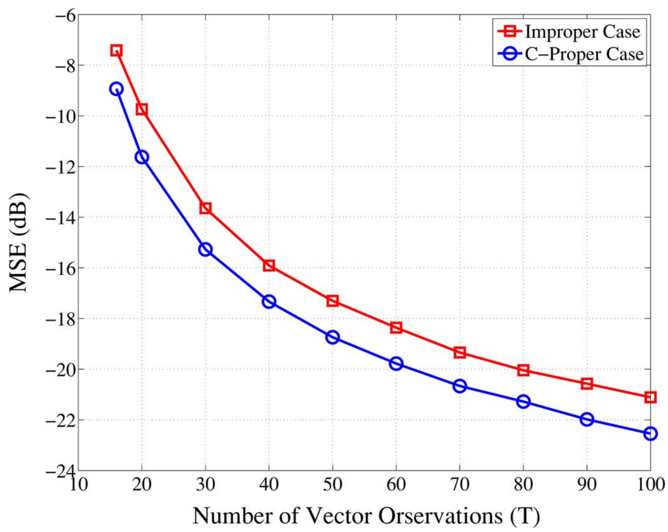

Fig. 3. Mean square error in the estimates of the principal $\mathbb{C}$-properness direction η after four iterations of the proposed algorithm.

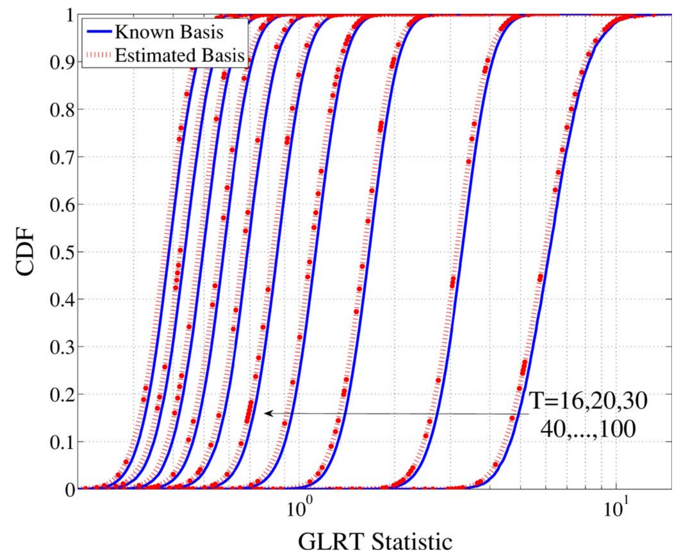

Fig. 5. Cumulative distribution function of the test statistic $\hat{\mathcal{P}}_{\mathbb{C}^\eta}$ under the null hypothesis $\mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta')}} = \mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta'')}} = \mathbf{0}_{n\times n}$ for $n = 4$.

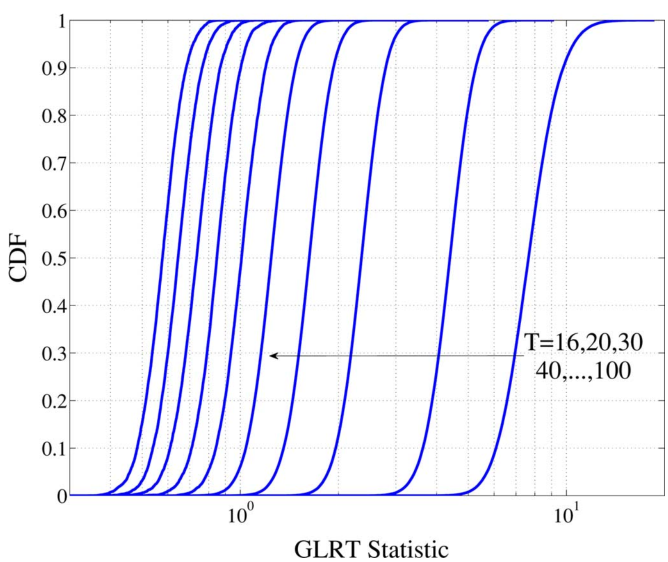

Fig. 4. Cumulative distribution function of the test statistic $\hat{\mathcal{P}}_{\mathbb{Q}}$ under the null hypothesis $\mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta)}} = \mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta')}} = \mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta'')}} = \mathbf{0}_{n\times n}$ for $n = 4$.

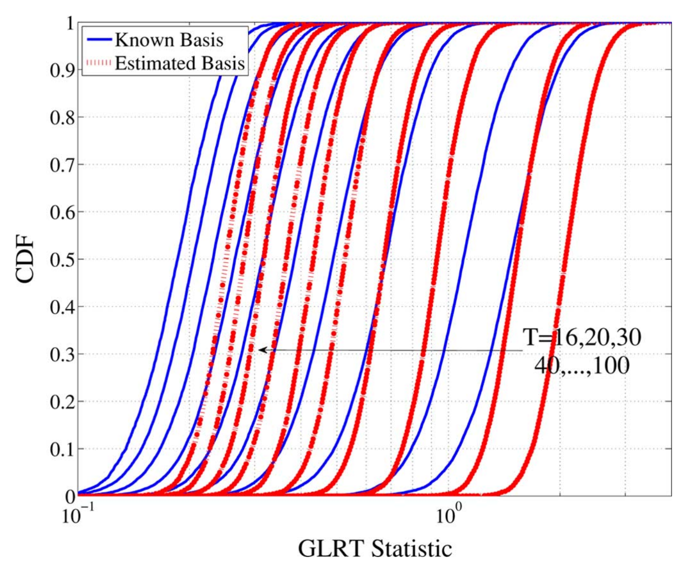

Fig. 6. Cumulative distribution function of the test statistic $\hat{\mathcal{P}}_{\mathbb{R}^\eta}$ under the null hypothesis $\mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta)}} = \mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta')}} = \mathbf{R}_{\mathbf{x},\mathbf{x}^{(\eta'')}} = \mathbf{0}_{n\times n}$ for $n = 4$.

shows the evolution of the $\mathbb{C}$-improperness measure $\hat{\mathcal{P}}_{\mathbb{C}}$ in four different scenarios, considering five independent examples for each scenario. In all the cases, the proposed algorithm has been initialized in 100 randomly generated values of η and, after a few iterations, we can see that the algorithm converges to the same solution. Based on these and other similar results, we have limited the proposed algorithm to four iterations. Finally, Fig. 3 shows the mean square error (MSE) in the estimate of the principal $\mathbb{C}$-properness direction, where we can see that the proposed algorithm provides reliable estimates both in the case of $\mathbb{C}$-proper and improper random vectors.

B. Cumulative Distribution Function of the GLRT Statistics Under the Null-Hypothesis

As we have previously pointed out, the invariances of the test statistics $\hat{\mathcal{P}}_{\mathbb{Q}}$, $\hat{\mathcal{P}}_{\mathbb{C}^\eta}$, and $\hat{\mathcal{P}}_{\mathbb{R}^\eta}$ can be exploited for obtaining their distribution under the null (proper) hypothesis by means of simulations. Figs. 4–6 show the numerically obtained cumulative distribution functions (CDFs) of the three test statistics, both in the case of known and unknown $\mathbb{C}$-properness direction. Additionally, Tables II–IV show the critical values of the thresholds for three different probabilities of false alarm P_{f} and several values of n and T. Finally, we must remember that the distribution of $\hat{\mathcal{P}}_{\mathbb{C}^\eta}$ for estimated η depends on the actual second-order statistics. Table III has been obtained using $\mathbf{R}_{\bar{\mathbf{x}},\bar{\mathbf{x}}} = \mathbf{I}_{4n}$ under the null distribution.

C. Receiver Operating Characteristic Curves

In this set of examples, we show the receiver operating characteristic (ROC) curves for the three proposed GLRTs. Specifically, Figs. 7–9 show the probability of miss (probability of accepting the null hypothesis when the alternative is true) as

TABLE II
CRITICAL VALUES $\gamma_{\mathbb{Q}}$ FOR THE $\mathbb{Q}$-PROPERNESS TEST

n	P_f	$T=40$	$T=60$	$T=80$	$T=100$
1	0.1	0.1917	0.1261	0.0935	0.0745
	0.05	0.2209	0.1459	0.1077	0.0860
	0.01	0.2824	0.1877	0.1379	0.1099
2	0.1	0.5480	0.3552	0.2623	0.2082
	0.05	0.5955	0.3861	0.2852	0.2265
	0.01	0.6942	0.4485	0.3313	0.2634
4	0.1	1.9308	1.1969	0.8693	0.6821
	0.05	2.0236	1.2540	0.9119	0.7145
	0.01	2.2028	1.3641	0.9956	0.7804
8	0.1	9.1629	4.8864	3.3752	2.5852
	0.05	9.4078	5.0073	3.4610	2.6508
	0.01	9.8833	5.2458	3.6201	2.7752

TABLE III
CRITICAL VALUES $\gamma_{\mathbb{C}}$ FOR THE $\mathbb{C}$-PROPERNESS TEST

n	P_f	η	$T=40$	$T=60$	$T=80$	$T=100$
1	0.1	known	0.1401	0.0918	0.0685	0.0545
		estimated	0.0596	0.0386	0.0288	0.0226
	0.05	known	0.1653	0.1087	0.0808	0.0644
		estimated	0.0732	0.0474	0.0352	0.0277
	0.01	known	0.2208	0.1451	0.1081	0.0860
		estimated	0.1043	0.0674	0.0497	0.0390
2	0.1	known	0.3972	0.2545	0.1870	0.1481
		estimated	0.2685	0.1711	0.1250	0.0988
	0.05	known	0.4380	0.2811	0.2065	0.1632
		estimated	0.2976	0.1899	0.1389	0.1094
	0.01	known	0.5232	0.3360	0.2470	0.1952
		estimated	0.3585	0.2284	0.1669	0.1319
4	0.1	known	1.3903	0.8473	0.6109	0.4775
		estimated	1.1615	0.7022	0.5030	0.3920
	0.05	known	1.4704	0.8972	0.6457	0.5045
		estimated	1.2301	0.7411	0.5304	0.4139
	0.01	known	1.6265	0.9932	0.7173	0.5584
		estimated	1.3661	0.8203	0.5854	0.4573
8	0.1	known	6.9536	3.5251	2.3907	1.8118
		estimated	6.5083	3.2444	2.1858	1.6500
	0.05	known	7.1785	3.6348	2.4621	1.8678
		estimated	6.7188	3.3403	2.2507	1.6985
	0.01	known	7.6232	3.8446	2.6028	1.9770
		estimated	7.1544	3.5094	2.3896	1.7895

TABLE IV
CRITICAL VALUES $\gamma_{\mathbb{R}}$ FOR THE $\mathbb{C}$- VERSUS $\mathbb{Q}$-PROPERNESS TEST

n	P_f	η	$T=40$	$T=60$	$T=80$	$T=100$
1	0.1	known	0.0801	0.0530	0.0393	0.0317
		estimated	0.1449	0.0955	0.0712	0.0569
	0.05	known	0.0997	0.0661	0.0492	0.0396
		estimated	0.1688	0.1115	0.0830	0.0663
	0.01	known	0.1437	0.0960	0.0713	0.0577
		estimated	0.2218	0.1465	0.1092	0.0868
2	0.1	known	0.2087	0.1372	0.1021	0.0811
		estimated	0.3100	0.2034	0.1507	0.1198
	0.05	known	0.2392	0.1568	0.1169	0.0928
		estimated	0.3428	0.2245	0.1666	0.1325
	0.01	known	0.3038	0.2008	0.1481	0.1173
		estimated	0.4121	0.2691	0.1994	0.1588
4	0.1	known	0.6524	0.4205	0.3097	0.2455
		estimated	0.8297	0.5327	0.3915	0.3104
	0.05	known	0.7054	0.4538	0.3348	0.2649
		estimated	0.8814	0.5652	0.4162	0.3293
	0.01	known	0.8156	0.5230	0.3823	0.3036
		estimated	0.9899	0.6297	0.4650	0.3671
8	0.1	known	2.4678	1.5146	1.0923	0.8552
		estimated	2.8341	1.7309	1.2475	0.9754
	0.05	known	2.5762	1.5799	1.1391	0.8913
		estimated	2.9322	1.7903	1.2900	1.0086
	0.01	known	2.7831	1.7079	1.2295	0.9636
		estimated	3.1207	1.9070	1.3731	1.0746

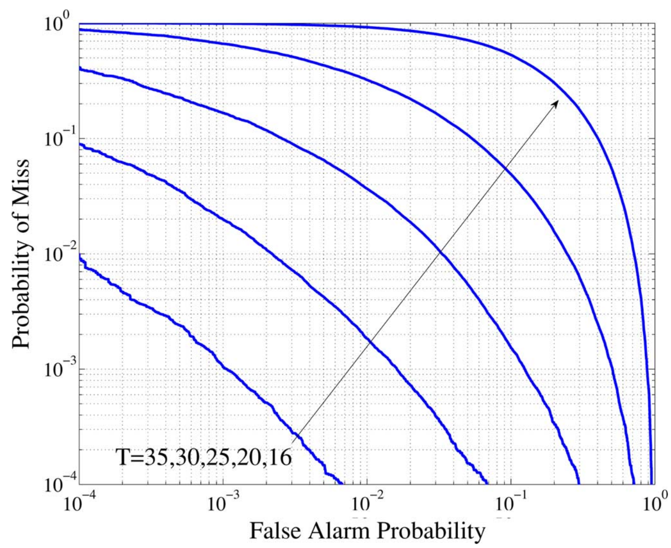

Fig. 7. Receiver operating characteristic. $\mathbb{Q}$-properness GLRT.

a function of the false alarm probability. Interestingly, the figures show that the third GLRT ($\mathbb{C}$-properness versus $\mathbb{Q}$-properness) is more affected by errors in the estimate of the principal $\mathbb{C}$-properness direction. As previously pointed out, this is due to the inconsistency of the estimates η under the $\mathbb{Q}$-proper hypothesis.

D. Practical Example

In the final example, we show a practical application of the derived GLRTs. In particular, we consider an optical communication system with dual polarization that, considering a single frequency, can be modeled as [36]–[38]

$$\mathbf{y} = \mathbf{U}\boldsymbol{\Gamma}_{\text{PDL}}\mathbf{V}\mathbf{x} + \mathbf{n} \tag{83}$$

where $\mathbf{x} = [x_1, x_2]^{\text{T}}$ (resp. $\mathbf{y} = [y_1, y_2]^{\text{T}}$) is a complex vector in the plane $\{1, \eta\}$ representing the transmitted (resp. received) signals in the two orthogonal principal states of polarization (PSP), $\mathbf{n} = [n_1, n_2]^{\text{T}}$ represents the i.i.d. circular complex Gaussian noise

$$\mathbf{U} = \begin{bmatrix} u_1 & -u_2^* \\ u_2 & u_1^* \end{bmatrix} \quad \mathbf{V} = \begin{bmatrix} v_1 & -v_2^* \\ v_2 & v_1^* \end{bmatrix} \tag{84}$$

are complex rotation matrices representing the PSP mismatch between the fiber and the transmitted signals, and

$$\boldsymbol{\Gamma}_{\text{PDL}} = \begin{bmatrix} \sqrt{1+\gamma_{\text{PDL}}} & 0 \\ 0 & \sqrt{1-\gamma_{\text{PDL}}} \end{bmatrix} \tag{85}$$

represents the polarization dependent losses (PDLs). In particular, the PDL factor is defined as $\text{PDL} = 10\log_{10}(1+\gamma_{\text{PDL}})/(1-\gamma_{\text{PDL}})$.

In the experiments, we consider a communication system transmitting QPSK symbols with signal-to-noise ratio (SNR) of

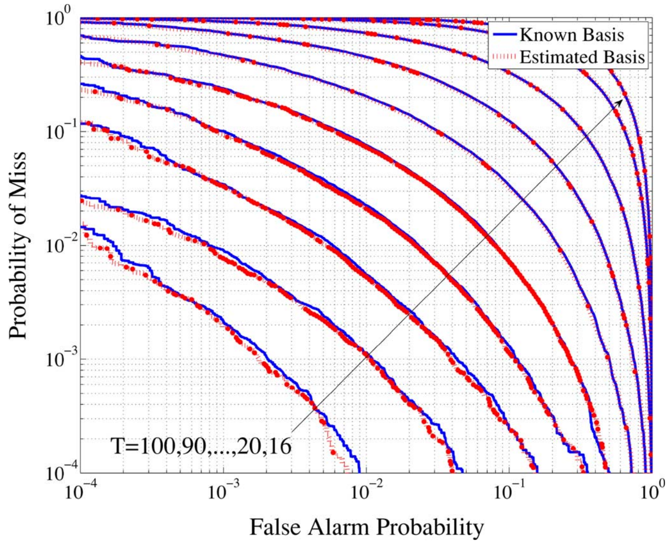

Fig. 8. Receiver operating characteristic. $\mathbb{C}^{\eta}$-properness GLRT.

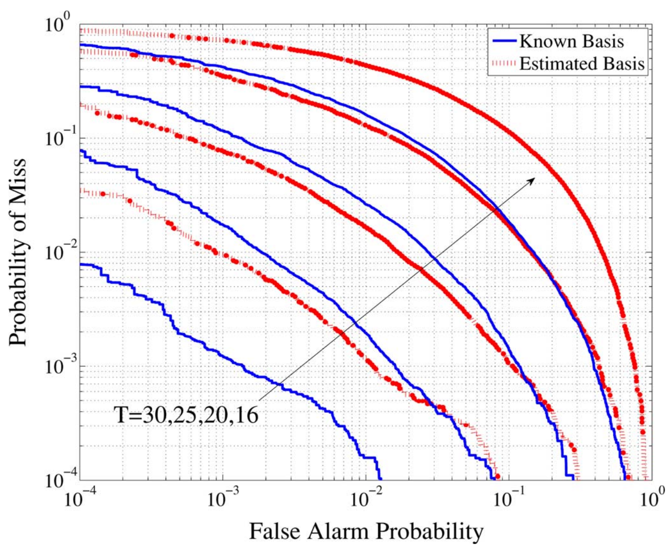

Fig. 9. Receiver operating characteristic. $\mathbb{C}^{\eta}$-properness versus $\mathbb{Q}$-properness GLRT.

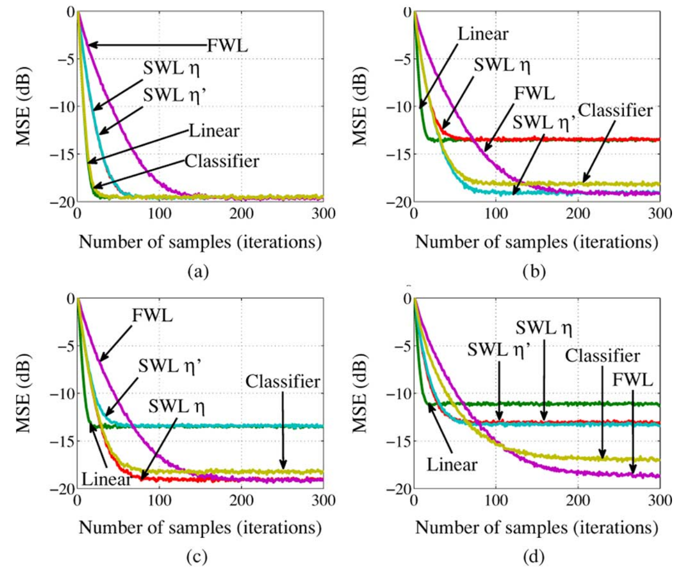

Fig. 10. Convergence of the quaternion LMS algorithm in the four different scenarios. The curves show the conventional linear LMS (Linear), the semi-widely linear LMSs (SWL η and SWL η'), the full-widely linear LMS (FWL), and the proposed technique (Classifier). (a) Balanced system. (b) System with IQ imbalance. (c) System with PDL. (d) System with PDL and IQ imbalance.

20 dB and random matrices $\mathbf{U}$ and $\mathbf{V}$. Furthermore, we consider a possible IQ imbalance in the transmitted signals x_1, x_2, where the IQ imbalance factor is defined as $\mathrm{IQ} = 10\log_{10}(\sigma_I^2)/(\sigma_Q^2)$, with σ_I^2 and σ_Q^2 representing the power in the *in-phase* and *quadrature* components of the transmitted signals. Thus, defining the quaternions $x = x_1 + x_2\eta'$ and $y = y_1 + y_2\eta'$, our goal consists in applying the linear, semi-widely linear, or full-widely linear quaternion LMS algorithm [9]–[11] for recovering x from y. In particular, the quaternion LMS updating rule can be written as

$$\mathbf{w} = \mathbf{w} + \mu(e\mathbf{z} - 2\mathbf{z}e^*) \tag{86}$$

where $\hat{x} = \mathbf{w}^{\mathrm{H}}\mathbf{z}$ is the estimate of x, $e = x - \hat{x}$ is the error, $\mathbf{w}$ is the system equalizer, and $\mathbf{z}$ is the input for the LMS algorithm. Here, we consider four different scenarios.

- Balanced system $(\mathrm{PDL} = \mathrm{IQ} = 0)$: In this case, the source quaternion x is $\mathbb{Q}$-proper, and the channel preserves the *quaternion structure*. Therefore, y is also $\mathbb{Q}$-proper, and the most appropriate processing is the conventional linear model. That is, we should select $\mathbf{z} = x$ for the quaternion LMS.
- System with IQ imbalance $(\mathrm{PDL} = 0, \mathrm{IQ} = 3)$: In this case, the source x and observation y are $\mathbb{C}^{\eta'}$-proper. The optimal linear processing is semi-widely linear in η', which means that we should select $\mathbf{z} = [x, x^{(\eta')}]^{\mathrm{T}}$.
- System with PDL $(\mathrm{PDL} = 3, \mathrm{IQ} = 0)$: Although the source quaternion is $\mathbb{Q}$-proper, the channel introduces a power imbalance, and the received quaternion y is $\mathbb{C}^{\eta}$-proper. Therefore, the most appropriate processing is semi-widely linear in η ($\mathbf{z} = [x, x^{(\eta)}]^{\mathrm{T}}$).
- System with PDL and IQ imbalance $(\mathrm{PDL} = \mathrm{IQ} = 3)$: The source is $\mathbb{C}^{\eta'}$-proper, but y is improper. Furthermore, the principal $\mathbb{C}$-properness direction and the $\mathbb{C}$-improperness measure depend on the particular value of $\mathbf{U}$ and $\mathbf{V}$. In general, the optimal linear processing is full-widely linear ($\mathbf{z} = [x, x^{(\eta)}, x^{(\eta')}, x^{(\eta'')}]^{\mathrm{T}}$).

The proposed GLRTs, including the principal $\mathbb{C}$-properness direction algorithm, are applied to a set of $T = 100$ observations for solving the multiple-hypotheses testing problem and selecting the most convenient kind of processing. In case of selecting the semi-widely linear processing, we use the estimated principal $\mathbb{C}$-properness direction ν, i.e., we use $\mathbf{z} = [x, x^{(\nu)}]^{\mathrm{T}}$ as input for the LMS algorithm. In order to achieve the same steady-state error in the $\mathbb{Q}$-proper case, the learning rate has been selected as $\mu = 0.1$ in the case of conventional linear processing, $\mu = 0.05$ for semi-widely linear processing, and $\mu = 0.025$ for full-widely linear processing. The results, averaged for 1000 independent simulations, are shown in Fig. 10. As can be seen, the performance of the proposed technique is always very close to that of the most convenient kind of processing, and the small deviations in the $\mathbb{C}$-proper and improper scenarios are due to classification mistakes and residual errors in the estimate of the principal $\mathbb{C}$-properness direction.

VII. Conclusion

In this paper, we have presented three generalized likelihood ratio tests (GLRTs) for testing the properness of a quaternion random vector. This is an important problem because the type of quaternion properness will determine the required kind of linear processing (*full-widely* linear, *semi-widely* linear, or *conventional* linear processing). The proposed tests have been derived under the Gaussian assumption, and they reduce to the estimation, and comparison to a fixed threshold, of three previously proposed improperness measures. Additionally, we have presented an algorithm for the estimation of the principal $\mathbb{C}$-properness direction, or equivalently, to decompose the quaternion vector into two complex vectors with the lowest improperness degree. The proposed technique is based on the successive convex approximations method, which guarantees the convergence to a solution satisfying the Karush–Kuhn–Tucker conditions. Finally, the performance and practical application of the proposed techniques have been illustrated by means of several simulation examples.

References

[1] L. Fortuna, G. Muscato, and M. Xibilia, "A comparison between HMLP and HRBF for attitude control," *IEEE Trans. Neural Netw.*, vol. 12, no. 2, pp. 318–328, Mar. 2001.

[2] J. B. Kuipers, *Quaternions and Rotation Sequences: A Primer With Applications to Orbits, Aerospace and Virtual Reality*. Princeton, NJ: Princeton Univ. Press, Aug. 2002.

[3] A. J. Hanson, *Visualizing Quaternions*. San Mateo, CA: Morgan Kaufmann, Dec. 2005.

[4] S.-C. Pei and C.-M. Cheng, "Color image processing by using binary quaternion-moment-preserving thresholding technique," *IEEE Trans. Image Process.*, vol. 8, no. 5, pp. 614–628, May 1999.

[5] C. Moxey, S. Sangwine, and T. Ell, "Hypercomplex correlation techniques for vector images," *IEEE Trans. Signal Process.*, vol. 51, no. 7, pp. 1941–1953, Jul. 2003.

[6] T. Bulow and G. Sommer, "Hypercomplex signals—A novel extension of the analytic signal to the multidimensional case," *IEEE Trans. Signal Process.*, vol. 49, no. 11, pp. 2844–2852, Nov. 2001.

[7] M. Felsberg and G. Sommer, "The monogenic signal," *IEEE Trans. Signal Process.*, vol. 49, no. 12, pp. 3136–3144, Dec. 2001.

[8] T. Ell and S. Sangwine, "Hypercomplex Fourier transforms of color images," *IEEE Trans. Image Process.*, vol. 16, no. 1, pp. 22–35, Jan. 2007.

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

[10] C. C. Took and D. P. Mandic, "A quaternion widely linear adaptive filter," *IEEE Trans. Signal Process.*, vol. 58, no. 8, pp. 4427–4431, Aug. 2010.

[11] C. C. Took and D. P. Mandic, "Quaternion-valued stochastic gradient-based adaptive IIR filtering," *IEEE Trans. Signal Process.*, vol. 58, no. 7, pp. 3895–3901, Jul. 2010.

[12] N. ur Rehman and D. P. Mandic, "Empirical mode decomposition for trivariate signals," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1059–-1068, Mar. 2010.

[13] 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.

[14] 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.

[15] 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.

[16] J.-C. Belfiore and G. Rekaya, "Quaternionic lattices for space-time coding," in *Proc. IEEE ITW*, 2003, pp. 267–270.

[17] J.-C. Belfiore, G. Rekaya, and E. Viterbo, "The golden code: A 2×2 full-rate space-time code with non-vanishing determinant," *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1432–1436, Apr. 2005.

[18] M.-Y. Chen, H.-C. Li, and S.-C. Pei, "Algebraic identification for optimal nonorthogonality 4×4 complex space-time block codes using tensor product on quaternions," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 324–330, Jan. 2005.

[19] C. Hollanti, J. Lahtonen, K. Ranto, R. Vehkalahti, and E. Viterbo, "On the algebraic structure of the silver code: A 2×2 perfect space-time block code," in *Proc. IEEE ITW*, May 2008, pp. 91–94.

[20] S. Sirianunpiboon, A. Calderbank, and S. Howard, "Bayesian analysis of interference cancellation for Alamouti multiplexing," *IEEE Trans. Inf. Theory*, vol. 54, no. 10, pp. 4755–4761, Oct. 2008.

[21] N. N. Vakhania, "Random vectors with values in quaternion Hilbert spaces," *Theory Probab. Appl.*, vol. 43, no. 1, pp. 99–115, 1999.

[22] P. Amblard and N. Le Bihan, "On properness of quaternion valued random variables," in *Proc. IMA Conf. Math. Signal Process.*, Cirencester, U.K., 2004, pp. 23–26.

[23] N. Le Bihan and P. O. Amblard, "Detection and estimation of Gaussian proper quaternion valued random processes," in *Proc. 7th IMA Conf. Math. Signal Process.*, Cirencester, U.K., Dec. 2006, pp. 23–26.

[24] N. Le Bihan and S. Buchholz, "Optimal separation of polarized signals by quaternionic neural networks," in *Proc. XIV EUSIPCO*, Florence, Italy, 2006.

[25] J. Vía, D. Ramírez, and I. Santamaría, "Properness and widely linear processing of quaternion random vectors," *IEEE Trans. Inf. Theory*, vol. 56, no. 7, pp. 3502–3515, Jul. 2010.

[26] F. Neeser and J. Massey, "Proper complex random processes with applications to information theory," *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1293–1302, Jul. 1993.

[27] B. Picinbono, "On circularity," *IEEE Trans. Signal Process.*, vol. 42, no. 12, pp. 3473–3482, Dec. 1994.

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

[29] A. van den Bos, "The multivariate complex normal distribution—A generalization," *IEEE Trans. Inf. Theory*, vol. 41, no. 2, pp. 537–539, Mar. 1995.

[30] 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.

[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] 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.

[33] 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.

[34] 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.

[35] J.-P. Delmas, A. Oukaci, and P. Chevalier, "Asymptotic distribution of GLR for impropriety of complex signals," in *Proc. IEEE ICASSP*, Mar. 2010, pp. 3594–3597.

[36] M. Karlsson and M. Petersson, "Quaternion approach to PMD and PDL phenomena in optical fiber systems," *J. Lightw. Technol.*, vol. 22, no. 4, pp. 1137–1146, Apr. 2004.

[37] D. Barros and J. Kahn, "Optimized dispersion compensation using orthogonal frequency-division multiplexing," *J. Lightw. Technol.*, vol. 26, no. 16, pp. 2889–2898, Aug. 2008.

[38] S. Mumtaz, G.-B. Othman, and Y. Jaouen, "Space-time codes for optical fiber communication with polarization multiplexing," in *Proc. IEEE ICC*, May 2010, pp. 1–5.

[39] B. H. Andersson, S. A. , and S. T. Jensen, "Distribution of eigenvalues in multivariate statistical analysis," *Ann. Statist.*, vol. 11, pp. 392–415, 1983.

[40] M. Chiang, C. W. Tan, D. Palomar, D. O'Neill, and D. Julian, "Power control by geometric programming," *IEEE Trans. Wireless Commun.*, vol. 6, no. 7, pp. 2640–2651, Jul. 2007.

[41] B. R. Marks and G. P. Wright, "A general inner approximation algorithm for nonconvex mathematical programs," *Oper. Res.*, vol. 26, no. 4, pp. 681–683, 1978.

[42] M. Avriel, *Advances in Geometric Programming*, M. Avriel, Ed. New York: Plenum, 1980.

[43] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, Mar. 2004.

[44] W. R. Hamilton, "On quaternions," in *Proc. Roy. Irish Acad.*, Nov. 11, 1844.

[45] J. P. Ward, *Quaternions and Cayley Numbers: Algebra and Applications*. Dordrecht, The Netherlands: Kluwer, 1997.

[46] B. Picinbono, "Second-order complex random vectors and normal distributions," *IEEE Trans. Signal Process.*, vol. 44, no. 10, pp. 2637–2640, Oct. 1996.

[47] H. S. M. Coxeter, "Quaternions and reflections," *Amer. Math. Monthly*, vol. 53, no. 3, pp. 136–146, 1946.

[48] E. Ollila and V. Koivunen, "Generalized complex elliptical distributions," in *Proc. IEEE SAM*, Jul. 2004, pp. 460–464.

[49] J. P. Delmas and H. Abeida, "Asymptotic distribution of circularity coefficients estimate of complex random variables," *Signal Process.*, vol. 89, no. 12, pp. 2670–2675, 2009.

[50] T. Cover and J. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[51] J. R. Kettenring, "Canonical analysis of several sets of variables," *Biometrika*, vol. 58, no. 3, pp. 433–451, 1971.

[52] 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.

[53] 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.

[54] E. Ollila and V. Koivunen, "Adjusting the generalized likelihood ratio test of circularity robust to non-normality," in *Proc. 10th IEEE SPAWC*, Jun. 2009, pp. 558–562.

[55] M. Novey, T. Adalı, and A. Roy, "Circularity and Gaussianity detection using the complex generalized Gaussian distribution," *IEEE Signal Process. Lett.*, vol. 16, no. 11, pp. 993–996, Nov. 2009.

[56] G. A. Young and R. L. Smith, *Essentials of Statistical Inference*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[57] G. E. P. Box, "A general distribution theory for a class of likelihood criteria," *Biometrika*, vol. 36, no. 3/4, pp. 317–346, Dec. 1949.

[58] D. Cox and D. Hinckley, *Theoretical Statistics*. New York: Chapman & Hall, 1974.

[59] S. Kay, "Asymptotically optimal detection in incompletely characterized non-Gaussian noise," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 5, pp. 627–633, May 1989.

[60] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume 2: Detection Theory*. Upper Saddle River, NJ: Prentice-Hall, 1993.

[61] J. Mauchly, "Significance test for sphericity of a normal n-variate distribution," *Ann. Math. Statist.*, vol. 11, pp. 204–209, 1940.

[62] S. J. Sangwine and N. Le Bihan, "Quaternion toolbox for MATLAB," 2005 [Online]. Available: http://qtfm.sourceforge.net/

Javier Vía (S'04–M'08) received the Telecommunication Engineer Degree and Ph.D. degree in electrical engineering from the University of Cantabria, Santander, Spain, in 2002 and 2007, respectively.

In 2002, he joined the Department of Communications Engineering, University of Cantabria, where he is currently an Assistant Professor. He has spent visiting periods with the Smart Antennas Research Group, Stanford University, Stanford, CA, and with the Department of Electronics and Computer Engineering, Hong Kong University of Science and Technology, Hong Kong. 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.

Daniel P. Palomar (S'99–M'03–SM'08) received the Electrical Engineering and Ph.D. degrees (both with honors) from the Technical University of Catalonia (UPC), Barcelona, Spain, in 1998 and 2003, respectively.

He is an Associate Professor with the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology (HKUST), Hong Kong, which he joined in 2006. He had previously held several research appointments, namely at King's College London (KCL), London, U.K.; UPC; Stanford University, Stanford, CA; the Telecommunications Technological Center of Catalonia (CTTC), Barcelona, Spain; the Royal Institute of Technology (KTH), Stockholm, Sweden; the University of Rome "La Sapienza," Rome, Italy; and Princeton University, Princeton, NJ. His current research interests include applications of convex optimization theory, game theory, and variational inequality theory to signal processing and communications.

Dr. Palomar is an Associate Editor of the IEEE TRANSACTIONS ON INFORMATION THEORY and has been an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING, a Guest Editor of the *IEEE Signal Processing Magazine* 2010 Special Issue on "Convex Optimization for Signal Processing," a Guest Editor of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS 2008 Special Issue on "Game Theory in Communication Systems," and the Lead Guest Editor of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS 2007 Special Issue on "Optimization of MIMO Transceivers for Realistic Communication Networks." He serves on the IEEE Signal Processing Society Technical Committee on Signal Processing for Communications (SPCOM). He was the General Co-Chair of the 2009 IEEE Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP). He is a recipient of a 2004–2006 Fulbright Research Fellowship; the 2004 Young Author Best Paper Award by the IEEE Signal Processing Society; the 2002–2003 best Ph.D. prize in Information Technologies and Communications by the UPC; the 2002–2003 Rosina Ribalta first prize for the Best Doctoral Thesis in Information Technologies and Communications by the Epson Foundation, and the 2004 prize for the Best Doctoral Thesis in Advanced Mobile Communications by the Vodafone Foundation and COIT.

Luis Vielva was born in Santander, Spain, in 1966. He received the Licenciado and Ph.D. degrees in physics from the University of Cantabria, Santander, Spain, in 1989 and 1997, respectively.

In 1989, he joined the Department of Communications Engineering, University of Cantabria, where he is currently an Associate Professor. His current research interests include systems biology, convex optimization, and geometric algebra.