Simultaneous diagonalisation of the covariance and complementary covariance matrices in quaternion widely linear signal processing

Min Xiang [a],[*], Shirin Enshaeifar [b], Alexander E. Stott [a], Clive Cheong Took [b], Yili Xia [c], Sithan Kanna [a], Danilo P. Mandic [a]

[a] *Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2BT, UK*
[b] *Department of Computer Science, University of Surrey, Surrey GU2 7XH, UK*
[c] *School of Information Science and Engineering, Southeast University, Nanjing 210018, China*

A B S T R A C T

Recent developments in quaternion-valued widely linear processing have established that the exploitation of complete second-order statistics requires consideration of both the standard covariance and the three complementary covariance matrices. Although such matrices have a tremendous amount of structure and their decomposition is a powerful tool in a variety of applications, the non-commutative nature of the quaternion product has been prohibitive to the development of quaternion uncorrelating transforms. To this end, we introduce novel techniques for a simultaneous decomposition of the covariance and complementary covariance matrices in the quaternion domain, whereby the quaternion version of the Takagi factorisation is explored to diagonalise symmetric quaternion-valued matrices. This gives new insights into the quaternion uncorrelating transform (QUT) and forms a basis for the proposed quaternion approximate uncorrelating transform (QAUT) which simultaneously diagonalises all four covariance matrices associated with improper quaternion signals. The effectiveness of the proposed uncorrelating transforms is validated by simulations on both synthetic and real-world quaternion-valued signals.

1. Introduction

Advances in sensing technology have enabled widespread recording from 3-D and 4-D data sources, such as measurements from seismometers [1], ultrasonic anemometers [2], and inertial body sensors [3]. It is convenient to express such measurements as vectors in the $\mathbb{R}^3$ and $\mathbb{R}^4$ fields of reals, however, the real vector algebra is not a division algebra and is therefore inadequate when modelling orientation and rotation [4]. Quaternions have advantages in representing 3-D and 4-D data, owing to their division algebra, a generic extension of the real and complex algebras. Quaternions also naturally account for mutual information between multiple data channels, provide a compact representation, and have proven to offer a physically meaningful interpretation for real-world applications in the fields of navigation, communication, and image processing [5,6]. A recent resurgence in research on quaternion signal processing spans the areas of filtering [7], independent component analysis (ICA) [8,9], neural networks [10,11], and Fourier transforms [12].

Diagonalisable matrices play a fundamental role in engineering applications, whereby the related uncorrelating transforms greatly simplify the analysis of complex problems, both in terms of enabling a solution and providing a methodological framework (e.g. performance bounds). Computational advantages associated with matrix diagonalisation include a reduction in the size of the parameter space from N^2 to N, while the statistical advantages include the possibility to match source properties (e.g. orthogonality) in blind source separation. In addition, joint diagonalisation of matrices in some blind source separation applications has a close link with the canonical decomposition problem for tensors [13,14]. The real linear algebra is a mature area, however, covariance matrices in widely linear signal processing [15] in the complex and quaternion division algebras and their structures [16] have recently received significant attention [17–21].

In the context of complex-valued signal processing, for a complex random vector $\mathbf{z}$, both its covariance, $\mathbf{C} = E\{\mathbf{z}\mathbf{z}^H\}$, and pseudo-covariance, $\mathbf{P} = E\{\mathbf{z}\mathbf{z}^T\}$, matrices are necessary to capture complete second-order statistical information [22]. To extract the statistical information embedded in the two covariance matrices, simultaneous diagonalisation of $\mathbf{C}$ and $\mathbf{P}$ is a prerequisite, while their respective individual diagonalisations can be performed via the eigendecomposition and the Takagi factorisation. Such decorrelations exist

* Corresponding author.
E-mail address: m.xiang13@ic.ac.uk (M. Xiang).

in their general forms in mathematics [23]. De Lathauwer and De Moor [15] and Eriksson and Koivunen [16] have given them a practical value through an important engineering contribution called the strong uncorrelating transform (SUT), together with its extension, the generalised uncorrelating transform [24]. Our own recent contributions include a computationally efficient approximate uncorrelating transform (AUT) [25], while to preserve the integrity of the original bivariate sources in the decorrelation process, we have proposed the correlation preserving transform [26]. These transforms have been used in performance analysis of complex-valued adaptive filters [27–30], blind separation of non-circular complex-valued sources [31,32], complex-valued subspace tracking [33], and separation of signal and noise components in subbands of harmonic signals [34].

Real- and complex-valued signal processing have benefited greatly from making sophisticated use of real- and complex-valued matrix algebras [35,36]. However, quaternion-valued signal processing, which is advantageous for 3-D and 4-D data, has been hindered by the underdevelopment of quaternion-valued matrix algebra [37], compared to the well-established real- and complex-valued matrix algebras [23,36]. This is due to the non-commutativity nature of quaternion multiplication and the enhanced degrees of freedom provided by quaternions. For example, a quaternion square matrix has right and left eigenvalues; the right eigenvalues have been well studied, while the left eigenvalues are less known and are not computationally well-posed [37,38]. The recent theory of quaternion matrix derivatives provides a systematic framework for the calculation of derivatives of quaternion matrix functions with respect to quaternion matrix variables [39]. Advances in structural quaternion matrix decompositions include tools to diagonalise quaternion matrices [40,41], while in the context of quaternion widely linear processing, it was shown that for a quaternion random vector $\mathbf{x}$, the diagonalisation of the Hermitian covariance matrix $\mathbf{C_x} = E\{\mathbf{xx}^H\}$ can be performed straightforwardly using the quaternion eigendecomposition, whereas the diagonalisation of the complementary covariance matrices ($\mathbf{C_{x^\iota}} = E\{\mathbf{xx}^{\iota H}\}$, $\mathbf{C_{x^J}} = E\{\mathbf{xx}^{JH}\}$, $\mathbf{C_{x^\kappa}} = E\{\mathbf{xx}^{\kappa H}\}$) can be computed via the quaternion singular value decomposition (SVD) [42]. Furthermore, the quaternion uncorrelating transform (QUT), a quaternion analogue of the SUT for complex data, was proposed in [43] to jointly diagonalise the covariance matrix and one of the three complementary covariance matrices. However, the simultaneous diagonalisation of all the four covariance matrices is still an open problem. It is important to notice that the relationship between the quaternion covariance matrices in the widely linear model is governed by Cheong Took and Mandic [21]

$$\mathbf{P_x} = \frac{1}{2}(\mathbf{C_{x^\iota}} + \mathbf{C_{x^J}} + \mathbf{C_{x^\kappa}} - \mathbf{C_x}) \tag{1}$$

which demonstrates that the diagonalisation of the pseudo-covariance matrix, $\mathbf{P_x} = E\{\mathbf{xx}^T\}$, requires us to address the following issues:

- There is no *closed-form* solution[1] to perform a simultaneous diagonalisation of the covariance matrix and the three complementary covariance matrices.
- Given the dimensionality of the problem and numerous practical applications, a simple approximate diagonalising transform which would apply to both the covariance and pseudo-covariance matrices would be beneficial.

In this paper, we therefore set out to propose solutions to these open problems, and support the analysis with illustrative examples.

The rest of this paper is organised as follows. Section 2 provides an overview of quaternion algebra and statistics. Section 3 proposes novel techniques for the diagonalisation of symmetric quaternion matrices and a simultaneous diagonalisation of two quaternion covariance matrices. Section 4 proposes an approximate simultaneous diagonalisation of the four quaternion covariance matrices. Simulation results are given in Section 5, and Section 6 concludes the paper. Throughout the paper, we use bold-face capital letters to denote matrices, $\mathbf{A}$, boldface lowercase letters for vectors, $\mathbf{a}$, and standard letters for scalar quantities, a. Superscripts $(\cdot)^T$, $(\cdot)^*$ and $(\cdot)^H$ denote the transpose, conjugate, and Hermitian (i.e. transpose and conjugate), respectively, $\mathbf{I}$ the identity matrix, and $E\{\cdot\}$ the statistical expectation operator.

2. Background

2.1. Quaternion algebra

The quaternion domain $\mathbb{H}$ is a 4-D vector space over the field of reals, spanned by the basis $\{1, \iota, J, \kappa\}$, where $\iota = (1, 0, 0)$, $J = (0, 1, 0)$ and $\kappa = (0, 0, 1)$ are three unit vectors in $\mathbb{R}^3$. A quaternion vector[2] $\mathbf{x} \in \mathbb{H}^{L \times 1}$ consists of a real (scalar) part $\Re[\cdot]$ and an imaginary (vector) part $\Im[\cdot]$ which comprises $\iota-$, $J-$, and $\kappa-$ imaginary components, so that

$$\begin{aligned} \mathbf{x} &= \Re[\mathbf{x}] + \Im[\mathbf{x}] \\ &= \Re[\mathbf{x}] + \Im_\iota[\mathbf{x}]\iota + \Im_J[\mathbf{x}]J + \Im_\kappa[\mathbf{x}]\kappa \\ &= \mathbf{x}_a + \mathbf{x}_b\iota + \mathbf{x}_c J + \mathbf{x}_d\kappa \end{aligned} \tag{2}$$

where $\Re[\mathbf{x}] = \mathbf{x}_a, \Im_\iota[\mathbf{x}] = \mathbf{x}_b, \Im_J[\mathbf{x}] = \mathbf{x}_c, \Im_\kappa[\mathbf{x}] = \mathbf{x}_d$ are $L \times 1$ real vectors, and ι, J, κ are orthogonal imaginary units with properties

$$\begin{aligned} \iota J = -J\iota = \kappa \qquad J\kappa = -\kappa J = \iota \qquad \kappa\iota = -\iota\kappa = J \\ \iota^2 = J^2 = \kappa^2 = -1 \end{aligned} \tag{3}$$

A quaternion variable x is called a pure quaternion if it satisfies $\Re[x] = 0$. The modulus of a quaternion variable $x \in \mathbb{H}$ is defined as

$$|x| = \sqrt{x_a^2 + x_b^2 + x_c^2 + x_d^2}$$

A quaternion variable x is called a unit quaternion if it satisfies $|x| = 1$. The product of two quaternions $x, y \in \mathbb{H}$ are defined as

$$xy = \Re[x]\Re[y] - \Im[x] \cdot \Im[y] + \Re[x]\Im[y] + \Re[y]\Im[x] + \Im[x] \times \Im[y]$$

where "$\cdot$" denotes the scalar product and "$\times$" the vector product. Notice that the presence of the vector product in the above expression causes the non-commutativity of the quaternion product, that is, $xy \neq yx$.

Another important notion in the quaternion domain is the so-called "quaternion involution" [44], which defines a self-inverse mapping analogous to the complex conjugate. The general involution of the quaternion vector $\mathbf{x}$ is defined as $\mathbf{x}^\alpha = -\alpha\mathbf{x}\alpha$, and represents the rotation of the vector part of $\mathbf{x}$ by π about a unit pure quaternion α. The special cases of involutions about the ι, J and κ imaginary axes are given by Ell and Sangwine [44]

$$\begin{aligned} \mathbf{x}^\iota &= -\iota\mathbf{x}\iota = \mathbf{x}_a + \mathbf{x}_b\iota - \mathbf{x}_c J - \mathbf{x}_d\kappa \\ \mathbf{x}^J &= -J\mathbf{x}J = \mathbf{x}_a - \mathbf{x}_b\iota + \mathbf{x}_c J - \mathbf{x}_d\kappa \\ \mathbf{x}^\kappa &= -\kappa\mathbf{x}\kappa = \mathbf{x}_a - \mathbf{x}_b\iota - \mathbf{x}_c J + \mathbf{x}_d\kappa \end{aligned} \tag{4}$$

This set of involutions, together with the original quaternion, $\mathbf{x}$, forms the most frequently used basis for augmented quaternion statistics [20,21]. While an involution represents a rotation along

---

[1] Iterative solutions were proposed in the context of quaternion ICA [8].

[2] Throughout this paper, the quaternion vector is assumed to be zero-mean.

a single imaginary axis, the quaternion conjugate operator $(\cdot)^*$ rotates the quaternion along all three imaginary axes, and is given by

$$\mathbf{x}^* = \Re[\mathbf{x}] - \Im[\mathbf{x}] = \mathbf{x}_a - \mathbf{x}_b \imath - \mathbf{x}_c \jmath - \mathbf{x}_d \kappa \tag{5}$$

A useful algebraic operation which conjugates only one imaginary component is given by

$$\mathbf{x}^{\imath*} = (\mathbf{x}^\imath)^* = (\mathbf{x}^*)^\imath = \mathbf{x}_a - \mathbf{x}_b \imath + \mathbf{x}_c \jmath + \mathbf{x}_d \kappa$$
$$\mathbf{x}^{\jmath*} = (\mathbf{x}^\jmath)^* = (\mathbf{x}^*)^\jmath = \mathbf{x}_a + \mathbf{x}_b \imath - \mathbf{x}_c \jmath + \mathbf{x}_d \kappa$$
$$\mathbf{x}^{\kappa*} = (\mathbf{x}^\kappa)^* = (\mathbf{x}^*)^\kappa = \mathbf{x}_a + \mathbf{x}_b \imath + \mathbf{x}_c \jmath - \mathbf{x}_d \kappa \tag{6}$$

Since the only difference between $\mathbf{x}$ and $\mathbf{x}^{\alpha*}$, for $\alpha \in \{\imath, \jmath, \kappa\}$, is the sign of the α-imaginary component, a quaternion vector $\mathbf{x}$ is called $(\cdot)^{\alpha*}$ invariant if $\Im_\alpha[\mathbf{x}] = 0$. For example, if $\mathbf{x}$ has a vanishing $\imath$-imaginary part, then $\mathbf{x}^{\imath*} = \mathbf{x}$. For rigour, the structure and general properties of quaternion matrices are given in Appendix A.

2.2. Quaternion covariance matrices: Structural insights

The recent success of complex augmented statistics is largely due to the simplicity and physical meaningfulness of the statistical descriptors in the form of the standard covariance, $\mathbf{C_x} = E\{\mathbf{xx}^H\}$, and the pseudo-covariance, $\mathbf{P_x} = E\{\mathbf{xx}^T\}$. In particular, the pseudo-covariance enables us to account for the improperness (channel power imbalance or correlation) of complex variables, while the symmetry of $\mathbf{P_x}$ implies that its decomposition can be computed using the Takagi factorisation as $\mathbf{P_x} = \mathbf{Q}\boldsymbol{\Sigma}\mathbf{Q}^T$, where $\mathbf{Q}$ is a complex unitary matrix and $\boldsymbol{\Sigma}$ a real non-negative diagonal matrix [23]. However, the pseudo-covariance of a quaternion random vector $\mathbf{x} = [x_1, x_2, \ldots, x_L]^T$ does not exhibit symmetry, owing to the non-commutativity of the quaternion product, whereby $x_m x_n \neq x_n x_m \, (m \neq n)$, yield an asymmetric matrix

$$\mathbf{P_x} = E\{\mathbf{xx}^T\} = \begin{bmatrix} E\{x_1 x_1\} & E\{x_1 x_2\} & \ldots & E\{x_1 x_L\} \\ E\{x_2 x_1\} & E\{x_2 x_2\} & \ldots & E\{x_2 x_L\} \\ \vdots & \vdots & \ddots & \vdots \\ E\{x_L x_1\} & E\{x_L x_2\} & \ldots & E\{x_L x_L\} \end{bmatrix} \tag{7}$$

The quaternion involution basis in (4) is at the core of the recently proposed widely linear processing [20,45], which provides theoretical and practical performance gains over traditional strictly linear processing [46]. For example, the widely linear minimum mean square error (MMSE) estimation of the quaternion signal $\mathbf{y}$ in terms of the observation $\mathbf{x}$ is performed as

$$\hat{\mathbf{y}} = E\{\mathbf{y}|\mathbf{x}, \mathbf{x}^\imath, \mathbf{x}^\jmath, \mathbf{x}^\kappa\} \tag{8}$$

which is analogous to the complex widely linear MMSE estimation [22], given by

$$\hat{\mathbf{y}} = E\{\mathbf{y}|\mathbf{x}, \mathbf{x}^*\}$$

Then, the involution basis for the quaternion widely linear processing provides a useful description of quaternion second-order statistics, represented by the standard covariance matrix and the $\imath-$, $\jmath-$, and $\kappa-$ complementary covariance matrices, which are given by Cheong Took and Mandic [21]

$$\mathbf{C}_{\mathbf{x}^\alpha} = E\{\mathbf{xx}^{\alpha H}\} = \begin{bmatrix} E\{x_1 x_1^{\alpha*}\} & E\{x_1 x_2^{\alpha*}\} & \cdots & E\{x_1 x_L^{\alpha*}\} \\ E\{x_2 x_1^{\alpha*}\} & E\{x_2 x_2^{\alpha*}\} & \cdots & E\{x_2 x_L^{\alpha*}\} \\ \vdots & \vdots & \ddots & \vdots \\ E\{x_L x_1^{\alpha*}\} & E\{x_L x_2^{\alpha*}\} & \cdots & E\{x_L x_L^{\alpha*}\} \end{bmatrix} \tag{9}$$

where $\alpha \in \{\imath, \jmath, \kappa\}$. The α-complementary covariance matrix is α-Hermitian, that is, $\mathbf{C}_{\mathbf{x}^\alpha} = (\mathbf{C}_{\mathbf{x}^\alpha})^{\alpha H}$, which stems from the fact that its diagonal entries are $(\cdot)^{\alpha*}$ invariant, whereas its off-diagonal entries are governed by the relationship $\mathbf{C}_{\mathbf{x}^\alpha}[m, n] = (\mathbf{C}_{\mathbf{x}^\alpha}[n, m])^{\alpha*}$, where $\mathbf{C}_{\mathbf{x}^\alpha}[m, n]$ denotes the entry of $\mathbf{C}_{\mathbf{x}^\alpha}$ in row m, column n. The

relationship between the pseudo-covariance and the three complementary covariances is given in (1).

Remark 1. The knowledge of both the standard covariance matrix and the three complementary covariance matrices is necessary to ensure the exploitation of complete second-order statistical information in the quaternion domain.

2.3. Quaternion properness

The notion of non-circularity (improperness) is unique to division algebras. While non-circularity refers to probability distributions which are not rotation-invariant, a proper complex random vector $\mathbf{z} = \mathbf{z}_r + \mathbf{z}_\imath \imath$ $(\mathbf{z}_r, \mathbf{z}_\imath \in \mathbb{R}^{L \times 1})$ has a vanishing pseudo-covariance $E\{\mathbf{zz}^T\} = \mathbf{0}$. In other words, its real and imaginary parts are uncorrelated and have equal variance, that is, $E\{\mathbf{z}_r \mathbf{z}_\imath^T\} = \mathbf{0}$ and $E\{\mathbf{z}_r \mathbf{z}_r^T\} = E\{\mathbf{z}_\imath \mathbf{z}_\imath^T\}$. Similarly, improperness in the quaternion domain is characterised by the degree of correlation and/or power imbalance between imaginary components relative to the real component. The additional degrees of freedom in the quaternion domain allow for two types of properness: $\mathbb{H}$-properness and $\mathbb{C}^\alpha$-properness [47].

Definition 1 ($\mathbb{H}$-properness). A quaternion random vector $\mathbf{x}$ is $\mathbb{H}$-proper if it is uncorrelated with its vector involutions, $\mathbf{x}^\imath$, $\mathbf{x}^\jmath$, and $\mathbf{x}^\kappa$, so that

$$\mathbf{C}_{\mathbf{x}^\imath} = E\{\mathbf{xx}^{\imath H}\} = \mathbf{0} \qquad \mathbf{C}_{\mathbf{x}^\jmath} = E\{\mathbf{xx}^{\jmath H}\} = \mathbf{0}$$
$$\mathbf{C}_{\mathbf{x}^\kappa} = E\{\mathbf{xx}^{\kappa H}\} = \mathbf{0} \tag{10}$$

Definition 2 ($\mathbb{C}^\alpha$-properness). A quaternion random vector $\mathbf{x}$ is $\mathbb{C}^\alpha$-improper[3] with respect to $\alpha = \imath, \jmath$ or κ if it is correlated only with the involution $\mathbf{x}^\alpha$, so that all the complementary covariances except for $\mathbf{C}_{\mathbf{x}^\alpha}$ vanish.

As shown in Appendix B, a $\mathbb{C}^\alpha$-improper quaternion random vector can be generated from two proper complex random vectors.

3. Diagonalisation of quaternion covariance matrices

Prior to introducing the diagonalisation of the quaternion pseudo-covariance matrix and the simultaneous diagonalisation of two quaternion covariance matrices, the following two observations establish results essential for subsequent analyses.

Observation 1. The eigendecomposition applied to the Hermitian quaternion covariance matrix, $\mathbf{C_x}$, gives $\mathbf{C_x} = \mathbf{Q}^H \boldsymbol{\Lambda_x} \mathbf{Q}$, where $\mathbf{Q}$ is a quaternion unitary matrix and $\boldsymbol{\Lambda_x}$ a real-valued diagonal matrix for which the entries are the eigenvalues[4] of $\mathbf{C_x}$ [48].

Observation 2. The α-complementary covariance matrix, $\mathbf{C}_{\mathbf{x}^\alpha}$, $\alpha \in \{\imath, \jmath, \kappa\}$, can be factorised as $\mathbf{C}_{\mathbf{x}^\alpha} = \mathbf{Q}_\alpha^H \boldsymbol{\Lambda}_\alpha \mathbf{Q}_\alpha^\alpha$, where $\mathbf{Q}_\alpha$ is a quaternion unitary matrix and $\boldsymbol{\Lambda}_\alpha$ is a real-valued non-negative diagonal matrix for which the diagonal entries are the singular values of $\mathbf{C}_{\mathbf{x}^\alpha}$ [42].

Observation 1 provides an algebraic tool to diagonalise the quaternion covariance matrix. Observation 2 enables the diagonalisation of quaternion complementary covariance matrices, and degenerates into the Takagi factorisation of complex symmetric matrices if the quaternion vector $\mathbf{x}$ has two vanishing imaginary parts [49].

---

[3] It is called $\mathbb{C}^\alpha$-proper in some literature, but we consider the term $\mathbb{C}^\alpha$-improper is more intuitive.

[4] They are both right and left eigenvalues, but are simply termed eigenvalues for conciseness in this paper.

3.1. Diagonalisation of symmetric matrices

The tools for the decomposition of quaternion pseudo-covariance matrices are still in their infancy, because:

- The tools for the analysis of the symmetric matrices in $\mathbb{C}$ cannot be readily generalised to $\mathbb{H}$.
- It is required to simultaneously diagonalise all the three complementary covariance matrices, a task with prohibitively many degrees of freedom.

We shall start our analysis with the diagonalisation of 2×2 symmetric quaternion matrices.

Proposition 1. *A 2×2 symmetric quaternion matrix $\mathbf{A}$ admits the factorisation $\mathbf{A} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^T$ if either its diagonal elements or off-diagonal elements are real-valued, where $\mathbf{U}$ is a quaternion unitary matrix and $\boldsymbol{\Lambda}$ is a real diagonal matrix with the singular values of $\mathbf{A}$ on the diagonal.*

Proof. Based on the SVD of quaternion matrices [37], $\mathbf{A} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{V}^H$, the matrix products $\mathbf{A}\mathbf{A}^*$ and $\mathbf{A}^*\mathbf{A}$ can be expressed as

$$\mathbf{A}\mathbf{A}^* = \mathbf{A}\mathbf{A}^H = (\mathbf{U}\boldsymbol{\Lambda}\mathbf{V}^H)(\mathbf{U}\boldsymbol{\Lambda}\mathbf{V}^H)^H = \mathbf{U}\boldsymbol{\Lambda}^2\mathbf{U}^H$$

$$\mathbf{A}^*\mathbf{A} = \mathbf{A}^H\mathbf{A} = (\mathbf{U}\boldsymbol{\Lambda}\mathbf{V}^H)^H(\mathbf{U}\boldsymbol{\Lambda}\mathbf{V}^H) = \mathbf{V}\boldsymbol{\Lambda}^2\mathbf{V}^H$$

The result in Appendix C shows that $(\mathbf{A}\mathbf{A}^*)^* = \mathbf{A}^*\mathbf{A}$ if either the diagonal or off-diagonal elements of $\mathbf{A}$ are real-valued. It then follows that $\mathbf{U}^*\boldsymbol{\Lambda}^2(\mathbf{U}^*)^H = \mathbf{V}\boldsymbol{\Lambda}^2\mathbf{V}^H$, where the commutativity of the product is valid because of the condition that the diagonal or off-diagonal elements of $\mathbf{A}$ are real-valued. Thus, we can obtain $\mathbf{V} = \mathbf{U}^*$, whereby the Takagi factorisation[5] of quaternion symmetric matrices is identical to that of complex symmetric matrices, $\mathbf{A} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^T$. □

A consequence of Proposition 1 is that the Takagi factorisation of a quaternion symmetric matrix $\mathbf{A}$ is possible if $(\mathbf{A}\mathbf{A}^*)^* = \mathbf{A}^*\mathbf{A}$ holds. However, in general, $(\mathbf{A}\mathbf{A}^*)^* \neq \mathbf{A}^*\mathbf{A}$, indicating that such a Takagi factorisation does not exist for general quaternion symmetric matrices.

3.2. Simultaneous diagonalisation of two covariance matrices

We next proceed to introduce a simultaneous diagonalisation of the covariance and complementary covariance matrices. The finding that a non-singular transform is sufficient to diagonalise[6] both $\mathbf{C_x}$ and $\mathbf{C_{x^\alpha}}$ will form a basis for the quaternion uncorrelating transform [43]. First, we shall clarify that the unitary transform which can simultaneously diagonalise $\mathbf{C_x}$ and $\mathbf{C_{x^\alpha}}$ for a general quaternion random vector $\mathbf{x}$ does not exist.

Proposition 2. *If $\mathbf{C_x}\mathbf{C_{x^\alpha}} \neq (\mathbf{C_x}\mathbf{C_{x^\alpha}})^{\alpha H}$, then there is no quaternion unitary matrix $\mathbf{U}$ such that both $\mathbf{U}^H\mathbf{C_x}\mathbf{U}$ and $\mathbf{U}^H\mathbf{C_{x^\alpha}}\mathbf{U}^\alpha$ are diagonal.*

Proof. Define $\boldsymbol{\Lambda}_\mathbf{x} = \mathbf{U}^H\mathbf{C_x}\mathbf{U}$, $\boldsymbol{\Lambda}_\alpha = \mathbf{U}^H\mathbf{C_{x^\alpha}}\mathbf{U}^\alpha$, where $\mathbf{U}$ is a quaternion unitary matrix. It is then obvious that $\boldsymbol{\Lambda}_\mathbf{x}$ is a Hermitian matrix, so that if $\boldsymbol{\Lambda}_\mathbf{x}$ is diagonal, then $\boldsymbol{\Lambda}_\mathbf{x}$ is real diagonal. Furthermore, if $\boldsymbol{\Lambda}_\alpha$ is also diagonal, we can obtain

$$\mathbf{C_x}\mathbf{C_{x^\alpha}} = \mathbf{U}\boldsymbol{\Lambda}_\mathbf{x}\mathbf{U}^H\mathbf{U}\boldsymbol{\Lambda}_\alpha\mathbf{U}^{\alpha H} = \mathbf{U}\boldsymbol{\Lambda}_\alpha\mathbf{U}^{\alpha H}\mathbf{U}^\alpha\boldsymbol{\Lambda}_\mathbf{x}\mathbf{U}^{\alpha H} = \mathbf{C_{x^\alpha}}\mathbf{C_x}^\alpha \qquad (11)$$

Because $\mathbf{C_x}$ is Hermitian and $\mathbf{C_{x^\alpha}}$ is α-Hermitian, we can readily obtain

$$\mathbf{C_{x^\alpha}}\mathbf{C_x}^\alpha = \mathbf{C_{x^\alpha}}^{\alpha H}\mathbf{C_x}^{\alpha H} = (\mathbf{C_x}\mathbf{C_{x^\alpha}})^{\alpha H} \qquad (12)$$

Combining (11) and (12) yields $\mathbf{C_x}\mathbf{C_{x^\alpha}} = (\mathbf{C_x}\mathbf{C_{x^\alpha}})^{\alpha H}$. □

Although a unitary transform for the simultaneous diagonalisation of $\mathbf{C_x}$ and $\mathbf{C_{x^\alpha}}$ is inapplicable, a non-unitary transform can be derived in a similar manner to the SUT in the complex domain. On the basis of Observation 1, $\mathbf{C_x}$ can be factorised as $\mathbf{C_x} = \mathbf{V}\boldsymbol{\Lambda}_\mathbf{x}\mathbf{V}^H$ where $\mathbf{V}$ is a quaternion unitary matrix, and $\boldsymbol{\Lambda}_\mathbf{x}$ is a real diagonal matrix. Define a whitening transform $\mathbf{D} = \mathbf{V}\boldsymbol{\Lambda}_\mathbf{x}^{-\frac{1}{2}}\mathbf{V}^H$, and denote $\mathbf{s} = \mathbf{D}\mathbf{x}$, to obtain the covariance matrix of $\mathbf{s}$ as

$$\mathbf{C_s} = \mathbf{D}\mathbf{C_x}\mathbf{D}^H = \mathbf{V}\boldsymbol{\Lambda}_\mathbf{x}^{-\frac{1}{2}}\mathbf{V}^H\mathbf{V}\boldsymbol{\Lambda}_\mathbf{x}\mathbf{V}^H\mathbf{V}\boldsymbol{\Lambda}_\mathbf{x}^{-\frac{1}{2}}\mathbf{V}^H = \mathbf{I}$$

Based on Observation 2, the α-complementary covariance matrix of $\mathbf{s}$ can be factorised as $\mathbf{C_{s^\alpha}} = \mathbf{W}\boldsymbol{\Lambda}_\alpha\mathbf{W}^{\alpha H}$, where $\mathbf{W}$ is a quaternion unitary matrix, and $\boldsymbol{\Lambda}_\alpha$ is a real diagonal matrix. Now, the non-singular uncorrelating transform, $\mathbf{Q} = \mathbf{W}^H\mathbf{D}$, and its application in the form $\mathbf{y} = \mathbf{Q}\mathbf{x}$ yield the simultaneous diagonalisation

$$\mathbf{C_y} = \mathbf{W}^H\mathbf{C_s}\mathbf{W} = \mathbf{W}^H\mathbf{I}\mathbf{W} = \mathbf{I}$$

$$\mathbf{C_{y^\alpha}} = \mathbf{W}^H\mathbf{C_{s^\alpha}}\mathbf{W}^\alpha = \mathbf{W}^H\mathbf{W}\boldsymbol{\Lambda}_\alpha\mathbf{W}^{\alpha H}\mathbf{W}^\alpha = \boldsymbol{\Lambda}_\alpha$$

We refer to the transform $\mathbf{Q}$ as the *quaternion uncorrelating transform* (QUT), which is summarised in Proposition 3 and Algorithm 1.

Algorithm 1 Quaternion Uncorrelating Transform (QUT).

1. Compute the eigendecomposition of the covariance matrix $\mathbf{C_x} = E\{\mathbf{x}\mathbf{x}^H\} = \mathbf{V}\boldsymbol{\Lambda}_\mathbf{x}\mathbf{V}^H$.
2. Compute the whitening matrix $\mathbf{D} = \mathbf{V}\boldsymbol{\Lambda}_\mathbf{x}^{-\frac{1}{2}}\mathbf{V}^H$.
3. Calculate the α-complementary covariance of the whitened data, $\mathbf{s} = \mathbf{D}\mathbf{x}$, as $\mathbf{C_{s^\alpha}} = E\{\mathbf{s}\mathbf{s}^{\alpha H}\}$.
4. Compute the factorisation of the α-complementary covariance matrix, $\mathbf{C_{s^\alpha}} = \mathbf{W}\boldsymbol{\Lambda}_\alpha\mathbf{W}^{\alpha H}$, where $\mathbf{W}$ is a quaternion unitary matrix and $\boldsymbol{\Lambda}_\alpha$ a real diagonal matrix.
5. The QUT matrix is then $\mathbf{Q} = \mathbf{W}^H\mathbf{D}$.

Proposition 3 (QUT)**.** *For a random quaternion vector $\mathbf{x}$ with finite second-order statistics, there exists a quaternion non-singular matrix $\mathbf{Q}$ for which the transformed vector $\mathbf{y} = \mathbf{Q}\mathbf{x}$ has the following covariance and α-complementary covariance matrices: $\mathbf{C_y} = \mathbf{I}$ $\mathbf{C_{y^\alpha}} = \boldsymbol{\Lambda}_\alpha$,   where $\alpha \in \{\iota, \jmath, \kappa\}$.*

Recall that for a $\mathbb{C}^\alpha$-improper vector $\mathbf{x}$, only the α-complementary covariance is non-vanishing while the other two complementary covariances vanish. Thus, the QUT of $\mathbb{C}^\alpha$-improper signals, for example, for the case $\alpha = \kappa$, satisfies

$$\mathbf{C_y} = \mathbf{I} \qquad \mathbf{C_{y^\iota}} = \boldsymbol{\Lambda}_\iota \qquad \mathbf{C_{y^\jmath}} = \boldsymbol{\Lambda}_\jmath \qquad \mathbf{C_{y^\kappa}} = \boldsymbol{\Lambda}_\kappa$$

where $\boldsymbol{\Lambda}_\kappa$ is a diagonal matrix and $\boldsymbol{\Lambda}_\iota = \boldsymbol{\Lambda}_\jmath = \mathbf{0}$. In other words, for a $\mathbb{C}^\alpha$-improper quaternion vector, the QUT can be regarded as the quaternion version of SUT [15,16]. For a general improper quaternion vector, however, the remaining two complementary covariance matrices are still not diagonalised by the QUT. In fact, so far there is no closed-form solution to joint diagonalisation of four covariance matrices for general improper quaternion vectors. To circumvent this problem, we shall next introduce an approximate way to simultaneously diagonalise all four covariance matrices, a practically useful tool in most applications of quaternions, considering the effectiveness of the AUT in complex-valued signal processing [28–30].

4. Simultaneous diagonalisation of all four covariance matrices

Following the approach in Section 3, we next propose the quaternion analogue of our proposed approximate uncorrelating transform (AUT) for complex-valued data [25]. This will enable the simultaneous diagonlistion of the covariance matrix and three complementary covariance matrices.

---

[5] The Takagi factorisation is a special case of the complex/quaternion SVD when $\mathbf{A} = \mathbf{A}^T$.

[6] Corollary 4.5.18 (c) in [23] is the corresponding result in the complex domain.

4.1. Univariate QAUT

Consider first the univariate quaternion random vector $\mathbf{x} = (x_1, \ldots, x_L)^T$. Our aim is to find a unitary transform $\mathbf{y} = \boldsymbol{\Phi}\mathbf{x}$ which simultaneously diagonalises the covariance matrix, $\mathbf{C_y} = E\{\mathbf{yy}^H\} = \boldsymbol{\Phi}\mathbf{C_x}\boldsymbol{\Phi}^H$, and the three complementary covariance matrices, $\mathbf{C}_{\mathbf{y}^\alpha} = E\{\mathbf{yy}^{\alpha H}\} = \boldsymbol{\Phi}\mathbf{C}_{\mathbf{x}^\alpha}\boldsymbol{\Phi}^{\alpha H}$, $\alpha \in \{\imath, \jmath, \kappa\}$. Notice that if $\mathbf{C_x}\mathbf{C_x}^H = \mathbf{C}_{\mathbf{x}^\alpha}\mathbf{C}_{\mathbf{x}^\alpha}^H$, we have $\mathbf{C_y}\mathbf{C_y}^H = \mathbf{C}_{\mathbf{y}^\alpha}\mathbf{C}_{\mathbf{y}^\alpha}^H$.

For N samples of $\mathbf{x}$, $N > L$, the covariance matrix, and the α-complementary covariance matrix can be estimated as

$$\mathbf{C_x} = \frac{1}{N}\sum_{n_1=1}^{N}\mathbf{x}(n_1)\mathbf{x}^H(n_1) \tag{13}$$

$$\mathbf{C}_{\mathbf{x}^\alpha} = \frac{1}{N}\sum_{n_1=1}^{N}\mathbf{x}(n_1)\mathbf{x}^{\alpha H}(n_1) \tag{14}$$

Their quadratic forms $\mathbf{C_x}\mathbf{C_x}^H$ and $\mathbf{C}_{\mathbf{x}^\alpha}\mathbf{C}_{\mathbf{x}^\alpha}^H$ are calculated as

$$\mathbf{C_x}\mathbf{C_x}^H = \left[\frac{1}{N}\sum_{n_1=1}^{N}\mathbf{x}(n_1)\mathbf{x}^H(n_1)\right]\left[\frac{1}{N}\sum_{n_2=1}^{N}\mathbf{x}(n_2)\mathbf{x}^H(n_2)\right]$$
$$= \frac{1}{N^2}\sum_{n_1=1}^{N}\sum_{n_2=1}^{N}\mathbf{x}(n_1)\left[\sum_{l=1}^{L}x_l^*(n_1)x_l(n_2)\right]\mathbf{x}^H(n_2)$$

$$\mathbf{C}_{\mathbf{x}^\alpha}\mathbf{C}_{\mathbf{x}^\alpha}^H = \left[\frac{1}{N}\sum_{n_1=1}^{N}\mathbf{x}(n_1)\mathbf{x}^{\alpha H}(n_1)\right]\left[\frac{1}{N}\sum_{n_2=1}^{N}\mathbf{x}^\alpha(n_2)\mathbf{x}^H(n_2)\right]$$
$$= \frac{1}{N^2}\sum_{n_1=1}^{N}\sum_{n_2=1}^{N}\mathbf{x}(n_1)\left[\sum_{l=1}^{L}x_l^*(n_1)x_l(n_2)\right]^\alpha\mathbf{x}^H(n_2)$$

Upon applying the approximation

$$\sum_{l=1}^{L}x_l^*(n_1)x_l(n_2) \approx \left[\sum_{l=1}^{L}x_l^*(n_1)x_l(n_2)\right]^\alpha \tag{15}$$

we obtain $\mathbf{C_x}\mathbf{C_x}^H \approx \mathbf{C}_{\mathbf{x}^\alpha}\mathbf{C}_{\mathbf{x}^\alpha}^H$, and hence $\mathbf{C_y}\mathbf{C_y}^H \approx \mathbf{C}_{\mathbf{y}^\alpha}\mathbf{C}_{\mathbf{y}^\alpha}^H$, which implies that if $\mathbf{C_y}$ is diagonal, so too is $\mathbf{C}_{\mathbf{y}^\alpha}$, and *vice versa*. Furthermore, if

$$\sum_{l=1}^{L}x_l^*(n_1)x_l(n_2) \in \mathbb{R} \tag{16}$$

holds approximately, then (15) holds for all $\alpha \in \{\imath, \jmath, \kappa\}$, and thus the simultaneous diagonalisation applies to all of the three complementary covariance matrices, producing the following *quaternion approximate uncorrelating transform* (QAUT).

Proposition 4 (QAUT). *If any of $\mathbf{C_x}$, $\mathbf{C}_{\mathbf{x}^\imath}$, $\mathbf{C}_{\mathbf{x}^\jmath}$ and $\mathbf{C}_{\mathbf{x}^\kappa}$ is diagonalised by the unitary matrix $\boldsymbol{\Phi}$, the structural similarity between the four covariance matrices, displayed in (15), makes possible the approximate diagonalisation of the other three covariance matrices by $\boldsymbol{\Phi}$. From (1), the pseudo-covariance matrix is also approximately diagonalised in this way.*

The transformation matrix $\boldsymbol{\Phi}$ can take the form of $\mathbf{Q}$ in Observation 1 or $\mathbf{Q}_\alpha$ in Observation 2. For example, the implementation of the transform $\mathbf{Q}$, obtained from the eigendecomposition of $\mathbf{C_x}$, yields

$$\mathbf{C_x} = \mathbf{Q}^H\boldsymbol{\Lambda_x}\mathbf{Q}$$
$$\mathbf{C}_{\mathbf{x}}^\imath \approx \mathbf{Q}^H\boldsymbol{\Lambda}_\imath\mathbf{Q}^\imath$$
$$\mathbf{C}_{\mathbf{x}}^\jmath \approx \mathbf{Q}^H\boldsymbol{\Lambda}_\jmath\mathbf{Q}^\jmath$$
$$\mathbf{C}_{\mathbf{x}}^\kappa \approx \mathbf{Q}^H\boldsymbol{\Lambda}_\kappa\mathbf{Q}^\kappa \tag{17}$$

4.2. Principle of QAUT

The QAUT is based on the approximation in (16). To examine the validity of this approximation for highly improper quaternion signals, assume that the four components of the quaternion variable x_l are perfectly correlated and consider the Cayley–Dickson construction [50]

$$x_l = x_a + x_b\imath + x_c\jmath + x_d\kappa = c_1 + c_2\jmath$$

where $c_1 = x_a + x_b\imath$ and $c_2 = x_c + x_d\imath$ are complex variables defined in the plane spanned by $\{1, \imath\}$. This gives

$$x_l^*(n_1)x_l(n_2) = c_1^*(n_1)c_1(n_2) + c_2(n_1)c_2^*(n_2)$$
$$+ [c_1^*(n_1)c_2(n_2) - c_2(n_1)c_1^*(n_2)]\jmath \tag{18}$$

Upon applying the statistical expectation operator, we have

$$E\{x_l^*(n_1)x_l(n_2)\} = E\{c_1^*(n_1)c_1(n_2)\} + E\{c_2(n_1)c_2^*(n_2)\}$$
$$+ E\{c_1^*(n_1)c_2(n_2) - c_2(n_1)c_1^*(n_2)\}\jmath \tag{19}$$

where

$$E\{c_1^*(n_1)c_1(n_2)\} = E\{x_a(n_1)x_a(n_2) + x_b(n_1)x_b(n_2)\}$$
$$+ E\{x_a(n_1)x_b(n_2) - x_a(n_2)x_b(n_1)\}\imath$$
$$\approx E\{x_a(n_1)x_a(n_2) + x_b(n_1)x_b(n_2)\}$$
$$+ E\{x_b(n_1)x_b(n_2) - x_b(n_2)x_b(n_1)\}\imath$$
$$= E\{x_a(n_1)x_a(n_2) + x_b(n_1)x_b(n_2)\}$$
$$\in \mathbb{R} \tag{20}$$

and likewise for $E\{c_2(n_1)c_2^*(n_2)\} \in \mathbb{R}$. Moreover, if c_1 and c_2 are highly correlated, they can be expressed as $c_2(t) \approx kc_1(t) + g$, where k and g are complex-valued constants, to give

$$E\{c_1^*(n_1)c_2(n_2) - c_2(n_1)c_1^*(n_2)\}$$
$$\approx E\{c_1^*(n_1)[kc_1(n_2) + g] - [kc_1(n_1) + g]c_1^*(n_2)\}$$
$$= 2kE\{\Im[c_1^*(n_1)c_1(n_2)]\} + E\{g\}E\{c_1^*(n_1)\}$$
$$- E\{g\}E\{c_1^*(n_2)\}$$
$$\approx 0 \tag{21}$$

Therefore, we obtain approximately

$$E\{x_l^*(n_1)x_l(n_2)\} \in \mathbb{R} \tag{22}$$

and the approximation (16) holds given the assumption

$$E\{x_l^*(n_1)x_l(n_2)\} \approx \frac{1}{L}\sum_{l=1}^{L}x_l^*(n_1)x_l(n_2)$$

From the above analysis, the QAUT holds exactly when the four components of the quaternion variable are perfectly correlated (a maximally improper x). For general quaternion signals, however, these components are not perfectly correlated, causing the approximation error impacted by the correlation coefficients, given by

$$\rho_{ab} = \frac{\text{cov}[x_a, x_b]}{\sigma_{x_a}\sigma_{x_b}}, \quad \rho_{ac} = \frac{\text{cov}[x_a, x_c]}{\sigma_{x_a}\sigma_{x_c}}, \quad \rho_{ad} = \frac{\text{cov}[x_a, x_d]}{\sigma_{x_a}\sigma_{x_d}},$$
$$\rho_{bc} = \frac{\text{cov}[x_b, x_c]}{\sigma_{x_b}\sigma_{x_c}}, \quad \rho_{bd} = \frac{\text{cov}[x_b, x_d]}{\sigma_{x_b}\sigma_{x_d}}, \quad \rho_{cd} = \frac{\text{cov}[x_c, x_d]}{\sigma_{x_c}\sigma_{x_d}}$$

with the range $[-1, 1]$, where 'cov' denotes the covariance and σ the standard deviation. From the diagonalisation condition in (15), the diagonalisation error of $\mathbf{C}_{\mathbf{x}^\alpha}$ arises from and is positively correlated with

$$x_l^*(n_1)x_l(n_2) - \left[x_l^*(n_1)x_l(n_2)\right]^\alpha$$
$$= 2\Im_\beta\left[x_l^*(n_1)x_l(n_2)\right]\beta + 2\Im_\gamma\left[x_l^*(n_1)x_l(n_2)\right]\gamma$$
$$\text{for distinct } \alpha, \beta, \gamma \in \{\imath, \jmath, \kappa\} \tag{23}$$

Note that the imaginary part of $x_l^*(n_1)x_l(n_2)$ is given by

$$
\begin{aligned}
\Im\big[x_l^*(n_1)x_l(n_2)\big] \\
= \ & [x_a(n_1)x_b(n_2) - x_a(n_2)x_b(n_1) + x_c(n_2)x_d(n_1) - x_c(n_1)x_d(n_2)]\imath \\
& + [x_a(n_1)x_c(n_2) - x_a(n_2)x_c(n_1) + x_b(n_1)x_d(n_2) - x_b(n_2)x_d(n_1)]\jmath \\
& + [x_a(n_1)x_d(n_2) - x_a(n_2)x_d(n_1) + x_b(n_2)x_c(n_1) - x_b(n_1)x_c(n_2)]\kappa
\end{aligned}
\tag{24}
$$

Similarly to the AUT for complex-valued data [25], Eq. (24) can be rewritten as

$$
\begin{aligned}
\Im\big[x_l^*(n_1)x_l(n_2)\big] = \ & \left(\xi_{ab}\sqrt{1-\rho_{ab}^2} + \xi_{cd}\sqrt{1-\rho_{cd}^2}\right) \\
& + \left(\xi_{ac}\sqrt{1-\rho_{ac}^2} + \xi_{bd}\sqrt{1-\rho_{bd}^2}\right)J \\
& + \left(\xi_{ad}\sqrt{1-\rho_{ad}^2} + \xi_{bc}\sqrt{1-\rho_{bc}^2}\right)\kappa
\end{aligned}
\tag{25}
$$

where $\xi_{ab}, \xi_{cd}, \xi_{ac}, \xi_{bd}, \xi_{ad}, \xi_{bc}$ are real-valued coefficients dependent on the statistics of x_l.

It is now obvious that when the components of quaternion data become more correlated, (25) decreases and the diagonalisation error hence decreases. Owing to the cumulative nature of the error in (16), the approximation error also increases with the data segment length L.

4.3. Multivariate QAUT

We next extend the univariate QAUT to the multivariate case when an M-variate quaternion signal is represented by a data matrix $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_M]^T$, where the m-th column random vector $\mathbf{x}_m = [x_m(1), \ldots, x_m(N)]^T$ $(N > M \geq m)$ represents N samples of the m-th variate. The covariance and α-covariance matrices of the M variates are then given by $\mathbf{C_X} = \frac{1}{N}E\{\mathbf{X}\mathbf{X}^H\}$ and $\mathbf{C_X^\alpha} = \frac{1}{N}E\{\mathbf{X}\mathbf{X}^{\alpha H}\}$. Our aim is to find a unitary transform

$$
\mathbf{Y} = \mathbf{\Phi}\mathbf{X}
\tag{26}
$$

which simultaneously diagonalises the covariance matrix, $\mathbf{C_Y} = E\{\mathbf{Y}\mathbf{Y}^H\} = \mathbf{\Phi}\mathbf{C_X}\mathbf{\Phi}^H$, and the three complementary covariance matrices, $\mathbf{C_{Y^\alpha}} = E\{\mathbf{Y}\mathbf{Y}^{\alpha H}\} = \mathbf{\Phi}\mathbf{C_{X^\alpha}}\mathbf{\Phi}^{\alpha H}$, $\alpha \in \{\imath, \jmath, \kappa\}$. Notice that if $\mathbf{C_X}\mathbf{C_X^H} = \mathbf{C_{X^\alpha}}\mathbf{C_{X^\alpha}^H}$, then $\mathbf{C_Y}\mathbf{C_Y^H} = \mathbf{C_{Y^\alpha}}\mathbf{C_{Y^\alpha}^H}$.

The covariance and α-complementary covariance matrices can be estimated from $\mathbf{X}$ as

$$
\mathbf{C_X} = \frac{1}{N}\mathbf{X}\mathbf{X}^H
\tag{27}
$$

$$
\mathbf{C_X^\alpha} = \frac{1}{N}\mathbf{X}\mathbf{X}^{\alpha H}
\tag{28}
$$

Their quadratic forms $\mathbf{C_X}\mathbf{C_X^H}$ and $\mathbf{C_{X^\alpha}}\mathbf{C_{X^\alpha}^H}$ are calculated as

$$
\mathbf{C_X}\mathbf{C_X^H} = \frac{1}{N^2}\mathbf{X}\mathbf{X}^H\mathbf{X}\mathbf{X}^H = \frac{1}{N^2}\mathbf{X}\Big(\sum_{m=1}^M \mathbf{x}_m^*\mathbf{x}_m^T\Big)\mathbf{X}^H
$$

$$
\mathbf{C_{X^\alpha}}\mathbf{C_{X^\alpha}^H} = \frac{1}{N^2}\mathbf{X}\mathbf{X}^{\alpha H}\mathbf{X}^\alpha\mathbf{X}^H = \frac{1}{N^2}\mathbf{X}\Big(\sum_{m=1}^M \mathbf{x}_m^*\mathbf{x}_m^T\Big)^\alpha\mathbf{X}^H
$$

Following the analysis in (18)–(21), we can prove that the $[n_1, n_2]$ element of the matrix $E\{\mathbf{x}_m^*\mathbf{x}_m^T\}$, $E\{x_m^*(n_1)x_m(n_2)\}$, is real-valued, for all $n_1, n_2 \in \{1, \ldots, N\}$, whereby $E\{\mathbf{x}_m^*\mathbf{x}_m^T\} \in \mathbb{R}^{N\times N}$ and

$$
\sum_{m=1}^M \mathbf{x}_m^*\mathbf{x}_m^T \in \mathbb{R}^{N\times N}
$$

Under this condition, we obtain $\mathbf{C_X}\mathbf{C_X^H} \approx \mathbf{C_{X^\alpha}}\mathbf{C_{X^\alpha}^H}$, so that $\mathbf{C_Y}\mathbf{C_Y^H} \approx \mathbf{C_{Y^\alpha}}\mathbf{C_{Y^\alpha}^H}$, which implies if $\mathbf{C_Y}$ is diagonal, so too is $\mathbf{C_{Y^\alpha}}$, and *vice*
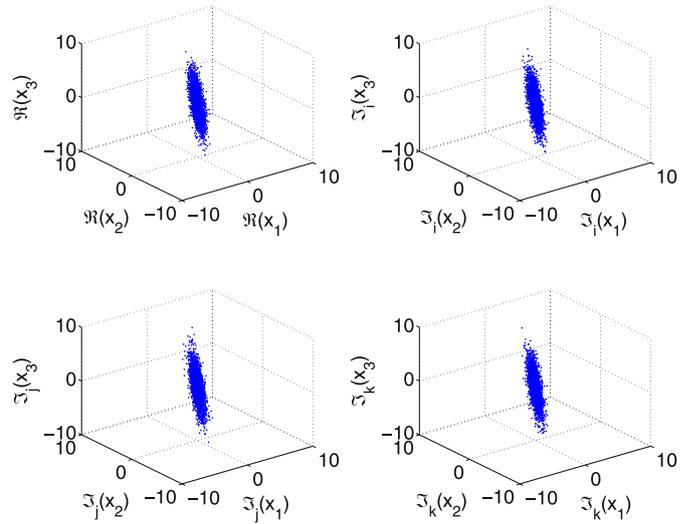
Fig. 1. Scatter plots, indicating the high correlation (improperness) of the three correlated $\mathbb{C}^\kappa$-improper quaternion signals, $\mathbf{x}_1$, $\mathbf{x}_2$ and $\mathbf{x}_3$.

versa. This means that the four covariance matrices of multivariate quaternion data can be simultaneously diagonalised via the QAUT, such as in (17).

Similarly to the univariate case, the diagonalisation error of the multivariate QAUT decreases with an increase in the correlation between the components of quaternion data, and increases with the number of variates, M.

Remark 2. The QAUT becomes exact for quaternion signals with special second-order statistical properties, such as $\mathbb{H}$-proper signals, the signals with fully correlated components, and the signals for which the covariance and complementary covariance matrices have the same singular vectors. This can be verified by exploring the underlying matrix structures.

5. Simulation studies

The effectiveness of the proposed QUT and QAUT techniques is illustrated via simulations on both synthetic and real-world quaternion signals. In the experiments, the covariance and complementary covariance matrices were estimated based on (13), (14), (27) and (28).

5.1. Performance of QUT for $\mathbb{C}^\eta$-improper signals

In the first experiment, the QUT was applied to the decorrelation of multivariate $\mathbb{C}^\kappa$-improper quaternion signals. First, three uncorrelated $\mathbb{C}^\kappa$-improper quaternion signals were generated through the approach introduced in Appendix B and were subsequently mixed through a 3×3 matrix for which the elements were drawn from a standard normal distribution. The mixed $\mathbb{C}^\kappa$-improper signals, $\mathbf{x}_1$, $\mathbf{x}_2$ and $\mathbf{x}_3$, were correlated in terms of the κ-complementary covariance, with the correlation coefficients 0.69 (between $\mathbf{x}_1$ and $\mathbf{x}_2$), 0.34 (between $\mathbf{x}_1$ and $\mathbf{x}_3$), and 0.91 (between $\mathbf{x}_2$ and $\mathbf{x}_3$). Fig. 1 shows 3-D scatter plots of the mixed signals, $\mathbf{x}_1$, $\mathbf{x}_2$ and $\mathbf{x}_3$, illustrating their high correlation, as indicated by elliptical shapes of the scatter plots at an angle to the coordinate axes. The QUT was used to decorrelate $\mathbf{x}_1$, $\mathbf{x}_2$ and $\mathbf{x}_3$ into $\mathbf{y}_1$, $\mathbf{y}_2$ and $\mathbf{y}_3$, as illustrated in Fig. 2, which shows much more circular nature of $\mathbf{y}_1$, $\mathbf{y}_2$ and $\mathbf{y}_3$, as desired.
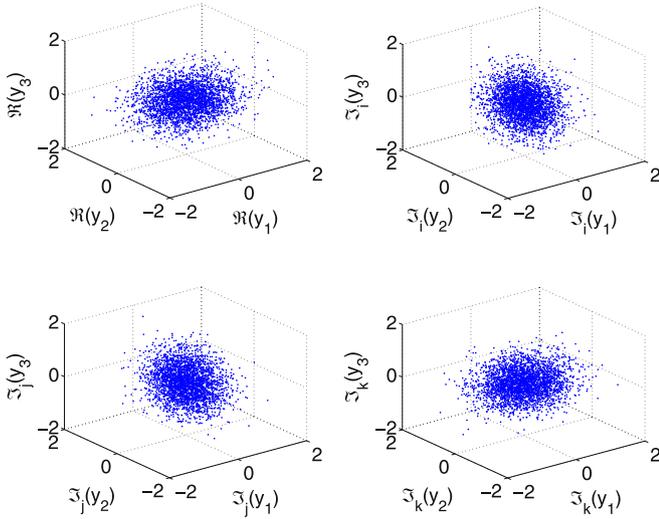
Fig. 2. Scatter plots, indicating the low correlation (properness) of the three $\mathbb{C}^\kappa$-improper quaternion signals, $\mathbf{y}_1$, $\mathbf{y}_2$ and $\mathbf{y}_3$, decorrelated via the QUT.

5.2. Performance of QAUT for synthetic improper signals

To assess the performance of QAUT, the squared diagonal error of the complementary covariance matrix $\mathbf{C}_{\mathbf{x}^\alpha}$ was measured by an average power ratio of the off-diagonal, $\lambda_{\alpha,ij}$, versus diagonal, $\lambda_{\alpha,ii}$, elements of the approximately diagonal matrix $\mathbf{\Lambda}_\alpha$, given by

$$\varepsilon_\alpha^2 = \frac{\sum_{i,j=1}^{L} \left| \lambda_{\alpha,ij} \right|^2}{(N-1)\sum_{i=1}^{L} \lambda_{\alpha,ii}^2} \times 100\%, \quad i \neq j$$

Synthetic quaternion signals with a varying degree of correlation between the real, $\imath$, $\jmath$ and κ components were considered. For simplicity, we assumed the six correlation coefficients of the four components to be equally distributed in $(0, 1]$ and denoted by ρ, where $\rho = 1$ indicates full correlation and $\rho = 0$ no correlation.

In the univariate simulations, a synthetic quaternion data vector with a varying ρ was generated from a quaternion white Gaussian signal. Then, the performance of the univariate QAUT was assessed comprehensively against the degree of correlation present in the data components, $\rho \in [0.5, 1]$, and the length of data segment, $L \in [5, 20]$. Conforming with the analysis, Fig. 3 shows that the diagonalisation error increased as ρ or L increased. The performance was excellent, for example, for a moderate correlation degree, $\rho = 0.5$, and a long data segment, $L = 20$, where the squared diagonalisation error was less than 1%, while for a high correlation degree, ρ is close to unit, where the error was negligible.

In the multivariate simulations, multiple channels of quaternion signals with a varying ρ were generated and then mixed using a matrix the elements of which were drawn from a standard normal distribution, to generate multivariate quaternion data. The performance of the multivariate QAUT was assessed against the degree of correlation present in the data components, $\rho \in [0.5, 1]$, and the number of variates, $M \in [5, 20]$. Fig. 4 shows that the diagonalisation error increased as ρ or M increased. Similarly to the univariate case, for a moderate correlation degree, $\rho = 0.5$, and a large number of data variates, $M = 20$, the squared diagonalisation error was less than 1%, while for a high correlation degree, ρ is close to unit, the error was negligible.
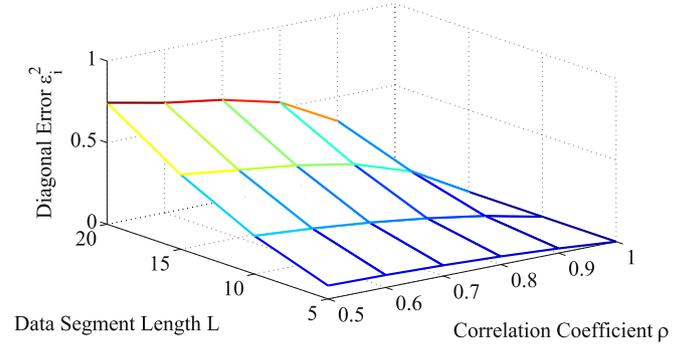
(a) Squared diagonal error of $\mathbf{C}_{\mathbf{x}^\imath}$
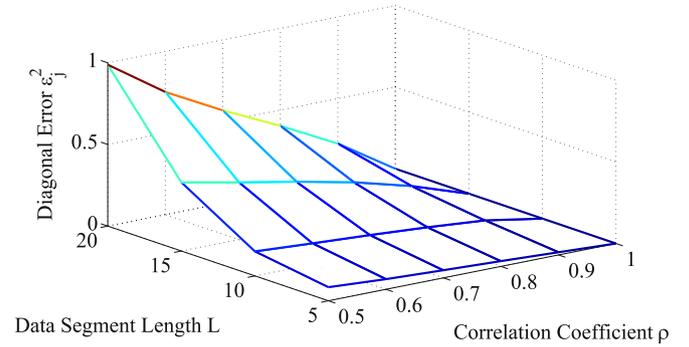
(b) Squared diagonal error of $\mathbf{C}_{\mathbf{x}^\jmath}$
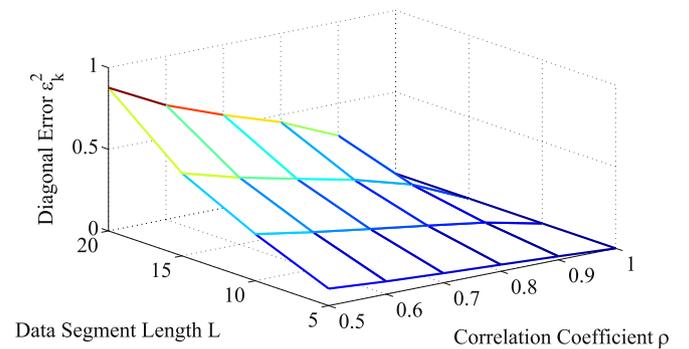
(c) Squared diagonal error of $\mathbf{C}_{\mathbf{x}^\kappa}$

Fig. 3. Squared diagonal error [in %] against the correlation degree and the data segment length for synthetic univariate data.

5.3. Performance of QAUT for real-world signals

The QAUT was also applied to 4-D real-world electroencephalogram (EEG) signals by combining four adjacent channels of real-valued EEG signals into a quaternion-valued signal. The four channels measuring adjacent brain regions exhibited strong correlation with each other while the channels measuring distant brain regions were relatively weakly correlated [51,52]. We tested the QAUT on such quaternion EEG signals with low, medium and high correlations between the four dimensions, with the data segment length L varying between 5 and 20. Fig. 5 illustrates that the diagonalisation error increased with either an increase in L or a decrease in the degree of correlation present in the data. The squared diagonal error was less than 0.2%.
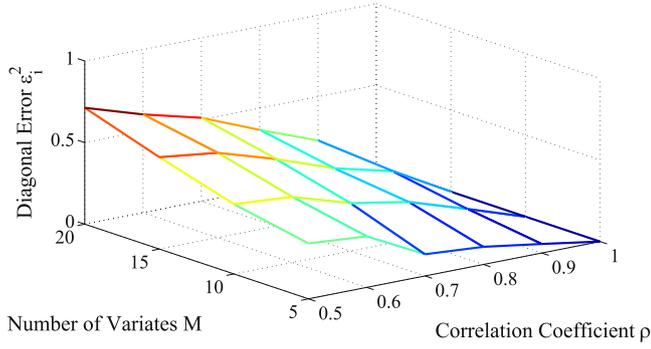
(a) Squared diagonal error of $\mathbf{C}_{\mathbf{x}^{\imath}}$
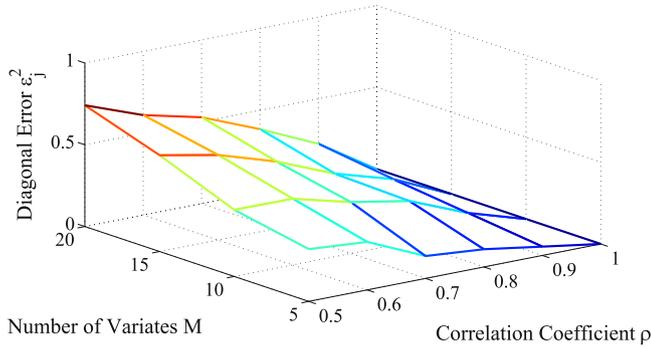
(b) Squared diagonal error of $\mathbf{C}_{\mathbf{x}^{\jmath}}$
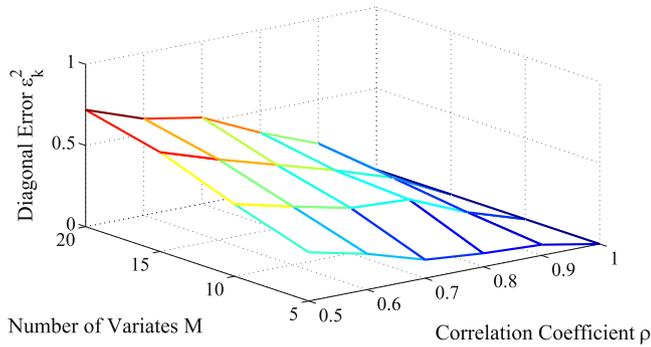
(c) Squared diagonal error of $\mathbf{C}_{\mathbf{x}^{\kappa}}$

Fig. 4. Squared diagonal error [in %] against the correlation degree and the number of variates for synthetic multivariate data.

(a) Squared diagonal error of $\mathbf{C}_{\mathbf{x}^{\imath}}$

(b) Squared diagonal error of $\mathbf{C}_{\mathbf{x}^{\jmath}}$

(c) Squared diagonal error of $\mathbf{C}_{\mathbf{x}^{\kappa}}$

Fig. 5. Squared diagonal error [in %] against the correlation degree and the data segment length for EEG data.

5.4. QAUT for rank reduction

In many practical applications the acquired data are highly correlated or even collinear, with redundant elements that do not contribute to the accuracy of processing but require excessive computational resources and may affect stability of algorithms through the associated singular correlation matrices. It is therefore desirable to use the minimum number of variables to describe the information within a data set; this is typically achieved through a low-rank matrix approximation. For data requiring rank reduction, the assumption that the components of variables are highly correlated is sensible, that is, the approximation conditions for the QAUT hold. The multivariate QAUT in (26) can be therefore rearranged to decompose the data matrix $\mathbf{X}$ into a sum of uncorrelated rank one vector outer products

$$\mathbf{X} = \mathbf{\Phi}^H \mathbf{Y} = \sum_{m=1}^{M} \boldsymbol{\phi}_m^H \mathbf{y}_m$$

where $\boldsymbol{\phi}_m$ and $\mathbf{y}_m$ are rows in the matrices $\mathbf{\Phi}$ and $\mathbf{Y}$, respectively, and the transformed variables, $\mathbf{y}_m$, are monotonically nonincreasing in the variance. The use of variables for which the variance is above a defined threshold and the disposal of the remaining variables, which account for noise, provides the reduced rank
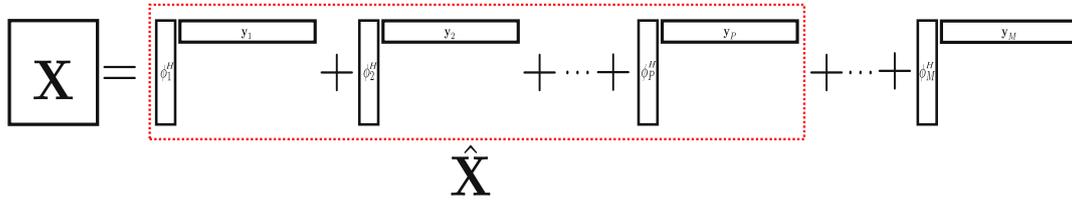
Fig. 6. Rank reduction using the QAUT.
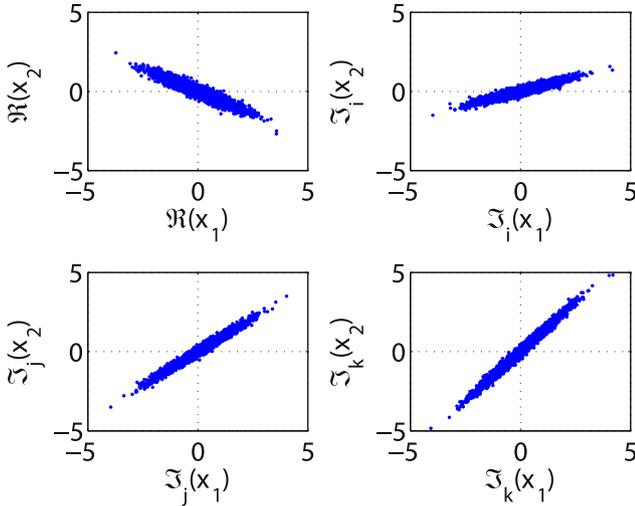
Fig. 7. Scatter plots showing high correlation between the correlated x_1 and x_2 variables.
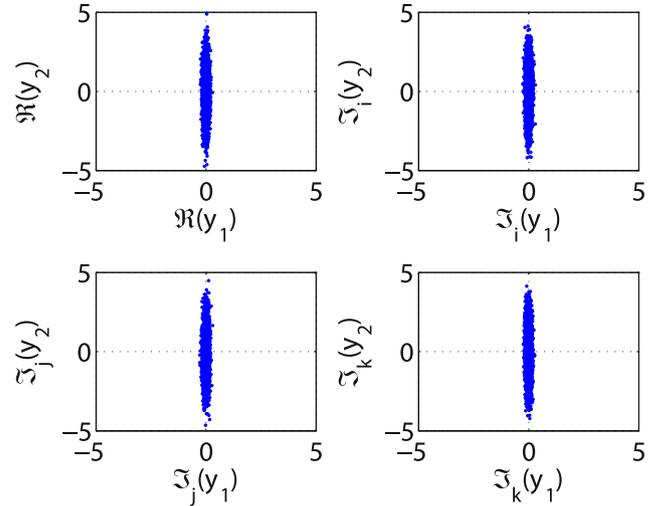
Fig. 8. Scatter plots showing no correlation between the transformed y_1 and y_2 variables.

approximation of **X** in the form

$$\hat{\mathbf{X}} = \sum_{m=1}^{P} \boldsymbol{\phi}_m^H \mathbf{y}_m, \qquad \text{where} \quad P < M$$

which is illustrated in Fig. 6. This is analogous to the real domain where the SVD provides a change of coordinates to yield uncorrelated variables for which the variance is monotonically non-increasing.

To illustrate the potential of the QAUT in rank reduction of quaternion data, a 2-D quaternion process, $\mathbf{x}_1$, $\mathbf{x}_2$, was generated where $\mathbf{x}_1$ and $\mathbf{x}_2$ were highly correlated as shown in Fig. 7, which plots the component relationship in the real, $\imath$, $\jmath$ and κ planes. The QAUT decorrelated $\mathbf{x}_1$, $\mathbf{x}_2$ into $\mathbf{y}_1$ and $\mathbf{y}_2$. Fig. 8 shows significant variation only in the $\mathbf{y}_2$ axis while the $\mathbf{y}_1$ variation corresponds to noise, indicating that the true dimensionality of the process is one rather than two.

Rank reduction is commonly used in chemometrics where many sets of measurements are required to examine a process and the important information may not be directly observable. For example, Near Infrared Spectra (NIR) are used to examine the chemical content of a data sample and are often highly correlated or collinear. We applied the QAUT to the NIR obtained from 80 samples of corn using three different spectrometers [53]. A pure quaternion variable was formed for each corn sample by combining its spectra from the three spectrometers. The QAUT diagonalised the data covariance matrices successfully to show that one quaternion variable was sufficient to express most of the information, as the largest diagonal element was 39 compared to 0.01 for the next largest, as shown in Fig. 9, which displays the logarithm of the ratio of variance explained by each variable, compared to the total variance in the data. The insert plot in Fig. 9 shows the cumulative ratio of variance explained by including further variables for the original and transformed data. Observe that the transformed
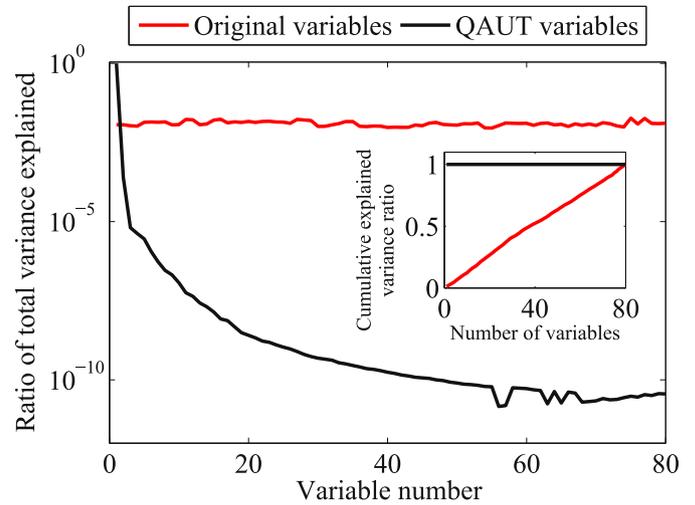
Fig. 9. The logarithm of the ratio of the variance explained by each variable to the total variance. *Insert diagram*: The cumulative ratio of variance against the number of variables for the original and decorrelated corn data.

data with much fewer variables explain the same information as the original, thus verifying the QAUT.

5.5. QAUT for imbalance detection in three-phase power systems

As shown in Section 4, when the components of data are highly correlated, the QAUT diagonalises all complementary covariance matrices at the expense of a small error, however, this error increases for lower degrees of correlation. We now show a way to exploit this data dependence of the QAUT and the structure of the
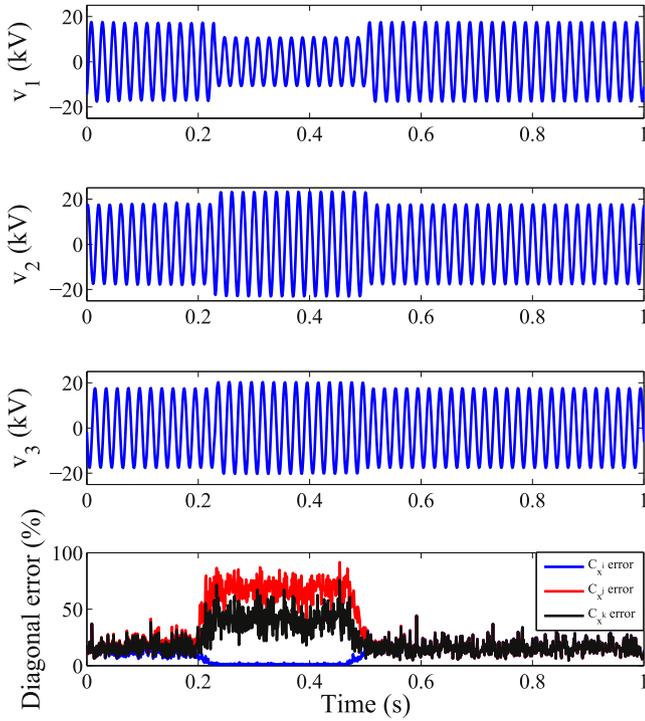
Fig. 10. Three-phase power signal and imbalance detection.

complementary covariance matrices to detect imbalance in a three-phase power system.

In general, the instantaneous voltages of a three-phase power system are given by Talebi and Mandic [54]

$$v_1(n) = A_1 \sin(2\pi f \Delta T n + \varphi_1)$$
$$v_2(n) = A_2 \sin(2\pi f \Delta T n + \varphi_2)$$
$$v_3(n) = A_3 \sin(2\pi f \Delta T n + \varphi_3) \quad (29)$$

where A_1, A_2, A_3 are the instantaneous amplitudes, φ_1, φ_2, φ_3 the instantaneous phases, f the system frequency, and ΔT the sampling interval. In a balanced three-phase system, $A_1 = A_2 = A_3$, $\varphi_2 - \varphi_1 = \frac{2\pi}{3}$, $\varphi_3 - \varphi_2 = \frac{2\pi}{3}$, and the correlation between each component is equivalent [55]. The power grid is usually designed to operate optimally under balanced conditions; however, faults in the power system can cause imbalanced operating conditions that propagate through the network, threatening its stability. Therefore, it is important in fault detection and mitigation applications to identify the incidences when the power grid is operating in an unbalanced fashion [56].

Pure quaternions have been used to deal with data recordings from all three phases simultaneously as $x(n) = v_1(n)\imath + v_2(n)\jmath + v_3(n)\kappa$ [54]. Here we organise the recorded data into a pure quaternion vector, $\mathbf{x} = \mathbf{v}_1\imath + \mathbf{v}_2\jmath + \mathbf{v}_3\kappa$, where $\mathbf{v}_m = [v_m(1), v_m(2), \ldots, v_m(n)]^T$, $m = 1, 2, 3$. We can then apply the univariate QAUT to $\mathbf{x}$ and employ the diagonalisation errors of the complementary covariance matrices as an indicator of power imbalance. This can be explained by the overall diagonalisation error given by

$$\begin{aligned}\Im\big[x_l^*(n_1)x_l(n_2)\big] = &[v_2(n_2)v_3(n_1) - v_2(n_1)v_3(n_2)]\imath \\ &+ [v_1(n_1)v_3(n_2) - v_1(n_2)v_3(n_1)]\jmath \\ &+ [v_1(n_2)v_2(n_1) - v_1(n_1)v_2(n_2)]\kappa \quad (30)\end{aligned}$$

which is a reduced form of (24) and is positively correlated with the diagonalisation error. When the system is balanced, the covariances $E\{\mathbf{v}_p\mathbf{v}_q^T\}$ are equivalent for all phases p, $q \in \{1, 2, 3\}$, so the three imaginary parts of $x_l^*(n_1)x_l(n_2)$ in (30) are equal in absolute

value, leading to equal diagonalisation errors of the three complementary covariance matrices. On the other hand, when the system becomes unbalanced, the correlation degrees between the three phase voltages are different, and hence the three imaginary parts of $x_l^*(n_1)x_l(n_2)$ differ, resulting in different diagonalisation errors of the three complementary covariance matrices. This allows the QAUT to be used in an imbalance detection application. To this end, we applied the QAUT to a sliding window of 0.04 seconds of a three-phase power system for which a fault occurred at 0.25 s lasting for 0.25 s before the system returned to a balanced state. During this fault, $\mathbf{v}_1$ decreased significantly in amplitude whereas $\mathbf{v}_2$ and $\mathbf{v}_3$ increased in amplitude ($\mathbf{v}_2 > \mathbf{v}_3$). Fig. 10 shows that the diagonalisation error for $\mathbf{C}_{\mathbf{x}^J}$ and $\mathbf{C}_{\mathbf{x}^\kappa}$ increased while the error for $\mathbf{C}_{\mathbf{x}^l}$ decreased and therefore analysing the difference between the errors is an appropriate detection method for the power imbalance. This can be explained via (30) and (23). The decrease in the amplitude of v_1 and the increase in the amplitudes of v_2 and v_3 induced an increased $\imath$-imaginary part and decreased $\jmath$-imaginary and κ-imaginary parts of $x_l^*(n_1)x_l(n_2)$, so that the diagonalisation error of $\mathbf{C}_{\mathbf{x}^l}$ declined and the diagonalisation errors of $\mathbf{C}_{\mathbf{x}^J}$ and $\mathbf{C}_{\mathbf{x}^\kappa}$ rose.

6. Conclusion

Novel simultaneous matrix factorisation techniques for the covariance and complementary covariance matrices, which arise in widely linear quaternion algebra, have been introduced. The diagonalisation of quaternion symmetric matrices has been first addressed, as the Takagi factorisation of complex symmetric matrices cannot be readily extended to the quaternion domain. We have shown that the symmetry is insufficient for a quaternion matrix $\mathbf{A}$ to be diagonalisable, and instead the condition $(\mathbf{AA}^*)^* = \mathbf{A}^*\mathbf{A}$ must also hold, which in general is not the case. The simultaneous diagonalisation of two quaternion covariance matrices, quaternion uncorrelating transform (QUT), has then been introduced as an analogue to the strong uncorrelating transform (SUT) in the complex domain. It has also been shown that for improper quaternion data, a typical case in quaternion signal processing applications, a single eigendecomposition of the covariance matrix is sufficient to diagonalise approximately the three complementary matrices. The analysis of the so introduced quaternion approximate uncorrelating transform (QAUT) has demonstrated its usefulness for improper quaternion signals, a typical case in practical applications. Simulation studies on synthetic and real-world signals support the proposed QUT and QAUT techniques.

Appendix A. Properties of quaternion matrices

For two general quaternion matrices, $\mathbf{A}$ and $\mathbf{B}$, the following properties hold [42]:

P1. $\mathbf{A}^{T*} = (\mathbf{A}^*)^T = (\mathbf{A}^T)^*$ and $\mathbf{A}^{\alpha*} = (\mathbf{A}^*)^\alpha = (\mathbf{A}^\alpha)^*$,

P2. $\mathbf{A}^{\alpha T} = (\mathbf{A}^\alpha)^T = (\mathbf{A}^T)^\alpha$ and $\mathbf{A}^{\alpha H} = (\mathbf{A}^\alpha)^H = (\mathbf{A}^H)^\alpha$,

P3. $(\mathbf{A}^\alpha)^\beta = (\mathbf{A}^\beta)^\alpha = \mathbf{A}^\gamma$ for distinct $\alpha, \beta, \gamma \in \{\imath, \jmath, \kappa\}$,

P4. $(\mathbf{AB})^* \neq \mathbf{A}^*\mathbf{B}^*$,

P5. $(\mathbf{AB})^T \neq \mathbf{B}^T\mathbf{A}^T$,

P6. $(\mathbf{AB})^H = \mathbf{B}^H\mathbf{A}^H$,

P7. $(\mathbf{AB})^\alpha = \mathbf{A}^\alpha\mathbf{B}^\alpha$,

P8. $(\mathbf{AB})^{\alpha H} = \mathbf{B}^{\alpha H}\mathbf{A}^{\alpha H}$.

Appendix B. $\mathbb{C}^\alpha$-improper quaternion random vectors

For distinct α, $\beta \in \{\imath, \jmath, \kappa\}$, the Cayley–Dickson construction allows for the complex representation of a quaternion random vector $\mathbf{x}$, as

$$\mathbf{x} = \mathbf{z}_1 + \mathbf{z}_2\beta$$

where $\mathbf{z}_1$ and $\mathbf{z}_2$ are complex random vectors defined in the plane spanned by $\{1, \alpha\}$. The quaternion random vector $\mathbf{x}$ is $\mathbb{C}^\alpha$-improper if and only if both of the following two conditions are fulfilled [47]:

1) $\mathbf{z}_1$ and $\mathbf{z}_2$ are proper complex vectors, which is achieved when their real and imaginary parts are uncorrelated and with the same variance,
2) $\mathbf{z}_1$ and $\mathbf{z}_2$ have different variances.

Appendix C. Symmetric quaternion matrices

Consider a 2×2 symmetric quaternion matrix $\mathbf{A}$, whereby

$$\mathbf{AA}^* = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} a_{11}^* & a_{12}^* \\ a_{21}^* & a_{22}^* \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}a_{11}^* + a_{12}a_{21}^* & a_{11}a_{12}^* + a_{12}a_{22}^* \\ a_{21}a_{11}^* + a_{22}a_{21}^* & a_{21}a_{12}^* + a_{22}a_{22}^* \end{bmatrix}$$

$$\mathbf{A}^*\mathbf{A} = \begin{bmatrix} a_{11}^* & a_{12}^* \\ a_{21}^* & a_{22}^* \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}^*a_{11} + a_{12}^*a_{21} & a_{11}^*a_{12} + a_{12}^*a_{22} \\ a_{21}^*a_{11} + a_{22}^*a_{21} & a_{21}^*a_{12} + a_{22}^*a_{22} \end{bmatrix}$$

By considering the symmetry $\mathbf{A} = \mathbf{A}^T$ ($a_{12} = a_{21}$), the matrices $(\mathbf{AA}^*)^*$ and $\mathbf{A}^*\mathbf{A}$ can be structured as

$$(\mathbf{AA}^*)^* = \begin{bmatrix} |a_{11}|^2 + |a_{12}|^2 & a_{12}a_{11}^* + a_{22}a_{12}^* \\ a_{11}a_{12}^* + a_{12}a_{22}^* & |a_{12}|^2 + |a_{22}|^2 \end{bmatrix}$$

$$\mathbf{A}^*\mathbf{A} = \begin{bmatrix} |a_{11}|^2 + |a_{12}|^2 & a_{11}^*a_{12} + a_{12}^*a_{22} \\ a_{12}^*a_{11} + a_{22}^*a_{12} & |a_{12}|^2 + |a_{22}|^2 \end{bmatrix}$$

Note that $(\mathbf{AA}^*)^* = \mathbf{A}^*\mathbf{A}$ if their off-diagonal elements are commutative, implying that either the diagonal or off-diagonal elements of $\mathbf{A}$ are real-valued.

References

[1] L. Krieger, F. Grigoli, Optimal reorientation of geophysical sensors: a quaternion-based analytical solution, Geophysics 80 (2) (2015) F19–F30.
[2] C. Cheong Took, G. Strbac, K. Aihara, D.P. Mandic, Quaternion-valued short--term joint forecasting of three-dimensional wind and atmospheric parameters, Renew. Energy 36 (6) (2011) 1754–1760.
[3] H. Fourati, N. Manamanni, L. Afilal, Y. Handrich, Complementary observer for body segments motion capturing by inertial and magnetic sensors, IEEE/ASME Trans. Mechatron. 19 (1) (2014) 149–157, doi:10.1109/TMECH.2012.2225151.
[4] J.B. Kuipers, Quaternions and Rotation Sequences: A Primer with Applications to Orbits, Aerospace and Virtual Reality, Princeton University Press, 2002.
[5] J. Tao, W. Chang, Adaptive beamforming based on complex quaternion processes, Math. Probl. Eng. 2014 (2014). Article ID 291249, 10 pages, doi: 10.1155/2014/291249.
[6] B. Chen, Q. Liu, X. Sun, X. Li, H. Shu, Removing gaussian noise for color images by quaternion representation and optimisation of weights in non-local means filter, IET Image Proc. 8 (10) (2014) 591–600.
[7] C. Cheong Took, D.P. Mandic, The quaternion LMS algorithm for adaptive filtering of hypercomplex processes, IEEE Trans. Signal Process. 57 (4) (2009) 1316–1327, doi:10.1109/TSP.2008.2010600.
[8] J. Vía, D. Palomar, L. Vielva, I. Santamaría, Quaternion ICA from second-order statistics, IEEE Trans. Signal Process. 59 (4) (2011) 1586–1600, doi:10.1109/TSP.2010.2101065.
[9] S. Javidi, C. Took, D.P. Mandic, Fast independent component analysis algorithm for quaternion valued signals, IEEE Trans. Neural Netw. 22 (12) (2011) 1967–1978, doi:10.1109/TNN.2011.2171362.
[10] Y. Xia, C. Jahanchahi, D.P. Mandic, Quaternion-valued echo state networks, IEEE Trans. Neural Netw. Learn. Syst. 26 (4) (2015) 663–673.
[11] T. Minemoto, T. Isokawa, H. Nishimura, N. Matsui, Feedforward neural network with random quaternionic neurons, Signal Process. 136 (2017) 59–68.
[12] T.A. Ell, N. Le Bihan, S.J. Sangwine, Quaternion Fourier Transforms for Signal and Image Processing, John Wiley & Sons, 2014.
[13] L. De Lathauwer, A link between the canonical decomposition in multilinear algebra and simultaneous matrix diagonalization, SIAM J. Matrix Anal. Appl. 28 (3) (2006) 642–666.
[14] A. Cichocki, D.P. Mandic, G. Zhou, Q. Zhao, C. Caiafa, H. Phan, L. De Lathauwer, Tensor decompositions for signal processing applications: from two-way to multiway component analysis, IEEE Signal Process. Mag. 32 (2) (2015) 145–163.
[15] L. De Lathauwer, B. De Moor, On the blind separation of non-circular sources, in: Proceedings of European Signal Processing Conference, 2002, pp. 1–4.
[16] J. Eriksson, V. Koivunen, Complex random vectors and ICA models: identifiability, uniqueness, and separability, IEEE Trans. Inf. Theory 52 (3) (2006) 1017–1029, doi:10.1109/TIT.2005.864440.
[17] P.J. Schreier, L.L. Scharf, Statistical Signal Processing of Complex-Valued Data: The Theory of Improper and Noncircular Signals, Cambridge University Press, 2010.
[18] T. Adalı, S. Haykin, Adaptive Signal Processing: Next Generation Solutions, Wiley-IEEE Press, 2010.
[19] A. Yeredor, Performance analysis of the strong uncorrelating transformation in blind separation of complex-valued sources, IEEE Trans. Signal Process. 60 (1) (2012) 478–483.
[20] J. Vía, D. Ramírez, I. Santamaría, Properness and widely linear processing of quaternion random vectors, IEEE Trans. Inf. Theory 56 (7) (2010) 3502–3515.
[21] C. Cheong Took, D.P. Mandic, Augmented second-order statistics of quaternion random signals, Signal Process. 91 (2) (2011) 214–224.
[22] D.P. Mandic, V.S.L. Goh, Complex Valued Nonlinear Adaptive Filters: Noncircularity, Widely Linear and Neural Models, Wiley-Blackwell, 2009.
[23] R.A. Horn, C.R. Johnson, Matrix Analysis, Cambridge University Press, 1990.
[24] E. Ollila, V. Koivunen, Complex ICA using generalized uncorrelating transform, Signal Process. 89 (4) (2009) 365–377.
[25] C. Cheong Took, S. Douglas, D.P. Mandic, On approximate diagonalization of correlation matrices in widely linear signal processing, IEEE Trans. Signal Process. 60 (3) (2012) 1469–1473.
[26] C. Cheong Took, S. Douglas, D.P. Mandic, Maintaining the integrity of sources in complex learning systems: intraference and the correlation preserving transform, IEEE Trans. Neural Netw. Learn. Syst. 26 (3) (2015) 500–509.
[27] S.C. Douglas, D.P. Mandic, Performance analysis of the conventional complex LMS and augmented complex LMS algorithms, in: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2010, pp. 3794–3797.
[28] D.P. Mandic, S. Kanna, S.C. Douglas, Mean square analysis of the CLMS and ACLMS for non-circular signals: the approximate uncorrelating transform approach, in: Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2015, pp. 3531–3535.
[29] Y. Xia, D.P. Mandic, Complementary mean square analysis of augmented CLMS for second-order noncircular Gaussian signals, IEEE Signal Process. Lett. 24 (9) (2017) 1413–1417.
[30] Y. Xia, D.P. Mandic, A full mean square analysis of CLMS for second-order noncircular inputs, IEEE Trans. Signal Process. 65 (21) (2017) 5578–5590.
[31] H. Shen, M. Kleinsteuber, Complex blind source separation via simultaneous strong uncorrelating transform, in: Latent Variable Analysis and Signal Separation. LVA/ICA 2010, in: Lecture Notes in Computer Science, 6365, Berlin, Heidelberg, 2010, pp. 287–294.
[32] D. Ramírez, P.J. Schreier, J. Vía, I. Santamaría, GLRT for testing separability of a complex-valued mixture based on the strong uncorrelating transform, in: Proceedings of IEEE International Workshop on Machine Learning for Signal Processing, 2012, pp. 1–6, doi:10.1109/MLSP.2012.6349785.
[33] S.C. Douglas, J. Eriksson, V. Koivunen, Adaptive estimation of the strong uncorrelating transform with applications to subspace tracking, in: Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IV, 2006, pp. 941–944, doi:10.1109/ICASSP.2006.1661125.
[34] G. Okopal, S. Wisdom, L. Atlas, Speech analysis with the strong uncorrelating transform, IEEE/ACM Trans. Audio Speech Lang. Process. 23 (11) (2015) 1858–1868, doi:10.1109/TASLP.2015.2456426.
[35] A. Bojanczyk, G. Cybenko, Linear Algebra for Signal Processing, Springer Science & Business Media, 1995.
[36] A. Hjørungnes, Complex-Valued Matrix Derivatives: With Applications in Signal Processing and Communications, Cambridge University Press, Cambridge, UK, 2011.
[37] F. Zhang, Quaternions and matrices of quaternions, Linear Algebra Appl. 251 (1997) 21–57.
[38] L. Zou, Y. Jiang, J. Wu, Location for the right eigenvalues of quaternion matrices, J. Appl. Math. Comput. 38 (1) (2012) 71–83.
[39] D. Xu, D.P. Mandic, The theory of quaternion matrix derivatives, IEEE Trans. Signal Process. 63 (3) (2015) 1543–1556.
[40] S. Sangwine, N.L. Bihan, Quaternion singular value decomposition based on bidiagonalization to a real or complex matrix using quaternion householder transformations, Appl. Math. Comput. 182 (2006) 727–738.
[41] N.L. Bihan, S. Sangwine, Jacobi method for quaternion matrix singular value decomposition, Appl. Math. Comput. 187 (2007) 1265–1271.
[42] C. Cheong Took, D.P. Mandic, F. Zhang, On the unitary diagonalisation of a special class of quaternion matrices, Appl. Math. Lett. 24 (11) (2011) 1806–1809.
[43] S. Enshaeifar, C. Cheong Took, S. Sanei, D.P. Mandic, Novel quaternion matrix factorisation, in: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2016, pp. 3946–3950.
[44] T.A. Ell, S.J. Sangwine, Quaternion involutions and anti-involutions, Comput. Math. Appl. 53 (1) (2007) 137–143.
[45] C. Cheong Took, D.P. Mandic, A quaternion widely linear adaptive filter, IEEE Trans. Signal Process. 58 (8) (2010) 4427–4431.
[46] Y. Xia, C. Jahanchahi, T. Nitta, D.P. Mandic, Performance bounds of quaternion estimators, IEEE Trans. Neural Netw. Learn. Syst. 26 (12) (2015) 3287–3292.
[47] P.O. Amblard, N. Le Bihan, On properness of quaternion valued random variables, in: Proceedings of the IMA International Conference on Mathematics in Signal Processing, 2004, pp. 23–26.

204

M. Xiang et al./Signal Processing 148 (2018) 193–204

[48] L. Rodman, Topics in Quaternion Linear Algebra, Princeton University Press, 2014.
[49] F. Zhang, The Schur Complement and its Applications, Springer, 2006.
[50] J.P. Ward, Quaternions and Cayley numbers: Algebra and Applications, Springer, 1997.
[51] C. Park, C. Cheong Took, D.P. Mandic, Augmented complex common spatial patterns for classification of noncircular EEG from motor imagery tasks, IEEE Trans. Neural Syst. Rehabil. Eng. 22 (1) (2014) 1–10, doi:10.1109/TNSRE.2013.2294903.
[52] S. Enshaeifar, C. Cheong Took, C. Park, D.P. Mandic, Quaternion common spatial patterns, IEEE Trans. Neural Syst. Rehabil. Eng. 25 (8) (2017) 1278–1286.
[53] NIR of corn samples for standardization benchmarking, 2005, http://www.eigenvector.com/data/Corn/.
[54] S.P. Talebi, D.P. Mandic, A quaternion frequency estimator for three-phase power systems, in: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2015, pp. 3956–3960.
[55] Y. Xia, S.C. Douglas, D.P. Mandic, Adaptive frequency estimation in smart grid applications: exploiting noncircularity and widely linear adaptive estimators, IEEE Signal Process. Mag. 29 (5) (2012) 44–54, doi:10.1109/MSP.2012.2183689.
[56] S. Kanna, D.H. Dini, Y. Xia, S.Y. Hui, D.P. Mandic, Distributed widely linear Kalman filtering for frequency estimation in power networks, IEEE Trans. Signal Inf. Process. Netw. 1 (1) (2015) 45–57.