

NOVEL QUATERNION MATRIX FACTORISATIONS

S. Enshaeifar, C. Cheong Took, S. Sanei, and D. P. Mandic

ABSTRACT

The recent introduction of η -Hermitian matrices $\mathbf{A} = \mathbf{A}^{\eta H}$ has opened a new avenue of research in quaternion signal processing. However, the exploitation of this matrix structure has been limited, perhaps due to the lack of joint diagonalisation methodologies of these matrices. As such, we propose novel decompositions of η -Hermitian matrices to address this shortcoming in the literature. As an application, we consider a blind source separation problem in the form of an Alamouti-based communication system. Simulation studies demonstrate the effectiveness of our proposed joint diagonalisation technique and indicate that our approach is particularly useful when the sources are correlated.

Index Terms— Joint diagonalisation, quaternion domain, uncorrelating transform.

1. INTRODUCTION

Diagonalisation of covariance matrices is a pivotal procedure in a number of the statistical signal processing algorithms, such as principal component analysis (PCA) and blind source separation (BSS) [1]. For example, joint diagonalisation of covariance matrices of multivariate data has been an instrument to decorrelate the data channels in BSS.

Recent advances in complex statistics have highlighted the necessity of widely linear modelling of signals, to exploit the power difference or correlation between the data channels. For this purpose, ‘augmented’ statistics have been established to incorporate the covariance and complementary covariance matrices, and to cater for the second order noncircular (improper) processes in statistical signal processing [2]. Important contributions to augmented statistics are the general uncorrelating transform (GUT) [3] and strong-uncorrelating transform (SUT) [4, 5] which diagonalise the covariance and pseudo-covariance matrices simultaneously.

The SUT has been established for both single and multichannel complex-valued processes. However, the advances in multidimensional sensor technologies have underlined the need for signal processing algorithms in the quaternion domain (\mathbb{H}) due to its potential for modelling of three- and four-dimensional data [6]. Similar to the complex domain, the use of augmented statistics and associated widely linear modelling is also required for quaternion signals to account for the full second order information. Recent developments include the analysis of quaternion data via augmented quaternion statistics; for example, the unitary diagonalisation of quaternion matrices was introduced in [7] and [8] established quaternion fast independent component analysis (Q-FastICA) for the blind separation of both proper and improper quaternion-valued processes. Although research on quaternion signal processing has revived interest in several applications, there are still some shortcomings in the quaternion formulations, e.g. there is a lack of closed form solutions to perform

the simultaneous diagonalisation of the quaternion covariance matrices. To this end, we consider the simultaneous diagonalisation of a pair of covariance matrices in quaternion domain, paving the way for the quaternion uncorrelating transform¹ (QUT) to allow for the covariance and a desired complementary covariance to be diagonalised simultaneously.

2. SECOND-ORDER STATISTICS OF QUATERNION

Augmented statistics have been established to incorporate the complementary covariance matrices and exploit the complete second-order information. In complex domain, a basis vector is augmented by its conjugate² variables as $\mathbf{z}^a = [\mathbf{z}, \mathbf{z}^*]^T \in \mathbb{C}$ [9]. Thus, the augmented covariance matrix, $\mathbf{C}^a = E\{\mathbf{z}^a \mathbf{z}^{aH}\}$, takes into account the complete second-order information. However, such convenient manipulation is not possible in quaternion domain - illustrated in the next section.

2.1. Quaternion statistics

Quaternions are hypercomplex numbers denoted by \mathbb{H} . A quaternion vector³ \mathbf{x} can be expressed in the Cartesian form as:

$$\mathbf{x} = \mathbf{x}_r + \iota \mathbf{x}_i + j \mathbf{x}_j + \kappa \mathbf{x}_k \quad (1)$$

where $\Re\{\mathbf{x}\} = \mathbf{x}_r$ is the scalar (real) part, $\Im\{\mathbf{x}\} = \mathbf{x} - \mathbf{x}_r$ is the vector (pure quaternion) part, and ι, j and κ are orthogonal unit vectors. An important notion for the quaternion domain is the so-called ‘quaternion involution’ which forms the basis for augmented quaternion statistics. An involution operator defines a self-inverse mapping about a unit quaternion $\eta \in \{\iota, j, \kappa\}$, given by [10, 11]

$$\begin{aligned} \mathbf{x}^\eta &= -\eta \mathbf{x} \eta \\ \text{e.g. } \mathbf{x}^\iota &= -\iota \mathbf{x} \iota = \mathbf{x}_r + \iota \mathbf{x}_i - j \mathbf{x}_j - \kappa \mathbf{x}_k \end{aligned} \quad (2)$$

Note that the involution designates a rotation along a single unit axis, while the quaternion conjugate operator $(\cdot)^*$ rotates along all three imaginary axes, and can be computed as

$$\begin{aligned} \mathbf{x}^* &= \Re\{\mathbf{x}\} - \Im\{\mathbf{x}\} = \mathbf{x}_r - \iota \mathbf{x}_i - j \mathbf{x}_j - \kappa \mathbf{x}_k \\ &= \frac{1}{2}(\mathbf{x}^\iota + \mathbf{x}^j + \mathbf{x}^\kappa - \mathbf{x}) \end{aligned} \quad (3)$$

According to (2) and (3), the correspondence between the elements of a quaternion variable in \mathbb{H} and the elements of a quadrivariate vector in \mathbb{R}^4 can be obtained as [9]:

$$\Re\{\mathbf{x}\} = \frac{1}{2}(\mathbf{x} + \mathbf{x}^*), \quad \Im_\eta\{\mathbf{x}\} = \frac{1}{2\eta}(\mathbf{x} - \mathbf{x}^{\eta*}) \quad \eta \in \{\iota, j, \kappa\} \quad (4)$$

Therefore, quaternion statistics should generally include all quaternion involutions $\mathbf{x}^\iota, \mathbf{x}^j$, and \mathbf{x}^κ to access to the complete second order statistical information. In other words, second order quaternion statistics should contain ι -, j -, and κ - covariance matrices as

¹QUT is analogous to the SUT for complex matrices [4].

²The operators $(\cdot)^T$, $(\cdot)^*$ and $(\cdot)^H$ represent transpose, conjugate and Hermitian (conjugate transpose) respectively.

³Throughout this paper, we assume zero mean quaternion variables with unit variances. This does not affect the generality of our results.

S. Enshaeifar, C. Cheong Took, and S. Sanei are with the Faculty of Engineering and Physical Sciences, University of Surrey, GU2 7XH, U.K.

D. P. Mandic is with the Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2AZ, U.K.

well as the standard covariance matrix. These matrices are also called the complementary covariance matrices which are obtained as $\mathbf{C}_{\mathbf{x}\eta} = E\{\mathbf{x}\mathbf{x}^{\eta H}\}$, $\eta \in \{\iota, j, \kappa\}$.

Remark#1: The standard covariance matrix $\mathbf{C}_{\mathbf{x}} = E\{\mathbf{x}\mathbf{x}^H\}$ is a Hermitian matrix, while a complementary covariance matrix $\mathbf{C}_{\mathbf{x}\eta} = E\{\mathbf{x}\mathbf{x}^{\eta H}\}$ is an η -Hermitian matrix.

Similar to the complex domain [5], diagonalisation of the above covariance matrices is fundamental in several statistical signal processing algorithms, which is next discussed.

3. SIMULTANEOUS DIAGONALISATION

To consider the simultaneous diagonalisation of the quaternion correlation matrices, we first consider the following propositions and lemma:

Proposition 1: If $\mathbf{A}, \mathbf{B} \in \mathbb{H}$ are η -Hermitian and \mathbf{A} is nonsingular, \mathbf{A}, \mathbf{B} are simultaneously diagonalisable if and only if $\mathbf{D} = \mathbf{A}^{-1}\mathbf{B}$ is normal.

Proof. Consider a unitary matrix \mathbf{M} such that $\mathbf{\Lambda}_a = \mathbf{M}^{\eta H}\mathbf{A}\mathbf{M}$ and $\mathbf{\Lambda}_b = \mathbf{M}^{\eta H}\mathbf{B}\mathbf{M}$ are both diagonal. Thus, $\mathbf{A}^{-1} = \mathbf{M}\mathbf{\Lambda}_a^{-1}\mathbf{M}^{\eta H}$ and $\mathbf{B} = \mathbf{M}^{\eta H}\mathbf{\Lambda}_b\mathbf{M}^H$ where $\mathbf{A}^{-1}\mathbf{B} = \mathbf{M}(\mathbf{\Lambda}_a^{-1}\mathbf{\Lambda}_b)\mathbf{M}^H$ is unitarily diagonalisable, i.e. normal. ■

Proposition 2: If $\mathbf{A} = \begin{bmatrix} \mathbf{B} & \mathbf{C} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \in \mathbb{H}$, then \mathbf{A} is normal if and only if \mathbf{B} is normal and $\mathbf{C} = \mathbf{0}$.

Proof. \mathbf{A} is normal if $\mathbf{A}\mathbf{A}^H = \mathbf{A}^H\mathbf{A}$ where:

$$\mathbf{A}\mathbf{A}^H = \begin{bmatrix} \mathbf{B}\mathbf{B}^* + \mathbf{C}\mathbf{C}^* & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \quad \mathbf{A}^H\mathbf{A} = \begin{bmatrix} \mathbf{B}^*\mathbf{B} & \mathbf{B}^*\mathbf{C} \\ \mathbf{C}^*\mathbf{B} & \mathbf{C}^*\mathbf{C} \end{bmatrix}$$

therefore, $\mathbf{A}\mathbf{A}^H = \mathbf{A}^H\mathbf{A}$ holds if $\mathbf{C} = \mathbf{0}$, and $\mathbf{B}\mathbf{B}^* = \mathbf{B}^*\mathbf{B}$, i.e. \mathbf{B} is normal. ■

Lemma 1: An η -Hermitian matrix \mathbf{A} can be factorised as $\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{\eta H}$, where \mathbf{Q} denotes a quaternion unitary matrix and $\mathbf{\Lambda}$ is a real-valued nonnegative diagonal matrix. This lemma is fully described in [7, 12].

3.1. Diagonalisation

The introduction of η -Hermitian matrices implies that there are three kinds of simultaneous diagonalisation of quaternion matrices, namely when both matrices are Hermitian or η -Hermitian and when one matrix is Hermitian and the other is η -Hermitian; we therefore consider such diagonalisations in the form of the following corollary.

Corollary 1. For given matrices $\mathbf{A}, \mathbf{B} \in \mathbb{H}$:

(a) If \mathbf{A} and \mathbf{B} are both Hermitian, there exists a matrix $\mathbf{M} \in \mathbb{H}$ such that $\mathbf{M}^H\mathbf{A}\mathbf{M}$ and $\mathbf{M}^H\mathbf{B}\mathbf{M}$ are both diagonal if and only if $\mathbf{A}\mathbf{B}$ is Hermitian, i.e. $\mathbf{A}\mathbf{B} = \mathbf{B}\mathbf{A}$.

(b) If \mathbf{A} and \mathbf{B} are both η -Hermitian, there exists a unitary matrix $\mathbf{M} \in \mathbb{H}$ such that $\mathbf{M}^{\eta H}\mathbf{A}\mathbf{M}$ and $\mathbf{M}^{\eta H}\mathbf{B}\mathbf{M}$ are both diagonal if and only if $\mathbf{A}\mathbf{B}^{\eta}$ is normal, i.e. $\mathbf{A}\mathbf{B}^{\eta}\mathbf{B}^{\eta H}\mathbf{A}^H = \mathbf{B}^{\eta H}\mathbf{A}^H\mathbf{A}\mathbf{B}^{\eta}$.

(c) If \mathbf{A} is Hermitian and \mathbf{B} is η -Hermitian, there exists a matrix $\mathbf{M} \in \mathbb{H}$ such that $\mathbf{M}^H\mathbf{A}\mathbf{M}$ and $\mathbf{M}^{\eta H}\mathbf{B}\mathbf{M}$ are both diagonal if and only if $\mathbf{B}\mathbf{A}$ is η -Hermitian, i.e. $\mathbf{B}\mathbf{A} = (\mathbf{B}\mathbf{A})^{\eta H} = \mathbf{A}^{\eta H}\mathbf{B}^{\eta H} = \mathbf{A}^{\eta}\mathbf{B}$.

3.1.1. Proof of Corollary 1

(a) If $\mathbf{M}^H\mathbf{A}\mathbf{M} = \mathbf{\Lambda}_a$ and $\mathbf{M}^H\mathbf{B}\mathbf{M} = \mathbf{\Lambda}_b$ are diagonal, we have $\mathbf{A} = \mathbf{M}\mathbf{\Lambda}_a\mathbf{M}^H$ and $\mathbf{B} = \mathbf{M}\mathbf{\Lambda}_b\mathbf{M}^H$. Thus:

$$\mathbf{A}\mathbf{B} = \mathbf{M}\mathbf{\Lambda}_a\mathbf{M}^H\mathbf{M}\mathbf{\Lambda}_b\mathbf{M}^H = \mathbf{M}\mathbf{\Lambda}_b(\mathbf{M}^H\mathbf{M})\mathbf{\Lambda}_a\mathbf{M}^H = \mathbf{B}\mathbf{A}$$

■

(b) If $\mathbf{M}^{\eta H}\mathbf{A}\mathbf{M} = \mathbf{\Lambda}_a$ and $\mathbf{M}^{\eta H}\mathbf{B}\mathbf{M} = \mathbf{\Lambda}_b$ are diagonal, then $\mathbf{A} = \mathbf{M}^{\eta}\mathbf{\Lambda}_a\mathbf{M}^{\eta H}$ and $\mathbf{B} = \mathbf{M}^{\eta}\mathbf{\Lambda}_b\mathbf{M}^{\eta H}$. Hence:

$$\mathbf{A}\mathbf{B}^{\eta} = \mathbf{M}^{\eta}\mathbf{\Lambda}_a\mathbf{M}^{\eta H}\mathbf{M}\mathbf{\Lambda}_b\mathbf{M}^{\eta H} = \mathbf{M}^{\eta}(\mathbf{\Lambda}_a\mathbf{\Lambda}_b)\mathbf{M}^{\eta H}$$

which is diagonalisable with a unitary matrix and it is normal. For the converse, consider the following scenarios:

(i) Suppose $\mathbf{A}\mathbf{B}^{\eta}$ is normal and \mathbf{A} is nonsingular. Since $\mathbf{A}\mathbf{B}^{\eta} = (\mathbf{A}^{-1})^{-1}\mathbf{B}^{\eta}$ is normal, \mathbf{A}^{-1} and \mathbf{B}^{η} are simultaneously unitary diagonalisable, see Proposition 1. Since \mathbf{A} and \mathbf{B} are η -Hermitian, according to Lemma 1, we have $\mathbf{A}^{-1} = \mathbf{M}\mathbf{\Lambda}_a^{-1}\mathbf{M}^{\eta H}$ and $\mathbf{B}^{\eta} = \mathbf{M}\mathbf{\Lambda}_b\mathbf{M}^{\eta H}$. Thus:

$$\mathbf{A} = \mathbf{M}^{\eta}\mathbf{\Lambda}_a\mathbf{M}^{\eta H} = \mathbf{M}^{\eta}\mathbf{\Lambda}_a(\mathbf{M}^{\eta})^{\eta H}$$

$$\mathbf{B} = \mathbf{M}^{\eta}\mathbf{\Lambda}_b\mathbf{M}^{\eta H} = \mathbf{M}^{\eta}\mathbf{\Lambda}_b(\mathbf{M}^{\eta})^{\eta H}$$

which are simultaneous diagonalisations of \mathbf{A} and \mathbf{B} .

(ii) Suppose $\mathbf{A}\mathbf{B}^{\eta}$ is normal and \mathbf{A} is singular. Hence, there is a unitary matrix $\mathbf{M} \in \mathbb{H}$ where $\mathbf{M}^{\eta H}\mathbf{A}\mathbf{M}$ is diagonal. The columns of \mathbf{M} can be permuted such that:

$$\mathbf{M}^{\eta H}\mathbf{A}\mathbf{M} = \begin{bmatrix} \mathbf{\Sigma} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \quad \mathbf{M}^{\eta H}\mathbf{B}\mathbf{M} = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{12}^{\eta H} & \mathbf{B}_{22} \end{bmatrix}$$

where $\mathbf{\Sigma}$ is a diagonal block matrix with vanishing η -components, and \mathbf{B}_{11} and \mathbf{B}_{22} have vanishing η -components in their diagonal elements. Thus, as \mathbf{M} is unitary:

$$\begin{aligned} (\mathbf{M}^{\eta H}\mathbf{A}\mathbf{M})(\mathbf{M}^{\eta H}\mathbf{B}\mathbf{M})^{\eta} &= \mathbf{M}^{\eta H}\mathbf{A}\mathbf{B}^{\eta}\mathbf{M}^{\eta} \\ &= \begin{bmatrix} \mathbf{\Sigma}\mathbf{B}_{11}^{\eta} & \mathbf{\Sigma}\mathbf{B}_{12}^{\eta} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \end{aligned}$$

Since $\mathbf{A}\mathbf{B}^{\eta}$ is normal, $\mathbf{\Sigma}\mathbf{B}_{12}^{\eta} = \mathbf{0}$, see Proposition 2. As $\mathbf{\Sigma}$ is nonsingular, $\mathbf{B}_{12}^{\eta} = \mathbf{0}$ and we have:

$$\mathbf{M}^{\eta H}\mathbf{A}\mathbf{M} = \begin{bmatrix} \mathbf{\Sigma} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \quad \mathbf{M}^{\eta H}\mathbf{B}\mathbf{M} = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_{22} \end{bmatrix}$$

Therefore, if there is a unitary matrix \mathbf{M} to diagonalise \mathbf{A} , it is also sufficient for \mathbf{B} . ■

(c) If $\mathbf{M}^H\mathbf{A}\mathbf{M} = \mathbf{\Lambda}_a$ and $\mathbf{M}^{\eta H}\mathbf{B}\mathbf{M} = \mathbf{\Lambda}_b$ are diagonal, we have $\mathbf{A} = \mathbf{M}\mathbf{\Lambda}_a\mathbf{M}^H$ and $\mathbf{B} = \mathbf{M}^{\eta}\mathbf{\Lambda}_b\mathbf{M}^{\eta H}$. Thus,

$$\begin{aligned} \mathbf{B}\mathbf{A} &= \mathbf{M}^{\eta}\mathbf{\Lambda}_b\mathbf{M}^{\eta H}\mathbf{M}\mathbf{\Lambda}_a\mathbf{M}^H \\ &= \mathbf{M}^{\eta}\mathbf{\Lambda}_a(\mathbf{M}^{\eta H}\mathbf{M}^{\eta})\mathbf{\Lambda}_b\mathbf{M}^{\eta H} = \mathbf{A}^{\eta}\mathbf{B} \end{aligned}$$

■

3.1.2. Derivation of matrix \mathbf{M} in Corollary 1

The matrix \mathbf{M} in Corollary 1 can be derived as follows⁴:

(a) Since $\mathbf{A} = \mathbf{U}\mathbf{S}_a\mathbf{U}^H$ and \mathbf{B} are both Hermitian:

$$\mathbf{D} = \mathbf{S}_a^{-\frac{1}{2}}\mathbf{U}^H \quad \text{so} \left\{ \begin{array}{l} \mathbf{D}\mathbf{A}\mathbf{D}^H = \mathbf{I} \\ \mathbf{D}\mathbf{B}\mathbf{D}^H = \mathbf{W}\mathbf{\Lambda}_b\mathbf{W}^H \end{array} \right\}$$

consider $\mathbf{M} = \mathbf{W}^H\mathbf{D}$, thus:

$$\mathbf{M}\mathbf{A}\mathbf{M}^H = \mathbf{W}^H\mathbf{D}\mathbf{A}\mathbf{D}^H\mathbf{W} = \mathbf{I} = \mathbf{\Lambda}_a$$

$$\mathbf{M}\mathbf{B}\mathbf{M}^H = \mathbf{W}^H(\mathbf{D}\mathbf{B}\mathbf{D}^H)\mathbf{W} = \mathbf{\Lambda}_b$$

(b) If \mathbf{A} and \mathbf{B} are both η -Hermitian and $\mathbf{A}\mathbf{B}^{\eta}$ is normal, a single unitary matrix \mathbf{M} is sufficient to diagonalise them simultaneously. In general, by applying quaternion singular value decomposition (SVD) [6, 13] on \mathbf{A} , it can be rewritten as $\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^H$. Using Lemma 1, \mathbf{A} can be expressed as $\mathbf{A} = \mathbf{Q}\mathbf{S}\mathbf{Q}^{\eta H}$ in which $\mathbf{Q} = \mathbf{U}(\mathbf{D}^{\eta})^{\frac{1}{2}}$ and $\mathbf{U} = \mathbf{V}^{\eta}\mathbf{D}$ [7]. Thus, $\mathbf{M} = \mathbf{Q}^H$ diagonalises \mathbf{A} and \mathbf{B} .

(c) Since $\mathbf{A} = \mathbf{U}\mathbf{S}_a\mathbf{U}^H$ is Hermitian and \mathbf{B} is η -Hermitian:

$$\mathbf{D} = \mathbf{S}_a^{-\frac{1}{2}}\mathbf{U}^H \quad \text{so} \left\{ \begin{array}{l} \mathbf{D}\mathbf{A}\mathbf{D}^H = \mathbf{I} \\ \mathbf{D}\mathbf{B}\mathbf{D}^{\eta H} = \mathbf{W}\mathbf{\Lambda}_b\mathbf{W}^{\eta H} \end{array} \right\}$$

⁴In this section, \mathbf{I} is the identity matrix.

Consider $\mathbf{M} = \mathbf{W}^H \mathbf{D}$, thus:

$$\begin{aligned} \mathbf{MAM}^H &= \mathbf{W}^H \mathbf{DAD}^H \mathbf{W} = \mathbf{I} = \Lambda_a \\ \mathbf{MBM}^{\eta H} &= \mathbf{W}^H (\mathbf{DBD}^{\eta H}) \mathbf{W} = \Lambda_b \end{aligned}$$

3.2. Quaternion Uncorrelating Transform

Note that Corollary 1(c) leads to the quaternion uncorrelating transform (QUT) as follows:

Corollary 2. For a random quaternion-valued vector \mathbf{x} , there exists a QUT matrix \mathbf{M} which simultaneously whitens the covariance $\mathbf{C}_{\mathbf{y}} = \mathbf{I}$ and diagonalises the η -covariance $\mathbf{C}_{\mathbf{y}\eta} = \Lambda_\eta$, where $\mathbf{y} = \mathbf{M}\mathbf{x}$. The MATLAB⁵ implementation of the proposed QUT algorithm is included in Table 1.

Table 1: Pseudo-MATLAB implementation of QUT.

```
function [M, y] = QUT(x)
if size(x, 1) > size(x, 2)
    x = x.';
end
n = length(x);
C_x = (x*x')/n;
% Decompose the standard covariance matrix
[U, S] = svd(C_x);
% Derive the whitening matrix D and whitened data q
D = diag(diag(S).^ - 0.5) * U';
q = D * x;
% Generate the i, j, and kappa-covariances of the whitened data q
% The eta-involution in (2) is implemented by the 'invijk' function
C_i = (q * invijk(q, 'i'))/n;
C_j = (q * invijk(q, 'j'))/n;
C_kappa = (q * invijk(q, 'kappa'))/n;
% Analyse complementary covariances to select the imaginary
% unit with maximum correlation
c = [ norm(C_i - diag(diag(C_i)))
      norm(C_j - diag(diag(C_j)))
      norm(C_kappa - diag(diag(C_kappa))) ];
[~, f] = max(c.');
% Define the desired unit axis using f, i.e if f = 1 -> eta = i,
% f = 2 -> eta = j, and f = 3 -> eta = kappa.
% Apply QTakagi factorisation on the selected eta-covariance,
% such that C_eta = W * S_1 * invijk(W, 'eta')'. See [7]
[U_1, S_1, V_1] = svd(C_eta);
P = invijk(V_1, 'eta') * U_1;
W = U_1 * diag(sqrt(diag(invijk(P, 'eta'))));
% Calculate the QUT matrix M and the decorrelated output y
M = W' * D;
y = M * x
```

3.3. Properness and Quaternion Strong Uncorrelating Transform

The notion of properness or improperness is an important statistical property in both complex and quaternion domains. Similar to the complex domain, quaternion properness is characterised by the degree of correlation and power ratio of imaginary components with respect to the real component. However, the additional degrees of

⁵To use the MATLAB code, the modified version of quaternion toolbox 'qtfm' is required, which is available at http://www.surrey.ac.uk/cs/people/clive_cheong_tool/index.htm.

freedom in the quaternion domain means that there are two types of properness: \mathbb{H} -properness and \mathbb{C}^η -properness.

A quaternion vector \mathbf{x} is second order \mathbb{H} -proper if and only if all the complementary covariance matrices vanish, i.e. the vector \mathbf{x} is uncorrelated with its vector involutions \mathbf{x}^i , \mathbf{x}^j , \mathbf{x}^κ . Note that for a given quaternion vector \mathbf{x} the degree of \mathbb{H} -properness can be defined in terms of a circularity coefficient ρ which is calculated based on the ratio of the complementary covariances to the standard covariance matrix given by [8]

$$\rho = \frac{|E\{\mathbf{x}\mathbf{x}^i H\}| + |E\{\mathbf{x}\mathbf{x}^j H\}| + |E\{\mathbf{x}\mathbf{x}^\kappa H\}|}{3E\{\mathbf{x}\mathbf{x}^H\}} \quad \rho \in [0, 1] \quad (5)$$

where $\rho = 0$ defines a \mathbb{H} -proper, and $\rho = 1$ corresponds to \mathbb{H} -improper sources.

On the other hand, \mathbf{x} is \mathbb{C}^η -proper for an imaginary unit $\eta \in \{i, j, \kappa\}$ if and only if the vector \mathbf{x} is only correlated with \mathbf{x}^η . Thus, for a \mathbb{C}^η -proper process \mathbf{x} , only the η -covariance $\mathbf{C}_{\mathbf{x}\eta} = E\{\mathbf{x}\mathbf{x}^{\eta H}\}$ exists and QUT offers a convenient way to diagonalise both the covariance and the η -covariance matrices, as the other two covariance matrices vanish. For instance, for a \mathbb{C}^η -proper process \mathbf{x} , QUT leads to $\mathbf{C}_{\mathbf{y}} = \mathbf{I}$, $\mathbf{C}_{\mathbf{y}^i} = \Lambda_i$, and $\mathbf{C}_{\mathbf{y}^j} = \mathbf{C}_{\mathbf{y}^\kappa} = \Lambda = 0$, where $\mathbf{y} = \mathbf{M}\mathbf{x}$.

Remark#2: In the context of \mathbb{C}^η -properness, QUT is regarded as the quaternion strong-uncorrelating (Q-SUT) transform - an analogue to the work in the complex domain by Eriksson and Koivunen [4]. However, without the \mathbb{C}^η -properness condition, QUT is generally not robust for decorrelating purposes.

4. SIMULATIONS AND DISCUSSION

4.1. ARMA Simulation

The first experiment evaluated the performance of the Q-SUT as a source separation method under different levels of additive white noise. Simulations were performed on multivariate widely linear autoregressive moving average (ARMA) data. To this end, three uncorrelated quaternion-valued ARMA sources \mathbf{s} were generated such that each source was an \mathbb{C}^i -proper signal. Then, they were mixed using a 3×3 random matrix \mathbf{A} drawn from the standard normal distribution to obtain three mixtures as $\mathbf{x} = \mathbf{A}\mathbf{s}$. Overall, 50 sets of data were generated to compute the averaged results.

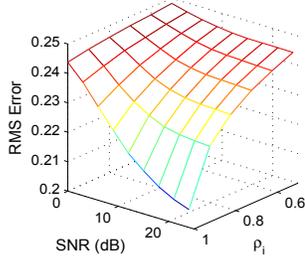
In order to assess the effect of noise on the performance accuracy, white noise was added so as to vary the signal to noise ratio (SNR) from 0 to 25 dB. The performance was also assessed in terms of the i -circularity coefficient, such that $E\{\mathbf{x}\mathbf{x}^H\} = 1$ was constant, and $\rho_i = |E\{\mathbf{x}\mathbf{x}^i H\}|$ was manually adjusted from 0.5 to 1, see (5).

To evaluate the separation accuracy, the root mean square (RMS) error was calculated among the original sources (\mathbf{s}) and the estimated sources ($\hat{\mathbf{x}}$) obtained via Q-SUT. Since Q-SUT suffers from the permutation problem, the 3-D sources (both original and estimated) were first averaged to provide a 1-D quaternion channel. Then, the RMS error was calculated among the corresponding sources.

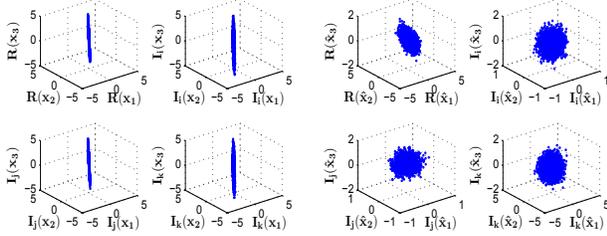
Fig. 1a illustrates the reconstruction error versus SNR and circularity coefficient. It was shown that the RMS error decreases for larger circularity coefficients⁶ and higher SNR values (lower noise level). As expected, after applying Q-SUT, the $\mathbf{C}_{\hat{\mathbf{x}}^i}$ was diagonal and $\mathbf{C}_{\hat{\mathbf{x}}^j} = \mathbf{C}_{\hat{\mathbf{x}}^\kappa} = 0$ for all simulations.

Furthermore, Fig. 1b and 1c represent the 3-D scatter plots of the generated signals \mathbf{x} versus the estimated signals $\hat{\mathbf{x}}$ obtained via Q-SUT. Note that the elliptical scatter plots of the original signals confirmed their high correlation, while circular nature of the estimated signals indicated successful decorrelation of Q-SUT.

⁶In this work, larger circularity was associated with the stronger correlation, regardless of the power.



(a) Reconstruction error against circularity coefficient and SNR. Note that the error was averaged over 30 sets of simulations.



(b) Scatter plots between three \mathbb{C}^1 -proper (c) Scatter plots between the estimated sources $\hat{\mathbf{x}}$ obtained via Q-SUT.

Fig. 1: Results of the proposed Q-SUT method for decorrelation and BSS of \mathbb{C}^1 -proper sources.

4.2. Alamouti-based Communication Systems

This section presents the application of the QUT for a practical communication problem of Alamouti coding [14]. Consider a 2×2 source separation problem in \mathbb{H} given by $\mathbf{x} = \mathbf{A}\mathbf{s}$

$$\begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \end{bmatrix} \quad (6)$$

in which $\mathbf{x} = \mathbf{x}_a + \mathbf{x}_b j$ and $\mathbf{s} = \mathbf{s}_a + \mathbf{s}_b j$ where \mathbf{x}_a and \mathbf{x}_b are two consecutive signals observed at the receiver; \mathbf{s}_a and \mathbf{s}_b are the transmitted complex sources to be recovered. While the work in [8], the above equation corresponds to:

$$\begin{aligned} \begin{bmatrix} \mathbf{x}_{1a} \\ \mathbf{x}_{1b} \end{bmatrix} &= \begin{bmatrix} \mathbf{s}_{1a} & -\mathbf{s}_{1b}^* \\ \mathbf{s}_{1b} & \mathbf{s}_{1a}^* \end{bmatrix} \begin{bmatrix} a_{11\alpha} \\ a_{11\beta} \end{bmatrix} + \begin{bmatrix} \mathbf{s}_{2a} & -\mathbf{s}_{2b}^* \\ \mathbf{s}_{2b} & \mathbf{s}_{2a}^* \end{bmatrix} \begin{bmatrix} a_{12\alpha} \\ a_{12\beta} \end{bmatrix} \in \mathbb{C} \\ \begin{bmatrix} \mathbf{x}_{2a} \\ \mathbf{x}_{2b} \end{bmatrix} &= \begin{bmatrix} \mathbf{s}_{2a} & -\mathbf{s}_{2b}^* \\ \mathbf{s}_{2b} & \mathbf{s}_{2a}^* \end{bmatrix} \begin{bmatrix} a_{22\alpha} \\ a_{22\beta} \end{bmatrix} + \begin{bmatrix} \mathbf{s}_{1a} & -\mathbf{s}_{1b}^* \\ \mathbf{s}_{1b} & \mathbf{s}_{1a}^* \end{bmatrix} \begin{bmatrix} a_{21\alpha} \\ a_{21\beta} \end{bmatrix} \in \mathbb{C} \end{aligned} \quad (7)$$

where $a_{ij\tau}$ represents the channel between the receiver of the i -th user and the τ -th transmit antenna of the j -th user, $\tau \in \{\alpha, \beta\}$.

In this study, the sources were selected as binary and quadrature phase shift keying (BPSK or QPSK) and 16-quadrature amplitude modulation (QAM). Initially, we considered a scenario with quaternion-valued mixing matrix \mathbf{A} , where the pair of symbols \mathbf{s}_{1a} and \mathbf{s}_{1b} of the first user were equal, shown in the left columns in Fig. 2. For the sake of comparison, we also considered two cases where the mixing matrix \mathbf{A} was complex-valued and \mathbf{s}_{1a} and \mathbf{s}_{1b} could be either equal or different, as in the middle and right columns in Fig. 2.

After generating the observations \mathbf{x} , QUT was applied to provide an estimation of the sources as $\hat{\mathbf{x}}$. To this end, QUT decorrelated the standard covariance and a unitary complementary covariance with the maximum correlation, i.e. the unit axis in which marginal variables had highest correlation. Furthermore, QUT has been compared with conventional complex SUT [4].

Note that when using a quaternion-valued mixing matrix and $\mathbf{s}_{1a} = \mathbf{s}_{1b}$, the conventional SUT failed to recover the sources, see Fig. 2 L(b). However, for these cases, QUT provided reasonable estimates of the sources in L(c).

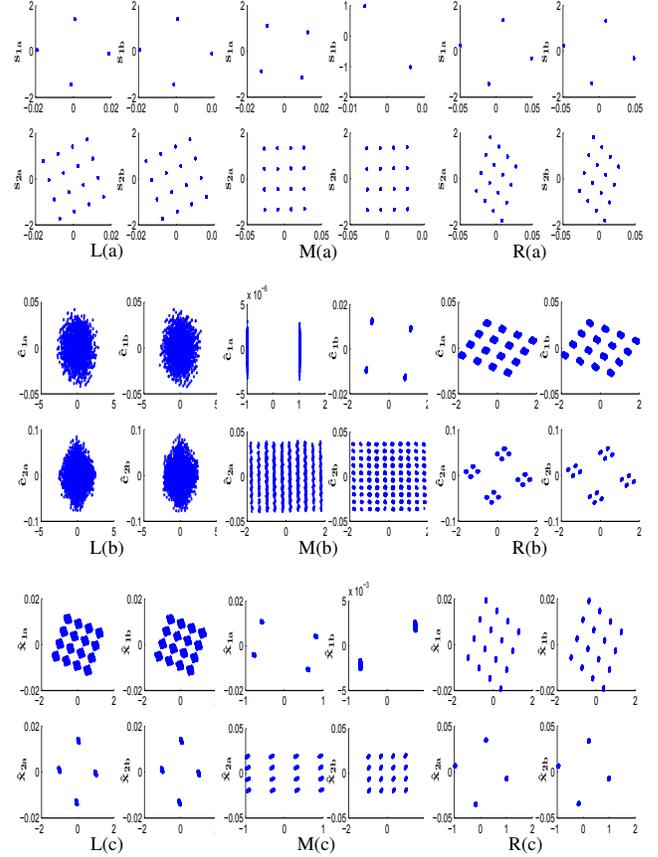


Fig. 2: Plots of communication constellations: (L) left column corresponds to the case $\mathbf{s}_{1a} = \mathbf{s}_{1b}$ with a quaternion mixing matrix, (M) middle column represents the $\mathbf{s}_{1a} \neq \mathbf{s}_{1b}$ scenario with a complex mixing matrix and (R) right column corresponds to $\mathbf{s}_{1a} = \mathbf{s}_{1b}$ with a complex mixing matrix. In each column, subplot (a) represents the original sources, (b) is the estimated sources using complex SUT and (c) is the estimated sources using QUT.

On the other hand, when using a complex-valued mixing matrix⁷ and $\mathbf{s}_{1a} \neq \mathbf{s}_{1b}$, the conventional SUT reconstructed a reasonable version of either BPSK or QPSK sources, however, it still could not recover the 16-QAM source, see Fig. 2 M(b). In contrast, QUT provided reconstruction of 16-QAM sources as well as the BPSK and QPSK sources in M(c). Furthermore, when using a complex-valued mixing matrix and $\mathbf{s}_{1a} = \mathbf{s}_{1b}$, QUT outperformed the Q-SUT by providing higher distinct output, compare R(b) and R(c) in Fig. 2. Recall that both QUT and Q-SUT suffer from the permutation problem and they estimate the sources in a random order - similar to most ICA algorithms.

5. CONCLUSION

We have proposed a novel set of matrix decompositions for the joint diagonalisation of η -Hermitian matrices. Of particular interest is the proposed quaternion uncorrelating transform (QUT), which showed the effectiveness of our method in an Alamouti-based source separation problem. Furthermore, we hope that our findings in joint diagonalisation of quaternion matrices which satisfy conditions such as normality of $\mathbf{A}\mathbf{B}^\eta$, or $\mathbf{A}\mathbf{B} = \mathbf{B}\mathbf{A}$, or $\mathbf{B}\mathbf{A} = \mathbf{A}^\eta\mathbf{B}$ lay the foundations for further research in quaternion signal processing.

⁷To this end, the channel response a_{ib} of the second transmitter antenna for each i -th user was set to zero, although this is unlikely in practice.

6. REFERENCES

- [1] L. De Lathauwer and B. De Moor, "On the blind separation of non-circular sources," *Proceedings of the IEEE European Signal Processing Conference*, pp. 1–4, 2002.
- [2] D. P. Mandic and V. S. L. Goh, "Complex valued nonlinear adaptive filters: noncircularity, widely linear and neural models," *John Wiley & Sons*, vol. 59, 2009.
- [3] E. Ollila and V. Koivunen, "Complex ICA using generalized uncorrelating transform," *Signal Processing*, vol. 89, no. 4, pp. 365–377, 2009.
- [4] J. Eriksson and V. Koivunen, "Complex random vectors and ICA models: Identifiability, uniqueness, and separability," *IEEE Transactions on Information Theory*, vol. 52, no. 3, pp. 1017–1029, 2006.
- [5] R. A. Horn and C. R. Johnson, *Matrix analysis*. Cambridge University Press, 2012.
- [6] N. Le Bihan and J. Mars, "Singular value decomposition of quaternion matrices: a new tool for vector-sensor signal processing," *Signal Processing*, vol. 84, no. 7, pp. 1177–1199, 2004.
- [7] C. C. Took, D. P. Mandic, and F. Zhang, "On the unitary diagonalisation of a special class of quaternion matrices," *Applied Mathematics Letters*, vol. 24, no. 11, pp. 1806–1809, 2011.
- [8] S. Javidi, C. C. Took, and D. P. Mandic, "Fast independent component analysis algorithm for quaternion valued signals," *IEEE Transactions on Neural Networks*, vol. 22, no. 12, pp. 1967–1978, 2011.
- [9] C. C. Took and D. P. Mandic, "Augmented second-order statistics of quaternion random signals," *Signal Processing*, vol. 91, no. 2, pp. 214 – 224, 2011.
- [10] T. A. Ell and S. J. Sangwine, "Quaternion involutions and anti-involutions," *Computers & Mathematics with Applications*, vol. 53, no. 1, pp. 137–143, 2007.
- [11] C. C. Took and D. P. Mandic, "Augmented second-order statistics of quaternion random signals," *Signal Processing*, vol. 91, no. 2, pp. 214–224, 2011.
- [12] R. A. Horn and F. Zhang, "A generalization of the complex Autonne–Takagi factorization to quaternion matrices," *Linear and Multilinear Algebra*, vol. 60, no. 11–12, pp. 1239–1244, 2012.
- [13] N. L. Bihan and S. Sangwine, "Jacobi method for quaternion matrix singular value decomposition," *Applied Mathematics and Computation*, vol. 187, pp. 1265–1271, 2007.
- [14] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1451–1458, 1998.