PERFORMANCE ADVANTAGE OF QUATERNION WIDELY LINEAR ESTIMATION: AN APPROXIMATE UNCORRELATING TRANSFORM APPROACH

Min Xiang[1], Sithan Kanna[1], Scott C. Douglas[2] and Danilo P. Mandic[1]

[1]Department of Electrical and Electronic Engineering, Imperial College London, SW7 2AZ, U.K.
[2]Department of Electrical Engineering, Southern Methodist University, Dallas, Texas 75275, USA
Emails: {m.xiang13, ssk08, d.mandic}@ic.ac.uk, douglas@lyle.smu.edu

ABSTRACT

Widely linear processing has been shown to be superior to the traditional strictly linear processing in quaternion minimum mean square error (MMSE) estimation. However, a quantifiable performance difference between strictly and widely linear processing and the relationship between the performance and quaternion impropriety are still lacking. To this end, we present a proof for the performance advantage of widely linear estimation and relate the performance bounds to signal properties by exploiting the approximate joint diagonalisation of quaternion covariance matrices. In that sense, this work can be seen as a generalisation of complex-valued MMSE estimation, and can thus also be applied to the complex-valued case. Simulations on synthetic signals support the analysis.

Index Terms— Quaternion estimation, widely linear modelling, minimum mean square error (MMSE), impropriety

1. INTRODUCTION

Quaternions have traditionally been used in aerospace and computer graphics in order to model 3-D rotations and orientations as their algebra avoids numerical problems associated with vector algebras, such as gimbal lock [1]. In recent years, with the introduction of the $\mathbb{HR}$ calculus [2, 3] and augmented quaternion statistics [4, 5], quaternion-valued signals have experienced a resurgence as they account for mutual information between data channels, provide compact representations, and offer physically meaningful interpretations for a number of 3-D and 4-D applications, such as wind prediction [6] and frequency estimation in smart grid [7]. The minimum mean square error (MMSE) estimation framework has also been extended to handle quaternion-valued signals [8, 9].

Conventional strictly linear quaternion MMSE estimators utilise standard covariance matrices and are optimal only for estimating second-order circular quaternion signals. Advances in quaternion statistics have established that in order to capture complete second-order statistics of quaternion signals, widely linear MMSE estimators that exploit three complementary covariances matrices in addition to the standard covariance matrix are needed [4, 5].

Due to the non-trivial nature of quaternion algebra, a closed-form performance comparison between strictly and widely linear MMSE estimators has not yet been presented. The only available result was achieved in an indirect fashion via the semi-widely linear estimator [10–12], which is a generalisation of the analysis for complex estimation [13]. In addition, the extent to which the performance bounds of these estimators are affected by impropriety of signals remains unclear.

This paper provides an analytical understanding and second-order performance comparison between strictly and widely linear estimators. This is achieved by examining the structure of the covariance and complementary covariance matrices of the input signal, and by employing the approximate uncorrelating transform, a tool to jointly decompose the covariance and complementary covariance matrices [14, 15]. In this way, the relationship between the MMSE and second-order statistics of the input signal is established, and the effect of signal properties on the performance is quantified. The analysis is supported by illustrative simulations.

Throughout the paper, we use lowercase letters to denote scalars, boldface lowercase letters for vectors, and boldface uppercase letters for matrices. Superscripts $(\cdot)^T$, $(\cdot)^*$, and $(\cdot)^H$ denote the transpose, conjugate, and Hermitian (i.e., transpose and conjugate), respectively. $E\{\cdot\}$ denotes the statistical expectation operator.

2. BACKGROUND

2.1. Quaternion algebra

The quaternion domain $\mathbb{H}$ is a 4-D vector space over the real field, spanned by the basis $\{1, \imath, \jmath, \kappa\}$. A quaternion vector $\mathbf{x}$ comprises of a scalar part $\Re\{\cdot\}$ and a vector part $\Im\{\cdot\}$ which consists of three imaginary components, so that

$$\mathbf{x} = \Re\{\mathbf{x}\} + \Im\{\mathbf{x}\} = \mathbf{x}_a + \imath\mathbf{x}_b + \jmath\mathbf{x}_c + \kappa\mathbf{x}_d \tag{1}$$

where $\imath$, $\jmath$, κ are orthogonal imaginary units with the properties

$$\begin{aligned} &\imath\jmath = -\jmath\imath = \kappa \qquad \jmath\kappa = -\kappa\jmath = \imath \qquad \kappa\imath = -\imath\kappa = \jmath \\ &\imath^2 = \jmath^2 = \kappa^2 = \imath\jmath\kappa = -1 \end{aligned} \tag{2}$$

The l_2-norm of a quaternion random vector $\mathbf{x} \in \mathbb{H}$ is defined as

$$||\mathbf{x}||_2 = (||\mathbf{x}_a||_2^2 + ||\mathbf{x}_b||_2^2 + ||\mathbf{x}_c||_2^2 + ||\mathbf{x}_d||_2^2)^{\frac{1}{2}}$$

while the product of two quaternions x and y is given by

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

where the symbol '$\cdot$' denotes the scalar product, and '$\times$' denotes the vector product. The presence of the vector product causes non-commutativity of the quaternion product, that is, $xy \neq yx$.

Another important notion is that of the quaternion involution [16], which defines a self-inverse quaternion mapping. The general involution of the quaternion vector $\mathbf{x}$ is defined as $\mathbf{x}^\alpha = -\alpha\mathbf{x}\alpha$, where α is a unit pure quaternion, while the special cases of involutions about the $\imath$, $\jmath$ and κ imaginary axes are given by

$$\begin{aligned} \mathbf{x}^\imath &= -\imath\mathbf{x}\imath = \mathbf{x}_a + \imath\mathbf{x}_b - \jmath\mathbf{x}_c - \kappa\mathbf{x}_d \\ \mathbf{x}^\jmath &= -\jmath\mathbf{x}\jmath = \mathbf{x}_a - \imath\mathbf{x}_b + \jmath\mathbf{x}_c - \kappa\mathbf{x}_d \\ \mathbf{x}^\kappa &= -\kappa\mathbf{x}\kappa = \mathbf{x}_a - \imath\mathbf{x}_b - \jmath\mathbf{x}_c + \kappa\mathbf{x}_d \end{aligned} \tag{3}$$

Note that an involution represents a rotation along a single unit axis, while the quaternion conjugate operator $(\cdot)^*$ rotates the quaternion variable along all three imaginary axes, and is given by

$$\mathbf{x}^* = \Re\{\mathbf{x}\} - \Im\{\mathbf{x}\} = \mathbf{x}_a - \imath\mathbf{x}_b - \jmath\mathbf{x}_c - \kappa\mathbf{x}_d \tag{4}$$

2.2. Second-order statistics

The notion of impropriety or second-order noncircularity, which refers to second-order moments being rotation dependent, is unique to division algebras. In the quaternion domain, impropriety is characterised by the degree of correlation and/or power imbalance between the imaginary components relative to the real component. The impropriety coefficients of the zero-mean quaternion random variable x can be defined as [17, 18]

$$r_\alpha = \left| \frac{E\{xx^{\alpha*}\}}{E\{xx^*\}} \right| \in [0,1] \qquad \alpha \in \{\imath, \jmath, \kappa\} \tag{5}$$

which reflect the correlation and between x and each of its involutions.

Definition 1. A random quaternion variable x is $\mathbb{H}$-proper iff its impropriety coefficients all vanish, that is, $r_\imath = r_\jmath = r_\kappa = 0$. A random quaternion variable x is maximally improper iff its impropriety coefficients are maximal, that is, $r_\imath = r_\jmath = r_\kappa = 1$.

For a random quaternion vector $\mathbf{x}$ with zero mean, the Hermitian standard covariance $\mathbf{C}_{\mathbf{xx}} = E\{\mathbf{x}\mathbf{x}^H\}$ is insufficient to exploit complete second-order statistical information about $\mathbf{x}$, and to that end it is a prerequisite to also exploit information conveyed by quaternion involutions represented by the $\imath$-, $\jmath$-, and κ-covariances, which are referred to as the complementary covariances and are given by [5]

$$\mathbf{C}_{\mathbf{x}\alpha} = E\{\mathbf{x}\mathbf{x}^{\alpha H}\} \qquad \alpha \in \{\imath, \jmath, \kappa\} \tag{6}$$

The $\alpha-$covariance matrices are $\alpha-$Hermitian, that is, $\mathbf{C}_{\mathbf{x}\alpha} = (\mathbf{C}_{\mathbf{x}\alpha})^{\alpha H}$.

Definition 2. A random quaternion vector $\mathbf{x}$ is $\mathbb{H}$-proper iff it is uncorrelated with its involutions $\mathbf{x}^\imath$, $\mathbf{x}^\jmath$ and $\mathbf{x}^\kappa$, so that all complementary covariances vanish, that is, $\mathbf{C}_{\mathbf{x}\imath} = \mathbf{C}_{\mathbf{x}\jmath} = \mathbf{C}_{\mathbf{x}\kappa} = \mathbf{0}$.

2.3. Augmented quaternion statistics

The set of involutions in (3), together with the original quaternion vector $\mathbf{x}$, forms the most frequently used basis for augmented quaternion statistics [4, 5]. The complete second-order information of random quaternion vectors can only be obtained by examining the augmented random vector $\mathbf{x}^a = \left[\mathbf{x}^T, \mathbf{x}^{\imath T}, \mathbf{x}^{\jmath T}, \mathbf{x}^{\kappa T}\right]^T$ and its corresponding augmented covariance matrix $\mathbf{C}_{\mathbf{x}^a\mathbf{x}^a} = E\{\mathbf{x}^a\mathbf{x}^{aH}\}$. For notational simplicity, we denote $\mathbf{x}^b = \left[\mathbf{x}^{\imath T}, \mathbf{x}^{\jmath T}, \mathbf{x}^{\kappa T}\right]^T$ which only includes the involutions of $\mathbf{x}$. The augmented covariance matrix can then be examined through the following four blocks

$$\begin{aligned}\mathbf{C}_{\mathbf{x}^a\mathbf{x}^a} &= \left[\begin{array}{c|ccc} \mathbf{C}_{\mathbf{xx}} & \mathbf{C}_{\mathbf{x}\imath} & \mathbf{C}_{\mathbf{x}\jmath} & \mathbf{C}_{\mathbf{x}\kappa} \\ \hline \mathbf{C}^\imath_{\mathbf{x}\imath} & \mathbf{C}^\imath_{\mathbf{xx}} & \mathbf{C}^\imath_{\mathbf{x}\kappa} & \mathbf{C}^\imath_{\mathbf{x}\jmath} \\ \mathbf{C}^\jmath_{\mathbf{x}\jmath} & \mathbf{C}^\jmath_{\mathbf{x}\kappa} & \mathbf{C}^\jmath_{\mathbf{xx}} & \mathbf{C}^\jmath_{\mathbf{x}\imath} \\ \mathbf{C}^\kappa_{\mathbf{x}\kappa} & \mathbf{C}^\kappa_{\mathbf{x}\jmath} & \mathbf{C}^\kappa_{\mathbf{x}\imath} & \mathbf{C}^\kappa_{\mathbf{xx}} \end{array}\right] \\ &= \left[\begin{array}{c|c} \mathbf{C}_{\mathbf{xx}} & \mathbf{C}_{\mathbf{xx}^b} \\ \hline \mathbf{C}_{\mathbf{x}^b\mathbf{x}} & \mathbf{C}_{\mathbf{x}^b\mathbf{x}^b} \end{array}\right]\end{aligned} \tag{7}$$

where

$$\mathbf{C}_{\mathbf{xx}^b} = E\left\{\mathbf{x}\mathbf{x}^{bH}\right\} = \left[\begin{array}{ccc} \mathbf{C}_{\mathbf{x}\imath} & \mathbf{C}_{\mathbf{x}\jmath} & \mathbf{C}_{\mathbf{x}\kappa} \end{array}\right] \tag{8}$$

$$\mathbf{C}_{\mathbf{x}^b\mathbf{x}} = E\left\{\mathbf{x}^b\mathbf{x}^H\right\} = \left[\begin{array}{ccc} \mathbf{C}^\imath_{\mathbf{x}\imath} & \mathbf{C}^\jmath_{\mathbf{x}\jmath} & \mathbf{C}^\kappa_{\mathbf{x}\kappa} \end{array}\right]^T \tag{9}$$

$$\mathbf{C}_{\mathbf{x}^b\mathbf{x}^b} = E\left\{\mathbf{x}^b\mathbf{x}^{bH}\right\} = \left[\begin{array}{ccc} \mathbf{C}^\imath_{\mathbf{xx}} & \mathbf{C}^\imath_{\mathbf{x}\kappa} & \mathbf{C}^\imath_{\mathbf{x}\jmath} \\ \mathbf{C}^{\imath H}_{\mathbf{x}\kappa} & \mathbf{C}^\jmath_{\mathbf{xx}} & \mathbf{C}^\jmath_{\mathbf{x}\imath} \\ \mathbf{C}^{\imath H}_{\mathbf{x}\jmath} & \mathbf{C}^{\jmath H}_{\mathbf{x}\imath} & \mathbf{C}^\kappa_{\mathbf{xx}} \end{array}\right] \tag{10}$$

These covariance matrices are Hermitian symmetrical, that is,

$$\mathbf{C}_{\mathbf{x}^a\mathbf{x}^a} = \mathbf{C}^H_{\mathbf{x}^a\mathbf{x}^a} \qquad \mathbf{C}_{\mathbf{x}^b\mathbf{x}^b} = \mathbf{C}^H_{\mathbf{x}^b\mathbf{x}^b} \qquad \mathbf{C}_{\mathbf{x}^b\mathbf{x}} = \mathbf{C}^H_{\mathbf{xx}^b}$$

2.4. Quaternion MMSE estimation

The goal of MMSE estimation is to find the optimal estimate $\hat{y} = E[y\,|\mathbf{x}]$ of a desired signal $y \in \mathbb{H}$ in terms of the input signal vector $\mathbf{x} \in \mathbb{H}^{L\times 1}$ by minimising MSE $= E\left\{|y - \hat{y}|^2\right\}$. The estimation mechanism, $\hat{y} = f(\mathbf{x})$, which represents the prior knowledge about the relation between y and $\mathbf{x}$, is crucial for estimation performance. Traditionally, the strictly linear (SL) model has been used, which has the form [8]

$$\hat{y} = \hat{\mathbf{h}}^H\mathbf{x} \tag{11}$$

where $\mathbf{h} \in \mathbb{H}^{L\times 1}$ is the weight vector. This model achieves optimal estimation for circular quaternion signals but is suboptimal for general quaternion signals. To address this issue, involutions of $\mathbf{x}$ can be incorporated to capture complete second-order statistical information, leading to the widely linear (WL) model given by [9]

$$\begin{aligned}\hat{y} &= \hat{\mathbf{h}}^H\mathbf{x} + \hat{\mathbf{g}}^H\mathbf{x}^\imath + \hat{\mathbf{u}}^H\mathbf{x}^\jmath + \hat{\mathbf{v}}^H\mathbf{x}^\kappa \\ &= \hat{\mathbf{h}}^H\mathbf{x} + \hat{\mathbf{w}}^{bH}\mathbf{x}^b \\ &= \hat{\mathbf{w}}^{aH}\mathbf{x}^a\end{aligned} \tag{12}$$

where $\hat{\mathbf{h}}, \hat{\mathbf{g}}, \hat{\mathbf{u}}, \hat{\mathbf{v}} \in \mathbb{H}^{L\times 1}$ are the estimated weight vectors which can be compactly represented in the augmented form as $\hat{\mathbf{w}}^a = \left[\hat{\mathbf{h}}^T, \hat{\mathbf{g}}^T, \hat{\mathbf{u}}^T, \hat{\mathbf{v}}^T\right]^T$. For notational simplicity, we shall also use $\hat{\mathbf{w}}^b = \left[\hat{\mathbf{g}}^T, \hat{\mathbf{u}}^T, \hat{\mathbf{v}}^T\right]^T$ to denote the complementary weight vector that is present in the WL model but not in the SL model.

3. PERFORMANCE ADVANTAGE OF WIDELY LINEAR ESTIMATION

3.1. Desired signal model

Consider the desired signal, y, which is generated by a WL system driven by noise, υ, to give

$$\begin{aligned}y &= \mathbf{h}^H\mathbf{x} + \mathbf{g}^H\mathbf{x}^\imath + \mathbf{u}^H\mathbf{x}^\jmath + \mathbf{v}^H\mathbf{x}^\kappa + \upsilon \\ &= \mathbf{h}^H\mathbf{x} + \mathbf{w}^{bH}\mathbf{x}^b + \upsilon \\ &= \mathbf{w}^{aH}\mathbf{x}^a + \upsilon\end{aligned} \tag{13}$$

where $\mathbf{w}^a = \left[\mathbf{h}^T, \mathbf{g}^T, \mathbf{u}^T, \mathbf{v}^T\right]^T$ denotes the true augmented weight vector of the WL system, while $\mathbf{w}^b = \left[\mathbf{g}^T, \mathbf{u}^T, \mathbf{v}^T\right]^T$ designates the true complementary weight vector of the WL system, and υ quaternion white Gaussian noise with variance σ_υ^2 that is statistically independent of $\mathbf{x}$.

3.2. MMSE in SL estimation

It has been proved that the optimal weight of the SL estimator in (11) is provided by the Wiener solution [11], given by

$$\begin{aligned}\hat{\mathbf{h}}_{\mathrm{o}} &= E\left\{\mathbf{x}\mathbf{x}^H\right\}^{-1} E\left\{\mathbf{x}y^*\right\} \\ &= E\left\{\mathbf{x}\mathbf{x}^H\right\}^{-1} E\left\{\mathbf{x}\left(\mathbf{h}^H\mathbf{x} + \mathbf{w}^{bH}\mathbf{x}^b + \upsilon\right)^*\right\} \\ &= \mathbf{h} + \mathbf{C}_{\mathbf{xx}}^{-1}\mathbf{C}_{\mathbf{xx}^b}\mathbf{w}^b \end{aligned} \quad (14)$$

Note that $\hat{\mathbf{h}}_{\mathrm{o}} \neq \mathbf{h}$ if $\mathbf{C}_{\mathbf{xx}}^{-1}\mathbf{C}_{\mathbf{xx}^b}\mathbf{w}^b \neq \mathbf{0}$, which occurs when the system is WL ($\mathbf{w}^b \neq \mathbf{0}$) and $\mathbf{x}$ is not $\mathbb{H}$-proper ($\mathbf{C}_{\mathbf{xx}^b} \neq \mathbf{0}$). The SL estimator yields the steady-state error

$$e_{\mathrm{sl}} = y - \hat{\mathbf{h}}_{\mathrm{o}}^H\mathbf{x} = \mathbf{w}^{bH}\mathbf{x}^b - \mathbf{w}^{bH}\mathbf{C}_{\mathbf{xx}^b}^H\mathbf{C}_{\mathbf{xx}}^{-1}\mathbf{x} + \upsilon$$

and the corresponding error power (MMSE) is given by

$$E\left\{|e_{\mathrm{sl}}|^2\right\} = \sigma_\upsilon^2 + \mathbf{w}^{bH}\mathbf{M}\mathbf{w}^b \quad (15)$$

where $\mathbf{M} = \mathbf{C}_{\mathbf{x}^b\mathbf{x}^b} - \mathbf{C}_{\mathbf{xx}^b}^H\mathbf{C}_{\mathbf{xx}}^{-1}\mathbf{C}_{\mathbf{xx}^b}$ is the Schur complement of the augmented covariance matrix $\mathbf{C}_{\mathbf{x}^a\mathbf{x}^a}$ in (7).

3.3. MMSE in WL estimation

The optimal weight vector of the WL estimator in (12) is given by the Wiener solution $\hat{\mathbf{w}}_{\mathrm{o}}^a = E\{\mathbf{x}^a\mathbf{x}^{aH}\}^{-1}E\{\mathbf{x}^a y^*\} = \mathbf{w}^a$, resulting in the steady-state error $e_{\mathrm{wl}} = \upsilon$, and the corresponding MMSE

$$E\left\{|e_{\mathrm{wl}}|^2\right\} = \sigma_\upsilon^2 \quad (16)$$

3.4. MMSE comparison between SL and WL models

The difference between the MMSEs of the above two estimators is obtained from (15) and (16) to yield

$$\delta e^2 = E\left\{|e_{\mathrm{sl}}|^2\right\} - E\left\{|e_{\mathrm{wl}}|^2\right\} = \mathbf{w}^{bH}\mathbf{M}\mathbf{w}^b \quad (17)$$

Since $\mathbf{C}_{\mathbf{x}^a\mathbf{x}^a}$ is positive semi-definite, its Schur complement, $\mathbf{M}$, is also positive semi-definite [19], and thus $\delta e^2 \geq 0$, indicating that the WL estimator has a smaller or equal MMSE compared to its SL counterpart. However, their MMSEs are equal, that is, $\delta e^2 = 0$, if at least one of the following three conditions is satisfied:

1. The system being modelled is strictly linear, that is, $\mathbf{w}^b = \mathbf{0}$.
2. The complementary weight vector $\mathbf{w}^b$ falls within the nullspace of $\mathbf{M}$, that is, $\mathbf{M}\mathbf{w}^b = \mathbf{0}$.
3. The input signal $\mathbf{x}$ is maximally improper, that is, $\mathbf{x} = \alpha\mathbf{x}^\imath = \beta\mathbf{x}^\jmath = \gamma\mathbf{x}^\kappa$ with probability 1 for constant α,β and γ. In this case, $\mathbf{M} = \mathbf{0}$.

Remark 1. WL processing has a lower MMSE than SL processing when estimating singals generated by WL systems, unless the second or third condition holds. Otherwise, their MMSEs are equal.

4. PERFORMANCE ANALYSIS WITH THE APPROXIMATE UNCORRELATING TRANSFORM

Recent advances in quaternion linear algebra include the approximate uncorrelating transform (AUT) which allows for a joint diagonalisation of the standard covariance matrix, $\mathbf{C}_{\mathbf{xx}}$, and the complementary covariance matrices, $\mathbf{C}_{\mathbf{x}\imath}$,$\mathbf{C}_{\mathbf{x}\jmath}$,$\mathbf{C}_{\mathbf{x}\kappa}$, from (7), using the same unitary matrix [14, 15]. The AUT will next be used to understand the effect of the matrix $\mathbf{M}$ on δe^2 in (17).

Formally, the AUT in the quaternion domain states that there exists a unitary matrix $\mathbf{Q}$ and diagonal matrices $\mathbf{\Lambda}_{\mathbf{x}}$,$\mathbf{\Lambda}_\imath$,$\mathbf{\Lambda}_\jmath$,$\mathbf{\Lambda}_\kappa$ with singular values of $\mathbf{C}_{\mathbf{xx}}$,$\mathbf{C}_{\mathbf{x}\imath}$,$\mathbf{C}_{\mathbf{x}\jmath}$,$\mathbf{C}_{\mathbf{x}\kappa}$ on the diagonal such that

$$\begin{aligned} \mathbf{C}_{\mathbf{xx}} &\approx \mathbf{Q}\mathbf{\Lambda}_{\mathbf{x}}\mathbf{Q}^H & \qquad \mathbf{C}_{\mathbf{x}\imath} &\approx \mathbf{Q}\mathbf{\Lambda}_\imath\mathbf{Q}^{\imath H} \\ \mathbf{C}_{\mathbf{x}\jmath} &\approx \mathbf{Q}\mathbf{\Lambda}_\jmath\mathbf{Q}^{\jmath H} & \qquad \mathbf{C}_{\mathbf{x}\kappa} &\approx \mathbf{Q}\mathbf{\Lambda}_\kappa\mathbf{Q}^{\kappa H} \end{aligned} \quad (18)$$

From (18), the block partitioned matrices from (8) and (10) can be expressed as

$$\mathbf{C}_{\mathbf{xx}^b} \approx \mathbf{Q}\begin{bmatrix} \mathbf{\Lambda}_\imath & \mathbf{\Lambda}_\jmath & \mathbf{\Lambda}_\kappa \end{bmatrix}\mathbf{U}^H \quad (19)$$

$$\mathbf{C}_{\mathbf{x}^b\mathbf{x}^b} \approx \mathbf{U}\begin{bmatrix} \mathbf{\Lambda}_{\mathbf{x}} & \mathbf{\Lambda}_\kappa^\imath & \mathbf{\Lambda}_\jmath^\imath \\ \mathbf{\Lambda}_\kappa^\jmath & \mathbf{\Lambda}_{\mathbf{x}} & \mathbf{\Lambda}_\imath^\jmath \\ \mathbf{\Lambda}_\jmath^\kappa & \mathbf{\Lambda}_\imath^\kappa & \mathbf{\Lambda}_{\mathbf{x}} \end{bmatrix}\mathbf{U}^H \quad (20)$$

where $\mathbf{U} = \mathsf{bdiag}\{\mathbf{Q}^\imath, \mathbf{Q}^\jmath, \mathbf{Q}^\kappa\}$ is a unitary block diagonal matrix with the involutions of the matrix $\mathbf{Q}$ as the block diagonal entries. Therefore, the Schur complement matrix $\mathbf{M}$ in (17) is factorised as $\mathbf{M} = \mathbf{U}\mathbf{S}\mathbf{U}^H$ where

$$\mathbf{S} = \begin{bmatrix} \mathbf{\Lambda}_{\mathbf{x}} - \mathbf{\Lambda}_{\mathbf{x}}^{-1}\mathbf{\Lambda}_\imath^*\mathbf{\Lambda}_\imath & \mathbf{\Lambda}_\kappa^\imath - \mathbf{\Lambda}_{\mathbf{x}}^{-1}\mathbf{\Lambda}_\imath^*\mathbf{\Lambda}_\jmath & \mathbf{\Lambda}_\jmath^\imath - \mathbf{\Lambda}_{\mathbf{x}}^{-1}\mathbf{\Lambda}_\imath^*\mathbf{\Lambda}_\kappa \\ \mathbf{\Lambda}_\kappa^\jmath - \mathbf{\Lambda}_{\mathbf{x}}^{-1}\mathbf{\Lambda}_\jmath^*\mathbf{\Lambda}_\imath & \mathbf{\Lambda}_{\mathbf{x}} - \mathbf{\Lambda}_{\mathbf{x}}^{-1}\mathbf{\Lambda}_\jmath^*\mathbf{\Lambda}_\jmath & \mathbf{\Lambda}_\imath^\jmath - \mathbf{\Lambda}_{\mathbf{x}}^{-1}\mathbf{\Lambda}_\jmath^*\mathbf{\Lambda}_\kappa \\ \mathbf{\Lambda}_\jmath^\kappa - \mathbf{\Lambda}_{\mathbf{x}}^{-1}\mathbf{\Lambda}_\kappa^*\mathbf{\Lambda}_\imath & \mathbf{\Lambda}_\imath^\kappa - \mathbf{\Lambda}_{\mathbf{x}}^{-1}\mathbf{\Lambda}_\kappa^*\mathbf{\Lambda}_\jmath & \mathbf{\Lambda}_{\mathbf{x}} - \mathbf{\Lambda}_{\mathbf{x}}^{-1}\mathbf{\Lambda}_\kappa^*\mathbf{\Lambda}_\kappa \end{bmatrix}$$

and (17) turns into $\delta e^2 = \mathbf{w}^{bH}\mathbf{U}\mathbf{S}\mathbf{U}^H\mathbf{w}^b$.

4.1. Impact of impropriety of the input signal

In order to clarify the relationship between the performance advantage of WL estimation and impropriety of $\mathbf{x}$, it is useful to discuss the following two special cases prior to the general case.

4.1.1. $\mathbb{H}$-proper signal

The performance difference (17) obeys

$$\delta e^2 = \mathbf{w}^{bH}\left(\mathbf{C}_{\mathbf{x}^b\mathbf{x}^b} - \mathbf{C}_{\mathbf{xx}^b}^H\mathbf{C}_{\mathbf{xx}}^{-1}\mathbf{C}_{\mathbf{xx}^b}\right)\mathbf{w}^b \leq \mathbf{w}^{bH}\mathbf{C}_{\mathbf{x}^b\mathbf{x}^b}\mathbf{w}^b$$

The second equality holds if $\mathbf{C}_{\mathbf{xx}^b} = \mathbf{0}$, that is, $\mathbf{x}$ is $\mathbb{H}$-proper, in which case $\mathbf{C}_{\mathbf{x}^b\mathbf{x}^b}$ is diagonal, and the maximum of δe^2 becomes

$$\begin{aligned} &\mathbf{w}^{bH}\mathbf{C}_{\mathbf{x}^b\mathbf{x}^b}\mathbf{w}^b \\ &= \mathbf{g}^H\mathbf{Q}^\imath\mathbf{\Lambda}_{\mathbf{x}}\mathbf{Q}^{\imath H}\mathbf{g} + \mathbf{u}^H\mathbf{Q}^\jmath\mathbf{\Lambda}_{\mathbf{x}}\mathbf{Q}^{\jmath H}\mathbf{u} + \mathbf{v}^H\mathbf{Q}^\kappa\mathbf{\Lambda}_{\mathbf{x}}\mathbf{Q}^{\kappa H}\mathbf{v} \end{aligned}$$

If $\mathbf{\Lambda}_{\mathbf{x}} = \lambda\mathbf{I}$, where $\mathbf{I}$ is the $L \times L$ identity matrix, then we have

$$\delta e^2 = \lambda\left(\|\mathbf{g}\|_2^2 + \|\mathbf{u}\|_2^2 + \|\mathbf{v}\|_2^2\right) = \lambda\|\mathbf{w}^b\|_2^2 \quad (21)$$

Remark 2. Without loss in generality, the performance difference between WL and SL estimation increases as the l_2-norm of the complementary weight vector, $\|\mathbf{w}^b\|_2$, increases. This statement is true regardless of the impropriety of the input signal since the complementary weight of the WL system cannot be modelled by the SL estimator.

4.1.2. Uncorrelated improper signal

For an improper random quaternion vector $\mathbf{x}$ for which the elements are uncorrelated and have equal power, the covariance and complementary covariance matrices are $\mathbf{C}_{\mathbf{xx}} = c_x\mathbf{I}$, $\mathbf{C}_{\mathbf{x}\imath} = c_\imath\mathbf{I}$, $\mathbf{C}_{\mathbf{x}\jmath} = c_\jmath\mathbf{I}$, $\mathbf{C}_{\mathbf{x}\kappa} = c_\kappa\mathbf{I}$, where $c_x, c_\imath, c_\jmath, c_\kappa$ are constants. The performance difference (17) therefore becomes

$$\delta e^2 = \mathbf{w}^{bH}\left(\begin{bmatrix} c_x - c_x^{-1}|c_\imath|^2 & c_\kappa^\imath - c_x^{-1}c_\imath^*c_\jmath & c_\jmath^\imath - c_x^{-1}c_\imath^*c_\kappa \\ c_\kappa^\jmath - c_x^{-1}c_\jmath^*c_\imath & c_x - c_x^{-1}|c_\jmath|^2 & c_\imath^\jmath - c_x^{-1}c_\jmath^*c_\kappa \\ c_\jmath^\kappa - c_x^{-1}c_\kappa^*c_\imath & c_\imath^\kappa - c_x^{-1}c_\kappa^*c_\jmath & c_x - c_x^{-1}|c_\kappa|^2 \end{bmatrix} \otimes \mathbf{I}\right)\mathbf{w}^b \quad (22)$$

where the symbol $\otimes$ denotes the Kronecker product.

4.1.3. *General improper signal*

To examine the effect of the matrix $\mathbf{M}$ has on the performance difference δe^2, we shall next inspect its trace. Since $\mathbf{U}$ is unitary, the trace of $\mathbf{M}$ is

$$\begin{aligned}\mathrm{Tr}\left[\mathbf{M}\right] &= \mathrm{Tr}\left[\boldsymbol{\Lambda}_{\mathbf{x}} - \boldsymbol{\Lambda}_{\mathbf{x}}^{-1}\boldsymbol{\Lambda}_{\imath}^{*}\boldsymbol{\Lambda}_{\imath} + \boldsymbol{\Lambda}_{\mathbf{x}} - \boldsymbol{\Lambda}_{\mathbf{x}}^{-1}\boldsymbol{\Lambda}_{\jmath}^{*}\boldsymbol{\Lambda}_{\jmath} + \boldsymbol{\Lambda}_{\mathbf{x}} - \boldsymbol{\Lambda}_{\mathbf{x}}^{-1}\boldsymbol{\Lambda}_{\kappa}^{*}\boldsymbol{\Lambda}_{\kappa}\right] \\ &= \sum_{n=1}^{L} \lambda_n \left(3 - \left|\frac{\lambda_{\imath,n}}{\lambda_n}\right|^2 - \left|\frac{\lambda_{\jmath,n}}{\lambda_n}\right|^2 - \left|\frac{\lambda_{\kappa,n}}{\lambda_n}\right|^2\right) \qquad (23)\end{aligned}$$

where $\lambda_n, \lambda_{\imath,n}, \lambda_{\jmath,n}, \lambda_{\kappa,n}$ are singular values of $\mathbf{C}_{\mathbf{xx}}$, $\mathbf{C}_{\mathbf{x}\imath}$, $\mathbf{C}_{\mathbf{x}\jmath}$, $\mathbf{C}_{\mathbf{x}\kappa}$. As the degree of signal impropriety, which is reflected in $\left|\frac{\lambda_{\imath,n}}{\lambda_n}\right|$, $\left|\frac{\lambda_{\jmath,n}}{\lambda_n}\right|$ and $\left|\frac{\lambda_{\kappa,n}}{\lambda_n}\right|$, increases, Tr $[\mathbf{M}]$ decreases, so do eigenvalues of $\mathbf{M}$, and hence the performance difference δe^2 decreases.

Remark 3. An $\mathbb{H}$-proper quaternion vector $\mathbf{x}$ and its involutions are uncorrelated, and the Wiener solution of SL estimation is equal to $\mathbf{h}$, the system weight of $\mathbf{x}$ in (13). In contrast, an $\mathbb{H}$-improper quaternion vector $\mathbf{x}$ and its involutions are correlated, therefore some information about the involutions that is embedded in $\mathbf{x}$ can be exploited using the SL model, in which case the low performance of SL estimation is compensated for by the additional term $\mathbf{C}_{\mathbf{xx}}^{-1}\mathbf{C}_{\mathbf{xx}^b}\mathbf{w}^b$ in (14). This could explain why WL estimation shows only a slight improvement in MSE performance over SL estimation for highly improper signals, such as the 4-D wind data in [6].

Furthermore, the relative performance advantage of WL estimation can be defined as

$$\epsilon = \frac{E\left\{|e_{\mathrm{sl}}|^2\right\} - E\left\{|e_{\mathrm{wl}}|^2\right\}}{E\left\{|e_{\mathrm{sl}}|^2\right\}} = \frac{\delta e^2}{\delta e^2 + \sigma_{\upsilon}^2}$$

Remark 4. As δe^2 is related to power of the input signal, $\mathbf{x}$, according to (23), the relative performance advantage ϵ increases with the signal-to-noise ratio of $\mathbf{x}$ to υ.

4.2. MMSE analysis for complex MMSE estimation

The above analysis reduces to the complex case when the quaternion signal reduces to the complex one, that is, $\mathbf{x} = \mathbf{x}_1 + \imath\mathbf{x}_2$, where $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^{L\times 1}$. It follows that $\mathbf{C}_{\mathbf{x}\imath} = \mathbf{C}_{\mathbf{xx}}$, $\mathbf{C}_{\mathbf{x}\jmath} = \mathbf{C}_{\mathbf{x}\kappa} = E\{\mathbf{xx}^T\}$. Define the pseudo-covariance of $\mathbf{x}$ as $\mathbf{P}_{\mathbf{xx}} = E\{\mathbf{xx}^T\}$. The difference between the MMSEs of SL and WL complex estimators [20] can now be obtained through a degeneration of (17), and is given by

$$\delta e^2 = E\left\{|e_{\mathrm{sl}}|^2\right\} - E\left\{|e_{\mathrm{wl}}|^2\right\} = \mathbf{g}^H\left(\mathbf{C}_{\mathbf{xx}}^{*} - \mathbf{P}_{\mathbf{xx}}^{*}\mathbf{C}_{\mathbf{xx}}^{-1}\mathbf{P}_{\mathbf{xx}}\right)\mathbf{g} \quad (24)$$

which conforms with the previous work [21], while (23) reduces to

$$\mathrm{Tr}\left[\mathbf{M}\right] = \mathrm{Tr}\left[\boldsymbol{\Lambda}_{\mathbf{x}} - \boldsymbol{\Lambda}_{\mathbf{x}}^{-1}\boldsymbol{\Lambda}_{p}^{*}\boldsymbol{\Lambda}_{p}\right] = \sum_{n=1}^{L}\lambda_n\left(1 - \left|\frac{\lambda_{p,n}}{\lambda_n}\right|^2\right) \quad (25)$$

where $\boldsymbol{\Lambda}_p$ is the diagonal matrix for which the diagonal element, $\lambda_{p,n}$, is the singular value of $\mathbf{P}_{\mathbf{xx}}$, and $\left|\frac{\lambda_{p,n}}{\lambda_n}\right|$ reflects the impropriety of the complex input signal $\mathbf{x}$. The performance advantage δe^2 of complex WL estimation increases with the widely linear nature of the system, propriety of $\mathbf{x}$, or power of $\mathbf{x}$; the relative advantage increases with the first two factors and the signal-to-noise ratio of $\mathbf{x}$ to υ. A higher degree of impropriety of $\mathbf{x}$ indicates higher correlation between $\mathbf{x}$ and its conjugate $\mathbf{x}^*$, meaning that some information about $\mathbf{x}^*$ that is embedded in $\mathbf{x}$ can be exploited via the SL model, which compensates for the performance loss of SL estimation.

5. SIMULATIONS

The analysis was conducted by averaging the steady state error power from 100 independent trials. A quaternion input signal $\mathbf{x}$ with a varying degree of impropriety was generated and the WL system in (13) was constructed, where the length of $\mathbf{x}$ was 4, and the complementary weights of the system were set as $\mathbf{g} = \mathbf{u} = \mathbf{v} = \beta\mathbf{h}$ with $\beta \in \mathbb{R}$ a varying coefficient reflecting the widely linear nature of the system. To avoid the excess MSE, the noise υ was neglected. For simplicity, elements of $\mathbf{x}$ were generated to be uncorrelated and with equal power, so the theoretical performance advantage of WL estimation could be calculated from (22). The impropriety coefficients $r_\imath, r_\jmath, r_\kappa$ of $\mathbf{x}$ were equalised to a unique r. The quaternion least mean square (LMS) algorithms based on the SL and WL models were employed to estimate y. As shown in Figure 1, the theoretical and simulated performance advantage of the widely linear LMS over the strictly linear LMS were in agreement, and they both increased with the norm of the complementary weight vector of the WL system, $\|\mathbf{w}^b\|_2$, and decreased as the impropriety coefficient of $\mathbf{x}$, r, increased. The curves in Fig. 1 (a) and (b) are parabolas, conforming with the quadratic forms in (17) and (23), thus supporting the analysis in Section 3 and Section 4.

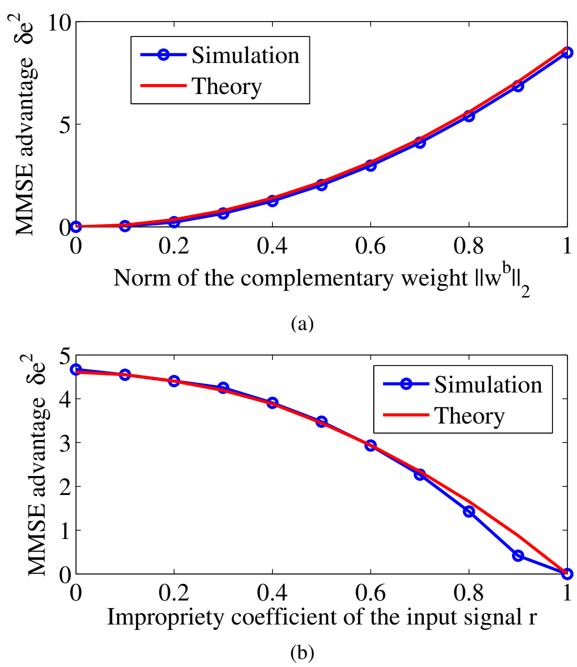

Fig. 1. Effects of (a) widely linear nature of the system and (b) impropriety of the input signal on the MMSE advantage of the quaternion widely linear LMS estimator.

6. CONCLUSION

We have examined the performance relationship between the SL and and WL models for MMSE estimation of quaternion signals. The WL model has been proved to achieve a lower or equal MSE compared with the SL model, and this performance advantage has been found to be dependent on the degree of widely linear nature of the system, impropriety and power of the input signal. Simulations on synthetic signals support the analysis. The excess MSE is out of scope of this paper, and is an interesting topic in further research.

7. REFERENCES

[1] R. Mukundan, "Quaternions: From classical mechanics to computer graphics, and beyond," in *Proc. 7th Asian Technol. Conf. Math.*, pp. 97–105, 2002.

[2] D. P. Mandic, C. Jahanchahi, and C. Cheong Took, "A quaternion gradient operator and its applications," *IEEE Signal Process. Lett.*, vol. 18, no. 1, pp. 47–50, 2011.

[3] D. Xu, C. Jahanchahi, C. Cheong Took, and D. P. Mandic, "Enabling quaternion derivatives: the generalized HR calculus," *Royal Society open science*, vol. 2, no. 8, p. 150255, 2015.

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

[5] C. Cheong Took and D. P. Mandic, "Augmented second-order statistics of quaternion random signals," *Signal Process.*, vol. 91, no. 2, pp. 214–224, 2011.

[6] C. Jahanchahi, C. Cheong Took, and D. P. Mandic, "On HR calculus, quaternion valued stochastic gradient, and adaptive three dimensional wind forecasting," in *Proc. 2010 Int. Joint Conf. Neural Netw. (IJCNN)*, pp. 1–5, 2010.

[7] S. P. Talebi and D. P. Mandic, "A quaternion frequency estimator for three-phase power systems," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, pp. 3956–3960, 2015.

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

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

[10] T. Nitta, "A theoretical foundation for the widely linear processing of quaternion-valued data," *Appl. Math.*, vol. 4, no. 12, pp. 1616–1620, 2013.

[11] C. Jahanchahi, S. Kanna, and D. P. Mandic, "A theoretical bound for the performance advantage of quaternion widely linear estimation," in *Proc. 19th Int. Conf. Digital Signal Process. (DSP)*, pp. 254–258, 2014.

[12] Y. Xia, C. Jahanchahi, T. Nitta, and D. P. Mandic, "Performance bounds of quaternion estimators," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 12, pp. 3287–3292, 2015.

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

[14] M. Xiang, S. Enshaeifar, A. Stott, C. Cheong Took, and D. P. Mandic, "Diagonalisation of covariance matrices in quaternion widely linear signal processing," submitted for publication.

[15] C. Cheong Took, S. Douglas, and D. P. Mandic, "On approximate diagonalization of correlation matrices in widely linear signal processing," *IEEE Trans. Signal Process.*, vol. 60, no. 3, pp. 1469–1473, 2012.

[16] T. A. Ell and S. J. Sangwine, "Quaternion involutions and anti-involutions," *Comput. Math. Applcat.*, vol. 53, no. 1, pp. 137–143, 2007.

[17] S. P. Talebi, S. Kanna, and D. P. Mandic, "Real-time estimation of quaternion impropriety," in *Proc. IEEE Int. Conf. Digital Signal Process. (DSP)*, pp. 557–561, 2015.

[18] J. Vía, D. Palomar, L. Vielva, and I. Santamaría, "Quaternion ICA from second-order statistics," *IEEE Trans. Signal Process.*, vol. 59, no. 4, pp. 1586–1600, 2011.

[19] R. Sznajder, M. S. Gowda, and M. M. Moldovan, "More results on Schur complements in Euclidean Jordan algebras," *J. Global Optimization*, vol. 53, no. 1, pp. 121–134, 2012.

[20] D. P. Mandic and V. S. L. Goh, *Complex Valued Nonlinear Adaptive Filters: Noncircularity, Widely Linear and Neural Models*. John Wiley & Sons, 2009.

[21] S. C. Douglas and D. P. Mandic, "Performance analysis of the conventional complex LMS and augmented complex LMS algorithms," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, pp. 3794–3797, 2010.