

# Cost-effective quaternion minimum mean square error estimation: From widely linear to four-channel processing



Min Xiang<sup>a,\*</sup>, Clive Cheong Took<sup>b</sup>, Danilo P. Mandic<sup>a</sup>

<sup>a</sup> Department of Electrical and Electronic Engineering, Imperial College London, United Kingdom

<sup>b</sup> Department of Computer Science, The University of Surrey, United Kingdom

## ARTICLE INFO

### Keywords:

Quaternion filter  
Mean square error  
Widely linear estimation  
Four-channel model  
Computational complexity

## ABSTRACT

Widely linear estimation plays an important role in quaternion signal processing, as it caters for both proper and improper quaternion signals. However, widely linear algorithms are computationally expensive owing to the use of augmented variables and statistics. To reduce the computation cost while maintaining the performance level, we propose a four-channel estimation framework as an efficient alternative to quaternion widely linear estimation. This is achieved by using four linear models to estimate the four components of quaternion signals. We also show that any of the four channels is able to replace a strictly linear quaternion estimator when estimating strictly linear systems. The proposed method is shown to reduce computational complexity and provide more flexible algorithms, while preserving the physical meaning inherent in the quaternion domain. The proposed framework is next applied to quaternion minimum mean square error estimation to yield the reduced-complexity versions of the quaternion least mean square (QLMS), quaternion recursive least squares (QRLS), and quaternion nonlinear gradient decent (QNGD) algorithms. For the proposed QLMS algorithm, an adaptive step-size strategy is also explored. The effectiveness of the so introduced estimation techniques is validated by simulations on synthetic and real-world signals.

## 1. Introduction

Recent advances in sensing technology have enabled ubiquitous recording from 3-D and 4-D data sources, such as measurements from seismometers [1], ultrasonic anemometers [2], and inertial body sensors [3]. Traditionally, these measurements have been considered as vectors in the  $\mathbb{R}^3$  and  $\mathbb{R}^4$  fields of reals, however, the vector algebra is not a division algebra and is inadequate when modelling orientation and rotation [4]. Owing to their division algebra, quaternions have inherent advantages in representing 3-D and 4-D data, and quaternion-valued algorithms are a generic extension of their real- and complex-valued counterparts. Quaternions also naturally account for mutual information between multiple data channels, provide a compact representation, and have proven to offer a physically meaningful interpretation to real-world applications, such as in navigation, communications, and image processing [5,6]. Recent resurgence in research on quaternion signal processing spans the areas of adaptive filtering [7], neural networks [8], independent component analysis [9], and spectral estimation [10].

When it comes to adaptive filtering, estimation of a set of signals based on information obtained from measurements of other signals,

which is widely used in signal enhancement and system identification, has been recently extended from the real and complex domains to the quaternion domain  $\mathbb{H}$  [7]. Traditional strictly linear quaternion estimators utilise the second-order statistics based on the standard covariance and are optimal only for estimating second-order circular (proper) signals. Advances in quaternion statistics have established that in order to capture complete second-order statistics of quaternion signals, widely linear estimators which use the complementary covariances, in addition to the standard covariance, are needed [11,12]. However, the widely linear approach requires four times the number of parameters to update, inducing higher computational complexity. To reduce the computation cost, efficient algorithms casting the computations from  $\mathbb{H}$  to  $\mathbb{R}$  have been proposed, for example, the reduced-complexity widely linear quaternion least mean square (RC-WL-QLMS) algorithm employs a quaternion-valued weight vector and a real-valued input vector [13], while the multi-channel LMS (MLMS) algorithm exploits the duality between  $\mathbb{R}^4$  and  $\mathbb{H}$  [14]. Despite computational efficiency, these algorithms have no means of exploring the physical meaning inherent in  $\mathbb{H}$ , as the input and weight vectors no longer reside in  $\mathbb{H}$ .

In order to obtain physically meaningful estimates of quaternion

\* Corresponding author.

E-mail address: [m.xiang13@ic.ac.uk](mailto:m.xiang13@ic.ac.uk) (M. Xiang).

signals at low computation cost, we here extend the recently proposed complex dual channel estimation [15] to a four-channel quaternion estimation framework, which comprises four sub-estimators for the four components of the quaternion. In this way, four degrees of freedom are provided to capture complete second-order statistical information; the so introduced estimation model and weight update rule reduce the computational cost compared with the conventional widely linear estimation, while achieving equivalent estimation performance. In this way, the physical meaning inherent to the quaternion domain is also preserved, together with enhanced flexibility as the convergence of the four sub-estimators can be controlled individually. When estimating strictly linear quaternion systems, the proposed four-channel estimation simplifies to a single-channel problem, which can also be used as an efficient alternative to conventional strictly linear estimation. For rigour, we derive a class of four-channel quaternion minimum mean square error (MMSE) estimation algorithms, including the four-channel quaternion least mean square (FC-QLMS), the four-channel quaternion recursive least squares (FC-QRLS), and the four-channel quaternion nonlinear gradient decent (FC-QNGD). The transient and steady-state behaviour and computational complexity of the FC-QLMS algorithm are established, and an adaptive step-size strategy based on the Barzilai–Borwein method is introduced. Simulations on synthetic and real-world multi-dimensional signals support the analysis.

The rest of this paper is organised as follows. Section 2 provides a background of quaternions and quaternion-valued signal estimation. The novel four-channel quaternion estimation framework is introduced in Section 3. Section 4 applies the four-channel estimation framework to quaternion MMSE estimation and introduces the corresponding algorithms. Section 5 analyses the performance of the FC-QLMS algorithm and introduces an adaptive step-size strategy. Simulation results are given in Section 6, and Section 7 concludes the paper. Throughout the paper, we use boldface capital letters to denote matrices,  $\mathbf{A}$ , boldface lowercase letters for vectors,  $\mathbf{a}$ , and italic 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 four-dimensional vector space over the real field  $\mathbb{R}$ , spanned by the basis  $\{1, i, j, \kappa\}$ . A random quaternion vector,  $\mathbf{x} \in \mathbb{H}^{L \times 1}$ , comprises a real part  $\Re\{\cdot\}$  and an imaginary part  $\Im\{\cdot\}$  which consists of three imaginary components,  $i\Im_i, j\Im_j, \kappa\Im_\kappa$ , so that

$$\mathbf{x} = \Re[\mathbf{x}] + \Im[\mathbf{x}] = \Re[\mathbf{x}] + i\Im_i[\mathbf{x}] + j\Im_j[\mathbf{x}] + \kappa\Im_\kappa[\mathbf{x}]$$

where  $\Re[\mathbf{x}], \Im_i[\mathbf{x}], \Im_j[\mathbf{x}], \Im_\kappa[\mathbf{x}] \in \mathbb{R}^{L \times 1}$ , and  $i, j, \kappa$  are the roots of -1 which satisfy

$$i^2 = j^2 = \kappa^2 = i\kappa = -1 \quad ij = -ji = \kappa \quad j\kappa = -\kappa j = i \quad \kappa i = -i\kappa = j$$

The conjugate of a random quaternion vector  $\mathbf{x}$  is defined as

$$\mathbf{x}^* = \Re[\mathbf{x}] - \Im[\mathbf{x}] = \Re[\mathbf{x}] - i\Im_i[\mathbf{x}] - j\Im_j[\mathbf{x}] - \kappa\Im_\kappa[\mathbf{x}]$$

The modulus of a quaternion variable  $x \in \mathbb{H}$  is defined as

$$|x| = \sqrt{\Re^2[x] + \Im_i^2[x] + \Im_j^2[x] + \Im_\kappa^2[x]}$$

and the product of two quaternions  $x, y \in \mathbb{H}$  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 ‘ $\times$ ’ the vector product. The presence of the vector product causes the non-commutativity of the quaternion product, that is,  $xy \neq yx$ . The quaternion

product has the following properties [16]:

$$|xy| = |x| |y|, \quad x^{-1} = \frac{x^*}{|x|^2}, \quad (xy)^{-1} = y^{-1}x^{-1}, \quad (xy)^* = y^*x^* \quad (1)$$

A quaternion variable  $x$  is called a unit quaternion if it satisfies  $|x| = 1$ . A quaternion variable  $x$  is called a pure quaternion if it satisfies  $\Re[x] = 0$ .

Another important notion is that of the quaternion involution [17], which defines a self-inverse mapping analogous to the complex conjugate. The general involution of the quaternion vector  $\mathbf{x}$  is defined as  $\mathbf{x}^\eta = -\eta\mathbf{x}\eta$ , and represents the rotation of the vector part of  $\mathbf{x}$  by  $\pi$  about a unit pure quaternion  $\eta$ . The involutions obey  $(\mathbf{x}^\eta)^\eta = \mathbf{x}$ , and the conjugate involutions are defined as  $\mathbf{x}^{\eta*} = (\mathbf{x}^\eta)^* = (\mathbf{x}^*)^\eta$ . The special cases of involutions about the  $i, j$  and  $\kappa$  imaginary axes are given by

$$\begin{aligned} \mathbf{x}^i &= -i\mathbf{x}i = \Re[\mathbf{x}] + i\Im_i[\mathbf{x}] - j\Im_j[\mathbf{x}] - \kappa\Im_\kappa[\mathbf{x}] \\ \mathbf{x}^j &= -j\mathbf{x}j = \Re[\mathbf{x}] - i\Im_i[\mathbf{x}] + j\Im_j[\mathbf{x}] - \kappa\Im_\kappa[\mathbf{x}] \\ \mathbf{x}^\kappa &= -\kappa\mathbf{x}\kappa = \Re[\mathbf{x}] - i\Im_i[\mathbf{x}] - j\Im_j[\mathbf{x}] + \kappa\Im_\kappa[\mathbf{x}] \end{aligned} \quad (2)$$

Through the above quaternion involutions, the four real-valued components of  $\mathbf{x}$  can be expressed as

$$\begin{aligned} \Re[\mathbf{x}] &= \frac{1}{4}(\mathbf{x} + \mathbf{x}^i + \mathbf{x}^j + \mathbf{x}^\kappa) \\ \Im_i[\mathbf{x}] &= \frac{1}{4i}(\mathbf{x} + \mathbf{x}^i - \mathbf{x}^j - \mathbf{x}^\kappa) \\ \Im_j[\mathbf{x}] &= \frac{1}{4j}(\mathbf{x} - \mathbf{x}^i + \mathbf{x}^j - \mathbf{x}^\kappa) \\ \Im_\kappa[\mathbf{x}] &= \frac{1}{4\kappa}(\mathbf{x} - \mathbf{x}^i - \mathbf{x}^j + \mathbf{x}^\kappa) \end{aligned} \quad (3)$$

### 2.2. Second-order statistics

The set of quaternion involutions in (2) and the original quaternion vector  $\mathbf{x}$  form the most frequently used basis for augmented quaternion statistics, which is at the core of the recently proposed widely linear processing methodology [18,19]. Benefiting from this involution basis, complete second-order quaternion statistics is described by the  $i$ -,  $j$ -, and  $\kappa$ - complementary covariance matrices, together with the standard Hermitian covariance matrix,  $\mathbf{C}_x = E\{\mathbf{x}\mathbf{x}^H\}$ . The complementary covariance matrices can be represented in a unified form as  $\mathbf{C}_{\mathbf{x}\mathbf{x}^\eta} = E\{\mathbf{x}\mathbf{x}^{\eta H}\}$ ,  $\eta \in \{i, j, \kappa\}$ , and every  $\eta$ -complementary covariance matrix is  $\eta$ - Hermitian, that is,  $\mathbf{C}_{\mathbf{x}\mathbf{x}^\eta} = (\mathbf{C}_{\mathbf{x}\mathbf{x}^\eta})^{\eta H}$ . The knowledge of both the covariance matrix and the complementary covariance matrices is necessary to ensure the utilisation of complete second-order statistical information in  $\mathbb{H}$ .

The notion of non-circularity (improperness) is unique to division algebras. For quaternion data, non-circularity refers to probability distributions which are not rotation-invariant, while improperness is characterised by the degree of correlation and/or power imbalance between the real and imaginary components. The additional degrees of freedom in the quaternion domain allow for types of properness:  $\mathbb{R}^\eta$ -properness,  $\mathbb{C}^\eta$ -properness, and  $\mathbb{H}$ -properness [18,19], defined below.

**Definition 1 ( $\mathbb{H}$ -properness).** A random quaternion vector  $\mathbf{x}$  is  $\mathbb{H}$ -proper if it is uncorrelated with the involutions  $\mathbf{x}^i, \mathbf{x}^j$  and  $\mathbf{x}^\kappa$ , so that  $\mathbf{C}_{\mathbf{x}\mathbf{x}^i} = \mathbf{C}_{\mathbf{x}\mathbf{x}^j} = \mathbf{C}_{\mathbf{x}\mathbf{x}^\kappa} = \mathbf{0}$ .

**Definition 2 ( $\mathbb{R}^\eta$ -properness).** A random quaternion vector  $\mathbf{x}$  is  $\mathbb{R}^\eta$ -proper with respect to an imaginary unit  $\eta \in \{i, j, \kappa\}$  if it is uncorrelated only with the involutions  $\mathbf{x}^\eta$ , so that  $\mathbf{C}_{\mathbf{x}\mathbf{x}^\eta} = \mathbf{0}$ .

**Definition 3 ( $\mathbb{C}^\eta$ -properness).** A random quaternion vector  $\mathbf{x}$  is  $\mathbb{C}^\eta$ -improper with respect to an imaginary unit  $\eta \in \{i, j, \kappa\}$  if it is correlated with only one of the involutions,  $\mathbf{x}^\eta$ , so that all complementary covariances except for  $\mathbf{C}_{\mathbf{x}\mathbf{x}^\eta}$  vanish.

### 2.3. Quaternion estimation

A fundamental problem in quaternion signal processing is to obtain the estimate,  $\hat{y}$ , of a desired signal,  $y \in \mathbb{H}$ , from a set of measurements,  $\mathbf{x} \in \mathbb{H}^{L \times 1}$ , which carry information about  $y$ . The estimation model,  $\hat{y} = f(\mathbf{x})$ , incorporates knowledge about the relationship between  $y$  and  $\mathbf{x}$ , is crucial for estimation performance. Traditionally, the strictly linear model has been used for this purpose and is given by [7]

$$\hat{y} = \hat{\mathbf{h}}^H \mathbf{x} \quad (4)$$

where  $\hat{\mathbf{h}}$  is the weight vector. This model achieves optimal estimation for proper quaternion signals but is suboptimal for general improper quaternion signals. To address this issue, the incorporation of the involution basis for  $\mathbf{x}$  yields the widely linear model [11]

$$\hat{y} = \hat{\mathbf{h}}^H \mathbf{x} + \hat{\mathbf{g}}^H \mathbf{x}^1 + \hat{\mathbf{u}}^H \mathbf{x}^j + \hat{\mathbf{v}}^H \mathbf{x}^\kappa = \hat{\mathbf{w}}^a H \mathbf{x}^a \quad (5)$$

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  $\hat{\mathbf{w}}^a = [\hat{\mathbf{h}}^T, \hat{\mathbf{g}}^T, \hat{\mathbf{u}}^T, \hat{\mathbf{v}}^T]^T$ , and  $\mathbf{x}^a = [\mathbf{x}^T, \mathbf{x}^{1T}, \mathbf{x}^{jT}, \mathbf{x}^{\kappa T}]^T$  is the corresponding augmented input vector. The widely linear estimation has proven to outperform the strictly linear estimation for general quaternion signals [20].

### 3. Four-channel quaternion estimation model

Consider a quaternion-valued desired signal,  $y$ , given by a widely linear system corrupted by independent noise,  $v$ , in the form:

$$y = \mathbf{h}^H \mathbf{x} + \mathbf{g}^H \mathbf{x}^1 + \mathbf{u}^H \mathbf{x}^j + \mathbf{v}^H \mathbf{x}^\kappa + v \quad (6)$$

where  $\mathbf{h}, \mathbf{g}, \mathbf{u}, \mathbf{v} \in \mathbb{H}^{L \times 1}$  are the true weight vectors. According to the relationship (3), the four real-valued components of  $y$  can be represented by four linear models, as

$$\begin{aligned} \Re\{y\} &= \Re\{(\mathbf{h} + \mathbf{g}^1 + \mathbf{u}^j + \mathbf{v}^\kappa)^H \mathbf{x} + v\} \\ \Im_1\{y\} &= \Im_1\{(\mathbf{h} + \mathbf{g}^1 - \mathbf{u}^j - \mathbf{v}^\kappa)^H \mathbf{x} + v\} \\ \Im_j\{y\} &= \Im_j\{(\mathbf{h} - \mathbf{g}^1 + \mathbf{u}^j - \mathbf{v}^\kappa)^H \mathbf{x} + v\} \\ \Im_\kappa\{y\} &= \Im_\kappa\{(\mathbf{h} - \mathbf{g}^1 - \mathbf{u}^j + \mathbf{v}^\kappa)^H \mathbf{x} + v\} \end{aligned} \quad (7)$$

Therefore, the desired signal (6) can be expressed as

$$y = \Re\{\mathbf{w}_1^H \mathbf{x}\} + \Im_1\{\mathbf{w}_1^H \mathbf{x}\} + \Im_j\{\mathbf{w}_j^H \mathbf{x}\} + \Im_\kappa\{\mathbf{w}_\kappa^H \mathbf{x}\} + v \quad (8)$$

where  $\mathbf{w}_1, \mathbf{w}_j, \mathbf{w}_\kappa$  are quaternion-valued weight vectors. This motivates us to separately estimate the four real-valued components,  $\Re\{\hat{y}_1\}$ ,  $\Im_1\{\hat{y}_1\}$ ,  $\Im_j\{\hat{y}_j\}$ ,  $\Im_\kappa\{\hat{y}_\kappa\}$ , of  $y$ , in the form:

$$\hat{y}_1 = \hat{\mathbf{w}}_1^H \mathbf{x}, \quad \hat{y}_j = \hat{\mathbf{w}}_j^H \mathbf{x}, \quad \hat{y}_\kappa = \hat{\mathbf{w}}_\kappa^H \mathbf{x} \quad (9)$$

which are respectively referred to as channel 1,  $j$  and  $\kappa$ . Then these components can be combined into a quaternion-valued estimate

$$\hat{y} = \Re\{\hat{y}_1\} + \Im_1\{\hat{y}_1\} + \Im_j\{\hat{y}_j\} + \Im_\kappa\{\hat{y}_\kappa\} \quad (10)$$

for which the estimation error is given by

$$e = y - \hat{y} = e_1 + \Im_1\{e_1\} + \Im_j\{e_j\} + \Im_\kappa\{e_\kappa\}$$

where

$$e_1 = \Re\{y\} - \Re\{\hat{y}_1\}, \quad \Im_1\{e_1\} = \Im_1\{y\} - \Im_1\{\hat{y}_1\}, \quad \Im_j\{e_j\} = \Im_j\{y\} - \Im_j\{\hat{y}_j\},$$

$$e_\kappa = \Im_\kappa\{y\} - \Im_\kappa\{\hat{y}_\kappa\}$$

are the respective estimation errors in the four channels (components). We refer to Eqs. (9) and (10) as the four-channel linear estimation model, which avoids the summation of the four quaternion quantities in the widely linear estimation model (5) and consequently reduces the computation cost.

### 4. Four-channel quaternion MMSE estimation

The MMSE estimation is a widely used estimation technique which aims to minimise the mean square error (MSE),  $E\{e^2\}$ . On the basis of the four-channel linear model in (9) and (10), we shall now obtain the optimal estimate by analysing individual channels. For example, let  $\hat{\mathbf{w}}_{1,o}$  denote the optimal estimate of the system weight  $\mathbf{w}_1$ . Then the optimal estimate of  $\Re\{y\}$  is given by

$$\Re\{\hat{\mathbf{w}}_{1,o}^H \mathbf{x}\} = \frac{1}{4}[\hat{\mathbf{w}}_{1,o}^H \mathbf{x} + \hat{\mathbf{w}}_{1,o}^{1H} \mathbf{x}^1 + \hat{\mathbf{w}}_{1,o}^{jH} \mathbf{x}^j + \hat{\mathbf{w}}_{1,o}^{\kappa H} \mathbf{x}^\kappa]$$

which is the projection of  $\Re\{y\}$  onto a Hilbert subspace spanned by the basis in (2), and obeys the orthogonality principle

$$e_1 \perp \mathbf{x} \quad e_1 \perp \mathbf{x}^1 \quad e_1 \perp \mathbf{x}^j \quad e_1 \perp \mathbf{x}^\kappa$$

where the symbol ‘ $\perp$ ’ designates that the error  $e_1$  is orthogonal to  $\mathbf{x}$ ,  $\mathbf{x}^1$ ,  $\mathbf{x}^j$  and  $\mathbf{x}^\kappa$ . This orthogonality condition is equivalent to

$$E\{\mathbf{x}e_1^*\} = E\{\mathbf{x}^1e_1^{*1}\} = E\{\mathbf{x}^je_1^{*j}\} = E\{\mathbf{x}^\kappa e_1^{*\kappa}\} = \mathbf{0} \quad (11)$$

Substituting  $e_1 = \Re\{(\mathbf{w}_1 - \hat{\mathbf{w}}_{1,o})^H \mathbf{x}\} + \Re\{v\}$  into (11) yields

$$E\{\mathbf{x}\mathbf{x}^{aH}\} \begin{bmatrix} \hat{\mathbf{w}}_{1,o}^1 \\ \hat{\mathbf{w}}_{1,o}^j \\ \hat{\mathbf{w}}_{1,o}^\kappa \end{bmatrix} = E\{\mathbf{x}\mathbf{x}^{aH}\} \begin{bmatrix} \mathbf{w}_1^1 \\ \mathbf{w}_1^j \\ \mathbf{w}_1^\kappa \end{bmatrix}$$

which implies  $\hat{\mathbf{w}}_{1,o} = \mathbf{w}_1$ . Similarly, the remaining three optimal estimated weights can be derived as  $\hat{\mathbf{w}}_{j,o} = \mathbf{w}_j$ ,  $\hat{\mathbf{w}}_{\kappa,o} = \mathbf{w}_\kappa$ .

For  $\mathbf{w}_1 = \mathbf{w}_j = \mathbf{w}_\kappa$ , the desired signal (8) is expressed as a strictly linear model  $y = \mathbf{w}_1^H \mathbf{x} + v$ , and the optimal estimated weight vectors of the four-channel estimator are  $\hat{\mathbf{w}}_{1,o} = \hat{\mathbf{w}}_{j,o} = \hat{\mathbf{w}}_{\kappa,o} = \hat{\mathbf{w}}_{\kappa,o} = \mathbf{w}_1$ . Therefore, we can use a single sub-channel to estimate the weight vector  $\mathbf{w}_1$ , and hence obtain

$$\hat{y} = \Re\{\hat{\mathbf{w}}_1^H \mathbf{x}\} + \Im_1\{\hat{\mathbf{w}}_1^H \mathbf{x}\} + \Im_j\{\hat{\mathbf{w}}_1^H \mathbf{x}\} + \Im_\kappa\{\hat{\mathbf{w}}_1^H \mathbf{x}\} = \hat{\mathbf{w}}_1^H \mathbf{x}$$

This reduced form of the four-channel estimation is referred to as single-channel estimation.

We next introduce three adaptive algorithms for quaternion MMSE estimation within the proposed four-channel framework.

#### 4.1. FC-QLMS algorithm

Based on the four-channel linear estimation model in (9) and (10), a FC-QLMS algorithm is next derived for quaternion estimation. As shown in Fig. 1, the desired signal at the  $n$ -th iteration,  $y_n$ , is estimated as

$$\hat{y}_n = \Re\{\hat{\mathbf{w}}_{1,n}^H \mathbf{x}_n\} + \Im_1\{\hat{\mathbf{w}}_{1,n}^H \mathbf{x}_n\} + \Im_j\{\hat{\mathbf{w}}_{j,n}^H \mathbf{x}_n\} + \Im_\kappa\{\hat{\mathbf{w}}_{\kappa,n}^H \mathbf{x}_n\}$$

and the weights are updated to minimise the cost function

$$J_n^{\text{LMS}} = |y_n - \hat{y}_n|^2 = e_{1,n}^2 + e_{1,n}^2 + e_{j,n}^2 + e_{\kappa,n}^2$$

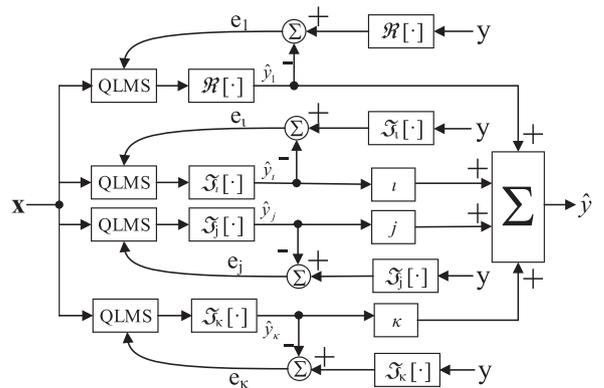


Fig. 1. Architecture of the FC-QLMS algorithm.

where

$$\begin{aligned} e_{1,n} &= \Re[y_n] - \Re[\widehat{\mathbf{w}}_{1,n}^H \mathbf{x}_n] \\ e_{i,n} &= \Im_i[y_n] - \Im_i[\widehat{\mathbf{w}}_{i,n}^H \mathbf{x}_n] \\ e_{j,n} &= \Im_j[y_n] - \Im_j[\widehat{\mathbf{w}}_{j,n}^H \mathbf{x}_n] \\ e_{\kappa,n} &= \Im_{\kappa}[y_n] - \Im_{\kappa}[\widehat{\mathbf{w}}_{\kappa,n}^H \mathbf{x}_n] \end{aligned}$$

As the maximum change of  $J_n^{\text{LMS}}$  occurs when the weight update is collinear with the conjugate derivative of  $J_n^{\text{LMS}}$  [21], the weight at the  $(n+1)$ -th iteration can be derived from the  $n$ -th iteration as

$$\widehat{\mathbf{w}}_{\eta,n+1} = \widehat{\mathbf{w}}_{\eta,n} - \mu_{\text{fc}} \nabla_{\widehat{\mathbf{w}}_{\eta,n}^*} J_n^{\text{LMS}} \quad (12)$$

where  $\eta \in \{1, i, j, \kappa\}$ , and  $\mu_{\text{fc}}$  is a positive step-size. The derivative in (12) is calculated based on the generalised HR calculus [22] as

$$\nabla_{\widehat{\mathbf{w}}_{\eta,n}^*} J_n^{\text{LMS}} = -\frac{1}{2} e_{\eta,n} \mathbf{x}_n \eta^{-1} \quad (13)$$

Therefore, the component-wise weight updates are given by

$$\widehat{\mathbf{w}}_{\eta,n+1} = \widehat{\mathbf{w}}_{\eta,n} + \mu_{\text{fc}} e_{\eta,n} \mathbf{x}_n \eta^{-1} \quad (14)$$

where the constant  $\frac{1}{2}$  in (13) is absorbed into the step-size  $\mu_{\text{fc}}$ .

If the desired signal obeys the strictly linear model, the four-channel estimation can be replaced with single-channel estimation, so that the FC-QLMS simplifies to a single-channel QLMS (SC-QLMS) algorithm. The SC-QLMS using channel 1 is therefore given by

$$\begin{aligned} \widehat{y}_n &= \widehat{\mathbf{w}}_{1,n}^H \mathbf{x}_n \\ \widehat{\mathbf{w}}_{1,n+1} &= \widehat{\mathbf{w}}_{1,n} + \mu_{\text{fc}} (\Re[y_n] - \Re[\widehat{y}_n]) \mathbf{x}_n \end{aligned} \quad (15)$$

The proposed model is also scalable, for example, when the quaternion input signal reduces to a complex one, that is,  $\mathbf{x} = \Re[\mathbf{x}] + i\Im[\mathbf{x}]$ , the FC-QLMS algorithm reduces to a complex algorithm, as

$$\begin{aligned} \widehat{y}_n &= \Re[\widehat{\mathbf{w}}_{1,n}^H \mathbf{x}_n] + i\Im_i[\widehat{\mathbf{w}}_{1,n}^H \mathbf{x}_n] \\ \widehat{\mathbf{w}}_{1,n+1} &= \widehat{\mathbf{w}}_{1,n} + \mu_{\text{fc}} \Re[y - \widehat{y}_n] \mathbf{x}_n \\ \widehat{\mathbf{w}}_{i,n+1} &= \widehat{\mathbf{w}}_{i,n} - \mu_{\text{fc}} \Im_i[y - \widehat{\mathbf{w}}_{i,n}^H \mathbf{x}_n] \mathbf{x}_n \end{aligned}$$

which is equivalent to the dual channel complex LMS algorithm in [15].

In addition to the reduction of computational cost owing to the four-channel estimation model, the FC-QLMS has a more efficient weight update rule than existing QLMS algorithms [22,23]. Another advantage of FC-QLMS is that the step-sizes of the four channels can be different, and thus the convergence rate of each channel can be controlled individually, which provides the algorithm with more flexibility in practical applications.

#### 4.2. FC-QRLS algorithm

The class of quaternion recursive least squares (QRLS) algorithms include the strictly linear QRLS (SL-QRLS) algorithm given by [24]

$$\begin{aligned} e_n^p &= y_n - \widehat{\mathbf{h}}_{n-1}^H \mathbf{x}_n \\ \mathbf{p}_n &= \mathbf{P}_{n-1} \mathbf{x}_n (\rho + \mathbf{x}_n^H \mathbf{P}_{n-1} \mathbf{x}_n)^{-1} \\ \widehat{\mathbf{h}}_n &= \widehat{\mathbf{h}}_{n-1} + \mathbf{p}_n e_n^{p*} \\ \mathbf{P}_n &= \rho^{-1} [\mathbf{P}_{n-1} - \mathbf{p}_n \mathbf{x}_n^H \mathbf{P}_{n-1}] \end{aligned}$$

and the widely linear QRLS (WL-QRLS) algorithm given by [24]

$$\begin{aligned} e_n^p &= y_n - \widehat{\mathbf{w}}_{n-1}^{aH} \mathbf{x}_n^a \\ \mathbf{p}_n^a &= \mathbf{P}_{n-1}^a \mathbf{x}_n^a (\rho + \mathbf{x}_n^{aH} \mathbf{P}_{n-1}^a \mathbf{x}_n^a)^{-1} \\ \widehat{\mathbf{w}}_n^a &= \widehat{\mathbf{w}}_{n-1}^a + \mathbf{p}_n^a e_n^{p*} \\ \mathbf{P}_n^a &= \rho^{-1} [\mathbf{P}_{n-1}^a - \mathbf{p}_n^a \mathbf{x}_n^{aH} \mathbf{P}_{n-1}^a] \end{aligned} \quad (16)$$

where  $\mathbf{P}_n \in \mathbb{H}^{L \times L}$  and  $\mathbf{P}_n^a \in \mathbb{H}^{4L \times 4L}$  are the estimated inverses of the covariance and augmented covariance matrices. We next introduce an FC-QRLS version of the WL-QRLS algorithm, which aims to minimise the cost function

$$\begin{aligned} J_n^{\text{RLS}} &= \sum_{m=0}^n \rho^{n-m} |e_m|^2 \\ e_m &= y_m - \Re[\widehat{\mathbf{w}}_{1,m}^H \mathbf{x}_m] - i\Im_i[\widehat{\mathbf{w}}_{i,m}^H \mathbf{x}_m] - j\Im_j[\widehat{\mathbf{w}}_{j,m}^H \mathbf{x}_m] - \kappa\Im_{\kappa}[\widehat{\mathbf{w}}_{\kappa,m}^H \mathbf{x}_m] \end{aligned}$$

where the weight update is derived from the WL-QRLS weight in (16) according to the isomorphism in (7), to yield

$$\widehat{\mathbf{w}}_{\eta,n} = \widehat{\mathbf{w}}_{\eta,n-1} + 4\mathcal{P}[e_n^{p*}] \mathbf{r}_n \eta \quad \eta \in \{1, i, j, \kappa\} \quad (17)$$

In (17),  $\mathbf{r}_n \in \mathbb{H}^{L \times 1}$  contains the first  $L$  entries of  $\mathbf{p}_n^a$ ,  $\mathcal{P}[\cdot] = \Re[\cdot]$  when  $\eta = 1$ ,  $\mathcal{P}[\cdot] = \Im_i[\cdot]$  when  $\eta = i$ ,  $\mathcal{P}[\cdot] = \Im_j[\cdot]$  when  $\eta = j$ ,  $\mathcal{P}[\cdot] = \Im_{\kappa}[\cdot]$  when  $\eta = \kappa$ . The weight calculation in (17) is more efficient than the weight update of WL-QRLS in (16). If the desired signal obeys the strictly linear model, a single sub-channel is sufficient for optimal estimation, leading to a more efficient single-channel QRLS (SC-QRLS) algorithm.

#### 4.3. FC-QNGD algorithm

The class of quaternion nonlinear gradient decent (QNGD) algorithms include the strictly linear QNGD (SL-QNGD) algorithm, given by [23,25]

$$\begin{aligned} q_n &= \widehat{\mathbf{h}}_n^H \mathbf{x}_n \\ \widehat{y}_n &= \Phi(q_n) \\ \widehat{\mathbf{h}}_{n+1} &= \widehat{\mathbf{h}}_n + \mu \sum_{\nu \in \{1, i, j, \kappa\}} \mathbf{x}_n \frac{\partial \Phi^{\nu}(q_n)}{\partial q_n} (y_n - \widehat{y}_n)^* \end{aligned}$$

and the widely linear QNGD (WL-QNGD) algorithm, given by [25]

$$\begin{aligned} q_n &= \widehat{\mathbf{w}}_n^{aH} \mathbf{x}_n^a \\ \widehat{y}_n &= \Phi(q_n) \\ \widehat{\mathbf{w}}_{n+1}^a &= \widehat{\mathbf{w}}_n^a + \mu \sum_{\nu \in \{1, i, j, \kappa\}} \mathbf{x}_n^a \frac{\partial \Phi^{\nu}(q_n)}{\partial q_n} (y_n - \widehat{y}_n)^* \end{aligned}$$

where  $\Phi(q_n)$  is a nonlinear function of  $q_n$ ,  $\mu$  is a positive step-size.

The FC-QNGD algorithm can also be derived on the basis of the four-channel estimation model. Consider the minimisation of the cost function

$$J_n^{\text{NGD}} = |e_n|^2 = |y_n - \Phi(q_n)|^2$$

where

$$q_n = \Re[\widehat{\mathbf{w}}_{1,n}^H \mathbf{x}_n] + i\Im_i[\widehat{\mathbf{w}}_{i,n}^H \mathbf{x}_n] + j\Im_j[\widehat{\mathbf{w}}_{j,n}^H \mathbf{x}_n] + \kappa\Im_{\kappa}[\widehat{\mathbf{w}}_{\kappa,n}^H \mathbf{x}_n]$$

The maximum change of  $J_n^{\text{NGD}}$  occurs when the weight update is collinear with the conjugate derivative of  $J_n^{\text{NGD}}$  [21], so the weights can be updated as

$$\widehat{\mathbf{w}}_{\eta,n+1} = \widehat{\mathbf{w}}_{\eta,n} - \mu_{\text{fc}} \nabla_{\widehat{\mathbf{w}}_{\eta,n}^*} J_n^{\text{NGD}}$$

where  $\eta \in \{1, i, j, \kappa\}$  and the derivative is calculated based on the generalised HR calculus [22] as

$$\nabla_{\widehat{\mathbf{w}}_{\eta,n}^*} J_n^{\text{NGD}} = \sum_{\nu \in \{1, i, j, \kappa\}} \frac{\partial |e_n|^2}{\partial e_n^{\nu}} \frac{\partial e_n^{\nu}}{\partial \widehat{\mathbf{w}}_{\eta,n}^*} = -\frac{1}{8} \sum_{\nu \in \{1, i, j, \kappa\}} e_n^{\nu*} \frac{\partial \Phi^{\nu}(q_n)}{\partial \mathcal{P}[q_n]} \mathbf{x}_n \eta^{-1} \quad (18)$$

Therefore, the weight update is given by

$$\widehat{\mathbf{w}}_{\eta,n+1} = \widehat{\mathbf{w}}_{\eta,n} + \mu_{\text{fc}} \sum_{\nu \in \{1, i, j, \kappa\}} e_n^{\nu*} \frac{\partial \Phi^{\nu}(q_n)}{\partial \mathcal{P}[q_n]} \mathbf{x}_n \eta^{-1} \quad (19)$$

where the constant  $\frac{1}{8}$  in (18) is absorbed into the step-size  $\mu_{\text{fc}}$ .

If  $\Phi(q)$  is a split quaternion nonlinear function  $\Phi(q) = \Phi_1(q) + i\Phi_i(q) + j\Phi_j(q) + \kappa\Phi_{\kappa}(q)$ , where  $\Phi_1(q)$  is a real-valued nonlinear function applied to  $\Re[q]$ ,  $\Phi_i(q)$  to  $\Im_i[q]$ ,  $\Phi_j(q)$  to  $\Im_j[q]$ , and  $\Phi_{\kappa}(q)$  to  $\Im_{\kappa}[q]$ , then the cost function can be written as

$J_n^{\text{NGD}} = \sum_{\eta \in \{1, j, \kappa\}} e_{\eta, n}^2$ , where  $e_{\eta, n} = \mathcal{P}[y_n] - \Phi_\eta(q_n)$  are split errors. The conjugate derivative of  $J_n^{\text{NGD}}$  is derived as

$$\nabla_{\widehat{\mathbf{w}}_\eta^*} J_n^{\text{NGD}} = \frac{\partial J_n^{\text{NGD}}}{\partial e_{\eta, n}} \frac{\partial e_{\eta, n}}{\partial \widehat{\mathbf{w}}_\eta^*} = -\frac{1}{2} e_{\eta, n} \frac{\partial \Phi_\eta(q_n)}{\partial \mathcal{P}[q_n]} \mathbf{x}_n \eta^{-1} \quad (20)$$

and the weights are updated as

$$\widehat{\mathbf{w}}_{\eta, n+1} = \widehat{\mathbf{w}}_{\eta, n} - \mu_{\text{fc}} \nabla_{\widehat{\mathbf{w}}_\eta^*} J_n^{\text{NGD}} = \widehat{\mathbf{w}}_{\eta, n} + \mu_{\text{fc}} e_{\eta, n} \frac{\partial \Phi_\eta(q_n)}{\partial \mathcal{P}[q_n]} \mathbf{x}_n \eta^{-1} \quad (21)$$

where the constant  $\frac{1}{2}$  in (20) is absorbed into the step-size  $\mu_{\text{fc}}$ .

The FC-QNGD weight update is obviously less complicated than the WL-QNGD weight update. When the desired signal obeys the strictly linear model, a single sub-channel is sufficient for optimal estimation, yielding a more efficient single-channel QNGD (SC-QNGD) algorithm.

## 5. Performance of FC-QLMS

The performance of the four-channel estimation with the FC-QLMS algorithm is next compared with existing QLMS algorithms.

### 5.1. Equivalence of the FC-QLMS and WL-QLMS

Recently proposed QLMS algorithms include the strictly linear quaternion least mean square (SL-QLMS) algorithm [23]

$$\begin{aligned} \widehat{y}_n &= \widehat{\mathbf{h}}_n^H \mathbf{x}_n \\ \widehat{\mathbf{h}}_{n+1} &= \widehat{\mathbf{h}}_n + \mu \mathbf{x}_n (y_n - \widehat{y}_n)^* \end{aligned} \quad (22)$$

and the widely linear quaternion least mean square (WL-QLMS) algorithm [22]

$$\begin{aligned} \widehat{y}_n &= \widehat{\mathbf{w}}_n^a H \mathbf{x}_n^a \\ \widehat{\mathbf{w}}_{n+1}^a &= \widehat{\mathbf{w}}_n^a + \mu \mathbf{x}_n^a (y_n - \widehat{y}_n)^* \end{aligned} \quad (23)$$

where  $\mu$  is a positive step-size.

Let  $\mathcal{P}[\cdot] \in \{\Re[\cdot], \Im[\cdot], \Im_j[\cdot], \Im_\kappa[\cdot]\}$ , then the four components of the FC-QLMS estimate have a unified form given by

$$\begin{aligned} \mathcal{P}[\widehat{y}_n] &= \mathcal{P}[\widehat{\mathbf{w}}_{\eta, n}^H \mathbf{x}_n] = \mathcal{P}[(\widehat{\mathbf{w}}_{\eta, 0}^H + \mu_{\text{fc}} e_{\eta, 0} \eta \mathbf{x}_0^H + \dots + \mu_{\text{fc}} e_{\eta, n-1} \eta \mathbf{x}_{n-1}^H) \mathbf{x}_n] \\ &= \mathcal{P}[\widehat{\mathbf{w}}_{\eta, 0}^H \mathbf{x}_n] + \mu_{\text{fc}} \Re[\mathbf{x}_0^H \mathbf{x}_n] e_{\eta, 0} + \dots + \Re[\mathbf{x}_{n-1}^H \mathbf{x}_n] e_{\eta, n-1} \end{aligned}$$

where  $\eta = 1$  when  $\mathcal{P}[\cdot] = \Re[\cdot]$ ,  $\eta = i$  when  $\mathcal{P}[\cdot] = \Im[\cdot]$ ,  $\eta = j$  when  $\mathcal{P}[\cdot] = \Im_j[\cdot]$ ,  $\eta = \kappa$  when  $\mathcal{P}[\cdot] = \Im_\kappa[\cdot]$ . Similarly, the components of the WL-QLMS estimate are represented as

$$\begin{aligned} \mathcal{P}[\widehat{y}_n] &= \mathcal{P}[\widehat{\mathbf{h}}_n^H \mathbf{x}_n] + \mathcal{P}[\widehat{\mathbf{g}}_n^H \mathbf{x}_n^1] + \mathcal{P}[\widehat{\mathbf{u}}_n^H \mathbf{x}_n^j] + \mathcal{P}[\widehat{\mathbf{v}}_n^H \mathbf{x}_n^\kappa] \\ &= \mathcal{P}[(\widehat{\mathbf{h}}_0^H + \mu e_0 \mathbf{x}_0^H + \dots + \mu e_{n-1} \mathbf{x}_{n-1}^H) \mathbf{x}_n] \\ &\quad + \mathcal{P}[(\widehat{\mathbf{g}}_0^H + \mu e_0 \mathbf{x}_0^H + \dots + \mu e_{n-1} \mathbf{x}_{n-1}^H) \mathbf{x}_n^1] \\ &\quad + \mathcal{P}[(\widehat{\mathbf{u}}_0^H + \mu e_0 \mathbf{x}_0^H + \dots + \mu e_{n-1} \mathbf{x}_{n-1}^H) \mathbf{x}_n^j] \\ &\quad + \mathcal{P}[(\widehat{\mathbf{v}}_0^H + \mu e_0 \mathbf{x}_0^H + \dots + \mu e_{n-1} \mathbf{x}_{n-1}^H) \mathbf{x}_n^\kappa] \\ &= \mathcal{P}[\widehat{\mathbf{h}}_0^H \mathbf{x}_n + \widehat{\mathbf{g}}_0^H \mathbf{x}_n^1 + \widehat{\mathbf{u}}_0^H \mathbf{x}_n^j + \widehat{\mathbf{v}}_0^H \mathbf{x}_n^\kappa] \\ &\quad + 4\mu \Re[\mathbf{x}_0^H \mathbf{x}_n] \mathcal{P}[e_0] + \dots + \Re[\mathbf{x}_{n-1}^H \mathbf{x}_n] \mathcal{P}[e_{n-1}] \end{aligned}$$

**Remark 1.** From the above two formulas, the estimates of FC-QLMS and WL-QLMS are identical if  $\mu_{\text{fc}} = 4\mu$ , and the initial values of estimated weights satisfy  $\widehat{\mathbf{h}}_0 = \widehat{\mathbf{g}}_0 = \widehat{\mathbf{u}}_0 = \widehat{\mathbf{v}}_0$ ,  $\widehat{\mathbf{w}}_{1,0} = 4\Re[\widehat{\mathbf{h}}_0]$ ,  $\widehat{\mathbf{w}}_{j,0} = 4i\Im_j[\widehat{\mathbf{h}}_0]$ ,  $\widehat{\mathbf{w}}_{\kappa,0} = 4\kappa\Im_\kappa[\widehat{\mathbf{h}}_0]$ . This indicates that the WL-QLMS can be replaced with the FC-QLMS without performance degradation.

### 5.2. Transient performance

For the FC-QLMS algorithm, the weight error vector in channel  $\eta$  at

the  $n$ -th iteration is defined as  $\widetilde{\mathbf{w}}_{\eta, n} \triangleq \mathbf{w}_\eta - \widehat{\mathbf{w}}_{\eta, n}$ , so that the estimation error in channel  $\eta$  becomes  $e_{\eta, n} = \mathcal{P}[\widetilde{\mathbf{w}}_{\eta, n}^H \mathbf{x}_n + v_n]$ , and the weight error vector at the next iteration can be expressed as

$$\widetilde{\mathbf{w}}_{\eta, n+1} = \widetilde{\mathbf{w}}_{\eta, n} - \mu_{\text{fc}} e_{\eta, n} \mathbf{x}_n \eta^{-1} \quad (24)$$

which yields

$$\begin{aligned} \widetilde{\mathbf{w}}_{1, n+1} &= \widetilde{\mathbf{w}}_{1, n} - \frac{1}{4} \mu_{\text{fc}} \mathbf{x}_n (\mathbf{x}_n^H \widetilde{\mathbf{w}}_{1, n} + \mathbf{x}_n^H \widetilde{\mathbf{w}}_{1, n}^i + \mathbf{x}_n^H \widetilde{\mathbf{w}}_{1, n}^j + \mathbf{x}_n^H \widetilde{\mathbf{w}}_{1, n}^\kappa + \Re[v_n]) \\ \widetilde{\mathbf{w}}_{j, n+1} &= \widetilde{\mathbf{w}}_{j, n} - \frac{1}{4} \mu_{\text{fc}} \mathbf{x}_n (\mathbf{x}_n^H \widetilde{\mathbf{w}}_{j, n} + \mathbf{x}_n^H \widetilde{\mathbf{w}}_{j, n}^i - \mathbf{x}_n^H \widetilde{\mathbf{w}}_{j, n}^j - \mathbf{x}_n^H \widetilde{\mathbf{w}}_{j, n}^\kappa + \Im_j[v_n]) \\ \widetilde{\mathbf{w}}_{\kappa, n+1} &= \widetilde{\mathbf{w}}_{\kappa, n} - \frac{1}{4} \mu_{\text{fc}} \mathbf{x}_n (\mathbf{x}_n^H \widetilde{\mathbf{w}}_{\kappa, n} - \mathbf{x}_n^H \widetilde{\mathbf{w}}_{\kappa, n}^i - \mathbf{x}_n^H \widetilde{\mathbf{w}}_{\kappa, n}^j + \mathbf{x}_n^H \widetilde{\mathbf{w}}_{\kappa, n}^\kappa + \Im_\kappa[v_n]) \end{aligned}$$

Upon applying the statistical expectation operator, we obtain

$$\begin{aligned} E\{\widetilde{\mathbf{w}}_{1, n+1}\} &= \left(\mathbf{I} - \frac{\mu_{\text{fc}}}{4} \mathbf{C}_{\text{xx}}\right) E\{\widetilde{\mathbf{w}}_{1, n}\} - \frac{\mu_{\text{fc}}}{4} (\mathbf{C}_{\text{xx}} E\{\widetilde{\mathbf{w}}_{1, n}^i\} + \mathbf{C}_{\text{xx}} E\{\widetilde{\mathbf{w}}_{1, n}^j\} \\ &\quad + \mathbf{C}_{\text{xx}} E\{\widetilde{\mathbf{w}}_{1, n}^\kappa\}) \\ E\{\widetilde{\mathbf{w}}_{j, n+1}\} &= \left(\mathbf{I} - \frac{\mu_{\text{fc}}}{4} \mathbf{C}_{\text{xx}}\right) E\{\widetilde{\mathbf{w}}_{j, n}\} - \frac{\mu_{\text{fc}}}{4} (\mathbf{C}_{\text{xx}} E\{\widetilde{\mathbf{w}}_{j, n}^i\} - \mathbf{C}_{\text{xx}} E\{\widetilde{\mathbf{w}}_{j, n}^j\} \\ &\quad - \mathbf{C}_{\text{xx}} E\{\widetilde{\mathbf{w}}_{j, n}^\kappa\}) \\ E\{\widetilde{\mathbf{w}}_{\kappa, n+1}\} &= \left(\mathbf{I} - \frac{\mu_{\text{fc}}}{4} \mathbf{C}_{\text{xx}}\right) E\{\widetilde{\mathbf{w}}_{\kappa, n}\} - \frac{\mu_{\text{fc}}}{4} (-\mathbf{C}_{\text{xx}} E\{\widetilde{\mathbf{w}}_{\kappa, n}^i\} + \mathbf{C}_{\text{xx}} E\{\widetilde{\mathbf{w}}_{\kappa, n}^j\} \\ &\quad + \mathbf{C}_{\text{xx}} E\{\widetilde{\mathbf{w}}_{\kappa, n}^\kappa\}) \end{aligned} \quad (25)$$

the solution to which is dependent on second-order statistics of  $\mathbf{x}$ . It is useful to discuss the following three special cases prior to considering the general case.

#### 5.2.1. $\mathbb{H}$ -proper signal

If  $\mathbf{x}$  is  $\mathbb{H}$ -proper, the three complementary covariances vanish, and the four weight error vectors obey the same recursion form:

$$E\{\widetilde{\mathbf{w}}_{\eta, n+1}\} = \left(\mathbf{I} - \frac{1}{4} \mu_{\text{fc}} \mathbf{C}_{\text{xx}}\right) E\{\widetilde{\mathbf{w}}_{\eta, n}\} \quad (26)$$

The equivalence between the evolution of the four weight vectors is a consequence of the four components of  $\mathbf{x}$  being uncorrelated and with the same variance. The weight error vectors converge if all the eigenvalues of the matrix  $(\mathbf{I} - \frac{1}{4} \mu_{\text{fc}} \mathbf{C}_{\text{xx}})$  are within  $(-1, 1)$ , that is,

$$0 < \mu_{\text{fc}} < \frac{8}{\lambda_{\text{max}}}$$

where  $\lambda_{\text{max}}$  is the maximum eigenvalue of  $\mathbf{C}_{\text{xx}}$ .

According to the expression for the SL-QLMS algorithm in (22), the weight error vector recursion of SL-QLMS is given by

$$E\{\widetilde{\mathbf{h}}_{n+1}\} = (\mathbf{I} - \mu \mathbf{C}_{\text{xx}}) E\{\widetilde{\mathbf{h}}_n\} \quad (27)$$

**Remark 2.** For an  $\mathbb{H}$ -proper quaternion signal  $\mathbf{x}$ , the ratio between the convergence rate of the FC-QLMS and the SL-QLMS is  $\frac{\mu_{\text{fc}}}{4\mu}$ . If  $\mu_{\text{fc}} = 4\mu$ , these two algorithms exhibit identical weight error evolution, which is illustrated in Fig. 2 showing the averaged weight trajectories along the error surfaces while estimating a moving average (MA) process.

#### 5.2.2. $\mathbb{C}^n$ -improper signal

For a  $\mathbb{C}^n$ -improper quaternion signal  $\mathbf{x}$ , we shall use  $\eta'$  and  $\eta''$  to denote any other two imaginary units among  $i, j, \kappa$ . Then the weight error vector recursions in (25) reduce to

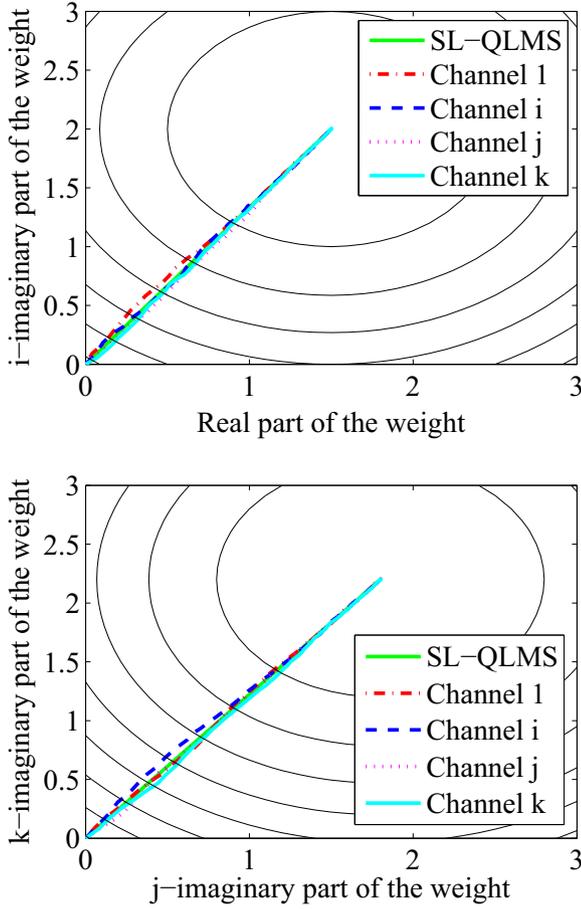


Fig. 2. Averaged weight trajectories for the estimation of a strictly linear MA(1) process driven by  $\mathbf{H}$ -proper white Gaussian noise.

$$\begin{aligned}
 E\{\bar{\mathbf{w}}_{1,n+1}\} &= (\mathbf{I} - \frac{\mu_{fc}}{4}\mathbf{C}_{xx})E\{\bar{\mathbf{w}}_{1,n}\} - \frac{\mu_{fc}}{4}\mathbf{C}_{xx}E\{\bar{\mathbf{w}}_{1,n}^{\eta}\} \\
 E\{\bar{\mathbf{w}}_{\eta,n+1}\} &= (\mathbf{I} - \frac{\mu_{fc}}{4}\mathbf{C}_{xx})E\{\bar{\mathbf{w}}_{\eta,n}\} - \frac{\mu_{fc}}{4}\mathbf{C}_{xx}E\{\bar{\mathbf{w}}_{\eta,n}^{\eta}\} \\
 E\{\bar{\mathbf{w}}_{j,n+1}\} &= (\mathbf{I} - \frac{\mu_{fc}}{4}\mathbf{C}_{xx})E\{\bar{\mathbf{w}}_{j,n}\} + \frac{\mu_{fc}}{4}\mathbf{C}_{xx}E\{\bar{\mathbf{w}}_{j,n}^{\eta}\} \\
 E\{\bar{\mathbf{w}}_{k,n+1}\} &= (\mathbf{I} - \frac{\mu_{fc}}{4}\mathbf{C}_{xx})E\{\bar{\mathbf{w}}_{k,n}\} + \frac{\mu_{fc}}{4}\mathbf{C}_{xx}E\{\bar{\mathbf{w}}_{k,n}^{\eta}\}
 \end{aligned}$$

Observe that the evolutions of  $\bar{\mathbf{w}}_1$  and  $\bar{\mathbf{w}}_{\eta}$  are the same, while  $\bar{\mathbf{w}}_j$  and  $\bar{\mathbf{w}}_k$  also exhibit identical evolution, and the paths of the four weight errors are symmetric about the weight error path of SL-QLMS in (27). This is illustrated in Fig. 3 showing the averaged weight trajectories in the estimation of an MA process. Fig. 3 can also be interpreted by the Cayley–Dickson construction of  $\mathbf{C}^1$ -improper quaternion vectors. A  $\mathbf{C}^1$ -improper  $\mathbf{x}$  can be expressed as  $\mathbf{x} = \mathbf{z}_1 + \mathbf{z}_2j$  where  $\mathbf{z}_1$  and  $\mathbf{z}_2$  are proper complex vectors defined in the plane spanned by  $\{1, j\}$  [26], that is, the real and imaginary parts of each of  $\mathbf{z}_1$  and  $\mathbf{z}_2$  are uncorrelated and with the same variance, so that channel 1 and i or channel j and k follow the same path, similarly to the  $\mathbf{H}$ -proper case in (26). On the other hand, if the two complex vectors  $\mathbf{z}_1$  and  $\mathbf{z}_2$  are related or with different powers, this results in the opposite directions of the two pairs of paths.

### 5.2.3. General signal with uncorrelated components

If the four components of  $\mathbf{x}$  are uncorrelated, the covariance and complementary covariance matrices of  $\mathbf{x}$  can be expressed as

$$\begin{aligned}
 \mathbf{C}_{xx} &= \mathbf{C}_1 + \mathbf{C}_i + \mathbf{C}_j + \mathbf{C}_k \\
 \mathbf{C}_{xx'} &= \mathbf{C}_1 + \mathbf{C}_i - \mathbf{C}_j - \mathbf{C}_k \\
 \mathbf{C}_{xx''} &= \mathbf{C}_1 - \mathbf{C}_i + \mathbf{C}_j - \mathbf{C}_k \\
 \mathbf{C}_{xx'''} &= \mathbf{C}_1 - \mathbf{C}_i - \mathbf{C}_j + \mathbf{C}_k
 \end{aligned}$$

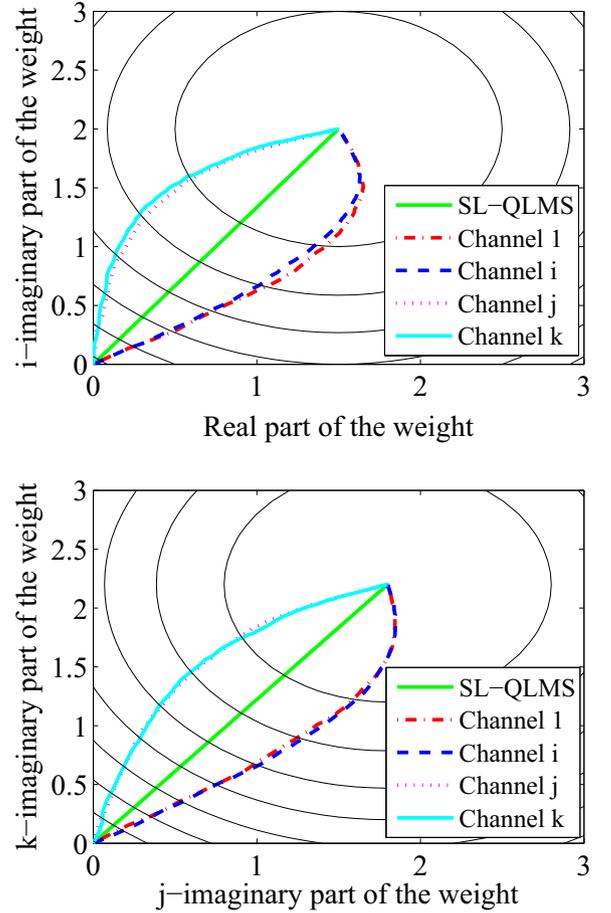


Fig. 3. Averaged weight trajectories for the estimation of a strictly linear MA(1) process driven by  $\mathbf{C}^1$ -improper white Gaussian noise.

where  $\mathbf{C}_1 = E\{\Re[\mathbf{x}]\Re[\mathbf{x}]^T\}$ ,  $\mathbf{C}_i = E\{\Im[\mathbf{x}]\Im[\mathbf{x}]^T\}$ ,  $\mathbf{C}_j = E\{\Im[\mathbf{x}]\Im[\mathbf{x}]^T\}$ , and  $\mathbf{C}_k = E\{\Im[\mathbf{x}]\Im[\mathbf{x}]^T\}$ . The corresponding weight error vectors are given by

$$\begin{aligned}
 E\{\bar{\mathbf{w}}_{1,n+1}\} &= (\mathbf{I} - \mu_{fc}\mathbf{C}_1)E\{\Re[\bar{\mathbf{w}}_{1,n}]\} + \iota(\mathbf{I} - \mu_{fc}\mathbf{C}_i)E\{\Im[\bar{\mathbf{w}}_{1,n}]\} \\
 &\quad + j(\mathbf{I} - \mu_{fc}\mathbf{C}_j)E\{\Im[\bar{\mathbf{w}}_{1,n}]\} + \kappa(\mathbf{I} - \mu_{fc}\mathbf{C}_k)E\{\Im[\bar{\mathbf{w}}_{1,n}]\} \\
 E\{\bar{\mathbf{w}}_{i,n+1}\} &= (\mathbf{I} - \mu_{fc}\mathbf{C}_i)E\{\Re[\bar{\mathbf{w}}_{i,n}]\} + \iota(\mathbf{I} - \mu_{fc}\mathbf{C}_1)E\{\Im[\bar{\mathbf{w}}_{i,n}]\} \\
 &\quad + j(\mathbf{I} - \mu_{fc}\mathbf{C}_k)E\{\Im[\bar{\mathbf{w}}_{i,n}]\} + \kappa(\mathbf{I} - \mu_{fc}\mathbf{C}_j)E\{\Im[\bar{\mathbf{w}}_{i,n}]\} \\
 E\{\bar{\mathbf{w}}_{j,n+1}\} &= (\mathbf{I} - \mu_{fc}\mathbf{C}_j)E\{\Re[\bar{\mathbf{w}}_{j,n}]\} + \iota(\mathbf{I} - \mu_{fc}\mathbf{C}_k)E\{\Im[\bar{\mathbf{w}}_{j,n}]\} \\
 &\quad + j(\mathbf{I} - \mu_{fc}\mathbf{C}_1)E\{\Im[\bar{\mathbf{w}}_{j,n}]\} + \kappa(\mathbf{I} - \mu_{fc}\mathbf{C}_i)E\{\Im[\bar{\mathbf{w}}_{j,n}]\} \\
 E\{\bar{\mathbf{w}}_{k,n+1}\} &= (\mathbf{I} - \mu_{fc}\mathbf{C}_k)E\{\Re[\bar{\mathbf{w}}_{k,n}]\} + \iota(\mathbf{I} - \mu_{fc}\mathbf{C}_j)E\{\Im[\bar{\mathbf{w}}_{k,n}]\} \\
 &\quad + j(\mathbf{I} - \mu_{fc}\mathbf{C}_i)E\{\Im[\bar{\mathbf{w}}_{k,n}]\} + \kappa(\mathbf{I} - \mu_{fc}\mathbf{C}_1)E\{\Im[\bar{\mathbf{w}}_{k,n}]\}
 \end{aligned}$$

and converge if all the eigenvalues of the matrices  $(\mathbf{I} - \mu_{fc}\mathbf{C}_1)$ ,  $(\mathbf{I} - \mu_{fc}\mathbf{C}_i)$ ,  $(\mathbf{I} - \mu_{fc}\mathbf{C}_j)$ , and  $(\mathbf{I} - \mu_{fc}\mathbf{C}_k)$  are within  $(-1, 1)$ , that is,

$$0 < \mu_{fc} < \frac{2}{\sigma_{\max}}$$

where  $\sigma_{\max}$  is the maximum eigenvalue of  $\mathbf{C}_1$ ,  $\mathbf{C}_i$ ,  $\mathbf{C}_j$  and  $\mathbf{C}_k$ .

### 5.2.4. General improper signal

Eq. (25) indicates that for a general improper quaternion signal,  $\mathbf{x}$ , the four weight error vectors obey different recursions which are symmetric about the weight error path of SL-QLMS, as shown in Fig. 4.

### 5.3. Steady-state performance

To quantify the steady-state performance of the FC-QLMS algo-

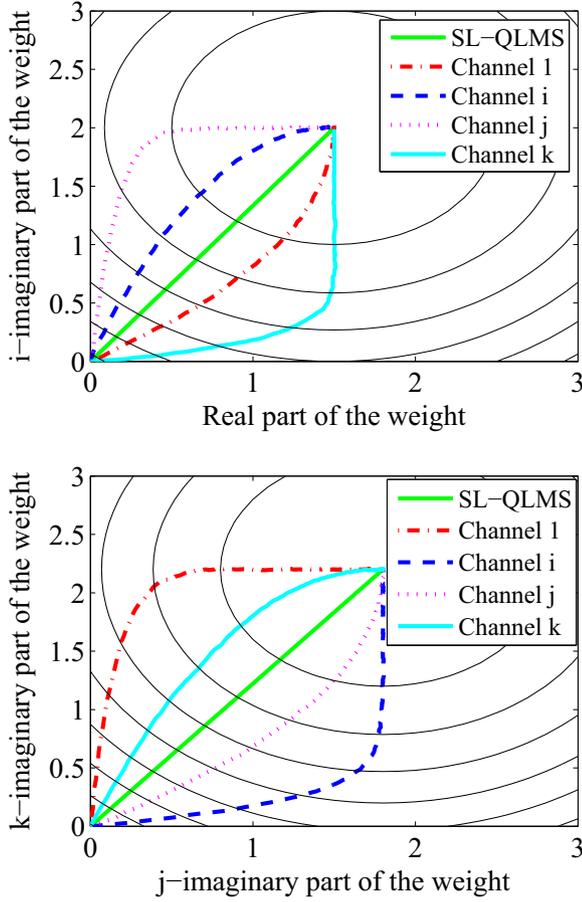


Fig. 4. Averaged weight trajectories for the estimation of a strictly linear MA(1) process driven by general improper white Gaussian noise.

riethm, we next analyse the MSE in the estimation of a desired signal in (6). The estimation error of FC-QLMS is given by

$$e_n = e_{1,n} + 1e_{1,n} + J e_{j,n} + \kappa e_{\kappa,n} = e_n^a + v_n$$

where  $e_n^a$  is the *a priori* error defined as

$$e_n^a = \Re\{\bar{\mathbf{w}}_{1,n}^H \mathbf{x}_n\} + 1\Im\{\bar{\mathbf{w}}_{j,n}^H \mathbf{x}_n\} + J\Im\{\bar{\mathbf{w}}_{j,n}^H \mathbf{x}_n\} + \kappa\Im\{\bar{\mathbf{w}}_{\kappa,n}^H \mathbf{x}_n\}$$

The steady-state MSE is therefore

$$\text{MSE} = \lim_{n \rightarrow \infty} E\{|e_n|^2\} = \lim_{n \rightarrow \infty} E\{|e_n^a|^2\} + \sigma_v^2$$

where  $\lim_{n \rightarrow \infty} E\{|e_n^a|^2\}$  is the excess MSE (EMSE) resulting from a mismatch between the estimated value and true value of the system weight vector. The MSEs in the four channels shall be analysed separately. The *a priori* and *a posteriori* errors in channel  $\eta$  are defined as  $e_{\eta,n}^a = \Psi\{\bar{\mathbf{w}}_{\eta,n}^H \mathbf{x}_n\}$  and  $e_{\eta,n}^p = \Psi\{\bar{\mathbf{w}}_{\eta,n+1}^H \mathbf{x}_n\}$ , which are related as

$$e_{\eta,n}^p = e_{\eta,n}^a - \mu_{fc} \|\mathbf{x}_n\|^2 e_{\eta,n} \quad (28)$$

Combining (24) and (28) yields

$$\bar{\mathbf{w}}_{\eta,n+1} + \frac{e_{\eta,n}^a}{\|\mathbf{x}_n\|^2} \mathbf{x}_n \eta^{-1} = \bar{\mathbf{w}}_{\eta,n} + \frac{e_{\eta,n}^p}{\|\mathbf{x}_n\|^2} \mathbf{x}_n \eta^{-1} \quad (29)$$

Upon evaluating the energies on both sides of (29) and applying the statistical expectation operator, we arrive at [27]

$$E\{\bar{\mathbf{w}}_{\eta,n+1}^2\} + E\left\{\frac{(e_{\eta,n}^a)^2}{\|\mathbf{x}_n\|^2}\right\} = E\{\bar{\mathbf{w}}_{\eta,n}^2\} + E\left\{\frac{(e_{\eta,n}^p)^2}{\|\mathbf{x}_n\|^2}\right\}$$

Assuming  $E\{\bar{\mathbf{w}}_{\eta,n+1}^2\} = E\{\bar{\mathbf{w}}_{\eta,n}^2\}$  at the steady state (as  $n \rightarrow \infty$ ), we obtain the steady-state condition

$$\begin{aligned} E\left\{\frac{(e_{\eta,n}^a)^2}{\|\mathbf{x}_n\|^2}\right\} &= E\left\{\frac{(e_{\eta,n}^p)^2}{\|\mathbf{x}_n\|^2}\right\} = E\left\{\frac{(e_{\eta,n}^a - \mu_{fc} e_{\eta,n} \|\mathbf{x}_n\|^2)^2}{\|\mathbf{x}_n\|^2}\right\} \\ &= E\left\{\frac{(e_{\eta,n}^a - \mu_{fc} e_{\eta,n}^a \|\mathbf{x}_n\|^2 - \mu_{fc} \Psi\{v_n\} \|\mathbf{x}_n\|^2)^2}{\|\mathbf{x}_n\|^2}\right\} \end{aligned} \quad (30)$$

As the noise  $v$  is independent of the input signal  $\mathbf{x}$ , expression (30) simplifies into

$$2E\{(e_{\eta,n}^a)^2\} = \mu_{fc} E\{(e_{\eta,n}^a)^2 \|\mathbf{x}_n\|^2\} + \mu_{fc} \text{Tr}\{\mathbf{C}_{\mathbf{xx}}\} E\{\Psi\{v_n\}^2\}$$

Considering  $|e_n^a|^2 = \sum_{\eta=1,i,j,\kappa} (e_{\eta,n}^a)^2$ , we next obtain

$$2E\{|e_n^a|^2\} = \mu_{fc} E\{|e_n^a|^2 \|\mathbf{x}_n\|^2\} + \mu_{fc} \text{Tr}\{\mathbf{C}_{\mathbf{xx}}\} \sigma_v^2$$

For a small step-size  $\mu_{fc}$ , the term  $\mu_{fc} E\{|e_n^a|^2 \|\mathbf{x}_n\|^2\}$  is negligible compared to  $\mu_{fc} \text{Tr}\{\mathbf{C}_{\mathbf{xx}}\} \sigma_v^2$ , and thus the EMSE and MSE are given by

$$\text{EMSE}_{fc} = \lim_{n \rightarrow \infty} E\{|e_n^a|^2\} = \frac{\mu_{fc}}{2} \text{Tr}\{\mathbf{C}_{\mathbf{xx}}\} \sigma_v^2$$

$$\text{MSE}_{fc} = \text{EMSE}_{fc} + \sigma_v^2 = \left(\frac{\mu_{fc}}{2} \text{Tr}\{\mathbf{C}_{\mathbf{xx}}\} + 1\right) \sigma_v^2$$

However, for a large step-size  $\mu_{fc}$ , the term  $\mu_{fc} E\{|e_n^a|^2 \|\mathbf{x}_n\|^2\}$  is not negligible. Assuming that  $\|\mathbf{x}_n\|^2$  is independent of  $|e_n^a|^2$ , the EMSE and MSE in this case are given by

$$\text{EMSE}_{fc} = \lim_{n \rightarrow \infty} E\{|e_n^a|^2\} = \frac{\mu_{fc}}{2 - \mu_{fc} \text{Tr}\{\mathbf{C}_{\mathbf{xx}}\}} \text{Tr}\{\mathbf{C}_{\mathbf{xx}}\} \sigma_v^2$$

$$\text{MSE}_{fc} = \text{EMSE}_{fc} + \sigma_v^2 = \frac{2}{2 - \mu_{fc} \text{Tr}\{\mathbf{C}_{\mathbf{xx}}\}} \sigma_v^2$$

As shown in the Appendix, the steady-state MSE of the WL-QLMS is given by

$$\text{MSE}_{wl} = (2\mu \text{Tr}\{\mathbf{C}_{\mathbf{xx}}\} + 1) \sigma_v^2 \text{ (for a small } \mu)$$

$$\text{MSE}_{wl} = \frac{\sigma_v^2}{1 - 2\mu \text{Tr}\{\mathbf{C}_{\mathbf{xx}}\}} \text{ (for a large } \mu)$$

**Remark 3.** The FC-QLMS and WL-QLMS algorithms have the same steady-state MSE if  $\mu_{fc} = 4\mu$ .

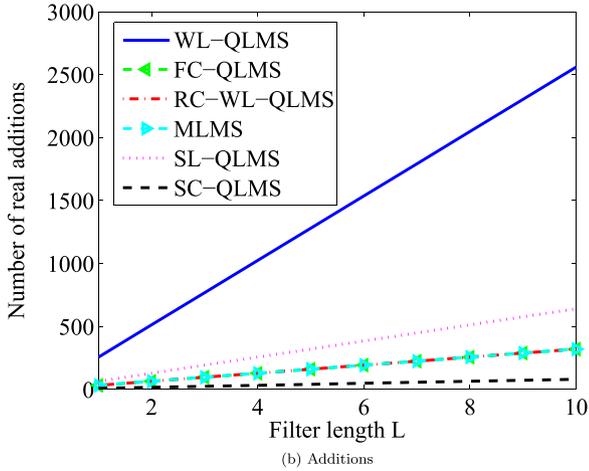
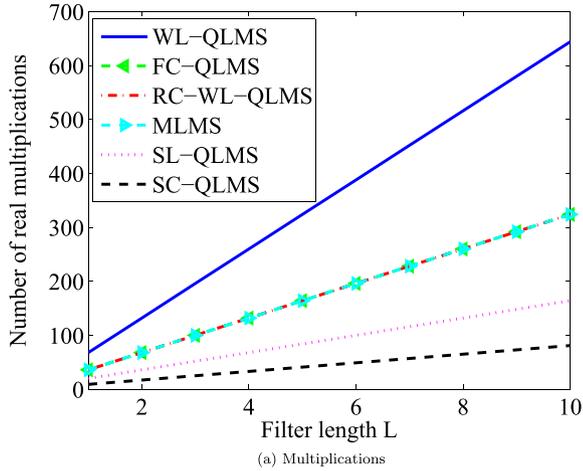
#### 5.4. Computational complexity

Compared with existing widely linear and strictly quaternion estimation approaches, the proposed four-channel estimation technique requires less computation cost. This is because the proposed estimation algorithm avoids the quaternion-quaternion addition in (9) and (10) while the proposed weight update rule in (14) replaces quaternion-quaternion multiplications, which cost 8 real multiplications and 28 real additions per multiplication [28], with real-quaternion multiplications costing only 4 real multiplications per multiplication. Table 1 and Fig. 5 compare the FC-QLMS and SC-QLMS algorithms with the existing QLMS algorithms in terms of the number of real-valued operations required per iteration. Observe that the FC-QLMS and SC-QLMS require only about half the number of multiplications and an eighth of the number of additions, as compared to the WL-QLMS and the SL-QLMS. Therefore, the four-channel estimation framework allows for an efficient formulation of WL-QLMS and SL-QLMS in the form of FC-QLMS and SC-QLMS. Although the RC-WL-QLMS [13] and the MLMS [14] also have such low computational complexity, they result in a loss of physical meaning inherent in the quaternion domain.

**Table 1**

The number of real-valued operations per iteration for the QLMS algorithms, for an adaptive filter of length  $L$ . The number of real-valued operations for quaternion-quaternion multiplications is based on [28].

Algorithm	Real multiplications	Real additions
Estimation of widely linear systems		
Proposed: FC-QLMS	$32L + 4$	$32L$
WL-QLMS [22]	$64L + 4$	$256L$
RC-WL-QLMS [13]	$32L + 4$	$32L$
MLMS [14]	$32L + 4$	$32L$
Estimation of strictly linear systems		
Proposed: SC-QLMS	$8L + 1$	$8L$
SL-QLMS [23]	$16L + 4$	$64L$



**Fig. 5.** The number of real-valued operations per iteration for the QLMS algorithms, for an adaptive filter of length  $L$ . The number of real-valued operations for quaternion-quaternion multiplications is based on [28].

### 5.5. Adaptive step-size

Standard adaptive algorithms might experience degraded convergence when processing non-stationary signals with large dynamical ranges. This degradation can be circumvented by algorithms that adaptively optimise the step-sizes. For the FC-QLMS algorithm, the four step-sizes in the four channels can be optimised separately as  $\mu_1$ ,  $\mu_j$  and  $\mu_k$ . We shall use the Barzilai–Borwein method [29] to obtain the optimal step-sizes at the  $n$ -th iteration, as

$$\mu_{\eta,n} = \arg \min_{\mu \in \mathbb{R}^+} \|s_{\eta,n} - \mu \Gamma_{\eta,n}\|_2^2 \quad \eta \in \{1, j, k\}$$

where

$$s_{\eta,n} = \widehat{\mathbf{w}}_{\eta,n} - \widehat{\mathbf{w}}_{\eta,n-1}$$

$$\Gamma_{\eta,n} = \frac{\partial \Psi[y_n - \widehat{\mathbf{w}}_{\eta,n}^H \mathbf{x}_n]}{\partial \widehat{\mathbf{w}}_{\eta,n}^*} - \frac{\partial \Psi[y_n - \widehat{\mathbf{w}}_{\eta,n-1}^H \mathbf{x}_n]}{\partial \widehat{\mathbf{w}}_{\eta,n-1}^*} = \Psi'[(\widehat{\mathbf{w}}_{\eta,n} - \widehat{\mathbf{w}}_{\eta,n-1})^H \mathbf{x}_n] \mathbf{x}_n^{-1}$$

By setting  $\nabla_{\mu_{\eta,n}} \|s_{\eta,n} - \mu_{\eta,n} \Gamma_{\eta,n}\|_2^2 = \mathbf{0}$ , we obtain the optimal step-size

$$\mu_{\eta,n} = \frac{\Re[s_{\eta,n}^H \Gamma_{\eta,n}]}{\Gamma_{\eta,n}^H \Gamma_{\eta,n}} = \|\mathbf{x}_n\|_2^{-2} \quad (31)$$

which is equivalent to a unit step-size after data normalisation [30].

## 6. Simulations

All estimation algorithms discussed were evaluated by simulations over synthetic and real-world signals. The performance index in the form of normalised MSE (NMSE) was calculated at each iteration through averaging the error power normalised by the signal power from 100 independent trials, to yield

$$\text{NMSE}_n = \frac{1}{100} \sum_{l=1}^{100} \frac{|y_n^{(l)} - \widehat{y}_n^{(l)}|^2}{|y_n^{(l)}|^2}$$

where  $y_n^{(l)}$  is the desired signal and  $\widehat{y}_n^{(l)}$  the estimate at the  $l$ -th trial. For a fair comparison, according to the theoretical analysis in Section 5, the values of step-sizes of the four-channel and single-channel algorithms proposed in the paper were set to four times those of their widely linear and strictly linear counterparts, and the weight vectors in the algorithms were initialised to zero vectors.

### 6.1. QLMS

The QLMS algorithms were implemented to identify a strictly linear MA system, given by

$$y_n = b_0 x_n + b_1 x_{n-1} + b_2 x_{n-2} + b_3 x_{n-3} + v_n \quad (32)$$

and the widely linear MA system

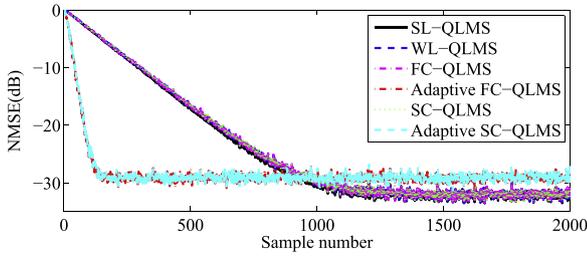
$$y_n = b_0 x_n + b_1 x_{n-1}^i + b_2 x_{n-2}^j + b_3 x_{n-3}^k + v_n \quad (33)$$

where  $b_0, b_1, b_2, b_3 \in \mathbb{H}$ , the signal  $x$  and noise  $v$  were  $\mathbb{H}$ -proper quaternion-valued white Gaussian data, and the signal-to-noise ratio (SNR) was 10 dB. Fig. 6a illustrates that when estimating the strictly linear MA system in (32), FC-QLMS, WL-QLMS, SC-QLMS, and SL-QLMS behaved identically, while the FC-QLMS and SC-QLMS with an adaptive step-size (31), which are referred to as the adaptive FC-QLMS and adaptive SC-QLMS, exhibited faster convergence. Fig. 6b shows that when estimating the widely linear MA system in (33), the FC-QLMS and the WL-QLMS behaved identically, and had a much smaller steady-state NMSE than the SL-QLMS, while the adaptive FC-QLMS converged fastest. Table 2 compares the steady-state EMSE and MSE of the FC-QLMS and WL-QLMS when estimating the strictly linear MA process in (32). The match between the theoretical and experimental values, and the equivalence between the results of the two algorithms, support the quantitative analysis in Section 5.3.

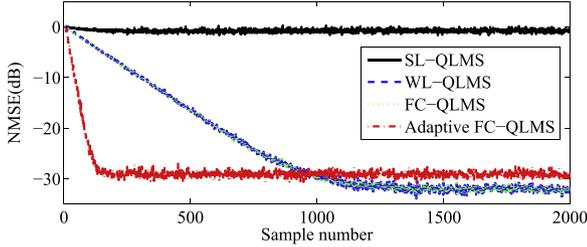
These QLMS algorithms were also used to forecast quaternion-valued wind data, the four components of which contain the wind speed in the north-south, east-west, and vertical directions, and the temperature. As shown in Fig. 7, the FC-QLMS and the WL-QLMS behaved identically, while the adaptive FC-QLMS had the fastest convergence.

### 6.2. QRLS

Figs. 8 and 9 illustrate the performance of QRLS algorithms in the identification of the MA systems in (32) and (33), and in the forecasting of quaternion-valued wind data. The FC-QRLS behaved identically to



(a) Strictly linear MA system



(b) Widely linear MA system

Fig. 6. NMSE curves of the QLMS algorithms for the estimation of MA systems.

Table 2

The MSE and EMSE of the FC-QLMS and WL-QLMS for the identification of a strictly linear MA system.

Index	$\sigma_v^2$	MSE <sub>fc</sub>	EMSE <sub>fc</sub>	MSE <sub>wl</sub>	EMSE <sub>wl</sub>
Simulations	0.1	0.140	0.040	0.140	0.040
Theory	0.1	0.132	0.032	0.132	0.032

the WL-QRLS for both synthetic and wind data, while when identifying the strictly linear MA system, the SC-QRLS achieved equivalent performance.

### 6.3. QNGD

The QNGD algorithms were employed in a one-step ahead prediction of the 4-D Saito's chaotic circuit data, which is governed by four state variables  $x_1, y_1, x_2, y_2$  and five parameters, as [31]

$$\begin{bmatrix} \frac{\partial x_1}{\partial \tau} \\ \frac{\partial y_1}{\partial \tau} \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ -\alpha_1 & -\alpha_1\beta_1 \end{bmatrix} \begin{bmatrix} x_1 - \gamma \frac{\beta_1}{1-\beta_1} h(z) \\ y_1 - \gamma \frac{1}{1-\beta_1} h(z) \end{bmatrix}$$

$$\begin{bmatrix} \frac{\partial x_2}{\partial \tau} \\ \frac{\partial y_2}{\partial \tau} \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ -\alpha_2 & -\alpha_2\beta_2 \end{bmatrix} \begin{bmatrix} x_2 - \gamma \frac{\beta_2}{1-\beta_2} h(z) \\ y_2 - \gamma \frac{1}{1-\beta_2} h(z) \end{bmatrix}$$

$$h(z) = \begin{cases} 1, & z \geq -1 \\ -1, & z \leq 1 \end{cases}$$

$$z = x_1 + x_2$$

where  $\tau$  is the time constant of the chaotic circuit. The parameter values were chosen as  $\gamma = 1.3, \alpha_1 = 7.5, \alpha_2 = 15, \beta_1 = 0.16, \beta_2 = 0.097$ . Fig. 10 shows that the FC-QNGD and WL-QNGD behaved identically and outperformed the SL-QNGD.

## 7. Conclusion

We have introduced a four-channel linear model for quaternion

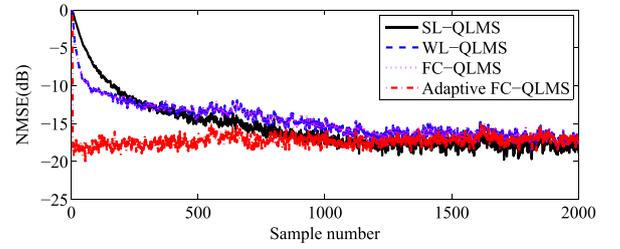
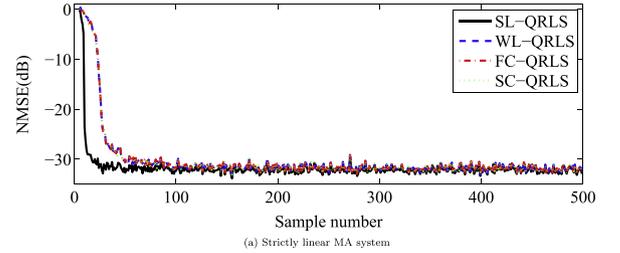
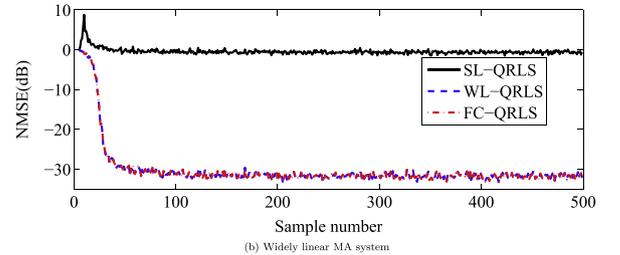


Fig. 7. NMSE curves of the QLMS algorithms for the wind forecasting.



(a) Strictly linear MA system



(b) Widely linear MA system

Fig. 8. NMSE curves of the QRLS algorithms for the estimation of MA systems.

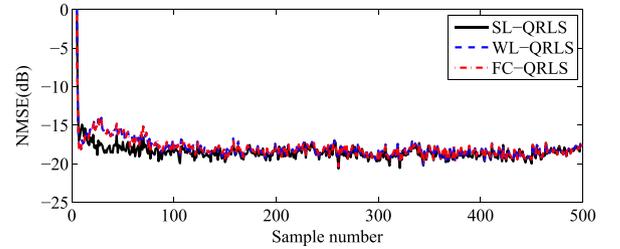
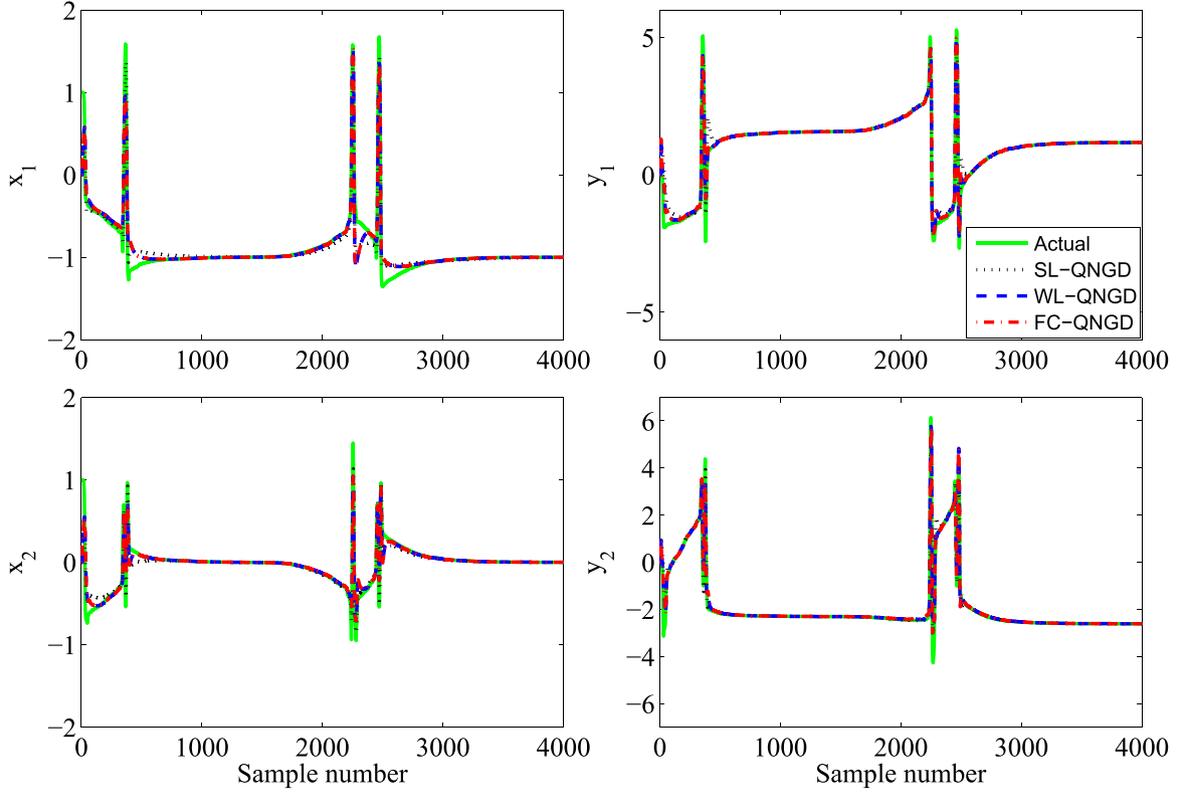
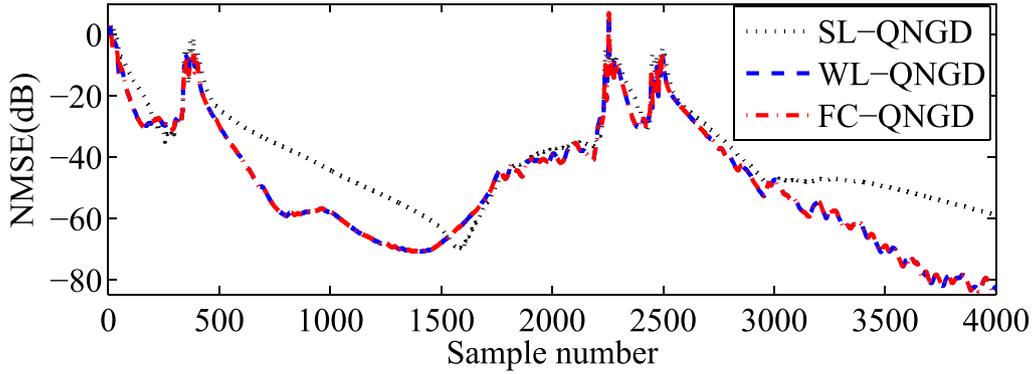


Fig. 9. NMSE curves of the QRLS algorithms for the wind forecasting.

signal estimation and have proposed the corresponding adaptive MMSE estimation algorithms, referred to as the FC-QLMS, FC-QRLS and FC-QNGD. These have been shown to be cost-effective alternatives to existing widely linear algorithms, while exhibiting identical performances. The proposed individual estimation of real and imaginary quaternion components has maintained the four degrees of freedom necessary for widely linear estimation, together with providing physical insight into quaternion estimation and enhanced flexibility through the estimation in each channel being controlled independently. In this way, the computational complexity has been dramatically reduced, while maintaining the performance of the original algorithms. Simulation studies on synthetic and real-world signals support the analysis. The proposed four-channel linear model is suitable for a wide range of applications in widely linear processing of quaternion signals, such as for widely linear series expansions [32].



(a) Estimate curves



(b) NMSE curves

Fig. 10. Performance of the QNGD algorithms for the one-step ahead prediction of Saito's chaotic circuit data.

### Appendix A. Steady-state performance of WL-QLMS

For the desired signal model (6), the system weight vector can be expressed in the augmented form as  $\mathbf{w}^a = [\mathbf{h}^T, \mathbf{g}^T, \mathbf{u}^T, \mathbf{v}^T]^T$ , and represents the optimal Wiener solution. The estimation error is then given by

$$e_n = y_n - \mathbf{w}_n^{aH} \mathbf{x}_n^a = e_n^a + v_n$$

where  $e_n^a$  is the *a priori* error defined as  $e_n^a = \tilde{\mathbf{w}}_n^{aH} \mathbf{x}_n^a$  where  $\tilde{\mathbf{w}}_n^a = \mathbf{w}^a - \hat{\mathbf{w}}_n^a$  is the weight error vector. The steady-state MSE is then given by

$$\text{MSE}_{\text{wl}} = \lim_{n \rightarrow \infty} E[|e_n|^2] = \lim_{n \rightarrow \infty} E[|e_n^a|^2] + \sigma_v^2$$

where  $\lim_{n \rightarrow \infty} E[|e_n^a|^2]$  is the EMSE. The weight error vectors at the  $n$ -th and the  $(n + 1)$ -th iteration are related by

$$\tilde{\mathbf{w}}_{n+1}^a = \tilde{\mathbf{w}}_n^a - \mu \mathbf{x}_n^a e_n^* \tag{A.1}$$

from which we obtain

$$e_n^p = e_n^a - 4\mu e_n \|\mathbf{x}_n\|^2 \tag{A.2}$$

where  $e_n^p = \tilde{\mathbf{w}}_{n+1}^{aH} \mathbf{x}_n$  is the *a posteriori* error. Combining (A.1) and (A.2) yields

$$\bar{\mathbf{w}}_{n+1}^a + \mathbf{x}_n \frac{e_n^{a*}}{4 \|\mathbf{x}_n\|^2} = \bar{\mathbf{w}}_n^a + \mathbf{x}_n \frac{e_n^{p*}}{4 \|\mathbf{x}_n\|^2} \quad (\text{A.3})$$

Evaluating the energies on both sides of (A.3) and applying the statistical expectation operator yields [27]

$$E\left\{\|\bar{\mathbf{w}}_{n+1}^a\|^2\right\} + E\left\{\frac{|e_n^{a}|^2}{16 \|\mathbf{x}_n\|^2}\right\} = E\left\{\|\bar{\mathbf{w}}_n^a\|^2\right\} + E\left\{\frac{|e_n^{p}|^2}{16 \|\mathbf{x}_n\|^2}\right\}$$

Assuming  $E\{\|\bar{\mathbf{w}}_{n+1}^a\|^2\} = E\{\|\bar{\mathbf{w}}_n^a\|^2\}$  at the steady state (as  $n \rightarrow \infty$ ), we obtain the steady-state condition as

$$E\left\{\frac{|e_n^{a}|^2}{16 \|\mathbf{x}_n\|^2}\right\} = E\left\{\frac{|e_n^{p}|^2}{16 \|\mathbf{x}_n\|^2}\right\} = E\left\{\frac{|e_n^a - 4\mu e_n \|\mathbf{x}_n\|^2|^2}{16 \|\mathbf{x}_n\|^2}\right\} = E\left\{\frac{|e_n^a - 4\mu(e_n^a + v_n)\|\mathbf{x}_n\|^2|^2}{16 \|\mathbf{x}_n\|^2}\right\} \quad (\text{A.4})$$

As the noise  $v$  is independent of  $\mathbf{x}$ , Eq. (A.4) simplifies into

$$E[|e_n^{a}|^2] = 2\mu E\{\|\mathbf{x}_n\|^2 |e_n^{a}|^2\} + 2\mu \text{Tr}[\mathbf{C}_{\mathbf{xx}}] \sigma_v^2$$

For a small step-size  $\mu$ , the term  $2\mu E\{\|\mathbf{x}_n\|^2 |e_n^{a}|^2\}$  is negligible compared to  $2\mu \text{Tr}[\mathbf{C}_{\mathbf{xx}}] \sigma_v^2$ , so that the EMSE is given by

$$\text{EMSE}_{\text{wl}} = \lim_{n \rightarrow \infty} E\{|e_n^{a}|^2\} = 2\mu \text{Tr}[\mathbf{C}_{\mathbf{xx}}] \sigma_v^2$$

and the steady-state MSE becomes

$$\text{MSE}_{\text{wl}} = \text{EMSE}_{\text{wl}} + \sigma_v^2 = (2\mu \text{Tr}[\mathbf{C}_{\mathbf{xx}}] + 1) \sigma_v^2$$

However, for a large step-size  $\mu$ , the term  $2\mu E\{\|\mathbf{x}_n\|^2 |e_n^{a}|^2\}$  is not negligible. Assuming that  $\|\mathbf{x}_n\|^2$  is independent of  $|e_n^{a}|^2$ , we arrive at

$$\text{EMSE}_{\text{wl}} = \lim_{n \rightarrow \infty} E\{|e_n^{a}|^2\} = \frac{2\mu \text{Tr}[\mathbf{C}_{\mathbf{xx}}] \sigma_v^2}{1 - 2\mu \text{Tr}[\mathbf{C}_{\mathbf{xx}}]}$$

$$\text{MSE}_{\text{wl}} = \text{EMSE}_{\text{wl}} + \sigma_v^2 = \frac{\sigma_v^2}{1 - 2\mu \text{Tr}[\mathbf{C}_{\mathbf{xx}}]}$$

## 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.
- [4] J.B. Kuipers, *Quaternions and Rotation Sequences: A Primer with Applications to Orbits, Aerospace and Virtual Reality*, Princeton University Press, Princeton, New Jersey, 1999.
- [5] J. Tao, W. Chang, Adaptive beamforming based on complex quaternion processes, *Math. Probl. Eng.* (2014), <http://dx.doi.org/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 Process.* 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.
- [8] Y. Xia, C. Jahanchahi, D.P. Mandic, Quaternion-valued echo state networks, *IEEE Trans. Neural Netw. Learn. Syst.* 26 (4) (2015) 663–673.
- [9] J. Via, D. Palomar, L. Vielva, I. Santamaría, Quaternion ICA from second-order statistics, *IEEE Trans. Signal Process.* 59 (4) (2011) 1586–1600.
- [10] T.A. Ell, N. Le Bihan, S.J. Sangwine, *Quaternion Fourier Transforms for Signal and Image Processing*, ISTE-Wiley, UK & USA, 2014.
- [11] C. Cheong Took, D.P. Mandic, A quaternion widely linear adaptive filter, *IEEE Trans. Signal Process.* 58 (8) (2010) 4427–4431.
- [12] J. Navarro-Moreno, R. Mari, J.C. Ruiz-Molina, A quaternion widely linear model for nonlinear gaussian estimation, *IEEE Trans. Signal Process.* 62 (24) (2014) 6414–6424.
- [13] F.G.A. Neto, V.H. Nascimento, A novel reduced-complexity widely linear QLMS algorithm, in: *Proceedings of IEEE Statistical Signal Process. Workshop (SSP)*, 2011, pp. 81–84.
- [14] C. Cheong Took, D.P. Mandic, J. Benesty, Study of the quaternion LMS and four-channel LMS algorithms, in: *Proceedings of IEEE International Conference on Acoustics, Speech, Signal Process. (ICASSP)*, 2009, pp. 3109–3112.
- [15] C. Jahanchahi, S. Kanna, D.P. Mandic, Complex dual channel estimation: cost effective widely linear adaptive filtering, *Signal Process.* 104 (2014) 33–42.
- [16] L. Rodman, *Topics in Quaternion Linear Algebra*, Princeton University Press, Princeton, New Jersey, 2014.
- [17] T.A. Ell, S.J. Sangwine, Quaternion involutions and anti-involutions, *Comput. Math. Appl.* 53 (1) (2007) 137–143.
- [18] J. Via, D. Ramirez, I. Santamaría, Properness and widely linear processing of quaternion random vectors, *IEEE Trans. Inf. Theory* 56 (7) (2010) 3502–3515.
- [19] C. Cheong Took, D.P. Mandic, Augmented second-order statistics of quaternion random signals, *Signal Process.* 91 (2) (2011) 214–224.
- [20] 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.
- [21] D.P. Mandic, C. Jahanchahi, C. Cheong Took, A quaternion gradient operator and its applications, *IEEE Signal Process. Lett.* 18 (1) (2011) 47–50.
- [22] D. Xu, C. Jahanchahi, C. Cheong Took, D.P. Mandic, Enabling quaternion derivatives: the generalized HR calculus, *R. Soc. Open Sci.* 2 (8) (2015) 150255.
- [23] D. Xu, Y. Xia, D.P. Mandic, Optimization in quaternion dynamic systems: Gradient, Hessian, and learning algorithms, *IEEE Trans. Neural Netw. Learn. Syst.* 27 (2) (2015) 249–261.
- [24] C. Jahanchahi, C. Cheong Took, D.P. Mandic, The widely linear quaternion recursive least squares filter, in: *Proceedings of IEEE International Cognitive Information Processing Workshop*, 2010, pp. 87–92.
- [25] B. Che Ujang, C. Cheong Took, D.P. Mandic, Quaternion-valued nonlinear adaptive filtering, *IEEE Trans. Neural Netw.* 22 (8) (2011) 1193–1206.
- [26] P.O. Amblard, N. Le Bihan, On properness of quaternion valued random variables, in: *Proceedings of IMA Conference on Mathematics in Signal Processing*, 2004, pp. 23–26.
- [27] N.R. Yousef, A.H. Sayed, A unified approach to the steady-state and tracking analyses of adaptive filters, *IEEE Trans. Signal Process.* 49 (2) (2001) 314–324.
- [28] A. Cariow, G. Cariowa, An unified approach for developing rationalized algorithms for hypercomplex number multiplication, *Electr. Rev.* 91 (2) (2015) 36–39.
- [29] J. Barzilai, J.M. Borwein, Two-point step size gradient methods, *IMA J. Numer. Anal.* 8 (1) (1988) 141–148.
- [30] B. Widrow, S.D. Stearns, *Adaptive Signal Processing*, Prentice-Hall Inc, Englewood, NJ, 1985.
- [31] K. Mitsubori, T. Saito, Torus doubling and hyperchaos in a five dimensional hysteresis circuit, in: *Proceedings of IEEE International Symposium on Circuits and Systems*, vol. 6, 1994, pp. 113–116.
- [32] J. Navarro-Moreno, R. Fernandez-Alcala, J. Ruiz-Molina, A quaternion widely linear series expansion and its applications, *IEEE Signal Process. Lett.* 19 (12) (2012) 868–871.