

# Adaptive Convex Combination Approach for the Identification of Improper Quaternion Processes

Bukhari Che Ujang, *Student Member, IEEE*, Cyrus Jahanchahi, *Student Member, IEEE*,  
Clive Cheong Took, *Member, IEEE*, and Danilo P. Mandic, *Senior Member, IEEE*

**Abstract**—Data-adaptive optimal modeling and identification of real-world vector sensor data is provided by combining the fractional tap-length (FT) approach with model order selection in the quaternion domain. To account rigorously for the generality of such processes, both second-order circular (proper) and noncircular (improper), the proposed approach in this paper combines the FT length optimization with both the strictly linear quaternion least mean square (QLMS) and widely linear QLMS (WL-QLMS). A collaborative approach based on QLMS and WL-QLMS is shown to both identify the type of processes (proper or improper) and to track their optimal parameters in real time. Analysis shows that monitoring the evolution of the convex mixing parameter within the collaborative approach allows us to track the improperness in real time. Further insight into the properties of those algorithms is provided by establishing a relationship between the steady-state error and optimal model order. The approach is supported by simulations on model order selection and identification of both strictly linear and widely linear quaternion-valued systems, such as those routinely used in renewable energy (wind) and human-centered computing (biomechanics).

**Index Terms**—Augmented quaternion statistics, fractional tap length, model order selection, noncircularity detection, nonstationarity, quaternion noncircularity, widely linear modeling, widely linear quaternion least mean square (WL-QLMS).

## I. INTRODUCTION

THE RECENT interest in quaternion-valued statistical signal processing stems from the enhanced accuracy, physical insight, mathematical rigor, and the convenience of representation that it provides in the modeling of three- and four-dimensional real-world data. Indeed, many 3-D phenomena (inertial body motor sensors, wind field) in our 3-D world are best represented as quaternion-valued, yet the algorithms for their identification are still lacking. These data sources are almost invariably nonstationary and with a time-varying model of the signal generating system. The progress in quaternion learning systems has been enabled by the developments in the

theory of quaternion gradient [1], nonlinear analytic quaternion function analysis [2], augmented quaternion statistics [3], [4], and the advances in quaternion independent component analysis [5], [6]. Model order selection for quaternion-valued systems underpins the identification of signals and systems but is still an open problem made even more complex by the need to differentiate between circular (those with rotation-invariant distributions) and various forms of noncircular signals (those with rotation-dependent distributions).

Model order identification is a routine procedure for stationary circular signals, whereas for systems with time-varying parameters it has been shown that a convenient and rigorous way to identify the model order of any univariate and multivariate system is by using a combination of adaptive filters and variable tap-length algorithms, making it possible to both operate in nonstationary environments and to optimize for the optimal model order online [7], [8]. The variable tap-length algorithm considered in this paper is of the fractional tap-length (FT) type because of its simplicity and robustness [8]. Notice that the FT algorithm was designed specifically for real-valued learning systems and was only recently extended to the widely linear complex-valued case in order to cater for both circular and noncircular data and linear and widely linear systems [9].

Real vectors in  $\mathbb{R}^3$  are not a division algebra and therefore have a number of mathematical shortcomings (gimbal lock for rotation) when modeling real-world data. Quaternions form a division algebra and are thus ideal for 3-D processes with high dynamics, such as 3-D wind modeling in renewable energy or 3-D body motion in human-centered computing. To deal with evolving environments, we here introduce a class of quaternion-valued FT algorithms; for rigor, this is achieved by considering the full second-order (augmented) quaternion statistics of the signal. The quaternion-valued algorithms considered are the recently introduced quaternion least mean square (QLMS) [10] and widely linear QLMS (WL-QLMS) [11], which cater, respectively, for strictly linear and widely linear quaternion-valued processes. We also consider a collaborative architecture comprising QLMS and WL-QLMS, which equips us with the ability to identify both the model order and the second-order circular (proper) or noncircular (improper) nature of general quaternion-valued systems and processes. The WL-QLMS is based on the widely linear model which captures the full second-order statistics of the quaternion signal, characterized by the standard covariance matrix  $C_q$  and three complementary covariance matrices termed the  $\iota$ -covariance  $C_{q\iota}$ ,  $j$ -covariance  $C_{qj}$ , and  $\kappa$ -covariance  $C_{q\kappa}$  [3], [4]

Manuscript received April 18, 2012; revised January 20, 2013; accepted February 10, 2013. Date of publication March 12, 2013; date of current version December 13, 2013.

B. Che Ujang is with the Department of Computer and Communication Systems Engineering, Universiti Putra Malaysia, UPM Serdang 43400, Malaysia (e-mail: bukhari@eng.upm.edu.my).

C. Jahanchahi and D. P. Mandic are with the Electrical and Electronic Engineering Department, Imperial College London, London SW7 2BT, U.K. (e-mail: cyrus.jahanchahi05@imperial.ac.uk; d.mandic@imperial.ac.uk).

C. Cheong Took is with the Department of Computing, University of Surrey, Surrey GU2 7XH, U.K. (e-mail: c.cheongtook@surrey.ac.uk).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TNNLS.2013.2248165

and is optimal for second-order noncircular (improper), whereas QLMS is optimal only for second-order circular (proper) data but is faster. The collaborative combination of QLMS (CCQLMS) and WL-QLMS, therefore, promises a flexible tool for modeling the generality quaternion-valued processes. We also show that the evolution of the convex mixing parameter within the proposed architecture indicates the nature of the underlying linear model of a given system.

The rest of this paper is organized as follows. Section II provides the basics of quaternion algebra. This is followed by an overview of augmented quaternion statistics in Section III. Section IV describes the operation of the proposed model order identification algorithms. The convergence of the convex mixing parameter is presented in Section V, followed by the steady-state analysis in Section VI. In Section VII, simulations supporting the proposed approaches are shown. This paper concludes in Section VIII.

## II. QUATERNION ALGEBRA

A quaternion variable  $q$  can be expressed as  $q = [q_a, \bar{q}] = q_a + q_b i + q_c j + q_d \kappa$ , where  $q_a, q_b, q_c, q_d \in \mathbb{R}$ ,  $\bar{q}$  is the vector part and  $i, j, \kappa$  are both orthogonal unit vectors and imaginary units. Quaternion algebra is a division algebra, and its unique property is the noncommutativity of multiplication, defined as

$$wx = [w_a, \bar{w}][x_a, \bar{x}] = [w_a x_a - \bar{w} \cdot \bar{x}, w_a \bar{x} + x_a \bar{w} + \bar{w} \times \bar{x}]$$

where the symbols “ $\cdot$ ” and “ $\times$ ” denote, respectively, the dot-product and cross-product. Observe that the quaternion multiplication is noncommutative because of the existence of the outer product between  $\bar{w}$  and  $\bar{x}$ .

Operators of crucial importance to this paper are the three quaternion involutions given by

$$\begin{aligned} q^i &= -i q i = q_a + q_b i - q_c j - q_d \kappa, \\ q^j &= -j q j = q_a - q_b i + q_c j - q_d \kappa, \\ q^\kappa &= -\kappa q \kappa = q_a - q_b i - q_c j + q_d \kappa. \end{aligned}$$

Fig. 1 illustrates the  $\kappa$ -involution; it can be seen that the  $\kappa$ -involution reflects the vector  $\zeta$  with respect to the imaginary axes defined by  $i$  and  $j$ . Similar visualizations can also be produced for the  $i$ - and  $j$ -involutions.

Two other operators used in this paper are the quaternion conjugate and the norm square, given by  $q^* = [q_a, \bar{q}]^* = [q_a, -\bar{q}]$  and  $\|q\|_2^2 = qq^* = q^*q$ . In the sequel, all the constants and variables are assumed quaternion-valued, unless stated otherwise.

## III. AUGMENTED QUATERNION STATISTICS

A real-valued mean square error (MSE) estimator is given by

$$\hat{y} = E[y|x]$$

where  $\hat{y}$  is the estimated process,  $x$  the observed variable, and  $E[\cdot]$  is the expectation operator.

For a jointly Gaussian  $x$  and  $y$ , the optimal solution is a linear estimator given by

$$\hat{y} = \mathbf{w}^T \mathbf{x}$$

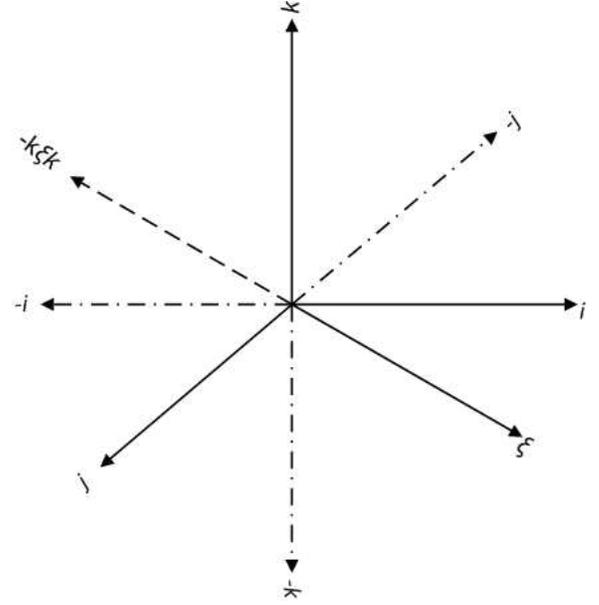


Fig. 1. Geometry of the  $\kappa$ -involution. The dotted line represents a plane in the opposite direction.

where  $\mathbf{w}$  and  $\mathbf{x}$  are, respectively, the real-valued coefficient and regressor vector. The symbol  $(\cdot)^T$  denotes the vector transpose operator.

For the MSE estimator in the complex domain  $\mathbb{C}$ , the standard solution is also the linear estimator given by

$$\hat{y} = \mathbf{w}^T \mathbf{x}$$

where  $\mathbf{w}$  and  $\mathbf{x}$  are the complex-valued coefficient and regressor vector, respectively. For insight, we can rewrite the complex domain MSE estimator componentwise, to yield

$$\hat{y}_\eta = E[y_\eta | x_r, x_i] \quad \eta \in \{r, i\}$$

and exploit the relationship between  $x_r$  and  $x_i$  given by

$$x_r = \frac{x + x^*}{2}; \quad x_i = \frac{x - x^*}{2}$$

to arrive at the complex-valued MSE estimator, given by

$$\hat{y} = E[y_r | x, x^*] + i E[y_i | x, x^*].$$

The optimal linear estimator becomes the widely linear model [12], [13]

$$\hat{y} = E[y|x, x^*] \Rightarrow y = \mathbf{h}^T \mathbf{x} + \mathbf{g}^T \mathbf{x}^*$$

where  $\mathbf{g}$  and  $\mathbf{h}$  are complex-valued regressor vectors.

Similarly, the current strictly linear quaternion-valued estimator is given by

$$\hat{y} = \mathbf{w}^T \mathbf{x}.$$

Upon expanding it componentwise, we have

$$\hat{y}_\eta = E[y_\eta | x_a, x_b, x_c, x_d] \quad \eta \in \{a, b, c, d\}.$$

Using the involutions in (1), the relationship between the input components  $x_a, x_b, x_c, x_d$  and  $x, x^I, x^J, x^K$  becomes

$$\begin{aligned} x_a &= \frac{1}{4}(x + x^I + x^J + x^K), \\ x_b &= \frac{1}{4}(x + x^I - x^J - x^K), \\ x_c &= \frac{1}{4}(x - x^I + x^J - x^K), \\ x_d &= \frac{1}{4}(x - x^I - x^J + x^K). \end{aligned}$$

A full capture of the second-order quaternion statistics available, following the same steps as in the complex case, results in the quaternion widely linear model given by [3], [4] and [14]

$\hat{y} = E[y|x, x^I, x^J, x^K] = \mathbf{w}^a \mathbf{x}^a = \mathbf{g}^T \mathbf{x} + \mathbf{h}^T \mathbf{x}^I + \mathbf{u}^T \mathbf{x}^J + \mathbf{v}^T \mathbf{x}^K$  where  $\mathbf{g}, \mathbf{h}, \mathbf{u}$ , and  $\mathbf{v}$  are the quaternion-valued regressor vectors and  $\mathbf{w}^a = [\mathbf{g}^T \ \mathbf{h}^T \ \mathbf{u}^T \ \mathbf{v}^T]^T$ ,  $\mathbf{x}^a = [\mathbf{x}^T \ \mathbf{x}^{I^T} \ \mathbf{x}^{J^T} \ \mathbf{x}^{K^T}]^T$  are the augmented weight vector and the augmented input vector.

Statistics based on the augmented input vector  $\mathbf{x}^a$  indicate that both the covariance matrix  $\mathcal{C}_{\mathbf{xx}} = E\{\mathbf{xx}^H\}$  and three other complementary covariance matrices need to be employed to fully describe a second-order noncircular signal in the quaternion domain  $\mathbb{H}$ . These complementary covariance matrices are termed the  $I$ -covariance  $\mathcal{C}_{\mathbf{x}^I} = E\{\mathbf{xx}^{IH}\}$ ,  $J$ -covariance  $\mathcal{C}_{\mathbf{x}^J} = E\{\mathbf{xx}^{JH}\}$ , and  $K$ -covariance  $\mathcal{C}_{\mathbf{x}^K} = E\{\mathbf{xx}^{KH}\}$  [3], [4].

Thus, the complete second-order characteristics of a quaternion random vector can be described by the augmented covariance matrix  $\mathcal{C}_{\mathbf{x}^a}^a$  of an augmented vector  $\mathbf{x}^a = [\mathbf{x}^T \ \mathbf{x}^{I^T} \ \mathbf{x}^{J^T} \ \mathbf{x}^{K^T}]^T$ , given by

$$\mathcal{C}_{\mathbf{x}^a}^a = E\{\mathbf{x}^a \mathbf{x}^{aH}\} = \begin{bmatrix} \mathcal{C}_{\mathbf{xx}} & \mathcal{C}_{\mathbf{x}^I} & \mathcal{C}_{\mathbf{x}^J} & \mathcal{C}_{\mathbf{x}^K} \\ \mathcal{C}_{\mathbf{x}^I}^H & \mathcal{C}_{\mathbf{x}^I \mathbf{x}^I} & \mathcal{C}_{\mathbf{x}^I \mathbf{x}^J} & \mathcal{C}_{\mathbf{x}^I \mathbf{x}^K} \\ \mathcal{C}_{\mathbf{x}^J}^H & \mathcal{C}_{\mathbf{x}^J \mathbf{x}^I} & \mathcal{C}_{\mathbf{x}^J \mathbf{x}^J} & \mathcal{C}_{\mathbf{x}^J \mathbf{x}^K} \\ \mathcal{C}_{\mathbf{x}^K}^H & \mathcal{C}_{\mathbf{x}^K \mathbf{x}^I} & \mathcal{C}_{\mathbf{x}^K \mathbf{x}^J} & \mathcal{C}_{\mathbf{x}^K \mathbf{x}^K} \end{bmatrix} \quad (1)$$

where the submatrices in (1) are calculated according to

$$\begin{aligned} \mathcal{C}_{\delta} &= E\{\mathbf{x}\delta^H\} & \mathcal{C}_{\alpha\beta} &= E\{\alpha\beta^H\} \\ \delta &\in \{\mathbf{x}^I, \mathbf{x}^J, \mathbf{x}^K\} & \alpha, \beta &\in \{\mathbf{x}, \mathbf{x}^I, \mathbf{x}^J, \mathbf{x}^K\}. \end{aligned}$$

#### IV. MODEL ORDER IDENTIFICATION

The proposed algorithms for the identification of the widely linear systems (and noncircularity of a signal) comprise two parts: an adaptive finite impulse response (FIR) filter which optimizes the adaptive weight coefficients, followed by the FT algorithm that adapts the tap length of the filter to an optimal length, all performed for any given time  $k$ . We first review the FIR filter weight updates and proceed to illustrate how the FT algorithm can be exploited within quaternion-valued adaptive systems.

##### A. Filter Weight Updates

Quaternion-valued adaptive filtering algorithms are based on optimizing a real-valued cost function of quaternion variables [15]

$$\begin{aligned} E(k) &= e_a^2(k) + e_b^2(k) + e_c^2(k) + e_d^2(k) \\ &= e(k)e^*(k) = \|e(k)\|_2^2 \end{aligned} \quad (2)$$

where the error  $e(k) = d(k) - y(k)$ , and  $d(k)$  and  $y(k)$ , respectively, are the desired and output signal. The terms  $e_a(k)$ ,  $e_b(k)$ ,  $e_c(k)$  and  $e_d(k)$  denote, respectively, the error component in the real part,  $I$  part,  $J$  part, and  $K$  part of a quaternion variable.

Quaternion-valued adaptive filtering algorithms minimize the cost function (2) through a gradient descent weight update specified by

$$\mathbf{w}(k+1) = \mathbf{w}(k) - \mu \nabla_{\mathbf{w}} E(k) \quad (3)$$

where  $\mu$  is a real-valued learning rate and the gradient  $\nabla_{\mathbf{w}} E(k)$  is given by [1], [16]

$$\nabla_{\mathbf{w}} E(k) = \frac{\partial E(k)}{\partial \mathbf{w}^*} = \frac{\partial E(k)}{\partial \mathbf{w}_a} + \frac{\partial E(k)}{\partial \mathbf{w}_b} I + \frac{\partial E(k)}{\partial \mathbf{w}_c} J + \frac{\partial E(k)}{\partial \mathbf{w}_d} K.$$

An alternative is to utilize  $\mathbb{H}$ -calculus to directly calculate the gradient [1]. Because of the noncommutativity aspect of quaternion algebra, the  $\mathbb{H}$ -calculus is a nontrivial extension of the complex Wirtinger's calculus [17], [18] to the quaternion domain  $\mathbb{H}$ .

The strictly linear QLMS and WL-QLMS are based on gradient descent (3) described by [10] and [11], respectively. Both algorithms are described in Algorithm 1.

The collaborative adaptive estimator, shown in Fig. 2, consists of two independent subfilters sharing the common input  $\mathbf{x}(k)$  and desired signal  $d(k)$ . In [19], it was shown that a hybrid combination of a collaborative strictly linear and widely linear estimators has the ability to identify in real time both the system order and the identification of widely linearity of the systems, in the context of complex-valued widely linear modeling. Following that approach, we here employ a convex combination of QLMS and WL-QLMS, termed the convex combination QLMS (CC-QLMS), to form the overall output  $y_{cc}(k)$  given by

$$y_{cc}(k) = \lambda(k)y_1(k) + (1 - \lambda(k))y_w(k)$$

where  $\lambda(k)$  is a real-valued convex mixing parameter, whose update is governed by

$$\lambda(k+1) = \lambda(k) - \mu_{\lambda} \nabla_{\lambda} E(k)$$

where  $\mu_{\lambda}$  and  $\nabla_{\lambda} E(k)$  represent the real-valued step size and the error gradient.

The error gradient  $\nabla_{\lambda} E(k)$  can be evaluated as

$$\begin{aligned} \nabla_{\lambda} E(k) &= e_{cc}(k) \frac{\partial e_{cc}^*(k)}{\partial \lambda(k)} + \frac{\partial e_{cc}(k)}{\partial \lambda(k)} e_{cc}^*(k) \\ &= e_{cc}(k) (y_1(k) - y_w(k))^* \\ &\quad + (y_1(k) - y_w(k)) e_{cc}^*(k) \\ &= 2\Re\{e_{cc}(k) (y_1(k) - y_w(k))^*\} \end{aligned}$$

where  $e_{cc}(k) = d(k) - y_{cc}(k)$  is the error of the CC-QLMS algorithm, and symbol  $\Re\{\cdot\}$  denotes the real part of a quaternion variable. Finally, the weight update of the convex mixing parameter  $\lambda(k)$  has the form

$$\lambda(k+1) = \lambda(k) - \mu_{\lambda} \left( 2\Re\{e_{cc}(k) (y_1(k) - y_w(k))^*\} \right).$$

The convex nature of the CC-QLMS increases the robustness of the collaborative filter by allowing the filter to have

fast convergence for the processing of proper signals and enhanced steady-state performance for improper signals. The mixing parameter  $\lambda$  is maintained real-valued, to reflect the identification of the strictly and widely linear filter, and their combination matches the data, and also, physically, to ensure convexity of the mixing  $\lambda \in [0, 1] \subset \mathbb{R}$ . Because of the convex nature of the CC-QLMS and given the range of the mixing parameter  $\lambda(k)$  (within  $[0, 1]$ ), the CC-QLMS converges as long as one of the subsystems in Fig. 2 converges [20]; with this in mind, in the simulations the value of  $\lambda(k)$  is hard bounded to  $\lambda \in [0, 1]$ .

### B. Tap-Length Adaptation

The adaptation of model order in real time, also termed ‘‘tap-length adaptation,’’ is enabled by the adaptive nature of the models considered and has the ability to identify the evolving changes in the nature of the signal for time-varying system parameters.

Based on the tap-length adaptation, the filter length is extended or truncated at every time instant. The tap-length adaptation is governed by the FT algorithm given by [8]

$$\eta_f(k+1) = (\eta_f(k) - \alpha) - \gamma \cdot \left[ \left( E_p(k) \right) - \left( E_{p-\Delta}(k) \right) \right]$$

where  $\eta_f$  is the pseudo-fractional tap length which can take only a positive real value,  $\alpha$  and  $\gamma$  are the leaky factor and tap-length learning rate and small positive real parameters that satisfy  $\alpha \ll \gamma$ . Symbols  $E_p(k)$  and  $E_{p-\Delta}(k)$  denote, respectively, the instantaneous square errors for the tap lengths of order  $p$  and  $(p - \Delta)$ , the symbol  $p(k)$  denotes the true tap length at a discrete time instant  $k$ , and  $\Delta$  is a real positive integer such that  $\min\{p(k) - \Delta\} > 0$ .

The instantaneous square output errors for filters of lengths  $p$  and  $(p - \Delta)$  are given by

$$\begin{aligned} E_p(k) &= (e_p(k))(e_p(k))^*, \\ E_{p-\Delta}(k) &= (e_{p-\Delta}(k))(e_{p-\Delta}(k))^* \end{aligned} \quad (4)$$

and are based on the errors  $e_p(k)$  and  $e_{p-\Delta}(k)$ . These errors can be evaluated as

$$e_q(k) = d(k) - y_q(k) = d(k) - \mathbf{w}_q^T(k) \mathbf{x}_q(k)$$

where  $1 \leq q \leq p$ , while  $\mathbf{w}_q(k)$  and  $\mathbf{x}_q(k)$  are the vectors consisting of the first  $q$  coefficients of  $\mathbf{w}(k)$  and  $\mathbf{x}(k)$ , respectively.

To calculate the optimal model length, the tap-length parameter  $p(k)$  is made adaptive according to [8]

$$p(k+1) = \begin{cases} \lfloor \eta_f(k) \rfloor, & |p(k) - \eta_f(k)| \geq \delta \\ p(k), & \text{otherwise} \end{cases}$$

where  $\delta$  is a predefined integer threshold and the symbol  $\lfloor \cdot \rfloor$  denotes the floor operator. The coefficient adaptation for the FT collaborative architecture in Fig. 2 is summarized in Algorithm 1.

---

### Algorithm 1: Adaptation of the Collaborative Adaptive System Identification Architecture

---

#### Filter Weight Algorithms

##### CC-QLMS, QLMS:

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu \left( 2e_1(k) \mathbf{x}^*(k) - \mathbf{x}^*(k) e_1^*(k) \right)$$

##### CC-QLMS, WL-QLMS:

$$\mathbf{w}^a(k+1) = \mathbf{w}^a(k) + \mu \left( 2e_w(k) \mathbf{x}^{a*}(k) - \mathbf{x}^{a*}(k) e_w^*(k) \right)$$

##### CC-QLMS:

$$\lambda(k+1) = \lambda(k) - \mu_\lambda \left( 2\Re\{e_{cc}(k)(y_1(k) - y_w(k))^*\} \right)$$

##### FT Algorithm

$$\eta_f(k+1) = (\eta_f(k) - \alpha) - \gamma \cdot \left[ \left( E_p(k) \right) - \left( E_{p-\Delta}(k) \right) \right]$$

$$p(k+1) = \begin{cases} \lfloor \eta_f(k) \rfloor, & |p(k) - \eta_f(k)| \geq \delta \\ p(k), & \text{otherwise} \end{cases}$$


---

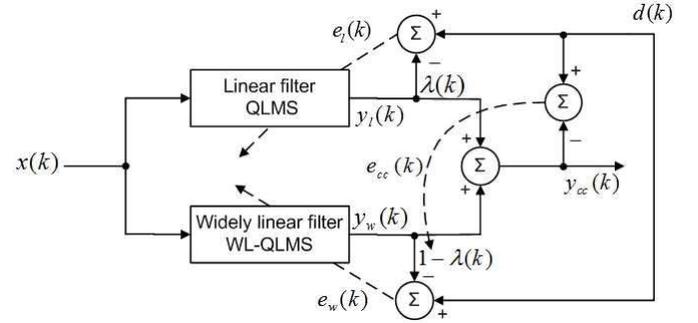


Fig. 2. Hybrid filter structure.

## V. CONVERGENCE OF THE MIXING PARAMETER

To address the convergence of the mixing parameter  $\lambda$  within CC-QLMS when processing both strictly proper and improper signals, consider first a strictly linear model given by

$$d(k) = \mathbf{w}^o T \mathbf{x}(k) + v(k)$$

where  $\mathbf{w}^o$  the optimal weight coefficients and  $v$  is a  $\mathbb{H}$ -circular quaternion-valued quadruply white Gaussian noise (WGN).

At the steady state, we employ the standard assumptions that  $\mathbf{w}^o = \mathbf{w} = \mathbf{g}$  and, due to circularity,  $\mathbf{h} = \mathbf{u} = \mathbf{v} = \mathbf{0}$ , which will lead to  $e_{cc}(k) = v(k)$ . The expression for the evolution of the mixing parameter  $\lambda$  can be rewritten as

$$\begin{aligned} \lambda(k+1) &= \lambda(k) - \mu_\lambda \left( 2\Re\{e_{cc}(k)(y_1(k) - y_w(k))^*\} \right) \\ &= \lambda(k) \end{aligned}$$

where the symbol  $\Re\{\cdot\}$  denotes the real part of a quaternion variable. Since the QLMS converges faster than the WL-QLMS (due to the smaller number of weight terms), the CC-QLMS will favor the QLMS-trained subfilter leading to  $\lambda(k) \rightarrow 1$ .

*Remark 1:* The mixing parameter  $\lambda \rightarrow 1$  within the CC-QLMS indicates a strictly linear quaternion-valued model, for which the power in all the components are equal. Current

models all assume properness, but this is not justified for real-world data (3-D inertial body sensors, wind).

Consider next the improper teaching signal (model) given by

$$d(k) = \mathbf{g}^{oT} \mathbf{x}(k) + \mathbf{h}^{oT} \mathbf{x}^l(k) + \mathbf{u}^{oT} \mathbf{x}^j(k) + \mathbf{v}^{oT}(k) \mathbf{x}^c(k) + v(k)$$

where  $\mathbf{g}^o$ ,  $\mathbf{h}^o$ ,  $\mathbf{u}^o$ , and  $\mathbf{v}^o$  are the optimal weight coefficients. At the steady state,  $\mathbf{g} = \mathbf{g}^o \neq \mathbf{w}$  and  $\mathbf{h} = \mathbf{h}^o$ ,  $\mathbf{u} = \mathbf{u}^o$ ,  $\mathbf{v} = \mathbf{v}^o$ , leading to  $e_{cc}(k) \neq \lambda(k)$ .

To simplify the analysis, we shall employ the following standard independence assumptions [21]:

- 1) the filter weights are independent of one another;
- 2) the error and the input vector are statistically independent of one another.

Enforcing these assumptions and applying the statistical expectation operator, the final expression is given as

$$\begin{aligned} E\{\lambda(k+1)\} &= E\{\lambda(k)\} \left( 1 - 2\mu_\lambda \|\mathbf{g}\|^2 \sigma_x^2 + \|\mathbf{h}\|^2 \sigma_x^2 \right. \\ &\quad \left. + \|\mathbf{u}\|^2 \sigma_x^2 + \|\mathbf{v}\|^2 \sigma_x^2 + \|\mathbf{w}\|^2 \sigma_x^2 \right) \\ &= E\{\lambda(k)\} \left( 1 - 2\mu_\lambda \Delta e^2 \right) \end{aligned}$$

where  $\sigma_x^2$  is the variance of the signal and  $\Delta e^2$  denotes the performance advantage of the widely linear model over the standard linear model. Since both the learning rate  $\mu_\lambda$  and  $\Delta e^2$  are positive, for the widely linear model or improper signal, the mixing parameter  $\lambda(k)$  converges towards zero, favoring the WL-QLMS subfilter within the collaborative architecture in Fig. 1.

*Remark 2:* The convergence of the mixing parameter within CC-QLMS to  $\lambda \rightarrow 0$  indicates a quaternion-valued widely linear model or equivalently a noncircular signal.

The above two remarks hold when the signal is stationary, the underlying system generating the signal is linear in the parameters, and no noise is present in the system. In the event that the first condition is violated, the mixing parameter  $\lambda$  does not have time to reach steady state and so  $\lambda$  can take any value between 0 and 1. Under this scenario,  $\lambda$  cannot be used to indicate whether the system is widely linear, as steady state is never reached. Therefore, we need to ensure that  $\lambda$  has enough time to converge, which is usually the case in practice. If the second or third condition is violated,  $\lambda$  will also lie between 0 and 1, lying closer to either extreme as the noise level approaches 0. In this case, the presence of noise brings closer together the learning curves of the QLMS and WL-QLMS, thus making it difficult for the CC-QLMS to identify with certainty which filter performs better. The further  $\lambda$  is from 0 or 1, the greater the uncertainty about the underlying model.

## VI. STEADY-STATE ANALYSIS OF FT-BASED ALGORITHMS

To illustrate the robustness of the proposed approach, and to convey the information about both the second-order model order and its linear or widely linear nature, we next provide a rigorous steady-state analysis of the of the FT-QLMS,

FT-WLQLMS, and FT-CCQLMS algorithms for two cases: 1) the desired system is widely linear, and 2) the desired system is strictly linear. We first consider the case of widely linear teaching signal and the FT-WLQLMS algorithm, where the desired (teaching) signal  $d(k)$  is defined as

$$d(k) = \mathbf{g}_{\text{Lopt}}^{oT} \mathbf{x}_{\text{Lopt}}(k) + \mathbf{h}_{\text{Lopt}}^{oT} \mathbf{x}_{\text{Lopt}}^l(k) + \mathbf{u}_{\text{Lopt}}^{oT} \mathbf{x}_{\text{Lopt}}^j(k) + \mathbf{v}_{\text{Lopt}}^{oT}(k) \mathbf{x}_{\text{Lopt}}^c(k) + v(k)$$

where  $\mathbf{g}_{\text{Lopt}}^o$ ,  $\mathbf{h}_{\text{Lopt}}^o$ ,  $\mathbf{u}_{\text{Lopt}}^o$ , and  $\mathbf{v}_{\text{Lopt}}^o$  are the optimal weight coefficients of the optimal tap lengths  $\text{Lopt}$  of the widely linear model, and  $v(k)$  is a  $\mathbb{H}$ -circular quaternion-valued quadruply WGN.

The output of the FT-WLQLMS algorithm is then given as

$$y_w(k) = \underbrace{\mathbf{g}^T(k) \mathbf{x}(k)}_{\text{standard part}} + \underbrace{\mathbf{h}^T(k) \mathbf{x}^l(k) + \mathbf{u}^T(k) \mathbf{x}^j(k) + \mathbf{v}^T(k) \mathbf{x}^c(k)}_{\text{augmented part}}. \quad (5)$$

Proceeding in a manner similar to the analysis in [22], the optimal weight vector coefficients can be split into three parts

$$\begin{aligned} \mathbf{g}_{\text{Lopt}}^o &= \begin{bmatrix} \mathbf{g}^{o'} \\ \mathbf{g}^{o''} \\ \mathbf{g}^{o'''} \end{bmatrix} & \mathbf{h}_{\text{Lopt}}^o &= \begin{bmatrix} \mathbf{h}^{o'} \\ \mathbf{h}^{o''} \\ \mathbf{h}^{o'''} \end{bmatrix} \\ \mathbf{u}_{\text{Lopt}}^o &= \begin{bmatrix} \mathbf{u}^{o'} \\ \mathbf{u}^{o''} \\ \mathbf{u}^{o'''} \end{bmatrix} & \mathbf{v}_{\text{Lopt}}^o &= \begin{bmatrix} \mathbf{v}^{o'} \\ \mathbf{v}^{o''} \\ \mathbf{v}^{o'''} \end{bmatrix} \end{aligned} \quad (6)$$

where  $\mathbf{g}^{o'}$ ,  $\mathbf{h}^{o'}$ ,  $\mathbf{u}^{o'}$ , and  $\mathbf{v}^{o'}$  are the coefficients modeled by tap length 1 to  $(p - \Delta)$ ;  $\mathbf{g}^{o''}$ ,  $\mathbf{h}^{o''}$ ,  $\mathbf{u}^{o''}$ , and  $\mathbf{v}^{o''}$  are the coefficients modeled by the tap length  $(p - \Delta + 1)$  to  $p$ ; and  $\mathbf{g}^{o'''}$ ,  $\mathbf{h}^{o'''}$ ,  $\mathbf{u}^{o'''}$ , and  $\mathbf{v}^{o'''}$  are the undermodeled coefficients.

For convenience, we denote the coefficient error vectors of the FT-WLQLMS as

$$\begin{aligned} \tilde{\mathbf{g}}(k) &= \mathbf{g}^o - \mathbf{g}_p(k); & \tilde{\mathbf{h}}(k) &= \mathbf{h}^o - \mathbf{h}_p(k), \\ \tilde{\mathbf{u}}(k) &= \mathbf{u}^o - \mathbf{u}_p(k); & \tilde{\mathbf{v}}(k) &= \mathbf{v}^o - \mathbf{v}_p(k) \end{aligned}$$

where  $\mathbf{g}_p(k)$ ,  $\mathbf{h}_p(k)$ ,  $\mathbf{u}_p(k)$ , and  $\mathbf{v}_p(k)$  are the weight vectors of length  $p$ .

Similar to (6), the weight error vectors can also be split up into three parts

$$\begin{aligned} \tilde{\mathbf{g}}(k) &= \begin{bmatrix} \tilde{\mathbf{g}}'(k) \\ \tilde{\mathbf{g}}''(k) \\ \tilde{\mathbf{g}}'''(k) \end{bmatrix} & \tilde{\mathbf{h}}(k) &= \begin{bmatrix} \tilde{\mathbf{h}}'(k) \\ \tilde{\mathbf{h}}''(k) \\ \tilde{\mathbf{h}}'''(k) \end{bmatrix} \\ \tilde{\mathbf{u}}(k) &= \begin{bmatrix} \tilde{\mathbf{u}}'(k) \\ \tilde{\mathbf{u}}''(k) \\ \tilde{\mathbf{u}}'''(k) \end{bmatrix} & \tilde{\mathbf{v}}(k) &= \begin{bmatrix} \tilde{\mathbf{v}}'(k) \\ \tilde{\mathbf{v}}''(k) \\ \tilde{\mathbf{v}}'''(k) \end{bmatrix}. \end{aligned}$$

In order to ensure mathematical tractability, we shall employ the following usual independence assumptions [22].

- 1) Both the input signal  $\mathbf{x}(k)$  and the noise  $v(k)$  are i.i.d. zero-mean white jointly Gaussian with the respective variances  $\sigma_x^2$  and  $\sigma_v^2$ .
- 2) At the steady state, the input signal  $\mathbf{x}(k)$  is independent of the weight vectors.

3) The tap-length parameter has converged at steady state, hence  $E\{\eta_f(k+1)\} = E\{\eta_f(k)\}$ , leading to the under-modeled error vectors vanishing.

We proceed by applying the statistical expectation operator to the steady-state MSE to yield

$$E\left\{\left(E_p(k) - \left(E_{p-\Delta}(k)\right)\right)\right\} < \left|\frac{\alpha}{\gamma}\right|. \quad (7)$$

Following the definition in (4), we can rewrite (7) to give

$$\begin{aligned} E\left\{\|\tilde{\mathbf{g}}''^T(k)\mathbf{x}''(k)\|_2^2 + \|\tilde{\mathbf{h}}''^T(k)\mathbf{x}''(k)\|_2^2\right. \\ \left. + \|\tilde{\mathbf{u}}''^T(k)\mathbf{x}''(k)\|_2^2 + \|\tilde{\mathbf{v}}''^T(k)\mathbf{x}''(k)\|_2^2\right. \\ \left. - \|\mathbf{g}''^T(k)\mathbf{x}''(k)\|_2^2 - \|\mathbf{h}''^T(k)\mathbf{x}''(k)\|_2^2\right. \\ \left. - \|\mathbf{u}''^T(k)\mathbf{x}''(k)\|_2^2 - \|\mathbf{v}''^T(k)\mathbf{x}''(k)\|_2^2\right\} < \left|\frac{\alpha}{\gamma}\right|. \quad (8) \end{aligned}$$

*Remark 3:* The FT-WLQLMS incorporates the errors from both the standard and augmented parts of the quaternion widely linear model in adapting the tap length, thus ensuring efficient modeling of the general widely linear quaternion-valued systems, i.e., incorporating both those with strictly and widely linear system dynamics.

In order to obtain the steady state of the FT-QLMS algorithm, we first set the augmented part in (5) to zero, which gives

$$y_1(k) = \mathbf{w}^T(k)\mathbf{x}(k).$$

Proceeding in a similar fashion to obtain the steady state of FT-WLQLMS yields

$$E\left\{\|\mathbf{w}''^T(k)\tilde{\mathbf{x}}''(k)\|_2^2 - \|\mathbf{g}''^T(k)\mathbf{x}''(k)\|_2^2\right\} < \left|\frac{\alpha}{\gamma}\right|.$$

*Remark 4:* The FT-QLMS only considers the error from the standard part of the quaternion widely linear model in adapting the tap length, proving to be inadequate for the modeling of widely linear quaternion-valued systems and improper quaternion-valued processes, i.e., those with different powers in the components, a typical case in real-world scenarios.

To derive the steady state of the FT-CCQLMS algorithm, consider the output of FT-CCQLMS, given by

$$\begin{aligned} y_{cc}(k) = \lambda(k)\mathbf{w}^T(k)\mathbf{x}(k) + \left(1 - \lambda(k)\right) \\ \times \left(\mathbf{g}^T(k)\mathbf{x}(k) + \mathbf{h}^T(k)\mathbf{x}'(k) + \mathbf{u}^T(k)\mathbf{x}^j(k)\right. \\ \left. + \mathbf{v}^T(k)\mathbf{x}^k(k)\right). \end{aligned}$$

Proceeding in a similar manner as above to obtain

FT-WLQLMS and FT-QLMS, the final steady state becomes

$$\begin{aligned} E\left\{\left(1 - \lambda(k)\right)\left(\|\tilde{\mathbf{g}}''^T(k)\mathbf{x}''(k)\|_2^2 + \|\tilde{\mathbf{h}}''^T(k)\mathbf{x}''(k)\|_2^2\right.\right. \\ \left.\left. + \|\tilde{\mathbf{u}}''^T(k)\mathbf{x}''(k)\|_2^2 + \|\tilde{\mathbf{v}}''^T(k)\mathbf{x}''(k)\|_2^2\right)\right. \\ \left. + \lambda(k) \cdot \|\tilde{\mathbf{w}}''^T(k)\mathbf{x}''(k)\|_2^2 - \|\mathbf{g}''^T(k)\mathbf{x}''(k)\|_2^2\right. \\ \left. - \|\mathbf{h}''^T(k)\mathbf{x}''(k)\|_2^2 - \|\mathbf{u}''^T(k)\mathbf{x}''(k)\|_2^2\right. \\ \left. - \|\mathbf{v}''^T(k)\mathbf{x}''(k)\|_2^2\right\} < \left|\frac{\alpha}{\gamma}\right|. \quad (9) \end{aligned}$$

As stated in Remark 2, for the processing of widely linear systems,  $\lambda \rightarrow 0$ , simplifying (9) to an expression similar to the steady state of FT-WLQLMS in (8).

*Remark 5:* For widely linear systems (improper signals), as  $\lambda \rightarrow 0$ , the steady state of FT-CCQLMS is similar to that of the FT-WLQLMS when identifying widely linear systems.

Consider next a strictly linear model given by

$$d(k) = \mathbf{w}_{\text{Lopt}}^{oT}\mathbf{x}_{\text{Lopt}}(k) + v(k)$$

which gives a similar steady-state expression for FT-QLMS and FT-WLQLMS, given by

$$\begin{aligned} \text{FT-QLMS} : E\left\{\|\tilde{\mathbf{w}}''^T(k)\mathbf{x}''(k)\|_2^2\right. \\ \left. - \|\mathbf{w}''^T(k)\mathbf{x}''(k)\|_2^2\right\} < \left|\frac{\alpha}{\gamma}\right| \end{aligned}$$

$$\begin{aligned} \text{FT-WLQLMS} : E\left\{\|\tilde{\mathbf{g}}''^T(k)\mathbf{x}''(k)\|_2^2\right. \\ \left. - \|\mathbf{w}''^T(k)\mathbf{x}''(k)\|_2^2\right\} < \left|\frac{\alpha}{\gamma}\right|. \end{aligned}$$

*Remark 6:* Both the FT-QLMS and FT-WLQLMS take into account the error from the standard (strictly linear) part of the quaternion linear model in adapting their tap lengths, demonstrating their suitability for the modeling of strictly linear quaternion-valued systems.

Similarly, the steady-state expression for FT-CCQLMS becomes

$$\begin{aligned} E\left\{\lambda(k)\|\tilde{\mathbf{w}}''^T(k)\mathbf{x}''(k)\|_2^2 + \left(1 - \lambda(k)\right)\right. \\ \left.\cdot \left(\|\tilde{\mathbf{g}}''^T(k)\mathbf{x}''(k)\|_2^2 + \|\tilde{\mathbf{h}}''^T(k)\mathbf{x}''(k)\|_2^2\right.\right. \\ \left.\left. + \|\tilde{\mathbf{u}}''^T(k)\mathbf{x}''(k)\|_2^2 + \|\tilde{\mathbf{v}}''^T(k)\mathbf{x}''(k)\|_2^2\right)\right. \\ \left. - \|\mathbf{w}''^T(k)\mathbf{x}''(k)\|_2^2\right\} < \left|\frac{\alpha}{\gamma}\right|. \end{aligned}$$

For optimal processing of the linear model,  $\lambda \approx 1$  (as stated by Remark 1), resulting in a similar expression to the steady state of FT-QLMS in (10).

*Remark 7:* The FT-CCQLMS will have similar behavior as FT-QLMS in (9) for the modeling of strictly linear quaternion-valued systems, when  $\lambda \rightarrow 1$ .

The analysis sets the scene for the assessment of the changes in some fundamental parameters of quaternion-valued systems and signals in real time and in time-varying environments, a typical case in real-world applications.

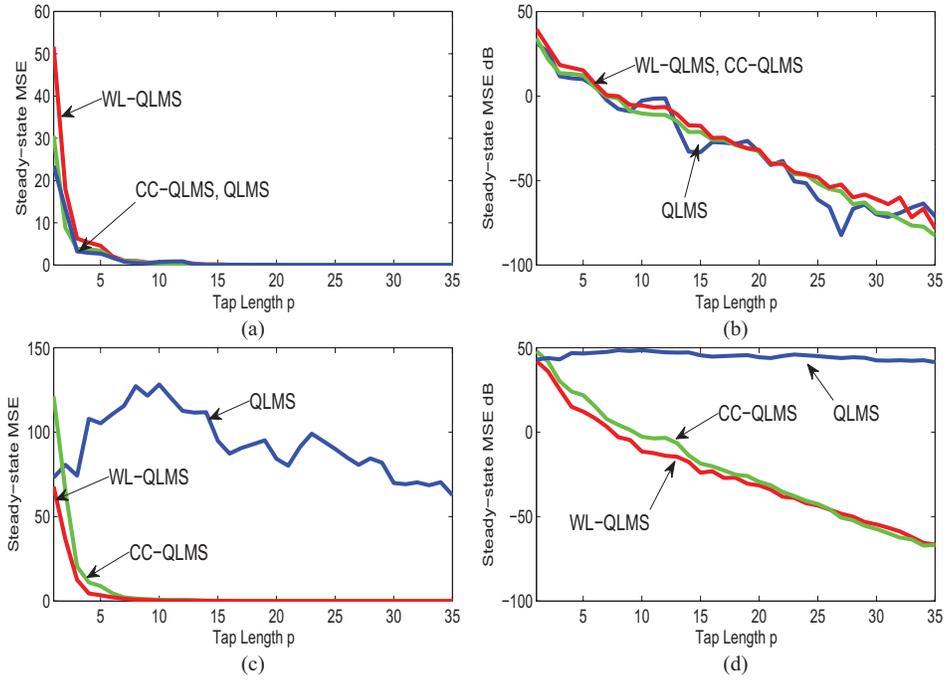


Fig. 3. Steady-state MSE for the circular process  $W_1$  and the noncircular process  $W_2$  with respect to the tap length (model order). (a) Steady-state MSE for the proper process  $W_1$ . (b) Steady-state MSE for the proper process  $W_1$  in dB. (c) Steady-state MSE for the improper process  $W_2$ . (d) Steady-state MSE for the improper process  $W_2$  in dB.

## VII. SIMULATIONS

Simulations were conducted in the system identification setting, and performances of FT-QLMS, FT-WLQLMS, and FT-CCQLMS were evaluated for a range of both benchmark and real-world systems. For benchmark systems, the quaternion quadruply white circular Gaussian noise (QWGN) defined by

$$\epsilon(k) = \epsilon_a(k) + \epsilon_b(k)\iota + \epsilon_c(k)j + \epsilon_d(k)\kappa$$

served as the driving input, where  $\epsilon_a$ ,  $\epsilon_b$ ,  $\epsilon_c$ , and  $\epsilon_d$  are realizations of real-valued independent WGNs.

The QWGN was first fed through a filter defined by  $A(k) = 0.35\epsilon(k) + \epsilon(k-1) + 0.35\epsilon(k-2)$ , which slows down the convergence of the algorithms. The output of  $A(k)$  was then fed into the systems defined by

$$\begin{aligned} W_1(k) &= 1.79W_1(k-1) - 1.85W_1(k-2) \\ &\quad + 1.27W_1(k-3) - 0.41W_1(k-4) + A(k), \\ W_2(k) &= 1.79W_2(k-1) - 1.85W_2(k-2) \\ &\quad + 1.27W_2(k-3) - 0.41W_2(k-4) + A(k) \\ &\quad + 0.5A^*(k) + 0.9A^*(k-1) \end{aligned}$$

where  $W_1$  is a strictly linear autoregressive (AR)(4) system [23] and  $W_2$  a widely linear WLAR(4) system [9], [24].

This way, the system  $W_1$  was linear (circular) and  $W_2$  widely linear (improper). The real-world process were D wind and 3-D inertial body motion data.

### A. Optimal Tap Length

The optimal tap lengths for both systems were determined by the steady-state MSE criterion [8]

$$\bar{\epsilon}(k) = \lambda_{\text{MSE}}\bar{\epsilon}(k-1) + (1 - \lambda_{\text{MSE}})E(k)$$

where  $\bar{\epsilon}$  is the estimated steady-state MSE and  $\lambda_{\text{MSE}} = 0.9$ .

Fig. 3 depicts the steady-state MSE for both the strictly linear system  $W_1$  and widely linear system  $W_2$ , using the QLMS, WLQLMS, and CCQLMS algorithms with  $\mu = 10^{-3}$ . From Fig. 3(a), it can be seen that all the three MSE curves were monotonically nonincreasing functions of the tap length, and as such the optimal tap lengths for all algorithms were found to be  $p_0 = \{15, 16\}$ . This is supported by Fig. 3(b), where the dB plot shows that the steady-state MSE values for all algorithms are decreasing with increasing tap length. Fig. 3(c) shows that the shape of the MSE curve for QLMS did not asymptotically converge, thus proving the inability of the strictly linear QLMS to model the widely linear system  $W_2$ . The standard linear QLMS is unable to process optimally  $W_2$  because of the widely linear nature of the system. Therefore, the CC-QLMS takes advantage of the WL-QLMS to accurately predict  $W_2$ , exhibiting a better performance than the linear QLMS. On the other hand, the MSE curves for the WL-QLMS and CC-QLMS converged, indicating their ability to model  $W_2$  for which the optimal tap length was found to be  $p_0 = \{21, 22\}$ . This can also be observed from Fig. 3(d), as the QLMS steady state remains almost constant at 40 dB. The optimal tap lengths for both systems were not a single integer because of the use of feedforward filters, which can only give approximations of the considered AR infinite impulse response (IIR) system.

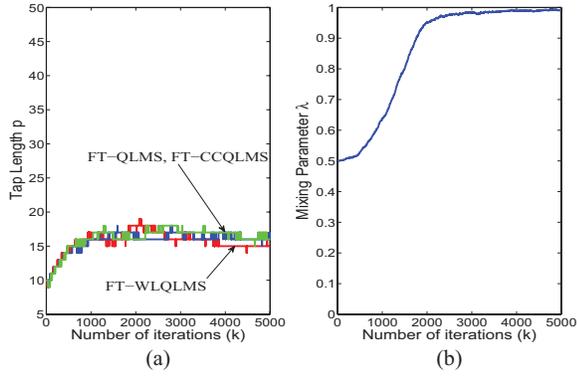


Fig. 4. Evolution of the optimal filter length parameter  $p$  and the mixing parameter  $\lambda$  for the modeling of the strictly linear system  $W_1$ . (a) Modelling of the strictly linear system  $W_1$ . (b) Modelling of the strictly linear system  $W_1$ .

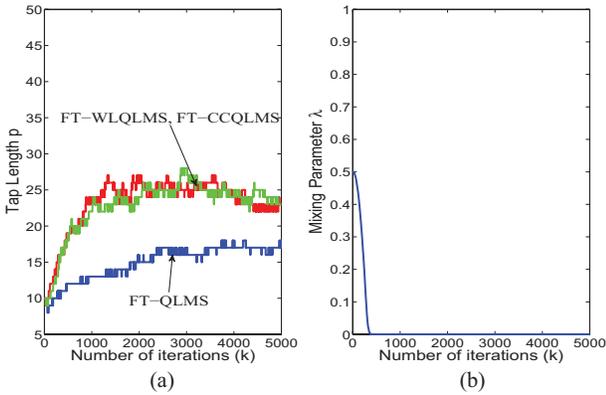


Fig. 5. Evolution of the optimal filter length parameter  $p$  and mixing parameter  $\lambda$  for the modeling of the widely linear system  $W_2$ . (a) Modelling of the widely linear system  $W_2$ . (b) Mixing Parameter  $\lambda$  of the widely linear system  $W_2$ .

### B. Modeling of Quaternion-Valued Systems

Fig. 4 depicts the evolution of the optimal tap length parameter  $p$  for the FT-QLMS, FT-WLQLMS, and FT-CCQLMS algorithms when employed for the modeling of strictly linear AR(4) system  $W_1$  along with the evolution of the mixing parameter  $\lambda$  of FT-CCQLMS. These algorithms were initialized with the following parameters:  $\alpha = 0.03$ ,  $\gamma = 1$ ,  $\delta = 1$ ,  $\Delta = 4$ ,  $\mu = 1 \times 10^{-5}$ ,  $\mu_1 = 5 \times 10^{-4}$ , the initial mixing parameter  $\lambda(0) = 0.5$ , and the initial tap length  $p(0) = 10$ . From Fig. 4(a), it is evident that the performances of all three algorithms considered were similar, as they converged to the optimal tap length at around the same number of iterations. This is in conformity with Remarks 6 and 7, which gives justification for their similar performances. Fig. 4(b) shows that the mixing parameter  $\lambda$  of FT-CCQLMS  $\lambda \rightarrow 1$  for the modeling of linear system  $W_1$ , conforming with Remark 1.

Similarly, Fig. 5 shows the results for the widely linear system  $W_2$ . Fig. 5(a) illustrates that the strictly linear FT-QLMS was unable to model the widely linear system  $W_2$ , whereas FT-WLQLMS and FT-CCQLMS converged to the optimal tap length. This is justified by Remarks 3–5. Fig. 5(b) illustrates that for the modeling of widely linear systems, value of  $\lambda \rightarrow 0$ , conforming with Remark 2.

TABLE I  
NONCIRCULAR QUADRUPLY WHITE QUATERNION  
GAUSSIAN NOISE USED

WGN	Noncircular
$\epsilon_a$	$\mathcal{N}(0, 1)$
$\epsilon_b$	$-0.6\epsilon_a + \mathcal{N}(0, 1)$
$\epsilon_c$	$0.8\epsilon_b + \mathcal{N}(0, 1)$
$\epsilon_d$	$0.8\epsilon_a - 0.4\epsilon_b + \mathcal{N}(0, 1)$

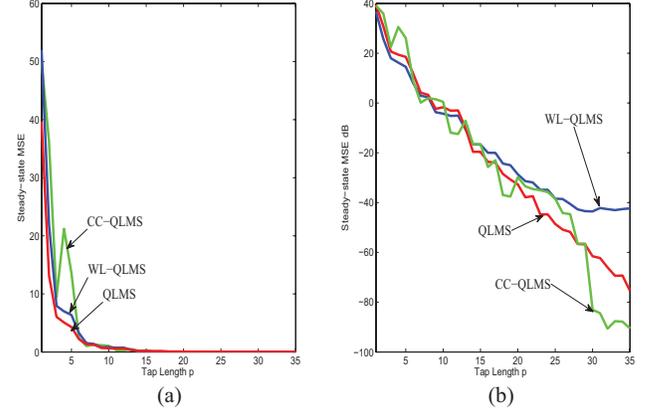


Fig. 6. Steady-state MSE for the process noncircular  $W_1$  with respect to tap length. (a) Steady-state MSE for the noncircular process  $W_1$ . (b) Steady-state MSE in dB for the noncircular process  $W_1$ .

### C. Evolution of Model Parameters in Nonstationary Systems

To illustrate the real-time tracking ability of the proposed algorithms, we next considered a system consisting of three separate subsystems of different natures. The first subsystem was the linear system  $W_1$  for the intervals of  $1 \leq k \leq 3000$ , followed by widely linear system  $W_2$  for  $3001 \leq k \leq 6000$ , and a linear noncircular  $W_1$  for the interval of  $6001 \leq k \leq 9000$ . The linear noncircular  $W_1$  was the original system  $W_1$  (10) fed with a noncircular QWGN as the driving input. The noncircular QWGN was constructed by referring to Table I.

Fig. 6 shows the steady-state MSE for the so-produced noncircular linear process  $W_1$  using the QLMS, WL-QLMS, and CCQLMS algorithms with  $\mu = 10^{-4}$ . As desired, all three of the MSE curves were monotonically nonincreasing functions of the tap length and the optimal tap length was found to be  $p_o = \{21, 22\}$ , illustrating the learning ability of the proposed architecture. In Fig. 6(a), for the processing of noncircular linear  $W_1$  process, both the QLMS and CC-QLMS performances were identical at tap lengths around the optimal value  $p_0$ . This is because, despite the process being noncircular, its generating mechanism is still strictly linear, thus allowing the optimal processing using the linear QLMS. The CC-QLMS curve in Fig. 6(a) shows an error spike at the tap length  $p = 5$ . This is because the value  $p = 4$  is a local minimum and the CC-QLMS is struggling to escape from it. This can also be observed from the reduced slope of the QLMS and WL-QLMS learning curves for tap lengths  $4 \leq p \leq 6$ . In Fig. 6(b), it can be seen that at larger tap lengths, the steady-state MSE for the CC-QLMS is lower than the QLMS despite sharing the same optimal tap length. This is because,

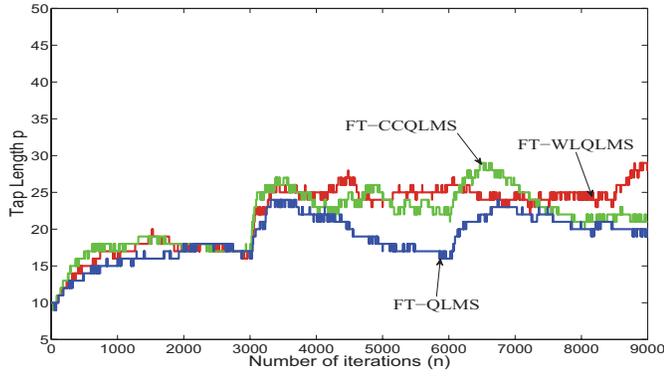


Fig. 7. Evolution of the optimal filter length parameter  $p$  for the modeling of the linear system  $W_1$  for the intervals  $1 \leq k \leq 3000$ , widely linear  $W_2$  for  $3001 \leq k \leq 6000$  and noncircular  $W_1$  for  $6001 \leq k \leq 9000$ .

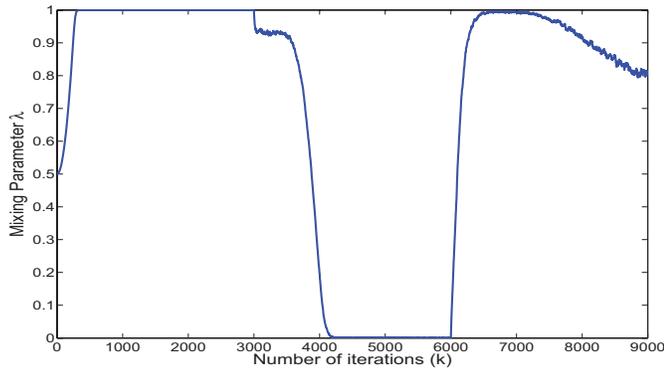


Fig. 8. Evolution of the mixing parameter  $\lambda$  for the modeling of the system  $W_1$  for the intervals  $1 \leq k \leq 3000$ ,  $W_2$  for  $3001 \leq k \leq 6000$ , and noncircular  $W_1$   $6001 \leq k \leq 9000$ .

for a noncircular linear process  $W_1$ , although the QLMS is dominant, the contribution from the WL-QLMS could not be completely neglected, as the value of the mixing parameter  $\lambda$  converges slightly below 1.

Fig. 7 shows the evolution of the optimal filter length parameter  $p$  for the FT-QLMS, FT-WLQLMS, and FT-CCQLMS employed for the modeling of the system  $W_1$  (interval of  $1 \leq k \leq 3000$ ),  $W_2$  (interval of  $3001 \leq k \leq 6000$ ), and the noncircular  $W_1$  (interval of  $6001 \leq k \leq 9000$ ). These algorithms were initialized as follows:  $\alpha = 0.03$ ,  $\gamma = 1$ ,  $\delta = 1$ ,  $\Delta = 4$ ,  $\mu = 1 \times 10^{-5}$ ,  $\mu_1 = 5 \times 10^{-4}$ , the initial mixing parameter  $\lambda(0) = 0.5$ , and the initial tap length  $p(0) = 25$ . From the figure, the FT-WLQLMS was able to converge to the optimal tap length of the system  $W_1$  for the interval  $1 \leq k \leq 3000$  and adapts to the system  $W_2$  for  $3001 \leq k \leq 6000$ . The FT-WLQLMS was unable to model efficiently the noncircular  $W_1$  for interval  $6001 \leq k \leq 9000$ . As for the FT-QLMS, it was incapable of adapting to the system  $W_2$  during the interval of  $3001 \leq k \leq 6000$ , but was able to model  $W_1$  and  $W_2$ . FT-CCQLMS was able to model all three systems owing to its theoretical advantages and the robust adaptation of the mixing parameter  $\lambda$ .

Fig. 8 depicts the evolution of the mixing parameter  $\lambda$  of FT-CCQLMS for the modeling of subsystems  $W_1$ ,  $W_2$ , and noncircular  $W_1$ . For the modeling of the linear system  $W_1$ ,

the parameter  $\lambda \rightarrow 1$  for the interval of  $1 \leq k \leq 3000$ , making FT-QLMS dominant over FT-WLQLMS. As for the widely linear system  $W_2$  in the interval  $3001 \leq k \leq 6000$ , the parameter  $\lambda \rightarrow 0$ , resulting in FT-WLQLMS to be superior. For the processing of the noncircular linear system  $W_1$ , the parameter  $\lambda \rightarrow 1$ , favoring the linear model. This corroborates with earlier findings in [25] and [26].

#### D. Evolution of the Circular Nature of Real-World 3-D Wind Field

In this simulation, a 3-D wind field was used as an input.<sup>1</sup> The wind data was initially sampled at 32 Hz and resampled at 10 Hz for simulation purposes. This wind data was divided into 80 sliding windows of length 1000 each. The motivation wind nonstationarity makes the circularity assessment difficult. However, by considering a segment of the wind, we can assume local stationarity, leading to the feasibility of measuring the circularity. The absolute complementary covariances were measured and normalized with respect to the covariance for each window segment, as shown in Fig. 9(a). According to [3], [4], [14], the complementary covariances vanish for circular region. Based on this understanding, we defined the low noncircularity region to be from window number 60 to 75 and the high noncircularity region from 30 to 45.

Fig. 9(b) shows the evolution of the mixing parameter  $\lambda$  for the high noncircularity region. Conforming with the analysis,  $\lambda$  converged to 0.2, favoring the WL-QLMS over the QLMS. The evolution of  $\lambda$  for the low noncircularity region is depicted in Fig. 9(c). Since the signal is predominantly circular in this region,  $\lambda$  approaches 0.8, signifying the dominance of the QLMS. These results corroborate with Remarks 1 and 2. This proves that the mixing parameter  $\lambda$  is able to track the underlying mechanisms of real-world signals. The values of  $\lambda$  do not take either extreme value since the three complementary covariances are never perfectly 0 or 1, hence the wind data is neither strictly linear nor widely linear. This is because parameter  $\lambda$  is inversely proportional to the absolute values of the complementary covariances.

#### E. Modeling of Real-World Inertial Motion Data

Five 3-D inertial body sensors were placed on the left arm, left hand, right arm, right hand, and the waist of an athlete performing Tai Chi movements and 3-D motion data were recorded using the XSense MTx 3DOF Orientation Tracker. The movement of the left arm was used as a pure quaternion input for this simulation. The optimal tap length cannot be determined through standard steady-state methods due to the nonstationary nature of human movements.

Fig. 10 shows the evolution of the optimal tap-length parameter  $p$  for the FT-QLMS, FT-WLQLMS, and FT-CCQLMS algorithms employed for the modeling of Tai Chi motion and the evolution of the mixing parameter  $\lambda$  of FT-CCQLMS. These algorithms were initialized as follows:  $\alpha = 0.03$ ,  $\gamma = 1$ ,  $\delta = 1$ ,  $\Delta = 4$ ,  $\mu = 1 \times 10^{-6}$ ,  $\mu_1 = 5 \times 10^{-4}$ , the initial mixing

<sup>1</sup>The wind data were sampled at 32 Hz and recorded by the 3-D windmaster anemometer provided by Gill instruments.

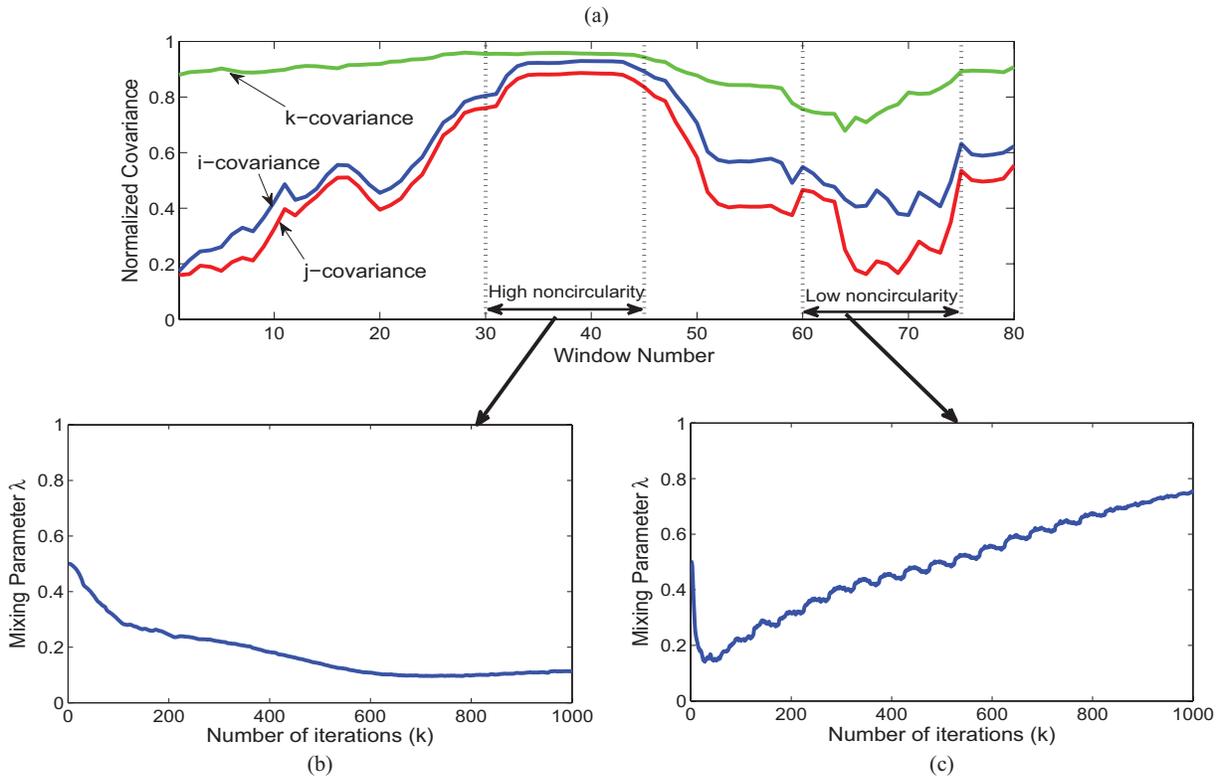


Fig. 9. Complementary covariances and the evolution of the mixing parameter  $\lambda$  for the modeling of a real-world wind field. (a) Normalized absolute complementary covariances of the 3-D wind data. (b) Mixing parameter  $\lambda$  of the high noncircularity region. (c) Mixing parameter  $\lambda$  of the low noncircularity region.

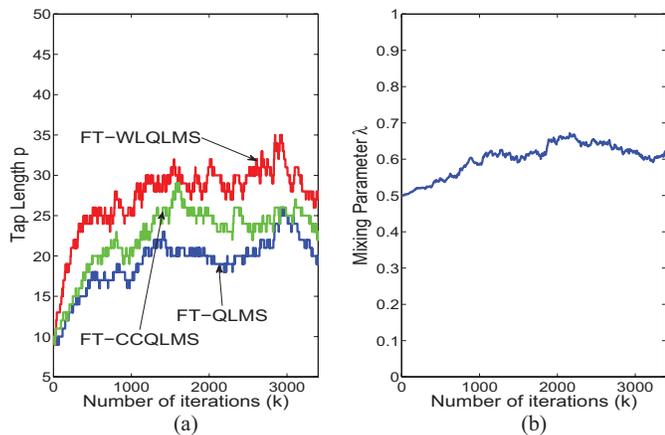


Fig. 10. Evolution of the optimal filter length parameter  $p$  and mixing parameter  $\lambda$  for the modelling of real-world Tai Chi motion. (a) Modelling of the real-world TaiChi motion. (b) Mixing Parameter  $\lambda$  of the real-world TaiChi motion.

parameter  $\lambda(0) = 0.5$ , and the initial tap length  $p(0) = 10$ . It is apparent from Fig. 10(a) that the optimal tap length for the three algorithms fluctuates because of the nonstationary nature of the signal. The FT-WLQLMS had the largest tap-length value, FT-QLMS the smallest, and FT-CCQLMS was in between. To support the results, the mixing parameter  $\lambda$  of the FT-CCQLMS was analyzed, as shown in Fig. 10(b). The values of  $\lambda$  converged to around 0.6, which means that the Tai Chi motion is neither strictly linear nor overly widely linear in nature. Hence, we can deduce that the expected

optimal tap length should be in between those estimated by the linear model (FT-QLMS) and the widely linear model (FT-WLQLMS), which is in agreement with the results obtained by FT-CCQLMS.

## VIII. CONCLUSION

We introduced an FT optimization into quaternion-valued adaptive modeling and demonstrated the advantages in model order selection and the identification of linear or widely linear nature of the system. The collaborative adaptive FT-CCQLMS showed to be able to model efficiently both widely linear and strictly linear quaternion-valued systems and to have a number of theoretical and practical advantages. The analysis of the convergence of the mixing parameter and the relationship between the steady-state error and tap length was established, giving a mathematical justification to the modeling capabilities of all algorithms. Simulations on model order selection and the identification of quaternion-valued circular natures for the problems in renewable energy and human-centered computing support the approach.

## REFERENCES

- [1] 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, Jan. 2011.
- [2] B. C. Ujang, C. C. Took, and D. P. Mandic, "Quaternion valued nonlinear adaptive filtering," *IEEE Trans. Neural Netw.*, vol. 22, no. 8, pp. 1193–1206, Aug. 2011.

- [3] J. Via, D. Ramirez, and I. Santamaria, "Properness and widely linear processing of quaternion random vectors," *IEEE Trans. Inform. Theory*, vol. 56, no. 7, pp. 3502–3515, Jul. 2010.
- [4] C. C. Took and D. P. Mandic, "Augmented second-order statistics of quaternion random process," *Signal Process.*, vol. 91, no. 2, pp. 214–224, 2011.
- [5] S. Javidi, C. C. Took, and D. P. Mandic, "Fast independent component analysis algorithm for quaternion valued signals," *IEEE Trans. Neural Netw.*, vol. 22, no. 12, pp. 1967–1978, Dec. 2011.
- [6] J. Via, D. P. Palomar, L. Vielva, and I. Santamaria, "Quaternion ICA from second-order statistics," *IEEE Trans. Signal Process.*, vol. 59, no. 4, pp. 1586–1600, Apr. 2011.
- [7] Z. Pritzker and A. Feuer, "Variable length stochastic gradient algorithm," *IEEE Trans. Signal Process.*, vol. 39, no. 4, pp. 997–1001, Apr. 1991.
- [8] Y. Gong and C. F. N. Cowan, "An LMS style variable tap-length algorithm for structure adaptation," *IEEE Trans. Signal Process.*, vol. 53, no. 7, pp. 2400–2407, Jul. 2005.
- [9] B. C. Ujang, C. C. Took, and D. P. Mandic, "Identification of improper processes by variable tap-length complex valued adaptive filters," in *Proc. Int. Joint Conf. Neural Netw.*, Jul. 2010, pp. 1–6.
- [10] C. C. Took and D. P. Mandic, "The quaternion LMS algorithm for adaptive filtering of hypercomplex processes," *IEEE Trans. Signal Process.*, vol. 57, no. 4, pp. 1316–1327, Apr. 2009.
- [11] C. C. Took and D. P. Mandic, "A quaternion widely linear adaptive filter," *IEEE Trans. Signal Process.*, vol. 58, no. 8, pp. 4427–4431, Aug. 2010.
- [12] B. Picinbono and P. Chevalier, "Widely linear estimation with complex data," *IEEE Trans. Signal Process.*, vol. 43, no. 8, pp. 2030–2033, Aug. 1995.
- [13] J. Navarro-Moreno, J. Moreno-Kayser, R. M. Fernandez-Alcala, and J. C. Ruiz-Molina, "Widely linear estimation algorithms for second-order stationary signals," *IEEE Trans. Signal Process.*, vol. 57, no. 12, pp. 4930–4935, Dec. 2009.
- [14] J. Via, D. P. Palomar, and L. Vielva, "Generalized likelihood ratios for testing the properness of quaternion Gaussian vectors," *IEEE Trans. Signal Process.*, vol. 59, no. 4, pp. 1356–1370, Apr. 2011.
- [15] P. Arena, L. Fortuna, G. Muscato, and M. G. Xibilia, *Neural Networks in Multidimensional Domains* (Lecture Notes in Control and Information Sciences), vol. 234, New York, USA: Springer-Verlag, 1998.
- [16] C. A. Deavours, "The quaternion calculus," *Amer. Math. Monthly*, vol. 80, no. 9, pp. 995–1008, 1973.
- [17] K. Kreutz Delgado, "Complex gradient operator and the CR-calculus," Dept. Electr. Comput. Eng., UC San Diego, San Diego, CA, USA, Tech. Rep. ECE275A, 2006.
- [18] D. Brandwood, "A complex gradient operator and its application in adaptive array theory," *IEEE F Commun., Radar Signal Process.*, vol. 130, no. 1, pp. 11–16, Feb. 1983.
- [19] B. Jelfs, D. P. Mandic, and S. C. Douglas, "An adaptive approach for the identification of improper complex signals," *Signal Process.*, vol. 92, no. 2, pp. 335–344, Feb. 2012.
- [20] J. Arenas-Garcia, A. H. Sayed, and A. R. Figueiras-Vidal, "Mean-square performance of a convex combination of two adaptive filters," *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 1078–1090, Mar. 2006.
- [21] S. Haykin, *Adaptive Filter Theory*. Englewood Cliffs, NJ, USA: Prentice Hall, 2002.
- [22] Y. Zhang, N. Li, J. A. Chambers, and A. H. Sayed, "Steady-state performance analysis of variable tap-length LMS algorithm," *IEEE Trans. Signal Process.*, vol. 56, no. 2, pp. 839–845, Feb. 2008.
- [23] D. P. Mandic and V. S. L. Goh, *Complex Valued Nonlinear Adaptive Filters: Noncircularity, Widely Linear and Neural Models*, New York, USA: Wiley, 2009.
- [24] J. Navarro-Moreno, "ARMA prediction of widely linear systems by using the innovations algorithm," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 3061–3068, Jul. 2008.
- [25] B. Jelfs, S. Javidi, P. Vayanos, and D. P. Mandic, "Characterisation of signal modality: Exploiting signal nonlinearity in machine learning and signal processing," *J. Signal Process. Syst.*, vol. 61, no. 1, pp. 105–115, 2010.
- [26] E. Ollila, "On the circularity of a complex random variable," *IEEE Signal Process. Lett.*, vol. 15, no. 11, pp. 841–844, Nov. 2008.

**Bukhari Che Ujang** (S'11) received the B.Eng. degree in electrical engineering from Multimedia University, Cyberjaya, Malaysia, in 2006, and the Ph.D. degree in adaptive signal processing from Imperial College London, London, U.K.

He is currently a Senior Lecturer with Universiti Putra Malaysia, Selangor, Malaysia. His current research interests include adaptive filters, nonlinear signal processing, and quaternion-valued analysis.

**Cyrus Jahanchahi** (S'11) received the M.Eng. degree in electrical and electronic engineering from Imperial College London, London, U.K., in 2009. He is currently pursuing the Ph.D. degree in adaptive signal processing at Imperial College London.

His current research interests include quaternion valued adaptive signal processing and widely linear modelling.

**Clive Cheong Took** (S'06–M'07) received the B.S. degree in telecommunication engineering from King's College London University, London, U.K., where he was the top departmental graduate in 2004, and the Ph.D. degree in blind signal processing from Cardiff University, Cardiff, U.K., in 2007.

He is currently a Lecturer with the University of Surrey, Surrey, U.K. His current research interests include adaptive, blind, and multidimensional signal processing.

**Danilo P. Mandic** (M'99–SM'03) is a Professor in signal processing with Imperial College London, London, U.K.

He has been working in the area of nonlinear adaptive signal processing and nonlinear dynamics. His publication record includes two research monographs titled *Recurrent Neural Networks for Prediction* (1st ed., 2001) and *Complex Valued Nonlinear Adaptive Filters: Noncircularity, Widely Linear and Neural Models* (1st ed., Wiley, 2009), an edited book titled *Signal Processing for Information Fusion* (Springer, 2008) and more than 200 publications on signal and image processing. He has been a Guest Professor at K.U. Leuven Belgium, TUAT Tokyo, Japan, and Westminster University, London, and a Frontier Researcher in RIKEN Japan.

Prof. Mandic has been a Member of the *IEEE Technical Committee on Machine Learning for Signal Processing*, an Associate Editor for the *IEEE Transactions on Circuits and Systems II*, the *IEEE TRANSACTIONS ON SIGNAL PROCESSING*, *IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS*, and the *International Journal of Mathematical Modelling and Algorithms*. He has produced award winning papers and products resulting from his collaboration with the industry. He is a member of the London Mathematical Society.