

# Performance Analysis of Quaternion-Valued Adaptive Filters in Nonstationary Environments

Min Xiang , Member, IEEE, Sithan Kanna , Student Member, IEEE, and Danilo P. Mandic , Fellow, IEEE

**Abstract**—Quaternion adaptive filters have been widely used for the processing of three-dimensional (3-D) and 4-D phenomena, but complete analysis of their performance is still lacking, partly due to the cumbersomeness of multivariate quaternion analysis. This causes difficulties in both understanding their behavior and designing optimal filters. Based on a thorough exploration of the augmented statistics of quaternion random vectors, this paper extends an analysis framework for real-valued adaptive filters to the mean and mean square convergence analyses of general quaternion adaptive filters in nonstationary environments. The extension is nontrivial, considering the noncommutative quaternion algebra, only recently resolved issues with quaternion gradient, and the multidimensional augmented quaternion statistics. Also, for rigor, in order to model a nonstationary environment, the system weights are assumed to vary according to a first-order random-walk model. Transient and steady-state performance of a general class of quaternion adaptive filters is provided by exploiting the augmented quaternion statistics. An innovative quaternion decorrelation technique allows us to develop intuitive closed-form expressions for the performance of quaternion least mean square (QLMS) filters with Gaussian inputs, which provide new insights into the relationship between the filter behavior and the complete second-order statistics of the input signal, that is, quaternion non-circularity. The closed-form expressions for the performance of strictly linear, semiwidely linear, and widely linear QLMS filters are investigated in detail, while numerical simulations for the three classes of QLMS filters with correlated Gaussian inputs support the theoretical analysis.

**Index Terms**—Quaternion adaptive filters, least mean square, mean square analysis, stability, steady-state performance, improper signals, noncircularity.

## I. INTRODUCTION

QUATERNIONS have traditionally been used in aerospace engineering and computer graphics in order to model three-dimensional rotations and orientations as their algebra avoids numerical problems associated with vector algebras [1]. The recently introduced augmented quaternion statistics [2], [3] and  $\mathbb{H}\mathbb{R}$  calculus [4], [5] have triggered a resurgence of research on quaternion-valued signal processing, owing to a

compact model of mutual information between data channels provided by quaternions, and the inherent physically meaningful interpretation for three-dimensional and four-dimensional problems. Recent research mainly focuses on adaptive filtering, neural networks, independent component analysis, and spectral estimation [6]–[11]. Quaternions have subsequently found new applications in communications, motion tracking, and biomedical engineering [12]–[15].

Recently, there have been recent extensive works on theory and applications of quaternion filters [16]–[20]. Traditional strictly linear quaternion filters based on the strictly linear model  $\hat{y} = \hat{\mathbf{h}}^H \mathbf{x}$  with the input vector  $\mathbf{x} \in \mathbb{H}^{M \times 1}$ , the weight vector  $\hat{\mathbf{h}} \in \mathbb{H}^{M \times 1}$ , and the output  $\hat{y} \in \mathbb{H}$ , utilise second-order quaternion statistics based on the standard covariance and are optimal only for estimating second-order circular (proper) quaternion signals [16]. Advances in quaternion statistics have established that widely linear quaternion filters based on the widely linear model  $\hat{y} = \hat{\mathbf{h}}^H \mathbf{x} + \hat{\mathbf{g}}^H \mathbf{x}^l + \hat{\mathbf{u}}^H \mathbf{x}^j + \hat{\mathbf{v}}^H \mathbf{x}^k$ , where  $\hat{\mathbf{g}}, \hat{\mathbf{u}}, \hat{\mathbf{v}} \in \mathbb{H}^{M \times 1}$  are complementary weight vectors, and  $\mathbf{x}^l, \mathbf{x}^j, \mathbf{x}^k$  are involutions of  $\mathbf{x}$ , exploit the three complementary covariances in addition to the standard covariance, and thus capture complete second-order statistical information in quaternion signals [17], [21]. For quaternion signals with special second-order statistical properties, the widely linear filters reduce to semi-widely linear filters based on the semi-widely linear model  $\hat{y}_n = \hat{\mathbf{h}}_n^H \mathbf{x}_n + \hat{\mathbf{g}}_n^H \mathbf{x}_n^l$  [22], [23]. Notice that, similar to the duality between complex filters and bivariate real filters [24], [25], quadrivariate real filters are isomorphic to widely linear quaternion filters, but are totally different from strictly and semi-widely linear quaternion filters.

Contrary to the research on real-valued adaptive filters [26]–[29], the complete performance analysis of quaternion adaptive filters is still an open problem, causing difficulties in understanding their behaviour and in the design of optimal filtering strategies. Since the quadrivariate real filters are not isomorphic to the whole class of quaternion filters, it is impossible to use the performance of the former to straightforwardly analyse the latter, and thus the performance analysis of the latter must be undertaken in the quaternion domain. The mean convergence analysis for the quaternion least mean square (QLMS) filters was proposed in [30], but the non-commutativity of quaternion products poses a challenge to the application of classical mean square convergence analysis methods for real-valued and complex-valued adaptive filters to the quaternion domain. For the mean square convergence analysis, the work in [31] discussed a simple case, the univariate strictly linear QLMS

Manuscript received May 2, 2017; revised August 21, 2017 and October 8, 2017; accepted November 24, 2017. Date of publication December 25, 2017; date of current version February 7, 2018. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Masahiro Yukawa. (Corresponding author: Min Xiang.)

The authors are with the Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2BT, U.K. (e-mail: m.xiang13@ic.ac.uk; ssk08@ic.ac.uk; d.mandic@ic.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/TSP.2017.2787102

algorithm, and derived the bounds of the step size. However, considering the difficulty in multivariate quaternion analysis due to the non-commutativity of quaternion products and the lack of quaternion matrix factorisation methods, it is challenging to extend the analysis in [31] to general multivariate QLMS algorithms.

Based on the analysis framework for real-valued adaptive filters developed in [27], [28], this paper analyses the mean and mean square convergence of general quaternion adaptive filters with general (proper or improper) quaternion inputs in non-stationary environments, and thus quantifies their transient and steady-state performance. The proposed analysis not only treats different adaptation approaches uniformly, but also caters for different linear models for quaternion signal estimation. In order to model a non-stationary environment, the system weights are assumed to vary according to a first-order random-walk model. By decorrelating the quaternion regressor vectors, the analysis of the QLMS filters with Gaussian inputs yields closed-form solutions relating to the second-order statistics of input data. The analytical results of the strictly linear QLMS (SL-QLMS), widely linear QLMS (WL-QLMS), and semi-widely linear QLMS (SWL-QLMS) filters are investigated in detail. The key contributions of our work are as follows: (i) a unified performance analysis for a general class of quaternion adaptive filters, encompassing various adaptation approaches and quaternion linear estimation models; (ii) the mean square performance of QLMS filters with Gaussian inputs is linked to the second-order statistics of quaternion inputs by representing a fourth-order quaternion moment matrix in terms of covariance and complementary covariance matrices; (iii) the application of a new decorrelation technique for random quaternion vectors to the Gaussian regressors of QLMS filters leads to concise and physically meaningful closed-form results. This work is therefore a non-trivial extension of the real-valued analysis [27], [28] to the quaternion domain. Compared to [27], [28], careful attention is paid to the non-commutative quaternion algebra, and quaternion analysis techniques are exploited to deal with augmented quaternion statistics. Compared to the mean square analysis of univariate SL-QLMS in [31], a general class of multivariate quaternion adaptive filters are considered by using a generic filter form and quaternion multivariate statistics. The so established analytical results bring new insights into the behaviour of the quaternion adaptive filters.

The rest of this paper is organised as follows. Section II provides an overview of quaternions and quaternion adaptive filters. Section III presents a convergence analysis framework for general quaternion adaptive filters to obtain the bounds of stability conditions and steady-state performances, from which the performance of QLMS filters with Gaussian regressors is deduced. In Section IV, a decorrelation technique for Gaussian regressor vectors of QLMS filters is exploited to derive concise and intuitive closed-form analytical results, which are exemplified with the SL-QLMS, SWL-QLMS and WL-QLMS. Simulations for QLMS filters with correlated improper Gaussian input vectors are presented in Section V, and Section VI 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) operators respectively,  $\text{diag}(\mathbf{A})$  creates a column vector containing the diagonal entries of matrix  $\mathbf{A}$ , and  $\text{Diag}(\mathbf{a})$  creates a diagonal matrix with the elements of vector  $\mathbf{a}$  on the diagonal. The symbol  $\mathbf{I}_M$  denotes an  $M \times M$  identity matrix,  $\mathbf{1}_M$  an  $M \times 1$  all-ones vector,  $[\mathbf{A}]_{a,b}$  the element on the  $a$ -th row and  $b$ -th column of matrix  $\mathbf{A}$ ,  $\|\mathbf{a}\|$  the Euclidean norm of vector  $\mathbf{a}$ ,  $E\{\cdot\}$  the statistical expectation operator, while  $\lambda(\mathbf{A})$ ,  $\lambda_{\max}(\mathbf{A})$ , and  $\lambda_{\min}(\mathbf{A})$  denote the eigenvalue, maximum eigenvalue, and minimum eigenvalue of matrix  $\mathbf{A}$ .

## II. BACKGROUND

### A. Quaternion Algebra

The quaternion domain  $\mathbb{H}$  is a four-dimensional vector space over the real field  $\mathbb{R}$ , spanned by the basis  $\{1, \iota, j, \kappa\}$ . A random quaternion variable  $x \in \mathbb{H}$  consists of a real part  $\Re[\cdot]$  and a imaginary part  $\Im[\cdot]$  which comprises three imaginary components, so that

$$x = \Re[x] + \Im[x] = \Re[x] + \Im_\iota[x]\iota + \Im_j[x]j + \Im_\kappa[x]\kappa \quad (1)$$

where  $\Re[x]$ ,  $\Im_\iota[x]$ ,  $\Im_j[x]$ ,  $\Im_\kappa[x]$  are real variables and  $\iota, j, \kappa$  are imaginary units with the properties

$$\begin{aligned} \iota^2 = j^2 = \kappa^2 = \iota j \kappa = -1, \quad \iota j = -j \iota = \kappa \\ j \kappa = -\kappa j = \iota, \quad \kappa \iota = -\iota \kappa = j \end{aligned}$$

The conjugate of  $x$  is defined as

$$x^* = \Re[x] - \Im[x] = \Re[x] - \Im_\iota[x]\iota - \Im_j[x]j - \Im_\kappa[x]\kappa \quad (2)$$

The modulus of  $x$  is given by

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

and the product of two quaternions  $x, y \in \mathbb{H}$  by

$$\begin{aligned} xy = \Re[x]\Re[y] - \Im[x] \cdot \Im[y] + \Re[x]\Im[y] + \Re[y]\Im[x] \\ + \Im[x] \times \Im[y] \end{aligned}$$

where “ $\cdot$ ” denotes the scalar product and “ $\times$ ” the vector product. The presence of the vector product causes non-commutativity of the quaternion product, that is,  $xy \neq yx$ . The quaternion product has the following properties:

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

A quaternion variable  $x$  is called a pure quaternion if  $\Re[x] = 0$ . A quaternion variable  $x$  is called a unit quaternion if  $|x| = 1$ . Similar to complex numbers, a quaternion  $x$  can be represented in the polar form as  $x = |x|e^\zeta$ , where  $\zeta$  is a pure quaternion, and the exponential can be represented by the sum  $e^\zeta = \sum_{n=0}^{\infty} \zeta^n / (n!)$  and satisfies  $(e^\zeta)^* = e^{-\zeta}$  [32].

Another important notion is the quaternion involution [33], which defines a self-inverse mapping, analogous to the complex conjugate. The general involution of the quaternion variable  $x$  is defined as  $x^\alpha = -\alpha x \alpha$ , which represents the rotation of the vector part of  $x$  by  $\pi$  about a unit pure quaternion  $\alpha$ . The quaternion involutions have the property:  $(x^\alpha)^\alpha = x$ . Accordingly,  $x^{\alpha*} = (x^\alpha)^* = (x^*)^\alpha$ . The three special cases of involutions

about the  $\iota$ ,  $j$  and  $\kappa$  imaginary axes are given by

$$x^\iota = -\iota x \iota = \Re[x] + \Im_\iota[x] \iota - \Im_j[x] j - \Im_\kappa[x] \kappa \quad (3)$$

$$x^j = -j x j = \Re[x] - \Im_\iota[x] \iota + \Im_j[x] j - \Im_\kappa[x] \kappa \quad (4)$$

$$x^\kappa = -\kappa x \kappa = \Re[x] - \Im_\iota[x] \iota - \Im_j[x] j + \Im_\kappa[x] \kappa \quad (5)$$

Due to the non-commutativity of quaternion products, quaternion matrices have different properties from real and complex matrices. For example, a quaternion square matrix has two types of eigenvalues [34].

*Definition 1:* Given  $\mathbf{A} \in \mathbb{H}^{M \times M}$ , if  $\mathbf{A}\mathbf{x} = \mathbf{x}\lambda$  for  $\lambda \in \mathbb{H}$  and some non-zero  $\mathbf{x} \in \mathbb{H}^{M \times 1}$ , then  $\lambda$  is called a right eigenvalue of  $\mathbf{A}$ , and  $\mathbf{x}$  is called a right eigenvector.

*Definition 2:* Given  $\mathbf{A} \in \mathbb{H}^{M \times M}$ , if  $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$  for  $\lambda \in \mathbb{H}$  and some non-zero  $\mathbf{x} \in \mathbb{H}^{M \times 1}$ , then  $\lambda$  is called a left eigenvalue of  $\mathbf{A}$ , and  $\mathbf{x}$  is called a left eigenvector.

Right eigenvalues have been well studied in literature, while left eigenvalues are less known and are not computationally well-posed [34]–[37]. Relevant to our analysis, a Hermitian quaternion matrix  $\mathbf{H} \in \mathbb{H}^{M \times M}$  has exactly  $M$  right eigenvalues, which are also left eigenvalues, and are real-valued. Note that  $\mathbf{H}$  can also have non-real left eigenvalues. There exists a unitary quaternion matrix  $\mathbf{U} \in \mathbb{H}^{M \times M}$ , which satisfies  $\mathbf{U}^H \mathbf{U} = \mathbf{U} \mathbf{U}^H = \mathbf{I}_M$ , such that  $\mathbf{H} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^H$  is the eigendecomposition of  $\mathbf{H}$  with the right eigenvalues of  $\mathbf{H}$  on the diagonal of a diagonal real matrix,  $\mathbf{\Lambda}$ . A Hermitian quaternion matrix is called positive definite, positive semi-definite, negative definite, or negative semi-definite if  $\mathbf{x}^H \mathbf{H} \mathbf{x}$  is respectively positive, non-negative, negative, or non-positive for any non-zero quaternion vector  $\mathbf{x} \in \mathbb{H}^{M \times 1}$ . This is the case if and only if all right eigenvalues of  $\mathbf{H}$  are positive, non-negative, negative, or non-positive.

In the subsequent analysis, without loss in generality, we will apply the eigendecomposition of Hermitian quaternion matrices,  $\mathbf{H} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^H$ , which only involves right eigenvalues. For conciseness, we will use the terminology, eigenvalues and eigenvectors, to denote right eigenvalues and right eigenvectors of Hermitian quaternion matrices [38].

Notice that the Weyl's inequality about the eigenvalues of Hermitian complex matrices also applies to Hermitian quaternion matrices. If  $\mathbf{A}$  and  $\mathbf{B}$  are Hermitian quaternion matrices with eigenvalues  $\lambda_1(\mathbf{A}) \geq \lambda_2(\mathbf{A}) \geq \dots \geq \lambda_M(\mathbf{A})$ ,  $\lambda_1(\mathbf{B}) \geq \lambda_2(\mathbf{B}) \geq \dots \geq \lambda_M(\mathbf{B})$ , then  $\mathbf{C} = \mathbf{A} + \mathbf{B}$  is also Hermitian with eigenvalues  $\lambda_1(\mathbf{C}) \geq \lambda_2(\mathbf{C}) \geq \dots \geq \lambda_M(\mathbf{C})$ , satisfying

$$\lambda_i(\mathbf{A}) + \lambda_j(\mathbf{B}) \leq \lambda_m(\mathbf{C}) \leq \lambda_k(\mathbf{A}) + \lambda_l(\mathbf{B}) \quad (6)$$

if  $k + l - 1 \leq m \leq i + j - 1$

$\forall i, j, k, l \in \{1, 2, \dots, M\}$  [39].

The singular value decomposition has also been recently extended to the quaternion domain [34]. For any  $\mathbf{A} \in \mathbb{H}^{M \times N}$ , of rank  $d$ , there exist unitary quaternion matrices  $\mathbf{U} \in \mathbb{H}^{M \times M}$  and  $\mathbf{V} \in \mathbb{H}^{N \times N}$  such that  $\mathbf{A} = \mathbf{U} \begin{pmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{V}$  where  $\mathbf{D} \in \mathbb{R}^{d \times d}$  is a diagonal matrix with  $d$  positive singular values of  $\mathbf{A}$  on the diagonal.

## B. Augmented Second-Order Quaternion Statistics

The set of involutions in (3)–(5), together with the original quaternion random variable  $x$ , forms the most frequently used basis for augmented quaternion statistics, which is at the core of the recently proposed widely linear processing methodology [2], [3]. Benefiting from the involution basis, augmented second-order statistics of a zero-mean random quaternion variable<sup>1</sup>  $x$  is exploited by  $\iota$ -,  $j$ -, and  $\kappa$ -covariances,  $C_{xx^\alpha} = E\{xx^{\alpha*}\}$ ,  $\alpha = \iota, j, \kappa$ , which are referred to as complementary covariances, together with the standard covariance  $C_{xx} = E\{xx^*\}$ . These covariances enable the characterisation of the quaternion impropriety (second-order non-circularity) which arises from the degree of correlation and/or power imbalance between imaginary components relative to the real component. The impropriety coefficients of a quaternion random variable  $x$  are defined as [31], [40]

$$r_\alpha = \left| \frac{C_{xx^\alpha}}{C_{xx}} \right|, \quad \alpha = \iota, j, \kappa \quad (7)$$

which reflect the correlation between  $x$  and each of its involutions, normalised by the signal power. Note that  $r_\alpha \in [0, 1]$ . The four degrees of freedom in the quaternion domain allow for different levels of properness:  $\mathbb{H}$ -properness,  $\mathbb{R}^\alpha$ -properness and  $\mathbb{C}^\alpha$ -properness.

*Definition 3 (Properness of a random quaternion variable):* A random quaternion variable  $x$  is  $\mathbb{H}$ -proper if it is uncorrelated with its involutions  $x^\iota$ ,  $x^j$  and  $x^\kappa$ , so that  $C_{xx^\iota} = C_{xx^j} = C_{xx^\kappa} = \mathbf{0}$ ;  $x$  is  $\mathbb{R}^\alpha$ -proper if it is only uncorrelated with the involution  $x^\alpha$ , so that only  $C_{xx^\alpha}$  among the three complementary covariances vanishes;  $x$  is  $\mathbb{C}^\alpha$ -improper<sup>2</sup> if it is only correlated with the involution  $x^\alpha$ , so that all complementary covariances except  $C_{xx^\alpha}$  vanish;  $x$  is maximally improper (rectilinear) if its impropriety coefficients are maximal, so that  $r_\iota = r_j = r_\kappa = 1$ .

Similarly, augmented second-order statistics of a zero-mean random quaternion column vector  $\mathbf{x}$  is exploited by  $\iota$ -,  $j$ -, and  $\kappa$ -covariance matrices,  $\mathbf{C}_{\mathbf{x}\mathbf{x}^\alpha} = E\{\mathbf{x}\mathbf{x}^{\alpha H}\}$ ,  $\alpha = \iota, j, \kappa$ , which are referred to as complementary covariance matrices, together with the standard Hermitian covariance matrix,  $\mathbf{C}_{\mathbf{x}\mathbf{x}} = E\{\mathbf{x}\mathbf{x}^H\}$ . The  $\alpha$ -complementary covariance matrix is  $\alpha$ -Hermitian, which means  $\mathbf{C}_{\mathbf{x}\mathbf{x}^\alpha} = (\mathbf{C}_{\mathbf{x}\mathbf{x}^\alpha})^{\alpha H}$ . The knowledge on both the covariance and the three complementary covariance matrices is necessary to ensure the utilisation of complete second-order statistical information. Using such knowledge, the semi-widely linear processing with the augmented signal vector  $\dot{\mathbf{x}} \triangleq [\mathbf{x}^T, \mathbf{x}^{\eta T}]^T$ , where  $\eta \in \{\iota, j, \kappa\}$ , and the widely linear processing with the augmented signal vector  $\ddot{\mathbf{x}} \triangleq [\mathbf{x}^T, \mathbf{x}^{\iota T}, \mathbf{x}^j T, \mathbf{x}^{\kappa T}]^T$  have been shown to achieve better performance for improper quaternion signals, compared to the traditional strictly linear processing, which is based on only  $\mathbf{x}$  [42]. The covariance

<sup>1</sup>Throughout the paper, we assume zero-mean quaternion variables. This does not affect the generality of our results.

<sup>2</sup>It is called  $\mathbb{C}^\alpha$ -proper in some literature [41], but we consider  $\mathbb{C}^\alpha$ -improper to be more intuitive.

matrices of  $\dot{\mathbf{x}}$  and  $\ddot{\mathbf{x}}$  can be represented by

$$\mathbf{C}_{\dot{\mathbf{x}}\dot{\mathbf{x}}} = E \{ \dot{\mathbf{x}}\dot{\mathbf{x}}^H \} = \begin{bmatrix} \mathbf{C}_{\mathbf{x}\mathbf{x}} & \mathbf{C}_{\mathbf{x}\mathbf{x}^i} \\ \mathbf{C}_{\mathbf{x}\mathbf{x}^j} & \mathbf{C}_{\mathbf{x}\mathbf{x}^k} \end{bmatrix}$$

$$\mathbf{C}_{\ddot{\mathbf{x}}\ddot{\mathbf{x}}} = E \{ \ddot{\mathbf{x}}\ddot{\mathbf{x}}^H \} = \begin{bmatrix} \mathbf{C}_{\mathbf{x}\mathbf{x}} & \mathbf{C}_{\mathbf{x}\mathbf{x}^i} & \mathbf{C}_{\mathbf{x}\mathbf{x}^j} & \mathbf{C}_{\mathbf{x}\mathbf{x}^k} \\ \mathbf{C}_{\mathbf{x}\mathbf{x}^i}^i & \mathbf{C}_{\mathbf{x}\mathbf{x}^i}^i & \mathbf{C}_{\mathbf{x}\mathbf{x}^i}^j & \mathbf{C}_{\mathbf{x}\mathbf{x}^i}^k \\ \mathbf{C}_{\mathbf{x}\mathbf{x}^j}^j & \mathbf{C}_{\mathbf{x}\mathbf{x}^j}^i & \mathbf{C}_{\mathbf{x}\mathbf{x}^j}^j & \mathbf{C}_{\mathbf{x}\mathbf{x}^j}^k \\ \mathbf{C}_{\mathbf{x}\mathbf{x}^k}^k & \mathbf{C}_{\mathbf{x}\mathbf{x}^k}^i & \mathbf{C}_{\mathbf{x}\mathbf{x}^k}^j & \mathbf{C}_{\mathbf{x}\mathbf{x}^k}^k \end{bmatrix}$$

Likewise, the complementary covariance matrices of  $\dot{\mathbf{x}}$  and  $\ddot{\mathbf{x}}$  can also be represented by block matrices built from the covariance and complementary covariance matrices of  $\mathbf{x}$ .

*Definition 4 (Properness of a random quaternion vector):*

A random quaternion vector  $\mathbf{x}$  is  $\mathbb{H}$ -proper if it is uncorrelated with its involutions  $\mathbf{x}^i$ ,  $\mathbf{x}^j$  and  $\mathbf{x}^k$ , so that  $\mathbf{C}_{\mathbf{x}\mathbf{x}^i} = \mathbf{C}_{\mathbf{x}\mathbf{x}^j} = \mathbf{C}_{\mathbf{x}\mathbf{x}^k} = \mathbf{0}$ ;  $\mathbf{x}$  is  $\mathbb{R}^\alpha$ -proper if it is only uncorrelated with the involution  $\mathbf{x}^\alpha$ , so that only  $\mathbf{C}_{\mathbf{x}\mathbf{x}^\alpha}$  among the three complementary covariances vanishes;  $\mathbf{x}$  is  $\mathbb{C}^\alpha$ -improper if it is only correlated with the involution  $\mathbf{x}^\alpha$ , so that all complementary covariances except  $\mathbf{C}_{\mathbf{x}\mathbf{x}^\alpha}$  vanish;  $\mathbf{x}$  is maximally improper (rectilinear) if it is maximally correlated with its involutions, so that the propriety coefficients of the elements of  $\mathbf{x}$  are all units.

### C. Quaternion Adaptive Filters

A fundamental problem in quaternion signal processing is to obtain the estimate,  $\hat{y}_n$ , of a desired signal,  $y_n \in \mathbb{H}$ , from a set of measurements,  $\mathbf{x}_n \in \mathbb{H}^{L \times 1}$ , which carries information about  $y_n$ , at time  $n$ . The estimation model  $\hat{y}_n = f(\mathbf{x}_n)$ , which incorporates the knowledge on the relationship between  $y_n$  and  $\mathbf{x}_n$ , is crucial for estimation performance. Three linear estimation models for quaternion signals have been proposed and can be represented in a unified form given by

$$\hat{y}_n = \hat{\mathbf{w}}_n^H \mathbf{s}_n \quad (8)$$

where  $\mathbf{s}_n \in \mathbb{H}^{M \times 1}$  is the regressor vector and  $\hat{\mathbf{w}}_n \in \mathbb{H}^{M \times 1}$  the weight vector. These three linear models arise from (8) as follows [2], [42]:

- Strictly linear model:

$$\mathbf{s}_n = \mathbf{x}_n, \quad M = L \quad (9)$$

- Semi-widely linear model:

$$\mathbf{s}_n = \dot{\mathbf{x}}_n, \quad M = 2L \quad (10)$$

- Widely linear model:

$$\mathbf{s}_n = \ddot{\mathbf{x}}_n, \quad M = 4L \quad (11)$$

The estimation model in (8) is based on the assumption that the desired signal arises from the following linear system:

$$y_n = \mathbf{w}_n^H \mathbf{s}_n + v_n \quad (12)$$

where the system noise  $v_n$  is i.i.d., zero-mean, independent of  $\mathbf{s}_n$ , and with variance  $\sigma_v^2$ . For the signal in (12), the optimal weight in minimum MSE estimation based on (8) can be calculated as

$$E \{ \mathbf{s}_n \mathbf{s}_n^H \}^{-1} E \{ \mathbf{s}_n y_n^* \} = E \{ \mathbf{w}_n \} \quad (13)$$

To explore the filter behaviour in non-stationary environments, we assume that the system weight vector  $\mathbf{w}_n \in \mathbb{H}^{M \times 1}$  varies

according to a widely used first-order random-walk model given by [26], [29], [43], [44]

$$\mathbf{w}_n = \mathbf{w}_{n-1} + \mathbf{q}_n \quad (14)$$

where  $\mathbf{q}_n \in \mathbb{H}^{M \times 1}$  is a random quaternion vector that is i.i.d., zero-mean and independent of  $\mathbf{s}_n$ . The expression in (14) is an approximation of a Markov model given by [45]

$$\mathbf{w}_n = \mathbf{A} \mathbf{w}_{n-1} + \mathbf{q}_n \quad (15)$$

where the coefficient matrix  $\mathbf{A}$  is assumed to be close to  $\mathbf{I}_M$  in a number of practical applications. For example, in modelling Rayleigh fading channels in a wireless communications environment, (15) is a first-order approximation of the variation of fading coefficients, and  $\mathbf{A}$  is usually close to  $\mathbf{I}_M$  [29]. In order to simplify the derivation during the convergence analysis, the real-valued adaptive filtering literature customarily assumes  $\mathbf{A} = \mathbf{I}_M$  to employ the first-order random-walk model in (14), which can be straightforwardly extended to the quaternion domain. Despite the simplified non-stationary scenarios studied here, the essence of the problem that is common to more complicated situations is retained.

A number of quaternion adaptive filtering algorithms, based on (8) and the stochastic gradient decent minimisation of MSE, assume the following form:

$$\begin{aligned} \hat{y}_n &= \hat{\mathbf{w}}_n^H \mathbf{s}_n \\ e_n &= y_n - \hat{y}_n \\ \hat{\mathbf{w}}_{n+1} &= \hat{\mathbf{w}}_n + \mu f[\mathbf{s}_n] \mathbf{s}_n e_n^* \end{aligned} \quad (16)$$

where  $\mu$  is the step size,  $f[\mathbf{s}_n]$  a real-valued scalar function of  $\mathbf{s}_n$ ,  $f[\mathbf{s}_n] : \mathbb{H}^{M \times 1} \rightarrow \mathbb{R}$ . Specific values of  $f[\mathbf{s}_n]$  correspond to specific filtering algorithms. For example,  $f[\mathbf{s}_n] = 1$  results in the QLMS,  $f[\mathbf{s}_n] = \|\mathbf{s}_n\|^{-2}$  results in the normalised QLMS, and  $f[\mathbf{s}_n] = (\epsilon + \|\mathbf{s}_n\|^2)^{-1}$  results in the  $\epsilon$ -normalised QLMS [43].

## III. PERFORMANCE ANALYSIS OF QUATERNION ADAPTIVE FILTERS

To analyse the convergence performance of the filters in the form of (16), define the weight-error vector as  $\tilde{\mathbf{w}}_n \triangleq \mathbf{w}_n - \hat{\mathbf{w}}_n$ , and express the filter error as  $e_n = \tilde{\mathbf{w}}_n^H \mathbf{s}_n + v_n$ . Then the weight-error vector recursion is expressed as

$$\begin{aligned} \tilde{\mathbf{w}}_{n+1} &= \tilde{\mathbf{w}}_n - \mu f[\mathbf{s}_n] \mathbf{s}_n e_n^* + \mathbf{q}_n \\ &= (\mathbf{I}_M - \mu f[\mathbf{s}_n] \mathbf{s}_n \mathbf{s}_n^H) \tilde{\mathbf{w}}_n - \mu f[\mathbf{s}_n] \mathbf{s}_n v_n^* + \mathbf{q}_n \end{aligned} \quad (17)$$

### A. Mean Analysis

Since  $v_n$  and  $\mathbf{q}_n$  are i.i.d., zero-mean and independent of  $\mathbf{s}_n$ , the expectation of (17) is given by

$$E \{ \tilde{\mathbf{w}}_{n+1} \} = E \{ (\mathbf{I}_M - \mu f[\mathbf{s}_n] \mathbf{s}_n \mathbf{s}_n^H) \tilde{\mathbf{w}}_n \} \quad (18)$$

A convenient change of coordinates is enabled by the eigen-decomposition,  $E \{ f[\mathbf{s}_n] \mathbf{s}_n \mathbf{s}_n^H \} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^H$ , where  $\mathbf{U}$  is a unitary quaternion matrix and  $\mathbf{\Lambda}$  is a real diagonal matrix containing eigenvalues of  $E \{ f[\mathbf{s}_n] \mathbf{s}_n \mathbf{s}_n^H \}$  on the diagonal. Define  $\mathbf{w}'_n = \mathbf{U}^H \tilde{\mathbf{w}}_n$ , then (18) becomes

$$E \{ \mathbf{w}'_{n+1} \} = E \{ (\mathbf{I}_M - \mu \mathbf{\Lambda}) \mathbf{w}'_n \} \quad (19)$$

from which we obtain that  $\lim_{n \rightarrow \infty} E\{\mathbf{w}'_n\} = \mathbf{0}$  if and only if the eigenvalues of the matrix  $E\{\mathbf{I}_M - \mu\mathbf{\Lambda}\}$  are all within  $(-1, 1)$ , or equivalently

$$0 < \mu < 2\lambda_{\max}^{-1} (E\{f[\mathbf{s}_n] \mathbf{s}_n \mathbf{s}_n^H\}) \quad (20)$$

### B. Mean Square Analysis

Define the *a priori* and *a posteriori* estimation errors as

$$e_{a,n} \triangleq \mathbf{w}_n^H \mathbf{s}_n - \hat{\mathbf{w}}_n^H \mathbf{s}_n = \tilde{\mathbf{w}}_n^H \mathbf{s}_n$$

$$e_{p,n} \triangleq \mathbf{w}_n^H \mathbf{s}_n - \hat{\mathbf{w}}_{n+1}^H \mathbf{s}_n = (\tilde{\mathbf{w}}_{n+1} - \mathbf{q}_n)^H \mathbf{s}_n$$

and the weighted *a priori* and *a posteriori* errors as  $e_{a,n}^\Sigma \triangleq \tilde{\mathbf{w}}_n^H \Sigma \mathbf{s}_n$  and  $e_{p,n}^\Sigma \triangleq (\tilde{\mathbf{w}}_{n+1} - \mathbf{q}_n)^H \Sigma \mathbf{s}_n$ , where  $\Sigma$  is a quaternion Hermitian positive definite weighting matrix. For notation conciseness, denote  $\mathbf{s}_n^H \Sigma \mathbf{s}_n$  by  $\|\mathbf{s}_n\|_\Sigma^2$ . Because of (17) and

$$e_{p,n}^\Sigma = e_{a,n}^\Sigma - \mu f[\mathbf{s}_n] e_n \mathbf{s}_n^H \Sigma \mathbf{s}_n = e_{a,n}^\Sigma - \mu f[\mathbf{s}_n] e_n \|\mathbf{s}_n\|_\Sigma^2 \quad (21)$$

we obtain

$$\tilde{\mathbf{w}}_{n+1} + \mathbf{s}_n \frac{e_{a,n}^{\Sigma*}}{\|\mathbf{s}_n\|_\Sigma^2} = \tilde{\mathbf{w}}_n + \mathbf{s}_n \frac{e_{p,n}^{\Sigma*}}{\|\mathbf{s}_n\|_\Sigma^2} + \mathbf{q}_n \quad (22)$$

Evaluating energies of both sides of (22) yields

$$\begin{aligned} \|\tilde{\mathbf{w}}_{n+1}\|_\Sigma^2 + \frac{|e_{a,n}^\Sigma|^2}{\|\mathbf{s}_n\|_\Sigma^2} &= \|\tilde{\mathbf{w}}_n\|_\Sigma^2 + \frac{|e_{p,n}^\Sigma|^2}{\|\mathbf{s}_n\|_\Sigma^2} + \|\mathbf{q}_n\|_\Sigma^2 \\ &\quad + 2\Re \left[ \mathbf{q}_n^H \Sigma \left( \tilde{\mathbf{w}}_n + \mathbf{s}_n \frac{e_{p,n}^{\Sigma*}}{\|\mathbf{s}_n\|_\Sigma^2} \right) \right] \end{aligned}$$

which can be rewritten by combining (21) as

$$\begin{aligned} \|\tilde{\mathbf{w}}_{n+1}\|_\Sigma^2 &= \|\tilde{\mathbf{w}}_n\|_{\Sigma_1}^2 + \mu^2 |v_n|^2 f^2[\mathbf{s}_n] \|\mathbf{s}_n\|_\Sigma^2 + \|\mathbf{q}_n\|_\Sigma^2 \\ &\quad + 2\Re \left[ \mathbf{q}_n^H \Sigma \left( \tilde{\mathbf{w}}_n + \mathbf{s}_n \frac{e_{p,n}^{\Sigma*}}{\|\mathbf{s}_n\|_\Sigma^2} \right) \right] \\ &\quad - 2\mu \Re \left[ f[\mathbf{s}_n] \left( e_{a,n}^\Sigma - \mu f[\mathbf{s}_n] \|\mathbf{s}_n\|_\Sigma^2 e_{a,n} \right) v_n \right] \\ \Sigma_1 &\triangleq \Sigma - \mu \left( f[\mathbf{s}_n] \mathbf{s}_n \mathbf{s}_n^H \Sigma + f[\mathbf{s}_n] \Sigma \mathbf{s}_n \mathbf{s}_n^H \right) \\ &\quad + \mu^2 f^2[\mathbf{s}_n] \|\mathbf{s}_n\|_\Sigma^2 \mathbf{s}_n \mathbf{s}_n^H \end{aligned}$$

The application of the statistical expectation operator to the above equation yields

$$E\left\{\|\tilde{\mathbf{w}}_{n+1}\|_\Sigma^2\right\} = E\left\{\|\tilde{\mathbf{w}}_n\|_{\Sigma_1}^2\right\} + E\left\{\|\mathbf{q}_n\|_\Sigma^2\right\} + \mu^2 \sigma_v^2 E\left\{f^2[\mathbf{s}_n] \|\mathbf{s}_n\|_\Sigma^2\right\} \quad (23)$$

For simplicity, we assume that the sequence of vectors  $\mathbf{s}_n$  is i.i.d.. Thus (23) becomes

$$E\left\{\|\tilde{\mathbf{w}}_{n+1}\|_\Sigma^2\right\} = E\left\{\|\tilde{\mathbf{w}}_n\|_{\Sigma'}^2\right\} + E\left\{\|\mathbf{q}_n\|_\Sigma^2\right\} + \mu^2 \sigma_v^2 E\left\{f^2[\mathbf{s}_n] \|\mathbf{s}_n\|_\Sigma^2\right\} \quad (24)$$

$$\begin{aligned} \Sigma' &\triangleq E\{\Sigma_1\} = \Sigma + \mu^2 E\left\{f^2[\mathbf{s}_n] \|\mathbf{s}_n\|_\Sigma^2 \mathbf{s}_n \mathbf{s}_n^H\right\} \\ &\quad - \mu \left( E\{f[\mathbf{s}_n] \mathbf{s}_n \mathbf{s}_n^H\} \Sigma + \Sigma E\{f[\mathbf{s}_n] \mathbf{s}_n \mathbf{s}_n^H\} \right) \end{aligned}$$

1) *Transient Analysis*: Consider now the transient behaviour of  $E\{\|\tilde{\mathbf{w}}_n\|^2\}$  by letting  $\Sigma = \mathbf{I}_M$  in (24). Following the analysis framework proposed in [27], we vectorise matrix  $\Sigma'$  into the vector  $\text{vec}(\Sigma')$  by stacking the columns of  $\Sigma$  on top of one another, and refer to  $\|\cdot\|_{\Sigma'}^2$  as  $\|\cdot\|_{\text{vec}(\Sigma')}$ . In this way, (24) can be rewritten as

$$E\left\{\|\tilde{\mathbf{w}}_{n+1}\|^2\right\} = E\left\{\|\tilde{\mathbf{w}}_n\|_{\text{vec}(\mathbf{I}_M)}^2\right\} + E\left\{\|\mathbf{q}_n\|^2\right\} + \mu^2 \sigma_v^2 E\left\{f^2[\mathbf{s}_n] \|\mathbf{s}_n\|^2\right\} \quad (25)$$

$$\mathbf{F} \triangleq \mathbf{I}_{M^2} - \mu\mathbf{A} + \mu^2\mathbf{B}$$

$$\mathbf{A} \triangleq \mathbf{I}_M \otimes E\{f[\mathbf{s}_n] \mathbf{s}_n \mathbf{s}_n^H\} + (E\{f[\mathbf{s}_n] \mathbf{s}_n \mathbf{s}_n^H\})^T \otimes \mathbf{I}_M$$

$$\mathbf{B} \triangleq E\left\{f^2[\mathbf{s}_n] (\mathbf{s}_n \mathbf{s}_n^H)^T \otimes_{\mathbb{R}} (\mathbf{s}_n \mathbf{s}_n^H)\right\}$$

where the symbol “ $\otimes$ ” denotes the left Kronecker product

$$\mathbf{A} \otimes \mathbf{B} \triangleq \begin{bmatrix} [\mathbf{A}]_{1,1} \mathbf{B} & \cdots & [\mathbf{A}]_{1,M} \mathbf{B} \\ \vdots & \ddots & \vdots \\ [\mathbf{A}]_{M,1} \mathbf{B} & \cdots & [\mathbf{A}]_{M,M} \mathbf{B} \end{bmatrix}$$

and “ $\otimes_{\mathbb{R}}$ ” the right Kronecker product [46]

$$\mathbf{A} \otimes_{\mathbb{R}} \mathbf{B} \triangleq \begin{bmatrix} \mathbf{B} [\mathbf{A}]_{1,1} & \cdots & \mathbf{B} [\mathbf{A}]_{1,M} \\ \vdots & \ddots & \vdots \\ \mathbf{B} [\mathbf{A}]_{M,1} & \cdots & \mathbf{B} [\mathbf{A}]_{M,M} \end{bmatrix}$$

Notice that the eigenvalues of  $\mathbf{F}$  are all real-valued as  $\mathbf{F}$  is a Hermitian quaternion matrix [38].

Based on the Cayley-Hamilton theorem, [27] has proved that (25) is stable for real variables if and only if all eigenvalues of  $\mathbf{F}$  are within  $(-1, 1)$ , which is equivalent to

$$0 < \mu < \min \left\{ \lambda_{\max}^{-1} (\mathbf{A}^{-1} \mathbf{B}), \max \left\{ \lambda \left( \begin{bmatrix} \mathbf{A}/2 & -\mathbf{B}/2 \\ \mathbf{I}_{M^2} & \mathbf{0} \end{bmatrix} \right) \in \mathbb{R}^+ \right\}^{-1} \right\} \quad (26)$$

This conclusion also holds for quaternion variables since the Cayley-Hamilton theorem applies to the quaternion domain [34], implying that (26) is also the mean square stability condition of quaternion adaptive filters. In contrast to real and complex analyses [25], [29], the steady-state performance of quaternion adaptive filters is difficult to quantify in a closed form via the above analysis approach, since the required vectorisation of quaternion matrices is hindered by the non-commutativity of quaternion products.

2) *Performance Bounds*: The sufficient and necessary stability condition shown in (26) is difficult to interpret owing to the complicated structure of matrices  $\mathbf{A}$  and  $\mathbf{B}$ . Next, we shall derive simple sufficient and necessary conditions for the mean square stability. Set  $\Sigma = \mathbf{I}_M$ , then the evolution of  $E\{\|\tilde{\mathbf{w}}_{n+1}\|^2\}$  can be deduced from (24) as

$$E\left\{\|\tilde{\mathbf{w}}_{n+1}\|^2\right\} = E\left\{\|\tilde{\mathbf{w}}_n\|_{\Sigma'}^2\right\} + E\left\{\|\mathbf{q}_n\|^2\right\} + \mu^2 \sigma_v^2 E\left\{f^2[\mathbf{s}_n] \|\mathbf{s}_n\|^2\right\} \quad (27)$$

$$\Sigma' = \mathbf{I}_M - 2\mu E\{f[\mathbf{s}_n] \mathbf{s}_n \mathbf{s}_n^H\} + \mu^2 E\left\{f^2[\mathbf{s}_n] \|\mathbf{s}_n\|^2 \mathbf{s}_n \mathbf{s}_n^H\right\}$$

Notice that the Cauchy–Schwarz inequality can be applied to the Hermitian quaternion matrix  $E\{f^2[\mathbf{s}_n]\|\mathbf{s}_n\|^2\mathbf{s}_n\mathbf{s}_n^H\}$  [47], yielding

$$\begin{aligned}\boldsymbol{\Sigma}' &\succeq \mathbf{I}_M - 2\mu E\{f[\mathbf{s}_n]\mathbf{s}_n\mathbf{s}_n^H\} + \mu^2 (E\{f[\mathbf{s}_n]\mathbf{s}_n\mathbf{s}_n^H\})^2 \\ &= (\mathbf{I}_M - \mu E\{f[\mathbf{s}_n]\mathbf{s}_n\mathbf{s}_n^H\})^2\end{aligned}\quad (28)$$

Observe that  $\boldsymbol{\Sigma}'$  is positive semi-definite, and its eigendecomposition is given by  $\boldsymbol{\Sigma}' = \mathbf{T}\boldsymbol{\Lambda}_{\boldsymbol{\Sigma}'}\mathbf{T}^H$ , where  $\mathbf{T}$  is a quaternion unitary matrix, and  $\boldsymbol{\Lambda}_{\boldsymbol{\Sigma}'}$  a real diagonal matrix with non-negative eigenvalues of  $\boldsymbol{\Sigma}'$  on the diagonal. Thus, we have

$$E\{\|\tilde{\mathbf{w}}_n\|_{\boldsymbol{\Sigma}'}^2\} = E\{\tilde{\mathbf{w}}_n^H \mathbf{T}\boldsymbol{\Lambda}_{\boldsymbol{\Sigma}'}\mathbf{T}^H \tilde{\mathbf{w}}_n\} \quad (29)$$

From (29), we next obtain

$$E\{\|\tilde{\mathbf{w}}_n\|_{\boldsymbol{\Sigma}'}^2\} \leq \lambda_{\max}(\boldsymbol{\Sigma}') E\{\|\tilde{\mathbf{w}}_n\|^2\} \quad (30)$$

where  $\lambda_{\max}(\boldsymbol{\Sigma}')$  is bounded by

$$\begin{aligned}\lambda_{\max}(\boldsymbol{\Sigma}') &\leq \xi_1 \triangleq 1 - 2\mu\lambda_{\min}(E\{f[\mathbf{s}_n]\mathbf{s}_n\mathbf{s}_n^H\}) \\ &\quad + \mu^2\lambda_{\max}(E\{f^2[\mathbf{s}_n]\|\mathbf{s}_n\|^2\mathbf{s}_n\mathbf{s}_n^H\})\end{aligned}$$

which is based on (6). This inequality, together with (27) and (30), yields

$$\begin{aligned}E\{\|\tilde{\mathbf{w}}_{n+1}\|^2\} &\leq \xi_1 E\{\|\tilde{\mathbf{w}}_n\|^2\} + E\{\|\mathbf{q}_n\|^2\} \\ &\quad + \mu^2\sigma_v^2 E\{f^2[\mathbf{s}_n]\|\mathbf{s}_n\|^2\}\end{aligned}\quad (31)$$

which is stable if  $\xi_1 < 1$ , thus implying the following sufficient condition for the mean square stability

$$0 < \mu < \frac{2\lambda_{\min}(E\{f[\mathbf{s}_n]\mathbf{s}_n\mathbf{s}_n^H\})}{\lambda_{\max}(E\{f^2[\mathbf{s}_n]\|\mathbf{s}_n\|^2\mathbf{s}_n\mathbf{s}_n^H\})} \quad (32)$$

From (29), we also have

$$E\{\|\tilde{\mathbf{w}}_n\|_{\boldsymbol{\Sigma}'}^2\} \geq \lambda_{\min}(\boldsymbol{\Sigma}') E\{\|\tilde{\mathbf{w}}_n\|^2\} \quad (33)$$

where  $\lambda_{\min}(\boldsymbol{\Sigma}')$  is bounded by

$$\begin{aligned}\lambda_{\min}(\boldsymbol{\Sigma}') &\geq \xi_2 \triangleq 1 - 2\mu\lambda_{\max}(E\{f[\mathbf{s}_n]\mathbf{s}_n\mathbf{s}_n^H\}) \\ &\quad + \mu^2\lambda_{\min}(E\{f^2[\mathbf{s}_n]\|\mathbf{s}_n\|^2\mathbf{s}_n\mathbf{s}_n^H\})\end{aligned}$$

This inequality, together with (27) and (33), yields

$$\begin{aligned}E\{\|\tilde{\mathbf{w}}_{n+1}\|^2\} &\geq \xi_2 E\{\|\tilde{\mathbf{w}}_n\|^2\} + E\{\|\mathbf{q}_n\|^2\} \\ &\quad + \mu^2\sigma_v^2 E\{f^2[\mathbf{s}_n]\|\mathbf{s}_n\|^2\}\end{aligned}\quad (34)$$

Thus, a necessary condition for the convergence of  $E\{\|\tilde{\mathbf{w}}_n\|^2\}$  is  $\xi_2 < 1$ , or equivalently

$$0 < \mu < \frac{2\lambda_{\max}(E\{f[\mathbf{s}_n]\mathbf{s}_n\mathbf{s}_n^H\})}{\lambda_{\min}(E\{f^2[\mathbf{s}_n]\|\mathbf{s}_n\|^2\mathbf{s}_n\mathbf{s}_n^H\})} \quad (35)$$

From (28), another necessary condition for the convergence of  $E\{\|\tilde{\mathbf{w}}_n\|^2\}$  is that the eigenvalues of the matrix  $(\mathbf{I}_M - \mu E\{f[\mathbf{s}_n]\mathbf{s}_n\mathbf{s}_n^H\})^2$  are all less than unit, which yields the same step size bound as in (20).

*Remark 1:* Any value of  $\mu$  that satisfies the mean square stability condition given in (26) also guarantees the mean stability.

Next, the steady-state performance of quaternion adaptive filters will be bounded through a similar approach. At the steady state, we have  $\lim_{n \rightarrow \infty} E\{\|\tilde{\mathbf{w}}_{n+1}\|^2\} = \lim_{n \rightarrow \infty} E\{\|\tilde{\mathbf{w}}_n\|^2\}$ ,

and thus the bounds of  $\text{MSD} = \lim_{n \rightarrow \infty} E\{\|\tilde{\mathbf{w}}_n\|^2\}$  can be derived from (31) and (34) as

$$\begin{aligned}\text{MSD} &\leq \frac{\mu^2\sigma_v^2\text{Tr}(\mathbf{C}_{\text{ss}}) + \text{Tr}(\mathbf{C}_{\text{qq}})}{2\mu\lambda_{\min}(E\{f[\mathbf{s}_n]\mathbf{s}_n\mathbf{s}_n^H\}) - \mu^2\lambda_{\max}(E\{f^2[\mathbf{s}_n]\|\mathbf{s}_n\|^2\mathbf{s}_n\mathbf{s}_n^H\})} \\ \text{MSD} &\geq \frac{\mu^2\sigma_v^2\text{Tr}(\mathbf{C}_{\text{ss}}) + \text{Tr}(\mathbf{C}_{\text{qq}})}{2\mu\lambda_{\max}(E\{f[\mathbf{s}_n]\mathbf{s}_n\mathbf{s}_n^H\}) - \mu^2\lambda_{\min}(E\{f^2[\mathbf{s}_n]\|\mathbf{s}_n\|^2\mathbf{s}_n\mathbf{s}_n^H\})}\end{aligned}\quad (36)$$

Since

$$E\{\|\tilde{\mathbf{w}}_n\|_{\mathbf{C}_{\text{ss}}}^2\} = E\{\tilde{\mathbf{w}}_n^H \mathbf{K}\boldsymbol{\Lambda}_{\text{ss}}\mathbf{K}^H \tilde{\mathbf{w}}_n\}$$

where  $\mathbf{K}\boldsymbol{\Lambda}_{\text{ss}}\mathbf{K}^H$  is the eigendecomposition of  $\mathbf{C}_{\text{ss}}$ , we obtain

$$\begin{aligned}\lambda_{\min}(\mathbf{C}_{\text{ss}}) E\{\|\tilde{\mathbf{w}}_n\|^2\} \\ \leq E\{\|\tilde{\mathbf{w}}_n\|_{\mathbf{C}_{\text{ss}}}^2\} \leq \lambda_{\max}(\mathbf{C}_{\text{ss}}) E\{\|\tilde{\mathbf{w}}_n\|^2\}\end{aligned}$$

so that  $\text{EMSE} = \lim_{n \rightarrow \infty} E\{\|\tilde{\mathbf{w}}_n\|_{\mathbf{C}_{\text{ss}}}^2\}$  is bounded by

$$\lambda_{\min}(\mathbf{C}_{\text{ss}}) \cdot \text{MSD} \leq \text{EMSE} \leq \lambda_{\max}(\mathbf{C}_{\text{ss}}) \cdot \text{MSD} \quad (37)$$

### C. Performance Analysis of QLMS

When  $f[\mathbf{s}_n] = 1$ , the general quaternion adaptive filtering algorithm in the form of (16) reduces to the QLMS. Therefore, the performance of QLMS filters can be obtained by substituting  $f[\mathbf{s}_n] = 1$  into the above analysis results for general quaternion adaptive filters. Specifically, the mean weight-error vector recursion in (18) becomes

$$E\{\tilde{\mathbf{w}}_{n+1}\} = (\mathbf{I}_M - \mu\mathbf{C}_{\text{ss}}) E\{\tilde{\mathbf{w}}_n\} \quad (38)$$

while the mean stability condition in (20) reduces to

$$0 < \mu < 2\lambda_{\max}^{-1}(\mathbf{C}_{\text{ss}}) \quad (39)$$

which is the same as the mean stability condition obtained in [30].

The mean square stability condition in (26), which is also the overall stability condition, becomes

$$\begin{aligned}0 < \mu < \min \left\{ \lambda_{\max}^{-1}(\bar{\mathbf{A}}^{-1}\bar{\mathbf{B}}), \right. \\ \left. \max \left\{ \lambda \left( \left[ \begin{array}{cc} \bar{\mathbf{A}}/2 & -\bar{\mathbf{B}}/2 \\ \mathbf{I}_{M^2} & \mathbf{0} \end{array} \right] \in \mathbb{R}^+ \right)^{-1} \right\} \right\} \\ \bar{\mathbf{A}} &\triangleq \mathbf{I}_M \otimes \mathbf{C}_{\text{ss}} + \mathbf{C}_{\text{ss}}^T \otimes \mathbf{I}_M \\ \bar{\mathbf{B}} &\triangleq E\{(\mathbf{s}_n\mathbf{s}_n^H)^T \otimes_{\mathbb{R}} (\mathbf{s}_n\mathbf{s}_n^H)\}\end{aligned}\quad (40)$$

Furthermore, if the regressor vector  $\mathbf{s}_n$  arises from a Gaussian distribution, the analysis results shown in (32), (35)–(37) can be simplified by expressing the fourth-order moment matrix  $E\{\|\mathbf{s}_n\|^2\mathbf{s}_n\mathbf{s}_n^H\}$  in terms of covariance and complementary covariance matrices of  $\mathbf{s}_n$  as

$$\begin{aligned}E\{\|\mathbf{s}_n\|^2\mathbf{s}_n\mathbf{s}_n^H\} \\ = \text{Tr}(\mathbf{C}_{\text{ss}})\mathbf{C}_{\text{ss}} + \frac{1}{2} \left( \mathbf{C}_{\text{ss}}^2 + \sum_{\alpha=\nu, j, \kappa} \mathbf{C}_{\text{ss}}^{\alpha} \mathbf{C}_{\text{ss}}^{\alpha H} \right)\end{aligned}\quad (41)$$

which is proved in the Appendix. Then, we have

$$\begin{aligned} \lambda_{\max} \left( E \left\{ \|\mathbf{s}_n\|^2 \mathbf{s}_n \mathbf{s}_n^H \right\} \right) &\leq \text{Tr}(\mathbf{C}_{\text{ss}}) \lambda_{\max}(\mathbf{C}_{\text{ss}}) \\ &+ \frac{1}{2} \left[ \lambda_{\max}^2(\mathbf{C}_{\text{ss}}) + \sum_{\alpha=\iota, j, \kappa} p_{\max}^2(\mathbf{C}_{\text{ss}^\alpha}) \right] \\ \lambda_{\min} \left( E \left\{ \|\mathbf{s}_n\|^2 \mathbf{s}_n \mathbf{s}_n^H \right\} \right) &\geq \text{Tr}(\mathbf{C}_{\text{ss}}) \lambda_{\min}(\mathbf{C}_{\text{ss}}) \\ &+ \frac{1}{2} \left[ \lambda_{\min}^2(\mathbf{C}_{\text{ss}}) + \sum_{\alpha=\iota, j, \kappa} p_{\min}^2(\mathbf{C}_{\text{ss}^\alpha}) \right] \end{aligned}$$

where  $p_{\max}(\mathbf{C}_{\text{ss}^\alpha})$  and  $p_{\min}(\mathbf{C}_{\text{ss}^\alpha})$  are maximum and minimum singular values of  $\mathbf{C}_{\text{ss}^\alpha}$  respectively. A sufficient stability condition is deduced from (32) as

$$0 < \mu <$$

$$\frac{4\lambda_{\min}(\mathbf{C}_{\text{ss}})}{2\text{Tr}(\mathbf{C}_{\text{ss}}) \lambda_{\max}(\mathbf{C}_{\text{ss}}) + \lambda_{\max}^2(\mathbf{C}_{\text{ss}}) + \sum_{\alpha=\iota, j, \kappa} p_{\max}^2(\mathbf{C}_{\text{ss}^\alpha})} \quad (42)$$

while a necessary stability condition is deduced from (35) as

$$0 < \mu <$$

$$\frac{4\lambda_{\max}(\mathbf{C}_{\text{ss}})}{2\text{Tr}(\mathbf{C}_{\text{ss}}) \lambda_{\min}(\mathbf{C}_{\text{ss}}) + \lambda_{\min}^2(\mathbf{C}_{\text{ss}}) + \sum_{\alpha=\iota, j, \kappa} p_{\min}^2(\mathbf{C}_{\text{ss}^\alpha})} \quad (43)$$

The upper and lower bounds of MSD are then deduced from (36), as in (44) shown at the bottom of this page. The upper and lower bounds of EMSE can be calculated based on the MSD bounds in (44), as in (37).

#### IV. MEAN SQUARE ANALYSIS OF QLMS WITH GAUSSIAN REGRESSORS USING QUATERNION DATA DECORRELATION

As shown in Section III-B and Section III-C, unlike the analysis for real and complex adaptive filters [25], [29], the mean square performance of quaternion adaptive filters is difficult to evaluate exactly in a closed form, owing to the non-commutative quaternion algebra and the lack of structural quaternion matrix decomposition methods. In order to obtain a concise closed-form solution to the mean square convergence equation (24) for QLMS with Gaussian regressors, we next use a novel quaternion data decorrelation technique to simplify the involved quaternion covariance matrices. The simplification is based on the following assumption on the augmented second-order statistics of quaternion Gaussian regressor vector  $\mathbf{s}_n$ .

*Assumption 1:* There exists a unitary transform  $\bar{\mathbf{s}}_n = \mathbf{G}^H \mathbf{s}_n$  such that

$$\mathbf{C}_{\bar{\text{ss}}} = \mathbf{\Lambda}_{\text{ss}}, \quad \mathbf{C}_{\bar{\text{ss}}^\alpha} = \mathbf{\Lambda}_{\text{ss}^\alpha}, \quad \alpha = \iota, j, \kappa \quad (45)$$

where  $\mathbf{\Lambda}_{\text{ss}}$  is a diagonal real matrix containing the eigenvalues of  $\mathbf{C}_{\text{ss}}$  on the diagonal, and  $\mathbf{\Lambda}_{\text{ss}^\alpha}$  are diagonal quaternion matrices.

It can be verified that Assumption 1 holds exactly when one of the following conditions on  $\mathbf{s}_n$  is met:

- 1) The covariance and complementary covariance matrices of  $\mathbf{s}_n$  are related by scalar multiplication. This condition holds for regressor vectors with specific second-order statistics, such as an  $\mathbb{H}$ -proper regressor vector ( $\mathbf{C}_{\text{ss}^\alpha} = 0\mathbf{C}_{\text{ss}}$ ) and a rectilinear regressor vector ( $\mathbf{C}_{\text{ss}^\alpha} = \rho_\alpha \mathbf{C}_{\text{ss}}$ , where  $\rho_\alpha$  is a unit quaternion).
- 2)  $\mathbf{s}_n$  can be expressed by applying a unitary transform to a white Gaussian random quaternion vector  $\mathbf{v}_n$ , that is,  $\mathbf{s}_n = \mathbf{G}\mathbf{v}_n$ , where  $\mathbf{G}$  is a unitary quaternion matrix.

For quaternion inputs that do not meet the above two conditions, Assumption 1 holds approximately as the result of the recently proposed approximate uncorrelating transform [48]–[50]. Under Assumption 1, we shall define  $\bar{\mathbf{w}}_n = \mathbf{G}^H \tilde{\mathbf{w}}_n$ ,  $\bar{\mathbf{q}}_n = \mathbf{G}^H \mathbf{q}_n$ ,  $\bar{\Sigma} = \mathbf{G}^H \Sigma \mathbf{G}$  and convert (24) into

$$\begin{aligned} E \left\{ \|\bar{\mathbf{w}}_{n+1}\|_{\bar{\Sigma}}^2 \right\} &= E \left\{ \|\bar{\mathbf{w}}_n\|_{\bar{\Sigma}'}^2 \right\} + \mu^2 \sigma_v^2 E \left\{ \|\bar{\mathbf{s}}_n\|_{\bar{\Sigma}}^2 \right\} \\ &+ E \left\{ \|\bar{\mathbf{q}}_n\|_{\bar{\Sigma}}^2 \right\} \\ \bar{\Sigma}' &= \bar{\Sigma} - \mu \left( E \left\{ \bar{\mathbf{s}}_n \bar{\mathbf{s}}_n^H \right\} \bar{\Sigma} + \bar{\Sigma} E \left\{ \bar{\mathbf{s}}_n \bar{\mathbf{s}}_n^H \right\} \right) \\ &+ \mu^2 E \left\{ \|\bar{\mathbf{s}}_n\|_{\bar{\Sigma}}^2 \bar{\mathbf{s}}_n \bar{\mathbf{s}}_n^H \right\} \end{aligned} \quad (46)$$

where

$$\begin{aligned} E \left\{ \|\bar{\mathbf{s}}_n\|_{\bar{\Sigma}}^2 \bar{\mathbf{s}}_n \bar{\mathbf{s}}_n^H \right\} &= \text{Tr}(\bar{\Sigma} \mathbf{\Lambda}_{\text{ss}}) \mathbf{\Lambda}_{\text{ss}} \\ &+ \frac{1}{2} \bar{\Sigma} \left( \mathbf{\Lambda}_{\text{ss}}^2 + \sum_{\alpha=\iota, j, \kappa} \mathbf{\Lambda}_{\text{ss}^\alpha} \mathbf{\Lambda}_{\text{ss}^\alpha}^* \right) \end{aligned} \quad (47)$$

is derived similarly to the derivation of (41). Let  $\bar{\Sigma}$  be a diagonal real matrix, then  $\bar{\Sigma}'$  is also diagonal and real, and hence these two matrices can be completely characterised by their diagonal entries. Defining  $\bar{\sigma} \triangleq \text{diag}(\bar{\Sigma})$ ,  $\bar{\sigma}' \triangleq \text{diag}(\bar{\Sigma}')$ ,  $\lambda_s \triangleq \text{diag}(\mathbf{\Lambda}_{\text{ss}})$ ,  $\lambda_q \triangleq \text{diag}(E\{\bar{\mathbf{q}}_n \bar{\mathbf{q}}_n^H\})$ , and referring to  $\|\cdot\|_{\bar{\Sigma}}^2$  and  $\|\cdot\|_{\bar{\Sigma}'}^2$  as  $\|\cdot\|_{\bar{\sigma}}^2$  and  $\|\cdot\|_{\bar{\sigma}'}^2$ , we obtain the compact forms of (46) and (47) as

$$\begin{aligned} E \left\{ \|\bar{\mathbf{w}}_{n+1}\|_{\bar{\sigma}}^2 \right\} &= E \left\{ \|\bar{\mathbf{w}}_n\|_{\bar{\sigma}'}^2 \right\} + (\mu^2 \sigma_v^2 \lambda_s + \lambda_q)^T \bar{\sigma} \\ \mathbf{F} &= \mathbf{I}_M - \mu \mathbf{A} + \mu^2 \mathbf{B} \\ \mathbf{A} &= 2\mathbf{\Lambda}_{\text{ss}} \\ \mathbf{B} &= \frac{1}{2} \left( \mathbf{\Lambda}_{\text{ss}}^2 + \sum_{\alpha=\iota, j, \kappa} \mathbf{\Lambda}_{\text{ss}^\alpha} \mathbf{\Lambda}_{\text{ss}^\alpha}^* \right) + \lambda_s \lambda_s^T \end{aligned} \quad (48)$$

Similar to the stability condition for (26), (48) is stable if and only if all eigenvalues of  $\mathbf{F}$  are within  $(-1, 1)$ . Since these

$$\begin{aligned} \text{MSD} &\leq \frac{2\mu^2 \sigma_v^2 \text{Tr}(\mathbf{C}_{\text{ss}}) + 2\text{Tr}(\mathbf{C}_{\text{qq}})}{4\mu \lambda_{\min}(\mathbf{C}_{\text{ss}}) - \mu^2 \left[ 2\text{Tr}(\mathbf{C}_{\text{ss}}) \lambda_{\max}(\mathbf{C}_{\text{ss}}) + \lambda_{\max}^2(\mathbf{C}_{\text{ss}}) + \sum_{\alpha=\iota, j, \kappa} p_{\max}^2(\mathbf{C}_{\text{ss}^\alpha}) \right]} \\ \text{MSD} &\geq \frac{2\mu^2 \sigma_v^2 \text{Tr}(\mathbf{C}_{\text{ss}}) + 2\text{Tr}(\mathbf{C}_{\text{qq}})}{4\mu \lambda_{\max}(\mathbf{C}_{\text{ss}}) - \mu^2 \left[ 2\text{Tr}(\mathbf{C}_{\text{ss}}) \lambda_{\min}(\mathbf{C}_{\text{ss}}) + \lambda_{\min}^2(\mathbf{C}_{\text{ss}}) + \sum_{\alpha=\iota, j, \kappa} p_{\min}^2(\mathbf{C}_{\text{ss}^\alpha}) \right]} \end{aligned} \quad (44)$$

eigenvalues are easily seen to be greater than  $-1$ , we only need to investigate how to guarantee that they are less than unit. It can be proved that they are upper bounded by unit if and only if  $0 < \mu < \lambda_{\max}^{-1}(\mathbf{A}^{-1}\mathbf{B})$ , which is verified to be equivalent to

$$0 < \sum_{m=1}^M \frac{2\mu\lambda_{s,m}}{4 - \mu\lambda_{s,m} \left(1 + \sum_{\alpha=i,j,\kappa} r_{\bar{s}\alpha,m}^2\right)} < 1 \quad (49)$$

where  $\lambda_{s,m}$  is the  $m$ -th element of  $\lambda_s$ , and  $r_{\bar{s}\alpha,m}$  is the  $\alpha$ -impropriety coefficient of the  $m$ -th element of  $\bar{s}_n$ . According to Remark 1, the condition in (49) is sufficient and necessary for the algorithm stability both in the mean and mean square sense.

*Remark 2:* The stability condition in (49) indicates that the upper bound of  $\mu$  decreases with the increase in the length of the regressor vector,  $M$ , the increase in the power of the regressor reflected by  $\lambda_{s,m}$ , and the increase in the impropriety degree of  $\bar{s}_n$  reflected by  $r_{\bar{s}i,m}, r_{\bar{s}j,m}, r_{\bar{s}\kappa,m}$ .

MSD and EMSE analysis. Iterating (48) gives

$$E \left\{ \|\bar{\mathbf{w}}_n\|_{\bar{\boldsymbol{\sigma}}}^2 \right\} = E \left\{ \|\bar{\mathbf{w}}_0\|_{\mathbf{F}^n \bar{\boldsymbol{\sigma}}}^2 \right\} + (\mu^2 \sigma_v^2 \lambda_s + \lambda_q)^T \sum_{l=0}^{n-1} \mathbf{F}^l \bar{\boldsymbol{\sigma}} \quad (50)$$

The MSD and EMSE at iteration  $n$  can now be derived from (50) by letting  $\bar{\boldsymbol{\sigma}} = \mathbf{1}_M$  and  $\bar{\boldsymbol{\sigma}} = \lambda_s$ , respectively. Equation (50) also yields the iteration

$$E \left\{ \|\bar{\mathbf{w}}_{n+1}\|_{\bar{\boldsymbol{\sigma}}}^2 \right\} = E \left\{ \|\bar{\mathbf{w}}_n\|_{\bar{\boldsymbol{\sigma}}}^2 \right\} + E \left\{ \|\bar{\mathbf{w}}_0\|_{\mathbf{F}^n (\mathbf{F}-\mathbf{I}_M) \bar{\boldsymbol{\sigma}}}^2 \right\} + (\mu^2 \sigma_v^2 \lambda_s + \lambda_q)^T \mathbf{F}^n \bar{\boldsymbol{\sigma}} \quad (51)$$

which characterises the evolution of MSD when  $\bar{\boldsymbol{\sigma}} = \mathbf{1}_M$ , and characterises the evolution of EMSE when  $\bar{\boldsymbol{\sigma}} = \lambda_s$ .

At the steady state,  $\lim_{n \rightarrow \infty} E \left\{ \|\bar{\mathbf{w}}_{n+1}\|_{\bar{\boldsymbol{\sigma}}}^2 \right\} = \lim_{n \rightarrow \infty} E \left\{ \|\bar{\mathbf{w}}_n\|_{\bar{\boldsymbol{\sigma}}}^2 \right\}$ , so that (48) yields

$$\lim_{n \rightarrow \infty} E \left\{ \|\bar{\mathbf{w}}_n\|_{(\mu\mathbf{A} - \mu^2\mathbf{B})\bar{\boldsymbol{\sigma}}}^2 \right\} = (\mu^2 \sigma_v^2 \lambda_s + \lambda_q)^T \bar{\boldsymbol{\sigma}}$$

Finally, the steady-state MSD and EMSE become

$$\begin{aligned} \text{MSD} &= \lim_{n \rightarrow \infty} E \left\{ \|\bar{\mathbf{w}}_n\|_{\mathbf{1}_M}^2 \right\} \\ &= (\mu^2 \sigma_v^2 \lambda_s + \lambda_q)^T (\mu\mathbf{A} - \mu^2\mathbf{B})^{-1} \mathbf{1}_M \end{aligned}$$

$$\begin{aligned} \text{EMSE} &= \lim_{n \rightarrow \infty} E \left\{ \|\bar{\mathbf{w}}_n\|_{\lambda_s}^2 \right\} \\ &= (\mu^2 \sigma_v^2 \lambda_s + \lambda_q)^T (\mu\mathbf{A} - \mu^2\mathbf{B})^{-1} \lambda_s \end{aligned}$$

which can be rewritten as

$$\text{MSD} = \frac{2 \sum_{m=1}^M \frac{\mu^2 \sigma_v^2 + \lambda_{q,m} \lambda_{s,m}^{-1}}{4 - \mu\lambda_{s,m} (1 + \sum_{\alpha=i,j,\kappa} r_{\bar{s}\alpha,m}^2)}}{\mu - 2\mu^2 \sum_{m=1}^M \frac{\lambda_{s,m}}{4 - \mu\lambda_{s,m} (1 + \sum_{\alpha=i,j,\kappa} r_{\bar{s}\alpha,m}^2)}} \quad (52)$$

$$\text{EMSE} = \frac{2 \sum_{m=1}^M \frac{\mu^2 \sigma_v^2 \lambda_{s,m} + \lambda_{q,m}}{4 - \mu\lambda_{s,m} (1 + \sum_{\alpha=i,j,\kappa} r_{\bar{s}\alpha,m}^2)}}{\mu - 2\mu^2 \sum_{m=1}^M \frac{\lambda_{s,m}}{4 - \mu\lambda_{s,m} (1 + \sum_{\alpha=i,j,\kappa} r_{\bar{s}\alpha,m}^2)}} \quad (53)$$

*Remark 3:* Observe from (52) and (53) that the steady-state EMSE and MSD increase with the length of the regressor vector,

TABLE I  
SUMMARY OF ANALYSIS RESULTS

	Stability condition			Steady state	
	S	N	N&S	MSD	EMSE
General quaternion adaptive filter	Eq.(32)	Eq.(35)	Eq.(26)	Eq.(36) (bounds)	Eq.(37) (bounds)
QLMS with Gaussian inputs	Eq.(42)	Eq.(43)	Eq.(40)	Eq.(44) (bounds)	Eq.(37) (bounds)
QLMS with Gaussian inputs under Assumption 1	Eq.(42)	Eq.(43)	Eq.(49)	Eq.(52)	Eq.(53)

S: Sufficient; N: Necessary.

$M$ , the power of the regressor reflected by  $\lambda_{s,m}$ , the impropriety degree of  $\bar{s}_n$  reflected by  $r_{\bar{s}i,m}, r_{\bar{s}j,m}, r_{\bar{s}\kappa,m}$ , and the variance of the system weight variation reflected by  $\lambda_{q,m}$ .

Table I summarises the sufficient stability conditions, necessary stability conditions, necessary and sufficient stability conditions, MSDs, and EMSEs for general quaternion adaptive filters and QLMS filters with Gaussian inputs. The closed-form representations of the performance of SL-QLMS, SWL-QLMS, and WL-QLMS with Gaussian regressors satisfying Assumption 1 can be obtained by replacing  $\mathbf{s}$  with  $\mathbf{x}_n, \dot{\mathbf{x}}_n$ , and  $\ddot{\mathbf{x}}_n$  respectively in (49), (52) and (53). Exploring the second-order statistics of input signal, the following three subsections further address the performance of the three classes of QLMS algorithms.

#### A. SL-QLMS

1) *Scalar Input:* The stability condition of the SL-QLMS with a scalar regressor  $x_n$  is obtained from (49) as

$$0 < \mu < \frac{4}{\sigma_x^2 (3 + r_{x_i}^2 + r_{x_j}^2 + r_{x_\kappa}^2)} \quad (54)$$

where  $\sigma_x^2$  is the variance of  $x_n$ . This condition is the same as the analysis result obtained in [31]. The steady-state performance can be derived from (52) and (53) as

$$\text{MSD} = \frac{2 (\mu^2 \sigma_v^2 \sigma_x^2 + \sigma_q^2)}{\mu \sigma_x^2 [4 - \mu \sigma_x^2 (3 + r_{x_i}^2 + r_{x_j}^2 + r_{x_\kappa}^2)]} \quad (55)$$

$$\text{EMSE} = \frac{2 (\mu^2 \sigma_v^2 \sigma_x^2 + \sigma_q^2)}{\mu [4 - \mu \sigma_x^2 (3 + r_{x_i}^2 + r_{x_j}^2 + r_{x_\kappa}^2)]} \quad (56)$$

where  $\sigma_q^2$  is the variance of the weight variation  $q_n$ .

2)  *$\mathbb{H}$ -Proper Input Vector:* The diagonalisation of  $\mathbf{C}_{\mathbf{x}\mathbf{x}}$  completely decorrelates an  $\mathbb{H}$ -proper quaternion input vector  $\mathbf{x}_n$ . For such inputs,  $r_{\bar{x}\alpha,m} = 0, \forall m, \alpha$ , so that the stability condition in (49) and the steady-state performance in (52) and (53) become

$$0 < \sum_{m=1}^M \frac{2\mu\lambda_{x,m}}{4 - \mu\lambda_{x,m}} < 1 \quad (57)$$



$$\text{MSD} = \frac{2\sum_{m=1}^M \frac{\mu^2 \sigma_v^2 + \lambda_{q,m} \lambda_{x,m}^{-1}}{4 - \mu \lambda_{x,m}}}{\mu \left(1 - \mu \sum_{m=1}^M \frac{2\lambda_{x,m}}{4 - \mu \lambda_{x,m}}\right)} \quad (58)$$

$$\text{EMSE} = \frac{2\sum_{m=1}^M \frac{\mu^2 \sigma_v^2 \lambda_{x,m} + \lambda_{q,m}}{4 - \mu \lambda_{x,m}}}{\mu \left(1 - \mu \sum_{m=1}^M \frac{2\lambda_{x,m}}{4 - \mu \lambda_{x,m}}\right)} \quad (59)$$

3) *Rectilinear Input Vector*: If the Gaussian input vector  $\mathbf{x}_n$  is rectilinear, it can be expressed as  $\mathbf{x}_n = e^\zeta \mathbf{z}_n$ , where  $\mathbf{z}_n$  is a random Gaussian real vector, and  $\zeta$  a constant pure quaternion scalar [51]. For such inputs, the covariance and complementary covariance matrices can be represented by  $\mathbf{C}_{\mathbf{x}\mathbf{x}} = \mathbf{C}_{\mathbf{z}\mathbf{z}}$ ,  $\mathbf{C}_{\mathbf{x}\mathbf{x}^\alpha} = e^\zeta (e^{-\zeta})^\alpha \mathbf{C}_{\mathbf{z}\mathbf{z}}$ ,  $\alpha = i, j, \kappa$ . Therefore,  $\mathbf{C}_{\mathbf{x}\mathbf{x}}$ ,  $\mathbf{C}_{\mathbf{x}\mathbf{x}^\alpha}$  can be simultaneously diagonalised via the diagonalisation of the real symmetric matrix  $\mathbf{C}_{\mathbf{z}\mathbf{z}}$ . Note  $r_{\bar{\mathbf{x}}\alpha, m} = 1, \forall m, \alpha$ , so that the stability condition in (49) and the steady-state performance in (52) and (53) become

$$0 < \sum_{m=1}^M \frac{\mu \lambda_{x,m}}{2 - 2\mu \lambda_{x,m}} < 1 \quad (60)$$

$$\text{MSD} = \frac{\sum_{m=1}^M \frac{\mu^2 \sigma_v^2 + \lambda_{q,m} \lambda_{x,m}^{-1}}{2 - 2\mu \lambda_{x,m}}}{\mu \left(1 - \mu \sum_{m=1}^M \frac{\lambda_{x,m}}{2 - 2\mu \lambda_{x,m}}\right)} \quad (61)$$

$$\text{EMSE} = \frac{\sum_{m=1}^M \frac{\mu^2 \sigma_v^2 \lambda_{x,m} + \lambda_{q,m}}{2 - 2\mu \lambda_{x,m}}}{\mu \left(1 - \mu \sum_{m=1}^M \frac{\lambda_{x,m}}{2 - 2\mu \lambda_{x,m}}\right)} \quad (62)$$

## B. SWL-QLMS

1) *H-Propper Input Vector*: If the random quaternion vector  $\mathbf{x}_n$  is H-proper,  $\mathbf{C}_{\bar{\mathbf{x}}\bar{\mathbf{x}}}$  reduces to a block diagonal matrix. Denote the eigendecomposition of  $\mathbf{C}_{\mathbf{x}\mathbf{x}}$  by  $\mathbf{C}_{\mathbf{x}\mathbf{x}} = \mathbf{Q}\mathbf{\Lambda}_{\mathbf{x}\mathbf{x}}\mathbf{Q}^H$ , where  $\mathbf{\Lambda}_{\mathbf{x}\mathbf{x}}$  is a diagonal real matrix containing all eigenvalues of  $\mathbf{C}_{\mathbf{x}\mathbf{x}}$  on the diagonal, and  $\mathbf{Q}$  is a unitary quaternion matrix containing eigenvectors of  $\mathbf{C}_{\mathbf{x}\mathbf{x}}$ . We can find a unitary quaternion matrix

$$\mathbf{G} = \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{Q} & -\mathbf{Q} \\ \mathbf{Q}^\eta & \mathbf{Q}^\eta \end{bmatrix}$$

so that  $\bar{\mathbf{s}}_n = \mathbf{G}^H \bar{\mathbf{x}}_n$  is a  $\mathcal{C}^\eta$ -improper vector and satisfies

$$\mathbf{C}_{\bar{\mathbf{s}}\bar{\mathbf{s}}} = \mathbf{I}_2 \otimes \mathbf{\Lambda}_{\mathbf{x}\mathbf{x}}, \quad \mathbf{C}_{\bar{\mathbf{s}}\bar{\mathbf{s}}^\eta} = \text{Diag}(1, -1) \otimes \mathbf{\Lambda}_{\mathbf{x}\mathbf{x}}$$

This reduces the stability condition in (49) and the steady-state performance in (52) and (53) to

$$0 < \sum_{m=1}^{M/2} \frac{2\mu \lambda_{x,m}}{2 - \mu \lambda_{x,m}} < 1 \quad (63)$$

$$\text{MSD} = \frac{\sum_{m=1}^{M/2} \frac{2\mu^2 \sigma_v^2 \lambda_{x,m} + \lambda_{q,m} + \lambda_{q,m} + M/2}{\lambda_{x,m} (2 - \mu \lambda_{x,m})}}{\mu \left(1 - 2\mu \sum_{m=1}^{M/2} \frac{\lambda_{x,m}}{2 - \mu \lambda_{x,m}}\right)} \quad (64)$$

$$\text{EMSE} = \frac{\sum_{m=1}^{M/2} \frac{2\mu^2 \sigma_v^2 \lambda_{x,m} + \lambda_{q,m} + \lambda_{q,m} + M/2}{2 - \mu \lambda_{x,m}}}{\mu \left(1 - 2\mu \sum_{m=1}^{M/2} \frac{\lambda_{x,m}}{2 - \mu \lambda_{x,m}}\right)} \quad (65)$$

2) *Rectilinear Input Vector*: As shown in Section IV-A3, the covariance matrix of a rectilinear vector  $\mathbf{x}_n$  is real-valued, so its eigendecomposition can be represented by  $\mathbf{C}_{\mathbf{x}\mathbf{x}} = \mathbf{Q}\mathbf{\Lambda}_{\mathbf{x}\mathbf{x}}\mathbf{Q}^T$ ,

where  $\mathbf{Q}$  is a real orthogonal matrix. We can find a quaternion unitary matrix

$$\mathbf{G} = \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{Q} & -e^\zeta (e^{-\zeta})^\eta \mathbf{Q} \\ (e^\zeta)^\eta e^{-\zeta} \mathbf{Q} & \mathbf{Q} \end{bmatrix}$$

so that  $\bar{\mathbf{s}}_n = \mathbf{G}^H \bar{\mathbf{x}}_n$  satisfies

$$\mathbf{C}_{\bar{\mathbf{s}}\bar{\mathbf{s}}} = \text{Diag}(2, 0) \otimes \mathbf{\Lambda}_{\mathbf{x}\mathbf{x}}$$

$$\mathbf{C}_{\bar{\mathbf{s}}\bar{\mathbf{s}}^\eta} = \text{Diag}(2e^\zeta (e^{-\zeta})^\alpha, 0) \otimes \mathbf{\Lambda}_{\mathbf{x}\mathbf{x}}, \quad \alpha = i, j, \kappa$$

which reduces the stability condition in (49) and the steady-state EMSE in (53) to

$$0 < \sum_{m=1}^{M/2} \frac{\mu \lambda_{x,m}}{1 - 2\mu \lambda_{x,m}} < 1 \quad (66)$$

$$\text{EMSE} = \frac{\sum_{m=1}^{M/2} \frac{2\mu^2 \sigma_v^2 \lambda_{x,m} + \lambda_{q,m}}{1 - 2\mu \lambda_{x,m}} + \sum_{m=1+M/2}^M \lambda_{q,m}}{2\mu \left(1 - \mu \sum_{m=1}^{M/2} \frac{\lambda_{x,m}}{1 - 2\mu \lambda_{x,m}}\right)} \quad (67)$$

In this case, the steady-state MSD is uncertain as the weight solution is not unique.

## C. WL-QLMS

1) *H-proper Input Vector*: If  $\mathbf{x}_n$  is H-proper,  $\mathbf{C}_{\bar{\mathbf{x}}\bar{\mathbf{x}}}$  reduces to a block diagonal matrix. Considering the eigendecomposition of  $\mathbf{C}_{\mathbf{x}\mathbf{x}}$ ,  $\mathbf{C}_{\mathbf{x}\mathbf{x}} = \mathbf{Q}\mathbf{\Lambda}_{\mathbf{x}\mathbf{x}}\mathbf{Q}^H$ , we can find such a quaternion unitary matrix given by

$$\mathbf{G} = \frac{1}{2} \begin{bmatrix} \mathbf{Q} & -\mathbf{Q} & -\mathbf{Q} & -\mathbf{Q} \\ \mathbf{Q}^i & -\mathbf{Q}^i & \mathbf{Q}^j & \mathbf{Q}^j \\ \mathbf{Q}^j & \mathbf{Q}^j & -\mathbf{Q}^j & \mathbf{Q}^j \\ \mathbf{Q}^\kappa & \mathbf{Q}^\kappa & \mathbf{Q}^\kappa & -\mathbf{Q}^\kappa \end{bmatrix}$$

that  $\bar{\mathbf{s}}_n = \mathbf{G}^H \bar{\mathbf{x}}_n$  satisfies

$$\mathbf{C}_{\bar{\mathbf{s}}\bar{\mathbf{s}}} = \mathbf{I}_4 \otimes \mathbf{\Lambda}_{\mathbf{x}\mathbf{x}}$$

$$\mathbf{C}_{\bar{\mathbf{s}}\bar{\mathbf{s}}^i} = \text{Diag}(1, 1, -1, -1) \otimes \mathbf{\Lambda}_{\mathbf{x}\mathbf{x}}$$

$$\mathbf{C}_{\bar{\mathbf{s}}\bar{\mathbf{s}}^j} = \text{Diag}(1, -1, 1, -1) \otimes \mathbf{\Lambda}_{\mathbf{x}\mathbf{x}}$$

$$\mathbf{C}_{\bar{\mathbf{s}}\bar{\mathbf{s}}^\kappa} = \text{Diag}(1, -1, -1, 1) \otimes \mathbf{\Lambda}_{\mathbf{x}\mathbf{x}}$$

The stability condition in (49) and the steady-state performance in (52) and (53) become

$$0 < \sum_{m=1}^{M/4} \frac{2\mu \lambda_{x,m}}{1 - \mu \lambda_{x,m}} < 1 \quad (68)$$

MSD =

$$\frac{\sum_{m=1}^{M/4} \frac{4\mu^2 \sigma_v^2 \lambda_{x,m} + \lambda_{q,m} + \lambda_{q,m} + M/4 + \lambda_{q,m} + M/2 + \lambda_{q,m} + 3M/4}{2\lambda_{x,m} (1 - \mu \lambda_{x,m})}}{\mu \left(1 - 2\mu \sum_{m=1}^{M/4} \frac{\lambda_{x,m}}{1 - \mu \lambda_{x,m}}\right)} \quad (69)$$

EMSE =

$$\frac{\sum_{m=1}^{M/4} \frac{4\mu^2 \sigma_v^2 \lambda_{x,m} + \lambda_{q,m} + \lambda_{q,m} + M/4 + \lambda_{q,m} + M/2 + \lambda_{q,m} + 3M/4}{2(1 - \mu \lambda_{x,m})}}{\mu \left(1 - 2\mu \sum_{m=1}^{M/4} \frac{\lambda_{x,m}}{1 - \mu \lambda_{x,m}}\right)} \quad (70)$$

2) *Rectilinear input vector*: If  $\mathbf{x}_n$  is rectilinear, the quaternion unitary matrix

$$\mathbf{G} = \frac{1}{2} \begin{bmatrix} \mathbf{Q} & -\mathbf{Q} & -\mathbf{Q} & -\mathbf{Q} \\ (e^\zeta)^i e^{-\zeta} \mathbf{Q} & -(e^\zeta)^i e^{-\zeta} \mathbf{Q} & (e^\zeta)^i e^{-\zeta} \mathbf{Q} & (e^\zeta)^i e^{-\zeta} \mathbf{Q} \\ (e^\zeta)^j e^{-\zeta} \mathbf{Q} & (e^\zeta)^j e^{-\zeta} \mathbf{Q} & -(e^\zeta)^j e^{-\zeta} \mathbf{Q} & (e^\zeta)^j e^{-\zeta} \mathbf{Q} \\ (e^\zeta)^k e^{-\zeta} \mathbf{Q} & (e^\zeta)^k e^{-\zeta} \mathbf{Q} & (e^\zeta)^k e^{-\zeta} \mathbf{Q} & -(e^\zeta)^k e^{-\zeta} \mathbf{Q} \end{bmatrix}$$

allows the transformed vector  $\bar{\mathbf{s}}_n = \mathbf{G}^H \ddot{\mathbf{x}}_n$  to satisfy

$$\mathbf{C}_{\bar{\mathbf{s}}\bar{\mathbf{s}}} = \text{Diag}(4, 0, 0, 0) \otimes \mathbf{\Lambda}_{\mathbf{x}\mathbf{x}}$$

$$\mathbf{C}_{\bar{\mathbf{s}}\bar{\mathbf{s}}^\alpha} = \text{Diag}(4e^\zeta (e^{-\zeta})^\alpha, 0, 0, 0) \otimes \mathbf{\Lambda}_{\mathbf{x}\mathbf{x}}, \quad \alpha = i, j, k$$

which reduces the stability condition in (49) and the steady-state EMSE in (53) to

$$0 < \sum_{m=1}^{M/4} \frac{2\mu\lambda_{\mathbf{x},m}}{1-4\mu\lambda_{\mathbf{x},m}} < 1 \quad (71)$$

$$\text{EMSE} = \frac{\sum_{m=1}^{M/4} \frac{4\mu^2\sigma_v^2\lambda_{\mathbf{x},m} + \lambda_{\mathbf{q},m}}{1-4\mu\lambda_{\mathbf{x},m}} + \sum_{m=1+M/4}^M \lambda_{\mathbf{q},m}}{2\mu \left(1 - 2\mu \sum_{m=1}^{M/4} \frac{\lambda_{\mathbf{x},m}}{1-4\mu\lambda_{\mathbf{x},m}}\right)} \quad (72)$$

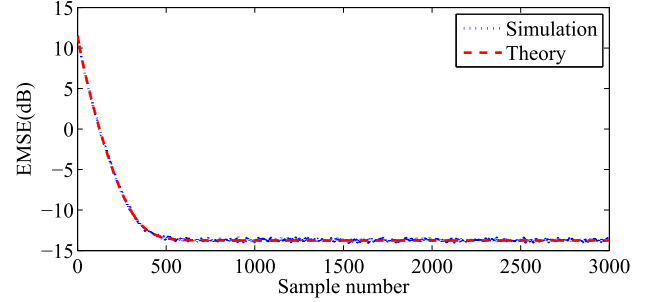
In this case, the steady-state MSD is uncertain as the weight solution is not unique.

## V. SIMULATIONS

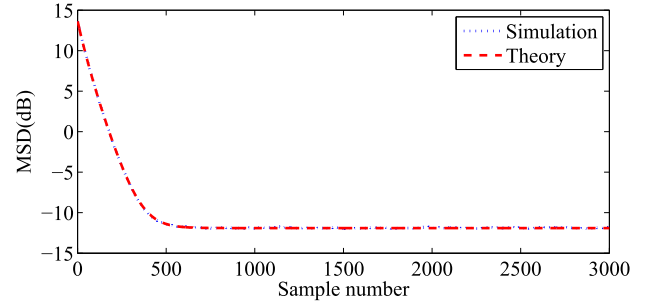
The performances of the three considered classes of QLMS algorithms with Gaussian inputs were evaluated by averaging results over 1000 independent numerical simulation trials. Throughout the simulations, the length of the synthetic correlated input vector  $\mathbf{x}_n$  was  $L = 4$ , and the eigenvalues of  $\mathbf{C}_{\mathbf{x}\mathbf{x}}$  were kept the same. For simplicity, the elements of  $\mathbf{x}_n$  were set to have an equal impropriety coefficient, that is,  $r_{m,i} = r_{m,j} = r_{m,k} = r$ .

### A. SL-QLMS

In the simulations for the SL-QLMS, the desired signal was generated from the strictly linear model in (9) and (12). The variance of the additive Gaussian noise  $v_n$  was  $\sigma_v^2 = 0.01$ , and the Euclidean norm of the Gaussian weight variation  $\mathbf{q}_n$  was 0.0016. Fig. 1 depicts the theoretical and simulated EMSE and MSD learning curves of the SL-QLMS for an improper input with  $r = 0.5$  and  $\mu = 0.02$ . These curves exhibit a good match between the theory and simulation results. Figs. 2 and 3 illustrate the theoretical and simulated steady-state EMSEs for different values of the step size  $\mu$  when  $\mathbf{x}_n$  is  $\mathbb{H}$ -proper and when  $\mathbf{x}_n$  is improper with  $r = 0.5$ , respectively. The theoretical bounds of  $\mu$  are also indicated. Both figures show that there was a good match between theory and simulations and their difference increased with  $\mu$ , which is similar to the behaviour of real-valued LMS filters [27]. Observe that the decrease in the stability bound for the improper input conforms with Remark 2. Fig. 8 illustrates a match between the theoretical and simulated steady-state EMSEs for inputs with various impropriety coefficients  $r \in [0, 1]$  and  $\mu = 0.2$ . The increase in the steady-state EMSEs with  $r$  conforms with Remark 3. As Remark 2, Figs. 2 and 3 show, the upper bound of  $\mu$  decreases with the increase in  $r$ .



(a) EMSE curves



(b) MSD curves

Fig. 1. Theoretical and simulated learning curves of SL-QLMS for an improper input with  $r = 0.5$  and  $\mu = 0.02$ .

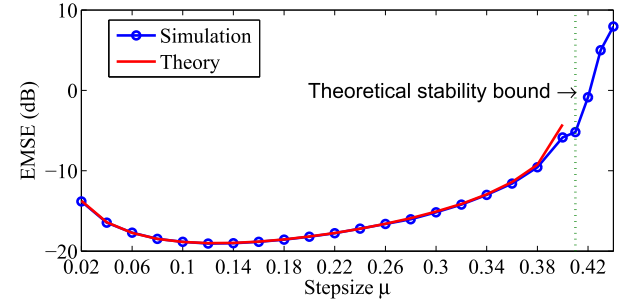


Fig. 2. Theoretical and simulated EMSEs of SL-QLMS as a function of the step size, for an  $\mathbb{H}$ -proper input.

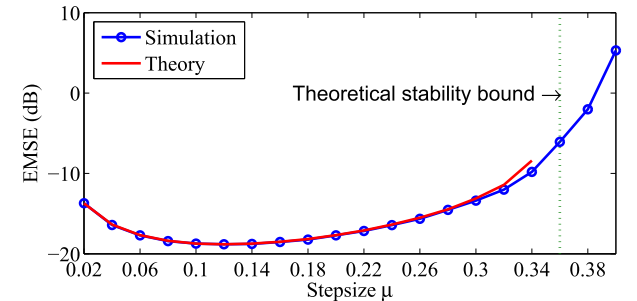


Fig. 3. Theoretical and simulated EMSEs of SL-QLMS as a function of the step size, for an improper input with  $r = 0.5$ .

So, for a fixed  $\mu$ , the theoretical EMSE for highly improper inputs is less accurate, as  $\mu$  is closer to the upper bound. This explains the difference between the theoretical and simulated steady-state EMSEs for a large  $r$  in Fig. 4 as well as in Figs. 8 and 12.

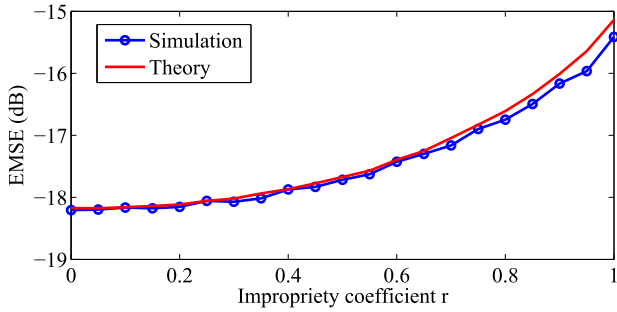


Fig. 4. Theoretical and simulated EMSEs of SL-QLMS as a function of the impropriety coefficient, for  $\mu = 0.2$ .

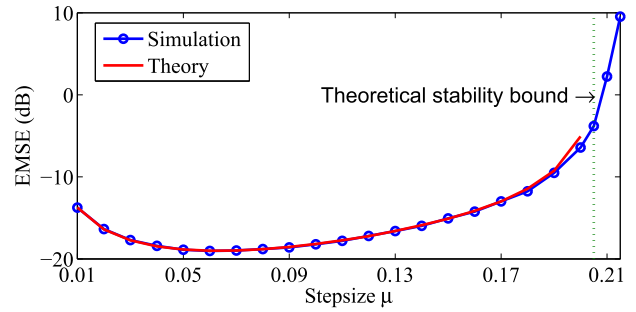


Fig. 6. Theoretical and simulated EMSEs of SWL-QLMS as a function of the step size, for an  $\mathbb{H}$ -proper input.

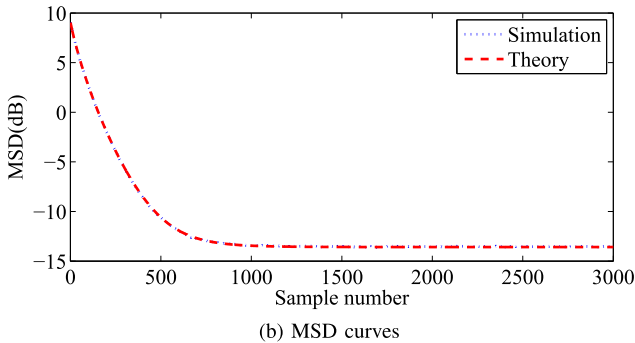
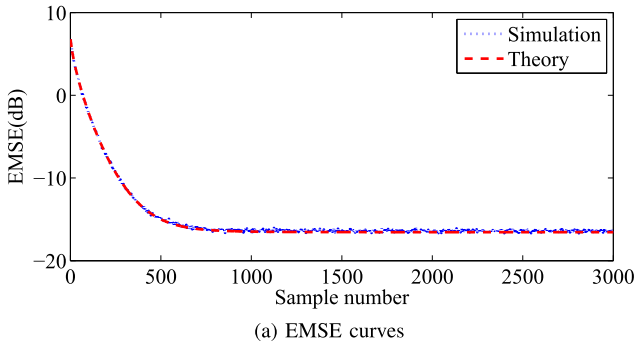


Fig. 5. Theoretical and simulated learning curves of SWL-QLMS for an improper input with  $r = 0.5$  and  $\mu = 0.02$ .

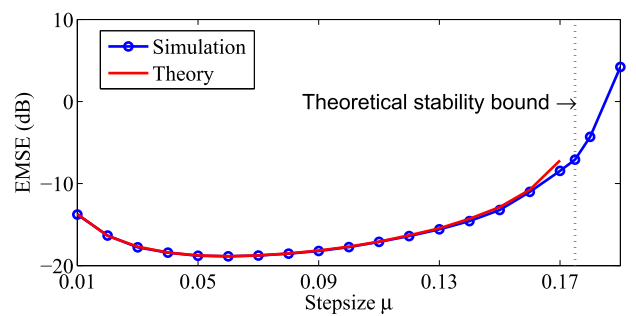


Fig. 7. Theoretical and simulated EMSEs of SWL-QLMS as a function of the step size, for an improper input with  $r = 0.5$ .

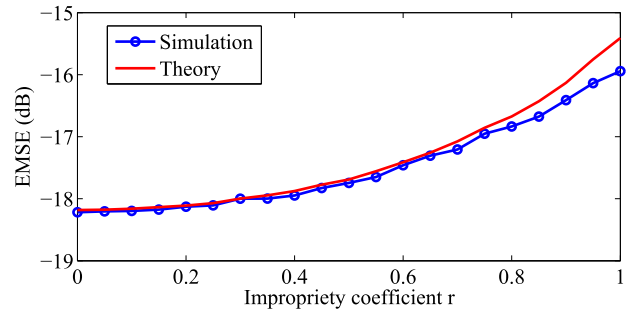


Fig. 8. Theoretical and simulated EMSEs of SWL-QLMS as a function of the impropriety coefficient, for  $\mu = 0.1$ .

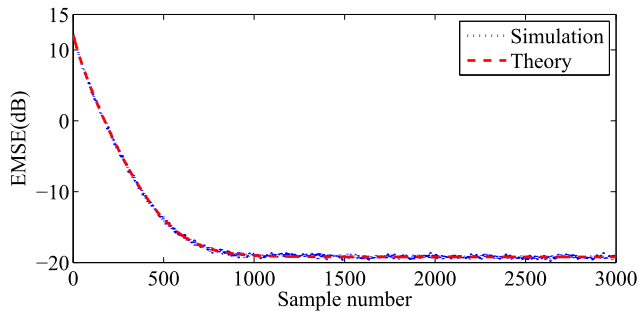
**B. SWL-QLMS**

In the simulations for the SWL-QLMS, the desired signal was generated from the semi-widely linear model in (10) and (12). The variance of the additive Gaussian noise  $v_n$  was  $\sigma_v^2 = 0.01$ , and the Euclidean norm of the Gaussian weight variation  $\mathbf{q}_n$  was 0.0008. Fig. 5 depicts theoretical and simulated EMSE and MSD learning curves of the SL-QLMS for an improper input with  $r = 0.5$  and  $\mu = 0.02$ . These curves exhibit a good match between the theory and simulation results. Figs. 6 and 7 show theoretical and simulated steady-state EMSEs for different values of the step size  $\mu$  when  $\mathbf{x}_n$  is  $\mathbb{H}$ -proper and when  $\mathbf{x}_n$  is improper with  $r = 0.5$ . The theoretical bounds of  $\mu$  are also indicated. Both figures illustrate a good match between theory and simulations and an increase in their difference with  $\mu$ . The decrease in the stability bound for the improper input conforms with Remark 2. Fig. 8 illustrates a match between theoretical

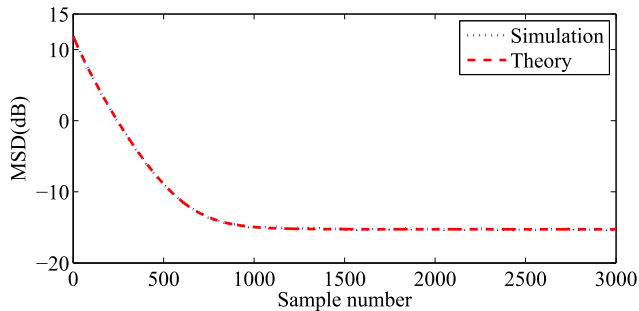
and simulated steady-state EMSEs for inputs with various impropriety coefficients  $r \in [0, 1]$  and  $\mu = 0.1$ . The increase in the steady-state EMSEs with  $r$  conforms with Remark 3.

**C. WL-QLMS**

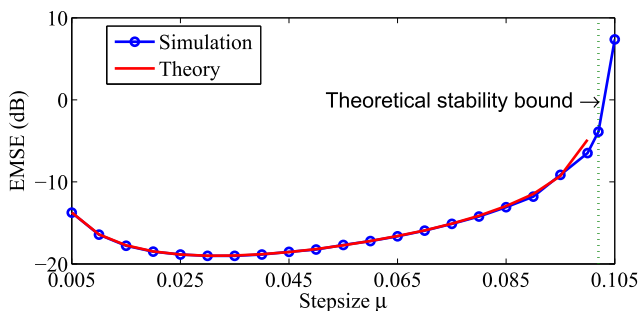
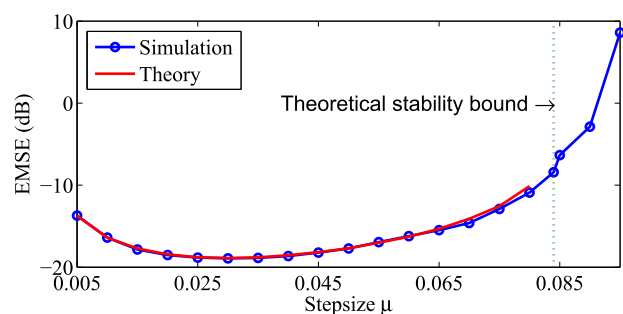
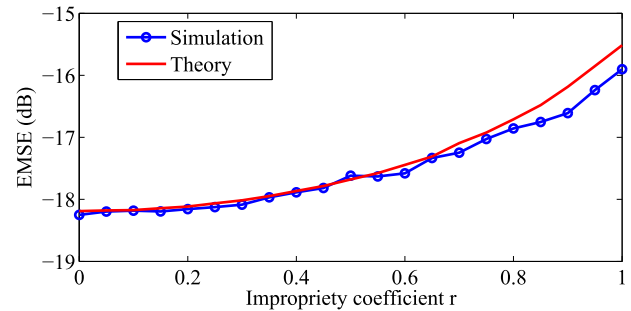
In the simulations for the WL-QLMS, the desired signal was generated from the widely linear model in (11) and (12). The variance of the additive Gaussian noise  $v_n$  was  $\sigma_v^2 = 0.01$ , and the Euclidean norm of the Gaussian weight variation  $\mathbf{q}_n$  was 0.0004. Fig. 9 depicts theoretical and simulated EMSE and MSD learning curves of the SL-QLMS for an improper input with  $r = 0.5$  and  $\mu = 0.02$ . These curves exhibit a good match between the theory and simulation results. Figs. 10 and 11 show theoretical and simulated steady-state EMSEs for different values of the step size  $\mu$  when  $\mathbf{x}_n$  is  $\mathbb{H}$ -proper and when  $\mathbf{x}_n$  is improper with  $r = 0.5$ . The theoretical bounds of  $\mu$  are also indicated. From both figures, we can observe a good match



(a) EMSE curves



(b) MSD curves

 Fig. 9. Theoretical and simulated learning curves of WL-QLMS for an improper input with  $r = 0.5$  and  $\mu = 0.02$ .

 Fig. 10. Theoretical and simulated EMSEs of WL-QLMS as a function of the step size, for an  $\mathbb{H}$ -proper input.

 Fig. 11. Theoretical and simulated EMSEs of WL-QLMS as a function of the step size, for an improper input with  $r = 0.5$ .

 Fig. 12. Theoretical and simulated EMSEs of WL-QLMS as a function of the impropriety coefficient, for  $\mu = 0.05$ .

between theory and simulations and an increase in their difference with  $\mu$ . The decrease in the stability bound for the improper input conforms with Remark 2. Fig. 12 illustrates a match between theoretical and simulated steady-state EMSEs for inputs with various impropriety coefficients  $r \in [0, 1]$  and  $\mu = 0.05$ . The increase in the steady-state EMSEs with  $r$  conforms with Remark 3.

## VI. CONCLUSION

Convergence of quaternion adaptive filtering algorithms in non-stationary environments has been analysed by extending an analysis framework for real-valued adaptive filters to the quaternion domain. For rigour, the first-order random-walk model has been employed to model the variation of system weights in a non-stationary environment. The bounds on transient and steady-state performance of general quaternion adaptive filters have been derived. For QLMS filters with Gaussian inputs, an innovative quaternion decorrelation technique has simplified the analysis and enabled the step size bounds, MSDs and EMSEs to be expressed in closed forms. The analytical results, which are shown to be influenced by the second-order statistics of input signals, provide new insights into the statistical behaviour of the considered filtering algorithms. Effectiveness of the proposed analysis has been verified by numerical simulations for SL-QLMS, SWL-QLMS and WL-QLMS filters with correlated Gaussian inputs. The proposed analysis framework may be extended to other division algebras, such as the analysis of octonion adaptive filters, a topic of future research.

## APPENDIX PROOF OF (41)

Based on Isserlis' theorem [52], the fourth-order moments of real and complex variables can be represented by their second-order moments. However, Isserlis' theorem does not apply to quaternion variables because the quaternion products are non-commutative. To circumvent this difficulty, we represent quaternion variables using the Cayley-Dickson construction, that is, the  $m$ -th element of  $\mathbf{s}_n$  is written as  $s_m = c_m + d_m j$ , where  $c_m$  and  $d_m$  are complex variables defined on the basis  $\{1, i\}$  [18]. By applying Isserlis' theorem to the constitutive complex variables, the element on the  $a$ -th row and  $b$ -th column of the

fourth-order moment matrix  $E\{\|\mathbf{s}_n\|^2 \mathbf{s}_n \mathbf{s}_n^H\}$  can be represented by the second-order statistics of  $\mathbf{s}_n$  as

$$\begin{aligned}
& \left[ E \left\{ \|\mathbf{s}_n\|^2 \mathbf{s}_n \mathbf{s}_n^H \right\} \right]_{a,b} \\
&= E \left\{ \sum_{m=1}^M s_m s_m^* s_a s_b^* \right\} \\
&= E \left\{ \sum_{m=1}^M (c_m + d_m j) (c_m^* - d_m j) (c_a + d_a j) (c_b^* - d_b j) \right\} \\
&= \sum_{m=1}^M (E\{c_m c_m^* c_a c_b^* + c_m c_m^* d_a d_b^* + d_m d_m^* c_a c_b^* \\
&\quad + d_m d_m^* d_a d_b^*\} + E\{c_m c_m^* d_a c_b + d_m d_m^* d_a c_b - c_m c_m^* c_a d_b \\
&\quad - d_m d_m^* c_a d_b\} j) \\
&= \sum_{m=1}^M E\{c_m c_m^* + d_m d_m^*\} E\{c_a c_b^* + d_a d_b^* + (d_a c_b - c_a d_b) j\} \\
&\quad + \sum_{m=1}^M (E\{c_m c_a\} E\{c_m^* c_b^*\} + E\{c_m c_b^*\} E\{c_m^* c_a\} \\
&\quad + E\{c_m d_a\} E\{c_m^* d_b^*\} + E\{c_m d_b^*\} E\{c_m^* d_a\} \\
&\quad + E\{d_m c_a\} E\{d_m^* c_b^*\} + E\{d_m c_b^*\} E\{d_m^* c_a\} \\
&\quad + E\{d_m d_a\} E\{d_m^* d_b^*\} + E\{d_m d_b^*\} E\{d_m^* d_a\}) \\
&\quad + \sum_{m=1}^M (E\{c_m d_a\} E\{c_m^* c_b\} + E\{c_m c_b\} E\{c_m^* d_a\} \\
&\quad + E\{d_m d_a\} E\{d_m^* c_b\} + E\{d_m c_b\} E\{d_m^* d_a\} \\
&\quad - E\{c_m c_a\} E\{c_m^* d_b\} - E\{c_m d_b\} E\{c_m^* d_a\} \\
&\quad - E\{d_m c_a\} E\{d_m^* d_b\} - E\{d_m d_b\} E\{d_m^* c_a\}) j
\end{aligned}$$

We also observe

$$\begin{aligned}
[\mathbf{C}_{\text{ss}}]_{a,b} &= E\{s_a s_b^*\} = E\{(c_a + d_a j)(c_b^* - d_b j)\} \\
&= E\{c_a c_b^* + d_a d_b^*\} + E\{d_a c_b - c_a d_b\} j \\
[\mathbf{C}_{\text{ss}^t}]_{a,b} &= E\{s_a s_b^{*t}\} = E\{(c_a + d_a j)(c_b^* + d_b j)\} \\
&= E\{c_a c_b^* - d_a d_b^*\} + E\{d_a c_b + c_a d_b\} j \\
[\mathbf{C}_{\text{ss}^j}]_{a,b} &= E\{s_a s_b^{*j}\} = E\{(c_a + d_a j)(c_b - d_b^* j)\} \\
&= E\{c_a c_b + d_a d_b\} + E\{d_a c_b^* - c_a d_b^*\} j \\
[\mathbf{C}_{\text{ss}^k}]_{a,b} &= E\{s_a s_b^{*k}\} = E\{(c_a + d_a j)(c_b + d_b^* j)\} \\
&= E\{c_a c_b - d_a d_b\} + E\{d_a c_b^* + c_a d_b^*\} j
\end{aligned}$$

From the above equations, we can verify

$$\begin{aligned}
\left[ E \left\{ \|\mathbf{s}_n\|^2 \mathbf{s}_n \mathbf{s}_n^H \right\} \right]_{a,b} &= \text{Tr}(\mathbf{C}_{\text{ss}}) [\mathbf{C}_{\text{ss}}]_{a,b} + \frac{1}{2} [\mathbf{C}_{\text{ss}}^2]_{a,b} \\
&\quad + \frac{1}{2} \sum_{\alpha=l,j,\kappa} [\mathbf{C}_{\text{ss}^\alpha} \mathbf{C}_{\text{ss}^\alpha}^*]_{a,b}
\end{aligned}$$

Therefore, (41) holds.

## REFERENCES

- [1] J. B. Kuipers, *Quaternions and Rotation Sequences: A Primer With Applications to Orbits, Aerospace and Virtual Reality*. Princeton, NJ, USA: Princeton Univ. Press, 1999.
- [2] J. Vía, D. Ramírez, and I. Santamaría, "Properness and widely linear processing of quaternion random vectors," *IEEE Trans. Inf. Theory*, vol. 56, no. 7, pp. 3502–3515, Jul. 2010.
- [3] C. C. Took and D. P. Mandic, "Augmented second-order statistics of quaternion random signals," *Signal Process.*, vol. 91, no. 2, pp. 214–224, 2011.
- [4] D. P. Mandic, C. Jahanchahi, and C. C. Took, "A quaternion gradient operator and its applications," *IEEE Signal Process. Lett.*, vol. 18, no. 1, pp. 47–50, Jan. 2011.
- [5] D. Xu, C. Jahanchahi, C. Cheong Took, and D. P. Mandic, "Enabling quaternion derivatives: The generalized HR calculus," *Roy. Soc. Open Sci.*, vol. 2, no. 8, 2015, Art. no. 150255.
- [6] F. Ortolani, D. Communiello, M. Scarpiniti, and A. Uncini, "Frequency-domain adaptive filtering in hypercomplex systems," in *Advances in Neural Networks*. New York, NY, USA: Springer, 2016, pp. 47–56.
- [7] C. Jahanchahi and D. P. Mandic, "A class of quaternion Kalman filters," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 25, no. 3, pp. 533–544, Mar. 2014.
- [8] T. Isokawa, T. Kusakabe, N. Matsui, and F. Peper, "Quaternion neural network and its application," in *Knowledge-Based and Intelligent Information and Engineering Systems (Lecture Notes in Computer Science)*, vol. 2774. Berlin, Germany: Springer-Verlag, 2003, pp. 318–324.
- [9] N. Le Bihan and S. Buchholz, "Quaternionic independent component analysis using hypercomplex nonlinearities," in *IMA Conf. Math. Signal Process.*, 2006.
- [10] J. Navarro-Moreno, R. M. Fernández-Alcalá, and J. C. Ruiz-Molina, "A quaternion widely linear model for nonlinear Gaussian estimation," *IEEE Trans. Signal Process.*, vol. 62, no. 24, pp. 6414–6424, Dec. 2014.
- [11] T. A. Ell, N. Le Bihan, and S. J. Sangwine, *Quaternion Fourier Transforms for Signal and Image Processing*. Hoboken, NJ, USA: Wiley, 2014.
- [12] S. Miron, N. L. Bihan, and J. I. Mars, "Quaternion-MUSIC for vector-sensor array processing," *IEEE Trans. Signal Process.*, vol. 54, no. 4, pp. 1218–1229, Apr. 2006.
- [13] J.-F. Gu and K. Wu, "Quaternion modulation for dual-polarized antennas," *IEEE Commun. Lett.*, vol. 21, no. 2, pp. 286–289, Feb. 2017.
- [14] H. Fourati, N. Manamanni, L. Afilal, and Y. Handrich, "Complementary observer for body segments motion capturing by inertial and magnetic sensors," *IEEE/ASME Trans. Mechatronics*, vol. 19, no. 1, pp. 149–157, Feb. 2014.
- [15] K. Adhikari, S. Tatinati, W. T. Ang, K. C. Veluvolu, and K. Nazarpour, "A quaternion weighted Fourier linear combiner for modeling physiological tremor," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 11, pp. 2336–2346, Nov. 2016.
- [16] C. Cheong Took and D. P. Mandic, "The quaternion LMS algorithm for adaptive filtering of hypercomplex processes," *IEEE Trans. Signal Process.*, vol. 57, no. 4, pp. 1316–1327, Apr. 2009.
- [17] C. Cheong Took and D. P. Mandic, "A quaternion widely linear adaptive filter," *IEEE Trans. Signal Process.*, vol. 58, no. 8, pp. 4427–4431, Aug. 2010.
- [18] T. Mizoguchi and I. Yamada, "An algebraic translation of Cayley-Dickson linear systems and its applications to online learning," *IEEE Trans. Signal Process.*, vol. 62, no. 6, pp. 1438–1453, Mar. 2014.
- [19] M. Jiang, W. Liu, and Y. Li, "Adaptive beamforming for vector-sensor arrays based on a reweighted zero-attracting quaternion-valued LMS algorithm," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 63, no. 3, pp. 274–278, Mar. 2016.
- [20] F. Ortolani, D. Communiello, M. Scarpiniti, and A. Uncini, "Frequency domain quaternion adaptive filters: Algorithms and convergence performance," *Signal Process.*, vol. 136, pp. 69–80, 2017.
- [21] J. D. Jiménez-López, R. M. Fernández-Alcalá, J. Navarro-Moreno, and J. C. Ruiz-Molina, "Widely linear estimation of quaternion signals with intermittent observations," *Signal Process.*, vol. 136, pp. 92–101, 2017.
- [22] J. Navarro-Moreno, R. M. Fernández-Alcalá, and J. C. Ruiz-Molina, "Semi-widely simulation and estimation of continuous-time  $\mathbb{C}^n$ -proper quaternion random signals," *IEEE Trans. Signal Process.*, vol. 63, no. 18, pp. 4999–5012, Sep. 2015.
- [23] J. Navarro-Moreno and J. C. Ruiz-Molina, "Semi-widely linear estimation of  $\mathbb{C}^n$ -proper quaternion random signal vectors under Gaussian and stationary conditions," *Signal Process.*, vol. 119, pp. 56–66, 2016.

[24] D. P. Mandic, S. Still, and S. C. Douglas, "Duality between widely linear and dual channel adaptive filtering," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2009, pp. 1729–1732.

[25] D. P. Mandic and V. S. L. Goh, *Complex Valued Nonlinear Adaptive Filters: Noncircularity, Widely Linear and Neural Models*. Hoboken, NJ, USA: Wiley, 2009.

[26] S. Haykin, *Adaptive Filter Theory*, 4th ed., Upper Saddle River, NJ, USA: Prentice-Hall, 2002.

[27] T. Y. Al-Naffouri and A. H. Sayed, "Transient analysis of data-normalized adaptive filters," *IEEE Trans. Signal Process.*, vol. 51, no. 3, pp. 639–652, Mar. 2003.

[28] N. R. Yousef and A. H. Sayed, "A unified approach to the steady-state and tracking analyses of adaptive filters," *IEEE Trans. Signal Process.*, vol. 49, no. 2, pp. 314–324, Feb. 2001.

[29] A. H. Sayed, *Adaptive Filters*. Hoboken, NJ, USA: Wiley, 2008.

[30] C. Cheong Took, C. Jahanchahi, and D. P. Mandic, "A unifying framework for the analysis of quaternion valued adaptive filters," in *Proc. IEEE Signals, Syst. Comput.*, 2011, pp. 1771–1774.

[31] S. P. Talebi, S. Kanna, and D. P. Mandic, "Real-time estimation of quaternion imprpropriety," in *Proc. IEEE Int. Conf. Digit. Signal Process.*, 2015, pp. 557–561.

[32] T. A. Ell and S. J. Sangwine, "Hypercomplex Fourier transforms of color images," *IEEE Trans. Image Process.*, vol. 16, no. 1, pp. 22–35, Jan. 2007.

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

[34] F. Zhang, "Quaternions and matrices of quaternions," *Linear Algebr. Appl.*, vol. 251, pp. 21–57, 1997.

[35] L. Huang and W. So, "On left eigenvalues of a quaternionic matrix," *Linear Algebr. Appl.*, vol. 323, no. 1–3, pp. 105–116, 2001.

[36] D. R. Farenick and B. A. Pidkowich, "The spectral theorem in quaternions," *Linear Algebr. Appl.*, vol. 371, pp. 75–102, 2003.

[37] L. Zou, Y. Jiang, and J. Wu, "Location for the right eigenvalues of quaternion matrices," *J. Appl. Math. Comput.*, vol. 38, no. 1, pp. 71–83, 2012.

[38] L. Rodman, *Topics in Quaternion Linear Algebra*. Princeton, NJ, USA: Princeton Univ. Press, 2014.

[39] W. Fulton, "Eigenvalues, invariant factors, highest weights, and Schubert calculus," *Bull. Amer. Math. Soc.*, vol. 37, no. 3, pp. 209–249, 2000.

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

[41] P. O. Amblard and N. Le Bihan, "On properness of quaternion valued random variables," in *Proc. IMA Conf. Math. Signal Process.*, 2004, pp. 23–26.

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

[43] A. H. Sayed, *Fundamentals of Adaptive Filtering*. Hoboken, NJ, USA: Wiley, 2003.

[44] N. J. Bershad, E. Eweda, and J. C. M. Bermudez, "Stochastic analysis of the LMS and NLMS algorithms for cyclostationary white Gaussian inputs," *IEEE Trans. Signal Process.*, vol. 62, no. 9, pp. 2238–2249, May 2014.

[45] B. Widrow, J. M. McCool, M. G. Larimore, and C. R. Johnson, "Stationary and nonstationary learning characteristics of the LMS adaptive filter," *Proc. IEEE*, vol. 64, no. 8, pp. 1151–1162, Aug. 1976.

[46] D. Schulz and R. S. Thomä, "Using quaternion-valued linear algebra," arXiv:1311.7488, 2013.

[47] Y. Tian and G. P. Styan, "Some inequalities for sums of nonnegative definite matrices in quaternions," *J. Inequalities Appl.*, vol. 2005, no. 5, 2005, Art. no. 385246.

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

[49] M. Xiang, S. Enshaeifar, A. E. Stott, C. Cheong Took, Y. Xia, S. Kanna, and D. P. Mandic, "Simultaneous diagonalisation of the covariance and complementary covariance matrices in quaternion widely linear signal processing," *Signal Process.*, accepted for publication.

[50] M. Xiang, S. Kanna, S. C. Douglas, and D. P. Mandic, "Performance advantage of quaternion widely linear estimation: An approximate uncorrelating transform approach," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2016, pp. 4348–4352.

[51] P. Chevalier and F. Pipon, "New insights into optimal widely linear array receivers for the demodulation of BPSK, MSK, and GMSK signals corrupted by noncircular interferences-application to SAIC," *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 870–883, Mar. 2006.

[52] L. Isserlis, "On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables," *Biometrika*, vol. 12, no. 1/2, pp. 134–139, 1918.



**Min Xiang** (M'17) received the B.Eng. and M.Eng. degrees from Beihang University, Beijing, China. He is currently working toward the Ph.D. degree in electrical engineering at the Imperial College London, London, U.K. His research interests include quaternion-valued statistical signal processing and adaptive filtering.



**Sithan Kanna** (S'13) received the M.Eng. degree in electrical and electronic engineering with management from the Imperial College London, London, U.K., in 2012. His research interests include complex-valued statistical signal processing and frequency estimation in the smart grid. He was the recipient of full scholarship by The Rector's Fund at the Imperial College London to pursue the Ph.D. degree in adaptive signal processing.



**Danilo P. Mandic** (M'99–SM'03–F'12) received the Ph.D. degree in nonlinear adaptive signal processing from the Imperial College London, London, U.K., in 1999. He is a Professor of Signal Processing with the Imperial College London, where he has been involved in nonlinear adaptive signal processing and nonlinear dynamics. He is a Deputy Director of the Financial Signal Processing Laboratory, Imperial College London. He has been a Guest Professor with Katholieke Universiteit Leuven, Leuven, Belgium, the Tokyo University of Agriculture and Technology, Tokyo, Japan, and Westminster University, London, U.K., and a Frontier Researcher with RIKEN, Wako, Japan. He has authored/coauthored two research monographs titled *Recurrent Neural Networks for Prediction: Learning Algorithms, Architectures and Stability* (Wiley, 2001) and *Complex Valued Nonlinear Adaptive Filters: Noncircularity, Widely Linear and Neural Models* (Wiley, 2009), an edited book titled *Signal Processing Techniques for Knowledge Extraction and Information Fusion* (Springer, 2008), and more than 200 publications on signal and image processing. He has been a member of the IEEE Technical Committee on Signal Processing Theory and Methods, and an Associate Editor for the IEEE SIGNAL PROCESSING MAGAZINE, the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II, the IEEE TRANSACTIONS ON SIGNAL PROCESSING, and the IEEE TRANSACTIONS ON NEURAL NETWORKS. He has produced award winning papers and products resulting from his collaboration with the industry.