STUDY OF THE QUATERNION LMS AND FOUR-CHANNEL LMS ALGORITHMS

Clive Cheong Took[†], Danilo Mandic[†], and Jacob Benesty[*]

[†]Electrical Engineering Department, Imperial College London SW7 2AZ, UK,
[*]Université du Québec, INRS-EMT, Montréal QC H5A 1K6, Canada,
E-mails: {c.cheong-took, d.mandic }@ic.ac.uk, benesty@emt.inrs.ca

ABSTRACT

The recently proposed quaternion least-mean-square (QLMS) algorithm for adaptive filtering of three- and four-dimensional signals has been analysed in the context of multi-step ahead prediction. For rigour, the relationship between multichannel LMS (MLMS) and QLMS is examined, and their differences are highlighted. This is achieved both in terms of the input-output relationship and in terms of the dynamics of weight updates. The convergence of QLMS is investigated and stability bounds confirm that QLMS and MLMS are fundamentally different. Simulations on both synthetic and real world multidimensional signals support the analysis.

Index Terms— Quaternion adaptive filtering, multidimensional systems, multistep prediction, wind modelling.

1. INTRODUCTION

The advance in the theory of multiple-input and multiple-output (MIMO) systems, such as in acoustic echo cancellation, has brought to light the corresponding adaptive filtering algorithms, such as the multi-channel Least-Mean-Square (MLMS) and its derivatives [1]. In particular, for a 4×4 case addressed in this paper, the output of the adaptive filter is given by

$$\mathbf{y}(n) = \mathbf{H}^T(n)\mathbf{x}(n) \tag{1}$$

$$= \begin{bmatrix} \mathbf{h}_{11}(n) & \mathbf{h}_{12}(n) & \mathbf{h}_{13}(n) & \mathbf{h}_{14}(n) \\ \mathbf{h}_{21}(n) & \mathbf{h}_{22}(n) & \mathbf{h}_{23}(n) & \mathbf{h}_{24}(n) \\ \mathbf{h}_{31}(n) & \mathbf{h}_{32}(n) & \mathbf{h}_{33}(n) & \mathbf{h}_{34}(n) \\ \mathbf{h}_{41}(n) & \mathbf{h}_{42}(n) & \mathbf{h}_{43}(n) & \mathbf{h}_{44}(n) \end{bmatrix}^T \begin{bmatrix} \mathbf{x}_1(n) \\ \mathbf{x}_2(n) \\ \mathbf{x}_3(n) \\ \mathbf{x}_4(n) \end{bmatrix}$$

where the output $\mathbf{y}(n) = [y_1(n), \dots, y_4(n)]^T$, and the pth input vector $\mathbf{x}_i(n) = [x_i(n), \dots, x_i(n - L)]^T$. The update for each coefficient vector $\mathbf{h}_{ij}(n)$ is given by [1]

$$\begin{aligned} \mathbf{h}_{ij}(n + 1) &= \mathbf{h}_{ij}(n) + \Delta\mathbf{h}_{ij}(n) \tag{2} \\ &= \mathbf{h}_{ij}(n) + \mu e_j(n)\mathbf{x}_i(n) \quad i, j = 1, \dots, 4 \end{aligned}$$

where μ is a stepsize parameter that controls the stability of the algorithm. The error $e_j(n) = d_j(n) - y_j(n)$ is a scalar instantaneous output error corresponding to the jth channel.

Prior to performing multidimensional signal processing, it is important to establish whether the direct multidimensional adaptive filtering has advantages over multichannel modelling. For instance, complex signal processing enables the exploitation of the phase information; its augmented statistics [2] and the fully complex model [3] has led to improved algorithms through the coupling of the two dimensions. The benefit of processing in $\mathbb{C}$ over that in $\mathbb{R}^2$ also includes a compact representation of "directional" processes (radar, sonar and wind).

A thorough study of the duality between complex LMS and two-channel MLMS is given in [4]. However, as complex signal processing is restricted to at most two channels[1], the quaternion (hypercomplex) domain provides a higher dimensional alternative. This paper examines the differences between QLMS with its counterpart MLMS; this is achieved for both the input/output relationship and the dynamics of weight updates.

The organisation of this paper is as follows: section 2 provides an overview of the algebraic properties of quaternions. Section 3.1 examines the static models of both QLMS and MLMS. This is followed by an analysis on their dynamics in section 3.2. Next, the QLMS stepsize range for convergence is derived and compared with that of MLMS in section 3.3. Simulation studies are given in section 4, complemented by the conclusion in section 5.

2. BACKGROUND

Quaternions can be considered as non-commutative extensions of complex numbers, and comprise of at most four variables [6]. A quaternion variable $q \in \mathbb{H}$ has a real part (denoted with subscript a), and three imaginary parts (denoted

---

[1]Except when Nehorai and Paldi adopted a long vector model, in which the three or four components are concatenated in a long vector [5].

with subscripts b, c, d), and can be expressed as:

$$
\begin{aligned}
q &= [q_a, \mathbf{q}] \\
&= [q_a, (q_b, q_c, q_d)] \\
&= q_a + q_b\imath + q_c\jmath + q_d\kappa \qquad \{q_a, q_b, q_c, q_d \in \mathbb{R}\} \ (3)
\end{aligned}
$$

The relationship between $\imath$, $\jmath$, κ describing the three imaginary dimensions (the vector part) of a quaternion are

$$
\begin{aligned}
\imath\jmath &= \kappa \quad \jmath\kappa = \imath \quad \kappa\imath = \jmath \\
\imath\jmath\kappa &= \imath^2 = \jmath^2 = \kappa^2 = -1 \qquad (4)
\end{aligned}
$$

Due to the non-commutativity of the quaternion, for example, $\jmath\imath \neq \imath\jmath$, instead $\jmath\imath = -\kappa$. Likewise, the product of quaternions p and q is given by

$$
\begin{aligned}
pq &= [p_a, \mathbf{p}][q_a, \mathbf{q}] \\
&= [p_a q_a - \mathbf{p} \cdot \mathbf{q}, \ p_a\mathbf{q} + q_a\mathbf{p} + \mathbf{p} \times \mathbf{q}] \qquad (5)
\end{aligned}
$$

where symbols "·" and "×" denote respectively the dot-product and the cross-product. It is evident that the quaternion product is non-commutative, due to the outer product. On the other hand, similarly to the complex numbers, the conjugate of a quaternion is given by $q^* = [q_a, \mathbf{q}]^* = [q_a, -\mathbf{q}]$ and the norm $||q||_2^2 = qq^*$. It is crucial to appreciate the difference between the real norm and the quaternion norm; for more detail, see [7].

Recently, the quaternion LMS (QLMS) was introduced for four dimensional adaptive filtering [8], and the heterogeneous fusion of three- four-dimensional signals [9]. The QLMS minimises the real valued cost function

$$
\mathcal{J}(n) = \tilde{e}(n)\tilde{e}^*(n) = \tilde{e_a}^2(n) + \tilde{e_b}^2(n) + \tilde{e_c}^2(n) + \tilde{e_d}^2(n) \tag{6}
$$

which is a function of the quaternion-valued error $\tilde{e}(n)$. Mathematically, the QLMS can be expressed as

$$
\begin{aligned}
\mathbf{w}(n+1) &= \mathbf{w}(n) + \Delta\mathbf{w}(n) \\
&= \mathbf{w}(n) + \mu\big(2\tilde{e}(n)\tilde{\mathbf{x}}^*(n) - \tilde{\mathbf{x}}^*(n)\tilde{e}^*(n)\big) \ (7)
\end{aligned}
$$

where the superscript $(\cdot)^*$ denotes the quaternion conjugation operator, $\mathbf{w}(n) = \mathbf{w}_a + \mathbf{w}_b\imath + \mathbf{w}_c\jmath + \mathbf{w}_d\kappa$ and $\tilde{\mathbf{x}}(n) = \tilde{\mathbf{x}}_a + \tilde{\mathbf{x}}_b\imath + \tilde{\mathbf{x}}_c\jmath + \tilde{\mathbf{x}}_d\kappa$ are quaternion vectors of length L. Based on (5) and $\tilde{y}(n) = \mathbf{w}^T(n)\tilde{\mathbf{x}}(n)$, the outputs of the quaternion valued linear filter are given by

$$
\begin{aligned}
\tilde{y}_a(n) &= \mathbf{w}_a^T(n)\tilde{\mathbf{x}}_a(n) - \mathbf{w}_b^T(n)\tilde{\mathbf{x}}_b(n) \\
&\quad - \mathbf{w}_c^T(n)\tilde{\mathbf{x}}_c(n) - \mathbf{w}_d^T(n)\tilde{\mathbf{x}}_d(n) \\
\tilde{y}_b(n) &= \mathbf{w}_b^T(n)\tilde{\mathbf{x}}_a(n) + \mathbf{w}_a^T(n)\tilde{\mathbf{x}}_b(n) \\
&\quad - \mathbf{w}_d^T(n)\tilde{\mathbf{x}}_c(n) + \mathbf{w}_c^T(n)\tilde{\mathbf{x}}_d(n) \\
\tilde{y}_c(n) &= \mathbf{w}_c^T(n)\tilde{\mathbf{x}}_a(n) + \mathbf{w}_d^T(n)\tilde{\mathbf{x}}_b(n) \\
&\quad + \mathbf{w}_a^T(n)\tilde{\mathbf{x}}_c(n) - \mathbf{w}_b^T(n)\tilde{\mathbf{x}}_d(n) \\
\tilde{y}_d(n) &= \mathbf{w}_d^T(n)\tilde{\mathbf{x}}_a(n) - \mathbf{w}_c^T(n)\tilde{\mathbf{x}}_b(n) \\
&\quad + \mathbf{w}_b^T(n)\tilde{\mathbf{x}}_c(n) + \mathbf{w}_a^T(n)\tilde{\mathbf{x}}_d(n) \qquad (8)
\end{aligned}
$$

3. ADAPTIVE FILTERING IN $\mathbb{H}$ AND $\mathbb{R}^4$

To compare the operation of adaptive filters in $\mathbb{R}^4$ and $\mathbb{H}$, consider their functional-valued progression given respectively in (8) and (1). Observe that the corresponding channels

$$
\begin{aligned}
\mathbf{x}_1(n) &= \tilde{\mathbf{x}}_a(n) \qquad \mathbf{x}_2(n) = \tilde{\mathbf{x}}_b(n) \\
\mathbf{x}_3(n) &= \tilde{\mathbf{x}}_c(n) \qquad \mathbf{x}_4(n) = \tilde{\mathbf{x}}_d(n) \qquad (9)
\end{aligned}
$$

and the errors,

$$
\begin{aligned}
e_1(n) &= \tilde{e}_a(n) \qquad e_2(n) = \tilde{e}_b(n) \\
e_3(n) &= \tilde{e}_c(n) \qquad e_4(n) = \tilde{e}_d(n) \qquad (10)
\end{aligned}
$$

3.1. The input-output relationship

A comparison between the two sets of equations (1) and (8) shows that MLMS can be regarded as a constrained version of QLMS, provided the adaptive coefficient matrix of MLMS in (1) satisfies

$$
\mathbf{H}(n) = \begin{bmatrix}
\mathbf{w}_a(n) & \mathbf{w}_b(n) & \mathbf{w}_c(n) & \mathbf{w}_d(n) \\
-\mathbf{w}_b(n) & \mathbf{w}_a(n) & \mathbf{w}_d(n) & -\mathbf{w}_c(n) \\
-\mathbf{w}_c(n) & -\mathbf{w}_d(n) & \mathbf{w}_a(n) & \mathbf{w}_b(n) \\
-\mathbf{w}_d(n) & \mathbf{w}_c(n) & -\mathbf{w}_b(n) & \mathbf{w}_a(n)
\end{bmatrix} \tag{11}
$$

This constraint is, however, very stringent. For detailed insight, we proceed with examining the dynamic characteristics of these multidimensional algorithms via their coefficient updates and their convergence bounds.

3.2. The dynamics of weight updates

We next investigate the relationship between the dynamics of QLMS and MLMS updates. In other words, each component of $\Delta\mathbf{w}(n)$ will be related to corresponding components of $\Delta\mathbf{h}(n)$. To this end, from (7) $\Delta\mathbf{w}(n)$ can be expanded dimension-wise as

$$
\begin{aligned}
\Delta\mathbf{w}_a(n) &= \tilde{e}_a(n)\tilde{\mathbf{x}}_a(n) + 3[\tilde{e}_b(n)\tilde{\mathbf{x}}_b(n) + \tilde{e}_c(n)\tilde{\mathbf{x}}_c(n) + \tilde{e}_d(n)\tilde{\mathbf{x}}_d(n)] \\
\Delta\mathbf{w}_b(n) &= 3\tilde{e}_b(n)\tilde{\mathbf{x}}_a(n) - \tilde{e}_a(n)\tilde{\mathbf{x}}_b(n) + \tilde{e}_d(n)\tilde{\mathbf{x}}_c(n) - \tilde{e}_c(n)\tilde{\mathbf{x}}_d(n) \\
\Delta\mathbf{w}_c(n) &= 3\tilde{e}_c(n)\tilde{\mathbf{x}}_a(n) - \tilde{e}_d(n)\tilde{\mathbf{x}}_b(n) - \tilde{e}_a(n)\tilde{\mathbf{x}}_c(n) + \tilde{e}_b(n)\tilde{\mathbf{x}}_d(n) \\
\Delta\mathbf{w}_d(n) &= 3\tilde{e}_d(n)\tilde{\mathbf{x}}_a(n) + \tilde{e}_c(n)\tilde{\mathbf{x}}_b(n) - \tilde{e}_b(n)\tilde{\mathbf{x}}_c(n) - \tilde{e}_a(n)\tilde{\mathbf{x}}_d(n)
\end{aligned} \tag{12}
$$

For space, the multiplicative factor μ has been omitted. From (2) and (12), $\Delta\mathbf{w}(n)$ can be related to $\Delta\mathbf{h}(n)$ as

$$
\begin{aligned}
\Delta\mathbf{w}_a(n) &= \Delta\mathbf{h}_{11}(n) + 3[\Delta\mathbf{h}_{22}(n) + \Delta\mathbf{h}_{33}(n) + \Delta\mathbf{h}_{44}(n)] \\
\Delta\mathbf{w}_b(n) &= 3\Delta\mathbf{h}_{12}(n) - \Delta\mathbf{h}_{21}(n) + \Delta\mathbf{h}_{34}(n) - \Delta\mathbf{h}_{43}(n) \\
\Delta\mathbf{w}_c(n) &= 3\Delta\mathbf{h}_{13}(n) - \Delta\mathbf{h}_{24}(n) - \Delta\mathbf{h}_{31}(n) + \Delta\mathbf{h}_{42}(n) \\
\Delta\mathbf{w}_d(n) &= 3\Delta\mathbf{h}_{14}(n) + \Delta\mathbf{h}_{23}(n) - \Delta\mathbf{h}_{32}(n) - \Delta\mathbf{h}_{41}(n)
\end{aligned} \tag{13}
$$

Observe that the dynamics of QLMS and MLMS are different. To provide further insight, we next examine another factor that influences the dynamics of these algorithms, that is, the stepsize.

3.3. On the bounds of the stepsize

The stepsize parameter plays a key role in stability and convergence of adaptive filters. In particular, the stepsize bounds are an important issue for convergence, both from a theoretical and practical point of view. According to Huang and Benesty, the stepsize range of the MLMS is given by [1]

$$0 < \mu < \frac{2}{\sum_{p=1}^{4} ||\mathbf{x}_p(n)||_2^2} \quad (14)$$

On the other hand, it is not straightforward to derive the QLMS stepsize range, due to the non-commutativity of the quaternion product. However, similarly to the case of adaptive IIR filters [10], it is possible to obtain an approximate range by local linearisation using the Taylor series expansion.

To introduce stability bounds for QLMS, consider the following a priori $\tilde{e}(n)$ and a posteriori $\bar{e}(n)$ errors

$$\tilde{e}(n) = d(n) - \mathbf{w}^T(n)\tilde{\mathbf{x}}(n) \quad (15)$$
$$\bar{e}(n) = d(n) - \mathbf{w}^T(n+1)\tilde{\mathbf{x}}(n) \quad (16)$$

The task is to estimate the range of values for μ, such that $||\bar{e}(n)|| < ||\tilde{e}(n)||$. Following the approach from [10], $||\bar{e}(n)||^2$ can be expressed in terms of $||\tilde{e}(n)||^2$ via first-order Taylor series

$$||\bar{e}(n)||^2 = ||\tilde{e}(n)||^2 + \Delta\mathbf{w}^H(n)\frac{\partial ||\tilde{e}(n)||^2}{\partial \mathbf{w}(n)} \quad (17)$$

where $\Delta\mathbf{w}(n)$ can be readily obtained from (7), and $\frac{\partial ||\tilde{e}(n)||^2}{\partial \mathbf{w}(n)}$ can be shown to be [9]

$$\frac{\partial ||\tilde{e}(n)||^2}{\partial \mathbf{w}(n)} = -2\big(2\tilde{e}(n)\tilde{\mathbf{x}}^*(n) - \tilde{\mathbf{x}}^*(n)\tilde{e}^*(n)\big) \quad (18)$$

Substituting (18) and $\Delta\mathbf{w}(n)$ from (7) into (17), the a posteriori error $||\bar{e}(n)||^2$ can be expanded as

$$||\bar{e}(n)||^2 = ||\tilde{e}(n)||^2 - 2\mu\bigg(5||\tilde{e}(n)||^2 \sum_{\ell=1}^{L} ||\tilde{x}(n-\ell)||^2$$
$$-2\big(\tilde{\mathbf{x}}^T(n)\tilde{e}^*(n)\tilde{\mathbf{x}}^*(n)\tilde{e}^*(n) + \tilde{e}(n)\tilde{\mathbf{x}}^T(n)\tilde{e}(n)\tilde{\mathbf{x}}^*(n)\big)\bigg) \quad (19)$$

Due to the non-commutativity of the quaternion product, it is difficult to express the terms $\tilde{\mathbf{x}}^T(n)\tilde{e}^*(n)\tilde{\mathbf{x}}^*(n)\tilde{e}^*(n)$ and $\tilde{e}(n)\tilde{\mathbf{x}}^T(n)\tilde{e}(n)\tilde{\mathbf{x}}^*(n)$ in terms of $||\tilde{e}(n)||^2$. However, when the QLMS has converged to the optimal solution, the orthogonality assumption between the error and the input allows us to drop these two terms, and (19) becomes

$$||\bar{e}(n)||^2 = ||\tilde{e}(n)||^2 - 2\mu(5||\tilde{e}(n)||^2 \sum_{\ell=1}^{L} ||\tilde{x}(n-\ell)||^2)$$
$$= \bigg(1 - 10\mu \sum_{\ell=1}^{L} ||\tilde{x}(n-\ell)||^2\bigg)||\tilde{e}(n)||^2 \quad (20)$$

which gives bounds for the stepsize parameter as

$$0 < \mu < \frac{1}{10\sum_{\ell=1}^{L} ||\tilde{x}(n-\ell)||^2} \quad (21)$$

Notice that $||\tilde{x}(n-\ell)||^2$ is the quaternion norm, and therefore it caters for all the four dimensions. This means that $\sum_{p=1}^{4} ||\mathbf{x}_p(n)||_2^2$ in the denominator of (14) is equivalent to $\sum_{\ell=1}^{L} ||\tilde{x}(n-\ell)||^2$ in the quaternion case (21).

The convergence analysis in (14) and (21) shows again that QLMS and MLMS do not have the same dynamic characteristics. Therefore, it is expected that QLMS and MLMS perform differently when the same parameters such as the filter length L, the stepsize μ, or the prediction horizon M are used. We shall next investigate their performance dependence of the performance of QLMS and MLMS on those parameters on both synthetic and real-world data.

4. SIMULATIONS

The performances of the multidimensional QLMS and MLMS algorithms were compared in a prediction setting. In the first experiment, the input was the 3D Lorenz system which is expressed mathematically as [11]

$$\frac{\partial x}{\partial t} = \alpha(y-x) \quad \frac{\partial y}{\partial t} = x(\rho-z) - y$$
$$\frac{\partial z}{\partial t} = xy - \beta z \quad (22)$$

where α, ρ, $\beta > 0$. For a chaotic behavior of Lorenz attractor, the parameters were selected as: $\alpha = 10$, $\rho = 28$, and $\beta = 8/3$.

The performance index was the standard prediction gain given by

$$R_p = 10\log(\sigma_{\mathbf{x}}^2(n)/\sigma_{\mathbf{e}}^2(n)) \quad (23)$$

where $\sigma_{\mathbf{x}}^2(n)$ and $\sigma_{\mathbf{e}}^2(n)$ denote respectively the variance of the input and the error. Fig. 1 illustrates the performances and confirms that the two algorithms do not exhibit similar performances, with QLMS outperforming MLMS. We next proceed to examining the performances of these algorithms on a 4D real-world wind data fusion model (3D wind field augmented with the fourth dimension as the wind temperature). Again, QLMS outperformed MLMS (see Fig. 2), over a range of parameters and prediction horizon M.

As our aim was to illustrate the difference between adaptive filtering in $\mathbb{R}^4$ and $\mathbb{H}$, the simulations were run on examples which, by design, facilitate QLMS. For space constraints, simulations for example, multichannel speech processing setting are omitted.
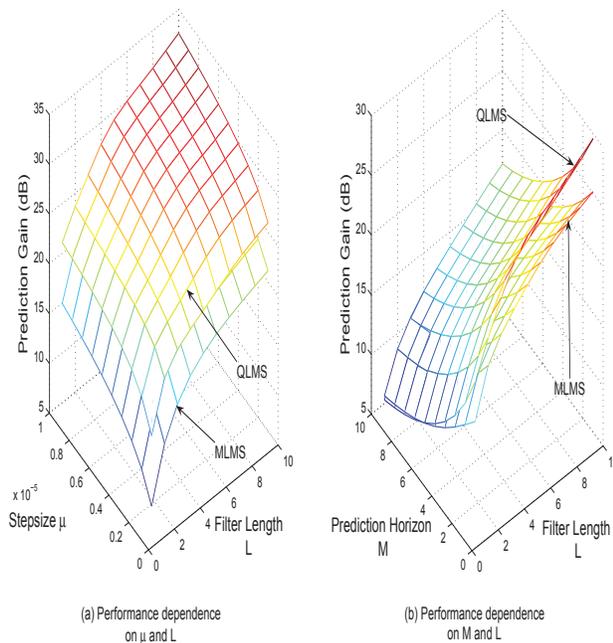
Fig. 1. The performance of Quaternion LMS and multi-channel LMS on the prediction of 3D Lorenz signal.

Fig. 2. The performance of Quaternion LMS and multi-channel LMS on the prediction of 3D wind field coupled with temperature.

5. CONCLUSIONS

We have provided a comparative analysis of the recently introduced quaternion LMS (QLMS) and the multichannel LMS (MLMS) algorithms. This has been achieved by both analysing the input/output mapping performed by the two filters and the dynamics of their updates. To this end, stability bounds for QLMS have been derived, and are valid for any QLMS formulation. It has been shown that the two algorithms are fundamentally different. MLMS has been derived for generic multichannel processes, whereas QLMS has been designed specifically for four dimensional processes. Simulation on four dimensional signals support the analysis.
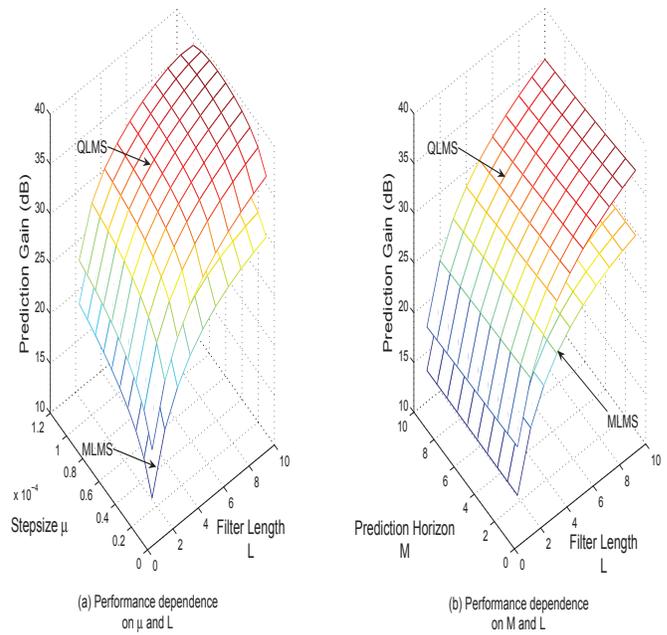
6. ACKNOWLEDGEMENT

7. REFERENCES

[1] Y. Huang and J. Benesty, *Audio Signal Processing for Next Generation Multimedia Communication Systems*, Kluwer Academic Publishers, 2004.

[2] P. J. Schreier and L. L. Scharf, "Second-order analysis of improper complex random vectors and processes," *IEEE Transactions on Signal Processing*, vol. 51, no. 3, pp. 714–725, 2003.

[3] T. Kim and T. Adali, "Approximation by fully complex multilayer perceptrons," *Neural Computation*, vol. 15, no. 7, pp. 1641 – 1666, 2003.

[4] D. P. Mandic and Vanessa Su Lee Goh, *Complex Valued Nonlinear Adaptive Filters: A Neural Network Approach*, John Wiley and Sons Ltd (to be published), 2008.

[5] A. Nehorai and E. Paldi, "Acoustic vector-sensor array processing," *IEEE Transactions on Signal Processing*, vol. 42, no. 9, pp. 2481–2491, 1994.

[6] W. Hamilton, *Elements of Quaternions*, Longmans, Green and Co, 1899.

[7] J. P. Ward, *Quaternions and Cayley Numbers: Algebra and Applications*, Kluwer Academic Publisher, 1997.

[8] C. Cheong-Took and D. Mandic, "The quaternion LMS algorithm for adaptive filtering of hypercomplex processes," *IEEE Transactions on Signal Processing [Accepted for future publication]*, Sept. 2008.

[9] C. Cheong-Took and D. Mandic, "Fusion of heterogeneous data sources: A quaternionic approach," *IEEE International Workshop on Machine Learning for Signal Processing*, Oct. 2008.

[10] P. A. Regalia, *Adaptive IIR Filtering in Signal Processing and Control*, Marcel Dekker, 1994.

[11] E. N. Lorenz, "Deterministic nonperiodic flow," *Journal of the Atmospheric Sciences*, vol. 20, no. 2, pp. 130141, 1963.