



Cite this article: Xu D, Gao H, Mandic DP. 2016

A new proof of the generalized
Hamiltonian–Real calculus. *R. Soc. open sci.*

3: 160211.

<http://dx.doi.org/10.1098/rsos.160211>

Received: 23 March 2016

Accepted: 8 August 2016

Subject Category:

Engineering

Subject Areas:

electrical engineering/systems
theory/mathematical modelling

Keywords:

GHR calculus, product and chain rules,
duality relationship, widely linear quaternion
least mean square

Author for correspondence:

Dongpo Xu

e-mail: dongpoxu@gmail.com


A new proof of the generalized Hamiltonian–Real calculus

Dongpo Xu¹, Hua Gao² and Danilo P. Mandic³

¹School of Mathematics and Statistics, Northeast Normal University,
Changchun 130024, People's Republic of China

²College of Science, Harbin Engineering University, Harbin 150001, People's Republic
of China

³Department of Electrical and Electronic Engineering, Imperial College London,
SW7 2AZ London, UK

 DX, 0000-0002-9663-9743

The recently introduced generalized Hamiltonian–Real (GHR) calculus comprises, for the first time, the product and chain rules that makes it a powerful tool for quaternion-based optimization and adaptive signal processing. In this paper, we introduce novel dual relationships between the GHR calculus and multivariate real calculus, in order to provide a new, simpler proof of the GHR derivative rules. This further reinforces the theoretical foundation of the GHR calculus and provides a convenient methodology for generic extensions of real- and complex-valued learning algorithms to the quaternion domain.

1. Introduction

Quaternions were introduced by William Hamilton in 1843, as an associative but not commutative algebra over reals, and have been used in many fields, such as physics, computer graphics and signal processing.

Because of the non-commutativity of quaternion product, definitions of quaternion derivative are very different from those for real and complex derivatives. For example, Sudbery [1] establishes that only linear quaternion functions fulfil the requirements of traditional derivative definition. However, such a derivative is too stringent for practical optimization problems, whereby the cost functions are often real-valued and therefore non-analytic. In order to relax the derivative condition, the recently introduced Hamiltonian–Real (HR) calculus [2] deals with both analytic and non-analytic quaternion functions within a unified framework. However, the HR calculus does not comprise effective derivative rules (product, chain) for dealing with complicated derivation operations in practical problems. By exploiting quaternion rotation, the recently introduced generalized HR (GHR) calculus embarks upon the

HR calculus to provide the powerful product and chain derivative rules that facilitate the calculation of quaternion gradient and Hessian in quaternion optimization problems [3]. In particular, we note that the product rule is a distinctive feature of the functional calculi. However, the original proof of the product and chain rules for the GHR calculus is quite involved and difficult to understand for non-experts.

The aim of this paper is to further elucidate the main ideas of the GHR calculus, and to give simpler proofs of the main results in [3]. The new proof is based on the duality between the GHR calculus and multivariate real calculus. One of the advantages of such an approach is that it allows for a direct treatment of quaternion-valued functions, without an intermediate transition to real functions. Based on the so-introduced relationships, we illustrate the effectiveness of such an approach by providing a dual version of the widely linear quaternion least mean square (WL-QLMS) adaptive learning algorithm [4,5], and show that it produces the same output as the primal WL-QLMS, but with a reduced computational complexity.

2. Preliminaries

In the following, we state some basic definitions of the GHR calculus [3,6,7].

Definition 2.1 (Quaternion rotation [8]). For any quaternion q , the quaternion rotation is defined as

$$q^\mu \triangleq \mu q \mu^{-1}, \quad (2.1)$$

where μ is any non-zero quaternion.

Definition 2.2 (Real-differentiability [1]). A function $f(q) = f_a(q) + if_b(q) + jf_c(q) + kf_d(q)$ is called real differentiable if its real-valued components $f_a(q)$, $f_b(q)$, $f_c(q)$ and $f_d(q)$ are differentiable functions with respect to four real components q_a, q_b, q_c and q_d of a quaternion variable $q = q_a + iq_b + jq_c + kq_d$.

Definition 2.3 (The GHR derivatives [3]). If $f: \mathbb{H} \rightarrow \mathbb{H}$ is real-differentiable, then, the GHR derivatives of the function f with respect to q^μ and $q^{\mu*}$ ($\mu \neq 0, \mu \in \mathbb{H}$) are defined as

$$\left. \begin{aligned} \frac{\partial f}{\partial q^\mu} &= \frac{1}{4} \left(\frac{\partial f}{\partial q_a} - \frac{\partial f}{\partial q_b} i^\mu - \frac{\partial f}{\partial q_c} j^\mu - \frac{\partial f}{\partial q_d} k^\mu \right) \\ \text{and} \quad \frac{\partial f}{\partial q^{\mu*}} &= \frac{1}{4} \left(\frac{\partial f}{\partial q_a} + \frac{\partial f}{\partial q_b} i^\mu + \frac{\partial f}{\partial q_c} j^\mu + \frac{\partial f}{\partial q_d} k^\mu \right) \end{aligned} \right\}, \quad (2.2)$$

where $q = q_a + iq_b + jq_c + kq_d$, quaternion components $q_a, q_b, q_c, q_d \in \mathbb{R}$, and $\partial f / \partial q_a, \partial f / \partial q_b, \partial f / \partial q_c$ and $\partial f / \partial q_d$ are the partial derivatives of f with respect to q_a, q_b, q_c and q_d , whereas the set $\{1, i^\mu, j^\mu, k^\mu\}$ is an orthogonal basis of \mathbb{H} .

Lemma 2.4 (Duality of quaternions and real vectors [9,10]). For any quaternion $q = q_a + iq_b + jq_c + kq_d$, where $q_l \in \mathbb{R}, l \in \{a, b, c, d\}$, the relationship between the augmented quaternion vector $\mathbf{q}^a = (q, q^i, q^j, q^k)^T \in \mathbb{H}^4$ and its quadrivariate real vector counterpart $\mathbf{q}^r = (q_a, q_b, q_c, q_d)^T \in \mathbb{R}^4$ is given by $\mathbf{q}^a = \mathbf{A} \mathbf{q}^r$, where

$$\mathbf{A} = \begin{pmatrix} 1 & i & j & k \\ 1 & i & -j & -k \\ 1 & -i & j & -k \\ 1 & -i & -j & k \end{pmatrix}, \quad \mathbf{A}^{-1} = \frac{1}{4} \mathbf{A}^H \quad (2.3)$$

and $(\cdot)^H$ denotes the quaternion conjugate transpose.

Lemma 2.5 (Duality of quaternion gradients and real gradient vectors [7]). For a real-differentiable function $f: \mathbb{H} \rightarrow \mathbb{R}$, the relation between the quadrivariate real gradient vector $\nabla_{\mathbf{q}}^r f = (\partial f / \partial q_a, \partial f / \partial q_b, \partial f / \partial q_c, \partial f / \partial q_d)^T$ and the augmented quaternion gradient vector $\nabla_{\mathbf{q}^a}^a f = (\partial f / \partial q^*, \partial f / \partial q^{i*}, \partial f / \partial q^{j*}, \partial f / \partial q^{k*})^T$ is given by $\nabla_{\mathbf{q}^a}^a f = \mathbf{A}^H \nabla_{\mathbf{q}^r}^r f$.

Lemma 2.6 (Duality of quaternion and real Jacobian matrices). For a real-differentiable function $f: \mathbb{H} \rightarrow \mathbb{H}$, the relation between the quadrivariate real Jacobian matrix $\mathbf{J}_q^r(f)$ and the augmented quaternion Jacobian

matrix $\mathbf{J}_q^a(f)$ is given by $\mathbf{J}_q^r(f) = \mathbf{A}^{-1}\mathbf{J}_{q^\mu}^a(f)\mathbf{A}^\mu$, where

$$\mathbf{J}_q^r(f) = \begin{pmatrix} \frac{\partial f_a}{\partial q_a} & \frac{\partial f_a}{\partial q_b} & \frac{\partial f_a}{\partial q_c} & \frac{\partial f_a}{\partial q_d} \\ \frac{\partial f_b}{\partial q_a} & \frac{\partial f_b}{\partial q_b} & \frac{\partial f_b}{\partial q_c} & \frac{\partial f_b}{\partial q_d} \\ \frac{\partial f_c}{\partial q_a} & \frac{\partial f_c}{\partial q_b} & \frac{\partial f_c}{\partial q_c} & \frac{\partial f_c}{\partial q_d} \\ \frac{\partial f_d}{\partial q_a} & \frac{\partial f_d}{\partial q_b} & \frac{\partial f_d}{\partial q_c} & \frac{\partial f_d}{\partial q_d} \end{pmatrix} \quad \text{and} \quad \mathbf{J}_{q^\mu}^a(f) = \begin{pmatrix} \frac{\partial f}{\partial q^\mu} & \frac{\partial f}{\partial q^{\mu i}} & \frac{\partial f}{\partial q^{\mu j}} & \frac{\partial f}{\partial q^{\mu k}} \\ \frac{\partial f^i}{\partial q^\mu} & \frac{\partial f^i}{\partial q^{\mu i}} & \frac{\partial f^i}{\partial q^{\mu j}} & \frac{\partial f^i}{\partial q^{\mu k}} \\ \frac{\partial f^j}{\partial q^\mu} & \frac{\partial f^j}{\partial q^{\mu i}} & \frac{\partial f^j}{\partial q^{\mu j}} & \frac{\partial f^j}{\partial q^{\mu k}} \\ \frac{\partial f^k}{\partial q^\mu} & \frac{\partial f^k}{\partial q^{\mu i}} & \frac{\partial f^k}{\partial q^{\mu j}} & \frac{\partial f^k}{\partial q^{\mu k}} \end{pmatrix} \quad (2.4)$$

with $f(q) = f_a(q) + if_b(q) + jf_c(q) + kf_d(q)$, $q = q_a + iq_b + jq_c + kq_d$ and $q_l \in \mathbb{R}, f_l(q) \in \mathbb{R}, l \in \{a, b, c, d\}$.

Proof. The proof follows directly from definition 2.3. ■

Lemma 2.7 (Duality of quaternion and real Hessian matrices [7]). For a real-differentiable function $f : \mathbb{H} \rightarrow \mathbb{R}$, the relation between the quadivariate real Hessian matrix $\mathbf{H}_q^r(f)$ and the augmented quaternion Hessian matrix $\mathbf{H}_q^a(f)$ is given by $\mathbf{H}_q^r(f) = \mathbf{A}^H \mathbf{H}_q^a(f) \mathbf{A}$.

3. Main results

Theorem 3.1 (Chain rule of GHR derivatives). Let $S \subseteq \mathbb{H}$ and let $g : S \rightarrow \mathbb{H}$ be real-differentiable at an interior point q of the set S . Let $T \subseteq \mathbb{H}$ be such that $g(q) \in T$ for all $q \in S$. Assume that $f : T \rightarrow \mathbb{H}$ is real-differentiable at an interior point $g(q) \in T$. Then, the composite function $h(q) = f(g(q))$ satisfies the following chain rule

$$\frac{\partial h(q)}{\partial q^\mu} = \frac{\partial f}{\partial g} \frac{\partial g}{\partial q^\mu} + \frac{\partial f}{\partial g^i} \frac{\partial g^i}{\partial q^\mu} + \frac{\partial f}{\partial g^j} \frac{\partial g^j}{\partial q^\mu} + \frac{\partial f}{\partial g^k} \frac{\partial g^k}{\partial q^\mu}, \quad (3.1)$$

where $\mu \in \mathbb{H}, \mu \neq 0, \partial f / \partial g^v = \partial f(g) / \partial g^v$ ($v = 1, i, j, k$) are the GHR derivatives.

Proof. Using the chain rule of multivariate real calculus, we have

$$\mathbf{J}_q^r(h) = \mathbf{J}_g^r(h) \mathbf{J}_q^r(g). \quad (3.2)$$

Multiplying both sides of (3.2) by \mathbf{A} and $(\mathbf{A}^\mu)^{-1}$, we have

$$\mathbf{A} \mathbf{J}_q^r(h) (\mathbf{A}^\mu)^{-1} = \mathbf{A} \mathbf{J}_g^r(h) \mathbf{A}^{-1} \mathbf{A} \mathbf{J}_q^r(g) (\mathbf{A}^\mu)^{-1}. \quad (3.3)$$

Using lemma 2.6 further yields

$$\mathbf{J}_{q^\mu}^a(h) = \mathbf{J}_g^a(h) \mathbf{J}_{q^\mu}^a(g). \quad (3.4)$$

Upon extracting the top left entry (i.e. (1,1)-entry) of the matrix equation (3.4), we arrive at

$$\frac{\partial f(g(q))}{\partial q^\mu} = \frac{\partial f(g)}{\partial g} \frac{\partial g(q)}{\partial q^\mu} + \frac{\partial f(g)}{\partial g^i} \frac{\partial g^i(q)}{\partial q^\mu} + \frac{\partial f(g)}{\partial g^j} \frac{\partial g^j(q)}{\partial q^\mu} + \frac{\partial f(g)}{\partial g^k} \frac{\partial g^k(q)}{\partial q^\mu}. \quad (3.5)$$

This completes the proof of theorem 3.1. ■

Theorem 3.2 (Product rule of GHR derivatives). If the functions $f, g : \mathbb{H} \rightarrow \mathbb{H}$ are real-differentiable, then the derivative of their product fg satisfies the following product rule

$$\frac{\partial (fg)}{\partial q^\mu} = f \frac{\partial g}{\partial q^\mu} + \frac{\partial f}{\partial q^{\mu \nu}} g, \quad (3.6)$$

where $\mu \in \mathbb{H}, \mu \neq 0, \partial f / \partial q^{\mu \nu}$ can be obtained by replacing μ with $g\mu$ in definition 2.3.

Proof. Denote $h(q) = f(q)g(q)$, $f = f_a + if_b + jf_c + kf_d$ and $g = g_a + ig_b + jg_c + kg_d$. Then,

$$\mathbf{h}^r = \begin{pmatrix} h_a \\ h_b \\ h_c \\ h_d \end{pmatrix} = \begin{pmatrix} f_a g_a - f_b g_b - f_c g_c - f_d g_d \\ f_a g_b + f_b g_a + f_c g_d - f_d g_c \\ f_a g_c - f_b g_d + f_c g_a + f_d g_b \\ f_a g_d + f_b g_c - f_c g_b + f_d g_a \end{pmatrix}. \tag{3.7}$$

Using the product rule of real calculus yields

$$\mathbf{J}_q^r(h) = \mathbf{F}\mathbf{J}_q^r(g) + \mathbf{G}\mathbf{J}_q^r(f), \tag{3.8}$$

where

$$\mathbf{F} = \begin{pmatrix} f_a & -f_b & -f_c & -f_d \\ f_b & f_a & -f_d & f_c \\ f_c & f_d & f_a & -f_b \\ f_d & -f_c & f_b & f_a \end{pmatrix} \quad \text{and} \quad \mathbf{G} = \begin{pmatrix} g_a & -g_b & -g_c & -g_d \\ g_b & g_a & g_d & -g_c \\ g_c & -g_d & g_a & g_b \\ g_d & g_c & -g_b & g_a \end{pmatrix} \tag{3.9}$$

Upon multiplying both sides of (3.9) by \mathbf{A} and $(\mathbf{A}^\mu)^{-1}$ and using lemma 2.6, we have

$$\begin{aligned} \mathbf{J}_{q^\mu}^a(h) &= \mathbf{A}\mathbf{J}_q^r(h)(\mathbf{A}^\mu)^{-1} = \mathbf{A}\mathbf{F}\mathbf{A}^{-1}\mathbf{A}\mathbf{J}_q^r(g)(\mathbf{A}^\mu)^{-1} + \mathbf{A}\mathbf{G}\mathbf{J}_q^r(f)(\mathbf{A}^\mu)^{-1} \\ &= \mathbf{F}^a\mathbf{J}_{q^\mu}^a(g) + (\mathbf{A}^T\mathbf{G}^a)^T\mathbf{J}_q^r(f)(\mathbf{A}^\mu)^{-1}, \end{aligned} \tag{3.10}$$

where

$$\mathbf{F}^a \triangleq \mathbf{A}\mathbf{F}\mathbf{A}^{-1} = \begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f^i & 0 & 0 \\ 0 & 0 & f^j & 0 \\ 0 & 0 & 0 & f^k \end{pmatrix} \quad \text{and} \quad \mathbf{G}^a \triangleq \begin{pmatrix} g & 0 & 0 & 0 \\ 0 & g^i & 0 & 0 \\ 0 & 0 & g^j & 0 \\ 0 & 0 & 0 & g^k \end{pmatrix} \tag{3.11}$$

Upon extracting the first row of the matrix equation (3.10), we arrive at

$$\begin{aligned} (\mathbf{J}_{q^\mu}^a(h))_1 &= f(\mathbf{J}_{q^\mu}^a(g))_1 + (\mathbf{A})_1\mathbf{G}\mathbf{J}_q^r(f)(\mathbf{A}^\mu)^{-1} = f(\mathbf{J}_{q^\mu}^a(g))_1 + (\mathbf{A})_1\mathbf{J}_q^r(f)g(\mathbf{A}^\mu)^{-1} \\ &= f(\mathbf{J}_{q^\mu}^a(g))_1 + (\mathbf{A})_1\mathbf{J}_q^r(f)(\mathbf{A}^{g^\mu})^{-1}g = f(\mathbf{J}_{q^\mu}^a(g))_1 + (\mathbf{J}_{q^{g^\mu}}^a(f))_1g, \end{aligned} \tag{3.12}$$

where $(\mathbf{X})_1$ denotes the first row of the matrix \mathbf{X} . The first element of the above vector equation finally yields

$$\frac{\partial h}{\partial q^\mu} = f \frac{\partial g}{\partial q^\mu} + \frac{\partial f}{\partial q^{g^\mu}} g. \tag{3.13}$$

This completes the proof of theorem 3.2. ■

4. Application example

The WL-QLMS algorithm is based on the quaternion widely linear model $y(n) = \mathbf{w}^T(n) \mathbf{q}(n)$, where $\mathbf{q}(n) = (\mathbf{x}^T(n), \mathbf{x}^{iT}(n), \mathbf{x}^{jT}(n), \mathbf{x}^{kT}(n))^T \in \mathbb{H}^{4N}$ is the augmented input vector and $\mathbf{w}(n) \in \mathbb{H}^{4N}$ the associated weight (parameter) vector. The cost function to be minimized is a real-valued function of quaternion variables, given by

$$J(n) = |e(n)|^2 = |d(n) - y(n)|^2, \tag{4.1}$$

where $e(n) = d(n) - y(n)$ is the error between the desired signal $d(n)$ and the filter output $y(n)$. The weight update of WL-QLMS is then given by [3]

$$\mathbf{w}(n + 1) = \mathbf{w}(n) + \alpha e(n) \mathbf{q}^*(n), \tag{4.2}$$

where $\alpha > 0$ is the step size. Similar to the scalar duality in lemma 2.4, the duality between the augmented quaternion vector $\mathbf{q}(n)$ and its dual quadrivariate real vector $\mathbf{r}(n) = (\mathbf{x}_a^T(n), \mathbf{x}_b^T(n), \mathbf{x}_c^T(n), \mathbf{x}_d^T(n))^T \in \mathbb{R}^{4N}$ is

Table 1. Computational complexity in terms of real operations per iteration (N is the filter length).

algorithm	+	×
QLMS	$32N$	$32N + 4$
WL-QLMS	$128N$	$128N + 4$
D-WL-QLMS	$32N$	$32N + 4$

given by [7]

$$\mathbf{q}(n) = \mathbf{J}\mathbf{r}(n), \quad \mathbf{q}^H(n) = \mathbf{r}^H(n)\mathbf{J}^H = \mathbf{r}^T(n)\mathbf{J}^H, \tag{4.3}$$

where $\mathbf{J} = \mathbf{A} \otimes \mathbf{I}_N$ denotes the Kronecker product of \mathbf{A} (cf. (2.3)) and the $N \times N$ identity matrix \mathbf{I}_N . The filter output $y(n)$ can now be rewritten as

$$y(n) = \mathbf{w}^T(n)\mathbf{q}(n) = \mathbf{w}^T(n)\mathbf{J}\mathbf{r}(n) = \mathbf{v}^T(n)\mathbf{r}(n), \tag{4.4}$$

where $\mathbf{v}^T(n) = \mathbf{w}^T(n)\mathbf{J} \in \mathbb{H}^{4N}$ is an alternative augmented quaternion weight vector. Multiplying both sides of (4.2) by \mathbf{J} and noting that $\mathbf{J}^H\mathbf{J} = 4\mathbf{I}_{4N}$, we have

$$\begin{aligned} \mathbf{w}^T(n+1)\mathbf{J} &= \mathbf{w}^T(n)\mathbf{J} + \alpha e(n)\mathbf{q}^H(n)\mathbf{J} = \mathbf{w}^T(n)\mathbf{J} + \alpha e(n)\mathbf{r}^T(n)\mathbf{J}^H\mathbf{J} \\ &= \mathbf{w}^T(n)\mathbf{J} + 4\alpha e(n)\mathbf{r}^T(n). \end{aligned} \tag{4.5}$$

Upon replacing $\mathbf{w}^T(n)\mathbf{J}$ with $\mathbf{v}^T(n)$, the dual version of the WL-QLMS (D-WL-QLMS) algorithm becomes

$$\mathbf{v}(n+1) = \mathbf{v}(n) + \alpha e(n)\mathbf{r}(n), \tag{4.6}$$

where the constant 4 in (4.5) is absorbed into the step size α .

Remark 4.1. From (4.4), we can see that the D-WL-QLMS and WL-QLMS have the same filter output. Therefore, the D-WL-QLMS has the same performance as the WL-QLMS, which is better than the strictly linear QLMS and the real LMS (RLMS) [5,11], when dealing with non-circular inputs. Table 1 shows that D-WL-QLMS also has a lower computational complexity than the WL-QLMS, owing to the use of real-valued input vector $\mathbf{r}(n)$ in the calculations. This is similar to the operation of the RLMS; however, the weight vector $\mathbf{v}(n)$ of the D-WL-QLMS is quaternion-valued, which is different from the RLMS.

Remark 4.2. Note that if we start from $y(n) = \mathbf{w}^H(n)\mathbf{q}(n)$, then an alternative form of the WL-QLMS is given by $\mathbf{w}(n+1) = \mathbf{w}(n) + \alpha \mathbf{q}(n)e^*(n)$ [3]. Denote $\mathbf{v}^T(n) = \mathbf{w}^H(n)\mathbf{J}$, then the filter output becomes $y(n) = \mathbf{w}^H(n)\mathbf{J} \mathbf{r}(n) = \mathbf{v}^T(n)\mathbf{r}(n)$, that is, the same as in (4.4). Finally, the proposed D-WL-QLMS keeps the same form of update rule in (4.6), and has the same computational complexity as that shown in table 1.

5. Conclusion

We have provided a new proof for the GHR calculus through the duality relationships between the GHR calculus and multivariate real calculus. These results complement the original proof given in [3], are easier to understand and are physically meaningful, and thus provide additional insights into the operation of the GHR calculus. An application example in adaptive learning theory demonstrates the advantages of the proposed approach.

Authors' contributions. D.X. designed the study, carried out the analysis and drafted the manuscript; H.G. checked the analysis results and helped with the WL-QLMS algorithm; D.P.M. conceived the study and helped with the analysis and drafting the manuscript. All authors gave final approval for publication.

Competing interests. The authors have no competing interests.

Funding. D.X. was supported by the National Natural Science Foundation of China (no. 61301202), by the Fundamental Research Funds for the Central Universities (no. 2412016KJ014) and by Science and Technology Development Planning of Jilin Province. D.P.M. was partially supported by EPSRC grant no. EP/H026266/1.

Acknowledgements. The authors thank the anonymous reviewers for their very helpful and insightful suggestions, and their expert advice on how to improve the proof of the two theorems.

References

1. Sudbery A. 1979 Quaternionic analysis. *Math. Proc. Camb. Phil. Soc.* **85**, 199–225. (doi:10.1017/S0305004100055638)
2. Mandic DP, Jahanchahi C, Took CC. 2011 A quaternion gradient operator and its applications. *IEEE Signal Proc. Lett.* **18**, 47–50. (doi:10.1109/LSP.2010.2091126)
3. Xu D, Jahanchahi C, Took CC, Mandic DP. 2015 Enabling quaternion derivatives: the generalized HR calculus. *R. Soc. open sci.* **2**, 150255. (doi:10.1098/rsos.150255)
4. Took CC, Mandic DP. 2010 A quaternion widely linear adaptive filter. *IEEE Trans. Signal Process.* **58**, 4427–4431. (doi:10.1109/TSP.2010.2048323)
5. Neto FGA, Nascimento VH. 2011 A novel reduced-complexity widely linear QLMS algorithm. *Statistical Signal Processing Workshop, Nice, France, June*, pp. 81–84. IEEE. (doi:10.1109/SSP.2011.5967831)
6. Xu D, Mandic DP. 2015 The theory of quaternion matrix derivatives. *IEEE Trans. Signal Process.* **63**, 1543–1556. (doi:10.1109/TSP.2015.2399865)
7. Xu D, Xia Y, Mandic DP. 2016 Optimization in quaternion dynamic systems: gradient, Hessian, and learning algorithms. *IEEE Trans. Neural Netw. Learn. Syst.* **27**, 249–261. (doi:10.1109/TNNLS.2015.2440473)
8. Ward JP. 1997 *Quaternions and Cayley numbers: algebra and applications*. London, UK: Kluwer Academic Publishers.
9. Via J, Ramirez D, Santamaria I. 2010 Properness and widely linear processing of quaternion random vectors. *IEEE Trans. Inf. Theory* **56**, 3502–3515. (doi:10.1109/TIT.2010.2048440)
10. Took CC, Mandic DP. 2011 Augmented second order statistics of quaternion random signals. *Signal Process.* **91**, 214–224. (doi:10.1016/j.sigpro.2010.06.024)
11. Xia Y, Jahanchahi C, Nitta T, Mandic DP. 2015 Performance bounds of quaternion estimators. *IEEE Trans. Neural Netw. Learn. Syst.* **26**, 3287–3292. (doi:10.1109/TNNLS.2015.2388782)