**Article:**

Jiang, M., Li, Y. and Liu, W. (2014) Properties and Applications of a Restricted HR Gradient Operator. (Unpublished)

# Properties and Applications of a Restricted HR Gradient Operator

Mengdi Jiang [a], Yi Li [b], Wei Liu [a]

[a] Communications Research Group, Department of Electronic and Electrical Engineering
University of Sheffield, Sheffield, S1 3JD, United Kingdom
{mjiang3, w.liu}@sheffield.ac.uk
[b] School of Mathematics and Statistics
University of Sheffield, Sheffield, S3 7RH, United Kingdom
yili@sheffield.ac.uk



***Abstract*—For quaternionic signal processing algorithms, the gradients of a quaternion-valued function are required for gradient-based methods. Given the non-commutativity of quaternion algebra, the definition of the gradients is non-trivial. The HR gradient operator provides a viable framework and has found a number of applications. However, the applications so far have been mainly limited to real-valued quaternion functions and linear quaternion-valued functions. To generalize the operator to nonlinear quaternion functions, we define a restricted version of the HR operator. The restricted HR gradient operator comes in two versions, the left and the right ones. We then present a detailed analysis of the properties of the operators, including several different product rules and chain rules. Using the new rules, we derive explicit expressions for the derivatives of a class of regular nonlinear quaternion-valued functions, and prove that the restricted HR gradients are consistent with the gradients in real domain.**

## I. Introduction

Quaternion calculus has been introduced in signal processing with application areas involving three or four-dimensional signals, such as color image processing [1]–[3], vector-sensor array systems [4]–[10] and wind profile prediction [11], [12]. Several quaternion-valued adaptive filtering algorithms have been proposed in [9]–[13]. Notwithstanding the advantages of the quaternionic algorithms, extra cares have to be taken in their developments, in particular when the derivatives of quaternion-valued functions are involved, due to the fact that quaternion algebra is non-commutative. A so-called HR gradient operator was proposed in [14], and has been applied in [15]. The interesting formulation appears to provide a general and flexible framework that could potentially have wide applications. However, it has only been applied to real-valued functions and linear quaternion-valued functions. In order to consider more general quaternion-valued functions, we propose a pair of restricted HR gradient operators, the left and the right restricted HR gradient operators, based on the previous work on the HR gradient operator [14] and our recent work [12].

To summarize, we make the following main contributions. Firstly, we give a detailed derivation of the relation between the gradients and the increment of a quaternion function, highlighting the difference between the left and the right gradients due to the non-commutativity of quaternion algebra. Secondly, we document several properties of the operators that have not been reported before, in particular several different versions of product rules and chain rules. Thirdly, we derive a general formula for the restricted HR derivatives of a wide class of regular quaternion-valued nonlinear functions, among which are the exponential, logarithmic, and the hyperbolic tangent functions. Finally, we prove that the restricted HR gradients are consistent with the usual definition for the gradient of a real function of a real variable. Its application to the derivation of quaternion-valued least mean squares (QLMS) adaptive algorithm is also briefly discussed.

The paper is organised as follows. The restricted HR gradient operator is developed in Section II, with its properties and rules introduced in Section III. Explicit expressions for the derivatives for a wide range of functions are derived in Section IV and results for the right restricted HR operator are summarised in Section V. The increment of a general quaternion function is discussed in Section VI with the QLMS adaptive algorithm revisited as a special case where the cost function is real-valued. Conclusions are drawn in Section VII.

## II. The restricted HR gradient operators

### A. Introduction of quaternion

Quaternion is a non-commutative extension of complex number. A quaternion $q$ is composed of four parts, i.e., $q = q_a + q_b i + q_c j + q_d k$, where $q_a$ is the real part, which is also denoted as $R(q)$. The other three terms constitute the imaginary part $I(q)$. $i$, $j$ and $k$ are the three imaginary units, which satisfy the following rules for multiplication: $ij = k$, $jk = i$, $ki = j$, $i^2 = j^2 = k^2 = -1$, and

$$ij = -ji, ki = -ik, kj = -jk. \quad (1)$$

Due to (1), in general the product of two quaternions depends on the order, i.e., $qp \neq pq$ where $p$ and $q$ are quaternions. However, the product commutes as long as at least one of the factors, say $q$, is real.

Let $v = |I(q)|$ and $\hat{\mathbf{v}} = I(q)/v$, the quaternion $q$ can also be written as $q = q_a + v\hat{\mathbf{v}}$. $\hat{\mathbf{v}}$ is a pure unit quaternion, which has the convenient property $\hat{\mathbf{v}}^2 := \hat{\mathbf{v}}\hat{\mathbf{v}} = -1$. The quaternionic conjugate of $q$ is $q^* = q_a - q_b i - q_c j - q_d k$, or $q^* = q_a - v\hat{\mathbf{v}}$. It is easy to show that $qq^* = q^*q = |q|^2$, and hence $q^{-1} = q^*/|q|^2$.

### B. Definition of the restricted HR gradient operators

Let $f : H \to H$ be a quaternion-valued function of a quaternion $q$, where $H$ is the non-commutative algebra of quaternions. We use the notation $f(q) = f_a + f_b i + f_c j + f_d k$, where $f_a, ..., f_d$ are the components of $f$. $f$ can also be viewed as a function of the four components of $q$, i.e., $f = f(q_a, q_b, q_c, q_d)$. In this view $f$ is a quaternion-valued function on $R^4$: $f : R^4 \to H$. To express the four real components of $q$, it is convenient to use its involutions $q^\nu := -\nu q \nu$ where $\nu \in \{i, j, k\}$ [16]. Explicitly, we have

$$q^i = -iqi = q_a + q_b i - q_c j - q_d k, \tag{2}$$
$$q^j = -jqj = q_a - q_b i + q_c j - q_d k, \tag{3}$$
$$q^k = -kqk = q_a - q_b i - q_c j + q_d k. \tag{4}$$

The real components can be recovered by

$$q_a = \frac{1}{4}(q + q^i + q^j + q^k), q_b = \frac{1}{4i}(q + q^i - q^j - q^k), \tag{5}$$
$$q_c = \frac{1}{4j}(q - q^i + q^j - q^k), q_d = \frac{1}{4k}(q - q^i - q^j + q^k). \tag{6}$$

Two useful relations are

$$q^* = \frac{1}{2}(q^i + q^j + q^k - q), \ \ q + q^i + q^j + q^k = 4R(q). \tag{7}$$

A so-called HR gradient of $f(q)$ was introduced in [14], which has been applied to real-valued functions and linear quaternion-valued functions. In order to find the gradients of more general quaternion-valued functions, we follow a similar approach to propose a 'restricted' HR gradient operator (some of the derivation was first presented in [12]). To motivate the definitions, we consider the differential $df(q)$ with respect to differential $dq := dq_a + dq_b i + dq_c j + dq_d k$. We observe that $df = df_a + idf_b + jdf_c + kdf_d$, where

$$df_a = \frac{\partial f_a}{\partial q_a} dq_a + \frac{\partial f_a}{\partial q_b} dq_b + \frac{\partial f_a}{\partial q_c} dq_c + \frac{\partial f_a}{\partial q_d} dq_d. \tag{8}$$

We have $dq_a = (dq + dq^i + dq^j + dq^k)/4$ according to (5). Making use of this and similar expressions for $dq_b$, $dq_c$ and $dq_d$, we find an expression for $df_a$ in terms of the differentials $dq$, $dq^i$, $dq^j$ and $dq^k$. Repeating the calculation for $idf_b$, $jdf_c$ and $kdf_d$, we finally arrive at

$$df = Ddq + D_i dq^i + D_j dq^j + D_k dq^k \tag{9}$$

where

$$D := \frac{1}{4}\left(\frac{\partial f}{\partial q_a} - \frac{\partial f}{\partial q_b} i - \frac{\partial f}{\partial q_c} j - \frac{\partial f}{\partial q_d} k\right), \tag{10}$$
$$D_i := \frac{1}{4}\left(\frac{\partial f}{\partial q_a} - \frac{\partial f}{\partial q_b} i + \frac{\partial f}{\partial q_c} j + \frac{\partial f}{\partial q_d} k\right), \tag{11}$$
$$D_j := \frac{1}{4}\left(\frac{\partial f}{\partial q_a} + \frac{\partial f}{\partial q_b} i - \frac{\partial f}{\partial q_c} j + \frac{\partial f}{\partial q_d} k\right), \tag{12}$$
$$D_k := \frac{1}{4}\left(\frac{\partial f}{\partial q_a} + \frac{\partial f}{\partial q_b} i + \frac{\partial f}{\partial q_c} j - \frac{\partial f}{\partial q_d} k\right). \tag{13}$$

More details are given in Appendix A. Thus one may define the partial derivatives of $f(q)$ as follows:

$$\frac{\partial f}{\partial q} := D, \quad \frac{\partial f}{\partial q^i} := D_i, \frac{\partial f}{\partial q^j} := D_j, \frac{\partial f}{\partial q^k} := D_k. \tag{14}$$

Introducing operators $\nabla_q := (\partial/\partial q, \partial/\partial q^i, \partial/\partial q^j, \partial/\partial q^k)$, and $\nabla_r := (\partial/\partial q_a, \partial/\partial q_b, \partial/\partial q_c, \partial/\partial q_d)$, equations (10-14) may be written as

$$\nabla_q f = \nabla_r f J^H \tag{15}$$

where the Jacobian matrix

$$J = \frac{1}{4}\begin{bmatrix} 1 & i & j & k \\ 1 & i & -j & -k \\ 1 & -i & j & -k \\ 1 & -i & -j & k \end{bmatrix} \tag{16}$$

and $J^H$ is the Hermitian transpose of $J$ [14]. Using $JJ^H = J^H J = 1/4$ [15], we may also write

$$\nabla_q f J = \frac{1}{4}\nabla_r f, \tag{17}$$

which is the inverse formulae for the derivatives.

We call the gradient operator defined by (15) the restricted HR gradient operator. The operator is closely related to the HR operator introduced in [14]. However, in the original definition of the HR operator, the Jacobian $J$ appears on the left-hand side of $\nabla_r f$, whereas in our definition it appears on the right (as the Hermitian transpose).

The differential $df$ is related to $\nabla_q f$ by

$$df = \frac{\partial f}{\partial q} dq + \frac{\partial f}{\partial q^i} dq^i + \frac{\partial f}{\partial q^j} dq^j + \frac{\partial f}{\partial q^k} dq^k. \tag{18}$$

Due to the non-commutativity of quaternion products, the order of the factors in the products of the above equation (as well as equations (10-13)) can not be swapped. In fact, one may call the above operator the left restricted HR gradient operator. As is shown in Appendix A, one may also define a right restricted HR gradient operator by

$$(\nabla_q^R f)^T := J^* (\nabla_r f)^T, \tag{19}$$

where

$$\nabla_q^R := (\partial^R/\partial q, \partial^R/\partial q^i, \partial^R/\partial q^j, \partial^R/\partial q^k),$$

and

$$\frac{\partial^R f}{\partial q} := \frac{1}{4}\left(\frac{\partial f}{\partial q_a} - i\frac{\partial f}{\partial q_b} - j\frac{\partial f}{\partial q_c} - k\frac{\partial f}{\partial q_d}\right), \tag{20}$$
$$\frac{\partial^R f}{\partial q^i} := \frac{1}{4}\left(\frac{\partial f}{\partial q_a} - i\frac{\partial f}{\partial q_b} + j\frac{\partial f}{\partial q_c} + k\frac{\partial f}{\partial q_d}\right), \tag{21}$$
$$\frac{\partial^R f}{\partial q^j} := \frac{1}{4}\left(\frac{\partial f}{\partial q_a} + i\frac{\partial f}{\partial q_b} - j\frac{\partial f}{\partial q_c} + k\frac{\partial f}{\partial q_d}\right), \tag{22}$$
$$\frac{\partial^R f}{\partial q^k} := \frac{1}{4}\left(\frac{\partial f}{\partial q_a} + i\frac{\partial f}{\partial q_b} + j\frac{\partial f}{\partial q_c} - k\frac{\partial f}{\partial q_d}\right). \tag{23}$$

The right restricted HR gradient operator is related to the differential $df$ by

$$df = dq\frac{\partial^R f}{\partial q} + dq^i\frac{\partial^R f}{\partial q^i} + dq^j\frac{\partial^R f}{\partial q^j} + dq^k\frac{\partial^R f}{\partial q^k}. \tag{24}$$

In general, the left and right restricted HR gradients are not the same. For example, even for the simplest linear function $f(q) = q_0 q$ with $q_0 \in H$ a constant, we have

$$\frac{\partial q_0 q}{\partial q} = q_0, \ \frac{\partial^R q_0 q}{\partial q} = R(q_0). \tag{25}$$

However, we will show later that the two gradients coincide for a class of functions. In particular, they are the same for real-valued quaternion functions.

The relation between the gradients and the differential is an important ingredient of gradient-based methods, which we will discuss further later.

## III. Properties and rules of the operator

We will now focus on the left restricted HR gradient and simply call it the restricted HR gradient unless stated otherwise. It can be easily calculated from the definitions, that

$$\frac{\partial q}{\partial q} = 1, \ \frac{\partial q^\nu}{\partial q} = 0, \ \frac{\partial q^*}{\partial q} = -\frac{1}{2}, \tag{26}$$

where $\nu \in \{i, j, k\}$. However, in order to find the derivatives for more complex quaternion functions, it is useful to first establish the rules of the gradient operators. We will see that some of the usual rules do not apply due to the non-commutativity of quaternion products.

1) Left-linearity: for arbitrary constant quaternions $\alpha$ and $\beta$, and functions $f(q)$ and $g(q)$, we have

$$\frac{\partial(\alpha f + \beta g)}{\partial q^\nu} = \alpha\frac{\partial f}{\partial q^\nu} + \beta\frac{\partial g}{\partial q^\nu} \tag{27}$$

for $\nu \in \{1, i, j, k\}$ with $q^1 := q$. However, linearity does not hold for right multiplications, i.e., in general

$$\frac{\partial f\alpha}{\partial q} \neq \frac{\partial f}{\partial q}\alpha. \tag{28}$$

This is because, according to the definition (10),

$$\frac{\partial f\alpha}{\partial q} = \frac{1}{4}\sum_{(\phi,\gamma)}\frac{\partial f}{\partial q_\phi}\alpha\gamma \tag{29}$$

for $(\phi, \gamma) \in \{(a, 1), (b, -i), (c, -j), (d, -k)\}$. However, $\alpha\gamma \neq \gamma\alpha$ in general. Therefore it is different from $(\partial f/\partial q)\alpha$, which is

$$\frac{1}{4}\left(\frac{\partial f}{\partial q_a} - \frac{\partial f}{\partial q_b}i - \frac{\partial f}{\partial q_c}j - \frac{\partial f}{\partial q_d}k\right)\alpha. \tag{30}$$

2) The first product rule: the following product rule holds:

$$\nabla_q(fg) = f\nabla_q g + [(\nabla_r f)g]J^H. \tag{31}$$

For example,

$$\frac{\partial fq}{\partial q} = f\frac{\partial g}{\partial q} + \frac{1}{4}\left(\frac{\partial f}{\partial q_a}g - \frac{\partial f}{\partial q_b}gi - \frac{\partial f}{\partial q_c}gj - \frac{\partial f}{\partial q_d}gk\right). \tag{32}$$

Thus the product rule in general is different from the usual one.

3) The second product rule: However, the usual product rule applies to differentiation with respect to real variables, i.e.,

$$\frac{\partial fg}{\partial q_\phi} = \frac{\partial f}{\partial q_\phi}g + f\frac{\partial g}{\partial q_\phi} \tag{33}$$

for $\phi = a, b, c$, or $d$.

4) The third product rule: The usual product rule also applies if at least one of the two functions $f(q)$ and $g(q)$ is real-valued, i.e.,

$$\frac{\partial fq}{\partial q} = f\frac{\partial g}{\partial q} + \frac{\partial f}{\partial q}g. \tag{34}$$

5) The first chain rule: For a composite function $f(g(q))$, $g(q) := g_a + g_b i + g_c j + g_d k$ being a quaternion-valued function, we have the following chain rule [15]:

$$\nabla_q f = (\nabla_q^g f)M \tag{35}$$

where $\nabla_q^g := (\partial/\partial g, \partial/\partial g^i, \partial/\partial g^j, \partial/\partial g^k)$ and $M$ is a $4 \times 4$ matrix with element $M_{\mu\nu} = \partial g^\mu/\partial q^\nu$ for $\mu, \nu \in \{1, i, j, k\}$ and $g^\mu = -\mu g \mu$ ($g^1$ is understood as the same as $g$). Explicitly, we may write

$$\frac{\partial f}{\partial q^\nu} = \sum_\mu \frac{\partial f}{\partial g^\mu}\frac{\partial g^\mu}{\partial q^\nu}. \tag{36}$$

The proof is outlined in Appendix C.

6) The second chain rule: The above chain rule uses $g$ and its involutions as the intermediate variables. It is sometimes convenient to use the real components of $g$ for that purpose instead. In this case, the following chain rule may be used:

$$\nabla_q f = (\nabla_r^g f)O \tag{37}$$

where $O$ is a $4 \times 4$ matrix with entry $O_{\phi\nu} = \partial g_\phi/\partial q^\nu$ with $\phi \in \{a, b, c, d\}$ and $\nu \in \{1, i, j, k\}$, and $\nabla_r^g := (\partial/\partial g_a, \partial/\partial g_b, \partial/\partial g_c, \partial/\partial g_d)$. Explicitly, we have

$$\frac{\partial f}{\partial q^\nu} = \sum_\phi \frac{\partial f}{\partial g_\phi}\frac{\partial g_\phi}{\partial q^\nu}. \tag{38}$$

7) The third chain rule: if the intermediate function $g(q)$ is real-valued, i.e., $g = g_a$, then from the second chain rule, we obtain

$$\frac{\partial f}{\partial q^\nu} = \frac{\partial f}{\partial g}\frac{\partial g}{\partial q^\nu}. \tag{39}$$

8) $f(q)$ is not independent of $q^i$, $q^j$ or $q^k$ in the sense that, in general,

$$\frac{\partial f(q)}{\partial q^i} \neq 0, \frac{\partial f(q)}{\partial q^j} \neq 0, \frac{\partial f(q)}{\partial q^k} \neq 0. \tag{40}$$

This can be illustrated by $f(q) = q^2$. Using the first product rule (equation (31)), we have

$$\frac{\partial q^2}{\partial q^i} = q\frac{\partial q}{\partial q^i} + \frac{1}{4}\sum_{(\phi,\nu)}\frac{\partial q}{\partial q_\phi}q\nu$$

for $(\phi, \nu) \in \{(a, 1), (b, i), (c, -j), (d, -k)\}$. It can then be shown that

$$\frac{\partial q^2}{\partial q^i} = q_b i, \ \frac{\partial q^2}{\partial q^j} = q_c j, \ \frac{\partial q^2}{\partial q^k} = q_d k. \quad (41)$$

This property demonstrates the intriguing difference between the HR derivative and the usual derivatives, although we can indeed show that

$$\frac{\partial q}{\partial q^\nu} = 0. \quad (42)$$

One implication of this observation is that, for a nonlinear algorithm involving simultaneously more than one gradients $\partial f/\partial q^\nu$, we have to take care to include all the terms.

## IV. Restricted HR derivatives for a class of regular functions

Using the above operation rules, we may find explicit expressions for the derivatives for a whole range of functions. We first introduce the following lemma:

**Lemma 1.** *The derivative of the power function $f(q) = (q - q_0)^n$, with integer $n$ and constant quaternion $q_0$, is*

$$\frac{\partial f(q)}{\partial q} = \frac{1}{2}\left(n\tilde{q}^{n-1} + \frac{\tilde{q}^n - \tilde{q}^{*n}}{\tilde{q} - \tilde{q}^*}\right), \quad (43)$$

*with $\tilde{q} = q - q_0$.*

*Remark.* The division in $(\tilde{q}^n - \tilde{q}^{*n})/(\tilde{q} - \tilde{q}^*)$ is understood as $(\tilde{q}^n - \tilde{q}^{*n})(\tilde{q} - \tilde{q}^*)^{-1}$ or $(\tilde{q} - \tilde{q}^*)^{-1}(\tilde{q}^n - \tilde{q}^{*n})$ which are the same since the two factors commute. The division operations in what follows are understood in the same way.

*Proof:* The lemma is obviously true for $n = 0$. Let $n \geq 1$, we apply the first product rule, and find

$$\frac{\partial (q - q_0)^n}{\partial q} = \tilde{q}\frac{\partial \tilde{q}^{n-1}}{\partial q} + R(\tilde{q}^{n-1}) \quad (44)$$

where $R(\tilde{q}^{n-1})$ is the real part of $\tilde{q}^{n-1}$. We then obtain by induction

$$\frac{\partial (q - q_0)^n}{\partial q} = \sum_{m=0}^{n-1} \tilde{q}^m R(\tilde{q}^{n-1-m}). \quad (45)$$

Using $R(\tilde{q}^{n-1-m}) = \frac{1}{2}(\tilde{q}^{n-1-m} + \tilde{q}^{*(n-1-m)})$, the summations can be evaluated explicitly, leading to equation (43).

For $n < 0$, we use the recurrent relation

$$\frac{\partial ((q - q_0)^{-n})}{\partial q} = \tilde{q}^{-1}\left[\frac{\partial \tilde{q}^{-(n-1)}}{\partial q} - R(\tilde{q}^{-n})\right] \quad (46)$$

and the result

$$\frac{\partial (q - q_0)^{-1}}{\partial q} = -\tilde{q}^{-1} R(\tilde{q}^{-1}). \quad (47)$$

Equation (43) is proven by using induction as for $n > 0$. More details are given in Appendix B. ■

**Theorem 1.** *Assuming $f : H \to H$ admits a power series representation $f(q) := g(\tilde{q}) := \sum_{n=-\infty}^{\infty} a_n \tilde{q}^n$, with $a_n$ being a quaternion constant and $\tilde{q} = q - q_0$, for $R_1 \leq |\tilde{q}| \leq R_2$ with $R_1, R_2 > 0$ being some constants, then*

$$\frac{\partial f(q)}{\partial q} = \frac{1}{2}\left[f'(q) + (g(\tilde{q}) - g(\tilde{q}^*))(\tilde{q} - \tilde{q}^*)^{-1}\right], \quad (48)$$

*where $f'(q)$ is the derivative in the usual sense, i.e.,*

$$f'(q) := \sum_{n=-\infty}^{\infty} n a_n \tilde{q}^{n-1} = \sum_{n=-\infty}^{\infty} n a_n (q - q_0)^{n-1}. \quad (49)$$

*Proof:* Using Lemma 1 and the left-linearity of HR gradients, we have

$$\begin{aligned} \frac{\partial f}{\partial q} &= \frac{1}{2}\sum_{n=-\infty}^{\infty} a_n[n\tilde{q}^{n-1} + (\tilde{q}^n - \tilde{q}^{*n})(\tilde{q} - \tilde{q}^*)^{-1}] \\ &= f'(q) + \frac{1}{2}\left[\sum_{n=\infty}^{\infty} a_n(\tilde{q}^n - \tilde{q}^{*n})\right](\tilde{q} - \tilde{q}^*)^{-1} \\ &= \frac{1}{2}[f'(q) + (g(\tilde{q}) - g(\tilde{q}^*))(\tilde{q} - \tilde{q}^*)^{-1}], \end{aligned}$$

proving the theorem. ■

The functions $f(q)$ form a class of regular functions on $H$. A full discussion of such functions is beyond the scope of this paper. However, we note that a similar class of functions have been discussed in [17]. A parallel development for the former is possible, and will be the topic of a future paper. Meanwhile, we observe that many useful elementary functions satisfy the conditions in Theorem 1. To illustrate the application of the theorem, we list below the derivatives of a number of such functions.

**Example 1.** Exponential function $f(q) = e^q$ has representation

$$e^q := \sum_{n=0}^{\infty} \frac{q^n}{n!}. \quad (50)$$

Applying Theorem 1 with $a_n = 1/n!$ and $q_0 = 0$, we have

$$\frac{\partial e^q}{\partial q} = \frac{1}{2}\left(e^q + \frac{e^q - e^{q^*}}{q - q^*}\right). \quad (51)$$

Making use of $e^q = e^{q_a}(\cos v + \hat{\mathbf{v}} \sin v)$ and $q = q_a + \hat{\mathbf{v}} v$, we have

$$\frac{\partial e^q}{\partial q} = \frac{1}{2}\left(e^q + e^{q_a} v^{-1} \sin v\right). \quad (52)$$

**Example 2.** The logarithmic function $f(q) = \ln q$ has representation

$$\ln q = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}(q - 1)^n. \quad (53)$$

with $a_n = (-1)^{n-1}/n$ and $q_0 = 1$. Since $q_0$ is a real number, $g(\tilde{q}^*) = f(q^*)$. Therefore, we have from Theorem 1

$$\frac{\partial \ln q}{\partial q} = \frac{1}{2}\left(q^{-1} + \frac{\ln q - \ln q^*}{q - q^*}\right). \quad (54)$$

Using representation $\ln q = \ln|q| + \hat{\mathbf{v}}\arccos(q_a/|q|)$, the expression can be simplified as

$$\frac{\partial \ln q}{\partial q} = \frac{1}{2}\left(q^{-1} + \frac{1}{v}\arccos\frac{q_a}{|q|}\right), \tag{55}$$

where $v = |I(q)|$.

**Example 3.** Hyperbolic tangent function $f(q) = \tanh q$ is defined as

$$\tanh q := \frac{e^q - e^{-q}}{e^q + e^{-q}} = q - \frac{q^3}{3} + \frac{2q^5}{15} - ... \tag{56}$$

Therefore, Theorem 1 applies. On the other hand, using the relation $e^q = e^{q_a}(\cos v + \hat{\mathbf{v}}\sin v)$, we can show that

$$\tanh q = \frac{1}{2}\frac{\sinh 2q_a + \hat{\mathbf{v}}\sin 2v}{\sinh^2 q_a + \cos^2 v}. \tag{57}$$

Then the second term in the expression given by Theorem 1 can be simplified. The final expression can be written as

$$\frac{\partial \tanh q}{\partial q} = \frac{1}{2}\left(\operatorname{sech}^2 q + \frac{v^{-1}\sin 2v}{\cosh 2q_a + \cos 2v}\right), \tag{58}$$

where $\operatorname{sech} q := 1/\cosh q$ is the quaternionic hyperbolic secant function.

*Remark.* Apparently, the derivatives for these functions can also be found by direct calculations without resorting to Theorem 1.

We now turn to a question of more theoretical interests. Even though it might not be obvious from the definitions, the following theorem shows that the restricted HR derivative is consistent with the derivative in the real domain for a class of functions, including those in the above examples.

**Theorem 2.** *For the function $f(q)$ in Theorem 1, if $q_0$ is a real number, then*

$$\frac{\partial f(q)}{\partial q} \to f'(q) \tag{59}$$

*when $q \to R(q)$, i.e., when $q$ approaches a real number.*

*Proof:* Using the polar representation, we write $\tilde{q} = |\tilde{q}|\exp(\hat{\mathbf{v}}\theta)$, where $\theta = \arcsin(v/|\tilde{q}|)$ is the argument of $\tilde{q}$ with $v = |I(\tilde{q})|$. Then $\tilde{q}^n = |\tilde{q}|^n\exp(n\hat{\mathbf{v}}\theta)$, and

$$(\tilde{q}^n - \tilde{q}^{*n})(\tilde{q} - \tilde{q}^*)^{-1} = \frac{I(\tilde{q}^n)}{I(\tilde{q})} = \frac{|\tilde{q}|^{n-1}\sin(n\theta)}{\sin\theta}. \tag{60}$$

For real $q_0$, $\tilde{q} \to q_a - q_0$ and $v \to 0$ when $q \to R(q)$. As a consequence, $\theta \to 0$ at the limit (or $\theta \to \pi$, which can be dealt with by slight modification), and

$$\frac{\sin(n\theta)}{\sin\theta} \sim \frac{\sin(n\theta)}{\theta} \to n, \quad |\tilde{q}|^{n-1} \to (q_a - q_0)^{n-1}. \tag{61}$$

Therefore,

$$(\tilde{q}^n - \tilde{q}^{*n})(\tilde{q} - \tilde{q}^*)^{-1} \to n\tilde{q}^{n-1} \tag{62}$$

and

$$[g(\tilde{q}) - g(\tilde{q}^*)](\tilde{q} - \tilde{q}^*)^{-1} \to \sum_{n=-\infty}^{\infty} n a_n \tilde{q}^{n-1} = f'(q). \tag{63}$$

Thus

$$\frac{\partial f(q)}{\partial q} \to \frac{1}{2}[f'(q) + f'(q)] = f'(q). \tag{64}$$

■

The functions in above three examples all satisfy the conditions in Theorem 2, hence we expect Theorem 2 applies. One can easily verify by direct calculations that the theorem indeed holds.

## V. The right restricted HR gradients

In this section, we briefly summarize the results for the right restricted HR gradients, and highlight the difference with left restricted HR gradients.

1) Right-linearity: for arbitrary quaternion constants $\alpha$ and $\beta$, and functions $f(q)$ and $g(q)$, we have

$$\frac{\partial^R(f\alpha + g\beta)}{\partial q^\nu} = \frac{\partial^R f}{\partial q^\nu}\alpha + \frac{\partial^R g}{\partial q^\nu}\beta. \tag{65}$$

However, linearity does not hold for left multiplications, i.e., in general

$$\frac{\partial^R \alpha f}{\partial q} \neq \alpha\frac{\partial^R f}{\partial q}. \tag{66}$$

2) The first product rule: for the right restricted HR operator, the following product rule holds:

$$[\nabla_q^R(fg)]^T = [(\nabla_q^R f)g]^T + J^*[f(\nabla_r g)^T]. \tag{67}$$

The second and third product rules are the same as for the left restricted operator.

3) The first chain rule: for the composite function $f(g(q))$, we have

$$(\nabla_q^R f)^T = M^T(\nabla_q^{gR} f)^T. \tag{68}$$

4) The second chain rule becomes:

$$(\nabla_q^R f)^T = O^T(\nabla_r^g f)^T. \tag{69}$$

5) The third chain rule becomes

$$\frac{\partial^R f}{\partial q^\nu} = \frac{\partial g}{\partial q^\nu}\frac{\partial f}{\partial g}. \tag{70}$$

Note that, $\partial g/\partial q^\nu = \partial^R g/\partial q^\nu$ since $g$ is real-valued. We thus have omitted the superscript $R$. Also, $\partial f/\partial g$ is a real derivative, so there is no distinction between left and right derivatives.

We can also find the right restricted HR gradients for common quaternion functions. First of all, Lemma 1 is also true for right derivatives:

**Lemma 2.** *For $f(q) = (q - q_0)^n$ with $n$ integer and $q_0$ a constant quaternion, we have*

$$\frac{\partial^R f(q)}{\partial q} = \frac{1}{2}\left(n\tilde{q}^{n-1} + \frac{\tilde{q}^n - \tilde{q}^{*n}}{\tilde{q} - \tilde{q}^*}\right), \tag{71}$$

*with $\tilde{q} = q - q_0$.*

*Remark.* To prove the lemma, we use the following recurrent relations:

$$\frac{\partial (q-q_0)^n}{\partial q} = \frac{\partial \tilde{q}^{n-1}}{\partial q}\tilde{q} + R(\tilde{q}^{n-1}) \quad (72)$$

$$\frac{\partial ((q-q_0)^{-n})}{\partial q} = \left[\frac{\partial \tilde{q}^{-(n-1)}}{\partial q} - R(\tilde{q}^{-n})\right]\tilde{q}^{-1}. \quad (73)$$

Using Lemma 2, We can prove the following result:

**Theorem 3.** *Assuming $f : H \to H$ admits a power series representation $f(q) := g(\tilde{q}) := \sum_{n=-\infty}^{\infty} \tilde{q}^n a_n$, with $a_n$ being a quaternion constant and $\tilde{q} = q - q_0$, for $R_1 \le |\tilde{q}| \le R_2$ with $R_1, R_2 > 0$ being some constants, then*

$$\frac{\partial^R f(q)}{\partial q} = \frac{1}{2}\left[f'(q) + (\tilde{q}-\tilde{q}^*)^{-1}(g(\tilde{q}) - g(\tilde{q}^*))\right], \quad (74)$$

*where $f'(q)$ is the derivative in the usual sense, i.e.,*

$$f'(q) := \sum_{n=-\infty}^{\infty} n\tilde{q}^{n-1}a_n = \sum_{n=-\infty}^{\infty} n(q-q_0)^{n-1}a_n. \quad (75)$$

Note that, the functions $f(q)$ in Theorem 3 in general form a different class of functions than the one in Theorem 1, because in the series representation $a_n$ appears on the right-hand side of the powers. However, if $a_n$ is a real number, then the two classes of functions coincide. Therefore, we have the following result:

**Theorem 4.** *If $a_n$ is real, then the left and right restricted HR gradients of $f(q)$ coincide.*

*Remark.* As a consequence, we can see immediately the right derivatives for the exponential, logarithmic and hyperbolic tangent functions are the same as the left ones.

Apparently, Theorem 2 is also true for the right derivatives. Hence, we have:

**Theorem 5.** *The right-restricted HR gradient is consistent with the real gradient in the sense of Theorem 2.*

## VI. The increment of a quaternion function

When $f(q)$ is a real-valued quaternion function, both left and right restricted HR gradients are coincident with the HR gradients. Besides, we have

$$\frac{\partial^R f}{\partial q^\nu} = \frac{\partial f}{\partial q^\nu} = \left(\frac{\partial f}{\partial q}\right)^\nu, \quad (76)$$

where $\nu \in i, j, k$. Thus only $\partial f/\partial q$ is independent. As a consequence (see also [14]),

$$\begin{aligned} df &= \sum_\nu \frac{\partial f}{\partial q^\nu} dq^\nu = \sum_\nu \left(\frac{\partial f}{\partial q}\right)^\nu dq^\nu \\ &= \sum_\nu \left(\frac{\partial f}{\partial q} dq\right)^\nu = 4R\left(\frac{\partial f}{\partial q} dq\right), \end{aligned} \quad (77)$$

where equation (76) has been used. Hence, $-(\partial f/\partial q)^*$ gives the steepest descent direction for $f$, and the increment is determined by $\partial f/\partial q$.

On the other hand, if $f$ is a quaternion-valued function, the increment will depend on all four derivatives. Taking $f(q) = q^2$ as an example, we have (see equations (41) and (43))

$$dq^2 = (q+q_a)dq + q_b i dq^i + q_c j dq^j + q_d k dq^k, \quad (78)$$

even though $f(q)$ appears to be independent of $q^i$, $q^j$ and $q^k$. It can be verified that the above expression is the same as the differential form given in terms of $dq_a$, $dq_b$, $dq_c$ and $dq_d$. Thus it is essential to include the contributions from $\partial f/\partial q^i$ etc.

We also note that, if the right gradient is used consistently, the same increment would result, since the basis of the definitions is the same, namely, the differential form in term of $dq_a$, $dq_b$, $dq_c$ and $dq_d$.

Now we apply the quaternion-valued restricted HR gradient operator to develop the QLMS algorithm as an application. This version of QLMS has been derived in [9], [12], [13], [15]. However, with the rules we have derived, some of the calculations can be simplified, as we will be showing below.

In terms of a standard adaptive filter, the output $y[n]$ and error $e[n]$ can be expressed as

$$y[n] = \mathbf{w}^T[n]\mathbf{x}[n] \quad (79)$$

$$e[n] = d[n] - \mathbf{w}^T[n]\mathbf{x}[n], \quad (80)$$

where $\mathbf{w}[n] = [w[1], w[2], \cdots, w[M]]^T$ is the quaternion adaptive weight coefficient vector with length $M$, $d[n]$ the reference signal, and $\mathbf{x}[n] = [x[n-1], x[n-2], \cdots, x[n-M]]^T$ the quaternion input sample sequence. The conjugate $\mathbf{e}^*[n]$ of the error signal $e[n]$ is

$$e^*[n] = d^*[n] - \mathbf{x}^H[n]\mathbf{w}^*[n]. \quad (81)$$

The cost function is defined as $J[n] = e[n]e^*[n]$ which is real-valued. According to the discussion above and [14], [18], the conjugate gradient $(\nabla_{\mathbf{w}} J[n])^*$ gives the maximum steepness direction for the optimization surface. Therefore it is used to update the weight vector. Specifically,

$$\mathbf{w}[n+1] = \mathbf{w}[n] - \mu(\nabla_{\mathbf{w}} J[n])^*, \quad (82)$$

where $\mu$ is the step size. To find $\nabla_{\mathbf{w}} J$ , we use the first product rule:

$$\begin{aligned} \nabla_{\mathbf{w}} &= \frac{\partial e[n]e^*[n]}{\partial \mathbf{w}} \\ &= e[n]\frac{\partial e^*[n]}{\partial \mathbf{w}} + \frac{1}{4}\Big(\frac{\partial e[n]}{\partial \mathbf{w}_a}e^*[n] - \frac{\partial e[n]}{\partial \mathbf{w}_b}e^*[n]i \\ &\quad - \frac{\partial e[n]}{\partial \mathbf{w}_c}e^*[n]j - \frac{\partial e[n]}{\partial \mathbf{w}_d}e^*[n]k\Big) \end{aligned} \quad (83)$$

After some algebra, we find

$$\nabla_{\mathbf{w}} J[n] = -\frac{1}{2}\mathbf{x}[n]e^*[n]. \quad (84)$$

Therefore we obtain the following update equation for the QLMS algorithm with a step size $\mu$

$$\mathbf{w}[n+1] = \mathbf{w}[n] + \mu(e[n]\mathbf{x}^*[n]). \tag{85}$$

Some simulation results have been reported in [12].

## VII. Conclusions

We have proposed a restricted HR gradient operator and discussed its properties, in particular several different versions of product rules and chain rules. Using the operator, we apply the rules to find the derivatives for a wide class of nonlinear quaternion-valued functions that admit a power series representation. The class includes the common elementary functions such as the exponential function, the logarithmic function, among others. The explicit expressions for the derivatives will be useful for nonlinear signal processing applications. We also prove for a wide class of functions, that the restricted HR gradient tends to the derivatives for real functions with respect to real variables, when the independent quaternion variable tends to the real axis, thus showing the consistency of the definition.

## Appendix A
## Definition of the Operators

We consider $df = df_a + idf_b + jdf_c + kdf_d$. By definition, we have $df_\gamma = \sum_\phi (\partial f_\gamma/\partial q_\phi) dq_\phi$, with $\gamma, \phi \in \{a, b, c, d\}$. Using the relations

$$dq_a = \frac{1}{4}(dq + dq^i + dq^j + dq^k), \tag{86}$$

$$dq_b = \frac{1}{4i}(dq + dq^i - dq^j - dq^k), \tag{87}$$

$$dq_c = \frac{1}{4j}(dq - dq^i + dq^j - dq^k), \tag{88}$$

$$dq_d = \frac{1}{4k}(dq - dq^i - dq^j + dq^k), \tag{89}$$

we may rewrite $df_\gamma$ as follows

$$\begin{aligned} df_\gamma = \quad & \frac{1}{4}(\frac{\partial f_\gamma}{\partial q_a} - i\frac{\partial f_\gamma}{\partial q_b} - j\frac{\partial f_\gamma}{\partial q_c} - k\frac{\partial f_\gamma}{\partial q_d})dq \\ & + \frac{1}{4}(\frac{\partial f_\gamma}{\partial q_a} - i\frac{\partial f_\gamma}{\partial q_b} + j\frac{\partial f_\gamma}{\partial q_c} + k\frac{\partial f_\gamma}{\partial q_d})dq^i \\ & + \frac{1}{4}(\frac{\partial f_\gamma}{\partial q_a} + i\frac{\partial f_\gamma}{\partial q_b} - j\frac{\partial f_\gamma}{\partial q_c} + k\frac{\partial f_\gamma}{\partial q_d})dq^j \\ & + \frac{1}{4}(\frac{\partial f_\gamma}{\partial q_a} + i\frac{\partial f_\gamma}{\partial q_b} + j\frac{\partial f_\gamma}{\partial q_c} - k\frac{\partial f_\gamma}{\partial q_d})dq^k \end{aligned}$$

which can be written as

$$df_\gamma = \frac{1}{4}\sum_\nu \left( \sum_{(\phi,\mu)} \frac{\partial f_\gamma}{\partial q_\phi}\mu^\nu \right) dq^\nu \tag{90}$$

where $(\phi, \mu) \in \{(a, 1), (b, -i), (c, -j), (d, -k)\}$, $\nu \in \{1, i, j, k\}$, and $\mu^\nu$ is the $\nu$-involution of $\mu$. Therefore

$$\begin{aligned} df &= df_a + idf_b + jdf_c + kdf_d \\ &= \frac{1}{4}\sum_\nu \left( \sum_{(\phi,\mu)} \frac{\partial(f_a + if_b + jf_c + kf_d)}{\partial q_\phi}\mu^\nu \right) dq^\nu \\ &= \frac{1}{4}\sum_\nu \left( \sum_{(\phi,\mu)} \frac{\partial f}{\partial q_\phi}\mu^\nu \right) dq^\nu \end{aligned} \tag{91}$$

which leads to the definitions (10-18) in the main text. Note that, because $\mu^\nu$ and $dq^\nu$ are quaternions, to obtain the last equation, we need to multiply $df_b$, $df_c$ and $df_d$ by $i$, $j$, and $k$ from the left.

On the other hand, we notice that the prefactors in (87-89) may be moved to the right-hand side of the other factors, i.e., we may write

$$dq_a = (dq + dq^i + dq^j + dq^k)\frac{1}{4}, \tag{92}$$

$$dq_b = (dq + dq^i - dq^j - dq^k)\frac{1}{4i}, \tag{93}$$

$$dq_c = (dq - dq^i + dq^j - dq^k)\frac{1}{4j}, \tag{94}$$

$$dq_d = (dq - dq^i - dq^j + dq^k)\frac{1}{4k}. \tag{95}$$

Using these relations, we may find another expression for $df_\gamma$ following the procedure above:

$$df_\gamma = \frac{1}{4}\sum_\nu dq^\nu \left( \sum_{(\phi,\mu)} \mu^\nu \frac{\partial f_\gamma}{\partial q_\phi} \right). \tag{96}$$

The expression is different from (90), in that the differentials $dq^\nu$ are on the left of $\mu^\nu$. Therefore, we derive

$$\begin{aligned} df &= df_a + df_b i + df_c j + df_d k \\ &= \frac{1}{4}\sum_\nu dq^\nu \left( \sum_{(\phi,\mu)} \mu^\nu \frac{\partial(f_a + f_b i + f_c j + f_d k)}{\partial q_\phi} \right) \\ &= \frac{1}{4}\sum_\nu dq^\nu \left( \sum_{(\phi,\mu)} \mu^\nu \frac{\partial f}{\partial q_\phi} \right), \end{aligned} \tag{97}$$

which is the basis for the definitions for the right restricted HR derivatives as given in the main text.

## Appendix B
## Additional Details for the Proof of Lemma 1

To prove Lemma 1, we have used the following relation

$$\frac{\partial q^{-1}}{\partial q} = -q^{-1}R(q^{-1}). \tag{98}$$

To show this result, we note $\partial(qq^{-1})/\partial q = \partial 1/\partial q = 0$. Thus

$$\begin{aligned} 0 &= q\frac{\partial q^{-1}}{\partial q} + \frac{1}{4}(q^{-1} - iq^{-1}i - jq^{-1}j - kq^{-1}k) \\ &= q\frac{\partial q^{-1}}{\partial q} + R(q^{-1}), \end{aligned} \tag{99}$$

from which the result follows. We have used equation (10) and the fact that

$$\frac{\partial q}{\partial q_a} = 1, \frac{\partial q}{\partial q_b} = i, \frac{\partial q}{\partial q_c} = j, \frac{\partial q}{\partial q_d} = k. \tag{100}$$

The proof also uses the following recurrent relation

$$\frac{\partial q^{-n}}{\partial q} = q^{-1}\left[\frac{\partial q^{-(n-1)}}{\partial q} - R(q^{-n})\right], \tag{101}$$

which can be shown as follows: using the first product rule, we have

$$\begin{aligned}\frac{\partial q^{-n}}{\partial q} &= q^{-1}\frac{\partial q^{-(n-1)}}{\partial q} + \frac{1}{4}\left(\frac{\partial q^{-1}}{\partial q_a}q^{-(n-1)} - \frac{\partial q^{-1}}{\partial q_b}q^{-(n-1)}i\right.\\ &\left.-\frac{\partial q^{-1}}{\partial q_c}q^{-(n-1)}j - \frac{\partial q^{-1}}{\partial q_d}q^{-(n-1)}k\right).\end{aligned} \tag{102}$$

Using the fact $\partial qq^{-1}/\partial q_\phi = 0$ and the second product rule, we can find

$$\frac{\partial q^{-1}}{\partial q_\phi} = -q^{-1}\frac{\partial q}{\partial q_\phi}q^{-1}. \tag{103}$$

Thus

$$\begin{aligned}\frac{\partial q^{-n}}{\partial q} &= q^{-1}\frac{\partial q^{-(n-1)}}{\partial q} - \frac{q^{-1}}{4}\left(q^{-n} - iq^{-n}i\right.\\ &\quad\left.-jq^{-n}j - kq^{-n}k\right)\\ &= q^{-1}\frac{\partial q^{-(n-1)}}{\partial q} - q^{-1}R(q^{-n}).\end{aligned} \tag{104}$$

## APPENDIX C
## DERIVATIONS OF THE FIRST CHAIN RULE

The function $f(g(q))$ may be view as a function of intermediate variables $g_a$, $g_b$, $g_c$ and $g_d$. Using the usual chain rule, we have

$$\frac{\partial f}{\partial q_\beta} = \sum_\phi \frac{\partial f}{\partial g_\phi}\frac{\partial g_\phi}{\partial q_\beta}, \tag{105}$$

with $\beta \in \{a, b, c, d\}$, which gives

$$\nabla_r f = (\nabla_r^g f)P \tag{106}$$

where $P$ is a $4 \times 4$ matrix with $P_{\phi\beta} = \partial g_\phi/\partial q_\beta$. With $(\nabla_r f)J^H = \nabla_q f$, and $\nabla_r^g f = 4(\nabla_q^g f)J$, the above equation leads to

$$\nabla_q f = 4(\nabla_q^g f)JPJ^H, \tag{107}$$

where it is easy to show that $4JPJ^H = M$.

## REFERENCES


[1] S.C. Pei and C.M. Cheng, "Color image processing by using binary quaternion-moment-preserving thresholding technique," *IEEE Transactions on Image Processing*, vol. 8, no. 5, pp. 614–628, 1999.

[2] SJ Sangwine, "The discrete fourier transform of a colour image," *Image Processing II Mathematical Methods, Algorithms and Applications*, pp. 430–441, 2000.

[3] M. Parfieniuk and A. Petrovsky, "Inherently lossless structures for eight- and six-channel linear-phase paraunitary filter banks based on quaternion multipliers," *Signal Processing*, vol. 90, no. 6, pp. 1755–1767, 2010.

[4] N. Le Bihan and J. Mars, "Singular value decomposition of quaternion matrices: a new tool for vector-sensor signal processing," *Signal Processing*, vol. 84, no. 7, pp. 1177–1199, 2004.

[5] S. Miron, N. Le Bihan, and J. I. Mars, "Quaternion-MUSIC for vector-sensor array processing," *IEEE Transactions on Signal Processing*, vol. 54, no. 4, pp. 1218–1229, April 2006.

[6] N. Le Bihan, S. Miron, and J. I. Mars, "MUSIC algorithm for vector-sensors array using biquaternions," *IEEE Transactions on Signal Processing*, vol. 55, no. 9, pp. 4523–4533, 2007.

[7] J. W. Tao and W. X. Chang, "A novel combined beamformer based on hypercomplex processes," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 49, no. 2, pp. 1276–1289, 2013.

[8] J.-W. Tao, "Performance analysis for interference and noise canceller based on hypercomplex and spatio-temporal-polarisation processes," *IET Radar, Sonar Navigation*, vol. 7, no. 3, pp. 277–286, 2013.

[9] J. W. Tao and W. X. Chang, "Adaptive beamforming based on complex quaternion processes," *Mathematical Problems in Engineering*, vol. 2014, 2014.

[10] X. R. Zhang, W. Liu, Y. G. Xu, and Z. W. Liu, "Quaternion-valued robust adaptive beamformer for electromagnetic vector-sensor arrays with worst-case constraint," *Signal Processing*, vol. 104, pp. 274–283, November 2014.

[11] Clive Cheong Took and Danilo P Mandic, "The quaternion LMS algorithm for adaptive filtering of hypercomplex processes," *IEEE Transactions on Signal Processing*, vol. 57, no. 4, pp. 1316–1327, 2009.

[12] M. D. Jiang, W. Liu, and Y. Li, "A general quaternion-valued gradient operator and its applications to computational fluid dynamics and adaptive beamforming," in *Proc. of the International Conference on Digital Signal Processing*, Hong Kong, August 2014 (http://arxiv.org/abs/1406.5719).

[13] M. D. Jiang, W. Liu, and Y. Li, "A zero-attracting quaternion-valued least mean square algorithm for sparse system identification," in *Proc. IEEE/IET International Sympoisum on Communication Systems, Networks and Digital Signal Processing*, Manchester, UK, July 2014 (http://arxiv.org/abs/1406.5721).

[14] D. P. Mandic, C. Jahanchahi, and C. C. Took, "A quaternion gradient operator and its applications," *IEEE Signal Processing Letters*, vol. 18, no. 1, pp. 47–50, 2011.

[15] D. P. Xu and D. P Mandic, "Quaternion gradient and hessian," *arXiv preprint arXiv:1406.3587*, 2014.

[16] T. A. Ell and S. J. Sangwine, "Quaternion involutions and anti-involutions," *Computers & Mathematics with Applications*, vol. 53, no. 1, pp. 137–143, 2007.

[17] G. Gentili, and D. C. Struppa, "A new theory of regular functions of a quaternionic variable," *Advances in Mathematics* , vol. 216, pp. 279–301, 2007.

[18] D. H. Brandwood, "A complex gradient operator and its application in adaptive array theory," *IEE Proceedings H (Microwaves, Optics and Antennas)*, vol. 130, no. 1, pp. 11–16, 1983.