About QLMS Derivations

Quentin Barthélemy, Anthony Larue, Jerome I. Mars

About QLMS derivations

Quentin Barthélemy, Anthony Larue and Jérôme I. Mars

***Abstract*—In this letter, a review of the quaternionic least mean squares (QLMS) algorithm is proposed. Three versions coming from three derivation ways exist: the original QLMS [1] based on componentwise gradients, $\mathbb{HR}$-QLMS [2] based on a quaternion gradient operator and iQLMS [3] based on an involutions-gradient. Noting and investigating the differences between the three QLMS formulations, we show that the original QLMS suffers from a mistake in the derivation calculus. Thus, we propose to derive rigorously the criterion following the first way, giving the correct version of QLMS. A comparison with the other QLMS versions validates these results on simulated data.**

***Index Terms*—Quaternionic signal processing; QLMS; adaptive filtering.**

I. Introduction

In signal processing, the least mean squares (LMS) algorithm [4] is well-used for several purposes and in particular for adaptive filtering. Filter weights are estimated to fit a least-squares criterion and are updated thanks to a stochastic gradient descent. This algorithm has been extended to complex in a first way in [5] by Widrow *et al.* who summed the componentwise gradients to derive the complex LMS (CLMS). Later, a gradient operator was introduced by Brandwood in [6]. Assuming $z \in \mathbb{C}$, the complex derivation rules are:

$$\frac{\partial z}{\partial z} = \frac{\partial z^*}{\partial z^*} = 1 \quad \text{and} \quad \frac{\partial z^*}{\partial z} = \frac{\partial z}{\partial z^*} = 0\,. \tag{1}$$

Additionally, he showed that the direction of maximum rate of change of a real-valued objective function $J = \|\epsilon\|^2$ with respect to z is $\partial J / \partial z^*$. Using these results, Brandwood retrieved exactly the CLMS by this second way, give or take a multiplicative constant.

Recently, the LMS has been extended to the quaternions by three different ways: the quaternionic LMS (QLMS) [1], the $\mathbb{HR}$-QLMS [2] and the iQLMS [3]. These algorithms are well-used in many recent works. The problem is that these three ways give three versions which are different.

In this letter, we examine rigorously the first way to derive the QLMS, investigating the work of Took and Mandic [1]. In the first section, quaternions are presented. Then, the three versions of QLMS are reviewed in Section III. The first derivation way is detailed in Section IV, giving a new version of the QLMS. A comparison on simulated data is made in Section VI to validate theoretical results.

II. Quaternion algebra

The quaternions algebra, denoted as $\mathbb{H}$, is an extension of the complex space $\mathbb{C}$ using three imaginary parts [7]. A quaternion $q \in \mathbb{H}$ is defined as:

$$q = q_a + q_b i + q_c j + q_d k\,, \tag{2}$$

with q_a, q_b, q_c, $q_d \in \mathbb{R}$ and with imaginary units defined as:

$$ij = k, jk = i, ki = j \quad \text{and} \quad i^2 = j^2 = k^2 = ijk = -1. \tag{3}$$

The quaternionic space is characterized by its noncommutativity: $q_1 q_2 \neq q_2 q_1$. The scalar part is $\Re(q) = q_a$, and the vectorial part is $\Im(q) = q_b i + q_c j + q_d k$. The conjugate q^* is defined as: $q^* = \Re(q) - \Im(q)$ and we have $(q_1 q_2)^* = q_2^* q_1^*$. The modulus is defined as $|q| = \sqrt{q q^*}$.

Concerning quaternionic vectors, $(.)^T$ denotes the transpose operator and $(.)^H$ the conjugate transpose operator.

III. QLMS derivations

In this section, the problem is presented and the three ways to derive QLMS are reviewed. For consistency, notations of the concerned articles [1]–[3] are kept.

A. Problem formulation

A linear model linking an input signal $x \in \mathbb{H}^N$ to an output signal $d \in \mathbb{H}^N$ is considered. We defined $d(n) \in \mathbb{H}$ as the instantaneous output data signal, $w(n) \in \mathbb{H}^L$ as the adaptive weights vector of length L, and $x(n) \in \mathbb{H}^L$ as the last L samples of the input data signal. The linear model, defined in [1], is written as:

$$d(n) = w(n)^T\, x(n) + e(n)\,, \tag{4}$$

with $e(n) \in \mathbb{H}$ the instantaneous error.

The goal of the QLMS is to estimate the optimal weights vector w which minimizes the least-squares criterion $J(n) = \|e(n)\|^2 = e(n)e^*(n)$. Weights are updated thanks to a stochastic gradient descent as:

$$w(n+1) = w(n) - \mu \cdot \nabla J(n)\,, \tag{5}$$

where $\mu > 0$ is a constant defining the descent step. The problem consists in calculating $\nabla J(n)$ with respect to the quaternionic vector w.

B. The QLMS versions

The sum of componentwise derivations of $J(n)$ was the first way to obtain a QLMS. It has been studied by Took and Mandic in [1], and the given expression is:

$$w(n+1) = w(n) + \mu\big(2e(n)x^*(n) - x^*(n)e^*(n)\big)\,. \tag{6}$$

This recent algorithm is well-used in quaternionic signal processing: for quaternionic adaptive filtering applied to image denoising [8] or wind forecasting [9], for second order statistics [10], [11], for gait recognition [12], etc.

Q. Barthélemy and A. Larue are with CEA, LIST, Data Analysis Tools Laboratory, 91191 Gif-sur-Yvette Cedex, France (e-mail: q.barthelemy@gmail.com, anthony.larue@cea.fr) .

J.I. Mars is with GIPSA-Lab, 38402 Saint Martin d'Hères, France (e-mail: jerome.mars@gipsa-lab.grenoble-inp.fr)

To give a mathematical foundation to this result, a quaternion gradient operator is given by Mandic *et al.* in [2], extending the complex gradient operator [6]. Considering a quaternion $q \in \mathbb{H}$, these new derivative rules are:

$$\frac{\partial q}{\partial q} = \frac{\partial q^*}{\partial q^*} = 1 \qquad \text{and} \qquad \frac{\partial q^*}{\partial q} = \frac{\partial q}{\partial q^*} = -1/2\,. \quad (7)$$

Using these rules, a $\mathbb{HR}$-QLMS is derived from $J(n)$ in [2]:

$$w(n+1) = w(n) + \mu\big(e(n)x^*(n) - \frac{1}{2}x(n)e^*(n)\big)\,. \quad (8)$$

These theoretical results have helped to derive for quaternions generalized gradient descent [13], independent component analysis [14], kernel adaptive filtering [15], affine projection algorithms [16], etc.

A third way is proposed by Took *et al.* in [3] using an involutions-gradient. A iQLMS is thus derived as:

$$w(n+1) = w(n) + \mu\frac{3}{2}e(n)x^*(n)\,. \quad (9)$$

For consistency, this expression has been multiplied by 2, since Took *et al.* minimize the criterion $\frac{1}{2}\,J$ in [3] and not J. To explain the differences between the three versions of the QLMS, a generic form is proposed in [3]:

$$\begin{aligned}\Re[\nabla J(n)] &= -\nu\Re\left[e(n)\Re[x(n)]\right] + \tau\Re\left[e(n)\Im[x(n)]\right]\\ \Im[\nabla J(n)] &= -\rho\Im\left[e(n)\Re[x(n)]\right] + \varsigma\Im\left[e(n)\Im[x(n)]\right]. \quad (10)\end{aligned}$$

It allows to express the three updates in the same form, with only different values for $\nu, \tau, \rho, \varsigma$. The different versions are thus said to be ”topologically similar” [17]. Remark that Jahanchahi *et al.* [17] detail the work [3], but with a different definition of the problem: $d(n) = w(n)^*\,x(n) + e(n)$, where w is conjugated instead of transposed as in Eq. (4).

C. Some problems in the QLMS versions

In this subsection, problems between these three non-similar QLMS versions are notified.

1. In fact, a mistake has been made during the derivation of QLMS of Eq. (6): in [1], the unit imaginary i has been commuted with $e^*(n)$ in Eq. (43). The same mistake has been done too with j in Eq. (44) and with k in Eq. (45).

2. QLMS in Eq. (6) and $\mathbb{HR}$-QLMS in Eq. (8) are not identical, even give or take a multiplicative factor: $x(n)$ is conjugated in the second member of Eq. (6) and not in Eq. (8).

3. Finally, the generic form (10) proposed in [3] to express the three QLMS versions is artificial and does not suppress the existing differences. A quaternionic scalar example ($L = 1$) is chosen, with $e(n) = x(n) = 1 + i$ (special case where quaternions are reduced to complex). For the QLMS, $\nabla J(n) = -4 - 2i$; for the $\mathbb{HR}$-QLMS, $\nabla J(n) = -1$ and for the iQLMS, $\nabla J(n) = -3$. The generic form is not convincing since it is not possible to pretend that these updates are ”topologically similar”.

To conclude this section, we propose to study the first derivation way to compute rigorously the QLMS expression.

IV. QLMS COMPONENTWISE DERIVATION

The proposed derivation is lengthy, but we choose to present the full detailed version to avoid calculus mistakes. The criterion to derive is:

$$\nabla_w J(n) = \nabla_w\big(e(n)e^*(n)\big)\,. \quad (11)$$

Indices n are suppressed from variables e, x and w to lightened calculus. The criterion J is now derived with respect to w:

$$\nabla_w(ee^*) = \nabla_{w_a}(ee^*) + \nabla_{w_b}(ee^*)i + \nabla_{w_c}(ee^*)j + \nabla_{w_d}(ee^*)k \quad (12)$$

$$\begin{aligned}&= e\nabla_{w_a}(e^*) + \nabla_{w_a}(e)e^* + e\nabla_{w_b}(e^*)i + \nabla_{w_b}(e)e^*i\\ &+ e\nabla_{w_c}(e^*)j + \nabla_{w_c}(e)e^*j + e\nabla_{w_d}(e^*)k + \nabla_{w_d}(e)e^*k. \quad (13)\end{aligned}$$

To compute easily the derivations of $e = d - w^Tx$ and $e^* = d^* - x^Hw^*$, the expressions w^Tx and x^Hw^* are expanded:

$$\begin{aligned}w^Tx \;&= {w_a}^Tx_a - {w_b}^Tx_b - {w_c}^Tx_c - {w_d}^Tx_d\\ &+ \big({w_a}^Tx_b + {w_b}^Tx_a + {w_c}^Tx_d - {w_d}^Tx_c\big)\,i\\ &+ \big({w_a}^Tx_c - {w_b}^Tx_d + {w_c}^Tx_a + {w_d}^Tx_b\big)\,j\\ &+ \big({w_a}^Tx_d + {w_b}^Tx_c - {w_c}^Tx_b + {w_d}^Tx_a\big)\,k\,, \quad (14)\end{aligned}$$

$$\begin{aligned}x^Hw^* \;&= {w_a}^Tx_a - {w_b}^Tx_b - {w_c}^Tx_c - {w_d}^Tx_d\\ &+ \big(-{w_a}^Tx_b - {w_b}^Tx_a - {w_c}^Tx_d + {w_d}^Tx_c\big)\,i\\ &+ \big(-{w_a}^Tx_c + {w_b}^Tx_d - {w_c}^Tx_a - {w_d}^Tx_b\big)\,j\\ &+ \big(-{w_a}^Tx_d - {w_b}^Tx_c + {w_c}^Tx_b - {w_d}^Tx_a\big)\,k. \quad (15)\end{aligned}$$

Using these expressions, the four componentwise gradients are now computed as:

$$\begin{aligned}\nabla_{w_a}(ee^*) =& e(-x^*) + (-x)e^*\\ =& -ex^* - xe^*\,, \quad (16)\end{aligned}$$

$$\begin{aligned}\nabla_{w_b}(ee^*)i =& e(x_b + x_ai - x_dj + x_ck)i\\ &+ (x_b - x_ai + x_dj - x_ck)e^*i\\ =& e(-x_a + x_bi + x_cj + x_dk)\\ &+ (x_b - x_ai + x_dj - x_ck)(e_b + e_ai - e_dj + e_ck)\\ =& -ex^* + (\alpha)\,, \quad (17)\end{aligned}$$

$$\begin{aligned}\nabla_{w_c}(ee^*)j =& e(x_c + x_di + x_aj - x_bk)j\\ &+ (x_c - x_di - x_aj + x_bk)e^*j\\ =& e(-x_a + x_bi + x_cj + x_dk)\\ &+ (x_c - x_di - x_aj + x_bk)(e_c + e_di + e_aj - e_bk)\\ =& -ex^* + (\beta)\,, \quad (18)\end{aligned}$$

$$\begin{aligned}\nabla_{w_d}(ee^*)k =& e(x_d - x_ci + x_bj + x_ak)k\\ &+ (x_d + x_ci - x_bj - x_ak)e^*k\\ =& e(-x_a + x_bi + x_cj + x_dk)\\ &+ (x_d + x_ci - x_bj - x_ak)(e_d - e_ci + e_bj + e_ak)\\ =& -ex^* + (\gamma)\,. \quad (19)\end{aligned}$$

So, summing the componentwise gradients (16), (17), (18) and (19), we obtain:

$$\nabla_w(ee^*) = -4ex^* - xe^* + (\alpha) + (\beta) + (\gamma)\,. \tag{20}$$

The three expressions are expanded:

$$\begin{aligned}(\alpha)\ &= x_b e_b + x_b e_a i - x_b e_d j + x_b e_c k - x_a e_b i + x_a e_a \\ &\quad + x_a e_d k + x_a e_c j + x_d e_b j - x_d e_a k + x_d e_d \\ &\quad + x_d e_c i - x_c e_b k - x_c e_a j - x_c e_d i + x_c e_c\,,\end{aligned} \tag{21}$$

$$\begin{aligned}(\beta)\ &= x_c e_c + x_c e_d i + x_c e_a j - x_c e_b k - x_d e_c i + x_d e_d \\ &\quad - x_d e_a k - x_d e_b j - x_a e_c j + x_a e_d k + x_a e_a \\ &\quad + x_a e_b i + x_b e_c k + x_b e_d j - x_b e_a i + x_b e_b\,,\end{aligned} \tag{22}$$

$$\begin{aligned}(\gamma)\ &= x_d e_d - x_d e_c i + x_d e_b j + x_d e_a k + x_c e_d i + x_c e_c \\ &\quad + x_c e_b k - x_c e_a j - x_b e_d j - x_b e_c k + x_b e_b \\ &\quad - x_b e_a i - x_a e_d k + x_a e_c j + x_a e_b i + x_a e_a\,.\end{aligned} \tag{23}$$

Summing these three expressions, we have:

$$\begin{aligned}(\alpha)+(\beta)+(\gamma) &= 3x_b e_b - x_b e_a i - x_b e_d j + x_b e_c k + x_a e_b i \\ &\quad + 3x_a e_a + x_a e_d k + x_a e_c j + x_d e_b j - x_d e_a k + 3x_d e_d \\ &\quad - x_d e_c i - x_c e_b k - x_c e_a j + x_c e_d i + 3x_c e_c\,.\end{aligned} \tag{24}$$

Reordering the terms, we obtain:

$$\begin{aligned}(\alpha)+(\beta)+(\gamma) &= 2\Big(e_a x_a - e_a x_b i - e_a x_c j - e_a x_d k \\ &\quad + e_b x_a i + e_b x_b - e_b x_c k + e_b x_d j + e_c x_a j + e_c x_b k \\ &\quad + e_c x_c - e_c x_d i + e_d x_a k - e_d x_b j + e_d x_c i + e_d x_d\Big) \\ &\quad + \Big(x_a e_a - x_a e_b i - x_a e_c j - x_a e_d k + x_b e_a i + x_b e_b \\ &\quad - x_b e_c k + x_b e_d j + x_c e_a j + x_c e_b k + x_c e_c - x_c e_d i \\ &\quad + x_d e_a k - x_d e_b j + x_d e_c i + x_d e_d\Big)\,.\end{aligned} \tag{25}$$

$$\begin{aligned}&(\alpha)+(\beta)+(\gamma) \\ &\quad = 2(e_a + e_b i + e_c j + e_d k)(x_a - x_b i - x_c j - x_d k) \\ &\qquad + (x_a + x_b i + x_c j + x_d k)(e_a - e_b i - e_c j - e_d k) \\ &\quad = 2ex^* + xe^*\,.\end{aligned} \tag{26}$$

From Eq. (20) and Eq. (26), we finally obtain:

$$\nabla_w(ee^*) = -4ex^* - xe^* + 2ex^* + xe^* = -2ex^*\,. \tag{27}$$

So, the QLMS expression is:

$$w(n+1) = w(n) + \mu\, 2\, e(n) x^*(n)\,. \tag{28}$$

V. Remarks

After this mathematical development, we give some comments about this exact derivative version. The different QLMS versions are summed up in Table I.

1. The expression of the componentwise CLMS [5] is *exactly* recovered by the QLMS given in Eq. (28). Consequently, the presented QLMS is a valid generalization of the real case (LMS [4] with $q_b = q_c = q_d = 0$) and of the complex case (CLMS [5] with $q_c = q_d = 0$).

TABLE I
Summary table for the QLMS versions for $J(n) = \|e(n)\|^2$.

QLMS version	$\nabla J(n)$
QLMS original	$-\big(2e(n)x^*(n) - x^*(n)e^*(n)\big)$
$\mathbb{HR}$-QLMS	$-\big(e(n)x^*(n) - \frac{1}{2}x(n)e^*(n)\big)$
iQLMS	$-\frac{3}{2}\, e(n)x^*(n)$
QLMS	$-2\, e(n)x^*(n)$

2. Considering Eq. (28), the iQLMS of [3] expressed in Eq. (9) is recovered too, give or take a multiplicative factor. In [17], QLMS, $\mathbb{HR}$-QLMS and iQLMS are compared with respect to their convergence speeds, and iQLMS was observed to be the most rapid. After the presented derivation, it seems normal that iQLMS has given the best results.

3. The differences between QLMS versions are not due to quaternions noncommutativity, as explained in [18]. The difference with the original QLMS [1] is caused by a commutativity mistake in the componentwise derivation (Section III-C).

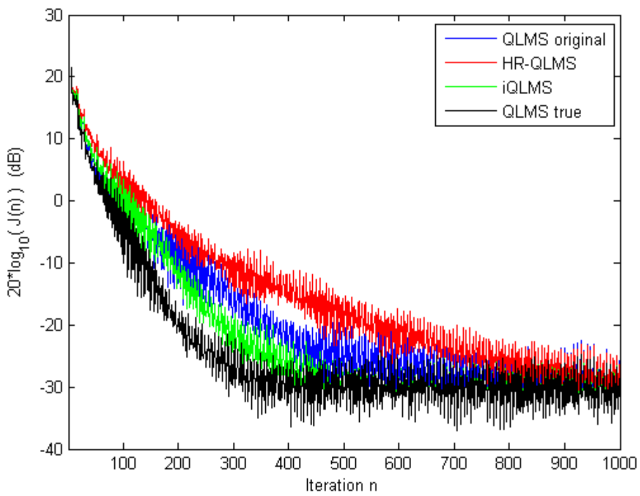
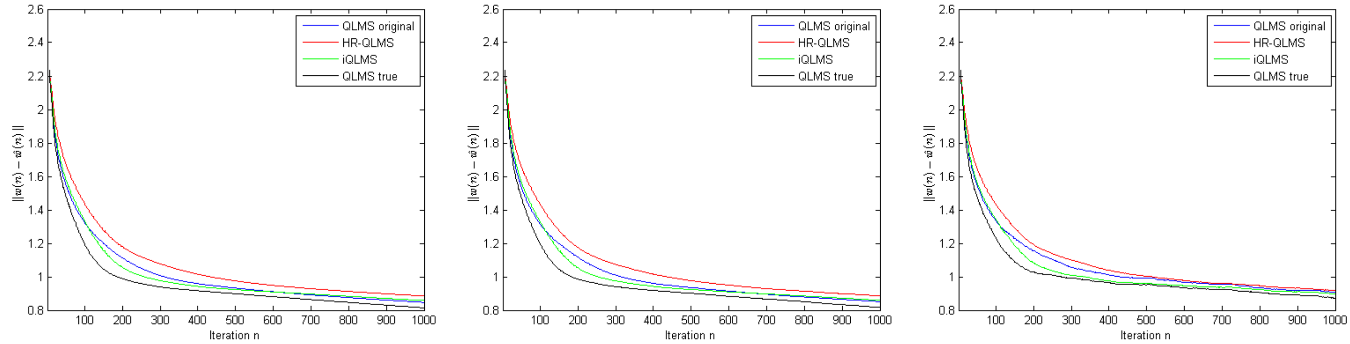


Fig. 1. Evolution of the criterion $J(n)$ in dB as a function of the iteration n, averaged 100 times, for the original QLMS, the $\mathbb{HR}$-QLMS, the iQLMS and the true QLMS.

VI. Comparison on Simulated Data

A comparison between the different QLMS versions is made in this section. A signal $x \in \mathbb{H}^N$ is created with $N = 1000$ samples, and a filter $w \in \mathbb{H}^L$ is composed of uniformly distributed unit quaternions, with $L = 5$ samples. The signal $d \in \mathbb{H}^N$ is formed using the model defined in Eq. (4). The different versions of the QLMS are used on these data: the original QLMS given in Eq. (6), the $\mathbb{HR}$-QLMS in Eq. (8), the iQLMS in Eq. (9) and the true QLMS in Eq. (28). The descent step is the same for the four versions, with $\mu = 0.01$. For each version, the criterion $J(n)$ is computed at each iteration/sample n. The error on the filter $\|w(n) - \hat{w}(n)\|$ is computed too, with $\hat{w}(n)$ defined as the estimated filter.

Results, averaged 100 times, are plotted in the following figures, with the original QLMS in blue, the $\mathbb{HR}$-QLMS in red,



(a) SNR: 30 dB (b) SNR: 20 dB (c) SNR: 10 dB

Fig. 3. Evolution of the estimation error $\|w(n) - \hat{w}(n)\|$ as a function of the iteration n, averaged 100 times, for the original QLMS, the $\mathbb{HR}$-QLMS, the iQLMS and the true QLMS. A uniformly distributed unit quaternionic noise is added at different SNR: 30 dB (a), 20 dB (b) and 10 dB (c).

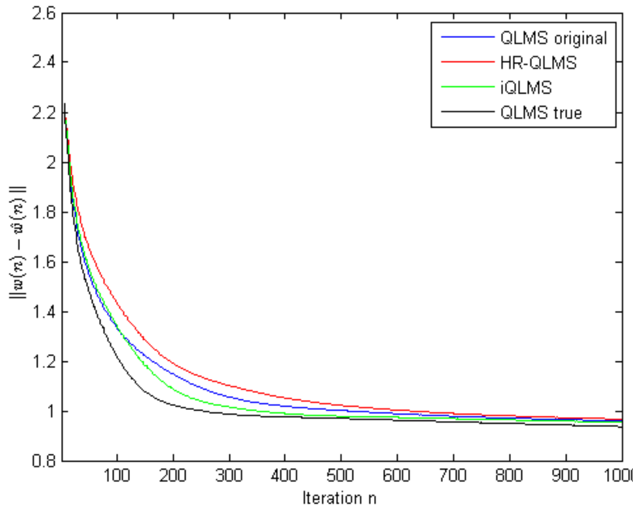


Fig. 2. Evolution of the estimation error $\|w(n) - \hat{w}(n)\|$ as a function of the iteration n, averaged 100 times, for the original QLMS, the $\mathbb{HR}$-QLMS, the iQLMS and the true QLMS.

the iQLMS in green and the true QLMS in black. In Fig. 1, the criterion $J(n)$ is plotted in dB as a function of the iteration n. We observe that the convergence of the true QLMS is faster than the original QLMS and the iQLMS ones, themselves faster than the $\mathbb{HR}$-QLMS one. In Fig. 2, the estimation error $\|w(n) - \hat{w}(n)\|$ is plotted as a function of the iteration n. These figures show the recovery property of the algorithms. As previously, we observe that the convergence of the true QLMS is always better than the original QLMS and the iQLMS ones which have similar behaviors, themselves better than the $\mathbb{HR}$-QLMS one.

Experimental protocol is now slightly changed since a uniformly distributed unit quaternionic noise is added to signals d. Different signal-to-noise ratios (SNR) are considered: 30, 20 and 10 dB. Results are respectively shown in Fig. 3(a), 3(b) and 3(c) and the curves are quite similar. The previous observations about algorithms behaviors are still verified with the added noise.

To conclude, this comparison highlights the optimality of the proposed QLMS, including the multiplicative factor 2 with respect to $3/2$ for the iQLMS.

VII. Conclusion

After a review of the different QLMS derivations, the componentwise way has been examined scrupulously. Since a mistake has been done in the original version of Took and Mandic [1], the derivation has been detailed and the correct expression has been proposed. Comparisons on simulated data have validated the theoretical results. Finally, this method has many applications, already cited in Section III.

Prospects are to investigate rigorously the quaternion gradient operator and the $\mathbb{HR}$ derivative rules [2], which had been written to support the incorrect formula of the original QLMS [1]. Especially as the quaternion gradient operator has been observed to be invalid in [19] for the derivation of quaternionic sparse pursuits.

References

[1] C. Took and D. Mandic, "The quaternion LMS algorithm for adaptive filtering of hypercomplex processes," *IEEE Trans. on Signal Processing*, vol. 57, pp. 1316–1327, 2009.

[2] D. Mandic, C. Jahanchahi, and C. Took, "A quaternion gradient operator and its applications," *IEEE Signal Processing Letters*, vol. 18, pp. 47–50, 2011.

[3] C. Took, C. Jahanchahi, and D. Mandic, "A unifying framework for the analysis of quaternion valued adaptive filters," in *Conf. Record of the Asilomar Conf. on Signals, Systems and Comput.*, 2011, pp. 1771–1774.

[4] B. Widrow, *Adaptive Filters*. Aspects of Network and System Theory, 1971, pp. 563–586.

[5] B. Widrow, J. McCool, and M. Ball, "The complex LMS algorithm," *Proceedings of the IEEE*, vol. 63, pp. 719–720, 1975.

[6] D. Brandwood, "A complex gradient operator and its application in adaptive array theory," *IEE Proceedings F - Communications, Radar and Signal Processing*, vol. 130, pp. 11–16, 1983.

[7] F. Zhang, "Quaternions and matrices of quaternions," *Linear Algebra and its Applications*, vol. 251, pp. 21–57, 1997.

[8] C. Took and D. Mandic, "Quaternion-valued stochastic gradient-based adaptive IIR filtering," *IEEE Trans. on Signal Processing*, vol. 58, pp. 3895–3901, 2010.

[9] C. Took, G. Strbac, K. Aihara, and D. Mandic, "Quaternion-valued short-term joint forecasting of three-dimensional wind and atmospheric parameters," *Renewable Energy*, vol. 36, pp. 1754–1760, 2011.

[10] F. Neto and V. Nascimento, "A novel reduced-complexity widely linear QLMS algorithm," in *Proc. IEEE Workshop on Statistical Signal Processing SSP '11*, 2011, pp. 81–84.

[11] ——, "Second-order analysis of the RC-WL-QLMS algorithm," in *Proc. Int. Symposium on Wireless Communication Systems ISWCS*, 2012, pp. 466–470.

[12] A. Bravi and A. Sabatini, "A multidimensional approach to postural sway modeling," in *Proc. IEEE Int. Workshop on Medical Measurements and Applications MeMeA*, 2010, pp. 121–124.

[13] M. Wang, C. Took, and D. Mandic, “A class of fast quaternion valued variable stepsize stochastic gradient learning algorithms for vector sensor processes,” in *Int. Joint Conf. on Neural Networks IJCNN*, 2011, pp. 2783–2786.

[14] S. Javidi, C. Took, and D. Mandic, “Fast independent component analysis algorithm for quaternion valued signals,” *IEEE Trans. on Neural Networks*, vol. 22, pp. 1967–1978, 2011.

[15] T. Ogunfunmi and T. Paul, “An alternative kernel adaptive filtering algorithm for quaternion-valued data,” in *Signal Information Processing Association Annual Summit and Conference APSIPA ASC*, 2012, pp. 1–5.

[16] C. Jahanchahi, C. Took, and D. Mandic, “A class of quaternion valued affine projection algorithms,” *Signal Process.*, vol. 93, pp. 1712–1723, 2013.

[17] ——, “On gradient calculation in quaternion adaptive filtering,” in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing ICASSP 2012*, 2012, pp. 3773 –3776.

[18] C. Jahanchahi and D. Mandic, “A unifying approach to quaternion adaptive filtering: Addressing the gradient and convergence,” *Preprint arXiv:1310.5612*, 2013.

[19] Q. Barthélemy, A. Larue, and J. Mars, “Sparse approximations for quaternionic signals,” *Advances in Applied Clifford Algebras*, in press, 2014.