MULTICHANNEL QUATERNION LEAST MEAN SQUARE ALGORITHM

Clive Cheong Took[*] *and Yili Xia*[†]

[*]Department of Electronic Engineering, Royal Holloway University of London
[†]School of Information Science and Engineering, Southeast University, China
Emails: clive.cheongtook@rhul.ac.uk, yili_xia@seu.edu.cn

ABSTRACT

Quaternion least-mean-quare (QLMS) algorithm has been thoroughly investigated in the past for the adaptive filtering of 3D and 4D signals. However, QLMS is restricted to one-channel processing of such 3D and 4D processes. To address this shortcoming, we proposed the multichannel QLMS so that we can exploit the additional information from the multitude of 3D-4D vector sensors. For rigour, we provide the stability bound of the multichannel QLMS and demonstrate its robustness against noise. Simulations on real world 4D signals support the analysis.

Index Terms— hypercomplex, quaternions, adaptive filters, multichannel processing

1. INTRODUCTION

There are many natural phenomena that are inherently three-dimensional (3D) and four-dimensional (4D), yet it is common to treat these phenomena as multichannel signals instead of 3D or 4D signals. Perhaps due to convention and convenience or perhaps due to lack of modelling in 3D and 4D. Learning in the hypercomplex domain enables us us to process 3D and 4D data as a single entity rather than modelling as a multichannel entity, hence preserving the integrity of the data. The cross-product $\mathbf{a} \times \mathbf{b}$ is well defined for $\mathbb{R}^3$. This simple yet useful mathematical operation indicates the suitability of quaternions for 3D modelling, since the quaternion product involves the cross product as

$$\begin{aligned} x_1 x_2 &= Sx_1 Sx_2 - Vx_1 Vx_2 + Sx_2 Vx_2 \\ &\quad + Sx_1 Vx_2 + \underbrace{Vx_1 \times Vx_2}_{\text{cross product}} \end{aligned} \qquad (1)$$

where S and V denote the scalar and vector part of the quaternion variable. To this end, quaternion signal processing has emerged as a candidate tool to cater specifically for these 3D and 4D processes. For instance, quaternion least mean square algorithm (QLMS) has been exploited in wind forecasting and 3D audio systems [1], quaternion principal component analysis for colour images [2], and weighted linear combiner for modelling 3D tremors [3].

The emergence of quaternion algorithms and theories has particularly sparked an interest in the adaptive signal processing community. Nascimento *et al.* have proposed low complexity adaptive filters [4, 5], and Liu proposed novel adaptive beamforming algorithms [6]. The analysis of these quaternion-valued adaptive filters have been investigated in [7], whereas Mars *et al.* discussed the derivations of quaternion adaptive filters [8]. However, the adaptive filtering of *multiple* 3D and 4D processes simultaneously is still lacking and is addressed in this work in the form of multichannel QLMS.

The remainder of this paper is organised as follows. Fundamentals of quaternions are revisited in Section 2. The multichannel QLMS is derived in Section 3 followed by some remarks regarding issues with the implementation of multichannel QLMS. Section 4 provides simulation results demonstrating the performance of the proposed algorithm and finally, Section 5 concludes this paper.

2. QUATERNIONS AND THE INVOLUTION-GRADIENT

A quaternion $x = x_a + \imath x_b + \jmath x_c + \kappa x_d$ is comprised of a real (scalar) part x_a and three imaginary components x_b, x_c, x_d in its vector part. The vector product of these three imaginary parts makes the quaternion product between x and y noncommutative, i.e.

$$xy \neq yx \qquad (2)$$

As in the complex case, the conjugate of a quaternion variable is obtained by negating the imaginary parts as

$$x = x_a - \imath x_b - \jmath x_c - \kappa x_d \qquad (3)$$

The additional degrees of freedom from complex to quaternion domain means that there exist other useful algebraic operations such as the quaternion involutions (self-inverse mappings)

$$x^\eta = -\eta x \eta, \qquad \eta \in \{\imath, \jmath, \kappa\} \qquad (4)$$

Alternatively, these involutions can be expressed as

$$\begin{aligned} x^{\imath} &= x_a + \imath x_b - \jmath x_c - \kappa x_d \\ x^{\jmath} &= x_a - \imath x_b + \jmath x_c - \kappa x_d \\ x^{\kappa} &= x_a - \imath x_b - \jmath x_c + \kappa x_d \end{aligned} \tag{5}$$

In this way, these involutions enable componentwise conjugation. For example, $x^{\imath *}$ means the conjugation of only the $\imath-$imaginary component of x. The relationships between these involutions are governed by

$$\begin{aligned} x^* &= \frac{1}{2}(x^{\imath} + x^{\jmath} + x^{\kappa} - x) \\ x &= \frac{1}{2}(x^{\imath *} + x^{\jmath *} + x^{\kappa *} - x^*) \end{aligned} \tag{6}$$

The pair of relationships in (6) will be used in the derivation of the multichannel QLMS algorithms.

2.1. The Involution Gradient

The so called pseudogradient is often employed to calculate gradients due to its straightforward formulation as the sum of componentwise gradients, i.e.

$$\frac{\partial J}{\partial x} = \frac{\partial J}{\partial x} + \imath\frac{\partial J}{\partial x_{\imath}} + \jmath\frac{\partial J}{\partial x_{\jmath}} + \kappa\frac{\partial J}{\partial x_{\kappa}} \tag{7}$$

However, the process of calculating componentwise gradients is tedious. To address this issue, we introduced the $\mathbb{HR}$-calculus [9], so that derivatives can be obtained directly in the quaternion space. For example, $(\partial x^*)/(\partial x^*) = 1$, whereas based on (6) we have

$$\begin{aligned} \frac{\partial x^*}{\partial x} &= \frac{\partial[0.5(x^{\imath} + x^{\jmath} + x^{\kappa} - x)]}{\partial x} \\ &= -0.5x \end{aligned} \tag{8}$$

Due to the relationships in (6), we can also take quaternion gradients with respect to the involutions (5), which was referred as the involution-gradient, i.e.

$$\nabla_{x^{\eta}} J = \sum_{\eta=\{\imath,\jmath,\kappa\}} \frac{\partial J}{\partial x^{\eta}} \tag{9}$$

It was shown in [10] that the steepness of the involution-gradient was marginally greater than the $\mathbb{HR}^*-$derivative, yet both lead to the same solution at the local minima. It is the involution-gradient that we will adopt to derive multichannel QLMS algorithms.

3. THE MULTICHANNEL QLMS ALGORITHM

For the multichannel QLMS algorithm, the qth output can be expressed in terms of its inputs $\mathbf{x}$ and filter coefficients $\mathbf{w}$ as

$$y_q(n) = \sum_{p=1}^{P} \mathbf{w}_{pq}^{\mathrm{H}} \mathbf{x}_p(n) \qquad q = 1, \ldots, Q \tag{10}$$

where

$$\begin{aligned} \mathbf{w}_{pq} &= [w_{pq,1} \ \ w_{pq,2} \ \cdots \ w_{pq,L}]^{\mathrm{T}} \\ \mathbf{x}_p(n) &= [x(n) \ \ x(n-1) \ \cdots \ x(n-L+1)]^{\mathrm{T}} \end{aligned}$$

The superscripts $(\cdot)^{\mathrm{H}}$ and $(\cdot)^{\mathrm{T}}$ denote the Hermitian and transpose operator respectively. More generally, the output vector $\mathbf{y}(n)$ can be expressed as

$$\mathbf{y}(n) = \mathbf{W}^{\mathrm{H}}\mathbf{x}(n)$$

$$\begin{bmatrix} y_1(n) \\ y_2(n) \\ \vdots \\ y_Q(n) \end{bmatrix} = \begin{bmatrix} \mathbf{w}_{11} & \mathbf{w}_{12} & \ldots & \mathbf{w}_{1Q} \\ \mathbf{w}_{21} & \mathbf{w}_{21} & \ldots & \mathbf{w}_{2Q} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{w}_{P1} & \mathbf{w}_{P2} & \ldots & \mathbf{w}_{PQ} \end{bmatrix}^{\mathrm{H}} \begin{bmatrix} \mathbf{x}_1(n) \\ \mathbf{x}_2(n) \\ \vdots \\ \mathbf{x}_P(n) \end{bmatrix} \tag{11}$$

The error for the qth channel can be expressed as

$$e_q(n) = d_q(n) - y_q(n) \qquad q = 1, ..., Q \tag{12}$$

The cost function of qth channel can be formulated as [11]

$$J_q(n) = e_q(n)e_q^*(n) \tag{13}$$

For the qth channel, the update for its coefficients can be computed using the steepest descent:

$$\mathbf{W}_q(n+1) = \mathbf{W}_q(n) - \mu\nabla J_q(n) \tag{14}$$

where $\mathbf{W}_q$ is given by

$$\begin{bmatrix} w_{1q,1} & w_{1q,2} & \ldots & w_{1q,L} \\ w_{2q,1} & w_{2q,2} & \ldots & w_{2q,L} \\ \vdots & \vdots & \ddots & \vdots \\ w_{Pq,1} & w_{Pq,2} & \ldots & w_{Pq,L} \end{bmatrix}$$

The derivative in (14) can be formulated based on the involution gradient as

$$\nabla J_q(n) = \sum_{\eta=\{\imath,\jmath,\kappa\}} e_q(n)\frac{\partial e_q^*(n)}{\partial \mathbf{W}_q^{\eta}(n)} + \frac{\partial e_q(n)}{\partial \mathbf{W}_q^{\eta}(n)} e_q^*(n) \tag{15}$$

As the conjugate of the error (12) is given by

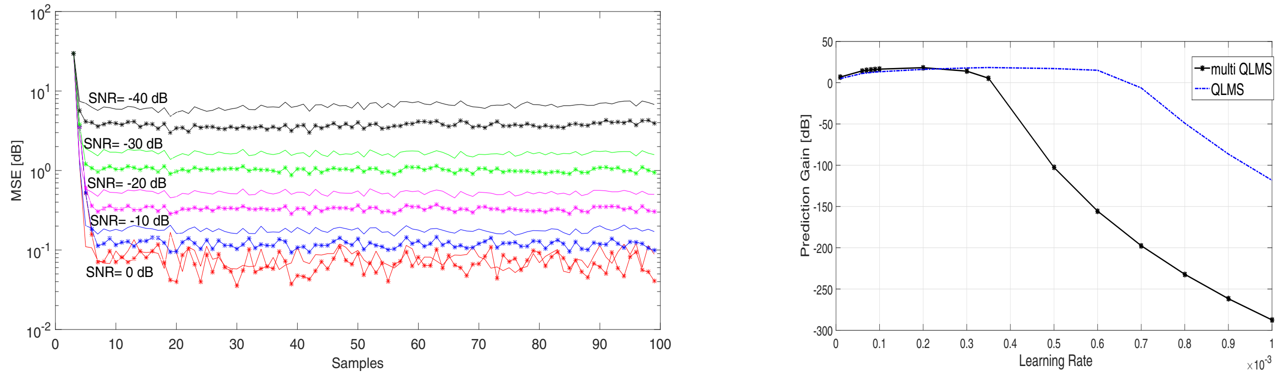
$$e_q^*(n) = d_q^*(n) - \sum_{p=1}^{P} \mathbf{x}_p^{\mathrm{H}}(n)\mathbf{w}_{pq} \qquad q = 1, ..., Q \tag{16}$$

the involution relationship (6) means that $\frac{\partial e_q^*(n)}{\partial \mathbf{W}_q^{\eta}(n)} = 0$, since Eq. (16) is dependent of $\mathbf{W}_q^{\eta *}(n)$, but independent of $\mathbf{W}_q^{\eta}(n)$. Similarly, the other derivative in (15) can be calculated as

$$\frac{\partial e_q(n)}{\partial \mathbf{W}_q^{\eta}(n)} = -0.5\mathbf{x}(n) \tag{17}$$

Based (14)-(15), the coefficients corresponding to the qth channel output of multichannel QLMS can be learnt as

$$\mathbf{W}_q(n+1) = \mathbf{W}_q(n) + \mu\mathbf{X}(n)e_q^*(n) \tag{18}$$



(a) (b)

Fig. 1. The left plot (a) shows the learning curves of QLMS ('-') and multichannel QLMS('-*'). The right plot (b) shows the stability range of the learning rate or stepsize.

where any constant has been absorbed in the learning rate/step size and $\mathbf{X}(n)$ is given by

$$\begin{bmatrix} x_1(n) & x_1(n-1) & \dots & x_1(n-L+1) \\ x_2(n) & x_2(n-1) & \dots & x_2(n-L+1) \\ \vdots & \vdots & \ddots & \vdots \\ x_P(n) & x_P(n-1) & \dots & x_P(n-L+1) \end{bmatrix}$$

Instead of updating each qth adaptive filter individually (18), we can generalise the update of the multichannel QLMS as

$$\mathbf{W}(n+1) = \mathbf{W}(n) + \mu \mathbf{x}(n)\mathbf{e}^{\mathrm{H}}(n) \tag{19}$$

where the variables $\mathbf{W}(n)$ and $\mathbf{x}(n)$ have the same dimensionality as in (11), and the error vector is given by

$$\mathbf{e}(n) = [e_1(n) \quad e_2(n) \quad \cdots \quad e_q(n)]^T$$

Each learning coefficient of (19) can be expressed as

$$\begin{aligned} w_{pq,l}(n+1) &= w_{pq,l}(n) + \Delta w_{pq,l}(n+1) & (20) \\ &= w_{pq,l}(n) + \mu x_p(n-l)e_q^*(n) \qquad l = 1, ..., L \end{aligned}$$

Remark 1: At first sight from (20), it looks like each update of the qth adaptive filter does not depend on the other filters, since its update depends only on the qth error. However, on expanding the qth error, we have

$$\Delta w_{pq,l}(n+1) = \mu x_p(n)[d_q^*(n) - \sum_{i=1}^{P} \mathbf{x}_i^{\mathrm{H}}(n)\mathbf{w}_{iq}(n)] \tag{21}$$

Notice that each learning coefficient is updated based on *all* input channels $\{x_p\}_{p=1}^{P}$. This leads to the 'dependency' between each channel adaptive filter, as observed in [11].

Remark 2: Remark 1 also demonstrates that the multichannel QLMS exploits cross-channel correlation between 3D or 4D processes, since the update depends on the correlation term $E\{x_p(n)\sum_{i=1}^{P}\mathbf{x}_i(n)\}$ upon taking its expectation.

Remark 3: The correlation term $E\{x_p(n)\sum_{i=1}^{P}\mathbf{x}_i\}$, however, does not guarantee that the coupling within each quaternion input channel $x_i(n)$ is exploited. To do so, we can take advantage of advances in quaternion statistics by augmenting the input vector with the three involutions, i.e. $x^a(n) = [x \;\; x^\imath \;\; x^\jmath \;\; x^\kappa]$ [12].

Remark 4: The blockbase implementation in (19) is faster than channelwise implementation in (18), if implemented in Matlab. The latter requires loop-based coding, whereas the former can be computed much faster due to the 'vectorisation' facility implemented in Matlab.

Remark 5: It is straightforward to show that the range of values of the learning rate for multichannel QLMS (19) is $0 < \mu < \frac{2}{\lambda_{\text{max}}}$, where λ_{max} is maximum eigenvalue of the *multichannel* correlation matrix $E\{\mathbf{x}\mathbf{x}^H\}$. A more practical bound of λ_{max} is $\frac{L}{N}\sum_{i=0}^{N}\sum_{p=1}^{P}x_p^2(n-i)$, where N is the number of samples that has been processed, and L is the filter length.

4. SIMULATION

The performance and convergence properties of our proposed multichannel QLMS was assessed against QLMS in a forecasting application with a prediction horizon of one step ahead. For this forecasting task, wind data which comprised of two anemometers (sensors) recordings of wind speeds (East-West, North-South, Vertical direction) and temperature (degree Celsius) was considered. Hence, the total number of channels for quaternion-valued processes were two. For rigour, two performance metrics were used, i.e. the traditional mean-square error (MSE) and the prediction gain R_p, which can computed as

$$R_p = 10\log_{10}(\sigma_x^2/\sigma_n^2) \quad [\text{dB}]$$

where σ_x^2 and σ_n^2 are the estimated sample variances of the input and the additive white Gaussian noise of all channels. To assess the robustness of the algorithms against noise, 50 Monte Carlo simulations were run and averaged out for each value of signal-noise-ratio (SNR) and shown in the left plot of Fig. 1. As expected, when the SNR was decreased, the performance of multichannel QLMS (denoted as '-*') and QLMS (denoted as '-') deteriorated.

Remark 6: The cross-correlation between the two quaternion-valued channels benefited the multi-channel QLMS and made it more robust against noise, whereas the QLMS could not exploit this correlation due to its inherent uni-channel processing. This was expected due to *Remark 2*.

To investigate the convergence and stability properties of QLMS and multichannel QLMS, an experiment was carried out to measure the prediction gain of both algorithms over a range of stepsizes/learning rates, and the results are shown in the right plot of Fig. 1. If we consider the performance of both algorithms to be 'reasonable' when the prediction gain is greater than 0 dB, then the range of stepsize for QLMS is

$$0 < \mu < 0.68 \times 10^{-3}$$

whereas for the multichannel QLMS, the range is

$$0 < \mu < 0.35 \times 10^{-3}$$

Remark 7: In our simulation, the range of values of stepsize for QLMS is approximately twice that of multichannel QLMS. This observation confirms *Remark 5* (i.e. $0 < \mu < \frac{2}{\lambda_{\max}}$), since the *two channels* of multichannel QLMS makes it likely that its $\lambda_{\max}$ is twice that of the *one channel* of QLMS.

5. CONCLUSION

We have proposed the multichannel QLMS algorithm that can process 3D and 4D processes simultaneously. It was shown that the additional information of the multichannel processing over the unichannel processing made our proposed algorithm more robust against noise. However, the cost of processing more channels means that the stability range of the multichannel QLMS is likely to be much less than that of QLMS. This is due to the dependency between channels, i.e. the error of one channel can affect the other channels as well, following *Remark 2*. Future work will involve the investigation of augmented statistics in deriving the multichannel QLMS, however, we expect the same results as in this paper, when comparing it against the widely linear QLMS.

6. REFERENCES

[1] F. Ortolani, D. Comminiello, and A. Uncini, "The widely linear block quaternion least mean square algorithm for fast computation in 3d audio systems," *IEEE International Workshop on Machine Learning for Signal Processing*, Sept. 2016.

[2] N. Le Bihan and S. J. Sangwine, "Quaternion principal component analysis of color images," *IEEE International Conference on Image Processing*, Sept. 2003.

[3] K. Adhikari, S. Tatinati, W. T. Ang, K. C. Veluvolu, and K. Nazarpour, "A quaternion weighted Fourier linear combiner for modeling physiological tremor," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 11, 2016.

[4] F. G. A. Neto and V. H. Nascimento, "A novel reduced-complexity widely linear QLMS algorithm," *IEEE Statistical Signal Processing Workshop*, June 2011.

[5] Vítor H. Nascimento and Rodrigo C. de Lamare, "A low-complexity strategy for speeding up the convergence of convex combinations of adaptive filters," *IEEE International Conference on Acoustics, Speech, and Signal Processing*, March 2012.

[6] X. Lan and W. Liu, "Fully quaternion-valued adaptive beamforming based on crossed-dipole arrays," *Electronics*, vol. 6, no. 2, pp. 34, 2017.

[7] Min Xiang, Sithan Kanna, and Danilo P. Mandic, "Performance analysis of quaternion-valued adaptive filters in nonstationary environments," *IEEE Transactions on Signal Processing*, vol. 66, no. 6, pp. 1566–1579, 2018.

[8] Q. Barthelemy, A. La Rue, and J. Mars, "About QLMS derivations," *IEEE Signal Processing Letters*, vol. 21, no. 2, pp. 240–243, 2014.

[9] D. 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.

[10] C. Jahanchahi, C. C. Took, and D. Mandic, "On gradient calculation in quaternion adaptive filtering," *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 3773–3776, 2012.

[11] J. Benesty, T. Gansler, Y. Huang, and M. Rupp, "Adaptive algorithms for MIMO acoustic echo cancellation," in *Audio Signal Processing for Next-Generation Multimedia Communication Systems*, Y. Huang and J. Benesty, Eds., pp. 119–147. Springer, 2007.

[12] C. Cheong Took and D. Mandic, "Augmented second-order statistics of quaternion random signals," vol. 2, pp. 214–224, 2011.