

THE WIDELY LINEAR QUATERNION RECURSIVE TOTAL LEAST SQUARES

Thiannithi Thanthawarittisai¹, Felipe Tobar², Anthony G. Constantinides¹, and Danilo P. Mandic¹

¹Department of Electrical and Electronic Engineering, Imperial College London

²Department of Engineering, University of Cambridge

Email: tt1513@imperial.ac.uk, fat25@cam.ac.uk, {a.constantinides, d.mandic}@imperial.ac.uk

ABSTRACT

A widely linear quaternion recursive total least squares (WL-QRTLS) algorithm is introduced for the processing of $\mathbb{Q}$ -improper processes contaminated by noise. The total least squares for quaternions (QTLS) is a generalisation of the real-valued total least squares and is introduced rigorously, starting from the existence condition for low-rank approximation of quaternion matrices. Then, a quaternion Rayleigh quotient (QRQ) is defined to establish the link between the QTLS solution and the minimisation of the QRQ. Finally, the rank-one update formula is employed to allow for fast iterative solution based on the QRQ. Through simulations, the WL-QRTLS was shown to exhibit superior performance, under perturbations on both input and output signals, to other adaptive filtering of the same class - the widely linear quaternion least mean squares (WL-QLMS) and the widely linear quaternion recursive least squares (WL-QRLS). The experiments on both synthetic and real-world $\mathbb{Q}$ -improper processes supported the analysis.

Index Terms— widely linear QRTLS, total least squares, Rayleigh quotient, low-rank approximation, quaternions

1. INTRODUCTION

Quaternion-valued signal processing is a very convenient alternative to 3-D and 4-D multichannel statistical signal processing [1], as conventional techniques employed typically do not fully accommodate the *coupled* nature of the information within the channels [2]. Most approaches based on channel-wise processing are often inadequate, as these data channels are typically correlated. Unlike channelwise processing, the quaternion domain allows for the joint modeling of 3-D and 4-D signals, and its algebraic structure naturally accounts for the coupling between the signal components [3].

The quaternion LMS (QLMS) adaptive filter was first introduced in [4] and was shown to outperform both bivariate complex LMS and quadrivariate LMS for multichannel data processing. The first adaptive algorithm based on the widely linear model was proposed in [5] termed the widely linear QLMS (WL-QLMS) to process quaternion signals, both $\mathbb{Q}$ -proper and $\mathbb{Q}$ -improper. The quaternion recursive least

squares algorithm based on the widely linear model (WL-QRLS) has also been developed [6]. Although these results provide the fundamental insight into quaternion-valued adaptive signal processing, none of them has explicitly taken into account the effect of noise on both input and output signals.

So far, in the quaternion domain, there exists only an offline method for detecting quaternion signals in additive noise [7], while methods based on quaternion regression are still lacking. The concept of total least squares (TLS) is known to yield a better approximate and robust solution to systems of linear equations, when the variables of both sides are contaminated by noise [8]. The technique was further improved to tackle real-time estimation via recursive total least squares (RTLS) [9]. It is therefore natural to consider an adaptive algorithm based on the TLS for quaternions (QTLS).

In this paper, we first briefly explain mathematical concepts which support the existence of QTLS. Then, we introduce quaternion Rayleigh quotient (QRQ) to express the QTLS solution in an analytical form. Finally, we propose the widely linear quaternion recursive total least squares (WL-QRTLS) based on the iterative QTLS solution tracking via QRQ, called quaternion Rayleigh quotient iteration (QRQI).

2. TOTAL LEAST SQUARES FOR QUATERNIONS

For the existence of TLS over quaternions, the low-rank approximation theorem needs to hold for quaternion matrices. In the complex-valued case, fundamental concepts supporting this theorem are the singular value decomposition and rank-nullity theorem. Analogously, in the quaternion domain, the singular value decomposition for matrices of quaternions (SVD $\mathbb{Q}$) exists [10, 11]. For the rank-nullity theorem, it was proved such a property holds for the *left* D -module of a division ring D [12], and since the *right* D -module of a ring D^{op} is the opposite division algebra of the *left* D -module, it immediately gives the result for the *right* D -module, and thereby the quaternion *right* Hilbert space, since a module is a generalised notion of a vector space [13, 14]. Note that there are many advantages of adopting the *right* topology for the analysis. For example, it retains the classical rules of matrix algebra [15]. Moreover, the singular values mathematically relate to the *right* eigenvalues [16], facilitating the analysis.

2.1. Low-Rank Approximation for Quaternions

For the case of real-valued matrices, the theorem below shows the best approximation of a particular matrix in the sense of Frobenius norm is its reduced-rank SVD [17]. The analogous statement for quaternion matrices is shown to hold as well.

Theorem 1. [17, 18] Let $\mathbf{A}_p \in \mathbb{H}^{m \times n}$ with rank $p < r$ be the rank- p SVDQ of $\mathbf{A} \in \mathbb{H}^{m \times n}$ with rank $r \leq \min(m, n)$. The SVDQs of $\mathbf{A}$ and $\mathbf{A}_p$ are respectively given by

$$\mathbf{A} = \sum_{i=1}^r \sigma_i \mathbf{u}_i \mathbf{v}_i^H \quad \text{and} \quad \mathbf{A}_p = \sum_{i=1}^p \sigma_i \mathbf{u}_i \mathbf{v}_i^H. \quad (1)$$

where $\sigma_i, \mathbf{u}_i$ and $\mathbf{v}_i \forall i$, are respectively the singular values (arranged in decreasing order), the left singular vectors and the right singular vectors associated with $\mathbf{A}$ and $\mathbf{A}_p$. If $\mathbf{B} \in \mathbb{H}^{m \times n}$ is an arbitrary matrix of rank p , then

$$\min_{\mathbf{B}} \|\mathbf{A} - \mathbf{B}\|_F = \|\mathbf{A} - \mathbf{A}_p\|_F = \sqrt{\sum_{i=p+1}^r \sigma_i^2} \quad (2)$$

2.2. Solution to the Total Least Squares for Quaternions

Generalising the real-valued TLS problem [8], the statement of the QTLS problem can be outlined as follows.

Consider an overdetermined set of linear equations $\mathbf{X}\mathbf{w} \approx \mathbf{d}$, where $\mathbf{X} \in \mathbb{H}^{m \times n}$ and $\mathbf{d} \in \mathbb{H}^{m \times 1}$ are the given data and the coefficients $\mathbf{w} \in \mathbb{H}^{n \times 1}$ are unknown. Now, consider the correction $\hat{\mathbf{X}}\hat{\mathbf{w}} = \hat{\mathbf{d}}$. The QTLS problem seeks to solve the system

$$\begin{aligned} &\text{minimise} \quad \|\mathbf{X}; \mathbf{d}\| - \|\hat{\mathbf{X}}; \hat{\mathbf{d}}\|_F \\ &\quad [\hat{\mathbf{X}}; \hat{\mathbf{d}}] \\ &\text{subject to} \quad \hat{\mathbf{d}} \in R(\hat{\mathbf{X}}) \end{aligned} \quad (3)$$

where $R(\hat{\mathbf{X}})$ denotes the *right* column space of the matrix $\hat{\mathbf{X}}$. Note that the statement $\hat{\mathbf{d}} \in R(\hat{\mathbf{X}})$ is equivalent to there exists $\hat{\mathbf{w}}$ such that $\hat{\mathbf{X}}\hat{\mathbf{w}} = \hat{\mathbf{d}}$. Once $[\hat{\mathbf{X}}; \hat{\mathbf{d}}]$ is found, the weight vector $\hat{\mathbf{w}}$ is called a *QTLS solution*. It is obvious that $[\mathbf{X}; \mathbf{d}]$ and $[\hat{\mathbf{X}}; \hat{\mathbf{d}}]$ are respectively of rank $(n+1)$ and n . Let the SVDQ of $[\mathbf{X}; \mathbf{d}]$ be given as the LHS of eq. (1) with $r = n+1$. By Theorem 1, for $\hat{\mathbf{w}}$ to be a QTLS solution, $[\hat{\mathbf{X}}; \hat{\mathbf{d}}]$ must be a rank- n SVDQ of $[\mathbf{X}; \mathbf{d}]$. It can be then shown that the QTLS solution, $\hat{\mathbf{w}}$, is

$$\begin{bmatrix} \hat{\mathbf{w}} \\ -1 \end{bmatrix} = -\frac{\mathbf{v}_{n+1} v_{n+1, n+1}^*}{\|v_{n+1, n+1}\|^2} \quad (4)$$

2.3. Quaternion Rayleigh Quotient

The quaternion Rayleigh quotient (QRQ) is defined as follows.

Definition 1. [8, 18] For a given Hermitian matrix $\mathbf{G} \in \mathbb{H}^{n \times n}$ and a nonzero vector $\mathbf{z} \in \mathbb{H}^n$, the quaternion Rayleigh quotient $\mathcal{R}(\mathbf{G}, \mathbf{z})$ is given by

$$\mathcal{R}(\mathbf{G}, \mathbf{z}) = \frac{\mathbf{z}^H \mathbf{G} \mathbf{z}}{\mathbf{z}^H \mathbf{z}} \quad (5)$$

The connection between the QRQ and the QTLS solution is then established through the following theorem.

Theorem 2. [8, 18] Consider an overdetermined set of linear equations $\mathbf{X}\mathbf{w} \approx \mathbf{d}$ where $\hat{\mathbf{w}}$ is the QTLS solution given in eq. (4). Let the SVDQ of $\mathbf{A} = [\mathbf{X}; \mathbf{d}]$ be given as the LHS of eq. (1) with $r = n+1$. Then, for a nonzero vector $\mathbf{z} \in \mathbb{H}^{n+1}$, the solution to $\min_{\mathbf{z}} \mathcal{R}(\mathbf{A}^H \mathbf{A}, \mathbf{z})$ is the QTLS solution, that is,

$$\min_{\mathbf{z}} \mathcal{R}(\mathbf{A}^H \mathbf{A}, \mathbf{z}) = \mathcal{R}\left(\mathbf{A}^H \mathbf{A}, \begin{bmatrix} \hat{\mathbf{w}} \\ -1 \end{bmatrix}\right) = \sigma_{n+1}^2 \quad (6)$$

3. THE WIDELY LINEAR QRTLS

3.1. Model Formulation

Let $d(n), x(n) \in \mathbb{H}$ denote respectively the desired (output) signal and the input signal. The MMSE estimator of $d(n)$ based on QWL model, denoted by $y(n)$, can be expressed as [19]

$$y(n) = \mathbf{w}^T(n-1) \mathbf{q}(n) \quad (7)$$

where the augmented weight vector $\mathbf{w}(n)$ and augmented input vector $\mathbf{q}(n)$ are defined respectively as

$$\begin{aligned} \mathbf{w}(n) &= [\mathbf{u}^T(n) \mathbf{v}^T(n) \mathbf{g}^T(n) \mathbf{h}^T(n)]^T \\ \mathbf{q}(n) &= [\mathbf{x}^T(n) \mathbf{x}^{iT}(n) \mathbf{x}^{jT}(n) \mathbf{x}^{kT}(n)]^T \end{aligned} \quad (8)$$

where $(\cdot)^i, (\cdot)^j, (\cdot)^k$ represent quaternion involutions [2], and $\mathbf{x}(n) = [x(n) x(n-1) \cdots x(n-p+1)]^T$, for a filter of order p . Note that the index of $\mathbf{w}$ is set to $(n-1)$ to aid the algorithmic implementation. Now, eq. (7) can be re-written in a *right* topological form as $y^*(n) = \mathbf{q}^H(n) \mathbf{w}^*(n-1)$. Thus, we can say that $\mathbf{w}^*(n)$ is the QTLS solution to $\mathbf{A} \in \mathbb{H}^{n \times (4p+1)}$, defined as

$$\mathbf{A}(n) = \mathbf{A}^{\frac{1}{2}}(n) \begin{bmatrix} \mathbf{q}^H(1); d^*(1) \\ \mathbf{q}^H(2); d^*(2) \\ \vdots \\ \mathbf{q}^H(n); d^*(n) \end{bmatrix} \quad (9)$$

where p denotes the order of an adaptive filter and $\mathbf{A}(n) = \text{diag}(\lambda^{n-1}, \lambda^{n-2}, \dots, \lambda, 1)$; $\lambda \in [0, 1)$ is the forgetting factor. By Theorem 2, $\mathbf{w}^*(n)$ also minimises the following cost function,

$$\mathcal{J}(n) = \mathcal{R}(\mathbf{G}(n), \mathbf{z}(n)) \quad (10a)$$

$$\mathbf{G}(n) = \mathbf{A}^H(n) \mathbf{A}(n) = \begin{bmatrix} \mathbf{R}_{qq}(n) & \mathbf{r}_{qd}(n) \\ \mathbf{r}_{qd}^H(n) & \sigma_d^2(n) \end{bmatrix} \quad (10b)$$

where $\mathbf{R}_{qq}(n) \in \mathbb{H}^{4p \times 4p}$, $\mathbf{r}_{qd}(n) \in \mathbb{H}^{4p \times 1}$, and $\sigma_d(n) \in \mathbb{R}$ are given by

$$\begin{aligned} \mathbf{R}_{qq}(n) &= \sum_{k=1}^n \lambda^{n-k} \mathbf{q}(k) \mathbf{q}^H(k) \\ &= \lambda \mathbf{R}_{qq}(n-1) + \mathbf{q}(n) \mathbf{q}^H(n) \end{aligned} \quad (11a)$$

$$\begin{aligned}\mathbf{r}_{qd}(n) &= \sum_{k=1}^n \lambda^{n-k} \mathbf{q}(k) d^*(k) \\ &= \lambda \mathbf{r}_{qd}(n-1) + \mathbf{q}(n) d^*(n)\end{aligned}\quad (11b)$$

$$\sigma_d^2(n) = \sum_{k=1}^n \lambda^{n-k} \|d(k)\|^2 = \lambda \sigma_d^2(n-1) + \|d(n)\|^2 \quad (11c)$$

3.2. Derivation of the QRQI-based algorithm

The derivation in this section follows that for the real-valued case in [9]. The aim of the QRTLS algorithm is to minimise the cost function in eq. (10a). This is achieved by employing the quaternion Rayleigh quotient to perform the optimisation in an iterative manner.

Consider the following rank-one update of $\mathbf{w}(n)$ [20]:

$$\mathbf{w}(n) = \mathbf{w}(n-1) + \mu(n) \mathbf{q}^*(n) \quad (12)$$

where $\mu(n) \in \mathbb{H}$ is the *adaptation gain*. Upon substituting eqs. (4) and (10b) into the inverse power formula presented in [21], yields

$$\begin{bmatrix} \mathbf{R}_{qq}(n) & \mathbf{r}_{qd}(n) \\ \mathbf{r}_{qd}^H(n) & \sigma_d^2(n) \end{bmatrix} \begin{bmatrix} \mathbf{w}^*(n) \\ -1 \end{bmatrix} = \begin{bmatrix} \mathbf{w}^*(n-1) \\ -1 \end{bmatrix} \varrho(n) \quad (13)$$

where $\varrho(n)$ is unknown and is defined for convenience. To find the parameters $\mu(n)$ and $\varrho(n)$, define the *augmented gain vector* $\mathbf{k}(n)$ as

$$\mathbf{k}(n) = \mathbf{R}_{qq}(n) \mathbf{q}(n) \quad (14)$$

A fast and robust recursive update for $\mathbf{k}(n)$ can be achieved via the method of shift structure as in [22, 23] and is summarised in Table 1. Now, let

$$\begin{aligned}\rho_0 &= \mathbf{k}^H(n) \mathbf{q}(n) \\ \rho_e &= \mathbf{r}_{qd}^H(n) \mathbf{q}(n) - \mathbf{w}^T(n-1) \mathbf{k}(n)\end{aligned}\quad (15)$$

$$L(n) = \|\mathbf{w}(n)\|^2 + 1 \quad (16)$$

$$\begin{aligned}\mathcal{P}^0 &= [\mathbf{w}^T(n-1); -1] \mathbf{G}(n) \begin{bmatrix} \mathbf{w}^*(n-1) \\ -1 \end{bmatrix} \\ \mathcal{P}(n) &= [\mathbf{w}^T(n); -1] \mathbf{G}(n) \begin{bmatrix} \mathbf{w}^*(n) \\ -1 \end{bmatrix}\end{aligned}\quad (17)$$

By virtue of eqs. (11), (12) and (15), $\mathcal{P}^0$ and $\mathcal{P}(n)$ in eq. (17) can be represented as coupled iterative formulae, that is,

$$\begin{aligned}\mathcal{P}^0 &= \lambda \mathcal{P}(n-1) + \|e(n)\|^2 \\ \mathcal{P}(n) &= \mathcal{P}^0 - 2\Re\{\rho_e \mu(n)\} + \rho_0 \mu(n)^2\end{aligned}\quad (18)$$

where $e(n) = d(n) - y(n)$ and $y(n)$ is as given in eq. (7). Substituting eq. (12) into eq. (13) and pre-multiplying eq. (13) by $[\mathbf{q}^H(n); 0]$ and $[\mathbf{w}^T(n-1); -1]$ yields

$$\begin{aligned}\rho_0 \mu^*(n) - y^*(n) \varrho(n) &= \rho_e^* \\ \rho_e \mu^*(n) + L(n-1) \varrho(n) &= \mathcal{P}^0\end{aligned}\quad (19)$$

which are linear equations of two unknowns. Finally, solving eq. (20) yields

$$\mu(n) = \frac{\Psi^* \Delta}{\|\Delta\|^2} \quad (20)$$

where

$$\Psi = \mathcal{P}^0 \tilde{y} + \rho_e^* \quad \Delta = \rho_0 + \tilde{y} \rho_e \quad (21a)$$

$$\tilde{y} = \frac{y^*(n)}{L(n-1)} \quad (21b)$$

4. EXPERIMENTAL RESULTS

For a fair comparison of three algorithms, the forgetting factor λ of both WL-QRLS and WL-QRTLS, and the step size of WL-QLMS were chosen such that all routines converged approximately at the same rate for the signal at hand. Also, in each experiment, arbitrary $\mathbb{Q}$ -improper white Gaussian noise (WGN) was added to the experimental signal as a perturbation. For simplification, the perturbations added to input and

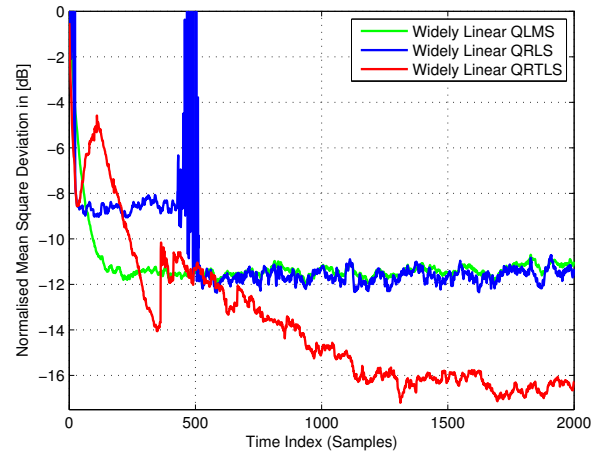


Fig. 1: Learning curves of WL-QLMS, WL-QRLS and WL-QRTLS, averaged over 100 realisations, for the identification of the QMA(2) process (eq. (22)) under perturbations.

Figure 1 compares the ability of the three algorithms to identify a $\mathbb{Q}$ -improper MA(2) process given by

$$\begin{aligned}y(n) &= (2.1 + 2j)\tilde{x}(n) + (1.4i + 0.8k)\tilde{x}(n-1) \\ &\quad + (2.9 + 0.4j + 1.6k)\tilde{x}^i(n) + (1.9i)\tilde{x}^i(n-1) \\ &\quad + (0.7 + 1.8i)\tilde{x}^j(n) + (0.9j + 3k)\tilde{x}^j(n-1) \\ &\quad + (2.5 + 2.8k)\tilde{x}^k(n) + (2.6i + 0.7j)\tilde{x}^k(n-1)\end{aligned}\quad (22)$$

where $\tilde{x}(n)$ is arbitrary $\mathbb{Q}$ -improper WGN. The $d(n)$ and $x(n)$ in eqs. (7) to (9), were obtained by adding WGN perturbation to $y(n)$ and $\tilde{x}(n)$, respectively. It is clear from the figure that the WL-QRTLS, though erratically fluctuating in the transient state, finally reached the lowest level of steady-state error, compared to the other two.

1. $\mathbf{k}_r(n)$: $\mathbf{r}_1$, $\mathbf{r}_p$ and ς_r initialised to nought	2. $\mathbf{k}_p(n)$: $\mathbf{p}_1$, $\mathbf{p}_p$ and ς_p initialised to nought
$\mathbf{r}_1(n) = \lambda \mathbf{r}_1(n-1) + \mathbf{x}(n-1)x^*(n)$ $\mathbf{r}_p(n) = \lambda \mathbf{r}_p(n-1) + \mathbf{x}(n)x^*(n-p)$ $\varsigma_r(n) = \lambda \varsigma_r(n-1) + x(n)x^*(n)$ $\mathbf{k}_r(n) = \left\{ \begin{bmatrix} \mathbf{r}_1^H(n)\mathbf{x}(n-1) + \varsigma_r(n)x(n) \\ \mathbf{k}_r(n-1) + \mathbf{r}_1(n)x(n) \end{bmatrix} \right\}_{1:p} - \mathbf{r}_p(n)x(n-p)$	$\mathbf{p}_1(n) = \lambda \mathbf{p}_1(n-1) + \mathbf{x}^i(n-1)x^*(n)$ $\mathbf{p}_p(n) = \lambda \mathbf{p}_p(n-1) + \mathbf{x}^i(n)x^*(n-p)$ $\varsigma_p(n) = \lambda \varsigma_p(n-1) + x^i(n)x^*(n)$ $\mathbf{k}_p(n) = \left\{ \begin{bmatrix} \mathbf{p}_1^{iH}(n)\mathbf{x}(n-1) + \varsigma_p(n)x(n) \\ \mathbf{k}_p(n-1) + \mathbf{p}_1(n)x(n) \end{bmatrix} \right\}_{1:p} - \mathbf{p}_p(n)x(n-p)$
3. $\mathbf{k}_s(n)$: $\mathbf{s}_1$, $\mathbf{s}_p$ and ς_s initialised to nought	4. $\mathbf{k}_t(n)$: $\mathbf{t}_1$, $\mathbf{t}_p$ and ς_t initialised to nought
$\mathbf{s}_1(n) = \lambda \mathbf{s}_1(n-1) + \mathbf{x}^j(n-1)x^*(n)$ $\mathbf{s}_p(n) = \lambda \mathbf{s}_p(n-1) + \mathbf{x}^j(n)x^*(n-p)$ $\varsigma_s(n) = \lambda \varsigma_s(n-1) + x^j(n)x^*(n)$ $\mathbf{k}_s(n) = \left\{ \begin{bmatrix} \mathbf{s}_1^{jH}(n)\mathbf{x}(n-1) + \varsigma_s(n)x(n) \\ \mathbf{k}_s(n-1) + \mathbf{s}_1(n)x(n) \end{bmatrix} \right\}_{1:p} - \mathbf{s}_p(n)x(n-p)$	$\mathbf{t}_1(n) = \lambda \mathbf{t}_1(n-1) + \mathbf{x}^k(n-1)x^*(n)$ $\mathbf{t}_p(n) = \lambda \mathbf{t}_p(n-1) + \mathbf{x}^k(n)x^*(n-p)$ $\varsigma_t(n) = \lambda \varsigma_t(n-1) + x^k(n)x^*(n)$ $\mathbf{k}_t(n) = \left\{ \begin{bmatrix} \mathbf{t}_1^{kH}(n)\mathbf{x}(n-1) + \varsigma_t(n)x(n) \\ \mathbf{k}_t(n-1) + \mathbf{t}_1(n)x(n) \end{bmatrix} \right\}_{1:p} - \mathbf{t}_p(n)x(n-p)$
5. $\mathbf{k}_u$, $\mathbf{k}_v$, $\mathbf{k}_g$ and $\mathbf{k}_h$	
$\mathbf{k}_u = \mathbf{k}_r(n) + \mathbf{k}_p^i(n) + \mathbf{k}_s^j(n) + \mathbf{k}_t^k(n)$ $\mathbf{k}_v = \mathbf{k}_p(n) + \mathbf{k}_r^i(n) + \mathbf{k}_t^j(n) + \mathbf{k}_s^k(n)$ $\mathbf{k}_g = \mathbf{k}_s(n) + \mathbf{k}_t^i(n) + \mathbf{k}_r^j(n) + \mathbf{k}_p^k(n)$ $\mathbf{k}_h = \mathbf{k}_t(n) + \mathbf{k}_s^i(n) + \mathbf{k}_p^j(n) + \mathbf{k}_r^k(n)$	
6. $\mathbf{k}(n)$	
$\mathbf{k}(n) = [\mathbf{k}_u^T \quad \mathbf{k}_v^T \quad \mathbf{k}_g^T \quad \mathbf{k}_h^T]^T$	

Table 1: Fast recursion of augmented gain vector $\mathbf{k}(n)$

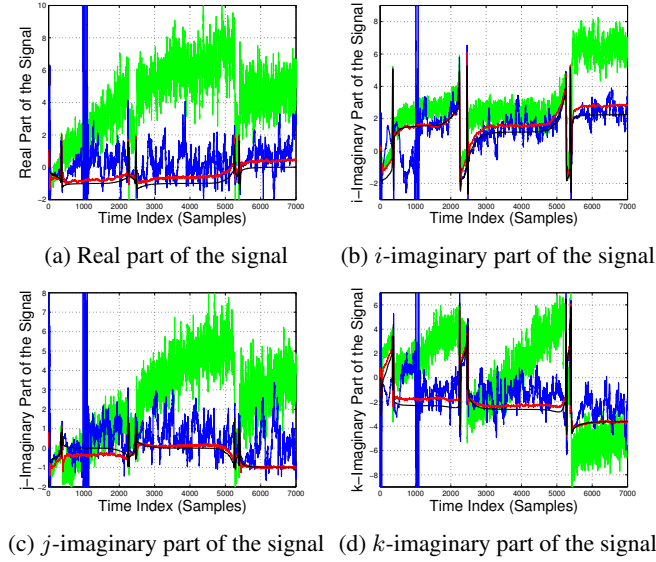


Fig. 2: Estimated output of WL-QLMS (green), WL-QRLS (blue) and WL-QRTLs (red) compared with the actual Saito's $\mathbb{Q}$ -improper chaotic time series (black), under perturbations.

Figure 2 illustrates the prediction of Saito's chaotic time series [24] by the three algorithms, compared with the actual signal. This was achieved under noisy conditions, and the WL-QRTLs exhibited superior performance and could more closely estimate the actual signal with little residual variance.

Figure 3 compares the learning curves of the three algorithms for the prediction of a $\mathbb{Q}$ -improper wind field¹. The WL-QRTLs was again a consistent estimator while the other two gave highly varying estimates. The WL-QRTLs also exhibited additional numerical robustness over the WL-QRLS which sometimes produced over-shoots.

¹<http://www.commsp.ee.ic.ac.uk/~mandic/>

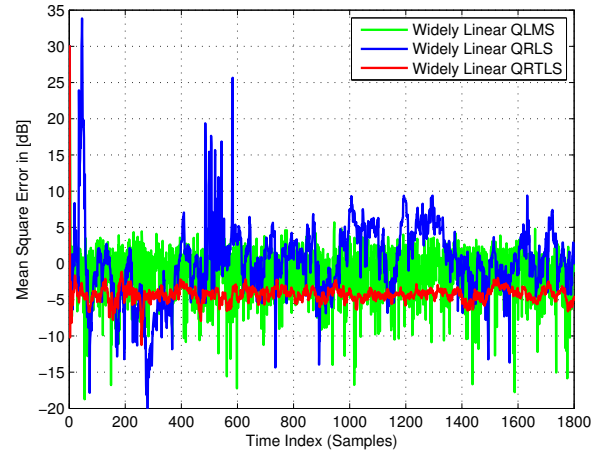


Fig. 3: Learning curves of WL-QLMS, WL-QRLS and WL-QRTLs for the prediction of 4-D wind field (3-D wind velocity and air temperature) [25] under perturbations due to additive $\mathbb{Q}$ -improper WGN.

5. CONCLUSION

We have introduced the widely linear QRTLs algorithm for the processing of $\mathbb{Q}$ -improper processes where both input and output are contaminated by noise. This has been achieved by extending the concept of total least squares to the quaternion domain, and then finalising the recursive solution via quaternion Rayleigh quotient iteration. Note that the WL-QRTLs is computationally faster than the WL-QRLS, thanks to no explicit matrix multiplication involved. The WL-QRTLs has been found to have superior steady state performance, a significant increase in noise tolerance, and quite fast computation, as compared to other algorithms of the same class (WL-QLMS and WL-QRLS). Simulations on benchmark $\mathbb{Q}$ -improper signals, and 4-D wind field supported the analysis.

6. REFERENCES

- [1] C. Took and D. Mandic, "Fusion of heterogeneous data sources: A quaternionic approach," in *Proc. of the MLSP, IEEE Workshop on Machine Learning for Signal Processing*, 2008, pp. 456–461.
- [2] —, "Augmented second-order statistics of quaternion random signals," *Signal Processing*, vol. 91, no. 2, pp. 214–224, 2011.
- [3] J. Via, D. Palomar, L. Vielva, and I. Santamaria, "Quaternion ICA from second-order statistics," *IEEE Transactions on Signal Processing*, vol. 59, no. 4, pp. 1586–1600, April 2011.
- [4] C. Took and D. Mandic, "The quaternion LMS algorithm for adaptive filtering of hypercomplex processes," *IEEE Transactions on Signal Processing*, vol. 57, no. 4, pp. 1316–1327, Apr 2009.
- [5] —, "A quaternion widely linear adaptive filter," *IEEE Transactions on Signal Processing*, vol. 58, no. 8, pp. 4427–4431, Aug 2010.
- [6] C. Jahanchahi, C. Took, and D. Mandic, "The widely linear quaternion recursive least squares filter," in *Proc. of the 2nd International Workshop on Cognitive Information Processing (CIP)*, Jun 2010, pp. 87–92.
- [7] J. Navarro-Moreno, J. C. Ruiz-Molina, A. Oya, and J. M. Quesada-Rubio, "Detection of continuous-time quaternion signals in additive noise," *EURASIP Journal on Advances in Signal Processing*, vol. 2012, p. 234, 2012.
- [8] S. Van Huffel and J. Vandewalle, *The Total Least Squares Problem: Computational Aspects and Analysis*, ser. Frontiers in applied mathematics. Philadelphia: Society for Industrial and Applied Mathematics, 1991, vol. 9.
- [9] D.-Z. Feng and W. X. Zheng, "Fast approximate inverse power iteration algorithm for adaptive total least-squares FIR filtering," *IEEE Transactions on Signal Processing*, vol. 54, no. 10, pp. 4032–4039, Oct 2006.
- [10] F. Zhang, "Quaternions and matrices of quaternions," *Linear Algebra and its Applications*, vol. 251, pp. 21–57, 1997.
- [11] 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.
- [12] P. Garret, "Algebras and involutions," Sch. of Math., Univ. of Minnesota, Tech. Rep., Feb. 2005.
- [13] G. Ehrlich, *Fundamental Concepts of Abstract Algebra*. Dover Publications, Feb 2012.
- [14] F. Tobar and D. Mandic, "Quaternion reproducing kernel hilbert spaces: Existence and uniqueness conditions," *IEEE Transactions on Information Theory*, vol. 60, no. 9, pp. 5736–5749, Sept 2014.
- [15] A. Baker, "Right eigenvalues for quaternionic matrices: A topological approach," *Linear Algebra and its Applications*, vol. 286, no. 13, pp. 303–309, 1999.
- [16] D. Farenick and B. Pidkowich, "The spectral theorem in quaternions," *Linear Algebra and its Applications*, vol. 371, pp. 75–102, 2003.
- [17] G. Stewart, *Matrix Algorithms: Volume 1, Basic Decompositions*, 1st ed. Society for Industrial and Applied Mathematics, 1998.
- [18] T. Thanthawarittisai, "An approach to hypercomplex adaptive filters and their real-world applications," M.Sc. thesis, Imperial College London, London, UK, Sep 2014.
- [19] J. Via, D. Ramirez, and I. Santamaria, "Properness and widely linear processing of quaternion random vectors," *IEEE Transactions on Information Theory*, vol. 56, no. 7, pp. 3502–3515, July 2010.
- [20] E. Chong and S. Zak, *An Introduction to Optimization*. John Wiley and Sons, 2013.
- [21] G. Golub and C. Van Loan, *Matrix Computations*. Baltimore, MD-US: Johns Hopkins University Press, 2013.
- [22] L. Ljung, M. Morf, and D. Falconer, "Fast calculation of gain matrices for recursive estimation schemes," *International Journal of Control*, vol. 27, no. 1, pp. 1–19, 1978.
- [23] C. Davila, "Line search algorithms for adaptive filtering," *IEEE Transactions on Signal Processing*, vol. 41, no. 7, pp. 2490–2494, Jul 1993.
- [24] P. Arena, L. Fortuna, G. Muscato, and M. Xibilia, *Neural Networks in Multidimensional Domains*. New York: Springer, 1998.
- [25] C. 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, no. 6, pp. 1754 – 1760, 2011.