THE WIDELY LINEAR QUATERNION RECURSIVE TOTAL LEAST SQUARES

Thiannithi Thanthawaritthisai1, 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 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 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 channelwise 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 Q-proper and 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 (SVDQ) 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 Dop 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 \( A_p \in \mathbb{H}^{m \times n} \) with rank \( p < r \) be the \( r \)-rank SVDQ of \( A \in \mathbb{H}^{m \times n} \) with rank \( r \leq \min(m, n) \). The SVDQs of \( A \) and \( A_p \) are respectively given by

\[
A = \sum_{i=1}^{r} \sigma_i u_i v_i^H \quad \text{and} \quad A_p = \sum_{i=1}^{p} \sigma_i u_i v_i^H. \tag{1}\]

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

\[
\min_B \| A - B \|_F = \| A - A_p \|_F = \sqrt{\sum_{i=p+1}^{r} \sigma_i^2} \tag{2}\]

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 \( Xw \approx d \), where \( X \in \mathbb{H}^{m \times n} \) and \( d \in \mathbb{H}^{m \times 1} \) are the given data and the coefficients \( w \in \mathbb{H}^{n \times 1} \) are unknown. Now, consider the correction \( X\hat{w} = d \). The QTLS problem seeks to solve the system

\[
\min_{[X; d]} \| [X; d] - [X; d] \|_F \quad \text{subject to} \quad \hat{d} \in R(X) \tag{3}\]

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

\[
\begin{bmatrix} \hat{w} \\ -1 \end{bmatrix} = -\frac{v_{n+1}^* v_{n+1,n+1}^*}{\|v_{n+1,n+1}\|^2} \tag{4}\]

2.3. Quaternion Rayleigh Quotient

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

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

\[
R(G, z) = \frac{z^H G z}{z^H z} \tag{5}\]

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 \( Xw \approx d \) where \( \hat{w} \) is the QTLS solution given in eq. (4). Let the SVDQ of \( A = [X; d] \) be given as the LHS of eq. (1) with \( r = n + 1 \). Then, for a nonzero vector \( z \in \mathbb{H}^{n+1} \), the solution to min \( R(A^H A, z) \) is the QTLS solution, that is,

\[
\min_z R(A^H A, z) = R(A^H A, \begin{bmatrix} \hat{w} \\ -1 \end{bmatrix}) = \sigma_{n+1}^2 \tag{6}\]

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]

\[
y(n) = w^T(n - 1) q(n) \tag{7}\]

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

\[
w(n) = [u^T(n) \ v^T(n) \ g^T(n) \ h^T(n)]^T \tag{8}\]

\[
q(n) = [x^T(n) \ x^{1T}(n) \ x^{2T}(n) \ x^{3T}(n)]^T \tag{9}\]

where \((\cdot)^*, (\cdot)^{1}, (\cdot)^{2}\) represent quaternion involutions [2], and \( x(n) = [x(n) x(n-1) \cdots x(n-p+1)]^T \), for a filter of order \( p \). Note that the index of \( w \) is set to \((n-1)\) to aid the algorithmic implementation. Now, eq. (7) can be re-written in a right-topological form as \( y^+(n) = q^H(n) w^+(n-1) \).

Thus, we can say that \( w^+(n) \) is the QTLS solution to \( A \in \mathbb{H}^{n \times (4p+1)} \), defined as

\[
A(n) = A^\frac{1}{2}(n) \begin{bmatrix} q^H(1); d^+(1) \\ q^H(2); d^+(2) \\ \vdots \\ q^H(n); d^+(n) \end{bmatrix} \tag{10}\]

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

\[
\mathcal{J}(n) = R\left( G(n), z(n) \right) \tag{11a}\]

where \( G(n) = A^H(n) A(n) = \begin{bmatrix} R_{qq}(n) & r_{qd}(n) \\ r_{dq}(n) & \sigma_d^2(n) \end{bmatrix} \tag{11b}\)

are given by

\[
R_{qq}(n) = \sum_{k=1}^{n} \lambda^{n-k} q(k) q^H(k) \tag{11a}\]

\[
= \lambda R_{qq}(n-1) + q(n) q^H(n) \tag{11b}\]
\[
\begin{align*}
    r_{qd}(n) &= \sum_{k=1}^{n} \lambda^{n-k} q(k)d^*(k) \\
    &= \lambda r_{qd}(n-1) + q(n)d^*(n) \\
    \sigma_q^2(n) &= \sum_{k=1}^{n} \lambda^{n-k} \|d(k)\|^2 = \lambda \sigma_q^2(n-1) + \|d(n)\|^2
  \end{align*}
\] (11b)
\[\sigma_q^2(n) = \sum_{k=1}^{n} \lambda^{n-k} \|d(k)\|^2 = \lambda \sigma_q^2(n-1) + \|d(n)\|^2 \] (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 QRQI 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 \( w(n) \) [20]:
\[
w(n) = w(n-1) + \mu(n)q^*(n) \] (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
\[
\begin{bmatrix}
  \mathbf{R}_{qq}(n) \\
  \mathbf{r}_{qd}(n) \\
  \sigma_q^2(n)
\end{bmatrix}
\begin{bmatrix}
  w^*(n) \\
  -1
\end{bmatrix} = \begin{bmatrix}
  w^*(n-1) \\
  -1
\end{bmatrix} \varrho(n) \] (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 \( k(n) \) as
\[
k(n) = \mathbf{R}_{qq}(n)q(n) \] (14)
A fast and robust recursive update for \( k(n) \) can be achieved via the method of shift structure as in [22, 23] and is summarised in Table 1. Now, let
\[
\begin{align*}
    \rho_0 &= k^H(n)q(n) \\
    \rho_e &= r_{qd}(n)q(n) - w^T(n-1)k(n) \\
    L(n) &= \|w(n)\|^2 + 1 \\
    \mathcal{P}^0 &= [w^T(n-1); -1]G(n) \begin{bmatrix}
  w^*(n-1) \\
  -1
\end{bmatrix} \\
    \mathcal{P}(n) &= [w^T(n); -1]G(n) \begin{bmatrix}
  w^*(n) \\
  -1
\end{bmatrix}
\end{align*}
\] (15)
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,
\[
\begin{align*}
    \mathcal{P}^0 &= \lambda \mathcal{P}(n-1) + \|e(n)\|^2 \\
    \mathcal{P}(n) &= \mathcal{P}^0 - 2\rho \{\rho_0 \mu(n)\} + \rho_0 \mu(n)^2
\end{align*}
\] (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 \( [q^H(n); 0] \) and \( [w^T(n-1); -1] \) yields
\[
\begin{align*}
    \rho_0 \mu^*(n) - y^*(n) \varrho(n) &= \rho_e^* \\
    \rho_e \mu^*(n) + L(n-1) \varrho(n) &= \mathcal{P}^0
\end{align*}
\] (19)
which are linear equations of two unknowns. Finally, solving eq. (20) yields
\[
\mu(n) = \frac{\Psi^* \Delta}{\|\Delta\|^2}
\] (20)
where
\[
\begin{align*}
    \Psi &= \mathcal{P}^0 \tilde{y} + \rho_e^* \\
    \Delta &= \rho_0 + \tilde{y} \rho_e \\
    \tilde{y} &= \frac{y^*(n)}{L(n-1)}
\end{align*}
\] (21a)
4. EXPERIMENTAL RESULTS

For a fair comparison of three algorithms, the forgetting factor \( \lambda \) 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 output signals in each experiment were arbitrary
\[
\begin{align*}
    d(n) &= \{2.1 + 2j\} \tilde{x}(n) + \{1.4i + 0.8k\} \tilde{x}(n-1) \\
    &+ \{2.9 + 0.4j + 1.6k\} \tilde{x}^2(n) + \{1.9i\} \tilde{x}^2(n-1) \\
    &+ \{0.7 + 1.8i\} \tilde{x}^3(n) + \{0.9j + 3k\} \tilde{x}^3(n-1) \\
    &+ \{2.5 + 2.8k\} \tilde{x}^4(n) + \{2.6i + 0.7j\} \tilde{x}^4(n-1)
\end{align*}
\] (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.

Figure 1 compares the ability of the three algorithms to identify a \( \mathbb{Q} \)-improper MA(2) process given by
\[
\begin{align*}
    y(n) &= \{2.1 + 2j\} \tilde{x}(n) + \{1.4i + 0.8k\} \tilde{x}(n-1) \\
    &+ \{2.9 + 0.4j + 1.6k\} \tilde{x}^2(n) + \{1.9i\} \tilde{x}^2(n-1) \\
    &+ \{0.7 + 1.8i\} \tilde{x}^3(n) + \{0.9j + 3k\} \tilde{x}^3(n-1) \\
    &+ \{2.5 + 2.8k\} \tilde{x}^4(n) + \{2.6i + 0.7j\} \tilde{x}^4(n-1)
\end{align*}
\] (22)
2. $k_r(n)$: $r_1$, $r_p$ and $ς_r$ initialised to nought

$r_1(n) = \lambda r_1(n-1) + x(n-1)x^*(n)$
$r_p(n) = r_p(n-1) + x(n)x^*(n-p)$$ς_r(n) = \lambda ς_r(n-1) + x(n)x^*(n)$
$k_r(n) = \left\{ \left[ r_1^H(n)x(n-1) + ς_r(n)x(n) \right] \right\}_{1:p} - r_p(n)x(n-p)$

3. $s_1$, $s_p$ and $ς_1$ initialised to nought

$s_1(n) = \lambda s_1(n-1) + x^*(n-1)x(n)$
$s_p(n) = s_p(n-1) + x^*(n)x^*(n-p)$
$ς_1(n) = \lambda ς_1(n-1) + x^*(n)x^*(n)$
$k_s(n) = \left\{ \left[ s_1^H(n)x(n-1) + ς_1(n)x(n) \right] \right\}_{1:p} - s_p(n)x(n-p)$

5. $k_u$, $k_v$, $k_p$ and $k_h$

$k_u = k_r(n) + k_v^T(n) + k_i^T(n) + k_j^T(n)$
$k_v = k_p(n) + k_i^T(n) + k_j^T(n) + k_h^T(n)$
$k_p = k_i(n) + k_i^T(n) + k_j^T(n) + k_h^T(n)$
$k_h = k(n) + k_i^T(n) + k_j^T(n) + k_h^T(n)$

$k(n) = [k_u^T, k_v^T, k_p^T, k_h^T]^T$

Table 1: Fast recursion of augmented gain vector $k(n)$

---

Figure 2 illustrates the prediction of Saito’s chaotic time series [24] by the three algorithms, compared with the actual Saito’s Q-improper chaotic time series (black), under perturbations.

![Real part of the signal](a)  
![i-imaginary part of the signal](b)  
![j-imaginary part of the signal](c)  
![k-imaginary part of the signal](d)  

Fig. 2: Estimated output of WL-QLMS (green), WL-QRLS (blue) and WI-QRTLS (red) compared with the actual Saito’s Q-improper chaotic time series (black), under perturbations.

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

![Mean Square Error in [dB]](a)  
![Mean Square Error in [dB]](b)  

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

5. CONCLUSION

We have introduced the widely linear QRTLS algorithm for the processing of 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 Q-improper signals, and 4-D wind field supported the analysis.

---

1http://www.commsp.ee.ic.ac.uk/~mandic/
6. REFERENCES


