A Theoretical Bound for the Performance Advantage of Quaternion Widely Linear Estimation

Cyrus Jahanchahi, Sithan Kanna*, Danilo Mandic
Communications and Signal Processing Group
Department of Electrical and Electronic Engineering
Imperial College London
London SW7 2AZ, United Kingdom
*Email: ssk08@ic.ac.uk

***Abstract*—The quaternion widely linear model was recently introduced for optimal second order estimation of noncircular 3D and 4D data. Its superiority over the standard strictly linear model was shown experimentally, however, a rigorous proof giving performance bounds has been lacking. To this end, we here present a mathematical proof for the degree of performance benefits obtained when using the widely linear model in the context of minimum mean square error estimation.**

***Index Terms*—Quaternion widely linear model, QLMS, WLQLMS, mean square error, quaternion noncircularity.**

I. Introduction

Standard real-valued algorithms used in multichannel statistical signal processing do cater for the 'coupled' nature of the available information across data channels, but the information in the correlation matrices in not easy to use, due to their multiple, and scattered block structures. In practice, we are largely dealing with three- and four-dimensional signals; when modelled using real vectors these produce e.g. 10 correlation matrices for 4D data. Also, accuracy may be compromised due to the deficiencies of the non-division vector algebra (gimbal lock). These problem, are largely mitigated when modeling three- and four-dimensional signals in the quaternion domain, since its division algebra naturally accounts for the coupling between the data channels and also provides parsimonious representation, e.g. with only four covariance matrices needed. For this reason, quaternion valued algorithms are rapidly gaining popularity in applications involving vector sensors (wind modeling, inertial body sensors, array processing [1].

The majority of minimum mean square error (MSE) estimators developed thus far in the quaternion domain are based on the so-called strictly linear model

$$\hat{y} = E[y|\mathbf{x}] = \mathbf{w}^H\mathbf{x} \tag{1}$$

where $\mathbf{x}$ is the observation vector and $\mathbf{w}$ coefficients. The minimum MSE solution for the optimum weight vector $\mathbf{w}$ is well known and takes the same form as the complex valued Wiener filter. Although optimal for second order circular signals that have probability density functions invariant to rotation (like in the complex domain [2]), to capture the complete second order statistics of noncircular signals it is not sufficient to consider only the covariance matrix (as is the case with the Wiener solution). Indeed, recent advances in quaternion statistics have shown that for noncircular signals, three pseudo-covariance matrices must also be employed and incorporated into minimum mean square error estimation solutions. This is achieved through the widely linear quaternion model [3] , given by

$$\hat{y} = \mathbf{u}^H\mathbf{x} + \mathbf{v}^H\mathbf{x}^{\imath} + \mathbf{g}^H\mathbf{x}^{\jmath} + \mathbf{h}^H\mathbf{x}^{\kappa} \tag{2}$$

Recently, several articles have demonstrated either implicitly or experimentally the superiority of the widely linear algorithms over the standard strictly linear ones, in modeling real world data [3] [4] . However, a mathematical proof for the performance bound that would facilitate a more widespread use of the widely linear model for quaternion estimators is still lacking. Using Picinbono's original proof for the advantage of complex widely linear model [5] as a basis for the analysis, we provide a formal proof for the benefits obtained in using the quaternion widely linear model.

II. Quaternion Algebra

Quaternions are an associative but noncommutative algebra over $\mathbb{R}$, defined as

$$\mathbb{H} = \{q_a + \imath q_b + \jmath q_c + \kappa q_d \,|\, q_a, q_b, q_c, q_d \in \mathbb{R}\}$$

where the imaginary units $\imath$, $\jmath$ and κ are also unit axis vectors, for which $\imath^2 = \jmath^2 = \kappa^2 = \imath\jmath\kappa = -1$. For any quaternion

$$q = q_a + \imath q_b + \jmath q_c + \kappa q_d = Sq + Vq \tag{3}$$

the scalar (real) part is denoted by $Sq = \Re(q)$, whereas the vector part (also called pure quaternion) $Vq = \Im(q) = \imath q_b + \jmath q_c + \kappa q_d$ comprises the three imaginary parts. The quaternion product is given by

$$q_1q_2 = Sq_1Sq_2 - Vq_1 \cdot Vq_2 + Sq_2Vq_1 + Sq_1Vq_2 + Vq_1 \times Vq_2 \tag{4}$$

where the symbol '$\cdot$' denotes the scalar product and '$\times$' the vector product. Due to the vector product in (4), the quaternion product is non-commutative, that is, $q_1q_2 \neq q_2q_1$ and e.g. $\imath\jmath = -\jmath\imath = \kappa$, $\jmath\kappa = -\kappa\jmath = \imath$, $\kappa\imath = -\imath\kappa = \jmath$. The scalar product $q_1 \cdot q_2 =< q_1, q_2 >, q_1, q_2 \in \mathbb{H}$ is defined as

$$q_1 \cdot q_2 = q_{1a}q_{2a} + q_{1b}q_{2b} + q_{1c}q_{2c} + q_{1d}q_{2d} = \Re\big(q_1q_2^*\big) = \Re\big(q_1^*q_2\big)$$

The quaternion conjugate is given by $q^* = Sq - Vq$, and the norm by $\| q \| = \sqrt{<q,q>} = \sqrt{qq^*}$, and thus $q^{-1} = \frac{q^*}{qq^*} = \frac{q^*}{\|q\|^2} = \frac{q^*}{<q,q>}$, and $\imath^{-1} = -\imath, \jmath^{-1} = -\jmath, \kappa^{-1} = -\kappa$.

A. Equivalence Relations and Involutions

The similarity relation, denoted by '$\sim$', between quaternions q_1 and q_2 implies that $\exists \mu \in \mathbb{H}$, such that

$$q_1 \sim q_2 \quad \Leftrightarrow \quad q_1 = \mu q_2 \mu^{-1}, \quad \mu \neq 0 \tag{5}$$

Similarity is an equivalence relation, and e.g. $q_1 \sim q_2$ implies $\| q_1 \| = \| q_2 \|$ and $q_{1a} = q_{2a}$. Also, the three imaginary units are similar, that is, $\imath \sim \jmath \sim \kappa$, and also $q \sim q^*$.

Equivalence relations of importance to this work are the quaternion involutions (self-inverse mappings) [6]

$$\begin{aligned} q^\imath &= -\imath q \imath = q_a + \imath q_b - \jmath q_c - \kappa q_d \\ q^\jmath &= -\jmath q \jmath = q_a - \imath q_b + \jmath q_c - \kappa q_d \\ q^\kappa &= -\kappa q \kappa = q_a - \imath q_b - \jmath q_c + \kappa q_d \end{aligned} \tag{6}$$

Notice that the quaternion conjugate is also an involution[1], that is, $(q^*)^* = q$. The four real components of a quaternion q can now be expressed based on its involutions as

$$\begin{aligned} q_a &= \frac{1}{4}[q + q^\imath + q^\jmath + q^\kappa] \qquad q_c = \frac{1}{4\jmath}[q - q^\imath + q^\jmath - q^\kappa] \\ q_b &= \frac{1}{4\imath}[q + q^\imath - q^\jmath - q^\kappa] \qquad q_d = \frac{1}{4\kappa}[q - q^\imath - q^\jmath + q^\kappa] \end{aligned} \tag{7}$$

allowing any (either quadrivariate or quaternion-valued) function $g(q_a, q_b, q_c, q_d)$ of the four real variables q_a, q_b, q_c, q_d to be expressed as a function of the quaternion variable q and its perpendicular involutions.

III. The Quaternion Widely Linear Model

Consider a real valued mean square error (MSE) estimator

$$\hat{y} = E[y \mid \mathbf{x}]$$

where $\hat{y}$ is the estimated process and $\mathbf{x}$ the observed variable (regressor). For jointly Gaussian $\mathbf{x}$ and y, the optimal solution is a linear estimator, given by

$$\hat{y} = \mathbf{w}^T \mathbf{x} \tag{8}$$

where $\mathbf{w}$ is the coefficient vector. For the standard, strictly linear, complex domain MSE estimator it is also assumed that $\hat{y} = E[y \mid \mathbf{x}]$, leading to the so called strictly linear model

$$\hat{y} = \mathbf{w}^H \mathbf{x} \tag{9}$$

However, observe that

$$\hat{y}_r = E[y_r \mid \mathbf{x}_r, \mathbf{x}_\imath] \qquad \hat{y}_\imath = E[y_\imath \mid \mathbf{x}_r, \mathbf{x}_\imath]$$

and since $\mathbf{x}_r = \frac{\mathbf{x}+\mathbf{x}^*}{2}$ and $\mathbf{x}_i = \frac{\mathbf{x}-\mathbf{x}^*}{2\imath}$, the complex *widely linear* model is given by [7] [8] [9]

$$\hat{y} = E[y \mid \mathbf{x}, \mathbf{x}^*] \qquad \Rightarrow \qquad \hat{y} = \mathbf{h}^H \mathbf{x} + \mathbf{g}^H \mathbf{x}^*$$

that is, it comprises both the 'strictly linear' part $\mathbf{h}^H \mathbf{x}$ and the 'conjugate' part $\mathbf{g}^H \mathbf{x}^*$, where $\mathbf{g}$ is a coefficient vector. Similarly, the existing (strictly linear) estimation model in the quaternion domain is given by

$$\hat{y} = \mathbf{w}^H \mathbf{x} \tag{10}$$

However, the quaternion components can be expressed as[2]

$$\hat{y}_\eta = E[y_\eta \mid \mathbf{x}_r, \mathbf{x}_\imath, \mathbf{x}_\jmath, \mathbf{x}_\kappa] \qquad \eta \in \{r, \imath, \jmath, \kappa\}$$

and using (7), these are readily re-written via the involutions, e.g. $\mathbf{x}_r = \frac{1}{4}(\mathbf{x} + \mathbf{x}^\imath + \mathbf{x}^\jmath + \mathbf{x}^\kappa)$, leading to the estimator

$$\hat{y_\eta} = E[y_\eta \mid \mathbf{x}, \mathbf{x}^\imath, \mathbf{x}^\jmath, \mathbf{x}^\kappa] \text{ and } \hat{y} = E[y \mid \mathbf{x}, \mathbf{x}^\imath, \mathbf{x}^\jmath, \mathbf{x}^\kappa]$$

Therefore, since every quaternion component is a function of the involutions, to capture the full second order information available we require the quaternion *widely linear* model

$$\hat{y} = \mathbf{u}^H \mathbf{x} + \mathbf{v}^H \mathbf{x}^\imath + \mathbf{g}^H \mathbf{x}^\jmath + \mathbf{h}^H \mathbf{x}^\kappa = \mathbf{w}^{aH} \mathbf{x}^a \tag{11}$$

where $\mathbf{w}^a = [\mathbf{u}^T, \mathbf{v}^T, \mathbf{g}^T, \mathbf{h}^T]^T$ is the augmented coefficient vector and the augmented regressor vector $\mathbf{x}^a = [\mathbf{x}^T, \mathbf{x}^{\imath T}, \mathbf{x}^{\jmath T}, \mathbf{x}^{\kappa T}]^T$. It should be mentioned that an alternative widely linear model could be equally obtained using a conjugate augmented regressor vector $\mathbf{x}^a = [\mathbf{x}^H, \mathbf{x}^{\imath H}, \mathbf{x}^{\jmath H}, \mathbf{x}^{\kappa H}]^T$, for more detail see [3].

Current statistical signal processing in $\mathbb{H}$ is largely based on strictly linear models, drawing upon the covariance matrix $\mathbf{R}_x = E[\mathbf{x}\mathbf{x}^H]$. However, based on 11 the modelling of both the second order circular (proper) and noncircular (improper) signals, is only possible using the augmented covariance matrix, given by [3]

$$\mathbf{R}_x^a = E[\mathbf{x}^a \mathbf{x}^{aH}] = \begin{bmatrix} \mathbf{R}_x & \mathbf{P}_x & \mathbf{S}_x & \mathbf{T}_x \\ \mathbf{P}_x^\imath & \mathbf{R}_x^\imath & \mathbf{T}_x^\imath & \mathbf{S}_x^\imath \\ \mathbf{S}_x^\jmath & \mathbf{T}_x^\jmath & \mathbf{R}_x^\jmath & \mathbf{P}_x^\jmath \\ \mathbf{T}_x^\kappa & \mathbf{S}_x^\kappa & \mathbf{P}_x^\kappa & \mathbf{R}_x^\kappa \end{bmatrix} \tag{12}$$

where $\mathbf{R}_x = E[\mathbf{x}\mathbf{x}^H]$, and the three pseudocovariances $\mathbf{P}_x = E[\mathbf{x}\mathbf{x}^{\imath H}]$, $\mathbf{S}_x = E[\mathbf{x}\mathbf{x}^{\jmath H}]$ and $\mathbf{T}_x = E[\mathbf{x}\mathbf{x}^{\kappa H}]$.

Proper (second order circular) signals, have probability distributions that are rotation invariant with respect to all the six possible pairs of axes (combinations of $\imath$, $\jmath$ and κ) [3], and thus equal powers in all the components, so that the three pseudocovariance matrices $\mathbf{P}_x$, $\mathbf{S}_x$ and $\mathbf{T}_x$ vanish.

Remark#1: The processing in $\mathbb{R}^4$ requires ten covariance matrices, as opposed to four in the quaternion domain since only $\mathbf{R}_x$, $\mathbf{P}_x$, $\mathbf{S}_x$ and $\mathbf{T}_x$ are needed to fully describe $\mathbf{R}_x^a$.

IV. MSE Analysis of the Widely Linear Model

To establish the extent to which the quaternion minimum MSE estimator based on the widely linear (WL) model yields superior performance over the strictly linear (SL) model, we adopt a two step approach. First we show that the semi-widely linear (SWL) model, given by ($\mathbf{c}$, $\mathbf{d}$ are coefficients)

$$\hat{y} = \mathbf{c}^H \mathbf{x} + \mathbf{d}^H \mathbf{x}^i \tag{13}$$

[1] Involutions can be considered as generalisations of the complex conjugate operator, when applied to every imaginary component of a quaternion.

[2] Throughout this paper, a vector $\mathbf{x}$ and its involutions are treated formally as independent variables. This is a usual formalism inherited from the complex domain, and the $\mathbb{CR}$-calculus.

attains a MSE than is smaller or equal to that of the strictly linear model $\hat{y} = \mathbf{w}^H\mathbf{x}$. We then show that for improper signals the fully-widely linear model in (11) offers better steady state performance than the semi-widely linear model in (13), hence outperforming the strictly linear model too.

A. Semi-Widely Linear Model vs Strictly Linear Model

To obtain expressions for the MSE and thus compare the minimum mean square error of the semi-widely linear model to that of the strictly linear model we must first obtain the optimum filter weight coefficients $\mathbf{c}$,$\mathbf{d}$ and $\mathbf{w}$. Similar to [7], starting from the orthogonality condition within the semi-widely linear model, we have

$$E[\mathbf{x}(y-\hat{y})^*] = 0 \rightarrow E[\mathbf{x}y^*] = E[\mathbf{x}\hat{y}^*] \quad (14)$$
$$E[\mathbf{x}^\imath(y-\hat{y})^*] = 0 \rightarrow E[\mathbf{x}^\imath y^*] = E[\mathbf{x}^\imath\hat{y}^*] \quad (15)$$

Substituting the estimator $\hat{y}$ in (13) into the above, we have

$$\mathbf{r} = \mathbf{R}\mathbf{c} + \mathbf{P}\mathbf{d} \qquad \mathbf{p} = \mathbf{P}^\imath\mathbf{c} + \mathbf{R}^\imath\mathbf{d} \quad (16)$$

where $\mathbf{r} = E[\mathbf{x}y^*]$ and $\mathbf{p} = E[\mathbf{x}^\imath y^*]$, while $\mathbf{R}$, and $\mathbf{P}$ are respectively the covariance matrix $E[\mathbf{x}\mathbf{x}^H]$ and i-pseudocovarince matrix $E[\mathbf{x}\mathbf{x}^{\imath H}]$. Solving for $\mathbf{c}$ and $\mathbf{d}$ we have

$$\mathbf{c} = [\mathbf{R} - \mathbf{P}\mathbf{R}^{-\imath}\mathbf{P}^i]^{-1}[\mathbf{r} - \mathbf{P}\mathbf{R}^{-\imath}\mathbf{p}] \quad (17)$$
$$\mathbf{d} = [\mathbf{R}^\imath - \mathbf{P}^\imath\mathbf{R}^{-1}\mathbf{P}]^{-1}[\mathbf{p} - \mathbf{P}^\imath\mathbf{R}^{-1}\mathbf{r}] \quad (18)$$

For the strictly linear model, the optimum filter weight $\mathbf{w}$ is given by the Wiener solution

$$\mathbf{w} = \mathbf{R}^{-1}\mathbf{r} \quad (19)$$

and the respective strictly linear and semi widely linear mean square errors, e_l^2 and e_{swl}^2, as

$$\begin{aligned} e_l^2 &= E[|y-\hat{y}|^2] = E[e(y - \mathbf{w}^H\mathbf{x})^H] \\ &= E[|y|^2] - \mathbf{r}^H\mathbf{R}^{-1}\mathbf{r} \\ e_{swl}^2 &= E[|y-\hat{y}|^2] = E[e(y - \mathbf{c}^H\mathbf{x} - \mathbf{d}^H\mathbf{x}^\imath)^H] \\ &= E[|y|^2] - \mathbf{c}^H\mathbf{r} - \mathbf{d}^H\mathbf{p} \end{aligned}$$

The squared error difference between the MSEs for the SL and SWL can then be written as (this also conforms with the complex case in [5])

$$\begin{aligned} \delta e_{swl}^2 = e_l^2 - e_{swl}^2 &= -\mathbf{r}^H\mathbf{R}^{-1} \quad (20) \\ &+ [\mathbf{r} - \mathbf{P}\mathbf{R}^{-\imath}\mathbf{p}]^H[\mathbf{R} - \mathbf{P}\mathbf{R}^{-\imath}\mathbf{P}^i]^{-1}\mathbf{r} \\ &+ [\mathbf{p} - \mathbf{P}^\imath\mathbf{R}^{-1}\mathbf{r}]^H[\mathbf{R}^\imath - \mathbf{P}^\imath\mathbf{R}^{-1}\mathbf{P}]^{-1}\mathbf{p} \end{aligned}$$

The Appendix shows that δe_{swl}^2 above can be written in the following compact form

$$\delta e_{swl}^2 = [\mathbf{p} - \mathbf{P}^\imath\mathbf{R}^{-1}\mathbf{r}]^H[\mathbf{R}^\imath - \mathbf{P}^\imath\mathbf{R}^{-1}\mathbf{P}]^{-1}[\mathbf{p} - \mathbf{P}^\imath\mathbf{R}^{-1}\mathbf{r}] \quad (21)$$

Remark#2: Observe that the term $\mathbf{R}^\imath - \mathbf{P}^\imath\mathbf{R}^{-1}\mathbf{P}$ is the Schur complement [10] of the semi-widely linear augmented covariance matrix

$$E[\mathbf{x}^a\mathbf{x}^{aH}] = \begin{bmatrix} \mathbf{R} & \mathbf{P} \\ \mathbf{P}^\imath & \mathbf{R}^\imath \end{bmatrix}$$

where $\mathbf{x}^a = [\mathbf{x}^T, \mathbf{x}^{\imath T}]^T$. Since the augmented covariance matrix is positive semi-definite and $\mathbf{P}^\imath = \mathbf{P}^H$, so too is its Schur complement $\mathbf{R}^\imath - \mathbf{P}^\imath\mathbf{R}^{-1}\mathbf{P}$, and its inverse and hence $\delta e_{swl}^2 \geq 0$.

Remark#3: The MSE of the SL model is always greater or equal to that of the SWL model, the equality only holds when $\mathbf{p} - \mathbf{P}^\imath\mathbf{R}^{-1}\mathbf{r} = \mathbf{0}$, that is, when the statistics of $\mathbf{x}$ is $\imath-$circular[3] and the input $\mathbf{x}^\imath$ is uncorrelated to the output y.

B. Widely Linear Model vs Semi-Widely Linear Model

To evaluate the MSE performance of the semi-widely linear model in (13)against the widely linear model in (11), we shall rewrite them as

$$\hat{y}_{swl} = \mathbf{f}\mathbf{x}^a \quad (22)$$
$$\hat{y}_{wl} = \mathbf{o}\mathbf{x}^a + \mathbf{l}\mathbf{x}^b \quad (23)$$

where $\mathbf{x}^a = [\mathbf{x}^T, \mathbf{x}^\imath T]^T$, $\mathbf{x^b} = [\mathbf{x}^{\jmath T}, \mathbf{x}^{\kappa T}]^T$ and $\mathbf{f} = [\mathbf{c}^T, \mathbf{d}^T]^T$, $\mathbf{o} = [\mathbf{u}^T, \mathbf{v}^T]^T$, and $\mathbf{l} = [\mathbf{g}^T, \mathbf{h}^T]^T$ are coefficient vectors. Written in this form, it is straightforward to follow the approach taken in the previous section to obtain the optimal weights $\mathbf{g}$ and $\mathbf{h}$, as

$$\mathbf{o} = [\mathbf{N} - \mathbf{M}\mathbf{N}^{-\jmath}\mathbf{M}^\jmath]^{-1}[\mathbf{n} - \mathbf{M}\mathbf{N}^{-\jmath}\mathbf{m}] \quad (24)$$
$$\mathbf{l} = [\mathbf{N}^\jmath - \mathbf{M}^\jmath\mathbf{N}^{-1}\mathbf{M}]^{-1}[\mathbf{m} - \mathbf{M}^\jmath N^{-1}\mathbf{n}] \quad (25)$$

where $\mathbf{N} = E[\mathbf{x}^a\mathbf{x}^{aH}]$ and $\mathbf{M} = E[\mathbf{x}^a\mathbf{x}^{bH}]$ are the corresponding augmented covariance matrices, and $\mathbf{n} = E[\mathbf{x}^a y^*]$ and $\mathbf{m} = E[\mathbf{x}^b y^*]$ augmented cross-correlations. The term 'augmented' is used because the input vectors $\mathbf{x}^a$ and $\mathbf{x}^b$ are each made up of two components, the involutions of $\mathbf{x}$. From the previous section, the difference in the MSE between the widely linear and semi widely linear models can be written as

$$\delta e_{wl}^2 = e_{swl}^2 - e_{wl}^2 \quad (26)$$
$$\delta e_{wl}^2 = [\mathbf{m} - \mathbf{M}^\jmath\mathbf{N}^{-1}\mathbf{m}]^H[\mathbf{N}^\jmath - \mathbf{M}^\jmath\mathbf{N}^{-1}\mathbf{M}]^{-1}[\mathbf{m} - \mathbf{M}^\jmath\mathbf{N}^{-1}\mathbf{m}]$$

From Remark 2, $[\mathbf{N}^\jmath - \mathbf{M}^\jmath\mathbf{R}^{-1}\mathbf{M}]^{-1}$ is positive semidefinite and hence $\delta e_{wl} \geq 0$.

Remark#4: The MSE of the semi-widely linear model is always greater or equal to that of the widely linear model, the equality holds only when $\mathbf{m} - \mathbf{M}^\jmath\mathbf{N}^{-1}\mathbf{m} = \mathbf{0}$, that is, when the statistics of $\mathbf{x}$ is $\jmath-$ and $\kappa-$circular and the input $\mathbf{x}^b$ is uncorrelated to the output y.

V. Simulations

To demonstrate experimentally the findings, we compared the MSEs of the quaternion LMS [11] and widely linear quaternion LMS [12] on one step ahead prediction of a process generated by the following AR(4) model

$$\begin{aligned} x(k) &= 1.79x(k-1) - 1.85x(k-2) + 1.27x(k-3) \\ &\quad -0.41x(k-4) + r(k) \end{aligned}$$

where $r(k) = n(k) + 0.9n^*(k-1)$ and $n(k)$ is circular white Gaussian noise. Due to the presence of $n^*(k-1)$ in the noise

[3]The term $\imath-$circular refers to the fact that the pseudocovariance $E[\mathbf{x}\mathbf{x}^{iH}] = \mathbf{0}$.

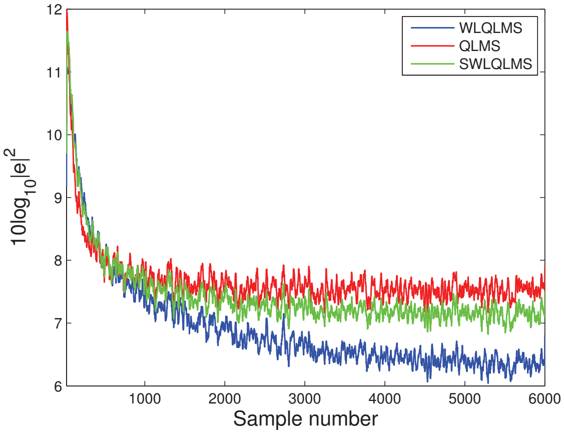

Fig. 1. Steady state performance of WLQLMS and QLMS for the prediction of an AR(4) process driven by noncircular white noise.

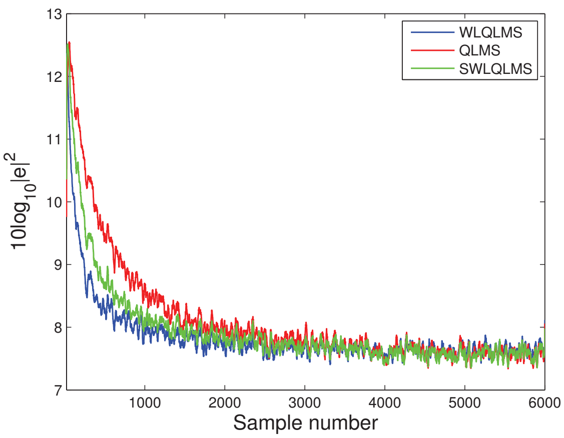

Fig. 2. Steady state performance of WLQLMS and QLMS for the prediction of an AR(4) process driven by circular white noise.

term $r(k)$, the widely linear model is needed to fully capture the cross-correlation between the output $x(k)$ and the input vector $[x(k-1), x(k-2), x(k-3), x(k-4)]$. Figure 1 shows the learning curves for the WLQLMS, SWLQLMS (semi-widely linear QLMS) and QLMS - as expected the widely linear WLQLMS achieved lower MSE than SWLQLMS, which itself achieved lower MSE than QLMS. We next investigated the prediction performance when the driving noise $r(k) = n(k)$ was circular. Figure 2 shows the evolution of the MSE for the QLMS and WLQLMS. Observe that since the driving noise is circular, the WLQLMS, SWLQLMS and QLMS achieved the same steady state performance. Table I compares the simulated e_l^2, e_{swl}^2 and δe_{wl}^2 to the theoretical value as measured by (21) and (27).

Note the close match between both the theoretical and experimental measures and how, as desired, the simulated MSE is a sum of the theoretical MSE and the excess MSE of the filter which is approximately $J_{ex} \approx \frac{1}{2}\mu J_{min} tr(\mathbf{R})$, where μ is the step size, $tr(\mathbf{R})$ is the trace of the input vector

TABLE I
THEORETICAL AND SIMULATED e^2 AND δe^2 FOR THE CIRCULAR AND NONCIRCULAR AR(4) MODEL

	e_l^2	e_{swl}^2	e_{wl}^2	δe_{swl}^2	δe_{wl}^2
Noncircular AR(4)					
simulated	7.58	7.27	6.23	0.31	1.04
theoretical	6.57	6.23	5.22	0.34	1.01
excess MSE	0.98	1.01	1.09	-	-
Circular AR(4)					
theoretical	7.67	7.67	7.68	0.00	0.00
simulated	6.54	6.54	6.54	0.00	0.00
excess MSE	1.12	1.11	1.10	-	-

covariance matrix and J_{min} is the theoretical MSE.

VI. CONCLUSIONS

We have proved that the MSE achieved by using the strictly linear model is always greater than or equal to that achieved by the widely linear model. A theoretical performance bound for noncircular signals is also established and the validity of the result is demonstrated on illustrative simulations.

APPENDIX: QUATERNION LINEAR MODEL

Apply the Woodbury matrix identity to (21) to give

$$\delta e_{swl}^2 = -\mathbf{r}^H\mathbf{R}^{-1} \qquad (27)$$
$$+\underbrace{[\mathbf{r} - \mathbf{P}\mathbf{R}^{-\imath}\mathbf{p}]^H[\mathbf{R}^{-1} - \mathbf{R}^{-1}\mathbf{P}(\mathbf{R}^{\imath} - \mathbf{P}^{\imath}\mathbf{R}^{-1}\mathbf{P})^{-1}\mathbf{P}^{\imath}\mathbf{R}^{-1}]\mathbf{r}}_{\alpha}$$
$$+\underbrace{[\mathbf{p} - \mathbf{P}^{\imath}\mathbf{R}^{-1}\mathbf{r}]^H[\mathbf{R}^{-\imath} - \mathbf{R}^{-\imath}\mathbf{P}^{\imath}(\mathbf{R} - \mathbf{P}\mathbf{R}^{-\imath}\mathbf{P}^{\imath})^{-1}\mathbf{P}\mathbf{R}^{-\imath}]\mathbf{p}}_{\beta}$$

Expanding the term α above we have

$$\alpha = \underbrace{(\mathbf{P}\mathbf{R}^{-\imath}\mathbf{p})^H\mathbf{R}^{-1}\mathbf{p}(\mathbf{R}^{\imath} - \mathbf{P}^{\imath}\mathbf{R}^{-1}\mathbf{p})^{-1}\mathbf{P}^{\imath}\mathbf{R}^{-1}\mathbf{r}}_{\theta}$$
$$-\underbrace{(\mathbf{P}\mathbf{R}^{-\imath}\mathbf{p})^H\mathbf{R}^{-1}\mathbf{r}}_{\phi} - \underbrace{\mathbf{r}^H\mathbf{R}^{-1}\mathbf{P}(\mathbf{R}^{\imath} - \mathbf{P}^{\imath}\mathbf{R}^{-1}\mathbf{P})^{-1}\mathbf{P}^{\imath}\mathbf{R}^{-1}\mathbf{r}}_{\omega}$$

Using $\mathbf{P}^{\imath} = \mathbf{P}^H$ we can rewrite the terms θ, ϕ and ω as

$$\begin{aligned} \theta &= \mathbf{p}^H[\mathbf{R}^{-\imath}\mathbf{P}\mathbf{R}^{-1}\mathbf{p}(\mathbf{R}^{\imath} - \mathbf{P}^{\imath}\mathbf{R}^{-1}\mathbf{p})]\mathbf{P}^{\imath}\mathbf{R}^{-1}\mathbf{r} \\ \phi &= \mathbf{p}^H\mathbf{R}^{-\imath}\mathbf{P}^{\imath}\mathbf{R}^{-1} \\ \omega &= (\mathbf{P}^{\imath}\mathbf{R}^{-1}\mathbf{r})^H(\mathbf{R}^{\imath} - \mathbf{P}^{\imath}\mathbf{R}^{-1}\mathbf{P})^{-1}(\mathbf{P}^{\imath}\mathbf{R}^{-1}\mathbf{r}) \end{aligned}$$

which allows us to express the term α as

$$[\mathbf{p}^H - (\mathbf{P}^{\imath}\mathbf{R}^{-1}\mathbf{r})^H][\mathbf{R}^{-\imath} - \mathbf{R}^{-\imath}\mathbf{P}^{\imath}(\mathbf{R} - \mathbf{P}\mathbf{R}^{-\imath}\mathbf{P}^{\imath})^{-1}\mathbf{P}\mathbf{R}^{-\imath}] \times[-\mathbf{P}^{\imath}\mathbf{R}^{-1}\mathbf{r}]$$

Making use of the common factor between α above and β in (27), we can rewrite (27) as

$$\delta e_{swl}^2 = [\mathbf{p} - \mathbf{P}^{\imath}\mathbf{R}^{-1}\mathbf{r}]^H[\mathbf{R}^{\imath} - \mathbf{P}^{\imath}\mathbf{R}^{-1}\mathbf{P}]^{-1}[\mathbf{p} - \mathbf{P}^{\imath}\mathbf{R}^{-1}\mathbf{r}]$$

REFERENCES

[1] S. Miron, N. Le Bihan, and J. Mars, "High resolution vector-sensor array processing based on biquaternions," in *2006 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP*, vol. 4, 2006, pp. IV–IV.

[2] J.-P. Delmas and H. Abeida, "On the degree of second-order non-circularity of complex random variables," in *IEEE International Conference on Acoustics, Speech and Signal Processing,ICASSP*, 2008, pp. 3905–3908.

[3] C. C. Took and D. P. Mandic, "Augmented second-order statistics of quaternion random signals," *Signal Processing*, vol. 91, no. 2, pp. 214–224, 2011.

[4] J. Navarro-Moreno, R. Fernandez-Alcala, and J. Ruiz-Molina, "A quaternion widely linear series expansion and its applications," *Signal Processing Letters, IEEE*, vol. 19, no. 12, pp. 868 –871, dec. 2012.

[5] B. Picinbono and P. Chevalier, "Widely linear estimation with complex data," *IEEE Transactions on Signal Processing*, vol. 43, no. 8, pp. 2030–2033, 1995.

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

[7] B. Picinbono and P. Bondon, "Second-order statistics of complex signals," *IEEE Transactions on Signal Processing*, vol. 45, no. 2, pp. 411 –420, 1997.

[8] B. Picinbono, "On circularity," *IEEE Transactions on Signal Processing*, vol. 42, no. 12, pp. 3473–3482, 1994.

[9] L. De Lathauwer and B. De Moor, "On the Blind Separation of Non-circular Sources," in *Proceedings of EUSIPCO*, vol. 2, 2002, pp. 99–102.

[10] R. Sznajder, M. Gowda, and M. Moldovan, "More results on schur complements in Euclidean Jordan algebras," *Journal of Global Optimization*, vol. 53, no. 1, pp. 121–134, 2012.

[11] C. Cheong Took and D. P. Mandic, "The quaternion LMS algorithm for adaptive filtering of hypercomplex processes," *IEEE Transactions on Signal Processing*, vol. 57, no. 3, pp. 1316–1327, 2009.

[12] C. C. Took and D. P. Mandic, "A quaternion widely linear adaptive filter," *IEEE Transactions on Signal Processing*, vol. 58, no. 8, pp. 4427–4431, 2010.