Signal Processing 93 (2013) 2573–2580

Contents lists available at SciVerse ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Prediction of wide-sense stationary quaternion random signals

Jesús Navarro-Moreno [a,*], Rosa M. Fernández-Alcalá [a], Clive Cheong Took [b], Danilo P. Mandic [c]

[a] University of Jaén, Department of Statistics and Operations Research, Campus Las Lagunillas, 23071 Jaén, Spain
[b] University of Surrey, Department of Computing, Guildford GU27XH, UK
[c] Imperial College, Department of Electrical and Electronic Engineering, London SW7 2AZ, UK

ARTICLE INFO

Article history:
Received 15 May 2012
Received in revised form
6 December 2012
Accepted 5 February 2013
Available online 13 February 2013

Keywords:
Quaternion random signals
Widely linear prediction

ABSTRACT

An efficient widely linear prediction algorithm is introduced for the class of wide-sense stationary quaternion signals. Specifically, using second order statistics information in the quaternion domain, a multivariate Durbin–Levison-like algorithm is derived. The proposed solution can be applied under a very general formulation of the problem, allowing for the estimation of a function of the quaternion signal which is observed through a system with both additive/multiplicative noises.

1. Introduction

The recent progress in sensor technology has brought to light several classes of multidimensional and multivariate signals, for which many established algorithms are used in an ad hoc fashion to handle these signals in the real domain. For three- and four-dimensional signals, this approach, though practical, does not offer the generality of quaternions, since e.g. real and complex numbers are special instances of quaternions. In other words, a quaternion-valued solution may solve a real- or complex-valued problem, but the reverse may not hold. This advantage has led to the growing popularity of quaternion in the signal processingcommunity [18].

Quaternions offer a new possibility to the signal processing community, to operate directly in a multidimensional domain. In doing so, we treat a multidimensional signal as a single variable, which offers generality and flexibility. For instance, quaternions facilitate the modeling of three- and four-dimensional signals, and account for the co-information between the data channels in a natural way. In terms of applications, quaternions have been employed in image processing [11], seismic processing [17], robotics [8], biomedicine [16], avionics [1], security [2], Kalman filtering [9], MUSIC spectrum estimation [15], and singular value decomposition for vector sensing [14], among others. Another benefit of quaternions comes from the fact that they allow correlation and coupling to prevail within the channels of a multivariate independent component, for which independent component analysis in the real domain [13] is not suitable.

One important paradigm in statistical signal processing is prediction. The problem can also be naturally addressed in the multidimensional domain where the correlation information between quaternion components is accounted for. Following this conventional or strictly linear (SL) processing, a quaternion least mean square (QLMS) prediction algorithm is proposed by Cheong Took and Mandic [4]. Moreover, benefits reported by a widely linear (WL)

* Corresponding author. Tel.: +34 953211912, +34 953212449; fax: +34 953212034.
E-mail addresses: jnavarro@ujaen.es (J. Navarro-Moreno), rmfernan@ujaen.es (R.M. Fernández-Alcalá), c.cheong-took@surrey.ac.uk (C. Cheong Took), d.mandic@imperial.ac.uk (D.P. Mandic).

0165-1684/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.sigpro.2013.02.007

processing in the complex field, characterized by taking both correlation and complementary functions into account, have motivated the authors to also derive an augmented QLMS (AQLMS) by considering an augmented vector formed by the signal and its conjugate. However, if this augmented vector is sufficient to obtain a complete description of the second order statistics of a complex signal, in the quaternion domain the three perpendicular quaternion involutions must be also considered for this purpose. In this framework, the augmented statistics and the corresponding WL processing were introduced by Cheong Took and Mandic [5] to formulate a widely linear QLMS (WL-QLMS) prediction algorithm. All these algorithms cater for non-stationary signals such as wind. The non-stationarity of the data makes the explicit computation of the correlation matrices redundant, hence the suitability of these adaptive predictors.

For stationary signals, the readily available of second order statistics implies that there is a high potential for enhanced accuracy of the prediction solution, especially when the Yule–Walker equations are considered. The most elegant way to solve these equations is to employ the Durbin–Levison (DL) recursion [3]. However, the derivation of the DL recursion in the quaternion domain is not straightforward. Although quaternions have a division algebra which is a generalization of that in the real and complex domains, the non-commutativity of quaternion product is a stumbling block that should be addressed in the derivation of the DL algorithm. Moreover for generality, the DL method has to be formulated in such a way that (i) it can account for the estimation of any function (linear/non-linear) of multivariate quaternion signals and (ii) it should operate under both additive and multiplicative noise.

This work aims to provide some fundamental tools for quaternion signal processing in the context of linear and widely linear prediction of stationary data. To this end, we first provide a one-stage prediction algorithm by following quaternion WL (QWL) processing in Section 3. Next, the computational efficiency of DL recursion is exploited and adapted to the non-commutative algebra of quaternions. Specifically, to make full use of statistical properties of the quaternion-valued signal, the so-called augmented statistics is incorporated in the derivation of the DL algorithm, through the consideration of perpendicular quaternion involutions. In Section 4, the quaternion DL algorithm is extended to a more general prediction problem where the estimation of a quaternion signal based upon noisy observations is addressed. Finally, simulations on both synthetic and real-world data support the proposed approach. For clarity, we start with a brief review of some properties of quaternion random vectors that will be used throughout the paper.

2. Preliminaries

In the following, all signals are assumed to be zero-mean. The superscripts '$*$', 'T' and 'H' represent the quaternion conjugate, transpose and quaternion conjugate transpose operators, respectively.

A quaternion random signal can be defined by $q_t = a_t + \mathrm{i}b_t + \mathrm{j}c_t + \mathrm{k}d_t$, where $a_t, b_t, c_t, d_t \in \mathbb{R}$ are random variables and $\mathrm{i}^2 = \mathrm{j}^2 = \mathrm{k}^2 = \mathrm{ijk} = -1$, which implies that $\mathrm{ij} = \mathrm{k}$, $\mathrm{ki} = \mathrm{j}$, and $\mathrm{jk} = \mathrm{i}$.

Let $\boldsymbol{x} = [x_1, \ldots, x_{n_1}]^{\mathrm{T}}$ and $\boldsymbol{y} = [y_1, \ldots, y_{n_2}]^{\mathrm{T}}$ be two n_1 and n_2-dimensional quaternion random vectors, where x_i and y_j are quaternion random signals. Define the outer product $\langle \boldsymbol{x}, \boldsymbol{y} \rangle_{\mathcal{Q}} = E[\boldsymbol{x}\boldsymbol{y}^{\mathrm{H}}]$ with

$$\boldsymbol{x}\boldsymbol{y}^{\mathrm{H}} = \begin{bmatrix} x_1 y_1^* & x_1 y_2^* & \ldots & x_1 y_{n_2}^* \\ \vdots & \vdots & \ldots & \vdots \\ x_{n_1} y_1^* & x_{n_1} y_2^* & \ldots & x_{n_1} y_{n_2}^* \end{bmatrix}$$

where $x_i y_j^*$ represents the quaternion product between the quaternions x_i and y_j^* (see [6] for more details).

The complete description of the second order statistics of the quaternion q_t in the quaternion domain $\mathbb{H}$ can be given by the augmented quaternion vector $\boldsymbol{q}_t = [q_t, q_t^*, q_t^{\mathrm{i}}, q_t^{\mathrm{j}}]^{\mathrm{T}}$, where q_t^*, q_t^{i} and q_t^{j} are defined by[1] [6]

$$q_t^* = a_t - \mathrm{i}b_t - \mathrm{j}c_t - \mathrm{k}d_t$$

$$q_t^{\mathrm{i}} = a_t + \mathrm{i}b_t - \mathrm{j}c_t - \mathrm{k}d_t$$

$$q_t^{\mathrm{j}} = a_t - \mathrm{i}b_t + \mathrm{j}c_t - \mathrm{k}d_t$$

Denote by $\overline{\mathrm{sp}}\{\boldsymbol{q}_1, \ldots, \boldsymbol{q}_n\}$ the closed span of $\{\boldsymbol{q}_1, \ldots, \boldsymbol{q}_n\}$.

Moreover, the concept of wide-sense stationarity (WSS) in the quaternion domain was introduced by Cheong Took and Mandic [6]. A centered quaternion random signal q_t is said to be WSS if and only if it fulfills the following two conditions:

1. The correlation and its complementary correlations are functions of only the lag τ, that is
   - correlation function: $C_q(\tau) = E[q_t q_{t+\tau}^*]$,
   - i-correlation function: $C_{q,q^{\mathrm{i}}}(\tau) = E[q_t q_{t+\tau}^{\mathrm{i}^*}]$,
   - j-correlation function: $C_{q,q^{\mathrm{j}}}(\tau) = E[q_t q_{t+\tau}^{\mathrm{j}^*}]$,
   - k-correlation function: $C_{q,q^{\mathrm{k}}}(\tau) = E[q_t q_{t+\tau}^{\mathrm{k}^*}]$.
2. The correlation function is finite, i.e., $C_q(\tau) < \infty$.

Note that, condition 2 is equivalent to assuming that the autocorrelation function of the augmented quaternion vector $\boldsymbol{q}_t$ is a function of only the lag τ, that is

$$\boldsymbol{\Gamma}_{\boldsymbol{q}}(\tau) = \langle \boldsymbol{q}_{t+\tau}, \boldsymbol{q}_t \rangle_{\mathcal{Q}} \tag{1}$$

On the other hand, for a quaternion random signal, there exist two main kinds of properness [18]: $\mathbb{Q}$-properness (or $\mathbb{H}$-circularity) and $\mathbb{C}^{\eta}$-properness (or $\mathbb{C}^{\eta}$-circularity). A quaternion random vector q_t is $\mathbb{Q}$-proper if and only if the three complementary functions $C_{q,q^{\mathrm{i}}}(t,s) = E[q_t q_s^{\mathrm{i}^*}]$, $C_{q,q^{\mathrm{j}}}(t,s) = E[q_t q_s^{\mathrm{j}^*}]$ and $C_{q,q^{\mathrm{k}}}(t,s) = E[q_t q_s^{\mathrm{k}^*}]$ vanish. Moreover, a quaternion random vector q_t is $\mathbb{C}^{\eta}$-proper, for any pure imaginary unit $\eta \in \{\mathrm{i},\mathrm{j},\mathrm{k}\}$ if and only if all the complementary functions vanish except $C_{q,q^{\eta}}(t,s) = E[q_t q_s^{\eta^*}]$. Note that, $\mathbb{Q}$-properness implies $\mathbb{C}^{\eta}$-properness.

[1] Note that this choice is not unique and any other combination of four elements of $\{q_t, q_t^*, q_t^{\mathrm{i}}, q_t^{\mathrm{j}}, q_t^{\mathrm{k}}\}$, with $q_t^{\mathrm{k}} = a_t - \mathrm{i}b_t - \mathrm{j}c_t + \mathrm{k}d_t$, or their conjugates can be used with the same effect, due to the relationship $q_t^* = \frac{1}{2}[q_t^{\mathrm{i}} + q_t^{\mathrm{j}} + q_t^{\mathrm{k}} - q_t]$ [6].

3. One-stage prediction algorithm

In this section, the aim is to provide an efficient computational algorithm for the one-stage prediction problem for the quaternion q_t, that is, we aim to predict the signal q_{n+1} on the basis of the set of observations $\{q_1,\ldots,q_n\}$, under a WL processing in the quaternion domain. This predictor will be denoted by $\hat{q}^{WL}(n+1/n)$.

This problem can be posed as estimating the correct projection of the augmented quaternion vector $\boldsymbol{q}_{n+1}$ onto the space $\overline{\mathrm{sp}}\{\boldsymbol{q}_1,\ldots,\boldsymbol{q}_n\}$. This projection is of the form

$$\hat{\boldsymbol{q}}(n+1/n)=\sum_{j=1}^{n}\boldsymbol{\Theta}_{n,j}\boldsymbol{q}_{n+1-j} \tag{2}$$

where the coefficients $\boldsymbol{\Theta}_{n,j}$ can be determined from the projection theorem, and thus, the QWL estimator of q_{n+1} is trivially

$$\hat{q}^{WL}(n+1/n)=\sum_{j=1}^{n}\boldsymbol{\theta}_{n,j}\boldsymbol{q}_{n+1-j}$$

with $\boldsymbol{\theta}_{n,j}=[1,0,0,0]\boldsymbol{\Theta}_{n,j}$ for which the associated mean-square error, termed WL-QMSE, is $\varepsilon_{n,1}=[1,0,0,0]\ \boldsymbol{\Sigma}_{n,1}\ [1,0,0,0]^{\mathrm{T}}$, where

$$\boldsymbol{\Sigma}_{n,1}=\langle\boldsymbol{q}_{n+1}-\hat{\boldsymbol{q}}(n+1/n),\boldsymbol{q}_{n+1}-\hat{\boldsymbol{q}}(n+1/n)\rangle_{\mathcal{Q}}$$

A priori, the computation of (2) proves to be more difficult than the direct computation of $\hat{q}^{WL}(n+1/n)$. However, exploiting the WSS quaternion properties leads to the efficient computation of the terms $\boldsymbol{\Theta}_{n,j}$ and $\boldsymbol{\Sigma}_{n,1}$. Hence, using the Hilbert space theory and property (1), a method similar to the DL algorithm is next proposed for the efficient computation of these terms.

This algorithm requires the simultaneous solution of two sets of equations, one arising from the computation of the predictor (2) and the other from the computation of the estimator

$$\hat{\boldsymbol{q}}(1/2,\ldots,n)=\sum_{j=2}^{n}\tilde{\boldsymbol{\Theta}}_{n-1,j-1}\boldsymbol{q}_j \tag{3}$$

where $\hat{\boldsymbol{q}}(1/2,\ldots,n)$ denotes the estimator $\boldsymbol{q}_1$ based on the observations $\boldsymbol{q}_2,\ldots,\boldsymbol{q}_n$, for which the mean-square error is given by

$$\tilde{\boldsymbol{\Sigma}}_{n-1,1}=\langle\boldsymbol{q}_1-\hat{\boldsymbol{q}}(1/2,\ldots,n),\boldsymbol{q}_1-\hat{\boldsymbol{q}}(1/2,\ldots,n)\rangle_{\mathcal{Q}} \tag{4}$$

In the next algorithm, the multivariate DL method [3] is generalized in order to operate in the quaternion domain. This extension is non-trivial, as the problem of the non-commutativity of the quaternion product is encountered and needs to be addressed.

Algorithm 1.

$$\begin{aligned}
\boldsymbol{\Theta}_{n,n}&=\boldsymbol{\Delta}_{n-1}\tilde{\boldsymbol{\Sigma}}_{n-1,1}^{-1}\\
\tilde{\boldsymbol{\Theta}}_{n,n}&=\boldsymbol{\Delta}_{n-1}^{\mathrm{H}}\boldsymbol{\Sigma}_{n-1,1}^{-1}\\
\boldsymbol{\Theta}_{n,j}&=\boldsymbol{\Theta}_{n-1,j}-\boldsymbol{\Theta}_{n,n}\tilde{\boldsymbol{\Theta}}_{n-1,n-j},\quad j=1,\ldots,n-1\\
\tilde{\boldsymbol{\Theta}}_{n,j}&=\tilde{\boldsymbol{\Theta}}_{n-1,j}-\tilde{\boldsymbol{\Theta}}_{n,n}\boldsymbol{\Theta}_{n-1,n-j},\quad j=1,\ldots,n-1
\end{aligned} \tag{5}$$

where

$$\boldsymbol{\Sigma}_{n,1}=\boldsymbol{\Sigma}_{n-1,1}-\boldsymbol{\Theta}_{n,n}\boldsymbol{\Delta}_{n-1}^{\mathrm{H}}$$
$$\tilde{\boldsymbol{\Sigma}}_{n,1}=\tilde{\boldsymbol{\Sigma}}_{n-1,1}-\tilde{\boldsymbol{\Theta}}_{n,n}\boldsymbol{\Delta}_{n-1}$$

$$\boldsymbol{\Delta}_n=\boldsymbol{\Gamma}_{\boldsymbol{q}}(n+1)-\sum_{j=1}^{n}\boldsymbol{\Theta}_{n,j}\boldsymbol{\Gamma}_{\boldsymbol{q}}(n+1-j) \tag{6}$$

with

$$\boldsymbol{\Sigma}_{0,1}=\tilde{\boldsymbol{\Sigma}}_{0,1}=\boldsymbol{\Gamma}_{\boldsymbol{q}}(0)$$
$$\boldsymbol{\Delta}_0=\boldsymbol{\Gamma}_{\boldsymbol{q}}(1)$$

4. A generic prediction algorithm

In this section, our purpose is to extend the previous results to a more general scope. In particular, we consider a quaternion signal x_t that cannot be observed directly but through a function of the signal $q_t=\mathcal{L}(x_t)$, in such a way that q_t is a WSS quaternion. Based on the quaternion observations q_t, we aim to estimate a function of x_t, $z_t=\mathcal{T}(x_t)$, such that z_t has correlation function $\gamma_z(k,l)=\langle z_k,z_l\rangle_{\mathcal{Q}}$ and cross-correlation function with the augmented quaternion $\boldsymbol{q}_t$, $\boldsymbol{\alpha}(k,l)=\langle z_k,\boldsymbol{q}_l\rangle_{\mathcal{Q}}$. Suppose that, for $k\geq l$, this function $\boldsymbol{\alpha}(k,l)$ only depends on the difference $k-l$, that is $\boldsymbol{\alpha}(\tau)=\langle z_{t+\tau},\boldsymbol{q}_t\rangle_{\mathcal{Q}}$.

The following results provide recursive algorithms for this estimation problem. Specifically, we focus on the prediction and filtering cases, that is, we aim to estimate z_{n+m} on the basis of the observations $\boldsymbol{q}_1,\ldots,\boldsymbol{q}_n$, for some $m=0,1,\ldots$. Thus, the value $m=0$ corresponds to the filtering problem and for $m>0$ we have the prediction problem. Define the QWL predictor of z_{n+m} as follows:

$$\hat{z}^{WL}(n+m/n)=\sum_{j=1}^{n}\boldsymbol{\pi}_{n,j}^{m}\boldsymbol{q}_{n+1-j} \tag{7}$$

and its associated WL-QMSE as

$$e_{n,m}=\langle z_{n+m}-\hat{z}^{WL}(n+m/n),z_{n+m}-\hat{z}^{WL}(n+m/n)\rangle_{\mathcal{Q}} \tag{8}$$

Observe that the terms $\boldsymbol{\pi}_{n,j}^{m}$ in (7) and $e_{n,m}$ in (8) depend on the fixed instant prediction ahead m considered in the estimation. However, in order to simplify the notation, for clarity we omit the index m in both expressions.

A recursive algorithm for computing $\boldsymbol{\pi}_{n,j}$ and e_n in (7) and (8), can be summarised as in Algorithm 2.

Algorithm 2.

$$\begin{aligned}
\boldsymbol{\pi}_{n,n}&=\boldsymbol{\lambda}_{n-1}\tilde{\boldsymbol{\Sigma}}_{n-1,1}^{-1}\\
\boldsymbol{\pi}_{n,j}&=\boldsymbol{\pi}_{n-1,j}-\boldsymbol{\pi}_{n,n}\tilde{\boldsymbol{\Theta}}_{n-1,n-j},\quad j=1,\ldots,n-1
\end{aligned} \tag{9}$$

where[2]

$$\boldsymbol{\lambda}_0=\boldsymbol{\alpha}(m)$$
$$\boldsymbol{\lambda}_n=\boldsymbol{\alpha}(n+m)-\sum_{j=1}^{n}\boldsymbol{\alpha}(n+m-j)\tilde{\boldsymbol{\Theta}}_{n,j}^{\mathrm{H}},\quad n\geq 1$$

and the matrices $\tilde{\boldsymbol{\Theta}}_{n,j}$ and $\tilde{\boldsymbol{\Sigma}}_{n,1}$ are computed from (5) and (6), respectively.

The error e_n is equal to

$$e_0=\gamma_z(m,m)$$

[2] $\boldsymbol{\lambda}_n$, $n\geq 0$, depends also of the value of m and, following the same argument, this is omitted.

$$e_n = e_{n-1} - \boldsymbol{\pi}_{n,n}\boldsymbol{\lambda}^{\mathrm{H}}_{n-1} + \gamma_z(n+m,n+m) - \gamma_z(n+m-1,n+m-1), \quad n \geq 1 \tag{10}$$

Proof. First of all, we need to compute $\boldsymbol{\pi}_{1,1}$. Consider $\hat{z}^{WL}(1+m/1) = \boldsymbol{\pi}_{1,1}\boldsymbol{q}_1$. Thus,

$$\boldsymbol{\pi}_{1,1} = \langle z_{1+m}, \boldsymbol{q}_1 \rangle_{\mathcal{Q}} \langle \boldsymbol{q}_1, \boldsymbol{q}_1 \rangle_{\mathcal{Q}}^{-1} = \boldsymbol{\alpha}(m)\boldsymbol{\Gamma}_{\boldsymbol{q}}^{-1}(0)$$

and then it is verified that $\boldsymbol{\lambda}_0 = \boldsymbol{\alpha}(m)$.

The error e_1 is given by

$$\begin{aligned} e_1 &= \langle z_{1+m} - \hat{z}^{WL}(1+m/1),\, z_{1+m} - \hat{z}^{WL}(1+m/1) \rangle_{\mathcal{Q}} \\ &= \langle z_{1+m} - \hat{z}^{WL}(1+m/1),\, z_{1+m} \rangle_{\mathcal{Q}} \\ &= \langle z_{1+m} - \boldsymbol{\pi}_{1,1}\boldsymbol{q}_1,\, z_{1+m} \rangle_{\mathcal{Q}} \\ &= \gamma_z(1+m,1+m) - \boldsymbol{\pi}_{1,1}\boldsymbol{\alpha}^{\mathrm{H}}(m) \end{aligned}$$

where we have used the orthogonality condition of $\hat{z}^{WL}(1+m/1)$. Hence, $e_0 = \gamma_z(m,m)$.

The following step is to prove (9) and (10) for $n>1$. Let $P_{\mathcal{K}}$ denotes the projection operator on an arbitrary set of quaternions $\mathcal{K}$. Let $\mathcal{K}_1^n = \overline{\mathrm{sp}}\{\boldsymbol{q}_2, \ldots, \boldsymbol{q}_n\}$ and $\mathcal{K}_2^n = \overline{\mathrm{sp}}\{\boldsymbol{q}_1 - P_{\mathcal{K}_1^n}\boldsymbol{q}_1\}$ be two subspaces of $\overline{\mathrm{sp}}\{\boldsymbol{q}_1, \ldots, \boldsymbol{q}_n\}$. It is easy to check that $\overline{\mathrm{sp}}\{\boldsymbol{q}_1, \ldots, \boldsymbol{q}_n\} = \mathcal{K}_1^n \oplus \mathcal{K}_2^n$, that is, $\overline{\mathrm{sp}}\{\boldsymbol{q}_1, \ldots, \boldsymbol{q}_n\}$ is the direct sum of $\mathcal{K}_1^n$ and $\mathcal{K}_2^n$, so that

$$\hat{z}^{WL}(n+m/n) = P_{\mathcal{K}_1^n} z_{n+m} + P_{\mathcal{K}_2^n} z_{n+m} \tag{11}$$

Next, $P_{\mathcal{K}_1^n} z_{n+m}$ and $P_{\mathcal{K}_2^n} z_{n+m}$ are obtained. Suppose that $P_{\mathcal{K}_1^n} z_{n+m}$ can be expressed in the form

$$P_{\mathcal{K}_1^n} z_{n+m} = \sum_{j=2}^{n} \boldsymbol{a}_{j-1}\boldsymbol{q}_{n+2-j}$$

then, it follows that

$$\boldsymbol{\alpha}(i+m-2) = \langle z_{n+m}, \boldsymbol{q}_{n+2-i} \rangle_{\mathcal{Q}} = \langle P_{\mathcal{K}_1^n} z_{n+m}, \boldsymbol{q}_{n+2-i} \rangle_{\mathcal{Q}} = \sum_{j=2}^{n} \boldsymbol{a}_{j-1}\boldsymbol{\Gamma}_{\boldsymbol{q}}(i-j)$$

for $i = 2, \ldots, n$. By solving this system of equations, we can check that $\boldsymbol{a}_j = \boldsymbol{\pi}_{n-1,j}$, $j = 1, \ldots, n-1$, and hence

$$P_{\mathcal{K}_1^n} z_{n+m} = \sum_{j=2}^{n} \boldsymbol{\pi}_{n-1,j-1}\boldsymbol{q}_{n+2-j} \tag{12}$$

On the other hand, by expressing $P_{\mathcal{K}_2^n} z_{n+m} = \boldsymbol{c}(\boldsymbol{q}_1 - P_{\mathcal{K}_1^n}\boldsymbol{q}_1)$, we obtain that

$$\boldsymbol{c} = \langle z_{n+m}, \boldsymbol{q}_1 - P_{\mathcal{K}_1^n}\boldsymbol{q}_1 \rangle_{\mathcal{Q}} \langle \boldsymbol{q}_1 - P_{\mathcal{K}_1^n}\boldsymbol{q}_1, \boldsymbol{q}_1 - P_{\mathcal{K}_1^n}\boldsymbol{q}_1 \rangle_{\mathcal{Q}}^{-1}$$

and taking (3) and (4) into account we get[3]

$$\boldsymbol{c} = \left(\boldsymbol{\alpha}(n+m-1) - \sum_{j=2}^{n} \boldsymbol{\alpha}(n+m-j)\tilde{\boldsymbol{\Theta}}^{\mathrm{H}}_{n-1,j-1} \right) \tilde{\boldsymbol{\Sigma}}^{-1}_{n-1,1}$$

Consequently, denoting

$$\boldsymbol{\lambda}_{n-1} = \boldsymbol{\alpha}(n+m-1) - \sum_{j=2}^{n} \boldsymbol{\alpha}(n+m-j)\tilde{\boldsymbol{\Theta}}^{\mathrm{H}}_{n-1,j-1}$$

it follows that

$$P_{\mathcal{K}_2^n} z_{n+m} = \boldsymbol{\lambda}_{n-1}\tilde{\boldsymbol{\Sigma}}^{-1}_{n-1,1}(\boldsymbol{q}_1 - P_{\mathcal{K}_1^n}\boldsymbol{q}_1) \tag{13}$$

and, from (11), (12), (3) and (13), we have that

$$\hat{z}^{WL}(n+m/n) = \boldsymbol{\lambda}_{n-1}\tilde{\boldsymbol{\Sigma}}^{-1}_{n-1,1}\boldsymbol{q}_1 + \sum_{j=1}^{n-1}(\boldsymbol{\pi}_{n-1,j} - \boldsymbol{\lambda}_{n-1}\tilde{\boldsymbol{\Sigma}}^{-1}_{n-1,1}\tilde{\boldsymbol{\Theta}}_{n-1,n-j})\boldsymbol{q}_{n+1-j}$$

Finally, by the uniqueness of $\hat{z}^{WL}(n+m/n)$ we prove (9), for $n>1$.

From (11) and the orthogonality condition of $\hat{z}^{WL}(n+m/n)$, we have that its associated error is given by

$$\begin{aligned} e_n &= \langle z_{n+m} - \hat{z}^{WL}(n+m/n), z_{n+m} - \hat{z}^{WL}(n+m/n) \rangle_{\mathcal{Q}} \\ &= \langle z_{n+m} - \hat{z}^{WL}(n+m/n), z_{n+m} \rangle_{\mathcal{Q}} \\ &= \langle z_{n+m} - P_{\mathcal{K}_1^n} z_{n+m}, z_{n+m} \rangle_{\mathcal{Q}} - \langle P_{\mathcal{K}_2^n} z_{n+m}, z_{n+m} \rangle_{\mathcal{Q}} \end{aligned} \tag{14}$$

We shall now compute the last two terms in (14). From (12), we obtain

$$\langle z_{n+m} - P_{\mathcal{K}_1^n} z_{n+m}, z_{n+m} \rangle_{\mathcal{Q}} = \gamma_z(n+m,n+m) - \sum_{j=2}^{n} \boldsymbol{\pi}_{n-1,j-1}\boldsymbol{\alpha}^{\mathrm{H}}(j+m-2) \tag{15}$$

On the other hand, since $\hat{z}^{WL}(n+m/n) = \sum_{j=1}^{n} \boldsymbol{\pi}_{n,j}\boldsymbol{q}_{n+1-j}$, we have

$$e_n = \langle z_{n+m} - \hat{z}^{WL}(n+m/n), z_{n+m} \rangle_{\mathcal{Q}} = \gamma_z(n+m,n+m) - \sum_{j=1}^{n} \boldsymbol{\pi}_{n,j}\boldsymbol{\alpha}^{\mathrm{H}}(j+m-1)$$

Thus, (15) can be expressed in the form

$$\langle z_{n+m} - P_{\mathcal{K}_1^n} z_{n+m}, z_{n+m} \rangle_{\mathcal{Q}} = \gamma_z(n+m,n+m) + e_{n-1} - \gamma_z(n+m-1,n+m-1) \tag{16}$$

As for the second term, by using (9) and (13), it can be calculated as

$$\langle P_{\mathcal{K}_2^n} z_{n+m}, z_{n+m} \rangle_{\mathcal{Q}} = \boldsymbol{\lambda}_{n-1}\tilde{\boldsymbol{\Sigma}}^{-1}_{n-1,1} \langle \boldsymbol{q}_1 - P_{\mathcal{K}_1^n}\boldsymbol{q}_1, z_{n+m} \rangle_{\mathcal{Q}} = \boldsymbol{\pi}_{n,n}\boldsymbol{\lambda}^{\mathrm{H}}_{n-1} \tag{17}$$

where we have taken into account that $\boldsymbol{\lambda}_{n-1} = \langle z_{n+m}, \boldsymbol{q}_1 - P_{\mathcal{K}_1^n}\boldsymbol{q}_1 \rangle_{\mathcal{Q}}$.

Finally, using (16) and (17) in (14) we have (10), for $n \geq 1$. □

Remark 1. Algorithm 2 runs in $\mathcal{O}(n^2)$ time, with the same order of computational complexity as the classical DL algorithm.

5. Numerical examples

The benefits of the proposed approach were validated over three sets of experiments. The first two examples show the simultaneous forecasting of three-dimensional (3D) wind fields[4] from datasets recorded by a three-dimensional

[3] Note that $P_{\mathcal{K}_1^n}\boldsymbol{q}_1 = \hat{\boldsymbol{q}}(1/2, \ldots, n)$.

[4] The wind data can be obtained from http://www.commsp.ee.ic.ac.uk/~mandic/research/Smart-Grid-and-Renewables.htm.

ultrasonic anemometer at Imperial College London (ICL) in a controlled/closed environment (Example 1) and an open space (Example 2). In the third simulation example, the three-dimensional velocity of an aircraft is predicted from noisy measurements of its three-dimensional position.

Additionally, to compare the performance of the proposed algorithm with standard methods, a new synthetic example was considered, where a signal was defined in such a way that it is WSS and obeys a state-space model, and our solution is compared with a quaternion Kalman filter.

5.1. Example 1. Wind data in a controlled environment

In this experiment, wind data were recorded at a sampling frequency of 32 Hz in a controlled environment at ICL by using three fans placed around the anemometer at a distance up to 1 m. The air temperature was also measured and then such four-dimensional data sets were considered; the three speed direction of the wind (north–south, east–west and vertical direction) which are taken as pure quaternion components, together with the air temperature represent the real dimension of the full quaternion.

Then, on the basis of a sampling of 100 observed data, we considered the problem of predicting the quaternion signal (the three speed directions and the temperature) at the prediction horizon of $m=10$ steps ahead. For this purpose, Algorithm 2 was applied for the computation of the QWL estimator as well as its mean-square error, the WL-QMSE in (8). This error was compared in Fig. 1 with the mean-square errors associated to the m steps ahead quaternion predictions derived from a strictly linear processing (SL-QMSE), which only considers the information supplied by the quaternion signal,[5] and an augmented quaternion mean-square error (AQMSE) which uses information derived from solely the correlation and complementary functions, that is, second order statistics of the quaternion signal and its conjugate.

As expected, although augmented quaternion (AQ) predictions outperformed quaternion SL (QSL) predictions, the proposed QWL algorithm outperforms both AQ and QSL prediction results.

5.2. Example 2. Wind data in an urban environment

In this example, we computed the wind acceleration from the wind speed observations. Specifically, the signal x_t was a pure quaternion model formed by the three orthogonal wind speed components (north–south, east–west and vertical direction). Thus, we considered a set of observations of the signal provided by the three-dimensional ultrasonic anemometer in an interval of time where conditions for applying our solution are verified. From this set of observations, our aim was to predict, one-stage ahead, its mean-square derivative $z_t=\dot{x}_t$.

[5] Note that the SL-QMSE coincides with that of the classical Wiener filter for quaternion signals.

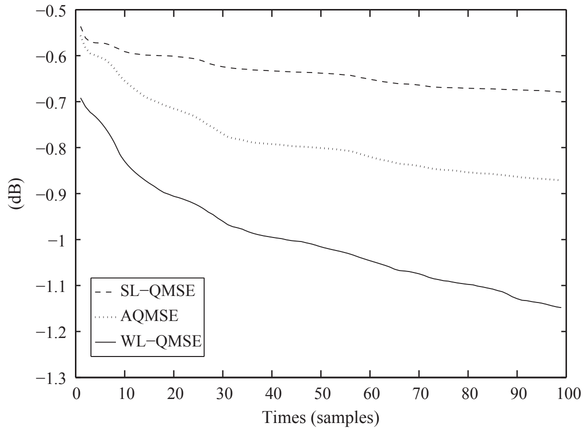

Fig. 1. WL-QMSE (solid line), AQMSE (dotted line) and SL-QMSE (dashed line) for the $m=10$ steps ahead predictor of the quaternion signal q_t.

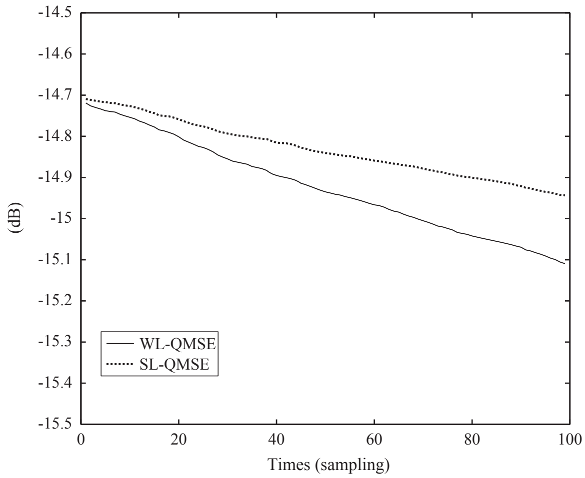

Fig. 2. WL-QMSE (solid line) and SL-QMSE (dotted line) associated with the one-stage predictor of z_t.

Fig. 2 compares the WL-QMSE associated to the QWL predictor $\hat{z}^{WL}(n+1/n)$, for $n=1,\ldots,100$, with the SL-QMSE obtained through a QSL processing. Again, we observe that the best prediction performance was obtained using a QWL processing.

Additionally, the advantage of the quaternion DL algorithm over the complex-valued version is shown in Table 1, where the mean-square values (MSVs) of the one-stage predictors $\hat{z}^{C}(n+1/n)$, $\hat{z}^{SL}(n+1/n)$ and $\hat{z}^{WL}(n+1/n)$ derived from the complex-valued, QSL and QWL DL algorithms, respectively, were calculated on the basis of two-dimensional speed directions at a time through the expression

$$\frac{1}{100}\sum_{n=1}^{100}(z_{n+1}-\hat{z}(n+1/n))(z_{n+1}-\hat{z}(n+1/n))^{*}$$

where $\hat{z}(n+1/n)$ is replaced by the corresponding predictor $\hat{z}^{C}(n+1/n)$, $\hat{z}^{SL}(n+1/n)$ or $\hat{z}^{WL}(n+1/n)$, in each case. As expected, in all cases, the results derived from a quaternion-valued DL algorithm outperforms complex-valued estimators, with QWL approach outperforming QSL solution.

5.3. Example 3. Aircraft trajectory tracking

In this example, we simulated the movement of an aircraft in a closed form orbit. Thus, the three-dimensional position was made quaternion valued (the real part was set to zero) and each experiment was conducted in the presence of both circular and non-circular additive quaternion quadruply white Gaussian noise (QWGN). Specifically, the observed signals were noisy measurements of the target position x_t obtained from the equation

$$q_t = x_t + w_t$$

where the QWGN w_t is of the form [7]

$$w_t = w_t^a + \mathrm{i}w_t^b + \mathrm{j}w_t^c + \mathrm{k}w_t^d \tag{18}$$

Table 1
MSVs of prediction using complex and quaternion-valued DL algorithms.

	Two-dimensional speed directions		
	East-north	East-vertical	North-vertical
MSV for $\hat{z}^C(n+1/n)$	0.0285	0.0281	0.0221
MSV for $\hat{z}^{SL}(n+1/n)$	0.0283	0.0279	0.0216
MSV for $\hat{z}^{WL}(n+1/n)$	0.0266	0.0258	0.0201

with w_t^a, w_t^b, w_t^c and w_t^d realizations of real-valued white Gaussian noises (WGN). As $\mathbb{H}$-circular noise, the four components of the QWGN (18) were considered to be independent real-valued WGN with a variance of 5. Moreover, a $\mathbb{C}^{\mathrm{i}}$-circular QWGN was considered by taking $w_t^a = z_a$, $w_t^b = z_b$, $w_t^c = 0.4w_t^a + 0.8w_t^b + z_c$ and $w_t^d = 0.8w_t^a - 0.4w_t^b + z_d$, where z_a, z_b, z_c and z_d are mutually independent real-valued WGN with variance of 5. Finally, as non-circular noise, the following noises were used in (18): $w_t^a = z_a$,

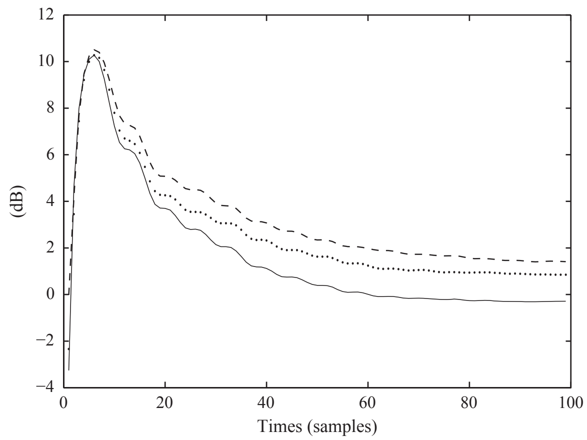

Fig. 4. Difference in performance (in dB) between SL-QMSE and WL-QMSE for the $m=5$ steps ahead predictor of the velocity of the aircraft in the presence of $\mathbb{H}$-circular (solid line), $\mathbb{C}^{\mathrm{i}}$-circular (dotted line) and non-circular (dashed line) noises.

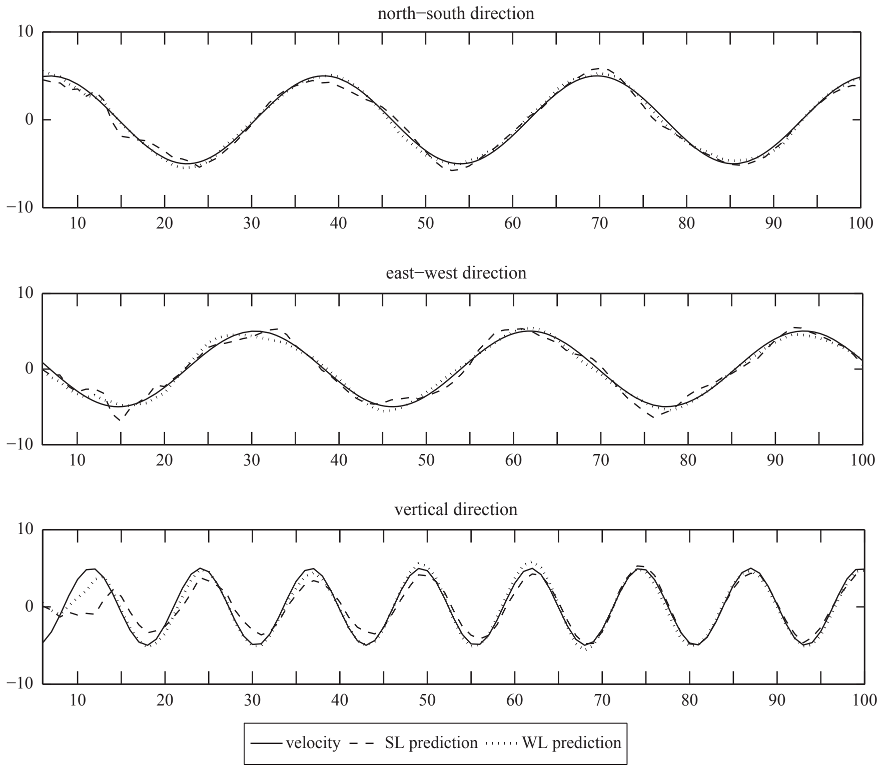

Fig. 3. Waveforms for the $m=5$ steps ahead predictions of the velocity of the aircraft using QSL and QWL processing.

$w_t^b = -0.6w_t^a + z_b$, $w_t^c = 0.8w_t^b + z_c$ and $w_t^d = 0.8w_t^a - 0.4w_t^b + z_d$, where z_a, z_b, z_c and z_d are independent real-valued WGN with variance of 5. In all cases, the signal-to-noise ratio SNR was at 15 dB.

The aim of the experiment was to predict the velocity of a moving target based on noisy observations of the target location with respect to a static sensor. Thus, the signal of interest was the mean-square derivative of the signal $z_t = \dot{x}_t$.

For the purpose of illustration, the $m=5$ steps ahead QWL predictor $\hat{z}^{WL}(n+m/n)$ was computed for $n=1,\ldots,100$. Fig. 3 displays a velocity trajectory (solid line) as well as the QWL (dotted line) and QSL (dashed line) prediction results, in the case of observations with non-circular QWGN. Observe that quaternion QWL predictor matches the target trajectory better than QSL predictor. Moreover, the faster convergence of the QWL predictor is noted. Observed that similar results were obtained with the $\mathbb{H}$-circular and $\mathbb{C}^{\mathrm{i}}$-circular noises.

Next, in order to analyze the advantage of the QWL estimator in terms of the improperty of the QWGN, we computed the difference between the SL-QMSE and the WL-QMSE for the different classes of QWGN considered here. Thus, Fig. 4 compares these error differences (dB) for the $\mathbb{H}$-circular, $\mathbb{C}^{\mathrm{i}}$-circular and non-circular QWGN. Notice that the superiority of the QWL predictor is even higher in the case of nonlinear noises. Further, $\mathbb{C}^{\mathrm{j}}$-circular and $\mathbb{C}^{\mathrm{k}}$-circular QWGN have been also used in our simulations and similar results with $\mathbb{C}^{\mathrm{i}}$-circular QWGN have been obtained.

5.4. Example 4. Proposed algorithm vs. Kalman filter

Let x_t be a WSS quaternion signal whose augmented quaternion vector $\boldsymbol{x}_t = [x_t, x_t^*, x_t^{\mathrm{i}}, x_t^{\mathrm{j}}]^{\mathrm{T}}$ is given by the state-space model

$$\boldsymbol{x}_n = \boldsymbol{F}\boldsymbol{x}_{n-1} + \boldsymbol{u}_n$$
$$\boldsymbol{q}_n = \boldsymbol{x}_n + \boldsymbol{v}_n \tag{19}$$

where $\boldsymbol{F} = \mathrm{diag}(0.5,0.5,0.5,0.5)$, the variance of the augmented quaternion signal at the initial state $\boldsymbol{P}_0 = \langle \boldsymbol{x}_0, \boldsymbol{x}_0^{\mathrm{H}} \rangle_{\mathcal{Q}}$ is

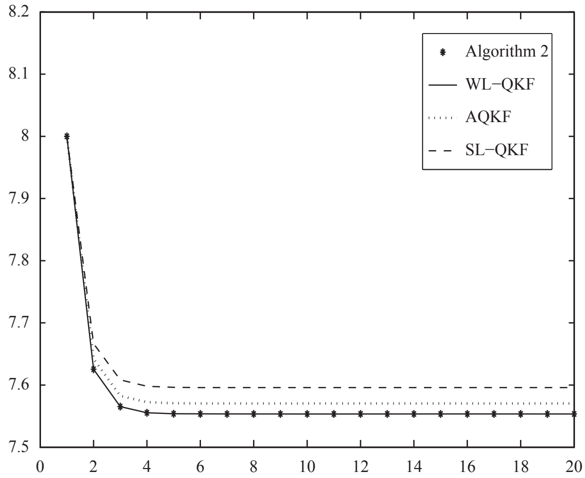

Fig. 5. QMSE for the one-stage prediction of x_t from Algorithm 2 (stars), WL-QKF (solid line), AQKF (dotted line), and SL-QKF (dashed line).

In this example, we aim to compare our prediction algorithm for WSS quaternion signals with solutions provided by other standard methods such as the Kalman filter. Specifically, the one-stage prediction of the signal $\hat{x}^{WL}(n+1/n)$, for $n=1,\ldots,20$, was performed on the basis of the augmented quaternion observations $\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_{20}$, by means of Algorithm 2 as well as the Kalman filter applied on the state-space model (19), called WL-QKF [12]. Additionally, we have also applied the classical complex Kalman filter [10] on the set of augmented observations $\{\boldsymbol{q}_n^c = [q_n, q_n^*]^{\mathrm{T}},\ n=1,\ldots,20\}$, that we name AQKF by similitude with Example 1, and the traditional Kalman filter (SL-QKF) on the quaternion observations $q_1, q_2, \ldots, q_{20}$. The obtained results are shown in Fig. 5, showing that when a signal can be modeled in state-space, the proposed algorithm coincides with the WL-QKF and outperforms both AQKF and SL-QKF. Therefore, the proposed solution is a viable alternative prediction approach, useful when the quaternion signal of interest is WSS and its state-space model is not readily accessible.

$$\boldsymbol{P}_0 = \begin{bmatrix} 8 & -4+2\mathrm{i}+\mathrm{j}+0.4\mathrm{k} & 0.6\mathrm{j}+1.2\mathrm{k} & 2.2\mathrm{i}-0.4\mathrm{k} \\ -4-2\mathrm{i}-\mathrm{j}-0.4\mathrm{k} & 8 & 4-2\mathrm{i}+0.4\mathrm{j}-0.8\mathrm{k} & 4-0.2\mathrm{i}-\mathrm{j}+0.8\mathrm{k} \\ -0.6\mathrm{j}-1.2\mathrm{k} & 4+2\mathrm{i}-0.4\mathrm{j}+0.8\mathrm{k} & 8 & 1.8\mathrm{i}-1.4\mathrm{j} \\ -2.2\mathrm{i}+0.4\mathrm{k} & 4+0.2\mathrm{i}+\mathrm{j}-0.8\mathrm{k} & -1.8\mathrm{i}+1.4\mathrm{j} & 8 \end{bmatrix}$$

Moreover, $\boldsymbol{u}_n$ and $\boldsymbol{v}_n$ are white noises uncorrelated with the signal and whose respective variances $\boldsymbol{Q} = \langle \boldsymbol{u}_n, \boldsymbol{u}_n^{\mathrm{H}} \rangle_{\mathcal{Q}}$ and $\boldsymbol{R} = \langle \boldsymbol{v}_n, \boldsymbol{v}_n^{\mathrm{H}} \rangle_{\mathcal{Q}}$ are

$$\boldsymbol{Q} = \begin{bmatrix} 6 & -3+1.5\mathrm{i}+0.75\mathrm{j}+0.9\mathrm{k} & 0.45\mathrm{j}+0.9\mathrm{k} & 1.65\mathrm{i}-0.3\mathrm{k} \\ -3-1.5\mathrm{i}-0.75\mathrm{j}-0.3\mathrm{k} & 6 & 3-1.5\mathrm{i}+0.3\mathrm{j}-0.6\mathrm{k} & 3-0.15\mathrm{i}-0.75\mathrm{j}+0.6\mathrm{k} \\ -0.45\mathrm{j}-0.9\mathrm{k} & 3+1.5\mathrm{i}-0.3\mathrm{j}+0.6\mathrm{k} & 6 & 1.35\mathrm{i}-1.05\mathrm{j} \\ -1.65\mathrm{i}+0.3\mathrm{k} & 3+0.15\mathrm{i}+0.75\mathrm{j}-0.6\mathrm{k} & -1.35\mathrm{i}+1.05\mathrm{j} & 6 \end{bmatrix}$$

$$\boldsymbol{R} = \begin{bmatrix} 40 & -20+0.04\mathrm{i}+0.08\mathrm{j}+0.02\mathrm{k} & 0.04\mathrm{j}+0.08\mathrm{k} & 0.06\mathrm{i}-0.04\mathrm{k} \\ -20-0.04\mathrm{i}-0.08\mathrm{j}-0.02\mathrm{k} & 40 & 20-0.04\mathrm{i}+0.04\mathrm{j}-0.06\mathrm{k} & 20-0.02\mathrm{i}-0.08\mathrm{j}+0.06\mathrm{k} \\ -0.04\mathrm{j}-0.08\mathrm{k} & 20+0.04\mathrm{i}-0.04\mathrm{j}+0.06\mathrm{k} & 40 & 0.02\mathrm{i}-0.12\mathrm{j} \\ -0.06\mathrm{i}+0.04\mathrm{k} & 20+0.02\mathrm{i}+0.08\mathrm{j}-0.06\mathrm{k} & -0.02\mathrm{i}+0.12\mathrm{j} & 40 \end{bmatrix}$$

with a SNR around -7 dB.

References

[1] H. Ahn, C. Won, Fast alignment using rotation vector and adaptive Kalman filter, IEEE Transactions on Aerospace and Electronic Systems 42 (1) (2006) 70–83.

[2] P. Bas, N. Le Bihan, J.M. Chassery, Color watermarking using quaternion Fourier transform, in: ICASSP, Hong-Kong, 2003.

[3] P.J. Brockwell, R.A. Davis, Time Series: Theory and Methods, 2nd ed. Springer-Verlag, New York, 1991.

[4] C. Cheong Took, D.P. Mandic, The quaternion LMS algorithm for adaptive filtering of hypercomplex real world processes, IEEE Transactions on Signal Processing 57 (4) (2009) 1316–1327.

[5] C. Cheong Took, D.P. Mandic, A quaternion widely linear adaptive filter, IEEE Transactions on Signal Processing 58 (8) (2010) 4427–4431.

[6] C. Cheong Took, D.P. Mandic, Augmented second-order statistics of quaternion random signals, Signal Processing 91 (2) (2011) 214–224.

[7] B. Che-Ujang, C. Cheong Took, D.P. Mandic, Quaternion-valued nonlinear adaptive filtering, IEEE Transactions on Neural Networks 22 (8) (2011) 1193–1206.

[8] J.C.K. Chou, Quaternion kinematic and dynamic differential equations, IEEE Transactions on Robotics and Automation 8 (1) (1992) 53–64.

[9] D. Choukkroun, I.Y. Bar-Itzhack, Y. Oshman, Novel quaternion Kalman filter, IEEE Transactions on Aerospace and Electronics Systems 42 (1) (2006) 174–190.

[10] D. Dini, D.P. Mandic, Class of widely linear complex Kalman filters, IEEE Transactions on Neural Networks and Learning Systems 23 (5) (2012) 775–786.

[11] T.A. Ell, S.J. Sangwine, Hypercomplex Fourier transforms of color images, IEEE Transactions on Image Processing 16 (1) (2007) 22–35.

[12] C. Jahanchahi, D.P. Mandic, A class of quaternion Kalman filters, IEEE Transactions on Neural Networks and Learning Systems, 2012, in press.

[13] S. Javidi, C. Cheong Took, D.P. Mandic, A fast independent component analysis algorithm for quaternion valued signals, IEEE Transactions on Neural Networks 22 (12) (2004) 1967–1978.

[14] N. Le Bihan, J. Mars, Singular value decomposition of quaternion matrices: a new tool for vector-sensor signal processing, Signal Processing 84 (7) (2004) 1177–1199.

[15] S. Miron, N. Le Bihan, J. Mars, Quaternion-MUSIC for vector-sensor array processing, IEEE Transactions on Signal Processing 54 (4) (2006) 1218–1229.

[16] A.M. Sabatini, Quaternion-based extended Kalman filter for determining orientation by inertial and magnetic sensing, IEEE Transactions on Biomedical Engineering 53 (7) (2006) 1346–1356.

[17] J. Shragge, Quaternion-based signal processing, in: Annual Meeting of Society of Exploration Geophysicists, 2006.

[18] J. Via, N. Le Bihan, D.P. Mandic, S.J. Sangwine, Quaternion signal processing: theory and applications, in: Tutorial in EUSIPCO, Barcelona, 2011.