www.ietdl.org

Published in IET Signal Processing
Received on 7th April 2011
Revised on 21st December 2011
doi: 10.1049/iet-spr.2011.0128

Special Issue on Multi-Sensor Signal Processing for Defence: Detection, Localisation & Classification

ISSN 1751-9675

Kalman filtering for widely linear complex and quaternion valued bearings only tracking

D.H. Dini   C. Jahanchahi   D.P. Mandic

Department of Electrical and Electronic Engineering, Imperial College London, Exhibition Road, London SW7 2BT, UK
E-mail: dahir.dini@imperial.ac.uk

Abstract: Bearings only target tracking is concerned with estimating the trajectory of an object from noise-corrupted bearing (phase) measurements. Traditionally this problem has been formulated as real valued for the Cartesian coordinate system or modified polar coordinate system. In this study, the authors introduce the bearings only tracking problem for the complex and quaternion domains to take advantage of the natural representation offered by these domains, for multivariate real signals, as well as the greater insights provided into the dynamics of tracking. Moreover, the authors introduce the augmented complex and quaternion extended Kalman filters for the modelling of second-order non-circular complex and quaternion valued signals, for which a widely linear model is shown to be more suitable than a strictly linear model.

1 Introduction

Bearings only tracking (BOT) paradigm is encountered in a variety of practical applications, including submarine tracking by passive sonar or aircraft surveillance by a radar in passive mode. The objective is to estimate online the kinematics (position and velocity) of a moving target using observer line of sight bearing (phase) measurements corrupted by noise. As the range measurements are not available and the bearings are not linearly related to the target state, the problem is hence inherently non-linear. Since a single static sensor is not able to track targets using bearings measurements only (because of the lack of range measurements [1, 2]), in order to estimate the range, the sensor has to manoeuvre. For two or more sensors, observability problem is not an issue, as the multiple bearing measurements can be used to form a range estimate.

Bearings only target motion analysis is generally carried out in either two dimensions for ocean environment or three dimensions for passive radar tracking [3, 4]. In both cases, the problem has been formulated as real valued in Cartesian or modified polar coordinate systems. However, the phase is inherently related to the complex (quaternion) domain, we aim to represent the two-dimensional (2D) and three-dimensional (3D) BOT scenarios more naturally as complex and quaternion valued problems.

The convenience of representation offered by the complex and quaternion domains is yet to be fully exploited by the research community within the context of bearings only target motion analysis. This is because the second-order statistics of complex and quaternion signals are not simple extensions of their real valued counterparts, and as such, conventional complex and quaternion valued statistical signal processing algorithms, such as the complex extended Kalman filter, are suboptimal for non-circular data and will generally not utilise all the available statistical information. In addition, the non-linear function relating the target position to the bearing measurement is non-holomorphic (non-analytic) in the Cauchy–Riemann and Cauchy–Riemann–Fueter sense, hence restricting the use of the extended Kalman filter (EKF) which uses the Taylor series expansion of the observation non-linearities.

The BOT problem is normally addressed from a state space point of view, lending itself to the use of the EKF [5]. In this work, we propose the use of the augmented complex and quaternion EKF in the context of BOT and prove their isomorphism (duality) with the real valued bivariate and quadvariate EKFs, thus enabling the use of the complex and quaternion domains as more convenient and natural alternatives to the real domain [6].

We first address 2D BOT using the widely linear (augmented) complex valued EKF [7]. Owing to recent advances in the so-called augmented complex statistics, for a second-order non-circular complex signal $z = z_r + jz_i$, that is, a signal with a rotation-dependent probability density function, the standard linear estimation model based on the covariance $\boldsymbol{R}_z = E\{zz^H\}$, where $E\{\cdot\}$ is the statistical expectation operator, is inadequate and a second moment function known as the pseudocovariance $\boldsymbol{P}_z = E\{zz^T\}$ is also required to fully capture the second-order statistics. This is carried out by use of the widely linear model, where z and its conjugate z^* are combined to form the augmented input vector $z^a = [z^T, z^H]^T$. The issue of non-analyticity of the non-linear observation function is addressed through the so-called $\mathbb{CR}$ calculus, which exploits the duality between augmented complex vectors and real valued vectors to allow for the derivatives of non-analytic functions to be computed.

For tracking objects in a 3D space, we propose to use a quaternion representation. To this end, we first introduce the quaternion Taylor series and derive the quaternion Kalman filter using the recently proposed $\mathbb{HR}$ calculus framework [8], and show that the quaternion Kalman filter has the same generic form as its complex counterpart. Similarly to the complex case, to utilise full second-order information, we employ the widely linear model to cater for non-circular quaternion valued signals. Simulations on 2 and 3D BOT support the analysis.

2 Background

2.1 Second order statistics of complex signals and widely linear modelling

The second-order statistical properties of a zero mean complex vector $\boldsymbol{z} = x + jy \in \mathbb{C}^N$ is normally charcaterised by its covariance matrix $\boldsymbol{R}_z = E\{\boldsymbol{z}\boldsymbol{z}^H\}$, where $E\{\cdot\}$ is the statistical expectation operator. In practice, this only holds for the special class of complex signals known as second-order circular or proper, that is, those with rotation invariant probability distributions [9].

Most real world processes are non-circular, either because of the different signal powers in the real and imaginary parts, or due to correlation between the real and imaginary parts [10]. Therefore for second-order non-circular (improper) signals we also need to include the pseudocovariance matrix $\boldsymbol{P}_z = E\{\boldsymbol{z}\boldsymbol{z}^{\mathrm{T}}\}$, in order to capture the full second-order information. For circular signals, $\boldsymbol{P}_z = 0$, while for non-circular signals, $\boldsymbol{P}_z \neq 0$, that is, the pseudocovariance cannot be ignored as it contains crucial information [9]. To cater for the generality of complex signals, we employ the widely linear model, which is linear in both the input $\boldsymbol{z}$ and its conjugate $\boldsymbol{z}^*$ [11], that is

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{z} + \boldsymbol{G}\boldsymbol{z}^* \tag{1}$$

where $\boldsymbol{y}$ is the output, $\boldsymbol{H}$ and $\boldsymbol{G}$ are complex coefficient matrices and $\boldsymbol{z}^a = [\boldsymbol{z}^{\mathrm{T}}, \boldsymbol{z}^{H}]^{\mathrm{T}}$ is the augmented or widely linear input. For $\boldsymbol{G} = 0$, (1) becomes the standard strictly linear model. The full second-order information is now contained in the augmented covariance matrix

$$\boldsymbol{R}_z^a = E\{\boldsymbol{z}^a\boldsymbol{z}^{a\mathrm{H}}\} = \begin{bmatrix} \boldsymbol{R}_z & \boldsymbol{P}_z \\ \boldsymbol{P}_z^* & \boldsymbol{R}_z^* \end{bmatrix} \tag{2}$$

which encompasses both $\boldsymbol{P}_z$ and $\boldsymbol{R}_z$.

2.2 BOT in $\mathbb{R}$

Conventionally, although the notion of phase is intimately related with the complex (quaternion) number system, the BOT problem has been formulated with a real valued state space representation. In order to estimate the trajectory of a target at every discrete time instant k, that is, its position (x_k, y_k) and velocity $(\dot{x}_k, \dot{y}_k)$, for a system with L observers located at $(x^{\mathrm{o}}_{i,k}, y^{\mathrm{o}}_{i,k})$, $i = 1, 2, \ldots, L$, the real valued state space model is given by

$$\boldsymbol{x}_k^r = \boldsymbol{F}^r\boldsymbol{x}_{k-1}^r + \boldsymbol{K}^r\boldsymbol{w}_k^r \tag{3a}$$

$$\boldsymbol{z}_k^r = \boldsymbol{h}^r[\boldsymbol{x}_k^r] + \boldsymbol{v}_k^r \tag{3b}$$

where the variables are defined as follows:

- $\boldsymbol{x}_k^r = \begin{bmatrix} x_k & \dot{x}_k & y_k & \dot{y}_k \end{bmatrix}^{\mathrm{T}}$ is the target state vector at time k,
- $\boldsymbol{F}^r$ and $\boldsymbol{K}^r$ are matrices defined as

$$\boldsymbol{F}^r = \begin{bmatrix} 1 & T & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad \boldsymbol{K}^r = \begin{bmatrix} \frac{T^2}{2} & 0 \\ T & 0 \\ 0 & \frac{T^2}{2} \\ 0 & T \end{bmatrix}$$

where T is the sampling interval.

- $\boldsymbol{z}_k^r$ is the observation vector and the vector $\boldsymbol{h}^r[\boldsymbol{x}_k^r] = \begin{bmatrix} \beta_{1,k} & \beta_{2,k} & \cdots & \beta_{L,k} \end{bmatrix}^{\mathrm{T}}$ contains the true target bearings with respect to the different sensors, where $\beta_{i,k} = \tan^{-1}(y_k - y^{\mathrm{o}}_{i,k}/x_k - x^{\mathrm{o}}_{i,k})$ is the true bearing at sensor i; the observation function $\boldsymbol{h}^r[\cdot]$ can be deduced from the given expression.
- $\boldsymbol{w}_k^r = \begin{bmatrix} \ddot{x}_k & \ddot{y}_k \end{bmatrix}^{\mathrm{T}}$ is the zero-mean state noise vector (the unknown target acceleration is modelled as noise) with covariance $\boldsymbol{R}_{w,k}^r$ and $\boldsymbol{v}_k = \begin{bmatrix} v_{1,k} & v_{2,k} & \cdots & v_{L,k} \end{bmatrix}^{\mathrm{T}}$ is the zero-mean observation noise vector with covariance $\boldsymbol{R}_{v,k}^r$.

The state space in (3) can be implemented using the real valued EKF to form online estimates of the target state. The EKF approximates the non-linearities in the state space by a first-order Taylor series expansion, in order to form linearised state and observation models. In the BOT problem the state equation is linear, however, the observation equation is non-linear and its first-order linearisation about the state estimate $\hat{\boldsymbol{x}}_{k|k-1}^r$ is given by

$$\begin{aligned} \boldsymbol{z}_k^r &= \boldsymbol{h}^r[\hat{\boldsymbol{x}}_{k|k-1}^r] + \boldsymbol{H}_k^r(\boldsymbol{x}_k^r - \hat{\boldsymbol{x}}_{k|k-1}^r) + \boldsymbol{v}_k^r \\ &= \boldsymbol{H}_k^r\boldsymbol{x}_k^r + \boldsymbol{v}_k^r + \boldsymbol{d}_k^r \end{aligned} \tag{4}$$

where $\boldsymbol{H}_k^r = (\partial\boldsymbol{h}^r/\partial\boldsymbol{x}_k^r)|_{\boldsymbol{x}_k^r = \hat{\boldsymbol{x}}_{k|k-1}^r}$ and $\boldsymbol{d}_k^r = \boldsymbol{h}^r[\hat{\boldsymbol{x}}_{k|k-1}^r] - \boldsymbol{H}_k^r\hat{\boldsymbol{x}}_{k|k-1}^r$ is a deterministic input to the model. The above model depends on the point about which the linearisation is carried out and this point should be as close as possible to the true state $\boldsymbol{x}_k^r$ in order to have a good approximation. The complete implementation of the BOT problem using the real valued EKF is shown in Algorithm 1.

Algorithm 1: Real valued EKF
Initialise with

$$\hat{\boldsymbol{x}}_{0|0}^r = E\{\boldsymbol{x}_0^r\} \quad \boldsymbol{M}_{0|0}^r = E\{\boldsymbol{x}_0^r\boldsymbol{x}_0^{r\mathrm{T}}\}$$

Model output

$$\hat{\boldsymbol{x}}_{k|k-1}^r = \boldsymbol{F}^r\hat{\boldsymbol{x}}_{k-1|k-1}^r \tag{5}$$

$$\boldsymbol{M}_{k|k-1}^r = \boldsymbol{F}^r\boldsymbol{M}_{k-1|k-1}^r\boldsymbol{F}^{r\mathrm{T}} + \boldsymbol{K}^r\boldsymbol{R}_{w,k}^r\boldsymbol{K}^{r\mathrm{T}} \tag{6}$$

Measurement output

$$\boldsymbol{G}_k^r = \boldsymbol{M}_{k|k-1}^r\boldsymbol{H}_k^{r\mathrm{T}}(\boldsymbol{H}_k^r\boldsymbol{M}_{k|k-1}^r\boldsymbol{H}_k^{r\mathrm{T}} + \boldsymbol{R}_{v,k}^r)^{-1} \tag{7}$$

$$\hat{\boldsymbol{x}}_{k|k}^r = \hat{\boldsymbol{x}}_{k|k-1}^r + \boldsymbol{G}_k^r(\boldsymbol{z}_k^r - \boldsymbol{h}^r[\hat{\boldsymbol{x}}_{k|k-1}^r]) \tag{8}$$

$$\boldsymbol{M}_{k|k}^r = (\boldsymbol{I} - \boldsymbol{G}_k^r\boldsymbol{H}_k^r)\boldsymbol{M}_{k|k-1}^r \tag{9}$$

The algorithm for 3D BOT in $\mathbb{R}^3$ follows directly from its

doi: 10.1049/iet-spr.2011.0128

2D counterpart, with the following differences accounting for the higher dimensionality:

- The 3D state vector becomes

$$\boldsymbol{x}_k^r = [\boldsymbol{x}_k \quad \dot{\boldsymbol{x}}_k \quad \boldsymbol{y}_k \quad \dot{\boldsymbol{y}}_k \quad \boldsymbol{z}_k \quad \dot{\boldsymbol{z}}_k]^{\mathrm{T}}$$

- The state transition matrix becomes

$$\boldsymbol{F}^r = \begin{bmatrix} 1 & T & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & T & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & T \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

- The state noise matrix becomes

$$\boldsymbol{K}^r = \begin{bmatrix} \frac{T^2}{2} & 0 & 0 \\ T & 0 & 0 \\ 0 & \frac{T^2}{2} & 0 \\ 0 & T & 0 \\ 0 & 0 & \frac{T^2}{2} \\ 0 & 0 & T \end{bmatrix}$$

- The observation function $\boldsymbol{h}^r[\boldsymbol{x}_k^r]$ contains the true azimuth angle θ and elevation ϕ angle with respect to the different sensors, that is

$$\boldsymbol{h}^r[x_k^r] = [\theta_1, \ldots, \theta_L, \phi_1, \ldots, \phi_L]^{\mathrm{T}}$$

where

$$\theta_i = \tan^{-1}\frac{y_k - y_i^0}{x_k - x_i^0}$$

$$\phi_i = \tan^{-1}\frac{z_k - z_i^0}{((x_k - x_i^0)^2 + (y_k - y_i^0)^2)^{1/2}}$$

- The state noise vector becomes

$$w_k^r = [\ddot{x}_k \quad \ddot{y}_k \quad \ddot{z}_k]^{\mathrm{T}}$$

The geometry of a system, where a target is tracked by a single sensor, is shown in Fig. 1.

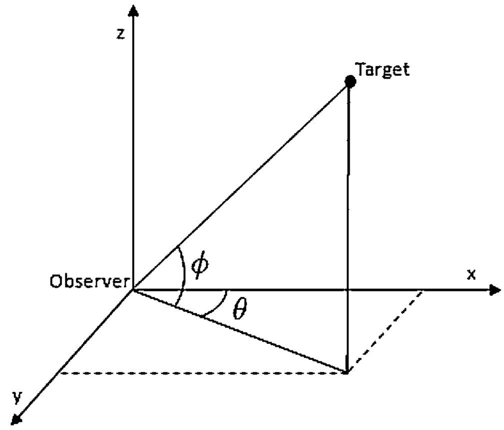

Fig. 1 *Single observer BOT system*

2.3 Quaternion algebra

Quaternions are an extension of complex numbers (forming an ordered pair) and comprise a real part (denoted by a subscript a) and three imaginary parts (denoted by subscripts b, c and d). A quaternion variable $q \in \mathbb{H}$ can be described as

$$q = q_a + iq_b + jq_c + kq_d$$

The unit vectors i, j and k in the quaternion domain $\mathbb{H}$ are also imaginary units, and obey the following rules

$$ij = k \quad jk = i \quad ki = j$$

$$i^2 = j^2 = k^2 = ijk = -1$$

Note that quaternion multiplication is not commutative, that is, $ij \neq ji = -k$.

A quaternion variable q can be conveniently written as [12]

$$q = Sq + Vq$$

where $Sq = q_a$ (denotes the scalar part of q) and $Vq = iq_b + jq_c + kq_d$ (denotes the vector part of q). The quaternion product can then be expressed as

$$\begin{aligned} q_1q_2 &= (Sq_1 + Vq_1)(Sq_2 + Sq_2) \\ &= Sq_1Sq_2 - Vq_1 \cdot Vq_2 + Sq_2Vq_1 + Sq_1Vq_2 + Vq_1 \times Vq_2 \end{aligned}$$

where the symbol '$\cdot$' denotes the dot-product and '$\times$' the cross-product in vector analysis. The quaternion conjugate, denoted by q^* is given by

$$q^* = Sq - Vq$$

the norm $\|q\|$ of a quaternion variable q is defined as

$$||q|| = \sqrt{qq^*} = \sqrt{q_a^2 + q_b^2 + q_c^2 + q_d^2}$$

The 3D vector part Vq is also called a pure quaternion, whereas the inclusion of the real part Sq gives a full quaternion. The unique algebraic structure of quaternions enables unified processing of 3D and 4D multivariate processes under one umbrella.

The key to widely linear modelling in the quaternion domains are the involutions, or self-inverse mappings defined as (Note that the quaternion conjugate is also an involution, that is $(q_1q_2)^* = (q_2^*q_1^*)$.) [13]

$$q^i = -iqi = q_a + iq_b - jq_c - kq_d$$

$$q^j = -jqj = q_a - iq_b + jq_c - kq_d$$

$$q^k = -kqk = q_a - iq_b - jq_c + kq_d$$

To show that involutions are self-inverse mappings, consider for instance $(q^i)^i = q$. The quaternion involutions have the following properties:

P1: $(q^{\text{inv}})^{\text{inv}} = q$ for inv $= i, j, k$
P2: $(q_1 q_2)^{\text{inv}} = q_1^{\text{inv}} q_2^{\text{inv}}$ for inv $= i, j, k$
P3: $(q^i)^j = (q^j)^i = q^k$

It is important to realise that involutions can be seen as a counterpart of the complex conjugate, as they allow the components of a quaternion variable to be expressed in terms of the actual variable and its involutions, that is,

$$
\begin{aligned}
q_a &= \frac{1}{4}[q + q^i + q^j + q^k] \quad q_b = \frac{1}{4i}[q + q^i - q^j - q^k] \\
q_c &= \frac{1}{4j}[q - q^i + q^j - q^k] \quad q_d = \frac{1}{4k}[q - q^i - q^j + q^k]
\end{aligned}
\tag{10}
$$

similar to the complex domain where x and y can be written as

$$
x = \frac{1}{2}(z + z^*) \quad y = \frac{1}{2i}(z - z^*) \tag{11}
$$

2.4 $\mathbb{HR}$ calculus

The concept of $\mathbb{CR}$ calculus has been recently extended to the quaternion domain [8] and can be summarised as follows

$$
\begin{bmatrix} \dfrac{\partial f(q, q^i, q^j, q^k)}{\partial q} \\ \dfrac{\partial f(q, q^i, q^j, q^k)}{\partial q^i} \\ \dfrac{\partial f(q, q^i, q^j, q^k)}{\partial q^j} \\ \dfrac{\partial f(q, q^i, q^j, q^k)}{\partial q^\kappa} \end{bmatrix} = \begin{bmatrix} 1 & -i & -j & -k \\ 1 & -i & j & k \\ 1 & i & -j & k \\ 1 & i & j & -k \end{bmatrix} \times \begin{bmatrix} \dfrac{\partial f(q_a, q_b, q_c, q_d)}{\partial q_a} \\ \dfrac{\partial f(q_a, q_b, q_c, q_d)}{\partial q_b} \\ \dfrac{\partial f(q_a, q_b, q_c, q_d)}{\partial q_c} \\ \dfrac{\partial f(q_a, q_b, q_c, q_d)}{\partial q_d} \end{bmatrix}
$$

$$
\begin{bmatrix} \dfrac{\partial f(q^*, q^{i*}, q^{j*}, q^{k*})}{\partial q^*} \\ \dfrac{\partial f(q^*, q^{i*}, q^{j*}, q^{k*})}{\partial q^{i*}} \\ \dfrac{\partial f(q^*, q^{i*}, q^{j*}, q^{k*})}{\partial q^{j*}} \\ \dfrac{\partial f(q^*, q^{i*}, q^{j*}, q^{k*})}{\partial q^{k*}} \end{bmatrix} = \begin{bmatrix} 1 & i & j & k \\ 1 & i & -j & -k \\ 1 & -i & j & -k \\ 1 & -i & -j & k \end{bmatrix} \times \begin{bmatrix} \dfrac{\partial f(q_a, q_b, q_c, q_d)}{\partial q_a} \\ \dfrac{\partial f(q_a, q_b, q_c, q_d)}{\partial q_b} \\ \dfrac{\partial f(q_a, q_b, q_c, q_d)}{\partial q_c} \\ \dfrac{\partial f(q_a, q_b, q_c, q_d)}{\partial q_d} \end{bmatrix}
$$

with the first set of equations called the $\mathbb{HR}$ derivatives and the second called the $\mathbb{HR}^*$ derivatives. This allows us to perform the derivatives of both the holomorphic and non-holomorphic non-linear functions. The rigid CRF conditions for quaternion differentiability have pervented development of non-linear algorithms, and the $\mathbb{HR}$ calculus is a breakthrough.

2.5 Quaternion widely linear model

In the complex domain, non-circular signals have non-vanishing pseudocovariance $E[\boldsymbol{x}\boldsymbol{x}^{\mathrm{T}}]$ and for complete second-order modelling both the pseudocovariance $E[\boldsymbol{x}\boldsymbol{x}^{\mathrm{T}}]$ as well as the covariance matrix $E[\boldsymbol{x}\boldsymbol{x}^{\mathrm{H}}]$ are required. In adaptive filtering problems this translates to deriving algorithms based on the widely linear model.

Recently, widely linear modelling has been extended to the quaternion domain [14], where to entirely describe the second-order statistics of quaternion non-circular random variables, the additional complementary covariance matrices $E[\boldsymbol{q}\boldsymbol{q}^{i\mathrm{H}}]$, $E[\boldsymbol{q}\boldsymbol{q}^{j\mathrm{H}}]$ and $E[\boldsymbol{q}\boldsymbol{q}^{k\mathrm{H}}]$ are employed. Analogously to the complex case, for a non-circular process, the model should comprise the involution terms q, q^i, q^j and q^k, to fully capture the so-called augmented statistics. The augmented covariance matrix of the quaternion augmented vector $\boldsymbol{q}^a = [\boldsymbol{q}^{\mathrm{T}}, \boldsymbol{q}^{i\mathrm{T}}, \boldsymbol{q}^{j\mathrm{T}}, \boldsymbol{q}^{k\mathrm{T}}]^{\mathrm{T}}$ can now be written as

$$
\boldsymbol{R}_q^a = E[\boldsymbol{q}^a \boldsymbol{q}^{a\mathrm{H}}] \begin{bmatrix} \boldsymbol{R}_q & \boldsymbol{P}_{q^i} & \boldsymbol{P}_{q^j} & \boldsymbol{P}_{q^k} \\ \boldsymbol{P}_{q^i}^i & \boldsymbol{R}_q^i & \boldsymbol{P}_{q^k}^i & \boldsymbol{P}_{q^j}^i \\ \boldsymbol{P}_{q^j}^j & \boldsymbol{P}_{q^k}^j & \boldsymbol{R}_q^j & \boldsymbol{P}_{q^i}^j \\ \boldsymbol{P}_{q^k}^k & \boldsymbol{P}_{q^j}^k & \boldsymbol{P}_{q^i}^k & \boldsymbol{R}_q^k \end{bmatrix}
$$

where $\boldsymbol{R}_q = E[\boldsymbol{q}\boldsymbol{q}^{\mathrm{H}}]$, $\boldsymbol{P}_{q^i} = E[\boldsymbol{q}\boldsymbol{q}^{i\mathrm{H}}]$, $\boldsymbol{P}_{q^j} = E[\boldsymbol{q}\boldsymbol{q}^{j\mathrm{H}}]$, $\boldsymbol{P}_{q^k} = E[\boldsymbol{q}\boldsymbol{q}^{k\mathrm{H}}]$.

2.6 Quaternion Taylor series

We next introduce quaternion valued Taylor series, where development has been hampered by the stringent CRF consitions. The use of the $\mathbb{HR}$ calculus allows us, for the first time, to write a generic expression for quaternion TSE. This is achieved by starting from the Taylor series of a real quadrivariate function $u(q_a, q_b, q_c, q_d)$ given in [15]

$$
\begin{aligned}
u(q_a, q_b, q_c, q_d) &= u(q_{a0}, q_{b0}, q_{c0}, q_{d0}) + \sum_{n=1}^{\infty} \frac{1}{n} \Bigg[\Delta q_a \frac{\partial}{\partial q_a} \\
&\quad + \Delta q_b \frac{\partial}{\partial q_b} + \Delta q_c \frac{\partial}{\partial q_c} + \Delta q_d \frac{\partial}{\partial q_d} + \Bigg]^n \\
&\quad \times u(q_{a0}, q_{b0}, q_{c0}, q_{d0})
\end{aligned}
\tag{12}
$$

where $\Delta q_a = q_a - q_{a0}$, $\Delta q_b = q_b - q_{b0}$, $\Delta q_c = q_c - q_{c0}$ and $\Delta q_d = q_d - q_{d0}$. By using the identity $(\partial f / \partial \boldsymbol{w}^{\mathrm{T}}) = (\partial f / \partial \boldsymbol{v}^{\mathrm{T}})\boldsymbol{A}$, this can be expanded as

$$
\begin{aligned}
&\begin{bmatrix} \dfrac{\partial f}{\partial q_a} & \dfrac{\partial f}{\partial q_b} & \dfrac{\partial f}{\partial q_c} & \dfrac{\partial f}{\partial q_d} \end{bmatrix} \\
&= \begin{bmatrix} \dfrac{\partial f}{\partial q} & \dfrac{\partial f}{\partial q^i} & \dfrac{\partial f}{\partial q^j} & \dfrac{\partial f}{\partial q^k} \end{bmatrix} \begin{bmatrix} 1 & i & j & k \\ 1 & i & -j & -k \\ 1 & -i & j & -k \\ 1 & -i & -j & k \end{bmatrix}
\end{aligned}
\tag{13}
$$

IET Signal Process., 2012, Vol. 6, Iss. 5, pp. 435–445
doi: 10.1049/iet-spr.2011.0128

For a real function of quaternion variables, we have $f(q, q^i, q^j, q^k) = u(q_a, q_b, q_c, q_d)$ and so substituting (13) into (12) and using the identities

$$\Delta q_a = \frac{1}{4}(\Delta q + \Delta q^i + \Delta q^j + \Delta q^k)$$

$$\Delta q_b = \frac{1}{4i}(\Delta q + \Delta q^i - \Delta q^j - \Delta q^k)$$

$$\Delta q_c = \frac{1}{4j}(\Delta q - \Delta q^i + \Delta q^j - \Delta q^k)$$

$$\Delta q_d = \frac{1}{4k}(\Delta q - \Delta q^i - \Delta q^j + \Delta q^k)$$

the quaternion Taylor series can be obtained as

$$f(q, q^i, q^j, q^k) = f(q_0, q_0^i, q_0^j, q_0^k) + \sum_{n=1}^{\infty} \frac{1}{n}\left[\Delta q \frac{\partial}{\partial q} + \Delta q^i \frac{\partial}{\partial q^i} + \Delta q^j \frac{\partial}{\partial q^j} + \Delta q^k \frac{\partial}{\partial q^k} + \right]^n f(q_0, q_0^i, q_0^j, q_0^k)$$

For a multivariate quaternion function, the Taylor series becomes

$$f(\boldsymbol{q}, \boldsymbol{q}^i, \boldsymbol{q}^j, \boldsymbol{q}^k) = f(\boldsymbol{q}_0, \boldsymbol{q}_0^i, \boldsymbol{q}_0^j, \boldsymbol{q}_0^k) + \sum_{n=1}^{\infty} \frac{1}{n}\left[\Delta \boldsymbol{q}^{\mathrm{T}} \frac{\partial}{\partial \boldsymbol{q}} + \Delta \boldsymbol{q}^{i\mathrm{T}} \frac{\partial}{\partial \boldsymbol{q}^i} + \Delta \boldsymbol{q}^{j\mathrm{T}} \frac{\partial}{\partial \boldsymbol{q}^j} + \Delta \boldsymbol{q}^{k\mathrm{T}} \frac{\partial}{\partial \boldsymbol{q}^k} + \right]^n f(\boldsymbol{q}_0, \boldsymbol{q}_0^i, \boldsymbol{q}_0^j, \boldsymbol{q}_0^k) \tag{14}$$

and will serve as a basis for the development of extended quaternion valued filters.

3 BOT in $\mathbb{C}$

3.1 Augmented complex extended Kalman filter (ACEKF)

The 2D BOT problem can be conveniently represented in the complex domain, where the complex BOT state space can be defined as

$$\boldsymbol{x}_k = \boldsymbol{F}\boldsymbol{x}_{k-1} + \boldsymbol{K}w_k \tag{15a}$$

$$\boldsymbol{z}_k = \boldsymbol{h}[\boldsymbol{x}_k] + \boldsymbol{v}_k \tag{15b}$$

and the model variables are described as follows:

- $\boldsymbol{x}_k = \begin{bmatrix} x_k + jy_k & \dot{x}_k + j\dot{y}_k \end{bmatrix}^{\mathrm{T}}$ is the complex target state vector
- $\boldsymbol{F}$ and $\boldsymbol{K}$ are matrices defined as

$$\boldsymbol{F} = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad \boldsymbol{K} = \begin{bmatrix} \frac{T^2}{2} \\ T \end{bmatrix}$$

- $\boldsymbol{z}_k$ is the complex observation vector and $\boldsymbol{h}[\boldsymbol{x}_k]$ is a complex vector function defined as

$$\boldsymbol{h}[\boldsymbol{x}_k] = \begin{bmatrix} \beta_{1,k} + j\beta_{(L/2)+1,k} \\ \beta_{2,k} + j\beta_{(L/2)+2,k} \\ \vdots \\ \beta_{(L/2),k} + j\beta_{L,k} \end{bmatrix}$$

- $w_k = \ddot{x}_k + j\ddot{y}_k$ is the zero mean state noise with variance $r_{w,k}$ and pseudovariance $p_{w,k}$, while

$$\boldsymbol{v}_k = \begin{bmatrix} v_{1,k} + jv_{L/2+1,k} & \cdots & v_{L/2,k} + jv_{L,k} \end{bmatrix}^{\mathrm{T}}$$

is the zero mean observation noise with covariance $\boldsymbol{R}_{v,k}$ and pseudocovariance $\boldsymbol{P}_{v,k}$.

The function $\boldsymbol{h}[\boldsymbol{x}_k]$ is complex valued and it is straightforward to show that this function does not satisfy the Cauchy–Riemann conditions, that is, $(\partial \boldsymbol{h}[\boldsymbol{x}_k]/\partial \boldsymbol{x}_k^*) \neq 0$ and is hence non-holomorphic. However, by utilising the $\mathbb{CR}$ calculus [16] framework, the Taylor series expansions of this non-holomorphic function is possible, and its first-order TSE about $\hat{\boldsymbol{x}}_{k|k-1}$ is given by

$$\boldsymbol{h}[\boldsymbol{x}_k] \simeq \boldsymbol{h}[\hat{\boldsymbol{x}}_{k|k-1}] + \boldsymbol{A}_k(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1}) + \boldsymbol{B}_k(\boldsymbol{x}_k^* - \hat{\boldsymbol{x}}_{k|k-1}^*) \tag{16}$$

where the matrices $\boldsymbol{A}_k$ and $\boldsymbol{B}_k$ are the Jacobians defined as

$$\boldsymbol{A}_k = \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{x}_k}\Big|_{\boldsymbol{x}_k=\hat{\boldsymbol{x}}_{k|k-1}} \quad \text{and} \quad \boldsymbol{B}_k = \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{x}_k^*}\Big|_{\boldsymbol{x}_k^*=\hat{\boldsymbol{x}}_{k|k-1}^*} \tag{17}$$

The observation equation can then be written as

$$\boldsymbol{z}_k = \boldsymbol{A}_k\boldsymbol{x}_k + \boldsymbol{B}_k\boldsymbol{x}_k^* + \boldsymbol{v}_k + \boldsymbol{d}_k \tag{18}$$

where $\boldsymbol{d}_k = \boldsymbol{h}[\hat{\boldsymbol{x}}_{k|k-1}] - \boldsymbol{A}_k\hat{\boldsymbol{x}}_{k|k-1} - \boldsymbol{B}_k\hat{\boldsymbol{x}}_{k|k-1}^*$. From (18) note that the observation equation is a widely linear function of the state, being a function of the state $\boldsymbol{x}_k$ and its conjugate $\boldsymbol{x}_k^*$. This is because the first-order TSE of non-holomorphic functions are widely linear, since the derivative with respect to the complex conjugate does not vanish, that is $\boldsymbol{B}_k \neq 0$ in (16).

The transformation $\boldsymbol{x}_k \to \boldsymbol{x}_k^*$ is a non-linear mapping, and as such, the first-order Taylor series approximation in (18) remains a non-linear function of the state $\boldsymbol{x}_k$. In order to overcome this non-linear relationship between the state and the observation, the state space needs to be redefined so that the augmented state vector $\boldsymbol{x}_k^a$ is utilised instead of the conventional state vector $\boldsymbol{x}_k$. However, by using an augmented state vector $\boldsymbol{x}_k^a$, the observation equation becomes a linear function of the state.

Therefore in order to cater for the covariances and pseudocovariances of the state and observation noises, both the state and observation equations have to be augmented, and can be implemented using the ACEKF. Earlier versions of the augmented complex Kalman filter (ACKF) and ACEKF were introduced for the training of recurrent neural networks [17, 18], where the gradient was calculated using the augmented RTRL [19]. The ACEKF derived here is different, as it is more general, and is able to cater for the full second-order statistics of both the state and observation noises. The fully augmented complex state space for BOT is defined as

$$x_k^a = F^a x_{k-1}^a + K^a w_k^a \tag{19a}$$

$$z_k^a = h^a[x_k^a] + v_k^a \tag{19b}$$

where

$$x_k^a = \begin{bmatrix} x_k \\ x_k^* \end{bmatrix}, \quad w_k^a = \begin{bmatrix} w_k \\ w_k^* \end{bmatrix}, \quad F^a = \begin{bmatrix} F & 0 \\ 0 & F^* \end{bmatrix},$$

$$K^a = \begin{bmatrix} K & 0 \\ 0 & K^* \end{bmatrix}, \quad z_k^a = \begin{bmatrix} z_k \\ z_k^* \end{bmatrix},$$

$$v_k^a = \begin{bmatrix} v_k \\ v_k^* \end{bmatrix} \quad \text{and} \quad h^a[x_k^a] = \begin{bmatrix} h[x_k] \\ h^*[x_k] \end{bmatrix}$$

and the linearised augmented state space can be expressed as

$$x_k^a = F^a x_{k-1}^a + K^a w_k^a$$

$$z_k^a = H_k^a x_k^a + v_k^a + d_k^a$$

where $H_k^a = (\partial h^a / \partial x_k^a) = \begin{bmatrix} A_k & B_k \\ B_k^* & A_k^* \end{bmatrix}$, $d_k^a = \begin{bmatrix} d_k \\ d_k^* \end{bmatrix}$ and $d_k = h[\hat{x}_{k|k-1}] - A_k \hat{x}_{k|k-1} - B_k \hat{x}_{k|k-1}^*$.

The ACEKF, which utilises the augmented state space, is summarised in Algorithm 2.

Algorithm 2: Augmented Complex EKF (ACEKF)
Initialise with

$$\hat{x}_{0|0}^a = E\{x_0^a\} \quad M_{0|0}^a = E\{x_0^a x_0^{aH}\}$$

Model output

$$\hat{x}_{k|k-1}^a = F^a \hat{x}_{k-1|k-1}^a \tag{20}$$

$$M_{k|k-1}^a = F^a M_{k-1|k-1}^a F^{aH} + K^a R_{w,k}^a K^{aH} \tag{21}$$

Measurement output

$$G_k^a = M_{k|k-1}^a H_k^{aH} (H_k^a M_{k|k-1}^a H_k^{aH} + R_{v,k}^a)^{-1} \tag{22}$$

$$\hat{x}_{k|k}^a = \hat{x}_{k|k-1}^a + G_k^a (z_k^a - h^a[\hat{x}_{k|k-1}^a]) \tag{23}$$

$$M_{k|k}^a = (I - G_k^a H_k^a) M_{k|k-1}^a \tag{24}$$

The ACEKF utilises the augmented state and observation noise covariance matrices, that is

$$R_{w,k}^a = E\{w_k^a w_k^{aH}\} = \begin{bmatrix} r_{w,k} & p_{w,k} \\ p_{w,k}^* & r_{w,k}^* \end{bmatrix} \tag{25}$$

$$R_{v,k}^a = E\{v_k^a v_k^{aH}\} = \begin{bmatrix} R_{v,k} & P_{v,k} \\ P_{v,k}^* & R_{v,k}^* \end{bmatrix} \tag{26}$$

where $p_{w,k}$ and $P_{v,k}$ are the state and observation noises, hence, ACEKF caters for the full second-order statistics.

4 Duality between the ACEKF and real valued EKF

Owing to the duality between the augmented complex space $\mathbb{C}^{2N}$ and the real space $\mathbb{R}^{2N}$, the ACEKF and the real valued EKF are equivalent and give the same estimate at every time instant. For any complex vector $u = u_r + ju_i$ it holds that

$$\underbrace{\begin{bmatrix} u \\ u^* \end{bmatrix}}_{u^a} = \underbrace{\begin{bmatrix} I & jI \\ I & -jI \end{bmatrix}}_{J} \underbrace{\begin{bmatrix} u_{\mathrm{r}} \\ u_{\mathrm{i}} \end{bmatrix}}_{u^r}$$

where I is the identity matrix and $J^{-1} = (1/2)J^H$. By utilising this isomorphism between the augmented and real spaces, the augmented complex state space (19) can be transformed into the corresponding real valued state space (3). By premultiplying (19) with J^{-1} and using $x_k^r = J^{-1} x_k^a$, $z_k^r = J^{-1} z_k^a$, $v_k^r = J^{-1} v_k^a$ and $w_k^r = J^{-1} w_k^a$, the duality between the two state spaces can be observed through

$$\begin{aligned} J^{-1} x_k^a &= J^{-1} F^a x_{k-1}^a + J^{-1} K^a w_k^a \\ &= J^{-1} F^a J x_{k-1}^r + J^{-1} K^a J w_k^r \\ &= F^r x_{k-1}^r + K^r w_k^r \\ J^{-1} z_k^a &= J^{-1} h^a[x_k^a] + J^{-1} v_k^a \\ &= h^r[x_k^r] + v_k^r \end{aligned} \tag{27}$$

where $F^r = J^{-1} F^a J$, $K^r = J^{-1} K^a J$ and $h^r[x_k^r] = J^{-1} h^a[x_k^a]$. The covariance matrices of the real valued state and observation noises, w_k^r and v_k^r, are given by

$$R_{w^r,k} = E\{w_k^r w_k^{r\mathrm{T}}\} = J^{-1} R_{w^a,k} J^{-\mathrm{H}}$$

$$R_{v^r,k} = E\{v_k^r v_k^{r\mathrm{T}}\} = J^{-1} R_{v^a,k} J^{-\mathrm{H}}$$

It can be shown that the real valued EKF has the exact same performance as the ACEKF for every time instant. Assume the ACEKF is initiated at time instant $(k-1)$, with initial state $\hat{x}_{k-1|k-1}^a$ and covariance matrix $M_{k-1|k-1}^a$. The real valued EKF is then initialised as follows

$$\hat{x}_{k-1|k-1}^r = J^{-1} \hat{x}_{k-1|k-1}^a \tag{28}$$

$$M_{k-1|k-1}^r = J^{-1} M_{k-1|k-1}^a J^{-\mathrm{H}} \tag{29}$$

It is now straightforward to show that the predicted states and covariance matrices of the two filters are related as

$$\hat{x}_{k|k-1}^r = J^{-1} \hat{x}_{k|k-1}^a \tag{30}$$

$$M_{k|k-1}^r = J^{-1} M_{k|k-1}^a J^{-\mathrm{H}} \tag{31}$$

The relationship between the real and augmented complex Jacobians H_k^r and H_k^a can be shown to be $H_{k-1}^r = J^{-1} H_k^a J$, while the Kalman gains in (22) and (7) are related as

$$\begin{aligned} G_k^a &= M_{k|k-1}^a H_k^{a\mathrm{H}} [H_k^a M_{k|k-1}^a H_k^{a\mathrm{H}} + R_{v^a,k}]^{-1} \\ &= J M_{k|k-1}^r J^{\mathrm{H}} J^{-\mathrm{H}} H_k^{r\mathrm{H}} J^{\mathrm{H}} \\ &\quad \times [J H_k^r J^{-1} J M_{k|k-1}^r J^{\mathrm{H}} J^{-\mathrm{H}} H_k^{r\mathrm{H}} J^{\mathrm{H}} + J R_{v^r,k} J^{\mathrm{H}}]^{-1} \\ &= J M_{k|k-1}^r H_k^{r\mathrm{H}} [H_k^r M_{k|k-1}^r H_k^{r\mathrm{H}} + R_{v^r,k}]^{-1} J^{-1} \\ &= J G_k^r J^{-1} \end{aligned} \tag{32}$$

doi: 10.1049/iet-spr.2011.0128

Consequently, the state estimates $\hat{\boldsymbol{x}}^a_{k|k}$ and $\hat{\boldsymbol{x}}^r_{k|k}$ have the following relationship

$$\begin{aligned}\hat{\boldsymbol{x}}^r_{k|k} &= \hat{\boldsymbol{x}}^r_{k|k-1} + \boldsymbol{G}^r_k(\boldsymbol{z}^r_k - \boldsymbol{h}^r[\hat{\boldsymbol{x}}^r_{k|k-1}]) \\ &= \boldsymbol{J}^{-1}\hat{\boldsymbol{x}}^a_{k|k-1} + \boldsymbol{J}^{-1}\boldsymbol{G}^a_k\boldsymbol{J}(\boldsymbol{y}^r_k - \boldsymbol{h}^r[\hat{\boldsymbol{x}}^r_{k|k-1}]) \\ &= \boldsymbol{J}^{-1}\hat{\boldsymbol{x}}^a_{k|k} \end{aligned} \tag{33}$$

and the covariance matrices are related as

$$\boldsymbol{M}^r_{k|k} = \boldsymbol{J}^{-1}\boldsymbol{M}^a_{k|k}\boldsymbol{J}^{-\mathrm{H}} \tag{34}$$

From (33), observe that the state estimates $\hat{\boldsymbol{x}}^a_{k|k}$ and $\hat{\boldsymbol{x}}^r_{k|k}$ are equivalent and are related by an invertible linear mapping. To show that ACKF and its dual real valued Kalman filter achieve the same mean square error (MSE), recall that the MSE for the real valued bivariate Kalman filter is given by (The use of the term MSE is a misnomer since the ACEKF and real valued EKF are only approximations to the MSE estimator, except for the case of linear state space models with uncorrelated Gaussian state and observation noises.)

$$\varepsilon^r_k = \mathrm{tr}\{\boldsymbol{M}^r_{k|k}\} \tag{35}$$

where the symbol tr{·} denotes the matrix trace operator. Similarly, the MSE corresponding to the augmented MSE matrix $\boldsymbol{M}^a_{n|n}$ is given by the trace of (34), that

$$\begin{aligned}\mathrm{Tr}\{\boldsymbol{M}^a_{k|k}\} &= \mathrm{tr}\{\boldsymbol{J}\boldsymbol{M}^r_{k|k}\boldsymbol{J}^{\mathrm{H}}\} \\ &= \mathrm{tr}\{\boldsymbol{M}^r_{k|k}\boldsymbol{J}^{\mathrm{H}}\boldsymbol{J}\} \\ &= 2\cdot\mathrm{tr}\{\boldsymbol{M}^r_{k|k}\}\end{aligned} \tag{36}$$

where the expression $\boldsymbol{J}^{\mathrm{H}} = 2\boldsymbol{J}^{-1}$ was utilised. At first, this result is misleading as it suggests that ACEKF achieves twice the error of its dual real valued KF. However, this is because the error term is counted twice by the trace of $\boldsymbol{M}^a_{k|k}$, owing to the block diagonal structure of the augmented MSE covariance matrix, and hence needs to be halved to express the true augmented MSE, that is

$$\varepsilon^a_k = \frac{1}{2}\mathrm{tr}\{\boldsymbol{M}^a_{k|k}\} = \varepsilon^r_k$$

Therefore the ACEKF and the its dual bivariate real valued KF are equivalent forms of the same state space models. They achieve the same state estimates and MSE at every time instant, regardless of the circularity of the processed signals.

5 3D BOT in $\mathbb{H}$

We next introduce the quaternion counterparts of the proposed widely linear complex Kalman filters for BOT.

5.1 Quaternion EKF

The state space representation for the quaternion EKF takes the same form as the quaternion Kalman filter (see the Appendix) and can be written as

$$\boldsymbol{x}_k = \boldsymbol{F}\boldsymbol{x}_{k-1} + \boldsymbol{K}w_k \tag{37}$$

$$z_k = \boldsymbol{h}[\boldsymbol{x}_k] + \boldsymbol{v}_k \tag{38}$$

where for 3D BOT, we define

- $\boldsymbol{x}_k = [ix_k + jy_k + kz_k \quad i\dot{x}_k + j\dot{y}_k + k\dot{z}_k]$
- the state transition matrix $\boldsymbol{F}$ and matrix $\boldsymbol{K}$ are defined as

$$\boldsymbol{F} = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}$$

$$\boldsymbol{K} = \begin{bmatrix} \frac{T^2}{2} & 0 \\ T & 0 \end{bmatrix}$$

- the function $\boldsymbol{h}[\boldsymbol{x}_k]$ is defined as

$$\boldsymbol{h}[\boldsymbol{x}_k] = \begin{bmatrix} \theta_1 + i\theta_{(L/2)+1} + j\phi_1 + k\phi_{(L/2)+1} \\ \theta_2 + i\theta_{(L/2)+2} + j\phi_2 + k\phi_{(L/2)+2} \\ \cdots \\ \theta_{L/2} + i\theta_L + j\phi_{L/2} + k\phi_L \end{bmatrix}$$

where

$$\theta_i = \tan^{-1}\frac{y_k - y_i^0}{x_k - x_i^0}$$

$$\phi_i = \tan^{-1}\frac{z_k - z_i^0}{((x_k - x_i^0)^2 + (y_k - y_i^0)^2)^{1/2}}$$

- $w_k = ix_k + jy_k + kz_k$ is the zero mean state noise.

Using the quaternion Taylor series derived in Section 2.6, in conjunction with the $\mathbb{HR}$ calculus, the function $\boldsymbol{h}[\boldsymbol{x}_k]$ can be linearised as

$$\begin{aligned}\boldsymbol{h}[\boldsymbol{x}_k] \simeq{}& \boldsymbol{h}[\hat{\boldsymbol{x}}_{k|k-1}] + \boldsymbol{A}_k(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1}) + \boldsymbol{B}_k(\boldsymbol{x}^i_k - \hat{\boldsymbol{x}}^i_{k|k-1}) \\ &+ \boldsymbol{C}_k(\boldsymbol{x}^j_k - \hat{\boldsymbol{x}}^j_{k|k-1}) + \boldsymbol{D}_k(\boldsymbol{x}^k_k - \hat{\boldsymbol{x}}^k_{k|k-1})\end{aligned} \tag{39}$$

where the matrices $\boldsymbol{A}_k$, $\boldsymbol{B}_k$, $\boldsymbol{C}_k$, $\boldsymbol{D}_k$ are defined as

$$\begin{aligned}\boldsymbol{A}_k &= \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{x}_k}, \quad \boldsymbol{B}_k = \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{x}^i_k} \\ \boldsymbol{C}_k &= \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{x}^j_k}, \quad \boldsymbol{D}_k = \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{x}^k_k}\end{aligned} \tag{40}$$

The observation equation can now be written as

$$\begin{aligned}z_k ={}& \boldsymbol{A}_k\boldsymbol{x}_k + \boldsymbol{B}_k\boldsymbol{x}^i_k + \boldsymbol{C}_k\boldsymbol{x}^j_k + \boldsymbol{D}_k\boldsymbol{x}^k_k + (\boldsymbol{h}[\hat{\boldsymbol{x}}_{k|k-1}] - \boldsymbol{A}_k\hat{\boldsymbol{x}}_{k|k-1} \\ &- \boldsymbol{B}_k\hat{\boldsymbol{x}}^i_{k|k-1} - \boldsymbol{C}_k\hat{\boldsymbol{x}}^j_{k|k-1} - \boldsymbol{D}_k\hat{\boldsymbol{x}}^k_{k|k-1})\end{aligned} \tag{41}$$

Similarly to the complex case, the observation equation is widely linear in $\boldsymbol{x}_k$, the state space model in (38) must be modified as

$$x_k^a = F^a x_{k-1}^a + K^a w_k^a \quad (42)$$

$$z_k = H_k x_k^a + v_k + d_k \quad (43)$$

where

$$x_k^a = [x_k \quad x_k^i \quad x_k^j \quad x_k^k]^T$$

$$w_k^a = [w_k \quad w_k^i \quad w_k^j \quad w_k^k]^T$$

$$F^a = \begin{bmatrix} F & 0 & 0 & 0 \\ 0 & F^i & 0 & 0 \\ 0 & 0 & F^j & 0 \\ 0 & 0 & 0 & F^k \end{bmatrix}$$

$$K^a = \begin{bmatrix} K & 0 & 0 & 0 \\ 0 & K^i & 0 & 0 \\ 0 & 0 & K^j & 0 \\ 0 & 0 & 0 & K^k \end{bmatrix}$$

$$H_k = [A_k \quad B_k \quad C_k \quad D_k]$$

$$d_k = h[\hat{x}_{k|k-1}] - A_k\hat{x}_{k|k-1} - B_k\hat{x}^i_{k|k-1} - C_k\hat{x}^j_{k|k-1} - D_k\hat{x}^k_{k|k-1}$$

5.2 Augmented quaternion extended Kalman filter (AQEKF)

Similar to the complex case, the state space for the quaternion BOT becomes

$$x_k^a = F^a x_{k-1}^a + K^a w_k^a \quad (44)$$

$$z_k^a = h^a[x_k^a] + v_k^a \quad (45)$$

where

$$z_k^a = [z_k \quad z_k^i \quad z_k^j \quad z_k^k]$$

$$v_k^a = [v_k \quad v_k^i \quad v_k^j \quad v_k^k]$$

$$h^a[x_k] = [h[x_k] \quad h^i[x_k] \quad h^j[x_k] \quad h_k^k[x_k]]$$

Using the Taylor series to linearise the non-linear function, the state model can be re-written as

$$x_k = F^a x_{k-1}^a + K^a w_k^a \quad (46)$$

$$z_k^a = H_k^a x_k^a + v_k^a + d_k^a \quad (47)$$

where

$$H_k^a = \begin{bmatrix} A_k & B_k & C_k & D_k \\ B_k^i & A_k^i & D_k^i & C_k^i \\ C_k^j & D_k^j & A_k^j & B_k^j \\ D_k^k & C_k^k & B_k^k & A_k^k \end{bmatrix}$$

The definitions above can be applied to the algorithm in (21)–(24) to obtain the widely linear quaternion BOT. The duality between the real and quaternion valued 3D BOT follows from the fact that the quaternion variable $\{q, q^i, q^j, q^k\}$ and the real variable $\{q_a, q_b, q_c, q_d\}$ are linearly related, as

$$\begin{bmatrix} q \\ q^i \\ q^j \\ q^k \end{bmatrix} = \begin{bmatrix} 1 & i & j & k \\ 1 & i & -j & -k \\ 1 & -i & j & -k \\ 1 & -i & -j & k \end{bmatrix} \begin{bmatrix} q_a \\ q_b \\ q_c \\ q_d \end{bmatrix} \quad (48)$$

This isomorphism between the augmented quaternion and real space is similar to that found between the complex and real space and can be used to show that the AQEKF and its dual real valued EKF are essentially the same filter.

6 Simulation examples

6.1 Simulation results for a 2D tracking scenario

To illustrate the potential of ACEKF within the bearings only target motion analysis context, we first consider a scenario with two static sensors located at (−1200, 1300) and (1000, 1500). The system described by (3.1) was simulated with a sampling interval of $T = 1$ s to generate 300 samples. The target was initially located at (200, 100) and was moving with an initial velocity (2, 1). The true target state vector was assumed to be unknown at initialisation, and as such the ACEKF algorithm were initialised with position (300, 300) and velocity (4, 4).

The system was simulated using second-order circular and non-circular state and observation noises, and the results are illustrated in Fig. 2. In the first set of experiments, the system was simulated using second-order circular white Gaussian noise processes w_k and v_k with distributions

$$w_k \sim \mathcal{N}(0, 2.5 \times 10^{-2}, 0)$$

$$v_k \sim \mathcal{N}(0, 5 \times 10^{-7}, 0)$$

where the last term in the distribution represents the pseudocovariance of the noise. Fig. 2*a* shows that the ACEKF converged, and that it was able to estimate the target velocity more accurately than its position. This is because the target experiences larger variations in its location (driven by the state noise process) than in velocity.

The next set of experiments were for a non-circular state noise and a circular observation noise, that is

$$w_k \sim \mathcal{N}(0, 2.5 \times 10^{-2}, 2.3 \times 10^{-2})$$

$$v_k \sim \mathcal{N}(0, 5 \times 10^{-7}, 0)$$

The results are shown in Fig. 2*b*, where again ACEKF was able to capture the target dynamics, as well as achieving a lower steady state error compared to the case with circular noises.

6.2 Simulation results for a 3D tracking scenario

For the 3D BOT, the two static sensors were located at (−1200,1300,0) and (1000,1500,100) while the target was initially located at (200,100,300) and was moving with an initial velocity (2,1,0.5). The initial target position and velocity were again assumed unknown and the Kalman filter was initiated with location (300,300,400) and velocity (4,4,4). The sampling interval used was the same as that in the 2D case. Similarly to the complex case, the system was first simulated using second-order circular white Gaussian

doi: 10.1049/iet-spr.2011.0128

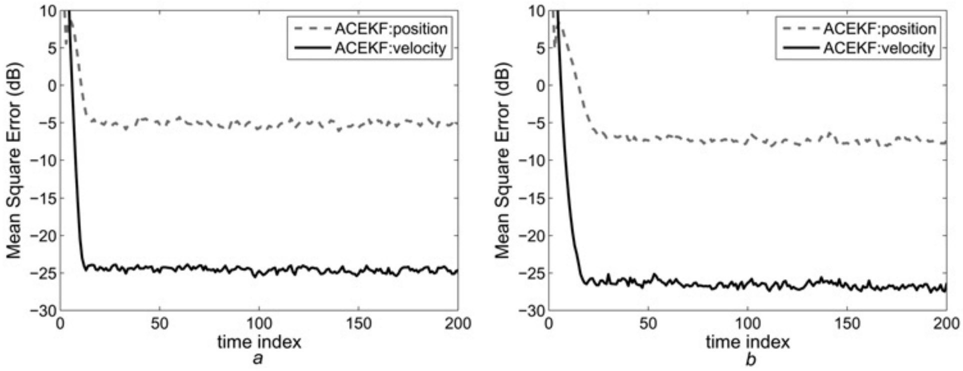

Fig. 2 *Performance of ACEKF for a 2D tracking scenario and system with second-order circular and non-circular state and observation noises*

a Circular state and observation noises
b Non-circular state noise and a circular observation noise

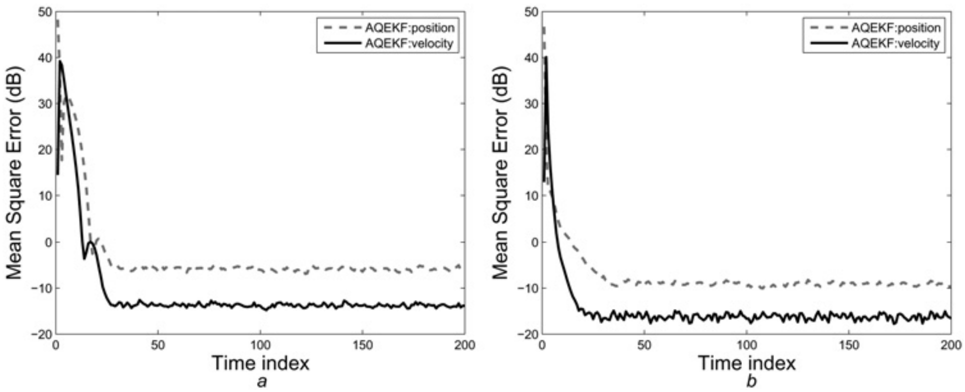

Fig. 3 *Performance of AQEKF for a system with second-order circular and non-circular state and observation noises*

a Circular state and observation noises
b Non-circular state noise and a circular observation noise

noise $\boldsymbol{w}_k$ and $\boldsymbol{v}_k$ with distributions

$$\boldsymbol{w}_n \sim \mathcal{N}(0, 2.5 \times 10^{-2}, 0, 0, 0)$$

$$\boldsymbol{v}_n \sim \mathcal{N}(0, 5 \times 10^{-7}, 0, 0, 0)$$

where the last three terms in the distribution represent the pseudovariance of the noise (respectively, $\boldsymbol{P}_q^i$, $\boldsymbol{P}_q^j$, $\boldsymbol{P}_q^k$). The MSE performance is shown in Fig. 3*a*. As expected, the AQEKF converged and was able to track the underlying dynamics of the target position and velocity. The same experiment is repeated but this time the state noise is made non-circular, that is,

$$\boldsymbol{w}_n \sim \mathcal{N}(0, 2.5 \times 10^{-2}, -2.5 \times 10^{-2}, -2.5 \times 10^{-2}, 2.5 \times 10^{-2})$$

$$\boldsymbol{v}_n \sim \mathcal{N}(0, 10^{-7}, 0, 0, 0)$$

Similarly to the case with circular state noise, the AQEKF converged, but to a lower steady state MSE. Fig. 4 shows the performance of AQEKF from a geometric view point. A geometric view of the converging performance of AQEKF is shown in Fig. 4, where the algorithm was initialised with the wrong target location and velocity.

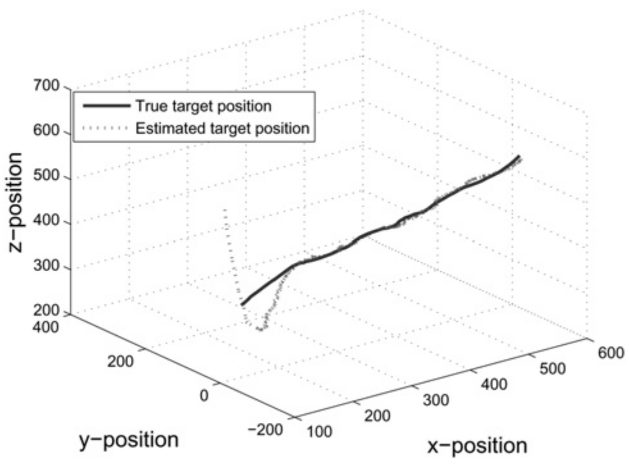

Fig. 4 *3D target tracking in the presence of non-circular noise*

7 Conclusion

The complex and quaternion domains are the natural representations of 2 and 3D motion trajectories as they are naturally suited to signals with amplitude and direction (angle). To this effect, the 2D and 3D BOT problem has been revisited, and the duality between the real valued EKF and the augmented complex and quaternion EKFs, ACEKF and AQEKF, has been illuminated. The analysis has shown that the observation model is non-holomorphic in the

Cauchy-Riemann sense, and by using the recently developed $\mathbb{CR}$ and $\mathbb{HR}$ calculus, it has been shown that its Taylor series approximation is a widely linear function of the state, hence lending itself to the ACEKF and AQEKF algorithms, which cater for the second-order statistics of the complex and quaternion signals, through their use of widely linear models. Simulation on circular and non-circular bearing only tracking problems in both 2D and 3D support the analysis, where both the complex and quaternion augmented EKFs have been shown to be able to accurately track target dynamics.

8 Acknowledgment

This work was part of the University Defence Research Centre (UDRC) at Imperial College London, supported by the MoD and DSTL.

9 References

1 Song, T.L.: 'Observability of target tracking with bearings-only measurements', *IEEE Trans. Aerosp. Electron. Syst.*, 1996, **32**, pp. 1468–1472
2 Nardone, S., Aidala, V.: 'Observability criteria for bearings-only target motion analysis', *IEEE Trans. Aerosp. Electron. Syst.*, 1981, **AES-17**, (2), pp. 162–166
3 Le Cadre, J.-P., Tremois, O.: 'Bearings-only tracking for maneuvering sources', *IEEE Trans. Aerosp. Electron. Syst.*, 1998, **34**, pp. 179–193
4 Nardone, S., Graham, M.: 'A closed-form solution to bearings-only target motion analysis', *IEEE J. Ocean. Eng.*, 1997, **22**, pp. 168–178
5 Aidala, V.: 'Kalman filter behavior in bearings-only tracking applications', *IEEE Trans. Aerosp. Electron. Syst.*, 1979, **AES-15**, (1), pp. 29–39
6 Mandic, D., Javidi, S., Souretis, G., Goh, V.: 'Why a complex valued solution for a real domain problem'. IEEE Workshop on Machine Learning for Signal Processing 2007, 2007, pp. 384–389
7 Dini, D.H., Mandic, D.P.: 'Widely linear complex extended Kalman filters'. Sensor Signal Processing for Defence Conf., 2011
8 Mandic, D., Jahanchahi, C., Took, C.: 'A quaternion gradient operator and its applications', *IEEE Signal Proc. Lett.*, 2011, **18**, (1), pp. 47–50
9 Picinbono, B., Bondon, P.: 'Second-order statistics of complex signals', *IEEE Trans. Signal Process.*, 1997, **45**, (2), pp. 411–420
10 Mandic, D.P., Goh, V.S.L.: 'Complex valued nonlinear adaptive filters: noncircularity, widely linear and neural models' (Wiley, 2009)
11 Picinbono, B.: 'Second-order complex random vectors and normal distributions', *IEEE Trans. Signal Process.*, 1996, **44**, (10), pp. 2637–2640
12 Ward, J.: 'Quaternions and cayley numbers' (Kluwer Academic Publishers, 1997)
13 Ella, T.A., Sangwine, S.J.: 'Quaternion involutions and anti-involutions', *Comput. Math. Appl.*, 2007, **53**, pp. 137–143
14 Cheong Took, C., Mandic, D.P.: 'Augmented second-order statistics of quaternion random signals', *Signal Process.*, 2011, **91**, pp. 214–224
15 Bayin, S.: 'Mathematical methods in science and engineering' (John Wiley and Sons, 2006)
16 Kreutz-Delgado, K.: 'The complex gradient operator and the CR-calculus'. Tech. Report ECE275CG-F05v1.3d, Electrical and Computer Engineering, Jacobs School of Engineering, University of California, San Diego, 2006
17 Goh, S., Mandic, D.P.: 'An augmented extended Kalman filter algorithm for complex-valued recurrent neural networks', *Neural Comput.*, 2007, **19**, pp. 1039–1055
18 Dini, D.H., Mandic, D.P., Julier, S.J.: 'A widely linear complex unscented Kalman filter', *IEEE Signal Process. Lett.*, 2011, **18**, pp. 623–626
19 Goh, S.L., Mandic, D.: 'Stochastic gradient-adaptive complex-valued nonlinear neural adaptive filters with a gradient-adaptive step size', *IEEE Trans. Neural Netw.*, 2007, **18**, (5), pp. 1511–1516

10 Appendix: derivation of the quaternion Kalman filter

To perform 3D BOT in the quaternion domain, the quaternion augmented Kalman filter was first derived. Following the complex domain derivation, we assume that the quaternion state $\boldsymbol{x}_k \in \mathbb{H}^{n\times 1}$ evolves according to

$$\boldsymbol{x}_k = \boldsymbol{F}_k\boldsymbol{x}_{k-1} + \boldsymbol{B}_k\boldsymbol{u}_k + \boldsymbol{w}_k \tag{49}$$

where $\boldsymbol{F}_k \in \mathbb{H}^{n\times n}$ is the state transition matrix, $\boldsymbol{B}_k \in \mathbb{H}^{n\times n}$ is the control-input matrix for the control input $\boldsymbol{u}_k \in \mathbb{H}^{n\times 1}$ and $\boldsymbol{w}_k \in \mathbb{H}^{n\times 1}$ is the model noise.

The state $\boldsymbol{x}_k$ cannot be observed directly but can be inferred from measuring the quantity $\boldsymbol{z}_k \in \mathbb{H}^{m\times 1}$ which relates to the state $\boldsymbol{x}_k$ according to

$$\boldsymbol{z}_k = \boldsymbol{H}_k\boldsymbol{x}_k + \boldsymbol{v}_k \tag{50}$$

where $\boldsymbol{H}_k \in \mathbb{H}^{m\times n}$ is the observation matrix and $\boldsymbol{v}_k \in \mathbb{H}^{m\times 1}$ is the measurement noise. Both the model noise and measurement noise are zero mean and Gaussian, that is,

$$\boldsymbol{w}_k \sim N(0, \boldsymbol{R}_{w,k}) \tag{51}$$

$$\boldsymbol{v}_k \sim N(0, \boldsymbol{R}_{v,k}) \tag{52}$$

The a priori state estimate $\hat{\boldsymbol{x}}_{k|k-1} \in \mathbb{H}^{n\times 1}$ (the estimate of the state $\boldsymbol{x}_k$ before obtaining the new measurement) can be obtained from the state model

$$\hat{\boldsymbol{x}}_{k|k-1} = \boldsymbol{F}_k\hat{\boldsymbol{x}}_{k-1|k-1} + \boldsymbol{B}_k\boldsymbol{u}_k \tag{53}$$

where $\hat{\boldsymbol{x}}_{k-1|k-1}$ is the a posteriori state estimate (from the previous state estimate).

Using the measurement $\boldsymbol{z}_k$, the estimate $\hat{\boldsymbol{x}}_{k|k-1}$ can be improved to obtain the a posteriori state estimate $\hat{\boldsymbol{x}}_{k|k}$.

$$\hat{\boldsymbol{x}}_{k|k} = \hat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{G}_k(\boldsymbol{z}_k - \boldsymbol{H}_k\hat{\boldsymbol{x}}_{k|k-1}) \tag{54}$$

where $\boldsymbol{z}_k - \boldsymbol{H}_k\hat{\boldsymbol{x}}_{k|k-1}$ is the innovation factor (the error between the estimate and the measurement). Following the same approach as for the derivation of the complex Kalman filter, an expression for the Kalman gain $\boldsymbol{G}_k$ can be obtained, leading to the quaternion Kalman filter, summarised in Algorithm 3

Algorithm 3: Quaternion Kalman Filter (QKF)
Model output

$$\hat{\boldsymbol{x}}_{k|k-1} = \boldsymbol{F}_k\hat{\boldsymbol{x}}_{k-1|k-1} + \boldsymbol{B}_k\boldsymbol{u}_k \tag{55}$$

$$\boldsymbol{M}_{k|k-1} = \boldsymbol{F}_k\boldsymbol{M}_{k-1|k-1}\boldsymbol{F}_k^{\mathrm{H}} + \boldsymbol{R}_{w,k} \tag{56}$$

Measurement output

$$\boldsymbol{S}_k = \boldsymbol{H}_k\boldsymbol{M}_{k|k-1}\boldsymbol{H}_k^{\mathrm{H}} + \boldsymbol{R}_{v,k} \tag{57}$$

$$\boldsymbol{G}_k = \boldsymbol{M}_{k|k-1}\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{S}_k^{-1} \tag{58}$$

$$\hat{\boldsymbol{x}}_{k|k} = \hat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{G}_k[\boldsymbol{z}_k - \boldsymbol{H}_k\hat{\boldsymbol{x}}_{k|k-1}] \tag{59}$$

$$\boldsymbol{M}_{k|k} = [\boldsymbol{I} - \boldsymbol{G}_k\boldsymbol{H}_k]\boldsymbol{M}_{k|k-1} \tag{60}$$

Note that the expressions above have the same generic form as those in the complex domain.

To cater for quaternion non-circularity, the widely linear model can be incorporated into the quaternion Kalman

doi: 10.1049/iet-spr.2011.0128

filter, giving the following structure to the filter

$$\begin{aligned} \boldsymbol{x}_k^a &= \boldsymbol{F}_k^a \boldsymbol{x}_{k-1}^a + \boldsymbol{B}_k^a \boldsymbol{u}_k^a + \boldsymbol{w}_k^a \\ \boldsymbol{z}_k^a &= \boldsymbol{H}_k^a \boldsymbol{x}_{k-1}^a + \boldsymbol{v}_k^a \end{aligned} \tag{61}$$

where

$$\boldsymbol{x}_k^a = [\boldsymbol{x}_k^{\mathrm{T}} \quad \boldsymbol{x}_k^{i\mathrm{T}} \quad \boldsymbol{x}_k^{j\mathrm{T}} \quad \boldsymbol{x}_k^{k\mathrm{T}}]^{\mathrm{T}}$$

$$\boldsymbol{w}_k^a = [\boldsymbol{w}_k^{\mathrm{T}} \quad \boldsymbol{w}_k^{i\mathrm{T}} \quad \boldsymbol{w}_k^{j\mathrm{T}} \quad \boldsymbol{w}_k^{k\mathrm{T}}]^{\mathrm{T}}$$

$$\boldsymbol{v}_k^a = [\boldsymbol{v}_k^{\mathrm{T}} \quad \boldsymbol{v}_k^{i\mathrm{T}} \quad \boldsymbol{v}_k^{j\mathrm{T}} \quad \boldsymbol{v}_k^{k\mathrm{T}}]^{\mathrm{T}}$$

$$\boldsymbol{F}_k^a = \begin{bmatrix} \boldsymbol{F}_1 & \boldsymbol{F}_2 & \boldsymbol{F}_3 & \boldsymbol{F}_4 \\ \boldsymbol{F}_2^i & \boldsymbol{F}_1^i & \boldsymbol{F}_4^i & F_3^i \\ \boldsymbol{F}_3 & \boldsymbol{F}_4^j & \boldsymbol{F}_1^j & \boldsymbol{F}_2^j \\ \mathbf{A}_4^k & \boldsymbol{F}_3^k & \boldsymbol{F}_2^k & \boldsymbol{F}_4^k \end{bmatrix}$$

$$\boldsymbol{H}_k^a = \begin{bmatrix} \boldsymbol{H}_1 & \boldsymbol{H}_2 & \boldsymbol{H}_3 & \boldsymbol{H}_4 \\ \boldsymbol{H}_2^i & \boldsymbol{H}_1^i & \boldsymbol{H}_4^i & H_3^i \\ \boldsymbol{H}_3^j & \boldsymbol{H}_4^j & \boldsymbol{H}_1^j & \boldsymbol{H}_2^j \\ \boldsymbol{H}_4^k & \boldsymbol{H}_3^k & \boldsymbol{H}_2^k & \boldsymbol{H}_1^k \end{bmatrix}$$

Note that the structure in $\boldsymbol{H}_k^a$ and $\boldsymbol{F}_k^a$ (the subscript k is dropped in the partitioned matrix for ease of representation) is necessary for the model in (61) to hold. If the model is linear in $\boldsymbol{x}$, then the off-diagonal elements will vanish. The only difference is notational, where every term has a superscript a. The augmented quaternion Kalman filter expressions are summarised in Algorithm 4.

Algorithm 4: Augmented Quaternion Kalman Filter (AQKF)
Model output

$$\hat{\boldsymbol{x}}_{k|k-1}^a = \boldsymbol{F}_k^a \hat{\boldsymbol{x}}_{k-1|k-1}^a + \boldsymbol{B}_k^a \boldsymbol{u}_k^a \tag{62}$$

$$\boldsymbol{M}_{k|k-1}^a = \boldsymbol{F}_k^a \boldsymbol{M}_{k-1|k-1}^a \boldsymbol{F}_k^{a\mathrm{H}} + \boldsymbol{R}_{w,k}^a \tag{63}$$

Measurement output

$$\boldsymbol{S}_k^a = \boldsymbol{H}_k^a \boldsymbol{M}_{k|k-1}^a \boldsymbol{H}_k^{aH} + \boldsymbol{R}_v^a \tag{64}$$

$$\boldsymbol{G}_k^a = \boldsymbol{M}_{k|k-1}^a \boldsymbol{H}_k^{a\mathrm{H}} \boldsymbol{S}^{a^{-1}} \tag{65}$$

$$\hat{\boldsymbol{x}}_{k|k}^a = \hat{\boldsymbol{x}}_{k|k-1}^a + \boldsymbol{G}_k^a [\boldsymbol{z}_k^a - \boldsymbol{H}^a \hat{\boldsymbol{x}}_{k|k-1}^a] \tag{66}$$

$$\boldsymbol{M}_{k|k}^a = [\boldsymbol{I} - \boldsymbol{G}_k^a \boldsymbol{H}_k^a] \boldsymbol{M}_{k|k-1}^a \tag{67}$$