A Class of Widely Linear Complex Kalman Filters

Dahir Dini and **Danilo P. Mandic**
Communications and Signal Processing Research Group,
Department of Electrical and Electronic Engineering,
Imperial College London, London, SW7 2AZ, U.K.
email: {dahir.dini, d.mandic}@imperial.ac.uk
URL: http://www.commsp.ee.ic.ac.uk/~mandic

Technical Report: TR-IC-UDRC-WL-Kalman/0257-09-10-2
Presented at the University Defense Research Centre (UDRC) quaterly meeting,
http://mod-udrc.org/
Created: November 2009, *Updated*: September 2010, *Last Updated*: June 2011

Abstract

Recently, a class of widely linear (augmented) complex valued Kalman filters, which utilise augmented complex statistics, have been proposed. For the sequential state space estimation of the generality of complex signals in the contex of neural network training [1]. This has allowed for a unified treatment of both second order circular and noncircular signals, that is, both those with rotation invariant and dependent distributions. In this paper, we revist the augmented complex Kalman filter (ACKF), augmented complex extended Kalman filter (ACEKF) and augmented complex unscented Kalman filter (ACUKF), in a more general contex and analyse their performances for different degrees of state and measurement noise noncircularity. A theoretical bound for the performance of the class of widely linear (augmented) Kalman filters over their strictly linear counterparts is provided. The analysis also address the duality with bivariate real valued Kalman filters. Simulations using both synthetic and real world proper and improper signals support the analysis.

Index Terms

Widely linear model, complex circularity, complex Kalman filter, extended Kalman filter, unscented Kalman filter, augmented complex Kalman filter

I. Introduction

Complex valued signals arise in a variety of applications such as in communications systems, radar and AC power systems. In addition, a complex representation of bivariate real valued signals, such as 2-dimensional wind data [1], might be chosen to provide a convenient representation for these signals as well as a natural way of preserving the characteristics of the signals and the transformations they undergo, such as the phase and magnitude distortion.

The second order statistical properties of complex signals are characterised by their second order moment variance and pseudocovariance functions. The covariance captures the information concerning the total power of the signal, while the pseudocovariance encapsulates the information about the power difference and cross-correlation between the real and imaginary parts of the signal. Conventional complex valued signal processing algorithms have generally been designed, explicitly or implicitly, to cater for second order circular (proper) complex signals, that is signals with rotation invariant probability distributions. These are characterised by a vanishing pseudocovariance, which makes them inadequate for most real world signals which are almost invariably second order noncircular due to the different signal powers in the real and imaginary parts, correlation of the real and imaginary parts, or due to nonstationarity [1].

The Kalman filter is a state space based estimation technique that has many applications in technology, and is an essential part of space and military technology development. The Kalman filter and its different extensions are commonly used for both real and complex valued scenarios. Complex valued Kalman filters have been used extensively in a variety of applications, including frequency estimation of time-varying signals [2], training of neural networks [3] and wireless localization [4]. However, the traditional implementation of the complex valued Kalman filter inherently assumes second order circular state and measurement noises as well as input data, and as such does not fully utilise the full available second order statistics of the complex signals.

The recent introduction of so called 'augmented complex statistics' [5] [1] has highlighted that for a general (improper) complex vector $\mathbf{x}$, estimation based solely on the covariance matrix $\mathbf{R}_{\mathbf{x}} = E\{\mathbf{x}\mathbf{x}^H\}$ is inadequate, and the pseudocovariance matrix $\mathbf{P}_{\mathbf{x}} = E\{\mathbf{x}\mathbf{x}^T\}$ is also required to fully capture the full second order statistics. To introduce an optimal second order estimator for the generality of complex signals, consider first the mean square error (MSE) estimator of a real valued random vector $\mathbf{y}$ in terms of an observed real vector $\mathbf{x}$, that is, $\hat{\mathbf{y}} = E\{\mathbf{y}|\mathbf{x}\}$. For zero-mean, jointly normal $\mathbf{y}$ and $\mathbf{x}$, the optimal estimator is linear, that is

$$\hat{\mathbf{y}} = \mathbf{H}\mathbf{x} \tag{1}$$

where $\mathbf{H}$ is a coefficient matrix. Standard, 'strictly linear' estimation in $\mathbb{C}$ assumes the same model but with complex valued $\mathbf{y}, \mathbf{x}$, and $\mathbf{H}$. Based on the fact that both the real $\mathbf{y}_r$ and imaginary $\mathbf{y}_i$ parts of the vector $\mathbf{y}$ are real valued, and

$$\hat{\mathbf{y}}_r = E\{\mathbf{y}_r|\mathbf{x}_r, \mathbf{x}_i\} \qquad \hat{\mathbf{y}}_i = E\{\mathbf{y}_i|\mathbf{x}_r, \mathbf{x}_i\} \tag{2}$$

Substituting $\mathbf{x}_r = (\mathbf{x} + \mathbf{x}^*)/2$ and $\mathbf{x}_i = (\mathbf{x} - \mathbf{x}^*)/2\jmath$ yields

$$\hat{\mathbf{y}}_r = E\{\mathbf{y}_r|\mathbf{x}, \mathbf{x}^*\} \qquad \hat{\mathbf{y}}_i = E\{\mathbf{y}_i|\mathbf{x}, \mathbf{x}^*\} \tag{3}$$

and using (1), we obtain the *widely linear* complex estimator[1]

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{G}\mathbf{x}^* = \mathbf{W}\mathbf{x}^a \tag{4}$$

where the matrix $\mathbf{W}$ comprises the coefficient matrices $\mathbf{H}$ and $\mathbf{G}$, and $\mathbf{x}^a = [\mathbf{x}^T, \mathbf{x}^H]^T$ is the 'augmented' input vector. The full second order information is thus contained in the augmented covariance matrix

$$\mathbf{R}^a_{\mathbf{x}} = E\{\mathbf{x}^a\mathbf{x}^{aH}\} = \begin{bmatrix} \mathbf{R}_{\mathbf{x}} & \mathbf{P}_{\mathbf{x}} \\ \mathbf{P}^*_{\mathbf{x}} & \mathbf{R}^*_{\mathbf{x}} \end{bmatrix} \tag{5}$$

and as such, estimation based on $\mathbf{R}^a_{\mathbf{x}}$ incorporates both the covariance and the pseudocovariance matrices.

Recently, the widely linear (augmented) complex Kalman filter (ACKF) [3], the augmented complex extended Kalman filter (ACEKF) [3] and the augmented complex unscented Kalman filter (ACUKF) [1], which are suitable for the generality of complex signals both second order circular and noncircular, have been introduced, and applied for the training of neural networks and have been shown to have superior performance, when compared with their corresponding conventional complex Kalman filters, for second order noncircular signals. However, the performance of these filters were not elaborated for the general case where the sources of improperness included both the input data and system parameters. Moreover, the effect of signal noncircularity on the mean square behavior of the conventional complex Kalman filter (CCKF), the complex extended Kalman filter (CEKF) and the complex unscented Kalman filter (CUKF) still needs further attention.

In this paper, we consolidate our recent work on a class of widely linear Kalman filters [1] and illuminate their performances under general widely linear state and observation noises, and for nonholomorphic state and observation models. We revisit the recently introduced class of widely linear Kalman filters, the augmented complex Kalman filter (ACKF), the augmented complex extended Kalman filter (ACEKF) and the augmented complex unscented Kalman filter (ACUKF), and show that the ACKF is superior to the CCKF in the mean square error (MSE) sense for the generality of complex signals. We also illustrate that the computational complexity of the ACKF can be significantly reduced by exploiting the isomorphism between the bivariate real and complex domains. A more general form of the ACEKF, is then introduced, which is able to cater to both analytic and nonanalytic state space models, in the Cauchy-Riemann sense. The effect of noncircular state and observation noises on the MSE behavior of the CEKF and CUKF are also analysed.

II. The Augmented Complex Kalman Filter (ACKF)

The Kalman filter is an optimal sequential state estimator for linear dynamical systems, in the sense that it achieves the minimum mean squared error (MMSE). Consider the conventional state space model given by [6]

$$\mathbf{x}_n = \mathbf{F}_{n-1}\mathbf{x}_{n-1} + \mathbf{w}_n \tag{6}$$

$$\mathbf{y}_n = \mathbf{H}_n\mathbf{x}_n + \mathbf{v}_n \tag{7}$$

where $\mathbf{x}_n$ is the state to be estimated (of dimension $p \times 1$), $\mathbf{y}_n$ is the noisy observation (of dimension $q \times 1$), and the vectors $\mathbf{w}_n$ and $\mathbf{v}_n$ are the state and measurement noises[2], with zero means and covariance matrices $\mathbf{Q}_n$ and $\mathbf{R}_n$ respectively. The matrix $\mathbf{F}$ is the state transition matrix (of dimension $p \times p$), whereas $\mathbf{H}$ is the observation matrix (of dimension $q \times p$). The corresponding augmented state space model can be written as

$$\begin{aligned} \mathbf{x}^a_n &= \mathbf{F}^a_{n-1}\mathbf{x}^a_{n-1} + \mathbf{w}^a_n \\ \mathbf{y}^a_n &= \mathbf{H}^a_n\mathbf{x}^a_n + \mathbf{v}^a_n \end{aligned} \tag{8}$$

where $\mathbf{x}^a_n = [\mathbf{x}^T_n, \mathbf{x}^H_n]^T$, $\mathbf{y}^a_n = [\mathbf{y}^T_n, \mathbf{y}^H_n]^T$,
$\mathbf{F}^a_n = \begin{bmatrix} \mathbf{F}_n & \mathbf{A}_n \\ \mathbf{A}^*_n & \mathbf{F}^*_n \end{bmatrix}$ and $\mathbf{H}^a = \begin{bmatrix} \mathbf{H}_n & \mathbf{B}_n \\ \mathbf{B}^*_n & \mathbf{H}^*_n \end{bmatrix}$.

The terms $\mathbf{A}$ and $\mathbf{B}$ in the augmented state transition matrix and augmented observation matrix, allow for the state and observation equations to be widely linear. If both $\mathbf{A}$ and $\mathbf{B}$ are zero, then the state and observation equations are strictly linear.

[1]The 'widely linear' model is associated with the signal generating system, whereas 'augmented statistics' describe statistical properties of measured signals. Both the terms 'widely linear' and 'augmented' are used to name the resulting algorithms - in our work we mostly use the term 'augmented'.

[2]In the derivation of the Kalman filter, the state and measurement noises are assumed to be Gaussian, white and uncorrelated.

Algorithm 1 The augmented complex Kalman filter (ACKF) algorithm

Initialise with:

$$\begin{aligned}\widehat{\mathbf{x}}^a_{0|0} &= E\{\mathbf{x}^a_0\}\\ \mathbf{M}^a_{0|0} &= E\{(\mathbf{x}^a_0 - E\{\mathbf{x}^a_0\})(\mathbf{x}^a_0 - E\{\mathbf{x}^a_0\})^H\}\end{aligned}$$

Prediction:

$$\widehat{\mathbf{x}}^a_{n|n-1} = \mathbf{F}^a_{n-1}\widehat{\mathbf{x}}^a_{n-1|n-1} \tag{11}$$

Minimum Prediction MSE Matrix:

$$\mathbf{M}^a_{n|n-1} = \mathbf{F}^a_{n-1}\mathbf{M}^a_{n-1|n-1}(\mathbf{F}^a_{n-1})^H + \mathbf{Q}^a_n \tag{12}$$

Kalman Gain Matrix:

$$\mathbf{G}^a_n = \mathbf{M}^a_{n|n-1}(\mathbf{H}^a_n)^H[\mathbf{H}^a_n\mathbf{M}^a_{n|n-1}(\mathbf{H}^a_n)^H + \mathbf{R}^a_n]^{-1} \tag{13}$$

Correction:

$$\widehat{\mathbf{x}}^a_{n|n} = \widehat{\mathbf{x}}^a_{n|n-1} + \mathbf{G}^a_n(\mathbf{y}^a_n - \mathbf{H}^a_n\widehat{\mathbf{x}}^a_{n|n-1}) \tag{14}$$

Minimum MSE Matrix:

$$\mathbf{M}^a_{n|n} = (\mathbf{I} - \mathbf{G}^a_n\mathbf{H}^a_n)\mathbf{M}^a_{n|n-1} \tag{15}$$

The covariance matrices of the augmented state and measurement noises, $\mathbf{w}^a_n = [\mathbf{x}^T_n, \mathbf{w}^H_n]^T$ and $\mathbf{v}^a_n = [\mathbf{v}^T_n, \mathbf{v}^H_n]^T$, can be written as

$$\mathbf{Q}^a_n = E\{\mathbf{w}^a_n\mathbf{w}^{aH}_n\} = \begin{bmatrix}\mathbf{Q}_n & \mathbf{P}_n\\ \mathbf{P}^*_n & \mathbf{Q}^*_n\end{bmatrix} \tag{9}$$

$$\mathbf{R}^a_n = E\{\mathbf{v}^a_n\mathbf{v}^{aH}_n\} = \begin{bmatrix}\mathbf{R}_n & \mathbf{U}_n\\ \mathbf{U}^*_n & \mathbf{R}^*_n\end{bmatrix} \tag{10}$$

where $E\{\cdot\}$ is the statistical expectation operator, and $\mathbf{P}_n$ and $\mathbf{U}_n$ are the pseudocovariance matrices of $\mathbf{w}_n$ and $\mathbf{v}_n$ respectively.

The MMSE estimator $\widehat{\mathbf{x}}^a_{n|n} = E[\mathbf{x}^a_n|\mathbf{y}^a_0, \mathbf{y}^a_1, ..., \mathbf{y}^a_n]$ of $\mathbf{x}^a_n$ based on the observations $\{\mathbf{y}^a_0, \mathbf{y}^a_1, ..., \mathbf{y}^a_n\}$ can then be computed sequentially using Algorithm 1 [7].

The ACKF estimate $\widehat{\mathbf{x}}^a_{n|n} = \left[\widehat{\mathbf{x}}^T_{n|n}, \widehat{\mathbf{x}}^H_{n|n}\right]^T$ and the CCKF estimate $\widehat{\mathbf{x}}^L_{n|n}$ are both optimal in the MMSE sense, if the state and observation signals are white and uncorrelated with a Gaussian distribution [8] [9]. However, the difference between the two filters lies in the fact that the ACKF utilises the augmented covariance matrices, which cater for the noncircularity of the signals. However, the ACKF and CCKF have identical performance, namely $\widehat{\mathbf{x}}_{n|n} = \widehat{\mathbf{x}}^L_{n|n}$, for circular state and observation noises, and strictly linear state and observation equations, that is

$$\begin{aligned}\mathbf{Q}^a_n = \mathbf{Q}^L_n = \begin{bmatrix}\mathbf{Q}_n & \mathbf{0}\\ \mathbf{0} & \mathbf{Q}^*_n\end{bmatrix}, \quad \mathbf{R}^a_n = \mathbf{R}^L_n = \begin{bmatrix}\mathbf{R}_n & \mathbf{0}\\ \mathbf{0} & \mathbf{R}^*_n\end{bmatrix},\\ \mathbf{F}^a_n = \mathbf{F}^L_n = \begin{bmatrix}\mathbf{F}_n & \mathbf{0}\\ \mathbf{0} & \mathbf{F}^*_n\end{bmatrix} \quad \text{and} \quad \mathbf{H}^a_n = \mathbf{H}^L_n = \begin{bmatrix}\mathbf{H}_n & \mathbf{0}\\ \mathbf{0} & \mathbf{H}^*_n\end{bmatrix}\end{aligned} \tag{16}$$

The duality between the CCKF and ACKF under these conditions follows from the fact that both filters attain the same Kalman gain at every time instant. To this end, consider the predicted MMSE matrix, which can be expressed as

$$\begin{aligned}\mathbf{M}^a_{n+1|n} = \mathbf{F}^a_{n-1}&\mathbf{M}^a_{n|n-1}(\mathbf{F}^a_{n-1})^H\\ &-\mathbf{F}^a_{n-1}\mathbf{M}^a_{n|n-1}(\mathbf{H}^a_n)^H[\mathbf{H}^a_n\mathbf{M}^a_{n|n-1}(\mathbf{H}^a_n)^H + \mathbf{R}^a_n]^{-1}\\ &\times(\mathbf{H}^a_n)\mathbf{M}^a_{n|n-1}(\mathbf{F}^a_{n-1})^H + \mathbf{Q}^a_n\end{aligned} \tag{17}$$

which is a Riccati recursion. Observe that the computations of $\mathbf{M}^a_{n|n-1}$ and $\mathbf{M}^a_{n|n}$ are independent of the observation vector and as such can be calculated before any observations are taken into account. By substituting equation (13) into (15) and using the matrix inversion lemma, The matrix $\mathbf{M}^a_{n|n}$ can be expressed as

$$\begin{aligned}\mathbf{M}^a_{n|n} &= \mathbf{M}^a_{n|n-1} - \mathbf{M}^a_{n|n-1}(\mathbf{H}^a_n)^H \\ &\quad \times \big[\mathbf{H}^a_n \mathbf{M}^a_{n|n-1}(\mathbf{H}^a_n)^H + \mathbf{R}^a_n\big]^{-1}\mathbf{H}^a_n\mathbf{M}^a_{n|n-1} \\ &= \big[(\mathbf{M}^a_{n|n-1})^{-1} + (\mathbf{H}^a_n)^H(\mathbf{R}^a_n)^{-1}\mathbf{H}^a_n\big]^{-1} \end{aligned} \tag{18}$$

Substituting (18) into (13) allows the Kalman gain to be written as

$$\begin{aligned}\mathbf{G}^a_n &= [(\mathbf{M}^a_{n|n-1})^{-1} + (\mathbf{H}^a_n)^H(\mathbf{R}^a_n)^{-1}\mathbf{H}^a_n]^{-1}(\mathbf{H}^a_n)^H(\mathbf{R}^a_n)^{-1} \\ &= \mathbf{M}^a_{n|n}(\mathbf{H}^a_n)^H(\mathbf{R}^a_n)^{-1}\end{aligned} \tag{19}$$

Assuming the CCKF and ACKF have the same initialisation MSE matrices, that is

$$\mathbf{M}^a_{0|0} = \begin{bmatrix}\mathbf{M}^L_{0|0} & \mathbf{0} \\ \mathbf{0} & \mathbf{M}^{L*}_{0|0}\end{bmatrix} \tag{20}$$

where $\mathbf{M}^L_{0|0}$ is the initial MSE for the CCKF, then substituting $\mathbf{Q}^L_n$, $\mathbf{R}^L_n$, $\mathbf{F}^L_n$ and $\mathbf{H}^L_n$ into (19), gives us

$$\mathbf{G}^a_n = \begin{bmatrix}\mathbf{G}^L_n & \mathbf{0} \\ \mathbf{0} & \mathbf{G}^{*L}_n\end{bmatrix} \tag{21}$$

where $\mathbf{G}^L_n = \mathbf{M}^L_{n|n}(\mathbf{H}_n)^H(\mathbf{R}_n)^{-1}$ is the Kalman gain for the CCKF at time instant n. From equation (21) it is clear that the CCKF and ACKF have the same Kalman gain, and by substituting (21) into (14) we can see that the two filters will yield identical estimates for the state $\mathbf{x}_n$.

Remark #1: Hence, when the state and observation noises are both circular, and the augmented state transition and observation matrices are block-diagonal, the ACKF has the same performance as the CCKF.

A. Performance analysis

In this section, we illuminate the mean square error (MSE) performances of the CCKF and ACKF in order to provide insight into the behavior of Kalman filters for the generality of complex signals, both second order circular and noncircular [10]. We start from the general state space model for the Kalman filter given by (6) and (7). The Kalman filter estimate $\widehat{\mathbf{x}}_{n|n}$ of the state $\mathbf{x}_n$ is based on the all observations up to time n, and can be written as a linear combination of the sequence of observations, $\mathbf{z}_n = \big[\mathbf{y}^T_1, \mathbf{y}^T_2, ..., \mathbf{y}^T_n\big]^T$, that is

$$\widehat{\mathbf{x}}_{n|n} = \mathbf{W}_n\mathbf{z}_n \tag{22}$$

where $\mathbf{W}_n$ is the minimum MSE weight matrix, which is the solution to the normal equation, that is

$$\mathbf{W}_n = \mathbf{R}_{\mathbf{xz},n,n}\mathbf{R}^{-1}_{\mathbf{z},n} \tag{23}$$

with $\mathbf{R}_{\mathbf{xz},n,n} = E\big\{(\mathbf{x}_n - E\{\mathbf{x}_n\})(\mathbf{z}_n - E\{\mathbf{z}_n\})^H\big\}$ and $\mathbf{R}_{\mathbf{z},n} = E\big\{(\mathbf{z}_n - E\{\mathbf{z}_n\})(\mathbf{z}_n - E\{\mathbf{z}_n\})^H\big\}$. The MSE is then given by

$$\begin{aligned}\mathbf{M}_{n|n} &= E\{(\mathbf{x}_n - \widehat{\mathbf{x}}_{n|n})(\mathbf{x}_n - \widehat{\mathbf{x}}_{n|n})^H\} \\ &= \mathbf{R}_{\mathbf{x},n} - \mathbf{R}_{\mathbf{xz},n,n}\mathbf{R}^{-1}_{\mathbf{z},n}\mathbf{R}^H_{\mathbf{xz},n,n}\end{aligned} \tag{24}$$

The Kalman filter is summarised by state estimate (mean) (22) and covariance estimate (24) expressions at each time instant, however, these expressions are not recursive and the computational complexity increases with time. Nonetheless, these expressions suffice for analysis of the MSE performances of the CCKF and the ACKF.

We can write the state equation (6) non-recursively as

$$\mathbf{x}_n = \mathbf{F}_{n:0}\mathbf{x}_0 + \sum_{i=1}^{n}\mathbf{F}_{n:i}\mathbf{w}_i \tag{25}$$

where $\mathbf{x}_0$ is the initial state value with the assumption that $E\{\mathbf{x}_0\} = \mathbf{0}$, while the state transition matrix has the properties

$$\mathbf{F}_{n:i} = \mathbf{F}_n\mathbf{F}_{n-1}\cdots\mathbf{F}_i, \quad \mathbf{F}_{i:i} = \mathbf{I} \quad \text{and} \quad \mathbf{F}_0 = \mathbf{I}$$

Now, the state covariance matrix can be written as

$$\mathbf{R}_{\mathbf{x},n} = \mathbf{F}_{n:0}\mathbf{R}_{\mathbf{x},0}\mathbf{F}^H_{n:0} + \sum_{i=1}^{n}\mathbf{F}_{n:i}\mathbf{R}_{\mathbf{w},i}\mathbf{F}^H_{n:i} \tag{26}$$

and the observation covariance becomes

$$\begin{aligned}\mathbf{R}_{\mathbf{y},n,m} &= E\{\mathbf{y}_n\mathbf{y}_m^H\}\\ &= \begin{cases}\mathbf{H}_n\mathbf{R}_{\mathbf{x},n}\mathbf{H}_n^H+\mathbf{R}_{\mathbf{v},n} & \text{if } n=m\\ \mathbf{H}_n\mathbf{R}_{\mathbf{x},n}\mathbf{H}_m^H & \text{if } n<m\\ \mathbf{H}_n\mathbf{R}_{\mathbf{x},m}\mathbf{H}_m^H & \text{if } n>m\end{cases}\end{aligned} \tag{27}$$

where we made the usual assumptions that the measurement noise $\mathbf{v}[n]$ is orthogonal to the current and previous states, and the state noise is white. The cross-correlation between the state and observation can then be written as

$$\begin{aligned}\mathbf{R}_{\mathbf{xy},n,m} &= E\{\mathbf{x}_n\mathbf{y}_m^H\} \qquad n\geq m\\ &= E\{\mathbf{x}_n(\mathbf{H}_m\mathbf{x}_m+\mathbf{v}_m)^H\}\\ &= \mathbf{R}_{\mathbf{x},m}\mathbf{H}_m^H\end{aligned} \tag{28}$$

while the cross-correlation between the state $\mathbf{x}_n$ and the observation sequence $\mathbf{z}_n$, and the covariance of the observation sequence are given by

$$\mathbf{R}_{xz}[n,n]=\begin{bmatrix}\mathbf{R}_{\mathbf{xy},n,1} & \mathbf{R}_{\mathbf{xy},n,2} & \cdots & \mathbf{R}_{\mathbf{xy},n,n}\end{bmatrix} \tag{29}$$

and

$$\mathbf{R}_z[n]=\begin{bmatrix}\mathbf{R}_{\mathbf{y},1} & \mathbf{R}_{\mathbf{y},1,2} & \cdots & \mathbf{R}_{\mathbf{y},1,n}\\ \mathbf{R}_{\mathbf{y},2,1} & \mathbf{R}_{\mathbf{y},2} & \cdots & \mathbf{R}_{\mathbf{y},2,n}\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{R}_{\mathbf{y},n,1} & \mathbf{R}_{\mathbf{y},n,2} & \cdots & \mathbf{R}_{\mathbf{y},n}\end{bmatrix} \tag{30}$$

From taking the expectations of (22) and (25), it can be seen that the estimate $\widehat{\mathbf{x}}_{n|n}$ is an unbiased estimator of $\mathbf{x}_n$, that is,

$$E\{\mathbf{e}[n|n]\} = E\{(\mathbf{x}[n]-\widehat{\mathbf{x}}[n|n])\}=\mathbf{0}$$

and, as such, the mean characteristics of the conventional complex Kalman filter do not change with noncircular state and observation signal.

Equation (24) shows that the mean square characteristics of the CCKF is dependent on the covariance matrices of the state and observation noises but not on their pseudocovariances.

Remark #2: The noncircularity of the state and observation noises does not affect the performance of the linear conventional complex Kalman filter.

For the augmented complex Kalman filter (ACKF), the state estimate and the MSE matrix are given by expressions similar to (22) and (24), that is

$$\widehat{\mathbf{x}}^a_{n|n} = \mathbf{W}^a_n\mathbf{z}^a_n=\mathbf{R}^a_{\mathbf{xz},n,n}(\mathbf{R}^a_{\mathbf{z},n})^{-1}\mathbf{z}^a_n \tag{31}$$

$$\begin{aligned}\mathbf{M}^a_{n|n} &= E\{(\mathbf{x}^a_n-\widehat{\mathbf{x}}^a_{n|n})(\mathbf{x}^a_n-\widehat{\mathbf{x}}^a_{n|n})^H\}\\ &= \mathbf{R}^a_{\mathbf{x},n}-\mathbf{R}^a_{\mathbf{xz},n,n}(\mathbf{R}^a_{\mathbf{z},n})^{-1}\mathbf{R}^{aH}_{\mathbf{xz},n,n}\end{aligned} \tag{32}$$

where the matrix form of the augmented (widely linear) mean square error $\mathbf{M}^a_{n|n}$ can be written as

$$\begin{aligned}\begin{bmatrix}\mathbf{M}_{wl,n|n} & \mathbf{P}_{wl,n|n}\\ \mathbf{P}^*_{wl,n|n} & \mathbf{M}^*_{wl,n|n}\end{bmatrix} &= \begin{bmatrix}\mathbf{R}_{\mathbf{x},n} & \mathbf{P}_{\mathbf{x},n}\\ \mathbf{P}^*_{\mathbf{x},n} & \mathbf{R}^*_{\mathbf{x},n}\end{bmatrix}-\begin{bmatrix}\mathbf{R}_{\mathbf{xz},n,n} & \mathbf{P}_{\mathbf{xz},n,n}\\ \mathbf{P}^*_{\mathbf{xz},n,n} & \mathbf{R}^*_{\mathbf{xz},n,n}\end{bmatrix}\\ &\quad\times\begin{bmatrix}\mathbf{R}_{\mathbf{z},n} & \mathbf{P}_{\mathbf{z},n}\\ \mathbf{P}^*_{\mathbf{z},n} & \mathbf{R}^*_{\mathbf{z},n}\end{bmatrix}^{-1}\begin{bmatrix}\mathbf{R}_{\mathbf{xz},n,n} & \mathbf{P}_{\mathbf{xz},n,n}\\ \mathbf{P}^*_{\mathbf{xz},n,n} & \mathbf{R}^*_{\mathbf{xz},n,n}\end{bmatrix}^H\end{aligned}$$

The terms $\mathbf{P}_{\mathbf{x},n}$ and $\mathbf{P}_{\mathbf{z},n}$ are the pseudo-cross-correlation of the state and observation sequence respectively, while $\mathbf{P}_{\mathbf{xz},n,n}=E\{\mathbf{x}_n\mathbf{z}_n^T\}$ is the pseudo-correlation between the state and observation sequence. Notice that the inverse of the augmented covariance matrix $(\mathbf{R}^a_{\mathbf{z},n})^{-1}$ can be expressed as

$$\begin{bmatrix}\mathbf{R}_{\mathbf{z},n} & \mathbf{P}_{\mathbf{z},n}\\ \mathbf{P}^*_{\mathbf{z},n} & \mathbf{R}^*_{\mathbf{z},n}\end{bmatrix}^{-1}=\begin{bmatrix}\mathbf{C} & \mathbf{D}\\ \mathbf{D}^* & \mathbf{C}^*\end{bmatrix}$$

where

$$\begin{aligned}\mathbf{C} &= (\mathbf{R}_{\mathbf{z},n} - \mathbf{P}_{\mathbf{z},n}\mathbf{R}_{\mathbf{z},n}^{*-1}\mathbf{P}_{\mathbf{z},n}^{*})^{-1} \\ \mathbf{D} &= -(\mathbf{R}_{\mathbf{z},n} - \mathbf{P}_{\mathbf{z},n}\mathbf{R}_{\mathbf{z},n}^{*-1}\mathbf{P}_{\mathbf{z},n}^{*})^{-1}\mathbf{P}_{\mathbf{z},n}\mathbf{R}^{*-1}\end{aligned}$$

and, the widely linear (augmented) MSE of the ACKF can be expressed as

$$\begin{aligned}\mathbf{M}_{wl,n|n} = &\mathbf{R}_{\mathbf{x},n} - \mathbf{R}_{\mathbf{xz},n,n}\mathbf{C}\mathbf{R}_{\mathbf{xz},n,n}^{H} - \mathbf{R}_{\mathbf{xz},n,n}\mathbf{D}\mathbf{P}_{\mathbf{xz},n,n}^{H} \\ - &\mathbf{P}_{\mathbf{xz},n,n}\mathbf{D}^{*}\mathbf{R}_{\mathbf{xz},n,n}^{H} - \mathbf{P}_{\mathbf{xz},n,n}\mathbf{C}^{*}\mathbf{P}_{\mathbf{xz},n,n}^{H}\end{aligned} \tag{33}$$

After some tedious algebraic manipulations, the difference between the CCKF and the ACKF is found to be [11]

$$\begin{aligned}\Delta\mathbf{M}_{n} &= \mathbf{M}_{n|n} - \mathbf{M}_{wl,n|n} \\ &= (\mathbf{P}_{\mathbf{xz},n,n} - \mathbf{R}_{\mathbf{xz},n,n}\mathbf{R}_{\mathbf{z},n}^{-1}\mathbf{P}_{\mathbf{z},n}) \\ &\quad \times(\mathbf{R}_{\mathbf{z},n}^{*} - \mathbf{P}_{\mathbf{z},n}^{*}\mathbf{R}_{\mathbf{z},n}^{-1}\mathbf{P}_{\mathbf{z},n})^{-1} \\ &\quad \times(\mathbf{P}_{\mathbf{xz},n,n} - \mathbf{R}_{\mathbf{xz},n,n}\mathbf{R}_{\mathbf{z},n}^{-1}\mathbf{P}_{\mathbf{z},n})^{H}\end{aligned} \tag{34}$$

Remark #3: The expression (34) is always positive semidefinite since the matrix $(\mathbf{R}_{\mathbf{z},n}^{*} - \mathbf{P}_{\mathbf{z},n}^{*}\mathbf{R}_{\mathbf{z},n}^{-1}\mathbf{P}_{\mathbf{z},n})$ is positive definite, and consequently $\Delta\mathbf{M}_{n} = \mathbf{0}$ only when $(\mathbf{P}_{\mathbf{xz},n,n} - \mathbf{R}_{\mathbf{xz},n,n}^{-1}\mathbf{R}_{\mathbf{z},n}\mathbf{P}_{\mathbf{z},n}) = \mathbf{0}$. Therefore, the ACKF always has the same or better MSE performance than the CCKF.

Remark #4: The CCKF and ACKF are equivalent if the observation sequence is circular, $\mathbf{P}_{\mathbf{z},n} = \mathbf{0}$, and the state and observation sequence are jointly circular, $\mathbf{P}_{\mathbf{xz},n,n} = \mathbf{0}$.

Remark #5: Because $\Delta\mathbf{M}_{n} \geq \mathbf{0}$ for all time instants n, the ACKF hence has the same or better convergence than the CCKF.

B. Duality Analysis

Due to the duality between augmented complex vectors and real vectors, the ACKF has a dual real valued Kalman filter (KF), which can be used to significantly reduce its computational complexity. For any complex vector $\mathbf{z} = \mathbf{z}_r + j\mathbf{z}_i$ it holds that

$$\mathbf{z}^{a} = \begin{bmatrix}\mathbf{z} \\ \mathbf{z}^{*}\end{bmatrix} = \underbrace{\begin{bmatrix}\mathbf{I} & j\mathbf{I} \\ \mathbf{I} & -j\mathbf{I}\end{bmatrix}}_{\equiv\mathbf{J}}\underbrace{\begin{bmatrix}\mathbf{z_r} \\ \mathbf{z_i}\end{bmatrix}}_{=\mathbf{z}^{r}} \tag{35}$$

where $\mathbf{I}$ is the identity matrix and the invertible mapping $\mathbf{J} : \mathbb{C} \rightarrow \mathbb{R}$ is given by $\mathbf{J}^{-1} = \frac{1}{2}\mathbf{J}^{H}$ [12]. Based on this isomorphism, the real valued state space corresponding to the augmented complex state space is given by

$$\begin{aligned}\mathbf{x}_{n}^{r} &= \mathbf{F}_{n-1}^{r}\mathbf{x}_{n-1}^{r} + \mathbf{w}_{n}^{r} \\ \mathbf{y}_{n}^{r} &= \mathbf{H}_{n}^{r}\mathbf{x}_{n}^{r} + \mathbf{v}_{n}^{r}\end{aligned} \tag{36}$$

where $\mathbf{x}_{n}^{r} = \mathbf{J}^{-1}\mathbf{x}_{n}^{a}$, $\mathbf{y}_{n}^{r} = \mathbf{J}^{-1}\mathbf{y}_{n}^{a}$, $\mathbf{F}_{n-1}^{r} = \mathbf{J}^{-1}\mathbf{F}_{n-1}^{a}\mathbf{J}$, $\mathbf{H}_{n}^{r} = \mathbf{J}^{-1}\mathbf{H}_{n}^{a}\mathbf{J}$, $\mathbf{w}_{n}^{r} = \mathbf{J}^{-1}\mathbf{w}_{n}^{a}$ and $\mathbf{v}_{n}^{r} = \mathbf{J}^{-1}\mathbf{v}_{n}^{a}$. The covariance matrices of the real valued state and observation noises, $\mathbf{w}_{n}^{r}$ and $\mathbf{v}_{n}^{r}$, are given by

$$\begin{aligned}\mathbf{Q}_{n}^{r} &= E\{\mathbf{w}_{n}^{r}\mathbf{w}_{n}^{rH}\} = \mathbf{J}^{-1}\mathbf{Q}_{n}^{a}\mathbf{J}^{-H} \\ \mathbf{R}_{n}^{r} &= E\{\mathbf{v}_{n}^{r}\mathbf{v}_{n}^{rH}\} = \mathbf{J}^{-1}\mathbf{R}_{n}^{a}\mathbf{J}^{-H}\end{aligned}$$

It can be shown that the ACKF and its dual real valued KF have the same performance at each time instant. Assuming that ACKF is initiated at time $(n-1)$, with initial state $\widehat{\mathbf{x}}_{n-1|n-1}^{a}$ and MSE matrix $\mathbf{M}_{n-1|n-1}^{a}$, the corresponding dual real valued KF initialisation is given by

$$\begin{aligned}\widehat{\mathbf{x}}_{n-1|n-1}^{r} &= \mathbf{J}^{-1}\widehat{\mathbf{x}}_{n-1|n-1}^{a} \\ \mathbf{M}_{n-1|n-1}^{r} &= \mathbf{J}^{-1}\mathbf{M}_{n-1|n-1}^{a}\mathbf{J}^{-H}\end{aligned} \tag{37}$$

It is now straightforward to show that the state and MSE matrix predictions of the ACKF and its dual real valued KF are related as

$$\begin{aligned}\widehat{\mathbf{x}}_{n|n-1}^{r} &= \mathbf{J}^{-1}\widehat{\mathbf{x}}_{n|n-1}^{a} \\ \mathbf{M}_{n|n-1}^{r} &= \mathbf{J}^{-1}\mathbf{M}_{n|n-1}^{a}\mathbf{J}^{-H}\end{aligned} \tag{38}$$

and that he augmented Kalman gain is related to its corresponding real valued Kalman gain by the following expression

$$\begin{aligned}\mathbf{G}_n^a &= \mathbf{M}_{n|n-1}^a\mathbf{H}_n^{aH}[\mathbf{H}_n^a\mathbf{M}_{n|n-1}^a\mathbf{H}_n^{aH}+\mathbf{R}_n^a]^{-1}\\ &= \mathbf{J}\mathbf{M}_{n|n-1}^r\mathbf{J}^H\mathbf{J}^{-H}\mathbf{H}_n^{rH}\mathbf{J}^H\\ &\quad\times[\mathbf{J}\mathbf{H}_n^r\mathbf{J}^{-1}\mathbf{J}\mathbf{M}_{n|n-1}^r\mathbf{J}^H\mathbf{J}^{-H}\mathbf{H}_n^{rH}\mathbf{J}^H+\mathbf{J}\mathbf{R}_n^r\mathbf{J}^H]^{-1}\\ &= \mathbf{J}\mathbf{M}_{n|n-1}^r\mathbf{H}_n^{rH}[\mathbf{H}_n^r\mathbf{M}_{n|n-1}^r\mathbf{H}_n^{rH}+\mathbf{R}_n^r]^{-1}\mathbf{J}^{-1}\\ &= \mathbf{J}\mathbf{G}_n^r\mathbf{J}^{-1}\end{aligned}\tag{39}$$

It can be shown that the state estimates $\widehat{\mathbf{x}}_{n|n}^a$ and $\widehat{\mathbf{x}}_{n|n}^r$ have the following relationship

$$\begin{aligned}\widehat{\mathbf{x}}_{n|n}^r &= \widehat{\mathbf{x}}_{n|n-1}^r+\mathbf{G}_n^r(\mathbf{y}_n^r-\mathbf{H}_n^r\widehat{\mathbf{x}}_{n|n-1}^r)\\ &= \mathbf{J}^{-1}\widehat{\mathbf{x}}_{n|n-1}^a+\mathbf{J}^{-1}\mathbf{G}_n^a\mathbf{J}(\mathbf{y}_n^r-\mathbf{H}_n^r\mathbf{J}^{-1}\widehat{\mathbf{x}}_{n|n-1}^a)\\ &= \mathbf{J}^{-1}\widehat{\mathbf{x}}_{n|n}^a\end{aligned}\tag{40}$$

and the MSE matrices are related as

$$\mathbf{M}_{n|n}^r=\mathbf{J}^{-1}\mathbf{M}_{n|n}^a\mathbf{J}^{-H}\tag{41}$$

From (40) it is clear that the state estimates $\widehat{\mathbf{x}}_{n|n}^a$ and $\widehat{\mathbf{x}}_{n|n}^r$ are equivalent and separated by an invertible linear mapping. The ACKF and ERKF can also be shown to achieve the same mean square error. Let mean square error for the ERKF be given by

$$\epsilon_n^r=tr\{\mathbf{M}_{n|n}^r\}\tag{42}$$

where $tr\{\cdot\}$ is the matrix trace operator. The mean square error corresponding to the augmented MSE matrix $\mathbf{M}_{n|n}^a$ is given by taking the trace of (41), that is

$$\begin{aligned}tr\{\mathbf{M}_{n|n}^a\} &= tr\{\mathbf{J}\mathbf{M}_{n|n}^r\mathbf{J}^H\}\\ &= tr\{\mathbf{M}_{n|n}^r\mathbf{J}^H\mathbf{J}\}\\ &= 2\cdot tr\{\mathbf{M}_{n|n}^r\}\end{aligned}\tag{43}$$

In (43), we used the fact that $\mathbf{J}^H=2\mathbf{J}^{-1}$. Equation (43) is misleading in that it suggests that ACKF has twice the error of its dual real valued KF. However, the error is counted twice by take the trace of $\mathbf{M}_{n|n}^a$, due to the block diagonal structure of augmented MSE covariance matrix, and hence needs to be halved to find the true augmented mean square error, that is

$$\epsilon_n^a = \frac{1}{2}tr\{\mathbf{M}_{n|n}^a\}=\epsilon_n^r$$

Therefore the ACKF and the its dual 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 signals. However, by using the dual real valued KF, the computational complexity of the ACKF is reduced, whereby the number of additions and multiplications required are approximately halved and quartered, respectively.

III. The Augmented Complex Extended Kalman Filter (ACEKF)

The extended Kalman filter (EKF) uses linear models to approximate nonlinear functions, and as such, the state and observation functions need not be linear but differentiable. Consider the state space model given by

$$\mathbf{x}_n = \mathbf{f}[\mathbf{x}_{n-1}]+\mathbf{w}_n\tag{44a}$$

$$\mathbf{y}_n = \mathbf{h}[\mathbf{x}_n]+\mathbf{v}_n\tag{44b}$$

where $\mathbf{f}[\cdot]$ and $\mathbf{h}[\cdot]$ are the nonlinear process and observations vector valued models respectively and the remaining variables are as defined above. The extended Kalman filter approximates these nonlinear functions by their first order Taylor series expansions (TSE) about certain desired points. However, calculating the complex derivative of a function requires the function to be analytic (differentiable) within the rigorous conditions set by the Cauchy-Riemann equations, though in practice, the functions $\mathbf{f}[\cdot]$ and $\mathbf{h}[\cdot]$ can be analytic or nonanalytic depending on the underlying physical model. Moreover, there is a large class of functions, such as real functions of complex variables, which do not satisfy the Cauchy-Riemann conditions thus severely restricting the set of allowable functions for nonlinear process and observations models.

By utilising the so called $\mathbb{CR}$ calculus framework [13], which exploits the isomorphism between the complex domain $\mathbb{C}$ and the real domain $\mathbb{R}$, the Taylor series expansions of both analytic and nonanalytic functions are still possible within the same framework. For instance, in $\mathbb{CR}$ calculus, the first order TSE of a function $f[\mathbf{z}]$ is given by

$$f[\mathbf{z}+\Delta\mathbf{z}]=f[\mathbf{z}]+\frac{\partial f}{\partial\mathbf{z}}\Delta\mathbf{z}+\frac{\partial f}{\partial\mathbf{z}^*}\Delta\mathbf{z}^*\tag{45}$$

whereby for analytic functions (in the the Cauchy-Riemann sense), the term $\frac{\partial f}{\partial \mathbf{z}^*}\Delta \mathbf{z}^*$ vanishes.

Consider the first order approximations of the state and observation equations, (44a) and (44b), about the estimates $\widehat{\mathbf{x}}_{n-1|n-1}$ and $\widehat{\mathbf{x}}_{n|n-1}$, that is

$$\mathbf{x}_n = \mathbf{F}_{n-1}\mathbf{x}_{n-1} + \mathbf{A}_{n-1}\mathbf{x}^*_{n-1} + \mathbf{w}_n + \mathbf{r}_{n-1} \tag{46}$$
$$\mathbf{y}_n = \mathbf{H}_n\mathbf{x}_n + \mathbf{B}_n\mathbf{x}^*_n + \mathbf{v}_n + \mathbf{z}_n \tag{47}$$

where the vectors $\mathbf{r}_n = \mathbf{f}[\widehat{\mathbf{x}}_{n-1|n-1}] - \mathbf{F}_{n-1}\widehat{\mathbf{x}}_{n-1|n-1} - \mathbf{A}_{n-1}\widehat{\mathbf{x}}^*_{n-1|n-1}$ and $\mathbf{z}_n = \mathbf{h}[\widehat{\mathbf{x}}_{n|n-1}] - \mathbf{H}_n\widehat{\mathbf{x}}_{n|n-1} - \mathbf{B}_n\widehat{\mathbf{x}}^*_{n|n-1}$, and the matrices $\mathbf{F}_{n-1}$, $\mathbf{A}_{n-1}$, $\mathbf{H}_n$ and $\mathbf{B}_n$ are the Jacobians defined as

$$\mathbf{F}_{n-1} = \frac{\partial \mathbf{f}}{\partial \mathbf{x}_{n-1}}\Big|_{\mathbf{x}_{n-1}=\widehat{\mathbf{x}}_{n-1|n-1}}, \ \mathbf{A}_{n-1} = \frac{\partial \mathbf{f}}{\partial \mathbf{x}^*_{n-1}}\Big|_{\mathbf{x}^*_{n-1}=\widehat{\mathbf{x}}^*_{n-1|n-1}},$$
$$\mathbf{H}_n = \frac{\partial \mathbf{h}}{\partial \mathbf{x}_n}\Big|_{\mathbf{x}_n=\widehat{\mathbf{x}}_{n|n-1}} \text{ and } \mathbf{B}_n = \frac{\partial \mathbf{h}}{\partial \mathbf{x}^*_n}\Big|_{\mathbf{x}^*_n=\widehat{\mathbf{x}}^*_{n|n-1}}$$

From (46) and (47), we observe that if $\mathbf{f}[\cdot]$ and $\mathbf{h}[\cdot]$ are nonanalytic, the linearised state and observation equations are widely linear (see (4)), and thus cannot be implemented using the standard complex extended Kalman filter (CEKF). However, the state space equations become strictly linear if these functions are analytic, since the derivatives with respect to the complex conjugates vanish, that is $\mathbf{A}_{n-1} = \mathbf{0}$ and $\mathbf{B}_n = \mathbf{0}$.

In order to provide deeper insight into the widely linear state and observation models, an 'augmented' state space representation is thus required; this will also cater for the full second order statistics of the process and measurement noises. To this end, consider the nonlinear augmented state space model given by

$$\mathbf{x}^a_n = \mathbf{f}^a[\mathbf{x}^a_{n-1}] + \mathbf{w}^a_n \tag{48a}$$
$$\mathbf{y}^a_n = \mathbf{h}^a[\mathbf{x}^a_n] + \mathbf{v}^a_n \tag{48b}$$

with $\mathbf{f}^a[\mathbf{x}^a_{n-1}] = \big[\mathbf{f}^T[\mathbf{x}^a_{n-1}], \mathbf{f}^H[\mathbf{x}^a_{n-1}]\big]^T$ and $\mathbf{h}^a[\mathbf{x}^a_n] = \big[\mathbf{h}^T[\mathbf{x}^a_n], \mathbf{h}^H[\mathbf{x}^a_n]\big]^T$. The linearised augmented state space can be expressed as

$$\mathbf{x}^a_n = \mathbf{F}^a_{n-1}\mathbf{x}^a_{n-1} + \mathbf{w}^a_n + \mathbf{r}^a_{n-1} \tag{49a}$$
$$\mathbf{y}^a_n = \mathbf{H}^a_n\mathbf{x}^a_n + \mathbf{v}^a_n + \mathbf{z}^a_n \tag{49b}$$

where $\mathbf{r}^a_n = \big[\mathbf{r}^T_n, \mathbf{r}^H_n\big]^T$, $\mathbf{z}^a_n = \big[\mathbf{z}^T_n, \mathbf{z}^H_n\big]^T$,

$$\mathbf{F}^a_n = \begin{bmatrix} \mathbf{F}_n & \mathbf{A}_n \\ \mathbf{A}^*_n & \mathbf{F}^*_n \end{bmatrix} \text{ and } \mathbf{H}^a = \begin{bmatrix} \mathbf{H}_n & \mathbf{B}_n \\ \mathbf{B}^*_n & \mathbf{H}^*_n \end{bmatrix}.$$

Note that $\mathbf{F}^a_n = \frac{\partial \mathbf{f}^a}{\partial \mathbf{x}^a_n}$ and $\mathbf{H}^a_n = \frac{\partial \mathbf{h}^a}{\partial \mathbf{x}^a_n}$.

Therefore, in contrast to the conventional CEKF, the ACEKF allows the state and observation models to be widely linear, and thus naturally caters for the noncircularity of the state and measurement noises. The derivation of the ACEKF follows from the derivation of the CEKF [1, Ch. 15.4], and utilises the augmented state space, and is summarised in Algorithm 2 [14].

The novelty of the ACEKF algorithm presented in this work is that it does not assume a specific state or observation models, that is $\mathbf{f}[\cdot]$ and $\mathbf{h}[\cdot]$, which makes it a more general form of the ACEKF presented in [3]. Moreover, by utilising the $\mathbb{CR}$ calculus framework, we have shown how the ACEKF can be used for the generality of complex state space models, both holomorphic and nonholomorphic.

A. *Duality Analysis of ACEKF and real valued EKF*

Similar to the ACKF, the ACEKF has a real valued EKF counterpart, which gives the same estimate at every time instant. Based on the isomorphism between augmented complex and real valued vectors, the nonlinear real valued state space corresponding to the augmented complex state space (48) is given by

$$\mathbf{x}^r_n = \mathbf{f}^r[\mathbf{x}^r_{n-1}] + \mathbf{w}^r_n \tag{55a}$$
$$\mathbf{y}^r_n = \mathbf{h}^r[\mathbf{x}^r_n] + \mathbf{v}^r_n \tag{55b}$$

where $\mathbf{x}^r_n = \mathbf{J}^{-1}\mathbf{x}^a_n$, $\mathbf{y}^r_n = \mathbf{J}^{-1}\mathbf{y}^a_n$, $\mathbf{v}^r_n = \mathbf{J}^{-1}\mathbf{v}^a_n$, $\mathbf{w}^r_n = \mathbf{J}^{-1}\mathbf{w}^a_n$, $\mathbf{f}^r[\mathbf{x}^r_n] = \mathbf{J}^{-1}\mathbf{f}^a[\mathbf{x}^a_n]$ and $\mathbf{h}^r[\mathbf{x}^r_n] = \mathbf{J}^{-1}\mathbf{h}^a[\mathbf{x}^a_n]$. The covariance matrices of the real valued state and observation noises, $\mathbf{w}^r_n$ and $\mathbf{v}^r_n$, are given by

$$\mathbf{Q}^r_n = E\{\mathbf{w}^r_n\mathbf{w}^{rT}_n\} = \mathbf{J}^{-1}\mathbf{Q}^a_n\mathbf{J}^{-H}$$
$$\mathbf{R}^r_n = E\{\mathbf{v}^r_n\mathbf{v}^{rT}_n\} = \mathbf{J}^{-1}\mathbf{R}^a_n\mathbf{J}^{-H}$$

Algorithm 2 The augmented complex extended Kalman filter (AEKF) algorithm

Initialise with:

$$\begin{aligned}\widehat{\mathbf{x}}^a_{0|0} &= E\{\mathbf{x}^a_0\}\\ \mathbf{M}^a_{0|0} &= E\{(\mathbf{x}^a_0 - E\{\mathbf{x}^a_0\})(\mathbf{x}^a_0 - E\{\mathbf{x}^a_0\})^H\}\end{aligned}$$

Prediction:

$$\widehat{\mathbf{x}}^a_{n|n-1} = \mathbf{f}^a[\widehat{\mathbf{x}}^a_{n-1|n-1}] \tag{50}$$

Prediction Covariance Matrix:

$$\mathbf{M}^a_{n|n-1} = \mathbf{F}^a_{n-1}\mathbf{M}^a_{n-1|n-1}(\mathbf{F}^a_{n-1})^H + \mathbf{Q}^a_n \tag{51}$$

Kalman Gain Matrix:

$$\mathbf{G}^a_n = \mathbf{M}^a_{n|n-1}(\mathbf{H}^a_n)^H[\mathbf{H}^a_n\mathbf{M}^a_{n|n-1}(\mathbf{H}^a_n)^H + \mathbf{R}^a_n]^{-1} \tag{52}$$

Correction:

$$\widehat{\mathbf{x}}^a_{n|n} = \widehat{\mathbf{x}}^a_{n|n-1} + \mathbf{G}^a_n(\mathbf{y}^a_n - \mathbf{h}^a[\widehat{\mathbf{x}}^a_{n|n-1}]) \tag{53}$$

Covariance Matrix:

$$\mathbf{M}^a_{n|n} = (\mathbf{I} - \mathbf{G}^a_n\mathbf{H}^a_n)\mathbf{M}^a_{n|n-1} \tag{54}$$

The relationship between the augmented complex and real Jacobians $\mathbf{F}^a_n$ and $\mathbf{F}^r_n$ is established by using $\mathbf{f}^r[\mathbf{x}^r_n] = \mathbf{J}^{-1}\mathbf{f}^a[\mathbf{x}^a_n]$ and by comparing the first order TSE of both sides, that is

$$\begin{aligned}\mathbf{f}^r[\mathbf{x}^r_n] &= \mathbf{J}^{-1}\mathbf{f}^a[\mathbf{x}^a_n]\\ &\approx \mathbf{J}^{-1}\mathbf{f}^a[\widehat{\mathbf{x}}^a_{n|n-1}] + \mathbf{J}^{-1}\mathbf{F}^a_n(\mathbf{x}^a_n - \widehat{\mathbf{x}}^a_{n|n-1})\\ &\approx \mathbf{J}^{-1}\mathbf{f}^a[\widehat{\mathbf{x}}^a_{n|n-1}] + \mathbf{J}^{-1}\mathbf{F}^a_n\mathbf{J}(\mathbf{x}^r_n - \widehat{\mathbf{x}}^r_{n|n-1})\\ &\approx \mathbf{f}^r[\widehat{\mathbf{x}}^r_{n|n-1}] + \mathbf{F}^r_{n-1}(\mathbf{x}^r_n - \widehat{\mathbf{x}}^r_{n|n-1})\end{aligned} \tag{56}$$

This shows that $\mathbf{F}^r_n = \mathbf{J}^{-1}\mathbf{F}^a_n\mathbf{J} = \frac{\partial \mathbf{f}^r}{\partial \mathbf{x}^r_n}\Big|_{\mathbf{x}^r_n = \widehat{\mathbf{x}}^r_{n|n}}$, and similarly $\mathbf{H}^r_n = \mathbf{J}^{-1}\mathbf{H}^a_n\mathbf{J} = \frac{\partial \mathbf{h}^r}{\partial \mathbf{x}^r_n}\Big|_{\mathbf{x}^r_n = \widehat{\mathbf{x}}^r_{n|n}}$.

The correspondence between the ACEKF and its dual real valued extended Kalman filter can be established [15] by showing that at every time instant n the following relationships hold:

$$\begin{aligned}\widehat{\mathbf{x}}^r_{n|n-1} &= \mathbf{J}^{-1}\widehat{\mathbf{x}}^a_{n|n-1}\\ \mathbf{M}^r_{n|n-1} &= \mathbf{J}^{-1}\mathbf{M}^a_{n|n-1}\mathbf{J}^{-H}\\ \mathbf{G}^r_n &= \mathbf{J}^{-1}\mathbf{G}^a_n\mathbf{J}\\ \widehat{\mathbf{x}}^r_{n|n} &= \mathbf{J}^{-1}\widehat{\mathbf{x}}^a_{n|n}\\ \mathbf{M}^r_{n|n} &= \mathbf{J}^{-1}\mathbf{M}^a_{n|n}\mathbf{J}^{-H}\end{aligned} \tag{57}$$

The ACEKF and its dual real valued extended Kalman filter essentially implement the same state space model, but operate in different domains. If the state space is naturally defined in the complex domain, it is generally desirable to keep all of the computations in the original complex domain in order to facilitate the understanding of the transformations the signal goes through, and to benefit from the notions of phase and circularity. However, the real valued equivalent extended Kalman filter provides means for reducing the computational complexity of the ACEKF, whereby similar to the ACKF, the number of additions and multiplications required are approximately halved and quartered, respectively.

IV. The Augmented Complex Kalman Filter Unscented (ACUKF)

The unscented Kalman filter (UKF) [16] has been proposed to address the problems arising from the first order approximation of nonlinearities withing EKFs, and approximates the statistical posterior distribution rather than approximating the nonlinearity [17]. The UKF uses a deterministic sampling technique to pick a set of sample points (known as sigma points) around the mean. These points are then propagated through the nonlinear state space models, from which the mean and covariance of the estimate are then recovered. This results in a filter which is able to more accurately capture the true state mean and covariance.

To illustrate the complex unscented transform (UT) and the augmented complex UT, consider the mapping

$$\mathbf{y} = \mathbf{f}[\mathbf{x}] = \mathbf{f}[\bar{\mathbf{x}} + \delta\mathbf{x}] \qquad \mathbf{x} \in \mathbb{C}^{p\times 1}, \mathbf{y} \in \mathbb{C}^{q\times 1} \tag{58}$$

where $\mathbf{f}[\cdot]$ is a nonlinear vector valued function, $\mathbf{y} = [y_1, \ldots, y_q]^T$ the output, $\mathbf{x} = [x_1, \ldots, x_p]^T$ is the input whose mean $\bar{\mathbf{x}} = E\{\mathbf{x}\}$, covariance $\mathbf{R}_\mathbf{x} = E\{(\mathbf{x}-\bar{\mathbf{x}})(\mathbf{x}-\bar{\mathbf{x}})^H\}$, pseudocovariance $\mathbf{P}_\mathbf{x} = E\{(\mathbf{x}-\bar{\mathbf{x}})(\mathbf{x}-\bar{\mathbf{x}})^T\} = \mathbf{0}$, and $\delta\mathbf{x} = \mathbf{x} - \bar{\mathbf{x}}$. The Taylor series expansion (TSE) of $\mathbf{y}$ about $\bar{\mathbf{x}}$ is given by

$$\mathbf{y} = \mathbf{f}[\bar{\mathbf{x}}] + \nabla_{\delta\mathbf{x}}\mathbf{f} + \frac{1}{2!}\nabla^2_{\delta\mathbf{x}}\mathbf{f} + \frac{1}{3!}\nabla^3_{\delta\mathbf{x}}\mathbf{f} + \cdots \tag{59}$$

where the ith order term in the TSE for $\mathbf{f}[\cdot]$ about $\bar{\mathbf{x}}$ is

$$\frac{1}{i!}\nabla^i_{\delta\mathbf{x}}\mathbf{f} = \frac{1}{i!}\Big(\sum_{k=1}^{p}\delta x_k \frac{\partial}{\partial x_k}\Big)^i \mathbf{f}[\mathbf{x}]\big|_{\mathbf{x}=\bar{\mathbf{x}}} \tag{60}$$

with δx_k being the kth component of $\delta\mathbf{x}$. The term above is an ith order polynomial in $\delta\mathbf{x}$ whose coefficients are given by the derivatives of $\mathbf{f}[\cdot]$. The mean of $\mathbf{y}$ can be expressed as

$$\begin{aligned}\bar{\mathbf{y}} &= E\{\mathbf{f}[\bar{\mathbf{x}} + \delta\mathbf{x}]\}\\ &= \mathbf{f}[\bar{\mathbf{x}}] + E\Big\{\nabla_{\delta\mathbf{x}}\mathbf{f} + \frac{1}{2!}\nabla^2_{\delta\mathbf{x}}\mathbf{f} + \frac{1}{3!}\nabla^3_{\delta\mathbf{x}}\mathbf{f} + \cdots\Big\}\end{aligned}$$

where the ith term is given by

$$\begin{aligned}E\Big\{\frac{1}{i!}\nabla^i_{\delta\mathbf{x}}\mathbf{f}\Big\} &= \frac{1}{i!}E\Big\{\Big(\sum_{k=1}^{p}\delta x_k\frac{\partial}{\partial x_k}\Big)^i\Big\}\mathbf{f}[\mathbf{x}]\big|_{\mathbf{x}=\bar{\mathbf{x}}}\\ &= \frac{1}{i!}\Big(m_{1,1,\cdots,1,1}\frac{\partial^i\mathbf{f}}{\partial x_1^i} + m_{1,1,\cdots,1,2}\frac{\partial^i\mathbf{f}}{\partial x_1^{i-1}\partial x_2} + \cdots\Big)\end{aligned}$$

The symbols $m_{a_1,a_2,\ldots,a_{i-1},a_i} = E\{\delta x_{a_1}\delta x_{a_2}\cdots\delta x_{a_{i-1}}\delta x_{a_i}\}$ denote the ith order central moments of the components $\mathbf{x}$ with $a_k \in [1, 2, \ldots, p]$. Observe that the ith order term in the series for $\bar{\mathbf{y}}$ is a function of the ith order central moment of $\mathbf{x}$ multiplied by the ith derivative of $\mathbf{f}[\cdot]$. Hence if the moments can be correctly evaluated up to the ith order, the mean $\bar{\mathbf{y}}$ will also be correct up to the ith order. The covariance matrix $\mathbf{R}_\mathbf{y} = E\{(\mathbf{y}-\bar{\mathbf{y}})(\mathbf{y}-\bar{\mathbf{y}})^H\}$ now becomes

$$\begin{aligned}\mathbf{R}_\mathbf{y} = \frac{\partial\mathbf{f}}{\partial\mathbf{x}}\mathbf{R}_\mathbf{x}\Big(\frac{\partial\mathbf{f}}{\partial\mathbf{x}}\Big)^H + E\Big\{&\frac{1}{3!}\nabla_{\delta\mathbf{x}}\mathbf{f}\big(\nabla^3_{\delta\mathbf{x}}\mathbf{f}\big)^H + \frac{1}{2!\times 2!}\nabla^2_{\delta\mathbf{x}}\mathbf{f}\big(\nabla^2_{\delta\mathbf{x}}\mathbf{f}\big)^H\\ &+ \frac{1}{3!}\nabla^3_{\delta\mathbf{x}}\mathbf{f}\big(\nabla_{\delta\mathbf{x}}\mathbf{f}\big)^H\Big\} - E\Big\{\frac{1}{2!}\nabla^2_{\delta\mathbf{x}}\mathbf{f}\Big\}E\Big\{\frac{1}{2!}\nabla^2_{\delta\mathbf{x}}\mathbf{f}\Big\}^H + \cdots\end{aligned}$$

and is correct if the ith central moment of $\mathbf{x}$ is correct. Within the complex unscented transform framework, the $p-$dimensional random variable $\mathbf{x}$ is approximated by a set $(2p+1)$ weighted (sigma) points $\{\mathcal{W}_i, \mathcal{X}_i\}_{i=0}^{2p+1}$, chosen so that their sample mean and covariance are equal to the true mean $\bar{\mathbf{x}}$ and covariance $\mathbf{R}_\mathbf{x}$. The nonlinear function $\mathbf{f}[\cdot]$ is then applied to each of these points to generate transformed points, $\mathcal{Y}_i = \mathbf{f}[\mathcal{X}_i]$, with a sample mean and covariance

$$\hat{\bar{\mathbf{y}}} = \sum_{i=0}^{2p}\mathcal{W}_i\mathcal{Y}_i \qquad \hat{\mathbf{R}}_\mathbf{y} = \sum_{i=0}^{2p}\mathcal{W}_i\big(\mathcal{Y}_i - \bar{\mathbf{y}}\big)\big(\mathcal{Y}_i - \bar{\mathbf{y}}\big)^H$$

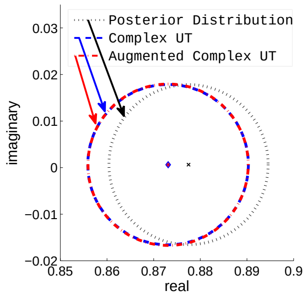
which are correct up to the second order. For a second order noncircular $\mathbf{y}$, the true output pseudocovariance $\mathbf{P}_\mathbf{y} = E\{(\mathbf{y} - \bar{\mathbf{y}})(\mathbf{y}-\bar{\mathbf{y}})^T\}$ is given by

$$\begin{aligned}\mathbf{P}_\mathbf{y} = \frac{\partial\mathbf{f}}{\partial\mathbf{x}}\mathbf{P}_\mathbf{x}\Big(\frac{\partial\mathbf{f}}{\partial\mathbf{x}}\Big)^T + E\Big\{&\frac{1}{3!}\nabla_{\delta\mathbf{x}}\mathbf{f}\big(\nabla^3_{\delta\mathbf{x}}\mathbf{f}\big)^T + \frac{1}{2!\times 2!}\nabla^2_{\delta\mathbf{x}}\mathbf{f}\big(\nabla^2_{\delta\mathbf{x}}\mathbf{f}\big)^T\\ &+ \frac{1}{3!}\nabla^3_{\delta\mathbf{x}}\mathbf{f}\big(\nabla_{\delta\mathbf{x}}\mathbf{f}\big)^T\Big\} - E\Big\{\frac{1}{2!}\nabla^2_{\delta\mathbf{x}}\mathbf{f}\Big\}E\Big\{\frac{1}{2!}\nabla^2_{\delta\mathbf{x}}\mathbf{f}\Big\}^T + \cdots\end{aligned}$$

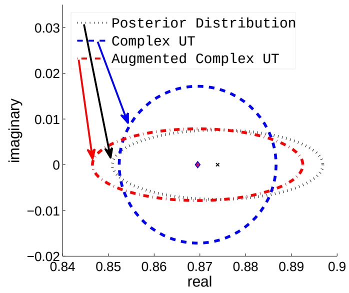
The standard complex unscented transform is does not cater for the input pseudocovariance and consequently the output pseudocovariance, due to the method used for generating the sigma points, which are calculated as

$$\mathcal{X}_0 = \bar{\mathbf{x}} \qquad \mathcal{X}_i = \bar{\mathbf{x}} \pm \Big(\sqrt{(p+\lambda)\mathbf{R}_\mathbf{x}}\Big)_i, i = 1, \ldots, 2p \tag{61}$$

where $\Big(\sqrt{(p+\lambda)\mathbf{R}_\mathbf{x}}\Big)_i$ is the ith column of the matrix square root and $\lambda = \alpha^2(2p+\kappa) - 2p$ is a scaling parameter, while α determines the spread of the sigma points around the mean and is usually set to a small positive value (e.g., 10^{-3}), κ is



(a) circular input



(b) noncircular input

Fig. 1. Performance of the complex UT and augmented complex UT

a secondary scaling parameter which is usually set to 0, and β is used to incorporate prior knowledge of the distribution (for Gaussian distributions, $\beta = 2$ is optimal). From (61) it is clear that the sigma points do not incorporate the pseudocovariance.

To incorporate the pseudocovariance into estimates of the output moments, consider the 'augmented' function

$$\mathbf{y}^a = \mathbf{f}^a[\mathbf{x}^a] = \mathbf{f}^a[\bar{\mathbf{x}}^a + \delta\mathbf{x}^a] \tag{62}$$

where $\mathbf{x}^a = [\mathbf{x}^T, \mathbf{x}^H]$, $\mathbf{y}^a = [\mathbf{y}^T, \mathbf{y}^H]$ and $\delta\mathbf{x}^a = [\delta\mathbf{x}^T, \delta\mathbf{x}^H]$. The sigma points corresponding to this model are given by

$$\mathcal{X}_0^a = \bar{\mathbf{x}}^a \qquad \mathcal{X}_i^a = \bar{\mathbf{x}}^a \pm \left(\sqrt{(p+\lambda)\mathbf{R}_{\mathbf{x}}^a}\right)_i, i = 1, \dots, 4p$$

and are functions of the mean of the augmented input and the augmented covariance matrix and can fully propagate the second order statistics of improper inputs.

To illustrate the benefits of the augmented complex UT over the standard UT, consider the system defined by $y_n = \cos[x_n]$ where the input x_n is a Gaussian doubly white circular random variable. Figure 1a shows that for a circular input $x_n \sim \mathcal{N}(\bar{x}, c_x, \rho_x) = \mathcal{N}(0.5, 0.01, 0)$ (c_x is the variance and ρ_x the pseudocovariance) the complex UT and the augmented complex UT had similar performance in capturing the distribution of the output y_n. Figure 1b illustrates that for a noncircular input $x_n \sim \mathcal{N}(0.5, 0.01, 0.008)$ the augmented complex UT captures the pseudocovariance of the output distribution closely, while the complex UT maintains a circular posterior distribution.

Based on (4) and (62), consider the 'augmented' model

$$\mathbf{x}_n^a = \mathbf{f}^a[\mathbf{x}_{n-1}^a] + \mathbf{w}_n^a \tag{63}$$

$$\mathbf{y}_n^a = \mathbf{h}^a[\mathbf{x}_n^a] + \mathbf{v}_n^a \tag{64}$$

The weights associated with the $(4p+1)$ augmented sigma points are then given by

$$\begin{aligned}\mathcal{W}_0^{(m)} &= \frac{\lambda}{2p+\lambda} \quad \mathcal{W}_0^{(c)} = \frac{\lambda}{2p+\lambda} + (1 - \alpha^2 + \beta) \\ \mathcal{W}_i^{(m)} &= \mathcal{W}_i^{(c)} = \frac{\lambda}{2(2p+\lambda)}, \qquad i = 1, \dots, 4p\end{aligned} \tag{65}$$

and the augmented complex unscented Kalman filter (ACUKF) is summarised in Algorithm 3 [18]. The novelty of the ACUKF algorithm presented in this work is that it does not assume a specific state or observation models which makes it a more general form of the ACUKF presented in [1].

A. Performance analysis

In this section we analyse the mean-square behavior of the the CUKF [16] for analytic state and observation functions. Consider the complex valued scalar state space given by

$$x_n = f[x_{n-1}] + w_n \tag{69}$$

$$y_n = h[x_n] + v_n \tag{70}$$

where $f[\cdot]$ and $h[\cdot]$ are the analytic nonlinear process and observation models respectively, x_n and y_n are the state and noisy observation, while w_n and v_n are uncorrelated zero-mean white complex-valued state (process) and observation (measurement) noises respectively. The process noise has variance $c_{w,n} = E\{w_n w_n^*\}$ and pseudocovariance $\rho_{w,n} = E\{w_n w_n\}$, while the measurement noise has a variance $c_{v,n} = E\{v_n v_n^*\}$ and pseudocovariance $\rho_{v,n} = E\{v_n v_n\}$.

Algorithm 3 The augmented unscented Kalman filter (AUKF) algorithm

Initialise with:

$$\begin{aligned}\widehat{\mathbf{x}}^a_{0|0} &= E\{\mathbf{x}^a_0\}\\ \mathbf{M}^a_{0|0} &= E\{(\mathbf{x}^a_0 - E\{\mathbf{x}^a_0\})(\mathbf{x}^a_0 - E\{\mathbf{x}^a_0\})^H\}\end{aligned}$$

Calculate sigma points for $i = 1, \ldots, 4p$

$$\begin{aligned}\mathcal{X}^a_{0,n-1} &= \widehat{\mathbf{x}}^a_{n-1|n-1}\\ \mathcal{X}^a_{i,n-1} &= \widehat{\mathbf{x}}^a_{n-1|n-1} \pm \left(\sqrt{(p+\lambda)\mathbf{M}^a_{n-1|n-1}}\right)_i\end{aligned} \tag{66}$$

Compute predictions:

$$\begin{aligned}\mathcal{X}^a_{i,n|n-1} &= \mathbf{f}^a[\mathcal{X}^a_{i,n-1}]\\ \widehat{\mathbf{x}}^a_{n|n-1} &= \sum_{i=0}^{4p} \mathcal{W}^{(m)}_i \mathcal{X}^a_{i,n|n-1}\\ \mathbf{M}^a_{n|n-1} &= \mathbf{Q}^a_n + \sum_{i=0}^{4p} \mathcal{W}^{(c)}_i \Big(\mathcal{X}^a_{i,n|n-1} - \widehat{\mathbf{x}}^a_{n|n-1}\Big)\Big(\mathcal{X}^a_{i,n|n-1} - \widehat{\mathbf{x}}^a_{n|n-1}\Big)^H\\ \mathcal{Y}^a_{i,n|n-1} &= \mathbf{h}^a[\mathcal{X}_{i,n|n-1}], \qquad i = 1, \ldots, 4p\\ \widehat{\mathbf{y}}^a_{n|n-1} &= \sum_{i=0}^{4p} \mathcal{W}^{(m)}_i \mathcal{Y}^a_{i,n|n-1}\end{aligned} \tag{67}$$

Measurement update:

$$\begin{aligned}\mathbf{R}^a_{\tilde{\mathbf{y}}^a,n|n-1} &= \mathbf{R}^a_n + \sum_{i=0}^{4p} \mathcal{W}^{(c)}_i \Big(\mathcal{Y}^a_{i,n|n-1} - \widehat{\mathbf{y}}^a_{n|n-1}\Big)\Big(\mathcal{Y}^a_{i,n|n-1} - \widehat{\mathbf{y}}^a_{n|n-1}\Big)^H\\ \mathbf{R}^a_{\mathbf{x}^a\mathbf{y}^a,n|n-1} &= \sum_{i=0}^{4p} \mathcal{W}^{(c)}_i \Big(\mathcal{X}^a_{i,n|n-1} - \widehat{\mathbf{x}}^a_{n|n-1}\Big)\Big(\mathcal{Y}^a_{i,n|n-1} - \widehat{\mathbf{y}}^a_{n|n-1}\Big)^H\\ \mathbf{G}^a_n &= \mathbf{R}^a_{\mathbf{x}^a\mathbf{y}^a,n|n-1}\Big(\mathbf{R}^a_{\tilde{\mathbf{y}}^a,n|n-1}\Big)^{-1}\\ \widehat{\mathbf{x}}^a_{n|n} &= \widehat{\mathbf{x}}^a_{n|n-1} + \mathbf{G}^a_n(\mathbf{y}^a_n - \widehat{\mathbf{y}}^a_{n|n-1})\\ \mathbf{M}^a_{n|n} &= \mathbf{M}^a_{n|n-1} - \mathbf{G}^a_n \mathbf{R}^a_{\tilde{\mathbf{y}}^a,n|n-1} \mathbf{G}^{aH}_n\end{aligned} \tag{68}$$

The unscented and extended Kalman filters use the same general update formula, given by (68) and (53), to compute the estimate of the state, that is

$$\widehat{x}_{n|n} = \widehat{x}_{n|n-1} + g_n(y_n - \widehat{y}_{n|n-1}) \tag{71}$$

where g_n is the Kalman gain. This equation shows that the estimate comprises of a prediction term, $\widehat{x}_{n|n-1}$, and a weighted innovation term, $(y_n - \widehat{y}_{n|n-1})$.

Substituting the state equation (69) in to the observation equation (70) gives

$$y_n = h\big[f[x_{n-1}] + w_n\big] + v_n \tag{72}$$

Let $z = f[x_{n-1}] + w_n$, then the TSE of the function $h[f[x_{n-1}] + w_n] = h[z]$ about $f[x_{n-1}]$ can be written as

$$h[f[x_{n-1}] + w_n] = h[f[x_{n-1}]] + \frac{\partial h}{\partial z} w_n + \frac{1}{2}\mathcal{H}_{zz} w_n^2 + \text{h.o.t.} \tag{73}$$

with the Jacobian $\frac{\partial h}{\partial z}$ and Hessian $\mathcal{H}_{zz} = \frac{\partial}{\partial z}\left(\frac{\partial h}{\partial z}\right)$ evaluated at $z = f[x_{n-1}]$. Now subtract the true state, x_n, from the estimate given in (71) to find the state estimation error

$$\begin{aligned}e_n &= x_n - \widehat{x}_{n|n}\\ &= (f[x_{n-1}] + w_n) - \widehat{x}_{n|n-1} - g_n(y_n - \widehat{y}_{n|n-1})\end{aligned} \tag{74}$$

Substituting (72) and (73) into (74) yields

$$
\begin{aligned}
e_n = &(f[x_{n-1}] + w_n) - \widehat{x}_{n|n-1} - g_n\Big(h[f[x_{n-1}]] \\
&+\frac{\partial h}{\partial z}w_n + \quad \frac{1}{2}\mathcal{H}_{zz}w_n^2 + \text{h.o.t.} + v_n - \widehat{y}_{n|n-1}\Big)
\end{aligned} \tag{75}
$$

Based on (75), the MSE, that is $E\{e_n e_n^*\}$, consists of a large number of terms, however, since we are only interested in the effect of circularity on the MSE, we shall only analyse terms related to the process and measurement noise pseudocovariances and these terms are

$$
\begin{aligned}
E\{e_n e_n^*\} = &-E\Big\{\frac{1}{2}g_n\mathcal{H}_{zz}w_n^2\Big(f[x_{n-1}] - \widehat{x}_{n|n-1}\Big)^*\Big\} \\
&- \quad E\Big\{\frac{1}{2}\Big(f[x_{n-1}] - \widehat{x}_{n|n-1}\Big)g_n^*\mathcal{H}_{zz}^*(w_n^*)^2\Big\} \\
&+ \quad E\Big\{\frac{1}{2}g_n\mathcal{H}_{zz}w_n^2\Big(g_n\big(h[f[x_{n-1}] - \widehat{y}_{n|n-1})\Big)^*\Big\} \\
&+ \quad E\Big\{\frac{1}{2}\Big(g_n\big(h[f[x_{n-1}] - \widehat{y}_{n|n-1})\Big)g_n^*\mathcal{H}_{zz}^*(w_n^*)^2\Big\} \\
&+ \quad (\text{otherterms \& h.o.t.}) \\
= &-\Re\Big\{E\Big\{g_n\mathcal{H}_{zz}\Big(f[x_{n-1}] - \widehat{x}_{n|n-1}\Big)^*\Big\}\rho_{w,n}\Big\} \\
&+ \quad \Re\Big\{E\Big\{\left|g_n\right|^2\mathcal{H}_{zz}\big(h[f[x_{n-1}]-\widehat{y}_{n|n-1})^*\Big\}\rho_{w,n}^*\Big\} \\
&+ \quad (\text{other terms \& h.o.t.})
\end{aligned} \tag{76}
$$

where $\Re\{\cdot\}$ is the real part of a complex quantity.

Remark #6: From (76) it can be seen that the MSE for the CUKF and CEKF are dependent on the pseudocovariance of the state noise, namely it is a function of $\rho_{w,n}$ and $\rho_{w,n}^*$, hence there mean square behavior are affected by the circularity of the state noise, if the observation equation is nonlinear.

Remark #7: If the state space model is a linear, then the Hessian term $\mathcal{H}_{zz}$ in (76) vanishes, as the second derivatives of h is zero, and as a consequence the four terms in the MSE (76) which are dependent on the pseudocovariances also vanish. Therefore, the mean square characteristic of the conventional linear complex Kalman filter does not depend on the circularity of the state or observation noises.

V. Application Examples

To illustrate the benefits of widely linear complex Kalman filters over conventional complex Kalman filters, we shall consider two case studies: (1) filtering for a noisy complex valued autoregressive process and (2) multistep ahead prediction for real-world noncircular and nonstationary wind data and the second order noncircular Lorenz attractor.

A. Complex autoregressive process

The performances of all the standard and widely linear Kalman filters discussed above were used to filter the first order complex autoregressive process, AR(1), given by [1] [19]

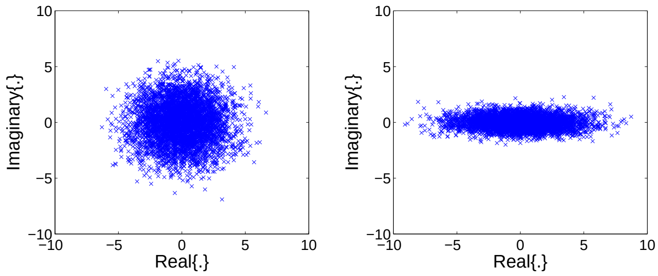
$$x_n = 0.9x_{n-1} + u_n$$

where the driving noise was u_n was a doubly white, Gaussian, zero-mean noise with variance and pseudocovariance defined as

$$
\begin{aligned}
E\{u_{n-i}u_{n-l}^*\} &= c_u\delta_{i-l} \\
E\{u_{n-i}u_{n-l}\} &= \rho_u\delta_{i-l}
\end{aligned}
$$

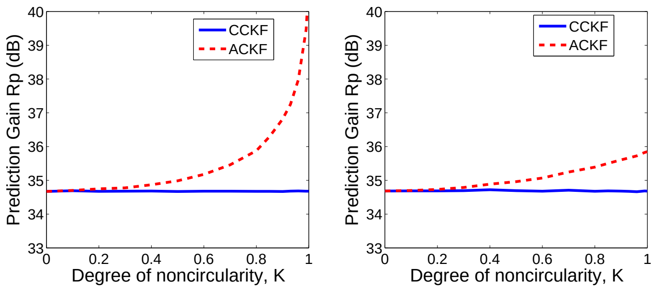
where δ is the discrete Dirac delta function. The observation equation for the linear filters, namely the CCKF and ACKF, were such that x_n was observed in the presence of an additive complex noncircular white noise, v_n, that is (see [9] for the Kalman filter implementation of an autoregressive process)

$$y_n = x_n + v_n$$



(a) A circular Gaussian distribution (b) A noncircular Gaussian distribution

Fig. 2. A geometric view of circularity via a real-imaginary scatter plot of the AR(1) process driven by (a) a circular ($K = 0$) and (b) a noncircular ($K = 0.9$) Gaussian distributions.



(a) Noncircular state noise (b) Noncircular observation noise

Fig. 3. Performance of CKF and ACKF for the AR(1) process with varying degrees of state and observation noises noncircularity

while, the observation equation corresponding to the nonlinear CEKF, CUKF and their corresponding augmented versions, was given by

$$y_n = \arctan[x_n] + v_n$$

The ratio of pseudocovariance magnitude to covariance, that is

$$K = \frac{|\rho|}{c}$$

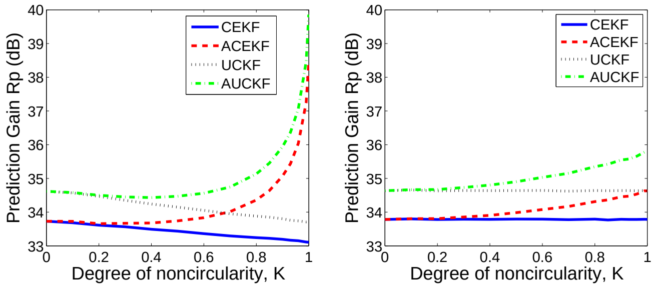
was used as a measure for the degree of noncircularity of the complex state and measurement noises [20], where a complex random variable is circular if $K = 0$ and maximally noncircular if $K = 1$. Figure 2 shows a real-imaginary scatter plot for two different realisations of the AR(1) process driven by Gaussian complex white variable with different levels of circularity. Note the circular symmetry for the circular signal and the non-circular shape for $K = 0.9$. For a quantitative assessment of the performance, the standard prediction gain $R_p = 10\log(\sigma_y^2/\sigma_e^2)$ was used, where σ_y^2 and σ_e^2 are the powers of the input signal and the output error.

Figure 3 shows the performances of the standard CCKF and its corresponding widely linear (augmented) version, the ACKF. Figure 3a illustrates the results for a circular observation noise and a state noise with various degrees of noncircularity, while Figure 3b shows the results for a noncircular observation noise with a circular state noise. For both sets of simulations, when the noises were circular the ACKF had the same performance as the CCKF, while for noncircular noises the ACKF had superior performance as the degree of noise noncircularity (K) increased.

Figure 4 shows the corresponding results for the nonlinear CEKF, CUKF and their corresponding augmented versions, ACEKF and ACUKF. Similar to the ACKF, the general behavior is that ACEKF and ACUKF outperform the CEKF and CUKF, respectively, if the either of the state or observation noises are noncircular, while for circular noises they have similar performances. However, when the state noise is noncircular, as illustrated in Figure 4a, the MSE behavior of CEKF and CUKF change with the circularity of the state noise.

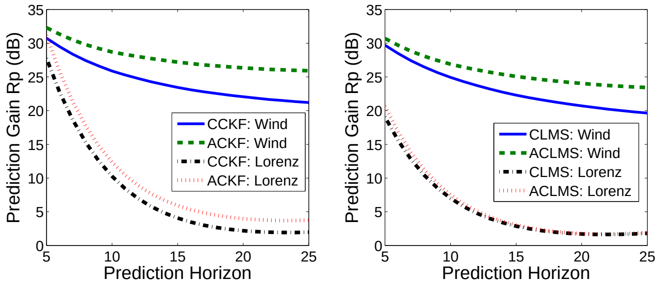
B. Multistep ahead prediction of different signals

The performances of the CCKF and ACKF were next assessed for the multistep ahead prediction of the noncircular *Lorenz* attractor and some real world noncircular and nonstationary *Wind* data using linear and widely linear autoregressive models. Simulations for the complex least mean square (CLMS) and its augmented version, the ACLMS, were also carried out to provide a performance comparison [21].



(a) Noncircular state noise (b) Noncircular observation noise

Fig. 4. Performance comparison between CEKF, CUKF and their corresponding augmented versions for the AR(1) process with varying degrees of state and observation noises noncircularity



(a) CCKF and ACKF (b) CLMS and ACLMS

Fig. 5. Multistep ahead prediction of real-world Wind data and the Lorenz attractor

Figure 5a summarises the prediction performances for the *Lorenz* and the *Wind* data. The ACKF was able to capture the underlying dynamics of the signals better than CCKF, which is indicated by its superior prediction performance. This can be attributed to the use of the widely linear 'augmented' model, which is better able to capture the second order statistics of the noncircular signals. Similarly, the prediction performances for the ACEKF and ACUKF were also superior to the CEKF and CUKF, but were not shown here in order to avoid repetition. Figure 5b shows the corresponding simulations for the CLMS and ACLMS, where the ACLMS is shown to have superior performance compared to the CLMS, but is worse off compared to the ACKF.

VI. Conclusion

The second order statistics of zero-mean complex signals are described by their covariance function and a second moment function known as the pseudocovariance. With the aim of fully utilising both these statistical moments, we have readdressed the augmented complex Kalman filter (ACKF) algorithm and have examined its performance in relation to the conventional complex Kalman filter (CCKF). The analysis has shown that the ACKF has the potential to offer significant performance gains over the CCKF for noncircular state or observation noises and the same performance as the CCKF for circular signals. Moreover, we have analysed a more general form of the augmented complex extended Kalman filter (ACEKF), by using the so called $\mathbb{CR}$ calculus framework, allowing us to deal with both analytic and nonanalytic state space models. We have also analysed the mean square characteristics of CCKF, the extended complex Kalman filter (CEKF), the unscented complex Kalman filter (CUKF), and have shown that the mean square behavior of the CCKF is unaffected by the noncircularity of the state and observation noises. The analysis shows that the mean square characteristics of the CEKF and CUKF are effected by state noise noncircularity, if the observation equation is nonlinear.

Appendix A
A consise summary of $\mathbb{CR}$ calculus

In what follows, the $\mathbb{CR}$ calculus framework, which was originally introduced by Wirtinger and is known as Wirtinger calculus within the German speaking engineering community, is briefly introduced. More recently Kreutz-Delgado [13] provided a comprehensive overview of the topic and came with the name $\mathbb{CR}$ calculus due to the dual real and complex perspective of complex functions within this framework.

In order for the complex derivative of a function of $\mathbf{z} = \mathbf{z}_r + j\mathbf{z}_i$

$$f[\mathbf{z}] = u[\mathbf{z}_r, \mathbf{z}_i] + jv[\mathbf{z}_r, \mathbf{z}_i] \tag{A-1}$$

to exist in the standard sense, the real partial derivatives of $u[\mathbf{z}_r, \mathbf{z}_i]$ and $v[\mathbf{z}_r, \mathbf{z}_i]$ must satisfy the Cauchy-Riemann equations given by

$$\frac{\partial u}{\partial \mathbf{z}_r} = \frac{\partial v}{\partial \mathbf{z}_i} \qquad\qquad \frac{\partial v}{\partial \mathbf{z}_r} = -\frac{\partial u}{\partial \mathbf{z}_i} \tag{A-2}$$

For example the function $f_1[z] = z^2$ is complex differentiable, where $\frac{\partial f}{\partial z} = \frac{\partial u}{\partial z_r} + j\frac{\partial v}{\partial z_i}$, while $f_2[z] = zz^*$ does not satisfy the Cauchy-Riemann equations and is not complex differentiable in this light. However by using $\mathbb{CR}$ calculus, which establishes a duality between the real- and complex-valued derivatives, allows for the Taylor series expansion (TSE) in $\mathbb{R}$ and $\mathbb{C}$.

The function $f[\mathbf{z}]$ can be seen as a function of both $\mathbf{z}$ and its complex conjugate $\mathbf{z}^*$, that is $f[\mathbf{z}, \mathbf{z}^*]$. Although $\mathbf{z}$ and $\mathbf{z}^*$ are not truly independent, the introduced methodology can be considered as a formalism whereby f is analytic in $\mathbf{z}$ for fixed $\mathbf{z}^*$ and vice versa where f is analytic in $\mathbf{z}^*$ for fixed $\mathbf{z}$. The variables $\mathbf{z}$ and $\mathbf{z}^*$ are called conjugate coordinates and the function $f[\mathbf{z}]$ can be expressed as

$$f[\mathbf{z}] = f[\mathbf{z}, \mathbf{z}^*] = g[\mathbf{z}_r, \mathbf{z}_i] = u[\mathbf{z}_r, \mathbf{z}_i] + jv[\mathbf{z}_r, \mathbf{z}_i] \tag{A-3}$$

The relations between the partial derivatives $\frac{\partial f}{\partial \mathbf{z}}$ and $\frac{\partial f}{\partial \mathbf{z}^*}$, and the partial derivatives $\frac{\partial f}{\partial \mathbf{z}_r}$ and $\frac{\partial f}{\partial \mathbf{z}_i}$ was proven by Brandwood in [12], and an alternative approach based on the total differential of f is shown below [1].

The total differential of the function $g[\mathbf{z}_r, \mathbf{z}_i]$ can be expressed as

$$dg[\mathbf{z}_r, \mathbf{z}_i] = \frac{\partial g}{\partial \mathbf{z}_r} d\mathbf{z}_r + \frac{\partial g}{\partial \mathbf{z}_i} d\mathbf{z}_i \tag{A-4}$$

and after some algebraic manipulations, it can be expressed as

$$dg[\mathbf{z}_r, \mathbf{z}_i] = \frac{1}{2}\Big(\frac{\partial g}{\partial \mathbf{z}_r} - j\frac{\partial g}{\partial \mathbf{z}_i}\Big)d\mathbf{z} + \frac{1}{2}\Big(\frac{\partial g}{\partial \mathbf{z}_r} + j\frac{\partial g}{\partial \mathbf{z}_i}\Big)d\mathbf{z}^* \tag{A-5}$$

or in a more compact form

$$df[\mathbf{z}] = \frac{\partial f}{\partial \mathbf{z}} d\mathbf{z} + \frac{\partial f}{\partial \mathbf{z}^*} d\mathbf{z}^* \tag{A-6}$$

This leads us to one on the important results of $\mathbb{CR}$ calculus, given by

$$\begin{aligned} &\mathbb{R}\text{-derivative:} && \frac{\partial f}{\partial \mathbf{z}}\Big|_{\mathbf{z}^*=constant} = \frac{1}{2}\Big(\frac{\partial f}{\partial \mathbf{z}_r} - j\frac{\partial f}{\partial \mathbf{z}_i}\Big) \\ &\mathbb{R}^*\text{-derivative:} && \frac{\partial f}{\partial \mathbf{z}^*}\Big|_{\mathbf{z}=constant} = \frac{1}{2}\Big(\frac{\partial f}{\partial \mathbf{z}_r} + j\frac{\partial f}{\partial \mathbf{z}_i}\Big) \end{aligned} \tag{A-7}$$

For these generalised derivatives it is assumed that $\mathbf{z}$ and $\mathbf{z}^*$ are mutually independent, namely $\frac{\partial \mathbf{z}}{\partial \mathbf{z}} = \frac{\partial \mathbf{z}^*}{\partial \mathbf{z}^*} = 1$ and $\frac{\partial \mathbf{z}}{\partial \mathbf{z}^*} = \frac{\partial \mathbf{z}^*}{\partial \mathbf{z}} = 0$. Hence it is possible to consider the derivatives of both analytic and non-analytic functions , where the functions can be either complex- or real-valued. The $\mathbb{R}$-derivative and $\mathbb{R}^*$-derivative for analytic functions satisfying the Cauchy-Riemann equations simplify to

$$\begin{aligned} &\mathbb{R}\text{-derivative:} && \frac{\partial f}{\partial \mathbf{z}}\Big|_{\mathbf{z}^*=constant} &= \frac{1}{2}\Big(2\frac{\partial f}{\partial \mathbf{z}_r} + j2\frac{\partial f}{\partial \mathbf{z}_i}\Big) = f^{'}[\mathbf{z}] \\ &\mathbb{R}^*\text{-derivative:} && \frac{\partial f}{\partial \mathbf{z}^*}\Big|_{\mathbf{z}=constant} &= 0 \end{aligned} \tag{A-8}$$

that is, for analytic functions the $\mathbb{R}$-derivative is equivalent to the standard complex derivative and the $\mathbb{R}^*$-derivative vanishes, namely the function is independent of $\mathbf{z}^*$. The Cauchy-Riemann condition can then be expressed as

$$\frac{\partial f}{\partial \mathbf{z}^*} = \mathbf{0} \tag{A-9}$$

Therein lies the beauty of the $\mathbb{CR}$ calculus framework: for when applied to an analytic function it is equal to the standard complex derivative and when applied to a non-analytic function it is equal to the pseudo-gradient ($\mathbb{R}^*$-derivative).

Next, the Taylor series expansion of $f[\mathbf{z}]$ up to the 2nd order is considered. This is facilitated by considering the equivalent forms of the function f, that is

$$f[\mathbf{z}] \longleftrightarrow f[\mathbf{z}, \mathbf{z}^*] \equiv f[\mathbf{z}^a] \longleftrightarrow f[\mathbf{z}_r, \mathbf{z}_i] \equiv f[\mathbf{r}] \tag{A-10}$$

and establishing the duality between the derivatives of the functions in $\mathbb{R}^{2N}$ and $\mathbb{C}^{2N}$. In (A-10) the augmented vector $\mathbf{z}^a = \begin{bmatrix}\mathbf{z}^T & \mathbf{z}^H\end{bmatrix} \in \mathbb{C}^{2N}$ and $\mathbf{r} = \begin{bmatrix}\mathbf{z}_r^T & \mathbf{z}_i^T\end{bmatrix} \in \mathbb{R}^{2N}$. By establishing the isomorphism between vectors in $\mathbb{R}^{2N}$ and $\mathbb{C}^{2N}$ and identifying the Jacobian of the transformation, the first and second order derivatives for the terms in the TSE can be readily computed [13].

From the fact that $\mathbf{z} = \mathbf{z}_r + j\mathbf{z}_i$ and $\mathbf{z}^* = \mathbf{z}_r - j\mathbf{z}_i$ it can be shown that

$$\mathbf{z}^a = \begin{bmatrix} \mathbf{z} \\ \mathbf{z}^* \end{bmatrix} = \begin{bmatrix} \mathbf{I} & j\mathbf{I} \\ \mathbf{I} & -j\mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{z_r} \\ \mathbf{z_i} \end{bmatrix} \tag{A-11}$$

where $\mathbf{I}$ is the identity matrix. Now define

$$\mathbf{J} \equiv \begin{bmatrix} \mathbf{I} & j\mathbf{I} \\ \mathbf{I} & -j\mathbf{I} \end{bmatrix} \tag{A-12}$$

then it is easy to show that $\mathbf{J}^{-1} = \frac{1}{2}\mathbf{J}^H$ and we have the following mappings

$$\mathbf{z}^a = \mathbf{J}\mathbf{r} \qquad \text{and} \qquad \mathbf{r} = \mathbf{J}^{-1}\mathbf{z}^a = \frac{1}{2}\mathbf{J}^H\mathbf{z}^a \tag{A-13}$$

Because the mapping between $\mathbb{R}^{2N}$ and $\mathbb{C}^{2N}$ is linear and one to one, then these two spaces can be considered isomorphic. The mappings in (A-13) therefore correspond to an admissible coordinate transformation between the $\mathbf{z}^a$ and $\mathbf{r}$ representations of $\mathbf{z}$. Using the chain rule and the mappings in (A-13) , the partial derivatives between the two space can now be written as

$$\begin{aligned} \frac{\partial}{\partial \mathbf{z}^a} &= \frac{1}{2}\frac{\partial}{\partial \mathbf{r}}\mathbf{J}^H \\ \frac{\partial}{\partial \mathbf{r}} &= \frac{\partial}{\partial \mathbf{z}^a}\mathbf{J} \end{aligned} \tag{A-14}$$

The TSE for $\mathbb{R}^{2N}$ up to the second order term can be expressed as

$$f[\mathbf{r} + \Delta\mathbf{r}] = f[\mathbf{r}] + \frac{\partial f}{\partial \mathbf{r}}\Delta\mathbf{r} + \frac{1}{2}\Delta\mathbf{r}^T\mathcal{H}_{\mathbf{rr}}\Delta\mathbf{r} \tag{A-15}$$

where $\mathcal{H}_{\mathbf{rr}} = \frac{\partial}{\partial \mathbf{r}}\left(\frac{\partial f}{\partial \mathbf{r}}\right)^T$ is the real-valued Hessian matrix. The corresponding first order term in the augmented complex space is given by

$$\frac{\partial f}{\partial \mathbf{r}}\Delta\mathbf{r} = \mathbf{J}\cdot\mathbf{J}^{-1}\Delta\mathbf{z}^a = \frac{\partial f}{\partial \mathbf{z}^a}\Delta\mathbf{z}^a \tag{A-16}$$

The augmented complex Hessian matrix is given by

$$\mathcal{H}^a_{\mathbf{zz}} = \frac{\partial}{\partial \mathbf{z}^a}\left(\frac{\partial f}{\partial \mathbf{z}^a}\right)^H = \begin{bmatrix} \mathcal{H}_{\mathbf{zz}} & \mathcal{H}_{\mathbf{z}^*\mathbf{z}} \\ \mathcal{H}_{\mathbf{zz}^*} & \mathcal{H}_{\mathbf{z}^*\mathbf{z}^*} \end{bmatrix}, \tag{A-17}$$

where $\mathcal{H}_{\mathbf{ab}} = \frac{\partial}{\partial \mathbf{a}}\left(\frac{\partial f}{\partial \mathbf{b}}\right)^H$ with $\mathbf{a}, \mathbf{b} \in [\mathbf{z}, \mathbf{z}^*]$. The relationship between $\mathcal{H}_{\mathbf{rr}}$ and $\mathcal{H}^a_{\mathbf{zz}}$ can be expressed as

$$\mathcal{H}_{\mathbf{rr}} = \mathbf{J}^H\mathcal{H}^a_{\mathbf{zz}}\mathbf{J}. \tag{A-18}$$

Hence, the second-order term in the augmented complex TSE is computed as

$$\frac{1}{2}\Delta\mathbf{r}^T\mathcal{H}_{\mathbf{rr}}\Delta\mathbf{r} = \frac{1}{2}\Delta\mathbf{z}^{aH}\mathcal{H}^a_{\mathbf{zz}}\Delta\mathbf{z}^a. \tag{A-19}$$

Combining the results so far allows the TSE in the $\mathbb{C}^{2N}$ up to the second-order term to be written as

$$f[\mathbf{z}^a + \Delta\mathbf{z}^a] = f[\mathbf{z}^a] + \frac{\partial f}{\partial \mathbf{z}^a}\Delta\mathbf{z}^a + \frac{1}{2}\Delta\mathbf{z}^{aH}\mathcal{H}^a_{\mathbf{zz}}\Delta\mathbf{z}^a \tag{A-20}$$

By expanding the complex augmented vector $\mathbf{z}^a$ in terms of the conjugate coordinates $\mathbf{z}$ and $\mathbf{z}^*$ and using that $\frac{\partial}{\partial \mathbf{z}^a} = [\frac{\partial}{\partial \mathbf{z}}, \frac{\partial}{\partial \mathbf{z}^*}]$, the complex TSE can be expressed in terms of $\mathbf{z}$ and $\mathbf{z}^*$ as

$$\begin{aligned} f[\mathbf{z} + \Delta\mathbf{z}] = f[\mathbf{z}] + \frac{\partial f}{\partial \mathbf{z}}\Delta\mathbf{z} + \frac{\partial f}{\partial \mathbf{z}^*}\Delta\mathbf{z}^* + \frac{1}{2}\Big(&\Delta\mathbf{z}^H\mathcal{H}_{\mathbf{zz}}\Delta\mathbf{z} + \Delta\mathbf{z}^H\mathcal{H}_{\mathbf{z}^*\mathbf{z}}\Delta\mathbf{z}^* + \Delta\mathbf{z}^{*H}\mathcal{H}_{\mathbf{zz}^*}\Delta\mathbf{z} \\ &+\Delta\mathbf{z}^{*H}\mathcal{H}_{\mathbf{z}^*\mathbf{z}^*}\Delta\mathbf{z}^*\Big) \end{aligned} \tag{A-21}$$

References

[1] D. P. Mandic and V. S. L. Goh, *Complex Valued Nonlinear Adaptive Filters: Noncircularity, Widely Linear and Neural Models*. Wiley, 2009.

[2] P. Dash, A. Pradhan, and G. Panda, "Frequency estimation of distorted power system signals using extended complex Kalman filter," *IEEE Transactions on Power Delivery*, vol. 14, pp. 761 –766, July 1999.

[3] V. S. L. Goh and D. P. Mandic, "An Augmented Extended Kalman Filter Algorithm for Complex-Valued Recurrent Neural Networks," *Neural Computation*, vol. 19, pp. 1039–1055, 2007.

[4] F. Cattivelli and A. Sayed, "Distributed nonlinear Kalman filtering with applications to wireless localization," in *IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)*, pp. 3522 –3525, 2010.

[5] B. Picinbono and P. Bondon, "Second-order Statistics of Complex Signals," *IEEE Transactions on Signal Processing*, vol. 45, no. 2, pp. 411–420, 1997.

[6] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice Hall International, Inc, 1993.

[7] D. H. Dini and D. P. Mandic, "Analysis of the widely linear complex Kalman filter," in *Sensor Signal Processing for Defence Conference*, 2010.

[8] S. Haykin, *Adaptive Filter Theory*. Prentice Hall International, 3rd ed. ed., 1996.

[9] M. H. Hayes, *Statistical Digital Signal Processing and Modeling*. John Wiley & Sons, 1996.

[10] D. H. Dini and D. P. Mandic, "A class of widely linear complex Kalman filters," in *IEEE Transactions on Neural Networks*, submitted.

[11] B. Picinbono and P. Chevalier, "Widely Linear Estimation with Complex Data," *IEEE Transactions on Signal Processing*, vol. 43, no. 8, pp. 2030–2033, 1995.

[12] D. H. Brandwood, "A Complex Gradient Operator and its Application in Adaptive Array Theory," tech. rep., Plessey Electronic Systems Research Ltd., Romsey, UK, 1983.

[13] K. Kreutz-Delgado, "The Complex Gradient Operator and the CR-Calculus," Tech. Rep. ECE275CG-F05v1.3d, Electrical and Computer Engineering, Jacobs School of Engineering, University of California, San Diego, 2006.

[14] D. H. Dini and D. P. Mandic, "Widely linear complex extended Kalman filters," in *Sensor Signal Processing for Defence Conference*, submitted.

[15] D. H. Dini, C. Jahanchahi, and D. P. Mandic, "Widely linear complex and quaternion valued bearings only tracking," in *IET Signal Processing Special Issue on: Multi-Sensor Signal Processing for Defence*, submitted, Feb 2011.

[16] S. Julier and J. Uhlmann, "A new extension of the Kalman filter to nonlinear systems," *Proceedings of AeroSense: The 11th International Symposium on Aerospace/Defense Sensing, Simulation and Controls*, pp. 182–193, 1997.

[17] S. Julier, J. Uhlmann, and H. Durrant-Whyte, "A New Method for the Nonlinear Transformation of Means and Covariances in Filters and Estimators," *IEEE Transactions on Automatic Control*, vol. 45, no. 3, pp. 477–482, 2000.

[18] D. H. Dini, D. P. Mandic, and S. J. Julier, "The widely linear complex unscented Kalman filter," in *IEEE Signal Processing Letters*, submitted.

[19] J. Navarro-Moreno, "ARMA prediction of widely linear systems by using the innovations algorithm," *IEEE Transactions on Signal Processing*, vol. 56, pp. 3061 –3068, july 2008.

[20] E. Ollila, "On the circularity of a complex random variable," *IEEE Signal Processing Letters*, vol. 15, pp. 841–844, 2008.

[21] S. Javidi and M. Pedzisz and V. S. L. Goh and D. P. Mandic, "The augmented complex least mean square algorithm with application to adaptive prediction problems," *Proc. 1st IARP Workshop Cogn. Inf. Process*, pp. 54–57, 2008.