# Class of Widely Linear Complex Kalman Filters

Dahir H. Dini, *Student Member, IEEE*, and Danilo P. Mandic, *Senior Member, IEEE*

***Abstract*— Recently, a class of widely linear (augmented) complex-valued Kalman filters (KFs), that make use of augmented complex statistics, have been proposed for sequential state space estimation of the generality of complex signals. This was achieved in the context of neural network training, and has allowed for a unified treatment of both second-order circular and noncircular signals, that is, both those with rotation invariant and rotation-dependent distributions. In this paper, we revisit the augmented complex KF, augmented complex extended KF, and augmented complex unscented KF in a more general context, and analyze their performances for different degrees of noncircularity of input and the state and measurement noises. For rigor, a theoretical bound for the performance advantage of widely linear KFs over their strictly linear counterparts is provided. The analysis also addresses the duality with bivariate real-valued KFs, together with several issues of implementation. Simulations using both synthetic and real world proper and improper signals support the analysis.**

***Index Terms*— Augmented complex Kalman filter, complex circularity, complex Kalman filter, extended Kalman filter, unscented Kalman filter, widely linear model.**

## I. Introduction

COMPLEX-VALUED signals arise in a variety of applications, from communication systems to radar and electric grid. A complex or quaternion representation of bivariate real-valued signals, such as in the case of multidimensional wind data [1]–[4], may also provide a convenient analysis framework, as well as natural means to understand some fundamental characteristics of such signals and the transformations they undergo, for instance phase and magnitude distortions.

The second-order statistical properties of complex signals are characterized by their covariance and pseudocovariance functions. The covariance captures the information concerning the total power of the signal, while the pseudocovariance conveys 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, based solely on the covariance, making them suited to only second-order circular (proper) complex signals, that is, those with rotation invariant probability distributions. Proper signals are characterized by a vanishing pseudocovariance, however, most real-world signals are almost invariably second-order noncircular (improper); exhibiting different signal powers in the real and imaginary parts, correlation of the real and imaginary parts [1], or finite sample size.

The Kalman filter (KF) is a sequential state space estimation technique, and is an essential part of space navigation, military technology development, neural network training [5]–[8], and sensor fusion [9]. Complex-valued KFs have also been used extensively [10]–[13], however, the traditional implementation of the complex-valued KF inherently assumes second-order circular (proper) state and measurement noises as well as input data, and as such does not fully utilize the full available second-order statistics of the complex signals. In our earlier work [1], [12], we introduced widely linear KFs in the context of neural network training, but did not elaborate on their convergence or operation in general augmented state spaces.

The recent introduction of so called ‘augmented complex statistics’ [1], [14] has highlighted that for a general (improper) complex vector $\mathbf{x}$, estimation based solely on the covariance matrix $\mathbf{R_x} = E\{\mathbf{x}\mathbf{x}^H\}$ is inadequate, and the pseudocovariance matrix $\mathbf{P_x} = E\{\mathbf{x}\mathbf{x}^T\}$ 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 $\mathbf{H}$. Since both the real $\mathbf{y}_r$ and imaginary $\mathbf{y}_i$ parts of the vector $\mathbf{y}$ are real valued, we have

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

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

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

Manuscript received June 29, 2011; revised February 23, 2012; accepted February 26, 2012. Date of publication April 3, 2012; date of current version May 2, 2012. This work was supported in part by the University Defence Research Centre at Imperial College London, sponsored by the MoD and DSTL.

The authors are with the Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2BT, U.K. (e-mail: dahir.dini@imperial.ac.uk; d.mandic@imperial.ac.uk).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TNNLS.2012.2189893

[1]The ‘widely linear’ model is associated with the signal generating system, whereas ‘augmented statistics’ describe the 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.’

and as such, estimation based on $\mathbf{R}^a_{\mathbf{x}}$ applies to both proper and improper data.

The recently introduced widely linear (augmented) complex Kalman filter (ACKF), augmented complex extended Kalman filter (ACEKF) [12], and augmented complex unscented Kalman filter (ACUKF) [1] have been shown to be suitable for the generality of complex signals, both second-order circular and noncircular. They were applied for the training of neural networks and exhibited superior performance when compared with their strictly linear counterparts for improper data. However, the properties and performances of these filters were not elaborated for the general case, and outside neural network training, where the sources of improperness include both the input data and system parameters, and for general data models.

In this paper, we consolidate and expand our recent work on a class of widely linear KFs [1] and illuminate their performances under general noncircular state and observation noises, and for nonholomorphic state and observation models. 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) is analyzed in depth, and the Cramer–Rao lower bound (CRLB) for the widely linear KFs is established. 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. Simulations on both benchmark and real-world noncircular data support the analysis.

## II. ACKF

Consider the linear state space model given by [15]

$$\mathbf{x}_n = \mathbf{F}_{n-1}\mathbf{x}_{n-1} + \mathbf{w}_n \tag{6a}$$
$$\mathbf{y}_n = \mathbf{H}_n\mathbf{x}_n + \mathbf{v}_n \tag{6b}$$

where $\mathbf{x}_n$ is the state to be estimated (of dimension $p \times 1$), $\mathbf{y}_n$ is the noisy observation (of dimension $q \times 1$), the vectors $\mathbf{w}_n$ and $\mathbf{v}_n$ are the zero-mean state and measurement noises[2] with covariances $\mathbf{Q}_n$ and $\mathbf{R}_n$ and pseudocovariances $\mathbf{P}_n$ and $\mathbf{U}_n$, respectively. The matrices $\mathbf{F}_n$ and $\mathbf{H}_n$ are the state transition and observation matrices. From (4), the corresponding widely linear state space model is defined as[3]

$$\mathbf{x}_n = \mathbf{F}_{n-1}\mathbf{x}_{n-1} + \mathbf{A}_{n-1}\mathbf{x}^*_{n-1} + \mathbf{w}_{n-1} \tag{7a}$$
$$\mathbf{y}_n = \mathbf{H}_n\mathbf{x}_n + \mathbf{B}_n\mathbf{x}^*_n + \mathbf{v}_n \tag{7b}$$

and can be expressed in a compact form using so called "augmented" complex vectors, such that [1]

$$\mathbf{x}^a_n = \mathbf{F}^a_{n-1}\mathbf{x}^a_{n-1} + \mathbf{w}^a_n \tag{8a}$$
$$\mathbf{y}^a_n = \mathbf{H}^a_n\mathbf{x}^a_n + \mathbf{v}^a_n \tag{8b}$$

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

[3]The state and observation noises can also be widely linear, in which case: $\mathbf{w}_n = \mathbf{C}_n\mathbf{u}_n + \mathbf{D}_n\mathbf{u}^*_n$ and $\mathbf{v}_n = \mathbf{E}_n\mathbf{n}_n + \mathbf{F}_n\mathbf{n}^*_n$, where $\mathbf{C}, \mathbf{D}, \mathbf{E}$, and $\mathbf{F}$ are coefficient matrices and $\mathbf{u}, \mathbf{n}$ are the noise models.

**Algorithm 1** The ACKF algorithm

*Initialize with:*

$$\hat{\mathbf{x}}^a_{0|0} = E\{\mathbf{x}^a_0\}$$
$$\mathbf{M}^a_{0|0} = E\left\{(\mathbf{x}^a_0 - E\{\mathbf{x}^a_0\})(\mathbf{x}^a_0 - E\{\mathbf{x}^a_0\})^H\right\}$$

*Prediction:*

$$\hat{\mathbf{x}}^a_{n|n-1} = \mathbf{F}^a_{n-1}\hat{\mathbf{x}}^a_{n-1|n-1} \tag{9}$$

*Minimum Prediction MSE Matrix:*

$$\mathbf{M}^a_{n|n-1} = \mathbf{F}^a_{n-1}\mathbf{M}^a_{n-1|n-1}\mathbf{F}^{aH}_{n-1} + \mathbf{Q}^a_n \tag{10}$$

*Kalman Gain Matrix:*

$$\mathbf{G}^a_n = \mathbf{M}^a_{n|n-1}\mathbf{H}^{aH}_n\left(\mathbf{H}^a_n\mathbf{M}^a_{n|n-1}\mathbf{H}^{aH}_n + \mathbf{R}^a_n\right)^{-1} \tag{11}$$

*Correction:*

$$\hat{\mathbf{x}}^a_{n|n} = \hat{\mathbf{x}}^a_{n|n-1} + \mathbf{G}^a_n\left(\mathbf{y}^a_n - \mathbf{H}^a_n\hat{\mathbf{x}}^a_{n|n-1}\right) \tag{12}$$

*Minimum MSE Matrix:*

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

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

The conjugate matrices $\mathbf{A}$ and $\mathbf{B}$ in (7) and (8) determine whether the state and observation equations are strictly linear or widely linear. For instance, if $\mathbf{A} = \mathbf{0}$ and $\mathbf{B} = \mathbf{0}$, the state space becomes strictly linear, however, the augmented state space representation should still be preferred over the linear state space, in order to cater for noncircular state and observation noises. This was not considered in earlier widely linear KFs.

Consider the augmented covariance matrices of $\mathbf{w}^a_n = [\mathbf{x}^T_n, \mathbf{w}^H_n]^T$ and $\mathbf{v}^a_n = [\mathbf{v}^T_n, \mathbf{v}^H_n]^T$, that is

$$\mathbf{Q}^a_n = E\{\mathbf{w}^a_n\mathbf{w}^{aH}_n\} = \begin{bmatrix}\mathbf{Q}_n & \mathbf{P}_n \\ \mathbf{P}^*_n & \mathbf{Q}^*_n\end{bmatrix} \tag{14}$$
$$\mathbf{R}^a_n = E\{\mathbf{v}^a_n\mathbf{v}^{aH}_n\} = \begin{bmatrix}\mathbf{R}_n & \mathbf{U}_n \\ \mathbf{U}^*_n & \mathbf{R}^*_n\end{bmatrix} \tag{15}$$

where the pseudocovariances are the off-diagonal block terms. The ACKF is a minimum MSE estimator $\hat{\mathbf{x}}^a_{n|n} = E[\mathbf{x}^a_n|\mathbf{y}^a_0, \mathbf{y}^a_1, \ldots, \mathbf{y}^a_n]$ of $\mathbf{x}^a_n$ based on the observations $\{\mathbf{y}^a_0, \mathbf{y}^a_1, \ldots, \mathbf{y}^a_n\}$, and is summarized in Algorithm 1.

### A. Convergence Analysis

When the state and observation noises are circular, and the state and observation equations strictly linear, the CCKF and ACKF are equivalent, that is, they yield the same state estimate and MSE at every time instant. These conditions can be summarized as follows:

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

The correspondence between CCKF and ACKF for proper data and under the same initialization (even when initial filter parameters are incorrect) can be illustrated as follows. Consider the predicted MMSE matrix, which can be expressed as a Riccati recursion, that is

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

Notice 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 measurements are observed. By substituting (11) into (13) and using the matrix inversion lemma, the augmented MSE matrix $\mathbf{M}^a_{n|n}$ can be expressed as

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

Substituting (18) into (11) allows for the Kalman gain to be expressed as

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

For the same state and MSE initialization of both CCKF and ACKF (even when these quantities are not necessarily the true values), we have

$$\begin{aligned}\widehat{\mathbf{x}}^a_{0|0} &= \left[\widehat{\mathbf{x}}^T_{0|0},\widehat{\mathbf{x}}^H_{0|0}\right]^T\\ \mathbf{M}^a_{0|0} &= \begin{bmatrix}\mathbf{M}_{0|0} & \mathbf{0}\\ \mathbf{0} & \mathbf{M}^*_{0|0}\end{bmatrix}\end{aligned}$$

where $\widehat{\mathbf{x}}_{0|0}$ and $\mathbf{M}_{0|0}$ are, respectively, the initial state and MSE for the strictly linear CCKF. Substituting the expressions in (16) into (19) yields

$$\mathbf{G}^a_n = \begin{bmatrix}\mathbf{G}_n & \mathbf{0}\\ \mathbf{0} & \mathbf{G}^*_n\end{bmatrix} \tag{20}$$

where $\mathbf{G}_n = \mathbf{M}_{n|n}(\mathbf{H}_n)^H(\mathbf{R}_n)^{-1}$ is the Kalman gain for the CCKF at time instant $n$. Observe that CCKF and ACKF have the same Kalman gain, even though for ACKF it takes a block-conjugate structure, by substituting (20) into (12) it follows that the two filters yield identical state estimates.

**Remark 1:** For circular state and observation noises with strictly linear state and observation equations (that is, the conditions illustrated in (16)), the ACKF and CCKF have identical performances, and consequently the same convergence properties, when identically initialized.

### *B. MSE Performance Analysis*

We next illuminate the MSE performances of the CCKF and ACKF in order to provide insight into the behavior of KFs for the generality of complex signals, both second-order circular and noncircular. From the state space model given by (6a) and (6b), the KF 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 observation sequence, $\mathbf{z}_n = \left[\mathbf{y}^T_1,\mathbf{y}^T_2,\ldots,\mathbf{y}^T_n\right]^T$, that is

$$\widehat{\mathbf{x}}_{n|n} = E\{\mathbf{x}_0\}+\mathbf{W}_n\mathbf{z}_n \tag{21}$$

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

$$\mathbf{W}_n = \mathbf{R}_{\mathbf{xz},n,n}\mathbf{R}^{-1}_{\mathbf{z},n} \tag{22}$$

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

$$\begin{aligned}\mathbf{M}_{n|n} &= E\left\{(\mathbf{x}_n-\widehat{\mathbf{x}}_{n|n})(\mathbf{x}_n-\widehat{\mathbf{x}}_{n|n})^H\right\}\\ &= \mathbf{R}_{\mathbf{x},n}-\mathbf{R}_{\mathbf{xz},n,n}\mathbf{R}^{-1}_{\mathbf{z},n}\mathbf{R}^H_{\mathbf{xz},n,n}.\end{aligned} \tag{23}$$

The KF is summarized by state estimate (mean) in (21) and covariance estimate in (23), although the computational complexity (dimensions) of these expressions increase with time, nonetheless, they are general and suffice for the analysis of the MSE performances of CCKF and ACKF.

Consider (6a) in its non-recursive form

$$\mathbf{x}_n = \mathbf{F}_{n:0}\mathbf{x}_0+\sum_{i=1}^{n}\mathbf{F}_{n:i}\mathbf{w}_i \tag{24}$$

where $\mathbf{x}_0$ is the initial state,[4] and the state transition matrix has the properties

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

This allows us to express the state covariance matrix as

$$\mathbf{R}_{\mathbf{x},n} = \mathbf{F}_{n:0}\mathbf{R}_{\mathbf{x},0}\mathbf{F}^H_{n:0}+\sum_{i=1}^{n}\mathbf{F}_{n:i}\mathbf{Q}_i\mathbf{F}^H_{n:i} \tag{25}$$

and the observation covariance as

$$\begin{aligned}\mathbf{R}_{\mathbf{y},n,m} &= E\{\mathbf{y}_n\mathbf{y}^H_m\}\\ &= \begin{cases}\mathbf{H}_n\mathbf{R}_{\mathbf{x},n}\mathbf{H}^H_n+\mathbf{R}_n, & \text{if } n=m\\ \mathbf{H}_n\mathbf{R}_{\mathbf{x},n}\mathbf{H}^H_m, & \text{if } n<m\\ \mathbf{H}_n\mathbf{R}_{\mathbf{x},m}\mathbf{H}^H_m, & \text{if } n>m.\end{cases}\end{aligned} \tag{26}$$

We have 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 expressed as

$$\begin{aligned}\mathbf{R}_{\mathbf{xy},n,m} &= E\left\{\mathbf{x}_n\mathbf{y}^H_m\right\}\qquad n\ge m\\ &= E\left\{\mathbf{x}_n(\mathbf{H}_m\mathbf{x}_m+\mathbf{v}_m)^H\right\}\\ &= \mathbf{R}_{\mathbf{x},m}\mathbf{H}^H_m\end{aligned} \tag{27}$$

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, respectively, given by

$$\mathbf{R}_{\mathbf{xz},n,n} = \left[\mathbf{R}_{\mathbf{xy},n,1}\ \mathbf{R}_{\mathbf{xy},n,2}\ \cdots\ \mathbf{R}_{\mathbf{xy},n,n}\right] \tag{28}$$

[4]Without loss of generality, we assume $E\{\mathbf{x}_0\}=\mathbf{0}$.

and

$$\mathbf{R}_{\mathbf{z},n} = \begin{bmatrix} \mathbf{R}_{\mathbf{y},1} & \mathbf{R}_{\mathbf{y},1,2} & \cdots & \mathbf{R}_{\mathbf{y},1,n} \\ \mathbf{R}_{\mathbf{y},2,1} & \mathbf{R}_{\mathbf{y},2} & \cdots & \mathbf{R}_{\mathbf{y},2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{R}_{\mathbf{y},n,1} & \mathbf{R}_{\mathbf{y},n,2} & \cdots & \mathbf{R}_{\mathbf{y},n} \end{bmatrix}. \tag{29}$$

Based on the expectations of (21) and (24), observe that the estimate $\widehat{\mathbf{x}}_{n|n}$ is unbiased, 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 KF do not change for noncircular state and observation signals. Equation (23) shows that the mean square characteristics of the CCKF are dependent on the covariance matrices of the state and observation noises but not on their pseudocovariances, leading to the following remark.

**Remark 2:** The noncircularity of the state and observation noises does not affect the performance of the linear conventional complex KF.

For the ACKF, the state estimate and the MSE matrix are given by expressions similar to (21) and (23), that is

$$\begin{aligned} \widehat{\mathbf{x}}^a_{n|n} &= E\{\mathbf{x}^a_0\} + \mathbf{W}^a_n\mathbf{z}^a_n = E\{\mathbf{x}^a_0\} + \mathbf{R}^a_{\mathbf{xz},n,n}(\mathbf{R}^a_{\mathbf{z},n})^{-1}\mathbf{z}^a_n \\ \mathbf{M}^a_{n|n} &= E\left\{(\mathbf{x}^a_n - \widehat{\mathbf{x}}^a_{n|n})(\mathbf{x}^a_n - \widehat{\mathbf{x}}^a_{n|n})^H\right\} \\ &= \mathbf{R}^a_{\mathbf{x},n} - \mathbf{R}^a_{\mathbf{xz},n,n}(\mathbf{R}^a_{\mathbf{z},n})^{-1}\mathbf{R}^{aH}_{\mathbf{xz},n,n} \end{aligned} \tag{30}$$

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

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

The terms $\mathbf{P}_{\mathbf{x},n}$ and $\mathbf{P}_{\mathbf{z},n}$ are the pseudocovariances 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. 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

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

and the widely linear (augmented) MSE for the ACKF can be written as

$$\begin{aligned} \mathbf{M}_{wl,n|n} &= \mathbf{R}_{\mathbf{x},n} - \mathbf{R}_{\mathbf{xz},n,n}\mathbf{C}\mathbf{R}^H_{\mathbf{xz},n,n} - \mathbf{R}_{\mathbf{xz},n,n}\mathbf{D}\mathbf{P}^H_{\mathbf{xz},n,n} \\ &\quad -\mathbf{P}_{\mathbf{xz},n,n}\mathbf{D}^*\mathbf{R}^H_{\mathbf{xz},n,n} - \mathbf{P}_{\mathbf{xz},n,n}\mathbf{C}^*\mathbf{P}^H_{\mathbf{xz},n,n}. \end{aligned} \tag{31}$$

After some tedious algebraic manipulations and following the approach in [16], the difference between the CCKF and the ACKF is found to be:

$$\begin{aligned} \Delta\mathbf{M}_n &= \mathbf{M}_{n|n} - \mathbf{M}_{wl,n|n} \\ &= (\mathbf{P}_{\mathbf{xz},n,n} - \mathbf{R}_{\mathbf{xz},n,n}\mathbf{R}^{-1}_{\mathbf{z},n}\mathbf{P}_{\mathbf{z},n}) \\ &\quad \times(\mathbf{R}^*_{\mathbf{z},n} - \mathbf{P}^*_{\mathbf{z},n}\mathbf{R}^{-1}_{\mathbf{z},n}\mathbf{P}_{\mathbf{z},n})^{-1} \\ &\quad \times(\mathbf{P}_{\mathbf{xz},n,n} - \mathbf{R}_{\mathbf{xz},n,n}\mathbf{R}^{-1}_{\mathbf{z},n}\mathbf{P}_{\mathbf{z},n})^H. \end{aligned} \tag{32}$$

**Remark 3:** Equation (32) is always positive semidefinite owing to the positive definiteness of the matrix $(\mathbf{R}^*_{\mathbf{z},n} - \mathbf{P}^*_{\mathbf{z},n}\mathbf{R}^{-1}_{\mathbf{z},n}\mathbf{P}_{\mathbf{z},n})$, and consequently $\Delta\mathbf{M}_n = \mathbf{0}$ only when $(\mathbf{P}_{\mathbf{xz},n,n} - \mathbf{R}^{-1}_{\mathbf{xz},n,n}\mathbf{R}_{\mathbf{z},n}\mathbf{P}_{\mathbf{z},n}) = \mathbf{0}$. Therefore, the widely linear ACKF always has the same or better MSE performance than the strictly linear CCKF.

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

### C. Duality Analysis of ACKF and Real-Valued KF

Due to the topological isomorphism between augmented complex vectors and bivariate real vectors, the ACKF has a dual bivariate real-valued KF, a property which can be used to significantly reduce the computational complexity of ACKF. For any complex vector $\mathbf{z} = \mathbf{z}_r + j\mathbf{z}_i \in \mathbb{C}^q$, the duality mapping is given by

$$\mathbf{z}^a = \begin{bmatrix} \mathbf{z} \\ \mathbf{z}^* \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{I} & j\mathbf{I} \\ \mathbf{I} & -j\mathbf{I} \end{bmatrix}}_{\equiv \mathbf{J}_\mathbf{z}} \underbrace{\begin{bmatrix} \mathbf{z}_\mathbf{r} \\ \mathbf{z}_\mathbf{i} \end{bmatrix}}_{=\mathbf{z}^r} \tag{33}$$

where $\mathbf{I}$ is the identity matrix and the invertible orthogonal mapping[5] $\mathbf{J}_\mathbf{z} : \mathbb{C} \to \mathbb{R}$ has the property $\mathbf{J}^{-1}_\mathbf{z} = (1/2)\mathbf{J}^H_\mathbf{z}$ [17]. Based on this isomorphism, the real bivariate state space corresponding to the augmented complex state space is given by

$$\begin{aligned} \mathbf{x}^r_n &= \mathbf{F}^r_{n-1}\mathbf{x}^r_{n-1} + \mathbf{w}^r_n \\ \mathbf{y}^r_n &= \mathbf{H}^r_n\mathbf{x}^r_n + \mathbf{v}^r_n \end{aligned} \tag{34}$$

where $\mathbf{x}^r_n = \mathbf{J}^{-1}_\mathbf{x}\mathbf{x}^a_n$, $\mathbf{y}^r_n = \mathbf{J}^{-1}_\mathbf{y}\mathbf{y}^a_n$, $\mathbf{F}^r_{n-1} = \mathbf{J}^{-1}_\mathbf{x}\mathbf{F}^a_{n-1}\mathbf{J}_\mathbf{x}$, $\mathbf{H}^r_n = \mathbf{J}^{-1}_\mathbf{y}\mathbf{H}^a_n\mathbf{J}_\mathbf{x}$, $\mathbf{w}^r_n = \mathbf{J}^{-1}_\mathbf{x}\mathbf{w}^a_n$, and $\mathbf{v}^r_n = \mathbf{J}^{-1}_\mathbf{y}\mathbf{v}^a_n$. The corresponding real-valued covariance matrices for the state and observation noises, $\mathbf{w}^r_n$ and $\mathbf{v}^r_n$, are given by

$$\begin{aligned} \mathbf{Q}^r_n &= E\{\mathbf{w}^r_n\mathbf{w}^{rH}_n\} = \mathbf{J}^{-1}_\mathbf{x}\mathbf{Q}^a_n\mathbf{J}^{-H}_\mathbf{x} \\ \mathbf{R}^r_n &= E\{\mathbf{v}^r_n\mathbf{v}^{rH}_n\} = \mathbf{J}^{-1}_\mathbf{y}\mathbf{R}^a_n\mathbf{J}^{-H}_\mathbf{y}. \end{aligned}$$

Following a similar analysis to that in [1], it can be shown that the ACKF and its dual real-valued KF have the same performance. Assuming that ACKF is initiated at time $(n-1)$, with initial state $\widehat{\mathbf{x}}^a_{n-1|n-1}$ and MSE matrix $\mathbf{M}^a_{n-1|n-1}$, the corresponding dual real-valued KF initialization is given by:

$$\begin{aligned} \widehat{\mathbf{x}}^r_{n-1|n-1} &= \mathbf{J}^{-1}_\mathbf{x}\widehat{\mathbf{x}}^a_{n-1|n-1} \\ \mathbf{M}^r_{n-1|n-1} &= \mathbf{J}^{-1}_\mathbf{x}\mathbf{M}^a_{n-1|n-1}\mathbf{J}^{-H}_\mathbf{x}. \end{aligned} \tag{35}$$

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

$$\begin{aligned} \widehat{\mathbf{x}}^r_{n|n-1} &= \mathbf{J}^{-1}_\mathbf{x}\widehat{\mathbf{x}}^a_{n|n-1} \\ \mathbf{M}^r_{n|n-1} &= \mathbf{J}^{-1}_\mathbf{x}\mathbf{M}^a_{n|n-1}\mathbf{J}^{-H}_\mathbf{x} \end{aligned} \tag{36}$$

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

[5] The matrix $\mathbf{J}_\mathbf{z}$ is of dimensions $2q \times 2q$.

expressions:

$$\begin{aligned}\mathbf{G}_n^a &= \mathbf{M}_{n|n-1}^a \mathbf{H}_n^{aH}[\mathbf{H}_n^a \mathbf{M}_{n|n-1}^a \mathbf{H}_n^{aH} + \mathbf{R}_n^a]^{-1} \\ &= \mathbf{J_x M}_{n|n-1}^r \mathbf{J_x}^H \mathbf{J_y}^{-H} \mathbf{H}_n^{rH} \mathbf{J_x}^H \\ &\quad \times [\mathbf{J_y H}_n^r \mathbf{J_x}^{-1} \mathbf{J_x M}_{n|n-1}^r \mathbf{J_x}^H \mathbf{J_x}^{-H} \mathbf{H}_n^{rH} \mathbf{J_y}^H + \mathbf{J_y R}_n^r \mathbf{J_y}^H]^{-1} \\ &= \mathbf{J_x M}_{n|n-1}^r \mathbf{H}_n^{rH} [\mathbf{H}_n^r \mathbf{M}_{n|n-1}^r \mathbf{H}_n^{rH} + \mathbf{R}_n^r]^{-1} \mathbf{J_y}^{-1} \\ &= \mathbf{J_x G}_n^r \mathbf{J_y}^{-1}. \end{aligned} \tag{37}$$

Consequently, for the state estimates $\widehat{\mathbf{x}}_{n|n}^a$ and $\widehat{\mathbf{x}}_{n|n}^r$ we have

$$\begin{aligned}\widehat{\mathbf{x}}_{n|n}^r &= \widehat{\mathbf{x}}_{n|n-1}^r + \mathbf{G}_n^r(\mathbf{y}_n^r - \mathbf{H}_n^r \widehat{\mathbf{x}}_{n|n-1}^r) \\ &= \mathbf{J_x}^{-1}\widehat{\mathbf{x}}_{n|n-1}^a + \mathbf{J_x}^{-1}\mathbf{G}_n^a \mathbf{J_y}(\mathbf{y}_n^r - \mathbf{H}_n^r \mathbf{J_x}^{-1}\widehat{\mathbf{x}}_{n|n-1}^a) \\ &= \mathbf{J_x}^{-1}\widehat{\mathbf{x}}_{n|n}^a \end{aligned} \tag{38}$$

and the MSE matrices are related as

$$\mathbf{M}_{n|n}^r = \mathbf{J_x}^{-1}\mathbf{M}_{n|n}^a \mathbf{J_x}^{-H}. \tag{39}$$

From (38), observe that the state estimates $\widehat{\mathbf{x}}_{n|n}^a$ and $\widehat{\mathbf{x}}_{n|n}^r$ are equivalent, and are related by an invertible linear mapping. To show that ACKF and its dual real-valued bivariate KF achieve the same MSE, recall that the MSE for the real-valued KF is given by

$$\epsilon_n^r = \mathrm{tr}\{\mathbf{M}_{n|n}^r\} \tag{40}$$

where the symbol $\mathrm{tr}\{\cdot\}$ denotes the matrix trace operator. Similarly, the MSE corresponding to the augmented MSE matrix $\mathbf{M}_{n|n}^a$ is given by the trace of (39), that is

$$\begin{aligned}\mathrm{tr}\{\mathbf{M}_{n|n}^a\} &= \mathrm{tr}\{\mathbf{J_x M}_{n|n}^r \mathbf{J_x}^H\} \\ &= \mathrm{tr}\{\mathbf{M}_{n|n}^r \mathbf{J_x}^H \mathbf{J_x}\} \\ &= 2 \cdot \mathrm{tr}\{\mathbf{M}_{n|n}^r\} \end{aligned} \tag{41}$$

where the expression $\mathbf{J_x}^H = 2\mathbf{J_x}^{-1}$ was utilized. At first, this result appears misleading as it suggests that ACKF achieves twice the error of its dual real-valued KF. However, this is because the error term is counted twice by the trace of $\mathbf{M}_{n|n}^a$, 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

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

**Remark 5:** The ACKF and the its dual bivariate real-valued KF are equivalent forms of the same state space models. They achieve identical MSE and state estimates at every time instant, regardless of the circularity of the processed signals.

By using the bivariate real KF, the computational complexity of ACKF is reduced, whereby the number of additions and multiplications required are approximately halved and quartered, respectively. Moreover, due to the duality between the bivariate real-valued KF and ACKF, the stability and convergence analysis for real-valued KFs also apply to the ACKF.

### D. Posterior Cramer–Rao Bound (PCRB)

For time invariant statistical models, the Cramer–Rao bound provides a theoretical performance bound for all unbiased estimators, by establishing the lowest attainable MSE. In time varying systems, such as the state space models where the state is driven by random noise, the PCRB provides a lower bound on the MSE performance of unbiased estimators [18].

For an unbiased estimator $\hat{\theta}[\mathbf{y}]$ of an $r$-dimensional random variable $\theta$, based on the observation $\mathbf{y}$, where the joint probability density of $\theta$ and $\mathbf{y}$ is given by $\mathcal{P}_{\theta,\mathbf{y}}[\mathbf{\Theta}, \mathbf{Y}]$, the PCRB on the estimation has the form [18]

$$\mathbf{\Gamma} = E\left\{(\theta - \hat{\theta}[\mathbf{y}])(\theta - \hat{\theta}[\mathbf{y}])^H\right\} \geq \mathbf{\Sigma}^{-1} \tag{42}$$

where $\mathbf{\Sigma}$ is the $r \times r$ dimensional Fisher information matrix with elements defined as

$$\Sigma_{lk} = -E\left\{\frac{\partial^2 \log \mathcal{P}_{\boldsymbol{\theta},\mathbf{y}}[\mathbf{\Theta}, \mathbf{Y}]}{\partial \mathbf{\Theta}_l \partial \mathbf{\Theta}_k}\right\} \qquad l, k = 1, \ldots, r. \tag{43}$$

The inequality in (42) implies that the difference $\Gamma - \Sigma^{-1}$ is positive semidefinite. Consider a general state space model of the form

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

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

where $\mathbf{f}$ and $\mathbf{h}$ are nonlinear, possibly time varying vector-valued functions, while $\mathbf{w}_n$ and $\mathbf{v}_n$ are independent white processes (not necessarily Gaussian). For this model, it was shown in [19] that the Fisher information matrix corresponding to the state $\mathbf{x}_{n+1}$ at time instant $(n+1)$ can be written in a computationally efficient recursive form, that is

$$\mathbf{\Sigma}_{n+1} = \mathbf{D}_n^{22} - \mathbf{D}_n^{21}(\mathbf{\Sigma}_n + \mathbf{D}_n^{11})^{-1}\mathbf{D}_n^{12} \tag{45}$$

where

$$\begin{aligned}\mathbf{D}_n^{11} &= E\{-\Delta_{\mathbf{x}_n}^{\mathbf{x}_n} \log \mathcal{P}[\mathbf{x}_{n+1}|\mathbf{x}_n]\} \\ \mathbf{D}_n^{12} &= E\{-\Delta_{\mathbf{x}_n}^{\mathbf{x}_{n+1}} \log \mathcal{P}[\mathbf{x}_{n+1}|\mathbf{x}_n]\} \\ \mathbf{D}_n^{21} &= E\{-\Delta_{\mathbf{x}_{n+1}}^{\mathbf{x}_n} \log \mathcal{P}[\mathbf{x}_{n+1}|\mathbf{x}_n]\} = (\mathbf{D}_n^{12})^T \\ \mathbf{D}_n^{22} &= E\{-\Delta_{\mathbf{x}_{n+1}}^{\mathbf{x}_{n+1}} \log \mathcal{P}[\mathbf{x}_{n+1}|\mathbf{x}_n]\} \\ &\quad + E\{-\Delta_{\mathbf{x}_{n+1}}^{\mathbf{x}_{n+1}} \log \mathcal{P}[\mathbf{y}_{n+1}|\mathbf{x}_{n+1}]\}\end{aligned}$$

and $\Delta_{\mathbf{b}}^{\mathbf{a}} = (\partial^2/\partial\mathbf{a}\partial\mathbf{b})$, while conditional probability densities $\mathcal{P}[\mathbf{x}_{n+1}|\mathbf{x}_n]$ and $\mathcal{P}[\mathbf{y}_{n+1}|\mathbf{x}_{n+1}]$ can be computed from (44).

Consider the application of the PCRB to the linear filtering problem characterized by the state space model in (8), where $\mathbf{w}_n$ and $\mathbf{v}_n$ are assumed zero-mean independent complex doubly white Gaussian noises with augmented covariance matrices $\mathbf{Q}_n^a$ and $\mathbf{R}_n^a$, respectively, [(14) and (15)], that is, $\mathbf{w}_n$ (and similarly $\mathbf{v}_n$) has a multivariate complex normal distribution defined as [20]

$$\begin{aligned}\mathcal{P}[\mathbf{w}_n] = &\frac{1}{\pi^r (\det \mathbf{Q}_n^a)^{1/2}} \\ &\times \exp\Big(-\frac{1}{2}(\mathbf{w}_n - E\{\mathbf{w}_n\})^H \mathbf{Q}_n^{-a}(\mathbf{w}_n - E\{\mathbf{w}_n\})\Big).\end{aligned}$$

It is straightforward to show that

$$\begin{aligned}\mathbf{D}_n^{11} &= \mathbf{F}_n^{aH}(\mathbf{Q}_{n+1}^a)^{-1}\mathbf{F}_n^a \\ \mathbf{D}_n^{12} &= -\mathbf{F}_n^{aH}(\mathbf{Q}_{n+1}^a)^{-1} \\ \mathbf{D}_n^{22} &= (\mathbf{Q}_{n+1}^a)^{-1} + \mathbf{H}_{n+1}^{aH}(\mathbf{R}_{n+1}^a)^{-1}\mathbf{H}_{n+1}^a.\end{aligned}$$

The distribution of the state estimate within the ACKF framework is Gaussian with a mean $\widehat{\mathbf{x}}_{n|n}^a$ and covariance $\mathbf{M}_{n|n}^a$, and it can be shown that the information matrix is the inverse of the state covariance matrix at every time instant, that is

$$\mathbf{M}^a_{n|n} = \boldsymbol{\Sigma}_n^{-1}. \tag{46}$$

**Remark 6:** The ACKF, like the bivariate real-valued KF, achieves the CRLB [19], [21], since it essentially estimates the state $\mathbf{x}^a_n$ as $\widehat{\mathbf{x}}^a_{n|n} = E[\mathbf{x}^a_n|\mathbf{y}^a_0, \mathbf{y}^a_1, \ldots, \mathbf{y}^a_n]$. However, CCKF only achieves the CRLB when the conditions in (16) are true for all time instances, which is generally not the case.

## III. ACEKF

The 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

$$\mathbf{x}_n = \mathbf{f}[\mathbf{x}_{n-1}] + \mathbf{w}_n \tag{47a}$$
$$\mathbf{y}_n = \mathbf{h}[\mathbf{x}_n] + \mathbf{v}_n \tag{47b}$$

where $\mathbf{f}[\cdot]$ and $\mathbf{h}[\cdot]$ are the nonlinear process and observation vector-valued functions, respectively, which may depend on the time $n$, and the remaining variables are as defined above. The EKF approximates these nonlinear functions by their first-order Taylor series expansions (TSE) about the state estimates. 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. For instance, a large class of functions, such as real functions of complex variables, do not satisfy the Cauchy–Riemann conditions thus severely restricting the set of allowable functions for nonlinear process and observations models.

The so-called $\mathbb{CR}$ calculus [22] exploits the isomorphism between the complex domain $\mathbb{C}$ and the real domain $\mathbb{R}^2$, and makes possible the TSE of both analytic and nonanalytic functions within the same framework. This way, the first-order TSE of a function $f[\mathbf{z}]$ is given by

$$f[\mathbf{z} + \Delta\mathbf{z}] = f[\mathbf{z}] + \frac{\partial f}{\partial \mathbf{z}}\Delta\mathbf{z} + \frac{\partial f}{\partial \mathbf{z}^*}\Delta\mathbf{z}^* \tag{48}$$

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

The first-order approximations of the state and observation equations, (47a) and (47b), about the estimates $\widehat{\mathbf{x}}_{n-1|n-1}$ and $\widehat{\mathbf{x}}_{n|n-1}$, are given by

$$\mathbf{x}_n = \mathbf{F}_{n-1}\mathbf{x}_{n-1} + \mathbf{A}_{n-1}\mathbf{x}^*_{n-1} + \mathbf{w}_n + \mathbf{r}_{n-1} \tag{49}$$
$$\mathbf{y}_n = \mathbf{H}_n\mathbf{x}_n + \mathbf{B}_n\mathbf{x}^*_n + \mathbf{v}_n + \mathbf{z}_n \tag{50}$$

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

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

From (49) and (50), observe that if $\mathbf{f}[\cdot]$ and $\mathbf{h}[\cdot]$ are nonanalytic, that is, $\mathbf{A} \neq \mathbf{0}$ and $\mathbf{B} \neq \mathbf{0}$, the linearized state and observation equations are widely linear [see (4)], and thus cannot be implemented using the standard 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 state and measurement noises. To this end, consider the nonlinear augmented state space model given by

$$\mathbf{x}^a_n = \mathbf{f}^a[\mathbf{x}^a_{n-1}] + \mathbf{w}^a_n \tag{51a}$$
$$\mathbf{y}^a_n = \mathbf{h}^a[\mathbf{x}^a_n] + \mathbf{v}^a_n \tag{51b}$$

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

$$\mathbf{x}^a_n = \mathbf{F}^a_{n-1}\mathbf{x}^a_{n-1} + \mathbf{w}^a_n + \mathbf{r}^a_{n-1} \tag{52a}$$
$$\mathbf{y}^a_n = \mathbf{H}^a_n\mathbf{x}^a_n + \mathbf{v}^a_n + \mathbf{z}^a_n \tag{52b}$$

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

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

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], however, utilizes the augmented state space model, and is summarized in Algorithm 2.

**Algorithm 2** The ACEKF algorithm

*Initialize with:*

$$\widehat{\mathbf{x}}^a_{0|0} = E\left\{\mathbf{x}^a_0\right\}$$
$$\mathbf{M}^a_{0|0} = E\left\{(\mathbf{x}^a_0 - E\{\mathbf{x}^a_0\})(\mathbf{x}^a_0 - E\{\mathbf{x}^a_0\})^H\right\}$$

*Prediction:*

$$\widehat{\mathbf{x}}^a_{n|n-1} = \mathbf{f}^a[\widehat{\mathbf{x}}^a_{n-1|n-1}] \tag{53}$$

*Prediction Covariance Matrix:*

$$\mathbf{M}^a_{n|n-1} = \mathbf{F}^a_{n-1}\mathbf{M}^a_{n-1|n-1}\mathbf{F}^{aH}_{n-1} + \mathbf{Q}^a_n \tag{54}$$

*Kalman Gain Matrix:*

$$\mathbf{G}^a_n = \mathbf{M}^a_{n|n-1}\mathbf{H}^{aH}_n\left(\mathbf{H}^a_n\mathbf{M}^a_{n|n-1}\mathbf{H}^{aH}_n + \mathbf{R}^a_n\right)^{-1} \tag{55}$$

*Correction:*

$$\widehat{\mathbf{x}}^a_{n|n} = \widehat{\mathbf{x}}^a_{n|n-1} + \mathbf{G}^a_n\left(\mathbf{y}^a_n - \mathbf{h}^a[\widehat{\mathbf{x}}^a_{n|n-1}]\right) \tag{56}$$

*Covariance Matrix:*

$$\mathbf{M}^a_{n|n} = (\mathbf{I} - \mathbf{G}^a_n\mathbf{H}^a_n)\mathbf{M}^a_{n|n-1} \tag{57}$$

The novelty of the ACEKF algorithm presented in this paper 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 [12]. Moreover, by utilizing the $\mathbb{CR}$ calculus, 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 in (33), the nonlinear real-valued state space corresponding to the augmented complex state space (51) is given by

$$\mathbf{x}_n^r = \mathbf{f}^r[\mathbf{x}_{n-1}^r] + \mathbf{w}_n^r \tag{58a}$$

$$\mathbf{y}_n^r = \mathbf{h}^r[\mathbf{x}_n^r] + \mathbf{v}_n^r \tag{58b}$$

where $\mathbf{f}^r[\mathbf{x}_n^r] = \mathbf{J}_\mathbf{x}^{-1}\mathbf{f}^a[\mathbf{x}_n^a]$, $\mathbf{h}^r[\mathbf{x}_n^r] = \mathbf{J}_\mathbf{y}^{-1}\mathbf{h}^a[\mathbf{x}_n^a]$ and the remaining variables are as defined above. The relationship between the augmented complex and real Jacobians $\mathbf{F}_n^a$ and $\mathbf{F}_n^r$ is established by comparing the first-order TSE of both sides of the relationship $\mathbf{f}^r[\mathbf{x}_n^r] = \mathbf{J}_\mathbf{x}^{-1}\mathbf{f}^a[\mathbf{x}_n^a]$, which leads to

$$\mathbf{F}_n^r = \mathbf{J}_\mathbf{x}^{-1}\mathbf{F}_n^a\mathbf{J}_\mathbf{x} = \frac{\partial \mathbf{f}^r}{\partial \mathbf{x}_n^r}\Big|_{\mathbf{x}_n^r=\widehat{\mathbf{x}}_{n|n}^r}$$

and similarly $\mathbf{H}_n^r = \mathbf{J}_\mathbf{y}^{-1}\mathbf{H}_n^a\mathbf{J}_\mathbf{x} = \frac{\partial \mathbf{h}^r}{\partial \mathbf{x}_n^r}\Big|_{\mathbf{x}_n^r=\widehat{\mathbf{x}}_{n|n-1}^r}$.

The duality between the ACEKF and the real-valued EKF can be established by showing that at every time instant $n$, the following relationships holds:

$$\begin{aligned}
\widehat{\mathbf{x}}_{n|n-1}^r &= \mathbf{J}_\mathbf{x}^{-1}\widehat{\mathbf{x}}_{n|n-1}^a\\
\mathbf{M}_{n|n-1}^r &= \mathbf{J}_\mathbf{x}^{-1}\mathbf{M}_{n|n-1}^a\mathbf{J}_\mathbf{x}^{-H}\\
\mathbf{G}_n^r &= \mathbf{J}_\mathbf{x}^{-1}\mathbf{G}_n^a\mathbf{J}_\mathbf{y}\\
\widehat{\mathbf{x}}_{n|n}^r &= \mathbf{J}_\mathbf{x}^{-1}\widehat{\mathbf{x}}_{n|n}^a\\
\mathbf{M}_{n|n}^r &= \mathbf{J}_\mathbf{x}^{-1}\mathbf{M}_{n|n}^a\mathbf{J}_\mathbf{x}^{-H}.
\end{aligned}$$

Therefore, the ACEKF and the bivariate real-valued EKF essentially implement the same state space model, but operate in different domains. However, if the state space is naturally defined in the complex domain, it is desirable to keep all of the computations in the original complex domain in order to facilitate understanding of the transformations the signal goes through, and to benefit from the notion of phase and circularity and more degrees-of-freedom in the estimation.

However, by using the real-valued bivariate EKF, the computational complexity of ACEKF is reduced, whereby the number of additions and multiplications required are approximately halved and quartered, respectively.

## IV. ACUKF

The unscented Kalman filter (UKF) [23] has been proposed to address the problems arising from the first-order approximation of nonlinearities within EKFs. It approximates the complex statistical posterior distribution rather than approximating the nonlinearity [24]. 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 recovered. This results in a filter that 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

$$\mathbf{y} = \mathbf{f}[\mathbf{x}] = \mathbf{f}[\bar{\mathbf{x}} + \delta\mathbf{x}] \qquad \mathbf{x} \in \mathbb{C}^{p\times 1},\ \mathbf{y} \in \mathbb{C}^{q\times 1} \tag{59}$$

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

$$\mathbf{y} = \mathbf{f}[\bar{\mathbf{x}}] + \nabla_{\delta\mathbf{x}}\mathbf{f} + \frac{1}{2!}\nabla_{\delta\mathbf{x}}^2\mathbf{f} + \frac{1}{3!}\nabla_{\delta\mathbf{x}}^3\mathbf{f} + \cdots \tag{60}$$

where the $i$th-order term in the TSE for $\mathbf{f}[\cdot]$ about $\bar{\mathbf{x}}$ is [24]

$$\frac{1}{i!}\nabla_{\delta\mathbf{x}}^i\mathbf{f} = \frac{1}{i!}\left(\sum_{k=1}^{p}\delta x_k\frac{\partial}{\partial x_k}\right)^i \mathbf{f}[\mathbf{x}]\big|_{\mathbf{x}=\bar{\mathbf{x}}} \tag{61}$$

with $\delta x_k$ being the $k$th component of $\delta\mathbf{x}$. The term above is an $i$th 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

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

where the $i$th term is given by

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

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

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

and is correct if the $i$th central moment of $\mathbf{x}$ is correct. Within the complex UT 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 $\mathbf{R_x}$. 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 TSE. For a second-order noncircular $\mathbf{y}$, the true output pseudocovariance $\mathbf{P_y} = E\{(\mathbf{y}-\bar{\mathbf{y}})(\mathbf{y}-\bar{\mathbf{y}})^T\}$ is given by

$$\mathbf{P_y} = \frac{\partial \mathbf{f}}{\partial \mathbf{x}} \mathbf{P_x} \Big(\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\Big)^T + E\left\{\frac{1}{3!}\nabla_{\delta\mathbf{x}}\mathbf{f}\big(\nabla^3_{\delta\mathbf{x}}\mathbf{f}\big)^T + \frac{1}{2!\times 2!}\nabla^2_{\delta\mathbf{x}}\mathbf{f}\big(\nabla^2_{\delta\mathbf{x}}\mathbf{f}\big)^T + \frac{1}{3!}\nabla^3_{\delta\mathbf{x}}\mathbf{f}\big(\nabla_{\delta\mathbf{x}}\mathbf{f}\big)^T\right\} - E\left\{\frac{1}{2!}\nabla^2_{\delta\mathbf{x}}\mathbf{f}\right\} E\left\{\frac{1}{2!}\nabla^2_{\delta\mathbf{x}}\mathbf{f}\right\}^T + \cdots$$

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

$$\mathcal{X}_0 = \bar{\mathbf{x}}, \qquad \mathcal{X}_i = \bar{\mathbf{x}} + \Big(\sqrt{(p+\lambda)\mathbf{R_x}}\Big)_i, \quad i = 1, \ldots, p$$
$$\mathcal{X}_i = \bar{\mathbf{x}} - \Big(\sqrt{(p+\lambda)\mathbf{R_x}}\Big)_i, \quad i = p+1, \ldots, 2p \qquad (62)$$

where $\Big(\sqrt{(p+\lambda)\mathbf{R_x}}\Big)_i$ is the $i$th column of the matrix square root[6] and $\lambda = \alpha^2(2p+\kappa) - 2p$ is a scaling parameter, while $\alpha$ determines the spread of the sigma points around the mean and is usually set to a small positive value (e.g., $10^{-3}$), $\kappa$ is a secondary scaling parameter which is usually set to 0, and $\beta$ is used to incorporate prior knowledge of the distribution (for Gaussian distributions, $\beta = 2$ is optimal). From (62) it is clear that the conventional sigma points do not incorporate the input pseudocovariance, and to overcome this issue, we consider the 'augmented' sigma points given by

$$\mathcal{X}_0^a = \bar{\mathbf{x}}^a, \ \mathcal{X}_i^a = \bar{\mathbf{x}}^a + \Big(\sqrt{(p+\lambda)\mathbf{R}_{\mathbf{x}}^a}\Big)_i, \quad i = 1, \ldots, 2p$$
$$\mathcal{X}_i^a = \bar{\mathbf{x}}^a - \Big(\sqrt{(p+\lambda)\mathbf{R}_{\mathbf{x}}^a}\Big)_i, \quad i = 2p+1, \ldots, 4p$$

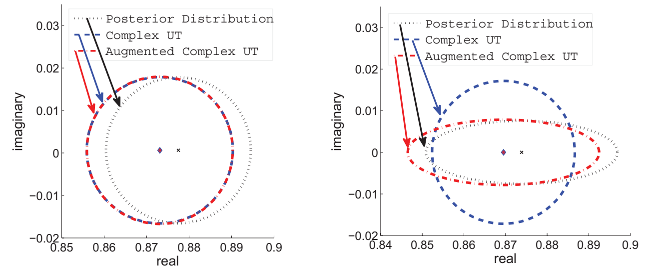
which are functions of the input mean, covariance and pseudocovariance, due to the use of the augmented covariance matrix, and can fully propagate the second-order statistics of improper inputs. The weights associated with the augmented sigma points are then given by

$$\mathcal{W}_0^{(m)} = \frac{\lambda}{2p+\lambda}, \qquad \mathcal{W}_0^{(c)} = \frac{\lambda}{2p+\lambda} + (1 - \alpha^2 + \beta)$$
$$\mathcal{W}_i^{(m)} = \mathcal{W}_i^{(c)} = \frac{\lambda}{2(2p+\lambda)}, \quad i = 1, \ldots, 4p \qquad (63)$$

where the output mean and covariance are computed using the $m$ and $c$ superscripted weights, respectively.

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. Fig. 1(a) 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 $\rho_x$ the pseudocovariance) the complex UT and the augmented complex UT had similar performance in capturing the distribution of the output $y_n$. Fig. 1(b) 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.

[6] If $\mathbf{L}$ is the matrix square root of $\mathbf{R_x} = \mathbf{L}\mathbf{L}^H$, then $\Big(\sqrt{(p+\lambda)\mathbf{R_x}}\Big)_i$ is the $i$th column of the matrix $\sqrt{(p+\lambda)}\mathbf{L}$.




(a) (b)

Fig. 1. Performance of the complex UT and augmented complex UT. (a) Circular input. (b) Noncircular input.

The ACUKF corresponding to the nonlinear state space model defined in (47) is summarized in Algorithm 3. Preliminary ACUKF algorithms have been introduced in [1] and [25]. The novelty of the ACUKF algorithm presented in this paper is that it does not assume a specific state or observation model, which makes it a more general form of the ACUKF presented in [1].

### A. Performance Analysis

In this section, we analyze the mean-square behavior of the the CEKF and CUKF [23] for analytic state and observation functions. Consider the complex-valued scalar state space given by

$$x_n = f[x_{n-1}] + w_n \qquad (64)$$
$$y_n = h[x_n] + v_n \qquad (65)$$

where $f[\cdot]$ and $h[\cdot]$ are holomorphic nonlinear state and observation models, respectively, $x_n$ and $y_n$ are the state and 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 $\rho_{v,n} = E\{v_n v_n\}$.

The unscented and EKF use the same general update expression, given by (72) and (56), to compute the estimate of the state, that is

$$\widehat{x}_{n|n} = \widehat{x}_{n|n-1} + g_n(y_n - \widehat{y}_{n|n-1}) \qquad (66)$$

where $g_n$ is the Kalman gain. This 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 (64) in to the observation equation (65) gives

$$y_n = h\big[f[x_{n-1}] + w_n\big] + v_n. \qquad (67)$$

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

$$h[f[x_{n-1}]+w_n]=h[f[x_{n-1}]]+\frac{\partial h}{\partial z}w_n+\frac{1}{2}\mathcal{H}_{zz}w_n^2+\text{h.o.t.} \tag{68}$$

with the Jacobian $(\partial h/\partial z)$ and Hessian $\mathcal{H}_{zz} = (\partial/\partial z)(\partial h/\partial z)$ evaluated at $z = f[x_{n-1}]$. Now subtract the true state, $x_n$, from the estimate given in (66) to find the state estimation error

$$\begin{aligned} e_n &= x_n - \widehat{x}_{n|n} \\ &= (f[x_{n-1}] + w_n) - \widehat{x}_{n|n-1} - g_n(y_n - \widehat{y}_{n|n-1}). \end{aligned} \tag{69}$$

**Algorithm 3** The ACUKF algorithm

*Initialize with:*

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

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

$$\begin{aligned} \mathcal{X}^a_{0,n-1} &= \widehat{\mathbf{x}}^a_{n-1|n-1} \\ \mathcal{X}^a_{i,n-1} &= \widehat{\mathbf{x}}^a_{n-1|n-1} \pm \left(\sqrt{(p+\lambda)\mathbf{M}^a_{n-1|n-1}}\right)_i \end{aligned} \tag{70}$$

*Compute predictions*:

$$\begin{aligned} \mathcal{X}^a_{i,n|n-1} &= \mathbf{f}^a[\mathcal{X}^a_{i,n-1}] \\ \widehat{\mathbf{x}}^a_{n|n-1} &= \sum_{i=0}^{4p} \mathcal{W}^{(m)}_i \mathcal{X}^a_{i,n|n-1} \\ \mathbf{M}^a_{n|n-1} &= \mathbf{Q}^a_n + \sum_{i=0}^{4p} \mathcal{W}^{(c)}_i \left(\mathcal{X}^a_{i,n|n-1} - \widehat{\mathbf{x}}^a_{n|n-1}\right)\left(\mathcal{X}^a_{i,n|n-1} - \widehat{\mathbf{x}}^a_{n|n-1}\right)^H \\ \mathcal{Y}^a_{i,n|n-1} &= \mathbf{h}^a[\mathcal{X}_{i,n|n-1}], \qquad i = 1, \ldots, 4p \\ \widehat{\mathbf{y}}^a_{n|n-1} &= \sum_{i=0}^{4p} \mathcal{W}^{(m)}_i \mathcal{Y}^a_{i,n|n-1} \end{aligned} \tag{71}$$

*Measurement update*:

$$\begin{aligned} \mathbf{R}^a_{\tilde{\mathbf{y}}^a,n|n-1} &= \mathbf{R}^a_n + \sum_{i=0}^{4p} \mathcal{W}^{(c)}_i \left(\mathcal{Y}^a_{i,n|n-1} - \widehat{\mathbf{y}}^a_{n|n-1}\right) \times \left(\mathcal{Y}^a_{i,n|n-1} - \widehat{\mathbf{y}}^a_{n|n-1}\right)^H \\ \mathbf{R}^a_{\mathbf{x}^a\mathbf{y}^a,n|n-1} &= \sum_{i=0}^{4p} \mathcal{W}^{(c)}_i \left(\mathcal{X}^a_{i,n|n-1} - \widehat{\mathbf{x}}^a_{n|n-1}\right) \times \left(\mathcal{Y}^a_{i,n|n-1} - \widehat{\mathbf{y}}^a_{n|n-1}\right)^H \\ \mathbf{G}^a_n &= \mathbf{R}^a_{\mathbf{x}^a\mathbf{y}^a,n|n-1}\left(\mathbf{R}^a_{\tilde{\mathbf{y}}^a,n|n-1}\right)^{-1} \\ \widehat{\mathbf{x}}^a_{n|n} &= \widehat{\mathbf{x}}^a_{n|n-1} + \mathbf{G}^a_n(\mathbf{y}^a_n - \widehat{\mathbf{y}}^a_{n|n-1}) \\ \mathbf{M}^a_{n|n} &= \mathbf{M}^a_{n|n-1} - \mathbf{G}^a_n \mathbf{R}^a_{\tilde{\mathbf{y}}^a,n|n-1} \mathbf{G}^{aH}_n \end{aligned} \tag{72}$$

Substituting (67) and (68) into (69) yields

$$\begin{aligned} e_n &= (f[x_{n-1}] + w_n) - \widehat{x}_{n|n-1} - g_n\Big(h[f[x_{n-1}]] \\ &\quad + \frac{\partial h}{\partial z}w_n + \frac{1}{2}\mathcal{H}_{zz}w_n^2 + \text{h.o.t.} + v_n - \widehat{y}_{n|n-1}\Big). \end{aligned} \tag{73}$$

Based on (73), 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 analyze terms related to the state and measurement noise pseudocovariances, these terms are

$$\begin{aligned} E\{e_n e_n^*\} &= -E\left\{\frac{1}{2}g_n\mathcal{H}_{zz}w_n^2\Big(f[x_{n-1}] - \widehat{x}_{n|n-1}\Big)^*\right\} \\ &\quad - E\left\{\frac{1}{2}\Big(f[x_{n-1}] - \widehat{x}_{n|n-1}\Big)g_n^*\mathcal{H}^*_{zz}(w_n^*)^2\right\} \\ &\quad + E\left\{\frac{1}{2}g_n\mathcal{H}_{zz}w_n^2\Big(g_n\big(h[f[x_{n-1}] - \widehat{y}_{n|n-1})\Big)^*\right\} \\ &\quad + E\left\{\frac{1}{2}\Big(g_n\big(h[f[x_{n-1}] - \widehat{y}_{n|n-1})\Big)g_n^*\mathcal{H}^*_{zz}(w_n^*)^2\right\} \\ &\quad + \text{(other terms \& h.o.t.)} \\ &= -\Re\left\{E\left\{g_n\mathcal{H}_{zz}\Big(f[x_{n-1}] - \widehat{x}_{n|n-1}\Big)^*\right\}\rho_{w,n}\right\} \\ &\quad + \Re\left\{E\left\{|g_n|^2\,\mathcal{H}_{zz}\big(h[f[x_{n-1}] - \widehat{y}_{n|n-1})^*\right\}\rho^*_{w,n}\right\} \\ &\quad + \text{(other terms \& h.o.t.)} \end{aligned} \tag{74}$$

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

**Remark 7:** From (74) observe that the MSEs for the CUKF and CEKF depend on the pseudocovariance of the state noise, a function of $\rho_{w,n}$ and $\rho^*_{w,n}$, and do not depend on the pseudocovariances of the observation noise. Hence, if the observation equation is nonlinear, their mean square behaviors are affected by the circularity of the state noise, regardless of whether the state equation is linear or nonlinear.

**Remark 8:** For linear state space models, the Hessian term $\mathcal{H}_{zz}$ in (74) vanishes, as the second derivatives of $h$ are zero, and so too do the four terms in the MSE expression (74) since they are all dependent on the pseudocovariance. Therefore, the mean square characteristics of the conventional linear complex Kalman filter (CCKF) does not depend on the circularity of the state or observation noises, but only on the total power (covariance).

## V. Application Examples

To illustrate the advantages of widely linear complex KFs over their conventional counterparts, we considered the following case studies: 1) filtering for a noisy complex-valued autoregressive process; 2) multistep ahead prediction for real-world noncircular and nonstationary wind data and the second-order noncircular Lorenz attractor; and 3) the nonlinear bearings only tracking.

### A. Complex Autoregressive Process

The performances of both the standard and widely linear KFs were examined using the first-order complex autoregressive process, AR(1), given by [1] and [26]

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

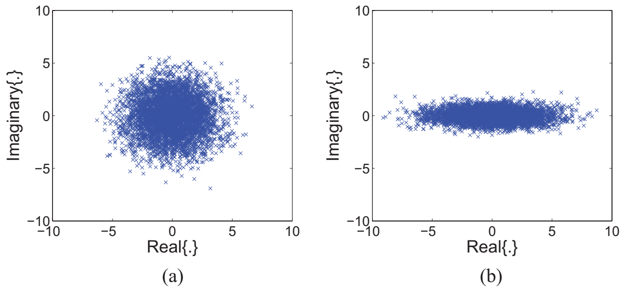



Fig. 2. Geometric view of circularity via a real–imaginary scatter plot of the AR(1) process driven by (a) circular ($K = 0$) and (b) noncircular ($K = 0.9$) white Gaussian distributions.

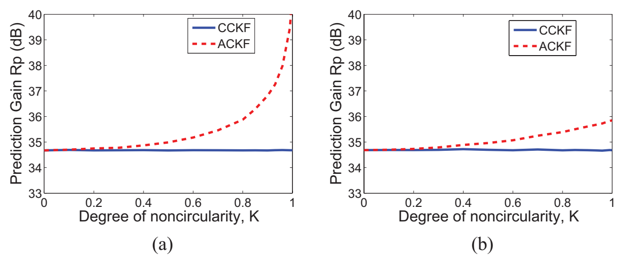



Fig. 3. Performance comparison between CCKF and ACKF for the AR(1) process with varying degrees of state and observation noise noncircularity. (a) Noncircular state noise. (b) Noncircular observation noise.

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

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

and $\delta$ is the discrete Dirac delta function. The estimates based on the linear filters, namely the CCKF and ACKF, were observed in the presence of additive complex noncircular doubly-white noise, $v_n$, that is (see [27] for the KF implementation of an autoregressive process)

$$y_n = x_n + v_n$$

while the observations corresponding to the nonlinear CEKF, CUKF and their corresponding augmented versions were performed as

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

The ratio of pseudocovariance to covariance, that is $K = (|\rho|/c)$, was used as a measure for the degree of noncircularity of the complex state and measurement noises [28], where a complex random variable is circular, if $K = 0$ and maximally noncircular if $K = 1$. Fig. 2 shows a real–imaginary scatter plot for two different realizations of the AR(1) process driven by Gaussian doubly white complex noise with different levels of circularity. For a quantitative assessment of the performance, the standard prediction gain $R_p = 10\log(\sigma_y^2/\sigma_e^2)$ was used, where $\sigma_y^2$ and $\sigma_e^2$ are the powers of the input signal and the output error.

Fig. 3 shows the performances of the standard CCKF and its corresponding widely linear (augmented) version, the ACKF. Fig. 3(a) illustrates the results for a circular observation noise and a state noise with various degrees of noncircularity, while

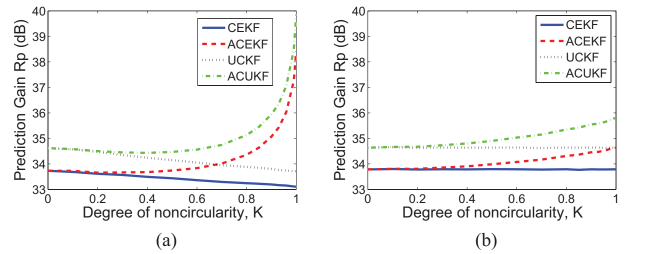



Fig. 4. Performance comparison between CEKF, CUKF and their corresponding widely linear (augmented) versions for the AR(1) process with varying degrees of state and observation noise noncircularity. (a) Noncircular state noise. (b) Noncircular observation noise.

Fig. 3(b) 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.

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

### B. Multistep Ahead Prediction

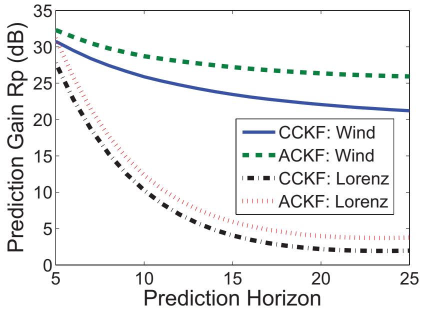
The performances of the CCKF and ACKF were next assessed for the multistep ahead prediction of the noncircular *Lorenz* signal and real-world noncircular and nonstationary *Wind* data. Simulations for the complex least mean square (CLMS) and its augmented version, the ACLMS, were also carried out to provide a performance comparison [29].

Fig. 5(a) summarizes the prediction performances for the *Lorenz* and the *Wind* data.[7] 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 suited to capturing the second-order statistics of 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. Fig. 5(b) 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 than that of the ACKF.

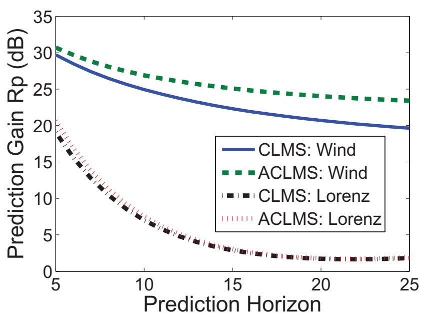
### C. Bearings Only Tracking (BOT)

BOT is a problem encountered in many practical applications, including submarine tracking by passive sonar or aircraft surveillance by a radar in passive mode. The objective is the

[7]The Wind signal ($x_n$), which has a magnitude (speed) ($v_n$) and direction ($\phi_n$), can be naturally represented as complex signal ($x_n = v_n e^{\phi_n}$).




(a)




(b)

Fig. 5. Multistep ahead prediction of real-world Wind data and the Lorenz attractor using CCKF, CLMS, and their corresponding widely linear versions. (a) CCKF and ACKF. (b) CLMS and ACLMS.

online estimation of the kinematics (position and velocity) of a moving target using observer line of sight noise-corrupted bearing (phase) measurements [30]. As the range measurements are not available and the bearings are not linearly related to the target state, the problem is inherently nonlinear. A single static sensor is not able to track targets using bearing measurements only (due to the lack of range measurements), and in order to estimate the range, the sensor has to maneuver. However, for two or more stationary sensors the observability problem is not an issue, as the multiple bearing measurements can be used to form a range estimate.

To estimate the trajectory of a target at time instant $n$, that is, its position $(x_n, y_n)$ and velocity $(\dot{x}_n, \dot{y}_n)$, for a system with $L$ observers located at $(x^o_{i,n}, y^o_{i,n})$, $i = 1, 2, \ldots, L$, the complex BOT state space is defined as

$$\mathbf{x}_n = \mathbf{F}\mathbf{x}_{n-1} + \mathbf{K}\omega_n, \qquad \mathbf{z}_n = \mathbf{h}[\mathbf{x}_n] + \mathbf{v}_n \tag{75}$$

with the variables defined as follows:

1) $\mathbf{x}_n = \left[x_n + jy_n \ \dot{x}_n + j\dot{y}_n\right]^T$ is the complex target state vector;
2) $\mathbf{F}$ and $\mathbf{K}$ are matrices defined as
$$\mathbf{F} = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad \mathbf{K} = \begin{bmatrix} \frac{T^2}{2} \\ T \end{bmatrix}$$
where $T$ is the sampling interval;
3) $\mathbf{z}_n$ is the observation vector and $\mathbf{h}[\mathbf{x}_n]$ is a vector function defined as
$$\mathbf{h}[\mathbf{x}_n] = \left[\beta_{1,n} \ \beta_{2,n} \ \cdots \ \beta_{L,n}\right]^T$$
where $\beta_i = \tan^{-1}(y_n - y^o_{i,n}/x_n - x^o_{i,n})$ is the target bearing at sensor $i$;

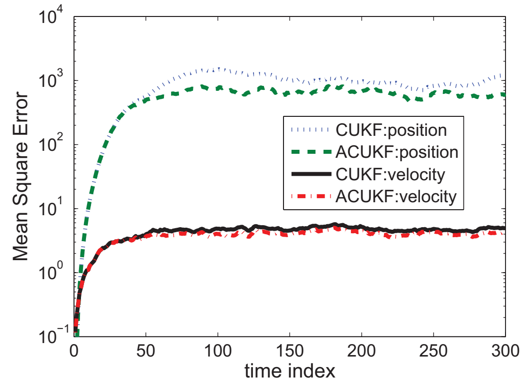



Fig. 6. Performances of CUKF and ACUKF with second-order noncircular state noise ($K = 0.9$).

4) $\omega_n = \ddot{x}_n + j\ddot{y}_n$ is the zero-mean state noise (used to account for the unknown target accelerations) with variance $r_{\omega,n}$ and pseudovariance $\rho_{\omega,n}$, while $\mathbf{v}_k = \left[\upsilon_{1,n} \ \upsilon_{2,n} \ \cdots \ \upsilon_{L,n}\right]^T$ is the zero mean real-valued observation noise with covariance $\mathbf{R}_{\upsilon,n}$.

The vector function $\mathbf{h}[\mathbf{x}_n]$ is real valued and it is straightforward to show that it does not satisfy the Cauchy–Riemann conditions, that is, $(\partial\mathbf{h}[\mathbf{x}_n]/\partial\mathbf{x}^*_n) \neq \mathbf{0}$, and is hence nonholomorphic.

To illustrate the benefits of ACUKF with over CUKF within the context of bearings only target motion analysis, consider a scenario with two static sensors located at $(-1200, 1300)$ and $(1000, 1500)$. The system described by (75) was simulated with a sampling interval of $T = 0.5$, and the MSE of the different algorithms was computed by averaging 100 independent simulations.

The performances of the CUKF and the ACUKF were compared using a second-order noncircular Gaussian state noise (with a degree of noncircularity of $K = 0.9$) with a distribution defined as

$$\omega_k \sim \mathcal{N}(0, 0.025), \qquad \upsilon_k \sim \mathcal{N}(0, 0.005).$$

The results, shown in Fig. 6, illustrate that the ACUKF had a lower MSE in estimating both the position and velocity of the target compared to CUKF.

## 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 utilizing both these statistical moments, we have readdressed the ACKF, and have examined its performance in relation to the CCKF. The analysis has shown that the ACKF offers significant performance gains over the CCKF for noncircular signals, and the same performance as the CCKF for circular signals. We have also introduced a more general form of the ACEKF, by using the so-called $\mathbb{CR}$ calculus, which enables the filter to operate with both analytic and nonanalytic state space models. Analysis of the mean square characteristics of CCKFs has shown that the mean square behavior of the CCKF does not account for the noncircularity of the state and observation noises, however, the mean square characteristics of the CEKF and CUKF were effected by state noise noncircularity, if the observation equation is nonlinear. The analysis was supported by simulations on both synthetic and real-world data.

## Acknowledgment

The authors would like to thank the anonymous reviewers for their helpful comments, and Gill Instruments, Lymington, U.K., for providing anemometers to record the real-world wind data.

## References

[1] D. P. Mandic and V. S. L. Goh, *Complex Valued Nonlinear Adaptive Filters: Noncircularity, Widely Linear and Neural Models*. New York: Wiley, 2009.
[2] Y. Xia, B. Jelfs, M. M. Van Hulle, J. C. Principe, and D. P. Mandic, "An augmented echo state network for nonlinear adaptive filtering of complex noncircular signals," *IEEE Trans. Neural Netw.*, vol. 22, no. 1, pp. 74–83, Jan. 2011.
[3] S. Javidi, C. C. Took, and D. P. Mandic, "Fast independent component analysis algorithm for quaternion valued signals," *IEEE Trans. Neural Netw.*, vol. 22, no. 12, pp. 1967–1978, Dec. 2011.
[4] B. C. Ujang, C. C. Took, and D. P. Mandic, "Quaternion-valued nonlinear adaptive filtering," *IEEE Trans. Neural Netw.*, vol. 22, no. 8, pp. 1193–1206, Aug. 2011.
[5] X. Wang and Y. Huang, "Convergence study in extended Kalman filter-based training of recurrent neural networks," *IEEE Trans. Neural Netw.*, vol. 22, no. 4, pp. 588–600, Apr. 2011.
[6] A. Y. Alanis, E. N. Sanchez, A. G. Loukianov, and M. A. Perez, "Real-time recurrent neural state estimation," *IEEE Trans. Neural Netw.*, vol. 22, no. 3, pp. 497–505, Mar. 2011.
[7] S. Saeedi, L. Paull, M. Trentini, and H. Li, "Neural network-based multiple robot simultaneous localization and mapping," *IEEE Trans. Neural Netw.*, vol. 22, no. 12, pp. 2376–2387, Dec. 2011.
[8] D. V. Prokhorov, "Training recurrent neurocontrollers for robustness with derivative-free Kalman filter," *IEEE Trans. Neural Netw.*, vol. 17, no. 6, pp. 1606–1616, Nov. 2006.
[9] R. Sharaf and A. Noureldin, "Sensor integration for satellite-based vehicular navigation using neural networks," *IEEE Trans. Neural Netw.*, vol. 18, no. 2, pp. 589–594, Mar. 2007.
[10] P. Dash, A. Pradhan, and G. Panda, "Frequency estimation of distorted power system signals using extended complex Kalman filter," *IEEE Trans. Power Delivery*, vol. 14, no. 3, pp. 761–766, Jul. 1999.
[11] G. V. Puskorius and L. A. Feldkamp, "Neurocontrol of nonlinear dynamical systems with Kalman filter trained recurrent networks," *IEEE Trans. Neural Netw.*, vol. 5, no. 2, pp. 279–297, Mar. 1994.
[12] V. S. L. Goh and D. P. Mandic, "An augmented extended Kalman filter algorithm for complex-valued recurrent neural networks," *Neural Comput.*, vol. 19, no. 4, pp. 1039–1055, 2007.
[13] F. Cattivelli and A. Sayed, "Distributed nonlinear Kalman filtering with applications to wireless localization," in *Proc. IEEE Int. Conf. Acoustics Speech Signal Process.*, Dallas, TX, Mar. 2010, pp. 3522–3525.
[14] B. Picinbono and P. Bondon, "Second-order statistics of complex signals," *IEEE Trans. Signal Process.*, vol. 45, no. 2, pp. 411–420, Feb. 1997.
[15] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ: Prentice Hall, 1993.
[16] B. Picinbono and P. Chevalier, "Widely linear estimation with complex data," *IEEE Trans. Signal Process.*, vol. 43, no. 8, pp. 2030–2033, Aug. 1995.
[17] D. H. Brandwood, "A complex gradient operator and its application in adaptive array theory," *IEEE Proc. Commun. Radar Signal Process.*, vol. 130, no. 1, pp. 11–16, Feb. 1983.
[18] H. L. van Trees, *Detection, Estimation and Modulation Theory*. New York: Wiley, 1968.
[19] P. Tichavsky, C. H. Muravchik, and A. Nehorai, "Posterior Cramer–Rao bounds for discrete-time nonlinear filtering," *IEEE Trans. Signal Process.*, vol. 46, no. 5, pp. 1386–1396, May 1998.
[20] A. Bos, "The multivariate complex normal distribution-a generalization," *IEEE Trans. Inform. Theory*, vol. 41, no. 2, pp. 537–539, Mar. 1995.
[21] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. Englewood Cliffs, NJ: Prentice-Hall, 1979.
[22] K. Kreutz-Delgado, "The complex gradient operator and the CR-calculus," Dept. Comput. Eng., Jacobs School Electr. Eng., Univ. California, San Diego, CA, Tech. Rep. ECE275CG-F05v1.3d, 2006.
[23] S. Julier and J. Uhlmann, "A new extension of the Kalman filter to nonlinear systems," in *Proc. Aerosen.: The 11th Int. Symp. Aerospace/Defense Sens. Simulat. Controls*, 1997, pp. 182–193.
[24] S. Julier, J. Uhlmann, and H. Durrant-Whyte, "A new method for the nonlinear transformation of means and covariances in filters and estimators," *IEEE Trans. Automat. Control*, vol. 45, no. 3, pp. 477–482, Mar. 2000.
[25] D. H. Dini, D. P. Mandic, and S. J. Julier, "The widely linear complex unscented Kalman filter," *IEEE Signal Process. Lett.*, vol. 18, no. 11, pp. 623–626, May 2011.
[26] J. Navarro-Moreno, "ARMA prediction of widely linear systems by using the innovations algorithm," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 3061–3068, Jul. 2008.
[27] M. H. Hayes, *Statistical Digital Signal Processing and Modeling*. New York: Wiley, 1996.
[28] E. Ollila, "On the circularity of a complex random variable," *IEEE Signal Process. Lett.*, vol. 15, pp. 841–844, Nov. 2008.
[29] S. Javidi, M. Pedzisz, V. S. L. Goh, and D. P. Mandic, "The augmented complex least mean square algorithm with application to adaptive prediction problems," in *Proc. 1st IARP Workshop Cogn. Inf. Process*, 2008, pp. 54–57.
[30] J.-P. Le Cadre and O. Tremois, "Bearings-only tracking for maneuvering sources," *IEEE Trans. Aerospace Electron. Syst.*, vol. 34, no. 1, pp. 179–193, Jan. 1998.

**Dahir H. Dini** (S'11) received the M.Eng. degree in Electrical and Electronic Engineering from Imperial College London, U.K., in 2009, and is currently pursuing a Ph.D. degree in adaptive signal processing at the same university.

His current research interests include multidimensional adaptive filtering, nonlinear signal processing and augmented complex statistics.

**Danilo P. Mandic** (M'99–SM'03) is a Professor in signal processing with Imperial College London, London, U.K., where he has been working in the area of nonlinear adaptive signal processing and nonlinear dynamics. He has been a Guest Professor with Katholieke Universiteit Leuven, Leuven, Belgium, the Tokyo University of Agriculture and Technology, Tokyo, Japan, and Westminster University, London, and a Frontier Researcher with RIKEN, Tokyo. His publication record includes two research monographs titled *Recurrent Neural Networks for Prediction* (West Sussex, U.K.: Wiley, August 2001) and *Complex-Valued Nonlinear Adaptive Filters: Noncircularity, Widely Linear and Neural Models* (West Sussex, U.K.: Wiley, April 2009), an edited book titled *Signal Processing for Information Fusion* (New York: Springer, 2008), and more than 200 publications on signal and image processing.

He has been a member of the IEEE Technical Committee on Machine Learning for Signal Processing, an Associate Editor for the IEEE Transactions on Circuits and Systems II, the IEEE Transactions on Signal Processing, the IEEE Transactions on Neural Networks, and the *International Journal of Mathematical Modeling and Algorithms*. He has produced award winning papers and products from his collaboration with the industry. He is a member of the London Mathematical Society.