

A Widely Linear Complex Unscented Kalman Filter

Dahir H. Dini, Danilo P. Mandic, and Simon J. Julier

Abstract—Conventional complex valued signal processing algorithms assume rotation invariant (circular) signal distributions, and are thus suboptimal for real world processes which exhibit rotation dependent distributions (noncircular). In nonlinear sequential state space estimation, noncircularity can arise from the data, state transition model, and state and observation noises. We provide further insight by revisiting the augmented complex unscented Kalman filter (ACUKF) and illuminating its operation in such scenarios. The analysis establishes a relationship between the estimation error and the degree of second order noncircularity (improperness) in the system for the conventional complex unscented Kalman filter (CUKF), and is supported by simulations on both synthetic and real world proper and improper signals.

Index Terms—Augmented complex UKF, complex circularity, improperness, unscented Kalman filter, widely linear Kalman filter, widely linear model.

I. INTRODUCTION

KALMAN filtering is a nonlinear sequential state estimation technique which has found a wide range of applications and extensions, both in real- and complex-valued scenarios [1]. Applications of complex valued Kalman filters include frequency estimation of time-varying signals [2] and training of neural networks [3]. Standard complex valued approaches have explicitly or implicitly assumed second order circularity (properness), that is, rotation invariant data distributions. When only second order statistics are considered, this amounts to equal powers in the real and imaginary parts of the data. However, real world signals have rotation dependent distributions and are second order noncircular (improper) [4].

Problems in nonlinear state space estimation, when dealing with noncircular signals, arise from the assumption that complex second order statistics (SOS) is a simple extension of real SOS (e.g., replace $E\{\mathbf{x}\mathbf{x}^T\}$ in \mathbb{R} with $E\{\mathbf{x}\mathbf{x}^H\}$ in \mathbb{C}). In addition, nonlinearities in state space models may not be differentiable in the standard Cauchy-Riemann sense, thus prohibiting the implementation of complex extended Kalman filters (CEKF)s. The issues with data improperness in this context are addressed by the widely linear model [1], [4], whereas the linearisation of nonholomorphic functions is performed

Manuscript received May 31, 2011; revised August 08, 2011; accepted August 08, 2011. Date of publication August 30, 2011; date of current version September 12, 2011. This work is part of the University Defence Research Centre (UDRC) at Imperial College London, supported by the MoD and DSTL. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Saeed Gazor.

D. H. Dini and D. P. Mandic are with Imperial College London, U.K. (e-mail: dahir.dini@ic.ac.uk; d.mandic@ic.ac.uk).

S. J. Julier is with University College London, U.K. (e-mail: S.Julier@cs.ucl.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/LSP.2011.2166259

using the $\mathbb{C}\mathbb{R}$ calculus [5]—leading to augmented (widely linear) complex Kalman filters¹ [1], [3].

Recent advances in the so called ‘augmented complex statistics’ [4] have highlighted that for a general (improper) complex vector \mathbf{x} , SOS based 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 second order statistics. To introduce an optimal second order estimator for the generality of complex signals, consider first the mean square estimator (MSE) 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

$$\hat{\mathbf{y}} = \mathbf{H}\mathbf{x} \quad (1)$$

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} . However, observe that both the real y_r and imaginary y_i parts of the vector \mathbf{y} are real valued, and

$$\hat{y}_r = E\{y_r|\mathbf{x}_r, \mathbf{x}_i\} \quad \hat{y}_i = E\{y_i|\mathbf{x}_r, \mathbf{x}_i\}. \quad (2)$$

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

$$\hat{y}_r = E\{y_r|\mathbf{x}, \mathbf{x}^*\} \quad \hat{y}_i = E\{y_i|\mathbf{x}, \mathbf{x}^*\} \quad (3)$$

and using (1) we obtain the *widely linear* complex estimator²

$$\hat{\mathbf{y}} = \mathbf{H}\mathbf{x} + \mathbf{G}\mathbf{x}^* = \mathbf{W}\mathbf{x}^a \quad (4)$$

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

$$\mathbf{R}_{\mathbf{x}}^a = E\{\mathbf{x}^a \mathbf{x}^{aH}\} = \begin{bmatrix} \mathbf{R}_{\mathbf{x}} & \mathbf{P}_{\mathbf{x}} \\ \mathbf{P}_{\mathbf{x}}^* & \mathbf{R}_{\mathbf{x}}^* \end{bmatrix}. \quad (5)$$

Recently introduced widely linear adaptive filters include the augmented complex least mean square (ACLMS) [1] and the augmented complex recursive least squares (ACRLS) [7]. The augmented complex Kalman filter (ACKF) and the augmented complex extended Kalman filter (ACEKF) were introduced in [3], while the augmented complex unscented Kalman filter (ACUKF) was introduced in [1, Ch. 15.4] within the context of neural network training using linear random walk state models. The advantages of ACUKF over CUKF when dealing with noncircular signals were illustrated, however, were not elaborated for a general case where sources of improperness include both the input data and system parameters.

To that end, we forgo the linear state model assumption, and introduce ACUKF for general state space models, analyze its performance for improper data, and illustrate its operation under general state and observation noise (proper and improper). Illustrative simulations support the analysis.

¹The Matlab toolbox for widely linear Kalman filters can be found at [6].

²The “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.”

II. THE AUGMENTED COMPLEX UKF

The unscented Kalman filter (UKF)[8] has been proposed to address the problems arising from the first order approximation of nonlinearities within EKFs, it approximates the statistical posterior distribution rather than approximating the nonlinearity within EKF [9]. To illustrate the conventional complex unscented transform (UT) and the augmented complex UT, consider the mapping

$$\mathbf{y} = \mathbf{f}[\mathbf{x}] = \mathbf{f}[\bar{\mathbf{x}} + \delta\mathbf{x}] \quad \mathbf{x} \in \mathbb{C}^{p \times 1}, \quad \mathbf{y} \in \mathbb{C}^{q \times 1} \quad (6)$$

where $\mathbf{f}[\cdot]$ is a nonlinear function (holomorphicity is assumed for clarity), $\mathbf{y} = [y_1, \dots, y_q]^T$ is the output, $\mathbf{x} = [x_1, \dots, x_p]^T$ 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 $\mathbf{P}_{\mathbf{x}} = E\{(\mathbf{x} - \bar{\mathbf{x}})(\mathbf{x} - \bar{\mathbf{x}})^T\} = \mathbf{0}$, and $\delta\mathbf{x} = \mathbf{x} - \bar{\mathbf{x}}$. The Taylor series expansion (TSE) of \mathbf{y} about $\bar{\mathbf{x}}$ is given by [5],

$$\mathbf{y} = \mathbf{f}[\bar{\mathbf{x}}] + \nabla_{\delta\mathbf{x}}\mathbf{f} + \frac{1}{2!}\nabla_{\delta\mathbf{x}}^2\mathbf{f} + \frac{1}{3!}\nabla_{\delta\mathbf{x}}^3\mathbf{f} + \dots \quad (7)$$

where the i th order term in the TSE for $\mathbf{f}[\cdot]$ about $\bar{\mathbf{x}}$ is [9]

$$\frac{1}{i!}\nabla_{\delta\mathbf{x}}^i\mathbf{f} = \frac{1}{i!}\left(\sum_{k=1}^p\delta x_k\frac{\partial}{\partial x_k}\right)^i\mathbf{f}[\mathbf{x}]_{|\mathbf{x}=\bar{\mathbf{x}}} \quad (8)$$

with δx_k being the k th component of $\delta\mathbf{x}$. This expression 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

$$\begin{aligned} \bar{\mathbf{y}} &= E\{\mathbf{f}[\bar{\mathbf{x}} + \delta\mathbf{x}]\} \\ &= \mathbf{f}[\bar{\mathbf{x}}] + E\left\{\frac{1}{2!}\nabla_{\delta\mathbf{x}}^2\mathbf{f} + \frac{1}{3!}\nabla_{\delta\mathbf{x}}^3\mathbf{f} + \dots\right\} \end{aligned} \quad (9)$$

where the i th term is given by

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

The symbols $m_{a_1, a_2, \dots, a_{i-1}, a_i} = E\{\delta x_{a_1}\delta x_{a_2}\dots\delta x_{a_{i-1}}\delta x_{a_i}\}$ denote the i th order central moments of the components \mathbf{x} with $a_k \in [1, 2, \dots, p]$. 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

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

and is correct if the i th 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

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

which are correct up to the second order TSEs. For a second order noncircular signal \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

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

The conventional complex unscented transform [2] 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

$$\begin{aligned} \mathcal{X}_0 &= \bar{\mathbf{x}}, \quad \mathcal{X}_i = \bar{\mathbf{x}} + \left(\sqrt{(p+\lambda)\mathbf{R}_{\mathbf{x}}}\right)_i, \quad i = 1, \dots, p \\ \mathcal{X}_i &= \bar{\mathbf{x}} - \left(\sqrt{(p+\lambda)\mathbf{R}_{\mathbf{x}}}\right)_i, \quad i = p+1, \dots, 2p \end{aligned} \quad (10)$$

where $(\sqrt{(p+\lambda)\mathbf{R}_{\mathbf{x}}})_i$ is the i th column of the matrix square root³ and λ is a scaling parameter. From (10) 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

$$\begin{aligned} \mathcal{X}_0^a &= \bar{\mathbf{x}}^a, \quad \mathcal{X}_i^a = \bar{\mathbf{x}}^a + \left(\sqrt{(p+\lambda)\mathbf{R}_{\mathbf{x}}^a}\right)_i, \quad i = 1, \dots, 2p \\ \mathcal{X}_i^a &= \bar{\mathbf{x}}^a - \left(\sqrt{(p+\lambda)\mathbf{R}_{\mathbf{x}}^a}\right)_i, \quad i = 2p+1, \dots, 4p \end{aligned}$$

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 by⁴

$$\begin{aligned} \mathcal{W}_0^{(m)} &= \frac{\lambda}{2p+\lambda}, \quad \mathcal{W}_0^{(c)} = \frac{\lambda}{2p+\lambda} + (1 - \alpha^2 + \beta) \\ \mathcal{W}_i^{(m)} &= \mathcal{W}_i^{(c)} = \frac{\lambda}{2(2p+\lambda)}, \quad i = 1, \dots, 4p. \end{aligned} \quad (11)$$

To illustrate the benefits of the augmented complex UT over the standard UT, consider the system defined by $y_n = \tanh[x_n]$, where the input x_n is a Gaussian doubly white 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 ρ_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

³If \mathbf{L} is the matrix square root of $\mathbf{R}_{\mathbf{x}} = \mathbf{L}\mathbf{L}^H$, then $(\sqrt{(p+\lambda)\mathbf{R}_{\mathbf{x}}})_i$ is the i th column of the matrix $\sqrt{(p+\lambda)\mathbf{L}}$.

⁴The variables α and β are as defined in [9], and the output mean and covariance are computed using the m and c superscripted weights, respectively.

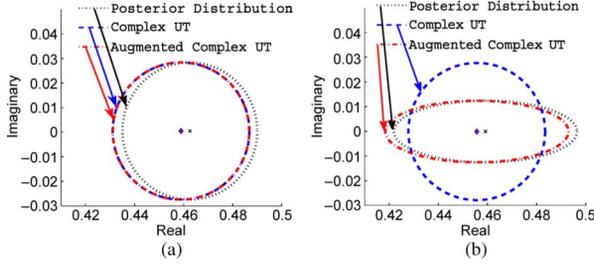


Fig. 1. Performance of the complex UT and augmented complex UT. (a) Circular input; (b) noncircular input.

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. Next, consider the state space

$$\mathbf{x}_n = \mathbf{f}[\mathbf{x}_{n-1}, \mathbf{w}_{n-1}], \quad \mathbf{y}_n = \mathbf{h}[\mathbf{x}_n, \mathbf{v}_n] \quad (12)$$

where $\mathbf{x}_n, \mathbf{y}_n, \mathbf{w}_n$ and \mathbf{v}_n are the state with covariance matrix \mathbf{M}_n , observed signal, state noise with covariance \mathbf{Q}_n and observation noise with covariance \mathbf{R}_n respectively. In the unscented Kalman filter framework, the vector $\underline{\mathbf{x}}_n = [\mathbf{x}_n^T, \mathbf{w}_n^T, \mathbf{v}_n^T]^T$ with covariance matrix $\underline{\mathbf{M}}_n = \text{diag}(\mathbf{M}_n, \mathbf{Q}_n, \mathbf{R}_n)$ is normally used as the extended state vector [8], while for ACUKF, summarised in Algorithm 1, we use the augmented vector $\underline{\mathbf{M}}_n^a = [\mathbf{x}_n^{aT}, \mathbf{w}_n^{aT}, \mathbf{v}_n^{aT}]^T$ with covariance $\underline{\mathbf{M}}_n^a = \text{diag}(\mathbf{M}_n^a, \mathbf{Q}_n^a, \mathbf{R}_n^a)$. The novelty of ACUKF algorithm presented in this work is that it does not assume a specific state or observation model which makes it more general than ACUKF presented in [1].

III. PERFORMANCE ANALYSIS OF CUKFs

Consider the complex valued scalar state space given by

$$x_n = f[x_{n-1}] + w_n \quad (13)$$

$$y_n = h[x_n] + v_n \quad (14)$$

where $f[\cdot]$ and $h[\cdot]$ are holomorphic nonlinear state 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 and observation noises respectively. The state 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 Kalman filter uses the update expression (15) to compute the estimate of the state,

Algorithm 1 The augmented complex unscented Kalman filter

Initialize with

$$\hat{\mathbf{x}}_{0|0}^a = E\{\mathbf{x}_0^a\}, \quad \mathbf{M}_{0|0}^a = E\left\{(\mathbf{x}_0^a - E\{\mathbf{x}_0^a\})(\mathbf{x}_0^a - E\{\mathbf{x}_0^a\})^H\right\}$$

$$\hat{\underline{\mathbf{x}}}_{0|0}^a = [\hat{\mathbf{x}}_{0|0}^a, \mathbf{0}, \mathbf{0}]^T, \quad \underline{\mathbf{M}}_{0|0}^a = \text{diag}(\mathbf{M}_{0|0}^a, \mathbf{Q}_0^a, \mathbf{R}_0^a).$$

Calculate sigma points

$$\underline{\mathbf{x}}_{n-1}^a = \left[\hat{\underline{\mathbf{x}}}_{n-1|n-1}^a, \hat{\underline{\mathbf{x}}}_{n-1|n-1}^a \pm \left(\sqrt{(p+\lambda)\underline{\mathbf{M}}_{n-1|n-1}^a} \right)_i \right]$$

where $\underline{\mathbf{x}}_{n-1}^a = [\mathcal{X}_{x,n-1}^{aT}, \mathcal{X}_{w,n-1}^{aT}, \mathcal{X}_{v,n-1}^{aT}]^T$.

Compute predictions

$$\mathbf{x}_{x,n|n-1}^a = \mathbf{f}[\mathbf{x}_{x,n-1}^a, \mathbf{x}_{w,n-1}^a],$$

$$\hat{\mathbf{x}}_n^a|_{n-1} = \sum_{i=0}^{4p} \mathcal{W}_i^{(m)} \mathcal{X}_{i,x,n|n-1}^a$$

$$\mathbf{M}_n^a|_{n-1} = \sum_{i=0}^{4p} \mathcal{W}_i^{(c)} (\mathcal{X}_{i,x,n|n-1}^a - \hat{\mathbf{x}}_n^a|_{n-1}) \times (\mathcal{X}_{i,x,n|n-1}^a - \hat{\mathbf{x}}_n^a|_{n-1})^H$$

$$\mathbf{y}_n^a|_{n-1} = \mathbf{f}[\mathbf{x}_{x,n|n-1}^a, \mathbf{x}_{v,n-1}^a],$$

$$\hat{\mathbf{y}}_n^a|_{n-1} = \sum_{i=0}^{4p} \mathcal{W}_i^{(m)} \mathcal{Y}_{i,n|n-1}^a.$$

Measurement update

$$\mathbf{R}_{\hat{\mathbf{y}}^a, n|n-1}^a = \sum_{i=0}^{4p} \mathcal{W}_i^{(c)} (\mathcal{Y}_{i,n|n-1}^a - \hat{\mathbf{y}}_n^a|_{n-1}) \times (\mathcal{Y}_{i,n|n-1}^a - \hat{\mathbf{y}}_n^a|_{n-1})^H$$

$$\mathbf{R}_{\mathbf{x}^a \mathbf{y}^a, n|n-1}^a = \sum_{i=0}^{4p} \mathcal{W}_i^{(c)} (\mathcal{X}_{i,x,n|n-1}^a - \hat{\mathbf{x}}_n^a|_{n-1}) \times (\mathcal{Y}_{i,n|n-1}^a - \hat{\mathbf{y}}_n^a|_{n-1})^H$$

$$\mathbf{G}_n^a = \mathbf{R}_{\mathbf{x}^a \mathbf{y}^a, n|n-1}^a (\mathbf{R}_{\hat{\mathbf{y}}^a, n|n-1}^a)^{-1}$$

$$\mathbf{M}_n^a|_n = \mathbf{M}_n^a|_{n-1} - \mathbf{G}_n^a \mathbf{R}_{\hat{\mathbf{y}}^a, n|n-1}^a \mathbf{G}_n^{aH}$$

$$\hat{\mathbf{x}}_n^a|_n = \hat{\mathbf{x}}_n^a|_{n-1} + \mathbf{G}_n^a (\mathbf{y}_n^a - \hat{\mathbf{y}}_n^a|_{n-1}) \quad (15)$$

that is

$$\hat{x}_n = \hat{x}_{n|n-1} + g_n (y_n - \hat{y}_{n|n-1}) \quad (16)$$

where g_n is the Kalman gain. The state estimate comprises of a prediction term, $\hat{x}_{n|n-1}$, and a weighted innovation term, $(y_n - \hat{y}_{n|n-1})$.

Substituting the state (13) in to the observation (14) gives

$$y_n = h[f[x_{n-1}] + w_n] + v_n \quad (17)$$

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

$$h[f[x_{n-1}] + w_n] = h[f[x_{n-1}]] + \frac{\partial h}{\partial z} w_n + \frac{1}{2} \mathcal{H}_{zz} w_n^2 + \text{h.o.t.} \quad (18)$$

where h.o.t. abbreviates higher order terms, and the Jacobian $\partial h/\partial z$ and Hessian $\mathcal{H}_{zz} = (\partial/\partial z)(\partial h/\partial z)$ are evaluated at $f[x_{n-1}]$. The estimation error, $e_n = x_n - \hat{x}_{n|n}$, is given by

$$e_n = (f[x_{n-1}] + w_n) - \hat{x}_{n|n-1} - g_n (y_n - \hat{y}_{n|n-1}). \quad (19)$$

Substituting (17) and (18) into (19) yields

$$e_n = (f[x_{n-1}] + w_n) - \hat{x}_{n|n-1} - g_n \left(h[f[x_{n-1}]] + \frac{\partial h}{\partial z} w_n + \frac{1}{2} \mathcal{H}_{zz} w_n^2 + \text{h.o.t.} + v_n - \hat{y}_{n|n-1} \right). \quad (20)$$

Observe from (20) that the mean square error (MSE) given by $E\{e_n e_n^*\}$, consists of a large number of terms, however, since we are only interested in the effect of circularity on the MSE, we restrict ourselves to analyzing the terms related to the state and observation noise pseudocovariances, obtained from

$$E\{e_n e_n^*\} = -\Re\left\{E\left\{g_n \mathcal{H}_{zz} (f[x_{n-1}] - \hat{x}_{n|n-1})^*\right\} \rho_{w,n}\right\} + \Re\left\{E\left\{|g_n|^2 \mathcal{H}_{zz} (h[f[x_{n-1}] - \hat{y}_{n|n-1}])^*\right\} \rho_{w,n}^*\right\} + (\text{other terms \& h.o.t.}) \quad (21)$$

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

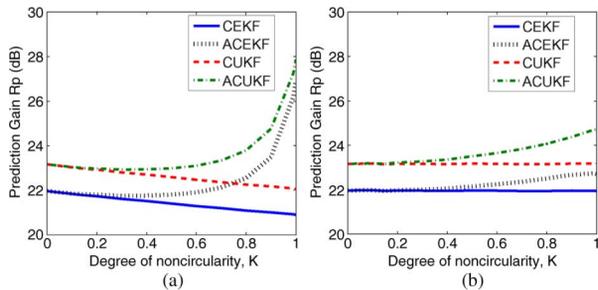


Fig. 2. Performance comparison between CEKF, CUKF and their corresponding augmented versions for the AR(4) process. (a) Noncircular state noise; (b) noncircular observation noise.

From (21) it can be seen that the MSE for the CUKF, for the state space defined by (13) and (14), is dependent on the pseudocovariance of the state noise, namely it is a function of $\rho_{w,n}$ and $\rho_{w,n}^*$, but does not depend on the pseudocovariance of the measurement noise. Hence its mean-square behavior is affected by the circularity of the state noise, if the observation equation is nonlinear, regardless of whether the state equation is linear or nonlinear.

For a linear state space model the Hessian term \mathcal{H}_{zz} in (21) vanishes, and so too do the four terms in the MSE expression (21) that depend on the pseudocovariance. Therefore, the mean-square characteristics of the conventional linear complex Kalman filter do not depend on the circularity of the state or observation noises, rendering it suboptimal.

IV. APPLICATION EXAMPLES

A. Complex Autoregressive Process

The performances of both the standard and widely linear complex unscented Kalman filters were examined using the 4th order complex autoregressive process given by [1]

$$\text{AR}(4): x_n = 1.79x_{n-1} - 1.85x_{n-2} + 1.27x_{n-3} - 0.41x_{n-4} + u_n$$

and driven by doubly white Gaussian, zero-mean noise with variance $E\{u_{n-i}u_{n-l}^*\} = c_u\delta_{i-l}$ and pseudocovariance $E\{u_{n-i}u_{n-l}\} = \rho_u\delta_{i-l}$. The observation equation was a nonlinear function of the state, $y_n = \tanh[x_n] + v_n$, and in the presence of additive noncircular complex white noise, v_n . The ratio of the magnitude of pseudocovariance to covariance, $K = |\rho|/c$, was used as a measure for the circularity of the noises [10], where a complex random variable is circular if $K = 0$ and maximally noncircular if $K = 1$. 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.

Fig. 2 shows the performances of the standard, strictly linear, CEKF and CUKF and their corresponding widely linear versions, ACEKF and ACUKF. Fig. 2(a) illustrates the results for a circular observation noise and a state noise with various degrees of noncircularity, while Fig. 2(b) shows the results for a noncircular observation noise with a circular state noise. For both sets of simulations, ACEKF and ACUKF had superior performances compared to CEKF and CUKF, as the degree of noise noncircularity (K) increased. The CEKF and CUKF had decreasing performance as the degree of noncircularity of the state noise

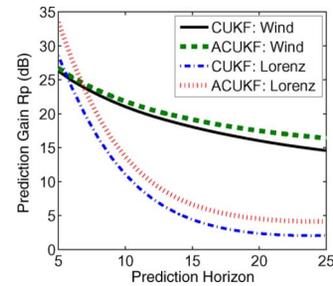


Fig. 3. Prediction performance of CUKF and ACUKF.

increased, while their performance was unaffected by observation noise noncircularity.

B. Multistep Ahead Prediction

The performances of the filters were next assessed for multistep ahead prediction of the noncircular *Lorenz* attractor and real world noncircular and nonstationary *Wind* data (the speed and direction are modeled as a complex number). Fig. 3 summarises the performances; in all the cases, the ACUKF was able to capture the underlying dynamics of the signals better than CUKF. This can be attributed to the use of the widely linear model, which caters for the full second order statistics of the signals.

V. CONCLUSION

We have revisited the widely linear ACUKF and have performed a rigorous analysis of its performance relative to the conventional, strictly linear, CUKF, showing that ACUKF exhibits superior performance for noncircular signals. The mean square error of CUKF has been shown to be affected by state noise noncircularity only if the observation equation is nonlinear and unaffected by observation noise noncircularity. Simulation results support the analysis.

REFERENCES

- [1] D. Mandic and V. S. L. Goh, *Complex Valued Nonlinear Adaptive Filters: Noncircularity, Widely Linear and Neural Models*. Hoboken, NJ: Wiley, 2009.
- [2] P. Dash, S. Hasan, and B. Panigrahi, "Adaptive complex unscented Kalman filter for frequency estimation of time-varying signals," *IET Sci. Meas. Technol.*, vol. 4, no. 2, pp. 93–103, 2010.
- [3] S. Goh and D. P. Mandic, "An augmented extended kalman filter algorithm for complex-valued recurrent neural networks," *Neural Comput.*, vol. 19, pp. 1039–1055, 2007.
- [4] B. Picinbono and P. Bondon, "Second-order statistics of complex signals," *IEEE Trans. Signal Process.*, vol. 45, no. 2, pp. 411–420, Feb. 1997.
- [5] K. Kreutz-Delgado, Electrical and Computer Engineering, Jacobs Sch. Eng., Univ. California-San Diego, The Complex Gradient Operator and the CR-Calculus Tech. Rep. ECE275CG-F05v1.3d, 2006.
- [6] [Online]. Available: <http://www.commsp.ee.ic.ac.uk/~mandic/complexbook/>
- [7] S. Douglas, "Widely-linear recursive least-squares algorithm for adaptive beamforming," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, 2009, 2009, pp. 2041–2044 [Online]. Available: <http://www.commsp.ee.ic.ac.uk/~mandic/complexbook/>
- [8] S. Julier and J. Uhlmann, "A new extension of the Kalman filter to nonlinear systems," in *Proc. AeroSense 11th Int. Symp. Aerospace/Defense Sensing, Simulation and Controls*, 1997, pp. 182–193.
- [9] 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. Contr.*, vol. 45, no. 3, pp. 477–482, 2000.
- [10] E. Ollila, "On the circularity of a complex random variable," *IEEE Signal Process. Lett.*, vol. 15, pp. 841–844, 2008.