State estimation of nonlinear systems using the Unscented Kalman Filter

João Almeida

***Abstract*— This paper addresses the problem of estimating the state of a nonlinear system from measurements that are perturbed by a random source of noise. The Extended Kalman Filter is a type of all-purpose filter that tries to solve this problem by dealing with a linearized version of the system. A new methodology proposed by Julier et al. in [1] named Unscented Kalman Filter is presented. It uses the so-called unscented transformation to better describe the stochastic evolution of the system's state. An illustrative example is given where the performances of the two filters are compared and discussed.**

I. Introduction

A fundamental problem in control is that of estimating the state of a general nonlinear system. The knowledge of the system's state is important by itself but it may also be of great interest to a controller that seeks to stabilize the system.

For linear systems that are observable and perturbed by additive zero-mean Gaussian white-noise, the problem has an optimal solution: the *Kalman Filter* (KF) [2]. The KF is optimal in a minimum mean squared sense and has a recursive form making it easy to implement. A straightforward extension for the nonlinear case is the *Extended Kalman Filter* (EKF). The EKF propagates the statistics that describe the state of the system (mean and covariance) by considering a linearized version of the original nonlinear system so that the standard KF equations can be applied. One consequence of this approach is that guarantees of both stability and optimality of the filter are lost.

In [1], a new approach is proposed to generalize the KF to nonlinear systems which consist of using the so-called *unscented transformation* to propagate the mean and covariance of the system's state in a different manner. The authors claim that this new filter, named *Unscented Kalman Filter* (UKF), has an expected performance superior to that of the EKF. Since the original paper was published, further developments and improvements have been made. These are summarized in [3] and [4].

More formally, consider a discrete-time nonlinear system or plant, described by

$$\begin{aligned} \boldsymbol{x}_{k+1} &= \boldsymbol{f}(\boldsymbol{x}_k, \boldsymbol{u}_k, \boldsymbol{\xi}_k) \\ \boldsymbol{z}_k &= \boldsymbol{g}(\boldsymbol{x}_k, \boldsymbol{\theta}_k) \end{aligned}$$

where $\boldsymbol{x}_k$ is the system's state at time index k, $\boldsymbol{u}_k$ is the input applied to the system, $\boldsymbol{\xi}_k$ is the process noise associated to disturbances, $\boldsymbol{z}_k$ is the measured output of the system, and $\boldsymbol{\theta}_k$ represents measurement or observation noise. The initial state of the system is assumed to be described by

$$\boldsymbol{x}_0 \sim \mathcal{N}(\bar{\boldsymbol{x}}_0, \boldsymbol{\Sigma}_0)$$

where $\bar{\boldsymbol{x}}_0 = \mathsf{E}\{\boldsymbol{x}_0\}$ and $\boldsymbol{\Sigma}_0 = \mathsf{E}\{(\boldsymbol{x}_0 - \bar{\boldsymbol{x}}_0)(\boldsymbol{x}_0 - \bar{\boldsymbol{x}}_0)^\top\}$.

The optimal solution of the estimation problem in the minimum mean squared error sense can not be described by a finite number of parameters. However, under some simplifying assumptions (namely, that the distribution of $\boldsymbol{x}_k$ is Gaussian at time k), only the mean and covariance need to be propagated and this propagation is given by the recursive equations

$$\begin{aligned} \hat{\boldsymbol{x}}_{k+1|k+1} &= (\text{prediction of } \boldsymbol{x}_{k+1}) \\ &\quad + \mathbf{H}_{k+1}[\boldsymbol{z}_{k+1} - (\text{prediction of } \boldsymbol{z}_{k+1})] \\ \mathbf{P}_{k+1|k+1} &= \mathbf{P}_{k+1|k} - \mathbf{H}_{k+1}\mathbf{P}_{\tilde{\boldsymbol{z}}_{k+1}}\mathbf{H}_{k+1}^\top \end{aligned}$$

where $\hat{\boldsymbol{x}}_{k|k}$ and $\mathbf{P}_{k|k}$ are the mean and covariance of $\boldsymbol{x}_k$, $\mathbf{P}_{k+1|k}$ is the prediction of the covariance of $\boldsymbol{x}_{k+1}$, $\mathbf{P}_{\tilde{\boldsymbol{z}}_k}$ is the covariance of $\tilde{\boldsymbol{z}}_k = \boldsymbol{z}_k - \hat{\boldsymbol{z}}_k$ with $\hat{\boldsymbol{z}}_k$ denoting the predicted value of $\boldsymbol{z}_k$, and $\mathbf{H}_k$ is the filter gain. The optimal terms in the recursion are given by

$$\begin{aligned} \hat{\boldsymbol{x}}_{k+1|k} &= \mathsf{E}\{\boldsymbol{f}(\hat{\boldsymbol{x}}_{k|k}, \boldsymbol{u}_k, \boldsymbol{\xi}_k)\} \\ \hat{\boldsymbol{z}}_{k+1} &= \mathsf{E}\{\boldsymbol{h}(\hat{\boldsymbol{x}}_{k+1|k}, \boldsymbol{\theta}_k)\} \\ \mathbf{H}_k &= \mathbf{P}_{\boldsymbol{x}_k\boldsymbol{z}_k}\mathbf{P}_{\tilde{\boldsymbol{z}}_k}^{-1} \end{aligned}$$

where $\hat{\boldsymbol{x}}_{k+1|k}$ and $\hat{\boldsymbol{z}}_{k+1}$ are the predictions of $\boldsymbol{x}_{k+1}$ and $\boldsymbol{z}_{k+1}$, respectively, and $\mathbf{P}_{\boldsymbol{x}_k\boldsymbol{z}_k}$ is cross-covariance of $\boldsymbol{x}_k$ and $\boldsymbol{z}_k$. For the linear case, the KF computes all these terms exactly. As we will see, the EKF and the UKF use different methods to find an approximate solution.

Although both the EKF and the UKF can handle the general nonlinear case where the noises have arbitrary distributions, for the sake of simplicity, in this paper we will only consider systems with process and measurement noise that are additive and independent zero-mean Gaussian white-noise sequences. That is, the system can be written as

$$\begin{aligned} \boldsymbol{x}_{k+1} &= \boldsymbol{f}(\boldsymbol{x}_k, \boldsymbol{u}_k) + \mathbf{L}_k\boldsymbol{\xi}_k \\ \boldsymbol{z}_k &= \boldsymbol{g}(\boldsymbol{x}_k) + \mathbf{D}_k\boldsymbol{\theta}_k \end{aligned}$$

and the noise signals satisfy

$$\left.\begin{aligned} \mathsf{E}\{\boldsymbol{\xi}_k\} &= \mathbf{0} \\ \mathsf{E}\{\boldsymbol{\theta}_k\} &= \mathbf{0} \end{aligned}\right\} \text{ for } k = 0, 1, 2, \ldots$$

$$\left.\begin{aligned} \mathsf{cov}[\boldsymbol{\xi}_i; \boldsymbol{\xi}_j] &= \boldsymbol{\Xi}\delta_{ij} \\ \mathsf{cov}[\boldsymbol{\theta}_i; \boldsymbol{\theta}_j] &= \boldsymbol{\Theta}\delta_{ij} \\ \mathsf{cov}[\boldsymbol{\xi}_i; \boldsymbol{\theta}_j] &= \mathbf{0} \end{aligned}\right\} \text{ for all } i, j$$

We will denote this as $\boldsymbol{\xi}_k \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Xi})$ and $\boldsymbol{\theta}_k \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Theta})$.

The paper is organized as follows. In Section II, we summarize the inner workings of the EKF. In Section III, we describe the unscented transformation and its role in

In partial fulfillment of the requirements for Dynamic Stochastic Filtering, Prediction, and Smoothing (2007/2008), taught by Prof. Paulo Oliveira at Instituto Superior Técnico.

E-mail: `jalmeida@isr.ist.utl.pt`

development of the UKF. Most of the theoretical content of Sections II and III is adapted from [4] with a few simplifications and changes in notation. For comparison purposes, in Section IV an illustrative example based on a simplified model of a real vehicle is introduced. Simulation results using both filters are presented and differences in performance are discussed. Finally, some concluding remarks are given in Section V.

II. The Extended Kalman Filter

In this case, the predictions are approximated as

$$\begin{aligned}\hat{\boldsymbol{x}}_{k+1|k} &= \boldsymbol{f}(\hat{\boldsymbol{x}}_{k|k}, \boldsymbol{u}_k)\\ \hat{\boldsymbol{z}}_{k+1} &= \boldsymbol{g}(\hat{\boldsymbol{x}}_{k+1|k})\end{aligned}$$

which is the same as assuming that the expected value operator $\mathsf{E}\{\cdot\}$ commutes with both $\boldsymbol{f}$ and $\boldsymbol{g}$. To propagate the covariance, the EKF resort to the linearization of the system about the current estimated state

$$\begin{aligned}\boldsymbol{x}_{k+1} &\approx \hat{\mathbf{A}}_k \boldsymbol{x}_k + \mathbf{L}_k \boldsymbol{\xi}_k\\ \boldsymbol{z}_{k+1} &\approx \hat{\mathbf{C}}_k \boldsymbol{x}_k + \mathbf{D}_k \boldsymbol{\theta}_k\end{aligned}$$

where

$$\hat{\mathbf{A}}_k = \left.\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}\right|_{\substack{\boldsymbol{x}=\hat{\boldsymbol{x}}_k\\ \boldsymbol{u}=\boldsymbol{u}_k}} \qquad \hat{\mathbf{C}}_k = \left.\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{x}}\right|_{\substack{\boldsymbol{x}=\hat{\boldsymbol{x}}_k\\ \boldsymbol{u}=\boldsymbol{u}_k}}$$

Note that these matrices must be computed at every time step. It is then a matter of applying the KF covariance propagation scheme to obtain the EKF equations.

Initialization: $\hat{\boldsymbol{x}}_{0|0} = \bar{\boldsymbol{x}}_0, \mathbf{P}_{0|0} = \boldsymbol{\Sigma}_0$
Main cycle: for $k = 0, 1, 2, \ldots$

1) Predict step:

$$\begin{aligned}\hat{\boldsymbol{x}}_{k+1|k} &= \boldsymbol{f}(\hat{\boldsymbol{x}}_{k|k}, \boldsymbol{u}_k)\\ \mathbf{P}_{k+1|k} &= \hat{\mathbf{A}}_k \mathbf{P}_{k|k} \hat{\mathbf{A}}_k^\top + \mathbf{L}_k \boldsymbol{\Xi} \mathbf{L}_k^\top\end{aligned}$$

2) Measurement update step:

$$\begin{aligned}\hat{\boldsymbol{z}}_{k+1} &= \boldsymbol{g}(\hat{\boldsymbol{x}}_{k+1|k})\\ \mathbf{H}_{k+1} &= \mathbf{P}_{k+1|k} \hat{\mathbf{C}}_k^\top \left[\hat{\mathbf{C}}_k \mathbf{P}_{k+1|k} \hat{\mathbf{C}}_k^\top + \mathbf{D}_k \boldsymbol{\Theta} \mathbf{D}_k^\top\right]^{-1}\\ \hat{\boldsymbol{x}}_{k+1|k+1} &= \hat{\boldsymbol{x}}_{k+1|k} + \mathbf{H}_{k+1}(\boldsymbol{z}_{k+1} - \hat{\boldsymbol{z}}_{k+1})\\ \mathbf{P}_{k+1|k+1} &= \left[\mathbf{I} - \mathbf{H}_{k+1}\hat{\mathbf{C}}_k\right] \mathbf{P}_{k+1|k}\end{aligned}$$

The need to compute Jacobians may be troublesome and computationally intensive for some nonlinear functions. Next, we present an alternative algorithm that although its main focus is on predicting the mean and covariance more accurately, its implementation is more straightforward.

III. The Unscented Kalman Filter

A. The unscented transformation

The unscented transformation (UT) is a method for calculating the statistics of a random variable which undergoes a nonlinear transformation. Consider propagating a random variable $\boldsymbol{x}$ (of dimension L) through a nonlinear function $\boldsymbol{z} = \boldsymbol{f}(\boldsymbol{x})$. Assume $\boldsymbol{x}$ has mean $\bar{\boldsymbol{x}}$ and covariance $\mathbf{P}_{\boldsymbol{x}}$. To calculate the statistics of $\boldsymbol{z}$, we start by generating a set of points, called sigma points, whose sample mean and covariance match those of $\boldsymbol{x}$. The nonlinear function $\boldsymbol{f}(\cdot)$ is then applied to each of these sigma points that yields a set of transformed points. The predicted mean and covariance of $\boldsymbol{z}$ are then computed from these transformed points. At first sight, this may resemble a Monte Carlo method. However, the sample points are not drawn at random: they are deterministically chosen so that they capture specific information about the distribution.

Formally, the UT consist of the following steps.

1) Form the set of $2L+1$ sigma points from the columns of the $L \times L$ matrix $\sqrt{(L+\lambda)\mathbf{P}_{\boldsymbol{x}}}$, as follows

$$\begin{aligned}\mathcal{X}_0 &= \bar{\boldsymbol{x}}\\ \mathcal{X}_i &= \bar{\boldsymbol{x}} + \left(\sqrt{(L+\lambda)\mathbf{P}_{\boldsymbol{x}}}\right)_i, \quad i = 1, \ldots, L\\ \mathcal{X}_i &= \bar{\boldsymbol{x}} - \left(\sqrt{(L+\lambda)\mathbf{P}_{\boldsymbol{x}}}\right)_{i-L}, \quad i = L+1, \ldots, 2L\end{aligned}$$

where $(\mathbf{X})_i$ denotes the ith column of matrix $\mathbf{X}$, and compute the associated weights

$$\begin{aligned}W_0^{(m)} &= \lambda/(L+\lambda)\\ W_0^{(c)} &= \lambda/(L+\lambda) + (1-\alpha^2+\beta)\\ W_i^{(m)} &= W_i^{(c)} = 1/\{2(L+\lambda)\}, \quad i = 1, \ldots, 2L\end{aligned}$$

2) Transform each of the sigma points as

$$\mathcal{Z}_i = \boldsymbol{f}(\mathcal{X}_i), \quad i = 0, \ldots, 2L$$

3) Mean and covariance estimates for $\boldsymbol{z}$ are computed as

$$\begin{aligned}\bar{\boldsymbol{z}} &= \sum_{i=0}^{2L} W_i^{(m)} \mathcal{Z}_i\\ \mathbf{P}_{\boldsymbol{z}} &= \sum_{i=0}^{2L} W_i^{(c)} \left(\mathcal{Z}_i - \bar{\boldsymbol{z}}\right)\left(\mathcal{Z}_i - \bar{\boldsymbol{z}}\right)^\top\end{aligned}$$

4) The cross-covariance of $\boldsymbol{x}$ and $\boldsymbol{z}$ is estimated as

$$\mathbf{P}_{\boldsymbol{x}\boldsymbol{z}} = \sum_{i=0}^{2L} W_i^{(c)} \left(\mathcal{X}_i - \bar{\boldsymbol{x}}\right)\left(\mathcal{Z}_i - \bar{\boldsymbol{z}}\right)^\top$$

The UT has several tuning parameters: $\lambda = \alpha^2(L+\kappa) - L$ is a scaling parameter; α determines the spread of the sigma points around $\bar{\boldsymbol{x}}$ and is usually between 1 and 10^{-4}; κ is a secondary scaling parameter which is usually set to 0 or $3 - L$; and, β is used to incorporate prior knowledge of the distribution of $\boldsymbol{x}$ (for Gaussian distributions, $\beta = 2$ is optimal).

The only computationally intensive operation in the UT is the computation of a matrix square root in Step 1. This operation can be efficiently executed by using a Cholesky decomposition.

Note that we can represent the unscented transformation as a function with the following syntax

$$[\bar{\boldsymbol{z}}, \mathbf{P}_{\boldsymbol{z}}, \mathbf{P}_{\boldsymbol{x}\boldsymbol{z}}] = \mathsf{UT}\{\boldsymbol{f}(\cdot), \bar{\boldsymbol{x}}, \mathbf{P}_{\boldsymbol{x}}\}$$

B. *UKF equations*

The UKF is a discrete-time filtering algorithm that uses the UT for computing approximate solutions to the state-estimation problem. The UT used twice: once in the predict step and again in the measurement update step in order to propagate the mean and covariance through functions $\boldsymbol{f}(\cdot)$ and $\boldsymbol{g}(\cdot)$, respectively. The initialization is the same as the one performed by the EKF.

Initialization: $\hat{\boldsymbol{x}}_{0|0} = \bar{\boldsymbol{x}}_0, \mathbf{P}_{0|0} = \boldsymbol{\Sigma}_0$
Main cycle: for $k = 0, 1, 2, \dots$

1) Predict step:

$$
\begin{aligned}
[\hat{\boldsymbol{x}}_{k+1|k}, \mathbf{P}_{k+1|k}] &= \mathsf{UT}\{\boldsymbol{f}(\cdot,\cdot), \hat{\boldsymbol{x}}_{k|k}, \mathbf{P}_{k|k}\} \\
\mathbf{P}_{k+1|k} &= \mathbf{P}_{k+1|k} + \mathbf{L}_k \boldsymbol{\Xi} \mathbf{L}_k^\top
\end{aligned}
$$

2) Measurement update step:

$$
\begin{aligned}
[\hat{\boldsymbol{z}}_{k+1}, \mathbf{P}_{\tilde{\boldsymbol{z}}}, \mathbf{P}_{\boldsymbol{x}\boldsymbol{z}}] &= \mathsf{UT}\{\boldsymbol{g}(\cdot), \hat{\boldsymbol{x}}_{k+1|k}, \mathbf{P}_{k+1|k}\} \\
\mathbf{P}_{\tilde{\boldsymbol{z}}} &= \mathbf{P}_{\tilde{\boldsymbol{z}}} + \mathbf{D}_k \boldsymbol{\Theta} \mathbf{D}_k^\top \\
\mathbf{H}_{k+1} &= \mathbf{P}_{\boldsymbol{x}\boldsymbol{z}} \mathbf{P}_{\tilde{\boldsymbol{z}}}^{-1} \\
\hat{\boldsymbol{x}}_{k+1|k+1} &= \hat{\boldsymbol{x}}_{k+1|k} + \mathbf{H}_{k+1}(\boldsymbol{z}_{k+1} - \hat{\boldsymbol{z}}_{k+1}) \\
\mathbf{P}_{k+1|k+1} &= \mathbf{P}_{k+1|k} - \mathbf{H}_{k+1} \mathbf{P}_{\tilde{\boldsymbol{z}}} \mathbf{H}_{k+1}
\end{aligned}
$$

The UKF uses the maps $\boldsymbol{f}$ and $\boldsymbol{g}$ directly and is therefore derivative-free which makes it easier to implement.

Like the EKF, there is also a continuous version of the UKF filter. See [5] for details.

IV. An illustrative example

Consider an autonomous marine surface craft (ASC) moving along straight lines. A simplified model that describes it movement is given by the nonlinear differential equation

$$
m\ddot{x}(t) + a_1 \dot{x}(t)|\dot{x}(t)| + a_2 \dot{x}^3(t) = b(u(t) + w(t))
$$

where m is the vehicle's mass, $x(t)$ and $\dot{x}(t)$ denote the vehicle's position and velocity, respectively, $u(t)$ represents the force generated by the propulsion system, a_1, a_2, and b are drag force and propulsion coefficients, and the term $w(t)$ represents an external disturbance, which is caused by waves, that affects the propulsion system.

A. *Plant model*

Let $x_1(t) = x(t)$ and $x_2(t) = \dot{x}(t)$. Then, the ASC dynamics can be written as

$$
\begin{aligned}
\dot{x}_1 &= x_2 \\
\dot{x}_2 &= -\frac{a_1}{m} x_2 |x_2| - \frac{a_2}{m} x_2^3 + \frac{b}{m}(u + w)
\end{aligned}
$$

The waves are described by their power spectral density

$$
\Phi_{ww}(\omega) = \frac{100\omega^2}{81\omega^4 + 18\omega^2 + 1}
$$

This kind of disturbance can be generated by applying zero-mean Gaussian white-noise to a LTI system described by the transfer function

$$
H(s) = \frac{10s}{9s^2 + 6s + 1}
$$

or equivalently, by a state-space representation of it such as

$$
\begin{aligned}
\dot{x}_{w1} &= x_{w2} \\
\dot{x}_{w2} &= -\frac{1}{9} x_{w1} - \frac{2}{3} x_{w2} + \frac{1}{9} \xi_w \\
w &= 10 x_{w2}
\end{aligned}
$$

where $\boldsymbol{x}_w(0) \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_{w0})$ with

$$
\boldsymbol{\Sigma}_{w0} = \begin{bmatrix} 0 & 0 \\ 0 & \frac{1}{81} \end{bmatrix}
$$

and $\xi_w(t) \sim \mathcal{N}(0, 1)$.

On board the ASC, there are two sensors: one measuring x_1 and another measuring x_2. The first sensor is modeled as

$$
z_1 = x_1 + \theta_1
$$

where $\theta_1(t) \sim \mathcal{N}(0, \sigma_1^2)$, and the second one as

$$
z_2 = x_2 + \theta_2
$$

where $\theta_2 \sim \mathcal{N}(\bar{\theta}_2, \sigma_2^2)$ with $\bar{\theta}_2 = 0.001\,\mathrm{m/s}$. Since θ_2 is not zero mean, we must write it as the output of the linear system

$$
\dot{x}_s = 0 \qquad \theta_2 = x_s + \sigma_2 \xi_s
$$

where $x_s(0) = \bar{\theta}_2$ and $\xi_s \sim \mathcal{N}(0, 1)$.

The parameters a_1, a_2, and b of the ASC are assumed unknown and must be estimated by the filter. For this purpose, we introduce three new state variables x_3, x_4, and x_5 that represent the parameters a_1, a_2, and b, respectively. That is, $\dot{x}_3(t) = \dot{x}_4(t) = \dot{x}_5(t) = 0$, $x_3(0) = a_1$, $x_4(0) = a_2$, and $x_5(0) = b$.

Merging plant, waves and sensors into one single system with state variable

$$
\boldsymbol{x} = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 & x_5 & x_{w1} & x_{w2} & x_s \end{bmatrix}^\top
$$

we get the dynamics

$$
\dot{\boldsymbol{x}} = \begin{bmatrix} x_2 \\ -\frac{1}{m}\left(x_3 x_2 |x_2| + x_4 x_2^3 + x_5(u + 10x_{w2})\right) \\ 0 \\ 0 \\ 0 \\ x_{w2} \\ -\frac{1}{9}x_{w1} - \frac{2}{3}x_{w2} + \frac{1}{9}\xi_w \\ 0 \end{bmatrix}
$$

with output function

$$
\boldsymbol{z} = \begin{bmatrix} x_1 + \sigma_1 \xi_{s1} \\ x_2 + x_s + \sigma_2 \xi_{s2} \end{bmatrix}
$$

where $\xi_w(t) \sim \mathcal{N}(0, 1)$ and $\boldsymbol{\xi}_s(t) \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_2)$.

Since the state is to be estimated by discrete-time filters, the model must be discretized. To this end, the original continuous-time (extended) plant

$$
\begin{aligned}
\dot{\boldsymbol{x}}(t) &= \boldsymbol{f}(\boldsymbol{x}(t), u(t), \xi_w(t)) \\
\boldsymbol{z}(t) &= g(\boldsymbol{x}(t), \boldsymbol{\xi}_s(t))
\end{aligned}
$$

is discretized using a step size (or sampling period) of h seconds,

$$
\boldsymbol{x}((k+1)h) = \boldsymbol{x}(kh) + h\boldsymbol{f}(\boldsymbol{x}(kh), u(kh), \xi_w(kh))
$$

$$\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\frac{1}{m}(2x_3|x_2| + 3x_4x_2^2) & -\frac{1}{m}x_2|x_2| & -\frac{1}{m}x_2^3 & \frac{1}{m}(u + 10x_{w2}) & 0 & \frac{10}{m}x_5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -\frac{1}{9} & -\frac{2}{3} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

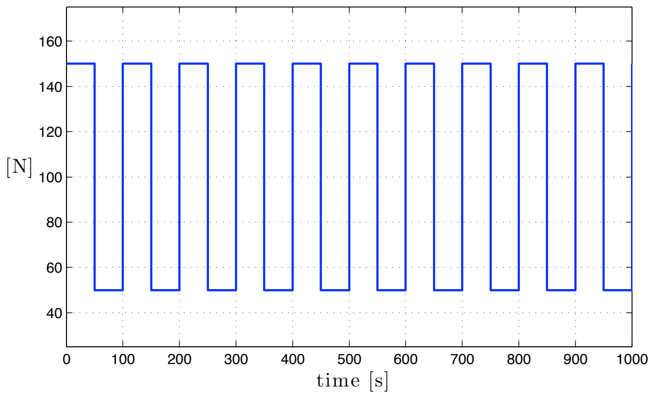

Fig. 1. Control input applied to the system.

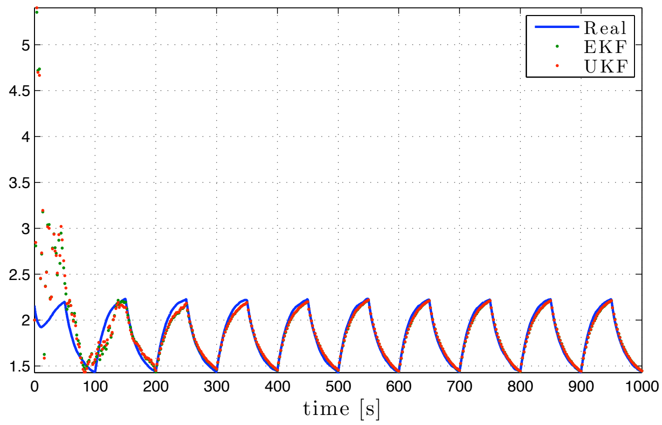

Fig. 2. Evolution of $\mathsf{E}\{x_2\}$ for configuration ($\mathcal{P}_2$, $\mathcal{S}_2$).

From here on, we will write $\boldsymbol{x}_k$ for $\boldsymbol{x}(kh)$. Applying this discretization procedure to our system, we get

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + h\boldsymbol{f}(\boldsymbol{x}_k, u_k, \xi_{w,k})$$

with output function

$$\boldsymbol{z}_k = \begin{bmatrix} x_{1,k} + \frac{\sigma_1}{h}\xi_{s1,k} \\ x_{2,k} + x_{s,k} + \frac{\sigma_2}{h}\xi_{s2,k} \end{bmatrix}$$

The matrices needed by the EKF, which are obtained by linearization, are

$$\hat{\mathbf{A}}_k = \mathbf{I}_8 + h\left.\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}\right|_{\substack{\boldsymbol{x}=\hat{\boldsymbol{x}}_k \\ u=u_k \\ \xi_w=0}}$$

$$\mathbf{L}_k = h\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{9} & 0 \end{bmatrix}^\top$$

$$\hat{\mathbf{C}}_k = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\mathbf{D}_k = \frac{1}{h}\begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix}, \boldsymbol{\Xi} = 1, \boldsymbol{\Theta} = \mathbf{I}_2$$

where $\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}$ is shown at the top of the page.

B. Simulation results

The sampling period was set to $h = 2\,\mathrm{s}$. The control input $u(t)$ applied to the system is shown in Fig. 1. The initial state of the system was randomly generated according to $\boldsymbol{x}(0) \sim \mathcal{N}(\bar{\boldsymbol{x}}_0, \boldsymbol{\Sigma}_0)$ where

$$\bar{\boldsymbol{x}}_0 = \begin{bmatrix} 0 & 2 & \hat{x}_3(0) & \hat{x}_4(0) & \hat{x}_5(0) & 0 & 0 & \bar{\theta}_2 \end{bmatrix}^\top$$

and

$$\boldsymbol{\Sigma}_0 = \mathrm{diag}(20, 5, 10, 5, 5, 0, \frac{1}{81}, 0)$$

We ran simulations for 4 different configurations of parameters and sensor covariances. The parameters and their initial estimates belong to one of the two following sets

$$\begin{aligned} \mathcal{P}_1 = \{&a_1 = 25, a_2 = 0, b = 1; \\ &\hat{x}_3(0) = -5, \hat{x}_4(0) = 0, \hat{x}_5(0) = 10\} \\ \mathcal{P}_2 = \{&a_1 = 25, a_2 = 2, b = 1; \\ &\hat{x}_3(0) = -5, \hat{x}_4(0) = -5, \hat{x}_5(0) = 10\} \end{aligned}$$

while the sensor covariances belong to one of the two following sets

$$\begin{aligned} \mathcal{S}_1 &= \{\sigma_1^2 = 1\,\mathrm{m}^2, \sigma_2^2 = 0.04\,(\mathrm{m/s})^2\} \\ \mathcal{S}_2 &= \{\sigma_1^2 = 10\,\mathrm{m}^2, \sigma_2^2 = 0.4\,(\mathrm{m/s})^2\} \end{aligned}$$

For each configuration, 10 independent simulations of duration $1000\,\mathrm{s}$ were carried out and the results were averaged.

The evolution of the excepted values of the states x_2, x_3, x_4, and x_5 for configuration $(\mathcal{P}_2, \mathcal{S}_2)$ are show in Fig. 2–5. The evolution of the expected value of the error squared is plotted in Fig. 6.

To compare the two filters, we computed for each error its root-mean-squared (RMS) value as

$$(\tilde{x})_{\mathrm{RMS}} = \sqrt{\frac{1}{N}\sum_{k=0}^{N-1} \mathsf{E}\{\tilde{x}_k^2\}}$$

for the scalar case and

$$(\tilde{\boldsymbol{x}})_{\mathrm{RMS}} = \sqrt{\frac{1}{N}\sum_{k=0}^{N-1} \mathsf{E}\{\tilde{\boldsymbol{x}}_k^\top \tilde{\boldsymbol{x}}_k\}}$$

for the vector case. We did this for each state variable and then for the full state. The values obtained are given in Table I.

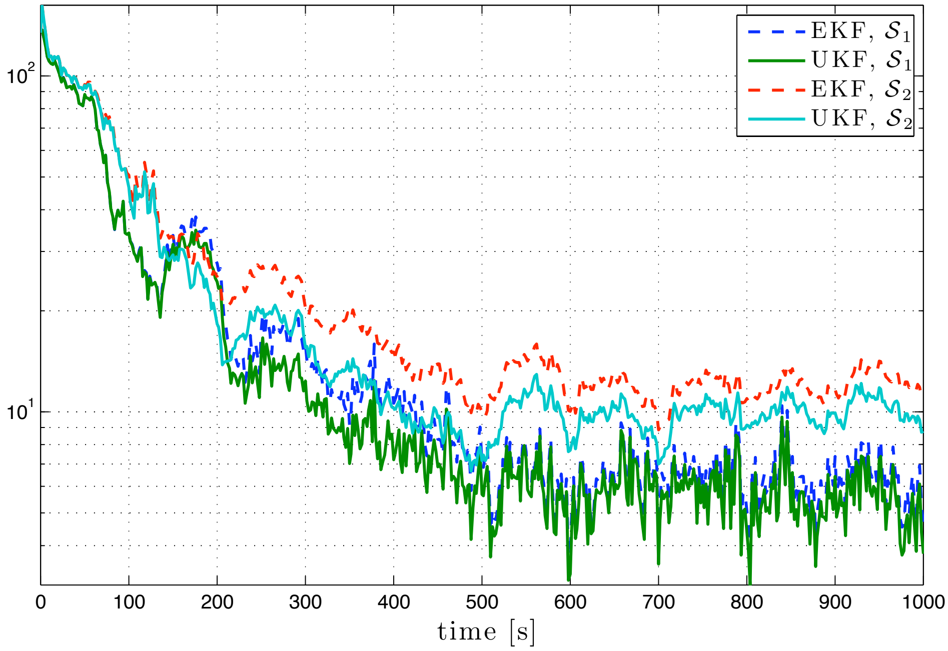

Fig. 6. Evolution of $\mathsf{E}\{\tilde{\boldsymbol{x}}^\top \tilde{\boldsymbol{x}}\}$ for both filters with parameters $\mathcal{P}_2$ under both sensor characteristics.

TABLE I

RMS VALUES OF BOTH FILTERS FOR DIFFERENT PARAMETER AND SENSOR CONFIGURATIONS.

Parameters	**Sensors**	**Filter**	**RMS values**							
			$\tilde{x}_1$	$\tilde{x}_2$	$\tilde{x}_3$	$\tilde{x}_4$	$\tilde{x}_5$	$\tilde{x}_{w1}$	$\tilde{x}_{w2}$	$\tilde{\boldsymbol{x}}$
$\mathcal{P}_1$	$\mathcal{S}_1$	EKF	4.092	0.187	7.556	—	0.739	4.390	1.319	**9.769**
		UKF	4.095	0.187	7.528	—	0.739	4.386	1.318	**9.747**
	$\mathcal{S}_2$	EKF	5.079	0.449	9.763	—	0.952	3.169	1.028	**11.547**
		UKF	5.064	0.449	9.700	—	0.950	3.169	1.028	**11.487**
$\mathcal{P}_2$	$\mathcal{S}_1$	EKF	3.882	0.198	9.939	3.016	0.753	4.384	1.328	**12.022**
		UKF	3.881	0.197	9.516	2.718	0.751	4.352	1.319	**11.588**
	$\mathcal{S}_2$	EKF	5.146	0.499	11.345	5.222	1.067	3.333	1.059	**14.003**
		UKF	5.029	0.497	10.610	4.450	1.045	3.278	1.050	**13.071**

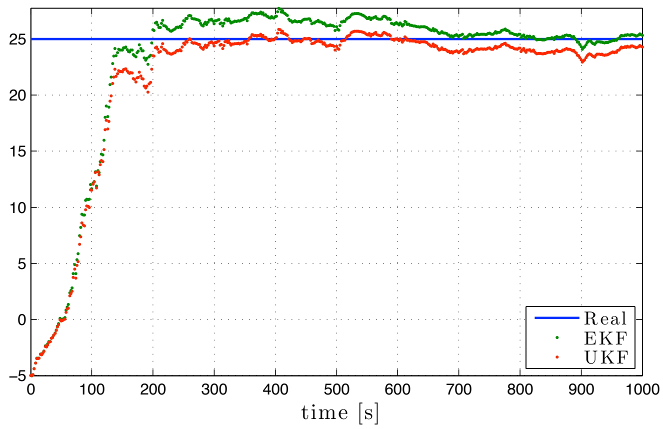

Fig. 3. Evolution of $\mathsf{E}\{x_3\}$ for configuration ($\mathcal{P}_2$, $\mathcal{S}_2$).

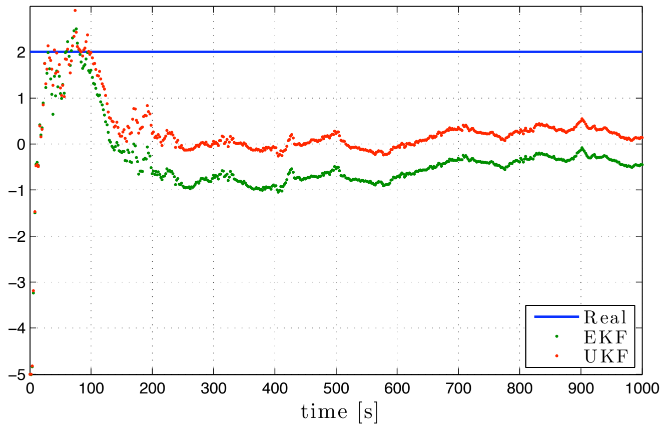

Fig. 4. Evolution of $\mathsf{E}\{x_4\}$ for configuration ($\mathcal{P}_2$, $\mathcal{S}_2$).

C. Discussion

By looking at Fig. 2, it is impossible to see a clear difference between the two filters, as both expected values seem to overlap as time progresses. When it comes to estimating the parameters, both filters have similar evolutions and although both of them correctly estimate a_1 (Fig. 3), the estimates of parameters a_2 (Fig. 4) and b (Fig. 5) are biased.

If we look at the evolution of the expected error squared (Fig. 6), we see that both filters perform better with lower noise covariances as expected, that is, lower RMS values are

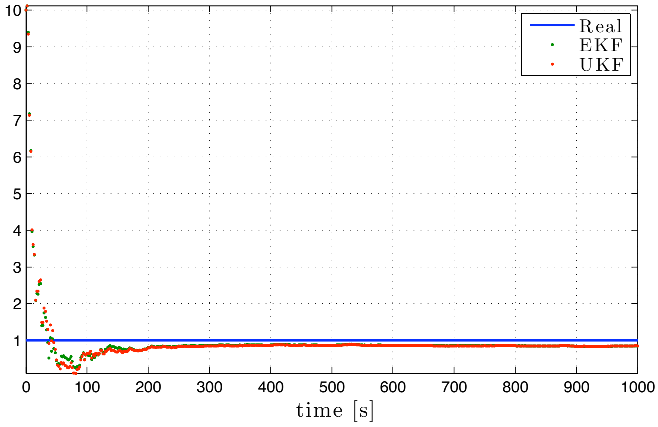

Fig. 5. Evolution of $\mathsf{E}\{x_5\}$ for configuration ($\mathcal{P}_2$, $\mathcal{S}_2$).

achieved using sensor characteristics $\mathcal{S}_1$. We also start to see some differences in the performance of both filters, has the evolution of the squared error of the UKF is most of the time below that of the EKF. Table I shows us that there is an actual decrease of the RMS values when the UKF filter is used. Although for some configurations the performance of both filter is basically the same, in others there is a clear overall improvement that does not seem to be evenly spread by all state variables.

A clear gain in performance by the UKF seems to be only visible in "very" nonlinear systems and under severe noise regimes. Only when the term x_2^3 is included in the dynamics of the ASC (that is, when $\mathcal{P}_2$ is selected), does the difference in performance between both filters stands out. If the measurement equations or output functions were also nonlinear, a more distinguishable change may have been visible.

V. Conclusion

We addressed the problem of estimating the state of a nonlinear system perturbed by additive sources of zero-mean Gaussian white-noise. The Extended Kalman Filter solves this problem by dealing with a linearized version of the system and then applying the standard Kalman Filter equations. A new filter called Unscented Kalman Filter (UKF) is introduced, that uses the unscented transformation to propagate the mean and covariance of the state in a different manner.

By resorting to a simplified model of a real vehicle, a nonlinear system is derived and used to compare the performance of both filters. It is shown that only when more nonlinear terms are included or under severe noise situations, does the gain in performance of the UKF stands out.

References

[1] S. Julier, J. Uhlmann, and H. F. Durrant-Whyte, "A new approach for filtering nonlinear systems," in *Proc. of Am. Contr. Conf.*, Seattle, WA, 21–23 Jun. 1995, pp. 1628–1632.

[2] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. New Jersey, USA: Prentice-Hall, 1979.

[3] S. Julier, J. Uhlmann, and H. F. Durrant-Whyte, "A new method for nonlinear transformation of means and covariances in filters and estimators," *IEEE Trans. Autom. Control*, vol. 45, no. 3, pp. 477–482, Mar. 2000.

[4] E. A. Wan and R. van der Merwe, *The Unscented Kalman Filter*, ser. Kalman Filtering and Neural Networks. John Wiley and Sons, Inc., 2001, ch. 7.

[5] S. Särkkä, "On unscented kalman filtering for state estimation of continuous-time nonlinear systems," *IEEE Trans. Autom. Control*, vol. 52, no. 9, pp. 1631–1641, Sep. 2007.