



**HAL**  
open science

## On the convergence of the unscented Kalman filter

Assia Daid, Eric Busvelle, Mohamed Aidene

► **To cite this version:**

Assia Daid, Eric Busvelle, Mohamed Aidene. On the convergence of the unscented Kalman filter. European Journal of Control, In press, 10.1016/j.ejcon.2020.05.003 . hal-02891724

**HAL Id: hal-02891724**

**<https://hal.science/hal-02891724>**

Submitted on 7 Jul 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# On the convergence of the unscented Kalman filter

Assia Daid<sup>a,b,\*</sup>, Eric Busvelle<sup>a</sup>, Mohamed Aidene<sup>b</sup>

<sup>a</sup>LIS, UMR CNRS 7020, Université de Toulon,  
Aix Marseille Univ., France

<sup>b</sup>Laboratoire de Conception et Conduite des Systèmes de Production,  
Université Mouloud Mammeri, Tizi-Ouzou, Algeria

---

## Abstract

A convergence analysis of the modified unscented Kalman filter (UKF), used as an observer for a class of nonlinear deterministic continuous time systems, is presented. Under certain conditions, the extended Kalman filter (EKF) is an exponential observer for non-linear systems, i.e., the dynamics of the estimation error is exponentially stable. It is shown that unlike the EKF, the UKF is not an exponentially converging observer. A modification of the UKF – the unscented Kalman observer – is proposed, which is a better candidate for an observer.

This paper is a first step towards a proof of the global convergence of the high-gain version of the UKO.

*Keywords:* Nonlinear observers, Unscented Kalman filter, Unscented Kalman Observer.

---

## 1. Introduction

$y(t) = AK(1 - e^{-\frac{t}{\tau}})H(t)$  For technological reasons (material, feasibility), reliability (failure of measurements) or economic (cost of sensors), in many applications the measurement of the whole state of a system is not possible, but the knowledge of the state at any moment is essential, for example: the synthesis of a command [2] or the diagnosis [44] and detection of defects in an industrial process [52].

State observers (sometimes called soft sensors or intelligent sensors) are a way to handle this problem. The task of an observer is to provide an estimate (in some sense, asymptotic or exponential) of the state of a studied system according to the information available about this system (the input control and output measurements and the dynamic model of the process, [51]).

The first observers dedicated to the estimation of the state of a linear system that are characterized by necessary and sufficient conditions, were published around the 1960s by Kalman (Kalman filter, [25]) in a stochastic framework and subsequently by Luenberger in a deterministic framework (Luenberger observer, [31, 32]). These two observers are widely used today, but linear systems cover only a small percentage of industrial processes. Indeed, most real applications have non-linear behavior, which has prompted the development of non-linear filters and observers ([7, 11, 47]).

Among these non-linear systems, we consider those non-linear dynamic systems in continuous time which have the form

$$\begin{cases} \frac{dx(t)}{dt} = f(x(t), t) \\ y(t) = h(x(t), t) \end{cases} \quad (1)$$

$x(t) \in \mathbb{R}^n$ ,  $y(t) \in \mathbb{R}^p$ . Since  $f$  and  $h$  depend explicitly on the time  $t$ , we omit writing explicitly the control variable  $u(t)$  without loss of generality. The extended Kalman filter (EKF) is an important and widely used tool for state estimation in such non-linear dynamic systems [8, 18]. It is based on a linearization of the system along the estimated trajectory. However, this linearization can lead to poor performance (and even the divergence of the filter for highly non-linear problems), and the required derivation of Jacobian matrices may not be easy in many applications (see, for instance, [27] and a way to work around the problem). Although the linearization around the estimated trajectory of the system usually prevents analytically proving the observer's convergence for any initial error, such proofs exist in

---

\*This work is carried out with the contribution of the Hubert Curien Tassili Research Program (PHC)

\*Corresponding author

the deterministic and stochastic cases when the initial estimation errors and disturbing noise are small enough (see [3, 6, 37, 39, 38, 45, 41, 40]).

A recent improvement to the EKF is the unscented Kalman filter (UKF) [20, 21]. The UKF approximates the probability density resulting from the non-linear transformation of a random variable instead of approximating the non-linear functions with a Taylor series expansion. The approximation is done by evaluating the non-linear function on a minimal set of carefully chosen sample points [42, 43, 22, 24]. Consequently, the UKF has received great attention in many areas such as control theory, aerospace navigation, eye tracking, and information fusion [53, 29]. This approach may be seen as a Monte Carlo approach, although the number of points used to propagate the law of probability remains small (see [15]).

To summarize, there exist two approaches to non-linear state estimation.

- The first one is non-linear filtering, which consists in studying the probability law of the state of a stochastic dynamic system conditionally on the observations. The extended Kalman filter is based upon an approximation by a normal distribution of a pair of random vectors  $X \in \mathbb{R}^n$  and  $Y \in \mathbb{R}^p$  such that  $X$  is a Gaussian random vector obeying the law  $N(m, P)$  and  $Y = g(X)$  where  $g$  is a non-linear transformation (see [36, 35]). The usual approximation of the law of the pair  $(X, Y)$  uses the first-order approximation

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim \mathcal{N} \left( \begin{pmatrix} m \\ g(m) \end{pmatrix}, \begin{pmatrix} P & PA^t \\ AP & APA^t \end{pmatrix} \right) \quad (2)$$

where  $A$  is the Jacobian matrix of  $g$  at the point  $m$ . To better take into account the non-linear transformation of a Gaussian vector, the unscented transformation has been developed, which consists of writing

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim \mathcal{N} \left( \begin{pmatrix} m \\ \mu_U \end{pmatrix}, \begin{pmatrix} P & C_U \\ C_U^t & S_U \end{pmatrix} \right) \quad (3)$$

where the vector  $\mu_U$  and the matrices  $C_U$  and  $S_U$  are approximated from  $g(m)$  and from the image by  $g$  of  $2n + 1$   $\sigma$ -points, ( $2n$  points are placed around  $m$ , we will explain further this construction in Section 2.1, (8), (9) and (10), see also [22, 48, 23, 1]). This non-linear version of the Kalman filter has several advantages, the first being a better consideration of the propagation of a Gaussian noise in a non-linear system. The second advantage is that this non-linear version does not require calculating two Jacobian matrices, which are sometimes complex and sources of numerical instabilities (see [12] for a concrete example).

- The second approach to the problem of estimating the state of a system is purely deterministic. It consists in constructing another dynamic system ('the observer'), which uses as input the available measurements and whose state converges asymptotically (and generally exponentially) to the state of the system [4, 14]. A classical way to construct an observer is to apply a stochastic filter in the deterministic framework. In this approach, the matrices  $Q$  and  $R$  that represent the covariance matrices of the state noise and the measurement noise are interpreted as tuning parameters (and corresponds to quadratic cost matrices in a quadratic cost minimization framework).

There are some techniques that show that the observer/filter (EKF, UKF), converges exponentially locally to the state of the system under certain conditions (observability). Some authors use the Lyapunov method [6, 40], or the contraction theory, developed by Lohmiller and Slotine [30, 33] for estimating the state of the stochastic or deterministic non-linear systems. In the stochastic framework, a novel stochastic analysis based on exponential concentration inequalities and uniform  $\chi$ -square type estimates for stochastic quadratic type systems, has been adapted to complement the above-mentioned techniques. For more details, see [10, 26].

Compared to the extended Kalman filter, little research has been done to analyse the convergence of the UKF. Some proofs have been established for non-linear stochastic discrete time systems (see [49, 28, 13, 17, 16]) and in the stochastic continuous-time case (see [50]). However, in the case of nonlinear stochastic systems with linear state equation, such as the tightly coupled INS/GPS integrated system [17], the use of the UKF causes a great amount of redundant computation in the prediction process, because the system state equation is used to propagate  $2n + 1$   $\sigma$ -points. For remedy this the authors have proposed a derivative unscented Kalman filter (DUKF) to reduce the computational complexity by uses the original Kalman filter (KF) in the prediction process and they have analyse the stochastic stability of the DUKF in [16].

Despite its interest (due to the fact that it does not require a priori Jacobian computation), UKF convergence has not been studied in a deterministic setting. Two reasons may explain this lack: the first one is that the UKF is an improvement of EKF only because it tackles Gaussian noise more efficiently. The second one is because without modification, it is not convergent, as we will see in Section 2.3. We claim that our study explains theoretically several experimental studies comparing EKF and UKF (see, e.g. [34, 46]) as well as what can be expected from the UKF when the noise is small (and how to tune the UKF in such cases).

The structure of this paper is as follows. In Section 2, we recall the unscented transformation and the unscented Kalman filter, and we show that the unscented Kalman filter does not converge in a deterministic setting, as soon as the system is non-linear. Therefore, in Section 3, we propose a modification which avoids the bias of the UKF. This new algorithm is called the unscented Kalman observer (UKO). In Section 4 we prove the convergence of the UKO, as a deterministic observer for continuous-time systems, when the initial estimation errors are small enough. The proof relies heavily on the existence of bounds for the solution of the Riccati-like observer equation. The bounds on the solution of this Riccati-like equation are obtained by small perturbation of a Riccati equation. For the sake of clarity, the proof of this technical result is relegated to Appendix A. It basically follows the ideas developed in [14] (to prove that the solution of the disturbed Riccati equation is bounded, we followed the ideas developed in [14] which require several lemmas. In our article, only the significantly different proofs are detailed). A simulation has been done to illustrate the UKO's performance against the UKF and EKF in Section 5.2.

## 2. Continuous-Time Unscented Kalman Filter

In this section, we present the unscented Kalman filter in matrix form (see [42], it having been first introduced by Julier and Uhlmann in [20]). It uses the unscented transformation which propagates  $\sigma$ -points, (denoted by  $X$ ) through the non-linear equation and uses the results to capture the means and covariance of the posterior distribution [23, 48]. In the following, we will recall the unscented transformation algorithm, tuning parameters, and weights.

### 2.1. The unscented transformation

We now introduce some notation that will be convenient for the calculations and proofs.

**Notation 1.** Let  $1_n$  and  $0^n$  denote, respectively, the row vector  $1_n = (1, 1, \dots, 1) \in \mathbb{R}^n$  and the column vector  $0^n = (0, 0, \dots, 0)^t \in \mathbb{R}^n$ . Let  $u$  denote the row vector  $u = [1, 1_n, 1_n] \in \mathbb{R}^{2n+1}$  and  $U$  the  $n \times (2n+1)$  matrix  $U = [0_n, Id_n, -Id_n]$  (The brackets  $[ ]$  denote matrix concatenation).

The procedure to implement the UKF in continuous-time is summarized as follows [42]

1. Choose  $2n + 1$  points in  $\mathbb{R}^n$  called  $\sigma$ -points:

$$X(t) = m(t)u + \sqrt{c} \sqrt{P(t)}U \quad (4)$$

- $X(t)$  is the matrix of  $\sigma$ -points (written in columns),  $X(t) \in \mathcal{M}_{n,2n+1}(\mathbb{R})$ , the set of matrices of dimension  $n \times (2n + 1)^1$ ,
- $k \geq 0$  and  $c \in ]0, n + k]$  are parameters to be set,

In the proofs, the parameter  $c$  will play a crucial role and will be used as a tuning parameter.  $\sqrt{P(t)}$  denotes the principal square root of  $P(t)$ , that is, the symmetric definite positive matrix such that  $P(t) = \sqrt{P(t)}^2$ . We remark that  $\sqrt{P(t)}$  is not the Cholesky decomposition of  $P(t)$ , as is often the case in the UKF literature (see, e.g. [42]). This does not change any basic results (since  $P(t) = \sqrt{P(t)}\sqrt{P(t)}^t$  as usual).

**Remark 1.** In [34], several alternatives for the square root in the UKF are considered, without any clear empirical or theoretical reasons for making a choice between the Cholesky or the principal square root (or anything else). In [46], a divergence problem is noted when using the Cholesky decomposition, and the principal square root was chosen, which performed well. In the present paper, we reach the same conclusion from purely theoretical considerations.

---

<sup>1</sup>More generally,  $\mathcal{M}_{q,p}(\mathbb{R})$  is the set of real-valued matrices of dimension  $q \times p$ .

2. Calculate the weights  $W_m = (W_m^{(0)}, W_m^{(1)}, \dots, W_m^{(2n)})^t \in \mathcal{M}_{2n+1,1}(\mathbb{R})$  associated with the  $\sigma$ -points:

$$\begin{aligned} W_m^{(0)} &= \frac{\lambda}{n + \lambda}; \\ W_m^{(i)} &= \frac{1}{2(n + \lambda)}, \quad i = 1, \dots, 2n \end{aligned} \quad (5)$$

and  $W_c = (W_c^{(0)}, W_c^{(1)}, \dots, W_c^{(2n)})^t \in \mathcal{M}_{2n+1,1}(\mathbb{R})$ :

$$\begin{aligned} W_c^{(0)} &= \frac{\lambda}{n + \lambda} + 1 + \beta - \frac{c}{n + k} \\ W_c^{(i)} &= \frac{1}{2(n + \lambda)}, \quad i = 1, \dots, 2n \end{aligned} \quad (6)$$

where  $\lambda$  is a scalar parameter defined by  $\lambda = c - n$ . Some authors set  $k = 0$ .  $\beta$  is a non negative weighting term which can be used to incorporate knowledge of the higher order moments of the distribution. For a Gaussian distribution the optimal choice is  $\beta = 2$  [19].

3. Propagate each  $\sigma$ -points via the non-linear transformation  $g$

$$Y = g(X) \quad (7)$$

where  $X \in \mathcal{M}_{n,2n+1}(\mathbb{R})$ ,  $Y \in \mathcal{M}_{p,2n+1}(\mathbb{R})$ , and  $g: \mathbb{R}^n \rightarrow \mathbb{R}^p$ , (7) means that the  $i^{\text{th}}$  column  $Y_i$  of the matrix  $Y$  is calculated from the  $i^{\text{th}}$  column of  $X$ , as follows:

$$Y_i = g(X_i),$$

i.e.  $g(X) = [g(X_0) \cdots g(X_{2n})]$ .

4. The mean and the covariance of  $g(X)$  are given by

$$E[g(X)] \approx \mu_u \stackrel{\text{def.}}{=} g(X)W_m = \sum_{i=0}^{2n} W_m^{(i)} g(X_i) \quad (8)$$

$$\text{Cov}(g(X)) \approx S_u \stackrel{\text{def.}}{=} \sum_{i=0}^{2n} W_c^{(i)} (g(X_i) - m)(g(X_i) - m)^t. \quad (9)$$

5. The cross-covariance of  $X$  and  $g(X)$  can estimated as

$$C_u \stackrel{\text{def.}}{=} \sum_{i=0}^{2n} W_c^{(i)} (X_i - m)(g(X_i) - \mu_u)^t. \quad (10)$$

Finally, we define the matrix  $W \in \mathcal{M}_{2n+1,2n+1}(\mathbb{R})$  by

$$W = (I - [W_m \ \cdots \ W_m]) \times \text{diag} \left( W_c^{(0)} \ \cdots \ W_c^{(2n)} \right) \times (I - [W_m \ \cdots \ W_m])^t. \quad (11)$$

## 2.2. The unscented Kalman filter algorithm

The equations for the UKF in the continuous case for the system (1) are derived from the unscented transformation and given in [42, 43]:

$$\begin{cases} K(t) = X(t)W h(X(t), t)^t R^{-1}(t) \\ \frac{dm(t)}{dt} = f(X(t), t)W_m + K(t)(y(t) - h(X(t), t)W_m) \\ \frac{dP(t)}{dt} = X(t)W f(X(t), t)^t + f(X(t), t)W X(t)^t + Q(t) - K(t)R(t)K(t)^t. \end{cases} \quad (12)$$

In this algorithm,  $Q(t)$  and  $R(t)$  are the covariance matrices of the state measurements and the noise, respectively. They are symmetric and positive definite. In the deterministic case, these two matrices will be considered as parameters to be set.

### 2.3. The non-convergence of UKF as an observer

The unscented Kalman filter is not a good candidate for an exponential observer. Indeed, assume that the function  $h$  (in (1)) is linear, and suppose that  $x(t) = 0$  is an equilibrium state. If  $m(t) = 0$ , one expects that the observer remains at the equilibrium state 0 (since the innovation term  $y(t) - h(X(t), t)W_m$  is equal to zero). The state equation in (12) becomes

$$\frac{dm(t)}{dt} = f(X(t), t)W_m = \frac{1}{2(n + \lambda)} \sum_{i=1}^n f(\sqrt{c}\sqrt{P(t)}_i, t) + f(-\sqrt{c}\sqrt{P(t)}_i, t). \quad (13)$$

The right-hand side of (13) is equal to zero if  $f$  is linear (or simply odd): if it were different from zero, 0 would not be an equilibrium state of (13).

For instance, and to be more concrete, set  $f(x, t) = -x(1 + (2x - 1)^2)$  and  $h(x, t) = x$  with  $x(0) = 0$ . This example is constructed so that 0 is a globally asymptotically stable point of the system, in which  $x \rightarrow f(x, t)$  is not odd function (obviously non-linear), as mentioned above. Indeed, if  $m(t) = 0$ , i.e. if the observer perfectly estimated the state, then (13) would become

$$\frac{dm(t)}{dt} = f(X(t), t)W_m = \frac{1}{2(n + \lambda)} 4c\sqrt{P(t)}^2 = 2P(t) \quad (14)$$

where  $P(t) > 0$ . Thus, this result contradicts the hypothesis that  $m(t)$  converges asymptotically to the solution.

The result of a numerical simulation of the example above is shown in Figure 1. A logarithmic scale was used for the time axis to deal with the exponential convergence rate.

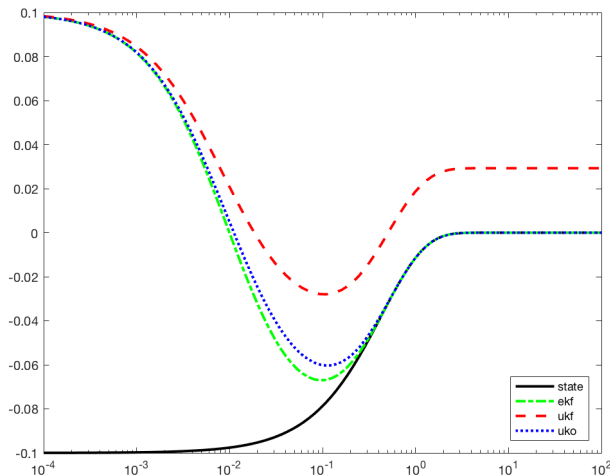


Figure 1: The biased estimation of UKF

Figure 1 shows that the UKF (dashed red curve) does not converge (unlike the EKF, the dash-dotted green curve), and that there remains an estimation bias, consistent with Eq. (14). Figure 1 also shows the performance of the UKO (dotted blue curve), which will be explained below.

## 3. Unscented Kalman Observer

### 3.1. Equations for the UKO

The UKF does not converge as an observer because of the differential equation for  $m(t)$  in (12). To resolve this, we introduce a variant using both the approximation of the first order (2) and the unscented approximation (3):

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim \mathcal{N} \left( \begin{pmatrix} m \\ g(m) \end{pmatrix}, \begin{pmatrix} P & C_U \\ C_U^t & S_U \end{pmatrix} \right). \quad (15)$$

Like the two previous ones, this transformation makes it possible to establish the equations of an observer for system (1), which we will call the unscented Kalman observer (UKO) (for which a simulation has been shown in Figure 1):

$$\begin{cases} K(t) = X(t)Wh^t(X(t),t)R^{-1}(t) \\ \frac{dm(t)}{dt} = f(m(t),t) + K(t)(y(t) - h(m(t),t)) \\ \frac{dP(t)}{dt} = X(t)Wf^t(X(t),t) + f(X(t),t)WX^t(t) + Q(t) - K(t)R(t)K^t(t). \end{cases} \quad (16)$$

### 3.2. Perturbed Riccati equation

In this section, we will prove that, unlike the UKF used as an observer, the UKO is a locally exponentially converging observer.

We first give a sketch of the proof. First of all, we will develop each term of the Riccati equation with respect to  $c$ . Recall that  $c$  is a scaling parameter representing the dispersion of  $\sigma$ -points. We will see that the limit when  $c$  goes to 0 of the matrix equation in (16) is the classical Kalman's Riccati equation. We will then prove, using a perturbation argument, that  $P(t)$  remains bounded if  $c$  is small enough (Appendix A).

**Remark 2.** *The fact that  $c$  should be small in our result may be considered as a strong hypothesis regarding the meaning of  $c$  in the UKF. However, it should be recalled that*

- *if  $c$  is large and non-linearities are too important, the quality of the unscented transformation may be very poor;*
- *we want to establish the exponential convergence of an observer, so we need to use an observability hypothesis to ensure convergence without noise. If  $c$  is not small enough, the solution of the Riccati-like equation may even be undefined.*

We will then conclude, using the classical Lyapunov approach, that the UKO converges exponentially.

Before calculating  $f(X(t),t)WX^t(t)$  (and  $h(X(t),t)WX^t(t)$ , which is similar), we remark that  $uW_m = 1$ ,  $UW_m = 0_n$  and  $UU^t = 2Id_n$  (see Notation 1 in section 2.1). We omit the time variable  $t$  in the equations below for lighten the notations in calculations.

$$\begin{aligned} WX^t &= (Id_{2n+1} - W_m u) \text{diag}(W_c)(Id_{2n+1} - W_m u)^t(mu + \sqrt{c}\sqrt{P}U)^t \\ &= (Id_{2n+1} - W_m u) \text{diag}(W_c)(mu + \sqrt{c}\sqrt{P}U - \underbrace{m u W_m u}_{=1} - \sqrt{c}\sqrt{P}\underbrace{UW_m u}_{=0_n})^t \\ &= (Id_{2n+1} - W_m u) \text{diag}(W_c)U^t \sqrt{c}\sqrt{P} \\ &= W_c^{(1)}(Id_{2n+1} - W_m u)U^t \sqrt{c}\sqrt{P} = \frac{1}{2c}U^t \sqrt{c}\sqrt{P} \end{aligned}$$

since

$$\text{diag}(W_c)U^t = \begin{pmatrix} W_c^0 & & \\ & \ddots & \\ & & W_c^{2n+1} \end{pmatrix} \begin{pmatrix} 0_n^t \\ Id_{2n+1} \\ -Id_{2n+1} \end{pmatrix} = W_c^{(1)}U^t.$$

We will write, temporarily,  $B = \sqrt{P}$  to lighten the notation:

$$\begin{aligned} f(X,t)WX^t &= \frac{1}{2\sqrt{c}}f(mu + \sqrt{c}BU, t)U^t B \\ &= \frac{1}{2\sqrt{c}} [f(m, t), f(m1_n + \sqrt{c}B, t), f(m1_n - \sqrt{c}B, t)] U^t B \\ &= \frac{1}{2\sqrt{c}} (f(m1_n + \sqrt{c}B, t) - f(m1_n - \sqrt{c}B, t)) B. \end{aligned}$$

We now remark that setting  $f = Id_n$ , we see that  $XWX^t = P$  and hence the UKO is, in the linear case, just a Kalman filter. Using Taylor's formula, develop each coefficient of the  $n \times (2n + 1)$  matrix up to second order  $f(m1_n + \sqrt{c}B, t) - f(m1_n - \sqrt{c}B, t)$

$$\begin{aligned} (f(m1_n + \sqrt{c}B, t) - f(m1_n - \sqrt{c}B, t))_{i,j} &= 2\sqrt{c} \sum_{k=1}^n \frac{\partial f_i}{\partial x_k}(m, t) B_{k,j} \\ &+ c \sum_{k,l=1}^n \frac{\partial^2 f_i(m + \tau_{i,j}\sqrt{c}B_{\cdot,j}, t)}{\partial x_k \partial x_l} B_{k,j} B_{l,j} \end{aligned} \quad (17)$$

where  $-1 \leq \tau_{i,j} \leq 1$  and  $B_{\cdot,j}$  denotes the  $j^{\text{th}}$  column of  $B$ . The approximation of  $f$  in matrix form is then

$$f(X, t)WX^t = \frac{1}{2\sqrt{c}} (f(m1_n + \sqrt{c}B, t) - f(m1_n - \sqrt{c}B, t)) B = AP + \frac{\sqrt{c}}{2} H_f \sqrt{P} \quad (18)$$

and similarly for  $h$ :

$$h(X, t)WX^t = \frac{1}{2\sqrt{c}} (h(m1_n + \sqrt{c}B, t) - h(m1_n - \sqrt{c}B, t)) B = CP + \frac{\sqrt{c}}{2} H_h \sqrt{P} \quad (19)$$

where  $A = A(t)$  and  $C = C(t)$  are time-varying matrices (the Jacobian of  $f$  and  $h$  respectively)

$$A(t) = \frac{\partial f}{\partial x}(m, t) \text{ and } C(t) = \frac{\partial h}{\partial x}(m, t), \quad (20)$$

$H_f \in \mathcal{M}_{n,n}(\mathbb{R})$  (and  $H_h \in \mathcal{M}_{n,n}(\mathbb{R})$  which is similar) being defined as

$$(H_f)_{i,j} = \sum_{k,l=1}^n \frac{\partial^2 f_i(m + \tau_{i,j}\sqrt{c}B_{\cdot,j}, t)}{\partial x_k \partial x_l} B_{k,j} B_{l,j} = B_{j,\cdot} \partial^2 f_i(m + \tau_{i,j}\sqrt{c}B_{\cdot,j}, t) B_{\cdot,j}. \quad (21)$$

We assumed that each coefficient of the Hessian matrices<sup>2</sup> are bounded from above by  $M_f > 0$  (respectively,  $M_h > 0$ ), that is to say,

$$M_f = \max_{1 \leq i,k,l \leq n} \sup_{x \in \mathbb{R}^n, t \geq 0} \left\| \frac{\partial^2 f_i(x, t)}{\partial x_k \partial x_l} \right\| \quad (22)$$

and the same with  $M_h$ :

$$M_h = \max_{1 \leq i,k,l \leq n} \sup_{x \in \mathbb{R}^n, t \geq 0} \left\| \frac{\partial^2 h_i(x, t)}{\partial x_k \partial x_l} \right\|. \quad (23)$$

Using the expressions in (18) and (19), the system (16) becomes

$$\left\{ \begin{array}{l} K(t) = P(t)C(t)^t R^{-1}(t) + \frac{\sqrt{c}}{2} \sqrt{P(t)} H_h(t)^t R^{-1}(t) \\ \frac{dm(t)}{dt} = f(m(t), t) + K(t)(y(t) - h(m(t), t)) \\ \frac{dP(t)}{dt} = P(t)A(t)^t + A(t)P(t) + Q(t) \\ \quad + \frac{\sqrt{c}}{2} (\sqrt{P(t)} H_f(t)^t + H_f(t) \sqrt{P(t)}) - P(t)C(t)^t R^{-1}(t) C(t) P(t) \\ \quad - \frac{\sqrt{c}}{2} (P(t)C(t)^t R^{-1}(t) H_h(t) \sqrt{P(t)} + \sqrt{P(t)} H_h(t)^t R^{-1}(t) C(t) P(t)) \\ \quad - \frac{c}{4} \sqrt{P(t)} H_h(t)^t R^{-1}(t) H_h(t) \sqrt{P(t)}. \end{array} \right. \quad (24)$$

<sup>2</sup> $H_f$  and  $H_h$  are not strictly speaking Hessian matrices, and can not be written as the product of Hessian matrices with  $B$ , since the coefficients are second derivatives of  $f$  (respectively,  $h$ ) but are not evaluated at the same points.



Rewriting the system (24) with  $S(t) = P(t)^{-1}$ , we obtain

$$\left\{ \begin{array}{l} K(t) = S^{-1}(t)C(t)^t R^{-1}(t) + \frac{\sqrt{c}}{2}\sqrt{P(t)}H_h(t)^t R^{-1}(t) \\ \frac{dm(t)}{dt} = f(m(t), t) + K(t)(y(t) - h(m(t), t)) \\ \frac{dS(t)}{dt} = -A(t)^t S(t) - S(t)A(t) - S(t)Q(t)S(t) \\ \quad - \frac{\sqrt{c}}{2}(S(t)\sqrt{P(t)}H_f(t)^t S(t) + S(t)H_f(t)\sqrt{P(t)}^t S(t)) + C(t)^t R^{-1}(t)C(t) \\ \quad + \frac{\sqrt{c}}{2}(C(t)^t R^{-1}(t)H_h(t)\sqrt{P(t)}^t S(t) + S(t)\sqrt{P(t)}H_h(t)^t R^{-1}(t)C(t)) \\ \quad + \frac{c}{4}S(t)\sqrt{P(t)}H_h(t)^t R^{-1}(t)H_h(t)\sqrt{P(t)}^t S(t). \end{array} \right. \quad (25)$$

#### 4. Convergence analysis

First, we will consider the matrix equation in (25). We will prove the following result, which is a crucial result for the proof of the convergence of the UKO. This result applies to the matrix equation of the UKO, which is the same matrix equation as for the UKF. Therefore, this result may be applied to establish a priori bounds on  $S(t)$  (or, equivalently, on  $P(t)$ ), as required in Theorem 3.1 in [50].

**Definition 2.** (See, for instance, [5]). We define the observability Gramian as

$$W_T(t) = \int_{t-T}^t \Psi^t(\tau, t) C^t(\tau) R^{-1}(\tau) C(\tau) \Psi(\tau, t) d\tau$$

for  $T > 0$  and  $t \geq T$ , where  $\Psi(\tau, t)$  is the state-transition matrix associated to the linear system defined by  $A(t)$ , that is,

$$\left\{ \begin{array}{l} \frac{d}{dt} \Psi(\tau, t) = A(\tau) \Psi(\tau, t) \\ \Psi(t, t) = Id_n. \end{array} \right. \quad (26)$$

If  $W_T(t)$  is symmetric positive definite and uniformly bounded for any  $t \geq 0$ , then the system characterized by the pair  $(A(t), C(t))$  and written as

$$\left\{ \begin{array}{l} \frac{dx(t)}{dt} = A(t)x(t) \\ y(t) = C(t)x(t) \end{array} \right. \quad (27)$$

is said to be completely uniformly observable.

Before proving that the proposed observer is an exponential observer, we need to expand the functions  $f$  and  $h$  of the system (1) up to first order, and make the following assumptions (see (20) for notation):

$$\begin{aligned} f(x(t), t) - f(m(t), t) &= A(t)(x(t) - m(t)) + \psi(x(t), m(t)), \\ h(x(t), t) - h(m(t), t) &= C(t)(x(t) - m(t)) + \phi(x(t), m(t)) \end{aligned} \quad (28)$$

where  $A(t)$  and  $C(t)$  have been defined in (20), and  $\psi(x(t), m(t))$ , and  $\phi(x(t), m(t))$  are the remaining non-linear terms.

**Assumption 1.** There are positive real numbers  $\kappa_\psi$ ,  $\kappa_\phi$ ,  $\epsilon_\psi$ ,  $\epsilon_\phi > 0$  such that the non-linearities are bounded by

$$\begin{aligned} \|\psi(x_1(t), x_2(t))\| &\leq \kappa_\psi \|x_1(t) - x_2(t)\|^2, \\ \|\phi(x_1(t), x_2(t))\| &\leq \kappa_\phi \|x_1(t) - x_2(t)\|^2. \end{aligned}$$

for  $x_1(t), x_2(t) \in \mathbb{R}^n$  such that  $\|x_1(t) - x_2(t)\| \leq \epsilon_\psi$  and  $\|x_1(t) - x_2(t)\| \leq \epsilon_\phi$ , respectively.

**Assumption 2.** There are positive real numbers  $\bar{c}$ ,  $r_{min}$  and  $q_{min} > 0$  such that  $Q(t)$ ,  $R(t)$  and  $C(t)$  are time-varying matrices bounded by

$$\begin{aligned} \|C(t)\| &\leq \bar{c} \\ \sqrt{n}r_{min} &\leq \|R(t)\| \\ q_{min}Id_n &\leq Q(t). \end{aligned}$$

with  $\bar{c}$  is the lower bound of the matrix  $C(t)$ ,  $r_{min}$  and  $q_{min}$  are the smallest eigenvalue of  $Q(t)$  and  $R(t)$  respectively.

**Theorem 3.** We now consider the UKO in continuous time, and the following Riccati-like equation,

$$\begin{cases} \frac{dS(t)}{dt} = -A(t)^t S(t) - S(t)A(t) - S(t)(Q(t) + \Omega(S(t)))S(t) + C(t)^t R^{-1}(t)C(t) \\ S(0) = S_0, \end{cases} \quad (29)$$

where

$$\begin{aligned} \Omega(S(t)) = & -\frac{\sqrt{c}}{2}(\sqrt{P(t)}H_f(t)^t + H_f(t)\sqrt{P(t)}^t) + \frac{\sqrt{c}}{2} \left( P(t)C(t)^t R^{-1}(t)H_h(t)\sqrt{P(t)}^t + \sqrt{P(t)}H_h(t)^t R^{-1}(t)C(t)P(t) \right) \\ & + \frac{c}{4}\sqrt{P(t)}H_h(t)^t R^{-1}(t)H_h(t)\sqrt{P(t)}^t \end{aligned} \quad (30)$$

and where  $(A(t), C(t))$  is a time-dependent observable pair (Definition 2). We assume that all the elements of the matrices  $A(t)$  and  $C(t)$  are uniformly bounded,  $Q(t)$  and  $R(t)$  are symmetric positive definite matrices and  $S(0)$  is a symmetric positive definite matrix taken in the compact set  $\mathcal{S} = \{S \in S_n^+, aId_n \leq S(0) \leq bId_n\}$ , with  $0 < a < b$ .

Assuming (22) and (23) hold for  $f$  and  $h$ , under Assumptions 1 and 2, there exist  $c_0$  and  $0 < \alpha < \beta$  such that for any  $c \leq c_0$ ,

$$\alpha Id_n \leq S(t) \leq \beta Id_n.$$

$\alpha$  and  $\beta$  are called the lower and upper bounds of  $S(t)$  respectively.

*Proof.* The proof is detailed in the Appendix. □

Let us define  $\varepsilon(t) = x(t) - m(t)$ , then

$$\dot{\varepsilon}(t) = \dot{x}(t) - \dot{m}(t) = f(x(t), t) - f(m(t), t) - K(t)(h(x(t), t) - h(m(t), t)). \quad (31)$$

**Definition 3.** The equilibrium point  $\varepsilon(t) = 0$  is an exponentially stable equilibrium point of Eq. (31) if there exist constants  $m$ ,  $\alpha > 0$  and  $\epsilon$  such that

$$\|\varepsilon(t)\| \leq \eta \|\varepsilon(0)\| e^{-\alpha t} \quad (32)$$

holds for every  $t \geq 0$  and for all  $\|\varepsilon(0)\| \leq \epsilon$ .

**Theorem 4.** Consider the non-linear system (1) and the observer (UKO) given by system (25) and suppose that all the assumptions of Theorem 3 are fulfilled, so that  $\alpha Id_n \leq S(t) \leq \beta Id_n$ .

Then, for all  $t \geq 0$ , there are positive constants  $\bar{k}, L, \delta$  and  $\eta$ , hence the following inequality is satisfied:

$$\|\varepsilon(t)\|^2 \leq \frac{\beta}{\alpha} \|\varepsilon(0)\|^2 e^{\left(-\alpha q_{min} + \frac{(2\bar{k} + \sqrt{c}\beta L)\epsilon_0}{\alpha} + \frac{\sqrt{c}}{\alpha}(\beta\delta + \sqrt{c}\beta\eta)\right)t} \quad (33)$$

**Remark 5.** As can be seen from (33),  $c$  is defined as in Section 2.1 and it must be small enough to ensure the negativity of  $-\alpha q_{min} + \frac{(2\bar{k} + \sqrt{c}\beta L)\epsilon_0}{\alpha} + \frac{\sqrt{c}}{\alpha}(\beta\delta + \sqrt{c}\beta\eta)$ . Exact expressions for  $\bar{k}, L, \delta$  and  $\eta$  appear at the end of the proof.

*Proof.* The proof of this theorem relies on an analysis of the dynamics (31) of the estimation error  $\varepsilon(t)$ .

Using the development (25) :

$$\begin{aligned} \dot{\varepsilon}(t) = & [A(t) - K(t)C(t)]\varepsilon(t) + \psi(x(t), m(t)) - K(t)\phi(x(t), m(t)) \\ = & A(t)\varepsilon(t) - P(t)C(t)^t R(t)^{-1}C(t)\varepsilon(t) - P(t)C(t)^t R(t)^{-1}\phi(x(t), m(t)) + \psi(x(t), m(t)) \\ & - \frac{\sqrt{c}}{2}\sqrt{P(t)}H_h(t)^t R(t)^{-1}C(t)\varepsilon(t) - \frac{\sqrt{c}}{2}\sqrt{P(t)}H_h(t)^t R(t)^{-1}\phi(x(t), m(t)). \end{aligned} \quad (34)$$

We choose the classical quadratic Lyapunov function

$$V(\varepsilon(t)) = \varepsilon(t)^t S(t)\varepsilon(t), \quad (35)$$

where  $S(t)$  is the solution of the Riccati equation (25), and we will compute its time derivative

$$\dot{V}(\varepsilon(t)) = 2\dot{\varepsilon}(t)^t S(t)\varepsilon(t) + \varepsilon^t \dot{S}(t)\varepsilon(t). \quad (36)$$

In the following, to ease the notation and as long as there is no ambiguity, we omit again the time variable  $t$  in the equations. Inserting  $\dot{\varepsilon}$ ,  $\dot{S}$  in (36), we obtain

$$\begin{aligned} \dot{V}(\varepsilon) = & 2[A\varepsilon - PC^t R^{-1}C\varepsilon - \frac{\sqrt{c}}{2}\sqrt{P}H_h^t R^{-1}C\varepsilon + \psi(x, m) - PC^t R^{-1}\phi(x, m) - \frac{\sqrt{c}}{2}\sqrt{P}H_h^t R^{-1}\phi(x, m)]^t S\varepsilon \\ & + \varepsilon^t [-A^t S - SA - SQS - \frac{\sqrt{c}}{2}(S\sqrt{P}H_f^t S + SH_f\sqrt{P}^t S) + C^t R^{-1}C \\ & + \frac{\sqrt{c}}{2}(C^t R^{-1}H_h\sqrt{P}^t S + S\sqrt{P}H_h^t R^{-1}C) + \frac{c}{4}S\sqrt{P}H_h^t R^{-1}H_h\sqrt{P}^t S]\varepsilon. \end{aligned} \quad (37)$$

With a few rearrangements,  $\dot{V}(\varepsilon)$  becomes

$$\begin{aligned}\dot{V}(\varepsilon) = & -\varepsilon^t S Q S \varepsilon - \varepsilon^t C^t R^{-1} C \varepsilon + 2\varepsilon^t S \psi(x, m) - 2\varepsilon^t C^t R^{-1} \phi(x, m) \\ & - \sqrt{c} \varepsilon^t S [\sqrt{P} H_h^t R^{-1} \phi(x, m) + \sqrt{P} H_f^t S \varepsilon - \frac{\sqrt{c}}{4} \sqrt{P} H_h^t R^{-1} H_h \sqrt{P}^t S \varepsilon].\end{aligned}\quad (38)$$

Using Theorem 3 and Assumptions 1 and 2 we have

$$\dot{V}(\varepsilon) \leq -\alpha q_{\min} \varepsilon^t S \varepsilon + \left( (2\bar{k} + \sqrt{c}\beta L) \|\varepsilon\| + (\sqrt{c}\beta\delta + c\beta\eta) \right) \|\varepsilon\|^2 \quad (39)$$

with  $\bar{k} = (\beta\kappa_\psi + \frac{\bar{c}\sqrt{n}\kappa_\phi}{r_{\min}})$ ,  $L = \frac{n^2 M_h \kappa_\phi n^{3/4}}{r_{\min} \alpha^{3/2}}$ ,  $\delta = \frac{n^{3/2} M_f \beta n^{3/4}}{\alpha^{3/2}}$  and  $\eta = \frac{n^5 M_h^2 \beta}{4\alpha^3 r_{\min}}$ .

For  $\|\varepsilon\| \leq \varepsilon_0$  with  $\varepsilon_0 \leq \min(\varepsilon_\psi, \varepsilon_\phi)$ ,

$$\dot{V}(\varepsilon) \leq -\alpha q_{\min} \varepsilon^t S \varepsilon + \left( (2\bar{k} + \sqrt{c}\beta L) \varepsilon_0 + (\sqrt{c}\beta\delta + c\beta\eta) \right) \|\varepsilon\|^2. \quad (40)$$

From (40) and  $\|\varepsilon\|^2 \leq \frac{1}{\alpha} \varepsilon^t S \varepsilon$  we obtain

$$\dot{V}(\varepsilon) \leq \left( -\alpha q_{\min} + \frac{(2\bar{k} + \sqrt{c}\beta L) \varepsilon_0}{\alpha} + \frac{\sqrt{c}}{\alpha} (\beta\delta + \sqrt{c}\beta\eta) \right) \varepsilon^t S \varepsilon \quad (41)$$

and hence

$$V(\varepsilon(t)) \leq V(\varepsilon(0)) e^{-\left( \alpha q_{\min} + \frac{(2\bar{k} + \sqrt{c}\beta L) \varepsilon_0}{\alpha} + \frac{\sqrt{c}}{\alpha} (\beta\delta + \sqrt{c}\beta\eta) \right) t}. \quad (42)$$

Since the Lyapunov function is bounded by

$$\alpha \|\varepsilon(t)\|^2 \leq V(\varepsilon(t)) \leq \beta \|\varepsilon(t)\|^2 \quad (43)$$

then

$$\|\varepsilon(t)\| \leq \sqrt{\frac{\beta}{\alpha}} \|\varepsilon(0)\| e^{\frac{1}{2} \left( -\alpha q_{\min} + \frac{(2\bar{k} + \sqrt{c}\beta L) \varepsilon_0}{\alpha} + \frac{\sqrt{c}}{\alpha} (\beta\delta + \sqrt{c}\beta\eta) \right) t} \quad (44)$$

Let  $\varepsilon_0$  and  $c$  be small enough such that

$$-\alpha q_{\min} + \frac{(2\bar{k} + \sqrt{c}\beta L) \varepsilon_0}{\alpha} + \frac{\sqrt{c}}{\alpha} (\beta\delta + \sqrt{c}\beta\eta) < 0$$

and let  $\|\varepsilon(0)\| \leq \sqrt{\frac{\beta}{\alpha}} \varepsilon_0$  such that

$$\|\varepsilon(t)\| \leq \varepsilon_0 e^{\frac{1}{2} \left( -\alpha q_{\min} + \frac{(2\bar{k} + \sqrt{c}\beta L) \varepsilon_0}{\alpha} + \frac{\sqrt{c}}{\alpha} (\beta\delta + \sqrt{c}\beta\eta) \right) t}. \quad (45)$$

This ensure that  $\|\varepsilon(t)\| \leq \varepsilon_0$ ,  $\forall t \geq 0$  and that the inequality (45) holds, which proves Theorem 4.  $\square$

## 5. Applications and simulations

### 5.1. Distillation column : the binary distillation process

To illustrate the theorem, let us consider the classical CMO (constant molar overflow) model of a binary distillation column. In the following, we will only consider a three plates model, which is very simple and easy to describe (see [11] for a more general presentation). The difference between this low-dimension model with the CMO model of a depropanizer or debutanizer is only quantitative. In particular, nonlinearities are exactly the same. The CMO model is only a material balance of the lightest component on each plate of the column:

$$\begin{cases} H_1 \frac{dx_1}{dt} = V(t)(k(x_2) - x_1) \\ H_2 \frac{dx_2}{dt} = F(Z_F - x_2) + L(t)(x_1 - x_2) + V(t)(k(x_3) - k(x_2)) \\ H_3 \frac{dx_3}{dt} = (L(t) + F)(x_2 - x_3) + V(t)(x_3 - k(x_3)) \end{cases} \quad (46)$$

where  $H_1, H_2, H_3$  (liquid holdups),  $F$  (feed flow) and  $Z_F$  (feed composition) are positive constants,  $k(x) = \frac{\alpha x}{1 + (\alpha - 1)x}$  with  $\alpha > 1$ , and  $L$  (liquid flow) and  $V$  (vapor flow) are smoothly varying positive functions of time (control variables). The state is  $(x_1, x_2, x_3)$  (composition in liquid phase) and the output is  $x_3$  (bottom composition). This system has been extensively used to illustrates high-gain observer approach because it is already in a canonical form of observability, see [11, 14]).

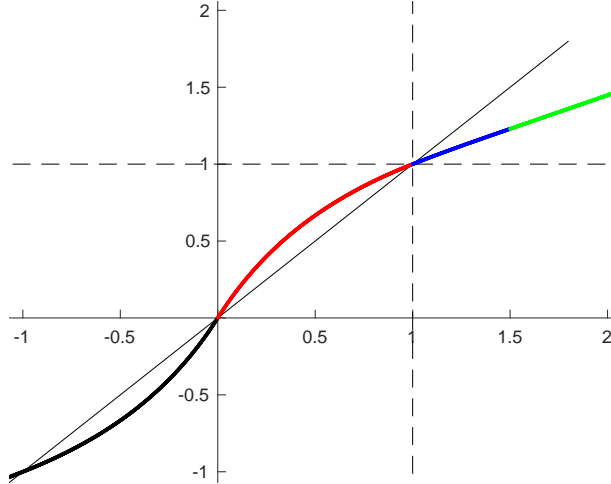


Figure 2:  $k$  (in red, between 0 and 1) and its odd (black)  $C^2$ -continuation, spline prolongation between 1 and  $1 + \varepsilon$  (blue), linear beyond (green).

The state of this system belongs to  $[0, 1]^3$  but the state of the filter may leave the state space so one should extend  $k$  such that it remains twice differentiable and Lipschitz. It is an easy task, see for instance Figure 2 where the function have been extended between 1 and  $1 + \varepsilon$  using a polynomial function of degree 3 (a spline), and extended linearly beyond  $1 + \varepsilon$  such that  $k$  remains Lipschitz. For negative values,  $k$  is extended so as to be an odd function. Therefore,  $k(x)$  is clearly Lipschitz (linear outside a compact interval). We have

$$A(t) = \begin{pmatrix} -\frac{V(t)}{H_1} & \frac{V(t)}{H_1} \frac{dk}{dx}(m_1(t)) & 0 \\ \frac{L(t)}{H_2} & \frac{1}{H_2}(-F - L(t) - V(t) \frac{dk}{dx}(m_2(t))) & \frac{V(t)}{H_2} \frac{dk}{dx}(m_3(t)) \\ 0 & \frac{L(t)+F}{H_3} & \frac{1}{H_3}(-L(t) - F + V(t)(1 - \frac{dk}{dx}(m_3(t)))) \end{pmatrix}$$

and  $C(t) = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix}$ . Therefore, matrix are bounded and second derivatives also (since  $k$  is Lipschitz twice differentiable on  $\mathbb{R}$ ). Moreover, the system is clearly uniformly observable since the following matrix is triangular and its diagonal is bounded from below by a positive constant:

$$\begin{pmatrix} C(t) \\ C(t)A(t) \\ C(t)A(t)^2 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & \frac{L(t)+F}{H_3} & \bullet \\ \frac{(L(t)+F)L(t)}{H_2H_3} & \bullet & \bullet \end{pmatrix}$$

(The coefficients below the diagonal do not have to be calculated) The non-linearities of the system being only on  $k$ , it is clear that Assumption 1 is satisfied, as well as (22) and (23), and since  $C$  is constant, Assumption 2 is also verified.

## 5.2. Numerical Simulations

Now, we simulate the model of the binary distillation column (46). For that we choose  $H_1 = 40$ ,  $H_2 = 10$ ,  $H_3 = 80$ ,  $F = 10$ ,  $Z_F = 0.4$ ,  $L = 13$ ,  $V = 17$  and  $\alpha$ , the parameter of the function  $k(x)$ , is set to 2. The initial conditions are  $x_0 = [0.5, 0.5, 0.5]^t$  for the system and  $m_0 = [1, 0.6, 0.3]^t$  for the observer. The initial condition of the Riccati differential equation (16) is set to  $P(0) = Id_3$ , while the positive definite matrix  $Q$  is set to  $Q = Id_3$  and  $R = 1$ . We have set  $\beta = 0$ ,  $k = 0$  and  $c = 0.03$  in UKO (see (4), (5) and (6)).

Figure 3 shows the evolution of the state and estimates of the system as a function of time for both algorithms UKO and EKF. We did not add noise, our objective being to compare the convergence rates.

Observers UKO and EKF converge locally and perform similarly. They only differ in the calculation made: EKF requires the calculation of the Jacobian of the non-linear functions  $f(\cdot)$  and  $h(\cdot)$ , rather than the propagation of the  $\sigma$ -points, as does the UKO.

## 6. Conclusion

In this paper, we have shown that the unscented Kalman filter is not an exponentially convergent observer, and we have proposed a slight modification of it which allows building an exponentially convergent observer, called the unscented Kalman observer.

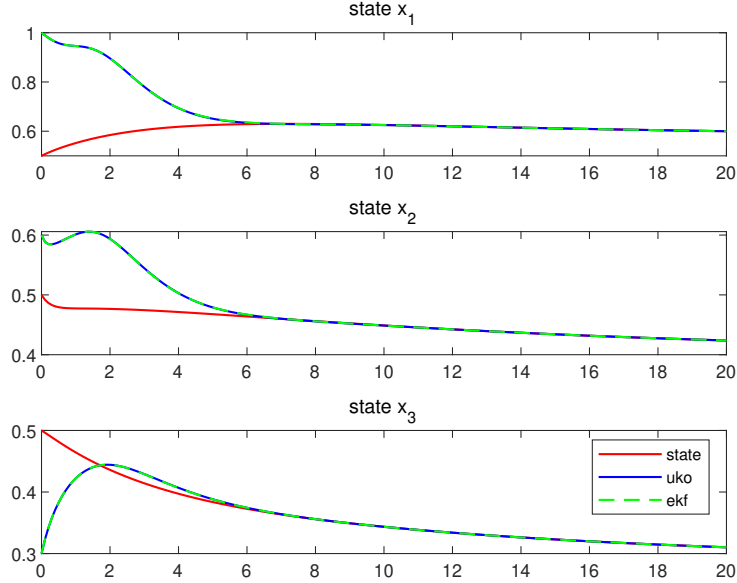


Figure 3: State estimation with UKO and EKF.

The convergence of this observer has been proved locally if the tuning parameter  $c$  of the observer is small enough.

In [9], we built a high-gain version of the unscented Kalman observer and performed some comparison of the performance of each observer (the unscented Kalman filter, the unscented Kalman observer, and the high-gain unscented Kalman filter). Moreover, if we assume that the solution of the Riccati equation is bounded, then the unscented Kalman observer becomes a globally converging observer. Using Theorem 3 of this paper and the proof in Appendix A, we should be able to prove the global exponential convergence of the high-gain unscented Kalman observer under the same observability conditions.

## Appendix A. Bounds on the solution of the Riccati equation

In this appendix, we prove Theorem 3, following [14]. However, the Riccati-like equation of UKO is different from the classical Kalman Riccati equation, and we interpret the difference as a perturbation of the Kalman Riccati equation. Only proofs notably different from those of [14] will be given in detail.

First of all, we introduce some notation and recall some basic properties.

### Notation 4.

- $\mathcal{M}_n$  is the set of  $(n \times n)$  matrices having real entries;
- $\mathcal{S}_n$  is the set of symmetric matrices of  $\mathcal{M}_n$ ;
- $\mathcal{S}_n^+$  is the set of positive definite matrices of  $\mathcal{S}_n$ ;
- $\text{Tr}(A)$  denotes the trace of the matrix  $A \in \mathcal{M}_n$ ;
- For any matrix  $A$ ,  $\|A\|_2$  denote the matrix norm induced from the Euclidean vector norm :  $\|A\|_2 = \sup_{\|x\|_2=1} \|Ax\|_2$ ;
- For any matrix  $A$ ,  $\|A\|_F$  denote the Frobenius norm:  $\|A\|_F = \sqrt{\text{Tr}(A^t A)}$ ;
- $\|\cdot\|$  denotes any matrix norm;
- For any matrix  $Q \in \mathcal{S}_n$ ,  $q_{\min}$  (respectively,  $q_{\max}$ ) denotes the smallest (respectively, largest) eigenvalue of  $Q$ .

We now recall some properties.

### Properties 6.

- $\frac{d}{dt} \text{Tr}(M) = \text{Tr} \frac{dM}{dt}$  for any  $M \in \mathcal{M}_n$ ;
- let  $A, B \in \mathcal{M}_n$ ,  $|\text{Tr}(AB)| \leq \sqrt{\text{Tr}(A^t A)} \sqrt{\text{Tr}(B^t B)}$ ;
- if  $S \in \mathcal{S}_n^+$ , then  $\frac{1}{n} \text{Tr}(S)^2 \leq \text{Tr}(S^2) \leq \text{Tr}(S)^2$ , and  $\text{Tr}(SQS) \geq \frac{q_{\min}}{n} \text{Tr}(S)^2$ ;

We consider the Riccati equation with a disturbance term depending on time as follows.

$$\begin{cases} \frac{dS(t)}{dt} = -A(t)^t S(t) - S(t)A(t) + C(t)^t R^{-1}(t)C(t) - S(t)(Q(t) + \omega(t))S(t) \\ S(0) = S_0 \end{cases} \quad (\text{A.1})$$

In addition, we assume that  $\omega(t)$  is symmetric and  $\|\omega(t)\|_F \leq \omega_0$  so that  $(Q(t) + \omega(t))$  remains symmetric positive definite. Since  $q_{\min} Id_n \leq Q(t) \leq q_{\max} Id_n$ , it is sufficient that  $\omega_0 < q_{\min}$ .

The proof is divided into three parts:

- In the first part, we prove the existence and the positive definiteness of  $S(t)$  and the existence of an upper bound on  $S(t)$ ;
- In the second part, we prove the existence of a lower bound on  $S(t)$ ;
- Finally, we deduce the result when  $\omega(t)$  is replaced by  $\Omega(S(t))$  as in Theorem 3.

*Appendix A.1. Upper bound with perturbation  $\omega(t)$*

**Lemma 7.** (Lemma 2.12 in [14]) For any  $\mu \in \mathbb{R}^*$  and any solution  $S : [0, T[ \rightarrow \mathcal{S}_n$  of (A.1) (possibly  $T = +\infty$ ), we have, for all  $t \in [0, T[$

$$\begin{aligned} S(t) &= e^{-\mu t} \Phi(t, 0) S_0 \Phi^t(t, 0) + \int_0^t e^{-\mu(t-\tau)} \Phi(t, \tau) C^t(\tau) R^{-1}(\tau) C(\tau) \Phi^t(t, \tau) d\tau \\ &\quad + \mu \int_0^t e^{-\mu(t-\tau)} \Phi(t, \tau) \left( S(\tau) - \frac{S(\tau)(Q(\tau) + \omega(\tau))S(\tau)}{\mu} \right) \Phi^t(t, \tau) d\tau, \end{aligned} \quad (\text{A.2})$$

*Proof.* Let  $\hat{S} : [0, T[ \rightarrow \mathcal{S}_n$ ,  $\hat{S}(t) = e^{\mu t} S(t)$ , and  $\frac{d\hat{S}(t)}{dt} = \mu e^{\mu t} S(t) + e^{\mu t} \dot{S}(t)$ . Then  $\hat{S}(t)$  satisfies

$$\begin{aligned} \frac{d\hat{S}(t)}{dt} &= \mu e^{\mu t} S(t) - e^{\mu t} A^t(t) S(t) - e^{\mu t} S(t) A(t) + e^{\mu t} C^t(t) R^{-1}(t) C(t) - e^{\mu t} S(t) (Q(t) + \omega(t)) S(t) \\ &= \mu \hat{S}(t) - A^t(t) \hat{S}(t) - \hat{S}(t) A(t) + e^{\mu t} C^t(t) R^{-1}(t) C(t) - e^{-\mu t} \hat{S}(t) (Q(t) + \omega(t)) \hat{S}(t). \end{aligned}$$

for  $t \in [0, T[$ . Applying the variation of constants formula to

$$\frac{\Lambda(t)}{dt} = -A(t)^t \Lambda(t) - \Lambda(t) A(t) + F(t).$$

and writing  $\Phi(t, \tau)$  for the state-transition matrix defined by  $\frac{d\Phi(t, \tau)}{d\tau} = -A(t)^t \Phi(t, \tau)$ ,  $\Phi(\tau, \tau) = Id_n$ , we get the relation (A.2), due to the fact that  $\Phi(t, \tau) = \Psi^t(\tau, t)$ . Where  $\Psi(\tau, t)$  is defined in (26) (Definition 2).  $\square$

**Lemma 8.** (Lemma 2.14 in [14]). If  $S : [0, T[ \rightarrow \mathcal{S}_n^+$  is a solution of Equation (A.1), then, for all  $t \in [0, T[$ ,

$$\frac{d \text{Tr}(S(t))}{dt} \leq -a \text{Tr}(S(t))^2 + 2b \text{Tr}(S(t)) + c \quad (\text{A.3})$$

where  $a = \frac{q_{\min} - \omega_0}{n}$ ,  $b = \sup_t \sqrt{\text{Tr}(A(t)^t A(t))}$ , and  $c = \frac{n}{r_{\min}} \sup_t (\text{Tr}(C(t)^t C(t)))$ .

Moreover,

$$\text{Tr}(S(t)) \leq \max(\text{Tr}(S(0)), s_1)$$

where  $s_1$  is the unique positive solution of  $-as^2 + 2bs + c = 0$ .

As a consequence,

$$\|S(t)\|_2 \leq \|S(t)\|_F = \sqrt{\text{Tr}(S(t)^2)} \leq \sqrt{\text{Tr}(S(t))^2} = \text{Tr}(S(t)) \leq \max(\text{Tr}(S(0)), s_1)$$

hence, setting  $\beta = \max(\text{Tr}(S(0)), s_1)$ , for all  $t < T$ ,  $S(t) \leq \beta Id_n$ .

*Proof.*

$$\begin{aligned} \frac{d}{dt} \text{Tr}(S(t)) &= -\text{Tr}(A(t)^t S(t)) - \text{Tr}(S(t) A(t)) - \text{Tr}(S(t)(Q(t) + \omega(t))S(t)) + \text{Tr}(C(t)^t R^{-1}(t)C(t)) \\ &\leq 2 |\text{Tr}(A(t)^t S(t))| - \text{Tr}(S(t)(Q(t) + \omega(t))S(t)) + \text{Tr}(C(t)^t R^{-1}(t)C(t)). \end{aligned}$$

The matrix is bounded due to Assumption 2, so using the properties of the trace (Properties 6),

$$\text{Tr}(S(t)(Q(t) + \omega(t))S(t)) \geq \frac{q_{\min} - \omega_0}{n} (\text{Tr}(S(t)))^2$$

$$|\text{Tr}(A(t)^t S(t))| \leq \sqrt{\text{Tr}(A(t)^t A(t))} \sqrt{\text{Tr}(S(t)^t S(t))} \leq \text{Tr}(A(t)^t A(t))^{\frac{1}{2}} [\text{Tr}(S(t)^2)]^{\frac{1}{2}} \leq \sup (\text{Tr}(A(t)^t A(t)))^{\frac{1}{2}} \text{Tr}(S(t))$$

and

$$|\text{Tr}(C(t)^t R^{-1}(t)C(t))| = |\text{Tr}(C(t)C(t)^t R^{-1}(t))| \leq \text{Tr}(C(t)C(t)^t) \text{Tr}(R^{-1}(t)) \leq \frac{n}{r_{\min}} \sup (\text{Tr}(C(t)^t C(t))).$$

Therefore, (A.3) holds. The fact is that  $\text{Tr}(S(t)) \leq \max(\text{Tr}(S(0)), s_1)$  is independent of the perturbation as soon as  $\|\omega(t)\|_F \leq \omega_0$ . Roughly speaking, this is a direct consequence of (A.3), by noting that

- $-as^2 + 2bs + c = 0$  has two roots  $s_0$  and  $s_1$  s.t.  $0 < s_0 < s_1$ ,
- $\frac{d\text{Tr}(S(t))}{dt}$  is negative once  $S(t) > s_1$ .

□

**Lemma 9.** (Lemma 2.15 in [14]). Let  $S : [0, e(S)[ \rightarrow \mathcal{S}_n^+$  be a maximal semi-solution of (A.1). If  $S(0) = S_0$  is positive definite, then  $e(S) = +\infty$  and  $S(t)$  is positive definite for all  $t \geq 0$  (and therefore,  $T = +\infty$  in Lemma 8).

*Proof.* Suppose that  $S$  is not always positive definite. Let  $\theta = \inf \{t | S(t) \notin \mathcal{S}_n^+\}$ . Then  $S(t) \in \mathcal{S}_n^+$  for all  $t \in [0, \theta[$ . Using lemma 2.13 in [14] and Lemma 8, we see that  $\text{Tr}(S(t)) \leq \max(\text{Tr}(S_0), s_1)$ . Then,  $\|S(t)\| = \sqrt{\text{Tr}(S(t)^2)} \leq \text{Tr}(S(t)) \leq \max(\text{Tr}(S_0), s_1)$ . Choose  $\mu > | (Q(t) + \omega(t)) | \max(\text{Tr}(S_0), s_1)$ . Apply Lemma 7 with  $t = \theta$ .  $S(\theta) = (I) + (II) + (III)$ ,

$$\begin{aligned} (I) &= e^{-\mu\theta} \Phi(\theta, 0) S_0 \Phi^t(\theta, 0), \\ (II) &= \int_0^\theta e^{-\mu(\theta-\tau)} \Phi(\theta, \tau) C^t(\tau) R(\tau)^{-1} C(\tau) \Phi^t(\theta, \tau) d\tau, \\ (III) &= \mu \int_0^\theta e^{-\mu(\theta-\tau)} \Phi(\theta, \tau) \left( S(\tau) - \frac{S(\tau)(Q(\tau) + \omega(\tau))S(\tau)}{\mu} \right) \Phi^t(\theta, \tau) d\tau. \end{aligned}$$

- (I) is obviously positive definite, hence  $S(0)$  is positive definite.
- (II)  $R(t)$  is symmetric positive definite and  $C(t)$  is bounded by  $\bar{c} > 0$  hence (II) is symmetric positive definite.
- (III) depends on  $S(\tau) - \frac{S(\tau)(Q(\tau) + \omega(\tau))S(\tau)}{\mu}$ , which can be rewritten

$$\sqrt{S(\tau)} \left( Id_n - \frac{\sqrt{S(\tau)}(Q(\tau) + \omega(\tau))\sqrt{S(\tau)}}{\mu} \right) \sqrt{S(\tau)}$$

Since  $S(t)$  is positive definite for  $0 < t < \theta$ . Therefore the positiveness of (III) depends on  $Id_n - \frac{\sqrt{S(\tau)}(Q(\tau) + \omega(\tau))\sqrt{S(\tau)}}{\mu}$ . From the definition of  $\mu$  we have

$$\frac{\sqrt{S(\tau)}(Q(\tau) + \omega(\tau))\sqrt{S(\tau)}}{\mu} < Id_n$$

and

$$| \sqrt{S(\tau)}(Q(\tau) + \omega(\tau))\sqrt{S(\tau)} | \leq | S(\tau) | \times | (Q(\tau) + \omega(\tau)) | \leq | (Q(\tau) + \omega(\tau)) | \max(\text{Tr}(S_0), s_1)$$

Hence  $S(\tau) - \frac{S(\tau)(Q(\tau) + \omega(\tau))S(\tau)}{\mu}$  is positive definite. So is (III). This is in contradiction with the supposition that  $S(\theta)$  is not positive definite.

Now by Lemma 8,  $\text{Tr}(S(t)) \leq \max(\text{Tr}(S_0), s_1)$  for all  $t \in [0, e(S)[$  and  $e(s) = +\infty$ . □

*Appendix A.2. Lower bound with perturbation  $\omega(t)$*

**Lemma 10.** For each time  $T_0 > 0$ , there exists a constant  $\gamma_0 > 0$  such that if  $S : [0, +\infty[ \rightarrow \mathcal{S}_n$  is a solution of (A.1) such that  $S(0) = S_0 \in \mathcal{S}_n^+$ , then

$$S(t) \geq \gamma_0 Id_n, \text{ for } t \geq T_0.$$

Moreover, if  $S(0) > 0$ , there exists  $\alpha$  such that  $S(t) \geq \alpha Id_n, \forall t \geq 0$ .

*Proof.* We use formula (A.2) from Lemma 7. The first and the third terms are positive (as soon as  $\mu$  is large enough, see the proof of Lemma 9). Now consider the second term,

$$\begin{aligned} \int_0^t e^{-\mu(t-\tau)} \Phi(t, \tau) C^t(\tau) R^{-1}(\tau) C(\tau) \Phi^t(t, \tau) d\tau &\geq \int_{t-T_0}^t e^{-\mu(t-\tau)} \Phi(t, \tau) C^t(\tau) R^{-1}(\tau) C(\tau) \Phi^t(t, \tau) d\tau \\ &\geq e^{-\mu T_0} \underbrace{\int_{t-T_0}^t \Phi(t, \tau) C^t(\tau) R^{-1}(\tau) C(\tau) \Phi^t(t, \tau) d\tau}_{W_{T_0}(t)} \\ &\geq e^{-\mu T_0} K(T_0) \end{aligned}$$

for any  $t \geq T_0$ , with  $K(T_0) > 0$  by definition of persistent observability (see [5]). We get the lemma taking  $\gamma_0 = e^{-\mu T_0} K(T_0)$ .

Therefore,  $S(t) > 0$  for all  $t \geq 0$  (Lemma 9) and  $S(t) \geq \gamma_0 Id_n \forall t \geq T_0$ . So,  $\alpha_0 = \inf_{0 \leq t \leq T_0} S(t)$  is strictly positive and  $\alpha = \inf(\alpha_0, \gamma_0)$  is such that  $S(t) \geq \alpha Id_n \forall t \geq 0$ , which proves the result. □

Appendix A.3. Bounds on the Riccati-like equation

Let us come back to the Riccati-like equation with perturbation  $\Omega(S(t))$  :

$$\frac{dS(t)}{dt} = -A(t)^t S(t) - S(t)A(t) + C(t)^t R^{-1}(t)C(t) - S(t)(Q(t) + \Omega(S(t)))S(t)$$

**Lemma 11.** *Suppose there exists  $\alpha_1$  and  $\beta_1$  such that  $\alpha_1 Id_n \leq S(t) \leq \beta_1 Id_n$ . Then there exists  $\Omega_0 > 0$ , and  $c_0 > 0$  small enough such that for any  $0 < c < c_0$ , we have*

$$\|\Omega(S(t))\|_F \leq \sqrt{c}\Omega_0$$

*Proof.* From (30), we have

$$\begin{aligned} \|\Omega(S(t))\|_F &\leq \frac{\sqrt{c}}{2}(2\|\sqrt{P(t)}\|_F\|H_f(t)\|_F) + \sqrt{c}\|P(t)\|_F\|C(t)\|\|R^{-1}(t)\|\|H_h(t)\|_F\|\sqrt{P(t)}\|_F \\ &\quad + \frac{c}{4}\|\sqrt{P(t)}\|_F^2\|H_h(t)\|_F^2\|R^{-1}(t)\|. \end{aligned} \quad (\text{A.4})$$

Under Assumption 1 on the second order derivatives of  $f$  and  $h$ , and Assumption 2, and the bound on  $S(t)$   $\|\sqrt{P(t)}\|_F^2 = \text{Tr}(P(t)) \leq n\|P(t)\|_F$ , then,  $\|\sqrt{P(t)}\|_F \leq \sqrt{n}\|P(t)\|_F^{\frac{1}{2}}$ , it follows that

$$\|\Omega(S(t))\|_F \leq \frac{\sqrt{c}n^{3/2}M_f n^{3/4}}{\alpha_1^{3/2}} + \frac{\sqrt{c}cn^2M_h n^{3/4}}{r_{\min}\alpha_1^{3/2}} + \frac{c}{4} \frac{n^5 M_h^2}{r_{\min}\alpha_1^3}, \quad (\text{A.5})$$

for  $c$  small enough and  $\|\Omega(S(t))\|_F \leq \sqrt{c}\Omega_0$ , with

$$\Omega_0 = \frac{n^{3/2}M_f n^{3/4}}{\alpha_1^{3/2}} + \frac{\bar{c}n^2M_h n^{3/4}}{r_{\min}\alpha_1^{3/2}} + \frac{\sqrt{c_0}}{4} \frac{n^5 M_h^2}{r_{\min}\alpha_1^3}. \quad (\text{A.6})$$

□

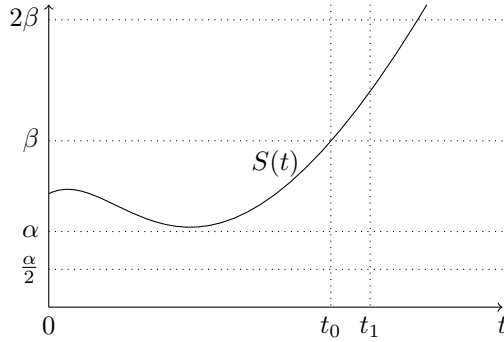


Figure A.4: Bounds on  $S(t)$

To prove the Theorem 3, which bounds the solution  $S(t)$  of the Riccati-like equation,

- We first choose  $\omega_0$  such that if  $\|\omega(t)\|_F \leq \omega_0$  then  $\alpha Id_n \leq S(t) \leq \beta Id_n$  (Lemmas 8, 9 and 10);
- Then we choose  $c$  such that if  $\frac{\alpha}{2} Id_n \leq S(t) \leq 2\beta Id_n$  then  $\|\Omega(S(t))\|_F \leq \sqrt{c}\Omega_0 \leq \omega_0$  (Lemma 11).

We have to prove that  $\forall t \geq 0$ ,  $\alpha Id_n \leq S(t) \leq \beta Id_n$ .

Using a reductio ad absurdum, let  $t_0$  be the largest time value such that,  $\forall t \leq t_0$ ,  $\alpha Id_n \leq S(t) \leq \beta Id_n$ . Then there exist  $t_1 > t_0$  (see Figure A.4) such that  $\frac{\alpha}{2} Id_n \leq S(t) \leq 2\beta Id_n$  because  $S(t)$  is continuous, and so  $\|\Omega(S(t))\|_F \leq \omega_0$  for  $t \in [t_0, t_1]$ , and therefore for any  $t_0 < t \leq t_1$ ,  $\alpha Id_n \leq S(t) \leq \beta Id_n$ . This contradicts the existence of  $t_0$ .



## References

- [1] J. Almeida, P. Oliveira, C. Silvestre and A. Pascoal. "State estimation of nonlinear systems using the Unscented Kalman Filter," *TENCON 2015 - 2015 IEEE Region 10 Conference*, Nov. 2015.
- [2] B.D.O. Anderson and J.B. Moore. *Optimal control: linear quadratic methods*. Prentice Hall, 1989.
- [3] J. S. Baras, A. Bensoussan and M. R. James. "Dynamic observers as asymptotic limits of recursive filters: Special cases", *SIAM Journal on Applied Mathematics*, vol. 48, no. 5, pp. 1147–1158, 1988.
- [4] N. Boizot and E. Busvelle. "Adaptive-gain observers and applications," In: *Nonlinear Observers and Applications*, Springer-Verlag, Berlin, 2007, pages 71–114.
- [5] G. Bornard, F. Celle-Couenne and G. Gilles. Observability and observers. *Nonlinear Systems*, pp. 173–216, 1995.
- [6] M. Boutayeb, H. Rafaralahy, and M. Darouach. Convergence analysis of the extended Kalman filter used as an observer for nonlinear deterministic discrete-time systems, *IEEE Transactions on Automatic Control* vol. 42, no. 4, pp. 581–586, Apr. 1997.
- [7] E. Busvelle, and J. -P. Gauthier, "Observation and identification tools for non-linear systems: application to a fluid catalytic cracker," *International Journal of Control*, vol. 78, no. 3, pp. 208–234, Feb. 2005.
- [8] C. K. Chui and G. Chen. *Kalman Filtering with Real-Time Applications*, 3rd ed. Springer-Verlag, Berlin, 1998.
- [9] A. Daid, E. Busvelle and M. Aidene. Unscented Kalman Observer, In: *8th International Conference on Systems and Control*, 24-25 Octobre 2019.
- [10] P. Del Moral, A. Kurtzmann, and J. Tugaut. "On the stability and the concentration of extended Kalman-Bucy filters", *Electronic Journal of Probability*, vol. 23, no. 0, 2018.
- [11] F. Deza, E. Busvelle, and J.-P. Gauthier. "Exponentially converging observers for distillation columns and internal stability of the dynamic output feedback," *Chemical Engineering Science*, vol. 47, no. 15–16, pp. 3935–3941, Oct. 1992.
- [12] M. Doumiati, A. Victorino, A. Charara and D. Lechner. Unscented Kalman filter for real-time vehicle lateral tire forces and sideslip angle estimation, In: *IEEE Intelligent Vehicles Symposium*, Jun. 2009.
- [13] G. Dymirkovsky. New results on stochastic stability of discrete-time Unscented Kalman Filter, In: *7th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, Jul. 2012.
- [14] J.-P. Gauthier and I. A. K. Kupka. *Deterministic Observation Theory and Applications*, Cambridge University Press, 2001.
- [15] F. Gustafsson and H. Gustaf. Some relations between extended and unscented Kalman filters, *IEEE Transactions on Signal Processing* vol. 60, no. 2, pp. 545–555, 2011
- [16] G. Hu, S. Gao, Y. Zhong, and B. Gao. "Stochastic stability of the derivative unscented Kalman filter," *Chinese Physics B*, vol. 24, no. 7, p. 070202, Jun. 2015.
- [17] G. Hu, S. Gao, and Y. Zhong. "A derivative UKF for tightly coupled INS/GPS integrated navigation," *ISA Transactions*, vol. 56, pp. 135–144, May 2015.
- [18] A. H. Jazwinski. *Stochastic Processes and Filtering Theory*. Dover Publications, N.Y., 1970.
- [19] S. J. Julier. "The Scaled Unscented Transformation", In: *Proceedings of the American Control Conference* (May 2002), vol. 6, pp. 4555–4559.
- [20] S. J. Julier, J. K. Uhlmann, and H. F. Durrant-Whyte. "A new approach for filtering nonlinear systems", In: *Proceedings of 1995 American Control Conference*, 1995.
- [21] S. Julier, J. Uhlmann, and H. F. Durrant-Whyte. "A new method for the nonlinear transformation of means and covariances in filters and estimators," *IEEE Transactions on Automatic Control* vol. 45, no. 3, pp. 477–482, Mar. 2000.
- [22] S. J. Julier and J. K. Uhlmann. New extension of the Kalman filter to nonlinear systems, In: *Signal processing, sensor fusion, and target recognition VI* International Society for Optics and Photonics vol. 3068, pp. 182–193, 1997.

- [23] S. J. Julier and J. K. Uhlmann. “Unscented Filtering and Nonlinear Estimation,” *Proceedings of the IEEE* vol. 92, no. 3, pp. 401–422, Mar. 2004.
- [24] S. J. Julier and J. K. Uhlmann. *A general method for approximating nonlinear transformations of probability distributions*. Technical report, Robotics Research Group, Department of Engineering Science, University of Oxford, 1996.
- [25] R. E. Kalman. A new approach to linear filtering and prediction problems, *Trans. ASME Ser.D., J. Basic Eng.* vol. 82, pp. 35-45, 1960.
- [26] T. Karvonen, S. Bonnabel, E. Moulines, and S. Särkkä. ”On stability of a class of filters for non-linear stochastic systems.” *arXiv preprint arXiv:1809.05667* (2018).
- [27] F. Lafont, E. Busvelle, and J.-P. Gauthier. An adaptive high-gain observer for wastewater treatment systems, *Journal of Process Control* vol. 21, no. 6, 2011, pp. 893–900.
- [28] L. Li and Y. Xia. ”Stochastic stability of the unscented Kalman filter with intermittent observations”, *Automatica* vol. 48, no. 5, pp. 978–981, May 2012.
- [29] X. Liu, Q. Gao, and X.-L. Li. ”Control of epileptiform spikes based on nonlinear unscented Kalman filter”, *Chinese Physics B*, vol. 23, no. 1, p. 010202, Jan. 2014.
- [30] W. Lohmiller and J.-J. E. Slotine. “On Contraction Analysis for Non-linear Systems”, *Automatica*, vol. 34, no. 6, pp. 683–696, Jun. 1998.
- [31] D. G. Luenberger. Observers for multivariable linear systems, *IEEE Trans. on Automatic Control* vol. 2, pp. 190-197, 1966.
- [32] D. Luenberger. ”An Introduction to Observer”, *IEEE Transactions on Automatic and Control*, AC-16 : pp. 596–602, 1971.
- [33] J. P. Maree, L. Imsland, and J. Jouffroy. ” On convergence of the unscented Kalman–Bucy filter using contraction theory,” *International Journal of Systems Science*, vol. 47, no. 8, pp. 1816–1827, Aug. 2014.
- [34] M. Rhudy, Y. Gu, J. Gross, and M. R. Napolitano. Evaluation of Matrix Square Root Operations for UKF within a UAV GPS/INS Sensor Fusion Application, *International Journal of Navigation and Observation* vol. 2011, 2011.
- [35] E. Pardoux. ”Filtrage non linéaire et équations aux dérivées partielles stochastiques associées”, In: *Ecole d’Eté de Probabilités de Saint-Flour XIX 1989*, Springer-Verlag, Berlin, 1991, pages 68–163.
- [36] J. Picard, ”Efficiency of the extended Kalman filter for nonlinear systems with small noise”, *SIAM Journal on Applied mathematics* vol. 51, no. 3, pp. 843–885, 1991.
- [37] K. Reif, S. Gunther, E. Yaz, and R. Unbehauen. Stochastic stability of the discrete-time extended Kalman filter *IEEE Transactions on Automatic Control* vol. 44, no. 4, pp. 714–728, Apr. 1999.
- [38] K. Reif, F. Sonnemann, and R. Unbehauen An EKF-Based Nonlinear Observer with a Prescribed Degree of Stability, *Automatica* vol. 34, no. 9, pp. 1119–1123, Sep. 1998.
- [39] K. Reif, F. Sonnemann, and R. Unbehauen Modification of the extended Kalman filter with an additive term of instability. In: *Proceedings of 35th IEEE Conference on Decision and Control*.
- [40] K. Reif and R. Unbehauen The extended Kalman filter as an exponential observer for nonlinear systems *IEEE Transactions on Signal Processing* vol. 47, no. 8, pp. 2324–2328, 1999.
- [41] K. Reif, R. Unbehauen, E. Yaz, and S. Günther, Stochastic stability of the continuous-time extended Kalman filter, *IEE Proceedings-Control Theory and Applications* vol. 147, no. 1, pp. 45–52, Jan. 2000
- [42] S. Särkkä. On Unscented Kalman Filtering for State Estimation of Continuous-Time Nonlinear Systems, *IEEE Transactions on Automatic Control* vol. 52, no. 9, pp. 1631–1641, Sep. 2007.
- [43] S. Särkkä, Bayesian Filtering and Smoothing, *Cambridge University Press*, 2009.
- [44] S. Simani, C. Fantuzzi, and R. J. Patton. *Model-based fault diagnosis in dynamic systems using identification techniques*, Springer-Verlag, Berlin, 2002.
- [45] Y. Song, and J. W. Grizzle. The Extended Kalman Filter as a Local Asymptotic Observer for Nonlinear Discrete-Time Systems. In: *American Control Conference*, pp.3365–3369, Jun. 1992.

- [46] N. Stastny, R. Bettinger, and F. Chavez. Comparison of the extended and unscented Kalman filters for angles based relative navigation. In: *AIAA/AAS Astrodynamics Specialist Conference and Exhibit*, p. 7083. 2008.
- [47] F. Viel, E. Busvelle, and J. -P. Gauthier, “Stability of polymerization reactors using I/O linearization and a high-gain observer,” *Automatica*, vol. 31, no. 7, pp. 971–984, Jul. 1995.
- [48] E. A. Wan and R. Van Der Merwe. The unscented Kalman filter for nonlinear estimation. In: *Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium AS-SPCC*, pp. 153–158, 2000
- [49] K. Xiong, H. Y. Zhang, and C. W. Chan. Performance evaluation of UKF-based nonlinear filtering, *Automatica* vol. 42, no. 2, pp. 261–270, Feb. 2006.
- [50] J. Xu, S. Wang, G. M. Dimirovski, and Y. Jing. Stochastic stability of the continuous-time unscented Kalman filter, *47th IEEE Conference on Decision and Control*, pp. 5110–5115, 2008.
- [51] Z. Yacine. *Observateurs pour l’Estimation de la Dynamique Latérale du véhicule et Application à la Détection de Situations Critiques*, Thèse de Doctorat, Université Mouloud Mammeri, Tizi Ouzou, Algeria, 2016.
- [52] Y.Zou, D.Du, B.Chang, L.Ji, and J.Pan. Automatic weld defect detection method based on Kalman filtering for real-time radiographic inspection of spiral pipe, *NDT & E International* vol. 72, pp. 1-9, 2015.
- [53] Z. Zu-Tao and Z. Jia-Shu. “Sampling strong tracking nonlinear unscented Kalman filter and its application in eye tracking,” *Chinese Physics B*, vol. 19, no. 10, p. 104601, Oct. 2010.