2009 American Control Conference
Hyatt Regency Riverfront, St. Louis, MO, USA
June 10-12, 2009

The Unscented Kalman Filtering in Extended Noise Environments

Yucheng Zhou, Jiahe Xu, Yuanwei Jing, Georgi M. Dimirovski, *Senior Member, IEEE*

Abstract—**This paper introduces an extended environment for the unscented Kalman filtering that considers also the presence of additive noise on input observations in order to solve the problem of optimal estimation of noise-corrupted input and output sequences. This environment includes as sub-cases both errors-in-variables filtering and unscented Kalman filtering. The unscented Kalman filtering to the presence of additive noise on input observations is considered, and is used to solve the problem of optimal estimation of noise-corrupted input and output sequences. A Monte Carlo simulation shows that the performance of the unscented Kalman filtering technique leads to the expected minimal variance estimates.**

I. Introduction

THE filtering for nonlinear dynamic system is an important research area and has attracted considerable interest. A large number of suboptimal approaches have been developed to solve the nonlinear filtering problem. These include extended Kalman filtering (EKF) [1]-[4], Gaussian sum filter [5], grid-based methods [6] and particle filters [7], [8]. Among these methods, very few algorithms have seen so many applications in different areas as the extended Kalman filtering that constitutes the standard for information retrieval from noisy data generated by known processes and affected by noise with known statistical properties. However, the EKF may diverge in the case of large initial estimation error or large external disturbances. The reason is that the EKF is obtained by first-order linearization of the nonlinear model, and the approximation error between the linearised model and the original nonlinear system may become significant when there are large deviations of the estimated state from the real one [9].

The unscented Kalman filter (UKF) [10] is proposed by the scholars of Oxford University in 1990s as an improvement to EKF. This method is based on the unscented transform (UT) technique, a mechanism for propagating mean and covariance through a nonlinear transformation. The state vector is represented by a minimal set of carefully chosen sample points, called sigma points, which approximate the posterior mean and covariance of the Gaussian random variable with a second order accuracy. In contrast, the linearization technique used in the EKF can only achieve first order accuracy [11], [12]. Further, it is not necessary to compute the Jacobian matrices in the UKF [13], [14]. The UKF is widely used in practice, ranging from multi-sensor fusion [15], target tracking [10], position determination [16], to training of neural networks [17].

A limit of the stochastic environment of Kalman filtering concerns its asymmetrical description of the uncertainties on the observations, while the output is considered as affected by additive noise, the input is assumed as exactly known. This condition is met in all control applications where the process input is generated by known laws but restrictive in other contexts. Symmetrical environments are at the basis of errors-in-variables (EIV) models that consider all system attributes as affected by unknown additive disturbances. This symmetry allows avoiding system orientation, such as the necessity of partitioning observations into inputs and outputs.

The UKF cannot be directly applied to EIV filtering problems, as discussed in [18], because of the optimal reconstruction of inputs and outputs of EIV models on the basis of their noise-corrupted observations. The problem of EIV filtering has been recently solved in [19] by making reference to both behavioral and state-space contexts, starting from the solution of optimal EIV interpolation. The numerical aspects of this problem have then been investigated in [20], where a high-efficiency algorithm has been described based on the properties of Cholesky factorization. Optimal EIV filtering can be approached also in a deterministic context, as an optimization problem along the line followed by Roorda and Heij [20]. A unified context for both Kalman and EIV filtering has subsequently been developed in [21] and [22] by extending Kalman filtering to the more general case of symmetrical noise environments in order to include both Kalman and EIV filtering, as particular cases.

This paper proposes an extended unified context for both Kalman and EIV filtering based on the unscented Kalman filtering and considers also time-varying processes and the possibility of mutual noise correlations. Section II is dedicated to a statement of the filtering problem while Section III considers its solution included as sub-cases both errors-in-variables filtering, which is the optimal estimate of inputs and outputs from noisy observations and the unscented Kalman filtering, which is the optimal estimate of state and output in presence of state and output noise. The evaluation of the expected performance of the filter is then considered in

This work is supported by the National Natural Science Foundation of China under Grant 60274009 and Specialized Research Fund for the Doctoral Program of Higher Education under Grant 20020145007, P. R. of China, as well as in part by Dogus University Fund for Science and Ministry of Education and Science of the Republic of Macedonia.

Yucheng Zhou is Faculty of Engineering, Department of Research Institute of Wood Industry Chinese Academy of Forestry, Beijing, 100091, P.R. of China. E-mail:zhouyc@ caf.ac.cn

Jiahe Xu and Yuanwei Jing are with Faculty of Information Science and Engineering, Northeastern University, 110004, Shenyang, Liaoning, P.R. of China. E-mail: ellipisis@163.com

G. M. Dimirovski is with Dogus University, Faculty of Engineering, Istanbul, TR-34722, R. of Turkey; and SS Cyril and Methodius University, Faculty of FEIT, Skopje, MK-1000, R. of Macedonia. E-mail: gdimirovski @dogus.edu.tr.

978-1-4244-4524-0/09/$25.00 ©2009 AACC

Section IV. The results of a Monte Carlo simulation are reported in Section V and some concluding remarks are finally reported in Section VI.

II. Statement of the Problem

The considered nonlinear system is represented by the state-space equations

$$x(t+1) = f(x(t),t) + B(t)\hat{u}(t) + G(t)w(t)\,,\ x(0) = x_0 \qquad (1)$$

$$\hat{y}(t) = h(x(t),t) + D(t)\hat{u}(t) \qquad (2)$$

where $x(t) \in R^n$, $\hat{u}(t) \in R^r$ and $y(t) \in R^m$ denote the state, input and output processes while $w(t)$ is the noise acting on the state. The true input and output are unknown and only the noisy observations are available.

$$u(t) = \hat{u}(t) + \tilde{u}(t) \qquad (3)$$

$$y(t) = \hat{y}(t) + \tilde{y}(t) \qquad (4)$$

where, $\tilde{u}(t)$ and $\tilde{y}(t)$ denote the additive noise on $\hat{u}(t)$ and $\hat{y}(t)$. In the sequel, we will assume that $w(t)$, $\tilde{u}(t)$ and $\tilde{y}(t)$ are zero mean white processes, uncorrelated with $\hat{u}(t)$ and with covariances as follows.

$$E[w(t)w^T(t-\tau)] = Q_w(t)\delta(\tau) \qquad (5)$$

$$E[\tilde{u}(t)\tilde{u}^T(t-\tau)] = \tilde{Q}_u(t)\delta(\tau) \qquad (6)$$

$$E[\tilde{y}(t)\tilde{y}^T(t-\tau)] = \tilde{Q}_y(t)\delta(\tau) \qquad (7)$$

$$E[\tilde{u}(t)\tilde{y}^T(t-\tau)] = \tilde{Q}_{uy}(t)\delta(\tau) \qquad (8)$$

$$E[w(t)\tilde{u}^T(t-\tau)] = 0 \qquad \forall\tau \qquad (9)$$

$$E[w(t)\tilde{y}^T(t-\tau)] = 0 \qquad \forall\tau \qquad (10)$$

where $\delta(\tau)$ denotes the Kronecker delta function. The initial state x_0 is a random vector with mean $\bar{x}_0$ and covariance matrix P_0, uncorrelated with $w(t)$, $\tilde{u}(t)$ and $\tilde{y}(t)$, $\forall t$. Concerning the uncorrelation between the state noise $w(t)$ and the measurement noise $\tilde{u}(t)$, $\tilde{y}(t)$, assumptions (9) and (10) have been introduced only for the sake of simplicity and can be easily removed.

The optimal filtering problem can be defined that given model (1)-(4), covariance matrices (5)-(10) and the input-output observations $\{u(0), y(0),\ldots, u(t), y(t)\}$, determine the optimal nonlinear estimates of $\hat{u}(t)$ and $\hat{y}(t)$, at every *t*.

III. Optimal Filtering

A. Errors-in-variables Filtering

In this section, a simple approach to present errors-in-variables filtering, which include optimal estimate of inputs and outputs from noisy observations, is given.

In this part, the errors-in-variables filter is obtained based on the following theorem.

Theorem 1. The optimal estimates $\hat{u}^*(t)$ and $\hat{y}^*(t)$ of $\hat{u}(t)$ and $\hat{y}(t)$ that can be obtained from $\{u(0), y(0),\ldots, u(t), y(t)\}$, under constraints (1)-(4) are given by

$$\hat{u}^*(t) = u(t) - \tilde{u}^*(t) = u(t) - E[\tilde{u}(t) \mid z(t)] \qquad (11)$$

$$\hat{y}^*(t) = y(t) - \tilde{y}^*(t) = y(t) - E[\tilde{y}(t) \mid z(t)] \qquad (12)$$

with

$$z(t) = y(t) - D(t)u(t) \qquad (13)$$

where $E[\bullet]$ denotes mathematical expectation and $E[x \mid y]$ is the nonlinear minimal variance estimator that coincides with the conditional expectation in the gaussian case.

Proof. For the sake of simplicity, we will make reference to a time-invariant system described by the matrices $\partial f(x(t))/\partial x$, B, G, $\partial h(x(t))/\partial x$, D and to stationary white processes described by the covariances $Q_w(t)$, $\tilde{Q}_u(t)$, $\tilde{Q}_y(t)$ and $\tilde{Q}_{uy}(t)$. As it will be clear later, it is not restrictive to assume $x_0 = 0$.

Let us define the following vectors:

$$\hat{u} = [\hat{u}(0)\ \ \hat{u}(1)\ \ldots\ \hat{u}(t)]^T \qquad (14)$$

$$\hat{y} = [\hat{y}(0)\ \ \hat{y}(1)\ \ldots\ \hat{y}(t)]^T \qquad (15)$$

$$w = [w(0)\ \ w(1)\ \ldots\ w(t)]^T \qquad (16)$$

It is easy to verify that, because of (1)-(2), they are linked by the relation

$$\hat{y} = M_u\hat{u} + M_w w \qquad (17)$$

where

$$M_u = \begin{bmatrix} D & 0 & \cdots & 0 \\ \dfrac{\partial h(x)}{\partial x}B & D & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ \dfrac{\partial h(x)}{\partial x}[\dfrac{\partial f(x)}{\partial x}]^{t-1}B & \dfrac{\partial h(x)}{\partial x}[\dfrac{\partial f(x)}{\partial x}]^{t-2}B & \cdots & D \end{bmatrix}$$

$$M_u = \begin{bmatrix} 0 & \cdots & \cdots & 0 \\ \dfrac{\partial h(x)}{\partial x}G & 0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ \dfrac{\partial h(x)}{\partial x}[\dfrac{\partial f(x)}{\partial x}]^{t-1}G & \cdots & \dfrac{\partial h(x)}{\partial x}G & 0 \end{bmatrix}$$

By denoting with I_k the $k \times k$ identity matrix and with $\otimes$ the Kronecker product, relation (17) can also be expressed in the form

$$M\hat{v} = 0 \qquad (18)$$

where

$$\hat{v} = [\hat{y}^T\ \ \hat{u}^T\ \ w^T]^T \qquad (19)$$

and

$$M = [I_{t+1} \otimes I_m\ \ -M_u\ \ -M_w] \qquad (20)$$

Define now the vectors

$$v = [y^T\ \ u^T\ \ \underbrace{0\ \ldots\ 0}_{(t+1)\times p}]^T \qquad (21)$$

$$\tilde{v} = [y^T\ \ u^T\ \ -w^T]^T \qquad (22)$$

where

$$u = [u(0)\ \ u(1)\ \ldots\ u(t)]^T \qquad (23)$$

$$\tilde{u} = [\tilde{u}(0)\ \ \tilde{u}(1)\ \ldots\ \tilde{u}(t)]^T \qquad (24)$$

$$y = [y(0)\ \ y(1)\ \ldots\ y(t)]^T \qquad (25)$$

$$\tilde{y} = [\tilde{y}(0)\ \ \tilde{y}(1)\ \ldots\ \tilde{y}(t)]^T \qquad (26)$$

From (3)-(4), it follows immediately:

$$v = \hat{v} + \tilde{v} \qquad (27)$$

so that (18) leads to

$$Mv = M\tilde{v} = \Psi \qquad (28)$$

Now we need find the optimal estimate $\hat{v}^*$ of $\hat{v}$ satisfying condition (18), starting from the data v, the model M and the covariances (5)-(10). The solution can rely on two different approaches whether $w(t)$, $\tilde{u}(t)$ and $\tilde{y}(t)$ are Gaussian processes or not. In the former case, the problem can be solved by maximizing the likelihood function

$$J(\hat{v}^*) = -\frac{1}{2}(v-\hat{v}^*)^T \tilde{Q}^{-1}(v-\hat{v}^*) \quad (29)$$

under the constraint (18), where

$$\tilde{Q} = E[\tilde{v}\tilde{v}^T] = \begin{bmatrix} I_{t+1}\otimes\tilde{Q}_y & I_{t+1}\otimes\tilde{Q}_{uy}^T & 0 \\ I_{t+1}\otimes\tilde{Q}_{uy} & I_{t+1}\otimes\tilde{Q}_u & 0 \\ 0 & 0 & I_{t+1}\otimes Q_w \end{bmatrix} \quad (30)$$

By introducing the Lagrange multipliers vector λ the solution can be obtained by minimizing the loss function

$$J(\hat{v}^*,\lambda) = (v-\hat{v}^*)^T \tilde{Q}^{-1}(v-\hat{v}^*) + \lambda^T M\hat{v}^* \quad (31)$$

By equating to zero the gradient vectors of (31) with respect to $\hat{v}^*$ and λ we obtain

$$\lambda = 2(M\tilde{Q}M^T)^{-1}Mv \quad (32)$$

$$\hat{v}^* = v - \frac{\tilde{Q}M^T\lambda}{2} \quad (33)$$

The maximum likelihood solution is thus given by

$$\hat{v}^* = v - \tilde{Q}M^T(M\tilde{Q}M^T)^{-1}Mv \quad (34)$$

Note that the same result can be obtained also by considering the minimal variance estimate of the noise vector $\tilde{v}$, given by the conditional expectation

$$\tilde{v}^* = E[\tilde{v}\,|\,\Psi] = E[\tilde{v}\;\Psi^T]E[\Psi\Psi^T]\Psi = \tilde{Q}M^T(M\tilde{Q}M^T)^{-1}\Psi \quad (35)$$

where Ψ is defined by (28); in fact

$$\hat{v}^* = v - \tilde{v}^* \quad (36)$$

In the non-Gaussian case, relation (35) can be obtained by solving a weighted least-squares problem and constitutes the best nonlinear estimate of $\tilde{v}$ that can be obtained from v under condition (28).

The generic samples $\tilde{u}^*(\tau)$, $\tilde{y}^*(\tau)$, ($0 \le \tau \le t$) obtained from (35) depend on both past and future data; the only exception concerns $\tilde{u}^*(0)$, $\tilde{y}^*(0)$ that do not depend on past samples and $\tilde{u}^*(t)$, $\tilde{y}^*(t)$ that do not depend on future samples. $\tilde{u}^*(t)$ and $\tilde{y}^*(t)$ are thus the filtered quantities required for the optimal filter. Because of (35), they are given by

$$\tilde{u}^*(t) = E[\tilde{u}(t)\,|\,\Psi] = E[\tilde{u}(t)\,|\,\psi(t),\ \psi(t-1),\ldots,\ \psi(0)] \quad (37)$$

$$\tilde{y}^*(t) = E[\tilde{y}(t)\,|\,\Psi] = E[\tilde{y}(t)\,|\,\psi(t),\ \psi(t-1),\ldots,\ \psi(0)] \quad (38)$$

where, from (28)

$$\psi(\tau) = y(\tau) - \sum_{k=0}^{\tau-1}\frac{\partial h(x)}{\partial x}[\frac{\partial f(x)}{\partial x}]^{t-1-k}Bu(k) - Du(\tau) \quad (39)$$

Since $\tilde{u}(t)$ and $\tilde{y}(t)$ are white and uncorrelated with $\hat{u}(t)$ they do not depend on the past values of $u(t)$ and $y(t)$ so that

$$\tilde{u}^*(t) = E[\tilde{u}(t)\,|\,\Psi] = E[\tilde{u}(t)\,|\,\psi(t)] = E[\tilde{u}(t)\,|\,y(t)-Du(t)] \quad (41)$$

$$\tilde{y}^*(t) = E[\tilde{y}(t)\,|\,\Psi] = E[\tilde{y}(t)\,|\,\psi(t)] = E[\tilde{y}(t)\,|\,y(t)-Du(t)] \quad (42)$$

By recalling (36), the filtered samples $\hat{u}^*(t)$, $\hat{y}^*(t)$ can finally be obtained by means of the relations

$$\hat{u}^*(t) = u(t) - \tilde{u}^*(t) = u(t) - E[\tilde{u}(t)\,|\,y(t)-Du(t)] \quad (43)$$

$$\hat{y}^*(t) = y(t) - \tilde{y}^*(t) = y(t) - E[\tilde{y}(t)\,|\,y(t)-Du(t)] \quad (44)$$

If the initial state x_0 is not zero, the term $M_x x_0$ will appear in (17).

where

$$M_x = \left[[\frac{\partial h(x)}{\partial x}]^T \quad [\frac{\partial f(x)}{\partial x}]^T[\frac{\partial h(x)}{\partial x}]^T \quad \cdots \quad [(\frac{\partial f(x)}{\partial x})^t]^T[\frac{\partial h(x)}{\partial x}]^T \right]^T$$

This quantity must be included in the constraint (18); it does not affect, however, the obtained results since x_0 has been assumed as uncorrelated with $\tilde{u}(t)$ and $\tilde{y}(t)$. □

B. Unscented Kalman Filtering

Expressions (11) and (12) can be computed recursively by taking advantage of the unscented Kalman filter. Note that relations (3) and (4) allow writing model (1)-(2) in the form

$$x(t+1) = f(x(t),t) + B(t)\hat{u}(t) - B(t)\tilde{u}(t) + G(t)w(t) \quad (45)$$

$$\hat{y}(t) = h(x(t),t) + D(t)\hat{u}(t) - D(t)\tilde{u}(t) + \tilde{y}(t) \quad (46)$$

By introducing the auxiliary white processes

$$n_x(t) = G(t)w(t) - B(t)\tilde{u}(t) \quad (47)$$

$$n_y(t) = \tilde{y}(t) - D(t)\tilde{u}(t) \quad (48)$$

with covariances

$$Q(t) = E[n_x(t)n_x^T(t)] = G(t)Q_w(t)G^T(t) + B(t)\tilde{Q}_u(t)B^T(t) \quad (49)$$

$$\begin{aligned} R(t) &= E[n_y(t)n_y^T(t)] \\ &= \tilde{Q}_y(t) - \tilde{Q}_{uy}^T(t)D^T(t) - D(t)\tilde{Q}_{uy}(t) + D(t)\tilde{Q}_u(t)D^T(t) \end{aligned} \quad (50)$$

$$S(t) = E[n_x(t)n_y^T(t)] = B(t)[\tilde{Q}_u(t)D^T(t) - \tilde{Q}_{uy}(t)] \quad (51)$$

It is possible to rewrite relations (45)-(46) as

$$x(t+1) = f(x(t),t) + B(t)u(t) + n_x(t) \quad (52)$$

$$z(t) = h(x(t),t) + n_y(t) \quad (53)$$

According to the procedure of the normal UKBF, the n-dimensional random variable $x(t)$ with mean $\hat{x}(t)$ and covariance $P(t)$ is approximated by the matrix of sigma points $X(t)$ selected using the following equations firstly.

$$X^{(0)}(t) = \hat{x}(t)$$

$$X^{(i)}(t) = \hat{x}(t) + \sqrt{cP},\ i = 1,\ldots,L$$

$$X^{(i)}(t) = \hat{x}(t) - \sqrt{cP},\ i = L+1,\ldots,2L$$

where $c = \alpha^2(n+l)$ is a tuning parameter. The opposite weight ω_m is as follow

$$\omega_m = [W_m^{(0)} \quad \cdots \quad W_m^{(2n)}]^{\mathrm{T}}$$

where

$$W_m^{(0)} = \frac{\lambda}{(n+\lambda)},\ W_m^{(i)} = \frac{1}{2(n+\lambda)}\ (i = 1,\cdots,2n).$$

We define the matrix W as follow

$$W = (I - [\omega_m \cdots \omega_m]) \times diag(W_c^{(0)} \cdots W_c^{(2n)}) \times (I - [\omega_m \cdots \omega_m])^{\mathrm{T}}$$

where

$$W_c^{(0)} = \frac{\lambda}{(n+\lambda)+(1-\alpha^2+\beta)},\ W_c^{(i)} = \frac{1}{2(n+\lambda)}\ (i = 1,\cdots,2n).$$

The parameter λ is a scaling parameter defined as $\lambda = \alpha^2(n+l) - n$. The positive constants α, β and l are used as parameters of the method.

In the previous formulation, $z(t)$ plays the role of measured output so that the corresponding Kalman filter equations can be defined as follows.

$$\begin{aligned} x(t+1\,|\,t) &= f(X(t),t)\omega_m + B(t)u(t) + K(t)\varepsilon(t) \\ x(0\,|\,-1) &= \tilde{x}_0 \end{aligned} \quad (54)$$

$$K(t) = [X(t)Wh^{\mathrm{T}}(X(t),t) + S(t)]Q_\varepsilon^{-1}(t) \quad (55)$$

$$\begin{aligned} P(t+1\,|\,t) &= X(t)Wf^{\mathrm{T}}(X(t),t) + f(X(t),t)WX^{\mathrm{T}}(t) \\ &\quad + Q(t) - K(t)Q_\varepsilon^{-1}(t)K^{\mathrm{T}}(t) \end{aligned} \quad (56)$$

$$P(0\,|\,-1) = P_0$$

where, $\varepsilon(t)$ and $Q_\varepsilon(t)$ denote the innovation of $z(t)$ and its covariance matrix, given by

$$\varepsilon(t)=[z(t)-h(X(t),t)\omega_m] \quad (57)$$

$$Q_\varepsilon^{-1}(t)=h(X(t),t)Wh^{\mathrm{T}}(X(t),t)+R(t) \quad (58)$$

To compute $\hat{u}^*(t)$ and $\hat{y}^*(t)$, we can replace $z(t)$ with its innovation

$$\hat{u}^*(t)=E[\tilde{u}(t)\,|\,\varepsilon(t)]=E[\tilde{u}(t)\varepsilon^T(t)]E[\varepsilon(t)\varepsilon^T(t)]^{-1}\varepsilon(t) \quad (59)$$

$$\hat{y}^*(t)=E[\tilde{y}(t)\,|\,\varepsilon(t)]=E[\tilde{y}(t)\varepsilon^T(t)]E[\varepsilon(t)\varepsilon^T(t)]^{-1}\varepsilon(t) \quad (60)$$

then, since $\tilde{u}(t)$ and $\tilde{y}(t)$ are uncorrelated with $x(t)-x(t\,|\,t-1)$ it is easy to show that

$$E[\tilde{u}(t)\varepsilon^T(t)]=E[\tilde{u}(t)n_y^T(t)]=\tilde{Q}_{uy}(t)-\tilde{Q}_u(t)D^T(t) \quad (61)$$

$$E[\tilde{y}(t)\varepsilon^T(t)]=E[\tilde{y}(t)n_y^T(t)]=\tilde{Q}_y(t)-\tilde{Q}_{uy}^T(t)D^T(t) \quad (62)$$

hence

$$\tilde{u}^*(t)=[\tilde{Q}_{uy}(t)-\tilde{Q}_u(t)D^T(t)]Q_\varepsilon^{-1}(t)\varepsilon(t) \quad (63)$$

$$\tilde{y}^*(t)=[\tilde{Q}_y(t)-\tilde{Q}_{uy}^T(t)D^T(t)]Q_\varepsilon^{-1}(t)\varepsilon(t) \quad (64)$$

Finally, by using (11) and (12), the minimal variance estimates of $\hat{u}(t)$ and $\hat{y}(t)$ can be written in the form

$$\hat{u}^*(t)=u(t)-[\tilde{Q}_{uy}(t)-\tilde{Q}_u(t)D^T(t)]Q_\varepsilon^{-1}(t)\varepsilon(t) \quad (65)$$

$$\hat{y}^*(t)=y(t)-[\tilde{Q}_y(t)-\tilde{Q}_{uy}^T(t)D^T(t)]Q_\varepsilon^{-1}(t)\varepsilon(t) \quad (66)$$

These equations constitute the solution of the optimal filtering problem. And the Theorem 2 is obtained as follow.

Theorem 2. An alternative expression for the filtered output $\hat{y}^*(t)$ is given by

$$\hat{y}^*(t)=h(x(t\,|\,t))+D(t)\hat{u}^*(t) \quad (67)$$

where $x(t\,|\,t)$ is the filtered state given by

$$x(t\,|\,t)=x(t\,|\,t-1)+X(t)Wh^T(X(t),t)Q_\varepsilon^{-1}(t)\varepsilon(t) \quad (68)$$

Proof. To verify the equivalence between (66) and (67) first replace $x(t\,|\,t)$ and $\hat{u}^*(t)$ in (67) with (65) and (68) in order to obtain

$$\begin{aligned}\hat{y}^*(t)=&h(x(t\,|\,t-1))+D(t)u(t)+[h(X(t),t)Wh^{\mathrm{T}}(X(t),t)\\&-D(t)\tilde{Q}_{uy}(t)+D(t)\tilde{Q}_u(t)D^T(t)]Q_\varepsilon^{-1}(t)\varepsilon(t)\end{aligned} \quad (69)$$

Now, by using (50) and (58), it can be shown that

$$\begin{aligned}\hat{y}^*(t)=&h(x(t\,|\,t-1))+D(t)u(t)+\varepsilon(t)\\&-[\tilde{Q}_y(t)-\tilde{Q}_{uy}^T(t)D^T(t)]Q_\varepsilon^{-1}(t)\varepsilon(t)\end{aligned} \quad (70)$$

Finally, by replacing $\varepsilon(t)$ with $z(t)-h(x(t\,|\,t-1))$ and $z(t)$ with $y(t)-D(t)u(t)$, (66) can be obtained in a straightforward way. □

Remark 1. When the system is purely dynamic, i.e., $D(t)=0,$, $\forall t$ the optimal estimate of $\hat{y}(t)$ is given by $h(x(t\,|\,t))$, as in UKF. Note also that if $\tilde{Q}_y(t)=\tilde{Q}_{uy}^T(t)D^T(t),\forall t$, the optimal estimate of $\hat{y}(t)$ coincides with its observation $y(t)$. Finally, when $D(t)=0,\forall t$ and $\tilde{Q}_{uy}(t)=0,\forall t$, the optimal estimate of the noiseless input $\hat{u}(t)$ coincides with its observation $u(t)$.

IV. Evaluations of the Expected Performance

The purpose of this section is to develop an expression for the expected performance of the filter (65) and (66), i.e., for the covariance matrices of the estimation errors

$$e_u(t)=\hat{u}(t)-\hat{u}^*(t)=-\tilde{u}(t)+[\tilde{Q}_{uy}(t)-\tilde{Q}_u(t)D^T(t)]Q_\varepsilon^{-1}(t)\varepsilon(t) \quad (71)$$

$$e_y(t)=\hat{y}(t)-\hat{y}^*(t)=-\tilde{y}(t)+[\tilde{Q}_y(t)-\tilde{Q}_{uy}^T(t)D^T(t)]Q_\varepsilon^{-1}(t)\varepsilon(t) \quad (72)$$

For simplicity, rewrite (71) and (72) as

$$e_u(t)=-\tilde{u}(t)+H_u(t)Q_\varepsilon^{-1}(t)\varepsilon(t) \quad (73)$$

$$e_y(t)=-\tilde{y}(t)+H_y(t)Q_\varepsilon^{-1}(t)\varepsilon(t) \quad (74)$$

where

$$H_u(t)=\tilde{Q}_{uy}(t)-\tilde{Q}_u(t)D^T(t) \quad (75)$$

$$H_y(t)=\tilde{Q}_y(t)-\tilde{Q}_{uy}^T(t)D^T(t) \quad (76)$$

By taking into account (58), (61), and (62) it is easy to show that

$$P_u(t)=E[e_u(t)e_u^T(t)]=\tilde{Q}_u(t)-H_u(t)Q_\varepsilon^{-1}(t)H_u^T(t) \quad (77)$$

$$P_y(t)=E[e_y(t)e_y^T(t)]=\tilde{Q}_y(t)-H_y(t)Q_\varepsilon^{-1}(t)H_y^T(t) \quad (78)$$

Remark 2. Consider now a time-invariant system described by the matrices $\partial f(x(t))/\partial x$, B, G, $\partial h(x(t))/\partial x$, D, whose state, input, and output noise are stationary stochastic processes characterized by covariance and cross-covariance matrices $Q_w(t)$, $\tilde{Q}_u(t)$, $\tilde{Q}_y(t)$ and $\tilde{Q}_{uy}(t)$. In this case, when the pair $(\partial f(x(t))/\partial x-SR^{-1}\partial h(x(t))/\partial x,\ \partial h(x(t))/\partial x)$ is detectable and the pair $(\partial f(x(t))/\partial x-SR^{-1}\partial h(x(t))/\partial x,\ \bar{Q})$ is stabilizable ($\bar{Q}\bar{Q}^T=Q-SR^{-1}S^T$); $P(t+1\,|\,t)$ converges to the unique solution P of the algebraic Riccati equation for $t\to\infty$.

$$\begin{aligned}P=&XWf^{\mathrm{T}}(X)+f(X)WX^{\mathrm{T}}+Q\\&-[XWh^{\mathrm{T}}(X)+S][h(X)Wh^{\mathrm{T}}(X)+R]^{-1}[XWh^{\mathrm{T}}(X)+S]^{\mathrm{T}}\end{aligned} \quad (79)$$

so that $P_u(t)$ and $P_y(t)$ converge to the matrices P_u and P_y given by

$$P_u=\lim_{t\to\infty}P_u(t)=\tilde{Q}_u-H_u[h(X)Wh^{\mathrm{T}}(X)+R]^{-1}H_u^T \quad (80)$$

$$P_y=\lim_{t\to\infty}P_y(t)=\tilde{Q}_y-H_y[h(X)Wh^{\mathrm{T}}(X)+R]^{-1}H_y^T \quad (81)$$

with

$$H_u(t)=\tilde{Q}_{uy}(t)-\tilde{Q}_u(t)D^T(t) \quad (82)$$

$$H_y(t)=\tilde{Q}_y(t)-\tilde{Q}_{uy}^T(t)D^T(t) \quad (83)$$

Moreover, the filter (54)-(56) is asymptotically stable for $t\to\infty$.

V. Numerical Simulations

The results obtained in previous sections have been numerically verified by means of a 100 runs Monte Carlo simulation performed on the two inputs-two outputs time-invariant system described by the following matrices.

$$f(x(t),t)=\begin{bmatrix}x_2(t)\\-x_1(t)+(x_1^2(t)+x_2^2(t)-1)x_3(t)\\x_1^2(t)+x_2^2(t)-x_3(t)\end{bmatrix}$$

$$B=\begin{bmatrix}0.8&-0.8\\0.17&-0.37\\1.09&1.1\end{bmatrix}\quad G=\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}$$

$$h(x(t),t)=\begin{bmatrix}\exp(c-x_1(t))\\\exp(c+x_3(t))\end{bmatrix}\quad D=\begin{bmatrix}1.7&1.5\\0.51&-1\end{bmatrix}$$

The number of samples is 500. The input sequences $\hat{u}_1(\bullet)$ and $\hat{u}_2(\bullet)$ have unit variance and are shown in Figs. 1 and 2 (last 200 samples). In every run, the noise sequences $w(\bullet)$,

$\tilde{u}(\bullet)$ and $\tilde{y}(\bullet)$ are characterized by the following covariance and cross-covariance matrices.

$$Q_w = \begin{bmatrix} 0.56 & 0.26 & 0.45 \\ 0.26 & 0.17 & 0.23 \\ 0.45 & 0.23 & 0.39 \end{bmatrix} \qquad \tilde{Q}_u = \begin{bmatrix} 0.12 & 0.15 \\ 0.15 & 0.25 \end{bmatrix}$$

$$\tilde{Q}_y = \begin{bmatrix} 1.3 & 1.7 \\ 1.7 & 2.4 \end{bmatrix} \qquad \tilde{Q}_{uy} = \begin{bmatrix} 0.38 & 0.51 \\ 0.46 & 0.7 \end{bmatrix}$$

The initial state x_0 is a random vector and (54) and (56) have been initialized as $x(0|-1) = 0$ and $P(0|-1) = I_n$.

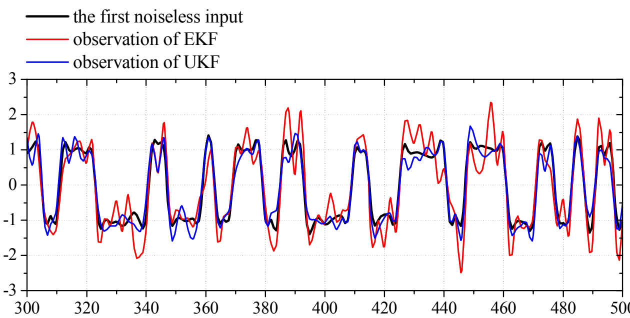

Fig.1. Comparison between the first noiseless input and its observation

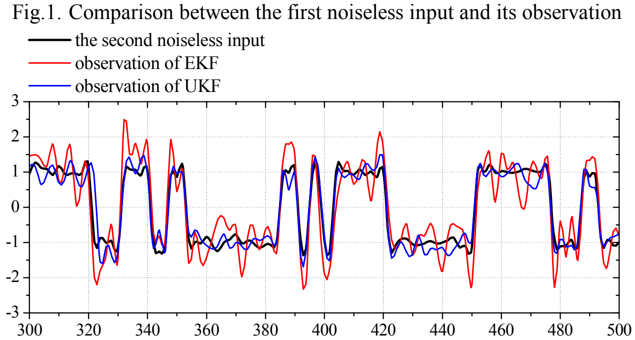

Fig.2. Comparison between the second noiseless input and its observation

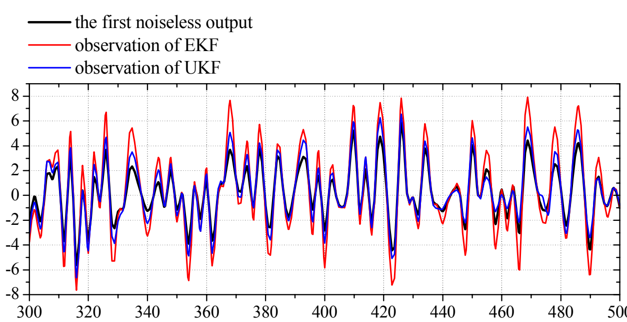

Fig.3. Comparison between the first noiseless output and its observation

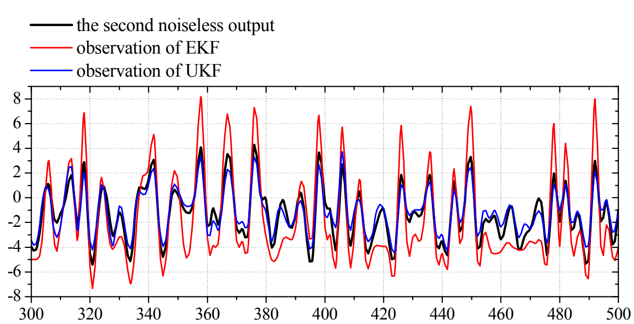

Fig.4. Comparison between the second noiseless output and its observation

Figs. 1-4 report the noiseless inputs and outputs, as well as the associated noisy observations in a typical case of the Monte Carlo simulation (last 200 samples). The effectiveness of the filter can be observed, in the same typical case, in Figs. 5-8, where the noiseless inputs and outputs are compared with the corresponding filtered quantities. The performance of the UKF is compared with the EKF in the simulation. It is shown that the UKF has better accuracy than the EKF in the observations and the optimal estimations.

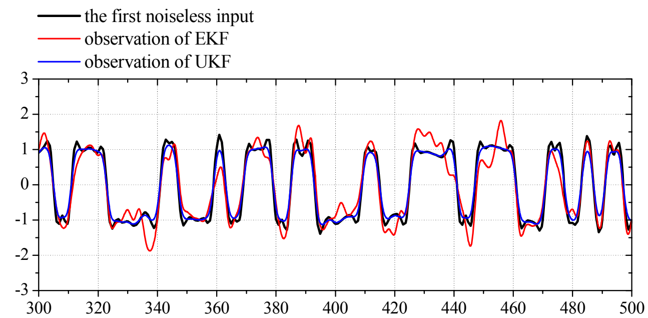

Fig.5. Comparison between the first noiseless input and its optimal estimate

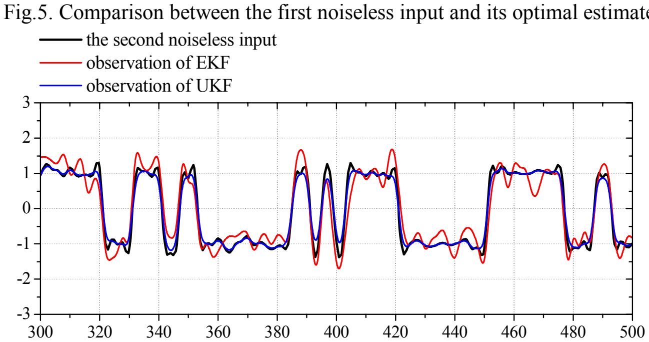

Fig.6. Comparison between the second noiseless input and its optimal estimate

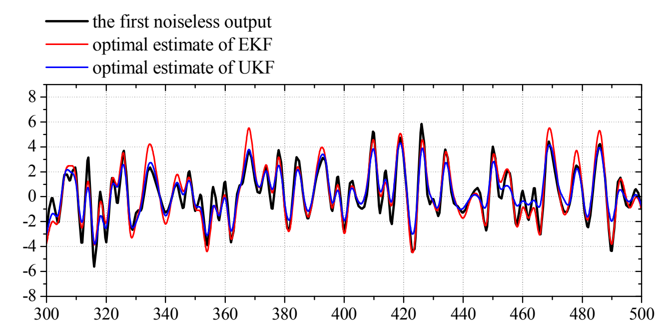

Fig.7. Comparison between the first noiseless output and its optimal estimate

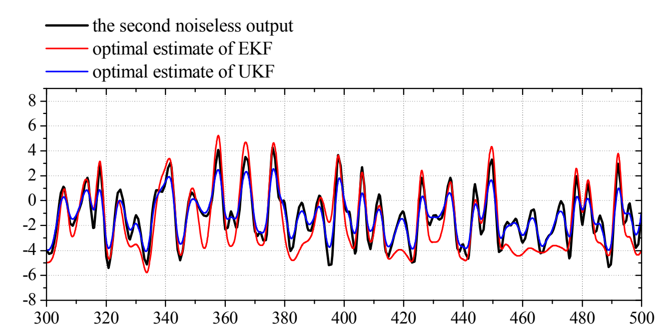

Fig.8. Comparison between the second noiseless output and its optimal estimate

In addition, the covariance matrices of the estimation errors P_u and P_y are obtained by means of the asymptotic relations (80) and (81), while the means of the actual values obtained in the 100 runs of the Monte Carlo simulation $\bar{P}_u$ and $\bar{P}_y$ are compared with P_u and P_y. The difference value between P_u, P_y and $\bar{P}_u$, $\bar{P}_y$ are calculated from ten independent trials and summarized. It verified the better accuracy of the UKF compared with the EKF. And the theoretical results are thus confirmed in a very accurate way by the numerical simulation.

$$(\bar{P}_u - P_u)_{\text{EKF}} = \begin{bmatrix} \pm 0.004665 & \pm 0.0071125 \\ \pm 0.0071125 & \pm 0.012225 \end{bmatrix}$$

$$(\bar{P}_u - P_u)_{\text{UKF}} = \begin{bmatrix} \pm 0.0019975 & \pm 0.00373 \\ \pm 0.00373 & \pm 0.004565 \end{bmatrix}$$

$$(\bar{P}_y - P_y)_{\text{EKF}} = \begin{bmatrix} \pm 0.0409575 & \pm 0.043275 \\ \pm 0.043275 & \pm 0.05693 \end{bmatrix}$$

$$(\bar{P}_y - P_y)_{\text{UKF}} = \begin{bmatrix} \pm 0.021645 & \pm 0.0347925 \\ \pm 0.0347925 & \pm 0.03535 \end{bmatrix}$$

From the example simulations, we can confirm that the proposed filtering and the unscented Kalman filter, as well as the related results are all effective.

VI. Conclusion

In this note, the extension of the stochastic context of unscented Kalman filtering to the presence of additive noise on input observations has been considered. The unscented Kalman filter has then been used to solve the problem of optimal estimation of noise-corrupted input and output sequences. A unified context for both the unscented Kalman filter and the errors-in-variables filter has thus been established. A Monte Carlo simulation has shown the effectiveness of the UKF and the excellent agreement between expected and observed performances.

References

[1] Gobbo D D, Napolitano M, Famouri P, Innocenti M. Experimental application of extended Kalman filtering for sensor validation. *IEEE Transactions on Control System Technology*, 2001, 9(2): 376-380.

[2] Goodwin G C, Sin K S. *Adaptive filtering prediction and control*. Englewood Cliffs, NJ: Prentice-Hall, 1984.

[3] La Scala B F, Bitmead R R, Quinn B G. An extended Kalman filter frequency tracker for high-noise environments. *IEEE Transactions on Signal Processing*, 1996(44): 431-434.

[4] Ljung L. Asymptotic behavior of the extended Kalman filter as a parameter estimator for linear systems. *IEEE Transactions on Automatic Control*, 1979, AC-24: 36-50.

[5] Alspach D L, Sorenson H W. Non-linear Bayesian estimation using Gaussian sum approximation. *IEEE Transactions on Automatic Control*, 1972, AC-20: 439-447.

[6] Kramer S C, Sorenson H W. Recursive Bayesian estimation using piece-wise constant approximations. *Automatica*, 1988(24): 789-801.

[7] Arulampalam M S, Maskell S, Gordon N, Clapp T. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian Tracking. *IEEE Transactions on Signal Processing*, 2002, 50(2): 174-188.

[8] Gordon N J, Salmond D J, Smith A F M. Novel approach to nonlinear/nonlinear-Gaussian Bayesian state estimation. *Proceedings of the Institute of Electrical Engineers*, F, 1993, 140(2): 107-113.

[9] Hall D L, Llinas J. *Handbook of multisensor data fusion*. Boca Raton, FL: CRC Press, 2001.

[10] Julier S J, Uhlmann J K, Durrant-Whyte H F. A new approach for filtering nonlinear systems. *Proceeding of the American Control Conference*, *Washington*: *Seattle*, 1995: 1628-1632.

[11] Julier S J, Uhlmann J K. A general method for approximating nonlinear transformations of probability distributions [EB/OL]. http: / / www. robots. ox. ac. uk/～siju/ work/ publications/ Unscented. zip, 1997-09-27.

[12] Julier S J, Uhlmann J K. A consistent, debiased method for converting between polar and Cartesian coordinate systems. *Proceeding of Aero Sense: 11th Int. Symposium on Aerospace/ Defense Sensing, Simulation and Controls*, *Orlando*, 1997: 110-121.

[13] Julier S J, Uhlmann J K, Durrant-Whyte H F. A new approach for the nonlinear transformation of means and covariances in filters and estimators. *IEEE Transaction on Automatic Control*, 2000, 45(3): 477-482.

[14] Lefebvre T, Bruyninckx H, De Schutter J. Comment on "a new method for the nonlinear transformation of means and covariances in filters and estimators". *IEEE Transaction on Automatic Control*, 2002, 47(8): 1406-1408.

[15] Ristic B, Farina A, Benvenuti D, Arulampalam M S. Performance bounds and comparison of nonlinear filters for tracking a ballistic object on re-entry. *IEE proceedings of the Radar Sonar Navigation*, 2003, 150(2): 65-70.

[16] Julier S J. The scaled unscented transformation. *Proceedings of the American Control Conference*, *Anchorage*, 2002: 4555- 4559.

[17] Wan E A, R Van Der Merwe. The unscented Kalman filter for nonlinear estimation. *Adaptive Systems for Signal Processing, Communications, and Control Symposium*, 2000: 153-158.

[18] Guidorzi R, Diversi R, Soverini U. Optimal errors-in-variables filtering. *Automatica*, 2003, 39(2): 281-289.

[19] Diversi R, Guidorzi R, Soverini U. Algorithms for optimal errors-in-variables filtering. *Letters in System Control*, 2003, 48(1): 1-13.

[20] Roorda B, Heij C, Global total least squares modeling of multivariable time series. *IEEE Transactions on Automatic Control*, 1995, 40(1): 50-63.

[21] Diversi R, Guidorzi R, Soverini U. Kalman filtering in symmetrical noise environments. *Proceeding of 11th IEEE Mediterranean Conference Control and Automation*, *Rhodes*: *Greece*, 2003.

[22] Diversi R, Guidorzi R, Soverini U. Kalman Filtering in Extended Noise Environments. *IEEE Transaction on Automatic Control*, 2005, 50(9): 1396-1402.