The Fifteenth International Symposium on Artificial Life and Robotics 2010 (AROB 15th '10),
B-Con Plaza, Beppu,Oita, Japan, February 4-6, 2010

# The Number of Unscented Transformations and the Effect of Noise Estimates in an Unscented Kalman Filtering Problem

Saifudin bin Razali, Keigo Watanabe, Shoichi Maeyama, and Kiyotaka Izumi†

Division of Industrial Innovation Sciences,
Graduate School of Natural Science and Technology, Okayama University
3-1-1 Tsushima-naka, Kita-ku, Okayama 700-8530, Japan
*Tel & Fax: +81-86-251-8064*
*E-mail: saifudin@usm.sys.okayama-u.ac.jp, {watanabe, maeyama}@sys.okayama-u.ac.jp*
† Department of Advanced Systems Control Engineering,
Graduate School of Science and Engineering, Saga University
1-Honjomachi, Saga 840-8502, Japan
*Tel: +81-952-28-8696, Fax: +81-952-28-8587*
*E-mail: izumi@me.saga-u.ac.jp*

***Abstract***: The unscented transformation is known as a technique to firstly generate a set of $2n+1$ sigma points and their weights, and secondly to propagate each sigma point value through a nonlinear function, where $n$ denotes the dimension of the random state variable. Note however that there are two cases in a discrete-time filtering problem: one is the case where such a transformation is applied two times to the nonlinear model function and the nonlinear measurement function separately by using different mean and covariance, whereas the other is the case where such a transformation is basically applied to the nonlinear model function and the same sigma point values are only propagated to the nonlinear measurement function. So, we here examine the performance difference between them in a particular estimation problem. In addition, it is sometimes to encounter the case where for an unscented Kalman filter, the original state is augmented with other system and measurement noises simultaneously as if the original state and measurement noises are included in nonlinear functions, even though they are actually to be additive to each model function. Therefore, we further check how much the performance improvement or degradation is, compared to the case where there is no inconsistency in the model assumptions.

***Keywords***: Nonlinear system, Unscented transformation, Sigma point.

## I. INTRODUCTION

The nonlinear filtering problem has been extensively studied and various methods are provided in literature. Among them, the most useful ones are the Extended Kalman Filter (EKF), the Ensemble Kalman Filter (EnKF), the Unscented Kalman Filter (UKF), and the Particle Filter (PF). Historically, the EKF is still the most widely adopted approach to solve the nonlinear estimation problem. It is based on the assumption that the nonlinear system dynamics can be accurately modeled by a first-order Taylor series expansion as shown by van de Merwe [1]. The EnKF introduced by Evensen [2] is a reduced rank filter which propagates the states through nonlinearity and updates a relatively small ensemble of samples from which an assumed Gaussian distribution captures the main characteristics in the uncertainty. The PF also uses a sampling approach to estimate the higher order moments of the posterior probability distribution by propagating and updating a number of particles, but without assuming Gaussian statistics as explained by Arulampalam et al. [3]. The UKF, which is a derivative-free alternative to EKF, overcomes the differentiation problem by using a deterministic sampling approach demonstrated by Julier and Uhlmann [4] and Wan and van der Merwe [5].

The state distribution is represented using a minimal set of carefully chosen sample points, called sigma points. This technique is used to linearize a nonlinear function of a random variable through a linear regression between $n$ points drawn from the prior distribution of the random variable. Since we are considering the spread of the random variable during linearization, the technique tends to be more accurate than Taylor series linearization used in the EKF, particularly in the presence of strong nonlinearities as proved by van de Merwe [1]. The $2n+1$ sigma points, are chosen based on a square-root decomposition of the prior covariance, where $n$ is the state dimension. These sigma points are propagated through the true nonlinear function, without approximation, and then a weighted mean and covariance is taken. This approach results in approximations that are accurate to the third order (Taylor series expansion) for Gaussian inputs for all nonlinearities. For non-Gaussian inputs, approximations are accurate to at least the second-order as mentioned by Julier and Uhlmann [4], whereas the linearization approach of the EKF results only in first order accuracy.

However, there are two cases in a discrete-time filtering problem: one is the case where such a transformation is applied two times to the nonlinear model function and the nonlinear measurement function separately by using different mean and covariance, whereas the other is the case where such a transformation is basically applied to the nonlinear model function and the same sigma point values are only propagated to the nonlinear measurement function. This paper examines the performance difference between them in a particular estimation problem. In addition, it is sometimes to encounter the case where for an unscented Kalman filter, the original state is augmented with other system and measurement noises simultaneously as if the original state and measurement noises are included in nonlinear functions, even though they are actually to be additive to each model function. Therefore, we further check how much the performance improvement or degradation is, compared to the case where there is no inconsistency in the model assumptions.

The structure of this paper as follows: In Section II we describe the problem statement for applying a difference set of sigma points of Kalman Filter to propagate nonlinear model function and update nonlinear measurement function. An example is presented in Section III, in which the effect of a different set of sigma points applied to nonlinear model and measurement functions is discussed. The paper is concluded in Section IV.

## II. UNSCENTED KALMAN FILTER

The basic framework for the UKF involves the estimation of the state for discrete-time nonlinear dynamic system:

$$x_{k+1} = f\left(x_k, u_k\right) + w_k \tag{1}$$

$$y_k = h\left(x_k, u_k\right) + v_k \tag{2}$$

where $x_k$ represents the $n$-dimensional unobserved state of the system, $u_k$ is a known exogenous input, and $y_k$ is the $m$-dimensional observed measurement signal. The system noise is represented by $w_k$ and the observation noise is given by $v_k$, where these noises are uncorrelated each other, but with their covariances $Q_k$ and $R_k$, respectively. The standard UKF implementation uses the following weight definitions of $\{w^{(i)}\}$, where $w_0^{(m)} = \frac{\lambda}{\lambda+n}$, $w_0^{(c)} = w_0^{(m)} + (1-\alpha^2+\beta)$ and $w_i^{(c)} = w_i^{(m)} = \frac{1}{2(n+\lambda)}$. $\lambda = \alpha^2 + (n+\kappa) - n$ and $\gamma = \sqrt{n+\lambda}$ are scaling factors. The constant $\alpha$ determines the spread of the sigma points around its mean and is usually set ranging from $10^{-4} \le \alpha \le 1$. $\kappa$ is a secondary scaling parameter (set to 0 for state estimation). The $\beta$ is used to incorporate prior knowledge of the distribution of $x_k$ and $\beta = 2$ for Gaussian distributions.

We now evaluate this nonlinear system using two different approaches of transformation of sigma point. In the first case, the transformation of sigma points is applied two times to the nonlinear model function and the nonlinear measurement function separately by using different mean and covariance, whereas the other is the case where such a transformation is basically applied to the nonlinear model function and the same sigma point values are only propagated to the nonlinear measurement function. In addition, we investigate the case where for an unscented Kalman filter, the original state is augmented with other system and measurement noises simultaneously as if the original state and measurement noises are included in nonlinear functions, even though they are actually to be additive to each model function.

### 1. Two times transformation of sigma points

In this case, generated sigma points are propagated two times through a nonlinear model function as well as nonlinear measurement function, respectively. The algorithm for this case can be summarized as Table 1.

Table I. UKF algorithm with two-times transformation of sigma points

$$\chi_{k-1} = [\hat{x}_{k-1} \cdots \hat{x}_{k-1}] + \gamma[0 \quad \sqrt{P_{x_{k-1}}} \quad -\sqrt{P_{x_{k-1}}}] \tag{3}$$

$$\chi_{i,k}^{*} = f(\chi_{i,k-1}, u_{k-1}) \tag{4}$$

$$\hat{x}_k^- = \sum_{i=0}^{2n} w_i^{(m)} \chi_{i,k}^{*} \tag{5}$$

$$P_{x_k}^- = Q_{k-1} + \sum_{i=0}^{2n} w_i^{(c)} (\chi_{i,k}^{*} - \hat{x}_k^-)(\chi_{i,k}^{*} - \hat{x}_k^-)^T \tag{6}$$

$$\chi_k^- = [\hat{x}_k^- \cdots \hat{x}_k^-] + \gamma[0 \quad \sqrt{P_{x_k}^-} \quad -\sqrt{P_{x_k}^-}] \tag{7}$$

$$\mathcal{Y}_{i,k} = h(\chi_{i,k}^-, u_k) \tag{8}$$

$$\hat{y}_k^- = \sum_{i=0}^{2n} w_i^{(m)} \mathcal{Y}_{i,k} \tag{9}$$

$$P_{\hat{y}_k} = R_k + \sum_{i=0}^{2n} w_i^{(c)} (\mathcal{Y}_{i,k} - \hat{y}_k^-)(\mathcal{Y}_{i,k} - \hat{y}_k^-)^T \tag{10}$$

$$P_{x_k y_k} = \sum_{i=0}^{2n} w_i^{(c)} (\chi_{i,k}^- - \hat{x}_k^-)(\mathcal{Y}_{i,k} - \hat{y}_k^-)^T \tag{11}$$

$$K_k = P_{x_k y_k} P_{\hat{y}_k}^{-1} \tag{12}$$

$$\hat{x}_k = \hat{x}_k^- + K_k(y_k - \hat{y}_k^-) \tag{13}$$

$$P_{x_k} = P_{x_k}^- - K_k P_{\hat{y}_k} K_k^T \tag{14}$$

### 2. Single transformation of sigma points

In this case, generated same sigma points are propagated through a nonlinear model function as well as nonlinear measurement function [8]. The algorithm of this case can be described as Table 2.

Table 2. UKF algorithm with single transformation of sigma points

$$\chi_{k-1} = [\hat{x}_{k-1} \cdots \hat{x}_{k-1}] + \gamma[0 \quad \sqrt{P_{x_{k-1}}} \quad -\sqrt{P_{x_{k-1}}}] \tag{15}$$

$$\chi_{i,k} = f(\chi_{i,k-1}, u_{k-1}) \tag{16}$$

$$\hat{x}_k^- = \sum_{i=0}^{2n} w_i^{(m)} \chi_{i,k} \tag{17}$$

$$P_{x_k}^- = Q_{k-1} + \sum_{i=0}^{2n} w_i^{(c)} (\chi_{i,k} - \hat{x}_k^-)(\chi_{i,k} - \hat{x}_k^-)^T \quad (18)$$

$$\mathcal{Y}_{i,k} = h(\chi_{i,k}, u_k) \quad (19)$$

$$\hat{y}_k^- = \sum_{i=0}^{2n} w_i^{(m)} \mathcal{Y}_{i,k} \quad (20)$$

$$P_{\hat{y}_k} = R_k + \sum_{i=0}^{2n} w_i^{(c)} (\mathcal{Y}_{i,k} - \hat{y}_k^-)(\mathcal{Y}_{i,k} - \hat{y}_k^-)^T \quad (21)$$

$$P_{x_k y_k} = \sum_{i=0}^{2n} w_i^{(c)} (\chi_{i,k} - \hat{x}_k^-)(\mathcal{Y}_{i,k} - \hat{y}_k^-)^T \quad (22)$$

$$K_k = P_{x_k y_k} P_{\hat{y}_k}^{-1} \quad (23)$$

$$\hat{x}_k = \hat{x}_k^- + K_k (y_k - \hat{y}_k^-) \quad (24)$$

$$P_{x_k} = P_{x_k}^- - K_k P_{\hat{y}_k} K_k^T \quad (25)$$

### 3. State augmented case

In this section, we investigate the effect when the original state is augmented with other system and measurement noises simultaneously [9] as if the original state and measurement noises are included in nonlinear functions, even though they are actually to be additive to each model function. Note however that a single set of generated sigma points are used to evaluate the result. The algorithm can be summarized as in Table 3.

Table 3. UKF algorithm with an augmented state vector

$$x_k^a = \begin{bmatrix} x_k^T & w_k^T & v_k^T \end{bmatrix}^T, \; P_{x_k}^a = \mathrm{diag}\,(P_{x_k}, Q_k, R_k) \quad (26)$$

$$\chi_{k-1}^a = [\hat{x}_{k-1}^a \cdots \hat{x}_{k-1}^a] + \gamma [0 \quad \sqrt{P_{x_{k-1}}^a} \quad -\sqrt{P_{x_{k-1}}^a}] \quad (27)$$

$$\chi_{i,k}^x = f(\chi_{i,k-1}^x, u_{k-1}) + \chi_{i,k-1}^w \quad (28)$$

$$\hat{x}_k^- = \sum_{i=0}^{2L} w_i^{(m)} \chi_{i,k}^x \quad (29)$$

$$P_{x_k}^- = \sum_{i=0}^{2L} w_i^{(c)} (\chi_{i,k}^x - \hat{x}_k^-)(\chi_{i,k}^x - \hat{x}_k^-)^T \quad (30)$$

$$\mathcal{Y}_{i,k} = h(\chi_{i,k}^x, u_k) + \chi_{i,k-1}^v \quad (31)$$

$$\hat{y}_k^- = \sum_{i=0}^{2L} w_i^{(m)} \mathcal{Y}_{i,k} \quad (32)$$

$$P_{\hat{y}_k} = \sum_{i=0}^{2L} w_i^{(c)} (\mathcal{Y}_{i,k} - \hat{y}_k^-)(\mathcal{Y}_{i,k} - \hat{y}_k^-)^T \quad (33)$$

$$P_{x_k y_k} = \sum_{i=0}^{2L} w_i^{(c)} (\chi_{i,k}^x - \hat{x}_k^-)(\mathcal{Y}_{i,k} - \hat{y}_k^-)^T \quad (34)$$

$$K_k = P_{x_k y_k} P_{\hat{y}_k}^{-1} \quad (35)$$

$$\hat{x}_k = \hat{x}_k^- + K_k (y_k - \hat{y}_k^-) \quad (36)$$

$$P_{x_k} = P_{x_k}^- - K_k P_{\hat{y}_k} K_k^T \quad (37)$$

Note here that $L$ denotes $2n+m$, $\chi_{i,k-1}^x$, $\chi_{i,k-1}^w$, and $\chi_{i,k-1}^v$ denote the $i$-th column of $\chi_{k-1}^x$, $\chi_{k-1}^w$, and $\chi_{k-1}^v$ which are also the components of $\chi_{k-1}^a = [(\chi_{k-1}^x)^T \quad (\chi_{k-1}^w)^T \quad (\chi_{k-1}^v)^T]^T$. Note also that $n$ in the weights appearing in the former algorithms should be replaced by $L$ for the present algorithm.

## III. AN EXAMPLE APPLICATION

In this section we consider the problem that a vehicle enters the atmosphere at high altitude and at a very high speed. The position of the body is to be tracked by a radar which accurately measures range and bearing. This type of problem has been identified by Mehra [10], Austin and Leondes [11] and Athans et al. [12] as being particularly stressful for filters and trackers because of the strong nonlinearities exhibited by the forces which act on the vehicle. There are three types of forces which act: i.e., (1) the most dominant is aerodynamic drag, which is a function of vehicle speed and has a substantial nonlinear variation in altitude; (2) gravity force which accelerates the vehicle towards the centre of the earth; and (3) random buffeting force terms. The effect of these forces gives a trajectory of the form that initially the trajectory is almost ballistic but as the density of the atmosphere increases, drag effects become important and the vehicle rapidly decelerates until its motion is almost vertical. The tracking problem is made more difficult by the fact that the drag properties of the vehicle might be only very crudely known.

This tracking system should be able to track an object which experiences a set of complicated, highly nonlinear forces. These depend on the current position and velocity of the vehicle as well as on certain characteristics which are not known precisely. The filter's state space consists of the position of the body ($x_1$ and $x_2$), its velocity ($x_3$ and $x_4$) and a parameter of its aerodynamic properties ($x_5$). The vehicle state dynamics are

$$\begin{aligned} \dot{x}_2(k) &= x_4(k) \\ \dot{x}_2(k) &= x_4(k) \\ \dot{x}_3(k) &= D(k)\,x_3(k) + G(k)\,x_1(k) + v_1(k) \\ \dot{x}_4(k) &= D(k)\,x_4(k) + G(k)\,x_2(k) + v_2(k) \\ \dot{x}_5(k) &= x_3(k) \end{aligned} \quad (38)$$

where $D(k)$ is the drag-related force term, $G(k)$ is the gravity-related force term and $v(k)$ are the process noise terms. Defining $R(k) = \sqrt{x_1^2(k) + x_2^2(k)}$ as the distance from the center of the Earth and $V(k) = \sqrt{x_3^2(k) + x_4^2(k)}$ as absolute vehicle speed then the drag and gravitational terms are $D(k) = -\beta(k) \exp\left\{\frac{R_0 - R(k)}{H_0}\right\} V(k)$, $G(k) = -\frac{G m_0}{r^3(k)}$ and $\beta(k) = \beta_0 \exp x_5(k)$.

For this example, the parameter values are $\beta_0 = -0.59783$, $H_0 = 13.406$, $Gm_0 = 3.9860 \times 10^5$ and

$R_0 = 6374$ and reflect typical environmental and vehicle characteristics. The parameterization of the ballistic coefficient, $\beta(k)$, reflects the uncertainty in vehicle characteristics. $\beta_0$ is the ballistic coefficient of a typical vehicle and it is scaled by $\exp x_5(k)$ to ensure that its value is always positive. This is vital for filter stability. The motion of the vehicle is measured by a radar which is located at $(x_r, y_r)$. It is able to measure range $r$ and bearing $\theta$ at a frequency of 10Hz, where

$$r_r(k) = \sqrt{(x_1(k) - x_r)^2 + (x_2(k) - y_r)^2} + w_1(k) \tag{39}$$

$$\theta(k) = \tan^{-1}\left[\frac{x_2(k) - y_r}{x_1(k) - x_r}\right] + w_2(k) \tag{40}$$

$w_1(k)$and $w_2(k)$ are zero mean uncorrelated noise processes with variances of $1\,\mathrm{m}$ and $17\,\mathrm{mrad}$ respectively. The high update rate and extreme accuracy of the sensor means that a large quantity of extremely high quality data is available for the filter. The true initial conditions for the vehicle are

$$x(0) = \begin{pmatrix} 6500.4 \\ 349.14 \\ -1.8093 \\ -6.7967 \\ 0.6932 \end{pmatrix},$$

$$P(0) = \mathrm{diag}(10^{-6}, 10^{-6}, 10^{-6}, 10^{-6}, 0)$$

In other words, the vehicle's coefficient is twice the nominal coefficient.

The vehicle is buffeted by random accelerations,

$$Q(k) = \mathrm{diag}(2.4046 \times 10^{-5}, 2.4046 \times 10^{-5}, 0)$$

The initial conditions assumed by the filter are,

$$x(0) = \begin{pmatrix} 6500.4 \\ 349.14 \\ -1.8093 \\ -6.7967 \\ 0 \end{pmatrix},$$

$$P(0) = \mathrm{diag}(10^{-6}, 10^{-6}, 10^{-6}, 10^{-6}, 1)$$

The filter uses the nominal initial condition and, to offset for the uncertainty, the variance on this initial estimate is 1. Both filters were implemented in discrete time and observations were taken at a frequency of $10\,\mathrm{Hz}$, but the integration step was set to be $50\,\mathrm{ms}$ which meant that two predictions were made per update.

## IV. CONCLUSION

In this paper, we have examined three cases related to sigma points of unscented transformation of nonlinear model and measurement functions. In the first case, the sigma points are applied two times to the nonlinear model and measurement functions separately by using different mean and covariance. In the second case, the transformation is once applied to the nonlinear model function and the same sigma point values are only propagated to the nonlinear measurement function. Lastly, the performance of an unscented Kalman filter was further evaluated for the case where the original state is augmented with other system and measurement noises simultaneously as if the original state and measurement noises are included in nonlinear functions, even though they are actually to be additive to each model function.

## REFERENCES

[1] van der Merwe R (2004), Sigma point Kalman filter for probabilistic inference in dynamic state-space models. PhD Thesis, Oregon Health & Science University, Portland

[2] Evensen G (2003), The ensemble Kalman filter: Theoretical formulation and practical implementation. Ocean Dynamics, 53: 343–367

[3] Arulampalam MS, Maskell S, Gordon N, and Clapp T (2002), A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. IEEE Trans. on Signal Processing, 50(2), 174–188

[4] Julier SJ and Uhlmann JK (2004), Unscented filtering and nonlinear estimation. Proceedings of the IEEE, 92(3): 401–422

[5] Wan E and van der Merwe R (2001), The Unscented Kalman Filter. Wiley Publishing

[6] Terejanu G, Singh T and Scott PD (2007), Unscented Kalman filter/smoother for a CBRN puff-based dispersion model. Proc. of the 10th Int. Conf. on Information Fusion, Quebec City, Quebec, July 9–12

[7] Sarkka S (2010), Continuous-time and continuous-discrete-time unscented Rauch-Tung-Striebel smoothers. Signal Processing, 90: 225–235

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

[9] Wang X and Zhang H (2007), A UPF-UKF framework for SLAM. Proc. of IEEE Int. Conf. on Robotics and Automation, pp. 1664–1669

[10] Mehra RK (1971). A comparison of several nonlinear filters for reentry vehicle tracking. IEEE Trans. on Automatic Control, AC-16(4): 307–319

[11] Austin JW and Leondes CT (1981), Statistically linearized estimation of reentry trajectories. IEEE Trans. on Aerospace and Electronic Systems, AES-17(1): 54–61

[12] Athans M, Wishner RP and Bertolini A (1968), Suboptimal state estimation for continuous-time nonlinear systems from discrete noisy measurements. IEEE Trans. on Automatic Control, AC-13(6): 504–518