UNSCENTED KALMAN FILTER FOR FAULT DETECTION

K. Xiong*, C. W. Chan and H. Y. Zhang***

**School of Automation Science and Electrical Engineering, Beihang University*
Beijing, 100083 China
e-mail: tobelove2001@tom.com

*** Department of Mechanical Engineering, The University of Hong Kong*
Pokfulam Road, Hong Kong, China
e-mail: mechan@hkucc.hku.hk

Abstract: In this paper, the approximation of nonlinear systems using unscented Kalman filter (UKF) is discussed, and the conditions for the convergence of the UKF are derived. The detection of faults from residuals generated by the UKF is presented. As fault detection often reduced to detecting irregularities in the residuals, such as the mean, the local approach, a powerful statistical technique to detect such changes, is used to detect fault from the residuals generated from the UKF. The properties of the proposed method are also presented. To illustrate the performance of the proposed method, it is applied to detect faults in the attitude sensors of a satellite.

Keywords: fault detection, nonlinear filters, Kalman filters, unscented transformation

1. INTRODUCTION

Fault detection for nonlinear system is an important research area attracting considerable interest. Model-based fault detection techniques are popular. For nonlinear systems with additive Gaussian noise, the extended Kalman filters (EKF) are used to generate residuals for fault detection (Gobbo, et al., 2001). However, the EKF suffers from two well-known drawbacks: 1) it is a first- order approximation of the nonlinear system, introducing large errors in the mean and covariance of the state vector, and even divergence of the filter, and 2) the derivation of the Jacobian matrices is nontrivial and can often lead to significant implementation difficulties.

Unscented Kalman filters (UKF) have been proposed recently for estimating the state of nonlinear systems. Another important method in state estimation for nonlinear systems is presented by (Ravn, etal., 2000). The UKF is derived using the unscented transformation (UT) involving a set of carefully chosen sample points, called the sigma points. It has shown that the UKF outperforms the EKF (Julier, *et al*., 2000), as it is able to approximate the posterior mean and covariance of the output variable with a second order accuracy instead of a first order accuracy in the EKF. Further, as it is not necessary to compute the Jacobians or Hessians, it is being widely used in applications, such as target tracking (Julier, *et al*., 2000) and multi-sensor fusion (Hall, *et al*., 2001).

Very few results are available in the literature on the convergence of the UKF. In this paper, the sufficient conditions for the convergence of the UKF are derived based on a new formulation of the unscented transform. Based on this result, fault detection for nonlinear systems is derived using the local approach, a statistical tool that transforms the fault detection problem into one that detects changes in the mean of a Gaussian random variable (Zhang, *et al*., 1998). The performance of the proposed technique is demonstrated by the attitude sensors of a satellite.

The paper is organized as follows. In Section II, a brief review of the UT is presented, followed by the derivation of the UKF, and the conditions for it to converge. In section III, the detection of faults from the residuals generated by the UKF using the local approach is derived, together with the miss-detection. In section IV, the performance of the proposed method is illustrated by applying it to detect attitude sensor faults of a satellite.

2. THE UNSCENTED KALMAN FILTER

2.1 The unscented transform

The UT is a method for calculating the statistics of a random variable that undergoes a nonlinear transformation (Hall, *et al*., 2001). A discrete distribution composed of a number of samples, referred to as the sigma points, are computed based on the known initial mean and covariance of the state variable. Then the nonlinear transformation is applied to each sample. As an example, the UT of a variable with dimension 2 is shown in Fig. 1. The sample mean and covariance of the transformed ensemble can then be used to compute the estimate of the nonlinear transformation of the original distribution. The computed mean and covariance is accurate up to second order (Julier, et al., 2000).

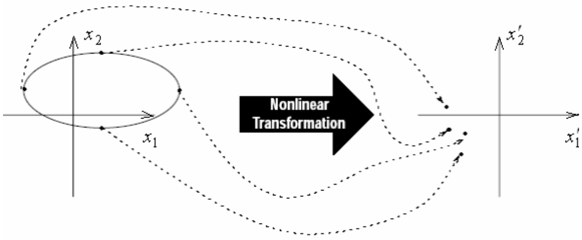

Fig. 1. The unscented transformation

Consider a random variable, $x \in R^L$. Let

$$y = f(x) \in R^L \tag{1}$$

where y is a nonlinear mapping of x, and $f(x)$ is a known nonlinear function. Denote the mean of x by $\bar{x}$, and the covariance by $P_x \in R^{L\times L}$. The statistics of y are computed from x using the sigma points $\chi = \{\chi_i,\ i = 0, 1, \ldots, 2L\}$, as given below.

$$\begin{cases} \chi_i = \bar{x} & i = 0 \\ \chi_i = \bar{x} + a(LP_x)_i^{1/2} & i = 1, \cdots, L \\ \chi_i = \bar{x} - a(LP_x)_{i-L}^{1/2} & i = L+1, \cdots, 2L \end{cases} \tag{2}$$

where a is the spread of the sigma points around $\bar{x}$, and $(LP_x)_i^{1/2}$ is the i^{th} column of the matrix square root of LP_x. The parameter a is to provide an extra degree of freedom to "fine tune" the higher order moments of the approximation, and is usually set to a small positive value. The sample mean and covariance of χ are (Wan and Merwe, 2000):

$$\bar{\chi} = \sum_{i=0}^{2L} w_i \chi_i \ ; \quad P_\chi = \sum_{i=0}^{2L} w_i (\chi_i - \bar{\chi})(\chi_i - \bar{\chi})^T \tag{3}$$

where $\begin{cases} w_i = 1 - \dfrac{1}{a^2} & i = 0 \\ w_i = \dfrac{1}{2La^2} & i = 1, \cdots, 2L \end{cases}$

Let $\gamma_i = f(\chi_i) \in R^L$, for $i = 0, 1, \ldots, 2L$. The mean and covariance for y can be approximated by the sample mean and covariance of γ, as given below.

$$\bar{\gamma} = \sum_{i=0}^{2L} w_i \gamma_i \ ; \quad P_\gamma = \sum_{i=0}^{2L} w_i (\gamma_i - \bar{\gamma})(\gamma_i - \bar{\gamma})^T \tag{4}$$

The properties of the UT are (Julier, *et al*., 2000):

Property 1: The mean and covariance of the set of sigma points given by (2) are identical to that of x.

Property 2: The approximation of the mean and covariance of y by $\bar{\gamma}$ and P_γ has a second order accuracy.

2.2 The unscented Kalman filter (UKF)

Consider the nonlinear system:

$$\begin{cases} x(k) = f(x(k-1)) + w(k) \\ y(k) = h(x(k)) + v(k) \end{cases} \tag{5}$$

where $f(.)$ and $h(.)$ are known nonlinear functions, $x(k)$ is the state vector, $y(k)$ is the output vector, $w(k)$ and $v(k)$ are normally distributed white noise with zero mean and covariance matrices: $E[w(k)w(k)^T] = Q(k)$ and $E[v(k)v(k)^T] = R(k)$. It is assumed that the output can be measured, but not the state. Similar to the Kalman filter, the UKF is obtained by minimizing the mean-squared error. The new state of the system $\hat{x}(k\,|\,k-1)$, the estimated output $\hat{y}(k)$ and the corresponding covariance matrices are computed recursively using the state after applying the UT. The procedure for implementing the UKF is as follows (Wan and Merwe, 2000),

Step 1 Calculate the sigma points from (2),

$$\begin{cases} \chi_i(k-1) = \hat{x}(k-1) & i = 0 \\ \chi_i(k-1) = \hat{x}(k-1) + a[LP(k-1)]_i^{1/2} & i = 1, \cdots, L \\ \chi_i(k-1) = \hat{x}(k-1) - a[LP(k-1)]_{i-L}^{1/2} & i = L+1, \cdots, 2L \end{cases}$$

Step 2 Compute the predicted mean, from (4),

$$\chi_i(k\,|\,k-1) = f(\chi_i(k-1)) \tag{6}$$

$$\hat{x}(k\,|\,k-1) = \sum_{i=0}^{2L} w_i^m \chi_i(k\,|\,k-1) \tag{7}$$

and the predicted covariance from (4),

$$\begin{aligned} & P(k\,|\,k-1) \\ & = \sum_{i=0}^{2L} w_i^c [\chi_i(k\,|\,k-1) - \hat{x}(k\,|\,k-1)][\chi_i(k\,|\,k-1) \\ & \quad - \hat{x}(k\,|\,k-1)] + Q(k) \end{aligned} \tag{8}$$

Step 3 The predicted observation is computed by

$$\gamma_i(k) = h(\chi_i(k\,|\,k-1)), \quad \hat{y}(k) = \sum_{i=0}^{2L} w_i^m \gamma_i(k)$$

the covariance and the cross correlation matrix by,

$$P_{\varepsilon\varepsilon} = \sum_{i=0}^{2L} w_i^c [\gamma_i(k) - \hat{y}(k)][\gamma_i(k) - \hat{y}(k)]^T + R(k)$$

$$P_{xy} = \sum_{i=0}^{2L} w_i^c [\chi_i(k\,|\,k-1) - \hat{x}(k\,|\,k-1)][\gamma_i(k) - \hat{y}(k)]^T$$

And the predicted state is computed using the classical Kalman filter,

$$\begin{cases} \hat{x}(k) = \hat{x}(k\,|\,k-1) + P_{xy} P_{\varepsilon\varepsilon}^{-1} [y(k) - \hat{y}(k)] \\ P(k) = P(k\,|\,k-1) - P_{xy} P_{\varepsilon\varepsilon}^{-1} P_{xy}^T \end{cases} \tag{9}$$

Step 4 Repeat steps 1 to 3 for the next sample.

Since the mean and covariance of $x(k)$ are accurate up to second order, and the same also applied to the computed mean and covariance of $y(k)$, the UKF can predict with a second-order accuracy, but without the need to compute the Jacobian or Hessian matrix. In contrast, the state vector computed by the EKF is only a first-order approximation of the nonlinear system, and hence can only achieve the first-order accuracy. Further, the computational load of UKF is only in the same order as that for the EKF.

2.3 Convergence analysis of the UKF

The convergence analysis of the UKF is derived using an approach similar to that of the EKF (Boutayeb, *et al.*, 1997). Denote the error of the estimated state by

$$\tilde{x}(k) = x(k) - \hat{x}(k) \tag{10}$$

and the prediction error of the state by

$$\tilde{x}(k \mid k-1) = x(k) - \hat{x}(k \mid k-1) \tag{11}$$

Assuming that $w(k)$ is neglectable, expanding $x(k)$ given by (5) by a Taylor Series about $\hat{x}(k-1)$ gives,

$$x(k) = f(\hat{x}(k-1)) + f'(\hat{x}(k-1))\tilde{x}(k-1) + \frac{1}{2} f''(\hat{x}(k-1))\tilde{x}(k-1)^2 + \cdots \tag{12}$$

Similarly, $\hat{x}(k \mid k-1)$ can be expressed as,

$$\hat{x}(k \mid k-1) = f(\hat{x}(k-1)) + \frac{1}{2} f''(\hat{x}(k-1))P(k-1) + \ldots \tag{13}$$

hence, $\tilde{x}(k \mid k-1)$ can be approximated by,

$$\tilde{x}(k \mid k-1) \approx F\tilde{x}(k-1) \tag{14}$$

where $F = f'(\hat{x}(k-1))$. Assuming that $v(k)$ is small, expanding $y(k)$ and $\hat{y}(k)$ about $\hat{x}(k \mid k-1)$ gives,

$$y(k) = h(\hat{x}(k \mid k-1)) + h'(\hat{x}(k \mid k-1))\tilde{x}(k \mid k-1) + \frac{1}{2}h''(\hat{x}(k \mid k-1))\tilde{x}(k \mid k-1)^2 + \cdots$$

$$\hat{y}(k) = h(\hat{x}(k \mid k-1)) + \frac{1}{2}h''(\hat{x}(k \mid k-1))P(k \mid k-1) + \cdots$$

Similarly, $\varepsilon(k) = y(k) - \hat{y}(k)$ can be approximated by,

$$\varepsilon(k) \approx H\tilde{x}(k \mid k-1) \tag{16}$$

where $H = h'(\hat{x}(k \mid k-1))$. In general, $\varepsilon(k)$ is not identically zero, as it is a second order approximation of $\tilde{x}(k \mid k-1)$. Hence, (14) and (16) are modified as,

$$\tilde{x}(k \mid k-1) = \beta(k)F\tilde{x}(k-1) \tag{17}$$

$$\varepsilon(k) = \alpha(k)H\tilde{x}(k \mid k-1) \tag{18}$$

where $\alpha(k) = diag(\alpha_1(k), \alpha_2(k), \ldots, \alpha_N(k))$ and $\beta(k) = diag(\beta_1(k), \beta_2(k), \cdots, \beta_L(k))$ are unknown diagonal matrices. The sufficient condition for the convergence of the UKF is given below.

Theorem 1: Assuming F is a nonsingular matrix, and $\beta(k)$ satisfies the following condition:

$$\left|\lambda_{\max}(\beta(k))\right| < \left[\frac{\lambda_{\min}(F^{-T}F^{-1})}{\lambda_{\max}(U)}\right]^{1/2} \tag{19}$$

where $U = I + H^T\alpha(k)P_{\varepsilon\varepsilon}^{-1}P_{xy}^T + P_{xy}P_{\varepsilon\varepsilon}^{-1}\alpha(k)H + H^T\alpha(k)P_{\varepsilon\varepsilon}^{-1}P_{xy}^T P_{xy}P_{\varepsilon\varepsilon}^{-1}\alpha(k)H$, and $\lambda_{\max}(\cdot)$ is the maximum eigenvalue of the matrix F, then

$$\lim_{k\to\infty} \tilde{x}(k) = 0 \tag{20}$$

Proof: Let $V(k) = \tilde{x}(k)^T \tilde{x}(k)$. From (9), (11), (17) and (18),

$$\begin{aligned} V(k) - V(k-1) &= \tilde{x}(k)^T \tilde{x}(k) - \tilde{x}(k-1)^T \tilde{x}(k-1) \\ &= [\tilde{x}(k \mid k-1) + P_{xy}P_{\varepsilon\varepsilon}^{-1}e(k)]^T[\tilde{x}(k \mid k-1) + P_{xy}P_{\varepsilon\varepsilon}^{-1}e(k)] - \tilde{x}(k-1)^T \tilde{x}(k-1) \\ &= \tilde{x}(k-1)^T[F^T\beta(k)U\beta(k)F - I]\tilde{x}(k-1) \end{aligned} \tag{21}$$

From the Rayleigh-Ritz theorem (Yu and Shi, 2004), for any vector $z \neq 0$,

$$\lambda_{\max}(U) = \max_{x\neq 0}\left(z^T Uz/(z^T z)\right)$$

$$\lambda_{\min}(F^{-T}F^{-1}) = \min_{x\neq 0}\left(z^T F^{-T}F^{-1}z/(z^T z)\right)$$

If assumption (19) hold, then

$$\left(\frac{z^T F^{-T}F^{-1}z}{z^T Uz}\right)^{1/2} \geq \left[\frac{\min_{z\neq 0}\left(z^T F^{-T}F^{-1}z/(z^T z)\right)}{\max_{z\neq 0}\left(z^T Uz/(z^T z)\right)}\right]^{1/2}$$

$$= \left[\frac{\lambda_{\min}(F^{-T}F^{-1})}{\lambda_{\max}(U)}\right]^{1/2} > \left|\lambda_{\max}(\beta(k))\right| \geq \beta_j(k) \tag{22}$$

where the subscript j denotes the j^{th} component of the diagonal matrix $\beta(k)$. From (22), the following inequality is obtained:

$$z^T[\beta_j(k)U\beta_j(k) - F^{-T}F^{-1}]z < 0 \tag{23}$$

As $z \neq 0$, hence

$$\beta(k)U\beta(k) - F^{-T}F^{-1} < 0 \tag{24}$$

For $\tilde{x}(k-1) \neq 0$, it follows that

$$\tilde{x}(k-1)^T[F^T\beta(k)U\beta(k)F - I]\tilde{x}(k-1) < 0 \tag{25}$$

From (25) and (21), $V(k) - V(k-1) < 0$, $V(k)$ is a decreasing sequence, and hence $\lim_{k\to\infty} V(k) = 0$. It follows that $\lim_{k\to\infty} \tilde{x}(k) = 0$. □

The unknown diagonal matrices $\alpha(k)$ and $\beta(k)$ are introduced to evaluate the UT of the state variables that propagates through the nonlinear model. If the magnitude of the eigenvalue of $\beta(k)$ is sufficiently small, the convergence of the UKF is ensured. If the magnitude of α_{ik} are small enough, the convergence of the UKF may be improved in the sense that the domain of $\lambda_{\max}(\beta(k))$ will be enlarged. Indeed, the sufficient conditions (19) mean that if the error introduced by the UT is small enough, $V(k)$ is a decreasing sequence. As $\alpha(k)$ and $\beta(k)$ are unknown factors, sigma points should be chosen properly to decrease the error of the UT so that (20) is fulfilled.

3. FAULT DETECTION BY LOCAL APPROACH

The measurement equation (5) can be rewritten as,

$$y(k) = \bar{h}(x(k)) + \psi(k) + v(k) \tag{26}$$

where $\bar{h}(x(k))$ is a measurement model, and $\psi(k) = h(x(k)) - \bar{h}(x(k))$ is the modelling error. Consider the predicted observation $\hat{y}(k)$ obtained from the UKF. Under normal operating condition, the residual of the UKF is,

$$\varepsilon(k) = y(k) - \hat{y}(k) = \bar{h}(x(k)) + \psi(k) + v(k) - \bar{h}(\hat{x}(k))$$
$$= \psi(k) + \phi(k) + v(k) \quad (27)$$

where $\hat{y}(k) = \bar{h}(\hat{x}(k))$ is the predicted observation and $\phi(k) = \bar{h}(x(k)) - \bar{h}(\hat{x}(k))$ is the estimation error. When there is a sensor fault, the residual becomes,

$$\varepsilon(k) = y(k) + b_f - \hat{y}(k)$$
$$= b_f + \psi(k) + \phi(k) + v(k) \quad (28)$$

where $b_f \neq 0$ is the output arising from the sensor fault. However, faults can only be detected if the term is large compared with the modelling errors and the system noise. For small faults, it is difficult to detect b_f from $\varepsilon(k)$.

The local approach is now applied to the residuals generated by the UKF. In the local approach, the cumulative sum of the residual D^m is computed for a window size of m samples (Wang and Chan, 2002),

$$D^m = \frac{1}{\sqrt{m}} \sum_{k=1}^{m} \varepsilon(k)$$
$$= \frac{1}{\sqrt{m}} \left(\sum_{k=1}^{m} \psi(k) + \sum_{k=1}^{m} \phi(k) + \sum_{k=1}^{m} v(k) \right) \quad (29)$$

Assuming the model is accurate and $\psi(k) = 0$, then the residual D^m can now be approximated by

$$D^m = \frac{1}{\sqrt{m}} \sum_{k=1}^{m} \phi(k) + \frac{1}{\sqrt{m}} \sum_{k=1}^{m} v(k) \quad (30)$$

From *Theorem 1*, $\lim_{k \to \infty} (x(k) - \hat{x}(k)) = 0$ holds under certain conditions. Assuming $h(.)$ is a continuous function, then

$$\lim_{k \to \infty} [\bar{h}(x(k)) - \bar{h}(\hat{x}(k))] = \lim_{k \to \infty} \phi(k) = 0 \quad (31)$$

Consequently, if the sufficient condition (19) is satisfied and k is sufficiently large, D^m is Gaussian distributed with zero mean. If there is a sensor fault, (30) becomes,

$$D^m = \frac{1}{\sqrt{m}} \left(\sum_{k=1}^{m} b_f + \sum_{k=1}^{m} \phi(k) + \sum_{k=1}^{m} v(k) \right) \quad (32)$$

As b_f is non-zero, D^m is also non-zero.

3.1 Fault detection method

The proposed fault detection scheme can be implemented on-line as follows:

Step 1 Select m, the window size for computing the cumulative sum of residual.

Step 2 Compute the mean of the residual generated from the UKF. This is necessary, as $\psi(k)$ is ignored in the above analysis.

$$b_i(0) = \frac{1}{N} \sum_{k=1}^{N} \varepsilon_i(k) \quad (33)$$

where N is a large positive integer, the subscript i denotes the i^{th} component of vector.

Step 3 At the k^{th} sampling period, the cumulative sum of residuals is computed from (29) as given below.

$$D_i^m(k) = \frac{1}{\sqrt{m}} \sum_{t=k-m+1}^{k} (\varepsilon_i(t) - b_i(0)) \quad (34)$$

where $k > m$. Normalizing the cumulative sum of the residual by its variance gives,

$$S_i^m(k) = [D_i^m(k)]^2 [P_{\varepsilon\varepsilon}^{ii}(k-m)]^{-1} \quad (35)$$

where $P_{\varepsilon\varepsilon}^{ii}$ is the i^{th} diagonal element $P_{\varepsilon\varepsilon}$.

Step 4 If $S_i^m(k) \le \lambda_i$, then there is no fault, but a fault otherwise. As $S_i^m(k)$ is χ^2-distributed, λ_i can be obtained from χ^2-table for a given confidence level.

Step 5 Repeat step 3 and 4.

3.2 Properties of the fault detection method

If there is no fault, $\varepsilon(k)$ is Gaussian distributed: $N(0, P_{\varepsilon\varepsilon})$. From (35), the expectation of $\varepsilon(k)$ and the covariance matrix of D^m are respectively:

$$E(D^m) = \frac{1}{\sqrt{m}} \sum_{t=k-m+1}^{k} E(\varepsilon(t)) = 0 \quad (36)$$

$$Cov(D^m) = \frac{1}{m} \sum_{t=k-m+1}^{k} Cov(\varepsilon(t)) = P_{\varepsilon\varepsilon} \quad (37)$$

where $E(.)$ and $Cov(.)$ are respectively the expectation and the covariance. Hence D^m is also Gaussian distributed: $N(0, P_{\varepsilon\varepsilon})$. If there is a fault, the distribution of $\varepsilon(k)$ is: $N(b_f, P_{\varepsilon\varepsilon})$, and the mean and covariance of D^m are:

$$E(D^m) = \sqrt{m} b_f, \ Cov(D^m) = P_{\varepsilon\varepsilon} \quad (38)$$

The distribution of D^m is: $N(\sqrt{m} b_f, P_{\varepsilon\varepsilon})$. The miss-detection of the proposed fault detection scheme is given in the following theorem. This result provides a guideline for choosing m and the probability of the miss-detection. The argument t and the subscript i are ignored for simplicity.

Theorem 2: Let λ be obtained for a given confidence level. A fault is detected, if $S^m = (D^m)^2 P_{\varepsilon\varepsilon}^{-1} > \lambda$. The false alarm P_F is independent of m, while the miss-detection P_M depends on m.

Proof: If there is no fault, the distribution of D^m is $N(0, P_{\varepsilon\varepsilon})$, and the probability density function (pdf) of D^m is:

$$p(D^m \mid H_0) = \frac{1}{\sqrt{2\pi P_{\varepsilon\varepsilon}}} \exp\left(-(D^m)^2 / (2P_{\varepsilon\varepsilon})\right) \quad (39)$$

Let the null hypothesis denoting no fault be H_0. From (35), $S^2 = (D^m)^2 P_{\varepsilon\varepsilon}^{-1}$, the pdf of S^m is given by,

$$p(S^m \mid H_0) = \frac{2}{\sqrt{2\pi P_{\varepsilon\varepsilon}}} \exp\left(-S^m/2\right) \frac{P_{\varepsilon\varepsilon}}{2\sqrt{P_{\varepsilon\varepsilon} S^m}}$$
$$= \frac{1}{\sqrt{2\pi S^m}} \exp\left(-S^m/2\right) \quad (40)$$

The false alarm P_F is defined by,

$$P_F = \int_{\lambda}^{\infty} p(S^m \mid H_0) dS^m \tag{41}$$

Since $p(S^m|H_0)$ is independent of m, P_F is also independent of m. If there is a fault, the distribution of D^m becomes $N(\sqrt{m} b_f, P_{\varepsilon\varepsilon})$, and the pdf is:

$$p(D^m \mid H_1) = \frac{1}{\sqrt{2\pi P_{\varepsilon\varepsilon}}} \exp\left[-\frac{(D^m - \sqrt{m} b_f)^2}{2P_{\varepsilon\varepsilon}} \right] \tag{42}$$

Let H_1 be the hypothesis that there is a fault. Then the pdf of S^m can be expressed as,

$$p(S^m \mid H_1) = \frac{1}{2\sqrt{2\pi S^m}} \{\exp[-\frac{1}{2}(\sqrt{S^m} - \sqrt{m/P_{\varepsilon\varepsilon}} b_f)^2] + \exp[-\frac{1}{2}(\sqrt{S^m} + \sqrt{m/P_{\varepsilon\varepsilon}} b_f)^2]\} \tag{43}$$

From (43), the miss-detection P_M is given by,

$$\begin{aligned} P_M(m) &= \int_0^{\lambda} p(S^m \mid H_1) dS^m \\ &= \int_0^{\sqrt{\lambda}} \frac{1}{\sqrt{2\pi}} \{\exp[-\frac{1}{2}(x - \sqrt{m/P_{\varepsilon\varepsilon}} b_f)^2] + \exp[-\frac{1}{2}(x + \sqrt{m/P_{\varepsilon\varepsilon}} b_f)^2]\} dx \\ &= \Phi\left(-\sqrt{m/P_{\varepsilon\varepsilon}} b_f + \sqrt{\lambda}\right) - \Phi\left(-\sqrt{m/P_{\varepsilon\varepsilon}} b_f\right) \\ &\quad + \Phi\left(\sqrt{m/P_{\varepsilon\varepsilon}} b_f + \sqrt{\lambda}\right) - \Phi\left(\sqrt{m/P_{\varepsilon\varepsilon}} b_f\right) \end{aligned} \tag{44}$$

where $\Phi = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp\left(-x^2/2\right) dx$. Therefore P_M depends on m.

The relation between P_M and m is shown in Fig. 2, where the shaded part of the curve is P_M. Clearly, if m is large, the miss-detection from (44) is small,

$$\lim_{m \to \infty} P_M(m) = 0 \tag{45}$$

However, if m is large, a longer time is required before faults are detected (Wang and Chan, 2002). If the false alarm and the miss-detection are chosen to be small, λ and m can be determined from (41) and (44), as illustrated in the example presented below.

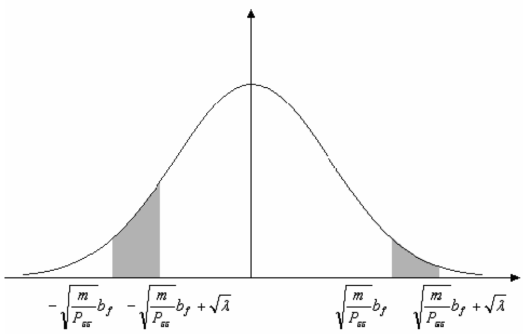

Fig. 2 The relation between P_M and m

4. SIMULATION EXAMPLE

4.1 Satellite attitude determination system

The satellite attitude determination system consists of the sun sensor, the earth sensor and the gyroscope, described by the following equation (Yang, 2002):

$$\begin{bmatrix} \dot{\gamma} \\ \dot{\theta} \\ \dot{\psi} \\ \dot{b}_{Ix} \\ \dot{b}_{Iy} \\ \dot{b}_{Iz} \\ \dot{d}_{Ix} \\ \dot{d}_{Iy} \\ \dot{d}_{Iz} \end{bmatrix} = \begin{bmatrix} A_{11} & -I_{3\times3} & -I_{3\times3} \\ 0_{3\times3} & 0_{3\times3} & 0_{3\times3} \\ 0_{3\times3} & 0_{3\times3} & A_{33} \end{bmatrix} \begin{bmatrix} \gamma \\ \theta \\ \psi \\ b_{Ix} \\ b_{Iy} \\ b_{Iz} \\ d_{Ix} \\ d_{Iy} \\ d_{Iz} \end{bmatrix} + \begin{bmatrix} \omega_{Ix} \\ \omega_{Iy} + \omega_0 \\ \omega_{Iz} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + W$$

$$\begin{bmatrix} m_\gamma \\ m_\theta \\ m_\psi \\ \gamma_h \\ \theta_h \end{bmatrix} = \begin{bmatrix} \dfrac{-\psi S_x^0 + S_y^0 + \gamma S_z^0}{\theta S_x^0 - \gamma S_y^0 + S_z^0} \\ \dfrac{S_x^0 + \psi S_y^0 - \theta S_z^0}{\theta S_x^0 - \gamma S_y^0 + S_z^0} \\ \dfrac{S_x^0 + \psi S_y^0 - \theta S_z^0}{-\psi S_x^0 + S_y^0 + \gamma S_z^0} \\ \gamma \\ \theta \end{bmatrix} + V \tag{46}$$

where

$$A_{11} = \begin{bmatrix} 0 & 0 & \omega_0 \\ 0 & 0 & 0 \\ -\omega_0 & 0 & 0 \end{bmatrix},$$

$$A_{33} = \begin{bmatrix} -1/\tau_{Ix} & 0 & 0 \\ 0 & -1/\tau_{Iy} & 0 \\ 0 & 0 & -1/\tau_{Iz} \end{bmatrix}$$

$I_{3\times3}$ is the identity matrix, $0_{3\times3}$, the zero matrix, γ, θ and ψ are the roll, the pitch and the heading of satellite, ω_0 is orbit angle velocity, ω_{Ix}, ω_{Iy}, ω_{Iz} are the measurement from gyroscope, b_{Ix}, b_{Iy}, b_{Iz}, d_{Ix}, d_{Iy}, d_{Iz} are the drifting errors of the gyroscope, τ_{Ix}, τ_{Iy}, τ_{Iz} are the first order Markov time constant, S_x^0, S_y^0, S_z^0 are the projections of sun vector onto the coordinate of the spacecraft, m_γ, m_θ, m_ψ are measurements of sun sensor, γ_h, θ_h are measurements of earth sensor, W and V are zero mean Gaussian white noise.

4.2 Simulation results

It is assumed in the simulation that the satellite is being stabilized relative to the earth. The initial values of θ, γ and ψ are set to zero. For a sampling interval of 0.1 second, the satellite given by (46) is simulated for 50 seconds. The drifting error of the parameters of the gyroscope is 10°/h, the measurement noises of sun sensor and earth sensor are zero-mean, uncorrelated noises with covariance given by constants $R^{ii} = 0.01^2$, for i = 1, …, 5. The proposed fault detection scheme is applied to detect the following fault in sun sensor, which occurred separately at 30s.

$$y_{f,1}(k) = y_1(k) + 0.02 \quad \text{, for} \quad k \geq 300 \text{ ;}$$

where the constant b_f = 0.02 represents a sensor fault is being added to the observation m_γ, and from (46) and (32), there is a drift in $\varepsilon_1(k)$. Following the procedure described in section 2.2, residuals are obtained from the UKF. The false alarm is set to: P_F = 0.1%, and the miss-detection P_M is expected to be not larger than 6%. From (41), λ_i obtained from the χ^2-table for a 0.1% false alarm is: λ_i = 10.8. When the fault occurs, b_f = 0.02, the miss-detection rate can be computed by (44). For m = 6, $P_M \approx 6\%$ from statistical table on Gaussian distribution as $P_{\varepsilon\varepsilon}^{11}$ is set to 0.01^2. So the requirement on miss-detection can be satisfied. If m = 1, the miss-detection is about 90%, and hence the requirement on the miss-detection is not satisfied. In this case, it is necessary to increase m to reduce the miss-detection.

When the fault occurs, the residuals $\varepsilon_1(k)$ and $\varepsilon_4(k)$ are shown in Fig. 3, showing a small step change in the mean of $\varepsilon_1(k)$, for k > 300. For m = 6, $S_1^m(k)$ and $S_4^m(k)$ are shown in Fig. 4. As only $S_1^m(k)$ is greater than the threshold for k > 302, a fault is detected in the component of the sun sensor, which corresponds to $S_1^m(k)$. This result agrees with the properties of the fault detection method presented in section 3.2, illustrating the ability of the local approach in detecting faults.

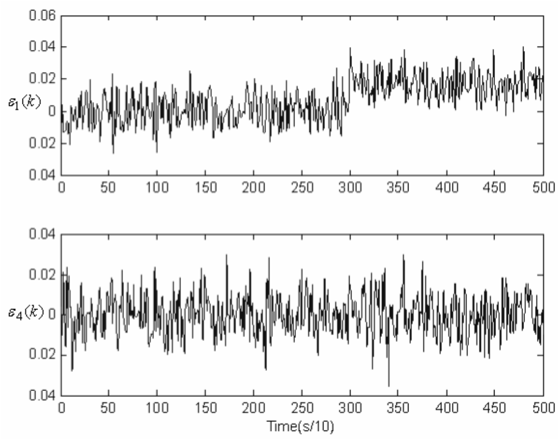

Fig. 3 $\varepsilon_1(k)$ and $\varepsilon_4(k)$

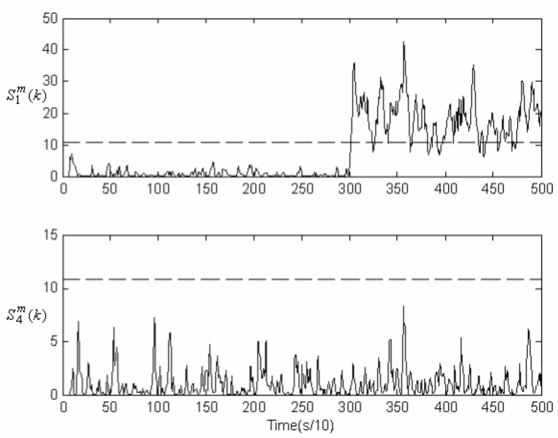

Fig. 4 $S_1^m(k)$ and $S_4^m(k)$

5. CONCLUSION

In this paper, a fault detection scheme for nonlinear systems is derived based on the UKF and the local approach. Since the UKF can approximate the mean and the covariance of a Gaussian random variable up to a second order accuracy, it is used here to generate residuals for detecting faults. The sufficient condition for the convergence of the UKF is presented. The local approach is applied to detect faults from the residuals, and properties of this method are derived. These properties are then used to devise guidelines for choosing the window size in the statistical test. The proposed method has been applied successfully to detect faults in the satellite attitude determination system.

ACKNOLEDGEMENT

This work was supported in part by China Natural Science Foundation (No. 60234010), and the HKSAR RGC Grant (HKU 7050/02E).

REFERENCES

Boutayeb, M., H. Rafaralahy, and M. Darouach (1997). Convergence analysis of the extended Kalman filter used as an observer for nonlinear deterministic discrete-time systems. *IEEE Trans. Auto. Contr.*, **42**, 581-586.

Del Gobbo, D., M. Napolitano, P. Famouri and M. Innocenti (2001). Experimental application of extended Kalman filtering for sensor validation. *IEEE Trans. Contr. Sys. Technology*, **9**, 376-380.

Hall, D. L. and J. Llinas (2001), *Handbook of Multisensor Data Fusion*, CRC Press.

Julier, S., J. Uhlmann and H. F. Durrant-Whyte (2000). A new method for the nonlinear transformation of means and covariances in Filters and Estimators. *IEEE Trans. Auto. Contr*. 45, 477-482.

Magnus Norgaard,Niels K. Poulsen, Ole Ravn(2000), New developments in state estimation for nonlinear systems, Automatic, 36, 1627-1638

Wan, E. A. and R. van der Merwe (2000). The unscented Kalman filter for nonlinear estimation, adaptive systems for signal processing. *Communication and Contr. Symposium*, 153-158.

Wang, Y. and C. W. Chan (2002). Asymptotic local approach in fault detection with faults modeled by neurofuzzy networks. *Proceedings 15th Triennial World Congress, Barcelona, Spain*.

Yang, J. (2002). *Attitude determination, positioning and fault diagnosis of GPS*. Ph.D. Dissertation, Beihang University.

Yu, S. X. and J. Shi (2004). Segmentation given partial grouping constraints. *IEEE Trans. on Pattern Analy. and Mach. Intellig.*, 26, 173-183.

Zhang, Q., M. Basseville and A. Benveniste (1994). Early warning of slight changes in system. *Automatica*, 30, 95-115.

Zhang, Q., M. Basseville and A. Benveniste (1998). Fault detection and isolation in nonlinear dynamic systems: a combined input-output and local approach. *Automatica*, **34**, 1359-1373.