

Outlier-Tolerant Kalman Filter of State Vectors in Linear Stochastic System

HU Shaolin

State Key Laboratory of Astronautics
Xi'an, 710043, P.O.Box 505-16, China;

Karl Meinke

CSC School, Royal Institute of Technology
SE-100 44, Stockholm, Sweden

Huajiang Ouyang

School of Engineering, The University of Liverpool
Liverpool, L69 3GH, UK;

SUN Guoji

System Engineering Institute, Xi'an Jiaotong University
Xi'an, 710043, China

Abstract—The Kalman filter is widely used in many different fields. Many practical applications and theoretical results show that the Kalman filter is very sensitive to outliers in a measurement process. In this paper some reasons why the Kalman Filter is sensitive to outliers are analyzed and a series of outlier-tolerant algorithms are designed to be used as substitutes of the Kalman Filter. These outlier-tolerant filters are highly capable of preventing adverse effects from outliers similar with the Kalman Filter in complexity degree and very outlier-tolerant in the case there are some outliers arisen in sampling data set of linear stochastic systems. Simulation results show that these modified algorithms are safe and applicable.

Keywords- Kalman filter; Outlier-tolerant; Outlier; Linear stochastic system.

I. INTRODUCTION

The Kalman filter is not only a widely used tool to estimate or to reconstruct states of a dynamic system in modern control but also famous powerful tool to extract useful information from noisy signals in signals processing. The Kalman filtering algorithms have many advantages: it is optimal in linear estimator set and suitable for online processing because of its recursive relationship; it can be used in stationary system and non-stationary system; and it can be used in multi-dimensional processes, etc. Since Kalman and Bucy put forward this linear optimal iterative filtering algorithm with the development of state-space theory in the early 1960s', its applications have become more and more wide-spread in many different engineering fields, such as process control, stochastic control and navigation, etc. It is also very useful in fault diagnosis of dynamic systems.

Although the Kalman filter possesses many advantages stated above, recent research has revealed that the Kalman filtering algorithm is not robust (see [1,3]) against perturbation in a model or observed/measured data. Practical experience in using Kalman filter to process signals also indicates that outliers in the measured data would degrade the performance of Kalman filter. How to improve the Kalman filtering algorithm is an open question in the cases when there are outliers in measurement data sequence because outliers are unavoidable

and could lead to a considerable deviation of the estimated target from the true system status when using Kalman-filter based algorithms. Durovic, et al. (1999) discussed robust estimation with unknown noise statistics; Nihal et al. (1991) and Chan et al. (2005) built a new robust Kalman filter algorithm with outliers respectively; Ting et al. (2007) reviewed the Kalman filter and suggested a kind of robust Kalman filtering with Bayesian weights so as to overcome negative effects from outliers.

This paper analyzes systematically the adverse effect of outliers on Kalman filter and establishes a series of outlier-tolerant filtering algorithms. Numerical results of simulated examples show the validity of the outlier-tolerant algorithms.

II. EFFECTS OF OUTLIERS ON KALMAN FILTER

The Kalman filter is used to estimate the state vector of the following linear stochastic dynamic-measurement system

$$\begin{cases} X_{k+1} = A_{k+1}X_k + \varepsilon_k \\ Y_{k+1} = H_{k+1}X_{k+1} + \eta_{k+1} \end{cases} \quad (X_k \in R^n, Y_k \in R^m) \quad (1)$$

If both the multi-dimensional dynamic noise process $\{\varepsilon_k\}$ and dynamic measurement error series $\{\eta_k\}$ possess the following properties:

- (a) $E\{\varepsilon_k\} = E\eta_k = 0$
- (b) $\text{cov}(\varepsilon_k, \eta_k) = 0$
- (c) $R_{\eta(k)} = \text{cov}(\eta_k, \eta_k)$, $R_{\varepsilon(k)} = \text{cov}(\varepsilon_k, \varepsilon_k)$

then the optimal estimator of the state vector X_k in model (1) can be expressed by the following recursive relationships:

$$\begin{cases} \hat{X}_{(k+1|k+1)} = A_{k+1}\hat{X}_{(k|k)} + K_{k+1}\hat{E}_{(k+1|k)} \\ \hat{X}_{(k+1|k)} = A_{k+1}\hat{X}_{(k|k)} \\ K_{k+1} = \sum_{k+1} H_{k+1}^T \{H_{k+1} \sum_{k+1} H_{k+1}^T + R_{\eta(k+1)}\}^{-1} \\ \sum_{k+1} = A_{k+1}(I - K_k H_k) \sum_k A_{k+1}^T + R_{\varepsilon(k)} \end{cases} \quad (2)$$

These recursive relationships can intuitively express that the best estimator $\hat{X}_{(k+1|k+1)}$ of the state X_{k+1} at time t_{k+1} is composed of two parts: the best estimator of the state X_k at time t_k and the sampling innovation as follows

$$\hat{E}_{(k+1|k)} = Y_{k+1} - H_{k+1}A_{k+1}\hat{X}_{(k|k)} \quad (3)$$

Obviously, if the additional samples Y_{k+1} are “normal” values, the sampling innovation $\hat{E}_{(k+1|k)}$ will make some correct modification to predictor $\hat{X}_{(k+1|k)}$ with ratio K_{k+1} to get the next optimal estimator $\hat{X}_{(k+1|k+1)}$. On the other hand, if the additional samples Y_{k+1} are outliers, the resultant innovation $\hat{E}_{(k+1|k)}$ will be abnormal and the abnormal information will result in an erroneous modification to prediction $\hat{X}_{(k+1|k)}$ by the same ratio K_{k+1} , which lead to the filtering estimators deviating from normal states.

In section V of this paper, an example is given to substantiate the negative influence of outliers on the Kalman filtering estimators of state vectors. Figure 1 plots the simulation results, which indicates that the Kalman filtering of state vectors is far from normal states when there are outliers in sampling data series.

For a practical measurement system or device, outliers are inevitable in sampling data because of faulty operations or recording errors. In some cases, there may be complicated abnormal measurement data existing in sampling processes, such as step-type jumps or patchy outliers, etc. For example, when the flight of a spacecraft was tracked by an impulse radar, 1%~2% (occasionally as high as 5%) of measurement data display serious deviations from the trend formed by most other samples. So it is very important to modify the Kalman filtering algorithms or reduce the negative effects of outliers on Kalman filtering estimators of state vectors.

III. TACTICS TO DEAL WITH OUTLIERS IN SAMPLE

Considering that it is very difficult to diagnose outliers in a large quantity of data and that the Kalman filtering is still widely used, a best effort is made in this paper to improve fault-tolerance of the linear optimal recursive filtering algorithm. The improved filters should possess the following properties:

- 1) The algorithm should be recursive and easy to use;
- 2) When there are a few abnormal samples in the measured data series, the filtering algorithm must have a strong ability to overcome negative effects from abnormal data, or must restrict the negative effects within prescribed bounds;
- 3) When there are no abnormal samples, the algorithm is capable of making full use of useful information to achieve high filtering accuracy.

Considering these three restrictive conditions stated above, a set of algorithms which are similar to the Kalman filter are established as follows

$$\begin{cases} \tilde{X}_{(k+1|k+1)} = A_{k+1}\tilde{X}_{(k|k)} + K_{k+1}\Phi_{k+1}(r_{k+1})\hat{E}_{(k+1|k)} \\ r_{k+1} = \hat{E}_{(k+1|k)}^T G_{k+1}^{-1} \hat{E}_{(k+1|k)} \end{cases} \quad (3)$$

where the function series $\{\Phi_{k+1}(\cdot)\}$ are segment-wise smooth and bounded, K_{k+1} is the gain and G_{k+1} is a weighting matrix.

When $\Phi_{k+1} = 1$ and $G_{k+1} = I$ (identity matrix) are selected, $\tilde{X}_{(k+1|k+1)}$ in equation (3) is reduced to the conventional Kalman filter. It is found that the main reason for outlier-tolerance of conventional Kalman filter is that the function sequence $\{\Phi_{k+1} = 1\}$ potentially treat all of the innovations (normal and abnormal) equally. This is the root cause why Kalman filter is unable to deal with outliers. In order to endow a filter anti-outliers capability, a sensible tactic is to select a suitable Φ_{k+1} which must decrease or diminish to zero when r_{k+1} increases.

IV. OPTIMAL SELECTION OF Φ_k

In order to reduce the negative effects of outliers in state filtering, some analyses of formulae (3) must be done as follows:

$$\begin{aligned} \Delta\tilde{X}_{(k+1|k+1)} &= \tilde{X}_{(k+1|k+1)} - \tilde{X}_{(k+1|k)} \\ &= \tilde{X}_{(k+1|k+1)} - A_{k+1}\tilde{X}_{(k|k)} \\ &= K_{k+1}\Phi_{k+1}(\|\omega_{(k+1|k)}\|^2)G_{k+1}^{1/2}\omega_{(k+1|k)} \end{aligned} \quad (4)$$

where $\omega_{(k+1|k)} = G_{k+1}^{-\frac{1}{2}}\hat{E}_{(k+1|k)}$ is one-step predicted weighted residual.

For threshold constant series $\{C_{k+1}\}$ used to control the difference between the filtered values and the predicted values of state vectors X_{k+1} , Φ_{k+1} should be suitably chosen so as to make sure that the following inequality holds

$$\begin{aligned} \|\Delta\tilde{X}_{(k+1|k+1)}\| &\leq \{\omega_{(k+1|k)}^T \omega_{(k+1|k)} \lambda_{k+1} \Phi_{k+1}^2(\|\omega_{(k+1|k)}\|^2)\}^{1/2} \\ &\leq C_{k+1} \end{aligned} \quad (5)$$

where λ_{k+1} is the maximum eigenvalue of matrix $K_{k+1}G_{k+1}K_{k+1}^T$.

It is easy to see that there are quite many different kinds of function series $\{\Phi_{k+1}\}$ satisfying inequality (5). All of the function series $\{\Phi_{k+1}\}$ are denoted as a set S :

$$S = \{\{\Phi_{k+1}\} : \Phi_{k+1} \leq 1, \sqrt{r}\Phi_{k+1}(r) \leq \frac{C_{k+1}}{\sqrt{\lambda_{k+1}}}, r \in [0, +\infty), k \in N\} \quad (6)$$

The way to select function series $\{\Phi_{k+1}\}$ from the set S is examined below.

Theorem 1 : For a linear stochastic system, when the noise series $\{x_0, \varepsilon_0, \varepsilon_1, \Lambda; \eta_1, \eta_2, \Lambda\}$ are stochastic sequences that possess a normal distribution and are mutually independent, if

the weighting matrix sequence are $G_{k+1} = H_{k+1}\Sigma_{k+1}H_{k+1}^T + R_{\eta(k+1)}$, then the function series $\{\Phi_{k+1}\} \in S$ which lead to the least errors of filtering estimation are as follows:

$$\Phi_k(r) = \begin{cases} 1, r \leq \frac{C_k^2}{\lambda_k} \\ \frac{C_k}{\sqrt{r\lambda_k}}, r > \frac{C_k^2}{\lambda_k} \end{cases} \quad (7)$$

where C_k is the suitably chosen threshold constant.

Proof : It follows from equations (2) and (3) that the filtering error can be expressed as follows:

$$\begin{aligned} E\|\tilde{X}_{(k+1|k+1)} - X_{k+1}\|^2 &= E\|\tilde{X}_{(k+1|k+1)} - \hat{X}_{(k+1|k+1)} + \hat{X}_{(k+1|k+1)} - X_{k+1}\|^2 \\ &= E\|\tilde{X}_{(k+1|k+1)} - \hat{X}_{(k+1|k+1)}\|^2 + E\|\hat{X}_{(k+1|k+1)} - X_{k+1}\|^2 \\ &\quad + 2E(\tilde{X}_{(k+1|k+1)} - \hat{X}_{(k+1|k+1)})^T (\hat{X}_{(k+1|k+1)} - X_{k+1}) \end{aligned}$$

Using the basic properties on conditional probability distribution, the following results may be deduced

$$\begin{aligned} E(\tilde{X}_{(k+1|k+1)} - \hat{X}_{(k+1|k+1)})^T (\hat{X}_{(k+1|k+1)} - X_{k+1}) &= E\{E(\tilde{X}_{(k+1|k+1)} - \hat{X}_{(k+1|k+1)})^T (\hat{X}_{(k+1|k+1)} - X_{k+1}) | Y_1, \Lambda, Y_{k+1}\} \\ &= E(\tilde{X}_{(k+1|k+1)} - \hat{X}_{(k+1|k+1)})^T E\{(\hat{X}_{(k+1|k+1)} - X_{k+1}) | Y_1, \Lambda, Y_{k+1}\} \\ &= E(\tilde{X}_{(k+1|k+1)} - \hat{X}_{(k+1|k+1)})^T \{ \hat{X}_{(k+1|k+1)} - E(X_{k+1} | Y_1, \Lambda, Y_{k+1}) \} \\ &= 0 \end{aligned}$$

Note that the project property of Kalman filtering has been used in the last equation of the above expression. So the following result can be obtained:

$$\begin{aligned} E\|\tilde{X}_{(k+1|k+1)} - X_{k+1}\|^2 &= E\|\tilde{X}_{(k+1|k+1)} - \hat{X}_{(k+1|k+1)}\|^2 + E\|\hat{X}_{(k+1|k+1)} - X_{k+1}\|^2 \end{aligned} \quad (8)$$

In order to investigate how to minimize the error in equation (8), equations (2) and (3) are used to deduce the following expression:

$$E\|\tilde{X}_{(k+1|k+1)} - \hat{X}_{(k+1|k+1)}\|^2 = E\{[1 - \Phi_{k+1}(r_{k+1})]^2 \|K_{k+1}\hat{E}_{(k+1|k)}\|^2\} \quad (9)$$

It is obvious that function series $\{\Phi_k\} \in S$ which were prescribed in the equation (7) minimize the equation (9). ■

In the following part of this section, we will discuss how to calculate weigh matrix sequence $\{G_{k+1}\}$ and threshold constants C_{k+1} ($k = 1, 2, \Lambda$).

Hypothesis 1. Suppose that $\{\varepsilon_k\}$ and $\{\eta_k\}$ are two independent noise sequences and they satisfy following relationships:

- a). $E\varepsilon_k = 0, E\eta_k = 0, Ex_0 = \bar{x}_0, Ex_0x_0^T = R_0 + \bar{x}_0\bar{x}_0^T$;
- b). $E\varepsilon_k\varepsilon_k^T = R_{\varepsilon(k)}\delta(k-k), E\eta_k\eta_k^T = R_{\eta(k)}\delta(k-k)$;
- c). $E\varepsilon_k\eta_k^T = 0, Ex_0\varepsilon_k^T = 0, Ex_0\eta_k^T = 0$

Hypothesis 2 : Suppose that $\{\varepsilon_k\}$ and $\{\eta_k\}$ are two independent noise sequences and they satisfy following distributions:

$$x_0 \sim N(\bar{x}_0, R_0); \varepsilon_i \sim N(0, R_{\varepsilon(i)}); \eta_i \sim N(0, R_{\eta(i)})$$

Lemma 1 : If a linear stochastic system defined by equation (1) satisfies hypothesis 1 and hypothesis 2, then stochastic sequence $\{x_0, \varepsilon_0, \Lambda, \varepsilon_k; \eta_1, \Lambda, \eta_k\}$ obeys the law of a multivariate normal distribution and their joint covariance matrix C_a is as follows:

$$C_{a(k)} = \begin{pmatrix} C_{a1}(k) & 0 \\ 0 & C_{a2}(k) \end{pmatrix} = C_{a1}(k) \& C_{a2}(k) \quad (10)$$

where operator “&” represents the block-diagonal matrix formed by the two block matrices on either side of the operator; $C_{a2}(k)$ is the covariance matrix of the measured noise sequence, which can be expressed as the inverse matrix of the tri-diagonal matrix $A = (a_{i,j})_{k \times k}$ which is defined as follows

$$\begin{cases} a_{1,1} = R_0^{-1} \& A_1^T R_{\varepsilon(1)}^{-1} A_1 \\ a_{i,i} = R_{\varepsilon(i-1)}^{-1} \& A_i^T R_{\varepsilon(i)}^{-1} A_i \\ a_{k,k} = R_{\varepsilon(k-1)}^{-1} \\ a_{i,i+1} = A_i^T R_{\varepsilon(i)}^{-1} \\ a_{i-1,i} = R_{\varepsilon(i)}^{-1} A_i \\ a_{i,j} = 0 \quad (|i-j| \geq 2, i, j = 1, \dots, k) \end{cases} \quad (11)$$

In formula (11) matrix $C_{a2}(k)$ is the covariance matrix of the measured noise sequence $\{\eta_k\}$.

$$C_{a2}(k) = \text{diag}\{R_{\eta(1)}, \Lambda, R_{\eta(k)}\}$$

Lemma 2 : If the linear stochastic system defined by equation (1) satisfies hypothesis 1 and hypothesis 2, the joint distributions of sequence $\{x_0, x_1, \Lambda, x_k; y_1, \Lambda, y_k\}$ are normal and their covariance matrix is given by the following relationships

$$C_b = \begin{bmatrix} C_{a1} & M & C_{a1}H^T \\ \Lambda & \Lambda & \Lambda \\ HC_{a1} & M & HC_{a1}H^T + C_{a1} \end{bmatrix}, H = \begin{bmatrix} 0 & H_1 & \Lambda & 0 \\ M & M & & M \\ 0 & 0 & \Lambda & H_k \end{bmatrix}$$

where the superscript τ denotes transpose of a matrix.

Analyzing these two lemmas described above, it is found that the joint distribution of sequence $\{y_1, \Lambda, y_k\}$ is also normal.

If the linear stochastic system defined by equation (1) satisfies those two hypotheses, denoting the joint covariance matrix of $\{y_1, \Lambda, y_k\}$ by $D_{(k)}$ and the covariance matrix of $\hat{E}_{(k+1|k)}$ by Ω_{k+1} , it is easy to verify the properties of Kalman filtering algorithms that

$$\Omega_{k+1} = d_{k+1} - D_{d(k)}{}^\tau D_{(k)}^{-1} D_{d(k)} \quad (13)$$

where $d_{k+1} = \text{COV}(y_{k+1}, y_{k+1})$, the auto-variance matrix of stochastic vector y_{k+1} ; matrix $D_{d(k)}$ is the covariance matrix of $\{y_1, \Lambda, y_k\}$ and y_{k+1} .

Weighting matrix G_{k+1} and weighted residual $\omega_{(k+1|k)}$ can be calculated as follows

$$\begin{cases} G_{k+1} = \Omega_{k+1} = d_{k+1} - D_{d(k)}{}^\tau D_{(k)}^{-1} D_{d(k)} \\ \omega_{(k+1|k)} = \Omega_{k+1}^{-1/2} \hat{E}_{(k+1|k)} \end{cases} \quad (14)$$

It can be proven that $\omega_{(k+1|k)}$ follows the standard multivariate normal distribution. Supposing that the measured information is an m -dimensional vector, the weighted residual $\omega_{(k+1|k)}$ is an m -dimensional standard normal variable, the norm of which is equal to r_{k+1} :

$$r_{k+1} = \hat{E}_{(k+1|k)}{}^\tau \Omega_{k+1}^{-1} \hat{E}_{(k+1|k)} \sim \chi^2(m) \quad (15)$$

where $\chi^2(m)$ denotes the χ -distribution of degrees-of-freedom m .

Since the random variable r_{k+1} satisfies the distribution $\chi^2(m)$, the threshold constant $c^\alpha(m)$ satisfies

$$P(r_{k+1} > c^\alpha(m)) = \alpha$$

Generally parameter α is taken as 0.05 or 0.025. If $r_{k+1} \leq c^\alpha(m)$, it is believed with $(1-\alpha) \times 100\%$ confidence that the additional information from the measured data y_{k+1} is reasonable; otherwise, when $r_{k+1} \geq c^\alpha(m)$, it is also believed with $\alpha \times 100\%$ confidence that the additional information from the measured data y_{k+1} is unreasonable and hence this additional information must be discarded.

From the above analysis and considering equation (6), it is known that C_{k+1} can be given by

$$C_{k+1} = (\lambda_{k+1} c^\alpha(m))^{1/2} \quad (16)$$

where λ_{k+1} is the maximum eigenvalue of matrix $K_{k+1} K_{k+1}{}^\tau$.

In specific applications in engineering, C_{k+1} may be chosen with our experience.

V. NUMERICAL SIMULATION

Supposing that the coefficient matrices and covariance matrix of errors defined by equation (1) are as follows

$$A_k = \begin{bmatrix} 1.0 & 0 & 0 \\ 0.2 & 1.0 & 0 \\ 0 & 0.3 & 1.0 \end{bmatrix}, \quad H_k = \begin{pmatrix} 0.9 & 0 \\ 0 & 0.2 \\ 0 & 0.7 \end{pmatrix}{}^\tau$$

and the covariance of the model errors are respectively equal to

$$\begin{cases} R_{\delta(k)} = \text{diag}\{0.5, 0.5, 0.5\} \\ R_{\eta(k)} = \text{diag}\{0.1, 0.1\} \end{cases}$$

Using Monte Carlo method and selecting the initial states of the system as the following

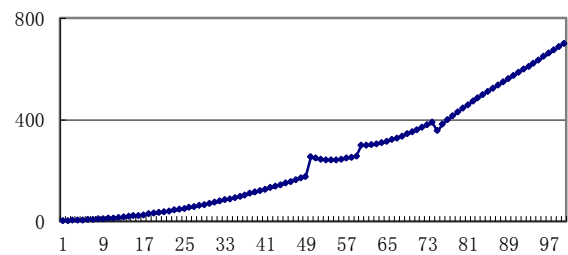
$$\begin{cases} x_{1(0|0)} = 1.3 \\ x_{2(0|0)} = 1.5 \\ x_{3(0|0)} = 2.3 \end{cases}$$

$$\Sigma_{(0|0)} = \begin{bmatrix} 0.01 & & \\ & 0.01 & \\ & & 0.01 \end{bmatrix}.$$

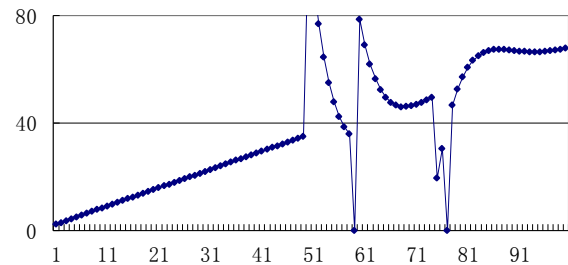
one hundred pieces of simulation data are generated and denoted by the set S . Let deviations of the 50th and 75th pieces of data be

$$\Delta y_1 = (-1)^i 100, \Delta y_2 = (-1)^{i+1} 5 \quad (i = 50, 75)$$

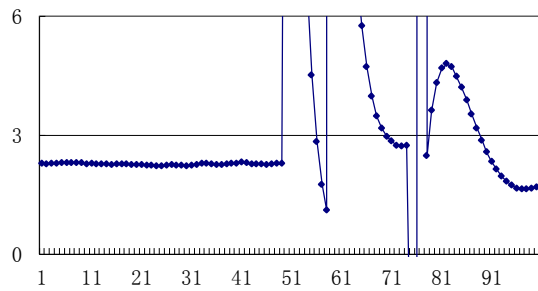
The new set with these two outliers is denoted by S^* . The estimates on S^* made by equations (1) and (2) are shown in figures 1 and 2.



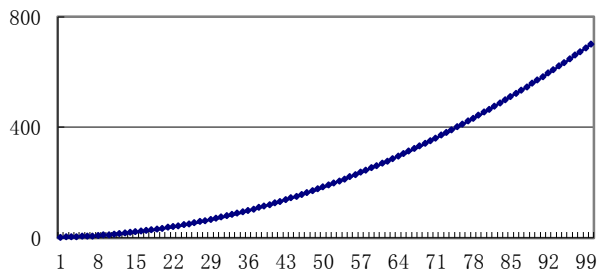
(a) Filtering Estimation of X_1



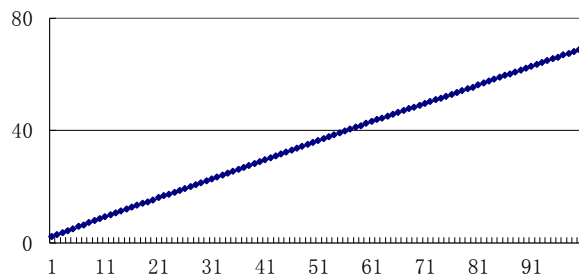
(b) Filtering Estimation of X_2



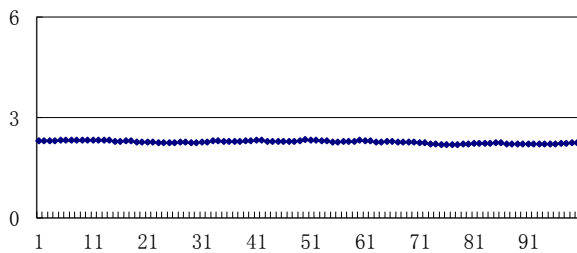
(c) Filtering Estimation of X_3
Figure 1 Kalman Filters of State Variables $X \in \mathbb{R}^3$



(a) Filtering Estimation of X_1



(b) Filtering Estimation of X_2



(c) Filtering Estimation of X_3

Figure 2 Fault-tolerant Filters of State Variables $X \in \mathbb{R}^3$

It can be seen clearly that the outliers have a very negative effect on conventional Kalman filtering and the outliers-tolerant modification method proposed in the paper is capable of overcoming this negative effect and is reliable.

ACKNOWLEDGEMENTS

The research was supported by the Sweden Institute Grant (SI 05483/2005-210), the National Natural Science Foundation of China (No.61074077) and the Tianyuan Fund of National Natural Science Foundation of China (No.11026224).

The first author would like to thank Dean Ingrid Melinder of the Computer Science and Communication (CSC) of the Royal Institute of Technology (KTH) for her friendly help when he visited the CSC of KTH of Sweden.

REFERENCES

- [1] Hu Shaolin, Fan Jingcheng. "Bounded Influence Filter of the Dynamic-Measurement System". *Control Theory and Applications*, 1993.10(1): 36-45
- [2] Han Congzhao, Wang Yuejuan, Wan Baiwu. *Stochastic System Theory*. Press of the Xi'an Jiaotong University, 1987
- [3] Martin D, Mintz. "Robust Filtering and Prediction for Linear System with Uncertain Dynamics — a game theoretic approach". *IEEE Trans. Automatic Control*, 1983, AC-18
- [4] Martin D. *Robust Method for Time Series*, Academic Press Inc, 1980
- [5] Hu Shaolin, Wang Xiaofeng, Karl Meinke, Huajiang Ouyang. Outlier-tolerant fitting and online diagnosis of outliers in dynamic process sampling data series. in *Artificial Intelligence and Computational Intelligence*, pp.195-204, LNAI 7004(Deng Miao, Lei Wang, eds) Springer Press, 2011
- [6] F Carlos, S Alcalá, J Qin. Unified Analysis of Diagnosis Methods for Process Monitoring. In: *Proc of the 7th IFAC Symp on Fault Detection, Supervision and Safety of Technical Processes*, Spain, 2009, 1007-1012
- [7] Nihal Y, Bovas A, MacGregor J. A kalman filter in the presence of outliers *Communications in Statistics - Theory and Methods* Volume 20, Issue 5-6, 1991
- [8] Durovic, Z.M., Kovacevic, B.D.: Robust estimation with unknown noise statistics. *IEEE Transactions on Automatic Control* 44, 1292-1296, 1999
- [9] Chan, S.C., Zhang, Z.G., Tse, K.W.: A new robust Kalman filter algorithm under outliers and system uncertainties. In: *IEEE International Symposium on Circuits and Systems*. IEEE, 4317-4320, 2005
- [10] Jo-Anne Ting, Evangelos Theodorou, Stefan Schaal. A Kalman Filter for Robust Outlier Detection. *Computational Learning & Motor Control Lab University of Southern California, IROS*, 2007