STOCHASTIC ANALYSIS
BANACH CENTER PUBLICATIONS, VOLUME 105
INSTITUTE OF MATHEMATICS
POLISH ACADEMY OF SCIENCES
WARSZAWA 2015

STATE ESTIMATION UNDER NON-GAUSSIAN LÉVY NOISE: A MODIFIED KALMAN FILTERING METHOD

XU SUN

School of Mathematics and Statistics, Huazhong University of Science and Technology
Wuhan, Hubei, 430074, China
E-mail: xsun15@gmail.com; xsun@hust.edu.cn

JINQIAO DUAN and XIAOFAN LI

Department of Applied Mathematics, Illinois Institute of Technology
Chicago, IL 60616, U.S.A.
E-mail: duan@iit.edu, lix@iit.edu

XIANGJUN WANG

School of Mathematics and Statistics, Huazhong University of Science and Technology
Wuhan, Hubei, 430074, China
E-mail: xjwang@hust.edu.cn

Abstract. The Kalman filter is extensively used for state estimation for linear systems under Gaussian noise. When non-Gaussian Lévy noise is present, the conventional Kalman filter may fail to be effective due to the fact that the non-Gaussian Lévy noise may have infinite variance. A modified Kalman filter for linear systems with non-Gaussian Lévy noise is devised. It works effectively with reasonable computational cost. Simulation results are presented to illustrate this non-Gaussian filtering method.

1. Introduction and statement of the problem. The Kalman filter, or the Kalman filtering method, provides an efficient way to estimate the state of a linear dynamical system subject to Gaussian white noise [4, 6, 8]. It has been widely used in applications such as target tracking, parameter estimation, control theory, signal processing, and other data assimilation tasks.

2010 *Mathematics Subject Classification*: Primary 93E11; Secondary 60G35.
Key words and phrases: Kalman filter, non-Gaussian white noise, Lévy process, state estimation.
The paper is in final form and no version of it will be published elsewhere.

DOI: 10.4064/bc105-0-14

The Kalman filter requires the noise be either Gaussian or with finite variance [4, 6], and thus it is not applicable to linear systems with non-Gaussian noise of infinite variance. α-stable noise and many other impulsive noises with heavy-tailed distributions [7, 10] are specific examples of Lévy noise which has infinite variance. As non-Gaussian Lévy noise with infinite variance exists ubiquitously [14, 12, 10], it is desirable to study the Kalman filtering problems under Lévy noise. Very little work has been done for this issue. Le Breton and Musiela [9] presented a scheme for Kalman filtering with noise of infinite variance, while assuming the contribution of the jumps are exactly known. The filter in [9] is nonlinear and recursive, and thus may greatly limit its application in practice. Ahn and Feldman [1] proposed to minimize the difference between the true state and the filtered observation in the L^μ-norm. However, as pointed in [12], this method does not really address the Kalman filtering problem which consists of combining forecasts and observations. The Kalman–Lévy filter proposed in [12] focuses on large errors and has a robust performance. As shown by the simulation results in [5], the Kalman–Lévy filter with a simple heavy-tailed model could achieve comparable performance to a well designed three model IMM filter. In [11] a new type of IMM estimator, which combines Kalman and Kalman–Lévy filters, is developed. However, the Kalman–Lévy filter has a very high computational cost due to the matrix diagonalization and the operation of the fractional power in each step [12, 5, 11]. This may be an obstacle for real-time implementation in practical applications. Note that in practice, each iteration step must be completed during every sampling period, and it is greatly desirable to make the algorithm as fast as possible.

In this letter, we will present a simple algorithm, which has the similar computational cost as that of the Kalman filter, but can be applied to linear systems with non-Gaussian Lévy noise of infinite variance.

We consider the discrete time model with the state equation

$$\mathbf{x}_{k+1} = \mathbf{F}_k\mathbf{x}_k + \mathbf{w}_k, \tag{1}$$

and the observation equation

$$\mathbf{z}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{v}_k, \tag{2}$$

where $\mathbf{x}_k$, an n-by-1 vector, is the state variable, and $\mathbf{z}_k$, an m-by-1 vector, is the measurement (or observation) variable, $\mathbf{w}_k$ represents the modeling error noise, $\mathbf{v}_k$ the measurement error noise, and $\mathbf{F}_k$ and $\mathbf{H}_k$ are n-by-n and m-by-n matrices, respectively. We only consider the cases where $\mathbf{w}_k$ is a Gaussian noise, and $\mathbf{v}_k$ is a non-Gaussian Lévy noise.

This letter is arranged as follows. In Section 2, the usual Kalman filter is briefly reviewed. The proposed modified Kalman filter is presented in Section 3. A simulation example is provided in Section 4 to illustrate the effectiveness of the modified Kalman filter.

2. Review of the conventional Kalman filter. A derivation of the Kalman filter is briefly reviewed in this section. Some equations and ideas presented in this section will be used to present our proposed modified Kalman filter in the next section. Derivations of the Kalman filter can be found in many references [4, 6, 8].

Consider the model as given in (1) and (2). The Kalman filtering assumes that both the modeling error noise $\mathbf{w}_k$ and the measurement disturbance $\mathbf{v}_k$ are Gaussian with the following covariance matrix,

$$E\big[\mathbf{w}_i\mathbf{w}_k^T\big] = \begin{cases} \mathbf{Q}_k, & \text{for} \quad i = k, \\ 0, & \text{for} \quad i \neq k. \end{cases} \tag{3}$$

$$E\big[\mathbf{v}_i\mathbf{v}_k^T\big] = \begin{cases} \mathbf{R}_k, & \text{for} \quad i = k, \\ 0, & \text{for} \quad i \neq k. \end{cases} \tag{4}$$

Let $\bar{\mathbf{x}}_k$ be the a priori estimate, which is the estimate of $\mathbf{x}_k$ given $\mathbf{z}_0, \mathbf{z}_1, \ldots, \mathbf{z}_{k-1}$, and let $\hat{\mathbf{x}}_k$ be the posterior estimate, which is the estimate of $\mathbf{x}_k$ given $\mathbf{z}_0, \mathbf{z}_1, \ldots, \mathbf{z}_k$. It is known that

$$E\{\bar{\mathbf{x}}_k\} = E\{\mathbf{x}_k\} \tag{5}$$

and

$$\bar{\mathbf{x}}_{k+1} = \mathbf{F}_k\hat{\mathbf{x}}_k, \tag{6}$$

where $E\{\cdot\}$ represents expectation and $\mathbf{F}_k$ is from (1).

The Kalman filter assumes that the posterior estimate is expressed as the a priori estimate corrected by the measurement data,

$$\hat{\mathbf{x}}_k = \bar{\mathbf{x}}_k + \mathbf{K}_k\left(\mathbf{z}_k - \mathbf{H}_k\bar{\mathbf{x}}_k\right), \tag{7}$$

for some n-by-m matrix $\mathbf{K}_k$ (so called Kalman gain). Note that $\mathbf{H}_k$ is from (2). The Kalman gain $\mathbf{K}_k$ is solved by minimizing $E\big[(\hat{\mathbf{x}}_k - \mathbf{x}_k)^2\big]$. Note that

$$E\big[(\hat{\mathbf{x}}_k - \mathbf{x}_k)^T(\hat{\mathbf{x}}_k - \mathbf{x}_k)\big] = \mathrm{Tr}\{\mathbf{P}_k\}, \tag{8}$$

where $\mathrm{Tr}\{\cdot\}$ represents the trace operator, and the n-by-n covariance matrix $\mathbf{P}_k$ is defined as follows

$$\mathbf{P}_k = E\big[(\mathbf{x}_k - \hat{\mathbf{x}}_k)(\mathbf{x}_k - \hat{\mathbf{x}}_k)^T\big]. \tag{9}$$

Define

$$\bar{\mathbf{P}}_k = E\big[(\mathbf{x}_k - \bar{\mathbf{x}}_k)(\mathbf{x}_k - \bar{\mathbf{x}}_k)^T\big]. \tag{10}$$

It follows from (3), (4), (9) and (10) that

$$\bar{\mathbf{P}}_{k+1} = \mathbf{F}_k\mathbf{P}_k\mathbf{F}_k^T + \mathbf{Q}_k. \tag{11}$$

Substituting (7) into (9), we get

$$\mathbf{P}_k = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\bar{\mathbf{P}}_k(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)^T + \mathbf{K}_k\mathbf{R}_k\mathbf{K}_k^T. \tag{12}$$

It follows from (12) that

$$\frac{d}{d\mathbf{K}_k}\,\mathrm{Tr}\{\mathbf{P}_k\} = -2\bar{\mathbf{P}}_k\mathbf{H}^T + 2\mathbf{K}_k(\mathbf{H}_k\bar{\mathbf{P}}_k\mathbf{H}_k^T + \mathbf{R}_k). \tag{13}$$

Solve $\mathbf{K}_k$ by letting $\frac{d}{d\mathbf{K}_k}\,\mathrm{Tr}\{\mathbf{P}_k\} = 0$, we get

$$\mathbf{K}_k = \bar{\mathbf{P}}_k\mathbf{H}_k^T(\mathbf{H}_k\bar{\mathbf{P}}_k\mathbf{H}_k^T + \mathbf{R}_k)^{-1} \tag{14}$$

By (12) and (14), $\mathbf{P}_k$ can be rewritten as

$$\mathbf{P}_k = (\mathbf{I} - \mathbf{F}_k\mathbf{H}_k)\bar{\mathbf{P}}_k. \tag{15}$$

Combining (6), (14) and (15), we thus have the conventional Kalman filter. This algorithm is shown in Figure 1.

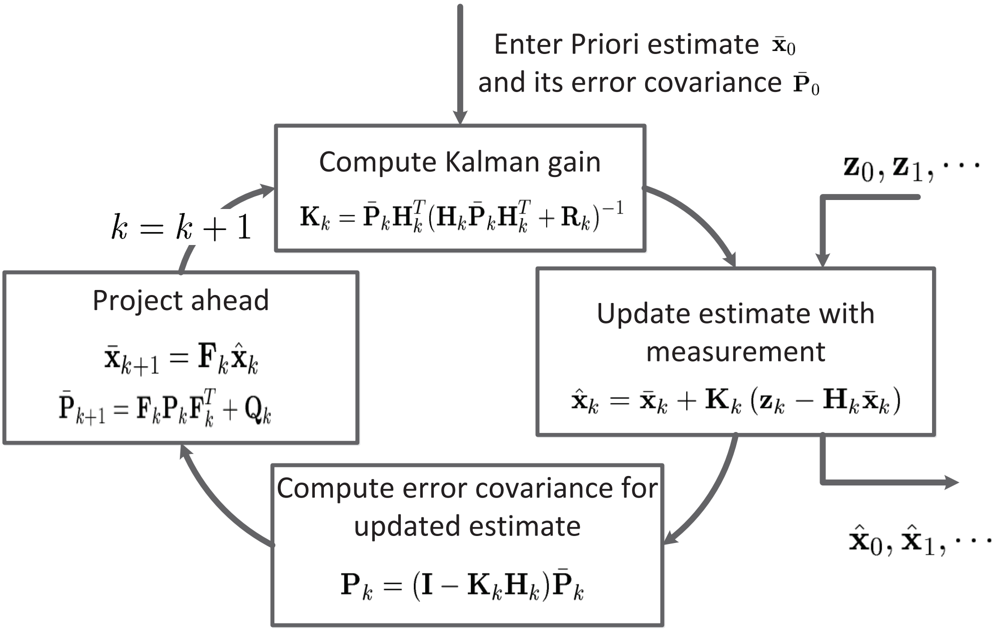

Fig. 1. The usual Kalman filtering algorithm

3. A modified Kalman filter. It is known that the discrete time Gaussian white noise can be approximated by the increments of Brownian motion per time step, and the non-Gaussian Lévy noise can be approximated by the increments of the corresponding Lévy process per time step. By Lévy–Itô theorem [2], a Lévy process can be decomposed into the sum of a Gaussian process and a pure jump process. It is shown in [3] that the small jumps of a Lévy process can be approximated by a Gaussian process. Therefore, we can approximately regard a Lévy process as combination of a Gaussian process and a process with big jumps. For more information about decomposition of a Lévy processes, see [3, 2]. These results enable us to decompose a non-Gaussian Lévy noise into a Gaussian noise plus some extremely large values.

In our proposed filtering method, we convert the original Lévy noise into a Gaussian noise by clipping off its extremely large values.

Let $\tilde{\mathbf{v}}_k$ represent the clipped version of the Lévy measurement disturbance $\mathbf{v}_k$, and let $\tilde{\mathbf{z}}_k$ represent the corresponding clipped observation. Thus

$$\tilde{\mathbf{z}}_k = \mathbf{H}_k \mathbf{x}_k + \tilde{\mathbf{v}}_k. \tag{16}$$

In practice, since the measurement noise, $\mathbf{v}_k$, is unknown, we propose to clip the observation $\mathbf{z}_k$ instead of $\mathbf{v}_k$ in a component-wise way by the following operation:

$$\begin{cases} \tilde{\mathbf{z}}_k^i = \sum_j \mathbf{H}_k^{i,j} \bar{\mathbf{x}}_k^j + C \cdot \operatorname{sign}\big(\mathbf{z}_k^i - \sum_j \mathbf{H}_k^{i,j} \bar{\mathbf{x}}_k^j\big) & \text{if} \quad \big|\mathbf{z}_k^i - \sum_j \mathbf{H}_k^{i,j} \bar{\mathbf{x}}_k^j\big| \geq C, \\ \tilde{\mathbf{z}}_k^i = \mathbf{z}_k^i & \text{if} \quad |\mathbf{z}_k^i - \sum_j \mathbf{H}_k^{i,j} \bar{\mathbf{x}}_k^j| < C, \end{cases} \tag{17}$$

where C is some positive threshold value, $\mathbf{z}_k^i$ and $\bar{\mathbf{x}}_k^i$ represent the i-th components of the vectors $\mathbf{z}_k$ and $\bar{\mathbf{x}}_k$, respectively, and $\sum_j \mathbf{H}_k^{i,j}\bar{\mathbf{x}}_k^j$ is the i-th component of the vector $\mathbf{H}_k\bar{\mathbf{x}}_k$. Note that C is determined by the statistical properties of the measurement noise $\mathbf{v}_k$. Replacing the observation value $\mathbf{z}_k$ in (7) with its clipped value, we get

$$\hat{\mathbf{x}}_k = \bar{\mathbf{x}}_k + \mathbf{K}_k\left(\tilde{\mathbf{z}}_k - \mathbf{H}_k\bar{\mathbf{x}}_k\right). \tag{18}$$

Repeating the same procedure as in Section 2 to solve the Kalman gain $\mathbf{K}_k$ by minimizing $E\{(\mathbf{x}_k - \hat{\mathbf{x}}_k)^2\}$, we get

$$\mathbf{K}_k = \bar{\mathbf{P}}_k\mathbf{H}_k^T(\mathbf{H}_k\bar{\mathbf{P}}_k\mathbf{H}_k^T + \tilde{\mathbf{R}}_k)^{-1}, \tag{19}$$

where $\tilde{\mathbf{R}}_k$ is the covariance matrix of $\tilde{\mathbf{v}}_k$ defined as

$$\tilde{\mathbf{R}}_k = E\big\{\tilde{\mathbf{v}}_k\tilde{\mathbf{v}}_k^T\big\}. \tag{20}$$

In the conventional Kalman filter, $\mathbf{Q}_k$ and $\mathbf{R}_k$ are assumed to be known, and as noted in [6], it is often a difficult task to estimate the covariance matrices $\mathbf{Q}_k$ and $\mathbf{R}_k$.

In the modified Kalman filter here, we only assume $\mathbf{Q}_k$ is known and suggest $\tilde{\mathbf{R}}_k$ be estimated as follows. It follows from (16) and (17) that

$$\begin{aligned}\tilde{\mathbf{R}}_k &= E\big\{\tilde{\mathbf{v}}_k\tilde{\mathbf{v}}_k^T\big\}\\ &= E\big\{\big[(\tilde{\mathbf{z}}_k - \mathbf{H}_k\bar{\mathbf{x}}_k) - \mathbf{H}_k(\mathbf{x}_k - \bar{\mathbf{x}}_k)\big]\big[(\tilde{\mathbf{z}}_k - \mathbf{H}_k\bar{\mathbf{x}}_k) - \mathbf{H}_k(\mathbf{x}_k - \bar{\mathbf{x}}_k)\big]^T\big\}\\ &= (\tilde{\mathbf{z}}_k - \mathbf{H}_k\bar{\mathbf{x}}_k)(\tilde{\mathbf{z}}_k - \mathbf{H}_k\bar{\mathbf{x}}_k)^T + \mathbf{H}_k\bar{\mathbf{P}}_k\mathbf{H}_k^T.\end{aligned} \tag{21}$$

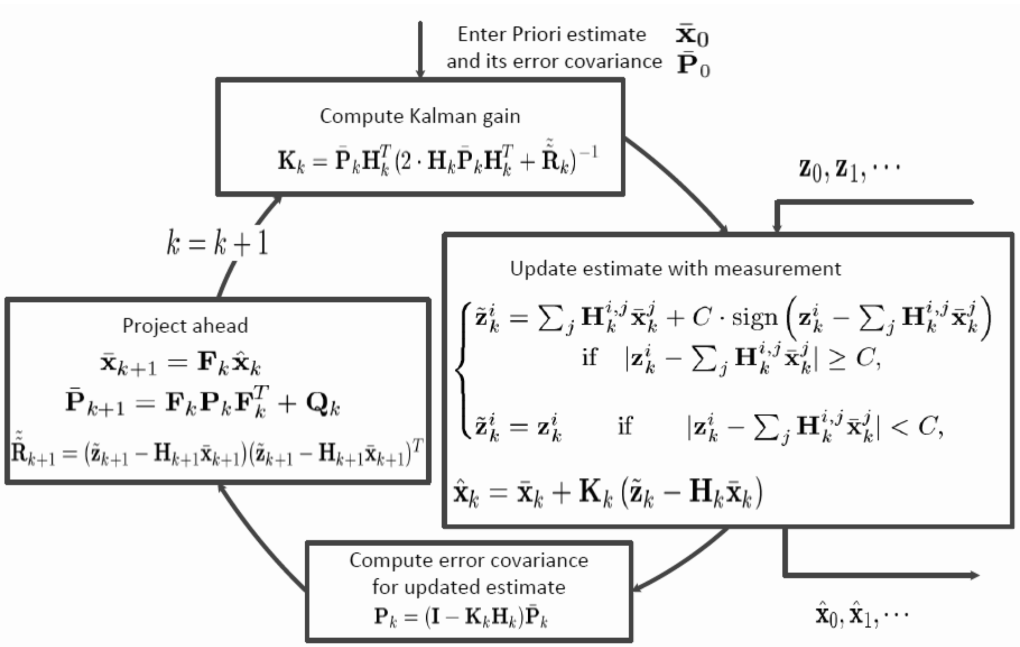

Fig. 2. The modified Kalman filtering algorithm

In deriving the last identity of (21), we have used the fact that $\tilde{\mathbf{z}}_k$ and $\bar{\mathbf{x}}_k$ are known values and

$$E\{(\tilde{\mathbf{z}}_k - \mathbf{H}_k\bar{\mathbf{x}}_k)(\mathbf{x}_k - \bar{\mathbf{x}}_k)^T\mathbf{H}_k^T\} = (\tilde{\mathbf{z}}_k - \mathbf{H}_k\bar{\mathbf{x}}_k)E\{(\mathbf{x}_k - \bar{\mathbf{x}}_k)^T\}\mathbf{H}_k^T = 0. \tag{22}$$

With (21), (19) can be rewritten as

$$\mathbf{K}_k = \bar{\mathbf{P}}_k\mathbf{H}_k^T(2\cdot\mathbf{H}_k\bar{\mathbf{P}}_k\mathbf{H}_k^T + \tilde{\tilde{\mathbf{R}}}_k)^{-1}, \tag{23}$$

where

$$\tilde{\tilde{\mathbf{R}}}_k = (\tilde{\mathbf{z}}_k - \mathbf{H}_k\bar{\mathbf{x}}_k)(\tilde{\mathbf{z}}_k - \mathbf{H}_k\bar{\mathbf{x}}_k)^T. \tag{24}$$

Combining equations (17), (18), (23), and (24), we obtain the modified Kalman filter, as shown graphically in Figure 2.

Compared with the conventional Kalman filter, the proposed filter has an moderately increased computational cost due to the following two operations: i) the clipping operation for $\mathbf{z}_k$; ii) the computation of $\tilde{\tilde{R}}_k$. The former operation is implemented by IF-ELSE sentence, and the latter is simply a vector-vector outer product.

Note that the prediction obtained in modified Kalman filter is not a conditional expectation.

4. Simulation results. In the simulation, we consider the fourth Navy tracking benchmark [11, 13] with emphasis on the first target of set 4. The model of the simulation is expressed by equations (1) and (2). The state vector at time k is given by

$$\mathbf{x}_k = \left[x_k^1, x_k^2, x_k^3, x_k^4, x_k^5, x_k^6\right]^T, \tag{25}$$

where x_k^1, x_k^3 and x_k^5 are the position coordinator of the target with respect to the radar, and x_k^2, x_k^4 and x_k^6 are the corresponding velocity components. The matrix F_k in equation (1) is given by

$$\mathbf{F}_k = \begin{pmatrix} 1 & T & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & T & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & T \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{26}$$

where T represents the sampling period, and the matrix $\mathbf{H}_k$ in (2) is given by

$$\mathbf{H}_k = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}. \tag{27}$$

In the simulation, the sampling period is taken as $T = 2$. It is assumed that the components of $\mathbf{w}_k$, denoted by w_k^1, w_k^2, w_k^3, w_k^4, w_k^5, w_k^6, are independent Gaussian white noises with zero mean and unit variance, and the components of $\mathbf{v}_k$, i.e., v_k^i $(i = 1, 2, \ldots, 6)$, are independent, and identically distributed noises consist of two components: i) a symmetric α-stable Lévy noises with the index of stability $\alpha = 1.3$ and the scale parameter $\sigma = 10$ (see [7]); ii) a Gaussian white noise with variance of 5. In the simulation, we are interested in estimating the position coordinate x_k^1, x_k^3, and x_k^5, which are the first, third, and fifth component of the state vector $\mathbf{x}_k$ as defined in (25), respectively. Since the measurement noises, $\mathbf{v}_k$, has infinite variances, the conventional Kalman filter cannot be applied to estimate the position coordinate x_k^1, x_k^3, and x_k^5. So we apply the modified Kalman filter proposed in the previous section.

Take the initial value of the state vector as $\mathbf{x}_0 = \begin{pmatrix}10 & 1 & 8 & 2 & 9 & 1\end{pmatrix}^T$, and apply the modified Kalman filtering method to estimate x_k, y_k and z_k. In the simulation, the initial a priori estimate of the state, $(\bar{\mathbf{x}}_k)$, is set to be equal to the observation at time 0, its error covariance, $\bar{P}_0$, is set to be unit matrix, and the threshold value C is set to be 40. The simulation results are shown in Figure 3, where the estimate position error, ER, defined by

$$ER_k = \sqrt{(\hat{x}_k^1 - x_k^1)^2 + (\hat{x}_k^2 - x_k^2)^2 + (\hat{x}_k^3 - x_k^3)^2}\,, \tag{28}$$

is compared with the observed position error, OR, defined by

$$OR_k = \sqrt{(z_k^1 - x_k^1)^2 + (z_k^2 - x_k^2)^2 + (z_k^3 - x_k^3)^2}\,. \tag{29}$$

The results in Figure 3 are calculated by averaging $10,000$ times of simulations. It is seen from this figure that the position estimation error is significantly improved by using our modified Kalman filter.

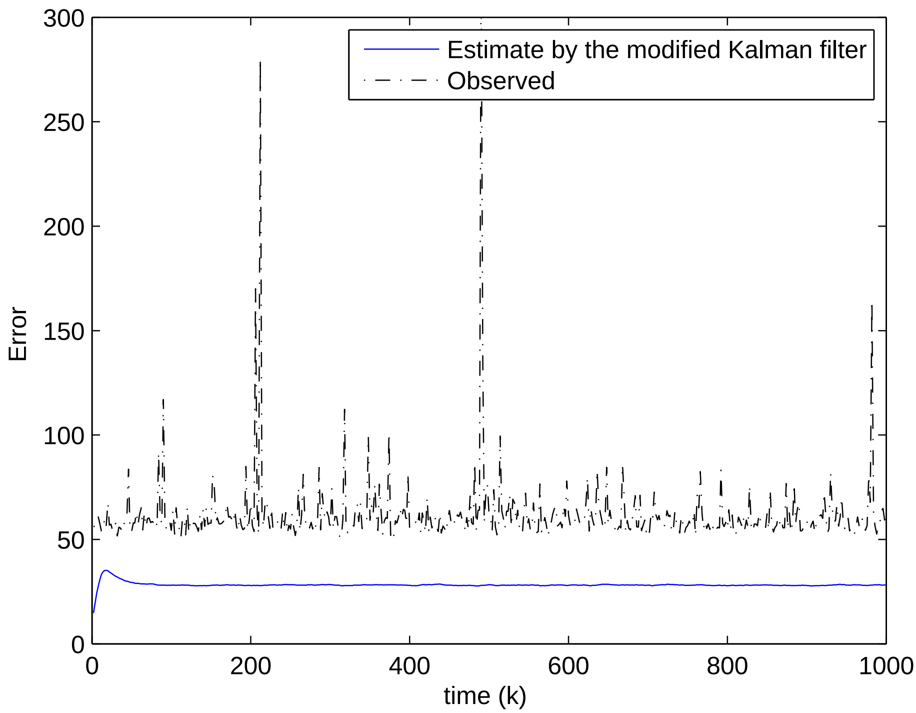

Fig. 3. The error of the modified Kalman filtering algorithm

In the simulation, we select C by the method of trial and error, and it is found that the threshold value C is not very picky and C can vary from 30 to 100 without significant effects on the performance of the modified Kalman filter. We propose to select C by based on the percentile of the estimated measurement errors. For example, it is effective in our simulation to select C such that 95th percent of the estimated errors, which is defined as $|\mathbf{z}_k^i - \sum_j \mathbf{H}_k^{i,j} \bar{\mathbf{x}}_k^j|$ in (17), is less than C.

References

[1] H. Ahn, R. Feldman, *Optimal filtering of a Gaussian signal in the presence of Lévy noise*, SIAM J. Appl. Math. 60 (2000), 359–369.

[2] D. Applebaum, *Lévy Processes and Stochastic Calculus*, second ed., Cambridge Stud. Adv. Math. 116, Cambridge Univ. Press, Cambridge 2009.

[3] S. Asmussen, J. Rosinski, *Approximation of small jumps of Lévy processes with a view towards simulation*, J. Appl. Probab. 38 (2001), 482–493.

[4] R. Brown, P. Hwang, *Introduction to Random Signals and Applied Kalman Filtering*, third ed., John Wiley & Sons, New York 1997.

[5] N. Gordon, J. Percival, M. Robinson, *The Kalman–Lévy filter and heavy-tailed models for tracking manoeuvring targets*, in: Information Fusion 2003, Proceedings of The Sixth International Conference of Information Fusion, IEEE, 2003, 1024–1031.

[6] M. Grewal, A. Andrews, *Kalman Filtering: Theory and Practice Using Matlab*, third ed., John Wiley & Sons, New York 2008.

[7] A. Janicki, A. Weron, *Simulation and Chaotic Behavior of α-Stable Stochastic Processes*, Monogr. Textbooks Pure Appl. Math. 178, Marcel Dekker, Inc., New York 1994.

[8] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*, Academic Press, New York 1970.

[9] A. Le Breton, M. Musiela, *A generalization of the Kalman filter to models with infinite variance*, Stochastic Process. Appl. 47 (1993), 75–94.

[10] C. L. Nikias, *Signal Processing with Alpha-Stable Distribution and Applications*, John Wiley & Sons, New York 2008.

[11] A. Sinha, T. Kirubarajan, Y. Bar-Shalom, *Application of the Kalman–Lévy filter for tracking maneuvering targets*, IEEE Transactions on Aerospace and Electronic Systems 43 (2007), 1099–1107.

[12] D. Sornette, K. Ide, *The Kalman–Lévy filter*, Physica D 151 (2001), 142–174.

[13] G. A. Watson, D. H. McCable, *Benchmark Problem with a Multisensor Framework for Radar Resource Allocation and the Tracking of Highly Maneuvering Targets, Closely Spaced Targets, and Targets in the Presence of Sea-Surface-Induced Multipath*, Naval Surface Warfare Center, Dahlgren, VA, Technical Report NSWC-DD/TR-99/32, Mar. 1999.

[14] W. A. Woyczynski, *Lévy processes in the physical sciences*, in: Lévy Processes: Theory and Applications, Birkhäuser, Boston 2001, 241–266.