

Research Article

Variational Inference of Kalman Filter and Its Application in Wireless Sensor Networks

Zijian Dong^{1,2} and Tiecheng Song²

¹ School of Electronic Engineering, Huaihai Institute of Technology, Lianyungang 222005, China

² School of Information Science and Engineering, Southeast University, Nanjing 210096, China

Correspondence should be addressed to Zijian Dong; dzjian@126.com

Received 28 September 2013; Accepted 15 October 2013

Academic Editor: Shuai Li

Copyright © 2013 Z. Dong and T. Song. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

An improved Kalman filter algorithm by using variational inference (VIKF) is proposed. With variational method, the joint posterior distribution of the states is approximately decomposed into several relatively independent posterior distributions. To avoid the difficulty of high-dimensional integrals, these independent posterior distributions are solved by using Kullback-Leibler divergence. The variational inference of Kalman filter includes two steps, the predict step and the update step, and an iterative process is included in the update step to get the optimized solutions of the posterior distribution. To verify the effectiveness of the proposed algorithm, VIKF is applied to the state estimation of discrete linear state space and the tracking problems in wireless sensor networks. Simulation results show that the variational approximation is effective and reliable for the linear state space, especially for the case with time-varying non-Gaussian noise.

1. Introduction

The Kalman filter (KF), also known as linear quadratic estimation (LQE), is one of the most used linear state estimation methods and has numerous applications [1], which include guidance, navigation, and control of vehicles [2–5]. Kalman filter is also widely applied in time series analysis in fields such as signal processing and econometrics [6, 7]. Other nonlinear methods, such as Extension-Kalman filter (EKF) [2] and unscented Kalman filter (UKF) [8], also use the Kalman filter to fix problems, combined with the linearization of nonlinear functions or the approximation of the probability density distribution, respectively. Practical implementation of the Kalman Filter is often difficult due to the inability of getting a good estimate of the noise covariance matrices. Extensive researches have been carried out in this field to estimate these covariances from data, such as the autocovariance least-squares (ALS) [9, 10], the modified Bryson-Frazier smoother [11], and the minimum-variance smoother [12, 13].

Variational inference for the Kalman filter in linear state space model was also considered in several literatures [14–16]. However, the authors in [14] just used the limited memory BFGS method to decrease the computational burden, both factorization and Kullback-Leibler (KL) divergence [15] (the

expectation of joint distribution with holding some variable distribution as constant) were not considered. Works in [16] indeed applied the variational approximation, and got the posterior distributions of the state variables and the measurement noise. However, in the update step of the Variational Bayesian approximation Adaptive Kalman Filter (VBAKF), the initial values of states or variances were used in all iterative steps; thus, the iterations would quickly reach “steady states.” Sometimes, the maximum number of the iterations is only 2 in the algorithm. That is to say that there are only two effective iterative loops performed in the update step of the Kalman filter, therefore, the performance of this algorithm is not good in such cases as the non-Gaussian and high-dimensional linear state space models. In this paper, we rederive the parameter expressions for the Kalman filter by using variational inference method (we call it VIKF in this paper), which can really form an iterative process for the update step of the Kalman filter.

Location and tracking are the keys in variety WSN applications [17–19], where the Kalman filter and its extension algorithms are widely used. In this paper, we will also consider the application of VIKF for WSN tracking problems, especially with time-varying and non-Gaussian noises.

This paper is organized as follows. We firstly introduce the linear state space model, then the derivation of Kalman filter by using variational inference (VIKF) follows, and the iterative process of VIKF is presented. In simulation section, the effectiveness of our proposed method will be demonstrated by examples. Finally, we will compare VIKF with standard KF and VBAKF in the cases of high-dimensional discrete state spaces, and the tracking in WSN with non-Gaussian noise.

2. The Linear State Space Model

Linear state space models are a widely used class of models for control, economics, and so on. The most general state-space representation of a discrete-time linear system with K samples, n state variables, and m outputs is written in the following form:

$$\mathbf{X}_k = \mathbf{M}_k \mathbf{X}_{k-1} + \boldsymbol{\varepsilon}_k^x, \quad (1)$$

$$\mathbf{Y}_k = \mathbf{N}_k \mathbf{X}_k + \boldsymbol{\varepsilon}_k^y, \quad (2)$$

where \mathbf{X}_k is an $n \times 1$ state vector of the system at time k , $k = 1, 2, \dots, K$. \mathbf{M}_k is the state transition matrix at time k . $\boldsymbol{\varepsilon}_k^x$ is the corresponding prediction error vector of \mathbf{X}_k , $\boldsymbol{\varepsilon}_k^x \sim N(0, \mathbf{Q}_k)$. $\boldsymbol{\varepsilon}_k^x$ is assumed to be independent and identically distributed (i.i.d), so \mathbf{Q}_k is a diagonal matrix. \mathbf{Y}_k is the observation vector, and we assume that there are m observations, so \mathbf{N}_k is an $m \times n$ linear observation operator matrix. $\boldsymbol{\varepsilon}_k^y$ is an $m \times 1$ observation error vector and assumed to be i.i.d; therefore, $\boldsymbol{\varepsilon}_k^y \sim N(0, \Gamma_k)$, $\Gamma_k = \text{diag}(\sigma_{k,1}^2, \dots, \sigma_{k,m}^2)$. Time is indexed by k .

In general, we assume that the initial state has a Gaussian prior distribution $\mathbf{X}_0 \sim N(\boldsymbol{\mu}_0, \mathbf{P}_0)$, and $\mathbf{M}_k, \mathbf{Q}_k, \mathbf{N}_k, \boldsymbol{\mu}_0$, and \mathbf{P}_0 are all known. In tracking problems, \mathbf{X}_k includes the information of positions and speeds of the moving target, $\mathbf{X}_k = [p_x, p_y, s_x, s_y]^T$, where (p_x, p_y) is the two-dimensional coordinate positions and (s_x, s_y) is the two-dimensional speeds.

3. Variational Inference for the Linear Space Model

Variational inference techniques have been extensively studied to solve problems in various fields [20–23]. There are two steps in the optimization for the standard Kalman filter [24]. The first step is to predict the state \mathbf{X}_k and Γ_k from the last state \mathbf{X}_{k-1} and the observations $\mathbf{Y}_{1:k-1}$. The second step is to update \mathbf{X}_k and Γ_k using information from new observations \mathbf{Y}_k .

3.1. Predict Step. This step calculates \mathbf{X}_k and Γ_k from $\mathbf{Y}_{1:k-1}$. Here we assume that $p(\mathbf{X}_k, \Gamma_k | \mathbf{Y}_{1:k-1})$ can be decomposed as

$$p(\mathbf{X}_k, \Gamma_k | \mathbf{Y}_{1:k-1}) = p(\mathbf{X}_k | \mathbf{Y}_{1:k-1}) p(\Gamma_k | \mathbf{Y}_{1:k-1}). \quad (3)$$

According to (1), we know that $p(\mathbf{X}_k | \mathbf{Y}_{1:k-1})$ is a Gaussian distribution:

$$p(\mathbf{X}_k | \mathbf{Y}_{1:k-1}) \sim N(\boldsymbol{\mu}'_k, \mathbf{P}'_k) \quad (4)$$

with

$$\boldsymbol{\mu}'_k = \mathbf{M}_k \boldsymbol{\mu}_{k-1}, \quad (5)$$

$$\mathbf{P}'_k = \mathbf{M}_k \mathbf{P}_{k-1} \mathbf{M}_k^T + \mathbf{Q}_k,$$

where \mathbf{P}_{k-1} is the variance of the distribution of last state \mathbf{X}_{k-1} .

$p(\Gamma_k | \mathbf{Y}_{1:k-1})$ is the posterior variance of observation errors. We can assume that $p(\Gamma_k | \mathbf{Y}_{1:k-1})$ equals $p(\Gamma_{k-1} | \mathbf{Y}_{1:k-1})$ temporarily in the prediction step, and it will be updated in the update step. We also assume that this posterior distribution is an inverse-Gamma distribution (also called Wald distribution). The inverse-Gamma distribution is often used in Bayesian inference, where the distribution arises as the marginal posterior distribution for the unknown variance of a normal distribution.

$p(\Gamma_k | \mathbf{Y}_{1:k-1})$ can be decomposed as

$$p(\Gamma_k | \mathbf{Y}_{1:k-1}) = \prod_{i=1}^m \text{IG}(\alpha'_{k,i}, \beta'_{k,i}), \quad (6)$$

where $\alpha'_{k,i}, \beta'_{k,i}$ are the shape and scale parameters of inverse Gamma distribution $\text{IG}(\cdot)$, and set

$$\alpha'_{k,i} = \alpha_{k-1,i}, \quad (7)$$

$$\beta'_{k,i} = \beta_{k-1,i}.$$

3.2. Update Step. This step updates \mathbf{X}_k and Γ_k based on the new observation \mathbf{Y}_k . According to the Bayes' rule,

$$p(\mathbf{X}_k, \Gamma_k | \mathbf{Y}_{1:k}) \propto p(\mathbf{Y}_k, \mathbf{X}_k, \Gamma_k | \mathbf{Y}_{1:k-1}). \quad (8)$$

We assume that $p(\mathbf{X}_k, \Gamma_k | \mathbf{Y}_{1:k})$ can be factored as

$$p(\mathbf{X}_k, \Gamma_k | \mathbf{Y}_{1:k}) = q(\mathbf{X}_k) q(\Gamma_k). \quad (9)$$

The variational method gets the approximation of each component by minimizing the Kullback-Leibler (KL) divergence between the estimated value and the true joint posterior. Each component can be obtained by performing the expectation operation to the joint distribution with respect to all other unknown parameters. So we have

$$\ln q_{\mathbf{X}_k}(\mathbf{X}_k) = E_{\Gamma_k}(\ln p(\mathbf{Y}_k, \mathbf{X}_k, \Gamma_k | \mathbf{Y}_{1:k-1})) + \text{const.}, \quad (10)$$

$$\ln q_{\Gamma_k}(\Gamma_k) = E_{\mathbf{X}_k}(\ln p(\mathbf{Y}_k, \mathbf{X}_k, \Gamma_k | \mathbf{Y}_{1:k-1})) + \text{const.} \quad (11)$$

According to the known priori, we calculate (10):

$$\begin{aligned} \ln q_{\mathbf{X}_k}(\mathbf{X}_k) &= E_{\Gamma_k}(\ln p(\mathbf{Y}_k, \mathbf{X}_k, \Gamma_k | \mathbf{Y}_{1:k-1})) + \text{const.} \\ &= E_{\Gamma_k}(\ln [p(\mathbf{Y}_k | \mathbf{X}_k, \Gamma_k) * p(\mathbf{X}_k | \mathbf{Y}_{1:k-1}) \\ &\quad * p(\Gamma_k | \mathbf{Y}_{1:k-1})]) + \text{const.} \\ &= E_{\Gamma_k}(\ln [p(\mathbf{Y}_k | \mathbf{X}_k, \Gamma_k) * p(\mathbf{X}_k | \mathbf{Y}_{1:k-1})]) \\ &\quad + \text{const.}, \end{aligned} \quad (12)$$

Input observation matrix $\mathbf{Y}_{1:m}$; set initial hyper parameters $\boldsymbol{\mu}_0, \mathbf{P}_0, \alpha_{0,i}, \beta_{0,i}, i = 1, \dots, m$, Set a small value to ξ , which is used to judge whether the loop of update step comes to an end or not.

Predict:

Calculate $\boldsymbol{\mu}'_k, \mathbf{P}'_k$ by (5);

Calculate $\alpha'_{k,i}, \beta'_{k,i}$ by (7)

Update:

Set $\boldsymbol{\mu}_k^{(0)} = \boldsymbol{\mu}'_k, \mathbf{P}_k^{(0)} = \mathbf{P}'_k, \Gamma' = \text{diag}(\beta_{k,1}/\alpha_{k,1}, \dots, \beta_{k,m}/\alpha_{k,m}), \alpha_{k,i} = 1/2 + \alpha'_{k,i}$;

Set Flag = 1;

While Flag

Update $\boldsymbol{\mu}_k$ by (15), Update \mathbf{P}_k by (16);

Update $\boldsymbol{\beta}_k$ by (20);

Calculate $\Gamma' = \text{diag}(\beta_{k,1}/\alpha_{k,1}, \dots, \beta_{k,m}/\alpha_{k,m})$ and $S = \sum_i (\boldsymbol{\mu}_{k,i}^{(s+1)} - \boldsymbol{\mu}_{k,i}^{(s)})^2$

If $S < \xi$, then set Flag = 0;

End flag

Output $\boldsymbol{\mu}_k$ in the last iteration as the inferred value \mathbf{X}_k .

ALGORITHM 1

where $E_{\Gamma_k}(\ln p(\Gamma_k | \mathbf{Y}_{1:k-1}))$ is a constant In (12), $p(\mathbf{Y}_k | \mathbf{X}_k, \Gamma_k)$ is a Gaussian distribution (see (2)), so we have

$$p(\mathbf{Y}_k | \mathbf{X}_k, \Gamma_k) \sim N(\mathbf{Y}_k - \mathbf{N}_k \mathbf{X}_k, \Gamma_k). \quad (13)$$

From the predict step, we know that $p(\mathbf{X}_k | \mathbf{Y}_{1:k-1})$ is also a Gaussian distribution $N(\boldsymbol{\mu}'_k, \mathbf{P}'_k)$ (see (4) and (5)). The product of two Gaussian functions is still a Gaussian, so

$$q_{\mathbf{X}_k}(\mathbf{X}_k) = N(\mathbf{X}_k | \boldsymbol{\mu}_k, \mathbf{P}_k). \quad (14)$$

Based on (4) and (13), we get the analytical expression of $\boldsymbol{\mu}_k$ and \mathbf{P}_k , and because variational method usually requires an iterative process to achieve convergence, so we expressed them as

$$\boldsymbol{\mu}_k^{(s+1)} = \boldsymbol{\mu}_k^{(s)} + \mathbf{P}_k^{(s)} \mathbf{N}_k^T (\mathbf{N}_k (\mathbf{P}_k^{(s)})^{-1} \mathbf{N}_k^T + \Gamma')^{-1} (\mathbf{Y}_k - \mathbf{N}_k \boldsymbol{\mu}_k^{(s)}), \quad (15)$$

$$\mathbf{P}_k^{(s+1)} = \mathbf{P}_k^{(s)} - \mathbf{P}_k^{(s)} \mathbf{N}_k^T (\mathbf{N}_k (\mathbf{P}_k^{(s)})^{-1} \mathbf{N}_k^T + \Gamma')^{-1} \mathbf{N}_k \mathbf{P}_k^{(s)}, \quad (16)$$

where the superscript (s) means the sth iteration, $\Gamma' = \text{diag}(\beta_{k,1}/\alpha_{k,1}, \dots, \beta_{k,m}/\alpha_{k,m}), \beta_{k,i}, \alpha_{k,i}, i = 1, \dots, m$ will be computed lately. At the beginning of the iteration, we use the results of the predict step, set $\boldsymbol{\mu}_k^{(0)} = \boldsymbol{\mu}'_k, \mathbf{P}_k^{(0)} = \mathbf{P}'_k$.

Similar results also appeared in [16] and the algorithm there was called VBAKF. In the iteration process, VBAKF computes $\boldsymbol{\mu}_k^{(s+1)}$ and $\mathbf{P}_k^{(s+1)}$ with $\boldsymbol{\mu}_k^{(s)}$ and $\mathbf{P}_k^{(s)}$, not $\boldsymbol{\mu}'_k$ and \mathbf{P}'_k ; therefore, VBAKF cannot form sufficient iteration during the update process. The maximum number of effective iterations of VBAKF is small, sometimes even only 2, so the performance cannot achieve the optimal solution of the variational method.

Similarly, we calculate $\ln q_{\Gamma_k}(\Gamma_k) = E_{\mathbf{X}_k}(\ln p(\mathbf{Y}_k, \mathbf{X}_k, \Gamma_k | \mathbf{Y}_{1:k})) + \text{const.}$ as

$$\begin{aligned} \ln q_{\Gamma_k}(\Gamma_k) &= E_{\mathbf{X}_k}(\ln p(\mathbf{Y}_k, \mathbf{X}_k, \Gamma_k | \mathbf{Y}_{1:k})) + \text{const.} \\ &= E_{\mathbf{X}_k}(\ln [p(\mathbf{Y}_k | \mathbf{X}_k, \Gamma_k) * p(\mathbf{X}_k | \mathbf{Y}_{1:k-1}) \\ &\quad * p(\Gamma_k | \mathbf{Y}_{1:k-1})]) + \text{const.} \quad (17) \\ &= E_{\mathbf{X}_k}(\ln [p(\mathbf{Y}_k | \mathbf{X}_k, \Gamma_k) * p(\Gamma_k | \mathbf{Y}_{1:k-1})]) \\ &\quad + \text{const.} \end{aligned}$$

Based on the known conditions, we know that $q_{\Gamma_k}(\Gamma_k)$ is still a Wald distribution, and can be expressed as

$$Q_{\Gamma_k}(\Gamma_k) = \prod_{i=1}^m \text{IG}(\sigma_{k,i}^2 | \alpha_{k,i}, \beta_{k,i}). \quad (18)$$

In the iterations, parameters are calculated as follows:

$$\alpha_{k,i}^{(s+1)} = \frac{1}{2} + \alpha_{k,i}^{(s)}, \quad (19)$$

$$\beta_{k,i}^{(s+1)} = \beta_{k,i}^{(s)} + \frac{1}{2} \left[(\mathbf{Y}_k - \mathbf{N}_k \mathbf{m}_k^{(s+1)})_i^2 + (\mathbf{N}_k \mathbf{P}_k^{(s+1)} \mathbf{N}_k^T)_{ii} \right]. \quad (20)$$

The initial values can be obtained from the predict step, so $\alpha_{k,i}^{(0)} = \alpha'_{k,i}, \beta_{k,i}^{(0)} = \beta'_{k,i}$.

We can see that (15), (16), (19), and (20) form an iterative process.

Variational method is an approximate calculation, not the best computation. Through simulations, we find that the performance of the algorithm is better if we set $\alpha_{k,i}^{(s+1)} = 1/2 + \alpha_{k,i}^{(0)}$. Therefore, $\alpha_{k,i}$ does not participate in the iterative process in our algorithm.

During the iteration, the iterative process should be executed with a sufficient number of times to reach a steady state. A simple stop condition for this is to test the value of

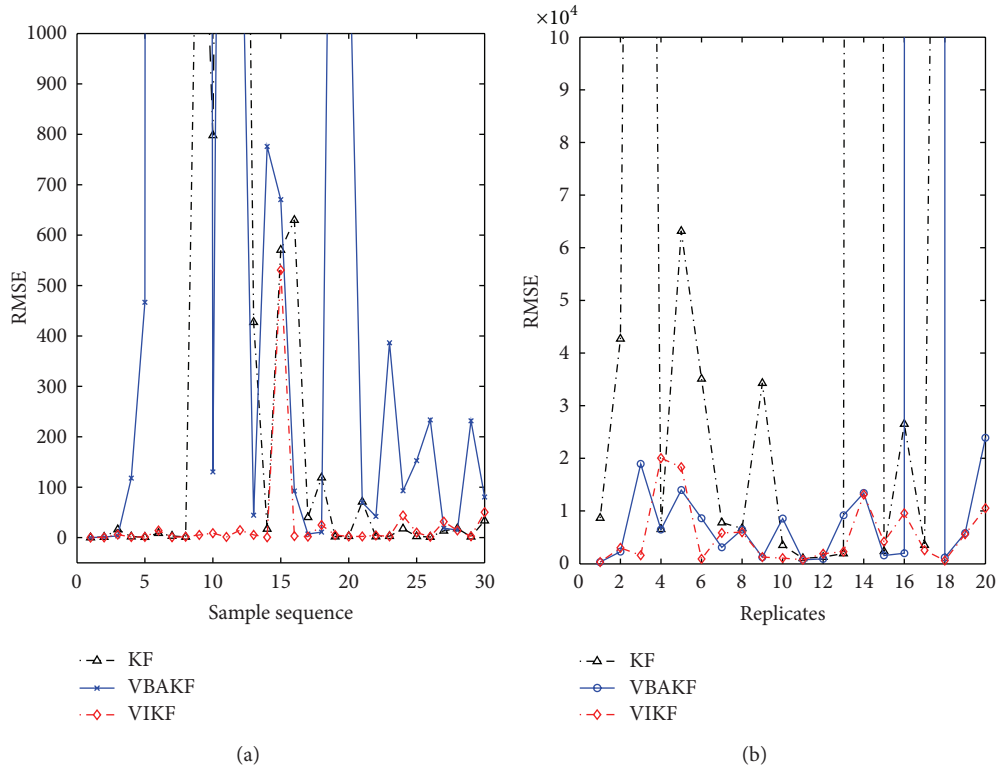


FIGURE 1: Performances of VIKF, VBAKF, and KF for the case of Student's t -distribution. (a) The RMSE of each state obtained by averaging the results of 20 replicates. (b) The accumulated RMSE of all 30 steps for each replicate.

the difference of inferred parameters between two successive iterations; that is, $S = \|\mu_k^{(t+1)} - \mu_k^{(t)}\|^2$. If the difference is small enough (say ξ), the iteration has reached a stable state. The selection for ξ can influence the accuracy of the algorithm. The smaller ξ is, the higher the accuracy is, at the cost of more iterations and more computational time needed.

3.3. The Algorithm of VIKF. VIKF includes two steps; the first step is to predict the present state with the initial information or the results of previous states, and the second step is the update step, which includes an iterative process to update the state with new observations.

The process of the algorithm is as shown in Algorithm 1.

4. Simulations

The performance of proposed algorithm (VIKF) was simulated and compared with that of the standard Kalman filter (KF) and the VBAKF. We first applied the three algorithms to high-dimensional discrete state space with the noise of Gaussian or Student's t -distributions. Residual mean square errors (RMSE) between the true state and the inferred state were calculated in all simulations and were used as the main evaluation criteria. We also applied these three algorithms to the tracking problems in the wireless sensor networks. In these simulations, time-varying parameters are considered for the noise of Student's t -distribution.

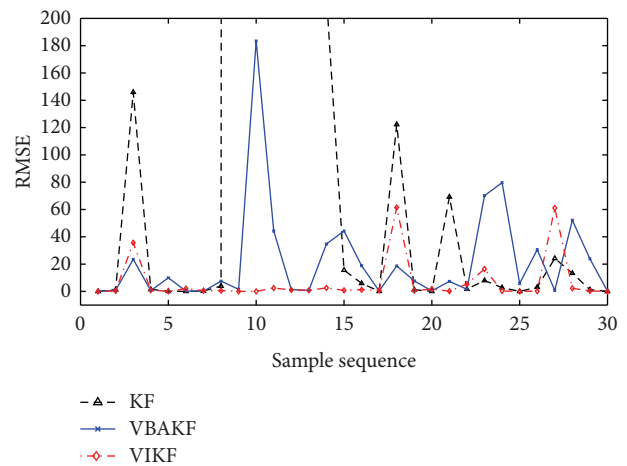


FIGURE 2: A typical simulation result with Student's t -distribution in high-dimensional linear state space.

4.1. High-Dimensional Discrete State Space Simulation. We considered a 40-dimensional discrete linear state space; the dimension of observation vector was 80, and there are 30 samples. State transition matrix and observation operator matrix were all sparse matrices. In each simulation case, 20 replicates were conducted and the results were averaged to get the performance of RMSE.

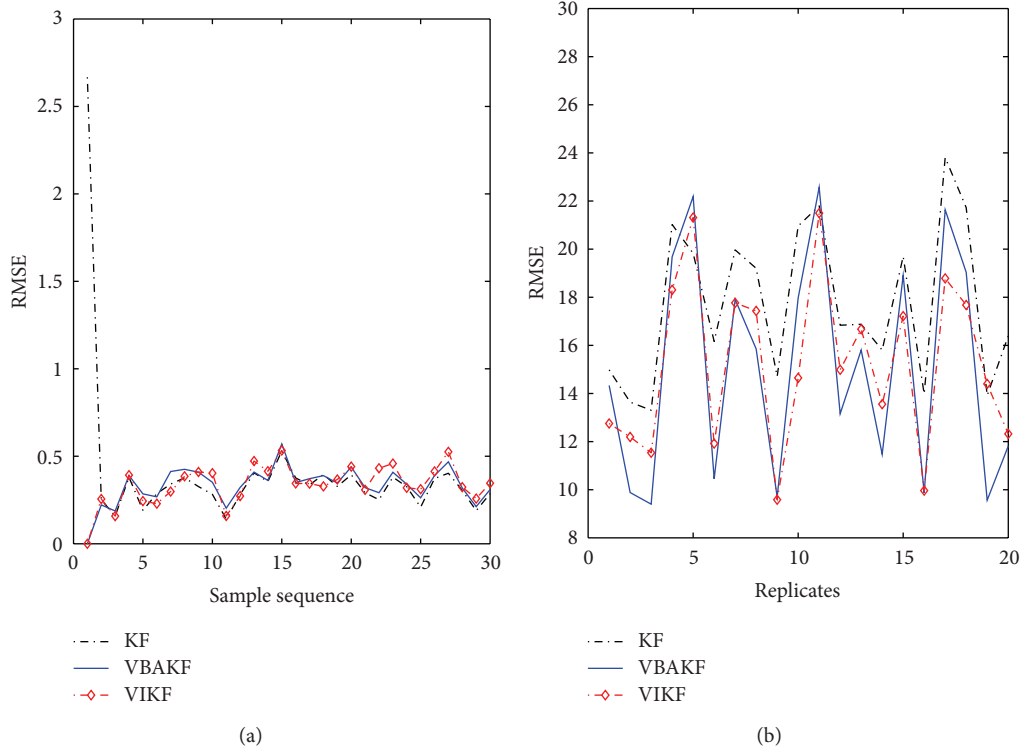


FIGURE 3: Performances of VIKF, VBAKF, and KF for the case of Gaussian distribution. (a) The RMSE of each state obtained by averaging the results of 20 replicates. (b) The accumulated RMSE of all 30 steps for each replicate.

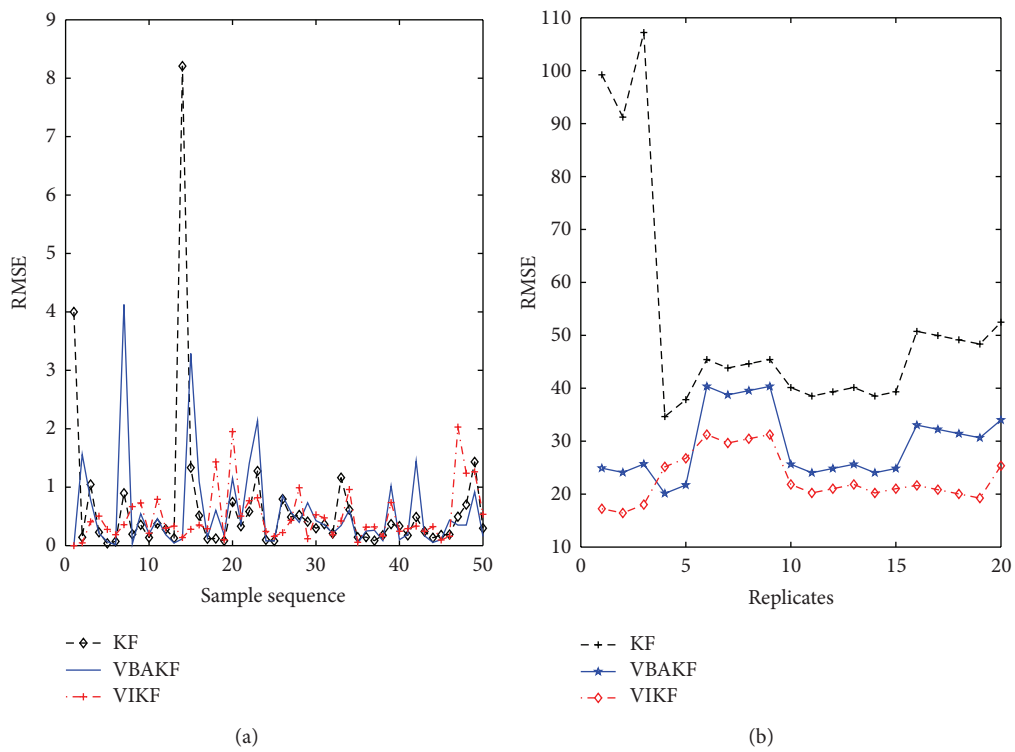


FIGURE 4: Performances of VIKF, VBAKF, and KF for the case of Student's t -distribution. (a) The RMSE of each state obtained by averaging the results of 20 replicates. (b) The accumulated RMSE of all 50 steps for each replicate.

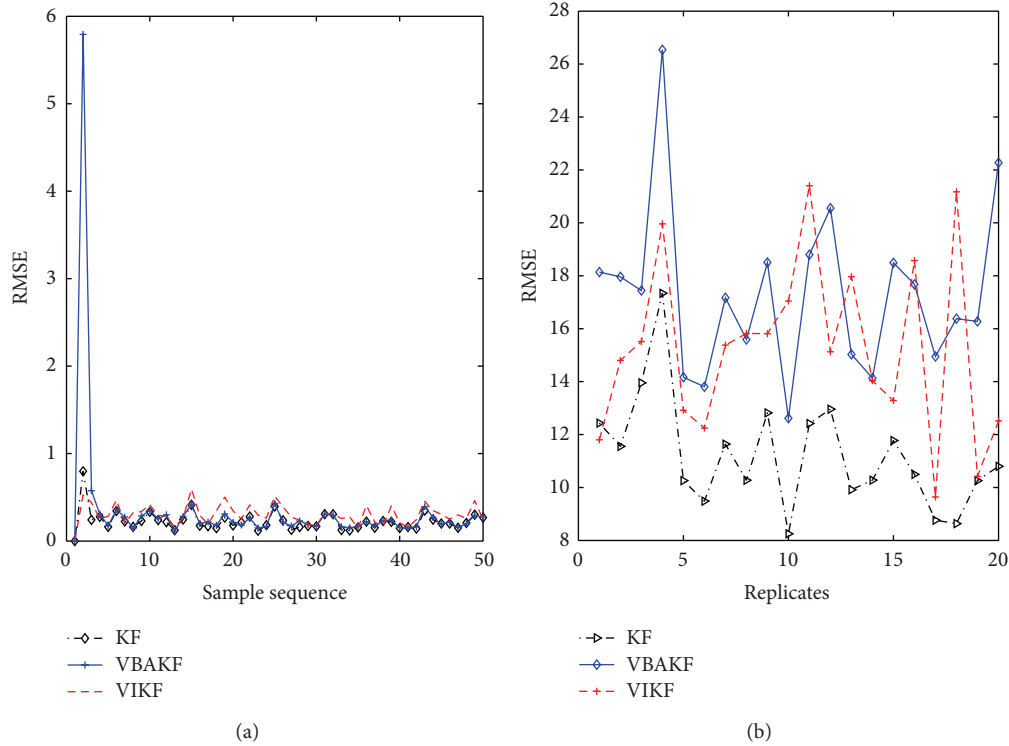


FIGURE 5: Performances of VIKF, VBAKF, and KF for the case of Gaussian distribution. (a) The RMSE of each state obtained by averaging the results of 20 replicates. (b) The accumulated RMSE of all 50 steps for the 20 replicates.

In the case of Gaussian noise, we assumed that the process noise and the observation noise were all subject to $N(0, 1)$. In the simulations, the freedom of Student's t -distribution was 2 for the process noise, while the freedom was selected randomly from set $(1, 2, 3, 4)$ for the observation noise. The noise intensities were 1 for all the cases in simulations.

The RMSE performances of VIKF, VBAKF, and KF are shown in Figure 1, Figure 2 (for Student's t -distribution), and Figure 3 (for Gaussian distribution). The results shown in Figure 1(a) show the RMSE of each state, which are obtained by averaging the results of the 20 replicates. Figure 1(b) shows the accumulated RMSE of all steps for each replicate. A typical simulation result is also shown in Figure 2. In these figures, we restrict the display range of the Y-axis to show the results clearly, so some points with high RMSE cannot be seen. We can see that the performance of VIKF is clearly higher than that of VBAKF and KF in the case of noise with Student's t -distribution. KF and VBAKF all have a higher probability of occurring extreme RMSE in certain circumstances, while the VIKF algorithm has a relatively lower probability. The reason that VIKF can decrease the probability of occurring large RMSE is that VIKF can perform sufficient iterations.

In the case of Gaussian noise, the performances of these three algorithms are close, as shown in Figure 3.

4.2. Tracking in WSN. The state space for the tracking problem in WSN was fixed as in Section 4, included two-dimensional coordinates and two-dimensional velocities. As the simulations in Section 4.1, we considered two kinds of

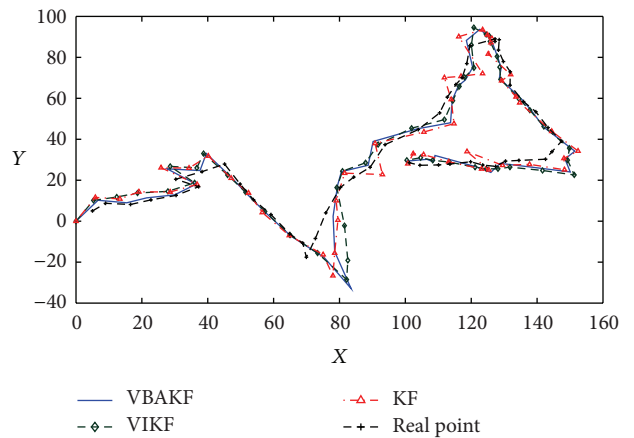


FIGURE 6: Tracking performances of VIKF, VBAKF, and KF with Student's t -distribution noise.

noise: Gaussian distribution and Student's t -distribution. We assume that there are 15 sensor nodes and 50 sampling points. Noise was set as that in Section 4.1.

As Figures 1 and 3, Figure 4 shows the RMSE performances with Student's t -distribution noise, where Figure 4(a) shows the RMSE of each state obtained by averaging the results of 20 replicates and Figure 4(b) shows the accumulated RMSE of all steps for each replicate. Figure 5 shows the corresponding results with Gaussian noise. Figure 6 shows

the tracking performance using the three algorithms with noise of Student's t -distribution.

With the noise of Student's t -distribution, the performance of VIKF in tracking is higher than that of VBAKF and KF. VIKF's probability of high RMSE is also lower than KF and VBAKF. In Gaussian cases, it seems that the performance of KF is better than that of VBIKF and VIKF.

5. Conclusions

This paper used variational approximation method to solve the problems arisen from high-dimensional discrete state space and the tracking problems in WSN with linear state model. We proposed a modified variational filtering algorithm, in which an iterative process was formed to get the optimization solutions.

The performance of variational methods is slightly lower than that of KF in a Gaussian environment. But in non-Gaussian and time-varying noise environment, the average performances of VIKF and VBAKF are better than that of KF. Since VIKF can carry out more sufficient iterations than the VBAKF in the update step, the performance of VIKF is better than that of VBAKF. VIKF algorithm also has a lower probability of high RMSE than KF and VBAKF in non-Gaussian environments.

Acknowledgments

This work was jointly supported by The National Natural Science Foundation of China (No. 61271207, No. 61174013), the Natural Science Foundation of Jiangsu Province, China (No. BK2011398), the Jiangsu Overseas Research & Training Program for University Prominent Young & Middle-aged Teachers and Presidents, and the Priority Academic Program Development of Jiangsu Higher Education Institutions.

References

- [1] S. L. Lauritzen, "Time series analysis in 1880: a discussion of contributions made by TN Thiele," *International Statistical Review*, vol. 49, pp. 319–331, 1981.
- [2] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation: Theory Algorithms and Software*, John Wiley & Sons, New York, NY, USA, 2004.
- [3] G. Einicke, J. Ralston, C. Hargrave, D. C. Reid, and D. Hainsworth, "Longwall mining automation: an application of minimum-variance smoothing [Applications of Control]," *IEEE Control Systems*, vol. 28, no. 6, pp. 28–37, 2008.
- [4] R. S. Bucy and P. D. Joseph, *Filtering for Stochastic Processes with Applications to Guidance*, vol. 326, The American Mathematical Society, 1987.
- [5] P. D. Groves, *Principles of GNSS, Inertial, and Multisensor Integrated Navigation Systems*, Artech House, 2008.
- [6] K. Paliwal and A. Basu, "A speech enhancement method based on Kalman filtering," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '87)*, pp. 177–180, 1987.
- [7] A. C. Harvey, "Application of the Kalman filter in econometrics," in *Advances in Econometrics: Fifth World Congress*, Cambridge University Press, Cambridge, Mass, USA, 1987.
- [8] E. A. Wan and R. van der Merwe, "The unscented Kalman filter," in *Kalman Filtering and Neural Networks*, pp. 221–280, 2001.
- [9] M. R. Rajamani, *Data-Based Techniques to Improve State Estimation in Model Predictive Control*, ProQuest, 2007.
- [10] M. R. Rajamani and J. B. Rawlings, "Estimation of the disturbance structure from data using semidefinite programming and optimal weighting," *Automatica*, vol. 45, no. 1, pp. 142–148, 2009.
- [11] G. J. Bierman, *Factorization Methods for Discrete Sequential Estimation*, Courier Dover, Mineola, NY, USA, 2006.
- [12] G. A. Einicke, "Optimal and robust noncausal filter formulations," *IEEE Transactions on Signal Processing*, vol. 54, no. 3, pp. 1069–1077, 2006.
- [13] G. A. Einicke, "Asymptotic optimality of the minimum-variance fixed-interval smoother," *IEEE Transactions on Signal Processing*, vol. 55, no. 4, pp. 1543–1547, 2007.
- [14] H. Auvinen, J. Bardsley, H. Haario, and T. Kauranne, "The variational Kalman filter and an efficient implementation using limited memory BFGS," *International Journal for Numerical Methods in Fluids*, vol. 64, no. 3, pp. 314–335, 2010.
- [15] J.-F. Cardoso, "Infomax and maximum likelihood for blind source separation," *IEEE Signal Processing Letters*, vol. 4, no. 4, pp. 112–114, 1997.
- [16] S. Sarkka and A. Nummenmaa, "Recursive noise adaptive Kalman filtering by variational Bayesian approximations," *IEEE Transactions on Automatic Control*, vol. 54, no. 3, pp. 596–600, 2009.
- [17] S. Li and F. Qin, "A dynamic neural network approach for solving nonlinear inequalities defined on a graph and its application to distributed, routing-free, range-free localization of WSNs," *Neurocomputing*, vol. 117, pp. 72–80, 2013.
- [18] S. Li, B. Liu, B. Chen, and Y. Lou, "Neural network based mobile phone localization using Bluetooth connectivity," *Neural Computing and Applications*, vol. 23, no. 3–4, pp. 667–675, 2013.
- [19] B. Kusý, I. Amundson, J. Sallai, P. Völgyesi, A. Lédeczi, and X. Koutsoukos, "RF doppler shift-based mobile sensor tracking and navigation," *ACM Transactions on Sensor Networks*, vol. 7, no. 1, article 1, 2010.
- [20] Z. Ghahramani and M. J. Beal, "Variational inference for Bayesian mixtures of factor analysers," in *Advances in Neural Information Processing Systems*, vol. 12, pp. 449–455, MIT Press, 2000.
- [21] C. W. Fox and S. J. Roberts, "A tutorial on variational Bayesian inference," *Artificial Intelligence Review*, vol. 38, no. 2, pp. 85–95, 2012.
- [22] M. W. Seeger, "Sparse linear models: variational approximate inference and Bayesian experimental design," *Journal of Physics: Conference Series*, vol. 197, no. 1, 2009.
- [23] M. J. Beal, *Variational algorithms for approximate Bayesian inference [Ph.D. thesis]*, University of London, 2003.
- [24] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*, Dover, Mineola, NY, USA, 2007.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

