2018 21st International Conference on Information Fusion (FUSION)

Improved Adaptive Kalman Filter With Unknown Process Noise Covariance

Jirong Ma[1,2], Hua Lan[1,2], Zengfu Wang[1,2], Xuezhi Wang[3], Quan Pan[1,2], and Bill Moran[4]

[1]School of Automation, Northwestern Polytechnical University, China
[2]Key laboratory of Information Fusion Technology, Ministry of Education, China
Email: mjr@mail.nwpu.edu.cn, {lanhua, wangzengfu, quanpan}@nwpu.edu.cn
[3]School of Engineering, RMIT University, Australia
Email: xuezhi.wang@rmit.edu.au
[4]Dept. Electrical and Electronic Engineering, University of Melbourne, Australia
Email: wmoran@unimelb.edu.au

***Abstract*—This paper considers the joint recursive estimation of the dynamic state and the time-varying process noise covariance for a linear state space model. The conjugate prior on the process noise covariance, the inverse Wishart distribution, provides a latent variable. A variational Bayesian inference framework is then adopted to iteratively estimate the posterior density functions of the dynamic state, process noise covariance and the introduced latent variable. The performance of the algorithm is demonstrated with simulated data in a target tracking application.**

***Index Terms*—Adaptive filtering, unknown process noise covariance, variational Bayesian, conjugate priors**

I. Introduction

Linear state-space models are almost ubiquitously used to represent real world dynamical systems, where possible system disturbance or time-evolution variation is modeled by system process noise. When the Kalman filter (KF) is exploited to estimate the state of such systems, a lack of knowledge of process noise statistics is known to result in estimation error [1]. Such a situation occurs, for example, in autonomous navigation and target tracking, where large uncertainty of system evolution may appear and the process noise covariance varies with time. Adaptive filtering, which estimates both the unknown model parameters and system dynamic state simultaneously from measurements, is an ideal approach to address this issue [2]. State-of-the-art adaptive filtering approaches may be categorized into four methods, Bayesian, maximum likelihood, correlation, and covariance matching [1]. State augmentation [3], interactive multiple models [4] and particle filters [5] are the most well-known Bayesian approaches. The safe-husa sage-husa adaptive KF of [6] is a covariance matching method, which estimates process noise covariance and measurement noise covariance sequentially.

Variational Bayesian inference (VB) is a closed-loop iterative method that is often able to turn difficult inference problems into optimization problems, and has lower computational cost compared to, for instance, a randomized sampling method [7], [8]. In recent years, VB has been applied to adaptive filtering through joint state and noise covariance estimation. For the linear state space model, Sakka and Nummenmaa [2] considered an unknown diagonal measurement noise covariance matrix, each element of which is assumed to follow an inverse Gamma distribution, as the conjugate prior for the variance of a Gaussian distribution. Then, an adaptive KF method based on VB (VBAKF) was used for the joint estimation of the dynamic state and the measurement noise covariance. This work was further extended in [9], where an inverse Wishart prior is used for a general (not diagonal) measurement covariance matrix with nonlinear system dynamics, and in [10], where a t-distribution rather than a Gaussian is used for the measurement distribution to improve outlier elimination. The latter paper also uses a Rauch-Tung-Striebel smoother for state estimation. However, extending the work of [9] to consider unknown process noise covariance is not easy since the process noise covariance does not appear in as straightforward conjugate prior form as the measurement noise [9].

Exactly!

Ardeshiri *et al.* [11] presented a batch-processing VB algorithm for joint estimation of the dynamic system state, the measurement noise covariance and the process noise covariance, with the noise covariance matrices being identified off-line. Instead of estimating the process noise covariance matrix directly, Huang *et al.* [12] proposed a novel VBAKF in which the inverse Wishart distribution was used as a prior for the predicted error covariance matrix and measurement noise covariance matrix, and inferred the system state with the unknown covariance matrices of process noise and measurement noise. The approach of [12] is an on-line recursive method but it needs a nominal process noise covariance matrix at each time step as the parameter of the algorithm.

In this paper, we develop a novel VB based adaptive KF algorithm, referred as VBAKF-Q, for state estimation with unknown process noise covariance. By introducing a new latent variable, the conjugate prior distribution of the underlying process noise covariance is assumed to follow the inverse Wishart distribution. Thus, the problem of joint

978-0-9964527-7-9 ©2018 ISIF

state estimation and process noise covariance identification is reformulated as the problem of estimating the system state together with the unknown process noise covariance and latent variable. The posterior probability density functions (PDFs) of the system state, the process noise covariance and the latent variable are updated iteratively in a closed-form VB for each measurement cycle. The proposed VBAKF-Q is compared with the VBAKF-P method of [12].

II. Problem Formulation

Consider the following linear state space model

向量与标量表示不明确

model

$$x_k = F_k x_{k-1} + w_k, \tag{1}$$
$$y_k = H_k x_k + v_k, \tag{2}$$

where $x_k \in \mathbb{R}^{n_x}$ is the system state and $y_k \in \mathbb{R}^{n_y}$ is the sensor measurement at time k, with n_x and n_y the dimensions of the state and the sensor measurement, respectively. The process noise w_k is a zero mean Gaussian process with unknown covariance matrix Q_k and the measurement noise v_k is a zeros mean Gaussian process with known covariance matrix R_k. The initial system state is assumed to follow a Gaussian distribution with known mean $x_{0|0}$ and covariance $P_{0|0}$. The state transition matrix F_k and the measurement matrix H_k are assumed to be known.

Optimal Bayesian filtering of the above linear state space model with unknown Q_k computes the joint posterior distribution $p(x_k, Q_k|y_{1:k})$, and consists of the following prediction and update steps.

The predictive distribution of the system state x_k and the process noise covariance Q_k is given by the Chapman-Kolmogorov equation

prediction

$$p(x_k, Q_k|y_{1:k-1}) = \int p(x_k|x_{k-1}, Q_k)p(Q_k|Q_{k-1}) \times p(x_{k-1}, Q_{k-1}|y_{1:k-1})\mathrm{d}x_{k-1}\mathrm{d}Q_{k-1}. \tag{3}$$

Given the measurement y_k, the predictive distribution of Eq. (3) is updated to the posterior distribution according to the Bayes rule:

update

$$p(x_k, Q_k|y_{1:k}) \propto p(y_k|x_k, Q_k)p(x_k, Q_k|y_{1:k-1}). \tag{4}$$

The main difficulty of state estimation with an unknown process noise covariance arises from the intractable integration of Eq. (3) that often has no closed-form analytical solution. VB provides an effective approximate solution for complicated integration problems but, to do this, requires appropriate conjugate prior models of latent variables.

The prior distribution for the system state is assumed to be Gaussian, i.e.,

$$p(x_k|y_{1:k-1}, Q_k) = \mathcal{N}(x_k|x_{k|k-1}, P_{k|k-1}), \tag{5}$$

where $x_{k|k-1} = F_k x_{k-1|k-1}$ and $P_{k|k-1} = F_k P_{k-1|k-1} F_k^T + Q_k$. The conjugate prior for the covariance of a Gaussian distribution is an inverse Wishart distribution, but Q_k is only a part of the covariance of the Gaussian distribution of x_k.

共轭先验分布

inverse Gamma仅适用于Qk为对角阵，inverse Wishart也适用于非对角阵。

假设$X \in R^{n*p}$，$X_i \backsim N_p(0,\Sigma)$，$\Sigma \backsim W_p^{-1}(m,\Omega)$
那么Σ后验分布：

$$\Sigma|data \backsim W_p^{-1}(m+n, A+\Omega), \quad A = \sum_i X_i^T X_i = nS$$

Wishart:多元正态分布的协方差矩阵
Inverse-Wishart分布常作为Bayes中多元正态分布的协方差阵的共轭先验分布

In order to decompose $P_{k|k-1}$, we introduce an intermediate latent variable θ_k, and rewrite Eq. (5) as

中间隐变量

$$\mathcal{N}(x_k|x_{k|k-1}, P_{k|k-1}) = \int \mathcal{N}(x_k|\theta_k, Q_k) \times \mathcal{N}(\theta_k|x_{k|k-1}, \Sigma_k)\mathrm{d}\theta_k, \tag{6}$$

预测满足正态分布

where $\Sigma_k = F_k P_{k-1|k-1} F_k^T$.

Now, Q_k is the covariance of the Gaussian distribution and we can use the inverse Wishart distribution as its prior, i.e.,

$$p(Q_k|y_{1:k-1}) = \mathrm{IW}(Q_k|u_{k|k-1}, U_{k|k-1}), \tag{7}$$

where $\mathrm{IW}(Q|u, U)$ signifies that Q follows an inverse Wishart distribution with degrees of freedom parameter u and inverse scale matrix U. The initial process noise covariance is also assumed to follow an inverse Wishart distribution $\mathrm{IW}(Q_0|u_{0|0}, U_{0|0})$. The mean value for Q_0 is [13]

$$\bar{Q}_0 = \frac{U_{0|0}}{u_{0|0} - n_x - 1}. \tag{8}$$

Thus, the initial parameters are chosen to be

$$u_{0|0} = \tau + n_x + 1, \quad U_{0|0} = \tau\bar{Q}_0, \tag{9}$$

where τ is the tuning parameter.

In a similar way to [9], [12] , the dynamic model $p(Q_k|Q_{k-1})$ is defined as

$$u_{k|k-1} = \rho(u_{k-1} - n_x - 1) + n_x + 1, \tag{10}$$
$$U_{k|k-1} = \rho U_{k-1}, \tag{11}$$

where the decreasing factor $\rho \in [0, 1]$.

In the following, we solve the state estimation with unknown process noise covariance by using VB to update the PDFs of x_k, θ_k and Q_k iteratively.

III. Variational Bayesian Approximation

In order to estimate the system state x_k, the unknown process noise covariance Q_k, and latent variable θ_k, we need to compute the joint posterior PDF $p(x_k, \theta_k, Q_k|y_{1:k})$. This is usually not analytically tractable, but VB, which uses a tractable PDF $q(\cdot)$ to approximate the intractable $p(\cdot)$, is able to solve the inference problem. By the mean-field approximation [8], we have

$$\begin{aligned} p(x_k, \theta_k, Q_k|y_{1:k}) &\approx q(x_k, \theta_k, Q_k) \\ &\approx q(x_k)q(\theta_k)q(Q_k). \end{aligned} \tag{12}$$

VB transforms the inference problem into the optimization problem of minimizing the Kullback-leibler (KL) divergence between $p(x_k, \theta_k, Q_k|y_{1:k})$ and $q(x_k, \theta_k, Q_k)$. Because of the non-negative property of the KL divergence, the minimization problem is equivalent to finding $q(x_k, \theta_k, Q_k)$ that maximizes the evidence lower bound (ELBO) $\mathcal{B}(q)$ of $\log p(y_k)$ [8], i.e., $q^*(x_k, \theta_k, Q_k) = \arg\max_{q(x_k,\theta_k,Q_k)} \mathcal{B}(q)$ where

$$\mathcal{B}(q) = \mathbb{E}\left[\log \frac{p(x_k, \theta_k, Q_k, y_k|y_{1:k-1})}{q(x_k, \theta_k, Q_k)}\right]. \tag{13}$$

Define $f \overset{c}{=} g$ if $f = g + c$, where c is an additive constant. By using coordinate ascent, we have

$$\log q(x_k) \overset{c}{=} \mathbb{E}_{q(\theta_k,Q_k)}\left[\log p(x_k,\theta_k,Q_k,y_k|y_{1:k-1})\right], \quad (14)$$

$$\log q(\theta_k) \overset{c}{=} \mathbb{E}_{q(x_k,Q_k)}\left[\log p(x_k,\theta_k,Q_k,y_k|y_{1:k-1})\right], \quad (15)$$

$$\log q(Q_k) \overset{c}{=} \mathbb{E}_{q(x_k,\theta_k)}\left[\log p(x_k,\theta_k,Q_k,y_k|y_{1:k-1})\right]. \quad (16)$$

The joint PDF $\mathcal{L}_k = \log p(x_k,\theta_k,Q_k,y_k|y_{1:k-1})$ can be decomposed as

$$\begin{aligned}\mathcal{L}_k =& \log p(x_k|y_{1:k-1},\theta_k,Q_k) + \log p(\theta_k|y_{1:k-1}) \\ &+ \log p(Q_k|y_{1:k-1}) + \log p(y_k|x_k) \\ =& \log \mathcal{N}(x_k|\theta_k,Q_k) + \log \mathcal{N}(\theta_k|x_{k|k-1},\Sigma_k) \\ &+ \log \mathrm{IW}(Q_k|u_{k|k-1},U_{k|k-1}) + \log \mathcal{N}(y_k|H_k x_k,R_k).\end{aligned} \quad (17)$$

Next we derive the posterior PDFs $q(x_k)$, $q(\theta_k)$ and $q(Q_k)$.

A. Derivations of the approximate posterior $q(x_k)$

By the definition in Eq. (14) and the joint PDF of Eq. (17), we have

$$\begin{aligned}\log q(x_k) \overset{c}{=}& \mathbb{E}_{q(\theta_k,Q_k)}[\log \mathcal{N}(x_k|\theta_k,Q_k) \\ &+ \log \mathcal{N}(y_k|H_k x_k,R_k)] \\ =& -\frac{1}{2}\mathrm{tr}\left(\mathbb{E}[Q_k^{-1}]\mathbb{E}[(x_k-\theta_k)(x_k-\theta_k)^T]\right) \\ &+ \log \mathcal{N}(y_k|H_k x_k,R_k) \\ \overset{c}{=}& -\frac{1}{2}\mathrm{tr}\left(\mathbb{E}[Q_k^{-1}](x_k-\mathbb{E}[\theta_k])(x_k-\mathbb{E}[\theta_k])^T\right) \\ &+ \log \mathcal{N}(y_k|H_k x_k,R_k) \\ \overset{c}{=}& \log \mathcal{N}\left(x_k|\mathbb{E}[\theta_k],\{\mathbb{E}[Q_k^{-1}]\}^{-1}\right)\mathcal{N}(y_k|H_k x_k,R_k).\end{aligned} \quad (18)$$

Let $A_k = \{\mathbb{E}[Q_k^{-1}]\}^{-1}$. Exponentiating both sides of Eq. (18) yields 仍是两正态分布的加权融合

$$\begin{aligned}q(x_k) =& \mathcal{N}(x_k|\mathbb{E}[\theta_k],A_k)\mathcal{N}(y_k|H_k x_k,R_k) \\ =& \mathcal{N}(x_k|x_{k|k},P_{k|k}).\end{aligned} \quad (19)$$

It is seen that the posterior PDF of x_k is also Gaussian with mean $x_{k|k}$ and covariance $P_{k|k}$, and this can be calculated by the KF as follows

$$K_{x,k} = A_k H_k^T\left(H_k A_k H_k^T + R_k\right)^{-1}, \quad (20)$$

$$x_{k|k} = \mathbb{E}[\theta_k] + K_{x,k}(y_k - H_k\mathbb{E}[\theta_k]), \quad (21)$$

$$P_{k|k} = A_k - K_{x,k}H_k A_k. \quad (22)$$

B. Derivations of the approximate posterior $q(\theta_k)$

From Eq. (15) and the joint PDF in Eq. (17), we have

$$\begin{aligned}\log q(\theta_k) \overset{c}{=}& \mathbb{E}_{q(x_k,\theta_k)}[\log \mathcal{N}(x_k|\theta_k,Q_k) \\ &+ \log \mathcal{N}(\theta_k|x_{k|k-1},\Sigma_k)] \\ =& -\frac{1}{2}\mathrm{tr}\left(\mathbb{E}[Q_k^{-1}]\mathbb{E}[(x_k-\theta_k)(x_k-\theta_k)^T]\right) \\ &+ \log \mathcal{N}(\theta_k|x_{k|k-1},\Sigma_k) \\ \overset{c}{=}& -\frac{1}{2}\mathrm{tr}\left(\mathbb{E}[Q_k^{-1}](\mathbb{E}[x_k]-\theta_k)(\mathbb{E}[x_k]-\theta_k)^T\right) \\ &+ \log \mathcal{N}(\theta_k|x_{k|k-1},\Sigma_k) \\ =& \log \mathcal{N}(\mathbb{E}[x_k]|\theta_k,A_k)\mathcal{N}(\theta_k|x_{k|k-1},\Sigma_k).\end{aligned} \quad (23)$$

Exponentiating both sides of Eq. (23) yields

$$\begin{aligned}q(\theta_k) =& \mathcal{N}(\mathbb{E}[x_k]|\theta_k,A_k)\mathcal{N}(\theta_k|x_{k|k-1},\Sigma_k) \\ =& \mathcal{N}(\theta_k|\theta_{k|k},P_{\theta,k|k}).\end{aligned} \quad (24)$$

Thus the posterior PDF of θ_k is also Gaussian with mean $\theta_{k|k}$ and covariance $P_{\theta,k|k}$, and we can again invoke the KF to calculate this:

$$K_{\theta,k} = \Sigma_k\left(\Sigma_k + A_k\right)^{-1}, \quad (25)$$

$$\theta_{k|k} = x_{k|k-1} + K_{\theta,k}(\mathbb{E}[x_k] - x_{k|k-1}), \quad (26)$$

$$P_{\theta,k|k} = \Sigma_k - K_{\theta,k}\Sigma_k. \quad (27)$$

C. Derivations of the approximate posterior $q(Q_k)$

In similar vein, from Eq. (16) and the joint PDF of Eq. (17), we have

$$\begin{aligned}\log q(Q_k) \overset{c}{=}& \mathbb{E}_{q(x_k,Q_k)}[\log \mathcal{N}(x_k|\theta_k,Q_k) \\ &+ \log \mathrm{IW}(Q_k|u_{k|k-1},U_{k|k-1})] \\ =& -\frac{1}{2}\log|Q_k| - \frac{1}{2}\mathrm{tr}\left(\mathbb{E}[(x_k-\theta_k)(x_k-\theta_k)^T]Q_k^{-1}\right) \\ &- \frac{1}{2}(u_{k|k-1}+n_x+1)\log|Q_k| - \frac{1}{2}\mathrm{tr}\left(U_{k|k-1}Q_k^{-1}\right) \\ =& -\frac{1}{2}(u_{k|k-1}+n_x+2)\log|Q_k| \\ &- \frac{1}{2}\mathrm{tr}\left((U_{k|k-1}+B_k)Q_k^{-1}\right),\end{aligned} \quad (28)$$

where

$$\begin{aligned}B_k =& \mathbb{E}[(x_k-\theta_k)(x_k-\theta_k)^T] \\ =& (\mathbb{E}[x_k]-\mathbb{E}[\theta_k])(\mathbb{E}[x_k]-\mathbb{E}[\theta_k])^T + \mathrm{cov}(x_k x_k^T) + \mathrm{cov}(\theta_k\theta_k^T) \\ =& (\mathbb{E}[x_k]-\mathbb{E}[\theta_k])(\mathbb{E}[x_k]-\mathbb{E}[\theta_k])^T + P_{k|k} + P_{\theta,k|k}.\end{aligned} \quad (29)$$

Removal of the non-essential normalization term [14], gives the inverse Wishart distribution used in this paper:

$$\mathrm{IW}(Q|u,U) \propto |Q|^{-(u+n+1)/2}\exp\left(-\frac{1}{2}\mathrm{tr}(UQ^{-1})\right), \quad (30)$$

where Q and U are $n \times n$ positive definite matrices.

Exponentiating both sides of Eq. (28) yields

$$q(Q_k) \propto \mathrm{IW}(Q_k|u_{k|k},U_{k|k}), \quad (31)$$

where

$$\begin{aligned}u_{k|k} =& u_{k|k-1} + 1, \\ U_{k|k} =& U_{k|k-1} + B_k\end{aligned} \quad (32)$$

Thus, the inverse Q_k^{-1} has a Wishart distribution $\mathrm{W}(Q_k^{-1}|u_{k|k},U_{k|k}^{-1})$, with mean value given by [13]

$$\mathbb{E}[Q_k^{-1}] = u_{k|k}U_{k|k}^{-1}. \quad (33)$$

D. Summary

The proposed VBAKF-Q algorithm, which performs joint estimation of the system state x_k, the process noise covariance Q_k, and the intermediate latent variable θ_k in a closed-loop iterative manner, is summarized as **Algorithm 1**.

IV. Simulations

We consider a target tracking simulation scenario in 2D cartesian coordinates with unknown and slowly time-varying process noise covariance. The proposed VBAKF-Q algorithm is compared with VBAKF-P of [12], which estimates the predicted error covariance $P_{k|k-1}$ instead of the process noise covariance. Note that although covariance matrices of process noise and measurement noise are both assumed to be unknown in [12], we impose that the latter takes the true value in VBAKF-P for comparison. We also compare VBAKF-Q with the KF with a nominal process noise covariance (KF-NC), and the KF with true process noise covariance (KF-TC). All

Algorithm 1 The variational Bayesian adaptive Kalman filter with unknown Q (VBAKF-Q) algorithm

Input: $x_{k-1|k-1}$, $P_{k-1|k-1}$, $u_{k-1|k-1}$, $U_{k-1|k-1}$, F_k, H_k, R_k, ρ, N;

Output: $x_{k|k}$, $P_{k|k}$, $u_{k|k}$, $U_{k|k}$;

One-step Time Prediction:

$$x_{k|k-1} = F_k x_{k-1|k-1} \tag{34}$$

$$\Sigma_k = F_k P_{k-1|k-1} F_k^T \tag{35}$$

$$u_{k|k-1} = \rho(u_{k-1|k-1} - n_x - 1) + n_x + 1 \tag{36}$$

$$U_{k|k-1} = \rho U_{k-1|k-1} \tag{37}$$

VB Measurement Update:

Iteration Initialization:

$$\theta^0_{k|k} = x_{k|k-1},\ u^0_{k|k} = u_{k|k-1},\ U^0_{k|k} = U_{k|k-1} \tag{38}$$

Iteration Loop:

for each iteration i = 1 : N **do**

$$A_k = \{\mathbb{E}^{i-1}[Q_k^{-1}]\}^{-1} = \frac{U^{i-1}_{k|k}}{u^{i-1}_{k|k}}$$

Update $q(x_k)$

$$K_{x,k} = A_k H_k^T \left(H_k A_k H_k^T + R_k\right)^{-1} \tag{39}$$

$$x^i_{k|k} = \theta^{i-1}_{k|k} + K_{x,k}(y_k - H_k \theta^{i-1}_{k|k}) \tag{40}$$

$$P^i_{k|k} = A_k - K_{x,k} H_k A_k \tag{41}$$

Update $q(\theta_k)$

$$K_{\theta,k} = \Sigma_k \left(\Sigma_k + A_k\right)^{-1} \tag{42}$$

$$\theta^i_{k|k} = x_{k|k-1} + K_{\theta,k}(x^i_{k|k} - x_{k|k-1}) \tag{43}$$

$$P^i_{\theta,k|k} = \Sigma_k - K_{\theta,k}\Sigma_k \tag{44}$$

Update $q(Q_k)$

$$u^i_{k|k} = u_{k|k-1} + 1 \tag{45}$$

$$U^i_{k|k} = U_{k|k-1} + (x^i_{k|k} - \theta^i_{k|k})(x^i_{k|k} - \theta^i_{k|k})^T + P^i_{k|k} + P^i_{\theta,k|k} \tag{46}$$

if $|x^i_{k|k} - x^{i-1}_{k|k}|/|x^i_{k|k}| < \delta_x$ **then**

Exit the iteration; 提前结束

end if

end for

of the algorithms are implemented by MATLAB R2015b on a PC with Intel core i7-3770 CPU.

1) Scenario parameters: The simulation scenario is similar to [12]. The target moves with constant velocity and the corresponding parameters are given by

$$F_k = \mathbf{I}_2 \otimes \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}, \quad H_k = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \tag{47}$$

$$Q_k = (65 + 5\cos(\pi k/T))\, q \mathbf{I}_2 \otimes \begin{bmatrix} \frac{T^3}{3} & \frac{T^2}{2} \\ \frac{T^2}{2} & T \end{bmatrix}, \tag{48}$$

$$R_k = (2 + \cos(\pi k/T))\, r \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}, \tag{49}$$

where $q = 1\ \mathrm{m^2/s^3}$, $r = 100\ \mathrm{m^2}$ and $T = 1$ s.

2) Algorithm parameters: All the algorithm parameters are chosen to be the same as [12]. The tuning parameter $\tau = 3$. The decreasing parameter for Q_k is $\rho = 1 - \exp(-4)$, the number of iterations $N = 10$, and the iterative terminated threshold $\delta_x = 10^{-6}$. The nominal process noise covariance matrices for VBAKF-P and KF-NC are selected to be $\bar{Q}_k = \alpha \mathbf{I}_4$. The initial parameters of the process noise covariance for VBAKF-Q are $u_{0|0} = \tau + n_x + 1$, $U_{0|0} = \tau \bar{Q}_0$.

3) Performance evaluation: The root mean square error (RMSE) of target position and velocity are

$$\mathrm{RMSE_{pos}} \triangleq \sqrt{\frac{1}{N_s}\sum_{N_s}[(p^x_k - p^x_{k,t})^2 + (p^y_k - p^y_{k,t})^2]}, \tag{50}$$

$$\mathrm{RMSE_{vel}} \triangleq \sqrt{\frac{1}{N_s}\sum_{N_s}[(v^x_k - v^x_{k,t})^2 + (v^y_k - v^y_{k,t})^2]}, \tag{51}$$

where N_s is the number of Monte Carlo runs, p^x_k, $p^x_{k,t}$, v^x_k and $v^x_{k,t}$ are the estimated position, true position, estimated velocity, and true velocity in the x direction, and p^y_k, $p^y_{k,t}$, v^y_k and $v^y_{k,t}$ are the estimated position, true position, estimated velocity, and true velocity in the y direction, respectively.

4) Simulation results: The RMSEs obtained by 1000 Monte carlo runs are shown in Figs. 1∼4 with different nominal process noise covariances. The simulations are implemented with $\alpha = 0.1, 10$. Note that the error between nominal process noise covariance and true process noise covariance is larger when $\alpha = 0.1$. When $\alpha = 10$, VBAKF-Q almost has the same estimation accuracy as VBAKF-P in position, while the former performs better than the latter in velocity. When $\alpha = 0.1$, VBAKF-Q is able to achieve nearly the same estimation accuracy as before while both VBAKF-P and KF-NC have large errors. This is because, with this value of α, the nominal process noise covariance affects the performance of VBAKF-P and KF-NC more strongly.

The initial process noise covariance matrices for VBAKF-Q are different with different α, and the results show that it has little effect on the performance. The results also indicate that VBAKF-Q has a better performance in velocity, and can nearly obtain the same estimation accuracy as KF-TC. This result is largely due to the way that the underlying system uncertainty is modeled.

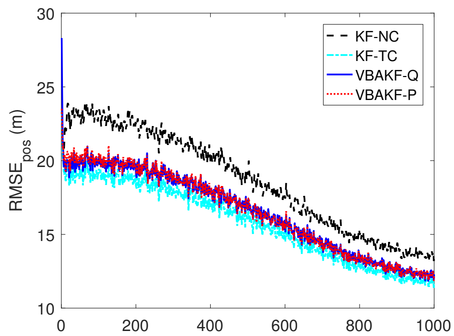

Fig. 1. $\mathrm{RMSE_{pos}}$ with $\alpha = 10$

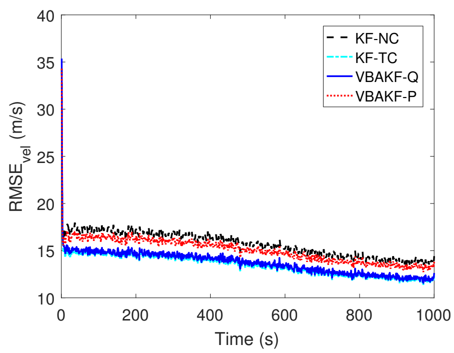

Fig. 2. $\mathrm{RMSE_{vel}}$ with $\alpha = 10$

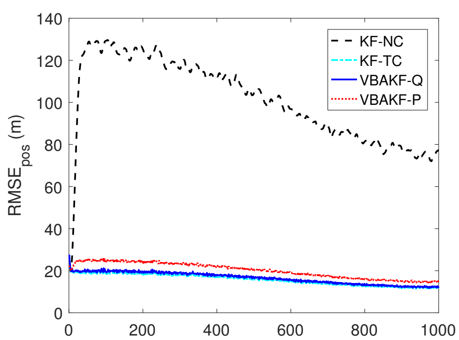

Fig. 3. $\mathrm{RMSE_{pos}}$ with $\alpha = 0.1$

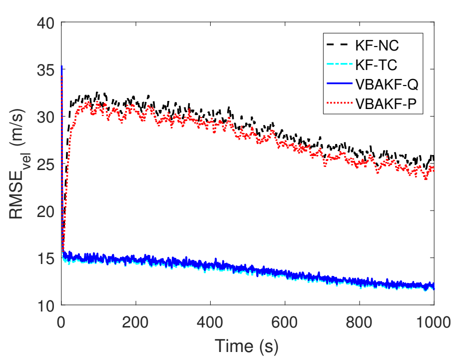

Fig. 4. $\mathrm{RMSE_{vel}}$ with $\alpha = 0.1$

V. Conclusion

In this work, the problem of state estimation with unknown process noise covariance is considered and a novel adaptive KF is proposed. The underlying conjugate prior distribution of the process noise covariance is the inverse Wishart distribution with a latent variable and we estimated the system state together with the unknown process noise covariance and the latent variable in a VB framework. For each measurement, the joint posterior PDF is updated iteratively by maximizing the ELBO. An improved estimation performance is confirmed by the simulation results presented.

Acknowledgment

This work was supported by the National Natural Science Foundation of China (Project No. 61501378, 61503305 and 61790552).

References

[1] R. Mehra, "Approaches to adaptive filtering," *IEEE Transactions on Automatic Control*, vol. 17, no. 5, pp. 693–698, 1972.

[2] 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.

[3] P. S. Maybeck, *Stochastic models, estimation, and control*, vol. 3. Academic press, 1982.

[4] X. Li and Y. Bar-Shalom, "A recursive multiple model approach to noise identification," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 30, no. 3, pp. 671–684, 1994.

[5] G. Storvik, "Particle filters for state-space models with the presence of unknown static parameters," *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 281–289, 2002.

[6] A. P. Sage and G. W. Husa, "Adaptive filtering with unknown prior statistics," in *Proc. Joint Automatic Control Conference*, no. 7, pp. 760–769, 1969.

[7] M. J. Beal, *Variational algorithms for approximate Bayesian inference*. London: University of London, 2003.

[8] D. M. Blei, A. Kucukelbir, and J. D. McAuliffe, "Variational inference: A review for statisticians," *Journal of the American Statistical Association*, vol. 112, no. 518, pp. 859–877, 2017.

[9] S. Sarkka and J. Hartikainen, "Non-linear noise adaptive Kalman filtering via variational Bayes," in *Proc. 2013 IEEE International Workshop on Machine Learning for Signal Processing*, pp. 1–6, 2013.

[10] R. Piché, S. Särkkä, and J. Hartikainen, "Recursive outlier-robust filtering and smoothing for nonlinear systems using the multivariate Student-t distribution," in *Proc. 2012 IEEE International Workshop on Machine Learning for Signal Processing*, pp. 1–6, 2012.

[11] T. Ardeshiri, E. Özkan, U. Orguner, and F. Gustafsson, "Approximate Bayesian smoothing with unknown process and measurement noise covariances," *IEEE Signal Processing Letters*, vol. 22, no. 12, pp. 2450–2454, 2015.

[12] Y. Huang, Y. Zhang, Z. Wu, N. Li, and J. Chambers, "A novel adaptive Kalman filter with inaccurate process and measurement noise covariance matrices," *IEEE Transactions on Automatic Control*, vol. 63, no. 2, pp. 594–601, 2018.

[13] K. V. Mardia, J. T. Kent, and J. M. Bibby, *Bayesian data analysis*. Academic press, 1979.

[14] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin, *Bayesian data analysis*, vol. 2. Boca Raton, FL: CRC press, 2014.