Article

Multi-Fading Factor and Updated Monitoring Strategy Adaptive Kalman Filter-Based Variational Bayesian

Chenghao Shan [1], **Weidong Zhou** [1,*], **Yefeng Yang** [2] **and Zihao Jiang** [1]

1   College of Intelligent Systems Science and Engineering, Harbin Engineering University,
    Harbin 150001, China; shanchenghao123@hrbeu.edu.cn (C.S.); jiangzihao@hrbeu.edu.cn (Z.J.)
2   Center for Control Theory and Guidance Technology, Harbin Institute of Technology, Harbin 150001, China;
    18B904013@stu.hit.edu.cn
*   Correspondence: zhouweidong@hrbeu.edu.cn; Tel.: +86-0451-82568107

Abstract: Aiming at the problem that the performance of adaptive Kalman filter estimation will be affected when the statistical characteristics of the process and measurement of the noise matrices are inaccurate and time-varying in the linear Gaussian state-space model, an algorithm of multi-fading factor and an updated monitoring strategy adaptive Kalman filter-based variational Bayesian is proposed. Inverse Wishart distribution is selected as the measurement noise model and the system state vector and measurement noise covariance matrix are estimated with the variational Bayesian method. The process noise covariance matrix is estimated by the maximum a posteriori principle, and the updated monitoring strategy with adjustment factors is used to maintain the positive semi-definite of the updated matrix. The above optimal estimation results are introduced as time-varying parameters into the multiple fading factors to improve the estimation accuracy of the one-step state predicted covariance matrix. The application of the proposed algorithm in target tracking is simulated. The results show that compared with the current filters, the proposed filtering algorithm has better accuracy and convergence performance, and realizes the simultaneous estimation of inaccurate time-varying process and measurement noise covariance matrices.

Keywords: variational Bayesian; multiple-fading factors; time-varying noise covariance matrices; inaccurate noise; target tracking; updated monitoring strategy

1. Introduction

In many practical engineering applications, the actual values of the required state variables are often not directly available. For example, when radar detects an air target, it can calculate the target distance based on information such as reflected waves. Still, there is random interference in the radar detection process, resulting in random noise in the observation signal. In this case, it is impossible to obtain the required state variables accurately, and these state variables can only be estimated or predicted based on the observed signal. In linear systems, the Kalman filter is the optimal filter [1]. With the development of computer technology, the calculation requirements and complexity of Kalman filtering no longer become obstacles to its application [2]. At present, the Kalman filtering theory has been widely used in tracking, navigation, guidance, and other areas [3–9].

The application of the Kalman filter requires prior knowledge of the mathematical model of the system and the statistical characteristics of noise. Still, in many practical application problems, these are unknown or only partially known [10–12]. If inaccurate mathematical models or statistical noise characteristics are used to design the Kalman filter, the performance of the filter will be degraded, resulting in larger estimation errors, and even filter divergence. To solve this problem, various adaptive Kalman filters (AKF) have been produced [13,14].

The Sage–Husa filter (SH-KF) is widely used because of its simple algorithm, which can estimate the first and second moments of noise online [15]. However, the Sage–

Husa adaptive noise estimator has problems such as a large amount of calculation and easy divergence of state estimation [16]. In addition, some literature has pointed out that the covariance matrices of process noise and observation noise cannot be estimated dynamically in real-time by the Sage–Husa adaptive estimator at the same time, which can only estimate the other noise covariance matrix when one noise covariance matrix is known [17,18]. The maximum-likelihood-based adaptive filtering method (ML-KF) can evaluate and correct the second-order moments of the noise statistics online, but it needs to rely on accurate innovation covariance estimation, and ML-KF requires a large sliding window to obtain a precise estimate of the noise covariance matrix, which theoretically makes it only for time-varying noise covariance matrix estimation [19,20]. Strong tracking Kalman filter (ST-KF) is an adaptive filtering algorithm based on the principle of residual orthogonality. It adjusts the weight of new measurement data by adding an estimate of the one-step predictive covariance matrix. It has strong robustness regarding model parameter mismatch, lower sensitivity to the statistical characteristics of noise and initial values, and a strong ability to track sudden changes. But its adjustment ability for each filtering channel is the same, and state predicted error covariance matrix (PECM) and measurement noise covariance matrix (MNCM) are not estimated [21,22]

In recent years, many scholars have introduced the variational Bayesian machine-learning method into the KF algorithm and proposed AKF algorithm based on the variational Bayesian approach (VB-KF), which is an approximation of the Bayesian method. By choosing a suitable conjugate prior distribution, the slow time-varying measurement noise covariance can be estimated [23–25]. Literature [26] proposed a variational adaptive Kalman filter (R-VBKF) but only estimated the system state vector and the measurement noise covariance matrix (MNCM); the accuracy is not satisfactory enough. In the algorithm presented in the literature [27], the state predicted error covariance matrix (PECM) and the measurement noise covariance matrix (MNCM) are estimated, but the process noise covariance matrix (PNCM) is not directly assessed. The specific estimated value of the PNCM cannot be obtained, and the estimation accuracy needs to be further improved.

Aiming at the linear Gaussian state–space model with slow time-varying covariance of process and measurement noise, taking into account the estimation accuracy, convergence performance, robustness, and the realization of simultaneous estimation of noise covariance matrices, the multi-fading factor and updated monitoring strategy, AKF-based variational Bayesian (MFMS-VBAKF) is proposed. Its feasibility is proved by simulation experiments.

The main contributions of the algorithm proposed in this paper are as follows: compared with the inverse-Gamma distribution applied by the R-VBKF algorithm in [25], the algorithm proposed in this paper selects the distribution of the measurement noise model as a more reasonable inverse-Wishart distribution in variational Bayesian approach, which improves the estimation accuracy of system state vector and MNCM; compared with SH-KF in [15], the algorithm proposed in this paper guarantees the positive semi-definiteness of the PNCM under the maximum posterior estimation by designing an updated monitoring strategy, and achieves the simultaneous estimation of the PNCM and MNCM.; and compared with ST-KF [21], the proposed algorithm has different adjustment abilities for each channel by introducing multi-fading factors and the PECM estimation accuracy is also improved.

The main structure is as follows: Section 2 illustrates the mathematical modeling of the problem. In Section 3, the multi-fading factor and updated monitoring strategy AKF-based variational Bayesian is derived. In Section 4, the proposed algorithm is compared with the existing algorithm through the simulation of the target tracking application, and the excellent performance of the proposed algorithm is proved. Section 5 summarizes the conclusions. Finally, Section 6 plans the future work.

2. Problem Modeling

Consider the following discrete linear stochastic system of the state–space model

$$X_k = \boldsymbol{\Phi}_{k-1} X_{k-1} + \boldsymbol{\omega}_{k-1}, \tag{1}$$

$$Z_k = H_{k-1}X_k + v_k, \tag{2}$$

where (1) and (2) are process and measurement equations, respectively. k is discrete-time, $X_k \in R^{n \times n}$ is the state vector of the system at time k, $Z_k \in R^{m \times m}$ is the measurement signal vector of the corresponding state. $\Phi_k \in R^{n \times n}$ is the state-transition matrix, $H_k \in R^{m \times n}$ is the measurement matrix. $\omega_k \in R^n$ and $v_k \in R^m$ are uncorrelated white Gaussian noise with zero mean vectors and covariance matrices Q_k and R_k, respectively. The initial state X_0 is assumed to be a Gaussian distribution with mean vector $\hat{X}_0$ and the covariance matrix P_0. X_0 is uncorrelated to ω_k and v_k at any time [1].

For linear Gaussian state–space models, the Kalman filter (KF) algorithm is an optimal estimation filter algorithm. If the noise covariance matrices Q_k and R_k are fully known, KF estimates the state vector X_k through the measurement information of $Z_{1:k}$, and the estimation accuracy is satisfactory. However, the performance of the KF algorithm overly depends on the prior knowledge of the noise statistics. If the time-varying noise covariance matrices Q_k and R_k are unknown or inaccurate, the accuracy of the KF algorithm will decrease, and even cause the estimation to diverge. Besides, when most existing AKF algorithms estimate the PNCM Q_k and MNCM R_k at the same time, the filtering will diverge. Therefore, a multi-fading factor and updated monitoring strategy AKF-based variational Bayesian with inaccurate time-varying PNCM and MNCM is proposed.

3. The Proposed Multi-Fading Factor and Updates Monitoring Strategy AKF-Based Variational Bayesian

In the VBAKF algorithm, the independent state vector X_k and the measurement noise covariance matrix R_k are regarded as the parameters to be estimated.

3.1. AKF-Based Variational Bayesian (VBAKF)

3.1.1. Prediction Process and Distribution Selection

In the traditional Kalman filter framework, the Gaussian distributions are selected as the distributions of one-step predicted probability density function (PDF) $P(X_k|Z_{1:k-1})$ and likelihood PDF $p(Z_k|X_k)$:

$$p(X_k|Z_{1:k-1}) = N\big(X_{k-1}; \hat{X}_{k:k-1}, P_{k:k-1}\big), \tag{3}$$

$$p(Z_k|X_k) = N\big(Z_k; H\hat{X}_{k:k-1}, R_k\big), \tag{4}$$

where $N(G; \mu, \Sigma)$ is the Gaussian distribution and μ and Σ represent the mean and variance of the distribution, respectively. The PDF of the Gaussian distribution is:

$$N(G; \mu, \Sigma) = \frac{1}{\sqrt{|2\pi\Sigma|}} exp^{-\frac{1}{2}(G-\mu)^T \Sigma^{-1}(G-\mu)}, \tag{5}$$

According to Equation (1), the predicted state vector $\hat{X}_{k:k-1}$ and the corresponding one-step predicted error covariance matrix (PECM) $P_{k:k-1}$ can be described as:

$$\hat{X}_{k:k-1} = \Phi_{k-1}\hat{X}_{k-1:k-1}, \tag{6}$$

$$P_{k:k-1} = \Phi_{k-1}P_{k-1:k-1}\Phi_{k-1}^T + Q_{k-1}, \tag{7}$$

where $\hat{X}_{k:k-1}$ and $P_{k:k-1}$ represent the state estimation at time $k-1$ and the corresponding estimation error covariance matrix, respectively. $(.)^T$ represents the transpose of the matrix. Among them, it is assumed that the real PNCM Q_k is unknown due to the complex environmental factors in real applications, which will lead to an inaccurate $P_{k:k-1}$ in Equation (7). The estimation methods for Q_k and $P_{k:k-1}$ will be given in the next two sections.

In this section, the aim is to infer X_k with R_k. For the purpose, the conjugate prior distributions need to be first selected for inaccurate MNCM R_k since the conjugacy can ensure that the posterior distribution and the prior distribution maintain the same functional form.

According to Bayesian statistical theory, if the Gaussian distribution has a known mean, the conjugate prior distribution of the covariance matrix can be regarded as the inverse Wishart (IW) distribution [28]. C^{-1} is the inverse matrix of a positive definite matrix C. If C^{-1} follows the Wishart-distribution $W\left(C^{-1}; \lambda, \Psi^{-1}\right)$, the matrix C follows the IW distribution:

$$IW(C; \lambda, \Psi) = \frac{|\Psi|^{\lambda/2}}{2^{\lambda d/2}\Gamma_d\left(\frac{\lambda}{2}\right)}|C|^{-(\lambda+d+1)/2}exp^{-\frac{1}{2}tr[\Psi C^{-1}]}. \tag{8}$$

In Equation (8), C is a symmetric positive definite random matrix, distribution parameter λ is a dof parameter, Ψ is a symmetric positive definite matrix, d is the dimension of C, $\Gamma_n(.)$ represents a multivariate gamma function, and $tr[.]$ is the matrix trace calculation. Additionally, if $\lambda > d + 1$ and $E\left[C^{-1}\right] \sim IW(C; \lambda, \Psi)$, then $E\left[C^{-1}\right] = (\lambda - d - 1)\Psi^{-1}$. $E(.)$ stands for mathematical expectation calculation [29]. Since R_k is the covariance matrix of the Gaussian PDF, the prior distribution of $P(R_k|Z_{1:k-1})$ can be written as IW distribution:

$$p(R_k|Z_{1:k-1}) = IW\left(R_{k:k-1}; \hat{t}_{k:k-1}, \hat{T}_{k:k-1}\right), \tag{9}$$

where $\hat{t}_{k:k-1}$ and $\hat{T}_{k:k-1}$ denote the degrees of freedom (dof) parameter and scale matrix of $p(R_k|Z_{k-1})$, respectively. Next, the values of $\hat{t}_{k:k-1}$ and $\hat{T}_{k:k-1}$ need to be assigned.

Owing to the Bayesian theorem, the prior distribution $p(R_k|Z_{k-1})$ can be written as:

$$p(R_k|Z_{1:k-1}) = \int p(R_k|R_{k-1})p(R_{k-1}|Z_{1:k-1})dR_{k-1}, \tag{10}$$

where $p(R_{k-1}|Z_{1:k-1}))$ is the posterior probability density function (PDF) of the MNCM R_{k-1}.

Utilizing (9), the distribution of posterior PDF $p(R_{k-1}|Z_{1:k-2})$ can be replaced as inverse Wishart distribution, due to the prior distribution $p(R_{k-1}|Z_{1:k-1})$ of MNCM R_{k-1} is selected as inverse Wishart distribution, and $p(R_{k-1}|Z_{1:k-1})$ can be written as:

$$p(R_{k-1}|Z_{1:k-1}) = IW\left(R_{k-1}; \hat{t}_{k-1:k-1}, \hat{T}_{k-1:k-1}\right). \tag{11}$$

To guarantee $p(R_k|Z_{1:k-1})$ also obeys the inverse Wishart distribution, a changing factor ρ is introduced to modify the one-step predicted values of the distribution parameters $\hat{t}_{k:k-1}$ and $\hat{T}_{k:k-1}$. The formulas are as follows:

$$\hat{t}_{k:k-1} = \rho(\hat{t}_{k-1} - m - 1) + m + 1, \tag{12}$$

$$\hat{T}_{k:k-1} = \rho\hat{T}_{k-1}, \tag{13}$$

among them, m is the dimension of Z_k, $\rho \in (0,1]$, the time-varying measurement noise covariance matrix can be changed with a certain probability distribution, and control the posterior and prior probability density functions to have the same distribution.

In addition, the initial PDF distribution of MNCM R_0 is also assumed as inverse Wishart distribution. $p(R_0) = IW\left(R_0; \hat{t}_{0:0}, \hat{T}_{0:0}\right)$. At the initial moment, to formulate the prior information of the measurement noise covariance matrix, the mean value of R_0 is considered as the initial fixed measurement noise covariance matrix $\widetilde{R}_0$, i.e.,

$$\widetilde{R}_0 = \frac{\hat{T}_{k:k-1}}{\hat{t}_{k:k-1}-m-1} \tag{14}$$

Assuming that the prior distribution of the joint probability density function of the state variable and the MNCM is the product of the Gaussian distribution and the inverse Wishart distribution, the prediction process can be defined as:

$$p(X, R_k|Z_{1:k-1}) = p(X_k|Z_{1:k-1})p(R_k|Z_{k-1}) = N\left(X_{k-1}; \hat{X}_{k:k-1}, P_{k:k-1}\right)IW\left(R_{k-1}; \hat{t}_{k-1:k-1}, \hat{T}_{k-1:k-1}\right) \tag{15}$$

3.1.2. Variational Update Process

Aiming at estimating the state X_k and the MNCM R_k, their joint posterior PDF $p(X_k, R_k|Z_{1:k})$ needs to be calculated. However, the analytical solution of this joint posterior PDF cannot be obtained directly. The variational Bayesian method is utilized to find an approximate PDF of a free form as follows [30]:

$$p(X_k, R_k|Z_{1:k}) \approx q(X_k)q(R_k), \tag{16}$$

where $q(.)$ means the approximate posterior PDF of $p(.)$.

In the standard VB method, Kullback–Leibler divergence (KLD) is used to measure the degree of approximation between the approximation posterior PDF and the true posterior PDF, and the optimal solution is obtained by minimizing KLD. The VB method can provide a closed form solution for the approximate posterior PDF. Minimizing the KLD between the approximation posterior PDF $q(X_k)$, $q(R_k)$ and the true joint posterior $p(X_k, R_k|Z_{1:k-1})$ is used to form the VB-approximation [30]:

$$\{q(X_k)q(R_k)\} = \arg min \, KLD(q(X_k), q(R_k) \, || \, p(X_k, R_k|Z_{1:k})), \tag{17}$$

The divergence function KLD(.) is defined as:

$$KLD(q(A) \, || \, p(A)) = \int q(A) \log \frac{q(A)}{p(A)} dA. \tag{18}$$

Combined with Equations (17) and (18), the optimal solution of Equation (16) is derived as:

$$\log q(X_k) = E_{R_k}[\log p(X_k, R_k, Z_{1:k})] + c_{X_k}, \tag{19}$$

$$\log q(R_k) = E_{X_k}[\log p(X_k, R_k, Z_{1:k})] + c_{R_k}, \tag{20}$$

where $\log(.)$ stands for natural logarithm calculation, $E_\varphi[.]$ denotes the expectation calculation of the approximate posterior PDF of the variable φ, c_{X_k} and c_{R_k} represent the constants of variable X_k and MNCM R_k, respectively. The solutions of Equations (19) and (20) cannot be solved directly since $q(X_k)$ and $q(R_k)$ are coupled. Therefore, the fixed-point iteration method is introduced to calculate the solution to these parameters.

The further form of Equation (20) can be derived as (See Appendix A for details):

$$\log q^{(i+1)}(R_k) = c_{R_k} - \frac{1}{2}(m + \hat{t}_{k:k-1} + 2)\log|R_k| - $$
$$\frac{1}{2}trace\left(\left(E^{(i)}\left[(Z_k - H_kX_k)(Z_k - H_kX_k)^T\right] + \hat{T}_{k:k-1}\right)R_k^{-1}\right), \tag{21}$$

where: $q^{(i)}(.)$ represents the approximate probability distribution of $q(.)$ at the i-th iteration, $trace(.)$ is the calculation of the matrix trace, C_{R_k} is a constant related to R_k which independent of the distribution form, and m is the dimension of the real observation matrix. The expectation part of Equation (21) is defined as $V_k^{(i)}$, and expanded as:

$$V_k^{(i)} = E^{(i)}\left[(Z_k - H_kX_k)(Z_k - H_kX_k)^T\right]$$
$$= \int (Z_k - H_kX_k)(Z_k - H_kX_k)^T + N\left(X_k; \hat{X}_{k:k}^i, P_{k:k}\right)dX_k \tag{22}$$
$$= \left(Z_k - H_k\hat{X}_k^i\right)\left(Z_k - H_k\hat{X}_k^i\right)^T + H_kP_kH_k^T,$$

it can be seen that $q^{(i+1)}(R_k)$ obeys a new inverse Wishart distribution form as follows,

$$q^{(i+1)}(R_k) = IW\left(R_k; \hat{t}_k^{(i+1)}, \hat{T}_k^{(i+1)}\right). \tag{23}$$

and the distribution parameters $\hat{t}_k^{(i+1)}$ and $\hat{T}_k^{(i+1)}$ are, respectively, as follows:

$$\hat{t}_k^{(i+1)} = \hat{t}_{k:k-1} + 1, \tag{24}$$

$$\hat{\boldsymbol{T}}_k^{(i+1)} = \boldsymbol{V}_k^{(i)} + \hat{\boldsymbol{T}}_k^{(i)}, \tag{25}$$

Similarly, the logarithmic expression of the approximate distribution of the system state $\boldsymbol{X}_k$ is as follows:

$$\log q^{(i+1)}(\boldsymbol{X}_k) = -\frac{1}{2}(\boldsymbol{Z}_k - \boldsymbol{H}_k \boldsymbol{X}_k)^T E^{(i+1)}\left[\boldsymbol{R}_k^{-1}\right](\boldsymbol{Z}_k - \boldsymbol{H}_k \boldsymbol{X}_k) - \\ \frac{1}{2}\left(\boldsymbol{X}_k - \hat{\boldsymbol{X}}_{k:k-1}\right)^T \boldsymbol{P}_{k:k-1}^{-1}\left(\boldsymbol{X}_k - \hat{\boldsymbol{X}}_{k:k-1}\right) + c_{\boldsymbol{X}_k}, \tag{26}$$

where $E^{(i+1)}\left[\boldsymbol{R}_k^{-1}\right]^{-1}$ is given by:

$$E^{(i+1)}\left[\boldsymbol{R}_k^{-1}\right]^{-1} = \left(\hat{t}_k^{(i+1)} - m - 1\right)^{-1}\hat{\boldsymbol{T}}_k^{(i+1)}. \tag{27}$$

The likelihood PDF $p(\boldsymbol{Z}_k|\boldsymbol{X}_k)$ in Equation (4) after updating the $(i+1)$-th iteration can be derived as follows:

$$p^{(i+1)}(\boldsymbol{Z}_k|\boldsymbol{X}_k) = N\left(\boldsymbol{Z}_k; \boldsymbol{H}_k \boldsymbol{X}_k, \hat{\boldsymbol{R}}_k^{(i+1)}\right), \tag{28}$$

The corrected measurement noise covariance matrix (MNCM) $\hat{\boldsymbol{R}}_k^{(i+1)}$ can be written as:

$$\hat{\boldsymbol{R}}_k^{(i+1)} = E^{(i+1)}\left[\boldsymbol{R}_k^{-1}\right]^{-1} \tag{29}$$

Since $q(\boldsymbol{X}_k)$ obeys the Gaussian distribution as $q(\boldsymbol{X}_k) = N\left(\boldsymbol{X}_k; \hat{\boldsymbol{X}}_k, \boldsymbol{P}_k\right)$. Combining with the standard Kalman filter framework, the gain matrix $\boldsymbol{K}_k^{(i+1)}$, system state $\hat{\boldsymbol{X}}_k^{(i+1)}$, and state covariance $\boldsymbol{P}_k^{(i+1)}$ in the variational measurement update are corrected as follows, respectively:

$$\boldsymbol{K}_k^{(i+1)} = \boldsymbol{P}_{k:k-1}\boldsymbol{H}_k^T\left(\boldsymbol{H}_k\boldsymbol{P}_{k:k-1}\boldsymbol{H}_k^T + \hat{\boldsymbol{R}}_k^{(i+1)}\right)^{-1}, \tag{30}$$

$$\hat{\boldsymbol{X}}_k^{(i+1)} = \hat{\boldsymbol{X}}_{k:k-1} + \boldsymbol{K}_k^{(i+1)}\left(\boldsymbol{Z}_k - \boldsymbol{H}_k\hat{\boldsymbol{X}}_{k:k-1}\right), \tag{31}$$

$$\boldsymbol{P}_k^{(i+1)} = \left(\boldsymbol{I} - \boldsymbol{K}_k^{(i+1)}\boldsymbol{H}_k\right)\boldsymbol{P}_{k:k-1}\left(\boldsymbol{I} - \boldsymbol{K}_k^{(i+1)}\boldsymbol{H}_k\right)^T + \boldsymbol{K}_k^{(i+1)}\hat{\boldsymbol{R}}_k^{(i+1)}\left(\boldsymbol{K}_k^{(i+1)}\right)^T \tag{32}$$

Analyzing the above derivation, it can be seen that the implicit solution of the variational update formula is constituted by the Equations (22), (24), (25), and (29)–(32). The expected maximum approach is used to iteratively calculate $q(\boldsymbol{X}_k)$ and $q(\boldsymbol{R}_k)$ to update the parameters $\boldsymbol{X}_k$ and $\boldsymbol{R}_k$ to be estimated continuously. When $q(\boldsymbol{X}_k)$ and $q(\boldsymbol{R}_k)$ are closer to $p(\boldsymbol{X}, \boldsymbol{R}_{k-1}|\boldsymbol{Z}_{1:k})$, the KLD value of Equation (17) is smaller, and the estimation results of the parameter to be estimated adaptively approach to the true value until the iteration of the variational update process is finished. At this time, the optimal estimation results of parameters $\hat{\boldsymbol{X}}_k$ and $\hat{\boldsymbol{R}}_k$ to be estimated at time k can be calculated as follows (N is the number of fixed-point iterations):

$$q(\boldsymbol{X}_k) \approx q^{(N)}(\boldsymbol{X}_k) = \left(\boldsymbol{X}_k; \hat{\boldsymbol{X}}_{k;k}^{(N)}, \boldsymbol{P}_{k;k}^{(N)}\right) = \left(\boldsymbol{X}_k; \hat{\boldsymbol{X}}_{k:k}, \boldsymbol{P}_{k:k}\right), \tag{33}$$

$$q(\boldsymbol{R}_k) \approx q^{(N)}(\boldsymbol{R}_k) = IW\left(\boldsymbol{R}_k; \hat{t}_k^{(N)}, \hat{\boldsymbol{T}}_k^{(N)}\right) = IW\left(\boldsymbol{R}_k; \hat{t}_{k|k}, \hat{\boldsymbol{T}}_{k|k}\right). \tag{34}$$

3.2. Updated Monitoring Strategy Based on Maximum a Posterior (MAP) for Estimating the PNCM Q_k

Some existing adaptive filtering methods estimate the process noise covariance matrix (PNCM) $\boldsymbol{Q}_k$ and measurement noise covariance matrix (MNCM) $\boldsymbol{R}_k$ at the same time, it is easy to cause the accuracy of the estimated value of the state $\boldsymbol{X}_k$ to decrease or even diverge. This is caused by the value of $\boldsymbol{Q}_k$ becoming negative definite matrix during the estimation process [17].

Aiming at realizing the simultaneous estimation of the PNCM Q_k, MNCM R_k and improving the estimation accuracy of the state vector X_k. An updated monitoring strategy based on maximum a posterior (MAP) for estimating the PNCM Q_k is proposed.

According to the state–space model as Equations (1) and (2), in paper [15], the maximum a posteriori suboptimal unbiased estimation method based on the noise statistics of measurement $\{Z_1, Z_2, Z_3, \cdots Z_k\}$ for estimating the PNCM Q_k is given as:

$$\hat{Q}_k = \frac{1}{k} * \sum_{s=0}^{k-1} [K_s \gamma_s \gamma_s^T K_s^T + P_{s:s} - \Phi_{s-1} P_{s-1:s-1} \Phi_{s-1}^T], \tag{35}$$

combining Equation (35) with the conclusion of Section 3.1, where $K_s = K_s^{(N)}$ is the optimal gain calculated by VBAKF through Nth variational iterations at time s, $P_{s:s} = P_s^{(N)}$ is the state covariance calculated after Nth variational iterations at time s and $\gamma_s = Z_k - \hat{X}_{k:k-1}$ is the residual.

From a statistical point of view, Equation (35) is an arithmetic average, and the weight coefficient in the formula is $1/(k+1)$. However, when estimating the time-varying process noise covariance matrix, the role of the latest information should be highlighted, which can be achieved by multiplying each item in $\sum_{s=0}^{k} [.]$ by a different weighting coefficient. Owing to the exponential weighting method, we can assign different weights to each item in $\sum_{s=0}^{k} [.]$, and increase the weight of the latest information, thereby improving the accuracy of $\hat{Q}_k$ estimation. The exponential weighting method is introduced, and the weighting coefficient $\{\vartheta_s\}$ is selected to satisfy:

$$\vartheta_s = \vartheta_{s-1} b; \ b \in (0,1); \ \sum_{s=0}^{k} \vartheta_s = 1. \tag{36}$$

further deduced as follows:

$$\vartheta_s = d_k b^s, \ d_k = \frac{(1-b)}{(1-b^{k+1})}; \ s \in [0,k], \tag{37}$$

where b is the attenuation factor. In $\sum_{s=0}^{k} [.]$ of Equation (35), each item is multiplied by the weight coefficient d_k, instead of the original weight coefficient. The time-varying process noise covariance matrix (PNCM) Q_k estimation method is obtained, and the recursive algorithm is derived as:

$$\hat{Q}_k = (1 - d_k) \hat{Q}_{k-1} + d_k \left(K_k^{(N)} \gamma_k \gamma_k^T K_k^{(N)^T} + P_k^{(N)} - \Phi_{k-1} P_{k-1}^{(N)} \Phi_{k-1}^T \right). \tag{38}$$

Equations (7), (30)–(32), and (38) constitutes the VBAKF algorithm that simultaneously estimates the PNCM Q_k, MNCM R_k, state vector X_k, and one-step predicted state error covariance matrix (PECM) $P_{k:k-1}$.

However, through simulation experiments, the estimation of the PNCM Q_k by the above algorithm is prone to abnormality; that is, it loses positive semi-definiteness, which leads to filtering divergence.

To solve this problem, based on not losing the information of the original process noise estimation algorithm (38), an updated monitoring strategy of process noise parameters is designed. Firstly, it is judged whether Q_k calculated by Equation (38) is a positive semi-definite matrix. Then the adjustment factor β is introduced to update the process noise estimation parameters to ensure that the corrected PNCM meets the requirements.

The right side of the Equation (38) is shifted as follows:

$$\hat{Q}_k = \hat{Q}_{k-1} + d_k \left(K_k^{(N)} \gamma_k \gamma_k^T K_k^{(N)^T} + D_k \right), \tag{39}$$

$$D_k = P_k^{(N)} - \Phi_k P_{k-1}^{(N)} \Phi_k^T - \hat{Q}_{k-1}. \tag{40}$$

Generally speaking, the selection of the initial value of the state error covariance matrix P_0 is imprecise, and the deviation from the ideal value is large in the initial stage of filtering, resulting in the absolute value of the theoretical process noise covariance matrix D_k determined by Equation (39) being much larger than $\gamma_k \gamma_k^T$. It is easy to cause $\hat{Q}_k$ to lose the positive semi-definiteness, and then the filtering will diverge. Therefore, it is necessary to introduce an adjustment factor to attenuate the effect of the state error covariance matrix at the initial moment, to avoid the indefiniteness of the estimated value of the PNCM, to prevent the filter from diverging. The updated monitoring strategy of the PNCM is as follows:

$$\hat{Q}_k = \hat{Q}_{k-1} + d_k \left(K_k^{(N)} \gamma_k \gamma_k^T K_k^{(N)^T} + (\beta_k)^p \left(P_k^{(N)} - \Phi_{k-1} P_{k-1}^{(N)} \Phi_{k-1}^T - \hat{Q}_{k-1} \right) \right), \tag{41}$$

where $p \geq 1$ is a positive integer (the initial value is 1), β_k is the adjustment factor, and the value of β_k is related to the state error variance matrix as follows:

$$\beta_k = exp^{-a_k}, \tag{42}$$

$$a_k = \frac{trace\left(P_k^{(N)} - \Phi_{k-1} P_{k-1}^{(N)} \Phi_{k-1}^T - \hat{Q}_{k-1} \right)}{trace\left(K_k^{(N)} \gamma_k \gamma_k^T K_k^{(N)^T} \right)}. \tag{43}$$

The specific process for the updated monitoring strategy is as follows: monitor the process noise covariance matrix calculated by Equation (38), and judge whether $\hat{Q}_k$ is a positive semi-definite matrix to determine whether $\hat{Q}_k$ needs to be updated. If not, output $\hat{Q}_k$. Otherwise, turn to Equation (41) and set $p = 1$; Equation (41) is used to update the process noise parameters. Continue the monitoring of the updated estimated process noise covariance matrix to determine whether it is necessary to continue to update. If it is necessary, take $p = p + 1$; Equation (41) is used to recalculate $\hat{Q}_k$. The loop is executed until $\hat{Q}_k$ is a positive semi-definite matrix. End the update of the process noise covariance matrix at the current moment. The flowchart of one time-step of the updated monitoring strategy is shown in Figure 1.

So far, combined with the traditional Kalman filter framework, the VBAKF algorithm with updated monitoring strategy is derived to estimate system state vector X_k, the process noise covariance matrix (PNCM) Q_k, measurement noise covariance matrix (MNCM) R_k, one-step predicted state error covariance matrix (PECM) $P_{k:k-1}$, and state error covariance matrix P_k at the same time.

3.3. Improved by Introducing Multiple Fading Factors

If the statistical characteristics of the process and measurement noise are time-varying, the convergence speed of VBAKF will slow down, and there will be a particular error in the estimation result, which will be reflected by the residual sequence γ_k [31].
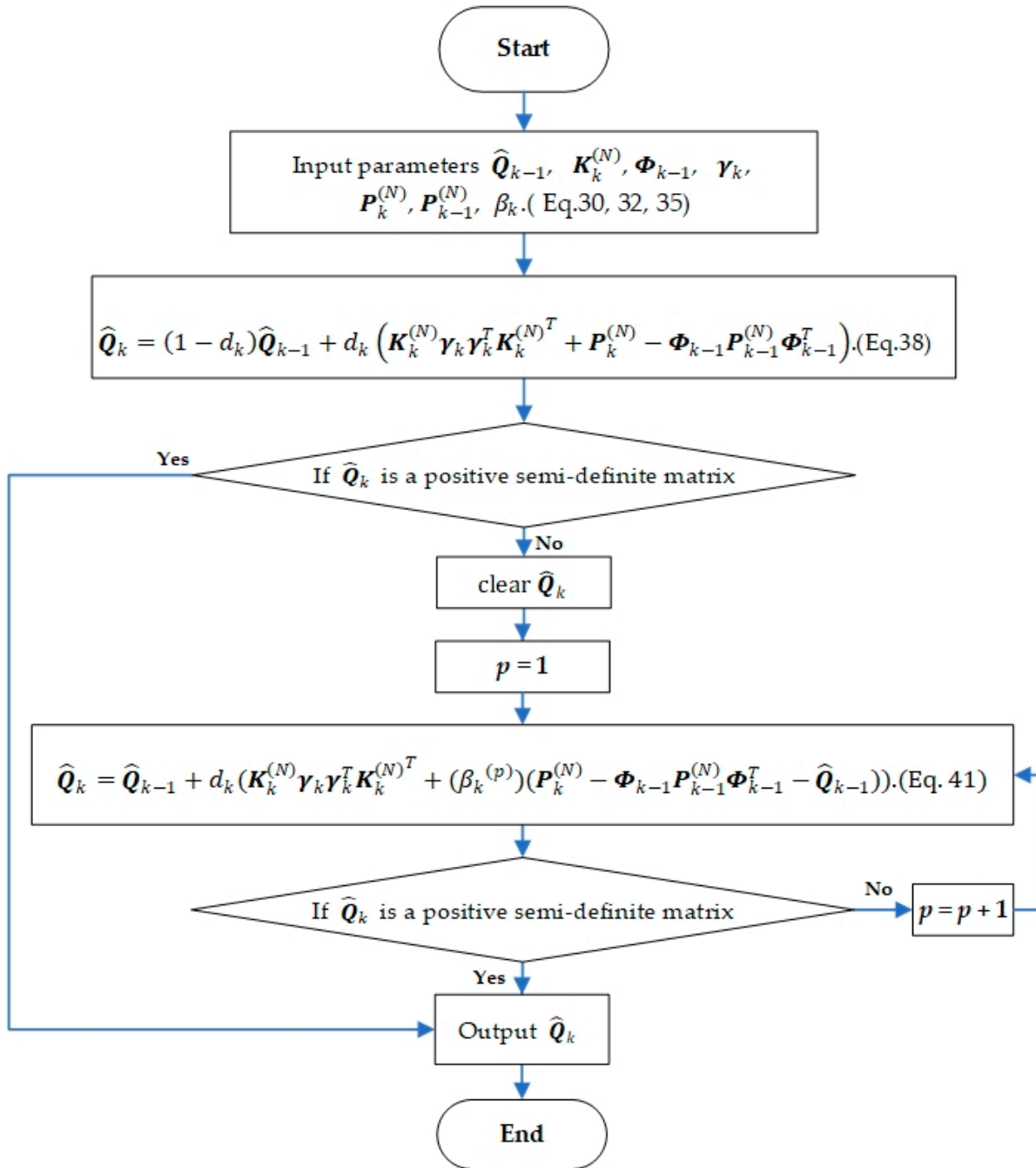
In view of this, to improve the accuracy of estimation, the multiple fading factor L_{md} is introduced to realize the correction of the one-step predicted error covariance matrix (PECM) $P_{k:k-1}$. Equation (7) can be rewritten as:

$$P_{k:k-1}^* = L_{md} \Phi_{k-1} P_{k-1} \Phi_{k-1}^T + \hat{Q}_{k-1}, \tag{44}$$

adjusting the gain $K_k^{(N)}$ in real-time to keep γ_k orthogonal, forcing the filter to keep track of the actual state of the system. The tracking ability is thereby improved [31]. The calculation method of multiple fading factors L_{md} is as follows:

$$L_{md} = diag\left[\lambda_k^1, \lambda_k^2, \cdots, \lambda_k^n \right], \tag{45}$$

Figure 1. The flowchart of one time-step of the updated monitoring strategy based on maximum a posterior (MAP) for estimating the process noise covariance matrix (PNCM).

$$\lambda_k^i = \begin{cases} \lambda_k^{0i}, & \lambda_k^{0i} > 1 \\ 1, & \lambda_k^{0i} \le 1 \end{cases} \quad i = 1, 2, \cdots, n, \tag{46}$$

where n is the dimension of the state vector X_k, $\lambda_k^{0i} = \alpha^i \cdot G_k$, the value of α^i is determined by the system prior information. The formula of G_k is as follows:

$$G_k = \frac{trace(N_k)}{\sum_{i=1}^n \alpha^i \cdot M_k^{ii}}, \tag{47}$$

where M_k^{ii} is the *i*-th element of the main diagonal of M_k, the calculation formulas of N_k and M_k are as follows:

$$N_k = B_k - \boldsymbol{H}_{k-1}\hat{\boldsymbol{Q}}_{k-1}\boldsymbol{H}_{k-1}^T - \tau\hat{\boldsymbol{R}}_k, \tag{48}$$

$$M_k = \boldsymbol{\Phi}_{k-1}P_{k:k}\boldsymbol{\Phi}_{k-1}^{T}\boldsymbol{H}_{k-1}^T\boldsymbol{H}_{k-1}. \tag{49}$$

In Equation (48), τ is the weakening factor and B_k is unknown and can be estimated by the following formula:

$$B_k = \begin{cases} \gamma_k\gamma_k^T, & k = 1 \\ \frac{\mu B_{k-1}+\gamma_k\gamma_k^T}{1+\mu}, & k > 1 \end{cases} \tag{50}$$

where $\mu \in (0, 1]$ is the forgetting factor.

Since L_{md} is used to correct the one-step predicted error covariance matrix (PECM) in the prediction step of the filtering algorithm, the initial value $\hat{\boldsymbol{R}}_1$ of the MNCM $\hat{\boldsymbol{R}}_k$ must be set in advance, assuming that $\hat{\boldsymbol{R}}_1$ also obeys the inverse Wishart distribution. The estimated value of measurement noise covariance matrix $\hat{\boldsymbol{R}}_k$ in the time-update step is defined as:

$$\hat{\boldsymbol{R}}_k = \begin{cases} \frac{\hat{T}_0}{(\hat{t}_0-m-1)}, & k = 1 \\ \hat{\boldsymbol{R}}_{k-1}, & k > 1 \end{cases}, \tag{51}$$

In Equation (51), the variation range of the slowly time-varying measurement noise covariance matrix is small, and the estimated value $\hat{\boldsymbol{R}}_{k-1}$ at the previous time still has a great reference value for the current time estimation. Therefore, $\hat{\boldsymbol{R}}_{k-1}$ estimated by the variational update recursively of the previous time is used as $\hat{\boldsymbol{R}}_k$ at time $k > 1$. $\hat{\boldsymbol{R}}_k$ is used as the time-varying parameter of L_{md} to modify the one-step predicted error covariance matrix (PECM) $\boldsymbol{P}_{k:k-1}^*$ more accurately, and the accurate $\boldsymbol{P}_{k:k-1}^*$ can affect the accuracy of the estimation result of the variational iteration recursion directly.

The multi-fading factor and updated monitoring strategy AKF-based variational Bayesian in this paper is composed of Equations (6), (12)–(14), (22)–(25), (27), (29)–(34), (41)–(51). The pseudo-code implementation of the proposed MFMS-VBAKF algorithm is listed in Algorithm 1.

Algorithm 1 One time-step of the proposed multi-fading factor and updated monitoring strategy AKF-based variational Bayesian

---

Inputs: $\hat{X}_{k-1:k-1}$, $P_{k-1:k-1}$, $\hat{t}_{k-1:k-1}$, $\hat{T}_{k-1:k-1}$, Φ, H, Z_k, m, n, ρ, b, μ, τ, α^i, N

Time update:

1. $\hat{X}_{k:k-1} = \Phi X_{k-1:k-1}$, $\gamma_k = Z_k - \hat{X}_{k:k-1}$

2. **if** $k = 1$ **then**

3. $\hat{R}_k = \hat{T}_0 / (\hat{t}_0 - m - 1)$, $Q_k = \tilde{Q}_0$

4. $B_k = \gamma_k \gamma_k^T$,

5. **else**

6. $\hat{R}_k = \hat{R}_{k-1}$, $Q_k = \hat{Q}_{k-1}$

7. $B_k = \mu B_{k-1} + \gamma_k \gamma_k^T / 1 + \mu$

8. $N_k = B_k - H_{k-1}\hat{Q}_{k-1}H_{k-1}^T - \tau\hat{R}_k$,

9. $M_k = \Phi_{k-1}P_{k:k}\Phi_{k-1}{}^T H_{k-1}^T H_{k-1}$,

10. $G_k = trace(N_k) / \sum\limits_{i=1}^{n}\left(\alpha^i \cdot M_k^{ii}\right)$, $\lambda_k^{0i} = \alpha^i \cdot G_k$

11. **if** $\lambda_k^{0i} > 1$

12. $\lambda_k^i = \lambda_k^{0i}$

13. **else**

14. $\lambda_k^i = 1$

15. $L_{md} = diag\left[\lambda_k^1, \lambda_k^2, \cdots, \lambda_k^n\right]$

16. $d_k = (1-b) / \left(1 - b^{k+1}\right)$

17. $\hat{Q}_k = (1-d_k)\hat{Q}_{k-1} + d_k\left(K_k\gamma_k\gamma_k^T K_k{}^T + P_k - \Phi_{k-1}P_{k-1}\Phi_{k-1}^T\right)$, $c = 1$

18. $\theta = min(eig(\hat{Q}_k))$,

19. $a_k = trace\left(P_k - \Phi_{k-1}P_{k-1}\Phi_{k-1}^T - \hat{Q}_{k-1}\right) / trace\left(K_k\gamma_k\gamma_k^T K_k{}^T\right)$

20. $\beta_k = exp^{-a_k}$

21. **while** $\theta < 0$ **do**

22. $\hat{Q}_k = \hat{Q}_{k-1} + d_k\left(K_k\gamma_k\gamma_k^T K_k{}^T + (\beta_k)^p\left(P_k - \Phi_{k-1}P_{k-1}\Phi_{k-1}^T - \hat{Q}_{k-1}\right)\right)$

23. $\theta = min(eig(\hat{Q}_k))$, $c = c + 1$

24. $P_{k:k-1}^* = L_{md}\Phi_{k-1}P_{k-1:k-1}\Phi_{k-1}{}^T + \hat{Q}_{k-1}$

Iterated measurement update

25. **Initialization**

26. $V_k^{(i)} = \left(Z_k - H_k\hat{X}_k^i\right)\left(Z_k - H_k\hat{X}_k^i\right)^T + H_k P_k H_k^T$

27. $\hat{t}_k^{(i+1)} = \hat{t}_{k:k-1} + 1$, $\hat{T}_k^{(i+1)} = V_k^{(i)} + \hat{T}_k^{(i)}$

for $i = 0 : N - 1$

Update $q^{(N)}(X_k) = (X_k; \hat{X}_{k:k}, P_{k:k})$ **given** $P_{k:k-1}^*$ **and** $q^{(i+1)}(R_k)$:

28. $E^{(i+1)}\left[R_k^{-1}\right]^{-1} = \left(\hat{t}_k^{(i+1)} - m - 1\right)^{-1}\hat{T}_k^{(i+1)}$

29. $\hat{R}_k^{(i+1)} = E^{(i+1)}\left[R_k^{-1}\right]^{-1}$

30. $K_k^{(i+1)} = P_{k:k-1}^* H_k^T\left(H_k P_{k:k-1}^* H_k^T + \hat{R}_k^{(i+1)}\right)^{-1}$

31. $\hat{X}_k^{(i+1)} = \hat{X}_{k:k-1} + K_k^{(i+1)}\left(Z_k - H_k\hat{X}_{k:k-1}\right)$

32. $\hat{P}_k^{(i+1)} = \left(I - K_k^{(i+1)}H_k\right)P_{k:k-1}\left(I - K_k^{(i+1)}H_k\right)^T + K_k^{(i+1)}\hat{R}_k^{(i+1)}\left(K_k^{(i+1)}\right)^T$

End for

33. $\hat{X}_{k:k} = \hat{X}_{k:k}^{(N)}$, $\hat{P}_{k:k} = \hat{P}_{k:k}^{(N)}$, $\hat{t}_{k:k} = \hat{t}_k^{(N)}$, $\hat{T}_{k:k} = \hat{T}_k^{(N)}$, $\hat{R}_k = \hat{R}_k^{(N)}$

Outputs: $\hat{X}_{k:k}$, $\hat{P}_{k:k}$, $\hat{t}_{k:k}$, $\hat{T}_{k:k}$, $\hat{R}_k$, $\hat{Q}_k$

---

4. Simulations and Results

The application of the proposed algorithm in target tracking is simulated. The target moves according to the continuous white noise accelerated motion model in the two-dimensional Cartesian coordinate system. Sensors are used to collect the target location. The system state is defined as $x_k = \left[x_k' \ \dot{x}_k' \ y_k' \ \dot{y}_k'\right]$, where x_k' and y_k' represent the Cartesian coordinates of the target at time k, $\dot{x}_k'$ and $\dot{y}_k'$ represent the velocity of the target at the

corresponding position [24,25]. The state transition matrix $\mathbf{\Phi}_{k-1}$ and the measurement matrix $\mathbf{H}_k$ are respectively set as:

$$\mathbf{\Phi}_{k-1} = \begin{bmatrix} \mathbf{I}_2 & \Delta t \mathbf{I}_2 \\ 0 & \mathbf{I}_2 \end{bmatrix}, \; \mathbf{H}_k = \begin{bmatrix} \mathbf{I}_2 & 0 \end{bmatrix}, \tag{52}$$

where the parameter $\Delta t = 1\ s$ represent the sampling interval, and $\mathbf{I}_n$ represent the n-dimensional unit matrix. Similar to [25], the true process noise covariance matrix (PNCM) and the measurement noise covariance matrix (MNCM) are set to slow time-varying models, which are:

$$\begin{cases} \mathbf{Q}_k = \left(9.5 + 0.5cos\frac{\pi k}{T}\right)q \begin{bmatrix} \frac{\Delta t^3}{3}\mathbf{I}_2 & \frac{\Delta t^2}{2}\mathbf{I}_2 \\ \frac{\Delta t^2}{2}\mathbf{I}_2 & \Delta t\mathbf{I}_2 \end{bmatrix} \\ \mathbf{R}_k = \left(0.1 + 0.05cos\frac{\pi k}{T}\right)r \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix} \end{cases} \tag{53}$$

The simulation environment is set as follows: $T = 1000$ s is the total simulation time, q is a parameter related to process noise, and r is a parameter related to measurement noise. The fixed PNCM and MNCM are set as $\widetilde{\mathbf{Q}}_k = \sigma \mathbf{I}_4$ and $\widetilde{\mathbf{R}}_k = \varepsilon \mathbf{I}_2$, respectively, where σ and ε are the prior confidence parameters used to adjust the initial fixed noise covariance matrix.

The parameters in the MFMS-VBAKF algorithm proposed in this paper are set as follows: $\sigma = 1$, $\varepsilon = 100$, changing factor $\rho = \exp(-4)$, the number of variational iterations $N = 10$, the initial value of the variational parameter $\hat{t}_0 = 1$, $\hat{\mathbf{T}}_0 = 300\mathbf{I}_2$, forgetting factor $\mu = 0.95$, the weakening factor $\tau = 0.4$, the parameter $\begin{bmatrix} \alpha^1 & \alpha^2 & \alpha^3 & \alpha^4 \end{bmatrix} = [1.7\ 1.1\ 1.7\ 1.1]$, and the attenuation factor $b = 0.96$.

This paper compares MFMS-VBAKF and true noise covariance matrix Kalman filter (TCMKF) [1], fixed noise covariance matrix Kalman filter (FCMKF) [1], SH-KF [15], ML-KF [20], ST-KF [21], and R-VBKF [26] algorithms. Table 1 lists the estimated parameters and parameter settings of the existing algorithms. All algorithms are programmed using MATLAB R2018a, and the simulation program runs on a computer with Intel® Core™ i5-6300HQ CPU at 2.30 GHz and 8GB of RAM.

Table 1. Estimated parameters and parameter settings of the existing algorithms.

Types of Filtering Algorithms	Estimated Parameters	Algorithm Parameters	Value of Algorithm Parameters
SH-KF	$\hat{X}_k$, $\hat{P}_{k:k-1}$, $\hat{P}_k$, $\hat{Q}_k$	Forgetting factor	0.96
ML-KF	$\hat{X}_k$, $\hat{P}_{k:k-1}$, $\hat{P}_k$, $\hat{Q}_k$	The size of sliding window	150
ST-KF	$\hat{X}_k$, $\hat{P}_{k:k-1}$, $\hat{P}_k$	Forgetting factor	0.94
R-VBKF	$\hat{X}_k$, $\hat{P}_{k:k-1}$, $\hat{P}_k$, $\hat{R}_k$	The number of iterations	10
		Forgetting factor	0.98

SHKF, Sage–Husa Kalman filter in [15]; MLKF, maximum likelihood Kalman filter in [20]; ST-KF, strong-tracking Kalman Filter in [21]; R-VBKF, the Kalman filter algorithm that uses variational iteration to recursively estimate $\hat{R}_k$ and $\hat{X}_k$ in [26].

With the aim of evaluating the accuracy of system state estimation, the root mean square error and average root mean square error of position and velocity are regarded as performance indicators, which are defined as follows:

$$E_{RMSE,pos} \triangleq \sqrt{\frac{1}{M}\sum_{S=1}^{M}\left(\left(x_k^s - \hat{x}_k^s\right)^2 + \left(y_k^s - \hat{y}_k^s\right)^2\right)} \tag{54}$$

$$E_{ARMSE,pos} \triangleq \sqrt{\frac{1}{MT} \sum_{k=1}^{T} \sum_{S=1}^{M} \left(\left(x_k^s - \hat{x}_k^s\right)^2 + \left(y_k^s - \hat{y}_k^s\right)^2 \right)} \tag{55}$$

where $\left(x_k', \hat{x}_k'\right)$ and $\left(y_k', \hat{y}_k'\right)$, respectively, represent the true value and estimated value of the position in the *s*-th Monte Carlo experiment and $M = 1000$ represents the total number of Monte Carlo experiment runs. Similarly, the calculation formulas of RMSE and ARMSE for the corresponding velocity can be obtained.

True and estimated trajectories of the target are shown in Figure 2. It can be seen that, compared to the existing adaptive filtering algorithm, the proposed filter maintains excellent target tracking performance throughout the entire process.

Figure 2. True and estimated trajectories of the target.

Figure 3a,b plot the RMSE variation curves of the position and velocity of the existing filter and the proposed MFMS-VBAKF, respectively. The RMSE of the estimation result of TCMKF algorithm is regarded as the benchmark. It can be seen from Figure 2 that compared with existing algorithms, the proposed algorithm has a faster convergence speed and higher accuracy. To further elaborate on the advantages of the proposed algorithm, Table 2 lists the average root mean square error (ARMSE) of different KF filtering algorithms:

(**a**) Position estimation

Figure 3. *Cont.*

(**b**) Velocity estimation

Figure 3. The root mean square errors (RMSEs) of the target position and velocity estimation.

Table 2. Average root mean square error (ARMSE) of various Kalman filter algorithms.

Filters	$E_{ARMSE,pos}$/m	$E_{ARMSE,vel}$/(m·s^{-1})
FCMKF	9.853	6.348
ML-KF	4.596	29.759
SH-KF	9.646	9.545
R-VBKF	9.352	5.961
ST-KF	6.157	5.653
MFMS-VBAKF	4.073	3.946
TCMKF	3.649	3.253

According to the data in Table 2, it can be found that in comparison with other AKF algorithms, the MFMS-VBAKF algorithm has the smallest ARMSE and the highest accuracy in estimating target position and velocity.

To compare the computational complexity with existing algorithms, Table 3 lists the single-step running time of each algorithm. It can be seen that, compared with the R-VBAKF using the variational Bayesian method, the proposed MFMS-VBAKF increases the single-step running time by 0.21 μs. The design of multi-fading factors and updated monitoring strategy has ensured a substantial increase in estimation accuracy while bringing higher computational complexity.

Table 3. Single-step running time of each algorithm.

Filters	Single-Step Running Time (μs)
FCMKF	0.28
ST-KF	0.41
SH-KF	0.46
ML-KF	0.54
R-VBKF	0.71
MFMS-VBAKF	0.92

To evaluate the accuracy of the estimation of one-step predicted state error covariance matrix PECM and the noise covariance matrices PNCM, MNCM, the square root of the normalized Frobenius norm (SRNFN) and the averaged SRNFN (ASRNFN) are used as the measure of error, which are defined as:

$$E_{SRNFN,P} \triangleq \left(\frac{1}{n^2 M} \sum_{s=1}^{M} ||\hat{P}_{k:k-1}^s - \hat{P}_{o,k:k-1}^s||^2 \right)^{1/4} \tag{56}$$

$$E_{ASRNFN,P} \triangleq \left(\frac{1}{n^2 MT} \sum_{k=1}^{T} \sum_{s=1}^{M} ||\hat{P}_{k:k-1}^s - \hat{P}_{o,k:k-1}^s||^2 \right)^{1/4} \tag{57}$$

where $\hat{P}^s_{o,k:k-1}$ and $\hat{P}^s_{k:k-1}$ represent the true value and estimated value of the noise covariance matrix or one-step predicted state error covariance matrix in the *s*-th Monte Carlo experiment, respectively. The SRNFN and ASRNFN of the estimation result of PECM are shown in Figure 4 and Table 4, respectively.
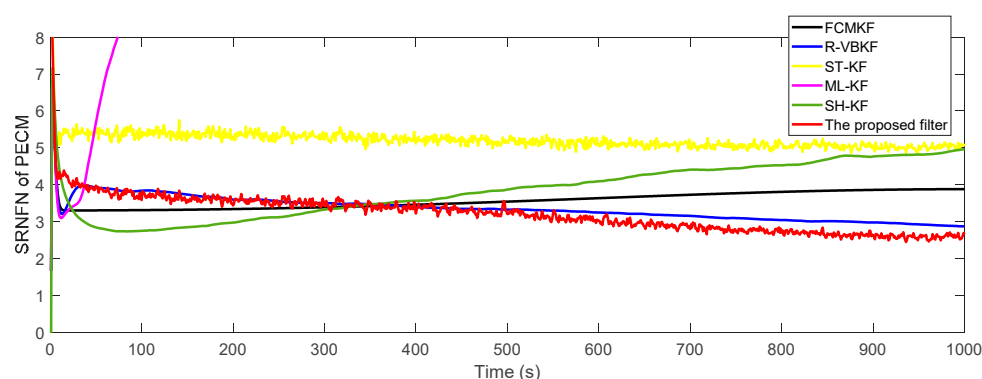

Figure 4. The SRNFN of PECM estimation.

Table 4. ASRNFN of PECM estimation of various Kalman filter algorithms.

Filters	$E_{ASRNFN,P}$
ML-KF	21.365
ST-KF	5.573
SH-KF	4.113
FCMKF	3.852
R-VBKF	3.115
MFMS-VBAKF	2.945

It can be clearly seen that, compared with the existing adaptive KF algorithm, if the noise covariance matrices are slowly time-varying, the SRNFN of the MFMS-VBAKF algorithm is smaller than the SRNFN of the current algorithm. Compared with R-VBKF with similar performance, the ASRNFN of MFMS-VBAKF is reduced by 5.45%.

Figure 5 shows the SRNFN of the measurement noise covariance matrix (MNCM) estimation. Obviously, the MFMS-VBAKF algorithm has the strongest tracking ability, the highest estimation accuracy and the fastest convergence speed of the slowly time-varying measurement noise covariance matrix estimation.
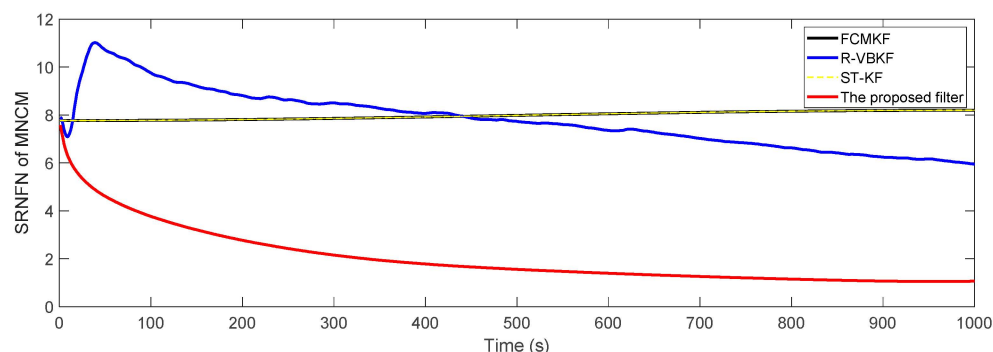

Figure 5. The square root of the normalized Frobenius norm (SRNFN) of the measurement noise covariance matrix (MNCM) estimation.

Figure 6 and Table 5 show respectively the SRNFNs and the ASRNFNs of the PNCM from the existing filters and MFMS-VBAKF algorithm. It can be seen that the SRNFN and ASRNFN of the proposed MFMS-VBAKF are both smaller than the current filters. Thus,

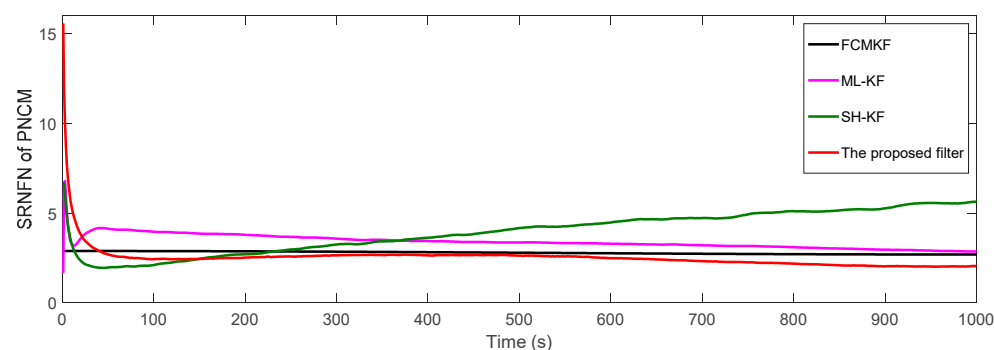the MFMS-VBAKF has better estimation accuracy and satisfactory convergence speed in PNCM estimation.

Figure 6. The SRNFN of the PNCM estimation.

Table 5. The averaged SRNFN (ASRNFN) of PNCM estimation of various Kalman filter algorithms.

Filters	$E_{ASRNFN,Q}$
SH-KF	4.915
ML-KF	3.688
FCMKF	3.357
MFMS-VBAKF	2.754

Next, we compare and analyze the influence of the values of four critical parameters (changing factor ρ, forgetting factor μ, weakening factor τ and attenuation factor b) in the MFMS-VBAKF algorithm on the estimated effect.

Figure 7 shows the RMSEs of position and velocity from the existing filters and MFMS-VBAKF in the case of $\rho = 0.85, 0.93, 0.95, 1 - \exp(-4), 1.0$. The MFMS-VBAKF with $\rho = 0.85, 0.93, 0.95, 1 - \exp(-4), 1.0$ has better estimation accuracy than existing adaptive KF filters. And when $\rho = 1 - \exp(-4)$, the MFMS-VBAKF algorithm has the best estimation accuracy and convergence performance.
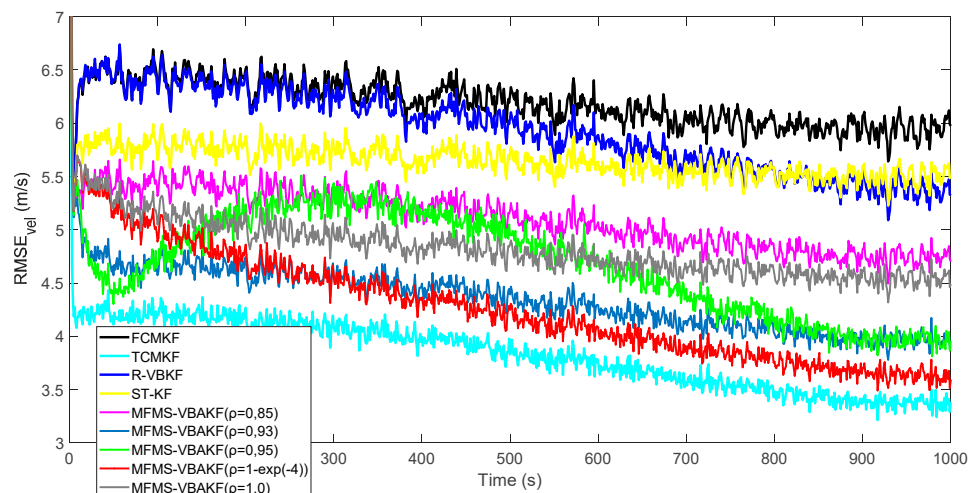
Figure 8 plots the RMSE curves of position and velocity from the existing filters and the MFMS-VBAKF in the case of $\mu = 0.65, 0.75, 0.85, 0.95, 1.0$. The MFMS-VBAKF with $\mu = 0.65, 0.75, 0.85, 0.95, 1.0$. has better estimation accuracy than existing adaptive KF filters, and when $\mu = 0.95$, the MFMS-VBAKF algorithm has the best estimation accuracy and convergence performance.

Figure 9 plots the RMSE curves of position and velocity from the existing filters and the MFMS-VBAKF in the case of $\tau = 0.2, 0.4, 0.6, 0.8, 1.0$. The MFMS-VBAKF with $\tau = 0.2, 0.4, 0.6, 0.8, 1.0$ has better estimation accuracy than existing adaptive KF filters, and when $\tau = 0.4$, the MFMS-VBAKF algorithm has the best estimation accuracy and convergence performance.

Figure 10 shows the RMSEs of position and velocity from the existing filters and MFMS-VBAKF in the case of $b = 0.76, 0.86, 0.96, 0.99$. The MFMS-VBAKF with $b = 0.76, 0.86, 0.96, 0.99$. has better estimation accuracy than existing adaptive KF filters. And when $b = 0.96$, MFMS-VBAKF algorithm has the best estimation accuracy and convergence performance.
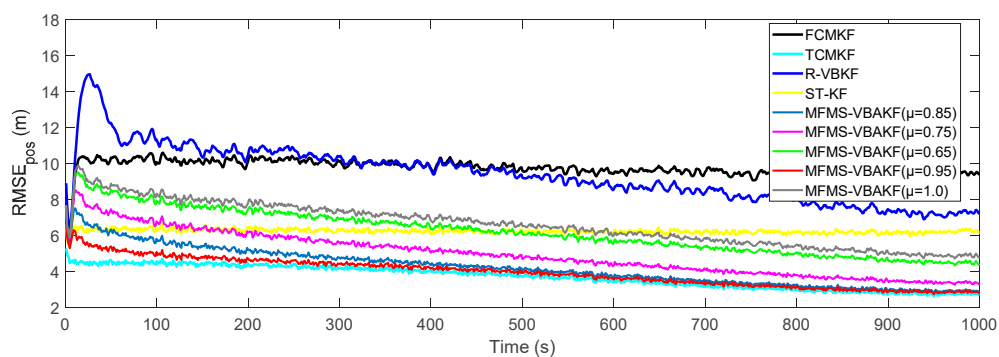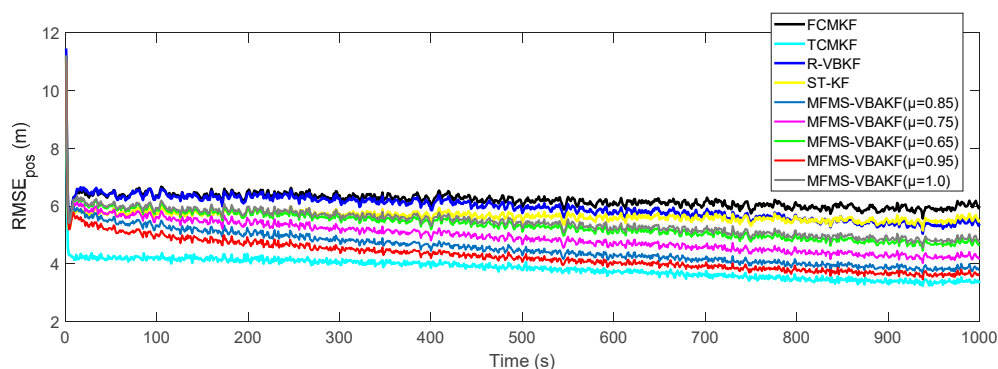
(**a**) Position estimation

(**b**) Velocity estimation

Figure 7. The RMSEs of the position and velocity estimation in the case of $\rho = 0.85,\ 0.93,\ 0.95,\ 1 - \exp(-4),\ 1.0$.
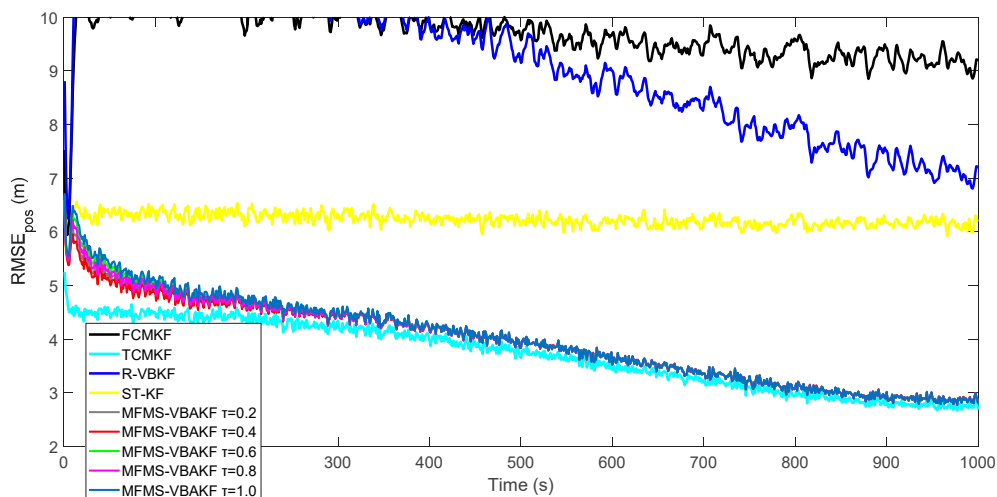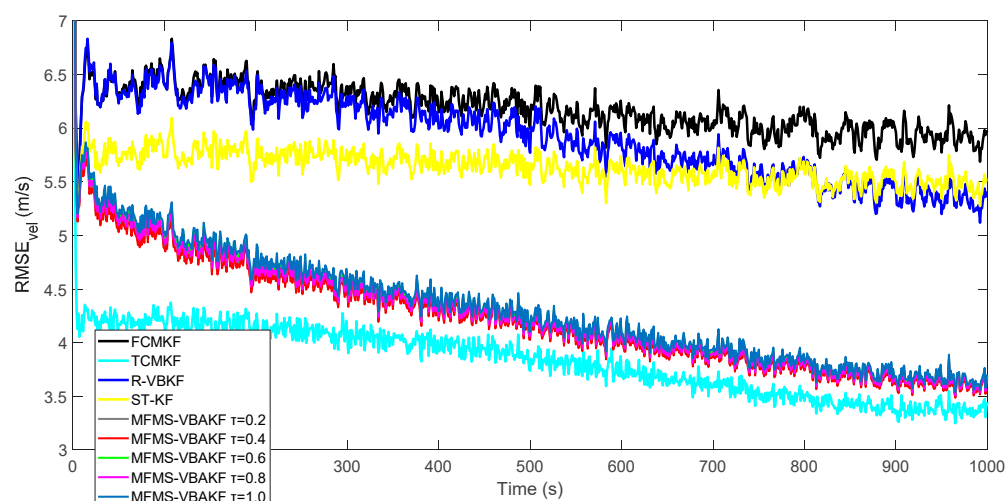

(**a**) Position estimation

Figure 8. *Cont.*
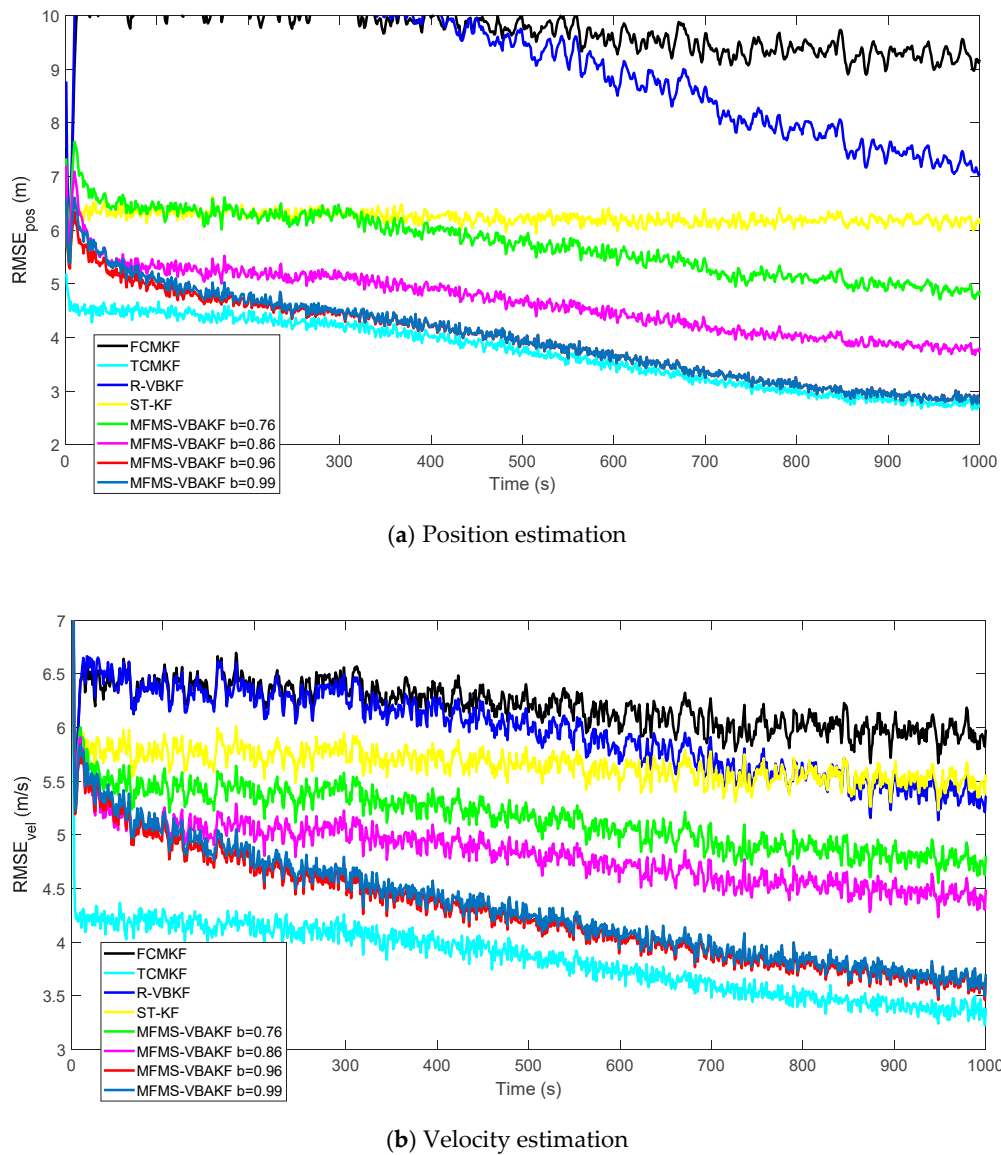
(**b**) Velocity estimation

Figure 8. The RMSE of the position and velocity estimation in the case of μ = 0.65, 0.75, 0.85, 0.95, 1.0.

(**a**) Position estimation
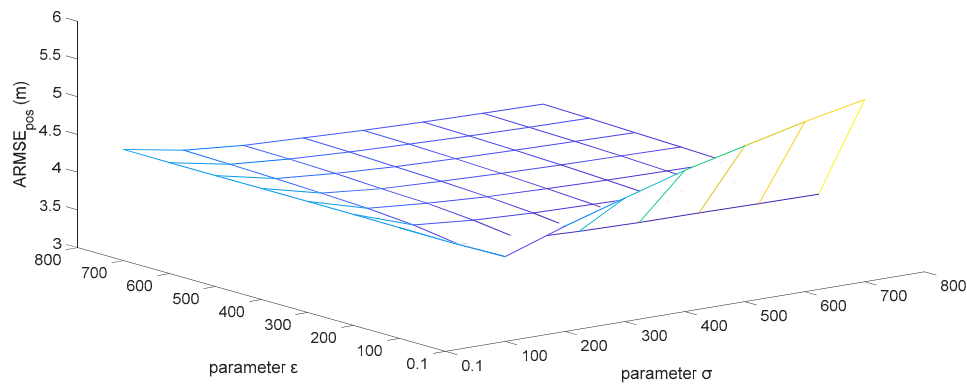

(**b**) Velocity estimation

Figure 9. The RMSE of the position and velocity estimation in the case of $\tau = 0.2$, 0.4, 0.6, 0.8, 1.0.

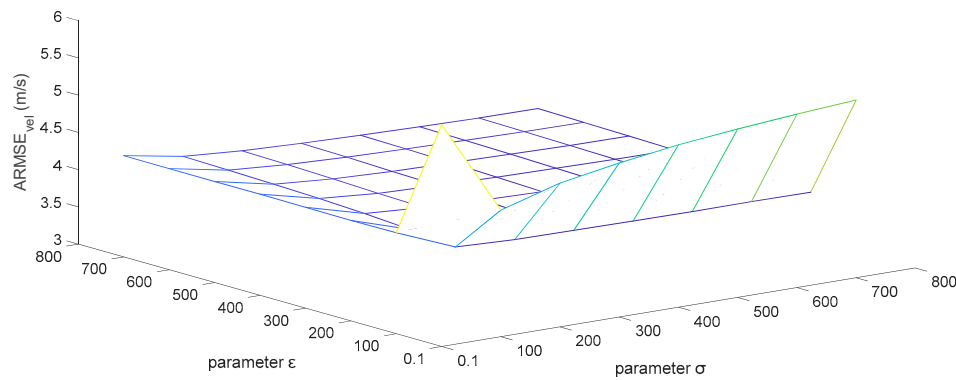(**a**) Position estimation

(**b**) Velocity estimation

Figure 10. The RMSE of the position and velocity estimation in the case of $b = 0.76, 0.86, 0.96, 0.99$.

For the sake of testing the robustness of the adaptive correction capability of the MFMS-VBAKF algorithm when the fixed noise covariance matrices are set to different initial values, the priori confidence parameters σ and ε are set to change in combination within the grid area of $(\sigma, \varepsilon) \in [0.1, 800] \times [0.1, 800]$. The ARMSEs estimated by the algorithm for position and velocity are displayed in Figure 10.

It can be analyzed from Figure 11 that the ARMSEs of position and velocity estimation are flat in a large area of the set grid, and the estimation results are close to the actual values. However, the initial setting values of the fixed noise covariance matrices in the extremely narrow area on the right edge of the grid are too different from the actual values, which leads to unsatisfactory performance of the estimation results. This is caused by the variational Bayesian method that can only guarantee local convergence. In general, the estimated effects of the MFMS-VBAKF algorithm can converge to near the actual values, with excellent robust performance.

(**a**) Position estimation

(**b**) Velocity estimation

Figure 11. The ARMSEs of the position and velocity estimation under a combination of $(\sigma, \varepsilon) \in [0.1, \ 800] \times [0.1, \ 800]$.

5. Conclusions

This paper presents a multi-fading factor and updated monitoring strategy AKF-based variational Bayesian with the inaccurate time-varying process and measurement noise covariance matrices. The model of measurement error is defined as the inverse Wishart distribution, and the variational Bayesian method is used to recursively estimate the measurement noise covariance matrix and the system state, the estimation results of the two can be approximated to the true value. The process noise covariance matrix is estimated by the maximum a posteriori principle, and the updated monitoring strategy with adjustment factors is used to guarantee the positive semi-definiteness of the updated matrix. The estimated value of measurement noise covariance obtained by the variational iteration recursion and the estimated value of process noise covariance obtained by updating the monitoring strategy are used as time-varying parameters of multiple fading factors, which can be corrected to obtain more accurate state predicted error covariance. Variational Bayesian and the updated monitoring strategy and multi-fading factors complement each other, which not only enhances the responsiveness of target tracking, but also improves the estimation accuracy of variational iteration recursion. The simulation results show that the proposed MFMS-VBAKF algorithm realizes the simultaneous estimation of the process noise covariance matrix and the measurement noise covariance matrix, and has achieved satisfactory results in terms of estimation accuracy, convergence performance, and robustness.

6. Future Work

Compared with existing filtering algorithms, the MFMS-VBAKF algorithm proposed in this paper exhibits excellent performance in the state estimation problem of linear systems with inaccurate time-varying processes and measurement noise covariance matrix. However, in real applications, the influence of the control input of the system should not be underestimated. Control input will cause large changes in noise, and even noise will obey non-Gaussian distribution. Moreover, most of the systems in real applications are nonlinear systems. In the above cases, MFMS-VBAKF will not be applicable. Therefore, in future work, the theory of this paper will be further extended to the problem of nonlinear adaptive Kalman filtering. The problem of nonlinear AKF with inaccurate state transition matrix and non-Gaussian noise distribution will be further studied.

Author Contributions: C.S. conceived the idea for this paper, designed the experiments and wrote the paper; Z.J. assisted in model designing and experiments; W.Z. and Y.Y. polished the language. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Derivation of (21)

The Bayesian estimation theory states that the joint PDF can be defined as:

$$p(X_k, R_k, Z_{1:k}) = p(Z_k|X_k, R_k)p(X_k|Z_{1:k-1})p(R_k|Z_{1:k-1})p(Z_{1:k-1}) \tag{A1}$$

Equations (3), (4), and (11) are substituted into (A1) to derive:

$$p(X_k, R_k, Z_{1:k}) = N(Z_k; H_k X_k, R_k)N(X_{k-1}; \hat{X}_{k:k-1}, P_{k:k-1})IW(R_{k-1}; \hat{t}_{k-1:k-1}, \hat{T}_{k-1:k-1})p(Z_{1:k-1}) \tag{A2}$$

Combined with Equation (5), $\log N(G; \mu, \Sigma)$ can be rewritten as:

$$\begin{aligned}
\log N(G; \mu, \Sigma) &= \log\left[\frac{1}{\sqrt{|2\pi\Sigma|}}exp^{-\frac{1}{2}(G-\mu)^T\Sigma^{-1}(G-\mu)}\right] \\
&= \log\frac{1}{\sqrt{2\pi}} - \frac{1}{2}\log|\Sigma| - \frac{1}{2}(G-\mu)^T\Sigma^{-1}(G-\mu) \\
&= -\frac{1}{2}\log|\Sigma| - \frac{1}{2}(G-\mu)^T\Sigma^{-1}(G-\mu) + c_G,
\end{aligned} \tag{A3}$$

where c_G denotes a constant related to the variable G.

Combined with Equation (8), $\log IW(C; \lambda, \Psi)$ can be rewritten as:

$$\begin{aligned}
\log IW(C; \lambda, \Psi) &= \log\left\{\frac{|\Psi|^{\lambda/2}}{2^{\lambda d/2}\Gamma_d\left(\frac{\lambda}{2}\right)}|C|^{-(\lambda+d+1)/2}exp^{-\frac{1}{2}tr[\Psi C^{-1}]}\right\} \\
&= \frac{\lambda}{2}\log|\Psi| - \frac{(\lambda+d+1)}{2}\log|C| - \frac{1}{2}tr[\Psi C^{-1}] - \frac{\lambda d}{2}\log 2 - \log\left|\Gamma_d\left(\frac{\lambda}{2}\right)\right| \\
&= -\frac{(\lambda+d+1)}{2}\log|C| - \frac{1}{2}tr[\Psi C^{-1}] + c_C,
\end{aligned} \tag{A4}$$

where c_C denotes the constant related to the variable C.

Combining (A2)–(A4), the logarithmic form of the joint PDF can be written as:

$$\log p(\boldsymbol{X}_k, \boldsymbol{R}_k, \boldsymbol{Z}_{1:k}) = \log\{N(\boldsymbol{Z}_k; \boldsymbol{H}_k\boldsymbol{X}_k, \boldsymbol{R}_k)N(\boldsymbol{X}_{k-1}; \hat{\boldsymbol{X}}_{k:k-1}, \boldsymbol{P}_{k:k-1})$$

$$\times IW(\boldsymbol{R}_{k-1}; \hat{\boldsymbol{t}}_{k-1:k-1}, \hat{\boldsymbol{T}}_{k-1:k-1})p(\boldsymbol{Z}_{1:k-1})$$

$$= -\tfrac{1}{2}\log|\boldsymbol{R}_k| - \tfrac{1}{2}(\boldsymbol{Z}_k - \boldsymbol{H}_k\boldsymbol{X}_k)^T\boldsymbol{R}_k^{-1}(\boldsymbol{Z}_k - \boldsymbol{H}_k\boldsymbol{X}_k) - \tfrac{1}{2}\log|\boldsymbol{P}_{k:k-1}|$$

$$-\tfrac{1}{2}(\boldsymbol{X}_k - \hat{\boldsymbol{X}}_{k:k-1})^T\boldsymbol{P}_{k:k-1}^{-1}(\boldsymbol{X}_k - \hat{\boldsymbol{X}}_{k:k-1}) - \tfrac{(\hat{t}_{k:k-1}+m+1)}{2}\log|\boldsymbol{R}_k| - \tfrac{1}{2}tr\left(\hat{\boldsymbol{T}}_{k:k-1}\boldsymbol{R}_k^{-1}\right) + c_\varnothing \qquad (A5)$$

$$= -\tfrac{(\hat{t}_{k:k-1}+m+2)}{2}\log|\boldsymbol{R}_k| - \tfrac{1}{2}(\boldsymbol{Z}_k - \boldsymbol{H}_k\boldsymbol{X}_k)^T\boldsymbol{R}_k^{-1}(\boldsymbol{Z}_k - \boldsymbol{H}_k\boldsymbol{X}_k) - \tfrac{1}{2}tr\left(\hat{\boldsymbol{T}}_{k:k-1}\boldsymbol{R}_k^{-1}\right)$$

$$-\tfrac{1}{2}\log|\boldsymbol{P}_{k:k-1}| - \tfrac{1}{2}(\boldsymbol{X}_k - \hat{\boldsymbol{X}}_{k:k-1})^T\boldsymbol{P}_{k:k-1}^{-1}(\boldsymbol{X}_k - \hat{\boldsymbol{X}}_{k:k-1}) + c_\varnothing$$

where $c_\varnothing$ denotes a constant related to the variables $\boldsymbol{X}_k$ and $\boldsymbol{R}_k$.

Bring (A5) to Equation (20) to get the further derivation of Equation (21):

$$\log q^{(i+1)}(\boldsymbol{R}_k) = E_{\boldsymbol{X}_k}^{(i)}[\log p(\boldsymbol{X}_k, \boldsymbol{R}_k, \boldsymbol{Z}_{1:k})] + c_{\boldsymbol{R}_k}$$

$$= -\tfrac{(\hat{t}_{k:k-1}+m+2)}{2}\log|\boldsymbol{R}_k| - \tfrac{1}{2}E^{(i)}\left[tr\left(\left(V_k^{(i)} + \hat{\boldsymbol{T}}_{k:k-1}\right)\boldsymbol{R}_k^{-1}\right)\right] - \tfrac{1}{2}E^{(i)}\left[tr\left(\hat{\boldsymbol{T}}_{k:k-1}\boldsymbol{R}_k^{-1}\right)\right]$$

$$-\tfrac{1}{2}E^{(i)}\left[(\boldsymbol{Z}_k - \boldsymbol{H}_k\boldsymbol{X}_k)^T\boldsymbol{R}_k^{-1}(\boldsymbol{Z}_k - \boldsymbol{H}_k\boldsymbol{X}_k)\right] - \tfrac{1}{2}E^{(i)}\left[(\boldsymbol{X}_k - \hat{\boldsymbol{X}}_{k:k-1})^T\boldsymbol{P}_{k:k-1}^{-1}(\boldsymbol{X}_k - \hat{\boldsymbol{X}}_{k:k-1})\right] \qquad (A6)$$

$$-E^{(i)}\left[\tfrac{1}{2}\log|\boldsymbol{P}_{k:k-1}|\right] + c_\varnothing$$

$$= -\tfrac{(\hat{t}_{k:k-1}+m+2)}{2}\log|\boldsymbol{R}_k| - \tfrac{1}{2}tr\left(\left(V_k^{(i)} + \hat{\boldsymbol{T}}_{k:k-1}\right)\boldsymbol{R}_k^{-1}\right) + c_{\boldsymbol{R}_k}$$

where

$$c_{\boldsymbol{R}_k} = -E^{(i)}\left[\tfrac{1}{2}\log|\boldsymbol{P}_{k:k-1}|\right] - \tfrac{1}{2}E^{(i)}\left[(\boldsymbol{Z}_k - \boldsymbol{H}_k\boldsymbol{X}_k)^T\boldsymbol{R}_k^{-1}(\boldsymbol{Z}_k - \boldsymbol{H}_k\boldsymbol{X}_k)\right] + c_\varnothing. \qquad (A7)$$

Equation (26) can be derived in a similar way.

References

1. Simon, D. The Discrete-Time Kalman Filter. In *Optimal State Estimation: Kalman, H∞, and Nonlinear Approaches*; John Wiley & Sons Publishing: Hoboken, NJ, USA, 2006; pp. 121–148.
2. Kawase, T.; Hideshi, T.; Ehara, N.; Sasase, I. A Kalman Tracker with a Turning Acceleration Estimator. *Electron. Commun. Jpn.* **2001**, *84*, 1–11. [CrossRef]
3. Kaur, H.; Sahambi, J.S. Vehicle Tracking in Video Using Fractional Feedback Kalman Filter. *IEEE Trans. Comput. Imaging* **2016**, *4*, 550–561. [CrossRef]
4. Meng, Z.; Zhou, W.; Gazor, S. Robust widely linear beamforming using estimation of extended covariance matrix and steering vector. *EURASIP J. Wirel. Commun. Netw.* **2020**, *1*, 1–20. [CrossRef]
5. Zhang, Y.G.; Dang, Y.F.; Li, N.; Huang, Y.L. An Integrated Navigation Algorithm Based on Distributed Kalman Filter. In Proceedings of the 2015 IEEE International Conference on Information and Automation, Lijiang, China, 8–10 August 2015; pp. 2132–2135.
6. Zhu, W.; Wang, W.; Yuan, G.N. An Improved Interacting Multiple Model Filtering Algorithm Based on the Cubature Kalman Filter for Maneuvering Target Tracking. *Sensors* **2016**, *16*, 805. [CrossRef] [PubMed]
7. Susmita, B.; Dinesh, M. Kalman Filter-Based Raim for Reliable Aircraft Positioning with Gps and Navic Constellations. *Sensors* **2020**, *20*, 6606.
8. Kumar, G.S.; Ghosh, R.; Ghose, D.; Vengadarajan, A. Guidance of Seekerless Interceptors Using Innovation Covariance Based Tuning of Kalman Filters. *J. Guid. Control Dyn.* **2017**, *40*, 603–614. [CrossRef]
9. Caccia, M.; Bibuli, M.; Bono, R.; Bruzzone, G. Basic Navigation, Guidance and Control of an Unmanned Surface Vehicle. *Auton. Robot.* **2008**, *25*, 349–365. [CrossRef]
10. Liu, Y.H.; Fan, X.Q.; Lv, C. An Innovative Information Fusion Method with Adaptive Kalman Filter for Integrated Ins/Gps Navigation of Autonomous Vehicles. *Mech. Syst. Signal Process.* **2018**, *100*, 605–616. [CrossRef]
11. Sun, F.C.; Hu, X.S.; Zou, Y.; Li, S.G. Adaptive Unscented Kalman Filtering for State of Charge Estimation of a Lithium-Ion Battery for Electric Vehicles. *Energy* **2011**, *36*, 3531–3540. [CrossRef]
12. Bisht, S.; Mahendra, P.S. An Adaptive Unscented Kalman Filter for Tracking Sudden Stiffness Changes. *Mech. Syst. Signal Process.* **2014**, *49*, 181–195. [CrossRef]
13. Mohamed, A.H. *Optimizing the Estimation Procedure in INS/GPS Integration for Kinematic Applications*; University of Calgary: Calgary, AB, Canada, 1999.

14. Xu, F.; Su, Y.; Liu, H. Research of Optimized Adaptive Kalman Filtering. In Proceedings of the 26th Chinese Control and Decision Conference (2014 CCDC), Changsha, China, 31 May–2 June 2014; pp. 1210–1214.

15. Sage, A.P.; Husa, G.W. Adaptive filtering with unknown prior statistics. In Proceedings of the Joint Automatic Control Conference, Boulder, CO, USA, 5–7 August 1969; pp. 760–769.

16. Mehra, R. Approaches to Adaptive Filtering. *IEEE Trans. Autom. Control* **1972**, *17*, 693–698. [CrossRef]

17. Zhu, Z.; Liu, S.R.; Zhang, B.T. An improved Sage-Husa adaptive filtering algorithm. In Proceedings of the 31st Chinese Control Conference, Hefei, China, 25–27 July 2012; pp. 5113–5117.

18. Xu, S.Q.; Zhou, H.Y.; Wang, J.Q.; He, Z.M.; Wang, D.Y. Sins/Cns/Gnss Integrated Navigation Based on an Improved Federated Sage-Husa Adaptive Filter. *Sensors* **2019**, *19*, 3812. [CrossRef] [PubMed]

19. Fu, M.Y.; de Souza, C.E.; Luo, Z.Q. Finite-Horizon Robust Kalman Filter Design. *IEEE Trans. Signal Process.* **2001**, *49*, 2103–2112.

20. Mohamed, A.; Schwarz, K. Adaptive Kalman Filtering for INS/GPS. *J. Geod.* **1999**, *73*, 193–203. [CrossRef]

21. Chen, X.Y.; Shen, C.; Zhang, W.B.; Tomizuka, M.; Xu, Y.; Chiu, K.L. Novel Hybrid of Strong Tracking Kalman Filter and Wavelet Neural Network for Gps/Ins During Gps Outages. *Measurement* **2013**, *46*, 3847–3854. [CrossRef]

22. Chang, G.B.; Liu, M. An Adaptive Fading Kalman Filter Based on Mahalanobis Distance. *Proc. Inst. Mech. Eng. Part G J. Aerosp. Eng.* **2015**, *229*, 1114–1123. [CrossRef]

23. Sarkka, S.; Nummenmaa, A. Recursive Noise Adaptive Kalman Filtering by Variational Bayesian Approximations. *IEEE Trans. Autom. Control* **2009**, *54*, 596–600. [CrossRef]

24. Huang, Y.; Zhang, Y.; Li, N.; Wu, Z.; Chambers, J.A. A Novel Robust Student's t-Based Kalman Filter. *IEEE Trans. Aerosp. Electron. Syst.* **2017**, *53*, 1545–1554. [CrossRef]

25. Tohid, A. Approximate Bayesian Smoothing with Unknown Process and Measurement Noise Covariances. *IEEE Signal Process. Lett.* **2015**, *22*, 2450–2454.

26. Särkkä, S.; Hartikainen, J. Variational Bayesian Adaptation of Noise Covariances in Non-Linear Kalman Filtering. *arXiv* **2013**, arXiv:1302.0681v1.

27. Li, Z.F.; Zhang, J.Y.; Wang, J.G.; Zhou, Q.S. Recursive Noise Adaptive Extended Object Tracking by Variational Bayesian Approximation. *IEEE Access* **2019**, *7*, 151168–151179. [CrossRef]

28. Berger, J.O. Prior Information and Subjective Probability. In *Statistical Decision Theory and Bayesian Analysis*; Springer: New York, NY, USA, 1985; pp. 74–117.

29. Lindley, D. Kendall's Advanced Theory of Statistics, Volume 2b, Bayesian Inference, 2nd Edn. *J. R. Stat. Soc. Ser. A* **2005**, *168*, 259–260. [CrossRef]

30. Tzikas, D.G.; Likas, A.C.; Galatsanos, N.P. The Variational Approximation for Bayesian Inference. *IEEE Signal Process. Mag.* **2008**, *25*, 131–146. [CrossRef]

31. Zhou, D.H.; Frank, P. Strong tracking Kalman filtering of nonlinear time-varying stochastic systems with coloured noise: Application to parameter estimation and empirical robustness analysis. *Int. J. Control* **1996**, *65*, 295–307. [CrossRef]