Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

A Robust Adaptive Sliding-Window Variational Bayesian Filter with Unknown Noise Covariance Statistics

Peng Gu
Jiangsu Vocational College of Information Technology
Zhongliang Jing
zljing@sjtu.edu.cn
Shanghai Jiao Tong University
Minzhe Li
Shanghai Jiao Tong University

Research Article

Keywords: Target state estimation, Variational Bayesian, Sliding-window technique, Square root implementation, Noise covariance matrices

Posted Date: November 16th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11498837/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

A Robust Adaptive Sliding-Window Variational Bayesian Filter with Unknown Noise Covariance Statistics

Peng Gu[1,2], Zhongliang Jing[2*], Minzhe Li[2]

[1]Jiangsu Vocational College of Information Technology, Wuxi, China

[2]School of Aeronautics and Astronautics, Shanghai Jiao Tong University, Shanghai, China

Abstract: Target tracking performance is severely limited under unknown noise covariance matrices, as conventional filtering algorithms fail to produce satisfactory optimal state estimates. To address this bottleneck, this work develops a novel adaptive robust sliding-window variational Bayesian filter. Within this framework, both measurement noise covariance matrix (MNCM) and process noise covariance matrix (PNCM) are modeled with inverse Wishart (IW) distributions. The variational Bayesian (VB) is adopted to jointly iterate over target state, MNCM and PNCM for simultaneous estimation. The obtained optimal estimations are fed as time varying inputs into multiple fading factors to compensate the prediction error covariance matrix (PECM), which boosts overall estimation precision. Besides, an enhanced adaptive sliding-window strategy is constructed to calibrate prior states from posterior knowledge, yielding better estimation accuracy together with higher computational efficiency. A square root implementation is further derived to guarantee numerical stability. It maintains symmetry and positive definiteness of covariance matrices and avoids filter divergence. Simulation experiments verify that the presented approach can effectively mitigate performance degradation induced by imprecise noise covariance, and outperforms mainstream existing filters in target state estimation.

Keywords: Target state estimation; Variational Bayesian; Sliding-window technique; Square root implementation; Noise covariance matrices

1 Introduction

Benefiting from broad application prospects in both military and civilian scenarios, optimal state estimation has drawn growing research interest within the target-tracking community [1-3]. Target-tracking techniques reconstruct target trajectories and infer time-varying motion parameters by mining useful information obtained from sensor measurements [4,5]. Tracking accuracy heavily depends on the performance of the filtering algorithm. Sensors are capable of providing statistical properties for measurement noise. Nevertheless, external interference may distort these statistics and make them less credible in practical scenarios [6]. Likewise, random outside disturbances hinder the acquisition of accurate process-noise statistics [7]. Moreover, maneuvering target tracking involves time-varying noise covariances, which conventional algorithms fail to adequately address [8]. Consequently, adaptive algorithms are required to handle the uncertainty associated with noise covariances [9].

Adaptive Kalman filters (AKF) represent a class of state estimation methods that operate without complete prior information. They mainly include Sage-Husa adaptive KF (SHAKF) [10,11], innovation-based adaptive KF (IAKF) [12,13], fading adaptive KF (FAKF) [14,15], and variational Bayesian adaptive KF (VBAKF) [16,17]. Relying on

the maximum a posteriori rule, SHAKF achieves optimal estimation for measurement-noise statistics, and a variant with gradual fading-memory property has also been developed [18]. Unfortunately, it is incapable of maintaining positive definiteness for the measurement noise covariance matrix, which may cause the filter to diverge. IAKF utilizes innovation sequences to address the low estimation accuracy caused by incorrect prior information [19]. Even so, it demands a comparatively long learning phase before generating reliable estimates for MNCM, which renders it poorly adapted to scenarios where MNCM varies in real-time. FAKF weakens the contribution of historical data via amplifying PECM [20]. Nevertheless, identical correction strength is imposed on every channel, which constrains the overall flexibility.

The VBAKF approach chooses suitable conjugate prior distributions and leverages variational Bayesian inference to achieve joint estimation of target states as well as noise covariance matrices [21,22]. To handle scenarios with slowly varying MNCM and PNCM, Huang et al. developed the novel VBAKF (NVBAKF) approach that yields favorable target-tracking performance while estimating target states simultaneously [23]. However, its time complexity remains high. Tan et al. introduced an ST-VBAKF algorithm that incorporates suboptimal fading factors and adaptively estimates MNCM, which greatly alleviates the adverse effects induced by time-varying PNCM [24]. To further enhance the tracking capability, Pan et al. developed a variational Bayesian adaptive filter designed for strong tracking, in which multiple fading factors are employed [25]. Although these improved VBAKFs achieve high-accuracy joint estimation of noise covariances and states, the accuracy improvement comes at the cost of fixed-point iterative computations.

To improve estimation accuracy while reducing time complexity, Huang et al. developed a novel sliding window variational AKF (SWVAKF) that jointly estimates PNCM, MNCM, and target states without iterative calculations [26]. To deal with the unknown MNCM that varies slowly over time, Xu et al. constructed an adaptive smoothing approach with fixed-lag based on VB inference [27]. Wang et al. proposed a variational Bayesian filtering solution for jump Markov systems with uncertain noise statistics. Their method is built upon an interacting multiple-model architecture and effectively handles maneuvering target tracking under both measurement and process noise uncertainty [28]. However, these algorithms still cannot avoid fixed-point iterations, nor can they guarantee the positive definiteness in the PECM. To overcome the above drawbacks, a robust adaptive sliding window variational Bayesian filter is proposed, which reduces computational complexity while preserving robustness.

The main contributions of this work are outlined below:

(1) We integrate multiple fading factors into the sliding-window variational Bayesian filter to calibrate the PECM, enabling channel-dependent adjustment capabilities to improve estimation performance.

(2) An improved sliding-window scheme is developed based on the SWVAKF, which eliminates fixed-point iterations. Under this framework, noise covariance matrices are assigned inverse-Wishart priors, and target states along with unknown noise covariances are jointly estimated via variational Bayesian inference. Prior states are further refined using subsequent-state information to enhance estimation accuracy and reduce computational complexity.

(3) Square-root decomposition techniques are

employed, which ensures that the covariance matrix stays positive definite, thereby preventing filter divergence and enhancing overall filter stability.

The structure of this paper is as follows. The square-root Kalman filter is introduced in Section 2. Section 3 is devoted to the proposed RASW-VBAKF, where the prior distributions for process and measurement noise are selected, multiple fading factors are constructed, and variational update equations are derived. In Section 4, simulation studies are conducted to validate the proposed method against existing filters. Concluding remarks and a summary of the proposed approach are given in Section 5.

2 System model and problem statement

This work focuses on a discrete linear system, modeled in state-space form as below.

$$x_k = F_{k-1}x_{k-1} + \omega_{k-1} \tag{1}$$

$$z_k = H_k x_k + \upsilon_k \tag{2}$$

where x_k is the state vector, z_k is the measurement vector, F_{k-1} is the state transition matrix, H_k is the measurement matrix, ω_{k-1} and υ_k are process noise and measurement noise with the covariance $Q_k = E(\omega_k \omega_k^T)$ and $R_k = E(\upsilon_k \upsilon_k^T)$, respectively. The initial state vector x_0 is assumed to follow Gaussian distribution with mean vector $\hat{x}_0$ and covariance matrix P_0. Furthermore, it is necessary to assume that x_0, ω_k, and υ_k are not correlated with each other at any time.

Kalman filter is susceptible to numerical errors during its operation, resulting in the propagation of non-positive definite covariance and divergence [29]. As a variant of the Kalman filter, the square-root Kalman filter iterates on the square root of the covariance matrix instead of the covariance itself, thereby mitigating divergence [30]. Such a square-root implementation improves numerical robustness, keeps the state covariance matrix positive definite, and yields notable gains in both estimation accuracy and stability [31].

$$\hat{x}_{k|k-1} = F_{k-1}\hat{x}_{k-1|k-1} \tag{3}$$

$$P_{k|k-1} = F_{k-1}P_{k-1|k-1}F_{k-1}^T + Q_{k-1} \tag{4}$$

$$P_{k-1|k-1} = S_{k-1|k-1}S_{k-1|k-1}^T \tag{5}$$

$$S_{k|k-1} = \begin{bmatrix} (F_{k|k-1}S_{k|k-1})^T \\ (Q_{k-1}^{0.5})^T \end{bmatrix}^T \tag{6}$$

$$A_k = S_{k|k-1}^T H_k^T \tag{7}$$

$$B_k = (A_k^T A_k + R_k)^{-1} \tag{8}$$

$$K_k = S_{k|k-1}A_k B_k \tag{9}$$

$$P_{k|k} = P_{k|k-1} - K_k H_k P_{k|k-1} \tag{10}$$

Under a linear Gaussian state-space model, the Kalman filter achieves the minimum mean square error (MMSE) provided that both Q_k and R_k are exactly known. In practice, imprecise Q_k and R_k can severely degrade state estimation, particularly in autonomous vehicle navigation and positioning systems [32-33].Since practical operating environments cannot be precisely characterized, pre-assumed noise statistics frequently deviate from those encountered in real-world scenarios. As shown in Eqs. (6) and (8), inaccuracies in PNCM and MNCM will deteriorate estimation accuracy and may even trigger filter divergence.

3 RASW-VBAKF approach

3.1 Choice of Prior Distributions and variational approximations

The variational Bayesian approach approximates the true posterior distribution using a factorized variational distribution, thereby constructing a variational lower bound associated with the log-marginal likelihood. As this bound is maximized repeatedly, the variational distribution converges progressively toward the true posterior [34]. The purpose of this paper is to simultaneously estimate the target state along with the unknown noise covariance matrices. This motivates the use of conjugate priors for the imprecise PNCM and MNCM, so that the posterior preserves the same functional form as the prior. Within the Bayesian framework, the inverse Wishart distribution is a natural conjugate prior for the covariance of a Gaussian distribution when the mean is known [35]. On this basis, the MNCM and PNCM priors can be expressed in terms of the inverse Wishart probability density function as given below [23]:

$$p(R_k \mid z_{1:k-1}) = IW(R_k; u_{k|k-1}, U_{k|k-1}) \tag{11}$$

$$p(Q_k \mid z_{1:k-1}) = IW(Q_k; t_{k|k-1}, T_{k|k-1}) \tag{12}$$

$$IW(R_k; u_{k|k-1}, U_{k|k-1}) = \frac{|U_{k|k-1}|^{0.5u_{k|k-1}} |R_k|^{-0.5(u_{k|k-1}+m+1)}}{2^{0.5mu_{k|k-1}}\Gamma(0.5u_{k|k-1})} * \exp(-0.5tr(U_{k|k-1}/R_k)) \tag{13}$$

$$IW(Q_k; t_{k|k-1}, T_{k|k-1}) = \frac{|T_{k|k-1}|^{0.5t_{k|k-1}} |Q_k|^{-0.5(t_{k|k-1}+n+1)} \exp(-0.5tr(T_{k|k-1}/Q_k))}{2^{0.5nt_{k|k-1}}\Gamma(0.5t_{k|k-1})} \tag{14}$$

At time step *k*, the joint prior corresponding to target states as well as noise covariance matrices is assumed to follow a factorized form composed of Gaussian and inverse-Wishart distributions:

$$p(x_{k-L}, R_k, Q_k \mid z_{1:k-1}) = q(x_{k-L})q(R_k)q(Q_k) \tag{15}$$

where $u_{k|k-1}$ denotes the degrees-of-freedom parameter and $U_{k|k-1}$ represents the scale matrix associated with the inverse-Wishart distribution of R_k, $t_{k|k-1}$ and $T_{k|k-1}$ are the corresponding parameters for Q_k; m is the dimension of R_k; $\Gamma(\cdot)$ denotes the multivariate gamma function; and $tr(\cdot)$ denotes the trace operator. For an inverse Wishart distribution, the expectation of its inverse is given by

$$E[R_k^{-1}] = (u_{k|k-1} - m - 1)/U_{k|k-1}$$

and similarly,

$$E[Q_k^{-1}] = (t_{k|k-1} - n - 1)/T_{k|k-1}.$$

In most practical cases, the noise covariance matrices do not change rapidly but evolve slowly over time. To capture this temporal variation, the prior PDF at the preceding time step is propagated by applying a forgetting factor $\rho \in (0,1]$. The degrees of freedom parameters are then updated as

$$\begin{cases} t_{k|k-1} = \rho(t_{k-1} - n - 1) + n + 1 \\ T_{k|k-1} = \rho T_{k-1} \end{cases} \tag{16}$$

$$\begin{cases} u_{k|k-1} = \rho(u_{k-1} - m - 1) + m + 1 \\ U_{k|k-1} = \rho U_{k-1} \end{cases} \tag{17}$$

Following the standard variational Bayesian (VB) framework, the true posterior $p(x_{k-L}, R_k, Q_k \mid z_{1:k})$ is replaced by an approximate distribution $q(x_{k-L})q(Q_k)q(R_k)$, and the Kullback–Leibler divergence (KLD) between these two distributions is computed iteratively so as to obtain the optimal estimate [36]. The KLD functional, denoted by KL(·), takes the form

$$KL[q(x_{L-k})q(R_k)q(Q_k) \| p(x_{L-k}, R_k, Q_k \mid z_{1:k}] = \arg\min \int q(x_{L-k})q(R_k)q(Q_k) \ln \frac{q(x_{L-k})q(R_k)q(Q_k)}{p(x_{L-k}, R_k, Q_k)} dx_{L-k} dR_k dQ_k \tag{18}$$

By variational derivation, it follows that the optimum of equation (18) must satisfy

$$\begin{cases} \ln q(\theta) = E_{\Xi^{(-\theta)}}\left[\log p(\Xi, z_{1:k})\right] + const \\ \Xi \square \{x_{k-L:k}, Q_k, R_k\} \end{cases} \tag{19}$$

where $E[\cdot]$ is the expected operation, $\log(\cdot)$ is the logarithmic function, $\Xi^{(-\theta)}$ represents all elements except for θ, and $const$ can be treated as a constant, which is not related to θ.

Owing to the interdependence of the variational parameters $x_{k-L:k}$, Q_k, and R_k, equation (19) cannot be solved directly. Accordingly, a mean-field approximation is invoked to decompose the joint posterior. More specifically, the expectations in equation (19) are evaluated from the posterior PDF at the (i+1)th iteration, and $q^{(i)}(\theta)$ is then updated to $q^{(i+1)}(\theta)$ using these expectations. However, this iterative scheme is computationally expensive. To reduce computational burden, the proposed scheme avoids fixed-point iterations and approximates the variational variables in the following form.

$$q(x_{k-L:k}) = N(x_{k-L:k} \mid \hat{x}_{k-L:k|k}, P^*_{k-L:k|k}) \tag{20}$$

$$q(Q_k) = IW(Q_k \mid t_{k|k}, T_{k|k}) \tag{21}$$

$$q(R_k) = IW(R_k \mid u_{k|k}, U_{k|k}) \tag{22}$$

The $q(R_k)$ distribution involves two parameters, i.e., the degree of freedom and the inverse scale matrix, which are specified as

$$\begin{cases} u_k = u_{k|k-1} + L \\ U_k = U_{k|k-1} + \sum_{i=k-L+1}^{(k)} \alpha_k^{(i)} \end{cases} \tag{23}$$

The $q(Q_k)$ distribution involves two parameters, i.e., the degree of freedom and the inverse scale matrix, which are specified as

$$\begin{cases} t_k = t_{k|k-1} + L \\ T_k = T_{k|k-1} + \sum_{i=k-L+1}^{(k)} \beta_k^{(i)} \end{cases} \tag{24}$$

The auxiliary matrix is represented as follows:

$$\alpha_k = E[(z_k - H_k x_k)(z_k - H_k x_k)^T] = (z_k - H_k \hat{x}_k)(z_k - H_k \hat{x}_k)^T + H_k P_{k|k} H_k^T \tag{25}$$

$$\beta_k = E[(x_k - \hat{x}_{k|k-1})(x_k - \hat{x}_{k|k-1})^T] = (\hat{x}_{k|k} - \hat{x}_{k|k-1})(\hat{x}_{k|k} - \hat{x}_{k|k-1})^T + P_{k|k} \tag{26}$$

The corrected PNCM and MNCM can be represented as follows

$$R_k = \{E[R_k^{-1}]\}^{-1} = (u_k - m - 1)^{-1} U_k \tag{27}$$

$$Q_k = \{E[Q_k^{-1}]\}^{-1} = (t_k - n - 1)^{-1} T_k \tag{28}$$

3.2 Multiple Fading Factors

When the PNCM is not constant but changes over time, the conventional Kalman filter exhibits slow convergence and significant estimation errors, as reflected in the residual sequence ε_k. To address this issue, the proposed approach corrects the PECM $P_{k|k-1}$ through multiple fading factors formulated on the basis of the strong tracking principle, i.e.,

$$P^*_{k|k-1} = \lambda_k F_{k-1} P_{k-1} F_{k-1}^T + Q_{k-1} \tag{29}$$

The diagonal entries of the PECM exert distinct influences on the estimation accuracy for each individual state component [37]. To achieve channel-dependent adjustment capability and further

improve state estimation performance, multiple fading factors are incorporated into the proposed approach, as specified below:

$$\lambda_k = dig(\lambda_k^1, \lambda_k^2, ..., \lambda_k^n) \tag{30}$$

$$\lambda_k^i = \mu_i c_k,\ i = 1, ..., n \tag{31}$$

where n, c_k, and μ_i respectively denoting the target state dimension, the fading factor at time k, and the weight coefficient of fading factors.

In accordance with the fading factor definition, c_k is given below:

$$c_k = \max(1, \frac{tr[N_k]}{\sum_{i=1}^{n} a_i M_k^{ii}}) \tag{32}$$

The covariance V_k of the output residual sequence is

$$V_k = \frac{(1-\xi)\varepsilon_k \varepsilon_k^T + (\xi - \xi^k) V_{k-1}}{1-\xi^k} \tag{33}$$

$$\varepsilon_k = z_k - \hat{z}_{k|k-1} \tag{34}$$

$$N_k = V_k - H_k Q_{k-1} H_k^T - \beta \hat{R}_k \tag{35}$$

$$M_k = H_k F_{k-1} P_{k-1|k-1} F_{k-1}^T H_k^T \tag{36}$$

Since the noise covariances vary slowly over time, estimates Q_{k-1} and R_{k-1} from the last time step still provide valuable information for the current iteration. By treating Q_k and R_k, as time-varying parameters, a refined PECM can be obtained by properly adjusting the fading factor λ_k. An accurate $P_{k|k-1}$ not only mitigates the adverse effects of incorrect PNCM and MNCM but also directly improves the state estimation accuracy.

3.3 Adaptive sliding-window variational Bayesian filter

The recursive procedure of RASW-VBAKF is described as follows.

Prediction:

$$x_{k|k-1} = F x_{k-1|k-1} \tag{37}$$

$$\varepsilon_k = z_k - H_k x_{k|k-1} \tag{38}$$

Calculate multiple fading factors

$$V_k = \frac{(1-\xi)\varepsilon_k \varepsilon_k^T + (\xi - \xi^k) V_{k-1}}{1-\xi^k} \tag{39}$$

$$R_k = \begin{cases} \dfrac{U_0}{u_0 - m - 1} & k = 1 \\ R_{k-1} & k > 1 \end{cases} \tag{40}$$

$$Q_k = \begin{cases} \dfrac{T_0}{t_0 - n - 1} & k = 1 \\ Q_{k-1} & k > 1 \end{cases} \tag{41}$$

$$N_k = V_k - H_k Q_{k-1} H_k^T - \beta R_k \tag{42}$$

$$M_k = H_k F_{k-1} P_{k-1|k-1} F_{k-1}^T H_k^T \tag{43}$$

$$\lambda_k = dig(\lambda_k^1, \lambda_k^2, ..., \lambda_k^n) \tag{44}$$

$$P_{k|k-1}^* = \lambda_k F_{k-1} P_{k-1} F_{k-1}^T + Q_k \tag{45}$$

Measurement update:

$$P_{k-1|k-1} = S_{k-1|k-1} S_{k-1|k-1}^T \tag{46}$$

$$S_{k|k-1} = \begin{bmatrix} (F_{k|k-1} S_{k|k-1})^T \\ (Q_{k-1}^{0.5})^T \end{bmatrix}^T \tag{47}$$

$$\varphi_k = S_{k|k-1}^T H_k^T \tag{48}$$

$$\gamma_k = [\varphi_k^T \varphi_k + R_k]^{-1} \tag{49}$$

$$K_k = S_{k|k-1} \varphi_k \gamma_k \tag{50}$$

$$x_{k|k} = x_{k|k-1} + K_k (z_k - H_k x_{k|k-1}) \tag{51}$$

$$P_{k|k} = (I - K_k H_k)\hat{P}_{k|k-1} \tag{52}$$

$$S_{k|k} = chol(P_{k|k}) \tag{53}$$

State Smoothing

$$G_{k|k} = P_{k-1|k-1} F_k^T \hat{P}_{k|k-1}^{-1} \tag{54}$$

$$P_{k-1,k} = G_{k-1} P_{k|k} \tag{55}$$

$$x_{k-1|k-1} = x_{k-1|k-1} + G_{k-1}(x_{k|k} - F_k x_{k-1|k-1}) \tag{56}$$

$$P_{k-1|k-1} = P_{k-1|k-1} + G_{k-1}(P_{k|k} - \hat{P}_{k|k-1})G_{k-1}^T \tag{57}$$

for $j = k:-1:k-L+1$

$$\alpha_j = (z_j - H_j \hat{x}_{j|j})(z_j - H_j \hat{x}_{j|j})^T + H_j P_{j|j} H_j^T \tag{58}$$

$$\begin{aligned}\beta_j &= P_{j|j} - F_j P_{k-1,k} - P_{k-1,k} F_j^T + F_j P_{j-1|j-1} F_j^T \\ &+ (\hat{x}_{j|j} - F_j \hat{x}_{j-1|j-1})(\hat{x}_{j|j} - F_j \hat{x}_{j-1|j-1})^T\end{aligned} \tag{59}$$

end for

Online estimation of R_k and Q_k

$$u_{k|k-1} = \rho(u_{k-1} - m - 1) + m + 1,$$

$$U_{k|k-1} = \rho U_{k-1} \tag{60}$$

$$U_{k|k} = U_{k|k-1} + \sum_{i=k-L+1}^{k} \alpha_i,$$

$$u_{k|k} = u_{k|k-1} + L \tag{61}$$

$$R_k = \{E[R_k^{-1}]\}^{-1} = (u_{k|k} - m - 1)^{-1} U_{k|k} \tag{62}$$

$$t_{k|k-1} = \rho(t_{k-1} - n - 1) + n + 1,$$

$$T_{k|k-1} = \rho T_{k-1} \tag{63}$$

$$T_{k|k} = T_{k|k-1} + \sum_{i=k-L+1}^{k} \beta_i,$$

$$t_{k|k} = t_{k|k-1} + L \tag{64}$$

$$Q_k = \{E[Q_k^{-1}]\}^{-1} = (t_{k|k} - n - 1)^{-1} T_{k|k} \tag{65}$$

4 Simulation

Based on the scenario in [28], a continuous white noise acceleration model governs the target motion in two-dimensional Cartesian coordinates, while sensor measurements capture its position. The state vector is given by $[x_k, y_k, \dot{x}_k, \dot{y}_k]$, which refer to the distance and velocity. The state transition matrix F_{k-1} and the observation matrix H_k are (66) and (67), respectively.

$$F_{k-1} = \begin{bmatrix} I_2 & \Delta t I_2 \\ 0 & I_2 \end{bmatrix} \tag{66}$$

$$H_k = \begin{bmatrix} I_2 & 0 \end{bmatrix} \tag{67}$$

where the parameter $\Delta = 1s$ denotes the sampling interval.

The real process and measurement noise covariance are modeled using a slowly time-varying model, i.e.,

$$\begin{cases} Q_k = (6.5 + 0.5\cos(\pi k / T))q \begin{bmatrix} \Delta t^3 I_2 / 3 & \Delta t^2 I_2 / 2 \\ \Delta t^2 I_2 / 2 & \Delta t I_2 \end{bmatrix} \\ R_k = (0.1 + 0.05\cos(\pi k / T))r \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix} \end{cases} \tag{68}$$

The simulation parameters are specified as follows. The simulation runs for $T = 2000s$, and the noise-related parameters are $q = 1m^2/s^3$ and $r = 100m^2$. Constant covariance matrices are adopted for both process and measurement noise, given by $Q = \tau I_4$ and $R_k = \sigma I_4$, where the symbols τ and σ refer to the prior confidence parameters. The approach parameters are set to $\tau = 1$, $\sigma = 100$, forgetting factor $\rho = 1 - \exp(-4)$, smoothing length $L = 3$, distribution parameters $t = 1$, $T = 300 I_2$, $u = 1$, $U = 300 I_2$, changing factor $\xi = 0.95$, and weakening factor $\beta = 0.8$.

The following filters are used for comparison:

KFTCM: KF based on the actual noise covariance matrices.

KFNCM: KF based on the actual noise covariance matrices.

R-KF: KF with variational recursive estimation of MNCM and state.

N-KF: KF with variational recursive estimation of PECM, MNCM, and state.

MST-AKF: KF combining strong tracking with VB to estimate PECM, MNCM, and state.

SWVAKF: Sliding window variational adaptive KF for noise covariance and state estimation.

STSQ-MSWVAF: sliding window variational adaptive square-root KF to estimate MNCM, PECM, and state.

For performance evaluation, two metrics are considered: the position root mean square error (RMSE) and the averaged RMSE (ARMSE). Their expressions are provided below.

$$\begin{cases} E_{RMSE,pos} = \sqrt{\dfrac{1}{M}\sum_{s=1}^{M}((x_k^s - \hat{x}_k^s)^2 + (y_k^s - \hat{y}_k^s)^2)} \\ E_{ARMSE,pos} = \sqrt{\dfrac{1}{MT}\sum_{k=1}^{T}\sum_{s=1}^{M}((x_k^s - \hat{x}_k^s)^2 + (y_k^s - \hat{y}_k^s)^2)} \end{cases} \tag{69}$$

where the symbols (x_k^s, y_k^s) and $(\hat{x}_k^s, \hat{y}_k^s)$ denote the true and estimated positions obtained at time k during the s-th Monte Carlo trial, and $M = 500$ stands for the total number of trials. Similar expressions hold for the velocity RMSE and ARMSE.

The accuracy of PECM and MNCM estimation is measured by two indicators, namely the square root of normalized Frobenius norm (SRNFN) and the averaged square root of normalized Frobenius norm (ASRNFN), which are formulated as follows:

$$\begin{cases} E_{SRNFN,pos} = \left(\dfrac{1}{n^2 M}\sum_{s=1}^{M} \| P_{k|k-1}^{*s} - P_{k|k-1}^{s} \|^2\right)^{0.25} \\ E_{ASRNFN,pos} = \left(\dfrac{1}{n^2 MT}\sum_{k=1}^{T}\sum_{s=1}^{M} \| P_{k|k-1}^{*s} - P_{k|k-1}^{s} \|^2\right)^{0.25} \end{cases} \tag{70}$$

where the symbols $P_{k|k-1}^{s}$ and $P_{k|k-1}^{*s}$ refer to the true and estimated PECM obtained at the s-th Monte Carlo run, with n being the state vector dimension. Analogous expressions hold for the SRNFN and ASRNFN of MNCM.

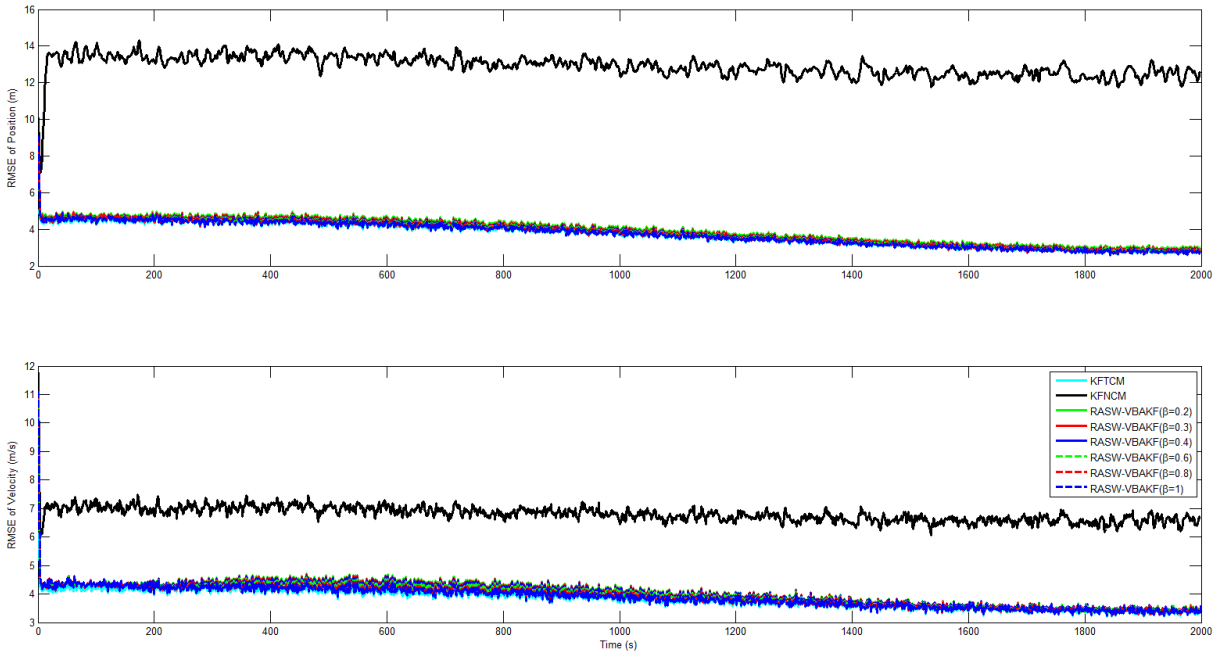


Figure 1. RMSE of distance and velocity with different β

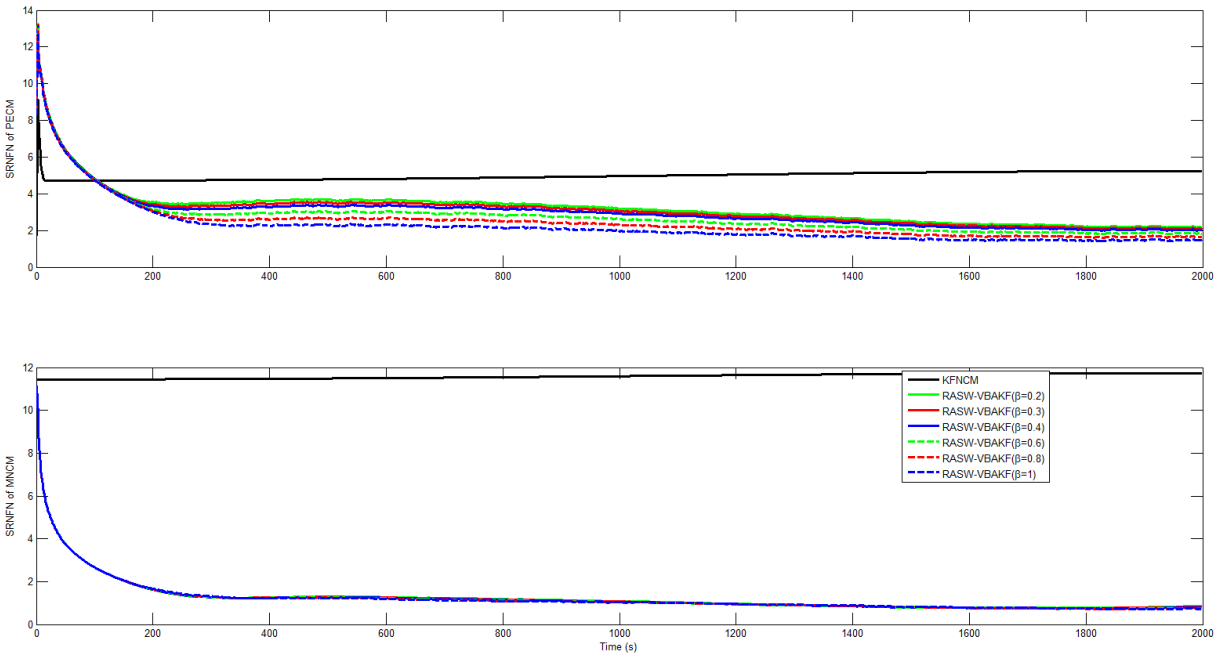


Figure 2. SRNFN of PECM and MNCM with different β

Figure 1 illustrates the position RMSE along with velocity obtained by KFTCM, KFNCM, and the RASW-VBAKF approaches at β=0.2, 0.3, 0.4, 0.6, 0.8, and 1.0. From Figure 1, it is evident that the RMSE of position obtained by RASW-VBAKF is significantly smaller than that of KFNCM. As β gradually decreases, the RMSE of RASW-VBAKF remains consistently higher in both position and velocity, converging to that

of KFTCM when β=1. Analogously to the position RMSE, the velocity RMSE is also smaller than that of KFNCM. Compared with velocity, the position is less prone to the effects of β because the corresponding fading factors carry more weight. Figure 2 plots the SRNFN of the PECM and MNCM, comparing KFNCM with RASW-VBAKF across different β values. From Figure 2, it can be seen that the SRNFN of PECM is inferior to that of KFNCM during the first 130 seconds. This is because inaccurate prior information is introduced into the PECM, causing the SRNFN to increase as β decreases.

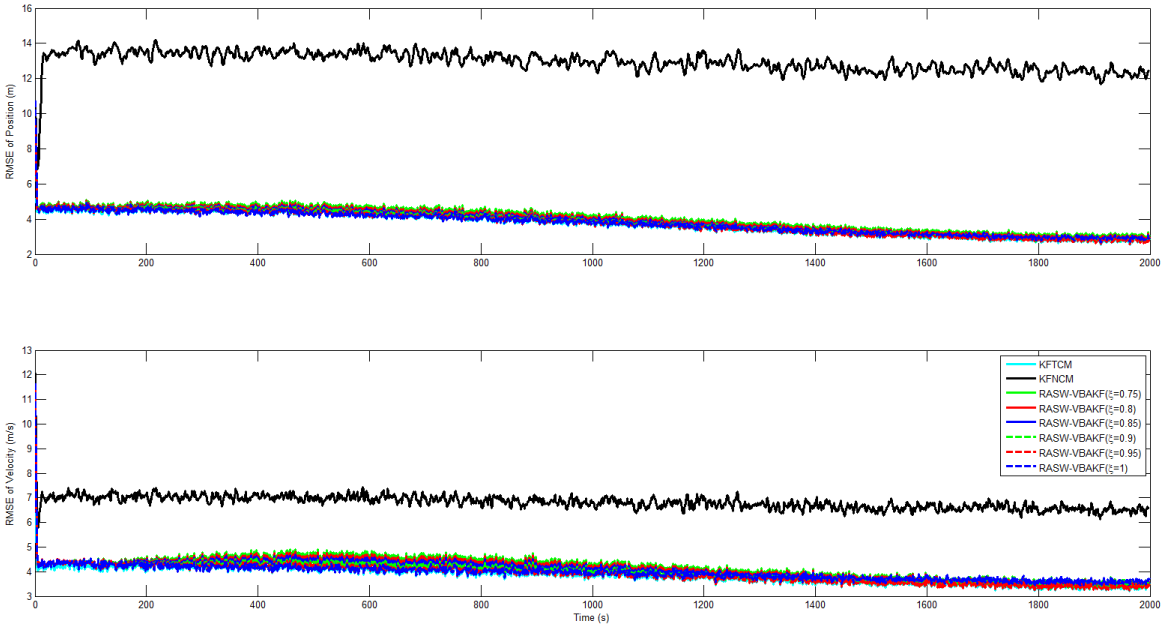


Figure 3. RMSE of distance and velocity with different ξ

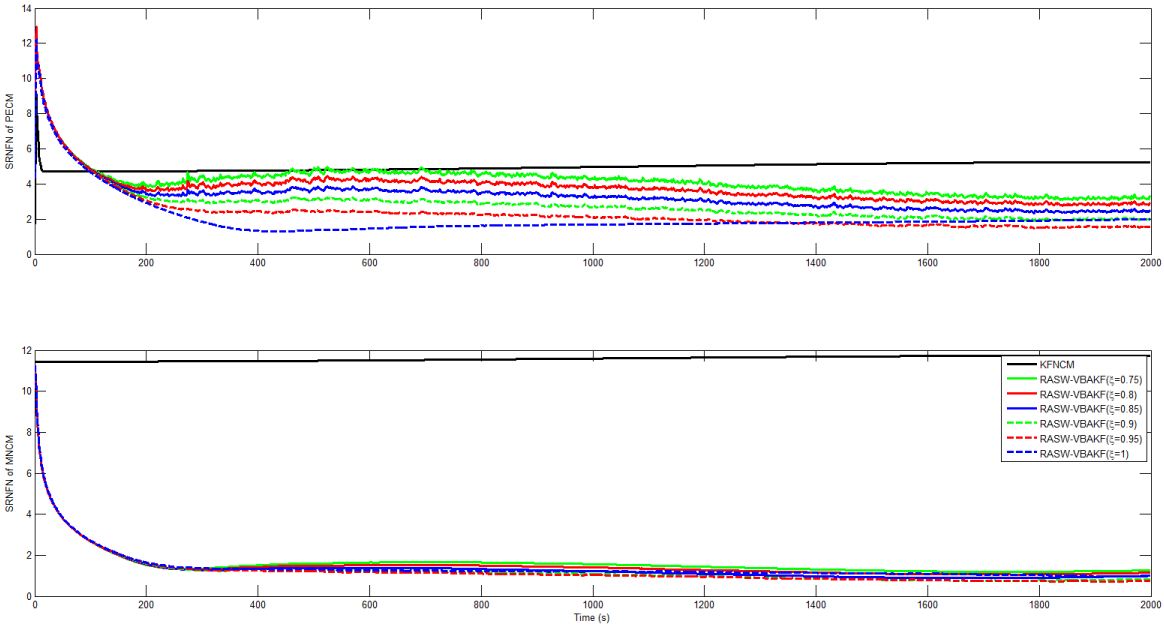


Figure 4. SRNFN of PECM and MNCM with different ξ

Figure 3 plots the position and velocity RMSE obtained by KFTCM, KFNCM, and RASW-VBAKF at ξ=0.75, 0.8, 0.85, 0.9, 0.95, and 1. From Figure 3, it can be seen that RASW-VBAKF exhibits higher estimation accuracy in terms of position and velocity compared to KFNCM. As ξ gradually increases, the RMSE of RASW-VBAKF initially exhibits a decreasing trend, reaching its minimum at $\xi = 0.95$. However, beyond this point (specifically at $\xi = 1$), the RMSE begins to increase and diverge significantly after $t = 1200$. This implies that an overly large ξ can introduce excessive reliance on historical residual data, causing the filter to lose stability over time. Therefore, $\xi = 0.95$ serves as the optimal trade-off point for state estimation accuracy and long term stability. To further analyze the impact of the varying factor ξ on the performance of target state estimation, Figure 4 illustrates the SRNFN of PECM and MNCM obtained by KFNCM and RASW-VBAKF at different ξ values. It is clearly observed that RASW-VBAKF exhibits a significant advantage in regulating the SRNFN of PECM compared to KFNCM. The SRNFN of RASW-VBAKF gradually decreases with an increase in ξ during the steady state phase; however, the curve for $\xi = 1$ shows a noticeable upward divergence in the later stage ($t > 1200$), consistent with the observed degradation in RMSE. This phenomenon confirms that while a larger ξ improves the utilization of historical residual data and reduces initial uncertainty, setting it too high ultimately compromises the filter's robustness and leads to long term divergence.

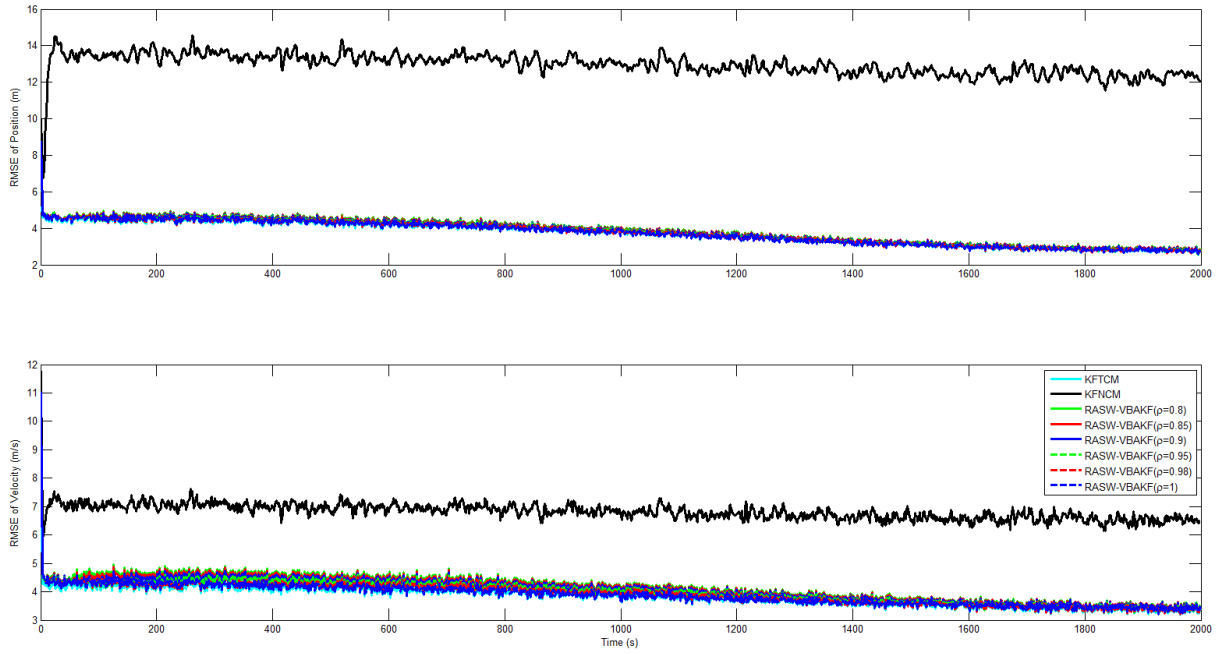


Figure 5. RMSE of position and velocity with different ρ

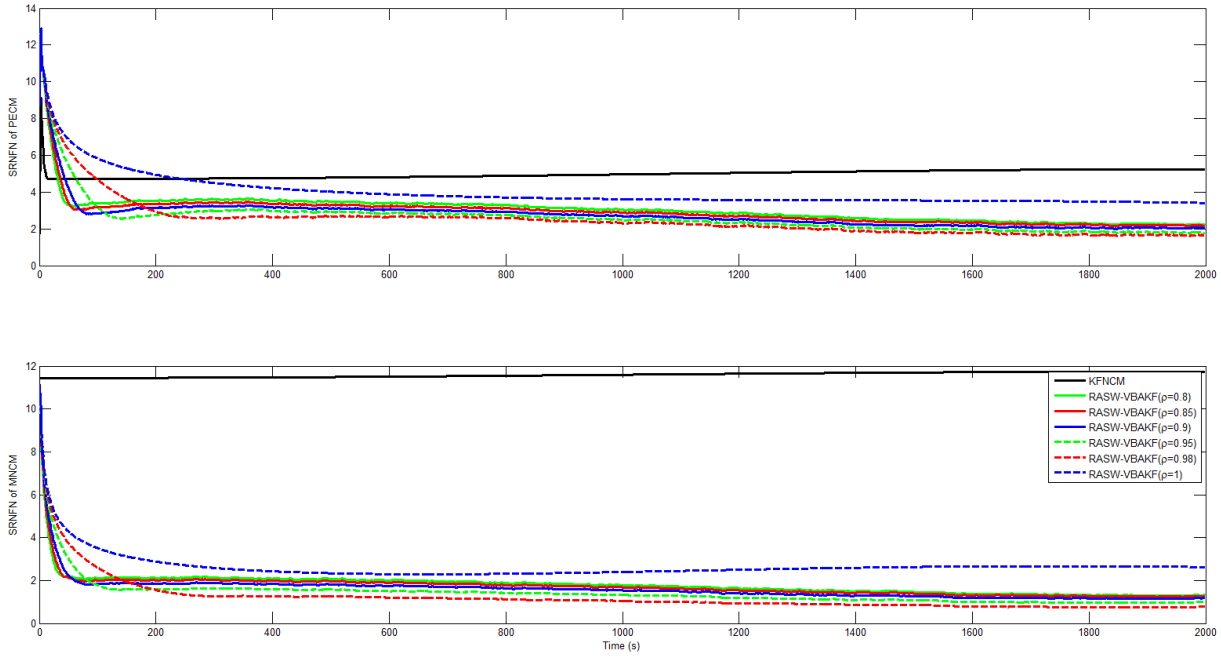


Figure 6. SRNFN of PECM and MNCM with different ρ

Figures 5-6 show the RMSE of the position and velocity obtained by KFNCM and the proposed RASW-VBAKF approach at $\rho = 0.8, 0.85, 0.9, 0.95, 0.98, 1$. From Figure 5, it can be seen that the proposed RASW-VBAKF exhibits better estimation accuracy than KFNCM across all settings. Furthermore, the estimation accuracy of the proposed RASW-VBAKF generally improves as ρ increases, primarily because a larger ρ enhances the utilization of historical residual data. However, a notable exception occurs at $\rho = 1$: the estimation accuracy is lower than that at intermediate values $\rho = 0.8, 0.85, 0.9, 0.95, 0.98$. This is mainly because the noise covariance matrices are regarded as constant when $\rho = 1$. As a result, the adaptability of filter to changing system conditions is weakened, leading to degraded estimation performance.

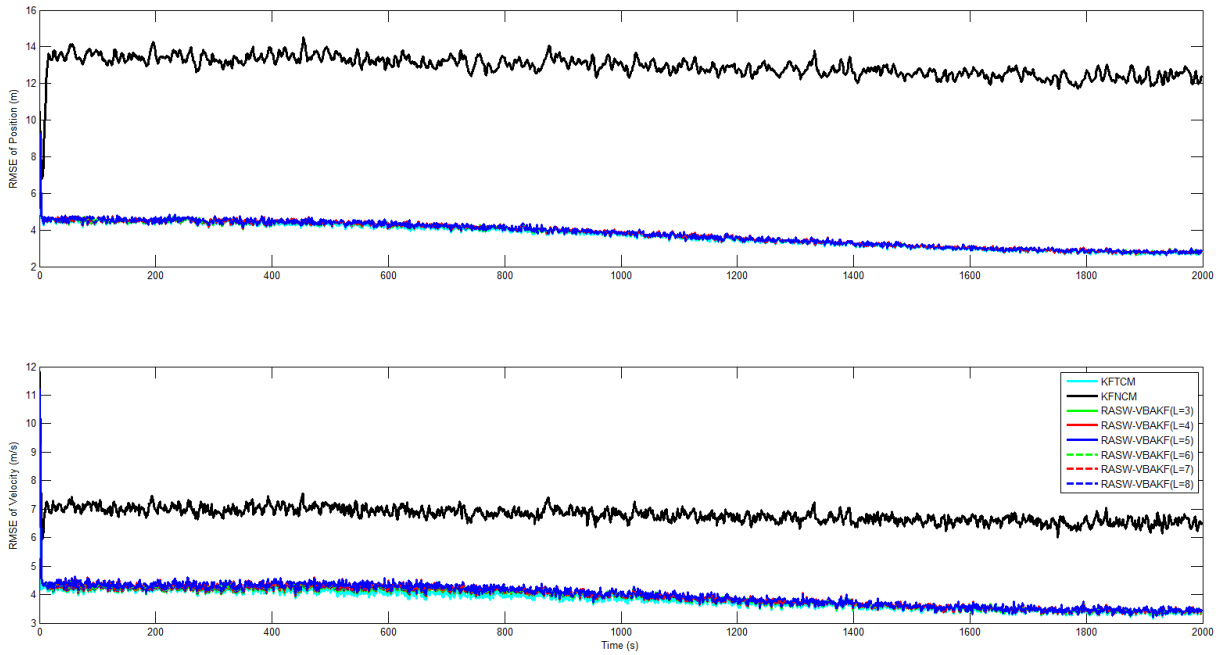


Figure 7. RMSE of position and velocity with different L

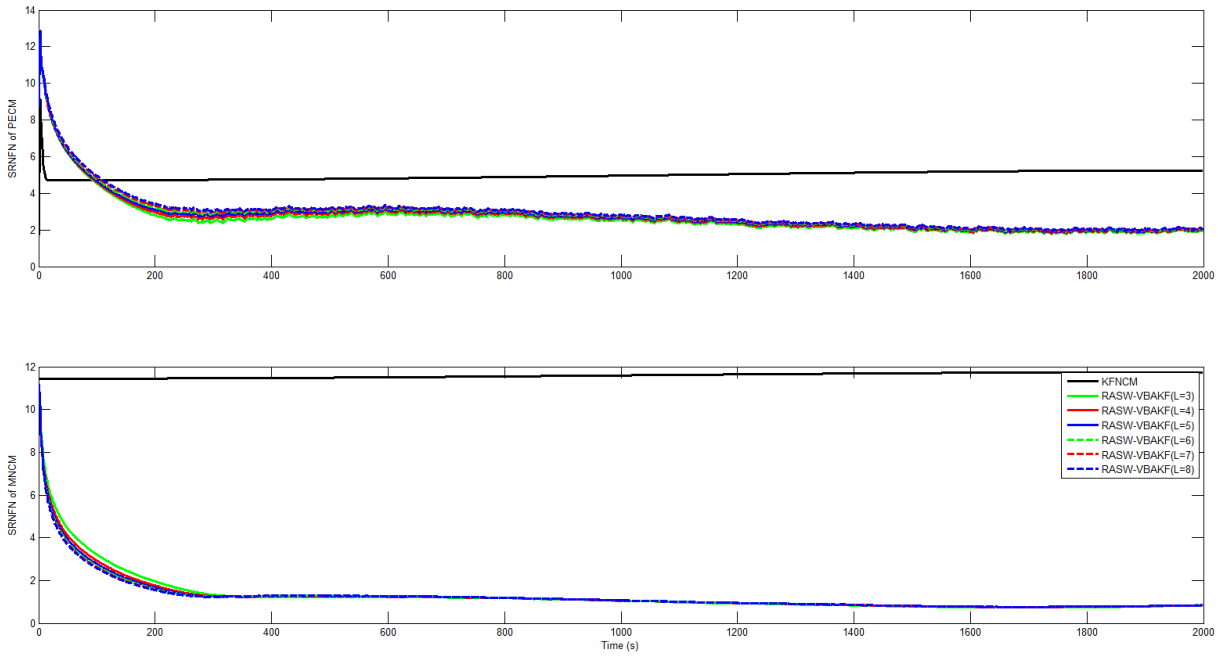


Figure 8. SRNFN of PECM and MNCM with different L

Figures 7 and 8 show the RMSE of position and velocity, as well as the SRNFN of PECM and MNCM, for the existing KFNCM and the proposed RASW-VBAKF approaches when $L = 3,4,5,6,7,8$, respectively. From Figures 7 and 8, it can be seen that the RASW-VBAKF exhibits significantly smaller RMSEs and SRNFNs than the KFNCM, and reaches steady-state convergence rapidly (within approximately 200 seconds). Therefore, the RASW-VBAKF demonstrates satisfactory convergence speeds with

respect to different sliding window sizes. Furthermore, the proposed algorithm exhibits highly consistent and robust estimation performance across all tested window lengths ($L = 3$ to 8), and consistently achieves higher estimation accuracy than the KFNCM baseline.

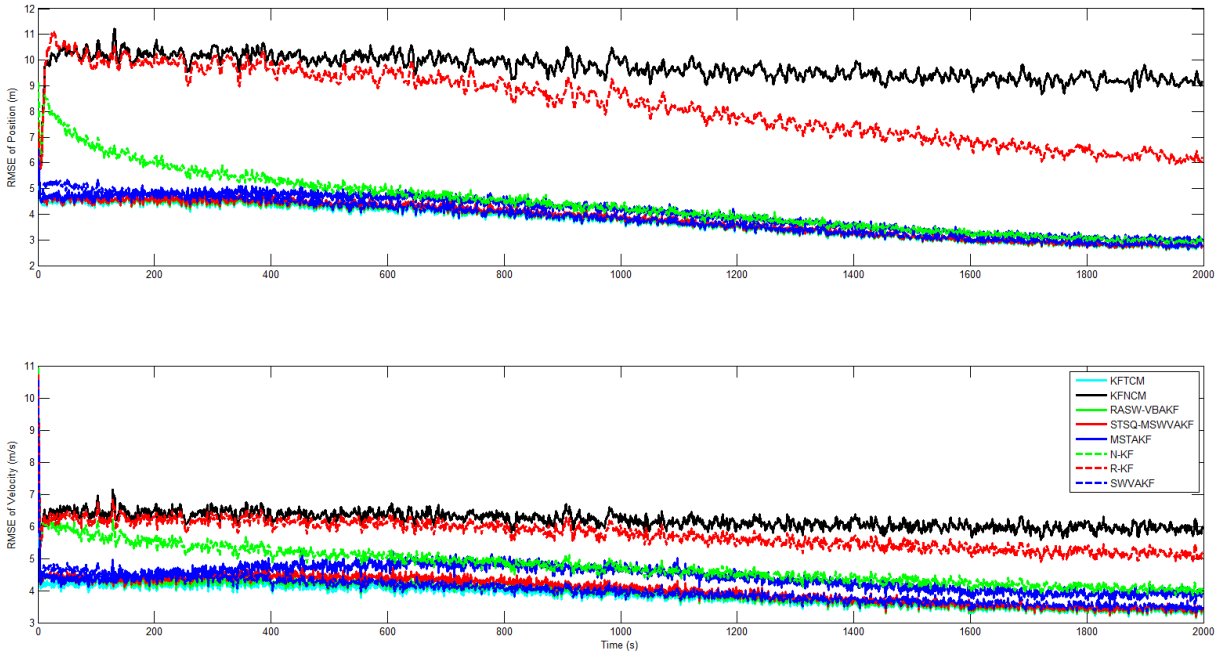


Figure 9. RMSE of distance and velocity with different filters

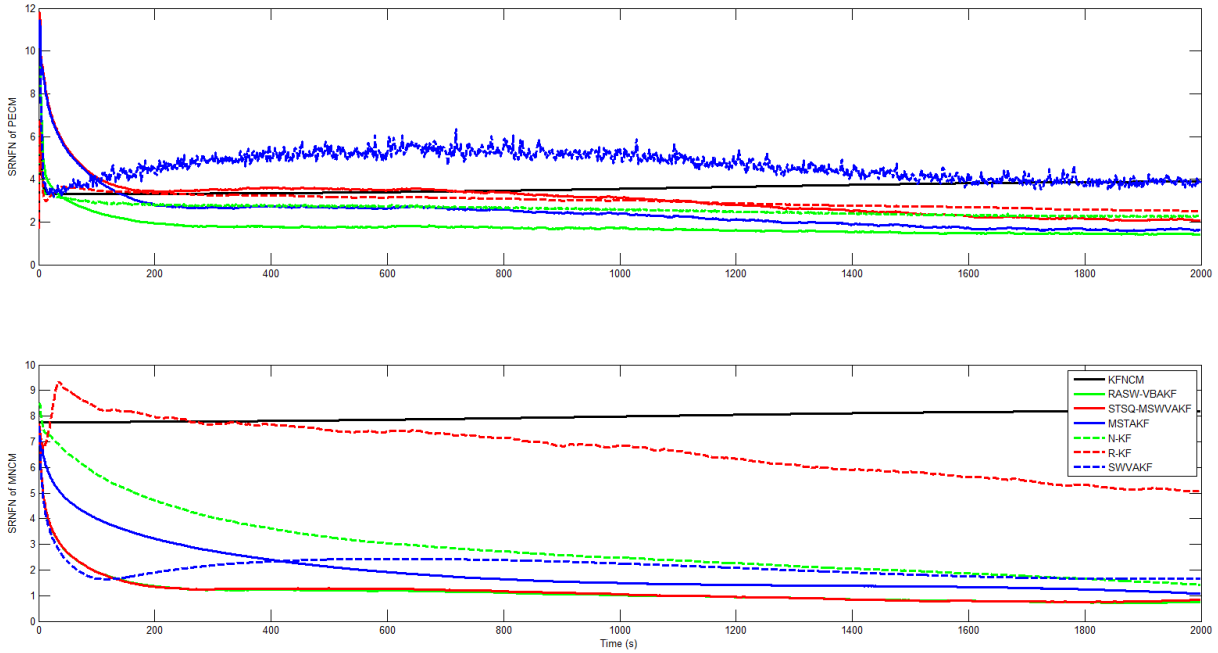


Figure 10. SRNFN of MNCM and PECM with different filters

Figure 9 shows the RMSE of position and velocity obtained from different filters. Because the proposed RASW-VBAKF utilizes multiple fading factors to correct different channels of the PECM and estimates the noise covariance matrices using the Variational Bayesian approach, its estimation of position and velocity is much closer to the ideal KFTCM benchmark compared to other existing filters. Table 1 lists the ARMSE of these filters. From Table 1, it can be quantitatively confirmed that RASW-VBAKF achieves the highest state estimation accuracy among all the evaluated methods. Specifically, while its ARMSE is slightly higher than the ideal KFTCM (3.75 vs 3.70 for position), it yields the lowest error among all practical filters, closely followed by the STSQ-MSWVAKF.

Figure 10 illustrates the SRNFN of PECM and MNCM obtained from different filters, with the corresponding ASRNFN listed in Table 2. It can be clearly observed that KFNCM exhibits poor accuracy due to its lack of real-time estimation and update capabilities for the noise covariance matrices and PECM. Additionally, the R-KF exhibits a relatively slow convergence speed; however, its SRNFN of MNCM remains significantly higher than that of KFNCM throughout the simulation (6.69 vs 7.98 in Table 2), indicating inferior long-term model uncertainty regulation. The MSTAKF achieves higher estimation accuracy than N-KF because it leverages the VB approach to estimate the measurement noise covariance as a time-varying parameter, which is updated with multiple fading factors, steering the state prediction covariance in a more accurate direction. Finally, the primary reason the proposed RASW-VBAKF outperforms STSQ-MSWVAKF and SWVAKF is that it not only uses the VB approach to adaptively estimate the noise covariance matrices but also employs multiple fading factors to individually correct each channel of the PECM. This synergistic mechanism results in the lowest ASRNFN values for both PECM (1.75) and MNCM (1.12), providing superior robustness and precision in state estimation.

Table 1. The ARMSE of position and velocity for different filters

Filters	ARMSE Position	ARMSE Velocity
KFTCM	3.70	3.83
KFNCM	9.72	6.19
STSQ-MSWVAF	3.78	4.02
MST-AKF	3.82	3.99
N-KF	4.39	4.76
R-KF	8.27	5.74
SWVAKF	4.05	4.42
RASW-VBAKF	3.75	3.89

Table 2. The ASRNFN of PECM and MNCM for different filters

Filters	ASRNFN PECM	ASRNFN MNCM
KFTCM	/	/
KFNCM	3.58	7.98
STSQ-MSWVAF	3.08	1.15
MST-AKF	2.44	1.91
N-KF	2.58	2.86
R-KF	2.98	6.69
SWVAKF	4.62	2.09
RASW-VBAKF	1.75	1.12

5 Conclusion

This paper proposes a novel and robust adaptive sliding window variational Bayesian filter (RASW-VBAKF). To ensure the positivity of the error covariance matrix and enhance overall robustness, a square root decomposition technique is first performed. Then, the Variational Bayesian (VB) approach is adopted to jointly estimate the noise covariance matrices. By substituting these estimated matrices into the calculation of multiple fading factors, the proposed method can adaptively correct different channels of the error covariance matrix, thereby significantly improving the accuracy of target

state estimation. Furthermore, by avoiding fixed point iteration, the RASW-VBAKF effectively reduces computational complexity while maintaining high estimation precision. Extensive simulation results demonstrate that, compared to existing filters, the proposed filter achieves superior state estimation accuracy and robustness, validating its effectiveness in practical applications.

Funding This work is partially supported by the National Natural Science Foundation of China (Grant No. 61673262), National GF Basic Research Program under JCKY2021110B134, the Fundamental Research Funds for the Central Universities, Jiangsu Province Industry-University-Research Cooperation Projects under BY20250395, and Key Research Project of Jiangsu Vocational College of Information Technology under JSITKY202401.

Data availability Not applicable.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Blair WD, Bar-Shalom Y (2023) MSE Design of Nearly Constant Velocity Kalman Filters for Tracking Targets with Deterministic Maneuvers. IEEE Trans Aerosp Electron Syst 59:4180-4191

2. Montañez OJ, Suarez MJ, Fernandez EA (2023) Application of Data Sensor Fusion Using Extended Kalman Filter Algorithm for Identification and Tracking of Moving Targets from LiDAR-Radar Data. Remote Sens 15:3396

3. Liu Y, Ning X, Li J, Wang B, Ma X (2021) Adaptive Central Difference Kalman Filter with Unknown Measurement Noise Covariance and Its Application to Airborne POS. IEEE Sens J 21:9927-9936

4. Nosratollahi M, Delalat M (2023) Comparison of Dynamic Models for Aerial Target Tracking Maneuvers Based on Stability and Measurement Loss. Iran J Sci Technol Trans Mech Eng 47:541-556

5. Kazimierski W, Zaniewicz G (2021) Determination of Process Noise for Underwater Target Tracking with Forward Looking Sonar. Remote Sens 13:1014

6. Mohd-Yasin F, Nagel DJ (2021) Noise as Diagnostic Tool for Quality and Reliability of MEMS. Sensors 21:1510

7. Chang G, Chen C, Zhang Q, Zhang S (2021) Variational Bayesian adaptation of process noise covariance matrix in Kalman filtering. J Franklin Inst 358:3980-3993

8. Chen W, Jiang S, Cao J, Sun R (2024) An adaptive EKF-SLAM algorithm for cooperative navigation of multi-aircrafts. IET Control Theory Appl 18(18):2823-2829

9. Yang XS, Li WY, Zhang WA (2025) Integer-constrained Progressive Gaussian Filtering Method for RTK Positioning. Acta Autom Sin 51(2):366-375

10. Gao S, Wei W, Zhong Y, Subic A (2015) Sage windowing and random weighting adaptive filtering method for kinematic model error. IEEE Trans Aerosp Electron Syst 51:1488-1500

11. Gao X, You D, Katayama S (2012) Seam Tracking Monitoring Based on Adaptive Kalman Filter Embedded Elman Neural Network During High-Power Fiber Laser Welding. IEEE Trans Ind Electron 59:4315-4325

12. Wang Y, Sun Y, Wei Z, Sun G (2018) Parameters estimation of electromechanical oscillation with incomplete measurement information. IEEE Trans Power Syst 33:5016-5028

13. Mohamed AH, Schwarz KP (1999) Adaptive Kalman filtering for INS/GPS. J Geod 73:193-203

14. Kiani M, Ahmadvand R (2022) The strong tracking innovation filter. IEEE Trans Aerosp Electron Syst 58:3261-3270

15. Liu X, Guan J, Jiang R (2022) Robust strong tracking unscented Kalman filter for non-linear systems with unknown inputs. IET Signal

Process 16:351-365
16. Ardeshiri T, Özkan E, Orguner U, Gustafsson F (2015) Approximate Bayesian Smoothing with Unknown Process and Measurement Noise Covariances. IEEE Signal Process Lett 22:2450-2454
17. Agamennoni G, Nieto JI, Nebot EM (2012) Approximate inference in state-space models with heavy-tailed noise. IEEE Trans Signal Process 60:5024-5037
18. Yan W, Ding Q, Chen J, Liu Y, Cheng SS (2021) Needle tip tracking in 2D ultrasound based on improved compressive tracking and adaptive Kalman filter. IEEE Rob Autom Lett 6:3224-3231
19. Yin X, Chai H, Xiang M, Du Z, Tian X (2021) A GEO-pivoted adaptive extended Kalman filtering method in low-latitude dense building environment. Measurement 173:108657
20. Hou B, Wang J, He Z, Qin Y, Zhou H (2020) Novel interacting multiple model filter for uncertain target tracking systems based on weighted Kullback-Leibler divergence. J Franklin Inst 357:13041-13084
21. Lan H, Zhao S, Hu J, Wang Z, Fu J (2025) Joint state estimation and noise identification based on variational optimization. IEEE Trans Autom Control 70(7):4500-4515
22. Sarkka S, Nummenmaa A (2009) Recursive noise adaptive Kalman filtering by variational Bayesian approximations. IEEE Trans Autom Control 54:596-600
23. Huang Y, Zhang Y, Wu Z, Li N, Chambers JA (2018) A novel adaptive Kalman filter with inaccurate process and measurement noise covariance matrices. IEEE Trans Autom Control 63:594-601
24. Tan F, Zhao J (2020) Strong tracking based variational Bayesian adaptive Kalman filtering algorithm. Electron Opt Control 27(12):12-16+36
25. Pan C, Gao J, Li Z, Qian N, Li F (2021) Multiple fading factors-based strong tracking variational bayesian adaptive Kalman filter. Measurement 176:109139
26. Huang Y, Zhu F, Jia G, Zhang Y (2020) A slide window variational adaptive Kalman filter. IEEE Trans Circuits Syst II Express Briefs 67:3552-3556
27. Xu H, Duan K, Yuan H, Xi W, Wang Y (2021) Adaptive fixed-lag smoothing algorithms based on the variational Bayesian method. IEEE Trans Autom Control 66:4881-4887
28. Wang G, Wang X, Zhang YG (2020) Variational Bayesian IMM filter for JMSs with unknown noise covariances. IEEE Trans Aerosp Electron Syst 56:1652-1661
29. Wang Z, Hou X, Dan Z, Fang K (2023) Stochastic modeling with robust Kalman filter for real-time kinematic GPS single-frequency positioning. GPS Solut 27:153
30. Closas P, Fernandez-Prades C, Vila-Valls J (2012) Multiple Quadrature Kalman Filtering. IEEE Trans Signal Process 60:6125-6137
31. Jiang C, Wang S, Wu B, Fernandez C (2021) A state-of-charge estimation method of the power lithium-ion battery in complex conditions based on adaptive square root extended Kalman filter. Energy 219:119603
32. Mu X, He B, Wu S, Zhang X, Song Y, Yan T (2021) A practical INS/GPS/DVL/PS integrated navigation algorithm and its application on Autonomous Underwater Vehicle. Appl Ocean Res 106:102441
33. Chen P, Li P, He B (2025) Integrated underwater navigation algorithm for AUV based on improved Sage-Husa and iterated maximum mixture correntropy unscented Kalman filter. Ocean Eng 317:122610
34. Zhang Y, Li X (2026) Variational Bayesian Joint Measurement Anomaly and Noise Inference Filter for Robust MEMS-INS Localization. IEEE Trans Instrum Meas 75:8509616
35. Yu K, Li H, Jiang L (2026) Variational Bayesian Kalman Filter Based on MPR Parameterization and Mode Switching Criteria. IEEE Trans Instrum Meas 75:6512117

36. Huang Y, Zhang Y, Chambers JA (2019) A novel Kullback-Leibler divergence minimization-based adaptive student's t-filter. IEEE Trans Signal Process 67:5417-5432
37. Wang J, Chen X, Yang P (2021) Adaptive H-infinite kalman filter based on multiple fading factors and its application in unmanned underwater vehicle. ISA Trans 108:295-304