Hindawi
International Journal of Aerospace Engineering
Volume 2024, Article ID 5599144, 15 pages
https://doi.org/10.1155/2024/5599144

Research Article

VB-Based Gaussian Sum Cubature Kalman Filter for Adaptive Estimation of Unknown Delay and Loss Probability

Ruipeng Wang,[1] **Xiaogang Wang**,[1] **and Haojie Zhang**[2]

[1]*School of Astronautics, Harbin Institute of Technology, Harbin, China*
[2]*Beijing Institute of Electronic System Engineering, Beijing, China*

Correspondence should be addressed to Xiaogang Wang; wangxiaogang@hit.edu.cn

Received 19 July 2023; Revised 17 December 2023; Accepted 26 December 2023; Published 25 January 2024

Academic Editor: Jinchao Chen

The traditional Kalman filter assumes that all measurements can be obtained in real time, which is invalid in practical engineering. Therefore, a variational Bayesian- (VB-) based Gaussian sum cubature Kalman filter is proposed to solve the nonlinear tracking problem of multistep random measurement delay and loss (MRMDL) with unknown probability. First, the measurement model with MRMDL is modified by Bernoulli random variables. Then, the expression of the likelihood function is reformulated as a mixture of multiple Gaussian distributions, and the cubature rule is used to improve the estimation accuracy under the framework of Gaussian sum filter in the process of time update. Finally, by constructing a hierarchical Gaussian model, the unknown and time-varying measurement delay and loss probability are estimated in real time with the state jointly using the VB method in the measurement update stage. The algorithm does not need to calculate the equivalent noise covariance matrix so as to avoid the possible division by zero operation, which improves the stability of the algorithm. Simulation results for a target tracking problem show that the proposed algorithm has a better performance in the presence of MRMDL and can estimate the unknown measurement delay and loss probability accurately.

1. Introduction

Kalman filter (KF) has been widely used in target tracking and integrated navigation system. As a real-time recursive state estimator, KF provides the analytical solution for minimum mean square error estimation of linear systems under Gaussian noise assumption [1, 2]. KF assumes that all measurements are obtained at the current time. However, when the signal transmission network is blocked or affected by the complex environment, the measurement data may be randomly delayed and lost [3–5], affecting the estimation accuracy of KF.

To deal with this situation, many improved KFs for measurement delay and loss are proposed. In [6], the extended KF (EKF) and unscented KF (UKF) have been generalized to one-step randomly delayed system by rewriting the measurement equation into a Gaussian mixture of the updated measurement and one-step delayed measurement. In [7], the augmented state KF (AS-KF) and augmented state probabilistic data association filter (AS-PDA) are developed to scenarios devoid of clutter or involving clutter with multistep random measurement delay, in which multistep random delay is processed by state augmentation. In [8], modified adaptive EKF is proposed by reorganizing the state update equation and combining the adaptive estimation algorithm, which avoids the augmentation of state variable or measurement equation. In [9, 10], the Gaussian approximate filters and smoothers are proposed for systems with one-step randomly delayed measurements. In [11], a modified likelihood cubature Kalman filter (MLCKF) is proposed to modify the likelihood function by marginalizing the delay variables, and the modified likelihood function is in the form of Gaussian mixture. When the measurement loss is known, the problem can be solved by using the intermittent KF (IKF) [12, 13]. In IKF, measurement loss is treated as a case where the measurement noise variance is infinite. The upper and lower bounds of the estimated error covariance are given in [12], and the stability is analyzed in [13]. In [14], particle filter is applied to the system with one-step random delay and measurement loss by modifying the expression of particle weights. Then, particle filter is used for multistep delay system [15]. In [16], an optimal linear filter is derived based on the known probability under the linear minimum variance criterion. However, in the above methods, the delay and loss probability is known and constant. In practice, the probability of measurement delay and loss is often unknown and time varying. Therefore, the estimation accuracy of the above algorithms suffers severely when inaccurate prior probability is used.

In order to deal with the problem of unknown probability, some methods have been proposed to identify the probability of delay and loss. In [17, 18], the expectation maximization (EM) approach is used to obtain the estimation of delay probability. However, because they are based on the smoothing framework, they cannot identify the time-varying probability accurately. In [19], a risk-sensitive KF (RSKF) is proposed by minimizing the expectation of the accumulated exponential quadratic error for system with one-step randomly delayed measurement with unknown probability and uncertain model. In recent years, the use of variational Bayesian (VB) method to identify unknown parameters has been widely studied [20–23], such as the identification of Student's t noise parameters [24–26]. In [27], an improved KF with one-step randomly delayed measurement is proposed. And the improved UKF (IUKF) version is presented in [28]. In a similar way, a VB-based adaptive KF with lost measurement (VBAKFLM) is proposed to achieve the estimation of a linear system with unknown and time-varying probability of measurement loss [29]. These filters use a Bernoulli variable to indicate whether the measurement is delayed or lost, and the likelihood function is transformed into a hierarchical Gaussian state model. Then, the VB method is used to estimate the probability. However, these methods cannot simultaneously estimate the state of nonlinear system with unknown and time-varying multistep delay and loss probability of measurements. In addition, there is a risk of division zero when obtaining the equivalent augmented measurement noise covariance matrix, which may cause the filter to crash.

In this article, we proposed a novel VB-based Gaussian sum cubature Kalman filter (CKF) for nonlinear system to estimate the state and identify the unknown delay and loss probability. By introducing two Bernoulli random variables, the measurement model with MRMDL is modified and the expression of the likelihood function is reformulated as a Gaussian mixture distribution. Therefore, the state estimation problem with MRMDL can be solved in the framework of Gaussian sum filter (GSF) while the cubature rule is used to improve the estimation accuracy. By constructing a hierarchical Gaussian model, the nonconjugated Gaussian mixture model is modified to conjugate exponential multiplication form by the introduced Bernoulli random variables. Then, the unknown and time-varying measurement delay and loss probability are estimated in real time with the state jointly using the VB method. Different from other VB-based filters, the proposed algorithm uses the estimated probability to update the weights without augmenting the measurement vector, which avoids the division by zero operation that may be involved in obtaining the equivalent noise covariance matrix, which improves the stability of the algorithm.

The remainder of this paper is listed as follows. In Section 2, the problem formulation is given. In Section 3, a hierarchical Gaussian state-space model for nonlinear system with MRMDL is constructed, and the VB method for joint posterior probability distribution function (PDF) is provided. In Section 4, simulation results and comparisons are shown. In Section 5, conclusions are given.

2. Problem Formulation

Consider the following discrete-time nonlinear dynamical system with MRMDL:

$$\begin{aligned} \boldsymbol{x}_k &= \boldsymbol{f}(\boldsymbol{x}_{k-1}) + \boldsymbol{w}_{k-1}, \\ \boldsymbol{z}_k &= \rho_k \boldsymbol{h}(\boldsymbol{x}_k) + \boldsymbol{v}_k, \\ \boldsymbol{y}_k &= \sum_{i=0}^{I} \lambda_k^i \boldsymbol{z}_{k-i}, \end{aligned} \tag{1}$$

where $\boldsymbol{x}_k \in \mathbb{R}^n$ is the n-dimensional state vector; $\boldsymbol{z}_k \in \mathbb{R}^m$ is the m-dimensional measurement vector at time index k; $\boldsymbol{f}(\cdot)$ and $\boldsymbol{h}(\cdot)$ are the known nonlinear state and measurement equation functions, respectively; and $\boldsymbol{w}_k \in \mathbb{R}^n$ and $\boldsymbol{v}_k \in \mathbb{R}^m$ are the zero mean process noise and measurement noise with known covariances $\boldsymbol{Q}_k$ and $\boldsymbol{R}_k$, respectively. ρ_k is a random binary variable which has a Bernoulli distribution, and $\rho_k = 0$ indicate that the measurement is lost, while $\rho_k = 1$ represents that a measurement is obtained by the sensor. Because of the unreliability of the communication link, the measurement received by the filter, $\boldsymbol{y}_k$, is different from the sensor's output $\boldsymbol{z}_k$. I is the maximum number of delay, and λ_k^i is an unknown binary variable in the set $\boldsymbol{\lambda}_k$ of which only one element is equal to 1 at time index k. In this paper, $\rho_k = 1$ and $\lambda_k^0 = 1$ imply that the current measurement is received at time k, while $\rho_k = 1$ and $\lambda_k^i = 1$indicate that the measurement with i-step delay is received. The probability of measurement loss can be expressed as $p(\rho_k = 0) = \gamma_k$, and the probability of receiving the measurement z_{k-i} is represented by $p(\boldsymbol{y}_k = \boldsymbol{z}_{k-i} | \rho_k = 1) = \boldsymbol{\mu}_k^i$; then, we have

$$p(\rho_k | \gamma_k) = {\gamma_k}^{(1-\rho_k)} (1 - \gamma_k)^{\rho_k}, \tag{2}$$

$$p(\boldsymbol{\lambda}_k | \boldsymbol{\mu}_k) = \prod_{i=0}^{I} \left(\mu_k^i\right)^{\lambda_k^i}. \tag{3}$$

3. VB-Based Gaussian Sum Cubature Kalman Filter

3.1. Modification of the Likelihood Function. Since the measurement obtained by the filter comes from $\boldsymbol{y}_k$ rather than $\boldsymbol{z}_k$, through Bayesian rules, we have

$$p(\boldsymbol{x}_{1:k}|\boldsymbol{y}_{1:k}) = \frac{p(\boldsymbol{y}_k|\boldsymbol{x}_{1:k})p(\boldsymbol{x}_k|\boldsymbol{x}_{k-1})}{p(\boldsymbol{y}_k|\boldsymbol{y}_{1:k-1})} \times p(\boldsymbol{x}_{1:k-1}|\boldsymbol{y}_{1:k-1}). \quad (4)$$

Due to the fact that the $\boldsymbol{y}_k$ is related to $\boldsymbol{x}_k, \boldsymbol{x}_{k-1}, \cdots, \boldsymbol{x}_{k-I}$, the likelihood function can be written as $p(\boldsymbol{y}_k|\boldsymbol{x}_{1:k}) = p(\boldsymbol{y}_k|\boldsymbol{x}_k, \boldsymbol{x}_{k-1}, \cdots, \boldsymbol{x}_{k-I})$. To obtain the posterior probability $p(\boldsymbol{x}_k|\boldsymbol{y}_{1:k})$, (4) is marginalized as follows [11]:

$$\begin{aligned} p(\boldsymbol{x}_k|\boldsymbol{y}_{1:k}) = \frac{1}{p(\boldsymbol{y}_k|\boldsymbol{y}_{1:k-1})} \int_{\boldsymbol{x}_{k-I}} \cdots \int_{\boldsymbol{x}_{k-1}} p(\boldsymbol{x}_k|\boldsymbol{x}_{k-1}) p(\boldsymbol{y}_k|\boldsymbol{x}_k, \boldsymbol{x}_{k-1}, \cdots, \boldsymbol{x}_{k-I}) \\ \times p(\boldsymbol{x}_{k-1}, \cdots, \boldsymbol{x}_{k-I}|\boldsymbol{y}_{1:k-1}) d\boldsymbol{x}_{k-1} d\boldsymbol{x}_{k-I}. \end{aligned} \quad (5)$$

Therefore, consider the following augmented system:

$$\begin{aligned} \boldsymbol{X}_k &= \boldsymbol{F}(\boldsymbol{X}_{k-1}) + \boldsymbol{B}\boldsymbol{w}_{k-1}, \\ \boldsymbol{y}_k &= \boldsymbol{v}_k \text{ or } \boldsymbol{y}_k = \boldsymbol{h}\left(\boldsymbol{C}^i \boldsymbol{X}_k\right) + \boldsymbol{v}_{k-i}, \end{aligned} \quad (6)$$

where $\boldsymbol{X}_k = [\boldsymbol{x}_k^T, \boldsymbol{x}_{k-1}^T, \cdots \boldsymbol{x}_{k-I}^T]^T$, $\boldsymbol{F}(\boldsymbol{X}_{k-1}) = [\boldsymbol{f}(\boldsymbol{x}_{k-1})^T, \boldsymbol{x}_{k-1}^T, \cdots \boldsymbol{x}_{k-I}^T]^T$, $\boldsymbol{C}^i \boldsymbol{X}_k = \boldsymbol{x}_{k-i}$, $\boldsymbol{B} = [\boldsymbol{I}_n, 0, \cdots, 0]^T$, and $\boldsymbol{I}_n$ is the n-dimensional identity matrix. Then, the likelihood function can be rewritten as

$$\begin{aligned} p(\boldsymbol{y}_k|\boldsymbol{X}_k) &= \sum_{i=0}^{I} p\left(\boldsymbol{y}_k, \lambda_k^i, \rho_k = 1|\boldsymbol{X}_k\right) + p(\boldsymbol{y}_k, \rho_k = 0|\boldsymbol{X}_k) \\ &= \sum_{i=0}^{I} p\left(\boldsymbol{y}_k|\lambda_k^i, \rho_k = 1, \boldsymbol{X}_k\right) p\left(\lambda_k^i|\rho_k = 1\right) \times p(\rho_k = 1) \\ &\quad + p(\boldsymbol{y}_k|\rho_k = 0, \boldsymbol{X}_k) p(\rho_k = 0) \\ &= \sum_{i=0}^{I} (1 - \gamma_k)\mu_k^i N(\boldsymbol{y}_k; \boldsymbol{h}(\boldsymbol{C}^i \boldsymbol{X}_k), \boldsymbol{R}_{k-i}) + \gamma_k N(\boldsymbol{y}_k; 0, \boldsymbol{R}_k). \end{aligned} \quad (7)$$

It can be seen that the likelihood function is a mixture of multiple Gaussian distribution, so the posterior probability can be updated in the way of GSF. The prior probability of the augmented state is approximated to be a single Gaussian distribution as

$$p(\boldsymbol{X}_{k-1}|\boldsymbol{y}_{1:k-1}) = N\left(\boldsymbol{X}_{k-1}; \widehat{\boldsymbol{X}}_{k-1|k-1}, \boldsymbol{P}_{k-1|k-1}\right), \quad (8)$$

and the prediction PDF can be obtained through the Chapman-Kolmogorov equation as

$$p(\boldsymbol{X}_k|\boldsymbol{y}_{1:k-1}) = \int_{\boldsymbol{X}_{k-1}} p(\boldsymbol{X}_k|\boldsymbol{X}_{k-1}) p(\boldsymbol{X}_{k-1}|\boldsymbol{y}_{1:k-1}) d\boldsymbol{X}_{k-1}. \quad (9)$$

According to the Bayesian rules, the augmented state posterior density $p(\boldsymbol{X}_k|\boldsymbol{y}_{1:k})$ can be received through GSF, i.e.,

$$\begin{aligned} p(\boldsymbol{X}_k|\boldsymbol{y}_{1:k}) &= \frac{p(\boldsymbol{y}_k|\boldsymbol{X}_k) p(\boldsymbol{X}_k|\boldsymbol{y}_{1:k-1})}{\int p(\boldsymbol{y}_k|\boldsymbol{X}_k) p(\boldsymbol{X}_k|\boldsymbol{y}_{1:k-1}) d\boldsymbol{X}_k} \\ &= \sum_{i=0}^{I+1} \widehat{\tau}_k^i N\left(\boldsymbol{X}_k; \widehat{\boldsymbol{X}}_{k|k}^i, \boldsymbol{P}_{k|k}^i\right), \end{aligned} \quad (10)$$

where $\widehat{\boldsymbol{X}}_{k|k}^i$ and $\boldsymbol{P}_{k|k}^i$ are the mean and covariance of state, respectively, which can be obtained through KF, and $\widehat{\tau}_k^i$ is the weight of Gaussian components at time index k, which is calculated by

$$\begin{cases} \tilde{\tau}_k^i = (1 - \gamma_k)\mu_k^i N\left(\boldsymbol{y}_k; \boldsymbol{h}\left(\boldsymbol{C}^i \widehat{\boldsymbol{X}}_{k|k-1}\right), \boldsymbol{R}_{k-i}\right), i = 0, \cdots, I, \\ \tilde{\tau}_k^i = \gamma_k N(\boldsymbol{y}_k; 0, \boldsymbol{R}_k), i = I + 1, \end{cases} \quad (11)$$

$$\widehat{\tau}_k^i = \frac{\tilde{\tau}_k^i}{\sum_{i=0}^{I+1} \tilde{\tau}_k^i}, \quad (12)$$

where $\widehat{\boldsymbol{X}}_{k|k-1}$ represents the prediction state. When the probability of measurement delay and loss is time varying and unknown, it will inevitably lead to poor accuracy of state estimate if inaccurate prior probability is used for approximation. Therefore, we use the VB approach to obtain the measurement delay and loss probability while estimating the state. However, the Gaussian mixture model of (7) is a nonconjugated prior distribution. In order to use the VB method to jointly estimate the state and parameter, we construct a Gaussian hierarchical model; that is, the likelihood function (7) is modified into the form of exponential multiplication by using the random variables $\boldsymbol{\lambda}_k$ and ρ_k, which represent measurement delay and loss, respectively, i.e.,

$$\begin{aligned} p(\boldsymbol{y}_k|\boldsymbol{X}_k, \rho_k, \boldsymbol{\lambda}_k) &= [N(\boldsymbol{y}_k; 0, \boldsymbol{R}_k)]^{(1-\rho_k)} \\ &\quad \times \left\{ \prod_{i=0}^{I} \left[N\left(\boldsymbol{y}_k; \boldsymbol{h}\left(\boldsymbol{C}^i \boldsymbol{X}_k\right), \boldsymbol{R}_{k-i}\right) \right]^{\lambda_k^i} \right\}^{\rho_k}. \end{aligned} \quad (13)$$

3.2. The Variational Bayesian Method. In order to infer the state $\boldsymbol{X}_k$ together with the Bernoulli random variable ρ_k, the binary random variable $\boldsymbol{\lambda}_k$, the loss probability γ_k, and the delay probability $\boldsymbol{\mu}_k$, the joint PDF $p(\Theta_k|\boldsymbol{y}_{1:k})$ needs to be calculated recursively which satisfies

$$p(\Theta_k|\boldsymbol{y}_{1:k}) = \frac{p(\boldsymbol{y}_k|\Theta_k) p(\Theta_k|\boldsymbol{y}_{1:k-1})}{p(\boldsymbol{y}_k|\boldsymbol{y}_{1:k-1})}, \quad (14)$$

where $\Theta_k = \{\boldsymbol{X}_k, \rho_k, \boldsymbol{\lambda}_k, \gamma_k, \boldsymbol{\mu}_k\}$. By making use of the mean field theory, we have

$$q(\Theta_k) \approx q(\boldsymbol{X}_k) q(\rho_k) q(\boldsymbol{\lambda}_k) q(\gamma_k) q(\boldsymbol{\mu}_k), \quad (15)$$

where $q(\cdot)$ is the approximate posterior PDF which can be calculated through VB approach as follows:

$$\log q(\phi_k) = \underset{\Theta_k - \phi_k}{E}[\log p(\Theta_k, \boldsymbol{y}_{1:k})] + c_{\phi_k}, \tag{16}$$

where ϕ_k is an arbitrary element of Θ_k, $\Theta_k - \phi_k$ denotes the set Θ_k without ϕ_k, c_{ϕ_k} is the constant related to ϕ_k, and $E[\cdot]$ represents the expectation operation. The joint distribution $p(\Theta_k, \boldsymbol{y}_{1:k})$ can be divided into

$$\begin{aligned} p(\Theta_k, \boldsymbol{y}_{1:k}) &= p(\boldsymbol{y}_k|\boldsymbol{X}_k, \rho_k, \boldsymbol{\lambda}_k)p(\boldsymbol{X}_k|\boldsymbol{y}_{1:k-1}) \\ &\times p(\rho_k|\gamma_k)p(\gamma_k|\boldsymbol{y}_{1:k-1})p(\boldsymbol{\lambda}_k|\boldsymbol{\mu}_k) \\ &\times p(\boldsymbol{\mu}_k|\boldsymbol{y}_{1:k-1})p(\boldsymbol{y}_{1:k-1}). \end{aligned} \tag{17}$$

Based on the VB inference, the conjugate prior distributions for γ_k and $\boldsymbol{\mu}_k$ are expressed as Beta distribution and Dirichlet distribution [30], i.e.,

$$p(\gamma_{k-1}|\boldsymbol{y}_{1:k-1}) = Be\left(\gamma_{k-1}\,;\hat{\alpha}_{k-1|k-1}, \hat{\beta}_{k-1|k-1}\right), \tag{18}$$

$$p(\boldsymbol{\mu}_{k-1}|\boldsymbol{y}_{1:k-1}) = D\left(\boldsymbol{\mu}_{k-1}\,;\hat{\boldsymbol{a}}_{k-1|k-1}\right). \tag{19}$$

Considering the unknown probability change over time, we adopt the Beta-Bartlett evolutionary model to obtain the predicted distribution of the parameter as [31]

$$\begin{aligned} \hat{\alpha}_{k|k-1} &= \varsigma\hat{\alpha}_{k-1|k-1}, \\ \hat{\beta}_{k|k-1} &= \varsigma\hat{\beta}_{k-1|k-1}, \end{aligned} \tag{20}$$

$$\hat{\boldsymbol{a}}_{k|k-1} = \varsigma\hat{\boldsymbol{a}}_{k-1|k-1}, \tag{21}$$

where ς is the forgetting factor whose value is usually between $[0.95\;1)$.

Substituting (2), (3), (9), (13), and (18)–(21) in (17), the $\log p(\Theta_k, y_{1:k})$ is given by

$$\begin{aligned} \log p(\Theta_k, \boldsymbol{y}_{1:k}) = &\sum_{i=0}^{I} \rho_k \lambda_k^i \left\{-0.5\left[\boldsymbol{y}_k - \boldsymbol{h}\left(\boldsymbol{C}^i\boldsymbol{X}_k\right)\right]^T \boldsymbol{R}_{k-i}^{-1}\left[\boldsymbol{y}_k - \boldsymbol{h}\left(\boldsymbol{C}^i\boldsymbol{X}_k\right)\right]\right\} \\ &+ (1-\rho_k)\left(-0.5\boldsymbol{y}_k^T\boldsymbol{R}_k^{-1}\mathrm{y}_k\right) \\ &- 0.5\left(\boldsymbol{X}_k - \hat{\boldsymbol{X}}_{k|k-1}\right)^T \boldsymbol{P}_{k|k-1}^{-1}\left(\boldsymbol{X}_k - \hat{\boldsymbol{X}}_{k|k-1}\right) \\ &+ \rho_k \log(1-\gamma_k) + (1-\rho_k)\log\gamma_k \\ &+ \left(\hat{\alpha}_{k|k-1} - 1\right)\log\gamma_k + \left(\hat{b}_{k|k-1} - 1\right)\log(1-\gamma_k) \\ &+ \sum_{i=0}^{I}\lambda_k^i \log\left(\mu_k^i\right) + \sum_{i=0}^{I}\left(\hat{a}_{k|k-1}^i - 1\right)\log\left(\mu_k^i\right) + C_{\Theta_k}. \end{aligned} \tag{22}$$

By updating one element of Θ_k while keeping the others' estimated value constant, the VB solution of (16) can be obtained iteratively.

Let $\phi_k = \boldsymbol{X}_k$, and substituting (22) in (16), it follows

$$\begin{aligned} \log q^{(d+1)}(\boldsymbol{X}_k) = &\sum_{i=0}^{I} E^{(d)}[\rho_k]E^{(d)}\left[\lambda_k^i\right] \\ &\times\left\{-0.5\left[\boldsymbol{y}_k - \boldsymbol{h}\left(\boldsymbol{C}^i\boldsymbol{X}_k\right)\right]^T\boldsymbol{R}_{k-i}^{-1}\left[\boldsymbol{y}_k - \boldsymbol{h}\left(\boldsymbol{C}^i\boldsymbol{X}_k\right)\right]\right\} \\ &- 0.5\left(\boldsymbol{X}_k - \hat{\boldsymbol{X}}_{k|k-1}\right)^T\boldsymbol{P}_{k|k-1}^{-1}\left(\boldsymbol{X}_k - \hat{\boldsymbol{X}}_{k|k-1}\right), \end{aligned} \tag{23}$$

where d represents the number of iterations. From (23), we can see that $q^{(d+1)}(\boldsymbol{X}_k) \propto (1/\tilde{c}_{\boldsymbol{X}_k})p(\boldsymbol{X}_k|\boldsymbol{y}_{1:k-1}) \times p^{(d+1)}(\boldsymbol{y}_k|\boldsymbol{X}_k)$, where the predicted PDF $p(\boldsymbol{X}_k|\boldsymbol{y}_{1:k-1})$ and the likelihood PDF $p^{(d+1)}(\boldsymbol{y}_k|\boldsymbol{X}_k)$ are defined as follows:

$$p(\boldsymbol{X}_k|\boldsymbol{y}_{1:k-1}) = N\left(\boldsymbol{X}_k\,;\hat{\boldsymbol{X}}_{k|k-1}, \boldsymbol{P}_{k|k-1}\right), \tag{24}$$

$$p^{(d+1)}(\boldsymbol{y}_k|\boldsymbol{X}_k) = \sum_{i=0}^{I} E^{(d)}[\rho_k]E^{(d)}\left[\lambda_k^i\right] \times N\left(\boldsymbol{y}_k\,;\boldsymbol{h}\left(\boldsymbol{C}^i\boldsymbol{X}_k\right), \boldsymbol{R}_{k-i}\right), \tag{25}$$

where $\hat{\boldsymbol{X}}_{k|k-1}$ and $\boldsymbol{P}_{k|k-1}$ can be calculated through CKF as

$$\hat{\boldsymbol{X}}_{k|k-1} = \frac{1}{2n}\sum_{j=1}^{2n}\boldsymbol{F}\left(\chi_{k-1|k-1}^j\right), \tag{26}$$

$$\boldsymbol{P}_{k|k-1} = \frac{1}{2n}\sum_{j=1}^{2n}\left[\left(\boldsymbol{F}\left(\chi_{k-1|k-1}^j\right) - \hat{\boldsymbol{X}}_{k|k-1}\right)(\cdot)^T\right] + \boldsymbol{Q}_{k-1}, \tag{27}$$

where $\chi_{k-1|k-1}^j, j = 1, 2, \cdots, 2n$ is the jth cubature point generated by

$$\chi_{k-1|k-1}^j = \hat{\boldsymbol{X}}_{k-1|k-1} + \boldsymbol{S}_{k-1}\zeta_j, \tag{28}$$

where $\boldsymbol{S}_{k-1}$ is the Cholesky decomposition of $\boldsymbol{P}_{k-1|k-1}$ which satisfy $\boldsymbol{P}_{k-1|k-1} = \boldsymbol{S}_{k-1}(\boldsymbol{S}_{k-1})^T$ and ζ_j is the jth element of the following set:

$$\sqrt{n}\left[\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}, \cdots, \begin{pmatrix}0\\0\\\vdots\\1\end{pmatrix}, \begin{pmatrix}-1\\0\\\vdots\\0\end{pmatrix}, \cdots, \begin{pmatrix}0\\0\\\vdots\\-1\end{pmatrix}\right]. \tag{29}$$

It is noted that when the measurement is lost, there is no effect on the measurement update of the state, so the likelihood function obtained according to the VB method does not contain the terms related to measurement loss, which also leads to that the sum of weights in (25) is less than 1. However, in order to satisfy the condition that the sum of probability is 1 when GSF is running, we refer to (7) to keep the parameter terms related to measurement loss in the likelihood function, i.e.,

$$p^{(d+1)}(\boldsymbol{y}_k|\boldsymbol{X}_k) = \sum_{i=0}^{I} E^{(d)}[\rho_k]E^{(d)}\left[\lambda_k^i\right]N\left(\boldsymbol{y}_k ; \boldsymbol{h}\left(\boldsymbol{C}^i\boldsymbol{X}_k\right), \boldsymbol{R}_{k-i}\right) + \left(1 - E^{(d)}[\rho_k]\right)N(\boldsymbol{y}_k ; 0, \boldsymbol{R}_k), \tag{30}$$

which means that the proposed algorithm only does time update when the measurement is lost. And we can obtain the augmented state estimates of mean and covariance of each component in the $d+1$th iteration, which is formulated as

$$\begin{cases} \boldsymbol{P}_{xy}^i = \dfrac{1}{2n}\sum_{j=1}^{2n}\left(\boldsymbol{F}\left(\chi_{k|k-1}^j\right) - \widehat{\boldsymbol{X}}_{k|k-1}\right)\left(\widehat{\boldsymbol{y}}_{k|k-1}^i - \boldsymbol{h}\left(\boldsymbol{C}^i\chi_{k|k-1}^j\right)^T\right), \\ \boldsymbol{P}_{yy}^i = \dfrac{1}{2n}\sum_{j=1}^{2n}\left(\widehat{\boldsymbol{y}}_{k|k-1}^i - \boldsymbol{h}\left(\boldsymbol{C}^i\chi_{k|k-1}^j\right)\right) \times \left(\widehat{\boldsymbol{y}}_{k|k-1}^i - \boldsymbol{h}\left(\boldsymbol{C}^i\chi_{k|k-1}^j\right)\right) + \boldsymbol{R}_{k-i}, \\ \boldsymbol{K}_k^i = \boldsymbol{P}_{xy}^i\left(\boldsymbol{P}_{yy}^i\right)^{-1}, \\ \widehat{\boldsymbol{X}}_{k|k}^i = \widehat{\boldsymbol{X}}_{k-1|k} + \boldsymbol{K}_k^i\left(\boldsymbol{y}_k - \widehat{\boldsymbol{y}}_{k-1|k}^i\right), \\ \boldsymbol{P}_{k|k}^i = \boldsymbol{P}_{k-1|k} - \boldsymbol{K}_k^i\boldsymbol{P}_{yy}^i\left(\boldsymbol{K}_k^i\right)^T, \\ \widehat{\boldsymbol{y}}_{k|k-1}^i = \dfrac{1}{2n}\sum_{j=1}^{2n}\boldsymbol{h}\left(\boldsymbol{C}^i\chi_{k|k-1}^j\right), \end{cases} \quad , i = 0, \cdots, I,$$

$$\widehat{\boldsymbol{X}}_{k|k}^{I+1} = \widehat{\boldsymbol{X}}_{k-1|k}, \quad \boldsymbol{P}_{k|k}^{I+1} = \boldsymbol{P}_{k-1|k}, \tag{31}$$

where

$$\chi_{k|k-1}^j = \widehat{\boldsymbol{X}}_{k|k-1} + \boldsymbol{S}_k\zeta_j, \tag{32}$$

and $\boldsymbol{S}_k$ satisfied $\boldsymbol{P}_{k|k-1} = \boldsymbol{S}_k(\boldsymbol{S}_k)^T$.

To reduce the computational burden, we approximate the posterior probability of the state estimate as a single Gaussian as

$$\begin{cases} \widehat{\boldsymbol{X}}_{k|k}^{(d+1)} = \sum_{i=0}^{I+1}\widehat{\tau}_k^i\widehat{\boldsymbol{X}}_{k|k}^i, \\ \boldsymbol{P}_{k|k}^{(d+1)} = \sum_{i=0}^{I+1}\widehat{\tau}_k^i\left[\boldsymbol{P}_{k|k}^i + \left(\widehat{\boldsymbol{X}}_{k|k}^{(d+1)} - \widehat{\boldsymbol{X}}_{k|k}^i\right)(\cdot)^T\right], \end{cases} \tag{33}$$

where $\widehat{\tau}_k^i$ can be calculated as (11).

Let $\phi_k = \rho_k$, and substituting (22) in (16), it follows

$$\log q^{(d+1)}(\rho_k) = \sum_{i=0}^{I}\rho_k E^{(d)}\left[\lambda_k^i\right]\left\{-0.5\left[\boldsymbol{y}_k - \boldsymbol{h}\left(\boldsymbol{C}^i\boldsymbol{X}_k\right)\right]^T\boldsymbol{R}_{k-i}^{-1} \times \left[\boldsymbol{y}_k - \boldsymbol{h}\left(\boldsymbol{C}^i\boldsymbol{X}_k\right)\right]\right\} + (1-\rho_k)\left(-0.5\boldsymbol{y}_k^T\boldsymbol{R}_k^{-1}\boldsymbol{y}_k\right) + \rho_k\log(1-\gamma_k) + (1-\rho_k)\log\gamma_k. \tag{34}$$

From (34), we can see that $q^{(d+1)}(\rho_k)$ is a Bernoulli distribution, i.e.,

$$\begin{cases} P_\rho^{(d+1)}(\rho_k = 1) = \exp\left\{E^{(d)}[\log(1-\gamma_k)] - \sum_{i=0}^{I}0.5E^{(d)}\left[\lambda_k^i\right]tr\left(\boldsymbol{A}_k^i\boldsymbol{R}_{k-i}^{-1}\right)\right\}, \\ P_\rho^{(d+1)}(\rho_k = 0) = \exp\left\{E^{(d)}[\log(\gamma_k)] - 0.5tr\left(\boldsymbol{B}_k\boldsymbol{R}_{k-i}^{-1}\right)\right\}. \end{cases} \tag{35}$$

where

$$\begin{cases} E^{(d+1)}[\log(\gamma_k)] = \psi\left(\widehat{\alpha}_{k|k}^{(d+1)}\right) - \psi\left(\widehat{\alpha}_{k|k}^{(d+1)} + \widehat{\beta}_{k|k}^{(d+1)}\right), \\ E^{(d+1)}[\log(1-\gamma_k)] = \psi\left(\widehat{\beta}_{k|k}^{(d+1)}\right) - \psi\left(\widehat{\alpha}_{k|k}^{(d+1)} + \widehat{\beta}_{k|k}^{(d+1)}\right), \end{cases} \tag{36}$$

and $\boldsymbol{A}_k^i$ and $\boldsymbol{B}_k$ are given by

$$\begin{aligned} \boldsymbol{A}_k^i &= E\left[\left[\boldsymbol{y}_k - \boldsymbol{h}\left(\boldsymbol{C}^i\boldsymbol{X}_k\right)\right][\cdot]^T\right] = \left[\boldsymbol{y}_k - \boldsymbol{h}\left(\boldsymbol{C}^i\widehat{\boldsymbol{X}}_{k|k}^{(d+1)}\right)\right][\cdot]^T \\ &\quad + \frac{1}{2n}\sum_{j=1}^{2n}\left(\widehat{\boldsymbol{y}}_{k|k-1}^i - \boldsymbol{h}\left(\boldsymbol{C}^i\chi_{k|k-1}^j\right)\right)(\cdot)^T, \\ \boldsymbol{B}_k &= \boldsymbol{y}_k\boldsymbol{y}_k^T. \end{aligned} \tag{37}$$

Hence, the expectation of ρ_k can be updated as

$$E^{(d+1)}[\rho_k] = \frac{P_\rho^{(d+1)}(\rho_k = 1)}{P_\rho^{(d+1)}(\rho_k = 1) + P_\rho^{(d+1)}(\rho_k = 0)}. \tag{38}$$

Let $\phi_k = \gamma_k$, and substituting (22) in (16), it follows

$$\log q^{(d+1)}(\gamma_k) = E^{(d+1)}[\rho_k]\log(1-\gamma_k) + \left(1 - E^{(d+1)}[\rho_k]\right)\log\gamma_k + \left(\widehat{\alpha}_{k|k-1} - 1\right)\log\gamma_k + \left(\widehat{b}_{k|k-1} - 1\right)\log(1-\gamma_k). \tag{39}$$

From (39), we can see that $q^{(d+1)}(\gamma_k)$ is a Beta distribution, i.e., $q^{(d+1)}(\gamma_k) = Be(\gamma_k ; \widehat{\alpha}_{k|k}^{(d+1)}, \widehat{\beta}_{k|k}^{(d+1)})$, where the parameters $\widehat{\alpha}_{k|k}^{(d+1)}$ and $\widehat{\beta}_{k|k}^{(d+1)}$ are updated as

$$\begin{cases} \widehat{\alpha}_{k|k}^{(d+1)} = \widehat{\alpha}_{k|k-1} + 1 - E^{(d+1)}[\rho_k], \\ \widehat{\beta}_{k|k}^{(d+1)} = \widehat{\beta}_{k|k-1} + E^{(d+1)}[\rho_k]. \end{cases} \tag{40}$$

Let $\phi_k = \lambda_k$, and substituting (22) in (16), it follows

$$\log q^{(d+1)}(\lambda_k) = \sum_{i=0}^{I}E^{(d+1)}[\rho_k]\lambda_k^i\left\{-0.5\left[\boldsymbol{y}_k - \boldsymbol{h}\left(\boldsymbol{C}^i\boldsymbol{X}_k\right)\right]^T\boldsymbol{R}_{k-i}^{-1}\left[\boldsymbol{y}_k - \boldsymbol{h}\left(\boldsymbol{C}^i\boldsymbol{X}_k\right)\right]\right\} + \sum_{i=0}^{I}\lambda_k^i\log\left(\mu_k^i\right). \tag{41}$$

From (41), we can see that $q^{(d+1)}(\lambda_k)$ is supposed to be a

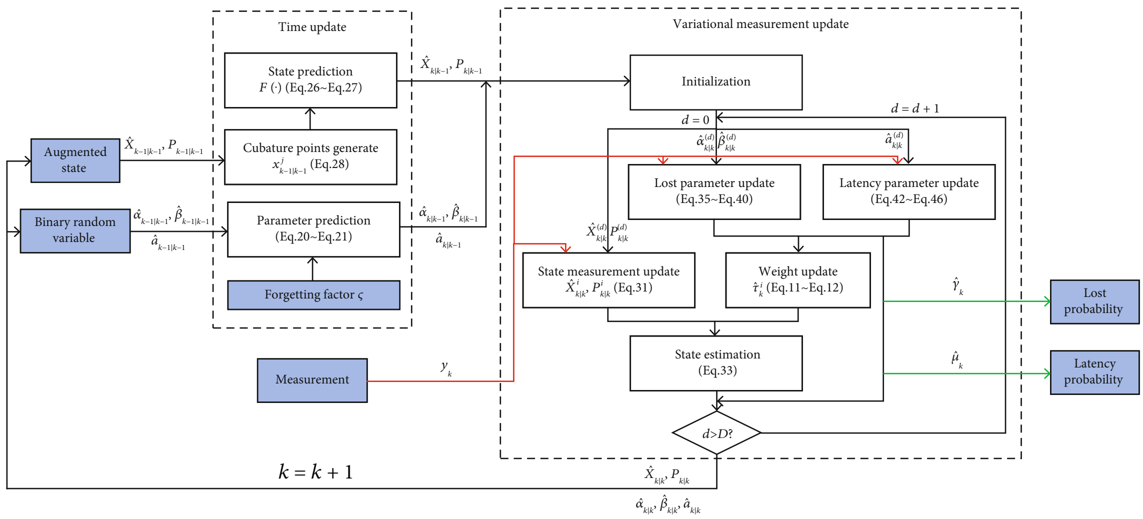

Figure 1: The flow chart of the algorithm.

Input: $\widehat{\mathbf{X}}_{k-1|k-1}, \mathbf{P}_{k-1|k-1}, \mathbf{f}(\cdot), \mathbf{h}(\cdot), \mathbf{y}_k, \mathbf{Q}_{k-1}, \mathbf{R}_{k-1}, \widehat{\alpha}_{k-1|k-1}, \widehat{\beta}_{k-1|k-1}, \widehat{\mathbf{a}}_{k-1|k-1}, \varsigma, I, N$

Time update:

Obtain $\widehat{\mathbf{X}}_{k|k-1}$ *and* $\mathbf{P}_{k|k-1}$ *using ((26)) and ((27))*

$\widehat{\alpha}_{k|k-1} = \varsigma\widehat{\alpha}_{k-1|k-1}, \widehat{\beta}_{k|k-1} = \varsigma\widehat{\beta}_{k-1|k-1}, \widehat{\mathbf{a}}_{k|k-1} = \varsigma\widehat{\mathbf{a}}_{k-1|k-1};$

Variational measurement update:

Initialization $\widehat{\mathbf{X}}^{(0)}_{k|k} = \widehat{\mathbf{X}}_{k|k-1}$, $\mathbf{P}^{(0)}_{k|k} = \mathbf{P}_{k|k-1}$, $\widehat{\alpha}^{(0)}_{k|k} = \widehat{\alpha}_{k|k-1}$, $\widehat{\beta}^{(0)}_{k|k} = \widehat{\beta}_{k|k-1}$, $\widehat{\mathbf{a}}^{(0)}_{k|k} = \widehat{\mathbf{a}}_{k|k-1}$, $E^{(0)}[\rho_k] = \widehat{\beta}^{(0)}_{k|k}/(\widehat{\alpha}^{(0)}_{k|k} + \widehat{\beta}^{(0)}_{k|k})$, $E^{(0)}[\boldsymbol{\lambda}_k] = \dfrac{\widehat{\mathbf{a}}^{(0)}_{k|k}}{\sum_{i=1}^{I}\widehat{a}^{(0),i}_{k|k}}$,

$E^{(0)}[\log(\gamma_k)] = \psi(\widehat{\alpha}^{(0)}_{k|k}) - \psi(\widehat{\alpha}^{(0)}_{k|k} + \widehat{\beta}^{(0)}_{k|k})$, $E^{(0)}[\log(1-\gamma_k)] = \psi(\widehat{\beta}^{(0)}_{k|k}) - \psi(\widehat{\alpha}^{(0)}_{k|k} + \widehat{\beta}^{(0)}_{k|k})$, $E^{(0)}[\log(\mu^i_k)] = \psi(\widehat{a}^{(0),i}_{k|k}) - \psi(\sum_{i=1}^{I}\widehat{a}^{(0),i}_{k|k})$;

for $d = 0$ **to** $D-1$ **do**

  Update $q^{(d+1)}(\mathbf{X}_k) = N(\mathbf{X}_k; \widehat{\mathbf{X}}^{(d+1)}_{k|k}, \mathbf{P}^{(d+1)}_{k|k})$ *given* $E^{(d)}[\rho_k]$ *and* $E^{(d)}[\lambda_k]$ *using ((33));*

  Update $E^{(d+1)}[\rho_k]$ *using ((38));*

  Update $\widehat{\alpha}^{(0)}_{k|k}$ *and* $\widehat{\beta}^{(0)}_{k|k}$ *using ((40));*

  Compute $E^{(d+1)}[\log(\gamma_k)]$ *and* $E^{(d+1)}[\log(1-\gamma_k)]$ *as ((36));*

  Update $E^{(d+1)}[\boldsymbol{\lambda}_k]$ *using ((44));*

  Update $\widehat{\mathbf{a}}^{(d+1)}_{k|k}$ *using ((46));*

  Compute $E^{(d+1)}[\log(\mu^i_k)]$ *as ((43));*

end

$\widehat{\mathbf{X}}_{k|k} = \widehat{\mathbf{X}}^{(d+1)}_{k|k}$, $\mathbf{P}_{k|k} = \mathbf{P}^{(d+1)}_{k|k}$, $\widehat{\alpha}_{k|k} = \widehat{\alpha}^{(d+1)}_{k|k}$, $\widehat{\beta}_{k|k} = \widehat{\beta}^{(d+1)}_{k|k}$, $\widehat{\mathbf{a}}_{k|k} = \widehat{\mathbf{a}}^{(d+1)}_{k|k}$;

Output: $\widehat{\mathbf{X}}_{k|k}, \mathbf{P}_{k|k}, \widehat{\alpha}_{k|k}, \widehat{\beta}_{k|k}, \widehat{\mathbf{a}}_{k|k}$,The estimated lost probability $\widehat{\gamma}_k = \widehat{\beta}_{k|k}/(\widehat{\alpha}_{k|k} + \widehat{\beta}_{k|k})$,The estimated latency probability $\widehat{\mu}_k = \widehat{\mathbf{a}}_{k|k}/\sum_{i=1}^{I}\widehat{a}^i_{k|k}$

Algorithm 1: The proposed VB-based Gaussian sum cubature Kalman filter algorithm.

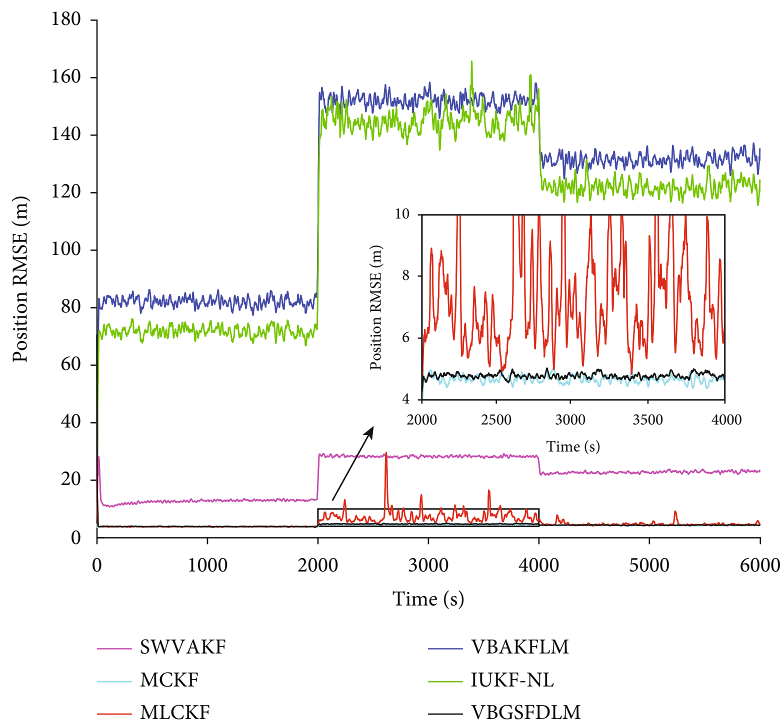

Figure 2: RMSEs of position of the VBAKFLM, IUKF-NL, SWVAKF, MLCKF, MCKF, and proposed VBGSFDLM.

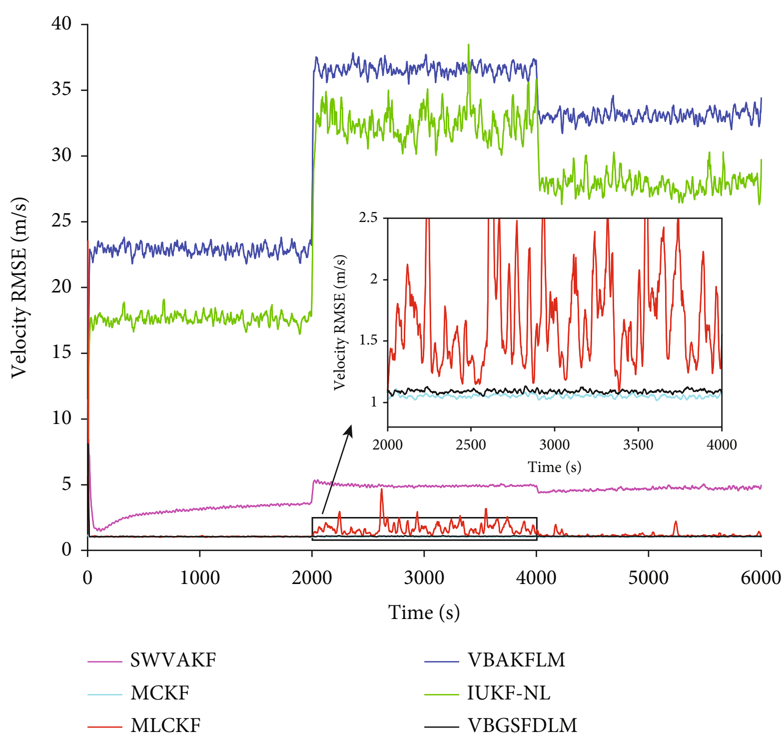

Figure 3: RMSEs of velocity of the VBAKFLM, IUKF-NL, SWVAKF, MLCKF, MCKF, and proposed VBGSFDLM.

multinomial distribution, i.e., $q^{(d+1)}(\lambda_k)=\prod_{i=0}^{I}(\widehat{\xi}_k^{(d+1),i})^{\lambda_k^i}$, where the parameter $\widehat{\xi}_k^{(d+1),i}$ can be updated as

$$\xi_k^{(d+1),i}\propto\left\{E^{(d)}\left[\log\left(\mu_k^i\right)\right]-0.5E^{(d+1)}[\rho_k]tr\left[\boldsymbol{A}_k^i\boldsymbol{R}_{k-i}^{-1}\right]\right\},$$
$$\widehat{\xi}_k^{(d+1),i}=\frac{\xi_k^{(d+1),i}}{\sum_{i=0}^{I}\xi_k^{(d+1),i}},\tag{42}$$

where

$$E^{(d+1)}\left[\log\left(\mu_k^i\right)\right]=\psi\left(\widehat{a}_{k|k}^{(d+1),i}\right)-\psi\left(\sum_{i=0}^{I}\widehat{a}_{k|k}^{(d+1),i}\right),\tag{43}$$

and $E^{(d+1)}[\lambda_k^i]$ is updated as

$$E^{(d+1)}\left[\lambda_k^i\right]=\widehat{\xi}_k^{(d+1)}.\tag{44}$$

Let $\phi_k=\boldsymbol{\mu}_k$, and substituting (22) in (16), it follows

$$\log q^{(d+1)}(\boldsymbol{\mu}_k)=\sum_{i=0}^{I}\left\{E^{(d+1)}\left[\lambda_k^i\right]+\left(\widehat{\boldsymbol{a}}_{k|k-1}^i-1\right)\right\}\log\left(\mu_k^i\right).\tag{45}$$

From (45), we can see $q^{(d+1)}(\boldsymbol{\mu}_k)$ is supposed to be a Dirichlet distribution, i.e., $q^{(d+1)}(\boldsymbol{\mu}_k)=D(\boldsymbol{\mu}_k;\widehat{a}_{k|k}^{(d+1)})$, where the parameter $\widehat{\boldsymbol{a}}_{k|k}^{(d+1)}$ can be updated as

$$\widehat{\boldsymbol{a}}_{k|k}^{(d+1)}=\widehat{\boldsymbol{a}}_{k|k-1}+\widehat{\xi}_k^{(d+1)}.\tag{46}$$

The flow chart and one time step of the proposed algorithm are summarized in Figure 1 and Algorithm 1, respectively.

Remark 1. The common way of dealing with (23) is augmenting measurement vector, and the modified matrix $\tilde{\boldsymbol{R}}_{k-i}=\boldsymbol{R}_{k-i}/(E^{(d)}[\rho_k]E^{(d)}[\lambda_k^i])$is given to the diagonally augmented measurement noise covariance matrix [27, 29, 32, 33]. However, in some cases, the probability of delay or loss of a step, $E[\lambda_k^i]$ or $E[\rho_k^i]$, may be close to zero. It will cause $\tilde{\boldsymbol{R}}_{k-i}$ to be infinite and lead to filter collapse. In this paper, this problem is avoided by using probability to update the weight of each Gaussian component. Therefore, the stability of the proposed algorithm is improved.

4. Simulations

In this section, the proposed VB-based Gaussian sum cubature Kalman filter with delay and loss measurement

Table 1: $ARMSE_{pos}$ and $ARMSE_{vel}$ of VBAKFLM, IUKF-NL, SWVAKF, MLCKF, MCKF, and proposed VBGSFDLM.

Filter	$ARMSE_{pos}$ (m)	$ARMSE_{vel}$ (m/s)
VBAKFLM	122.5286	30.9837
IUKF-NL	113.7579	26.1438
SWVAKF	21.3963	4.2468
MLCKF	5.3894	1.3123
VBGSFDLM	4.3503	1.0769
MCKF	4.2964	1.0448

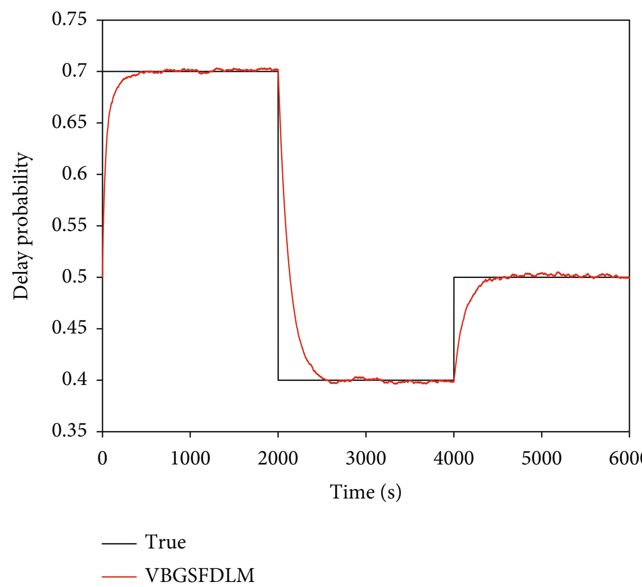

Figure 4: True and estimated delay probability.

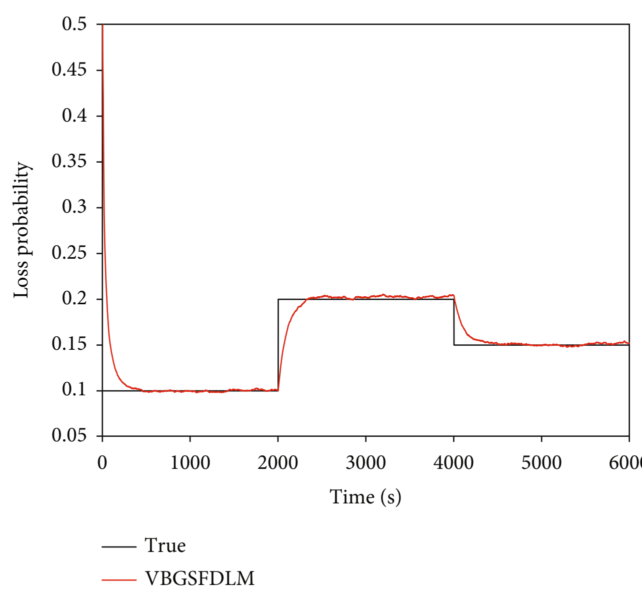

Figure 5: True and estimated loss probability.

Table 2: Computational complexity of VBAKFLM, IUKF-NL, SWVAKF, MLCKF, and proposed VBGSFDLM.

Filter	Computational complexity
VBAKFLM	$o\left(2n^3 + n^2 + D\left(n^3 + 4n^2m + 3nm^2 + 3nm + 2m^3 + 2m^2\right)\right)$
IUKF-NL	$o\left(M_n\left(2N_x^2 + N_x + M_x\right) + 2N_x^2 + D\left(4N_z^3 + 3N_z^2 + 3N_z^2N_x + 2N_x^2N_z + 3N_xN_z + M_z + M_n\left(N_x^2 + N_z^2 + N_z + N_xN_z + M_z\right)\right)\right)$
SWVAKF	$o\left(3N_x^3 + N_x^2 + 3N_xN_z + 4N_x^2N_z + 3N_xN_z^2 + N_z^2 + L\left(11N_x^3 + 4N_x^2 + N_x^2N_z + N_xN_z^2 + N_xN_z\right)\right)$
MLCKF	$o\left(M_n\left(3N_x^2 + N_x + M_x + N_xN_z + M_z + N_z^2 + N_z\right) + 2N_x^2 + 2N_z^2 + 2N_xN_z + (I+1)\left(2N_xN_z^2 + N_x^2N_z + N_xN_z\right)\right)$
VBGSFDLM	$o\left(M_n\left(2N_x^2 + N_x + M_x\right) + 2N_x^2 + D\left(N_z^3 + 4N_z^2 + 2N_xN_z + M_z + N_xN_z^2 + N_x^2N_z + M_n\left(N_x^2 + N_xN_z + M_z + N_z^2 + N_z\right) + (I+1)\left(2N_xN_z^2 + N_x^2N_z + N_xN_z + 4N_z^3\right)\right)\right)$

(VBGSFDLM) effectiveness is verified by a tracking simulation applied in air-traffic control system with unknown probability of MRMDL [11]. The maximum number of delays is set as 3. And the target dynamic is modeled by a constant turn rate model, i.e.,

$$\boldsymbol{x}_k = \begin{bmatrix} 1 & \frac{\sin(\Omega T)}{\Omega} & 0 & -\frac{1-\cos(\Omega T)}{\Omega} & 0 \\ 0 & \cos(\Omega T) & 0 & -\sin(\Omega T) & 0 \\ 0 & \frac{1-\cos(\Omega T)}{\Omega} & 1 & \frac{\sin(\Omega T)}{\Omega} & 0 \\ 0 & \sin(\Omega T) & 0 & \cos(\Omega T) & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \boldsymbol{x}_{k-1} + \boldsymbol{w}_{k-1}, \tag{47}$$

where $T = 1$ s denotes the sampling interval and Ω represents turn rate; the state vector $\boldsymbol{x}_k$ consists of position and velocity in the x and y directions and Ω, which is defined as $\boldsymbol{x}_k = [x_k^T, \dot{x}_k^T, y_k^T, \dot{y}_k^T, \Omega]^T$, while the nonlinear measurement equation is composed of range r_k and bearing θ_k:

$$\begin{aligned} \boldsymbol{z}_k &= \rho_k \begin{bmatrix} r_k \\ \theta_k \end{bmatrix} + \boldsymbol{v}_k = \rho_k \begin{bmatrix} \sqrt{x_k^2 + y_k^2} \\ \arctan\left(\frac{y_k}{x_k}\right) \end{bmatrix} + \boldsymbol{v}_k, \\ \boldsymbol{y}_k &= \lambda_k^0 \boldsymbol{z}_k + \lambda_k^1 \boldsymbol{z}_{k-1} + \lambda_k^2 \boldsymbol{z}_{k-2} + \lambda_k^3 \boldsymbol{z}_{k-3}, \end{aligned} \tag{48}$$

where ρ_k and λ_k are both random variables with probability as

$$\begin{aligned} p(\rho_k = 0) &= \begin{cases} 0.1, k < 2000, \\ 0.2, 2000 \le k < 4000, \\ 0.15, 4000 \le k \le 6000, \end{cases} \\ p(\lambda_k^0 = 0) &= \begin{cases} 0.7, k < 2000, \\ 0.4, 2000 \le k < 4000, \\ 0.5, 4000 \le k \le 6000. \end{cases} \end{aligned} \tag{49}$$

In the simulation, we consider that each step has the same delay probability, so the probability distribution of λ_k can be expressed as $p(\lambda_k^0 = 1) = 1 - p(\lambda_k^0 = 0)$, $p(\lambda_k^1 = 1) = (1 - p(\lambda_k^0 = 0))p(\lambda_k^0 = 0)$, $p(\lambda_k^2 = 1) = (1 - p(\lambda_k^0 = 0))(p(\lambda_k^0 = 0))^2$, and $p(\lambda_k^2 = 1) = (p(\lambda_k^0 = 0))^3$. The initial state estimatation $\hat{\boldsymbol{x}}_{0|0}$ is generated randomly from the true value $\boldsymbol{x}_0 = [0,20,0,0,0.15]^T$ and $\boldsymbol{P}_0 = \mathrm{diag}\,[10000,100,10000,100,0.01]$. The covariances $\boldsymbol{Q}$ and $\boldsymbol{R}$ are given as $\boldsymbol{Q} = \mathrm{diag}\,[\boldsymbol{q}_1, \boldsymbol{q}_1, T]$ and $\boldsymbol{R} = \mathrm{diag}\,[\delta_r^2, \delta_\theta^2]$ where

Table 3: Single step running time of SWVAKF, MLCKF, IUKF-NL, VBAKFLM, and proposed VBGSFDLM.

Filter	Single step running time (ms)
VBAKFLM	3.32*e*-1
IUKF-NL	1.37
SWVAKF	1.53*e*-1
MLCKF	3.76*e*-1
VBGSFDLM	3.81

$$\begin{aligned} \boldsymbol{q}_1 &= \begin{bmatrix} \frac{T^3}{3} & \frac{T^2}{2} \\ \frac{T^2}{2} & T \end{bmatrix}, \\ \delta_r &= 5\ \mathrm{m}, \\ \delta_\theta &= 0.0017\ \mathrm{rad}. \end{aligned} \tag{50}$$

To evaluate the accuracy of the state estimation, the root mean square errors (RMSEs) of position and velocity are considered as the metric:

$$\begin{aligned} \mathrm{RMSE}_p &= \sqrt{\frac{1}{M}\sum_{m=1}^{M}\left[(x^m - \hat{x}^m)^2 + (y^m - \hat{y}^m)^2\right]}, \\ \mathrm{RMSE}_v &= \sqrt{\frac{1}{M}\sum_{m=1}^{M}\left[\left(\dot{x}^m - \hat{\dot{x}}^m\right)^2 + \left(\dot{y}^m - \hat{\dot{y}}^m\right)^2\right]}. \end{aligned} \tag{51}$$

The number of Monte Carlo runs is set as $M = 500$.

The proposed filter is compared with the MLCKF [11], the matched CKF (MCKF), the SWVAKF [23], the IUKF with a nominal lost probability (IUKF-NL) [28], and the VBAKFLM [29]. The nominal latency probability of MLCKF is set as $\mu_k = [0.5, 0.25, 0.125, 0.125]^T$. The window length is set as $L = 5$ in SWVAKF. In the IUKF-NL and VBAKFLM, the initial shape parameters are selected as $\hat{\alpha}_{0|0} = 10$, $\hat{\beta}_{0|0} = 10$. The nominal loss probability of measurement is set as 0.5. In the proposed VBGSFDLM, $\hat{\alpha}_{0|0} = 10$, $\hat{\beta}_{0|0} = 10$, and $\hat{\mathbf{a}}_{0|0} = [16, 8, 4, 4]^T$, which is the same as the weight of MLCKF. And the number of iterations is set as $D = 10$, forgetting factor $\varsigma = 0.97$ for SWVAKF, IUKF-NL, VBAKFLM, and VBGSFDLM. In MCKF, the delay and loss of measurement are precisely known. Therefore, the estimated result of MCKF can be regarded as the optimal result.

Figures 2 and 3 show the RMSE_p and RMSE_v of the VBAKFLM, IUKF-NL, SWVAKF, MLCKF, MCKF, and VBGSFDLM. And the average RMSEs (ARMSEs) are given in Table 1. It can be seen that the ARMSEs of position of the proposed VBGSFDLM is reduced by 96.45%, 96.18%, 79.67%, and 19.27%, respectively, compared with VBAKFLM, IUKF-NL, SWVAKF, and MLCKF while the ARMSEs of velocity is reduced by 96.52%, 95.88%, 74.64%, and 17.94%. Due to the fact that the IUKF-NL and VBAKFLM can only

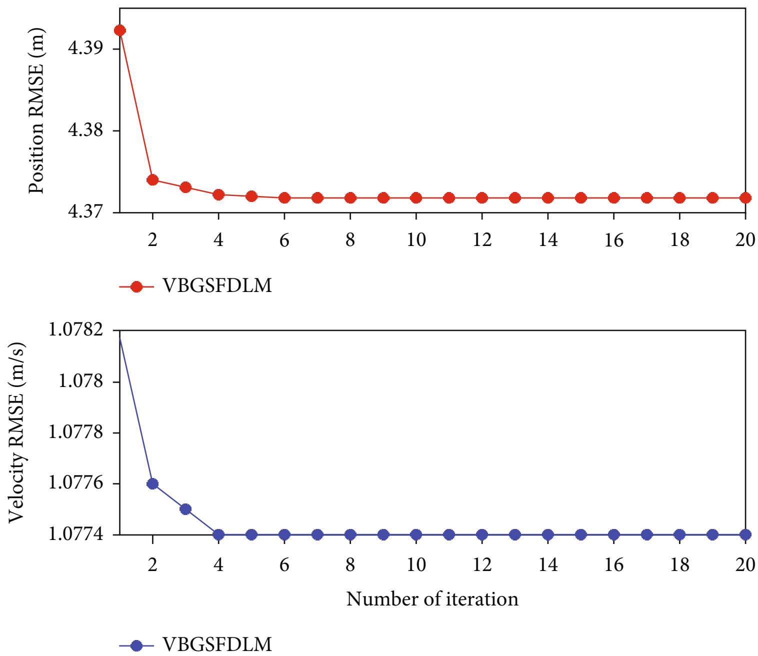

Figure 6: ARMSEs of the position and velocity when $D = 1, 2, \cdots, 20$.

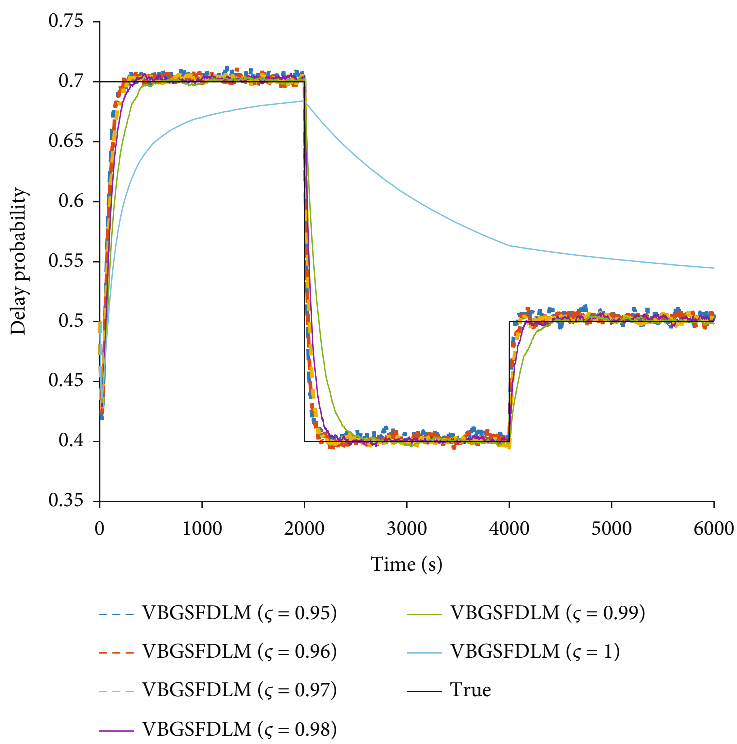

Figure 7: True and estimated delay probability when $\varsigma = 0.95, 0.96, 0.97, 0.98, 0.99$, and 1.

estimate the probability of one-step delay and measurement loss, respectively. The state estimation accuracy of above two filters will be affected when measurement loss and multistep randomly delay exist at the same time. Although the SWVAKF uses a sliding window to reduce the influence of multistep random delay, the estimation accuracy is still affected by time-varying delay probability because it is designed to identify time-varying Gaussian noise. While MLCKF updates the weight of each state estimation component by combining likelihood probability and multistep delay probability, it achieves some degree of robustness against inaccurate prior delay and loss probability under the framework of

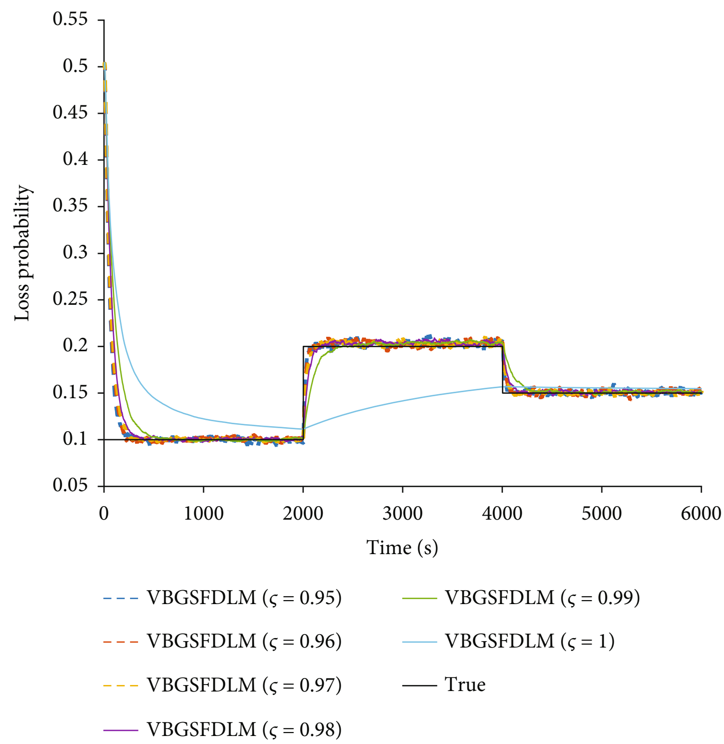

Figure 8: True and estimated loss probability when $\varsigma = 0.95, 0.96, 0.97, 0.98, 0.99$, and 1.

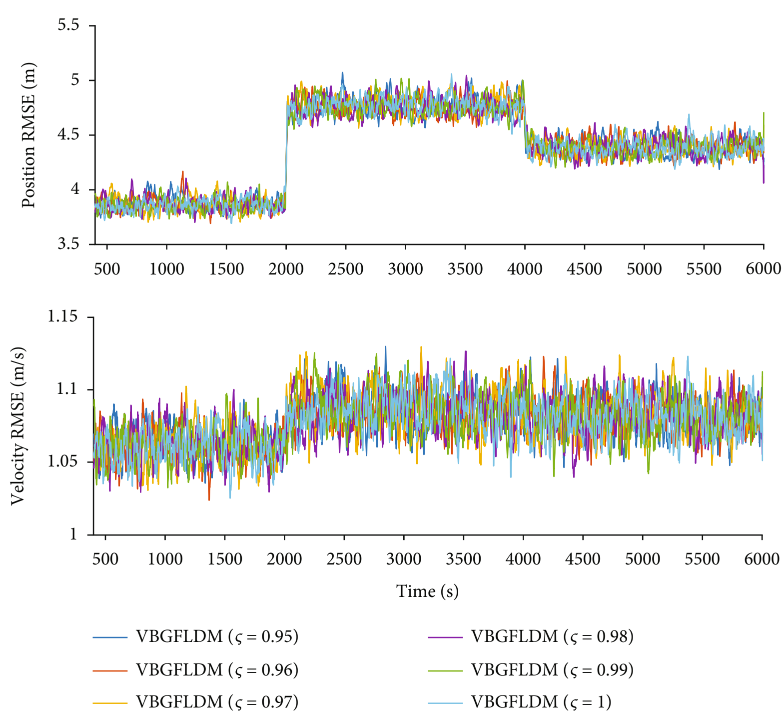

Figure 9: RMSEs of position and velocity of VBGSFDLM when $\varsigma = 0.95, 0.96, 0.97, 0.98, 0.99$, and 1.

GSF. However, MLCKF does not have the ability to identify the time-varying measurement delay and loss probability. In the proposed VBGSFDLM, the augmented state includes the state with multistep random delay and a Gaussian component with the predicted state is added to the GSF to deal with the case of measurement loss. In addition, the augmented state and parameters are estimated jointly by VB method to obtain accurate identification of the unknown measurement delay and lost probability which improves the estimation accuracy of the state. Therefore, the estimation accuracy is the closest

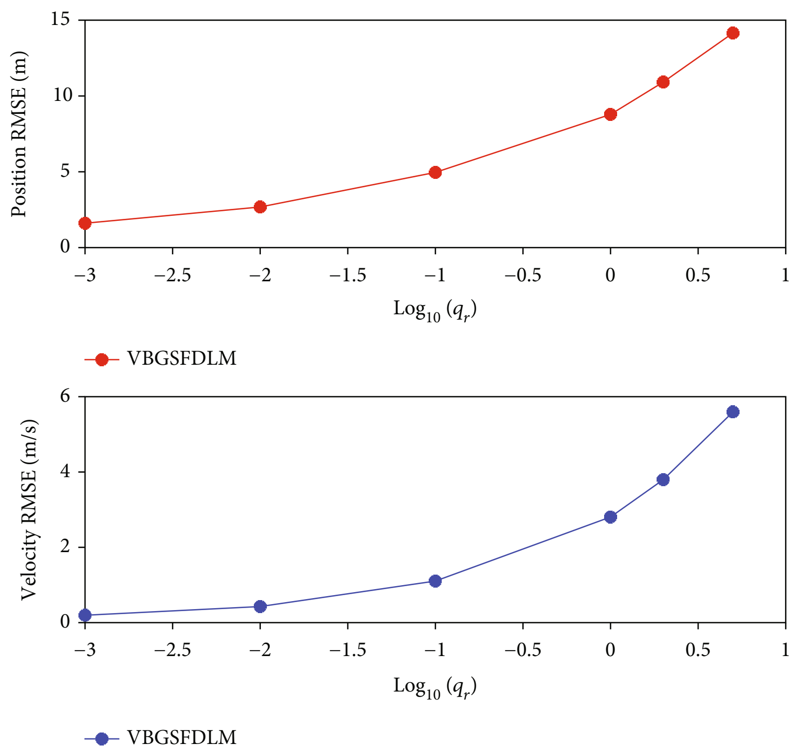

Figure 10: RMSEs of position and velocity of VBGSFDLM when $q_r = 0.001, 0.01, 0.1, 1, 2,$ and 5.

to MCKF. Furthermore, the true and estimated delay and loss probability are shown in Figures 4 and 5, which illustrate that the proposed VBGSFDLM can accurately estimate the multi-step random delay and loss probability.

Besides, the computational complexity of VBGSFDLM is compared with VBAKFLM, IUKF-NL, SWVAKF, MLCKF, and VBGSFDLM. The number of multiplications in each algorithm is considered as the evaluation criterion. The results are shown in Table 2. And the single step running time of each filters is shown in Table 3, where $N_x = (I+1)n$ and $N_z = (I+1)m$ are the dimension of the augmented state and measurement vector, respectively. $M_n = 2N_x$ is the number of sampling points. M_x and M_z represent the number of multiplications in nonlinear transformations $\boldsymbol{F}(\cdot)$ and $\boldsymbol{h}(\cdot)$ in augmented system. It can be seen that the significant improvement of estimation accuracy is at the cost of computation.

Further, we discuss the effect of the number of iterations on VBGSFDLM. Figure 6 shows the ARMSEs of position and velocity when $D = 1, 2, \cdots, 20$. It can be seen from Figure 6 that the VBGSFDLM has well estimation accuracy when $d \geq 4$ and converges when $d \geq 8$.

Next, we study the forgetting factor effect of the time-varying MRMDL on the performance of VBGSFDLM. Figures 7 and 8 show the true and estimated delay and loss probability. And Figure 9 shows the RMSEs of position and velocity of VBGSFDLM when $\varsigma = 0.95, 0.96, 0.97, 0.98, 0.99,$ and 1. It can be seen from Figures 7 and 8 that VBGSFDLM with $\varsigma = 0.95, 0.96, 0.97, 0.98,$ and 0.99 has essentially consistent estimation performance in probability estimation. However, VBGSFDLM with $\varsigma = 1.0$ converges slowly because $\varsigma = 1.0$ corresponds to the case of constant probability of MRMDL so that the estimation performance degrades when the actual probability is slowly varying. Besides, Figure 9 shows that the RMSEs of position and velocity are almost the same in different ς.

Finally, in order to investigate the influence of target maneuvering on the tracking performance of the VBGSFDLM, different values of process noise $\mathbf{Q}_k = q_r \mathbf{Q}$ are taken in the simulation. In Figure 10, the RMSEs of position and velocity are shown with $q_r = 0.001, 0.01, 0.1, 1, 2,$ and 5. Obviously, it can be seen in Figure 10 that the tracking RMSE increases when the process noise increases. This shows that the proposed algorithm needs a precise state equation to achieve accurate estimation of unknown delay and loss probabilities. And process noise will significantly affect the estimation results of the algorithm.

5. Conclusion

In this paper, we proposed a VB-based Gaussian sum filter to obtain the estimation of state for nonlinear systems with MRMDL with unknown probability. By introducing two random variables, the Gaussian mixture distribution is rewritten into an exponential multiplication form and VB method is used to estimate the state and the unknown measurement delay and lost probability jointly. Simulations show that the proposed VBGSFDLM has a better performance in state estimation and probability identification in the presence of unknown and time-varying delay and loss probability.

Data Availability

The data that support the findings of this study are available from the corresponding author.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

References

[1] Y. Huang, Y. Zhang, N. Li, Z. Wu, and J. A. Chambers, “A novel robust Student's t-based Kalman filter,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 53, no. 3, pp. 1545–1554, 2017.

[2] H. Xu, W. Xie, H. Yuan, K. Duan, W. Liu, and Y. Wang, “Fixed-point iteration Gaussian sum filtering estimator with unknown time-varying non-Gaussian measurement noise,” *Signal Processing*, vol. 153, pp. 132–142, 2018.

[3] W. Qin, X. Wang, Y. Bai, and N. Cui, “Arbitrary-step randomly delayed robust filter with application to boost phase tracking,” *Acta Astronautica*, vol. 145, pp. 304–318, 2018.

[4] N. Zhu, J. Marais, D. Bétaille, and M. Berbineau, “GNSS position integrity in urban environments: a review of literature,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 9, pp. 2762–2778, 2018.

[5] R. Caballero-Águila, A. Hermoso-Carazo, and J. Linares-Pérez, “Networked distributed fusion estimation under uncertain outputs with random transmission delays, packet losses and multipacket processing,” *Signal Processing*, vol. 156, pp. 71–83, 2019.

[6] A. Hermoso-Carazo and J. Linares-Pérez, “Extended and unscented filtering algorithms using one-step randomly delayed observations,” *Applied Mathematics and Computation*, vol. 190, no. 2, pp. 1375–1393, 2007.

[7] S. Challa, R. J. Evans, and X. Wang, “A Bayesian solution and its approximations to out-of-sequence measurement problems,” *Information Fusion*, vol. 4, no. 3, pp. 185–199, 2003.

[8] Y. Fei, T. Meng, and Z. Jin, “Nano satellite attitude determination with randomly delayed measurements,” *Acta Astronautica*, vol. 185, pp. 319–332, 2021.

[9] X. Wang, Y. Liang, Q. Pan, and C. Zhao, “Gaussian filter for nonlinear systems with one-step randomly delayed measurements,” *Automatica*, vol. 49, no. 4, pp. 976–986, 2013.

[10] X. Wang, Q. Pan, Y. Liang, and F. Yang, “Gaussian smoothers for nonlinear systems with one-step randomly delayed measurements,” *IEEE Transactions on Automatic Control*, vol. 58, no. 7, pp. 1828–1835, 2013.

[11] R. Esmzad and R. M. Esfanjani, “Bayesian filter for nonlinear systems with randomly delayed and lost measurements,” *Automatica*, vol. 107, pp. 36–42, 2019.

[12] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. I. Jordan, and S. S. Sastry, “Kalman filtering with intermittent observations,” *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1453–1464, 2004.

[13] K. You, F. Minyue, and L. Xie, “Mean square stability for kalman filtering with Markovian packet losses,” *Automatica*, vol. 47, no. 12, pp. 2647–2657, 2011.

[14] L. Xu, K. Ma, W. Li, Y. Liu, and F. E. Alsaadi, “Particle filtering for networked nonlinear systems subject to random one-step sensor delay and missing measurements,” *Neurocomputing*, vol. 275, pp. 2162–2169, 2018.

[15] Y. Chen, Z. Yan, and X. Zhang, “A novel particle filtering for nonlinear systems with multistep randomly delayed measurements,” *Applied Mathematical Modelling*, vol. 100, pp. 282–302, 2021.

[16] S. Sun, “Optimal linear filters for discrete-time systems with randomly delayed and lost measurements with/without time stamps,” *IEEE Transactions on Automatic Control*, vol. 58, no. 6, pp. 1551–1556, 2013.

[17] Y. Huang, Y. Zhang, and N. Li, “Latency probability estimation of non-linear systems with one-step randomly delayed measurements,” *IET Control Theory and Applications*, vol. 10, no. 7, pp. 843–852, 2016.

[18] Y. L. XiaoxuWang, Q. Pan, and Y. Wang, “Measurement random latency probability identification,” *IEEE Transactions on Automatic Control*, vol. 61, no. 12, pp. 4210–4216, 2016.

[19] R. K. Tiwari and S. Bhaumik, “Risk sensitive filtering with randomly delayed measurements,” *Automatica*, vol. 142, article 110409, 2022.

[20] G. Jia, Y. Zhang, M. Bai, N. Li, and J. Qian, “A novel robust students t-based Gaussian approximate filter with one-step randomly delayed measurements,” *Signal Processing*, vol. 171, article 107496, 2020.

[21] Y. Zhao, A. Fatehi, and B. Huang, “Robust estimation of arx models with time varying time delays using variational Bayesian approach,” *IEEE Transactions on Cybernetics*, vol. 48, no. 2, pp. 532–542, 2018.

[22] M. Rongjun, S. Bingzhi, J. Chen, Y. Li, and N. Cui, “Multiple-step randomly delayed adaptive robust filter with application to INS/VNS integrated navigation on asteroid missions,” *IEEE Access*, vol. 8, pp. 118853–118868, 2020.

[23] Y. Huang, F. Zhu, G. Jia, and Y. Zhang, “A slide window variational adaptive Kalman filter,” *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 67, no. 12, pp. 3552–3556, 2020.

[24] C. Shan, W. Zhou, Z. Jiang, and H. Shan, “A new Gaussian approximate filter with colored non-stationary heavy-tailed measurement noise,” *Digital Signal Processing*, vol. 122, article 103358, 2022.

[25] H. Zhu, K. Zou, Y. Li, and H. Leung, “Robust sensor fusion with heavy-tailed noises,” *Signal Processing*, vol. 175, article 107659, 2020.

[26] C. Shan, W. Zhou, Y. Yang, and H. Shan, “A new Gaussian-Student's t mixing distribution-based Kalman filter with unknown measurement random delay rate,” *International Journal of Adaptive Control and Signal Processing*, vol. 36, no. 7, pp. 1724–1745, 2022.

[27] Z. Wang, Y. Huang, Y. Zhang, G. Jia, and J. Chambers, “An improved Kalman filter with adaptive estimate of latency probability,” *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 67, no. 10, pp. 2259–2263, 2020.

[28] Y. Tong, Z. Zheng, W. Fan, Q. Li, and Z. Liu, “An improved unscented Kalman filter for nonlinear systems with one-step randomly delayed measurement and unknown latency probability,” *Digital Signal Processing*, vol. 121, article 103324, 2022.

[29] G. Jia, Y. Huang, Y. Zhang, and J. Chambers, “A novel adaptive Kalman filter with unknown probability of measurement loss,” *IEEE Signal Processing Letters*, vol. 26, no. 12, pp. 1862–1866, 2019.

[30] J. Forster and A. OHagan, *Kendalls Advanced Theory of Statistics: Bayesian Inference: 2b (Kendalls Library of Statistics)*, Arnold, 2004.

[31] A. Assa and K. N. Plataniotis, "Adaptive Kalman filtering by covariance sampling," *IEEE Signal Processing Letters*, vol. 24, no. 9, pp. 1288–1292, 2017.

[32] X. Lai, J. Huang, C. Ye, F. Sun, and Y. Liu, "Adaptive multinoulli-based Kalman filter with randomly unknown delayed and lost measurements," *Digital Signal Processing*, vol. 129, article 103653, 2022.

[33] H. Zhu, J. Mi, Y. Li, K.-V. Yuen, and H. Leung, "VB-Kalman based localization for connected vehicles with delayed and lost measurements: theory and experiments," *IEEE/ASME Transactions on Mechatronics*, vol. 27, no. 3, pp. 1370–1378, 2022.