HIGH TECHNOLOGY LETTERS | Vol. 28 No. 2 | June 2022 | pp. 181-189

doi:10.3772/j.issn.1006-6748.2022.02.008

Gaussian-Student's t mixture distribution PHD robust filtering algorithm based on variational Bayesian inference①

HU Zhentao(胡振涛)*, YANG Linlin*, HU Yumei②**, YANG Shibo*

(*School of Artificial Intelligence, Henan University, Zhengzhou 450046, P. R. China)

(**School of Automation, Northwestern Polytechnical University, Xi'an 710029, P. R. China)

Abstract

Aiming at the problem of filtering precision degradation caused by the random outliers of process noise and measurement noise in multi-target tracking (MTT) system, a new Gaussian-Student's t mixture distribution probability hypothesis density (PHD) robust filtering algorithm based on variational Bayesian inference (GST-vbPHD) is proposed. Firstly, since it can accurately describe the heavy-tailed characteristics of noise with outliers, Gaussian-Student's t mixture distribution is employed to model process noise and measurement noise respectively. Then Bernoulli random variable is introduced to correct the likelihood distribution of the mixture probability, leading hierarchical Gaussian distribution constructed by the Gaussian-Student's t mixture distribution suitable to model non-stationary noise. Finally, the approximate solutions including target weights, measurement noise covariance and state estimation error covariance are obtained according to variational Bayesian inference approach. The simulation results show that, in the heavy-tailed noise environment, the proposed algorithm leads to strong improvements over the traditional PHD filter and the Student's t distribution PHD filter.

Key words: multi-target tracking (MTT), variational Bayesian inference, Gaussian-Student's t mixture distribution, heavy-tailed noise

0 Introduction

Multi-target tracking (MTT) technique based on point measurements is used to real-time estimate the number of targets, status, trajectory, and other attribute information with the processing of measurement information. The traditional implementation of MTT generally adopts the data association strategies, such as joint probabilistic data association (JPDA)[1], multi-hypothesis tracking (MHT)[2], and probabilistic multi-hypothesis tracking (PMHT)[3]. However, these above methods cannot deal well with the time-varying characteristics of the target state, i.e. the time-varying number of targets makes it difficult to achieve an effective correlation between the state set and the measurement set of the target. Recently, since bypassing the complex data association, the MTT based on random finite set (RFS) theory and its improvements have attracted extensive attention[4]. Specifically, their complexity and track ability are better than those methods using data association strategy. A typical implementation mentioned above is the probability hypothesis density (PHD) filter which recursively solves the state posterior first-order statistical moments, thus gives the first engineering implementation of RFS[5]. The existing PHD filter implementation strategies mainly include sequential Monte Carlo PHD (SMC-PHD)[6] and Gaussian mixture PHD (GM-PHD)[7-8].

In practical engineering applications, the noise outlier induced by electromagnetic interference, aging of the sensor, and uncertainty of the dynamic model will deteriorate PHD filter tracking accuracy. Besides, the outlier-containing noise usually exhibits heavy-tailed characteristics. However, traditional GM-PHD suffers poor robustness at heavy-tailed process noise and measurement noise existing[9]. Under the condition of Gaussian distribution, the SMC-PHD filter may partially relieve the above problem with high computational cost. Huber's M-estimation theory can be used

① Supported by the National Natural Science Foundation of China (No. 61976080), the Science and Technology Key Project of Science and Technology Department of Henan Province (No. 212102310298), and the Innovation and Quality Improvement Project for Graduate Education of Henan University (No. SYL20010101).

② To whom correspondence should be addressed. E-mail: huym_henu@163.com.

Received on Aug. 24, 2021

to improve the GM-PHD filter's performance when outliers exist in the measurement model, but it cannot deal with outliers in process noisy. Moreover, since it is based on Gaussian distribution approximation, GM-PHD filter may induce biased estimates on the state and number of targets, thus is unsuitable to handle non-Gaussian noise system model with noisy outliers[10]. Existing literatures show that heavy-tailed noise may not be efficiently tackled in Gaussian noise hypothetical scenario, so heavy-tailed noise modeling becomes the key to deal with multi-target tracking problem with noise outliers[11].

Since Student's t distribution exhibits heavier tail than Gaussian distribution and converges to Gaussian distribution as its freedom increasing, it may be suitable for modeling non-Gaussian noise with significant heavy-tailed. Assuming the measurement noise follows Student's t distribution, Li et al.[12] proposed a robust PHD filter which used VB to update the posterior likelihood function, but the method is unsuitable for noisy outliers. Liu et al.[13] presented a robust Student's t mixture PHD filter by recursively propagating the intensity as a mixture of Student's t components in PHD filtering framework. In addition, to alleviate the unfavorable effects on filtering performance induced by heavy-tailed noise, Liu re-weighted on true measurement, outliers and clutter according to their value, and proposed M-estimation based dual-gating strategy to construct a Student's t mixture distribution. With approximately regarding the process noise and measurement noise as the Student's t distribution, Hong et al.[14] proposed a Student's t mixture particle PHD (STMP-PHD) filter. They argued that the intensity of the multi-target may be approximated by using a Student's t mixture model, while Monte Carlo is utilized to calculate the Student's t function integral, leading to a closed Student's t hybrid recursive framework. However, few literatures above focus on improving the filtering robustness by using variational Bayesian inference. Zhang et al.[15] designed a robust Student's t based labeled multi-Bernoulli (RSTLMB) filter through modeling the Student's t distribution with the state prediction probability density and the measurement likelihood function of individual targets. Moreover, a closed recursion filter is proposed to jointly estimate the target state and the parameters of the Student's t distribution. Due to the random occurrence of the outliers in noise, RSTLMB hardly model non-stationarity of noise by using one single Student's t distribution.

Obviously, using fixed inverse scale matrix or Student's t distribution can hardly model random noise with heavy-tailed outliers. To address the above problem, a new Gaussian-Student's t mixture distribution PHD robust filtering algorithm based on variational Bayesian inference (GST-vbPHD) is proposed here. The main contributions are summarized as follows.

(1) Random outliers existing in process noise and measurement noise are modeled as Gaussian-Student's t mixture distribution, in addition, the parameters of the mixture distribution and kinematic state are integrated in the augmentation matrix.

(2) Bernoulli random variables are introduced to transform the mixture distribution model of noise outliers into a hierarchical Gaussian form in which parameters including targets states and weights are updated by variational Bayesian inference.

(3) In different experiment scenarios, two types of performance indicators, the optimal subpattern assignment (OSPA) distance and the accuracy of the target number estimation, are used to verify the feasibility and validity of the proposed algorithm. The experiment results demonstrate that the proposed algorithm outperforms the comparison methods on tracking accuracy.

1 Gaussian mixed PHD filter

Suppose $M(k)$ and $N(x)$ denote respectively the numbers of target states and measurements in the monitoring area at time k. The set of multi-target states and the set of multi-target observations are denoted as $\boldsymbol{X}_k = \{\boldsymbol{x}_{k,1}, \cdots, \boldsymbol{x}_{k,M(k)}\}$ and $\boldsymbol{Z}_k = \{\boldsymbol{z}_{k,1}, \cdots, \boldsymbol{z}_{k,N(k)}\}$, respectively. Let $\boldsymbol{X}_{k-1}$ be the set of multi-target states at time $k-1$, then $\boldsymbol{X}_k$ and $\boldsymbol{Z}_k$ can be expressed as

$$\boldsymbol{X}_k = [\bigcup_{x \in X_{k-1}} S_{k|k-1}(\boldsymbol{x})] \cup [\bigcup_{x \in X_{k-1}} B_{k|k-1}(\boldsymbol{x})] \cup \Gamma_k \tag{1}$$

$$\boldsymbol{Z}_k = (\bigcup_{x \in X_k} \Theta(\boldsymbol{x})) \cup \boldsymbol{\kappa}_k \tag{2}$$

where $S_{k|k-1}(\boldsymbol{x})$ and $B_{k|k-1}(\boldsymbol{x})$ denote the random finite sets of survival targets and spawned targets from $\boldsymbol{X}_{k-1}$ at time k, respectively. Γ_k is the random finite sets of birth targets at time k. $\Theta(\boldsymbol{x})$ and $\boldsymbol{\kappa}_k$ denote respectively the observed random sets generated by targets and clutters at time k.

The PHD filter estimates the states of targets and its number by iteratively propagating the posterior intensity, which is a first order statistic of the random finite set[7]. The linear Gaussian MTT system develops Gaussian mixture implementation in finding the analytic solution of the Bayesian integral, the process of which clearly demonstrates how the Gaussian components propagate analytically to the next moment. Assume that the prior intensity function $\boldsymbol{v}_{k-1}$ at time $k-1$ obeys the Gaussian distribution

$$\boldsymbol{v}_{k-1} = \sum_{i=1}^{J_{k-1}} w_{k-1}^{(i)} N(\boldsymbol{x};\ \boldsymbol{m}_{k-1}^{(i)},\ \boldsymbol{P}_{k-1}^{(i)}) \tag{3}$$

J_{k-1} is the Gaussian component at time $k-1$, $w_{k-1}^{(i)}$ is the weight corresponding to the i-th Gaussian component. $N(\ :;\boldsymbol{m},\ \boldsymbol{P})$ denotes the Gaussian density function with the mean $\boldsymbol{m}$ and covariance $\boldsymbol{P}$. $\boldsymbol{m}_{k-1}^{(i)}$ and $\boldsymbol{P}_{k-1}^{(i)}$ denote the target state and the state error covariance, respectively. At time k, the target predicted intensity function $\boldsymbol{v}_{k|k-1}(\boldsymbol{x})$ and updated intensity function $\boldsymbol{v}_k(\boldsymbol{x})$ are obtained from the Bayesian recursive estimation as follows

$$\boldsymbol{v}_{k|k-1}(\boldsymbol{x}) = \sum_{i=1}^{J_{k|k-1}} w_{k|k-1}^{(i)} N(\boldsymbol{x};\ \boldsymbol{m}_{k|k-1}^{(i)},\ \boldsymbol{P}_{k|k-1}^{(i)}) \tag{4}$$

$$\boldsymbol{v}_k(\boldsymbol{x}) = (1 - P_{D,k}) v_{k|k-1}(\boldsymbol{x}) + \sum_{z \in Z_k} \sum_{i=1}^{J_{k|k-1}} w_k^{(i)}(\boldsymbol{z}) N(\boldsymbol{x};\ \boldsymbol{m}_k^{(i)}(\boldsymbol{z}),\ \boldsymbol{P}_k^{(i)}) \tag{5}$$

where $J_{k|k-1}$ and $w_{k|k-1}^{(i)}$ denote respectively the predicted Gaussian components and the corresponding prediction weights. The target state prediction estimates is represented as $\boldsymbol{m}_{k|k-1}^{(i)}$ and the state prediction error covariance is marked as $\boldsymbol{P}_{k|k-1}^{(i)}$. $P_{D,k}$ refers to the detection probability. $w_k^{(i)}$, $\boldsymbol{m}_k^{(i)}$ and $\boldsymbol{P}_k^{(i)}$ denote the updated target weights, state estimates and state estimation error covariance at time k, respectively.

2 GST-vbPHD filter

2.1 Gaussian-Student's t mixture distribution

The outliers of the process noise and measurement noise may appear at different moments in practical engineering application, resulting in the non-stationarity characteristic of the non-Gaussian noise. The modeling of noise with outliers, by adopting the mixed probability η as a mixed Gaussian-Student's t distribution, is as follows

$$p(\boldsymbol{x} \mid \eta) = \eta N(\boldsymbol{x};\ \boldsymbol{m},\ \boldsymbol{P}) + (1-\eta) St(\boldsymbol{x};\ \boldsymbol{m},\ \boldsymbol{P},\ \nu) \tag{6}$$

$$p(\eta) = Be(\eta;\ e,\ 1-e) \tag{7}$$

where $St(\ :;\ \boldsymbol{m},\ \boldsymbol{P},\ \nu)$ denotes the Student's t density function, the parameters contain the mean $\boldsymbol{m}$, scale matrix $\boldsymbol{P}$ and degree of freedom (DOF) parameter ν. Assuming that η is unknown and obeys the beta distribution $Be(\ :;\ e,\ 1-e)$, e refers to the prior shape parameter. In order to model the noise outlier, the Bernoulli random variable ε is introduced as the conjugate prior distribution of the beta distribution

$$p(\varepsilon \mid \eta) = \eta^{\varepsilon}(1-\eta)^{1-\varepsilon} \quad \text{s.t.}\ \ \varepsilon \in \{0,1\} \tag{8}$$

The auxiliary variable λ is introduced to transform the mixed Gaussian-Student's t distribution into the following hierarchical Gaussian form[16]

$$p(\boldsymbol{x} \mid \varepsilon, \lambda) = [N(\boldsymbol{x};\ \boldsymbol{m},\ \boldsymbol{p})]^{\varepsilon} [N(\boldsymbol{x};\ \boldsymbol{m},\ \boldsymbol{p}/\lambda)]^{(1-\varepsilon)} \tag{9}$$

$$p(\lambda) = G\left(\lambda;\ \frac{v}{2},\ \frac{v}{2}\right) \tag{10}$$

where $G\left(\ :;\ \frac{v}{2},\ \frac{v}{2}\right)$ is the gamma density function with DOF parameter v.

2.2 GST-vbPHD filting

2.2.1 Predict

Combined with the model constructed by Eq. (9) and Eq. (10), the implementation of GST-vbPHD is derived for linear multi-target systems. According to Bayesian probability theory, a beta distribution is selected as the conjugate prior distribution of unknown mixing probability η_k[16]. The likelihood distribution of the η_k is expressed as Bernoulli distribution, and Bernoulli component ε_k is introduced to select Gaussian or Student's t distribution. It is well known that the Student's t distribution can be expressed as the product of gamma distribution and Gaussian distribution after introducing auxiliary variables λ_k.

Suppose the augmented state $\boldsymbol{\vartheta}_k$ of one single target, which contains one single target state and a set of parameters for constructing the distribution, can be represented as $\boldsymbol{\vartheta}_k \triangleq (\boldsymbol{x}_k, \eta_k, \varepsilon_k, \boldsymbol{\lambda}_k)$, where η_k, ε_k and $\boldsymbol{\lambda}_k$ respectively refer to the mixed probability, Bernoulli random variables and auxiliary variables, and they are mutually independent of $\boldsymbol{x}_k$. The components of the predicted intensity are the same as Eq. (1), so the mixed distribution model of joint probability density is expressed as

$$\boldsymbol{v}_{k|k-1}(\boldsymbol{\vartheta}_{k|k-1}) = \boldsymbol{v}_{S,k|k-1}(\boldsymbol{\vartheta}_{S,k|k-1}) + \boldsymbol{v}_{\beta,k|k-1}(\boldsymbol{\vartheta}_{\beta,k|k-1}) + \boldsymbol{\delta}_k(\boldsymbol{\vartheta}_k) \tag{11}$$

$$\boldsymbol{v}_{S,k|k-1}(\boldsymbol{\vartheta}_{S,k|k-1}) = P_{S,k} \sum_{j=1}^{J_{k-1}} w_{k-1}^{(j)} [N(\boldsymbol{x};\ \boldsymbol{m}_{S,k|k-1}^{(j)},\ \boldsymbol{P}_{S,k|k-1}^{(j)})] \tag{12}$$

where $P_{S,k}$ and $\boldsymbol{v}_{S,k|k-1}(\boldsymbol{\vartheta}_{S,k|k-1})$ are the target survival probability and the intensity of the target surviving to time k, respectively. To obtain the priori model state information, by one-step nominal prediction state vector $\boldsymbol{m}_{S,k|k-1}^{(j)}$ and the corresponding error covariance matrix $\boldsymbol{P}_{S,k|k-1}^{(j)}$, the predicted probability density functions $\boldsymbol{m}_{S,k|k-1}^{*(j)}$ and $\boldsymbol{P}_{S,k|k-1}^{*(j)}$ are approximated as

$$\boldsymbol{m}_{S,k|k-1}^{(j)} \approx \boldsymbol{m}_{S,k|k-1}^{*(j)} = \boldsymbol{F}_{k-1} \boldsymbol{m}_{k-1}^{(j)} \tag{13}$$

$$\boldsymbol{P}_{S,k|k-1}^{(j)} \approx \boldsymbol{P}_{S,k|k-1}^{*(j)} = \boldsymbol{F}_k \boldsymbol{P}_{k-1}^{(j)} \boldsymbol{F}_k^{\mathrm{T}} + \boldsymbol{Q}_{k-1} \tag{14}$$

where $\boldsymbol{Q}_{k-1}$ represents the nominal process covariance matrix. And $\boldsymbol{v}_{\beta,k|k-1}(\boldsymbol{\vartheta}_{\beta,k|k-1})$ is given by

$$\boldsymbol{v}_{\beta,k|k-1}(\boldsymbol{\vartheta}_{\beta,k|k-1}) = \sum_{j=1}^{J_{k-1}}\sum_{l=1}^{J_{\beta,k}} w_{k-1}^{(j)} w_{\beta,k}^{(l)} [N(\boldsymbol{x};\boldsymbol{m}_{\beta,k|k-1}^{(j,l)},\boldsymbol{P}_{\beta,k|k-1}^{(j,l)})] \tag{15}$$

$$\boldsymbol{m}_{\beta,k|k-1}^{(j,l)} \approx \boldsymbol{m}_{\beta,k|k-1}^{*(j,l)} = \boldsymbol{F}_{\beta,k-1}^{(l)}\boldsymbol{m}_{k-1}^{(j)} + \boldsymbol{d}_{\beta,k-1}^{(l)} \tag{16}$$

$$\boldsymbol{P}_{\beta,k|k-1}^{(j,l)} \approx \boldsymbol{P}_{\beta,k|k-1}^{*(j,l)} = \boldsymbol{F}_{\beta,k-1}^{(l)}\boldsymbol{P}_{\beta,k-1}^{(j)}(\boldsymbol{F}_{\beta,k-1}^{(l)})^{\mathrm{T}} + \boldsymbol{Q}_{\beta,k-1}^{(l)} \tag{17}$$

The birth target intensity can be written as

$$\boldsymbol{\delta}_k(\boldsymbol{\vartheta}_k) = \sum_{i=1}^{J_{\delta,k}} w_k^{(i)}[N(\boldsymbol{x};\boldsymbol{m}_{\delta,k}^{(i)},\boldsymbol{P}_{\delta,k}^{(i)})] \tag{18}$$

where, $\boldsymbol{v}_{\beta,k|k-1}(\boldsymbol{\vartheta}_{\beta,k|k-1})$ is the spawning target intensity and $\boldsymbol{\delta}_k(\boldsymbol{\vartheta}_k)$ is birth target intensity, both of which have the same mixed distribution form with $\boldsymbol{v}_{S,k|k-1}(\boldsymbol{\vartheta}_{S,k|k-1})$.

2.2.2 Variational Bayesian approximation

Considering the measurement noise and process noise both exist outliers, the state one-step prediction PDF and the measurement likelihood PDF are modeled by the above model at the same time, and the same distribution type is used in the modeling process. Since the parameters in the mixed distribution are unknown, the posterior PDF $p(\boldsymbol{\vartheta}_k \mid \boldsymbol{z}_{1:k})$ is difficult to obtain an analytical solution. To simplify calculations, the VB method[17] is adopted to solve the approximate distribution

$$p(\boldsymbol{x}_k,\eta_k,\varepsilon_k,\lambda_k \mid z_{1:k}) \approx q(\boldsymbol{x}_k)q(\eta_k)q(\varepsilon_k)q(\lambda_k) \tag{19}$$

where $q(\cdot)$ is the approximate posterior PDF of $p(\cdot)$. After minimizing the KLD scatter[18] between the approximate posterior PDF and the true posterior PDF, there is

$$\{q(\boldsymbol{x}_k),q(\eta_k),q(\varepsilon_k),q(\lambda_k)\} = \operatorname{argmin} KLD\,(q(\boldsymbol{x}_k)q(\eta_k)q(\varepsilon_k)q(\lambda_k) \parallel p(\boldsymbol{x}_k,\eta_k,\varepsilon_k,\lambda_k \mid \boldsymbol{z}_{1:k})) \tag{20}$$

With the VB approximation strategy, the posterior PDF satisfies the following equation

$$\log q(\theta) = E_{\boldsymbol{\vartheta}_k^{(-\theta)}}[\log p(\boldsymbol{\vartheta}_k,\boldsymbol{z}_{1:k})] + c_\theta \tag{21}$$

where $E(\cdot)$ is the expectation operation and $\log(\cdot)$ is logarithmic operation. θ is any element in the set $\boldsymbol{\vartheta}_k$, and $\boldsymbol{\vartheta}_k^{(-\theta)}$ is all elements in the set of $\boldsymbol{\vartheta}_k$ except θ. c_θ is the constant associated with θ[19].

The posterior PDF is estimated for joint fixed-point iterations using the VB approximation, which can be used to obtain

$$q^{(n+1)(j)}(\boldsymbol{x}_k) = N(\boldsymbol{x}_k;\boldsymbol{m}_{k|k}^{(n+1)(j)},\boldsymbol{P}_{k|k}^{(n+1)(j)}) \tag{22}$$

$$q^{(n+1)(j)}(\eta_{i,k}) = Be(\eta_{i,k};e_{i,k}^{(n+1)(j)},t_{i,k}^{(n+1)(j)}) \tag{23}$$

$$q^{(n+1)(j)}(\lambda_{i,k}) = G(\lambda_{i,k};w_{i,k}^{(n+1)(j)},h_{i,k}^{(n+1)(j)}) \tag{24}$$

where $i = [1,2]$. Since outliers are assumed to appear in the measurement noise, it may also appear in the process noise. Two sets of parameters are used to correct measurement noise $\boldsymbol{R}_k$ and state prediction error covariance $\boldsymbol{P}_{k|k-1}^{(j)}$ respectively. When $i = 1$, the $\boldsymbol{P}_{k|k-1}^{(j)}$ is corrected, correction for $\boldsymbol{R}_k$ at $i = 2$.

According to Eq. (17), the PDFs of $\eta_{i,k}$ are updated as beta distributions by using the mixture distribution model of the state prediction PDF and the measurement likelihood PDF, where the shape parameters are updated as follow

$$e_{1,k}^{(n+1)(j)} = e_1 + E^{(n+1)}[\varepsilon_{1,k}] \tag{25}$$

$$t_{1,k}^{(n+1)(j)} = 2 - e_1 - E^{(n+1)}[\varepsilon_{1,k}] \tag{26}$$

$$e_{2,k}^{(n+1)(j)} = e_2 + E^{(n+1)}[\varepsilon_{2,k}] \tag{27}$$

$$t_{2,k}^{(n+1)(j)} = 2 - e_2 - E^{(n+1)}[\varepsilon_{2,k}] \tag{28}$$

and the shape parameters and the rate parameters of the gamma distributions are updated as

$$w_{1,k}^{(n+1)(j)} = 0.5n(1 - E^{(n+1)}[\varepsilon_{1,k}]) + 0.5\omega_1 \tag{29}$$

$$w_{2,k}^{(n+1)(j)} = 0.5m(1 - E^{(n+1)}[\varepsilon_{2,k}]) + 0.5\omega_2 \tag{30}$$

$$h_{1,k}^{(n+1)(j)} = 0.5\operatorname{tr}(\boldsymbol{A}_k^{(n+1)}(\boldsymbol{P}_{k|k-1}^{(j)})^{-1})(1 - E^{(n+1)}[\varepsilon_{1,k}]) + 0.5\omega_1 \tag{31}$$

$$h_{2,k}^{(n+1)(j)} = 0.5\operatorname{tr}(\boldsymbol{B}_k^{(n+1)}\boldsymbol{R}_k^{-1})(1 - E^{(n+1)}[\varepsilon_{2,k}]) + 0.5\omega_2 \tag{32}$$

When the Bernoulli parameters $\varepsilon_{1,k}$ and $\varepsilon_{2,k}$ take 0 and 1 respectively, the probabilities are given as

$$Pr^{(n+1)(j)}(\varepsilon_{1,k} = 1) = \Delta_1^{(n+1)}\exp\{E^{(n)}[\log\eta_{1,k}] - 0.5\operatorname{tr}(\boldsymbol{A}_k^{(n+1)}(\boldsymbol{P}_{k|k-1}^{(j)})^{-1})\} \tag{33}$$

$$Pr^{(n+1)(j)}(\varepsilon_{2,k} = 1) = \Delta_2^{(n+1)}\exp\{E^{(n)}[\log\eta_{2,k}] - 0.5\operatorname{tr}(\boldsymbol{B}_k^{(n+1)}\boldsymbol{R}_k^{-1})\} \tag{34}$$

$$Pr^{(n+1)(j)}(\varepsilon_{1,k} = 0) = \Delta_1^{(n+1)}\exp\{E^{(n)}[\log(1-\eta_{1,k})] + 0.5n\,E^{(n)}[\log\lambda_{1,k}] - 0.5\,E^{(n)}[\lambda_{1,k}]\operatorname{tr}(\boldsymbol{A}_k^{(n+1)}(\boldsymbol{P}_{k|k-1}^{(j)})^{-1})\} \tag{35}$$

$$Pr^{(n+1)(j)}(\varepsilon_{2,k} = 0) = \Delta_2^{(n+1)}\exp\{E^{(n)}[\log(1-\eta_{2,k})] + 0.5m\,E^{(n)}[\log\lambda_{2,k}] - 0.5\,E^{(n)}[\lambda_{2,k}]\operatorname{tr}(\boldsymbol{B}_k^{(n+1)}\boldsymbol{R}_k^{-1})\} \tag{36}$$

where $\Delta_1^{(n+1)}$ and $\Delta_2^{(n+1)}$ are normalization constants and $\operatorname{tr}(\cdot)$ is trace operation on matrix. $\boldsymbol{m}$, $\boldsymbol{n}$ are the dimension of the state vector and the measurement vector, respectively. $\boldsymbol{A}_k^{(n+1)}$ and $\boldsymbol{B}_k^{(n+1)}$ are the auxiliary parameters.

$$\boldsymbol{A}_k^{(n+1)} = \boldsymbol{P}_k^{(j)(n+1)} + (\boldsymbol{m}_k^{(j)(n+1)} - \boldsymbol{m}_{k|k-1}^{(j)})(\boldsymbol{m}_k^{(j)(n+1)} - \boldsymbol{m}_{k|k-1}^{(j)})^{\mathrm{T}} \tag{37}$$

$$\boldsymbol{B}_k^{(n+1)} = \boldsymbol{H}_k\boldsymbol{P}_k^{(j)(n+1)}\boldsymbol{H}_k^{\mathrm{T}} + (\boldsymbol{z}_k - \boldsymbol{H}_k\boldsymbol{m}_k^{(j)(n+1)})$$

$$(\boldsymbol{z}_k - \boldsymbol{H}_k \boldsymbol{m}_k^{(j)(n+1)})^{\mathrm{T}} \tag{38}$$

The expectations of the mixture probability parameters $\log\eta_{i,k}$ and $\log(1-\eta_{i,k})$ are

$$E^{n+1}[\log\eta_{i,k}] = \psi(e^{(n+1)}) - \psi(e^{(n+1)} + t^{(n+1)}) \tag{39}$$

$$E^{n+1}[\log(1-\eta_{i,k})] = \psi(t^{(n+1)}) - \psi(e^{(n+1)} + t^{(n+1)}) \tag{40}$$

At time k, $e_{1,k}^{(n+1)(j)}$, $t_{1,k}^{(n+1)(j)}$ and $\varepsilon_{1,k}$ denote the parameters of the state prediction PDF for the j-th target at the $n+1$ variational iteration. $e_{2,k}^{(n+1)(j)}$, $t_{2,k}^{(n+1)(j)}$ and $\varepsilon_{2,k}$ denote the parameters of the measurement likelihood PDF under same conditions.

2.2.3 Update

At time k, the target posterior intensity function is updated with

$$\boldsymbol{v}_k(\boldsymbol{\vartheta}_k) = (1-p_{D,k})\boldsymbol{v}_{k|k-1}(\boldsymbol{\vartheta}_{k|k-1}) + \sum_{z\in Z_k}\boldsymbol{v}_{D,k}(\boldsymbol{\vartheta}_{D,k};\boldsymbol{z}) \tag{41}$$

where

$$\boldsymbol{v}_{D,k}(\boldsymbol{\vartheta}_k;\boldsymbol{z}) = \sum_{j=1}^{J_{k-1}} w_k^{(j)}(\boldsymbol{z})[N(\boldsymbol{x}_k;\boldsymbol{m}_k^{(j)}(\boldsymbol{z}),\boldsymbol{P}_k^{(j)})^{\varepsilon_{1,k}} N(\boldsymbol{x}_k;\boldsymbol{m}_k^{(j)}(\boldsymbol{z}),\boldsymbol{P}_k^{(j)}/\lambda_{1,k})^{(1-\varepsilon_{1,k})}]G\left(\lambda_{1,k};\frac{\omega_1}{2},\frac{\omega_1}{2}\right) (\eta_{1,k}^{\varepsilon_{1,k}})(1-\eta_{1,k}^{\varepsilon_{1,k}})Be(\eta_{1,k};e_1,1-e_1) \tag{42}$$

In addition to integrating VB inference into the update step, which is different from the traditional PHD framework, the calculation of target weight is also improved. The corrected state prediction error covariance $\tilde{\boldsymbol{P}}_{k|k-1}^{(j)(n)}$ and the corrected measurement noise covariance $\tilde{\boldsymbol{R}}_k^{(j)(n)}$ are used to update the target weights, which improves the accuracy of calculating Gaussian Student's t mixture components and reduces the error caused by the interference of colored noise.

$$w_k^{(j)(n+1)}(\boldsymbol{z}) = \frac{P_{D,k}w_{k|k-1}^{(j)}q_k^{(j)(n+1)}(\boldsymbol{z})}{\kappa_k(\boldsymbol{z}) + P_{D,k}\sum_{l=1}^{J_{k|k-1}} w_{k|k-1}^{(l)}q_k^{(l)(n+1)}(\boldsymbol{z})} \tag{43}$$

$$q_k^{(j)(n+1)}(\boldsymbol{z}) = N(\boldsymbol{z}_k;\boldsymbol{H}_k\boldsymbol{m}_{k|k-1}^{(j)},\tilde{\boldsymbol{R}}_k^{(j)(n)} + \boldsymbol{H}_k\tilde{\boldsymbol{P}}_{k|k-1}^{(j)(n)}\boldsymbol{H}_k^{\mathrm{T}}) \tag{44}$$

In the standard Kalman filtering measurement update framework, the target state estimates $\boldsymbol{m}_k^{(j)(n+1)}$ and its corresponding state estimation error covariance matrix $\boldsymbol{P}_k^{(j)(n+1)}$ are updated iteratively

$$\boldsymbol{K}_k^{(j)(n+1)} = \tilde{\boldsymbol{P}}_{k|k-1}^{(j)(n)}\boldsymbol{H}_k^{\mathrm{T}}(\boldsymbol{H}_k\tilde{\boldsymbol{P}}_{k|k-1}^{(j)(n)}\boldsymbol{H}_k^{\mathrm{T}} + \tilde{\boldsymbol{R}}_k^{(j)(n)})^{-1} \tag{45}$$

$$\boldsymbol{m}_k^{(j)(n+1)} = \boldsymbol{m}_{k|k-1}^{(j)} + \boldsymbol{K}_k^{(j)(n+1)}(\boldsymbol{z}_k - \boldsymbol{H}_k\boldsymbol{m}_{k|k-1}^{(j)}) \tag{46}$$

$$\boldsymbol{P}_k^{(j)(n+1)} = [\boldsymbol{I} - \boldsymbol{K}_k^{(j)(n+1)}\boldsymbol{H}_k]\tilde{\boldsymbol{P}}_{k|k-1}^{(j)(n)} \tag{47}$$

where $\boldsymbol{K}_k^{(j)(n+1)}$ is the Kalman filter gain. With the nominal state error covariance and the nominal measurement noise covariance, $\tilde{\boldsymbol{P}}_{k|k-1}^{(j)(n)}$ and $\tilde{\boldsymbol{R}}_k^{(j)(n)}$ are respectively used to correct as follows

$$\tilde{\boldsymbol{P}}_{k|k-1}^{(j)(n)} = \frac{\boldsymbol{P}_{k|k-1}^{(j)(0)}}{E^{(n)}[\varepsilon_{1,k}] + (1-E^{(n)}[\varepsilon_{1,k}])E^{(n)}[\lambda_{1,k}]} \tag{48}$$

$$\tilde{\boldsymbol{R}}_k^{(j)(n)} = \frac{\boldsymbol{R}_k}{E^{(n)}[\varepsilon_{2,k}] + (1-E^{(n)}[\varepsilon_{2,k}])E^{(n)}[\lambda_{2,k}]} \tag{49}$$

To achieve the expectation solutions for the parameters $\varepsilon_{1,k}$ and $\varepsilon_{2,k}$, $\lambda_{1,k}$ and $\lambda_{2,k}$, variational fixed-point iteration is performed as follows

$$E^{(n+1)}[\varepsilon_{i,k}] = \frac{Pr^{(n+1)}(\varepsilon_{i,k}=1)}{Pr^{(n+1)}(\varepsilon_{i,k}=1) + Pr^{(n+1)}(\varepsilon_{i,k}=0)} \tag{50}$$

$$E^{(n+1)}[\lambda_{i,k}] = \frac{w^{(n+1)}}{h^{(n+1)}} \tag{51}$$

$$E^{(n+1)}[\log\lambda_{i,k}] = \psi(w^{(n+1)}) - \log h^{(n+1)} \tag{52}$$

where $n\in[1,N]$ and N is the maximum number of variational iterations. If $(\boldsymbol{m}_k^{(j)(n+1)} - \boldsymbol{m}_k^{(j)(n)}) \leqslant \varepsilon$ then stop the variational iteration, else the loop continues. The updated $\boldsymbol{m}_k^{(j)}$, $\boldsymbol{P}_k^{(j)}$ and $w_k^{(j)}(\boldsymbol{z})$ are output, and all the outputs values are used as the input values of the clipping step and merging step. As the sampling time grows, the Gaussian Student's t mixture components increases exponentially, so each update time requires clipping of the mixture terms with existence probability below the threshold[20]. Furthermore, the remaining terms with merging distance less than the threshold are output. Finally, the target state is extracted.

To summarize the update steps visualization of the GST-vbPHD algorithm, the pseudocode of implementation flow is given in Algorithm 1.

Algorithm 1 The variational iteration process for each Gaussian Student's t mixture components

Inputs: $\boldsymbol{m}_{k|k-1}^{(j)}$, $\boldsymbol{P}_{k|k-1}^{(j)}$, $w_{k|k-1}^{(j)}$, $\boldsymbol{H}_k$, $\boldsymbol{R}_k$, z_k, $P_{D,k}$, r, e_1, e_2, ω_1, ω_2, N

1. Initialization: $E^{(0)}[\gamma]=1$, $E^{(0)}[\log\gamma]=0$, $E^{(0)}[\varepsilon]=1$,
   $E^{(0)}[\log(1-\sigma)] = n\psi(1-e) - \psi(1)$
   $E^{(0)}[\log\sigma] = \psi(e) - \psi(1)$,
2. $\boldsymbol{m}_k^{(j)(0)} = \boldsymbol{m}_{k|k-1}^{(j)}$, $\boldsymbol{P}_k^{(j)(0)} = \boldsymbol{P}_{k|k-1}^{(j)}$
3. for $n=0:N-1$ do
4. Calculate $\tilde{\boldsymbol{P}}_{k|k-1}^{(j)(n)}$ and $\tilde{\boldsymbol{R}}_k^{(j)(n)}$ using Eqs(48) and (49)
5. Calculate $\boldsymbol{m}_k^{(j)(n+1)}$ and $\boldsymbol{P}_k^{(j)(n+1)}$ using Eqs(45) – (47)
6. Calculate $w_k^{(j)(n+1)}(\boldsymbol{z})$ using Eqs(43) – (44)
7. Calculate $\boldsymbol{A}_k^{(n+1)}$, $\boldsymbol{B}_k^{(n+1)}$ using Eqs(39) – (40) Update $q^{n+1}(\varepsilon_{1,k})$, $q^{n+1}(\varepsilon_{2,k})$ as Bernoulli distributions

8. Calculate $Pr^{n+1}(\varepsilon_{1,k}=1)$, $Pr^{n+1}(\varepsilon_{1,k}=0)$ and $Pr^{n+1}(\varepsilon_{2,k}=1)$, $Pr^{n+1}(\varepsilon_{2,k}=0)$ using Eqs(33) - (36)
9. Calculate $E^{n+1}[\varepsilon_{1,k}]$, $E^{n+1}[\varepsilon_{2,k}]$ using Eq. (50)
   Update $q^{n+1}(\gamma_{1,k})$, $q^{n+1}(\gamma_{2,k})$ as Gamma distributions
10. Calculate $w_{1,k}^{n+1}$, $h_{1,k}^{n+1}$ and $w_{2,k}^{n+1}$, $h_{2,k}^{n+1}$ using Eqs(29) - (32)
11. Calculate $E^{n+1}[\gamma_{1,k}]$, $E^{n+1}[\gamma_{2,k}]$ and $E^{n+1}[\log\gamma_{1,k}]$, $E^{n+1}[\log\gamma_{2,k}]$ using Eqs(51) - (52)
   Update $q^{n+1}(\sigma_{1,k})$, $q^{n+1}(\sigma_{2,k})$ as Bate distributions
12. Calculate $e_{1,k}^{n+1}$, $t_{1,k}^{n+1}$ and $e_{2,k}^{n+1}$, $t_{2,k}^{n+1}$ using Eqs(25) - (28)
13. Calculate $E^{n+1}[\log\sigma_{1,k}]$, $E^{n+1}[\log\sigma_{2,k}]$ and $E^{n+1}[\log(1-\sigma_{1,k})]$, $E^{n+1}[\log(1-\sigma_{2,k})]$ using Eqs(39) - (40)
14. if $\boldsymbol{m}_k^{(j+1)(n+1)} - \boldsymbol{m}_k^{(j+1)(n)} \leqslant \varepsilon$ then
15. Stop the iteration
16. end if
17. end for

Outputs: $\boldsymbol{m}_k^{(j)}$, $\boldsymbol{P}_k^{(j)}$ and $w_k^{(j)}$

3 Simulation results and analysis

3.1 Scenario design

To verify the tracking performance of the proposed algorithm, two simulation scenarios are designed in the 2-D plane, i. e. the scenario of the measurement noise with outliers and the scenario of outliers in both process noise and measurement noise. In addition, the different probabilities of generating outliers are compared in the two scenarios. For comparison, the tracking performance of the Gaussian mixture PHD filter (GM-PHD), the robust Student's t based PHD filter (RST-PHD) and the GST-vbPHD are employed.

Assuming that there are four targets in the surveillance range, they are present at time [1 8 12 26](s) until time [20 25 25 40](s) in turn disappears, with uniform motion during the survival period. The real trajectories of all targets are plotted in Fig. 1.

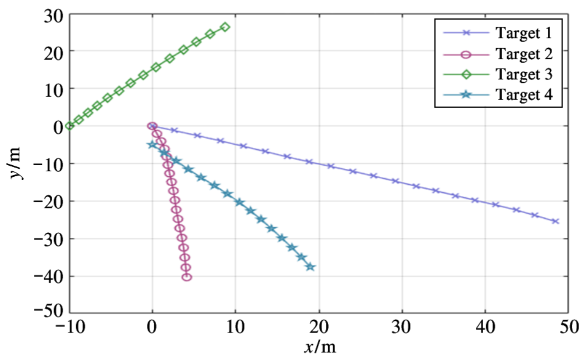

Fig. 1 The true trajectory of multiple targets

3.2 Simulation parameters

A total of 40 steps are running in the simulation process, and the simulation results are the average after 300 Monte Carlo (MC) trials. What is more, the objective survival probability $P_{S,k} = 0.99$, detection probability $P_{D,k} = 0.98$ and the clutter rate $\lambda = 3$. The state and measurement equations of the targets are modeled as the following form

$$\boldsymbol{x}_k = \boldsymbol{F}\boldsymbol{x}_{k-1} + \boldsymbol{Q}_k$$
$$\boldsymbol{z}_k = \boldsymbol{H}\boldsymbol{x}_k + \boldsymbol{R}_k$$

where the state transfer matrix $\boldsymbol{F} = I_2 \otimes \begin{pmatrix} 1 & T \\ 0 & 1 \end{pmatrix}$ and the observation matrix $\boldsymbol{H} = I_2 \otimes (1 \quad 0)$, $\otimes$ represent Cartesian product operators.

Same as RST-PHD, DOF parameter and variational iteration number are all set to be 10 in proposed algorithm. The clipping threshold and the merging threshold for the target mixture are set as $T = 10^{-5}$ and $U = 4$, respectively. The performance of the three filters is compared by using the two types of performance metrics that are OPSA error and target number error. The OSPA distance has two subsets of X and Z with dimensions m and n respectively.

$$d_p^c(X, Z) = \left(\frac{1}{n}\left(\min_{\pi\in\prod_n}\sum_{i=1}^{m} d^{(c)}(x_i, z_{\pi(i)})^p + c^p(n-m)\right)\right)^{1/p}$$

where the distance sensitivity parameter satisfies $1 \leq P \leq \infty$, $m \leq n$, and the association sensitivity parameter is usable at $c > 0$. In this paper, we choose $p = 2$ and $c = 100$. $\prod_n$ refers to all permutations in $\{1, \cdots, n\}$. If $m > n$, then

$$\bar{d}_p^{(c)}(X, Z) = \bar{d}_p^{(c)}(Z, X)$$

3.3 Results and analysis

3.3.1 Scenario 1

To observe the performance of MTT at measurement outliers existing, the measurement noise covariance is constructed with outlier according to Ref. [21].

$$\boldsymbol{v}_k \sim \begin{cases} N(0, \boldsymbol{R}_0) & p_1 \\ N(0, 50\boldsymbol{R}_0) & 1-p_1 \end{cases}$$

where, $p1$ is the probability of the measurement noise without outliers and is set to be in a range of 5 - 30 s during the multi-target motion.

Fig. 2 shows the OSPA distance errors for the three filters with probability $p1 = 0.98$. Due to the measurement outliers, it can be observed that the GM-PHD has significantly inferior tracking performance to the other two filters. Specifically, when there are outliers in the measurement, because of the light-weight tail property of Gaussian distribution, the weight of Gaussian components tends to be a small value or even zero in some cases, which leads to a larger OSPA distance. In contrast, although RST-PHD takes into account the heavy-tailed feature of noise, it uses a fixed Student's t distribution for modeling, which lacks robustness to randomly occurring outliers. The OSPA distance curve of GST-vbPHD is lower than that of GM-PHD and RST-PHD, which demonstrates that the tracking performance of GST-vbPHD surpasses the other two algorithms. Due to the random characteristic of the measurement noise outliers, the noise cannot always remain in a heavy-tailed or Gaussian distribution state. GST-vbPHD employs the model with mixture distribution to better estimate the target weights, which helps to track the target without loss. The results of three algorithms for estimating the number of targets are given in Fig. 3, and it can be seen that GST-vbPHD significantly outperforms GM-PHD and RST-PHD.

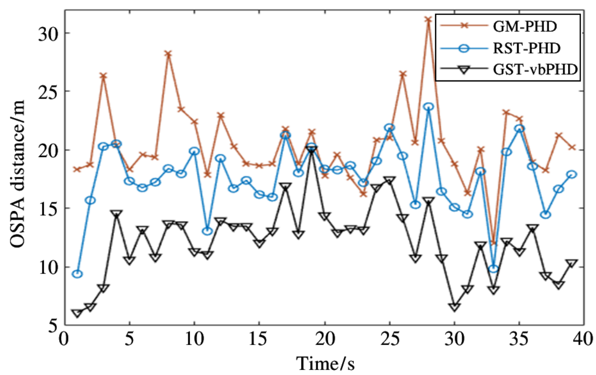

Fig. 2 The OSPA distance

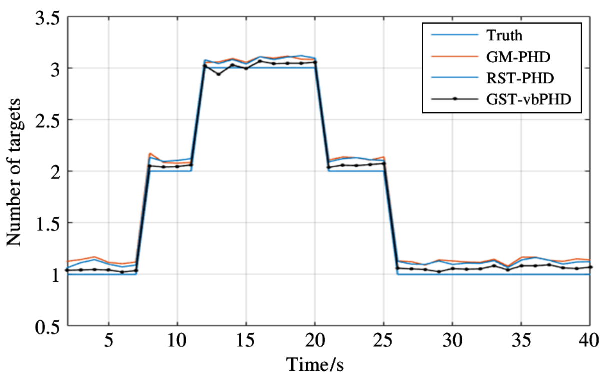

Fig. 3 The number of targets

To further analyze the impact of the probability p_1 on the filter performance, the statistical analysis on the average OSPA distance is shown in Fig. 4. When the probability of the measurement noise without outliers is 0.98, 0.96, 0.94, 0.92 and 0.9 respectively, OSPA average distance all decrease with increasing of the probability of the measurement noise, and that is because the effect of measurement noise outliers on the system significantly weakens. GST-vbPHD has a lower OSPA average distance than GM-PHD and RST-PHD, and has better tracking accuracy under lighter-tailed measurements or even heavy-tailed measurements.

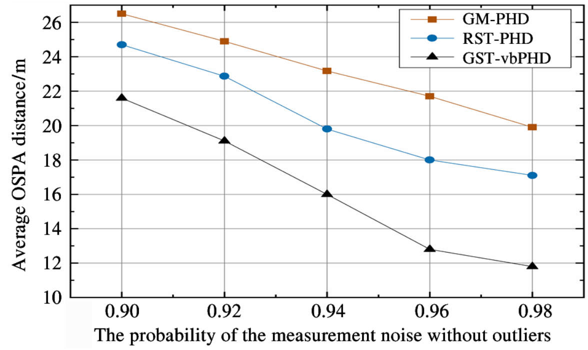

Fig. 4 Average OSPA with different p_1

3.3.2 Scenario 2

To evaluate the performance of MTT at outlier both existing in process noise and measurement noise, a new experiment Scenario 2 is constructed. The measurement noise outliers can be generated according to Scenario 1, and the process noise covariance with outliers is shown as follows.

$$\boldsymbol{w}_k \sim \begin{cases} N(0, \boldsymbol{Q}_0) & p_2 \\ N(0, 10\boldsymbol{Q}_0) & 1 - p_2 \end{cases}$$

where p_2 is the probability of the process noise without outliers. Assume that the time period of outliers in the noise is the same as Scenario 1.

For outliers of the process noise and the measurement noise with the same probability, Fig. 5 shows the OSPA distance comparison of three filters. From Fig. 5 and Fig. 6, it can be found that GST-vbPHD shows a better tracking result for both the tracking accuracy and the estimation of target number. The process noise outliers may be induced by target maneuvers, while the GM-PHD filter cannot capture the target due to the light tail of the Gaussian distribution, and RST-PHD lacks adaptability to random outliers. The proposed algorithm utilizes the mixture distribution model of noise

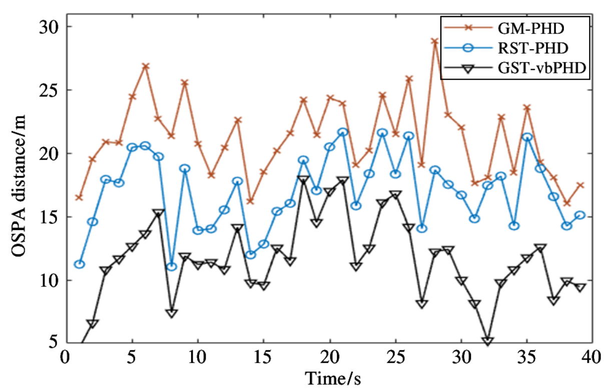

Fig. 5 The OSPA distance

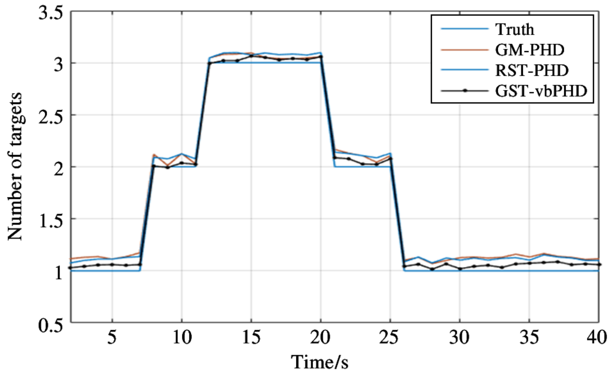

Fig. 6 The number of targets

to correct state error covariance, effectively eliminates the adverse effects induced by process noise outliers. Overall, if there are outliers in both process noise and measurement noise, the GST-vbPHD can achieve reliable and effective performance in MTT.

In order to deeply analyze the tracking performance of the filters under different probabilities of outliers, the additional experiments are executed. First, p_2 is fixed, while p_1 is 0.9, 0.92, 0.94, 0.96, and 0.98 respectively. After that, the average OSPA distance is given in Fig. 7. On the contrary, when p_1 is fixed, p_2 changes and the corresponding simulation results are shown in Fig. 8. The average OSPA distance of the three filters gradually decreases, which means that the poor tracking performance with the occurrence probability of outliers increases. In Fig. 8, the average OSPA distance of the three filters decreases less obviously than that in Fig. 7. The results can be attributed to the different multiples of setting outliers in the dynamic model, and the filter is more sensitive to the process noise. As shown in Fig. 7 and Fig. 8, the proposed algorithm achieves relatively stable tracking performance for different probabilities of outliers in process noise and measurement noise.

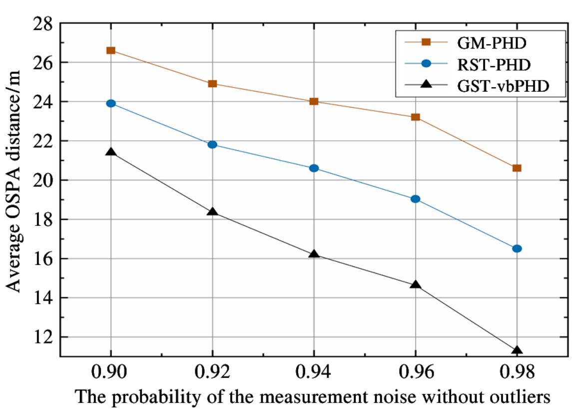

Fig. 7 Average OSPA with different p_1

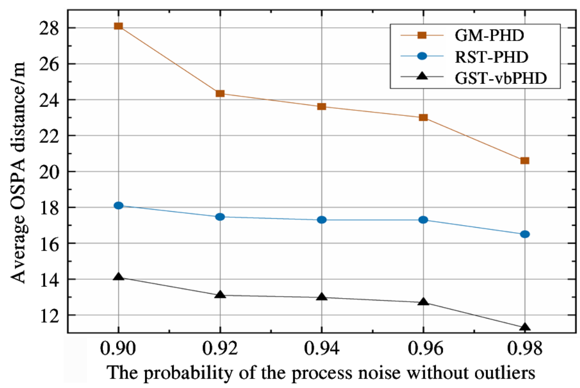

Fig. 8 Average OSPA with different p_2

The computation time of the RST-PHD and the GST-vbPHD in this paper is 1.4 s and 3.5 s respectively when both the variational approximation iterations are used. The proposed algorithm not only improves the accuracy of tracking and estimating the number of targets, but also increases the operation time. This is because we consider that both process noise and measurement noise may have outliers. Two sets of parameters are used to modify the measurement noise covariance and state error covariance respectively, and participate in the variational iteration. The contrast algorithm only considers the heavy tail characteristics of noise outliers, but ignores the associate the nonstationarity. It simply uses student t distribution to model the noise. Therefore, the proposed algorithm in this paper does not perform well on the evaluation index of operation time.

4 Conclusions

In this paper, a new Gaussian-Student's t mixture distribution PHD robust filtering algorithm is proposed based on variational Bayesian inference, which models the one-step state prediction PDF and the measurement likelihood PDF as the hierarchical Gaussian forms. Concretely, the hierarchical Gaussian form is employed to correct state error covariance matrix and measurement noise covariance matrix, eliminating the adverse effects of process noise and measurement noise both with outliers on the tracking performance. In addition, the parameters in the mixed distribution term are iteratively optimized by variational inference to obtain the target posterior probability density. The simulation results show that the proposed algorithm can achieve competitive performance with the traditional Gaussian hybrid PHD filter and the Student's t PHD filter on tracking accuracy in MTT. Future work will focus on how to construct a hierarchical Gaussian noise distribution for nonlinear systems to effectively solve the influence of noise outliers.

References

[1] AINSLEIGH P L, LUGINBUHL T E, WILLETT P K. A sequential target existence statistic for joint probabilistic data association[J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2021,57 (1): 371-381

[2] BLACKMAN S S. Multiple hypothesis tracking for multiple target tracking[J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2004,19(1):5-8

[3] WIENEKE M, KOCH W. A PMHT approach for extended objects and object groups[J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2012, 48(3): 2349-2370

[4] VO B T, VO B N. Labeled random finite sets and multi-object conjugate priors[J]. *IEEE Transactions on Signal Processing*, 2013,61(13): 3460-3475

[5] VO B N, SINGH S, DOUCET A. Sequential Monte Carlo methods for multi-target filtering with random finite sets [J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2005,41(4):1224-45

[6] DANIYAN A, GONG Y, LAMBOTHARAN S, et al. Kalman-gain aided particle PHD filter for multitarget tracking[J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2017,53(5):2251-2265

[7] VO B N, MA W K. The Gaussian mixture probability hypothesis density filter[J]. *IEEE Transactions on Signal Process*, 2006,54(11):4091-104

[8] GRANSTRM K, LUNDQUIST C, ORGUNER U. A Gaussian mixture PHD filter for extended target tracking [C] // International Conference on Information Fusion, Edinburgh, UK, 2010:1-8

[9] LIU B, CHENG S, SHI Y. Improving estimation of distribution algorithms with heavy-tailed student's t distributions[C]//International Conference on Advanced Computational Intelligence, Xiamen, China, 2018:11-6

[10] LIU Z, CHEN S, WU H, et al. A Student's t mixture probability hypothesis density filter for multi-target tracking with outliers[J]. *Sensors*, 2018,18(4): 1095

[11] PAN Q, HU Y M, LAN H, et al. Information fusion progress: joint optimization based on variational Bayesian theory[J]. *Acta Automatica Sinica*, 2019,45(7):1207-1223

[12] LI P, XU C, WANG W, et al. Robust Student's T distribution based PHD/CPHD filter for multiple targets tracking using variational Bayesian approach[J]. *Radioengineering*, 2020,29(3):529-539

[13] LIU Z, CHEN S, WU H, et al. Robust Student's t mixture probability hypothesis density filter for multi-target tracking with heavy-tailed noises [J]. *IEEE Access*, 2018,6:39208-39219

[14] HONG L, CHEN SX, WU H, et al. Student't particle implementations of probability hypothesis density filter [J]. *Application Research of Computers*, 2020,37(6): 58-62

[15] ZHANG W Y, LIANG Y, YANG F, et al. A robust student's t-based labeled multi-Bernoulli filter[C]//International Conference on Information Fusion, Ottawa, Canada,2019, 2-5

[16] HUANG Y, ZHANG Y, ZHAO Y, et al. A novel robust Gaussian-Student's t mixture distribution based Kalman filter[J]. *IEEE Transactions on Signal Processing*, 2019, 67(13):3606-3620

[17] TZIKAS D G, LIKAS A C, GALATSANOS N P. The variational approximation for Bayesian inference[J]. *IEEE Signal Processing Magazine*, 2018,25(6):131-146

[18] HU Y M, WANG X Z, LAN H, et al. An iterative nonlinear filter using variational Bayesian optimization[J]. *Sensors*, 2018,18(12): 4222

[19] HUANG Y, ZHANG Y, WU Z, et al. A novel adaptive Kalman filter with inaccurate process and measurement noise covariance matrices[J]. *IEEE Transactions on Automatic Control*, 2017,63(2):594-601

[20] VO B T, VO B N, CANTONI A. Analytic implementations of the cardinalized probability hypothesis density filter[J]. *IEEE Transactions on Signal Processing*, 2007, 55(7):3553-67

[21] ROTH M, OZKAN E, GUSTAFSSON F. A Student's t filter for heavy tailed process and measurement noise[C] //International Conference on Acoustics, Speech and Signal Processing, Vancouver, Canada, 2013: 5770-5774

HU Zhentao, born in 1979. He received his Ph. D degree in Control Science and Engineering from Northwestern Polytechnical University in 2010. He also received his B. S. and M. S. degrees from Henan University in 2003 and 2006 respectively. Now, he is a professor of School of Artificial Intelligence, Henan University. His research interests include complex system modeling and estimation, target tracking and particle filter.