IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received May 28, 2018, accepted July 13, 2018, date of publication July 17, 2018, date of current version August 7, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2856847

Robust Student's t Mixture Probability Hypothesis Density Filter for Multi-Target Tracking With Heavy-Tailed Noises

ZHUOWEI LIU[1], SHUXIN CHEN[1], HAO WU[1], AND KUN CHEN[2]
[1]Department of Communication Theory, Air Force Engineering University, Xi'an 710077, China
[2]Department of Aviation Theory, Air Force Harbin Flight Academy, Harbin 150000, China

Corresponding author: Hao Wu (wuhaostudy@163.com)

This work was supported by the National Natural Science Foundation of China under Grant 61703420 and Grant 61673392.

ABSTRACT In order to improve filtering accuracy and restrain the degradation of filtering performance caused by the heavy-tailed process and measurement noises in multi-target tracking, this paper proposes a robust Student's t mixture probability hypothesis density (PHD) filter. In the proposed method, a Student's t mixture is implemented to the PHD filter, which recursively propagates the intensity as a mixture of Student's t components in PHD filtering framework. Furthermore, with the advantage of a designed judging and re-weighting mechanism, an M-estimation-based dual-gating strategy is designed for the Student's t mixture implementation to suppress the negative effect of the heavy-tailed noises. Our proposed approach not only utilizes the Student's t distribution to match the real heavy-tailed non-Gaussian noise well but also enhances the robustness of the Student's t mixture-based approach via the designed dual-gating strategy. The simulation results verify that the proposed algorithm can keep good filtering accuracy in the presence of the process and measurement outliers simultaneously.

INDEX TERMS Multi-target tracking, PHD filter, student's t mixture, heavy-tailed noises, dual-gating strategy, robustness.

I. INTRODUCTION

Multi-target tracking (MTT) technique has been widely used in the civilian and military applications such as air traffic control, remote sensing, ballistic missile guidance and computer vision [1], [2]. Traditional data association based MTT approaches, such as the joint probabilistic data association (JPDA) [3] and the multiple hypothesis tracking (MHT) [4], cannot match the real-time requirement since they suffer from high computation complexity arising from data association. Recently, much attention has been drawn by the random finite set (RFS) frameworks, such as probability hypothesis density (PHD) filters [5], [6], multi-target multi-Bernoulli filters [2], [7] and labeled RFS-based multi-Bernoulli filters [8], [9], which are effective to relief the computational burden in many real applications since they need no data association.

This paper focuses on the computationally efficient PHD filter with simple filtering framework and low computation complexity. The PHD filter realizes the joint estimation of the target states and the number of target by propagating the first-order statistical moment of multi-target posterior probability density. To obtain closed form solution to the PHD filter, the sequential Monte Carlo (SMC) implementation [10] and the Gaussian mixture (GM) implementation [11] are proposed, respectively. Although the SMC implementation can address the nonlinear problems well, it suffers from high computational load. Compared with the SMC implementation, the GM implementation based on linear Gaussian assumption has the advantage of low computational cost. Furthermore, some nonlinear extensions [11]–[14] made the GM-PHD filter have ability to handle mild nonlinear problems. And the GM-PHD filter has also been widely used in many applications [15]–[17].

In MTT applications, the process and measurement models are usually accompanied with the heavy-tailed non-Gaussian noise (also called outlier in this paper). The heavy-tailed noise is usually caused by various reasons such as electromagnetic interference, sensor's own unreliability and unknown dynamic model. The performance of the PHD filter would be severely degraded by these heavy-tailed noises.

The standard GM implementation lacks of robustness to the heavy-tailed noise, especially the case that both heavy-tailed process and measurement noises exist simultaneously. There are some existing approaches [18]–[21] which can handle the heavy-tailed noises in some degree. For example, the inflating approach [19] resists the heavy-tailed noise interference through inflating covariance of the measurement noise. The adaptive approach [20], [21], which utilizes variational Bayesian (VB) approximation to estimate the multi-target state and the measurement noise covariance jointly, has the ability to cope with the mismatch between the actual heavy-tailed measurement noise and the Gaussian-assumed measurement noise. The multiple-model (MM) approach [18], which utilizes the multiple models to match the target states in different dynamic stages, can handle the mismatch between the actual heavy-tailed process noise and the Gaussian-assumed process noise. However, the inflating approach and the adaptive approach cannot address the heavy-tailed process noise, and the MM approach is not suitable to handle the heavy-tailed measurement noise. Some robust filtering methods [22], [23] based on M-estimation theory can cope with the heavy-tailed process and measurement noises simultaneously by re-weighting scheme, which are often used in the Gaussian assumption based Bayesian filtering framework. However, the capability of the method to handle heavy-tailed non-Gaussian noises is limited because of the Gaussian assumption. Some recent studies focus on coping with this difficulty. For example, [24] proposes a Student's t based approach for the linear system, which can address the process and measurement outliers well through utilizing the heavy-tailed Student's t distribution. Further, a generalized Student's t based filter for the nonlinear system [25] and some corresponding numerical approaches [26]–[29] are also proposed and developed. Although the performance of the Student's t based filter may degrades due to inappropriate choice of the degree of freedom parameter, it still shows the superiority in terms of handling the heavy-tailed non-Gaussian noises compared to the Gaussian approximation based filter.

In this paper, a robust implementation of the PHD filter is proposed based on Student's t mixture approximation, intending to improve the estimation accuracy in terms of the target states and the target number in the presence of the heavy-tailed process and measurement noises. The proposed implementation recursively propagates the intensity as a mixture of Student's t components in PHD filtering framework. Based on this implementation, an M-estimation based dual-gating strategy is designed in the updating step of the Student's t mixture PHD (STM-PHD) recursion for improving its capability to handle outliers further. Firstly, the proposed gating method separates the measurements into normal measurements, outliers and clutters by two predetermined thresholds. Then, it assigns different weights to the corresponding types of the measurements with a three-piece weight function, making a balance between utilizing the useful information contained in normal measurements and outliers and suppressing the harmful information contained in outliers and clutters. The main contributions of our work are that utilize the Student's t distribution to match the real heavy-tailed non-Gaussian noise well and combine gating technique with M-estimation theory to enhance the robustness of the proposed Student's t mixture based filtering algorithm further.

The remainder of this paper is organized as follows. Section II provides an overview of the PHD filter and some properties of Student's t filtering. Section III presents the Student's t mixture implementation of the PHD filter and Section IV presents the robust Student's t mixture PHD filter with dual-gating strategy. Simulation results are displayed in Section V, and conclusions are given in Section VI.

II. BACKGROUND

A. THE PHD FILTER

Under the random finite set (RFS) framework [2], multi-target states and measurements at time k are expressed as finite sets $X_k = \{x_{k,1}, \ldots, x_{k,M(k)}\} \in F(\mathbf{X})$ and $Z_k = \{z_{k,1}, \ldots, z_{k,N(k)}\} \in F(\mathbf{Z})$, respectively, where $\mathbf{X}(\mathbf{X} \subseteq \mathbb{R}^{d_x})$ is the state space, $\mathbf{Z}(\mathbf{Z} \subseteq \mathbb{R}^{d_z})$ is the measurement space, $F(\bullet)$ is the collection of all the finite subsets of the corresponding space, $M(k), N(k) \in \mathbb{N}$, and the superscript d_x and d_z are the dimensions of $\mathbf{X}$ and $\mathbf{Z}$, respectively. Let

$$X_k = \left[\bigcup_{x \in X_{k-1}} S_{k|k-1}(x)\right] \cup \left[\bigcup_{x \in X_{k-1}} B_{k|k-1}(x)\right] \cup \Gamma_k, \quad (1)$$

$$Z_k = \left[\bigcup_{x \in X_k} G_k(x)\right] \cup \Omega_k, \quad (2)$$

where $S_{k|k-1}(x)$ and $B_{k|k-1}(x)$ denote the finite sets of survival targets and spawned targets at time k, and Γ_k is the finite set of birth targets at time k. $G_k(x)$ and Ω_k respectively represent the finite sets of measurements originated from targets and clutters at time k.

On the basis above, the MTT problem can be recast into Bayesian filtering framework. Then a tractable approach, namely the PHD filter, was proposed in [5]. In PHD recursion, multi-target intensity $D(x)$ which is the first-order moment of multi-target posterior probability density $p(X_k|Z_k)$, is recursively propagated in time. Given the PHD $D_{k-1}(x)$ at time $k-1$, the predicted PHD $D_{k|k-1}(x)$ and the updated PHD $D_k(x)$ are calculated by (3) and (4),

$$D_{k|k-1}(x) = \int (p_s(u) f_{k|k-1}(x|u) + \beta_{k|k-1}(x|u)) D_{k-1}(u) du + \gamma_k(x), \quad (3)$$

$$D_k(x) = [1 - p_D(x) + \sum_{z \in Z_k} \frac{p_D(x) g_k(z|x)}{\kappa_k(z) + \int p_D(\zeta) g_k(z|\zeta) D_{k|k-1}(\zeta) d\zeta}] \times D_{k|k-1}(x), \quad (4)$$

where $f_{k|k-1}(x|u)$ and $\beta_{k|k-1}(x|u)$ respectively are the transition density function associated with survival targets and spawned targets from state u at time $k-1$ to state x at time k, $\gamma_k(x)$ and $\kappa_k(x)$ respectively are the intensity function of birth targets and clutters at time k, $p_S(x)$ and $p_D(x)$ are

survival probability of each target and detection probability, respectively. $g_k(z|x)$ is the measurement likelihood function at time k.

B. IMPORTANT PROPERTIES OF STUDENT'S T FILTERING

The probability density function (PDF) of Student's t distribution can be presented by [30]

$$p(x) = \frac{\Gamma(\frac{\upsilon+2}{2})}{\Gamma(\frac{\upsilon}{2})} \frac{1}{(\upsilon\pi)^{d/2}} \frac{1}{\sqrt{\det(P)}} (1 + \frac{\Delta^2}{\upsilon})^{-(\frac{\upsilon+2}{2})}, \quad (5)$$

where $\Gamma(\bullet)$ is the Gamma function and $\Delta^2 = (x-m)^T P^{-1}(x-m)$. The abbreviated form of PDF is $St(x; m, P, \upsilon)$ with mean m, scale matrix P and degree of freedom parameter υ.

Student's t distribution has a heavy tail, the behavior of which is severely influenced by its degrees of freedom υ. The heavy-tailed level has a negative correlation with the parameter υ. Because of this characteristic, Student's t distribution is often used to approximate the heavy-tailed non-Gaussian noise in many scenarios. In Bayesian filtering framework, two important Lemmas for the nonlinear models derived based on the Student's t approximation are shown as follows [25].

Lemma 1: Given that the jointly PDF of the current state and one-step ahead state vectors is Student's t and Q, m and P of appropriate dimensions and that Q and P are positive definite.

$$\int St(x; f(u), Q, \upsilon_1) St(u; m, P, \upsilon_3) du = St(x; \bar{m}, \bar{P}, \upsilon_3), \quad (6)$$

where

$$\bar{m} = \int f(u) St(u; m, P, \upsilon_3) du \quad (7)$$

$$\bar{P} = \frac{\upsilon_3 - 2}{\upsilon_3} \int f(u) f^T(u) St(u; m, P, \upsilon_3) du - \frac{\upsilon_3 - 2}{\upsilon_3} \bar{m}\bar{m}^T + \frac{\upsilon_1(\upsilon_3 - 2)}{(\upsilon_1 - 2)\upsilon_3} Q. \quad (8)$$

Lemma 2: Given that the jointly PDF of the state and measurement vectors is Student's t and R, m, P of appropriate dimensions and that R and P are positive definite.

$$St(z; h(x), R, \upsilon_2) St(x; m, P, \upsilon_3) = St(z; \eta, P_{zz}, \upsilon_3) St(x; m', P', \upsilon_3'), \quad (9)$$

where

$$m' = m + K(z - \eta) \quad (10)$$

$$P' = \frac{\upsilon_3 + \Delta^2}{\upsilon_3 + d_z}(P - KP_{zz}K^T) \quad (11)$$

$$\upsilon_3' = \upsilon_3 + d_z \quad (12)$$

$$K = P_{xz}(P_{zz})^{-1} \quad (13)$$

$$\Delta^2 = (z - \eta)^T (P_{zz})^{-1}(z - \eta) \quad (14)$$

$$\eta = \int h(x) St(x; m, P, \upsilon_3) dx \quad (15)$$

$$P_{zz,k|k-1} = \frac{\upsilon_3 - 2}{\upsilon_3} \int h(x) h^T(x) St(x; m, P, \upsilon_3) dx - \frac{\upsilon_3 - 2}{\upsilon_3} \eta\eta^T + \frac{\upsilon_2(\upsilon_3 - 2)}{(\upsilon_2 - 2)\upsilon_3} R \quad (16)$$

$$P_{xz,k|k-1} = \frac{\upsilon_3 - 2}{\upsilon_3} \int x h^T(x) St(x; m, P, \upsilon_3) dx_k - \frac{\upsilon_3 - 2}{\upsilon_3} m\eta^T. \quad (17)$$

The detailed derivations of Lemma 1 and Lemma 2 can be seen from [25]. Note that the Student's t integrals in (7), (8) and (15)-(17) need to be solved. Numerical solutions of the Student's t integrals can be found from [26]–[29].

III. STUDENT'S T MIXTURE PHD FILTER

A. MODELS AND ASSUMPTIONS

Consider the system models with additional noise shown in (18) and (19).

$$x_k = f_k(x_{k-1}) + \xi_k \quad (18)$$

$$z_k = h_k(x_k) + v_k, \quad (19)$$

where x_k and z_k are the state vector and the measurement vector at time k, respectively, $f_k(\cdot)$ is the state transition, $h_k(\cdot)$ is the measurement function, ξ_k and v_k are process noise vector and measurement noise vector, respectively.

Different from the Gaussian assumption based GM implementation, the proposed Student's t mixture (STM) implementation of the PHD filter is derived based on the following Student's t assumptions.

Assumption 1: The process noise and the measurement noise are assumed as the Student's t distribution.

$$p(\xi_k) = St(\xi_k; 0, Q_k, \upsilon_1) \quad (20)$$

$$p(v_k) = St(v_k; 0, R_k, \upsilon_2). \quad (21)$$

Assumption 2: Each target follows the nonlinear state and measurement models as (18)-(19), i.e.,

$$f_{k|k-1}(x|u) = St(x; f_k(u), Q_k, \upsilon_1) \quad (22)$$

$$g_k(z|x) = St(z; h_k(x), R_k, \upsilon_2). \quad (23)$$

Assumption 3: The survival probability and detection probability are state independent, i.e.,

$$p_S(x) = p_S \quad (24)$$

$$p_D(x) = p_D. \quad (25)$$

Assumption 4: Assume the birth target intensity can be approximated as a Student's t mixture of the form.

$$\gamma_k(x) = \sum_{i=1}^{J_{\gamma,k}} w_{\gamma,k}^{(i)} St\left(x; m_{\gamma,k}^{(i)}, P_{\gamma,k}, \upsilon_3\right). \quad (26)$$

Assumption 3 is commonly used in GM-PHD filters. Assumption 1, 2 and 4 are given for the actual applications with the process and measurement outliers, i.e., tracking some agile targets with unreliable sensors.

B. STUDENT'S T MIXTURE PHD RECURSION

Based on the Assumptions 1-4 and the Lemma 1 and Lemma 2 (in section II.B), the Student's t mixture based closed form recursion to PHD (3)-(4) is derived as follows [31].

1) PREDICTION STEP

Suppose that Assumptions 1-4 hold and the posterior intensity of multi-target RFS at time $k-1$ is a Student's t mixture with the form shown in (27),

$$D_{k-1}(x)=\sum_{j=1}^{J_{k-1}} w_{k-1}^{(j)} St(x; m_{k-1}^{(j)}, P_{k-1}^{(j)}, \upsilon_3). \tag{27}$$

Then the predicted intensity $D_{k|k-1}(x)$ can be given by

$$D_{k|k-1}(x)=D_{S,k|k-1}(x)+\gamma_k(x), \tag{28}$$

where the birth target intensity $\gamma_k(x)$ is shown as (26) and the survival target intensity $D_{S,k|k-1}(x)$ is approximated as a Student's t mixture of the form (29),

$$D_{S,k|k-1}(x)=\sum_{j=1}^{J_{k|k-1}} w_{S,k|k-1}^{(j)} St(x; m_{S,k|k-1}^{(j)}, P_{S,k|k-1}^{(j)}, \upsilon_3), \tag{29}$$

with

$$w_{S,k|k-1}^{(j)}=p_S w_{k-1}^{(j)} \tag{30}$$

$$\begin{aligned} m_{S,k|k-1}^{(j)} &= \int f_k(x) St(x; m_{k-1}^{(j)}, P_{k-1}^{(j)}, \upsilon_3) dx \\ &= \sum_{l=1}^{L} \mu^{(j,l)} f_k(x_{k-1}^{(j,l)}) \end{aligned} \tag{31}$$

$$\begin{aligned} P_{S,k|k-1}^{(j)} &= \frac{\upsilon_3-2}{\upsilon_3}\int f_k(x) f_k^T(x) St(x; m_{k-1}^{(j)}, P_{k-1}^{(j)}, \upsilon_3) dx \\ &\quad - \frac{\upsilon_3-2}{\upsilon_3} m_{S,k|k-1}^{(j)} (m_{S,k|k-1}^{(j)})^T \\ &\quad + \frac{\upsilon_1(\upsilon_3-2)}{(\upsilon_1-2)\upsilon_3} Q_{k-1} \\ &= \frac{\upsilon_3-2}{\upsilon_3}\sum_{l=1}^{L} \mu^{(j,l)} (f_k(x_{k-1}^{(j,l)}) - m_{S,k|k-1}^{(j)}) \\ &\quad \times (f_k(x_{k-1}^{(j,l)}) - m_{S,k|k-1}^{(j)})^T + \frac{\upsilon_1(\upsilon_3-2)}{\upsilon_3(\upsilon_1-2)} Q_{k-1}, \end{aligned} \tag{32}$$

where $\left\{x_{k-1}^{(j,l)}, \mu^{(j,l)}\right\}_{l=1}^{L}$ are the set of cubature points and weights corresponding to $St(x; m_{k-1}^{(j)}, P_{k-1}^{(j)}, \upsilon_3)$, computation of which can be found from [27].

2) UPDATING STEP

Suppose that Assumptions 1–4 hold and the predicted PHD $D_{k|k-1}(x)$ is a Student's t mixture,

$$D_{k|k-1}(x)=\sum_{i=1}^{J_{k|k-1}} w_{k|k-1}^{(j)} St(x;\ m_{k|k-1}^{(j)}, P_{k|k-1}^{(j)}, \upsilon_3). \tag{33}$$

Then, the posterior intensity $D_k(x)$ at time k can be approximated by a Student's t mixture,

$$D_k(x)=(1-p_D)D_{k|k-1}(x)+\sum_{z\in Z_k} D_k(x; z), \tag{34}$$

with

$$D_k(x; z)=\sum_{j=1}^{J_{k|k-1}} w_k^{(j)}(z) St(x; m_{k|k}^{(j)\prime}(z), P_{k|k}^{(j)\prime}(z), \upsilon_3') \tag{35}$$

$$\upsilon_3' = \upsilon_3 + d_z \tag{36}$$

$$w_k^{(j)}(z)=\frac{p_D w_{k|k-1}^{(j)} q_k^{(j)}(z)}{\kappa_k(z)+p_D\sum_{l=1}^{J_{k|k-1}} w_{k|k-1}^{(l)} q_k^{(l)}(z)} \tag{37}$$

$$m_{k|k}^{(j)\prime}(z) = m_{k|k-1}^{(j)} + K_k^{(j)}(z-\eta_{k|k-1}^{(j)}) \tag{38}$$

$$\begin{aligned} P_{k|k}^{(j)\prime}(z) &= \frac{\upsilon_3+\Delta_k^2}{\upsilon_3+d_z} \\ &\quad \times (P_{k|k-1}^{(j)} - P_{xz,k|k-1}^{(j)}(P_{zz,k|k-1}^{(j)})^{-1}(P_{xz,k|k-1}^{(j)})^T) \end{aligned} \tag{39}$$

$$q_k^{(j)}(z) = St\left(z; \eta_{k|k-1}^{(j)}, P_{zz,k|k-1}^{(j)}, \upsilon_3\right) \tag{40}$$

$$K_k^{(j)} = P_{xz,k|k-1}^{(j)}(P_{zz,k|k-1}^{(j)})^{-1} \tag{41}$$

$$\Delta^2 = (z-\eta_{k|k-1}^{(j)})^T (P_{zz,k|k-1}^{(j)})^{-1}(z-\eta_{k|k-1}^{(j)}) \tag{42}$$

$$\begin{aligned} \eta_{k|k-1}^{(j)} &= \int h_k(x) St(x; m_{k|k-1}^{(j)}, P_{k|k-1}^{(j)}, \upsilon_3) dx \\ &= \sum_{l=1}^{L} \mu^{(j,l)} h_k(x_{k|k-1}^{(j,l)}) \end{aligned} \tag{43}$$

$$\begin{aligned} P_{zz,k|k-1}^{(j)} &= \frac{\upsilon_3-2}{\upsilon_3}\int h_k(x) h_k^T(x) St(x; m_{k|k-1}^{(j)}, P_{k|k-1}^{(j)}, \upsilon_3) dx \\ &\quad - \frac{\upsilon_3-2}{\upsilon_3}\eta_{k|k-1}^{(j)}(\eta_{k|k-1}^{(j)})^T + \frac{\upsilon_2(\upsilon_3-2)}{(\upsilon_2-2)\upsilon_3} R_k \\ &= \frac{\upsilon_3-2}{\upsilon_3}\sum_{l=1}^{L} \mu^{(j,l)} (h_k(x_{k|k-1}^{(j,l)}) - \eta_{k|k-1}^{(j)}) \\ &\quad \times (h_k(x_{k|k-1}^{(j,l)}) - \eta_{k|k-1}^{(j)})^T + \frac{\upsilon_2(\upsilon_3-2)}{\upsilon_3(\upsilon_2-2)} R_k \end{aligned} \tag{44}$$

$$\begin{aligned} P_{xz,k|k-1}^{(j)} &= \frac{\upsilon_3-2}{\upsilon_3}\int x h_k^T(x) St(x; m_{k|k-1}^{(j)}, P_{k|k-1}^{(j)}, \upsilon_3) dx \\ &\quad - \frac{\upsilon_3-2}{\upsilon_3} m_{k|k-1}^{(j)}(\eta_{k|k-1}^{(j)})^T \\ &= \frac{\upsilon_3-2}{\upsilon_3}\sum_{l=1}^{L} \mu^{(j,l)} (x_{k|k-1}^{(j,l)} - m_{k|k-1}^{(j)})(h_k(x_{k|k-1}^{(j,l)}) \\ &\quad - \eta_{k|k-1}^{(j)})^T, \end{aligned} \tag{45}$$

where $\left\{x_{k|k-1}^{(j,l)}, \mu^{(j,l)}\right\}_{l=1}^{L}$ are the set of cubature points and weights corresponding to $St(x; m_{k|k-1}^{(j)}, P_{k|k-1}^{(j)}, \upsilon_3)$.

Based on (36), the degree of freedom υ_3' will tend to infinity with recursion performing. It means that the Student's t mixture converges to the Gaussian mixture [24]. It results in that the Student's t mixture PHD filter loses robustness against outliers. To avoid such degradation of performance,

correction to (35) is performed with the moment matching method shown in (46),

$$D_k(x;z)=\sum_{j=1}^{J_k} w_k^{(j)}(z)St(x;m_{k|k}^{(j)}(z),P_{k|k}^{(j)}(z),\upsilon_3) \quad (46)$$

where

$$m_{k|k}^{(j)}(z)=m_{k|k}^{(j)\prime}(z), \quad (47)$$

$$P_{k|k}^{(j)}(z)=\frac{\upsilon_3-2}{\upsilon_3}\frac{\upsilon_3'}{\upsilon_3'-2}P_{k|k}^{(j)\prime}(z). \quad (48)$$

After the corrected posterior intensity $D_k(x)$ derived, the target number and their states can be computed just like the GM-PHD filter [11]. The proposed STM implementation is a robust method, which can address the heavy-tailed noises well through choosing the appropriate value of the degree of freedom parameters (are generally set [3], [10] according to [25]).

Remark 1: Like the GM-PHD recursion, the Student's t mixture PHD (STM-PHD) recursion also needs the process of merging and pruning. And the concrete process is almost the same as the GM case (can refer to [11]). The different point is that scale matrix is presented in Student's t component while covariance matrix is presented in Gaussian component. So the scale matrix P should be converted to the covariance matrix Σ ($\Sigma=\frac{\upsilon}{\upsilon-2}P$ *[30]) in the merging process of the STM-PHD recursion.*

IV. DUAL-GATING STRATEGY BASED STUDENT'S T MIXTURE PHD RECURSION

Although the Student's t mixture PHD filter is robust to the heavy-tailed noises, the selection of the degree of freedom parameters is still a difficult task. The mismatch between the noise model and the real noise induced by the inappropriate degree of freedom parameter will degrade the performance of the algorithm, especially for the scenarios with high dense outliers. Because the M-estimation theory [32] has good robustness against outliers and its wide applications in robust filtering, this paper proposes an M-estimation based dual-gating strategy in the updating step of the STM-PHD filter, which improves the robustness of the algorithm further.

A. DUAL-GATING STRATEGY BASED PHD UPDATING

The contribution of a measurement to the posterior intensity of the PHD filter depends on the distance between the measurement projection onto the state space and the estimate. Further with the increasing of the distance, the contribution of the measurement is decreasing. Once a certain distance limit is exceeded, the contribution of the measurement can be ignored [33]. This illustrates that the almost all contribution to the posterior intensity just orients from the nearby measurements. Based on this basic principle, the likelihood function can be given as follows.

$$\hat{g}_k(z|x)=\begin{cases} g_k^1(z|x) & \text{if } d(z,h_k(x))\le I_1 \\ g_k^2(z|x) & \text{if } I_1<d(z,h_k(x))\le I_2 \\ 0 & \text{otherwise,}\end{cases} \quad (49)$$

where I_1 and I_2 ($I_1<I_2$) are two elliptical gate thresholds, respectively, which are set for separating the measurements into normal measurements, outliers and clutters according to measurements noise variance. $d(z,h_k(x))$ denotes a distance between the measurement z and the predicted measurement $h_k(x)$. The judging condition $d(z,h_k(x))\le I_1$ indicates that the measurement is normal measurement, corresponding to the likelihood $g_k^1(z|x)$. Under the condition of $I_1<d(z,h_k(x))\le I_2$, the measurement is regarded as outlier, corresponding to the likelihood $g_k^2(z|x)$. And $g_k^2(z|x)$ is always lower than $g_k^1(z|x)$. The likelihood corresponding to clutter is set zero. Such strategy makes a trade-off between utilizing the useful information and resisting the harmful information from the measurements.

Similar to the judging mechanism of typical gating techniques [4], [13], [33], [34], dual-gating strategy reduces the validation region of observation through presetting gate thresholds. The reduction of the validation region will reduce the number of clutters used for updating. In theory, the clutter intensity needs to be corrected with the change of validation region. However, the contribution of discarded clutters is so small that the correction of the clutter intensity can be ignored.

In our approach, the dual-gating strategy based PHD corrector with the modified likelihood $\hat{g}_k(z|x)$ can be formally written as

$$\begin{aligned} &D_k(x)\\ &=[1-p_{D,k}(x)\\ &\quad+\sum_{z\in Z_k}\frac{p_D(x)\hat{g}_k(z|x)}{\kappa_k(z)+\int p_D(\zeta)\hat{g}_k(z|\zeta)D_{k|k-1}(\zeta)d\zeta}]D_{k|k-1}(x). \end{aligned} \quad (50)$$

B. DUAL-GATING STRATEGY BASED STUDENT'S T MIXTURE IMPLEMENTATION

The prediction stage of the dual-gating STM-PHD filter is the same as the standard STM-PHD recursion (can be seen in section III.B). The main difference is mainly located in the updating step.

The key point of the dual-gating strategy is how to construct the form of the likelihood reasonably in update step of the STM-PHD filter. Inspired by the idea [22] that the likelihood calculation can be converted to the design of an equivalent weight function, we replace the scale matrix R_k in (44) with the inverse equivalent weight matrix ρ_k^{-1}. With this change, the equivalent innovation scale matrix $\bar{P}_{zz,k|k-1}^{(j)}$ is given by

$$\begin{aligned} \bar{P}_{zz,k|k-1}^{(j)}=&\frac{\upsilon_3-2}{\upsilon_3}\sum_{l=1}^{L}\mu^{(j,l)}(h_k(x_{k|k-1}^{(j,l)})-\eta_{k|k-1}^{(j)})\\ &\times(h_k(x_{k|k-1}^{(j,l)})-\eta_{k|k-1}^{(j)})^T+\frac{\upsilon_3-2}{\upsilon_3}\frac{\upsilon_2}{\upsilon_2-2}\rho_k^{-1}. \end{aligned} \quad (51)$$

Then, the updated intensity shown as (34) changes to

$$D_k(x) = (1 - p_{D,k})D_{k|k-1}(x) + \sum_{z \in Z_k} \bar{D}_k(x; z), \tag{52}$$

where

$$\bar{D}_k(x; z) = \sum_{j=1}^{J_k} \bar{w}_k^{(j)}(z) St(x; \bar{m}_{k|k}^{(j)\prime}(z), \bar{P}_{k|k}^{(j)\prime}(z), \upsilon_3') \tag{53}$$

$$\upsilon_3' = \upsilon_3 + d_z \tag{54}$$

$$\bar{w}_k^{(j)}(z) = \frac{p_{D,k} w_{k|k-1}^{(j)} \bar{q}_k^{(j)}(z)}{\kappa_k(z) + p_{D,k} \sum_{l=1}^{J_{k|k-1}} w_{k|k-1}^{(l)} \bar{q}_k^{(l)}(z)} \tag{55}$$

$$\bar{m}_{k|k}^{(j)\prime}(z) = m_{k|k-1}^{(j)} + \bar{K}_k^{(j)}(z - \eta_{k|k-1}^{(j)}) \tag{56}$$

$$\bar{P}_{k|k}^{(j)\prime}(z) = \frac{\upsilon_3 + \bar{\Delta}^2}{\upsilon_3 + d_z} \times (P_{k|k-1}^{(j)} - P_{xz,k|k-1}^{(j)} (\bar{P}_{zz,k|k-1}^{(j)})^{-1} (P_{xz,k|k-1}^{(j)})^T) \tag{57}$$

$$\bar{q}_k^{(j)}(z) = St\left(z; z_{k|k-1}^{(j)}, \bar{P}_{zz,k|k-1}^{(j)}, \upsilon_3\right) \tag{58}$$

$$\bar{K}_k^{(j)} = P_{xz,k|k-1}^{(j)} (\bar{P}_{zz,k|k-1}^{(j)})^{-1} \tag{59}$$

$$\bar{\Delta}^2 = (z - \eta_{k|k-1}^{(j)})^T (\bar{P}_{zz,k|k-1}^{(j)})^{-1} (z - \eta_{k|k-1}^{(j)}). \tag{60}$$

The correction of the posterior intensity in the STM-PHD filter shown as (46)-(48) should be presented by

$$D_k(x; z) = \sum_{j=1}^{J_k} \bar{w}_k^{(j)}(z) St(x; m_{k|k}^{(j)}(z), P_{k|k}^{(j)}(z), \upsilon_3), \tag{61}$$

with

$$m_{k|k}^{(j)}(z) = \bar{m}_{k|k}^{(j)\prime}(z) \tag{62}$$

$$P_{k|k}^{(j)}(z) = \frac{\upsilon_3 - 2}{\upsilon_3} \frac{\upsilon_3'}{\upsilon_3' - 2} \bar{P}_{k|k}^{(j)\prime}(z), \tag{63}$$

where $\bar{\bullet}$ denotes equivalent formation of the corresponding vector or matrix. With the equivalent weight function being introduced, a response to outlier is made in updating step of the PHD recursion by propagating ρ_k^{-1}. From comparison of (50) with (55), it indicates that $\bar{q}_k^{(j)}(z)$ is the concrete expression for the likelihood $\hat{g}_k(z|x)$ in the dual-gating STM-PHD filter. So calculation of $\hat{g}_k(z|x)$ can be converted to the design of the equivalent weight matrix ρ_k.

The equivalent weight matrix ρ_k can be designed to be [35]

$$\rho_k = \begin{cases} R_k^{-1} & \lambda_k \le I_1 \\ R_k^{-1} \frac{I_1}{\lambda_k} (\frac{\lambda_k - I_1}{I_2 - I_1})^2 & I_1 < \lambda_k \le I_2 \\ O_{d_z} & \lambda_k > I_2, \end{cases} \tag{64}$$

where λ_k is a judgment variable which is used to judge the type of the measurement at time k. According to [36], we take the squared Mahalanobis distance between the predicted measurement and the actual measurement as the judgment variable, that is, $\lambda_k = (z_k - \eta_{k|k-1}) P_{zz,k|k-1}^{-1} (z_k - \eta_{k|k-1})^T$. I_1 and I_2 are the same thresholds as shown in (49).

When λ_k satisfies the judgment condition $I_1 < \lambda_k \le I_2$, the corresponding weight $R_k^{-1} \frac{I_1}{\lambda_k} (\frac{\lambda_k - I_1}{I_2 - I_1})^2$ is assigned to the present measurement which is regarded as the outlier. Such assignment can successfully reach the expectation that just utilizes the useful information contained in outliers for updating. Considering $\rho_k = O_{d_z}$ might lead to the divergence of algorithm, the evaluation of the state, scale matrix and corresponding weight should be as follows.

$$\bar{m}_k^{(j)}(z) = m_{k|k-1}^{(j)} \tag{65}$$

$$\bar{P}_k^{(j)} = P_{k|k-1}^{(j)} \tag{66}$$

$$\bar{w}_k^{(j)}(z) = 0. \tag{67}$$

C. GATE SIZE ISSUES

Usage of gating technique in STM-PHD filter will produce many gate regions, which correspond to different Student's t components, respectively. The union of all these gate regions forms validation region for the measurements. The volume of the validation region is given as

$$V_V = \bigcup_{i=1}^{n} V_i. \tag{68}$$

Based on [13], given the gate threshold I, the i-th gate volume V_i can be expressed by (69)

$$V_i = c_{d_z} |P_{zz}|^{1/2} I^{1/2}, \tag{69}$$

where c_{d_z} is a constant determined by the dimension of the measurement vector. Equation (69) indicates that the gate volume depends on both innovation scale matrix and gate threshold.

Combining (69) with (51) and (64), the gate volume under the dual-gating strategy is given as

$$\begin{cases} V_1 = c_{d_z} |P_{zz0} + R_k|^{1/2} I_1^{1/2} \\ V_2 = c_{d_z} |P_{zz0} + \beta R_k|^{1/2} I_2^{1/2}, \end{cases} \tag{70}$$

with

$$P_{zz0} = \frac{\upsilon_3 - 2}{\upsilon_3} \sum_{l=1}^{L} \mu^{(j,l)} \times (h_k(x_{k|k-1}^{(j,l)}) - \eta_{k|k-1}^{(j)})(h_k(x_{k|k-1}^{(j,l)}) - \eta_{k|k-1}^{(j)})^T \tag{71}$$

$$\beta = \frac{\lambda_k}{I_1} (\frac{I_2 - I_1}{\lambda_k - I_1})^2. \tag{72}$$

Corresponding to gate regions of the proposed method shown in Figure 1, V_1 and V_2 denote the volume of the regions corresponding to gate 1 and gate 2, respectively, and the predicted measurement $\eta_{k|k-1}$ is the center point. The region G1 is the region where the measurement $z_{1,k}$ is viewed as normal measurement. The region G2 is an important region where the measurement $z_{2,k}$ is regarded as outlier. Its volume can be expressed as $V_2 - V_1$ whose size determines the number of outliers contained in this region. (70) and (72) indicate that the gate thresholds I_1 and I_2 not only directly affect the volume of region G2, but also indirectly affect the volume of region G2 through the factor β. It means that the

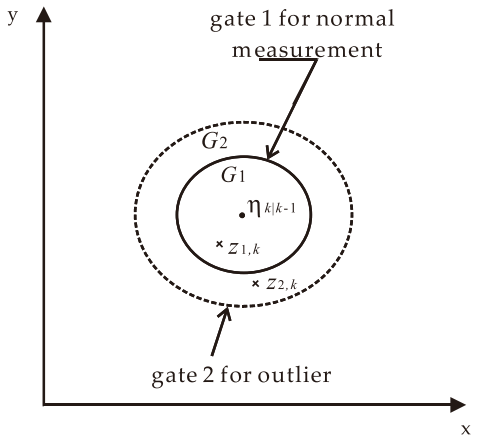


FIGURE 1. Gate region of dual-gating strategy.

choice of gate thresholds I_1 and I_2 is the key point in the dual-gating strategy.

Under the assumption of Student's t distribution, the judgment variable λ_k approximately follows the F-distribution with d_z degree of freedom. Consequently, I_1 and I_2 can be obtained according to the different confidence level α_1 and α_2 as follows.

$$p(\lambda_k > I_1) = 1 - \alpha_1 \tag{73}$$
$$p(\lambda_k > I_2) = 1 - \alpha_2, \tag{74}$$

where $p(\cdot)$ is probabilistic computation operator, I_1 and I_2 actually represent α_1 quantile and α_2 quantile, respectively, which can be chosen from F-distribution tables. Besides, the value of I_1 and I_2 can be also obtained with the empirical method according practical conditions.

Remark 2: Compared with the traditional gating techniques with only one gate threshold, the proposed gating strategy with two gate thresholds is more feasible in terms of the gate size. Concretely speaking, in traditional gating technique, the big gate size will lead more clutters to be used for updating, which not only increases the computation load but also reduces the estimation accuracy. The small gate size is likely to miss the target-oriented measurements, leading to the degradation of filtering performance. On the contrary, the dual-gating strategy can make a trade-off between the computation load and the estimation accuracy. Especially in the presence of outliers, it can effectively utilize the information contained in outliers and obtain more accurate estimation.

V. SIMULATIONS AND RESULTS

To demonstrate the performance of the proposed approach, simulation examples are designed in the linear range-only scenario and the nonlinear range-bearing scenario, respectively. In the simulation, we choose optimal subpattern assignment (OSPA) distance as the metric to evaluate the performance of filter, because OSPA distance can comprehensively measure the estimation errors of the multi-target filtering algorithms [37]. The OSPA distance between arbitrary finite sets $X = \{x_1, \ldots, x_m\}$ and $Y = \{y_1, \ldots, y_n\}$ $(m, n \in \mathbb{N}_0)$ is defined as

$$\bar{d}_p^{(c)}(X, Y) := \left(\frac{1}{n} \left(\min_{\pi \in \Pi_n} \sum_{i=1}^{m} d^{(c)}\left(x_i, y_{\pi(i)}\right)^p + c^p (n - m) \right) \right)^{\frac{1}{p}} \tag{75}$$

if $m < n$; and $\bar{d}_p^{(c)}(X, Y) := \bar{d}_p^{(c)}(Y, X)$ if $m > n$, where $d^{(c)}(x, y) := \min(c, \|x - y\|)$. In our simulation examples, the parameters p and c are set as 2 and 100, respectively.

A. LINEAR RANGE-ONLY SCENARIO

Consider a two-dimensional scenario with 10 targets which respectively appear and disappear over the observation region $[-1000, 1000]$ m $\times$ $[-1000, 1000]$ m during the interval of 100 s. The linear state model and the measurement model are the same as [11] with

$$F = \begin{bmatrix} 1 & T & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T \\ 0 & 0 & 0 & 1 \end{bmatrix}$$
$$Q = \begin{bmatrix} T^4/4 & T^3/2 & 0 & 0 \\ T^3/2 & T^2 & 0 & 0 \\ 0 & 0 & T^4/4 & T^3/2 \\ 0 & 0 & T^3/2 & T^2 \end{bmatrix} \sigma_u^2,$$
$$H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad R = \begin{bmatrix} \sigma_v^2 & 0 \\ 0 & \sigma_v^2 \end{bmatrix},$$

where $T = 1$ s, $\sigma_\xi = 5$ m/s^2 and $\sigma_v = 10$ m. The state $x_k = [p_{x,k}, p'_{x,k}, p_{y,k}, p'_{y,k}]^T$ of each target consist of position $[p_{x,k}, p_{y,k}]$ and velocity $[p'_{x,k}, p'_{y,k}]$ at time k. Table 1 presents initial conditions for each target.

TABLE 1. A list of initial target states.

Target Index	Life time (s)	Initial states (m, m/s, m, m/s)
#1	(1,70)	[0,0,0,-10]
#2	(1,100)	[400,-10,-600,5]
#3	(1,70)	[-800,20,-200,-5]
#4	(20,100)	[400,-7,-600,-4]
#5	(20,100)	[400,-2.5,-600,10]
#6	(20,100)	[0,7.5,0,-5]
#7	(40,100)	[-800,12,-200,7]
#8	(40,100)	[-200,-3,800,-10]
#9	(60,100)	[-800,3,-200,15]
#10	(60,100)	[-200,-3,800,-15]
#11	(80,100)	[0,-20,0,-15]
#12	(80,100)	[-200,15,800,-5]

The process and measurement noises with heavy tails are given as [25].

$$\xi_k \sim \begin{cases} N(0, Q) & \text{with probability } 0.96 \\ N(0, 25Q) & \text{with probability } 0.04 \end{cases} \tag{76}$$

$$v_k \sim \begin{cases} N(0, R) & \text{with probability } 0.96 \\ N(0, 25R) & \text{with probability } 0.04. \end{cases} \tag{77}$$

From (76) and (77), four percent (namely contaminated rate $\varepsilon = 0.04$ [36]) of the process noise and measurement noise generate from Gaussians with severely high covariance.

Without considering spawned target, the birth process of target is regarded as the Poisson point process with intensity

$$\gamma_k(x) = 0.03 \sum_{i=1}^{4} St(x; m_\gamma^{(i)}, P_\gamma, \upsilon_3), \tag{78}$$

where

$$\begin{aligned} m_\gamma^{(1)} &= [0\text{ m}, 0\text{ m/s}, 0\text{ m}, 0\text{ m/s}]^T \\ m_\gamma^{(2)} &= [400\text{ m}, 0\text{ m/s}, -600\text{ m}, 0\text{ m/s}]^T \\ m_\gamma^{(3)} &= [-800\text{ m}, 0\text{ m/s}, -200\text{ m}, 0\text{ m/s}]^T \\ m_\gamma^{(4)} &= [-200\text{ m}, 0\text{ m/s}, 800\text{ m}, 0\text{ m/s}]^T \\ \frac{\upsilon_3}{\upsilon_3 - 2} P_\gamma &= \text{diag}([10\text{ m}, 10\text{ m/s}, 10\text{ m}, 10\text{ m/s}]^2) \\ \upsilon_3 &= 10. \end{aligned}$$

The true trajectories of each target are shown in Figure 2. The detection probability is $p_{D,k} = 0.98$ and the survival probability is $p_{S,k} = 0.99$. The parameters related to pruning and merging are set as $T_\text{p} = 10^{-5}$, $U = 4$ and $J_\text{max} = 100$, respectively. The judgment thresholds are set as $I_1 = 10$ and $I_2 = 32$. For simplification, assume $\upsilon_1 = \upsilon_2 = \upsilon_3 = 10$.

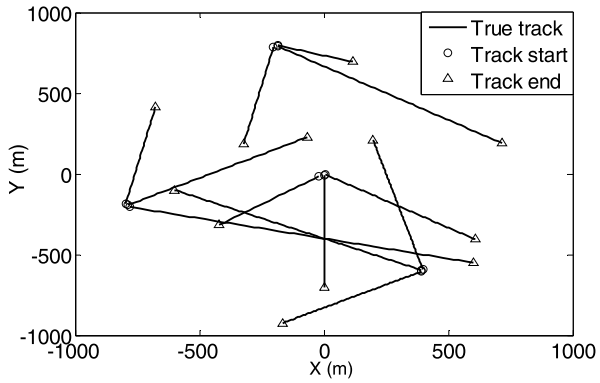


FIGURE 2. True trajectories of each target.

Case 1: To evaluate the performance of the proposed algorithm, 100 Monte Carlo (MC) trails are performed with fixed clutter rate λc=20 (an average of 20 clutter returns per scan). We make comparison among the GM-PHD filter, the Student's t mixture PHD (STM-PHD) filter and the dual-gating Student's t mixture PHD (DGSTM-PHD) filter in terms of cardinality and OSPA distance. The results are shown in Figure 3 and Figure 4, respectively.

The result in Figure 3 shows that the DGSTM-PHD filter and the STM-PHD filter have noticeable improvement in terms of the cardinality estimation accuracy compared with the GM-PHD filter. Moreover, compared with the STM-PHD filter, the DGSTM-PHD filter has obvious superiority in terms of the cardinality estimation accuracy. However, some biased cardinality estimates still appear for the DGSTM-PHD

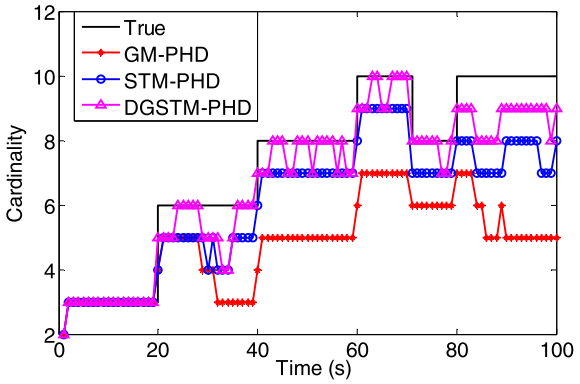


FIGURE 3. Comparison of cardinality estimation of three filters with fixed clutter rate (λc=20).

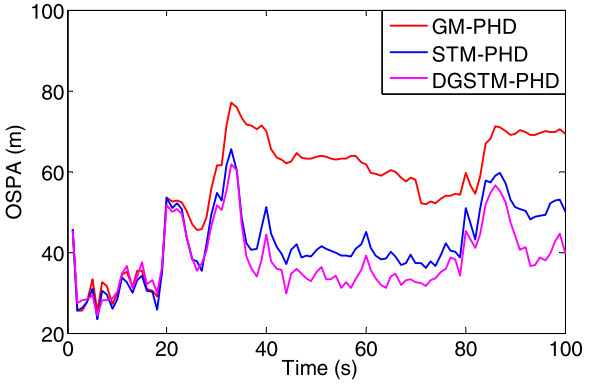


FIGURE 4. Comparison of OSPA distance of three filters with fixed clutter rate (λc=20).

filter and the STM-PHD filter with the existence of heavy-tailed process and measurement noises. The main reason is that the negative influence of the heavy-tailed process and measurement noises is too strong to be completely conquered by the DGSTM-PHD filter and the STM-PHD filter.

In Figure 4 the OSPA distance of the DGSTM-PHD filter is the lowest in three filters, and the OSPA distance of the STM-PHD filter is lower than that of the GM-PHD filter. Especially 20s later, when the outliers appear, the superiority of the DGSTM-PHD filter is more noticeable in terms of the OSPA distance. The main reason is that the DGSTM-PHD filter has more accurate cardinality estimate.

Case 2: To evaluate the performance of the proposed algorithm sufficiently, simulation is executed over 100 MC trials with different contaminated rate from $\varepsilon = 0$ to $\varepsilon = 0.06$. Then time averaged OSPA distances of three filters are shown in Figure 5.

From Figure 5, it can be seen that the time averaged OSPA distance of the STM-PHD filter is lower than that of the GM-PHD filter no matter how many the contaminated rate is. Moreover, the time averaged OSPA distance of the DGSTM-PHD filter is lower than that of the STM-PHD filter when $\varepsilon > 0.03$. This result indicates that the performance of the STM-PHD filter is superior to the GM-PHD filter. Because the Student's t noise model in the proposed approach

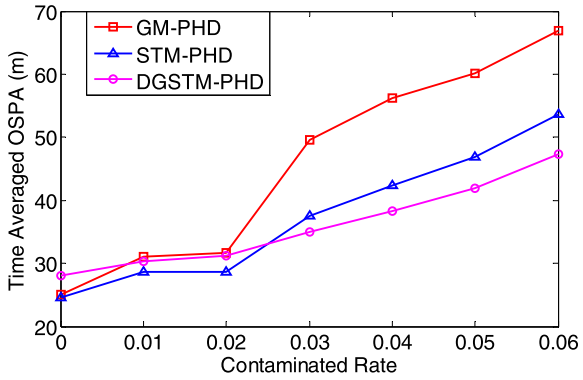


FIGURE 5. Comparison of OSPA distance of two filters with different contaminated rate (λc=20).

can match the heavy-tailed non-Gaussian noise well. Furthermore, the DGSTM-PHD filter outperforms the STM-PHD filter when more outliers appear. Because the proposed dual-gating strategy can effectively suppress the negative effect of high dense outliers, which is helpful for improving the cardinality estimation accuracy further. Additionally, at $\varepsilon = 0$, the OSPA distance for the STM-PHD filter is the same as the GM-PHD filter. It indicates that the STM-PHD filter and the GM-PHD filter have the same tracking performance when no outliers exist.

B. NONLINEAR RANGE-BEARING SCENARIO

We consider a two-dimensional scene with 12 targets which respectively appear and disappear over the observation region $[-\pi, \pi]$ rad $\times$ [0, 2000] m. Each of targets moves as a nearly constant turn dynamic model [19].

$$\begin{aligned} x_k &= F(\omega_{k-1})x_{k-1} + G\xi_k \\ \omega_k &= \omega_{k-1} + T\varsigma_k, \end{aligned} \tag{79}$$

with

$$F(\omega) = \begin{bmatrix} 1 & \sin\omega T/\omega & 0 & -(1-\cos\omega T)/\omega \\ 0 & \cos\omega T & 0 & -\sin\omega T \\ 0 & (1-\cos\omega T)/\omega & 1 & \sin\omega T/\omega \\ 0 & \sin\omega T & 0 & \cos\omega T \end{bmatrix}$$

$$G = \begin{bmatrix} T^2/2 & 0 \\ T & 0 \\ 0 & T^2/2 \\ 0 & T \end{bmatrix}$$

$$Q = \begin{bmatrix} \sigma_\xi T^2/2 & 0 & 0 \\ \sigma_\xi T & 0 & 0 \\ 0 & \sigma_\xi T^2/2 & 0 \\ 0 & \sigma_\xi T & 0 \\ 0 & 0 & \sigma_\varsigma T \end{bmatrix} \times (\bullet)^T,$$

where $F(\omega)$ and G respectively represent the transition matrix and the input matrix with sample interval $T = 1$ s. The state vector $\tilde{x}_k = [x_k^T, \omega_k]^T$ consists of positions and velocities $x_k = [p_{x,k}, p'_{x,k}, p_{y,k}, p'_{y,k}]^T$ along with turn rate ω_k. Q denotes the covariance matrix of the process noise vector $\tilde{\xi}_k = [\xi_k^T, \varsigma_k]^T$. with $\sigma_\xi = 5$ m/s^2 and $\sigma_\varsigma = \pi/180$ rad/s. $(\bullet)^T$ denotes the transpose operation for the matrix before product sign.

The range and bearing measurement vector $z_k = [r_k, \theta_k]$ admits the following model

$$z_k = \begin{bmatrix} \sqrt{p_{x,k}^2 + p_{y,k}^2} \\ \arctan\dfrac{p_{y,k}}{p_{x,k}} \end{bmatrix} + v_k. \tag{80}$$

with measurement variance $R = \text{diag}([10\text{ m}, 2(\pi/180)\text{ rad}]^2)$.

Like the linear scenario, the outliers-contaminated process and measurement noises can be given by

$$\tilde{\xi}_k \sim \begin{cases} N(0, Q) & \text{with probability } 0.96 \\ N(0, 25Q) & \text{with probability } 0.04 \end{cases} \tag{81}$$

$$v_k \sim \begin{cases} N(0, R) & \text{with probability } 0.96 \\ N(0, 25R) & \text{with probability } 0.04, \end{cases} \tag{82}$$

In the simulation, spawned target is not under consideration. The birth process of target is regarded as the Poisson point process with intensity

$$\begin{aligned} \gamma_k(x) = 0.02 \sum_{i=1}^{2} St(x; m_\gamma^{(i)}, P_\gamma, \upsilon_3) \\ + 0.03 \sum_{i=3}^{4} St(x; m_\gamma^{(i)}, P_\gamma, \upsilon_3), \end{aligned} \tag{83}$$

with

$$\begin{aligned} m_\gamma^{(1)} &= [-1500\text{ m}, 0\text{ m/s}, 250\text{ m}, 0\text{ m/s}, 0\text{ rad}]^T \\ m_\gamma^{(2)} &= [-250\text{ m}, 0\text{ m/s}, 1000\text{ m}, 0\text{ m/s}, 0\text{ rad}]^T \\ m_\gamma^{(3)} &= [250\text{ m}, 0\text{ m/s}, 750\text{ m}, 0\text{ m/s}, 0\text{ rad}]^T \\ m_\gamma^{(4)} &= [1000\text{ m}, 0\text{ m/s}, 1500\text{ m}, 0\text{ m/s}, 0\text{ rad}]^T \\ \frac{\upsilon_3}{\upsilon_3 - 2} P_\gamma &= \text{diag}([50\text{ m}, 50\text{ m/s}, 50\text{ m}, 50\text{ m/s}, 6\pi/180\text{ rad}]^2) \\ \upsilon_3 &= 6. \end{aligned}$$

In nonlinear example, parameters of filters are the same as the linear example except for the degree of freedom parameters $\upsilon_1 = \upsilon_2 = \upsilon_3 = 6$. The target trajectories are shown as Figure 3 and the initial target states can be seen in Table 2.

Case 3: To evaluate the performance of the proposed algorithm in nonlinear system, we also compared the DGSTM-PHD filter along with the STM-PHD and the GM-PHD filter in terms of cardinality and the OSPA distance. Simulation is performed over 100 MC trials under condition λc=30 and the average results are shown in Figure 7 and Figure 8.

From Figure 7, it can be seen that the DGSTM-PHD filter is superior to the STM-PHD filter and the GM-PHD

TABLE 2. A list of initial target states.

Target Index	Life time (s)	Initial states (m, m/s, m, m/s, rad)
#1	(1, 100)	[1000, −10,1500,−10,2π/(180×8)]
#2	(10, 100)	[−250, 20, 1000, 3, −2π/(180×3)]
#3	(10, 100)	[−1500, 11, 250, 10,−2π/(180×2)]
#4	(10, 66)	[−1500, 43, 250, 0, 0]
#5	(20, 80)	[250, 11, 750, 5, 2π/(180×4)]
#6	(40, 100)	[−250,−12,1000,−12,2π/(180×2)]
#7	(40, 100)	[1000, 0, 1500, −10, 2π/(180×4)]
#8	(40, 80)	[250, −50, 750, 0, −2π/(180×4)]
#9	(60, 100)	[1000, −50, 1500, 00,−2π/180×4]
#10	(60, 100)	[250, −40, 750, 25, 2π/(180×4)]

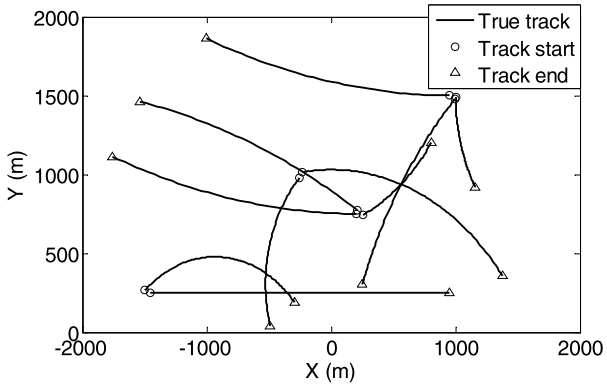


FIGURE 6. True trajectories of each target.

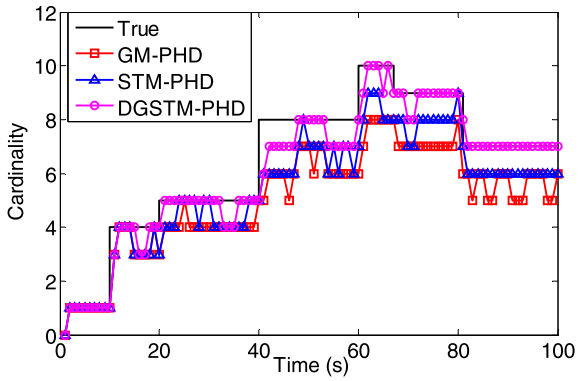


FIGURE 7. Comparison of cardinality estimation of three filters with fixed clutter rate (λc=30).

filter in terms of cardinality estimation accuracy, although the DGSTM-PHD filter has some cardinality bias. The result in Figure 7 also shows that the STM-PHD filter outperforms the GM-PHD filter.

In Figure 8, the OSPA distance of the DGSTM-PHD filter is the lowest in three filters and the OSPA distance of the STM-PHD filter is also strikingly lower than that of the GM-PHD filter. This result matches the result in Figure 7 well. It indicates that the Student's t mixture based approach can improve the estimation accuracy in the presence of the outliers and the proposed dual-gating strategy can further enhance the performance of the STM-PHD filter.

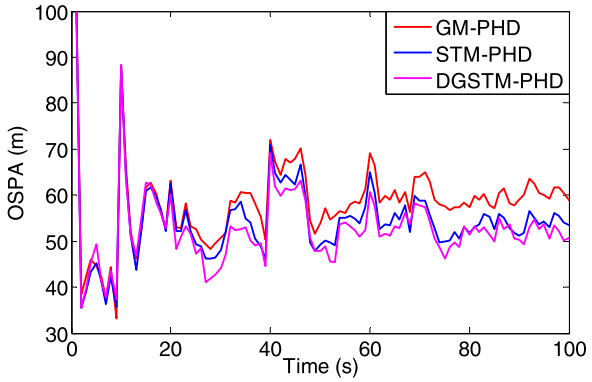


FIGURE 8. Comparison of OSPA distance of three filters with fixed clutter rate (λc=30).

Case 4: To evaluate the performance of the proposed filter sufficiently, simulation is performed over 100 MC trials with different contaminated rate from $\varepsilon = 0$ to $\varepsilon = 0.1$. Then the time averaged OSPA distances of three filters are shown in Figure 9.

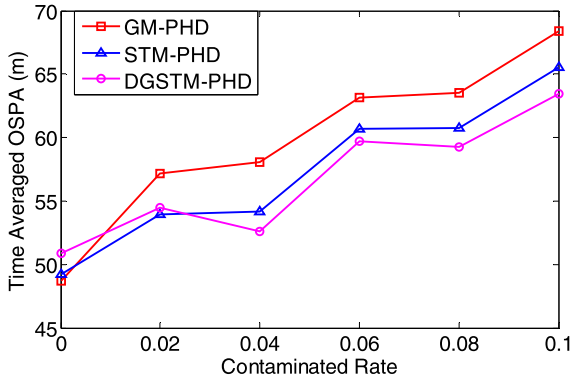


FIGURE 9. Comparison of OSPA distance of three filters with different contaminated rate.

From Figure 9, it shows the results analogous to the linear scenario as the Figure 5. Except for $\varepsilon = 0$, the time averaged OSPA distances of the DGSTM-PHD filter and the STM-PHD filter are almost lower than that of the GM-PHD filter. When $\varepsilon > 0.02$, the time averaged OSPA distance of the DGSTM-PHD filter becomes lower than that of the STM-PHD filter. This result verifies that the Student's t mixture based approach can still keep a good estimation accuracy in the nonlinear scenario with the outliers and the dual-gating strategy can improve the performance of the algorithm further. The difference is that the gaps of the time averaged OSPA distances among the three filters are not noticeable like the linear scenario. The main reason is that a relatively big approximation error is induced in nonlinear scenario for the Student's t mixture based implementation.

In summary, when the process and measurement outliers exist simultaneously, the linear scenario results in Figures 3-5 and the nonlinear scenario results in Figures 7-9 together indicate that the proposed Student's t mixture based approach

has better estimation accuracy and robustness than the Gaussian mixture based approach. In particular, the proposed dual-gating strategy improves the performance of the Student's t mixture based algorithm further in scenarios with relatively high dense outliers.

VI. CONCLUSIONS

To solve the problem that the process and measurement outliers degrade the performance of the PHD filter, this paper proposes a novel robust implementation of the PHD filter. Firstly, a Student's t mixture based analytic solution to the PHD filter is derived, which propagates the intensity as the Student's t mixture of the form. Furthermore, we propose an M-estimation based dual-gating strategy, which can resist the influence of the outliers through a judging and re-weighting mechanism. The proposed dual-gating strategy enhances the robustness of the Student's t mixture PHD filter against the outliers further. The numerical examples show that compared with the GM-based algorithm, the proposed Student's t mixture algorithm can obtain more accurate and robust estimations when the heavy-tailed process and measurement noises exist simultaneously; besides, the dual-gating strategy can effectively improve the estimation accuracy in the scenes with high dense outliers. These examples can also indicate that the proposed robust algorithm is more suitable for the real engineering applications than the GM-PHD filter. In the future research, we will try to improve the robustness of the other RFS-based filters against heavy-tailed noises.

REFERENCES

[1] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation With Applications to Tracking and Navigation: Theory Algorithms and Software*. New York, NY, USA: Wiley, 2001.

[2] R. P. S. Mahler, *Advances in Statistical Multisource-Multitarget Information Fusion*. Norwood, MA, USA: Artech House, 2014.

[3] Y. Bar-Shalom and T. E. Fortmann, *Tracking and Data Association*. New York, NY, USA: Academic, 1988.

[4] S. S. Blackman and R. Popoli, *Design and Analysis of Modern Tracking Systems*. Norwood, MA, USA: Artech House, 1999.

[5] R. P. S. Mahler, "Multitarget Bayes filtering via first-order multitarget moments," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 39, no. 4, pp. 1152–1178, Oct. 2003.

[6] R. P. S. Mahler, "PHD filters of higher order in target number," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 43, no. 4, pp. 1523–1543, Oct. 2007.

[7] B.-T. Vo, B.-N. Vo, and A. Cantoni, "The cardinality balanced multi-target multi-Bernoulli filter and its implementations," *IEEE Trans. Signal Process.*, vol. 57, no. 2, pp. 409–423, Feb. 2009.

[8] B.-T. Vo and B.-N. Vo, "Labeled random finite sets and multi-object conjugate priors," *IEEE Trans. Signal Process.*, vol. 61, no. 13, pp. 3460–3475, Jul. 2013.

[9] S. Reuter, B. T. Vo, B. N. Vo, and K. Dietmayer, "The labeled multi-Bernoulli filter," *IEEE Trans. Signal Process.*, vol. 62, no. 12, pp. 3246–3260, Dec. 2014.

[10] B.-N. Vo, S. Singh, and A. Doucet, "Sequential Monte Carlo methods for multitarget filtering with random finite sets," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 41, no. 4, pp. 1224–1245, Oct. 2005.

[11] B. N. Vo and W. K. Ma, "The Gaussian mixture probability hypothesis density filter," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4091–4104, Nov. 2006.

[12] D. Macagnano and G. T. F. De Abreu, "Multitarget tracking with the Cubature Kalman probability hypothesis density filter," in *Proc. Conf. Rec. 44th Asilomar Conf. Signals, Syst. Comput.*, Nov. 2010, pp. 1455–1459.

[13] D. Macagnano and G. T. F. De Abreu, "Adaptive gating for multitarget tracking with Gaussian mixture filters," *IEEE Trans. Signal Process.*, vol. 60, no. 3, pp. 1533–1538, Mar. 2012.

[14] C. Fu and J. Zhang, "Observation bootstrapping CDK-GMPHD filter based on consensus fusion strategy," *Optik-Int. J. Light. Electron. Opt.*, vol. 136, pp. 301–311, May 2017.

[15] Q. Zhang and T. L. Song, "Improved bearings-only multi-target tracking with GM-PHD filtering," *Sensors*, vol. 16, no. 9, p. 1469, Sep. 2016.

[16] T. Kutschbach, E. Bochinski, V. Eiselein, and T. Sikora, "Sequential sensor fusion combining probability hypothesis density and kernelized correlation filters for multi-object tracking in video data," in *Proc. 14th Int. Conf. Adv. Video Signal Based Surveill. (AVSS)*, Aug./Sep. 2017, pp. 1–5.

[17] L. Zhang, T. Wang, F. Zhang, and D. Xu, "Cooperative localization for multi-AUVs based on GM-PHD filters and information entropy theory," *Sensors*, vol. 17, no. 10, p. 2286, Oct. 2017.

[18] S. A. Pasha, B.-N. Vo, H. D. Tuan, and W.-K. Ma, "A Gaussian mixture PHD filter for jump Markov system models," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 45, no. 3, pp. 919–936, Jul. 2009.

[19] B.-T. Vo, B.-N. Vo, R. Hoseinnezhad, and R. P. S. Mahler, "Robust multi-Bernoulli filtering," *IEEE J. Sel. Topics Signal Process.*, vol. 7, no. 3, pp. 399–409, Jun. 2013.

[20] T. Ardeshiri and E. Özkan, "An adaptive PHD filter for tracking with unknown sensor characteristics," in *Proc. 16th Int. Conf. Inf. Fusion (FUSION)*, Jul. 2013, pp. 1736–1743.

[21] G. Zhang, F. Lian, C. Han, and S. Han, "An improved PHD filter based on variational Bayesian method for multi-target tracking," in *Proc. 17th Int. Conf. Inf. Fusion (FUSION)*, Jul. 2014, pp. 1–6.

[22] H. Wu, S. Chen, B. Yang, and K. Chen, "Robust derivative-free cubature Kalman filter for bearings-only tracking," *J. Guid., Control, Dyn.*, vol. 39, no. 8, pp. 1866–1871, Aug. 2016.

[23] H. Zhou, H. Huang, H. Zhao, X. Zhao, and X. Yin, "Adaptive unscented Kalman filter for target tracking in the presence of nonlinear systems involving model mismatches," *Remote Sens.*, vol. 9, no. 7, p. 657, Jun. 2017.

[24] M. Roth, E. Özkan, and F. Gustafsson, "A Student's t filter for heavy tailed process and measurement noise," in *Proc. 38th Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2013, pp. 5770–5774.

[25] Y. Huang, Y. Zhang, N. Li, and J. Chambers, "Robust Student's t based nonlinear filter and smoother," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 52, no. 5, pp. 2586–2596, Oct. 2017.

[26] F. Tronarp, R. Hostettler, and S. Särkkä, "Sigma-point filtering for nonlinear systems with non-additive heavy-tailed noise," in *Proc. 19th Int. Conf. Inf. Fusion (FUSION)*, Jul. 2016, pp. 1859–1866.

[27] Y. Huang, Y. Zhang, N. Li, S. M. Naqvi, and J. Chambers, "A robust Student's t based cubature filter," in *Proc. Int. Conf. Inf. Fusion (FUSION)*, Jul. 2016, pp. 9–16.

[28] O. Straka and J. Duník, "Stochastic integration Student's-t filter," in *Proc. 20th Int. Conf. Inf. Fusion (FUSION)*, Jul. 2017, pp. 1–8.

[29] Y. Huang and Y. Zhang, "Robust Student's t-based stochastic cubature filter for nonlinear systems with heavy-tailed process and measurement noises," *IEEE Access*, vol. 5, pp. 7964–7974, 2017.

[30] S. Kotz and S. Nadarajah, *Multivariate T-Distributions and Their Applications*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[31] Z. Liu, S. Chen, H. Wu, R. He, and L. Hao, "A Student's t mixture probability hypothesis density filter for multi-target tracking with outliers," *Sensors*, vol. 18, no. 4, p. 1095, Apr. 2018.

[32] P. J. Huber and E. M. Ronchetti, *Robust Statistics*. Hoboken, NJ, USA: Wiley, 2009.

[33] T. Li, S. Sun, and T. P. Sattar, "High-speed sigma-gating SMC-PHD filter," *Signal Process.*, vol. 93, no. 9, pp. 2586–2593, Sep. 2013.

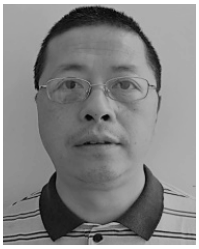
[34] H. Zhang, Z. Jing, and S. Hu, "Gaussian mixture CPHD filter with gating technique," *Signal Process.*, vol. 89, no. 8, pp. 1521–1530, 2009.

[35] L. Yang, Y. Shen, and L. Lou, "Equivalent weight robust estimation method based on median parameter estimates," (in Chinese), *Acta Geodaetica Cartographica Sin.*, vol. 40, no. 1, pp. 28–32, Feb. 2011. .

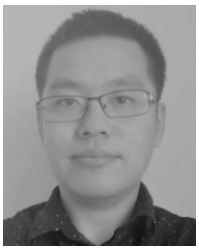
[36] G. Chang, "Robust Kalman filtering based on Mahalanobis distance as outlier judging criterion," *J. Geodesy*, vol. 88, pp. 391–401, Jan. 2014.

[37] D. Schuhmacher, B.-T. Vo, and B.-N. Vo, "A consistent metric for performance evaluation of multi-object filters," *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3447–3457, Aug. 2008.

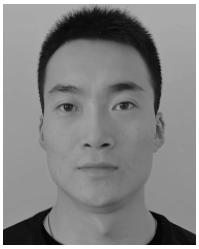
ZHUOWEI LIU was born in 1989. He received the B.S. and M.S. degrees from the Department of Flight Vehicle Propulsion Engineering, Air Force Engineering University, Xi'an, China, in 2012 and 2014, respectively, where he is currently pursuing the Ph.D. degree in information and communication engineering. His current research interests include target tracking and signal processing.



SHUXIN CHEN was born in 1965. He received the B.S. degree from the Department of Electronic Technology, National University of Defense Technology, Changsha, China, in 1987, and the M.S. degree from the Department of Electronic Engineering, Xidian University, Xi'an, China, in 1993, and the Ph.D. degree from the Department of Electronic Engineering, Northwestern Polytechnical University, Xi'an, in 2002. He is currently a Professor of communication theory with Air Force Engineering University. His research interests are signal processing and communication system modeling and simulation.



HAO WU was born in 1988. He received the B.S. and M.S. degrees from the Department of Navigation Engineering, Air Force Engineering University, Xi'an, China, in 2010 and 2012, respectively, and the Ph.D. degree in information and communication engineering from Air Force Engineering University in 2016. He is currently a Lecturer with Air Force Engineering University. His research interests are signal processing and bearings-only tracking.



KUN CHEN was born in 1989. He received the B.S. and M.S. degrees from the Department of Navigation Engineering, Air Force Engineering University, Xi'an, China, in 2011 and 2013, respectively, and the Ph.D. degree in information and communication engineering from Air Force Engineering University, in 2017. He is currently a Lecturer with Air Force Harbin Flight Academy. His research interests are signal processing and navigation technology.