

Article

A Student's t Mixture Probability Hypothesis Density Filter for Multi-Target Tracking with Outliers

Zhuowei Liu ¹, Shuxin Chen ¹, Hao Wu ^{1,*}, Renke He ¹ and Lin Hao ²

¹ Information and Navigation College Air Force Engineering University, Xi'an 710077, China; lzwlovef1@163.com (Z.L.); chenshuxin68@163.com (S.C.); lnzrds@163.com (R.H.)

² Unit 93786, Chinese People's Liberation Army (PLA), Zhangjiakou 075000, China; dearvic@126.com

* Correspondence: wuhaostudy@163.com; Tel.: +86-29-8479-1523

Received: 28 February 2018; Accepted: 29 March 2018; Published: 4 April 2018



Abstract: In multi-target tracking, the outliers-corrupted process and measurement noises can reduce the performance of the probability hypothesis density (PHD) filter severely. To solve the problem, this paper proposed a novel PHD filter, called Student's t mixture PHD (STM-PHD) filter. The proposed filter models the heavy-tailed process noise and measurement noise as a Student's t distribution as well as approximates the multi-target intensity as a mixture of Student's t components to be propagated in time. Then, a closed PHD recursion is obtained based on Student's t approximation. Our approach can make full use of the heavy-tailed characteristic of a Student's t distribution to handle the situations with heavy-tailed process and the measurement noises. The simulation results verify that the proposed filter can overcome the negative effect generated by outliers and maintain a good tracking accuracy in the simultaneous presence of process and measurement outliers.

Keywords: multi-target tracking; PHD filter; Student's t mixture; outliers; robustness

1. Introduction

Multi-target tracking (MTT) plays an important role in many sensing systems, such as infrared, radar, sonar, etc., which uses the sensor data to jointly estimate the target state and the number of targets. Nowadays it is widely used in civilian and military applications such as air traffic control, remote sensing, ballistic missile guidance, and computer vision [1,2]. In MTT, the time-varying number of targets causes a problem in that the associations between state and measurement sets of targets are hard to know, which makes the traditional data association-based multi-target tracking methods problematic. In recent years, the random finite set (RFS) theory-based multi-target tracking filters, such as probability hypothesis density (PHD) filter [3], cardinalized PHD (CPHD) filter [4], multi-target multi-Bernoulli (MeMber) filter [1] and cardinality-balanced MeMber (CBMeMber) filter [5], have attracted much more attention since they can avoid the combinatorial problem that arises from data association. Moreover, some labeled RFS-based multi-Bernoulli filters [6,7] which can accommodate target tracks were proposed.

The focus of this paper is the PHD filter, which has relatively simple recursion, making it suitable for the applications demanding real time results. The PHD filter provides a tractable sub-optimal strategy for jointly estimating the number and the state of a variable number of targets by propagating the first-order statistical moment of multi-target posterior probability density in time. It has two basic implementations: the sequential Monte Carlo (SMC) method [8] and the Gaussian mixture (GM) method [9] which can solve the problem of computationally intractable multiple integrals involved in the PHD recursions. Compared to SMC implementation, GM implementation of the PHD filter has the advantages of simple state extraction and low computational cost, which is suitable for the requirement of real-time scenes. Moreover, some nonlinear extensions [9–11] and improvements [12–14] extend the scope of applications for the GM-PHD filter.

In MTT, noise, as the important part of measurement uncertainty, is an inevitable problem that reduces estimation accuracy of the PHD filter. Vo [15] thought that the setting of a reasonably large noise variance can accommodate noise interference in most situations. However, this method is only suitable for Gaussian noise. In real applications, it is hard for the measurement noise from sensor data to follow the Gaussian distribution because of electromagnetic interference or sensors' own unreliability. Such measurements with outliers, which often express heavy-tailed character, degrade the performance of the PHD filter strikingly. What's worse, in some real applications, such as tracking some agile targets with unreliable sensors, outliers may appear in not only the measurement model, but also the process model. This situation with simultaneous heavy-tailed process and measurement noises reduces the performance of the PHD filter severely, and can even make it break down. Although the SMC-PHD filter can deal with the problem to a certain degree, it has to pay a high computational cost, especially in high dimensions. For the GM-PHD filter, its foundational Gaussian approximation limits the capability to handle heavy-tailed non-Gaussian noises. Although Huber's M-estimation theory [16] or variational Bayesian (VB) method [17] can be utilized to improve the performance of the GM-PHD filter against outliers in the measurement model, they both cannot handle the outliers in the process model. More importantly, the two methods above do not change the foundation of the Gaussian approximation-based GM-PHD filter. This means that the noise model still cannot match the outliers-corrupted process and measurement noises well, leading to biased estimates of the target state and the number of targets. Obviously, the Gaussian noise model cannot handle the heavy-tailed non-Gaussian noise, so how to model the heavy-tailed noise becomes the key point. As [18] said, Student's *t* distribution, which has a heavy tail characteristic, is a good choice to match the heavy-tailed non-Gaussian noise. Under the Bayesian filtering framework, the Student's *t* approximation-based closed form recursions are obtained for the linear system [18]. Further, the Student's *t* approximation-based approach also can be used in the nonlinear system [19–22]. However, up to the present, Student's *t* approximation-based approaches to approximate a PHD filter with heavy-tailed process and measurement noises do not exist.

In this paper, a novel 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 heavy-tailed process and measurement noises. The proposed approach models the process noise and the measurement noise as a Student's *t* distribution, meanwhile, the multi-target prior intensity is approximated as a mixture of the Student's *t* distributions. Then, the Student's *t* mixture-approximated predicted intensity and posterior intensity are obtained through utilizing Student's *t* approximation, forming a closed form recursion of the PHD filter. The Student's *t* mixture implementation is proposed in RFS-based MTT algorithms for the first time. Compared to the GM case, it is a Student's *t*-based implementation, which propagates a mixture of Student's *t* components. Because it utilizes the heavy-tailed characteristic of Student's *t* distribution, the proposed filter has better accuracy and robustness in MTT scenes with heavy-tailed process and measurement noises. Moreover, the proposed filter also has relatively low computing cost just like the GM-PHD filter. The above advantages of the proposed approach are verified by simulations designed in linear scenario and nonlinear scenario, respectively.

The remainder of this paper is organized as follows: Section 2 presents an overview of the PHD filter and some properties of the Student's *t* distribution. Section 3 presents the proposed filter for linear system and extends the proposed filter to the nonlinear system. Simulation results are given in Section 4, and conclusions are drawn in Section 5.

2. Background

2.1. The PHD Filter

The random finite set (RFS) approach [1] provides a mathematically elegant treatment for the difficult problem that is how to estimate the time-varying number of targets and their states jointly

in MTT scenarios. According to RFS theory, the collections of target states and measurements at time k can be represented as finite sets $X_k = \{x_{k,1}, \dots, x_{k,M(k)}\} \in F(\mathbf{X})$ ($\mathbf{X} \subseteq \mathbb{R}^{d_x}$) and $Z_k = \{z_{k,1}, \dots, z_{k,N(k)}\} \in F(\mathbf{Z})$ ($\mathbf{Z} \subseteq \mathbb{R}^{d_z}$) where $M(k)$ and $N(k)$ are the number of targets and the number of measurements, respectively, $F(\mathbf{X})$ and $F(\mathbf{Z})$ are the collections of all finite subsets of target states and measurements, respectively, d_x and d_z are the dimensions of the state space \mathbf{X} and the measurement space \mathbf{Z} , respectively. Let:

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

where $S_{k|k-1}(x)$ is the RFS of survival targets at time k from target states X_{k-1} and $B_{k|k-1}(x)$ is the RFS of spawned targets at time k from target states X_{k-1} , and Γ_k is the RFS of birth targets at time k .

The RFS of measurements Z_k is defined as:

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

where $G_k(x)$ and K_k are the RFS of measurements coming from target and clutter at time k , respectively.

Based on the RFS approach, the MTT problem was recast into a Bayesian filtering framework, but optimal Bayesian recursions are computationally intractable due to their sets integral. Then the PHD filter, which recursively propagates the first-order moment of multi-target posterior probability density, was proposed as an approximate solution in [3]. Let $D_{k-1}(x)$ denote the posterior intensity function at time $k - 1$, then the prediction step of the PHD filter is given by

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

where $f_{k|k-1}(\bullet|\bullet)$ is transition density function in a Markov process from time $k - 1$ to time k , $p_{s,k}(x)$ is the survival probability of each target at time k , $\beta_{k|k-1}(x|u)$ is the intensity function of the RFS $B_{k|k-1}(x|u)$ of targets spawning from previous state u , $\gamma_k(x)$ is the intensity function of birth targets at time k .

Given the predicted intensity $D_{k|k-1}(x)$, the update intensity $D_k(x)$ can be given by:

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

where $p_{D,k}(x)$ is the probability of detection at time k , $g_k(z|x)$ denotes the measurement likelihood function at time k , $\kappa_k(x)$ is the intensity function of clutter. The integration domain of the integral functions in (3) and (4) is the state space $\mathbf{X} \subseteq \mathbb{R}^{d_x}$.

2.2. Review of the GM-PHD Filter

The GM-PHD filter approximates the intensity function of multi-target as Gaussian mixture components under the assumption of the linear Gaussian dynamical model and measurement model. In addition, it still needs to satisfy the following assumptions: (1) the survival and detection probabilities are state-independent; (2) the intensities of the birth and spawn RFSs are Gaussian mixtures of the form. Then a closed-form solution of the PHD filter can be obtained as follows.

Prediction: The Gaussian mixture formulation of the PHD recursion at time $k - 1$ is given as:

$$D_{k-1}(x) = \sum_{j=1}^{J_{k-1}} w_{k-1}^{(j)} N(x; m_{k-1}^{(j)}, P_{k-1}^{(j)}), \quad (5)$$

where $N(x; \bullet, \bullet)$ denotes that random vector x follows Gaussian distribution, $m_{k-1}^{(j)}$ and $P_{k-1}^{(j)}$ are the mean and the covariance matrix of the j th item in all Gaussian mixture components, respectively, $w_{k-1}^{(j)}$ is the corresponding weight, J_{k-1} is the number of Gaussian components.

Assume that the intensities of survival targets, spawned targets and birth targets are Gaussian mixtures of the form, the predicted intensity $D_{k|k-1}(x)$ is given by:

$$D_{k|k-1}(x) = \sum_{j=1}^{J_{k|k-1}} w_{k|k-1}^{(j)} N(x; m_{k|k-1}^{(j)}, P_{k|k-1}^{(j)}), \quad (6)$$

where:

$$w_{k|k-1}^{(j)} = p_{S,k} w_{k-1}^{(j)}, \quad (7)$$

$$m_{k|k-1}^{(j)} = F_{k-1} m_{k-1}^{(j)}, \quad (8)$$

$$P_{k|k-1}^{(j)} = F_{k-1} P_{k-1}^{(j)} F_{k-1}^T + Q_{k-1}. \quad (9)$$

Update: Given the PHD predictor as Equation (6), the PHD updater is formulated as:

$$D_k(x) = (1 - p_{D,k}) D_{k|k-1}(x) + \sum_{z \in Z_k} \sum_{j=1}^{J_{k|k-1}} w_k^{(j)}(z) N(x; m_k^{(j)}(z), P_k^{(j)}(z)), \quad (10)$$

where:

$$w_k^{(j)}(z) = \frac{p_{D,k} w_{k|k-1}^{(j)} q_k^{(j)}(z)}{\kappa_k(z) + p_{D,k} \sum_{l=1}^{J_{k|k-1}} w_{k|k-1}^{(l)} q_k^{(l)}(z)}, \quad (11)$$

$$q_k^{(j)}(z) = N(z; H_k m_{k|k-1}^{(j)}, H_k P_{k|k-1}^{(j)} H_k^T + R_k), \quad (12)$$

$$m_k^{(j)}(z) = m_{k|k-1}^{(j)} + K_k^{(j)} (z - H_k m_{k|k-1}^{(j)}), \quad (13)$$

$$P_k^{(j)} = P_{k|k-1}^{(j)} - K_k^{(j)} H_k P_{k|k-1}^{(j)}, \quad (14)$$

$$K_k^{(j)} = P_{k|k-1}^{(j)} (P_{k|k-1}^{(j)})^{-1}, \quad (15)$$

in the above, F_{k-1} and H_k are the state transition matrix and observation matrix, respectively, Q_{k-1} and R_k are the process and measurement noise covariance, respectively.

2.3. Student's t Distribution

Let a random vector $x \in \mathbb{R}^d$ admit the Student's t distribution; its probability density function (PDF) can then be expressed as [23]:

$$p(x) = \frac{\Gamma(\frac{v+2}{2})}{\Gamma(\frac{v}{2})} \frac{1}{(v\pi)^{d/2}} \frac{1}{\sqrt{\det(P)}} \left(1 + \frac{\Delta^2}{v}\right)^{-\frac{(v+2)}{2}}, \quad (16)$$

where $\Gamma(\bullet)$ denotes Gamma function and $\Delta^2 = (x - m)^T P^{-1} (x - m)$. The above PDF abbreviated by $St(x; m, P, v)$ with mean m , scale matrix P and degrees of freedom parameter v . The corresponding covariance matrix can be calculated as $\frac{v}{v-2} P$ ($v > 2$) [23].

The tail behavior of the Student's t distribution is very much influenced by the degrees of freedom parameter v . The smaller v is, the heavier the tail is, and vice versa. In addition, the Student's t distribution reduces to the Gaussian as v tends to infinity, thus includes it as a special case. A number

of convenient properties are shared by both and can be easily derived. As for the PDF of affine mappings of Student's t variables [23] we have that $z = Ax + b$, with appropriate A and b , admits:

$$p(z) = St\left(z; Am + b, APA^T, v\right), \quad (17)$$

The degrees of freedom parameter remains unaltered. Turning to random vectors $x_1 \in \mathbb{R}^{d_1}$ and $x_2 \in \mathbb{R}^{d_2}$ that are jointly Student's t distributed with:

$$p(x_1, x_2) = St\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}; \begin{bmatrix} m_1 \\ m_2 \end{bmatrix}, \begin{bmatrix} P_{11} & P_{12} \\ P_{12}^T & P_{22} \end{bmatrix}, v\right), \quad (18)$$

The marginal PDF of x_1 can be computed by applying a linear transformation with $A = [I \ 0]$ to (18):

$$p(x_1) = St(x_1; m_1, P_{11}, v) \quad (19)$$

From (18) and (19), the conditional PDF is also a Student's t as follows:

$$p(x_1|x_2) = St\left(x_{1|2}; m_{1|2}, P_{1|2}, v_{1|2}\right) \quad (20)$$

with:

$$m_{1|2} = m_1 + P_{12}P_{22}^{-1}(x_2 - m_2), \quad (21)$$

$$P_{1|2} = \frac{v + \Delta_2^2}{v + d_2}(P_{11} - P_{12}P_{22}^{-1}P_{12}^T), \quad (22)$$

$$v_{1|2} = v + d_2, \quad (23)$$

$$\Delta_2^2 = (x_2 - m_2)^T P_{22}^{-1}(x_2 - m_2), \quad (24)$$

It can be seen that the conditional mean (21) corresponds to the Gaussian conditional mean. The matrix parameter (22) is a scaled version of the conditional covariance in the Gaussian case, which is recovered as v tends to infinity. In contrast to the Gaussian, $P_{1|2}$ depends on x_2 . Also, the degrees of freedom parameter in (23) increases. Above properties of Student's t distribution like (17) and (20) are the foundations of Lemmas 1 and 2 (in Section 3.3) that are the key of our proposed approach.

3. Student's t Mixture PHD Recursion

Differing from the GM-PHD recursion, the proposed approach approximates the intensity of multi-target RFS as the Student's t mixture of the form, which is propagated in the closed-form recursions. The following gives the novel algorithm for linear system at first and then extends it to nonlinear system.

3.1. Basic Assumptions for Linear Model

As for the linear system, some foundational assumptions are given as follows.

Assumption 1. Given that the process and measurement noises admit the zero-mean t distribution with scale matrix Q_{k-1} and R_k , respectively, and the initial state vector also follows a t distribution, each target follows a linear Student's t dynamical model and the sensor has a linear Student's t measurement model, i.e.:

$$f_{k|k-1}(x|\zeta) = St(x; F_{k-1}\zeta, Q_{k-1}, v_1), \quad (25)$$

$$g_k(z|x) = St(z; H_k x, R_k, v_2), \quad (26)$$

where F_{k-1} and H_k denote the transition matrix and measurement matrix, respectively. And assume the matrices F_{k-1} , H_k , Q_{k-1} and R_k are known.

Assumption 2. The survival and detection probabilities are state independent, i.e.:

$$p_{S,k}(x) = p_{S,k}, \quad (27)$$

$$p_{D,k}(x) = p_{D,k}. \quad (28)$$

Assumption 3. The intensities of the birth and spawn RFSs are Student's t mixtures of the form:

$$\gamma_k(x) = \sum_{i=1}^{J_{\gamma,k}} w_{\gamma,k}^{(i)} \text{St}\left(x; m_{\gamma,k}^{(i)}, P_{\gamma,k}, v_3\right), \quad (29)$$

$$\beta_{k|k-1}(x|\zeta) = \sum_{j=1}^{J_{\beta,k}} w_{\beta,k}^{(j)} \text{St}\left(x; F_{\beta,k-1}^{(j)}\zeta + d_{\beta,k-1}^{(j)}, Q_{\beta,k-1}^{(j)}, v_3\right). \quad (30)$$

Assumption 1 is given according the affine transition nature of Student's t variables [19]. Assumption 2 is commonly used in GM-PHD filters. Assumption 3 is given in the presence of process and measurement noises with heavy tails, so it is reasonable in actual applications, i.e., when using unreliable sensors or sensors suffer from electromagnetic interference while tracking some agile targets, where both process noise and measurement noise are prone to show a heavy-tailed character. For illustrating this, a one-dimensional heavy-tailed noise distribution and Gaussian noise distribution are given as Figure 1.

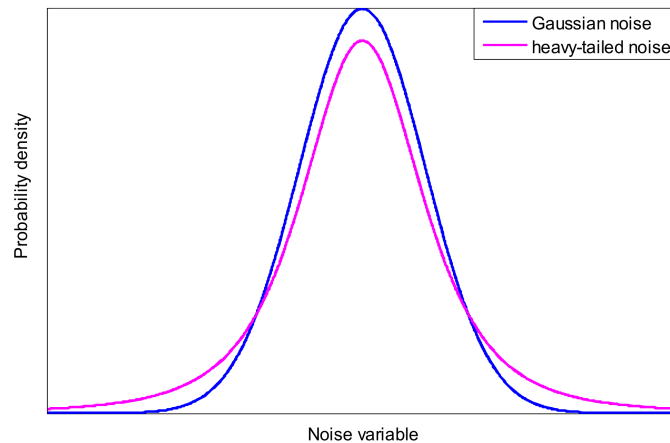


Figure 1. Illustration of heavy-tailed noise distribution and Gaussian noise distribution.

3.2. Student's t Mixture PHD Recursion

Based on the Assumptions 1–3, the closed solution to the PHD recursion (3) and (4) is presented as two Propositions. The proposed Propositions show how the Student's t components of the posterior intensity are analytically propagated to the next time, analogous to the GM-PHD recursion.

Proposition 1. Suppose that Assumptions 1–3 hold and that the posterior intensity at time $k - 1$ is a Student's t mixture of the form:

$$D_{k-1}(x) = \sum_{j=1}^{J_{k-1}} w_{k-1}^{(j)} \text{St}\left(x; m_{k-1}^{(j)}, P_{k-1}^{(j)}, v_3\right), \quad (31)$$

Then, the predicted intensity at time k is also a Student's t mixture and is given by:

$$D_{k|k-1}(x) = D_{S,k|k-1}(x) + D_{\beta,k|k-1}(x) + \gamma_k(x), \quad (32)$$

where:

$$D_{S,k|k-1}(x) = p_{S,k} \sum_{j=1}^{J_{k-1}} w_{k-1}^{(j)} St(x; m_{S,k|k-1}^{(j)}, P_{S,k|k-1}^{(j)}, v_3), \quad (33)$$

$$m_{S,k|k-1}^{(j)} = F_{k-1} m_{k-1}^{(j)}, \quad (34)$$

$$P_{S,k|k-1}^{(j)} = F_{k-1} P_{k-1}^{(j)} F_{k-1}^T + \frac{v_3 - 2}{v_3} \frac{v_1}{v_1 - 2} Q_{k-1}, \quad (35)$$

$$D_{\beta,k|k-1}(x) = \sum_{j=1}^{J_{k-1}} \sum_{l=1}^{J_{\beta,k}} w_{k-1}^{(j)} w_{\beta,k}^{(l)} St(x; m_{\beta,k|k-1}^{(j,l)}, P_{\beta,k|k-1}^{(j,l)}, v_3), \quad (36)$$

$$m_{\beta,k|k-1}^{(j,l)} = F_{\beta,k-1}^{(l)} m_{k-1}^{(j)} + d_{\beta,k-1}^{(l)}, \quad (37)$$

$$P_{\beta,k|k-1}^{(j,l)} = F_{\beta,k-1}^{(l)} P_{\beta,k-1}^{(j)} (F_{\beta,k-1}^{(l)})^T + \frac{v_3 - 2}{v_3} \frac{v_1}{v_1 - 2} Q_{\beta,k-1}. \quad (38)$$

Proposition 2. Suppose that Assumptions 1–3 hold and that the predicted intensity for time k is a Student's t mixture of the form:

$$D_{k|k-1}(x) = \sum_{j=1}^{J_{k|k-1}} w_{k|k-1}^{(j)} St(x; m_{k|k-1}^{(j)}, P_{k|k-1}^{(j)}, v_3), \quad (39)$$

Then, the posterior intensity at time k is also a Student's t mixture and is given by:

$$D_k(x) = (1 - p_{D,k}) D_{k|k-1}(x) + \sum_{z \in Z_k} D_k(x; z), \quad (40)$$

where:

$$D_k(x; z) = \sum_{j=1}^{J_{k|k-1}} w_k^{(j)}(z) St(x; m_{k|k}^{(j)'}(z), P_{k|k}^{(j)'}(z), v_3'), \quad (41)$$

$$w_k^{(j)}(z) = \frac{p_{D,k} w_{k|k-1}^{(j)} q_k^{(j)}(z)}{\kappa_k(z) + p_{D,k} \sum_{l=1}^{J_{k|k-1}} w_{k|k-1}^{(l)} q_k^{(l)}(z)}, \quad (42)$$

$$m_{k|k}^{(j)'}(z) = m_{k|k-1}^{(j)} + K_k^{(j)}(z - H_k m_{k|k-1}^{(j)}), \quad (43)$$

$$P_{k|k}^{(j)'}(z) = \frac{v_3 + \Delta^2}{v_3 + d_z} (P_{k|k-1}^{(j)} - K_k^{(j)} H_k P_{k|k-1}^{(j)}), \quad (44)$$

$$v_3' = v_3 + d_z, \quad (45)$$

$$q_k^{(j)}(z) = St(z; H_k m_{k|k-1}^{(j)}, P_{zz,k|k-1}^{(j)}, v_3), \quad (46)$$

$$K_k^{(j)} = P_{xz,k|k-1}^{(j)} (P_{zz,k|k-1}^{(j)})^{-1}, \quad (47)$$

$$\Delta^2 = (z - H_k m_{k|k-1}^{(j)})^T (P_{zz,k|k-1}^{(j)})^{-1} (z - H_k m_{k|k-1}^{(j)}), \quad (48)$$

$$P_{zz,k|k-1}^{(j)} = H_k P_{k|k-1}^{(j)} H_k^T + \frac{v_3 - 2}{v_3} \frac{v_2}{v_2 - 2} R_k, \quad (49)$$

$$P_{xz,k|k-1}^{(j)} = P_{k|k-1}^{(j)} H_k^T. \quad (50)$$

Aforementioned Propositions 1 and 2 can be established by applying the following approximated results for Student's t functions.

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

$$\int St(x; F\zeta, Q, v_1)St(\zeta; m, P, v_3)d\zeta = St(x; Fm, FPF^T + \frac{(v_3 - 2)v_1}{(v_1 - 2)v_3}Q, v_3). \quad (51)$$

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

$$St(z; Hx, R, v_2)St(x; m, P, v_3) = q(z)St(x; m', P', v_3'), \quad (52)$$

where:

$$q(z) = St(z; Hm, HPH^T + \frac{(v_3 - 2)v_2}{(v_2 - 2)v_3}R, v_3), \quad (53)$$

$$m' = m + K(z - Hm), \quad (54)$$

$$P' = \frac{v_3 + \Delta^2}{v_3 + d_z}(P - KHP), \quad (55)$$

$$v_3' = v_3 + d_z, \quad (56)$$

$$K = PH^T(HPH^T + \frac{(v_3 - 2)v_2}{(v_2 - 2)v_3}R)^{-1}, \quad (57)$$

$$\Delta^2 = (z - Hm)^T(HPH^T + \frac{(v_3 - 2)v_2}{(v_2 - 2)v_3}R)^{-1}(z - Hm). \quad (58)$$

Lemmas 1 and 2 are derived based on the properties of Student's t distribution listed in Section 2.3. The detailed derivations of Lemmas 1 and 2 can be seen [19].

Proposition 1 is established by substituting (25), (27) and (29)–(31) into the PHD prediction (3), and replacing integrals of the form (51) by appropriate Student's t as given by Lemma 1. Similarly, Proposition 2 is established by substituting (26), (28) and (39) into the PHD update (4), and then replacing integrals of the form (51) and product of Student's t of the form (52) by appropriate Student's t as given by Lemmas 1 and 2 respectively. The concrete proofs of Propositions 1 and 2 are given in Appendix B.

From (45), we know that the degree of freedom v_3' will increase infinitely with recursion performing, which results in that the Student's t mixture degrades the Gaussian mixture according to [18]. It means that the robustness against outliers for the Student's t mixture PHD filter will be lost with time going by. To solve the problem, we adapt the moment matching approach [18] to obtain the correction of posterior intensity at time k given as (59)–(61). A pseudocode of main process of the proposed algorithm is given by Table A1 in Appendix A:

$$D_k(x; z) = \sum_{j=1}^{J_k} w_k^{(j)}(z)St(x; m_{k|k}^{(j)}(z), P_{k|k}^{(j)}(z), v_3), \quad (59)$$

$$m_{k|k}^{(j)}(z) = m_{k|k}^{(j)'}(z), \quad (60)$$

$$P_{k|k}^{(j)}(z) = \frac{v_3 - 2}{v_3} \frac{v_3'}{v_3' - 2} P_{k|k}^{(j)'}(z). \quad (61)$$

Remark 1. The degree of freedom parameters v_1, v_2, v_3 for the process noise model, measurement noise model and the initial multi-target state intensity are different in general. As in the recursion performed from (31) to (39), there is a problem that how to select the degree of freedom parameter between v_1 and v_3 as the degree of freedom parameter of predicted multi-target intensity to be propagated in the next recursive step. A valid method [18] is that choose the minimized value between v_1 and v_3 to be propagated. The same problem, existing in the recursion from (39) to (40), also can be handled by selecting the minimized value between v_2 and v_3 . For simplicity, this paper assumes that the degree of freedom parameters v_1, v_2, v_3 are equal.

Remark 2. Analogous to the derivation of GM implementation of the CPHD filter in [24], the proposed Student's t mixture implementation shown as in Propositions 1 and 2 can also be used in the CPHD filter, forming the corresponding Student's t mixture CPHD recursion.

3.3. Implementation Issues

Like the GM-PHD filter, the Student's t mixture PHD (STM-PHD) filter also suffers from the computation problem that the number of Student's t components increases endlessly with recursive time, so a pruning procedure and a merging procedure are necessary for the STM-PHD filter. The concrete procedures are similar to the procedures in the GM-PHD filter (readers can refer to [9]). The different point is that the Student's t components give the scale matrix not the covariance matrix, so the covariance matrix, calculated by $\frac{v}{v-2}P$, should be used in the merging procedure of the STM-PHD recursion. The estimated number of targets is obtained by summing up the weights of all the Student's t components. This procedure for the STM-PHD filter is no different from the GM-PHD filter. Again, the state extraction is also analogous to the GM-PHD filter as to select the means of the Student's t components that have weights greater than some threshold (generally set as 0.5 [9]).

In addition, for the scene with high clutter density, gating strategy is always used to reduce the computing cost for the GM-PHD filter. It is easy to know that the gating strategy is also suitable for the STM-PHD filter. The core principle to select the measurement can be expressed as:

$$z_k^{(i)} \begin{cases} \in \tilde{Z}_k & \text{if } \exists (i, j) | (z_k^{(i)} - H_k m_{k|k-1}^{(j)})^T (P_{zz,k|k-1}^{(j)})^{-1} (z_k^{(i)} - H_k m_{k|k-1}^{(j)}) \leq T \\ \notin \tilde{Z}_k & \text{otherwise.} \end{cases} \quad (62)$$

where \tilde{Z}_k denotes the reduced set of the measurement at time k , $P_{zz,k|k-1}^{(j)}$ is the innovation covariance matrix corresponding to the j th predicted measurement, and T is the gate threshold. In contrast to the Gaussian mixture case, the judging variable $(z_k^{(i)} - H_k m_{k|k-1}^{(j)})^T (P_{zz,k|k-1}^{(j)})^{-1} (z_k^{(i)} - H_k m_{k|k-1}^{(j)})$ for the Student's t mixture case follows an F-distribution not a chi-squared distribution. It means that the gate threshold for the STM-PHD filter, which can be chosen from F-distribution table, is different from the GM-PHD filter at the same probability regions.

3.4. Extension to Nonlinear Model

This section considers the situation that process and measurement models are nonlinear. The models in Assumption 1 change to become:

$$x_k = f_k(x_{k-1}) + w_{k-1}, \quad (63)$$

$$z_k = h_k(x_k) + v_k, \quad (64)$$

where f_k and h_k are known nonlinear functions, w_{k-1} and v_k are additional noises, which follow zero-mean Student's t distribution with scale matrix Q_{k-1} and R_k , respectively.

Different from the GM-PHD filter to cope with nonlinear problems via giving the numerical solution to Gaussian integrals, the core problem of the STM-PHD filter for nonlinear systems is

how to compute the Student’s t integrals. Some researchers have given numerical solutions to Student’s t integrals based on Taylor linearization, unscented transform or cubature rule [19,21,22]. Compared with Taylor linearization and unscented transform, the cubature rule-based filter is a derivative-free and vigorous method [25], so this paper utilizes the cubature rule to extend the STM-PHD filter to nonlinear system according to [22]. The proposed algorithm for handle nonlinearity is given as Table A2 in Appendix A.

4. Simulations and Results

To illustrate the performance of the proposed filter, simulation examples are designed to compare with standard GM-PHD filter in linear and nonlinear scenarios, respectively. To compare the performance of two filters, we choose the Optimal Sub-pattern Assignment (OSPA) distance as the metric, which can comprehensively measure the cardinality and localization errors [26]. The OSPA distance is defined as follows. Let $d^{(c)}(x, y) := \min(c, \|x - y\|)$ for $x, y \in W$, and Π_k denotes the set of permutations on $\{1, 2, \dots, k\}$ for any $k \in \mathbb{N} = \{1, 2, \dots\}$. For $p \geq 1$, $c > 0$, and arbitrary finite subsets $X = \{x_1, \dots, x_m\}$ and $Y = \{y_1, \dots, y_n\}$ belong to W , where $m, n \in \mathbb{N}_0 = \{0, 1, 2, \dots\}$:

$$\bar{d}_p^{(c)}(X, Y) := \left(\frac{1}{n} \left(\min_{\pi \in \Pi_n} \sum_{i=1}^m d^{(c)}(x_i, y_{\pi(i)})^p + c^p(n - m) \right) \right)^{\frac{1}{p}}, \tag{65}$$

If $m < n$, and $\bar{d}_p^{(c)}(X, Y) := \bar{d}_p^{(c)}(Y, X)$ if $m > n$; and $\bar{d}_p^{(c)}(X, Y) = \bar{d}_p^{(c)}(Y, X) = 0$ if $m = n = 0$. p is the order parameter that determines the sensitivity to outliers and c is the cut-off parameter that determines the relative weighting of the penalties assigned to cardinality and localization errors. The details to choose the parameters p and c can be seen in [26]. In our simulation examples, we set $p = 2$ and $c = 100$.

4.1. Linear Scenario

Consider a two-dimensional scenario where there are twelve targets over region $[-1000, 1000] \times [-1000, 1000]$ during the interval of 100 s. Assuming no target spawning and each target moves as a constant velocity model similar to [9] with:

$$F = \begin{bmatrix} 1 & T & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T \\ 0 & 0 & 0 & 1 \end{bmatrix}, Q = \begin{bmatrix} T^3/4 & T^3/2 & 0 & 0 \\ T^3/2 & T & 0 & 0 \\ 0 & 0 & T^3/4 & T^3/2 \\ 0 & 0 & T^3/2 & T \end{bmatrix} q^2,$$

where $T = 1$ and $q = 5$. 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 . Their corresponding initial state and life time of each target are given as Table 1.

The noisy measurement model is the same as [9] with:

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, R = \begin{bmatrix} 100 & 0 \\ 0 & 100 \end{bmatrix},$$

The process and measurement noises with heavy tails are given as (66) and (67):

$$w_k \sim \begin{cases} N(0, Q) & \text{with probability 0.99} \\ N(0, 25Q) & \text{with probability 0.01} \end{cases}, \tag{66}$$

$$v_k \sim \begin{cases} N(0, R) & \text{with probability 0.99} \\ N(0, 25R) & \text{with probability 0.01} \end{cases}, \tag{67}$$

For the process and measurement noises in (66) and (67), about one percent of process and measurement noise values are drawn from Gaussian with severely high covariance. This percentage is also called contaminated rate which can be denoted by ε [27].

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]

Assuming no spawned target and birth targets appear spontaneously according to a Poisson point process with intensity function:

$$\gamma_k(x) = 0.03 \sum_{i=1}^4 St(x; m_{\gamma}^{(i)}, P_{\gamma}, v_3), \quad (68)$$

where $m_{\gamma}^{(1)} = [0, 0, 0, 0]^T$, $m_{\gamma}^{(2)} = [400, 0, -600, 0]^T$, $m_{\gamma}^{(3)} = [-800, 0, -200, 0]^T$ and $m_{\gamma}^{(4)} = [-200, 0, 800, 0]^T$, $\frac{v_3}{v_3-2} P_{\gamma} = \text{diag}([100, 100, 100, 100]^T)$ and $v_3 = 10$. The true trajectories of each target are shown in Figure 2, while Figure 3 plots these trajectories with Gaussian measurements and heavy-tailed measurements over time (not plot clutter in figure). From Figure 3, it is can be seen that the individual heavy-tailed measurements obviously bias the true position compared with the corresponding Gaussian measurements, which may degrade the estimation accuracy.

The detection probability and target survival probability are $p_{D,k} = 0.98$ and $p_{S,k} = 0.99$, respectively. Truncated threshold, merged threshold and the maximum Student's t components related to pruning and merging process are $T_p = 10^{-5}$, $U = 4$ and $J_{\max} = 100$, respectively. For simplification, set $v_1 = v_2 = v_3 = 10$.

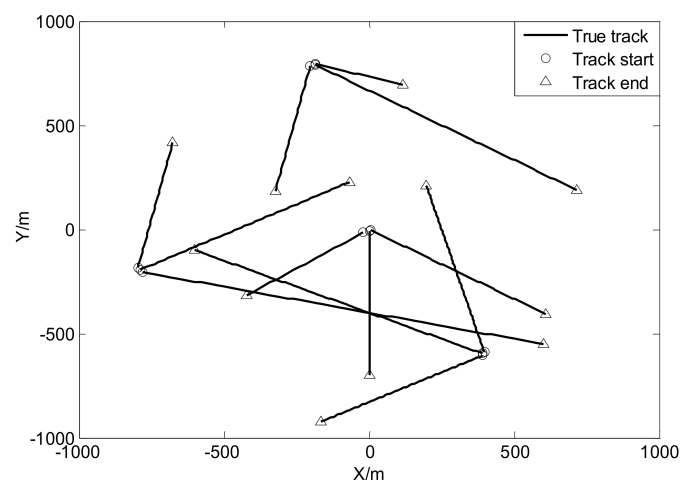


Figure 2. True trajectories of each target.

To evaluate the performance of the STM-PHD filter, we compare it with GM-PHD filter over 100 Monte Carlo (MC) trails with fixed clutter density. Under the uniform distribution assumption, the clutter density can be given by clutter rate λ_c with the relationship $\kappa_k(z) = \lambda_c/V$. In this simulation, we set $\lambda_c = 20$ (giving an average of 20 clutter returns per scan). Figures 4 and 5 respectively show the estimated cardinality and the OSPA distance for two filters. The result in Figure 4 shows that the STM-PHD filter provides a noticeable improvement in terms of cardinality estimation accuracy compared with the GM-PHD filter, although some biased cardinality estimates appear for the STM-PHD filter.

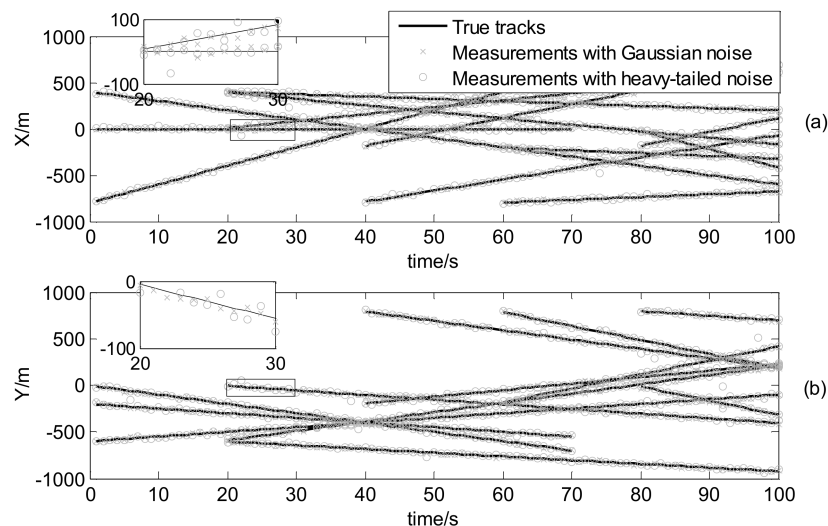


Figure 3. Measurements and true target positions versus time: (a) in x coordinate; (b) in y coordinate.

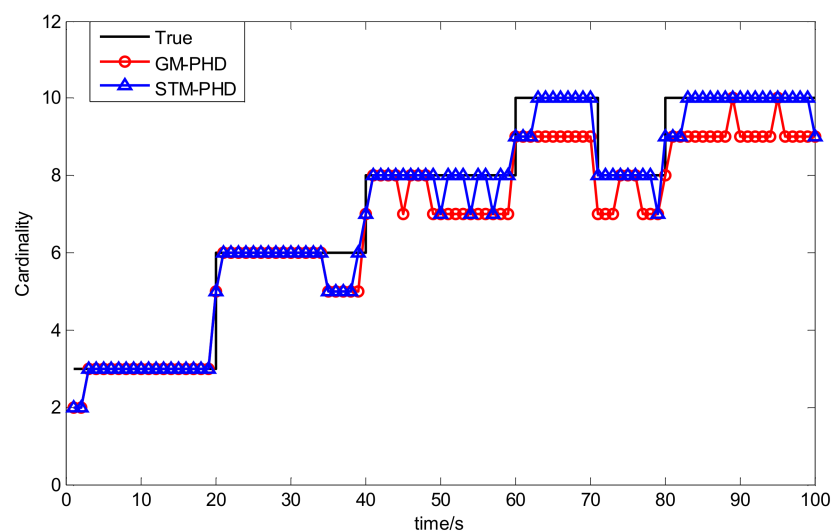


Figure 4. Comparison of cardinality estimation of two filters with fixed clutter rate ($\lambda_c = 20$).

In Figure 5 the OSPA distance of the STM-PHD filter is lower than that of the GM-PHD filter. Especially after 40 s more targets appear, difference of the OSPA distance between two filters is more noticeable. The main reason is that the STM-PHD filter has more accurate cardinality estimation.

To evaluate the performance of the proposed filter sufficiently, a simulation is executed over 100 MC trials with different contamination rates from $\varepsilon = 0$ to $\varepsilon = 0.05$. Then the time averaged OSPA distance of the STM-PHD filter and the GM-PHD filter, respectively, are shown in Figure 6.

From Figure 6, it can be seen that the time averaged OSPA distance of the STM-PHD filter is lower than that of the GM-PHD filter overall. The time averaged OSPA distances of two filters increases with the increasing contamination rate. Remarkably, the gap of OSPA distance between the STM-PHD filter and the GM-PHD filter changes wider from $\varepsilon = 0$ to $\varepsilon = 0.05$. It means that the STM-PHD filter has strong robustness against the negative effect of outliers, especially for high contamination rates. This is due to the fact the Student's *t* noise model in the proposed approach can match the heavy-tailed non-Gaussian noise well. On the contrary, the Gaussian-based GM approach matches such a non-Gaussian noise worse and worse with the increasing of contaminated rate. 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 outliers do not exist.

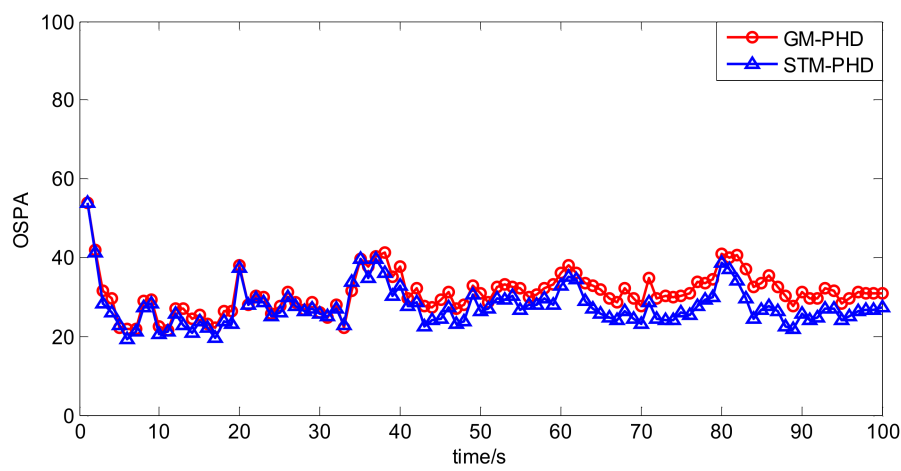


Figure 5. Comparison of OSPA distance of two filters with fixed clutter rate ($\lambda_c = 20$).

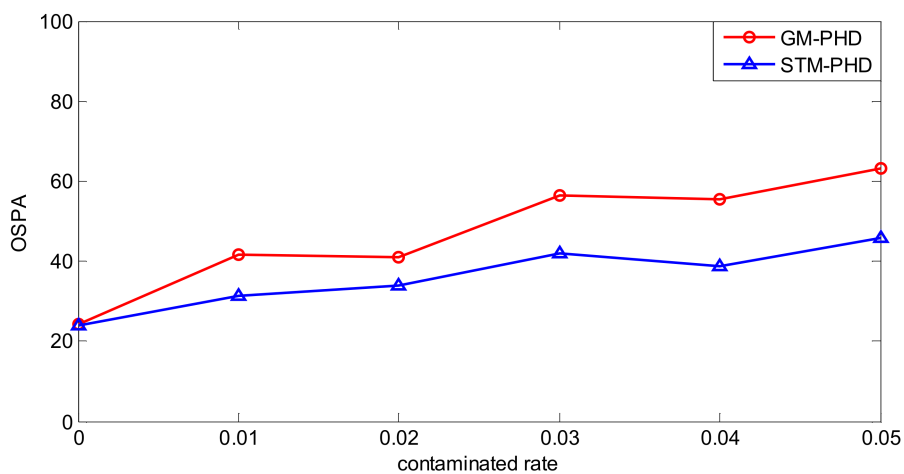


Figure 6. Comparison of OSPA distance of two filters with different contaminated rate.

To further evaluate the performance of the proposed filter, a simulation is performed over 100 MC trails with different clutter rates from $\lambda_c = 0$ to $\lambda_c = 50$. The resulting time averaged OSPA distances of the proposed filter and the GM-PHD filter are shown in Figure 7. It can be seen that the time averaged OSPA distances of two filters increase with the increasing clutter rate and the time averaged OSPA distance of the STM-PHD filter is always lower than that of the GM-PHD filter under different clutter rates. This means that the STM-PHD filter generally outperforms the GM-PHD filter when outliers exist, no matter what the clutter rate is.

In addition, the computational cost for the STM-PHD filter lies at the same level as that of the GM-PHD filter for the linear system. Running on a computer with an Intel(R) Core(TM) i5-4570 CPU at 3.2 GHz, the average computing times per execution of the GM-PHD filter and the STM-PHD filter with different clutter rate are given in Table 2.

Table 2. Average computing time with different clutter rates.

	Clutter Rate					
	0	10	20	30	40	50
GM-PHD	0.9217 s	0.9917 s	1.0782 s	1.1253 s	1.1938 s	1.2068 s
STM-PHD	0.9146 s	0.9961 s	1.0780 s	1.1417 s	1.2576 s	1.2486 s

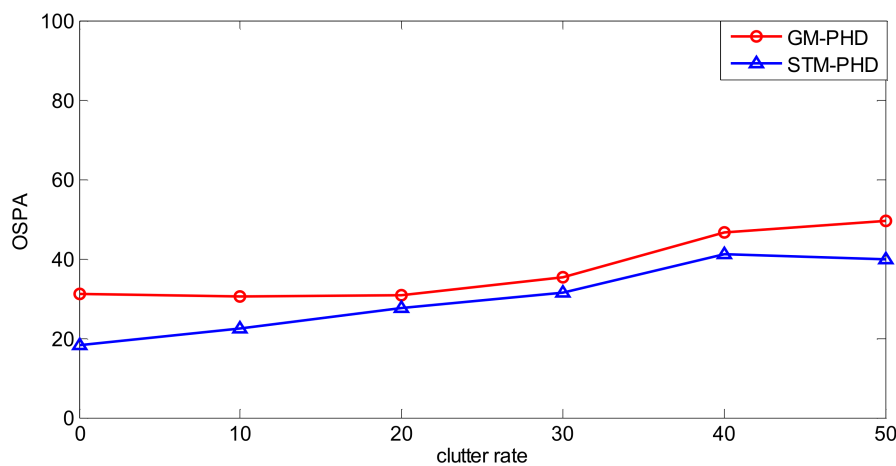


Figure 7. Comparison of OSPA distance for two filters with different clutter rate.

4.2. Nonlinear Scenario

In this example, we assume a maximum of ten targets appears on the observation region $[-\pi, \pi] \times [0, 2000]$ and a nearly constant turn state model and nonlinear bearings and range measurement model are considered according to [9]. The state $\tilde{x}_k = [x_k^T, \omega_k]^T$ consists of position and velocity $x_k = [p_{x,k}, p'_{x,k}, p_{y,k}, p'_{y,k}]^T$ as well as the turn rate ω_k . The state model is given by:

$$\begin{aligned} x_k &= F(\omega_{k-1})x_{k-1} + Gw_{k-1} \\ \omega_k &= \omega_{k-1} + Tu_{k-1} \end{aligned} \quad (69)$$

where:

$$F(\omega) = \begin{bmatrix} 1 & \sin \omega T / \omega & 0 & -(1 - \cos \omega T) / \omega \\ 0 & \cos \omega T & 0 & -\sin \omega T \\ 0 & (1 - \cos \omega T) / \omega & 1 & \sin \omega T / \omega \\ 0 & \sin \omega T & 0 & \cos \omega T \end{bmatrix}, G = \begin{bmatrix} T^2/2 & 0 \\ T & 0 \\ 0 & T^2/2 \\ 0 & T \end{bmatrix}, T = 1,$$

The noisy measurement model with range and bearing measurement $z_k = [r_k, \theta_k]$ is given by:

$$z_k = \begin{bmatrix} \sqrt{p_{x,k}^2 + p_{y,k}^2} \\ \arctan \frac{p_{y,k}}{p_{x,k}} \end{bmatrix} + v_k \quad (70)$$

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

$$w_k \sim \begin{cases} N(0, Q) & \text{with probability } 0.97 \\ N(0, 25Q) & \text{with probability } 0.03 \end{cases} \quad (71)$$

$$v_k \sim \begin{cases} N(0, R) & \text{with probability } 0.97 \\ N(0, 25R) & \text{with probability } 0.03 \end{cases} \quad (72)$$

with $Q = \text{diag}([5, 5, \pi/180]^T)^2$, $R = \text{diag}([10, 2(\pi/180)]^T)^2$.

In the simulation, we assume no spawned target and that the birth target is Poisson with intensity:

$$\gamma_k(x) = 0.02 \sum_{i=1}^2 St(x; m_{\gamma}^{(i)}, P_{\gamma}, v_3) + 0.03 \sum_{i=3}^4 St(x; m_{\gamma}^{(i)}, P_{\gamma}, v_3), \quad (73)$$

where $m_{\gamma}^{(1)} = [-1500, 0, 250, 0, 0]^T$, $m_{\gamma}^{(2)} = [-250, 0, 1000, 0, 0]^T$, $m_{\gamma}^{(3)} = [250, 0, 750, 0, 0]^T$ and $m_{\gamma}^{(4)} = [1000, 0, 1500, 0, 0]^T$, $\frac{v_3}{v_3-2} P_{\gamma} = \text{diag}([50, 50, 50, 50, 6(\pi/180)]^T)^2$ and $v_3 = 6$. (The unit of distance, angle and time in this paper are meter, radian and second, respectively.)

The initial target states are given by Table 3 and the true trajectories of each target are shown as Figure 8. In addition, Figure 9 plots corresponding measurements with Gaussian noise and heavy-tailed noise respectively over time (not plot clutter in figure). In Figure 9, it also shows the results analogous to the linear case as shown in Figure 3.

Table 3. A list of initial target states.

Target Index	Life Time (s)	Initial States (m, m/s, m, m/s, rad/s)
#1	(1, 100)	[1000, -10, 1500, -10, $2\pi/(180 \times 8)$]
#2	(10, 100)	[-250, 20, 1000, 3, $-2\pi/(180 \times 3)$]
#3	(10, 100)	[-1500, 11, 250, 10, $-2\pi/(180 \times 2)$]
#4	(10, 66)	[-1500, 43, 250, 0, 0]
#5	(20, 80)	[250, 11, 750, 5, $2\pi/(180 \times 4)$]
#6	(40, 100)	[-250, -12, 1000, -12, $2\pi/(180 \times 2)$]
#7	(40, 100)	[1000, 0, 1500, -10, $2\pi/(180 \times 4)$]
#8	(40, 80)	[250, -50, 750, 0, $-2\pi/(180 \times 4)$]
#9	(60, 100)	[1000, -50, 1500, 00, $-2\pi/180 \times 4$]
#10	(60, 100)	[250, -40, 750, 25, $2\pi/(180 \times 4)$]

To evaluate the performance of the CKF based STM-PHD filter to cope with nonlinear problem, we compare it with CKF based GM-PHD filter [11] over 100 Monte Carlo (MC) trails with fixed clutter rate $\lambda_c = 20$. Figures 10 and 11 respectively show the estimated cardinality and the OSPA distance of two filters. From Figure 10, it can be seen that the STM-PHD filter is superior to the GM-PHD filter in terms of cardinality estimation accuracy, although the STM-PHD filter has cardinality bias when the number of targets increases. The main reason for generating cardinality bias is that some necessary approximations for coping with nonlinear problems induce errors.

In Figure 11 the OSPA distances of the STM-PHD filter and the GM-PHD filter are at the same level before 65 s and later on the OSPA distance of the STM-PHD filter is obviously lower than that of the GM-PHD filter. This result matches the result in Figure 10 well. It indicates that the difference of OSPA distance between two filters is due to the difference of cardinality estimation accuracy.

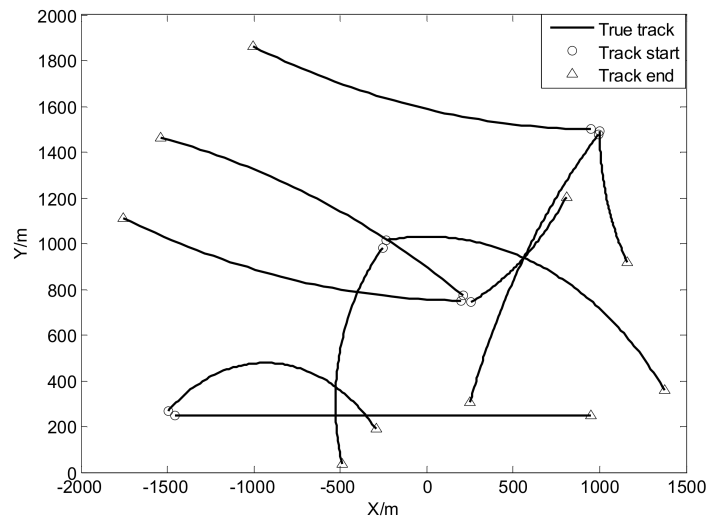


Figure 8. True trajectories of each target.

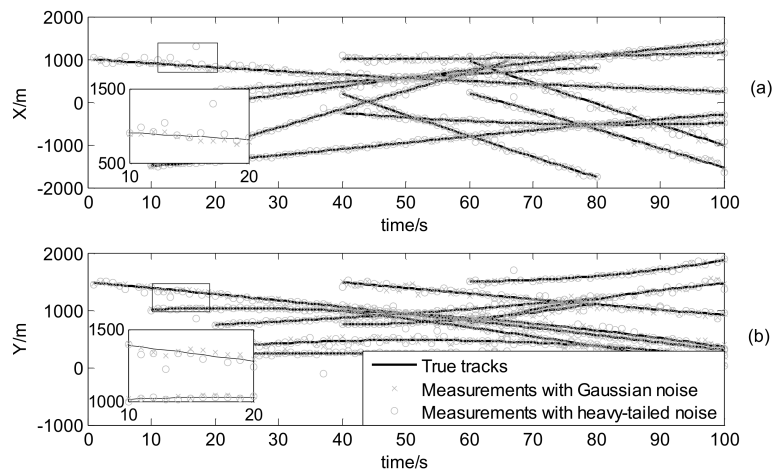


Figure 9. Measurements and true target positions versus time: (a) in x coordinate; (b) in y coordinate.

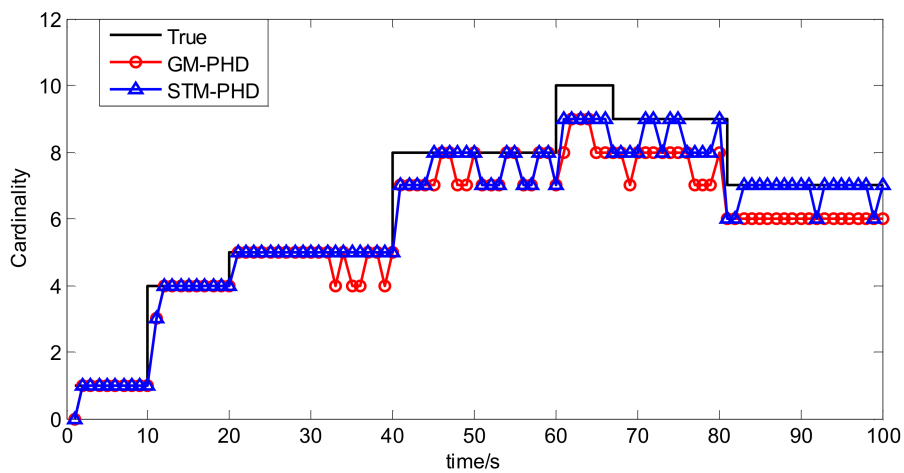


Figure 10. Comparison of cardinality estimation of two filters with fixed clutter rate ($\lambda_c = 20$).

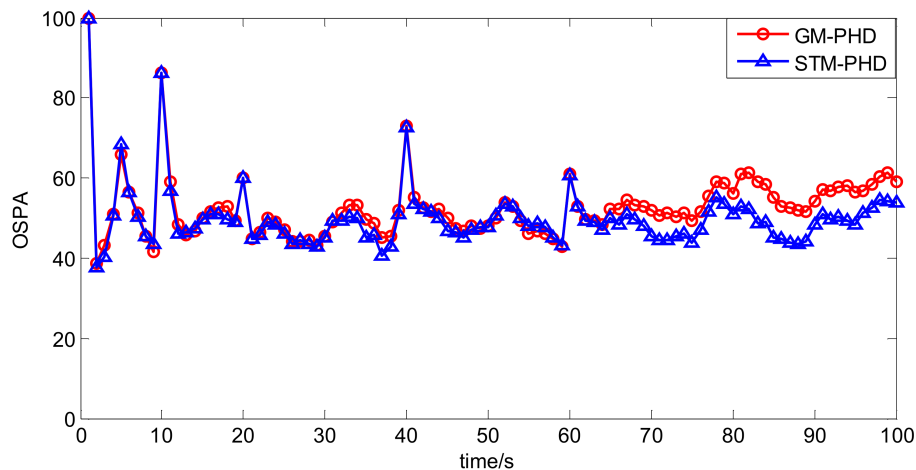


Figure 11. Comparison of OSPA distance of two filters with fixed clutter rate ($\lambda_c = 20$).

To evaluate the performance of the STM-PHD filter sufficiently, simulation is performed over 100 MC trials with different contaminated rate from $\varepsilon = 0$ to $\varepsilon = 0.05$. Then the time averaged OSPA distances of the STM-PHD filter and the GM-PHD filter are shown in Figure 12.

Analogous to the linear scenario, it can be seen that the time averaged OSPA distance of the STM-PHD filter is lower than that of the GM-PHD filter at almost all contaminated rates. The gap of the time averaged OSPA distance between the STM-PHD filter and the GM-PHD filter also changes widely as the contamination rates increase. Nevertheless, the gap is not noticeable like in the linear scenario. This is due to the relatively big approximation error induced in the nonlinear scenario. The result indicates that the STM-PHD filter still outperforms the GM-PHD filter to cope with outliers for nonlinear systems.

To further evaluate the performance of the proposed filter, 100 MC trails are performed from $\lambda_c = 0$ to $\lambda_c = 50$. The time averaged OSPA distances versus varying clutter rate for the STM-PHD filter and the GM-PHD filter are shown in Figure 13. It can be seen that the time averaged OSPA distance of the STM-PHD filter is lower than that of the GM-PHD filter at different clutter rates, and the trend for two filters goes up with the increase of clutter rate. Generally speaking, the STM-PHD filter is superior to the GM-PHD filter but the superiority is not noticeable like in the linear scenario. This is due to the fact the approximation error induced in a nonlinear scenario is bigger than that in a linear scenario. Nevertheless, the results still indicate that the STM-PHD filter is valid to handle the outliers.

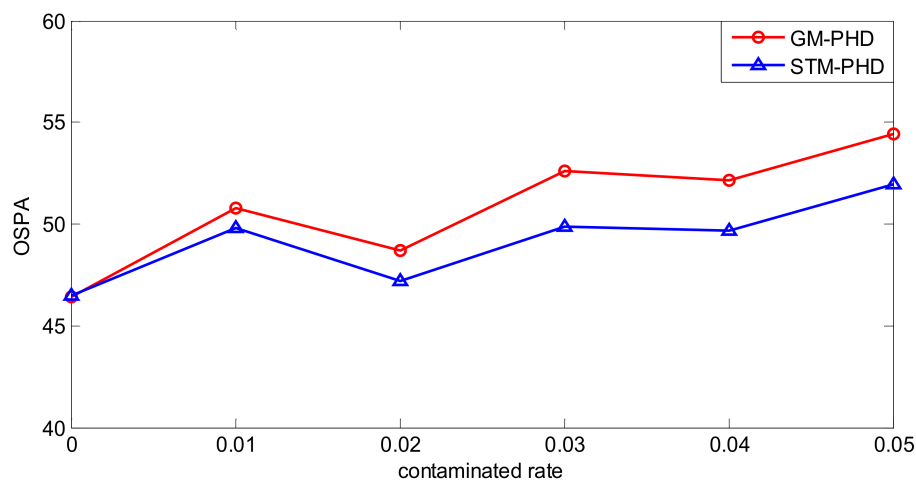


Figure 12. Comparison of OSPA distance of two filters with different contamination rates.

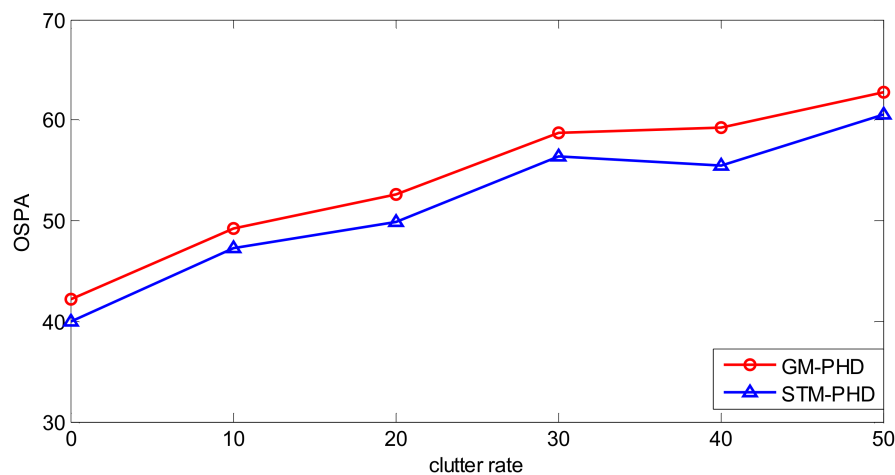


Figure 13. Comparison of OSPA distance of two filters with different clutter rates.

Again, the computing time for each filter under the different clutter rate is given in Table 4. It shows that the STM-PHD filter has a higher computing cost compared to the GM-PHD filter and the higher the clutter rate is, much more time is consumed for the STM-PHD filter. The reason is that computing nonlinear Student's *t* integrals is more complex than computing nonlinear Gaussian integrals.

Table 4. Average computing time with different clutter rate.

	Clutter Rate					
	0	10	20	30	40	50
GM-PHD	1.0142 s	1.3780 s	1.8481 s	2.4988 s	3.3321 s	3.9051 s
STM-PHD	1.8534 s	3.3179 s	5.4081 s	8.5213 s	12.9163 s	16.6113 s

5. Conclusions

To solve the problem that the process and measurement outliers degrade the performance of PHD filter, this paper proposes a Student's *t* mixture-based PHD filter, where the prior multi-target intensities are approximated as a mixture of the Student's *t* components to be propagated in time and the Student's *t* mixture approximated posterior multi-target intensity are obtained through a Student's *t* approximation-based recursion. The proposed approach can efficiently suppress the negative impact caused by the process and measurement outliers, while still maintaining good estimation accuracy. The advantages of the proposed filter are verified sufficiently by simulation. The simulation results also imply that the proposed approach outperforms the GM-based approach in scenarios where outliers appear due to electromagnetic interference or sensors' own unreliability. In addition, the proposed approach is also suitable for the CPHD filter. However, the proposed approach has relatively high computational cost for nonlinear systems. In further study, we will try to improve this. Moreover, how to combine the proposed Student's *t* mixture implementation with the multi-Bernoulli filters (including CBMeMber filter and labeled RFS based multi-Bernoulli filter) is also our focus in the next work.

Acknowledgments: This work was partly supported by the National Natural Science Foundation of China (Grants No. 61673392 and Grants No. 61703420).

Author Contributions: Zhuowei Liu performed the simulations and wrote the paper; Shuxin Chen and Hao Wu revised this paper and offered some useful suggestions in methodologies and simulations; Renke He and Lin Hao offered revision during the whole process.

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

Appendix A

Table A1. Pseudocode for the STM-PHD filter.

given $\{w_{k-1}^{(i)}, m_{k-1}^{(i)}, P_{k-1}^{(i)}\}_{i=1}^{J_{k-1}}$, the measurement set Z_k and the degree of freedom parameters v_1, v_2, v_3 .

step 1. (prediction for birth targets)

$i = 0.$
 for $j = 1, \dots, J_{\gamma,k}$
 $i := i + 1.$
 $w_{k|k-1}^{(i)} = w_{\gamma,k'}^{(j)}, m_{k|k-1}^{(i)} = m_{\gamma,k'}^{(j)}, P_{k|k-1}^{(i)} = P_{\gamma,k'}^{(j)}$
 end
 for $j = 1, \dots, J_{\beta,k}$
 for $l = 1, \dots, J_{k-1}$
 $i := i + 1.$
 $w_{k|k-1}^{(i)} = w_{k-1}^{(l)} w_{\beta,k'}^{(j)}, m_{k|k-1}^{(i)} = F_{\beta,k-1}^{(j)} m_{k-1}^{(l)} + d_{\beta,k-1}^{(j)},$
 $P_{k|k-1}^{(i)} = F_{\beta,k-1}^{(j)} P_{k-1}^{(l)} (F_{\beta,k-1}^{(j)})^T + \frac{v_1(v_3-2)}{v_3(v_1-2)} Q_{\beta,k-1},$
 end
 end

step 2. (prediction for existing targets)

for $j = 1, \dots, J_{k-1}$
 $i := i + 1.$
 $w_{k|k-1}^{(i)} = p_{S,k} w_{k-1}^{(j)},$
 $m_{k|k-1}^{(i)} = F_{k-1}^{(j)} m_{k-1}^{(j)}, P_{k|k-1}^{(i)} = F_{k-1}^{(j)} P_{k-1}^{(j)} (F_{k-1}^{(j)})^T + \frac{v_1(v_3-2)}{v_3(v_1-2)} Q_{k-1},$
 end
 $J_{k|k-1} = i.$

step 3. (construction of PHD update components)

for $j = 1, \dots, J_{k|k-1}$
 $\eta_{k|k-1}^{(j)} = H_k m_{k|k-1}^{(j)}, P_{zz,k|k-1}^{(j)} = H_k P_{k|k-1}^{(j)} H_k^T + \frac{v_2(v_3-2)}{v_3(v_2-2)} R_k,$
 $K_k^{(j)} = P_{k|k-1}^{(j)} H_k^T [P_{zz,k|k-1}^{(j)}]^{-1}, P_{k|k}^{(j)} = P_{k|k-1}^{(j)} - K_k^{(j)} H_k P_{k|k-1}^{(j)},$
 end

step 4. (update)

for $j = 1, \dots, J_{k|k-1}$
 $w_k^{(j)} = (1 - p_{D,k}) w_{k|k-1}^{(j)},$
 $m_k^{(j)} = m_{k|k-1}^{(j)}, P_k^{(j)} = P_{k|k-1}^{(j)},$
 end
 $l = 0.$
 for each $z \in Z_k$
 $l := l + 1.$
 for $j = 1, \dots, J_{k|k-1}$
 $\Delta^{(l, J_{k|k-1}+j)} = \sqrt{(z - \eta_{k|k-1}^{(j)})^T (P_{zz,k|k-1}^{(j)})^{-1} (z - \eta_{k|k-1}^{(j)})}$
 $w_k^{(l, J_{k|k-1}+j)} = p_{D,k} w_{k|k-1}^{(j)} St(z; \eta_{k|k-1}^{(j)}, S_k^{(j)}, v_3),$
 $m_k^{(l, J_{k|k-1}+j)} = m_{k|k-1}^{(j)} + K_k^{(j)} (z - \eta_{k|k-1}^{(j)}),$
 $P_k^{(l, J_{k|k-1}+j)} = \frac{v_3 + (\Delta^{(l, J_{k|k-1}+j)})^2}{v_3 + d_z} P_{k|k-1}^{(j)},$
 end
 $w_k^{(l, J_{k|k-1}+j)} := \frac{w_k^{(l, J_{k|k-1}+j)}}{\kappa_k(z) + \sum_{i=1}^{J_{k|k-1}+j} w_k^{(i, J_{k|k-1}+j)}},$ for $j = 1, \dots, J_{k|k-1}.$
 end
 $v_3' = v_3 + d_z, J_k = l J_{k|k-1} + J_{k|k-1}.$

step 5. (moment matching)

for $j = 1, \dots, J_k$
 $m_k^{(j)} = m_k^{(j)}, P_k^{(j)} = \frac{v_3'(v_3-2)}{v_3(v_3'-2)} P_k^{(j)},$
 end

output $\{w_k^{(i)}, m_k^{(i)}, P_k^{(i)}\}_{i=1}^{J_k}$

Table A2. Pseudocode for the cubature rule based STM-PHD filter.

given $\{w_{k-1}^{(i)}, m_{k-1}^{(i)}, P_{k-1}^{(i)}\}_{i=1}^{J_{k-1}}$, the measurement set Z_k and the degree of freedom parameters v_1, v_2, v_3 .

step 1. (prediction for birth targets)
 follow Step 1. of table.
 for $j = 1, \dots, i$
 – use the cubature rule suitable for Student's t integral (see [22]) with $m_{k-1}^{(j)}$ and $P_{k-1}^{(j)}$
 to obtain a set of cubature points and weights, denoted by $\{x_{k|k-1}^{(l)}, \mu^{(l)}\}_{l=1}^L$
 – compute

$$x_{k|k-1}^{(l)} := f_{\beta,k}(x_{k-1}^{(l)}), \quad l = 1, \dots, L,$$

$$m_{k|k-1}^{(j)} = \sum_{l=1}^L \mu^{(l)} x_{k|k-1}^{(l)},$$

$$P_{k|k-1}^{(j)} = \frac{v_3-2}{v_3} \sum_{l=1}^L \mu^{(l)} (x_{k|k-1}^{(l)} - m_{k|k-1}^{(j)})(x_{k|k-1}^{(l)} - m_{k|k-1}^{(j)})^T + \frac{v_1(v_3-2)}{v_3(v_1-2)} Q_{\beta,k-1},$$
 end

step 2. (prediction for existing targets)
 for $j = 1, \dots, J_{k-1}$
 $i := i + 1.$
 $w_{k|k-1}^{(j)} = p_{S,k} w_{k-1}^{(j)}$
 – use the cubature rule suitable for Student's t integral with $m_{k-1}^{(j)}$ and $P_{k-1}^{(j)}$
 to obtain a set of cubature points and weights, denoted by $\{x_{k-1}^{(l)}, \mu^{(l)}\}_{l=1}^L$
 – compute

$$x_{k|k-1}^{(l)} := f_k(x_{k-1}^{(l)}), \quad l = 1, \dots, L,$$

$$m_{k|k-1}^{(j)} = \sum_{l=1}^L \mu^{(l)} x_{k|k-1}^{(l)},$$

$$P_{k|k-1}^{(j)} = \frac{v_3-2}{v_3} \sum_{l=1}^L \mu^{(l)} (x_{k|k-1}^{(l)} - m_{k|k-1}^{(j)})(x_{k|k-1}^{(l)} - m_{k|k-1}^{(j)})^T + \frac{v_1(v_3-2)}{v_3(v_1-2)} Q_{k-1},$$
 end
 $J_{k|k-1} = i.$

step 3. (construction of PHD update components)
 for $j = 1, \dots, J_{k|k-1}$
 – compute

$$z_{k|k-1}^{(l)} := h_k(x_{k|k-1}^{(l)}), \quad l = 1, \dots, L,$$

$$\eta_{k|k-1}^{(j)} = \sum_{l=1}^L \mu^{(l)} z_{k|k-1}^{(l)},$$

$$P_{zz,k|k-1}^{(j)} = \frac{v_3-2}{v_3} \sum_{l=1}^L \mu^{(l)} (z_{k|k-1}^{(l)} - \eta_{k|k-1}^{(j)})(z_{k|k-1}^{(l)} - \eta_{k|k-1}^{(j)})^T + \frac{v_2(v_3-2)}{v_3(v_2-2)} R_k,$$

$$P_{xz,k|k-1}^{(j)} = \frac{v_3-2}{v_3} \sum_{l=1}^L \mu^{(l)} (x_{k|k-1}^{(l)} - m_{k|k-1}^{(j)})(z_{k|k-1}^{(l)} - \eta_{k|k-1}^{(j)})^T$$

$$K_k^{(j)} = P_{xz,k|k-1}^{(j)} [P_{zz,k|k-1}^{(j)}]^{-1},$$

$$P_{k|k}^{(j)} = P_{k|k-1}^{(j)} - P_{xz,k|k-1}^{(j)} [P_{zz,k|k-1}^{(j)}]^{-1} [P_{xz,k|k-1}^{(j)}]^T,$$
 end

step 4. (update) **and step 5.** (moment matching)
 follow Step 4. and Step 5. of table to obtain $\{w_k^{(i)}, m_k^{(i)}, P_k^{(i)}\}_{i=1}^{J_k}$.

output $\{w_k^{(i)}, m_k^{(i)}, P_k^{(i)}\}_{i=1}^{J_k}$.

Appendix B

Proof of Proposition 1:

$$\begin{aligned}
 D_{k|k-1}(x) &= \int (p_{s,k} St(x; F_{k-1}u, Q_{k-1}, v_1) + \sum_{l=1}^{J_{\beta,k}} w_{\beta,k}^{(l)} St(x; F_{\beta,k-1}^{(l)}u + d_{\beta,k-1}^{(l)}, Q_{\beta,k-1}^{(l)}, v_3)) \sum_{j=1}^{J_{k-1}} w_{k-1}^{(j)} St(u; m_{k-1}^{(j)}, P_{k-1}^{(j)}, v_3) du \\
 &\quad + \sum_{i=1}^{J_{\gamma,k}} w_{\gamma,k}^{(i)} St(x; m_{\gamma,k}^{(i)}, P_{\gamma,k}, v_3) \\
 &= p_{s,k} \sum_{j=1}^{J_{k-1}} w_{k-1}^{(j)} \int St(x; F_{k-1}u, Q_{k-1}, v_1) St(u; m_{k-1}^{(j)}, P_{k-1}^{(j)}, v_3) du \\
 &\quad + \sum_{j=1}^{J_{k-1}} \sum_{l=1}^{J_{\beta,k}} w_{k-1}^{(j)} w_{\beta,k}^{(l)} \int St(x; F_{\beta,k-1}^{(l)}u + d_{\beta,k-1}^{(l)}, Q_{\beta,k-1}^{(l)}, v_3) St(u; m_{k-1}^{(j)}, P_{k-1}^{(j)}, v_3) du \\
 &\quad + \sum_{i=1}^{J_{\gamma,k}} w_{\gamma,k}^{(i)} St(x; m_{\gamma,k}^{(i)}, P_{\gamma,k}, v_3) \\
 &= p_{s,k} \sum_{j=1}^{J_{k-1}} w_{k-1}^{(j)} St(x; F_{k-1}m_{k-1}^{(j)}, F_{k-1}P_{k-1}^{(j)} F_{k-1}^T + \frac{v_3-2}{v_3} \frac{v_1-2}{v_1-2} Q_{k-1}, v_3) \\
 &\quad + \sum_{j=1}^{J_{k-1}} \sum_{l=1}^{J_{\beta,k}} w_{k-1}^{(j)} w_{\beta,k}^{(l)} St(x; F_{\beta,k-1}^{(l)}m_{k-1}^{(j)} + d_{\beta,k-1}^{(l)}, F_{\beta,k-1}^{(l)}P_{\beta,k-1}^{(l)} (F_{\beta,k-1}^{(l)})^T + \frac{v_3-2}{v_3} \frac{v_1-2}{v_1-2} Q_{\beta,k-1}, v_3) \\
 &\quad + \sum_{i=1}^{J_{\gamma,k}} w_{\gamma,k}^{(i)} St(x; m_{\gamma,k}^{(i)}, P_{\gamma,k}, v_3).
 \end{aligned}$$

Proof of Proposition 2:

$$\begin{aligned}
 D_k(x) &= (1 - p_{D,k}) D_{k|k-1}(x) \\
 &\quad + \sum_{z \in Z_k} \frac{p_{D,k} St(z; H_k x, R_k, v_2) \sum_{j=1}^{J_{k|k-1}} w_{k|k-1}^{(j)} St(x; m_{k|k-1}^{(j)}, P_{k|k-1}^{(j)}, v_3)}{\kappa_k(z) + \int p_{D,k} St(z; H_k x, R_k, v_2) \sum_{l=1}^{J_{k|k-1}} w_{k|k-1}^{(l)} St(u; m_{k|k-1}^{(l)}, P_{k|k-1}^{(l)}, v_3) du} \\
 &= (1 - p_{D,k}) D_{k|k-1}(x) \\
 &\quad + \sum_{z \in Z_k} \frac{p_{D,k} \sum_{j=1}^{J_{k|k-1}} w_{k|k-1}^{(j)} St(z; H_k x, R_k, v_2) St(x; m_{k|k-1}^{(j)}, P_{k|k-1}^{(j)}, v_3)}{\kappa_k(z) + p_{D,k} \sum_{l=1}^{J_{k|k-1}} w_{k|k-1}^{(l)} \int St(z; H_k x, R_k, v_2) St(u; m_{k|k-1}^{(l)}, P_{k|k-1}^{(l)}, v_3) du},
 \end{aligned}$$

Using Lemma 2:

$$\begin{aligned}
 &St(z; H_k x, R_k, v_2) St(x; m_{k|k-1}^{(j)}, P_{k|k-1}^{(j)}, v_3) \\
 &= St\left(z; H_k m_{k|k-1}^{(j)}, H_k P_{k|k-1}^{(j)} H_k^T + \frac{(v_3-2)v_2}{(v_2-2)v_3} R, v_3\right) \\
 &St\left(x; m_{k|k-1}^{(j)} + K_k^{(j)}(z - H_k m_{k|k-1}^{(j)}), \frac{v_3 + \Delta^2}{v_3 + d_z} (P_{k|k-1}^{(j)} - K_k^{(j)} H_k P_{k|k-1}^{(j)}), v_3 + d_z\right) \\
 &= q_k^{(j)}(z) St(x; m_{k|k}^{(j)'}(z), P_{k|k}^{(j)'}(z), v_3'),
 \end{aligned}$$

with:

$$\begin{aligned}
 K_k^{(j)} &= P_{k|k-1}^{(j)} H_k^T (H_k P_{k|k-1}^{(j)} H_k^T + \frac{v_3-2}{v_3} \frac{v_2}{v_2-2} R_k)^{-1}, \\
 \Delta^2 &= (z - H_k m_{k|k-1}^{(j)})^T (P_{zz,k|k-1}^{(j)})^{-1} (z - H_k m_{k|k-1}^{(j)}),
 \end{aligned}$$

Using Lemma 1:

$$\int St(z; H_k x, R_k, v_2) St(u; m_{k|k-1}^{(l)}, P_{k|k-1}^{(l)}, v_3) du = St\left(z; H_k m_{k|k-1}^{(l)}, H_k P_{k|k-1}^{(l)} H_k^T + \frac{(v_3-2)v_2}{(v_2-2)v_3} R, v_3\right) = q_k^{(l)}(z),$$

so:

$$\begin{aligned}
 D_k(x) &= (1 - p_{D,k}) D_{k|k-1}(x) \\
 &\quad + \sum_{z \in Z_k} \sum_{j=1}^{J_{k|k-1}} \frac{p_{D,k} w_{k|k-1}^{(j)} q_k^{(j)}(z) St(x; m_{k|k}^{(j)'}(z), P_{k|k}^{(j)'}(z), v_3')}{\kappa_k(z) + p_{D,k} \sum_{l=1}^{J_{k|k-1}} w_{k|k-1}^{(l)} q_k^{(l)}(z)}.
 \end{aligned}$$

References

1. Mahler, R.P.S. *Advances in Statistical Multisource-Multitarget Information Fusion*; Artech House: Norwood, MA, USA, 2014.
2. Bar-Shalom, Y.; Kirubarajan, T.; Li, X.R. *Estimation with Applications to Tracking and Navigation*; Wiley: Hoboken, NJ, USA, 2001; pp. 993–999.
3. Mahler, R.P.S. Multitarget Bayes Filtering via First-Order Multitarget Moments. *IEEE Trans. Aerosp. Electron. Syst.* **2004**, *39*, 1152–1178. [[CrossRef](#)]
4. Mahler, R.P.S. PHD Filters of Higher Order in Target Number. *IEEE Trans. Aerosp. Electron. Syst.* **2007**, *43*, 1523–1543. [[CrossRef](#)]
5. Vo, B.T.; Vo, B.N.; Cantoni, A. The Cardinality Balanced Multi-Target Multi-Bernoulli Filter and its Implementations. *IEEE Trans. Signal Process.* **2009**, *57*, 409–423.
6. Vo, B.T.; Vo, B.N. Labeled Random Finite Sets and Multi-Object Conjugate Priors. *IEEE Trans. Signal Process.* **2013**, *61*, 3460–3475. [[CrossRef](#)]
7. Reuter, S.; Vo, B.T.; Vo, B.N.; Dietmayer, K. The Labeled Multi-Bernoulli Filter. *IEEE Trans. Signal Process.* **2014**, *62*, 3246–3260.
8. Vo, B.N.; Singh, S.; Doucet, A. Sequential Monte Carlo Methods for Multitarget Filtering with Random Finite Sets. *IEEE Trans. Aerosp. Electron. Syst.* **2005**, *41*, 1224–1245.
9. Vo, B.N.; Ma, W.K. The Gaussian Mixture Probability Hypothesis Density Filter. *IEEE Trans. Signal Process.* **2006**, *54*, 4091–4104. [[CrossRef](#)]
10. Macagnano, D.; De Abreu, G.T.F. Adaptive Gating for Multitarget Tracking with Gaussian Mixture Filters. *IEEE Trans. Signal Process.* **2012**, *60*, 1533–1538. [[CrossRef](#)]
11. Macagnano, D.; De Abreu, G.T.F. Multitarget Tracking with the Cubature Kalman Probability Hypothesis Density Filter. In Proceedings of the International Conference on Signals, Systems and Computers, Pacific Grove, CA, USA, 7–10 November 2010; pp. 1455–1459.
12. Zhang, L.; Wang, T.; Zhang, F.; Xu, D. Cooperative Localization for Multi-Auvs Based On GM-PHD Filters and Information Entropy Theory. *Sensors* **2017**, *17*, 2286. [[CrossRef](#)] [[PubMed](#)]
13. Zhang, Q.; Song, T.L. Improved Bearings-Only Multi-Target Tracking with GM-PHD Filtering. *Sensors* **2016**, *16*, 1469. [[CrossRef](#)] [[PubMed](#)]
14. Yang, F.; Wang, Y.; Chen, H.; Zhang, P.; Liang, Y. Adaptive Collaborative Gaussian Mixture Probability Hypothesis Density Filter for Multi-Target Tracking. *Sensors* **2016**, *16*, 1666. [[CrossRef](#)] [[PubMed](#)]
15. Vo, B.T.; Vo, B.N.; Hoseinnezhad, R.; Mahler, R.P.S. Robust Multi-Bernoulli Filtering. *IEEE J. Sel. Top. Signal Process.* **2013**, *7*, 399–409. [[CrossRef](#)]
16. Huber, P.J.; Ronchetti, E.M. *Robust Statistics*, 2nd ed.; John Wiley & Sons: Hoboken, NJ, USA, 2009.
17. Šmídl, V.; Quinn, A. *The Variational Bayes Method in Signal Processing*; Springer: Heidelberg, Germany, 2006.
18. Roth, M.; Özkan, E.; Gustafsson, F. A Student's T Filter for Heavy Tailed Process and Measurement Noise. In Proceedings of the 38th International Conference on Acoustics, Speech and Signal Processing (ICASSP), Vancouver, BC, Canada, 26–31 May 2013; pp. 5770–5774.
19. Huang, Y.; Zhang, Y.; Li, N.; Chambers, J. Robust Student's T Based Nonlinear Filter and Smoother. *IEEE Trans. Aerosp. Electron. Syst.* **2017**, *52*, 2586–2596. [[CrossRef](#)]
20. Straka, O.; Dunik, J. Stochastic Integration Student's-T Filter. In Proceedings of the 20th International Conference on Information Fusion (FUSION), Xi'an, China, 10–13 July 2017; pp. 1–8.
21. Tronarp, F.; Hostettler, R.; Särkkä, S. Sigma-Point Filtering for Nonlinear Systems with Non-Additive Heavy-Tailed Noise. In Proceedings of the 19th International Conference on Information Fusion (FUSION), Heidelberg, Germany, 5–8 July 2016; pp. 1859–1866.
22. Huang, Y.; Zhang, Y.; Li, N.; Naqvi, S.M.; Chambers, J. A Robust Student's T Based Cubature Filter. In Proceedings of the 19th International Conference on Information Fusion (FUSION), Heidelberg, Germany, 5–8 July 2016; pp. 9–16.
23. Kotz, S.; Nadarajah, S. *Multivariate T-Distributions and their Applications*; Cambridge University Press: Cambridge, UK, 2004.
24. Vo, B.T.; Vo, B.N.; Cantoni, A. Analytic Implementations of the Cardinalized Probability Hypothesis Density Filter. *IEEE Trans. Signal Process.* **2007**, *55*, 3553–3567. [[CrossRef](#)]

25. Arasaratnam, I.; Haykin, S. Cubature Kalman Filters. *IEEE Trans. Autom. Control* **2009**, *54*, 1254–1269. [[CrossRef](#)]
26. Schuhmacher, D.; Vo, B.T.; Vo, B.N. A Consistent Metric for Performance Evaluation of Multi-Object Filters. *IEEE Trans. Signal Process.* **2008**, *56*, 3447–3457. [[CrossRef](#)]
27. Chang, G. Robust Kalman Filtering Based on Mahalanobis Distance as Outlier Judging Criterion. *J. Geod.* **2014**, *88*, 391–401. [[CrossRef](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).