sensors

MDPI

Article

A Robust SMC-PHD Filter for Multi-Target Tracking with Unknown Heavy-Tailed Measurement Noise

Yang Gong * and Chen Cui

Institute of Electronic Countermeasure, National University of Defense Technology, Hefei 230037, China; ky13285650152@163.com

* Correspondence: gongyang17@nudt.edu.cn; Tel.: +86-131-5652-7915

Abstract: In multi-target tracking, the sequential Monte Carlo probability hypothesis density (SMC-PHD) filter is a practical algorithm. Influenced by outliers under unknown heavy-tailed measurement noise, the SMC-PHD filter suffers severe performance degradation. In this paper, a robust SMC-PHD (RSMC-PHD) filter is proposed. In the proposed filter, Student-t distribution is introduced to describe the unknown heavy-tailed measurement noise where the degrees of freedom (DOF) and the scale matrix of the Student-t distribution are respectively modeled as a Gamma distribution and an inverse Wishart distribution. Furthermore, the variational Bayesian (VB) technique is employed to infer the unknown DOF and scale matrix parameters while the recursion estimation framework of the RSMC-PHD filter is derived. In addition, considering that the introduced Student- t distribution might lead to an overestimation of the target number, a strategy is applied to modify the updated weight of each particle. Simulation results demonstrate that the proposed filter is effective with unknown heavy-tailed measurement noise.

Keywords: multi-target tracking; SMC-PHD filter; student-t distribution; variational Bayesian

Citation: Gong, Y.; Cui, C. A Robust SMC-PHD Filter for Multi-Target Tracking with Unknown Heavy-Tailed Measurement Noise. *Sensors* **2021**, *21*, 3611. https://doi.org/10.3390/s21113611

Academic Editor: Taek Lyul Song

Received: 7 March 2021
Accepted: 20 May 2021
Published: 22 May 2021

1. Introduction

Multi-target tracking (MTT) involves the estimate of target number and states from noisy measurements [1]. Influenced by clutter and detection uncertainty, the traditional MTT algorithms based on data association, such as joint probabilistic data association (JPDA) [2] and multiple hypotheses tracking (MHT) [3], are burdened with high computation complexity with the increase of target number and clutter number. Under the framework of finite set statistics (FISST), multi-target states and measurements are modeled as random finite sets (RFS) to avoid data association. Unfortunately, the MTT approaches based on RFS are still computationally intractable because of the multiple integrals. To obtain a computational tractable solution, the probability hypothesis density (PHD) filter [4] which propagates the first order moment approximation of multi-target posterior density is proposed and realizes the estimate of multi-target number and states. The two main implementations of the PHD filter are the Gaussian mixture PHD (GM-PHD) filter [5] and the sequential Monte Carle PHD (SMC-PHD) [6]. In recent years, the PHD filter has generated substantial interest in visual tracking [7], radar tracking [8], etc.

However, the above algorithms only achieve good performance with the known Gaussian measurement noise variance. When the measurement noise variance is unknown and inconsistent with the true variance, the performance of these algorithms decreases. In [9], Sarkka considers the unknown measurement noise variance problem under the linear Gaussian system and models noise variance as an inverse Gamma (IG) distribution. Then, the joint estimate of target state and noise variance is obtained by variational Bayesian (VB) technique. Since the inverse Gamma distribution is only suitable for the situation where the noise variance is diagonal, the inverse Wishart (IW) distribution is further used to model the full noise variance [10,11]. Following this, the VB technique was introduced into the PHD filter [12–14] for MTT.

In some applications, the measurement noise might be non-Gaussian. For example, due to the electromagnetic interference and the sensor's unreliability, the measurement model is usually accompanied with outliers and the measurement noise can be treated as non-Gaussian noise with a heavy-tailed probability density function [15]. The outliers will result in the performance degradation of conventional tracking algorithms. In [15], the combination of a Laplacian distribution and a Gaussian distribution is used to describe the heavy-tailed measurement noise. In [16,17], the heavy-tailed measurement noise is characterized as a mixture of two Gaussian distributions. Due to the advantage of the heavy-tailed characteristic, Student-t distribution was widely used to deal with outliers in [18–20]. However, the relevant parameters of Student-t distribution need to be known in these applications. Similar to [9–14], the VB technique was employed in the situation where the Student-t distribution parameters are unknown in [21–24]. The joint estimate of target and unknown parameters is obtained in a linear system. However, the above approaches work worse in a nonlinear system. The nonlinear transformation algorithms, such as extended Kalman filter (EKF) and unscented Kalman filter (UKF), can be used to cope with the nonlinear system to some extent. Unfortunately, the EKF depends on the nature of nonlinearities and the UKF relies on the Gaussian characteristic of the system. They often give an unreliable estimate if the nonlinearities are severe or the system is non-Gaussian, which leads to target missing.

With heavy-tailed measurement noise, the robust PHD filter in a nonlinear system has been much less studied than that in a linear system. Compared with nonlinear transformations, the particle filter is more suitable for practical applications because it can deal with nonlinear and non-Gaussian dynamics. Motivated by [23], we extend the Student-t distribution into SMC-PHD filter and propose a robust SMC-PHD (RSMC-PHD) filter with unknown heavy-tailed measurement noise. In the proposed filter, the Student-t distribution where the degrees of freedom (DOF) and the scale matrix are both unknown is used to model the measurement noise. The Gamma distribution and IW distribution are respectively used to model the DOF and the scale matrix. Then, the augmented state of the target involves the target state, DOF, and the scale matrix. Further, the target state is estimated by particle approximations and the DOF and the scale matrix are updated by the VB technique within the SMC-PHD filter framework. In addition, considering the problem that the introduced Student-t distribution might lead to an overestimation of the target number, we apply a strategy to modify the updated weight of each particle. Simulation results show that the proposed RSMC-PHD filter can achieve higher estimate accuracy than the compared filter under unknown heavy-tailed measurement noise.

This paper is organized as follows. Section 2 gives a brief overview of the tracking model, measurement noise model, PHD filter, and SMC-PHD filter. Section 3 presents the detailed implementation of the proposed filter. Section 4 evaluates the performance of the proposed filter under two scenarios. Finally, Section 5 displays the conclusions.

2. Background

2.1. Tracking Model and Measurement Noise Model

Consider a state space system given by

$$\boldsymbol{x}_k = f(\boldsymbol{x}_{k-1}) + \boldsymbol{w}_k \tag{1}$$

$$\boldsymbol{z}_k = h(\boldsymbol{x}_k) + \boldsymbol{\varepsilon}_k \tag{2}$$

where $\boldsymbol{x}_k$ is referred to as the target state and $\boldsymbol{z}_k$ is referred to as the measurement at time k. $\boldsymbol{w}_k$ is the process noise assumed to be Gaussian white noise with covariance matrix $\boldsymbol{Q}_k$. $f(\cdot)$ and $h(\cdot)$ are the dynamic motion model function and measurement model function, respectively. $\boldsymbol{\varepsilon}_k$ is the measurement noise.

The target dynamic motion model function is given by

$$f(\boldsymbol{x}_k|\boldsymbol{x}_{k-1}) = \mathcal{N}(\boldsymbol{x}_k; f(\boldsymbol{x}_{k-1}), \boldsymbol{Q}_k) \tag{3}$$

where $\mathcal{N}(\boldsymbol{x};\hat{\boldsymbol{x}},\boldsymbol{P})$ denotes a Gaussian probability density function with mean $\hat{\boldsymbol{x}}$ and covariance matrix $\boldsymbol{P}$. Assume the measurement noise has a heavy tail and can be described by a Student-t distribution as [18]

$$\begin{aligned} p(\boldsymbol{\varepsilon}_k) &= \mathrm{St}(\boldsymbol{\varepsilon}_k; 0_d, \boldsymbol{R}_k, \nu_k) \\ &= \frac{\Gamma\left(\frac{\nu_k+d}{2}\right)}{\Gamma\left(\frac{\nu_k}{2}\right)} \frac{1}{(\nu_k\pi)^{\frac{d}{2}}} \frac{1}{|\boldsymbol{R}_k|^{\frac{1}{2}}} \left(1+\frac{1}{\nu_k}\boldsymbol{\varepsilon}_k^T \boldsymbol{R}_k{}^{-1}\boldsymbol{\varepsilon}_k\right)^{-\frac{\nu_k+d}{2}} \end{aligned} \tag{4}$$

where $\mathrm{St}(\cdot)$ denotes the Student-t distribution, 0_d denotes d dimension zeros vector, d is the dimension of $\boldsymbol{z}_k$, ν_k denotes DOF, $\boldsymbol{R}_k$ is a symmetric, positive defined $d \times d$ covariance matrix and $\Gamma(\cdot)$ denotes the Gamma function. Different from the Gaussian distribution, the Student-t distribution has a thicker tail away from the mean value. When $\nu_k \to \infty$, the Student-t distribution is equivalent to a Gaussian distribution. The probability density function in Equation (4) can be also formulated as a mixture of Gamma distribution and Gaussian distribution denoted as [25]

$$\mathrm{St}(\boldsymbol{\varepsilon}_k; 0, \boldsymbol{R}_k, \nu_k) = \int_0^\infty \mathcal{N}(\boldsymbol{\varepsilon}_k; 0, \boldsymbol{R}_k/u_k)\mathrm{G}\left(u_k; \frac{\nu_k}{2}, \frac{\nu_k}{2}\right) du_k \tag{5}$$

where $\mathrm{G}(\cdot)$ denotes the gamma distribution, u_k is an auxiliary variable. The likelihood of the measurement can be written as

$$g_k(\boldsymbol{z}_k|\boldsymbol{x}_k, \boldsymbol{R}_k, \nu_k) = \mathrm{St}(\boldsymbol{z}_k; h(\boldsymbol{x}_k), \boldsymbol{R}_k, \nu_k) \tag{6}$$

According to Equation (5), we can obtain

$$g_k(\boldsymbol{z}_k|\boldsymbol{x}_k, \boldsymbol{R}_k, \mu_k) = \mathcal{N}(\boldsymbol{z}_k; h(\boldsymbol{x}_k), \boldsymbol{R}_k/u_k) \tag{7}$$

where u_k is distributed as

$$p(u_k|\nu_k) = \mathrm{G}(u_k; \nu_k/2, \nu_k/2) \tag{8}$$

In this paper, the DOF ν_k and scale matrix $\boldsymbol{R}_k$ are both unknown. Considering the conjugate prior to Gaussian distribution with unknown variance is the IW distribution, we model $\boldsymbol{R}_k$ as the IW distribution. Similarly, the Gamma distribution is assumed to be the conjugate prior to the DOF ν_k. Then the joint probability density of $\boldsymbol{R}_k$, u_k, and ν_k can be expressed as

$$\begin{aligned} p(\boldsymbol{R}_k, u_k, \nu_k) &= \mathrm{IW}(\boldsymbol{R}_k; \upsilon_k, \boldsymbol{V}_k)\mathrm{G}(u_k; \nu_k/2, \nu_k/2) \\ &\quad \times \mathrm{G}(\nu_k; \alpha_k, \beta_k) \end{aligned} \tag{9}$$

where $\mathrm{IW}(\cdot)$ denotes the IW distribution, υ_k, $\boldsymbol{V}_k$ are the sufficient statistics of $\boldsymbol{R}_k$, α_k and β_k are the sufficient statistics of ν_k.

2.2. PHD Filter

Assume that $\boldsymbol{X}_k = \left\{\boldsymbol{x}_k^1, \cdots, \boldsymbol{x}_k^{N_k}\right\}$ is the multi-target states set and $\boldsymbol{Z}_k = \left\{\boldsymbol{z}_k^1, \cdots, \boldsymbol{z}_k^{M_k}\right\}$ is the measurements set at time k. $\boldsymbol{x}_k^n$ is referred to as the nth target state and $\boldsymbol{z}_k^m$ is referred to as the mth measurement. N_k and M_k are the number of targets and measurements, respectively. Then the Bayesian filter recursion can be expressed as

$$p_{k|k-1}(\boldsymbol{X}_k|\boldsymbol{Z}_{1:k-1}) = \int f(\boldsymbol{X}_k|\boldsymbol{X}_{k-1})p_{k-1}(\boldsymbol{X}_{k-1}|\boldsymbol{Z}_{1:k-1})\mu_s d\boldsymbol{X}_{k-1} \tag{10}$$

$$p_k(\boldsymbol{X}_k|\boldsymbol{Z}_{1:k}) = \frac{g_k(\boldsymbol{Z}_k|\boldsymbol{X}_k)p_{k|k-1}(\boldsymbol{X}_k|\boldsymbol{Z}_{1:k-1})}{\int g_k(\boldsymbol{Z}_k|\boldsymbol{X}_k)p_{k|k-1}(\boldsymbol{X}_k|\boldsymbol{Z}_{1:k-1})\mu_s d\boldsymbol{X}_k} \tag{11}$$

where $g_k(\cdot)$ is the likelihood function. $p_{k|k-1}(\cdot)$ and $p_k(\cdot)$ denote multi-target predicted density and multi-target posterior density, respectively. μ_s is the reference measure.

The recursion Equations (10) and (11) involve multiple integrals, which are computationally intractable. To alleviate the computational intractability, the PHD filter provides a sub-optimal solution by propagating the first-order statistical moment of the multiple posterior densities. Without the spawning targets, the predicted and updated equations of the PHD filter are given by [5]

$$D_{k|k-1}(\boldsymbol{x}_k) = \int p_{S,k} f(\boldsymbol{x}_k|\boldsymbol{x}_{k-1}) D_{k-1}(\boldsymbol{x}_{k-1}) d\boldsymbol{x}_{k-1} + \gamma_k(\boldsymbol{x}_k) \tag{12}$$

$$D_k(\boldsymbol{x}_k) = (1 - p_{D,k}) D_{k|k-1}(\boldsymbol{x}_k) + \sum_{z \in \boldsymbol{Z}_k} \frac{p_{D,k} g_k(z|\boldsymbol{x}_k) D_{k|k-1}(\boldsymbol{x}_k)}{\kappa_k(z) + \int p_{D,k} g_k(z|\boldsymbol{x}_k) D_{k|k-1}(\boldsymbol{x}_k) d\boldsymbol{x}_k} \tag{13}$$

where $D_{k|k-1}(\boldsymbol{x}_k)$ and $D_{k|k}(\boldsymbol{x}_k)$ are the predicted intensity and posterior intensity, respectively. $\gamma_k(\boldsymbol{x}_k)$ is the birth targets intensity, $p_{S,k}$ is the target survival probability, $p_{D,k}$ is the detection probability. $\kappa_k(z) = \lambda c(z)$ is the clutter intensity with average clutter number λ and spatial distribution $c(z)$.

2.3. SMC-PHD Filter

Assume that the multi-target posterior intensity can be denoted by weighted particles set at time $k-1$, i.e., $D_{k-1}(\boldsymbol{x}_{k-1}|\boldsymbol{Z}_{1:k-1}) = \sum_{i=1}^{L_{k-1}} \omega_{k-1}^{(i)} \delta(\boldsymbol{x}_{k-1} - \boldsymbol{x}_{k-1}^{(i)})$. L_{k-1} denotes the particle number at time $k-1$, $\boldsymbol{x}_{k-1}^{(i)}$ is the ith particle state and $\omega_{k-1}^{(i)}$ is the corresponding weight. Then, the SMC-PHD filter can be summarized by the following steps.

Step 1 Prediction

$$\boldsymbol{x}_k^{(i)} \sim \begin{cases} q_k(\cdot \left| \boldsymbol{x}_{k-1}^{(i)}, \boldsymbol{Z}_k \right.), i = 1, \ldots, L_{k-1} \\ p_k(\cdot|\boldsymbol{Z}_k), i = L_{k-1}+1, \ldots, L_{k-1}+J_k \end{cases} \tag{14}$$

$$\omega_{k|k-1}^{(i)} \sim \begin{cases} \omega_{k-1}^{(i)} \frac{p_{S,k} f_{k|k-1}(\boldsymbol{x}_k^{(i)} \left| \boldsymbol{x}_{k-1}^{(i)} \right.)}{q_k(\boldsymbol{x}_k^{(i)} \left| \boldsymbol{x}_{k-1}^{(i)}, \boldsymbol{Z}_k \right.)}, i = 1, \ldots, L_{k-1} \\ \frac{1}{J_k} \frac{\gamma_k(\boldsymbol{x}_k^{(i)})}{p_k(\boldsymbol{x}_k^{(i)}|\boldsymbol{Z}_k)}, i = L_{k-1}+1, \ldots, L_{k-1}+J_k \end{cases} \tag{15}$$

$$D_{k|k-1}(\boldsymbol{x}_k|\boldsymbol{Z}_{1:k-1}) = \sum_{i=1}^{L_{k-1}+J_k} \omega_{k|k-1}^{(i)} \delta(\boldsymbol{x}_k - \boldsymbol{x}_k^{(i)}) \tag{16}$$

where $q_k\left(\cdot \left| \boldsymbol{x}_{k-1}^{(i)}, \boldsymbol{Z}_k \right.\right)$ is the proposal density of survival targets and $p_k(\cdot|\boldsymbol{Z}_k)$ is the proposal density of birth targets. J_k is the particles number of birth targets.

Step 2 Update

$$\omega_k^{(i)} = (1 - p_{D,k}) \omega_{k|k-1}^{(i)} + \sum_{z \in \boldsymbol{Z}_k} \frac{p_{D,k} g_k(z \left| \boldsymbol{x}_k^{(i)} \right.)}{\kappa_k(z) + C_k(z)} \omega_{k|k-1}^{(i)} \tag{17}$$

$$C_k(z) = \sum_{j=1}^{L_{k-1}+J_k} p_{D,k} g_k(z \left| \boldsymbol{x}_k^{(i)} \right.) \omega_{k|k-1}^{(i)} \tag{18}$$

$$D_k(\boldsymbol{x}_k|\boldsymbol{Z}_{1:k}) = \sum_{i=1}^{L_{k-1}+J_k} \omega_k^{(i)} \delta(\boldsymbol{x}_k - \boldsymbol{x}_k^{(i)}) \tag{19}$$

Step 3 Resampling

Target number is estimated by $\hat{N}_k = \text{ceil}\left(\sum_{i=1}^{L_{k-1}+J_k} \omega_k^{(i)}\right)$, and resample $\left\{\boldsymbol{x}_k^{(i)}, \omega_k^{(i)}\right\}_{i=1}^{L_{k-1}+J_k}$. $\text{ceil}(x)$ denotes the smallest integer more than x, $L_k = N_{particle} \cdot \hat{N}_k$, $N_{particle}$ is the number of particles per expected target.

Step 4 State extraction

Extract the target states $\boldsymbol{X}_k = \left\{\hat{\boldsymbol{x}}_k^1, \cdots, \hat{\boldsymbol{x}}_k^{\hat{N}_k}\right\}$ from the resampling particles set, where $\hat{\boldsymbol{x}}_k^1, \cdots, \hat{\boldsymbol{x}}_k^{\hat{N}_k}$ denote the estimated target states.

3. The Robust SMC-PHD Filter

3.1. VB Approximation

The aforementioned SMC-PHD filter performs well under Gaussian measurement noise. However, the outliers in measurements will have a negative effect on the SMC-PHD filter. To solve the problem, the Student-t distribution is introduced to model the heavy-tailed measurement noise. Since the parameters of the Student-t distribution are unknown, the joint posterior probability density of target state $\boldsymbol{x}_k$, scale matrix $\boldsymbol{R}_k$, auxiliary variable u_k, and DOF ν_k need to be estimated. Therefore, the augmented target state can be expressed as

$$\ddot{\boldsymbol{x}}_k = (\boldsymbol{x}_k, \boldsymbol{R}_k, u_k, \nu_k) \tag{20}$$

Since there is no explicit analytical solution existing for the joint posterior density $p(\ddot{\boldsymbol{x}}_k|\boldsymbol{Z}_{1:k})$, we utilize the VB inference [9] to obtain the approximate solution. As for an approximate method, VB inference is widely used in searching for a factored approximate posterior density as follows

$$\begin{aligned} p(\ddot{\boldsymbol{x}}_k|\boldsymbol{Z}_{1:k}) &= p(\boldsymbol{x}_k, \boldsymbol{R}_k, u_k, \nu_k|\boldsymbol{Z}_{1:k}) \\ &\approx q(\ddot{\boldsymbol{x}}_k|\boldsymbol{Z}_{1:k}) \triangleq q_x(\boldsymbol{x}_k) q_{\boldsymbol{R}}(\boldsymbol{R}_k) q_u(u_k) q_v(\nu_k) \end{aligned} \tag{21}$$

where $q(\ddot{\boldsymbol{x}}_k|\boldsymbol{Z}_{1:k})$ is the approximate joint posterior density, $q_{\boldsymbol{x}}(\cdot)$, $q_{\boldsymbol{R}}(\cdot)$, $q_u(\cdot)$,and $q_v(\cdot)$ are the approximate posterior densities. Equation (21) has the implicit assumption that the target state and the Student-t distribution parameters are independent. By minimizing the Kullback–Leibler (KL) divergence between $p(\ddot{\boldsymbol{x}}_k|\boldsymbol{Z}_{1:k})$ and $q(\ddot{\boldsymbol{x}}_k|\boldsymbol{Z}_{1:k})$, the estimate $\hat{q}_{\boldsymbol{x}}(\boldsymbol{x}_k)$, $\hat{q}_{\boldsymbol{R}}(\boldsymbol{R}_k)$, $\hat{q}_u(u_k)$, and $\hat{q}_v(\nu_k)$ can be easily obtained by

$$\{\hat{q}_{\boldsymbol{x}}, \hat{q}_{\boldsymbol{R}}, \hat{q}_u, \hat{q}_v\} = \underset{q_{\boldsymbol{x}}, q_{\boldsymbol{R}}, q_u, q_v}{\operatorname{argmin}} KL\left[q(\{\ddot{\boldsymbol{x}}_k\}|\boldsymbol{Z}_{1:k}) \middle\| p(\{\ddot{\boldsymbol{x}}_k\}|\boldsymbol{Z}_{1:k})\right] \tag{22}$$

where $KL(q(x)|p(x)) \triangleq \int q(x) \log\left(\frac{q(x)}{p(x)}\right) dx$. The optimal solution of Equation (22) can be obtained by [9,21,26]

$$\ln \hat{q}(\phi) = E_{\ddot{\boldsymbol{x}}_k^{(-\phi)}}\left[\ln p(\ddot{\boldsymbol{x}}_k, \boldsymbol{Z}_k|\boldsymbol{Z}_{1:k-1})\right] + c_\phi \tag{23}$$

where ϕ denotes an arbitrary element in $\ddot{\boldsymbol{x}}_k$, $E[\cdot]$ denotes the expectation, $\ddot{\boldsymbol{x}}_k^{(-\phi)}$ denotes the elements set in $\ddot{\boldsymbol{x}}_k$ except for ϕ, and c_ϕ is an irrelevant constant for element ϕ.

3.2. The Update of Unknown Student-t Distribution Parameters

Based on the assumption that the target state is independent of the Student-t distribution parameters, the joint posterior density of the augmented target state can be factorized as

$$p(\ddot{\boldsymbol{x}}_k|\boldsymbol{Z}_{1:k}) = p(\boldsymbol{R}_k, u_k, \nu_k|\boldsymbol{Z}_{1:k}) p(\boldsymbol{x}_k|\boldsymbol{Z}_{1:k}) \tag{24}$$

where the unknown Student-t distribution parameters are integrated out. Based on (24), the target state can be estimated by particle samples and the Student-t distribution parameters can be updated by the VB approximation method.

In this subsection, we derive the update of Student-t distribution parameters for each particle. At time k, let $\ddot{\boldsymbol{x}}_k = \left\{\ddot{\boldsymbol{x}}_k^{(i)}\right\}_{i=1}^N = \left\{\boldsymbol{x}_k^{(i)}, \boldsymbol{R}_k^{(i)}, u_k^{(i)}, \nu_k^{(i)}\right\}_{i=1}^N$ is the augmented target state where $\boldsymbol{R}_k^{(i)}$, $u_k^{(i)}$, $\nu_k^{(i)}$ are the Student-t distribution parameters with respect to the ith

particle, and N is the number of particles. Assume the predicted density of $\ddot{\boldsymbol{x}}_k^{(i)}$ at time k can be expressed as

$$\begin{aligned} p(\ddot{\boldsymbol{x}}_k^{(i)}|\boldsymbol{Z}_{1:k-1}) &= \omega_{k|k-1}^{(i)}\delta(\boldsymbol{x}_k-\boldsymbol{x}_k^{(i)})\mathrm{IW}(\boldsymbol{R}_k^{(i)};v_{k|k-1}^{(i)},\boldsymbol{V}_{k|k-1}^{(i)}) \\ &\quad\times\mathrm{G}\left(u_k^{(i)};\tfrac{\nu_k^{(i)}}{2},\tfrac{\nu_k^{(i)}}{2}\right)\mathrm{G}(\nu_k^{(i)};\alpha_{k|k-1}^{(i)},\beta_{k|k-1}^{(i)}) \end{aligned} \tag{25}$$

where $v_{k|k-1}^{(i)}$, $\boldsymbol{V}_{k|k-1}^{(i)}$, $\alpha_{k|k-1}^{(i)}$ and $\beta_{k|k-1}^{(i)}$ are the predicted sufficient statistics with respect to $\boldsymbol{R}_k^{(i)}$ and $\nu_k^{(i)}$. Then the joint density of $p(\ddot{\boldsymbol{x}}_k^{(i)},\boldsymbol{z}_k|\boldsymbol{Z}_{1:k-1})$ can be expressed as

$$\begin{aligned} p(\ddot{\boldsymbol{x}}_k^{(i)},\boldsymbol{z}_k|\boldsymbol{Z}_{1:k-1}) = \quad & \omega_{k|k-1}^{(i)}\delta(\boldsymbol{x}_k-\boldsymbol{x}_k^{(i)})\mathcal{N}\left(\boldsymbol{z}_k;h(\boldsymbol{x}_k^{(i)}),\tfrac{\boldsymbol{R}_k^{(i)}}{u_k^{(i)}}\right) \\ & \times\mathrm{IW}(\boldsymbol{R}_k^{(i)};v_{k|k-1}^{(i)},\boldsymbol{V}_{k|k-1}^{(i)})\mathrm{G}\left(u_k^{(i)};\tfrac{\nu_k^{(i)}}{2},\tfrac{\nu_k^{(i)}}{2}\right) \\ & \times\mathrm{G}(\nu_k^{(i)};\alpha_{k|k-1}^{(i)},\beta_{k|k-1}^{(i)}) \end{aligned} \tag{26}$$

where $\boldsymbol{z}_k$ is the measurement with respect to the target represented by particle $\boldsymbol{x}_k^{(i)}$. Based on Equation (26), $\ln p(\ddot{\boldsymbol{x}}_k^{(i)},\boldsymbol{z}_k|\boldsymbol{Z}_{1:k-1})$ can be expressed as

$$\begin{aligned} \ln p(\ddot{\boldsymbol{x}}_k^{(i)},\boldsymbol{z}_k|\boldsymbol{Z}_{1:k-1}) &= -\tfrac{1}{2}\ln\left|\boldsymbol{R}_k^{(i)}\right| + \tfrac{d}{2}\ln u_k^{(i)} - \tfrac{u_k^{(i)}}{2}tr\left((\boldsymbol{R}_k^{(i)})^{-1}\boldsymbol{\Xi}_k^{(i)}\right) - \tfrac{v_{k|k-1}^{(i)}}{2}\ln\left|\boldsymbol{R}_k^{(i)}\right| \\ &\quad -\tfrac{1}{2}tr\left((\boldsymbol{R}_k^{(i)})^{-1}\boldsymbol{V}_{k|k-1}^{(i)}\right) + \tfrac{\nu_k^{(i)}}{2}\ln\tfrac{\nu_k^{(i)}}{2} - \ln\Gamma(\tfrac{\nu_k^{(i)}}{2}) + (\tfrac{\nu_k^{(i)}}{2}-1)\ln u_k^{(i)} \\ &\quad -\tfrac{\nu_k^{(i)}}{2}u_k^{(i)} + (\alpha_{k|k-1}^{(i)}-1)\ln\nu_k^{(i)} - \beta_{k|k-1}^{(i)}\nu_k^{(i)} + c_{\ddot{\boldsymbol{x}}_k^{(i)}} \end{aligned} \tag{27}$$

where $\boldsymbol{\Xi}_k^{(i)} = (\boldsymbol{z}_k - h(\boldsymbol{x}_k^{(i)}))(\boldsymbol{z}_k - h(\boldsymbol{x}_k^{(i)}))^T$ denotes the innovation covariance matrix. $tr(\cdot)$ denotes the trace operation. $c_{\ddot{\boldsymbol{x}}_k^{(i)}}$ denotes an irrelevant constant for variables $\boldsymbol{x}_k^{(i)}$, $\boldsymbol{R}_k^{(i)}$, $u_k^{(i)}$, and $\nu_k^{(i)}$.

Letting $\phi = \boldsymbol{R}_k^{(i)}$ and using Equation (27) in Equation (23), we can obtain

$$\ln\hat{q}(\boldsymbol{R}_k^{(i)}) = -\frac{v_{k|k-1}^{(i)}+1}{2}\ln\left|\boldsymbol{R}_k^{(i)}\right| - \frac{1}{2}\left\{tr\left((\boldsymbol{R}_k^{(i)})^{-1}\left[\boldsymbol{V}_{k|k-1}^{(i)} + u_k^{(i)}\boldsymbol{\Xi}_k^{(i)}\right]\right)\right\} + c_{\boldsymbol{R}_k^{(i)}} \tag{28}$$

The posterior distribution $\hat{q}(\boldsymbol{R}_k^{(i)})$ of $\boldsymbol{R}_k^{(i)}$ is an IW distribution, i.e.,

$$\hat{q}(\boldsymbol{R}_k^{(i)}) = \mathrm{IW}(\boldsymbol{R}_k^{(i)};v_k^{(i)},\boldsymbol{V}_k^{(i)}) \tag{29}$$

with parameters

$$v_k^{(i)} = v_{k|k-1}^{(i)} + 1 \tag{30}$$

$$\boldsymbol{V}_k^{(i)} = \boldsymbol{V}_{k|k-1}^{(i)} + E[u_k^{(i)}]\boldsymbol{\Xi}_k^{(i)} \tag{31}$$

Letting $\phi = \nu_k^{(i)}$ and using Equation (27) in Equation (23), we can obtain

$$\begin{aligned} \ln\hat{q}(\nu_k^{(i)}) &= \tfrac{\nu_k^{(i)}}{2}\ln\tfrac{\nu_k^{(i)}}{2} - \ln\Gamma(\tfrac{\nu_k^{(i)}}{2}) + \tfrac{\nu_k^{(i)}}{2}\ln u_k^{(i)} \\ &\quad -\tfrac{\nu_k^{(i)}}{2}u_k^{(i)} + (\alpha_{k|k-1}^{(i)}-1)\ln\nu_k^{(i)} - \beta_{k|k-1}^{(i)}\nu_k^{(i)} + c_{\nu_k^{(i)}} \end{aligned} \tag{32}$$

Using Stirling's approximation in [27]

$$\ln\Gamma(\frac{\nu_k^{(i)}}{2}) \approx \frac{\nu_k^{(i)}-1}{2}\ln\frac{\nu_k^{(i)}}{2} - \frac{\nu_k^{(i)}}{2} \tag{33}$$

And $\ln\hat{q}(\nu_k^{(i)})$ can be expressed as

$$\begin{aligned}\ln\hat{q}(\nu_k^{(i)}) &= (\alpha_{k|k-1}^{(i)} + \tfrac{1}{2} - 1)\ln\nu_k^{(i)} \\ &-(\beta_{k|k-1}^{(i)} - \tfrac{1}{2}(1 + E[\ln u_k^{(i)}] - E[u_k^{(i)}]))\nu_k^{(i)}\end{aligned} \tag{34}$$

The posterior distribution $\hat{q}(\nu_k^{(i)})$ of $\nu_k^{(i)}$ is a Gamma distribution, i.e.,

$$\hat{q}(\nu_k^{(i)}) = \mathrm{G}(\nu_k^{(i)}; \alpha_k^{(i)}, \beta_k^{(i)}) \tag{35}$$

with parameters

$$\alpha_k^{(i)} = \alpha_{k|k-1}^{(i)} + \frac{1}{2} \tag{36}$$

$$\beta_k^{(i)} = \beta_{k|k-1}^{(i)} - \frac{1}{2}(1 + E[\ln u_k^{(i)}] - E[u_k^{(i)}]) \tag{37}$$

Letting $\phi = u_k^{(i)}$ and using Equation (27) in Equation (23), we can obtain

$$\ln\hat{q}(u_k^{(i)}) = (\frac{\nu_k^{(i)}+d}{2} - 1)\ln u_k^{(i)} - \frac{1}{2}[\nu_k^{(i)} + tr\left((\boldsymbol{R}_k^{(i)})^{-1}\boldsymbol{\Xi}_k^{(i)}\right)]u_k^{(i)} + c_{u_k^{(i)}} \tag{38}$$

The posterior distribution $\hat{q}(u_k^{(i)})$ of $u_k^{(i)}$ is a Gamma distribution, i.e.,

$$\hat{q}(u_k^{(i)}) = \mathrm{G}(u_k^{(i)}; \frac{\nu_k^{(i)}+d}{2}, \frac{1}{2}[\nu_k^{(i)} + tr\left((\boldsymbol{R}_k^{(i)})^{-1}\boldsymbol{\Xi}_k^{(i)}\right)]) \tag{39}$$

Then we can obtain

$$E[u_k^{(i)}] = a_k^{(i)}/b_k^{(i)} \tag{40}$$

$$E[\ln u_k^{(i)}] = \varphi(a_k^{(i)}) - \ln b_k^{(i)} \tag{41}$$

where $a_k^{(i)} = \frac{E[\nu_k^{(i)}]+d}{2}$, $b_k^{(i)} = \frac{1}{2}\left\{E[\nu_k^{(i)}] + tr\left((\boldsymbol{R}_k^{(i)})^{-1}\boldsymbol{\Xi}_k^{(i)}\right)\right\}$. Since the parameters in Equations (30), (31), (36), (37), (40) and (41) are coupled, it can be solved iteratively by the fixed point method in [21].

3.3. The Implementation of RSMC-PHD Filter

In this subsection, we introduce the Student-t distribution into the SMC-PHD filter and present an implementation of the robust SMC-PHD filter with unknown heavy-tailed measurement noise. Assume that the multi-target posterior intensity is represented by a weighted particles set $\left\{\omega_{k-1}^{(i)}, \boldsymbol{x}_{k-1}^{(i)}, \boldsymbol{R}_{k-1}^{(i)}, u_{k-1}^{(i)}, \nu_{k-1}^{(i)}\right\}_{i=1}^{L_{k-1}}$ at time $k-1$, i.e.,

$$\begin{aligned}D_{k-1}(\ddot{\boldsymbol{x}}_{k-1}|\boldsymbol{Z}_{1:k-1}) &= \sum_{i=1}^{L_{k-1}} \omega_{k-1}^{(i)}\delta(\boldsymbol{x}_{k-1} - \boldsymbol{x}_{k-1}^{(i)})\mathrm{IW}(\boldsymbol{R}_{k-1}^{(i)}; v_{k-1}^{(i)}, \boldsymbol{V}_{k-1}^{(i)}) \\ &\times\mathrm{G}\left(u_k^{(i)}; \frac{\nu_{k-1}^{(i)}}{2}, \frac{\nu_{k-1}^{(i)}}{2}\right)\mathrm{G}(\nu_{k-1}^{(i)}; \alpha_{k-1}^{(i)}, \beta_{k-1}^{(i)})\end{aligned} \tag{42}$$

Assume $\boldsymbol{R}_k$, u_k, and ν_k are independent and the predicted intensity at time k can be expressed as

$$\begin{aligned} D_{k|k-1}(\ddot{\boldsymbol{x}}_k|\boldsymbol{Z}_{1:k-1}) &= \sum_{i=1}^{L_{k-1}+J_k} \omega_{k|k-1}^{(i)} \delta(\boldsymbol{x}_k - \boldsymbol{x}_k^{(i)}) \mathrm{G}\left(u_k^{(i)}; \tfrac{\nu_k^{(i)}}{2}, \tfrac{\nu_k^{(i)}}{2}\right) \\ &\times \mathrm{IW}(\boldsymbol{R}_k^{(i)}; v_{k|k-1}^{(i)}, \boldsymbol{V}_{k|k-1}^{(i)}) \mathrm{G}(\nu_k^{(i)}; \alpha_{k|k-1}^{(i)}, \beta_{k|k-1}^{(i)}) \end{aligned} \tag{43}$$

where $\omega_{k|k-1}^{(i)}$ and $\boldsymbol{x}_k^{(i)}$ are given in Equations (14) and (15). There are L_{k-1} particles for survival targets and J_k particles for birth targets. Generally, the dynamic of measurement noise is unknown and we use the similar heuristic dynamic in [9] for the predicted sufficient statistics. For the particles of survival targets $i = 1, 2, \ldots, L_{k-1}$, the predicted sufficient statistics are given by

$$v_{k|k-1}^{(i)} = \rho(v_{k-1}^{(i)} - d - 1) + d + 1, \boldsymbol{V}_{k|k-1}^{(i)} = \rho \boldsymbol{V}_{k-1}^{(i)} \tag{44}$$

$$\alpha_{k|k-1}^{(i)} = \rho \alpha_{k-1}^{(i)}, \beta_{k|k-1}^{(i)} = \rho \beta_{k-1}^{(i)} \tag{45}$$

where $\rho \in (0, 1]$ is the decreasing factor. For the particles of birth targets $i = L_{k-1} + 1, L_{k-1} + 2, \ldots, L_{k-1} + J_k$, the predicted sufficient statistics are given by

$$v_{k|k-1}^{(i)} = v_{\gamma,k}^{(i)}, \boldsymbol{V}_{k|k-1}^{(i)} = \boldsymbol{V}_{\gamma,k}^{(i)} \tag{46}$$

$$\alpha_{k|k-1}^{(i)} = \rho \alpha_{k-1}^{(i)}, \beta_{k|k-1}^{(i)} = \rho \beta_{k-1}^{(i)} \tag{47}$$

where $v_{\gamma,k}^{(i)}$, $\boldsymbol{V}_{\gamma,k}^{(i)}$, $\alpha_{\gamma,k}^{(i)}$ and $\beta_{\gamma,k}^{(i)}$ are sufficient statistics with respect to the birth targets.

The posterior intensity can be expressed as

$$\begin{aligned} D_k(\ddot{\boldsymbol{x}}_k|\boldsymbol{Z}_{1:k}) &= \sum_{i=1}^{L_{k-1}+J_k} \omega_k^{(i)} \delta(\boldsymbol{x}_k - \boldsymbol{x}_k^{(i)}) \mathrm{G}\left(u_k^{(i)}; \tfrac{\nu_k^{(i)}}{2}, \tfrac{\nu_k^{(i)}}{2}\right) \\ &\times \mathrm{IW}(\boldsymbol{R}_k^{(i)}; v_k^{(i)}, \boldsymbol{V}_k^{(i)}) \mathrm{G}(\nu_k^{(i)}; \alpha_k^{(i)}, \beta_k^{(i)}) \end{aligned} \tag{48}$$

where $\omega_k^{(i)}$ is given in (17). It is worth mentioning that the likelihood function is $g_k(\boldsymbol{z}|\ddot{\boldsymbol{x}}_k) = \mathrm{St}(\boldsymbol{z}; h(\boldsymbol{x}_k^{(i)}), \boldsymbol{R}_k^{(i)}, \nu_k^{(i)})$. Then, the sufficient statistics $v_k^{(i)}$, $\boldsymbol{V}_k^{(i)}$, $\alpha_k^{(i)}$, and $\beta_k^{(i)}$ are updated by

$$\boldsymbol{R}_k^{(i)(j)} = \boldsymbol{V}_k^{(i)(j-1)} / (v_k^{(i)} - d - 1) \tag{49}$$

$$\widetilde{\boldsymbol{R}}_k^{(i)(j)} = \boldsymbol{R}_k^{(i)(j)} / u_k^{(i)(j-1)} \tag{50}$$

$$\gamma_k^{(i)(j)} = tr(\boldsymbol{\Xi}_k^{(i)} / \widetilde{\boldsymbol{R}}_k^{(i)(j)}) \tag{51}$$

$$a_k^{(i)(j)} = (\nu_k^{(i)(j-1)} + d)/2 \tag{52}$$

$$b_k^{(i)(j)} = (\nu_k^{(i)(j-1)} + \gamma_k^{(i)(j)})/2 \tag{53}$$

$$u_k^{(i)(j)} = a_k^{(i)(j)} / b_k^{(i)(j)} \tag{54}$$

$$\beta_k^{(i)(j)} = \beta_{k|k-1}^{(i)} - \frac{1}{2} - \frac{\varphi(a_k^{(i)(j)}) - \ln(b_k^{(i)(j)})}{2} + \frac{u_k^{(i)(j)}}{2} \tag{55}$$

$$\boldsymbol{V}_k^{(i)(j)} = \boldsymbol{V}_{k|k-1}^{(i)} + \boldsymbol{\Xi}_k^{(i)} u_k^{(i)(j)} \tag{56}$$

$$\nu_k^{(i)(j)} = \alpha_k^{(i)} / \beta_k^{(i)(j)} \tag{57}$$

where $(\cdot)^{(i)(j)}$ denotes the update of parameter $(\cdot)^{(i)}$ at the jth iteration. The initial values are given by $\boldsymbol{V}_k^{(i)(0)} = \boldsymbol{V}_{k|k-1}^{(i)}$, $v_k^{(i)} = v_{k|k-1}^{(i)} + 1$, $\alpha_k^{(i)} = \alpha_{k|k-1}^{(i)} + 1/2$, $\beta_k^{(i)(0)} = \beta_{k|k-1}^{(i)}$, $\nu_k^{(i)(0)} = \alpha_k^{(i)} / \beta_k^{(i)(0)}$.

It should be noted that the innovation covariance matrix $\boldsymbol{\Xi}_k^{(i)} = (\boldsymbol{z}_k - h(\boldsymbol{x}_k^{(i)}))(\boldsymbol{z}_k - h(\boldsymbol{x}_k^{(i)}))^T$ for particle $\boldsymbol{x}_k^{(i)}$ is needed in each iteration. Since there is no explicit association between measurements and targets, the measurement origin is ambiguous. Considering the fact that the measurement with larger likelihood value is more likely generated by a potential target, we choose the measurement with the largest likelihood value to calculate the innovation covariance matrix as follows

$$\boldsymbol{\Xi}_k^{(i)} = (\boldsymbol{z}_k^l - h(\boldsymbol{x}_k^{(i)}))(\boldsymbol{z}_k^l - h(\boldsymbol{x}_k^{(i)}))^T \tag{58}$$

where $l = \underset{p \in I}{\operatorname{argmax}} g_k^p(\ddot{\boldsymbol{x}}_k^{(i)})$, $I = \{1, \ldots, |\boldsymbol{Z}_k|\}$. $g_k^p(\ddot{\boldsymbol{x}}_k^{(i)})$ is the normalized likelihood function value given by

$$g_k^p(\ddot{\boldsymbol{x}}_k^{(i)}) = \frac{p_{D,k} g_k(\boldsymbol{z}_k^p | \ddot{\boldsymbol{x}}_k^{(i)})}{\kappa_k(\boldsymbol{z}_k^p) + C_k(\boldsymbol{z}_k^p)} \tag{59}$$

$$C_k(\boldsymbol{z}_k^p) = \sum_{j=1}^{L_{k-1}+J_k} p_{D,k} g_k(\boldsymbol{z}_k^p | \ddot{\boldsymbol{x}}_k^{(i)}) \omega_{k|k-1}^{(i)} \tag{60}$$

3.4. The Modification of Particles Weight

The estimate of target number relies on the updated weights of particles in the SMC-PHD filter. In our paper, the weight is updated by Equation (17) for each particle. The likelihood function is $g_k(\boldsymbol{z}|\ddot{\boldsymbol{x}}_k) = \mathrm{St}(\boldsymbol{z}; h(\boldsymbol{x}_k^{(i)}), \boldsymbol{R}_k^{(i)}, \nu_k^{(i)})$, and the updated weight of the ith particle is given by

$$\omega_k^{(i)} = (1 - p_{D,k})\omega_{k|k-1}^{(i)} + \sum_{p=1}^{|\boldsymbol{Z}_k|} g_k^p(\ddot{\boldsymbol{x}}_k^{(i)}) \omega_{k|k-1}^{(i)} \propto \mathrm{St}(\boldsymbol{z}_k^p; h(\boldsymbol{x}_k^{(i)}), \boldsymbol{R}_k^{(i)}, \nu_k^{(i)}) \tag{61}$$

Figure 1 gives the illustration of the measurement likelihood function. We can see that the Student-t distribution has a heavy-tailed characteristic and it can deal with outliers well by increasing the weights of outliers. However, the weight of clutter will increase at the same time. Hence, some clutter may be regarded as target-generated measurements, resulting in the overestimate of the target number at the current time.

To solve this problem, a strategy is adopted to modify the updated weight in this subsection. In the PHD filter, one target can produce at most one measurement and only one measurement can be attributed to one target per time step [4], named as the 'one-to-one' assumption. For each particle, the normalized measurement likelihood value $g_k^p(\ddot{\boldsymbol{x}}_k^{(i)})$ can be obtained by Equations (59) and (60). Let $J = \left\{ p = 1, \ldots, |\boldsymbol{Z}_k| \middle| g_k^p(\ddot{\boldsymbol{x}}_k^{(i)}) \geq \sigma \right\}$, σ is a preset threshold. If $|J| \geq 2$, it indicates there might be multiple measurements associated with particle $\boldsymbol{x}_k^{(i)}$. In this case, the $g_k^p(\ddot{\boldsymbol{x}}_k^{(i)})$ is modified as

$$\tilde{g}_k^p(\ddot{\boldsymbol{x}}_k^{(i)}) = \begin{cases} g_k^p(\ddot{\boldsymbol{x}}_k^{(i)}), p = l \ \ or \ \ p \notin J \\ g_k^q(\ddot{\boldsymbol{x}}_k^{(i)}), p \neq l \ \ \& \ \ p \in J \end{cases} \tag{62}$$

where $q = \underset{p \in I}{\operatorname{argmin}}\, g_k^p(\ddot{\boldsymbol{x}}_k^{(i)})$, & is a logic and operation and *or* is a logic or operation. In our paper, the measurement with the largest normalized likelihood value is regarded as the corresponding measurement generated from the target corresponding to particle $\boldsymbol{x}_k^{(i)}$, and other measurements are regarded as clutter. The proposed strategy retains the

maximum likelihood value and penalizes the other larger likelihood value to hold the 'one-to-one' assumption.

Then the modified weight can be expressed as

$$\omega_k^{(i)} = (1 - p_{D,k})\omega_{k|k-1}^{(i)} + \sum_{p=1}^{|\boldsymbol{Z}_k|} \widetilde{g}_k^p(\ddot{\boldsymbol{x}}_k^{(i)})\omega_{k|k-1}^{(i)} \tag{63}$$

The robust SMC-PHD filter is summarized in Algorithm 1.

Remark 1: *Although the heavy-tailed characteristic of the Student-t distribution can deal with outliers, the target might still miss when the weights of outliers are lower than the extracted threshold and the target number would be underestimated. In our paper, the weight modification of the particles is mainly aimed to deal with the target number overestimate problem caused by clutter at the current time.*

Algorithm 1. The robust SMC-PHD filter.

Step 1 Prediction

for $i = 1, 2, \ldots, L_{k-1}$

$\boldsymbol{x}_k^{(i)} \sim q_k(\cdot \left| \boldsymbol{x}_{k-1}^{(i)}, \boldsymbol{Z}_k \right.), \omega_{k|k-1}^{(i)} \sim \omega_{k-1}^{(i)} \frac{p_{S,k} f_{k|k-1}(\boldsymbol{x}_k^{(i)} \left| \boldsymbol{x}_{k-1}^{(i)} \right.)}{q_k(\boldsymbol{x}_k^{(i)} \left| \boldsymbol{x}_{k-1}^{(i)}, \boldsymbol{Z}_k \right.)}$

$v_{k|k-1}^{(i)} = \rho(v_{k-1}^{(i)} - d - 1) + d + 1, \boldsymbol{V}_{k|k-1}^{(i)} = \rho \boldsymbol{V}_{k-1}^{(i)}$

$\alpha_{k|k-1}^{(i)} = \rho \alpha_{k-1}^{(i)}, \beta_{k|k-1}^{(i)} = \rho \beta_{k-1}^{(i)}$

end for

for $i = L_{k-1} + 1, L_{k-1} + 2, \ldots, L_{k-1} + J_k$

$\boldsymbol{x}_k^{(i)} \sim p_k(\cdot | \boldsymbol{Z}_k), \omega_{k|k-1}^{(i)} \sim \frac{1}{J_k} \frac{\gamma_k(\boldsymbol{x}_k^{(i)})}{p_k(\boldsymbol{x}_k^{(i)} | \boldsymbol{Z}_k)}$

$v_{k|k-1}^{(i)} = v_{\gamma,k}^{(i)}, \boldsymbol{V}_{k|k-1}^{(i)} = \boldsymbol{V}_{\gamma,k}^{(i)}$

$\alpha_{k|k-1}^{(i)} = \alpha_{\gamma,k}^{(i)}, \beta_{k|k-1}^{(i)} = \beta_{\gamma,k}^{(i)}$

end for

Step 2 Update

For $i = 1, 2, \ldots, L_{k-1} + J_k$

$\boldsymbol{R}_k^{(i)} = \boldsymbol{V}_{k|k-1}^{(i)} / (v_{k|k-1}^{(i)} - d - 1), \nu_k^{(i)} = \alpha_{k|k-1}^{(i)} / \beta_{k|k-1}^{(i)}$

$g_k(z \left| \ddot{\boldsymbol{x}}_k^{(i)} \right.) = \mathrm{St}(z; h(\boldsymbol{x}_k^{(i)}), \boldsymbol{R}_k^{(i)}, \nu_k^{(i)})$

$g_k^p(\ddot{\boldsymbol{x}}_k^{(i)}) = \frac{p_{D,k} g_k(z_k^p \left| \ddot{\boldsymbol{x}}_k^{(i)} \right.)}{\kappa_k(z_k^p) + C_k(z_k^p)}$

$C_k(z_k^p) = \sum_{j=1}^{L_{k-1}+J_k} p_{D,k} g_k(z_k^p \left| \ddot{\boldsymbol{x}}_k^{(i)} \right.) \omega_{k|k-1}^{(i)}$

$I = \{1, \ldots, |\boldsymbol{Z}_k|\}, J = \left\{ p = 1, \ldots, |\boldsymbol{Z}_k| \left| g_k^p(\ddot{\boldsymbol{x}}_k^{(i)}) \geq \sigma \right. \right\}$

$l = \underset{p \in I}{\mathrm{argmax}} g_k^p(\ddot{\boldsymbol{x}}_k^{(i)}), q = \underset{p \in I}{\mathrm{argmin}} g_k^p(\ddot{\boldsymbol{x}}_k^{(i)})$

$\widetilde{g}_k^p(\ddot{\boldsymbol{x}}_k^{(i)}) = \begin{cases} g_k^p(\ddot{\boldsymbol{x}}_k^{(i)}), p = l \ \ or \ \ p \notin J \\ g_k^q(\ddot{\boldsymbol{x}}_k^{(i)}), p \neq l \ \ \& \ \ p \in J \end{cases}$

$\omega_k^{(i)} = (1 - p_{D,k})\omega_{k|k-1}^{(i)} + \sum_{p=1}^{|\boldsymbol{Z}_k|} \widetilde{g}_k^p(\ddot{\boldsymbol{x}}_k^{(i)})\omega_{k|k-1}^{(i)}$

Perform iteration initialization, set the iteration number τ, and update the parameters of measurement noise by Equations (49)–(60).

end for

Step 3 Resampling

Step 4 State extraction

Step 3 and step 4 are the same as that in the standard SMC-PHD filter.

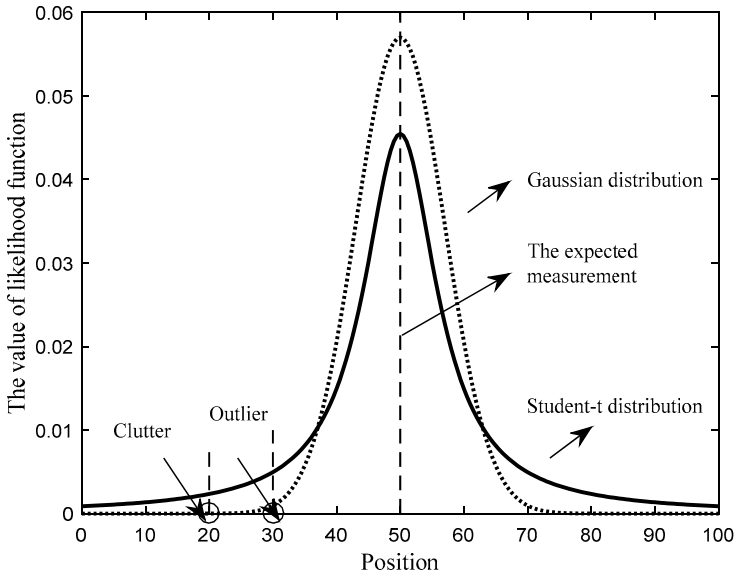

Figure 1. The illustration of measurement likelihood function.

4. Simulation Results

To reveal the performance of the proposed robust SMC-PHD (RSMC-PHD) filter, we compare the Gaussian (Normal) Gamma inverse Wishart Gamma PHD (NGIWG-PHD) filter [23] and the SMC-PHD filter [6] with the proposed filter. For each simulation, 100 Monte Carlo trials are performed. The target number and optimal sub-pattern assignment (OSPA) [28] distance are chosen as the metric, where the parameters of OSPA are $p = 2$ and $c = 100$.

4.1. The Choice of Threshold σ

In Section 3.4, threshold σ is utilized to distinguish the target-generated measurements from clutter. In the SMC-PHD filter, the underlying target relies on the updated weight of each particle, and the likelihood value in Equation (59) can directly reflect the contribution of a given measurement or clutter to the underlying target. In most cases, the target-generated measurements usually have the largest likelihood value for each particle. If σ is larger, some clutter might be still regarded as target-generated measurements, and the overestimate of the target number cannot be solved. By contrast, if σ is smaller, the likelihood values of some clutter which contribute a small proportion to the updated weight will be modified. However, this modification will have no obvious effect on the updated weight of the particle except for unnecessary complexity. As we all know, $N_{particle}$ is the number of particles for each expected target in the SMC-PHD filter and it is used to extract the weighted particles to approximate the posterior intensity. To some extent, $1/N_{particle}$ can approximately reflect the average contribution of each particle for an underlying target. Therefore, we choose $\sigma = 1/N_{particle}$ as the threshold.

4.2. Linear Scenario

The target state of each target is $\boldsymbol{x}_k = [x_k, \dot{x}_k, y_k, \dot{y}_k]^T$, consisting of the position $[x_k, y_k]^T$ and velocity $[\dot{x}_k, \dot{y}_k]^T$. The target motion model is given by

$$\boldsymbol{x}_k = \boldsymbol{F}\boldsymbol{x}_{k-1} + \boldsymbol{G}\boldsymbol{w}_k \tag{64}$$

where

$$\boldsymbol{F} = \begin{bmatrix} 1 & \Delta & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \Delta \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{65}$$

$$\boldsymbol{G} = \begin{bmatrix} \Delta & 1 & 0 & 0 \\ 0 & 0 & \Delta & 1 \end{bmatrix}^T \tag{66}$$

$\boldsymbol{w}_k \sim \mathcal{N}(\cdot; 0_2, \boldsymbol{Q}_k)$, $\boldsymbol{Q}_k = \mathrm{diag}([\sigma_w^2, \sigma_w^2]^T)$, $\sigma_w = 5$ m/s^2 is the standard deviation of the process noise, $\Delta = 1$ s is the sampling period.

The observation model is formulated by

$$\boldsymbol{z}_k = \boldsymbol{H}\boldsymbol{x}_k + \boldsymbol{\varepsilon}_k \tag{67}$$

where

$$\boldsymbol{H} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \tag{68}$$

The heavy-tailed measurement noise corrupted by outliers is represented as [21,23]

$$\boldsymbol{\varepsilon}_k \sim \begin{cases} \mathcal{N}(\boldsymbol{\varepsilon}_k; 0_2, \boldsymbol{R}_0) & 1-\eta \\ \mathcal{N}(\boldsymbol{\varepsilon}_k; 0_2, \boldsymbol{R}_n) & \eta \end{cases} \tag{69}$$

where η denotes the probability of the measurement noise outlier. $\boldsymbol{R}_0 = \mathrm{diag}([10^2, 10^2]^T)$ is the nominal noise variance, and $\boldsymbol{R}_n = 20\boldsymbol{R}_0$.

The model of the birth target is a Poisson RFS with intensity

$$\gamma_k(\boldsymbol{x}) = \sum_{j=1}^{3} 0.1\mathcal{N}(\boldsymbol{x}; \boldsymbol{m}_\gamma^{(j)}, \boldsymbol{P}_\gamma) \tag{70}$$

where $\boldsymbol{m}_\gamma^{(1)} = [1200, 0, -400, 0]^T$, $\boldsymbol{m}_\gamma^{(2)} = [600, 0, 700, 0]^T$, $\boldsymbol{m}_\gamma^{(3)} = [600, 0, -300, 0]^T$,$\boldsymbol{P}_\gamma = \mathrm{diag}([100, 100, 100, 100]^T)$. The survival probability is $p_{S,k} = 0.99$ and the detection probability is $p_{D,k} = 0.98$. The surveillance region is $[0, 2000]$ m $\times$ $[-1000, 1000]$ m. The clutter is uniformly distributed over the surveillance region, and the average clutter number is $\lambda = 10$. For the particles of birth targets, $\boldsymbol{V}_{\gamma,k}^{(i)} = \mathrm{diag}([10^2, 10^2]^T)$, $v_{\gamma,k}^{(i)} = 4$, $\alpha_{\gamma,k}^{(i)} = 3$, $\beta_{\gamma,k}^{(i)} = 1$. The number of iteration is $\tau = 5$, the decreasing factor is $\rho = 0.98$, and the particles number is $N_{particle} = 800$.

The trajectories of the targets are shown in Figure 2, where circles are the beginning positions and triangles are the ending positions.

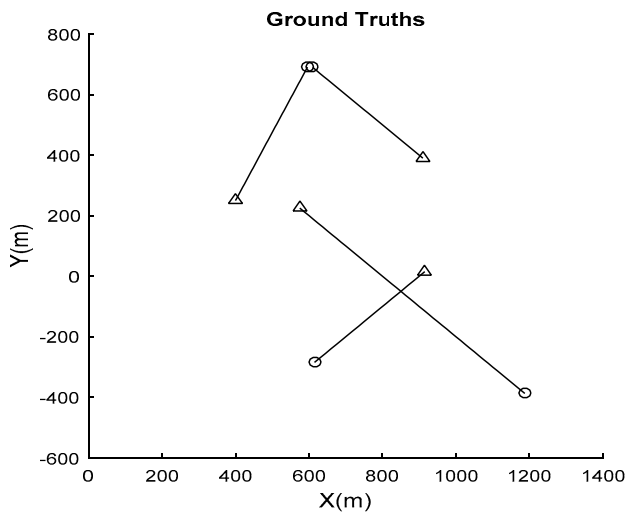

Figure 2. Trajectories of the targets.

To illustrate the performance of the proposed filter under the Gaussian measurement noise, η is set to 0. Figures 3 and 4 show the estimate of target number and OSPA distance of

the three filters, respectively. From Figure 3, we can see that all the filters achieve unbiased estimate of the target number and have a comparable estimate accuracy under Gaussian measurement noise. This is because the SMC-PHD filter utilizes the true measurement noise covariance. For the NGIWG-PHD filter and the RSMC-PHD filter, the DOF and scale matrix of the Student-t distribution can be adjusted automatically and the target tracking under Gaussian measurement noise realized. Figure 4 shows the OSPA distance of the three filters. Obviously, all the filters have an approximate OSPA distance.

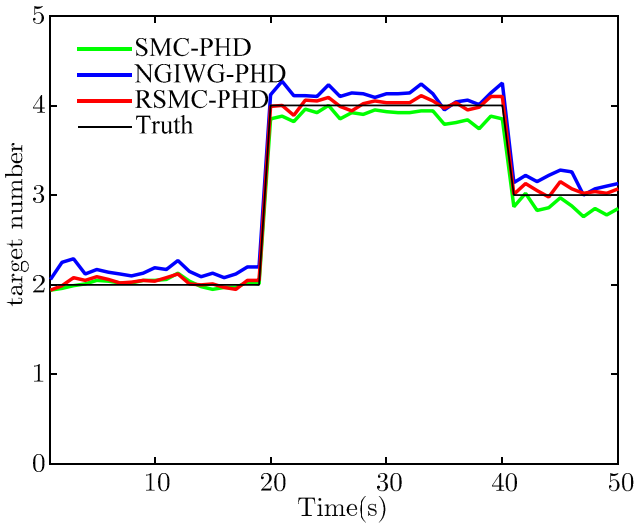

Figure 3. The estimate of target number.

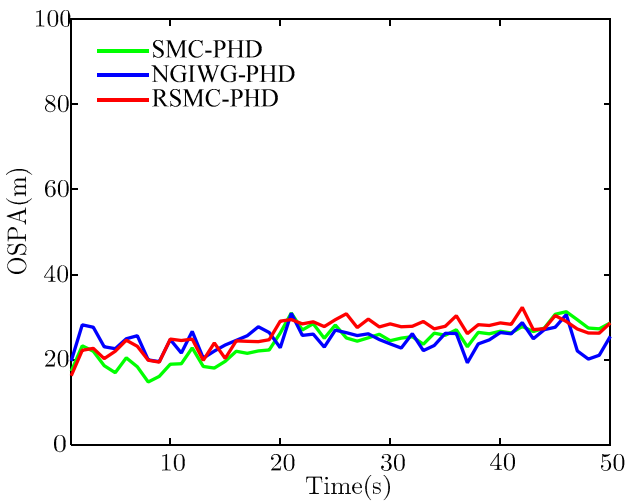

Figure 4. The estimate of OSPA.

To verify the multi-tracking performance under heavy-tailed measurement noise, η is set to 0.2. The remaining parameters are the same as above.

Figures 5 and 6 show the estimate of target number and OSPA distance of the three filters, respectively. Under heavy-tailed measurement noise, the SMC-PHD filter will diverge because the assumed measurement noise is not matched with the true measurement noise. For the NGIWG-PHD filter and the RSMC-PHD filter, the two filters will produce a weight which is non-negligible for filter update when there are outliers in measurements. Thus, the estimate accuracy of the target number of the NGIWG-PHD filter and the RSMC-PHD filter is higher than that of the SMC-PHD filter. From Figure 6, it can be observed that the RSMC-PHD filter has the best performance on OSPA distance.

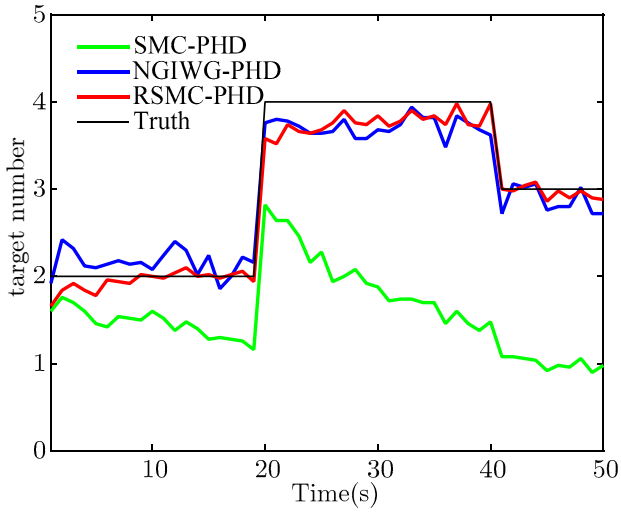

Figure 5. The estimate of target number.

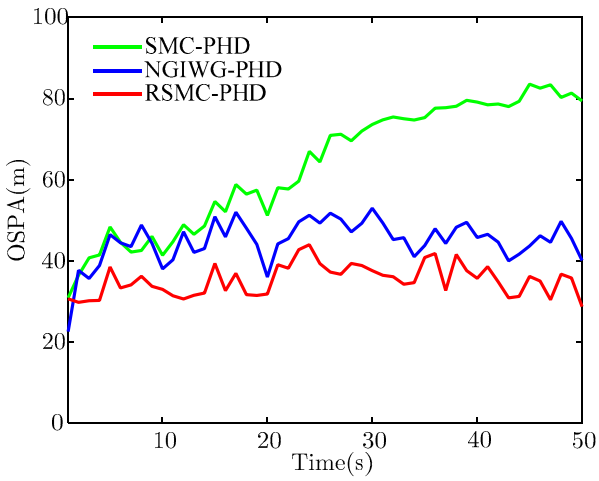

Figure 6. The estimate of OSPA.

To verify the validity of the weight modification method, we compare the proposed RSMC-PHD filter with the proposed robust SMC-PHD filter without weight modification (RSMC-PHD(M1)). Figures 7 and 8 show the mean absolute error (MAE) of the target number and the average OSPA (AOSPA) distance for the RSMC-PHD(M1) filter and the RSMC-PHD filter under different clutter number. The MAE of the target number and AOSPA are given by

$$\mathrm{MAE}_{\mathrm{number}} = \frac{1}{M_C T} \sum_{j=1}^{M_C} \sum_{k=1}^{T} \left| N_k^j - \hat{N}_k^j \right| \tag{71}$$

$$\mathrm{AOSPA} = \frac{1}{M_C T} \sum_{j=1}^{M_C} \sum_{k=1}^{T} \mathrm{OSPA}_k^j \tag{72}$$

where M_C is the number of Monte Carle trials, T is the simulation tracking time, N_k^j and $\hat{N}_k^j$ are the true and the estimated target numbers at time k in the jth Monte Carle trial, OSPA_k^j is the corresponding OSPA distance.

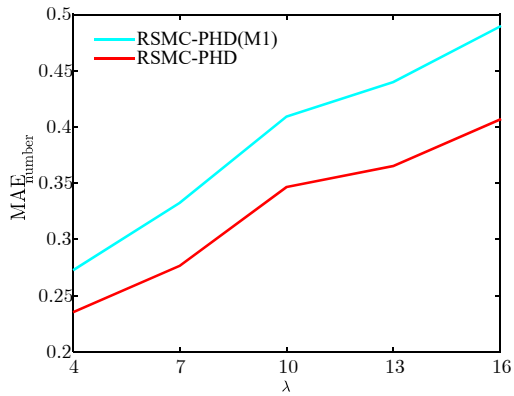

Figure 7. The MAE estimate of target number.

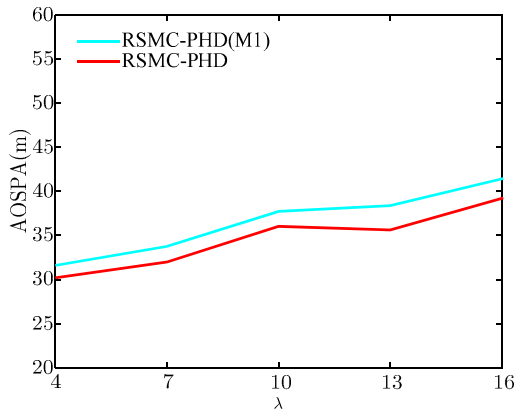

Figure 8. The estimate of AOSPA.

Figure 7 illustrates that the MAE estimate of target number for the RSMC-PHD filter is lower than that of the RSMC-PHD(M1) filter. The more the clutter number, the higher is the probability that the clutter is regarded as the target-generated measurement, resulting in an overestimate of the target number at the current time. From Figure 8, it can be seen that the RSMC-PHD filter also outperforms the RSMC-PHD(M1) filter on AOSPA distance.

4.3. Nonlinear Scenario

To further evaluate the performance of the proposed filter, we consider a nonlinear scenario. For the NGIWG-PHD filter, EKF is used to handle nonlinearity.

The target state transition function is given by

$$\boldsymbol{x}_k = \boldsymbol{F}_{ct}\boldsymbol{x}_{k-1} + \boldsymbol{G}_{ct}\boldsymbol{w}_k \tag{73}$$

where

$$\boldsymbol{F}_{ct} = \begin{bmatrix} 1 & \sin\omega\Delta/\omega & 0 & -(1-\cos\omega\Delta)/\omega & 0 \\ 0 & \cos\omega\Delta & 0 & -(1-\cos\omega\Delta)/\omega & 0 \\ 0 & (1-\cos\omega\Delta)/\omega & 1 & \sin\omega\Delta/\omega & 0 \\ 0 & \sin\omega\Delta/\omega & 0 & \cos\omega\Delta & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{74}$$

$$\boldsymbol{G}_{ct} = \begin{bmatrix} \Delta^2/2 & \Delta & 0 & 0 & 0 \\ 0 & 0 & \Delta^2/2 & \Delta & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}^T \tag{75}$$

The target state is $\boldsymbol{x}_k = [x_k, \dot{x}_k, y_k, \dot{y}_k, \omega]^T$, ω is the turn rate, $\boldsymbol{w}_k \sim \mathcal{N}(\cdot; 0_2, \boldsymbol{Q}_k)$, $\boldsymbol{Q}_k = \mathrm{diag}([\sigma_x^2, \sigma_y^2, \sigma_\omega^2]^T)$, $\sigma_x = \sigma_y = 5\ \mathrm{m/s^2}$, $\sigma_\omega = 5(\pi/180)\ \mathrm{rad/s}$, $\Delta = 1\ \mathrm{s}$.

The observation model is formulated by

$$z_k = \begin{bmatrix} \sqrt{x_k^2 + y_k^2} \\ \arctan(y_k / x_k) \end{bmatrix} + \varepsilon_k \tag{76}$$

Assume the nominal measurement noise variance is $\boldsymbol{R}_0 = \mathrm{diag}([10^2, (\pi/180)^2]^T)$, and the true measurement noise is represented as (69) with $\boldsymbol{R}_n = 20\boldsymbol{R}_0$ and $\eta = 0.2$.

The model of birth target is a Poisson RFS with intensity

$$\gamma_k(\boldsymbol{x}) = \sum_{j=1}^{3} 0.02\mathcal{N}(\boldsymbol{x}; \boldsymbol{m}_\gamma^{(j)}, \boldsymbol{P}_\gamma) \tag{77}$$

where $\boldsymbol{P}_\gamma = \mathrm{diag}([900, 900, 900, 900, (6 \times \pi/180)^2]^T)$, $\boldsymbol{m}_\gamma^{(1)} = [-400, 0, 600, 0, 0]^T$, $\boldsymbol{m}_\gamma^{(2)} = [-250, 0, 250, 0, 0]^T$, and $\boldsymbol{m}_\gamma^{(3)} = [250, 0, 450, 0, 0]^T$. The surveillance region is $[0, \pi]$rad $\times$ $[0, 1500]$m. The clutter is uniformly distributed over the surveillance region and the average clutter number is $\lambda = 10$. For the particles of birth targets, $\boldsymbol{V}_{\gamma,k}^{(i)} = \mathrm{diag}([10^2, (\pi/180)^2]^T)$, $v_{\gamma,k}^{(i)} = 4$, $\alpha_{\gamma,k}^{(i)} = 3$, $\beta_{\gamma,k}^{(i)} = 1$. The remaining parameters are the same as that in Section 4.2.

The trajectories of the targets are shown in Figure 9, where circles are the beginning positions and triangles are the ending positions.

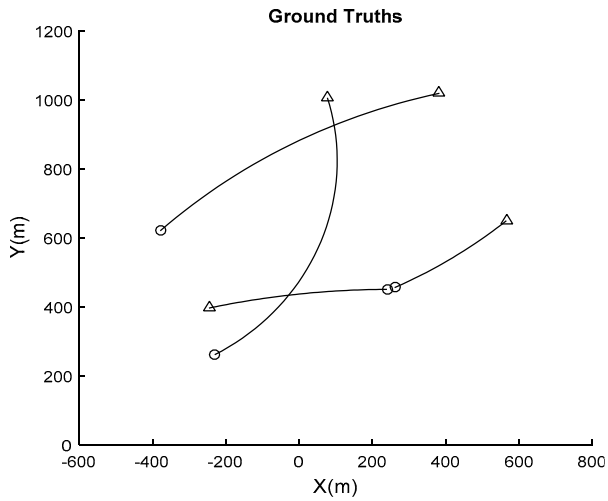

Figure 9. Trajectories of the targets.

Figures 10 and 11 show the estimate of target number and OSPA distance for the two filters, respectively. It indicates that the NGIWG-PHD filter performs worse in the nonlinear system. This is because the NGIWG-PHD filter uses EKF to linearize the nonlinear model, and the estimate error of the target state will become large. Under the influence of heavy-tailed measurement noise, the target will miss and the target number will be underestimated. On the contrary, the RSMC-PHD filter is not subject to nonlinear constraint, and has a better performance in the nonlinear system.

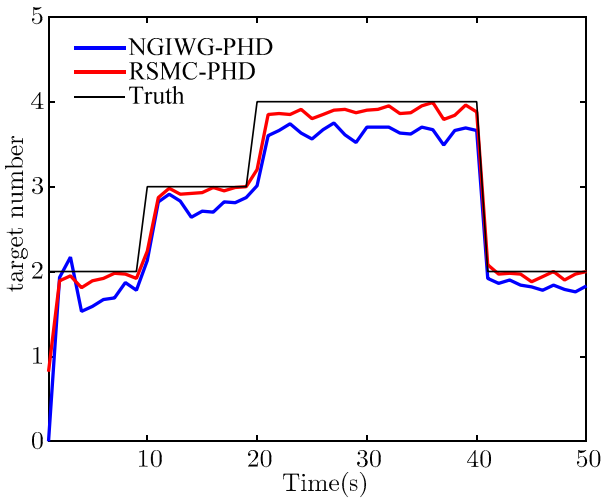

Figure 10. The estimate of target number.

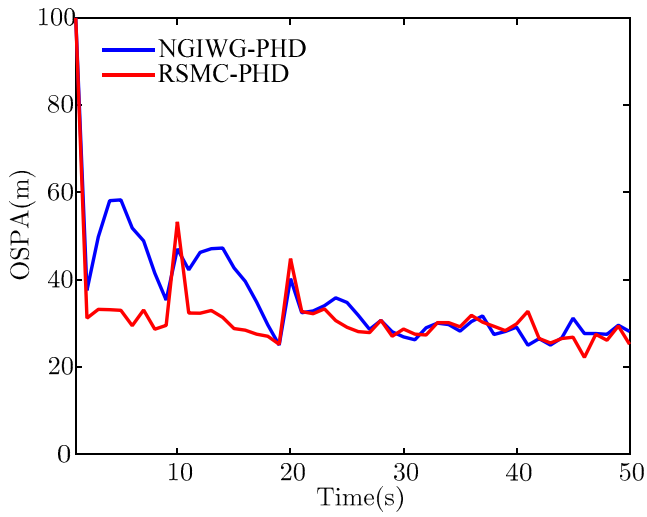

Figure 11. The estimate of OSPA.

5. Conclusions

In this paper, a robust SMC-PHD filter with unknown heavy-tailed measurement noise is proposed. The proposed filter uses the Student-t distribution to model the unknown measurement noise. The IW distribution is used to model the unknown scale matrix parameter. Meanwhile, the Gamma distribution is used to model the unknown DOF parameter. Then, the unknown noise parameters are updated by the VB technique within the SMC-PHD filter framework. In addition, a strategy is applied to solve the overestimate of the target number induced by the Student-t distribution. Simulation results under the linear system and nonlinear system show that the proposed filter has a better performance than the compared filters with the unknown heavy-tailed measurement noise, especially in a nonlinear system.

Author Contributions: Conceptualization, Y.G. and C.C.; methodology, Y.G.; software, Y.G.; validation, Y.G., C.C.; formal analysis, C.C.; investigation, Y.G.; resources, Y.G.; data curation, Y.G.; writing—original draft preparation, Y.G.; writing—review and editing, Y.G.; visualization, Y.G.; supervision, C.C.; project administration, Y.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

1. Mahler, R.P.S. *Statistical Multisource-Multitarget Information Fusion*; Artech House: Norwood, MA, USA, 2007.
2. Blom, H.; Bloem, E. Interacting multiple model joint probabilistic data association avoiding track coalescence. In Proceedings of the 41st IEEE Conference on Decision and Control, Las Vegas, NV, USA, 10–13 December 2002; pp. 3408–3415.
3. Blackman, S. Multiple hypothesis tracking for multiple target tracking. *IEEE Aerosp. Electron. Syst. Mag.* **2004**, *19*, 5–18. [CrossRef]
4. Mahler, R.P.S. Multitarget bayes filtering via first-order multitarget moments. *IEEE Trans. Aerosp. Electron. Syst.* **2003**, *39*, 1152–1178. [CrossRef]
5. Vo, B.-N.; Ma, W.-K. The Gaussian Mixture Probability Hypothesis Density Filter. *IEEE Trans. Signal Process.* **2006**, *54*, 4091–4104. [CrossRef]
6. Vo, B.-N.; Singh, S.S.; Boucet, A. Sequential monte carlo methods for multi-target filtering with random finite sets. *IEEE Trans. Aerosp. Electron. Syst.* **2005**, *41*, 1224–1245. [CrossRef]
7. Zhou, X.; Li, Y.; He, B. Entropy distribution and coverage rate-based birth intensity estimation in GM-PHD filter for multi-target visual tracking. *Signal Process.* **2014**, *94*, 650–660. [CrossRef]
8. Yan, B.; Xu, N.; Xu, L.; Li, M.Q.; Cheng, P. An improved partitioning algorithm based on FCM algorithm for extended target tracking in PHD filter. *Digit. Signal Process.* **2019**, *90*, 54–70. [CrossRef]
9. Sarkka, S.; Nummenmaa, A. Recursive Noise Adaptive Kalman Filter ing by Variational Bayesian Approximations. *IEEE Trans. Autom. Control* **2009**, *54*, 596–600. [CrossRef]
10. Sarkka, S.; Hartikainen, J. Non-linear noise adaptive Kalman filtering via variational Bayes. In Proceedings of the 2013 IEEE International Workshop on Machine Learning for Signal Processing (MLSP), Southampton, UK, 22–25 September 2013; pp. 1–6.
11. Huang, Y.; Zhang, Y.; Wu, Z.; Li, N.; Chambers, J.A. A Novel Adaptive Kalman Filter With Inaccurate Process and Measurement Noise Covariance Matrices. *IEEE Trans. Autom. Control* **2018**, *63*, 594–601. [CrossRef]
12. Wu, X.; Huang, G.M.; Gao, J. Adaptive noise variance identification for probability hypothesis density-based multi-target filter by variational Bayesian approximations. *IET Radar Sonar Navig.* **2013**, *7*, 895–903. [CrossRef]
13. Yang, J.; Ge, H. Adaptive probability hypothesis density filter based on variational Bayesian approximation for multi-target tracking. *IET Radar Sonar Navig.* **2013**, *7*, 959–967. [CrossRef]
14. Wu, X.; Huang, G.; Gao, J. Particle filters for probability hypothesis density filter with the presence of unknown measurement noise covariance. *Chin. J. Aeronaut.* **2013**, *26*, 1517–1523. [CrossRef]
15. Bilik, I.; Tabrikian, J. Maneuvering Target Tracking in the Presence of Glint using the Nonlinear Gaussian Mixture Kalman Filter. *IEEE Trans. Aerosp. Electron. Syst.* **2010**, *46*, 246–262. [CrossRef]
16. Kim, J.; Tandale, M.; Menon, P.; Ohlmeyer, E. Particle Filter for Ballistic Target Tracking with Glint Noise. *J. Guidan. Control Dynam.* **2010**, *33*, 1–4. [CrossRef]
17. Li, K.; Chang, L.; Hu, B. A Variational Bayesian-Based Unscented Kalman Filter with Both Adaptivity and Robustness. *IEEE Sensors J.* **2016**, *16*, 6966–6976. [CrossRef]
18. Piche, R.; Sarkka, S.; Hartikainen, J. Recursive outlier-robust filtering and smoothing for nonlinear systems using the multivariate student-t distribution. In Proceedings of the 2012 IEEE International Workshop on Machine Learning for Signal Processing, Santander, Spain, 23–26 September 2012; pp. 1–6.
19. Huang, Y.; Zhang, Y.; Li, N.; Chambers, J. Robust student's t based nonlinear filter and smoother. *IEEE Trans. Aerosp. Electron. Syst.* **2016**, *52*, 2586–2596. [CrossRef]
20. Huang, Y.; Zhang, Y. Robust Student's t-Based Stochastic Cubature Filter for Nonlinear Systems with Heavy-Tailed Process and Measurement Noises. *IEEE Access* **2017**, *5*, 7964–7974. [CrossRef]
21. Huang, Y.; Zhang, Y.; Wu, Z.; Li, N.; Chambers, J. A Novel Robust Student's t-Based Kalman Filter. *IEEE Trans. Aerosp. Electron. Syst.* **2017**, *53*, 1545–1554. [CrossRef]
22. Li, W.; Jia, Y.; Du, J.; Zhang, J. PHD filter for multi-target tracking with glint noise. *Signal Process.* **2014**, *94*, 48–56. [CrossRef]
23. Dong, P.; Jing, Z.; Leung, H.; Shen, K.; Li, M. The Labeled Multi-Bernoulli Filter for Multitarget Tracking with Glint Noise. *IEEE Trans. Aerosp. Electron. Syst.* **2018**, *55*, 2253–2268. [CrossRef]
24. Li, P.; Xu, C.; Wang, W.; Su, S. Robust Student's T Distribution Based PHD/CPHD Filter for Multiple Targets Tracking Using Variational Bayesian Approach. *Radioengineering* **2020**, *29*, 529–539. [CrossRef]
25. Xu, D.; Shen, C.; Shen, F. A Robust Particle Filtering Algorithm with Non-Gaussian Measurement Noise Using Student-t Dis-tribution. *IEEE Signal Process. Letters* **2013**, *21*, 30–34. [CrossRef]
26. Zhu, H.; Leung, H.; He, Z. A variational Bayesian approach to robust sensor fusion based on Student-t distribution. *Inf. Sci.* **2013**, *221*, 201–214. [CrossRef]

27. Christmas, J.; Everson, R. Robust Autoregression: Student-t Innovations Using Variational Bayes. *IEEE Trans. Signal Process.* **2010**, *59*, 48–57. [CrossRef]
28. Schuhmacher, D.; Vo, B.-T. A Consistent Metric for Performance Evaluation of Multi-Object Filters. *IEEE Trans. Signal Process.* **2008**, *56*, 3447–3457. [CrossRef]