RESEARCH ARTICLE

An extended target tracking algorithm based on trajectory feedback constrained spatial information

Yuting Zhang[1], **Hantao Li**[2]*, **Jiye Li**[1]

1 The School of Computer Engineering, Chengdu Technological University, Chengdu, SiChuan, P.R. China, **2** The School of Artificial Intelligence, Chongqing University of Education, Chongqing, P.R. China

* liht@cque.edu.cn

Check for updates

OPEN ACCESS

Citation: Zhang Y, Li H, Li J (2026) An extended target tracking algorithm based on trajectory feedback constrained spatial information. PLoS One 21(9): e0357813. https://doi.org/10.1371/journal.pone.0357813

Editor: Xuebo Zhang, Northwest Normal University, CHINA

Received: April 13, 2026

Accepted: August 21, 2026

Published: September 18, 2026

Copyright: © 2026 Zhang et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data availability statement: The data underlying the results of this study are available from Figshare (https://doi.org/10.6084/m9.figshare.32930585).

Funding: This study was financially supported by the Central Government Guided Local S&T Development Fund for Sichuan Province in the form of a grant awarded to YZ (2025ZYDF108).

Abstract

Accurately estimating the state of extended targets is a major challenge because measurement numbers and distributions change significantly, especially when tracking targets are close or overlapping. To solve the track fragmentation problem caused by spatial measurement ambiguity, a novel Trajectory Situation Feedback-based Gaussian Mixture Model Expectation Maximization (TSF-GMM-EM) method is proposed within the framework of the Gaussian Mixture Probability Hypothesis Density (GM-PHD) filter. The proposed method feeds historical trajectory information back to guide current measurement partitioning and motion estimation while providing prior-guided initialization for the EM process, thereby improving partitioning reliability during target intersections and accelerating clustering convergence. Experimental results show that the TSF-GMM-EM method limits the peak OSPA error to approximately 12 m during target intersections. It also maintains a measurement partitioning accuracy above 0.946 in overlapping target scenarios.

1. Introduction

Multiple target tracking (MTT) has been used to track multiple targets using various sensors [1]. Point target tracking (PTT) and extended target tracking (ETT) are fundamental components in numerous real-world applications [2–4]. With the advancement of high-resolution sensors [5–9], an extended target typically generates multiple measurements per scan. These multiple measurements allow for the estimation of additional target states, such as shape and orientation [10,11]. However, the presence of false alarms [12] and the increasing number of measurements introduce significant challenges to multi-extended target tracking (MTT). Specifically, the complexity of data association is exacerbated when measurement numbers and locations are time-varying, or when targets are in close proximity or overlapping. In such scenarios, incorrect measurement partitioning leads to degraded tracking accuracy or even track loss. Moreover, estimating target motion states from noisy measurements remains

This study was also financially supported by the Meteorological Information and Signal Processing Key Laboratory of Sichuan Higher Education Institutes of Chengdu University of Information Technology in the form of a grant awarded to YZ (CXHCL202507). Additional funding was provided by the Chengdu Technological University Teaching Reform and Quality Improvement Project in the form of a grant awarded to YZ (2025QNJG015). The Science and Technology Research Program of Chongqing Municipal Education Commission also provided financial support in the form of a grant awarded to HL (KJQN202601606). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

a fundamental challenge, as clutter interference inherently degrades the precision of state estimation [13,14]. Therefore, effective measurement partitioning remains a critical issue for enhancing MTT performance.

To address the inherent complexities of multi-target tracking problems, the Random Finite Set (RFS) framework has been extensively adopted, different from traditional data association-based methods, such as Joint Probabilistic Data Association (JPDA) filter and multiple hypotheses tracking (MHT). However, these methods incur a significant computational overhead as the number of targets and measurements increases. By modeling both target states and measurements as random sets, the RFS-based method effectively circumvents the combinatorial explosion associated with explicit data association [15–17]. A cornerstone of this framework is the Probability Hypothesis Density (PHD) filter proposed by Mahler [18,19], which propagates the first-order moment of the multi-target posterior density. Although the original Poisson-based recursions provided a solid theoretical foundation, their implementation remained computationally prohibitive for real-time applications until the development of closed-form solutions like the Gaussian Mixture (GM-PHD) filter. To maintain computational efficiency, the GM-PHD filter incorporates pruning and merging operations that effectively constrain the recursive expansion of Gaussian components. By eliminating explicit data association, the GM-PHD framework delivers reliable tracking performance with minimal computational cost.

Over the past decades, the advancement of Extended Target Tracking (ETT) has evolved through several distinct methodology paradigms, primarily focusing on how to jointly estimate target kinematics and spatial extensions [20,21]. Early pioneered works heavily relied on spatial shape parameterization, most notably the Random Matrix (RM) approach introduced by Koch [11], which models the target extension as a positive-definite matrix following an inverted Wishart distribution. This framework was later extended to multi-target scenarios by Granström et al. [22] and adapted for maneuvering non-ellipsoidal shapes by Liang et al. [3] using sub-random matrices. Concurrently, non-parametric approaches such as Monte Carlo methods [10], alongside spatiotemporal Gaussian Processes (GP) [23,24], were introduced to approximate highly irregular geometric contours.

To bridge the gap between contour tracking and intent inference, recent many researchers have shifted toward Joint Tracking and Classification (JTC). Notably, Cheng et al. [25] proposed an extended target JTC filter based on GPs and the Labeled Multi-Bernoulli (LMB) framework, while class-enhanced GP-PHD variants [26] optimized rectangular constraints for complex vehicle profiles. More recently, to explicitly capture multi-frame temporal tracks, the Trajectory Probability Hypothesis Density (TPHD) filter was integrated with Random Matrix Models [27], enabling the simultaneous recursive estimation of target states, classes, and explicit trajectory sets. While these filtering frameworks offer solid mathematical solutions for shape tracking, their state recursion accuracy remains fundamentally bounded by the data association layer, specifically the quality of measurement partitioning in dense clutter environments. Moreover, Al-Haddad et al. [28] have significantly elevated the

robustness of trajectory control against non-linear disturbances, and it is crucial to connect these filtering breakthroughs with state-of-the-art autonomous applications.

To feed the multi-target RFS update recursions—such as the Extended Target GM-PHD filter [22] and the interactive multiple-model (IMM-PHD) filter for coexisting point and extended targets proposed by Li et al. [29]—measurement partitioning acts as a critical front-end clustering mechanism that groups unlabelled point clouds into individual target-originated cells. Existing partitioning strategies can be broadly classified into two categories: spatially-driven heuristics and statistical mixture optimization models.

Within the spatially-driven category, Distance Partitioning (DP) [22] and Density-Based Spatial Clustering (DBSCAN) [30] are widely utilized due to their intuitive implementation based on spatial proximity. However, DP suffers from severe computational overhead because it necessitates iterating through an extensive range of predefined distance thresholds to handle varying target ranges [31], frequently generating redundant partitions. Similarly, DBSCAN is notoriously sensitive to its hyper-parameters, exhibiting poor robustness when dealing with heterogeneous target sizes. To alleviate threshold sensitivity, scalable association-based non-homogeneous Poisson process (AbNHPP) frameworks [32] have emerged in 2026 to dynamically model time-varying measurement rates using Generalized Inverse Gaussian distributions.

To circumvent the rigid thresholding of heuristic partitioning, optimization-based soft clustering approaches utilizing Gaussian Mixture Models and the Expectation-Maximization (GMM-EM) algorithm have been introduced [33,34]. These methods partition measurements based on continuous posterior association probabilities rather than hard spatial boundaries. Nevertheless, standard GMM-EM formulations operate purely in the spatial domain of a single time-frame, leaving them highly vulnerable to the structural initializations of cluster centers. Consequently, in complex scenarios where multiple extended targets closely approach, merge, or overlap, the standard EM iteration inherently traps into uninformative local maxima [20]. This directly manifests as severe cardinality estimation errors, track fragmentation, or catastrophic track loss.

Existing partitioning strategies, such as distance partitioning (DP) [22] and DBSCAN [30], rely on the spatial closeness of measurements. However, DP often requires predefined distance thresholds, which can lead to redundant partitions and high computational costs [31]. Similarly, DBSCAN's performance is sensitive to its radius and minimum-point parameters, showing poor robustness in environments with diverse target types. Alternatively, the Expectation-Maximization (EM) algorithm and Gaussian Mixture Models (GMM) [33,34] partition measurements based on posterior probabilities. Nevertheless, standard GMM-EM often converges to local maxima, resulting in uninformative partitions—especially when targets are close or converging.

A key limitation of the aforementioned methods is their reliance on spatial information from a single time step. In reality, target features exhibit strong spatio-temporal correlations across frames [23,24,35]. Exploiting the target's trajectory—which inherently contains prior information about target numbers and measurement structures—could break the current bottleneck in partitioning and state estimation.

1.1 Contributions

Based on the GM-PHD filter, this paper proposes a Trajectory Situation Feedback-based Gaussian Mixture Model Expectation Maximization (TSF-GMM-EM) method. The main contributions and advantages are summarized as follows:

- **A closed-loop trajectory feedback framework:** We establish a bidirectional interaction between trajectory evolution and measurement partitioning. Different from the open-loop methods, the proposed TSF-GMM-EM leverages temporal trajectory priors to guide spatial partitioning, effectively resolving data association ambiguity in complex scenarios.
- **A prior-guided EM initialization strategy:** By leveraging predicted target states to drive the measurement partitioning process, this strategy ensures spatio-temporal consistency between target trajectories and measurement clusters.

Moreover, this method effectively avoids local optima in the GMM-EM framework and prevents cardinality errors during target intersections, thereby enhancing both tracking precision.

1.2 Organization of the paper

The remainder of this paper is structured as follows. Section 2 establishes the extended target tracking mathematical model within the Random Finite Set framework and details the proposed TSF-GMM-EM method, including trajectory-guided measurement partitioning and intersection situation judgment. Section 3 presents the comprehensive experimental framework. Finally, Section 4 concludes the paper and outlines potential directions for future research.

2. Method

In this section, the extended target tracking problem is formulated within the Random Finite Set model. The Gaussian Mixture Probability Hypothesis Density filter is adopted as the baseline, where the target states and measurements are modeled as RFS to bypass explicit data association. For extended targets, the update step of the GM-PHD filter relies heavily on the quality of measurement partitioning. However, traditional partitioning methods often fail to provide accurate subsets when targets are in close scene. To address this, a feedback mechanism is introduced to bridge the gap between trajectory estimation and measurement clustering.

2.1 Model

A number of measurements from an extended target spread over a region at every instant, and their distribution changes over time [36]. In essence, multi-extended targets tracking is to estimate the number of targets and every target's state $\mathbf{X}_k = \left\{\mathbf{x}_k^{(i)}\right\}_{i=1}^{N_{x,k}}$ from the measurements $\mathbf{Z}_k = \left\{\mathbf{z}_k^{(i)}\right\}_{i=1}^{N_{z,k}}$ up to instant k, where $N_{x,k}$ denotes the unknown number produced by targets and $N_{z,k}$ denotes the available number of measurements. The dynamic evolution of each target state $\mathbf{x}_k^{(i)}$ in the RFS $\mathbf{X}_k$ can be modeled by a linear Gaussian dynamic model, and its motion state equation could be denoted as

$$\mathbf{x}_{k+1}^{(i)} = \mathbf{T_k}\mathbf{x}_k^{(i)} + \mathbf{G_k}\mathbf{w}_k^{(i)} \tag{1}$$

where $\mathbf{T_k}$ and $\mathbf{G_k}$ are the known state transition matrix, $\mathbf{w}_k^{(i)}$ is Gaussian white noise with covariance matrix $\mathbf{Q}_k^{(i)}$. It is worth noting that each target state is independent between others.

The number of measurements generated by an extended target is assumed by a Poisson-distributed random variable with rate $\gamma\left(\mathbf{x}_k^{(i)}\right)$ at every instant, where $\gamma(\cdot)$ is a known non-negative function. For an extended target, the detection probability could be defined as

$$P_{D,\text{eff}} = \left(1 - e^{-\gamma\left(\mathbf{x}_k^{(i)}\right)}\right) P_D\left(\mathbf{x}_k^{(i)}\right) \tag{2}$$

where $P_D\left(\mathbf{x}_k^{(i)}\right)$ denotes the probability of $i-th$ target, where $P_D(\cdot)$ is a known non-negative function.

The measurements originating from the $i-th$ target are usually modeled as

$$\mathbf{z}_k^{(i)} = \mathbf{H_k}\mathbf{x}_k^{(i)} + \mathbf{e}_k^{(i)} \tag{3}$$

where $\mathbf{H_k}$ is a known measurement matrix, and $\mathbf{e}_k^{(i)}$ denotes white Gaussian noise with covariance matrix $\mathbf{R}_k^{(i)}$. The noise vectors $\mathbf{w}_k^{(i)}$ and $\mathbf{e}_k^{(i)}$ are statistically independent.

In addition to the measurements generated by the extended targets, there are also false alarm in the measurement space. The number of false alarm measurements at instant k is also a Poisson distributed random variable with rate $\beta_{\mathrm{FA},k}$, and the mean number of false alarm measurements is $\beta_{\mathrm{FA},k}V_S$, where V_S is the surveillance volume. Moreover, the false alarm measurements is assumed as uniform distribution in the spatial domain.

In the ET-GM-PHD filter, the measurement update formula can be represented as

$$D_{k|k}\left(\mathbf{x}|\mathbf{Z}\right) = L_{\mathbf{z}_k}\left(\mathbf{x}\right) D_{k|k-1}\left(\mathbf{x}|\mathbf{Z}\right) \tag{4}$$

where the pseudo-likelihood function $L_{\mathbf{z}_k}$ is defined as

$$\begin{aligned} L_{\mathbf{z}_k}(\mathbf{x}) =& 1 - \left(1 - e^{-\gamma(\mathbf{x})}\right) P_D(\mathbf{x}) \\ &+ e^{-\gamma(\mathbf{x})} P_D(\mathbf{x}) \times \sum_{\mathrm{p}\angle\mathbf{z}_k} \omega_{\mathrm{p}} \sum_{W\in\mathrm{p}} \frac{\gamma(\mathbf{x})^{|W|}}{d_W} \\ &\cdot \prod_{z\in W} \frac{\phi_{\mathbf{z}_k}(\mathbf{x})}{\lambda_k c_k(\mathbf{z}_k)} \end{aligned} \tag{5}$$

where $\phi_{\mathbf{z}_k}\left(\mathbf{x}\right) = p\left(\mathbf{z_k}|\mathbf{x}\right)$ is the likelihood function of the measurement generated from a single target, λ_k denotes the mean number of false alarm measurements, $c_k\left(\mathbf{z}_k\right) = 1/V_S$ denotes the spatial distribution of the false alarm over the surveillance volume, $p\angle\mathbf{Z}_k$ denotes the sub-set p of measurement $\mathbf{Z}_k$ by means of measurement partition, and each nonempty cells W contains a potential measurement result of the extended target.

2.2 The proposed method

Inspired by the strong short-term spatio-temporal correlations exhibited by extended target measurements and motion trajectories, the proposed method leverages temporal information from previous frames to guide the current spatial partitioning. Specifically, it is structured into two primary modules: trajectory-guided measurement partitioning and intersection situation judgment.

2.2.1 Trajectory-guided measurement partitioning. The core idea of the expectation maximization is to find certain parameter for the complete data by maximizing the expectation likelihood [28]. The Gaussian mixture model EM (GMM-EM) algorithm estimates the probability that a measurement belongs to each cluster by the GMM parameters [37,38].

$$p\left(\mathbf{z}^{(i)}|\theta\right) = \sum_{c=1}^{C} \pi_c \mathcal{N}\left(\mathbf{z}^{(i)}|\boldsymbol{\mu}_c, \boldsymbol{P}_c\right) \tag{6}$$

where $\mathbf{z}^{(i)}$ is the measurements, $\boldsymbol{\theta} = \left\{\pi_c, \boldsymbol{\mu}_c, \boldsymbol{P}_c\right\}_{c=1}^{C}$ denotes the parameters of GMM, C is the number of Gaussian sub-models, $\boldsymbol{\mu}_c \in \mathbb{R}^{2\mathrm{x}1}$ and $\boldsymbol{P}_c \in \mathbb{R}^{2\mathrm{x}2}$ are the mean vector and the covariance matrix of the $c-th$ Gaussian component, respectively. Moreover, π_c is the component weight of the $c-th$ Gaussian component, and it abides by $0 < \pi_c \leq 1$ and $\sum_{c=1}^{C} \pi_c = 1$.

In order to simplify the calculation, the log calculation $L\left(\boldsymbol{\theta}\right)$ is performed on the maximum likelihood function to search the optimal parameters $\boldsymbol{\theta}^*$.

$$L\left(\boldsymbol{\theta}\right) = \sum_{i=1}^{N_z} \log\left(p\left(\mathbf{z}^{(i)}|\boldsymbol{\theta}\right)\right) \tag{7}$$

$$\theta^* = \arg\max_{\theta} L(\theta) \tag{8}$$

where N_z is the amount of all measurements, and the symbol argmax denotes maximizing the likelihood function to get the optimal parameters θ^*.

However, conventional GMM-EM typically operates in a purely spatial domain, often neglecting the temporal continuity of target motion. To bridge this gap, the proposed method extends the conventional likelihood maximization framework into the spatio-temporal domain to achieve a more robust partition. Specifically, let $\mathbf{F}_k$ represent the temporal series information at instant k, encompassing the target count, motion states, and measurement distributions. Assuming that the target's characteristic information remains quasi-stationary over short intervals, TSF-GMM-EM identifies the optimal parameter set $\theta^* = \{\pi_k^*, \mu_k^*, \mathbf{P}_k^*\}$ by maximizing the likelihood function conditioned on the temporal series $\mathbf{F}_k$. Here, the optimal parameters θ^* denote the estimated Gaussian components of multiple extended targets. Consequently, the historical information $\mathbf{F}_{k-1}$ accumulates the optimal estimation parameters up to instant $k-1$, which can be formally expressed as:

$$\mathbf{F}_{k-1} = \left\{\theta_1^*, \theta_2^*, ..., \theta_{k-1}^*\right\} \tag{9}$$

As the covariance of an extended target, $\boldsymbol{P} \in \mathbb{R}^{2x2}$ represent the shape information of Gaussian surface, and it is denoted as

$$\boldsymbol{P} = \begin{bmatrix} \sigma_{xx}^2 & \sigma_{xy}^2 \\ \sigma_{yx}^2 & \sigma_{yy}^2 \end{bmatrix} \tag{10}$$

where σ_{xx}^2 and σ_{yy}^2 are the variances at horizontal direction and vertical direction variance, respectively, and $\sigma_{xy}^2 = \sigma_{yx}^2$ represents the variance in horizontal biased vertical direction.

Normally, the state of an extended target including the measurements and the motion is consistent and correlative until multiple targets approach and encounter during tracking process. Meanwhile, all measurements usually are disturbed by noise. Therefore, TSF-GMM-EM assumes the estimated parameters $\hat{\boldsymbol{P}}_c^*$ obeys a one-dimensional Gaussian distribution as

$$\hat{\boldsymbol{P}}_c^*(k) \sim \mathcal{N}\left(\mathcal{M}_c, \sigma_c^2\right), c = 1, 2, ..., C \tag{11}$$

where $\mathcal{M}_c$ and σ_c^2 denotes the mean and the variance of the $c-th$ target's true covariance, respectively; $\hat{\boldsymbol{P}}_c^*(k)$ denotes the estimated optimal feature of $c-th$ target at instant k.

It is widely known that the states including approaching and meet of multiple extended targets can be predicted in advance by means of their trajectory. Therefore, it is possible to accurately estimate the measurements partition for extended targets by means of predicting the state even at the situation that their measurements are mixed up when they approach or encounter. When the situation that multiple targets approach or encounter is predicted, TSF-GMM-EM would use predicted state to constrain the measurements partition as:

$$\boldsymbol{P}_c^* = \frac{1}{k-1}\sum_{i=1}^{k-1} \hat{\boldsymbol{P}}_c^*(i), \boldsymbol{S} \in \text{Encounter} \tag{12}$$

where $\boldsymbol{S} \in$ Encounter denotes that multiple targets encounter. Statistically, when spatial measurements collapse into extreme ambiguity during close encounters, the current-frame observational likelihood becomes highly non-informative.

Under the hypothesis of short-term temporal quasi-stationarity of target trajectories, the formulation in Eq. (12) serves as a localized empirical expectation estimator. By leveraging the arithmetic sample mean over a short historical window preceding the intersection, this mechanism functions as a Minimum Mean Square Error (MMSE) predictor that minimizes individual frame noise variance, thereby providing a stable, time-continuous prior constraint to regularize the soft-clustering optimization.

For the Expectation step (E-step), we have the posterior probability $\alpha_c^{\zeta+1}\left(\mathbf{z}^{(i)}\right)$ oriented from the parameters $\boldsymbol{\theta}^{\zeta} = \left\{\hat{\pi}_c^{\zeta}, \hat{\boldsymbol{\mu}}_c^{\zeta}, \hat{\boldsymbol{P}}_c^{\zeta}\right\}_{c=1}^{C}$ as follows:

$$\alpha_c^{\zeta+1}\left(\mathbf{z}^{(i)}\right) = \frac{\hat{\pi}_c^{\zeta} \cdot \mathcal{N}\left(\mathbf{z}^{(i)} | \hat{\boldsymbol{\mu}}_c^{\zeta}, \hat{\boldsymbol{P}}_c^{\zeta}\right)}{\sum_{c=1}^{C} \hat{\pi}_c^{\zeta} \mathcal{N}\left(\mathbf{z}^{(i)} | \hat{\boldsymbol{\mu}}_c^{\zeta}, \hat{\boldsymbol{P}}_c^{\zeta}\right)} \tag{13}$$

where $\hat{\pi}_c^{\zeta}$, $\hat{\boldsymbol{\mu}}_c^{\zeta}$ and $\hat{\boldsymbol{P}}_c^{\zeta}$ denotes the estimated weight, mean vector and covariance matrix of $c-th$ Gaussian sub-model at step ζ, respectively. The maximization step (M-step) of TSF-GMM-EM method can be denoted as

$$\hat{\pi}_c^{\zeta+1} = \frac{\sum_{i=1}^{N_{z,k}} \alpha_c^{\zeta+1}\left(\mathbf{z}^{(i)}\right)}{N_{z,k}}, c = 1, 2, ..., C \tag{14}$$

$$\hat{\boldsymbol{\mu}}_c^{\zeta+1} = \frac{\sum_{i=1}^{N_{z,k}} \left(\mathbf{z}^{(i)} \alpha_c^{\zeta+1}\left(\mathbf{z}^{(i)}\right)\right)}{\sum_{i=1}^{N_{z,k}} \alpha_c^{\zeta+1}\left(\mathbf{z}^{(i)}\right)}, c = 1, 2, ..., C \tag{15}$$

$$\hat{\boldsymbol{P}}_c^{\zeta+1} = \begin{cases} \frac{1}{k-1} \sum_{i=1}^{k-1} \left(\hat{\boldsymbol{P}}_c^{*}(i)\right), \mathbf{S} \in \text{Encounter} \\ \frac{\sum_{i=1}^{N_{z,k}} \left(\left(\mathbf{z}^{(i)} - \hat{\boldsymbol{\mu}}_c^{\zeta+1}\right)\left(\mathbf{z}^{(i)} - \hat{\boldsymbol{\mu}}_c^{\zeta+1}\right)^{\mathrm{T}} \alpha_c^{\zeta+1}\left(\mathbf{z}^{(i)}\right)\right)}{\sum_{i=1}^{N_{z,k}} \alpha_c^{\zeta+1}\left(\mathbf{z}^{(i)}\right)}, \text{otherwise} \end{cases} \tag{16}$$

where ζ is the iteration step, $\alpha^{\zeta+1}$ denotes the posterior probability at step $\zeta + 1$. And $\alpha^{\zeta+1}$ is estimated by the parameters $\boldsymbol{\theta}^{\zeta} = \left\{\hat{\pi}_c^{\zeta}, \hat{\boldsymbol{\mu}}_c^{\zeta}, \hat{\boldsymbol{P}}_c^{\zeta}\right\}_{c=1}^{C}$.

The EM algorithm easily converge to the local optimal solution for it is sensitive to the initial parameters [20,39]. To ensure an optimal result, TSF-GMM-EM uses the predicted location $\boldsymbol{\mu}_k^*$ as the initialization cluster center after the target trajectory is successfully established. At step $\zeta + 1$, the likelihood is updated as

$$L\left(\hat{\theta}^{\zeta+1}\right) = \sum_{i=1}^{N_{z,k}} L\left(\hat{\theta}^{\zeta}\right) \tag{17}$$

The iterative process terminates when the log-likelihood difference $|\mathcal{L}(\hat{\theta}^{\zeta+1}) - \mathcal{L}(\hat{\theta}^{\zeta})|$ falls below a predefined tolerance ε. Subsequently, let $\alpha_c^{\zeta+1}(\mathbf{z}^{(i)}|\boldsymbol{\theta}_{EM}^*)$ represent the posterior probability that measurement $\mathbf{z}^{(i)}$ belongs to cluster c. The optimal measurement partition is determined by maximizing this posterior probability as follows:

$$W = \arg\max_{c} \alpha_c^{\zeta+1}\left(\mathbf{z}^{(i)} \mid \theta_{EM}^{*}\right), \quad i = 1, 2, ..., N_{z,k}; \quad c = 1, 2, ..., C \tag{18}$$

where W denotes the measurements partition results of all measurements, c_i^* denotes the cluster index of the i-th measurement. Accordingly, the final measurement partition set W is constructed by grouping all measurements $\mathbf{z}^{(i)}$ associated with the same cluster index. These partitioned cells are then incorporated into the GM-PHD filter to update the states and weights of the multiple extended targets.

2.2.2 Intersection situation judgment. Target tracking algorithms estimate kinematic states and geometric attributes such as shape and orientation of multiple extended targets by establishing consistent associations between measurements and trajectories. Beyond ensuring motion consistency, these trajectories provide critical situational awareness in identifying target proximity and potential intersections. Inspired by the strong short-term spatio-temporal correlations exhibited by extended target motion, the proposed method leverages the feedback from predicted trajectories to quantify the interaction between targets.

Let $\boldsymbol{G}_k^{(c)}$ and $\boldsymbol{G}_{k+1}^{(n)}$ denote the centroids of the c-th measurement set at instant k and the n-th measurement set at $k+1$, respectively. To evaluate whether two measurement sets at different time instants originate from the same target, a distance-based metric is employed as follows [40]:

$$d\left(\boldsymbol{G}_k^{(c)}, \boldsymbol{G}_{k+1}^{(c)}\right) = \max\left(0, \boldsymbol{G}_{k+1}^{(c)} - \boldsymbol{G}_k^{(c)} - v_{\max}k\right) + \max\left(0, -\boldsymbol{G}_{k+1}^{(c)} + \boldsymbol{G}_k^{(c)} + v_{\min}k\right) \tag{19}$$

where $v_{\max}$ and $v_{\min}$ denote the maximum and minimum velocity of target state change. Generally, the covariance of the zero-mean Gaussian white noise is $\mathbf{Y}_k$, and the normalized square is denoted as follows:

$$D\left(\boldsymbol{G}_k^{(c)}, \boldsymbol{G}_{k+1}^{(c)}\right) = d\left(\boldsymbol{G}_k^{(c)}, \boldsymbol{G}_{k+1}^{(c)}\right)^{\mathrm{T}} \cdot \left[\mathbf{Y}_k + \mathbf{Y}_{k+1}\right]^{-1} d\left(\boldsymbol{G}_k^{(c)}, \boldsymbol{G}_{k+1}^{(c)}\right) \tag{20}$$

where $D\left(\boldsymbol{G}_k^{(c)}, \boldsymbol{G}_{k+1}^{(c)}\right)$ is random variables with χ^2 distribution and the degree of freedom is p. In consequence, the deviation threshold φ can be determined in the χ^2 distribution table. And $D\left(\boldsymbol{G}_k^{(c)}, \boldsymbol{G}_{k+1}^{(c)}\right) \leq \varphi$ denotes the measurements of the same target at adjacent moments. Therefore, the predicted location $\boldsymbol{\mu}_k^*$ from the movement estimation could be replaced as follows:

$$\boldsymbol{\mu}_k^* = \boldsymbol{\mu}_{k-1}^* + \frac{1}{k-3}\sum_{k=3}^{k-1} \|\hat{\boldsymbol{\mu}}_{k-1}^* - \hat{\boldsymbol{\mu}}_{k-2}^*\|_2 \tag{21}$$

where $\boldsymbol{\mu}_k^*$ denotes the predicted location of target at instant k, and $\|\hat{\boldsymbol{\mu}}_{k-1}^* - \hat{\boldsymbol{\mu}}_{k-2}^*\|_2$ denotes the euclidean distance between estimated optimal locations of targets at adjacent instants.

The target's shape at every instant is estimated as:

$$\boldsymbol{TR}^{(i)} = \left[\hat{\boldsymbol{\mu}}^{*(i)}; \hat{\boldsymbol{P}}^{*(i)}\right] \tag{22}$$

where $\boldsymbol{TR}^{(i)}$ denoted the $i-th$ target's trajectory including the location and shape information, and $\hat{\boldsymbol{P}}^{*(i)}$ denotes the $i-th$ target's shape information at every instant.

Once target's trajectories are successfully initiated, the possible location of the tracking targets should be the predicted location that estimated from the spatio-temporal domain information

$$\hat{\boldsymbol{\mu}}_c^{\zeta} = \hat{\boldsymbol{\mu}}_k^*, \boldsymbol{S} \in \text{Track establishment} \tag{23}$$

Meanwhile, the target's shape information should remain short-time stationarity when targets meet or approach, therefore, the target being tracked should keep shape characteristics itself before the intersection as follows:

$$\hat{\boldsymbol{P}}_c^{\zeta+1} = \frac{1}{k-1}\sum_{i=1}^{k-1}\hat{\boldsymbol{P}}_c^*(i), \boldsymbol{S} \in \text{Encounter} \tag{24}$$

Due to the characteristics of the extended target, the proposed method considers both the location and shape information to judge whether multiple targets encounter occurs. For example, let t_1 and t_2 be two extended target, respectively, and the proposed method judge targets approach or meet as follows:

$$\begin{aligned} \boldsymbol{S} &= \left\|\hat{\boldsymbol{\mu}}_{k+1}^{t_1} - \hat{\boldsymbol{\mu}}_{k+1}^{t_2}\right\|_2 - \left(\left|\max\sqrt{\hat{\boldsymbol{P}}_{k+1}^{t_1}}\right| + \left|\max\sqrt{\hat{\boldsymbol{P}}_{k+1}^{t_2}}\right|\right) \\ \text{s.t.} &\quad \left(\left\|\hat{\boldsymbol{\mu}}_k^{t_1} - \hat{\boldsymbol{\mu}}_k^{t_2}\right\|_2 - \left\|\hat{\boldsymbol{\mu}}_{k-1}^{t_1} - \hat{\boldsymbol{\mu}}_{k-1}^{t_2}\right\|_2\right) < 0 \end{aligned} \tag{25}$$

where $\left|\max\sqrt{\hat{\boldsymbol{P}}_{k+1}^{t_1}}\right|$ is the maximum value of shape information among target No.1. That the constraint condition is less than 0 means that the two targets are gradually approaching rather than away. Therefore, the proposed method considers that $\boldsymbol{S} < 0$ represents the targets meet.

3. Experiment results and discussion

This section evaluates the performance of the proposed TSF-GMM-EM method through a comprehensive experimental framework. First, the simulation scenarios and evaluation metrics are described. Subsequently, two specialized experiments are conducted to perform parameter sensitivity analysis and validate the effectiveness of the measurement partitioning algorithm. The overall tracking performance of the proposed method is then verified across two distinct scenarios as shown in Fig 1. All simulations are implemented in MATLAB R2021a and executed on an Apple M1 processor.

3.1 Experiment model and parameters

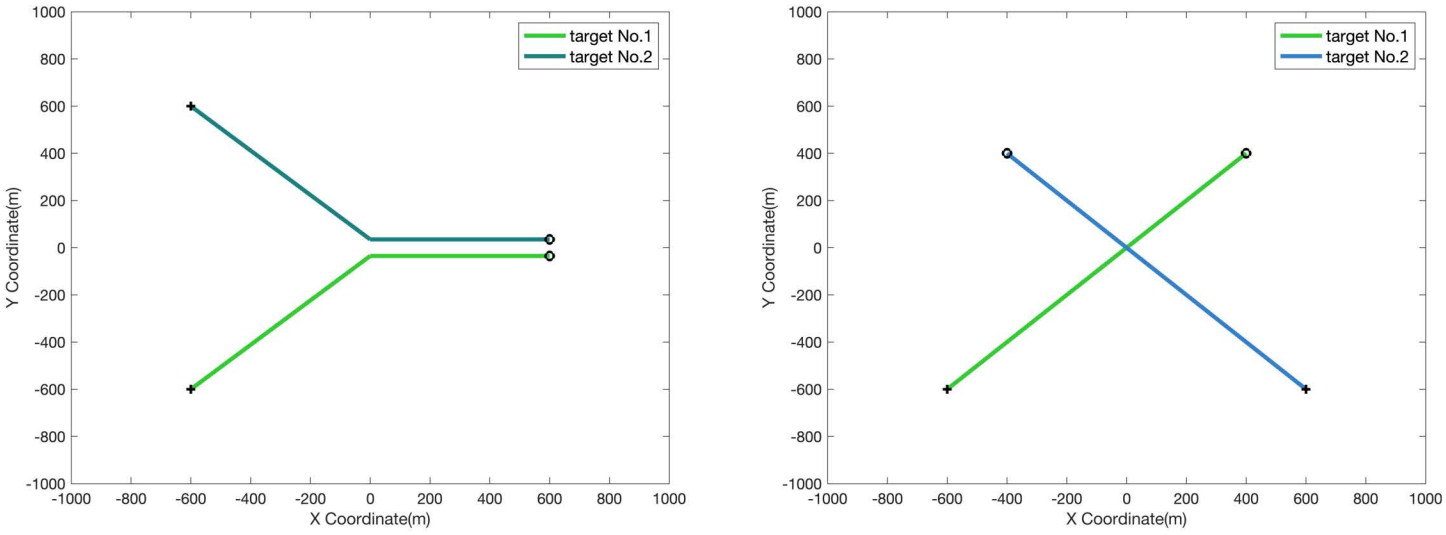
In both simulation scenarios, the number of measurements generated by an extended target is assumed to be Poisson-distributed at each time step, alongside the presence of clutter. The target and measurement states are formulated as follows:

$$\mathbf{x}_k = \left[x_k \;\; y_k \;\; v_k^x \;\; v_k^y\right]^{\mathrm{T}} \tag{26}$$

$$\mathbf{z}_k^{(i)} = \left[x_k^{(i)}, y_k^{(i)}\right]^{\mathrm{T}} \tag{27}$$

where x_k and y_k denotes the truth location of targets at instant k, the $x_k^{(i)}$ and $y_k^{(i)}$ denotes the $i-th$ measurements in batches of Cartesian x and y coordinates.

The state model are established as aforementioned linear Gaussian:



(a) Simulation scenario 1. (b) Simulation scenario 2.

Fig 1. Ground truths of simulation scenarios. (a) Simulation scenario 1. (b) Simulation scenario 2.

https://doi.org/10.1371/journal.pone.0357813.g001

$$\mathbf{x}_{k+1}^{(i)} = \mathbf{T}_k \mathbf{x}_k^{(i)} + \mathbf{G}_k \mathbf{Q}_k^{(i)} \tag{28}$$

where $\boldsymbol{Q}_k = \sigma_v^2 \mathbf{I}_2$ denotes the process noise, and $\mathbf{I}_2$ is a 2x2 identity matrix, $\sigma_v = 2$ m/s^2. The parameters of the model are set as:

$$\mathbf{T}_k = \begin{bmatrix} 1 & 0 & T & 0 \\ 0 & 1 & 0 & T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{29}$$

$$\mathbf{G}_k = \begin{bmatrix} T^2/2 & 0 & T & 0 \\ 0 & T^2/2 & 0 & T \end{bmatrix}^{\mathrm{T}} \tag{30}$$

where T is the sampling period, and it is set as one second in these experiments.

The measurements model is also a linear Gaussian model

$$\mathbf{z}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{e}_k^{(i)} \tag{31}$$

where $\mathbf{H}_k = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$. And the measurements occur with a ellipsoidal covariance that follows the Gaussian distribution "three-sigma" principle [41], the probability of value distribution in the range $[\mu - 3\sigma\ ,\ \mu + 3\sigma]$ is 99.74%.

In the two simulation scenarios, the probability of survival $P_S\left(\mathbf{x}_k^{(i)}\right)$ and the probability of detection $P_D\left(\mathbf{x}_k^{(i)}\right)$ for each target is the same as 0.99, and the number of clutter appear in measurement space at every instant follows a Poisson distribution with a mean of 10, i.e., $\gamma = 10$. The target birth follows a Poisson model with the intensity as follows:

$$v_b = \sum_{i=1}^{J_b} 0.1 \cdot \mathcal{N}\left(x; \boldsymbol{m}_b^{(i)}, \boldsymbol{K}_b\right) \tag{32}$$

where $\boldsymbol{K}_b = \text{diag}\left([100, 100, 25, 25]\right)$, J_b and $\boldsymbol{m}_b$ are different from the three simulation scenarios, then the values are set in the specific scenario, respectively. The pruning threshold is set to $T_{th} = 1e-10$, the merging threshold is set to $U = 8$, weight threshold is set to $\omega_{Th} = 0.5$ and maximum number of Gaussian terms is set to $J_{\max} = 200$ in ET-GM-PHD. The OSPA metric is used to evaluate the tracking performance [42].

$$\text{OSPA}_{p,c}(X_t, \hat{X}_t) = \left[\frac{1}{|\hat{X}_t|}\left(\min_{\pi \in \Pi_{|\hat{X}_t|}} \sum_{i=1}^{|X_t|} d_c(x^i, \hat{x}^{\pi(i)})^p\right) + \frac{c^p \cdot (|\hat{X}_t| - |X_t|)}{|\hat{X}_t|}\right]^{1/p} \tag{33}$$

where p and c are the order and the cutoff value of OSPA. $d_c\left(\mathbf{x}, \hat{\mathbf{x}}\right) = \min\left(c, d\left(\mathbf{x}, \hat{\mathbf{x}}\right)\right)$ is the cutoff distance between the cutoff parameter c and the euclidean distance between the states $\mathbf{x}$ and $\hat{\mathbf{x}}$.

3.2 Analysis of tracking performance

To evaluate the statistical consistency of the proposed method, 100 Monte Carlo trials are conducted for each simulation. In the following figures, ground truth trajectories are depicted as solid lines, while measurements—comprising both target originated points and clutter are denoted by the "×" marker. Additionally, the "O" notation along the trajectories indicates the spatial extent of the extended targets.

It should be noted that the benchmark methods vary across both scenarios. Specifically, the sub-partition distance method is specifically designed to handle target intersections. In scenarios without intersections, such as Scenario 1, the standard distance partitioning and sub-partition methods exhibit identical tracking performance; therefore, only the primary benchmarks are presented for clarity.

(a) **Simulation scenario 1** J_b is set to 2, the parameter $\boldsymbol{m}_b$ is set to $\boldsymbol{m}_b^1 = \left[-600, 600, 0, 0\right]^{\mathsf{T}}$ and $\boldsymbol{m}_b^2 = \left[-600, -600, 0, 0\right]^{\mathsf{T}}$, respectively. The target's shape are $\sigma_{e_1} = \sigma_{e_1} = 20$ m, corresponding to the mean of Poisson distribution $\gamma = 30$. Then, the birth intensity can be calculated as:

$$v_b = 0.1 \cdot \mathcal{N}\left(x; \boldsymbol{m}_{b,}^{(1)}, \boldsymbol{K}_b\right) + 0.1 \cdot \mathcal{N}\left(x; \boldsymbol{m}_{b,}^{(2)}, \boldsymbol{K}_b\right) \tag{34}$$

Figs 2 and 3 show the OSPA distance and the estimated target number of the evaluated methods over time. As illustrated in Fig 2, during the pre-intersection phase (0–50 seconds), all three methods track the targets accurately, maintaining low and stable OSPA errors.

However, a significant performance divergence occurs during the target crossing interval (50–55 seconds). Due to severe measurement association ambiguity in close proximity, both the Distance-Partitioning method and the PMBM method experience a pronounced error spike, with their OSPA distances climbing up to approximately 15. In contrast, the proposed TSF-GMM-EM-Partitioning framework effectively suppresses this transient error, keeping the peak OSPA distance below 8, which directly demonstrates its superior robustness against intersection ambiguities through the closed-loop trajectory feedback mechanism.

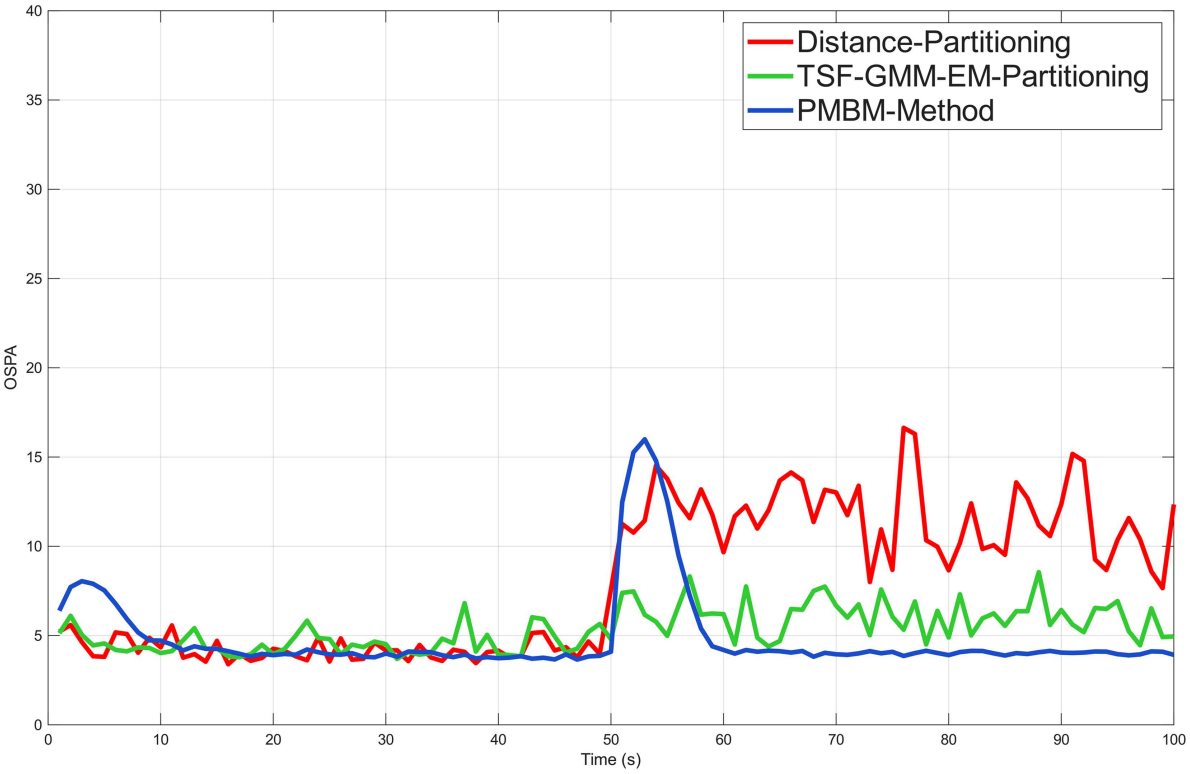


Fig 2. OSPA comparison in simulation scenario 1.

https://doi.org/10.1371/journal.pone.0357813.g002

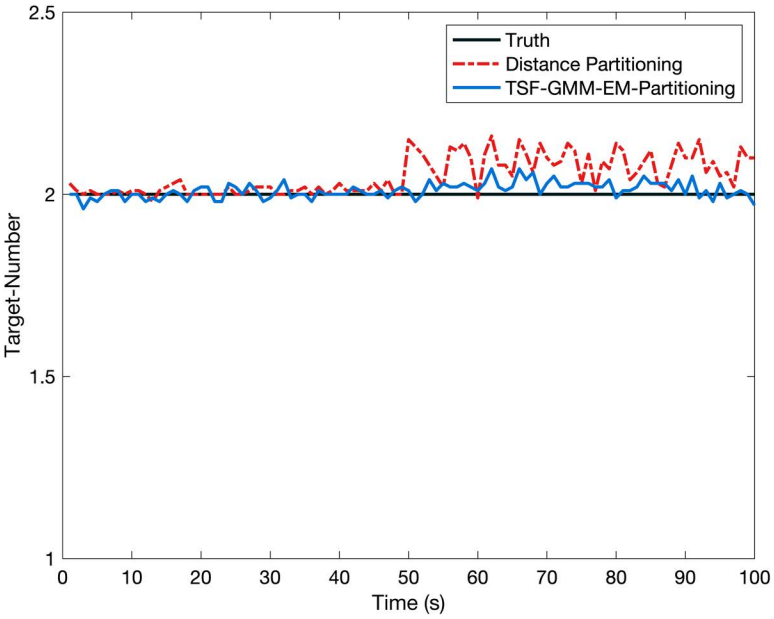


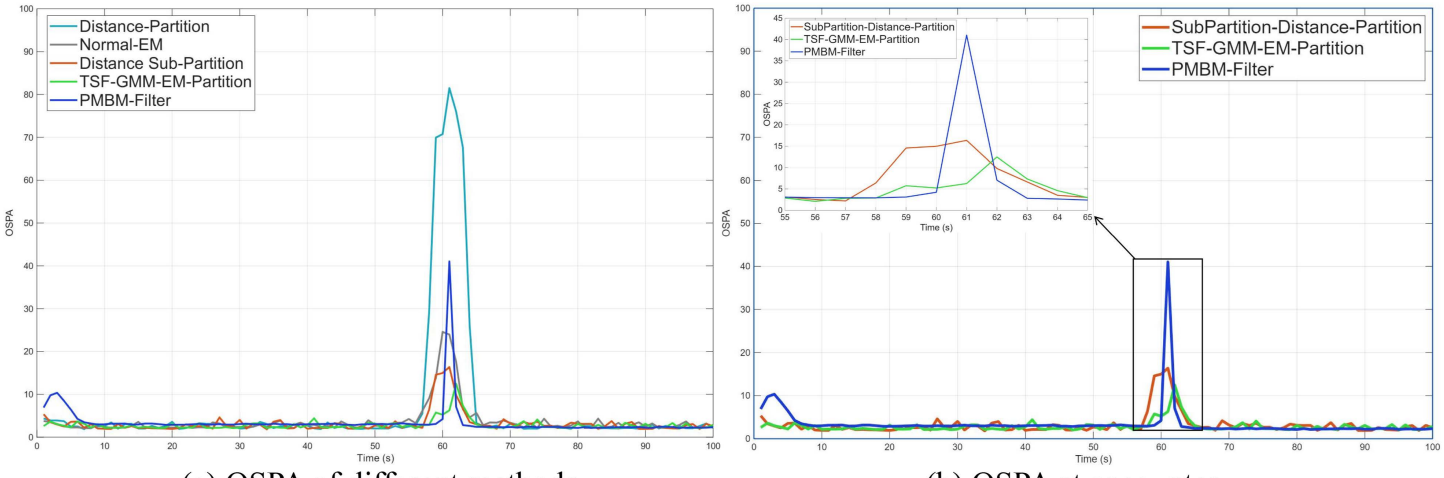
Fig 3. The comparison of number estimation in scenario 1.

https://doi.org/10.1371/journal.pone.0357813.g003

In the post-intersection phase (55–100 seconds), the traditional Distance-Partitioning method fails to recover, exhibiting a persistent and high OSPA error between 10 and 15 due to tracking divergence. Although its steady-state performance is slightly lower than that of the PMBM method, it still significantly outperforms the traditional Distance-Partitioning method, confirming its effectiveness in long-term tracking.

(b) **Simulation scenario 2** The parameters are set to σ_{e_1} = 15 m and σ_{e_2} = 6m, respectively, corresponding to the mean of Poisson distribution γ_1 = 30 and γ_2 = 15. J_b, $\boldsymbol{m}_b$ and the birth intensity are the same as in simulation scenario No.1. The two extended targets encounters during the period of 58–64 seconds, TSF-GMM-EM tracks the two targets accurately and continuously under clutter interference when they encounter.

OSPAs estimated by various methods are shown in Fig. 4. The OSPA of normal distance partition method is more than 80 metres, and it loses the targets when they encounters. The reason to failed track target can be attributed to no constraint on the target number. Normal EM method imposes certain constraints on the target number to reduce the tracking errors, however, it just achieve an OSPA of 24.5801 due to the absence of target shape information constraints. Since targets of different sizes produce different numbers of measurements, the existing method calculates the maximum likelihood function to determine sub-partition, then the target might be lost, therefore, this judgment way is not adopted in this paper. However, the target characteristics obtained by the proposed spatiotemporal information feedback are used in the distance sub-partition strategy, the measurement cells exceeding a certain number need to be sub-partition at the intersection moment. The sub-partition distance partition method and TSF-GMM-EM method achieve smaller OSPAs. In order to analyze the tracking performance in detailed, the OSPA curves during the period of 55–65 seconds are magnified and shown in Fig 4(b). During the period of 58–64 seconds, the total OSPAs of sub-partition distance method and TSF-GMM-EM method are 72.2747 and 44.5618, respectively. Constrained by the priori information that the track and the number of multiple targets should be consistent in a short period, TSF-GMM-EM method achieve an OSPA of 5.2318 at 60 second, and it tracks targets more accurately than other methods. The PMBM filter exhibits unstable target tracking at the intersection time of 61 seconds, leading to an increase in OSPA. However, it can quickly adjust and maintain effective tracking in both preceding and subsequent conditions. In the post-intersection phase (65–100 seconds), all methods successfully recover, with their steady-state OSPA errors converging uniformly to a very low level.



(a) OSPA of different methods. (b) OSPA at encounter.

Fig 4. OSPA comparison by different method in simulation scenario 2. (a) OSPA of different methods. (b) OSPA at encounter.

https://doi.org/10.1371/journal.pone.0357813.g004

The target number estimated by various methods is shown in Fig 5. When the two extended targets encounters, the target numbers estimated by normal distance partition method, normal EM method and TSF-GMM-EM are 1, 1.7 and 1.8, respectively. However, the target number estimated by sub-partition distance partition method and TSF-GMM-EM method is very close to 2, and the estimated fluctuation of the latter is smaller than the former.

3.3 Discussion

In this section, two aspects are considered: the robustness and the performance of measurement partition.

1) **Performance of measurement partition**

In order to evaluate the performance of measurement partition algorithm, the measurement partition accuracy is defined as

$$\rho = \frac{S_N}{N_{z,k}} \tag{35}$$

where S_N represents the cardinality of correctly identified and associated target-originated measurement points, and $N_{z,k}$ denotes the total number of available measurements received by the sensor at time step k. The maximum theoretical value of ρ is 1.

The measurement partition accuracy of 100 Monte Carlo runs is shown in Fig 6. It can be seen that the measurement partition accuracy decrease as the clutter density increases. Nevertheless, the measurement partition accuracy ρ is not lower than 0.946 under the test range of the clutter density, and the proposed method gets outperformance.

In this discussion section, two extended targets are modeled as originating from initial coordinates (244.9 m, 35 m) and $(244.9\,\text{m}, -35\,\text{m})$, respectively, with their spatial extensions expanding via a Gaussian covariance to an ellipsoidal boundary of $3\sigma_{e_1} = 3\sigma_{e_2} = 45\,\text{m}$. To aggressively challenge the algorithm's resilience against initialization errors, a stochastic

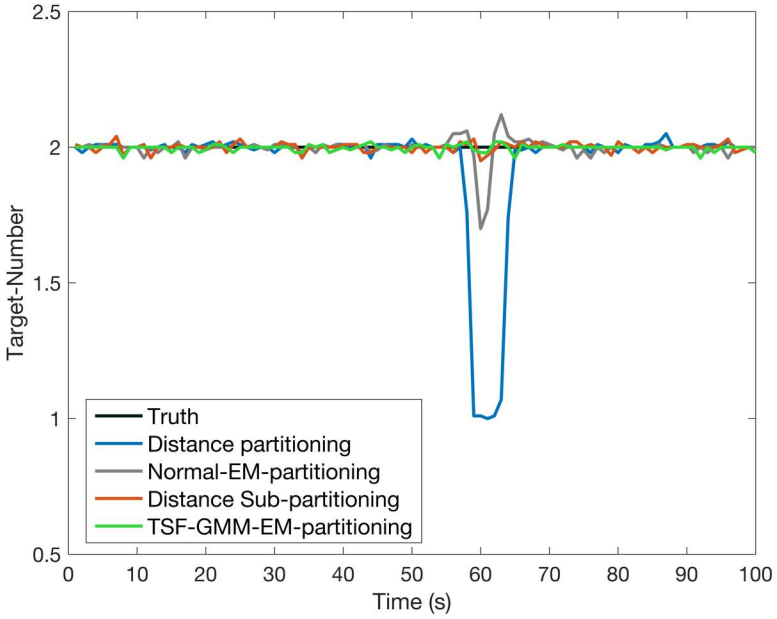


Fig 5. The comparison of number estimation in simulation scenario 2.

https://doi.org/10.1371/journal.pone.0357813.g005

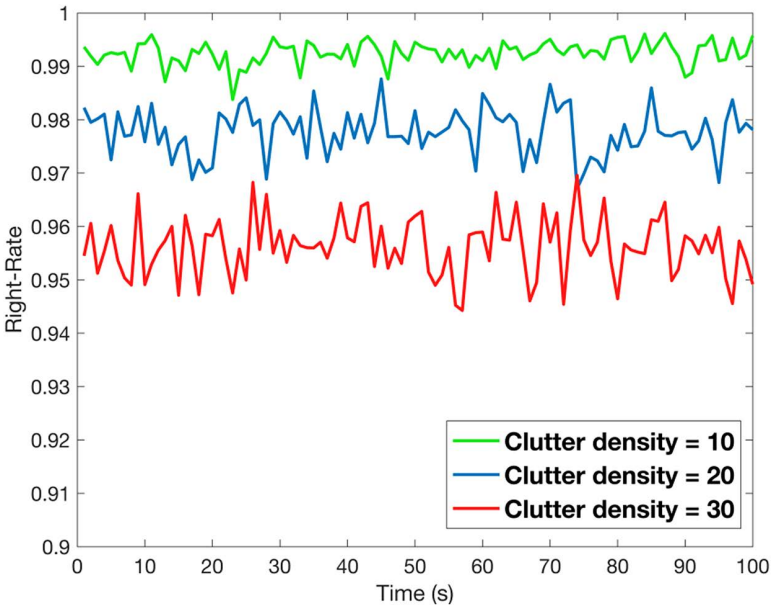


Fig 6. The cluster accuracy of TSF-GMM-EM with different clutter density.

https://doi.org/10.1371/journal.pone.0357813.g006

perturbation of $\pm 10\%$ is injected into the initial predicted locations. This forces the seeding cluster centers to deviate randomly within the bounded rectangular region of $[244.9 \pm 10\%, 35 \pm 10\%]$ at the onset of the tracking pipeline.

The empirical partitioning accuracy aggregated over 100 Monte Carlo runs is illustrated in Fig 6. An insightful phenomenon revealed by the curves is that the average clustering accuracy exhibits a slight, monotonic degradation as the ambient clutter density scales up from $\gamma = 10$ to $\gamma = 30$. This performance drop is physically attributed to the heightened probability of dense false alarms falling inside the overlapping tracking validation regions, which naturally injects noise into the posterior association weights. Nevertheless, because our trajectory situation feedback loop continuously feeds long-term temporal priors back to anchor the clustering centers, the conventional GMM-EM liability of trapping into uninformative local maxima is completely bypassed. Consequently, the accuracy curves remain remarkably flat and stable across the entire 100-second horizon without suffering from error propagation. Statistically, the partitioning accuracy ρ is strictly maintained above 0.946 even under the harshest clutter condition ($\gamma = 30$), fully confirming the exceptional spatial-temporal discriminative power of the proposed architecture.

2) **Robustness**

As illustrated by the clutter-sensitivity curves in Fig 7 (compiled over the critical 51–100 seconds interval), a global characteristic shared by both frameworks is that the OSPA distance elevates monotonically as the clutter density increases from $\gamma = 5$ to $\gamma = 40$. From a filtering standpoint, this upward trend is mathematically inevitable: an inflation in the false alarm rate forces more clutter points to penetrate the tracking validation gates, which directly corrupts the recursive pseudo-likelihood updates within the GM-PHD filter and biases the centroid estimation.

Nevertheless, the proposed TSF-GMM-EM method demonstrates anti-noise robustness, with its green curve consistently maintaining a lower tracking error than that of the traditional distance partitioning (DP) method across all noise thresholds. When the extended targets closely approach or encounter spatial intersections, the traditional DP method degrades drastically because it operates purely on a single-frame spatial snapshot; lacking cross-frame memory, it easily merges close multi-target point clouds and adjacent clutter into a single cell, leading to track fragmentation. In sharp

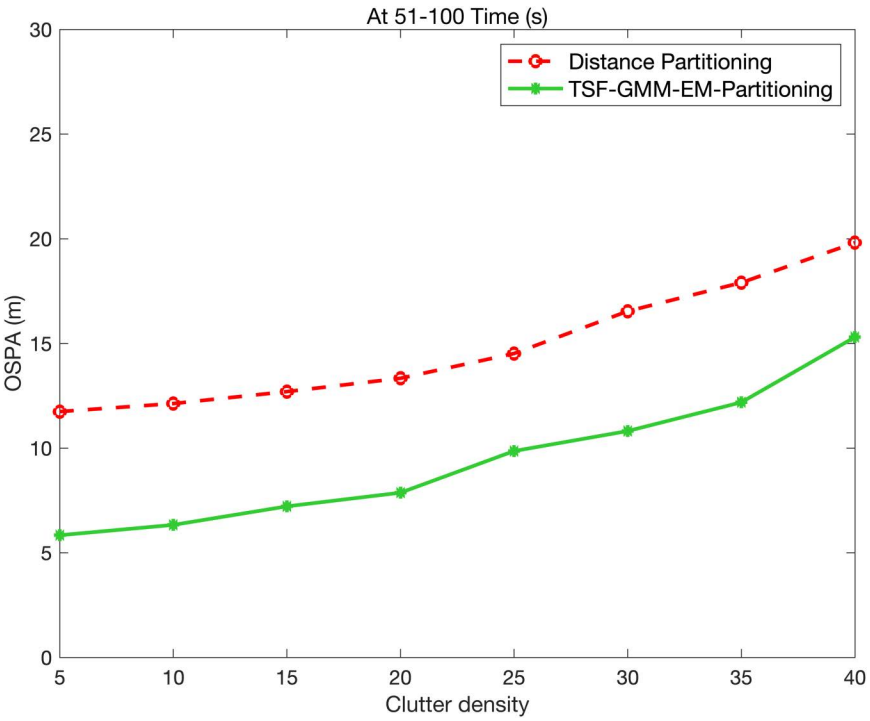


Fig 7. Performance comparison in terms of different clutter rates.

https://doi.org/10.1371/journal.pone.0357813.g007

contrast, by leveraging the predicted states derived from the temporal continuity of target trajectories, our proposed TSF-GMM-EM method enforces strict spatio-temporal consistency constraints onto the mixture model initialization. This closed-loop interaction effectively safeguards the spatial partitioning layer against high-density clutter interference, ensuring that the proposed method achieves robust outperformance in complex noise tracking scenes.

4. Conclusion

In this paper, we developed a TSF-GMM-EM method within the framework of the GM-PHD filter to tackle the persistent degradation challenges in multi-extended target tracking. By establishing a closed-loop trajectory feedback framework, this approach successfully realized a bidirectional interaction between cross-frame temporal trajectory evolution and single-frame spatial measurement partitioning, and effectively resolved severe data association ambiguities in dense clutter environments. Furthermore, the introduced prior-guided EM initialization strategy strategically leveraged predicted target kinematic states to seed the soft-clustering process. This mechanism enforced strict spatio-temporal consistency between established trajectories and incoming point clouds, which successfully prevented the conventional GMM-EM pipeline from trapping into uninformative local maxima and drastically mitigated tracking loss or cardinality errors during target intersections. Experimental results demonstrate that the proposed method achieves robust performance in complex tracking scenarios, particularly during target intersections.

Future investigations will focus on expanding the capability of the TSF-GMM-EM approach to accommodate highly non-linear target maneuvers and irregular, complex shape profiles by integrating spatiotemporal Gaussian Processes. Furthermore, we plan to explore alternative probabilistic and optimization paradigms to enrich the theoretical foundations of our framework. Specifically, we aim to extend the current closed-loop filtering structure into a formal Bayesian smoothing framework or a structured MAP estimator with temporal priors to mathematically bound the long-term tracking

performance. Investigating variational inference (VI) as an alternative to standard EM partitioning also represents a highly promising avenue, particularly for handling complex scenarios with completely unknown and dynamic target counts through automatic model selection. Additionally, validating the algorithm's scalability and its adaptive clutter-rate estimation capabilities under heterogeneous clutter profiles using real-world maritime or automotive sensor datasets remains a vital and promising direction for our future research.

Author contributions

Conceptualization: Yuting Zhang.

Project administration: Jiye Li.

Software: Jiye Li.

Supervision: Hantao Li.

Validation: Hantao Li.

Writing – original draft: Yuting Zhang.

Writing – review & editing: Yuting Zhang.

References

1. Kim J. Tracking multiple underwater targets using adaptive Gaussian mixture probability hypothesis density filter with unknown clutter rate. IEEE Trans Aerosp Electron Syst. 2024;60(6):9154–62. https://doi.org/10.1109/taes.2024.3437343
2. Indhumathi S, Clement C. Convex-based lightweight feature descriptor for Augmented Reality Tracking. PLoS One. 2024;19(7):e0305199. https://doi.org/10.1371/journal.pone.0305199 PMID: 39024253
3. Liang Z, Liu F, Gao J. Improved GGIW-PHD filter for maneuvering non-ellipsoidal extended targets or group targets tracking based on sub-random matrices. PLoS One. 2018;13(2):e0192473. https://doi.org/10.1371/journal.pone.0192473 PMID: 29444144
4. Zhu Y, Zhou S, Gao G, Zou H, Lei L. Extended emitter target tracking using GM-PHD filter. PLoS One. 2014;9(12):e114317. https://doi.org/10.1371/journal.pone.0114317 PMID: 25490206
5. Guo Y, Li Y, Tharmarasa R, Kirubarajan T, Efe M, Sarikaya B. GP-PDA filter for extended target tracking with measurement origin uncertainty. IEEE Trans Aerosp Electron Syst. 2018;55(4):1725–42.
6. Li F, Zhou M, Ding Y. An adaptive online co-search method with distributed samples for dynamic target tracking. IEEE Trans Contr Syst Technol. 2018;26(2):439–51. https://doi.org/10.1109/tcst.2017.2669154
7. Ma Y, Wang Z, Yang H, Yang L. Artificial intelligence applications in the development of autonomous vehicles: a survey. IEEE/CAA J Autom Sin. 2020;7(2):315–29. https://doi.org/10.1109/jas.2020.1003021
8. Kebria PM, Khosravi A, Salaken SM, Nahavandi S. Deep imitation learning for autonomous vehicles based on convolutional neural networks. IEEE/CAA J Autom Sin. 2020;7(1):82–95. https://doi.org/10.1109/jas.2019.1911825
9. Kaulbersch H, Honer J, Baum M. EM-based extended target tracking with automotive radar using learned spatial distribution models. 2019 22th International Conference on Information Fusion (FUSION). IEEE; 2019. p. 1–8.
10. Angelova D, Mihaylova L. Extended object tracking using Monte Carlo methods. IEEE Trans Signal Process. 2008;56(2):825–32. https://doi.org/10.1109/tsp.2007.907851
11. Koch JW. Bayesian approach to extended object and cluster tracking using random matrices. IEEE Trans Aerosp Electron Syst. 2008;44(3):1042–59. https://doi.org/10.1109/taes.2008.4655362
12. Zhang H, Ge H, Yang J, Yuan Y. A GM-PHD algorithm for multiple target tracking based on false alarm detection with irregular window. Signal Process. 2016;120:537–52. https://doi.org/10.1016/j.sigpro.2015.10.007
13. Zhao B, Peng G, Cheng F, Zhang W, Lin H, Wang J. A novel transformer-aided multiple model algorithm for maneuvering target tracking. IEEE Sensors J. 2024;24(23):39578–92. https://doi.org/10.1109/jsen.2024.3482563
14. Chang W, Chen D, Cao H, Bu L, Wang C, Fu T. Simulation of deep learning-based multitarget track association for ballistic target groups. IEEE Sens J. 2025;25(19):37429–44. https://doi.org/10.1109/jsen.2025.3601590
15. Qin Z, Liang Y, Li K, Zhou J. Measurement-driven sequential random sample consensus GM-PHD filter for ballistic target tracking. Mech Syst Signal Process. 2021;155:107407. https://doi.org/10.1016/j.ymssp.2020.107407
16. Wei J, Luo F, Chen S, Qi J. Robust fusion of GM-PHD filters based on geometric average. Signal Process. 2023;206:108912. https://doi.org/10.1016/j.sigpro.2022.108912

17. Vo B-N, Ma W-K. The Gaussian mixture probability hypothesis density filter. IEEE Trans Signal Process. 2006;54(11):4091–104. https://doi.org/10.1109/tsp.2006.881190
18. Mahler R. Statistical multisource-multitarget information fusion. Artech; 2007.
19. Mahler R. PHD filters for nonstandard targets, I: Extended targets. 2009 12th International Conference on Information Fusion. IEEE; 2009. p. 915–21.
20. Granstrom K, Orguner U. A phd filter for tracking multiple extended targets using random matrices. IEEE Trans Signal Process. 2012;60(11):5657–71. https://doi.org/10.1109/tsp.2012.2212888
21. Baisa NL, Wallace A. Development of a N-type GM-PHD filter for multiple target, multiple type visual tracking. J Vis Commun Image Represent. 2019;59:257–71. https://doi.org/10.1016/j.jvcir.2019.01.026
22. Granstrom K, Lundquist C, Orguner O. Extended target tracking using a Gaussian-mixture PHD filter. IEEE Trans Aerosp Electron Syst. 2012;48(4):3268–86. https://doi.org/10.1109/taes.2012.6324703
23. Akbari B, Zhu H. Tracking dependent extended targets using multi-output spatiotemporal Gaussian processes. IEEE Trans Intell Transp Syst. 2022;23(10):18301–14. https://doi.org/10.1109/tits.2022.3154926
24. Wahlström N, Özkan E. Extended target tracking using Gaussian processes. IEEE Trans Signal Process. 2015;63(16):4165–78.
25. Cheng X, Ji H, Zhang Y. Multiple extended target joint tracking and classification based on GPs and LMB filter. IEEE Signal Process Lett. 2024;31:1139–43. https://doi.org/10.1109/lsp.2024.3383954
26. Yang J, Xiong Y, Cao X, Peng C, Yi W. Joint tracking and classification of vehicles with the PHD filter and Gaussian processes. 2024 27th International Conference on Information Fusion (FUSION). IEEE; 2024. p. 1–7.
27. Li Z, Zhan R, Zhuang Z. Multiple extended target joint tracking and classification using RMM and TPHD filter. IEEE Signal Process Lett. 2025;32:4164–8. https://doi.org/10.1109/lsp.2025.3625895
28. Al-Haddad LA, Jaber AA, Dhahir MK, Nagim HY, Algburi ZI. Characterization and prediction of femtosecond laser induced tracks in silver-containing zinc phosphate glass. In: SYSTEM. 2024. p. 10–9.
29. Li Y, Du Yt, Cui L, Qiu W. Interactive multiple-model probability hypothesis density filter for coexisting point and extended targets tracking. Fifth International Conference on Information Science and Education (ICISE-IE 2025). vol. 13559. SPIE; 2025. 1355917 p.
30. Qin Z, Kirubarajan T, Liang Y. Application of an efficient graph-based partitioning algorithm for extended target tracking using GM-PHD filter. IEEE Trans Aerosp Electron Syst. 2020;56(6):4451–66. https://doi.org/10.1109/taes.2020.2990803
31. Sun L, Yu H, Fu Z, He Z, Tao F. Tracking of multiple closely spaced extended targets based on prediction-driven measurement sub-partitioning algorithm. Appl Sci. 2020;10(14):5004. https://doi.org/10.3390/app10145004
32. Qiu W, Li Y, Cui L, Du Y. A scalable multiple extended targets tracking based on the non-homogeneous poisson process and PHD filter. 2025 17th International Conference on Signal Processing Systems (ICSPS). IEEE; 2025. p. 165–70.
33. Wang W, Xi J, Zhao D. Learning and inferring a driver's braking action in car-following scenarios. IEEE Trans Veh Technol. 2018;67(5):3887–99. https://doi.org/10.1109/tvt.2018.2793889
34. Cao M, Wang R, Chen N, Wang J. A learning-based vehicle trajectory-tracking approach for autonomous vehicles with LiDAR failure under various lighting conditions. IEEE/ASME Trans Mechatron. 2022;27(2):1011–22. https://doi.org/10.1109/tmech.2021.3077388
35. Aftab W, Hostettler R, De Freitas A, Arvaneh M, Mihaylova L. Spatio-temporal gaussian process models for extended and group object tracking with irregular shapes. IEEE Trans Veh Technol. 2019;68(3):2137–51. https://doi.org/10.1109/tvt.2019.2891006
36. Granström K, Lundquist C, Orguner U. A Gaussian mixture PHD filter for extended target tracking. 2010 13th International Conference on Information Fusion. IEEE; 2010. p. 1–8.
37. Wang W, Xi J, Hedrick JK. A learning-based personalized driver model using bounded generalized Gaussian mixture models. IEEE Trans Veh Technol. 2019;68(12):11679–90. https://doi.org/10.1109/tvt.2019.2948911
38. Li X, Wang W, Roetting M. Estimating driver's lane-change intent considering driving style and contextual traffic. IEEE Trans Intell Transp Syst. 2019;20(9):3258–71. https://doi.org/10.1109/tits.2018.2873595
39. Altilio R, Di Lorenzo P, Panella M. Distributed data clustering over networks. Pattern Recognit. 2019;93:603–20. https://doi.org/10.1016/j.patcog.2019.04.021
40. Zhang C, Li Z, Zhu Y, Luo Z, Qin T. Improved GM-PHD filter with birth intensity and spawned intensity estimation based on trajectory situation feedback control. Remote Sens. 2022;14(7):1683. https://doi.org/10.3390/rs14071683
41. Jeong J, Park E, Han WS, Kim K, Choung S, Chung IM. Identifying outliers of non-Gaussian groundwater state data based on ensemble estimation for long-term trends. Journal of Hydrology. 2017;548:135–44. https://doi.org/10.1016/j.jhydrol.2017.02.058
42. Schuhmacher D, Vo B-T, Vo B-N. A consistent metric for performance evaluation of multi-object filters. IEEE Trans Signal Process. 2008;56(8):3447–57. https://doi.org/10.1109/tsp.2008.920469