From Target Tracking to Targeting Track: A Data-Driven Yet Analytical Approach to Joint Target Detection and Tracking

Tiancheng Li[a], Yan Song[a], Hongqi Fan[b]

[a]*Key Laboratory of Information Fusion Technology (Ministry of Education), School of Automation, Northwestern Polytechnical University, Xi'an 710072, China. E-mail: t.c.li@nwpu.edu.cn.*

[b]*National Key Laboratory of Science and Technology on ATR, National University of Defense Technology, Changsha 410073, Hunan, China. E-mail: fanhongqi@nudt.edu.cn*

Abstract

This paper addresses the problem of real-time detection and tracking of a non-cooperative target in the challenging scenario with almost no a-priori information about target birth, death, dynamics and detection probability. Furthermore, there are false and missing data at an unknown yet low rate in the measurements. The only information given in advance is about the target-measurement model and the constraint that there is no more than one target in the scenario. To solve these challenges, we model the movement of the target by using a polynomial trajectory function of time (T-FoT), which aims to estimate the *continuous-time trajectory* of the target rather than a series of discrete-time point estimates as is done in most existing filters/trackers. Data-driven T-FoT initiation and termination strategies are proposed for identifying the (re-)appearance and disappearance of the target. During the existence of the target, real target measurements are distinguished from clutter if the target indeed exists and is detected, in order to update the T-FoT at each scan for which we design a least-squares estimator. Overall, our approach is Markov-free, data-driven yet analytical. Simulations using either linear or nonlinear systems are conducted to demonstrate the effectiveness of our approach in comparison with the Bayes optimal Bernoulli filters. The results show that our approach is comparable to the perfectly-modeled filters, even outperforms them in some cases while requiring much less a-priori information and computing much faster.

Keywords: Target tracking, target detection, trajectory fitting, least-squares estimation, Bayes inference

1. Introduction

Target tracking that involves the online estimation of the trajectory of a target has been a long standing research topic and plays a significant role in aerospace, traffic, defence, robotics, etc. [1] In this work, we focus on an important class of targets with simple and smooth trajectory such as airplane, train, ship and so on. Here, the smoothness of the trajectory is closely related with the dynamical equations of the target based on differential calculation. The standard approach since the ground-breaking Kalman filter (KF) [2], is to design a state space model (SSM) consisting of a Markov-jump model to describe the dynamics of the target and a measurement model to relate the measurement of the sensors with the state of the target. In addition, models are needed to characterize the background clutter-measurements and misdetection events. However, in many cases especially those for non-cooperative targets [3], these models except the measurement function are unavailable in advance and intractable to be identified accurately online. This leads to a great challenge to the use of any filters.

Instead of estimating the *discrete-time state* of the target based on a sophistically designed SSM, we are actually more interested in estimating the *continuous-time trajectory/track* of the target in the context of target tracking. That is, we seek direct estimating the trajectory. Once such a trajectory is obtained, the position (and also the velocity and acceleration which correspond to the first and second order derivatives

of the position against time) of the target at a particular time can be easily inferred from the trajectory. Moreover, high-level important information such as the class of the target, the motion model/pattern/feature and so on may be able to be inferred from the continuous-time trajectory but not from a series of discrete point estimates. Existing approaches to target trajectory estimation can be categorized in the following two groups, depending on modeling the trajectory whether as a discrete-time series of position points or as a continuous-time curve. The latter which describes the movement of the target in continuous time is the focus of this paper.

1.1. Discrete-time Trajectory Estimation

There are various SSM-based studies that recursively estimate the discrete-time-series state set based on the measurement sequence. Typically, it is given by a trajectory of a prescribed dynamical model such that the output of the model best fits a series of measurements. This is often referred to generally as data assimilation which have found many applications [4, 5, 6]. Different from the stochastic modeling of the state process, deterministic Markov-jump models, namely using no random variable $\mathbf{w}_k$ in (1), have been used in the so-called shadowing filters/smoothers [7, 8], moving horizon estimator [9, 10] and Gauss-Newton filter [11, 12]. In particular, the Gauss-Newton filter that models the state transition by a deterministic differential equation is Cramér-Rao consistent (providing minimum variance) [11]. These approaches, mainly designed for some specific systems, avoid the difficulties to accurately model the process noises and are advantageous in handling constraints and measurement singularities. They, however, still rely heavily on the Markov-jump model, which is essentially vulnerable to target maneuvering and unknown input (e.g., non-zero acceleration in the context of target tracking). To handle unknown inputs/noises in the motion of the target, an alternative is the minimum model error estimator [13, 14] which, different from the classic KF, requires no a-priori statistics on the form of the model error but determines it as a part of the solution.

In the multi-target case, by adding an unique label to each target, the target trajectory can be consequently given by the time-series labelled states [15, 16]. Despite its promising performance, an insightful discussion on the optimality of linking discrete-time points is given in [17]. Another relevant approach models the trajectories of a random number of targets as a random finite set (RFS) to be estimated [18]. Compared to the point-state RFS, the trajectory RFS requires much higher computation and needs to solve the problem of trajectories of different lengths in the same RFS.

1.2. Continuous-time Trajectory Fitting

Continuous-time trajectory curve described by a function contains more information than the discrete-time point set. For example, one can learn more feature or class information of the target from the continuous-time trajectory rather than from the discrete-time point set. In fact, signal processing stems from the interpolation and extrapolation of a sequence of data that were viewed as a realization of a stochastic process [19]. Data regression or fitting is a self-contained mathematical problem and a prosperous research theme by its own, which has proven to be a powerful and universal method for pattern learning and time series data prediction, and has the inherent advantages of dealing with outliers or missed detections [20], especially when adequate analytical solutions may not exist. In particular, there have been a large number of efforts devoted to trajectory curve fitting among different dimensions of the state space in various ways, e.g., [21, 22, 23, 24, 25, 26, 15, 27, 28], most of which, however, are *non-recursive* over time and not designed

for online tracking. Differently, as the few attempts that assume a spatio-temporal trajectory for tracking, [29, 30, 31, 32, 33] designed fitting cardinal trajectory-splines for each dimension. The resulting trajectory is a continuous-time function of state, the same to our approach [34, 35, 36]; see Section 2. Even more advanced, machine learning and neural networks were employed for trajectory fitting [37, 38, 39, 40], which, however, typically require an even larger amount of training data generated from proper a-priori models; a cutting-edge review of Bayesian learning for data analysis can be found in [41]. These approaches are usually not analytical and suffer from lack of interpretability of the parameters in the networks. Either way, existing trajectory fitting approaches have not taken into account realistic tracking issues such as clutter measurements, missed detection and target death, etc. as we will do in this paper.

What is more, it is necessary to identify in real-time whether the target is present or absent, especially when the target is non-cooperative. This task is often referred to separately as "target detection" and is a prerequisite for tracking. As such, a reasonable solution is to address target detection and tracking jointly, as required in the context of realistic target tracking. Following this thinking, the terminology of "tracking" means more than "filtering" does.

1.3. Joint Detection and Tracking (JDT)

An indispensable task in realistic target tracking is target detection that involves identifying whether the target is present (or how many targets are present in the multi-target case) since the target may randomly appear in or disappear from the field of view. There have been a plurality of SSM-based JDT approaches, e.g., [42, 43, 44, 45]. One of the most attractive theories for this purpose is the RFS [46] of which the most known single-target detection and tracking filter is the Bernoulli filter (BF) that is exact Bayes optimal [46, 47, 48]. Given exact statistics about the target birth, death, dynamics and detection probability and clutter models and rate, such a single target tracking problem can be properly solved by the BF. This, however, is often too ideal to be true for non-cooperative targets as these statistical information are generally unknown, time/space-varying and can hardly be exactly characterized. In contrast, it becomes much challenging when there is little a-priori statistical information about the target which is just the actual need in many problems. While many effects have been devoted to online estimating one or two of these issues, e.g., fruitful achievements for noise identification alone [49, 50], no work addresses all in one. This motivates our work in this paper.

1.4. Contribution and Paper Organization

In this paper, the target may randomly appear, disappear and re-appear anytime and anywhere in the surveillance scenario while the number of targets is no more than one. The only information available a-priori is the statistics of the measurements. For non-cooperative target tracking with little a-priori information, we earlier proposed to model the target movement by using a trajectory function of time (T-FoT) which reformulates the tracking as an online curve fitting problem [34, 35] for which the least-squares (LS) approach plays a key role. Issues of false and missing data, as well as the connection of the approach with the classic KF, have been further addressed in [36], based on the prerequisite that the target birth/initial position, velocity and starting time are given in advance and that the target always exists in the scenario. In this work, we relax these restrictive assumptions and further address the random (in both the state space and

time domain) birth, death and even re-appearance of the target, improving the algorithm for better identifying false and missing measurements. All aim at a computationally efficient, data-driven JDT approach that suits the challenging, unknown background.

The remainder of the paper is organized as follows. Preliminary work is briefly addressed in Section 2. Scenario assumptions and models we consider are given in Section 3. The proposed approach is given in Section 4. Simulation results of our approach are given in Section 5, in comparison with the state-of-the-art, properly modelled BF [47, 48]. The paper is concluded in Section 6.

2. Preliminaries

2.1. Conservational Models and An Alternative

In this work, we focus on the single target detection and tracking problem. This problem is usually solved based on the SSM with additive noises which can be described as follows

$$\mathbf{x}_k = f_k(\mathbf{x}_{k-1}) + \mathbf{w}_k, \tag{1}$$

$$\mathbf{y}_k = h_k(\mathbf{x}_k) + \mathbf{v}_k. \tag{2}$$

where $k \in \mathbb{N}$ indicates the time-instant, $\mathbf{x}_k \in \mathbb{R}^{D_\mathbf{x}}, \mathbf{y}_k \in \mathbb{R}^{D_\mathbf{y}}$, $\mathbf{w}_k \in \mathbb{R}^{D_\mathbf{x}}$ and $\mathbf{v}_k \in \mathbb{R}^{D_\mathbf{y}}$ denote the $D_\mathbf{x}$-dimensional state, the $D_\mathbf{y}$-dimensional measurement and their respective noises [49, 50] at time k, respectively.

The SSM facilitates recursive optimal/suboptimal filtering [51, 52] in the Bayesian fashion whose performance, however, heavily relies on the match between the Markov model (1) and the true target dynamics which is intractable to be identified [53], as the terminology *non-cooperative* indicates. From time to time, non-cooperative target may maneuvers [54, 55] or there is unknown input [52] and so on, adding difficulties to precisely capture its kinematic model as in (1). This leads to a fundamental challenge in practice.

To eschew this challenge, we consider modeling the ground truth of the target motion by an engineering-friendly, spatio-temporal T-FoT $\mathbf{x}_t = f(t)$, where the target state $\mathbf{x}_t$ is limited to its coordinate position in the following unless otherwise stated. Moreover, the velocity and acceleration can be obtained from the derivatives of the T-FoT over time. In particular, we parameter the T-FoT $f(t)$ by $F(t;\mathbf{C})$ using a set of coefficients $\mathbf{C}$ that determine the trajectory based on the available measurements. This may be expressed as follows

$$\mathbf{x}_t = F(t;\mathbf{C}) + \epsilon(t), \tag{3}$$

where $t \in \mathbb{R}^+$ indicates the (positive) continuous-time, $\epsilon(t)$ denotes the T-FoT residual function-of-time, namely the approximate error of $F(t;\mathbf{C})$ to the real T-FoT $f(t)$.

By using the T-FoT, what is estimated now is the trajectory parameters $\mathbf{C}$. This formulation is naturally appealing to an important type of targets with smooth motion patterns such as aircraft, train, cruise ship, etc. The trajectory as given by (3) may still be smooth even the target maneuvers [34]. A smooth trajectory basically indicates a small process noise in the classic Markov motion model. More formally, we have the following definition on the smoothness of a curve function: the T-FoT is twice continuously differentiable and is said to be β-smooth in a time-window $[k', k]$ if $\forall k' \leq t \leq k$,

$$\| \nabla^2 F(t;\mathbf{C})\|_\mathrm{s} \leq \beta, \tag{4}$$

where $\|\cdot\|_s$ denotes the spectral norm (i.e. the largest singular value in different dimensions).

The constraint (4) upper bounds the acceleration of the target by β. In the case we consider, β is relatively insignificant as compared with the uncertainty of measurements.

2.2. Sliding Time-window LS T-FoT Fitting

To account for a-priori model information, one may define a constraint or penalty/regularization factor $\Omega_F(\mathbf{C})$ within the optimization [34, 35, 56]. Constrained fitting has actually been intensively studied [57]. Since we focus on non-cooperative targets in this work, we do neither assume constraints nor use any a-priori target models. Alternatively, a penalty factor $\Omega_F(\mathbf{C})$ may be used for measuring the disagreement between the fitting function and the constraint. For example, $\Omega_F(\mathbf{C}) := (F(t;\mathbf{C})-\mathbf{x}_t)^{\mathrm{T}}(F(t;\mathbf{C})-\mathbf{x}_t)$ measures the mismatch between the fitting trajectory and the state $\mathbf{x}_t$ that the target passes by at time t.

Obviously, the *best* T-FoT $F(t;\mathbf{C})$ should minimize the approximate error $|\epsilon(t)|$. But unfortunately, $\mathbf{x}_t$ is time-varying, unknown and is just what we want to estimate. We only have noisy measurements incoming in series at discrete-time-instants as shown in (2). A natural idea is determining $\mathbf{C}$ as those which best fit the time series measurements in a sliding time-window upto the current time k, denoted hereafter as

$$K := [k', k]$$

where $k' = \max(1, k - T)$, and T is the length of the time-window.

Disregarding false and missing data issues temporally here, the T-FoT at time k can be estimated by

$$\hat{\mathbf{C}}_K = \underset{\mathbf{C}}{\operatorname{argmin}} \sum_{i=k'}^{k} \mathcal{D}_i(\mathbf{C}). \tag{5}$$

In this work, the data fitting error $\mathcal{D}_i(\mathbf{C})$ is given in the LS sense using the Mahalanobis distance, i.e.,

$$\mathcal{D}_i(\mathbf{C}) := \|\mathbf{y}_i - \hat{\mathbf{y}}_i\|^2_{\Sigma^{-1}_{\mathbf{e}_i}} = (\mathbf{y}_i - \hat{\mathbf{y}}_i)^{\mathrm{T}}\Sigma^{-1}_{\mathbf{e}_i}(\mathbf{y}_i - \hat{\mathbf{y}}_i),$$

with a shorthand notation

$$\hat{\mathbf{y}}_i = h_i\left(F(i;\mathbf{C})\right) + \bar{\mathbf{v}}_i \tag{6}$$

where $\bar{\mathbf{v}}_i$ gives the mean of the measurement noise $\mathbf{v}_i$ at discrete time i, and the fitting error is given as

$$\mathbf{e}_i := \mathbf{y}_i - \hat{\mathbf{y}}_i, \tag{7}$$

which accounts for two sources of uncertainties including the measurement noise $\mathbf{v}_i$ and the T-FoT error $\epsilon(i)$ and $\Sigma_{\mathbf{e}_i}$ denotes the covariance of $\mathbf{e}_i$; assumption or simplification will be made on the statistics of the fitting error; see (20).

We have analyzed in [36] that in a linear Gaussian system, the penalty factor can take into account the Markov dynamic model information if available such that the optimization amounts to that behind the KF. In other words, in the case of a deterministic Markov model (namely no process noise, $E(\mathbf{w}_k^{\mathrm{T}}\mathbf{w}_k) = 0$), the optimization function of the KF will reduce to that of the T-FoT-oriented approach. The advantage of the T-FoT formulation is that it does not assume a Markov model for modeling the movement of the target, nor imposes requirements on state temporal independence and on chronological sensor data. Different from existing modeling of the kinematic model error such as [58], our formulation (5) treats the model error in a flexible regularization-based manner.

Based on the T-FoT parameter estimate $\hat{\mathbf{C}}_K$ conditioned on the measurements $\mathbf{y}_{k':k}$ received during time $[k', k]$, the state at any time t can be inferred as follows

$$\hat{\mathbf{x}}_t|_{\mathbf{y}_{k':k}} = F(t; \hat{\mathbf{C}}_K)\,. \tag{8}$$

More specifically, the inference is referred to as smoothing when $t < k$, as prediction or forecasting when $t > k$ and as online filtering when $t = k$. We focus on filtering in this work, i.e., estimate of the target position will be updated whenever new measurements are available.

2.3. Online Estimating $\mathbf{C}$

For easy computation, linear parameter dependence is assumed in each dimension and smooth piecewise polynomial fitting function is then used, i.e.,

$$F^{(d)}(t; \mathbf{C}) = c_0^{(d)} + c_1^{(d)} t + \cdots + c_\gamma^{(d)} t^\gamma, \tag{9}$$

where γ is referred to as the order of the fitting function and d indicates the dimension in the position space.

Parameters $c_0^{(d)}, c_1^{(d)}$ and $c_2^{(d)}$ indicate the initial position (at time $t = 0$), velocity and acceleration of the target in dimension d, respectively. As a rule of thumb, $\gamma = 1$ and $\gamma = 2$ are suitable to model the (near) constant velocity (CV) and constant acceleration (CA) [34], respectively. In the case that the target has a definite originating position (denoted by $\mathrm{p}_0^{(d)}$ in dimension d), constant velocity ($\mathrm{v}^{(d)}$ in dimension d), or constant acceleration ($\mathrm{a}^{(d)}$ in dimension d), then the respective constraint $c_0^{(d)} = \mathrm{p}_0^{(d)}$, $c_1^{(d)} = \mathrm{v}^{(d)}$, or $c_2^{(d)} = \mathrm{a}^{(d)}$ needs to be satisfied in (5), respectively. Therefore, the parameters of the T-FoT approach are theoretically explicable, rendering the approach analytical.

For the above linear systems, $\mathbf{C}$ can be calculated analytically and is unique; we will address this in the LS error sense in Section 4.1. For a nonlinear system that corresponds to a non-convex optimization problem, one has to resort to iterative/numerical approximation methods on the basis of an initial guess of the parameters for iterative searching. To speed up the optimization searching, it is important to start from the parameters $\hat{\mathbf{C}}_{[k'-1,k-1]}$ yielded at time $k-1$ for approaching the optimal $\hat{\mathbf{C}}_K$ at time k. This is reasonable since the two corresponding trajectories are actually overlapped in the piecewise $[k', k-1]$. This amounts to assuming a recursion on the latent T-FoT parameters which is critical for ensuring online implementation for non-convex optimization. Recently, Pacholska *et al.* [59] relaxed the problem and addressed the sufficient conditions for the guaranteed fitting optimality/uniqueness, where the measurement considered in particular is the distance between the sensor and the target. Even more recently, the polynomial fitting has been used for fitting measurement in tracking [60]. A sensor selection approach based on the T-FoT approach has been proposed [61] for online activating a finite number of sensors in a sensor network with communication bandwidth constraints.

As addressed so far, existing T-FoT-based approaches have not take into account false and missing data at unknown rates/ratios, or the birth and death of the target, which are essentially lying in the core of realistic target tracking.

3. Model and Scenario Assumptions

There is no more than one target in the scenario. If it exists and is detected, the position measurement $\mathbf{y}_k := \begin{bmatrix} y_k^{(1)} \\ y_k^{(2)} \end{bmatrix}$ is given by

$$\mathbf{y}_k = \begin{bmatrix} p_{x,k} \\ p_{y,k} \end{bmatrix} + \mathbf{v}_k, \tag{10}$$

where $[p_{x,k}, p_{y,k}]^{\mathrm{T}}$ is the position of the target, and $\mathbf{v}_k$ is the measurement noise.

We also consider the position-relevant measurements that can be converted to positions at each scan, i.e., the measurement function is injective [62] or multiple sensors being used [63, 64]. For example, the range-bearing measurement model can be written as follows

$$\begin{bmatrix} r_k \\ \theta_k \end{bmatrix} = \begin{bmatrix} \sqrt{p_{x,k}^2 + p_{y,k}^2} \\ \tan^{-1}\left(\frac{p_{y,k}}{p_{x,k}}\right) \end{bmatrix} + \mathbf{v}_k, \tag{11}$$

which can be converted to the position measurements by

$$\mathbf{y}_k = \begin{bmatrix} r_k \cos(\theta_k) \\ r_k \sin(\theta_k) \end{bmatrix}. \tag{12}$$

The above conversion has a bias in the mean of the converted measurement, which can be removed by multiplicating a factor either $\exp(-\frac{\Sigma_\theta}{2})$ or $\exp(\frac{\Sigma_\theta}{2})$ [65]. Further on, the converted covariance $\Sigma_{\mathbf{y}_k}$ can also be calculated numerically via Monte Carlo sampling [62] for any type of noises. Discussion on the information gain/loss due to measurement conversion can be found in [66] and the reference therein.

The measurements at time k are denoted by a set $\mathbf{Y}_k$ which are composed of both real measurement of the target (if detected) and clutter at time k. The following assumptions are required in our work.

A1. The measurement noise $\mathbf{v}_k$ which might not be Gaussian has a zero-mean and variance $\Sigma_{\mathbf{v}_k}$ given a-priori.

A2. Clutter is independently generated at different times, (near-)uniformly distributed over the surveillance region and independent to the measurement of the real target;

A3. Clutter rate r_c is relatively low as required by an algebraic manipulation analysis given in Appendix Appendix A. In our simulations, $r_c < 5$.

A4. The target detection probability p^{D} at any time is fairly high, e.g., larger than threshold $p^{\mathrm{D}} = 0.9$.

Further on, as the particular scenario in which the T-FoT approach performs best, the trajectory of the target is "smooth" having a minor β in (4), i.e., the target moves with relevantly insignificant acceleration and process noise as compared with its velocity. This together with the above assumptions are representative and match the real case of many realistic problems particularly for space target tracking (e.g., air traffic/aviation supervision).

4. Proposal: T-FoT Initiation, Maintenance and Termination

This section addresses the details of our approach for T-FoT initiation, maintenance and termination, respectively.

4.1. Trajectory Initiation

The lack of a-priori information about the target birth tangled with false and missing data lead to a significant challenge for initializing the tracker. In this work, we propose to cluster the time-series measurements of the sliding time-window K for detecting the birth of the target. This requires that the real measurements of the target in successive scans are significantly closer to each other on average than those between different clutter points and between the target and clutters. This is guaranteed by the above assumptions A1−A3. To this end, the density-based clustering method called DBSCAN [67] is readily available that distinguishes regions of high data density from low-density regions and does not need to be specified with the number of clusters in advance. For this purpose, two key parameters are required as to be addressed next: 1) the neighborhood radius ε: the maximum distance between two neighbor measurements in the cluster, and 2) the minimum cluster size T_{s}.

In addition, the following *point-target constraint* needs to be followed to guide the clustering: data-points in the cluster corresponds to different time-instants. This constraint can be formulated as a cannot link rule [68]: the measurements generated at the same time-instant cannot not belong to the same cluster. However, we note that, it is possible yet rare that two clusters are formed at the same time, at least one of which must be false alarm. In this case, we increase both the minimum cluster size T_{s} and the length of the time-window gradually and re-do the clustering till there is only one cluster to meet the single-target assumption.

4.1.1. Neighborhood radius ε

The distribution of real measurements of the target depends on two factors: 1) the measurement noise; see the analysis given in Appendix Appendix B, and 2) the velocity of the target [69, 45] to account for the movement of the target in one sensing period. In the extreme case that both the initial speed of the target and the statistics of the measurement noise are unknown, one have to learn ε from the data. The problem of whether two measurements are generated by the same target can be modelled as a composite binary-hypothesis testing problem as follows

$$\begin{cases} \mathcal{H}_0 : \text{at least one measurement is clutter} \\ \mathcal{H}_1 : \text{both measurements are from the target} \end{cases}$$

The hypothesis testing is carried out by comparing the Mahalanobis distances between measurements with the neighborhood radius ε, i.e.,

$$\|\mathbf{y}_i - \mathbf{y}_j\|^2_{\Sigma^{-1}_{\mathbf{v}_i}} \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\gtrless}} \varepsilon := \tau_1^2 . \tag{13}$$

Mahalanobis distance between two independently identically distributed (IID) variables is analyzed in Appendix Appendix B. However, the target measurements at two time-instants are not IID due to the varying of the target state. To compensate for the velocity of the target, we use a threshold τ_1 in (13) larger than necessarily required when the target is stationary as analyzed in Appendix B, e.g., $\tau_1 = 3$ in our simulation. Obviously, this works only when the initial velocity of the target is small.

4.1.2. Minimum cluster size T_s

The proposed track initiation follows a "m out of n" logic [70, 71, 72, 73] but our approach is free of the knowledge of the process noise statistics, i.e., a cluster/track is confirmed if and only if at least T_{s} detections

are clustered over K scans. We now address the choice of $T_{\rm s}$. First of all, it must be larger than the number of parameters of the fitting function to avoid underfit and is no larger than the average number of detections in the time-window, that is,

$$T_{\rm s} \in [\gamma + 1, Tp^{\rm D}]. \tag{14}$$

Note that as long as the *point-target constraint* holds, target detection can be earliest made $t \geq T_{\rm s}$ filtering steps after the target appears. So, $T_{\rm s}$ indicates how much latency/delay we have to endure at least for identifying the target; analytical results on the statistics of the detection delay have been given in [70, 71, 72, 73]. Assume that the probability for all clutter points generated at a particular sampling scan falling further than the threshold (corresponding to the radius of the cluster) to any given point in the surveillance area is p_r, c.f. (A.2); see the analysis given in Appendix A. Then, the probability for causing a false alarm (FA) due to the clustering of at least T_s clutter points (obtained at different time-instants) can be approximately estimated by

$$\begin{aligned} p_{\rm FA}(T_s) &\approx \sum_{t=T_s}^{T} C_T^t (1-p_r)^t p_r^{T-t} \\ &= C_T^{T_s}(1-p_r)^{T_s} p_r^{T-T_s} + p_{\rm FA}(T_s+1), \end{aligned} \tag{15}$$

where C_m^n stands for the number of combinations of n elements taken from a set of size m.

As shown, $p_{\rm FA}(T_s)$ decreases with the increase of T_s, that is, larger T_s implies smaller FA rate but also greater target-detection delay. A tradeoff is required here for which $3, 4$ are our recommendation.

4.1.3. T-FoT initiation Based on Weighted LS fitting

Denote the measurements in the confirmed cluster as $\{\mathbf{y}_i\}_{i\in\tilde{K}} \subseteq \{\mathbf{Y}_i\}_{i=k'}^{k}$ where $\tilde{K} \subseteq \{k', k'+1, \ldots, k\}$. According to (9), we assume the T-FoT in each position dimension as follows

$$\begin{bmatrix} y_i^{(1)} \\ y_i^{(2)} \end{bmatrix} = \begin{bmatrix} c_0^{(1)}, c_1^{(1)}, \cdots, c_\gamma^{(1)} \\ c_0^{(2)}, c_1^{(2)}, \cdots, c_\gamma^{(2)} \end{bmatrix} \begin{bmatrix} 1 \\ i \\ \cdots \\ i^\gamma \end{bmatrix} + \begin{bmatrix} e_i^{(1)} \\ e_i^{(2)} \end{bmatrix}, \tag{16}$$

where $e_i^{(d)}$ is the fitting error at time $i \in \tilde{K}$ in dimension $d = 1, 2$; see (7) and (20).

The T-FoT parameters $\mathbf{C}_{\tilde{K}}^{(d)} := \left[c_0^{(d)}, c_1^{(d)}, \cdots, c_\gamma^{(d)}\right]^{\rm T}, d = 1, 2$ can be estimated in the LS sense, i.e.,

$$\hat{\mathbf{C}}_{\tilde{K}}^{(d)} = \underset{c_0^{(d)}, c_1^{(d)}, \cdots, c_\gamma^{(d)}}{\operatorname{argmin}} \sum_{i\in\tilde{K}} \|y_i^{(d)} - \hat{y}_i^{(d)}\|^2_{\Sigma^{-1}_{e_i^{(d)}}}. \tag{17}$$

Given that $\mathbf{A}^T\Sigma_{\tilde{K}}^{-1}\mathbf{A}$ is non-singular (which is easy to be satisfied when $|\tilde{K}| > \gamma$ and when $\tilde{K}$ has no identical elements), the exact solution to (17) is given by

$$\hat{\mathbf{C}}_{\tilde{K}}^{(d)} = (\mathbf{A}^{\rm T}\Sigma_{\tilde{K}}^{-1}\mathbf{A})^{-1}\mathbf{A}^{\rm T}\Sigma_{\tilde{K}}^{-1}\mathbf{Y}_{\tilde{K}}^{(d)}, \tag{18}$$

where $\Sigma_{\tilde{K}} = {\rm Cov}(\mathbf{e}_{\tilde{K}}^{(d)})$, $\mathbf{e}_{\tilde{K}}^{(d)} = [e_{k'}^{(d)}, e_{k'+1}^{(d)}, \cdots, e_k^{(d)}]^{\rm T}$ and

$$\mathbf{A} = \begin{bmatrix} 1 & k' & \cdots & (k')^\gamma \\ 1 & k'+1 & \cdots & (k'+1)^\gamma \\ & \vdots & & \\ 1 & k & \cdots & (k)^\gamma \end{bmatrix}, \mathbf{Y}_{\tilde{K}}^{(d)} = \begin{bmatrix} y_{k'}^{(d)} \\ y_{k'+1}^{(d)} \\ \vdots \\ y_k^{(d)} \end{bmatrix}.$$

Assume fitting errors having zero mean, namely $\mathbf{E}(e_i^{(d)}) = 0, \forall i \in \tilde{K}$. Then, the Gauss-Markov theorem [74, pp.34] [75, pp.86] indicates that the LS estimate $\hat{\mathbf{C}}_{\tilde{K}}^{(d)}$ as given in (18) is the best linear unbiased estimate (BLUE) of $\mathbf{C}_{\tilde{K}}^{(d)}$, i.e., $\mathbf{E}(\hat{\mathbf{C}}_{\tilde{K}}^{(d)}) = \mathbf{C}_{\tilde{K}}^{(d)}$, and

$$\mathbf{Cov}(\hat{\mathbf{C}}_{\tilde{K}}^{(d)}) = (\mathbf{A}^{\mathrm{T}}\Sigma_{\tilde{K}}^{-1}\mathbf{A})^{-1}. \tag{19}$$

See the proof for Theorem 2.1 of [74, pp.34] and for Theorem 1 of [75, pp.86]. The BLUE is also referred to as minimum variance unbiased estimator in [75].

The key difference between our proposed T-FoT estimator and the classic Markov-Bayes estimator can be illustrated in Fig 1. Both approaches basically fit the real target trajectory with a prescribed model such that the output of the model best fits a series of measurements. To this end, the T-FoT-based LS fitting approach fits the ground truth by a continuous-time curve function while the KF assumes a Markov-jump model. They are based on the Markov- Bayes theorem and the Gauss-Markov theorem, respectively. Arguably speaking, they suit non-cooperative and cooperative target, respectively. More discussion on the relationship of the LS approach and the KF-type estimator are available in [76, 77, 78, 36].

4.1.4. Approximate Fitting Error $\mathbf{e}_i$

Obviously, $\mathbf{e}_i$ plays a key role in either (5) or (18). As addressed, it accounts for two sources of uncertainties including measurement noise $\mathbf{v}_i$ and T-FoT error $\epsilon(i)$ which, however, is generally unknown. Since the cost function uniformly accounts for the fitting errors at different time instants in the concerning time-window K, a reasonable assumption is that the fitting errors $\epsilon(i)$ is insensitive to time i. Then,

$$\Sigma_{\mathbf{e}_i}^{(d)} \propto \Sigma_{\mathbf{v}_i}^{(d)}. \tag{20}$$

This substitution brings much convergence for calculation in either (5) or (18). If further $\Sigma_{\mathbf{v}_i}$ is time-invariant, the weighted LS reduces to the ordinary LS.

4.2. Trajectory Maintenance

The main idea of our trajectory maintenance algorithm is to update (18) with time k increases. The key challenge is still from the misdetection and clutter. To enable the sliding time-window fitting, one has to identify the real measurements of the target (and their corresponding time-instants) from the clutter. To be more specific, the proposed trajectory maintenance scheme comprises the following three steps.

4.2.1. Measurement one-step prediction

Given the identified real measurements $\{\mathbf{y}_i\}_{i\in\tilde{K}} \subseteq \{\mathbf{Y}_i\}_{i=k'}^{k}$ where $\tilde{K} = \{\tilde{k}', \cdots, \tilde{k}\} \subseteq \{k', k'+1, \ldots, k\}$, the measurement of the target at time $k+1$ can be predicted following the approach of "predictive LS" [79, 80], if there are no missed detections. Alternatively and to accommodate potentially missed detection at any time instant, following (8), the estimate of the state of the target at time $k+1$ can be inferred by using the T-FoT parameter $\hat{\mathbf{C}}_{\tilde{K}}$ obtained at time k as follows

$$\hat{\mathbf{x}}_{k+1|k} = F(k+1; \hat{\mathbf{C}}_{\tilde{K}}). \tag{21}$$

Then, a pseudo measurement is calculated by using the measurement function and the predicted state (21) by

$$\hat{\mathbf{y}}_{k+1} = h_{k+1}\left(\hat{\mathbf{x}}_{k+1|k}\right) + \bar{\mathbf{v}}_{k+1}. \tag{22}$$

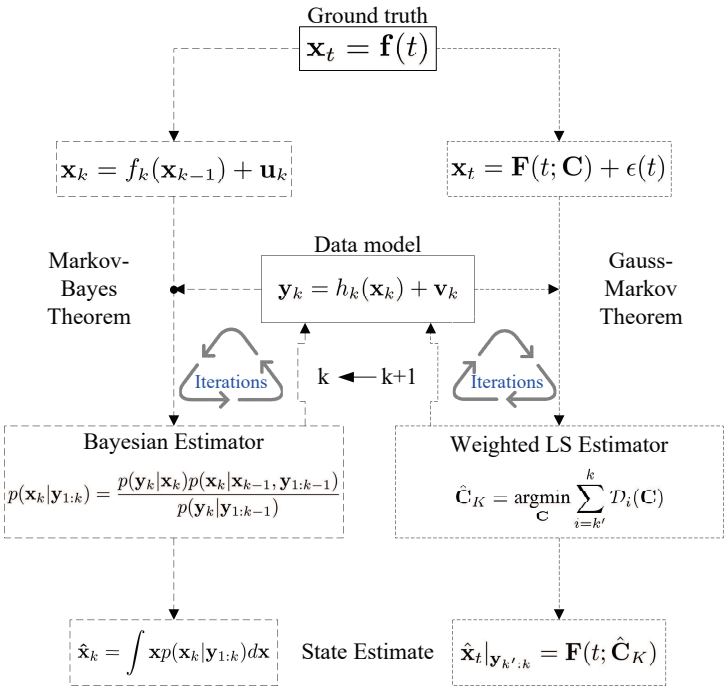

Figure 1: Comparison between the Markov-Bayes estimator based on the SSM and the weighted LS estimator based on the T-FoT.

4.2.2. Distinguishing real measurement from clutter

Denote the measurement set with I elements received at time $k+1$ as $\mathbf{Y}_{k+1} = \{\mathbf{y}_{k+1}^1, \mathbf{y}_{k+1}^2, ..., \mathbf{y}_{k+1}^I\}$. A potentially real measurement will be identified based on their (Mahalanobis) distance to the pseudo measurement as follows:

$$\tilde{\mathbf{y}}_{k+1} = \underset{\mathbf{y}_{k+1}^i \in \mathbf{Y}_{k+1}}{\operatorname{argmin}} \|\mathbf{y}_{k+1}^i - \hat{\mathbf{y}}_{k+1}\|_{\Sigma_{\mathbf{e}_k}^{-1}} \tag{23}$$

which yields the same result as the maximal likelihood criterion when the fitting error belongs to the exponential family [81].

To identify whether $\tilde{\mathbf{y}}_{k+1}$ is a real measurement of the target at time $k+1$ or actually there is no real measurement at all (namely the target is missed in detection), the Mahalanobis distance $\|\tilde{\mathbf{y}}_{k+1} - \hat{\mathbf{y}}_{k+1}\|^2_{\Sigma_{\mathbf{v}_{k+1}}^{-1}}$ can be used [82]. That is, if it is not larger than threshold τ_2^2, $\tilde{\mathbf{y}}_{k+1}$ will be identified as the real measurement otherwise, misdetection is identified. That is

$$\begin{cases} \mathbf{y}_{k+1} := \tilde{\mathbf{y}}_{k+1}, & \text{if } \|\tilde{\mathbf{y}}_{k+1} - \hat{\mathbf{y}}_{k+1}\|^2_{\Sigma_{\mathbf{v}_{k+1}}^{-1}} \le \tau_2^2 \\ \{\mathbf{y}_{k+1}\} := \emptyset, & \text{otherwise.} \end{cases} \tag{24}$$

However, when the target trajectory is not so smooth (corresponding to a large β in (4)), the one-step prediction based on (21) may be inaccurate. For example, when the target makes a sudden manoeuvre and a clutter is generated around by coincidence, the clutter which turns out to more likely from the target than the real measurement can easily be confused with the real measurement. To address this dilemma which challenges the traditional filter too, a potential solution is to carry out the density-based clustering scheme on the time-series measurements $\{\mathbf{Y}_i\}_{i=k-T_s+1}^k$ over a sliding time window from $k-T_s+1$ to the current time k for identifying the potentially real measurements in the time window. The clustered measurements can then be used for fitting a 2-D spatial trajectory in the format of $y = f'(x)$. Then, the measurement in the

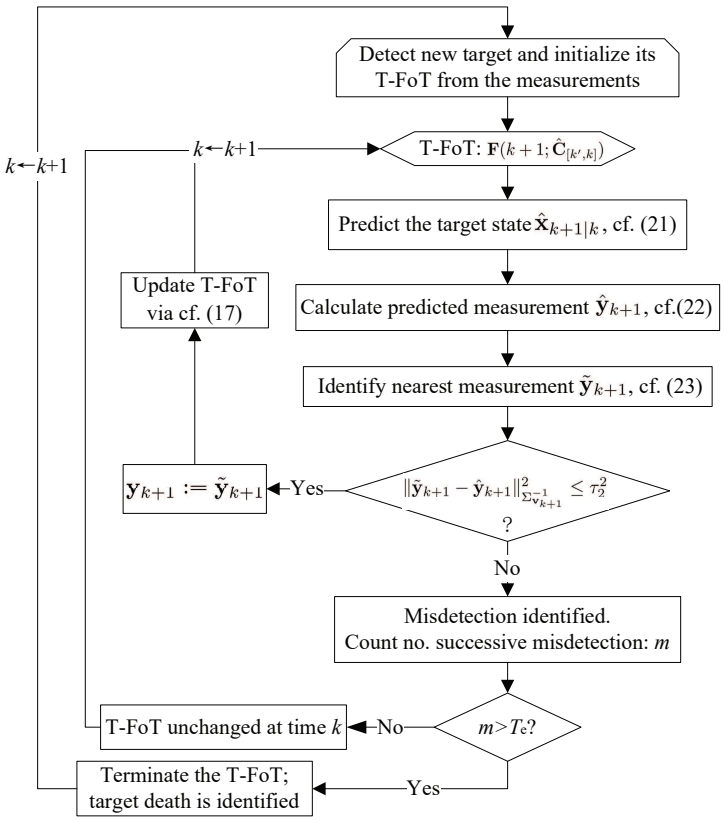

Figure 2: Procedure of the proposed JDT approach

random finite set $\mathbf{Y}_k$ that matches best $y = f'(x)$ will be identified as the potentially real measurement of the target at time k, instead of using (21)-(24). However, the price that has to be paid for this is a higher computation requirement.

4.2.3. T-FoT updating

Denote by set $K' \subseteq [k', k]$ all the time instants having real measurements, where $k' = \max(1, k - T)$, T is the length of the time-window. Based on these measurements $\mathbf{y}_i, i \in K'$, the parameters of T-FoT is now ready to be updated as shown in (18) for $d = 1, 2$, respectively.

4.3. Trajectory Termination and Potential Re-start

If the number of successive misdetections, m, is larger than T_{e}, target death will be declared and the trajectory should be terminated. The target may re-appear in the area and so the algorithm needs to re-do target detection (and then tracking) after terminating one T-FoT. This is also helpful when the algorithm terminates the trajectory wrongly due to more than T_{e} successive misdetections. Since our proposed approach does not assume the target birth model, the target re-birth model is not needed. If the target disappears shortly, it would be helpful to initialize the clustering procedure around the area where the target disappears. This can speed up the clustering calculation and lead to better accuracy. However, to ensure the best generality, we do not make such an assumption. We do not consider simultaneously detecting other targets while tracking one since it is assumed that there is no more than one target in the scenario. We leave the more challenging multiple target tracking issues to the future study.

Table 1: Parameters for Our Proposed Approach

Parameters	Purpose/meaning	Value Used
T_{s}	least latency for target detection	4
T_{e}	least latency for confirming target death	4
τ_1	used in clustering for target detection	3
τ_2	used for identifying missing detection	5
γ	order of the fitting function	1
T	length of the time window for fitting	10

4.4. Overview of All Parameters Needed

The whole procedure of the proposed JDT approach is given in Fig 2. In addition to the measurement function $h_k(\cdot)$ and the measurement noise statistics $\Sigma_{\mathbf{v}_i}^{(d)}$ which are needed in our algorithm, we summarize all the parameters needed in our approach in Table 1 as well as the favorable value for each parameter used in our simulation.

- Both τ_1 and τ_2 are Mahalanobis distance-based and reasonably related to the uncertainty of the measurements, namely the magnitude of the measurement noise.
- γ depends on the "smoothness" of the trajectory [34] for which $\gamma = 1, 2$ suit CV and CA targets, respectively.
- $T, T_{\mathrm{s}}, T_{\mathrm{e}}$ are determined more or less in a heuristic manner. As a rule of thumb, we specify $T_{\mathrm{s}}, T_{\mathrm{e}}$ in the scope of $[3, 5]$ and T in $[8, 12]$.

We note that in almost all trackers, confidence-based thresholds (relative to, e.g., target existing probability in the BF [47, 48]; see also [83, Ch.2] and [84, 85, 86]) are needed, whether explicitly or implicitly, to initialize and terminate the target track. This indicates the heuristic nature of T_{s} and T_{e}, as well as the time-window length $T = 10$, which depend on the practitioner. As a rule of thumb, $T = 10$ sensing steps turn out to be a close-to-best choice for the length of the sliding time-window in most of the scenarios we have tested.

5. Simulations

This section evaluates the performance of our approach and compare it with the Bayes-optimal BF [47, 48] in two different scenarios. One uses a CV target model and a position measurement model while the other uses a coordinated turn (CT) target model and a range-bearing measurement model. They are referred to as linear and nonlinear systems, respectively. In each system, different parameters are used for generating different ground truths and scenarios. The simulation for each scenario is performed for 1000 Monte Carlo runs, each run using different target-trajectories and measurement series, all randomly generated based on the specified statistical models and parameters. As usual, we still use the SSM for describing the kinematic model of the target. This is necessary for running a filter but not for our approach. While our approach uses no a-priori information about the target, the comparison BFs are provided with perfect a-priori information about the birth, death, and dynamics of the target, for their best possible performance. That is, the target is non-cooperative to our approach but cooperative to the BFs. This gives an edge to the latter.

The performance of the filter and the proposed T-FoT approach is evaluated by the optimal subpattern assignment error (OSPA) of the position estimation with cut-off $c = 1000$m and order $\rho = 2$ [87]. The OSPA between two RFSs $\hat{\mathbf{X}}$ and $\mathbf{X}$ is defined as follows, for $|\hat{\mathbf{X}}| \geq |\mathbf{X}|$,

$$d_{\text{ospa}}^{(c,p)}(\hat{\mathbf{X}}, \mathbf{X}) = \left(\frac{1}{|\hat{\mathbf{X}}|} \left(d_{\text{Loc}}(\hat{\mathbf{X}}, \mathbf{X}) + d_{\text{Card}}(\hat{\mathbf{X}}, \mathbf{X}) \right) \right)^{\frac{1}{p}} \tag{25}$$

which consists of two components accounting for the localization error and cardinality error, respectively, i.e., $d_{\text{Loc}}(\hat{\mathbf{X}}, \mathbf{X}) = \min_{\pi \in \Pi_{|\hat{\mathbf{X}}|}} \sum_{i=1}^{|\mathbf{X}|} d^{(c)}(\mathbf{x}_i, \hat{\mathbf{x}}_{\pi(i)})^p$, $d_{\text{Card}}(\hat{\mathbf{X}}, \mathbf{X}) = c^p(|\hat{\mathbf{X}}| - |\mathbf{X}|)$. Here, π and Π_n are a permutation and the set of all permutations on $\{1, \ldots, n\}$, and $d^{(c)}(\mathbf{x}, \mathbf{y}) = \min(d(\mathbf{x}, \mathbf{y}), c)$ is a metric between $\mathbf{x}$ and $\mathbf{y}$ cut-off at c. If $|\hat{\mathbf{X}}| < |\mathbf{X}|$, $d_{\text{ospa}}^{(c,p)}(\hat{\mathbf{X}}, \mathbf{X}) = d_{\text{ospa}}^{(c,p)}(\mathbf{X}, \hat{\mathbf{X}})$.

5.1. Linear System

In this simulation lasting for 100s, both the dynamic model and the measurement model are linear. Denote the target state as $\mathbf{x}_k = [p_{k,x}, \dot{p}_{k,x}, p_{k,y}, \dot{p}_{k,y}]^{\mathrm{T}}$, which is composed of position $[p_{x,k}, p_{y,k}]^{\mathrm{T}}$ and velocity $[\dot{p}_{x,k}, \dot{p}_{y,k}]^{\mathrm{T}}$. The target birth Bernoulli intensity function is given as

$$\gamma_k(\mathbf{x}_k) = 0.01\mathcal{N}(\mathbf{x}_k; \mathbf{m}, \mathbf{P}), \tag{26}$$

where $\mathbf{m} = [-500\text{m}, 10\text{m/s}, -500\text{m}, 10\text{m/s}]^{\mathrm{T}}$, $\mathbf{P} = \text{diag}(100\text{m}, 10\text{m/s}, 100\text{m}, 10\text{m/s})^2$.

The target appears at time $k = 10$s with a random initial state according to the above newborn target model and disappears at time $k = 80$s. To model this, the target survival probability is set as 0.99. During time $k \in [10\text{s}, 80\text{s}]$, the state of the target evolve according to a nearly CV model, i.e., $\mathbf{x}_k = F\mathbf{x}_{k-1} + \mathbf{G}\mathbf{u}_{k-1}$, with

$$F = \mathbf{I}_2 \otimes \begin{bmatrix} 1 & \Delta \\ 0 & 1 \end{bmatrix}, \mathbf{G} = \begin{bmatrix} \mathbf{I}_2 \otimes \frac{\Delta^2}{2} \\ \mathbf{I}_2 \otimes \Delta \end{bmatrix},$$

where $\Delta = 1$s and $\mathbf{u}_{k-1}$ is zero-mean Gaussian noise with time-invariant covariance $\mathbf{Q}$ of unit m^2/s^4.

The target detection probability denoted by p^{D} is constant. Once the target is detected, it generates a position measurement as in (10) with $\mathbf{v}_k$ being a zero-mean Gaussian noise with covariance $\Sigma_{\mathbf{v}_k} = \text{diag}(100\,\text{m}^2, 100\,\text{m}^2)$. The clutter is uniformly distributed over the planar area $[-1000\text{m}, 1000\text{m}] \times [-1000\text{m}, 1000\text{m}]$ with an average of r_c points per scan. We note here that, the target may fly out of the mentioned area in some runs. The clutter density is $r_c/2000^2\text{m}^{-2}$ in the cluttered area $[-1000\text{m}, 1000\text{m}] \times [-1000\text{m}, 1000\text{m}]$ and zero outside. However, the BF assumes a constant clutter intensity at $r_c/2000^2\text{m}^{-2}$. In our simulation, we test different values for p^{D}, r_c, and $\mathbf{Q}$.

Our approach uses the first order polynomial T-FoT in $x-$ and $y-$ dimensions, respectively, as follows

$$p_{x,t} = c_0^{(1)} + c_1^{(1)} t + \epsilon^{(1)}(t), \tag{27}$$

$$p_{y,t} = c_0^{(2)} + c_1^{(2)} t + \epsilon^{(2)}(t), \tag{28}$$

where the parameters $\mathbf{C}_{K'} := \{c_0^{(1)}, c_1^{(1)}, c_0^{(2)}, c_1^{(2)}\}$ are calculated by carrying out the LS optimization as in (5) over the time window $K' \subseteq [k', k]$, for $d = 1, 2$, respectively. Here, $k' = \max(1, k - T)$, $T = 10$s is the length of the time-window, and K' contains all the time-instants when there is a target detection in $[k', k]$.

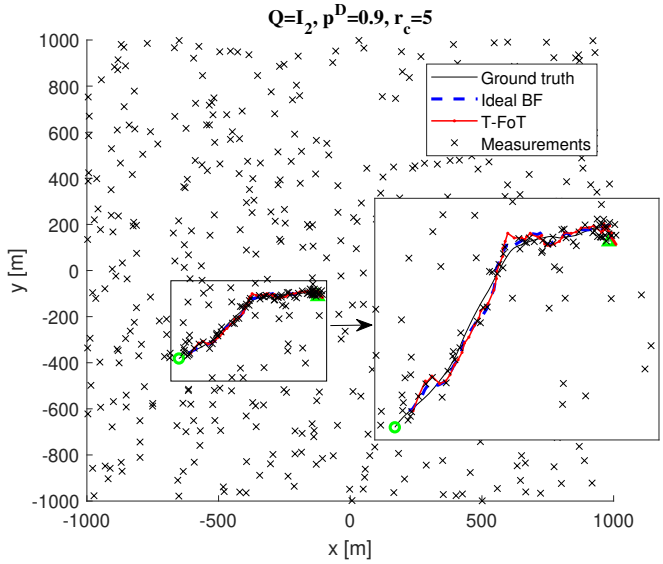

Figure 3: Target trajectory and measurements generated over 100 seconds in the linear system using process noise covariance $\mathbf{Q} = \mathbf{I}_2$, target detection probability $p^{\mathrm{D}} = 0.9$ and clutter rate $r_c = 5$ in one run. The green circle and triangle indicate the start and end of the trajectory, respectively.

Applying (20) with the time-invariant statistics of the measurement noise $\mathbf{v}_k$ leads to the time-invariant statistics of the fitting error $\mathbf{e}_t$. So, $\Sigma_{\mathbf{e}_t}$ can be removed from (5), i.e.,

$$\{c_0^{(d)}, c_1^{(d)}\} = \underset{\{c_0,c_1\}}{\operatorname{argmin}} \sum_{t\in K'} \left(y_t^{(d)} - (c_1 t + c_0)\right)^2, \tag{29}$$

Our approach does not use any mentioned target birth, death and dynamics, or clutter rate information but only need to use $\Sigma_{\mathbf{v}_k}$ in the initial clustering operation as shown in (13) for detecting the target. The values of all parameters needed are given in Table 1. The estimated position of the target at time k is simply given by substituting t by k in (27) and (28).

As the comparison approach, a Gaussian mixture (GM) BF [47, 48] is designed using all the mentioned necessary statistics information about the target, clutter, and the sensor. Furthermore, the BF used at most 50 Gaussian components (GCs), pruned GCs with weights below 10^{-5} and merged GCs with Mahalanobis distance below 4.

The average results of 1000 Monte Carlo runs are shown in Fig. 4 and Fig. 6 for process noise covariance $\mathbf{Q} = \mathbf{I}_2$ and $\mathbf{Q} = 4 \times \mathbf{I}_2$, respectively. In each figure, the subfigures show the results for different target detection probabilities $p^{\mathrm{D}} = 0.9, 0.95$ and different clutter rates $r_c = 2, 5$. The tracking scenario of one run using $p^{\mathrm{D}} = 0.95$ and $r_c = 5$ is illustrated in Fig 3 for $\mathbf{Q} = \mathbf{I}_2$ and in Fig 5 for $\mathbf{Q} = 4 \times \mathbf{I}_2$. The computing time of each updating step for all scenarios is given in Table 2. These results show that our proposed T-FoT approach computes much faster, yields higher track accuracy (when the target exists) but is less accurate in identifying the birth and death of the target as compared with the optimal BF. In particular, we notice that

1. The T-FoT approach suffers from target detection and track termination latency more than the BF as its OSPA reduces down to a stable low level slower than the BF does. This is mainly due to the lack of information about target birth and death. To combat this, one may reduce τ_1 for faster track initiation and reduce T_e for faster track termination. These, however, come with the price of causing false alarms and with the risk of premature killing the track. How to better trade-off both, which depends on the realistic need, remains an open problem.

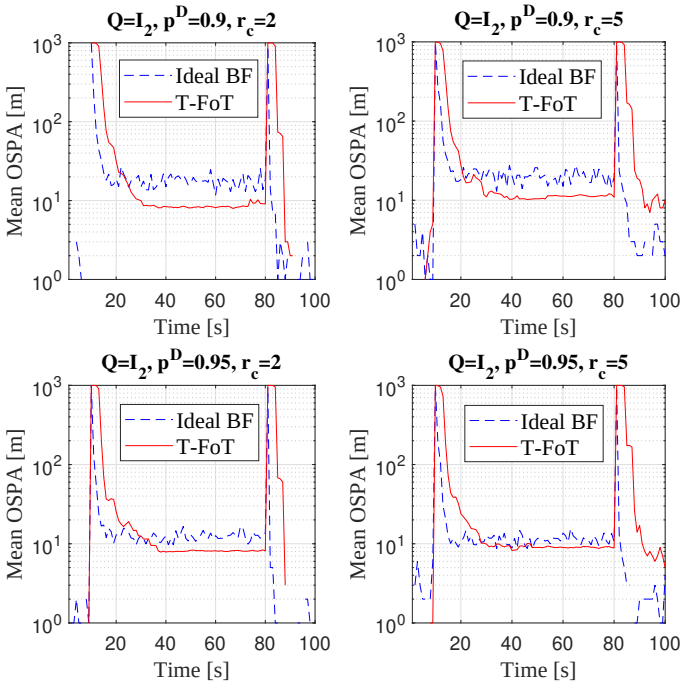

Figure 4: OSPA of the BF and our T-FoT approach in the linear system using $\mathbf{Q} = \mathbf{I}_2$, and different target detection probabilities p^{D} and clutter rates r_c.

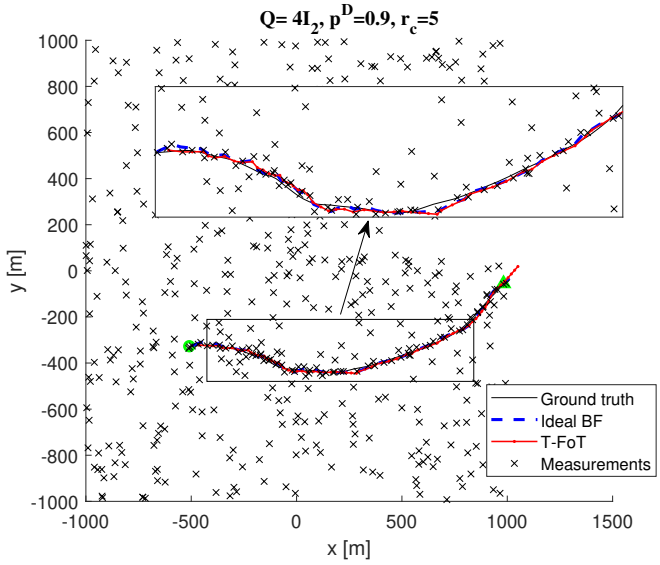

Figure 5: Target trajectory and measurements generated over 100 seconds in the linear system using process noise covariance $\mathbf{Q} = 4 \times \mathbf{I}_2$, target detection probability $p^{\mathrm{D}} = 0.9$ and clutter rate $r_c = 5$ in one run. The green circle and triangle indicate the start and end of the trajectory, respectively.

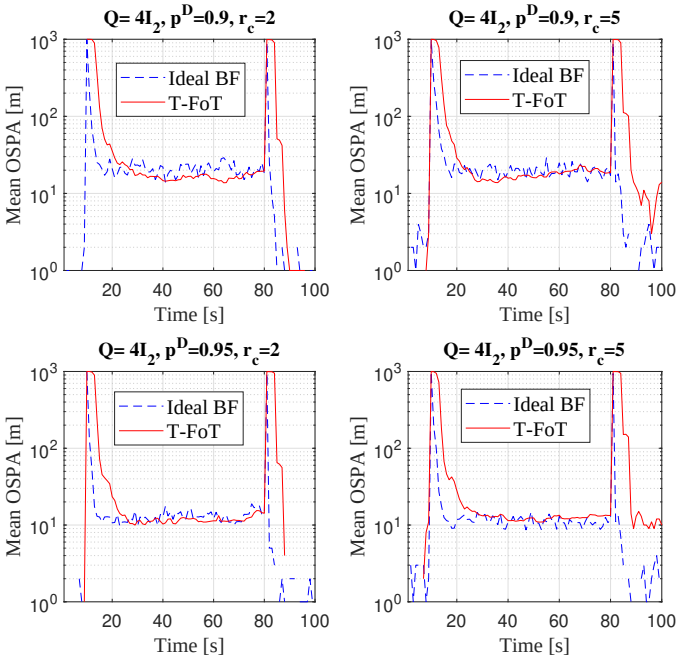

Figure 6: OSPA of the BF and our T-FoT approach in the linear system using $\mathbf{Q} = 4 \times \mathbf{I}_2$, and different detection probabilities p^{D} and clutter rates r_c.

2. For all cases except for $\mathbf{Q} = 4 \times \mathbf{I}_2, p^{\mathrm{D}} = 0.95, r_c = 5$, our proposed T-FoT approach outperforms the BF in mean OSPA during time $k \in [30\mathrm{s}, 80\mathrm{s}]$. This may be attributed to two reasons: First, the GM implementation of the BF that unavoidably bears approximation errors prevents the BF from achieving the theoretically best performance. To say the least, its Bayes-optimization does not equal minimum OSPA as the latter is no more than a specific yet common point-based metric. Second, our proposed approach is based on sliding time-window fitting which is not as sensitive to clutter or misdetection at any particular time-instant as the filter. For the scene using $\mathbf{Q} = 4 \times \mathbf{I}_2, p^{\mathrm{D}} = 0.95, r_c = 5$, both perform similar to each other. But in all cases, the T-FoT approach has an average OSPA varying over time in a smaller scope when the target exists and is tracked during $k \in [30\mathrm{s}, 80\mathrm{s}]$.
3. It is obvious that the performance of the proposed T-FoT approach is much better when $\mathbf{Q} = \mathbf{I}_2$ (which corresponds to a smoother trajectory) than when $\mathbf{Q} = 4 \times \mathbf{I}_2$. In contrast, the performance of the BFs does not change much with the change of $\mathbf{Q}$ or r_c, but it becomes much better if the target detection probability is increased from $p^{\mathrm{D}} = 0.9$ to $p^{\mathrm{D}} = 0.95$.
4. A higher clutter rate renders our approach prone to declaring appearance of the target (e.g., at time $k \in [85\mathrm{s}, 100\mathrm{s}]$). This is reasonable as it used no a-priori information about the target birth and so the algorithm can easily be misled by the locally clustered clutter in successive scans when the clutter rate is high. In contrast, the BF has a proper target birth model to prevent it from wrongly initializing a new track that does not match closely to the specified target-birth model.

5.2. Nonlinear System

In this simulation of a length of 150s, we consider an CT target motion model and a range-bearing measurement model. The target state is denoted as $\mathbf{x}_k = [p_{x,k}\ \dot{p}_{x,k}\ p_{y,k}\ \dot{p}_{x,k}\ \omega_k]^{\mathrm{T}}$ with time-varying turn rate ω_k. The target birth intensity function is given as

$$\gamma_k(\mathbf{x}_k) = 0.01 \times \mathcal{N}(\mathbf{x}_k; \mathbf{m}_1, \mathbf{P}_1), \tag{30}$$

Table 2: Computing Time for Each Step in Simulation 1 [second]

Scenario	BF	T-FoT
$\mathbf{Q} = \mathbf{I}_2, p^{\mathrm{D}} = 0.9, r_c = 2$	0.041	0.016
$\mathbf{Q} = \mathbf{I}_2, p^{\mathrm{D}} = 0.9, r_c = 5$	0.049	0.017
$\mathbf{Q} = \mathbf{I}_2, p^{\mathrm{D}} = 0.95, r_c = 2$	0.039	0.016
$\mathbf{Q} = \mathbf{I}_2, p^{\mathrm{D}} = 0.95, r_c = 5$	0.047	0.017
$\mathbf{Q} = 4\mathbf{I}_2, p^{\mathrm{D}} = 0.9, r_c = 2$	0.043	0.017
$\mathbf{Q} = 4\mathbf{I}_2, p^{\mathrm{D}} = 0.9, r_c = 5$	0.051	0.017
$\mathbf{Q} = 4\mathbf{I}_2, p^{\mathrm{D}} = 0.95, r_c = 2$	0.041	0.016
$\mathbf{Q} = 4\mathbf{I}_2, p^{\mathrm{D}} = 0.95, r_c = 5$	0.048	0.017

where $\mathbf{m}_1 = [100\mathrm{m}, 10\mathrm{m/s}, 100\mathrm{m}, 10\mathrm{m/s}, 0.01\mathrm{rad}]^{\mathrm{T}}$ and $\mathbf{P}_1 = \mathrm{diag}(100\mathrm{m}, 10\mathrm{m/s}, 100\mathrm{m}, 10\mathrm{m/s}, 0.01\mathrm{rad})$.

One target appears at time $k = 10$s with a random initial state according to the above newborn target model and disappears at time $k = 80$s. The probability of target survival is $P_k^{\mathrm{S}} = 0.99$ for the BF. The survival single-target movement follows a CT model with a sampling period of 1s and Markov transition function

$$f_{k|k-1}(\mathbf{x}_k|\mathbf{x}_{k-1}) = \mathcal{N}(\mathbf{x}_k; F(\omega_k)\mathbf{x}_k, \mathbf{Q}), \tag{31}$$

where $\mathbf{Q} = \mathrm{diag}([\mathbf{I}_2 \otimes \mathbf{G}, \sigma_u^2])$,

$$F(\omega) = \begin{bmatrix} 1 & \frac{\sin\omega}{\omega} & 0 & -\frac{1-\cos\omega}{\omega} & 0 \\ 0 & \cos\omega & 0 & -\sin\omega & 0 \\ 0 & \frac{1-\cos\omega}{\omega} & 1 & \frac{\sin\omega}{\omega} & 0 \\ 0 & \sin\omega & 0 & \cos\omega & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \mathbf{G} = \begin{bmatrix} \frac{\sigma_w^2}{4} & \frac{\sigma_w^2}{2} \\ \frac{\sigma_w^2}{2} & \sigma_w^2 \end{bmatrix}$$

with $\sigma_w = 2\mathrm{m/s}^2$, and $\sigma_u = (\pi/180)\mathrm{rad/s}$.

Further, we consider the re-appearance of the target. It is simulated that the target re-appears at time $k = 90$s and disappears again at time $k = 110$s. (Sure, this can be viewed as another new target). The initial state $\mathbf{x}_{90}$ is given by $\mathbf{x}_{90} \sim \mathcal{N}(\mathbf{x}_k; \mathbf{m}_2, \mathbf{P}_2)$, where $\mathbf{m}_2 = [500\mathrm{m}, 10\mathrm{m/s}, 500\mathrm{m}, 10\mathrm{m/s}, 0.01\mathrm{rad}]^{\mathrm{T}}$ and $\mathbf{P}_2 = \mathrm{diag}(100\mathrm{m}, 10\mathrm{m/s}, 100\mathrm{m}, 10\mathrm{m/s}, 0.01\mathrm{rad})$.

The range-bearing measurement model is given as in (11) where $v_{r,k}$ and $v_{\theta,k}$ are, individually, independent identical distributed zero-mean Gaussian with standard deviation $\sigma_r = 10$m and $\sigma_\theta = (\pi/90)$rad, respectively.

The target detection probability is state-related as given by

$$p_k^{\mathrm{D}}(\mathbf{x}_k) = p_{\max}^{\mathrm{D}} \cdot \frac{\mathcal{N}(\mu_{\mathrm{D}}(\mathbf{x}_k); \mathbf{0}, 2000^2\mathbf{I}_2)}{\mathcal{N}(\mathbf{0}; \mathbf{0}, 2000^2\mathbf{I}_2)}. \tag{32}$$

Here, $\mu_{\mathrm{D}}(\mathbf{x}_k) \triangleq [|p_{x,k}|, |p_{y,k}|]^{\mathrm{T}}$.

The clutter measurements are uniformly distributed over a disk of radius 2000m around the origin with an average number of r_{c} clutter measurements per time step. We considered two different clutter rates $r_c = 2, 5$ and two different maximal detection probabilities $p_{\max}^{\mathrm{D}} = 0.9, 0.95$.

In our approach for T-FoT fitting, the range-bearing measurements $\mathbf{y}_k$ are converted to position measurements as in (12) and the converted noise covariances are calculated based on linearization [65]. The only a-priori statistical information used in our approach is about σ_r and σ_θ. Then, both clustering and

fitting can be performed over the converted position measurement as in the last simulation, cf. (27) and (28). However, we note that, differently from the last simulation, the converted covariance $\Sigma_{\mathbf{y}_k}$ is obviously state-related, and therefore should be explicitly taken into account in the T-FoT optimization as shown in (5). The multi-dimensional optimization, however, is computationally much more complicated (and did not yield obviously better results as we found). Therefore, we over this difficulty by omitting the cross-correlation and carrying out individual dimension fitting. Furthermore, we use $\Sigma_{\mathbf{e}_i}^{(d)} \propto \Sigma_{\mathbf{y}_i}^{(d)}$. Then, we get the following weighted LS formulation

$$\{c_0^{(d)}, c_1^{(d)}\} = \underset{\{c_0,c_1\}}{\operatorname{argmin}} \sum_{i\in K'} \Sigma_{\mathbf{y}_i^{(d)}}^{-1} \big(y_i^{(d)} - (c_1 i + c_0)\big)^2. \tag{33}$$

For comparison, the local GM-BFs are implemented based on either the extended KF (EKF), unscented KF (UKF) or the particle filter (PF). In the former two cases, the filter used at most 50 GCs, pruned GCs with weights below 10^{-5} and merged GCs with Mahalanobis distance below 4 while in the latter, 2000 particles are allocated to the filter and 1000 particles are assigned to the target born process; see [47, 48] for the detail of these algorithms. Again, these filters make full use of all available models and parameters.

The target trajectory and estimates by different filters and the proposed T-FoT approach in one run using $p^{\mathrm{D}} = 0.95$ and $r_c = 5$ are illustrated in Fig 7. The average OSPA of these comparison filters and the proposed T-FoT approach over 1000 Monte Carlo runs are given in Fig. 8. The computing time of each updating step for all scenarios is given in Table 3. These results show that in this nonlinear system:

1. The T-FoT approach suffers again from target detection and track termination latency more than the BFs, similar as in the first simulation. Also, the results after the disappearance of the second target $k > 120$s show that a higher clutter rate ($r_c = 5$) can easier lead to false alarm to the T-FoT approach (as compared with $r_c = 2$). The same reasons hold.
2. The accuracy of the proposed T-FoT approach is comparable to the UKF/PF BFs (and outperforms the EKF-BF) during time $k \in [20\text{s}, 40\text{s}]$ but later on its accuracy decreases much faster than the BFs do. The accuracy decrease is due to the fact that in most runs, the target moves away from the origin and so the converted position accuracy of the range-bearing measurement as in (11) is reducing and also the target detection probability is reducing as indicated by (32). As can be easily illustrated, the position error corresponding to bearing error "σ_θ" at range distance of 100m is almost 10 times smaller than that at a range distance of 1000m. Moreover, the converted position cross-correlation that we omitted in our fitting increases as the range increases. This differs from the last simulation where the position measurements are of constant quality disregarding the state of the target and are not cross-correlated.
3. For the second, "unexpected" target appearing at time $k = 90$s, all BFs that do not set a corresponding model for it have missed it to a large degree while the T-FoT approach can quickly detect and then track it accurately the same as the first target. Arguably, both the success (i.e., quick detection) and failure (i.e., misdetection) of the filter depend on whether the specified models match perfectly the truth.
4. The proposed T-FoT approach is computationally more efficient than all BFs. The T-FoT approach only takes slightly more computation due to the measurement conversion as compared with that in the first simulation, while the EKF/UKF BFs take almost two times computation time for each filtering

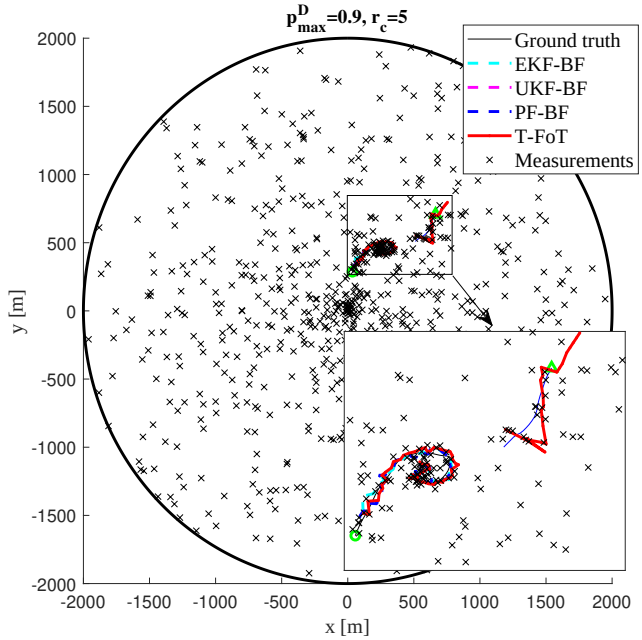

Figure 7: Target trajectory and measurements over 150 seconds generated by the nonlinear systemusing $p^D_{\max} = 0.9$ and $r_c = 5$ in one run. Green circle and triangle indicate the start and end of the trajectory, respectively.

Table 3: Computing Time for Each Step in Nonlinear System [second]

Scenario	EKF-BF	UKF-BF	PF-BF	T-FoT
$p^{\mathrm{D}} = 0.9, r_c = 2$	0.094	0.112	0.170	0.036
$p^{\mathrm{D}} = 0.9, r_c = 5$	0.152	0.169	0.230	0.035
$p^{\mathrm{D}} = 0.95, r_c = 2$	0.087	0.11	0.167	0.029
$p^{\mathrm{D}} = 0.95, r_c = 5$	0.138	0.160	0.226	0.029

step in order to deal with the nonlinearity in filtering. The computational cost of the PF-BF is even higher.

5.3. Discussion

We must reiterate that the scenarios we considered here are limited to (near-)uniformly distributed clutter with a low rate ($r_c \leq 5$) and high target detection probability ($p^{\mathrm{D}} \geq 0.9$). Relation of these limitations is valuable and deserves to be investigated. As shown in both simulations, the proposed T-FoT approach is comparable with, or even better than, the BFs in computing efficiency and track maintenance accuracy although it suffers more from track confirmation and termination latency due to the lack of a-prior information about the target birth and death. What are additionally notable include

1. The T-FoT approach computes efficiently thanks to the exact solution given for linear fitting as shown in (18).
2. The fitting calculation over data of a time-window is resilience to the false/missing data at a particular time-instant as compared with the recursive filtering.
3. The Bayes-optimization of the BF does not equal minimum OSPA. Obviously, the OSPA values depends on the choices of the parameters c and p.
4. Although the BF is Bayes optimal, it generally has no analytic solution and has to resort to approximative implementation using either GM or particles. The number of GCs or particles used and the

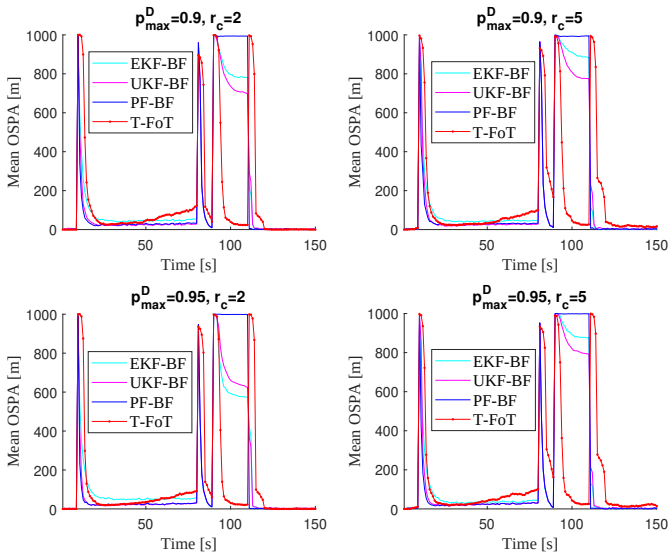

Figure 8: OSPA of BFs and our T-FoT approach in the nonlinear system using different target detection probabilities p^{D} and clutter rates r_c.

ineluctable merging/pruning/resampling operations can all have a significant effect on the filtering accuracy. These are attributed to the discrepancy of the realistic performance of the BF to the desired Bayes optimality.

6. Conclusion

We have presented a Markov-free, data-driven approach to JDT of a non-cooperative target that randomly appears and disappears in the presence of false and missing data. While the clutter rate is less than 5 per scan on average and the target detection probability is higher than 90%, their exact statistics, as well as those of the birth, death, and dynamics of the target are unknown and maybe time-varying. Our approach based on the T-FoT overcomes these challenges by only making use of measurements for joint target detection and continuous-time trajectory estimation including initiation, maintenance, termination and even re-detection. Simulation results have demonstrated that our approach performs comparable to the properly modelled Bernoulli filter provided with all required model and scenario information and even outperforms them in some cases while computing more efficiently. Therefore, the T-FoT approach provides a promising alternative to the classic state space model and accommodates favorably intelligent learning methods.

The future work can be threefold: The first is to extend the T-FoT approach to the scenario of an unknown number of targets, i.e., find multiple T-FoTs that fit best the measurement sequences over time. Such an extension is nontrivial as measurement-to-track association is involved, which is challenging whenever target tracks are interacting with or approaching each other. The second is to extend the T-FoT approach for extended target or even swarm target tracking by utilizing extension/swarm feature estimation approaches. The third is to consider a decentralized/large-scale sensor network for which an interesting issue would be continuous-time trajectory fusion [88]. All of these aim at a systematic data-driven and self-contained approach to multi-sensor multitarget detection and tracking.

Appendix A. Bound on Uniformly-distributed Clutter Rate

Disregarding the boundary issue of the surveillance region, the probability for a clutter point falling within a distance d_o to the target in the region (which corresponds to a circle around the target) is simply given by

$$p_1 = S^{-1}\pi d_o^2, \tag{A.1}$$

where S is the area of the entire surveillance region and πd_o^2 gives the area of the circle with radius d_o.

Now, consider r_c clutter points that are independently, uniformly generated over the region. The probability for all clutter points falling further than d_o to the target is given by

$$p_r = (1 - p_1)^{r_c} = (1 - S^{-1}\pi d_o^2)^{r_c}. \tag{A.2}$$

To ensure no clutter generated within a distance d_o to the target in order to avoid a FA within clustering under the confident probability p_r, the clutter rate needs to satisfy

$$r_c \leq \log_{1-S^{-1}\pi d_o^2}(p_r). \tag{A.3}$$

For instance, if $p_r = 0.95, p_1 = 0.01$ (corresponding to an area that is 1% of the whole area), then $\log_{0.99}(0.95) \approx 5.10$. That is, we have 95% confidence that uniformly-distributed clutter points should not lie in any region whose area is 1% of the whole area, when the clutter rate $r_c < 5.1$.

Appendix B. Probabilistic Distance between IID Variables

For D-dimensional Gaussian-random variables $\mathbf{a}$ and $\mathbf{b}$ which are independently identically distributed with covariance $\mathbf{R}$. The probability that they have distance within τ times the standard deviation of their distribution is given by [89]

$$\Pr[(\mathbf{a}-\mathbf{b})^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{a}-\mathbf{b}) \leq \tau^2] \leq \gamma\Big(\frac{D}{2}, \frac{\tau^2}{2}\Big), \tag{B.1}$$

where $\gamma(a, b) := \int_0^b t^{a-1}e^{-t}dt$ is the lower incomplete Gamma function.

When the variables are distribution free (i.e., unknown and probably non-Gaussian) then the probability that the variable lies within the standard deviation of the state estimate can still be bounded using, for example, the Chebyshev inequality as follows

$$\Pr[(\mathbf{c})^{\mathrm{T}}(\mathbf{c}) \geq \epsilon] \leq \frac{\Sigma_{\mathbf{c}}}{\epsilon^2}, \tag{B.2}$$

where $\mathbf{c}$ is a zero-mean random variable with variance $\Sigma_{\mathbf{c}}$ and $\epsilon > 0$.

A tighter bound is given by the Chebyshev-type inequality, Vysochanskiï-Petunin inequality [90], when it is known that the distribution is unimodal, i.e.,

$$\Pr[(\mathbf{a}-\mathbf{b})^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{a}-\mathbf{b}) \leq \tau^2] \geq 1 - \frac{4D}{9\tau^2}. \tag{B.3}$$

Acknowledgment

This work was supported in part by the National Natural Science Foundation of China under Grant 62071389, by the Key Laboratory Foundation of National Defence Technology under Grant JKWATR-210504, by the Natural Science Basic Research Program of Shaanxi and by the JWKJW Foundation . The authors would like to thank Prof. P. Djurić and Prof. F. Hlawatsch for their insightful discussion on the work.

References

[1] Y. Bar-Shalom, X.-R. Li, T. Kirubarajan, Estimation with Applications to Tracking and Navigation: Theory, Algorithms, and Software, John Wiley & Sons, 2002.

[2] R. Kalman, A new approach to linear filtering and prediction problems, J. basic Engineering 82 (1960) 35–45.

[3] D. Blacknell, H. Griffiths, Radar Automatic Target Recognition (ATR) and Non-Cooperative Target Recognition (NCTR), Institution of Engineering and Technology, London, UK, 2013.

[4] H. Wang, J. Sun, X. Zhang, X.-Y. Huang, T. Auligné, Radar data assimilation with wrf 4d-var. part i: System development and preliminary testing, Monthly Weather Review 141 (2013) 2224–2244.

[5] W. S. Rosenthal, S. Venkataramani, A. J. Mariano, J. M. Restrepo, Displacement data assimilation, J comput. Phys. 330 (2017) 594–614.

[6] A. Carrassi, M. Bocquet, L. Bertino, G. Evensen, Data assimilation in the geosciences: An overview of methods, issues, and perspectives, Wiley Interdiscip Rev Clim Change 9 (2018) 1–50.

[7] K. Judd, T. Stemler, Failures of sequential Bayesian filters and the successes of shadowing filters in tracking of nonlinear deterministic and stochastic systems, Phys. Rev. E 79 (2009) 066206.

[8] A. Zaitouny, T. Stemler, S. D. Algar, Optimal shadowing filter for a positioning and tracking methodology with limited information, Sensors 19 (2019).

[9] H. Michalska, D. Q. Mayne, Moving horizon observers and observer-based control, IEEE Trans. Autom. Control. 40 (1995) 995–1006.

[10] C. V. Rao, J. B. Rawlings, D. Q. Mayne, Constrained state estimation for nonlinear discrete-time systems: stability and moving horizon approximations, IEEE Trans. Autom. Contr. 48 (2003) 246–258.

[11] N. Morrison, Tracking Filter Engineering: The Gauss-Newton and Polynomial Filters, Electromagnetics and Radar Series, Institution of Engineering and Technology, 2012.

[12] R. Nadjiasngar, M. Inggs, Gauss-Newton filtering incorporating levenberg-marquardt methods for tracking, Dig. Signal Process. 23 (2013) 1662–1667.

[13] D. J. Mook, J. L. Junkins, Minimum model error estimation for poorly modeled dynamic systems, J Guid Control Dynam 11 (1988) 256–261.

[14] J. L. Crassidis, F. L. Markley, Predictive filtering for nonlinear systems, J Guid Control Dynam 20 (1997) 566–572.

[15] A. Delong, A. Osokin, H. N. Isack, Y. Boykov, Fast approximate energy minimization with label costs, Int. J. Comput. Vision 96 (2012) 1–27.

[16] B. Vo, B. Vo, A multi-scan labeled random finite set model for multi-object state estimation, IEEE Trans. Signal Process. 67 (2019) 4948–4963.

[17] L. Chen, From labels to tracks: It's complicated, in: I. Kadar (Ed.), Proc. SPIE 2018, volume 10646, pp. 1–6.

[18] A. F. García-Fernández, L. Svensson, M. R. Morelande, Multiple target tracking based on sets of trajectories, IEEE Transactions on Aerospace and Electronic Systems 56 (2020) 1685–1707.

[19] N. D. Singpurwalla, N. G. Polson, R. Soyer, From least squares to signal processing and particle filtering, Technometrics 60 (2018) 146–160.

[20] J. Fan, Q. Yao, Nonlinear time series: nonparametric and parametric methods, Springer, New York, USA, 2003.

[21] F. El-Hawary, Y. Jing, Robust regression-based EKF for tracking underwater targets, IEEE J Oceanic Eng. 20 (1995) 31–41.

[22] I. Dimatteo, C. R. Genovese, R. E. Kass, Bayesian curve-fitting with free-knot splines, Biometrika 88 (2001) 1055–1071.

[23] M. Hadzagic, H. Michalska, A Bayesian inference approach for batch trajectory estimation, in: Proc. FUSION 2011, Chicago, Illinois, USA, pp. 1–8.

[24] X. Wang, C. Han, Turn rate estimation based on curve fitting in maneuvering target tracking, in: Proc. 42nd Southeastern Sym. Syst. Theory, Tyler, TX, USA, pp. 194–196.

[25] Y. Liu, J. Suo, H. R. Karimi, X. Liu, A filtering algorithm for maneuvering target tracking based on smoothing spline fitting, Abstr. Appl. Anal. 2014 (2014) 1–6.

[26] P. Furgale, C. H. Tong, T. D. Barfoot, G. Sibley, Continuous-time batch trajectory estimation using temporal basis functions, Int. J. Rob. Res. 34 (2015) 1688–1710.

[27] A. Milan, K. Schindler, S. Roth, Multi-target tracking by discrete-continuous energy minimization, IEEE Trans. Pattern Anal. Mach. Intell. 38 (2016) 2054–2068.

[28] T. Li, J. Prieto, J. M. Corchado, Fitting for smoothing: A methodology for continuous-time target track estimation, in: Proc. IPIN 2016, Alcala de Henares, Madrid, Spain, pp. 1–8.

[29] Z. Wang, H. Zhou, Mathmetical processing to tracking data of range and range rate, Chinese Space Science and Technology 3 (1994) 17–24.

[30] R. Anderson-Sprecher, R. V. Lenth, Spline estimation of paths using bearings-only tracking data, J Am Stat Assoc. 91 (1996) 276–283.

[31] Z. Wang, J. Zhu, Reduced parameter model on trajectory tracking data with applications, Science in China Series E 42 (1999) 190–199.

[32] J. Liu, J. Zhu, M. Xie, Trajectory estimation with multi-range-rate system based on sparse representation and spline model optimization, Chinese Journal of Aeronautics 23 (2010) 84 – 90.

[33] D. Zheng, S. Wang, Q. Meng, Dynamic programming track-before-detect algorithm for radar target detection based on polynomial time series prediction, IET Radar, Sonar Navigation 10 (2016) 1327–1336.

[34] T. Li, H. Chen, S. Sun, J. M. Corchado, Joint smoothing and tracking based on continuous-time target trajectory function fitting, IEEE Trans. Autom. Sci. Eng. 16 (2019) 1476–1483.

[35] T. Li, Single-road-constrained positioning based on deterministic trajectory geometry, IEEE Comm. Lett. 23 (2019) 80–83.

[36] T. Li, X. Wang, Y. Liang, J. Yan, H. Fan, A track-oriented approach to target tracking with random finite set observations, in: Proc. ICCAIS 2019, Chengdu, China, pp. 1–6.

[37] M. G. Hamed, D. Gianazza, M. Serrurier, N. Durand, Statistical prediction of aircraft trajectory: regression methods vs point-mass model, in: Proc. ATM 2013, Chicago, US, pp. 1–10.

[38] K. Thormann, F. Sigges, M. Baum, Learning an object tracker with a random forest and simulated measurements, in: Proc. FUSION 2017, pp. 1–4.

[39] C. Gao, J. Yan, S. Zhou, B. Chen, H. Liu, Long short-term memory-based recurrent neural networks for nonlinear target tracking, Signal Process. 164 (2019) 67–73.

[40] J. Pinto, G. Hess, W. Ljungbergh, Y. Xia, L. Svensson, H. Wymeersch, Next generation multitarget trackers: Random finite set methods vs transformer-based deep learning, in: Proceedings of FUSION 2021, Sun City, South Africa, pp. 1–8.

[41] L. Cheng, F. Yin, S. Theodoridis, S. Chatzis, T.-H. Chang, Rethinking bayesian learning for data analysis: The art of prior and inference in sparsity-aware modeling, arXiv:2205.14283 (2022).

[42] D. Musicki, R. Evans, S. Stankovic, Integrated probabilistic data association, IEEE Trans. Autom. Contr. 39 (1994) 1237–1241.

[43] S. M. Tonissen, R. J. Evans, Peformance of dynamic programming techniques for track-before-detect, IEEE Trans. Aerosp. Electr. Syst. 32 (1996) 1440–1451.

[44] X. Jiang, K. Harishan, R. Tharmarasa, T. Kirubarajan, T. Thayaparan, Integrated track initialization and maintenance in heavy clutter using probabilistic data association, Signal Process. 94 (2014) 241 – 250.

[45] M. Mallick, Y. Bar-Shalom, T. Kirubarajan, M. Moreland, An improved single-point track initiation using GMTI measurements, IEEE Trans. Aerosp. Electr. Syst. 51 (2015) 2697–2714.

[46] R. P. S. Mahler, Statistical Multisource-Multitarget Information Fusion, Artech House, Norwood, MA, USA, 2007.

[47] B.-T. Vo, C. See, N. Ma, W. Ng, Multi-sensor joint detection and tracking with the Bernoulli filter, IEEE Trans. Aerosp. Electr. Syst. 48 (2012) 1385–1402.

[48] B. Ristic, B.-T. Vo, B.-N. Vo, A. Farina, A tutorial on Bernoulli filters: Theory, implementation and applications, IEEE Trans. Signal Process. 61 (2013) 3406–3430.

[49] J. Duník, O. Sträkä, O. Kost, J. Havlík, Noise covariance matrices in state-space models: A survey and comparison of estimation methods. part i, Internat. J. Adapt. Contr. Signal Process. 31 (2017) 1505–1543.

[50] L. Zhang, D. Sidoti, A. Bienkowski, K. Pattipati, Y. Bar-Shalom, D. Kleinman, On the Identification of Noise Covariances and Adaptive Kalman Filtering: A New Look at a 50 Year-old Problem, IEEE Access 8 (2020) 59362–59388.

[51] S. Sarkka, Bayesian Filtering and Smoothing, Cambridge University Press, New York, NY, USA, 2013.

[52] T. Li, J. Su, W. Liu, J. M. Corchado, Approximate Gaussian conjugacy: Recursive parametric filtering under nonlinearity, multimodality, uncertainty, and constraint, and beyond, Front. Inf. Technol. Electron. Eng. 18 (2017) 1913–1939.

[53] F. Tobar, P. M. Djurić, D. P. Mandic, Unsupervised state-space modeling using reproducing kernels, IEEE Trans. Signal Process. 63 (2015) 5210–5221.

[54] X. R. Li, V. P. Jilkov, Survey of maneuvering target tracking. Part I. Dynamic models, IEEE Trans. Aerosp. Electron. Syst.

39 (2003) 1333–1364.

[55] X. R. Li, V. P. Jilkov, Survey of maneuvering target tracking. part v. multiple-model methods, IEEE Trans. Aerosp. Electron. Syst. 41 (2005) 1255–1321.

[56] J. Zhou, T. Li, X. Wang, L. Zheng, Target tracking with equality/inequality constraints based on trajectory function of time, IEEE Signal Processing Letters 28 (2021) 1330–1334.

[57] S. Boyd, L. Vandenberghe, Convex Optimization, Cambridge University Press, 2014.

[58] B. Gao, G. Hu, Y. Zhong, X. Zhu, Cubature rule-based distributed optimal fusion with identification and prediction of kinematic model error for integrated uav navigation, Aerospace Science and Technology 109 (2021) 106447.

[59] M. Pacholska, F. Dümbgen, A. Scholefield, Relax and recover: Guaranteed range-only continuous localization, IEEE Robotics and Automation Letters 5 (2020) 2248–2255.

[60] Z. Tian, K. Yang, M. Danino, Y. Bar-Shalom, B. Milgrom, Launch point estimation with a single passive sensor without trajectory state estimation, IEEE Transactions on Aerospace and Electronic Systems 58 (2022) 318–327.

[61] C. Liu, K. Di, T. Li, V. Elvira, A sensor selection approach to maneuvering target tracking based on trajectory function of time, EURASIP J. Adv. Signal Process. 72 (2022).

[62] T. Li, J. M. Corchado, J. Bajo, S. Sun, J. F. D. Paz, Effectiveness of Bayesian filters: An information fusion perspective, Inf. Sci. 329 (2016) 670–689.

[63] T. Li, J. M. Corchado, S. Sun, J. Bajo, Clustering for filtering: Multi-object detection and estimation using multiple/massive sensors, Inf. Sci. 388-389 (2017) 172–190.

[64] T. Li, J. M. Corchado, H. Chen, Distributed flooding-then-clustering: A lazy networking approach for distributed multiple target tracking, in: Proc. FUSION 2018, Cambridge, UK, pp. 2415–2422.

[65] S. Bordonaro, P. Willett, Y. Bar-Shalom, Decorrelated unbiased converted measurement Kalman filter, IEEE Trans. Aerosp. Electron. Syst. 50 (2014) 1431–1444.

[66] J. Lan, X. R. Li, Nonlinear estimation based on conversion-sample optimization, Automatica 121 (2020) 109160.

[67] M. Ester, H.-P. Kriegel, J. Sander, X. Xu, A density-based algorithm for discovering clusters in large spatial databases with noise, in: Proc. KDD'96, AAAI Press, 1996, pp. 226–231.

[68] T. Li, F. D. Prieta, J. M. Corchado, J. Bajo, Multi-source homogeneous data clustering for multi-target detection from cluttered background with misdetection, Appl. Soft Comput. 60 (2017) 436–446.

[69] D. Musicki, T. L. Song, Track initialization: Prior target veloaddress and acceleration moments, IEEE Trans. Aerosp. Electr. Syst. 49 (2013) 665–670.

[70] F. R. Castella, Sliding window detection probabilities, IEEE Trans. Aeros. Electr. Syst. AES-12 (1976) 815–819.

[71] Y. Bar-Shalom, K. C. Chang, H. M. Shertukde, Performance evaluation of a cascaded logic for track formation in clutter, IEEE Trans. Aerosp. Electr. Syst. 25 (1989) 873–878.

[72] Z. Hu, H. Leung, M. Blanchette, Statistical performance analysis of track initiation techniques, IEEE Trans. Signal Process. 45 (1997) 445–456.

[73] R. Worsham, The probabilities of track initiation and loss using a sliding window for track acquisition, in: 2010 IEEE Radar Conference, Arlington, VA, USA, pp. 1270–1275.

[74] T. Kariya, H. Kurata, Generalized Least Squares Estimators, John Wiley & Sons, Ltd, 2004.

[75] D. Luenberger, Optimization by Vector Space Methods, Wiley, New York, 1968.

[76] H. W. Sorenson, Least-squares estimation: from Gauss to Kalman, IEEE Spectrum 7 (1970) 63–68.

[77] A. H. Sayed, T. Kailath, A state-space approach to adaptive RLS filtering, IEEE Signal Process. Mag. 11 (1994) 18–60.

[78] J. Humpherys, J. West, Kalman filtering with Newton's method [lecture notes], IEEE Control Systems Magazine 30 (2010) 101–106.

[79] J. Rissanen, A Predictive Least-Squares Principle, IMA Journal of Mathematical Control and Information 3 (1986) 211–222.

[80] M. Niedźwiecki, M. Ciołek, Akaike's final prediction error criterion revisited, in: Proc. ICTSP 2017, Barcelona, Spain, pp. 237–242.

[81] R. L. Streit, T. E. Luginbuhl, Maximum likelihood method for probabilistic multihypothesis tracking, in: O. E. Drummond (Ed.), Signal and Data Processing of Small Targets 1994, volume 2235, International Society for Optics and Photonics, SPIE, 1994, pp. 394 – 405.

[82] B. GAO, W. LI, G. HU, Y. ZHONG, X. ZHU, Mahalanobis distance-based fading cubature kalman filter with augmented mechanism for hypersonic vehicle ins/cns autonomous integration, Chinese Journal of Aeronautics 35 (2022) 114–128.

[83] H. V. Poor, An Introduction to Signal Detection and Estimation, Springer, New York, NY, USA, 1994.

[84] H. Leung, Z. Hu, M. Blanchette, Evaluation of multiple target track initiation techniques in real radar tracking environments, IEE Proceedings - Radar, Sonar and Navigation 143 (1996) 246–254.

[85] A. G. Tartakovsky, G. Sokolov, Y. Bar-Shalom, Nearly optimal adaptive sequential tests for object detection, IEEE Trans. Signal Processing 68 (2020) 3371–3384.

[86] K. Li, G. Zhou, T. Kirubarajan, A general model-based filter initialization approach for linear and nonlinear dynamic systems, Digital Signal Processing 111 (2021) 102978.

[87] D. Schuhmacher, B.-T. Vo, B.-N. Vo, A consistent metric for performance evaluation of multi-object filters, IEEE Trans. Signal Process. 56 (2008) 3447–3457.

[88] T. Li, K. Da, H. Fan, B. Yu, Multisensor random finite set information fusion: Advances, challenges, and opportunities, in: Y. Cao, O. Kaiwartya, T. Li (Eds.), Secure and Digitalized Future Mobility: Shaping the Ground and Air Vehicles Cooperation, CRC Press, 2022.

[89] J. C. Ye, Y. Bresler, P. Moulin, Asymptotic global confidence regions in parametric shape estimation problems, IEEE Trans. Inf. Theory 46 (2000) 1881–1895.

[90] D. F. Vysochanskii, Y. L. Petunin, Justification of the 3σ rule for unimodal distributions, Theory Prob. Math. Stat. 21 (1980) 25 – 36.