Trajectory PHD and CPHD filters

Ángel F. García-Fernández, Lennart Svensson

Abstract—This paper presents the probability hypothesis density filter (PHD) and the cardinality PHD (CPHD) filter for sets of trajectories, which are referred to as the trajectory PHD (TPHD) and trajectory CPHD (TCPHD) filters. Contrary to the PHD/CPHD filters, the TPHD/TCPHD filters are able to produce trajectory estimates from first principles. The TPHD filter is derived by recursively obtaining the best Poisson multi-trajectory density approximation to the posterior density over the alive trajectories by minimising the Kullback-Leibler divergence. The TCPHD is derived in the same way but propagating an independent identically distributed (IID) cluster multitrajectory density approximation. We also propose the Gaussian mixture implementations of the TPHD and TCPHD recursions, the Gaussian mixture TPHD (GMTPHD) and the Gaussian mixture TCPHD (GMTCPHD), and the L-scan computationally efficient implementations, which only update the density of the trajectory states of the last L time steps.

Index Terms—Multitarget tracking, random finite sets, sets of trajectories, PHD, CPHD.

I. INTRODUCTION

The probability hypothesis density (PHD) and cardinality PHD (CPHD) filters are widely used random finite set (RFS) algorithms for multitarget filtering, which aims to estimate the state of the targets at the current time based on a sequence of measurements [1]–[7]. These filters have been successfully used in different applications such as multitarget tracking [1], distributed multi-sensor fusion [8], [9], robotics [10], [11], computer vision [12], [13], road mapping [14] and sensor control [15].

The PHD/CPHD filters fit into the assumed density filtering framework and propagate a certain type of multitarget density on the current set of targets through the prediction and update steps [16]. The PHD filter considers a Poisson multitarget density, in which the cardinality of the set is Poisson distributed and, for each cardinality, its elements are independent and identically distributed (IID). On the other hand, the CPHD filter considers an IID cluster multitarget density, in which the cardinality distribution of the set is arbitrary and, for each cardinality, its elements are IID. If the output of either the prediction or the update step is no longer Poisson/IID cluster, the PHD/CPHD filters obtain the best Poisson/IID cluster approximation by minimising the Kullback-Leibler divergence (KLD).

The most important benefit of the PHD/CPHD filters is their low computational burden, as they avoid the measurement-to-target association problem. However, their main drawbacks

are their relatively low performance in some scenarios [1], [17] and the fact that they do not build tracks, which denote sequences of target states that belong to the same target. The smoother versions of these filters [1], [18], [19] do not solve these drawbacks.

Despite the fact that the PHD/CPHD filters are unable to provide tracks in a mathematically rigorous way, several track building procedures have been proposed [20]–[24]. A track building procedure for PHD/CPHD filters was proposed in [25] by adding labels [26], [27] to the target states. Nonetheless, in the resulting labelled Poisson and labelled IID cluster densities, there is total confusion in the label-to-target association so they are not useful for track formation [25, Sec. III.B] [28, Sec. II.B]. To solve this issue in [25], apart from the unique labels, unique tags are added to the PHD components, as in [22], and the original PHD/CPHD recursions are applied. However, in the considered posterior density, the tags are not part of the target state and are marginalised out. Therefore, the posterior is still distributed as labelled Poisson or labelled IID cluster and, theoretically, it does not have information to infer tracks. While tagging PHD components works well in some scenarios, each PHD component does not generally represent information about a unique target, as the corresponding number of targets is Poisson distributed. In fact, adding tags to the PHD components and reporting estimates with unique tags to build trajectories, can lead to track switches, missed detections and false targets when there is more than one target represented by the same tag.

In this paper, we address the intrinsic inability of standard PHD/CPHD filters to infer trajectories by developing PHD/CPHD filters that provide tracks from first principles, without adding labels or tags. We propose the trajectory PHD (TPHD) and trajectory CPHD (TCPHD) filters, which follow the same assumed density filtering scheme as the PHD/CPHD filters [29] with a fundamental difference: instead of using a set of targets as the state variable, they use a set of trajectories [28], [30].

The TPHD filter propagates a Poisson multitrajectory density on the space of sets of trajectories through the prediction and update steps, with a KLD minimisation after the update step. A diagram of the resulting Bayesian recursion is given in Figure 1. Similarly, the TCPHD filter propagates an IID cluster multitrajectory density and performs a KLD minimisation after the prediction and update steps, see Figure 2. Due to the widespread use of PHD/CPHD filters, this paper covers an important gap in the literature, as we show how PHD/CPHD filtering can be endowed with the ability to infer trajectories in a rigorous and principled way. In particular, the TPHD and TCPHD filters are able to estimate the trajectories of the alive targets by propagating a Poisson and an IID cluster
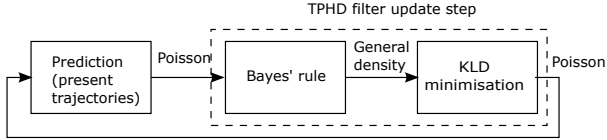
Figure 1: TPHD filter diagram for estimating the present trajectories at the current time. The TPHD filter assumes that the multitrajectory densities involved are Poisson (on the space of sets of trajectories). The output of Bayes' rule is no longer Poisson but, in order to be able to perform the Bayesian recursion, it obtains the best Poisson approximation to the filtering density by minimising the KLD.

multitrajectory density through the filtering recursion using KLD minimisations. The TPHD and TCPHD filters also avoid the above-mentioned drawbacks of trajectory estimation based on labelling/tagging the PHD. Apart from theoretically sound track formation, the proposed filters also have the advantage, compared to previous track building procedures used in PHD/CPHD filters, that they can update the information regarding past states of the trajectories.

In this paper, we also propose Gaussian mixture implementations of the TPHD/TCPHD filters, which follow the spirit of the Gaussian mixture PHD/CPHD filters [3], [5]. The resulting Gaussian mixture TPHD (GMTPHD) and TCPHD (GMTCPHD) filters build trajectories under a Poisson or IID cluster approximation, whose PHD is represented by a Gaussian mixture. In this setting, a Gaussian component of the GMTPHD/GMTCPHD filter represents information over entire trajectories, while a Gaussian component in the Gaussian mixture PHD/CPHD filters (tagged or not) represents information over current target states. It is therefore straightforward to extract trajectory estimates from the GMT-PHD/GMTCPHD filters. Additionally, we propose a version of the GMTPHD/GMTCPHD filters with lower computational burden called the L-scan GMTPHD/GMTCPHD filters. In practice, these filters only update the multitrajectory density of the trajectory states of the last L time instant leaving the rest unaltered, which is quite efficient for implementation. The theoretical foundation of the L-scan GMTPHD filter is also based on the assumed density filtering framework and KLD minimisations. Preliminary results of this paper covering the TPHD filter were presented in [29].

The remainder of the paper is organised as follows. Section II presents background material on sets of trajectories. In Section III, we introduce the Poisson and IID cluster multitrajectory densities and some of their properties. The TPHD and TCPHD filters are derived in Sections IV and V, respectively. Their Gaussian mixture implementations are provided in Section VI. Simulation results are provided in Section VII. Finally, conclusions are drawn in Section VIII.

II. BACKGROUND

In this section, we describe some background material on multiple target tracking using sets of trajectories [28]. We review the considered variables, the set integral and cardinality distribution for sets of trajectories in Sections II-A, II-B and II-C, respectively. Finally, we introduce the PHD for sets of trajectories in Section II-D.

A. Variables

A single target state $x \in \mathbb{R}^{n_x}$ contains information of interest about the target, e.g., its position and velocity. A set of single target states $\mathbf{x}$ belongs to $\mathcal{F}(\mathbb{R}^{n_x})$ where $\mathcal{F}(\mathbb{R}^{n_x})$ denotes the set of all finite subsets of $\mathbb{R}^{n_x}$. We are interested in estimating all target trajectories, where a trajectory consists of a sequence of target states that can start at any time step and end any time later on. Mathematically, a trajectory is represented as a variable $X = (t, x^{1:i})$ where t is the initial time step of the trajectory, i is its length and $x^{1:i} = (x^1, ..., x^i)$ denotes a sequence of length i that contains the target states at consecutive time steps of the trajectory.

We consider trajectories up to the current time step k. As a trajectory $(t, x^{1:i})$ exists from time step t to $t + i - 1$, variable (t, i) belongs to the set $I_{(k)} = \{(t, i) : 0 \leq t \leq k \text{ and } 1 \leq i \leq k - t + 1\}$. A single trajectory X up to time step k therefore belongs to the space $T_{(k)} = \biguplus_{(t,i) \in I_{(k)}} \{t\} \times \mathbb{R}^{in_x}$, where $\biguplus$ stands for disjoint union, which is used to highlight that the sets are disjoint. Similarly to the set $\mathbf{x}$ of targets, we denote a set of trajectories up to time step k as $\mathbf{X} \in \mathcal{F}(T_{(k)})$.

Given a trajectory $X = (t, x^{1:i})$, the set $\tau^{k'}(X)$, which can be empty, denotes the corresponding target state at a time step k'. Given a set $\mathbf{X}$ of trajectories, the set $\tau^{k'}(\mathbf{X})$ of target states at time k' is $\tau^{k'}(\mathbf{X}) = \bigcup_{X \in \mathbf{X}} \tau^{k'}(X)$.

B. Set integral

Given a real-valued function $\pi(\cdot)$ on the single trajectory space $T_{(k)}$, its integral is [28]

$$\int \pi(X) dX = \sum_{(t,i) \in I_{(k)}} \int \pi(t, x^{1:i}) dx^{1:i}. \tag{1}$$

This integral goes through all possible start times, lengths and target states of the trajectory. Given a real-valued function $\pi(\cdot)$ on the space $\mathcal{F}(T_{(k)})$ of sets of trajectories, its set integral is [28]

$$\int \pi(\mathbf{X}) \delta \mathbf{X} = \sum_{n=0}^{\infty} \frac{1}{n!} \int \pi(\{X_1, ..., X_n\}) dX_{1:n} \tag{2}$$

where $X_{1:n} = (X_1, ..., X_n)$. A function $\pi(\cdot)$ is a multitrajectory density if $\pi(\cdot) \geq 0$ and its set integral is one.

C. Cardinality distribution

Given a multitrajectory density $\pi(\cdot)$, its cardinality distribution is

$$\rho_\pi(n) = \frac{1}{n!} \int \pi(\{X_1, ..., X_n\}) dX_{1:n}, \tag{3}$$

which is analogous to the case where there is a set of targets.

D. Probability hypothesis density

The PHD [1] of a multitrajectory density $\pi(\cdot)$ is

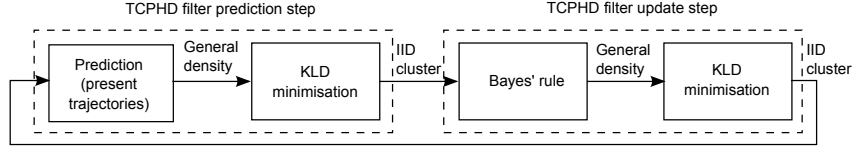$$D_\pi(X) = \int \pi(\{X\} \cup \mathbf{X}) \delta \mathbf{X}. \tag{4}$$

Figure 2: TCPHD filter diagram for estimating the present trajectories at the current time. The TCPHD filter assumes that the multitrajectory densities involved are IID cluster. The output of the prediction and Bayes' rule are no longer IID clusters but the TCPHD filter obtains the best IID cluster approximation by minimising the KLD.

As in the PHD for RFS of targets, integrating the PHD in a region $A \subseteq T_{(k)}$ gives us the expected number of trajectories in this region [1, Eq. (4.76)]:

$$
\hat{N}_A = \int_A D_\pi(X) dX
$$
$$
= \sum_{(t,i) \in I_{(k)}} \int 1_A \left(t, x^{1:i} \right) D_\pi(t, x^{1:i}) dx^{1:i} \quad (5)
$$

where $1_A (\cdot)$ is the indicator function of a subset A: $1_A (z) = 1$ if $z \in A$ and $1_A (z) = 0$ otherwise. Therefore, the expected number of trajectories up to time step k is given by substituting $A = T_{(k)}$ into (5).

Example 1. Let us consider a multitrajectory density $\nu (\cdot)$ with PHD

$$
D_\nu \left(1, x^1 \right) = \mathcal{N} \left(x^1; 10, 1 \right) + \mathcal{N} \left(x^1; 1000, 1 \right) \quad (6)
$$
$$
D_\nu \left(1, x^{1:2} \right) = \mathcal{N} \left(x^{1:2}; (10, 10.1), \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \right), \quad (7)
$$

and $D_\nu (X) = 0$ for $X \neq \left(1, x^1 \right)$ and $X \neq \left(1, x^{1:2} \right)$, where $\mathcal{N} (\cdot; m, P)$ is a Gaussian density with mean m and covariance matrix P. The expected number of trajectories that start at time one with length 1 is given by substituting $A = \{1\} \times \mathbb{R}^{n_x}$ into (5) so

$$
\hat{N}_A = \int D_\nu \left(1, x^1 \right) dx^1 = 2.
$$

The expected number of trajectories up to time step $k = 2$ is $\hat{N}_{T_{(k)}} = 3$. $\square$

III. POISSON AND IID CLUSTER TRAJECTORY RFSs

In this section, we explain the Poisson and IID cluster trajectory RFSs.

A. Multitrajectory densities

1) Poisson RFS: For a Poisson RFS, the cardinality of the set is Poisson distributed and, for each cardinality, its elements are IID. A Poisson multitrajectory density $\nu (\cdot)$ has the form

$$
\nu (\{X_1, ..., X_n\}) = e^{-\lambda_\nu} \lambda_\nu^n \prod_{j=1}^n \breve{\nu} (X_j) \quad (8)
$$

where $\breve{\nu} (\cdot)$ is a single trajectory density, which implies

$$
\int \breve{\nu} (X) dX = 1, \quad (9)
$$

and $\lambda_\nu \geq 0$. A Poisson multitrajectory density is characterised by either its PHD $D_\nu(X) = \lambda_\nu \breve{\nu} (X)$ or by λ_ν and $\breve{\nu} (\cdot)$ [1].

As a result, using (5), the expected number of trajectories is $\hat{N}_{T_{(k)}} = \lambda_\nu$. Further, its cardinality distribution is given by

$$
\rho_\nu (n) = \frac{1}{n!} \int \nu (\{X_1, ..., X_n\}) dX_{1:n} = \frac{1}{n!} e^{-\lambda_\nu} \lambda_\nu^n. \quad (10)
$$

Example 2. We consider a Poisson RFS with the PHD of Example 1. Using (10), its cardinality distribution is Poisson with $\lambda_\nu = 3$ and, therefore, its single trajectory density is $\breve{\nu} (X) = D_\nu (X) / 3$. $\square$

2) IID cluster RFS: For an IID cluster RFS with multitrajectory density $\nu (\cdot)$, the cardinality is distributed according to the probability mass function $\rho_\nu (\cdot)$ and, for each cardinality, its elements are IID according to a single trajectory density $\breve{\nu} (\cdot)$. The resulting multitrajectory density is

$$
\nu (\{X_1, ..., X_n\}) = \rho_\nu (n) n! \prod_{j=1}^n \breve{\nu} (X_j). \quad (11)
$$

As $\breve{\nu} (\cdot)$ is a single trajectory density, it meets (9). The PHD of (11) is given by [1]

$$
D_\nu (x) = \breve{\nu} (x) \sum_{n=0}^\infty n \rho_\nu (n) \quad (12)
$$

where the second factor corresponds to the expected number of trajectories. An IID cluster density can be characterised either by $\rho_\nu (\cdot)$ and $\breve{\nu} (\cdot)$, or by $\rho_\nu (\cdot)$ and $D_\nu (\cdot)$. How to draw samples from an IID cluster trajectory RFS, which includes the Poisson trajectory RFS as a particular case, is explained in Appendix A in the supplementary material.

B. KLD minimisation

In this subsection, we provide two KLD minimisation theorems for Poisson and IID cluster multitrajectory densities, which will be used to derive the trajectory PHD/CPHD filters.

The KLD $D (\pi \| \nu)$ between multitrajectory densities $\pi (\cdot)$ and $\nu (\cdot)$ is given by [1]

$$
D (\pi \| \nu) = \int \pi (\mathbf{X}) \log \frac{\pi (\mathbf{X})}{\nu (\mathbf{X})} \delta \mathbf{X}. \quad (13)
$$

Then, the following theorems hold:

Theorem 3. *Given a multitrajectory density $\pi (\cdot)$, the Poisson multitrajectory density $\nu (\cdot)$ that minimises the KLD $D (\pi \| \nu)$ is characterised by the PHD $D_\nu (\cdot) = D_\pi (\cdot)$.*

Theorem 4. *Given a multitrajectory density $\pi (\cdot)$, the IID cluster multitrajectory density $\nu (\cdot)$ that minimises the KLD*

$D\left(\pi\,\|\,\nu\right)$ *is characterised by the PHD* $D_\nu\left(\cdot\right)=D_\pi\left(\cdot\right)$ *and the cardinality distribution* $\rho_\nu\left(\cdot\right)=\rho_\pi\left(\cdot\right)$.

Theorem 3 is proved in Appendix A in [29]. The analogous theorem for sets of targets was proved in [31]. Theorem 4 is proved in Appendix B in the supplementary material. The analogous theorem for sets of targets was proved in [16], [32]. It should be noted that, as a Poisson RFS is a special type of IID cluster RFS, the best fitting IID cluster RFS always has a lower or equal KLD than the best fitting Poisson RFS.

C. Inference only on alive trajectories

In this section, we explain why the TPHD and TCPHD filters are mainly useful to approximate the posterior multi-trajectory density over alive trajectories, but not the posterior over all trajectories, which also include dead trajectories. This serves as a motivation to present the TPHD and TCPHD filters for tracking only the alive trajectories in the next sections.

Let us first explain why the TPHD filter, which considers a Poisson approximation, is only useful for the alive trajectories [29, Sec. V.B], though it was derived in [29] for dead and alive trajectories. In the prediction step, the part of the PHD that represents a trajectory that dies at the current time step is multiplied by the probability of death (one minus the probability of survival) [29, Thm. 5], which is usually low. As time goes on, the part of the PHD that represents dead trajectories never changes. As a result, even if a trajectory exists with a very high probability at some point in time, once it dies, the TPHD filter over all trajectories indicates that it existed with a very low probability. Therefore, the TPHD does not contain accurate information about dead trajectories, though it does contain useful information about alive trajectories.

In the following, we argue with an example why the TCPHD filter, which considers an IID cluster approximation, should only consider alive trajectories, as the TPHD filter.

Example 5. Let us consider that the posterior $\pi^k\left(\cdot\right)$ over the set of trajectories at time k has m trajectories with probability 1 so $\rho_{\pi^k}\left(m\right)=1$. In addition, $\pi^k\left(\cdot\right)$ indicates that there are m_d dead trajectories with independent (single trajectory) densities $\breve{d}_1\left(\cdot\right),...,\breve{d}_{m_d}\left(\cdot\right)$, and m_a alive trajectories with independent densities $\breve{a}_1\left(\cdot\right),...,\breve{a}_{m_a}\left(\cdot\right)$, where $m_d+m_a=m$. Note that we can obtain this kind of true posterior, without TPHD/TCPHD approximations, if the probability of detection is one, there is no clutter, targets are born independently and they are far from each other at all time steps. In Appendix C (see supplementary material), we compute the best IID cluster density approximation $\nu^k\left(\cdot\right)$ to $\pi^k\left(\cdot\right)$ using Theorem 4 and show that the cardinality distribution of the alive targets in $\nu^k\left(\cdot\right)$ is

$$\rho_a\left(n\right)=\left(\begin{array}{c}m\\n\end{array}\right)\left(\frac{m_a}{m}\right)^n\left(1-\frac{m_a}{m}\right)^{m-n}, \quad (14)$$

where $n\in\{0,1,...,m\}$. As the filtering recursion continues, the total number m of trajectories can only increase. On the contrary, m_a does not necessarily increase so after a sufficiently long time $\frac{m_a}{m}$ may become very small. Then, using the Poisson limit theorem, the cardinality distribution of the

alive targets can be approximated as Poisson with parameter m_a [33]. Therefore, even in this simple example in which the cardinality of the alive targets is known, the best IID cluster approximation of the whole trajectory posterior approximates the cardinality of the alive targets as a Poisson distribution. □

The conclusion of the previous example is that, in the long run, an IID cluster RFS is not necessarily better than a Poisson RFS, both considered over all trajectories, to approximate the cardinality distribution of the alive targets. In most applications, the cardinality of the alive trajectories is considerably more important than the cardinality of the total number of trajectories. In this paper, we therefore focus on an IID cluster approximation of the alive trajectories to develop the TCPHD filter. This implies that the TCPHD filter has an arbitrary cardinality distribution for the alive targets, as the CPHD filter.

IV. TRAJECTORY PHD FILTER

In this section, we derive the TPHD filter for tracking the alive targets. The TPHD propagates the multitrajectory density of a Poisson RFS through the filtering recursion. In the update step, the TPHD filter uses Bayes' rule followed by a KLD minimisation to approximate the posterior as Poisson, see Figure 1. In Section IV-A, we present the Bayesian filtering recursion for sets of trajectories. The prediction and update steps of the TPHD filter are given in Sections IV-B and IV-C, respectively.

A. Bayesian filtering recursion

The posterior multitrajectory density $\pi^k\left(\cdot\right)$ at time k, which denotes the density of set of trajectories present at time k given all measurements up to time k, is calculated via the prediction and update steps:

$$\omega^k\left(\mathbf{X}\right)=\int f\left(\mathbf{X}\,|\,\mathbf{Y}\right)\pi^{k-1}\left(\mathbf{Y}\right)\delta\mathbf{Y} \quad (15)$$

$$\pi^k\left(\mathbf{X}\right)=\frac{\ell^k\left(\mathbf{z}^k|\tau^k\left(\mathbf{X}\right)\right)\omega^k\left(\mathbf{X}\right)}{\ell^k\left(\mathbf{z}^k\right)} \quad (16)$$

where $f\left(\cdot\,|\,\cdot\right)$ is the transition density, $\omega^k\left(\cdot\right)$ is the predicted density at time k, $\mathbf{z}^k$ is the set of measurements at time k, $\ell^k\left(\cdot|\tau^k\left(\mathbf{X}\right)\right)$ is the density of the measurements given the current RFS of targets and

$$\ell^k\left(\mathbf{z}^k\right)=\int\ell^k\left(\mathbf{z}^k|\tau^k\left(\mathbf{X}\right)\right)\omega^k\left(\mathbf{X}\right)\delta\mathbf{X}$$

is the density of the measurements given the predicted density $\omega^k\left(\cdot\right)$. The predicted density at time k is the density of the set of trajectories present at time step k given the measurements up to time step $k-1$. As we only take into account the present trajectories, the only term that changes in (15)-(16) with respect to considering all trajectories is $f\left(\cdot\,|\,\cdot\right)$, see [34, Sec. IV.A] for a detailed explanation. The description of these models will be given in Sections IV-B and IV-C.

B. Prediction

We make the following assumptions in the prediction step:

- P1 Given the current set $\mathbf{x}$ of targets, each target $x \in \mathbf{x}$ survives with probability $p_S(x)$ and moves to a new state with a transition density $g(\cdot|x)$, or dies with probability $1 - p_S(x)$.
- P2 The multitarget state at the next time step is the union of the surviving targets and new targets, which are born independently with a Poisson multitarget density $\beta_\tau(\cdot)$.
- P3 The multitrajectory density $\pi^{k-1}(\cdot)$ of the trajectories present at time $k-1$ represents a Poisson RFS.

Note that we use subindex τ in densities on RFS of targets, as in $\beta_\tau(\cdot)$. Let $\mathbb{N}_k = \{1, ..., k\}$. Then, the relation between predicted PHD at time k and the PHD of the posterior at time $k-1$ is given by the following theorem.

Theorem 6 (TPHD filter prediction)**.** *Under Assumptions P1-P3, the predicted PHD $D_{\omega^k}(\cdot)$ of the trajectories present at time k is*

$$D_{\omega^k}(X) = D_{\xi^k}(X) + D_{\beta^k}(X) \tag{17}$$

where

$$D_{\beta^k}\left(t, x^{1:i}\right) = D_{\beta_\tau}\left(x^1\right) 1_{\{k\}}(t)$$
$$D_{\xi^k}\left(t, x^{1:i}\right) = p_S\left(x^{i-1}\right) g\left(x^i \left| x^{i-1}\right.\right)$$
$$\times D_{\pi_{k-1}}\left(t, x^{1:i-1}\right) 1_{\mathbb{N}_{k-1}}(t)$$

if $t + i - 1 = k$ or zero otherwise.

This theorem is proved in [29] for a more general case in which dead trajectories are considered. As mentioned in Section III-C, in this paper, we only present the results for alive trajectories, as the results are mainly useful in this case. The predicted PHD is the sum of the PHD $D_{\beta^k}(\cdot)$ of the trajectories born at time step k and the PHD $D_{\xi^k}(\cdot)$ of the surviving trajectories. The end time of trajectory $\left(t, x^{1:i}\right)$ is $t + i - 1$ so $D_{\omega^k}\left(t, x^{1:i}\right)$ is zero if $t + i - 1 \neq k$. For the surviving trajectories, we multiply the PHD by the transition density and the survival probability. Note that the provided PHD characterises the Poisson RFS that represents the predicted density.

C. Update

We make the following assumptions in the update step:

- U1 For a given multi-target state $\mathbf{x}$ at time k, each target state $x \in \mathbf{x}$ is either detected with probability $p_D(x)$ and generates one measurement with density $l(\cdot|x)$, or missed with probability $1 - p_D(x)$.
- U2 The measurement $\mathbf{z}^k$ is the union of the target-generated measurements and Poisson clutter with density $c(\cdot)$.
- U3 The multitrajectory density $\omega^k(\cdot)$ represents a Poisson RFS.

Let Ξ_{n,n_z} denote the set that contains all the vectors $\sigma = (\sigma_1, ..., \sigma_n)$ that indicate associations of n_z measurements to n targets, which can be either detected or undetected. If $\sigma \in \Xi_{n,n_z}$, $\sigma_i = j \in \{1, ..., n_z\}$ indicates that measurement j is associated with target i and $\sigma_i = 0$ indicates that target i

has not been detected. Under Assumptions U1 and U2, which define the standard measurement model, the density of the measurement given the state is [1, Eq. (7.21)]

$$\ell^k\left(\{z_1, ..., z_{n_z}\} | \{x_1, ..., x_n\}\right)$$
$$= e^{-\lambda_c} \left[\prod_{i=1}^{n_z} \lambda_c \breve{c}(z_i)\right] \left[\prod_{i=1}^n \left(1 - p_D(x_i)\right)\right]$$
$$\times \sum_{\sigma \in \Xi_{n,n_z}} \prod_{i:\sigma_i > 0} \frac{p_D(x_i) l(z_{\sigma_i}|x_i)}{\left(1 - p_D(x_i)\right) \lambda_c \breve{c}(z_{\sigma_i})}. \tag{18}$$

where λ_c and $\breve{c}(\cdot)$ characterise $c(\cdot)$, see (8).

Let $L_{\mathbf{z}^k}(\cdot)$ denote the PHD filter pseudolikelihood function, which is given by [1, Sec. 8.4.3]

$$L_{\mathbf{z}^k}(x) = 1 - p_D(x) + p_D(x)$$
$$\times \sum_{z \in \mathbf{z}^k} \frac{l(z|x)}{\lambda_c \breve{c}(z) + \int p_D(y) l(z|y) D_{\omega_\tau^k}(y) \, dy}$$

with $D_{\omega_\tau^k}(\cdot)$ representing the PHD of the targets at time k of density $\omega^k(\cdot)$, which is given by [29]

$$D_{\omega_\tau^k}(y) = \sum_{t=1}^k \int D_{\omega^k}\left(t, x^{1:k-t}, y\right) dx^{1:k-t}. \tag{19}$$

Then, the TPHD filter update step is given by the following theorem:

Theorem 7 (TPHD filter update)**.** *Under Assumptions U1-U3, the updated PHD $D_{\pi^k}(\cdot)$ at time k is*

$$D_{\pi^k}\left(t, x^{1:i}\right) = D_{\omega^k}\left(t, x^{1:i}\right) L_{\mathbf{z}^k}\left(x^i\right)$$

if $t + i - 1 = k$ or zero, otherwise.

This theorem is proved in [29] for a more general case in which dead trajectories are included. It should be noted that Bayes' update (16) uses a likelihood (18) that involves a summation over all target-to-measurement associations in the multitarget space. In contrast, the TPHD filter update is similar to the PHD filter update in the sense that it uses a pseudolikelihood function $L_{\mathbf{z}^k}(\cdot)$ which is defined on the single target space and only involves associations between a single target and the measurements.

V. TRAJECTORY CPHD FILTER

In this section we present the trajectory CPHD (TCPHD) filter for tracking the alive targets. The TCPHD propagates the multitrajectory density of an IID cluster RFS through the filtering recursion. In the prediction and update steps, the TCPHD filter makes use of a KLD minimisation to obtain an IID cluster approximation, see Figure 2.

Prior to deriving the TCPHD filter, we provide some notation. Given two sequences $a(n)$ and $b(n)$, $n \in \mathbb{N} \cup \{0\}$, we denote

$$\langle a, b \rangle = \sum_{n=0}^\infty a(n) b(n).$$

Given a set $\mathbf{z}$, the elementary symmetric function of order j is [5]

$$e_j(\mathbf{z}) = \sum_{\mathbf{s} \subseteq \mathbf{z}, |S| = j} \left(\prod_{\zeta \in \mathbf{s}} \zeta \right) \quad (20)$$

with $e_0(Z) = 1$ by convention. We also use $\backslash$ to denote set subtraction.

A. Prediction

The TCPHD filter prediction is obtained under Assumptions P1 and the additional assumptions

- P4 The multitarget state at the next time step is the union of the surviving targets and new targets, which are born independently with an IID cluster multitarget density $\beta_\tau(\cdot)$.
- P5 The multitrajectory density $\pi^{k-1}(\cdot)$ represents an IID cluster RFS.

Under Assumption P4, the set of new born trajectories at time k has cardinality $\rho_{\beta^k}(\cdot) = \rho_{\beta_\tau}(\cdot)$ and PHD

$$D_{\beta^k}\left(t, x^{1:i}\right) = \begin{cases} D_{\beta_\tau}\left(x^1\right) & t = k, \, i = 1 \\ 0 & \text{otherwise.} \end{cases}$$

The TCPHD filter prediction consists of applying the usual prediction step plus a KLD minimisation, which is performed by calculating the cardinality distribution and PHD of the predicted density, see Figure 2. The result is given in the following theorem.

Theorem 8 (TCPHD filter prediction). *Under Assumptions P1, P4 and P5, the PHD of the predicted density is the same as in the PHD filter, see Theorem 6. The cardinality distribution of the predicted density is*

$$\rho_{\omega^k}(m) = \sum_{j=0}^{m} \rho_{\beta^k}(m-j) \sum_{n=j}^{\infty} \binom{n}{j} \rho_{\pi^{k-1}}(n)$$
$$\times \frac{\left[\int (1 - p_S(x)) D_{\pi_\tau^{k-1}}(x) dx\right]^{n-j}}{\left[\int D_{\pi_\tau^{k-1}}(x) dx\right]^{n}}$$
$$\times \left[\int p_S(x) D_{\pi_\tau^{k-1}}(x) dx\right]^{j} \quad (21)$$

where $D_{\pi_\tau^{k-1}}(\cdot)$ is the PHD of the targets at time $k-1$ according to $\pi^{k-1}(\cdot)$, which is calculated as in (19).

Theorem 8 is proved in Appendix E (see supplementary material). The TCPHD filter prediction updates the cardinality distribution as the CPHD filter. The TCPHD does not integrate out past states of the trajectories in the PHD to keep trajectory information, while the CPHD filter does.

B. Update

The TCPHD filter update is derived under Assumption U1 and the additional assumptions

- U4 The measurement $\mathbf{z}^k$ is the union of the target-generated measurements and IID cluster clutter with density $c(\cdot)$.

- U5 The multitrajectory density $\omega^k(\cdot)$ represents an IID cluster RFS.

As indicated in Figure 2, the TCPHD update consists of applying Bayes' rule, see (16), followed by a KLD minimisation, which is performed as indicated by Theorem 4. We first consider the distribution of the present targets at the current time. Under Assumption U5, it is direct to obtain that the distribution of the targets present at time k is also an IID cluster with cardinality distribution $\rho_{\omega^k}(\cdot)$ and PHD (19). The resulting TCPHD filter update is given in the following theorem.

Theorem 9 (TCPHD filter update). *Under Assumptions U1, U4 and U5, the cardinality distribution and the PHD of the posterior at time k are*

$$\rho_{\pi^k}(n) = \frac{\Upsilon^0\left[D_{\omega_\tau^k}, \mathbf{z}^k\right](n) \rho_{\omega^k}(n)}{\left\langle \Upsilon^0\left[D_{\omega_\tau^k}, \mathbf{z}^k\right], \rho_{\omega^k}\right\rangle} \quad (22)$$

$$D_{\pi^k}\left(t, x^{1:i}\right) = \frac{\left\langle \Upsilon^1\left[D_{\omega_\tau^k}, \mathbf{z}^k\right], \rho_{\omega^k}\right\rangle}{\left\langle \Upsilon^0\left[D_{\omega_\tau^k}, \mathbf{z}^k\right], \rho_{\omega^k}\right\rangle}$$
$$\times \left(1 - p_D\left(x^i\right)\right) D_{\omega^k}\left(t, x^{1:i}\right)$$
$$+ \sum_{z \in \mathbf{z}^k} \frac{\left\langle \Upsilon^1\left[D_{\omega_\tau^k}, \mathbf{z}^k \backslash \{z\}\right], \rho_{\omega^k}\right\rangle}{\left\langle \Upsilon^0\left[D_{\omega_\tau^k}, \mathbf{z}^k\right], \rho_{\omega^k}\right\rangle}$$
$$\times \frac{l\left(z|x^i\right)}{\check{c}(z)} p_D\left(x^i\right) D_{\omega^k}(t, x^{1:i}) \quad (23)$$

if $t + i - 1 = k$ or $D_{\pi^k}\left(t, x^{1:i}\right) = 0$, otherwise, and

$$\Upsilon^u\left[D_{\omega_\tau^k}, \mathbf{z}^k\right](n) = \sum_{j=0}^{\min\left(|\mathbf{z}^k|, n-u\right)} \left(|\mathbf{z}^k| - j\right)! \rho_c\left(|\mathbf{z}^k| - j\right)$$
$$\times \frac{\left[\int (1 - p_D(x)) D_{\omega_\tau^k}(x) dx\right]^{n-(j+u)}}{\left[\int D_{\omega_\tau^k}(x) dx\right]^{n}}$$
$$\times \frac{n!}{(n-j-u)!} e_j\left(\Xi\left(D_{\omega_\tau^k}, \mathbf{z}^k\right)\right) \quad (24)$$

$$\Xi\left(D_{\omega_\tau^k}, \mathbf{z}^k\right) = \left\{\int \frac{l(z|x)}{\check{c}(z)} p_D(x) D_{\omega_\tau^k}(x) dx : z \in \mathbf{z}^k\right\}.$$

Theorem 9 is proved in Appendix D (see supplementary material). The update step of the TCPHD filter is equivalent to the CPHD filter, with the main difference that the updated PHD contains information about previous states of the trajectories.

VI. GAUSSIAN MIXTURE IMPLEMENTATIONS

In this section, we present the Gaussian mixture implementations of the TPHD and TCPHD filters. We use the notation

$$\mathcal{N}\left(t, x^{1:i}; t^k, m^k, P^k\right) = \begin{cases} \mathcal{N}\left(x^{1:i}; m^k, P^k\right) & t = t^k, \, i = i^k \\ 0 & \text{otherwise} \end{cases}$$
$$(25)$$

where $i^k = \dim\left(m^k\right)/n_x$. Equation (25) represents a single trajectory Gaussian density with start time t^k, duration i^k, mean $m^k \in \mathbb{R}^{i^k n_x}$ and covariance matrix $P^k \in \mathbb{R}^{i^k n_x \times i^k n_x}$ evaluated at $\left(t, x^{1:i}\right)$. We use $\otimes$ to indicate the Kronecker product and $0_{m,n}$ is the $m \times n$ zero matrix.

We make the additional assumptions

- A1 The probabilities p_S and p_D are constants.
- A2 $g\left(x^i \left| x^{i-1}\right.\right) = \mathcal{N}\left(x^i; F x^{i-1}, Q\right)$.
- A3 $l\left(z|x\right) = \mathcal{N}\left(z; Hx, R\right)$.
- A4 The PHD of the birth density $\beta^k\left(\cdot\right)$ is

$$D_{\beta^k}\left(X\right) = \sum_{j=1}^{J_\beta^k} w_{\beta,j}^k \mathcal{N}\left(X; k, m_{\beta,j}^k, P_{\beta,j}^k\right) \qquad (26)$$

where $J_\beta^k \in \mathbb{N}$ is the number of components, $w_{\beta,j}^k$ is the weight of the jth component, $m_{\beta,j}^k \in \mathbb{R}^{n_x}$ its mean and $P_{\beta,j}^k \in \mathbb{R}^{n_x \times n_x}$ its covariance matrix.

It should be noted that $F \in \mathbb{R}^{n_x \times n_x}$ is the single-target transition matrix, $Q \in \mathbb{R}^{n_x \times n_x}$ is the covariance matrix of the single-target process noise, $H \in \mathbb{R}^{n_z \times n_x}$ is the single-measurement matrix and $R \in \mathbb{R}^{n_z \times n_z}$ is the covariance matrix of the single-measurement noise. In addition, the models provided by A1-A4 could be time varying but time is omitted for notational convenience. In the rest of this section, we present the Gaussian mixture implementations of the TPHD and TCPHD filters in Sections VI-A and VI-B, respectively. The L-scan versions of the filters and trajectory estimation are addressed in Sections VI-C and VI-D. Finally, a discussion is provided in Section VI-E.

A. Gaussian mixture TPHD filter

Under Assumptions A1-A4, P1-P3 and U1-U3, we can calculate the TPHD filter in closed form giving rise to the GMTPHD filter, whose prediction and update steps are provided in the following propositions.

Proposition 10 (GMTPHD filter prediction). *Assume $\pi^{k-1}\left(\cdot\right)$ has a PHD*

$$D_{\pi^{k-1}}\left(X\right) = \sum_{j=1}^{J^{k-1}} w_j^{k-1} \mathcal{N}\left(X; t_j^{k-1}, m_j^{k-1}, P_j^{k-1}\right)$$

where $t_j^{k-1} + i_j^{k-1} - 1 = k - 1$ with $i_j^{k-1} = \dim\left(m_j^{k-1}\right)/n_x$. Then, the PHD of $\omega^k\left(\cdot\right)$ is

$$D_{\omega^k}\left(X\right) = D_{\beta^k}\left(X\right) + p_S \sum_{j=1}^{J^{k-1}} w_j^{k-1} \mathcal{N}\left(X; t_j^{k-1}, m_{\omega,j}^k, P_{\omega,j}^k\right)$$
$$(27)$$

where

$$m_{\omega,j}^k = \left[\left(m_j^{k-1}\right)^T, \left(\dot{F}_j m_j^{k-1}\right)^T\right]^T,$$
$$\dot{F}_j = \left[0_{1,i_j^{k-1}-1}, 1\right] \otimes F,$$
$$P_{\omega,j}^k = \begin{bmatrix} P_j^{k-1} & P_j^{k-1} \dot{F}_j^T \\ \dot{F}_j P_j^{k-1} & \dot{F}_j P_j^{k-1} \dot{F}_j^T + Q \end{bmatrix}.$$

Proposition 10 is a consequence of Theorem 6 and conventional properties of Gaussian densities. Compared to the GMPHD filter prediction, the main difference is that previous states are not integrated out.

Proposition 11 (GMTPHD filter update). *Assume $\omega^k\left(\cdot\right)$ has a PHD*

$$D_{\omega^k}\left(X\right) = \sum_{j=1}^{J_\omega^k} w_{\omega,j}^k \mathcal{N}\left(X; t_{\omega,j}^k, m_{\omega,j}^k, P_{\omega,j}^k\right). \qquad (28)$$

Then, the PHD of $\pi^k\left(\cdot\right)$ is

$$D_{\pi^k}\left(X\right) = \left(1 - p_D\right) D_{\omega^k}\left(X\right)$$
$$+ \sum_{z \in \mathbf{z}^k} \sum_{j=1}^{J^k} w_j\left(z\right) \mathcal{N}\left(X; t_{\omega,j}^k, m_j^k\left(z\right), P_j^k\right) \qquad (29)$$

where

$$w_j\left(z\right) = \frac{p_D w_{\omega,j}^k q_j\left(z\right)}{\lambda_c \check{c}\left(z\right) + p_D \sum_{l=1}^{J_\omega^k} w_{\omega,l}^k q_l\left(z\right)}$$
$$\overline{z}_j = \dot{H}_j m_{\omega,j}^k, \quad S_j = \dot{H}_j P_{\omega,j}^k \dot{H}_j^T + R$$
$$\dot{H}_j = \left[0_{1,i_{\omega,j}^k-1}, 1\right] \otimes H$$
$$q_j\left(z\right) = \mathcal{N}\left(z; \overline{z}_j, S_j\right)$$
$$m_j^k\left(z\right) = m_{\omega,j}^k + P_{\omega,j}^k \dot{H}^T S_j^{-1}\left(z - \overline{z}_j\right)$$
$$P_j^k = P_{\omega,j}^k - P_{\omega,j}^k \dot{H}^T S_j^{-1} \dot{H} P_{\omega,j}^k.$$

and $i_{\omega,j}^k = \dim\left(m_{\omega,j}^k\right)/n_x$

Proposition 11 is a consequence of Theorem 7 and the Kalman filter update equations [35]. Also, the GMTPHD filter update is similar to the GMPHD filter update. The main differences is that the GMTPHD updates the whole trajectories.

B. Gaussian mixture TCPHD filter

The GMTCPHD filtering recursion requires Assumptions A1-A4, P1, P4, P5, U1, U4 and U5, and is given by the following propositions.

Proposition 12 (GMTCPHD filter prediction). *Assume the posterior $\pi^{k-1}\left(\cdot\right)$ has a cardinality distribution $\rho_{\pi^{k-1}}\left(\cdot\right)$ and a PHD*

$$D_{\pi^{k-1}}\left(X\right) = \sum_{j=1}^{J^{k-1}} w_j^{k-1} \mathcal{N}\left(X; t_j^{k-1}, m_j^{k-1}, P_j^{k-1}\right). \qquad (30)$$

Then, the TCPHD filter prediction yields

$$\rho_{\omega^k}\left(m\right) = \sum_{j=0}^m \rho_{\beta^k}\left(m - j\right)$$
$$\times \sum_{n=j}^\infty \binom{n}{j} \rho_{\pi^{k-1}}\left(n\right) \left(1 - p_S\right)^{n-j} p_S^j$$

$$D_{\omega^k}\left(X\right) = D_{\beta^k}\left(X\right) + p_S \sum_{j=1}^{J^{k-1}} w_j^{k-1} \mathcal{N}\left(X; t_j^{k-1}, m_{\omega,j}^k, P_{\omega,j}^k\right)$$

where $m_{\omega,j}^k$ and $P_{\omega,j}^k$ are given in Proposition 10.

Proposition 12 is a consequence of Theorem 8. The main difference with the GMCPHD filter is that the GMTCPHD filter does not integrate out previous states in the PHD. It should also be noted that the GMTPHD and the GMTCPHD

propagate the PHD in the same way. The difference is that the GMTCPHD also considers the cardinality of $\omega^k(\cdot)$.

Proposition 13 (GMTCPHD filter update). *Assume the prior $\omega^k(\cdot)$ has a cardinality distribution $\rho_{\omega^k}(\cdot)$ and a PHD*

$$D_{\omega^k}(X) = \sum_{j=1}^{J_\omega^k} w_{\omega,j}^k \mathcal{N}\left(X; t_{\omega,j}^k, m_{\omega,j}^k, P_{\omega,j}^k\right). \tag{31}$$

Then, the TCPHD filter update yields

$$\rho_{\pi^k}(n) = \frac{\Psi^0\left[w_\omega^k, \mathbf{z}^k\right](n) \rho_{\omega^k}(n)}{\left\langle \Psi^0\left[w_\omega^k, \mathbf{z}^k\right], \rho_{\omega^k}\right\rangle}$$

$$D_{\pi^k}(X) = \frac{\left\langle \Psi^1\left[w_\omega^k, \mathbf{z}^k\right], \rho_{\omega^k}\right\rangle}{\left\langle \Psi^0\left[w_\omega^k, \mathbf{z}^k\right], \rho_{\omega^k}\right\rangle}(1 - p_D) D_{\omega^k}(X)$$

$$+ \sum_{z \in \mathbf{z}^k} \sum_{j=1}^{J_\omega^k} w_j(z) \mathcal{N}\left(X; t_{\omega,j}^k, m_j^k(z), P_j^k\right)$$

where

$$\Psi^u\left[w_\omega^k, \mathbf{z}^k\right](n) = \sum_{j=0}^{\min(|\mathbf{z}^k|, n-u)} \left(|\mathbf{z}^k| - j\right)! \rho_c\left(|\mathbf{z}^k| - j\right)$$

$$\times \frac{(1 - p_D)^{n-(j+u)}}{\langle 1, w_\omega^k\rangle^{j+u}} \frac{n!}{(n-j-u)!}$$

$$\times e_j\left(\Lambda\left(w_\omega^k, \mathbf{z}^k\right)\right)$$

$$\Lambda\left(w_\omega^k, \mathbf{z}^k\right) = \left\{\frac{p_D}{\check{c}(z)}\left(w_\omega^k\right)^T q(z) : z \in \mathbf{z}^k\right\}$$

$$w_\omega^k = \left[w_{\omega,1}^k, ..., w_{\omega,J_\omega^k}^k\right]^T$$

$$q(z) = \left[q_1(z), ..., q_{J_\omega^k}(z)\right]^T$$

$$w_j(z) = \frac{p_D w_{\omega,j}^k q_j(z) \left\langle \Psi^1\left[w_\omega^k, \mathbf{z}^k \setminus \{z\}\right], \rho_{\omega^k}\right\rangle}{\check{c}(z) \left\langle \Psi^0\left[w_\omega^k, \mathbf{z}^k\right], \rho_{\omega^k}\right\rangle}$$

and $q_j(z)$, $m_j^k(z)$ and P_j^k are given in Proposition 11.

Proposition 13 is a consequence of Theorem 9 and the Kalman filter update [35]. The GMTCPHD filter update is analogous to the GMCPHD filter [5], with the difference that previous states of the target trajectories are also included.

C. L-scan implementations

In this section, we propose the use of pruning and absorption to limit the number of components in the Gaussian mixture and a computationally efficient implementation of the Gaussian mixture filters: the L-scan GMTPHD and GMTCPHD filters.

The PHD of the GMTPHD/GMTCPHD filters has an increasing number of components as time progresses and, to limit complexity, we need to bound the number of components. We use the following techniques: pruning with threshold Γ_p, setting a maximum number J_{max} of components and absorption [29]. Absorption consists of removing components of the PHD whose distribution of the current target state is close to the distribution of the current target state of another component with a higher weight, and adding the weights of the removed components to the weight of the component that has not been removed. Absorption is motivated by the fact that

if two components have a very similar distribution over the current target state, based on a Mahalanobis distance criterion, future measurements will affect both component weights and future states in a similar way. Therefore, without absorption, we would have two components with practically the same Gaussian densities for the trajectory states corresponding to recent time steps, for which we would be repeating the same calculations. It should also be noted that single trajectory densities can be quite different in the past even if they are similar for the current target state. Therefore, a direct use of merging [36] for single trajectory densities, which would use moment matching at all time steps, can provide poor results and absorption is preferred. The steps of the pruning and absorption algorithms for the GMTPHD/GMTCPHD filters are given in Algorithm 1, where we use the notation $\Phi_j^k = \left(w_j^k, t_j^k, m_j^k, P_j^k\right)$.

Algorithm 1 Pruning and absorption for the GMTPHD and GMTCPHD filters

---

Input: Posterior PHD parameters $\left\{\Phi_j^k\right\}_{j=1}^{J^k}$, pruning threshold Γ_p, absorption threshold Γ_a, maximum number of terms J_{max}.
Output: Pruned posterior PHD parameters $\left\{\Phi_{o,j}^k\right\}_{j=1}^{\hat{j}^k}$

- Set $l = 0$ and $I = \left\{j \in \{1, ..., J^k\} : w_j^k > \Gamma_p\right\}$.
while $I \neq \emptyset$ **do**
   - Set $l \leftarrow l + 1$.
   - $j = \arg\max_{i \in I} w_i^k$.
   - $L = \left\{i \in I : \left(\hat{m}_i^k - \hat{m}_j^k\right)^T \left(\hat{P}_j^k\right)^{-1}\left(\hat{m}_i^k - \hat{m}_j^k\right) \leq \Gamma_a\right\}$
with $\hat{m}_j^k \in \mathbb{R}^{n_x}$ and $\hat{P}_j^k \in \mathbb{R}^{n_x \times n_x}$ denoting the mean and covariance matrix of the state at the current time step of the PHD component indexed by j.
   - $\Phi_{o,l}^k = \Phi_j^k$ with weight $w_{o,l}^k = \sum_{i \in L} w_i^k$.
   - $I \leftarrow I \setminus L$.
end while
- If $l > J_{max}$, only keep the J_{max} components with highest weight.

In addition, as time progresses, the lengths of the trajectories increase so, eventually, it is not computationally feasible to implement the proposed filters directly. In order to address this problem, we propose the L-scan implementations that propagate the joint density of the states of the last L time steps and independent densities for the previous states for each component of the PHD. This approach has a Kullback-Leibler divergence interpretation [29] and is motivated by the fact that measurements at the current time step only have a significant impact on the trajectory state estimates for recent time steps.

The L-scan GMTPHD/GMTCPHD filters are implemented as the GMTPHD/GMTCPHD with a minor modification in the prediction step, where we discard the correlations of states that happened L time steps before the current time step. Given the predicted PHD $D_{\omega^k}(\cdot)$ in Gaussian mixture form, see (27), its L-scan version is given by approximating the covariance matrices $P_{\omega,j}^k$ as

$$P_{\omega,j}^k \approx \text{diag}\left(\tilde{P}_j^{t_{\omega,j}^k}, \tilde{P}_j^{t_{\omega,j}^k+1}, ..., \tilde{P}_j^{k-L}, \tilde{P}_j^{k-L+1:k}\right) \tag{32}$$

where matrix $\tilde{P}_j^{k-L+1:k} \in \mathbb{R}^{L \cdot n_x \times L \cdot n_x}$ represents the joint covariance of the L last time instants, obtained from $P_{\omega,j}^k$,

and $\tilde{P}_j^k \in \mathbb{R}^{n_x \times n_x}$ represents the covariance matrix of the target state at time k, obtained from $P_{\omega,j}^k$. Therefore, we have independent Gaussian densities to represent the states outside the L-scan window, and a joint Gaussian density for the states in the L-scan window, as in [37]. The steps of the L-scan GMTPHD and GMTCPHD filters are summarised in Algorithm 2. It should be noted that the cardinality distribution is not affected by the choice of L in both filters.

Algorithm 2 L-scan Gaussian mixture TPHD/TCPHD filter steps

- Initialisation:
   - For TPHD: $D_{\pi^0}(\cdot) = 0$.
   - For TCPHD: $D_{\pi^0}(\cdot) = 0$, $\rho_{\pi^0}(0) = 1$.
for $k = 1$ to *final time step* **do**
   - Prediction:
     - For TPHD: use Proposition 10.
     - For TCPHD: use Proposition 12.
     - Approximate $P_{\omega,j}^k$ using (32), which discards correlations outside the L-scan window.
   - Update:
     - For TPHD: use Proposition 11.
     - For TCPHD: use Proposition 13.
   - Perform pruning/absorption using Algorithm 1.
   - Estimate the alive trajectories, see Section VI-D.
end for

D. Estimation

In this section, we adapt two commonly used estimators of the GMPHD and GMCPHD filters to the GMTPHD and GMTCPHD filters. We have observed via simulations that better performance is obtained if these estimators are applied after the pruning/absorption step, as indicated in Algorithm 2.

1) GMTPHD: We adapt the estimator for the GMPHD filter described in [1, Sec. 9.5.4.4] for sets of trajectories. First, the number of trajectories is estimated as

$$\hat{N}^k = \text{round}\left(\sum_{j=1}^{J^k} w_j^k\right). \quad (33)$$

Then, the estimated set of trajectories corresponds to $\left\{\left(t_{l_1}^k, m_{l_1}^k\right), ..., \left(t_{l_{\hat{N}^k}}^k, m_{l_{\hat{N}^k}}^k\right)\right\}$ where $\{l_1, ..., l_{\hat{N}^k}\}$ are the indices of the PHD components with highest weights.

2) GMTCPHD: We adapt the estimator for the GMCPHD filter described in [1, Sec. 9.5.5.4] for sets of trajectories. The estimated cardinality at time step k is obtained as

$$\hat{N}^k = \underset{n \in \mathbb{N} \cup \{0\}}{\arg\max} \, \rho_{\pi^k}(n). \quad (34)$$

Then, the estimated set of trajectories are given by $\left\{\left(t_{l_1}^k, m_{l_1}^k\right), ..., \left(t_{l_{\hat{N}^k}}^k, m_{l_{\hat{N}^k}}^k\right)\right\}$ where $\{l_1, ..., l_{\hat{N}^k}\}$ are the indices of the PHD components with highest weights.

E. Discussion

In this section, we discuss some of the prominent aspects of the proposed filters. The TPHD/TCPHD filters have a similar structure as the PHD/CPHD filters, with the additional benefit of providing trajectory estimates for the alive targets. That is, at each time step, these filters can estimate the trajectories

of each of the targets, including its time of birth. The main difference between the trajectory PHD/CPHD filters and their target counterparts is that the trajectory filters do not integrate out previous states of the trajectories.

The L-scan GMTPHD/GMTCPHD filters are computationally efficient implementations, which allow the propagation of the posterior for long time sequences. In fact, the 1-scan versions ($L = 1$) of the GMTPHD/GMTCPHD filters perform the same computations as the GMPHD/GMCPHD, the only difference being that the trajectory versions store the mean and covariance of the trajectory state at each time instant for each component of the PHD.

It should be noted that the computations of the cardinality probability mass functions and elementary symmetric functions are the same as in the GMCPHD filter. To calculate the elementary symmetric functions, we use the recursive formula in [38, Eq. (2.3)]. We would also like to note that we have presented the filters considering linear and Gaussian models for ease of exposition. Nevertheless, we can still apply the GMTPHD and GMTCPHD filters for nonlinear measurement and dynamic models by first linearising the system and then applying the prediction and update steps with the linearised model. This is the usual procedure in nonlinear Gaussian filtering, for example, as in the extended Kalman filter, the unscented Kalman filter or the iterated posterior linearisation filter [35], [39].

VII. Simulations

We proceed to assess the performance of the two proposed filters in comparison with the previous track building procedures for PHD/CPHD filters based on tagging each PHD component [22]. In this track building approach, we estimate the number $\hat{n}_k$ of targets as indicated in Section VI-D and take the $\hat{n}_k$ highest components of the PHD with different tags. These estimates are then appended to the estimated trajectories with the same tag at the previous time step. We refer to these algorithms as tagged PHD/CPHD filters. All units of the quantities in this section are given in the international system.

We consider a target state $x = [p_x, \dot{p}_x, p_y, \dot{p}_y]^T$, which contains position and velocity. The parameters of the single-target dynamic process are

$$F = I_2 \otimes \begin{pmatrix} 1 & \tau \\ 0 & 1 \end{pmatrix}, \quad Q = qI_2 \otimes \begin{pmatrix} \tau^3/3 & \tau^2/2 \\ \tau^2/2 & \tau \end{pmatrix}$$

where $\tau = 0.5$ is the sampling time and $q = 3.24$ is a parameter. We also set $p_S = 0.99$. The parameters of the measurement model are

$$H = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad R = \sigma^2 I_2,$$

where $\sigma^2 = 4$, and $p_D = 0.9$. The clutter intensity is $D_c(z) = \lambda_c \cdot u_A(z)$ where $u_A(z)$ is a uniform density in region $A = [0, 2000] \times [0, 2000]$ and $\lambda_c = 50$ is the average number of clutter measurements per scan. The birth process is Poisson with a PHD that is represented by a Gaussian mixture with parameters: $J_\beta^k = 3$, $w_{\beta,j}^k = 0.1$, $P_{\beta,j}^k = \text{diag}([225, 100, 225, 100])$ for $j \in \{1, 2, 3\}$,
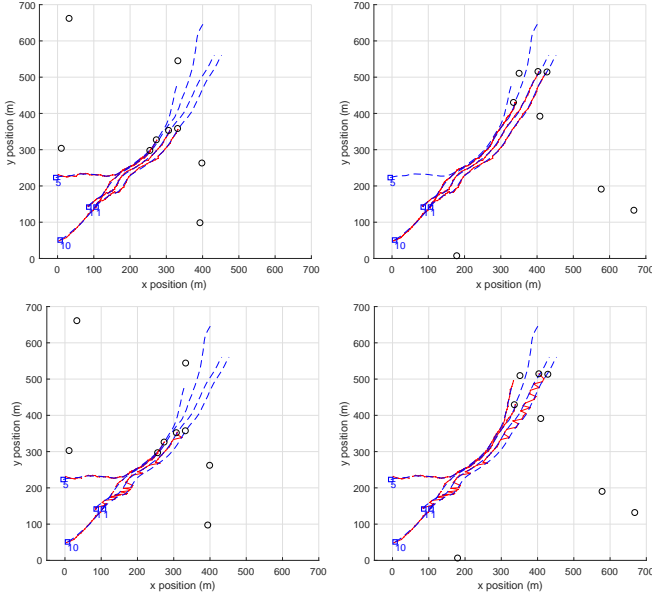
Figure 3: Exemplar outputs of the TPHD filter (top) and tagged PHD filter (bottom) at time steps 50 (left) and 70 (right), shown in a subregion of the surveillance area. The dashed blue lines represent the true trajectories. The blue squares and the numbers next to them denote the starting positions and starting times of different trajectories. There are four trajectories, two start at time step 1, close to each other, and end at time step 79. The other two start at time steps 5 and 10 and end at time steps 69 and 94, respectively. Black circles represent the current measurements. The TPHD filter is able to estimate the alive trajectories at each time step. The tagged PHD filter does not work well, as the estimator selects the highest peaks with distinct tags and there are two targets born from the same PHD component at the same time, which leads to track switching, false and missed targets. At time step 50, the tagged PHD filter estimates four trajectories, with one of length one in the birth location $[85, 140]^T$. The tagged PHD filter at time step 70 estimates three trajectories, though one of them is no longer alive.

$m_{\beta,1}^k = [85, 0, 140, 0]^T$, $m_{\beta,2}^k = [-5, 0, 220, 0]^T$ and $m_{\beta,3}^k = [7, 0, 50, 0]^T$.

We have implemented the L-scan GMTPHD and GMTCPHD filters with $L \in \{1, 2, 5\}$ in a scenario with $N_s = 100$ time steps. We use a pruning threshold $\Gamma_p = 10^{-4}$, absorption threshold $\Gamma_a = 4$ and limit the number of components to 30. An exemplar output of the 1-scan TPHD filter, the tagged PHD filter and the considered ground truth are shown in Figure 3. At each time step, TPHD and TCPHD filters provide an estimate of the set of present trajectories at the current time. The tagged PHD filter shows considerably worse performance as two targets are born at the same time step from the same PHD component. The start and end times of an estimated trajectory do not depend on the choice of L so the output for any other L has the same start time and duration, but with a different error.

In the following, we evaluate the performance of the filters by Monte Carlo simulation with $N_{mc} = 500$ runs. At each time step k, we measure the error between the set $\mathbf{X}_a^k$ of alive trajectories and its estimate $\hat{\mathbf{X}}_a^k$. In order to do so, we use the metric for sets of trajectories based on linear programming in [40] with parameters $p = 2$, $c = 10$ and $\gamma = 1$, which we denote here as $d(\cdot, \cdot)$. In this case, $d^2(\cdot, \cdot)$ can be decomposed into the square costs: $c_m^2(\cdot, \cdot)$ for missed targets,

$c_f^2(\cdot, \cdot)$ for false targets, $c_l^2(\cdot, \cdot)$ for the localisation error of properly detected targets, and $c_t^2(\cdot, \cdot)$ for track switches, as in performance evaluation in traditional multitarget tracking [41, Sec. 13.6]. That is, we have

$$d^2\left(\mathbf{X}_a^k, \hat{\mathbf{X}}_a^k\right) = c_l^2\left(\mathbf{X}_a^k, \hat{\mathbf{X}}_a^k\right) + c_m^2\left(\mathbf{X}_a^k, \hat{\mathbf{X}}_a^k\right)$$
$$+ c_f^2\left(\mathbf{X}_a^k, \hat{\mathbf{X}}_a^k\right) + c_t^2\left(\mathbf{X}_a^k, \hat{\mathbf{X}}_a^k\right). \quad (35)$$

This decomposition is useful to analyse the performances of the filters, as done in this section. In our results, we only use the position elements to compute the error and normalise the error by the considered time window such that the squared error at time k is $d^2\left(\mathbf{X}_a^k, \hat{\mathbf{X}}_a^k\right)/k$. The root mean square (RMS) error at a given time step is calculated as

$$d(k) = \sqrt{\frac{1}{N_{mc}k} \sum_{i=1}^{N_{mc}} d^2\left(\mathbf{X}_a^k, \hat{\mathbf{X}}_{a,i}^k\right)}, \quad (36)$$

where $\hat{\mathbf{X}}_{a,i}^k$ is the estimate of the alive trajectories at time k in the ith Monte Carlo run. In this scenario, the TPHD and TCPHD filters estimates do not have track switches and the resulting errors using $d\left(\mathbf{X}_a^k, \hat{\mathbf{X}}_{a,i}^k\right)$ are the same as the errors computed by the sum of the generalised optimal sub-pattern assignment (GOSPA) metric ($\alpha = 2$) [42] between the true targets and their estimates across all time steps. The tagged PHD and CPHD filters show track switches and therefore, the trajectory metric and GOSPA provide different errors.

The RMS trajectory errors for the algorithms are plotted in Figure 4. As expected, for the TPHD and TCPHD, increasing L improves estimation performance and lowers the error. The TCPHD filter shows lower errors than the TPHD filter, except when targets disappear. The TCPHD filter also estimates the cardinality more accurately than the TPHD filter, except just after a target disappears, as can be seen in Figure 5. The tagged PHD and CPHD filters provide a considerably higher error, as track estimation is done in a manner that does not work well if two targets are born from the same PHD component at the same time.

In order to analyse the results more thoroughly, we make use of the metric decomposition indicated in (35). We compute the RMS costs in (35) at each time step normalised by k, as in (36). The results are shown in Figure 6. Tagged PHD and CPHD filters have a higher cost for missed targets and track switches than TPHD and TCPHD filters, as there are two targets born from the same PHD component and tagging does not work well in this case. We can see that increasing L does not change the costs for missed and false targets and track switches for the trajectory algorithms, it mainly improves the localisation costs. We can see that the main advantage of using the TCPHD filter over the TPHD filter is the reduction of the number of missed targets. Both filters behave quite similarly in terms of localisation costs, missed targets and track switches. We can also see that the TCPHD filter has a higher cost for false targets than the PHD filter at time steps 70, 80 and 95, just after a target disappears.

The running times of our Matlab implementations of the filters in a computer with a 3.5 GHz Intel Xeon E5 processor
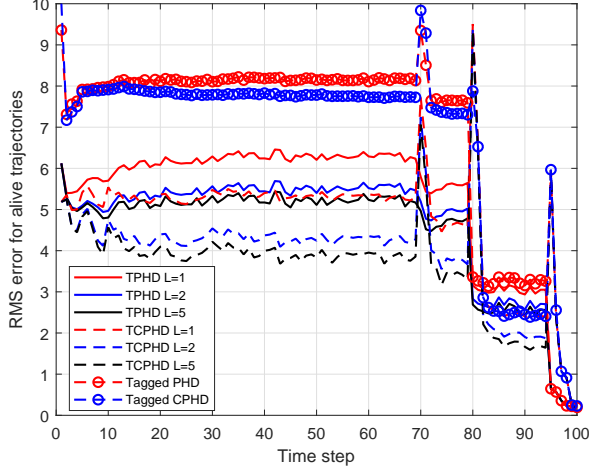
Figure 4: RMS trajectory metric error (36) of the alive trajectories for the TPHD/TCPHD filter and tagged PHD/CPHD filters. Filters based on sets of trajectories have a much higher performance than tagged filters. Increasing L lowers the error.
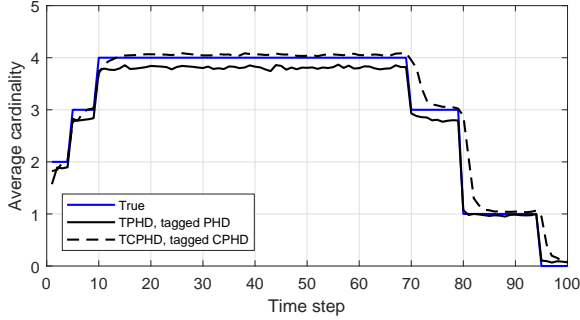

Figure 5: Estimated cardinality against time for the filters. CPHD-based filters estimate the cardinality more accurately except just after a target disappears.

are shown in Table I. The running times of the TCPHD filter are roughly 1 second longer than the TPHD for all values of L. For visualisation clarity, we have only shown performance results for $L \in \{1, 2, 5\}$ in Figures 4 and 6. For any $L \in \{10, 20, 30\}$, performance is similar and slightly better than for $L = 5$. In particular, the RMS error (36) considering all time steps

$$d_T = \sqrt{1/N_s \cdot \sum_{k=1}^{N_s} d^2(k)} \qquad (37)$$

using the trajectory metric is decreased from 4.68 (TPHD) and 3.90 (TCPHD) with $L = 5$ to 4.66 (TPHD) and 3.87 (TCPHD) with $L \in \{10, 20, 30\}$. Tagged filters have a higher computational burden than TPHD and TCPHD with low L, due to the estimation process that links the current target state estimates with the previously estimated trajectories using the tags.

We also show the RMS error (37) for other simulation parameters in Table II. In this table, we also include results where $d(\cdot, \cdot)$ is the sum of the OSPA error [43], [44], with the same p and c, between target states and their estimates
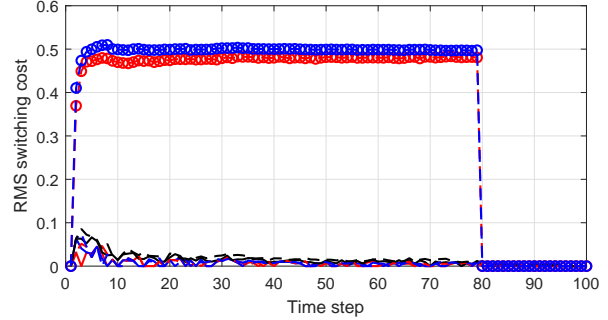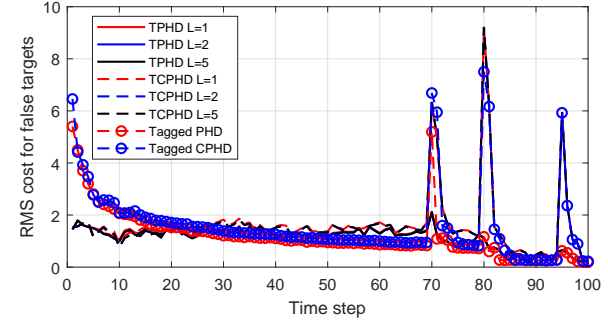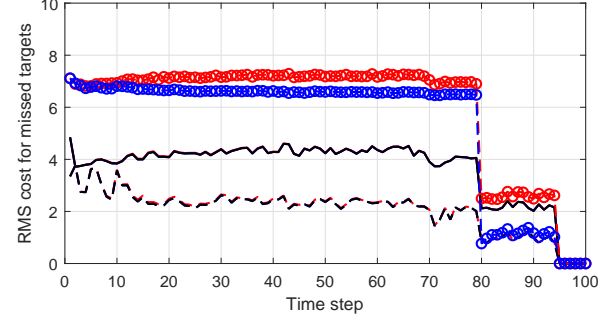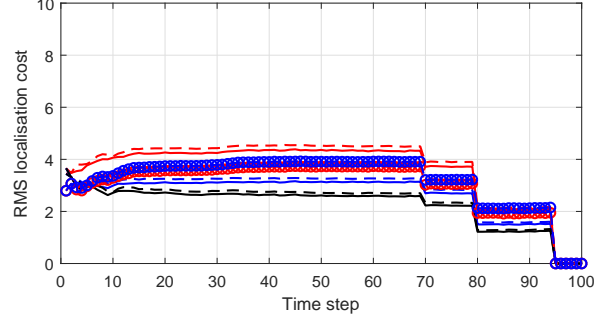

Figure 6: RMS costs for localisation errors for properly detected targets, missed targets, false targets and track switches for the alive trajectories at every time step. The scale of the y axis changes in the figure for the switching cost. The RMS costs for missed targets and false targets do not change with the considered value of L. Increasing L, decreases the localisation cost for trajectory-based algorithms. TCPHD misses fewer targets than TPHD. Switching cost is negligible for TPHD/TCPHD filters. The tagged PHD and CPHD filters have a high cost for missed targets and show track switching.

Table I: Running times of the algorithms in seconds

L	1	2	5	10	20	30
TPHD	1.1	1.1	1.2	1.7	3.4	6.0
TCPHD	2.0	2.0	2.1	2.6	4.3	6.9
Tagged PHD			2.2			
Tagged CPHD			3.2			

at each time step. The trajectory metric and GOSPA return the same errors for TPHD/TCPHD filters, for all simulation parameters considered, which implies that the cost of track switches is negligible. Tagged filters show track switches so the trajectory metric and GOSPA do not coincide. GOSPA errors are not shown due to space constraints, but they are generally 0.02 lower than the trajectory metric error for the tagged filters. According to all the metrics and all scenarios, the TCPHD filter is the best performing filter, followed by the TPHD filter. For both filters, all metrics and scenarios, the error decreases as L increases. The errors of the tagged filters are, in general, considerably higher than for the set of trajectories filters. As expected, performance improves for all filters if the probability of detection increases, clutter intensity decreases and the measurement noise variance decreases. We have also shown results with a different probability of survival and birth intensity to show that the results are consistent with different simulation parameters.

VIII. Conclusions

In this paper we have developed the trajectory PHD and CPHD filters based on KLD minimisations and sets of trajectories. The TPHD and TCPHD filters propagate a Poisson multitrajectory density and an IID cluster multitrajectory density through the filtering recursion to make inference on the set of alive trajectories. The theory presented in this paper endows the PHD/CPHD filters with the capability of estimating trajectories from first principles, which can span their already widespread use to more applications where tracks are required.

We have also proposed a Gaussian mixture implementation of the filters. In particular, the parameter L of the L-scan versions governs the accuracy of the estimation of past states of the trajectories. We have analysed the proposed filters in terms of localisation errors, missed targets, false targets and track switches, and also based on OSPA. Increasing L mainly improves the localisation error of past states of the trajectory for both filters. In general, the TCPHD filter outperforms the TPHD filter, and both provide better trajectory estimates than the ones provided by tagging the PHD and CPHD filters.

References

[1] R. P. S. Mahler, *Advances in Statistical Multisource-Multitarget Information Fusion*. Artech House, 2014.

[2] K. Granström, C. Lundquist, and O. Orguner, "Extended target tracking using a Gaussian-mixture PHD filter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 48, no. 4, pp. 3268–3286, Oct. 2012.

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

[4] C. Lundquist, K. Granström, and U. Orguner, "An extended target CPHD filter and a gamma Gaussian inverse Wishart implementation," *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 3, pp. 472–483, June 2013.

[5] B.-T. Vo, B.-N. Vo, and A. Cantoni, "Analytic implementations of the cardinalized probability hypothesis density filter," *IEEE Transactions on Signal Processing*, vol. 55, no. 7, pp. 3553–3567, July 2007.

[6] N. Whiteley, S. Singh, and S. Godsill, "Auxiliary particle implementation of probability hypothesis density filter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 46, no. 3, pp. 1437–1454, July 2010.

[7] O. Erdinc, P. Willett, and Y. Bar-Shalom, "The bin-occupancy filter and its connection to the PHD filters," *IEEE Transactions on Signal Processing*, vol. 57, no. 11, pp. 4232–4246, Nov. 2009.

[8] M. Üney, D. E. Clark, and S. J. Julier, "Distributed fusion of PHD filters via exponential mixture densities," *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 3, pp. 521–531, June 2013.

[9] G. Battistelli, L. Chisci, C. Fantacci, A. Farina, and A. Graziano, "Consensus CPHD filter for distributed multitarget tracking," *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 3, pp. 508–520, June 2013.

[10] C. S. Lee, S. Nagappa, N. Palomeras, D. Clark, and J. Salvi, "SLAM with SC-PHD filters: An underwater vehicle application," *IEEE Robotics & Automation Magazine*, vol. 21, no. 2, pp. 38–45, June 2014.

[11] J. Mullane, B.-N. Vo, M. D. Adams, and B.-T. Vo, "A random-finite-set approach to Bayesian SLAM," *IEEE Transactions on Robotics*, vol. 27, no. 2, pp. 268–282, April 2011.

[12] E. Maggio, M. Taj, and A. Cavallaro, "Efficient multitarget visual tracking using random finite sets," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 18, no. 8, pp. 1016–1027, Aug. 2008.

[13] T. M. Wood, C. A. Yates, D. A. Wilkinson, and G. Rosser, "Simplified multitarget tracking using the PHD filter for microscopic video data," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 5, pp. 702–713, May 2012.

[14] C. Lundquist, L. Hammarstrand, and F. Gustafsson, "Road intensity based mapping using radar measurements with a probability hypothesis density filter," *IEEE Transactions on Signal Processing*, vol. 59, no. 4, pp. 1397–1408, April 2011.

[15] B. Ristic, B.-N. Vo, and D. Clark, "A note on the reward function for PHD filters with sensor control," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 47, no. 2, pp. 1521–1529, April 2011.

[16] A. F. García-Fernández and B.-N. Vo, "Derivation of the PHD and CPHD filters based on direct Kullback-Leibler divergence minimization," *IEEE Transactions on Signal Processing*, vol. 63, no. 21, pp. 5812–5820, Nov. 2015.

[17] D. Franken, M. Schmidt, and M. Ulmke, ""Spooky action at a distance" in the cardinalized probability hypothesis density filter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 45, no. 4, pp. 1657–1664, Oct. 2009.

[18] N. Nadarajah, T. Kirubarajan, T. Lang, M. McDonald, and K. Punithakumar, "Multitarget tracking using probability hypothesis density smoothing," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 47, no. 4, pp. 2344–2360, Oct. 2011.

[19] S. Nagappa, E. D. Delande, D. E. Clark, and J. Houssineau, "A tractable forward-backward CPHD smoother," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 53, no. 1, pp. 201–217, Feb. 2017.

[20] L. Lin, Y. Bar-Shalom, and T. Kirubarajan, "Track labeling and PHD filter for multitarget tracking," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 42, no. 3, pp. 778–795, July 2006.

[21] K. Panta, B.-N. Vo, and S. Singh, "Novel data association schemes for the probability hypothesis density filter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 43, no. 2, pp. 556–570, April 2007.

[22] K. Panta, D. Clark, and B.-N. Vo, "Data association and track management for the Gaussian mixture probability hypothesis density filter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 45, no. 3, pp. 1003–1016, July 2009.

[23] E. Pollard, B. Pannetier, and M. Rombaut, "Hybrid algorithms for multitarget tracking using MHT and GM-CPHD," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 47, no. 2, pp. 832–847, April 2011.

[24] Z. Lu, W. Hu, Y. Liu, and T. Kirubarajan, "A new cardinalized probability hypothesis density filter with efficient track continuity and extraction," in *21st International Conference on Information Fusion*, 2018.

[25] Z. Lu, W. Hu, and T. Kirubarajan, "Labeled random finite sets with moment approximation," *IEEE Transactions on Signal Processing*, vol. 65, no. 13, pp. 3384–3398, July 2017.

[26] A. F. García-Fernández, J. Grajal, and M. R. Morelande, "Two-layer particle filter for multiple target detection and tracking," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 49, no. 3, pp. 1569–1588, July 2013.

[27] B. T. Vo and B. N. Vo, "Labeled random finite sets and multi-object conjugate priors," *IEEE Transactions on Signal Processing*, vol. 61, no. 13, pp. 3460–3475, July 2013.

[28] A. F. García-Fernández, L. Svensson, and M. R. Morelande, "Multiple target tracking based on sets of trajectories," *accepted for publication in IEEE Transactions on Aerospace and Electronic Systems*, 2019. [Online]. Available: https://arxiv.org/abs/1605.08163

Table II: Error in alive trajectories averaged over all time steps for the trajectory metric (TM) and OSPA

	TPHD						TCPHD						Tagged PHD		Tagged CPHD	
	TM/GOSPA			OSPA			TM/GOSPA			OSPA			TM	OSPA	TM	OSPA
L	1	2	5	1	2	5	1	2	5	1	2	5	-	-	-	-
No change	5.54	4.89	4.68	3.94	3.64	3.55	4.98	4.17	3.90	3.39	3.01	2.89	7.31	5.55	7.12	5.34
$\sigma^2 = 16$	9.31	8.50	8.04	5.83	5.40	5.17	8.27	8.39	7.88	5.61	5.14	4.87	8.95	6.37	8.94	6.29
$\sigma^2 = 1$	4.48	4.18	4.13	3.43	3.30	3.28	3.71	3.30	3.23	2.77	2.59	2.56	6.81	5.33	6.55	5.09
$\lambda_c = 70$	5.61	4.98	4.78	4.01	3.72	3.63	5.09	4.33	4.07	3.50	3.15	3.03	7.27	5.53	7.05	5.30
$\lambda_c = 30$	5.49	4.84	4.63	3.90	3.60	3.51	4.89	4.08	3.81	3.31	2.93	2.81	7.28	5.52	7.06	5.29
$p_D = 0.99$	4.15	3.34	3.09	2.63	2.21	2.09	4.07	3.23	2.97	2.57	2.14	2.01	7.17	5.36	7.20	5.35
$p_D = 0.85$	5.88	5.19	4.93	4.15	3.84	3.73	5.48	4.68	4.39	3.80	3.43	3.30	7.28	5.54	7.15	5.38
$p_D = 0.75$	6.77	6.07	5.76	4.84	4.53	4.41	6.40	5.59	5.23	4.47	4.11	3.95	7.59	5.81	7.35	5.55
$p_S = 0.95$	5.57	4.93	4.72	3.97	3.68	3.59	5.08	4.31	4.06	3.49	3.12	3.01	7.35	5.60	7.16	5.38
$w_{\beta,j}^k = 0.05$	5.56	4.91	4.70	3.97	3.68	3.59	5.00	4.21	3.94	3.45	3.08	2.96	7.33	5.58	7.15	5.40

[29] A. F. García-Fernández and L. Svensson, "Trajectory probability hypothesis density filter," in *21st International Conference on Information Fusion*, 2018.

[30] L. Svensson and M. Morelande, "Target tracking based on estimation of sets of trajectories," in *17th International Conference on Information Fusion*, 2014.

[31] R. P. S. Mahler, "Multitarget Bayes filtering via first-order multitarget moments," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 39, no. 4, pp. 1152–1178, Oct. 2003.

[32] J. L. Williams, "An efficient, variational approximation of the best fitting multi-Bernoulli filter," *IEEE Transactions on Signal Processing*, vol. 63, no. 1, pp. 258–273, Jan. 2015.

[33] A. Papoulis and S. U. Pillai, *Probability, Random Variables and Stochastic Processes*. McGraw-Hill, 2002.

[34] K. Granström, L. Svensson, Y. Xia, J. L. Williams, and A. F. García-Fernández, "Poisson multi-Bernoulli mixture trackers: continuity through random finite sets of trajectories," in *21st International Conference on Information Fusion*, 2018.

[35] S. Särkkä, *Bayesian Filtering and Smoothing*. Cambridge University Press, 2013.

[36] D. Salmond, "Mixture reduction algorithms for point and extended object tracking in clutter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 45, no. 2, pp. 667–686, April 2009.

[37] W. Koch and F. Govaers, "On accumulated state densities with applications to out-of-sequence measurement processing," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 47, no. 4, pp. 2766–2778, 2011.

[38] H. Oruç and H. K. Akmaz, "Symmetric functions and the Vandermonde matrix," *Journal of Computational and Applied Mathematics*, vol. 172, no. 1, pp. 49–64, Nov. 2004.

[39] A. F. García-Fernández, L. Svensson, M. R. Morelande, and S. Särkkä, "Posterior linearization filter: principles and implementation using sigma points," *IEEE Transactions on Signal Processing*, vol. 63, no. 20, pp. 5561–5573, Oct. 2015.

[40] A. S. Rahmathullah, A. F. García-Fernández, and L. Svensson, "A metric on the space of finite sets of trajectories for evaluation of multi-target tracking algorithms," 2016. [Online]. Available: http://arxiv.org/abs/1605.01177

[41] S. Blackman and R. Popoli, *Design and Analysis of Modern Tracking Systems*. Artech House, 1999.

[42] A. S. Rahmathullah, A. F. García-Fernández, and L. Svensson, "Generalized optimal sub-pattern assignment metric," in *20th International Conference on Information Fusion*, 2017.

[43] D. Schuhmacher and A. Xia, "A new metric between distributions of point processes," vol. 40, no. 3, pp. 651–672, Sep. 2008.

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

[45] R. P. S. Mahler, *Statistical Multisource-Multitarget Information Fusion*. Artech House, 2007.

[46] D. Stoyan, W. S. Kendall, and J. Mecke, *Stochastic geometry and its applications*. John Wiley & Sons, 1995.

Ángel F. García-Fernández received the telecommunication engineering degree (with honours) and the Ph.D. degree from Universidad Politécnica de Madrid, Madrid, Spain, in 2007 and 2011, respectively.

He is currently a Lecturer in the Department of Electrical Engineering and Electronics at the University of Liverpool, Liverpool, UK. He previously held postdoctoral positions at Universidad Politécnica de Madrid, Chalmers University of Technology, Gothenburg, Sweden, Curtin University, Perth, Australia, and Aalto University, Espoo, Finland. His main research activities and interests are in the area of Bayesian nonlinear inference, with emphasis on dynamic systems and multiple target tracking. He has received paper awards at the International Conference on Information Fusion in 2017 and 2019.

Lennart Svensson was born in Älvängen, Sweden in 1976. He received the M.S. degree in electrical engineering in 1999 and the Ph.D. degree in 2004, both from Chalmers University of Technology, Gothenburg, Sweden.

He is currently Professor of Signal Processing, again at Chalmers University of Technology. His main research interests include machine learning and Bayesian inference in general, and nonlinear filtering, deep learning and tracking in particular. He has organised a massive open online course on multiple object tracking, available on edX and YouTube, and received paper awards at the International Conference on Information Fusion in 2009, 2010, 2017 and 2019.

Supplementary material of "Trajectory PHD and CPHD filters"

APPENDIX A

In this appendix, we indicate how to draw samples from an IID cluster trajectory RFS, which includes the Poisson trajectory RFS as a particular case. Given a single trajectory density $\breve{\nu}(\cdot)$, the probability that the trajectory starts at time t and has duration i is

$$P_{\breve{\nu}}(t,i) = \int \breve{\nu}\left(t, x^{1:i}\right) dx^{1:i}. \tag{38}$$

Given the start time t and duration i, the density of the target states of the trajectory is

$$\breve{\nu}\left(x^{1:i}|t,i\right) = \frac{\breve{\nu}\left(t, x^{1:i}\right)}{P_{\breve{\nu}}(t,i)}. \tag{39}$$

Therefore, we can draw samples from an IID cluster trajectory RFS using (38) and (39), as indicated in Algorithm 3.

Algorithm 3 Sampling from an IID cluster multitrajectory density

---

Input: IID cluster multitrajectory density $\nu(\cdot)$.
Output: Sample $X \sim \nu(\cdot)$.
 - Set $X = \emptyset$ and sample $n \sim \rho_\nu(\cdot)$.
 for $j = 1$ to n **do**
   - Sample $(t,i) \sim P_{\breve{\nu}}(\cdot)$ and $x^{1:i} \sim \breve{\nu}(\cdot|t,i)$, see (38) and (39).
   - Set $X \leftarrow X \cup \left\{\left(t, x^{1:i}\right)\right\}$.
 end for

APPENDIX B

In this appendix we prove Theorem 4. A multitrajectory density $\pi(\cdot)$ can be written as [31]

$$\pi\left(\{X_1, ..., X_n\}\right) = \rho_\pi(n)\, n!\, \pi_n\left(X_1, ..., X_n\right) \tag{40}$$

where $\pi_n(\cdot)$ is a permutation invariant ordered trajectory density such that

$$\int \pi_n\left(X_1, ..., X_n\right) dX_{1:n} = 1.$$

The marginal density of this density is written as

$$\tilde{\pi}_n(X) = \int \pi_n\left(X, X_2..., X_n\right) dX_{2:n}$$
$$= \frac{1}{\rho_\pi(n)\, n!} \int \pi\left(\{X, X_2, ..., X_n\}\right) dX_{2:n}.$$

Substituting (11) and (40) into (13), we get

$$\mathrm{D}\left(\pi\,\|\,\nu\right) = \sum_{n=0}^{\infty} \rho_\pi(n) \log \frac{\rho_\pi(n)}{\rho_\nu(n)}$$
$$+ \sum_{n=0}^{\infty} \rho_\pi(n) \int \pi_n\left(X_1, ..., X_n\right)$$
$$\times \log \frac{\pi_n\left(X_1, ..., X_n\right)}{\prod_{j=1}^{n} \breve{\nu}(X_j)} dX_{1:n}. \tag{41}$$

The objective is to find $\rho_\nu(\cdot)$ and $\breve{\nu}(\cdot)$ that minimise $\mathrm{D}\left(\pi\,\|\,\nu\right)$. From KLD minimisation over discrete variables, $\rho_\nu(\cdot) =$

$\rho_\pi(\cdot)$ minimises the KLD. Minimizing $\mathrm{D}\left(\pi\,\|\,\nu\right)$ w.r.t. $\breve{\nu}(\cdot)$ is equivalent to minimising the functional

$$L[\breve{\nu}] = -\int \sum_{n=0}^{\infty} \rho_\pi(n)\, n \tilde{\pi}_n(X) \log \breve{\nu}(X)\, dX.$$

which is minimised by [29]

$$\breve{\nu}(X) = \frac{D_\pi(X)}{\sum_{n=0}^{\infty} \rho_\pi(n)\, n},$$

or equivalently, $D_\nu(\cdot) = D_\pi(\cdot)$.

APPENDIX C

In this appendix, we prove (14). The IID cluster multitrajectory density $\nu^k(\cdot)$ that minimises the KLD has the same cardinality distribution and PHD as $\pi^k(\cdot)$. The cardinality distribution of $\pi^k(\cdot)$ is such that $\rho_{\pi^k}(m) = 1$, so we compute its PHD. The posterior can be written as [45, Eq. (12.91)]

$$\pi^k\left(\{X_1, ..., X_m\}\right) = \sum_{\sigma \in \Xi_m} \left[\prod_{i=1}^{m_a} \breve{a}_i\left(X_{\sigma_i}\right)\right] \left[\prod_{i=1}^{m_d} \breve{d}_i\left(X_{\sigma_{i+m_a}}\right)\right] \tag{42}$$

where Ξ_m is the set that contains all permutations of $(1, ..., m)$ and σ_i is the ith component of σ.

Substituting (42) into (4), the PHD of the posterior is

$$D_{\pi^k}(X) = \sum_{n=0}^{\infty} \frac{1}{n!} \int \pi^k\left(\{X, X_1, ..., X_n\}\right) dX_{1:n}$$
$$= \sum_{i=1}^{m_a} \breve{a}_i(X) + \sum_{i=1}^{m_d} \breve{d}_i(X),$$

which is equal to $D_{\nu^k}(X)$. Then, using (12), we have that the single trajectory density $\breve{\nu}^k(\cdot)$ is

$$\breve{\nu}^k(X) = \frac{D_{\nu^k}(X)}{m}$$
$$= \frac{\sum_{i=1}^{m_a} \breve{a}_i(X) + \sum_{i=1}^{m_d} \breve{d}_i(X)}{m}.$$

A single trajectory with this density exists at the current time k with probability m_a/m. As we have m IID trajectories with this distribution, the number of alive trajectories follows the binomial distribution in (14).

APPENDIX D

In this appendix, we prove Theorem 9 (TCPHD update). The proof is quite similar to the CPHD update [16].

A. Density of the measurement

As the current set of targets is an IID cluster with cardinality $\rho_{\omega^k}(\cdot)$ and PHD $D_{\omega_T^k}(\cdot)$, the density of the measurements, under Assumptions U1, U4 and U5, is the same as in the CPHD filter [16, Eq. (32)].

$$\ell^k\left(\mathbf{z}^k\right) = \left\langle \rho_{\omega^k}, \Upsilon^0\left[D_{\omega_T^k}, \mathbf{z}^k\right]\right\rangle \prod_{z \in \mathbf{z}^k} \breve{c}(z). \tag{43}$$

B. Cardinality of the posterior

As the set of trajectories present at the current time has the same cardinality as the set of targets at the current time, the updated cardinality distribution is the same in both cases so the TCPHD cardinality update is the same as the CPHD filter cardinality update, which is given in (22).

C. PHD of the posterior

The PHD of the posterior is

$$
\begin{aligned}
D_{\pi^k}(X) \\
= \int \pi^k \left(\{X\} \cup \mathbf{X} \right) \delta \mathbf{X} \\
= \frac{1}{\ell^k \left(\mathbf{z}^k \right)} \int \ell^k \left(\mathbf{z}^k \left| \tau^k \left(\{X\} \cup \mathbf{X} \right) \right) \omega^k \left(\{X\} \cup \mathbf{X} \right) \delta \mathbf{X} \right. \\
= \frac{1}{\ell^k \left(\mathbf{z}^k \right)} \sum_{n=0}^{\infty} (n+1) \int \ell^k \left(\mathbf{z}^k | \tau^k \left(\{X, X_1, ..., X_n\} \right) \right) \\
\times \rho_{\omega^k} (n+1) \breve{\omega}^k (X) \left[\prod_{j=1}^{n} \breve{\omega}^k (X_j) \right] dX_{1:n}.
\end{aligned}
$$

We apply the decomposition [16, Eq. (14)]

$$
\begin{aligned}
\ell^k \left(\mathbf{z}^k \left| \tau^k \left(\{X\} \right) \cup \tau^k \left(\mathbf{X} \right) \right) \right. \\
= \left(1 - p_D \left(\tau^k \left(\{X\} \right) \right) \right) \ell^k \left(\mathbf{z}^k \left| \tau^k \left(\mathbf{X} \right) \right) \right. \\
+ p_D \left(\tau^k \left(\{X\} \right) \right) \sum_{z \in \mathbf{z}^k} l \left(z | \tau^k \left(\{X\} \right) \right) \ell^k \left(\mathbf{z}^k \setminus \{z\} \left| \tau^k \left(\mathbf{X} \right) \right) \right.
\end{aligned}
$$

and obtain $D_{\pi^k}(t, x^{1:i})$

$$
\begin{aligned}
= \frac{\left(1 - p_D \left(x^i \right) \right) \breve{\omega}^k \left(t, x^{1:i} \right)}{\ell^k \left(\mathbf{z}^k \right)} \sum_{n=0}^{\infty} (n+1) \rho_{\omega^k} (n+1) \\
\times \int \ell^k \left(\mathbf{z}^k | \tau^k \left(\{X_1, ..., X_n\} \right) \right) \left[\prod_{j=1}^{n} \breve{\omega}^k (X_j) \right] dX_{1:n} \\
+ \frac{p_D \left(x^i \right) \breve{\omega}^k \left(t, x^{1:i} \right)}{\ell^k \left(\mathbf{z}^k \right)} \sum_{z \in \mathbf{z}^k} l \left(z | x^i \right) \sum_{n=0}^{\infty} (n+1) \rho_{\omega^k} (n+1) \\
\times \int \ell^k \left(\mathbf{z}^k \setminus \{z\} \left| \tau^k \left(\{X_1, ..., X_n\} \right) \right) \left[\prod_{j=1}^{n} \breve{\omega}^k (X_j) \right] dX_{1:n} \right. \\
= \frac{\left(1 - p_D \left(x^i \right) \right) \breve{\omega}^k \left(t, x^{1:i} \right)}{\ell^k \left(\mathbf{z}^k \right)} \sum_{n=0}^{\infty} (n+1) \rho_{\omega^k} (n+1) \\
\times \int \ell^k \left(\mathbf{z}^k | \{x_1, ..., x_n\} \right) \left[\prod_{j=1}^{n} \breve{\omega}_\tau^k (x_j) \right] dx_{1:n} \\
+ \frac{p_D \left(x^i \right) \breve{\omega}^k \left(t, x^{1:i} \right)}{\ell^k \left(\mathbf{z}^k \right)} \sum_{z \in \mathbf{z}^k} l \left(z | x^i \right) \sum_{n=0}^{\infty} (n+1) \rho_{\omega^k} (n+1) \\
\times \int \ell^k \left(\mathbf{z}^k \setminus \{z\} | \{x_1, ..., x_n\} \right) \left[\prod_{j=1}^{n} \breve{\omega}_\tau^k (x_j) \right] dx_{1:n}.
\end{aligned}
$$
(44)

From [16, Eq. (39)], we know that

$$
\begin{aligned}
\sum_{n=0}^{\infty} (n+1) \rho_{\omega^k} (n+1) \\
\times \int \ell^k \left(\mathbf{z}^k | \{x_1, ..., x_n\} \right) \left[\prod_{j=1}^{n} \breve{\omega}_\tau^k (x_j) \right] dx_{1:n} \\
= \int D_{\omega_\tau^k}(x) dx \left\langle \rho_{\omega^k}, \Upsilon^1 \left[D_{\omega_\tau^k}, \mathbf{z}^k \right] \right\rangle \prod_{z \in \mathbf{z}^k} \breve{c}(z).
\end{aligned}
$$
(45)

Substituting (43) and (45) into the first and second term of (44) completes the proof of (23).

APPENDIX E

In this appendix, we prove Theorem 8 (TCPHD prediction). We proceed to calculate the cardinality and PHD of the predicted multitrajectory density. As in the CPHD filter, the assumptions of the TCPHD filter prediction (P1, P4, P5) are similar to the TCPHD filter update (U1, U4, U5) so the proof of the TCPHD filter update, which is given in Appendix D, is useful for the TCPHD filter prediction, as in the CPHD filter [16].

A. Cardinality

The cardinality distribution of the predicted multitrajectory density for alive targets is analogous to the cardinality distribution $\rho_{\ell^k} (\cdot)$ of the measurement, which can be calculated from (43) in the same way as in [16, Sec. V.C]. The result is

$$
\begin{aligned}
\rho_{\ell^k} (m) = \sum_{j=0}^{m} \rho_c (m-j) \sum_{n=j}^{\infty} \binom{n}{j} \rho_{\omega^k} (n) \\
\times \frac{\left[\int \left(1 - p_D(x) \right) D_{\omega_\tau^k}(x) dx \right]^{n-j}}{\left[\int D_{\omega_\tau^k}(x) dx \right]^n} \\
\times \left[\int p_D(x) D_{\omega_\tau^k}(x) dx \right]^j.
\end{aligned}
$$

By changing $\rho_c (\cdot)$, $p_D (\cdot)$ and $\omega^k (\cdot)$ for $\rho_{\beta^k} (\cdot)$, $p_S (\cdot)$ and $\pi^k (\cdot)$, respectively, we obtain (21).

B. PHD

The result for the PHD of the TCPHD prediction, which is the same as for the TPHD prediction, see (17), can be established directly based on thinning, displacement and superposition of point processes, as in the CPHD filter for targets [16]. Nevertheless, for completeness, we provide more details regarding its calculation here.

We have a multitrajectory density $\pi^{k-1} (\cdot)$ with a PHD $D_{\pi_{k-1}} (\cdot)$ and we want to compute the PHD of the multitrajectory density $\omega^k (\cdot)$, which is $D_{\omega^k} (\cdot)$. We first compute the PHD of $\xi^k (\cdot)$, which represents the multitrajectory density of the surviving trajectories and is given by

$$
\xi^k (\mathbf{X}) = \int f_s (\mathbf{X} | \mathbf{Y}) \pi^{k-1} (\mathbf{Y}) \delta \mathbf{Y}
$$
(46)

where $f_s (\cdot | \cdot)$ is the transition multitrajectory density for surviving trajectories. We proceed to explain a decomposition of $f_s (\cdot | \cdot)$, which will be useful for the proof.

We first write the probability of survival and single trajectory transition density for alive trajectories at time k as a function of trajectories. These correspond to

$$p'_S \left(t, x^{1:i} \right) = p_S \left(x^i \right) \tag{47}$$

$$g' \left(t, x^{1:i} | t', z^{1:i'} \right) = \delta \left[t - t' \right] g \left(x^i | z^{i'} \right) \delta \left(x^{1:i-1} - z^{1:i'} \right) \tag{48}$$

where $t + i - 1 = k$ and $t' + i' - 1 = k - 1$, otherwise, $p'_S \left(\cdot \right)$ and $g' \left(\cdot | \cdot \right)$ are zero. Also, $\delta \left[\cdot \right]$ and $\delta \left(\cdot \right)$ denote Kronecker and Dirac delta, respectively.

We use the following decomposition for the transition density of surviving trajectories

$$f_s \left(\{X\} \cup \mathbf{X} | \mathbf{Y} \right) = \sum_{Y \in \mathbf{Y}} g' \left(X | Y \right) p'_S \left(Y \right) f_s \left(\mathbf{X} | \mathbf{Y} \setminus \{Y\} \right).$$

Note that this decomposition results from the fact that given the surviving trajectories $\{X\} \cup \mathbf{X}$, trajectory X must have survived from a trajectory Y that belongs to $\mathbf{Y}$ and the rest of the surviving trajectories $\mathbf{X}$ have survived from the rest of the trajectories at the previous time step $\mathbf{Y} \setminus \{Y\}$.

The PHD of the surviving trajectories is obtained by substituting (46) into (4)

$$D_{\xi^k} \left(X \right) = \int \xi^k \left(\mathbf{X} \cup \{X\} \right) \delta \mathbf{X}$$

$$= \int \left[\int f_s \left(\{X\} \cup \mathbf{X} | \mathbf{Y} \right) \delta \mathbf{X} \right] \pi^{k-1} \left(\mathbf{Y} \right) \delta \mathbf{Y}. \tag{49}$$

We simplify the inner integral in (49) as

$$\int f_s \left(\{X\} \cup \mathbf{X} | \mathbf{Y} \right) \delta \mathbf{X}$$

$$= \int \sum_{Y \in \mathbf{Y}} g' \left(X | Y \right) p'_S \left(Y \right) f_s \left(\mathbf{X} | \mathbf{Y} \setminus \{Y\} \right) \delta \mathbf{X}$$

$$= \sum_{Y \in \mathbf{Y}} g' \left(X | Y \right) p'_S \left(Y \right).$$

Substituting this equation into (49) and applying Campbell's theorem [1, Sec. 4.2.12]

$$D_{\xi^k} \left(X \right) = \int \left(\sum_{Y \in \mathbf{Y}} g' \left(X | Y \right) p'_S \left(Y \right) \right) \pi^{k-1} \left(\mathbf{Y} \right) \delta \mathbf{Y}$$

$$= \int g' \left(X | Y \right) p'_S \left(Y \right) D_{\pi^{k-1}} \left(Y \right) dY. \tag{50}$$

Substituting (47) and (48) into (50) yields $D_{\xi^k} \left(\cdot \right)$ in (17). By the superposition theorem of point processes [46], the PHD of $\omega^k \left(\cdot \right)$ is the sum of the PHDs of the surviving trajectories, which is given by $D_{\xi^k} \left(\cdot \right)$, and the PHD of the new born trajectories, which is given by $D_{\beta^k} \left(\cdot \right)$. Therefore, the TCPHD prediction equation for the PHD is the same as for the TPHD filter, see (17).