Multitarget Tracking

Ba-Ngu Vo, Mahendra Mallick, Yaakov Bar-Shalom, Stefano Coraluppi, Richard Osborne, III, Ronald Mahler, and Ba-Tuong Vo

***Abstract*—Multitarget tracking (MTT) refers to the problem of jointly estimating the number of targets and their states or trajectories from noisy sensor measurements. MTT has a long history spanning over 50 years, with a plethora of applications in many fields of study. While numerous techniques have been developed, the three most widely used approaches to MTT are the joint probabilistic data association filter (JPDAF), multiple hypothesis tracking (MHT), and random finite set (RFS). The JPDAF and MHT have been widely used for more than two decades, while the random finite set (RFS) based MTT algorithms have received a great deal of attention during the last decade. In this article, we provide an overview of MTT and succinct summaries of popular state-of-the-art MTT algorithms.**

List of mathematical symbols:

$\delta(\cdot)$	Dirac delta
$\delta_n[m]$	Kronecker delta: $\delta_n[m] = \begin{cases} 1, \text{ if } m = n \\ 0, \text{ otherwise} \end{cases}$
$\emptyset$	empty set
$1_A(\cdot)$	indicator function on the set A
$\langle \cdot, \cdot \rangle$	inner product between functions/sequences
$\lvert X \rvert$	cardinality (number of elements) of the set X
C_j^n	number of j-combinations of n: $\frac{n!}{j!(n-j)!}$
P_j^n	number of j-permutations of n: $\frac{n!}{(n-j)!}$
$\mathcal{X}$	state space
$\mathbf{x}$	state vector
$\mathcal{F}(\mathcal{X})$	collection of finite subsets of $\mathcal{X}$
$\mathcal{Z}$	observation space
$\mathbf{z}$	observation vector
$\mathbf{z}_{l:k}$	$(\mathbf{z}_l, \mathbf{z}_{l+1}, ..., \mathbf{z}_k)$
$\mathbf{z}^k$	observation history $\mathbf{z}_{1:k}$
$\Pr(A)$	probability of event A
$f_{k\lvert k-1}(\cdot\lvert\cdot)$	state transition density
$g_k(\cdot\lvert\cdot)$	likelihood function
$p_{0:k}(\cdot\lvert\cdot)$	posterior density
$p_{k\lvert k-1}(\cdot\lvert\cdot)$	prediction density
$p_k(\cdot\lvert\cdot)$	filtering density
$\mathcal{N}(\cdot; \mathbf{m}, \mathbf{P})$	Gaussian probability density with mean $\mathbf{m}$ and covariance $\mathbf{P}$
$\mathbf{F}_{k,k-1}$	state transition matrix
$\mathbf{H}_k$	measurement matrix
$(\cdot)'$	transpose of a vector/matrix
$\mathbf{1}$	column vector of ones $[1, 1, ..., 1]'$
$P_{S,k\lvert k-1}(\mathbf{x})$	probability of survival of a target at time k given its previous state $\mathbf{x}$
$P_{D,k}(\mathbf{x})$	probability of target detection at time k given its state $\mathbf{x}$

B.-N. Vo and B.-T. Vo are with Curtin University, Bentley, WA, Australia.
M. Mallick is with Vellore Institute of Technology, Vellore, India.
Y. Bar-Shalom and R. Osborne, III are with University of Connecticut, Storrs, CT, USA.
S. Coraluppi is with STR, Woburn, MA, USA.
R. Mahler is an Independent Consultant, Minnesota, USA.

I. Introduction

In a multitarget scenario the number of targets and their trajectories vary with time due to targets appearing and disappearing. For example, the location, velocity and bearing of commercial planes at an airport, ships in a harbour, or pedestrians on the street. multitarget tracking (MTT) refers to the problem of jointly estimating the number of targets and their trajectories from sensor data. Driven by aerospace applications in the 1960's, MTT has a long history spanning over 50 years. During the last decade, advances in MTT techniques, along with sensing and computing technologies, have opened up numerous research venues as well as application areas. Today, MTT has found applications in diverse disciplines, including, air traffic control, surveillance, defence, space applications, oceanography, autonomous vehicles and robotics, remote sensing, computer vision, and biomedical research, see for example the texts [10], [15], [22], [73], [86], [96], [99], [144]. The goal of this article is to discuss the challenges in MTT and present the state-of-the-art techniques.

In this article we only consider the standard setting where sensor measurements at each instance have been preprocessed into a set of points or detections. The multitarget tracker receives a random number of measurements due to detection uncertainty and false alarms (FAs). Consequently, apart from process and measurement noises, the multitarget tracker has to contend with much more complex sources of uncertainty, such as measurement origin uncertainty, false alarm, missed detection, and births and deaths of targets. Moreover, in the multi-sensor setting, a multitarget tracker needs to process measurements from multiple heterogeneous sensors such as radar, sonar, electro-optical, infrared, camera, unattended ground sensor etc.

A number of MTT algorithms are used at present in various tracking applications, with the most popular being the joint probabilistic data association filter (JPDAF) [10], multiple hypothesis tracking (MHT) [15], and random finite set (RFS) based multitarget filters [86], [96]. This article focuses on summarizing JPDAF, MHT and RFS as the three main approaches to MTT. The JPDAF and MHT approaches are very well established and make up the bulk of the multitarget tracking literature, while the RFS approach is an emerging paradigm. JPDAF and MHT as well as many traditional MTT solutions, are formulated via data association followed by (single-target) filtering. Data association refers to the partitioning of the measurements into potential tracks and false alarms while filtering is used to estimate the state of the target given its measurement history (note that algorithms that operate on pre-detection signals do not involve data association). The distinguishing feature of the RFS approach is that, instead of

focusing on the data association problem, the RFS formulation directly seeks both optimal and suboptimal estimates of the multitarget state. Indeed some RFS-based algorithms do not require data association at all.

We begin by reviewing the fundamental principles of Bayesian estimation and summarizing some of the commonly used (single-target) filters for tracking in Section II. Section III presents some background on the MTT problem and describes the main challenges, setting the scene for the rest of the article. The JPDAF, MHT, and RFS approaches to MTT are presented in chronological order of developments in Sections IV, V, and VI respectively, with JPDAF being the earliest and RFS being the most recent. Nonetheless, Sections IV, V, and VI can be read independently from each other.

II. Bayesian Dynamic State Estimation

During the last two decades significant progress has been made in nonlinear filtering. This section provides a brief overview of the Bayesian paradigm for nonlinear filtering.

A. Bayesian Estimation

Consider the problem of estimating the state or parameter $\mathbf{x} \in \mathcal{X}$ from an observation $\mathbf{z} \in \mathcal{Z}$, where the state and observation spaces $\mathcal{X}$ and $\mathcal{Z}$ are assumed to be finite dimensional vector spaces in this article. The relationship between the observation and the state is described by the *likelihood function* $p(\mathbf{z}|\mathbf{x})$, the likelihood of the observation $\mathbf{z}$ given a state $\mathbf{x}$. Note that for each $\mathbf{x} \in \mathcal{X}$, $p(\cdot|\mathbf{x})$ is a probability density on $\mathcal{Z}$, i.e. for any $B \subseteq \mathcal{Z}$,

$$\Pr(\mathbf{z} \in B|\mathbf{x}) = \int_B p(\mathbf{z}|\mathbf{x})d\mathbf{z}.$$

In the Bayesian paradigm, prior information about the state is given by a *prior* probability density (or simply prior) p on $\mathcal{X}$, i.e. for any $A \subseteq \mathcal{X}$,

$$\Pr(\mathbf{x} \in A) = \int_A p(\mathbf{x})d\mathbf{x}.$$

All information about the state given the observation is contained in the *posterior* probability density (or simply posterior), which can be computed from the prior and likelihood function using Bayes rule

$$p(\mathbf{x}|\mathbf{z}) = \frac{p(\mathbf{z}|\mathbf{x})p(\mathbf{x})}{\int p(\mathbf{z}|\mathbf{x})p(\mathbf{x})d\mathbf{x}}. \quad (1)$$

An estimator of the state is a function $\hat{\mathbf{x}}$ that assigns the observation $\mathbf{z}$ a value $\hat{\mathbf{x}}(\mathbf{z}) \in \mathcal{X}$. A cost $C(\hat{\mathbf{x}}(\mathbf{z}), \mathbf{x})$ is associated with using $\hat{\mathbf{x}}(\mathbf{z})$ to estimate $\mathbf{x}$, and the *Bayes risk* $R(\hat{\mathbf{x}})$ is the expected cost over all possible realizations of the observation and state, i.e

$$R(\hat{\mathbf{x}}) = \int\int C(\hat{\mathbf{x}}(\mathbf{z}), \mathbf{x})p(\mathbf{z}|\mathbf{x})p(\mathbf{x})d\mathbf{x}d\mathbf{z}.$$

A *Bayes optimal estimator* is any estimator that minimizes the Bayes risk [70], [128]. The most common estimators are the *expected a posteriori* (EAP) or conditional mean and *maximum a posterior* (MAP) estimators given respectively by [1], [8], [70]

$$\begin{aligned} \hat{\mathbf{x}}^{\text{EAP}} &= \int \mathbf{x}p(\mathbf{x}|\mathbf{z})d\mathbf{x}, \\ \hat{\mathbf{x}}^{\text{MAP}} &= \arg\sup_{\mathbf{x}} p(\mathbf{x}|\mathbf{z}). \end{aligned}$$

These estimators minimize the Bayes risks for certain costs and are *consistent* in the sense that they converge almost surely to the true state as the number of data points increases. The EAP estimate is the minimum mean squared error estimate [8] and corresponds to the case where $C(\hat{\mathbf{x}}(\mathbf{z}), \mathbf{x}) = \|\hat{\mathbf{x}}(\mathbf{z}) - \mathbf{x}\|^2$.

B. The Bayes Recursion

Target tracking is a dynamic state estimation problem, in which the state varies with time. The dynamic model of a target can be described by a discrete-time model [1], [8], [64] or a continuous-time stochastic differential equation [8], [64]. This article only considers the discrete-time models.

The target state $\mathbf{x}_k$ evolves in time according to the *state transition equation*

$$\mathbf{x}_k = \mathbf{f}_{k,k-1}(\mathbf{x}_{k-1}, \mathbf{v}_{k-1}), \quad (2)$$

where $\mathbf{f}_{k,k-1}(\cdot,\cdot)$ is a nonlinear transformation and $\mathbf{v}_{k-1}$ is the process noise. In general, the state transition equation can be described by a *Markov transition density*.

$$f_{k|k-1}(\mathbf{x}_k|\mathbf{x}_{k-1}), \quad (3)$$

i.e. the probability density of a transition to the state $\mathbf{x}_k$ at time k given a state $\mathbf{x}_{k-1}$ at time $k-1$. Note that for each $\mathbf{x} \in \mathcal{X}$, $f_{k|k-1}(\cdot|\mathbf{x})$ is a probability density on $\mathcal{X}$. Some commonly used dynamic models are the nearly constant velocity, nearly constant acceleration, nearly coordinated turn, Ornstein-Uhlenbeck, and Singer models [8], [15], see the survey [78] for more details.

At each time k, the state $\mathbf{x}_k$ generates an observation $\mathbf{z}_k$ according to the *observation equation*

$$\mathbf{z}_k = \mathbf{g}_k(\mathbf{x}_k, \mathbf{w}_k), \quad (4)$$

where $\mathbf{g}_k(\cdot,\cdot)$ is a nonlinear transformation and $\mathbf{w}_k$ is observation noise. In general, the observation equation can be described by the *likelihood function*

$$g_k(\mathbf{z}_k|\mathbf{x}_k), \quad (5)$$

i.e. the probability density of receiving the observation $\mathbf{z}_k \in \mathcal{Z}$ given a state $\mathbf{x}_k$. For compactness we denote an array of variables $(\mathbf{y}_l, ..., \mathbf{y}_k)$ by $\mathbf{y}_{l:k}$. It is further assumed that the probability density of the observation history $\mathbf{z}_{1:k}$ condition on $\mathbf{x}_{1:k}$ is given by

$$p_{1:k}(\mathbf{z}_{1:k}|\mathbf{x}_{1:k}) = g_k(\mathbf{z}_k|\mathbf{x}_k)g_{k-1}(\mathbf{z}_{k-1}|\mathbf{x}_{k-1})...g_1(\mathbf{z}_1|\mathbf{x}_1).$$

All information about the state history to time k is encapsulated in the *posterior* density $p_{0:k}(\cdot|\mathbf{z}^k)$, where $\mathbf{z}^k = \mathbf{z}_{1:k}$ denotes the observation history. The posterior density can be computed recursively any for $k \geq 1$, starting from an initial prior p_0, via the Bayes recursion:

$$\begin{aligned} p_{0:k}(\mathbf{x}_{0:k}|\mathbf{z}^k) \propto \\ g_k(\mathbf{z}_k|\mathbf{x}_k)f_{k|k-1}(\mathbf{x}_k|\mathbf{x}_{k-1})p_{0:k-1}(\mathbf{x}_{0:k-1}|\mathbf{z}^{k-1}). \end{aligned} \quad (6)$$

The *filtering* density $p_k(\cdot|\mathbf{z}^k)$, is a marginal of the posterior density, which is defined as the probability density of the state at time k given the observation history $\mathbf{z}^k$. From an initial density p_0, the filtering density at time k can be computed recursively using the Bayes (filtering) recursion, which consists of the Chapman-Kolmogorov equation and the Bayes update:

$$p_{k|k-1}(\mathbf{x}_k|\mathbf{z}^{k-1}) = \int f_{k|k-1}(\mathbf{x}_k|\mathbf{x}) p_{k-1}(\mathbf{x}|\mathbf{z}^{k-1}) d\mathbf{x}, \quad (7)$$

$$p_k(\mathbf{x}_k|\mathbf{z}^k) = \frac{g_k(\mathbf{z}_k|\mathbf{x}_k) p_{k|k-1}(\mathbf{x}_k|\mathbf{z}^{k-1})}{\int g_k(\mathbf{z}_k|\mathbf{x}) p_{k|k-1}(\mathbf{x}|\mathbf{z}^{k-1}) d\mathbf{x}}, \quad (8)$$

where $p_{k|k-1}(\cdot|\mathbf{z}^{k-1})$ is called the *prediction* density. The *smoothing* density $p_{k|k+l}(\cdot|\mathbf{z}_{1:k+l})$, the probability density of the state at time k given the observation history $\mathbf{z}_{1:k+l}$, is another marginal of the posterior density. Smoothing can yield significantly better estimates than filtering by delaying the decision time and using data at a later time [105], [58], [46], [161].

C. The Kalman Filter

The Kalman filter (KF) is a closed form solution to the Bayes (filtering) recursion for linear Gaussian models [1], [8], [59], [64], [69], [126]. Specifically, the dynamical and observation models are linear transformations with additive Gaussian noise

$$\begin{aligned} \mathbf{x}_k &= \mathbf{F}_{k,k-1}\mathbf{x}_{k-1} + \mathbf{v}_{k-1}, \\ \mathbf{z}_k &= \mathbf{H}_k\mathbf{x}_k + \mathbf{w}_k, \end{aligned}$$

where $\mathbf{F}_{k,k-1}$ is the (square) transition matrix, $\mathbf{H}_k$ is the observation matrix, and $\mathbf{v}_{k-1}$ and $\mathbf{w}_k$ are independent zero-mean Gaussian noise variables with covariance matrices $\mathbf{Q}_{k-1}$ and $\mathbf{R}_k$ of appropriate dimensions. Thus, the transition density and likelihood function are

$$f_{k|k-1}(\mathbf{x}_k|\mathbf{x}_{k-1}) = \mathcal{N}(\mathbf{x}_k; \mathbf{F}_{k,k-1}\mathbf{x}_{k-1}, \mathbf{Q}_{k-1}), \quad (9)$$

$$g_k(\mathbf{z}_k|\mathbf{x}_k) = \mathcal{N}(\mathbf{z}_k; \mathbf{H}_k\mathbf{x}_k, \mathbf{R}_k), \quad (10)$$

where $\mathcal{N}(\cdot; \mathbf{m}, \mathbf{P})$ denotes a Gaussian density with mean and covariance $\mathbf{m}$ and $\mathbf{P}$ respectively. For example the nearly constant velocity model:

$$\mathbf{x}_k = \begin{bmatrix} x_k \\ y_k \\ \dot{x}_k \\ \dot{y}_k \end{bmatrix}, \qquad \mathbf{F}_{k,k-1} = \begin{bmatrix} \mathbf{I}_2 & \Delta\mathbf{I}_2 \\ \mathbf{0}_2 & \mathbf{I}_2 \end{bmatrix},$$

$$\mathbf{Q}_{k-1} = \sigma_w^2 \begin{bmatrix} \frac{\Delta^4}{4}\mathbf{I}_2 & \frac{\Delta^3}{2}\mathbf{I}_2 \\ \frac{\Delta^3}{2}\mathbf{I}_2 & \Delta^2\mathbf{I}_2 \end{bmatrix}, \qquad \mathbf{H}_k = \begin{bmatrix} \mathbf{I}_2 & \mathbf{0}_2 \end{bmatrix},$$

$\mathbf{R}_k = \sigma_v^2\mathbf{I}_2$, where $\mathbf{I}_n$ and $\mathbf{0}_n$ denote the $n \times n$ identity and zero matrices respectively, Δ is the sampling period, σ_w and σ_v are respectively the standard deviations of the process and measurement noise.

Under these assumptions, suppose that initial prior is a Gaussian $p_0 = \mathcal{N}(\cdot; \mathbf{m}_0, \mathbf{P}_0)$, then all subsequent filtering densities are Gaussians. Moreover, if at time $k-1$, the filtering density is a Gaussian of the form

$$p_{k-1}(\mathbf{x}_{k-1}|\mathbf{z}^{k-1}) = \mathcal{N}(\mathbf{x}_{k-1}; \mathbf{m}_{k-1}, \mathbf{P}_{k-1}),$$

then the predicted density to time k is a Gaussian

$$p_{k|k-1}(\mathbf{x}_k|\mathbf{z}^{k-1}) = \mathcal{N}(\mathbf{x}_k; \mathbf{m}_{k|k-1}, \mathbf{P}_{k|k-1}),$$

where

$$\begin{aligned} \mathbf{m}_{k|k-1} &= \mathbf{F}_{k,k-1}\mathbf{m}_{k-1}, \\ \mathbf{P}_{k|k-1} &= \mathbf{Q}_{k-1} + \mathbf{F}_{k,k-1}\mathbf{P}_{k-1}\mathbf{F}'_{k,k-1}, \end{aligned}$$

and the filtering density at time k is a Gaussian

$$p_k(\mathbf{x}_k|\mathbf{z}^k) = \mathcal{N}(\mathbf{x}_k; \mathbf{m}_k(\mathbf{z}_k), \mathbf{P}_k),$$

where

$$\begin{aligned} \mathbf{m}_k(\mathbf{z}_k) &= \mathbf{m}_{k|k-1} + \mathbf{K}_k(\mathbf{z}_k - \mathbf{H}_k\mathbf{m}_{k|k-1}), \\ \mathbf{P}_k &= [\mathbf{I} - \mathbf{K}_k\mathbf{H}_k]\mathbf{P}_{k|k-1}, \\ \mathbf{K}_k &= \mathbf{P}_{k|k-1}\mathbf{H}'_k\,\mathbf{S}_k^{-1}, \\ \mathbf{S}_k &= \mathbf{R}_k + \mathbf{H}_k\mathbf{P}_{k|k-1}\mathbf{H}'_k\,. \end{aligned}$$

The matrix $\mathbf{K}_k$ is referred to as the *Kalman gain*, the residual $\mathbf{z}_k - \mathbf{H}_k\mathbf{m}_{k|k-1}$ is referred to as the *innovation* and the matrix $\mathbf{S}_k$ is the *innovation covariance*.

The dynamic and measurement models in many real-world problems such as the bearing-only tracking, angle-only tracking, radar tracking, video tracking, etc. [10], [15], [100], [101], [126] are nonlinear. The process noise and measurement noise can also be non-additive and non-Gaussian. The Kalman filter is not applicable to these problems and in general, closed-form solutions are not possible. A number of approximate filtering algorithms such as the extended extended Kalman filter (EKF) [1], [8], [50], [64], [126], unscented Kalman filter (UKF) [67], [68], Gaussian sum filter [143], particle filter [4], [21], [44]–[46], [55], [126], quadrature filter, quasi Monte Carlo, grid based filter, cubature Kalman filter [3], [65], particle flow filter (PFF) [37], [42] have been proposed.

The EKF is a first order approximation to the Kalman filter based on linearization using the Taylor series expansion. The UKF uses the deterministic sampling principles of the unscented transform (UT) to propagate the first and second moments of the predicted and updated densities. The particle filter uses the sequential Monte Carlo (SMC) approach to approximates the posterior density using random sample points or *particles*. Next, we present two important approximate filters, the Gaussian sum filter and the particle filter.

D. The Gaussian Sum Filter

The Gaussian sum filter is a generalization of the Kalman filter to Gaussian mixture models [143]. Suppose that at time $k-1$, the filtering density is a Gaussian mixture of the form

$$p_{k-1}(\mathbf{x}_{k-1}|\mathbf{z}^{k-1}) = \sum_{i=1}^{N} w_{k-1}^{(i)} \mathcal{N}(\mathbf{x}_{k-1}; \mathbf{m}_{k-1}^{(i)}, \mathbf{P}_{k-1}^{(i)}).$$

Then the predicted density to time k and filtering density at time k are Gaussian mixtures

$$p_{k|k-1}(\mathbf{x}_k|\mathbf{z}^{k-1}) = \sum_{i=1}^{N} w_{k-1}^{(i)} \mathcal{N}(\mathbf{x}_k; \mathbf{m}_{k|k-1}^{(i)}, \mathbf{P}_{k|k-1}^{(i)}), \quad (11)$$

$$p_k(\mathbf{x}_k|\mathbf{z}^k) = \sum_{i=1}^{N} w_{k-1}^{(i)} \mathcal{N}(\mathbf{x}_k; \mathbf{m}_k^{(i)}(\mathbf{z}_k), \mathbf{P}_k^{(i)}), \quad (12)$$

where

$$\begin{aligned}
\mathbf{m}_{k|k-1}^{(i)} &= \mathbf{F}_{k,k-1}\mathbf{m}_{k-1}^{(i)}, && (13)\\
\mathbf{P}_{k|k-1}^{(i)} &= \mathbf{Q}_{k-1} + \mathbf{F}_{k,k-1}\mathbf{P}_{k-1}^{(i)}\mathbf{F}_{k|k-1}', && (14)\\
\mathbf{m}_{k}^{(i)}(\mathbf{z}_k) &= \mathbf{m}_{k|k-1}^{(i)} + \mathbf{K}_k^{(i)}(\mathbf{z}_k - \mathbf{H}_k\mathbf{m}_{k|k-1}^{(i)}), && (15)\\
\mathbf{P}_{k}^{(i)} &= [\mathbf{I} - \mathbf{K}_k^{(i)}\mathbf{H}_k]\mathbf{P}_{k|k-1}^{(i)}, && (16)\\
\mathbf{K}_k^{(i)} &= \mathbf{P}_{k|k-1}^{(i)}\mathbf{H}_k' \left[\mathbf{S}_{k|k-1}^{(i)}\right]^{-1}, && (17)\\
\mathbf{S}_{k|k-1}^{(i)} &= \mathbf{R}_k + \mathbf{H}_k\mathbf{P}_{k|k-1}^{(i)}\mathbf{H}_k'. && (18)
\end{aligned}$$

For clarity, we have presented the Gaussian sum filter prediction and update for linear Gaussian model. In the more general case where the transition density and/or likelihood function are Gaussian mixtures, the predicted density (11) involves an additional sum over the components of the Gaussian mixture transition density, and/or the filtering density (12) involves an additional sum over the components of the Gaussian mixture likelihood function [143]. The number of Gaussians required to represent the exact filtering density increases exponentially with time and Gaussian mixture reduction techniques are required to manage memory and computational load [135], [134], [137].

E. The Particle Filter

The particle or sequential Monte Carlo (SMC) method is a class of approximate numerical solutions to the Bayes recursion that are applicable to nonlinear non-Gaussian dynamic and observation models. The basis of the particle method is the use of random samples (particles) to approximate probability distributions of interest [4], [21], [44]–[46], [55], [126].

Consider N independently and identically distributed (i.i.d.) samples $\{\mathbf{x}^{(i)}\}_{i=1}^N$ from an arbitrary probability density p of $\mathbf{x}$. For any function h of $\mathbf{x}$, the (finite) expectation of h can be approximated by the empirical expectation, i.e.,

$$\int h(\mathbf{x})p(\mathbf{x})d\mathbf{x} \approx \frac{1}{N}\sum_{i=1}^{N} h(\mathbf{x}^{(i)}).$$

The empirical expectation is unbiased and tends to the true expectation almost surely as N tends to infinity. Moreover, the rate of convergence is not dependent on the dimension of the integral, but primarily on N, the number of independent samples. Hence, we can regard the samples $\{\mathbf{x}^{(i)}\}_{i=1}^N$ as a point mass approximation of p, i.e.,

$$p(\mathbf{x}) \approx \frac{1}{N}\sum_{i=1}^{N} \delta(\mathbf{x} - \mathbf{x}^{(i)}),$$

where δ denotes the Dirac delta.

Now consider the case where the density p is only known up to a normalizing constant, i.e. $p(\mathbf{x}) \propto \tilde{p}(\mathbf{x})$, such as in the Bayes recursion where the normalizing constant is difficult to compute. Since it is difficult to sample from p, we draw N i.i.d. samples $\{\mathbf{x}^{(i)}\}_{i=1}^N$ from a known density q, referred to as the *proposal* or *importance density*, and then weight these samples accordingly so as to obtain a weighted point mass approximation to p. More concisely, for any function h, the (finite) expectation of h can be approximated by the empirical expectation, i.e.

$$\int h(\mathbf{x})p(\mathbf{x})d\mathbf{x} \approx \sum_{i=1}^{N} w^{(i)}h(\mathbf{x}^{(i)}),$$

where

$$\begin{aligned}
w^{(i)} &= \frac{\tilde{w}(\mathbf{x}^{(i)})}{\sum_{j=1}^N \tilde{w}(\mathbf{x}^{(j)})},\\
\tilde{w}^{(i)}(\mathbf{x}^{(i)}) &= \frac{p(\mathbf{x}^{(i)})}{q(\mathbf{x}^{(i)})},
\end{aligned}$$

are known as the *normalized importance weights* and *importance weights* respectively. A "good" proposal is one such that the weights $\{w^{(i)}\}_{i=1}^N$ all have roughly the same value. For this so-called *importance sampling* approximation, the empirical expectation is biased. Nonetheless, it still tends to the true expectation almost surely as N tends to infinity. Hence, we can regard the weighted samples $\{(w^{(i)}, \mathbf{x}^{(i)})\}_{i=1}^N$ as a weighted point mass approximation of p, i.e.,

$$p(\mathbf{x}) \approx \sum_{i=1}^{N} w^{(i)}\delta(\mathbf{x} - \mathbf{x}^{(i)}).$$

The key operation in particle filtering is the sequential application of importance sampling to recursively approximate the posterior. This is known as *sequential importance sampling* (SIS) [55], [4], [44], [45], [126] and is described as follows:

Suppose that the posterior density $p_{0:k-1}$, at time $k-1$, is represented as a set of weighted particles $\{(w_{k-1}^{(i)}, \mathbf{x}_{0:k-1}^{(i)})\}_{i=1}^N$, i.e.,

$$p_{0:k-1}(\mathbf{x}_{0:k-1}|\mathbf{z}^{k-1}) \approx \sum_{i=1}^{N} w_{k-1}^{(i)}\delta(\mathbf{x}_{0:k-1} - \mathbf{x}_{0:k-1}^{(i)}),$$

and given a proposal density $q_k(\cdot|\mathbf{x}_{k-1}^{(i)}, \mathbf{z}_k)$ that we can easily sample from. Then the posterior density $p_{0:k}$, at time k, is represented as a new set of weighted particles $\{(w_k^{(i)}, \mathbf{x}_{0:k}^{(i)})\}_{i=1}^N$, i.e.,

$$p_{0:k}(\mathbf{x}_{0:k}|\mathbf{z}^{k}) \approx \sum_{i=1}^{N} w_{k}^{(i)}\delta(\mathbf{x}_{0:k} - \mathbf{x}_{0:k}^{(i)}),$$

where

$$\begin{aligned}
\mathbf{x}_{0:k}^{(i)} &= (\mathbf{x}_{0:k-1}^{(i)}, \mathbf{x}_k^{(i)})\\
\mathbf{x}_k^{(i)} &\sim q_k(\cdot|\mathbf{x}_{k-1}^{(i)}, \mathbf{z}_k),\\
w_k^{(i)} &= \tilde{w}_k^{(i)} / \sum\nolimits_{i=1}^{N} \tilde{w}_k^{(i)},\\
\tilde{w}_k^{(i)} &= w_{k-1}^{(i)} \frac{g_k(\mathbf{z}_k|\mathbf{x}_k^{(i)}) f_{k|k-1}(\mathbf{x}_k^{(i)}|\mathbf{x}_{k-1}^{(i)})}{q_k(\mathbf{x}_k^{(i)}|\mathbf{x}_{k-1}^{(i)}, \mathbf{z}_k)}.
\end{aligned}$$

The selection of optimal proposals along with practical strategies for constructing good proposals to this are considered in [44], [80], [118]. If we are only interested in the filtering density then only the most recent component of the samples are kept, i.e. the filtering density is represented by the weight samples $\{(w_k^{(i)}, \mathbf{x}_k^{(i)})\}_{i=1}^N$.

The basic SIS algorithm suffers from particle depletion or degeneracy where the variance of the importance weights

increases over time, thereby degrading the quality of the particle approximation. Particle depletion is generally mitigated by resampling the weighted particles $\{(w_k^{(i)}, \mathbf{x}_{0:k}^{(i)})\}_{i=1}^N$ to generate more replicas of particles with high weights and eliminate those with low weights [55]. There are many resampling schemes available, and the choice of resampling scheme affects the computational load as well as the quality of the particle approximation, see for example [19], [43], [46], [107]. An additional Markov Chain Monte Carlo (MCMC) step can then be used to rejuvenate particle diversity [25], [52] if necessary. Relevant convergence results for particle filtering can be found in [32], [41].

Various extensions of the particle filtering methodology have been proposed to improve performance. Rao-Blackwellization techniques can be incorporated with the particle filter (PF) [45], [126] to improve performance for particular classes of state space models, e.g. the Mixture Kalman Filter (MKF) [24]. The underlying idea is to partition the state vector into a linear Gaussian component and a nonlinear non-Gaussian component. Then, the former is solved analytically using a Kalman filter and the latter with a particle filter so that the computational effort is appropriately focused. Continuous approximations to the posterior density can be obtained with kernel smoothing techniques. Examples of this approach are the convolution or regularized particle filter [45], [126]. Related approaches are the Gaussian particle and Gaussian sum particle filters [74], [75].

F. Filtering Algorithms for Maneuvering Targets

The filtering algorithms discussed previously use a single dynamic model and hence are known as single-model filters. The motion of a maneuvering target involves multiple dynamic models. For example, an aircraft can fly with a nearly constant velocity motion, accelerated/decelerated motion, and coordinated turn [8], [10]. The multiple model approach is an effective filtering algorithm for maneuvering targets in which the continuous kinematic state and discrete mode or model are estimated. This class of problems are known as *jump Markov* or *hybrid state estimation* problems. The discrete-time dynamic and measurement models for the hybrid state estimation problem [8], [10], [126] are given, respectively, by

$$\begin{aligned} \mathbf{x}_k &= \mathbf{f}_{k,k-1}(\mathbf{x}_{k-1}, \mu_k, \mathbf{v}_{k-1}), \\ \mathbf{z}_k &= \mathbf{g}_k(\mathbf{x}_k, \mu_k, \mathbf{w}_k), \end{aligned}$$

where μ_k is the mode in effect from time $k-1$ to k. The interacting multiple model (IMM) and variable-structure IMM (VS-IMM) estimators [8], [10], [77], [79], [104] are two well known filtering algorithms for maneuvering targets. The number of modes in the IMM is kept fixed, whereas in the VS-IMM the number of modes are adaptively selected from a fixed set of modes for improved estimation accuracy and computational efficiency.

III. Multitarget Tracking

This section provides some background on the MTT problem and the main challenges, setting the scene for the rest of the article.

A. Multitarget Systems

Driven by aerospace applications, MTT was originally developed for tracking targets from radar measurements. Fig. 1 shows a typical scenario describing the measurements by a radar in which five true targets are present in the radar dwell volume (the volume of the measurement space sensed by a sensor at a scan time) and six measurements are collected by the radar. We see from Fig. 1 that three target-originated measurements and three false alarms (FAs) are generated, one target is not detected by the radar, and two closely spaced targets are not resolved. This type of information regarding the nature and origin of measurements is not known for real radar measurements due to *measurement origin uncertainty*. At each discrete dwell/scan time t_j, a set of noisy radar measurements with measurement origin uncertainty is sent to a tracker, as shown in Fig. 2.

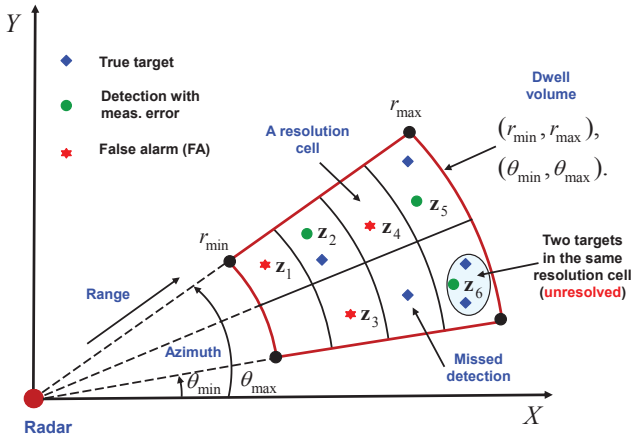

Fig. 1. A typical radar measurement scenario.

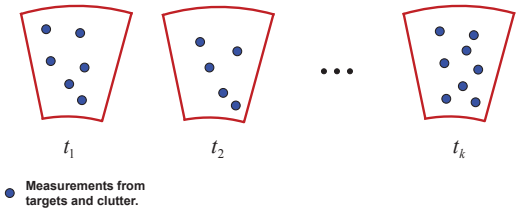

Fig. 2. Varying number of noisy radar measurements in dwells.

In a general multitarget system, not only do the states of the targets vary with time, but the number of targets also changes due to targets appearing and disappearing as illustrated in Fig. 3. The targets are observed by a sensor (or sensors) such as radar, sonar, electro-optical, infrared, camera etc. The sensor signals at each time step are preprocessed into a set of points or detections. It is important to note that existing targets may not be detected and that FAs (due to clutter) may occur. As a result, at each time step the multitarget observation is a set of detections, only some of which are generated by targets and there is no information on which targets generated which detections (see Fig. 3).

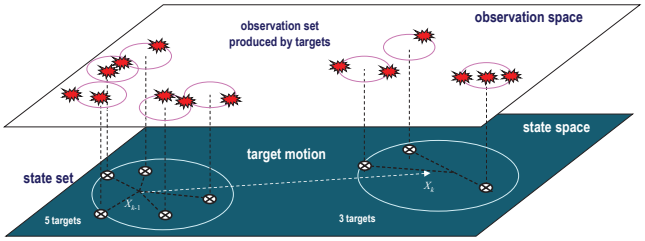

Fig. 3. Multiple-target system model: the number of targets changes from 5 to 3, targets generate at each time a random number of measurements.

Most MTT algorithms assume a *standard multitarget transition model*, in which each existing target $\mathbf{x}_{k-1}$, at time $k-1$, either continues to exist at time k with probability $P_{S,k|k-1}(\mathbf{x}_{k-1})$ and moves to a new state $\mathbf{x}_k$ with probability density $f_{k|k-1}(\mathbf{x}_k|\mathbf{x}_{k-1})$, or dies with probability $1-P_{S,k|k-1}(\mathbf{x}_{k-1})$. In addition, a random number of new targets can appear from random locations in the state space at time k. Each target is assumed to appear and evolve independently from others. Different multitarget tracking approaches employ different models for target births and deaths.

In a *standard multitarget observation model*, each target $\mathbf{x}_k$, at time k, is either detected with probability $P_{D,k}(\mathbf{x}_k)$ and generates an observation $\mathbf{z}_k$ with likelihood $g_k(\mathbf{z}_k|\mathbf{x}_k)$, or missed with probability $1-P_{D,k}(\mathbf{x}_k)$. In addition to the detections, the tracker also receives a random number of FAs from random locations in the measurement space. It is assumed that each target generates observations independently from other targets and FAs and that each detection can only be generated from at most one target. The standard multitarget observation model is the most widely used. Other models include: the merged or unresolved measurement model [23], [71], [17], [91], [150], [14], where two or more targets can share a detection; extended target/group measurement model, where each target/group can generate multiple detections [51], [72], [90] [114], [53], [48], [56], [81], [108]; track-before-detect/image measurement model [136], [39], [160], [62], [63], [49], and the superpositional measurement model [87], [111], where the observed signal is a superposition of observations generated by each of the targets present. This article only considers the standard multitarget measurement model.

B. The MTT Problem

The objective of MTT is to jointly estimate, at each observation time, the number of targets and their trajectories from sensor data. Even at a conceptual level, MTT is a non-trivial extension of single-target tracking. Indeed MTT is far more complex in both theory and practice.

The concept of estimation error between a reference quantity and its estimated values plays a fundamental role in any estimation problem. In (single-target) filtering the system state is a vector and the notion of state estimation error is taken for granted. For example, the EAP estimator minimizes the expected squared Euclidean distance $\|\hat{\mathbf{x}}-\mathbf{x}\|^2$ between the estimated state vector $\hat{\mathbf{x}}$ and true state vector $\mathbf{x}$. However, the concept of Euclidean distance is not suitable for the multitarget case. To see this consider the scenario depicted in Fig. 4. Suppose that the multitarget state is formed by stacking individual states into a single vector with the ground truth represented by X and the estimate represented by $\hat{X}$. The estimate is correct but the Euclidean distance is $||\hat{X}-X||=2$. Moreover, when the estimated number of targets is different from the true number the Euclidean distance is not defined.

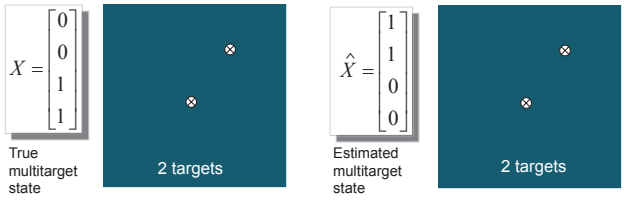

Fig. 4. A possible vector representation of multi-target states when the estimated and true multi-target states have the same number of targets.

Central to Bayesian state estimation is the concept of Bayes risk/optimality [70], [128]. A Bayes optimal solution is not simply one that invokes Bayes rule. Criteria for optimality for the single-target case such as the squared Euclidean distance is not appropriate. In addition, the concept of consistency (of an estimator) cannot be taken for granted since it is not clear what is the notion of convergence in the multitarget realm.

From a practical point of view, MTT is not a simple extension of classical (single-target) filtering. Even for the simple special case with exactly one target in the scene, classical filtering methods (described in Section II) cannot be directly applied due to false detection, missed detection, and measurement origin uncertainty. The simplest solution is the *nearest neighbor* (NN) filter which applies the Bayes filter to the measurement that is closest to the predicted measurement [7], [10], [15]. A more sophisticated yet intuitively appealing solution is the Probabilistic Data Association filter (PDAF) which applies the Bayes filter to the average of all measurements weighted according to their association probabilities [7], [10]. The solution based on enumerating association hypotheses, proposed in [141], coincides with the Bayes optimal filter in the presence of false detections, missed detections, and measurement origin uncertainty proposed in [158]. In the multitarget setting, even for the special case where all targets are detected and no false detections occur, classical filtering methods are not directly applicable since there is no information on which target has generated which measurements.

The simplest multitarget filter is the *global nearest neighbor* (GNN) tracker, an extension of the NN filter to the multiple target case. The GNN tracker searches for the unique joint association of measurements to targets that minimizes/maximizes a total cost, such as a total distance or likelihood. The GNN filter then performs standard Bayes filtering for each target using these associated measurements directly. Although the GNN scheme is intuitively appealing and simple to implement, it is susceptible to track loss and consequently exhibits poor performance when targets are not well separated [15].

The JPDAF [7], [10] is an extension of the PDAF to a fixed and known number of targets. The JPDAF uses joint association events and joint association probabilities in order to avoid conflicting measurement to track assignments in the presence of multiple targets. The complexity of the calculation for joint association probabilities grows exponentially with the number of targets and the number of measurements. Several approximation approaches have been proposed such as the deterministic strategies in [130], [131], [103], [61], [12], [169], [132] and the Markov Chain Monte Carlo (MCMC) based strategies in [112]. Moreover, since the basic JPDAF can only accommodate a fixed and known number of targets, several novel extensions have been proposed to accommodate an unknown and time varying number of targets, such as the joint integrated PDAF (JIPDAF) [109] along with an efficient implementation [110], and automatic track formation (ATF) [6]. Further detail on the JPDAF is given in Section IV.

MHT [123], [76], [15], [16], [10], [101] is a deferred decision approach to data association based MTT. At each observation time, the MHT algorithm attempts to propagate and maintain a set of association hypotheses with high posterior probability or track score. When a new set of measurements arrives, a new set of hypotheses is created from the existing hypotheses and their posterior probabilities or track scores are updated using Bayes rule. In this way, the MHT approach inherently handles initiation and termination of tracks, and hence accommodates an unknown and time-varying number of targets. Based on the best hypothesis, a standard Bayes (or Kalman when the models are linear Gaussian) filter can be used on the measurements in each track to estimate the trajectories of individual targets. The total number of possible hypotheses increases exponentially with time and heuristic pruning/merging of hypotheses is performed to reduce computational requirements. Further details on the MHT approach is given in Section V.

Related deferred decision approaches, based on an intuitive and explicit formulation with hidden Markov models and subsequent state estimation by application of the Viterbi algorithm, can also be found in [171]. An innovative and completely different approach proposed in [112] casts the problem of finding the hypothesis with the highest posterior probability as a combinatorial optimization problem, which is solved using reversible jump Markov Chain Monte Carlo (RJ-MCMC) techniques in order to generate samples from the posterior density.

The probabilistic MHT (PMHT) is a tractable approach that operates over several frames, reducing the complexity by formulating the problem as one of maximum likelihood estimation and applying Expectation-Maximization (EM) [146], [168], [33]. The computation is simplified at a sacrifice in performance, by removing the requirement for each target to have a single measurement. An efficient implementation of PMHT termed the turbo PMHT was proposed in [133] based on the idea of turbo coding, which exhibits good tracking performance with a very low computational complexity. PMHT with track maintenance is decribed in [38].

The RFS approach retains the same Bayesian estimation methodology for single-target (in Section II), by representing the multitarget state as a finite set [86], [96], which admits suitable distances between multitarget states (see [60], [140]). This framework provides appropriate notions of multitarget probability density, that enables concepts such as state space model, Bayes recursion, Bayes optimality to be directly translated to the multitarget case. Moreover it covers more complex multitarget tracking problems such as non-Poisson, non-homogeneous FAs, state dependent probability of detection [93], [95], extended targets [86], [56], merged measurements [14], non-standard measurement (including image, fuzzy and Dempster-Shafer) [86], distributed multitarget tracking [11], etc. under one single umbrella without any ad hoc modifications [96]. Further detail on the RFS approach is given in Section VI.

Two types of tracking architectures, centralized and distributed, are used in multisensor multitarget tracking (MTT) [10], [15], [101]. This article only addresses centralized tracking. The three data association based MTT algorithms GNN, JPDAF, MHT [10], [15], [101], have been widely used for more than three decades, while the RFS-based algorithms [86], [96] developed during the last decade have received considerable interest. The computational cost of the MHT is much higher than that of the GNN or JPDAF. Numerous studies have shown that the MHT works significantly better than the GNN and JPDAF for tracking scenarios with low signal-to-noise ratio (SNR) and closely spaced targets [15, Section 6.8.1]. Recent independent studies [147], [149] demonstrated that a sub-optimal RFS-based filter called the cardinalized probability hypothesis density filter [85], [156] has comparable performance to MHT which much lower computational cost.

IV. Joint Probabilistic Data Association Filter

List of mathematical symbols:

N_T	number of targets
$\mathbf{Z}_k$	vector of observations at time k
$\mathbf{Z}^k$	$(\mathbf{Z}_1, \mathbf{Z}_2..., \mathbf{Z}_k)$
θ_{jt_j}	event that measurement j originated from target t_j
$\boldsymbol{\theta}$	joint association event $[\theta_{jt_j}], j = 1, \ldots, m$
$\boldsymbol{\Omega}$	validation matrix: $[\omega_{jt}], j = 1, \ldots, m; t = 0, 1, \ldots, N_T$
$\hat{\boldsymbol{\Omega}}(\boldsymbol{\theta})$	event matrix: $[\hat{\omega}_{jt}(\boldsymbol{\theta})], j = 1, \ldots, m; t = 0, 1, \ldots, N_T$
$\delta_t(\boldsymbol{\theta})$	detection indicator for event $\boldsymbol{\theta}$
$\delta(\boldsymbol{\theta})$	vector of target detection indicators for event $\boldsymbol{\theta}$
$\tau_j(\boldsymbol{\theta})$	measurement association indicator for event $\boldsymbol{\theta}$
$\phi(\boldsymbol{\theta})$	number of false measurements in event $\boldsymbol{\theta}$
$\mu_F(\phi)$	prior pmf of the number of false measurements
V	volume of the surveillance region
λ	spatial density of false measurements
m_k	number of measurements in the union of the validation regions
P_D^t	detection probability of target t
$j_t(\boldsymbol{\theta})$	index of the measurement associated with target t in event $\boldsymbol{\theta}$
β_{jt}	marginal assocation probability

A. Overview

The joint probabilistic data association filter (JPDAF) is the multitarget extension of the probabilistic data association filter (PDAF) for single target tracking [10].

1) Assumptions:

- There is a known number of established targets N_T in clutter.
- Measurements from one target can fall in the validation region of a neighboring target — this can happen over several sampling times and acts as a *persistent interference*.
- The past is summarized by an approximate sufficient statistic — state estimates (approximate conditional means) and covariances for each target.
- The states are assumed Gaussian distributed with the above means and covariances.
- The models for the various targets do not have to be the same.
- The targets are resolved — there are no unresolved (merged) measurements.

2) The Approach:

- The measurement to target association probabilities are computed *jointly across the targets*.
- The association probabilities are computed only for the latest set (scan) of measurements.[1]
- The state estimation is done
  - separately for each target as in the PDAF (decoupled), or
  - in a coupled manner using a stacked state vector in the JPDA coupled filter (JPDACF) (see [10] for details).

B. The Key Feature of the JPDAF

The evaluation of the conditional probabilities of the following *joint association events* $\boldsymbol{\theta}$ pertaining to the current time k (the time index k is omitted for simplicity where it does not cause confusion) consisting of θ_{jt_j}, the event that measurement j originated from target t_j, $j = 1, \ldots, m$, $t = 0, 1, \ldots, N_T$; t_j is the index of the target to which measurement j is associated in the event θ_{jt_j} under consideration.

1) Remark: For the purpose of deriving the joint probabilities, no individual validation gates will be assumed for the various targets. Instead, each measurement will be assumed validated for each target, i.e., every validation gate coincides with the entire surveillance region.

This approach is adopted in order to have the pdf of each false measurement the *same*, i.e., uniformly distributed in the *entire validation region*.

C. The Feasible Joint Events

Validation gates are used for the selection of the *feasible joint events* but not in the evaluation of their probabilities. This logic avoids considering events whose probabilities are negligible and thus has a negligible effect on the other probabilities.

1) The Validation Matrix: Define the *validation matrix*

$$\boldsymbol{\Omega} = [\omega_{jt}] \qquad j = 1, \ldots, m; \quad t = 0, 1, \ldots, N_T \tag{19}$$

with binary elements that indicate if measurement j lies in the validation gate of target t. The index $t = 0$ stands for "none of the targets" and the corresponding column of $\boldsymbol{\Omega}$ has all units since each measurement could have originated from clutter or false alarm.

2) The Event Matrix: A *joint association event* $\boldsymbol{\theta}$ is represented by the *event matrix*

$$\hat{\boldsymbol{\Omega}}(\boldsymbol{\theta}) = [\hat{\omega}_{jt}(\boldsymbol{\theta})] \tag{20}$$

consisting of the units in $\boldsymbol{\Omega}$ corresponding to the associations in $\boldsymbol{\theta}$, with $\hat{\omega}_{jt}(\boldsymbol{\theta}) = 1$ if $\theta_{jt} \in \boldsymbol{\theta}$ and 0 otherwise.

A *feasible association event* is one where

(i) a measurement can have only one source,
(ii) at most one measurement can originate from a target, for which the detection indicator is denoted as $\delta_t(\boldsymbol{\theta})$.

3) Generation of the Feasible Joint Association Events: The generation of the event matrices $\hat{\boldsymbol{\Omega}}$ corresponding to feasible events can be done by scanning $\boldsymbol{\Omega}$ and picking

(i) one unit per row, and
(ii) one unit per column except for $t = 0$ where the number of units (which is the number of false measurements) is not restricted.

The binary variable $\delta_t(\boldsymbol{\theta})$ is called the *target detection indicator* since it is unity if one of the m measurements is associated to target t in event $\boldsymbol{\theta}$, i.e., target t has been detected.

It is also convenient to define another binary variable, called the *measurement association indicator* $\tau_j(\boldsymbol{\theta})$ to indicate if measurement j is associated with a target in event $\boldsymbol{\theta}$.

With this definition, the number of false (unassociated) measurements in event $\boldsymbol{\theta}$ is

$$\phi(\boldsymbol{\theta}) = \sum_{j=1}^{m} [1 - \tau_j(\boldsymbol{\theta})]. \tag{21}$$

D. Evaluation of the Joint Probabilities

The *joint association event probabilities* are, with Bayes' formula,

$$\Pr(\boldsymbol{\theta}_k | \mathbf{Z}^k) = \frac{1}{c} p(\mathbf{Z}_k | \boldsymbol{\theta}_k, m_k, \mathbf{Z}^{k-1}) \Pr(\boldsymbol{\theta}_k | m_k), \tag{22}$$

where c is the normalization constant.

1) Assumption: The states of the targets conditioned on the past observations are mutually *independent*.[2]

2) The Likelihood Function of a Joint Association Event: The *likelihood function of the joint association event* on the r.h.s. of (22) is

$$p(\mathbf{Z}_k | \boldsymbol{\theta}_k, m_k, \mathbf{Z}^{k-1}) = \prod_{j=1}^{m_k} p(\mathbf{z}_{j,k} | \boldsymbol{\theta}_{k,jt_j}, \mathbf{Z}^{k-1}), \tag{23}$$

[1] This is in view of the fact that, if a sufficient statistic is available, then there is no need to consider the past (previous measurements). However, it should be recalled that the Gaussian sufficient statistic is an approximation.

[2] This assumption can be relaxed, and results in the JPDACF [10].

where m_k is the number of measurements in the union of the validation regions at time k. The product form of (23) follows from the above assumption.

The conditional pdf of a measurement given its origin is

$$p(\mathbf{z}_{j,k}|\theta_{k,jt_j}, \mathbf{Z}^{k-1}) = \begin{cases} f_{t_j}(\mathbf{z}_{j,k}) & \text{if } \tau_j(\boldsymbol{\theta}_k) = 1, \\ V^{-1} & \text{if } \tau_j(\boldsymbol{\theta}_k) = 0, \end{cases} \tag{24}$$

where

$$f_{t_j}(\mathbf{z}_{j,k}) = \mathcal{N}(\mathbf{z}_{j,k}; \hat{\mathbf{z}}^{t_j}_{k|k-1}, \mathbf{S}^{t_j}_k) \tag{25}$$

and $\hat{\mathbf{z}}^{t_j}_{k|k-1}$ is the predicted measurement for target t_j, with associated innovation covariance $\mathbf{S}^{t_j}_k$.

Measurements not associated with a target are assumed uniformly distributed in the surveillance region of volume V.

Using (24), the pdf (23) can be written as follows

$$p(\mathbf{Z}_k|\boldsymbol{\theta}_k, m_k, \mathbf{Z}^{k-1}) = V^{-\phi} \prod_j [f_{t_j}(\mathbf{z}_{j,k})]^{\tau_j}. \tag{26}$$

In the above V^{-1} is raised to power $\phi(\boldsymbol{\theta})$, the total number of false measurements in event θ_k and the indicators $\tau_j(\boldsymbol{\theta})$ select the single measurement densities according to their associations in event $\boldsymbol{\theta}_k$.

3) The Prior Probability of a Joint Association Event: The prior (to time k) probability of an event $\theta(k)$, the last term in (22), is obtained next. Denote by $\delta(\boldsymbol{\theta})$ the vector of target detection indicators corresponding to event $\boldsymbol{\theta}_k$.

The joint probability can be written as

$$\begin{aligned} \Pr(\boldsymbol{\theta}_k|m_k) = \; & \Pr(\boldsymbol{\theta}_k|\delta(\boldsymbol{\theta}), \phi(\boldsymbol{\theta}), m_k) \\ & \cdot \Pr(\delta(\boldsymbol{\theta}), \phi(\boldsymbol{\theta})|m_k). \end{aligned} \tag{27}$$

The first term on the r.h.s. of the above is obtained from the following reasoning based on *combinatorics*:

(i) In event $\boldsymbol{\theta}_k$ the set of targets assumed detected consists of $m_k - \phi$ targets.

(ii) The number of measurement to target assignment events $\boldsymbol{\theta}_k$ in which the same set of targets is detected is given by the number of *permutations* of the m_k measurements taken as $m_k - \phi$, the number of targets to which a measurement is assigned under the same detection event.

Therefore, assuming each such event a priori equally likely, one has

$$\Pr(\boldsymbol{\theta}_k|\delta(\boldsymbol{\theta}), \phi(\boldsymbol{\theta}), m_k) = \left(\frac{m_k!}{\phi!}\right)^{-1}. \tag{28}$$

After some manipulations [34] and assuming δ and ϕ independent, the last term in (27) becomes

$$\Pr(\delta(\boldsymbol{\theta}), \phi(\boldsymbol{\theta})|m_k) = \left[\prod_t (P^t_D)^{\delta_t}(1 - P^t_D)^{1-\delta_t}\right] \frac{\mu_F(\phi)}{\Pr(m_k)}, \tag{29}$$

where P^t_D is the detection probability of target t and $\mu_F(\phi)$ is the prior pmf of the number of false measurements (the clutter model). The indicators $\delta_t(\boldsymbol{\theta})$ have been used in (29) to select the probabilities of detection and no detection events according to the event $\boldsymbol{\theta}_k$ under consideration. The term $\Pr(m_k)$ in (29) will be absorbed in the normalization constant in (31).

Combining (28) and (29) into (27) yields the *prior probability of a joint association event* $\boldsymbol{\theta}_k$ as

$$\Pr(\boldsymbol{\theta}_k|m_k) = \frac{\phi!\mu_F(\phi)}{m_k! \Pr(m_k)} \prod_t (P^t_D)^{\delta_t}(1 - P^t_D)^{1-\delta_t}. \tag{30}$$

4) The Posterior Probability of a Joint Association Event: Combining (26) and (30) into (22) yields the *posterior probability of a joint association event* $\boldsymbol{\theta}_k$ as

$$\begin{aligned} \Pr(\boldsymbol{\theta}_k|\mathbf{Z}^k) = \frac{1}{c}\frac{\phi!}{m_k!}\mu_F(\phi)V^{-\phi} & \prod_j [f_{t_j}(\mathbf{z}_{j,k})]^{\tau_j} \\ & \cdot \prod_t (P^t_D)^{\delta_t}(1 - P^t_D)^{1-\delta_t}, \end{aligned} \tag{31}$$

where ϕ , δ_t and τ_j are all functions of the event $\boldsymbol{\theta}_k$ under consideration.

The above still needs the specification of the pmf of the number of false measurements $\mu_F(\phi)$, carried out in the next section.

E. The Parametric and Nonparametric JPDAF

As in the case of the PDAF, the JPDAF has two versions, according to the model used for the pmf $\mu_F(\phi)$ of the number of false measurements.

1) The Parametric JPDAF: The *parametric JPDAF* uses the Poisson pmf $\mu_F(\phi)$ with parameter λV which requires the spatial density λ of the false measurements.

Using the Poisson pmf in (31) leads to the cancellation of V^ϕ and $\phi!$. Furthermore, each term contains $e^{-\lambda V}$ and $m_k!$, which also cancel since they appear in the denominator c of (31), which is the sum of all the numerators.

Thus the joint association probabilities of the parametric JPDAF are

$$\Pr(\boldsymbol{\theta}_k|\mathbf{Z}^k) = \frac{\lambda^\phi}{c_1} \prod_j [f_{t_j}(\mathbf{z}_{j,k})]^{\tau_j} \prod_t (P^t_D)^{\delta_t}(1 - P_D)^{1-\delta_t}, \tag{32}$$

where c_1 is the appropriate normalization constant.

Since m_k is a fixed number, the joint association probabilities can be rewritten as

$$\begin{aligned} P(\boldsymbol{\theta}_k|\mathbf{Z}^k) = \frac{1}{c_1\lambda^{-m_k}} & \prod_j [\lambda^{-1} f_{t_j}(\mathbf{z}_{j,k})]^{\tau_j} \\ & \cdot \prod_t (P^t_D)^{\delta_t}(1 - P_D)^{1-\delta_t} \end{aligned} \tag{33}$$

by defining a new normalization constant.

Each term in the first product above is the likelihood ratio of the corresponding measurement having originated from a particular target vs. from clutter. The denominator of these likelihood ratios is the spatial density of the clutter, which plays the role of the pdf of clutter originated measurements. This is a consequence of the Poisson prior.

2) The Nonparametric JPDAF: The *nonparametric JPDAF* uses the *diffuse prior*

$$\mu_F(\phi) = \epsilon \qquad \forall \phi, \tag{34}$$

which does not require the parameter λ.

With this, (31) becomes after canceling the constant ϵ and $m_k!$, which appear in each expression,

$$\Pr(\boldsymbol{\theta}_k|\mathbf{Z}^k) = \frac{1}{c_3}\frac{\phi!}{V^\phi}\prod_j [f_{t_j}(\mathbf{z}_{j,k})]^{\tau_j} \cdot \prod_t (P_D^t)^{\delta_t}(1-P_D)^{1-\delta_t}, \quad (35)$$

where c_3 is the appropriate normalization constant.

Similarly to the nonparametric case, the joint association probabilities can be rewritten as

$$\Pr(\boldsymbol{\theta}_k|\mathbf{Z}^k) = \frac{\phi!}{c_3 V^{-m_k}}\prod_j [V f_{t_j}(\mathbf{z}_j(k))]^{\tau_j} \cdot \prod_t (P_D^t)^{\delta_t}(1-P_D)^{1-\delta_t}. \quad (36)$$

As it can be seen from (35), the *nonparametric JPDAF* expressions contain a term that can be called *pseudo sample spatial measurement density* $\phi!/V^\phi$ in place of λ^ϕ in the *parametric JPDAF*.

F. The State Estimation

1) Assumption: The states of the targets conditioned on the past observations are mutually *independent*.[3]

In this case one needs the *marginal association probabilities*, which are obtained from the joint probabilities by summing over all the joint events in which the marginal event of interest occurs. This summation can be written as follows

$$\beta_{jt} := \sum_{\boldsymbol{\theta}:\theta_{jt}\in\boldsymbol{\theta}} \Pr(\boldsymbol{\theta}|\mathbf{Z}^k). \quad (37)$$

The state estimation equations are then exactly the same as in the standard PDAF.

2) Standard PDAF Estimation Equations: The PDAF updates the target state by combining the predicted state with the *combined innovation* multiplied by the filter gain $W(k)$. The *combined innovation* is the summation of the individual innovations, weighted by the *marginal association probabilities*, i.e.,

$$\boldsymbol{\nu}_k = \sum_{i=1}^{m_k} \beta_{i,k}\left[\mathbf{z}_{i,k} - \hat{\mathbf{x}}_{k|k-1}\right]. \quad (38)$$

The covariance associated with the updated state is

$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} + [\beta_{0,k}-1]\mathbf{W}_k\mathbf{S}_k\mathbf{W}_k' + \tilde{\mathbf{P}}_k, \quad (39)$$

where $\mathbf{S}_k$ is the innovation covariance and the *spread of the innovations* term is[4]

$$\tilde{\mathbf{P}}_k := \mathbf{W}_k\left[\sum_{i=1}^{m_k}\beta_{i,k}\left[\mathbf{z}_{i,k}-\hat{\mathbf{x}}_{k|k-1}\right]\cdot\left[\mathbf{z}_{i,k}-\hat{\mathbf{x}}_{k|k-1}\right]' - \boldsymbol{\nu}_k\boldsymbol{\nu}_k'\right]\mathbf{W}_k'. \quad (40)$$

[3]Considering the targets' states, given the past, as *correlated* — characterized by means, covariances as well as cross-covariances — leads to *coupled estimation* for the targets under consideration — the *JPDA Coupled Filter* (JPDACF) (See [10] Section 6.2.7 for details).

[4]This assumes that all the measurements have the same noise covariance and, hence, the same filter gain. The generalization to the case where each measurement has a different covariance is straightforward.

Since it is not known which of the validated measurements is correct, the term $\tilde{P}$, which is positive semidefinite, increases the covariance of the updated state — this is the effect of the measurement origin uncertainty.

G. A Modification of the JPDAF: Coupled Filtering

In Sections IV-D and IV-E the JPDAF was developed assuming that, conditioned on the past, the target states (and, thus, the target originated measurements) are *independently distributed*. Consequently, the joint association was followed by *decoupled* filtering of the targets' states — this is an approximation that simplifies the resulting algorithm.

For targets that "share" measurements (in the JPDAF sense) for several sampling times, a dependence of their estimation errors ensues and this can be taken into account by calculating the resulting error correlations.

The resulting algorithm, called *JPDA Coupled Filter (JPDACF),* does the filtering in a *coupled* manner for the targets with "common" measurements, yielding a covariance matrix with off-diagonal blocks — *cross-covariances* — that reflect the correlation between the targets' state estimation errors.

The conditional probability for a joint association event (31) becomes

$$\Pr(\boldsymbol{\theta}_k|\mathbf{Z}^k) = \frac{1}{c}\frac{\phi!\mu_F(\phi)}{m_k!V^{-\phi}} f_{t_{j_1},t_{j_2},\ldots}(\mathbf{z}_{j,k}, j:\tau_j=1) \cdot \prod_t (P_D^t)^{\delta_t}(1-P_D^t)^{1-\delta_t}, \quad (41)$$

where $f_{t_{j_1},t_{j_2},\ldots}$ is the joint pdf of the measurements of the targets under consideration; t_{j_1} is the target to which $z_{j_1}(k)$ is associated in event θ.

The joint probabilities are not reduced to the marginal association probabilities as in (37) for use in decoupled PDA filters. Instead, these joint probabilities are used directly in a *coupled filter*.

1) The JPDACF: Denote the stacked vector of the predicted states of the targets under consideration (assumed here to be 2) and the associated covariance matrix

$$\hat{\mathbf{x}}^{\mathcal{T}}_{k|k-1} = \begin{bmatrix}\hat{\mathbf{x}}^1_{k|k-1}\\ \hat{\mathbf{x}}^2_{k|k-1}\end{bmatrix}, \quad (42)$$

$$\mathbf{P}^{\mathcal{T}}_{k|k-1} = \begin{bmatrix}\mathbf{P}^{11}_{k|k-1} & \mathbf{P}^{12}_{k|k-1}\\ \mathbf{P}^{21}_{k|k-1} & \mathbf{P}^{22}_{k|k-1}\end{bmatrix}, \quad (43)$$

where $\hat{\mathbf{x}}^t$ and $\mathbf{P}^{tt}$ correspond to target t; $\mathbf{P}^{t_1t_2}$ is the *cross-covariance* between targets t_1 and t_2 (it will be zero before these targets become "coupled").

The coupled filtering is done as follows

$$\hat{\mathbf{x}}^{\mathcal{T}}_{k|k} = \hat{\mathbf{x}}^{\mathcal{T}}_{k|k-1} + \mathbf{W}^{\mathcal{T}}_k\sum_{\boldsymbol{\theta}}\Pr(\boldsymbol{\theta}_k|\mathbf{Z}^k)[\mathbf{z}^{\mathcal{T}}_k(\theta) - \hat{\mathbf{z}}^{\mathcal{T}}_{k|k-1}], \quad (44)$$

where

$$\mathbf{z}^{\mathcal{T}}_k(\boldsymbol{\theta}) = \begin{bmatrix}\mathbf{z}_{j_1(\boldsymbol{\theta}),k}\\ \mathbf{z}_{j_2(\boldsymbol{\theta}),k}\end{bmatrix}, \quad (45)$$

and $j_t(\theta)$ is the index of the measurement associated with target t in event θ at time k.

The filter gain in (44) is

$$\mathbf{W}_k^{\mathcal{T}} = \mathbf{P}_{k|k-1}^{\mathcal{T}} \left(\mathbf{H}_k^{\mathcal{T}}\right)' \left[\mathbf{H}_k^{\mathcal{T}} \mathbf{P}_{k|k-1}^{\mathcal{T}} \left(\mathbf{H}_k^{\mathcal{T}}\right)' + \mathbf{R}_k^{\mathcal{T}}\right]^{-1}, \quad (46)$$

where

$$\mathbf{H}_k^{\mathcal{T}} = \begin{bmatrix} \delta_\theta^1 \mathbf{H}_k^1 & 0 \\ 0 & \delta_\theta^2 \mathbf{H}_k^2 \end{bmatrix}, \quad (47)$$

$$\mathbf{R}_k^{\mathcal{T}} = \begin{bmatrix} \mathbf{R}_k^1 & 0 \\ 0 & \mathbf{R}_k^2 \end{bmatrix}, \quad (48)$$

are the (block diagonal) measurement matrix and noise covariance matrix, respectively, for the two targets under consideration. The (binary) detection indicator variables δ_θ^t above take care of the situation when only one of the targets is detected in event θ [129]. The predicted stacked measurement vector is

$$\hat{\mathbf{z}}_{k|k-1}^{\mathcal{T}} = \mathbf{H}_k^{\mathcal{T}} \hat{\mathbf{x}}_{k|k-1}^{\mathcal{T}}. \quad (49)$$

The update of the covariance of the (stacked) state is as in (39).

H. Extensions

1) Multiple Source Measurements: One can have an *unresolved (merged) measurement* from, e.g., two nearby targets.

The *JPDAM* — JPDA with Merged measurement includes a special model for a merged measurement (see [10] Section 6.4). A version of this is the *JPDAMCF* — JPDA with Merged measurement and Coupled Filter.

2) Splitting Target: A possible situation of interest is a platform that launches a weapon, where one has the situation of a *splitting target*.

The JPDAF has been extended to cover such a situation by using multiple models with the IMM configuration:

- there is a single non-maneuvering target.
- there is a single maneuvering target.
- the target splits into two targets.

This provides a "warm start" for the new target. For details, see Chapter 4 in [7].

3) The Coalescence Problem: Track coalescence can occur for the JPDAF if the tracks are close to each other for an extended time. Modifications of the JPDAF that counter the coalescence tendency are available [18], [35], [148], [170].

I. The JPDAF — Summary

Assumptions of the JPDAF:

- There are *several targets* to be tracked in the presence of *false measurements.*
- The number of targets is known.
- The track of each target has been initialized.
- The state equations of the targets are not necessarily the same.
- The validation regions of these targets can intersect and have "common" measurements.
- A target can give rise to at most one measurement — no multipath.
- The detection of a target occurs independently over time and from other targets according to a known probability.
- A measurement could have originated from at most one target (or none) — no unresolved measurements are considered here.
- The conditional pdf of each target's state given the past measurements is assumed *Gaussian* (a quasi-sufficient statistic that summarizes the past) and *independent across targets*, with means and covariances available from the previous cycle of the filter.

With the past summarized by an approximate sufficient statistic, the association probabilities are computed (only for the latest measurements) *jointly across the measurement and the targets*.

1) The JPDAF Steps:

- A validation matrix that indicates all the possible sources of each measurement is set up
- From this validation matrix all the *feasible joint association events* are obtained according to the rules
  - one source for each measurement,
  - one measurement (or none) from each target.
- The probabilities of these joint events are evaluated according to the assumptions
  - Target originated measurements are Gaussian distributed around the predicted location of the corresponding target's measurement,
  - False measurements are uniformly distributed in the surveillance region,
  - The number of false measurements is distributed according to
    - Poisson prior — *Parametric JPDAF,*
    - Diffuse prior — *Nonparametric JPDAF.*
- *Marginal* (individual measurement to target) association probabilities are obtained from the joint association probabilities.
- The target states are estimated by *separate (uncoupled)* PDA filters using these marginal probabilities.

V. Multiple Hypothesis Tracking

The MHT algorithm is described in a number of excellent papers [16], [76], [123], and books [10], [15], [101]. We will explain key concepts of the MHT algorithm through examples while keeping the mathematics to a minimum. The interested reader is encouraged to refer to books and papers mentioned in this section.

A number of terms such as target, track, track hypothesis, hypothesis, global hypothesis, association hypothesis, etc. [10], [15], [16], [76], [101] are commonly used in the MTT literature which are not clearly explained. Often a target and a track are used interchangeably. In order to remove such ambiguities we first explain these terms. To the best of our knowledge, a standard taxonomy does not exist.

1) *Target*: A target refers to the true object.
2) *True trajectory*: A true trajectory of a target is a time history of the true states $\{\mathbf{x}_k\}$ of the target.
3) *Track or track hypothesis*: A track represents an estimated trajectory of a target.
4) *Track label or identity* (ID): A distinct label or ID, usually a positive integer to uniquely identify a track.

5) *Compatible tracks*: A number of tracks are said compatible, if they do not have common measurements.
6) *Target tree*: In Kurien's track-oriented MHT (TOMHT) [76], an estimated target is represented by a target tree with a *resolved track*, a *root node*, and a number of tracks (branches) originating from the root node. The resolved track is a single branch from the first node to the current root node in a target tree. The first node is created when a new track is created using a measurement. The resolved tracks for a number of target trees are compatible and the branches in a target tree are not compatible. The location of the root node is moved forward by one scan when a new scan of measurements is processed.
7) *Gating*: A process which defines a volume in the measurement space to determine if a measurement can be associated with a predicted measurement corresponding to a predicted track.
8) *Association hypothesis*: An association hypothesis is generated when a measurement is associated with a predicted track.
9) *Global hypothesis or hypothesis*: A global hypothesis is a collection of compatible tracks representing a number of estimated trajectories.
10) *Assignment algorithm*: An algorithm which generates a number of compatible tracks using association between a number of tracks and measurements in one or more scans.

A. Single Hypothesis and Multiple Hypothesis Tracking

A simple multitarget tracking scenario is shown in Fig. 5, to illustrate the single and multiple hypothesis tracking methods where two resolved tracks T_1 and T_2 are present at scan $k-1$. We have a single global hypothesis $G_1 = \{T_1, T_2\}$ at scan $k-1$ and at scans k and $k+1$, the tracker receives measurement sets $\{\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3\}$ and $\{\mathbf{z}_4, \mathbf{z}_5\}$, respectively. It is not known if a measurement is from a target or due to clutter. Secondly, it is not known which measurement originates from which target. Thirdly, it is not known if missed detection events have occurred due to less than unity probability of detection. This phenomenon is known as the *measurement origin uncertainty.* In order to limit the number of candidate measurement-to-

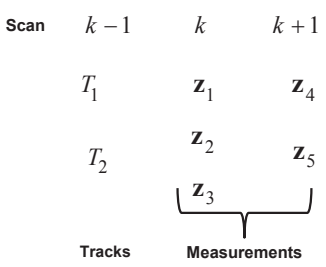

Fig. 5. A multitarget tracking scenario.

track associations (M2TAs), a data association based MTT algorithm uses *gating* [10], [15], [76], [123]. A *coarse gating* followed by a *fine gating* (ellipsoidal gating) is commonly used. A *coarse gating* is based on rectangular gating with a large value along each measurement coordinate. The coarse gating eliminates many unlikely M2TAs for computational efficiency. In Fig. 6, we assume that the measurements are 2D position measurements. The tracks T_1 and T_2 at scan $k-1$ are predicted to scan time k to obtain predicted tracks $\bar{T}_1$ and $\bar{T}_2$, respectively. Fig. 6 shows that measurements $\{\mathbf{z}_2, \mathbf{z}_3\}$ and $\{\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3\}$ can be associated with predicted tracks $\bar{T}_1$ and $\bar{T}_2$, respectively, by gating. Fig. 7 shows the generation of

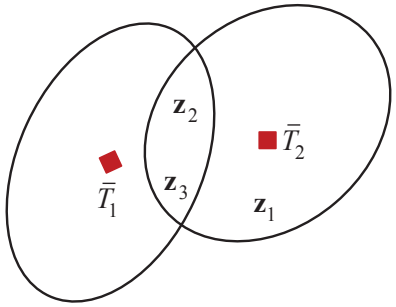

Fig. 6. Measurement-to-track association at scan k.

five M2TA hypotheses based on gating in Fig. 6, two missed detection hypotheses, and three *new tracks* T_3, T_4 and T_5 corresponding to measurements $\mathbf{z}_1, \mathbf{z}_2$ and $\mathbf{z}_3$, respectively. We note that two estimated targets are represented by two target trees with root nodes at scan $k-1$. The track from the first node at scan 1 to the root node at scan $k-1$ represents a *resolved track* for T_1 or T_2. We shall see in later discussion that this representation is used in the TOMHT first proposed by Kurien [76]. Table I shows ten possible global hypotheses corresponding to track hypotheses in Fig. 7. It is a coincidence that the number of track hypotheses and global hypotheses are the same for this scenario. In the *single*

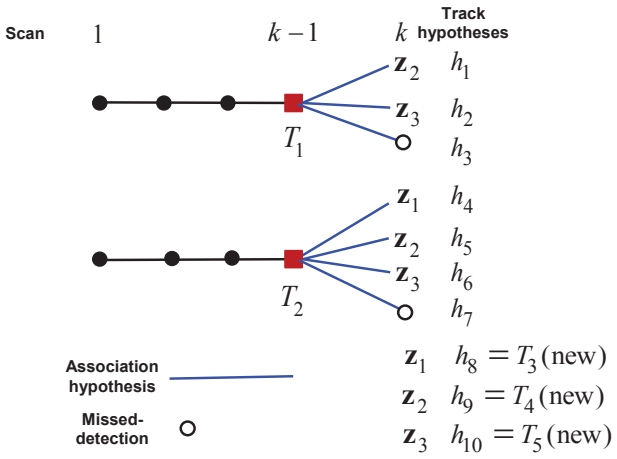

Fig. 7. Generation of track hypotheses at scan k.

hypothesis tracking (SHT) algorithm such as the GNN, the *best global hypothesis* (BGH) based on the maximum *total track score* or the minimum *total cost* is selected using the existing set of tracks and the current scan of measurements and all other M2TAs are discarded for future consideration. Suppose G_3 is the best global hypothesis. Then only the M2TAs $\{T_1 - \mathbf{z}_3, T_2 - \mathbf{z}_1, T_4 - \mathbf{z}_2\}$ contained in G_3 are kept and all other M2TAs are discarded in the GNN algorithm.

An MHT algorithm uses a *deferred decision logic* [15], [16] by allowing more than one scan of measurements to be used in the M2TA process. It is hoped that measurements in more than one scan can provide more accurate M2TA than those in a single scan. The assignment of measurements in one scan to one set of tracks is known as the *2D assignment* problem [10], [15], [117]. The assignment of measurements in $s-1, s>3$ scans to a set of tracks is known as the $s-$D assignment (also known as the multi-frame assignment (MFA) or multi-dimensional assignment (MDA)) problem [10], [15], [16], [40], [116], [117].

The number of association hypotheses or tracks in a TOMHT can grow exponentially as measurements in scans are processed sequentially. An MHT algorithm usually uses a number of techniques such as *clustering, gating, N-scan pruning*, and *track-score based pruning* to limit this exponential growth. These will be described in Section V-C.

For simplicity, the example described here has not considered false alarm hypotheses (i.e. all measurements are accounted for in track hypotheses), nor have we considered undetected target birth hypotheses. The same is true of the example discussed later in Sec. VII.C. As discussed in [30], it is beneficial to decouple data association and track extraction processes; the latter discards spurious returns. Thus, it is sufficient to consider hypotheses that account for all returns. Further, a recent MHT generalization that accounts for undetected target births is in [31].

Additionally, again for simplicity, we have not considered target death events. Generally, MHT hypothesis generation logic spawns only a target missed detection or target death track hypothesis, the latter after a sufficient number of missed detections.

B. Types of MHT Algorithms

There are two types of MHT, the hypothesis-oriented MHT (HOMHT) [10], [15], [123] and TOMHT [10], [15], [76], [101], [139], [149], [151]. Reid first proposed the HOMHT [123]. There are two different types of the TOMHT, tree based [29], [76], [101], [149] and non-tree based [138], [151]. Both types of TOMHT solve the same binary MDA problem and can yield the same binary (0-1) solutions. In the TOMHT approach, measurements in the last *s*-1 scans are associated with a number of tracks in the previous scan. The difference is due to the representation by which tracks are represented. Kurien first formulated a computationally efficient version of the tree based TOMHT [76] in which a hypothesized target is represented by a target tree. The non-tree based TOMHT [10], [138], [151] does not use a target tree. When a hypothesized target is represented by a target tree, $N-$scan pruning [15], [29], [76], [101] can be performed to reduce the number of tracks. The non-tree based TOMHT cannot perform the $N-$scan pruning. Therefore, the total number of tracks will be different in these two TOMHT implementations. Subsections V-C and V-D present tree based TOMHT and non-tree based TOMHT, respectively.

We model targets as points and assume that a tracker receives measurements in scans. Each scan contains the scan time, sensor state related information (e.g. sensor position, velocity, etc.), measurements and associated measurement error covariances, and sensor probability of detection. A conventional tracker is based on the fundamental assumption that a point target generates at most one measurement per scan [10], [15]. Multiple detections per scan for a point target arising in the over the horizon radar (OTHR) tracking problem [57], [121], [139] requires advanced algorithms where this fundamental assumption can be relaxed.

C. Tree Based TOMHT

The HOMHT keeps a number of global hypotheses between consecutive scans whereas tree based TOMHT only maintains a number of target trees, each containing a number of tracks which are not compatible. In tree based TOMHT, the best global hypothesis is formed from the existing set of tracks and the $N-$scan pruning and track-score based pruning are used to limit the number of tracks from growing exponentially. In tree based TOMHT, there are many more tracks than the number of tracks in the best global hypothesis. For large-scale tracking problems, there may be several thousand comparable global hypotheses from several hundred tracks in a cluster [16]. From practical experience, several hundred tracks can be easily handled by tree based TOMHT. Therefore, this tree based TOMHT has computational advantage over HOMHT. Based on software architecture development, maintenance, debugging, and cost effectiveness, most tracking groups at present use the tree based TOMHT.

A block diagram in Fig. 8 shows various processing steps of a tree based TOMHT. When the first scan is received, the measurements are partitioned into a number of clusters [123] first using a coarse method and then using the location and measurement error covariances. The use of clustering in MTT was first proposed by Reid [123] to partition the tracking problem to a number of sub-problems so that an MHT algorithm can be applied to each cluster for computational efficiency. For each measurement a new track is initiated using a single-point (SP) track initiation algorithm [8], [97], [98], [100] which calculates the initial state estimate and associated covariance. Additionally, the track score for a new track, which generates a new target tree is also calculated. A sensor usually collects measurements in a region of the measurement space known as the dwell or scan volume. For a radar measuring range and azimuth in a plane, the dwell volume (area in this case) can be specified by the minimum and maximum values of the range and azimuth. Thus a sensor has no information about targets outside the dwell volume.

Assumptions:

1) The number of FAs in the dwell volume is Poisson distributed [10]. Let λ_{FA} denote the expected number of

TABLE I
TEN POSSIBLE GLOBAL HYPOTHESES.

Global hypothesis	Structure of a global hypothesis
G_1	$\{T_1 - \mathbf{z}_2, T_2 - \mathbf{z}_1, T_5 - \mathbf{z}_3\}$
G_2	$\{T_1 - \mathbf{z}_2, T_2 - \mathbf{z}_3, T_3 - \mathbf{z}_1\}$
G_3	$\{T_1 - \mathbf{z}_3, T_2 - \mathbf{z}_1, T_4 - \mathbf{z}_2\}$
G_4	$\{T_1 - \mathbf{z}_3, T_2 - \mathbf{z}_2, T_3 - \mathbf{z}_1\}$
G_5	$\{T_1 - \varnothing, T_2 - \mathbf{z}_1, T_4 - \mathbf{z}_2, T_5 - \mathbf{z}_3\}$
G_6	$\{T_1 - \varnothing, T_2 - \mathbf{z}_2, T_3 - \mathbf{z}_1, T_5 - \mathbf{z}_3\}$
G_7	$\{T_1 - \varnothing, T_2 - \mathbf{z}_3, T_3 - \mathbf{z}_1, T_4 - \mathbf{z}_2\}$
G_8	$\{T_1 - \mathbf{z}_2, T_2 - \varnothing, T_3 - \mathbf{z}_1, T_5 - \mathbf{z}_3\}$
G_9	$\{T_1 - \mathbf{z}_3, T_2 - \varnothing, T_3 - \mathbf{z}_1, T_4 - \mathbf{z}_2\}$
G_{10}	$\{T_1 - \varnothing, T_2 - \varnothing, T_3 - \mathbf{z}_1, T_4 - \mathbf{z}_2, T_5 - \mathbf{z}_3\}$

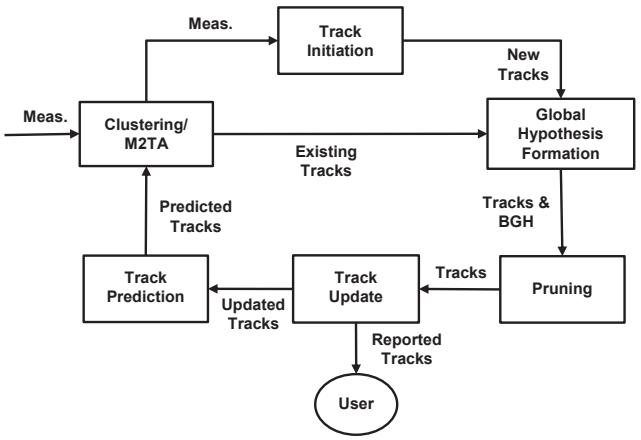

Fig. 8. Processing steps of a TOMHT.

FAs per unit volume of the measurement space, known as the spatial density of FAs.

2) The number of new targets appearing in the dwell volume is also Poisson distributed [10]. Let λ_{new} denote the expected number of new targets per unit volume of the measurement space, known as the spatial density of new targets.

1) Track Score: Tracks can have different number of detections and missed detections. In order to treat the tracks in a normalized manner, the likelihood ratio (LR), normalized by the FA probability density is used [9], [15], which is dimensionless. For computational convenience, the logarithm of the LR (LLR) is used for the track score [9], [15]. High and low track scores represent high and low quality tracks, respectively. Let LLR_k denote the track score at scan k. Assuming that measurements at different scans are independent given the state, the LLR_k is related to the LLR_{k-1} by

$$\text{LLR}_k = \text{LLR}_{k-1} + \text{L}\Delta\text{LR}_k, \tag{50}$$

where $\text{L}\Delta\text{LR}_k$ is the incremental log-likelihood ratio or incremental track score at scan k. As described in Section V-A, three possible cases arise; a measurement $\mathbf{z}_k$ can be associated with a track, a track can have a missed detection, and a new track corresponding to $\mathbf{z}_k$ can be created. Let P_{D} and P_{G} denote the probability of detection and gate probability, respectively. The LLR for a new track is given by [9], [76]

$$\text{LLR}_{\text{new}} = \log \frac{\lambda_{\text{new}}}{\lambda_{\text{FA}}}. \tag{51}$$

Let $\mathbf{Z}^{k-1}$ denote the measurements associated with a track up to time t_{k-1}. For generality, we assume that $\mathbf{Z}^{k-1}$ includes detections and missed detections. Then the $\text{L}\Delta\text{LR}_k$s for the association of $\mathbf{z}_k$ with the track and missed detection event, respectively, are given by [9], [76]

$$\text{L}\Delta\text{LR}_k = \log \frac{P_{\text{D}} p(\mathbf{z}_k|\mathbf{Z}^{k-1})}{\lambda_{\text{FA}}}, \tag{52}$$

$$\text{L}\Delta\text{LR}_k = \log(1 - P_{\text{D}} P_{\text{G}}). \tag{53}$$

2) Best Global Hypothesis Generation using MFA and Pruning: As seen in Section V-A, the number of tracks can grow exponentially as scans of measurements are processed sequentially. This can lead to a serious computational problem when a few hundred targets are present in the surveillance area. In order to have practical solutions for real-world tracking problems, the tree based TOMHT uses a number of pruning methods to delete tracks with low track score while keeping tracks with high scores. Next we describe the formation of the best global hypothesis using the example shown in Fig. 9, described in [101]. The notations and symbols used here are slightly different from those in [101]. This example is different from the example in Fig. 6.

We assume that at scan $k-1$ we have two resolved tracks T_1, T_2 and at scans k and $k+1$, the tracker receives measurement sets $\{\mathbf{z}_1, \mathbf{z}_2\}$ and $\{\mathbf{z}_3\}$, respectively. Following the procedure described in Section V-A, ten tracks $\{h_1, h_2, ..., h_{10}\}$ are generated at scan $k+1$. The problem shown in Fig. 9 is a 3-dimensional ($s = 3$) assignment problem. Let N and M denote the number of tracks at scan $k+1$ and the sum of the number of resolved tracks at the root node and the number of measurements in the last $s-1$ scans, respectively. For our example, $M = 5$, $N = 10$. Let $\mathbf{a}$ and $\mathbf{u}$ be $N-$ dimensional column vectors, where each element of $\mathbf{a}$ is zero or one and $u_i = \text{LLR}_i$, $i = 1, 2, ...N$. Let $\mathbf{b}$ be an $M-$ dimensional column vector where each element of $\mathbf{b}$ is one. We refer to $\mathbf{a}$ and $\mathbf{u}$ as the assignment and utility vectors, respectively. The best global hypothesis is determined by solving the *binary programing* problem where $\{a_i \in [0, 1]\}$ are determined by maximizing the total utility as shown in Fig. 10.

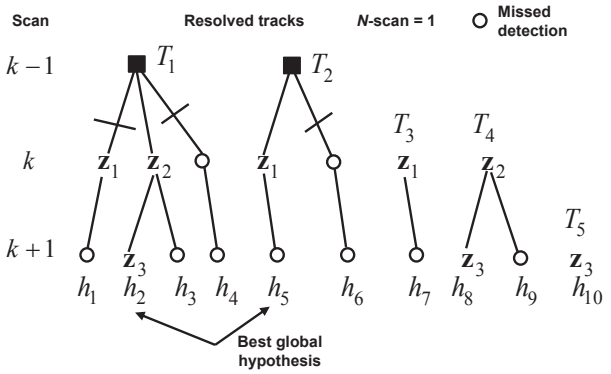

Fig. 9. Formation of best global hypothesis and N-scan pruning.

$$\arg\max_{\mathbf{a}} \ \mathbf{u}'\mathbf{a}$$

subject to $\mathbf{A}\mathbf{a} = \mathbf{b}$, with $b_j = 1$, for $j = 1, 2, ..., M$, and

$$a_i \in [0,1], \text{ for } i = 1, 2, ..., N.$$

$$\mathbf{A} = \begin{array}{c} \begin{matrix} h_1 & h_2 & h_3 & h_4 & h_5 & h_6 & h_7 & h_8 & h_9 & h_{10} \end{matrix} \\ \left[\begin{array}{cccccccccc} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \end{array}\right] \end{array} \begin{array}{l} \\ \left.\begin{array}{l} \\ \\ \end{array}\right\} \text{for tracks } T_1, T_2 \\ \left.\begin{array}{l} \\ \\ \\ \end{array}\right\} \text{for meas. } \mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3 \end{array}$$

Fig. 10. Multi-frame assignment problem.

The Lagrangian relaxation algorithm [117], [119], [120] and approximate linear programming (LP) [29], [145] can be used to solve the MFA problem.

Suppose the tracks h_2 and h_5 are included in the BGH and we choose N-scan as 1. In N-scan pruning, we move one scan back from the current scan $k+1$ to scan k and for target trees T_1 and T_2, we delete branches not included in the BGH. Thus, tracks h_1, h_4, h_6, h_7, h_8, and h_9 are deleted by N-scan pruning. In addition to N-scan pruning, track score based pruning is used to delete tracks with low track score.

In order to remove spurious tracks and keep good tracks, the status of a track is specified by *new, tentative*, and *confirmed*. A number of track-confirmation logic described in [15], [30] are used to keep good tracks. Generally, track management is performed in a sliding window and is either logic-based or score-based, the latter making use of the sequential probability ratio test (SPRT) for quickest change detection [15].

D. Non-tree Based TOMHT

Details of non-tree based TOMHT are presented in [10], [138], [151]. There are a number of variants of this type of non-tree based TOMHT. To illustrate this approach, we consider the example in Fig. 9 and use the method described in [151]. This multi-frame assignment problem is a 3D assignment problem. Following the approach in [151], we have depicted this 3D assignment problem in Fig. 11, where two resolved tracks T_1 and T_2 at scan $k-1$ and three measurements $\{\mathbf{z}_1, \mathbf{z}_2\}$ and $\{\mathbf{z}_3\}$ at scans k and $k+1$ are shown. Secondly, a dummy track corresponding to an "extraneous" measurement (new target or false alarm) at scan $k-1$ and dummy measurements corresponding to missed detections at scans k and $k+1$ are also shown in Fig. 11.

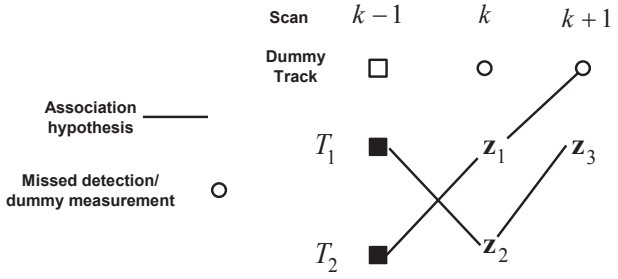

Fig. 11. Multi-frame assignment problem in non-tree based TOMHT.

Track initiation, track score computation, and track confirmation in this type of non-tree based TOMHT are similar to those in tree based TOMHT. An $s-$D assignment is used to determine best global hypothesis. Two tracks present in the best global hypothesis are shown in Fig. 11 by solid lines originating from resolved tracks T_1 and T_2.

E. Track Filtering

A single model filter (e.g KF, EKF, UKF, PF, PFF) and multiple model filter (e.g. IMM, VS-IMM) for a non-maneuvering and maneuvering target, respectively, are used in an MTT system. The dynamic and measurement models in a filter can be linear or nonlinear. Some of these filters have been described in Subsections II-C, II-D, and II-E. A detailed overview of the dynamic and measurement models and filters for non-maneuvering and maneuvering targets are presented in [101].

F. Applications of MHT

The MHT has been successfully used for solving many real-world problems in ground target tracking, maritime tracking, air target tracking, missile defense systems, computer vision systems, video tracking, persistence surveillance, and space object tracking (SOT) [15], [16], [101]. Computer hardware and software have advanced significantly during the last two decades. As a result, large-scale real-world problems involving thousands of targets can now be solved by the TOMHT using high performance computing (HPC) and cluster computing. As mentioned in [16], due to military applications and company proprietary policies, many of these studies are not available in the open literature.

G. Future Work

Future areas of research include multiple detection systems in which multiple measurements for a target can arise in a scan. These problems arise in over the horizon radar (OTHR) tracking [57], [121], [139], tracking using high range resolution radar (HRR), and passive coherent location system (PCL) [152]. Multiple detections per scan can also arise for an extended target [81]. Large-scale real-world problems which will involve HPC and cluster computing are SOT [102] and giga-pixel video surveillance system [2]. An important area of research is the comparative evaluation of the TOMHT and RFS based algorithms. Research in this area is quite limited [147]. Optimal solution to a complex MTT system is not possible. A key goal of an advanced MTT system is to have numerically efficient, near-optimal, robust, and scalable algorithms and software.

VI. The Random Finite Set Approach

List of mathematical symbols:

h^X	multitarget exponential: $\prod_{\mathbf{x}\in X} h(\mathbf{x})$, $h^{\emptyset}=1$
X_k	(set-valued) multitarget state at time k
Z_k	(set-valued) multitarget observation at time k
Z^k	multitarget observation history $(Z_1, Z_2, ..., Z_k)$
$\pi(\cdot)$	multi-target density
$v(\cdot)$	Probability Hypothesis Density (PHD)
$\rho(\cdot)$	cardinality distribution
$\mathbb{L}_k$	space of labels for targets born at time k
$\mathbb{L}_{0:k}$	space of labels for targets up to time k
$\mathcal{L}(X)$	the set of labels of X
$\Delta(X)$	distinct label indicator: $\delta_{\lvert X\rvert}[\lvert\mathcal{L}(X)\rvert]$
θ	association map taking $\mathbb{L}_{0:k}$ to $\{0,1,...,\lvert Z_k\rvert\}$ such that $\theta(\ell)=\theta(\ell')>0$ implies $\ell=\ell'$
Θ_k	space of association maps at time k
$\Theta_k(L)$	subset of Θ_k with domain L

The RFS approach represents the *multitarget state* as a finite set of single-target states, and the MTT problem is formulated as a dynamic multitarget state estimation problem, analogous to the single-target case in Section II.

A. Random Finite Set

An RFS X, of $\mathcal{X}$, is a random variable taking values in $\mathcal{F}(\mathcal{X})$, the collection of all finite subsets of $\mathcal{X}$. While $\mathcal{F}(\mathcal{X})$ does not inherit the usual Euclidean notion of probability density from $\mathcal{X}$, a measure-theoretic notion of probability density on $\mathcal{F}(\mathcal{X})$ is available [154]. However, we adopt the *Finite Set Statistic* (FISST) notion of density since it is convenient and by-passes measure theoretic constructs [54], [86].

The FISST density of an RFS X is a non-negative function π on $\mathcal{F}(\mathcal{X})$ such that for any region $S\subseteq\mathcal{X}$,

$$\Pr(X\subseteq S)=\int_S \pi(X)\delta X,$$

where the integral above is a *set integral* defined by [54], [84]

$$\int_S \pi(X)\delta X=\sum_{i=0}^{\infty}\frac{1}{i!}\int_{S^i}\pi(\{\mathbf{x}_1,...,\mathbf{x}_i\})d(\mathbf{x}_1,...,\mathbf{x}_i),$$

i.e. the set integral of the FISST density over a region S, yields the probability that X is contained in S. Although π is not a probability density, the function defined by $\pi(X)K^{|X|}$ is, where K denotes the unit of hyper-volume on $\mathcal{X}$ [154].

A *Bernoulli* RFS X has probability $1-r$ of being empty, and probability r of being a singleton whose element is distributed according to a probability density p (on $\mathcal{X}$). The density of a Bernoulli (RFS) is given by

$$\pi(X)=\begin{cases}1-r, & X=\emptyset,\\ rp(\mathbf{x}), & X=\{\mathbf{x}\}.\end{cases}$$

A *multi-Bernoulli* RFS is a union of independent Bernoulli's.

The cardinality distribution of an RFS is defined by

$$\rho(n)=\Pr(|X|=n).$$

An *i.i.d. cluster* RFS X has elements i.i.d. according to a probability density p (on $\mathcal{X}$), and is completely characterized by ρ and p [36]. Its density is given by

$$\pi(\{\mathbf{x}_1,...,\mathbf{x}_n\})=n!\rho(n)\prod_{i=1}^{n}p(\mathbf{x}_i),$$

with $\pi(\emptyset)=\rho(0)$. A *Poisson* RFS is a special case of *i.i.d. cluster* RFS with Poisson cardinality.

The first moment of an RFS is the *Probability Hypothesis Density* (PHD) also known as the *intensity function* [84], [154]. The PHD is a non-negative function v (on $\mathcal{X}$) whose integral over any region $S\subseteq\mathcal{X}$ gives the expected number of elements of the RFS that are in S, i.e.

$$\mathbb{E}\left[|X\cap S|\right]=\int_S v(\mathbf{x})d\mathbf{x}. \tag{54}$$

The PHD is computed from the multitarget density by [84]

$$v(\mathbf{x})=\int\pi(\{\mathbf{x}\}\cup X)\delta X. \tag{55}$$

The local maxima of the PHD are points in $\mathcal{X}$ with the highest local concentration of expected number of elements. Intuitively, we can use $\hat{n}=\mathbb{E}\left[|X|\right]$ or $\hat{n}=\arg\max_n\rho(n)$ as the estimated number of targets, and the $\hat{n}$ highest local maxima of the PHD as the estimated target states.

When a multitarget state, with *prior* density π, is observed as Z (e.g. a set of points, an image, or a function) modelled by the *likelihood function* $\pi(Z|X)$, all information about the multitarget state given the observation is contained in the multitarget *posterior* density, given by Bayes rule (cf. (1))

$$\pi(X|Z)=\frac{\pi(Z|X)\pi(X)}{\int\pi(Z|X)\pi(X)\delta X}. \tag{56}$$

Bayes optimal multitarget estimators can be formulated by minimizing the Bayes risk as in the single-target case. One such estimator is the Marginal multitarget estimator [86]: $\hat{X}=\arg\sup_{X:|X|=\hat{n}}\pi(X|Z)$, where $\hat{n}=\arg\max_n\rho(n|Z)$.

B. Multitarget State Space Model

In a standard multitarget transition model (see Section III), at time $k-1$, each target $\mathbf{x}_{k-1}$ of a multitarget state X_{k-1}, generates a Bernoulli RFS $S_{k|k-1}(\mathbf{x}_{k-1})$ at time k. New targets at time k are modeled by an RFS of spontaneous births

Γ_k. Thus, the multitarget state X_k generated by X_{k-1} is given by the *multitarget state transition equation* (cf. (2))

$$X_k = \bigcup_{\mathbf{x}_{k-1} \in X_{k-1}} S_{k|k-1}(\mathbf{x}_{k-1}) \cup \Gamma_k. \tag{57}$$

In general the multitarget transition equation can be described by a Markov *multitarget transition density* (cf. (3))

$$\phi_{k|k-1}(X_k|X_{k-1}), \tag{58}$$

i.e. the probability density that a given multitarget state X_{k-1} evolves to X_k. The multitarget transition density captures the underlying models of target motion, births and deaths.

In a standard multitarget observation model (see Section III), each target $\mathbf{x}_k$ of a multitarget state X_k, generates a Bernoulli RFS $D_k(\mathbf{x}_k)$. The observation Z_k generated by X_k is given by the *multitarget observation equation* (cf. (4))

$$Z_k = \bigcup_{\mathbf{x}_k \in X_k} D_k(\mathbf{x}_k) \cup F_k, \tag{59}$$

where F_k is an RFS of false detections. In general the multitarget observation model can be expressed as the *multitarget likelihood function* (cf. (5))

$$\varphi_k(Z_k|X_k), \tag{60}$$

i.e. the likelihood that the observation Z_k is generated by the multitarget state X_k. In a standard observation model, the multitarget observation likelihood captures underlying models of target detections, observation noise, and FAs. Unlike traditional techniques, this framework accommodates non-homogeneous, non-Poisson FAs, and state-dependent probability of detection in a principled way. Non-standard observations such as images, functions, etc. can also be described by the multitarget observation likelihood.

C. Multitarget Bayes Recursion

All information about the multitarget state history to time k is encapsulated in the *multitarget posterior* density $\pi_{0:k}(\cdot|Z_{1:k})$, which can be computed recursively from an initial prior π_0, via the multitarget Bayes recursion (cf. (6))

$$\pi_{0:k}(X_{0:k}|Z^k) \propto \varphi_k(Z_k|X_k)\phi_{k|k-1}(X_k|X_{k-1})\pi_{0:k-1}(X_{0:k-1}|Z^{k-1}). \tag{61}$$

Target trajectories are accommodated by incorporating a label in each target's state vector [54], [86], [163], [166]. The multitarget posterior, thus contains all information on the RFS of target trajectories, given the observation history.

The *multitarget filtering* density $\pi_k(\cdot|Z^k)$, is a marginal of the posterior density at time k, which is of interest for on-line multitarget tracking. The multitarget filtering density can be computed recursively using the multitarget Bayes prediction and update equations (cf. (7), (8))

$$\pi_{k|k-1}(X|Z^{k-1}) = \int \phi_{k|k-1}(X|Y)\pi_{k-1}(Y|Z^{k-1})\delta Y, \tag{62}$$

$$\pi_k(X|Z^k) = \frac{\varphi_k(Z_k|X)\pi_{k|k-1}(X|Z^k)}{\int \varphi_k(Z_k|Y)\pi_{k|k-1}(Y|Z^k)\delta Y}. \tag{63}$$

A generic particle implementation of the multitarget recursions (61) and (62)-(63) was given in [154]. Multitarget trackers based on incorporating labels in the target states include the generalized labeled multi-Bernoulli filter [163], [164], which solves the filtering recursion (62)-(63) analytically, and the particle marginal Metropolis-Hasting tracker [166], which simulates the posterior (61). Algorithms that only estimate the multitarget state include the PHD, Cardinalized PHD and multi-Bernoulli filters, [84]–[86], [159], [160], which are analytic approximations of the filtering recursion (62)-(63).

D. The PHD Filter

The PHD filter is a computationally inexpensive approximation of the multitarget Bayes filter derived by Mahler using FISST [84]. An alternative derivation of the PHD filter based on classical point process theory was given in [142], while an intuitive interpretation was given in [47].

Instead of propagating the multitarget filtering density $\pi_k(\cdot|Z^k)$, the PHD filter propagates its first moment, the filtered PHD $v_k(\cdot|Z^k)$. In addition to the standard multitarget state space model with Poisson FAs, the PHD recursion assumes that the updated and predicted multitarget RFSs are Poisson. For compactness, we drop the dependence on Z^k, and denote by $\langle \alpha, \beta \rangle$ the *inner product* $\int \alpha(\zeta)\beta(\zeta)d\zeta$ when α, β are functions, or $\sum_{\ell=0}^{\infty} \alpha(\ell)\beta(\ell)$ when α, β are sequences.

The PHD recursion consists of a prediction and an update

$$v_{k|k-1}(\mathbf{x}) = \langle P_{S,k|k-1} f_{k|k-1}(\mathbf{x}|\cdot), v_{k-1} \rangle + \gamma_k(\mathbf{x}), \tag{64}$$

$$v_k(\mathbf{x}) = [1 - P_{D,k}(\mathbf{x})]\, v_{k|k-1}(\mathbf{x}) + \sum_{\mathbf{z} \in Z_k} \frac{\psi_k(\mathbf{z};\mathbf{x}) v_{k|k-1}(\mathbf{x})}{\lambda_{F,k} + \langle \psi_k(\mathbf{z};\cdot), v_{k|k-1} \rangle}, \tag{65}$$

where γ_k is the PHD of the RFS of new targets,

$$\psi_k(\mathbf{z};\mathbf{x}) = P_{D,k}(\mathbf{x}) g_k(\mathbf{z}|\mathbf{x}) / p_{F,k}(\mathbf{z}) \tag{66}$$

is the detection-to-FA ratio of $\mathbf{z}$ given a target $\mathbf{x}$, $\lambda_{F,k}$ is expected number of FAs at time k, and $p_{F,k}$ is the probability density of each FA.

The PHD recursion (64)-(65) admits a closed form solution called the Gaussian Mixture PHD (GM-PHD) filter [155] under the *linear Gaussian multitarget* model: linear Gaussian single-target model, i.e. (9)-(10); constant survival and detection probabilities, i.e. $P_{S,k|k-1}(\mathbf{x}) = P_{S,k|k-1}$, and $P_{D,k}(\mathbf{x}) = P_{D,k}$ (Gaussian mixture $P_{S,k|k-1}(\mathbf{x})$ and $P_{D,k}(\mathbf{x})$ can also be accommodated); and Gaussian mixture birth PHD

$$\gamma_k(\mathbf{x}) = \sum\nolimits_{j=1}^{J_{\Gamma,k}} w_{\Gamma,k}^{(j)} \mathcal{N}(\mathbf{x}; \mathbf{m}_{\Gamma,k}^{(j)}, \mathbf{P}_{\Gamma,k}^{(j)}).$$

In this case, if v_{k-1} is a Gaussian mixture of the form

$$v_{k-1}(\mathbf{x}) = \sum\nolimits_{i=1}^{J_{k-1}} w_{k-1}^{(i)} \mathcal{N}(\mathbf{x}; \mathbf{m}_{k-1}^{(i)}, \mathbf{P}_{k-1}^{(i)}), \tag{67}$$

then the predicted PHD to time k is given by

$$v_{k|k-1}(\mathbf{x}) = \gamma_k(\mathbf{x}) + \sum_{i=1}^{J_{k-1}} w_{k|k-1}^{(i)} \mathcal{N}(\mathbf{x}; \mathbf{m}_{k|k-1}^{(i)}, \mathbf{P}_{k|k-1}^{(i)}), \tag{68}$$

where $w_{k|k-1}^{(i)} = P_{S,k|k-1} w_{k-1}^{(i)}$, and $\mathbf{m}_{k|k-1}^{(i)}$, $\mathbf{P}_{k|k-1}^{(i)}$ are given by the Gaussian sum filter prediction (13), (14), respectively. If we rewrite $v_{k|k-1}$ as

$$v_{k|k-1}(\mathbf{x}) = \sum\nolimits_{i=1}^{J_{k|k-1}} w_{k|k-1}^{(i)} \mathcal{N}(\mathbf{x}; \mathbf{m}_{k|k-1}^{(i)}, \mathbf{P}_{k|k-1}^{(i)}) \quad (69)$$

then, the updated PHD at time k is given by

$$\begin{aligned} v_k(\mathbf{x}) &= (1 - P_{D,k}) v_{k|k-1}(\mathbf{x}) \\ &\quad + P_{D,k} \sum_{\mathbf{z} \in Z_k} \sum_{i=1}^{J_{k|k-1}} w_k^{(i)}(\mathbf{z}) \mathcal{N}(\mathbf{x}; \mathbf{m}_k^{(i)}(\mathbf{z}), \mathbf{P}_k^{(i)}) \end{aligned} \quad (70)$$

where

$$w_k^{(i)}(\mathbf{z}) = \frac{w_{k|k-1}^{(i)} q_k^{(i)}(\mathbf{z})}{\lambda_{F,k} + P_{D,k} \sum_{\ell=1}^{J_{k|k-1}} w_{k|k-1}^{(\ell)} q_k^{(\ell)}(\mathbf{z})} \quad (71)$$

$$q_k^{(i)}(\mathbf{z}) = \mathcal{N}(\mathbf{z}; \mathbf{H}_k \mathbf{m}_{k|k-1}^{(i)}, \mathbf{S}_{k|k-1}^{(i)}) / p_{K,k}(\mathbf{z}), \quad (72)$$

with $\mathbf{m}_k^{(i)}(\mathbf{z})$, $\mathbf{P}_k^{(i)}$, $\mathbf{S}_{k|k-1}^{(i)}$ given by the Gaussian sum filter update (15)-(18) respectively.

Mixture reduction by pruning negligible components and merging similar components are needed to manage the growing the number of components [155]. Multitarget state estimation in the GM-PHD filter involves first estimating the number of targets from the sum of the weights, and then extracting the corresponding number of components with the highest weights from the PHD as state estimates. Alternatively, we can choose the means of components whose weights exceed a prescribed threshold.

For highly nonlinear problems, v_k can be approximated by a set of weighted particles $\{(w_k^{(i)}, \mathbf{x}_k^{(i)})\}_{i=1}^{L_k}$. A generic particle PHD filter is given in the algorithm below [154]. Convergence results similar result to that in [32] also hold for the particle PHD filter under standard assumptions [154], [26], [66]. Multitarget state estimation for the particle PHD filter requires the clustering of particles into groups, which involves additional processing. Techniques such as the auxiliary particle PHD filter [167], the measurement driven particle PHD filter [127] provide partial solutions to this problem.

Algorithm: Particle PHD filter

For $i = 1, ..., L_{k-1}$, sample $\widetilde{\mathbf{x}}_k^{(i)} \sim q_k(\cdot | \mathbf{x}_{k-1}^{(i)}, Z_k)$ and compute

$$\widetilde{w}_{k|k-1}^{(i)} = \frac{P_{S,k|k-1}(\mathbf{x}_{k-1}^{(i)}) f_{k|k-1}(\widetilde{\mathbf{x}}_k^{(i)} | \mathbf{x}_{k-1}^{(i)})}{q_k(\widetilde{\mathbf{x}}_k^{(i)} | \mathbf{x}_{k-1}^{(i)}, Z_k)} w_{k-1}^{(i)}.$$

For $i = 1, ..., J_k$, sample $\widetilde{\mathbf{x}}_k^{(i+L_{k-1})} \sim r_k(\cdot | Z_k)$ and compute

$$\widetilde{w}_{k|k-1}^{(i+L_{k-1})} = \frac{\gamma_k(\widetilde{\mathbf{x}}_k^{(i+L_{k-1})})}{J_k r_k(\widetilde{\mathbf{x}}_k^{(i+L_{k-1})} | Z_k)}.$$

For each $z \in Z_k$, compute

$$C_k(\mathbf{z}) = \sum\nolimits_{j=1}^{L_{k-1}+J_k} \psi_k(\mathbf{z}; \widetilde{\mathbf{x}}_k^{(j)}) \widetilde{w}_{k|k-1}^{(j)}.$$

For $i = 1, ..., L_{k-1} + J_k$, update weights

$$\widetilde{w}_k^{(i)} = \left[1 - P_{D,k}(\widetilde{\mathbf{x}}_k^{(i)}) + \sum\nolimits_{\mathbf{z} \in Z_k} \frac{\psi_k(\mathbf{z}; \widetilde{\mathbf{x}}_k^{(i)})}{\lambda_{F,k} + C_k(\mathbf{z})} \right] \widetilde{w}_{k|k-1}^{(i)}.$$

Resample to get $\{(w_k^{(i)}, \mathbf{x}_k^{(i)})\}_{i=1}^{L_k}$.

E. The Cardinalized PHD Filter

The cardinalized PHD (CPHD) filter is a generalization of the PHD filter that jointly propagates the PHD v_k and cardinality distribution ρ_k to provide better performance albeit at higher computational complexity [85]. In addition to the standard multitarget state space model with i.i.d. cluster FAs, the CPHD recursion assumes that the prior and predicted multitarget densities are i.i.d. cluster. We use the alternative form of the CPHD recursion given in [156] since it facilitates implementations.

The CPHD prediction is the same as the PHD prediction except for the additional calculation of the predicted cardinality distribution $\rho_{k|k-1}$, which is the convolution:

$$\rho_{k|k-1}(n) = \sum\nolimits_{j=0}^{n} \rho_{\Gamma,k}(n-j) \rho_{S,k|k-1}(j),$$

of the birth cardinality distributions $\rho_{\Gamma,k}$ (given in the birth model) and surviving target cardinality distribution

$$\begin{aligned} \rho_{S,k|k-1}(j) &= \sum_{\ell=j}^{\infty} C_j^{\ell} \bar{P}_{S,k|k-1}^{j} (1 - \bar{P}_{S,k|k-1})^{\ell - j} \rho_{k-1}(\ell) \\ \bar{P}_{S,k|k-1} &= \langle P_{S,k|k-1}, v_{k-1} \rangle / \langle 1, v_{k-1} \rangle \end{aligned}$$

The CPHD update is given by [85], [156]

$$\rho_k(n) = \frac{\Upsilon_k^{(0)}[v_{k|k-1}, Z_k](n) \rho_{k|k-1}(n)}{\langle \Upsilon_k^{(0)}[v_{k|k-1}, Z_k], \rho_{k|k-1} \rangle},$$

$$\begin{aligned} v_k(\mathbf{x}) = &\frac{\langle \Upsilon_k^{(1)}[v_{k|k-1}, Z_k], \rho_{k|k-1} \rangle (1 - P_{D,k}(\mathbf{x})) v_{k|k-1}(\mathbf{x})}{\langle \Upsilon_k^{(0)}[v_{k|k-1}, Z_k], \rho_{k|k-1} \rangle} \\ &+ \sum_{\mathbf{z} \in Z_k} \frac{\langle \Upsilon_k^{(1)}[v_{k|k-1}, Z_k - \{\mathbf{z}\}], \rho_{k|k-1} \rangle \psi_k(\mathbf{z}; \mathbf{x}) v_{k|k-1}(\mathbf{x})}{\langle \Upsilon_k^{(0)}[v_{k|k-1}, Z_k], \rho_{k|k-1} \rangle} \end{aligned}$$

where

$$\begin{aligned} \Upsilon_k^{(u)}[v, Z](n) &= \sum_{S \subseteq Z} \frac{P_u^{n-|S|}}{\langle 1 - P_{D,k}, v \rangle^u} \Xi_k[v, Z, S | n] \\ \Xi_k[v, Z, S | n] &= P_{|S|}^{n} [\bar{\psi}_k]^S (1 - \bar{P}_{D,k})^{n-|S|} \rho_{F,k}(|Z - S|) |Z - S|! \\ \bar{\psi}_k(\mathbf{z}) &= \langle \psi_k(\mathbf{z}; \cdot), v \rangle / \langle 1, v \rangle \\ \bar{P}_{D,k} &= \langle P_{D,k}, v \rangle / \langle 1, v \rangle. \\ \rho_{F,k} &= \text{cardinality distribution of FAs} \end{aligned}$$

Multitarget state estimation for the CPHD filter is similar to that for the PHD filter. In addition, the number of targets can be estimated using $\arg\max \rho_k(\cdot)$.

The bottleneck of the CPHD update is the evaluation of the *elementary symmetric function*

$$e_j(Y) = \sum_{S \subseteq Y, |S| = j} \prod_{\zeta \in S} \zeta,$$

for a finite subset Y of real numbers, with $e_0(Y) = 1$ by convention [20]. Using the Newton-Girard formulae (or Vieta's Theorem) $e_j(Y)$ can be evaluated efficiently by expanding out the polynomial with roots given by the elements of Y, and collect the coefficient of the $|Y| - j$ power [156]. Both the PHD and CPHD filter are linear in the number of targets. The PHD filter is linear in the number of observations while the CPHD filter has a complexity of $\mathcal{O}\left(|Z_k|^2 \log |Z_k|\right)$ [156].

Under the linear Gaussian multitarget model, the CPHD recursion also admits a closed form solution. If v_{k-1} is a Gaussian mixture of the form (67), then $v_{k|k-1}$ is the same as the GM-PHD prediction (68), and $\rho_{k|k-1}$ is the convolution of $\rho_{\Gamma,k}$ and $\rho_{S,k|k-1}$ with $\bar{P}_{S,k|k-1} = P_{S,k|k-1}$. If $v_{k|k-1}$ is a Gaussian mixture of the form (69), then

$$\rho_k(n) = \frac{\Upsilon_k^{(0)}[\mathbf{w}_{k|k-1}, Z_k](n)\rho_{k|k-1}(n)}{\langle \Upsilon_k^{(0)}[\mathbf{w}_{k|k-1}, Z_k], \rho_{k|k-1}\rangle},$$

$$v_k(\mathbf{x}) = (1-P_{D,k})\frac{\langle \Upsilon_k^{(1)}[\mathbf{w}_{k|k-1}, Z_k], \rho_{k|k-1}\rangle}{\langle \Upsilon_k^{(0)}[\mathbf{w}_{k|k-1}, Z_k], \rho_{k|k-1}\rangle} v_{k|k-1}(\mathbf{x}) + P_{D,k}\sum_{\mathbf{z}\in Z_k}\sum_{i=1}^{J_{k|k-1}} w_k^{(i)}(\mathbf{z})\mathcal{N}(\mathbf{x}; \mathbf{m}_k^{(i)}(\mathbf{z}), \mathbf{P}_k^{(i)}),$$

where

$$w_k^{(i)}(\mathbf{z}) = \frac{\langle \Upsilon_k^{(1)}[\mathbf{w}_{k|k-1}, Z_k - \{\mathbf{z}\}], \rho_{k|k-1}\rangle q_k^{(i)}(\mathbf{z}) w_{k|k-1}^{(i)}}{\langle \Upsilon_k^{(0)}[\mathbf{w}_{k|k-1}, Z_k], \rho_{k|k-1}\rangle}$$

$$\mathbf{w}_{k|k-1} = [w_{k|k-1}^{(1)}, ..., w_{k|k-1}^{(J_{k|k-1})}]',$$

$$\Upsilon_k^{(u)}[\mathbf{w}, Z](n) = \sum_{S\subseteq Z}\frac{P_u^{n-|S|}}{[(1-P_{D,k})\mathbf{1}'\mathbf{w}]^u}\Xi_k[\mathbf{w}, Z, S|n]$$

$$\Xi_k[\mathbf{w}, Z, S|n] = P_{|S|}^n[\bar{\psi}_k]^S(1-P_{D,k})^{n-|S|}\rho_{F,k}(|Z-S|)|Z-S|!$$

$$\bar{\psi}_k(\mathbf{z}) = P_{D,k}\mathbf{w}'\mathbf{q}_k(\mathbf{z})/\mathbf{1}'\mathbf{w}$$

$$\mathbf{q}_k(\mathbf{z}) = [q_k^{(1)}(\mathbf{z}), ..., q_k^{(J_{k|k-1})}(\mathbf{z})]',$$

and $\mathbf{m}_k^{(i)}(\mathbf{z})$, $\mathbf{P}_k^{(i)}$ and $q_k^{(i)}(\mathbf{z})$ are the same as in the GM-PHD filter.

As with the Kalman filter, the extended PHD/CPHD filters are GM-PHD/CPHD filters with linearized state space equations, and the unscented PHD/CPHD filters are GM-PHD/CPHD filters with unscented transform approximations [155]. The particle implementation of the CPHD filter follows that for the PHD filter [154]. The PHD/CPHD filters can be applied to jointly estimate the FA parameters, state-dependent detection probability, and the multitarget state [13], [93], [95]. They have also been extended to multiple models [115], [94], extended targets [90], [28], [56], superpositional measurements [111], multiple sensors [87], [89], [92], [28] and distributed multitarget filtering [11], [153]. We refer the reader to the text [96] for more details on advances in PHD filtering.

F. The Generalized Labeled Multi-Bernoulli Tracker

In addition to the PHD/CPHD filters, other approximations of the multitarget Bayes filter include the multi-Bernoulli filters [159], [160] and their extensions [122] [113], [162], [170]. While these filters are not formulated to output tracks, their generalization, the Generalized Labeled Multi-Bernoulli (GLMB) filter, is [163], [164].

Targets are labeled by an ordered pair of integers $\ell = (k, i)$, where k is the time of birth, and i is a unique index to distinguish targets born at the same time. Figure 12 illustrates the assignment of labels to target trajectories. The label space for targets born at time k is denoted as $\mathbb{L}_k$ and the label space for targets at time k (including those born prior to k) is denoted as $\mathbb{L}_{0:k}$. Note that $\mathbb{L}_{0:k} = \mathbb{L}_{0:k-1} \cup \mathbb{L}_k$.

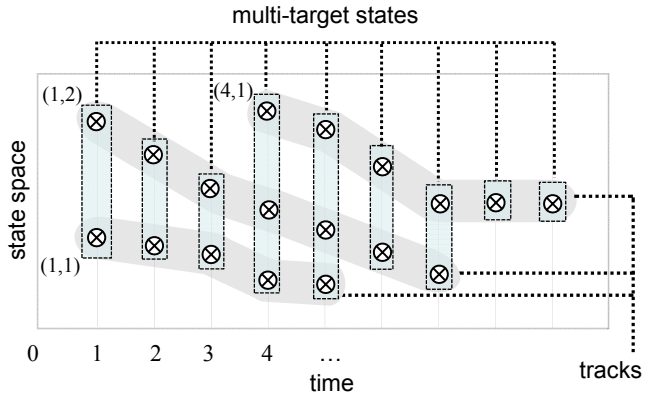

Fig. 12. A label assignment example: The two tracks born at time 1 are given labels (1,1) and (1,2), while the track born at time 4 is given label (4,1).

An existing target at time k has state $(\mathbf{x}, \ell)$ consisting of the kinematic/feature $\mathbf{x} \in \mathbb{X}$ and label $\ell \in \mathbb{L}_{0:k}$, i.e. single-target state space $\mathcal{X}$ is the Cartesian product $\mathbb{X}\times\mathbb{L}_{0:k}$. To ensure that the labels of a multitarget state $X \subset \mathbb{X}\times\mathbb{L}_{0:k}$ are distinct, we require X and the set of labels of X, denoted as $\mathcal{L}(X)$, to have the same cardinality. The function $\Delta(X) \triangleq \delta_{|X|}[|\mathcal{L}(X)|]$ is called the *distinct label indicator*.

An *association map* at time k is a function $\theta : \mathbb{L}_{0:k} \to \{0, 1, ..., |Z|\}$ such that $\theta(\ell) = \theta(\ell') > 0$ implies $\ell = \ell'$. Such a function can be regarded as an assignment of labels to measurements, with undetected labels assigned to 0. The set of all such association maps is denoted as Θ_k; the subset of association maps with domain L is denoted by $\Theta_k(L)$; and $\Theta_{0:k} \triangleq \Theta_0 \times ... \times \Theta_k$.

In the GLMB filter, the multitarget filtering density at time $k-1$ is a GLMB of the form

$$\pi_{k-1}(X|Z^{k-1}) = \Delta(X)\sum_{\xi\in\Theta_{0:k-1}} w_{k-1}^{(\xi)}(\mathcal{L}(X))[p_{k-1}^{(\xi)}]^X, \quad (73)$$

where each $p_{k-1}^{(\xi)}(\cdot, \ell)$ is a probability density, and each weight $w_{k-1}^{(\xi)}(L)$ is non-negative with

$$\sum_{L\in\mathcal{F}(\mathbb{L})}\sum_{\xi\in\Theta_{0:k-1}} w_{k-1}^{(\xi)}(L) = 1.$$

The cardinality distribution of the GLMB in (73) is given by

$$\rho_{k-1}(n) = \sum_{L\in\mathcal{F}(\mathbb{L})}\sum_{\xi\in\Theta_{0:k-1}} \delta_n(|L|) w_{k-1}^{(\xi)}(L). \quad (74)$$

Each $\xi = (\theta_0, ..., \theta_{k-1}) \in \Theta_{0:k-1}$ represents a history of association maps up to time $k-1$, and contains the history of target labels encapsulating both births and deaths. A tractable suboptimal multitarget estimate is obtained by the following proceedure: determine the MAP cardinality estimate n^*; determine the label set L^* and ξ^* with highest weight $w_{k-1}^{(\xi^*)}(L^*)$ among those with cardinality n^*; determine the expected values of the states from $p_{k-1}^{(\xi^*)}(\cdot, \ell)$, $\ell \in L^*$ [163].

The set of targets born at time k is modelled by a GLMB with one term: $f_{\Gamma,k}(X) = \Delta(X)w_{\Gamma,k}(\mathcal{L}(X))p_{\Gamma,k}^X$ (a full

GLMB birth can also be easily accommodated) [163]. Since the label of a target does not evolve with time, we have

$$f_{k|k-1}(\mathbf{x}_k,\ell_k|\mathbf{x}_{k-1},\ell_{k-1})=f_{k|k-1}(\mathbf{x}_k|\mathbf{x}_{k-1},\ell_{k-1})\delta_{\ell_{k-1}}[\ell_k].$$

The GLMB density is a conjugate prior with respect to the standard multitarget likelihood function and is also closed under the multitarget prediction. Under the standard multitarget model, if the multitarget filtering density, at the previous time, π_{k-1} is a GLMB of the form (73), then the multitarget prediction density $\pi_{k|k-1}$ is a GLMB given by [163]

$$\pi_{k|k-1}(X|Z^{k-1}) = \Delta(X)\sum_{\xi\in\Theta_{0:k-1}} w^{(\xi)}_{k|k-1}(\mathcal{L}(X))[p^{(\xi)}_{k|k-1}]^X,$$

where

$$\begin{aligned}
w^{(\xi)}_{k|k-1}(L) &= w^{(\xi)}_{S,k|k-1}(L\cap\mathbb{L}_{0:k-1})w_{\Gamma,k}(L\cap\mathbb{L}_k),\\
p^{(\xi)}_{k|k-1}(\mathbf{x},\ell) &= 1_{\mathbb{L}_{0:k-1}}(\ell)p^{(\xi)}_{S,k|k-1}(\mathbf{x},\ell)+1_{\mathbb{L}_k}(\ell)p_{\Gamma,k}(\mathbf{x},\ell),\\
w^{(\xi)}_{S,k|k-1}(L) &= [\bar{P}^{(\xi)}_{S,k|k-1}]^L\sum_{I\supseteq L}[1-\bar{P}^{(\xi)}_{S,k|k-1}]^{I-L}w^{(\xi)}_{k-1}(I),\\
\bar{P}^{(\xi)}_{S,k|k-1}(\ell) &= \left\langle P_{S,k|k-1}(\cdot,\ell),p^{(\xi)}_{k-1}(\cdot,\ell)\right\rangle,\\
p^{(\xi)}_{S,k|k-1}(\mathbf{x},\ell) &= \frac{\left\langle P_{S,k|k-1}(\cdot,\ell)f_{k|k-1}(\mathbf{x}|\cdot,\ell),p^{(\xi)}_{k-1}(\cdot,\ell)\right\rangle}{\bar{P}^{(\xi)}_{S,k|k-1}(\ell)},
\end{aligned}$$

and the multitarget filtering density π_k is a GLMB given by

$$\pi_k(X|Z^k) = \Delta(X)\sum_{\xi\in\Theta_{0:k-1}}\sum_{\theta\in\Theta_k} w^{(\xi,\theta)}_k(\mathcal{L}(X)|Z_k)[p^{(\xi,\theta)}(\cdot|Z_k)]^X,$$

where

$$\begin{aligned}
w^{(\xi,\theta)}_k(L|Z) &\propto 1_{\Theta_k(L)}(\theta)[\bar{\Psi}^{(\xi,\theta)}_{Z,k}]^L w^{(\xi)}_{k|k-1}(L),\\
\bar{\Psi}^{(\xi,\theta)}_{Z,k}(\ell) &= \left\langle \Psi^{(\theta)}_{Z,k}(\cdot,\ell),p^{(\xi)}_{k|k-1}(\cdot,\ell)\right\rangle,\\
\Psi^{(\theta)}_{Z,k}(\mathbf{x},\ell) &= \begin{cases}\frac{\psi_k(\mathbf{z}_{\theta(\ell)};(\mathbf{x},\ell))}{\lambda_{F,k}}, & \text{if } \theta(\ell)>0\\ 1-P_{D,k}(\mathbf{x},\ell), & \text{if } \theta(\ell)=0\end{cases}\\
p^{(\xi,\theta)}_k(\mathbf{x},\ell|Z) &= \frac{\Psi^{(\theta)}_{Z,k}(\mathbf{x},\ell)p^{(\xi)}_{k|k-1}(\mathbf{x},\ell)}{\bar{\Psi}^{(\xi,\theta)}_{Z,k}(\ell)}.
\end{aligned}$$

The GLMB recursion above is the first exact close form solution to the Bayes multitarget filter. Truncating the GLMB sum is needed to manage the growing the number of components in the GLMB filter. In [164] an implementation of the GLMB filter based on discarding "insignificant" components was detailed, and it was shown that such truncation minimizes the L_1 error in the multitarget density. This algorithm has a worst case complexity that is cubic in the number of observations.

A one term approximation to the GLMB filter, known as the LMB filter [124], was used to track thousands of targets simultaneously in relatively dense FAs on a laptop computer [165]. Moreover, it has been deployed as a real-time multitarget tracker in automative safety systems [125]. Recently, the GLMB filter was extended to the more realistic and very challenging problem of multitarget tracking with merged measurements [14].

ACKNOWLEDGMENT

The work on the Joint Probabilistic Data Association Filter (JPDAF) has been supported by ARO W911NF-10-1-0369. The work of the first author is supported by the Australian Research Council under Discovery Project DP130104404.

REFERENCES

[1] B. Anderson and J. Moore, *Optimal Filtering*, Prentice Hall, 1979.
[2] 1.8 gigapixel ARGUS-IS. World's highest resolution video surveillance platform by DARPA, https://www.youtube.com/watch?v=QGxNyaXfJsA.
[3] I. Arasaratnam and S. Haykin, "Cubature Kalman filters," *IEEE Trans. Automatic Control*, vol. 54, no. 6, pp. 1254–1269, 2009.
[4] M. Arulampalam, S. Maskell, N. Gordon and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Trans. Signal Processing*, vol. 50, no. 2, pp. 174–188, 2002.
[5] Y. Bar-Shalom and T. E. Fortmann, *Tracking and Data Association*. Academic Press, San Diego, 1988.
[6] Y. Bar-Shalom, Y. Chang, and H. Blom, "Automatic track formation in clutter with a recursive algorithm," in Y. Bar-Shalom (ed.) *Multisensor Multitarget Tracking: Advanced Applications*, pp. 25–42, 1992.
[7] Y. Bar-Shalom, Ed., *Multitarget-Multisensor Tracking: Applications and Advances*, vol. II, Norwood, MA, Artech House, 1992, Reprinted by YBS Publishing, 1998.
[8] Y. Bar-Shalom, X. Li and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation*, Wiley, New York, 2001.
[9] Y. Bar-Shalom, S. S. Blackman, and R. J. Fitzgerald, "Dimensionless score function for multiple hypothesis tracking," *IEEE Trans. Aerospace & Electronic Systems*, vol. 43, no. 1, pp. 392-400, 2007.
[10] Y. Bar-Shalom, P. Willett, and X. Tian, *Tracking and Data Fusion: A Handbook of Algorithms*, YBS Publishing, 2011.
[11] G. Battistelli, L. Chisci, C. Fantacci, A. Farina, and A. Graziano, "Consensus CPHD filter for distributed multitarget tracking," *IEEE J. Selected Topics in Signal Processing*, vol. 7, no. 3, pp. 508-520, 2013.
[12] M. Baum, P. Willett, Y. Bar-Shalom, and U. Hanebeck, "Approximate calculation of marginal association probabilities using a hybrid data association model," *in SPIE Signal and Data Processing of Small Targets*, vol. 9092, 2014.
[13] M. Beard, B.-T. Vo, and B.-N. Vo, "Multi-target filtering with unknown clutter density using a bootstrap GM-CPHD filter", *IEEE Signal Processing Letters*, vol. 20, no. 4, pp. 323-326, 2013.
[14] M. Beard, B.-T. Vo, and B.-N. Vo, "Bayesian multitarget tracking with merged measurements using labelled random finite sets," *IEEE Trans. Signal Processing*, vol. 63, no. 6, pp. 1433-1447, 2015.
[15] S. Blackman and R. Popoli, *Design and Analysis of Modern Tracking Systems*, Artech House, 1999.
[16] S. Blackman, "Multiple hypothesis tracking for multiple target tracking," *IEEE Aerospace & Electronic Systems Magazine*, vol. 19, no. 1, pp. 5–18, 2004.
[17] H. Blom and E. Bloem, "Bayesian tracking of two possibly unresolved maneuvering targets," *IEEE Trans. Aerospace & Electronic Systems*, vol. 43, no. 2, pp. 612-627, Apr. 2007.
[18] H. A. P. Blom and E. A. Bloem, "Probabilistic data association avoiding track coalescence," *IEEE Trans. Aerospace & Electronic Systems*, vol. 45, no. 2, pp. 247–259, 2000.
[19] M. Bolic, P. Djuric, and S. Hong, "Resampling algorithms and architectures for distributed particle filters," *IEEE Trans. Signal Processing*, vol. 53, no. 7, pp. 2442–2450, 2005.
[20] P. Borwein and T. Erdélyi, *Newton's Identities Section 1.1.E.2 in Polynomials and Polynomial Inequalities*. Springer-Verlag, New York, 1995.
[21] O. Cappé, S. J. Godsill, and Eric Moulines, "An overview of existing methods and recent advances in sequential Monte Carlo," *Proc. IEEE*, vol. 95, no. 5, pp. 899–924, 2007.
[22] S. Challa, M.Morelande, D.Musicki, and R. Evans, *Fundamentals of Object Tracking*, Cambridge University Press, 2011.
[23] K.-C. Chang and Y. Bar-Shalom, "Joint probabilistic data association for multitarget tracking with possibly unresolved measurements and maneuvers," *IEEE Trans. Automatic Control*, vol. 29, no. 7, pp. 585-594, Jul. 1984.
[24] R. Chen and J. Liu, "Mixture Kalman filters," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 62, part 3, pp. 493-508, 2000.

[25] N. Chopin, "A sequential particle filter method for static models," *Biometrika*, vol. 89, no. 3, pp. 539–551, Aug 2002.

[26] D. Clark and J. Bell, "Convergence results for the particle PHD filter," *IEEE Trans. Signal Processing*, vol. 54, no. 7, pp. 2652–2661, 2006.

[27] D. Clark and B.-N. Vo, "Convergence analysis of the Gaussian mixture PHD filter," *IEEE Trans. Signal Processing*, vol. 55, No. 4, pp. 1204–1212, 2007.

[28] D. Clark, and R. Mahler "Generalized PHD filters via a general chain rule," *Proc. Int'l. Conf. Information Fusion*, Singapore, July 2012.

[29] S. Coraluppi, C. Carthel, M. Luettgen, and S. Lynch, "All-source track and identity fusion," *Proc. Nat'l. Symp. Sensor & Data Fusion*, San Antonio TX, June 2000.

[30] S. Coraluppi and C. Carthel, "Modified scoring in multiple-hypothesis tracking," *J. Advances in Information Fusion*, vol 7, no. 2, pp. 153–164, 2012.

[31] S. Coraluppi and C. Carthel, "If a tree falls in the woods, it does make a sound: Multiple-Hypothesis Tracking with undetected target births," *IEEE Trans. Aerospace & Electronic Systems*, vol 50, no. 3, pp. 2379–2388, 2014.

[32] D. Crisan, "Particle filters-A theoretical perspective," in *Sequential Monte Carlo Methods in Practice*, Doucet A., de Freitas N. and Gordon N. J., (eds.), pp. 17-41, Springer-Verlag, May 2001.

[33] D. Crouse, M. Guerriero, and P. Willett, "A critical look at the PMHT," *Journal of Advances in Information Fusion*, vol. 4, no. 2, pp. 93-116, Dec. 2009.

[34] D. Crouse, Private Communication, 2010.

[35] D. Crouse, Y. Bar-Shalom, P. K. Willett, and L. Svensson, "The JPDAF in practical systems: Computation and snake oil," *Proc. SPIE Conf. Signal & Data Processing of Small Targets*, vol. 7698 , Orlando, FL, April 2010.

[36] D. Daley and D. Vere-Jones, *An introduction to the theory of point processes*. Springer-Verlag, 1988.

[37] F. Daum and J. Huang, "How to avoid normalization of particle flow for nonlinear filters, Bayesian decisions, and transport," *Proc. SPIE Conf. Defense & Security*, pp. 90920B–90920B, International Society for Optics and Photonics, 2014.

[38] S. Davey and D. Gray, "Integrated track maintenance for the PMHT via the hysteresis model," *IEEE Trans. Aerospace & Electronic Systems*, vol. 43, no. 1, pp. 93-111, Jan. 2007.

[39] S. Davey, M. Rutten, and B. Cheung, "A comparison of detection performance for several Track-Before-Detect algorithms," *EURASIP J. Advances in Signal Processing*, vol. 2008, no. 1, article 41, 2008.

[40] S. Deb, M. Yeddanapudi, K. Pattipati, and Y. Bar-Shalom, "A generalized S-D assignment algorithm for multisensor-multitarget state estimation," *IEEE Trans. Aerospace & Electronic Systems*, vol. 33, no. 2, pp. 523-538, Apr. 1997.

[41] P. Del Moral, *Mean field simulation for Monte Carlo integration*, Chapman & Hall/CRC Monographs on Statistics & Applied Probability, 2013.

[42] T. Ding and M. J. Coates, "Implementation of the Daum-Huang exact-flow particle filter," *Proc. IEEE Statistical Signal Processing Workshop (SSP)*, pp. 257–260, 2012.

[43] R. Douc, O. Cappé, and E. Moulines, "Comparison of resampling schemes for particle filtering," *Proc. Int'l Symp. Image and Signal Processing and Analysis*, Istanbul, Turkey, Sep 2005.

[44] A. Doucet, S. Godsill, and C. Andrieu, "On sequential Monte Carlo sampling methods for Bayesian filtering," *Statist. Comput.*, vol. 10, no. 3, pp. 197–208, 2000.

[45] A. Doucet, N.. de Freitas, and N. J. Gordon, "An introduction to sequential Monte Carlo methods," in *Sequential Monte Carlo Methods in Practice*, A. Doucet, N.. de Freitas, and N. J. Gordon, Eds. New York: Springer-Verlag, 2001.

[46] A. Doucet and A. M. Johansen, "A tutorial on particle filtering and smoothing: Fifteen years later," *Handbook of Nonlinear Filtering*, vol. 12, pp. 656–704, 2009.

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

[48] M. Feldmann, D. Franken, and W. Koch, "Tracking of extended objects and group targets using random matrices," *IEEE Trans. Signal Processing*, vol. 59, no. 4, pp. 1409-1420, Apr. 2011.

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

[50] A. Gelb, Editor, *Applied Optimal Estimation*, MIT Press, 1974.

[51] K. Gilholm and D. Salmond, "Spatial distribution model for tracking extended objects," *IEE Proceedings on Radar, Sonar and Navigation*, vol. 152, no. 5, pp. 364-371, Oct. 2005.

[52] W. R. Gilks and C. Berzuini, "Following a moving target–Monte Carlo inference for dynamic Bayesian models," *J. R. Statist. Soc. B*, vol. 63, pp. 127–146, 2001.

[53] A. Gning, L. Mihaylova, S. Maskell, S. Pang, and S. Godsill, "Group object structure and state estimation with evolving networks and Monte Carlo methods," *IEEE Trans. Signal Processing*, vol. 59, no. 4, pp. 1383-1396, Apr. 2011.

[54] I. Goodman, R. Mahler, and H. Nguyen, *Mathematics of Data Fusion*. Kluwer Academic Publishers, 1997.

[55] N. J. Gordon, D. J. Salmond, and A. F. M. Smith, "Novel approach to nonlinear/non-Gaussian Bayesian state estimation," *IEE Proceedings-F*, vol. 140, no. 2, pp. 107–113, 1993.

[56] K. Granstrom, and U. Orguner, "A PHD filter for tracking multiple extended targets using random matrices" *IEEE Trans. Signal Processing*, vol. 60, no. 11, pp. 5657–5671, 2012.

[57] B. K. Habtemariam, R. Tharmarasa, T. Thayaparan, M. Mallick, and T. Kirubarajan, "A Multiple detection probabilistic data association filter," *IEEE J. Selected Topics in Signal Processing*, vol. 7, no. 3, pp. 461–471, June 2013.

[58] A. C. Harvey, *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge University Press, 1989.

[59] Y-C. Ho and R. C. K. Lee, "A Bayesian approach to problems in stochastic estimation and control," *IEEE Trans. Automatic Control*, vol. 9, no. 4, pp. 333–339, 1964.

[60] J. Hoffman and R. Mahler, "Multitarget miss distance via optimal assignment," *IEEE Trans. Sys., Man, and Cybernetics-Part A*, vol. 34, no. 3, pp. 327–336, 2004.

[61] P. Horridge and S. Maskell, "Real-time tracking of hundreds of targets with efficient exact JPDAF implementation," *Proc. Int'l. Conf. Information Fusion*, Florence, Italy, July 2006.

[62] R. Hoseinnezhad, B.-N.Vo, B.-T. Vo, and D. Suter, "Visual tracking of numerous targets via multi-Bernoulli filtering of image data," *Pattern Recognition*, vol. 45, no. 10, pp. 3625-3635, 2012.

[63] R. Hoseinnezhad, B.-N. Vo and B.-T. Vo, "Visual Tracking in Background Subtracted Image Sequences via Multi-Bernoulli Filtering," *IEEE Trans. Signal Processing*, vol. 61, no. 2, pp. 392-397, 2013.

[64] A. Jazwinski, *Stochastic Processes and Filtering Theory*, Academic Press, 1970.

[65] B. Jia, M. Xin, and Y. Cheng, "High-degree cubature Kalman filter," *Automatica*, vol. 49, no. 2, pp. 510–518, 2013.

[66] A. Johansen, S. Singh, A. Doucet, and B.-N Vo, "Convergence of the SMC implementation of the PHD filter," *Methodology & Computing in Applied Probability*, vol. 8, no. 2, pp. 265–291, 2006.

[67] S. Julier, J. Uhlmann and H.F. Durrant-Whyte, "A new method for the nonlinear transformation of means and covariances in filters and estimators," *IEEE Trans. Automatic Control*, vol. AC-45, no. 3, pp. 477–482, March 2000.

[68] S. J. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation", *Proc. IEEE*, vol. 92, no. 3, pp. 401–422, March 2004.

[69] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Trans. ASME J. Basic Engineering*, vol. 82, pp. 35–45, 1960.

[70] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Prentice Hall, Englewood Cliffs, New Jersey, 1993.

[71] W. Koch and G. van Keuk, "Multiple hypothesis track maintenance with possibly unresolved measurements," *IEEE Trans. Aerospace & Electronic Systems*, vol. 33, no. 3, pp. 883-892, Jul. 1997.

[72] J. Koch, "Bayesian approach to extended object and cluster tracking using random matrices," *IEEE Trans. Aerospace & Electronic Systems*, vol. 44, no. 3, pp. 1042-1059, Jul. 2008.

[73] W. Koch, *Tracking and Sensor Data Fusion: Methodological Framework and Selected Applications*, Springer, Heidelberg, 2014.

[74] J. H. Kotecha and P. M. Djuric, "Gaussian particle filtering," *IEEE Trans. Signal Processing*, vol. 51, no. 10, pp. 2592–2601, 2003.

[75] J. H. Kotecha and P. M. Djuric, "Gaussian sum particle filtering," *IEEE Trans. Signal Processing*, vol. 51, no. 10, pp. 2602–2612, 2003.

[76] T. Kurien, "Issues in the design of practical multitarget tracking algorithms," Chapter 3 in *Multitarget-Multisensor Tracking: Advanced Applications*, Ed. Y. Bar-Shalom, Artech House, pp. 43–83, 1990.

[77] X. R. Li, "Engineer's guide to variable-structure multiple-model estimation for tracking," Chapter 10, in *Multitarget-Multisensor Tracking: Applications and Advances*, Volume III, Ed. Y. Bar-Shalom and W. D. Blair, pp. 449–567, Aetech House, 2000.

[78] X. Li and V. Jilkov, "Survey of maneuvering target tracking, Part I: Dynamic models," *IEEE Trans. Aerospace & Electronic Systems*, vol. 39, no. 4, pp. 1333-1364, 2003.

[79] X. R. Li and V. P. Jilkov, "A survey of maneuvering target tracking, Part V: Multiple-Model methods," *IEEE Trans. Aerospace & Electronic Systems*, vol. 41, no. 4, pp. 1255–1321, 2005.

[80] J. S. Liu and R. Chen, "Sequential Monte Carlo methods for dynamical systems," *J. Amer. Statist. Assoc.*, vol. 93, pp. 1032–1044, 1998.

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

[82] R. Mahler, "Global integrated data fusion," *Proc. 7th Nat'l. Symp. Sensor Fusion*, vol. 1, (Unclassified) Sandia National Laboratories, Albuquerque, ERIM Ann Arbor MI, pp. 187–199, 1994.

[83] R. Mahler, "A theoretical foundation for the Stein-Winter Probability Hypothesis Density (PHD) multitarget tracking approach," *Proc. MSS Nat'l. Symp. Sensor & Data Fusion*, vol. I (Unclassified), San Antonio TX, June 2000.

[84] R. Mahler, "Multitarget Bayes filtering via first-order multitarget moments," *IEEE Trans. Aerospace & Electronic Systems*, vol. 39, no. 4, pp. 1152–1178, 2003.

[85] R. Mahler, "PHD filters of higher order in target number," *IEEE Trans. Aerospace & Electronic Systems*, vol. 43, no. 4, pp. 1523–1543, 2007.

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

[87] R. Mahler, "CPHD filters for superpositional sensors," *Proc. SPIE Signal & Data Processing of Small Targets*, O. E. Drummond (ed.), vol. 7445, 2009.

[88] R. Mahler, "The multisensor PHD filter: I. General solution via multitarget calculus," *Proc. SPIE Signal Processing, Sensor Fusion, and Target Recognition XVIII*, vol. 7336, pp. 73360E-12, May 2009.

[89] R. Mahler, "The multisensor PHD filter: II. Erroneous solution via Poisson magic," *Proc. SPIE Signal Processing, Sensor Fusion, and Target Recognition XVIII*, vol. 7336, pp. 73360D-12, May 2009.

[90] R. Mahler, "PHD filters for nonstandard targets, I: Extended targets," *Proc. Int'l. Conf. Information Fusion*, Seattle, July 2009.

[91] R Mahler, "PHD filters for nonstandard targets, II: Unresolved targets," *Proc. Int'l. Conf. Information Fusion*, Seattle, July 2009.

[92] R. Mahler, "Approximate multisensor CPHD and PHD filters," *Proc. Int'l. Conf. Information Fusion*, Edinburg, UK, July, 2010.

[93] R. Mahler, B.-T. Vo and B.-N. Vo "CPHD filtering with unknown clutter rate and detection profile," *IEEE Trans. Signal Processing*, vol. 59, No. 8, pp. 3497–3513, 2011.

[94] R. Mahler, "On multitarget jump-Markov filters," *Proc. Int'l Conf. on Information Fusion*, pp. 149–156, Singapore, July 2012.

[95] R. Mahler, and B.-T. Vo, "An improved CPHD filter for unknown clutter backgrounds," *Proc. SPIE Defense & Security*, 2014.

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

[97] M. Mallick and S. Arulampalam, "Comparison of nonlinear filtering algorithms in ground moving target indicator (GMTI) target tracking," *Proc. SPIE*, vol. 5204, San Diego, CA, August 2003.

[98] M. Mallick and B. F. La Scala, "Comparison of single-point and two-point difference track initiation algorithms using position measurements," *Proc. Int'l Colloquium on Information Fusion*, Xian, China, August 2007.

[99] M. Mallick, V. Krishnamurthy, and B.-N. Vo, Eds., *Integrated Tracking, Classification, and Sensor Management: Theory and Applications*, Wiley/IEEE, 2012.

[100] M. Mallick, M. Morelande, L. Mihaylova, S. Arulampalam, and Y. Yan, "Angle-only filtering in three dimensions," Ch. 1, in *Integrated Tracking, Classification, and Sensor Management: Theory and Applications*, M. Mallick, V. Krishnamurthy, and B.-N. Vo, Eds., Wiley/IEEE, pp. 3–42, December 2012.

[101] M. Mallick, S. Coraluppi, and C. Carthel, "Multitarget tracking using multiple hypotheses tracking," Chapter 5, in *Integrated Tracking, Classification, and Sensor Management: Theory and Applications*, M. Mallick, V. Krishnamurthy, and B.-N. Vo, Eds., Wiley/IEEE, pp. 165–201, December 2012.

[102] M. Mallick, S. Rubin, B.-N. Vo, "An introduction to force and measurement modeling for space object tracking," *Proc. Int'l. Conf. Information Fusion*, Istanbul, Turkey, July 9-12, 2013.

[103] S. Maskell, M. Briers, and R. Wright, "Fast mutual exclusion," *in SPIE Signal and Data Processing of Small Targets*, vol. 5428, 2004.

[104] E. Mazor, A. Averbuch, Y. Bar-Shalom, and J. Dayan, "Interacting Multiple Model methods in target tracking: A survey," *IEEE Trans. Aerospace & Electronic System.*, vol. 34, no. 1, pp. 103–123, 1998.

[105] J. S. Meditch, "A survey of data smoothing for linear and nonlinear dynamic systems," *Automatica* , vol. 9, no. 2, pp. 151–162, 1973.

[106] R. J. Meinhold and N. D. Singpurwalla, "Understanding the Kalman filter," *The American Statistician*, vol. 37, no. 2, pp. 123–127, 1983.

[107] J. Miguez, "Analysis of parallelizable resampling algorithms for particle filtering," *Elsevier Signal Processing*, vol. 87, no. 12, pp. 3155–3174, Dec 2007.

[108] L. Mihaylova, A. Carmi, F. Septier, A. Gning, S. Pang, and S. Godsill, "Overview of Bayesian sequential Monte Carlo methods for group and extended object tracking," *Digital Signal Processing*, vol. 25, 2014.

[109] D. Musicki and R. Evans, "Joint integrated probabilistic data association," *IEEE Trans. Automatic Control*, vol. AC-40, no. 3, pp. 1093–1099, 2004.

[110] D. Musicki and B. La Scala, "Multi-target tracking in clutter without measurement assignment," *IEEE Trans. Aerospace & Electronic Systems*, vol. 44, no. 3, pp. 877-896, Jul. 2008.

[111] S. Nannuru, M. Coates, and R. Mahler, "Computationally-tractable approximate PHD and CPHD filters for superpositional sensors, " *IEEE J. Selected Topics in Signal Processing*, vol. 7, no. 3, pp. 410–420, 2013.

[112] S. Oh, S. Russell, and S. Sastry, "Markov chain Monte Carlo data association for multitarget tracking," *IEEE Trans. Automatic Control*, vol. 54., no. 3, 481-497, 2009.

[113] C. Ouyang, H. Ji, and C. Li, "Improved multi-target multi-Bernoulli filter," *IET Radar, Sonar Navigation*, vol. 6, no. 6, pp. 458-464, Jul. 2012.

[114] S. Pang, J. Li, and S. Godsill, "Detection and tracking of coordinated groups," *IEEE Trans. Aerospace & Electronic Systems*, vol. 47, no. 1, pp. 472-502, Jan. 2011.

[115] A. Pasha, B.-N. Vo, H. D. Tuan, and W. K. Ma, "A Gaussian mixture PHD filter for jump Markov system model," *IEEE Trans. Aerospace & Electronic Systems*, vol. 45, no.3, pp. 919–936, 2009.

[116] K. Pattipati, S. Deb, Y. Bar-Shalom, and R. Washburn, "A new relaxation algorithm and passive sensor data association," *IEEE Trans. Automatic Control*, vol. 37, no. 2, pp. 198-213, Feb. 1992.

[117] K. R. Pattipati, R. L. Popp, and T. T. Kirubarajan, "Survey of assignment techniques for multitarget tracking," Chapter 2, in *Multitarget-Multisensor Tracking: Applications and Advances*, vol. III, Eds. Y. Bar-Shalom and D. Blair, pp. 77–159, Artech House, Norwood, MA, USA, 2000.

[118] M. Pitt and N. Shephard, "Filtering via simulation: Auxiliary particle filters," *J. Amer. Statist. Assoc.*, vol. 94, no. 446, pp. 590–599, 1999.

[119] A. Poore and N. Rijavec, "A Lagrangian relaxation algorithm for multidimensional assignment problems arising from multitarget tracking," *SIAM J. Optimization*, vol. 3, no. 3, pp. 544-563, August 1993.

[120] A. B. Poore and A. J. Robertson, "A new Lagrangian relaxation based algorithm for a class of multidimensional assignment problems," *Computational Optimization & Applications*, vol. 8, no. 2, pp. 129–150, 1997.

[121] G. W. Pulford and R. J. Evans, "A multipath data association tracker for over-the-horizon radar," *IEEE Trans. Aerospace & Electronic Systems*, vol. 34, no. 4, pp. 1165–1183, 1998.

[122] V. Ravindra, L. Svensson, L. Hammarstrand, and M. Morelande, "A cardinality preserving multitarget multi-Bernoulli RFS tracker," *Proc. Int'l. Conf. Information Fusion*, Singapore, July 2012.

[123] D. Reid, "An algorithm for tracking multiple targets," *IEEE Trans. Automatic Control*, vol. 24, no. 6, pp. 843–854, 1979.

[124] S. Reuter, B.-T. Vo, B.-N. Vo, and K. Dietmayer, "The labelled multi-Bernoulli filter," *IEEE Trans. Signal Processing*, vol. 62, no. 12, pp. 3246–3260, 2014.

[125] S. Reuter, *Multi-Object Tracking Using Random Finite Sets*, Diss. Zugl.: Ulm, Universität Ulm, 2014.

[126] B. Ristic, S. Arulampalam, and N. J. Gordon, *Beyond the Kalman Filter: Particle Filters for Tracking Applications*. Artech House, 2004.

[127] B. Ristic, D. Clark, B.-N.Vo, and B.-T. Vo "Adaptive target birth intensity in PHD and CPHD filters," *IEEE Trans. Aerospace & Electronic Systems*, vol. 48, no. 2, pp. 1656–1668, 2012.

[128] C. P. Robert, *The Bayesian Choice*, Second Edition, Springer, New York, 2007.

[129] A. Rodningsby, Y. Bar-Shalom, O. Hallingstad and J. Glattetre, "Multitarget tracking in the presence of wakes," *J. Advances in Information Fusion*, vol. 4, no. 2, pp. 117–145, 2009.

[130] J. Roecker, "Suboptimal joint probabilistic data association," *IEEE Trans. Aerospace & Electronic Systems*, vol. 29, no. 2, pp. 510–517, 1993.

[131] J. Roecker, "A class of near optimal JPDA algorithms," *IEEE Trans. Aerospace & Electronic Systems*, vol. 30, no. 2, pp. 504–510, 1994.

[132] K. Romeo, D. Crouse, Y. Bar-Shalom, and P. Willett, "The JPDAF in practical systems: approximations," *in SPIE Signal and Data Processing of Small Targets*, vol. 7698, 2010.

[133] Y. Ruan and P. Willett, "The turbo PMHT," *IEEE Trans. Aerospace & Electronic Systems*, vol. 40, no. 4, pp. 1388–1398, 2004.

[134] A. Runnalls, "Kullback-Leibler approach to Gaussian mixture reduction," *IEEE Trans. Aerospace & Electronic Systems*, vol. 43, no. 3, pp. 989-999, Jul. 2007.

[135] D. Salmond, "Mixture reduction algorithms for target tracking in clutter," in SPIE Signal and Data Processing of Small Targets, vol. 1305, 1990.

[136] D. J. Salmond and H. Birch, "A particle filter for track-before-detect," *Proc. American Control Conference*, vol. 5, pp. 3755–3760, Arlington, VA, USA, June 2001.

[137] D. Salmond, "Mixture reduction algorithms for point and extended object tracking in clutter," *IEEE Trans. Aerospace & Electronic Systems*, vol. 45, no. 2, pp. 667-686, Apr. 2009.

[138] T. Sathyan, A. Sinha, T. Kirubarajan, M. McDonald, and T. Lang, "MDA-based data association with prior track information for passive multitarget tracking," *IEEE Trans. Aerospace & Electronic Systems*, vol. 47, no. 1, pp. 539–556, 2011.

[139] T. Sathyan, T-J. Chin, S. Arulampalam, and D. Suter, "A multiple hypothesis tracker for multitarget tracking with multiple simultaneous measurements," *IEEE J. Selected Topics in Signal Processing*, vol. 7, no. 3, pp. 448–460, 2013.

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

[141] R. Singer and J. Stein, "An optimal tracking filter for processing sensor data of imprecisely determined origin in surveillance systems," *Proc. IEEE Conf. Decision & Control*, Florida, USA, pp. 171–175, Dec 1971.

[142] S. Singh, B.-N. Vo, A. Baddeley, and S. Zuev, "Filters for spatial point processes," *SIAM J. Control and Optimization*, vol. 48, no. 4, pp. 2275–2295, 2009.

[143] H. W. Sorenson and D. L. Alspach, "Recursive Bayesian estimation using Gaussian sum," *Automatica*, vol. 7, pp. 465–479, 1971.

[144] L. Stone, R. Streit, T. Corwin, and K. Bell, *Bayesian Multiple Target Tracking*, 2nd ed. Artech House 2013

[145] P. Storms and F. Spieksma, "An LP-based algorithm for the data association problem in multitarget tracking," *Proc. Int'l. Conf. Information Fusion*, Paris, France, July 2000.

[146] R. Streit and T. Luginbuhl, "Probabilistic multiple hypothesis tracking," *Technical Report, NUWC-NPT*, vol. 428, no. 10, Feb 1995.

[147] D. Svensson, J. Wintenby, and L. Svensson, Performance evaluation of MHT and GMCPHD in a ground target tracking scenario, *Proc. Int'l. Conf. Information Fusion*, Seattle, Washington, USA, July, 2009.

[148] L. Svensson, D. Svensson, M. Guerriero, and P. Willett, "Set JPDA filter for multitarget tracking," *IEEE Trans. Signal Processing*, vol. 59, no. 10, pp. 4677-4691, Oct. 2011.

[149] D. Svensson, *Target Tracking in Complex Scenarios*, Ph.D. Thesis, Chalmers University of Technology, Göteberg, Sweden, 2010.

[150] D. Svensson, M. Ulmke and L. Hammarstrand, "Multitarget Sensor Resolution Model and Joint Probabilistic Data Association", *IEEE Trans. Aerospace & Electronic Systems*, vol. 48, no. 4, pp. 3418-3434, Oct. 2012.

[151] R. Tharmarasa, S. Sutharsan, T. Kirubarajan, and T. Lang, "Multiframe assignment tracker for MSTWG data," *Proc. Int'l Conf. Information Fusion*, Seattle, WA, USA, pp. 1837–1844, July 2009.

[152] R. Tharmarasa, M. Subramaniam, N. Nadarajah, T. Kirubarajan, and M. McDonald, "Multitarget passive coherent location with transmitter origin and target-altitude uncertainties," *IEEE Trans. Aerospace & Electronic Systems*, vol. 48, no. 3, pp. 2530–2550, 2012.

[153] M. Uney, D. Clark, S. Julier, "Distributed fusion of PHD filters via exponential mixture densities, Selected Topics in Signal Processing," *IEEE J. Selected Topics in Signal Processing*, vol. 7, no. 3, pp. 521-531, 2013.

[154] B.-N. Vo, S. Singh, and A. Doucet, "Sequential Monte Carlo methods for multitarget filtering with random finite sets," *IEEE Trans. Aerospace & Electronic Systems*, vol. 41, no. 4, pp. 1224–1245, 2005.

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

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

[157] B.-T. Vo, *Random Finite Sets in Multi-Object Filtering*, Ph.D Thesis, University of Western Australia, 2008.

[158] B.-T. Vo, B.-N. Vo, and A. Cantoni, "Bayesian filtering with random finite set observations," *IEEE Trans. Signal Processing*, vol. 56, no. 4, pp. 1313–1326, 2008.

[159] B.-T. Vo, B.-N. Vo, and A. Cantoni, "The Cardinality Balanced Multitarget Multi-Bernoulli filter and its implementations," *IEEE Trans. Signal Processing*, vol. 57, no. 2, pp. 409–423, Feb. 2009.

[160] B.-N. Vo, B.-T. Vo, N.-T. Pham and D. Suter, "Joint detection and estimation of multiple objects from image observations," *IEEE Trans. Signal Procesing*, vol. 58, no. 10, pp. 5129–5241, 2010.

[161] B.-N. Vo, B.-T. Vo and R. Mahler, "Closed form solutions to forward-backward smoothing," *IEEE Trans. Signal Processing*, vol. 60, no. 1, pp. 2–17, 2012.

[162] B.-T. Vo, B.-N. Vo, R. Hoseinnezhad, and R. Mahler, "Robust multi-Bernoulli filtering," *IEEE J. Sel. Topics Signal Processing*, vol. 7, no. 3, pp. 399-409, Jun. 2013.

[163] B.-T. Vo, and B.-N. Vo, "Labeled Random Finite Sets and multi-object conjugate priors," *IEEE Trans. Signal Processing*, vol. 61, no. 13, pp. 3460–3475, 2013.

[164] B.-N. Vo, B.-T. Vo, and D. Phung, "Labeled random finite sets and the Bayes multitarget tracking filter," *IEEE Trans. Signal Processing*, vol. 62, no. 24, pp. 6554–6567, 2014.

[165] B.-N. Vo, B.-T. Vo, S. Reuter, Q. Lam, and K. Dietmayer, "Towards large scale multi-target tracking," *Proc. SPIE Defense & Security*, pp. 90850W-90850W, June, 2014.

[166] T. Vu, B.-N. Vo and R.J. Evans "A Particle Marginal Metropolis-Hastings multitarget tracker," *IEEE Trans. Signal Processing*, vol. 62, no. 15, pp. 3953–3964, 2014.

[167] N. Whiteley, S. Singh, and S. Godsill, "Auxiliary Particle implementation of the Probability Hypothesis Density filter," *IEEE Trans. Aerospace & Electronic Systems*, vol. 46, no. 3, pp. 1437–1454, 2010.

[168] P. Willett, Y. Ruan, and R. Streit, "PMHT: Some problems and solutions," *IEEE Trans. Aerospace & Electronic Systems*, vol. 38, no. 3, pp. 738–754, 2002.

[169] J. Williams and R. Lau, "Approximate evaluation of marginal association probabilities with belief propagation," *IEEE Trans. Aerospace & Electronic System*, vol. 50, no. 4, Oct. 2014.

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

[171] X. Xie and R. Evans, "Multiple-target tracking and multiple frequency line tracking using hidden Markov models," *IEEE Trans. Signal Processing*, vol. 39, no. 12, pp. 2659–2676, 1991.