The Para-Normal Bayes Multi-Target Filter and the Spooky Effect

Ba Tuong Vo and Ba Ngu Vo
School of Electrical, Electronic and Computer Engineering
The University of Western Australia, Crawley, WA 6009, Australia.
E-mail: {ba-tuong.vo,ba-ngu.vo}@uwa.edu.au

***Abstract*—The Probability Hypothesis Density (PHD) and Cardinalized PHD (CPHD) filters exhibit a counter intuitive behaviour called the "spooky effect" where upon a missed detection, the PHD mass in the vicinity of the undetected target is shifted to the vicinity of the detected targets, regardless of the distance between the targets. This raises the question of whether spookiness is an artifact of the random finite set formulation of the multi-target filtering problem. Using a para-normal implementation of the labeled multi-target Bayes filter, we show that this filter does not exhibit the "spooky effect" observed in the PHD/CPHD filters.**

***Index Terms*—Random sets, Multi-Target Bayes filter, PHD, CPHD, Multi-Bernoulli, Conjugate prior, Tracking**

I. Introduction

Multi-target filtering is the problem of jointly estimating the number of targets present and their individual dynamic states, from a sequence of finite set observations. To date, three major approaches have emerged as the main solution 'paradigms'. These are, Multiple Hypotheses Tracking (MHT) [2], Joint Probabilistic Data Association (JPDA) [1], and Random Finite Set (RFS) [5].

The random finite set (RFS) approach is the newest of these paradigms pioneered by Mahler. In essence, the collection of target states at any given time is treated as a set-valued multi-target state [6], [12]. The rationale behind this representation traces back to a fundamental consideration in estimation theory–estimation error [16]. The centerpiece of the RFS approach is the multi-object Bayes filter [5], which propagates the posterior density of the multi-object state recursively in time. Due to the high numerical complexity of multi-object Bayes filter, the Probability Hypothesis Density (PHD) [6] and Cardinalized PHD (CPHD) [7] filters have been developed as approximations based on first moments and cardinality distributions while the Multi-Bernoulli (MB) filter is a parameterized approximation which admits intuitive track interpretations [15], [16]. Beyond moments and parameterization are implementations of the multi-object Bayes filter based on the notion of a conjugate prior introduced in [17].

The subject of this discussion is the "spooky effect" observed from the PHD/CPHD filters. In [3], [4] an intuitive interpretation of the PHD/CPHD filters is presented alongside the revelation of a counter intuitive behavior, whereby upon a missed detection, the PHD mass in the vicinity of the undetected target is shifted to the vicinity of the detected targets, regardless of the distance between the targets. The shifted mass and, hence, the remaining mass of the undetected part both depend on the total target number. The phenomenon is analytically traceable in the case of zero false alarm rate. This phenomenon was coined as "spookiness" due to the analogy to the "spooky action at a distance" in quantum entanglement. We also note from our own work that the multi-Bernoulli filter [15], to a lesser extent, also exhibits a similar behavior. This raises the question of whether spookiness is an artifact of the RFS formulation of the multi-target filtering problem.

The simplest way to answer this question is to implement the full multi-target Bayes filter and test it for spookiness. While there are implementations of the multi-target Bayes filter in the literature [11], [12], [9], the high computational requirements make Monte Carlo verification of experiments difficult. Fortunately, the discovery of a conjugate prior for the multi-target likelihood function based on the notion of a labeled RFS can substantially reduce the computational load via closed form solutions to the multi-target Bayes recursion [17]. In this work we use a para-Gaussian or para-normal implementation of the closed form labeled-multi-target Bayes recursion. Unlike the Gaussian implementations for PHD/CPHD filters, in the multi-target Bayes recursion the associated normal densities do not necessarily live in the same (single-target) state space, but live on the family of product spaces, hence the name para-Gaussian or para-normal. We demonstrate that the labeled multi-target Bayes filter does not exhibit the "spooky effect" observed in the PHD filters. Hence the "spooky effect" is not an artifact of the random finite set framework.

The paper is organized as follows. Section II recapitulates the random finite set formulation of the multi-target filtering problem. Section III describes the PHD, CPHD filters along with the spooky effect. In section IV we introduce labeled RFS and describe the closed form solution to the labeled multi-target Bayes recursion. Section V reports experiments with the spooky effect for the Gaussian mixture CPHD and para-Normal labeled multi-target Bayes filters, and conclusions are given in VI.

II. Background

A. Random Finite Sets

In a multi-target system, the number of targets varies in time due to the appearance and disappearance of targets.

Furthermore, the number of measurements received at each time step does not necessarily match the number of targets due to possible missed detections as well as clutter or false alarms. The objective of multi-target filtering is to jointly estimate the number of targets present and their individual states from the accumulated observations.

Suppose that at time k, there are $N(k)$ target states $x_{k,1},\ldots,x_{k,N(k)}$, each taking values in a state space $\mathbb{X}\subseteq\mathbb{R}^{n_x}$, and $M(k)$ observations $z_{k,1},\ldots,z_{k,M(k)}$ each taking values in an observation space $\mathbb{Z}\subseteq\mathbb{R}^{n_z}$. The traditional practice of stacking up individual states to form a vector of multi-target state does not admit a meaningful and mathematically consistent notion of a multi-target miss-distance, while a finite set representation does, since the distance between sets is a well understood concept [10]. Define the finite sets

$$X_k=\{x_{k,1},\ldots,x_{k,N(k)}\}\subset\mathbb{X},$$
$$Z_k=\{z_{k,1},\ldots,z_{k,M(k)}\}\subset\mathbb{Z},$$

to be the *multi-target state* and *multi-target observation* respectively.

Since the multi-target state and multi-target observation are finite sets, the concept of a *random finite set* (RFS) is required to cast the multi-target estimation problem into a Bayesian estimation framework. In plain terms, an RFS is a random variable with realizations as (unordered) finite sets, i.e. a *finite-set-valued random variable*. Mahler's Finite Set Statistics (FISST) provides the mathematical tools for characterizing and inferencing on RFSs [5], [6]. In summary, the notions of probability densities and generating functionals etc., i.e. calculus like tools applicable to RFSs are catered for with the theory of *finite set statistics* [5]–[7]. Connections and differences between the FISST and standard measure theoretic formulations of probabilities are discussed in [12].

Before proceeding to discuss common classes of RFSs, two important summary statistics for a RFS with density π are defined as follows:

- The *probability hypothesis density* (PHD) or *intensity function* of an RFS X on $\mathbb{X}$ is the function $v:\mathbb{X}\rightarrow[0,\infty)$ such that

  $$\mathbb{E}[|X\cap B|]=\int_B v(x)dx$$

  for any $B\subseteq\mathbb{X}$, where $|X|$ denotes the cardinality of a set X. The PHD at a point x gives the density of the expected number of points of X occurring at x. The PHD v is indeed the first moment of π.
- The *cardinality distribution* ρ of an RFS X on $\mathcal{X}$ is the discrete distribution characterizing the random number of points of X, i.e. for any $n\in\mathbb{N}$

  $$\rho(n)=\mathbb{P}(|X|=n)$$

B. Common Classes of RFSs

The following is a summary of common classes or distributions of RFSs which are commonly encountered in multi-target filtering.

1) Poisson RFSs: An RFS X on $\mathbb{X}$ is said to be *Poisson* with a given *intensity function* v (defined on $\mathbb{X}$) if its cardinality ,denoted as $|X|$, is Poisson distributed, with mean $\bar{N}=\int v(x)dx$, and for any finite cardinality, the elements x of X are independently and identically distributed (i.i.d.) according to the probability density $v(\cdot)/\bar{N}$ [8]. A Poisson RFS is completely characterized by its intensity function, also known in the tracking literature as the *Probability Hypothesis Density* (PHD). A Poisson RFS with intensity function v has probability density[1] (see [5] pp. 366).

$$\pi(X)=e^{-\bar{N}}v^X$$

where for any $h:\mathbb{X}\rightarrow\mathbb{R}$, $h^X\equiv\prod_{x\in X}h(x)$ with $h^{\emptyset}=1$ by convention.

2) I.I.D. Cluster RFS: An *independent and identically distributed* (i.i.d) cluster RFS X on $\mathbb{X}$ is uniquely characterized by its cardinality distribution $\rho(\cdot)$ and matching intensity function $v(\cdot)$ [8]. The cardinality distribution must satisfy $N=\sum_{n=0}^{\infty}n\rho(n)=\int v(x)dx$ but can otherwise be arbitrary, and for a given cardinality the elements of X are each independent and identically distributed (i.i.d) with probability density $v(\cdot)/N$. The probability density of an i.i.d. cluster RFS is

$$\pi(\{x_1,...,x_n\})=n!\rho(n)\left[\frac{v(\cdot)}{N}\right]^X.$$

Note that an i.i.d. cluster RFS essentially captures the spatial randomness of the Poisson RFS without the restriction of a Poisson cardinality.

3) Bernoulli RFS: A *Bernoulli* RFS on $\mathbb{X}$ has probability $1-r$ of being empty, and probability r of being a singleton whose (only) element is distributed according to a probability density p (defined on $\mathbb{X}$). The cardinality distribution of a Bernoulli RFS is a Bernoulli distribution with parameter r. The probability density of a Bernoulli RFS is given by (see [5] pp. 368)

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

4) Multi-Bernoulli RFS: A *multi-Bernoulli* RFS X on $\mathbb{X}$ is a union of a fixed number of independent Bernoulli RFSs $X^{(i)}$ with existence probability $r^{(i)}\in(0,1)$ and probability density $p^{(i)}$ (defined on $\mathbb{X}$), $i=1,...,M$, i.e. $X=\bigcup_{i=1}^{M}X^{(i)}$. . A multi-Bernoulli RFS is thus completely described by the multi-Bernoulli parameter set $\{(r^{(i)},p^{(i)})\}_{i=1}^{M}$. The probability density π is (see [5] pp. 368)

$$\pi(\{x_1,...,x_n\})=n!\;\pi(\emptyset)\sum_{\substack{\{i_1,...,i_n\}\\ \in\mathcal{F}_n(\{1,...,M\})}}\prod_{j=1}^{n}\frac{r^{(i_j)}p^{(i_j)}(x_j)}{1-r^{(i_j)}}.$$

where $\pi(\emptyset)=\prod_{j=1}^{M}\left(1-r^{(j)}\right)$ and $\mathcal{F}_n(\mathbb{X})$ denotes the collection of finite subsets of $\mathbb{X}$ with exactly n elements.

[1]For simplicity, in this paper, we shall not distinguish a FISST set derivative of a belief functional and a probability density. While the former is not a probability density [5], it is, equivalent to a probability density relative to the distribution of a Poisson RFS with unit intensity (see [12]).

C. Multi-Target Dynamical Model

Given a multi-target state X_{k-1} at time $k-1$, each target $x_{k-1} \in X_{k-1}$ either continues to exist at time k with probability $p_{S,k}(x_{k-1})$ and moves to a new state x_k with probability density $f_{k|k-1}(x_k|x_{k-1})$, or dies with probability $q_{S,k} = 1 - p_{S,k}(x_{k-1})$ and takes on the value $\emptyset$. Thus, given a state $x_{k-1} \in X_{k-1}$ at time $k-1$, its behavior at time k is modeled by the Bernoulli RFS

$$S_{k|k-1}(x_{k-1}).$$

Assuming that, conditional on X_{k-1}, each of the above Bernoulli RFSs are mutually independent, then the survival or death of all existing targets from time $k-1$ to time k is modeled by the multi-Bernoulli RFS

$$T_{k|k-1}(X_{k-1}) = \bigcup_{x_{k-1} \in X_{k-1}} S_{k|k-1}(x_{k-1}).$$

The appearance of new targets at time k is modeled by an RFS of spontaneous births Γ_k (with cardinality distribution $\rho_{\Gamma,k}$) which is usually specified as an i.i.d. cluster RFS with intensity function γ_k and cardinality distribution $\rho_{\Gamma,k}$. Consequently, the RFS multi-target state X_k at time k is given by the union

$$X_k = T_{k|k-1}(X_{k-1}) \cup \Gamma_k.$$

D. Multi-Target Measurement Model

Given a multi-target state X_k at time k, each target $x_k \in X_k$, at time k, is either detected with probability $p_{D,k}(x_k)$ and generates an observation z_k with likelihood $g_k(z_k|x_k)$, or missed with probability $q_{D,k}(x_k) = 1 - p_{D,k}(x_k)$ and generates the value $\emptyset$, i.e. each target $x_k \in X_k$ generates a Bernoulli RFS

$$D_k(x_k).$$

Assuming that, conditional on X_k, the above Bernoulli RFSs are mutually independent, then the generation and detection of measurements at time k is modeled by the multi-Bernoulli RFS

$$\Theta_k(X_k) = \bigcup_{x_k \in X_k} D_k(x_k).$$

In addition, the sensor receives a set of false/spurious measurements or clutter, modeled by an RFS K_k, which is usually specified as an i.i.d. cluster RFS with intensity function κ_k (and normalized intensity function c_k) and cardinality distribution $\rho_{K,k}$. Consequently, at time k, the multi-target measurement Z_k generated by the multi-target state X_k is formed by the union

$$Z_k = \Theta_k(X_k) \cup K_k.$$

E. Multi-Target Bayes Recursion

Analogous to the standard Bayes recursion which propagates the posterior density of a random vector, the multi-target Bayes recursion propagates the posterior density of the multiple target state recursively in time, via a time prediction step involving the multiple target transition density and a data update step involving the multiple target likelihood. Under the above assumptions, the Bayes multiple target recursion is given by [5]–[7]:

$$\pi_{k|k-1}(X_k|Z_{1:k-1}) = \int \breve{f}_{k|k-1}(X_k|X)\pi_{k-1}(X|Z_{1:k-1})\delta X,$$

$$\pi_k(X_k|Z_{1:k}) = \frac{\breve{g}_k(Z_k|X_k)\pi_{k|k-1}(X_k|Z_{1:k-1})}{\int \breve{g}_k(Z_k|X)\pi_{k|k-1}(X|Z_{1:k-1})\delta X},$$

where $\breve{f}_{k|k-1}(\cdot|\cdot)$ and $\breve{g}_k(\cdot|\cdot)$ are the multiple target transition and likelihood respectively (see [5]), and integrals are by the *FISST set integrals* given by

$$\int f(X)\delta X = \sum_{i=0}^{\infty} \frac{1}{i!} \int f(\{x_1, ..., x_i\}) dx_1 \cdots dx_i.$$

III. PHD and CPHD Filters

The PHD and CPHD filters are based on propagating important summary statistics of the multi-target posterior, in an attempt to alleviate the computational complexity associated with propagating the full posterior via the Bayes multi-target recursion. The following is a summary of the PHD filter which propagates only the first moment of the posterior density [6], [13], and the CPHD filter which jointly propagates the first moment and cardinality distribution [7], [14]. The essential difference between the two is that the former propagates cardinality information with only one parameter whereas the latter propagates and feeds back full cardinality information. The PHD filter is computationally less expensive, whereas the CPHD filter has more accurate cardinality estimation.

A. PHD Filter

If v_{k-1} is posterior intensity at time $k-1$, then $v_{k|k-1}$, the predicted intensity $v_{k|k-1}$ to time $k-1$ is

$$v_{k|k-1}(x) = \gamma_k(x) + \int p_{S,k}(\zeta) f_{k|k-1}(x|\zeta) v_{k-1}(\zeta) d\zeta$$

and if Z_k is the measurement set received at time k,then the updated or posterior intensity v_k at time k is

$$v_k(x) = q_{D,k}(x) v_{k|k-1}(x) + \sum_{z \in Z_k} \frac{p_{D,k}(x) g_k(z|x) v_{k|k-1}(x)}{\kappa_k(z) + \int p_{D,k}(\xi) g_k(z|\xi) v_{k|k-1}(\xi) d\xi}.$$

Tractable implementations of the PHD filter have been proposed via particle representations in [12] and Gaussian mixture closed form solutions in [13].

B. CPHD Filter

If v_{k-1} is posterior intensity and ρ_{k-1} is the posterior cardinality distribution at time $k-1$, then the prediction to time k is given by

$$\rho_{k|k-1}(n) = \sum_{j=0}^{n} \rho_{\Gamma,k}(n-j) \sum_{\ell=j}^{\infty} C_j^{\ell} \frac{\langle p_{S,k}, v_{k-1} \rangle^j \langle q_{S,k}, v_{k-1} \rangle^{\ell-j}}{\langle 1, v_{k-1} \rangle^{\ell}} \rho_{k-1}(\ell),$$

$$v_{k|k-1}(x) = \gamma_k(x) + \int p_{S,k}(\zeta) f_{k|k-1}(x|\zeta) v_{k-1}(\zeta) d\zeta,$$

and if Z_k is the measurement set received at time k,then the updated at time k is given by

$$\rho_k(n) = \frac{\mathcal{P}_k(Z|n)\rho_{k|k-1}(n)}{\mathcal{P}_k(Z)},$$

$$v_k(x) = \frac{\left[q_{D,k}(x)\mathcal{P}_k(Z|\bar{D}) + p_{D,k}(x)\mathcal{P}_k(Z|D)\right] v_{k|k-1}(x)}{\mathcal{P}_k(Z)}$$

where

$$\mathcal{P}_k(Z|n) = \Upsilon_k^0[v_{k|k-1}, Z_k](n)$$

$$\mathcal{P}_k(Z) = \langle \mathcal{P}(Z|\cdot), \rho_{k|k-1}\rangle$$

$$\mathcal{P}_k(Z|\bar{D}) = \langle \Upsilon_k^1[v_{k|k-1}, Z_k], \rho_{k|k-1}\rangle$$

$$\mathcal{P}_k(Z|D) = \sum_{z\in Z_k} \frac{g_k(z|x)\langle \Upsilon_k^1[v_{k|k-1}, Z_k\backslash\{z\}], \rho_{k|k-1}\rangle}{\kappa_k(z)/\langle 1, \kappa_k\rangle}$$

$$\Upsilon_k^u[v, Z](n) = \sum_{j=0}^{\min(|Z|,n)} (|Z|-j)!\rho_{K,k}(|Z|-j)P_{j+u}^n \times \frac{\langle q_{D,k}, v\rangle^{n-(j+u)}}{\langle 1, v\rangle^n} e_j(\Xi_k(v,Z)),$$

$$\Xi_k(v, Z) = \left\{ \frac{\langle p_{D,k}g_k(z|\cdot), v\rangle}{\kappa_k(z)/\langle 1, \kappa_k\rangle} : z \in Z \right\}$$

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

with $e_0(Z) = 1$ by convention.

Tractable implementations of the CPHD filter via particle representations can be obtained by straightforward extension of the work in [12] as well as via Gaussian mixture closed form solutions in [14].

C. The Spooky Effect

PHD/CPHD filters exhibit a counter intuitive behavior called the spooky effect illustrated in Figure 1. Suppose that we are given a predicted PHD for a multi-target state with two targets. Consider the updated PHD with the following two scenarios. In scenario 1 both targets are detected while in scenario 2 only the far target is detected. We assume that the measurement set is the same for both scenarios apart from the misdetected measurement.

The upper and lower subplots illustrate the updated PHDs for scenarios 1 and 2 respectively. The lower plot shows a decreased PHD mass in the vicinity of the miss-detected target. This is expected of the PHD update which tries to explain a prior of two targets with measurements that only indicated one target. However, the PHD in the vicinity of the detected of target increases even though there is only one target there. This transfer of PHD mass in the vicinity of undetected target to the detected target occurs regardless of the distance between the targets. The shifted mass and the remaining mass of the undetected target depend on the total target number. There appears to be some kind of interaction of the targets at a distance

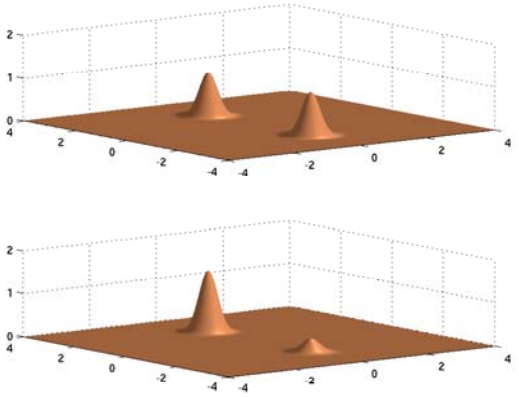

Figure 1. Illustration of the spooky effect: the PHD mass is transferred from the miss-detected target to the detected target

This phenomena was first observed by Franken et. al. in [3] and is analytically traceable in the case of zero false alarm rate. Franken et. al. illustrated using the GM-CPHD filter on a simple example consisting of two well separated tracks under zero false alarm, that the undetected track can lose 50% of its correctly updated weight, and the lost weight transferred towards the detected track regardless of how far apart the tracks are. They also note that the (uncardinalized) PHD filter exhibits a similar behavior. Depending on the target number and the value of probability of detection, the effect can be more or less pronounced. The CPHD filter update for the target number is correct while the PHD filter update for the target number is incorrect. Their overall observations are:

1) A missed detection of one track can increase the weights of the detected tracks. The lower the estimated number of targets in the surveillance region, the higher the weight gain becomes.

2) The updated weight of the miss-detected track decreases in accordance with the total number of targets in the surveillance region. The higher the estimated number of targets in the region, the lower the weight becomes.

Our own experience with the multi-Bernoulli filter [15] suggests, to a lesser extent, a similar behavior. So far there is no proper explanation for the spooky effect. It is also not known whether spookiness is an artifact of the RFS formulation of the multi-target filtering problem or if it is due to the approximations that are made with the PHD, CPHD and multi-Bernoulli filters. To answer this question we used a closed form solution of the multi-target Bayes recursion to be described next.

IV. Labeled Multi-target Bayes Filter

This section describes the conjugate prior for the multi-target likelihood function and the resulting closed form solution to the multi-target Bayes recursion [17]. The solution is underpinned by the notion of a labeled RFS, which assigns explicit unique identities to individual target states.

A. Generalized (Distinctly) Labeled Multi-Bernoulli RFS

Augment a label $\ell \in \mathbb{L}$, where $\mathbb{L}$ is the (discrete) space of labels, to the state $x \in \mathbb{X}$. For any $\tilde{X} \subset \mathbb{X}\times\ \mathbb{L}$, let $\mathcal{L}(\tilde{X})$ denote the set of labels of $\tilde{X}$, i.e. $\mathcal{L}(\tilde{X}) \triangleq \{\ell : (x,\ell) \in \tilde{X}\}$.

A labeled RFS with state space $\mathbb{X}$ and (discrete) label space $\mathbb{L}$ is an RFS on $\mathbb{X}\times\ \mathbb{L}$ such that for each realization $\tilde{X}$ we have $|\mathcal{L}(\tilde{X})| = |\tilde{X}|$.

Example 1: A labeled Poisson RFS $\tilde{X}$ with state space $\mathbb{X}$ and label space $\mathbb{L} = \mathbb{N}$ is a Poisson RFS X on $\mathbb{X}$ with intensity v tagged/augmented with labels from $\mathbb{L}$. A sample from a labeled Poisson RFS with intensity v can be generated by the following procedure:

Sampling a labeled Poisson RFS

- initialize $\tilde{X} = \emptyset$;
- sample $n \sim \text{Poiss}(\langle v, 1\rangle)$;
- for $i = 1 : n$
- sample $x \sim v(\cdot)/\langle v, 1\rangle$;
- set $\tilde{X} = \tilde{X} \cup \{(x, i)\}$;
- end;

Example 2: A labeled multi-Bernoulli $\tilde{X}$ with state space $\mathbb{X}$ and label space $\mathbb{L} = \mathbb{N}$ is a multi-Bernoulli RFS X on $\mathbb{X}$ augmented with labels corresponding to individual Bernoulli's which generate singletons. A sample from a labeled multi-Bernoulli RFS is generated from the multi-Bernoulli parameters $\{(r^{(i)}, p^{(i)})\}_{i=1}^{M}$, by the following procedure:

Sampling a labeled Multi-Bernoulli RFS

- initialize $\tilde{X} = \emptyset$;
- for $i = 1 : M$
- sample $u \sim \text{Uniform}[0, 1]$;
- if $u \leq r^{(i)}$,
- sample $x \sim p^{(i)}(\cdot)$;
- set $\tilde{X} = \tilde{X} \cup \{(x, i)\}$;
- end;
- end;

Note that the set integral for a function $\tilde{f} : \mathcal{F}(\mathbb{X} \times \mathbb{L}) \to \mathbb{R}$ is defined by

$$\int \tilde{f}(\tilde{X})\delta\tilde{X} = \sum_{i=0}^{\infty} \frac{1}{i!} \int \sum_{(\ell_1,...\ell_i)\in\mathbb{L}^i} \tilde{f}(\{(x_1,\ell_1),...,(x_i,\ell_i)\})dx_1 \cdots dx_i.$$

and that a labeled RFS is related to its unlabeled version via

$$\pi(\{x_1,...,x_n\}) = \sum_{(\ell_1,...\ell_n)\in\mathbb{L}^n} \tilde{\pi}(\{(x_1,\ell_1),...,(x_n,\ell_n)\}).$$

A generalized (distinctly) labeled multi-Bernoulli RFS is an RFS on $\mathbb{X}\times\ \mathbb{L}$ with probability density of the form

$$\tilde{\pi}(\tilde{X}) = \delta_{|\tilde{X}|}(|\mathcal{L}(\tilde{X})|) \sum_{c\in\mathbb{C}} w^{(c)} \left(\mathcal{L}(\tilde{X})\right) \left[\tilde{p}^{(c)}\right]^{\tilde{X}}$$

where $\mathbb{C}$ is a discrete index set, $w^{(c)}$ and $\tilde{p}^{(c)}$ satisfy

$$\begin{aligned} \int \tilde{p}^{(c)}(x,\ell)dx &= 1 \\ \sum_{c\in\mathbb{C}} \sum_{L\subseteq\mathbb{L}} w^{(c)}(L) &= 1, \end{aligned}$$

and $\delta_{|\tilde{X}|}(|\mathcal{L}(\tilde{X})|)$ is the distinct-label indicator of $\tilde{X}$, i.e. takes on the value 1 if the labels of $\tilde{X}$ are distinct and zero otherwise

Remark: The cardinality distribution of a generalised labeled multi-Bernoulli RFS is given by

$$\rho(n) = \sum_{c\in\mathbb{C}} \sum_{L\in\mathcal{F}_n(\mathbb{L})} w^{(c)}(L).$$

Remark: The PHD of the unlabeled version of a generalised labeled multi-Bernoulli RFS is given by

$$v(x) = \sum_{c\in\mathbb{C}} \sum_{\ell\in\mathbb{L}} p^{(c)}(x,\ell) \sum_{L\subseteq\mathbb{L}} 1_L(\ell) w^{(c)}(L).$$

B. Labeling Convention

In the context of multi-target systems, the labeling scheme adopted in this solution is described as follows. Each target is identified by a label $\ell = (k, i)$, where k is the time of birth and i is the index of the state of the target when it is born. targets born from labeled Poisson or labeled multi-Bernoulli are guaranteed to have unique labels. Using the following notations for the space of labels,

$$\begin{aligned} \mathbb{L}_k &= \{k\} \times \mathbb{N}, \\ \mathbb{L}_{0:k} &= \mathbb{L}_{0:k-1} \cup \mathbb{L}_k, \\ \mathbb{L}_k(M) &= \{k\} \times \{1,...M\}, \end{aligned}$$

a target born at time k, has state $\tilde{x} \in \mathbb{X}\times\mathbb{L}_k$ and a multi-target state at time k, is a finite subset of $\mathbb{X}\times\mathbb{L}_{0:k}$. The model for target dynamics is essentially identical to the unlabeled case, except that new born targets are labeled uniquely according to time and index of creation, and surviving targets retain their unique label. The model for observations is practically identical to the unlabeled case in the sense that measurements are statistically independent of target labels.

C. Closed Form Prediction

Following [5], the multi-target transition for surviving targets is given by

$$\tilde{f}_{S,k|k-1}(\tilde{X}|\tilde{X}') = \delta_{|\tilde{X}|}(|\mathcal{L}(\tilde{X})|) 1_{\mathcal{L}(\tilde{X}')}(\mathcal{L}(\tilde{X})) \left[\Phi_{k|k-1}(\tilde{X};\cdot)\right]^{\tilde{X}'}$$

where

$$\begin{aligned} \Phi_{k|k-1}(\tilde{X};x',\ell') = &\sum_{(x,\ell)\in\tilde{X}} \delta_{\ell'}(\ell) p_{S,k|k-1}(x',\ell') f_{k|k-1}(x|x',\ell') \\ &+ \left(1 - 1_{\mathcal{L}(\tilde{X})}(\ell')\right) q_{S,k|k-1}(x',\ell'). \end{aligned}$$

Further assume that the new born target set at any time is modeled as labeled Poisson RFS with probability density

$$f_{B,k}(\tilde{X}) = \delta_{\mathbb{L}_k(|\tilde{X}|)}(\mathcal{L}(\tilde{X})\}) \frac{e^{-\langle\gamma_k,1\rangle}}{|\tilde{X}|!} \prod_{(x,\ell)\in\tilde{X}} \gamma_k(x)$$

where γ_k is the intensity or PHD and $\mathbb{L}_k$ is the new label space at time k.

Consequently, if the prior for the multi-target state at time $k-1$ is a generalized labeled multi-Bernoulli density of the form,

$$\tilde{\pi}_{k-1}(\tilde{X}') = \delta_{|\tilde{X}'|}(|\mathcal{L}(\tilde{X}')|)\sum_{c\in\mathbb{C}} w_{k-1}^{(c)}\left(\mathcal{L}(\tilde{X}')\right)\left[\tilde{p}_{k-1}^{(c)}\right]^{\tilde{X}'}$$

Then the surviving multi-target state at time k is also a generalized labeled multi-Bernoulli with density given by

$$\tilde{\pi}_{S,k|k-1}(\tilde{X}) = \delta_{|\tilde{X}|}(|\mathcal{L}(\tilde{X})|)\sum_{c\in\mathbb{C}}\sum_{L'\subseteq\mathbb{L}_{0:k-1}} 1_{L'}(\mathcal{L}(\tilde{X}))\times$$

$$w_{k-1}^{(c)}(L')\left[q_{S,k|k-1}^{(c)}\right]^{L'}\left[\frac{K_{S,k|k-1}^{(c)}}{q_{S,k|k-1}^{(c)}}\right]^{\mathcal{L}(\tilde{X})}\left[\tilde{p}_{S,k|k-1}^{(c)}\right]^{\tilde{X}}$$

where

$$\tilde{p}_{S,k|k-1}^{(c)}(x,\ell) = \frac{\left\langle f_{k|k-1}(x|\cdot,\ell), p_{S,k|k-1}(\cdot,\ell)\tilde{p}_{k-1}^{(c)}(\cdot,\ell)\right\rangle}{K_{S,k|k-1}^{(c)}(\ell)}$$

$$K_{S,k|k-1}^{(c)}(\ell) = \int\left\langle f_{k|k-1}(x|\cdot,\ell), p_{S,k|k-1}(\cdot,\ell)\tilde{p}_{k-1}^{(c)}(\cdot,\ell)\right\rangle dx$$

$$q_{S,k|k-1}^{(c)}(\ell') = \left\langle q_{S,k|k-1}(\cdot,\ell'), \tilde{p}_{k-1}^{(c)}(\cdot,\ell')\right\rangle$$

Observe that the multi-target state at time k is the union of the surviving targets and new born targets, the multi-target transition is the convolution

$$\begin{aligned}\tilde{f}_{k|k-1}\left(\tilde{X}|\tilde{X}'\right) &= \sum_{W\subseteq\tilde{X}}\tilde{f}_{B,k}(\tilde{X}-W)\tilde{f}_{S,k|k-1}(W|\tilde{X}')\\ &= \tilde{f}_{B,k}(\tilde{X}-\tilde{X}_S)\tilde{f}_{S,k|k-1}(\tilde{X}_S|\tilde{X}')\end{aligned}$$

where $\tilde{X}_S = \{(x,\ell)\in\tilde{X} : \mathcal{L}(\tilde{X})\subset\mathbb{L}_{0:k-1}\}$. This can be verified by noting that $f_{S,k|k-1}(W|\tilde{X}')$ in the convolution sum is zero, except possibly for $W\subseteq\tilde{X}_S$, moreover, if $W\subset\tilde{X}_S, W\neq\emptyset$ then $f_{B,k}(\tilde{X}-W)$ is zero because $\tilde{X}-W$ has labels belonging to $\mathbb{L}_{0:k-1}$. Note that since the label space $\mathbb{L}_k$ of new births and the label space $\mathbb{L}_{0:k-1}$ of the surviving targets are mutually exclusive, the superposition of the birth and surviving targets is indeed a labeled RFS (this is not true in general).

The predicted multi-target density from time $k-1$ to k is consequently also a generalized (distinctly) labeled multi-Bernoulli density given by

$$\begin{aligned}\tilde{\pi}_{k|k-1}(\tilde{X}) &= \int\tilde{f}_{k|k-1}(\tilde{X}|\tilde{X}')\tilde{\pi}_{k-1}(\tilde{X}')\delta\tilde{X}'\\ &= \tilde{f}_{B,k}(\tilde{X}-\tilde{X}_S)\tilde{\pi}_{S,k|k-1}(\tilde{X}_S)\end{aligned}$$

D. Closed Form Update

From [5], note that the multi-target likelihood is given by

$$\tilde{g}(Z|\tilde{X}) = e^{-\langle\kappa,1\rangle}\kappa^Z q_D^{\tilde{X}}\sum_{\theta}[\psi_Z(\cdot;\theta)]^{\tilde{X}}$$

where the mapping $\theta : \mathcal{L}(\tilde{X})\to\{0,1,...,|Z|\}$, such that $\theta(\ell)=\alpha(\ell')>0$ implies $\ell=\ell'$, and

$$\psi_Z(x,\ell;\theta) = \begin{cases}\frac{p_D(x)}{q_D(x)}\frac{g(z_{\theta(\ell)}|x)}{\kappa(z_{\theta(\ell)})} & \theta(\ell)>0\\ 1 & \theta(\ell)=0\end{cases}$$

(recalling the assumption that measurements are statistically independent of target labels).

Consequently, if the prior distribution is a generalized labeled multi-Bernoulli of the form,

$$\tilde{\pi}(\tilde{X}) = \delta_{|\tilde{X}|}(|\mathcal{L}(\tilde{X})|)\sum_{c\in\mathbb{C}} w^{(c)}\left(\mathcal{L}(\tilde{X})\right)\left[\tilde{p}^{(c)}\right]^{\tilde{X}}$$

Then, the posterior distribution is also a generalized labeled multi-Bernoulli given by

$$\pi(\tilde{X}|Z) = \delta_{|\tilde{X}|}(|\mathcal{L}(\tilde{X})|)\sum_{c\in\mathbb{C}}\sum_{\theta}\tilde{w}^{(c,\theta)}(\mathcal{L}(\tilde{X})|Z)\left[\tilde{p}^{(c,\theta)}(\cdot|Z)\right]^{\tilde{X}}$$

where

$$\tilde{w}^{(c,\theta)}(\mathcal{L}(\tilde{X})|Z) \propto w^{(c)}(\mathcal{L}(\tilde{X}))\left[K_D^{(c,\theta)}(\cdot|Z)\right]^{\mathcal{L}(\tilde{X})}$$

$$\tilde{p}^{(c,\theta)}(x,\ell|Z) = \frac{\tilde{p}^{(c)}(x,\ell)q_D(x,\ell)\psi_Z(x,\ell;\theta)}{K_D^{(c,\theta)}(\ell|Z)}$$

$$K_D^{(c,\theta)}(\ell|Z) = \left\langle\tilde{p}^{(c)}(\cdot,\ell)q_D(\cdot,\ell),\psi_Z(\cdot,\ell;\theta)\right\rangle$$

Implementations of the prediction and update steps of the generalized labeled multi-Bernoulli filter via particle or Gaussian mixture techniques are discussed in [17]. Unlike the Gaussian-based implementations for PHD/CPHD filters, in the labeled multi-target Bayes recursion the associated Gaussians do not necessarily live in the same (single-target) state space, but live on the family of product spaces, hence the name para-Gaussian or para-Normal multi-target Bayes.

V. Experiments

In this section, we illustrate the so called "spooky effect" for the CPHD filter, and demonstrate that the para-Normal multi-target Bayes filter is immune to such effects. Consider a simple two target scenario on the two dimensional region $[-1000,1000]m\times[-1000,1000]m$. Track 1 starts in middle of the upper right quadrant, and travels directly west with a constant velocity of $-10ms^{-1}$. Track 2 starts in the middle of the lower left quadrant, and travels directly east with a constant velocity of $+10ms^{-1}$. The duration of the scenario is $K=100s$. The region and tracks are shown in Figure 2.

The kinematic target state is a vector of planar position and velocity $x_k=[\ p_{x,k},p_{y,k},\dot{p}_{x,k},\dot{p}_{y,k}\ ]^T$. Measurements are noisy vectors of planar position only $z_k=[\ z_{x,k},z_{y,k}\ ]^T$. The single-target state space model is linear Gaussian according to transition density $f_{k|k-1}(x_k|x_{k-1})=\mathcal{N}(x_k;F_kx_{k-1},Q_k)$ and likelihood $g_k(z_k|x_k)=\mathcal{N}(z_k;H_kx_k,R_k)$ with parameters

$$F_k=\begin{bmatrix}I_2 & \Delta I_2\\ 0_2 & I_2\end{bmatrix}\quad Q_k=\sigma_\nu^2\begin{bmatrix}\frac{\Delta^4}{4}I_2 & \frac{\Delta^3}{2}I_2\\ \frac{\Delta^3}{2}I_2 & \Delta^2I_2\end{bmatrix}$$

$$H_k=\begin{bmatrix}I_2 & 0_2\end{bmatrix}\qquad R_k=\sigma_\varepsilon^2I_2$$

where I_n and 0_n denote the $n\times n$ identity and zero matrices respectively, $\Delta=1s$ is the sampling period, $\sigma_\nu=5m/s^2$ and $\sigma_\varepsilon=10m$ are the standard deviations of the process noise and measurement noise. The survival probability is $p_{S,k}=0.99$ and detection probability is $p_{D,k}=0.98$. The birth model is a Poisson RFS with intensity

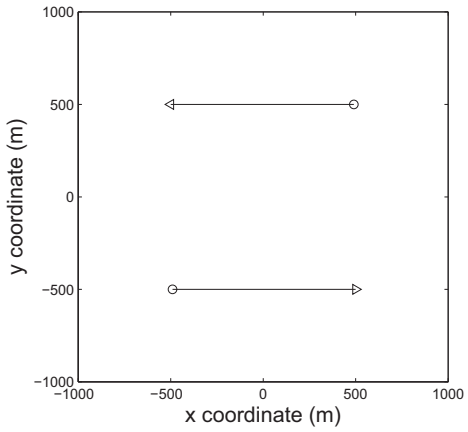

Figure 2. Trajectories in the xy plane. Start/Stop positions for each track are shown with ○/△.

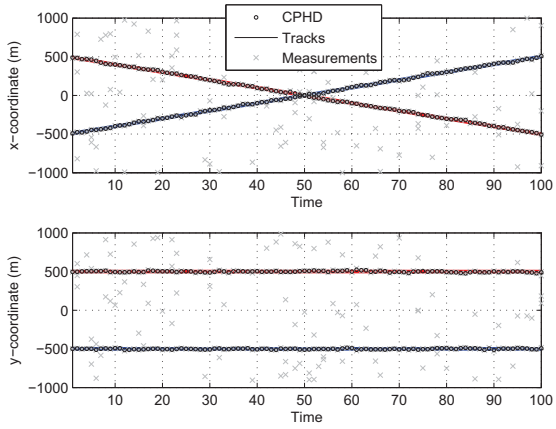

Figure 3. CPHD filter sample run output.

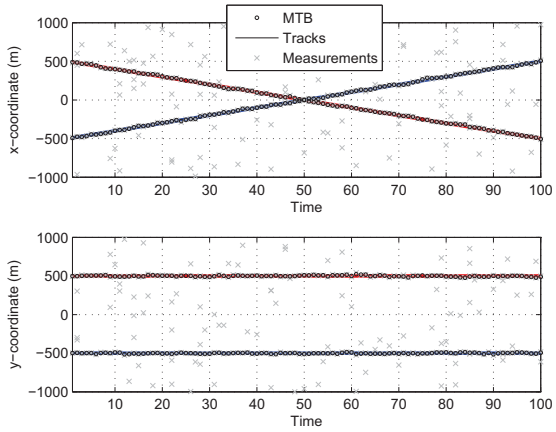

Figure 4. Para-Normal para-Normal multi-target Bayes filter sample run output.

$\gamma(x) = \sum_{i=1}^{2} w_\gamma \mathcal{N}(x; m_\gamma^{(i)}, P_\gamma)$,where $w_\gamma = 0.01$, $m_\gamma^{(1)} = [\ 500, -10, 500, 0\ ]^T$, $m_\gamma^{(2)} = [\ -500, 10, -500, 0\ ]^T$, and $P_\gamma = \text{diag}([\ 10,\ 10,\ 10,\ 10\ ]^T)^2$. Light clutter is present throughout the scenario according to a Poisson RFS with intensity $\kappa(z) = \mathcal{U}(\mathbb{Z})$ where $\mathcal{U}(\mathbb{Z})$ denotes a uniform density on the observation space. Gaussian Mixture implementations of the CPHD and para-Normal multi-target Bayes filters are used throughout.

In order to elucidate the spooky effect, two missed detections on track 1 only are deliberately inserted at times $k = 25, 75$, coinciding with the times that each tracks is approaching and departing the vertical axis. Both tracks are detected with noise at all other time instants. Figures 3 and 4 respectively show a single sample result for the CPHD and para-Normal multi-target Bayes filters. On the surface, both filters appear to track the targets effectively.

However, closer examination reveals the full effects of the CPHD approximation, as well as the relative advantages of the para-Normal multi-target Bayes filter. Over 100 Monte Carlo trials, Figure 5 shows the total PHD weight as estimated by the CPHD filter in time. The estimated number of targets is correct at 2, except at the times of missed detections $k = 25, 75$ where the total PHD weight drops slightly. Figure 6 further shows the total weight of tracks 1 and 2 as estimated by the CPHD filter. While both tracks are established and maintained, the spooky effect is clearly demonstrated, at the times of missed detections $k = 25, 75$. In this scenario, the CPHD filter almost drops track 1 pushing its weight towards zero, but shifts some weight over to track 2 pushing the total well above unity. It can be clearly seen that even though the targets are far apart, upon missed detection, the weight of track 1 is shifted and to that of track 2.

On the other hand, Figures 7 and 8 respectively show the total PHD weight as well as the individual weights for tracks 1 and 2 as estimated by the para-Normal multi-target Bayes filter. In this scenario, it can be seen that the full filter maintains the correct estimate of the total number of targets. Furthermore, even in the presence of missed detections at times $k = 25, 75$, the integrity of the tracks is maintained and the weights of both tracks remain below unity at all times. Thus the para-Normal multi-target Bayes filter does not exhibit the so-called spooky effect and does not arbitrarily transfer track weights during missed detections at times $k = 25, 75$. The explanation here is that the CPHD filter relies on an i.i.d. spatial approximation and cannot enforce a one target per track constraint, thus rendering it susceptible to the arbitrary mass shifting or the spooky effect. The para-Normal multi-target Bayes filter relies on a more complex parameterized approximation, which effectively enforces a one target per track constraint, and hence is immune to such effects.

VI. Conclusions

Our experiments have shown that the para-Normal multi-target Bayes filter does not suffer from the so-called spooky effect observed in the PHD and CPHD filters. Hence, the spooky effect is not an artifact of the random finite set formulation of the multi-target tracking problem. In view of this conclusion, the interesting question is what actually causes the spooky effect? Is it the indistinguishable nature of the elements in unlabeled random finite set? Or is it merely artifacts of the particular approximations used in the PHD, CPHD and multi-Bernoulli filters?

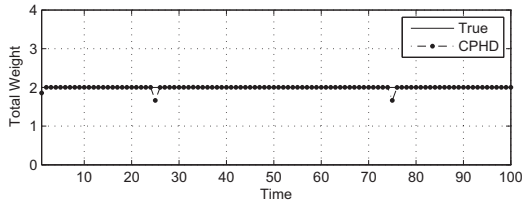

Figure 5. CPHD Filter: total estimated weight. Missed detections encountered at $k = 25, 75$. Drops in total weight observed.

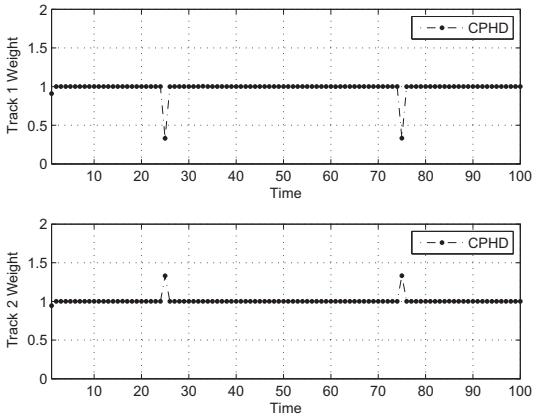

Figure 6. CPHD filter: estimated weight for individual tracks. Missed detections encountered at $k = 25, 75$, resulting in track 1 losing mass and track 2 gaining mass. This is the spooky effect.

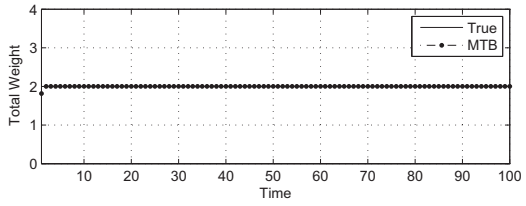

Figure 7. Para-Normal multi-target Bayesfilter: total estimated weight. Missed detections encountered at $k = 25, 75$. No effects observed.

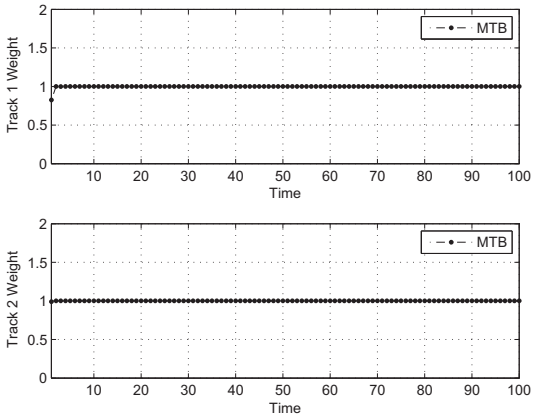

Figure 8. Para-Normal multi-target Bayes filter: estimated weight for individual tracks. Missed detections encountered at $k = 25, 75$, but no mass shifting is observed. Filter appears to be immune to the spooky effect.

REFERENCES

[1] Y. Bar-Shalom and T. E. Fortmann, *Tracking and Data Association*. Academic Press, San Diego, 1988.

[2] S. Blackman, *Multiple Target Tracking with Radar Applications*. Artech House, Norwood, 1986.

[3] D. Franken, M. Schmidt, and M. Ulmke, " "Spooky Action at a Distance" in the Cardinalized Probability Hypothesis Density Filter," *IEEE Trans. Aerospace and Electronic Systems*, vol. 45, no. 4, pp. 1657–1664, 2009.

[4] O. Erdinc, P. Willett, 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 , Nov. 2009.

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

[6] R. Mahler, "Multi-target Bayes filtering via first-order multi-target moments," *IEEE Trans. Aerospace and Electronic Systems*, vol. 39, no. 4, pp. 1152–1178, 2003.

[7] R. Mahler, "PHD filters of higher order in target number," *IEEE Trans. Aerospace & Electronic Systems*, Vol. 43, No. 3, July 2007.

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

[9] S Reuter, and K. Dietmayer, "Pedestrian Tracking Using Random Finite Sets," *Proc. Int'l Conf. Information Fusion,* Chicago IL, July 2011.

[10] 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, Aug. 2008.

[11] H. Sidenbladh, and S.-L. Wirkander, "Tracking random sets of vehicles in terrain," *Proc. 2003 IEEE Workshop on Multi-Object Tracking*, Madison WI, June, 2003.

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

[13] 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, Nov. 2006.

[14] 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, July. 2007.

[15] B.-T. Vo, B.-N. Vo, and A. Cantoni, "The cardinality balanced multi-target multi-Bernoulli filter and its implementations," *IEEE Trans. Signal Processing*, Vol. 57, No. 2, pp. 409–423, Feb. 2009.

[16] 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.

[17] B.-T. Vo, and B.-N. Vo, "A Random Finite Set Conjugate Prior and Application to Multi-Target Tracking," *Proc. IEEE Intl. Conf. Intelligent Sensors, Sensor Netorks, and Information Processing (ISSNIP)*, Adelaide, Australia, Dec. 2011.

[18] B.-T. Vo, and B.-N. Vo, "Conjugate Priors for random finite sets and Application to Multi-Target Tracking," *IEEE Trans. Signal Processing*, Submitted 2012.