Computationally-tractable approximate PHD and CPHD filters for superpositional sensors

Santosh Nannuru *Student Member, IEEE*, Mark Coates, *Senior Member, IEEE*, and Ronald Mahler

Abstract—In this paper we derive computationally-tractable approximations of the Probability Hypothesis Density (PHD) and Cardinalized Probability Hypothesis Density (CPHD) filters for superpositional sensors with Gaussian noise. We present implementations of the filters based on auxiliary particle filter approximations. As an example, we present simulation experiments that involve tracking multiple targets using acoustic amplitude sensors and a radio-frequency tomography sensor system. Our simulation study indicates that the CPHD filter provides promising tracking accuracy with reasonable computational requirements.

Index Terms—Moment based filters, PHD, CPHD, superpositional sensors, random set theory, multi-target tracking.

I. INTRODUCTION

WE address the problem of tracking multiple targets within a surveillance region based on measurements obtained from monitoring sensors. The targets can possibly enter and leave the region over time and we would like to accurately estimate the location and number of targets present at any given time. Sensors periodically provide measurements of the network which are used for the estimation of targets. Although we focus on the problem of target tracking, some other filtering problems, for example, wireless channel estimation, can also be formulated in a similar fashion [1] and the solutions discussed here can be extended to solve them.

Much of the multi-target tracking literature, particularly that component involving moment-based filters, employs the following modelling assumptions: (i) each target causes either one or no measurement; and (ii) each measurement is either caused by a single target or clutter. We refer to sensors that satisfy these assumptions to within a reasonable approximation as standard sensors [2]. These sensors form an important class of sensors but are not exhaustive. We are interested in sensors where measurements are functions of all the targets present rather than one of them. Specifically, in this environment (i) each target can contribute to any number of measurements; (ii) each measurement is potentially affected by multiple targets in an additive fashion; and (iii) measurements are not independent. We refer to this class of sensors as superpositional sensors [2]. Examples of systems belonging to this class are direction-of-arrival sensors for linear antenna arrays [3], antenna arrays in multi-user detection for wireless communication networks [1], acoustic amplitude sensors [4], and radio frequency (RF) tomographic tracking systems [5].

The recursive Bayes filter provides a systematic approach to solve the multi-target tracking problem optimally for standard or superpositional sensors. But the mathematical intractability and computational complexity of the solutions render them practically infeasible. To model and analyze the multi-target tracking problem, Mahler introduced the random finite set notation and finite set statistics (FISST) [6]. The system state is modeled as a random finite set instead of a random vector. The Bayes recursion is still intractable within this framework, but approximate solutions can be derived based on the first order moments of the posterior distribution defined over the random finite sets. The Probability Hypothesis Density (PHD) filter [7] and the Cardinalized Probability Hypothesis Density (CPHD) filter [8] are the first moment based filters derived by Mahler for the case of standard sensors. Different implementations of the PHD and CPHD filters have been proposed and applied in numerous practical applications [9]–[13].

The original PHD and CPHD filter equations are valid for standard sensors but are not applicable to superpositional sensors. CPHD filter equations for superpositional sensors were first derived in [2], but the equations involve multi-dimensional integrals, which are computationally intractable. The sensors used by Balakumar et al. in [3] for direction-of-arrival tracking are superpositional sensors, but the approach adopted is to approximate the model to obtain separable observations and then use the PHD filter for standard sensors.

In this paper we summarize our earlier derivations of computationally tractable approximations of the PHD and CPHD filters for superpositional sensors [18], [20]. In [18], Thouin, Nannuru and Coates derived a computationally tractable approximation of the PHD filter for superpositional sensors with Gaussian noise[1]. Mahler and El-Fallah derived the CPHD generalization of the ALM filter in [20]. In addition to the derivations, we present implementations based on auxiliary particle filters. Simulation experiments indicate that the approximate CPHD filter performs promisingly.

The paper is organized as follows. We formulate the problem of multiple target tracking with superpositional sensors in Section II. Section III summarizes the derivations of the approximate PHD and CPHD filters for superpositional sensors. Section IV discusses the auxiliary particle filter based implementation. Section V describes simulation experiments of multi-target tracking using acoustic amplitude sensors and a radio-frequency tomography sensor system. Section VI summarizes the current work and discusses future research directions.

---

[1]An error in the main update equation of ALM filter in [18] was corrected in an errata [19]; the correct equations were also presented in [20].

II. PROBLEM FORMULATION

The state of the system is the collection of individual target states $\mathbf{x}_{k,i} \in R^{n_{\mathbf{x}}}$ and is denoted by the random finite set $X_k = \{\mathbf{x}_{k,1}, \ldots \mathbf{x}_{k,n_k}\}$ where $n_k \geq 0$ is number of targets present at time k. We assume that the individual target dynamics are specified according to the Markovian model of the form $\mathbf{x}_{k+1,i} = f_{k+1|k}(\mathbf{x}_{k,i}, \mathbf{u}_k)$ where $\mathbf{u}_k$ is the noise. An example of the Markovian transition kernel $f_{k+1|k}$ will be presented in Section V.

Information about the state of the system is available from the sensors as an M-dimensional measurement vector $\mathbf{z}_k = [z_k^1 \ldots z_k^M]$. The relationship between the true state of the system and the sensor observation is given by the likelihood function $h_{\mathbf{z}_k}(X_k)$. In the case of superpositional sensors it has the following form,

$$
\begin{aligned}
h_{\mathbf{z}_k}(X_k) &= h_{\mathbf{z}_k}\left(r(X_k)\right) \\
&= h_{\mathbf{z}_k}\left(\sum_{\mathbf{x} \in X_k} g(\mathbf{x})\right)
\end{aligned}
\tag{1}
$$

where $h_{\mathbf{z}_k}$ is a real-valued function and g and r are (potentially non-linear) functions mapping to vectors of reals. The function r operates on the random finite set whereas the function g operates on the target states that are members of the set. Below we present two examples of the functions r and g which are based on empirical models for acoustic amplitude sensors and radio-frequency sensors, respectively.

In this paper we focus on the case where the likelihood function has the following Gaussian form:

$$
h_{\mathbf{z}_k}\left(\sum_{\mathbf{x} \in X_k} g(\mathbf{x})\right) = \mathcal{N}_{\Sigma_r}\left(\mathbf{z}_k - \sum_{\mathbf{x} \in X_k} g(\mathbf{x})\right).
\tag{2}
$$

where the notation $\mathcal{N}_{\Sigma}(\mathbf{x})$ denotes evaluation at $\mathbf{x}$ of a zero-mean Gaussian distribution with covariance matrix Σ. We will use the notation Σ_r throughout to denote the covariance of the measurement noise. Although the Gaussian noise assumption is not essential for deriving update equations, it is an important contributing factor for computational tractable approximations.

A. Example: acoustic amplitude sensors

The likelihood model of acoustic amplitude sensors is adapted from [4]. These sensors can be used in an active tracking system in which each target emits an acoustic signal of known amplitude A. Every sensor in the surveillance region receives the signal. Sensor j at location ζ_j receives the acoustic signal which has a reduced strength of $g^j(\mathbf{x}) = A/\|\mathbf{x} - \zeta_j\|^\kappa$ where $\mathbf{x}$ denotes the location of target and κ is the path loss exponent. Here g^j refers to the jth component of the vector-valued function g. When multiple targets are present, the strength of the combined signal received by each of the sensors is the sum of the strength of the signals due to each of the individual targets. The observation when multiple targets are present is given by:

$$
\begin{aligned}
z_k^j &= r^j(X_k) + \mathbf{v}_k^j \tag{3} \\
&= \sum_{\mathbf{x} \in X_k} \frac{A}{\|\mathbf{x} - \zeta_j\|^\kappa} + \mathbf{v}_k^j \tag{4}
\end{aligned}
$$

where $\mathbf{v}_k^j$ is the measurement noise, assumed to be Gaussian.

B. Example: radio-frequency tomography

Radio-frequency (RF) tomography systems strive to passively track moving objects within a surveillance region by iteratively transmitting RF signals from each sensor and measuring the attenuations at all of the other sensors. At every time instant k, the sensors communicate with each other and record the Received Signal Strength (RSS) values. The N sensors form a total of $M = N(N-1)/2$ communication links generating M measurements in every time step. During a period with an empty surveillance region, the sensor system learns background RSS values for each link. The objective of RF tomography is to use the measured deviations from these background RSS values to track moving targets.

In [22], a single target measurement model was proposed for RF tomography based on experimental analysis. The mean RSS attenuation on the link j due to a target at position $\mathbf{x}$ is modelled as:

$$
g^j(\mathbf{x}) = \phi \exp\left(-\frac{\lambda_j(\mathbf{x})}{\sigma_\lambda}\right)
\tag{5}
$$

where $\lambda_j(x)$ is an elliptical distance measure between a target located at $\mathbf{x}$ and link j (see [22] for more details); ϕ and σ_λ are fixed parameters based on physical properties of the sensors that have been learned empirically. A target located far from link j has a large associated $\lambda_j(\mathbf{x})$ and hence causes minimal additional attenuation $g^j(\mathbf{x})$; when the target is close to the link it has a much greater impact.

Multi-target extensions of this model have been successfully used for tracking based on measurements collected from real sensor networks. Tracking of up to four targets was demonstrated for an outdoor RF sensor network deployment in [23]; and successful tracking was achieved, with a slightly modified model, for three targets using indoor sensor networks [24]. The multi-target models for outdoor and indoor environments differ in the choice of g but have the same superpositional form:

$$
z_k^j = r^j(X_k) + \mathbf{v}_k^j = \sum_{\mathbf{x} \in X_k} g^j(\mathbf{x}) + \mathbf{v}_k^j
\tag{6}
$$

where $\mathbf{v}_k^j$ is the sensor noise for link j, assumed Gaussian.

III. PHD AND CPHD FILTERS FOR SUPERPOSITIONAL SENSORS

At time k, given all the measurements up to time k, the complete system state is specified by the multi-target posterior $p_{k|k}(X_k|Z^{[k]})$ where $Z^{[k]} = [\mathbf{z}_1, \ldots \mathbf{z}_k]$. The optimal Bayes filter solves for the posterior distribution in a recursive manner:

$$
p_{k+1|k}(X_{k+1}|Z^{[k]}) = \int f_{k+1|k}(X_k|W) p_{k|k}(W|Z^{[k]}) \, \delta W
\tag{7}
$$

$$
p_{k+1|k+1}(X_{k+1}|Z^{[k+1]}) = \frac{h_{\mathbf{z}_{k+1}}(X_{k+1}) p_{k+1|k}(X_{k+1}|Z^{[k]})}{\int h_{\mathbf{z}_{k+1}}(X) p_{k+1|k}(X|Z^{[k]}) \, \delta X}
\tag{8}
$$

The above equations involve evaluation of set integrals and cannot be analytically solved except for very few special cases. To alleviate this problem Mahler suggested propagating the first-order moment of the multi-target posterior rather than

the complete posterior. The first-order moment of the multi-target posterior, also called the Probability Hypothesis Density (PHD) is defined as the following set integral:

$$D_{k+1|k+1}(\mathbf{x}|Z^{[k+1]}) = \int p_{k+1|k+1}(\{\mathbf{x}\} \cup W|Z^{[k+1]})\,\delta W \quad (9)$$

The filter equations which recursively solve for the PHD are known as the PHD filter. When the distribution of the cardinality of system state, i.e., $p_{k|k}^c(|X_k| = n)$, is also propagated along with the PHD $D_{k|k}$, the resulting filter is known as the CPHD filter. The PHD [7] and the CPHD [8] filters were derived for the case of standard sensors by Mahler. For the case of superpositional sensors, computationally tractable approximations of the PHD [19] and CPHD [20] filters have been recently developed.

For brevity of results, in the following subsections, we drop the explicit notation of conditional dependence on the observations. For example, we write:

$$D_{k|k}(\mathbf{x}) := D_{k|k}(\mathbf{x}|Z^{[k]}), \quad (10)$$

$$D_{k+1|k}(\mathbf{x}) := D_{k+1|k}(\mathbf{x}|Z^{[k]}) \quad (11)$$

$$p_{k|k}^c(n) := p_{k|k}^c(n|Z^{[k]}) \quad (12)$$

$$p_{k+1|k}^c(n) := p_{k+1|k}^c(n|Z^{[k]}) \quad (13)$$

We also define the normalized PHD $s(\mathbf{x}) = D(\mathbf{x})/\int D(\mathbf{x})\,d\mathbf{x}$. For the predictive PHD, this becomes:

$$s_{k+1|k}(\mathbf{x}) = \frac{D_{k+1|k}(\mathbf{x})}{N_{k+1|k}} \quad (14)$$

$$N_{k+1|k} = \int D_{k+1|k}(\mathbf{x})d\mathbf{x}. \quad (15)$$

The superpositional assumption on the likelihood model does not affect the time prediction step of the filter. Hence we can apply Mahler's general law of motion for PHDs to compute the predictive PHD [7],

$$D_{k+1|k}(\mathbf{x}) = b_{k+1|k}(\mathbf{x}) + \int p_s(\mathbf{w})f_{k+1|k}(\mathbf{x}|\mathbf{w})D_{k|k}(\mathbf{w})d\mathbf{w}$$

where $p_s(\mathbf{w})$ is the target survival probability located at $\mathbf{w}$ and $b_{k+1|k}(\mathbf{x})$ is target birth intensity at $\mathbf{x}$. Spawning of targets is not considered in our analysis.

The cardinality prediction equations for superpositional sensors and standard sensors are also the same since the likelihood function has no role. From [8]:

$$p_{k+1|k}^c(n) = \sum_{j=0}^n p_b(n-j)$$
$$\times \left(\sum_{l=j}^\infty \binom{l}{j} \frac{\langle\!\langle p_s, D_{k+1|k}\rangle\!\rangle^j \langle\!\langle 1-p_s, D_{k+1|k}\rangle\!\rangle^{l-j}}{\langle\!\langle 1, D_{k+1|k}\rangle\!\rangle^l} p_{k|k}^c(l) \right)$$

where

$$p_b(j) = \text{birth probability of } j \text{ new targets} \quad (16)$$

$$p_s(\mathbf{x}) = \text{survival probability of target at } \mathbf{x} \quad (17)$$

$$\langle\!\langle a, b\rangle\!\rangle = \int a(\mathbf{x}) \times b(\mathbf{x})\,d\mathbf{x} \quad (18)$$

When the target survival probability is constant, $p_s(\mathbf{x}) = p_s$, the above equation reduces to

$$p_{k+1|k}^c(n) = \sum_{j=0}^n p_b(n-j)\left(\sum_{l=j}^\infty \binom{l}{j}p_s^j(1-p_s)^{l-j}p_{k|k}^c(l) \right)$$

For the derivations provided in this paper, we assume that there exists some $n_0 \geq 0$ such that for all $n > n_0$, we have $p_{k|k-1}(n) < 1/n$. This assumption holds in the common case when there is a bound on the maximum number of targets.

A. Key ingredients

The primary steps in the derivations of computationally tractable approximations of the filter update equations are the application of (i) a change of variables formula; and (ii) Campbell's theorem [20], [25]. The PHD and CPHD filter update equations for superpositional sensors can be expressed as set integrals, but these are computationally intractable. Application of suitable change of variables formulae allows us to transform these set integrals into ordinary integrals.

For a real-valued function T and a multi-target distribution $f(X)$ we can apply the following change of variables formula [26, Prop. 4, p. 180]

$$\int T(r(X)) \cdot f(X)\delta X = \int T(\mathbf{y}) \cdot P(\mathbf{y})d\mathbf{y} \quad (19)$$

where g has the superpositional form $r(X) = \sum_{\mathbf{x}\in X} g(\mathbf{x})$ and $P(\mathbf{y})$ is the distribution induced by the change of variables $\mathbf{y} = r(X)$. Note that the left hand side of (19) is a set integral whereas the right hand side is an ordinary integral.

Although much simpler, evaluation of the resultant ordinary integrals remains an unpalatable computational challenge. Application of Campbell's theorem (the linear case for the PHD and the more general quadratic version for the CPHD) allows us to evaluate the mean and variance of the distribution $P(\mathbf{y})$. We can then use a Gaussian distribution as an approximation of $P(\mathbf{y})$. When combined with Gaussian sensor noise, this leads to approximate update equations that involve much less computational overhead.

In deriving update equations for the PHD filter, we apply the linear version of Campbell's theorem. Let $\boldsymbol{\mu}$ and $\mathbf{C}$ denote the mean and variance of the distribution $P(\mathbf{y})$ above. When the multi-target distribution $f(X)$ corresponds to a Poisson point process [25], we have the following relations

$$\boldsymbol{\mu} = \int g(\mathbf{x})D(\mathbf{x})\,d\mathbf{x} \quad (20)$$

$$\mathbf{C} = \int g(\mathbf{x})g(\mathbf{x})^T D(\mathbf{x})\,d\mathbf{x} \quad (21)$$

Derivation of the CPHD filter update equations requires the quadratic version of Campbell's theorem [20]. As before, let $\mathbf{C}$ denote the variance of the distribution $P(\mathbf{y})$ above. The general quadratic Campbell's theorem gives the relation:

$$\mathbf{C} = \int g(\mathbf{x})g(\mathbf{x})^T D(\mathbf{x})\,d\mathbf{x}$$
$$+ \iint g(\mathbf{x}_1)g(\mathbf{x}_2)^T \cdot [D_2(\mathbf{x}_1,\mathbf{x}_2) - D(\mathbf{x}_1)D(\mathbf{x}_1)]\,d\mathbf{x}_1\,d\mathbf{x}_2 \quad (22)$$

where $D(\mathbf{x}_1)$ and $D_2(\mathbf{x}_1, \mathbf{x}_2)$ are, respectively, the PHD and the second factorial moment density of the multi-target distribution $f(X)$. The latter is defined as

$$D_2(\mathbf{x}_1, \mathbf{x}_2) = \int f(\{\mathbf{x}_1, \mathbf{x}_2\} \cup W) \cdot \delta W \qquad (23)$$

In the special case when the multi-target distribution $f(X)$ is Poisson, $D_2(\mathbf{x}_1, \mathbf{x}_2) = D(\mathbf{x}_1)D(\mathbf{x}_1)$, which leads back to the linear version of Campbell's theorem.

Building on the generalized form of Campbell's theorem, Mahler and El-Fallah derive expressions for the mean and variance of $P(\mathbf{y})$ in (19) for three key multi-target distributions in [20]. The results are summarized in Table I.

The first result is for a general i.i.d.c. multi-target distribution, i.e., one that can be written in the form $f(X) = |X|! \cdot p(X) \cdot s^X$. Here $s(\mathbf{x}) = D(\mathbf{x})/\int D(\mathbf{x})\,d\mathbf{x}$ and s^X is the power functional of s, with $s^X = 1$ if $X = \varnothing$ and $s^X = \prod_{\mathbf{x} \in X} s(\mathbf{x})$ otherwise. In Table I, $\widetilde{\mu}$, σ^2, and $G(x)$ are the expected value, variance and probability generating function (p.g.f.) of the cardinality distribution $p(n)$ of $f(X)$. We define:

$$\hat{\mathbf{o}} = \int g(\mathbf{x}) s(\mathbf{x}) \, d\mathbf{x} \qquad (24)$$

$$\hat{\mathbf{O}} = \int g(\mathbf{x}) g(\mathbf{x})^T \cdot s(\mathbf{x}) \, d\mathbf{x} \qquad (25)$$

for the measurement function $g(\mathbf{x})$.

The second result is for a multi-target distribution derived from an i.i.d.c. $f(X)$ but with fixed cardinality, $f_n(X) = \frac{1}{p^c(n)} \cdot \delta(|X| = n) \cdot f(X)$. The third result is for the multi-target distribution $f_x(X) = \frac{f(\{\mathbf{x}\} \cup X)}{D(\mathbf{x})}$, again derived from the i.i.d.c. $f(X)$, where $\mathbf{x}$ is a fixed value.

Multi-target distribution	Mean	Variance
$f(X) = |X|! \cdot p^c(|X|) \cdot s^X$	$\widetilde{\mu} \cdot \hat{\mathbf{o}}$	$\widetilde{\mu} \cdot \hat{\mathbf{O}} + (\sigma^2 - \widetilde{\mu}) \cdot \hat{\mathbf{o}}\hat{\mathbf{o}}^T$
$f_n(X) = \frac{1}{p^c(n)}$ $\cdot \delta(|X| = n) \cdot f(X)$	$n \cdot \hat{\mathbf{o}}$	$n \cdot (\hat{\mathbf{O}} - \hat{\mathbf{o}}\hat{\mathbf{o}}^T)$
$f_x(X) = \frac{f(\{\mathbf{x}\} \cup X)}{D(\mathbf{x})}$	$\frac{G^2(1)}{\widetilde{\mu}} \cdot \hat{\mathbf{o}}$	$\frac{G^2(1)}{\widetilde{\mu}} \cdot \hat{\mathbf{O}} +$ $\left(\frac{G^3(1)}{\widetilde{\mu}} - \frac{G^2(1)^2}{\widetilde{\mu}^2} \right) \hat{\mathbf{o}}\hat{\mathbf{o}}^T$

TABLE I
MEAN $\boldsymbol{\mu}$ AND VARIANCE C FOR $P(y)$ IN (19) FOR THREE MULTI-TARGET DISTRIBUTIONS.

B. Approximate PHD filter update derivation

We now derive the update formulae for PHD filter. By definition, the PHD at time $k + 1$ is given by

$$D_{k+1|k+1}(\mathbf{x}) = \int p_{k+1|k+1}(\mathbf{x} \cup W) \delta W \qquad (26)$$

Applying Bayes rule we have:

$$D_{k+1|k+1}(\mathbf{x}) = \int \frac{p_{k+1}(\mathbf{z}_{k+1}|\{\mathbf{x}\} \cup W) p_{k+1|k}(\{\mathbf{x}\} \cup W)}{p(\mathbf{z}_{k+1}|Z^{[k]})} \delta W$$

$$= K^{-1} \int p_{k+1}(\mathbf{z}_{k+1}|\{\mathbf{x}\} \cup W) p_{k+1|k}(\{\mathbf{x}\} \cup W) \, \delta W \quad (27)$$

where the normalization constant K is given by

$$K = p(\mathbf{z}_{k+1}|Z^{[k]}) \qquad (28)$$

$$= \int p_{k+1}(\mathbf{z}_{k+1}|W) p_{k+1|k}(W) \delta W \qquad (29)$$

If we now assume that the predictive PHD $p_{k+1|k}(\{\mathbf{x}\} \cup W)$ is a Poisson process, we have

$$D_{k+1|k+1}(\mathbf{x}) = K^{-1} \times D_{k+1|k}(\mathbf{x}) \times$$
$$\int p_{k+1}(\mathbf{z}_{k+1} - g(\mathbf{x})|W) p_{k+1|k}(W) \delta W \quad (30)$$

Thus we have the pseudolikelihood function as

$$L_{\mathbf{z}_{k+1}}(\mathbf{x}) = \frac{\int p_{k+1}(\mathbf{z}_{k+1} - g(\mathbf{x})|W) p_{k+1|k}(W) \, \delta W}{\int p_{k+1}(\mathbf{z}_{k+1}|W) p_{k+1|k}(W) \, \delta W} \qquad (31)$$

Under the assumption of Gaussian sensor noise,

$$p_{k+1}(\mathbf{z}_{k+1} - g(\mathbf{x})|W) = \mathcal{N}_{\Sigma_r}(\mathbf{z}_{k+1} - g(\mathbf{x}) - r(W))$$

and applying the change of variable $\mathbf{y} = r(W)$ leads to:

$$L_{\mathbf{z}_{k+1}}(\mathbf{x}) = \frac{\int \mathcal{N}_{\Sigma_r}(\mathbf{z}_{k+1} - g(\mathbf{x}) - \mathbf{y}) P(\mathbf{y}) \, d\mathbf{y}}{\int \mathcal{N}_{\Sigma_r}(\mathbf{z}_{k+1} - \mathbf{y}) P(\mathbf{y}) \, d\mathbf{y}} \qquad (32)$$

If we approximate $P(\mathbf{y})$ using a Gaussian distribution, then using the linear version of Campbell's theorem:

$$P(\mathbf{y}) \approx \mathcal{N}_{N_{k+1|k} \hat{\Sigma}_k}(\mathbf{y} - N_{k+1|k} \hat{\mu}_k)$$

where

$$\hat{\Sigma}_k = \int g(\mathbf{x}) g(\mathbf{x})^T \cdot s_{k+1|k}(\mathbf{x}) \qquad (33)$$

$$\hat{\mu}_k = \int g(\mathbf{x}) \cdot s_{k+1|k}(\mathbf{x}) \qquad (34)$$

The pseudo-likelihood can then be simplified as:

$$L_{\mathbf{z}_{k+1}}(\mathbf{x})$$
$$= \frac{\int \mathcal{N}_{\Sigma_r}(\mathbf{z}_{k+1} - g(\mathbf{x}) - \mathbf{y}) \mathcal{N}_{N_{k+1|k} \hat{\Sigma}_k}(\mathbf{y} - N_{k+1|k} \hat{\mu}_k) d\mathbf{y}}{\int \mathcal{N}_{\Sigma_r}(\mathbf{z}_{k+1} - \mathbf{y}) \mathcal{N}_{N_{k+1|k} \hat{\Sigma}_k}(\mathbf{y} - N_{k+1|k} \hat{\mu}_k) d\mathbf{y}}$$
$$\qquad (35)$$

$$= \frac{\mathcal{N}_{\Sigma_r + N_{k+1|k} \hat{\Sigma}_k}(\mathbf{z}_{k+1} - g(\mathbf{x}) - N_{k+1|k} \hat{\mu}_k)}{\mathcal{N}_{\Sigma_r + N_{k+1|k} \hat{\Sigma}_k}(\mathbf{z}_{k+1} - N_{k+1|k} \hat{\mu}_k)} \qquad (36)$$

The update equation for the PHD filter is then

$$D_{k+1|k+1}(\mathbf{x}) = L_{\mathbf{z}_{k+1}}(\mathbf{x}) \cdot D_{k+1|k}(\mathbf{x}). \qquad (37)$$

C. Approximate CPHD filter update derivation

The Cardinalized PHD filter propagates a probability distribution of the cardinality of the random set representing the state of the system along with the PHD. The derivation of the approximate CPHD filter update equations proceeds along similar lines as that of the PHD filter, although a more general form of Campbell's theorem must be applied. Here we provide an outline of the derivation; for more detail, see [20].

The cardinality distribution is defined as:

$$p_{k+1|k+1}^c(n) = \int_{|X|=n} p_{k+1|k+1}(X) \delta X$$

$$= \frac{\int_{|X|=n} p_{k+1}(\mathbf{z}_{k+1}|X) \cdot p_{k+1|k}(X) \delta X}{p(\mathbf{z}_{k+1}|Z^{[k]})} \qquad (38)$$

Define:

$$\Sigma_k = N_{k+1|k} \cdot \hat{\Sigma}_k + (\sigma^2_{k+1|k} - N_{k+1|k}) \cdot \hat{\boldsymbol{\mu}}_k \hat{\boldsymbol{\mu}}_k^T \quad (39)$$

Using the expression for the variance in the first row of Table I, with $\widetilde{\mu} = N_{k+1|k}$, $\hat{\mathbf{O}} = \hat{\Sigma}_k$, $\sigma^2 = \sigma^2_{k|k-1}$ and $\hat{\mathbf{o}} = \hat{\boldsymbol{\mu}}_k$, we can perform a change of variables $\mathbf{y} = r(W)$ to approximate the denominator of (38):

$$p(\mathbf{z}_{k+1}|Z^{[k]}) = \int p_{k+1}(\mathbf{z}_{k+1}|W)p_{k+1|k}(W|Z^{[k]})\,\delta W$$

$$= \int \mathcal{N}_{\Sigma_r}(\mathbf{z}_{k+1} - r(W))p_{k+1|k}(W|Z^{[k]})\,\delta W$$

$$= \int \mathcal{N}_{\Sigma_r}(\mathbf{z}_{k+1} - \mathbf{y})P(\mathbf{y})\,d\mathbf{y}$$

$$\approx \mathcal{N}_{\Sigma_r + \Sigma_k}(\mathbf{z}_{k+1} - N_{k+1|k}\hat{\boldsymbol{\mu}}_k) \quad (40)$$

The numerator of (38) can be expressed as

$$\int_{|X|=n} p_{k+1}(\mathbf{z}_{k+1}|X) \cdot p_{k+1|k}(X)\delta X$$

$$= p^c_{k+1|k}(n) \int p_{k+1}(\mathbf{z}_{k+1} - \mathbf{y}) \cdot P_n(\mathbf{y})d\mathbf{y} \quad (41)$$

The distribution $P_n(\mathbf{y})$ is approximated as a Gaussian and from the results in the second row in Table I it is of the form $P_n(\mathbf{y}) \approx \mathcal{N}_{\Sigma_k^n}(\mathbf{y} - n\hat{\boldsymbol{\mu}}_k)$, where

$$\Sigma_k^n = n \cdot (\hat{\Sigma}_k - \hat{\boldsymbol{\mu}}_k \hat{\boldsymbol{\mu}}_k^T) \quad (42)$$

Thus the approximate update expression for the cardinality distribution is

$$p^c_{k+1|k+1}(n)$$

$$\approx K_c^{-1} p^c_{k+1|k}(n) \frac{\int \mathcal{N}_{\Sigma_r}(\mathbf{z}_{k+1} - \mathbf{y}) \cdot \mathcal{N}_{\Sigma_k^n}(\mathbf{y} - n\hat{\boldsymbol{\mu}}_k)d\mathbf{y}}{\mathcal{N}_{\Sigma_r + \Sigma_k}(\mathbf{z}_{k+1} - N_{k+1|k}\hat{\boldsymbol{\mu}}_k)}$$

$$= K_c^{-1} p^c_{k+1|k}(n) \frac{\mathcal{N}_{\Sigma_r + \Sigma_k^n}(\mathbf{z}_{k+1} - n\hat{\boldsymbol{\mu}}_k)}{\mathcal{N}_{\Sigma_r + \Sigma_k}(\mathbf{z}_{k+1} - N_{k+1|k}\hat{\boldsymbol{\mu}}_k)} \quad (43)$$

Here K_c is a normalizing factor, included to ensure that the updated cardinality distribution sums to 1. We have:

$$K_c = \sum_{n \geq 0} p^c_{k+1|k}(n) \frac{\mathcal{N}_{\Sigma_r + \Sigma_k^n}(\mathbf{z}_{k+1} - n\hat{\boldsymbol{\mu}}_k)}{\mathcal{N}_{\Sigma_r + \Sigma_k}(\mathbf{z}_{k+1} - N_{k+1|k}\hat{\boldsymbol{\mu}}_k)} \quad (44)$$

The assumption that there is an $n_0 \geq 0$ such that $p^c_{k+1|k}(n) < 1/n$ for all $n > n_0$ ensures that the sum converges and K_c is finite, as detailed in [20].

From (27) the PHD update is given by the expression:

$$D_{k+1|k+1}(\mathbf{x}) = \int \frac{p_{k+1}(\mathbf{z}_{k+1}|\{\mathbf{x}\} \cup W)p_{k+1|k}(\{\mathbf{x}\} \cup W)}{p(\mathbf{z}_{k+1}|Z^{[k]})}\delta W \quad (45)$$

The approximate denominator was obtained in (40). Applying the change of variable $\mathbf{y} = r(W)$, the numerator can be expressed as:

$$\int p_{k+1}(\mathbf{z}_{k+1} - g(\mathbf{x})|W)p_{k+1|k}(W)\,\delta W$$

$$= \int \mathcal{N}_{\Sigma_r}(\mathbf{z}_{k+1} - g(\mathbf{x}) - \mathbf{y})P(\mathbf{y})\,d\mathbf{y}. \quad (46)$$

Approximating $P(\mathbf{y})$ as a Gaussian, we see from the third row of Table I that it has the form $P(\mathbf{y}) \approx \mathcal{N}_{\Sigma_k^o}(\mathbf{y} - \boldsymbol{\mu}_k^o)$, where

$$\boldsymbol{\mu}_k^o = \frac{G_{k+1|k}^{(2)}(1)}{N_{k+1|k}} \cdot \hat{\boldsymbol{\mu}}_k \quad (47)$$

$$\Sigma_k^o = \frac{G_{k+1|k}^{(2)}(1)}{N_{k+1|k}} \cdot \hat{\Sigma}_k + \left(\frac{G_{k+1|k}^{(3)}(1)}{N_{k+1|k}} - \frac{G_{k+1|k}^{(2)}(1)^2}{N_{k+1|k}^2} \right) \cdot \hat{\boldsymbol{\mu}}_k \hat{\boldsymbol{\mu}}_k^T \quad (48)$$

Here $\sigma^2_{k+1|k}$, $G_{k+1|k}^{(2)}(1)$ and $G_{k+1|k}^{(3)}(1)$ are the variance, second factorial moment and third factorial moment of the predicted cardinality distribution $p^c_{k+1|k}(n)$, with

$$G_{k+1|k}(t) = \sum_{n \geq 0} p^c_{k+1|k}(n) \cdot t^n \quad (49)$$

$$G_{k+1|k}^{(n)}(t) = \frac{d^n G_{k+1|k}}{dt^n}(t). \quad (50)$$

The approximate CPHD update equations for superpositional sensors are then:

$$p^c_{k+1|k+1}(n) = l_{k+1}(n) \cdot p^c_{k+1|k}(n) \quad (51)$$

$$D_{k+1|k+1}(\mathbf{x}) = L_{\mathbf{z}_{k+1}}(\mathbf{x}) \cdot D_{k+1|k}(\mathbf{x}) \quad (52)$$

where the pseudo-likelihood functions are given by

$$l_{k+1}(n) = \frac{\mathcal{N}_{\Sigma_r + \Sigma_k^n}(\mathbf{z}_{k+1} - n\hat{\boldsymbol{\mu}}_k)}{\mathcal{N}_{\Sigma_r + \Sigma_k}(\mathbf{z}_{k+1} - N_{k+1|k} \cdot \hat{\boldsymbol{\mu}}_k)} \quad (53)$$

$$L_{\mathbf{z}_{k+1}}(\mathbf{x}) = \frac{\mathcal{N}_{\Sigma_r + \Sigma_k^o}(\mathbf{z}_{k+1} - g(\mathbf{x}) - \boldsymbol{\mu}_k^o)}{\mathcal{N}_{\Sigma_r + \Sigma_k}(\mathbf{z}_{k+1} - N_{k+1|k} \cdot \hat{\boldsymbol{\mu}}_k)} \quad (54)$$

and where expressions for Σ_k, Σ_k^n, $\boldsymbol{\mu}_k^o$ and Σ_k^o are provided in (39), (42), (47) and (48), respectively.

IV. PHD AND CPHD IMPLEMENTATIONS

Equations (37), (52), and (51) give approximate expressions for the time update of the PHD and cardinality when new observation data become available. Although the equations specify how the update should be performed, there are in general no explicit formulae to express the PHD or cardinality at every time step in known standard forms which enable easy computational processing. Hence we consider the particle based implementation of the filters, propagating over time a weighted particle approximation of the PHD (which can be seen as a scaled density). The basic bootstrap particle filter implementation struggles when new targets arrive. We therefore implement an auxiliary particle filter, which, with its look-ahead property, is able to address new target arrivals.

A. Particle implementation of PHD filter

At every time step k, the PHD is approximated by a weighted set of particles,

$$D_{k|k} \approx \sum_{i=1}^{N_p} w_k^{(i)} \delta(\mathbf{x}_k^{(i)}) \quad (55)$$

The particle PHD filter algorithm is described in Figure 1. The algorithm first calculates $\hat{N}_{p,k}$, the number of particles used to track the targets that were identified at the previous

timestep. This is set to product of the estimated number of targets from the previous timestep, $\hat{N}_{k-1}$ (with $\hat{N}_0 = 0$), and the number of particles allocated to each target, N_{ppt} (an algorithmic parameter).

The auxiliary particle filter applies the approximate PHD filter update twice at each timestep. In the first execution of the PHD filter, the existing particles are propagated, with survival probability $p_s(\mathbf{x})$, according to the dynamics. In addition, J_p new particles are added by drawing from an importance sampling distribution $p_b(\mathbf{x}_k)$ (this distribution could depend on the measurements $\mathbf{z}_k$, but it is in general difficult to construct a meaningful distribution for superpositional sensors). We assume that we can specify an intensity function $\gamma_k(\mathbf{x})$ for the spontaneous birth process. For the propagated particles, the predictive weights $w_{k|k-1}$ are set to the weights from the previous timestep multiplied by the survival probability, $p_s(\mathbf{x}_{k-1}^{(i)})w_{k-1}$. The new particles are assigned weight $\frac{\gamma_k(\mathbf{x}_k^{(i)})}{J_p p_b(\mathbf{x}_k^{(i)})}$. Using these weighted particles, we approximate the integrals in equations (33) and (34) (lines 13 and 14) and estimate $\hat{\Sigma}_k$ and $\hat{\boldsymbol{\mu}}_k$. These estimates are used to perform the PHD update (lines 17 and 18).

The weighted particle set thus obtained is used in the second execution of the PHD filter to construct an alternative sampling distribution $q(\mathbf{x}_k)$ for the particles associated with potential new targets. In our experiments, we use a q formed by drawing particles with probability $(1-p)$ from a prior proportional to the birth intensity function γ_k, and with probability p from a Gaussian mixture distribution, $GM\{w_k^{(i)}, \mathbf{x}_k^{(i)}, \Sigma_v\}_{i=\hat{N}_{p,k}+1}^{i=\hat{N}_{p,k}+J_p}$ formed by placing a weighted zero-mean Gaussian with co-variance matrix Σ_v at each particle location. The weights in the mixture are the (normalized) particle weights obtained from the first PHD filter execution.

Since the PHD has the property that its integral over the complete observation space is equal to the expected number of targets, we should have $\sum_{i=1}^{N_k} w_k^{(i)} \approx E(|X_K|)$. Due to the approximations made in order to arrive at a computationally tractable filter, however, the error can be substantial. Hence to normalize the weights appropriately, we need to estimate the number of targets from the particles. We use the Silhouette method [27] to obtain the target number estimate.

Resampling is performed to obtain an unweighted set of particles. The k-means algorithm is used to cluster the particles into groups. The number of groups is varied from 2 to N_0. The partition which has maximum of the silhouette is declared to be the cardinality estimate $\hat{N}_k$. The particles are then resampled to update the number of particles and clustering is performed to obtain target location estimates.

B. Particle implementation of the CPHD filter

Summarized in Figure 2 is the particle implementation of the auxiliary CPHD filter. The cardinality distribution is assumed to have a finite support with $p^c(n) = 0, \quad n > N_0$. The implementation is much the same as the PHD filter, but we employ the weight update equations for the CPHD filter and the cardinality distribution is also updated. The maximum a posteriori estimate of the cardinality is used as the estimate of the number of targets.

1: Initialize particles $\{w_0^{(i)}, x_0^{(i)}\}_{i=1}^{i=J_p}$
2: **for** $k = 1$ to T **do**
3: $\quad \hat{N}_{p,k} = \hat{N}_{k-1} \times N_{ppt}$
4: $\quad$ **PHD first run**
5: $\quad$ **for** $i = 1$ to $\hat{N}_{p,k}$ **do**
6: $\quad\quad$ proposal: $\mathbf{x}_k^{(i)} \sim f_{k|k-1}(\mathbf{x}_k|\mathbf{x}_{k-1}^{(i)})$
7: $\quad\quad$ $w_{k|k-1}^{(i)} = p_s(\mathbf{x}_{k-1}^{(i)})w_{k-1}^{(i)}$
8: $\quad$ **end for**
9: $\quad$ **for** $i = \hat{N}_{p,k}+1$ to $\hat{N}_{p,k}+J_p$ **do**
10: $\quad\quad$ proposal: $\mathbf{x}_k^{(i)} \sim p_b(\mathbf{x}_k)$
11: $\quad\quad$ $w_{k|k-1}^{(i)} = \frac{\gamma_k(\mathbf{x}_k^{(i)})}{J_p p_b(\mathbf{x}_k^{(i)})}$
12: $\quad$ **end for**
13: $\quad$ $\hat{\boldsymbol{\mu}}_k = \sum_j w_{k|k-1}^{(j)} g(\mathbf{x}_k^{(j)})$
14: $\quad$ $\hat{\Sigma}_k = \sum_j w_{k|k-1}^{(j)} g(\mathbf{x}_k^{(j)}) g^T(\mathbf{x}_k^{(j)})$
15: $\quad$ **for** $i = 1$ to $\hat{N}_{p,k}+J_p$ **do**
16: $\quad\quad$ weight update:
17: $\quad\quad$ $L_{\mathbf{z}_k}(\mathbf{x}_k^{(i)}) = \frac{\mathcal{N}_{\hat{\Sigma}_k+\Sigma_r}(\mathbf{z}_k - g(x_k^{(i)}) - \hat{\boldsymbol{\mu}}_k)}{\mathcal{N}_{\hat{\Sigma}_k+\Sigma_r}(\mathbf{z}_k - \hat{\boldsymbol{\mu}}_k)}$
18: $\quad\quad$ $w_k^{(i)} = w_{k|k-1}^{(i)} \times L_{\mathbf{z}_k}(\mathbf{x}_k^{(i)})$
19: $\quad$ **end for**
20: $\quad$ **PHD second run**
21: $\quad$ **for** $i = \hat{N}_{p,k}+1$ to $\hat{N}_{p,k}+J_p$ **do**
22: $\quad\quad$ proposal: $\mathbf{x}_k^{(i)} \sim q(\mathbf{x}_k)$
23: $\quad\quad$ $w_{k|k-1}^{(i)} = \frac{\gamma_k(\mathbf{x}_k^{(i)})}{J_p q(\mathbf{x}_k^{(i)})}$
24: $\quad$ **end for**
25: $\quad$ Normalize $\{w_{k|k-1}^{(i)}\}_{i=\hat{N}_{p,k}+1}^{i=\hat{N}_{p,k}+J_p}$
26: $\quad$ $\hat{\boldsymbol{\mu}}_k = \sum_j w_{k|k-1}^{(j)} g(\mathbf{x}_k^{(j)})$
27: $\quad$ $\hat{\Sigma}_k = \sum_j w_{k|k-1}^{(j)} g(\mathbf{x}_k^{(j)}) g^T(\mathbf{x}_k^{(j)})$
28: $\quad$ **for** $i = 1$ to $N \times N_{ppt}+J_p$ **do**
29: $\quad\quad$ weight update:
30: $\quad\quad$ $L_{\mathbf{z}_k}(x_k^{(i)}) = \frac{\mathcal{N}_{\hat{\Sigma}_k+\Sigma_r}(\mathbf{z}_k - g(x_k^{(i)}) - \hat{\boldsymbol{\mu}}_k)}{\mathcal{N}_{\hat{\Sigma}_k+\Sigma_r}(\mathbf{z}_k - \hat{\boldsymbol{\mu}}_k)}$
31: $\quad\quad$ $w_k^{(i)} = w_{k|k-1}^{(i)} \times L_{\mathbf{z}_k}(x_k^{(i)})$
32: $\quad$ **end for**
33: $\quad$ target number estimation:
34: $\quad$ $\hat{N}_k = \text{Silhouette}(\{w_k^{(i)}, x_k^{(i)}\}_{i=1}^{i=\hat{N}_{p,k}+J_p})$
35: $\quad$ resample step:
36: $\quad$ $\{w_k^{(i)}, x_k^{(i)}\}_{i=1}^{i=\hat{N}_{p,k}+J_p} \rightarrow \{x_k^{(i)}\}_{i=1}^{i=\hat{N}_k \times N_{ppt}}$
37: $\quad$ clustering step:
38: $\quad$ $\{\hat{x}_{k,n}\}_{n=1}^{\hat{N}_k} = \text{cluster}(\{x_k^{(i)}\}_{i=1}^{i=\hat{N}_k \times N_{ppt}}, \hat{N}_k)$
39: **end for**

Fig. 1. Auxiliary particle filter implementation of approximate PHD filter for superpositional sensors.

C. Computational complexity

We obtain theoretical computational complexities for the PHD and CPHD algorithms. The major steps in the algorithm implementation are particle propagation, weight update, cardinality prediction and update, resampling and clustering.

For the PHD filter the dominant costs are the weight update and the identification of the number of targets. The

1: Initialize particles $\{w_0^{(i)}, \mathbf{x}_0^{(i)}\}_{i=1}^{i=J_p}$
2: **for** $k = 1$ to T **do**
3:     $\hat{N}_{p,k} = \hat{N}_{k-1} \times N_{ppt}$
4:     Cardinality prediction
5:     $\overline{p_{k|k-1}^c(n) = \sum_{j=0}^n p_b(n-j) \times}$
6:     $\left(\sum_{l=j}^\infty \binom{l}{r} p_s^j (1-p_s)^{l-j} p_{k-1|k-1}^c(l) \right)$
7:     **PHD first run**
8:     **for** $i = 1$ to $\hat{N}_{p,k}$ **do**
9:       $\underline{\text{proposal: } \mathbf{x}_k^{(i)} \sim f_{k-1}(\mathbf{x}_k | \mathbf{x}_{k-1}^{(i)})}$
10:       $w_{k|k-1}^{(i)} = p_s(\mathbf{x}_{k-1}^{(i)}) w_{k-1}^{(i)}$
11:     **end for**
12:     **for** $i = \hat{N}_{p,k} + 1$ to $\hat{N}_{p,k} + J_p$ **do**
13:       $\underline{\text{proposal: } \mathbf{x}_k^{(i)} \sim p_b(\mathbf{x}_k)}$
14:       $w_{k|k-1}^{(i)} = \frac{\gamma_k(\mathbf{x}_k^{(i)})}{J_p p_b(\mathbf{x}_k^{(i)})}$
15:     **end for**
16:     $\hat{N}_{k|k-1} = \sum_j w_{k|k-1}^{(j)}$
17:     $\hat{\boldsymbol{\mu}}_k = \frac{1}{\hat{N}_{k|k-1}} \sum_j w_{k|k-1}^{(j)} g(\mathbf{x}_k^{(j)})$
18:     $\hat{\Sigma}_k = \frac{1}{\hat{N}_{k|k-1}} \sum_j w_{k|k-1}^{(j)} g(\mathbf{x}_k^{(j)}) g^T(\mathbf{x}_k^{(j)})$
19:     **for** $i = 1$ to $\hat{N}_{p,k} + J_p$ **do**
20:       weight update:
21:       $L_{\mathbf{z}_k}(x_k^{(i)}) = \frac{\mathcal{N}_{\hat{\Sigma}_k^o + \Sigma_r}(\mathbf{z}_k - g(x_k^{(i)}) - \hat{\boldsymbol{\mu}}_k^o)}{\mathcal{N}_{\hat{\Sigma}_k + \Sigma_r}(\mathbf{z}_k - \hat{N}_{k|k-1} \cdot \hat{\boldsymbol{\mu}}_k)}$
22:       $w_k^{(i)} = w_{k|k-1}^{(i)} \times L_{\mathbf{z}_k}(x_k^{(i)})$
23:     **end for**
24:     **PHD second run**
25:     **for** $i = \hat{N}_{p,k} + 1$ to $\hat{N}_{p,k} + J_p$ **do**
26:       $\underline{\text{proposal: } \mathbf{x}_k^{(i)} \sim q(\mathbf{x}_k)}$
27:       $w_{k|k-1}^{(i)} = \frac{\gamma_k(\mathbf{x}_k^{(i)})}{J_p q(\mathbf{x}_k^{(i)})}$
28:     **end for**
29:     $\hat{N}_{k|k-1} = \sum_j w_{k|k-1}^{(j)}$
30:     $\hat{\boldsymbol{\mu}}_k = \frac{1}{\hat{N}_{k|k-1}} \sum_j w_{k|k-1}^{(j)} g(x_k^{(j)})$
31:     $\hat{\Sigma}_k = \frac{1}{\hat{N}_{k|k-1}} \sum_j w_{k|k-1}^{(j)} g(x_k^{(j)}) g^T(x_k^{(j)})$
32:     **for** $i = 1$ to $N \times N_{ppt} + J_p$ **do**
33:       weight update:
34:       $L_{\mathbf{z}_k}(x_k^{(i)}) = \frac{\mathcal{N}_{\hat{\Sigma}_k^o + \Sigma_r}(\mathbf{z}_k - g(x_k^{(i)}) - \hat{\boldsymbol{\mu}}_k^o)}{\mathcal{N}_{\hat{\Sigma}_k + \Sigma_r}(\mathbf{z}_k - \hat{N}_{k|k-1} \cdot \hat{\boldsymbol{\mu}}_k)}$
35:       $w_k^{(i)} = w_{k|k-1}^{(i)} \times L_{\mathbf{z}_k}(x_k^{(i)})$
36:     **end for**
37:     Cardinality update
38:     $\overline{p_{k|k}^c(n) = p_{k|k-1}^c(n) \frac{\mathcal{N}_{\Sigma_r + \Sigma_k^n}(\mathbf{z}_k - n \cdot \hat{\boldsymbol{\mu}}_k)}{\mathcal{N}_{\Sigma_r + \Sigma_k}(\mathbf{z}_k - \hat{N}_{k|k-1} \cdot \hat{\boldsymbol{\mu}}_k)}}$
39:     target number estimation:
40:     $\overline{\hat{N}_k = \text{MAP}(p_{k|k}^c(n))}$
41:     resample step:
42:     $\overline{\{w_k^{(i)}, x_k^{(i)}\}_{i=1}^{i=\hat{N}_{p,k}+J_p} \rightarrow \{x_k^{(i)}\}_{i=1}^{i=\hat{N}_k \times N_{ppt}}}$
43:     clustering step:
44:     $\{\hat{x}_{k,n}\}_{n=1}^{\hat{N}_k} = \text{cluster}(\{x_k^{(i)}\}_{i=1}^{i=\hat{N}_k \times N_{ppt}}, \hat{N}_k)$
45: **end for**

Fig. 2. Auxiliary particle filter implementation of the approximate CPHD filter.

weight update involves estimating an $M \times M$ covariance matrix, which has complexity $\mathcal{O}(N_p M^2)$ (where N_p is the number of propagated particles), and computing its inverse, which has complexity $\mathcal{O}(M^3)$. The Silhouette method, used to estimate the number of targets, performs multiple k-means clusterings with complexity $\mathcal{O}(N_p)$ and calculates silhouettes with complexity $\mathcal{O}(N_p^2)$. If the maximum number of targets is N_0, then the combined complexity of these operations is $\mathcal{O}(N_0^2 N_p^2)$. Thus the overall complexity for one iteration of PHD filter is $\mathcal{O}(N_p M^2 + M^3 + N_0^2 N_p^2)$.

The CPHD filter additionally propagates the cardinality, but the computational requirements are minor. The CPHD weight update equations require computing multiple covariance matrices and their inverse, which is $\mathcal{O}(N_p M^2 + M^3)$. The CPHD cardinality update involves multiple matrix inversion and is $\mathcal{O}(N_0 M^3)$. Clustering is performed only once using the k-means algorithm and has computational complexity $\mathcal{O}(N_0 N_p)$. The overall complexity for one iteration of the CPHD filter is $\mathcal{O}(N_p M^2 + N_0 M^3 + N_0 N_p)$. From the expressions it can be seen that the CPHD filter computation is dominated by matrix inversions and clustering, whereas the PHD filter has additional computational requirements for estimation of the number of targets.

V. APPLICATION TO MULTI-TARGET TRACKING

We compare and demonstrate the PHD and CPHD filters for the application of multiple target tracking in superpositional sensor environments. Targets can randomly appear and disappear within the monitored region. We consider the scenarios of acoustic amplitude sensors and a radio-frequency tomography sensor system. The measurement models are discussed in Section II. The targets move within the boundaries of the monitoring area according to linear Gaussian dynamics [10]:

$$\mathbf{x}_{k+1,i} = \begin{bmatrix} 1 & 0 & T & 0 \\ 0 & 1 & 0 & T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \mathbf{x}_{k,i} + \begin{bmatrix} \frac{T^2}{2} & 0 \\ 0 & \frac{T^2}{2} \\ T & 0 \\ 0 & T \end{bmatrix} \begin{bmatrix} u_x \\ u_y \end{bmatrix}$$

where T is the sampling period and u_x, u_y are zero-mean Gaussian white noise with respective variance σ_{u_x} and σ_{u_y}. In this model, the state of each object i at time k, $\mathbf{x}_{k,i}$, is represented by a four-dimensional vector: position on the x-axis and y-axis, velocity on the x-axis and y-axis.

For simulating the reference target motion, the model parameters are set to $T = 0.25s$, $\sigma_{u_x}^2 = \sigma_{u_y}^2 = 0.35$. Figure 5(a) shows the motion of targets in the region monitored by the sensors and Figure 3 (top panel) shows how the target number evolves over time. The simulation is run for 35 time steps covering a total duration of $35 \times 0.25 = 8.75s$.

We now discuss the algorithm implementation choices. The probabilities of birth of new targets and survival of existing targets are assumed to be constant for the purpose of simulations. We use the values $p_b = 0.2$ and $p_s = 0.9$. The two velocity components of the new particles are initialized using a standard normal distribution $N(0, 1)$. In the auxiliary implementation of filters, probability $p = 0.9$ and $\Sigma_v = \sigma_v^2 \cdot I_{2 \times 2}$ where, $\sigma_v^2 = 0.25$. For practical purposes we need to assume
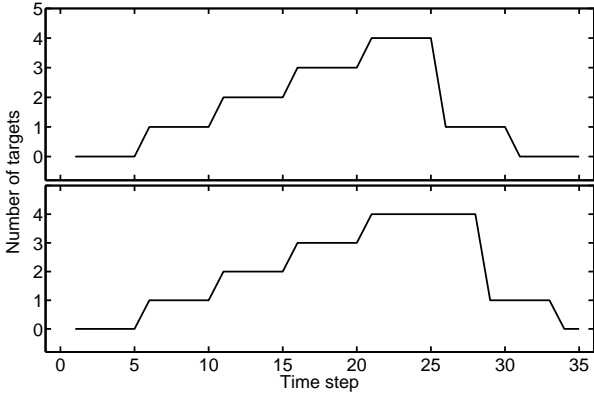
Fig. 3. Variation of number of targets over time. Top panel: the number of targets for the simulation experiments with separated target trajectories. Bottom panel: the number of targets for the simulation experiment with overlapping target trajectories.

a limit on the maximum number of targets that could be present at any give time. This limit can be chosen much higher than the true number of targets. We use a maximum of 6 targets for PHD filter and 9 targets for the CPHD filter. A smaller value is used for PHD filter as it significantly affects the computational time required for processing. We use the Silhouette method [27] to estimate the number of clusters for the PHD filter. The number of clusters is varied from 2 to 6 and the choice which maximizes the silhouette gives the estimate of the number of targets present. The method cannot identify when only a single target is present.

Identifying the number of targets in the CPHD filter implementation is much easier and maximum a posteriori (MAP) estimation is used. The peak in the cardinality distribution provides the target number estimate and computational requirements are minimal. The k-means clustering technique is then applied to group the particles into clusters and the centroids of the clusters are the target state estimates.

We compare the approximate PHD and CPHD filters with an MCMC filter that tracks the joint marginal posterior [28]. For a detailed discussion on the implementation of the MCMC filter, see [28] and [24]. The maximum number of targets is limited to 6. The burn-in is 1000 and the thinning factor is 3.

In order to compare the performance of different algorithms, we need an error metric to quantify the difference between the set of true targets present in the network and estimated target set. Since sets are involved, a root mean squared type of metric cannot be applied. We use the optimal subpattern assignment (OSPA) error metric [29] which is specifically designed for performance evaluation of multi-object filters. The OSPA metric penalizes the cardinality error in the estimates using the cardinality penalty factor c. When there are n targets and we estimate m targets then for $m \leq n$ the OSPA metric is defined as

$$d_p^{(c)}(X,Y) = \left(\frac{1}{n} \min_{\pi \in \Pi} \sum_{i=1}^{m} d^{(c)}(x_i, y_{\pi(i)})^p + c^p(n-m) \right)^{1/p}$$
(56)

where Π is the set of possible permutations of $\{1, 2, \ldots, n\}$, $d(x,y)$ is the Euclidean distance between x and y and

$d^{(c)}(x,y) = \min\{d(x,y), c\}$. $X = \{x_1, \ldots, x_n\}$ and $Y = \{y_1, \ldots, y_n\}$ are arbitrary sets and p is a fixed parameter. We use the value $p = 2$. When $m > n$, we calculate $d_p^{(c)}(Y, X)$. The OSPA metric finds the best permutation of the larger set which minimizes its distance from the smaller set and assigns a fixed penalty for each cardinality error.

The simulations are repeated multiple times with different random initializations and the average error is reported over all the simulations. The target trajectory is the same for each random initialization. A set of 20 different measurements are generated and each is processed with 5 different random initializations for all the algorithms. Thus the average error is reported over $20 \times 5 = 100$ simulations in order to reduce the variability introduced due to the stochastic nature of processing. The number of particles per target is set to $N_{ppt} = 500$ and $J_p = N_{ppt}$.

A. Acoustic amplitude sensors

The acoustic sensor likelihood model is discussed in Section II-A. The moving targets are monitored by 25 acoustic sensors distributed in a uniform grid. The targets emit a signal which has amplitide $A = 10$ at unit distance from the target. The sensors have a path loss exponent of $\kappa = 1$. When the targets lie within $d_0 = 0.2m$ distance of any sensor, the sensors record the same amplitude of A/d_0. This avoids any singularities in the measurements. The sensors are assumed to have a Gaussian noise variance of $\sigma_r^2 = 0.05$.

Table II presents the average error over 100 random initializations for the target tracks as shown above. The methods of CPHD, PHD and MCMC [28] are used for tracking. The error values are reported for different values of cardinality penalty factor ($c = 1, 2.5, 5$). The CPHD filter has the lowest OSPA error at all values of c indicating very few cardinality errors and accurate target location estimates.

Track. 1	OSPA error		
Algorithm	$c = 1$	$c = 2.5$	$c = 5$
CPHD	0.34	0.44	0.47
PHD	0.71	1.44	2.61
MCMC	0.50	0.80	0.99

TABLE II
ACOUSTIC AMPLITUDE SENSORS: AVERAGE OSPA ERROR.

Figure 4 shows the box-and-whisker plot of the error over time for the various methods. The PHD filter has a high error when the number of targets is one because the Silhoutte method used to find the number of clusters from the particles cannot estimate a single cluster. The accurate cardinality prediction using the CPHD filter is able to effectively mitigate this problem. Figures 5(a) and 5(b) show the true target trajectories and example estimated target locations as obtained using the various methods employed.

B. Radio-frequency tomography sensor system

The radio-frequency tomography sensor system is described in Section II-B. 24 radio frequency sensors are placed on the periphery of the monitoring region to form a sensor network.
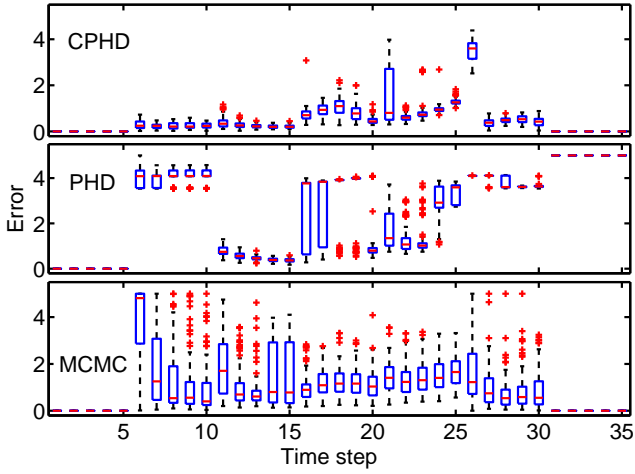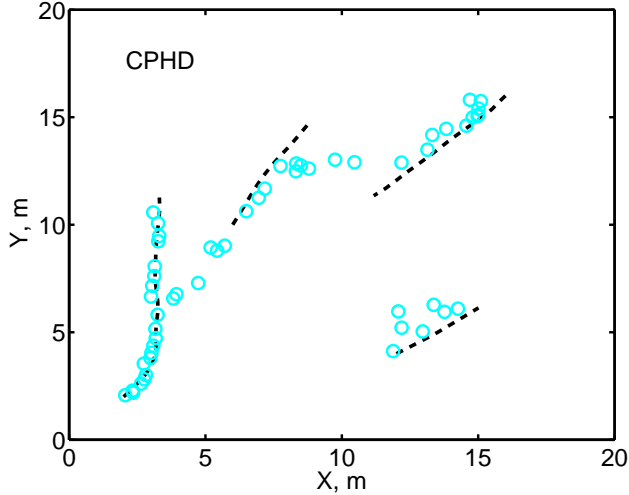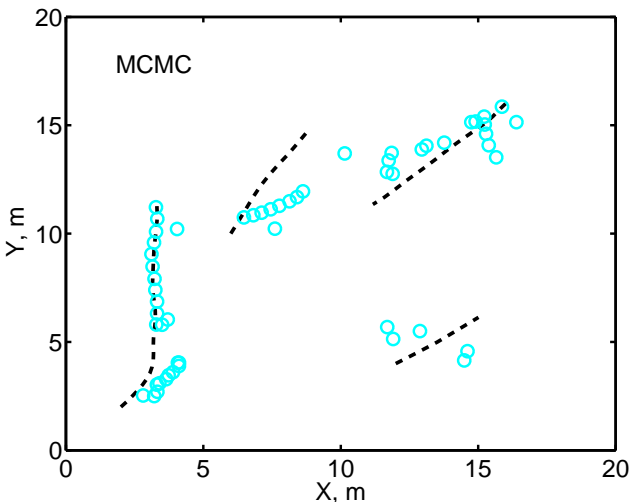
Fig. 4. Acoustic amplitude sensors: Box-and-whisker plot of the error over time for the CPHD, PHD and MCMC methods with $c = 5$. Boxes indicate 25-75 interquartile range; whiskers extend 1.5 times the range and '+' symbols indicate outliers lying beyond the whiskers.

(a) CPHD

(b) MCMC

Fig. 5. Acoustic amplitude sensors: True target tracks and example target location estimates (circles) obtained using the CPHD and MCMC methods.
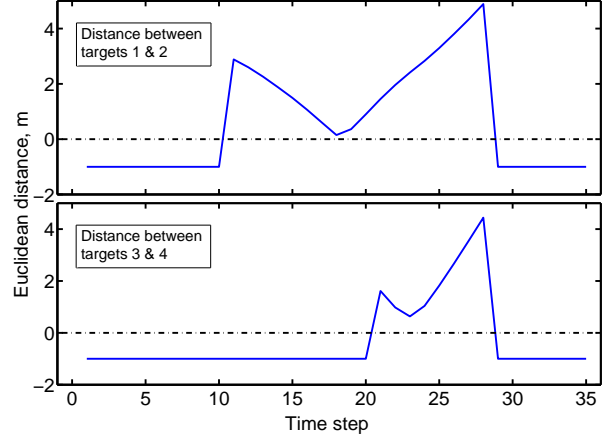

Fig. 6. Plot of Euclidean distance vs. time for target pairs 1,2 and 3,4. When either target in the pair is absent, the distance is indicated as -1.

The 24 sensors give rise to a total of 276 unique bidirectional links. The measurement parameters are $\phi = 5$ and $\sigma_\lambda = 0.2$. The Gaussian measurement noise has variance $\sigma_r^2 = 0.25$.

We use the RF tomography sensor system simulations to analyse the performance of filters when two or more targets are close in both space and time. We perform a simulation where the targets 1 and 2 and the targets 3 and 4 approach each other and then diverge (Figure 8(a)). The target number variation is shown in Figure 3 (bottom panel). Figure V-A illustrates the proximity of the target pairs, showing the Euclidean distance as a function of time for the target pairs 1, 2 and 3, 4.

The simulated observations are used to track the targets using the CPHD, PHD and MCMC algorithms. A summary of the average OSPA error, performed over 100 random simulations, is provided in Table III. The error values are reported for different values of cardinality error penalty ($c = 1, 2.5, 5$).

Overlapping trajectories and closely-spaced targets lead to higher average errors for all the algorithms, but the measurement dimension and the signal-to-noise ratio are much higher for the RF tomography setup, so the average errors are smaller than in the acoustic sensor case.

Track. 1	OSPA error		
Algorithm	$c = 1$	$c = 2.5$	$c = 5$
CPHD	0.16	0.20	0.23
PHD	0.57	1.29	2.48
MCMC	0.34	0.43	0.48

TABLE III
RADIO-FREQUENCY SENSORS: AVERAGE OSPA ERROR.

A detailed error behaviour over time can be seen from the box and whisker plot in the Figure 7. At time = 6 we observe that the MCMC filter has a much higher error median indicating difficulty in identifying the appearance of first target within the network. Also since tracking in the joint target state domain is difficult the mean error at subsequent times is higher when compared with the CPHD and PHD filters. Figures 8(a) and 8(b) plot example trajectory estimates using the different algorithms for the case of crossing targets.
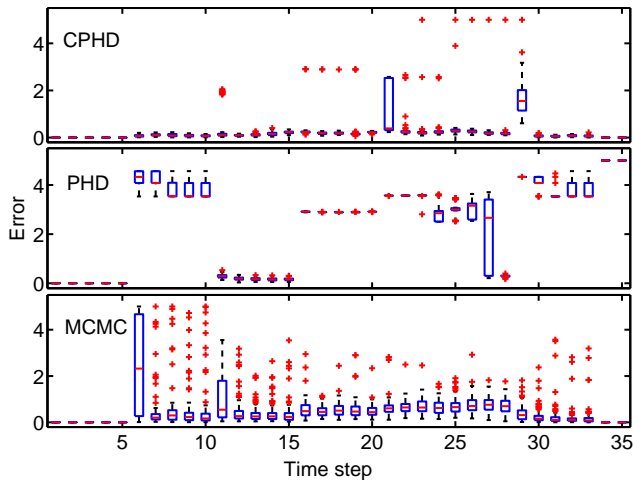
Fig. 7. Radio-frequency sensors: Box-and-whisker plot of error over time for the methods of CPHD, PHD and MCMC with $c = 5$ for overlapping target trajectories. Boxes indicate 25-75 interquartile range; whiskers extend 1.5 times the range and '+' symbols indicate outliers lying beyond the whiskers.

C. Computational requirements

Table IV summarises the computational time required[2] for each of the algorithms. The CPHD filter is the fastest and also the most accurate filter. A significant portion of the PHD filter computational time is spent towards identifying the number of targets. This could certainly be reduced substantially by adoption of an alternative technique, but in light of the good performance of the CPHD filter, we have not been motivated to conduct further exploration beyond the examination of various methods that was conducted in [21]. The MCMC filter is the slowest owing to the sequential nature of the algorithm and the fact that it operates in multi-target state space.

Algorithm	Acoustic sensors	RF Tomography
CPHD	4.15 ± 0.18	24.66 ± 0.75
PHD	108.79 ± 6.77	340.65 ± 16.57
MCMC	350.45 ± 2.52	774.45 ± 6.31

TABLE IV
CPU TIME REQUIRED IN SECONDS FOR DIFFERENT ALGORITHMS.

(a) CPHD

(b) MCMC

Fig. 8. Radio-frequency sensors: True target tracks and the target location estimates (circles) obtained from the CPHD, PHD and MCMC methods for the case of overlapping targets.
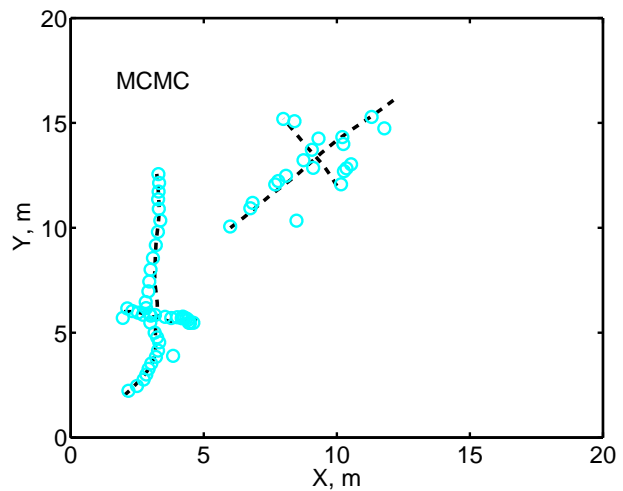
VI. CONCLUSION

We summarized the derivations of computationally-tractable approximations of the PHD and CPHD filters for superpositional sensors. The key steps in the filter derivations are the application of a change of variables and Campbell's theorem. The former allows us to shift our analysis from random sets to random variables; the latter allows us to express the first and second moments of the transformed random variables using the PHD and the second factorial moment of the multi-target distribution. We proposed auxiliary particle filter based implementations of the approximate filters and performed a simulation-based analysis of the filters using models of acoustic amplitude sensors and radio-frequency

tomography sensor systems. The CPHD filter accurately tracks the target locations and the number of targets, significantly outperforming the PHD filter which suffers from an unreliable cardinality estimate. It also outperforms a more computationally demanding joint-posterior MCMC filter. In common with most particle-based implementations of (C)PHD filters, the algorithms presented here rely on a clustering procedure to form a final estimate of the target states. This limitation motivates further investigation into more elegant solutions. Adaptation of the multi-Bernoulli filter, which was proposed by Mahler [14] and modified by Vo et al. in [15], is one promising avenue of research. Steps have been taken in this direction in [16], [17], where the multi-Bernoulli filer was applied to image observations, which have a similar likelihood structure.

[2]All the simulations were performed using algorithms implemented in Matlab on Two Xeon 4-core 2.5GHz, 14GB RAM computers.

REFERENCES

[1] D. Angelosante, E. Biglieri, and M. Lops, "Multiuser detection in a dynamic environment: Joint user identification and parameter estimation," in *Proc. IEEE Int. Symp. Inf. Theory*, Nice, France, Jun. 2007.

[2] R. Mahler, "CPHD filters for superpositional sensors," in *Proc. SPIE Signal and Data Proc. for Small Targets*, vol. 7445, San Diego, CA, USA, Aug. 2009.

[3] B. Balakumar, A. Sinha, T. Kirubarajan, and J. Reilly, "PHD filtering for tracking an unknown number of sources using an array of sensors," in *Proc. Workshop Stat. Sig. Proc.*, Bordeaux, France, Jul. 2005.

[4] O. Hlinka, O. Sluciak, F. Hlawatsch, P. M. Djuric, and M. Rupp, "Likelihood consensus and its application to distributed particle filtering," *IEEE Trans. Sig. Proc.*, vol. 60, no. 8, pp. 4334–4349, Aug. 2012.

[5] J. Wilson and N. Patwari, "Radio tomographic imaging with wireless networks," *IEEE Trans. Mobile Computing*, vol. 9, no. 5, pp. 621–632, Jan. 2010.

[6] R. Mahler, ""Statistics 101" for multisensor, multitarget data fusion," *IEEE Aerospace and Electronic Systems Magazine*, vol. 19, no. 1, pp. 53–64, Jan. 2004.

[7] ——, "Multitarget Bayes filtering via first-order multitarget moments," *IEEE Trans. Aerospace and Electronic Systems*, vol. 39, no. 4, pp. 1152–1178, Oct. 2003.

[8] ——, "PHD filters of higher order in target number," *IEEE Trans. Aerospace and Electronic Systems*, vol. 43, pp. 1523–1543, Oct. 2007.

[9] B.-N. Vo and W.-K. Ma, "The Gaussian mixture probability hypothesis density filter," *IEEE Trans. Signal Proc.*, vol. 54, pp. 4091–4104, Nov. 2006.

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

[11] D. Clark, I. Ruiz, Y. Petillot, and J. Bell, "Particle PHD filter multiple target tracking in sonar images," *IEEE Trans. Aerospace and Electronic Systems*, vol. 43, pp. 409–416, Jan. 2007.

[12] B.-T. Vo, B.-N. Vo, and A. Cantoni, "Analytic implementations of the cardinalized probability hypothesis density filter," *IEEE Trans. Sig. Proc.*, vol. 55, pp. 3553–3567, Jul. 2007.

[13] M. Ulmke, O. Erdinc, and P. Willett, "Gaussian mixture cardinalized PHD filter for ground moving target tracking," in *Proc. Int. Conf. Information Fusion*, Québec, Canada, Jul. 2007.

[14] R. Mahler, *Statistical multisource-multitarget information fusion*. Artech House, Boston, 2007.

[15] B. Vo, B. Vo, and A. Cantoni, "The cardinality balanced multi-target multi-bernoulli filter and its implementations," *IEEE Trans. Signal Proc.*, vol. 57, pp. 409–423, Feb. 2009.

[16] B. Vo, B. Vo, N. Pham, and D. Suter, "Joint detection and estimation of multiple objects from image observations," *IEEE Trans. Signal Proc.*, vol. 58, pp. 5129–5141, Oct. 2010.

[17] R. Hoseinnezhad, B. Vo, and B. Vo, "Visual tracking in background subtracted image sequences via multi-bernoulli filtering," *to appear, IEEE Trans. Signal Proc.*, 2012.

[18] F. Thouin, S. Nannuru, and M. Coates, "Multi-target tracking for measurement models with additive contributions," in *Proc. Int. Conf. Information Fusion*, Chicago, U.S.A., July 2011.

[19] ——. Multi-target tracking for measurement models with additive contributions. [Online]. Available: http://networks.ece.mcgill.ca/node/189

[20] R. Mahler and A. El-Fallah, "An approximate CPHD filter for superpositional sensors," in *Proc. SPIE Int Conf. Sig. Proc., Sensor Fusion, Target Recog.*, Baltimore, MD, USA, Apr. 2012.

[21] F. Thouin, "Bayesian inference in networks," Ph.D. dissertation, McGill University, 2011.

[22] X. Chen, A. Edelstein, Y. Li, M. Coates, M. Rabbat, and A. Men, "Sequential Monte Carlo for simultaneous passive device-free tracking and sensor localization using received signal strength measurements," in *Proc. Int. Conf. Inf. Proc. Sens. Networks*, Chicago, U.S.A., April 2011.

[23] S. Nannuru, Y. Li, M. Coates, and B. Yang, "Multi-target device-free tracking using radio frequency tomography," in *Proc. Int. Conf. Intelligent Sensors, Sensor Networks and Information Processing*, Adelaide, Australia, Dec. 2011.

[24] S. Nannuru, Y. Li, Y. Zeng, M. Coates, and B. Yang. (2012) Radio frequency tomography for passive indoor multi-target tracking. To appear, IEEE Trans. Mobile Computing. [Online]. Available: http://networks.ece.mcgill.ca/node/196

[25] J. Kingman, *Poisson Processes*. Oxford, UK: Clarendon Press, 1993.

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

[27] P. J. Rousseeuw, "Silhouettes: A graphical aid to the interpretation and validation of cluster analysis," *J. Computational and Applied Mathematics*, vol. 20, no. 27, pp. 53–65, Nov. 1987.

[28] F. Septier, S. Pang, A. Carmi, and S. Godsill, "On MCMC-based particle methods for Bayesian filtering: Application to multitarget tracking," in *Proc. IEEE Int. Work. Comp. Adv. Multi-Sensor Adaptive Proc.*, Aruba, Dutch Antilles, Dec. 2009.

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