A PHD filter for tracking multiple extended targets using random matrices

Karl Granström and Umut Orguner

Linköping University Post Print

N.B.: When citing this work, cite the original article.

A PHD filter for tracking multiple extended targets using random matrices

Karl Granström, *Member, IEEE,* and Umut Orguner, *Member, IEEE*

Abstract—This paper presents a random set based approach to tracking of an unknown number of extended targets, in the presence of clutter measurements and missed detections, where the targets' extensions are modeled as random matrices. For this purpose, the random matrix framework developed recently by Koch *et al.* is adapted into the extended target PHD framework, resulting in the Gaussian inverse Wishart PHD (GIW-PHD) filter. A suitable multiple target likelihood is derived, and the main filter recursion is presented along with the necessary assumptions and approximations. The particularly challenging case of close extended targets is addressed with practical measurement clustering algorithms. The capabilities and limitations of the resulting extended target tracking framework are illustrated both in simulations and in experiments based on laser scans.

Index Terms—Target tracking, extended target, PHD filter, random set, Gaussian distribution, inverse Wishart distribution, random matrix, laser sensor, occlusion, probability of detection.

I. INTRODUCTION

Early target tracking often made the assumption that each target can produce at most one measurement at a given time step, see e.g. [1]. With modern and more accurate sensors, the targets may occupy multiple resolution cells of the sensor, thus potentially producing more than one measurement at a given time step. Such targets are denoted extended, and tracking of extended targets has received increasing research attention over the past decade. Examples of extended target tracking include vehicle tracking using automotive radar, tracking of sufficiently close airplanes or ships with ground or marine radar stations, and person tracking using laser range sensors.

Assuming that the received target measurements are Poisson distributed in number, Gilholm and Salmond presented an approach to extended target tracking [2]. Their approach is illustrated with two examples, one in which the target is modeled as a point that may generate more than one measurement, and another example in which the target is an infinitely thin stick of length l. An inhomogeneous Poisson point process measurement model is suggested in [3], where a Poisson distributed number of measurements is distributed around the target. The model implies that the target is sufficiently far away from the sensor for the measurements to resemble a cluster rather than a geometric structure.

Karl Granström is with the Department of Electrical Engineering, Division of Automatic Control, Linköping University, Linköping, SE-581 83, Sweden, e-mail: karl@isy.liu.se.

Umut Orguner is with the Department of Electrical and Electronics Engineering, Middle East Technical University, 06531, Ankara, Turkey, e-mail: umut@eee.metu.edu.tr.

Another approach to extended target modeling is the random hypersurface model [4], which has been used to estimate elliptic targets [5]. Measurements of target down-range extent are used to aid track retention in [6]. Further approaches to estimating the target extensions, as ellipses, rectangles, or more general shapes, are given in [7]–[10].

With finite set statistics (FISST), Mahler introduced a set theoretic approach in which targets and measurements are modeled using random finite sets (RFS). The approach allows multiple target tracking in the presence of clutter and with uncertain associations to be cast in a Bayesian framework [11], resulting in an optimal multi-target Bayes filter. An important contribution of FISST is the statistical moments of the RFS, which enable practical implementation of the optimal multi-target Bayes filter. The first order moment of an RFS is called the *probability hypothesis density* (PHD), and is an intensity function defined over the target state space. The PHD filter propagates the target set's PHD in time [11], [12], and represents an approximation to the optimal multi-target Bayes filter. By approximating the PHD with a Gaussian mixture (GM), a practical implementation of the PHD filter is obtained, called the Gaussian mixture PHD (GM-PHD) filter [13]. An extension of the PHD filter to handle extended targets of the type presented in [3] is given in [14]. For the closely related area of group target tracking, in which several targets move in unison, an approach using the Gaussian mixture PHD filter, where groups are identified as targets with similar position or velocity estimates, is presented in [15]. The individual targets in a group are predicted together using a leader-follower model. A random finite set formulation of single extended target tracking is given in [16], a particle implementation is given for the general case and a closed form solution is shown for the linear Gaussian case.

A Gaussian mixture implementation of the extended target PHD filter [14], called the ET-GM-PHD-filter, has been presented in [17], with an early version given in [18]. In both of the works [17] and [18], only the kinematic properties of the targets' centroids are estimated. Estimating the targets' extents is omitted to reduce the complexity of the presentation, however this also leads to some drawbacks. In this paper, the case where the target extents are explicitly modeled and estimated along with the kinematic target states is investigated. For this purpose, we give an extended target PHD filter implementation where the target extents are represented with symmetric positive definite random matrices, i.e. the extensions are elliptical.

Using random matrices to track extended objects and groups of targets was suggested by Koch in 2008 [19]. The target

kinematical states are modeled using a Gaussian distribution, while the target extension is modeled using an inverse Wishart distribution. Using random matrices to track group targets under kinematical constraints is discussed in [20]. Modifications and improvements to the Gaussian-inverse Wishart model of [19] have been suggested in [21], and the model [19] has also been integrated into a Probabilistic Multi-Hypothesis Tracking (PMHT) framework in [22]. A comparison of random matrices and the random hypersurface model under single target assumption is given in [23].

The random matrix approach [19], to the best of our knowledge, has previously not been used in a framework for tracking an unknown number of multiple extended targets, in the presence of missed detections and clutter. The extended target PHD filter presented in this paper is capable of estimating both the kinematic states and the extents of multiple targets, in scenarios where both missed detections and clutter are allowed. At each time step, we first assume that the last estimated PHD is approximated with an unnormalized mixture of Gaussian inverse Wishart (GIW) distributions (i.e. the weights do not have to sum up to unity). We then show how the prediction and the measurement updates can be performed as was done in the single target case in [19], and also give a likelihood function suitable to handle multiple extended targets. The extended target PHD filter [14] requires all the partitions of the measurement set. As a feasible approximation, as in [17] we use only a subset of all partitions. In order to better handle spatially close targets, two additional approaches to measurement set partitioning are suggested. The resulting filter, called the Gaussian inverse Wishart PHD filter (GIW-PHD filter), is tested in simulations and in experiments based on laser scans.

The paper is organized as follows. Section II clearly specifies the extended targets of interest considered in this work and the selected extent modeling methodology. We mathematically describe the addressed target tracking problem in Section III. Section IV first lists the assumptions made, then gives the extended target PHD filter prediction and correction equations for the GIW-PHD filter, and finally presents a merging and pruning scheme for the GIW components. In this work, due to space considerations, we are not able to give all the details about the main partitioning algorithm described originally in [17]. For this reason, Section V presents only the required modifications and additions to the measurement partitioning method of [17]. Results from simulations and experiments are presented in Section VI and Section VII, and the paper is finalized with conclusions and future work in Section VIII.

II. Modeling the target extension

The extended targets considered in this work are characterized by a number of reflection points spread over their extents. Early examples of extended target tracking assume fixed measurement sources on the target, which can be tracked individually to estimate the overall lumped behavior of the extended target [24]. In many practical cases such an approach might fail, because the location of the measurement sources usually change fast according to the target sensor geometry.

Having few measurements from a single source might not be sufficient to generate good quality individual tracks. For these reasons, we avoid such an explicit estimation of the measurement sources, and instead model the global behavior of the measurements over the target extent.

As a general and simple model for the target extensions, we use ellipsoids represented by positive definite matrices, proposed by Koch in his pioneering work [19]. As admitted by Koch in [19], "*ellipsoidal object shapes are certainly a major simplification in view of large target groups which can be irregular in shape and in target density*". This remark might be considered to be true for extended targets when the targets are very close to the sensor. In this case the target features form clusters of sensor reports that are too structured to be represented accurately by ellipsoids. Nevertheless, in many real-life target tracking scenarios, the targets are neither sufficiently far from the sensors to generate only a single measurement, nor are they sufficiently close to the sensors such that their features are clearly articulated.

In this work, the targets of interest are those sufficiently far away from the sensor so that their measurements resemble a cluster of points. In Figure 1 we give an example of the ellipsoidal model applied to real laser range data. The figure shows two plots with measurements of a bicyclist and a pedestrian. While neither bikes nor humans are shaped as ellipses, we see that, given the measurements, the random matrix model is a reasonable approximation of the extensions of the bicyclist and pedestrian. In the results section of this work we also present results from experiments where multiple humans are tracked in laser range data using the ellipsoidal models.

When the target extents are modeled as ellipsoids, clearly there are many different ways to estimate the parameters of the ellipses. Classically, the target tracking problem is considered in a Bayesian framework utilizing state estimators such as Kalman filters, its extensions, and particle filters. We follow this tradition and use Koch's Bayesian random matrix methodology, where the random matrices are inverse-Wishart distributed. The inverse Wishart probability density function is a convenient prior for the considered types of measurements, and iterative update formulae for the inverse-Wishart parameters are obtained. The Bayesian framework used conveniently supplies probabilistic uncertainty measures to describe the extension estimates.

III. Target Tracking Problem Formulation

The set of extended targets at time k is denoted

$$\mathbf{X}_k = \left\{ \xi_k^{(i)} \right\}_{i=1}^{N_{x,k}}, \quad \xi_k^{(i)} \triangleq \left(\mathbf{x}_k^{(i)}, X_k^{(i)} \right), \quad (1)$$

where $N_{x,k}$ is the unknown number of targets, and, in accordance with [19], $\mathbf{x}_k^{(i)}$ is referred to as the kinematical state of the i:th target, and $X_k^{(i)}$ is referred to as the extension state. We denote the augmented state composed of the kinematic and extension states by $\xi_k^{(i)}$. Let the operation $|\cdot|$ denote set cardinality, i.e. $|\mathbf{X}_k| = N_{x,k}$. The target dynamic motion model is defined as [19]

$$\mathbf{x}_{k+1}^{(i)} = \left(F_{k+1|k} \otimes \mathbf{I}_d \right) \mathbf{x}_k^{(i)} + \mathbf{w}_{k+1}^{(i)} \quad (2)$$
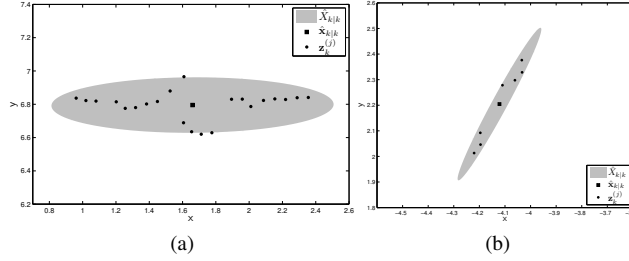
Fig. 1. The ellipsoidal extension model applied to laser range data. Measurements of a bicyclist (a) and a pedestrian (b). Both legs of the pedestrian are measured, explaining the two distinct clusters of three and four measurements, respectively. The measurements $\mathbf{z}_k^{(j)}$ are shown as black dots, the kinematical state estimates $\hat{\mathbf{x}}_{k|k}$ are shown as a black squares, and the representative extension state estimates $\hat{X}_{k|k}$ are shown as gray ellipses.

where $\mathbf{w}_{k+1}^{(i)}$ is zero mean Gaussian process noise with co-variance $\Delta_{k+1|k}^{(i)} = \mathbf{Q}_{k+1|k} \otimes X_{k+1}^{(i)}$ and d is the dimension of the target extent, i.e. $X_k^{(i)}$ is a $d \times d$ symmetric positive definite matrix and $\mathbf{I}_d$ is an identity matrix of dimension d. The notation $A \otimes B$ denotes the Kronecker product of matrices A and B. The object kinematics are modeled up to the $(s-1)$:th derivative, i.e. the length of the kinematic state vector is $n_x = s \times d$. Here $s = 3$, and $F_{k+1|k}$ and $\mathbf{Q}_{k+1|k}$ are given by [19]

$$F_{k+1|k} = \begin{bmatrix} 1 & T_s & \frac{1}{2}T_s^2 \\ 0 & 1 & T_s \\ 0 & 0 & e^{-T_s/\theta} \end{bmatrix}, \tag{3a}$$

$$\mathbf{Q}_{k+1|k} = \Sigma^2 \left(1 - e^{-2T_s/\theta} \right) \mathrm{diag} \left([0\ 0\ 1] \right), \tag{3b}$$

where T_s is the sampling time, Σ is the scalar acceleration standard deviation and θ is the maneuver correlation time.

The set of measurements obtained at time k is denoted

$$\mathbf{Z}_k = \left\{ \mathbf{z}_k^{(j)} \right\}_{j=1}^{N_{z,k}} \tag{4}$$

where $N_{z,k} = |\mathbf{Z}_k|$ is the number of measurements. The measurement model is defined as [19]

$$\mathbf{z}_k^{(j)} = (H_k \otimes \mathbf{I}_d) \mathbf{x}_k^{(i)} + \mathbf{e}_k^{(j)}, \tag{5}$$

where $\mathbf{e}_k^{(j)}$ is white Gaussian noise with covariance given by the target extension matrix $X_k^{(i)}$, and $H_k = [1\ 0\ 0]$ as in [19]. Each target generates a Poisson distributed number of measurements, where the Poisson rate $\gamma(\xi_k)$ is a function of the augmented state.

Clutter measurements are modeled as being Poisson distributed in number, with rate parameter $\beta_{FA,k}$ clutter measurements per surveillance volume per scan. With surveillance volume $\mathcal{S}$, the mean number of clutter measurements is $\beta_{FA,k}\mathcal{S}$ clutter measurements per scan. The clutter measurements are modeled as being uniformly distributed over the surveillance area.

The goal at each time step is to estimate the set of targets $\mathbf{X}_K$ given the sets of measurements $\mathbf{Z}^K = \{\mathbf{Z}_k\}_{k=1}^K$. This is achieved by propagating the predicted and updated PHDs of the set of targets $\mathbf{X}_k$, denoted $D_{k|k-1}(\cdot)$ and $D_{k|k}(\cdot)$, respectively, using the extended target PHD filter presented in [14].

IV. THE GAUSSIAN INVERSE WISHART PHD FILTER

For the multi-target tracking problem described in Section III, the extended target PHD filter prediction equations are given as follows [12].

$$D_{k+1|k}(\xi_{k+1}) = \int p_S(\xi_k) p_{k+1|k}(\xi_{k+1}|\xi_k) D_{k|k}(\xi_k) \, d\xi_k \\ + D_{k+1}^b(\xi_{k+1}), \tag{6}$$

where we omitted new target spawning[1], and

- $p_S(\cdot)$ is the probability of survival as a function of the augmented target state;
- $p_{k+1|k}(\cdot|\cdot)$ is the state transition density, describing the transition from state ξ_k to state ξ_{k+1};
- $D_k^b(\cdot)$ is the birth PHD, representing new targets.

The correction equations for the extended target PHD filter has the following form [14],

$$D_{k|k}(\xi_k|\mathbf{Z}^k) = L_{\mathbf{Z}_k}(\xi_k) D_{k|k-1}(\xi_k|\mathbf{Z}^{k-1}). \tag{7}$$

The measurement pseudo-likelihood function $L_{\mathbf{Z}_k}(\cdot)$ in (7) is defined as

$$L_{\mathbf{Z}_k}(\xi_k) \triangleq \left(1 - e^{-\gamma(\xi_k)} \right) p_D(\xi_k) + e^{-\gamma(\xi_k)} p_D(\xi_k) \tag{8}$$
$$\times \sum_{\mathrm{p} \angle \mathbf{Z}_k} \omega_\mathrm{p} \sum_{W \in \mathrm{p}} \frac{\gamma(\xi_k)^{|W|}}{d_W} \prod_{\mathbf{z}_k \in W} \frac{\phi_{\mathbf{z}_k}(\xi_k)}{\lambda_k c_k(\mathbf{z}_k)},$$

where

- $\lambda_k \triangleq \beta_{FA,k}\mathcal{S}$ is the mean number of clutter measurements;
- $c_k(\mathbf{z}_k) = 1/\mathcal{S}$ is the spatial distribution of the clutter over the surveillance volume;
- the notation $\mathrm{p} \angle \mathbf{Z}_k$ denotes that p partitions the measurement set $\mathbf{Z}_k$ into non-empty cells W. When used under a summation sign, the summation is over all possible partitions;
- the notation $W \in \mathrm{p}$ denotes that the set W is a cell in the partition p. When used under a summation sign, the summation is over all sets in the partition;

[1]More details on target spawning are given in Section VI-D.

- the quantities ω_{p} and d_W are non-negative coefficients defined, for each partition p and cell W respectively, as

$$\omega_{\mathrm{p}} = \frac{\prod_{W \in \mathrm{p}} d_W}{\sum_{\mathrm{p}' \angle \mathbf{z}_k} \prod_{W' \in \mathrm{p}'} d_{W'}}, \tag{9}$$

$$d_W = \delta_{|W|,1} + D_{k|k-1}\left[p_{\mathrm{D}} \gamma^{|W|} e^{-\gamma} \prod_{\mathbf{z}_k \in W} \frac{\phi_{\mathbf{z}_k}(\cdot)}{\lambda_k c_k(\mathbf{z}_k)} \right], \tag{10}$$

where $\delta_{i,j}$ is the Kronecker delta and the notation $f[g]$ denotes the integral $\int f(x)g(x)dx$.

- $\phi_{\mathbf{z}_k}(\xi_k) \triangleq p(\mathbf{z}_k | \xi_k)$ is the likelihood function for a single target generated measurement. Under the measurement model (5) it is given as

$$\phi_{\mathbf{z}_k}(\xi_k) = \mathcal{N}\left(\mathbf{z}_k\,;\,(H_k \otimes \mathbf{I}_d)\,\mathbf{x}_k, X_k\right). \tag{11}$$

In the following subsections, we are going to assume that we are at an intermediate stage of estimation at time t_k and the current estimated PHD $D_{k|k}(\cdot)$ can be approximated as an unnormalized mixture of Gaussian inverse Wishart (GIW) distributions as follows.

$$D_{k|k}(\xi_k) \approx \sum_{j=1}^{J_{k|k}} w_{k|k}^{(j)} \mathcal{N}\left(\mathbf{x}_k\,;\,m_{k|k}^{(j)}, P_{k|k}^{(j)} \otimes X_k\right)$$
$$\times \mathcal{IW}\left(X_k\,;\,\nu_{k|k}^{(j)}, V_{k|k}^{(j)}\right), \tag{12}$$

where

- $J_{k|k}$ is the number of components;
- $w_{k|k}^{(j)}$ is the weight of the j:th component;
- $m_{k|k}^{(j)}$ and $P_{k|k}^{(j)} \otimes X_k$ are the Gaussian mean and covariance of the j:th component;
- $\nu_{k|k}^{(j)}$ and $V_{k|k}^{(j)}$ are the inverse Wishart degrees of freedom and inverse scale matrix of the j:th component;
- the notation $\mathcal{N}(x\,;\,m, P)$ denotes a Gaussian distribution defined over the variable x with mean m and covariance P;
- the notation $\mathcal{IW}(X\,;\,\nu, V)$ denotes an inverse Wishart distribution defined over the variable X with degrees of freedom ν and inverse scale matrix V.

Further, let $\xi_{k|k}^{(j)}$ be an abbreviation of the sufficient statistics of the j:th GIW component, i.e.

$$\xi_{k|k}^{(j)} \triangleq \left(m_{k|k}^{(j)}, P_{k|k}^{(j)}, \nu_{k|k}^{(j)}, V_{k|k}^{(j)}\right). \tag{13}$$

Note that the distribution for the kinematical state x_k depends on the extension state X_k. Estimates of the kinematic state uncertainty and of the target extent are obtained as in [19],

$$\hat{P}_{k|k}^{(j)} = \frac{P_{k|k}^{(j)} \otimes V_{k|k}^{(j)}}{\nu_{k|k}^{(j)} + s - sd - 2}, \tag{14}$$

$$\hat{X}_{k|k}^{(j)} = \frac{V_{k|k}^{(j)}}{\nu_{k|k}^{(j)} - 2d - 2}, \tag{15}$$

for $\nu_{k|k}^{(j)}$ such that the denominators are positive. In the following, we give the assumptions made in the derivation of the GIW-PHD filter in Section IV-A. The prediction and

update formulas for the PHD representation in (12) are then presented in Section IV-B and Section IV-C. Finally, GIW mixture reduction using a pruning and merging scheme is addressed in Section IV-D.

A. Assumptions

In order to derive prediction and correction equations for the GIW-PHD filter, a number of assumptions are made. The first four assumptions are standard in most target tracking applications, see e.g. [1].

Assumption 1: Each target evolves and generates observations independently of all other targets. □

Assumption 2: Each target's kinematical part follows a linear Gaussian dynamical model, and the sensor has a linear Gaussian measurement model. □

Assumption 3: Clutter is Poisson distributed in number, and independent of target-originated measurements. □

Assumption 4: The survival probability is state independent, i.e. $p_{\mathrm{S}}(\xi_k) = p_{\mathrm{S}}$. □

The next assumption is reasonable in scenarios where target interactions are negligible [13].

Assumption 5: The predicted multi-target RFS is Poisson. □

In [13], [17] the PHD is represented as a mixture of Gaussian distributions, here a different assumption is made to accommodate the random matrix model.

Assumption 6: The intensity of the birth RFS is a mixture of GIW distributions. □

The following assumption is inherited from [19], where it is noted that it implies restrictions that can be justified in many practical cases.

Assumption 7: The target augmented state transition density satisfies

$$p_{k+1|k}(\xi_{k+1}|\xi_k) \approx p_{k+1|k}^1(\mathbf{x}_{k+1}|X_{k+1}, \mathbf{x}_k)\, p_{k+1|k}^2(X_{k+1}|X_k) \tag{16}$$

for all ξ_k and ξ_{k+1}. □

In addition to these, two more assumptions are made concerning the probability of detection $p_{\mathrm{D}}(\cdot)$ and the rate $\gamma(\cdot)$ that governs each target's measurement generation. These assumptions require a bit more elaboration.

Assumption 8: The following approximation about $p_{\mathrm{D}}(\cdot)$ holds for all ξ_k

$$p_{\mathrm{D}}(\xi_k) \mathcal{N}\left(\mathbf{x}_k\,;\,m_{k|k-1}^{(j)}, P_{k|k-1}^{(j)} \otimes X_k\right)$$
$$\times \mathcal{IW}\left(X_k\,;\,\nu_{k|k-1}^{(j)}, V_{k|k-1}^{(j)}\right)$$
$$\approx p_{\mathrm{D}}\left(\xi_{k|k-1}^{(j)}\right) \mathcal{N}\left(\mathbf{x}_k\,;\,m_{k|k-1}^{(j)}, P_{k|k-1}^{(j)} \otimes X_k\right)$$
$$\times \mathcal{IW}\left(X_k\,;\,\nu_{k|k-1}^{(j)}, V_{k|k-1}^{(j)}\right). \tag{17}$$

Let $p_{\mathrm{D}}^{(j)} \triangleq p_{\mathrm{D}}\left(\xi_{k|k-1}^{(j)}\right)$ abbreviate the probability of detection for the j:th GIW component. □

In Assumption 8 the approximation (17) is trivially satisfied when $p_{\mathrm{D}}(\cdot) = p_{\mathrm{D}}$, i.e. when $p_{\mathrm{D}}(\cdot)$ is constant. In general, Assumption 8 holds approximately when the function $p_{\mathrm{D}}(\cdot)$ does not vary much in the uncertainty zone of a target in the

augmented state space ξ_k. This is true either when $p_D(\cdot)$ is a sufficiently smooth function, or when the signal to noise ratio (SNR) is high enough such that the uncertainty zone is sufficiently small. A similar approach to variable probability of detection has been taken in order to model the *clutter notch* in ground moving target indicator target tracking [25].

For the expected number of measurements from the targets, represented by $\gamma(\cdot)$, similar remarks apply and the following assumption is made.

Assumption 9: The following approximation about $\gamma(\cdot)$ holds for all ξ_k, $j = 1, \ldots, J_{k|k-1}$ and all integers $n \geq 1$,

$$
\begin{aligned}
& e^{-\gamma(\xi_k)}\gamma^n(\xi_k)\mathcal{N}\left(\mathbf{x}\,;\,m_{k|k-1}^{(j)}, P_{k|k-1}^{(j)} \otimes X_k\right) \\
& \quad \times \mathcal{IW}\left(X_k\,;\,\nu_{k|k-1}^{(j)}, V_{k|k-1}^{(j)}\right) \\
& \approx e^{-\gamma\left(\xi_{k|k-1}^{(j)}\right)}\gamma^n\left(\xi_{k|k-1}^{(j)}\right)\mathcal{N}\left(\mathbf{x}\,;\,m_{k|k-1}^{(j)}, P_{k|k-1}^{(j)} \otimes X_k\right) \\
& \quad \times \mathcal{IW}\left(X_k\,;\,\nu_{k|k-1}^{(j)}, V_{k|k-1}^{(j)}\right).
\end{aligned}
\tag{18}
$$

Let $\gamma^{(j)} \triangleq \gamma\left(\xi_{k|k-1}^{(j)}\right)$ abbreviate the expected number of measurements for the j:th GIW component. $\square$

The trivial situation $\gamma(\cdot) = \gamma$, i.e. when $\gamma(\cdot)$ is constant, is again a special case where Assumption 9 is satisfied. In general, satisfying Assumption 9 is more difficult than Assumption 8. Nevertheless Assumption 9 is expected to hold approximately either when $\gamma(\cdot)$ is a sufficiently smooth function or when the signal to noise ratio (SNR) is high enough such that the uncertainty zone of a target in the augmented state space ξ_k is sufficiently small.

B. Prediction

Utilizing Assumptions 4 and 7, the prediction of existing targets can be written as

$$
\begin{aligned}
& p_S \int p_{k+1|k}^1\left(\mathbf{x}_{k+1}|X_{k+1}, \mathbf{x}_k\right) \\
& \quad \times p_{k+1|k}^2\left(X_{k+1}|X_k\right) D_{k|k}\left(\mathbf{x}_k, X_k\right) \mathrm{d}\mathbf{x}_k \mathrm{d}X_k \\
& = p_S \sum_{j=1}^{J_{k|k}} w_{k|k}^{(j)} \underbrace{\int \begin{array}{c} \mathcal{N}\left(\mathbf{x}_k\,;\,m_{k|k}^{(j)}, P_{k|k}^{(j)} \otimes X_{k+1}\right) \\ \times p_{k+1|k}^1\left(\mathbf{x}_{k+1}|X_{k+1}, \mathbf{x}_k\right) \end{array} \mathrm{d}\mathbf{x}_k}_{\text{Kinematical part}} \\
& \quad \times \underbrace{\int \mathcal{IW}\left(X_k\,;\,\nu_{k|k}^{(j)}, V_{k|k}^{(j)}\right) p_{k+1|k}^2\left(X_{k+1}|X_k\right) \mathrm{d}X_k}_{\text{Extension part}}.
\end{aligned}
\tag{19}
$$

Using the linear Gaussian model given in (2), the prediction for the kinematical part becomes [19]

$$
\begin{aligned}
& \int \mathcal{N}\left(\mathbf{x}_k\,;\,m_{k|k}^{(j)}, P_{k|k}^{(j)} \otimes X_{k+1}\right) \\
& \quad \times p_{k+1|k}^1\left(\mathbf{x}_{k+1}|X_{k+1}, \mathbf{x}_k\right) \mathrm{d}\mathbf{x}_k \\
& = \mathcal{N}\left(\mathbf{x}_{k+1}\,;\,m_{k+1|k}^{(j)}, P_{k+1|k}^{(j)} \otimes X_{k+1}\right),
\end{aligned}
\tag{20}
$$

where

$$
m_{k+1|k}^{(j)} = \left(F_{k+1|k} \otimes \mathbf{I}_d\right) m_{k|k}^{(j)},
\tag{21a}
$$

$$
P_{k+1|k}^{(j)} = F_{k+1|k} P_{k|k}^{(j)} F_{k+1|k}^{\mathrm{T}} + \mathbf{Q}_{k+1|k}.
\tag{21b}
$$

The extension part is less straightforward. Here, we apply the same heuristic approach as in [19], i.e. we make the approximation

$$
\begin{aligned}
& \int \mathcal{IW}\left(X_k\,;\,\nu_{k|k}^{(j)}, V_{k|k}^{(j)}\right) p_{k+1|k}^2\left(X_{k+1}|X_k\right) \mathrm{d}X_k \\
& \approx \mathcal{IW}\left(X_{k+1}\,;\,\nu_{k+1|k}^{(j)}, V_{k+1|k}^{(j)}\right)
\end{aligned}
\tag{22}
$$

where the predicted degrees of freedom and inverse scale matrix are approximated by

$$
\nu_{k+1|k}^{(j)} = e^{-T_s/\tau}\nu_{k|k}^{(j)},
\tag{23a}
$$

$$
V_{k+1|k}^{(j)} = \frac{\nu_{k+1|k}^{(j)} - d - 1}{\nu_{k|k}^{(j)} - d - 1}V_{k|k}^{(j)},
\tag{23b}
$$

where τ is a temporal decay constant. Thus, the PHD corresponding to predicted existing targets is

$$
\begin{aligned}
& \sum_{j=1}^{J_{k|k}} w_{k+1|k}^{(j)}\mathcal{N}\left(\mathbf{x}_{k+1}\,;\,m_{k+1|k}^{(j)}, P_{k+1|k}^{(j)} \otimes X_{k+1}\right) \\
& \quad \times \mathcal{IW}\left(X_{k+1}\,;\,\nu_{k+1|k}^{(j)}, V_{k+1|k}^{(j)}\right),
\end{aligned}
\tag{24}
$$

where $w_{k+1|k}^{(j)} = p_S w_{k|k}^{(j)}$, the Gaussian mean $m_{k+1|k}^{(j)}$ and covariance $P_{k+1|k}^{(j)}$ are given in (21), and the inverse Wishart degrees of freedom $\nu_{k+1|k}^{(j)}$ and inverse scale matrix $V_{k+1|k}^{(j)}$ are given in (23).

The birth PHD

$$
\begin{aligned}
D_k^b(\xi_k) = & \sum_{j=1}^{J_{b,k}} w_{b,k}^{(j)}\mathcal{N}\left(\mathbf{x}_k\,;\,m_{b,k}^{(j)}, P_{b,k}^{(j)} \otimes X_k\right) \\
& \times \mathcal{IW}\left(X_k\,;\,\nu_{b,k}^{(j)}, V_{b,k}^{(j)}\right),
\end{aligned}
\tag{25}
$$

represents new targets that appear at time step k. The full predicted PHD $D_{k+1|k}(\xi_{k+1})$ is the sum of the PHD of predicted existing targets (24) and the birth PHD (25), and contains a total of $J_{k+1|k} = J_{k|k} + J_{b,k+1}$ GIW components.

C. Correction

The corrected PHD is a GIW mixture given by

$$
D_{k|k}(\xi_k) = D_{k|k}^{\mathrm{ND}}(\xi_k) + \sum_{p \angle \mathbf{Z}_k} \sum_{W \in p} D_{k|k}^{\mathrm{D}}(\xi_k, W),
\tag{26}
$$

where $D_{k|k}^{\mathrm{ND}}(\cdot)$, handling the no detection cases, is given by

$$
\begin{aligned}
D_{k|k}^{\mathrm{ND}}(\xi_k) = & \sum_{j=1}^{J_{k|k-1}} w_{k|k}^{(j)}\mathcal{N}\left(\mathbf{x}_k\,;\,m_{k|k}^{(j)}, P_{k|k}^{(j)}\right) \\
& \times \mathcal{IW}\left(X_k\,;\,\nu_{k|k}^{(j)}, V_{k|k}^{(j)}\right),
\end{aligned}
\tag{27a}
$$

$$
w_{k|k}^{(j)} = \left(1 - \left(1 - e^{-\gamma^{(j)}}\right)p_D^{(j)}\right)w_{k|k-1}^{(j)},
\tag{27b}
$$

$$
\xi_{k|k}^{(j)} = \xi_{k|k-1}^{(j)}.
\tag{27c}
$$

The GIW mixture $D_{k|k}^{\mathrm{D}}(\xi_k, W)$, handling the detected target cases, requires the likelihood of the measurements in each cell

W,

$$\prod_{\mathbf{z}_k \in W} \frac{\phi_{\mathbf{z}_k}(\xi_k)}{\lambda_k c_k(\mathbf{z}_k)} = \beta_{FA,k}^{-|W|} \prod_{\mathbf{z}_k \in W} \mathcal{N}\left(\mathbf{z}_k^{(i)}; (H_k \otimes \mathbf{I}_d)\,\mathbf{x}_k, X_k\right),$$

$$(28)$$

multiplied with the predicted GIW components,

$$\mathcal{N}\left(\mathbf{x}_k; m_{k|k-1}^{(j)}, P_{k|k-1}^{(j)} \otimes X_k\right) \mathcal{IW}\left(X_k; \nu_{k|k-1}^{(j)}, V_{k|k-1}^{(j)}\right).$$

$$(29)$$

The product of (28) and (29) can be rewritten as

$$\beta_{FA,k}^{-|W|} \mathcal{L}_k^{(j,W)} \mathcal{N}\left(\mathbf{x}_k; m_{k|k}^{(j,W)}, P_{k|k}^{(j,W)} \otimes X_k\right)$$
$$\times \mathcal{IW}\left(X_k; \nu_{k|k}^{(j,W)}, V_{k|k}^{(j,W)}\right).$$

$$(30)$$

The details behind the derivation are given in Appendix A. The corrected Gaussian mean and covariance and inverse Wishart degrees of freedom and inverse scale matrix in (30) are given by

$$m_{k|k}^{(j,W)} = m_{k|k-1}^{(j)} + \left(K_{k|k-1}^{(j,W)} \otimes \mathbf{I}_d\right) \varepsilon_{k|k-1}^{(j,W)}, \tag{31a}$$

$$P_{k|k}^{(j,W)} = P_{k|k-1}^{(j)} - K_{k|k-1}^{(j,W)} S_{k|k-1}^{(j,W)} \left(K_{k|k-1}^{(j,W)}\right)^{\mathrm{T}}, \tag{31b}$$

$$\nu_{k|k}^{(j,W)} = \nu_{k|k-1}^{(j)} + |W|, \tag{31c}$$

$$V_{k|k}^{(j,W)} = V_{k|k-1}^{(j)} + N_{k|k-1}^{(j,W)} + Z_k^W, \tag{31d}$$

where the centroid measurement, scatter matrix, innovation factor, gain matrix, innovation vector and innovation matrix are defined as

$$\bar{\mathbf{z}}_k^W = \frac{1}{|W|} \sum_{\mathbf{z}_k^{(i)} \in W} \mathbf{z}_k^{(i)}, \tag{32a}$$

$$Z_k^W = \sum_{\mathbf{z}_k^{(i)} \in W} \left(\mathbf{z}_k^{(i)} - \bar{\mathbf{z}}_k^W\right)\left(\mathbf{z}_k^{(i)} - \bar{\mathbf{z}}_k^W\right)^{\mathrm{T}}, \tag{32b}$$

$$S_{k|k-1}^{(j,W)} = H_k P_{k|k-1}^{(j)} H_k^{\mathrm{T}} + \frac{1}{|W|}, \tag{32c}$$

$$K_{k|k-1}^{(j,W)} = P_{k|k-1}^{(j)} H_k^{\mathrm{T}} \left(S_{k|k-1}^{(j,W)}\right)^{-1}, \tag{32d}$$

$$\varepsilon_{k|k-1}^{(j,W)} = \bar{\mathbf{z}}_k^W - (H_k \otimes \mathbf{I}_d)\, m_{k|k-1}^{(j)}, \tag{32e}$$

$$N_{k|k-1}^{(j,W)} = \left(S_{k|k-1}^{(j,W)}\right)^{-1} \varepsilon_{k|k-1}^{(j,W)} \left(\varepsilon_{k|k-1}^{(j,W)}\right)^{\mathrm{T}}. \tag{32f}$$

The likelihood in (30) is given by

$$\mathcal{L}_k^{(j,W)} = \frac{1}{\left(\pi^{|W|}|W| S_{k|k-1}^{(j,W)}\right)^{\frac{d}{2}}} \frac{\left|V_{k|k-1}^{(j)}\right|^{\frac{\nu_{k|k-1}^{(j)}}{2}}}{\left|V_{k|k}^{(j,W)}\right|^{\frac{\nu_{k|k}^{(j,W)}}{2}}} \frac{\Gamma_d\left(\frac{\nu_{k|k}^{(j,W)}}{2}\right)}{\Gamma_d\left(\frac{\nu_{k|k-1}^{(j)}}{2}\right)}.$$

$$(33)$$

where $|V|$ denotes the determinant of the matrix V, and $|W|$ is the number of measurements in the cell W. The updated GIW component weight is given by

$$w_{k|k}^{(j,W)} = \frac{\omega_{\mathrm{p}}}{d_W} e^{-\gamma^{(j)}} \left(\frac{\gamma^{(j)}}{\beta_{FA,k}}\right)^{|W|} p_{\mathrm{D}}^{(j)} \mathcal{L}_k^{(j,W)} w_{k|k-1}^{(j)}, \quad (34)$$

where

$$d_W = \delta_{|W|,1} + \sum_{\ell=1}^{J_{k|k-1}} e^{-\gamma^{(\ell)}} \left(\frac{\gamma^{(\ell)}}{\beta_{FA,k}}\right)^{|W|} p_{\mathrm{D}}^{(\ell)} \mathcal{L}_k^{(\ell,W)} w_{k|k-1}^{(\ell)}.$$

$$(35)$$

Finally, the coefficients ω_{p} can be calculated by (9). The corrected PHD is of the form given in (12) with weights given by (34), and Gaussian and inverse Wishart parameters given in (31). Let $|\mathrm{p}_p|$ denote the number of cells W in the p:th partition, and let the set of partitions contain P unique partitions. The corrected PHD then has $J_{k|k} = J_{k|k-1} + J_{k|k-1} \sum_{p=1}^{P} |\mathrm{p}_p|$ GIW components.

D. Pruning and merging

From the prediction and correction, one quickly realizes that as time progresses, the number of GIW components increases rapidly. To keep the number of components at a tractable level, pruning and merging of GIW components is performed similarly to [13]. Empirically we have found that the merging threshold U must be chosen conservatively to avoid merging GIW components which correspond to multiple spatially close targets, because merging such components may cause cardinality error. The details of the implemented pruning and merging scheme are given below in Table I. Note that calculation of the merged covariance $\tilde{P}_{k|k}^{(\ell)}$ does not include the spread of means, the reason is that the means $m_{k|k}^{(i)}$ and covariances $P_{k|k}^{(i)}$ are of different dimensions ($s \times d$ and s, respectively). However, with a conservative merging threshold U, the spread of means is typically quite small and is thus negligible.

We also alert the reader about the very simple approach to merging of inverse-Wishart parameters in Table I. This procedure is sufficient when a conservative threshold is used, and the GIW-PHD filter is not very sensitive to changes in the merging algorithm. Nevertheless, finding a better method for GIW component merging is a potential subject for future research.

E. Implementation of the GIW-PHD filter

To facilitate implementation, we give pseudo code for the GIW-PHD filter, and address implementation issues and computational complexity, in a Technical Report [26, Online].

V. PARTITIONING THE MEASUREMENT SET

A central part of the correction equation given in (7), (8), is the partitioning of the set of measurements $\mathbf{Z}_k$ into partitions p containing non-empty cells W with measurements $\mathbf{z}_k^{(j)}$. For a given partition, the cells can be understood as containing measurements that are all from the same source, either a single target or a clutter source.

The measurement pseudo-likelihood (8) requires a summation over all possible partitions, which quickly becomes intractable because the number of possible partitions increases very rapidly as the size of $\mathbf{Z}_k$ increases [14], [17]. It has been noted that the full set of partitions can be approximated with

TABLE I
PSEUDO-CODE FOR GIW-PHD FILTER PRUNING AND MERGING

1: **input:** GIW components $\left\{ w_{k|k}^{(j)}, \xi_{k|k}^{(j)} \right\}_{j=1}^{J_{k|k}}$, a truncation threshold T, a merging threshold U and a maximum allowable number of GIW components $J_{\max}$.

2: **initialize:** Set $\ell \leftarrow 0$ and $I \leftarrow \left\{ i = 1, \ldots, J_{k|k} \middle| w_{k|k}^{(i)} > T \right\}$.

3: **repeat**

4:      $\ell \leftarrow \ell + 1$

5:      $j \leftarrow \arg\max\limits_{i \in I} w_{k|k}^{(i)}$

6:      Compute $\hat{P}_{k|k}^{(j)}$ using (14).

7:      $L \leftarrow \left\{ i \in I \middle| \left(m_{k|k}^{(i)} - m_{k|k}^{(j)} \right)^{\mathrm{T}} \left(\hat{P}_{k|k}^{(j)} \right)^{-1} \left(m_{k|k}^{(i)} - m_{k|k}^{(j)} \right) \leq U \right\}$

8:      $\tilde{w}_{k|k}^{(\ell)} \leftarrow \sum_{i \in L} w_{k|k}^{(i)}$

9:      $\tilde{m}_{k|k}^{(\ell)} \leftarrow \frac{1}{\tilde{w}_{k|k}^{(\ell)}} \sum_{i \in L} w_{k|k}^{(i)} m_{k|k}^{(i)}$,

10:     $\tilde{P}_{k|k}^{(\ell)} \leftarrow \frac{1}{\tilde{w}_{k|k}^{(\ell)}} \sum_{i \in L} w_{k|k}^{(i)} P_{k|k}^{(i)}$

11:     $\tilde{\nu}_{k|k}^{(\ell)} \leftarrow \frac{1}{\tilde{w}_{k|k}^{(\ell)}} \sum_{i \in L} w_{k|k}^{(i)} \nu_{k|k}^{(i)}$,

12:     $\tilde{V}_{k|k}^{(\ell)} \leftarrow \frac{1}{\tilde{w}_{k|k}^{(\ell)}} \sum_{i \in L} w_{k|k}^{(i)} V_{k|k}^{(i)}$

13:     $I \leftarrow I \backslash L$

14: **until** $I = \emptyset$

15: If $\ell > J_{\max}$ then replace $\left\{ \tilde{w}_{k|k}^{(j)}, \tilde{m}_{k|k}^{(j)}, \tilde{P}_{k|k}^{(j)}, \tilde{\nu}_{k|k}^{(j)}, \tilde{V}_{k|k}^{(j)} \right\}_{j=1}^{\ell}$ by those of the $J_{\max}$ GIW components with largest weights.

16: **output:** $\left\{ \tilde{w}_{k|k}^{(j)}, \tilde{\xi}_{k|k}^{(j)} \right\}_{j=1}^{\ell}$, $\tilde{\xi}_{k|k}^{(j)} = \left(\tilde{m}_{k|k}^{(j)}, \tilde{P}_{k|k}^{(j)}, \tilde{\nu}_{k|k}^{(j)}, \tilde{V}_{k|k}^{(j)} \right)$

a subset of partitions, so long as this subset contains the most likely ones among all of the possible partitions [17], [18]. A method called Distance Partition was suggested in [18], and it was augmented with the Sub-Partition algorithm in [17] to better handle the case of spatially close targets.

Distance Partition is based on the fundamental insight that measurements that are caused by the same extended target are spatially close to each other. Partitions are computed such that spatially close measurements are put into the same cell. However, a method based only on this places measurements from multiple targets in the same cell if two or more targets are spatially close, which may cause cardinality errors. In the Sub-Partition algorithm presented in [17], this problem was solved by generating additional partitions by considering the number of measurements in each cell $|W|$, and comparing it to the expected number of measurements from a single target. Given a maximum likelihood estimate K of the number of targets that caused the measurements in the cell, Sub-Partition uses K-means++ clustering to split the cell into K sub-cells. A new partition, that includes the sub-cells instead of the original cell, is added to the list of partitions. Though this method solves the cardinality issues in many practical cases, it is noted in [17] that it is only a first order solution to the problem.

Initial simulations with extended targets modeled using random matrices showed that Distance Partitioning with Sub-Partition was insufficient to handle some instances of multiple extended targets that are spatially close. The phenomenon is best explained with an example. Consider the two different sized and spatially close extended targets, with corresponding measurements, in Figure 2a. Distance Partition would place all measurements in the same cell, due to the spatial proximity of the measurements. Compare to the division of

the measurements using K-means++ in Figure 2b, which is the algorithm used in Sub-Partition. The result from K-means++ is typical, because for this type of scenario the K-means++ loss function profits much more by dividing the measurements by a vertical line, rather than a horizontal one. Because such a resulting additional partition will get a relatively lower likelihood, compared to the partition which assigns all the measurements to a single target (obtained initially by Distance Partitioning), the additional partition would not improve performance. Despite using Sub-Partition, the result would typically be a cardinality error in the filtering.

In order to be able to handle this type of true target scenario, in this paper two additional partitioning methods are suggested. The first is a method called Prediction Partition, which is based on the predicted GIW-PHD components. The second method, called EM Partition, is based on the expectation maximization (EM) algorithm [27]. Both methods are based on the intuition that in order to solve the problem for situations as in Figure 2, one has to incorporate the predicted kinematic and extent states of the targets into the partitioning process.

A. Prediction Partition

This partitioning method uses the predicted GIW components. For components with weight $w_{k+1|k}^{(j)} > 0.5$, a d-dimensional extension estimate $\hat{X}_{k+1|k}^{(j)}$ is computed as in (15). A corresponding position mean is obtained by taking the d first components of $m_{k+1|k}^{(j)}$, denoted $m_{k+1|k}^{(j),d}$. A partition is obtained by iterating over the components, in the order of decreasing weight, and putting all measurements $\mathbf{z}_k^{(i)}$ that fulfill

$$\left(\mathbf{z}_k^{(i)} - m_{k+1|k}^{(j),d} \right)^{\mathrm{T}} \left(\hat{X}_{k+1|k}^{(j)} \right)^{-1} \left(\mathbf{z}_k^{(i)} - m_{k+1|k}^{(j),d} \right) < \Delta_d(p) \tag{36}$$

into the same cell. Here, $\Delta_d(p)$ is computed using the inverse cumulative χ^2 distribution with d degrees of freedom, for probability $p = 0.99$. If a measurement falls into two or more extension estimates, it is only put into the cell corresponding to the component with highest weight. The measurements that do not fulfill (36) for any GIW component are placed in individual cells containing only one measurement.

This method works well when the true target motion can be well modeled by the dynamic motion model (2). However, when the targets maneuver the method is expected to be insufficient because the target predictions will be significantly erroneous.

B. EM *Partition*

The reason that K-means++ were successful in some scenarios in [17] was that the targets were mainly of the same size and circular (i.e. as opposed to elliptical). A typical extended target scenario can have targets of quite different sizes generating significantly different numbers of measurements, and the targets' measurement distributions can be significantly skewed, rather than circular. When there are targets of different sizes, the K-means++ algorithm, which does not use any measure of the clusters' physical sizes,

Fig. 2. Illustration of Sub-Partition. (a) Two spatially close extended targets, with corresponding measurements in black and gray. (b) Sub-cells resulting from K-means++, shown in black and gray. Ideally, the measurements should be split into two sub-cells along the $y = 15$ line.

often fails. The Expectation Maximization (EM) algorithm for Gaussian mixtures, which is a generalization of the K-means++ algorithm (see e.g. chapter 9 of [28]), incorporates both cluster sizes and number of measurements in each cluster via the covariances and the mixing coefficients. The specifics of the EM algorithm for Gaussian mixtures can be found in e.g. [28].

In the EM Partition algorithm, the Gaussian mixture parameters are initialized with means $\mu_\ell = m_{k+1|k}^{(j),d}$, covariances $\Sigma_\ell = \hat{X}_{k+1|k}^{(j)}$ and mixing coefficients $\pi_\ell \propto \gamma\left(\xi_{k+1|k}^{(j)}\right)$ for components j with weight $w_{k+1|k}^{(j)} > 0.5$. An additional mixture component is added with mean μ_ℓ at the center of the surveillance area, circular covariance Σ_ℓ scaled such that the corresponding 99% probability volume approximately covers the surveillance area, and mixing coefficient $\pi_\ell = 10^{-9}$. The mixing coefficients π_ℓ are normalized to satisfy $\sum_\ell \pi_\ell = 1$ before the first E-step.

The additional mixture component is added to capture the clutter measurements, such that the mixture components corresponding to the target estimates can converge approximately to the true partitioning. Note that for a given set of initial Gaussian mixture components, the EM algorithm will converge to the closest local maximum of the likelihood function, i.e. there is no guarantee that EM converges to the global maximum. Because EM Partition is initialized using the predicted GIW components, similarly to Prediction Partition it is sensitive to maneuvers that are modeled poorly by the motion model. However, because of the adaptation capability of the EM-iterations, EM Partition is slightly less sensitive than Prediction Partition.

C. Discussion

It is important to note that each of the three partitioning methods[2] used in this work have its respective failure modes. A problem with Distance Partition with Sub-Partition was highlighted in Figure 2. Prediction Partition relies on the prediction of the GIW components, this method sometimes returns a non-informative partition when targets are maneuvering. EM Partition can converge to a local maximum of the likelihood function that yields a non-informative partition. For this reason, it is a better choice to use all three methods, rather than just one method on its own. The more partitions that are used, the better the full set of partitions is approximated. Indeed, it is possible that adding further partitioning methods

[2]Distance Partition with Sub-Partition [17], Prediction Partition and EM Partition

would improve performance, however this should also be balanced against the fact that considering more partitions requires more computations.

VI. SIMULATION RESULTS

This section presents results from extended target tracking simulations. The target tracking setup is presented in the next section, followed by the results from four different extended target tracking scenarios.

A. Target tracking setup

Four different scenarios were simulated, each with two targets. The true tracks are shown in Figure 3. The true target extensions are given by

$$X_k^{(i)} = \mathsf{R}_k^{(i)} \mathrm{diag}\left(\left[A_i^2 \; a_i^2\right]\right)\left(\mathsf{R}_k^{(i)}\right)^{\mathrm{T}}, \qquad (37)$$

where $\mathsf{R}_k^{(i)}$ is a rotation matrix applied such that the i:th extension's major axis is aligned with the i:th target's direction of motion at time step k, and A_i and a_i are the length of the major and minor axes, respectively. In all four scenarios, the major and minor axes are $(A_1, a_1) = (20, 5)$ and $(A_2, a_2) = (10, 2.5)$ for the two targets, respectively.

The expected number of measurements generated by the targets is assumed to be a function of the extended target volume $\mathcal{V}_k^{(i)} \triangleq \pi\sqrt{\left|X_k^{(i)}\right|} = \pi A_i a_i$. This assumption is reasonable in many real world scenarios, where a smaller target would occupy fewer of the sensor's resolution cells than a larger target, thus yielding fewer measurements. Here we adopt the following simple model for the expected number of measurements that the targets generate,

$$\gamma_k^{(i)} = \left\lfloor \sqrt{\frac{4}{\pi}\mathcal{V}_k^{(i)}} + 0.5 \right\rfloor = \left\lfloor 2\sqrt{A_i a_i} + 0.5 \right\rfloor, \qquad (38)$$

where $\lfloor \cdot \rfloor$ is the floor function and $\lfloor x + 0.5 \rfloor$ rounds x to the nearest integer. This model is equivalent to assuming a uniform expected number of measurements per square root of surveillance area. In a typical real world scenario, the number of target measurements may also depend on the distance between the target and the sensor, i.e. depend on the kinematical target state $\mathbf{x}_k^{(i)}$. This case can easily be handled with a modified expected number of measurements model. For the sake of simplicity, this case is not included in this paper, and the readers are referred to [17] for such an example.

The motion model parameters are set to $T_s = 1$s, $\theta = 1$s, $\Sigma = 0.1\mathrm{m/s}^2$ and $\tau = 5$s. In three of four scenarios, the

parameters of the $J_{b,k} = 2$ birth PHD components are set as follows, $w_{b,k}^{(j)} = 0.1$, and

$$m_{b,k}^{(j)} = \left[\left(\mathbf{x}_0^{(j)}\right)^{\mathrm{T}} \mathbf{0}_4^{\mathrm{T}}\right]^{\mathrm{T}}, \quad P_{b,k}^{(j)} = \mathrm{diag}\left(\left[100^2\ 25^2\ 25^2\right]\right),$$
$$\nu_{b,k}^{(j)} = 7, \qquad\qquad V_{b,k}^{(j)} = \mathrm{diag}\left(\left[1\ 1\right]\right). \qquad (39)$$

The mean vectors $m_{b,k}^{(j)}$ are set such that they correspond to the starting points of the true targets. In the fourth scenario, there is $J_{b,k} = 1$ birth component, with mean vector set to the mean of the two targets' starting points. Knowing the starting points of the targets *a priori* is naturally not possible in many real world scenarios. In the experiment section, we elaborate further on how the birth PHD can be constructed in a real scenario.

A total of 100 Monte Carlo simulations were preformed for each scenario, with a clutter rate of 10 clutter measurements per time step. The results are presented in terms of the multi-target measure optimal subpattern assignment metric (OSPA) [29], cardinality and length of the estimated major and minor axes of the extension matrices.

B. Crossing tracks

In this scenario, the target tracks cross at close distance, see Figure 3a. The results are shown in Figure 4. The plots clearly show that straight line motion can be readily handled by the presented filter, even when the targets are spatially close. Noteworthy are the estimates of the major and minor axes, A_i and a_i, respectively. The results show that extensions which do not change over time (e.g. do not grow, shrink or rotate) can be estimated with low error.

C. Parallel tracks

In the second scenario, the two targets move closer and then move in parallel, before separating again, see Figure 3b. While moving in parallel, the true target extensions' three standard deviation ellipses (corresponding to the 99% probability volume) are separated by 2.5m. The results are shown in Figure 5. The mean sum of weights is close to the true value, however there is a downward trend while the targets are moving in parallel. This is caused mainly by missed detections, which often causes the PHD filter to lose the target estimate corresponding to the target that was not detected. In the subsequent time steps, when the target is detected again, the measurements from both targets are typically treated as being caused by one target.

This can also be seen in the OSPA value, which increases during parallel motion, and also has larger standard deviation. The estimates of A_i and a_i are slightly worse than in the previous scenario. In the beginning and end of the simulation this is caused by the rotation of the extension matrices X_k. This result is intuitive – the turning makes the extension more difficult to track since the prediction (23) does not account for the rotation of the extension.

D. Separating tracks

In the third scenario, the two targets start such that their respective three standard deviation ellipsoids, computed from the true extensions X_k, are touching. First the targets move in parallel, after about half the scenario they separate, see Figure 3c. For this scenario, a birth PHD with $J_{b,k} = 1$ component was used. The results are shown in Figure 6. For the first half of this scenario, when the target extents are touching, the filter incorrectly estimates one large extended target, instead of two smaller ones, in about 60% of the Monte Carlo simulations. This is what causes the mean sum of weights to be around 1.4. The targets start to separate at time 52, and from time 57 the cardinality is estimated correctly. Because the cardinality is underestimated in about 60% of the Monte Carlo simulations, the estimated major and minor axes, A_i and a_i, are difficult to interpret for the first half of the scenario. When the GIW-PHD filter estimates only one target, the major axis of the extracted target is estimated to be slightly lower than $A_1 = 20\mathrm{m}$. We had expected the major axis to be estimated as 20m or more when the targets are combined. However, this "underestimation" appears to be a property of the particular prediction and correction equations used for the inverse Wishart parameters. The major axes of the targets are, in a way, averaged to obtain a smaller estimate than 20m. For the second half, when the targets are separated, the results are better.

Furthermore, the results show that there is little need for a specific model for target spawning. As soon as the two targets are slightly separated, the partitioning algorithm (Distance Partition) automatically starts to generate a partition that suits the spawning event. This partition then dominates the other partitions, which can be seen e.g. in terms of the partition weights ω_p. This process is evident also from the cardinality estimates, which are corrected shortly after the targets separate.

E. Closely spaced targets

In the fourth scenario, the targets move closer and then move in parallel, both in straight lines and through a curve, before separating again, see Figure 3d. Estimating cardinality correctly becomes increasingly difficult as multiple targets move close to each other. Early tests showed that Distance Partition with Sub-Partition was insufficient to handle some cases of spatially close extended targets modeled as random matrices. To improve performance, Prediction Partition and EM Partition was implemented.

To test the PHD-filters capability of tracking multiple closely spaced targets, the scenarios in Figure 3b and Figure 3d were simulated when the targets' extents were separated by a distance d. The tracks in Figure 3b were simulated for separating distances $d = 0, 0.5, 1, \ldots, 5$ [m], and the mean sum of weights is shown in Figure 7a. When rounded to the nearest integer there is no cardinality error at any distance d, however estimating cardinality correctly becomes increasingly difficult at closer distances, which is shown by the lower mean value for $d < 2.5\mathrm{m}$. Without Prediction Partition
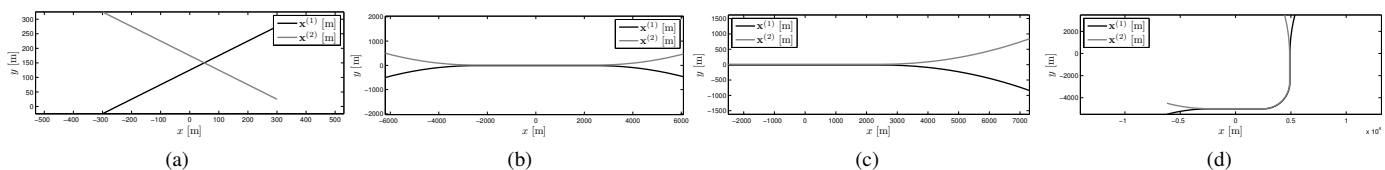
Fig. 3.   True target tracks used in simulations. (a) Crossing tracks. (b) Parallel tracks. (c) Separating tracks. (d) Turning tracks.
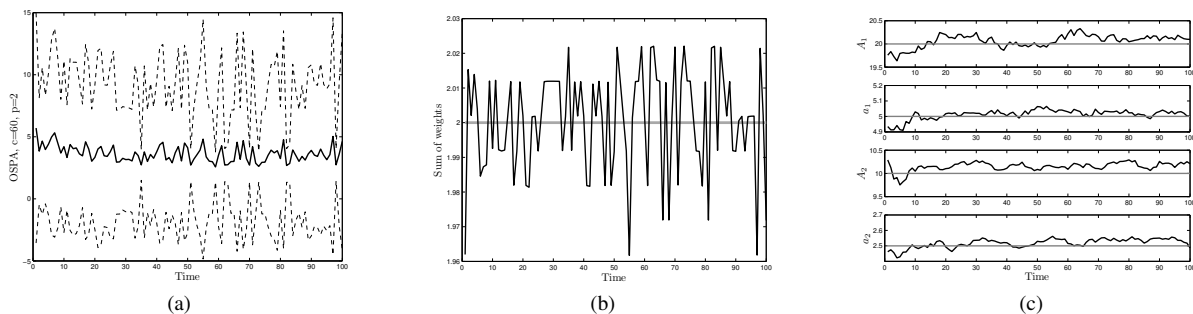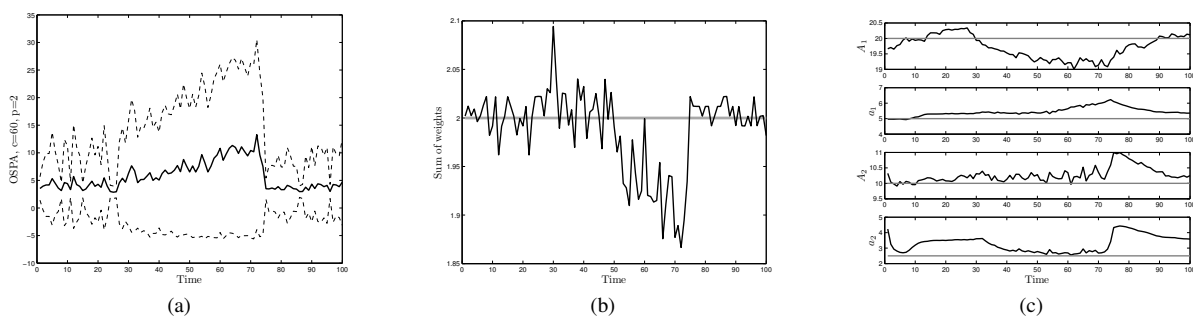

Fig. 4.   (a) Mean OSPA (solid line) $\pm$ one standard deviation (dashed lines). (b) Cardinality estimate, taken as the sum of weights (black), compared to true cardinality (gray). (c) Estimates of the major and minor axes, A_i and a_i, respectively. Mean estimates (black) compared to true value (gray).
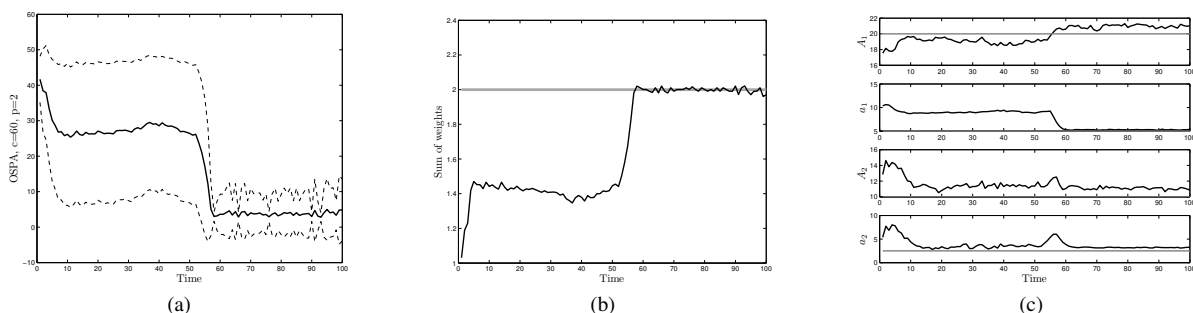

Fig. 5.   (a) Mean OSPA (solid line) $\pm$ one standard deviation (dashed lines). (b) Cardinality estimate, taken as the sum of weights (black), compared to true cardinality (gray). (c) Estimates of the major and minor axes, A_i and a_i, respectively. Mean estimates (black) compared to true value (gray).

Fig. 6.   (a) Mean OSPA (solid line) $\pm$ one standard deviation (dashed lines). (b) Cardinality estimate, taken as the sum of weights (black), compared to true cardinality (gray). (c) Estimates of the major and minor axes, A_i and a_i, respectively. Mean estimates (black) compared to true value (gray).

and EM Partition, simulations show that the cardinality is underestimated for distances $d < 7$m.

The tracks in Figure 3d contain a turn, making prediction of the extended target estimates harder, because the dynamic motion model is constant velocity and predicts target motion in a straight line. This scenario was simulated at two different speeds, 125m/s and 62.5m/s. The separating distances were $d = 0, 0.5, 1, \ldots, 25$ [m] for the faster speed and $d = 0, 0.5, 1, \ldots, 5$ [m] for the slower speed.

At the higher speed, target prediction is more difficult, especially during the turn, and subsequently Prediction Partition and EM Partition fails to compute informative partitions more often. This is a common cause of cardinality error. However, at the lower speed, the true target motion per time step is smaller, and the linear constant velocity prediction is good enough for Prediction Partition and EM Partition to compute informative partitions. The mean sum of weights at both speeds is shown in Figure 7b and Figure 7c, respectively. The filter can handle

the maneuver, i.e. there is no cardinality error when the mean sum of weights is rounded to the nearest integer, when the targets are separated by $d \geq 21$m at the higher speed. At the lower speed only $d \geq 2$m separation is needed.

To conclude, the results show that when a constant velocity motion model is used, the presented extended target tracking filter can handle all scenarios except the ones where multiple spatially close targets are maneuvering quickly. In scenarios where the target maneuvers are dominant, the use of interacting multiple models (IMM) [30] for motion prediction seems to be a reasonable solution, e.g. this was done in Section V of [21]. The presented tracking filter can easily be generalized to use an IMM filter, however this was considered to be beyond the scope of this paper.
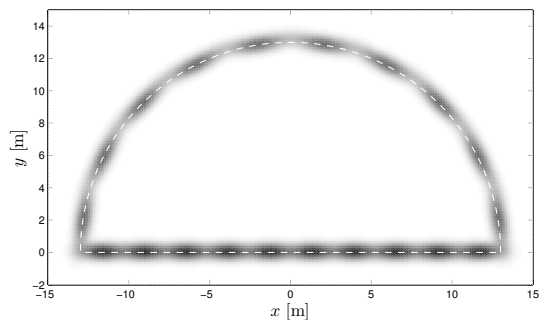

Fig. 8. Birth PHD used in experiments. The dark areas are locations which the birth PHD models as likely locations for new targets to appear. The edge of the surveillance area is shown as a dashed white line.

VII. Experiment results

This section presents results from experiments based on data from a laser range sensor. Measurements were collected using a SICK LMS laser range sensor, which measures range every $0.5°$ over a $180°$ surveillance area. Ranges shorter than 13m were converted to (x, y) measurements using a polar to Cartesian transformation. The two data sets contain 411 and 400 laser range sweeps, respectively. Human targets entered the surveillance area at different times, and were measured by the sensor at waist level. There is no ground truth available for the data, however by examining the measurements the true cardinality can be observed.

These two data sets have previously been used in [17], where only the kinematical part of the target state is tracked. A comparison of the results for the presented GIW-PHD filter to those for the ET-GM-PHD filter from [17] is performed.

A. Target tracking setup

The sensor's sampling time is $T_s = 0.2$s. The motion model parameters are set to $\theta = 1$s, $\Sigma = 2$m/s^2 and $\tau = 5$s. For the sensor used, new targets will appear somewhere along the edge of the semi circular surveillance area. Therefore, the birth PHD has $J_{b,k} = 20$ components located along the edge of the surveillance area, the intensity $D_k^b(\cdot)$ in the (x, y) dimension is shown in Figure 8. The birth components' weights are set to $w_{b,k}^{(j)} = 0.1/J_{b,k}$ and the inverse Wishart parameters are set to $\nu_{b,k}^{(j)} = 7$ and $V_{b,k}^{(j)} = \text{diag}\left([0.25^2 \ 0.1^2]\right)$.

For the sensor used here, the expected number of target generated measurements γ varies rapidly with the distance between the target and the sensor. We have found that the correction weight update (34) is not sensitive to setting the corresponding filter parameter constant, however Sub-Partition needs a reasonable estimate of γ in order to compute a maximum likelihood estimate K of the number of targets that generated the measurements in a cell W. To facilitate this, in the Sub-Partition algorithm we have estimated γ by assuming that a 50cm wide target is located at the particular cell's centroid $\bar{z}_k^W$. A width of 50cm roughly corresponds to the size of an average person, who is facing the sensor. This simple heuristic works well for the particular experiments presented here, however it remains within future work to design a method which does not rely on *a priori* information

of the tracking scenario. A study of the extended target PHD filter's performance for incorrect values of the filter parameter corresponding to γ is given in [17].

B. Experiment with close targets

This data set contains 411 laser range scans. Two humans walked through the surveillance area, repeatedly moving towards and away from each other, both in the same direction and in the opposite direction. Thus, the data set contains situations where the targets are spatially close for both longer and shorter periods of time. The positions of the extracted targets are shown in Figure 9a, the number of extracted targets are compared to the ground truth in Figure 9b and the sum of weights is shown in Figure 9c. There is no cardinality error for the entire length of the experiment, however at time 164 there is an unexpected increase in the sum of weights to 2.4. The sum of weights increases because the target generated measurements for one of the targets, at that time step, resemble two small clusters rather than one larger cluster. The GIW-PHD filter interprets this as an increased likelihood of an additional target being present. These results are a small improvement over the results in [17], where the ET-GM-PHD filter underestimates the cardinality for three consecutive time steps when the targets are close and moving in the same direction.

C. Experiment with occlusion

This data set contains 400 laser range scans. Four humans walked through the surveillance area, however at most three humans were present at any one time. The first target stands still at the position $(x, y) \approx (0.4, 6)$ for most of the experiment. The second target walks behind the first target, causing the second target to be fully occluded (i.e. the second target is not measured), and also walks in front of the first target, causing the first target to be partially occluded (i.e. only parts of the first target are measured). With a constant probability of detection, the occlusion would cause target loss. To handle the occlusion without target loss, a state dependent probability of detection is implemented. The variable probability is based on a simple understanding of the sensor – objects that are located behind other objects cannot be measured by the sensor, therefore the probability of detection behind a target should
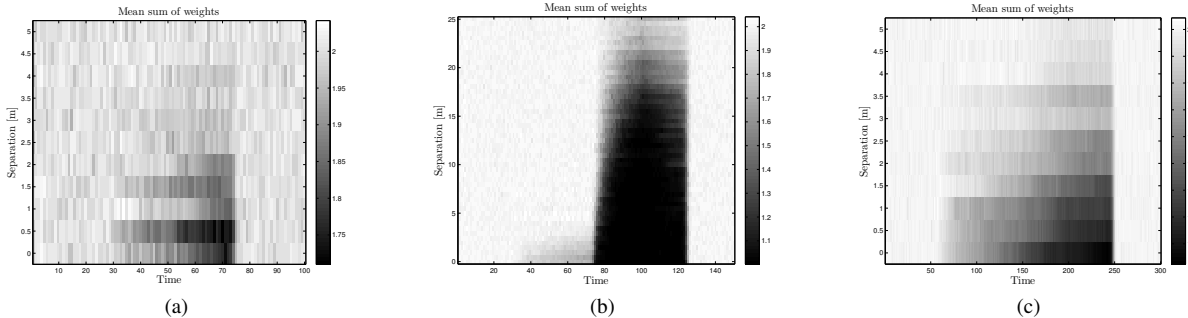
Fig. 7. Mean sum of weights for closely spaced targets, at various separating distance. The true cardinality is 2 for all three plots. (a) Tracks in Figure 3b, note that if the mean sum of weights is rounded to the nearest integer there is no cardinality error. (b) Tracks in Figure 3d at speed 125m/s. At distances $d \geq 21$m the cardinality is estimated correctly. (c) Tracks in Figure 3d at speed 62.5m/s. At distances $d \geq 2$m the cardinality is estimated correctly.
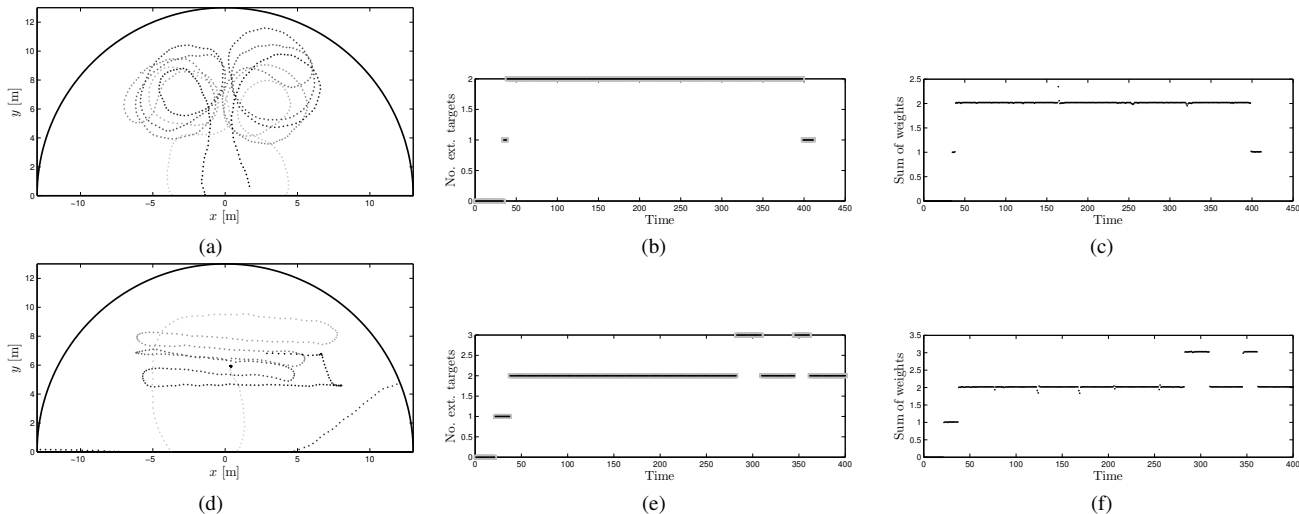

Fig. 9. Results from experiments, the top row shows results for the experiment with close targets (Section VII-B), the bottom row shows results for the experiment with occlusion (Section VII-C). (a) and (d) show the positions of the extracted targets' kinematical states. Light gray points corresponds to earlier time steps, dark gray corresponds to later time steps. (b) and (e) show the number of extracted targets in black, compared to the true cardinality in gray. (c) and (f) show the sum of weights over time for the two experiments.

be (close to) zero. The details of the variable probability of detection are given in Appendix B.

The positions of the extracted targets are shown in Figure 9d, the number of extracted targets are compared to the ground truth in Figure 9e and the sum of weights is shown in Figure 9f. At time 345, the time step when the fourth target enters the surveillance area, the cardinality is underestimated by 1. At this time step, the fourth target only generates one measurement, which the GIW-PHD filter interprets as clutter.

These results are a considerable improvement over the results in [17], where the ET-GM-PHD filter underestimated the cardinality in two situations where two targets are spatially close, such that one target is partially occluded. In [17], a variable probability of detection was also used. However with the target centroid occluded, the probability of detection of the partially occluded target is incorrectly set close to zero, causing cardinality error. With an estimate of the target extension, the partially occluded target can still be found to be detectable using the variable probability of detection in Appendix B. Thus, the experiment shows that for the data used, the GIW-PHD filter can handle occlusion and spatially close targets simultaneously. The experiment also shows the

benefit of estimating both the kinematical and extent states, compared to only estimating the kinematical state.

D. Discussion

The two experiments above are not an exhaustive evaluation of the GIW-PHD filter, but they serve as a proof of concept and a potential application (e.g. person tracking for mobile robots). The comparison to the results in [17], in which the target extent is not explicitly estimated, support the intuitive hypothesis that estimating the size of the extended targets improves the tracking performance. Initial steps have been taken toward including extension parameters in the target state in the ET-GM-PHD filter [7]. Thus, more experiments that compare the ET-GM-PHD filter and the GIW-PHD filter are needed, e.g. for data that contains more clutter than typical laser data does.

VIII. CONCLUSIONS AND FUTURE WORK

This paper presented a PHD filter for multiple extended target tracking, in the presence of clutter and missed detection. The target extensions are modeled as random matrices [19], and a suitable likelihood function is derived. The PHD is

approximated using Gaussian inverse Wishart distributions, and the assumptions necessary to obtain a computationally tractable PHD filter are presented. Two methods for measurement set partitioning are suggested to be added to the methods presented in [17]. The first method is based on the predicted Gaussian inverse Wishart PHD components and the second method is based on the well known EM algorithm. Adding the two partitioning methods improves tracking of multiple targets of different sizes when they are spatially close. A simulation study confirms that the presented PHD filter can handle spatially close targets, with the exception of when the targets maneuver quickly. It is further shown that target spawning can be handled without the use of a specific spawning model. The spawning is instead implicitly handled by the measurement partitioning. A potential application, person tracking using laser range sensors, is presented. Two experiments show the benefit of estimating the size of the target extents, compared to only tracking the kinematical states of the target centroid, as is performed in [17].

The presented target tracking filter estimates the target extent as a random matrix, giving an elliptical extended target shape. Alternatively, the target ellipse could be explicitly parametrized, included in the state vector, and estimated with the kinematical states. Such an example is given in [7], where elliptical shapes are tracked using a laser range sensor and a ET-GM-PHD filter. Ellipse tracking is also performed using random hypersurface models in [5], a comparison between random matrices and random hypersurface models for the single target case is given in [23]. A comparison of the presented GIW-PHD filter, and elliptic random hypersurface models using the ET-GM-PHD filter, could be interesting for the multiple target case.

The importance of measurement set partitioning was highlighted in the paper, and the case of close targets maneuvering quickly was shown to be difficult to handle. This could possibly be improved through the use of additional partitioning methods, or via modeling different type of motions using an IMM type filter. A prediction model that allows transformations of the extension, e.g. rotations, would possibly improve the filter performance.

Target spawning is not explicitly modeled in this work, however, it could be handled implicitly via the partitioning methods used. The targets must be separated sufficiently for the spawning to be detected, and the spawning event is therefore detected with a small time delay. It is not obvious how a Gaussian inverse Wishart distribution can be split into two Gaussian inverse Wishart distributions, however devising such a method could possibly improve performance for target spawning events.

A heuristic for determining the parameter γ in the Sub-Partition algorithm was suggested. A method which does not rely on either assumptions or *a priori* knowledge of the tracking scenario would be useful in the general case. Finally, an improved method for merging of GIW components is needed.

APPENDIX A
DERIVATION OF THE CORRECTION

Under the measurement model (5), the likelihood of n measurements $\mathbf{z}_j$ is

$$\prod_{j=1}^{n} \mathcal{N}\left(\mathbf{z}_j \,;\, (H \otimes \mathbf{I}_d)\,\mathbf{x}, X\right) = \prod_{j=1}^{n} \mathcal{N}\left(\mathbf{z}_j \,;\, \tilde{H}\mathbf{x}, X\right)$$
$$= (2\pi)^{-nd/2} |X|^{-n/2}$$
$$\times \text{etr}\left(-\frac{1}{2}\left(\sum_{j=1}^{n}\left(\mathbf{z}_j - \tilde{H}\mathbf{x}\right)\left(\mathbf{z}_j - \tilde{H}\mathbf{x}\right)^{\mathrm{T}}\right)X^{-1}\right),$$
(40)

where $\text{etr}\left(\cdot\right) = \exp\left(\text{Tr}\left(\cdot\right)\right)$ is exponential trace. Define the centroid measurement as $\bar{\mathbf{z}} \triangleq \frac{1}{n}\sum_{j=1}^{n}\mathbf{z}_j$ and the scatter matrix as $Z \triangleq \sum_{j=1}^{n}\left(\mathbf{z}_j - \bar{\mathbf{z}}\right)\left(\mathbf{z}_j - \bar{\mathbf{z}}\right)^{\mathrm{T}}$, and rewrite the summation as

$$\sum_{j=1}^{n}\left(\mathbf{z}_j - \tilde{H}\mathbf{x}\right)\left(\mathbf{z}_j - \tilde{H}\mathbf{x}\right)^{\mathrm{T}} = Z + n\left(\bar{\mathbf{z}} - \tilde{H}\mathbf{x}\right)\left(\bar{\mathbf{z}} - \tilde{H}\mathbf{x}\right)^{\mathrm{T}}.$$
(41)

Inserting (41) into (40) gives

$$\prod_{j=1}^{n} \mathcal{N}\left(\mathbf{z}_j \,;\, \tilde{H}\mathbf{x}, X\right) \tag{42a}$$

$$= (2\pi)^{-nd/2} |X|^{-n/2}\text{etr}\left(-\frac{1}{2}ZX^{-1}\right)$$
$$\times \text{etr}\left(-\frac{1}{2}\left(\bar{\mathbf{z}} - \tilde{H}\mathbf{x}\right)\left(\bar{\mathbf{z}} - \tilde{H}\mathbf{x}\right)^{\mathrm{T}}\left(\frac{X}{n}\right)^{-1}\right) \tag{42b}$$

$$= (2\pi)^{-(n-1)d/2} |X|^{-(n-1)/2} n^{-d/2}$$
$$\times \text{etr}\left(-\frac{1}{2}ZX^{-1}\right)\mathcal{N}\left(\bar{\mathbf{z}} \,;\, \tilde{H}\mathbf{x}, \frac{X}{n}\right) \tag{42c}$$

$$= \mathcal{L}_{\text{aux}}\mathcal{N}\left(\bar{\mathbf{z}} \,;\, \tilde{H}\mathbf{x}, \frac{X}{n}\right). \tag{42d}$$

Now, let the predicted target distribution be

$$\mathcal{N}\left(\mathbf{x} \,;\, m, P \otimes X\right)\mathcal{IW}\left(X \,;\, \nu, V\right). \tag{43}$$

The product of the measurement likelihood (42) and the predicted distribution (43) is

$$\mathcal{L}_{\text{aux}}\mathcal{N}\left(\bar{\mathbf{z}} \,;\, \tilde{H}\mathbf{x}, \frac{X}{n}\right)\mathcal{N}\left(\mathbf{x} \,;\, m, P \otimes X\right)\mathcal{IW}\left(X \,;\, \nu, V\right)$$
$$= \mathcal{N}\left(\mathbf{x} \,;\, m_+, P_+ \otimes X\right)\mathcal{N}\left(\bar{\mathbf{z}} \,;\, \tilde{H}m, SX\right)\mathcal{IW}\left(X \,;\, \nu, V\right)\mathcal{L}_{\text{aux}} \tag{44}$$

where we have

$$m_+ = m + (K \otimes \mathbf{I}_d)\left(\bar{\mathbf{z}} - \tilde{H}m\right), \tag{45a}$$
$$P_+ = P - KSK^{\mathrm{T}}. \tag{45b}$$

with innovation factor $S = HPH^{\mathrm{T}} + 1/n$ and gain matrix $K = PH^{\mathrm{T}}S^{-1}$.

This result is easy to derive using the product formula for Gaussians, and using some basic properties of the Kronecker

product. We thus have the corrected Gaussian distribution with mean and covariance (45), multiplied with

$$\mathcal{N}\left(\bar{\mathbf{z}}\,;\,\tilde{H}m,SX\right)\mathcal{IW}\left(X\,;\,\nu,V\right)\mathcal{L}_{\mathrm{aux}} \tag{46a}$$

$$=(2\pi)^{-d/2}\,|SX|^{-1/2}\frac{|V|^{\nu/2}\,|X|^{-(\nu+d+1)/2}}{2^{\nu d/2}\Gamma_d\left(\nu/2\right)}$$
$$\times\mathrm{etr}\left(-\frac{1}{2}\left(\bar{\mathbf{z}}-\tilde{H}m\right)\left(\bar{\mathbf{z}}-\tilde{H}m\right)^{\mathrm{T}}(SX)^{-1}\right)$$
$$\times\mathrm{etr}\left(-\frac{1}{2}VX^{-1}\right)\mathrm{etr}\left(-\frac{1}{2}ZX^{-1}\right)$$
$$\times(2\pi)^{-(n-1)d/2}\,|X|^{-(n-1)/2}n^{-d/2} \tag{46b}$$

$$=(2\pi)^{-nd/2}\,(nS)^{-d/2}\,\frac{|V|^{\nu/2}}{2^{\nu d/2}\Gamma_d\left(\nu/2\right)}$$
$$\times\frac{2^{(\nu+n)d/2}\Gamma_d\left((\nu+n)/2\right)}{|V+N+Z|^{(\nu+n)/2}}$$
$$\times\frac{|V+N+Z|^{(\nu+n)/2}\,|X|^{-(\nu+n+d+1)/2}}{2^{(\nu+n)d/2}\Gamma_d\left((\nu+n)/2\right)}$$
$$\times\mathrm{etr}\left(-\frac{1}{2}\left(V+N+Z\right)X^{-1}\right) \tag{46c}$$

$$=(\pi^n nS)^{-d/2}\,\frac{|V|^{\nu/2}}{|V+N+Z|^{(\nu+n)/2}}\frac{\Gamma_d\left((\nu+n)/2\right)}{\Gamma_d\left(\nu/2\right)}$$
$$\times\mathcal{IW}\left(X\,;\,\nu_+,V_+\right) \tag{46d}$$
$$=\mathcal{L}\times\mathcal{IW}\left(X\,;\,\nu_+,V_+\right) \tag{46e}$$

where we have $N=S^{-1}\left(\bar{\mathbf{z}}-\tilde{H}m\right)\left(\bar{\mathbf{z}}-\tilde{H}m\right)^{\mathrm{T}}$ and

$$\nu_+=\nu+n, \tag{47a}$$
$$V_+=V+N+Z, \tag{47b}$$

and the likelihood function $\mathcal{L}$ is defined as

$$\mathcal{L}=(\pi^n nS)^{-d/2}\,\frac{|V|^{\nu/2}}{|V_+|^{\nu_+/2}}\frac{\Gamma_d\left(\nu_+/2\right)}{\Gamma_d\left(\nu/2\right)}. \tag{48}$$

The likelihood function can be shown to be proportional to a generalized matrix variate beta type II distribution [31]. We have thus shown how the likelihood of n measurements $\mathbf{z}_j$ multiplied with a predicted GIW distribution can be rewritten as a corrected GIW distribution multiplied with a likelihood function.

APPENDIX B
VARIABLE PROBABILITY OF DETECTION FOR THE LASER RANGE SENSOR

The variable probability of detection used here is similar to the one presented in [17], however it relies on less assumptions, and instead utilizes the estimated target extensions. The idea is to decrease the probability of detection behind (i.e. at larger range from the sensor) each GIW component. In doing so, the function considers the component weight, the size of the estimated extension and the uncertainty in bearing (i.e. the polar angle from the sensor to the component).

For a given point (x,y) in the surveillance area, the probability of detection is computed as

$$p_{\mathrm{D}}\left(\mathsf{x},\mathsf{y}\right)=\max\left(p_{\mathrm{D,min}}\,,\,p_{\mathrm{D,0}}-\tilde{p}_{\mathrm{D}}\right), \tag{49a}$$
$$\tilde{p}_{\mathrm{D}}=\sum_{i:r>r^{(i)}}w^{(i)}\max\left(\mathcal{G}_1,\mathcal{G}_2,\mathcal{A}\right), \tag{49b}$$
$$\mathcal{G}_g=\exp\left(-\frac{\left(\varphi-\varphi^{(i)}+(-1)^g 2\sigma_{\varphi,e}\right)^2}{0.01\sigma_{\varphi,p}}\right), \tag{49c}$$
$$\mathcal{A}=\left|\varphi-\varphi^{(i)}\right|<2\sigma_{\varphi,e}, \tag{49d}$$

where

- $p_{\mathrm{D,min}}$ is the minimum probability of detection value allowed;
- $p_{\mathrm{D,0}}$ is the nominal probability of detection of targets which are not occluded;
- $r=\sqrt{\mathsf{x}^2+\mathsf{y}^2}$ and $\varphi=\tan^{-1}\left(\mathsf{y}/\mathsf{x}\right)$ is the range and bearing to the point (x,y);
- $w^{(i)}$, $r^{(i)}$ and $\varphi^{(i)}$ is the weight, range and bearing to the i:th GIW component's kinematical state;
- $\sigma_{\varphi,e}$ is the cross range size of the i:th component, computed by Cartesian to polar conversion of the extent matrix estimated as in (15);
- $\sigma_{\varphi,p}$ is the bearing standard deviation of the i:th components kinematical state, computed by Cartesian to polar conversion of the position uncertainty estimated as in (14).

To obtain the probability of detection for a GIW component $\xi_{k|k-1}^{(i)}$, the ellipsoid corresponding to two standard deviations of the estimated extent (15) is discretized into points (x,y), and for each discrete point along the extent a probability of detection is computed. The probability of detection for the GIW component $p_{\mathrm{D}}\left(\xi_{k|k-1}^{(i)}\right)$ is then given as the maximum of the probabilities computed along the discretized extent. In comparison, the variable probability of detection is computed only for the kinematical state position in [17]. Taking the maximum probability along the extent is what enables the GIW-PHD filter to handle partial target occlusion.

REFERENCES

[1] Y. Bar-Shalom and T. E. Fortmann, *Tracking and data association*, ser. Mathematics in Science and Engineering. San Diego, CA, USA: Academic Press Professional, Inc., 1987, vol. 179.

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

[3] K. Gilholm, S. Godsill, S. Maskell, and D. Salmond, "Poisson models for extended target and group tracking," in *Proceedings of Signal and Data Processing of Small Targets*, vol. 5913. San Diego, CA, USA: SPIE, Aug. 2005, pp. 230–241.

[4] M. Baum and U. D. Hanebeck, "Random hypersurface models for extended object tracking," in *IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*, Ajman, United Arab Emirates, Dec. 2009, pp. 178–183.

[5] M. Baum, B. Noack, and U. D. Hanebeck, "Extended Object and Group Tracking with Elliptic Random Hypersurface Models," in *Proceedings of the International Conference on Information Fusion*, Edinburgh, UK, Jul. 2010.

[6] D. J. Salmond and M. C. Parr, "Track maintenance using measurements of target extent," *IEE Proceedings - Radar, Sonar and Navigation*, vol. 150, no. 6, pp. 389–395, Dec. 2003.

[7] K. Granström, C. Lundquist, and U. Orguner, "Tracking Rectangular and Elliptical Extended Targets Using Laser Measurements," in *Proceedings of the International Conference on Information Fusion*, Chicago, IL, USA, Jul. 2011, pp. 592–599.

[8] C. Lundquist, K. Granström, and U. Orguner, "Estimating the Shape of Targets with a PHD Filter," in *Proceedings of the International Conference on Information Fusion*, Chicago, IL, USA, Jul. 2011, pp. 49–56.

[9] M. Baum and U. D. Hanebeck, "Shape Tracking of Extended Objects and Group Targets with Star-Convex RHMs," in *Proceedings of the International Conference on Information Fusion*, Chicago, IL, USA, Jul. 2011, pp. 338–345.

[10] H. Zhu, C. Han, and C. Li, "An extended target tracking method with random finite set observations," in *Proceedings of the International Conference on Information Fusion*, Chicago, IL, USA, Jul. 2011, pp. 73–78.

[11] R. Mahler, *Statistical Multisource-Multitarget Information Fusion*. Norwood, MA, USA: Artech House, 2007.

[12] ——, "Multitarget Bayes filtering via first-order multi target moments," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 39, no. 4, pp. 1152–1178, Oct. 2003.

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

[14] R. Mahler, "PHD filters for nonstandard targets, I: Extended targets," in *Proceedings of the International Conference on Information Fusion*, Seattle, WA, USA, Jul. 2009, pp. 915–921.

[15] D. Clark and S. Godsill, "Group target tracking with the gaussian mixture probability hypothesis density filter," in *International Conference on Intelligent Sensors, Sensor Networks and Information (ISSNIP)*, Melbourne, Australia, Dec. 2007, pp. 149–154.

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

[17] K. Granström, C. Lundquist, and U. Orguner, "Extended Target Tracking using a Gaussian Mixture PHD filter," *IEEE Transactions on Aerospace and Electronic Systems*.

[18] ——, "A Gaussian mixture PHD filter for extended target tracking," in *Proceedings of the International Conference on Information Fusion*, Edinburgh, UK, Jul. 2010.

[19] J. W. Koch, "Bayesian approach to extended object and cluster tracking using random matrices," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 44, no. 3, pp. 1042–1059, Jul. 2008.

[20] J. W. Koch and M. Feldmann, "Cluster tracking under kinematical constraints using random matrices," *Robotics and Autonomous Systems*, vol. 57, no. 3, pp. 296–309, Mar. 2009.

[21] M. Feldmann, D. Fränken, and J. W. Koch, "Tracking of extended objects and group targets using random matrices," *IEEE Transactions on Signal Processing*, vol. 59, no. 4, pp. 1409–1420, Apr. 2011.

[22] W. Wieneke and J. W. Koch, "Probabilistic tracking of multiple extended targets using random matrices," in *Proceedings of SPIE Signal and Data Processing of Small Targets*, Orlando, FL, USA, Apr. 2010.

[23] M. Baum, M. Feldmann, D. Fränken, U. D. Hanebeck, and J. W. Koch, "Extended object and group tracking: A comparison of random matrices and random hypersurface models," in *Proceedings of the IEEE ISIF Workshop on Sensor Data Fusion: Trends, Solutions, Applications (SDF)*, Leipzig, Germany, Oct. 2010.

[24] D. Salmond and N. Gordon, "Group and extended object tracking," in *IEE Colloquium on Target Tracking: Algorithms and Applications*, 1999, pp. 16/1–16/4.

[25] M. Ulmke, O. Erdinc, and P. Willett, "Gaussian Mixture Cardinalized PHD Filter for Ground Moving Target Tracking," in *Proceedings of the International Conference on Information Fusion*, Quebec City, Canada, Jul. 2007, pp. 1–8.

[26] K. Granström and U. Orguner, "Implementation of the GIW-PHD filter," Department of Electrical Engineering, Linköping University, SE-581 83 Linköping, Sweden, Tech. Rep. LiTH-ISY-R-3046, Mar. 2012. [Online]. Available: http://www.control.isy.liu.se/publications/doc?id=2508

[27] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the em algorithm," *Journal of the Royal Statistical Society, Series B (Methodological)*, vol. 39, no. 1, pp. 1–38, 1977.

[28] C. M. Bishop, *Pattern recognition and machine learning*. New York, USA: Springer, 2006.

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

[30] H. Blom and Y. Bar-Shalom, "The interacting multiple model algorithm for systems with Markovian switching coefficients," *IEEE Transactions on Automatic Control*, vol. 33, no. 8, pp. 780–783, aug 1988.

[31] A. K. Gupta and D. K. Nagar, *Matrix variate distributions*, ser. Chapman & Hall/CRC monographs and surveys in pure and applied mathematics. Chapman & Hall, 2000.

Karl Granström received the M.Sc. degree in Applied Physics and Electrical Engineering in 2008, and the Tech.Lic. degree in Automatic Control in 2011, both from Linköping University, Linköping, Sweden. At present he is working towards a Ph.D. degree at the Division of Automatic Control at the Department of Electrical Engineering, Linköping University.

His research interests include sensor fusion, target tracking, mapping and localisation, especially for mobile robots.

Umut Orguner is an Assistant Professor at Department of Electrical & Electronics Engineering of Middle East Technical University, Ankara, Turkey. He got his B.S., M.S. and Ph.D. degrees all in Electrical Engineering from the same department in 1999, 2002 and 2006 respectively. He held a postdoctoral position between 2007 and 2009 at the Division of Automatic Control at the Department of Electrical Engineering of Linköping University, Linköping, Sweden. Between 2009 and 2012 he was an Assistant Professor at the Division of Automatic Control at the Department of Electrical Engineering of Linköping University, Linköping, Sweden.

His research interests include estimation theory, multiple-model estimation, target tracking, and information fusion.