A Note on the Reward Function for PHD Filters with Sensor Control

Branko Ristic, Ba-Ngu Vo, Daniel Clark

Abstract

The context is sensor control for multi-object Bayes filtering in the framework of partially observed Markov decision processes. The current information state is represented by the multi-object probability density function (PDF), while the reward function, associated with each sensor control (action), is the information gain measured by the alpha or Rényi divergence. Assuming that both the predicted and updated state can be represented by independent identically distributed (IID) cluster random finite sets (RFS's) or, as a special case, the Poisson RFS's, the paper derives the analytic expressions of the corresponding Rényi divergence based information gains. The implementation of Rényi divergence via the sequential Monte Carlo method is presented. The performance of the proposed reward function is demonstrated by a numerical example, where a moving range-only sensor is controlled to estimate the number and the states of several moving objects using the PHD filter.

Index Terms

Random finite sets, Finite Set Statistics, multi-object filtering, sequential Monte Carlo method, PHD filter, particle filter, sensor management, Rényi divergence.

I. Introduction

Sensors in modern surveillance systems are increasingly controllable. For example, a sensor can be operated in a different mode, pointed to a different direction or tasked to move to a new location. Sensor management is the on-line control of individual sensors for the purpose of maximizing the overall utility of the surveillance system. Sensor management thus represents sequential decision making, where each decision generates new observations that provide additional information. The decisions are made in the presence of uncertainty (both in the state and the measurement space) using only the past observations. This class of problems has been studied in the framework of partially observed Markov decision processes (POMDPs) [1]. The elements of a POMDP include the current (uncertain) information state, a set of admissible sensor actions and the reward function associated with each action.

B. Ristic is with ISR Division, Defence Science and Technology Organisation, Melbourne, Australia

B.-N. Vo is with the School of EECE, The University of Western Australia, Crawley, Australia

D. Clark is with Joint Research Institute in Signal and Image Processing, Heriot-Watt University, Edinburgh, UK

◇ The manuscript is prepared for submission to IEEE Trans. Aerospace and Electronic Systems. Date: March 8, 2011.

The paper adopts the information theoretic approach to sensor management, where the uncertain information state is represented by a probability density function, while the reward function is a measure of the *information gain* associated with each action. The particular reward function adopted for this purpose is the Rényi or α divergence [2], which for different values of parameter α reduces to the well known measures of mutual information, e.g. the Kullback-Leibler divergence and the Hellinger affinity.

The paper considers sensor management for the purpose of Bayesian multi-object filtering. In this context, the number of existing objects and their positions in the state-space are stochastically varying in time according to a priori known stochastic models. The sensors are imperfect in the sense that existing targets may or may not not be detected, while false detections are also reported with some non-zero probability. The objective of multi-object filtering is to sequentially estimate both the number of objects in the surveillance volume of interest and their states, using the sequence of noisy and cluttered measurement sets collected by sensors. Finite set statistics (FISST) [3] has provided the first numerically tractable solution to this problem. FISST uses the Bayesian framework to recursively propagate (predict and update) a multi-object probability density function using a set-valued measurement received at each time step. The success of FISST for multi-object filtering has been demonstrated by the PHD filter [4], a linear complexity algorithm which has already attracted significant interest with applications in robotics [5], [6], computer vision [7], traffic monitoring [8], etc.

Sensor management in the context of the FISST multi-object Bayes filtering has been considered earlier. Mahler put forward the idea of using the Csiszár or f-divergence as the reward function for sensor management [9]. However, he did not further develop or implement this reward function; instead he introduced and focused on theoretical foundations and applications of another reward function, referred to as the posterior expected number of targets (PENT) [10], [11]. Witkoskie et al. [12] applied the Bayesian multi-object filter implemented via the Gaussian mixture approximation, in conjunction with the "restless bandit" approach to sensor management, to track moving vehicles in a road network.

This paper can be seen as a sequel to [13], in which multi-object filtering was carried out using the (full) multi-object Bayes filter, while the reward function associated with each sensor action was computed for one step ahead (myopic) via the Rényi divergence between the multi-object prior and the multi-object posterior FISST PDF. The computation of both the filter and the reward function was carried out using the sequential Monte Carlo method. While the methodology proposed in [13] is general and accurate, it becomes computationally intractable even for a small number of objects. Hence the current paper is devoted to sensor management (in particular, the reward function) for PHD filters: the computationally efficient but principled approximations of the multi-object Bayes filter. The paper derives analytical expressions for the Rényi divergence under the assumption of independent identically distributed (IID) cluster random finite sets (RFS's) and Poisson RFS's. The IID cluster RFS is the basis of the cardinalized PHD filter [14], while its special case, the Poisson RFS, is the basis of the PHD filter [4]. Both the PHD filter and its appropriate Rényi divergence for myopic sensor management are implemented using the sequential Monte Carlo method. A numerical example involving multiple moving objects and a controllable moving range-only measuring sensor is presented in support of the proposed reward function.

The paper is organised as follows. Sec. II presents a brief background on finite-set statistics and the PHD filter. Sec.III describes the derivation and implementation of the Rényi divergence as the reward function for PHD filters with controllable sensors. Sec.IV demonstrates the performance of the proposed reward function in the context of multi-object nonlinear filtering with a controllable moving range-only sensor. The conclusions are drawn in Sec.V.

II. Finite set statistics and the PHD filter

The section presents a brief overview of multi-object Bayes filtering following [3]. A multi-object density function $f(\mathbf{X})$ is a positive real-valued function of a random finite-set variable $\mathbf{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$, characterized by a random number of points (objects) $n \in \mathbb{N}_0$ and random spatial locations of objects $\mathbf{x}_1, \ldots, \mathbf{x}_n \in \mathcal{X}$. The multi-object density function is normalised, that is, its set integral equals unity:

$$\int_{\mathcal{F}(\mathcal{X})} f(\mathbf{X})\delta\mathbf{X} := f(\emptyset) + \sum_{n=1}^{\infty} \frac{1}{n!} \int_{\underbrace{\mathcal{X} \times \cdots \times \mathcal{X}}_{n}} f(\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}) d\mathbf{x}_1, \ldots, d\mathbf{x}_n = 1. \tag{1}$$

Here $\mathcal{F}(X)$ is the space of finite subsets of $\mathcal{X}$, while $f(\{\mathbf{x}_1, \ldots, \mathbf{x}_n\})$ is defined as the joint distribution scaled by its cardinality distribution $\rho(n) = Pr\{|\mathbf{X}| = n\}$, i.e. $f(\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}) := n! \cdot \rho(n) \cdot f(\mathbf{x}_1, \ldots, \mathbf{x}_n)$, with $\sum_{n=0}^{\infty} \rho(n) = 1$ and $\int f(\mathbf{x}_1, \ldots, \mathbf{x}_n) d\mathbf{x}_1, \ldots, d\mathbf{x}_n = 1$.

Let k be the discrete-time index. The objective of the multi-object Bayes filter [3] is to determine at each time step k the *posterior* multi-object PDF $f_{k|k}(\mathbf{X}_k|\mathbf{Z}_{1:k})$, where $\mathbf{Z}_{1:k} = (\mathbf{Z}_1, \ldots, \mathbf{Z}_k)$ denotes the accumulated measurement set sequence up to time k, and $\mathbf{X}_k$ is the RFS representing the multi-object state at time k. Measurement set at time k, $\mathbf{Z}_k = \{\mathbf{z}_{k,1}, \ldots, \mathbf{z}_{k,m_k}\}$ is also modelled by a RFS; typically it contains some measurements due to objects $\mathbf{X}_k = \{\mathbf{x}_{k,1}, \ldots, \mathbf{x}_{k,n_k}\}$, but may also include other spurious detections, referred to as clutter.

The multi-object posterior can be computed sequentially via the cycle of prediction and update steps. Suppose that $f_{k-1|k-1}(\mathbf{X}_{k-1}|\mathbf{Z}_{1:k-1})$ is known and that a new set of measurements $\mathbf{Z}_k$, corresponding to time k, has become available. Then the predicted and updated multi-object posterior densities are calculated as follows [3]:

$$f_{k|k-1}(\mathbf{X}_k|\mathbf{Z}_{1:k-1}) = \int \Pi_{k|k-1}(\mathbf{X}_k|\mathbf{X}_{k-1})\, f_{k-1|k-1}(\mathbf{X}_{k-1}|\mathbf{Z}_{1:k-1})\delta\mathbf{X}_{k-1} \tag{2}$$

$$f_{k|k}(\mathbf{X}_k|\mathbf{Z}_{1:k}) = \frac{\varphi_k(\mathbf{Z}_k|\mathbf{X}_k) f_{k|k-1}(\mathbf{X}_k|\mathbf{Z}_{1:k-1})}{\int \varphi_k(\mathbf{Z}_k|\mathbf{X}) f_{k|k-1}(\mathbf{X}|\mathbf{Z}_{1:k-1})\delta\mathbf{X}}, \tag{3}$$

where $\Pi_{k|k-1}(\mathbf{X}_k|\mathbf{X}_{k-1})$ is a multi-object transition density and $\varphi_k(\mathbf{Z}_k|\mathbf{X}_k)$ is a multi-object likelihood function. Eq. (2) is a Chapman-Kolmogorov equation for multi-object densities, while (3) follows from the Bayes rule.

Since $f_{k|k}(\mathbf{X}_k|\mathbf{Z}_{1:k})$ is defined over $\mathcal{F}(\mathcal{X})$, the computational complexity of the multi-object Bayes filter grows exponentially with the number of objects and hence all practical implementations are limited to a small number of objects [13], [15].

In order to overcome this limitation, Mahler [4] proposed to propagate only the first-order statistical moment of $f_{k|k}(\mathbf{X}|\mathbf{Z}_{1:k})$, the intensity function or the PHD $D_{k|k}(\mathbf{x}|\mathbf{Z}_{1:k}) = \int \delta_{\mathbf{X}}(\mathbf{x}) f_{k|k}(\mathbf{X}|\mathbf{Z}_{1:k})\delta\mathbf{X}$, defined over $\mathcal{X}$.

Here $\delta_{\mathbf{X}}(\mathbf{x}) = \sum_{\mathbf{w}\in\mathbf{X}} \delta_{\mathbf{w}}(\mathbf{x})$ and $\delta_{\mathbf{w}}(\mathbf{x})$ is the Dirac delta function concentrated at $\mathbf{w}$. Note that the intensity function completely characterizes the Poisson RFS. Thus if $\mathbf{X}$ is a Poisson RFS, its intensity function is given by $D(\mathbf{x}) = \lambda \cdot s(\mathbf{x})$, while its multi-object PDF can be written as:

$$f(\mathbf{X}) = e^{-\lambda} \prod_{\mathbf{x}\in\mathbf{X}} \lambda \cdot s(\mathbf{x}). \tag{4}$$

The meaning of λ and $s(\mathbf{x})$ is as follows: λ is the Poisson average number of objects in the RFS, each distributed according to the spatial single-object density $s(\mathbf{x})$. For a given $D(\mathbf{x})$ one can work out $\lambda = \int D(\mathbf{x})d\mathbf{x}$ and $s(\mathbf{x}) = D(\mathbf{x})/\lambda$ and therefore (4) is completely determined.

PHD filter predicts and updates sequentially the intensity function. Using abbreviation $D_{k|k}(\mathbf{x}|\mathbf{Z}_{1:k}) \overset{\text{abbr}}{=} D_{k|k}(\mathbf{x})$ for the posterior intensity function at time k, the prediction equation of the PHD filter is given by [4]:

$$D_{k|k-1}(\mathbf{x}) = \gamma_{k|k-1}(\mathbf{x}) + \left\langle\, p_S\, D_{k-1|k-1},\ \pi_{k|k-1}(\mathbf{x}|\cdot)\, \right\rangle \tag{5}$$

where

- $\gamma_{k|k-1}(\mathbf{x})$ is the PHD of object births between time $k-1$ and k;
- $p_S(\mathbf{x}') \overset{\text{abbr}}{=} p_{S,k|k-1}(\mathbf{x}')$ is the probability that a target with state $\mathbf{x}'$ at time $k-1$ will survive until time k;
- $\pi_{k|k-1}(\mathbf{x}|\mathbf{x}')$ is the single-object transition density from time $k-1$ to k;
- $\langle g, f\rangle = \int f(\mathbf{x})\, g(\mathbf{x})\, d\mathbf{x}$.

The first term on the RHS of (5) refers to the newborn objects, while the second represents the persistent objects. Upon receiving the measurement set $\mathbf{Z}_k$ at time k, the update step of the PHD filter is computed according to:

$$D_{k|k}(\mathbf{x}) = [1 - p_D(\mathbf{x})]\, D_{k|k-1}(\mathbf{x}) + \sum_{\mathbf{z}\in\mathbf{Z}_k} \frac{p_D(\mathbf{x}) g_k(\mathbf{z}|\mathbf{x})\, D_{k|k-1}(\mathbf{x})}{\kappa_k(\mathbf{z}) + \left\langle p_D\ g_k(\mathbf{z}|\cdot), D_{k|k-1}\right\rangle} \tag{6}$$

where

- $p_D(\mathbf{x}) \overset{\text{abbr}}{=} p_{D,k}(\mathbf{x})$ is the probability that an observation will be collected at time k from a target with state $\mathbf{x}$;
- $g_k(\mathbf{z}|\mathbf{x})$ is the single-object measurement likelihood at time k;
- $\kappa_k(\mathbf{z})$ is the PHD of clutter at time k.

In order to derive a closed form solution for the update step, the cardinalized PHD (CPHD) filter makes the assumption that the RFS is an IID cluster [16]. An IID cluster RFS $\mathbf{X}$ is completely specified by its cardinality distribution $\rho(n)$ and the spatial single-object density $s(\mathbf{x})$ defined on $\mathcal{X}$. The multi-object PDF of $\mathbf{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ with $|\mathbf{X}| = n$ is then:

$$f(\mathbf{X}) := n! \cdot \rho(n) \cdot s(\mathbf{x}_1) \cdots s(\mathbf{x}_n) \tag{7}$$

with the intensity function given by

$$D(\mathbf{x}) = s(\mathbf{x}) \sum_{n=1}^{\infty} n \cdot \rho(n). \tag{8}$$

The prediction and update steps of the CPHD filter are omitted for brevity, with full details given in [14], [17].

III. The reward function for Sensor control

A. Adopted framework

The elements of a POMDP include the current (uncertain) information state, a set of admissible sensor actions and the reward function associated with each action. Let $\mathbf{u}_k(i_k) \in \mathbb{U}_k(i_k)$ denote the control vector applied at time k to observer $i_k \in \{1, \ldots, N_o\}$ where $\mathbb{U}_k(i)$ is the set of admissible control vectors at time k for observer i. The current information state is described by the predicted multi-object posterior PDF at time k, $f_{k|k-1}(\mathbf{X}_k|\mathbf{Z}_{1:k-1})$, where the sequence of measurement sets $\mathbf{Z}_{1:k-1}$ was collected using the control vector sequence $\mathbf{u}_{0:k-1} = (\mathbf{u}_0(i_0), \ldots, \mathbf{u}_{k-1}(i_{k-1}))$. After processing the first $k-1$ measurement sets, the control vector $\mathbf{u}_k(i_k)$, to be applied at time k to observer i_k, is selected as

$$\mathbf{u}_k = \arg\max_{\mathbf{v}\in\mathbb{U}_k} \mathbb{E}[\mathcal{R}(\mathbf{v}, f_{k|k-1}(\mathbf{X}_k|\mathbf{Z}_{1:k-1}), \mathbf{Z}_k(\mathbf{v}))] \tag{9}$$

where $\mathcal{R}(\mathbf{v}, f, \mathbf{Z})$ is the real-valued reward function associated with the control $\mathbf{v}$, at the time when the information state is represented by the multi-object PDF f and when the application of control $\mathbf{v}$ would result in the (future) measurement set $\mathbf{Z}$. Since the goal is to decide on the future action without actually applying any action (and collecting any measurement sets prior to the decision), the expectation $\mathbb{E}$ in (9) is taken with respect to the measurement set $\mathbf{Z}_k$. For computational simplicity, eq.(9) is restricted to maximization of the reward based on a single future step only (the "myopic" policy).

The quality of estimation depends to a large extent on the choice of the reward function $\mathcal{R}(\mathbf{v}, f, \mathbf{Z})$ in (9). In the information theoretic context, this function is typically based on the *information gain*, measured from the current information state. For this purpose we adopt the Rényi divergence between the predicted multi-object posterior density $f_{k|k-1}(\mathbf{X}_k|\mathbf{Z}_{1:k-1})$ and the future updated posterior multi-object density $f_{k|k}(\mathbf{X}_k|\mathbf{Z}_{1:k-1}, \mathbf{Z}_k(\mathbf{u}_k))$, which is computed using the new measurement set $\mathbf{Z}_k$, obtained after sensor i_k has been controlled to take action $\mathbf{u}_k(i_k)$. The reward function can then be written as (in order to simplify notation we suppress the second and third argument of $\mathcal{R}$):

$$\mathcal{R}(\mathbf{u}_k) = \frac{1}{\alpha-1}\log\int [f_{k|k}(\mathbf{X}_k|\mathbf{Z}_{1:k-1}, \mathbf{Z}_k(\mathbf{u}_k))]^{\alpha}\,[f_{k|k-1}(\mathbf{X}_k|\mathbf{Z}_{1:k-1})]^{1-\alpha}\delta\mathbf{X}_k. \tag{10}$$

Here $0 < \alpha < 1$ is a parameter which determines how much we emphasize the tails of two distributions in the metric. By taking the limit $\alpha \to 1$ or by setting $\alpha = 0.5$, the Rényi divergence becomes the Kullback-Leibler divergence and the Hellinger affinity [18], respectively [2].

B. Rényi divergence for IID cluster RFS and Poisson RFS

Let us rewrite (10) using a simpler notation:

$$\mathcal{R}(\mathbf{u}) = \frac{1}{\alpha-1}\log\int f_1(\mathbf{X};\mathbf{u})^{\alpha} f_0(\mathbf{X})^{1-\alpha}\delta\mathbf{X} \tag{11}$$

Suppose that both the predicted and updated multi-object PDFs are IID cluster PDFs, i.e.

$$f_0(\mathbf{X}) = n!\,\rho_0(n) \prod_{\mathbf{x}\in\mathbf{X}} s_0(\mathbf{x}), \tag{12}$$

$$f_1(\mathbf{X};\mathbf{u}) = n!\,\rho_1(n;\mathbf{u}) \prod_{\mathbf{x}\in\mathbf{X}} s_1(\mathbf{x};\mathbf{u}). \tag{13}$$

This applies to the recursions of the cardinalized PHD filter. The Rényi divergence (11) can then be written as:

$$\mathcal{R}(\mathbf{u}) = \frac{1}{\alpha-1}\log\sum_{n=0}^{\infty}\rho_1(n;\mathbf{u})^\alpha\,\rho_0(n)^{1-\alpha}\left[\int s_1(\mathbf{x};\mathbf{u})^\alpha\; s_0(\mathbf{x})^{1-\alpha}d\mathbf{x}\right]^n \tag{14}$$

Note that the sum in (14) is the probability generating function of the discrete distribution $p_\alpha(n;\mathbf{u}) = \rho_1(n;\mathbf{u})^\alpha\,\rho_0(n)^{1-\alpha}$ with parameter $z_\alpha(\mathbf{u}) = \int s_1(\mathbf{x};\mathbf{u})^\alpha\; s_0(\mathbf{x})^{1-\alpha}d\mathbf{x}$.

Let us assume now that the predicted and the updated cardinality distributions are both Poisson, with means λ_0 and $\lambda_1(\mathbf{u})$, respectively, i.e.

$$\rho_0(n) = \frac{e^{-\lambda_0}\lambda_0}{n!}, \qquad \rho_1(n;\mathbf{u}) = \frac{e^{-\lambda_1(\mathbf{u})}\lambda_1(\mathbf{u})}{n!}. \tag{15}$$

This corresponds to the assumption that both the predicted and updated RFSs are Poisson RFSs, and applies to the recursions of the PHD filter. As stated earlier, in this case

$$\lambda_0 = \int D_0(\mathbf{x})d\mathbf{x}, \qquad \lambda_1(\mathbf{u}) = \int D_1(\mathbf{x};\mathbf{u})d\mathbf{x}, \tag{16}$$

which follows from (8) since both $s_0(\mathbf{x})$ and $s_1(\mathbf{x};\mathbf{u})$ are normalised densities. The Rényi divergence (14) for Poisson RFSs is then:

$$\mathcal{R}(\mathbf{u}) = \frac{1}{\alpha-1}\left\{-\lambda_1(\mathbf{u})\alpha - \lambda_0(1-\alpha) + \log\sum_{n=0}^{\infty}\frac{1}{n!}\left[\int \lambda_1(\mathbf{u})^\alpha\, s_1(\mathbf{x};\mathbf{u})^\alpha\,\lambda_0^{1-\alpha}\; s_0(\mathbf{x})^{1-\alpha}d\mathbf{x}\right]^n\right\} \tag{17}$$

which using identity $e^x = \sum_{n=0}^{\infty}\frac{x^n}{n!}$ simplifies to:

$$\mathcal{R}(\mathbf{u}) = \lambda_0 + \frac{\alpha}{1-\alpha}\lambda_1(\mathbf{u}) + \frac{\lambda_1(\mathbf{u})^\alpha\lambda_0^{1-\alpha}}{\alpha-1}\int s_1(\mathbf{x};\mathbf{u})^\alpha\; s_0(\mathbf{x})^{1-\alpha}\,d\mathbf{x}. \tag{18}$$

Let us now summarize the main steps of the PHD filter with sensor management. After performing the PHD update according to (5), the next sensor control vector is selected as:

$$\mathbf{u}_k = \arg\max_{\mathbf{v}\in\mathbb{U}_k}\mathbb{E}[\mathcal{R}(\mathbf{v})] \tag{19}$$

where

$$\mathcal{R}(\mathbf{v}) = \int D_{k|k-1}(\mathbf{x})d\mathbf{x} + \frac{\alpha}{1-\alpha}\int D_{k|k}(\mathbf{x};\mathbf{v})d\mathbf{x} - \frac{1}{1-\alpha}\int D_{k|k}(\mathbf{x};\mathbf{v})^\alpha\; D_{k|k-1}(\mathbf{x})^{1-\alpha}\,d\mathbf{x} \tag{20}$$

and $D_{k|k}(\mathbf{x};\mathbf{v})$ is a shortened notation for $D_{k|k}(\mathbf{x}|\mathbf{Z}_{1:k-1},\mathbf{Z}_k(\mathbf{v}))$, computed via the PHD filter update equation (6). Note that the adopted framework for sensor control is directly applicable to multiple observers if the fusion architecture is centralised and the measurements from different sensors/observers are asynchronous (non-simultaneous)[1].

[1]For simultaneous multi-sensor measurements the PHD filter update has no closed form solution, see [19].

Observe from (20) that if the predicted and updated intensity functions are equal, i.e. $D_{k|k-1}(\mathbf{x}) = D_{k|k}(\mathbf{x};\mathbf{u})$, then $\mathcal{R}(\mathbf{v}) = 0$. For $0 < \alpha < 1$, the third integral on the RHS of (20) is limited as follows:

$$0 \leq \int D_{k|k}(\mathbf{x};\mathbf{u})^{\alpha}\; D_{k|k-1}(\mathbf{x})^{1-\alpha}\; d\mathbf{x} \leq \left[\int D_{k|k}(\mathbf{x};\mathbf{v}) d\mathbf{x}\right]^{\alpha} \left[\int D_{k|k-1}(\mathbf{x}) d\mathbf{x}\right]^{1-\alpha} \tag{21}$$

Hence, the Rényi divergence in (20) will have the maximum value if the support of the predicted PHD and the support of the updated PHD do not overlap, that is $\int D_{k|k}(\mathbf{x};\mathbf{u})^{\alpha}\; D_{k|k-1}(\mathbf{x})^{1-\alpha}\; d\mathbf{x} \approx 0$. The theoretical asymptotic analysis of the locally optimum value of α [20] shows that $\alpha = 0.5$ provides the best discrimination between the two densities under consideration.

C. *SMC implementation*

The sequential Monte Carlo (SMC) method provides a general framework for the implementation of PHD filter recursions in (5) and (6). The resulting class of SMC-PHD filters have received considerable interest recently, see for example [21]–[25]. Assuming the PHD filter has been implemented by the SMC method, this section provides a brief description of the SMC implementation of the sensor control aspect, in particular eqs. (19) and (20).

Assume a set of weighted random samples (particles) $\{w^i_{k|k-1}, \mathbf{x}^i_{k|k-1}\}_{i=1}^N$ approximates the predicted intensity function:

$$D_{k|k-1}(\mathbf{x}) \approx \sum_{i=1}^{N} w^i_{k|k-1} \delta_{\mathbf{x}^i_{k|k-1}}(\mathbf{x}) \tag{22}$$

where $\delta_{\mathbf{x}_0}(\mathbf{x})$ is the Dirac delta function. For a "future" measurement set $\mathbf{Z}_k(\mathbf{v})$, obtained upon the execution of action $\mathbf{v} \in \mathbb{U}_k$, the updated intensity function is approximated as:

$$D_{k|k}(\mathbf{x}) \approx \sum_{i=1}^{N} w^i_{k|k} \delta_{\mathbf{x}^i_{k|k-1}}(\mathbf{x}) \tag{23}$$

where the weights $w^i_{k|k}$ are computed according to (6) as:

$$w^i_{k|k} = [1 - p_D(\mathbf{x}^i_{k|k-1})]\, w^i_{k|k-1} + \sum_{\mathbf{z}\in\mathbf{Z}_k(\mathbf{v})} \frac{p_D(\mathbf{x}^i_{k|k-1})\, g_k(\mathbf{z}|\mathbf{x}^i_{k|k-1})\, w^i_{k|k-1}}{\kappa_k(\mathbf{z}) + \sum_{i=1}^N p_D(\mathbf{x}^i_{k|k-1})\; g_k(\mathbf{z}|\mathbf{x}^i_{k|k-1})\; w^i_{k|k-1}}. \tag{24}$$

Note from (22) and (23) that an identical set of particles is used in approximating both $D_{k|k-1}(\mathbf{x})$ and $D_{k|k}(\mathbf{x})$ - only particle weights are different. This is particularly convenient for the computation of the Rényi divergence in (20), as its Monte Carlo approximation can then be written as:

$$\mathcal{R}(\mathbf{v}) \approx \sum_{i=1}^{N} w^i_{k|k-1} + \frac{\alpha}{1-\alpha}\sum_{i=1}^{N} w^i_{k|k} - \frac{1}{1-\alpha}\sum_{i=1}^{N}\left(w^i_{k|k}\right)^{\alpha}\;\left(w^i_{k|k-1}\right)^{1-\alpha}. \tag{25}$$

While the first two terms on the RHS of (25) are straightforward, in the light of the recent discussion in [26], the third term deserves further explanation, see Appendix.

The question remains how to carry out the expectation $\mathbb{E}$ over the future measurement sets $\mathbf{Z}_k(\mathbf{v})$ in (19). One approach is to generate an ensemble of $\mathbf{Z}_k(\mathbf{v})$, for given clutter intensity $\kappa_k(\mathbf{z})$, probability of detection p_D and the measurement likelihood $g_k(\mathbf{z}|\mathbf{x})$. This, however, would be computationally very intensive. Instead, the *predicted* *ideal measurement set* approach [9] is adopted here. The idea is to generate only one future measurement set for each action $\mathbf{v}$, the one which would result under the ideal conditions of no measurement noise, no clutter and $p_D = 1$. Suppose a measurement due to a single object can be expressed as $\mathbf{z} = h(\hat{\mathbf{x}}) + \boldsymbol{\zeta}$, where $\mathbf{z} \in \mathbf{Z}_k(\mathbf{v})$, $\hat{\mathbf{x}} \in \hat{\mathbf{X}}_{k|k-1}$ is the predicted state of a detected object (hence $\hat{\mathbf{X}}_{k|k-1}$ is the predicted multi-object state), and $\boldsymbol{\zeta}$ is measurement noise. Then for each action $\mathbf{v}$ an ideal measurement set at time k would be [9]:

$$\mathbf{Z}_k(\mathbf{v}) = \cup_{\hat{\mathbf{x}} \in \hat{\mathbf{X}}_{k|k-1}} \{h(\hat{\mathbf{x}})\} \tag{26}$$

This approach requires the estimate $\hat{\mathbf{X}}_{k|k-1}$ (the estimate of the number of predicted objects and their state vectors). An efficient (fast and accurate) method for multi-object estimation in the SMC framework is presented in [24].

IV. Numerical example

In order to demonstrate the proposed reward function, consider the problem where a controllable moving observer, equipped with a range-only sensor, is tasked to estimate the number of objects in a specified surveillance area and their positions and velocities. A single-object state is then a vector $\mathbf{x} = [\mathbf{p}^\intercal \;\; \mathbf{v}^\intercal]^\intercal$, where $\mathbf{p} = [x \;\; y]^\intercal$ is the position vector, $\mathbf{v} = [\dot{x} \;\; \dot{y}]^\intercal$ is the velocity vector, and $^\intercal$ denotes the matrix transpose.

Suppose the position of the observer is $\mathbf{u} = [\chi \;\; \gamma]^\intercal$. The sensor is capable of detecting an object at location $\mathbf{p}$ with the probability:

$$p_D(\mathbf{p}) = \begin{cases} 1, & \text{if } ||\mathbf{p} - \mathbf{u}|| \leq R_0 \\ \max\{0, 1 - (||\mathbf{p} - \mathbf{u}|| - R_0)\hbar\}, & \text{if } ||\mathbf{p} - \mathbf{u}|| > R_0 \end{cases} \tag{27}$$

where $||\mathbf{p} - \mathbf{u}|| = \sqrt{(x-\chi)^2 + (y-\gamma)^2}$ is the distance between the observer and object. In simulations we adopt a surveillance area to be a square of sides $a = 1000$m, with $R_0 = 320$m and $\hbar = 0.00025\text{m}^{-1}$. The range measurement originating from an object at location $\mathbf{p}$ is then $z = ||\mathbf{p} - \mathbf{u}|| + \zeta$, where ζ is zero-mean white Gaussian measurement noise, with standard deviation given by $\sigma_\zeta = \sigma_0 + \beta||\mathbf{p} - \mathbf{u}||^2$. In simulations we adopt $\sigma_0 = 1$m and $\beta = 5 \cdot 10^{-5}\text{m}^{-1}$. Then the single-object likelihood function is $g_k(z|\mathbf{x}) = \mathcal{N}(z; ||\mathbf{p} - \mathbf{u}||, \sigma_\zeta^2)$. The sensor reports false range measurements, modelled by a Poisson RFS. The intensity function of false measurements $\kappa(z) = \mu \cdot c(z)$ is specified by uniform density from zero to $\sqrt{2}a$, i.e. $c(z) = \mathcal{U}[0, \sqrt{2}a])$, with mean $\mu = 0.5$.

There are five objects in the surveillance area, positioned relatively close to each other. Their initial state vectors are (the units are omitted): $[800, 600, -1, 0]^\intercal$, $[650, 500, 0.3, 0.6]^\intercal$, $[620, 700, 0.25, -0.45]^\intercal$, $[750, 800, 0, -0.6]^\intercal$, and $[700, 700, 0.2, -0.6]^\intercal$. The objects move according to the constant velocity model. The observer enters the surveillance area at position $(10, 10)$m. Figure 1 show the location of objects (indicated by asterisks) and the observer path, after $k = 8$ time steps. Current measurements are represented by arcs (dashed blue lines).

The described scenario for numerical simulations is designed for a good reason. In order to achieve a high level of estimation accuracy using the PHD filter, the observer (sensor) needs to move towards the objects and then to remain in their vicinity, because in doing so, p_D will be high and σ_ζ will be low. A sensor control policy that does not follow this trend will result in a poor PHD filter error performance.

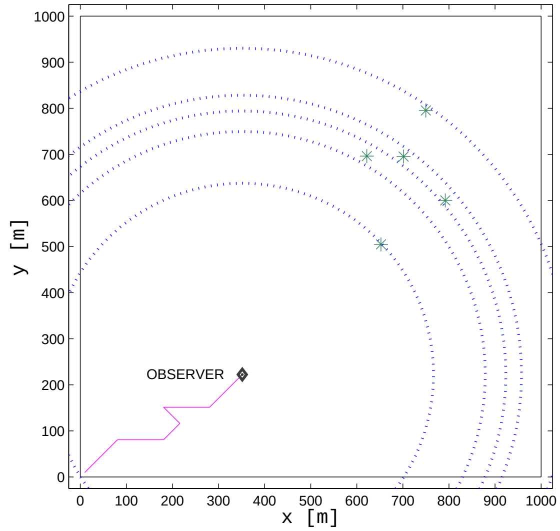


Fig. 1. *The position of objects and the observer at time $k = 8$; current measurements are represented by arcs (dashed blue lines)*

The details of the SMC-PHD filter for multi-object filtering are described in [24], [27]. The number of particles is 200 per estimated number of objects. The probability of survival is set to $p_S = 0.99$, and the transition density is $\pi_{k|k-1}(\mathbf{x}|\mathbf{x}') = \mathcal{N}(\mathbf{x}; \mathbf{F}\mathbf{x}', \mathbf{Q})$, where

$$\mathbf{F} = \begin{bmatrix} 1 & 0 & T & 0 \\ 0 & 1 & 0 & T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad \mathbf{Q} = q \begin{bmatrix} T^3/3 & 0 & T^2/2 & 0 \\ 0 & T^3/3 & 0 & T^2/2 \\ T^2/2 & 0 & T & 0 \\ 0 & T^2/2 & 0 & T \end{bmatrix}$$

with $q = 0.05$ and $T = 1$s. The object birth intensity is driven by measurements [24].

The set of admissible control vectors $\mathbb{U}_k$ is computed as follows. If the current position of the observer is $\mathbf{v}_k = [\chi_k \;\; \gamma_k]^T$, its one-step ahead future admissible locations are:

$$\mathbb{U}_k = \left\{ (\chi_k + j\Delta_R \cdot \cos(\ell\Delta_\theta), \gamma_k + j\Delta_R \cdot \sin(\ell\Delta_\theta)) \, ; \; j = 0, \ldots, N_R; \ell = 1, \ldots, N_\theta \right\} \tag{28}$$

where $\Delta_\theta = 2\pi/N_\theta$ and Δ_R is a conveniently selected radial step size. In this way the observer can stay in its current position $(j = 0)$ or move radially in incremental steps. The following values were adopted in simulations: $N_R = 2$, $N_\theta = 8$ and $\Delta_R = 50$m. In this way 17 control options are considered at each time step and for each of them, the reward function is computed as described in Sec. III-C. If a control vector $\mathbf{u}_k \in \mathbb{U}_k$ is outside the specified surveillance area, its reward is set to $-\infty$; in this way the observer is always kept inside the area of interest.

The comparative error performance of the SMC-PHD filter using three different reward functions for observer control is discussed next. In addition to the Rényi divergence based reward, we considered (1) a uniform random scheme; (2) the PENT criterion. The uniform random selects the control vector randomly from the set $\mathbb{U}_k$. The PENT was introduces for the control of sensors with a finite field-of-view (FoV), for the purpose of selecting the action which will maximize the number of objects to be seen by the sensor [10]. Since our sensor does not have a finite field of view, this reward function turns out to be independent of action and hence is expected to perform poorly.

The adopted performance error metric is the optimal sub-pattern assignment (OSPA) metric [28]. The OSPA measures the error between the estimated and the true multi-target states, $\hat{\mathbf{X}}_k$ and $\mathbf{X}_k$, respectively. Fig.2.(a) shows the temporal evolution of the OSPA error (order parameter $p = 2$ and cutoff $c = 100$m) averaged over 200 Monte Carlo runs for the three contesting reward functions. The corresponding localisation error and cardinality errors are shown in Fig.2.(b) and (c) respectively. Initially the OSPA error is large (for all three reward functions) reflecting the initial uncertainty about the number of objects and their states. As the time progresses and measurements are collected, the PHD filter using the Rényi divergence ($\alpha = 0.5$, $\alpha \to 1$) as the reward function performs superior to the other two schemes. This is so because the Rényi divergence immediately controls the observer to move towards the objects and then to zigzag in their vicinity in order to perform accurate localisation based on range-only measurements. Comparing the two OSPA error curves for Rényi divergence, note that the $\alpha = 0.5$ case converges quicker to the steady-state error (which is approximately attained for $k > 25$) than the $\alpha \to 1$ case.

V. Conclusion

The paper considered information theoretic sensor management for multi-object Bayes filtering. In this context, analytic expressions for Rényi divergence between two Poisson random finite sets and between two IID cluster random finite sets, have been derived. These expressions can serve as measures of information gain between the predicted and updated multi-object posterior densities in the PHD filter or the cardinalized PHD filter, respectively. In the context of partially observed Markov decision processes, they represent the rewards associated with each future sensor action. The numerical example demonstrated the effectiveness of the proposed reward function for PHD filtering with a controllable moving observer. Future work will investigate computationally efficient solutions for PHD filtering with (1) multiple steps ahead sensor control and (2) sensor control in distributed fusion architecture.

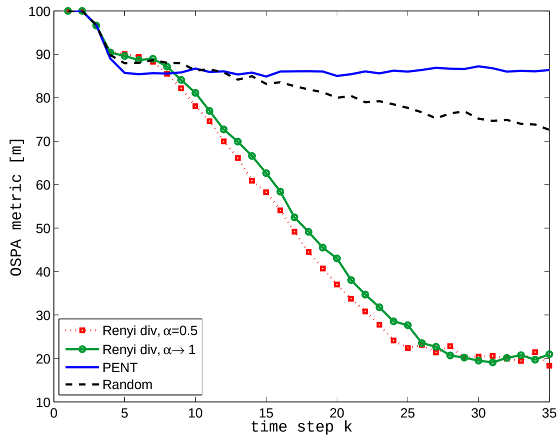
Appendix

The problem is Monte Carlo approximation of the integral

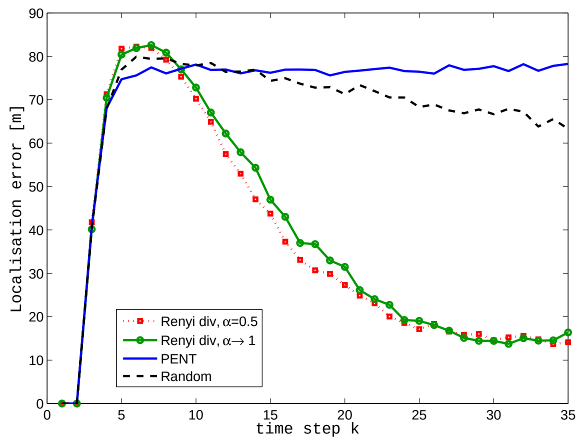
$$\mathcal{I} = \int D_{k|k}(\mathbf{x})^{\alpha}\ D_{k|k-1}(\mathbf{x})^{1-\alpha}\ d\mathbf{x},$$

which features in the third term on the RHS of (20). We adopted approximations (22) and (23) for $D_{k|k-1}(\mathbf{x})$ and $D_{k|k}(\mathbf{x})$, respectively. Note that in both of these approximations, the same supporting particle set $\{\mathbf{x}^i_{k|k-1}; i = 1, \ldots, N\}$ is used. Then we have:

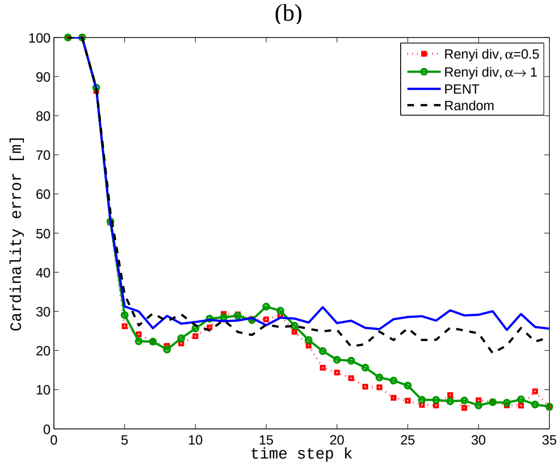
$$\mathcal{I} \approx \int \left[\sum_{i=1}^{N} w^i_{k|k}\ \delta_{\mathbf{x}^i_{k|k-1}}(\mathbf{x})\right]^{\alpha} \left[\sum_{i=1}^{N} w^i_{k|k-1}\ \delta_{\mathbf{x}^i_{k|k-1}}(\mathbf{x})\right]^{1-\alpha} d\mathbf{x} \tag{29}$$



(a)



(b)



(c)

Fig. 2. Error performance of the three contesting reward functions (averaged over 200 Monte Carlo runs): (a) OSPA metrics; (b) Localisation error; (c) Cardinality error

Note that when a sum of delta functions is raised to an exponent, the "cross-terms" vanish due to the different support of delta functions. Hence we can write:

$$\mathcal{I} \approx \int \left(\sum_{i=1}^{N} \left(w_{k|k}^{i}\right)^{\alpha} \, [\delta_{\mathbf{x}_{k|k-1}^{i}}(\mathbf{x})]^{\alpha} \right) \left(\sum_{i=1}^{N} \left(w_{k|k}^{i}\right)^{1-\alpha} \, [\delta_{\mathbf{x}_{k|k-1}^{i}}(\mathbf{x})]^{1-\alpha} \right) d\mathbf{x} \tag{30}$$

When the two sums of delta functions are multiplied, the “cross terms” again vanish for the same reason and the approximation simplifies to:

$$\mathcal{I} \approx \int \sum_{i=1}^{N} \left(w^i_{k|k}\right)^{\alpha} \left(w^i_{k|k-1}\right)^{1-\alpha} \delta_{\mathbf{x}^i_{k|k-1}}(\mathbf{x})\, d\mathbf{x} \tag{31}$$

$$= \sum_{i=1}^{N} \left(w^i_{k|k}\right)^{\alpha} \left(w^i_{k|k-1}\right)^{1-\alpha}. \tag{32}$$

The last expression features in the third term on the RHS of (25).

REFERENCES

[1] D. A. Castanón and L. Carin, “Stochastic control theory for sensor management,” in *Foundations and Applications of Sensor Management*, A. O. H. III, D. A. Castanón, D. Cochran, and K. Kastella, Eds. Springler, 2008, ch. 2, pp. 7–32.

[2] A. O. Hero, C. M. Kreucher, and D. Blatt, “Information theoretic approaches to sensor management,” in *Foundations and applications of sensor management*, A. O. Hero, D. Castanòn, D. Cochran, and K. Kastella, Eds. Springer, 2008, ch. 3, pp. 33–57.

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

[4] R. P. S. Mahler, “Multi-target bayes filtering via first-order multi-target moments,” *IEEE Trans. Aerospace & Electronic Systems*, vol. 39, no. 4, pp. 1152–1178, 2003.

[5] J. Mullane, B.-N. Vo, and M. D. Adams, “Rao-Blackwellised PHD SLAM,” in *Proc. IEEE Int. Conf. Robotics and Automation*, Alaska, USA, May 2010.

[6] A. N. Bishop and P. Jensfelt, “Global robot localization with random finite set statistics,” in *Proc. 13th Int. Conf. Information Fusion*, Edinburgh, UK, July 2010.

[7] E. Maggio, M. Taj, and A. Cavallaro, “Efficient multitarget visual tracking using random finite sets,” *IEEE Trans. Circuits & Systems for Video Technology*, vol. 18, no. 8, pp. 1016–1027, 2008.

[8] G. Battistelli, L. Chisci, S. Morrocchi, F. Papi, A. Benavoli, A. D. Lallo, A. Farina, and A. Graziano, “Traffic intensity estimation via PHD filtering,” in *Proc. 5th European Radar Conf.*, Amsterdam, The Netherlands, Oct. 2008, pp. 340–343.

[9] R. Mahler, “Multitarget sensor management of dispersed mobile sensors,” in *Theory and algorithms for cooperative systems*, D. Grundel, R. Murphey, and P. Pardalos, Eds. World Scientific books, 2004, ch. 12, pp. 239–310.

[10] R. Mahler and T. Zajic, “Probabilistic objective functions for sensor management,” in *Proc. SPIE*, vol. 5429, 2004, pp. 233–244.

[11] A. Zatezalo, A. El-Fallah, R. Mahler, R. K. Mehra, and K. Pham, “Joint search and sensor management for geosynchronous satellites,” in *Proc. SPIE*, vol. 6968, 2008.

[12] J. Witkoskie, W. Kukunski, S. Theophanis, and M. Otero, “Random set tracker experiment on a road constrained network with resource management,” in *Proc. 9th Int. Conf. Information Fusion*, Florence, Italy, 2006.

[13] B. Ristic and B.-N. Vo, “Sensor control for multi-object state-space estimation using random finite sets,” *Automatica*, 2010, (To be published).

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

[15] M. Vihola, “Rao-Blackwellised particle filtering in random set multitarget tracking,” *IEEE Trans. Aerospace & Electronic Systems*, vol. 43, no. 2, pp. 689–705, 2007.

[16] D. J. Daley and D. Vere-Jones, *An introduction to the Theory of Point Processes*. Springer, 1988.

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

[18] D. Pollard, *A user’s guide to measure theoretic probability*. Cambridge University Press, 2001.

[19] R. Mahler, “The multisensor PHD filter II: Errorneous solution via “Poisson magic”,” in *Proc. SPIE*, vol. 7336, 2009.

[20] A. O. Hero, B. Ma, O. Michel, and J. Gorman, “Alpha-divergence for classification, indexing and retrieval,” Comm. and Sig. Proc. Lab., Dept. EECS, The University of Michigan, Tech. Rep. CSPL-328, 2002.

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

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

[23] N. P. Whiteley, S. S. Singh, and S. J. Godsill, “Auxiliary particle implementation of the probability hypothesis density filter,” Cambridge University Engineering Department, Tech. Report CUED/F-INFENG/TR-592, Dec. 2007.

[24] B. Ristic, D. Clark, and B.-N. Vo, “Improved SMC implementation of the PHD filter,” in *Proc. 13th Int. Conf. Information Fusion*, Edinburgh, UK, July 2010.

[25] M. Morelande, “A sequential Monte Carlo method for PHD approximation with conditionally linear/Gaussian models,” in *Proc. 13th Int. Conf. Information Fusion*, Edinburgh, UK, July 2010.

[26] Y. Boers, H. Driessen, A. Bagchi, and P. Mandal, “Particle filter based entropy,” in *Proc. 13th Int. Conf. Information Fusion*, Edinburgh, UK, July, 2010 2010.

[27] B. Ristic, D. Clark, B.-N. Vo, and B.-T. Vo, “Adaptive target birth intensity in PHD and CPHD filters,” *IEEE Trans. Aerospace and Electronic Systems*, 2010, (In review).

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