

Manuscript version: Author's Accepted Manuscript

The version presented in WRAP is the author's accepted manuscript and may differ from the published version or Version of Record.

Persistent WRAP URL:

<http://wrap.warwick.ac.uk/142105>

How to cite:

Please refer to published version for the most recent bibliographic citation information. If a published version is known of, the repository item page linked to above, will contain details on accessing it.

Copyright and reuse:

The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions.

© 2020 Elsevier. Licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International <http://creativecommons.org/licenses/by-nc-nd/4.0/>.



Publisher's statement:

Please refer to the repository item page, publisher's statement section, for further information.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk.

Target tracking in the framework of possibility theory: The possibilistic Bernoulli filter

Branko Ristic^{a,*}, Jeremie Houssineau^b, Sanjeev Arulampalam^c

^a*RMIT University, Australia*

^b*University of Warwick, United Kingdom*

^c*Defence Science and Technology Group, Australia*

Abstract

The Bernoulli filter is a Bayes filter for joint detection and tracking of a target in the presence of false and miss detections. This paper presents a mathematical formulation of the Bernoulli filter in the framework of possibility theory, where uncertainty is represented using *possibility* functions, rather than *probability* distributions. Possibility functions model the uncertainty in a non-additive manner, and have the capacity to deal with partial (incomplete) problem specification. Thus, the main advantage of the possibilistic Bernoulli filter, derived in this paper, is that it can operate even in the absence of precise measurement and/or dynamic model parameters. This feature of the proposed filter is demonstrated in the context of target tracking using multi-static Doppler shifts as measurements.

Keywords: Target tracking; Possibility functions; Partially known probabilistic models.

*Corresponding author: B. Ristic, RMIT University, GPO Box 2476, Melbourne VIC 3001, Australia; tel: +61 3 9925 3302.

Email addresses: branko.ristic@rmit.edu.au (Branko Ristic),
jeremie.houssineau@warwick.ac.uk (Jeremie Houssineau),
sanjeev.arulampalam@dst.defence.gov.au (Sanjeev Arulampalam)

1. Introduction

Estimation of stochastic dynamic systems (stochastic filtering) is typically carried out using the sequential Bayesian estimation framework [1]. Assuming the dynamic system is fully characterised by its (hidden) state, the application of the sequential Bayesian estimation method requires the specification of two stochastic models: the dynamic model, which describes the evolution of the (hidden) state, and the observation model, which specifies the relationship between the sensor measurements and the (hidden) state. Practical applications of Bayes filtering are widespread, including target tracking, communications, navigation, field robotics, bio-informatics, finance, ecology, etc.

In situations where the dynamics and/or observation models are only partially known, using the sequential Bayesian estimation method is not straightforward. If, for example, some parameters of the model(s) are unknown, one approach would be to estimate (learn) their values sequentially from the data [2]. This, of course, has its limitations, because of limitations in available computational power or observability issues. Different methods such as Bayesian non-parametric models [3] allow for acknowledging that all the parameters in the selected dynamical and observation processes might not be perfectly known, however, these often involve even more parameters in order to describe what is the uncertainty on the original ones, thus only offering a partial solution to the problem. If there is no available model to describe the received data then neural networks can be used, see e.g. [4]; however, we consider the situation where some information is available about the dynamics of the target and about the corresponding observation process.

Research into reasoning under uncertainty in artificial intelligence (AI) is mainly focused on representation and explanation of uncertainty and inference rules for derivation of (uncertain) conclusions. Uncertainty in this context is classified either as aleatory (due to the random effects) or epistemic (due to imprecision, or partial knowledge) [5]. The research community in AI has recognised for some time that probability distributions are perfect to represent aleatory uncertainty, but inappropriate to capture effectively the uncertainty caused by ignorance, imprecision or partial knowledge [6, 7]. Alternative modelling of uncertainty have been proposed by different generalisations of probability theory, such as fuzzy logic [8], imprecise probabilities [9], possibility theory [10, 11] and Dempster-Shafer theory [12, 13]. Most of these approaches offer the ability to model a complete absence of information, but do not provide a general way of dealing with stochastic filtering. In addition, reasoning under uncertainty in AI is typically restricted to discrete state spaces.

In this paper we develop a stochastic filter for joint detection and state estimation of a dynamic object in the presence of false and miss detections, using exclusively *possibility* functions. The filter is referred to as the *possibilistic Bernoulli filter* (PBF), because it is the possibility theoretic analogue of the standard Bernoulli filter [14], originally derived by Mahler [15] using probability distributions and random finite sets, as the optimal Bayesian filter in the presence of false and miss detections. The motivation for using possibility functions, as non-additive models of uncertainty [16], instead of the probabilistic framework, is to provide an alternative representation of uncertainty, capable of handling, in a rigorous mathematical manner, the situations of ignorance or partial knowledge. Derivation of the PBF follows from the recently proposed framework for stochastic filtering using a class of outer measures [17, 18, 19, 20]. Bayes filtering style analytic expressions for prediction and update of outer measures have been formulated and implemented using numerical approximations, such as the grid-based and sequential Monte Carlo (SMC) methods.

The paper is organised as follows. Following [20], Sec. 2 reviews the standard Bayes filter and the possibilistic stochastic filter. The PBF is derived in Sec. 3 for the multi-sensor case and an application to target tracking using multi-static Doppler measurements is presented in Sec. 4. The emphasis in this application is that the probability of detection of each sensor is only partially known, that is, as an interval value. The findings in this article are summarised in Sec. 5.

2. Background

2.1. The standard Bayes filter

The stochastic filtering problem in the Bayesian framework can be formulated as follows [1]. Let us introduce a random variable $\mathbf{x}_k \in \mathcal{X} \subseteq \mathbb{R}^{n_x}$, referred to as the state-vector, as the complete specification of the state of a dynamic system at time t_k . Here \mathcal{X} is the state space, while k is the discrete-time index corresponding to time t_k . The problem is specified with two equations [21]:

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

$$\mathbf{z}_k = \mathbf{h}_k(\mathbf{x}_k) + \mathbf{w}_k, \quad (2)$$

referred to as the *dynamics equation* and the *observation (or measurement) equation*, respectively. The function $\mathbf{f}_{k-1} : \mathbb{R}^{n_x} \rightarrow \mathbb{R}^{n_x}$ is a nonlinear transition function defining the evolution of the state vector as a first-order Markov process. The random process $\mathbf{v}_k \in \mathbb{R}^{n_x}$ is independent identically distributed (IID) according to the probability density function (PDF) $p_{\mathbf{v}}$; and \mathbf{v}_k is referred to

as *process noise*. Its role is to model random disturbances affecting the state evolution model. The function $\mathbf{h}_k : \mathbb{R}^{n_x} \rightarrow \mathbb{R}^{n_z}$ defines the relationship between the state \mathbf{x}_k and the measurement $\mathbf{z}_k \in \mathcal{Z}$, where $\mathcal{Z} \subseteq \mathbb{R}^{n_z}$ is the measurement space. The random process $\mathbf{w}_k \in \mathbb{R}^{n_x}$, independent of \mathbf{v}_k , is also IID with PDF $p_{\mathbf{w}}$, and referred to as *measurement noise*.

In the formulation (1)-(2), functions \mathbf{f}_k and \mathbf{h}_k , as well as PDFs $p_{\mathbf{v}}$ and $p_{\mathbf{w}}$ are known. Equations (1) and (2) effectively define two probability distributions, the *transitional density* $p_{k|k-1}(\mathbf{x}_k|\mathbf{x}_{k-1})$ and the likelihood function $\ell_k(\mathbf{z}_k|\mathbf{x}_k)$, respectively. Given the transitional density, the likelihood function, and the initial density of the state (at $k = 0$), $p_0(\mathbf{x}_0)$, the goal of stochastic Bayesian filtering is to estimate recursively the posterior PDF of the state, denoted $p_{k|k}(\mathbf{x}_k|\mathbf{z}_{1:k})$, where $\mathbf{z}_{1:k} \stackrel{\text{abbr}}{=} \mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_k$.

The solution is usually presented as a two step procedure. Let $p_{k-1|k-1}(\mathbf{x}|\mathbf{z}_{1:k-1})$ denote the posterior PDF at $k - 1$. The first step *predicts* the density of the state to (the future) time k via the Chapman - Kolmogorov equation [1]:

$$p_{k|k-1}(\mathbf{x}|\mathbf{z}_{1:k-1}) = \int p_{k|k-1}(\mathbf{x}|\mathbf{x}')p(\mathbf{x}'|\mathbf{z}_{1:k-1})d\mathbf{x}'. \quad (3)$$

The second step applies the Bayes rule to *update* the predicted PDF using a measurement \mathbf{z}_k which becomes available at time k :

$$p_{k|k}(\mathbf{x}|\mathbf{z}_{1:k}) = \frac{\ell_k(\mathbf{z}_k|\mathbf{x}) p_{k|k-1}(\mathbf{x}|\mathbf{z}_{1:k-1})}{\int \ell_k(\mathbf{z}_k|\mathbf{x}') p_{k|k-1}(\mathbf{x}'|\mathbf{z}_{1:k-1})d\mathbf{x}'}. \quad (4)$$

Knowing the posterior $p_{k|k}(\mathbf{x}|\mathbf{z}_{1:k})$, one can compute the point estimates of the state, such as the expected a posteriori (EAP) estimate or the maximum a posteriori (MAP) estimate.

2.2. The possibilistic stochastic filter

Instead of random variables, we now consider *uncertain variable* [22] in order to enable imprecision of the probabilistic model to be considered (epistemic uncertainty). Uncertain variables can be described by outer measures of a certain form, however we consider the case where all the uncertainty is modelled as epistemic uncertainty, so that these outer measures simplify to possibility measures, as introduced in the seminal paper [10].

Let \mathbf{A} be a subset of the state space \mathcal{X} and let Π be a possibility measure associated with $\mathbf{x} \in \mathcal{X}$. Then the possibility measure of \mathbf{A} takes a value in the interval $[0, 1]$ representing the possibility that $\mathbf{x} \in \mathbf{A}$, and is defined as $\Pi(\mathbf{A}) = \sup_{\mathbf{x} \in \mathbf{A}} \pi(\mathbf{x})$, where $\pi(\mathbf{x})$ is the possibility function (or distribution [23]) corresponding to Π . The possibility function $\pi : \mathcal{X} \rightarrow [0, 1]$ is the primitive

object of possibility theory [23, 24], which assigns to each $\mathbf{x} \in \mathcal{X}$ a degree of possibility of being the true value of the state. It is normalised in the sense that $\sup_{\mathbf{x} \in \mathcal{X}} \pi(\mathbf{x}) = 1$.

Each possibility function can be seen as a membership function determining the fuzzy restriction of minimal specificity¹ about \mathbf{x} [10]. It can represent the two extremes of information specification, i.e. *complete knowledge* and *total ignorance*. Complete knowledge corresponds to the situation where it is known that the true value of the state is $\mathbf{x}_0 \in \mathcal{X}$. We can express this information as $\pi(\mathbf{x}_0) = 1$ and $\pi(\mathbf{x}) = 0$ for all $\mathbf{x} \in \mathcal{X}$ such that $\mathbf{x} \neq \mathbf{x}_0$. Complete ignorance is represented with $\pi(\mathbf{x}) = 1$ for all $\mathbf{x} \in \mathcal{X}$.

Note that a dual notion to *possibility*, referred to as *necessity*, can be introduced. A necessity of $\mathbf{A} \subseteq \mathcal{X}$ is defined as $N(\mathbf{A}) = \min_{\mathbf{x} \notin \mathbf{A}} [1 - \pi(\mathbf{x})]$. Duality of possibility and necessity can be expressed by $N(\mathbf{A}) = 1 - \Pi(\mathbf{A}^c)$, where \mathbf{A}^c is the complement of \mathbf{A} in \mathcal{X} . Possibility function π induces both Π and N , that is, knowing π is sufficient to calculate both. One can interpret the interval $[N, \Pi]$ as the probability interval, from *lower* to *upper* probability [24].

Any bounded PDF $p(\mathbf{x})$ can be transformed into a possibility function $\pi(\mathbf{x})$, and conversely, any integrable possibility function can be transformed into a PDF. An example of transformations is:

$$\pi(\mathbf{x}) = \frac{p(\mathbf{x})}{\sup_{\mathbf{x}' \in \mathcal{X}} p(\mathbf{x}')}, \quad (5)$$

$$p(\mathbf{x}) = \frac{\pi(\mathbf{x})}{\int_{\mathcal{X}} \pi(\mathbf{x}') d\mathbf{x}'}. \quad (6)$$

Other transformations are discussed in [24].

Considering that in majority of applications, both process noise and measurement noise are modelled by Gaussian distributions, we can focus on the Gaussian possibility function:

$$\bar{\mathcal{N}}(\mathbf{x}; \boldsymbol{\mu}, \mathbf{P}) = \exp \left(-\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^\top \mathbf{P}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right) \quad (7)$$

for some $\boldsymbol{\mu} \in \mathbb{R}^d$ and for some $d \times d$ positive definite matrix \mathbf{P} with real coefficients. With abuse of language, we will refer to $\boldsymbol{\mu}$ and \mathbf{P} as to the mean² and the covariance matrix of the Gaussian possibility function $\bar{\mathcal{N}}(\mathbf{x}; \boldsymbol{\mu}, \mathbf{P})$.

The goal of the possibilistic stochastic filter is to estimate sequentially the posterior possibility function $\pi_{k|k}(\mathbf{x}|\mathbf{z}_{1:k})$. Suppose the posterior possibility

¹In the sense that any hypothesis not known to be impossible cannot be ruled out.

²The possibilistic mean value has been defined as a closed interval [25], although other interpretations exist.

function at time $k-1$, $\pi_{k-1|k-1}(\mathbf{x}|\mathbf{z}_{1:k-1})$, is available. The prediction equation explains how to compute the possibility function of the state at time k using the transitional possibility function $\rho_{k|k-1}(\mathbf{x}_k|\mathbf{x}_{k-1})$, and is given by [17]:

$$\pi_{k|k-1}(\mathbf{x}|\mathbf{z}_{1:k-1}) = \sup_{\mathbf{x}' \in \mathcal{X}} \rho_{k|k-1}(\mathbf{x}|\mathbf{x}') \pi_{k-1|k-1}(\mathbf{x}'|\mathbf{z}_{1:k-1}). \quad (8)$$

The transitional possibility function $\rho_{k|k-1}(\mathbf{x}|\mathbf{x}')$ can be, for example, obtained from $p_{k|k-1}(\mathbf{x}|\mathbf{x}')$ using transformation (5). Note that (8) is an analogue of the Chapman-Kolmogorov equation of the standard Bayes filter (3). The two expressions differ only as follows: (i) the integral in (3) is replaced by the supremum in (8); (ii) the PDFs in (3) are replaced with possibility functions in (8).

The update step of the possibility filter “corrects” the predicted possibility function $\pi_{k|k-1}(\mathbf{x}|\mathbf{z}_{1:k-1})$ using the information contained in the new measurement \mathbf{z}_k . The update equation is given by [17]:

$$\pi_{k|k}(\mathbf{x}|\mathbf{z}_{1:k}) = \frac{g(\mathbf{z}_k|\mathbf{x}) \pi_{k|k-1}(\mathbf{x}|\mathbf{z}_{1:k-1})}{\sup_{\mathbf{x} \in \mathcal{X}} [g(\mathbf{z}_k|\mathbf{x}) \pi_{k|k-1}(\mathbf{x}|\mathbf{z}_{1:k-1})]}, \quad (9)$$

where $g(\mathbf{z}|\mathbf{x})$ represents the likelihood function expressed as a possibility function. Note that (9) is an analogue of the Bayes’ update (4). Again, the two expressions differ only in the following: (i) the supremum replaces the integral; (ii) the probability distributions are replaced with the possibility functions.

It was demonstrated in [17] that the predicted and posterior mean and variance in the recursion (8)-(9) are the ones of the Kalman filter in the linear-Gaussian case. In the non-linear case, a comparison (in the context of bearings-only tracking) between the standard Bayes filter and the possibilistic stochastic filter [20] revealed that: (a) in the absence of a model mismatch, the two filters perform identically; (b) in the presence of a (dynamic or observation) model mismatch, the possibilistic filter consistently results in a lower probability of divergence, which indicates a more robust performance.

3. Formulation of the PBF

The Bernoulli filter is a Bayes-type filter, designed for dynamic systems that are capable of switching on and off. In the target tracking context, this means that the target can appear/disappear from the region of interest. The Bernoulli filter estimates recursively the posterior probability of target presence, in addition to the posterior density of the target state. By monitoring the probability of target presence, one can effectively *detect* when the target appears or disappears.

Derivation of the Bernoulli filter³ [15, 14] was carried out in the framework of random finite set (RFS) theory. By replacing the concept of a random variable with an uncertain variable [17], and a concept of a RFS with an uncertain finite set (UFS), referred to as an *uncertain counting measure* in [27], we will next derive a possibilistic analogue of the standard Bernoulli filter.

3.1. Uncertain finite sets

An UFS $\mathbf{X} \in \mathcal{F}(\mathcal{X})$ is an uncertain variable that takes values as unordered finite sets [15] on \mathcal{X} . Here $\mathcal{F}(\mathcal{X})$ denotes the set of all finite subsets of \mathcal{X} . Both cardinality and the spatial distribution of the elements of \mathcal{X} are uncertain. An UFS is completely characterised by:

- (i) a cardinality distribution, modelled by a discrete possibility function $c(n) = \Pi\{|\mathbf{X}| = n\}$, where $n \in \mathbb{N}_0$, $\Pi\{A\}$ is a possibility of event A . Due to normalisation, $\max_{n \geq 0} c(n) = 1$;
- (ii) a family of symmetric possibility functions $\pi_n(\mathbf{x}_1, \dots, \mathbf{x}_n)$, with $n \in \mathbb{N}_0$, $\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathcal{X}$. Due to normalisation, $\sup_{\mathbf{x}_1, \dots, \mathbf{x}_n} \pi_n(\mathbf{x}_1, \dots, \mathbf{x}_n) = 1$.

The possibility function of an UFS is defined as:

$$f(\{\mathbf{x}_1, \dots, \mathbf{x}_n\}) = c(n)\pi_n(\mathbf{x}_1, \dots, \mathbf{x}_n) \quad (10)$$

for $n \in \mathbb{N}_0$. For a special class of UFSs, with a property that points $\mathbf{x}_1, \dots, \mathbf{x}_n$ are independently described by a single possibility function $\pi(\mathbf{x})$, the possibility function can be expressed as:

$$f(\mathbf{X}) = c(|\mathbf{X}|) \prod_{\mathbf{x} \in \mathbf{X}} \pi(\mathbf{x}). \quad (11)$$

Example. A Poisson discrete possibility function, with a parameter $\lambda > 0$, can be obtained simply as follows:

$$c(n) = \frac{1}{\beta} \frac{\lambda^n e^{-\lambda}}{n!} \quad (12)$$

where

$$\beta = \max_{n \geq 0} \frac{\lambda^n e^{-\lambda}}{n!} = \frac{\lambda^{\lfloor \lambda \rfloor} e^{-\lfloor \lambda \rfloor}}{\lfloor \lambda \rfloor!} \quad (13)$$

because $\lfloor \lambda \rfloor$ is the mode of this distribution. Fig. 1 illustrates the Poisson probability (red) and possibility function (blue), for parameter $\lambda = 4.2$. ■

³Which generalises the integrated probabilistic data association (IPDA) filter of [26].

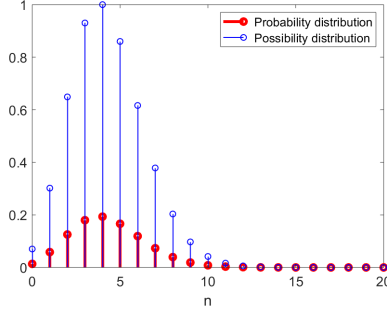


Figure 1: Illustration of the Poisson probability distribution and possibility function (parameter $\lambda = 4.2$)

As a possibility function on $\mathcal{F}(\mathcal{X})$, $f(\cdot)$ has a supremum equal to one on this domain. This is satisfied by construction since

$$\begin{aligned} \sup_{\mathbf{X} \in \mathcal{F}(\mathcal{X})} f(\mathbf{X}) &= \max \left\{ c(0), c(1) \sup_{\mathbf{x}_1 \in \mathcal{X}} \pi_1(\mathbf{x}_1), c(2) \sup_{(\mathbf{x}_1, \mathbf{x}_2) \in \mathcal{X} \times \mathcal{X}} \pi_2(\mathbf{x}_1, \mathbf{x}_2), \dots \right\} \\ &= \max_{n \geq 0} c(n) = 1. \end{aligned}$$

A Bernoulli UFS \mathbf{X} is a set whose cardinality can be 0 or 1. Its UFS possibility function is then:

$$f(\mathbf{X}) = \begin{cases} q^0, & \text{if } \mathbf{X} = \emptyset \\ q^1 \pi(\mathbf{x}), & \text{if } \mathbf{X} = \{\mathbf{x}\}, \end{cases} \quad (14)$$

where

- $q^0 = c(0)$ is the possibility that $\mathbf{X} = \emptyset$,
- $q^1 = c(1)$ is the possibility that $|\mathbf{X}| = 1$,
- $\pi(\mathbf{x})$ is the possibility function over \mathcal{X} , given that $|\mathbf{X}| = 1$.

Due to normalisation, it holds that $\max\{q^0, q^1 \sup_{\mathbf{x} \in \mathcal{X}} \pi(\mathbf{x})\} = 1$. By definition, it also holds that $\sup_{\mathbf{x} \in \mathcal{X}} \pi(\mathbf{x}) = 1$ and thus the normalisation constraint is simply $\max\{q^0, q^1\} = 1$.

3.2. Dynamic model and the prediction step

If the target is present at time $k-1$ and k , then its dynamics is characterised by the transitional possibility function $\rho_{k|k-1}(\mathbf{x}|\mathbf{x}')$, introduced earlier, see eq. (8). In order to model object appearance and disappearance, it is convenient

to introduce a binary uncertain variable $\epsilon_k \in \{0, 1\}$ referred to as the target *presence*. The convention is that $\epsilon_k = 1$ means that the target is present at k (and conversely, $\epsilon_k = 0$ means that it is absent). Let the dynamics of ϵ_k be modelled by a two-state Markov chain with a (time-invariant) transitional possibility matrix (TPM)

$$\mathcal{T} = \begin{bmatrix} \tau_{00} & \tau_{01} \\ \tau_{10} & \tau_{11} \end{bmatrix} \quad (15)$$

where τ_{ij} is the possibility of transition from $\epsilon_{k-1} = i$ to $\epsilon_k = j$, for $i, j \in \{0, 1\}$. Due to normalisation, $\max\{\tau_{i0}, \tau_{i1}\} = 1$, for $i = 0, 1$. We also need to specify the initial possibility (at $k = 0$) of target absence and presence, i.e. q_0^0 and q_0^1 , respectively, such that $\max\{q_0^0, q_0^1\} = 1$. If a target appears at time k , the possibility function describing the possibility of its appearance over the state space \mathcal{X} is denoted by $b_{k|k-1}(\mathbf{x})$.

Next we introduce the transitional possibility function of a Bernoulli UFS, from time $k - 1$ to k . Let us denote this possibility function as $\phi_{k|k-1}(\mathbf{X}|\mathbf{X}')$. If the target was not present at time $k - 1$, then

$$\phi_{k|k-1}(\mathbf{X}|\emptyset) = \begin{cases} \tau_{00}, & \text{if } \mathbf{X}_k = \emptyset \\ \tau_{01}b_{k|k-1}(\mathbf{x}), & \text{if } \mathbf{X} = \{\mathbf{x}\}. \end{cases} \quad (16)$$

If the target was present at time $k - 1$ and in state $\mathbf{x}' \in \mathcal{X}$, then

$$\phi_{k|k-1}(\mathbf{X}|\{\mathbf{x}'\}) = \begin{cases} \tau_{10}, & \text{if } \mathbf{X}_k = \emptyset \\ \tau_{11}\rho_{k|k-1}(\mathbf{x}|\mathbf{x}'), & \text{if } \mathbf{X} = \{\mathbf{x}\}. \end{cases} \quad (17)$$

Let the set of measurements at time k be denoted \mathbf{Z}_k . This set may contain false detections, while the true target detection may be missing due to imperfect target detection process. The target state at time k is represented by a Bernoulli UFS \mathbf{X}_k . The uncertainty of the target state at k is represented by the posterior possibility function $f_{k|k}(\mathbf{X}|\mathbf{Z}_{1:k})$, where $\mathbf{Z}_{1:k} \stackrel{\text{abbr}}{=} \mathbf{Z}_1, \dots, \mathbf{Z}_k$. In order to simplify notation, we will use abbreviation: $f_{k|k}(\mathbf{X}|\mathbf{Z}_{1:k}) \stackrel{\text{abbr}}{=} f_{k|k}(\mathbf{X})$.

The prediction equation of the PBF is then as follows. Suppose the posterior possibility function of a Bernoulli UFS at $k - 1$, that is $f_{k-1|k-1}(\mathbf{X})$, is available and expressed according to (14) as:

$$f_{k-1|k-1}(\mathbf{X}) = \begin{cases} q_{k-1|k-1}^0, & \text{if } \mathbf{X} = \emptyset \\ q_{k-1|k-1}^1 \pi_{k-1|k-1}(\mathbf{x}), & \text{if } \mathbf{X} = \{\mathbf{x}\}. \end{cases} \quad (18)$$

Prediction of this possibility function to time k is carried out using the transitional possibility function $\phi_{k|k-1}(\mathbf{X}|\mathbf{X}')$. Analogue to (8), it can be written

as:

$$f_{k|k-1}(\mathbf{X}) = \sup_{\mathbf{X}' \in \mathcal{F}(\mathcal{X})} [\phi_{k|k-1}(\mathbf{X}|\mathbf{X}') f_{k-1|k-1}(\mathbf{X}')] . \quad (19)$$

When we work out (19) for $\mathbf{X} = \emptyset$, we obtain (see Appendix A.1) the prediction equation for the possibility of target being absent at time k :

$$q_{k|k-1}^0 = \max \left\{ \tau_{00} q_{k-1|k-1}^0, \tau_{10} q_{k-1|k-1}^1 \right\} . \quad (20)$$

By solving (19) for $\mathbf{X} = \{\mathbf{x}\}$, we obtain (see Appendix A.1) the predicted possibility of target being present:

$$q_{k|k-1}^1 = \max \left\{ \tau_{01} q_{k-1|k-1}^0, \tau_{11} q_{k-1|k-1}^1 \right\} \quad (21)$$

and the predicted possibility function over \mathcal{X} :

$$\pi_{k|k-1}(\mathbf{x}) = \frac{1}{q_{k|k-1}^1} \max \left\{ \tau_{01} q_{k-1|k-1}^0 b_{k|k-1}(\mathbf{x}), \right. \\ \left. \tau_{11} q_{k-1|k-1}^1 \sup_{\mathbf{x}' \in \mathcal{X}} [\rho_{k|k-1}(\mathbf{x}|\mathbf{x}') \pi_{k-1|k-1}(\mathbf{x}')] \right\} . \quad (22)$$

It can be easily verified (see Appendix A.1) that $\max\{q_{k|k-1}^0, q_{k|k-1}^1\} = 1$ and that $\sup_{\mathbf{x} \in \mathcal{X}} \pi_{k|k-1}(\mathbf{x}) = 1$. Also, if the target is present and there are no presence/absence transitions, that is $q_{k-1|k-1}^1 = 1$, $\tau_{01} = 0$ and $\tau_{11} = 1$, then (22) reduces to (8).

3.3. Measurement model and the update step

Let us assume that M sensors are simultaneously collecting and reporting target measurements. At time k , sensor $i \in \{1, \dots, M\}$ reports a (finite) set of measurements (detections) $\mathbf{Z}_k^{(i)} = \{\mathbf{z}_{k,1}^{(i)}, \mathbf{z}_{k,2}^{(i)}, \dots, \mathbf{z}_{k,m_k^i}^{(i)}\}$, where both the cardinality of the set $m_k^i \in \mathbb{N}_0$, and the location of the points in $\mathbf{Z}_k^{(i)}$ in the measurement space $\mathcal{Z} \subset \mathbb{R}^{n_z}$, are uncertain.

The sensor detector is imperfect in the sense that: (i) the true target originated measurement may not be present in $\mathbf{Z}_k^{(i)}$, and (ii) $\mathbf{Z}_k^{(i)}$ may contain false detections. Suppose the target at time k is in the state \mathbf{x} and is detected by receiver $i \in \{1, \dots, M\}$, resulting in a measurement $\mathbf{z} \in \mathbf{Z}_k^{(i)}$. The likelihood function of i th sensor, $g_i(\mathbf{z}|\mathbf{x})$, was introduced in (9). It is expressed as a possibility function over \mathcal{Z} because it specifies the uncertain relationship between the measurement and the target state.

In accordance with (9), the update step of the Bernoulli filter consists of two stages:

1. The predicted Bernoulli possibility function $f_{k|k-1}(\mathbf{X})$ is multiplied with the likelihood function for all measurement sets $\mathbf{Z}_k \stackrel{\text{abbr}}{=} \mathbf{Z}_k^{(1)}, \dots, \mathbf{Z}_k^{(M)}$, given that the target is in the state \mathbf{X} ; this likelihood is denoted $\varphi(\mathbf{Z}_k|\mathbf{X})$;
2. Normalisation of the product computed in stage 1.

Mathematically, the update step can be expressed as:

$$f_{k|k}(\mathbf{X}) = \frac{\varphi(\mathbf{Z}_k|\mathbf{X}) f_{k|k-1}(\mathbf{X})}{\sup_{\mathbf{X} \in \mathcal{F}(\mathcal{X})} [\varphi(\mathbf{Z}_k|\mathbf{X}) f_{k|k-1}(\mathbf{X})]} \quad (23)$$

where $f_{k|k-1}(\mathbf{X})$, specified by the triplet $(q_{k|k-1}^0, q_{k|k-1}^1, \pi_{k|k-1}(\mathbf{x}))$, is in the form (14). The terms in the triplet can be computed via (20), (21) and (22), respectively.

Next we derive the likelihood function $\varphi(\mathbf{Z}_k|\mathbf{X})$. Assuming the sensors are independent, we can express this likelihood as a product [28, Def.4]:

$$\varphi(\mathbf{Z}_k|\mathbf{X}) = \prod_{i=1}^M \varphi_i(\mathbf{Z}_k^{(i)}|\mathbf{X}). \quad (24)$$

Note that an UFS $\mathbf{Z}_k^{(i)}$ (collected by i th sensor at time k) can be seen as a union of two independent UFSs $\mathbf{Z}_k^{(i)} = \mathbf{C}_k^{(i)} \cup \mathbf{W}_k^{(i)}$, where $\mathbf{C}_k^{(i)}$ is the UFS of false detections and $\mathbf{W}_k^{(i)}$ is a Bernoulli UFS modeling the detection from the target [15]. The target may not be detected, and hence the possibility function of $\mathbf{W}_k^{(i)}$, given that target state is $\mathbf{X} = \{\mathbf{x}\}$, according to (14) can be expressed as:

$$\eta_i(\mathbf{W}_k^{(i)}|\{\mathbf{x}\}) = \begin{cases} d_i^0, & \text{if } \mathbf{W}_k^{(i)} = \emptyset \\ d_i^1 g_i(\mathbf{z}|\mathbf{x}), & \text{if } \mathbf{W}_k^{(i)} = \{\mathbf{z}\}, \end{cases} \quad (25)$$

where d_i^0 and d_i^1 denote the possibility of target non-detection and detection by sensor i , respectively. Due to normalisation, $\max\{d_i^0, d_i^1\} = 1$.

When the target is absent (i.e. $\mathbf{X} = \emptyset$), the target originated detection is also absent (i.e. $\mathbf{W}_k^{(i)} = \emptyset$), hence the possibility function of $\mathbf{Z}_k^{(i)}$ equals the possibility function of false detections only, given in the form of (11):

$$\varphi_i(\mathbf{Z}_k^{(i)}|\emptyset) = \nu_i(|\mathbf{Z}_k^{(i)}|) \prod_{\mathbf{z} \in \mathbf{Z}_k^{(i)}} \mu_i(\mathbf{z}) \stackrel{\text{abbr}}{=} \kappa_i(\mathbf{Z}_k^{(i)}). \quad (26)$$

Here $\nu_i(n)$ is a discrete possibility function of the count of clutter measurements and $\mu_i(\mathbf{z})$ is the possibility function on \mathcal{Z} describing the clutter.

If the target is present, the likelihood function $\varphi_i(\mathbf{Z}_k^{(i)}|\mathbf{X})$ can be expressed

as follows:

$$\varphi_i(\mathbf{Z}_k^{(i)}|\{\mathbf{x}\}) = \max_{\mathbf{W} \subseteq \mathbf{Z}_k^{(i)}} \left[\eta_i(\mathbf{W}|\{\mathbf{x}\}) \kappa_i(\mathbf{Z}_k^{(i)} \setminus \mathbf{W}) \right] \quad (27a)$$

$$= \max \left\{ \eta_i(\emptyset|\{\mathbf{x}\}) \kappa_i(\mathbf{Z}_k^{(i)}), \max_{\mathbf{z} \in \mathbf{Z}_k^{(i)}} \left[\eta_i(\{\mathbf{z}\}|\{\mathbf{x}\}) \kappa_i(\mathbf{Z}_k^{(i)} \setminus \{\mathbf{z}\}) \right] \right\} \quad (27b)$$

$$= \max \left\{ d_i^0 \kappa_i(\mathbf{Z}_k^{(i)}), \max_{\mathbf{z} \in \mathbf{Z}_k^{(i)}} \left[d_i^1 g_i(\mathbf{z}|\mathbf{x}) \kappa_i(\mathbf{Z}_k^{(i)} \setminus \{\mathbf{z}\}) \right] \right\} \quad (27c)$$

where the sign \setminus in (27b) and (27c) denotes the set-minus operation. Note that (27a) represents the convolution formula [15] for UFSs, while (27b) is its simplification because $\eta_i(\mathbf{W}|\{\mathbf{x}\}) = 0$ whenever $|\mathbf{W}| > 1$.

Next we substitute expressions for $\varphi_i(\mathbf{Z}_k^{(i)}|\emptyset)$ and $\varphi_i(\mathbf{Z}_k^{(i)}|\{\mathbf{x}\})$, given by (26) and (27c), respectively, in the update equation (23). This leads to the multi-sensor update equations of the PBF (full derivation is given in Appendix A.2). The posterior possibility of target absence and presence are given by:

$$q_{k|k}^0 = \frac{q_{k|k-1}^0}{\max \{q_{k|k-1}^0, \alpha q_{k|k-1}^1\}} \quad (28)$$

$$q_{k|k}^1 = \frac{\alpha q_{k|k-1}^1}{\max \{q_{k|k-1}^0, \alpha q_{k|k-1}^1\}} \quad (29)$$

respectively, where

$$\alpha = \prod_{i=1}^M R_i(\mathbf{Z}_k^{(i)}) \quad (30)$$

and

$$R_i(\mathbf{Z}) = \max \left\{ d_i^0, d_i^1 \max_{\mathbf{z} \in \mathbf{Z}} \left[\frac{\kappa_i(\mathbf{Z} \setminus \{\mathbf{z}\})}{\kappa_i(\mathbf{Z})} \sup_{\mathbf{x} \in \mathcal{X}} [g_i(\mathbf{z}|\mathbf{x}) \pi_{k|k-1}(\mathbf{x})] \right] \right\}. \quad (31)$$

The update equation for the spatial possibility function is:

$$\pi_{k|k}(\mathbf{x}) = \prod_{i=1}^M \frac{L_i(\mathbf{Z}_k^{(i)}|\mathbf{x})}{R_i(\mathbf{Z}_k^{(i)})} \pi_{k|k-1}(\mathbf{x}) \quad (32)$$

where

$$L_i(\mathbf{Z}|\mathbf{x}) = \max \left\{ d_i^0, d_i^1 \max_{\mathbf{z} \in \mathbf{Z}} \left[\frac{\kappa_i(\mathbf{Z} \setminus \{\mathbf{z}\})}{\kappa_i(\mathbf{Z})} g_i(\mathbf{z}|\mathbf{x}) \right] \right\}. \quad (33)$$

It can be easily verified that $\max\{q_{k|k}^0, q_{k|k}^1\} = 1$ and that $\sup_{\mathbf{x} \in \mathcal{X}} \pi_{k|k}(\mathbf{x}) = 1$. Furthermore, if there is no clutter, if the possibility of detection is $d_i^1 = 1$ and the possibility of non-detection is $d_i^0 = 0$, then $\mathbf{Z}_k^{(i)}$ contains only one measurement, which must be due to the target. This leads to $q_{k|k}^1 = 1$ and $q_{k|k}^0 = 0$, while (32) reduces to (9).

4. Application: target tracking using multi-static Doppler shifts

4.1. Problem description

The problem description follows [29]. The state of the moving target in the two-dimensional surveillance area at time t_k is represented by the state vector

$$\mathbf{x}_k = [x_k \quad \dot{x}_k \quad y_k \quad \dot{y}_k]^\top, \quad (34)$$

where \top denotes the matrix transpose and k is the discrete-time index. Target position and velocity vector are denoted $\mathbf{p}_k = [x_k \quad y_k]^\top$ and $\dot{\mathbf{p}}_k = [\dot{x}_k \quad \dot{y}_k]^\top$, respectively. Uncertain target motion is described by the possibility function

$$\rho_{k|k-1}(\mathbf{x}|\mathbf{x}') = \bar{\mathcal{N}}(\mathbf{x}; \mathbf{F}\mathbf{x}', \mathbf{Q}) \quad (35)$$

where

$$\mathbf{F} = \mathbf{I}_2 \otimes \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}, \quad \mathbf{Q} = \mathbf{I}_2 \otimes q \begin{bmatrix} \frac{T^3}{3} & \frac{T^2}{2} \\ \frac{T^2}{2} & T \end{bmatrix}. \quad (36)$$

Here \otimes is the Kroneker product, $T = t_k - t_{k-1}$ is the constant sampling interval, q is the noise intensity and \mathbf{I}_2 is the 2×2 identity matrix.

Target tracking is carried out using Doppler shifts measured at spatially distributed receivers, as illustrated in Fig. 2. A transmitter T at a known position $\mathbf{t} = [x_0 \quad y_0]^\top$, illuminates the target at location \mathbf{p}_k by a sinusoidal waveform of a known carrier frequency f_c . The receivers in Fig. 2 are denoted by R_i , $i \in \{1, \dots, M\}$.

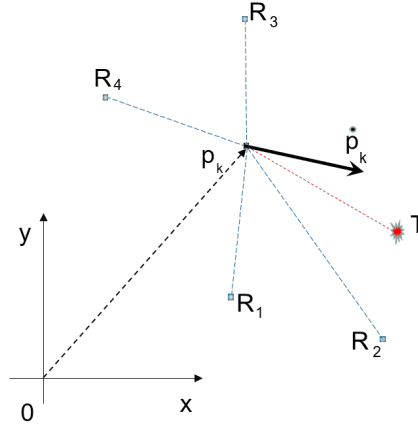


Figure 2: Multi-static Doppler-only surveillance network: T - transmitter; R_i - i th receiver; \mathbf{p}_k and $\dot{\mathbf{p}}_k$ are target position and velocity vector, respectively.

If the target at time k is in the state \mathbf{x}_k , and is detected by receiver $i \in \{1, \dots, M\}$ placed at a known location $\mathbf{r}_i = [x_i \ y_i]^\top$, then the receiver will report a Doppler-shift (measurement) $\mathbf{z}_k \in \mathcal{Z} = [-f_0, +f_0]$, described by the likelihood function expressed as a possibility function:

$$g_i(\mathbf{z}_k|\mathbf{x}_k) = \bar{\mathcal{N}}(\mathbf{z}_k; h_i(\mathbf{x}_k), \sigma_i^2). \quad (37)$$

The frequency f_0 (the maximum value of the Doppler shift), is assumed known. The nonlinear measurement function $h_i(\cdot)$ in (37) represents the true value of Doppler shift and is given by [29]:

$$h_i(\mathbf{x}_k) = -\dot{\mathbf{p}}_k^\top \left[\frac{\mathbf{p}_k - \mathbf{r}_i}{\|\mathbf{p}_k - \mathbf{r}_i\|} + \frac{\mathbf{p}_k - \mathbf{t}}{\|\mathbf{p}_k - \mathbf{t}\|} \right] \frac{f_c}{c} \quad (38)$$

where c is the speed of light. In accordance with the comment below (7), we refer to σ_i^2 as to the variance of the Gaussian possibility function (37).

The distribution of false detections over the measurement space \mathcal{Z} is assumed time invariant and independent of the target state. The number of false detections per scan is assumed to be Poisson distributed, with the mean value λ_i for receiver i .

Target originated Doppler shift measurement is detected by receiver i with the probability of detection $P_d^i(\mathbf{x}_k) \leq 1$. In general, the probability of detection is a function of the distance between the target at position \mathbf{p}_k and the receiver at location \mathbf{r}_i , i.e. $d_k^i = \|\mathbf{p}_k - \mathbf{r}_i\|$. For illustration, we adopt a formula $P_d^i(\mathbf{x}_k) = \exp[-(d_k^i/\beta)^4]$, where for d_k^i in meters $\beta = 12 \cdot 10^3$ (also in meters). Then, P_d^i is a monotonically decreasing function of distance, equaling 1 at $d_k^i = 0$, and dropping to 1/2 at $d_k^i \approx 8320\text{m}$. In simulations, the Doppler-shift measurements were generated using this formula for the probability of detection.

We argue that in practice, the probability of detection available to the filter, cannot be as precise as specified above, because in reality it would depend on the signal to noise ratio, which is unknown. For the same reasons, learning the functional form of P_d^i , $i = 1, \dots, M$, from the data would also be fairly difficult. The main advantage of the PBF over the standard Bernoulli filter [14] (formulated using the probability distributions and based on precise specification of all parameters, including P_d^i), in this application would be that it needs only a partial knowledge of P_d^i , via d_i^0 and d_i^1 , see Sec. 3.3. The pair (d_i^0, d_i^1) , where $\max\{d_i^0, d_i^1\} = 1$, effectively defines the interval of detection probability⁴, that is

⁴The possibility of detection is the upper probability of detection, while the lower probability of detection is the *necessity*, defined as one minus the possibility of the complement of detection (i.e. non-detection), see Sec. 2.2.

$P_d^i \in [1 - d_i^0, d_i^1]$. In the case of the total ignorance about P_d^i , we set $d_i^0 = d_i^1 = 1$.

4.2. Implementation of the PBF

We developed a computer implementation of the PBF based on an adaptation of the SMC method. Note that one cannot sample directly from a possibility function [18]. Instead, for a given possibility function π , samples must be drawn from a PDF p which is induced by π . While there is an infinite number of ways one can construct p from π (one being (6)), the natural solution is the one that results in the least informative p . Practical details of an SMC method for possibility functions can be found in [18].

Random samples or particles are propagated over time only as the support points of the posterior possibility function $\pi_{k|k}(\mathbf{x})$, mimicking an *adaptive grid* over the state space \mathcal{X} . The weights, associated with these particles, are computed using the PBF equations (22) and (32). Prediction of the posterior possibilities $q_{k|k}^0$ and $q_{k|k}^1$ is based on the straightforward application of equations (20) and (21). In the update step, equations (28) and (29), the SMC representation of $\pi_{k|k-1}(\mathbf{x})$ is used in the computation of α via (30) and (31).

The point estimate $\hat{\mathbf{x}}_{k|k}$ is computed as a weighted mean of the particles approximating $\pi_{k|k}(\mathbf{x})$.

4.3. Numerical results

The following values were used in simulations. The location of the transmitter: $\mathbf{t} = [0\text{m}, 0\text{m}]^\top$; $M = 5$ receivers, placed at $\mathbf{r}_1 = [-8000\text{m}, -1000\text{m}]^\top$, $\mathbf{r}_2 = [-9000\text{m}, 11000\text{m}]^\top$, $\mathbf{r}_3 = [-2000\text{m}, 2000\text{m}]^\top$, $\mathbf{r}_4 = [1000\text{m}, 11000\text{m}]^\top$ and $\mathbf{r}_5 = [9000\text{m}, 9000\text{m}]^\top$. Other parameters were: $f_c = 900$ MHz, $T = 2$ s, $q = 0.1$, $f_0 = 200$ Hz, $\sigma_i = 2.5$ Hz and $\lambda_i = 0.5$ for $i = 1, \dots, M$. False detections were uniformly distributed across \mathcal{Z} . The initial target state (at $k = 1$): $\mathbf{x}_1 = [-4000\text{m}, 30\text{m/s}, 7000\text{m}, -12\text{m/s}]^\top$. The observation interval is 140 seconds (i.e. $k = 1, \dots, 70$).

The parameters used in the SMC approximation of the PBF were as follows. The number of particles used was 10000. The target birth distribution $b_{k|k-1}(\mathbf{x}) = \tilde{\mathcal{N}}(\mathbf{x}; \boldsymbol{\mu}_b, \mathbf{P}_b)$, where the mean is $\boldsymbol{\mu}_b = [0 \ 0 \ 0 \ 0]^\top$, that is placed at the location of the transmitter, with zero target velocity. The covariance matrix was set to $\mathbf{P}_b = \text{diag}[(4\text{km})^2 \ (30\text{m/s})^2 \ (4\text{km})^2 \ (30\text{m/s})^2]$. Furthermore, $\tau_{00} = \tau_{11} = 1$ and $\tau_{01} = \tau_{10} = 0.01$. The initial possibilities of target presence and absence were set to $q_{1|1}^1 = 1$ and $q_{1|1}^0 = 1$, respectively. This corresponds to the total ignorance about target presence, i.e. its probability is in the interval

$[0, 1]$. A track is confirmed when the difference $q_{k|k}^1 - q_{k|k}^0 \geq 0.5$, corresponding to the probability of target presence being in the interval $[0.5, 1]$.

A single run of the PBF for the described simulation scenario is available, as an avi movie, in the Supplementary material. Fig. 3 shows a typical set of Doppler-frequency measurements Z_1, Z_2, \dots, Z_{70} obtained during a single run. Notice the effect of time-varying probability of detection and false Doppler measurements.

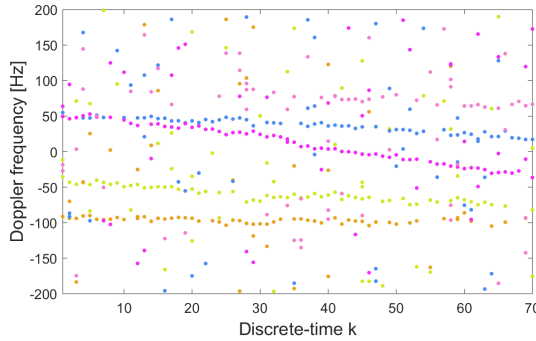


Figure 3: A typical set of Doppler-shift measurements over time (the same coloured measurements originate from the same receiver)

Joint detection and tracking performance is measured using the mean optimal sub-pattern assignment (OSPA) metric [30]. The parameters used in the computation of the OSPA metric were $p = 1$ and $c = 5000\text{m}$. Fig. 4 presents the mean OSPA errors (in position) [30] obtained by averaging over 500 Monte Carlo runs of the PBF. The three OSPA error curves shown in Fig. 4 correspond to the three different intervals of probability of detection used in the PBF: green line for $P_d^i \in [0.4, 1.0]$, red line for $P_d^i \in [0.6, 1.0]$ and the blue line for $P_d^i \in [0.8, 1.0]$. Fig. 4 demonstrates that the PBF works. The best performance is achieved for $P_d^i \in [0.6, 1.0]$, because this interval captures most accurately the spatio-temporal variation of the probability of detection for all five receivers. By setting $P_d^i \in [0.8, 1.0]$, the track is established quicker, however, the track maintenance is less reliable, resulting in occasional breaks in the track. Finally, with $P_d^i \in [0.4, 1.0]$, the track is not established in about 15% of the runs.

Next we compare the standard (probabilistic) Bernoulli filter with the PBF, using the same simulation setup and parameters. The only difference is in the specification of the probability of detection: for the PBF, $P_d \in [0.6, 1.0]$; for the standard Bernoulli filter, because we cannot specify the probability interval,

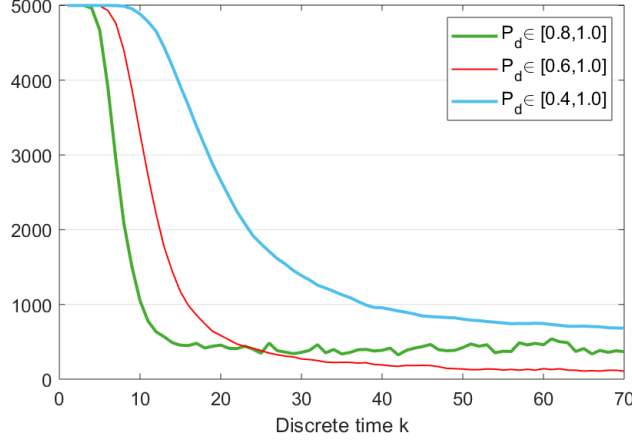


Figure 4: Mean OSPA error (position) for different specifications of the detection probability intervals in the PBF

we choose the mid-point of the interval, that is $P_d = 0.8$. Fig. 5 displays the mean OSPA error for the two Bernoulli filters. The PBF is more cautious and establishes the track with some delay in comparison with the standard Bernoulli filter. However, the established track is significantly more accurate: after $k = 25$, its mean OSPA error is at least twice lower. In summary, in the considered example, characterised by partially known (due to spatio-temporal variations) probability of detection, the PBF is a more reliable alternative to the standard Bernoulli filter.

5. Summary

To our best knowledge, the paper presented the first target tracking algorithm completely derived in the framework of possibility theory. The algorithm, referred to as the possibilistic Bernoulli filter, is characterised by the Bayesian filtering style analytic expressions for prediction and update. The motivation for using the possibility functions, instead of the probabilistic framework, is a more generalised representation of uncertainty, capable of handling, in a rigorous mathematical manner, the situations of ignorance or partial knowledge. The PBF has been demonstrated in the context of an application, where the true (but unknown) probability of detection was varying across the space and time. Compared to the standard (probabilistic) Bernoulli filter, the PBF is characterised by more cautious, but significantly more accurate performance. Finally, the computation time of the two Bernoulli filters is practically the same.

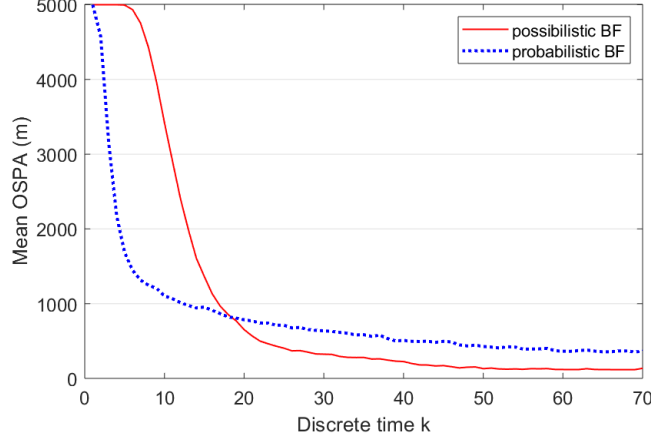


Figure 5: The standard (probabilistic) Bernoulli filter using $P_d = 0.8$ versus the PBF using $P_d \in [0.6, 1.0]$: the mean OSPA error curves (position).

Future research will consider theoretical formulations of other tracking algorithms in the framework of possibility theory.

Appendix A. Derivations

Appendix A.1. Derivation of prediction equations in Sec. 3.2

First we derive equation (20). Let us start with (19), i.e.

$$\begin{aligned}
 f_{k|k-1}(\mathbf{X}) &= \sup_{\mathbf{X}' \in \mathcal{F}(\mathcal{X})} [\phi_{k|k-1}(\mathbf{X}|\mathbf{X}') f_{k-1|k-1}(\mathbf{X}')] \\
 &= \max \left\{ \phi_{k|k-1}(\mathbf{X}|\emptyset) f_{k-1|k-1}(\emptyset), \right. \\
 &\quad \left. \sup_{\mathbf{x}' \in \mathcal{X}} \phi_{k|k-1}(\mathbf{X}|\{\mathbf{x}'\}) f_{k-1|k-1}(\{\mathbf{x}'\}) \right\} \quad (\text{A.1})
 \end{aligned}$$

For $\mathbf{X} = \emptyset$ we have:

$$\begin{aligned}
 f_{k|k-1}(\emptyset) &= \max \left\{ \phi_{k|k-1}(\emptyset|\emptyset) f_{k-1|k-1}(\emptyset), \sup_{\mathbf{x}' \in \mathcal{X}} \phi_{k|k-1}(\emptyset|\{\mathbf{x}'\}) f_{k-1|k-1}(\{\mathbf{x}'\}) \right\} \\
 &= \max \left\{ \tau_{00} q_{k-1|k-1}^0, \tau_{10} q_{k-1|k-1}^1 \sup_{\mathbf{x}' \in \mathcal{X}} \pi_{k-1|k-1}(\mathbf{x}') \right\} \quad (\text{A.2})
 \end{aligned}$$

Note that $\sup_{\mathbf{x}' \in \mathcal{X}} \pi_{k-1|k-1}(\mathbf{x}')$, which features on the right-hand side of (A.2), equals to 1 due to normalisation. Furthermore, since \mathbf{X} is a Bernoulli UFS, $f_{k|k-1}(\mathbf{X})$ can be expressed in form (14), i.e. as

$$f_{k|k-1}(\mathbf{X}) = \begin{cases} q_{k|k-1}^0, & \text{if } \mathbf{X} = \emptyset \\ q_{k|k-1}^1 \pi_{k|k-1}(\mathbf{x}), & \text{if } \mathbf{X} = \{\mathbf{x}\}. \end{cases} \quad (\text{A.3})$$

Hence $f_{k|k-1}(\emptyset)$, which appears on the left hand side of (A.2), represents the predicted possibility that the target is absent, i.e. $q_{k|k-1}^0$. Then from (A.2) follows (20), i.e.

$$q_{k|k-1}^0 = \max \left\{ \tau_{00} q_{k-1|k-1}^0, \tau_{10} q_{k-1|k-1}^1 \right\}. \quad (\text{A.4})$$

Next we derive equations (21) and (22). First we solve (A.1) for $\mathbf{X} = \{\mathbf{x}\}$:

$$\begin{aligned} f_{k|k-1}(\{\mathbf{x}\}) &= \max \left\{ \phi_{k|k-1}(\{\mathbf{x}\}|\emptyset) f_{k-1|k-1}(\emptyset), \right. \\ &\quad \left. \sup_{\mathbf{x}' \in \mathcal{X}} [\phi_{k|k-1}(\{\mathbf{x}\}|\{\mathbf{x}'\}) f_{k-1|k-1}(\{\mathbf{x}'\})] \right\} \\ &= \max \left\{ \tau_{01} q_{k-1|k-1}^0 b_{k|k-1}(\mathbf{x}), \right. \\ &\quad \left. \tau_{11} q_{k-1|k-1}^1 \sup_{\mathbf{x}' \in \mathcal{X}} [\rho_{k|k-1}(\mathbf{x}|\mathbf{x}') \pi_{k-1|k-1}(\mathbf{x}')] \right\} \end{aligned} \quad (\text{A.5})$$

From (A.3) we have $f_{k|k-1}(\{\mathbf{x}\}) = q_{k|k-1}^1 \pi_{k|k-1}(\mathbf{x})$, which leads to:

$$\sup_{\mathbf{x} \in \mathcal{X}} f_{k|k-1}(\{\mathbf{x}\}) = q_{k|k-1}^1 \underbrace{\sup_{\mathbf{x} \in \mathcal{X}} \pi_{k|k-1}(\mathbf{x})}_{=1} \quad (\text{A.6})$$

From (A.6), using (A.5), we obtain (21):

$$\begin{aligned} q_{k|k-1}^1 &= \sup_{\mathbf{x} \in \mathcal{X}} f_{k|k-1}(\{\mathbf{x}\}) \\ &= \max \left\{ \tau_{01} q_{k-1|k-1}^0, \tau_{11} q_{k-1|k-1}^1 \right\} \end{aligned} \quad (\text{A.7})$$

using the fact that $\sup_{\mathbf{x} \in \mathcal{X}} b_{k|k-1}(\mathbf{x}) = 1$ and

$$\sup_{\mathbf{x} \in \mathbf{X}} \left[\sup_{\mathbf{x}' \in \mathbf{X}} [\rho_{k|k-1}(\mathbf{x}|\mathbf{x}') \pi_{k-1|k-1}(\mathbf{x}')] \right] = 1.$$

From (A.3) we have:

$$\pi_{k|k-1}(\mathbf{x}) = \frac{1}{q_{k|k-1}^1} f_{k|k-1}(\{\mathbf{x}\}). \quad (\text{A.8})$$

Upon the substitution of $f_{k|k-1}(\{\mathbf{x}\})$ in (A.8) with the expression (A.5) we obtain (22).

Let us show that $\max\{q_{k|k-1}^0, q_{k|k-1}^1\} = 1$. From (A.4) and (A.7) we have:

$$\begin{aligned}
& \max\{q_{k|k-1}^0, q_{k|k-1}^1\} \\
&= \max\left\{\max\{\tau_{00} q_{k-1|k-1}^0, \tau_{10} q_{k-1|k-1}^1\}, \max\{\tau_{01} q_{k-1|k-1}^0, \tau_{11} q_{k-1|k-1}^1\}\right\} \\
&= \max\left\{\max\{\tau_{00} q_{k-1|k-1}^0, \tau_{01} q_{k-1|k-1}^0\}, \max\{\tau_{10} q_{k-1|k-1}^1, \tau_{11} q_{k-1|k-1}^1\}\right\} \\
&= \max\left\{\underbrace{\max\{\tau_{00}, \tau_{01}\}}_{=1} q_{k-1|k-1}^0, \underbrace{\max\{\tau_{10}, \tau_{11}\}}_{=1} q_{k-1|k-1}^1\right\} \\
&= \max\{q_{k-1|k-1}^0, q_{k-1|k-1}^1\} \\
&= 1.
\end{aligned}$$

Appendix A.2. Derivation of update equations in Sec. 3.3

The BPF update equation (23) for independent sensors, see (24), is given by:

$$f_{k|k}(\mathbf{X}) = \frac{\prod_{i=1}^M \varphi_i(\mathbf{Z}_k^{(i)}|\mathbf{X}) f_{k|k-1}(\mathbf{X})}{\sup_{\mathbf{X} \in \mathcal{F}(\mathcal{X})} \left[\prod_{i=1}^M \varphi_i(\mathbf{Z}_k^{(i)}|\mathbf{X}) f_{k|k-1}(\mathbf{X}) \right]} \quad (\text{A.9})$$

where

$$\varphi_i(\mathbf{Z}_k^{(i)}|\mathbf{X}) = \begin{cases} \kappa_i(\mathbf{Z}_k^{(i)}), & \text{if } \mathbf{X} = \emptyset \\ \max\left\{d^0 \kappa_i(\mathbf{Z}_k^{(i)}), \max_{\mathbf{z} \in \mathbf{Z}_k^{(i)}} [d^1 g_i(\mathbf{z}|\mathbf{x}) \kappa_i(\mathbf{Z}_k^{(i)} \setminus \{\mathbf{z}\})]\right\} & \text{if } \mathbf{X} = \{\mathbf{x}\} \end{cases} \quad (\text{A.10})$$

and $\kappa_i(\mathbf{Z})$ is defined by (26). Let us first focus on the denominator of (A.9), which we denote by D for brevity.

$$\begin{aligned}
D &= \sup_{\mathbf{X} \in \mathcal{F}(\mathcal{X})} \left[\prod_{i=1}^M \varphi_i(\mathbf{Z}_k^{(i)} | \mathbf{X}) f_{k|k-1}(\mathbf{X}) \right] \\
&= \max \left\{ \prod_{i=1}^M \varphi_i(\mathbf{Z}_k^{(i)} | \emptyset) f_{k|k-1}(\emptyset), \sup_{\mathbf{x} \in \mathcal{X}} \prod_{i=1}^M \varphi_i(\mathbf{Z}_k^{(i)} | \{\mathbf{x}\}) f_{k|k-1}(\{\mathbf{x}\}) \right\} \\
&= \max \left\{ q_{k|k-1}^0 \prod_{i=1}^M \kappa_i(\mathbf{Z}_k^{(i)}), \sup_{\mathbf{x} \in \mathcal{X}} \left[q_{k|k-1}^1 \pi_{k|k-1}(\mathbf{x}) \right. \right. \\
&\quad \left. \left. \times \prod_{i=1}^M \max \left\{ d^0 \kappa_i(\mathbf{Z}_k^{(i)}), \max_{\mathbf{z} \in \mathbf{Z}_k^{(i)}} [d^1 g_i(\mathbf{z} | \mathbf{x}) \kappa_i(\mathbf{Z}_k^{(i)} \setminus \{\mathbf{z}\})] \right\} \right] \right\} \\
&= \max \left\{ q_{k|k-1}^0 \prod_{i=1}^M \kappa_i(\mathbf{Z}_k^{(i)}), \right. \\
&\quad \left. q_{k|k-1}^1 \prod_{i=1}^M \max \left\{ d^0 \kappa_i(\mathbf{Z}_k^{(i)}), \max_{\mathbf{z} \in \mathbf{Z}_k^{(i)}} [d^1 \kappa_i(\mathbf{Z}_k^{(i)} \setminus \{\mathbf{z}\}) \sup_{\mathbf{x} \in \mathcal{X}} [g_i(\mathbf{z} | \mathbf{x}) \pi_{k|k-1}(\mathbf{x})]] \right\} \right\} \\
&= \prod_{i=1}^M \kappa_i(\mathbf{Z}_k^{(i)}) \max \{ q_{k|k-1}^0, \alpha q_{k|k-1}^1 \} \tag{A.11}
\end{aligned}$$

where

$$\alpha = \prod_{i=1}^M \max \left\{ d^0, d^1 \max_{\mathbf{z} \in \mathbf{Z}_k^{(i)}} \left[\frac{\kappa_i(\mathbf{Z}_k^{(i)} \setminus \{\mathbf{z}\})}{\kappa_i(\mathbf{Z}_k^{(i)})} \sup_{\mathbf{x} \in \mathcal{X}} [g_i(\mathbf{z} | \mathbf{x}) \pi_{k|k-1}(\mathbf{x})] \right] \right\} \tag{A.12}$$

Let us now write (A.9) for $\mathbf{X} = \emptyset$, recalling that $f_{k|k}(\emptyset) = q_{k|k}^0$ and $f_{k|k-1}(\emptyset) = q_{k|k-1}^0$ and using (A.11):

$$q_{k|k}^0 = \frac{\prod_{i=1}^M \kappa_i(\mathbf{Z}_k^{(i)}) q_{k|k-1}^0}{\prod_{i=1}^M \kappa_i(\mathbf{Z}_k^{(i)}) \max \{ q_{k|k-1}^0, \alpha q_{k|k-1}^1 \}} \tag{A.13}$$

After canceling the term $\prod_{i=1}^M \kappa_i(\mathbf{Z}_k^{(i)})$ in (A.13) we obtain (28), i.e.

$$q_{k|k}^0 = \frac{q_{k|k-1}^0}{\max \{ q_{k|k-1}^0, \alpha q_{k|k-1}^1 \}} \tag{A.14}$$

Next we derive (29). Because \mathbf{X} , after the update, remains a Bernoulli UFS, $f_{k|k}(\mathbf{X})$ of (A.9) will be in the form given by (14). Thus, for the case $\mathbf{X} = \{\mathbf{x}\}$,

we have $f_{k|k}(\{\mathbf{x}\}) = q_{k|k}^1 \pi_{k|k}(\mathbf{x})$. In accordance with the reasoning in (A.6):

$$q_{k|k}^1 = \sup_{\mathbf{x} \in \mathcal{X}} f_{k|k}(\{\mathbf{x}\}) \quad (\text{A.15})$$

$$= \frac{\sup_{\mathbf{x} \in \mathcal{X}} \left[\prod_{i=1}^M \varphi(\mathbf{Z}_k^{(i)} | \{\mathbf{x}\}) \right] f_{k|k-1}(\{\mathbf{x}\})}{D} \quad (\text{A.16})$$

where D is given by (A.11). Following derivation of D , it can easily be shown that

$$q_{k|k}^1 = \frac{\prod_{i=1}^M \varphi(\mathbf{Z}_k^{(i)}) \alpha q_{k|k-1}^1}{\prod_{i=1}^M \kappa_i(\mathbf{Z}_k^{(i)}) \max[q_{k|k-1}^0, \alpha q_{k|k-1}^1]} \quad (\text{A.17})$$

After canceling the term $\prod_{i=1}^M \kappa_i(\mathbf{Z}_k^{(i)})$ in (A.17) we obtain (29), i.e.

$$q_{k|k}^1 = \frac{\alpha q_{k|k-1}^1}{\max[q_{k|k-1}^0, \alpha q_{k|k-1}^1]}. \quad (\text{A.18})$$

Finally, (32) can be obtained from (A.9) in the case $\mathbf{X} = \{\mathbf{x}\}$. Recall that $f_{k|k}(\{\mathbf{x}\}) = q_{k|k}^1 \pi_{k|k}(\mathbf{x})$ and $f_{k|k-1}(\{\mathbf{x}\}) = q_{k|k-1}^1 \pi_{k|k-1}(\mathbf{x})$. Then using (A.11), (A.9) can be written as:

$$q_{k|k}^1 \pi_{k|k}(\mathbf{x}) = \frac{\prod_{i=1}^M \kappa_i(\mathbf{Z}_k^{(i)}) \prod_{i=1}^M L_i(\mathbf{Z}_k^{(i)} | \mathbf{x}) q_{k|k-1}^1 \pi_{k|k-1}(\mathbf{x})}{\prod_{i=1}^M \kappa_i(\mathbf{Z}_k^{(i)}) \max\{q_{k|k-1}^0, \alpha q_{k|k-1}^1\}} \quad (\text{A.19})$$

where $L_i(\mathbf{Z}_k^{(i)} | \mathbf{x})$ was defined in (33). After canceling the term $\prod_{i=1}^M \kappa_i(\mathbf{Z}_k^{(i)})$ and rearranging we have:

$$\pi_{k|k}(\mathbf{x}) = \frac{q_{k|k-1}^1}{q_{k|k}^1 \max\{q_{k|k-1}^0, \alpha q_{k|k-1}^1\}} \prod_{i=1}^M L_i(\mathbf{Z}_k^{(i)} | \mathbf{x}) \pi_{k|k-1}(\mathbf{x}) \quad (\text{A.20})$$

Using (A.18) we simplify (A.20) to:

$$\pi_{k|k}(\mathbf{x}) = \frac{1}{\alpha} \prod_{i=1}^M L_i(\mathbf{Z}_k^{(i)} | \mathbf{x}) \pi_{k|k-1}(\mathbf{x}) \quad (\text{A.21})$$

which is identical to (32).

References

- [1] A. H. Jazwinski. *Stochastic Processes and Filtering Theory*. Academic Press, 1970.

- [2] B. Ristic, S. Arulampalam, and X. Wang. Measurement variance ignorant target motion analysis. *Information Fusion*, 43:27–32, 2018.
- [3] T. S. Ferguson. A Bayesian analysis of some nonparametric problems. *The annals of statistics*, pages 209–230, 1973.
- [4] A. Milan, S. H. Rezatofighi, A. Dick, I. Reid, and K. Schindler. Online multi-target tracking using recurrent neural networks. In *Thirty-First AAAI Conference on Artificial Intelligence*, 2017.
- [5] T. O’Hagan. Dicing with the unknown. *Significance*, 1(3):132–133, 2004.
- [6] G. J. Klir. *Uncertainty and information: foundations of generalized information theory*. John Wiley & Sons, 2005.
- [7] T. Denoeux. Introduction to belief functions. Fifth School on Belief Functions and their Applications (https://bfasociety.org/BFTA2019/Lectures/BFTA2019_Thierry_Lecture1.pdf), 2019.
- [8] L. A. Zadeh. Fuzzy sets. *Information and control*, 8(3):338–353, 1965.
- [9] P. Walley. *Statistical reasoning with imprecise probabilities*. Chapman and Hall, 1991.
- [10] L. A. Zadeh. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, 1:3–28, 1978.
- [11] D. Dubois and H. Prade. Possibility theory and its applications: Where do we stand? In J. Kacprzyk and W. Pedrycz, editors, *Springer Handbook of Computational Intelligence*, pages 31–60. Springer Berlin Heidelberg, 2015.
- [12] A. P. Dempster. Upper and lower probabilities induced by a multivalued mapping. *The Annals of Mathematical Statistics*, 38(2):325–339, 1967.
- [13] G. Shafer. *A mathematical theory of evidence*, volume 42. Princeton university press, 1976.
- [14] B. Ristic, B.-T. Vo, B.-N. Vo, and A. Farina. A tutorial on Bernoulli filters: theory, implementation and applications. *IEEE Trans on Signal Processing*, 61(13):3406–3430, 2013.
- [15] R. Mahler. *Statistical Multisource Multitarget Information Fusion*. Artech House, 2007.
- [16] F. Hampel. Nonadditive probabilities in statistics. *Journal of Statistical Theory and Practice*, 3(1):11–23, 2009.

- [17] J. Houssineau and A.N. Bishop. Smoothing and filtering with a class of outer measures. *SIAM Journal on Uncertainty Quantification*, 6(2):845–866, 2018.
- [18] J. Houssineau and B. Ristic. Sequential Monte Carlo algorithms for a class of outer measures. *arXiv preprint arXiv:1708.06489*, 2017.
- [19] A. N. Bishop, J. Houssineau, D. Anglely, and B. Ristic. Spatio-temporal tracking from natural language statements using outer probability theory. *Information Sciences*, 463-464:56–74, 2018.
- [20] B. Ristic, J. Houssineau, and S. Arulampalam. Robust target motion analysis using the possibility particle filter. *IET Radar Sonar Navigation*, 13(1), 2019.
- [21] B. Ristic, S. Arulampalam, and N. Gordon. *Beyond the Kalman filter: Particle filters for tracking applications*. Artech House, 2004.
- [22] J. Houssineau. Parameter estimation with a class of outer probability measures. *arXiv preprint arXiv:1801.00569*, 2018.
- [23] D. Dubois. Possibility theory and statistical reasoning. *Computational statistics & data analysis*, 51(1):47–69, 2006.
- [24] D. Dubois and H. Prade. Possibility theory and its applications: Where do we stand? In *Springer Handbook of Computational Intelligence*, pages 31–60. Springer, 2015.
- [25] C. Carlsson and R. Fullér. Possibilistic mean value and variance of fuzzy numbers: Some examples of application. In *IEEE International Conference on Fuzzy Systems*, pages 587–592, 2009.
- [26] D. Musicki, R. Evans, and S. Stankovic. Integrated probabilistic data association. *IEEE Transactions on automatic control*, 39(6):1237–1241, 1994.
- [27] J. Houssineau. Detection and estimation of partially-observed dynamical systems: an outer-measure approach. *arXiv preprint arXiv:1801.00571*, 2018.
- [28] P. P. Shenoy. Using possibility theory in expert systems. *Fuzzy Sets and Systems*, 52(2):129–142, 1992.
- [29] B. Ristic and A. Farina. Target tracking via multi-static Doppler shifts. *IET Radar, Sonar & Navigation*, 7(5):508–516, 2013.
- [30] D. Schuhmacher, B.-T. Vo, and B.-N. Vo. A consistent metric for performance evaluation of multi-object filters. *IEEE Trans. Signal Processing*, 56(8):3447–3457, Aug. 2008.