Least Committed Basic Belief Density Induced by a Multivariate Gaussian: Formulation with Applications

François Caron [a,*], Branko Ristic [b], Emmanuel Duflos [a], Philippe Vanheeghe [a]

[a] *LAGIS, Ecole Centrale de Lille, Cité scientifique, BP48, 59651 Villeneuve d'Ascq Cedex, France*

[b] *ISR Division, DSTO, PO Box 1500, Edinburgh SA 5111, Australia*

Abstract

We consider here the case where our knowledge is partial and based on a betting density function which is n-dimensional Gaussian. The explicit formulation of the least committed basic belief density (bbd) of the multivariate Gaussian pdf is provided in the transferable belief model (TBM) framework. Beliefs are then assigned to hyperspheres and the bbd follows a χ^2 distribution. Two applications are also presented. The first one deals with model based classification in the joint speed-acceleration feature space. The second is devoted to joint target tracking and classification: the tracking part is performed using a Rao-Blackwellized particle filter, while the classification is carried out within the developed TBM scheme.

Key words: Belief function theory, Transferable Belief Model, evidential theory, multivariate Gaussian pdf, target classification, Target Tracking, Particle Filtering
PACS:

1 Introduction

The interpretation of the belief function theory within the Transferable Belief Model [1] (TBM) [1] has been initially defined for discrete frames of discernment. Smets re-

* Corresponding author
Email addresses: `Francois.Caron@ec-lille.fr` (François Caron), `Branko.Ristic@dsto.defence.gov.au` (Branko Ristic), `Emmanuel.Duflos@ec-lille.fr` (Emmanuel Duflos), `Philippe.Vanheeghe@ec-lille.fr` (Philippe Vanheeghe).
URL: `syner.ec-lille.fr/∼caron/` (François Caron).

[1] This interpretation of belief functions does not assume any underlying probability measure. Other interpretations of Dempster-Shafer theory exist, such as the lower probabilities

cently defined new tools for extending the belief functions to the set of reals [2]. In this model, beliefs are quantified as basic belief densities (bbd's) and focal elements are closed intervals of $\mathbb{R}$. This emergent theory has found some developments in the literature [3], [4].

We assume our knowledge is partial and represented by some betting probability function on the observation, which belongs to the continuous domain. From this betting probability function one can build the least committed bbd, so that the general tools of the belief function theory (such as the Generalized Bayesian Theorem [5], [6] and combination rules) can be applied for reasoning. Explicit solutions were given to find the least committed bbd induced by an univariate and unimodal probability density function (pdf) in [2]. Applications of this approach to model based classification have been presented in [7] and [8]. The resulting classifier is more cautious and its decisions arguably more meaningful than those obtained using the corresponding Bayesian classifier, due to the least commitment principle.

Although the belief function theory on reals is conceptually valid for $\mathbb{R}^n$, no explicit analytic solutions have been proposed in [2] for multi-dimensional spaces. In this paper, we go a step further and provide an explicit formulation of the least committed bbd induced by an n-dimensional Gaussian pdf, the "engineers favorite", to deal with possibly correlated multi-dimensional data. With this new theoretical tool in hand, we revisit model based target classification problems discussed in [8] and [9].

2 Review of the Transferable Belief Model

This section summarizes the main concepts of belief functions on discrete sets and on the set of reals. For proofs and a more thorough study, the reader should refer to [1] and [2].

The following notations are used throughout the paper for the basic belief densities (bbd) and basic belief assignment (bba) m and its related functions bel, pl and q:

$$m^{domain}[condition](subset)$$

- *domain*: the set of elements on which the bbd or bba is assigned.
- *condition*: the conditions which are assumed to hold true when the belief holder assesses the bba/bbd m. In this paper, the condition is given by observations provided by sensors.
- *subset*: any subset of the domain.

---

or the hints models.

2.1 Belief functions on a discrete frame

Consider a discrete set of n mutually exclusive events, called the frame of discernment

$$\Theta = \{\theta_1, \theta_2, \dots, \theta_n\}. \tag{1}$$

The belief functions are defined on the set of subsets of Θ, called 2^Θ and defined as $2^\Theta = \{A | A \subseteq \Theta\}$. The belief is represented by a so-called *basic belief assignment* (bba) $m : 2^\Theta \to [0,1]$ such that $\sum_{A \subseteq \Theta} m(A) = 1$. $m(A)$ represents the amount of belief that the actual solution is exactly committed to A, and due to lack of knowledge cannot be transferred to any more specific event. The subsets A with a non zero mass $m(A)$ are referred as the focal sets. The state of complete ignorance is represented by the so-called vacuous bba defined by $m(A) = 1$ if $A = \Theta$ and 0 otherwise.

2.1.1 Belief functions

The belief function bel, plausibility function pl and the commonality function q are other functions to quantify beliefs that are in one-to-one correspondance with the bba m. They are defined, for all $A \subseteq \Theta$ by

$$bel(A) = \sum_{\{B | B \subseteq A, B \neq \emptyset\}} m(B) \tag{2}$$

$$pl(A) = \sum_{\{B | A \cap B \neq \emptyset\}} m(B) \tag{3}$$

$$q(A) = \sum_{\{B | A \subseteq B\}} m(B) \tag{4}$$

2.1.2 Conjunctive combination

Let m_1 and m_2 be two bbas defined on the same frame of discernment Θ. Suppose the bbas result from two distinct pieces of evidence. The conjunctive combination of these two pieces of evidence is given, for all $A \subseteq \Theta$, by

$$m_{12}(A) = (m_1 \bigcirc\!\!\!\!\cap\, m_2)(A) = \sum_{\{A_1, A_2 | A_1 \cap A_2 = A\}} m_1(A_1) m_2(A_2) \tag{5}$$

2.1.3 Pignistic probability

In order to take a decision on the set of exclusive hypotheses Θ, one has to operate in the probabilistic framework by assigning probabilities to each singleton θ_i. This is done by the pignistic transform which is given for each $\theta_i \in \Theta$ by

$$BetP(\theta_i) = \sum_{\{A \subseteq \Theta | \theta_i \in A\}} \frac{m(A)}{|A|(1 - m(\emptyset))} \tag{6}$$

$BetP$ is called the pignistic probability and it is the probability measure used for decision making.

2.1.4 Generalised Bayesian Theorem

Let z be a measure on a space $\mathcal{Z}$. Suppose that one knows the conditionnal plausibilities $pl^{\mathcal{Z}}[\theta_i](z)$ for $i = 1, \ldots, n$. Then the Generalised Bayesian Theorem (GBT) provides a way to compute for all $A \subseteq \Theta$ the conditional bba $m^{\Theta}[z](A)$ as follows [5]

$$m^{\Theta}[z](A) = \prod_{\{\theta_i|\theta_i \in A\}} pl^{\mathcal{Z}}[\theta_i](z) \prod_{\{\theta_i|\theta_i \in \overline{A}\}} (1 - pl^{\mathcal{Z}}[\theta_i](z)) \tag{7}$$

2.2 Belief functions on $\mathbb{R}$

Under the transferable belief model on reals, basic belief masses become basic belief densities and positive bbds are only assigned to the intervals of $\mathbb{R}$ [2].

Consider the set $\mathcal{T}$ =$\{(x, y)|x \leq y\}$. Intervals of $\mathbb{R}$ are represented as points of the triangle $\mathcal{T}$ (see Figure 1). Let $f^{\mathcal{T}} : \mathcal{T} \rightarrow [0, +\infty[$ be an unnormalized probability density function on $\mathcal{T}$. The bbd allocated to the interval $[a, b]$ is

$$m^{\mathbb{R}}([a, b]) = f^{\mathcal{T}}(a, b) \tag{8}$$

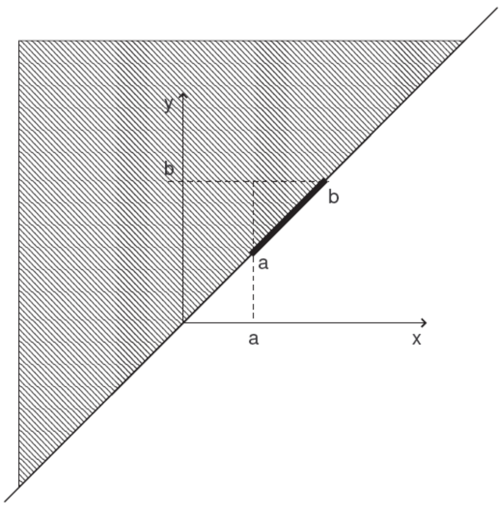

Fig. 1. Representation on $\mathcal{T}$ of the interval $[a, b] \subset \mathbb{R}$

2.2.1 Belief functions

The bbd $m([a, b])$ is the part of the belief holder that supports exactly $[a, b]$ i.e., that the actual world is in $[a, b]$ and that, due to lack of information, does not support any strict subset of $[a, b]$.

The degree of belief of $[a, b]$, $bel([a, b])$, quantifies the total amount of justified support given to $[a, b]$. $bel([a, b])$ is the sum of the masses given to the subsets of $[a, b]$. The degree of plausibility of $[a, b]$, $pl([a, b])$, quantifies the maximum amount of potential specific support that could be given to $[a, b]$. It is defined as the sum of the masses given to intervals $[a_i, b_i]$ such that $[a_i, b_i] \cap [a, b] \neq 0$.
The commonality $q([a, b])$ is another measure of belief which is useful for calculus in belief combination. It is defined as the sum of the masses given to the intervals $[a_i, b_i] \supseteq [a, b]$.

Graphical representations[2] on $\mathcal{T}$ of belief, commonality and plausibility are represented in the Figure 2. Each one is the integral of the bbds allocated to the gray areas.

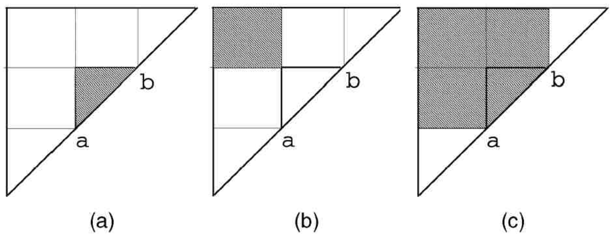

Fig. 2. Graphical representation on $\mathcal{T}$ of (a) belief (b)commonality and (c) plausibility

2.2.2 *Consonant bbd*

A consonant bbd is a bbd whose focal elements are nested. In this case, there exists an index u such that the focal elements can be labeled as I_u, with $I_u \subseteq I_{u'}$ when $u' > u$. If a bbd $m^{\mathbb{R}}$ is consonant, there exists an unormalized probability density function $h : [0, +\infty[\rightarrow [0, +\infty[$ such that

$$m^{\mathbb{R}}(I_u) = h(u), \; u \geq 0. \tag{9}$$

2.2.3 *Pignistic probability*

The pignistic density function $Betf$ is derived from $f^{\mathcal{T}}$ according to

$$Betf(a) = \lim_{\epsilon \rightarrow 0} \int_{x=-\infty}^{x=a} \int_{y=a+\epsilon}^{y=\infty} \frac{f^{\mathcal{T}}(x, y)}{y - x} dxdy, \; a \in \mathbb{R} \tag{10}$$

For consonant bbds, the last formula reduces to

$$Betf(a) = \int_{\min_{\{u' | a \in I_{u'}\}}(u')}^{\infty} \frac{h(u)}{l_u} du \tag{11}$$

where l_u is the length of the interval I_u.

[2] The vertical and horizontal lines in each triangle of this figure are shown only to make the pictures look nicer. Indeed they extend to infinity both upwards and to the left.

3 The least committed bbd of an univariate pdf

Suppose that your knowledge on the domain is partial and based on some potential betting behaviour, represented by the pignistic density function $Betf$. One wants to determine a basic belief density that induces this pignistic probability. Since the pignistic transform is a many-to-one transformation, an infinite number of bbd, called isopignistic bbd, can induce $Betf$. The least commitment principle [10], [11] suggests to choose, in the set of all isopignistic densities, the bbd that maximizes the commonality function q, named q-Least Committed (q-LC). As in the discrete case [12], the q-LC isopignistic is a consonant bbd, i.e. all focal sets are nested.

3.1 LC bbd of an univariate "bell-shaped" pdf

Smets [2] provided an explicit formulation of the LC isopignistic bbd for univariate "bell-shaped" pignistic probabilities. Let μ be the mode of the pdf

$$\mu = \arg\max_{x \in \mathbb{R}} (Betf(x)) \tag{12}$$

The focal sets of the least committed bbd are intervals $I_b = [a, b]$ such that $Betf(a) = Betf(b)$. $Betf$ being a bell-shaped density, a is uniquely defined by a function γ such as $a = \gamma(b)$. The bbd is defined by

$$m^{\mathbb{R}}(I_b) = \theta(b)\delta(a - \gamma(b)), \; b \geq \mu \tag{13}$$

with

$$\theta(b) = (\gamma(b) - b)\frac{\partial Betf(b)}{\partial b}, \; b \geq \mu. \tag{14}$$

and $\delta(x)$ is the Dirac's delta function.

3.2 LC bbd of an univariate Gaussian pdf

Deducing it from equations (13) and (14), Smets [2], [7] defined the LC isopignistic basic belief density for univariate Gaussian pignistic functions. We provide a new different formulation in the following result.

Theorem 1 *The q-LC isopignistic bbd of an univariate Gaussian pdf* $p(x) = \mathcal{N}(x : \mu, \sigma^2)$ *of mean* μ *and standard deviation* σ*, is defined by*

$$m^{\mathbb{R}}(I_\alpha) = \theta(\alpha), \; \alpha \geq 0 \tag{15}$$

where I_α *(*$\alpha \geq 0$*) are the nested focal elements (closed intervals) defined by*

$$\begin{aligned} I_\alpha &= \{x \in \mathbb{R} | \frac{(x-\mu)^2}{\sigma^2} \leq \alpha\} \\ &= [\mu - \sqrt{\alpha}\sigma; \mu + \sqrt{\alpha}\sigma] \end{aligned} \tag{16}$$

and $\theta(\alpha)$ is the degree of belief assigned to I_α which is shown to be a χ^2 probability density function with 3 degrees of freedoms, defined by

$$\theta(\alpha) = \frac{\sqrt{\alpha}}{\sqrt{2\pi}} \exp(-\frac{1}{2}\alpha), \alpha \geq 0. \tag{17}$$

The proof is given in appendix.

The plausibility of a point mass x is given by

$$pl^{\mathbb{R}}(x \in \mathbb{R}) = \int_{u=\left(\frac{x-\mu}{\sigma}\right)^2}^{\infty} \theta(u)du \tag{18}$$

Next we provide an n-dimensional generalization of the last result.

4 LC bbd induced by a multivariate Gaussian pdf

4.1 Intuitive understanding in 2-dimensional space

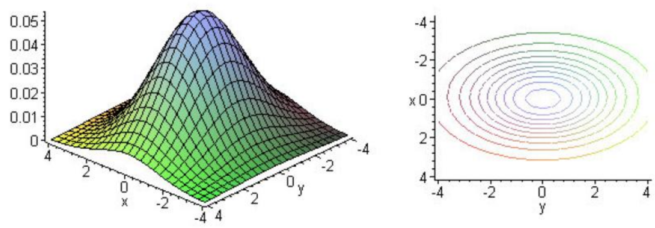

Fig. 3. Representation of a 2-dimensional Gaussian pdf

In two dimensions, isoprobability points of the multivariate Gaussian pdf of mean $\mu = \begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix}$ and covariance matrix $\Sigma = \begin{pmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{pmatrix}$ are ellipses of center (μ_x, μ_y) and semi-axis $(\alpha\sigma_x^2, \alpha\sigma_y^2)$ ($\alpha \in [0, +\infty[$). Such an ellipse is defined by the parametric equation

$$(\mathbf{x} - \mu)^T \Sigma^{-1} (\mathbf{x} - \mu) = \alpha$$

Masses of the bbd whose pignistic density is this multivariate Gaussian pdf are assigned to nested surfaces delimited by the different isoprobability ellipses. The set of all ellipses is defined by

$$E = \{E_\alpha,\ \alpha \in [0, +\infty[\}$$

with

$$E_\alpha = \{\mathbf{x} \in \mathbb{R}^2 | (\mathbf{x} - \mu)^T \Sigma^{-1} (\mathbf{x} - \mu) = \alpha\}$$

Masses are then assigned to surfaces delimited by these ellipses, i.e. to subsets S_α such that

$$S_\alpha = \{\mathbf{x} \in \mathbb{R}^2 | (\mathbf{x} - \mu)^T \Sigma^{-1} (\mathbf{x} - \mu) \leq \alpha\}.$$

4.2 LC isopignistic bbd of a n-dimensional Gaussian pdf

Let us consider the multivariate Gaussian pdf of mean $\mu \in \mathbb{R}^n$ and covariance matrix Σ

$$\begin{aligned} p(\mathbf{x}|\mu, \Sigma) &= \mathcal{N}(\mathbf{x} : \mu, \Sigma) \\ &= \frac{1}{(2\pi)^{\frac{n}{2}} \sqrt{\det(\Sigma)}} \exp(-\frac{1}{2} (\mathbf{x} - \mu)^T \Sigma^{-1} (\mathbf{x} - \mu)) \end{aligned} \tag{19}$$

with $\mathbf{x} \in \mathbb{R}^n$.

In dimension n, focal elements are the nested sets HV_α (hypervolumes) enclosed by the isoprobability hyperconics $HC_\alpha = \{\mathbf{x} \in \mathbb{R}^n | (\mathbf{x} - \mu)^T \Sigma^{-1} (\mathbf{x} - \mu) = \alpha\}$

$$HV_\alpha = \{\mathbf{x} \in \mathbb{R}^n | (\mathbf{x} - \mu)^T \Sigma^{-1} (\mathbf{x} - \mu) \leq \alpha\} \tag{20}$$

The bbd m is defined as

$$m^{\mathbb{R}^n}(HV_\alpha) = \theta(\alpha), \alpha \geq 0. \tag{21}$$

The volume of the hypersphere HV_α is expressed as (volume enclosed by an hyperconic of dimension n and axes $\alpha\Sigma$) :

$$V_\alpha = \int_{\mathbf{x} \in HV_\alpha} d\mathbf{x} = \frac{\pi^{\frac{n}{2}} \alpha^{\frac{n}{2}} \sqrt{\det(\Sigma)}}{\Gamma(\frac{n}{2} + 1)}. \tag{22}$$

Theorem 2 *The q-LC isopignistic bbd induced by a n-multivariate Gaussian pdf* $\mathcal{N}(\mathbf{x} : \mu, \Sigma)$ *of mean* μ *and covariance matrix* Σ *is the bbd defined by*

$$m^{\mathbb{R}^n}(HV_\alpha) = \theta(\alpha), \ \alpha \geq 0 \tag{23}$$

where

$$HV_\alpha = \{\mathbf{x} \in \mathbb{R}^n | (\mathbf{x} - \mu)^T \Sigma^{-1} (\mathbf{x} - \mu) \leq \alpha\} \tag{24}$$

and

$$\theta(\alpha) = \frac{\alpha^{\frac{n+2}{2} - 1}}{2^{\frac{n+2}{2}} \Gamma(\frac{n+2}{2})} \exp(-\frac{1}{2}\alpha) \tag{25}$$

is a χ^2 *distribution with* $n + 2$ *degrees of freedom.*
Plausibility of a point mass $\mathbf{x}$ *is then defined by*

$$\begin{aligned} pl^{\mathbb{R}^n}(\mathbf{x} \in \mathbb{R}^n) &= \int_{\alpha = (\mathbf{x} - \mu)^T \Sigma^{-1} (\mathbf{x} - \mu)}^{\alpha = +\infty} \theta(\alpha) d\alpha \qquad (26) \\ &= 1 - F_{n+2}((\mathbf{x} - \mu)^T \Sigma^{-1} (\mathbf{x} - \mu)) \qquad (27) \end{aligned}$$

where F_p is the cumulative density function[3] *of a χ^2 distribution with p degrees of freedom, defined by [13]:*

$$F_p(\chi^2) = \int_0^{\chi^2} \frac{u^{\frac{p}{2}-1}}{2^{\frac{p}{2}}\Gamma(\frac{p}{2})} \exp(-\frac{u}{2})du \tag{28}$$

The proof is given in appendix.

4.3 LC isopignistic bbd of a mixture of Gaussian pdfs

The last result can be generalized to probability density functions expressed as mixtures of Gaussian pdfs, in order to handle multivariate non-Gaussian densities. Let us assume that our knowledge is based on a finite mixture of M Gaussian pdfs:

$$p(\mathbf{x}) = \sum_{k=1}^{M} \beta_k \mathcal{N}(\mathbf{x} : \mu_k, \Sigma_k) \tag{29}$$

where $\mathbf{x} \in \mathbb{R}^n$, μ_k and Σ_k are the mean and covariance matrix of the kth Gaussian mixture component, with $k = 1, \ldots, M$, $\beta_k \geq 0$ and $\sum_{k=1}^{M} \beta_k = 1$.

Let us consider the sets $HV_{\alpha,k}$ for $\alpha \geq 0$ and $k = 1, \ldots, M$:

$$HV_{\alpha,k} = \{\mathbf{x} \in \mathbb{R}^n | (\mathbf{x} - \mu_k)^T \Sigma_k^{-1} (\mathbf{x} - \mu_k) \leq \alpha\}. \tag{30}$$

For a given component k of the mixture, the sets $HV_{\alpha,k}$, $\alpha \geq 0$, are nested. Parts of our belief are assigned to these sets such that, for $\alpha \geq 0$ and $k = 1, \ldots, M$:

$$m^{\mathbb{R}^n}(HV_{\alpha,k}) = \beta_k \theta(\alpha) \tag{31}$$

where $\theta(\alpha)$ is a χ^2 pdf with $n + 2$ degrees of freedom.

The plausibility of a point mass $x \in \mathbb{R}^n$ is given by

$$\begin{aligned} pl^{\mathbb{R}^n}(\mathbf{x} \in \mathbb{R}^n) &= \sum_{k=1}^{M} \beta_k \int_{\alpha=(\mathbf{x}-\mu_k)^T \Sigma_k^{-1} (\mathbf{x}-\mu_k)}^{\alpha=+\infty} \theta(\alpha) d\alpha \\ &= 1 - \sum_{k=1}^{M} \beta_k F_{n+2}\left((\mathbf{x} - \mu_k)^T \Sigma_k^{-1} (\mathbf{x} - \mu_k)\right) \end{aligned} \tag{32}$$

where F_p is the cumulative density function of a χ^2 distribution with p degrees of freedom.

[3] The cumulative density function F_p of Eq.(28) can be implemented using chi2cdf.m script in MATLAB.

5 Application to model-based classification

The problem is classification of non-cooperative flying objects in the surveillance volume. Many types of target features may be available for classification, such as the target shape, kinematic behaviour and Electro-Magnetic emissions [14]. In the following, we assume that 2-dimensional (2D) speed and acceleration data are at our disposal for classification of targets into one of three categories [9]:

- Class 1: Commercial planes,
- Class 2: Large military aircrafts (such as transporters, bombers),
- Class 3: Light and agile military aircrafts (fighter planes).

Our (incomplete) knowledge about acceleration and speed profiles for each of the three classes is typically described as shown in Table 1 [8].

Table 1
Speed and acceleration intervals for three air platform categories

Target class	Speed [km/h]		Acceleration [g]	
	Min	Max	Min	Max
Commercial (c_1)	560	885	-1	1
Bomber (c_2)	400	725	-4	4
Fighter (c_3)	525	950	-7	7

In our analysis we will consider the general case where the measured speed and acceleration are correlated. The Bayesian classifier will be compared to the belief function classifier.

5.1 Bayesian probabilistic analysis

Based on the information provided in Table 1, we define the corresponding 2D-Gaussian likelihoods for each class, in order to apply the Bayesian theorem. These 2D class-conditioned likelihoods $p(v, a|c_i)$ are represented in Figure 4 for each class c_i, $i = 1..3$.

The hypothesis space is $C = \{c_1, c_2, c_3\}$. Choosing a uniform prior for classes, the posterior is expressed as

$$\Pr(c_i|v, a) \propto p(v, a|c_i) \text{ for } i = 1..3$$

5.2 Belief function analysis

Considering that our knowledge about acceleration and speed profiles is very scarce and incomplete, one don't want to put to much confidence in this information. We

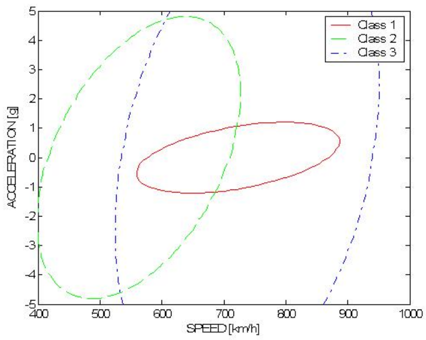

Fig. 4. Representation of the 2D Gaussian likelihoods $p(v, a|c_i)$ for each class c_i, $i = 1..3$. Each Gaussian likelihood of mean μ and covariance matrix Σ is represented by the set of points $\mathbf{x} = (v, a)$ such that $(\mathbf{x} - \mu)^T \Sigma^{-1} (\mathbf{x} - \mu) = 1$.

therefore would like to consider the Gaussian likelihood $p(v, a|c_i)$ as the pignistic transform of an underlying (least committed) bbd and to apply the General Bayesian Theorem so as to make the combination in the credal domain. The belief function scheme is the following [2], [8]:

(1) We start from the same pdfs $p(v, a|c_i)$ adopted for the Bayesian classifier.
(2) Since our knowledge is very scarce and incomplete, these pdfs are considered as pignistic probabilities. We then construct for each class c_i its LC bbd $m^{\mathbb{R}^2}[c_i]$, or directly its plausibility $pl^{\mathbb{R}^2}[c_i](v, a)$ using equation (26).
(3) Then we apply the General Bayesian Theorem(GBT) [5], [6]. It yields for every subset $A \subseteq C$ the following bba:

$$m^C[v, a](A) = \prod_{c_i \in A} pl^{\mathbb{R}^2}[c_i](v, a) \prod_{c_i \in \overline{A}} \left[1 - pl^{\mathbb{R}^2}[c_i](v, a)\right] \tag{33}$$

(4) The last step is to apply the pignistic transform to $m^C[v, a]$ in order to get the pignistic class probabilities

$$BetP^C[v, a](c_i) = \sum_{\{A \subseteq C | c_i \in A\}} \frac{m^C[v, a](A)}{|A| \left[1 - m^C[v, a](\emptyset)\right]} \tag{34}$$

where $|A|$ is the cardinal of the set A.

Posterior probability and pignistic probability, respectively given by the Bayesian classifier and the Belief function classifier, are represented as a function of speed, for $a = 0$, on Figures 6 and 7.

The mass allocated to the total ignorance $m^C[v, a](C)$ is represented in Figure 5. Observe that ignorance is high for $600 < v < 700$ and $-0.5 < a < 0.5$. This explains why in Figure 7, the belief function classifier is quite undecided between the three classes for a=0 and $600 < v < 700$. However, the Bayesian classifier largely favours class 1 in the same interval. Being undecided makes more sense here, considering that most likely observations of speed and acceleration fall in that region.

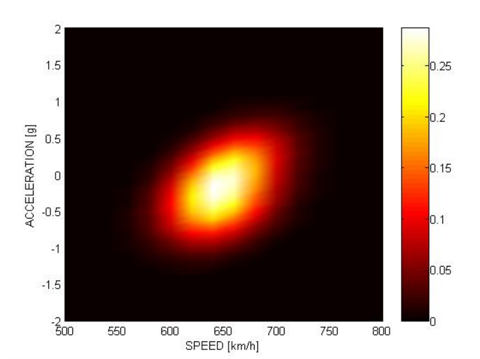

Fig. 5. Mass allocated to the total ignorance $m^C[v,a](C)$ in function of speed v and acceleration a. White corresponds to a high ignorance and black to a low ignorance. Ignorance is high for $600 < v < 700$ and $-0.5 < a < 0.5$.

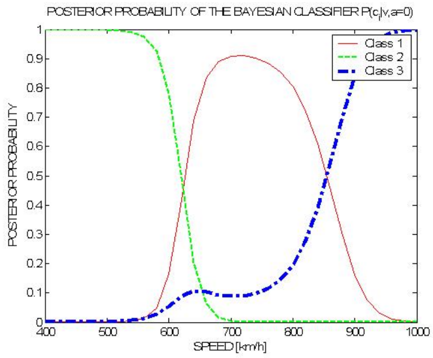

Fig. 6. Posterior class probabilities $\Pr(c_i|v,a)$ in function of v for $a = 0$

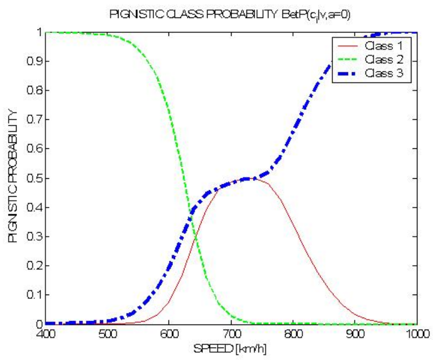

Fig. 7. Pignistic class probabilities $BetP^C[v,a](c_i)$ in function of v for $a = 0$

6 Application to joint target tracking and classification

Next we address the problem of joint tracking and classification of targets. The problem formulation and jump Markov statistical models are the same as those used in [9]. However, the algorithms both for tracking and classification are different. Instead of using an interacting multiple model (IMM) to perform target tracking, we use here a simulation-based method known as the Rao-Blackwellised particle filter [15–18]. Furthermore, we perform classification in the TBM framework, to better deal with underlying imprecision on target classes. Due to the nature of the tracker output, classification is based on results of Section 4.3.

6.1 Statistical model

Let us consider a target of unknown time-invariant class $c \in \{1, 2, 3\}$. We aim at estimating the state vector $\mathbf{x}_t = \begin{bmatrix} x\ \dot{x}\ y\ \dot{y} \end{bmatrix}^{\mathrm{T}}$ composed of the position and acceleration of the target in a 2-dimensional frame, as well as the type c of the target from a set of measurements $\mathbf{z}_1, \ldots, \mathbf{z}_t$ obtained sequentially. The target may experience different types of movements that are represented by a discrete mode $s_t(c)$. The target evolution model and sensor observation model, which define a Jump Markov Linear Model (JMLS), are defined as [9,19]

$$\mathbf{x}_{t+1} = F\mathbf{x}_t + G\mathbf{a}(s_{t+1}(c)) + G\mathbf{v}_t \tag{35}$$

$$\mathbf{z}_t = H\mathbf{x}_t + \mathbf{w}_t \tag{36}$$

where $F = \begin{pmatrix} 1 & T & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T \\ 0 & 0 & 0 & 1 \end{pmatrix}$, $G = \begin{pmatrix} \frac{T^2}{2} & 0 \\ T & 0 \\ 0 & \frac{T^2}{2} \\ 0 & T \end{pmatrix}$ and T is the sampling time. $\mathbf{v}_t$ is a white centered Gaussian noise of known covariance matrix Q. $\mathbf{a}(s_{t+1}(c))$ is the input acceleration vector. In order to simplify analysis, the observation model is supposed to be linear, with observation matrix $H = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$. $\mathbf{w}_t$ is a white centered Gaussian noise of known covariance matrix R. The evolution model has several modes which are selected using the discrete latent variable $s_t(c) \in \{1, \ldots, d(c)\}$ where $d(c)$ is the number of possible modes. As explained in [9], the number of possible modes depends on the target class value $c \in \{1, 2, 3\}$

$$\begin{cases} s_t(c) \in \{1\} \text{ if } c = 1 \\ s_t(c) \in \{1, \ldots, 5\} \text{ if } c = 2 \\ s_t(c) \in \{1, \ldots, 13\} \text{ if } c = 3 \end{cases} \tag{37}$$

The discrete latent variable $s_t(c)$ is supposed to follow a Markov model defined by

$$\Pr(s_t(c) | s_{t-1}(c), c) = \pi_c(s_{t-1}(c), s_t(c)) \tag{38}$$

where $\Pi_c = (\pi_c(i,j))$ is the transition matrix of the class c, $\Pi_1 = \begin{pmatrix} 1 & 0_{1\times 12} \\ 0_{12\times 1} & 0_{12\times 12} \end{pmatrix}$,

$\Pi_2 = \begin{pmatrix} A & 0_{5\times 7} \\ 0_{7\times 5} & 0_{7\times 7} \end{pmatrix}$ with $A = \begin{pmatrix} .9 & .025 & .025 & .025 & .025 \\ .1 & .9 & 0 & 0 & 0 \\ .1 & 0 & .9 & 0 & 0 \\ .1 & 0 & 0 & .9 & 0 \\ .1 & 0 & 0 & 0 & .9 \end{pmatrix}$. Π_3 is a 13×13 matrix which has a value of 0.9 along the diagonal, while the remaining $1-p$ are equally distributed across the non-zero elements in each row. Figure 8 summarizes the possible values of acceleration for each class and the corresponding values for s_t. Arrows indicate non-zero elements in Π.

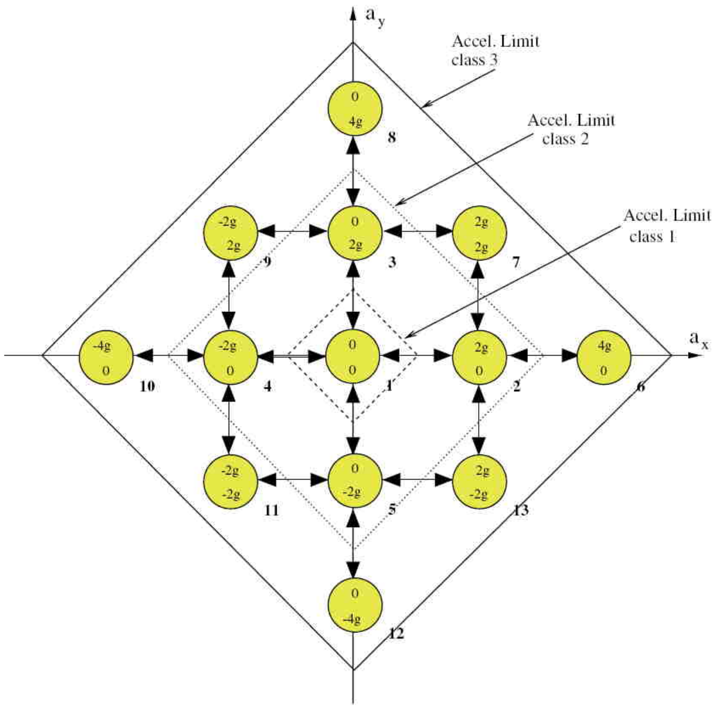

Fig. 8. The digraph used in the class-conditioned Jump Markov Linear model [9]. The bold numbers 1 to 13 correspond to the mode s_t, whose possible values are class-conditioned, as given by Equation (37). The values within the circles correspond to the input acceleration vector $\mathbf{a}(\cdot)$ associated with the mode s_t. For example, $\mathbf{a}(2) = [2g\ 0]$ and $\mathbf{a}(9) = [-2g\ 2g]$.

6.2 Target tracking

In the following we use notation: $\mathbf{a}_{u:v} = \{a_u, a_{u+1}, \ldots, a_v\}$ for $v > u$. The objective of filtering, for each class, is to estimate recursively the posterior pdf

$$p(\mathbf{x}_t|\mathbf{z}_{1:t}, c) = \int p(\mathbf{x}_t, \mathbf{s}_{1:t}(c)|\mathbf{z}_{1:t}, c)d\mathbf{s}_{1:t}(c) \tag{39}$$

$$= \sum_{\mathbf{s}_{1:t}(c)\in\{1,\ldots,d(c)\}^t} p(\mathbf{x}_t|\mathbf{s}_{1:t}(c), \mathbf{z}_{1:t}, c)\,\Pr(\mathbf{s}_{1:t}(c)|\mathbf{z}_{1:t}, c) \tag{40}$$

Here $\mathbf{s}_{1:t}(c)$ denotes a mode sequence, for example if $t = 5$ and $c = 2$, a mode sequence can take value $\mathbf{s}_{1:5}(2) = \{1, 1, 2, 2, 1\}$. As the number of possible values taken by $\mathbf{s}_{1:t}(c)$ grows exponentially with time if $d(c) > 1$ (i.e. for target classes 2 and 3), the last integral cannot be computed analytically. Traditionally the problem is solved using analytic approximate methods, such as the IMM. Instead, we will use a simulation-based approach, known as the Rao-Blackwellized particle filter [15–18] (RBPF). Contrary to the IMM, the RBPF performs a Monte Carlo approximation of the posterior density. This approximation converges to the true posterior as the number of random samples (or particles) increases. The RBPF is a more accurate, although also a computationally more intensive approach than the IMM. The idea is to approximate the last integral using a set of N weighted particles, indexed by i, each representing a random realisation of the mode sequence. The particles and their normalised weights are denoted by $\widetilde{\mathbf{s}}_{1:t}^{(i)}(c)$ and $\widetilde{w}_t^{(i)}(c)$, respecively. Then we can approximate the last term in (40) by [20–22]

$$\Pr(\mathbf{s}_{1:t}|\mathbf{z}_{1:t}, c) \simeq \sum_{i=1}^{N} \widetilde{w}_t^{(i)}(c)\delta(\mathbf{s}_{1:t} - \widetilde{\mathbf{s}}_{1:t}^{(i)}(c)) \tag{41}$$

where $\delta(x) = 1$ if $x = 0$ and 0 otherwise. The whole posterior is thus approximated by

$$p(\mathbf{x}_t|\mathbf{z}_{1:t}, c) \simeq \sum_{i=1}^{N} \widetilde{w}_t^{(i)}(c)p(\mathbf{x}_t|\widetilde{\mathbf{s}}_{1:t}^{(i)}(c), \mathbf{z}_{1:t}) \tag{42}$$

Conditionally on $\mathbf{s}_{1:t}(c)$, the system is linear and Gaussian, thus the conditional posterior $p(\mathbf{x}_t|\widetilde{\mathbf{s}}_{1:t}^{(i)}(c), \mathbf{z}_{1:t})$ can be computed recurrently with a Kalman filter. Hence

$$p(\mathbf{x}_t|\mathbf{z}_{1:t}, c) \simeq \sum_{i=1}^{N} \widetilde{w}_t^{(i)}(c)\mathcal{N}\left(\mathbf{x}_t : \widehat{\mathbf{x}}_{t|t}(\widetilde{\mathbf{s}}_{1:t}^{(i)}(c)), \Sigma_{t|t}(\widetilde{\mathbf{s}}_{1:t}^{(i)}(c))\right) \tag{43}$$

where $\widehat{\mathbf{x}}_{t|t}(\widetilde{\mathbf{s}}_{1:t}^{(i)}(c))$ and $\Sigma_{t|t}(\widetilde{\mathbf{s}}_{1:t}^{(i)}(c))$ are the state estimate and its covariance matrix, respectively, computed by a Kalman filter step for each particle i.

For each class c, the MMSE estimate (the tracking output) is then given by

$$\widehat{\mathbf{x}}_{t|t}^{\text{MMSE}}(c) = \sum_{i=1}^{N} \widetilde{w}_t^{(i)}(c)\widehat{\mathbf{x}}_{t|t}(\widetilde{\mathbf{s}}_{1:t}^{(i)}(c)) \tag{44}$$

6.3 Target classification

Target classification is carried out by following the same steps as in Section 5.2. In step 1, we have to get an estimate of the complete class-conditioned likelihood $p(\mathbf{z}_{1:t}|c)$. This pdf can be expressed as

$$p(\mathbf{z}_{1:t}|c) = \sum_{\mathbf{s}_{1:t}\in\{1,\dots,d(c)\}^t} p(\mathbf{z}_{1:t}|\mathbf{s}_{1:t}, c)\Pr(\mathbf{s}_{1:t}|c) \tag{45}$$

$$\simeq \sum_{i=1}^{N} \overline{w}_t^{(i)}(c)p(\mathbf{z}_{1:t}|\widetilde{\mathbf{s}}_{1:t}^{(i)}(c)) \tag{46}$$

where $\overline{w}_t^{(i)}(c) \propto \frac{\widetilde{w}_t^{(i)}(c)}{p(\mathbf{z}_{1:t}|\widetilde{\mathbf{s}}_{1:t}^{(i)}(c))}$ and $\sum_{i=1}^{N} \overline{w}_t^{(i)}(c) = 1$ because the probability $\Pr(\mathbf{s}_{1:t}|c)$ can be approximated by $\sum_{i=1}^{N} \overline{w}_t^{(i)}(c)\delta(\mathbf{s}_{1:t} - \widetilde{\mathbf{s}}_{1:t}^{(i)}(c))$. The system being conditionally linear and Gaussian, the pdf $p(\mathbf{z}_{1:t}|\widetilde{\mathbf{s}}_{1:t}^{(i)}(c))$ is Gaussian, of mean $\mu(\widetilde{\mathbf{s}}_{1:t}^{(i)}(c))$ and covariance matrix $P(\widetilde{\mathbf{s}}_{1:t}^{(i)}(c))$, and thus $p(\mathbf{z}_{1:t}|c)$ is a finite mixture of Gaussian pdfs, which falls in the framework described in Section 4.3. The pdfs $p(\mathbf{z}_{1:t}|\widetilde{\mathbf{s}}_{1:t}^{(i)}(c))$ can be recurrently computed with

$$p(\mathbf{z}_{1:t}|\widetilde{\mathbf{s}}_{1:t}^{(i)}(c)) = p(\mathbf{z}_{1:t-1}|\widetilde{\mathbf{s}}_{1:t-1}^{(i)}(c))p(\mathbf{z}_t|\mathbf{z}_{1:t-1}, \widetilde{\mathbf{s}}_{1:t}^{(i)}(c)) \tag{47}$$

where $p(\mathbf{z}_t|\mathbf{z}_{1:t-1}, \widetilde{\mathbf{s}}_{1:t}^{(i)}(c)) = \mathcal{N}(\mathbf{z}_t : \widehat{\mathbf{z}}_{t|t-1}(\widetilde{\mathbf{s}}_{1:t}^{(i)}(c)), S_t(\widetilde{\mathbf{s}}_{1:t}^{(i)}(c)))$ and $\widehat{\mathbf{z}}_{t|t-1}(\widetilde{\mathbf{s}}_{1:t}^{(i)}(c))$ and $S_t(\widetilde{\mathbf{s}}_{1:t}^{(i)}(c))$ are recurrently computed with the Kalman filter. To compute the plausibility, we need to get the quadratic scalar value

$$r_t(\mathbf{z}_{1:t}, \widetilde{\mathbf{s}}_{1:t}^{(i)}(c)) = (\mathbf{z}_{1:t} - \mu(\widetilde{\mathbf{s}}_{1:t}^{(i)}(c)))^{\mathrm{T}} P(\widetilde{\mathbf{s}}_{1:t}^{(i)}(c))^{-1}(\mathbf{z}_{1:t} - \mu(\widetilde{\mathbf{s}}_{1:t}^{(i)}(c)))$$

From Eq. (47), it is recurrently obtained by (see the proof in appendix C)

$$\begin{aligned} r_t(\mathbf{z}_{1:t}, \widetilde{\mathbf{s}}_{1:t}^{(i)}(c)) = {} & r_{t-1}(\mathbf{z}_{1:t-1}, \widetilde{\mathbf{s}}_{1:t-1}^{(i)}(c)) \\ & + (\mathbf{z}_t - \widehat{\mathbf{z}}_{t|t-1}(\widetilde{\mathbf{s}}_{1:t}^{(i)}(c)))^{\mathrm{T}} S_t(\widetilde{\mathbf{s}}_{1:t}^{(i)}(c))^{-1}(\mathbf{z}_t - \widehat{\mathbf{z}}_{t|t-1}(\widetilde{\mathbf{s}}_{1:t}^{(i)}(c))) \end{aligned} \tag{48}$$

Once we get $r_t(\mathbf{z}_{1:t}, \widetilde{\mathbf{s}}_{1:t}^{(i)}(c))$, we can use Eq. (32) to compute the plausibility

$$pl[c](\mathbf{z}_{1:t}) = 1 - \sum_{i=1}^{N} \overline{w}_t^{(i)} F_{\dim(\mathbf{z}_{1:t})+2}(r_t(\mathbf{z}_{1:t}, \widetilde{\mathbf{s}}_{1:t}^{(i)}(c))) \tag{49}$$

where F_p is the cumulative density function of a χ^2 distribution with p degrees of freedom, defined by Eq. (28). Although the dimension of the observation vector $\mathbf{z}_{1:t}$ increases with time, the class-conditioned plausibility is a scalar weighting sum of N terms where both the scalar point values $r_t(\mathbf{z}_{1:t}, \widetilde{\mathbf{s}}_{1:t}^{(i)}(c))$ and weights $\overline{w}_t^{(i)}$ are computed recurrently. From these plausibilities, one applies the GBT of Eq. (7) to obtain the bbas and finally the pignistic transform of Eq. (6) to compute the pignistic class probabilities.

The complete steps for joint tracking and classification are given in algorithm 1 on the following page. Particles $\widetilde{\mathbf{s}}_{1:t-1}^{(i)}(c)$ are extended from time $t-1$ to time t using the *importance distribution* denoted $q(s_t(c)|\mathbf{s}_{1:t-1}^{(i)}(c), c)$.

6.4 Numerical results

The numerical values used are those taken in [9].

The initial state is $\mathbf{x}_0 = [\,15km\ 220m/s\ 45km\ 10m/s\,]^T$. The trajectory consists of three constant velocity segments and two manoeuvres. The first turn is performed between scans 26 and 31 with acceleration $\mathbf{a} = [\,0\ 2.1g\,]$. The second turn is performed

Algorithm 1 Rao-Blackwellised particle filter for joint tracking and classification in the TBM framework

Initialization: For $c = 2, 3$ and for $i = 1, \ldots, N$

- Sample $s_0^{(i)}(c) \sim p_0(s_0)$
- Set $\left(\widehat{\mathbf{x}}_{0|0}(s_0^{(i)}(c)), \Sigma_{0|0}(s_0^{(i)}(c))\right) = (x_0, \Sigma_0)$
- Set $p(\mathbf{z}_0|s_0^{(i)}(c)) = 1$, $r_0(\mathbf{z}_0, s_0^{(i)}(c)) = 0$ and $w^{(i)}(c) = \frac{1}{N}$

Iterations: For $t = 1, 2, \ldots$

- For target class $c = 1$,
  - · Set $s_t(c) = 1$
  - % **Target tracking**
  - · Compute $\widehat{\mathbf{x}}_{t|t}(\mathbf{s}_{1:t}(c)), \Sigma_{t|t}(\mathbf{s}_{1:t}(c)), \widehat{\mathbf{z}}_{t|t-1}(\mathbf{s}_{1:t}(c))$ and $S_t(\mathbf{s}_{1:t}(c))$ with a Kalman filter step from $\widehat{\mathbf{x}}_{t-1|t-1}(\mathbf{s}_{1:t-1}(c))$ and $\Sigma_{t-1|t-1}(\mathbf{s}_{1:t-1}(c))$
  - % **Computation of plausibility**
  - · Compute $r_t(\mathbf{z}_{1:t}, \mathbf{s}_{1:t}(c))$ using Eq.(C.1) in appendix C
  - · Compute the plausibility $pl[c](\mathbf{z}_{1:t}) = 1 - F_{\dim(\mathbf{z}_{1:t})+2}\left(r_t(\mathbf{z}_{1:t}, \mathbf{s}_{1:t}(c))\right)$

- For each target class $c = 2, 3$
  - % **Target tracking**
  - · For $i = 1, \ldots, N$
    - Sample $\widetilde{s}_t^{(i)}(c)$ from the importance distribution $q(s_t(c)|\mathbf{s}_{1:t-1}^{(i)}(c), c)$
    - Compute $\widehat{\mathbf{x}}_{t|t}(\mathbf{s}_{1:t-1}^{(i)}(c), \widetilde{s}_t^{(i)}(c)), \Sigma_{t|t}(\mathbf{s}_{1:t-1}^{(i)}(c), \widetilde{s}_t^{(i)}(c)), \widehat{\mathbf{z}}_{t|t-1}(\mathbf{s}_{1:t-1}^{(i)}(c), \widetilde{s}_t^{(i)}(c))$ and $S_t(\mathbf{s}_{1:t-1}^{(i)}(c), \widetilde{s}_t^{(i)}(c))$ with a Kalman filter step from $\widehat{\mathbf{x}}_{t-1|t-1}(\mathbf{s}_{1:t-1}^{(i)}(c))$ and $\Sigma_{t-1|t-1}(\mathbf{s}_{1:t-1}^{(i)}(c))$
  - · For $i = 1, \ldots, N$, update the weights

$$\widetilde{w}_t^{(i)}(c) \propto \widetilde{w}_{t-1}^{(i)}(c)\mathcal{N}(\mathbf{z}_t : \mathbf{z}_{t|t-1}(\mathbf{s}_{1:t-1}^{(i)}(c), \widetilde{s}_t^{(i)}(c)), S_t(\mathbf{s}_{1:t-1}^{(i)}(c), \widetilde{s}_t^{(i)}(c))) \times \frac{\Pr(\widetilde{s}_t^{(i)}(c)|\mathbf{s}_{t-1}^{(i)}(c), c)}{q(\widetilde{s}_t^{(i)}(c)|\mathbf{s}_{1:t-1}^{(i)}(c), c)} \quad \text{with} \quad \sum_{i=1}^{N} \widetilde{w}_t^{(i)}(c) = 1.$$

  - · Compute the MMSE state estimate $\widehat{\mathbf{x}}_{t|t}^{\text{MMSE}}(c) = \sum_{i=1}^{N} \widetilde{w}_t^{(i)}(c)\widehat{\mathbf{x}}_{t|t}(\mathbf{s}_{1:t-1}^{(i)}(c), \widetilde{s}_t^{(i)}(c))$
  - · Resampling: Duplicate particles of high weight and delete those of low weight to obtain N new particles, named without $\widetilde{\cdot}$. Set $\widetilde{w}_t^{(i)} = \frac{1}{N}$.
  - % **Computation of plausibility**
  - · For $i = 1, \ldots, N$, compute

$$p(\mathbf{z}_{1:t}|\mathbf{s}_{1:t}^{(i)}(c)) = p(\mathbf{z}_{1:t-1}|\mathbf{s}_{1:t-1}^{(i)}(c))\mathcal{N}\left(\mathbf{z}_t : \widehat{\mathbf{z}}_{t|t-1}(\mathbf{s}_{1:t}^{(i)}(c)), S_t(\mathbf{s}_{1:t}^{(i)}(c))\right)$$

  - · For $i = 1, \ldots, N$, compute $r_t(\mathbf{z}_{1:t}, \widetilde{\mathbf{s}}_{1:t}^{(i)}(c))$ using Eq.(48)
  - · For $i = 1, \ldots, N$, compute $\overline{w}_t^{(i)}(c) \propto \frac{\widetilde{w}_t^{(i)}(c)}{p(\mathbf{z}_{1:t}|\widetilde{\mathbf{s}}_{1:t}^{(i)}(c))}$ with $\sum_{i=1}^{N} \overline{w}_t^{(i)}(c) = 1$
  - · Compute the plausibility $pl[c](\mathbf{z}_{1:t}) = 1 - \sum_{i=1}^{N} \overline{w}_t^{(i)} F_{\dim(\mathbf{z}_{1:t})+2}\left(r_t(\widetilde{\mathbf{s}}_{1:t}^{(i)}(c))\right)$

% **Target classification within the TBM framework**

- Apply the GBT from $pl[c = 1](\mathbf{z}_{1:t})$, $pl[c = 2](\mathbf{z}_{1:t})$ and $pl[c = 3](\mathbf{z}_{1:t})$ to obtain $m[\mathbf{z}_{1:t}](A)$, $\forall A \subseteq \{1, 2, 3\}$
- Apply the pignistic transform from $m[\mathbf{z}_{1:t}](A)$ to obtain $BetP(c|\mathbf{z}_{1:t})$ for $c = 1, 2, 3$

between scans 53 and 58 with acceleration $\mathbf{a} = [\,0\ -4.2g\,]$. The target trajectory is shown in Figure 9. The parameters are $T = 3s$, $Q = I_2$, $R = 10^4 I_2$, $\widehat{\mathbf{x}}_{0|0} = \mathbf{x}_0$, $\Sigma_{0|0} = 10^4 I_4$, and $N = 250$, where I_p is the $p \times p$ identity matrix. The importance distribution $q(s_t(c)|\mathbf{s}^{(i)}_{1:t-1}(c), c)$ used to propagate particles is the evolution Markov model $\Pr(s_t(c)|\mathbf{s}^{(i)}_{1:t-1}(c), c)$.

Plausibilities $pl[c](\mathbf{z}_{1:t})$ for each class c=1,2,3 are represented in Fig. 10. The pignistic probabilities $BetP(c|\mathbf{z}_{1:t})$ for target classes $c = 1, 2, 3$ used for decision making are shown in Fig. 11. The belief classifier is undecided between the three classes during the first CV segment. After the first turn, the pignistic probability of class 1 is zero, and the classifier is undecided between class 2 and 3. Finally, after the second turn, the pignistic probability of class 3 is almost equal to 1. These results make more sense than the Bayesian ones that are reported in [9]. In this paper, it has been shown that the Bayesian classifier tends to classify the target in class 1 during the first CV segment, and in class 2 during the second.

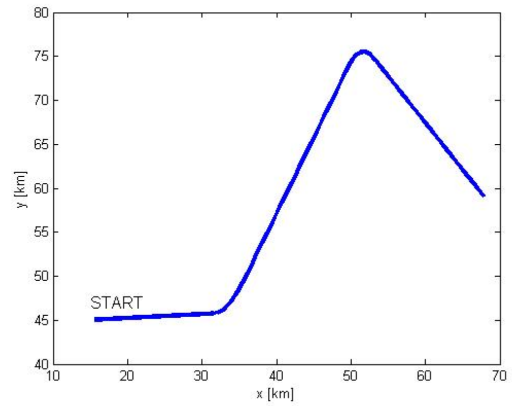

Fig. 9. Target trajectory. The trajectory consists of three constant velocity segments and two manoeuvres. The first turn is performed between scans 26 and 31 with acceleration $\mathbf{a} = [0\ 2.1g]^T$. The second turn is performed between scans 53 and 58 with acceleration $\mathbf{a} = [0\ \ -4.2g]^T$.

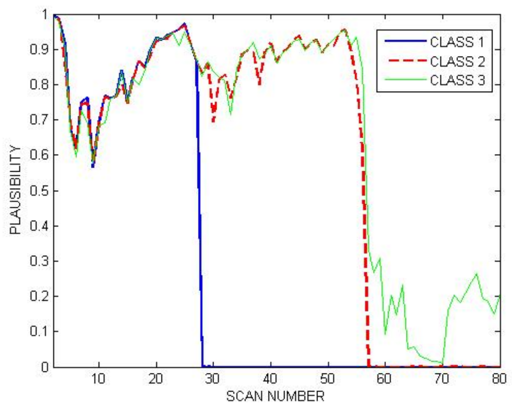

Fig. 10. Plausibility $pl[c](\mathbf{z}_{1:t})$ for each class c=1,2,3. $pl[c = 1](\mathbf{z}_{1:t})$ goes to 0 after the first turn (scans 26-31) and $pl[c = 2](\mathbf{z}_{1:t})$ goes to 0 after the second turn (scans 53-58).

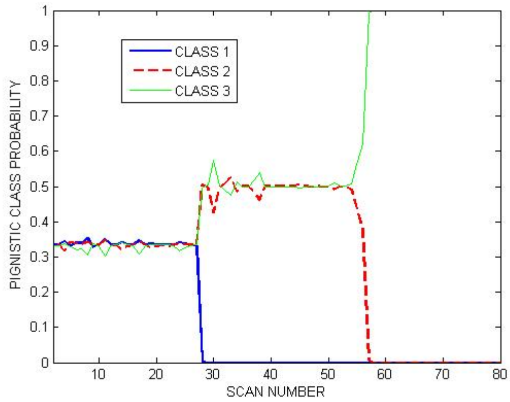

Fig. 11. Pignistic probability $BetP(c|\mathbf{z}_{1:t})$ for target classes $c = 1, 2, 3$. Before the first turn (scans 26-31) the three classes are equiprobable. After this first turn and before the second turn (scans 53-58) the pignistic probability of class 1 is zero while two others have the same value. After the second turn, the pignistic probability of class 3 is almost 1.

7 Discussion

It has to be noted that in this paper, the frame of discernment is discrete while the observations take their values in a continuous domain. Therefore, any combination is made using usual finite TBM tools. The generalization to a continuous frame of discernment is not straightforward. Actually, the bbd defined by Equation (23) has focal elements enclosed by hyperellipses. In the general case, the combination of two such bbds gives a bbd whose focal elements are not enclosed by hyperellipses and invariance is not guaranteed. We do not have this problem here because GBT is directly applied to the q-LC bbd and any combination is made in the discrete frame of discernment.

With the approach proposed in this paper, one has to adhere to a few assumptions:

- The available conditional probability distributions are viewed as pignistic probabilities from some unknown belief density function
- This underlying belief density is calculated according to the least commitment principle
- After the GBT is applied, the resulting belief assignment in the discrete domain is transformed into a probability mass according to the pignistic transform

The proposed approach starts from the same models of class-conditioned feature densities as the Bayesian, when considering multi-dimensional pdfs. However, it treats them as subjective (pignistic) rather than true models. The results obtained by these two approaches can be strikingly different, as illustrated by numerical examples. Simulations performed show that, in general, the Bayesian classifier tends to make quick decisions (whether right or wrong), while the proposed approach is more cautious (which is a characteristic of the least commitment principle), as already pointed out in [8].

8 Conclusion

This paper presented a generalization of the least committed bbd of a Gaussian pdf defined by Smets [2], for multivariate Gaussian pdfs. In this formulation the masses are assigned to hyperspheres, and the bbd is expressed as a well-known χ^2 pdf, allowing easy computations; more precisely, the plausibility of a point mass is simply the cumulative density function of a χ^2 pdf. These equations have also been extended to mixtures of Gaussians, allowing to handle multivariate non Gaussian pdfs.

Two applications of the proposed formulation have been presented. The first deals with model based target classification (similar to the one used in [7] and [8]). However, having a tool to deal with multivariate pdfs, in this paper we considered a more realistic case of two-dimensional correlated target feature measurements of *speed* and *acceleration*. The second application was devoted to the problem of joint target tracking and classification. The tracking part was carried out using a Rao-Blackwellised particle filter, while the classification part was performed within the TBM framework. As in [7] and [8], the belief function classifier gives arguably more meaningful results than the Bayesian one.

A Proof of Theorem 1

The pignistic function is defined as

$$Betf(x \in \mathbb{R}) = \int_{\{u|x \in I_u\}} \frac{\theta(u)}{l_u} du$$

where $l_u = 2\sqrt{u}\sigma$ is the length of the interval I_u. As we consider nested intervals, this relation is expressed as

$$Betf(x \in \mathbb{R}) = \int_{u=\left(\frac{x-\mu}{\sigma}\right)^2}^{\infty} \frac{\theta(u)}{2\sqrt{u}\sigma} du$$

Differentiation with respect to $\alpha = \left(\frac{x-\mu}{\sigma}\right)^2$ gives

$$\frac{\partial Betf(x \in \mathbb{R})}{\partial \alpha} = -\frac{\theta(\alpha)}{2\sqrt{\alpha}\sigma}$$

We also have that $Betf(x \in \mathbb{R}) = \mathcal{N}(x : \mu, \sigma^2)$; thus

$$\frac{\partial Betf(x \in \mathbb{R})}{\partial \alpha} = \frac{1}{\sqrt{2\pi}\sigma} \times (-\frac{1}{2}) \exp(-\frac{1}{2}\alpha)$$

It then follows that

$$\theta(\alpha) = \frac{\sqrt{\alpha}}{\sqrt{2\pi}} \exp(-\frac{1}{2}\alpha),\ \alpha \geq 0$$

B Proof of Theorem 2

To derive the pignistic probability $Betf(\mathbf{x})$ associated to the bbd m, we integrate on all the sets HV_α which contain x, i.e. the sets HV_α such that $\alpha \geq (\mathbf{x}-\mu)^T\Sigma^{-1}(\mathbf{x}-\mu)$, each belief mass $m(HV_\alpha)$ being divided by its volume V_α. Therefore we obtain

$$\begin{aligned} Betf(\mathbf{x} \in \mathbb{R}^n) &= \int_{u=(\mathbf{x}-\mu)^T\Sigma^{-1}(\mathbf{x}-\mu)}^{\infty} \frac{\theta(u)}{V_u} du \\ &= \int_{u=(\mathbf{x}-\mu)^T\Sigma^{-1}(\mathbf{x}-\mu)}^{\infty} \frac{\theta(u)}{\frac{\pi^{\frac{n}{2}} u^{\frac{n}{2}} \sqrt{\det(\Sigma)}}{\Gamma(\frac{n}{2}+1)}} du \end{aligned} \tag{B.1}$$

By differentiation with respect to $\alpha = (\mathbf{x}-\mu)^T\Sigma^{-1}(\mathbf{x}-\mu)$ one obtains

$$\frac{\partial Betf(\mathbf{x} \in \mathbb{R}^n)}{\partial \alpha} = -\frac{\theta(\alpha)}{\frac{\pi^{\frac{n}{2}} \alpha^{\frac{n}{2}} \sqrt{\det(\Sigma)}}{\Gamma(\frac{n}{2}+1)}} \tag{B.2}$$

We assume that $\mathcal{N}(\mathbf{x} : \mu, \Sigma)$ is the pignistic probability induced by the bbd m and thus

$$Betf(\mathbf{x} \in \mathbb{R}^n) = \frac{1}{(2\pi)^{\frac{n}{2}} \sqrt{\det(\Sigma)}} \exp(-\frac{1}{2}(\mathbf{x}-\mu)^T\Sigma^{-1}(\mathbf{x}-\mu)) \tag{B.3}$$

Differentiation of (B.3) w.r.t. $\alpha = (\mathbf{x}-\mu)^T\Sigma^{-1}(\mathbf{x}-\mu)$ we obtain

$$\frac{\partial Betf(\mathbf{x} \in \mathbb{R}^n)}{\partial \alpha} = \frac{1}{(2\pi)^{\frac{n}{2}} \sqrt{\det(\Sigma)}} (-\frac{1}{2}) \exp(-\frac{1}{2}\alpha) \tag{B.4}$$

and the final result follows

$$\theta(\alpha) = \frac{\alpha^{\frac{n}{2}}}{2^{\frac{n}{2}+1}\Gamma(\frac{n}{2}+1)} \exp(-\frac{1}{2}\alpha). \tag{B.5}$$

C Proof of Eq. (48)

Suppose that [4]

$$\begin{aligned} p(\mathbf{z}_{1:t-1}|\mathbf{s}_{1:t-1}) &= \mathcal{N}(\mathbf{z}_{1:t-1} : \mu(\mathbf{s}_{1:t-1}), P(\mathbf{s}_{1:t-1})) \\ &= \frac{1}{|2\pi P(\mathbf{s}_{1:t-1})|^{-\frac{1}{2}}} \exp\left(-\frac{1}{2}(\mathbf{z}_{1:t-1} - \mu(\mathbf{s}_{1:t-1}))^T P(\mathbf{s}_{1:t-1})^{-1}(\mathbf{z}_{1:t-1} - \mu(\mathbf{s}_{1:t-1}))\right) \\ &= \frac{1}{|2\pi P(\mathbf{s}_{1:t-1})|^{-\frac{1}{2}}} \exp(-\frac{1}{2} r_{t-1}(\mathbf{z}_{1:t-1}, \mathbf{s}_{1:t-1})) \end{aligned}$$

[4] Superscripts (i) and class-conditioning are omitted for the sake of clarity

and

$$p(\mathbf{z}_t|\mathbf{z}_{1:t-1},\mathbf{s}_{1:t}) = \mathcal{N}(\mathbf{z}_t : \widehat{\mathbf{z}}_{t|t-1}(\mathbf{s}_{1:t}), S_t(\mathbf{s}_{1:t})))$$
$$= \frac{1}{|2\pi S_t(\mathbf{s}_{1:t})|^{-\frac{1}{2}}} \exp\left(-\frac{1}{2}(\mathbf{z}_t - \widehat{\mathbf{z}}_{t|t-1}(\mathbf{s}_{1:t}))^T S_t(\mathbf{s}_{1:t})^{-1}(\mathbf{z}_t - \widehat{\mathbf{z}}_{t|t-1}(\mathbf{s}_{1:t}))\right)$$

Then $p(\mathbf{z}_{1:t}|\mathbf{s}_{1:t}) = p(\mathbf{z}_{1:t-1}|\mathbf{s}_{1:t-1})p(\mathbf{z}_t|\mathbf{z}_{1:t-1},\mathbf{s}_{1:t})$ is also a Gaussian pdf of mean $\mu(\mathbf{s}_{1:t})$ and covariance matrix $P(\mathbf{s}_{1:t})$, defined by

$$p(\mathbf{z}_{1:t}|\mathbf{s}_{1:t}) = \frac{1}{|2\pi P(\mathbf{s}_{1:t-1})|^{-\frac{1}{2}}} \exp(-\frac{1}{2} r_t(\mathbf{z}_{1:t},\mathbf{s}_{1:t}))$$

By applying Eq. (47), the exponent part gives

$$r_t(\mathbf{z}_{1:t},\mathbf{s}_{1:t}) = r_{t-1}(\mathbf{z}_{1:t-1},\mathbf{s}_{1:t-1}) + (\mathbf{z}_t - \widehat{\mathbf{z}}_{t|t-1}(\mathbf{s}_{1:t}))^T S_t(\mathbf{s}_{1:t})^{-1}(\mathbf{z}_t - \widehat{\mathbf{z}}_{t|t-1}(\mathbf{s}_{1:t})) \quad \text{(C.1)}$$

References

[1] P. Smets and R. Kennes. The Transferable Belief Model. *Artificial intelligence*, 66:191–234, 1994.

[2] P. Smets. Belief functions on real numbers. *International journal of approximate reasoning*, 40(3):181–223, 2005.

[3] P. Smets and B. Ristic. Kalman filter and joint tracking and classification based on belief functions in the TBM framework. *Information Fusion*, 2005. In Press.

[4] F. Caron, P. Smets, E. Duflos, and P. Vanheeghe. Multisensor data fusion in the frame of the TBM on reals. Application to land vehicle positioning. In *International Conference on Information Fusion*, Philadelphia, 2005.

[5] P. Smets. Belief functions: the disjunctive rule of combination and the generalized Bayesian theorem. *International Journal of Approximate Reasoning*, 9:1–35, 1993.

[6] F. Delmotte and P. Smets. Target identification based on the transferable belief interpretation of Dempster-Shafer model. *IEEE Transactions on Systems, Man and Cybernetics*, 34(4):457–471, 2004.

[7] B. Ristic and P. Smets. Belief function theory on the continuous space with an application to model based classification. In *IPMU 2004*, pages 1119–1126, Paris, 2004.

[8] B. Ristic and P. Smets. Target classification approach based on the belief function theory. *IEEE Transactions on Aerospace and Electronic Systems*, 41(2):574–583, 2005.

[9] B. Ristic, N. Gordon, and A. Bessell. On target classification using kinematic data. *Information Fusion*, 5(4):15–21, 2004.

[10] D. Dubois, H. Prade, and P. Smets. New semantics for quantitative possibility theory. In *ISIPTA*, Ithaca, NY, 2001.

[11] D. Dubois and H. Prade. *The principle of minimum specificity as a basis for evidential reasoning*, pages 75–84. Springer-Verlag, 1987.

[12] P. Smets. Quantified epistemic possibility theory seen as a hyper cautious Transferable Belief Model. In *Rencontres francophones sur la logique floue et ses applications*, pages 343–353, 2000.

[13] M. Abramowitz and I.A. Stegun. *Handbook of mathematical functions, with formulas, graphs, and mathematical tables.* Dover: NewYork, 1965.

[14] S. Blackman and R. Popoli. *Design and analysis of modern tracking systems.* Artech House, 1999.

[15] A. Doucet, S.J. Godsill, and C. Andrieu. On sequential monte carlo sampling methods for Bayesian filtering. *Statistical Computing*, 10(3):197–208, 2000.

[16] R. Chen and J.S. Liu. Mixture Kalman filters. *J. R. Statist. Soc. B*, 62(3):493–508, 2000.

[17] A. Doucet, N.J. Gordon, and V. Krishnamurthy. Particle filters for state estimation of jump Markov linear systems. *IEEE Transactions on Signal Processing*, 49(3):613–624, 2001.

[18] N. de Freitas. Rao-Blackwellised particle filtering for fault diagnosis. In *Proc. Aerospace Conference*, 2002.

[19] Y. Bar-Shalom, X. Rong Li, and T. Kirubajan. *Estimation with applications to tracking and navigation.* Editions Wiley-Interscience, 2001.

[20] N.J. Gordon, D.J. Salmond, and A.F.M. Smith. A novel approach to nonlinear/non-Gaussian Bayesian state estimation. In *Proc. Inst. Elect. Eng. Radar Signal Process.*, volume 140, pages 107–113, 1993.

[21] A. Doucet, N. de Freitas, and N. Gordon, editors. *Sequential Monte Carlo Methods in practice.* Springer-Verlag, 2001.

[22] B. Ristic, S. Arulampalam, and N. Gordon. *Beyond the Kalman Filter: Particle Filters for Tracking Applications.* Artech House, 2004.