Mixture of Uniform Probability Density Functions for non Linear State Estimation using Interval Analysis

Amadou Gning, Lyudmila Mihaylova
Department of Communication Systems,
InfoLab21, South Drive, Lancaster University
Lancaster LA1 4WA, United Kingdom
{e.gning, mila.mihaylova}@lancaster.ac.uk

Fahed Abdallah
HEUDIASYC, UMR CNRS 6599,
Université de Technologie de Compiègne
60205 Compiègne Cedex, France
fahed.abdallah@hds.utc.fr

Abstract – *In this work, a novel approach to nonlinear non-Gaussian state estimation problems is presented based on mixtures of uniform distributions with box supports. This class of filtering methods, introduced in* [1] *in the light of interval analysis framework, is called* Box Particle Filter (BPF). *It has been shown that weighted boxes, estimating the state variables, can be propagated using interval analysis tools combined with Particle filtering ideas. In this paper, in the light of the widely used Bayesian inference, we present a different interpretation of the BPF by expressing it as an approximation of posterior probability density functions, conditioned on available measurements, using mixture of uniform distributions. This interesting interpretation is theoretically justified. It provides derivation of the BPF procedures with detailed discussions.*

Keywords: Non linear System, Bayesian Filters, Uniform distribution, Monte Carlo Methods, Kalman Filters, Interval Analysis.

1 Introduction

The problem of state estimation of complex nonlinear stochastic systems, in the presence of non-Gaussian noisy measurements, has been of paramount importance for many application areas and has been a subject of intensive research. The most popular approach to tackle this problem is the Bayesian inference approach which aims are to provide an approximation of the posterior probability density function (pdf) of the state of interest conditioned on the available measurements. Among the Bayesian approaches, the Extended Kalman Filter (EKF) is a popular approach used for nonlinear system filtering [10]. This approach is based on applying the Kalman Filter (KF) algorithm on linearised state and measurement functions of the state model by using a first-order Taylor series expansion. At each time period, the state's posterior pdf is approximated using a Gaussian distribution which is propagated analytically through the first-order linearisation of the nonlinear system. Instead of linearising using Jacobian matrices, the Unscented Kalman Filter (UKF) uses a deterministic sampling strategy to capture the mean and covariance with a reasonably small set of carefully selected points called *sigma points* [13]. However, both the EKF and the UKF are limited to Gaussian and unimodal type of pdfs. A nice extension to more sophisticated and multimodal pdf is introduced in [3] by using mixtures of Gaussians. Using the approximation theory [5], it can be shown that the Gaussian family function is dense in the space of continuous functions and therefore, most of the Kalman type filters have been extended to mixtures of Gaussians.

Particle Filtering (PF) methods have recently emerged as a useful tool for more complex problems, in the presence of high nonlinearities or/and high dimensional systems, where the propagation of Gaussian mixtures is not always feasible. PF [4, 6] is a sequential Monte Carlo Bayesian estimator which represents the posterior pdf through sampling. Nevertheless, PF methods suffer from some drawbacks. These methods are very sensitive to unconsistent measurements or high measurement errors. In fact, the efficiency and accuracy of PFs depend mostly on the number of particles, and on the proposal functions used for the importance sampling. For instance, if the imprecision in the available information is high, the number of particles has to be chosen very large in order to explore a significant part of the state space and this fact induces real-time implementation issues. Several works attempt to use statistical approaches in order to overcome these shortcomings e.g., in [8], the sample set is dynamicallly updated during the estimation process.

The interval analysis framework is a promising methodology to model measurements with unknown or complex statistical bounded errors. Initially introduced to propagate rounding errors in mathematical computations earlier in the fifties [15], applications to state estimations have been recently investigated. For instance, in [12, 11] and [9] bounded-error observers, based on a predictor/estimator mechanism combined with various well known interval analysis, have been proposed. These approaches are different to the classical estimation methods, since the main objective is to estimate dynamically, in a guaranteed way, optimised boxes containing the hidden states.

In [1], again the interval framework is used for state estimation, but without focusing on obtaining estimates inside guaranteed boxes. A Box Particle Filter (BPF) is proposed consisting of propagated weighted boxes in a sequential way. Instead of using point samples, weighted boxes are used to approximate desired moments of the posterior pdfs. The key idea of the BPF stems from two possible understandings or interpretations of boxes : i) a box represents an infinite number of particles continuously distributed within the box, ii) a box represents a particle imprecisely located in the box. In this paper, the first interpretation is of interest. Each box is considered to be the support of an unknown pdf. In particular, the uniform pdf is a candidate that fits well with the BPF, since a common interpretation for a box enclosing a solution set is that all possible points inside the box have the same probability of belonging to the solution set. By deriving the BPF as a mixture of uniform pdfs, we show how the BPF fits to the Bayesian methodology. More over providing a sequential sampling method for the posterior pdf, the BPF also provides an approximation of the posterior pdf, in a similar way to the Gaussian mixture method.

The article is organised as follows. Section 2 formulates the principle of Bayesian filtering for nonlinear models. Section 3 briefly presents interval analysis principles and some relevant tools used in the bounded error approaches are introduced. Next, the BPF ideas and principles are given in Section 4. Results of the BPF applied to a dynamic localisation problem using GPS data merged to gyro and odometric data are also shown. The main contribution of this paper is presented in Section 5. The new Bayesian filter based on the BPF interpretation to a uniform pdfs mixture is presented. Finally, in Section 6 summarises our conclusions on the new Bayesian nonlinear state estimation method.

2 Bayesian Filtering

Consider the following system:

$$\left\{ \begin{array}{ll} \mathbf{x}_{k+1} = & f(\mathbf{x}_k, \nu_{k+1}) \\ \mathbf{y}_{k+1} = & h(\mathbf{x}_{k+1}, \omega_{k+1}) \end{array} \right. \quad (1)$$

where $f : \mathbb{R}^{n_x} \times \mathbb{R}^{n_\nu} \longrightarrow \mathbb{R}^{n_x}$ is a possibly non-linear transition function defining the state vector at time $k+1$ from the previous state $\mathbf{x}_k$ and from ν_k an independent identically distributed (iid) process noise sequence; n_x and n_ν denote, respectively, the dimensions of the state, the input and process noise vectors. Note that for simplicity an entry $\mathbf{u}_k$ is not added in the model.

The function $h : \mathbb{R}^{n_x} \times \mathbb{R}^{n_\omega} \longrightarrow \mathbb{R}^{n_y}$ defines the relation between the state and measurement vectors at time k, ω_k is an iid measurement noise sequence; n_y, n_ω are, respectively, dimensions of the measurement and measurement noise vectors. The states and the measurements up to time k are represented, respectively, by $\mathbf{X}_k = \{\mathbf{x}_i, i = 1, \cdots, k\}$ and $\mathbf{Y}_k = \{\mathbf{y}_i, i = 1, \cdots, k\}$.

In the Bayesian inference context, given the measurements $\mathbf{Y}_{k+1}$ up to time $k+1$, the posterior $p(\mathbf{X}_{k+1}|\mathbf{Y}_{k+1})$ provides a complete description of the state up to time instant $k+1$. In many applications, the marginal of the posterior pdf $p(\mathbf{x}_{k+1}|\mathbf{Y}_{k+1})$, at time instant $k+1$, also provides sufficient information and is given by:

$$p(\mathbf{x}_{k+1}|\mathbf{Y}_{k+1}) = \frac{1}{\alpha_{k+1}} p(\mathbf{x}_{k+1}|\mathbf{Y}_k) p(\mathbf{y}_{k+1}|\mathbf{x}_{k+1}), \quad (2)$$

$$p(\mathbf{x}_{k+1}|\mathbf{Y}_k) = \int p(\mathbf{x}_{k+1}|\mathbf{x}_k) p(\mathbf{x}_k|\mathbf{Y}_k) d\mathbf{x}_k, \quad (3)$$

where $\alpha_{k+1} = \int p(\mathbf{y}_{k+1}|\mathbf{x}_{k+1}) p(\mathbf{x}_{k+1}|\mathbf{Y}_k) d\mathbf{x}_{k+1}$ is a normalisation factor.

The recursion is initialised with a prior knowledge $p(x_0)$ e.g. a uniform pdf over some region of the state space. Equation (3) is the time update step and equation (2) is the measurement update step.

3 Elements of Interval Analysis

A real interval, $[x] = [\underline{x}, \overline{x}]$ is defined as a closed and connected subset of the set $\mathbb{R}$ of real numbers. In a vector form, a box $[\mathbf{x}]$ of $\mathbb{R}^{n_x}$ is defined as a Cartesian product of n_x intervals: $[\mathbf{x}] = [x_1] \times [x_2] \cdots \times [x_n] = \times_{i=1}^{n_x} [x_i]$. In this paper, the operator $|[.]|$ denotes the size $|[\mathbf{x}]|$ of a box $[\mathbf{x}]$. The underlying concept of interval analysis is to deal with intervals of real numbers instead of dealing with real numbers. For that purpose, elementary arithmetic operations, e.g., $+, -, *, \div$, etc., as well as operations between sets of $\mathbb{R}^n$, such as $\subset, \supset, \cap, \cup$, etc., have been naturally extended to interval analysis context. In addition, a lot of research has been performed with the so called *inclusion functions* [12, 2]. An *inclusion function* $[f]$ of a given function f is defined such that the image of a box $[\mathbf{x}]$ is a box $[f]([\mathbf{x}])$ containing $f([\mathbf{x}])$. The goal of inclusion functions is to work only with intervals, to optimise the interval enclosing the real image set and, then, to decrease the pessimism when intervals are propagated.

4 Box Particles Filtering

In this section the original idea of the BPF is briefly presented. Validation on real experiments detailed in [1] is also shown.

4.1 Sketch of the Box Particles Filtering

The aims of the BPF is to generalise particle filtering to the bounded error context. Instead of propagating

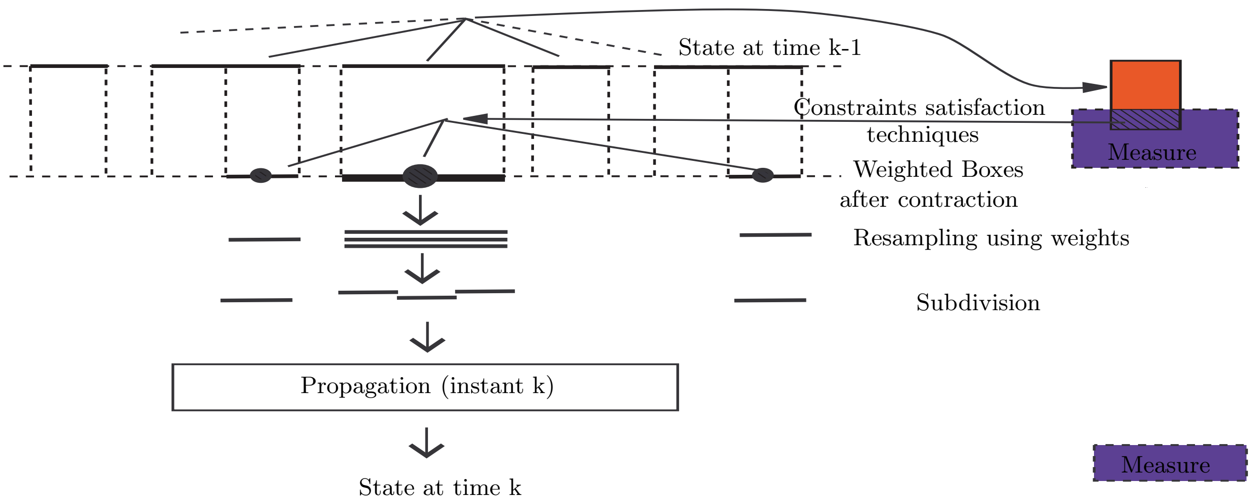

Figure 1: Scenarios of the Box Particle Filter.

weighted point samples of the posterior pdf, the key idea of the BPF is to propagate weighted box particles through bounded error models and using interval analysis. Figure 1 shows a picture of the BPF algorithm. The BPF steps can be described as follow:

- **Box particle initialisation**. This stage consists in splitting a prior bounded state space region into N mutually disjoint boxes $\{[\mathbf{x}^{(i)}]\}_{i=1}^{N}$ and associated weights for each of them. A first advantage, appearing with this initialisation using boxes, is the possibility of exploring a large initial prior region with few boxes particles.
- **Time update state step**. Using interval analysis tools, knowing the set of box particles $\{[\mathbf{x}_k^{(i)}]\}_{i=1}^{N}$ and assuming that the noise is known to be enclosed in $[\nu_{(k+1)}]$, the boxes at step $k+1$ are built using the propagation equations: $[\mathbf{x}_{k+1}^i] = [f]([\mathbf{x}_k^i], [\nu_{(k+1)}])$, where $[f]$ is an inclusion function for f (see Section 3). From this step, one interesting property of the BPF stems from the fact that, instead of propagating each particle using one realisation of the noise $\nu_{(k+1)}$, the uncertainty on the noise is also propagated for each box particle.
- **Measurement update**. As for the PF, the weights of the predicted box particles have to be updated using the new measurement at time $k+1$. For this purpose likelihood factors need to be calculated using innovation quantities. For the PF, the innovation for the N particles are $\mathbf{r}_{k+1}^i = \mathbf{y}_{k+1} - \mathbf{y}_{k+1}^i$, for $i = 1, \ldots, N$, where $\mathbf{y}_{k+1}^i = h(\mathbf{x}_{k+1}^i, \omega_{k+1}^i)$ is the i^{th} measurement simulation and ω_{k+1}^i is a noise realisation. For the BPF, the main difference consists in predicting for each box particle, a box enclosing the possible measurements which in turn will be compared to a box representing the bounded measurement $[\mathbf{y}_{k+1}]$ available at time $k+1$. The innovation for the i^{th} box particle should indicate the proximity between the measured box and the predicted box measurements. Thus, in the bounded error framework, it can be represented using the intersection between the two boxes. For all box particles, $i = 1 \cdots N$, the predicted box measurements have the expression $[\mathbf{y}_{k+1}^i] = [h]([\mathbf{x}_{k+1}^i])$, where $[h]$ is an inclusion function for h. The innovation is represented as: $[\mathbf{r}_{k+1}^i] = [\mathbf{y}_{k+1}^i] \cap [\mathbf{y}_{k+1}]$.

Next for the PF, using probabilistic models p_ω for the measurement noise ν, the likelihood of each particle is calculated as: $p(\mathbf{y}_{k+1}|\mathbf{x}_{k+1}^i) = p_\omega(\mathbf{y}_{k+1} - \mathbf{y}_{k+1}^i) = p_\omega(\mathbf{r}_{k+1}^i)$. For the BPF, in the bounded error context, the likelihood is calculated using the idea that a box particle with a corresponding predicted measurement without an intersection with the measured box has a likelihood factor equal to zero. In contrast, a box particle for which the predicted measure is included in the measured box has a likelihood close to one. This leads to a measure of the *box likelihood* in the form: $L_k^{(i)} = \prod_1^{n_y} L_k^{(i),j}$, where $L_k^{(i),j} = \frac{|[\mathbf{r}_{k+1}^i(j)]|}{|[\mathbf{y}_{k+1}^i(j)]|}$.

In addition, to the weights update, the BPF incorporates a new step. In the PF algorithm, each particle is propagated without any information about the variance of its position. In contrast, in the bounded error context, each box particle takes into account the imprecision caused by the model errors. This box correction is thus similar to the variance matrix measurement update step of Kalman filtering. Therefore, in order to preserve an appropriate size of each box, contraction algorithms [12] are performed which allows to eliminate the non consistent part of the box particles with respect to the measured box (see [1] for more details).

4.2 Application to a Dynamical Localisation Problem

In this section, a real experiment for localisation of a land vehicle is considered (see [1] for further details about the experiment). The mobile frame origin M is chosen at the middle of rear axle. The elementary rotation $\delta_{\theta,k}$ and the elementary displacement $\delta_{S,k}$ measurements between two time steps at the point M are obtained from a fiber optic gyro and the two rears wheels odometry sensors

$$\left\{\begin{array}{ll} \delta_{S,k} = & \frac{\delta_{RR,k}+\delta_{RL,k}}{2} \\ \delta_{\theta,k} = & \delta_{\theta,k}^{gyro} \end{array}\right. \tag{4}$$

where $\delta_{RR,k}$ and $\delta_{RL,k}$ denote respectively the measured right and left rear wheels displacements between two samples, and $\delta_{\theta,k}^{gyro}$ is a measure of the elementary rotation given by the gyro. The state $\mathbf{x}_k = x_k \times y_k \times \theta_k$ constituted by the position and the heading angle of the vehicle at time instant k, is propagated through the model

$$\left\{\begin{array}{ll} x_{k+1} = & x_k + \delta_{S,k}\cos(\theta_k + \frac{\delta_{\theta,k}}{2}) \\ y_{k+1} = & y_k + \delta_{S,k}\sin(\theta_k + \frac{\delta_{\theta,k}}{2}) \\ \theta_{k+1} = & \theta_k + \delta_{\theta,k} \end{array}\right. \tag{5}$$

The measurement consists in 2D position provided by a Global Position System (GPS) which is (x_{GPS}, y_{GPS}). The "longitude, latitude" estimated point of the GPS is converted in a Cartesian local frame. The GPS measurement box can be quantified using the standard deviation σ_{GPS} provided in real time by the GPS receiver. Thus,

$$\left\{\begin{array}{ll} [x_{GPS}] = & [x_{GPS} - 3\sigma_{GPS}, x_{GPS} + 3\sigma_{GPS}] \\ [y_{GPS}] = & [y_{GPS} - 3\sigma_{GPS}, y_{GPS} + 3\sigma_{GPS}] \end{array}\right. \tag{6}$$

The GPS measurement $([x_{GPS}], [y_{GPS}])$ is used to initialise the box state position $([x_1], [y_1])$ at instant t_1. The heading angle of the vehicle is initialised by $[\theta_1] = [-\infty, +\infty]$. In addition, a ground truth is available through a Thales Navigation GPS receiver used in a Post-Processed Kinematic mode, working at 1 Hz sampling rate with a local base (a Trimble 7400) and providing a few centimeters of accuracy. The data of the sensors have been time stamped and logged during several tests. We report hereafter the analysis of 4.7 Km path with a mean speed of 50 Km/h using a 3GHz Pentium 4 and a Matlab implementation. The two filters provide outputs at the frequency of the GPS (5Hz).

Table 1 shows the mean square error for the GPS alone, the BPF and PF. As a conclusion for this problem, the BPF and the PF give equivalent filtering performances. Nevertheless, for the BPF running, we use only 10 box particles comparing with 3000 particles for the PF. This is a good motivation for further developing the BPF since one can reduce significatively the particles numbers (for this application, the factor is about

	GPS	PF	BPF
mean square error for x(m)	0.134	0.129	0.119
mean square error for y(m)	0.374	0.217	0.242
particle number	-	3000	10
one step running time (ms)	-	666	149

Table 1: Comparison of PF and BPF. The table shows the mean square error for GPS, PF and BPF. The particle and box particle numbers are given for PF and BPF. We give also the mean of the running time of one step for each algorithm.

300). Table 1 gives the average one step computational time for each algorithm. Since the output frequency of each filter is 5 HZ, the running time for BPF satisfies real time constraints.

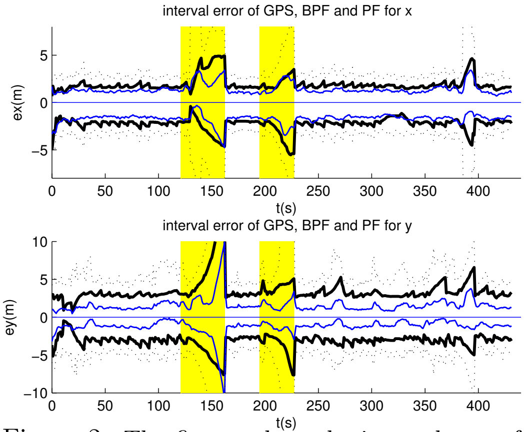

Figure 2: The figures show the interval error for x and y estimated for GPS (dashed black), BPF (bold black) and PF (solid blue).

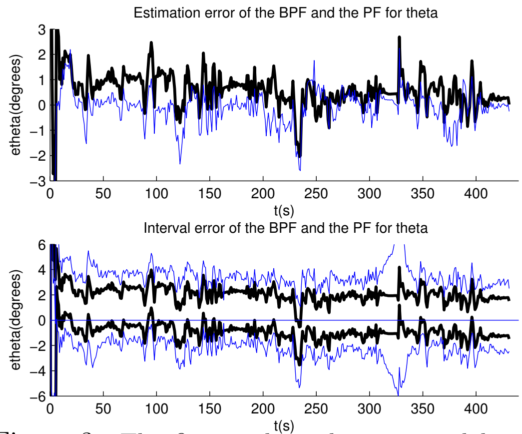

Figure 3: The figures show the estimated heading error and the interval errors, in degrees, for BPF (bold black) and PF (solid blue).

Figure 2 shows the interval error for x and y estimated for GPS (dashed black), BPF (bold black) and PF (solid blue). The coloured rectangles (yellow) correspond to the DGPS corrections lost periods of time. For PF, the interval error is calculated by using 3σ errors bounds around the point estimate. It can be seen that for this nonlinear problem, the two filters are consistent. Note that the interval errors contains "0" which is equivalent to an interval containing the ground truth value. Figure 3 plots the heading estimated error in addition to the interval errors, in degrees, for BPF (bold black) and PF (solid blue). The errors on the heading estimation angles provided by the BPF and the PF are of the same magnitude. One can conclude that the BPF is also able to observe a non directly measured variable.

5 Approximation Using a Mixture of Uniform PDFs

In the linear case, with a Gaussian prior and Gaussian independent noises, the Bayesian solution corresponds to the Kalman filter. However, an exact propagation of the posterior pdf using the two Bayesian steps (2) and (3) is unfortunately not feasible in general. Approximation of the posterior pdf using a weighted combination of a chosen family of pdf is a natural solution. Among these families, mixture of Gaussians [3] is the most popular approach since it has the advantages that each component of the mixture can be propagated using the well established Kalman steps. However, in the presence of strong non-linearities the propagation of the mixture is the main difficulty and is sometime intractable.

Uniform pdfs represent another attractive family. In addition to the natural simplicity of these pdfs, by choosing box supports, interval analysis theory offers a variety of tools to propagate these supports through linear, nonlinear functions and even through differential equations. Moreover, sum of uniform functions (or piecewise constant functions) with box supports has been widely used and is, for instance, the basis of Riemann integration theory [7]. One crucial result for using a sum of uniform pdfs as approximation of a continuous function is that, as the number of components increases toward infinity and the measure of the support tends to zero, any continuous real valued function defined in a compact space can be uniformly approximated by a piecewise constant function. In other words, similarly to the Gaussian family, the piecewise functions are dense in the space of continuous functions.

5.1 Time Update Step

First, let us denote $U_{[\mathbf{x}]}$ the uniform pdf with the box $[\mathbf{x}]$ as support. The uniform pdf sum representation of a random variable $\mathbf{x}$ is written as $p(\mathbf{x}) = \sum_{i=1}^{l} w^{(i)} U_{[\mathbf{x}^{(\mathbf{i})}]}(\mathbf{x})$, where l denotes the number of components, the $[\mathbf{x}^{(\mathbf{i})}]$ denote the box supports, the $w^{(i)}$ denote a set of normalised weights: $\sum_{i=1}^{l} w^{(i)} = 1$ and $\forall\, i,\ w^{(i)} \geq 0$.

Assume that, at time instant k, an approximation for the previous time pdf $p(\mathbf{x}_k|\mathbf{Y}_k)$ by a mixture of l_k uniform pdfs with box supports $[\mathbf{x}_\mathbf{k}^{(\mathbf{i})}]$ is available (7) and that the time update step is to be performed

$$p(\mathbf{x}_k|\mathbf{Y}_k) = \sum_{i=1}^{l_k} w_k^{(i)} U_{[\mathbf{x}_k^{(i)}]}(\mathbf{x}_k). \tag{7}$$

Inserting (7) into the time update equation (3) (introduced in Section 2) gives:

$$\begin{aligned} p(\mathbf{x}_{k+1}|\mathbf{Y}_k) &= \int p(\mathbf{x}_{k+1}|\mathbf{x}_k) \textstyle\sum_{i=1}^{l_k} w_k^{(i)} U_{[\mathbf{x}_k^{(i)}]}(\mathbf{x}_k) d\mathbf{x}_k \\ &= \textstyle\sum_{i=1}^{l_k} w_k^{(i)} \int p(\mathbf{x}_{k+1}|\mathbf{x}_k) U_{[\mathbf{x}_k^{(i)}]}(\mathbf{x}_k) d\mathbf{x}_k \\ &= \textstyle\sum_{i=1}^{l_k} w_k^{(i)} \frac{1}{|[\mathbf{x}_k^{(i)}]|} \int_{[\mathbf{x}_k^{(i)}]} p(\mathbf{x}_{k+1}|\mathbf{x}_k) d\mathbf{x}_k. \end{aligned} \tag{8}$$

Consider an inclusion function $[f]$ (see Section 3) for the transition model f, and let assume that the noise ν_{k+1}, at time instant $k+1$, is bounded in the box $[\nu_{k+1}]$. Then, by definition of the inclusion functions, $\forall\, i = 1, \ldots, l_k$, if $\mathbf{x}_k \in [\mathbf{x}_k^{(i)}]$ then $\mathbf{x}_{k+1} \in [f]([\mathbf{x}_k^{(i)}], [\nu_{k+1}])$. Thus, for all $i = 1\ \ldots, l_k$ we can write

$$p(\mathbf{x}_{k+1}|\mathbf{x}_k) U_{[\mathbf{x}_\mathbf{k}{}^{(i)}]}(\mathbf{x}_k) = 0 \quad \forall\, \mathbf{x}_{k+1} \notin [f]([\mathbf{x}_\mathbf{k}^{(\mathbf{i})}], [\nu_{k+1}]). \tag{9}$$

Equation (9) shows that for any transition function f, using interval analysis techniques, the support for the pdf terms $\int p(\mathbf{x}_{k+1}|\mathbf{x}_k) U_{[\mathbf{x}_k^{(i)}]}(\mathbf{x}_k) d\mathbf{x}_k$ can be approximated by $[f]([\mathbf{x}_\mathbf{k}^{(\mathbf{i})}], [\nu_{k+1}])$. In addition, it can be seen that, in the BPF algorithm, each pdf term $\int p(\mathbf{x}_{k+1}|\mathbf{x}_k) U_{[\mathbf{x}_k^{(i)}]}(\mathbf{x}_k) d\mathbf{x}_k$ in (8) is modeled using one uniform pdf component having as support the interval $[f]([\mathbf{x}_\mathbf{k}^{(\mathbf{i})}], [\mathbf{v}_{k+1}])$. The BPF strategy of approximating each pdf $\int p(\mathbf{x}_{k+1}|\mathbf{x}_k) U_{[\mathbf{x}_k^{(i)}]}(\mathbf{x}_k) d\mathbf{x}_k$ using one uniform pdf component seems to be not enough accurate (although it can be sufficient to approximate the first moments of the pdf as it has been shown experimentally in Section 4.2 and in [1] with more details). Alternately, a mixture of uniform pdfs can be used to properly approximate this pdf (note that for the Gaussian mixture method, similar suggestion is also performed in [3] with more than one component to properly model $p(\mathbf{x}_{k+1}|\mathbf{Y}_k)$ when the noise ν_{k+1} variance is large). For this purpose, for a general case, assume that the noise ν_{k+1} pdf is also approximated with a mixture of uniform pdfs: $p(\nu_{k+1}) = \sum_{i=1}^{q_k} \lambda_{k+1}^{(i)} U_{[\nu_{k+1}^{(i)}]}(\nu_{k+1})$ (with only one component if the only information about the noise is that is can be bounded). The transition probability $p(\mathbf{x}_{k+1}|\mathbf{x}_k)$ can be further developed into:

$$\begin{aligned} &p(\mathbf{x}_{k+1}|\mathbf{x}_k) = \\ &\int p(\mathbf{x}_{k+1}|\mathbf{x}_k, \nu_{k+1}) \textstyle\sum_{i=1}^{q_k} \lambda_{k+1}^{(i)} U_{[\nu_{\mathbf{k+1}}^{(\mathbf{i})}]}(\nu_{k+1}) d\nu_{k+1} \\ &= \textstyle\sum_{i=1}^{q_k} \lambda_{k+1}^{(i)} \frac{1}{|[\nu_\mathbf{k}^{(\mathbf{i})}]|} \int_{[\nu_{\mathbf{k+1}}^{(\mathbf{i})}]} p(\mathbf{x}_{k+1}|\mathbf{x}_k, \nu_{k+1}) d\nu_{k+1}. \end{aligned} \tag{10}$$

Combining (8) and (10) leads to the expression

$$\begin{aligned} p(\mathbf{x}_{k+1}|\mathbf{Y}_k) = \textstyle\sum_{i=1}^{l_k} \sum_{j=1}^{q_k} w_k^{(i)} \lambda_{k+1}^{(j)} \frac{1}{|[\mathbf{x}_k^{(i)}]|} \frac{1}{|[\nu_{\mathbf{k+1}}^{(\mathbf{j})}]|}. \\ (\int_{[\mathbf{x}_k^{(i)}]} \int_{[\nu_{\mathbf{k+1}}^{(\mathbf{i})}]} p(\mathbf{x}_{k+1}|\mathbf{x}_k, \nu_{k+1}) d\mathbf{x}_\mathbf{k} d\nu_{k+1}). \end{aligned} \tag{11}$$

Furthermore, denote μ the Lebesgue measure on $\mathbb{R}^{n_x+n_\nu}$ and I_A the indicator function for a given set A. The term $\int_{[\mathbf{x}_k^{(i)}]}\int_{[\nu_{\mathbf{k+1}}^{(\mathbf{i})}]} p(\mathbf{x}_{k+1}|\mathbf{x}_k,\nu_{k+1})d\mathbf{x}_k d\nu_{k+1}$ in (11) can be written:

$$
\begin{aligned}
&\int_{[\mathbf{x}_k^{(i)}]}\int_{[\nu_{\mathbf{k+1}}^{(\mathbf{i})}]} p(\mathbf{x}_{k+1}|\mathbf{x}_k,\nu_{k+1})d\mathbf{x}_k d\nu_{k+1}\\
&=\int_{[\mathbf{x}_k^{(i)}]}\int_{[\nu_{\mathbf{k+1}}^{(\mathbf{i})}]} I_{f(d\mathbf{x}_k,d\nu_{k+1})}(\mathbf{x}_{k+1})d\mathbf{x}_k d\nu_{k+1}\\
&=\mu(\{(\mathbf{x}_k,\nu_{k+1})\in[\mathbf{x}_k]\times[\nu_{k+1}]\mid f(\mathbf{x}_k,\nu_{k+1})=\mathbf{x}_{k+1}\})
\end{aligned}
\tag{12}
$$

In order to approximate the term $\mu(\{(\mathbf{x}_k,\nu_{k+1})\in[\mathbf{x}_k]\times[\nu_{k+1}]\mid f(\mathbf{x}_k,\nu_{k+1})=\mathbf{x}_{k+1}\}$ in (12), let denote Φ the real valued function (13) defined on $\mathbb{R}^{n_x}$

$$\Phi(\mathbf{z})=\mu(\{(\mathbf{x},\nu)\in[\mathbf{x}]\times[\nu]\mid f(\mathbf{x},\nu)=\mathbf{z}\}). \tag{13}$$

Using interval analysis tools, the function Φ can be approximated by a sum of constant functions with box supports. For that purpose, let introduce two lemmas.

Lemma 5.1. *Consider a set Ξ of mutually disjoints boxes $([\mathbf{z}_i])_{i=1\ldots N}$ on $\mathbb{R}^{n_x}$ such that $\bigcup_{i=1}^N[\mathbf{z}_i]\supseteq f([\mathbf{x}_k],[\nu_{k+1}])$ and such that $\bigcup_{i=1}^N[\mathbf{z}_i]\subseteq[f]([\mathbf{x}_k],[\nu_{k+1}])$ with $[f]$ an inclusion function of f. Let us define on $\mathbb{R}^{n_x}$ the real valued function Φ^Ξ (14).*

$$
\begin{aligned}
&\forall\,\mathbf{z}\notin\textstyle\bigcup_{i=1}^N[\mathbf{z}_i],\quad \Phi^\Xi(\mathbf{z})=0,\\
&\text{for } i=1,\ldots,N \text{ and } \forall\,\mathbf{z}\in[\mathbf{z}_i]\\
&\quad\Phi^\Xi(\mathbf{z})=\tfrac{1}{|[\mathbf{z}_i]|}\mu(\{(\mathbf{x},\nu)\in[\mathbf{x}]\times[\nu]\mid f(\mathbf{x},\nu)\in[\mathbf{z}_i]\}.
\end{aligned}
\tag{14}
$$

When $\max_{i=1,\ldots,N}(|[\mathbf{z}_i]|)$ tends to zero (and, consequently, N also tends to infinity), the functions Φ^Ξ defined in (14) tend to Φ.

Proof. *See the Appendices, section A.*

Lemma 5.1 is the first step allowing to approximate the function Φ using piecewise constant functions Φ^Ξ. Next, the lemma 5.2 is also introduced allowing to design an algorithm, based on interval analysis, in order to approximate the function Φ^Ξ.

Lemma 5.2. *Consider a set Ξ of mutually disjoints boxes $([\mathbf{z}_i])_{i=1\ldots N}$ on $\mathbb{R}^{n_x}$ such that $\bigcup_{i=1}^N[\mathbf{z}_i]\supseteq f([\mathbf{x}_k],[\nu_{k+1}])$ and consider a set Λ of mutually disjoints boxes $([\mathbf{x}_i]\times[\nu_i])_{i=1\ldots M}$ on $\mathbb{R}^{n_x+n_\nu}$ such that $\bigcup_{i=1}^M([\mathbf{x}_i]\times[\nu_i])=[\mathbf{x}]\times[\nu]$. Then, let us define on $\mathbb{R}^{n_x+n_\nu}$ the real valued function $\Phi^{\Xi,\Lambda}$ such that*

$$
\begin{aligned}
&\forall\,\mathbf{z}\notin\textstyle\bigcup_{i=1}^N[\mathbf{z}_i],\quad \Phi^{\Xi,\Lambda}(\mathbf{z})=0\\
&\text{for } i=1,\ldots,N \text{ and } \forall\,\mathbf{z}\in[\mathbf{z}_i]\\
&\quad\Phi^{\Xi,\Lambda}(\mathbf{z})=\tfrac{1}{|[\mathbf{z}_i]|}\sum_{j\,\in\, I^{[\mathbf{z}_i]}}|[\mathbf{x}_j]\times[\nu_j]|,
\end{aligned}
\tag{15}
$$

where $I^{[\mathbf{z}_i]}$ is a set of indexes defined as: $I^{[\mathbf{z}_i]}=\{j\in\{1,\cdots,M\}\mid f([\mathbf{x}_j],[\nu_j])\subseteq[\mathbf{z}_i]\}$. When $\max_{j=1,\ldots,M}(|[\mathbf{x}_j]\times[\nu_j]|)$ tends to zero (and consequently M also tends to infinity), the functions $\Phi^{\Xi,\Lambda}$ defined in (15) tend to Φ^Ξ

Proof. *See the Appendices, section B.*

By appropriately choosing first the set Ξ and secondly the set Λ, the combination of lemma 5.1 and lemma 5.2 allows to approximate the desired function Φ at a level of accuracy that will increase with the number of boxes. As an illustration, consider a simple summation of two uniform pdfs with respectively $[1,2]$ and $[4,6]$ as supports. Figure 4 shows the approximation using $\Phi^{\Xi,\Lambda}$ introduced in the two lemmas, with respectively 20 and 2000 intervals in Λ and Ξ.

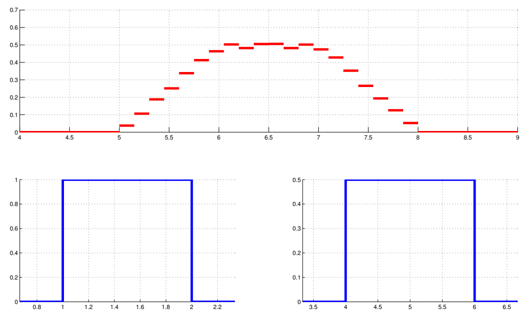

Figure 4: The figure shows a simple illustration of the time update step performed with a uniform pdfs propagation algorithm. At the top, the approximation of the resulting sum of the two uniform pdfs at the bottom.

5.2 Measurement Update Step

Let assume that, at time instant $k+1$, an approximation of the time update pdf $p(\mathbf{x}_{k+1}|\mathbf{Y}_k)$ by a mixture of $l_{k+1|k}$ uniform pdfs with interval supports $[\mathbf{x}_{\mathbf{k+1|k}}^{(\mathbf{i})}]$ and weights $w_{k+1|k}$ is available and that the measurement update step is to be performed. For the BPF algorithm, the measurement likelihood function is taken to be one component with uniform distribution. However, for a general case, a mixture can be considered. Assume that the likelihood function has the expression (16) with s_{k+1} components weighted with the normalised coefficient $\beta_{k+1}^{(j)}$

$$p(\mathbf{y}_{k+1}|\mathbf{x}_{k+1})=\sum_{j=1}^{s_{k+1}}\beta_{k+1}^{(j)}U_{[\mathbf{y}_{k+1}^{(j)}]}(h(\mathbf{x}_{k+1})), \tag{16}$$

where the $[\mathbf{y}_{k+1}^{(j)}]$ denote a set of box supports satisfying $\cup_{j=1}^{s_{k+1}}[\mathbf{y}_{k+1}^{(j)}]=[\mathbf{y}_{k+1}]$. The measurement update has the expression

$$
\begin{aligned}
p(\mathbf{x}_{k+1}|\mathbf{Y}_{k+1})&=\tfrac{1}{\alpha_{k+1}}p(\mathbf{y}_{k+1}|\mathbf{x}_{k+1})p(\mathbf{x}_{k+1}|\mathbf{Y}_k)\\
&=\tfrac{1}{\alpha_{k+1}}\textstyle\sum_{j=1}^{s_{k+1}}\beta_{k+1}^{(j)}U_{[\mathbf{y}_{k+1}^{(j)}]}(h(\mathbf{x}_{k+1}))\\
&\qquad\textstyle\sum_{i=1}^{l_{k+1|k}}w_{k+1|k}^{(i)}U_{[\mathbf{x}_{\mathbf{k+1|k}}^{(\mathbf{i})}]}(\mathbf{x}_{k+1})\\
&=\tfrac{1}{\alpha_{k+1}}\textstyle\sum_{j=1}^{s_{k+1}}\sum_{i=1}^{l_{k+1|k}}\beta_{k+1}^{(j)}w_{k+1|k}^{(i)}\\
&\qquad U_{[\mathbf{y}_{k+1}^{(j)}]}(h(\mathbf{x}_{k+1}))U_{[\mathbf{x}_{\mathbf{k+1|k}}^{(\mathbf{i})}]}(\mathbf{x}_{k+1}).
\end{aligned}
\tag{17}
$$

The term $U_{[\mathbf{y}_{k+1}^{(j)}]}(h(\mathbf{x}_{k+1}))U_{[\mathbf{x_{k+1|k}^{(i)}}]}(\mathbf{x}_{k+1})$ is also a constant function with a support being the set $\{\mathbf{x_{k+1}} \in [\mathbf{x_{k+1|k}^{(i)}}] \mid \exists\, \omega_{k+1} \in [\omega_{k+1}]$ such that $h(\mathbf{x}_{k+1}, \omega_{k+1}) \in [\mathbf{y}_{k+1}^{(j)}]\}$. We can deduce that, using consistency algorithms (see [12] for an introduction to these techniques), the predicted supports $[\mathbf{x_{k+1|k}^{(i)}}]$ for the time update pdf $p(\mathbf{x}_{k+1}|\mathbf{Y}_k)$ approximation have to be contracted with respect to the new measurement. These contraction steps give the new support for the posterior pdf $p(\mathbf{x}_{k+1}|\mathbf{Y}_{k+1})$ at time instant $k+1$. This is an interesting result since the contraction steps which have been heuristically introduced in the BPF are derived theoretically through the posterior pdf expression (17). We can see here that, using a Bayesian formulation through mixtures of uniform pdf, the BPF procedures are theoretically justified. Furthermore, these procedures can be extended in order to have a better approximation of the posterior pdfs. Indeed the weights update resulting from (17) is more correct and more elaborated than the BPF likelihood heuristic presented in Section (4). The time update is also, surprisingly, the most complicated step to derive and can be performed using the two lemmas introduced in Section 5.1.

6 Conclusions

In this paper, the BPF algorithm is studied as a Bayesian method through and interpretation using mixtures of uniform pdfs with box supports. This study provides theoretical justifications of the BPF procedures and, in addition, it is proven that mixture of uniform pdfs can be used to approximate the posteriors pdf for a state estimation problem. Furthermore, this study also provides more sophisticated procedures, to sequentially propagate and correct the mixture of uniform pdfs, in comparison with the original BPF procedures.

The method based on mixture of uniform pdfs can be classified between the PF method and the Gaussian mixture method. In general, a mixture of uniform pdfs needs more components than a mixture of Gaussians to approximate one given pdf but less components than a set of samples for a PF. In addition, the methods with mixtures of uniform pdfs share common property with both PFs and Gaussian mixture methods. For instance, for both PF and mixture of uniform pdfs with box supports, linearisation of the model is not necessary since interval analysis offers powerful tools to propagate intervals through continuous functions. On the other hand, mixture of uniform pdfs and mixture of Gaussian pdfs share the property of providing posterior pdfs approximation. Indeed, PFs are designed to provide samples allowing to approximate moments of the distribution but are not designed to provide a direct approximation of the posterior pdfs (this can be done, however, indirectly by using kernel distributions).

This study of the BPF, in the light of the Bayesian framework, opens numerous challenges. Current works are focused on using the new procedures in the prediction step and in the correction step, introduced in Section 5, to implement a more sophisticated non linear and real time filter. We expect a more consuming algorithm than the BPF but a better accuracy in addition to a direct approximation of the posterior pdfs as improvements of the BPF. We are also interested, as in the Gaussian mixture method, to introduce new steps such as merging of the mixture components in order to adapt the method to more complex problems.

Appendices

A Proof Lemma 5.1

The real valued function Φ^{Ξ} for a given set Ξ of mutually disjoints boxes $([\mathbf{z}_i])_{i=1\ldots N}$ is defined as:

$$\begin{aligned}&\Phi^{\Xi}(\mathbf{z}) = 0 \quad \forall\, \mathbf{z} \notin \textstyle\bigcup_{i=1}^{N}[\mathbf{z}_i], \text{ and}\\ &\text{for } i = 1,\ldots,N\ , \forall\, \mathbf{z} \in [\mathbf{z}_i],\\ &\Phi^{\Xi}(\mathbf{z}) = \tfrac{1}{||[\mathbf{z}_i]||}\mu(\{(\mathbf{x},\nu) \in [\mathbf{x}] \times [\nu] \mid f(\mathbf{x},\nu) \in [\mathbf{z}_i]\}.\end{aligned}$$

Since Φ is defined as $\Phi(z) = \mu(\{(\mathbf{x},\nu) \in [\mathbf{x}] \times [\nu] \mid f(\mathbf{x},\nu) = \mathbf{z}\})$, for $i = 1,\ldots,N$ and $\forall\, \mathbf{z} \in [\mathbf{z}_i]$ we can write:

$$\Phi^{\Xi}(\mathbf{z}) = \frac{1}{||[\mathbf{z}_i]||}\int_{[\mathbf{z}_i]}\Phi(\xi)d\xi.$$

In addition, Φ is a continuous function over $\mathbb{R}^{n_x}$ (Φ is derived from the equation (12) which is an integration of a transition probability). Since the support $f([\mathbf{x}],[\nu])$ of Φ is a compact, Φ is also uniformly continuous: $\forall \epsilon, \exists\, \eta \,|\, \forall\, z_1, z_2 \in \mathbb{R}^{n_x}$ with $||z_1, z_2|| < \eta, ||\Phi(z_1), \Phi(z_2)|| < \epsilon$ (here $||.||$ is the Euclidian norm). Then, if the $[\mathbf{z}_i]$ are chosen such that $\max\limits_{i=1,\ldots,N}(||[\mathbf{z}_i]|| < \eta)$:

$$\begin{aligned}&\text{for } i = 1,\ldots,N \text{ and } \forall\, \mathbf{z} \in [\mathbf{z}_i]\\ &||\Phi^{\Xi}(\mathbf{z}) - \Phi(\mathbf{z})|| = ||\tfrac{1}{||[\mathbf{z}_i]||}\textstyle\int_{[\mathbf{z}_i]}(\Phi(\xi) - \Phi(\mathbf{z}))d\xi||\\ &\leq \tfrac{1}{||[\mathbf{z}_i]||}\textstyle\int_{[\mathbf{z}_i]}||\Phi(\xi) - \Phi(\mathbf{z})||d\xi \leq \tfrac{1}{||[\mathbf{z}_i]||}\int_{[\mathbf{z}_i]}\epsilon d\xi \quad = \quad \epsilon.\end{aligned}$$

Finally, considering the hypothesis that the sets $([\mathbf{z}_i])_{i=1\ldots N}$ are mutually disjoints and that the common support $f([\mathbf{x}_k],[\nu_{k+1}])$ of Φ^{Ξ} and Φ is included in $\bigcup_{i=1}^{N}[\mathbf{z}_i]$, we can write

$$\begin{aligned}||\Phi^{\Xi} - \Phi|| &= \int\limits_{\bigcup_{i=1}^{N}[\mathbf{z}_i]}||\Phi^{\Xi}(\mathbf{z}) - \Phi(\mathbf{z})||d\mathbf{z}\\ &= \textstyle\sum_{i=1}^{N}\int_{[\mathbf{z}_i]}||\Phi^{\Xi}(\mathbf{z}) - \Phi(\mathbf{z})||d\mathbf{z}\\ &\leq \epsilon\textstyle\sum_{i=1}^{N}||[\mathbf{z}_i]|| \leq \epsilon|[f]([\mathbf{x}_k],[\nu_{k+1}])|.\end{aligned}$$

This proves Lemma 5.1.

B Proof Lemma 5.2

For simplicity, let us prove the lemma using two points.

- First, assume that the set Ξ is constituted by only one box $[\mathbf{z}]$. Φ^{Ξ} is defined as:

  $\forall\, \mathbf{z} \notin [\mathbf{z}], \quad \Phi^{\Xi}(\mathbf{z}) = 0$ and
  $\forall\, \mathbf{z} \in [\mathbf{z}],$
  $\Phi^{\Xi}(\mathbf{z}) = \frac{1}{|[\mathbf{z}]|}\mu(\{(\mathbf{x},\nu) \in [\mathbf{x}] \times [\nu] \mid f(\mathbf{x},\nu) \in [\mathbf{z}]\}$
  $= \frac{1}{|[\mathbf{z}]|}\mu(f^{-1}([\mathbf{z}]) \cap [\mathbf{x}] \times [\nu]).$

  For Λ a set of mutually disjoints boxes $([\mathbf{x}_i] \times [\nu_i])_{i=1\ldots M}$ on $\mathbb{R}^{n_x+n_\nu}$,$\Phi^{\Xi,\Lambda}$ is defined as:

  $\forall\, \mathbf{z} \notin [\mathbf{z}] \quad , \Phi^{\Xi,\Lambda}(\mathbf{z}) = 0$ and
  $\forall\, \mathbf{z} \in [\mathbf{z}]$
  $\Phi^{\Xi,\Lambda}(\mathbf{z}) = \frac{1}{|[\mathbf{z}]|} \sum_{j \,\in\, I^{[\mathbf{z}]}} |[\mathbf{x}_j] \times [\nu_j]|$.
  with $I^{[\mathbf{z}]} = \{j \in \{1, \cdots, M\} \mid f([\mathbf{x}_j], [\nu_j]) \subseteq [\mathbf{z}]\}$
  $= \{j \in \{1, \cdots, M\} \mid [\mathbf{x}_j] \times [\nu_j] \subseteq f^{-1}([\mathbf{z}])\}$

  Lemma 5.2 can be proven by using previous results introduced in [14]. A well known algorithm *Set Inversion Via Interval Analysis* (SIVIA) exists, in interval analysis, for inverting sets through large classes of functions, among which, the continuous functions. From a given prior box, and an image box, SIVIA provides two subpavings (union of disjoints intervals), inside the prior box, both converging to the true solution set. An inner subpaving included in the solution set and and outer subpaving enclosing the solution set can be obtained by SIVIA. In [14] the convergence of SIVIA has been derived. The property we are interested in and that can be found in [14] (see in particular section 4 and 5) is that when the size of the boxes inside the subpaving tends toward 0, the inner subpaving tend toward the solution set. For our problem, the set $([\mathbf{x}_i] \times [\nu_i])_{I^{[\mathbf{z}]}}$ constitutes the inner subpaving, of an initial prior $[\mathbf{x}]\times[\nu]$, approximating the solution set $f^{-1}([\mathbf{z}]) \cap [\mathbf{x}] \times [\nu]$. Then we can conclude that when the size of $[\mathbf{z}]$ tends to 0, $\sum_{j \,\in\, I^{[\mathbf{z}]}} |[\mathbf{x}_j] \times [\nu_j]|$ tends to $\mu(f^{-1}([\mathbf{z}]) \cap [\mathbf{x}] \times [\nu]$ and consequently $\Phi^{\Xi,\Lambda}$ tends to Φ^{Ξ}.

- Now if we assume that Ξ is a set of mutually disjoints $([\mathbf{z}_i])_{i=1\ldots N}$ on $\mathbb{R}^{n_x}$ on $\mathbb{R}^{n_x+n_\nu}$, the convergence is proven (from above) for each $[\mathbf{z}_i]$ taken separately. Consequently, since M is a finite number, this proves Lemma 5.2.

Acknowledgements. We acknowledge the support of the EPSRC project EP/E027253/1 and from the [European Community's] Seventh Framework Programme [FP7/2007-2013] under grant agreement No 238710 (Monte Carlo based Innovative Management and Processing for an Unrivalled Leap in Sensor Exploitation).

References

[1] F. Abdallah, A. Gning, and P. Bonnifait. Box particle filtering for nonlinear state estimation using interval analysis. *Automatica*, 44(3):807–815, 2008.

[2] G. Alefeld and J. Herzberger. *Introduction to Interval Computations.* Academic Press, 1983.

[3] D.L. Alspach and H.W. Sorenson. Nonlinear Bayesian estimation using Gaussian sum approximation. *IEEE Trans. Aut. Contr.*, 17(4):439–448, 1972.

[4] M. Arulampalam, S. Maskell, N. Gordon, and T. Clapp. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Trans. on Signal Proc.*, 50(2):174–188, 2002.

[5] E. W. Cheney. *Introduction to Approximation Theory.* New York: Chelsea, second edition, 1982.

[6] A. Doucet, N. Freitas, and Eds. N. Gordon. *Sequential Monte Carlo Methods in Practice.* New York: Springer-Verlag, 2001.

[7] R. E. Edwards. *What is the Riemann integral?* Dept. of Pure Mathematics, Dept. of Mathematics, Australian National University, Canberra, 1974.

[8] D. Fox. Adapting the sample size in particle filters through KLD-sampling. *Intl. J. of Robotics Research*, 22(12):985–1003, Dec 2003.

[9] A. Gning and Ph. Bonnifait. Constraints propagation techniques on intervals for a guaranteed localization using redundant data. *Automatica*, 42(7):1167–1175, 2006.

[10] A. Jaswinski. *Stochastic Processes and Filtering Theory.* New York: Academic Press, 1970.

[11] L. Jaulin. Nonlinear bounded-error state estimation of continuous-time systems. *Automatica*, 38(6):1079 – 1082, 2002.

[12] L. Jaulin, M. Kieffer, O. Didrit, and E. Walter. *Applied Interval Analysis.* Springer-Verlag, 2001.

[13] S. Julier, J. Uhlmann, and H. Durrant-White. A new approach for filtering nonlinear systems. In *Proc. of the Amer. Control Conf.*, Washington, DC, 1995.

[14] Jaulin L. and Walter E. Set inversion via interval analysis for nonlinear bounded-error estimation. *Automatica*, 29(4):1053–1064, 1993.

[15] R. E. Moore. *Interval Analysis.* Prentice-Hall, Englewood Cliffs, NJ, 1966.