Introduction to the Box Particle Filtering

Amadou Gning*, Branko Ristic§, Lyudmila Mihaylova*, Abdallah Fahed†

I. Scope

Resulting from the synergy between the sequential Monte Carlo (SMC) method [1] and interval analysis [2], the box particle filtering is a recently emerged approach aimed at solving a general class of nonlinear filtering problems. This approach is particularly appealing in practical situations involving imprecise stochastic measurements, thus resulting in very broad posterior densities. It relies on the concept of a *box particle*, which occupies a small and controllable rectangular region having a non-zero volume in the state space. Key advantages of the box particle filter (Box-PF) against the standard particle filter (PF) are in its reduced computational complexity and its suitability for distributed filtering. Indeed, in some applications where the sequential importance resampling (SIR) PF may require thousands of particles to achieve an accurate and reliable performance, the Box-PF can reach the same level of accuracy with just a few dozens of box particles. This lecture note presents a quick introduction to this new methodology and emphasizes the need for interval tools and its Bayesian interpretation. Some familiarity with particle filters is assumed.

II. Relevance

State estimation of stochastic dynamic systems plays a crucial role in many engineering systems, from navigation, autonomous vehicles and guidance to finance and bio-informatics [3]. The nonlinear nature of dynamic and observation models results in the non-Gaussian character of the posterior density which is difficult to represent accurately. In recent years, sequential Bayesian estimation has become the dominant framework for recursive state estimation. This trend is mainly due to the invention of a *particle filter* [1], which provides a numerical (simulation based) solution for a large class of nonlinear filtering problems. The traditional Bayesian estimation deals with stochastic but *precise* measurements and measurement models. Interval measurements are convenient for modeling bounded errors with unknown distributions and unknown measurement biases [4], [5]. In circumstances where the measurements are intervals, the

*School of Computing & Communications, Lancaster University, Lancaster, United Kingdom, email: e.gning@lancaster.ac.uk

§Defence Science and Technology Organisation, ISR Division, email: branko.ristic@dsto.defence.gov.au

*School of Computing & Communications, Lancaster University, Lancaster, United Kingdom, email: mila.mihaylova@lancaster.ac.uk, Corresponding author

†HEUDIASYC, UMR CNRS 6599, Université de Technologie de Compiègne, France, email: fahed.abdallah@hds.utc.fr

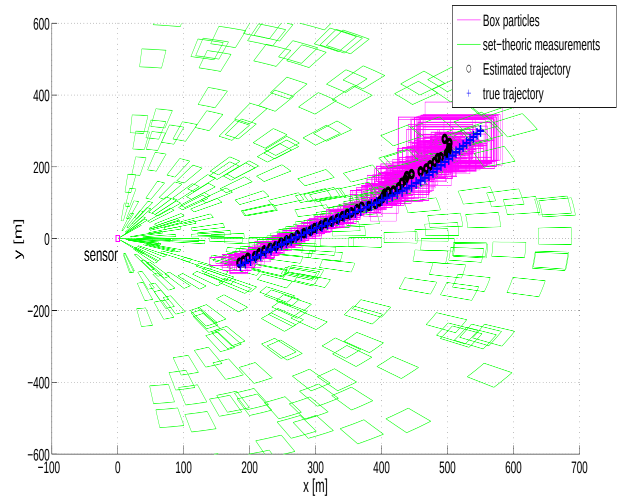

Fig. 1. Snapshot of an example of trajectory of the Box-PF. Box particles over the time are shown along with the estimated trajectory and the true one (details of the experiment and the Box-PF can be seen in [8])
.

optimal Bayes filter for state estimation can be formulated using the random set theory [6]. The SMC methods then require a massive number of particles to approximate the posterior state probability density function (pdf). In order to reduce the number of particles, the concept of a *box-particle* was introduced recently in [7]. The key idea is to replace a particle by a multi-dimensional interval or box of non-zero volume in the state space. Whilst the potential to reduce the required number of particles to approximate the posterior is a strong motivation, we will see that the use of box particles also introduces new challenges. An illustration of the Box-PF in action is given in Figure 1.

III. Bayesian Filtering

Consider the following system:

$$\begin{cases} \mathbf{x}_{k+1} = & \mathbf{f}(\mathbf{x}_k) + \mathbf{v}_k, \\ \mathbf{z}_{k+1} = & \mathbf{g}(\mathbf{x}_{k+1}) + \mathbf{w}_k, \end{cases} \tag{1}$$

where $\mathbf{f} : \mathbb{R}^{n_x} \times \mathbb{R}^{n_v} \longrightarrow \mathbb{R}^{n_x}$ is in general a nonlinear transition function defining the state vector $\mathbf{x}_{k+1}$ at time $k+1$ from the previous state $\mathbf{x}_k$ and from an independent identically distributed (iid) process noise sequence $\mathbf{v}_k$; n_x and n_v denote, respectively, the dimensions of the state and the process noise; $\mathbf{g} : \mathbb{R}^{n_x} \times \mathbb{R}^{n_w} \longrightarrow \mathbb{R}^{n_z}$ defines the relationship between the state and the measurement $\mathbf{z}_{k+1}$, with $\mathbf{w}_k$ being an iid measurement noise sequence; n_z, n_w are dimensions of the measurement and measurement noise, respectively. The states and the measurements up to time k are represented, respectively, by $\mathbf{X}_k = \{\mathbf{x}_\ell, \ell = 1, \ldots, k\}$ and $\mathbf{Z}_k = \{\mathbf{z}_\ell, \ell = 1, \ldots, k\}$. Within the Bayesian framework, the posterior state pdf $p(\mathbf{X}_{k+1}|\mathbf{Z}_{k+1})$ provides a complete description of the state up to time instant $k+1$, given the measurements $\mathbf{Z}_{k+1}$. In many applications, the marginal of the posterior pdf $p(\mathbf{x}_{k+1}|\mathbf{Z}_{k+1})$, also provides sufficient information and is given by:

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

$$p(\mathbf{x}_{k+1}|\mathbf{Z}_k) = \int_{\mathbb{R}^{nx}} p(\mathbf{x}_{k+1}|\mathbf{x}_k) p(\mathbf{x}_k|\mathbf{Z}_k) d\mathbf{x}_k, \quad (3)$$

where $p(\mathbf{x}_{k+1}|\mathbf{x}_k)$ is the transitional density, $p(\mathbf{z}_{k+1}|\mathbf{x}_{k+1})$ is the likelihood function and

$$\alpha_{k+1} = \int_{\mathbb{R}^{nx}} p(\mathbf{z}_{k+1}|\mathbf{x}_{k+1}) p(\mathbf{x}_{k+1}|\mathbf{Z}_k) d\mathbf{x}_{k+1}$$

is a normalisation factor. The recursion is initialised with a prior pdf $p(\mathbf{x}_0)$, e.g. with a uniform pdf over some region of the state space. Equation (3) corresponds to the time update or the prediction step while (2) represents the measurement update step.

IV. Box Particle Filtering

Figure 2 presents the similarities and difference between the PF and the Box-PF. Figure 2.a) illustrates the 4 main steps of the PF while Figure 2.b) illustrates the 5 main steps of the Box-PF. The likelihood step 2 of the PF is replaced by steps 2 and 3 in the Box-PF. Here, the additional step that appears in the Box-PF removes the box particles values that are not consistent with the measurement. Another visible difference is the use of *inclusion functions*, which are necessary in the Box-PF because, when propagated via a continuous nonlinear function, the image of the box particle is not necessarily a box.

The next sections give details about the 5 Box-PF steps in comparison with the PF, emphasizing the interval analysis tools needed in the Box-PF algorithm and also give a Bayesian interpretation of each step.

A. Box-PF Time Update Step

In the time update step of the PF, the posterior pdf at time k, represented by a set of N weighted particles, denoted by $\{(w_k^i, \mathbf{x}_k^i)\}_{i=1}^N$, is propagated to the next time $k+1$. In the importance sampling scheme, the particles are sampled according to a proposal pdf. If the proposal pdf is chosen to be the transition prior, the updated particles at time $k+1$ are sampled according to: $\{\mathbf{x}_{k+1|k}^i = \mathbf{f}(\mathbf{x}_k^i) + \mathbf{v}_k^i\}_{i=1}^N$, where $\mathbf{v}_k^i$ is a noise realisation corresponding to the particle $\mathbf{x}_k^i$. The new weights are equal to the previous ones, *i.e.*, $\left\{w_{k+1|k}^i = w_k^i\right\}_{i=1}^N$. For simplicity in the rest of the paper w_k^i is used instead if $w_{k+1|k}^i$.

For the Box-PF, the posterior at time k is represented by a set of N weighted box particles denoted by $\{(w_k^i, [\mathbf{x}_k^i])\}_{i=1}^N$. In the time update step, each box particle $[\mathbf{x}_k^i]$ is propagated through the transition prior using the tools of interval analysis: the interval arithmetic and the concept of inclusion function.

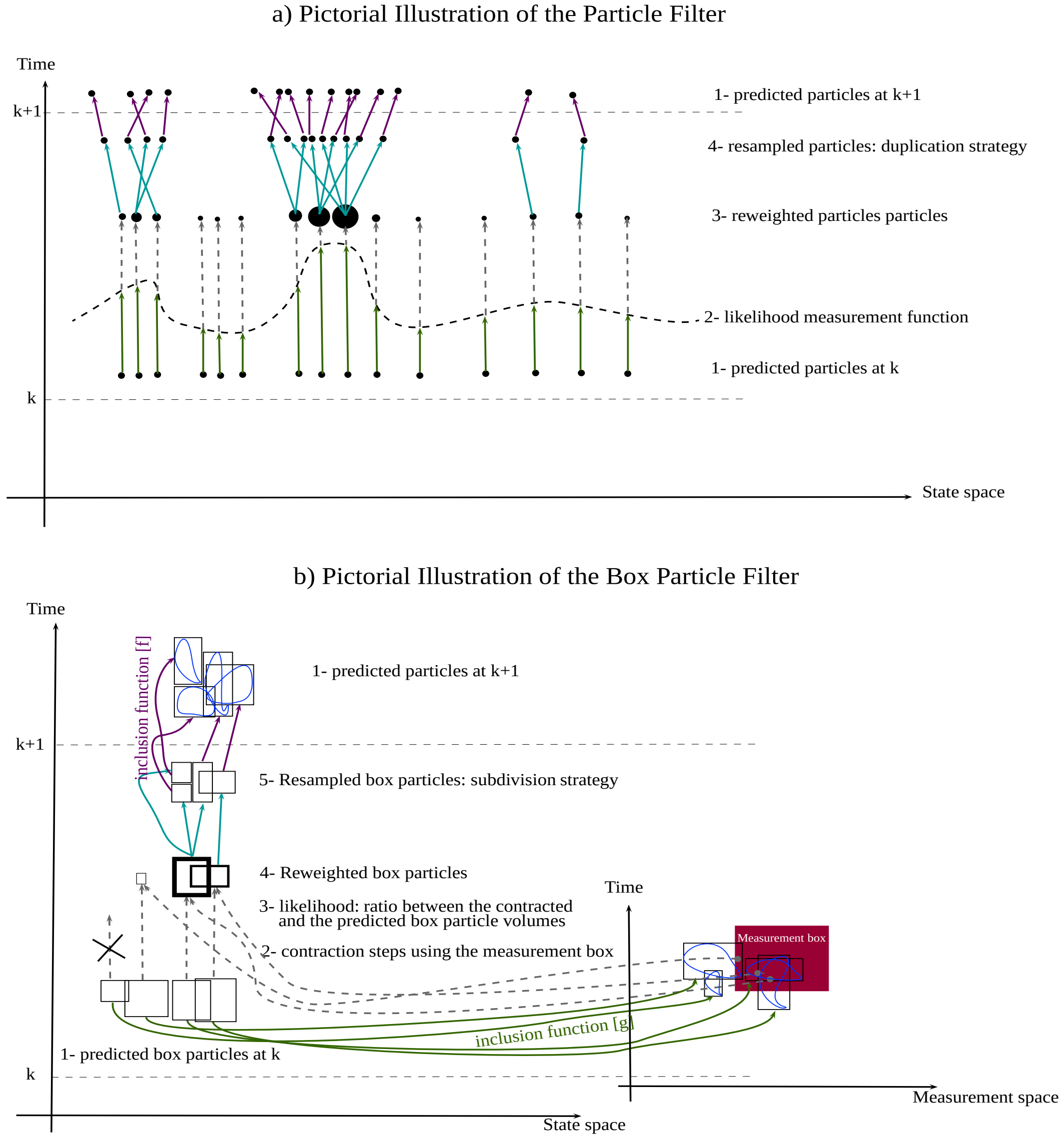

Fig. 2. This figure shows pictorial representations of the PF in a) and the Box-PF in b).

*1) **Interval Arithmetic and Inclusion Functions**:* In Table I, notations, basic definitions and interval tools are summarized. More details can be find in [2]. Additionally to the definition of the division operator given in Table I, in the case when 0 belongs to $[y]$, the division operator is straightforwardly defined by extending $\mathbb{R}$ to $\mathbb{R} \cup \{-\infty, \infty\}$ and by defining intervals in the form $[-\infty, \overline{x}]$, $[\underline{x}, \infty]$ and $[\infty, \infty]$.

The concept of inclusion function has a key role in the derivation of the Box-PF. For the Box-PF, the time update step is similar to the corresponding one in the generic Sampling Importance Resampling (SIR) particle filter, with the difference that the transition function has to be applied to a box. However, if a continuous function is applied to a box, in general there is no guarantee that the image of the function will also be a box (see Figure 3 for an illustration). In order to cope with this issue, the concept of inclusion function is applied; it guarantees that the prediction of box particles always results in (new) box particles.

TABLE I
INTERVAL ANALYSIS TOOLS

	Notations	Definitions
Interval	$[x]=[\underline{x},\overline{x}]=\{x\in\mathbb{R}\mid \underline{x}\leq x\leq,\overline{x}\}$	closed and connected subset of $\mathbb{R}$
The set of intervals	$\mathbb{IR}=\{[x]\subset\mathbb{R}\}$	
Interval Length $\|.\|$	$\|[x]\|=\overline{x}-\underline{x}$	
Interval Hull [.]	$[S]$, for any set S in $\mathbb{R}$	the smallest interval enclosing S
Set Theoretic Operator $\triangle$	$[x]\triangle[y]=[\{[x]\triangle[y]\}]$	interval hull of the resulting set
$\cap$	$[x]\cap[y]$	
Union $\sqcup$	$[x]\sqcup[y]=[[x]\cup[y]]$	*interval hull* of $[x]\cup[y]$
Binary Operator $\diamond$	$[x]\diamond[y]=[\{x\diamond y\mid x\in[x],y\in[y]\}]$	
$+$	$[x]+[y]=[\underline{x}+\underline{y},\overline{x}+\overline{y}]$	
$-$	$[x]-[y]=[\underline{x}-\overline{y},\overline{x}-\underline{y}]$	
$\times$	$[x]\times[y]=[\min(\underline{x}\underline{y},\underline{x}\overline{y},\overline{x}\underline{y},\overline{x}\overline{y}),\max(\underline{x}\underline{y},\underline{x}\overline{y},\overline{x}\underline{y},\overline{x}\overline{y})]$	
$/$	$[x]/[y]=[x]\times[1/\overline{y},1/\underline{y}]$	If the value $0\notin[y]$
Box	$[\mathbf{x}]=[x_1]\times\ldots\times[x_n]$	set of vectors of $\mathbb{R}^n$
Box volume $\|.\|$	$\|[\mathbf{x}]\|$	
Box Set Theoretic Operator $\triangle$	$[\mathbf{x}]\triangle[\mathbf{y}]=([x_1]\triangle[y_1])\times\ldots\times([x_n]\triangle[y_n])$	
Box Binary Operator $\diamond$	$[\mathbf{x}]\diamond[\mathbf{y}]=([x_1]\diamond[y_1])\times\ldots\times([x_n]\diamond[y_n])$	

Definition 1 *Let* $\mathbf{f}$ *be a function from* $\mathbb{R}^n$ *to* $\mathbb{R}^m$*. An "interval function"* $[\mathbf{f}]$ *from* $\mathbb{IR}^n$ *to* $\mathbb{IR}^m$ *is said to be an inclusion function for* $\mathbf{f}$ *if:* $\mathbf{f}([\mathbf{x}])\subseteq[\mathbf{f}]([\mathbf{x}]),\quad\forall[\mathbf{x}]\in\mathbb{IR}^n.$

Inclusion functions may be very pessimistic, as shown on Figure 3. The inclusion function $[\mathbf{f}]$ is minimal if, for any $[\mathbf{x}]$, $[\mathbf{f}]([\mathbf{x}])$ is the interval hull of $\mathbf{f}([\mathbf{x}])$. The minimal inclusion function for $\mathbf{f}$ is unique and is denoted by $[\mathbf{f}]^*$. The minimal inclusion function $[\mathbf{f}]^*$ satisfies: $[\mathbf{f}]([\mathbf{x}])=[\{\mathbf{f}(\mathbf{x})|\mathbf{x}\in[\mathbf{x}]\}]$.

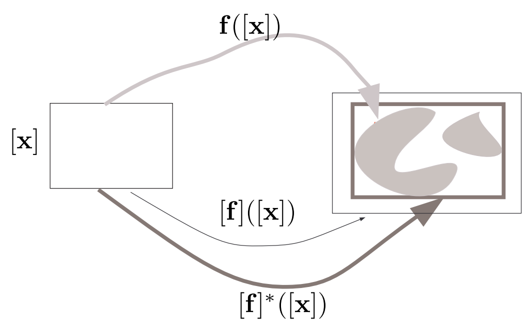

Fig. 3. This Figure shows inclusion functions obtained after applying the function $\mathbf{f}$ to a box $[\mathbf{x}]$. The resulting image is not necessarily a box. A pessimistic inclusion function $[\mathbf{f}]$ and the minimal inclusion function $[\mathbf{f}]^*$ are also presented.

Finding inclusion functions that can be evaluated with a convenient computational time and such that $[\mathbf{f}]([\mathbf{x}])$ is as close as possible to $[\mathbf{f}]^*([\mathbf{x}])$, for most $[\mathbf{x}]$, is one of the main purposes of interval analysis [2]. Different algorithms have been proposed to reduce the size of boxes enclosing $\mathbf{f}([\mathbf{x}])$, e.g., inclusion functions for elementary functions such as $\exp$, $\ln$, $\tan$ $\cos$ and $\sin$, are well studied and known. One interesting property is that, if $\mathbf{f}$ is continuous and monotonic, then $[\mathbf{f}]([\mathbf{x}])$ is simply equal to $\mathbf{f}([\mathbf{x}])$ (for instance, $[\exp]([x])=[\exp(\underline{x}),\exp(\overline{x})]$). For non-monotonic continuous functions, however, the computation of $[\mathbf{f}]$ usually is not straightforward.

*2) **Box Particles Prediction Using Inclusion Functions**:* Knowing the set of box particles $\{[\mathbf{x}_k^i]\}_{i=1}^N$ at time step k and assuming that the system noise is known to be enclosed in $[\mathbf{v}_{k+1}]$, the boxes are propagated using an inclusion function $[\mathbf{f}]$ of the transition function $\mathbf{f}$ in (1) i.e. $\{[\mathbf{x}_{k+1|k}^i] = [\mathbf{f}]([\mathbf{x}_k^i]) + [\mathbf{v}_k]\}_{i=1}^N$. This step brings an interesting property of the Box-PF: instead of propagating each particle using one realisation of the noise $\mathbf{v}_k$, the uncertainty due to noise is also propagated for each box particle.

*3) **Bayesian Interpretation for the Box-PF Time Update**:* In [9] it is shown that the Box-PF can be seen as a Bayesian filter by interpreting each box particle as a uniform pdf (with the box particle as support). The set of box particles is then interpreted as a mixture of uniform pdfs.

TABLE II
MIXTURE OF UNIFORM PDFS

Notations	Definitions
$U_{[\mathbf{x}]}$	the uniform pdf with the box $[\mathbf{x}]$ as support
$p(\mathbf{x}) = \sum_{i=1}^N w^i U_{[\mathbf{x}^i]}(\mathbf{x})$,	mixture of N components with $([\mathbf{x}^i])_{i=1}^N$ the box supports and $(w^i)_{i=1}^N$ a set of normalised weights
$p(\mathbf{x}_k \vert \mathbf{Z}_k) = \sum_{i=1}^N w_k^i U_{[\mathbf{x}_k^i]}(\mathbf{x}_k)$	previous time posterior pdf $p(\mathbf{x}_k \vert \mathbf{Z}_k)$ approximated by a mixture of uniform pdfs

Assume the posterior pdf at time k is approximated by a mixture of uniform pdfs. With the notations in Table II, the time update equation (3) can be written:

$$p(\mathbf{x}_{k+1}|\mathbf{Z}_k) \;=\; \int_{\mathbb{R}^{n_x}} p(\mathbf{x}_{k+1}|\mathbf{x}_k) \sum_{i=1}^N w_k^i U_{[\mathbf{x}_k^i]}(\mathbf{x}_k) d\mathbf{x}_k = \sum_{i=1}^N w_k^i \int_{[\mathbf{x}_k^i]} p(\mathbf{x}_{k+1}|\mathbf{x}_k) U_{[\mathbf{x}_k^i]}(\mathbf{x}_k) d\mathbf{x}_k. \tag{4}$$

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

$$p(\mathbf{x}_{k+1}|\mathbf{x}_k) \cdot U_{[\mathbf{x_k}^i]}(\mathbf{x}_k) = 0 \quad \forall\, \mathbf{x}_{k+1} \notin [\mathbf{f}]([\mathbf{x}_k^i]) + [\mathbf{v}_k]. \tag{5}$$

Equation (5) shows that for any transition function $\mathbf{f}$, using interval analysis techniques, the support for the pdf terms $\int_{[\mathbf{x}_k^i]} p(\mathbf{x}_{k+1}|\mathbf{x}_k) U_{[\mathbf{x}_k^i]}(\mathbf{x}_k) d\mathbf{x}_k$ can be approximated by $[\mathbf{f}]([\mathbf{x}_k^i]) + [\mathbf{v}_k]$. In addition, it can be seen that, in the Box-PF algorithm, each pdf term $\int_{[\mathbf{x}_k^i]} p(\mathbf{x}_{k+1}|\mathbf{x}_k) U_{[\mathbf{x}_k^i]}(\mathbf{x}_k) d\mathbf{x}_k$ in (4) is modeled using one uniform pdf component having as support the interval $[\mathbf{f}]([\mathbf{x}_k^i], [\mathbf{v}_k])$, *i.e.*,

$$\int_{[\mathbf{x}_k^i]} p(\mathbf{x}_{k+1}|\mathbf{x}_k) U_{[\mathbf{x}_k^i]}(\mathbf{x}_k) d\mathbf{x}_k \approx U_{[\mathbf{f}]([\mathbf{x}_k^i]) + [\mathbf{v}_k]}. \tag{6}$$

Combining (4) and (6) gives

$$p(\mathbf{x}_{k+1}|\mathbf{Z}_k) \approx \sum_{i=1}^N w_k^i U_{[\mathbf{f}]([\mathbf{x}_k^i]) + [\mathbf{v}_k]}. \tag{7}$$

The Box-PF strategy of approximating each pdf $\int_{[\mathbf{x}_k^i]} p(\mathbf{x}_{k+1}|\mathbf{x}_k) U_{[\mathbf{x}_k^i]}(\mathbf{x}_k) d\mathbf{x}_k$ using one uniform pdf component may not be accurate enough (however, as for the PF, it is sufficient to approximate the first moments of the pdf as shown experimentally in [7]). Alternately, a mixture of uniform pdfs can be used to better approximate this pdf as shown in [9] .

B. Box-PF Measurement Update Step

Similarly to the PF, the weights of the predicted box particles have to be updated using the new measurement. For this purpose, likelihood factors need to be calculated using the innovation quantities. In the case of the standard PF, the innovation for particle $i = 1, \ldots, N$ is $\mathbf{r}^i_{k+1} = \mathbf{z}_{k+1} - \mathbf{z}^i_{k+1}$, where $\mathbf{z}^i_{k+1} = \mathbf{g}(\mathbf{x}^i_{k+1|k})$ is the ith predicted measurement. Next, using the probabilistic model $p_{\mathbf{w}}$ for the measurement noise $\mathbf{w}_k$, the likelihood of each particle is calculated as: $p(\mathbf{z}_{k+1}|\mathbf{x}^i_{k+1|k}) = p_{\mathbf{w}}(\mathbf{z}_{k+1} - \mathbf{z}^i_{k+1}) = p_{\mathbf{w}}(\mathbf{r}^i_{k+1})$.

A reasonable assumption for the box-PF is that the measurement likelihood function has a bounded support that we call here *measurement likelihood box*. In the bounded error context, the likelihood for each box particle is calculated using the idea that a box particle with a corresponding predicted measurement without an intersection with the measurement likelihood box has a likelihood factor equal to zero. In contrast, a box particle for which the predicted measurement is included in the measurement likelihood box has a likelihood close to one. In order to calculate such a likelihood, the Box-PF incorporates a new step called the contraction 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. Therefore, in order to preserve an appropriate size of each box, a contraction step is performed which allows to eliminate the inconsistent part of the box particles with respect to the measured box. The contraction step for the box is an analog to the variance matrix measurement update step that appears in the Kalman filtering. This contraction step utilises interval analysis methods described in the next section.

1) ***Interval Contraction Methods****:* A major challenge for interval methods, is to solve systems of equations involving initial conditions falling into boxes. The next Section yields a formulation of such classes of problems. Table III introduces notations and definitions needed for the introduction of interval

TABLE III
CONSTRAINT SATISFACTION PROBLEMS

	Notations	Definitions
Real valued constraint f on $\mathbb{R}^n$	$f(\mathbf{x}) = f(x_1, \ldots, x_n) = 0$, with $\mathbf{x} = (x_1, \ldots, x_n)$	n variables x_i in $\mathbb{R}$, $i \in \{1, \ldots, n\}$ linked by f
Set of constraints	$\mathbf{f} = (f_1, f_2, \cdots, f_m)^T$	
Constraint Satisfaction Problem (CSP) $\mathcal{H}$	$\mathcal{H} : (\mathbf{f}(\mathbf{x}) = \mathbf{0}, \mathbf{x} \in [\mathbf{x}])$	problem of finding the smallest $[\mathbf{x}']$ enclosing the set of all $\mathbf{x}$ in $[\mathbf{x}]$ satisfying $\mathbf{f}$
The solution set of $\mathcal{H}$	$\mathbf{S} = \{\mathbf{x} \in [\mathbf{x}] \mid \mathbf{f}(\mathbf{x}) = \mathbf{0}\}$	

contraction concepts. In Table III, the set $\mathbf{S}$ is not necessary a box. Within the interval framework, solving a *Constraint Satisfaction Problem* (CSP) $\mathcal{H}$ implies finding the smallest box $[\mathbf{x}]' \subset [\mathbf{x}]$ constituting an outer approximation of $\mathbf{S}$, such that $\mathbf{S} \subseteq [\mathbf{x}]' \subseteq [\mathbf{x}]$.

Definition 2 *Contracting $\mathcal{H}$ means replacing $[\mathbf{x}]$ by a smaller domain $[\mathbf{x}]'$ such that $\mathbf{S} \subseteq [\mathbf{x}]' \subseteq [\mathbf{x}]$. A contractor $\mathcal{C}$ for $\mathcal{H}$ is any operator that can be used to contract $\mathcal{H}$.*

Several methods for building contractors are described in [2, Chapter 4], including Gauss elimination, Gauss-Seidel algorithm and linear programming. Each of these methods can be suitable for different types of CSP. An attractive contractor method is the so called *Constraints Propagation* (CP) technique [2]. The main advantage of the CP method is its efficiency in the presence of high redundancy of data and constraints. The CP method is also simple and, most importantly, independent of nonlinearities. The CP method proceeds by contracting $\mathcal{H}$ with respect to each variable, appearing in each constraint, until convergence to a minimal domain. Hereafter, a simple illustration is given though Example 1. A detailed description of the CP algorithm can be found in [2].

Example 1. Consider a three dimensional CSP with a single constraint $z = x\exp(y)$ and an initial domain $[z] = [0, 3]$, $[x] = [1, 7]$ and $[y] = [0, 1]$. The CP algorithm alternates between two phases commonly called forward propagation (FP) and backward propagation (BP).

- FP_1: $[z] \leftarrow [z] \cap ([x] \times [\exp]([y]) = [0, 3] \cap [1, 7] \times [1, e] = [1, 3]$.
- BP_2: $[x] \leftarrow [x] \cap ([z]/[\exp]([y])) = [1, 7] \cap [1, 3]/[1, e] = [1, 3]$
- BP_3: $[y] \leftarrow [y] \cap [\ln]([z]/[x]) = [0, 1] \cap [\ln][1, 3]/[1, 3]) = [0, 1]$.

FP_1 above contracted the domain of z, while BP_2 and BP_3, using an inversion of the constraint, contracted the domains of x and y. Thus after one forward-backward propagation cycle, the domains of the variables have been reduced to $[z] = [1, 3]$, $[x] = [1, 3]$ and $[y] = [0, 1]$.

*2) **Bayesian interpretation of the contraction step**:* Assume that, at time instant $k+1$, an approximation of the time update pdf $p(\mathbf{x}_{k+1}|\mathbf{Z}_k)$ by a mixture of N uniform pdfs with interval supports $[\mathbf{x}^{(i)}_{k+1|k}]$ and weights w^i_k is available and that the measurement update step is to be performed. Next, for the Box-PF, a probabilistic model $p_\mathbf{w}$ for the measurement noise $\mathbf{w}_k$ is also available. It is assumed in general that $p_\mathbf{w}$ can be expressed using a mixture of uniform pdfs. For simplicity and without loss of generality, $p_\mathbf{w}$ is considered here to be one uniform pdf over a box denoted $[\mathbf{z}_{k+1}]$, *i.e.*: $p(\mathbf{z}_{k+1}|\mathbf{x}_{k+1}) = U_{[\mathbf{z}_{k+1}]}(\mathbf{g}(\mathbf{x}_{k+1}))$. According to (2), the measurement update is:

$$\begin{aligned} p(\mathbf{x}_{k+1}|\mathbf{Z}_{k+1}) &= \frac{1}{\alpha_{k+1}} p(\mathbf{z}_{k+1}|\mathbf{x}_{k+1}) p(\mathbf{x}_{k+1}|\mathbf{Z}_k) = \frac{1}{\alpha_{k+1}} U_{[\mathbf{z}_{k+1}]}(\mathbf{g}(\mathbf{x}_{k+1})) \sum_{i=1}^{N} w^i_k \, U_{[\mathbf{x}^i_{k+1|k}]}(\mathbf{x}_{k+1}) \\ &= \frac{1}{\alpha_{k+1}} \sum_{i=1}^{N} w^i_k \, U_{[\mathbf{z}_{k+1}]}(\mathbf{g}(\mathbf{x}_{k+1})) \, U_{[\mathbf{x}^i_{k+1|k}]}(\mathbf{x}_{k+1}). \end{aligned} \tag{8}$$

Each of the terms $U_{[\mathbf{z}_{k+1}]}(\mathbf{g}(\mathbf{x}_{k+1}))U_{[\mathbf{x}^i_{k+1|k}]}(\mathbf{x}_{k+1})$ is also a constant function with a support being the following set $S_i \subset \mathbb{R}^{n_x}$

$$S_i = \left\{\mathbf{x}_{k+1} \in [\mathbf{x}^i_{k+1|k}] \mid \mathbf{g}(\mathbf{x}_{k+1}) \in [\mathbf{z}_{k+1}]\right\}. \tag{9}$$

Equation (9) defines a CSP and from its expression we can deduce that predicted supports $[\mathbf{x}^i_{k+1|k}]$, from the time update pdf $p(\mathbf{x}_{k+1}|\mathbf{Z}_k)$ approximation, have to be contracted with respect to the new measurement $[\mathbf{z}_{k+1}]$. These contraction steps result in the new box-particles denoted $[\tilde{\mathbf{x}}^i_{k+1}]$, for the posterior pdf $p(\mathbf{x}_{k+1}|\mathbf{Z}_{k+1})$ at time instant $k+1$.

*3) **Bayesian derivation of the likelihood**:* Following the definition of the sets S_i in (9), we can write

$$U_{[\mathbf{z}_{k+1}]}(\mathbf{g}(\mathbf{x}_{k+1}))U_{[\mathbf{x}^i_{k+1|k}]}(\mathbf{x}_{k+1}) = U_{[\mathbf{z}_{k+1}]}(\mathbf{g}(\mathbf{x}_{k+1}))\frac{1}{|[\mathbf{x}^i_{k+1|k}]|}||S_i||U_{S_i}(\mathbf{x}_{k+1}). \tag{10}$$

By combining eqs. (8) and (10), and keeping in mind that $[\tilde{\mathbf{x}}^i_{k+1}] = [S_i]$ (i.e. by definition $[\tilde{\mathbf{x}}^i_{k+1}]$ is the smallest box containing S_i),

$$\begin{aligned}
p(\mathbf{x}_{k+1}|\mathbf{Z}_{k+1}) &= \frac{1}{\alpha_{k+1}}\sum_{i=1}^{N} w^i_k \frac{1}{|[\mathbf{z}_{k+1}]|}\frac{1}{|[\mathbf{x}^i_{k+1|k}]|}||S_i||U_{S_i}(\mathbf{x}_{k+1}) \\
&\approx \frac{1}{\alpha_{k+1}}\sum_{i=1}^{N} w^i_k \frac{1}{|[\mathbf{z}_{k+1}]|}\frac{1}{|[\mathbf{x}^i_{k+1|k}]|}|[\tilde{\mathbf{x}}^i_{k+1}]|U_{|[\tilde{\mathbf{x}}^i_{k+1}]|}(\mathbf{x}_{k+1}) \\
&\propto \sum_{i=1}^{N} w^i_k \frac{|[\tilde{\mathbf{x}}^i_{k+1}]|}{|[\mathbf{x}^i_{k+1|k}]|}U_{|[\tilde{\mathbf{x}}^i_{k+1}]|}(\mathbf{x}_{k+1})
\end{aligned} \tag{11}$$

C. Weight update

In the SIR particle filter, each particle weight is updated by a factor equal to the likelihood $p(\mathbf{z}_k|\mathbf{x}^i_{k+1|k})$, followed by normalisation of weights. In the Box-PF this step is very similar: after contracting each box particle $[\mathbf{x}^i_{k+1|k}]$ into $[\tilde{\mathbf{x}}^i_{k+1}]$, according to (11) the weights are updated by the ratio $L^i_k = \frac{|[\tilde{\mathbf{x}}^i_{k+1}]|}{|[\mathbf{x}^i_{k+1|k}]|}$. In summary, $p(\mathbf{x}_{k+1}|\mathbf{Z}_{k+1})$ is approximated by $\{(\tilde{w}^i_{k+1}, [\tilde{\mathbf{x}}^i_{k+1}])\}_{i=1}^N$, where $\tilde{w}^i_{k+1} \propto w^i_k \cdot L^i_k$.

D. Resampling Box Particles

Similarly to the SIR PF algorithm, the resampling step is added to the box-PF to prevent degeneracy of box-particles. Different resampling algorithms can be used [10]. In the standard PF algorithm, the particles characterised by high weights have a good chance to multiply during the resampling step, and then to propagate to the future time with artificial noise added to increase their diversity. The same strategy can be used for box particles, with artificial noise added to each box to increase diversity. However, alternative techniques in the resampling step of the Box-PF are also possible. For instance, in order to increase the "resolution" in the regions of the state space where the posterior pdf has high values, a box particle which has been selected n times during resampling, can be partitioned into n disjoint smaller boxes. The efficiency of this box-particle resampling strategy is empirically confirmed in [7].

V. Lessons Learned, Further Reading and Future Avenues

This lecture note summarises a new method for sequential nonlinear estimation based on a combination of particle filtering and interval analysis. The method is based on a new concept of box particles for the purpose of reducing the number of random samples required by the standard particle filter. The Box-PF algorithm is presented in the prism of the Bayesian inference using mixtures of uniform pdfs with boxed supports.

More details about the Box-PF and its implementation can be found in [7], [9], [8].

Numerous challenges remain for future work. Many theoretical results are still missing, such as convergence results, better proposal densities, theoretical justification of resampling with partition. The significant reduction of the number of particles opens perspectives for distributed nonlinear and non parametric state estimation problems.

Acknowledgements. The authors acknowledges the support of 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] A. Doucet, J. F. G. de Freitas, and N. J. Gordon, Eds., *Sequential Monte Carlo Methods in Practice*, Springer, 2001.

[2] L. Jaulin, M. Kieffer, O. Didrit, and È. Walter, *Applied Interval Analysis*, Springer, 2001.

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

[4] M. Milanese and A. Vicino, "Optimal estimation theory for dynamic systems with set membership uncertainty: An overview," *Automatica*, vol. 27, no. 6, pp. 997–1009, 1991.

[5] U. D. Hanebeck, J. Horn, and G. Schmidt, "On combing statistical and set-theoretic estimation," *Automatica*, vol. 35, pp. 1101–1109, 1999.

[6] R. Mahler, *Statistical Multisource-multitarget Information Fusion*, Artech House, Boston, 2007.

[7] F. Abdallah, A. Gning, and P. Bonnifait, "Box particle filtering for nonlinear state estimation using interval analysis," *Automatica*, vol. 44, pp. 807–815, 2008.

[8] A. Gning, Ristic B., and L. Mihaylova, "A box particle filter for stochastic set-theoretic measurements with association uncertainty," in *Proc. of the International Conf. on Information Fusion*, Chicago, USA, 2011.

[9] A. Gning, L. Mihaylova, and F. Abdallah, "Mixture of uniform probability density functions for non linear state estimation using interval analysis," in *Proc. of the International Conf. on Information Fusion*, Edinburgh, UK, 2010.

[10] E. Wan and R. van der Merwe, *The Unscented Kalman Filter, Ch. 7: Kalman Filtering and Neural Networks. Edited by S. Haykin*, pp. 221–280, Wiley Publishing, September 2001.