Amadou Gning, Branko Ristic,
Lyudmila Mihaylova, and
Fahed Abdallah

An Introduction to Box Particle Filtering

Resulting from the synergy between the sequential Monte Carlo (SMC) method [1] and interval analysis [2], box particle filtering is an approach that has recently emerged [3] and is aimed at solving a general class of nonlinear filtering problems. This approach is particularly appealing in practical situations involving imprecise stochastic measurements that result in very broad posterior densities. It relies on the concept of a box particle that occupies a small and controllable rectangular region having a nonzero volume in the state space. Key advantages of the box particle filter (box-PF) against the standard particle filter (PF) are its reduced computational complexity and its suitability for distributed filtering. Indeed, in some applications where the sampling importance resampling (SIR) PF may require thousands of particles to achieve accurate and reliable performance, the box-PF can reach the same level of accuracy with just a few dozen box particles. Recent developments [4] also show that a box-PF can be interpreted as a Bayes' filter approximation allowing the application of box-PF to challenging target tracking problems [5].

Although the box-PF has appealing properties, its widespread use seems to be hindered by the signal processing community's unfamiliarity with interval analysis techniques. The aim of this lecture note is to present the box-PF in a concise, simple, and appealing way, via Bayesian formulation and a gradual introduction to interval analysis. Readers should have some familiarity with PFs.

RELEVANCE

State estimation of stochastic dynamic systems plays a crucial role in many engineering systems, from navigation, autonomous vehicles, and guidance to finance and bioinformatics [6]. The nonlinear nature of dynamic and observation models and the presence of non-Gaussian noises makes it difficult to accurately represent the posterior density.

In recent years, sequential Bayesian estimation has become the dominant framework for recursive state estimation. This trend is mainly due to the invention of the PF [1], which provides a numerical (simulation-based) solution for a large class of nonlinear filtering problems. 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 [7], [8]. In circumstances where the measurements are intervals, the optimal Bayes' filter for state estimation can be formulated using random set theory [9]. The SMC methods then require a massive number of particles to approximate the posterior state probability density function (pdf). To reduce the number of particles, the concept of a box-PF was recently introduced in [3]. The key idea is to replace a particle by a multidimensional interval or box of nonzero volume in the state space. While 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-PFs also introduces new challenges. An illustration of the box-PF in action is given in Figure 1.
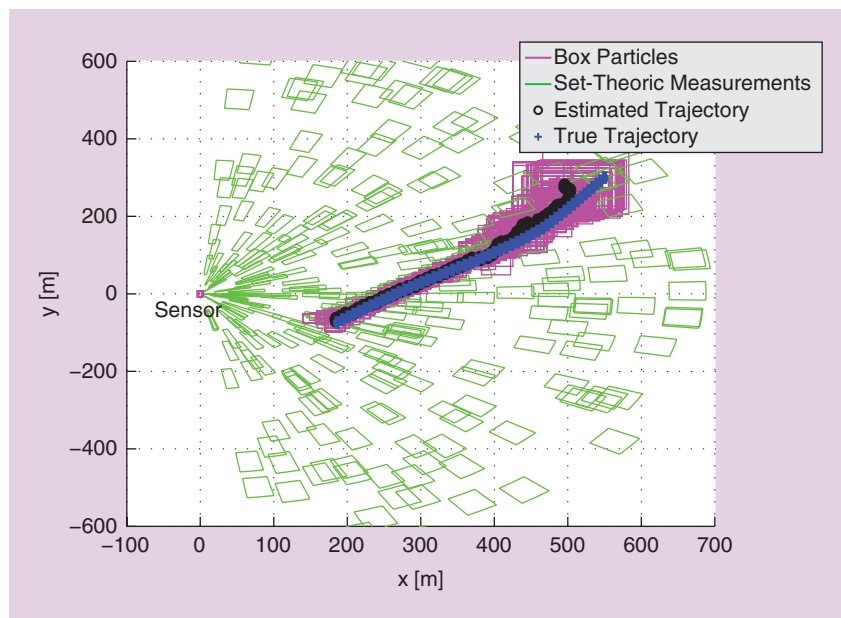
[FIG1] Snapshot of an example of the trajectory of the box-PF. The box particles over time show the estimated and true trajectories. Details of the experiment and the box-PF can be seen in [5].

PREREQUISITES

This article assumes only a familiarity with state-space representation, discrete time signal processing, probability theory- and Bayesian estimation.

PROBLEM STATEMENT

Consider the following system:

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

where $\mathbf{f}: \mathbb{R}^{n_x} \longrightarrow \mathbb{R}^{n_x}$ is in general a nonlinear transition function defining the state vector $\mathbf{x}_{k+1}$ at time $k + 1$ from the previous state $\mathbf{x}_k$ perturbed by an additive independent identically distributed (iid) process noise sequence $\mathbf{v}_k$ and n_x is the dimensions of the state; $\mathbf{g}: \mathbb{R}^{n_x} \longrightarrow \mathbb{R}^{n_z}$ defines the relationship between the state and the measurement $\mathbf{z}_{k+1}$, with $\mathbf{w}_k$ being an additive iid measurement noise sequence; n_z is the dimension of the measurement. 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

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

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

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

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

is a normalization factor. The recursion is initialized 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.

SOLUTION

In general, the recursive equations (2) and (3) of the posterior density cannot be determined analytically, e.g., in presence of a highly nonlinear model or non-Gaussian multimodal pdfs. In the next section, an approximation solution is proposed based on box particles.

BOX PARTICLE FILTERING ILLUSTRATION

Figure 2 presents the similarities and differences between the PF and the box-PF. Figure 2(a) illustrates the four main steps of the PF, while Figure 2(b) illustrates the five 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 values in the box particles that are not consistent with the measurement. Another important difference is that inclusion functions 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 several sections give details about the five 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.

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 realization corresponding to the particle $\mathbf{x}_k^i$. The new weights are equal to the previous weights, i.e., $\{w_{k+1|k}^i = w_k^i\}_{i=1}^N$. For simplicity, w_k^i is used instead of $w_{k+1|k}^i$. for the remainder of this article.

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: interval arithmetic and inclusion functions.

INTERVAL ARITHMETIC AND INCLUSION FUNCTIONS

A real interval, $[x] = [\underline{x}, \bar{x}]$, is defined as a closed and connected subset of the set $\mathbb{R}$ of real numbers with $\underline{x}$ and $\bar{x}$ denoting respectively the lower and the upper bounds of $[x]$. In Table 1, notations, basic definitions, and interval tools are summarized. More details can be found in [2]. In the case when 0 belongs to $[y]$, the division operator, given in Table 1, is straightforwardly extended by replacing $\mathbb{R}$ with $\mathbb{R} \cup \{-\infty, \infty\}$ and by defining intervals of the form $[-\infty, \bar{x}], [\underline{x}, \infty]$ and $[\infty, \infty]$.

Inclusion functions have a key role in the derivation of the box-PF. For the box-PF, the time update step is similar to the corresponding time update step of the generic SIR PF, 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 result of the function will also be a box (see Figure 3 for an illustration). To cope with this issue, inclusion functions are applied; they guarantee that the prediction of box particles always results in (new) box particles.

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 $\mathbf{f}([\mathbf{x}]) \subseteq [\mathbf{f}]([\mathbf{x}])$, for $\forall [\mathbf{x}] \in \mathbb{IR}^n$.

Inclusion functions may be very "pessimistic" in the sense that the resulting bounding box is larger than necessary, as shown in Figure 3. An 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 $[\mathbf{f}]([\mathbf{x}]) = [\{\mathbf{f}(\mathbf{x}) | \mathbf{x} \in [\mathbf{x}]\}]$.

Finding inclusion functions that can be evaluated with a convenient computational time and such that $[\mathbf{f}]([\mathbf{x}])$
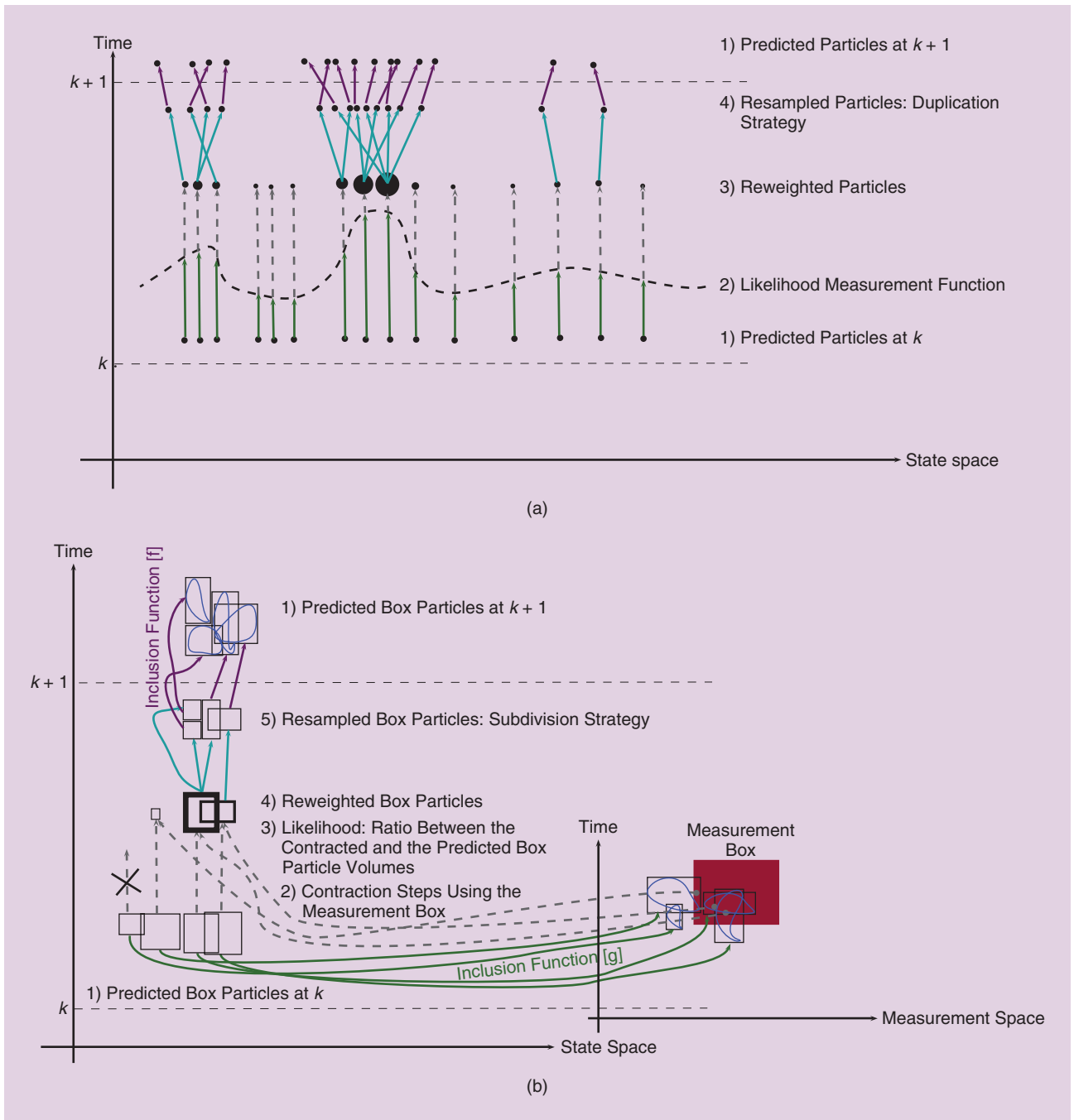
[FIG2] Pictorial representations of the (a) PF and (b) box-PF.

is as close as possible to $[f]^*([x])$, for most $[x]$, is one of the main purposes of interval analysis [2]. Different algorithms have been proposed to reduce the size of boxes enclosing $f([x])$, 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 f is continuous and monotonic, then $[f]([x])$ is simply equal to $f([x])$ (for

instance, $[\exp]([x]) = [\exp(\underline{x}), \exp(\overline{x})])$. For nonmonotonic continuous functions, however, the computation of $[f]$ is usually not straightforward.

BOX PARTICLE PREDICTION USING INCLUSION FUNCTIONS

Knowing the set of box particles $\{[x_k^i]\}_{i=1}^N$ at time step k and assuming that the system noise is known to be

enclosed in $[v_k]$, the boxes are propagated using an inclusion function $[f]$ of the transition function f in (1), i.e., $\{[x_{k+1|k}^i] = [f]([x_k^i]) + [v_k])\}_{i=1}^N$. This step brings an interesting property of the box-PF: instead of propagating each particle using one realization of the noise v_k, the uncertainty due to noise is also propagated for each box particle.

[TABLE 1] INTERVAL ANALYSIS TOOLS.

	NOTATIONS	DEFINITIONS
INTERVAL	$[x] = [\underline{x}, \bar{x}] = \{x \in \mathbb{R} \mid \underline{x} \le x \le \bar{x}\}$	CLOSED AND CONNECTED SUBSET OF $\mathbb{R}$
THE SET OF INTERVALS	$\mathbb{IR} = \{[x] \subset \mathbb{R}\}$	
INTERVAL LENGTH $\mid . \mid$	$\mid[x]\mid = \bar{x} - \underline{x}$	
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 $\diamond$	$[x] \diamond [y] = [\{x \diamond y \mid x \in [x], y \in [y]\}]$	
$+$	$[x] + [y] = [\underline{x} + \underline{y}, \bar{x} + \bar{y}]$	
$-$	$[x] - [y] = [\underline{x} - \bar{y}, \bar{x} - \underline{y}]$	
$\times$	$[x] \times [y] = [\min(\underline{x}\underline{y}, \underline{x}\bar{y}, \bar{x}\underline{y}, \bar{x}\bar{y}), \max(\underline{x}\underline{y}, \underline{x}\bar{y}, \bar{x}\underline{y}, \bar{x}\bar{y})]$	IF THE VALUE $0 \notin [y]$
/	$[\mathbf{x}] = [x_1] \times \cdots \times [x_n]$	SET OF VECTORS OF $\mathbb{R}^n$
BOX	$[x]/[y] = [x] \times [1/\bar{y}, 1/\underline{y}]$	
BOX VOLUME $\mid . \mid$	$\mid[\mathbf{x}]\mid$	
BOX SET THEORETIC OPERATOR $\triangle$	$[\mathbf{x}] \triangle [\mathbf{y}] = ([x_1] \triangle [y_1]) \times \cdots \times ([x_n] \triangle [y_n])$	
BOX BINARY OPERATOR $\diamond$	$[\mathbf{x}] \diamond [\mathbf{y}] = ([x_1] \diamond [y_1]) \times \cdots \times ([x_n] \diamond [y_n])$	

BAYESIAN INTERPRETATION FOR THE BOX-PF TIME UPDATE

In [4], it is shown that the box-PF can be seen as an approximation of the Bayes' filter by interpreting each box particle as a uniform pdf (with the limits of the box particle defining its support). The set of box particles is then interpreted as 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 2, the time update equation (3) can be approximated in the form

$$
\begin{aligned}
&p(\mathbf{x}_{k+1} \mid \mathbf{Z}_k) \\
&\approx \int_{\mathbb{R}^{n_x}} p(\mathbf{x}_{k+1} \mid \mathbf{x}_k) \sum_{i=1}^{N} w_k^i U_{[\mathbf{x}_k^i]}(\mathbf{x}_k) \, d\mathbf{x}_k \\
&= \sum_{i=1}^{N} w_k^i \int_{[\mathbf{x}_k^i]} p(\mathbf{x}_{k+1} \mid \mathbf{x}_k) U_{[\mathbf{x}_k^i]}(\mathbf{x}_k) \, d\mathbf{x}_k.
\end{aligned}
\tag{4}
$$

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

$$
\begin{aligned}
p(\mathbf{x}_{k+1} \mid \mathbf{x}_k) \cdot U_{[\mathbf{x}_k^i]}(\mathbf{x}_k) = 0 \\
\text{for } \forall \mathbf{x}_{k+1} \notin [f]([\mathbf{x}_k^i]) + [\mathbf{v}_k]. \quad (5)
\end{aligned}
$$

Equation (5) shows that for any transition function f, using interval analysis techniques, the support for the pdf terms $\int_{[\mathbf{x}_k^i]} p(\mathbf{x}_{k+1} \mid \mathbf{x}_k) U_{[\mathbf{x}_k^i]}(\mathbf{x}_k) \, d\mathbf{x}_k$ can be approximated by $[f]([\mathbf{x}_k^i]) + [\mathbf{v}_k]$. In the box-PF algorithm, each pdf term

$\int_{[\mathbf{x}_k^i]} p(\mathbf{x}_{k+1} \mid \mathbf{x}_k) U_{[\mathbf{x}_k^i]}(\mathbf{x}_k) \, d\mathbf{x}_k$ in (4) is modeled by one uniform pdf component having as support the interval $[f]([\mathbf{x}_k^i], [\mathbf{v}_k])$, i.e.,

$$
\begin{aligned}
&\int_{[\mathbf{x}_k^i]} p(\mathbf{x}_{k+1} \mid \mathbf{x}_k) U_{[\mathbf{x}_k^i]}(\mathbf{x}_k) \, d\mathbf{x}_k \\
&\approx U_{[f]([\mathbf{x}_k^i]) + [\mathbf{v}_k]}(\mathbf{x}_{k+1}).
\end{aligned}
\tag{6}
$$

Combining (4) and (6) gives

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

The box-PF strategy of approximating each pdf $\int_{[\mathbf{x}_k^i]} p(\mathbf{x}_{k+1} \mid \mathbf{x}_k) U_{[\mathbf{x}_k^i]}(\mathbf{x}_k) \, d\mathbf{x}_k$ 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 [3]. If a more accurate representation is required, then each term can be approximated as a mixture of uniform pdfs as shown in [4].

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 is $\mathbf{r}_{k+1}^i = \mathbf{z}_{k+1} - \mathbf{z}_{k+1}^i$, where $i = 1, \ldots, N$, and $\mathbf{z}_{k+1}^i = \mathbf{g}(\mathbf{x}_{k+1|k}^i)$ is the ith predicted measurement. Next, using the probabilistic model p_w for the measurement noise $\mathbf{w}_k$, the likelihood of each particle is calculated as $p(\mathbf{z}_{k+1} \mid \mathbf{x}_{k+1|k}^i) = p_w(\mathbf{z}_{k+1} - \mathbf{z}_{k+1}^i) = p_w(\mathbf{r}_{k+1}^i)$.

A reasonable assumption for the box-PF is that the measurement likelihood function has a bounded support that we call here the measurement likelihood box. In the bounded error context, the likelihood for each box particle is calculated using the following idea: a box particle whose corresponding predicted measurement does not intersect with the measurement likelihood box should be assigned a likelihood factor equal to zero. In contrast, a box particle whose corresponding predicted measurement is
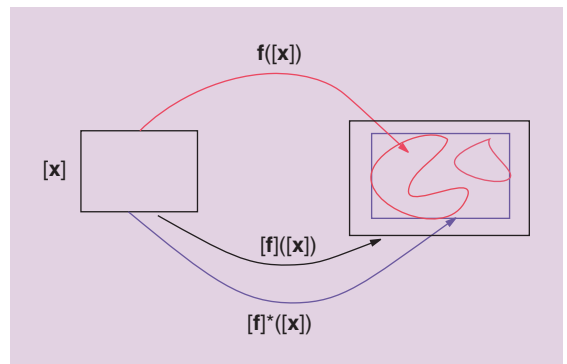
[FIG3] Inclusion functions obtained after applying the function **f** to a box **[x]**. The resulting image is not necessarily a box. A pessimistic inclusion function **[f]** and the minimal inclusion function **[f]***are presented.

[TABLE 2] 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 UNIFORM PDF COMPONENTS SPECIFIED BY BOXES $\{[\mathbf{x}^i]\}_{i=1}^{N}$ AND NORMALIZED WEIGHTS $\{w^i\}_{i=1}^{N}$
$p(\mathbf{x}_k\mid\mathbf{Z}_k) \approx \sum_{i=1}^{N} w_k^i U_{[\mathbf{x}_{k,i}^i]}(\mathbf{x}_k)$	POSTERIOR PDF $p(\mathbf{x}_k\mid\mathbf{Z}_k)$ APPROXIMATED AT TIME k BY A MIXTURE OF UNIFORM PDFs

included in the measurement likelihood box should have a likelihood equal to one. 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, to preserve an appropriate size of each box, a contraction step is performed that allows the elimination of the inconsistent part of the box particles with respect to the measured box. The contraction step for the box is analogous to the variance matrix measurement update step that appears in Kalman filtering. This contraction step utilizes interval analysis methods described in the next section.

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 3 introduces notations and definitions needed for the introduction of interval contraction concepts. In Table 3, the set 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 S, such that $S \subseteq [\mathbf{x}'] \subseteq [\mathbf{x}]$.

DEFINITION 2
Contracting $\mathcal{H}$ means replacing $[\mathbf{x}]$ by a smaller domain $[\mathbf{x}]'$ such that $S \subseteq [\mathbf{x}'] \subseteq [\mathbf{x}]$. 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, Ch. 4], including the Gauss elimination, Gauss-Seidel algorithms, 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. A simple illustration is given in the section "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 $[y] = [0, 1]$.

BAYESIAN INTERPRETATION OF THE CONTRACTION STEP
Assume that, at time instant $k + 1$, an approximation of the time update pdf $p(\mathbf{x}_{k+1}\mid\mathbf{Z}_k)$ by a mixture of N uniform pdfs with interval supports $[\mathbf{x}_{k+1|k}^{(i)}]$ and weights w_k^i is available and that the measurement update step is to be performed. Next, for the box-PF, a probabilistic model p_w for the measurement noise $\mathbf{w}_k$ is also available. It is assumed in general that p_w can be expressed by using a mixture of uniform pdfs. For simplicity and without loss of generality, p_w is considered here to be a single uniform pdf, such that the box measurement $[\mathbf{z}_{k+1}]$ contains all realizations of $\mathbf{g}(\mathbf{x}_{k+1}) + \mathbf{w}_k$. Then we have $p(\mathbf{z}_{k+1}\mid\mathbf{x}_{k+1}) = U_{[\mathbf{z}_{k+1}]}(\mathbf{g}(\mathbf{x}_{k+1}))$ and according to (2), the measurement update is

$$p(\mathbf{x}_{k+1}\mid\mathbf{Z}_{k+1})$$
$$= \frac{1}{\alpha_{k+1}} p(\mathbf{z}_{k+1}\mid\mathbf{x}_{k+1})p(\mathbf{x}_{k+1}\mid\mathbf{Z}_k)$$
$$= \frac{1}{\alpha_{k+1}} U_{[\mathbf{z}_{k+1}]}(\mathbf{g}(\mathbf{x}_{k+1}))\sum_{i=1}^{N} w_k^i U_{[\mathbf{x}_{k+1|k}^i]}(\mathbf{x}_{k+1})$$
$$= \frac{1}{\alpha_{k+1}}\sum_{i=1}^{N} w_k^i U_{[\mathbf{z}_{k+1}]}(\mathbf{g}(\mathbf{x}_{k+1})) U_{[\mathbf{x}_{k+1|k}^i]}(\mathbf{x}_{k+1}).$$
$$(8)$$

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

[TABLE 3] 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	$\mathbf{f} = (f_1, f_2, \ldots, f_m)^T$	
CONSTRAINT SATISFACTION PROBLEM $\mathcal{H}$	$\mathcal{H}: (\mathbf{f}(x) = \mathbf{0}, \mathbf{x} \in [\mathbf{x}])$	PROBLEM OF FINDING THE SMALLEST $[\mathbf{x}']$ ENCLOSING THE SET OF ALL $\mathbf{x}$ IN $[\mathbf{x}]$ SATISFYING $\mathbf{f}$
THE SOLUTION SET OF $\mathcal{H}$	$S = \{\mathbf{x} \in [\mathbf{x}] \mid f(\mathbf{x}) = \mathbf{0}\}$	

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

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 measurement $[\mathbf{z}_{k+1}]$. These contraction steps result in the new box particles denoted $[\tilde{\mathbf{x}}^i_{k+1}]$, which approximate the posterior pdf $p(\mathbf{x}_{k+1} | \mathbf{Z}_{k+1})$ at time $k+1$.

BAYESIAN DERIVATION OF THE LIKELIHOOD
Following the definition of the sets S_i in (9), we can write

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

By combining (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),

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

WEIGHT UPDATE
In the SIR PF, each particle weight is updated by a factor equal to the likelihood $p(\mathbf{z}_k | \mathbf{x}^i_{k+1|k})$, followed by normalization of weights. In the box-PF, this step is very similar, i.e., 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 $L^i_k = |[\tilde{\mathbf{x}}^i_{k+1}]| / |[\mathbf{x}^i_{k+1|k}]|$. In summary, $p(\mathbf{x}_{k+1} | \mathbf{Z}_{k+1})$ is approximated by $\{(\tilde{w}^i_{k+1}, [\tilde{\mathbf{x}}^i_{k+1}])\}^N_{i=1}$, where $\tilde{w}^i_{k+1} \propto w^i_k \cdot L^i_k$.

RESAMPLING BOX PARTICLES
Similar 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

characterized by high weights have a good chance of multiplying 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, to increase the "resolution" in the regions of the state space where the posterior pdf has high values, a box particle that has been selected n times during resampling can be partitioned into n disjoint smaller boxes. In fact, the manipulation of interval data always yields a very pessimistic solution as a result of the basic rules of the interval arithmetic and the wrapping effect when boxes are propagated via models [2]. Note that the wrapping effect will be more important if one has to propagate large-size boxes. Thus, the division of a box into several subboxes will make possible to refine the solution for the following steps.

The efficiency of this box-particle resampling strategy is empirically confirmed in [3].

LESSONS LEARNED, FURTHER READING, AND FUTURE AVENUES
This lecture note summarizes 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 through 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 [3]–[5].

Numerous challenges remain for future work. Many theoretical results are still missing, such as convergence results, better proposal densities and theoretical justification of resampling with partitioning. The significant reduction of the number of particles opens opportunities for distributed nonlinear and nonparametric state estimation problems. One important question that should be answered in future works is how to determine the number of

uniform pdfs involved in the approximation of the posterior, and how to merge and combine closed uniform distributions.

AUTHORS
Amadou Gning (e.gning@lancaster. ac.uk) is a research associate at University College London, United Kingdom.

Branko Ristic (branko.ristic@dsto. defence.gov.au) is a research scientist at the Defense Science and Technology Organization, ISR Division, Australia.

Lyudmila Mihaylova (mila.mihaylova@lancaster.ac.uk) is an associate professor at the School of Computing and Communications, Lancaster University, United Kingdom.

Fahed Abdallah (fahed.abdallah@hds. utc.fr) is an associate professor at HEU-DIASYC, UMR CNRS 6599, Université de Technologie de Compiègne, France.

REFERENCES
[1] A. Doucet, J. F. G. de Freitas, and N. J. Gordon, Eds., *Sequential Monte Carlo Methods in Practice.* New York: Springer-Verlag, 2001.

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

[3] F. Abdallah, A. Gning, and P. Bonnifait, "Box particle filtering for nonlinear state estimation using interval analysis," *Automatica*, vol. 44, no. 3, pp. 807–815, 2008.

[4] A. Gning, L. Mihaylova, F. Abdallah, and B. Ristic, "Particle filtering combined with interval methods for tracking applications," in *Integrated Tracking, Classification, and Sensor Management: Theory and Applications,* M. Mallick, V. Krishnamurthy, and B.-N. Vo, Eds. Hoboken, NJ: Wiley, 2012, ch. 2.

[5] A. Gning, B. Ristic, and L. Mihaylova, "Bernoulli/box-particle filters for detection and tracking in the presence of triple measurement uncertainty," *IEEE Trans. Signal Processing*, vol. 60, no. 5, pp. 2138–2151, 2012.

[6] A. H. Jazwinski, *Stochastic Processes and Filtering Theory.* New York: Academic, 1970.

[7] 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.

[8] U. D. Hanebeck, J. Horn, and G. Schmidt, "On combining statistical and set-theoretic estimation," *Automatica*, vol. 35, no. 6, pp. 1101–1109, 1999.

[9] R. Mahler, *Statistical Multisource-Multitarget Information Fusion.* Norwood, MA: Artech House, 2007.

[10] E. Wan and R. van der Merwe, "The unscented Kalman filter," in *Kalman Filtering and Neural Networks,* S. Haykin, Ed. New York: Wiley, Sept. 2001, ch. 7, pp. 221–280.

SP