Reinforced Likelihood Box Particle Filter

Quoc Hung Lu, Soheib Fergani, Carine Jauberthie

***Abstract*—This paper is concerned with the development of a general scheme for box particle filtering. It is based on the likelihood computation, the most crucial step of the estimation strategy. The proposed filter takes advantages from strong aspects of various existing box particle filters and adds an interesting reinforced likelihood computation method that enhances the estimation results. An overview on Box Particle Filters and discussions from assumptions used in the literature to the filters performance evaluation approach are presented. Also, comparative study of the obtained results by performing several scenarios on the illustration example is provided to highlight the efficiency of the proposed estimation strategy.**

***Index Terms*—State filtering and Estimation, Nonlinear System, Box Particle Filter, Interval Analysis.**

I. Introduction

In State Estimation or Filtering problems, when dealing with a linear Gaussian state-space model, analytical expressions computing the state estimates according to posterior distributions can be derived by the well known and widespread Standard Kalman Filter (SKF) [1]. Many extension of SKF are then provided by numerous researches in different contexts [2]–[4]. For nonlinear model without Gaussian measurement assumption, Particle Filters (PF) have been applied successfully to a variety of state estimation problems [5], [6]. The PF efficiency and accuracy depend mostly on the number of particles used in the estimation which may require a high computation time.

One of the famous extensions of PF to set membership approach is the Box Particle Filter (BPF) [7]. BPF handles box (interval vector of) states and bounded errors by using interval computation and constraint satisfaction techniques. This method has been shown to control quite efficiently the number of required particles, hence reducing the computational cost and providing good results in several experiments.Therefore, it has been applied in many applications, included aerial [8] and ground vehicle [9] estimations.

Since then, numerous variants of BPF are developed to deal with measurements bounded uncertainty, measurements stochastic uncertainty or measurements mixed uncertainty. Various techniques and theories have been proposed to address the diversity of requirements in these contexts, e.g. weight updating using Bayesian filtering technique extending to box particle case [10] or belief function theory with different methods [11], [12].

In the present work, regarding this large variety of BPF, a scheme is proposed to give a generalized description that highlights the specificity of this class of filters. An analysis of the likelihood computation methodology is investigated. Theory background is also provided to support the development of a novel filter benefiting advantages of existing methods via a numerous reinforcement techniques: score function, reduction percentage, exponential weighting, backward estimate.

The paper is organized as follows. The problem formulation is presented in Section II with discussions about assumptions used in the literature. Section III presents the general scheme of BPF. The likelihood computation methodology is analyzed in Section IV which also provides the notion of likelihood distinguishability. The proposed method is developed in Section V and simulations of the method are provided in Section VI. Section VII is the paper conclusion.

Date of first submission for review: April 19, 2022.

Authors are affiliated at LAAS-CNRS, Université de Toulouse, CNRS, UPS, Toulouse, France. Emails: qhlu,sfergani,cjauber@laas.fr

II. Problem formulation

A. Notations and definitions

A real interval matrix $[X]$ of dimension $p\times q$ is a matrix with real interval components $[x_{ij}]$, $i \in \{1,...,p\}$, $j \in \{1,...,q\}$. Write $X \in [X]$ to indicate a point matrix $X = (x_{ij})$ belonging element-wise to $[X]$. Define:

$$\begin{aligned} \overline{X} &\equiv \sup([X]) \triangleq (\sup([x_{ij}])), \\ \underline{X} &\equiv \inf([X]) \triangleq (\inf([x_{ij}])), \end{aligned}$$

and then $\text{mid}([X]) \triangleq (\overline{X}+\underline{X})/2$, $\text{rad}([X]) \triangleq (\overline{X}-\underline{X})/2$, $\text{width}([X]) \triangleq \overline{X}-\underline{X}$. Denote also $\overline{X} = \sup([X])$, $\underline{X} = \inf([X])$, $[X] = [\underline{X}, \overline{X}]$ and define the (convex) hull of two interval matrices $[X_1]$, $[X_2]$ of the same dimension as $\text{hull}\{[X_1],[X_2]\} \triangleq [\min\{\underline{X}_1,\underline{X}_2\}, \max\{\overline{X}_1,\overline{X}_2\}]$.

Basic interval operators $\diamond \in \{+,-,\times,\div\}$ defined in [13] can be used to compute directly all operations $[u]\diamond[v]$ and $\alpha\diamond[u]$, for real intervals $[u]$, $[v]$ and $\alpha \in \mathbb{R}$, without any further approximation algorithm. Then, interval matrix computations are defined similarly to matrix computations using the basic operators while more general operators are constructed by meant of inclusion function [13]. In practice, the package Intlab developed for Matlab is used for these computations.

B. Assumptions and discussions

Consider the following dynamical system:

$$(\Sigma): \begin{cases} x_k &= f(x_{k-1},u_k,w_k) \\ y_k &= h(x_k,u_k,v_k) \end{cases}, \quad k \in \mathbb{N}^*, \tag{1}$$

where $x_k \in \mathbb{R}^{n_x}$ and $y_k \in \mathbb{R}^{n_y}$ are respectively state and measurement output, $u_k \in \mathbb{R}^{n_u}$ input, $w_k \in \mathbb{R}^{n_w}$ state dynamic disturbance and $v_k \in \mathbb{R}^{n_y}$ measurement noise.

***Assumption (A)*: State Process Uncertainty**

u_k and w_k are unknown and belong to known intervals $[u_k]$ and $[w_k]$ respectively.

***Assumption (B)*: Measurement Bounded Uncertainty**

(B1) v_k (unknown) belongs to known interval $[v_k]$.

(B2) The measurements are intervals $[y_k]$.

(B3) The measurements are assumed to be accurate in the sense that $[y_k] \ni h(x_k, u_k) \equiv h(x_k, u_k, 0)$ (the zero noise case), where x_k is the real state.

***Assumption (C)*: Measurement Stochastic Uncertainty**

(C1) v_k are additive noises with known density p_v.

(C2) The measurements are point values y_k.

***Assumption (D)*: Measurement Mixed Uncertainty**

(D1) v_k are additive Gaussian noises with unknown mean $\mu_k \in \mathbb{R}^{n_y}$ and covariance $\Sigma_k \in \mathbb{R}^{n_y \times n_y}$.

(D2) μ_k, Σ_k belong to known intervals $[\mu_k]$, $[\Sigma_k]$.

(D3) The measurements are point values y_k.

Assumption (A) is used in [7], [10]–[12].

Assumptions (B) are under study in [7], [11]. In [7], the BPF is introduced and becomes standard for many extensions or variants with essential steps: *initialization, box propagation, contraction, likelihood (weight) computation, state estimation and resampling*. In [11], the Belief State Estimation algorithm is developed using the belief function theory. It may require some techniques for the construction and computation of masses, but after being normalized, these masses become likelihoods in the probability sense. Therefore, we also call likelihood computation as an essential step of this method.

Assumptions (C) are used in [10]. The method proposed therein includes a different approach to weight the box particles as well as a resampling procedure based on repartitioning the box enclosing the propagated states. There is no contraction step in this method.

Assumptions (D) are used in [12], in which (D1) is a special case of (C1) with a slight relaxation by adding bounded uncertainties to Gaussian parameters μ_k and Σ_k. In [12], the belief function theory is used with continuous mass functions to represent these kinds of uncertainties and to compute box particle likelihoods. The proposed approach therein leads to the so-called Evidential Box Particle Filter (EBPF) including all essential steps of the standard BPF.

Remark 1: (B3) is the implicit assumption deriving the consistency between the predicted measurement boxes $[h]([x_k^i], [u_k])$, $i \in \{1, ..., M\}$ (M the number of partitioned boxes), and the real measurement box $[y_k]$. This consistency is used in the contraction step and the likelihood computation by penalizing all particle boxes with which the intersections $[h]([x_k^i], [u_k]) \cap [y_k]$ are empty. □

Remark 2: Assumptions (D3) and (C2) are coincided. They can be transformed into (B2) with a slight relaxation of (B3). That is, knowing the density of v_k, we deduce its confidence intervals $[v_k]$ with some significant level α and define $[y_k] \triangleq y_k - [v_k]$. Then (B3) is relaxed in the sense that the observed measurements $[y_k]$ do not contain $h(x_k, u_k)$ with certainty but with only a probability $(1 - \alpha)$. □

III. General scheme of Box Particle Filter

In general, although applying different background theories, the proposed methods in [7], [10]–[12] study State Estimation in a framework of stochastic uncertainties and/or bounded uncertainties with two main objectives :

Objective 1: Reduce the width of box particles to penalize the conservatism due to interval computations.

Objective 2: Quantify (compute) box particle likelihoods, as best as possible, to enhance the accuracy of the estimates.

The methods used in these references can be considered as variants of BPF and be summarized by Scheme 1 which is applied in a mostly similar manner across them.

Scheme 1 General Scheme of Box Particle Filtering

STEP 1 : Initialization. At a time instance $k_0 \geq 0$, (re)partition the interval $[x_{k_0}]$ or resample the set $\{[x_{k_0}^j], w_j\}_{j=1:N_{k_0}}$ into M disjoints equal-volume sub-boxes with the same weights: $\{[x_{k_0}^i], w_i = 1/M\}_{i=1:M}$.

while a predetermined Condition C is still satisfied **do**

STEP 2 : Propagation. Get a new set of box particles $\{[x_{k_0+1}^i] = [f]([x_{k_0}^i], [u_{k_0}])\}_{i=1,...,M}$ estimating the box containing the real state $x_{k_0+1} = f(x_{k_0}, u_{k_0})$ with or without a contraction step.

STEP 3 : Likelihood computation

a) Compute (and normalize) the likelihoods of box particles $\{[x_{k_0+1}^i]\}_{i=1:M}$ being the box containing the real state x_{k_0+1}. This computation bases on the consistency between the estimated measurement $[h]([x_{k_0+1}^i], [u_{k_0}])$'s and the obtained measurement $[y_{k_0+1}]$ using different criteria and methods.

By this step, the following set of box particles with updated weights is obtained : $\{[x_{k_0+1}^i], w_i\}_{i=1:M}$.

b) Some techniques can be applied at this step to get a more "efficient" set of box particles, e.g. discarding the boxes with small weights (smaller than some predetermined threshold) and with/without replicating the box associated with the greatest weight,...

From this, the set of box particles becomes $\{[x_{k_0+1}^i], w_i\}_{i=1:N_{k_0+1}}$, $1 \leq N_{k_0+1} \leq M$.

STEP 4 : Estimation

$$\text{Interval estimate : } [x_{k_0+1}] = \sum_{i=1}^{N_{k_0+1}} w_i.[x_{k_0+1}^i] \quad (2)$$

$$\text{Point estimate : } x_{k_0+1} = \sum_{i=1}^{N_{k_0+1}} w_i.\text{mid}([x_{k_0+1}^i]) \quad (3)$$

STEP 5 : $k_0 = k_0 + 1$

end while

STEP 6 : Restarting at STEP 1.

Remark 3: In this Scheme, for a general presentation, the observed measurements are denoted as intervals since the point values are considered as special cases of intervals. N_{k_0} in the initialization step takes value in $\{1, ..., M\}$. It is the number of box particles obtained at the end of the likelihood computation step at the previous time instant $(k_0 - 1)$. For $k_0 = 0$, the initialization concerns only the partition of $[x_0]$ and not the resampling. The Condition C of the while loop in the Scheme may differ depending on the used method, e.g. in [7], this condition is $N_{k_0} \geq N_{eff}$ where N_{eff} is a threshold calculated at each iteration of the loop. □

IV. Likelihood computation analysis

A. Likelihood computation methodology

The diagram in Fig.1 is provided to explain the methodology

of LCMs used in Scheme 1.

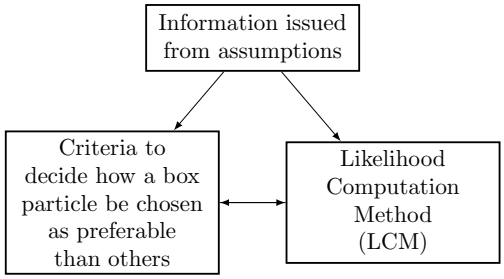


Fig. 1. Likelihood computation methodology diagram

The assumptions of the system under consideration supply the information, or Info for short, needed to build the likelihood. The information can be, for instance:

- **Info (a)**: The intersection between $[y_k]$ and the box $[h]([x_k^i],[u_k])$ containing the real value y_k must be non empty, ($[h]([x_k^i],[u_k]) \cap [y_k] \neq \emptyset$).
- **Info (b)**: The distribution of v_k and hence of $r_k = y_k - h(x_k, u_k)$ is Gaussian (for additive noise v_k).

The information can be directly an assumption or a deduction of the later. In bounded-error context, only Info (a) is treated [7] while in the mixed uncertainty case, both Info (a) and Info (b) are taken into account [12]. Criteria and methods are then chosen to exploit the information. Once a criterion is chosen, different methods can be used to calculate the likelihood. Inversely, a calculation method may correspond to many criteria. More precisely, a likelihood computation criterion is an application of the supplied information to establish an *order of preferability* of related events/objects that we want to compute their likelihoods, while a LCM is a method using mathematical formulae to represent and quantify the preferability by giving scores and normalizing them. Definition 1 formalizes these concepts.

Definition 1: Any set $\{s_i > 0\}_{i=1:n}$ can be a set of scores. This set has an increasing (decreasing) order of preferability if the greater (smaller) score is preferable. The set $\{L^i\}_{i=1:n}$ is a set of likelihoods if $L^i \in [0,1]$, $i \in \{1,...,n\}$ and $\sum_{i=1}^n L^i = 1$. A normalizing function is a function $\Phi : \mathbb{R}^n \to [0,1]^n$ transforming scores to likelihoods so that a preferable score corresponds to the greater likelihood.

Let's illustrate the above ideas via Example 1.

Example 1: To exploit Info (a), one can use either:

- **Criterion 1**: The particle $[x_k^i]$ giving a "bigger" intersection determined by $[\hat{z}_k^i] = [h]([x_k^i],[u_k]) \cap [y_k]$ must be preferable,
- **Criterion 2**: The particle $[x_k^i]$ making $[\hat{y}_k^i] = [h]([x_k^i],[u_k])$ "closer" to $[y_k]$ must be preferable.

How to represent "bigger" (size) or "closer" (closeness) notions and to calculate the corresponding likelihoods depend on the choice of LCMs. Criterion 1 is used in [7]. The corresponding LCM uses the volume Vol(.) determined by

$$\text{Vol}([x]) \triangleq \prod_{j=1}^{n_x} \text{width}([x_j]), \quad [x] = ([x_1],...,[x_{n_x}])^T, \quad (4)$$

to represent the box size and computes the likelihoods $L_1 = (L_1^1,...,L_1^M)$ by :

$$s_1^i = \frac{\text{Vol}([z_k^i])}{\text{Vol}([\hat{y}_k^i])}\ ,\ i \in \{1,...,M\}\ ,\ L_1^i = \frac{s_1^i}{\sum_{i=1}^M s_1^i}, \quad (5)$$

where $\{s_1^i\}$ are scores with increasing order of preferability. However, other methods can be used to calculate the likelihoods, e.g. with scores $s_j = (s_j^1,...,s_j^M)$, $j \in \{2,3,4\}$:

$$s_2^i = \text{Vol}([z_k^i]),\ s_3^i = \|\text{width}([z_k^i])\|_\infty,\ s_4^i = \|\text{width}([z_k^i])\|_2,$$

in which s_3 and s_4 use distances between the two bounds of the intersection to represent its size. Criterion 2 can be applied with different LCMs using a distance between $[\hat{y}_k^i]$ and $[y_k]$ to represent their closeness. Criterion 2 can also be used to exploit Info (b) as in [12] via the central tendency of the Gaussian vector $[r_k^i] = y_k - [\hat{y}_k^i]$ along with the belief function theory. A more detail analysis of the method used in [12] is found in section IV-A. □

It is worth to note that, in some cases, it is difficult to distinguish clearly between criterion and LCM as illustrated by Fig.1, e.g. in [10] with Interval Bayes filtering approach or in [11] and [12] with the belief theory. The reason is that the criteria are implied under complexes theories.

B. Distinguishability of computed likelihoods

In order to deal with **Objective 1**, in the literature, contractors are usually applied based on the Constraint Satisfaction Problem technique. However, this is not the most crucial step of BPFs using Scheme 1, e.g. this step is skipped in [10]. Furthermore, partition a box into M disjoint equal-volume sub-boxes and then compute the expected interval by (2) also help to reduce the conservatism due to interval computations. The most crucial step that differs one method from another in this class of BPFs is the Likelihood computation focusing on **Objective 2**. This is thus the main discussion of this section. Let's begin by introducing the definition of distinguishability of likelihoods (Definition 2).

Definition 2: Let $\{L^i\}_{i=1:n}$ be a set of likelihoods, define the distinguishability of two likelihoods of the set by $\delta_{i,j} \equiv \delta_{j,i} \triangleq |L^i - L^j|$, $1 \leq i \neq j \leq n$, and the total distinguishability of the set by $\delta_T \triangleq \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} \delta_{i,j} \equiv \sum_{i<j} \delta_{i,j}$.

In general, the box likelihoods are computed at every time instance k. The more they can represent the ability of a box containing the real value, the better estimate is obtained by Estimation step. A *weak distinguishability* means that most of the computed likelihoods are quasi equal and hence not useful for distinguishing between box particles. In the remainder of this section, two representative groups of criteria and LCMs used in the literature are analyzed to show their major disadvantage which is the weak distinguishability.

Group I : Apply Criterion 1 with LCMs using the box volume for the box size representation.

This criterion is used implicitly in the Contraction step of all BPF algorithms including it and applied in [7] with the LCM L_1 defined by (5).

Consider real interval vector $[y_k] = [\underline{y}_k, \overline{y}_k]$ and real point vector δ such that $0 \leq \delta \leq \overline{y}_k$ (element-wise), $\delta \in \mathbb{R}^p$. Let

$T = \text{diag}\{t_1, ..., t_p\}$ be a diagonal matrix with diagonal entries $\{t_r \geq 0\}_{r=1:p}$. Then, all boxes $[\hat{y}_k^i]$'s with forms:

$$\left[\underline{y}_k - T\delta, \underline{y}_k + \delta\right] \quad \text{or} \quad [\overline{y}_k - \delta, \overline{y}_k + T\delta]$$

have the same likelihoods $L_1^i = 1/\prod_{r=1}^{p}(1+t_r)$ (using (5)). In this case, these likelihoods have null distiguishability.

There are many other cases in which likelihoods are quasi equal and thus making the corresponding boxes $[\hat{y}_k^i]$ weakly distinguishable. For instance, M boxes $[\hat{y}_k^i]$'s may have likelihoods $L^i = 1/M + \delta_i$, $\delta_i \in (-\epsilon, \epsilon)$ with a small $\epsilon > 0$. In this case, the benefit of the likelihood computation step could be insignificant.

Group II : Use Criterion 2 to exploit Info (b).

The LCM used in [12] is investigated as the representative method of this group to deal with stochastic or mixed uncertainties and with additive Gaussian measurement noises. In this method, the innovation term $r_k = y_k - h(x_k, u_k)$ is Gaussian with $\mu_k \in [\mu_k]$ and $\Sigma_k \in [\Sigma_k]$. It belongs to some of intervals $[r_k^i] = y_k - [\hat{y}_k^i]$, $i \in \{1, ..., M\}$. One defines:

$$[HV_\alpha] = \left[\underline{\mu}_k - \sqrt{\alpha \text{Diag}(\overline{\Sigma}_k)}, \overline{\mu}_k + \sqrt{\alpha \text{Diag}(\overline{\Sigma}_k)}\right],$$

where $\alpha \geq 0$, $\sqrt{(.)}$ is an element-wise operator and $\text{Diag}(X)$ returns the diagonal of matrix X as a vector. Then, the belief $bel(.)$ and plausibility $pl(.)$ of $[r_k^i]$ are computed and considered as lower and upper bound of the probability of $[r_k^i]$ containing the real value r_k as follows:

$$bel([r_k^i]) = F_{n_y+2}(\alpha_{bel}^i) \text{ and } pl([r_k^i]) = 1 - F_{n_y+2}(\alpha_{pl}^i),$$
$$\alpha_{bel}^i = \max\{\alpha : [HV_\alpha] \subseteq [r_k^i]\},$$
$$\alpha_{pl}^i = \min\{\alpha : [HV_\alpha] \cap [r_k^i] \neq \emptyset\},$$
$$bel([r_k^i]) \leq \text{Probability}\left([r_k^i] \ni r_k\right) \leq pl([r_k^i]),$$

where F_{n_y+2} is the cumulative distribution function of the χ^2 distribution with $n_y + 2$ degrees of freedom. At this stage, Criterion 2 is applied based on the central tendency of the Gaussian vector r_k: the more $[\hat{y}_k^i]$ is close to y_k (equivalently $[r_k^i]$ is close to $[\mu_k]$), the greater belief and plausibility the box particles $[x_k^i]$ (that yield $[\hat{y}_k^i]$ via the function h) attains. Finally, the likelihood of each particle $[x_k^i]$ is computed thanks to the Generalized Bayes theorem (GBT) and Pignistic transformation [12].

The weak distinguishability of the method is shown via the following critical point. All boxes $[r_k^i]$ intersecting $[\mu_k]$ have the plausibility 1. So, these boxes are not distinctive regarding their plausibilities. They are distinguished only by their beliefs, in which:

- For the boxes that intersect $[\mu_k]$ but do not contain it, their beliefs are 0. A zero information can be issued about these boxes.
- For the ones containing $[\mu_k]$, their beliefs are characterized by the greatest focal element $[HV_{\alpha_{bel}^i}]$ they contain. The greater $[HV_{\alpha_{bel}^i}]$ a box can contain, the more belief it gets. It is quite similar to apply the rule: *"the more $[r_k^i]$ is centralized (having a bigger intersection with $[\mu_k]$) and has a bigger volume, the greater likelihood it gets"*. Other LCMs can be applied using that rule with a lightened calculation strategy and background theory.

Therefore, the likelihoods computed in the next step using GBT and Pignistic transformation are weakly distinguishable. The computation formulae are as follows [12]:

$$L_k^i = \sum_{A \subset \Omega, A \neq \emptyset} \frac{m(A|y_k)}{|A|} \cdot \mathbf{1}([x_k^i] \in A), \ \forall [x_K^i] \in \Omega, \quad (6)$$

$$m(A|y_k) = \eta \prod_{[x_k^i] \in A} pl\left([r_k^i]\right) \prod_{[x_k^j] \notin A} \left[1 - pl\left([r_k^j]\right)\right], \quad (7)$$

where A is a subset of $\Omega = \{[x_k^i], i = 1 : M\}$ with cardinality $|A|$, $[r_k^i] = y_k - [h]\left([x_k^i], [u_k]\right)$, $\mathbf{1}(x) = 1$ if x holds true and vanishes otherwise and $\eta = 1 - \prod_{[x_k^i] \in \Omega}\left[1 - pl\left([r_k^i]\right)\right]$.

Example 2: Let $\Omega = \{[x_k^1], [x_k^2], [x_k^3] \in \mathbb{IR}^2\}$ and put $[r_k^i] = y_k - [h]([x_k^i], [u_k])$, $i = 1, 2, 3$. Assume that $v_k \sim \mathcal{N}(\mu_k, \Sigma_k)$, $\mu_k \in [\mu_k]$, $\Sigma_k \in [\Sigma_k]$, $[r_k^i] \cap [\mu_k] \neq \emptyset$, $i = 1, 3$, and $[r_k^2] \cap [\mu_k] = \emptyset$. Then, we get $pl([r_k^i]) = 1$ for $i = 1, 3$ and $0 < pl([r_k^2]) < 1$. Using (6) and (7), we obtain :

$$L_k^1 = L_k^3 = \frac{1 - pl([r_k^2])}{2} + \frac{pl([r_k^2])}{3}, \quad L_k^2 = \frac{pl([r_k^2])}{3}.$$

Having the same likelihood, $[x_k^1]$ and $[x_k^3]$ are thus indistinguishable. If furthermore $pl([x_k^2])$ is close to 1 then all the three likelihoods are quasi equal. □

V. Reinforced Likelihood Box Particle Filter

A. Assumptions

Consider system (Σ) with **Assumptions (A), (B2), (D1)** and **(D2)**. Under these assumptions, Info (a) and Info (b) are concerned for likelihood computation.

Remark 4: The above measurement assumptions concern sensor errors (B2) and model (stochastic) uncertainties (D1). By (B2), the measurements are intervals $[y_k]$. Regarding Remark 2, it is necessary to replace $[y_k] \leftarrow [y_k] - [v_k]$ where $[v_k]$ is the 99.7% confidence interval of v_k determined by

$$[v_k] = \left[\underline{\mu}_k - 3\sqrt{\text{diag}(\overline{\Sigma}_k)}, \overline{\mu}_k + 3\sqrt{\text{diag}(\overline{\Sigma}_k)}\right], \quad (8)$$

as proposed in [12]. This treatment generalizes Remark 2. □

B. Method and Algorithm

The proposed method is named by Reinforced Likelihood Box Particle Filter (RLBPF) aiming to benefit advantages of existing criteria and LCMs and also attaining a gain in computation time. The method is based on the following two Propositions.

Proposition 1: Let $\{s_i\}_{i=1:n}$ be a set of scores so that $1 \leq s_1 \leq s_2 \leq ... \leq s_n$ and having an increasing order of preferability. Let Φ, $\tilde{\Phi}$ be normalizing functions so that $(L^1, ..., L^n) = \Phi(s_1, ..., s_n) = (s_1, ..., s_n)/\sum_{k=1}^n s_k$ and $(\tilde{L}^1, ..., \tilde{L}^n) = \tilde{\Phi}(s_1, ..., s_n) = (e^{s_1}, ..., e^{s_2})/\sum_{k=1}^n e^{s_k}$. Then the total distinguishability of the set $\{\tilde{L}^i\}_{i=1:n}$ is greater than the total distinguishability of the set $\{L^i\}_{i=1:n}$.

Proof 1: One must prove

$$\frac{\sum_{i<j}(s_j - s_i)}{\sum_{k=1}^n s_k} < \frac{\sum_{i<j}(e^{s_j} - e^{s_i})}{\sum_{k=1}^n e^{s_k}}.$$

One can expresses:

$$\begin{aligned} LHS &= \frac{\sum_{k=2}^{n} s_k - (n-1)s_1 + \sum_{k=3}^{n} s_k - (n-2)s_2 + }{\sum_{k=1}^{n} s_k} \\ &+ \frac{... + s_n - s_{n-1}}{\sum_{k=1}^{n} s_k} \\ &= (n-1) - 2\sum_{i=1}^{n-1}(n-i)\frac{s_i}{\sum_{k=1}^{n} s_k} \end{aligned}$$

and the RHS has a similar expression form with replacing s_i by e^{s_i}, $i = 1, ..., n$. So :

$$\begin{aligned} RHS - LHS &= 2\sum_{i=1}^{n-1}(n-i)\left(\frac{s_i}{\sum_{k=1}^{n} s_k} - \frac{e^{s_i}}{\sum_{k=1}^{n} e^{s_k}}\right) \\ &= 2\sum_{i=1}^{n-1}\frac{(n-i)\sum_{k\neq i}(e^{s_k}s_i - e^{s_i}s_k)}{\sum_{k=1}^{n} s_k \sum_{k=1}^{n} e^{s_k}} \\ &= \frac{2\sum_{i=1}^{n-1}\sum_{k=i+1}^{n}(k-i)(e^{s_k}s_i - e^{s_i}s_k)}{\sum_{k=1}^{n} s_k \sum_{k=1}^{n} e^{s_k}} \end{aligned}$$

For $k > i$, put $\Delta_{ik} = s_k - s_i$, then

$$\begin{aligned} e^{s_k}s_i - e^{s_i}s_k &= e^{s_i}s_i(e^{\Delta_{ik}} - 1) - e^{s_i}\Delta_{ik} \\ &> e^{s_i}\Delta_{ik}(s_i - 1) \geq 0 \end{aligned}$$

since $e^{\Delta_{ik}} - 1 > \Delta_{ik}$ and using the assumption $s_i \geq 1$, $\forall i = 1, ..., n$. The proposition is proved.

Proposition 2: Let Φ be a normalizing function so that $\Phi(a_1, ..., a_m) = (a_1, ..., a_m)/\sum_{k=1}^{m} a_k$, $\forall\{a_i\}_{i=1:m}$. Let $\{s_i\}_{i=1:n}$ be a set of scores with an increasing order of preferability so that $s_1 \leq ... \leq s_n$ and $\{s_i\}_{i=k:n}$ its truncated set with $k \geq 1$. Let $(L^1, ..., L^n) = \Phi(s_1, ..., s_n)$ and $(\tilde{L}^1, ..., \tilde{L}^{n-k+1}) = \Phi(s_k, ..., s_n)$. Then

$$L^i \leq \tilde{L}^{i-k+1}, \quad k \leq i \leq n, \tag{9}$$

and the distinguishability of the truncated set is so that

$$L^j - L^i \leq \tilde{L}^{j-k+1} - \tilde{L}^{i-k+1}, \quad k \leq i < j \leq n. \tag{10}$$

Proof 2: The proposition is proved by applying directly related definitions.

Principle: The scores $\{J_k^1, ..., J_k^M\}$ are associated to M considered particles as follows:

$$J_k^i = \left(d_{k,1}^i + d_{k,2}^i\right)V_k^i, \quad i \in \{1, ..., M\}, \tag{11}$$

where

- $d_1^i = d_H([y_k], [\hat{y}_k^i])$ (Hausdorff distance),
- $d_2^i = \|\text{mid}([y_k]) - \text{mid}([\hat{y}_k^i])\|_2$ (Euclidean norm),
- $V_k^i = \dfrac{\text{Vol}([\hat{y}_k^i] \setminus [y_k])}{\text{Vol}([\hat{y}_k^i])} = 1 - \dfrac{\text{Vol}([\hat{y}_k^i] \cap [y_k])}{\text{Vol}([\hat{y}_k^i])}$,

where Vol(.) is defined by (4). Thereby, J_k^i's measure the closeness between $[\hat{y}_k^i]$'s and $[y_k]$ via both a kind of maximum distance $d_{k,1}^i$ and a kind of concentric tendency measure $d_{k,2}^i$. J_k^i's also take into account the size of intersections $[\hat{y}_k^i] \cap [y_k]$ via the volume proportions V_k^i's. Consequently, J_k^i's exploit at the same time Info (a) and Info (b) and meets both Criterion 1 and Criterion 2. The particles having small scores are preferable, or equivalently the set $\{J_k^i\}_{i=1:M}$ has a decreasing order of preferability.

Then, the following strategy is applied:

+ $\{J_k^i\}_{i=1:M}$ is sorted in an ascending direction and a reduction percentage $R\%$ is applied, i.e. $N_h = \lfloor(100 - R)\%M\rfloor$ particles corresponding to the scores $\{J_k^i\}_{i=1:N_h}$ of the sorted set $\{J_k^i\}_{i=1:M}$ are retained.
+ the likelihoods (or weights) are computed as:

$$W_k^i = \frac{\exp\{-J_k^i + c\}}{\sum_{p=1}^{N_h}\exp\{-J_k^i + c\}}, \; i = 1 : N_h, \tag{12}$$

where $c \geq J_k^{N_h} + 1$ to let $-J_k^i + c \geq 1$, $i = 1 : N_h$.

This strategy ensures an increasing of the distinguishability of the computed likelihoods thanks to Proposition 1 and 2.

After computing estimate $[x_{k+1}]$ according to (2), a backward estimate is added as follows:

$$[x_k] = \text{hull}\{[x_k^i]\}, \tag{13}$$

for those $[x_k^i]$'s correspond to W_k^i's just computed to reduce the conservatism due to interval computation.

Algorithm 1 Reinforced Likelihood Box Particle Filter

Initialization: $[x_0] \equiv [\hat{x}_0]$, $R\%$, M, $[u_k]$, $[w_k]$, $[y_k]$, $[\mu_k]$, $[\Sigma_k]$, $k = 1, ..., N$, $N_h = \lfloor(100 - R)\%M\rfloor$.
for $k = 1, 2, 3, ...N$ **do**
  Partition $[\hat{x}_{k-1}]$ to M disjoint boxes $\{[\hat{x}_{k-1}^i]\}_{i=1:M}$
  $[\hat{x}_k^i] = [f]([\hat{x}_{k-1}^i], [u_k], [w_k])$, $i = 1, ..., M$
  $[\hat{y}_k^i] = [h]([\hat{x}_k^i], [u_k])$, $i = 1, ..., M$
  Compute $J_k = (J_k^1, ..., J_k^M)$ using (11).
  Sort J_k in ascending direction: $J_k = \text{sort}(J_k)$
  Hold N_h first values: $J_k = J_k(1 : N_h)$
  Compute W_k^i, $i = 1 : N_h$, using (12).
  $[\hat{x}_k] = \sum_{i=1}^{N_h} W_k^i.[\hat{x}_k^i]$; $[\hat{x}_{k-1}] = \text{hull}(\{[\hat{x}_k^i]\}_{i=1:N_h})$
end for

Remark 5: BPFs often use a non large number of particles to gain computation time and reduce the loss of a guaranteed estimation. Consequently, the resampling/repartition step happens at every or only after a few iterations. Thus, the fact that we hold previous weights and update them has no significant effect while this effect might not be quantified easily. Furthermore, conditions under which this procedure is implemented base usually on some heuristic choice of a threshold. It is also an issue of discussion but out of the scope of the present paper. Therefore, the RLBPF uses a reasonable small number of particles, performs the repartition at each iteration and strengthens the likelihood computation and the estimation by more efficient strategies. □

VI. Simulation

Consider the following nonlinear system which was used as an illustration example in [12]. It is used in the present work to compare the proposed method (RLBPF) to the one (EBPF) in the reference, for the system states estimation.

$$\begin{aligned} x_{k+1} &= \begin{pmatrix} \alpha_{k,1} & 1 \\ 1 - \alpha_{k,1} & \alpha_{k,2} \end{pmatrix} x_k + \text{diag}\{\beta_{k,1}, \beta_{k,2}\}u_k \\ &+ \text{diag}\{20, 10\}w_k, \\ y_k &= x_{k,2}x_k/10 + v_k, \end{aligned} \tag{14}$$

with $x_k = (x_{k,1}, x_{k,2})^T$, $u_k \in [u_k] = 5([15, 17], [-7, -5])^T$, $w_k \in [w_k] = ([-1, 1], [-1, 1])^T/10^2$ and for $i \in \{1, 2\}$: $\beta_{k,i} = 0.5 + e_i^T x_k/20^2$, $\alpha_{k,i} = (0.2 + e_i^T\delta_k/20)(2\beta_{k,i} - 0.5)$, $e_1 = (-1, 1)^T$, $e_2 = (1, 2)^T$, $\delta_k \in [\delta_k] = 10[w_k]$. The initial

state is $x_0 = [90, 80]^T$ with $[x_0] = ([85, 103], [75, 91])^T$, the number of iteration $N = 10^4$ and $v_k \sim \mathcal{N}(\mu_k, \Sigma_k)$ where $\mu_k \in [\mu_k] = ([-1, 1], [-1, 1])^T$ and $\Sigma_k \in [\Sigma_k] = [90, 200].\text{diag}\{1, 1\}$.

In order to evaluate how a LCM of a filter (e.g. RLBPF or EBPF) sharing Scheme 1 improves the efficiency of the estimation, we propose to compare the result of the filter with those of the following basic scenarios of Scheme 1:

- **Scenario 1**: Use the contraction step without partition (1 box);
- **Scenario 2**: Use equi-likelihood $\frac{1}{M}$ without contraction step (M boxes);
- **Scenario 3**: Use equi-likelihood $\frac{1}{M}$ with contraction step (M boxes).

These basic scenarios can be considered as simple BPFs. The reason of this proposition is that, in some applications, using solely the contraction step, the algorithm performance has been rather good and the efficiency brought by the LCM might be insignificant. The same manner might happen for the other scenarios. In addition, following indexes, proposed in [12], are used for performance evaluations:

$\overline{RMSE}_j = \sup\left(\sum_{k=1}^{N}(x_{k,j} - [\hat{x}_{k,j}])^2/N\right)^{1/2}$,
$E_j = \sum_{k=1}^{N} \text{width}([\hat{x}_{k,j}])/N$,
$O_j = \left(\sum_{k=1}^{N} \mathbf{1}(x_{k,j} \in [\hat{x}_{k,j}])/N\right) 100\%$, $j = 1, ..., n_x$,

where $\overline{RMSE}$ is the root mean squared error upper bound.

Let's consider the three basic scenarios previously defined. Table I shows that using only the contraction step gives no good performance results in terms of $\overline{RMSE}$ and E indexes (Scenario 1). Comparing Scenarios 2 and 3, it is shown that the contraction step brings a poor efficiency to the use of equi-likelihood.

TABLE I
THE THREE BASIC SCENARIOS

	Scenario 1		Scenario 2		Scenario 3	
j	1	2	1	2	1	2
$\overline{RMSE}_j$	11.04	19.22	4.72	8.06	4.70	7.85
$O_j(\%)$	100	100	99.80	99.92	99.80	99.80
E_j	18.86	31.27	6.88	11.92	6.86	11.68
Time(s)	46.93		49.12		81.98	

TABLE II
RLBPF VERSUS EBPF

	RLBPF		EBPF	
	$j = 1$	$j = 2$	$j = 1$	$j = 2$
$\overline{RMSE}_j$	4.67	7.43	5.62	8.90
O_j (%)	99.76	99.97	99.99	99.92
E_j	6.85	11.81	8.72	14.41
Time (s)	67.49		190.06	

Next, the simulation of RLBPF and EBPF applied to system (14) is considered. Since, in [12], the point value measurements y_k are given, in this simulation we also use such an assumption and get $[y_k] = y_k - [v_k]$ where $[v_k]$ is given by (8). The reduction percentage $R = 20\%$ is applied for RLBPF and the particle number $M = 9$ is applied for both methods. Table II shows the better performance of RLBPF versus EBPF in terms of $\overline{RMSE}$, E indexes and the computation time (reduced more than 60%). In addition, EBPF performance is not better than those of Scenarios 2 and 3 in all indexes and computation time. In contrast, the $\overline{RMSE}$ and E indexes provided by RLBPF are better than those of Scenario 2 and 3, while its computation time is a compromise between those of the two Scenarios.

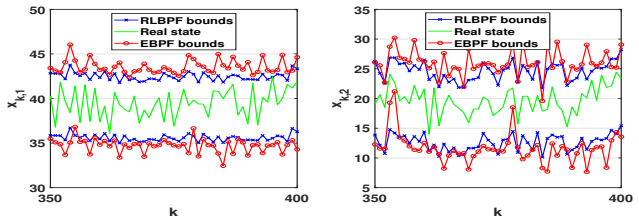


Fig. 2. RLBPF versus EBPF

VII. Conclusion

A general scheme is provided to generalize the specificity of BPFs. The likelihood computation methodology is investigated. This analysis point out the disadvantages of existing filters and opens a way to improve the computed likelihoods by making them more reliable using a reinforcement method. A strategy is proposed to evaluate the performance of this class of filters. The simulation highlights the efficiency of the RLBPF in gain of computation time and evaluation indexes.

References

[1] R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems," *Journal of Basic Engineering*, vol. 82, pp. 35–45, 1960.

[2] S. M. K. Mohamed and S. Nahavandi, "Robust finite-horizon kalman filtering for uncertain discrete-time systems," *IEEE Transactions on Automatic Control*, vol. 57, no. 6, pp. 1548–1552, 2012.

[3] C. Combastel, "Merging kalman filtering and zonotopic state bounding for robust fault detection under noisy environment," *IFAC-PapersOnLine*, vol. 48, no. 21, pp. 289–295, 2015.

[4] Q. H. Lu, S. Fergani, C. Jauberthie, and F. Le Gall, "Optimally bounded interval kalman filter," in *2019 IEEE 58th Conference on Decision and Control (CDC)*, 2019, pp. 379–384.

[5] N. J. Gordon, D. J. Salmond, and A. F. M. Smith, "Novel approach to nonlinear/non-gaussian bayesian state estimation," *IEE Proceedings F - Radar and Signal Processing*, vol. 140, no. 2, pp. 107–113, 1993.

[6] A. Doucet, N. de Freitas, and N. Gordon, Eds., *Sequential Monte Carlo Methods in Practice*. Springer, 2001.

[7] 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.

[8] A. De Freitas, L. Mihaylova, A. Gning, D. Angelova, and V. Kadirkamanathan, "Autonomous crowds tracking with box particle filtering and convolution particle filtering," *Automatica*, vol. 69, pp. 380–394, 2016.

[9] H. Tehrani Niknejad, A. Takeuchi, S. Mita, and D. McAllester, "On-road multivehicle tracking using deformable object model and particle filter with improved likelihood estimation," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 2, pp. 748–758, 2012.

[10] J. Blesa, F. Le Gall, C. Jauberthie, and L. Travé-Massuyès, "State estimation and fault detection using box particle filtering with stochastic measurements," in *26th International Workshop on Principles of Diagnosis (DX-15)*, Paris, France, Aug. 2015, pp. 67–73.

[11] G. Nassreddine, F. Abdallah, and T. Denoux, "State estimation using interval analysis and belief-function theory: Application to dynamic vehicle localization," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 40, no. 5, pp. 1205–1218, 2010.

[12] T. A. Tran, C. Jauberthie, F. Le Gall, and L. Travé-Massuyès, "Evidential box particle filter using belief function theory," *International Journal of Approximate Reasoning*, vol. 93, pp. 40–58, 2018.

[13] L. Jaulin, M. Kieffer, O. Didrit, and E. Walter, *Applied Interval Analysis, with Examples in Parameter and State Estimation, Robust Control and Robotics*. London: Springer-Verlag, 2001.