

# A Box Regularized Particle Filter for state estimation with severely ambiguous and non-linear measurements

Merlinge, NJA, Dahia, K, Piet-Lahanier, H, Brusey, J & Horri, N

Author post-print (accepted) deposited by Coventry University's Repository

## Original citation & hyperlink:

Merlinge, NJA, Dahia, K, Piet-Lahanier, H, Brusey, J & Horri, N 2019, 'A Box Regularized Particle Filter for state estimation with severely ambiguous and non-linear measurements' *Automatica*, vol. 104, pp. 102-110.

<https://dx.doi.org/10.1016/j.automatica.2019.02.033>

DOI 10.1016/j.automatica.2019.02.033

ISSN 0005-1098

Publisher: Elsevier

**NOTICE: this is the author's version of a work that was accepted for publication in *Automatica*. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in**

***Automatica*, [104] (2019) DOI: 10.1016/j.automatica.2019.02.033**

© 2019, Elsevier. Licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International

<http://creativecommons.org/licenses/by-nc-nd/4.0/>

Copyright © and Moral Rights are retained by the author(s) and/ or other copyright owners. A copy can be downloaded for personal non-commercial research or study, without prior permission or charge. This item cannot be reproduced or quoted extensively from without first obtaining permission in writing from the copyright holder(s). The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the copyright holders.

This document is the author's post-print version, incorporating any revisions agreed during the peer-review process. Some differences between the published version and this version may remain and you are advised to consult the published version if you wish to cite from it.

# A Box Regularized Particle Filter for state estimation with severely ambiguous and non-linear measurements

Nicolas Merlinge<sup>\*,\*\*</sup> Karim Dahia<sup>\*</sup> H       Piet-Lahanier<sup>\*</sup>  
James Brusey<sup>\*\*</sup> Nadjim Horri<sup>\*\*</sup>

<sup>\*</sup> ONERA—The French Aerospace Lab  
(e-mail: first name . last name @onera.fr)

<sup>\*\*</sup> Coventry University, United Kingdom  
(e-mail: first name . last name @coventry.ac.uk)

---

**Abstract:** This paper presents a Box Regularized Particle Filter (BRPF). This state estimator aims to be robust to ambiguous (non-injective) and uncertain measurements. Based on previous works on box particle filters, we present a more general and improved formulation of the algorithm, with two innovations: a generalized box resampling step and a kernel smoothing method, which is shown to be optimal in terms of Mean Integrated Square Error. Monte-Carlo simulations demonstrate the efficiency of BRPF on a severely ambiguous and non-linear estimation problem, the Terrain Aided Navigation. BRPF is compared to the Sample Importance Resampling Particle Filter (SIR-PF), the Markov Chain Monte Carlo approach (MCMC), and the original Box Particle Filter (BPF). The algorithm is demonstrated to outperform existing methods in terms of Root Mean Square Error (e.g., improvement up to 42% in geographical position estimation with respect to the BPF) for a large initial uncertainty. The BRPF yields a computational load reduction of 73% with respect to the SIR-PF and of 90% with respect to MCMC for similar RMSE orders of magnitude. The present work offers an accurate (in terms of RMSE) and robust (in terms of divergence rate) way to tackle state estimation from ambiguous measurements while requiring a significantly lower computational load than classic Monte-Carlo and particle filtering methods.

---

Description	Notation
2-Norm	$\ \cdot\ $
Frobenius norm	$\ \cdot\ _F$
Identity matrix	$\mathbf{I}_n \in \mathbb{R}^{n \times n}$
Real set (dimension $\delta$ )	$\mathbb{R}^\delta$
Set, such that (e.g.)	$\{\mathbf{S}\} = \{\mathbf{x} \in \mathbb{R}^\delta \mid \mathbf{x}^T \mathbf{x} > c\}$
Real interval set	$\mathbb{IR}$
Real box set (dimension $\delta$ )	$\mathbb{IR}^\delta = \mathbb{IR} \times \dots \times \mathbb{IR}$
Interval	$[a] \in \mathbb{IR}$
Box	$[\mathbf{a}] \in \mathbb{IR}^\delta$
Density of $\mathbf{x}$ relative to $\mathbf{y}$	$p(\mathbf{x} \mathbf{y})$
Indicator function on box $[\mathbf{a}]$	$\mathbf{1}_{[\mathbf{a}]}(\mathbf{x}) = \begin{cases} 1 & \text{if } \mathbf{x} \in [\mathbf{a}] \\ 0 & \text{else} \end{cases}$
Uniform kernel on box $[\mathbf{a}]$	$\mathcal{U}_{[\mathbf{a}]}(\mathbf{x})$
Normal density	$\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\sigma}^2)$
Probabilistic expectation	$\mathbb{E}[\cdot]$
Probabilistic covariance	$\text{Cov}(\cdot)$
State and measurements	$\mathbf{x}_k \in \mathbb{R}^d$ and $\mathbf{m}_k \in \mathbb{R}^{d_m}$

## 1. INTRODUCTION

A state estimation problem is commonly formalized by: a dynamical state model, which represents the state evolution, and an observation model, which links the state with the measurements. These two models usually include uncertainty to account for unmodeled dynamics, perturbations and measurement errors. Two main frameworks can be used to model uncertainty, namely: probabilistic filters and set-based observers.

In the probabilistic scheme, uncertainty is modeled by a probability density function. The most commonly used probabilistic filter is the Kalman Filter, originally designed for linear problems. It was extended to non-linear problems through the Extended Kalman Filter (EKF) and the Unscented Kalman Filter (UKF). Kalman filters assume that the uncertainty probability distributions are gaussian. In comparison, particle filter methods, which are based on Monte-Carlo methods, handle non-gaussian and strongly non-linear systems. Ristic (2004) provides a review of probabilistic state estimation techniques, from Kalman Filters to particle filters. In order to enhance the accuracy of particle filter estimation, a large variety of methods were proposed. In particular, the kernel smoothing regularization (Regularized Particle Filter, Musso et al. (2001)) was proposed to make the estimated density function fit the target density in an optimal way, based on the kernel theory of Silverman (1986). In what follows, *kernel* refers to a given probability density function. However, probabilistic techniques may diverge when strong measurement ambiguity arises. *Measurement ambiguity* refers to the case when the observation model is non-injective, i.e., when one measurement corresponds to several possible states.

In the set-based framework, uncertainty probability distributions are assumed to be unknown but bounded and are represented by bounded sets, as introduced by Schweppe (1968). Several set representations have been proposed, including: intervals and boxes (Jaulin, 2009),

ellipsoids (Maksarov and Norton, 1996), and polytopes (Piet-Lahanier and Walter, 1994). The output of those algorithms is thus not seen as a point estimate associated with an uncertainty, as with probabilistic filters, but is considered as a bounded set that is guaranteed to contain the real state. However, set-based techniques are often limited to linear or somewhat non-linear systems, especially with regard to the observation model.

In order to produce an estimator able to handle strong non-linearities with minimal complexity, Abdallah et al. (2007) introduced the Box Particle Filter (BPF). The algorithm is similar to a conventional particle filter, except that each particle is a weighted box. Each measurement is also represented by a box. At each measurement time-step, the box particles are contracted, which means that they are replaced by a subset of themselves consistent with the measurements. The interval analysis framework (Jaulin, 2001) makes box manipulations locally simple, while the resulting union of all box particles can approximate a complex state density.

The BPF has the advantage of requiring significantly less computation than the classic Sample Importance Resampling Particle Filter (SIR-PF) for a similar estimation performance. Indeed, the SIR-PF requires a large number of point-wise sampled particles to approach the state density. On the contrary, a  $d$ -dimensional box particle cloud description yields a better covering of the state space. As a result, far fewer particles are needed for the same estimation performance compared with a conventional particle filter. BPF was used for various cases, e.g., for tracking applications (Gning et al., 2012a), extended object tracking (Petrov et al., 2012), and crowd tracking (De Freitas et al., 2016).

However, BPF still suffers some drawbacks. First, modelling *a priori* the boxes as uniform kernels leads to a loss of generality. Although this simple probabilistic assumption is of practical interest in the case of unknown uncertainty probability density functions, it excludes the case when other assumptions could be available on those densities. The present paper proposes a more general theoretical description of BPF. In addition, BPF tends to be quite conservative for severely ambiguous problems. Indeed, the filter often yields excessively large confidence bounds around the estimate. An example of ambiguous measurements is Terrain Aided Navigation (TAN), for which BPF fails to accurately converge (Merlinge et al., 2016). This is mostly due to the resampling step, which aims to avoid the *degeneracy phenomenon*. The resampling is triggered when too many particles have a weight that is close to zero. They are considered inconsistent with respect to the measurements and are replaced by subdivisions of the more highly weighted particles. The subdivision is performed along one dimension of the state space and consists of a repaving of the particle into a number of new particles proportional to the original particle weight. The total number of particles remains unchanged. Although many problem-dependent subdivision approaches exist, there is no general method. In addition, the box resampling leads to an increasing number of box overlaps, which results in a lack of accuracy in terms of density approximation.

This paper presents two main contributions leading to a more general and accurate formulation of BPF (Sections 2 and 3) called the Box Regularized Particle Filter (BRPF) (Section 4):

- (1) A new box particle subdivision approach for resampling, with an analytic solution based on an observability-linked criterion (Section 4.1). A significant accuracy enhancement compared to the conventional BPF is demonstrated by numerical simulations, in terms of Root Mean Square Error.
- (2) The optimal smoothing of the posterior conditional density estimation using bounded kernel regularization expressed in terms of box parameters (Section 4.2). The optimal smoothing kernel is determined analytically by minimization of the Mean Integrated Square Error (MISE) criterion. The resulting performance enhancement compared to BPF is demonstrated by simulation.

Numerical simulation results are presented in Section 5.

## 2. PROBLEM STATEMENT

The system evolution is modeled by a discrete dynamical model:

$$\mathbf{x}_k = f(\mathbf{x}_{k-1}, \mathbf{u}_k) + \mathbf{w}_k \quad (1)$$

with  $\mathbf{w}_k \in \mathbb{R}^d$  the process noise and  $\mathbf{u}_k \in \mathbb{R}^{d_c}$  a control input. For the sake of brevity, the control input will be omitted. The observation equation is as follows:

$$\mathbf{m}_k = h(\mathbf{x}_k) + \mathbf{v}_k \quad (2)$$

where  $h$  is the observation model (potentially non-injective), and  $\mathbf{v}_k \in \mathbb{R}^{d_m}$  is the measurement noise. Process and measurement uncertainties can be described using two different frameworks. The first one consists in associating a probabilistic density function (pdf) to  $\mathbf{v}_k$  and  $\mathbf{w}_k$ . However when the selection of a suitable structure for the pdf proves difficult for highly non-linear models, an alternative consists in representing these uncertainties using the unknown but bounded representation which relies on providing only a description of the bounds within which they vary. These different descriptions lead to different estimation methods that are briefly recalled hereafter.

### 2.1 Probabilistic observers framework

In the probabilistic scheme, the process noise and the measurement noise are associated with some probability density function: the transition density  $p(\mathbf{x}_k|\mathbf{x}_{k-1})$  and the likelihood  $p(\mathbf{m}_k|\mathbf{x}_k)$ , with  $\mathbf{M}_k = \{\mathbf{m}_1, \dots, \mathbf{m}_k\}$ .

The prediction step corresponds to the dynamical propagation of the conditional density. The conditional density is obtained by a convolution of the prior conditional density  $p(\mathbf{x}_{k-1}|\mathbf{M}_{k-1})$  with the state transition density  $p(\mathbf{x}_k|\mathbf{x}_{k-1})$ , through the Chapman-Kolmogorov equation:

$$p(\mathbf{x}_k|\mathbf{M}_{k-1}) = \int p(\mathbf{x}_k|\mathbf{x}_{k-1})p(\mathbf{x}_{k-1}|\mathbf{M}_{k-1})d\mathbf{x}_{k-1} \quad (3)$$

The correction step corresponds to the predicted conditional density  $p(\mathbf{x}_k|\mathbf{M}_{k-1})$  updated by the measurements density  $p(\mathbf{m}_k|\mathbf{x}_k)$  and is obtained by the Bayes rule, under the assumption of statistically independent measurements:

$$p(\mathbf{x}_k | \mathbf{M}_k) = \frac{p(\mathbf{m}_k | \mathbf{x}_k) p(\mathbf{x}_k | \mathbf{M}_{k-1})}{\int p(\mathbf{m}_k | \mathbf{x}_k) p(\mathbf{x}_k | \mathbf{M}_{k-1}) d\mathbf{x}_k} \quad (4)$$

## 2.2 Set-based observers framework

In the set-based observer scheme, the process and measurement noises are modeled by bounded sets  $\{\mathbf{w}_k\}$  and  $\{\mathbf{v}_k\}$ . Therefore, let the measurement set be defined as  $\{\mathbf{m}_k\} = \{\mathbf{y} \in \mathbb{R}^{d_m} | \mathbf{y} - \mathbf{m}_k \in \{\mathbf{v}_k\}\}$ . Let  $\{\mathbf{x}_{k-1}\}$  be the previous state set estimation and  $\{\mathbf{x}_k\}$  the current state set estimation.

The prediction step corresponds to the propagation of all elements of the previous estimated set, plus the process uncertainty set:

$$\{\mathbf{x}_{k|k-1}\} = \{\mathbf{x} \in \mathbb{R}^d \mid \mathbf{x} - f(\mathbf{x}') \in \{\mathbf{w}_k\} \forall \mathbf{x}' \in \{\mathbf{x}_{k-1}\}\} \quad (5)$$

The correction step corresponds to the measurement update obtained by intersecting the predicted set:

$$\{\mathbf{x}_k\} = \{\mathbf{x} \in \{\mathbf{x}_{k|k-1}\} \mid h(\mathbf{x}) \in \{\mathbf{m}_k\}\} \quad (6)$$

Several methods have been introduced to describe these sets and their evolution. In the present work, the algorithm developed relies on the use of specific set descriptions that are boxes and intervals. The interval framework is described hereafter.

## 2.3 Interval analysis

This paragraph briefly recalls the interval analysis formalism described in Jaulin (2001). An interval  $[a]$  of  $\mathbb{R}$  is defined as  $[a] = [\underline{a}, \bar{a}] = \{x \in \mathbb{R}, \underline{a} \leq x \leq \bar{a}\} \in \mathbb{IR}$ , where  $\mathbb{IR}$  denotes the real interval space. Equivalently, it can be described by its center and its diameter as  $[a] = (c, \delta) = \{x \in \mathbb{R}, c - \delta/2 \leq x \leq c + \delta/2\} \in \mathbb{IR}$ .

A box is noted  $[\mathbf{a}] = [\underline{\mathbf{a}}, \bar{\mathbf{a}}] \in \mathbb{IR}^d$  with  $\underline{\mathbf{a}} \in \mathbb{R}^d$  and  $\bar{\mathbf{a}} \in \mathbb{R}^d$  the lower and upper bounds of  $[\mathbf{a}]$ . It describes a hyperrectangle and can be written as a vector of intervals:  $[\mathbf{a}] = [a_1] \times [a_2] \times \dots \times [a_d] \in \mathbb{IR}^d$ . Equivalently, it can be described by its center and its diameter  $[\mathbf{a}] = (\mathbf{c}_{[\mathbf{a}]}, \boldsymbol{\delta}_{[\mathbf{a}]})$ . Table 1 describes the most usual interval analysis operations. One can notice that the volume of

Table 1. Interval Analysis operations (Jaulin, 2001)

Description	Notation
Operation $\odot$ (+, -, *, /)	$[a] \odot [b] = [\{x \odot y \mid x \in [a], y \in [b]\}]$
Intersection	$[a] \cap [b] = [\max(\underline{a}, \underline{b}), \min(\bar{a}, \bar{b})]$
Diameter	$\delta_{[a]} =  [a]  = \bar{a} - \underline{a} \quad (\in \mathbb{R})$ By convention, $ \emptyset  = 0$
Volume of a box	$ [\mathbf{a}]  = \prod_{j=1}^d  [a_j]  \quad (\in \mathbb{R})$
Diameter of a box	$\boldsymbol{\delta}_{[\mathbf{a}]} = [ a_1 , \dots,  a_d ]^T \quad (\in \mathbb{R}^d)$
Center of a box	$\mathbf{c}_{[\mathbf{a}]} = \frac{1}{2}(\underline{\mathbf{a}} + \bar{\mathbf{a}}) \quad (\in \mathbb{R}^d)$

a box corresponds to the Lebesgue measure. Functions of  $\mathbb{R}^{\delta_1} \rightarrow \mathbb{R}^{\delta_2}$  ( $\delta_1, \delta_2 \in \mathbb{N}^*$ ) have also to be adapted to interval framework. This leads us to define the inclusion function concept. An inclusion function of a function  $\psi$  from  $\mathbb{R}^{\delta_1}$  to  $\mathbb{R}^{\delta_2}$  is defined as  $[\psi]$  from  $\mathbb{IR}^{\delta_1}$  to  $\mathbb{IR}^{\delta_2}$ . The output of  $[\psi]$  is a box  $[\psi]([\mathbf{a}]) \in \mathbb{IR}^{\delta_2}$  that contains the

output set  $\{\psi([\mathbf{a}])\} = \{\mathbf{y} \in \mathbb{R}^{\delta_2} \mid \mathbf{y} = \psi(\mathbf{x}) \forall \mathbf{x} \in [\mathbf{a}]\}$  of box  $[\mathbf{a}]$  by  $\psi$ . An inclusion function is minimal if its image is the smallest box containing  $\{\psi([\mathbf{a}])\}$ . In what follows, in order to limit the pessimism of box-based outer-approximations of output sets, inclusion functions are chosen minimal (Jaulin, 2001). Thus, (1) and (2) can be rewritten as follows:

$$\begin{cases} [\mathbf{x}_k] &= [f]([\mathbf{x}_{k-1}]) + [\mathbf{w}_k] \\ [\mathbf{m}_k] &= [h]([\mathbf{x}_k]) + [\mathbf{v}_k] \end{cases} \quad (7)$$

## 3. BOX PARTICLE FILTER (BPF)

The Box Particle Filter was initially proposed in (Abdallah et al., 2007) as a first bridge between Monte-Carlo methods and set-based approaches.

This section details the general BPF description by deriving the Optimal Filter's equations (3) and (4). A generic formulation of BPF is introduced. The prior conditional state density at time  $k-1$  is defined by a mixture of  $N$  kernels bounded by box particles  $[\mathbf{x}_{k-1}^i] \in \mathbb{IR}^d$  and weighted by weights  $w_k^i$  whose sum is unity:

$$p(\mathbf{x}_{k-1} | \mathbf{M}_{k-1}) = \sum_{i=1}^N w_{k-1}^i \pi_{k-1}^i(\mathbf{x}_{k-1}) \mathbf{1}_{[\mathbf{x}_{k-1}^i]}(\mathbf{x}_{k-1}) \quad (8)$$

where each box kernel  $\pi_{k-1}^i : \mathbb{R}^d \rightarrow \mathbb{R}$  refers to a known probability density function and satisfies  $\int_{[\mathbf{x}_{k-1}^i]} \pi_{k-1}^i(\mathbf{x}) d\mathbf{x} = 1$ . The assumption is made that the dynamical model  $f : \mathbb{R}^d \rightarrow \mathbb{R}^d$  and the observation model  $h : \mathbb{R}^d \rightarrow \mathbb{R}^{d_m}$  are continuous on their domain. Initial box kernels  $\pi_0^i : \mathbb{R}^d \rightarrow \mathbb{R}$  are also assumed to be continuous. Likewise, the measurement density  $\pi_k^m : \mathbb{R}^d \rightarrow \mathbb{R}^{d_m}$  is assumed to be continuous at each time-step. For the sake of brevity, kernel arguments may be omitted in what follows. A solution is provided for the specific case of uniform kernels (Abdallah et al., 2007; Gning et al., 2013), which is computationally efficient.

### 3.1 Prediction step

Applying the Chapman-Kolmogorov equation (3) to the prior density (8) yields:

$$\begin{aligned} p(\mathbf{x}_k | \mathbf{M}_{k-1}) &= \int_{\mathbb{R}^d} \pi_{k|k-1}^x \sum_{i=1}^N w_{k-1}^i \pi_{k-1}^i \mathbf{1}_{[\mathbf{x}_{k-1}^i]} d\mathbf{x}_{k-1} \\ &= \sum_{i=1}^N \left( w_{k-1}^i \int_{[\mathbf{x}_{k-1}^i]} \pi_{k|k-1}^x \pi_{k-1}^i \mathbf{1}_{[\mathbf{x}_{k-1}^i]} d\mathbf{x}_{k-1} \right) \end{aligned} \quad (9)$$

Define the  $i^{th}$  predicted kernel, whose support is included in box particle  $[\mathbf{x}_{k|k-1}^i]$ :

$$\pi_{k|k-1}^i(\mathbf{x}_k) \triangleq \int_{[\mathbf{x}_{k-1}^i]} \pi_{k|k-1}^x(\mathbf{x}_k - f(\mathbf{x}_{k-1})) \pi_{k-1}^i(\mathbf{x}_{k-1}) d\mathbf{x}_{k-1} \quad (10)$$

The predicted conditional density can be written as:

$$p(\mathbf{x}_k | \mathbf{M}_{k-1}) = \sum_{i=1}^N w_{k-1}^i \pi_{k|k-1}^i \mathbf{1}_{[\mathbf{x}_{k|k-1}^i]} \quad (11)$$

with

$$[\mathbf{x}_{k|k-1}^i] \triangleq [f]([\mathbf{x}_{k-1}^i]) + [\mathbf{w}_k] \quad (12)$$

As stated in Gning et al. (2013), for uniform kernels, the assumption can be made that:

$$\pi_{k|k-1}^i = \int_{\mathbb{R}^d} \mathcal{U}_{[\mathbf{x}_{k-1}^i]} p(\mathbf{x}_k | \mathbf{x}_{k-1}) d\mathbf{x}_{k-1} \approx \mathcal{U}_{[f]([\mathbf{x}_{k-1}^i]) + [\mathbf{w}_k]} \quad (13)$$

### 3.2 Correction step

The correction step determines the posterior conditional distribution of the state with respect to the predictive distribution (11) and the measurement density. The measurement noise density  $p(\mathbf{m}_k | \mathbf{x}_k) = \pi_k^m$  is assumed to be bounded by the box  $[\mathbf{m}_k]$ . Bayes' rule (4) leads to:

$$\begin{aligned} p(\mathbf{x}_k | \mathbf{M}_k) &= \frac{1}{q_k} \left( \sum_{i=1}^N w_{k-1}^i \pi_{k|k-1}^i \mathbf{1}_{[\mathbf{x}_{k|k-1}^i]} \right) \pi_k^m \\ &= \frac{1}{q_k} \sum_{i=1}^N w_{k-1}^i \left( \pi_{k|k-1}^i \mathbf{1}_{[\mathbf{x}_{k|k-1}^i]} \pi_k^m \mathbf{1}_{[\mathbf{m}_k]} \right) \end{aligned} \quad (14)$$

Therefore, a new box particle can be defined as the consequence of the product of the two indicator functions:

$$[\mathbf{x}_k^i] = \left[ \left\{ \mathbf{x} \in \mathbb{R}^d \mid \mathbf{1}_{[\mathbf{x}_{k|k-1}^i]}(\mathbf{x}) \mathbf{1}_{[\mathbf{m}_k]}(\mathbf{m}_k - h(\mathbf{x})) \neq 0 \right\} \right] \quad (15)$$

which is equivalent to a Constraints Satisfaction Problem defined by  $\left[ \left\{ \mathbf{x}_k \in [\mathbf{x}_{k|k-1}^i] \mid h(\mathbf{x}_k) \in [\mathbf{m}_k] \right\} \right]$  (Jaulin, 2009). Therefore, new kernel supports  $[\mathbf{x}_k^i]$  can be computed by interval contraction:

$$p(\mathbf{x}_k | \mathbf{M}_k) = \frac{1}{q_k} \sum_{i=1}^N w_{k-1}^i \left( \pi_k^m \pi_{k|k-1}^i \mathbf{1}_{[\mathbf{x}_k^i]} \right) \quad (16)$$

However, the term  $\pi_k^m \pi_{k|k-1}^i \mathbf{1}_{[\mathbf{x}_k^i]}$  no longer sums to unity and is therefore not a pdf. Indeed, the support of the kernel  $\pi_{k|k-1}^i \mathbf{1}_{[\mathbf{x}_{k|k-1}^i]}$  has been truncated by  $[\mathbf{x}_k^i] \subset [\mathbf{x}_{k|k-1}^i]$  whose volume is lower or equal to that of  $[\mathbf{x}_{k|k-1}^i]$ . Furthermore, it is multiplied by the measurement kernel  $\pi_k^m$ , which leads to a new kernel proportional to  $\pi_k^m \pi_{k|k-1}^i$ . Therefore, it has to be normalised by  $\int_{[\mathbf{x}_k^i]} \pi_k^m \pi_{k|k-1}^i d\mathbf{x}$ , which yields,

$$p(\mathbf{x}_k | \mathbf{M}_k) = \frac{1}{q_k} \sum_{i=1}^N \left( w_{k-1}^i \int_{[\mathbf{x}_k^i]} \pi_k^m \pi_{k|k-1}^i d\mathbf{x} \right) \pi_k^i \mathbf{1}_{[\mathbf{x}_k^i]} \quad (17)$$

where  $\pi_k^i \triangleq \frac{1}{\int_{[\mathbf{x}_k^i]} \pi_k^m \pi_{k|k-1}^i d\mathbf{x}}$  is the updated kernel. This kernel is only defined if  $\int_{[\mathbf{x}_k^i]} \pi_k^m \pi_{k|k-1}^i d\mathbf{x} \neq 0$ , i.e. if box  $i$  is consistent with the measurement density. Else, it can be set by convention to  $w_k^i = 0$ .

Finally, the posterior conditional density is obtained:

$$p(\mathbf{x}_k | \mathbf{M}_k) = \sum_{i=1}^N w_k^i \pi_k^i \mathbf{1}_{[\mathbf{x}_k^i]} \quad (18)$$

where the update of weights is, for  $k \geq 1$ :

$$w_k^i \triangleq \frac{1}{q_k} w_{k-1}^i \int_{[\mathbf{x}_k^i]} \pi_k^m (\mathbf{m}_k - h(\mathbf{x})) \pi_{k|k-1}^i(\mathbf{x}) d\mathbf{x} \quad (19)$$

The integral term in (19) can be interpreted as a consistency term between the box kernel  $\pi_{k|k-1}^i(\mathbf{x})$  whose support is restricted to box  $[\mathbf{x}_{k|k-1}^i]$  and the measurement kernel

$\pi_k^m(\mathbf{m}_k - h(\mathbf{x}))$ . This integral term belongs to interval  $[0, 1]$ . The normalisation quotient is:

$$q_k = \sum_i w_{k-1}^i \int_{[\mathbf{x}_k^i]} \pi_k^m \pi_{k|k-1}^i d\mathbf{x} \quad (20)$$

In the case of uniform kernels and a uniform measurement noise (Gning et al., 2013), the updated weights are obtained by:

$$w_k^i = \frac{1}{q_k} \frac{|[\mathbf{x}_k^i]|}{|[\mathbf{x}_{k|k-1}^i]|} w_{k-1}^i \quad (21)$$

### 3.3 Resampling step

As in a conventional particle filter, a resampling step is added to avoid the degeneracy phenomenon, when only a small number of box particles are consistent with the measurements and all others have a near-zero weight. When the resampling step is triggered, each particle is duplicated in  $n^i \in [0, N]$  instances. A zero  $n^i$  yields the particle deletion. The total number of particles usually remains the same after resampling, i.e.  $\sum_i n^i = N$ . The most common method used to determine  $n^i$  is the multinomial resampling. The duplication number per particle  $n^i$  relies on the particle weight, such that high weighted particles are more likely to be duplicated than low weighted particles. A survey on resampling techniques for particles filters can be found in Li et al. (2015). The resampling step is not processed at every time-step and is triggered by a threshold based on the particle weights. The most commonly used is the  $N_{effective}$  criterion:

$$\frac{1}{\sum_{i=1}^N w_k^i{}^2} < \theta_{eff} N \quad (22)$$

with  $\theta_{eff} \in [0, 1]$  a normalized tuning threshold. This criterion is designed to reflect the case where the number of particles with a low weight value exceed a given threshold.

The resampling step can be derived for the BPF scheme. Instead of duplicating the box particles, the BPF subdivides them along one dimension  $d_k^{cut,i} \in [1, d]$  to increase the resolution of the state density estimation. Several strategies were proposed to choose the  $d_k^{cut,i}$  dimension along which to subdivide the current box particle. However, to the best of the author's knowledge, they are either limited to a specific state representation, e.g. Abdallah et al. (2008); Merlinge et al. (2016); De Freitas et al. (2016), fully observable states (Luo and Qin, 2018) or are not deterministic (Gning et al., 2012b).

### 3.4 State estimation

A state estimate  $\hat{\mathbf{x}}_k$  can be derived from the box particles cloud such that:

$$\hat{\mathbf{x}}_k \triangleq E[\mathbf{x}_k \sim p(\mathbf{x}_k | \mathbf{M}_k)] \approx \sum_{i=1}^N w_{k-1}^i \mathbf{c}_k^i \quad (23)$$

For the sake of brevity, the box particles centers are denoted  $\mathbf{c}_k^i \triangleq \mathbf{c}_{[\mathbf{x}_k^i]}$ . An empirical covariance matrix can be defined by:

$$\hat{\mathbf{P}}_k = Cov(\mathbf{x}_k \sim p(\mathbf{x}_k | \mathbf{M}_k)) \approx \sum_{i=1}^N w_k^i (\mathbf{c}_k^i - \hat{\mathbf{x}}_k)(\mathbf{c}_k^i - \hat{\mathbf{x}}_k)^T \quad (24)$$

#### 4. BOX REGULARIZED PARTICLE FILTER (BRPF)

In the previous section, a general probabilistic formulation of BPF was provided. However, as mentioned in 3.3, the box-resampling dimension subdivision does not have a generic definition in the literature and is often problem dependent. In this section, a generic box subdivision technique is described, based on a variance evaluation.

Moreover, the resulting box particle cloud does not approximate the actual posterior conditional density in an optimal way with respect to the Mean Integrated Square Error (MISE) metric. Therefore, a kernel smoothing step is added after each resampling step to obtain an optimal box particles breakdown with respect to the MISE. We call the derived algorithm the Box Regularized Particle Filter.

##### 4.1 Box particles subdivision for resampling

The BPF box-resampling method aims to enhance the resolution of the state space exploration in its most likely regions by replacing low-weighted boxes by subdivisions of high-weighted ones. For each particle, the resampling algorithm provides an integer value indicating how many subdivisions will be performed on it. In this section, we consider a box  $i$  which has to be subdivided in  $n^i$  sub-boxes. A common approach is subdividing each box particle  $i$  along a single state dimension  $d_k^{cut,i} \in [1, d]$ , where  $d$  is the number of state variables. Various methods were proposed in the literature. However, they are often too application-specific, not deterministic, or limited to fully observed states.

The objective of this section is providing a generic and deterministic formulation for the subdivision dimension  $d_k^{cut,i}$  for each particle. Our approach relies on an edge normalization by a lowest expected variance in the sense of the maximum likelihood estimator.

*Proposition 1.* During the box subdivision resampling, each box can be subdivided along the edge that is the most pessimistic compared to a lowest expected variance. Provided that:

- The measurement noise (associated with the likelihood density) has a covariance  $\mathbf{R}_k \in \mathbb{R}^{d_m \times d_m}$ ,
- The observation model  $h$  is differentiable

The subdivision dimension  $d^{cut}$  is chosen by picking the largest coefficient of a normalised box particle's diameter  $\tilde{\delta}_k^i = [\tilde{\delta}_k^{i,1}, \dots, \tilde{\delta}_k^{i,d}]^T$ :

$$d_k^{cut,i} = \underset{j \in [1, d]}{\operatorname{argmax}} (\tilde{\delta}_k^{i,j}) \quad (25)$$

The normalised diameter is computed from the box particle's prior diameter  $\delta_k^i \triangleq \delta_{[\mathbf{x}_k^i]} \in \mathbb{R}^d$  and the inverse of the square root (e.g., Cholesky decomposition) of the lowest theoretical covariance  $\Sigma_k$ :

$$\tilde{\delta}_k^i \triangleq \left( \sqrt{\Sigma_k} \right)^{-1} \delta_k^i \in \mathbb{R}^d \quad (26)$$

where  $\Sigma_k \in \mathbb{R}^{d \times d}$  can be evaluated by:

$$\Sigma_k \triangleq \hat{\mathbf{P}}_k + \frac{1}{2} \sum_{i=1}^N w_k^i (\Sigma_k^i + \mathbf{V}_k^i) \quad (27)$$

where

$$\Sigma_k^i = \operatorname{Vec}^{-1} \left( (\lambda \mathbf{T}_k^i \otimes \mathbf{T}_k^i + (1 - \lambda) \mathbf{I}_{d^2})^{-1} \operatorname{Vec} \left( \lambda \mathbf{H}_k^i{}^T \mathbf{R}_k \mathbf{H}_k^i + (1 - \lambda) \Delta_k^i \right) \right) \quad (28)$$

where  $\mathbf{V}_k^i$  is obtained from the polar decomposition of  $\Sigma_k^i$ ,  $\mathbf{T}_k^i = \left( \frac{\partial h}{\partial \mathbf{x}} \right)^T \left( \frac{\partial h}{\partial \mathbf{x}} \right)$ ,  $\operatorname{Vec}$  is the vectorization operator, and  $\Delta_k^i$  is a diagonal matrix whose diagonal terms are the squared box's diameters  $\delta_k^{i,j^2}$  ( $j \in [1, d]$ ). The box can be subdivided in  $n^i$  new boxes along the dimension that maximises the normalised diameter:

$$d^{cut} = \underset{j \in [1, d]}{\operatorname{argmax}} (\tilde{\delta}_j) \quad (29)$$

*Proof* In the following, the analytic expression of  $\Sigma_k$  is derived, leading to equation (27).

The measurement density  $p(\mathbf{m}_k | \mathbf{x}_k) = \pi_k^m(\mathbf{m}_k - h(\mathbf{x}_k))$  is assumed to have a single maximum  $\hat{\mathbf{x}}_k^i$  inside each subset  $[\mathbf{x}_k^i]$  which satisfies:

$$\hat{\mathbf{x}}_k^i = \underset{\mathbf{x}_k}{\operatorname{argmax}} (p(\mathbf{m}_k | \mathbf{x}_k)) \quad (30)$$

Thus, one can link  $h(\hat{\mathbf{x}}_k^i)$  and  $\mathbf{m}_k$  as  $h(\hat{\mathbf{x}}_k^i) = \mathbf{m}_k$ . The measurement variance  $\operatorname{Cov}[\mathbf{m}_k] = \mathbf{R}_k \in \mathbb{R}^{d_m \times d_m}$  is assumed to be known. As a result, the maximum likelihood satisfies, for all  $i$ :

$$\operatorname{Cov} [h(\hat{\mathbf{x}}_k^i)] = \mathbf{R}_k \quad (31)$$

where the measurement covariance  $\mathbf{R}_k$  does not depend on  $i$ . On the other hand, the observation function  $h$  can be locally linearised to first order as follows:

$$h(\hat{\mathbf{x}}_k^i) = h(\mathbf{x}_k^i) + \mathbf{H}_k^i (\hat{\mathbf{x}}_k^i - \mathbf{x}_k^i) + o(\hat{\mathbf{x}}_k^i - \mathbf{x}_k^i) \quad (32)$$

with  $\mathbf{H}_k^i \triangleq \left. \frac{\partial h}{\partial \mathbf{x}} \right|_{\mathbf{x}=\mathbf{x}_k^i} \in \mathbb{R}^{d_m \times d}$  and  $\mathbf{x}_k^i = E[\pi_k^i]$ .

This implies that:

$$h(\hat{\mathbf{x}}_k^i) - h(\mathbf{x}_k^i) \approx \mathbf{H}_k^i (\hat{\mathbf{x}}_k^i - \mathbf{x}_k^i) \quad (33)$$

and that

$$\begin{aligned} & \left( h(\hat{\mathbf{x}}_k^i) - h(\mathbf{x}_k^i) \right) \left( h(\hat{\mathbf{x}}_k^i) - h(\mathbf{x}_k^i) \right)^T \\ & \approx \mathbf{H}_k^i (\hat{\mathbf{x}}_k^i - \mathbf{x}_k^i) (\hat{\mathbf{x}}_k^i - \mathbf{x}_k^i)^T \mathbf{H}_k^i{}^T \end{aligned} \quad (34)$$

Thus, by taking the expectancy of both hand sides:

$$\begin{aligned} & E \left[ \left( h(\hat{\mathbf{x}}_k^i) - h(\mathbf{x}_k^i) \right) \left( h(\hat{\mathbf{x}}_k^i) - h(\mathbf{x}_k^i) \right)^T \right] \\ & \approx E \left[ \mathbf{H}_k^i (\hat{\mathbf{x}}_k^i - \mathbf{x}_k^i) (\hat{\mathbf{x}}_k^i - \mathbf{x}_k^i)^T \mathbf{H}_k^i{}^T \right] \\ & = \mathbf{H}_k^i E \left[ (\hat{\mathbf{x}}_k^i - \mathbf{x}_k^i) (\hat{\mathbf{x}}_k^i - \mathbf{x}_k^i)^T \right] \mathbf{H}_k^i{}^T = \mathbf{H}_k^i \operatorname{Cov}[\hat{\mathbf{x}}_k^i] \mathbf{H}_k^i{}^T \end{aligned} \quad (35)$$

Thus, the local maximum likelihood must satisfy:

$$\operatorname{Cov} [h(\hat{\mathbf{x}}_k^i)] \approx \mathbf{H}_k^i \Sigma_k^i \mathbf{H}_k^i{}^T \quad (36)$$

with  $\Sigma_k^i \triangleq \operatorname{Cov}[\hat{\mathbf{x}}_k^i]$ . Therefore, by combining (31) and (36), one can write:

$$\mathbf{H}_k^i \Sigma_k^i \mathbf{H}_k^i{}^T \approx \mathbf{R}_k \quad (37)$$

In practice, since the actual state  $\mathbf{x}_k$  is unknown,  $\mathbf{H}_k^i$  can be evaluated from  $\mathbf{H}_k^i \approx \left. \frac{\partial h}{\partial \mathbf{x}} \right|_{\mathbf{x}=\mathbf{c}_k^i}$  with  $\mathbf{c}_k^i$  the center of the  $i^{th}$  box particle. Equation (37) imposes a constraint on  $\Sigma_k^i$  which depends on the observation equation.

However, if the rank of  $\mathbf{R}_k$  is less than the rank of  $\Sigma_k^i$ , some additional information needs to be added. Indeed, (37)

only affects the coefficients of  $\Sigma_k^i$  that explicitly depend on the measurement in the observation function  $h$ . In order to calculate the other coefficients, which are linked together through the dynamical model  $f()$ , a possible solution is to introduce a dynamical information with  $\Delta_k^i \triangleq \text{Diag}(\delta_k^{i2})$ . Function  $\text{Diag}()$  transforms a vector of  $\mathbb{R}^n$  into a diagonal matrix of  $\mathbb{R}^{n \times n}$ . Such techniques are known as *regularisation* techniques (Neumaier, 1998). Therefore, the problem to solve is a trade-off between the observation constraint and the state constraint:

$$\begin{cases} \Sigma_k^i = \text{argmin}(J(\Sigma)) \\ J(\Sigma) = \lambda \left\| \mathbf{H}_k^i \Sigma \mathbf{H}_k^{iT} - \mathbf{R}_k \right\|_F^2 + (1 - \lambda) \left\| \Sigma - \Delta_k^i \right\|_F^2 \\ \Sigma > 0 \end{cases} \quad (38)$$

where  $\lambda \in (0, 1)$  is a tuning coefficient. Equation (38) can be solved by computing the derivative of  $J$ :

$$\begin{aligned} \frac{\partial J}{\partial \Sigma} &= \frac{\partial}{\partial \Sigma} \left[ \lambda \text{tr} \left( (\mathbf{H}_k^i \Sigma \mathbf{H}_k^{iT} - \mathbf{R}_k)^T (\mathbf{H}_k^i \Sigma \mathbf{H}_k^{iT} - \mathbf{R}_k) \right) \right. \\ &\quad \left. + (1 - \lambda) \text{tr} \left( (\Sigma - \Delta_k^i)^T (\Sigma - \Delta_k^i) \right) \right] \\ &= 2\lambda \frac{\partial}{\partial \Sigma} \text{tr} \left( (\mathbf{H}_k^i \Sigma \mathbf{H}_k^{iT})^T (\mathbf{H}_k^i \Sigma \mathbf{H}_k^{iT}) \right) \\ &\quad - 2\lambda \frac{\partial}{\partial \Sigma} \text{tr} \left( \mathbf{H}_k^i \Sigma \mathbf{H}_k^{iT} \mathbf{R}_k \right) \\ &\quad + (1 - \lambda) \frac{\partial}{\partial \Sigma} \text{tr}(\Sigma^2) - 2(1 - \lambda) \frac{\partial}{\partial \Sigma} \text{tr}(\Sigma \Delta_k^i) \\ &= 2\lambda \left( \mathbf{H}_k^i{}^T \mathbf{H}_k^i \Sigma \mathbf{H}_k^i{}^T \mathbf{H}_k^i \right) - 2\lambda \left( \mathbf{H}_k^i{}^T \mathbf{R}_k \mathbf{H}_k^i \right) \\ &\quad + 2(1 - \lambda) \Sigma - 2(1 - \lambda) \Delta_k^i \end{aligned} \quad (39)$$

Finally, the equation  $\frac{\partial J}{\partial \Sigma} = 0$  is equivalent to:

$$\lambda \mathbf{T}_k^i \Sigma \mathbf{T}_k^i + (1 - \lambda) \Sigma = \Omega_k^i \quad (40)$$

with  $\mathbf{T}_k^i = \mathbf{H}_k^i{}^T \mathbf{H}_k^i$  and  $\Omega_k^i \triangleq \lambda \mathbf{H}_k^i{}^T \mathbf{R}_k \mathbf{H}_k^i + (1 - \lambda) \Delta_k^i$ . Using the Kronecker product properties, denoted  $\otimes$ , derive (40) as follows:

$$(\lambda \mathbf{T}_k^i \otimes \mathbf{T}_k^i + (1 - \lambda) \mathbf{I}_{d^2}) \text{Vec}(\Sigma) = \text{Vec}(\Omega_k^i) \quad (41)$$

where  $\text{Vec}() : \mathbb{R}^{d \times d} \rightarrow \mathbb{R}^{d^2}$  stands for the column-wise concatenation of a matrix.

As a result, the solution  $\Sigma_k^i$  can be obtained by:

$$\Sigma_k^i = \text{Vec}^{-1} \left( (\lambda \mathbf{T}_k^i \otimes \mathbf{T}_k^i + (1 - \lambda) \mathbf{I}_{d^2})^{-1} \text{Vec}(\Omega_k^i) \right) \quad (42)$$

where  $\text{Vec}^{-1}() : \mathbb{R}^{d^2} \rightarrow \mathbb{R}^{d \times d}$  gives a  $d \times d$  matrix representation of a  $d^2$  vector whose elements are taken column-wise.

However, the resulting  $\Sigma_k^i$  matrix might not be positive definite. It can then be approximated by the nearest positive definite matrix, in terms of Frobenius norm, by  $\bar{\Sigma}_k^i = \frac{1}{2} (\Sigma_k^i + \mathbf{V}_k^i)$ , where  $\mathbf{V}_k^i \in \mathbb{R}^{d \times d}$  is obtained from the polar decomposition of  $\Sigma_k^i$ , i.e.  $\Sigma_k^i = \mathbf{U}_k^i \mathbf{V}_k^i$  with  $\mathbf{U}_k^i{}^T \mathbf{U}_k^i = \mathbf{I}_d$ . This theorem is developed in the work of Higham (1988).

The Maximum Likelihood covariance can therefore be approached by:

$$\Sigma_k \triangleq \hat{\mathbf{P}}_k + \sum_{i=1}^N w_k^i \bar{\Sigma}_k^i \quad (43)$$

Each box diameter is normalised by the square root of  $\Sigma_k$  using (26) and the choice of  $d^{cut}$  is done using (29).

#### 4.2 Bounded Kernel Smoothing by Regularization

In the previous section, a deterministic way to subdivide box particles was described. However, whatever be the subdivision method, this operation ends with a high correlation between the box particles parameters (centers and diameters). As a result, several particles may exactly overlap and the posterior density  $p(\mathbf{x}_k | \mathbf{M}_k)$  is often poorly approximated, which leads to a biased state estimation. A possible solution to enhance the posterior density approximation is smoothing the box particles parameters distributions by adding a stochastic bounded noise. This process, called kernel regularization, was presented in (Musso et al., 2001) to improve the accuracy of conventional particles filters. It is based on the theory of Silverman (1986) on kernel smoothing and density estimation. A kernel refers to a given probability density function, as introduced in Section 1. We introduced a first adaptation of this approach to the Box Particle Filter in Merlinge et al. (2016). However, this formulation was performed by approximating each box particle by a uniform expectancy, which corresponds to the center of the box. As a result, the regularization formulation was equivalent to a conventional particle filter regularization.

The objective of this section is to determine an optimal smoothing kernel applied to the box parameters, in terms of the Mean Integrated Square Error (MISE) criterion defined in (48). Recall that the regularisation is only performed when a resampling step is triggered by condition (22). We propose a new adaptation of the kernel regularization method which relies on the whole boxes description. The regularization takes place after the resampling operation and the correction step that we have previously presented.

Each box particle  $[\mathbf{x}_k^i]$  is characterized by a vector of  $\mathbb{R}^{2d}$  which consists of their center  $\mathbf{c}_k^i \in \mathbb{R}^d$  and their diameter  $\delta_k^i \triangleq \delta_{[\mathbf{x}_k^i]} \in \mathbb{R}^d$ . A vector description of a box particle is:

$$\xi_k^i{}^T = [\mathbf{c}_k^i{}^T, \delta_k^i{}^T] \in \mathbb{R}^{2d} \quad (44)$$

The boxes parameters (center and diameter) can be associated to a random vector  $\xi_k$ . Then, a new expression of the density's approximation can be written, using some kernels  $K$  centred on each box particle  $\xi_k^i$ .

$$\hat{p}(\xi_k | \mathbf{M}_k) \approx \sum_{i=1}^N w_k^i K_h(\xi_k - \xi_k^i) \quad (45)$$

where

$$\begin{cases} K_h : \mathbb{R}^{2d} \mapsto \mathbb{R} \\ K_h(\xi) = \frac{1}{h^{2d}} K\left(\frac{1}{h} \xi\right) \end{cases} \quad (46)$$

in the re-scaled kernel density  $K(\cdot)$ ,  $h \in \mathbb{R}^{+*}$  is the kernel bandwidth. The kernel density is a symmetric probability density function such that:

$$\int \xi K(\xi) d\xi = 0, \int \|\xi\|^2 K(\xi) d\xi < \infty \quad (47)$$

The kernel  $K(\cdot)$  and bandwidth  $h$  are chosen to minimize the Mean Integrated Square Error (MISE) between the hypothetical posterior density and the corresponding regularized filter's representation, defined as:

$$\text{MISE}(\hat{p}) = E \left[ \int (\hat{p}(\xi_k | \mathbf{M}_k) - p(\xi_k | \mathbf{M}_k))^2 d\xi_k \right] \quad (48)$$

where  $\hat{p}(\xi_k|\mathbf{M}_k)$  denotes the approximation to  $p(\xi_k|\mathbf{M}_k)$  given by (45). If all the box particles have the same weight, during the resampling step, an optimal choice of the kernel is the bounded Epanechnikov kernel (Silverman, 1986).

$$K_{opt}(\xi) = \begin{cases} \frac{2d+2}{2c_{2d}} \left(1 - \|\xi\|^2\right) & \text{if } \|\xi\| < 1 \\ 0 & \text{otherwise} \end{cases} \quad (49)$$

where  $c_{2d}$  is the volume of the unit hypersphere in  $\mathbb{R}^{2d}$ . The kernel support length is expressed as:

$$\begin{cases} h_{opt} &= \mu A(K) N^{-\frac{1}{2d+4}} \\ A(K) &= [8c_{2d}^{-1}(2d+4)(2\sqrt{\pi})^{2d}]^{\frac{1}{2d+4}} \end{cases} \quad (50)$$

where  $\mu \in [0, 1]$  is a tuning parameter, introduced to avoid an over-smoothing of the density, which would produce divergences.

After kernel smoothing regularization, the box particle kernels mixture is guaranteed to better fit the optimal posterior density. In practice, this will result in an improved estimation accuracy.

#### 4.3 BRPF algorithm

The above developments are summarized in Algorithm 1.

---

#### Algorithm 1 Box Regularised Particle Filter

---

- 1: Generate  $N$  box particles  $\{[\mathbf{x}_0^i]\}_{i \in \llbracket 1, N \rrbracket}$  of empty intersection, associated to weights  $w_0^i = 1/N$ .
  - 2: **for** each time-step  $k$  **do**
  - 3:   Propagate box particles using (12).
  - 4:   Contract box particles using (15) (if a measurement is available).
  - 5:   Update weights using (21).
  - 6:   Normalize weights using (20).
  - 7:   **if** (22) is satisfied **then**
  - 8:     Use a resampling method (e.g., multinomial resampling) to determine the number of new boxes where  $n^i \in [0, N]$  per existing box particle.
  - 9:     Chose one subdivision dimension  $d_k^{cut,i}$  per box particle using (29), (26), and (27).
  - 10:    Subdivide each box in  $n_k^i$  new boxes along its edge  $d_k^{cut,i}$ .
  - 11:    Reset all weights to  $w_k^i = 1/N$ .
  - 12:    Regularize the box particle cloud by noising each box parameter (44) using the optimal smoothing kernel defined by (49) and (50).
  - 13:   **end if**
  - 14:   Estimate state  $\hat{\mathbf{x}}_k$  (23) and its confidence  $\hat{\mathbf{P}}_k$  (24).
  - 15: **end for**
- 

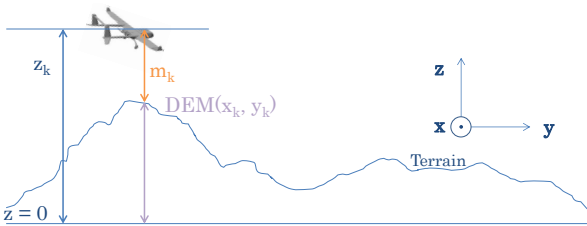


Fig. 1. Elevation measurement  $m_k^i$  in terrain navigation

## 5. NUMERICAL RESULTS, APPLICATION TO TERRAIN AIDED NAVIGATION

To illustrate the behaviour of the resulting filter with non-injective and uncertain measurements, an application to Terrain Aided Navigation (TAN) is presented. The TAN problem is a very non-linear and ambiguous estimation problem. To begin, the observation equation involves a non-analytic Digital Elevation Model (DEM) map, which can hardly be linearized. As a result, the Kalman filters (EKF, UKF) are not suitable. Then, the DEM often contains several similar patterns (e.g., peaks and valleys), which makes the relationship  $h$  between the state and the measurement non-injective. In addition, the measurement is a scalar value, which does not provide much information about the state. In practice, conventional Particle Filters often fail to solve this problem (Merlinge et al., 2016).

The state vector represents an aerial vehicle coordinate in a geographical frame at time-step  $k$ , it consists of:

$$\mathbf{x}_k^T = [\mathbf{p}_k^T, \mathbf{v}_k^T] \in \mathbb{R}^6 \quad (51)$$

where  $\mathbf{p}_k = [x_k, y_k, z_k]^T$  is the position vector and  $\mathbf{v}_k \in \mathbb{R}^3$  is the velocity vector.

The vehicle dynamical model is assumed to be linear:

$$\mathbf{x}_k = \mathbf{F} \mathbf{x}_{k-1} + \mathbf{w}_k \quad (52)$$

where  $\mathbf{F} = \begin{bmatrix} \mathbf{I}_3 & dt \mathbf{I}_3 \\ \mathbf{0}_3 & \mathbf{I}_3 \end{bmatrix}$  with  $\mathbf{0}_n \in \mathbb{R}^{n \times n}$  the null matrix and  $dt$  a time-step value. A radar altimeter provides elevation measurements (the relative height  $m^i$ , see Fig. 1) along the vehicle trajectory at discrete time values. By comparing on board these elevations with a DEM, it is possible to reconstruct the absolute position of the aircraft. The DEM gives the absolute elevation as a function of the geographical coordinates  $(x_k, y_k)$ . The measurement equation is, as illustrated in Figure 1:

$$m_k = z_k - DEM(x_k, y_k) + v_k \in \mathbb{R} \quad (53)$$

where  $DEM: \mathbb{R}^2 \rightarrow \mathbb{R}$  is the embedded terrain map and  $v_k \in \mathbb{R}$  is the measurement noise. There is no analytic description of  $DEM$ , which is assumed to be obtained from an embedded terrain map. In this paper, for reproducibility purpose only, the terrain map is analytically generated by the following equation, corresponding to the MATLAB<sup>®</sup> `peaks()` function plus a Fourier series:

$$\begin{cases} DEM(x, y) : \mathbb{R}^2 \mapsto \mathbb{R} \\ z = peaks(qx, qy) + \sum_{i=1}^6 a_i \sin(\omega_i qx) \cos(\varpi_i qy) \end{cases} \quad (54)$$

with  $a_i = \{300, 80, 60, 40, 20, 10\}$ ,  $\omega_i = \{5, 10, 20, 30, 80, 150\}$ ,  $\varpi_i = \{4, 10, 20, 40, 90, 150\}$ ,  $q = 3/(2.96 \times 10^4)$  a scale factor, and  $peaks(x, y) = 200(3(1-x)^2 e^{-x^2-(y+1)^2} - 10(x/5 - x^3 - y^5) e^{-x^2-y^2} - (1/3) e^{-(x+1)^2+y^2})$ .

#### 5.1 Simulations overview

In this section, the proposed BRPF algorithm is compared to the SIR-PF (Gordon et al., 1993), the Monte-Carlo Markov Chains (MCMC, Andrieu et al. (2010)) and the BPF (Gning et al., 2013). The individual effects of the two improvements presented in section 4 are numerically illustrated, namely the *box particles subdivision for resampling* (4.1) and the *kernel regularization smoothing* (4.2).

Table 2. Simulation configuration

Kinematics	Value
Initial position	$[-3.0, -19.2, 1.1] \times 10^3 \text{m}$
Initial velocity	$[211.5, 215.3, 0] \text{m/s}$
Initial state estimate uncertainty (st.d.)	$[1.0, 1.0, 0.1] \times 10^3 \text{m}$
Actual process noise (st.d.)	$[3.0, 3.0, 1.0] \text{m/s}$
	$[0.1, 0.1, 0.3] \text{m}$
	$[1.45, 2.28, 11.5] \times 10^{-2} \text{m/s}$
Time-step $dt$	100 ms
Final time $T$	100 s
Measurements	
Radar-altimeter error (support)	$v_k \in [-45, +45] \text{m}$
Radar-altimeter update rate	$\Delta t_{RA} = 100 \text{ms}$
BPF and BRPF	
Resampling coefficient $\theta_{eff}$	0.7
Normalization parameter $\lambda$	0.5
Process noise	None
SIR-PF	
Resampling coefficient $\theta_{eff}$	0.5
Process noise (standard deviation)	$[10, 10, 1] \text{m}$
	$[1, 1, 1] \times 10^{-1} \text{m/s}$
MCMC	
Burn-in samples	$5N$
Process noise (standard deviation)	$[10, 10, 5] \text{m}$
	$[1, 1, 10] \times 10^{-1} \text{m/s}$

Table 3. Simulation results (final time-step)

	SIR-PF	MCMC	BPF	BRPF		
$N$	$5 \times 10^4$	$5 \times 10^3$	$10^3$	$10^3$		
$\mu$	-	-	-	0	0.1	0.3
$RMSE_{\mathbf{x}}$ (m)	404	292	171	121	100	99
$RMSE_{\mathbf{v}}$ (m/s)	3.7	5.5	3.2	2.8	2.6	2.2
$\sigma_{\mathbf{x}}$ (m)	71	187	482	217	207	148
$\sigma_{\mathbf{v}}$ (m/s)	4.1	8.6	4.6	4.0	3.8	3.5
Resamp. rate (%)	20	-	2	2	2	3
Div (%)	4	15	0	0	0	1
Time (ms)	123	356	11	33	33	33

Table 2 describes the simulation parameters. The number of particles for each filter was chosen to get a similar computation load. The computation time is obtained on a 1600 MHz CPU running MATLAB®. A hundred Monte-Carlo simulations are run. The first evaluation criterion is the Root Mean Square Error (RMSE) defined by  $RMSE_{\mathbf{x}}(k) = \sqrt{\frac{1}{N_{MC}} \sum_{j=1}^{N_{MC}} \|\hat{\mathbf{x}}_{k,j} - \mathbf{x}_{k,j}\|^2}$ , where  $N_{MC} = 100$  is the number of runs. Vector  $\hat{\mathbf{x}}_{k,j}$  stands for estimate position or velocity at time-step  $k$  for run  $j$ . Vector  $\mathbf{x}_{k,j}$  stands for actual vector to be estimated. The second evaluation criterion is the mean estimate uncertainty, defined by  $\sigma_{\mathbf{x}}(k) = \frac{1}{N_{MC}} \sum_{j=1}^{N_{MC}} \|\hat{\sigma}_{\mathbf{x}}^j(k)\| \in \mathbb{R}^3$ , where  $\hat{\sigma}_{\mathbf{x}}^j(k)$  contains the diagonal coefficients of the square root of  $\hat{\mathbf{P}}_{\mathbf{x}}^j(k)$ , with  $\hat{\mathbf{P}}_{\mathbf{x}}^j(k) \in \mathbb{R}^{3 \times 3}$  the considered filter estimate covariance confidence of vector  $\mathbf{x} = \mathbf{p}_k$  or  $\mathbf{x} = \mathbf{v}_k$  (position or velocity) for simulation  $j$  at time-step  $k$ . The third criterion is a divergence counter (%), defined by  $Div(k) = \frac{100}{N_{MC}} \sum_{j=1}^{N_{MC}} (\|\hat{\mathbf{x}}_{k,j} - \mathbf{x}_{k,j}\| > 3\sigma_{\mathbf{x}}(k))$ .

## 5.2 Numerical results

Fig 2 shows the individual effects of each improvement (variance-based box particle subdivision in green and kernel smoothing regularization in red) on the RMSE criterion in comparison to the original BPF (black curve). The variance-based box particle resampling presented in 4.1 allows the filter to restrict its uncertainty and results in a

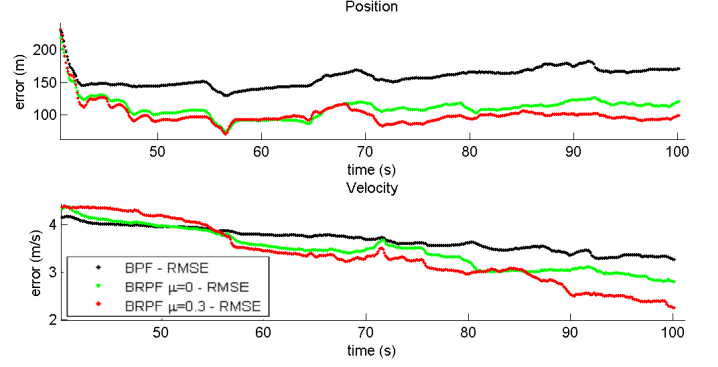


Fig. 2. RMSE for BPF and for two configurations of BRPF

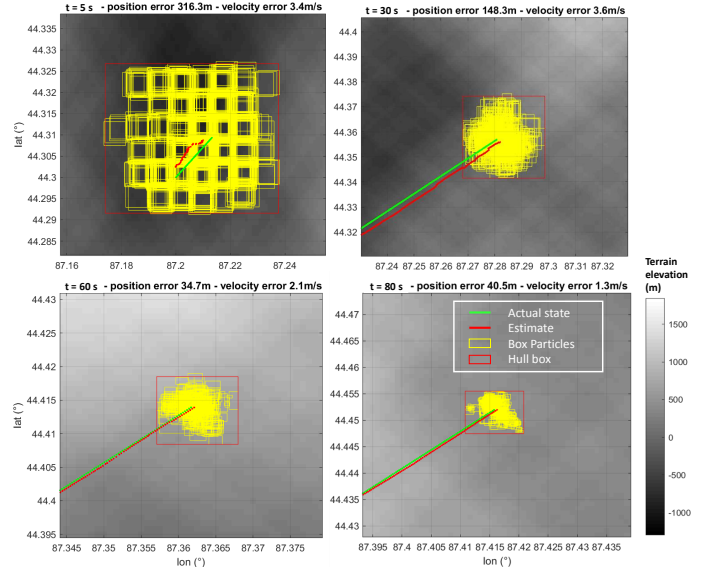


Fig. 3. BRPF simulation example at different times-steps.

lower error (green curve). The kernel smoothing regularization presented in 4.2 helps to converge more accurately. The combination of these modifications results in a lower estimation error (red curve). Figure 3 illustrates the filter's behavior for one run (BRPF,  $\mu = 0.3$ ).

Table 3 presents the performance obtained with SIR-PF, MCMC, BPF, and BRPF for several values of  $\mu$ . For a similar RMSE order of magnitude, the SIR-PF requires 50 times more particles than the BPF, which yields a significantly higher computation time than the BPF and the BRPF. The MCMC method appears more accurate than the SIR-PF with less samples but requires a higher computational load due to the burn-in period (see Andrieu et al. (2010)). SIR-PF and MCMC require a greater computation time per time-step than the desired 100 ms time-step imposed by the measurement rate (10 Hz). With a computation time of only 11 ms and 33 ms, the BPF and the BRPF yield a significantly lower RMSE. BPF and BRPF also produce a lower divergence rate ( $\leq 1\%$ ) than the other methods. This illustrates their ability to robustly tackle real-time applications with non-linear measurements. However, the original BPF yields a conservative final covariance ( $\sigma_{\mathbf{x}} = 482 \text{ m}$  and  $\sigma_{\mathbf{v}} = 4.8 \text{ m/s}$ ). As expected, the results obtained using the covariance-based subdivision resampling strategy (BRPF with  $\mu = 0$ ) are more accurate and less conservative

than those obtained using the random subdivision strategy (BPF), resulting in a lower RMSE (121 m versus 171 m) and a lower estimation confidence (217 m versus 482 m). The use of the Kernel Smoothing strategy  $\mu > 0$  makes it possible to enhance the estimate accuracy (for  $\mu = 0.1$  and  $\mu = 0.3$  the RMSE falls from 121 m to 99 m in position and from 2.8 m/s to 2.2 m/s in velocity). As a result, BRPF appears to be able to tackle ambiguous measurements (non-injective observation model) in a more accurate way than previous approaches while meeting real-time requirements.

## 6. CONCLUSION

In this work, a state estimation algorithm named Box Regularized Particle Filter (BRPF) is introduced. The algorithm description is more general than the ones presented in previous box particle filters approaches, in terms of dynamical model, observation model and involved uncertainties (Section 3). Two main improvements are introduced in the BPF classic formulation (Section 4). The first one provides a general way of subdividing the box particles during the box resampling step (Section 4.1). This formulation is analytic and deterministic. It can be applied to any dynamical and observation models. It leads the box particles to take the shape of the smallest expected variance. As a result, the estimation accuracy is significantly improved compared to the original BPF (see Section 5). The second one is a stochastic posterior kernel smoothing method (Section 4.2). It allows the estimated state density to fit the theoretical state density in an optimal way in the sense of the Mean Integrated Square Error criterion. The performance enhancement resulting from both contributions is demonstrated by numerical simulations on a very ambiguous and non-linear problem, the Terrain Aided Navigation (see Section 5). Higher accuracy is obtained by the BRPF compared to the BPF in terms of Root Mean Square Error. It is also far more robust than the conventional SIR-PF and the MCMC approach in terms of divergence rate. As a result, BRPF achieves an efficient trade-off between robustness and accuracy.

## REFERENCES

- Abdallah, F., Gning, A., and Bonnifait, P. (2007). Box particle filtering for non linear state estimation using interval analysis. *Automatica*.
- Abdallah, F., Gning, A., and Bonnifait, P. (2008). Box particle filtering for nonlinear state estimation using interval analysis. *Automatica*, 44(3), 807–815.
- Andrieu, C., Doucet, A., and Holenstein, R. (2010). Particle markov chain monte carlo methods. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 72(3), 269–342.
- De Freitas, A., Mihaylova, L., Gning, A., Angelova, D., and Kadirkamanathan, V. (2016). Autonomous crowds tracking with box particle filtering and convolution particle filtering. *Automatica*, 69, 380–394.
- Gning, A., Mihaylova, L., and Abdallah, F. (2012a). Particle filtering combined with interval methods for tracking applications. *Integrated Tracking, Classification, and Sensor Management: Theory and Applications*.
- Gning, A., Ristic, B., Mihaylova, L., and Abdallah, F. (2013). An introduction to box particle filtering. *IEEE Signal Processing Magazine*, 166–171.
- Gning, A., Ristic, B., and Mihaylova, L. (2012b). Bernoulli particle/box-particle filters for detection and tracking in the presence of triple measurement uncertainty. *IEEE Transactions on Signal Processing*, 60(5), 2138–2151.
- Gordon, N.J., Salmond, D.J., and Smith, A.F. (1993). Novel approach to nonlinear/non-gaussian bayesian state estimation. In *IEEE Proceedings F (Radar and Signal Processing)*, volume 140, 107–113. IET.
- Higham, N.J. (1988). Computing a nearest symmetric positive semidefinite matrix. *Linear algebra and its applications*, 103, 103–118.
- Jaulin, L. (2009). Robust set-membership state estimation. *Automatica*, 45, 202–206.
- Jaulin, L. (2001). *Applied interval analysis: with examples in parameter and state estimation, robust control and robotics*, volume 1. Springer Science & Business Media.
- Li, T., Bolic, M., and Djuric, P.M. (2015). Resampling methods for particle filtering: classification, implementation, and strategies. *IEEE Signal Processing Magazine*, 32(3), 70–86.
- Luo, J. and Qin, S. (2018). A fast algorithm of simultaneous localization and mapping for mobile robot based on ball particle filter. *IEEE Access*.
- Maksarov, D. and Norton, J. (1996). State bounding with ellipsoidal set description of the uncertainty. *International Journal of Control*, 65(5), 847–866.
- Merlinge, N., Dahia, K., and Piet-Lahanier, H. (2016). A box regularized particle filter for terrain navigation with highly non-linear measurements. *IFAC-PapersOnLine*, 49(17), 361–366.
- Musso, C., Oudjane, N., and LeGland, F. (2001). Improving regularized particle filters. *Sequential Monte Carlo Methods in Practice*, 12, 247–271.
- Neumaier, A. (1998). Solving ill-conditioned and singular linear systems: A tutorial on regularization. *SIAM review*, 40(3), 636–666.
- Petrov, N., Gning, A., Mihaylova, L., and Angelova, D. (2012). Box particle filtering for extended object tracking. In *Information Fusion (FUSION), 2012 15th International Conference on*, 82–89. IEEE.
- Piet-Lahanier, H. and Walter, E. (1994). Exact description of feasible parameter sets and minimax estimation. *International journal of adaptive control and signal processing*, 8(1), 5–14.
- Ristic, B. (2004). *Beyond the Kalman filter: Particle filters for tracking applications*. Artech House, Boston.
- Schweppe, F. (1968). Recursive state estimation: unknown but bounded errors and system inputs. *IEEE Trans. on Autom. Contr.*, 13, 22–28.
- Silverman, B. (1986). *Density Estimation for Statistics and Data Analysis*. Chapman & Hall, London.