Fixed-Lag Smoothing for Bayes Optimal Exploitation of External Knowledge

Francesco Papi, Mélanie Bocquel, Martin Podt, Yvo Boers
THALES Nederland B.V., Sensors-TBU Radar Engineering
Hengelo, The Netherlands, Email: francesco.papi@nl.thalesgroup.com

***Abstract*—Particle Filters (PFs) nowadays represent the state of art in nonlinear filtering. In particular, their high flexibility makes PFs particularly suited for Bayes optimal exploitation of possibly available external knowledge.**

In this paper we propose a new method for optimal processing of external knowledge that can be formalized in terms of hard constraints on the system dynamics. In particular, we are interested in the tracking performance improvements attainable when forward processing of external knowledge is performed over a moving window at every time step. That is, the one step ahead prediction of each particle is obtained through a *Fixed-Lag Smoothing* procedure, which uses *Pseudo-Measurements* to evaluate the level of adherence between each particle trajectory and the knowledge over multiple scans.

A proof of improvements is presented by utilizing *differential entropy* [1] as a measure of uncertainty. That is, we show that the differential entropy of the posterior PDF targeted by the proposed approach is always lower or equal to the differential entropy of the posterior PDF usually targeted in *constrained filtering*. Thus, for a sufficiently large number of particles, a PF implementation of the proposed *Knowledge-Based Fixed-Lag Smoother* can only improve the track accuracy upon classical algorithms for *constrained filtering*. Preliminary simulations show that the proposed approach guarantees substantial improvements when compared to the *Standard SISR-PF* and to the *Pseudo-Measurements PF*.

I. Introduction

In the last twenty years, the necessity of increasing the detection and tracking performance of civil and military surveillance systems has dictated the use of more accurate models. In turn, this has required the inclusion of nonlinearity and non-Gaussianity in the equations used for estimation purposes. In this way, the system and measurement equations can better represent the targets and sensors behaviors, especially in the case of strongly nonlinear sensors and/or highly maneuverable targets. However, such sophisticated models violate the assumptions of the Kalman Filter (KF), so that it is necessary to resort to a *more general Bayesian formulation* in order to optimally solve the *filtering problem*. Any other KF inspired solution to the *nonlinear filtering problem*, e.g., Extended Kalman Filter (EKF) or Unscented Kalman Filter (UKF), will yield suboptimal performance, independently from the available computational power.

The Bayesian approach to the *filtering problem* aims at constructing the a posteriori *probability density function* (PDF) of the state given all the available information, and then numerically approximating such a PDF. The *nonlinear filtering* problem is then recursively solved by using a PF, which is a Monte Carlo based approximation of the Bayesian recursion, and nowadays represents the state of art in nonlinear target tracking [2]. Such filters operate by propagating particles that are distributed according to the approximately true PDF of the state, and convergence to the true a posteriori distribution is guaranteed for a sufficiently large number of particles [3]. If computational power is available, a particle filtering based solution to the nonlinear filtering problem can theoretically achieve optimal detection and tracking performance.

Oftentimes, additional information about the target and/or the environment is available and can be formalized in terms of constraints on the target dynamics. Such information can be optimally processed by means of particle filtering. For instance, in [4] the use of inequality constraints in multi-target tracking is considered. The authors propose three different PF algorithms for processing of knowledge in the prediction and in the update steps of a PF, and show that the filters achieve good performance for the considered scenarios.

A refined model with state dependent detection probability and clutter notch information is proposed in [5] for airborne GMTI based ground tracking. Equality and inequality constraints are used to model the known road network. Hard inequality state constraints are used in [6] to represent the known flight envelope (i.e., minimum and maximum velocities), and a *Rejection-Sampling* approach is proposed to enforce the constraints. The method is intrinsically correct since all the predicted particles effectively respect the constraints, but computationally demanding. However, in [7] it has been shown that processing of external knowledge in the prediction or in the update step are equivalent, so that we can consider the *Pseudo-Measurements PF* as a correct way of achieving Bayes optimal exploitation of external knowledge, at least for a sufficiently large number of particles.

In this work, we propose a new method for processing of external knowledge that is described by *hard constraints* on the target dynamics. The idea is to use a *Knowledge-Based (KB) Fixed-Lag Smoothing procedure* at time $k-1$ in order to determine the predicted cloud of particles at time k. The delay L in showing the results, which is usually required when performing on-line fixed-lag smoothing, is not necessary in our case since we perform smoothing with respect to the constraints at successive time steps. In particular, the *KB Smoothing* procedure is performed within the standard prediction step of the filtering recursion.

A theoretical proof of improvements is presented by using *differential entropy* as a measure of uncertainty in continuous probability densities. That is, we show that the differential entropy of the posterior PDF targeted by the proposed approach is always lower or equal to the differential entropy of the posterior PDF usually targeted in *constrained filtering*. Thus, for a sufficiently large number of particles, a Sequential Monte Carlo (SMC) approximation of the proposed method can only improve the track accuracy upon classical algorithms for *single-scan constrained filtering*.

The paper is organized as follows: in section II we recall the *constrained filtering problem* in the case of *hard constraints*, and its Bayes optimal solution by means of the *Pseudo-Measurements PF*; in section III we consider particle filtering methods for Bayesian smoothing. We then present our approach for knowledge-based fixed-lag smoothing in section IV, along with a proof of optimality by means of entropy minimization. Simulation results are reported in section V, while our conclusions are discussed in section VI.

II. Constrained Bayesian Filtering

In this section, we briefly recall the *constrained Bayesian nonlinear filtering* problem and its optimal solution through *Pseudo-Measurements* based processing of knowledge. Hence, let us consider the following nonlinear state-space model:

$$\mathbf{x}_{k+1} = \mathbf{f}_k(\mathbf{x}_k) + \mathbf{w}_k \tag{1}$$

$$\mathbf{z}_k = \mathbf{h}_k(\mathbf{x}_k) + \mathbf{v}_k \tag{2}$$

where $\mathbf{x}_k \in \mathbb{R}^{n_x}$ is the system state, $\mathbf{z}_k \in \mathbb{R}^{n_z}$ the measurement vector, $\mathbf{w}_k \sim p_{\mathbf{w}_k}(\mathbf{w})$ the process noise, and $\mathbf{v}_k \sim p_{\mathbf{v}_k}(\mathbf{v})$ the measurement noise. The Markov property holds for the system (1)-(2), i.e.,

$$p(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{x}_{k-2}, \dots, \mathbf{x}_0) = p(\mathbf{x}_k|\mathbf{x}_{k-1}) \tag{3}$$

where $p(\mathbf{x}_k|\mathbf{x}_{k-1})$ is known as the *transition model (Kernel)*.

Let $\mathbf{z}_{0:k} \triangleq \{ \mathbf{z}_1 \quad \mathbf{z}_2 \quad \dots \quad \mathbf{z}_k \}$ be the sequence of measurements up to and including time k. Assume the measurement $\mathbf{z}_k$ at time k is independent from past states, i.e.,

$$p(\mathbf{z}_k|\mathbf{z}_{k-1}, \dots, \mathbf{z}_1, \mathbf{x}_k, \mathbf{x}_{k-1}, \dots, \mathbf{x}_0) = p(\mathbf{z}_k|\mathbf{x}_k) \tag{4}$$

where $p(\mathbf{z}_k|\mathbf{x}_k)$ is known as the *likelihood function*.

An expression for the *joint smoothing distribution* of states and observations is obtained by the probability chain rule, i.e.,

$$p(\mathbf{x}_{0:k}|\mathbf{z}_{0:k}) = p(\mathbf{x}_0) \left(\prod_{i=1}^{k} p(\mathbf{x}_i|\mathbf{x}_{i-1}) \right) \left(\prod_{0=1}^{k} p(\mathbf{z}_i|\mathbf{x}_i) \right) \tag{5}$$

where $\mathbf{x}_{0:k} \triangleq \{ \mathbf{x}_0 \quad \mathbf{x}_1 \quad \dots \quad \mathbf{x}_k \}$. A primary concern is the sequential estimation of the *filtering distribution* $p(\mathbf{x}_k|\mathbf{z}_{0:k})$. Given $p(\mathbf{x}_{k-1}|\mathbf{z}_{0:k-1})$ and $\mathbf{z}_k$, the *Bayesian Filtering* problem is solved using the following two step recursion:

- ***Prediction Step***

$$p(\mathbf{x}_k|\mathbf{z}_{0:k-1}) = \int p(\mathbf{x}_k|\mathbf{x}_{k-1}) \; p(\mathbf{x}_{k-1}|\mathbf{z}_{0:k-1}) \; d\mathbf{x}_{k-1} \tag{6}$$

where $p(\mathbf{x}_k|\mathbf{z}_{0:k-1})$ is the *predictive density* at time k.

- ***Update Step***

$$p(\mathbf{x}_k|\mathbf{z}_{0:k}) = \frac{p(\mathbf{z}_k|\mathbf{x}_k) \; p(\mathbf{x}_k|\mathbf{z}_{0:k-1})}{p(\mathbf{z}_k|\mathbf{z}_{0:k-1})} \tag{7}$$

where $p(\mathbf{z}_k|\mathbf{z}_{0:k-1})$ is the *Bayes normalization constant*.

Various state estimators are obtained from the posterior PDF $p(\mathbf{x}_k|\mathbf{z}_{0:k})$, like for instance the *minimum variance* (MV) estimator or the *maximum a posteriori* (MAP) estimator. In general, if $\boldsymbol{\phi}(\mathbf{x}_k) : \mathbb{R}^{n_x} \rightarrow \mathbb{R}^{n_\phi}$ is a function of the state we want to estimate, most estimation algorithms compute an approximation of the conditional expectation:

$$\mathbf{E}\left(\boldsymbol{\phi}(\mathbf{x}_k)|\mathbf{z}_{0:k}\right) = \int \boldsymbol{\phi}(\mathbf{x}_k) \; p(\mathbf{x}_k|\mathbf{z}_{0:k}) \; d\mathbf{x}_k \tag{8}$$

The particle filter computes an approximation of (8) using the empirical filtering density, i.e.,

$$\hat{p}_N(\mathbf{x}_k|\mathbf{z}_{0:k}) = \sum_{i=1}^{N} w_k^i \; \delta_{\mathbf{x}_k^i}(\mathbf{x}_k) \tag{9}$$

where each particle $\mathbf{x}_k^i$ has an importance weight w_k^i associated to it, and $\delta_{\mathbf{x}_k^i}(\cdot)$ denotes the delta-Dirac mass located at $\mathbf{x}_k^i$. Convergence results for PFs are surveyed in [3].

Algorithm 1: Pseudo-Measurements Particle Filter

Input: $\left\{\mathbf{x}_{k-1}^i, w_{k-1}^i\right\}_{i=1}^N$ and the new measurement $\mathbf{z}_k$
Output: $\{\mathbf{x}_k^i, w_k^i\}_{i=1}^N$
while $i = 1, 2, \dots, N$ ***(Prediction Step)*** **do**
  Generate a New Particle: $\mathbf{x}_k^i \sim p_k(\mathbf{x}_k^i|\mathbf{x}_{k-1}^i)$
end
while $i = 1, 2, \dots, N$ ***(Update Step)*** **do**
  Compute Weights: $\tilde{w}_k^i = w_{k-1}^i \; p(\mathbf{z}_k|\mathbf{x}_k^i) \; p(\mathcal{C}_k|\mathbf{x}_k)$;
end
Normalization Step: $w_k^i = \tilde{w}_k^i / \sum_{i=1}^N \tilde{w}_k^i \quad \forall \; i$;
Effective Sample Size: $N_{eff} = 1/\sum_{i=1}^N (w_k^i)^2$;
if $N_{eff} \leq \beta N$ ***(Resampling Step)*** **then**
  New Particles $\left\{\tilde{\mathbf{x}}_k^i, 1/N\right\}_{i=1}^N$ s.t. $P(\tilde{\mathbf{x}}_k^i = \mathbf{x}_k^i) = w_k^i$
end

Oftentimes additional information about the state is available, and can be formalized in terms of *hard constraints* on the system dynamics. This happens for instance in the tracking of ground vehicles moving on a road network, or when tracking a ship that is traveling in a canal. Hence, let $\mathbf{C}_k : \mathbb{R}^{n_x} \rightarrow \mathbb{R}^{n_c}$ be a function describing the constraints, and $\mathcal{C}_k$ the set of all states satisfying inequality constraints, i.e.,

$$\mathcal{C}_k \triangleq \{\mathbf{x}_k : \mathbf{x}_k \in \mathbb{R}^{n_x}, \mathbf{a}_k \leq \mathbf{C}_k(\mathbf{x}_k) \leq \mathbf{b}_k\} \tag{10}$$

Let $\mathcal{C}_{0:k} \triangleq \{\mathcal{C}_0, \mathcal{C}_1, \dots, \mathcal{C}_k\}$ be the sequence of $\mathcal{C}_k$ up to time k. From a Bayesian viewpoint, exploitation of external knowledge boils down to finding an approximation of

$$p(\mathbf{x}_k|\mathbf{z}_{0:k}, \mathcal{C}_{0:k}) \propto \begin{cases} p(\mathbf{x}_k|\mathbf{z}_{0:k}, \mathcal{C}_{0:k-1}), & \text{if } \mathbf{x}_k \in \mathcal{C}_k \\ 0, & \text{otherwise} \end{cases} \tag{11}$$

The external knowledge contained in $\mathcal{C}_{0:k}$ can be exploited in both the prediction or update steps. The methods are equivalent from a Bayesian viewpoint, and lead to the definition of two particles filters which are shown to be numerically equivalent for an increasing number of particles [7].

Let us assume that processing of the new information at time k, described by $\mathcal{C}_k$, is performed in the update step. Hence, the Bayesian filtering recursion takes the form:

- ***Prediction Step***

$$\begin{aligned} &p(\mathbf{x}_k|\mathbf{z}_{0:k-1}, \mathcal{C}_{0:k-1}) \\ &\quad = \int p(\mathbf{x}_k|\mathbf{x}_{k-1})\ p(\mathbf{x}_{k-1}|\mathbf{z}_{0:k-1}, \mathcal{C}_{0:k-1})\ d\mathbf{x}_{k-1} \end{aligned} \tag{12}$$

- ***Update Step***

$$\begin{aligned} &p(\mathbf{x}_k|\mathbf{z}_{0:k}, \mathcal{C}_{0:k}) \\ &\quad = \frac{p(\mathbf{z}_k|\mathbf{x}_k)\ p(\mathcal{C}_k|\mathbf{x}_k)\ p(\mathbf{x}_k|\mathbf{z}_{0:k-1}, \mathcal{C}_{0:k-1})}{p(\mathbf{z}_k|\mathbf{z}_{0:k-1}, \mathcal{C}_{0:k})\ p(\mathcal{C}_k|\mathcal{C}_{0:k-1})} \end{aligned} \tag{13}$$

where $p(\mathcal{C}_k|\mathbf{x}_k^i)$ is the *Pseudo-Measurements likelihood*, i.e.,

$$p(\mathcal{C}_k|\mathbf{x}_k^i) = \begin{cases} 1, & \text{if } \mathbf{a}_k \leq \mathbf{C}_k(\mathbf{x}_k^i) \leq \mathbf{b}_k \\ 0, & \text{otherwise} \end{cases} \tag{14}$$

which leads to the definition of the Bayes optimal *Pseudo-Measurements PF* described in Algorithm 1.

III. Particle Filtering for Bayesian Smoothing

In this section we review the particle filtering approach to *Bayesian Smoothing* with a specific interest in fixed-lag smoothing. The first thing to notice is that the basic filtering recursion provides an approximation of the joint smoothing distribution [8]. In fact, using a PF from time 0 to k, the stored particle trajectories $\left\{\mathbf{x}_{0:k}^i\right\}_{i=1}^N$ along with their weights $\left\{w_k^i\right\}_{i=1}^N$ can be viewed as weighted samples drawn from the *joint smoothing distribution* $p(\mathbf{x}_{0:k}|\mathbf{z}_{0:k})$, i.e.,

$$p(\mathbf{x}_{0:k}|\mathbf{z}_{0:k}) \approx \sum_{i=1}^N w_k^i\ \delta_{\mathbf{x}_{0:k}^i}(\mathbf{x}_{0:k}) \tag{15}$$

with $\sum_{i=1}^N w_k^i = 1$, $w_k^i \geq 0$, is a correct empirical approximation of the *joint smoothing distribution*.

Given a smoothing lag L, from the joint draws $\left\{\mathbf{x}_{0:k}^i, w_k^i\right\}_{i=1}^N$ one can readily extract an approximation of the *fixed-lag smoothing distribution* $p(\mathbf{x}_{0:k-L}|\mathbf{z}_{0:k})$ as follows:

$$p(\mathbf{x}_{0:k-L}|\mathbf{z}_{0:k}) \approx \sum_{i=1}^N w_k^i\ \delta_{\mathbf{x}_{0:k-L}^i}(\mathbf{x}_{0:k-L}) \tag{16}$$

Similarly, given a positive interval $\Delta = L - L_0$, an approximation to the *fixed-interval smoothing distribution* is:

$$\begin{aligned} &p(\mathbf{x}_{k-L_0+1:k-L}|\mathbf{z}_{0:k}) \approx \\ &\qquad \sum_{i=1}^N w_k^i\ \delta_{\mathbf{x}_{k-L_0+1:k-L}^i}(\mathbf{x}_{k-L_0+1:k-L}) \end{aligned} \tag{17}$$

which is of interest when studying the cross-correlation over time between the state variables.

In practice, if *independent identically distributed* (iid) samples $\left\{\mathbf{x}_{0:k-1}^i, 1/N\right\}_{i=1}^N$ are available, the standard factorization is used to move from time k to $k+1$, i.e.,

$$p(\mathbf{x}_{0:k}|\mathbf{z}_{0:k}) = p(\mathbf{x}_{0:k-1}|\mathbf{z}_{0:k-1})\ \frac{p(\mathbf{z}_k|\mathbf{x}_k)\ p(\mathbf{x}_k|\mathbf{x}_{k-1})}{p(\mathbf{z}_k|\mathbf{z}_{0:k-1})} \tag{18}$$

and the importance weights are updated as usual, i.e.,

$$w_k \propto \frac{p(\mathbf{z}_k|\mathbf{x}_k)\ p(\mathbf{x}_k|\mathbf{x}_{k-1})}{q(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{z}_k)} \tag{19}$$

where $q(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{z}_k)$ is a suitably chosen *proposal distribution*. A new set of iid samples is obtained by drawing N trajectories at random with replacement from $\left\{\mathbf{x}_{0:k}^i, w_k^i\right\}_{i=1}^N$, [9]. Thus, one of the approximations in eqs. (15) to (17) can be used to solve the smoothing problem at hand. This approach to smoothing is known as *trajectory-based smoothing*, or *ancestry tree*. However, due to the *resampling step*, for large values of L and Δ the approximation to the smoothing distribution turns out to be strongly depleted and inaccurate.

A. Forward Filtering Backward Smoothing

We are now interested in improving the above basic scheme. This can be done by extending the Kalman forward-backward smoothing recursion [10] to nonlinear problems. Such technique is known as *forward filtering backward smoothing (FFBS)*, or the *Baum-Welch algorithm* [11].

The joint distribution can be factorized as follows:

$$\begin{aligned} p\left(\mathbf{x}_{0:T}|\mathbf{z}_{0:T}\right) &= p\left(\mathbf{x}_T|\mathbf{z}_{0:T}\right) \prod_{k=0}^{T-1} p\left(\mathbf{x}_k|\mathbf{x}_{k+1:T}, \mathbf{z}_{0:T}\right) \\ &= p\left(\mathbf{x}_T|\mathbf{z}_{0:T}\right) \prod_{k=0}^{T-1} p\left(\mathbf{x}_k|\mathbf{x}_{k+1}, \mathbf{z}_{0:T}\right) \end{aligned} \tag{20}$$

where the term in the product can be expressed as

$$p\left(\mathbf{x}_k|\mathbf{x}_{k+1}, \mathbf{z}_{0:T}\right) = \frac{p\left(\mathbf{x}_k|\mathbf{z}_{0:k}\right)\ p\left(\mathbf{x}_{k+1}|\mathbf{x}_k\right)}{\int p\left(\mathbf{x}_k|\mathbf{z}_{0:k}\right)\ p\left(\mathbf{x}_{k+1}|\mathbf{x}_k\right)\ d\mathbf{x}_k} \tag{21}$$

$$\propto p\left(\mathbf{x}_k|\mathbf{z}_{0:k}\right)\ p\left(\mathbf{x}_{k+1}|\mathbf{x}_k\right) \tag{22}$$

Assume that filtering has been performed up to time T, leading to a particle representation $\left\{\mathbf{x}_{0:k}^i, w_k^i\right\}_{i=1}^N$ for $k = 0, 1, \ldots, T$. A particle approximation to $p\left(\mathbf{x}_k|\mathbf{x}_{k+1}, \mathbf{z}_{0:T}\right)$ in (22) is obtained straightforwardly from the weighted samples, i.e.,

$$p(\mathbf{x}_k|\mathbf{x}_{k+1}, \mathbf{z}_{0:T}) \approx \sum\nolimits_{i=1}^N \rho_k^i(\mathbf{x}_{k+1})\ \delta_{\mathbf{x}_k^i}\left(\mathbf{x}_k\right) \tag{23}$$

where the *backward weights* are evaluated as [12]:

$$\rho_k^i(\mathbf{x}_{k+1}) \triangleq \frac{w_k^i\ p\left(\mathbf{x}_{k+1}|\mathbf{x}_k^i\right)}{\sum_{j=1}^N w_k^j\ p\left(\mathbf{x}_{k+1}|\mathbf{x}_k^j\right)} \tag{24}$$

Using the sampling importance resampling idea, the particle-based distribution in (23) can be used to generate samples backwards in time. However, one is oftentimes interested in the marginal smoothing distribution $p\left(\mathbf{x}_k|\mathbf{z}_{0:T}\right)$ for some time $k<T$. Hence, the following backward recursion over a fixed interval 0 to T can be used, i.e.,

$$\begin{aligned} & p\left(\mathbf{x}_k|\mathbf{z}_{0:T}\right) \\ & \quad = p\left(\mathbf{x}_k|\mathbf{z}_{0:k}\right)\int\frac{p\left(\mathbf{x}_{k+1}|\mathbf{z}_{0:T}\right)\ p\left(\mathbf{x}_{k+1}|\mathbf{x}_k\right)}{\int p\left(\mathbf{x}|\mathbf{z}_{0:k}\right)\ p\left(\mathbf{x}_{k+1}|\mathbf{x}_k\right)\ d\mathbf{x}}d\mathbf{x}_{k+1} \\ & \quad = \int p\left(\mathbf{x}_{k+1}|\mathbf{z}_{0:T}\right)\ p(\mathbf{x}_k|\mathbf{x}_{k+1},\mathbf{z}_{0:T})\ d\mathbf{x}_{k+1} \qquad (25) \end{aligned}$$

where $p(\mathbf{x}_k|\mathbf{x}_{k+1},\mathbf{z}_{0:T})$ simplifies as in (23).

In a Monte Carlo implementation of eq. (25), one recursively obtains particle estimates of the marginal smoothing distribution at the next time instant, i.e., $p\left(\mathbf{x}_{k+1}|\mathbf{z}_{0:T}\right)$, and combines these with the particle filtering estimates of $p\left(\mathbf{x}_k|\mathbf{z}_{0:k}\right)$ in (25). A drawback of this method is that a Monte Carlo estimate is required to approximate the term $p\left(\mathbf{x}_{k+1}|\mathbf{z}_{0:k}\right)$ at denominator in (21). Proceeding, if we now approximate the term $p\left(\mathbf{x}_{k+1}|\mathbf{z}_{0:T}\right)$ using the weighted samples $\left\{\mathbf{x}^i_{k+1}, w^i_{k+1|0:T}\right\}_{i=1}^N$, an approximation of the *marginal smoothing distribution* is obtained as follows:

$$p\left(\mathbf{x}_k|\mathbf{z}_{0:T}\right)\approx\sum\nolimits_{i=1}^N w^i_{k|0:T}\ \delta_{\mathbf{x}^i_k}(\mathbf{x}_k) \qquad (26)$$

where the new weights are given by:

$$w^i_{k|0:T}=w^i_k\ \left(\sum_{j=1}^N\frac{w^j_{k+1|0:T}\ p\left(\mathbf{x}^j_{k+1}|\mathbf{x}^i_k\right)}{\sum_{l=1}^N w^l_k\ p\left(\mathbf{x}^j_{k+1}|\mathbf{x}^l_k\right)}\right) \qquad (27)$$

Other forms of marginal smoothing can be obtained using the so-called two filter formula [13]. In this case, an approximation of the marginal smoothed posterior distribution is obtained by combining the output of two independent filters. One of the filters recurses in the forwards time direction and calculates $p(\mathbf{x}_k|\mathbf{z}_{0:k-1})$, while the other recurses in the backwards time direction calculating $p(\mathbf{z}_{k:T}|\mathbf{x}_k)$. The distribution $p(\mathbf{x}_k|\mathbf{z}_{0:T})$ is then given by:

$$\begin{aligned} p(\mathbf{x}_k|\mathbf{z}_{0:T}) &= \frac{p(\mathbf{x}_k|\mathbf{z}_{0:k-1})\ p(\mathbf{z}_{k:T}|\mathbf{z}_{0:k-1},\mathbf{x}_k)}{p(\mathbf{z}_{k:T}|\mathbf{z}_{0:k-1})} \qquad (28) \\ &\approx p(\mathbf{x}_k|\mathbf{z}_{0:k}\ p(\mathbf{z}_{k+1:T}|\mathbf{x}_k) \qquad (29) \end{aligned}$$

The algorithm defined in this way will generally have a lower computational load for certain modeling assumptions. In addition, since the forward and backward recursions are based on different particle approximations, the effects of depletion should be reduced with respect to the FFBS approach.

IV. Knowledge-Based Fixed-Lag Smoother

In this section we introduce our new approach for processing of possibly available external knowledge. The filter is described in Algorithm 2 and is called *Knowledge-Based (KB) Fixed-Lag Smoother*. However, notice that since at time k we already know the mathematical structure of the sets $\mathcal{C}_k,\mathcal{C}_{k+1},\ldots,\mathcal{C}_{k+L}$ defining the constraints at successive time instants, the algorithm behaves like a standard filter and can be implemented on-line without need of any delay. We use the word *Smoother* to point out that a smoothing procedure is performed to process the knowledge within the prediction step of the standard filtering recursion.

Before going into the mathematical details, let us introduce the idea behind our method. At time k, before processing of the new measurement $\mathbf{z}_k$, it is required to predict the states from time $k-1$ to time k. If while doing so we exploit the constraints at time k, i.e., we predict constrained particles, the intrinsic uncertainty in the posterior PDF will be reduced. If we extend this concept and determine the predicted cloud at time k while exploiting the constraints at times $k,k+1,\ldots,k+L$, i.e., we predict constrained particles that generate constrained trajectories, the intrinsic uncertainty in the posterior PDF will be further reduced. Consider for instance the tracking of a ship that is moving in a canal. If the predicted particles cloud at time k is obtained respecting the canal topology at successive time instants, the track accuracy can only be improved.

Algorithm 2: Knowledge-Based Fixed-Lag Smoother

Input: $\left\{\mathbf{x}^i_{k-1}, w^i_{k-1}\right\}_{i=1}^N$ and the new measurement $\mathbf{z}_k$

Output: $\{\mathbf{x}^i_k, w^i_k\}_{i=1}^N$

Forward Recursion

while $n=k,k+1,\ldots,k+L$ **do**
  while $i=1,2,\ldots,N$ **do**
    Predict Particles: $\mathbf{x}^i_n\sim p_k(\mathbf{x}_n|\mathbf{x}^i_{n-1})$;
    Evaluate Weights: $\tilde{w}^i_n=w^i_{n-1}\ p(\mathbf{c}_n|\mathbf{x}^i_n)$;
  end
  Normalization: $w^i_n=\tilde{w}^i_n/\sum_{i=1}^N\tilde{w}^i_n,\quad\forall i=1,2,\ldots,N$;
  Resampling if necessary:
  if $\left(\sum_{i=1}^N(w^i_n)^2\right)^{-1}\le\gamma N$ **then**
    $\left\{\tilde{\mathbf{x}}^i_n,1/N\right\}\sim\left\{\mathbf{x}^i_n,w^i_n\right\}\quad\forall i=1,2,\ldots,N$;
  end
end

Backward Recursion

while $n=k+L,k+L-1,\ldots,k$ **do**
  Evaluate Backward Weights:
  $w^i_{n|k:T}=w^i_n\ \left(\sum_{j=1}^N\frac{w^j_{n+1|k:T}\ p(x^j_{n+1}|x^i_n)}{\sum_{l=1}^N w^l_n\ p(x^j_{n+1}|x^l_n)}\right)$;
end

$\left(\text{We now have } p(\mathbf{x}_k|\mathbf{c}_{k:T})\approx\sum_{i=1}^N w_{k|k:T}\ \delta_{\mathbf{x}^i_k}(\mathbf{x})\right)$;

Standard Update Step

while $i=1,2,\ldots,N$ **do**
  Evaluate Weights: $\tilde{w}^i_k=w^i_{k|k:T}\ p(\mathbf{z}_k|\mathbf{x}^i_k)$;
end

Normalization: $w^i_k=\tilde{w}^i_k/\sum_{i=1}^N\tilde{w}^i_k\quad\forall i=1,2,\ldots,N$;

Resampling if necessary:

if $\left(\sum_{i=1}^N(w^i_k)^2\right)^{-1}\le\beta N$ **then**
  $\left\{\tilde{\mathbf{x}}^i_k,1/N\right\}\sim\left\{\mathbf{x}^i_k,w^i_n\right\}\quad\forall i=1,2,\ldots,N$;
end

In order to give a theoretical justification to our approach, we are interested in showing that in the case of *hard con-*

straints, a *fixed-lag smoothing* procedure for processing of the knowledge can only reduce the intrinsic uncertainty in the posterior PDF. In particular, by using *differential entropy* as a measure of uncertainty in continuous probability densities, we show that the intrinsic uncertainty in the posterior PDF targeted by the proposed approach is less than or equal to the intrinsic uncertainty in the posterior PDF usually targeted in *single-scan constrained filtering*. Such property holds in general if averaging with respect to the measurements, i.e.,

$$h(\mathbf{x}) \leq h(\mathbf{x}|\mathbf{z}) \tag{30}$$

where $h(\mathbf{x})$ is the entropy of $p(\mathbf{x})$ and $h(\mathbf{x}|\mathbf{z})$ is the entropy of $p(\mathbf{x}|\mathbf{z})$, $\mathbf{x}$ and $\mathbf{z}$ being two random variables that are *not* conditionally independent. However, given a realization $\bar{\mathbf{z}}$ of $\mathbf{z}$ the property does not hold in general [14, p. 43]. We will show that for the case of *hard constraints* the property holds for any realization of the *Pseudo-Measurements* used to exploit the constraints. That is, for a sufficiently large number of particles, a Sequential Monte Carlo (SMC) approximation of such procedure will always yield better or equal tracking performance than *single-scan constrained filtering*.

***Theorem** 4.1:* Let $\mathbf{c}_k, \mathbf{c}_{k+1}, \ldots, \mathbf{c}_T$ be the sequence of *Pseudo-Measurements* describing the system constraints from time k to time $T = k + L,\ L > 0$ being the fixed-lag.

Given a single realization $\bar{\mathbf{c}}_k, \bar{\mathbf{c}}_{k+1}, \ldots, \bar{\mathbf{c}}_T$, the entropy of the fixed-lag smoothing PDF, $p(\mathbf{x}_k|\bar{\mathbf{c}}_{k:T})$, is less or equal than the entropy of the filtering PDF, $p(\mathbf{x}_k|\bar{\mathbf{c}}_k)$, i.e.,

$$h(\mathbf{x}_k|\bar{\mathbf{c}}_{k:T}) \leq h(\mathbf{x}_k|\bar{\mathbf{c}}_k),\ \forall \bar{\mathbf{c}}_{k:T} \in S_{\mathbf{c}} \tag{31}$$

Proof:
The entropy of the *filtering PDF* is defined as:

$$h(\mathbf{x}_k|\bar{\mathbf{c}}_k) = -\int\limits_{S_{\mathbf{x}_k}} p(\mathbf{x}_k|\bar{\mathbf{c}}_k)\ \log p(\mathbf{x}_k|\bar{\mathbf{c}}_k)\ d\mathbf{x}_k \tag{32}$$

while the entropy of the *smoothing PDF* is defined as:

$$h(\mathbf{x}_k|\bar{\mathbf{c}}_{k:T}) = -\int\limits_{S_{\mathbf{x}_k}} p(\mathbf{x}_k|\bar{\mathbf{c}}_{k:T})\ \log p(\mathbf{x}_k|\bar{\mathbf{c}}_{k:T})\ d\mathbf{x}_k \tag{33}$$

Let us define $S_{\mathbf{x}}^{\geq 1} = \{\mathbf{x} \in S_{\mathbf{x}} : p(\mathbf{x}) \geq 1\}$ and $S_{\mathbf{x}}^{<1} = \{\mathbf{x} \in S_{\mathbf{x}} : p(\mathbf{x}) < 1\}$, where $p(\mathbf{x})$ is a generic pdf. Given $\alpha(\cdot)$ such that $\alpha(\mathbf{x}) \geq 1,\ \forall \mathbf{x} \in S_{\mathbf{x}}$, the following properties hold, i.e.,

$$\begin{aligned} -\int_{S_{\mathbf{x}}^{\geq 1}} &\alpha(\mathbf{x})p(\mathbf{x}) \log \alpha(\mathbf{x})p(\mathbf{x}) dp(\mathbf{x}) \\ &\leq -\int_{S_{\mathbf{x}}^{\geq 1}} p(\mathbf{x}) \log p(\mathbf{x}) dp(\mathbf{x}) \leq 0 \end{aligned} \tag{34}$$

$$\begin{aligned} -\int_{S_{\mathbf{x}}^{<1}} &\alpha(\mathbf{x})p(\mathbf{x}) \log \alpha(\mathbf{x})p(\mathbf{x}) dp(\mathbf{x}) \\ &\leq -\int_{S_{\mathbf{x}}^{<1}} p(\mathbf{x}) \log p(\mathbf{x}) dp(\mathbf{x}) \end{aligned} \tag{35}$$

as well as the two filter formula, i.e.,

$$\begin{aligned} p(\mathbf{x}_k|\bar{\mathbf{c}}_{k:T}) &= \frac{p(\bar{\mathbf{c}}_{k+1:T}|\bar{\mathbf{c}}_k, \mathbf{x}_k)}{p(\bar{\mathbf{c}}_{k+1:T}|\bar{\mathbf{c}}_k)}\ p(\mathbf{x}_k|\bar{\mathbf{c}}_k) \\ &= \alpha(\mathbf{x}_k)\ p(\mathbf{x}_k|\bar{\mathbf{c}}_k) \end{aligned} \tag{36}$$

We now want to show that $\alpha(\mathbf{x}_k) \geq 1$ holds independently from $\mathbf{x}_k$ and $\bar{\mathbf{c}}_k, \ldots, \bar{\mathbf{c}}_T$. Consider first the numerator, i.e.,

$$\begin{aligned} &p(\bar{\mathbf{c}}_{k+1:T}|\bar{\mathbf{c}}_k, \mathbf{x}_k) \\ &\quad= \int\limits_{\times_{l=k+1}^{T} S_{\mathbf{x}_l}} p(\bar{\mathbf{c}}_{k+1:T}, \mathbf{x}_{k+1:T}|\bar{\mathbf{c}}_k, \mathbf{x}_k)\ d\mathbf{x}_{k+1:T} \\ &\quad= \int\limits_{\times_{l=k+1}^{T} S_{\mathbf{x}_l}} p(\bar{\mathbf{c}}_{k+1:T}, \mathbf{x}_{k+1:T})\ p(\mathbf{x}_{k+1:T}|\mathbf{x}_k)\ d\mathbf{x}_{k+1:T} \\ &\quad= \int\limits_{\times_{l=k+1}^{T} S_{\mathbf{x}_l}} \prod_{j=1}^{L} p(\bar{\mathbf{c}}_{k+j}|\mathbf{x}_{k+j}) \prod_{j=0}^{L-1} p(\mathbf{x}_{k+j+1}|\mathbf{x}_{k+j}) d\mathbf{x}_{k+1:T} \\ &\quad= \int\limits_{\times_{l=k+1}^{T} \mathcal{C}_l} \prod_{j=0}^{L-1} p(\mathbf{x}_{k+j+1}|\mathbf{x}_{k+j})\ d\mathbf{x}_{k+1:T} \end{aligned} \tag{37}$$

Where we use the notation $\times_{l=k+1}^{T} \mathcal{C}_l = \mathcal{C}_{k+1} \times \ldots \times \mathcal{C}_T$ for the Cartesian product. Consider now the denominator, i.e.,

$$\begin{aligned} &p(\bar{\mathbf{c}}_{k+1:T}|\bar{\mathbf{c}}_k) \\ &\quad= \int\limits_{\times_{l=k+1}^{T} S_{\mathbf{x}_l}} p(\bar{\mathbf{c}}_{k+1:T}, \mathbf{x}_{k+1:T}|\bar{\mathbf{c}}_k)\ d\mathbf{x}_{k+1:T} \\ &\quad= \int\limits_{\times_{l=k+1}^{T} S_{\mathbf{x}_l}} p(\bar{\mathbf{c}}_{k+1:T}, \mathbf{x}_{k+1:T})\ p(\mathbf{x}_{k+1:T}|\bar{\mathbf{c}}_k)\ d\mathbf{x}_{k+1:T} \\ &\quad= \int\limits_{\times_{l=k+1}^{T} \mathcal{C}_l} \int\limits_{S_{\mathbf{x}_k}} p(\mathbf{x}_{k+1:T}|\mathbf{x}_k) p(\mathbf{x}_k|\bar{\mathbf{c}}_k) d\mathbf{x}_k\ d\mathbf{x}_{k+1:T} \end{aligned} \tag{38}$$

Consider the integral over $S_{\mathbf{x}_k}$ in the above formula, i.e.,

$$\begin{aligned} &\int\limits_{S_{\mathbf{x}_k}} p(\mathbf{x}_{k+1:T}|\mathbf{x}_k) p(\mathbf{x}_k|\bar{\mathbf{c}}_k) d\mathbf{x}_k \\ &\qquad= \int\limits_{S_{\mathbf{x}_k}} \prod_{j=0}^{L-1} p(\mathbf{x}_{k+j+1}|\mathbf{x}_{k+j}) p(\mathbf{x}_k|\bar{\mathbf{c}}_k) d\mathbf{x}_k \\ &\qquad\leq \prod_{j=0}^{L-1} p(\mathbf{x}_{k+j+1}|\mathbf{x}_{k+j}) \end{aligned} \tag{39}$$

Compare now equations (37) and (38) by means of the inequality in (39). Thus, the following holds true, i.e.,

$$\alpha(\mathbf{x}_k) \geq 1,\ \forall \mathbf{x}_k \in S_{\mathbf{x}_k}, \mathbf{c}_k \in S_{\mathbf{z}_k}, \ldots, \mathbf{c}_T \in S_{\mathbf{z}_T} \tag{40}$$

The entropy of the smoothed density is given by:

$$h(\mathbf{x}_k|\bar{\mathbf{c}}_{k:T}) = -\int\limits_{S_{\mathbf{x}_k}} \alpha(\mathbf{x}_k) p(\mathbf{x}_k|\bar{\mathbf{c}}_k) \log \alpha(\mathbf{x}_k) p(\mathbf{x}_k|\bar{\mathbf{c}}_k)\ d\mathbf{x}_k \tag{41}$$

We can rewrite the entropies in (32) and (41) as follows

$$h(\mathbf{x}_k|\bar{\mathbf{c}}_k) = h\left(\mathbf{x}_k|\bar{\mathbf{c}}_k, S_{\mathbf{x}}^{\geq 1}\right) + h\left(\mathbf{x}_k|\bar{\mathbf{c}}_k, S_{\mathbf{x}}^{<1}\right) \tag{42}$$

$$h(\mathbf{x}_k|\bar{\mathbf{c}}_{k:T}) = h\left(\mathbf{x}_k|\bar{\mathbf{c}}_{k:T}, S_{\mathbf{x}}^{\geq 1}\right) + h\left(\mathbf{x}_k|\bar{\mathbf{c}}_{k:T}, S_{\mathbf{x}}^{<1}\right) \tag{43}$$

and using eqs. (34) and (35) we prove our claim, i.e.

$$h(\mathbf{x}_k|\bar{\mathbf{c}}_{k:T}) \leq h(\mathbf{x}_k|\bar{\mathbf{c}}_k), \ \forall \mathbf{c}_k \in S_{\mathbf{z}_k}, \ldots, \mathbf{c}_T \in S_{\mathbf{z}_T} \quad (44)$$

■

In summary, a fixed-lag smoothing approach for processing of external knowledge can only reduce the uncertainty with respect to *classical single-step constrained filtering*, in the sense of the entropy of the posterior PDF.

V. Simulation Results

The improvements in estimating the posterior PDF of a target will be demonstrated in a challenging Radar based surveillance scenario with low Signal-to-Noise Ratio (SNR). Classical systems are based on processing of *plot measurements*, i.e., point measurements obtained using a detection procedure. When the SNR is low, *missed detections* may happen, this way increasing the probability of track loss to the point that a target might not be detected. This is obviously unacceptable for defense applications. To overcame such problem, the Track-Before-Detect (TBD) paradigm was proposed in order to perform tracking directly on the basis of the raw power reflections collected by the Radar at each scan. The idea is to integrate over time the information contained in the measurements, this way delaying the moment at which the target is declared as *detected*, while sensibly reducing the probability of not detecting a target [15].

The system state vector is chosen as $\mathbf{X}_k = [\ \mathbf{x}_k^\intercal \quad \rho_k\ ]^\intercal$, where $\mathbf{x}_k \in \mathbb{R}^4$ contains the position and velocity vectors in 2D Cartesian coordinates, and $\rho_k \in \mathbb{R}$ is the unknown modulo of the target complex amplitude. Additional knowledge is available in terms of constraints on the target dynamics, i.e.,

$$\mathcal{C}_k \triangleq \left\{ \mathbf{X}_k \in \mathbb{R}^5 : \ \mathbf{x}_k \in \mathfrak{P}_k, \sqrt{\dot{x}_k^2 + \dot{y}_k^2} \leq v_{max} \right\} \quad (45)$$

where $\mathfrak{P}_k$ is a polygon representing the knowledge on the canal. Thus, the knowledge-based likelihood $p(\mathbf{c}_k|\mathbf{X}_k)$, to be used in both the *Pseudo-Measurements PF* and the *KB Fixed-Lag Smoother*, is defined straightforwardly, i.e.,

$$p(\mathbf{c}_k|\mathbf{X}_k) = \begin{cases} 1, & \text{if } \mathbf{x}_k \in \mathcal{C}_k \\ 0, & \text{otherwise} \end{cases} \quad (46)$$

A Nearly Constant Velocity (NCV) model is used to describe the system dynamics, and a Gaussian random walk models the fluctuations of the target complex amplitude, i.e.,

$$\mathbf{X}_{k+1} = F\ \mathbf{X}_k + \mathbf{v}_k, \quad \mathbf{v}_k \sim \mathcal{N}(0; Q) \quad (47)$$

where $\mathbf{v}_k$ is a zero-mean Gaussian process noise, T_s the Radar sampling time, and the matrices are defined as follows:

$$F = \text{diag}(F_1, F_1, 1), \quad F_1 = \begin{bmatrix} 1 & T_s \\ 0 & 1 \end{bmatrix} \quad (48)$$

$$Q = \text{diag}(a_x Q_1, a_y Q_1, a_\rho T_s), \quad Q_1 = \begin{bmatrix} \frac{T_s^3}{3} & \frac{T_s^2}{2} \\ \frac{T_s^2}{2} & T_s \end{bmatrix} \quad (49)$$

where a_x and a_y determine the process noise intensity for the Cartesian motion, and a_ρ describes the amplitude fluctuation.

Table I
Parameters Used in Simulation

Parameter	Symbol	Value
Signal to Noise Ratio	SNR	$10dB$
Range Quant Size	R	$10m$
Doppler Bin Size	D	$2m/s$
Beam Spacing	B	$1\deg$
Radar Sampling Time	T_s	$\{4,1\}s$
Maximum Target Speed	v_{max}	$15m/s$
Filters Process Noise	$\tilde{a}_{max}$	$\{.1,1\}m/s^2$

A Radar positioned at the Cartesian origin collects a measurement $\mathbf{z}_k$ at every time instant k. Such measurement consists of $N_r \times N_d \times N_b$ power measurements $\mathbf{z}_k^{lmn}$, one for each range-Doppler-bearing Radar cell, i.e.,

$$\mathbf{z}_k^{lmn} = \left|\mathbf{z}_{\rho,k}^{lmn}\right|^2, \quad k \in \mathbb{N} \quad (50)$$

where $\mathbf{z}_{\rho,k}^{lmn}$ is the complex signal, i.e.,

$$\mathbf{z}_{\rho,k}^{lmn} = \rho_k \exp\{i\phi_k\} h^{lmn}(\mathbf{x}_k) + n_k^{lmn}, \quad \phi_k \in [0, 2\pi) \quad (51)$$

where n_k^{lmn} is complex Gaussian noise with variance $2\sigma_n^2$, i.e., the real and imaginary components are zero-mean Gaussian with variance σ_n^2. The term $h^{lmn}(\mathbf{x}_k)$ in eq. (51) is the target reflection form and is defined as:

$$\begin{aligned} & h^{lmn}(\mathbf{x}_k) \\ & \triangleq \exp\left\{ -\frac{(r_l - r_k)^2}{2R} - \frac{(d_m - d_k)^2}{2D} - \frac{(b_n - b_k)^2}{2B} \right\} \quad (52) \\ & l = 1, \ldots, N_r; \ m = 1, \ldots, N_d; \ n = 1, \ldots, N_b; \ k \in \mathbb{N} \end{aligned}$$

and the following equations are used to map the state space into the measurements space, i.e.,

$$r_k = \sqrt{x_k^2 + y_k^2} \quad (53)$$

$$d_k = \frac{x_k \dot{x}_k + y_k \dot{y}_k}{r_k} \quad (54)$$

$$b_k = \arctan\left(\frac{y_k}{x_k}\right) \quad (55)$$

Notice that the reflection form in (52) describes the intensity of the target signal in the Radar cells surrounding the target. The parameters R, D, and B are related to the size of a range, doppler, and bearing cell, respectively. All the important parameters used in simulation are reported in Table I.

We consider a simplified version of the Track-Before-Detect (TBD) problem, where a single-target moves within the surveillance area and never leaves the sensor Field-of-View (FoV). In particular, we focus on the two slightly different scenarios depicted in fig. 1. Both scenarios model the realistic case of a ship traveling within a canal at a constant speed: *(Case 1)* is a straight line trajectory and is tracked using a sampling time $T_s = 4\,\text{sec}$, while *(Case 2)* is a slowly maneuvering trajectory ($a_{max} = 0.5m/\sec^2$) and is tracked using a sampling time $T_s = 1\,\text{sec}$.

We tested three PF algorithms and compared the proposed *KB Fixed-Lag Smoother* versus the standard *Unconstrained SISR-PF*, which does not exploit the additional knowledge, and

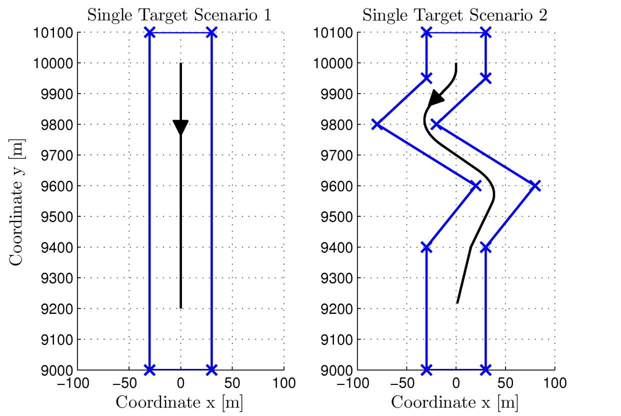

Figure 1. Scenarios considered in simulation: *(1)* straight line trajectory; *(2)* slowly maneuvering target ($a_{max} = 0.5m/\sec^2$). Both cases model a ship traveling within a canal (*shipping lane*) at a constant speed of $10m/s$.

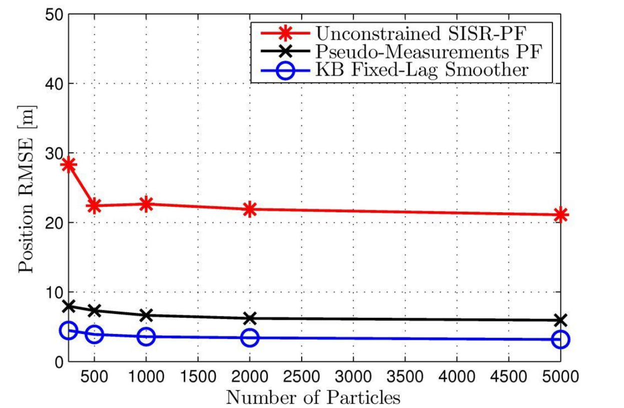

Figure 2. Time averaged Position RMSE for a varying number of particles. All filters use the same number of particles, and the *KB Fixed-Lag Smoother* is implemented through extraction of the ancestry tree ($L = 4$).

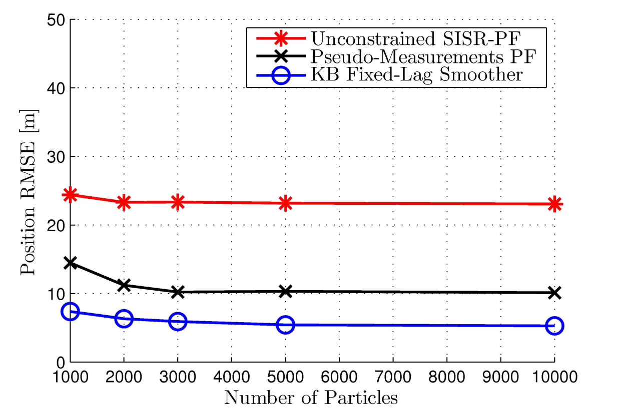

Figure 3. Time averaged Position RMSE for a varying number of particles. The *KB Fixed-Lag Smoother* is implemented through the *forward filtering backward smoothing* recursion, and uses 10% of the particles ($L = 3$).

the *Pseudo-Measurements PF*, which exploits the knowledge on the constraints at time k. Notice that for both scenarios, the extraction of the conditional mean is correct, as well as the use of the *position Root Mean Squared Error (RMSE)* for performance comparisons, since there cannot be multi-modalities in the posterior PDF and the target trajectory is symmetric with respect to the constraints. If one of this conditions is not verified, a MAP estimator should be used to extract estimates, and the Kullback-Leibler Divergence with respect to the true posterior PDF should be used for performance comparisons.

In fig. 4 we report the empirical posterior PDF and the conditional mean for the three filters over a single trial for *(Case 1)*. Notice that for this scenario, the *KB Fixed-Lag Smoother* is implemented following the *trajectory-based* approach to smoothing, i.e., through extraction of the ancestry tree. It is immediate to verify that the posterior PDF is less uncertain and pickier than the one obtained by classical *constrained filtering*. This justifies the improvements in terms of average *Position RMSE* reported in fig. (2). The same performance metrics are evaluated for *(Case 2)*, and reported in figs. (3) and (5). Here the *KB Fixed-Lag Smoother* is implemented through the *forward filtering backward smoothing* recursion. Notice that in order to limit the computational burden, the proposed filter only uses 10% of the particles, i.e., the x-axis values in fig. (3) have to be scaled down by a factor 10 when checking the performance of the *KB Fixed-Lag Smoother*.

VI. Conclusions

We propose a new method for Bayes optimal processing of possibly available external knowledge that can be formalized in terms of *hard constraints* on the system dynamics. The new algorithm is based on the idea of using a *fixed-lag smoother* to perform forward processing of knowledge. That is, the one step ahead prediction of the particles cloud at time k is such that most of the particle trajectories verify the constraints at successive time instants.

We formally show that in this way the uncertainty in the posterior PDF can only reduce in terms of the entropy, allowing for sensible improvements in tracking performance. Preliminary simulations show that such improvements can be of great interest when tracking targets with low SNR in the Track-Before-Detect (TBD) framework.

Future research will consider in-depth parametric analyses with respect to the *fixed-lag* L. We expect a dependence between the attainable improvements by increasing L and the minimal number of particles. Furthermore, we will investigate more efficient implementations of the proposed algorithm by means of the two-filter formula for smoothing and N-body methods to approximate the sum-kernel problem.

VII. Acknowledgment

The research leading to these results has received funding from the EU's Seventh Framework Programme under grant agreement n^o 238710. The research has been carried out in the MC IMPULSE project: https://mcimpulse.isy.liu.se

References

[1] T. Cover and J. Thomas, *Elements of Information Theory*. John Wiley & Sons, Inc., Hoboken, New Jersey., 2006.

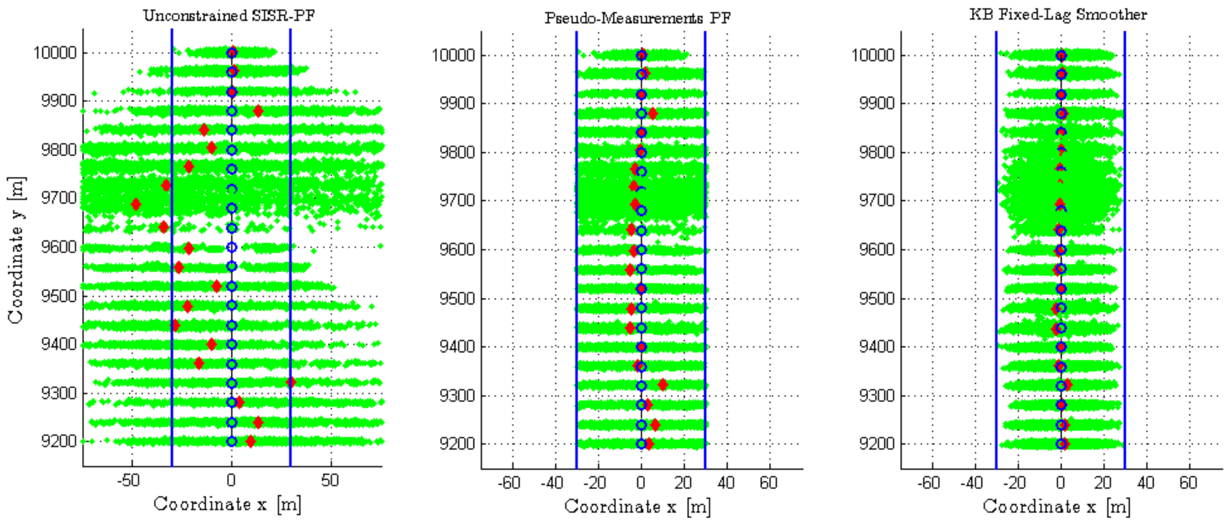

Figure 4. Empirical posterior distribution and conditional mean over a single trial. Each filter uses 10000 particles, and the *KB Fixed-Lag Smoother* is implemented through extraction of the ancestry tree (*trajectory-based*). It is immediate to verify that the proposed method reduces the intrinsic uncertainty in the posterior PDF when compared to classical *single-scan constrained filtering*. This justifies the improvements in terms of *Position RMSE*.

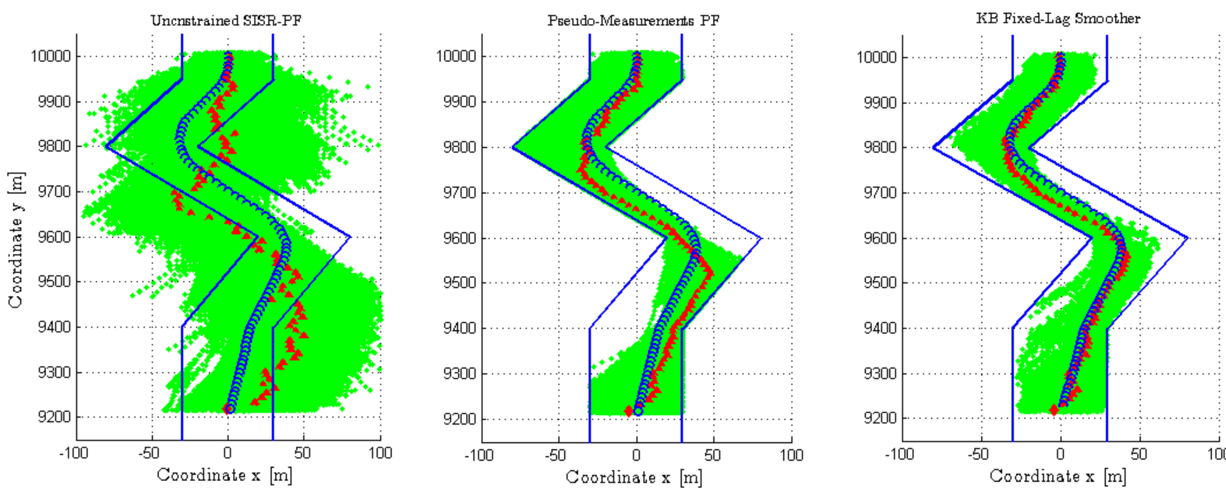

Figure 5. Empirical posterior distribution and conditional mean over a single trial. The *KB Fixed-Lag Smoother* is implemented through the *forward filtering backward smoothing* recursion, and the processing is performed using 1000 particles. The *Unconstrained SISR-PF* and the *Pseudo-Measurements PF* perform processing using 10000 particles. The *blue circles* represent the true target position, the *green dots* the particles, and the *red diamonds* the conditional mean.

[2] F. Gustafsson, "Particle filter theory and practice with positioning applications," *IEEE Transaction on Aerospace and Electornic Systems*, vol. 25(7), pp. 53–81, 2010.

[3] A. Doucet and D. Crisan, "A survey of convergence results on particle filtering methods," *IEEE Transactions on Singal Processing*, vol. 50(3), pp. 736–746, 2002.

[4] I. Kyriakides, D. Morrell, and A. Papandreou-Suppappola, "Sequential monte carlo methods for tracking multiple targets with deterministic and stochastic constraints," *IEEE Transactions on signal processing*, vol. 56(3), pp. 937–948, 2008.

[5] W. Koch and M. Ulmke, "Road-map assisted ground moving target tracking," *IEEE Transaction on Aerospace and Electornic Systems*, vol. 42(4), pp. 1264–1274, 2006.

[6] S. Challa and N. Bergman, "Target tracking incorporating flight envelope information," *Proc. of the 3th Int. Conf. on Information Fusion, FUSION 2000*, vol. 2, pp. THC2/22–THC2/27, Paris, France, July 2000.

[7] F. Papi, M. Podt, Y. Boers, G. Battistello, and M. Ulmke, "Bayes optimal knowledge exploitation for hard-constrained target tracking," *Proc. of the 9th IET Data Fusion & Target Tracking Conference*, London, UK 2012.

[8] O. Cappé, E. Moulines, and T. Ryden, *Inference in Hidden Markov Models*. Springer Series in Statistics, 2005.

[9] N. Gordon, D. Salmond, and A. Smith, "Novel approach to nonlinear/non-gaussian bayesian state estimation," *IEE Proceedings F on Radar and Signal Processing*, vol. 140(2), p. 107–113, 1993.

[10] B. Anderson and J. Moore, *Optimal Filtering*. Prentice-Hall, 1979.

[11] L. Baum, T. Petrie, G. Soules, and N. Weiss, "A maximization technique occurring in the statistical analysis of probabilistic functions of markov chains," *Ann. Math. Statist.*, vol. 41, p. 164–171, 1970.

[12] S. J. Godsill, A. Doucet, and M. West, "Monte carlo smoothing for nonlinear time series," *Journal of the American Statistical Association*, vol. 99, pp. 156–168, 2004.

[13] M. Klaas, M. Briers, N. de Freitas, A. Doucet, S. Maskell, and D. Lang, "Fast particle smoothing: If i had a millino particles," *Proc. of the* $23^r d$ *International Conference on Machine Learning*, Pittsburg, PA, USA 2006.

[14] J. Williams, "Information theoretic sensor management," Ph.D. dissertation, Massachusetts Institute of Technology, 2007.

[15] Y. Boers and J. Driessen, "Multitarget particle filter track before detect application," *IEE Proceedings on Radar, Sonar and Navigation*, vol. 151, pp. 351–357, 2004.