Saturated Particle Filter: Almost Sure Convergence and Improved Resampling

***Abstract*— Nonlinear stochastic dynamical systems are widely used to model physical processes. In many practical applications, the state variables are defined on a compact set of the state space, i.e., they are bounded or saturated. To estimate the states of systems with saturated variables, the Saturated Particle Filter (SPF) has recently been developed. This filter exploits the structure of the saturated system using a specific importance sampling distribution. In this paper we investigate the asymptotic properties of the filter, in particular its almost sure convergence to the true posterior PDF. Furthermore, an improved SPF is developed that uses a novel resampling procedure to overcome the practical shortcomings of the original SPF. We prove that this new filter also converges almost surely to the true posterior PDF. Both versions of the SPF are presented in easy to implement algorithmic forms.**

I. INTRODUCTION

Various practical problems require extracting information of interest from an uncertain or dynamically changing environment. Such problems are often represented in a state space form as Stochastic Dynamical Systems (SDS). In the SDS framework the required information is modeled as an unknown state or parameter of the system that needs to be estimated from available measurements. If the unknown property changes as the system evolves, the estimate has to be updated each time a measurement becomes available. For this purpose the Bayesian dynamic filter has been developed. The Bayesian dynamic filter is a recursive algorithm that uses prior knowledge of the system and the most current measurement to compute the posterior estimate. This is achieved by utilizing the Bayes theorem (Chapter 1 of [1]). The posterior estimate is a probability density function (PDF) of the state of the system.

The posterior PDF can be analytically computed only for systems with linear dynamics and additive Gaussian noises [2], [3], i.e., for a very restricted class of systems. In case of nonlinear or/and non-Gaussian noises, in general it is impossible to compute the exact form of the posterior PDF. Therefore, one needs to rely on approximations. Two types of approximations can be distinguished: *parametric* and *non-parametric*. The posterior PDF is parametrically approximated if it can be characterized by a function that depends on a finite number of parameters. In other words, the approximation of the posterior PDF belongs to a finitely-parameterized family of functions, e.g., Gaussians [4], [5], [6], [7], sum of Gaussians [8], [9], [10], Pearson [11]. If such a parameterization is impossible, the posterior PDF can be approximated with a non-parametric PDF of an arbitrary shape [12], [13], [14]. Among the non-parametric methods, the Particle Filter (PF) has gained the most popularity. The PF algorithm is based on the Monte Carlo method originally proposed by Ulam and von Neuman (see [15] and the references therein).

The properties of the PF have been extensively studied in recent years [12], [16], [17], and many versions of PF have been developed for specific types of problems [16], [18], [19], [20], [21], [22]. In particular, the Saturated Particle Filter (SPF) has been proposed in [23] as a method to estimate the states of a Saturated Stochastic Dynamical System (SSDS). A system belongs to the SSDS class if at each time step k for at least one of the state variables the probability that it saturates is strictly positive [23]. The point that belongs to the boundary of such a set is called the saturation point. These systems are frequently found both in industrial applications [24], [25], and in theoretical research [20], [26].

Under certain conditions, the states of the SSDS can be estimated by a parametric Unscented Kalman Filter [27], [28], [29]. Nevertheless, the flexibility of the nonparametric PF makes it a more suitable algorithm for the severely nonlinear SSDS [30], [20], [31], [23]. On the downside, however, the PF is based on the Monte Carlo approximation, hence it might require a large number of samples to achieve an accurate estimate. This makes the algorithm computationally expensive, and consequently, limits its on-line applicability. The choice of the importance sampling density is a crucial step toward reducing the computational costs, and therefore making the filter feasible for on-line applications. This has been an active field of research in recent years. In [20], [31] the constrained PF has been proposed that produces a state estimate that does not violate the physical constraints of the system. This is done by discarding unsuitable particles [31], or by projecting them on a constraint set [20]. The SPF combines the projection approach with a novel sampling method to effectively detect the saturation moment, and forces the particles to rapidly jump to the part of the state space close to the saturation point [23]. This makes the SPF a powerful tool for solving the estimation problem for SSDS.

Since the PF approximates the true posterior PDF by a set of N discrete samples the question that naturally arises is whether the approximation converges to the true posterior PDF as $N \to \infty$, and if yes in what sense? Extensive studies on the convergence properties of the PF have been conducted in [1], [17] and [32]. In both [1] and [17] two types of convergence have been discussed, 1) almost sure convergence and 2) convergence in the mean square error sense, and conditions that guarantee either type of convergence have been derived. The focus of [32] is more

on investigating relations between the sample size N and the time step k. Moreover, [32] presents a number of interesting results regarding asymptotic behavior of the variance of the estimator.

In this paper we consider the almost sure convergence of the SPF to the true posterior PDF. First, we derive sufficient conditions for the almost sure convergence of the SPF defined in [23]. Next, we propose an improved version of the SPF with a novel resampling procedure and we prove that it also converges almost surely to the true posterior PDF. In both cases we discuss the practical meaning of the constraints that ensure the filter's convergence properties. Furthermore, both algorithms are presented in an easy-to-implement algorithmic form.

The paper is structured as follows. In Section II we present background information regarding the PF, the SPF, and the stochastic systems. In Section III we discuss asymptotic properties of the originally proposed SPF. The improved version of the SPF is derived in Section IV, where we also prove the almost sure convergence of the new SPF to the true posterior PDF and discuss the practical properties of the improved SPF. Section V concludes the paper.

II. PRELIMINARIES

In this section we give an overview of the basic definitions and properties concerning the asymptotic behavior of Particle Filters (PF). This section presents background material, therefore the theorems are given without proofs. For detailed proofs, see [17], [33].

A. Particle Filter

Definition 1 (Stochastic Dynamical System): The *Stochastic Dynamical System* (SDS) is a process defined recursively by

$$x_{k+1} = f_k\left(x_k, w_k\right), \tag{1}$$

$$y_k = h_k\left(x_k, v_k\right), \tag{2}$$

$$x_0 \sim p_0\left(\cdot\right), \tag{3}$$

where w_k and v_k are mutually independent random variables, f_k is a (possibly nonlinear) function that describes the state evolution, h_k is a (possibly nonlinear) function that establishes the observation model, and p_0 is a *probability density function* (PDF) of the initial state x_0.

From the model (1)–(2) we derive the *transition probability kernel* $K_{k-1}(x_k|x_{k-1})$ defined by

$$K_{k-1}(x_k|x_{k-1}) := \mathbb{P}_w\left(f_{k-1}\left(x_{k-1}, w_{k-1}\right) = x_k\right), \tag{4}$$

i.e., the conditional PDF of the variable x_k given the previous state x_{k-1}, and the *likelihood function* $g_k(y_k|x_k)$ defined by

$$g_k(y_k|x_k) := \mathbb{P}_v\left(h_k\left(x_k, v_k\right) = y_k\right), \tag{5}$$

i.e., the conditional PDF of the variable y_k given the current state x_k.

The PF represents the posterior PDF $\pi_{k|k}$ of the state x_k by N random samples (particles) $\left\{x_k^i\right\}_{i=1}^N$ with their associated weights $\left\{\omega_k^i\right\}_{i=1}^N$, normalized so that $\sum_{i=1}^N \omega^i = 1$. At time instant k, the previous posterior PDF $\pi_{k-1|k-1}$ is represented by N samples $\left\{x_{k-1}^i\right\}_{i=1}^N$ and the corresponding weights $\left\{\omega_{k-1}^i\right\}_{i=1}^N$. To approximate the true posterior $\pi_{k|k}$, new samples $\left\{x_k^i\right\}_{i=1}^N$ and weights $\left\{\omega_k^i\right\}_{i=1}^N$ are generated using the *Sequential Importance Sampling* (SIS) method [34], [16]. The SIS method is a recursive algorithm that uses the most recent observation y_k to compute $\left\{\left(x_k^i, \omega_k^i\right)\right\}_{i=1}^N$ in two steps. First, for every $i = 1, ..., N$, sample x_k^i is drawn from a (chosen) *importance kernel* $\tilde{K}_{k-1}(x_k^i|x_{k-1}^i, y_k)$. Next, using the most recent observation y_k, the weights ω_k^i are updated according to the Bayes rule

$$\tilde{\omega}_k^i = \omega_{k-1}^i \frac{g_k(y_k|x_k^i)\, K_{k-1}(x_k^i|x_{k-1}^i)}{\tilde{K}_{k-1}(x_k^i|x_{k-1}^i, y_k)} \tag{6}$$

and normalized

$$\omega_k^i = \frac{\tilde{\omega}_k^i}{\sum_{j=1}^N \tilde{\omega}_k^j}\,. \tag{7}$$

The posterior PDF $\pi_{k|k}$ is represented by the set of weighted samples, conventionally denoted by:

$$\pi_{k|k} \approx \pi_{k|k}^N := \sum_{i=1}^N \omega_k^i \delta(x_k - x_k^i), \tag{8}$$

where δ denotes the Dirac delta at zero.

In theory, the best possible importance density is the posterior PDF $\pi_{k|k}$ itself. For every other choice of the importance density the variance of the weights $\left\{\omega_k^i\right\}_{i=1}^N$ increases over time [34]. Since it is generally impossible to sample from the posterior PDF $\pi_{k|k}$ one needs to rely on suboptimal importance densities in practice. It has been shown (see [34] and the references therein) that the importance density that minimizes the variance of the weights $\left\{\omega_k^i\right\}_{i=1}^N$ conditional upon simulated trajectories $\left\{x_j^i\right\}_{j=1}^{k-1}$ and the observations $\left\{y_j\right\}_{j=1}^k$ is equal to $\mathbb{P}\left(x_k|x_{k-1}^i, y_k\right)$, i.e., the PDF of the state x_k conditional upon immediately preceding realization x_{k-1}^i of the simulated trajectory and the most current observation y_k. $\mathbb{P}\left(x_k|x_{k-1}^i, y_k\right)$ is an optimal importance density in the aforementioned sense. Unfortunately, apart from a restricted class of systems [34], [35], [36], sampling from $\mathbb{P}\left(x_k|x_{k-1}^i, y_k\right)$ is practically impossible. Therefore, various suboptimal importance densities have been proposed during the recent years. The simplest choice is to use an importance density that is fixed over the time [37]. A more popular approach is utilized in the *bootstrap particle filter* (BPF) [38], [39]. The BPF samples points $\left\{w_{k-1}^i\right\}_{i=1}^N$ from the noise distribution of w_{k-1}, then propagates them together with the particles $\left\{x_{k-1}^i\right\}_{i=1}^N$, which approximate the previous posterior PDF $\pi_{k-1|k-1}$, through the prediction model (1) to obtain the particles $x_k^i := f_{k-1}\left(x_{k-1}^i, w_{k-1}^i\right)$ that approximate the predicted PDF $\pi_{k|k-1}$. A slightly different formulation of the BPF, with particles $\left\{x_{k-1}^i\right\}_{i=1}^N$ sampled directly from the transition probability $\mathbb{P}\left(x_k|x_{k-1}^i\right)$, is presented in [40]. Another variation of the BPF [17]

allows sampling particles $\left\{x_{k-1}^i\right\}_{i=1}^N$ from the weighted transition probability $\frac{1}{N}\sum_{j=1}^N \mathbb{P}\left(x_k|x_{k-1}^j\right)$. More information on other types of importance densities can be found in [32], [41].

A common problem of PF is the particle degeneracy: after several iterations, all but few particles will have negligible weights. This does not come as a surprise since the variance of the weights $\left\{\omega_k^i\right\}_{i=1}^N$ can only increase over time. When that occurs most of the computational power is wasted on updating negligible weights and the accuracy of the algorithm strongly deteriorates since the posterior PDF $\pi_{k|k}$ is approximated only by a small set of significant particles. The degeneracy phenomenon can be circumvented by monitoring the weights and resampling the particles, e.g., with Algorithm 1, after the degeneracy is detected [12], [42], [43], [44]. A common measure of the degeneracy is the *effective sample size* N_{eff}, computed by [12], [45]:

$$N_{\text{eff}} = \frac{1}{\sum_{i=1}^{N}(\omega_k^i)^2}. \tag{9}$$

Alternatively one can test a Kullback-Leibler distance between the sets of weights obtained in the consecutive iterations [43], or simply measure the maximum weight at each iteration before resampling [43], [42].

Algorithm 1 Resampling

Output: $\{(x_{\text{new}}^i, \omega_{\text{new}}^i)\}_{i=1}^N$
  for $i = 1, 2, \ldots, N$ **do**
    Compute cumulative sum of weights: $\omega_c^i = \sum_{j=1}^i \omega_k^j$
  end for
  Draw u_1 from $\mathcal{U}(0, \frac{1}{N})$
  for $i = 1, 2, \ldots, N$ **do**
    Find x^{+i}, the first sample such that $\omega_c^i \geq u_i$.
    Replace particle i: $x_{\text{new}}^i = x^{+i}$, $\omega_{\text{new}}^i = \frac{1}{N}$
    $u_{i+1} = u_i + \frac{1}{N}$
  end for

The PF that monitors the degeneracy using N_{eff} is summarized in Algorithm 2.

B. Saturated Particle Filter

We start with a definition of the class of systems we analyze. The goal of this paper is to investigate the asymptotic properties of the SPF with a new resampling algorithm. Thus, to make the argument simpler, we restrict our considerations to the SPF defined on one dimensional SDS. To properly extend the SPF method to the higher-dimensional systems we have to identify right conditions that need to be imposed on the dynamical system (1)–(3), e.g., the constraints of the state variables need to be convex. These are subject of ongoing research that is beyond the scope of this paper. However, the general concepts of the SPF, new resampling and the asymptotic properties of the algorithm are well captured by the one-dimensional case hereby discussed. Another assumption, which greatly simplifies the calculations that follow, is

Algorithm 2 Particle filter

Input: $K_{k-1}(x_k|x_{k-1})$, $\tilde{K}_{k-1}(x_k|x_{k-1}, y_k)$, $g_k(y_k|x_k)$, $p_0(x_0)$, N, N_{T}
  Initialize:
  for $i = 1, 2, \ldots, N$ **do**
    Draw a new particle: $x_0^i \sim p(x_0)$
    Assign weight: $\omega_0^i = \frac{1}{N}$
  end for
  At every time step $k = 1, 2, 3, \ldots$
  for $i = 1, 2, \ldots, N$ **do**
    Draw particle from importance distribution: $x_k^i \sim \tilde{K}_{k-1}(x_k^i|x_{k-1}^i, y_k)$
    Use measured y_k to update the weight: $\tilde{\omega}_k^i = \omega_{k-1}^i \frac{g_k(y_k|x_k^i)\, K_{k-1}(x_k^i|x_{k-1}^i)}{\tilde{K}_{k-1}(x_k^i|x_{k-1}^i, y_k)}$
  end for
  Normalize weights: $\omega_k^i = \frac{\tilde{\omega}_k^i}{\sum_{j=1}^N \tilde{\omega}_k^j}$
  if $\frac{1}{\sum_{i=1}^N (\omega_k^i)^2} < N_{\text{T}}$ **then**
    Resample using Algorithm 1.
  end if

considering only systems bounded on the positive real line. This condition is of purely technical nature as any variable bounded in $\mathbb{R}$, by an affine change of variables, can be transformed into a variable constrained in $\mathbb{R}_+$.

Definition 2 (Saturated Stochastic Dynamical System): Let us consider a real-valued SDS with the observation model given by

$$y_k = h_k\left(x_k\right) + v_k, \tag{10}$$

where $\mathbb{E}v_k = 0$. Such SDS is called *Saturated Stochastic Dynamical System* (SSDS) if there exists a function $C : \mathbb{R}_+ \to \mathbb{R}_+$ and a function $F_k : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}_+$ such that for each $k \geq 1$, the state equation takes the form:

$$x_{k+1} = \min\left(F_k\left(x_k, w_k\right), C\left(x_k\right)\right). \tag{11}$$

In the above definition the function F_k corresponds to the transition function of the "unsaturated" system, i.e., the system with the state that is not bounded by the function C. An example of SSDS is discussed in Section IV-D.3.

In order to online estimate the states of the SSDS the *Saturated Particle Filter* (SPF) has been developed [23]. The SPF is a SIS-type of algorithm that samples particles $\left\{x_{k+1}^i\right\}_{i=1}^N$ from a special importance kernel $\tilde{K}\left(\cdot|x_k^i, y_{k+1}\right)$ that is chosen in such a way that the distribution of these particles is "closer" to the posterior PDF $\pi_{k+1|k+1}$ than the distribution of the particles obtained by the BPF, i.e., by sampling from the transition probability kernel $K\left(\cdot|x_k^i\right)$. The SPF's improved estimation performance comes from the use of the *detection function* whose purpose is to quickly detect whether the saturation occurred by comparing the measurements with the state constraints. In other words, the detection function is intended as a "pseudometric" between y_{k+1} and

$h_{k+1}\left(C\left(x_k^i\right)\right)$. This extra information is used to force the particles to move to the "appropriate" region already in the sampling step of the algorithm. In general, the choice of detection function depends on the dynamics of the system under consideration. For the class of systems presented in Definition 2 let us introduce:

Definition 3 (Detection function): Let $\alpha : \mathbb{R} \to \mathbb{R}$ be a function for which the following conditions are satisfied:

1) α is non-decreasing,
2) there exists $y_0 \in \mathbb{R}$ such that $\alpha(y_0) = 0$.

Then, the mapping

$$\left(y_{k+1}, x_k^i\right) \mapsto \alpha\left(y_{k+1} - h_{k+1}\left(C\left(x_k^i\right)\right)\right) \tag{12}$$

is called a *detection function*.

This is not the only possible way of defining detection functions. Nevertheless, throughout this paper we stick to this definition because for the one dimensional SSDS it illustrates well the idea behind the SPF method.

Since there is one to one relation between the detection function defined by (12) and α, in what follows, we do not make a distinction between these two objects.

With the use of the detection function α we define the probabilities of saturation [23]:

Definition 4 (Probabilities of saturation): For every $i = 1, ..., N$ the *predicted probability of saturation* q_i is given by

$$q_i = \int_{C\left(x_k^i\right)}^{+\infty} \mathbb{P}\left(F_k\left(x_k, w_k\right) = z | x_k^i\right) dz, \tag{13}$$

and the *updated probability of saturation* q_i^α is given by

$$q_i^\alpha := \begin{cases} 1 \quad \textbf{if} \quad q_i + \alpha\left(y_{k+1} - h_{k+1}\left(C\left(x_k^i\right)\right)\right) > 1, \\ 0 \quad \textbf{if} \quad q_i + \alpha\left(y_{k+1} - h_{k+1}\left(C\left(x_k^i\right)\right)\right) < 0, \\ q_i + \alpha\left(y_{k+1} - h_{k+1}\left(C\left(x_k^i\right)\right)\right) \quad \textbf{otherwise}. \end{cases} \tag{14}$$

Using q_i^α defined in (14), and the detection function α, we define the importance density $\tilde{K}_k$ of the SPF by:

$$\begin{aligned} &\tilde{K}_k\left(x | x_k^i, y_{k+1}\right) := q_i^\alpha \delta\left(C\left(x_k^i\right) - x\right) \\ &+ \frac{1 - q_i^\alpha}{1 - q_i} \mathbb{P}\left(F_k\left(x_k, w_k\right) = x | x_k^i\right) \mathbf{1}_{\left[0, C\left(x_k^i\right)\right)}(x), \end{aligned} \tag{15}$$

where δ denotes the Dirac delta function, and $\mathbf{1}_{\left[0, C\left(x_k^i\right)\right)}$ is an indicator function on the interval $\left[0, C\left(x_k^i\right)\right)$.

It can be easily seen that $\tilde{K}_k$ defines a probability kernel. The importance density of the BPF filter is a special case of $\tilde{K}_k$ with $\alpha \equiv 0$.

Given the particle x_k^i, a new particle x_{k+1}^i is drawn from the importance density $\tilde{K}_k\left(\cdot | x_k^i, y_{k+1}\right)$. By (15) the particle x_{k+1}^i saturates, i.e., $x_{k+1}^i = C\left(x_k^i\right)$, with the probability q_i^α, and with probability $1 - q_i^\alpha$ it is drawn from

$$\frac{1}{1 - q_i} \mathbb{P}\left(F_k\left(x_k, w_k\right) = x | x_k^i\right) \mathbf{1}_{\left[0, C\left(x_k^i\right)\right)}(x) \tag{16}$$

The associated weights ω_{k+1}^i are computed using (6). If x_{k+1}^i saturates, i.e., $x_{k+1}^i = C\left(x_k^i\right)$, then, by the definitions of q_i, and q_i^α, the weight ω_{k+1}^i follows the formula:

$$\omega_{k+1}^i \propto \omega_k^i \frac{q_i}{q_i^\alpha} g_{k+1}\left(y_{k+1} | x_{k+1}^i\right), \tag{17}$$

if x_{k+1}^i does not saturate, the weight ω_{k+1}^i is updated by:

$$\omega_{k+1}^i \propto \omega_k^i \frac{1 - q_i}{1 - q_i^\alpha} g_{k+1}\left(y_{k+1} | x_{k+1}^i\right). \tag{18}$$

The SPF algorithm, developed in [23], is summarized in Algorithm 3. Note that the updated probability of saturation depends on the choice of the detection function α. Therefore, through q_i^α, the SPF also depends on α.

Algorithm 3 Saturated PF for a given α

Input: $\left\{\left(x_k^i, \omega_k^i\right)\right\}_{i=1}^N, y_{k+1}$
Output: $\left\{\left(x_{k+1}^i, \omega_{k+1}^i\right)\right\}_{i=1}^N$
for $i = 1, 2, \ldots, N$ **do**
  Compute the probability q_i according to (13)
  Compute the probability q_i^α according to (14)
  Draw $u \sim \mathcal{U}(0, 1)$
  if $u \leq q_i^\alpha$ **then**
    Particle x_{k+1}^i saturates:

$$\begin{aligned} x_{k+1}^i &:= C\left(x_k^i\right) \\ \omega_{k+1}^i &\propto \omega_k^i \frac{q_i}{q_i^\alpha} g_{k+1}\left(y_{k+1} | x_{k+1}^i\right) \end{aligned}$$

  else
    Particle x_{k+1}^i does not saturate:

$$\begin{aligned} x_{k+1}^i &\sim \frac{1 - q_i^\alpha}{1 - q_i} \mathbb{P}\left(F_k\left(x_k, w_k\right) = \bullet | x_k^i\right) \mathbf{1}_{\left[0, C\left(x_k^i\right)\right)}(\bullet) \\ \omega_{k+1}^i &\propto \omega_k^i \frac{1 - q_i}{1 - q_i^\alpha} g_{k+1}\left(y_{k+1} | x_{k+1}^i\right) \end{aligned}$$

  end if
end for

Up to this point we have defined the framework of SSDS and we have formulated the SPF algorithm solving the filtering problem for SSDS. More details concerning these topics can be found in [23]. Next, we present several concepts from stochastic systems theory.

C. Stochastic Systems Theory

We start this section with giving the basic definitions that are further used to establish the asymptotic properties of the SPF. Then, we recall the theorems dealing with the convergence properties of the Particle Filters. A comprehensive overview of the presented topics can be found in [17], [33], and [1].

Definition 5 (Feller kernel): The transition probability kernel $K(-|\cdot)$ on $(X, \Sigma_X, \mathbb{P})$ has the *Feller property*[1] if, for every continuous and bounded function φ, the function

$$z \to \int_X \varphi(x) K(dx|z) \tag{19}$$

[1]Weak Feller property by the definition of [33].

is continuous and bounded [1].

In the following, we show that the PF can be defined as an operator on the space of probabilistic measures $\mathcal{P}(X)$, where X is a given vector space. The construction of such an abstract operator requires some extra effort, but it allows us to derive simple conditions that guarantee good asymptotic behavior of a generic PF.

We start by introducing the operators [17] that will be used in the proof of convergence.

Let K_k be a transition probability kernel on the probability space $(X, \Sigma_X, \mathbb{P})$ defined by (4) and let $\tilde{K}_k$ be an arbitrary probability kernel that is absolutely continuous with respect to K_k. Furthermore, let g_k be a likelihood function defined by (5) and let ω_k be a *weighted likelihood function* defined by:

$$\omega_k(y, x|x_k) := \frac{g_{k+1}(y|x) K_k(x|x_k)}{\tilde{K}_k(x|x_k, y)}. \tag{20}$$

Definition 6 (Prediction operator): The *prediction operator* b_k maps the probabilistic measure $\nu \in \mathcal{P}(X)$ into the probabilistic measure $b_k(\nu) \in \mathcal{P}(X)$, defined by

$$[b_k(\nu)](A) := \int_X \tilde{K}_k(A|x_k, y_{k+1}) \nu(dx_k), \tag{21}$$

for every $A \in \Sigma_X$.

Definition 7 (Update operator): For given y_{k+1} and x_k, the *update operator* a_k maps the probabilistic measure $\nu \in \mathcal{P}(X)$ into the probabilistic measure $a_k(\nu) \in \mathcal{P}(X)$, defined by

$$\int_X \varphi(x) [a_k(\nu)](dx) := \left[\int_X \varphi(x) \omega_k(y_{k+1}, x|x_k) \nu(dx)\right] \times \left[\int_X \omega_k(y_{k+1}, x|x_k) \nu(dx)\right]^{-1} \tag{22}$$

for every continuous and bounded function φ.

Definition 8 (Multinomial sampling operator): The *multinomial sampling operator* c^N assigns to the probabilistic measure $\nu \in \mathcal{P}(X)$ its random discrete approximation $c^{N,x}(\nu)$ according to:

$$c^{N,x}(\nu) := \frac{1}{N} \sum_{j=1}^{N} \delta_{\{V_j(x)\}}, \tag{23}$$

where $N > 0$, $x \in X$, V_j, $j = 1, ..., N$ are i.i.d. random variables on X with the common distribution ν.

Equation (23) formally defines the empirical approximation of an arbitrary distribution ν by means of Monte Carlo sampling.

A special case of importance sampling, which is employed in the BPF, is when it is desirable to directly draw samples from the transition kernel $K_k(-|\cdot)$. In such a case $\tilde{K}_k(-|\cdot) := K_k(-|\cdot)$, and ω_k reduces to g_k.

Definition 9 (Particle Filter): Let $\bar{c}^N$ be the *resampling operator* on $\mathcal{P}(X)$ that maps the measure ν into a random measure $\bar{c}^{N,\cdot} \in \mathcal{P}(X)$ composed of N discrete random measures. The *Particle Filter* is an operator k_k^N that transforms the empirical measure $\pi_{k|k}^N$, which approximates the state of the system at time k, into the empirical measure $\pi_{k+1|k+1}^N$ at time $k+1$:

$$\pi_{k+1|k+1}^N = k_k^N\left(\pi_{k|k}^N\right) := \left[\bar{c}^N \circ a_k \circ c^N \circ b_k\right]\left(\pi_{k|k}^N\right). \tag{24}$$

The relation between this abstract definition and the standard formulation of the PF is the following:

1) *Prediction stage:* First, the predicted state density is computed by applying the operator b_k, defined in Definition 6, to the empirical measure $\pi_{k|k}^N$. Then the predicted state density is approximated by N random samples obtained by applying the sampling operator c^N (Definition 8) to $b_k\left(\pi_{k|k}^N\right)$.
2) *Update stage:* After the prediction stage, the updated state density is computed as the output of the update operator a_k (Definition 7) applied to $c^N \circ b_k\left(\pi_{k|k}^N\right)$. Finally, applying the operator $\bar{c}^N$ to the updated state density corresponds to the resampling step of the PF.

The asymptotic properties of the PF can be established from the following theorem:

Theorem 1 (Convergence of the generic PF [17]): Let us assume that for each k the importance kernel $\tilde{K}_k$ is Feller, and the likelihood function ω_k is bounded, continuous, and strictly positive. Furthermore, let $\bar{c}^N$ be a resampling operator such that for every bounded function φ, there exists a constant L such that:

$$\mathbb{E}\left[\left(\int_X \varphi(x)\left[\bar{c}^{N,\cdot}(\nu)\right](dx) - \int_X \varphi(x)\nu(dx)\right)^4\right] \leq \frac{L}{N^2}. \tag{25}$$

Then, as $N \to \infty$, the empirical measure $\pi_{k+1|k+1}^N$ defined by (24) converges almost surely towards the true posterior PDF $\pi_{k+1|k+1}$.

III. ASYMPTOTIC PROPERTIES OF THE SPF UNDER STANDARD RESAMPLING

In this section we investigate asymptotic properties of the SPF with respect to Theorem 1. First, we prove the theoretical convergence of the SPF. Second, we discuss the practical consequences of the conditions derived in the first part.

A. Theoretical results

We start with formulating the SPF algorithm in terms of the operator notation introduced in Section II-C.

Definition 10 (Saturated Particle Filter): Consider the SSDS defined by (1)-(3) and (11), and let α be an arbitrary detection function satisfying Definition 3. Furthermore, let K_k be the transition probability kernel and g_k be the likelihood function corresponding to the state model (1) and the observation model (2) respectively. The *Saturated*

Particle Filter (SPF) is a PF with the transition probability kernel $\tilde{K}_k$ defined by:

$$\tilde{K}_k\left(x|x_k, y_{k+1}\right) := q_{k+1}^{\alpha}\left(x_k, y_{k+1}\right)\delta_{\{C(x_k)\}}\left(x\right) + \frac{1-q_{k+1}^{\alpha}\left(x_k, y_{k+1}\right)}{1-q_{k+1}\left(x_k\right)}K_k\left(x|x_k\right), \quad (26)$$

and with the weighted likelihood function defined by

$$\tilde{\omega}_k\left(y, x|x_{k-1}\right) := g_k\left(y|x\right)\left(\frac{1-q_k\left(x_{k-1}\right)}{1-q_k^{\alpha}\left(x_{k-1}, y\right)}\mathbf{1}_{[0,C(x_{k-1}))}(x) + \frac{q_k\left(x_{k-1}\right)}{q_k^{\alpha}\left(x_{k-1}, y\right)}\delta_{C(x_{k-1})}(x)\right), \quad (27)$$

where the predicted probability of saturation q_k and the updated probability of saturation q_k^{α} are defined as:

$$q_k\left(x\right) := \int_{C(x)}^{+\infty} K_k\left(dz|x\right), \quad (28)$$

$$q_k^{\alpha}\left(x, y\right) := \begin{cases} 1 & \textbf{if} \quad q_k\left(x\right) + \alpha\left(y - h_{k+1}\left(C\left(x\right)\right)\right) > 1, \\ 0 & \textbf{if} \quad q_k\left(x\right) + \alpha\left(y - h_{k+1}\left(C\left(x\right)\right)\right) < 0, \\ q_k\left(x\right) + \alpha\left(y - h_{k+1}\left(C\left(x\right)\right)\right) & \textbf{otherwise}. \end{cases} \quad (29)$$

It is easy to see that Definition 10 is an abstraction of Algorithm 3.

In what follows we derive sufficient conditions that, if satisfied, ensure the almost sure convergence of the SPF to the true posterior PDF $\pi_{k+1|k+1}$. We start with two technical lemmas:

Lemma 1: Let K_k be a bounded Feller kernel and let $C : \mathbb{R}_+ \to \mathbb{R}_+$ be a continuous function. Then, for every bounded and continuous function φ, the function

$$z \to \int_0^{C(z)} \varphi(x) K_k(dx|z) \quad (30)$$

is continuous.

Proof: Let $\{z_n\}$ be a sequence in $\mathbb{R}_+$ such that $z_n \to z_0$. We have:

$$\left|\int_0^{C(z_n)} \varphi(x)K_k(dx|z_n) - \int_0^{C(z_0)} \varphi(x)K_k(dx|z_0)\right| \leq \left|\int_0^{C(z_n)} \varphi(x)K_k(dx|z_n) - \int_0^{C(z_0)} \varphi(x)K_k(dx|z_n)\right| \quad (31a)$$

$$+ \left|\int_0^{C(z_0)} \varphi(x)K_k(dx|z_n) - \int_0^{C(z_0)} \varphi(x)K_k(dx|z_0)\right|. \quad (31b)$$

The term (31b) converges to zero by the Feller property of K_k. Thus, let us focus on (31a). We have:

$$\left|\int_0^{C(z_n)} \varphi(x)K_k(dx|z_n) - \int_0^{C(z_0)} \varphi(x)K_k(dx|z_n)\right| = \left|\int_{C(z_0)}^{C(z_n)} \varphi(x)K_k(dx|z_n)\right| \quad (32)$$

$$\leq \|\varphi\|_{\infty}\left\|K_k(\cdot|z_n)\right\|_{\infty}\left|\int_{C(z_0)}^{C(z_n)} dx\right| \quad (33)$$

$$= \|\varphi\|_{\infty}\left\|K_k(\cdot|z_n)\right\|_{\infty}\left|C\left(z_n\right) - C\left(z_0\right)\right|. \quad (34)$$

The term (34) converges to zero by the continuity of C. ■

Lemma 2: Assume that K_k is a bounded Feller kernel. Furthermore, let α, C, $z \to K_k\left(z|\cdot\right)$ and h_{k+1} be continuous functions. Then, the kernel $\tilde{K}_k$ defined by (26) has the Feller property.

Proof: First, let us observe that by Lemma 1 the function q_k is continuous. Consequently, the function q_k^{α} is continuous by the continuity of q_k, α, C and $h_{k+1}\left(\cdot, 0\right)$.

Let φ be a bounded continuous function on $\mathbb{R}$. Then the following holds:

$$\int_{\mathbb{R}} \varphi(x)\tilde{K}_k\left(dx|z, y\right) = q_{k+1}^{\alpha}\left(z, y\right)\varphi\left(C\left(z\right)\right) + \frac{1-q_{k+1}^{\alpha}(z, y)}{1-q_{k+1}(z)}\int_0^{C(z)} \varphi(x)K_k\left(dx|z\right). \quad (35)$$

The continuity of the functions in (35) follows by the continuity of q_k, q_k^{α}, α, C and by Lemma 1. ■

Lemma 3: Assume that the likelihood function g_k is bounded, continuous and strictly positive. Also, let α, C, $z \to K_k\left(z|\cdot\right)$ and h_{k+1} be continuous functions. Furthermore, let us assume that there exist positive constants M_1 and M_2 such that for every $x \in \mathbb{R}$ it holds:

$$0 < M_1 \leq q_k(x) \leq M_2 < 1. \quad (36)$$

Then, if the detection function α is chosen so that it satisfies the condition:

$$\forall x \in \mathbb{R}: \quad -M_1 < \alpha(x) < 1 - M_2, \quad (37)$$

the weighted likelihood function ω_k defined by (27) is continuous, bounded and strictly positive.

Proof: Continuity of ω_k follows from the continuity of g_k, q_k, q_k^{α}, and C.

By (37) there exists $\epsilon > 0$ such that:

$$-M_1 + \epsilon \leq \alpha(x) \leq 1 - M_2 - \epsilon \quad (38)$$

holds for every $x \in \mathbb{R}$. Hence, for every $x, y \in \mathbb{R}$ we have

$$\begin{aligned} q_k^{\alpha}(x, y) &= q_k(x) + \alpha\left(y - h_{k+1}\left(C\left(x\right)\right)\right) \\ &\leq M_2 + 1 - M_2 - \epsilon \\ &\leq 1 - \epsilon. \end{aligned} \quad (39)$$

Similarly, we deduce that for every $x, y \in \mathbb{R}$ it holds:

$$q_k^{\alpha}(x, y) \geq \epsilon. \quad (40)$$

Therefore, by (27), we have

$$\|\omega_k\|_\infty \leq \|g_k\|_\infty \frac{1}{\epsilon}, \tag{41}$$

hence, ω_k is bounded.

Finally, strict positivity of ω_k follows by the strict positivity of g_k and by (36). ■

The asymptotic properties of the SPF are described by the next theorem:

Theorem 2 (Convergence of the SPF: c^N–resampling): Let us consider a SPF k_k^N with the resampling operator c^N defined by Definition 8. If

1) K_k is a bounded Feller kernel,
2) g_k is bounded, continuous and strictly positive,
3) α, C, $z \to K_k(z|\cdot)$ and h_{k+1} are continuous functions,
4) conditions (36) and (37) are satisfied,

then $k_k^N\left(\pi_{k|k}^N\right)$ converges almost surely towards the true posterior PDF $\pi_{k+1|k+1}$.

Proof: It has been proven in [17] that the multinomial sampling operator c^N satisfies (25). Furthermore, by Lemmas 1–3, the kernel $\tilde{K}_k$ is Feller, and the weighted likelihood function ω_k is bounded, continuous and strictly positive. Therefore, by Theorem 1, $k_k^N\left(\pi_{k|k}^N\right)$ converges almost surely towards the true posterior PDF $\pi_{k+1|k+1}$. ■

B. Practical considerations

Let us now discuss the meaning of conditions 1)–4) of Theorem 2 from the practical perspective.

Assumptions 1)–3) of Theorem 2 ensure that the model is "appropriately smooth", which is the case in most real life applications. Therefore, we can safely conclude that conditions 1)–3) are not very restrictive from the practical point of view.

Assumption 4) of Theorem 2 is more problematic. In particular, ensuring that (36) is satisfied is not trivial, because often the function q_k cannot be evaluated analytically [24]. Fortunately, in practice, we do not need to compute the values $q_k(x)$ for every $x \in \mathbb{R}$. It is sufficient to check whether (36) holds for every particle x_{k-1}^i, i.e., we need to check whether there exist positive constants M_1 and M_2 such that for every $i = 1, ..., N$

$$0 < M_1 \leq q_k^i \leq M_2 < 1 \tag{42}$$

holds.

Obviously, (42) is satisfied if and only if

$$\min_i \left\{q_k^i\right\} > 0, \tag{43a}$$

$$\max_i \left\{q_k^i\right\} < 1. \tag{43b}$$

Hence, for each particle the saturation event is possible, but not certain.

If conditions in (43) are satisfied, we can choose α such that

$$-\min_i\{q_k^i\}(1-\epsilon) \leq \alpha \leq \left(1-\max_i\{q_k^i\}\right)(1-\epsilon), \tag{44}$$

where the $\epsilon > 0$ is small enough so that α is nontrivial.

The advantage of choosing M_1, M_2 and α as in (42) and (44) respectively, is the low computational complexity of determining M_1, M_2, and α that satisfy condition 4) of Theorem 2. However, this approach has two shortcomings that need to be tackled.

1) α *becomes recursive*. By (44) we see that the conditions that α needs to satisfy depend on the values of q^i at time step k. Thus, α is not any more defined for all time steps k, but it becomes a recursive function α_k that needs to be updated at each iteration of the algorithm.
2) α *becomes negligible*. Since both $\min$ and $\max$ are monotonic, with the increasing number of particles the image of α_k becomes narrower. This means that the influence of α_k becomes negligible, hence the SPF becomes undistinguishable from the BPF.

Both issues are addressed in the next section.

IV. ASYMPTOTIC BEHAVIOR OF THE SPF UNDER NEW RESAMPLING

In this section we derive an improved SPF algorithm that allows for a recursive computation of the detection function α_k. Furthermore, by introducing a new resampling procedure, we make sure that at each time step k the influence of α_k is not trivial.

A. Motivation

As it was indicated in Section III, if there exists a particle x^i such that $q^i \approx 0$ or $q^i \approx 1$, then α becomes approximately zero, and therefore its influence becomes negligible. If the weight ω^i of such a particle x^i is close to one it means that the uncertainty associated with the estimate is very small. Therefore, in the next filtering step, it is reasonable to "trust" the model and limit the influence of the noisy measurement on the subsequent estimate. In general, the same reasoning holds if there exists a "small" ϵ such that in one of the intervals $[0, \epsilon]$ or $[1-\epsilon, 1]$ there are enough q^is so that the weights of the associated particles almost sum up to 1.

The situation is fundamentally different when we encounter a low weighted particle x^i for such that either $q^i \approx 0$ or $q^i \approx 1$. The probability of such an event is very high, especially when we use a large set of particles, yet such a particle does not give us any important information about the system. Nevertheless, by (44), the existence of such a particle significantly decreases the influence of α. To avoid this undesirable situation we need to discard the low weighted particles such that the corresponding q^is lay in either of the intervals $[0, \epsilon]$ or $[1-\epsilon, 1]$, and in their place resample an equal number of particles in the "high probability" regions. Such resampling only slightly influences the posterior PDF $\pi_{k|k}^N$ that approximates the true posterior PDF $\pi_{k|k}$. This is because the discussed resampling procedure cuts only the "light" tails, i.e., tails with negligible probability mass, thus also the resampled particles add insignificant weights to the approximation $\pi_{k|k}^N$. Nevertheless, by applying such resampling algorithm we are sure that α is not trivial, and that the interval $[-\epsilon, \epsilon]$ is in the

image of α, i.e., $([-\epsilon,\epsilon]\subset\alpha(\mathbb{R}))$. Furthermore, the number of particles remains constant throughout the filtering.

B. New resampling

Let us now formalize the heuristic approach described in Section IV-A. Following the convention described in Section II we introduce a new resampling procedure by defining a new resampling operator $\bar{c}^N$ acting on the space of probability measures on $\left(\mathbb{R}_+,\Sigma_{\mathbb{R}_+}\right)$. To define such an operator we first need to introduce a concept of the ϵ–set:

Definition 11 (ϵ–set): Consider the SPF setting according to Definition 10, and let q_k be a function defined by (28). For a given ϵ we define the ϵ–set Ω_ϵ by:

$$\Omega_\epsilon := \{x\in\mathbb{R}_+ : 1-\epsilon > q_k(x) > \epsilon\}. \tag{45}$$

Let us now consider an arbitrary probability measure ν on $\left(\mathbb{R}_+,\Sigma_{\mathbb{R}_+}\right)$. For a given ϵ we define a new probabilistic measure ν_ϵ as a measure ν conditioned on Ω_ϵ, i.e., for every $A\in\Sigma_{\mathbb{R}_+}$ it holds:

$$\nu_\epsilon(A) := \nu(A|\Omega_\epsilon). \tag{46}$$

Definition 12 (Resampling operator for the SPF): Let ν be a probabilistic measure on $\left(\mathbb{R}_+,\Sigma_{\mathbb{R}_+}\right)$, and let $\tilde{\epsilon}>0$ be a given constant. Let $\epsilon_0>0$ be the maximal positive constant such that the ν-measure on Ω_{ϵ_0} is greater or equal to $1-\tilde{\epsilon}$, i.e., ϵ_0 is given by:

$$\epsilon_0 := \max\{\epsilon : \nu(\Omega_\epsilon)\geq 1-\tilde{\epsilon}\}. \tag{47}$$

The new resampling operator $\bar{c}^N$ assigns to every probabilistic measure $\nu\in\mathcal{P}\left(\mathbb{R}_+,\Sigma_{\mathbb{R}_+}\right)$ its random approximation $\bar{c}^{N,x}(\nu)$ given by

$$\bar{c}^{N,x}(\nu) := \frac{1}{N}\sum_{j=1}^{N}\delta_{\tilde{V}_j(x)}, \tag{48}$$

where the $\{\tilde{V}_j\}_{j=1}^N$ is a set of N i.i.d. random variables distributed according to ν_{ϵ_0}.

Thus, the newly defined resampling operator $\bar{c}^N$, when applied to a measure ν, returns a probabilistic measure concentrated on the set that has a ν-measure close to $1-\tilde{\epsilon}$. When used in the recursive framework of the SPF, the operator $\bar{c}^N$ guarantees that the interval $[-\epsilon_0,\epsilon_0]$ is in the image of the detection function α, i.e., $([-\epsilon_0,\epsilon_0]\subset\alpha(\mathbb{R}))$.

Note that the operator $\bar{c}^N$ depends on both the sample size N and the constant $\tilde{\epsilon}>0$. In what follows we show that the choice of $\tilde{\epsilon}$ is not arbitrary but it is strictly determined by the sample size N. This is why in Definition 12 we did not use any symbol indicating the dependency of $\bar{c}^N$ on $\tilde{\epsilon}$.

C. Almost sure convergence of the SPF under the new resampling

In this section we prove that as the number of samples N increases, the SPF with the resampling operator $\bar{c}^N$ from Definition 12 converges almost surely to the true posterior PDF $\pi_{k+1|k+1}$.

First, we prove the following two lemmas.

Lemma 4: Let ν be an arbitrary probabilistic measure on $\left(\mathbb{R}_+,\Sigma_{\mathbb{R}_+}\right)$ and let $\tilde{\epsilon}>0$ be a given positive constant. Furthermore, let $\left\{\tilde{V}_j\right\}_{j=1}^N$ be a set of i.i.d. variables with a common distribution ν_{ϵ_0}, with ϵ_0 defined in (47). Finally, let φ be a continuous and bounded function on $\mathbb{R}_+$. Then, the following holds:

$$\mathbb{E}_\nu\left(\varphi\left(\tilde{V}_j\right)-\int_{\mathbb{R}_+}\varphi(x)d\nu(x)\right)\leq 2\|\varphi\|_\infty\tilde{\epsilon} \tag{49}$$

where $\|\cdot\|_\infty$ is a supremum norm on $\mathbb{R}_+$.

Proof: By (46) the distribution ν_{ϵ_0} of the variables $\tilde{V}_j$ is a measure ν conditioned on the set Ω_{ϵ_0}. Therefore, we have:

$$\begin{aligned}&\left|\mathbb{E}_\nu\left(\varphi\left(\tilde{V}_j\right)-\int_{\mathbb{R}_+}\varphi(x)d\nu(x)\right)\right|\\ &=\left|\mathbb{E}_\nu\left(\varphi\left(\tilde{V}_j\right)\right)-\int_{\mathbb{R}_+}\varphi(x)d\nu(x)\right| &(50)\\ &=\left|\int_{\mathbb{R}_+}\varphi(x)d\nu_{\epsilon_0}(x)-\int_{\mathbb{R}_+}\varphi(x)d\nu(x)\right|. &(51)\end{aligned}$$

By the definition of ν_{ϵ_0} we can write the first integral in (51) as

$$\int_{\mathbb{R}_+}\varphi(x)d\nu_{\epsilon_0}(x)=\int_{\mathbb{R}_+\cap\Omega_{\epsilon_0}}\varphi(x)\nu(\Omega_{\epsilon_0})^{-1}d\nu(x). \tag{52}$$

Let us split the second integral in (51) into two integrals over $\mathbb{R}_+\cap\Omega_{\epsilon_0}$ and $\mathbb{R}_+\cap\Omega_{\epsilon_0}^c$ respectively, i.e.,

$$\begin{aligned}&\int_{\mathbb{R}_+}\varphi(x)d\nu(x)\\ &=\int_{\mathbb{R}_+\cap\Omega_{\epsilon_0}}\varphi(x)d\nu(x)+\int_{\mathbb{R}_+\cap\Omega_{\epsilon_0}^c}\varphi(x)d\nu(x). &(53)\end{aligned}$$

Then, by (52)–(53), the expression (51) is bounded from above by

$$\begin{aligned}&\left|\int_{\mathbb{R}_+\cap\Omega_{\epsilon_0}}\varphi(x)\nu(\Omega_{\epsilon_0})^{-1}d\nu(x)-\int_{\mathbb{R}_+\cap\Omega_{\epsilon_0}}\varphi(x)d\nu(x)\right|\\ &+\left|\int_{\mathbb{R}_+\cap\Omega_{\epsilon_0}^c}\varphi(x)d\nu(x)\right| &(54)\\ &\leq\|\varphi\|_\infty\left(|1-\nu(\Omega_{\epsilon_0})|+\nu\left(\Omega_{\epsilon_0}^c\right)\right) &(55)\\ &\leq\|\varphi\|_\infty(\tilde{\epsilon}+\tilde{\epsilon}) &(56)\\ &=2\tilde{\epsilon}\|\varphi\|_\infty &(57)\end{aligned}$$

■

Lemma 5: Let ν be an arbitrary probabilistic measure on $\left(\mathbb{R}_+,\Sigma_{\mathbb{R}_+}\right)$ and let $\tilde{\epsilon}>0$ be a given positive constant. Furthermore, let $\left\{\tilde{V}_j\right\}_{j=1}^N$ be a set of i.i.d. variables with a common distribution ν_{ϵ_0}, with ϵ_0 defined in (47). Finally,

let φ be a continuous and bounded function on $\mathbb{R}_+$. Then, the following holds:

$$\mathbb{E}_\nu\left(\left(\frac{1}{N}\sum_{j=1}^{N}\varphi\left(\tilde{V}_j\right)-\int_{\mathbb{R}_+}\varphi(x)d\nu(x)\right)^4\right) \leq \frac{16}{N^2}\|\varphi\|_\infty^4\left(3+4\tilde{\epsilon}+6N\tilde{\epsilon}^2+N^2\tilde{\epsilon}^4\right), \quad (58)$$

where $\|\cdot\|_\infty$ is a supremum norm on $\mathbb{R}_+$.

Proof:

$$\mathbb{E}_\nu\left(\left(\frac{1}{N}\sum_{j=1}^{N}\varphi\left(\tilde{V}_j\right)-\int_{\mathbb{R}_+}\varphi(x)d\nu(x)\right)^4\right) \quad (59)$$

$$=\frac{1}{N^4}\mathbb{E}_\nu\left(\left(\sum_{j=1}^{N}\left(\varphi\left(\tilde{V}_j\right)-\int_{\mathbb{R}_+}\varphi(x)d\nu(x)\right)\right)^4\right) \quad (60)$$

$$=\frac{1}{N^4}\sum_{j_1,\ldots,j_4=1}^{N}\mathbb{E}_\nu\left(\prod_{k=1}^{4}\left(\varphi\left(\tilde{V}_{j_k}\right)-\int_{\mathbb{R}_+}\varphi(x)d\nu(x)\right)\right) \quad (61)$$

Because the variables are mutually independent, the sum in (61) can be decomposed into the summation of the even terms:

$$\sum_{j=1}^{N}\mathbb{E}_\nu\left(\varphi\left(\tilde{V}_j\right)-\int_{\mathbb{R}_+}\varphi(x)d\nu(x)\right)^4 \quad (62)$$

$$+6\sum_{j_1>j_2=1}^{N}\mathbb{E}_\nu\left(\varphi\left(\tilde{V}_{j_1}\right)-\int_{\mathbb{R}_+}\varphi(x)d\nu(x)\right)^2 \times\mathbb{E}_\nu\left(\varphi\left(\tilde{V}_{j_2}\right)-\int_{X}\varphi(x)d\nu(x)\right)^2, \quad (63)$$

and the odd terms:

$$4\sum_{j_1\neq j_2=1}^{N}\mathbb{E}_\nu\left(\varphi\left(\tilde{V}_{j_1}\right)-\int_{\mathbb{R}_+}\varphi(x)d\nu(x)\right) \times\mathbb{E}_\nu\left(\varphi\left(\tilde{V}_{j_2}\right)-\int_{\mathbb{R}_+}\varphi(x)d\nu(x)\right)^3 \quad (64)$$

$$+12\sum_{\substack{j_1>j_2\geq 1\\ j_1,j_2\neq j_3\geq 1}}\prod_{k=1}^{2}\mathbb{E}_\nu\left(\varphi\left(\tilde{V}_{j_k}\right)-\int_{\mathbb{R}_+}\varphi(x)d\nu(x)\right) \times\mathbb{E}_\nu\left(\varphi\left(\tilde{V}_{j_3}\right)-\int_{\mathbb{R}_+}\varphi(x)d\nu(x)\right)^2 \quad (65)$$

$$+24\sum_{j_1>\ldots>j_4=1}^{N}\prod_{k=1}^{4}\left(\mathbb{E}_\nu\left(\varphi\left(\tilde{V}_{j_k}\right)-\int_{\mathbb{R}_+}\varphi(x)d\nu(x)\right)\right) \quad (66)$$

Since ν is a probabilistic measure, the integral $\left|\int_{\mathbb{R}_+}\varphi(x)d\nu(x)\right|$ is bounded from above by $\|\varphi\|_\infty$. Thus, for every $j\in\{1,\ldots,N\}$ and every $k\in\mathbb{N}$ we have:

$$\mathbb{E}_\nu\left(\varphi\left(\tilde{V}_j\right)-\int_{\mathbb{R}_+}\varphi(x)d\nu(x)\right)^k\leq 2^k\|\varphi\|_\infty^k. \quad (67)$$

Furthermore, given Lemma 4, the first-order terms are bounded by $2\|\varphi\|_\infty\tilde{\epsilon}$. Therefore, by (62)–(66) and (67), the expression (59) is bounded from above by:

$$\frac{2^4}{N^4}\|\varphi\|_\infty^4\times\left(N+6\binom{N}{2}+4N(N-1)\tilde{\epsilon}\right. \left.+12N\binom{N-1}{2}\tilde{\epsilon}^2+24\binom{N}{4}\tilde{\epsilon}^4\right) \quad (68)$$

$$\leq\frac{2^4}{N^2}\|\varphi\|_\infty^4\left(3+4\tilde{\epsilon}+6N\tilde{\epsilon}^2+N^2\tilde{\epsilon}^4\right) \quad (69)$$

■

Theorem 3 (Convergence of the SPF: $\bar{c}^N$–resampling): Let us set $\tilde{\epsilon}=\frac{1}{\sqrt{N}}$ and let $\bar{c}^N$ be the resampling operator introduced in Definition 12. Furthermore, let k_k^N be the SPF, with the resampling operator $\bar{c}^N$ such that the following hold:

1) K_k is a bounded Feller kernel,
2) g_k is bounded, continuous and strictly positive,
3) α, C, $z\to K_k(z|\cdot)$ and h_{k+1} are continuous functions.

Then $k_k^N\left(\pi_{k|k}^N\right)$ converges almost surely towards the true posterior PDF $\pi_{k+1|k+1}$.

Proof: For a given $\nu\in\mathcal{P}\left(\mathbb{R}_+,\Sigma_{\mathbb{R}_+}\right)$, by the definition of $\bar{c}^N$ we have

$$\mathbb{E}_\nu\left[\left(\int_{\mathbb{R}_+}\varphi(x)\left[\bar{c}^{N,\cdot}(\nu)\right](dx)-\int_{\mathbb{R}_+}\varphi(x)\nu(dx)\right)^4\right] =\mathbb{E}_\nu\left(\left(\frac{1}{N}\sum_{j=1}^{N}\varphi\left(\tilde{V}_j\right)-\int_{\mathbb{R}_+}\varphi(x)d\nu(x)\right)^4\right), \quad (70)$$

where $\left\{\tilde{V}_j\right\}_{j=1}^N$ is a set of i.i.d. random variables distributed according to ν_{ϵ_0}. Since $\tilde{\epsilon}=\frac{1}{\sqrt{N}}$ then, by Lemma 5, we have:

$$\mathbb{E}_\nu\left(\left(\frac{1}{N}\sum_{j=1}^{N}\varphi\left(\tilde{V}_j\right)-\int_{\mathbb{R}_+}\varphi(x)d\nu(x)\right)^4\right)\leq\frac{176}{N^2}\|\varphi\|_\infty^4, \quad (71)$$

thus the resampling operator $\bar{c}^N$ satisfies (25). Furthermore, by Lemmas 1–2, the kernel $\tilde{K}_k$ is Feller. Moreover, by the definition of the operator $\bar{c}^N$, (36)–(37) are satisfied (with $M_1=\epsilon_0$ and $M_2=1-\epsilon_0$) hence, by Lemma 3, the weighted likelihood function ω_k is bounded, continuous and strictly positive. Therefore, by Theorem 1, $k_k^N\left(\pi_{k|k}^N\right)$ converges almost surely towards the true posterior PDF $\pi_{k+1|k+1}$. ■

D. Practical considerations

Let us now discuss the practical properties of the improved SPF (iSPF), i.e., the SPF with the resampling operator $\bar{c}^N$, in view of Theorem 3.

1) Implementation: The conditions 1)–3) of Theorem 3 are consistent with the conditions 1)–3) of Theorem 2 and play exactly the same role, i.e., they ensure an "appropriate smoothness" of the model. Thanks to the construction of the operator $\bar{c}^N$, the condition 4) of Theorem 2 is not necessary anymore in Theorem 3. Indeed, the conditions (36)–(37) are always satisfied with $M_1 = \epsilon_0$ and $M_2 = 1 - \epsilon_0$.

In practical applications at each filtering iteration we need to compute ϵ_0 according to (47). Since the real SPF algorithm approximates the true PDF by the set of N samples $\left\{\left(x^i, \omega^i, q^i\right)\right\}_{i=1}^N$, given $\tilde{\epsilon}$, we compute ϵ_0 by:

$$\epsilon_0 := \max\left\{\epsilon : \sum_{i:1-\epsilon \geq q^i \geq \epsilon} \omega^i \geq 1 - \tilde{\epsilon}\right\}. \tag{72}$$

A possible solution to the optimization problem (72) is presented in Algorithm 4.

Algorithm 4 Computation of ϵ_0

Input: $\left\{\left(\omega^i, q^i\right)\right\}_{i=1}^N, \tilde{\epsilon}$
Initialize $\epsilon = 0$
repeat
$\epsilon = \epsilon + \frac{1}{N}$
until $\sum_{i:1-\epsilon > q^i > \epsilon} \omega^i \leq 1 - \tilde{\epsilon}$
$\epsilon_0 = \epsilon - \frac{1}{N}$

With ϵ_0 chosen, the resampling procedure, represented abstractly by the operator $\bar{c}^N$, proceeds as follows. First, the algorithm discards the particles $\left(x^i, \omega^i, q^i\right)$ such that either $1 - \epsilon_0 < q^i$ or $q^i < \epsilon_0$. Next, the scaled degeneracy measure N'_{eff} is computed by

$$N'_{\text{eff}} := \frac{1}{\sum_{i=1}^{N'} (\omega_k^i)^2}, \tag{73}$$

where N' is a number of particles remaining after the discarding step. If N'_{eff} drops below a specified threshold

$$N'_{\text{T}} := N_{\text{T}} \cdot N'/N, \tag{74}$$

which means that the particles that were not discarded degenerate, all the particles are resampled according to Algorithm 1.

If the degeneracy does not occur the algorithm resamples N' particles from the conditional distribution ν_{ϵ_0}, which is approximated by the empirical PDF

$$\left\{\left(x^i, \frac{\omega^i}{\sum_{j=1}^{N'} \omega^j}\right)\right\}_{i=1}^{N'}. \tag{75}$$

Note that this resampling method has all the properties desired from the resampling algorithm discussed in Section IV-A. The overall resampling procedure is summarized in Algorithm 5.

As indicated in Section III-B, for the proper definition of the SPF the detection function α needs to be defined

Algorithm 5 SPF resampling

Input: $\left\{\left(x^i, \omega^i, q^i\right)\right\}_{i=1}^N, \epsilon_0, N_{\text{T}}$
Discarding step:
Discard the particles such that one of the following holds:

$$1 - \epsilon_0 < q^i \qquad \text{or} \qquad q^i < \epsilon_0$$

Compute the degeneracy measure for the remaining N' particles:

$$N'_{\text{eff}} := \frac{1}{\sum_{i=1}^{N'} (\omega^i)^2}$$

Resampling step:
if $N'_{\text{eff}} < N'_{\text{T}}$ **then**
resample $\left\{\left(x^i, \omega^i\right)\right\}_{i=1}^N$ according to Algorithm 1
else
for $i = 1$ **to** $N - N'$ **do**
draw x^i from

$$x^i \sim \left\{\left(x^i, \frac{\omega^i}{\sum_{i=1}^{N'} \omega^i}\right)\right\}_{i=1}^{N'}$$

$\omega^i = \frac{1}{N-N'}\left(1 - \sum_{i=1}^{N'} \omega^i\right)$
end for
end if

recursively. Indeed, from (44) we can see that α depends on the set of probabilities of saturation $\left\{q_k^i\right\}_{i=1}^N$. Therefore, for the recursive selection of the appropriate detection function α we proceed as follows: First, we choose an arbitrary detection function α_0, and a constant $\epsilon > 0$; Then at each filtering step $k = 1, 2, ...$ we compute the minimum and maximum of the set $\left\{q_k^i\right\}_{i=1}^N$ and set α_k to be equal to

$$\alpha_k(z) := \begin{cases} \alpha_0(z) \min_i\left\{q_k^i\right\}(1-\epsilon) & \textbf{for} \quad z < \alpha_{k-1}^{-1}(0) \\ \alpha_0(z)\left(1 - \max_i\left\{q_k^i\right\}\right)(1-\epsilon) & \textbf{otherwise} \end{cases} \tag{76}$$

This is summarized in Algorithm 6.

Algorithm 6 Update of the detection function α_k

Input: $\left\{q_k^i\right\}_{i=1}^N, \alpha_0, \epsilon$
Compute minimum and maximum:

$$\text{MIN} := \min_{i=1,\ldots,N}\left\{q_k^i\right\} \quad \text{MAX} := \max_{i=1,\ldots,N}\left\{q_k^i\right\}$$

if $z < \alpha_0^{-1}(0)$ **then**
$\alpha_k(z) := \alpha_0(z)\text{MIN}\,(1-\epsilon)$
else
$\alpha_k(z) := \alpha_0(z)\,(1 - \text{MAX})\,(1-\epsilon)$
end if

Note that the constant ϵ used in Algorithm 6 can be chosen arbitrarily from the interval $(0, 1)$. This gives us a degree of freedom in choosing between the stronger influence of α (for small values of ϵ) and a stricter upper bound for the weighted likelihood function ω_k (for bigger values of ϵ). The exact relation between these two properties will be investigated in

our further research.

The overall iSPF is summarized in Algorithm 7.

Algorithm 7 improved SPF

Input: $\left\{\left(x_k^i, \omega_k^i, q_k^i\right)\right\}_{i=1}^N, \alpha_k, y_{k+1}, \epsilon, N_{\mathrm{T}}$
Output: $\left\{\left(x_{k+1}^i, \omega_{k+1}^i, q_{k+1}^i\right)\right\}_{i=1}^N, \alpha_{k+1}$
Prediction:
for $i = 1$ **to** N **do**
Compute $q_i^\alpha := q_k^i + \alpha_k\left(y_{k+1} - h_{k+1}\left(C\left(x_k^i\right)\right)\right)$
Compute x_{k+1}^i according to Algorithm 3
Compute ω_{k+1}^i according to Algorithm 3
Compute $q_{k+1}^i := \int_{C(x_{k+1}^i)}^{+\infty} \mathbb{P}\left(F\left(x_k, w_k\right) = z | x_{k+1}^i\right) dz$
end for
Resampling:
Compute ϵ_0 according to Algorithm 4
Resample particles according to Algorithm 5
for $i = 1$ **to** N **do**
Compute $q_{k+1}^i := \int_{C(x_{k+1}^i)}^{+\infty} \mathbb{P}\left(F\left(x_k, w_k\right) = z | x_{k+1}^i\right) dz$
end for{C}ompute the predicted probabilities of saturation for the resampled particles
Update of the detection function:
Compute α_{k+1} according to Algorithm 6

2) Detection function: Let us analyze the problem of α becoming negligible, mentioned in Section III-B. By the definition of the resampling operator $\bar{c}^N$ and by (76), for each time step $k = 1, 2, ...$ the image of α_k contains the interval $[-\epsilon_0(1-\epsilon), \epsilon_0(1-\epsilon)]$, i.e., $[-\epsilon_0(1-\epsilon), \epsilon_0(1-\epsilon)] \subset \alpha_k(\mathbb{R})$. Therefore, α_k is never trivial. However, the value of ϵ_0 depends on the value of $\tilde{\epsilon}$ (see (47)), that in view of Theorem 3, decreases with the rate $\frac{1}{\sqrt{N}}$ when $N \to \infty$. This means that the measure ν_{ϵ_0} of the set Ω_{ϵ_0} increases, hence by (45)–(47) ϵ_0 decreases, and therefore the image of α_k becomes narrower.

The rate of decrease of ϵ_0 depends on the particular shapes of both the distribution of the probabilities of saturation $\left\{q_k^i\right\}_{i=1}^N$ and the distribution of weights $\left\{\omega_k^i\right\}_{i=1}^N$. The following example illustrates this dependency. First, let us assume that both sets $\left\{q_k^i\right\}_{i=1}^N$ and $\left\{\omega_k^i\right\}_{i=1}^N$ are uniformly distributed on the interval $[0, 1]$. Then, for a given number of particles N, we expect (in the statistical sense) to discard $\sqrt{N}$ particles and the expected value of ϵ_0 is equal to $\frac{1}{2\sqrt{N}}$. Second, let us assume that the weights $\left\{\omega_k^i\right\}_{i=1}^N$ are again distributed uniformly on the interval $[0, 1]$, but the set $\left\{q_k^i\right\}_{i=1}^N$ is approximated by the Gaussian distribution[2] with mean $\frac{1}{2}$ and the standard deviation $\frac{1}{10}$. In such a case, we still expect to discard $\sqrt{N}$ particles at each step, but this time the value of ϵ_0 is given by $\frac{1}{2} + \frac{\sqrt{2}}{10}\mathrm{erf}^{-1}\left(\frac{2}{\sqrt{N}} - 1\right)$. Table I compares the two cases for three different values of N.

As we can see the expected value of ϵ_0 is strongly dependent on the distribution of $\left\{q_k^i\right\}_{i=1}^N$. Therefore, the influence of α_k, which is determined by the value of ϵ_0, also depends on the shape of $\left\{q_k^i\right\}_{i=1}^N$. Let us explain a nature of this dependency by analyzing the examples from Table I.

[2]The set $\left\{q_k^i\right\}_{i=1}^N$ is bounded, therefore by saying that it is approximated by Gaussian distribution we mean a Gaussian with truncated tails and appropriately rescaled.

TABLE I
PROPERTIES OF THE SPF FOR DIFFERENT DISTRIBUTIONS OF PROBABILITIES OF SATURATION $\left\{q_k^i\right\}_{i=1}^N$ AND WEIGHTS $\left\{\omega_k^i\right\}_{i=1}^N$

$\left\{q_k^i\right\}_{i=1}^N, \left\{\omega_k^i\right\}_{i=1}^N \approx \mathcal{U}(0,1)$			
	$N = 10^2$	$N = 10^4$	$N = 10^6$
expected number of discarded particles	10	100	1000
expected value of ϵ_0	0.05	0.005	0.0005
$\left\{q_k^i\right\}_{i=1}^N \approx \mathcal{N}\left(\frac{1}{2}, \frac{1}{10}\right)$, $\left\{\omega_k^i\right\}_{i=1}^N \approx \mathcal{U}(0,1)$			
	$N = 10^2$	$N = 10^4$	$N = 10^6$
expected number of discarded particles	10	100	1000
expected value of ϵ_0	0.372	0.267	0.191

In the first case, where $\left\{q_k^i\right\}_{i=1}^N \approx \mathcal{U}(0,1)$, the model assigns the same probability to all the possible values of the probabilities of saturation q_k^i. In such situations the standard Bayesian update procedure should be more than sufficient in obtaining an accurate estimate. Thus, the small values of ϵ_0, and therefore the low influence of α_k is acceptable.

In the second case, where $\left\{q_k^i\right\}_{i=1}^N \approx \mathcal{N}\left(\frac{1}{2}, \frac{1}{10}\right)$, the values of the majority of the probabilities of saturation q_k^i are close to $\frac{1}{2}$. That means that the model is very uncertain in predicting whether the saturation will occur or not. In such cases the standard Bayesian approach is slow in detecting extreme changes of the system (e.g., saturation). Thus, for this example, it is strongly recommended to enforce the update procedure. For that we require a strong influence of α_k, hence relatively large values of ϵ_0.

Note that in both cases the expected number of discarded particles is small compared to the total number of particles. Therefore, the problem of losing diversity of the samples [16] is avoided.

3) Numerical example: We finish this section with a comparison of the SPF developed in [23], the iSPF derived in this paper, and the BPF. To illustrate the abilities of all three methods we have chosen the SSDS given by

$$x_{k+1} = \min\left(x_k + w_k, C\left(x_k\right)\right), \tag{77}$$

$$y_k = x_k + v_k, \tag{78}$$

where w_k is a random variable distributed according to the exponential distribution, with parameter θ, i.e., with the expected value $\mathbb{E}w_k = \theta^{-1}$. The variable v_k is a zero-mean Gaussian variable with standard deviation σ_v. The boundary

function $C(\cdot)$ is defined by:

$$C\left(x\right) := x + \log(2)/\theta. \tag{79}$$

The state model (77) is nonlinear and non-Gaussian, whereas the observation model (78) is both linear and conditionally Gaussian. The stochastic process (77) is a Lindley-type process, i.e., it is a modification of the celebrated *Lindley's recursion*, one of the most studied stochastic processes in applied probability [26], [46]. These type of processes are extensively used in queueing theory [46], [47].

To simulate the process from the initial state $x_0 = 1$, we used $\theta = 1$, $\sigma_v = 1$. The length of the simulation is 20 time steps. Note, that because the variable w_k is exponentially distributed, the cumulative density function of the random variable x_{k+1} is known. Thus, the integral in (13) can be computed analytically:

$$q_i = \exp\left(-\theta\left(C\left(x_k^i\right) - x_k^i\right)\right). \tag{80}$$

Figures 1 and 2 compare the results obtained by applying the BPF, the iSPF, and the SPF with 10 and 1000 particles respectively. All three filters use the model (77)–(78) with true parameters. The offset of 0.5 is introduced by setting the initial state p_0 for all three filters to $p_0(\cdot) = \mathcal{N}\left(\cdot; 0.5, 0.1\right)$. The resampling threshold N_T is set to 30% of the number of particles. Both the iSPF and the SPF are using the same detection function α_0 given by:

$$\alpha\left(z\right) = \begin{cases} 1 & \textbf{if} \quad z > 2, \\ -1 & \textbf{if} \quad z < 0, \\ z/2 - 1 & \textbf{otherwise}, \end{cases} \tag{81}$$

where $z = y_{k+1} - h_{k+1}\left(C\left(x_k^i\right)\right)$ is evaluated at each time step.

Figures 1 and 2 present the average of ten independent filters of each type applied to the simulated signal. From both figures we can conclude that introducing new resampling procedure improves the performance of the SPF introduced in [23]. However, with the growing number of samples the difference between the SPF and the iSPF becomes smaller. This is not surprising since in the view of Theorems 2 and 3 both filters converge to the same distribution as the number of samples increase.

Both simulations present results for a relatively small number of particles. This is because, as it was previously explained the influence of the detection function is the most visible when there are few particles. In this example we can observe that both the iSPF and the SPF outperform the BPF. This is confirmed by comparing the *mean square errors*[3] (MSE) of the three methods. Table III reports values of such errors and their dependency on the number of particles.

Finally, Figure 3 shows the standard deviations of the MMSE estimates of three discussed filters. It might be noticed that the spread of the MMSE estimates of filters using ten particles is much higher than the spread of the MMSE estimates of filters using thousand particles. In the latter case the spread is negligible, and in the former case it is considerable only for the BPF during the first few steps of the simulation.

[3]By the MSE we understand the average squared deviation of the estimate from the true value of the state, i.e., $\sum_{k=1}^{20}\left(\hat{x}(k) - x(k)\right)^2/20$.

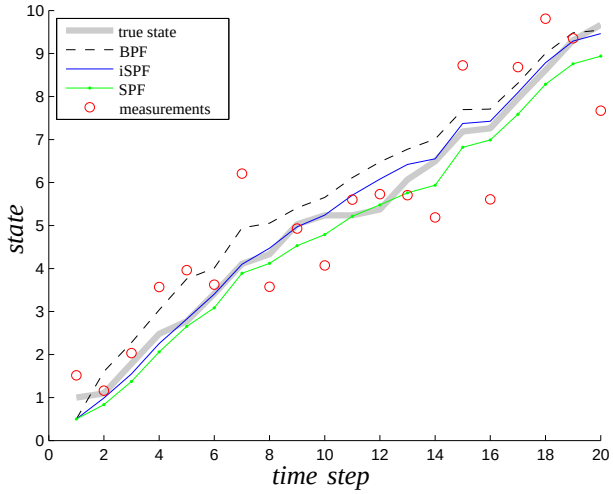

Fig. 1. The BPF, the iSPF and the SPF with 10 particles applied to the system (77)–(78). The thick solid line is the true value of the state, the circles denote the measurements, the thin dashed line denotes the MMSE estimate obtained by the BPF, the thin solid line represents the MMSE estimate of the state obtained by the iSPF and the thin solid line with dots represents the MMSE estimate of the state obtained by the SPF.

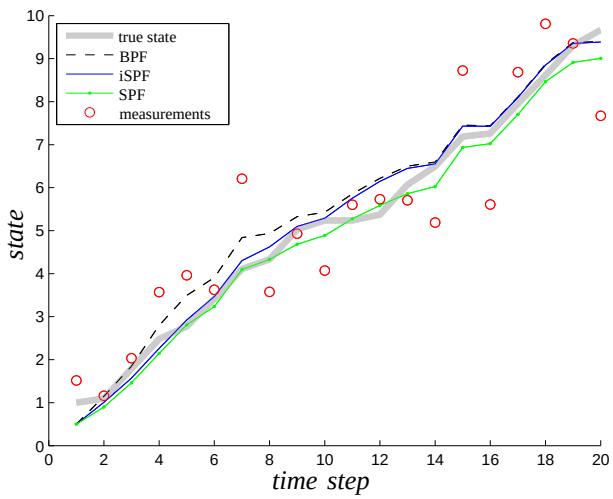

Fig. 2. The BPF, the iSPF and the SPF with 1000 particles applied to the system (77)–(78). The thick solid line is the true value of the state, the circles denote the measurements, the thin dashed line denotes the MMSE estimate obtained by the BPF, the thin solid line represents the MMSE estimate of the state obtained by the iSPF and the thin solid line with dots represents the MMSE estimate of the state obtained by the SPF.

V. SUMMARY

Saturated Stochastic Dynamical Systems (SSDS) are severely nonlinear models that are often met in real life problems. Due to their complicated dynamical structure the states or the parameters of the SSDS can be accurately estimated only by non-parametric filters such as Particle

TABLE II

COMPARISON OF THE MEAN SQUARE ERRORS

	$N = 10$	$N = 100$	$N = 1000$
BPF	0.3726	0.1808	0.1848
SPF	0.1516	0.0990	0.0946
iSPF	0.0693	0.0933	0.0880

TABLE III

THE MSE OF THREE FILTERS: THE BPF, THE SPF AND THE ISPF OBTAINED FOR DIFFERENT NUMBERS OF PARTICLES $N = 10, 100, 1000$.

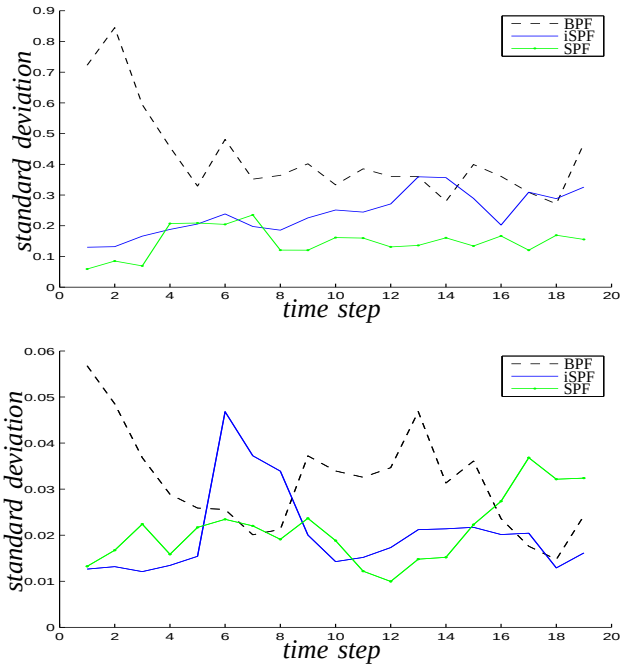

Fig. 3. The standard deviations of 10 MMSE estimates obtained by: the BPF (dashed line), the iSPF (solid line) and the SPF (solid-dotted line) with 10 particles (above) and 1000 particles (below) applied to the system (77)–(78). Note that the standard deviations of the filter employing 1000 particles is one order of magnitude lower than the standard deviations of filters using 10 particles.

Filters (PF).

The Saturated Particle Filter (SPF) exploits dynamical systems with a specific structure that allows the state to saturate. A characteristic feature of the SPF is the incorporation of the measurement in the prediction step of the Bayesian filtering. This is done through the use of the detection function α. In this paper we have derived the *improved Saturated Particle Filter* (iSPF) with the novel resampling algorithm and discussed asymptotic properties of the SPF and the iSPF.

In Section II we have described the standard SPF as it was introduced in [23]. Next, in Section III we have formulated the SPF as an operator acting on the spaces of probability measures $\mathcal{P}(\mathbb{R}_+)$. Such an abstraction enables the proof of the almost sure convergence of the posterior PDF given by the standard SPF to the true posterior PDF. We have also discussed the practical advantages and shortcomings of the standard SPF.

In Section IV we have derived the iSPF. This new algorithm is different from the standard SPF in two aspects:

1) The detection function α_k is updated recursively at each time step k,
2) A novel resampling procedure $\bar{c}^N$ is used to discard the low weights particles x^i such that the corresponding probability of saturation q^i achieve extreme values ($q^i \approx 0$ or $q^i \approx 1$).

We have shown in Theorem 3 that the iSPF also converges almost surely to the true posterior PDF.

From Theorem 3 we have concluded that as the number of samples N grows the influence of the detection function α_k declines. This is not a surprise since the BPF can be considered as a SPF with a detection function $\alpha^{BPF} = 0$. Furthermore, both the iSPF and the BPF converge almost surely to the same distribution. Thus, with the growing number of samples N the difference between those two estimators becomes smaller, hence the distance between α_k and α^{BPF} also converges to zero.

The influence of the detection function α_k is most noticeable when the number of samples N is relatively small. The strength of the influence always depends on the distribution of the weights $\{\omega_k^i\}_{i=1}^N$ and the distribution of the probabilities of saturation $\{q_k^i\}_{i=1}^N$. An analysis of this dependency has been illustrated on an example described in Table I. We have concluded that the influence of α_k is much stronger in the case when the set $\{q_k^i\}_{i=1}^N$ is concentrated around $\frac{1}{2}$, i.e., $\{q_k^i\}_{i=1}^N \approx \mathcal{N}\left(\frac{1}{2}, \frac{1}{10}\right)$, than in the case when the distribution of the probabilities of saturation $\{q_k^i\}_{i=1}^N$ is heavy tailed, i.e., $\{q_k^i\}_{i=1}^N \approx \mathcal{U}(0,1)$. Furthermore, in Section IV-D.3 the illustrative example of SSDS is used to compare the performance of the iSPF, the SPF and the BPF.

In general, the influence of the detection function α_k can be further modified by an appropriate choice of the constant $\epsilon > 0$ in Algorithm 6. This must be done carefully because, by (41), the value of ϵ determines the upper bound for the weighted likelihood function ω_k, hence also the variance of the weights $\{\omega_k^i\}_{i=1}^N$. The smaller the ϵ is the wider the image of α_k but, at the same time, the bigger is the upper bound for ω_k. The exact nature of this relation will be investigated in our further research.

ACKNOWLEDGEMENT

This research is funded by the dredging company IHC Systems B. V. P. O. Box 41, 3360 AA Sliedrecht, the Netherlands.

REFERENCES

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

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

[3] H. Tanizaki, *Nonlinear Filters: Estimation and Applications*. Springer Berlin, 1996.
[4] S. Julier and J. Uhlmann, "A General Method for Approximating Nonlinear Transformations of Probability Distributions," Robotics Research Group, Department of Engineering Science, University of Oxford, Tech. Rep., 1996.
[5] T. Lefebvre, H. Bruyninckx, and J. Schutter, "Kalman Filters for Nonlinear Systems: A Comparison of Performance," *International Journal of Control*, vol. 77, pp. 639–653, 2004.
[6] S. Julier and J. Uhlmann, "Unscented Filtering and Nonlinear Estimation," *Proceedings of the IEEE*, vol. 92, pp. 401–422, 2004.
[7] M. Nørgaard, N. Poulsen, and O. Ravn, "New Developments in State Estimation for Nonlinear Systems," *Automatica*, vol. 36, pp. 1627–1638, 2000.
[8] D. Alspach and H. Sorenson, "Recursive Bayesian Estimation Using Gaussian Sums," *Automatica*, vol. 7, pp. 465–479, 1971.
[9] I. Arasaratnam, S. Haykin, and R. Elliott, "Discrete-Time Nonlinear Filtering Algorithms Using Gauss-Hermite Quadrature," *Proceedings of the IEEE*, vol. 95, pp. 953–977, 2007.
[10] K. Ito and K. Xiong, "Gaussian Filters for Nonlinear Filtering Problems," *IEEE Transactions on Automatic Control*, vol. 45, pp. 910–927, 2000.
[11] G. Kitagawa, "Non-Gaussian State-Space Modeling of Nonstationary Time Series," *Journal of the American Statistical Association*, vol. 82, pp. 1032–1041, 1987.
[12] B. Ristic, S. Arulampalam, and N. Gordon, *Beyond the Kalman Filter: Particle Filters for Tracking Application*. Artech House, 2004.
[13] J. Rawlings and B. Bakshi, "Particle Filtering and Moving Horizon Estimation," *Computers and Chemical Engineering*, vol. 30, pp. 1529–1541, 2006.
[14] G. Kitagawa, "A Self-Organizing State-Space Model," *Journal of the American Statistical Association*, vol. 93, pp. 1203–1215, 1998.
[15] N. Metropolis and S. Ulam, "The Monte Carlo Methods," *Journal of the American Statistical Association*, vol. 44, pp. 335–341, 1949.
[16] S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A Tutorial on Particle Filters for Online Nonlinear/Non- Gaussian Bayesian Tracking," *IEEE Transactions on Signal Processing*, vol. 50, pp. 174–188, 2002.
[17] D. Cristian and A. Doucet, "A Survey of Convergence Results on Particle Filtering Methods for Practitioners," *IEEE Transactions on Signal Processing*, vol. 50, pp. 736–746, 2002.
[18] S. Kim, G. Eyink, J. Restrepo, F. Alexander, and G. Johnson, "Ensemble Filtering for Nonlinear Dynamics," *Monthly Weather Review*, vol. 131, pp. 2586–2594, 2003.
[19] K. Okuma, A. Taleghani, N. de Freitas, J. Little, and D. Lowe, "A Boosted Particle Filter: Multitarget Detection and Tracking," *Lecture Notes in Computer Science*, vol. 3021, pp. 28–39, 2004.
[20] X. Shao, B. Huang, and J. M. Lee, "Constrained Bayesian State Estimation - A Comparative Study and a New Particle Filter Based Approach," *Journal of Process and Control*, vol. 20, pp. 143–157, 2010.
[21] L. Lang, W. Chen, B. R. Bakshi, P. K. Goel, and S. Ungarala, "Bayesian Estimation via Sequential Monte Carlo Sampling - Constrained Dynamic Systems," *Automatica*, vol. 43, pp. 1615–1622, 2007.
[22] J. Prakash, S. C. Patwardhan, and S. L. Shah, "Constrained Nonlinear State Estimation Using Ensemble Kalman Filters," *Industrial and Engineering Chemistry Research*, vol. 49, pp. 2242–2253, 2010.
[23] P. Stano, Zs. Lendek, and R. Babuška, "Saturated Particle Filters," in *Proceedings of the American Control Conference*, San Francisco, California, 2011, pp. 1819–1824.
[24] P. Stano, Zs. Lendek, J. Braaksma, R. Babuška, and C. de Keizer, "Particle Filters for Estimating Average Grain Diameter of Material Excavated by Hopper Dredger," in *Proceedings of IEEE Conference on Control Applications*, Yokohama, Japan, 2010, pp. 292 – 297.
[25] P. Vachhani, S. Narasimhan, and R. Rengaswamy, "Robust and Reliable Estimation via Unscented Recursive Nonlinear Dynamic Data Reconciliation," *Journal of Process Control*, vol. 16, pp. 1075–1086, 2006.
[26] W. Stadje, "A New Approach to the Lindley Recursion," *Statistics and Probability Letters*, vol. 31, pp. 169–175, 1997.
[27] B. O. S. Teixeira, L. A. B. Tôrres, L. A. Aguirre, , and D. S. Bernstein, "Unscented Filtering for Interval-Constrained Nonlinear Systems," in *Proceedings of the 47th IEEE Conference on Decision and Control*, Cancun, Mexico, 2008.
[28] B. O. Teixeira, L. A. Tôrres, L. A. Aguirre, and D. S. Bernstein, "On Unscented Kalman Filtering with State Interval Constraints," *Journal of Process Control*, vol. 20, pp. 45–57, 2010.
[29] O. Straka, J. Dun'k, and M. Šimandl, "Truncation Nonlinear Filters for State Estimation with Nonlinear Inequality Constraints," *Automatica*, vol. 48, pp. 273–286, 2012.
[30] O. Straka, J. Dunik, and M. Simandl, "Truncated Unscented Particle Filter," in *Proceedings of the American Control Conference*, San Francisco, California, 2011, pp. 1825–1830.
[31] I. Kyriakides, D. Morrel, and A. Papandreou-Suppappola, "A Particle Filtering Approach to Constrained Motion Estimation in Tracking Multiple Targets," in *The 39th Asilomar Conference on Signals, Systems and Computers*, vol. 28, Pacific Grove, California, 2005, pp. 94–98.
[32] H. Künsch, "Recursive Monte Carlo Filters: Algorithms and Theoretical Analysis," *The Annals of Statistics*, vol. 33, pp. 1983–2021, 2005.
[33] S. Meyn and R. Tweedie, *Markov Chains and Stochastic Stability*. Springer-Verlang London, 1993.
[34] A. Doucet, S. Godsill, and C. Andrieu, "On Sequential Monte Carlo Sampling Methods for Bayesian Filtering," *Statistics and Computing*, vol. 10, pp. 197–208, 2000.
[35] J. H. Kotecha and P. M. Djurić, "Gaussian Particle Filters," *IEEE Transactions on Signal Processing*, vol. 51, pp. 2592–2601, 2003.
[36] ——, "Gaussian Sum Particle Filters," *IEEE Transactions on Signal Processing*, vol. 51, pp. 2602–2612, 2003.
[37] H. Tanizaki and R. S. Mariano, "Prediction Filtering and Smoothing in Non-Linear and Non-Normal Cases Using Monte Carlo Integration," *Journal of Applied Econometrics*, vol. 9, pp. 163–179, 1994.
[38] N. J. Gordon, D. J. Salmond, and A. F. M. Smith, "Novel Approach to Nonlinear/Non-Gaussian Bayesian State Estimation," *IEE-Proceedings-F*, vol. 140, pp. 107–113, 1993.
[39] G. Kitagawa, "Monte Carlo Filter and Smoother for Non-Gaussian Nonllinear State Space Models," *Journal of Computational and Graphical Statistics*, vol. 5, pp. 1–25, 1996.
[40] H. Tanizaki and R. S. Mariano, "Nonlinear and Non-Gaussian State Space Modeling with Monte-Carlo Simulations," *Journal of Econometrics*, vol. 83, pp. 263–290, 1998.
[41] J. Carpenter, P. Clifford, and P. Fearnhead, "Improved Particle Filter for Nonlinear Problems," *IEE-Proceedings Radar, Sonar and Navigation*, vol. 146, pp. 2–7, 1999.
[42] X. Fu and Y. Jia, "An Improvement on Resampling Algorithms of Particle Filters," *IEEE Transactions on Signal Processing*, vol. 58, pp. 5414–5420, 2010.
[43] D. S. Lee and N. K. K. Chia, "A Particle Algorithm for Sequential Bayesian Parameter Estimation and Model Selection," *IEEE Transactions on Signal Processing*, vol. 50, pp. 326–336, 2002.
[44] A. Kong, J. S. Liu, and W. H. Wong, "Sequential Imputations and Bayesian Missing Data Problems," *Journal of the American Statistical Association,*, vol. 89, pp. 278–288, 1994.
[45] J. S. Liu and R. Chen, "Sequential Monte Carlo Methods for Dynamic Systems," *Journal of the American Statistical Association*, vol. 93, pp. 1032–1044, 1998.
[46] M. Vlasiou, I. Adan, and J. Wessels, "A Lindley-Type Equation Arising from a Carousel Problems," *Journal of Applied Probability*, vol. 41, pp. 1171–1181, 2004.
[47] F. Simonot, "Convergence Rate for the Distributions of GI/M/1/n and M/GI/1/n as n Tends to Infinity," *Journal of Applied Probability*, vol. 34, pp. 1049–1060, 1997.

LIST OF SYMBOLS

a_k	the update operator at time step k
b_k	the prediction operator at time step k
c^N	the multinomial sampling operator
$\bar{c}^N$	the resampling operator
f_k	state model at time step k
F_k	unsaturated state model at time step k
g_k	likelihood function at time step k
h_k	observation model at time step k
k_k^N	the particle filter at time step k
p_0	distribution of the initial state

q_i	the predicted probability of saturation of the *i*-th particle
q_i^α	the updated probability of saturation of the *i*-th particle
w_k	process noise at time step k
v_k	observation noise at time step k
x_k	state variable at time step k
x_k^i	the state estimate at time step k
$\hat{x}_k$	importance kernel at time step k
y_k	observation variable at time step k
C	saturation function
K_k	transition probability kernel at time step k
$\tilde{K}_k$	importance kernel at time step k
N	sample size
N_{eff}	effective sample size
N_{T}	resampling threshold
V_j	random variable on X
$\mathbb{P}_w$	probability with respect to state noise w
$\mathbb{P}_v$	probability with respect to observation noise v
α	detection function
$\delta(\cdot)$	Dirac delta
φ	characteristic function (continuous and bounded)
ν	probabilistic measure on X
ν_ϵ	probabilistic measure conditioned on Ω_ϵ
ω_k	weighted likelihood function at time step k
$\tilde{\omega}_k^i$	unnormalized weight associated with the *i*-th particle at time step k
ω_k^i	normalized weight associated with the *i*-th particle at time step k
$\pi_{k\|k}$	posterior measure of the state at time step k
$\pi_{k\|k}^N$	empirical approximation of size N of the posterior measure of the state at time step k
Ω_ϵ	ϵ-set, see Definition 11
Σ_X	sigma algebra on a vector space X
$\mathcal{P}(X)$	the space of probabilistic measures on (X, Σ_X)
$\mathcal{N}(\mu, \sigma^2)$	Gaussian distribution with mean μ and variance σ^2
$\mathcal{U}(a, b)$	uniform distribution on interval $[a, b]$
$\|\cdot\|_\infty$	supremum norm
erf	error function defined by $\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$