On the Convergence of Constrained Particle Filters

Nesrine Amor, Nidhal Carla Bouaynaya, Roman Shterenberg and Souad Chebbi

***Abstract*—The power of particle filters in tracking the state of non-linear and non-Gaussian systems stems not only from their simple numerical implementation but also from their optimality and convergence properties. In particle filtering, the posterior distribution of the state is approximated by a discrete mass of samples, called particles, that stochastically evolve in time according to the dynamics of the model and the observations. Particle filters have been shown to converge almost surely toward the optimal filter as the number of particles increases. However, when additional constraints are imposed, such that every particle must satisfy these constraints, the optimality properties and error bounds of the constrained particle filter remain unexplored. This paper derives performance limits and error bounds of the constrained particle filter. We show that the estimation error is bounded by the area of the state posterior density that does not include the constraining interval. In particular, the error is small if the target density is "well-localized" in the constraining interval.**

***Index Terms*—Constrained Particle Filter; Convergence.**

I. Introduction

MANY real-world applications, such as target tracking, (electric and renewable) power grids, navigation and chemical processes, can be formulated as a state-space model, where the state of the dynamical system is subject to additional constraints that arise from physical laws, natural phenomena or model restrictions [1]–[3]. These constraints cannot be incorporated into the state-space model easily [4]–[6].

The particle filter (PF) has been proven a powerful Monte Carlo approach for solving nonlinear and non-Gaussian state estimation problems [7]. The PF approximates the state posterior density using a set of weighted samples, called *particles*. This approximation converges, almost surely and in mean-square error, to the optimal posterior density of the state when the number of particles increases [8]. In comparison with standard approximation methods, such as the Extended Kalman Filter [1], the principle advantage of PF is that they do not rely on any local linearisation technique or any crude functional approximation. However, PF do not handle additional constraints on the state.

There are two major approaches that handle constraints within the PF framework: the acceptance/rejection method [9]–[11] and the constrained distribution approach [12], [13]. The acceptance/rejection approach enforces the constraints by simply rejecting the particles violating them [9], [11]. Although the acceptance/rejection procedure does not make any assumptions on the distributions and therefore maintains the generic properties of the particle filter, it is computationally inefficient as resources are wasted in drawing particles that may be rejected later on. Moreover, the number of samples will be reduced and hence the estimation accuracy may decrease, especially with a poor choice of the sampling density. An extreme example is when most (or all) of the particle violate the constraint and the algorithm fails [10]. Also, unconstrained sampling from a density followed by verification against constraints (especially nonlinear) may be computationally more demanding than sampling directly from the constrained region [11].

An alternative way to impose state constraints within the particle filter framework is to impose the constraints on all particles or equivalently sample from a constrained importance distribution [11], [13]–[15]. Assuming interval-type constraints, constraining every particle to be within the interval is equivalent to constraining the support of the posterior distribution to this interval. For this reason, this approach has been termed POintwise DEnsity Truncation (PoDeT) in [16], and we will be adopting this nomenclature in this paper. Although most constrained particle filtering methods adopt the PoDeT approach [12], [13], [17]–[19], there are no mathematical grounds, including optimality properties and convergence results, of PoDeT.

In this letter, we will investigate the optimality properties and the estimation error of the PoDeT approach. We will derive performance limits and errors bounds of this approach. In particular, we will show that if the posterior density is not "well-localized" within the constraining interval, then PoDeT will result in a large estimation error. On the other hand, if most of the posterior density lies within the constraining interval, then PoDeT will result in a bounded estimation error. Simulation resuls will support our fundamental theorems.

II. CONSTRAINED PARTICLE FILTERING

A. *Unconstrained State-Space Model*

Consider a discrete-time state-space model defined by a state transition and observation models:

$$\begin{aligned} \boldsymbol{x}_t &= f(\boldsymbol{x}_{t-1}) + \boldsymbol{u}_t, & (1)\\ \boldsymbol{y}_t &= h(\boldsymbol{x}_t) + \boldsymbol{v}_t, & (2) \end{aligned}$$

where f and h are, respectively, the state and observation potentially non-linear functions. $\boldsymbol{u}_t$ and $\boldsymbol{v}_t$ are the zero-mean transition and observation noise with known probability

N. Amor is with the National Superior School of Engineers of Tunis (ENSIT), University of Tunis, Tunis, Tunisia. When this work was done, she was a visiting scholar with the Department of Electrical and Computer Engineering, Rowan University, New Jersey, USA, e-mail: (nisrine.amor@hotmail.fr).

N. C. Bouaynaya is with the Department of Electrical and Computer Engineering, Rowan University, New Jersey, USA, e-mail: (bouaynaya@rowan.edu)

R. Shterenberg is with the Department of Mathematics, University of Alabama at Birmingham, Birmingham, AL, USA, e-mail: (shterenb@uab.edu).

S. Chebbi is with the National Superior School of Engineers of Tunis (ENSIT), University of Tunis, Tunis, Tunisia, e-mail: (chebbi.souad@gmail.com).

density functions, respectively, K_t and g_t. To simplify the notations, we will omit the time dependence t and write K, g.

Let (Ω, F, P) be a probability space where we define the stochastic processes $X = \{X_t, t \in \mathbb{N}\}$ and $Y = \{Y_t, t \in \mathbb{N} \setminus \{0\}\}$. let n_x and n_y be the dimensions of the state space of X and Y, respectively. The process X is a Markov process with initial distribution $X_0 \sim \mu(dx_0)$ and probability transition kernel $K(\boldsymbol{x}_t|\boldsymbol{x}_{t-1})$. The history of observations up to time t is denoted by $\boldsymbol{y}_{1:t} = [\boldsymbol{y}_1, \cdots, \boldsymbol{y}_t]$. In a Bayesian context, inference of $\boldsymbol{x}_t$ given a realization of the observations $\boldsymbol{y}_{1:t}$ relies upon the posterior density $p(\boldsymbol{x}_t|\boldsymbol{y}_{1:t})$. Using the Bayesian rule, we obtain the following two-step Bayesian recursion:

Prediction step:

$$p(\boldsymbol{x}_t|\boldsymbol{y}_{1:t-1}) = \int p(\boldsymbol{x}_{t-1}|\boldsymbol{y}_{1:t-1})\ K(\boldsymbol{x}_t|\boldsymbol{x}_{t-1})\ d\boldsymbol{x}_{t-1} \quad (3)$$

Update step:

$$p(\boldsymbol{x}_t|\boldsymbol{y}_{1:t}) = \frac{g(\boldsymbol{y}_t|\boldsymbol{x}_t)\ p(\boldsymbol{x}_t|\boldsymbol{y}_{1:t-1})}{\int g(\boldsymbol{y}_t|\boldsymbol{x}_t)\ p(\boldsymbol{x}_t|\boldsymbol{y}_{1:t-1})\ d\boldsymbol{x}_t}. \quad (4)$$

Unfortunately, for the nonlinear case, it is impossible to evaluate equations (3)-(4) in a closed-form expression, due to the fact that the integrals are generally intractable.

B. Unconstrained Particle Filtering

The *particle filter* (PF) is a sequential Monte Carlo method that estimates the posterior density of the state without making any assumptions about the probability density functions (pdfs) or the linearity of the system model [7], [20]. The particle filter approximates the posterior pdf by an ensemble of particles $\{\boldsymbol{x}_t^{(i)}\}_{i=1}^N$ and their associated weights $\{w_t^{(i)}\}$, i.e.,

$$p^N(\boldsymbol{x}_t|\boldsymbol{y}_{1:t}) = \sum_{i=1}^N w_t^{(i)} \delta(\boldsymbol{x}_t - \boldsymbol{x}_t^{(i)}) \quad (5)$$

where δ is the dirac delta function.

The particles are sampled from an *importance distribution*, also called a *proposal distribution*, $q(\boldsymbol{x}_t|\boldsymbol{x}_{t-1}, \boldsymbol{x}_t)$ because the true posterior is not available. To make up the difference between the proposal distribution and the posterior density, the importance weights are computed as [7], [21]

$$\tilde{w}_t^{(i)} = w_{t-1}^{(i)} \frac{g(\boldsymbol{y}_t|\boldsymbol{x}_t^{(i)}) K(\boldsymbol{x}_t^{(i)}|\boldsymbol{x}_{t-1}^{(i)})}{q(\boldsymbol{x}_t^{(i)}|\boldsymbol{x}_{t-1}^{(i)}, \boldsymbol{y}_t)}, \quad (6)$$

The weights are then normalized such that: $\sum_{i=1}^N w_t^{(i)} = 1$. The conditional mean estimate of the state is given by the weighted sample mean as:

$$\hat{\boldsymbol{x}}_t = E[\boldsymbol{x}_t|\boldsymbol{y}_{1:t}] \approx \sum_{i=1}^N w_t^{(i)} \boldsymbol{x}_t^{(i)}. \quad (7)$$

Moreover, for any integrable function φ, it is easy to approximate the integral (φ, p) such as

$$(\varphi, p) \approx (\varphi, p^N) = \sum_{i=1}^N w_t^{(i)} \varphi(\boldsymbol{x}_t^{(i)}) \quad (8)$$

C. Constrained Particle Filtering

Consider the state-space model given in (1)-(2) with the following additional constraint on the state

$$a_t \leq \phi_t(\boldsymbol{x}_t) \leq b_t. \quad (9)$$

where ϕ_t denotes the constraint function at time t. Without loss of generality, we can assume $a_t = a$ and $b_t = b$.

III. Optimality Properties

A. Error Bounds of Empirical Measures

Lemma 1. *Let μ be a probability measure on Ω, and I a set such that $\mu(I) \geq 1 - \eta$, $0 \prec \eta \prec \frac{1}{2}$. We denote by μ_I the truncation of μ onto I, i,e. for any set A, we have*

$$\widehat{\mu}(A) = \mu_I(A) = \frac{\mu(A \cap I)}{\mu(I)}. \quad (10)$$

Then, the variation of the signed measure $\mu_I - \mu$ satisfies

$$|\widehat{\mu} - \mu|(\Omega) \leq 2\eta. \quad (11)$$

Proof. We have

$$\begin{aligned} |\widehat{\mu} - \mu|(\Omega) &= |\widehat{\mu} - \mu|(I) + |\widehat{\mu} - \mu|(\Omega \setminus I) \\ &= 1 - \mu(I) + \mu(\Omega \setminus I) \leq 2\eta. \quad (12) \end{aligned}$$

□

Denote by $E = \mathcal{P}(\mathbb{R}^{n_x})$ the space of probability measures over $\mathbb{R}^{n_x}$. Let $C_b(\mathbb{R}^n)$ be the set of all continuous bounded functions on $\mathbb{R}^{n_x}$. We endow the space E with the topology of weak convergence. In this topology, we say that a sequence of probability measures μ_N converge (weakly) to μ if, for any $\varphi \in C_b(\mathbb{R}^{n_x})$, we have [8]

$$\lim_{N\to\infty} (\mu_N, \varphi_i) = (\mu, \varphi_i), \quad (13)$$

where $(\mu, \varphi) = \int \mu\varphi$.

Let $(a_t)_{t=1}^{\infty}$ and $(b_t)_{t=1}^{\infty}$ be two sequences of continuous functions on this space. We define $b_t : \mathcal{P}(\mathbb{R}^{n_x}) \to \mathcal{P}(\mathbb{R}^{n_x})$ to be the mapping such that for any $\nu \in \mathcal{P}(\mathbb{R}^{n_x})$, $b_t(\nu)$ is a probability measure defined as

$$(b_t(\nu), \varphi) = \int_{\mathbb{R}^{n_x}} \int_{\mathbb{R}^{n_x}} \varphi(\boldsymbol{x}_t) K(d\boldsymbol{x}_t|\boldsymbol{x}_{t-1}) \nu(d\boldsymbol{x}_{t-1}) = (\nu, K\varphi), \quad (14)$$

for any $\varphi \in C_b(\mathbb{R}^{n_x})$. $a_t : \mathcal{P}(\mathbb{R}^{n_x}) \to \mathcal{P}(\mathbb{R}^{n_x})$ is defined as the mapping such that

$$(a_t(\nu), \varphi) = (\nu, g)^{-1} (\nu, \varphi g), \quad (15)$$

for any $\varphi \in C_b(\mathbb{R}^{n_x})$.

Moreover, let $k_t = a_t \circ b_t$ be another sequence of continuous functions, where "$\circ$" denotes the composition of functions operator. Consider

$$\widehat{k}_t = T \circ a_t \circ T \circ b_t, \quad (16)$$

where T is the operator truncating every probability measure μ onto some set I, where $\mu(I) \geq 1 - \eta$. Note that I depends on μ.

Let $c^N : E \to E$ be a function (not necessarily continuous) that approximates a measure by N of its random samples.

$$\widehat{k}_t^N = c^N \circ T \circ a_t \circ c^N \circ T \circ b_t. \quad (17)$$

B. Optimal Stochastic Filtering

In the stochastic filtering framework, b_t denotes the map that takes $p(\boldsymbol{x}_{t-1}|\boldsymbol{y}_{1:t-1})$ to $p(\boldsymbol{x}_t|\boldsymbol{y}_{1:t-1})$, and a_t is the map that takes $p(\boldsymbol{x}_t|\boldsymbol{y}_{1:t-1})$ to $p(\boldsymbol{x}_t|\boldsymbol{y}_{1:t})$. Thus, k_t maps $p(\boldsymbol{x}_{t-1}|\boldsymbol{y}_{1:t-1})$ to $p(\boldsymbol{x}_t|\boldsymbol{y}_{1:t})$.

If the likelihood $g(\boldsymbol{y}_t|.)$ is a continuous bounded strictly positive function, then it can be easily shown that a_t is continuous. If the transition kernel K is Feller, i.e., for φ a continuous bounded function, $K\varphi$ is also a continuous bounded function, then b_t will be continuous [8]. We will make these assumptions on g and K in the remaining of the paper. In addition, we will assume that g is bounded from below by a strictly positive constant, i.e., $g \geq c_g > 0$ for some real $c_g > 0$. Subsequently, $k_t = a_t \circ b_t$ is also continuous. Intuitively, the continuity condition states that a slight perturbation of the previous posterior distribution of the signal $\boldsymbol{x}_{t-1}$ will also result in a small variation in the current posterior distribution of the signal $\boldsymbol{x}_t$.

In the context of stochastic filtering, the perturbation operator c^N, which maps a distribution to its discrete approximation in (5), is a random one. Let $c^{N,\omega}, \omega \in \Omega$ be the perturbation given by

$$c^{N,\omega}(\mu) = \frac{1}{N}\sum_{i=1}^{N} \delta_{\{\boldsymbol{z}_j(\omega)\}}, \tag{18}$$

where $\{\boldsymbol{z}_j\}$ are i.i.d. random variables distributed according to μ. It can be shown that that if $c^{N,\omega}$ is defined as in (18), then for almost all $\omega \in \Omega$, $c^{N,\omega}$ converges uniformly to the identity function [8].

The following Lemma states that a_t and b_t are not just continuous but Lipschitz continuous.

Lemma 2. *Consider two measures ν and μ such that $|((\nu-\mu),\varphi)| \leq \delta\|\varphi\|$ for any $\varphi \in C_b(\mathbb{R}^{n_x})$, then*

$$|(b_t(\nu) - b_t(\mu), \varphi)| \leq \delta\ \|\varphi\|\ C(K), \tag{19}$$

and

$$|(a_t(\nu) - a_t(\mu), \varphi)| \leq \delta\ \|\varphi\|\ C(g), \tag{20}$$

where $C(K)$ and $C(g)$ are constants that depend, respectively, on the kernel K and likelihood g.

Proof. Consider two measures ν and μ as stated in the Lemma. By definition of the mapping b_t in Eq. (14), we have

$$\begin{aligned} |(b_t(\nu) - b_t(\mu)), \varphi)| &= |(\nu, K\varphi) - (\mu, K\varphi)| && (21)\\ &\leq \delta\|K\varphi\| \leq \delta\|K\|\|\varphi\| && (22) \end{aligned}$$

where the first inequality in (22) follows from the fact that $|((\nu-\mu),\varphi)| \leq \delta\|\varphi\|$ for any $\varphi \in C_b(\mathbb{R}^{n_x})$. By letting $C(K) = \|K\|$, we obtain the Lipschitz property for b_t.

By definition of the mapping a_t in Eq. (15), we have

$$\begin{aligned} &|(a_t(\nu) - a_t(\mu)), \varphi)| \\ &= \left|\frac{(\nu,\varphi g)}{(\nu,g)} - \frac{(\mu,\varphi g)}{(\mu,g)}\right| && (23)\\ &\leq \left|\frac{(\nu-\mu,\varphi g)}{(\nu,g)}\right| + \left|\frac{(\mu,\varphi g)}{(\nu,g)(\mu,g)}\right| |(\nu-\mu),g)| && (24) \end{aligned}$$

We have $(\nu-\mu,\varphi g) \leq \delta\|\varphi g\|$. Since g is strictly positive and bounded from below by some strictly positive constant, $g \geq c_g > 0$, we have $(\nu, g) \geq c_g$. Hence, the first term in (24) is bounded by $\frac{\delta\|\varphi\|\|g\|}{c_g}$. Similarly, we have $\left|\frac{(\mu,\varphi g)}{(\nu,g)(\mu,g)}\right| \leq \frac{\|\varphi\|\|g\|}{c_g^2}$. We also have $|(\nu-\mu),g)| \leq \delta\|g\|$. Thus,

$$\left|\frac{(\nu-\mu,\varphi g)}{(\nu,g)}\right| + \left|\frac{(\mu,\varphi g)}{(\nu,g)(\mu,g)}\right| |(\nu-\mu),g)| \leq \delta C(g)\|\varphi\|, \tag{25}$$

where $C(g) = \frac{\|g\|}{c_g}\left[1 + \frac{\|g\|}{c_g}\right]$. □

Theorem 1. *Assuming that the transition kernel K is Feller and that the likelihood function g is continuous and bounded from below by a strictly positive constant, and considering a truncation operator T that truncates any probability distribution to a set I such that $\mu(I) \geq 1-\eta$. Then, for every $\varphi \in C_b(\mathbb{R}^{n_x})$, we have*

$$\limsup_{N\to\infty} |((\widehat{\nu}_t^N - \nu_t), \varphi)| \leq \eta\ C_t\ \|\varphi\|, \tag{26}$$

where C_t is a time-dependent constant.

Proof. The proof proceeds by induction. Assume that at time $t-1$, we have for every $\varphi \in C_b(\mathbb{R}^{n_x})$,

$$\limsup_{N\to\infty} |((\widehat{\nu}_{t-1}^N - \nu_{t-1}), \varphi)| \leq \eta\ C_{t-1}\ \|\varphi\|. \tag{27}$$

We notice that

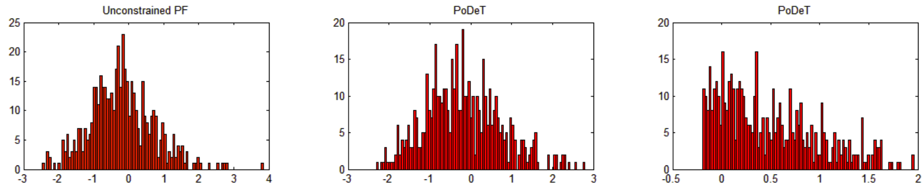
$$\begin{aligned} &|(\widehat{\nu}_t^N - \nu_t, \varphi)| \leq \\ +\ &|((c^N - id)\circ T \circ a_t \circ c^N \circ T \circ b_t(\widehat{\nu}_{t-1}^N), \varphi)| && (28)\\ +\ &|(id \circ (T - id) \circ a_t \circ c^N \circ T \circ b_t(\widehat{\nu}_{t-1}^N), \varphi)| && (29)\\ +\ &|(a_t \circ c^N \circ T \circ b_t(\widehat{\nu}_{t-1}^N) - a_t \circ id \circ T \circ b_t(\widehat{\nu}_{t-1}^N), \varphi)| \\ +\ &|(a_t \circ T \circ b_t(\widehat{\nu}_{t-1}^N) - a_t \circ id \circ b_t(\widehat{\nu}_{t-1}^N), \varphi)| && (30)\\ +\ &|(a_t \circ b_t(\widehat{\nu}_{t-1}^N) - a_t \circ b_t(\nu_{t-1}), \varphi)| && (31) \end{aligned}$$

From the uniform convergence of c^N to the identity id, we have that for any sequence of measures ν_N, $\lim_{N\to\infty} |(c^{N,\omega}(\nu_N),\varphi) - (\nu_N,\varphi))| = 0$ for almost all $\omega \in \Omega$ [8]. Hence, the first term in (28) goes to 0 as $N \to \infty$. The third term in (30) also goes to zero by the uniform convergence of c^N and the Lipschitz continuity of a_t in Lemma 2. From Lemma 1, $|(T - id)(\mu)| \leq 2\eta$ for any probability measure μ; hence, the second term in (29) is bounded by $2\eta\|\varphi\|$. From the Lipschitz continuity of a_t in Lemma 2 and the fact that $|(T-id)(\mu)| \leq 2\eta$, the fourth term in (30) is bounded by $2\eta C(g)\|\varphi\|$. From the induction assumption and Lemma 2, the fifth term (31) is bounded by $\eta C(g)C(K)C_{t-1}\|\varphi\|$. Thus, we obtain that

$$\limsup_{N\to\infty} |((\widehat{\nu}_t^N - \nu_t), \varphi)| \leq \eta\ C_t\ \|\varphi\|, \tag{32}$$

where $C_t = [2 + 2C(g) + C_{t-1}C(g)C(K)]\,\eta$. □

In Theorem (1), observe that η denotes the area of the state posterior density that does not include the constraining interval. Put simply, the PoDeT approach results in a bounded estimation error to the posterior density of the state if the target density is well-localized in the constraining interval $I = [a, b]$. In the one-dimensional case, a characterization of the localization of a distribution with respect to an interval I can be given in terms of the probability of the interval I: if $Pr\{[a, b]\} \geq 1 - \eta$, where $0 \leq \eta \ll 1$ is a small number,



(a) (b) (c)

Figure 1. (a): The unconstrained density of the system in (36) at time $n = 8$; the mean is -0.094; (b) First test case: the PoDeT posterior density, where the constraint interval is $[-2.8, 2.8]$; the mean is -0.0838. (c) Second test case: the PoDeT posterior density, where the constraint interval is $[-0.2, 2]$; the mean is 0.5124. Notice that, in the two test cases, the unconstrained mean naturally satisfies the constraints.

then the density is said to be well-localized. In particular, an important parameter that controls the estimation error of PoDeT is the area under the pdf delimited by the interval $[a, b]$. Intuitively, if high probability regions of the density are within the constraining interval, then the conditional mean estimate will be close to the truncated density at the support. In this case, the error in estimating the posterior distribution is small and can be quantified using the area under the tails of the well-localized density, i.e., the pdf area in the interval $]-\infty, a[\cup]b, \infty[$. Gaussian densities are obviously well-localized for properly chosen intervals; thus PoDeT may work well on truncation of Gaussian densities [1], [22], [23] provided the constraining interval occupies most of the density.

In the following Theorem, we establish the error estimate from below. We show that if the constraining interval is not mostly contained within the true density, then the PoDeT error will be bounded from below. Let $k_{1:t} = k_t \circ k_{t-1} \circ \cdots \circ k_1$ and $\widehat{k}^N_{1:t} = \widehat{k}^N_t \circ \widehat{k}^N_{t-1} \circ \cdots \circ \widehat{k}^N_1$. Denote by $\mu_t = k_{1:t}(\mu)$.

Theorem 2. *Consider a set I and let $\mu_t(I) \leq \eta$, then there exists a function $\varphi \in C_b(\mathbb{R}^{n_x})$ such that*

$$|((\widehat{k}^N_{1:t} - k_{1:t})(\mu), \varphi)| \geq \frac{1-\eta}{2}\|\varphi\|. \tag{33}$$

Proof. Notice that $\widehat{k}^N_{1:t}(\mu)$ is a probability measure supported on I. Thus, for every bounded continuous function φ with support outside of I, we have

$$|((\widehat{k}^N_{1:t} - k_{1:t})(\mu), \varphi)| = |(k_{1:t}(\mu), \varphi)| = |(\mu_t, \varphi)|. \tag{34}$$

Since by assumption $\mu_t(I) \leq \eta$, we have

$$\sup_{\varphi \in C_b(\mathbb{R}^{n_x}) \text{ supported outside of } I} \frac{|(\mu_t, \varphi)|}{\|\varphi\|} \geq 1 - \eta. \tag{35}$$

Hence, there exists a bounded continuous function φ supported outside of I such that $\frac{|(\mu_t,\varphi)|}{\|\varphi\|} \geq \frac{1-\eta}{2}$ or equivalently $|(\mu_t, \varphi)| \geq \frac{1-\eta}{2}\|\varphi\|$. ☐

Theorem (2) states that for non well-localized densities, the error of the PoDeT estimated density will be bounded from below. In particular, if the constraining interval covers a small area $\eta < 1/2$, then the density estimation error will be large, i.e., $\frac{1-\eta}{2} > 1/4$.

IV. Simulation Results

We consider the following non linear dynamical system:

$$\begin{cases} x_{t+1} = \frac{x_t}{2} + 25\frac{x_t}{1+x_t^2} + 8\cos(1.2t) + u_t, \\ y_t = \frac{x_t^2}{20} + v_t, \quad a_t \prec x_t \prec b_t. \end{cases} \tag{36}$$

This example is severely nonlinear [24], [25]. It was shown that the Extended Kalman Filter (EKF) fails in estimating the true state value of the unconstrained system [20], [26].

To assess the performance of PoDeT, we choose the constraint interval $[a_n, b_n]$, where the mean of the unconstrained posterior density naturally satisfies the constraint. We consider two cases: (i) most of the unconstrained posterior density lies within the constraint interval, thus well-localized; (ii) a high probability mass of the unconstrained posterior distribution lies outside of the constraint interval, thus not well-localized. We consider the posterior density at time $n = 8$. Test case (i): we choose the constraining interval $[a_8, b_8] = [-2.8, 2.8]$ (see Fig 1(b)). The unconstrained posterior density has mean $x_{true} \approx x_{unconstrained} = -0.094$ and the PoDeT mean estimate is $x_{PoDeT} = -0.0838$. Test case (ii): the constraining interval is chosen as $[a_8, b_8] = [-0.2, 2]$ (see Fig. 1(c)). PoDeT results in a truncated density with mean $= 0.5124$, which is further from the true mean (-0.094). PoDeT was able to estimate the mean of the well-localized case with a smaller error compared to the non-localized case. A real-world application on brain source localization from EEG data with additional constraints on the expected value of the state is presented in [16].

V. Conclusion

This paper addressed the optimality properties of PoDeT for constrained particle filtering. We discussed the error introduced when the particles are constrained to satisfy the boundary constraints, whereas the true density is not necessarily supported by the constraining interval. We showed that the PoDeT approach results in a bounded estimation error when the target density is "well localized" in the constraining interval (Theorem 1). On the other hand, PoDeT may lead to a large estimation error if the posterior density of the target is not well-localized (Theorem 2). In particular, unlike the unconstrained system, there are no convergence results of the PoDeT method. We hope that this paper incites more research into the performance limits of constrained particle filtering as well as the development of more algorithms that constrain the state estimate rather than the density itself.

VI. Acknowledgement

This work was supported by the National Science Foundation under Award Numbers NSF CCF-1527822 and NSF ACI-1429467.

REFERENCES

[1] D. Simon, *Optimal State Estimation: Kalman, H_∞, and Nonlinear Approaches*. Wiley, July 2006, p. 552.

[2] J. Hua, Z. Wang, B. Shen, and H. Gaoa, "Quantized recursive filtering for a class of nonlinear systems with multiplicative noises and missing measurements," *International Journal of Control*, vol. 86, no. 4, pp. 650–663, January 2013.

[3] J. Hua, Z. Wang, and H. Gaoa, "Recursive filtering with random parameter matrices, multiple fading measurements and correlated noises," *Automatica*, vol. 49, no. 11, pp. 3440–3448, November 2013.

[4] C. S. Agate and K. J. Sullivan, "Road-constrained target tracking and identification using a particle filter," in *Proceedings of SPIE*, vol. 5204, January 2004.

[5] C. Yang, M. Bakich, and E. Blasch, "Nonlinear constrained tracking of targets on roads," in *International Conference on Information Fusion*, February 2006.

[6] Y.-F. Huang, S. Werner, J. Huang, N. Kashyap, and V. Gupta, "State estimation in electric power grids: Meeting new challenges presented by the requirements of the future grid," *IEEE Signal Processing Magazine*, vol. 29, no. 5, pp. 33 – 43, September 2012.

[7] A. Doucet and A. M. Johansen, *Handbook of Nonlinear Filtering*, 2009, vol. 12, ch. A tutorial on particle filtering and smoothing: Fifteen years later, pp. 656–704.

[8] D. Crisan and A. Doucet, "A survery of convergence results on particle filtering methods for practitioners," *IEEE Transaction on Signal Processing*, vol. 50, no. 3, march 2002.

[9] L. Lang, W. S.Chen, B. R. Bakshi, P. K. Goel, , and S. Ungarala, "Bayesian estimation via sequential monte carlo sampling constrained dynamic systems," *Automatica*, vol. 43, no. 9, pp. 1615 – 622, September 2007.

[10] 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 Control*, vol. 20, no. 2, pp. 143–157, 2010.

[11] S. Ungarala, "A direct sampling particle filter from approximate conditional density function supported on constrained state space," *Computers & Chemical Engineering*, vol. 35, no. 6, p. 11101118, June 2011.

[12] F. Papi, M. Podt, Y. Boers, and G. Battistello, "On constraints exploitation for particle filtering based target tracking," in *International Conference on Information Fusion*, Singapore, July 2012, pp. 455 – 462.

[13] Z. Zhao, B. Huang, , and F. Liu, *Constrained Particle Filtering Methods for State Estimation of Nonlinear Process*. Wiley, June 2014, vol. 60, no. 6.

[14] O. Straka, J. Dunik, and M. Simandl, "Truncation nonlinear filters for state estimation with nonlinear inequality constraints," *Automatica*, vol. 48, no. 2, pp. 273–286, February 2012.

[15] J. Prakash, S. C. Patwardhan, and S. L. Shah, "On the choice of importance distributions for unconstrained and constrained state estimation using particle filter," *Journal of Process Control*, vol. 21, no. 1, pp. 3–16, January 2011.

[16] N. Amor, N. Bouaynaya, P. Georgieva, R. Shterenberg, and S. Chebbi, "EEG dynamic source localization using constrained particle filtering," in *International Conference on Symposium Series on Computational Intelligence (SSCI)*, Athens, Greece, December 2016.

[17] X. Wu and G. Huang, "Application of particle filter algorithm in nonlinear constraint optimization problem," in *International Conference on Natural Computation (ICNC)*, 2008.

[18] M. Yu, W. Chen, and J. Chambers, "Truncated unscented particle filter for dealing with non-linear inequality constraints," in *Sensor Signal Processing for Defence (SSPD)*, 2014.

[19] V. Pirard and E. Sviestins, "A robust and efficient particle filter for target tracking with spatial constraints," in *Information Fusion (FUSION)*, July 2013.

[20] N. Gordon, D. Salmond, and A. Smith, "Novel approach to nonlinear/non-Gaussian Bayesian state estimation," *IEE Proceedings in Radar and Signal Processing*, vol. 140, no. 2, pp. 107 – 113, 1993.

[21] N. Bouaynaya and D. Schonfeld, "On the optimality of motion-based particle filtering," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 19, no. 7, pp. 1068 – 1072, July 2009.

[22] D. Simon, "Kalman filtering with state constraints: a survey of linear and nonlinear algorithms," *IET Control Theory & Applications*, vol. 4, no. 8, pp. 1303 – 1318, August 2010.

[23] D. Simon and D. L. Simon, "Constrained Kalman filtering via density function truncation for turbofan engine health estimation," *International Journal of Systems Science*, vol. 41, no. 2, pp. 159–171, February 2010.

[24] M. 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, no. 2, pp. 174–188, February 2002.

[25] G. Kitagawa, "Monte carlo filter and smoother for non-gaussian nonlinear state space models," *Journal of Computational and Graphical Statistics*, vol. 5, no. 1, pp. 1–25, 1996.

[26] D. Simon and T. Chia, "Kalman filtering with state equality constraints," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 39, pp. 128–136, January 2002.