ON THE CONVERGENCE OF CONSTRAINED PARTICLE FILTERS

Rowan University

Tunis University

Nesrine Amor[1], Nidhal Carla Bouaynaya[2], Roman Shterenberg[3] and Souad Chebbi[4]

1,4 The Laboratory of Technologies of Information and Communication and Electrical Engineering (LaTICE) National Superior School of Engineers of Tunis (ENSIT), Tunis, Tunisia

2 Department of Electrical and Computer Engineering, Rowan University, New Jersey, USA

3 Department of Mathematics, University of Alabama at Birmingham, Birmingham, AL, USA

nisrine.amor@hotmail.fr, bouaynaya@rowan.edu, shterenb@uab.edu, chebbi.souad@gmail.com

Background

❖ 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. These constraints cannot be incorporated into the state-space model easily.

❖ ***The Particle Filter (PF)***

The particle filter (PF) has been proven a powerful Monte Carlo approach for solving nonlinear and non-Gaussian state estimation problems.

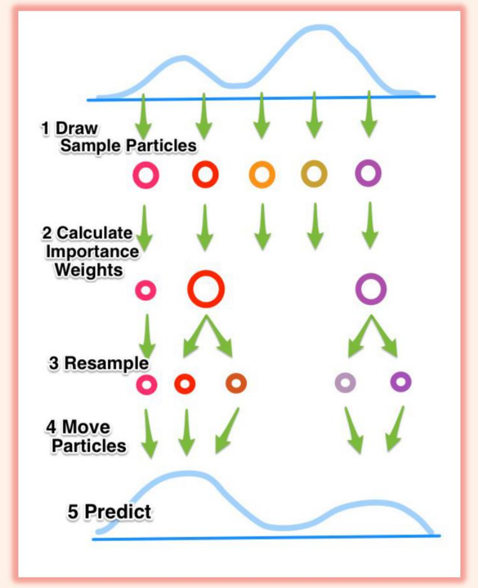
Main idea: Approximate the posterior density function of the state using a set of random samples, called *particles*, and their associated weights.

- The posterior pdf is approximated by a set of samples:

$$P(x_k \mid y_{1:k}) = \sum_{i=1}^{N} w_k^{(i)} \delta\left(x_k - x_k^{(i)}\right)$$

- The importance weights are given by:

$$w_k^{(i)} = w_{k-1}^{(i)} \frac{p(y_k \mid x_k^{(i)})\, p(x_k^{(i)} \mid x_{k-1}^{(i)})}{q(x_k^{(i)} \mid x_{k-1}^{(i)}, y_{1:k})}$$



How to integrate constraints within the PF?

The very numerical nature of the particle filters, which constitutes their strength for multidimensional numerical integration, becomes their major weakness in handling constraints on the state.

❖ ***State-of-the-art***

There are two major approaches that handle constraints within the PF framework:

- The acceptance/rejection approach enforces the constraints by simply rejecting the particles violating them. 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. 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.
- 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.

➡ ***Impose the constraints on all particles of the PF. This approach, however, is valid for hard constraints only. It actually constrains the posterior density of the state rather than its mean.***

Pointwise Density Truncation method (PoDeT)

➡ ***Although most constrained particle filtering methods adopt the PoDeT approach, there are no mathematical grounds, including optimality properties and convergence results, of PoDeT.***

❖ ***Proposed work***

- We investigated the optimality properties and the estimation error of the PoDeT approach.
- We derived performance limits and errors bounds of this approach.
- In particular, we showed 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.

❖ ***The constrained state-space model***

$$\begin{cases} x_t = f(x_{t-1}) + u_t \\ y_t = h(x_t) + v_t \end{cases}$$
$$a_t \le \Phi_t(\hat{x}_t) \le b_t$$

where ϕn is the constraint function at time *n* given by:

$$\phi_t(\hat{x}_t) = \phi_t\left(E[x_t \mid y_t]\right) \approx \phi_t\left(\sum_{i=1}^{N} w_t^{(i)} x_t^{(i)}\right)$$

Optimality Properties

Error Bounds of Empirical Measures

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

$$\hat{\mu}(A) = \mu_I(A) = \frac{\mu(A \cap I)}{\mu(I)}$$

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

$$|\hat{\mu} - \mu|(\Omega) \le 2\eta$$

Optimal Stochastic Filtering

In the stochastic filtering framework, b_t denotes the map that takes $p(x_{t-1} \mid y_{1:t-1})$ to $p(x_t \mid y_{1:t-1})$, and a_t is the map that takes $p(x_t \mid y_{1:t-1})$ to $p(x_t \mid y_{1:t})$. Thus, k_t maps $p(x_{t-1} \mid y_{1:t-1})$ to $p(x_t \mid y_{1:t})$.

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

$$|(b_t(\nu) - b_t(\mu), \varphi)| \le \delta \|\varphi\| C(K),$$

and

$$|(a_t(\nu) - a_t(\mu), \varphi)| \le \delta \|\varphi\| C(g).$$

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

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 $\mu(I) \ge 1-\eta$. Then, for every $\varphi \in C_b(\Re^{n_x})$, we have*

$$\lim_{N \to \infty} \sup \left|\left(\left(\hat{v}_t^N - v_t\right), \varphi\right)\right| \le \eta\, C_t \|\varphi\|$$

where C_t is a time-dependent constant.

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 $\Pr\{[a,b]\} \ge 1-\eta$, where $0 \le \eta << 1$ is a small number, 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 $]-\infty, a[\ \ ]b, \infty[$.

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 $k_{1:t} = k_t \ \ k_{t-1} \ \ ...k_1$ and $\hat{k}_{1:t}^N = k_t^N \ \ k_{t-1}^N \ \ ...k_1^N$. Denoted by $\mu_t = k_{1:t}(\mu)$.

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

$$\left|\left(\left(\hat{k}_{1:t}^N - k_{1:t}\right)(\mu), \varphi\right)\right| \ge \frac{1-\eta}{2}\|\varphi\|$$

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., $\frac{1-\eta}{2} > 1/4$.

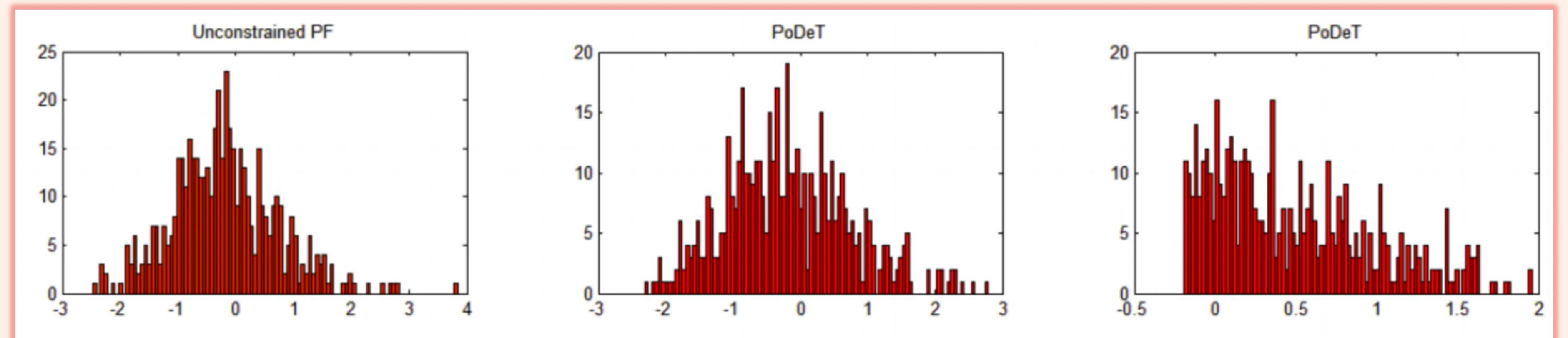
Simulation Results

➢ We consider the following non linear dynamical system

$$\begin{cases} x_{t+1} = \dfrac{x_t}{2} + 25\dfrac{x_t}{1+x_t^2} + 8\cos(1.2\,t) + u_t, \\ y_t = \dfrac{x_t^2}{20} + v_t, \qquad a_t \le x_t \le b_t. \end{cases}$$

➢ This example is severely nonlinear. It was shown that the Extended Kalman Filter (EKF) fails in estimating the true state value of the unconstrained system.

➢ 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.



(a) (b) (c)

❖ We consider the posterior density at time n = 8. Notice that, in the two test cases, the unconstrained mean naturally satisfies the constraints. 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 $x_{true} \approx x_{unconstraied} = -0.094$ 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.

Conclusion

➢ This work 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.

Acknowledgement

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