BAYESIAN NONPARAMETRIC STATE AND IMPULSIVE MEASUREMENT NOISE DENSITY ESTIMATION IN NONLINEAR DYNAMIC SYSTEMS

Nouha Jaoua[1,2], *Emmanuel Duflos*[1,2], *Philippe Vanheeghe*[1,2], *François Septier*[1,3]

[1] LAGIS UMR CNRS 8219, 59651 Villeneuve d'Ascq, France.
[2] Ecole Centrale de Lille, 59651 Villeneuve d'Ascq, France.
[3] Institut Mines-Télécom / Télécom Lille 1, 59658 Villeneuve d'Ascq, France.

ABSTRACT

In this paper, we address the problem of online state and measurement noise density estimation in nonlinear dynamic state-space models. We are especially interested in making inference in the presence of impulsive and multimodal noise. The proposed method relies on the introduction of a flexible Bayesian nonparametric noise model based on Dirichlet Process mixtures. A novel approach based on sequential Monte Carlo methods is proposed to perform the optimal online estimation. Simulation results demonstrate the efficiency and the robustness of this approach.

***Index Terms*—** Bayesian nonparametric, Dirichlet Process Mixture, particle filter, impulsive noise, α-stable process

1. INTRODUCTION

In signal processing literature, noise sources are often assumed to be Gaussian. However, in many fields the conventional Gaussian noise assumption is inadequate and can lead to the loss of resolution and/or accuracy. This is particularly the case of noise that exhibits impulsive nature. The latter is found in various areas [1–4]. In fact, impulsive noise tends to produce large amplitude excursions from the average value more frequently than Gaussian signals. It contains sharp spikes or occasional bursts. As a result, its probability density function (pdf) decays in the tails less rapidly than Gaussian pdf [5]. α-stable distribution is suitable for modeling this type of noise since its tails are heavier than those of Gaussian distribution.

Stable distributions stem from the generalized central limit theorem. Thus, they can be seen as a generalization of the Gaussian distribution. One difficulty is that they have no closed-form expressions for their pdf. They can be most conveniently described by their characteristic function [5]:

$$\varphi(t) = \begin{cases} \exp\left(i\mu t - \gamma^\alpha |t|^\alpha \left[1 - i\beta\,\mathrm{sgn}(t)\tan\frac{\alpha\pi}{2}\right]\right), & \alpha \neq 1 \\ \exp\left(i\mu t - \gamma |t|^\alpha \left[1 + i\beta\,\mathrm{sgn}(t)\frac{2}{\pi}\log|t|\right]\right), & \alpha = 1 \end{cases}$$

Thus, the stable distribution is completely determined by four parameters α, β, γ and μ. To denote an α-stable distribution with the parameters α, β, γ and μ, we will use the following notation $S_\alpha(\beta, \gamma, \mu)$. A more detailed description of α-stable distributions can be found in [5, 6].

In this work, we focus on the challenging task of the online estimation of the state and the unknown measurement noise density in nonlinear dynamic systems using sequential Monte Carlo (SMC) methods. The main difficulty of the problem arises from the fact that the estimation of the state depends on the unknown noise statistics and vice versa. In addition, we consider here that the measurement noise is a mixture of α-stable distributions. In such context, the considered task is even more difficult due to the lack of an analytical expression for the pdf of α-stable distributions. Therefore, only few works consider this issue and, furthermore most of these studies consider only unimodal noise [7, 8]. Studies on finite mixture of α-stable distributions are very limited [9–11]. These works focus on making inference on density parameters of the α-stable mixture without considering the state estimation. In addition, they treat the problem by using parametric methods with a finite number of parameters. In such models, the main difficulty is the choice of the number of mixture components. In [9], this number is supposed known which is generally not true in real-world applications contexts, but in [10, 11], the number of components is assumed unknown *a priori* and is estimated using Reversible Jump Markov Chain Monte Carlo (RJMCMC). It is theoretically desirable to consider models that are not limited to finite parameterizations. Our method is based on a flexible Bayesian nonparametric Dirichlet Process Mixture (DPM) model which allows straightforward estimation of the number of components without requiring RJMCMC-like computational approaches. Furthermore, by using such a model, the problem of the non-existence of an analytical expression for the pdf of α-stable distributions is surmounted. However, approaches proposed in [10, 11] require the numerical evaluation of the α-stable distribution at every iteration. In that case, one should resort to computationally intensive techniques to perform the inversion of the characteristic function of the α-stable distribution via the fast Fourier transform (FFT). In literature, studies involving SMC methods and nonparametric density estimation using DPMs have been conducted [12, 13]. In [12], this issue is addressed considering a dynamic linear system. The proposed particle filter (PF) is not able to change the mean and the covariance after they are initialized. In our formulation, we overcome this problem using an efficient importance density which takes into account the current observation. Moreover, in both [13] and [12], the unknown noise is assumed to be a mixture of Gaussian. In this paper, we consider a mixture of α-stable distribution which is a more general distribution including the Gaussian and the Cauchy ones.

This paper is organized as follows: in section 2 we briefly review Bayesian nonparametric density estimation using DPMs. In section 3, we introduce the dynamic model as well as the measurement noise modeling. Section 4 is devoted to the description of the proposed particle filter (PF). Simulation results are presented in section 5 and conclusions are drawn through section 6.

2. BAYESIAN NONPARAMETRIC DENSITY ESTIMATION USING DPMS

Consider a set of observations $\{z_k\}_{k=1}^n$ statistically distributed according to an unknown pdf F such as $z_k \sim F(.)$, $k = 1, ..., n$. We are interested in estimating the pdf $F(.)$ based on the sequence of observations $\{z_k\}_{k=1}^n$. To this purpose, we consider the following nonparametric model

$$F(y) = \int_{\Theta} f(z|\theta) d\mathbb{G}(\theta) \quad (1)$$

where $\theta \in \Theta$ is called the latent variable or cluster, $f(.|\theta)$ is the mixed pdf and $\mathbb{G}(.)$ is the mixing distribution. Under a Bayesian framework, $\mathbb{G}$ is assumed to be a Random Probability Measure (RPM) distributed according to a prior distribution. In this paper, we will select as a RPM the Dirichlet Process (DP) prior.

The DP was introduced by Ferguson [14] as a probability measure in the space of probability measures. It is characterized by two parameters: a base distribution $\mathbb{G}_0$ and a concentration parameter α; we denote it by $DP(\mathbb{G}_0, \alpha)$. The DP has a number of properties which make inference based on this nonparametric prior computationally tractable. An important property is that the realizations of a DP are discrete with probability one. Sethuraman [15] showed that with probability one, a random draw $\mathbb{G} \sim DP(\mathbb{G}_0, \alpha)$ can be expressed as:

$$\mathbb{G} = \sum_{k=1}^{\infty} \pi_k \delta_{\theta_k} \quad (2)$$

with $\theta_k \sim \mathbb{G}$, $\pi_k = \beta_k \prod_{j=1}^{k-1}(1-\beta_j)$ and $\beta_k \sim \mathcal{B}(1, \alpha)$ where $\mathcal{B}(.,.)$ denotes the beta distribution. Using (1), the unknown density of interest can be rewritten as:

$$F(z) = \sum_{k=1}^{\infty} \pi_k f(z|\theta_k) \quad (3)$$

The discreteness property is very attractive since it allows straightforward estimation of the number of components.
Another appealing feature of the DP is the so-called Polya urn representation that results from integrating over the underlying RPM $\mathbb{G}$ and provides the predictive distribution as follows [16]:

$$\theta_{n+1}|\theta_{1:n} \sim \frac{\alpha}{\alpha+n}\mathbb{G}_0 + \frac{1}{\alpha+n}\sum_{i=1}^{n-1}\delta_{\theta_i} \quad (4)$$

This representation is very useful in practice since it allows sampling observations from a DP without explicitly constructing the RPM $\mathbb{G} \sim DP(\mathbb{G}_0, \alpha)$. In fact, conditionally on the latent variable previously sampled $\theta_{1:n}$, a new sample is equal to an existing one with probability proportional to the number of times this parameter was previously observed, or is independently sampled from $\mathbb{G}_0$ with probability proportional to α.

The DPM model is based on DP prior for the mixing distribution. Such a model assumes that the RPM $\mathbb{G}$ itself is aleatory, drawn from a DP. The DPM defines the following hierarchical Bayesian structure:

$$\begin{aligned}
&\mathbb{G} \sim DP(\mathbb{G}_0, \alpha), \\
&\theta_k|\mathbb{G} \sim \mathbb{G}, \quad k = 1, ..., n \\
&z_k|\theta_k \sim f(.|\theta_k), \quad k = 1, ..., n
\end{aligned}$$

DPM is a widely used model for density estimation and is among the most successful ways of modeling multimodal distributions in a nonparametric Bayesian framework. In fact, the parameter with which an observation is associated implicitly clusters the data. In addition, the DP provides a prior that makes it more likely to associate an observation with a parameter to which other observations have already been associated. This reinforcement property is essential for inferring finite mixture models. It can be shown under mild conditions that if the data were generated by a finite mixture, then the DP posterior is guaranteed to converge (in distribution) to that finite set of mixture parameters [17].

3. DPM NOISE MODEL

Consider the following generic nonlinear dynamic system given in state-space form:

$$\begin{cases} x_{t+1} = g_t(x_t, w_t) \\ y_t = h_t(x_t, v_t) \end{cases} \quad (5)$$

where t is the time index, x_t is the state variable, y_t is the measurement, g_t and h_t are respectively the state and the observation functions, and w_t and v_t are mutually independent i.i.d noise processes. Such nonlinear dynamic systems are widely used to model systems across many areas in signal processing such as target tracking, communications, etc. Here, we assume that the distribution of the process noise is known. The measurement noise is assumed to be impulsive, skewed and multimodal with an unknown distribution.

In this paper, we suppose that the measurement noise distribution F is a DPM with a heavy-tailed kernel that is the Cauchy distribution. The mixed pdf $f(.|\theta_t)$ is thus assumed to be a Cauchy distribution with location parameter m_t and scale parameter c_t denoted $\mathcal{C}(m_t, c_t)$. We place a Gaussian prior with mean ξ and covariance κ^{-1}, $\mathcal{N}(\xi, \kappa^{-1})$, on the parameter m_t. For the parameter c_t, an inverse Gamma prior with shape a and scale b, $\mathcal{IG}(a, b)$, is chosen. We denote α the scale parameter of the DPM. The base distribution $\mathbb{G}_0$ can be defined as the product of the parameters priors:

$$\mathbb{G}_0 \sim \mathcal{N}(\xi, \kappa^{-1}) \times \mathcal{IG}(a, b) \quad (6)$$

Finally, we obtain the following DPM model for the measurement noise distribution:

$$\begin{aligned}
&\mathbb{G}|\Phi \sim DP(\mathbb{G}_0, \alpha) \\
&\theta_t|\mathbb{G} \sim \mathbb{G} \\
&v_t|\theta_t \sim \mathcal{C}(m_t, c_t)
\end{aligned} \quad (7)$$

where $\theta_t = \{m_t, c_t\}$ is the latent variable giving at each time index t the location and the scale of the cluster and $\Phi = \left\{\alpha, \xi, \kappa^{-1}, a, b\right\}$ denotes the set of hyperparameters. We assume here that these hyperparameters are pre-specified and fixed. This model can be rewritten as $v_t \sim F(v_t)$ where $F(v_t)$ is the measurement noise distribution expressed as follows:

$$F(v_t) = \int \mathcal{C}(v_t; m, c) d\mathbb{G}(m, c) \quad (8)$$

4. PARTICLE FILTER FOR SEQUENTIAL STATE AND NOISE DENSITY ESTIMATION

In this paper, our main goal is to jointly estimate the state x_t and the latent variable θ_t at each time t conditional on the observations $y_{1:t}$. Within a Bayesian framework, we need to compute the joint posterior pdf $p(x_t, \theta_t|y_{1:t}, \Phi)$. Unfortunately, this pdf is analytically intractable. Therefore, we propose to use SMC methods in order to find an estimate of the required posterior pdf. The set of hyperparameters Φ is assumed to be known, therefore it is omitted in the following. The posterior pdf $p(x_t, \theta_t|y_{1:t})$ is approximated by a PF:

$$p(x_t, \theta_t|y_{1:t}) \simeq \sum_{i=1}^{N} \omega_t^{(i)} \delta_{x_t^{(i)}, \theta_t^{(i)}}(x_t, \theta_t) \quad (9)$$

where δ is the Dirac delta function, $x_t^{(i)}$ and $\theta_t^{(i)}$ are respectively the state and the cluster particles drawn from the importance density $q(x_t, \theta_t | x_{0:t-1}, \theta_{1:t-1,}, y_{1:t})$ and $\omega_t^{(i)}$ is the normalized importance weight asociated to the ith particle.

Once the posterior density function of interest is identified, the remaining task is the simulation of state and cluster particles from the importance density. The choice of the importance density is crucial because it determines the efficiency as well as the complexity of the PF. In this paper, we consider the optimal importance density [18]. In our context, it is expressed as

$$q(x_t, \theta_t | x_{0:t-1}, \theta_{1:t-1,}, y_{1:t}) = p(x_t, \theta_t | x_{0:t-1}, \theta_{1:t-1,}, y_t) \quad (10)$$

This importance density is interesting because it incorporates information on the current observation. Consequently, the particles tend to cluster in regions of high probability mass of the posterior pdf.

The sampling of $x_t^{(i)}$ and $\theta_t^{(i)}$ from (10) requires the analytical expression of the optimal importance density. However, this pdf is analytically intractable. Using Bayes'theorem, the considered importance density can be written as

$$\begin{aligned} &p(x_t, \theta_t | x_{0:t-1}, \theta_{1:t-1}, y_t) \\ &\quad = \frac{p(y_t | x_{0:t}, \theta_{1:t}) p(x_t | \theta_{1:t}, x_{0:t-1}) p(\theta_t | \theta_{1:t-1})}{p(y_t | x_{0:t-1}, \theta_{1:t-1})} \\ &\quad \propto p(y_t | x_t, \theta_t) p(x_t | \theta_t, x_{t-1}) p(\theta_t | \theta_{1:t-1}) \end{aligned} \quad (11)$$

Thus, an approximation of the optimal importance density can be obtained using Monte Carlo method and importance sampling. For this purpose, we consider, for each trajectories pair $\left(x_{0:t-1}^{(i)}, \theta_{1:t-1}^{(i)}\right)$, a set of N_{IS} auxiliary particles $\left\{(\breve{x}_{t,i}^{(j)}, \breve{\theta}_{t,i}^{(j)})\right\}_{j=1}^{N_{IS}}$ where state particles $\breve{x}_{t,i}^{(j)}$ are sampled using the transition density $p(x_t | x_{0:t-1}^{(i)})$:

$$\breve{x}_{t,i}^{(j)} \sim p(x_t | x_{t-1}^{(i)}) \quad (12)$$

and latent variable particles $\breve{\theta}_{t,i}^{(j)}$ are sampled using the predictive distribution $p(\breve{\theta}_{t,i}^{(j)} | \theta_{1:t-1}^{(i)})$ which admits the following Polya urn representation (cf. section 2):

$$\breve{\theta}_{t,i}^{(j)} | \theta_{1:t-1}^{(i)} \sim \frac{\alpha}{\alpha + t} G_0 + \frac{1}{\alpha + t} \sum_{k=1}^{t-1} \delta_{\theta_k^{(i)}}(\theta_k) \quad (13)$$

Using this set of particles, the optimal importance density can be approximated by the following empirical distribution:

$$p(x_t, \theta_t | x_{0:t-1}^{(i)}, \theta_{1:t-1,}^{(i)}, y_t) \simeq \sum_{j=1}^{N_{IS}} \frac{\breve{\omega}_{t,i}^{(j)}}{S_{\breve{\omega}}} \delta_{\breve{x}_{t,i}^{(j)}, \breve{\theta}_{t,i}^{(j)}}(x_t, \theta_t) \quad (14)$$

where $\breve{\omega}_{t,i}^{(j)}$ is the unnormalized weight associated to the jth pair of particles $(\breve{x}_{t,i}^{(j)}, \breve{\theta}_{t,i}^{(j)})$ defined as

$$\breve{\omega}_{t,i}^{(j)} = p(y_t | \breve{x}_{t,i}^{(j)}, \breve{\theta}_{t,i}^{(j)}) \quad (15)$$

and $S_{\breve{\omega}}$ is the sum of unnormalized weights $S_{\breve{\omega}} = \sum_{j=1}^{N_{IS}} \breve{\omega}_{t,i}^{(j)}$. In order to sample the ith pair of particles $(x_t^{(i)}, \theta_t^{(i)})$ from the approximate optimal importance density given by (14), we just need to pick one particle from the set $\left\{(\breve{x}_{t,i}^{(j)}, \breve{\theta}_{t,i}^{(j)})\right\}_{j=1}^{N_{IS}}$ using weights $\left\{\breve{\omega}_{t,i}^{(j)}\right\}_{j=1}^{N_{IS}}$ as probabilities of selection. This can be done as follows:

$$J \sim Multinomial\left(\frac{\breve{\omega}_{t,i}^{(1)}}{S_{\breve{\omega}}}, \frac{\breve{\omega}_{t,i}^{(2)}}{S_{\breve{\omega}}}, ..., \frac{\breve{\omega}_{t,i}^{(N_{IS})}}{S_{\breve{\omega}}}\right) \quad (16)$$

Thus, the ith pair of particles $(x_t^{(i)}, \theta_t^{(i)})$ is given by

$$\left(x_t^{(i)}, \theta_t^{(i)}\right) = \left(\breve{x}_{t,i}^{(j=J)}, \breve{\theta}_{t,i}^{(j=J)}\right) \quad (17)$$

Using the importance density given by (14), the importance weights are updated according to the following relation:

$$\omega_t^{(i)} \propto \omega_{t-1}^{(i)} p(y_t | x_{0:t-1}^{(i)}, \theta_{1:t-1,}^{(i)}) \quad (18)$$

The pdf $p(y_t | x_{0:t-1}^{(i)}, \theta_{1:t-1,}^{(i)})$ can be written as:

$$\begin{aligned} &p(y_t | x_{0:t-1}, \theta_{1:t-1}) \\ &\quad = \int p(y_t | x_t, \theta_t) p(x_t | \theta_{1:t}, x_{0:t-1}) p(\theta_t | \theta_{1:t-1}) dx_t d\theta_t \end{aligned} \quad (19)$$

This integral can be approximated using the weighted set of particles of the importance sampling strategy. In doing so, $p(y_t | x_{0:t-1}^{(i)}, \theta_{1:t-1}^{(i)})$ is given by the sum of unnormalized weights $S_{\breve{\omega}}$.

The proposed PF for joint state and noise density estimation denoted by PF-JSNDE is summarized in algorithm 1.

Initialization
for $i = 1$ **to** N **do**
  Sample $x_0^{(i)} \sim p_0(x_0)$;
  Sample $\theta_0^{(i)} \sim \mathbb{G}_0$;
  Initialize $\omega_0^{(i)} = 1/N$;
end
for $t = 1$ **to** T **do**
  for $i = 1$ **to** N **do**
    for $j = 1$ **to** N_{IS} **do**
      Sample $\theta_t^{(j)} \sim p(\theta_t | \theta_{0:t-1}^{(i)})$ using (13);
      Sample $x_t^{(i)} \sim p(x_t | x_{0:t-1}^{(i)})$ using (12) ;
      Compute weights: $\breve{\omega}_{t,i}^{(j)}$ using (15) ;
    end
    Compute: $S_{\breve{\omega}} = \sum_{j=1}^{N_{IS}} \breve{\omega}_{t,i}^{(j)}$;
    Normalize weights: $\breve{\omega}_{t,i}^{(j)} = \breve{\omega}_{t,i}^{(j)} / S_{\breve{\omega}}$;
    Select a particle indice $J \in \{1, 2, ..., N_{IS}\}$ according to weights $\left\{\breve{\omega}_{t,i}^{(j)}\right\}_{j=1}^{N_{IS}}$ using (16);
    Set $x_t^{(i)} = \breve{x}_{t,i}^{(J)}$ and $\theta_t^{(i)} = \breve{\theta}_{t,i}^{(J)}$;
    Compute importance weights: $\omega_t^{(i)} \propto \omega_{t-1}^{(i)} S_{\breve{\omega}}$;
  end
  Normalize importance weights
  $\omega_t^{(i)} = \omega_t^{(i)} / \sum_{j=1}^{N} \omega_t^{(i)}, \quad i = 1, ..., N$;
  if $N_{eff} < \eta$ **then** Resampling step;
end

Algorithm 1: PF-JSNDE algorithm

5. SIMULATIONS

The performance of the proposed method is studied considering the following nonlinear time series model which has been extensively used in literature for benchmarking numerical filtering techniques [19–21]:

$$\begin{cases} x_{t+1} = 0.5x_t + 25\dfrac{x_t}{1+x_t^2} + 8\cos(1.2(t+1)) + w_t \\ y_t = \frac{x_t^2}{20} + v_t \end{cases}$$

This model has been simulated with the following parameters: $x_0 \sim \mathcal{N}(0,10)$, $w_t \sim \mathcal{N}(0,1)$, $F = 0.3\mathcal{S}_{1.3}(0,2,-10) + 0.7\mathcal{S}_{1.6}(0.5,1.5,0)$. The hyperparameters of the base distribution ξ, κ^{-1}, a, b are respectively set to 0, 50, 5 and 4. We fixed the scale parameter of the DPM α to 3. The proposed PF has been implemented with $N = 200$ particles and $N_{IS} = 100$ auxiliary particles. Results are illustrated in the different plots of Fig. 1, Fig. 2 and Fig. 3. Fig. 1 shows the estimated signal as well as the true and the observed ones. We also plot the measurement noise signal and the estimation error between the true and the estimated signals. From these plots, it can be seen that despite the fact that the noise is important, the state x_t is correctly estimated. Fig. 2 depicts the estimated measurement noise density as well as the true one. We can observe that the estimated pdf is close to the true one. In Fig. 3, the evolution over time of the Kullback-Leibler distance between the true and the estimated noise density is reported for each time index. It can be remarked that the Kullback-Leibler distance converges to a low value close to zero.

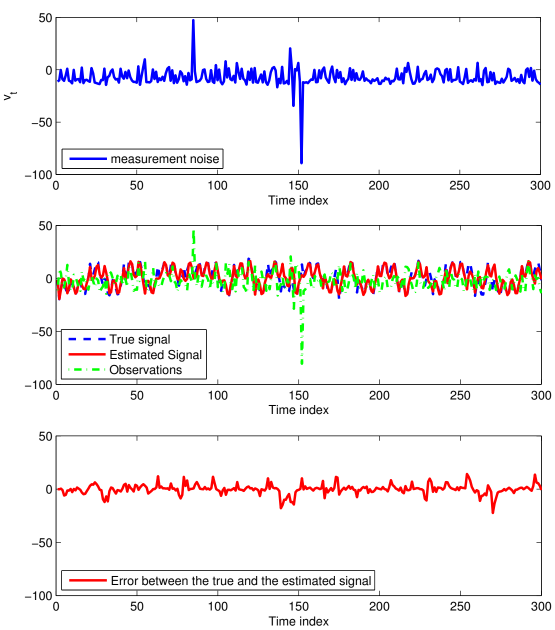

Fig. 1. Top picture: Measurement noise signal. Middle picture: True (dashed line), estimated (solid line) and observed (dash-dotted line) signals. Bottom picture: Error between the true and the estimated states.

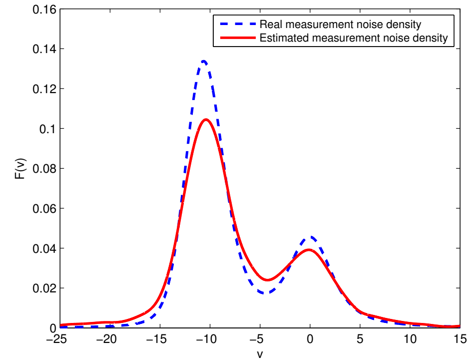

Fig. 2. True (dashed line) and estimated (solid line) noise density at time t=300.

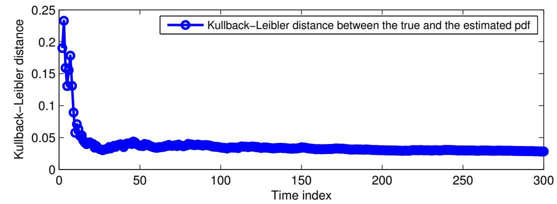

Fig. 3. Evolution over time of the Kullback-Leibler distance between the true and the estimated noise density.

6. CONCLUSIONS

In this paper, we present a novel approach allowing the joint estimation of state and measurement noise density in nonlinear dynamic systems using SMC methods. The measurement noise considered here is an α-stable process. A flexible Bayesian nonparametric noise model based on DPMs is introduced. The originality of this work consists in using DPM to model an α-stable process as an infinite mixture of Cauchy distributions. Based on the performed simulations, it can be concluded that the state estimation is possible even if the measurement noise density is unknown. Furthermore, we have shown that a DPM of Cauchy distributions is well suited to model an α-stable process. This result is very interesting since it provides a solution to the problem of the non-existence of an analytical expression for the pdf of α-stable distributions. Finally, it should be noted that the proposed method is not limited to the case of noise distributed according to a mixture of α-stable distributions and can be applied to other types of noise. In future works, we plan to make inference on the hyperparameters of the base distribution.

7. REFERENCES

[1] M. Shinde and S. Gupta, "Signal detection in the presence of atmospheric noise in tropics," *IEEE Transactions on Communications*, vol. 22, no. 8, pp. 1055–1063, aug 1974.

[2] D. Middleton, "Statistical-physical models of electromagnetic interference," *IEEE Transactions on Electromagnetic Compatibility*, vol. EMC-19, no. 3, pp. 106 –127, aug. 1977.

[3] M. Zimmermann and K. Dostert, "Analysis and modeling of impulsive noise in broad-band powerline communications," *IEEE Transactions on Electromagnetic Compatibility*, vol. 44, no. 1, pp. 249–258, feb 2002.

[4] J. Ilow and D. Hatzinakos, "Impulsive noise modeling with stable distributions in fading environments," in *8th IEEE Signal Processing Workshop on Statistical Signal and Array Processing*, jun 1996, pp. 140 –143.

[5] C.L. Nikias and M. Shao, *Signal Processing with Alpha-Stable Distributions and Applications*, John Wiley and Sons, 1995.

[6] N. Jaoua, E. Duflos, P. Vanheeghe, L. Clavier, and F. Septier, "Impulsive interference mitigation in ad hoc networks based on alpha-stable modeling and particle filtering," in *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, may 2011, pp. 3548 –3551.

[7] J. Vil-Valls, C. Fernandez Prades, P. Closas Gmez, and J. Fernndez Rubio, "Bayesian filtering for nonlinear state-space models in symmetric alpha-stable measurement noise," in *European Signal Processing Conference*, 2011 September, pp. 674 – 678.

[8] M. Lombardi and S.J. Godsill, "On-line Bayesian estimation of signals in symmetric alpha-stable noise," *IEEE Transactions on Signal Processing*, vol. 54, pp. 775–779, February 2006.

[9] R. Casarin, "Bayesian inference for mixtures of stable distributions," 2004.

[10] D. Salas-Gonzalez, E.E. Kuruoglu, and D.P. Ruiz, "Finite mixture of alpha-stable distributions," *Digit. Signal Process.*, vol. 19, no. 2, pp. 250–264, Mar. 2009.

[11] D. Salas-Gonzalez, E.E. Kuruoglu, and D.P. Ruiz, "Modelling with mixture of symmetric stable distributions using gibbs sampling," *Signal Processing*, vol. 90, no. 3, pp. 774 – 783, 2010.

[12] F. Caron, M. Davy, A. Doucet, E. Duflos, and P. Vanheeghe, "Bayesian inference for linear dynamic models with dirichlet process mixtures," *IEEE Transactions on Signal Processing*, vol. 56, no. 1, pp. 71 –84, jan. 2008.

[13] E. Ozkan, S. Saha, F. Gustafsson, and V. Smidl, "Non-parametric bayesian measurement noise density estimation in non-linear filtering," in *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, may 2011, pp. 5924 –5927.

[14] T.S. Ferguson, "A Bayesian Analysis of Some Nonparametric Problems," *The Annals of Statistics*, vol. 1, no. 2, pp. 209–230, 1973.

[15] J. Sethuraman, "A constructive definition of Dirichlet priors," *Statistica Sinica*, vol. 4, pp. 639–650, 1994.

[16] D. Blackwell and J.B. Macqueen, "Ferguson distributions via Pólya urn schemes," *The Annals of Statistics*, vol. 1, pp. 353–355, 1973.

[17] H. Ishwaran and M. Zarepour, "Dirichlet prior sieves in finite normal mixtures.," Statistica Sinica 12, 941-963 (2002)., 2002.

[18] V. Zaritskii, V. Svetnik, and L. Shimelevich, "Monte carlo technique in problems of optimal data processing," *Automation and Remote Control*, vol. 12, pp. 95 –103, 1975.

[19] 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.

[20] M. West, "Mixture models, Monte Carlo, Bayesian updating and dynamic models.," *Computing Science and Statistics*, vol. 24, pp. 325–333.

[21] N.J. Gordon, D.J. Salmond, and A.F.M. Smith, "Novel approach to nonlinear/non-Gaussian Bayesian state estimation," in *Radar and Signal Processing, IEE Proceedings F*, Apr. 1993, vol. 140, pp. 107–113.