RELEVANCE OF DIRICHLET PROCESS MIXTURES FOR MODELING INTERFERENCES IN UNDERLAY COGNITIVE RADIO

Vincent Pereira, Guillaume Ferré, Audrey Giremus and Eric Grivel

Bordeaux 1 University - IPB - UMR CNRS 5218 IMS
351 cours de la Libération 33405, TALENCE Cedex, France

ABSTRACT

In the field of underlay cognitive radio communications, the signal transmitted by the secondary user is disturbed by incoming signals from primary users. Thus, it is necessary to compensate for this secondary-link degradation at the receiver level. In this paper we use Dirichlet process mixtures (DPM) to relax *a priori* assumptions on the characteristics of the primary user-induced interference. DPM allow us to model the probability density function of the interference. The latter is estimated jointly with the symbols and the channel of the secondary link by using marginalized particle filtering. Our approach makes it possible to improve the symbol error rate compared with an algorithm that simply models the interference as a Gaussian noise.

Index Terms— Dirichlet Process, Cognitive radio, Particle filtering

1. INTRODUCTION

Cognitive radio (CR) is a possible concept for the fifth generation of mobile network [1]. CR can be viewed as a smart management of the radio frequency (RF) spectrum. The objective is to allow the mobile terminals, that use different communication standards, to coexist on the same frequency bands. With CR, a new class of users appears: the secondary users (SUs) as opposed to the primary users (PUs), who are the users of the current communication standards (GSM, UMTS, LTE, etc.). In contrast with PUs, the SU frequency bands are allocated dynamically. SUs are able to reuse the PU frequency bands. However, the interference level created by the SU signal on the PU received signal must be null or below a certain threshold. In order to use the PU frequency bands, different methods have been defined [2]: 1/Interweave, 2/overlay and 3/underlay spectrum access. The interweave technique consists in using, over a given time, one or several frequency bands left idle by PUs. In that case, the spectral efficiency of the SU communication depends on the *a priori* frequency band left unused by PUs. Underlay and overlay techniques allow a SU to reuse a frequency band used by a PU while guaranteeing a minimum level of interference on the PU signal [3, 4]. The main difference between underlay and overlay concerns the interference level created by the SU on the PU received signal. On the one hand, in underlay, the interference level of the transmitted SU signal must be maintained under an acceptable *a priori* known threshold. For example, spread spectrum access techniques are well appropriated in that case. On the other hand, in overlay there is no power constraint at the SU transmitter. Indeed, based on some information such as PU channel state information, a precoding and/or post-coding can be applied on the SU transmitted and/or received signals [6, 7, 8, 9]. Although the achievable secondary system capacity for these techniques is relatively low compared to the interweave case, they allow the CR to overcome the errors of PU detection and provide near-continuous availability of the secondary link. However, the interference level created by the PU on the SU received signal is not negligible. It may be reduced based on some *a priori* information on the PU, but generally they are difficult to be obtained in practice. Moreover, specific communication protocols, which reduce the time to exchange useful information as well as the spectral efficiency of the secondary link, are required.

In this paper we propose to reduce the PU interference on a SU received signal when overlay or underlay communications are considered. The proposed algorithm has the advantage of not requiring any *a priori* information about the PU at the SU receiver. Indeed, in the literature, many papers focus on the primary system performance, whereas few contributions have been proposed concerning the secondary link optimization. This is even truer, when the secondary system has no *a priori* information on the primary system. More specifically, our objective is to jointly estimate the propagation channel and the symbols coming from the SU by reducing the effect of interference due to the PU. Our contribution is to model the pdf[1] by a non-parametric model based on Dirichlet Process Mixtures (DPM). These highly flexible models, based on infinite mixtures of Gaussian distributions, can represent numerous probability density functions [5]. They have been recently exploited in various applications including the modeling of the multipath errors in GPS navigation [10, 11]. When using DPM in CR, it is unnecessary to know the number of interfering PU signals and their statistical characteristics. This is

[1] pdf standing for probability density function.

hence a great advantage.
The paper is organized as follows: Section 2 provides a DPM overview. Section 3 describes the Bayesian modeling of the problem. Section 4 describes the principle of the particle filter used for estimating the model variables. Finally, Section 5 provides the simulation results.

2. THE DIRICHLET PROCESS MIXTURE

Let $\{v_t\}_{t=1}^T$ be a T-length sequence of random variables with unknown distribution F.
When considering the problem as non-parametric, F can be written in the form

$$F(v_t) = \int_{\Theta} f(v_t|\boldsymbol{\theta}_t) dG(\boldsymbol{\theta}_t), \quad (1)$$

with $\boldsymbol{\theta}_t \in \Theta$ being a latent variable which contains the parameters of the user-defined mixed pdf f. Usually, f is a Gaussian pdf with mean μ_t and variance ϕ_t. $\boldsymbol{\theta}_t$ is then defined as $\boldsymbol{\theta}_t = [\mu_t, \phi_t]^T$ and one writes $f(v_t|\boldsymbol{\theta}_t) = \mathcal{N}(v_t; \mu_t, \phi_t)$. G is the mixing distribution, assumed to be random. In a Bayesian context, its prior distribution must be defined by the user. The Dirichlet process (DP) appears as a possible solution. It can be interpreted as a distribution on the space of the probability distributions. When G follows a DP distribution with base distribution G_0 and scale parameter α, one writes: $G \sim \mathcal{D}P(G_0, \alpha)$. Note that these distributions are discrete but infinite. In addition the "stick breaking" representation allows us to express G as an infinite mixture of Dirac measures as follows [12]:

$$G(\boldsymbol{\theta}_t) = \sum_{j=1}^{+\infty} \pi_j \delta_{\mathbf{U}_j}(\boldsymbol{\theta}_t) \text{ et } \pi_j = \beta_j \prod_{l=1}^{j-1} (1 - \beta_l), \quad (2)$$

with $\mathbf{U}_j \overset{iid}{\sim} G_0$[2] representing the j^{th} so-called cluster and $\delta_{\mathbf{U}_j}(\boldsymbol{\theta}_t)$ being the measure of Dirac with argument $\boldsymbol{\theta}_t$ located in $\boldsymbol{\theta}_t = \mathbf{U}_j$. π_j is the j^{th} weight, sequentially defined with $\beta_j \overset{iid}{\sim} \mathcal{B}(1, \alpha)$ with $\mathcal{B}$[3] being the Beta distribution. Combining (1) and (2) gives the unknown distribution F expressed as follows:

$$F(v_t) = \sum_{j=1}^{+\infty} \pi_j f(v_t|\mathbf{U}_j). \quad (3)$$

Thus F is a infinite mixture of pdfs $f(v_t|\mathbf{U}_j)$ the parameters of which are contained in vector $\mathbf{U}_j$.

Moreover DPs have interesting properties for inference. They can be easily sampled by using the so-called Polya urn model consisting in marginalizing G [13]. As a consequence, the latent variables $\{\boldsymbol{\theta}_t\}_{t=1}^T$ can be simulated sequentially according to the following laws:

$$p(\boldsymbol{\theta}_t|\boldsymbol{\theta}_{1:t-1}) = \frac{1}{\alpha + t - 1} \sum_{j=1}^{t-1} \delta_{\boldsymbol{\theta}_j}(\boldsymbol{\theta}_t) + \frac{\alpha}{\alpha + t - 1} G_0(\boldsymbol{\theta}_t), \quad (4)$$

with $\boldsymbol{\theta}_{1:t-1} = \{\boldsymbol{\theta}_1, ...\boldsymbol{\theta}_{t-1}\}$. It is clear that the high values of α ease the appearance of numerous different "clusters".
The next section exploits DPM for modeling PU signals in cognitive radios.

3. BAYESIAN MODELISATION OF THE PROBLEM

3.1. Received signal model

We consider a downlink communication[4] where the primary users and the SU transmit their information through an Orthogonal Frequency Division Multiplexing (OFDM) modulation. In this scenario, in addition to the classical additive white Gaussian noise (AWGN), the incoming signal at the SU receiver is disturbed by an interference term. The latter is the sum of the PU transmitted signals. Thus, the SU received signal can be written as follows:

$$y_n = h_{n,t} * x_n + \underbrace{i_n + n_n}_{v_n} \text{ with } i_n = \sum_{p=1}^{P} h_{n,t}^p * x_n^p, \quad (5)$$

where $h_{n,t}$ is the impulse response (IR) of the propagation channel between the secondary base station (SBS) and the SU. It should be noted that it may vary between two consecutive OFDM symbols. The subscripts t and n refer to the OFDM symbol index in an OFDM frame and the sample index, respectively.
Furthermore, we use the following definitions:

- x_n is the OFDM signal transmitted by the SBS,
- P is the number of the primary signals that interfer with the SU,
- $h_{n,t}^p$ represents the IR of the propagation channel between the p^{th} primary transmit antenna and the SU,
- x_n^p is the OFDM signal coming from the p^{th} antenna,
- n_n is the thermal noise (assumed AWGN) introduced by the secondary receiver.

At the ouput of the SU OFDM demodulator, the signal corresponding to the k^{th} sub-carrier, $k = 1, \ldots, K$, is expressed as follows[5]:

$$Y_t^k = H_t^k s_t^k + V_t^k \text{ with } V_t^k = I_t^k + N_t^k, \quad (6)$$

[2] iid standing for independant and identically distributed.
[3] The Beta pdf of parameter a, b and argument $\beta \in [0, 1]$ is defined as $\mathcal{B}(\beta; a, b) = \frac{\beta^{a-1}(1-\beta)^{b-1}}{B(a,b)}$ where $B(a, b)$ is the Beta function.
[4] From the base station to the mobile terminal.
[5] We assume the received secondary signal is time and frequency synchronized.

where Y_t^k, H_t^k, s_t^k, I_t^k and N_t^k are the k^{th} discrete Fourier transform (DFT) coefficients of y_n, $h_{n,t}$, x_n, i_n and n_n, respectively.

Given the above considerations we aim at addressing the joint estimation of s_t^k, H_t^k and the unknown pdf of the additive term V_t^k for a given sub-carrier. In the following, for the sake of simplicity we omit the superscript k.

In a Bayesian framework, this estimation is based on the posterior pdf of the unknown variables conditionnally upon the observations $Y_{1:t} = \{Y_1, \cdots, Y_t\}$. This distribution depends both on (6) and on the prior distributions defined in the next subsection.

3.2. Prior distributions

- *Information symbols s_t:* without loss of generality, the symbols are assumed to belong to a BPSK[6] modulation and to be equally distributed.

- *DFT of the propogation channel $H_t = H_t^{\mathcal{R}} + jH_t^{\mathcal{I}}$:* we consider a slowly time-varying channel. For that purpose, we model the evolution of both the real and imaginary parts by non-centered first-order autoregressive (AR) processes:

$$\begin{aligned} p(H_t^{\mathcal{R}}|H_{t-1}^{\mathcal{R}}) &= \mathcal{N}(H_t^{\mathcal{R}}; -a_1 H_{t-1}^{\mathcal{R}} + \gamma, \phi_H), \\ p(H_t^{\mathcal{I}}|H_{t-1}^{\mathcal{I}}) &= \mathcal{N}(H_t^{\mathcal{I}}; -a_1 H_{t-1}^{\mathcal{I}} + \gamma, \phi_H), \end{aligned} \quad (7)$$

where a_1 is the AR parameter, ϕ_H the variance of the driving process and γ a parameter related to the mean μ of the AR process by the formula $\mu = \gamma/(1 - a_1)$.

Furthermore, the equation (7) can be rewritten more compactly by introducing the following vectors and matrices: $\mathbf{x}_t = [H_t^{\mathcal{R}}, H_t^{\mathcal{I}}]^T$, $\mathbf{u} = [\gamma, \gamma]^T$, $\mathbf{F} = -a_1\mathbf{I}_2$ and $\mathbf{Q} = \phi_H\mathbf{I}_2$, where $\mathbf{I}_2$ represents the identity matrix of size 2×2:

$$p(\mathbf{x}_t|\mathbf{x}_{t-1}) = \mathcal{N}(\mathbf{x}_t; \mathbf{F}\mathbf{x}_{t-1} + \mathbf{u}, \mathbf{Q}). \quad (8)$$

- *Interferences and noise $V_t = V_t^{\mathcal{R}} + jV_t^{\mathcal{I}}$.* In a general manner, V_t can be written:

$$V_t = \sum_{p=1}^{P} \sum_{n=1}^{K} \Phi_n(x_{1:K}^p) h_{n,t}^p + N_t, \quad (9)$$

where the Φ_n are linear functions.

Classically, the propagation channels $h_{n,t}^p$ are assumed to be complex Gaussian. Thus, conditionally upon the primary signals received by the SU, the real and imaginary parts of V_t are also Gaussian. However, these information are usually not available. Therefore, we propose to represent the distributions of $V_t^{\mathcal{R}}$ and $V_t^{\mathcal{I}}$ by a DPM as defined by (1).

In the following, we define the vector $\underline{\boldsymbol{\theta}}_t = [(\boldsymbol{\theta}_t^{\mathcal{R}})^T, (\boldsymbol{\theta}_t^{\mathcal{I}})^T]^T$. It is composed of the latent variables corresponding to the real and imaginary parts assumed to be independent. Using the Polya urn representation, the evolution of the vector $\underline{\boldsymbol{\theta}}_t$ is described by the following distribution:

$$p(\underline{\boldsymbol{\theta}}_t|\underline{\boldsymbol{\theta}}_{1:t-1}) = p(\boldsymbol{\theta}_t^{\mathcal{R}}|\boldsymbol{\theta}_{1:t-1}^{\mathcal{R}})p(\boldsymbol{\theta}_t^{\mathcal{I}}|\boldsymbol{\theta}_{1:t-1}^{\mathcal{I}}) \quad (10)$$

Besides, we select for the DPM a Normal Inverse Gamma (NIG) base distribution. The pdf of the NIG law of arguments $\boldsymbol{\theta} = [\mu, \phi]^T$ and parameters μ_0, κ_0, α_0 et β_0 is defined by:

$$\mathcal{NIG}(\boldsymbol{\theta}; \mu_0, \kappa_0, \alpha_0, \beta_0) = \mathcal{N}(\mu; \mu_0, \frac{\phi}{\kappa_0})\mathcal{IG}(\phi; \alpha_0, \beta_0) \quad (11)$$

with $\mathcal{IG}(\phi; \alpha_0, \beta_0)$ the pdf of the Inverse Gamma law of argument ϕ and parameters α_0 and β_0. The latter makes it possible to define jointly a prior model for the mean and the variance. Taking advantage of the Polya urn representation in (4), the pdf estimation problem is equivalent to the computation of the joint posterior pdf of the latent variables $p(\underline{\boldsymbol{\theta}}_{1:t}|Y_{1:t})$. Based on the above Bayesian model, the extended state vector:

$$\mathbf{X}_t = [s_t, \mathbf{x}_t^T, (\underline{\boldsymbol{\theta}}_t)^T]^T \quad (12)$$

must be recursively estimated from the set of measurements $Y_{1:t}$. By using Bayes rule and conditional independences, the transition distribution of the extended state vector can be factorized as follows:

$$p(\mathbf{X}_t|\mathbf{X}_{1:t-1}) = \Pr[s_t; p_1]\ p(\mathbf{x}_t|\mathbf{x}_{t-1})\ p(\underline{\boldsymbol{\theta}}_t|\underline{\boldsymbol{\theta}}_{1:t-1}). \quad (13)$$

The model described in this section allows us to apply the marginalized particle filter [14] introduced in the next section.

4. MARGINALIZED PARTICLE FILTER

The marginalized particle filter is based on the factorization of the posterior pdf:

$$p(\mathbf{X}_{1:t}|Y_{1:t}) = p(\mathbf{x}_{1:t}|\underline{\boldsymbol{\theta}}_{1:t}, s_{1:t}, Y_{1:t})p(\underline{\boldsymbol{\theta}}_{1:t}, s_{1:t}|Y_{1:t}). \quad (14)$$

In our case, the evolution model of $\mathbf{x}_t$ defined in (8) is Gaussian and linear. Furthermore, conditionally upon to $\underline{\boldsymbol{\theta}}_{1:t}$ and $s_{1:t}$, the measurement model defined in (6) becomes linear Gaussian. Thus, given $\underline{\boldsymbol{\theta}}_{1:t}$ et $s_{1:t}$, the distribution $p(\mathbf{x}_{1:t}|\underline{\boldsymbol{\theta}}_{1:t}, s_{1:t}, Y_{1:t})$ in (14) is Gaussian and the Kalman filter can be used as an optimal estimator of the sequence $\mathbf{x}_{1:t}$ in the sense that it minimizes the minimum mean square error (MMSE). The latent variables $\underline{\boldsymbol{\theta}}_{1:t}$ and the symbols $s_{1:t}$ are estimated using a particle filter as follows:

$$\hat{P}_N(\underline{\boldsymbol{\theta}}_{1:t}, s_{1:t}|Y_{1:t}) = \sum_{i=1}^{N} w_t^{(i)} \delta_{\underline{\boldsymbol{\theta}}_{1:t}^{(i)}, s_{1:t}^{(i)}}(\underline{\boldsymbol{\theta}}_{1:t}, s_{1:t}), \quad (15)$$

where $\{\underline{\boldsymbol{\theta}}_{1:t}^{(i)}, s_{1:t}^{(i)}\}_{i=1}^N$ represent the particles, $\{w_t^{(i)}\}_{i=1}^N$ are the weights and N the number of used particles. Each particle is associated with a Kalman filter that computes recursively the estimate of the posterior mean of $\mathbf{x}_t$, denoted

[6] The proposed method can be applied whatever the constellation.

$\hat{\mathbf{x}}_{t|t}(\underline{\boldsymbol{\theta}}_{1:t}^{(i)}, s_{1:t}^{(i)})$, as well as the covariance matrix of the estimation error $\mathbf{P}_{t|t}(\underline{\boldsymbol{\theta}}_{1:t}^{(i)}, s_{1:t}^{(i)})$.
Finally, the marginal posterior pdf of $\mathbf{x}_t$ is estimated as a mixture of Gaussian distribution:

$$\hat{P}_N(\mathbf{x}_t|Y_{1:t}) = \sum_{i=1}^{N} w_t^{(i)} \mathcal{N}(\mathbf{x}_t; \hat{\mathbf{x}}_{t|t}(\underline{\boldsymbol{\theta}}_{1:t}^{(i)}, s_{1:t}^{(i)}), \mathbf{P}_{t|t}(\underline{\boldsymbol{\theta}}_{1:t}^{(i)}, s_{1:t}^{(i)}))$$

5. SIMULATION RESULTS

We test our algorithm on simulated data corresponding to an OFDM frame composed of $T = 500$ OFDM symbols. We consider $P = 6$ primary signals (for example transmitted from 3 primary base stations equipped with 2 transmit antennas).
The symbols carried by the primary signals belong to a BPSK constellation. The propagation channel of the SU is simulated according to the model (7) where the parameters are set to introduce a high temporal correlation: $\gamma = 0,02$, $a_1 = -0,98$ and $\phi_H = 0,004$. The primary channels are simulated according to a Rayleigh distribution without introducing temporal correlation. Figures 1.(a) and 1.(b) correspond to the histograms of $V_t^{\mathcal{R}}$ associated to two different OFDM frame. They are obtained from T OFDM symbols by taking into account the independency of the noises and the stationnarity of their distributions. The histograms show that the statistics of the interferences are highly multimodal. The pdf estimation of $V_t^{\mathcal{R}}$ obtained with our algorithm is superimposed on figure 1.(a) and 1.(b). Thus, although the estimated pdf are stationnary, the particle filter does not have convergence issues as it is usually the case with static parameters [14]. This is due to the Polya urn representation of the DP which makes it possible to integrate out analytically the mixing distribution. In this way only the dynamic latent variables have to be sampled.
Figures 1.(c) et 1.(d) show the estimation (in solid line) of the channel coefficients (in dashed line) $H_t^{\mathcal{R}}$ and $H_t^{\mathcal{I}}$. Finally, in table 1 are reported the estimation error percentages of the symbols s_t as a function of the signal to interference ratio (SIR). The first row corresponds to our algorithm whereas the second row to an algorithm that only uses an overbounding Gaussian distribution for the interference term. We can notice that the symbol error rate decreases more specifically for low RSI.

6. CONCLUSION

As a perspective, we can notice that the current model makes the assumption that the interferences between consecutive time instants are independent. Also widely spread in the literature, this assumption is not always satisfied in practice. To take into account this temporal correlation, we are currently studying new models based on time varying DPM.

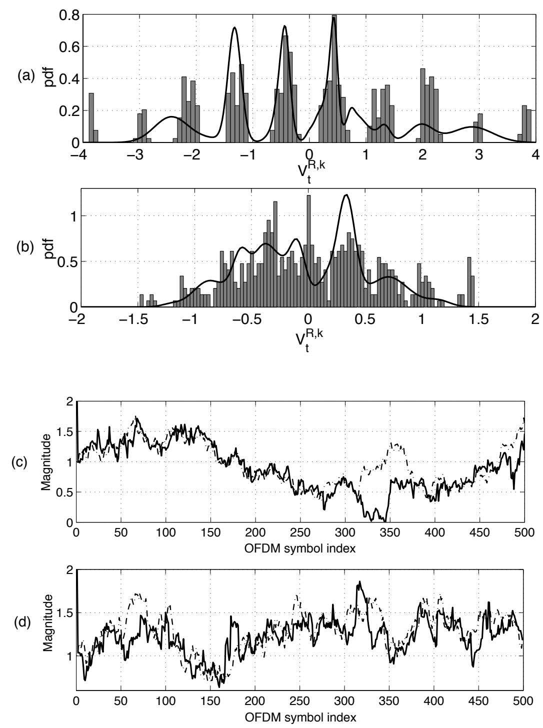

Fig. 1. (a)-(b) histograms and estimated pdf of $V_t^{\mathcal{R}}$, (c)-(d) real and imaginary part evolution of H_t.

Table 1. Symbol error rate (in %)

SIR (dB)	-5	-4	-3	-2	-1
without DPM	23,4%	26,8%	20,6%	14%	13,4%
with DPM	20,8%	20,6%	18,4%	13%	11,4%
SIR (dB)	0	1	2	3	4
without DPM	11,6%	7,8%	3,2%	3,6%	1,6%
with DPM	11%	3,9%	0,4%	0,6%	0,8%

REFERENCES

[1] J. Mitola III. *Cognitive Radio*: *An Integrated Agent Architecture for Software Defined Radio*. PhD Thesis, 2000.

[2] A. Goldsmith, S. Jafar, I. Maric and S. Srinivasa. *Breaking Spectrum Gridlock with Cognitive Radios*: *An Information Theoretic Perspective*. Proceedings of the IEEE, vol. 97, no. 5, pp. 894-914, 2009.

[3] Q. Spencer, A. Swindlehurst and M. Haardt. *Zero-Forcing Methods for Downlink Spatial Multiplexing in Multiuser MIMO Channels*. IEEE Transactions on Signal Processing, vol. 52, no. 2, pp. 461-471, 2004.

[4] L. Cardoso, M. Kobayashi, Y. Ryan and M. Debbah. *Vandermonde Frequency Division Multiplexing for*

Cognitive Radio. IEEE Signal Processing Advances in Wireless Communications, pp. 421-425, 2008.

[5] 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, 2008.

[6] J. Wu, T. Yang, H. Chou. *MIMO Active Interference Alignment for Underlay Cognitive Radio.* IEEE International Conference on Communications Workshops, pp.1-5, 23-27 May 2010.

[7] S. Perlaza, N. Fawaz, S. Lasaulce, M. Debbah. *From Spectrum Pooling to Space Pooling: Opportunistic Interference Alignment in MIMO Cognitive Networks.* IEEE Transactions on Signal Processing, vol.58, no.7, pp.3728-3741, July 2010.

[8] H. Zhu, B. Li, W. Chen. *A Practical and Efficient Algorithm for Distributed Interference Alignment Based on Cognitive Radio.* International Conference on Wireless Communications Networking and Mobile Computing, pp.1-4, 23-25 Sept. 2010.

[9] C. Shen, M. Fitz. *Opportunistic Spatial Orthogonalization and Its Application in Fading Cognitive Radio Networks.* IEEE Journal of Selected Topics in Signal Processing, vol.5, no.1, pp.182-189, Feb. 2011.

[10] A. Rabaoui, N. Viandier, E. Duflos, P. Vanheeghe and J. Marais. *Dirichlet Process Mixtures for Density Estimation in Dynamic Nonlinear Modeling*: *Application to GPS Positioning in Urban Canyons*. IEEE Transactions on Signal Processing, vol. 60, no. 4, pp. 1638-1655, 2012.

[11] V. Pereira, A. Giremus, A. Rabaoui and E. Grivel. *Online EM Estimation of the Dirichlet Process Mixtures Scale Parameter to Model the GPS Multipath Error*. Proceedings of the 38th International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2013.

[12] J. Sethuraman. *A Constructive Definition of Dirichlet Priors*. Statistica Sinica, vol. 4, no. 4, pp. 639-650, 1994.

[13] D. Blackwell and J. B. MacQueen. *Ferguson Distributions Via Polya Urn Schemes*. The annals of statistics, vol. 1, no. 2, pp. 353-355, 1973.

[14] A. Doucet, N. de Freitas and N. Gordon. *Sequential Monte Carlo Methods in Practice*. Springer, 2001.