GNSS pseudorange error density tracking using Dirichlet Process Mixture

Nicolas Viandier,
Juliette Marais,
Univ Lille Nord de France
Lille, France
INRETS, LEOST
Villeneuve d'Ascq, France
nicolas.viandier@inrets.fr
juliette.marais@inrets.fr

Asma Rabaoui
Univ Lille Nord de France
Lille, France
LAGIS UMR CNRS 8146
Ecole Centrale de Lille
Villeneuve d'Ascq, France
asma.rabaoui@ec-lille.fr

Emmanuel Duflos
Univ Lille Nord de France
Lille, France
LAGIS UMR CNRS 8146
Ecole Centrale de Lille
INRIA Futurs : SequeL Team
Villeneuve d'Ascq, France
emmanuel.duflos@ec-lille.fr

Abstract – *In satellite navigation system, classical localization algorithms assume that the observation noise is white-Gaussian. This assumption is not correct when the signal is reflected on the surrounding obstacles. That leads to a decrease of accuracy and of continuity of service.*
To enhance the localization performances, a better observation noise density can be use in an adapted filtering process. This article aims to show how the Dirichlet Process Mixture can be employed to track the observation density on-line. This sequential estimation solution is adapted when the noise is non-stationary. The approach will be tested under a simulation scenario with multiple propagation conditions. Then, this density modeling will be used in Rao-Blackwellised Particle Filter.

Keywords: GNSS, density estimation, Dirichlet Process Mixture, pseudodrange noise.

1 Introduction

Today, Global Navigation Satellite Systems (GNSS) have penetrated the transport field through applications such as monitoring of containers. These kind of applications do not request a high performance level of the positioning system. For safety applications (as complete guidance of the vehicle), performances require to be more stringent. The American system GPS (Global Positioning System) is the only fully operational solution for the moment. This monopole reduces the possibilities of measurement redundancy and diversity, thus limits the reachable performances.
Unfortunately, most of all these transport applications are mainly used in dense urban environments, highly constraining for signal propagation. This reduces the chances of good reception conditions. The consequences of environmental obstructions are unavailability of the service and reception of reflected signals that degrades in particular the accuracy of the positioning. Indeed, NLOS (Non Line Of Sight) signals - i.e. signals received after reflections on the surrounding obstacles - frequently occur in dense environments and degrade localization accuracy because of the delays observed on the propagation time measurement creating additional error on pseudorange estimation. The worst case of reception is the alternate path. In this case the LOS (Line Of Sight) signal from a satellite cannot reach the antenna and the receiver tracks only reflected signals.

In constricted environment, the pseudorange error becomes non-stationary because of signal reflexions. Consequently, classical localization methods (e.g. EKF: Extended Kalman filter), which assume that the observation noise is a white Gaussian noise, are inaccurate. Indeed, this assumption becomes unrealistic in such environments. Thus, to enhance the localization accuracy in case of alternate path reception, the filtering part of the receiver (after correlators) must be improved. Furthermore, in order to limit costs, we have chosen to work only with GNSS signals. In a goal of enhanced position accuracy, we propose a new statistical filtering method based on a better definition (and use) of the observation noise for each satellite signal. Moreover, in a very constricted environment (like urban environment or canyon) where reflected signals are frequent, the pseudorange noise density takes an unknown form. Consequently, we need to estimate this density, and we chose to a use mixture model. To compensate the lack of performances of GNSS positionning in complex propagation environments, we propose to use better density estimation methods in a particle filter. In our previous works, we have estimated the pseudorange error density by a finite Gaussian Mixture [1]. In these works, we have shown that using a more accurate density estimation enhances accuracy and continuity of service. But finite mixture have some drawbacks and cannot estimate with precision the real pseudorange error density.

Consequently, we decide to work with a more accurate and adapted estimation method: the Dirichlet Process Mixture.

This paper is organized as follow. First, the context of the GNSS application will be introduced and particularly the pseudorange measurement and the observation noise definition. Then, to enhance the density estimation, we use a Dirichlet Process Mixture. In the second part, this method will be described before developing how to adapt it for pseudorange error density estimation. In the last section, simulation results will illustrate the performance of the algorithm.

2 Context

In multi-sensors based systems, each sensor transmits a signal (an information) to the receiver (or antenna). These sensors can work according to three different operation modes which are:

- Failure mode, when the sensor cannot send its information.
- Degraded mode, when the information sent by the sensor is not accurate or is noisy.
- Normal mode, when the information sent by the sensor can be considered as accurate.

In this paper, the GNSS constellation is considered as a sensor network. Consequently, each GNSS satellite is considered as a sensor. The failure mode appears when the antenna cannot receive a signal from a satellite because of the local mask. In the normal mode, the satellite signal reaches the antenna in LOS. Finally, the degraded mode occurs when the signal reaches the antenna after one or more reflections. To compute a position in 3D by trilateration, the receiver needs to track signals from at least four different satellites. The distance between the satellite and the receiver is called pseudorange. This measure is deduced from the signal propagation time and is expressed as follow:

$$\rho_t^s = d_t^s - c.(\delta t_u + \delta t_s) + I_t^s + T_t^s + m_t^s + w_t \qquad (1)$$

Where ρ_t^s is the pseudorange between the satellite s and the receiver at the time t; d_t^s is the true satellite-receiver distance; c is the celerity; δt_u and δt_s are the receiver and the satellite clock offsets; I_t^s is the ionospheric error; T_t^s is the tropospheric error; m_t^s is the potential error due to the signal reflections and w_t is the receiver noise.

The atmospheric propagation errors and the satellites clock bias can be corrected by correction models ([2], [3] and [4]). Consequently, after corrections, (1) can be rewrite as (2):

$$\rho_t^s = d_t^s - c.\delta t_u + m_t^s + w_t = d_t^s - b_t + m_t^s + w_t \qquad (2)$$

Where b_t is equal to the receiver clock bias multiplied by the celerity.

From (2), the pseudorange error is only expressed in function of two different error sources as:

$$\epsilon_t^s = m_t^s + w_t \qquad (3)$$

According to the reception conditions, ϵ_t^s can switch between different observation models as expressed in 4.

$$\begin{cases} If LOS : m_t^s = 0, \epsilon_t^s \sim \mathcal{N}(0,\sigma) \\ Else : m_t^s \neq 0, \epsilon_t^s \nsim \mathcal{N}(0,\sigma) \end{cases} \qquad (4)$$

In the case when a signal is received in LOS, the pseudorange error distribution is considered white-Gaussian. In the second case, the pseudorange error distribution is unknown. In consticted environment, the density form can change ponctually or during a long period according to the obstacle nature. Moreover, several obstacles (other vehicules, people, vegetation, etc) have a random attitude and signal reflections (consequently error density location and form) cannot be predicted. Consequently, the pseudorange error become non-stationary because the mean and the standard deviation are not constant along time. To estimate the pseudorange error density with accuracy, the Dirichlet Process Mixture is used.

3 Dirichlet Process Mixture for the non-parametric density estimation

In our previous works [1], we have used a finite Gaussian Mixture to estimate the pseudorange error density. But this kind of method is not efficient in case of ponctual error. Indeed, with finite mixture, we must observe the error through a sliding window to estimate the density and the algorithm cannot detect immediatly a variation in the propagation condition. Moreover, the finite mixtures are not adapted when the noise is non-stationnary. That is why we decided to use the Dirichlet Process Mixture (DPM) for the pseudorange error density tracking. The DPM is an infinite mixture and is considered as a very flexible model [5].

In this section, the density problem will be described before introducing the Dirichlet Process and the Dirichlet Process Mixture.

3.1 Density Estimation Problem

Let us consider $y_1, y_2, \ldots, y_n$ a set of vectors statistically drawn from an unknown probability density F, such as $y_k \sim F$. In a Bayesian framework, the mixture model allowing to estimate F according to $y_{1:n}$ is given by (5):

$$F(y) = \int_\Theta f(y/\theta) d\mathbb{G}(\theta) \qquad (5)$$

Where $\theta \in \Theta$ is a latent variable (or cluster), $f(./\theta)$ is the mixed density and $\mathbb{G}$ is the mixing distribution.

$\mathbb{G}$ is a Random Probability Measure (RPM). In this paper, we choose $\mathbb{G}$ as being a Dirichlet Process (DP) because it deals with some interesting properties.

3.2 Dirichlet process

The DP was introduced by Ferguson in [6] as a probability measurement on the space of the probability measures. The DP is defined from two parameters which are a concentration parameter α and a base distribution $\mathbb{G}_0$. Let us consider $\mathbb{G}$ drawn from a DP with parameters α and $\mathbb{G}_0$, i.e. $\mathbb{G} \sim DP(\mathbb{G}_0, \alpha)$, the random variables $\theta_{1:n}$ are drawn from $\mathbb{G}$ as follows:

$$\theta_k/\mathbb{G} \sim \mathbb{G} \tag{6}$$

An important property of the DP is that $\mathbb{G}$ is discrete with probability one. Consequently, the stick breaking representation can be introduced as an infinite sum of Dirac measures [7]:

$$\mathbb{G} = \sum_{k=1}^{\infty} \pi_k.\delta_{\theta_k} \tag{7}$$

With $\theta \sim \mathbb{G}_0$, $\pi_k = \beta_k \prod_{j-1}^{k-1}(1-\beta_j)$ and $\beta_k \sim \mathcal{B}(1,\alpha)$ where $\mathcal{B}$ is the standard Beta distribution and δ_{θ_k} is a Dirac located in θ_k.
Moreover, marginalizing out $\mathbb{G}$, the distribution of $\{\theta_1, \dots, \theta_n\}$ follows a Polya urn scheme [8] as follows:

$$\theta_{n+1}/\theta_{1:n} \sim \frac{1}{\alpha+n}\sum_{k=1}^{n}\delta_{\theta_k} + \frac{\alpha}{\alpha+n}\mathbb{G}_0 \tag{8}$$

Using the Polya urn representation, the new sample θ_{n+1} can take the value of a previous sample with the probability $\frac{n}{\alpha+n}$ or can take a random value from $\mathbb{G}_0$ with the probability $\frac{\alpha}{\alpha+n}$.
From (7), we can express (5) to estimate the unknown distribution F as follows:

$$F(y) = \sum_{k=1}^{\infty} \pi_k.f(y/\theta_k) \tag{9}$$

3.3 Dirichlet Process Mixture

The Dirichlet Process Mixture (DPM) can be explained by the hierarchical model represented in fig.1. DPM are attractive when the parametric models impose assumptions on distributions that are too restrictive. The hierarchical model is expressed by (10):

$$\begin{array}{l} y_k/\theta_k \sim f(y/\theta_k) \\ \theta_k \sim \mathbb{G} \\ \mathbb{G} \sim DP(\mathbb{G}_0, \alpha) \end{array} \tag{10}$$

The RPM $\mathbb{G}$ is drawn from a Dirichlet Process, $DP(\mathbb{G}_0, \alpha)$. The clusters θ_k are distributed from $\mathbb{G}$. $f(./\theta)$ is a Gaussian mixed density.

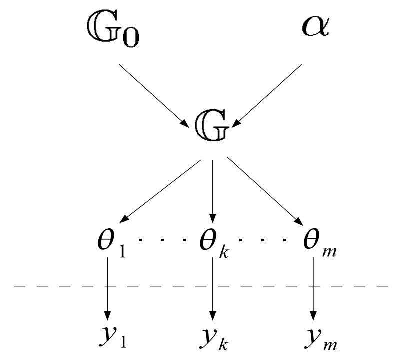

Figure 1: DPM hierarchical model representation

An important instance of the general model is the Normal-inverse Wishart DPM model, given by:

$$\begin{array}{c} p_{\mu,\Sigma}(y_i) = \mathcal{N}(y_i, \mu, \Sigma) \\ \mathbb{G}_0(\mu, \Sigma) = \mathcal{N}(\mu, m, B).\mathcal{W}^{-1}(\Sigma, r, R) \end{array} \tag{11}$$

Where $\mathcal{N}$ denotes the Normal distribution and $\mathcal{W}^{-1}$ denotes the inverse Wishart distribution.
A key feature of the DP in this DPM model is that the θ_i are marginally sampled from $\mathbb{G}_0$, and there is a positive probability that some of the θ_i are identical. This is due to the discreteness of the random measure $\mathbb{G}$. This discreteness of $\mathbb{G}$ is the main impediment to an efficient estimation. Posterior integration, and thereby most inferences, are made difficult because of a combinatorial explosion in the number of terms in the posterior distribution. That stems from the need to account for all possible configurations on how the θ_i's are identical and distinct. However, implementation of Gibbs sampling is almost straightforward when one marginalizes over $\mathbb{G}$ and works directly with the θ_i. Except for a difficulty which arises when resampling θ_i conditional on all other parameters. The new value of θ_i can either be one of the θ_h's, $h \neq i$, or θ_i could be a new draw from $\mathbb{G}_0$.

4 Error Density tracking using DPM for GNSS applications

To apply DPM to GNSS localization, the algorithm must be adapted. In our application of on-line pseudorange error density estimation, the latent variables θ_k are the mean and the variance of each Gaussian distribution included in the infinite mixture, i.e. $\theta_k = \{\mu_k, \sigma_k^2\}$. At each step N particles θ_k are computed. Then, the estimated pseudorange error density is obtained by computing the ponderate sum of Gaussian densities with parameters θ_k.
The problem with GNSS applications is that signal propagation can be considered in three different modes (LOS, NLOS and blocked). The NLOS case is the worst reception condition. It impacts on pseudorange error

properties and the noise amplitude and variation become very difficult to anticipate. In LOS reception, the pseudorange error noise follows a zero-mean Gaussian distribution with a standard deviation σ inferior or equal to 1 (actually it depends of the receiver used). In NLOS reception, the pseudorange error can be higher than 100 m. The same base distribution $\mathbb{G}_0$ cannot be applied in all situations. In the litterature, algorithms adapting the density estimation by mean jump [9] or variance jump can be found [10]. Here, we propose to adapt both, each time. To optimize the DPM for the GNSS applications, we propose to use an additional parameter r_t^s for each satellite s at each time t, as shown in fig.2.

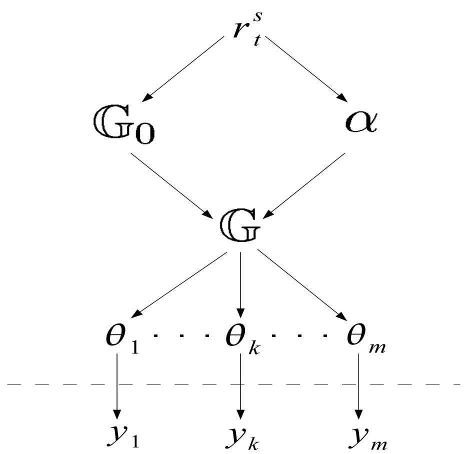

Figure 2: DPM hierarchical model adaptation for GNSS application

r_t^s is an estimation of the propagation state (LOS, NLOS or blocked). It updates the parameters of $\mathbb{G}_0$ and the concentration parameters α. In practice, when a LOS reception is detected the hyperparameters distributed by $\mathbb{G}_0$ are drawn as follow:

$$\begin{array}{c} \sigma_k \sim \mathcal{W}^{-1}(\Sigma_1, dof_1) \\ \mu_k \sim \mathcal{N}(\mu_0, \frac{\sigma_k}{k_1}) \end{array} \tag{12}$$

Where μ_0 is the mean of the Normal distribution, k is a scale parameter, Σ and dof are respectively the covariance and the degree of freedom for the inverse Wishart distribution.
However, when a NLOS reception is detected the distribution of the hyperparameters is adapted and:

$$\begin{array}{c} \sigma_k \sim \mathcal{W}^{-1}(\Sigma_2, dof_2) \\ \mu_k \sim \mathcal{N}(\mu_0 + res_t^s, \frac{\sigma_k}{k_2}) \end{array} \tag{13}$$

With $\Sigma_2 > \Sigma_1$, $dof_2 > dof_1$, res_t^s the estimation of the pseudorange error and $k_2 < k_1$.
We also adapt the concentration parameter α in such way that, in NLOS reception, θ_k has a strong probability to be drawn from $\mathbb{G}_0$ rather than to be equal to a previous cluster.

5 Density Tracking Performances

In this section, DPM is tested on simulation data performed with Ergospace®. First, the simulation setup is described. Then the DPM performances are compared with a finite Gaussian Mixture and the impact of density estimation on the localization performances is proved.

5.1 Simulations

The Ergospace® software simulates electromagnetic signal propagation in 3D realistic constricted environments. This software is developed as a support for the development of GNSS applications. The deterministic method of Ray Tracing determines possible paths of received rays. Different propagation models (corresponding to different types of environment) characterize interactions between the signal and the studied environment. Receiver behavior (mobile or other) and its performance are also taken into account in calculations. The software provides output raw data files in a Matlab format that we process in order to compute the pseudorange error and the position with our algorithms. Thus, in the simulation, the reception state r_t^s is well-known as well as the pseudorange error. Moreover, this software allows anticipating the Galileo constellation.
Simulations have been performed with a 3D scene of the town of Rouen, France (fig.3). The simulated route corresponds to a bus line. The mobile runs at a constant speed of 50 km/h, during 10 minutes and covers a distance of 8236 meters. GNSS signals can be received with a maximum of 3 reflections. The antenna passes trough different propagation environments (clear, urban, suburban,...) thus signals are not received in the same conditions according to each satellite position.

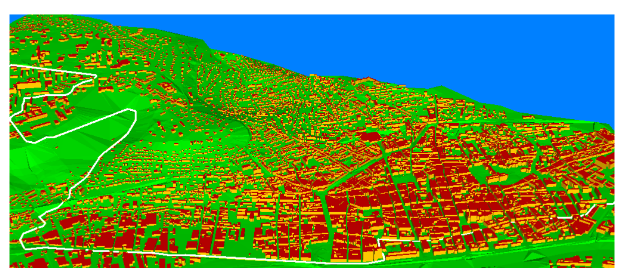

Figure 3: Bus line simulation on the 3D scene of town of Rouen

In this paper, we will consider the pseudorange measurements from two different satellites to test the algorithm performances. These satellites are the GPS satellites 18 and 13. Fig.4 and fig.5 show the timing diagram of the reception state for the two satellites.

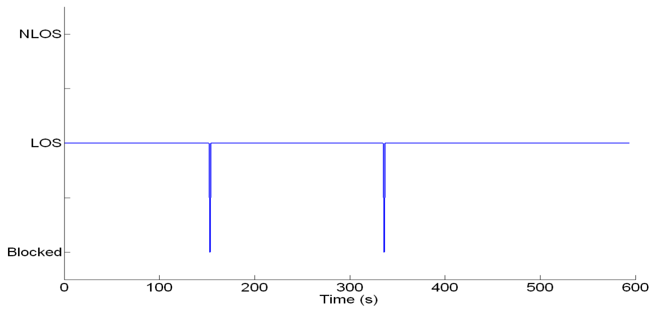

Figure 4: Timing diagram of the reception state for satellite 18 during the simulation

Signals coming from satellite 18 are always received in LOS. Consequently, the pseudorange error from this satellite is a white-Gaussian noise ($\mathcal{N}(0,1)$). For this satellite, the maximum of the pseudorange error amplitude (absolute value) is equal to 3.16m and the minimum is equal to 0.0014m.

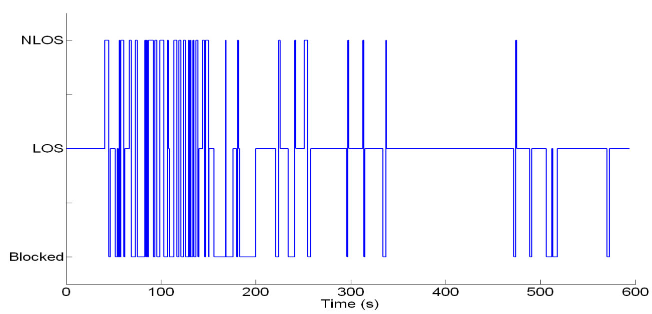

Figure 5: Timing diagram of the reception state for satellite 13 during the simulation

Signals coming from satellite 13 are received according to different reception states. Consequently, the pseudorange error from this satellite is a non-stationary noise. For this satellite, the maximum of the pseudorange error amplitude is equal to 102m and the minimum is equal to 0.0063m.

5.2 Density Tracking Results

In this section, the pseudorange error density is tested with the DPM algorithm. The DPM is compared to a finite Gaussian Mixture (GM). To test the DPM performances for the density tracking, we check at each simulation step if the pseudorange error is included in the estimated density.

To test this assumption, the mean and the variance of a mixture distribution must be known. The mean of a mixture distribution is given by:

$$\mu = \sum_{k=1}^{N} w_k.\mu_k \tag{14}$$

Where μ is the mean of the mixture, N is the particle number, w_k are the weights of the mixture and μ_k are the means of each Gaussian of the mixture.

And the variance of a mixture distribution is given by:

$$\sigma^2 = \sum_{k=1}^{N} w_k.(\sigma_k^2 + \mu_k^2) - \mu^2 \tag{15}$$

Where σ^2 is the variance of the mixture and σ_k^2 is the variance of each Gaussian of the mixture.

Fig.6 and fig.7 represent the true pseudorange error and the estimated density parameters (mean μ and standard deviation σ) for the satellites 18 and 13.

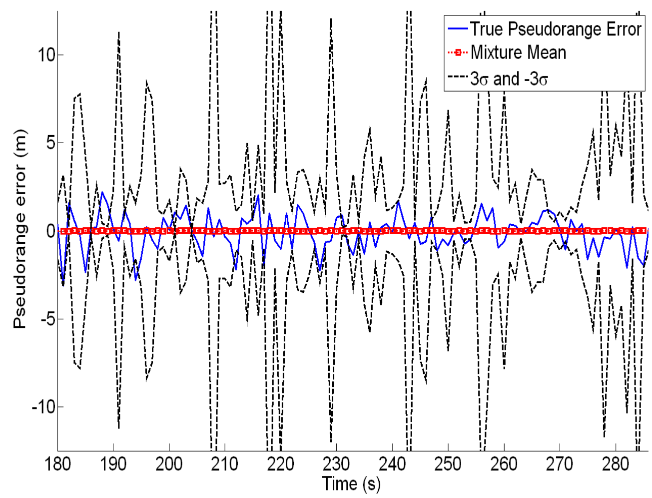

Figure 6: Results at 3-σ of pseudorange error density tracking for satellite 18

In presence of a white-Gaussian noise, DPM efficiently estimates the pseudorange error density for satellite 18. The majority of the true errors are included in the 3-σ area.

In presence of a non-stationary noise, DPM immediatly switches to the better density estimation as represented in fig.7. The part of the simulation composed of white-Gaussian noises is correctly estimated as well as the part composed of reflected signals. In the case of NLOS reception the variance value is higher.

Table 1 gives the percentages of acceptance of the pseudorange error tracking for the DPM and GM methods according to σ levels. These results confirm that the DPM model is efficient to caracterize pseudorange error density at 88% for a non-stationary noise and at 84% for a white-Gaussian noise.

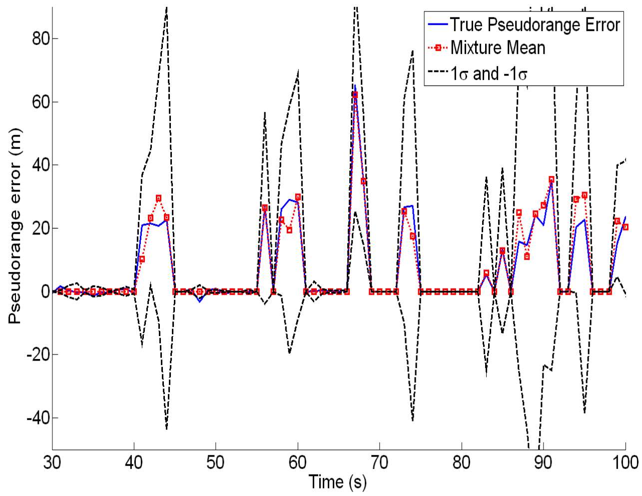

Figure 7: Results at 1-σ of pseudorange error density tracking for satellite 13

Table 1: Efficiency of the pseudorange error tracking for the DPM and GM methods according to σ level

	1-σ	2-σ	3-σ
DPM tracking with a non-stationnary noise (satellite 13)	71.04%	82.49%	88.38%
GM tracking with a non-stationnary noise (satellite 13)	76.26%	88.09%	90.90%
DPM tracking with a white-Gaussian noise (satellite 18)	56.90%	76.09%	83.84%
GM tracking with a white-Gaussian noise (satellite 18)	65.32%	87.54%	92.76%

Even if the higher percentages of acceptance is given by GM, fig.8 shows that this method is less accurate compared with DPM. Fig.8 draws the GM density tracking results for satellite 13. On the top, the results are given for a NLOS reception part. We can observe that with GM estimation the sudden and ponctual error variations are not correctly estimated. Moreover, the estimation is less accurate as we can see on the bottom part on the figure which squares with a LOS part. Indeed, the standard deviation is widely overestimated. We can explain this bad modelization for GM by the fact that the finite Gaussian Mixture is estimated through a time window. Consequently, the past values do not allow to update immediatly the density estimation. For the DPM, this problem does not exist.

To conclude this section, the localization performances are given for the different density tracking methods. The White-Gaussian noise (W.G.) is considered by applying an EKF. The GM and the DPM are used in

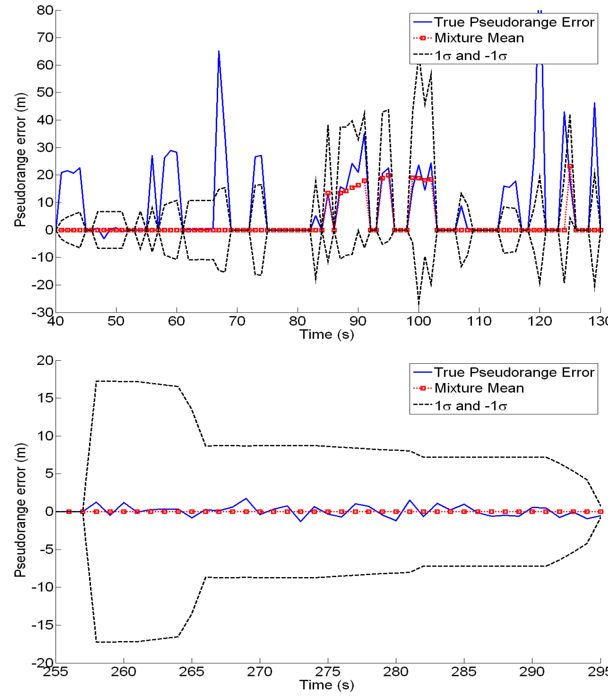

Figure 8: MG results at 1-σ of pseudorange error density tracking for the satellite 13. On the top, the results are given for a NLOS reception parts and on the bottom, for a LOS reception part.

a Rao-Blackwellised Particle Filter (RBPF) [11]. Table 2 shows the localization performances for the three noise models in terms of mean error, availability of the navigation solution with a threshold of 5 meters, MUT (Mean Up Time), MDT (Mean Down Time) and MTBF (Mean Time Between Failure).

From the results shown in Table 2, we can confirm that classical methods which assume that the noise is white and Gaussian are less accurate. Moreover, this kind of methods has a period of failure (when the position error is superior than 5 m) longer and a period of good performances shorter than the methods which do not make the same assumption about noise.

Table 2: Localization performances for different localization methods and different density models

	Mean error	Avail. at 5m	MUT	MDT	MTBF
EKF (W.G.)	6.02 m	62%	10.6 s	6.4 s	17.0 s
RBPF (GM)	2.93 m	87%	14.6 s	2.1 s	16.7 s
RBPF (DPM)	2.76 m	90%	14 s	1.6 s	15.6 s

The DPM enhances the accuracy of the solution and reduces the period of failure. The DPM performances can be ameliorated by using more accurate law parameters for $\mathbb{G}_0$.

6 Conclusions

In order to improve the pseudorange error density tracking on-line, an adapted Dirichlet Process Mixture (DPM) is proposed.

The main contribution of this paper is to show how to adapt DPM for GNSS applications and to show the added-value of this method on the pseudorange error tracking. First, we have explained the context of the work. Then, we have introduced the DPM principles and explain how to adapt the algorithm for GNSS. Finally, we have shown that DPM is efficient in comparison with a finite Gaussian Mixture estimation.

Future works will focus on applying DPM on real GNSS data in very constricted environment and in presence of severe NLOS receptions. Another perspective of work is to find the better $\mathbb{G}_0$ law parameters. Mainly, the scale parameter α should be estimated by applying a flexible method. A last perspective will be to use the DPM model as a posteriori model for mapping applications.

References

[1] N. Viandier, D. F. Nahimana, J. Marais and E. Duflos, “GNSS performance enhancement in urban environment based on pseudo-range error model,” *PLANS 2008*, IEEE/Institute Of Navigation, Monterey, California.

[2] H.S. Hopfield, “Two quadratic tropospheric refractivity profile for correction satellite data,” *Journal of Geophysical research*, April 1969.

[3] J. A. Klobuchar, “Ionospheric Effects on GPS,” *Global Positioning System: Theory and Applications, Volume 1*, American Institute of Aeronautics and Astronautics, Inc., Norwood, 1996.

[4] B. W. Parkinson and J. J. Spilker, *Global Positioning System : Theory and Applications, Volume 1*, American Institute of Aeronautics and Astronautics, Inc., Washington, 1996.

[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] T. Ferguson, “A Bayesian analysis of some nonparametric problems,” *The Annals of Statistics*, Vol 1, No. 2, pp. 209–230, 1973.

[7] J. Sethuraman, “A constructive definition of Dirichlet priors,” *Statistica Sinica*, Vol 4, pp. 639-650, 1994.

[8] D. Blackwell and J. MacQueen, “Ferguson distributions via Polya urn schemes,” *The Annals of Statistics*, Vol 1, No. 2, pp. 353-355, 1973.

[9] A. Giremus, J.-Y. Tourneret, and V. Calmettes, “A particle filter approach for joint detection/estimation of multipath effects on GPS measurements,” *IEEE Transactions on Signal Processing*, Vol 55, No. 4, pp. 1275–1285, 2007.

[10] M. Spangenberg, “Safe navigation for vehicles,” *Ph.D. dissertation*, Toulouse University, Toulouse, France, 2009.

[11] A. Rabaoui, N. Viandier, J. Marais and E. Duflos, “Using Dirichlet Process Mixtures for the modelling of GNSS pseudorange errors in urban canyon,” *ION GNSS 2009*, Institute Of Navigation, Savannah, Georgia.