COMPENSATING POWER AMPLIFIER DISTORTION IN COGNITIVE RADIO SYSTEMS WITH ADAPTIVE INTERACTING MULTIPLE MODEL

Mouna Ben Mabrouk, Eric Grivel, Clément Magnant, Guillaume Ferré and Nathalie Deltimple

University of Bordeaux, Bordeaux INP, IMS Laboratory UMR CNRS 5218,
351 Cours de la Libération, 33405 Talence, France

ABSTRACT

This work aims at improving the power amplifier (PA) efficiency in uplink OFDM-based cognitive radio (CR) communications. Unlike the traditional approaches, we suggest transmitting a non-linearly amplified signal without any filtering and addressing the OFDM sample estimation from the distorted signal at the receiver. The proposed post-distortion and detection technique is based on a Volterra model for the PA and the channel. As the transmission can switch from one sub-band to another, the CR-PA behavior varies over time and the Volterra kernels can be constant or suddenly change. Therefore, an interactive multiple model (IMM) combining extended Kalman filters is considered. The transition probability matrix, which plays a key role in the IMM, is also sequentially estimated. The resulting uplink system has various advantages: it learns from the observations and a part of the computational load is exported to the receiver, which is not battery driven unlike the mobile terminal.

***Index Terms*—** Power amplifier, digital post/pre-distortion, cognitive radio, Volterra modeling, interacting multiple model, transition probability matrix estimation.

1. INTRODUCTION

The cognitive radio (CR) offers a solution to the spectral crowding problem by introducing the opportunistic usage of frequency bands that are not heavily occupied by licensed users. CRs can be used as secondary systems of current allocation of licensed users. In this case, cognitive users detect the unused spectrum to exploit it [1]. Hence, the CR must have the ability to sense and to be aware of its operational environment. It must dynamically adjust its radio operating parameters accordingly.

Multicarrier transmission such as orthogonal frequency division multiplexing (OFDM), filter bank multicarrier (FBMC) and generalized frequency division multiplexing (GFDM) have been recently proposed as possible waveforms for CR thanks to their potential to fulfill the aforementioned requirements. However, these multicarrier modulations exhibit high peak-to-average power ratio (PAPR).This involves that power amplifiers (PAs) operate in their linear regions most of the time, where efficiency is low. To save power on terminal equipments, PAs must operate in their non-linear regions, causing signal compression. This leads to signal waveform distortion and adjacent-channel interference [2]. Power back-off and PAPR reduction techniques reduce the non-linear distortion effects but result in low power efficiency [3]. Otherwise, a linearization technique can be used.

In this paper, a PA linearization technique for uplink communication is developed. Unlike the existing approaches such as [3], we propose 1/ to transmit an amplified signal without any filtering and which has been non-linearly distorted. This can be done because the CR can ensure that the harmonics induced by the non-linearity do not disturb the licensed users. 2/ to export the computational load to the receiver, where a post-distortion technique is developed. The proposed approach has the following features:

1. It is based on a Volterra model to take into account both the PA behavior and the channel. As the equivalent channel can be constant during a communication in one sub-band but can suddenly change when one switches to another frequency band, the Volterra kernels can remain unchanged or be time-varying. This hence leads to a Markov-jump system.
2. At the receiver, a joint estimation of both the input signal and the Volterra kernels is proposed. Since this is a non-linear issue, an extended Kalman filter (EKF) can be considered. As at least two models must be used to represent the time evolutions of the Volterra kernels, an IMM structure combining different EKFs is proposed. However, the transition probability matrix (TPM), illustrating the switching between the two types of the Volterra-kernel evolutions, must be defined to use this IMM.
   Few papers deal with the selection or the estimation of the TPM. In addition, most of the time, the application is in the field of target or aircraft tracking [4]. A first solution is to set the TPM at a predefined value. As pointed out by Bloomer in [5], this *a priori* selection can be done provided that the mean sojourn time in each model can be determined. This is what we did in [6]. This choice is however too specific; it requires additional devices to estimate the mean sojourn times. In addition, the predefined TPM is not well suited if the system is non-stationary. As an

978-0-9928626-3-3/15/$31.00 ©2015 IEEE

alternative, Eun *et al.* suggest using fuzzy logic approach to adjust the state transition probabilities [4]. Other authors have proposed to estimate the TPM [7, 8]. In [7], Jilkov *et al.* first give the state of the art on this topic which includes Tugnait's work [8]. Then, they propose four algorithms: the so-called moment-based algorithm, the second-order algorithm, the quasi-Bayesian TPM estimate and the numerical integration. After doing a comparative study between these four methods, we decided to use the numerical integration algorithm as the others may sometimes diverge. Although the TPM estimation increases the computational cost compared to a predefined choice of the TPM, the whole estimation process is more flexible. As the TPM is adaptively estimated, this leads to an adaptive IMM.

The proposed approach has hence the advantage of increasing the PA efficiency at the transmitter. In uplink, it makes it possible to reduce the consumed power at the mobile terminal, which is usually battery driven unlike the receiver.

The paper is organized as follows: the system model is described in section 2. The proposed post-distortion and detection technique for non-time varying PA behavior as well as the time-varying case are detailed in section 3. In the simulation-result part, the performance and the limits of the proposed algorithm are then discussed. More particularly, we focus our attention on the TPM estimation.

In the following, $(.)^*$ is the complex conjugate, $\mathbf{I}_N$ denotes the identity matrix whose size is $N \times N$, $\mathbf{0}_{N\times M}$ is the zero matrix whose size is $N \times M$, the bold variables are vectors or matrices and the non-bold ones are scalars. $Re(\cdot)$ and $Im(\cdot)$ are the real and imaginary parts respectively.

2. SYSTEM MODEL

The system model considered in this part is depicted in Fig. 1. y_k is the output sample of the equivalent channel composed of the CR-PA and the multipath channel, n_k is an additive zero-mean white Gaussian noise (AWGN) with variance σ_n^2, z_k is the received sample, $\hat{u}_k$ is the estimate of the CR-PA OFDM input sample u_k. The latter can be expressed as follows:

$$u_k = \frac{1}{\sqrt{L}} \sum_{l=0}^{L-1} S_l \exp(j2\pi kl/L) \tag{1}$$

where L is the IFFT size and S_l for $l \in \{0, \cdots, L-1\}$ are the symbols assumed to be uniformly distributed. The distribution of u_k can be approximated by a Gaussian distribution with zero-mean and a variance proportional to L.

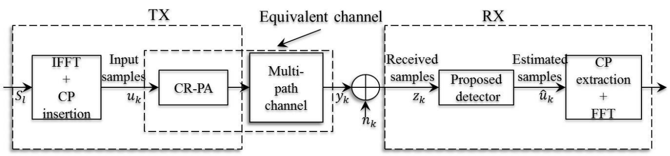

Fig. 1. System model.

Here, we suggest modeling the equivalent channel by a P^{th}-order Volterra model with memory depth M. y_k is hence written as follows [10]:

$$y_k = \sum_{n=1}^{P} \frac{(2n+1)!}{(n+1)!n!2^{2n}} \sum_{\tau_1=0}^{M-1} \cdots \sum_{\tau_n=0}^{M-1} h_n(\tau_1, \ldots, \tau_{2n+1}, k) \times \prod_{s=0}^{n} u^*_{k-\tau_s} \prod_{r=n+1}^{2n+1} u_{k-\tau_r} \tag{2}$$

where $h_n(\tau_1, \tau_2, \ldots, \tau_{2n+1}, k)$ for $n \in \{0, \cdots, P\}$ denote the Volterra kernels which depend on time k.

3. KALMAN ALGORITHMS FOR DIGITAL POST-DISTORTION AND DETECTION

In this section, let us first present the KF-based detector when the properties of the PA do not vary over time. Then, the time-varying PA behavior case is studied. This happens in a CR system when one switches from one sub-band to another.

3.1. Non time-varying PA behavior case

In order to estimate the Volterra kernels and the input signal, let us introduce the state-space representation (SSR) of the system depicted in Fig. 1 and defined by (2). Firstly, we propose to store the N Volterra kernels[1] in a column vector $\mathbf{C}_k$. When the PA behavior does not change over time during the communication, the Volterra kernels remain constant:

$$\mathbf{C}_k = \mathbf{C}_{k-1} = \mathbf{C}_{k-1} + \mathbf{w}_k(m_{st}) \tag{3}$$

where $\mathbf{w}_k(m_{st})$ is an AWGN with zero-mean and covariance $\overline{\mathbf{Q}}(m_{st}) = \mathbf{0}_{N\times N}$, the index st stands for static behavior.

To describe the way the OFDM input samples evolve over time, the real and the imaginary parts of the M last OFDM samples are stored in a column vector $\mathbf{D}_k$ of size $2M$.

$$\mathbf{D}_k = \left[\; Re(u_k), Im(u_k) \cdots Re(u_{k-M-1}), Im(u_{k-M-1}) \;\right]^T = \mathbf{F}\mathbf{D}_{k-1} + \mathbf{G}\left[\; Re(u_k) \quad Im(u_k) \;\right]^T \tag{4}$$

with $\mathbf{F} = \begin{pmatrix} \mathbf{0}_{2\times 2M} & \\ \mathbf{I}_{2M-2} & \mathbf{0}_{2M-2\times 2} \end{pmatrix}$ and $\mathbf{G} = \begin{pmatrix} \mathbf{I}_2 \\ \mathbf{0}_{2M-2\times 2} \end{pmatrix}$.

Given (3)-(4), the time evolution of the state vector is:

$$\mathbf{x}_k = \left[\; \mathbf{D}_k^T \quad \mathbf{C}_k^T \;\right]^T = \mathbf{F}'\mathbf{x}_{k-1} + \mathbf{G}'\mathbf{v}_k(m_{st}) \tag{5}$$

where $\mathbf{F}' = \begin{pmatrix} \mathbf{F} & \mathbf{0}_{2M\times N} \\ \mathbf{0}_{N\times 2M} & \mathbf{I}_N \end{pmatrix}$,

$\mathbf{G}' = \begin{pmatrix} \mathbf{G} & \mathbf{0}_{2M\times N} \\ \mathbf{0}_{N\times 2} & \mathbf{I}_N \end{pmatrix}$ and $\mathbf{v}_k(m_{st}) = \begin{pmatrix} Re(u_k) \\ Im(u_k) \\ \mathbf{w}_k(m_{st}) \end{pmatrix}$.

In addition, given (2), the state vector and the column vector $\mathbf{z}_k$ containing the real and the imaginary parts of z_k satisfy:

$$\mathbf{z}_k = \left[\; Re(y_k) \quad Im(y_k) \;\right]^T + \mathbf{n}_k = h(\mathbf{x}_k) + \mathbf{n}_k \tag{6}$$

where $\mathbf{n}_k$ is a zero-mean white Gaussian column vector with covariance matrix $\sigma_n^2\mathbf{I}_2$. Given the Gaussian assumptions on

[1] N depends on both M and P.

u_k and $\mathbf{n}_k$ and due to the non-linearity relation (6), the state vector can be sequentially estimated from the noisy observations from time 1 to k denoted as $z_{1:k}$. Its *a posteriori* estimate at time k satisfies:

$$\hat{\mathbf{x}}^{st}_{k|k} = E[\mathbf{x}_k|z_{1:k}] = \int \mathbf{x}_k p(\mathbf{x}_k|z_{1:k})d\mathbf{x}_k \qquad (7)$$

One of the following algorithms can be used: 1/ an EKF or a second-order EKF. 2/ sigma-point KFs that include the unscented KF, the central difference KF, the cubature and quadrature KFs [12]. In this application, as a 1^{st}-order Taylor expansion is a sufficient approximation of the non-linear measurement function $h(.)$, the EKF is a good compromise between computational cost and estimation accuracy.

In the following, we show how to address the joint estimations of the Volterra kernels and the PA input samples when the PA behavior evolves over time.

3.2. Time-varying behavior case: a CR-PA behavior

3.2.1. Combining two EKFs in an IMM structure

The Volterra kernels could not be tracked by the EKF based on (3). If they were modeled by random walks, i.e. if the model noise $\mathbf{w}_k$ in (3) was no longer zero but a zero-mean AWGN with a non-zero variance, the EKF could track these variations but it cannot provide accurate estimates of static parameters.

For these reasons, two EKFs are used. They are based on different *a priori* modeling of the Volterra-kernel time evolution. The SSR of the first one corresponds to the one presented above whereas the SSR of the second one is similar to (5), but the covariance matrix of the model noise $\mathbf{w}_k$ becomes equal to $\overline{\mathbf{Q}}(m_{ld}) = \sigma^2_{ld}\mathbf{I}_N$ where the index ld stands for large dynamics. Then, the estimators based on both assumptions, labeled m_{st}-EKF and m_{ld}-EKF, are combined in an IMM structure. In this IMM, each EKF provides an estimate of the state vector denoted as $\{\hat{\mathbf{x}}^j_{k|k}\}_{j=st,ld}$ and its associated error covariance matrix $\{\mathbf{P}^j_{k|k}\}_{j=st,ld}$. Then, $\hat{\mathbf{x}}^j_{k|k}$ is weighted by μ^j_k which is the probability that the system corresponds to the mode m_j at time k, where $m_j = m_{st}, m_{ld}$. All the weighted estimates are mixed together to get an estimation of the state vector $\hat{\mathbf{x}}_{k|k}$.

Designing an IMM seems quite an easy task but there are some issues to be addressed, especially the definitions of the transition probabilities between both SSRs.

3.2.2. About the transition probabilities

In the IMM, to update the mixing probabilities and the *a posteriori* mode probabilities, respectively denoted as $\mu^{l|j}_{k-1|k-1} = P\{M_{k-1} = m_l|M_k = m_j\}$ and μ^j_{k-1}, the system is assumed to be a Markov chain. The transition probabilities between the modes m_l at time $k-1$ and m_j at time k denoted by $p_{l,j}$ with $l,j = st, ld$ are stored in the TPM denoted as Π as follows:

$$\Pi = \begin{bmatrix} p_{st,st} & p_{st,ld} = 1 - p_{st,st} \\ p_{ld,st} = 1 - p_{ld,ld} & p_{ld,ld} \end{bmatrix} \qquad (8)$$

Concerning its setting, two cases can be considered:

On the one hand, Π can be *a priori* defined. As it depends on the two probabilities to stay in the states, namely $p_{st,st}$ and $p_{ld,ld}$, and that the latter are related to the mean sojourn time in a state[2], one could make an assumption on the mean sojourn time in each mode. Nevertheless, we would have to collect the spectral resource availability in a database at each cognitive terminal for a given time and localization. This is what we suggest in [6]. Nevertheless, the properties of this Markov chain could change after a certain time. In this case, Π would have to be modified.

On the other hand, Π can be jointly estimated with the state vector. In [7], four Bayesian algorithms are presented. They all aim at recursively updating the probability density function (pdf) $p(\Pi|z_{1:k})$ by using the following expression:

$$p(\Pi|z_{1:k}) = \frac{\mu_{k-1}\Pi\Lambda_k}{\mu_{k-1}\hat{\Pi}_{k-1}\Lambda_k} p(\Pi|z_{1:k-1}) \qquad (9)$$

where $\mu_{k-1} = [\mu^{st}_{k-1}, \mu^{ld}_{k-1}]^T$ and Λ_k are column vectors containing the *a posteriori* probabilities and the EKFs likelihoods at times $k-1$ and k respectively. In addition, $\hat{\Pi}_{k-1} = E[\Pi|z_{1:k-1}]$ is the estimation of Π given $z_{1:k-1}$. In the numerical integration method detailed in [7], $\hat{\Pi}_{k-1}$ is expressed as a weighted sum of a set of predefined matrices. The weights are recursively estimated with these steps:

- Choosing a set of q predefined matrices $\{\Pi^{(s)}\}_{s=1,\cdots,q}$, with transition probabilities $p^{(s)}_{st,st}$ and $p^{(s)}_{ld,ld}$,
- Calculating the weight $p^{(s)}_k$ assigned to each predefined matrix by using the *a posteriori* mode probabilities and the EKF likelihood functions.

Details about the estimations of the TPM and the state vector are given in **Algorithm 1**.

4. SIMULATIONS AND RESULTS

4.1. Case 1: Π never changes

Simulation protocol

The OFDM system in Fig.1 with a 16-QAM constellation and 128 subcarriers is considered. The non-linearity order P and the memory depth M of the Volterra model are respectively equal to 3 and 2. In addition, $\sigma^2_{ld} = 1$. Then, let us define the CR-PA model path (CMP) which is a set of Volterra-parameter values recorded during a communication.

A toy example is considered where the CMP is generated by a two-state Markov chain whose TPM is denoted Π^{CMP}, with transition probabilities $p^{CMP}_{st,st}$ and $p^{CMP}_{ld,ld}$. Both states

[2] $p_{j,j} = 1 - 1/E[t_j]$ for $j = st, ld$ where $E[t_j]$ is the mean sojourn time in mode j.

correspond to two random walks with covariance matrix $\overline{\mathbf{Q}}(m_{st})$ or $\overline{\mathbf{Q}}(m_{ld})$. Using the CMP, the PA output samples y_k are generated using (2) and then disturbed by the AWGN channel n_k to get a given signal-to-noise ratio (SNR). The TPM is estimated as well as the state vector by using **Algorithm 1**. In this case, $q = 225$ matrices $\Pi^{(s)}$ are used. The latter are defined by the transition probabilities $p^{(s)}_{ld,ld}$ and $p^{(s)}_{st,st}$ which can take 15 values from 0.1 to 0.999 corresponding to different mean sojourn times varying approximately from 1 to 1000 sampling time. Our approach is then compared with the m_{st}-EKF alone, the m_{ld}-EKF alone and [6] where the TPM is set to two different predefined values:
1/ $\Pi = \Pi^{CMP}$. It is labelled "true predefined value". The results we will obtain hence serve as a reference. However, this ideal case does not reflect the way we could use the approach in practice.
2/ $\Pi = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}$. In this case, no *a priori* information is provided. This is labelled "bad predefined value".

Comments on the online estimation of Π

Let us analyze the case if Π is accurately estimated. According to our simulations in the first protocol, the estimation performance depends on the way the CMP has been generated (and more particularly on Π^{CMP}) and on the set of the predefined matrices $\{\Pi^{(s)}\}_{s=1,\cdots,q}$.

1/ If Π^{CMP} is among the predefined matrices, two cases may happen:
When $p^{CMP}_{st,st}$ is not too close to 1, for instance when $\Pi^{CMP} = \begin{bmatrix} 0.9 & 0.1 \\ 0.5 & 0.5 \end{bmatrix}$, $\hat{\Pi}_k$ converges to Π^{CMP} after a few tens of samples.
When $p^{CMP}_{st,st}$ is close to 1, for instance when $\Pi^{CMP} = \begin{bmatrix} 0.999 & 0.001 \\ 0.9 & 0.1 \end{bmatrix}$, the first row of Π_k converges to the first row of Π^{CMP}, but the estimation of the second row is not necessarily reliable. This phenomenon can be explained by the following reason: in the steady state of the Markov chain, it can be easily shown that the probabilities π^{CMP}_{st} and π^{CMP}_{ld} to be in the states st and ld satisfy:

$$\pi^{CMP}_{st} = \frac{1 - p^{CMP}_{ld,ld}}{2 - p^{CMP}_{st,st} - p^{CMP}_{ld,ld}}$$

$$\pi^{CMP}_{ld} = \frac{1 - p^{CMP}_{st,st}}{2 - p^{CMP}_{st,st} - p^{CMP}_{ld,ld}}$$

Therefore, π^{CMP}_{st} is close to 1 and does not change much as long as $p^{CMP}_{ld,ld}$ is not close to $p^{CMP}_{st,st}$. As a consequence, the algorithm cannot really estimate $p^{CMP}_{ld,ld}$ and it selects all the predefined matrices having $p_{st,st} = 0.999$.

2/ If Π^{CMP} is not among the predefined matrices and if $p^{CMP}_{st,st}$ is close to 1, only the predefined matrices $\Pi^{(s)}$ whose transition probabilities $p_{st,st}$ are close to $p^{CMP}_{st,st}$ have a predominant role in the estimation of Π.

Algorithm 1 Adaptive IMM including TPM estimation algorithm - $(*)$ points out the additional steps regarding [6]

- $(*)$ Initially, setting the weights $p^{(s)}_0 = \frac{1}{q}$
- Computing the mixing probabilities, for $j, l = st, ld$:
$$\mu^{l|j}_{k-1|k-1} = \frac{1}{c^j} p_{l,j} \mu^l_{k-1} \text{ with } c^j = \sum_{l=st,ld} p_{l,j}\mu^l_{k-1}$$
- Deducing the merged means from the EKF estimates:
$$\hat{\mathbf{x}}^{0j}_{k-1|k-1} = \sum_{l=st,ld} \hat{\mathbf{x}}^l_{k-1|k-1}\mu^{l|j}_{k-1|k-1} \quad j = st, ld$$
- Computing the estimation error covariance matrices:
$$\mathbf{P}^{0j}_{k-1|k-1} = \sum_{l=st,ld} \mu^{l|j}_{k-1|k-1}\{\mathbf{P}^l_{k-1|k-1} + [\hat{\mathbf{x}}^l_{k-1|k-1} - \hat{\mathbf{x}}^{0j}_{k-1|k-1}][\hat{\mathbf{x}}^l_{k-1|k-1} - \hat{\mathbf{x}}^{0j}_{k-1|k-1}]^T\}$$
- Expressing the likelihood functions from each EKF:
$$\Lambda^j_k = \mathcal{N}(z_k; h(\hat{\mathbf{x}}^j_{k|k-1}), S^j_k) \quad j = st, ld$$
where $h(\hat{\mathbf{x}}^j_{k|k-1})$ is the predicted observation in the m_j-EKF by using $\hat{\mathbf{x}}^{0j}_{k-1|k-1}$.
- Updating the mode probabilities:
$$\mu^j_k = \frac{1}{c}\Lambda^j_k c^j, \quad j = st, ld \text{ where } c = \sum_{l=st,ld} \Lambda^j_k c^j$$
- (*) Updating the weights:
$$p^{(s)}_k = \frac{\mu_{k-1}\Pi^{(s)}\Lambda_k}{\mu_{k-1}\hat{\Pi}_{k-1}\Lambda_k} p^{(s)}_{k-1} \text{ with } s = 1, \cdots, q$$
- (*) Estimating Π: $\hat{\Pi}_k = \sum_{s=1}^{q} \Pi^{(s)} p^{(s)}_k$
- Mixing the EKF estimates to get a final state estimate:
$$\hat{\mathbf{x}}_{k|k} = \sum_{j=st,ld} \mu^j_k \hat{\mathbf{x}}^j_{k|k}$$

Comments on the state-vector estimation

The Volterra-parameter tracking performance are presented in Fig. 2 for one realization. For the sake of clarity, only one parameter is represented in the figure. When only using a m_{st}-EKF, the Volterra parameters can be estimated at the beginning. Then, their variations cannot be tracked. Using a m_{ld}-EKF, parameter tracking is faster but the estimate oscillates much. However, the time spent to reach the new Volterra value and the uncertainties on the estimates increase the BER. When using the IMM, both the convergence rapidity and the parameter estimate accuracy are obtained. Fig. 4 and Fig. 3 show that the IMM with TPM estimation performance is close to those of the IMM based detector when $\Pi = \Pi_{CMP}$. The TPM estimation provides a SNR gain of $0.86dB$ for BER equal to 10^{-3} compared to the case when no *a priori* information is provided.

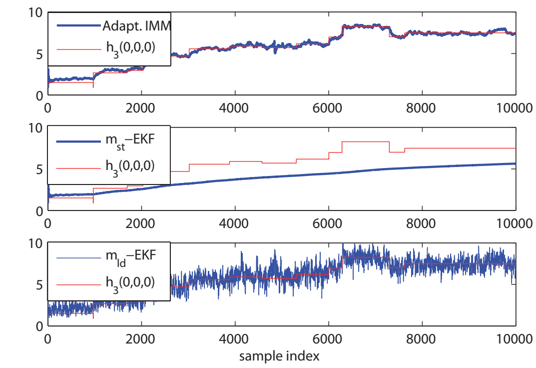

Fig. 2. Volterra parameter estimation comparison between the adaptive IMM, the m_{st}-EKF and the m_{ld}-EKF.

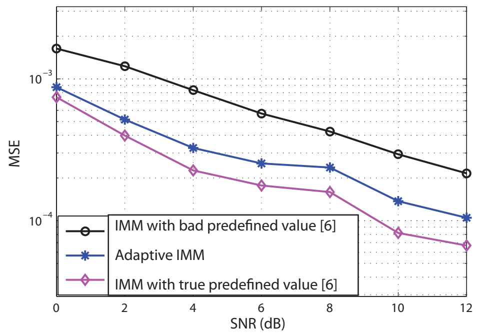

Fig. 3. MSE performance comparison.

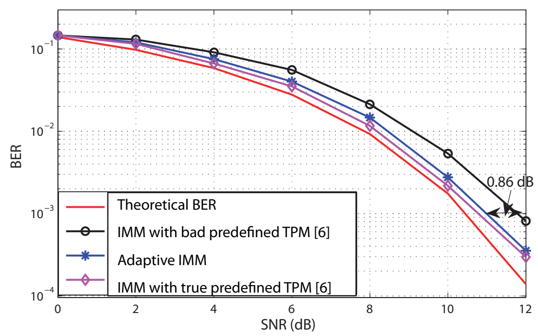

Fig. 4. BER performance comparison.

4.2. Case 2: when Π varies over time

In this case, we aim at testing if the TPM parameters can be tracked over time. The only difference between this second simulation protocol and the first one is that the CMP is now generated according to two successive TPMs, namely $\begin{bmatrix} 0.6 & 0.4 \\ 0.5 & 0.5 \end{bmatrix}$ and $\begin{bmatrix} 0.9 & 0.1 \\ 0.15 & 0.85 \end{bmatrix}$. According to Fig. 5, the transition probabilities can be tracked thanks to the adaptive IMM. It adapts itself to the changes of the mean sojourn time of each state over time whereas the approach we propose in [6] cannot. It provides a SNR gain of a few tenths dB for any BER compared to [6].

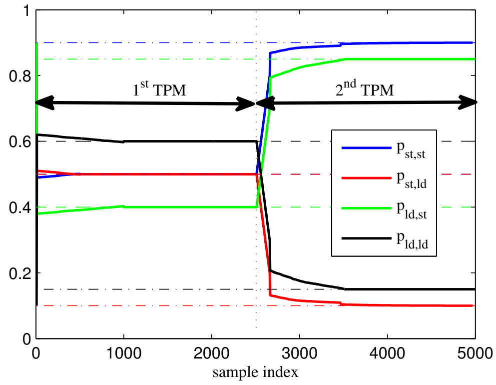

Fig. 5. TPM variation tracking.

5. CONCLUSIONS

To track the change of the CR-PA behavior, an IMM is used to jointly estimate the CR-PA model parameters and to restore the CR-PA input symbols. As the TPM choice impacts the performance of the algorithm in terms of BER, it is proposed to jointly estimate it with the state vector using the numerical integration method. This leads to an adaptive IMM. Our paper is also a complementary study to Jilkov's work [7] since we address particular cases of TPM and provide justifications of the performance of the estimation algorithm.

REFERENCES

[1] A. Goldsmith, S. A. 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.

[2] J. Vuolevi and T. Rahkonen, *Distortion in RF power amplifiers*, Artech House, 2003.

[3] S. C. Cripps, *Advanced techniques in RF power amplifier design*, Artech House, 2002.

[4] Y. Eun and D. Jeon, "Fuzzy inference-based dynamic determination of IMM mode transition probability for multi-radar tracking," Proceedings of the Intern. Conf. on Information Fusion, pp. 1520-1525, 2013.

[5] L. Bloomer and J. E. Gray, " Are more models better?: the effect of the model transition matrix on the IMM filter," Proceedings of the Thirty-Fourth Southeastern Symposium on System Theory, pp. 20-25, 2002.

[6] M. B. Mabrouk, G. Ferré, E. Grivel and N. Deltimple, "Interacting Multiple Model Based Detector to Compensate Power Amplifier Distortions in Cognitive Radio," IEEE Transactions on Communications, vo. 63, issue. 05, pp. 1580-1593, 2015.

[7] V. P. Jilkov and X. R. Li, "Online Bayesian estimation of transition probabilities for Markovian jump systems," IEEE Trans. on Signal Processing, vol. 52, no. 6, pp. 1620-1630, 2004.

[8] J. K. Tugnait, "Adaptive estimation and identification for discrete systems with Markov jump parameters," IEEE Trans. on Automatic Control, vol. 27, no. 05, pp. 1054-1065, 1982.

[9] C. E. Seah, I. Hwang, "State estimation for stochastic linear hybrid systems with continuous-state-dependent transitions: an IMM approach," IEEE Trans. on Aerospace and Electronic Systems, vol. 45, no. 1, 2009.

[10] W. J. Rugh, *Nonlinear system theory: the Volterra/ Wiener approach*, Chapter 1, The Johns Hopkins University Press, 1981.

[11] Y. Bar-Shalom, X. R. Li and T. Kirubarajan, *Estimation with applications to tracking and navigation*, John Wiley & Sons, 2001.

[12] I. Arasaratnam and S. Haykin, "Cubature Kalman filters," IEEE Trans. on Automatic Control, vol. 54, no. 6, pp. 1254-1269, 2009.