Use of Bayesian Changepoint Detection for Spectrum Sensing in Mobile Cognitive Radio

Akpaki Steaven V. Chede*†, François Rottenberg*, Michel Dossou†, and Jérôme Louveaux*
*ICTEAM, Université catholique de Louvain, Belgium
{akpaki.chede, francois.rottenberg, jerome.louveaux}@uclouvain.be
†LETIA, Ecole Polytechnique d'Abomey-Calavi (EPAC), Université d'Abomey-Calavi (UAC), Benin
michel.dossou@epac.uac.bj

***Abstract*—One important problem in spectrum sensing is to detect a noisy and unknown signal, while keeping the risk of detection error as low as possible. This problem may increase in mobile environments due to fast situation changes. In this paper, we consider a mobile cognitive radio scenario, and try to evaluate whether some knowledge about the environment and the mobility parameters of the user can help in improving the detection of changes in the spectrum occupancy. To do so, we assume that the mobility parameters can be summarized in some a priori knowledge on the average time of spectrum change and we use Bayesian changepoint detection methods. Considering that the power of the signal to be detected is usually unknown, a low-complexity algorithm is proposed that does not rely on this knowledge. It is then compared with the existing algorithms in the literature. Finally, a new metric is introduced to jointly evaluate the costs of interference and spectrum waste induced by the changepoint detection algorithms, in a time-limited communication context. Results reveal that the derived algorithm outperforms its non-Bayesian equivalent at low signal to noise ratio (SNR).**

***Index Terms*—mobile cognitive radio, spectrum sensing, changepoint detection**

I. Introduction

Cognitive radio (CR) is intended to increase the spectral efficiency in wireless communication systems, by allowing opportunistic access to spectrum holes. Spectrum sensing is a critical step in CR because it directly impacts the amount of interference caused on the existing primary users (PUs) and the amount of spectrum that will be available for the secondary user (SU). In mobile environments, the spectrum occupancy becomes more dynamic as the SU is moving. One may wonder if it is possible to use parameters such as speed and mobility patterns to make better decisions.

Classical detection paradigms are hypotheses tests that try to know which distribution the observed samples belong to, either sample by sample (sequential test) or considering a block of samples. Unlike those paradigms, the sequential changepoint detection (or quickest detection) wants to detect, as quickly as possible, a change in the distribution of the observed samples. It may be a better alternative to classical detection for spectrum sensing in mobility scenarios because of the need to quickly detect a PU (or another SU) appearance or disappearance when moving. Several approaches or frameworks of quickest detection are found in the literature. On the one hand, the *minimax approach* assumes that the changepoint is deterministic and unknown. On the other hand, the *Bayesian approach* assumes that the changepoint is a random variable with a known prior distribution.

The theory of changepoint detection includes many algorithms such as the cumulative sum (CUSUM) procedure, which is proved to be optimal in the minimax framework for independent and identically distributed (i.i.d) observations [1], [2], and the Shiryaev procedure which is optimal in the Bayesian framework. Comparing them, it is shown in [3] that CUSUM loses its asymptotic optimality (in terms of minimum average detection delay) in the Bayesian framework, when the changepoint has a prior distribution with exponentially decreasing tail, but it remains optimal for heavy-tailed priors. Moreover, authors in [4] showed that Shiryaev algorithm performs much better than CUSUM when the difference between the pre-change and the post-change distributions – measured as the Kullback-Leibler divergence – decreases. In the CR context, this corresponds to a low power signal to detect in noise, or in other words, a low SNR. In practice, the level of the signal to detect (or equivalently the SNR) is usually not known in advance. Various solutions help to deal with this. The generalized likelihood ratio (GLR) test, for instance, replaces the unknown parameter of the distribution by its maximum likelihood estimate [5]. The mixture-based test averages the decision statistic over the unknown parameter, assuming that the prior distribution of this one is known. Other approaches include non-parametric tests, and the interesting M-Shiryaev algorithm [6], [7], which is used when the unknown post-change parameter belongs to a discrete set of values.

Many works, in the literature, apply quickest detection to CR. For example, authors of [8] developed a changepoint detection framework, using CUSUM, GLR-based CUSUM and a nonparametric test, for different prior information available at the SU. Wideband quickest spectrum sensing is studied in [9], for independent channels, and in [10], for correlated channels. Impact of fading, multiple antennas [11] and collaborative schemes [12] over CUSUM is also covered.

In this paper, we investigate the possibility of using changepoint detection in mobile cognitive radio. We postulate that the prior distribution of the changepoint could be, in practice, inferred from mobility parameters. Therefore, we consider the Shiryaev algorithm which is optimal in the Bayesian framework. Given that the SU does not usually know the signal power of PUs or other SUs using the spectrum, we use the GLR approach. As the GLR-based Shiryaev involves a complex optimization problem, we derive a low-complexity GLR test (termed as *LC GLR-Shiryaev*) for Shiryaev algorithm. Our goal is to compare the derived algorithm with other existing algorithms (CUSUM, GLR-CUSUM, Shiryaev, M-Shiryaev) for realistic CR scenarios, in order to know in which situations it may offer the best performance. Moreover, unlike previous works, we consider that both interference and spectrum waste durations are important to evaluate the performance of detection algorithms in a CR context. Hence, we introduce a new metric, called *global penalty*, which

978-1-7281-8964-2/21/$31.00 ©2021 IEEE

considers both aspects. The simulations results show that the derived LC GLR-Shiryaev algorithm is a good choice for a practical mobile CR, when the SU does not have any *a priori* knowledge about the SNR and could be required to detect very low SNR signals. A new way of setting the optimal threshold based on the global penalty is also proposed.

II. System Model

Let us assume a system with a PU transmitting in a frequency band and a SU willing to opportunistically use the same band. It is assumed that a first spectrum sensing has been performed and that the SU starts using the available band. However, the SU may experience mobility. While moving, the spectrum occupancy may thus change, due to getting in range of a PU or another SU. For simplicity, this potentially interfering user is called the PU below, but the case of another SU is completely identical. In order to detect potential changes in a given band, the SU is continuously measuring the activity in this band. The observed signal sample at time m is denoted by $y[m]$. When there is no PU activity, it is given by $y[m] = w[m]$, where $w[m]$ is the Additive White Gaussian Noise (AWGN) sample. When the PU transmits a signal, the SU observes $y[m] = s[m] + w[m]$, $s[m]$ being the transmitted signal affected by fading. It is assumed that $s[m]$ and $w[m]$ are independent circularly symmetric complex Gaussian, $s[m] \sim \mathcal{CN}(0, \sigma_s^2)$, $w[m] \sim \mathcal{CN}(0, \sigma_w^2)$, and independent of each other.

We assume that, initially, the PU is not transmitting (or is not visible) and the observed samples follow a Gaussian distribution with variance σ_w^2. At an unknown time sample τ, the PU appears while the SU is moving, and the distribution changes to a Gaussian distribution with variance $\sigma_s^2 + \sigma_w^2$. At each time sample n, the SU tries to detect the possible change in the distribution, based on the received sequence $y_1^n = y[1], ..., y[n]$ up to the current time sample. The SU should thus distinguish between the following two hypotheses

$$\begin{cases} \mathcal{H}_0: & y[m] = w[m], \qquad m = 1, ..., n \\ \mathcal{H}_1: & \exists\, \tau \in [0, n], \text{ such that} \\ & y[m] = w[m], \qquad m = 1, ..., \tau \\ & y[m] = s[m] + w[m], \quad m = \tau + 1, ..., n \end{cases}$$

where $\mathcal{H}_0$ means there has not been any change in the received sequence distribution up to now, and $\mathcal{H}_1$ means there has been a change at time sample τ. As we consider the Bayesian framework, it is assumed that τ is a random variable independent of the observations, with a geometric prior. This choice is motivated by the fact that the experience of finding the changepoint can be viewed as a sequence of trials with a certain number of failures (no change) before the first success (changepoint). Furthermore, we postulate that the parameter p of this geometric prior can be known with a certain degree of accuracy, if the speed of the SU and its environment are known. Intuitively, a high speed of the SU may correspond to a high value of p because the spectrum situation may change more frequently.

III. Classical changepoint detection algorithms

In this section, the Bayesian approach of changepoint detection and the most popular changepoint detection algorithms are reviewed. In general, the stopping time of an algorithm (time sample when a change is detected) will be denoted by T.

A. Deterministic approach - CUSUM algorithms

1) CUSUM: At each time sample n, the CUSUM algorithm compares a decision statistic C_n with a threshold α and an alarm is raised as soon as C_n exceeds α. Its stopping time is thus given by [8]

$$T_{\text{CUSUM}} = \inf\{n \geq 1 : C_n \geq \alpha\}, \text{with } C_n = \max_{k \leq n} \sum_{m=k+1}^{n} \ln L_m, \tag{1}$$

where L_m is the likelihood ratio of the m^{th} observation, given by $L_m = \frac{\sigma_w^2}{\sigma_s^2+\sigma_w^2} \exp\left(\frac{|y(m)|^2\, \sigma_s^2}{\sigma_w^2\,(\sigma_s^2+\sigma_w^2)}\right)$.

2) GLR-CUSUM: When σ_s^2 is unknown, it is replaced by the value that maximizes C_n and the decision statistic becomes

$$C_{n,\text{GLR}} = \max_{k \leq n} \sup_{\sigma_s^2} \sum_{m=k+1}^{n} \left(\frac{|y(m)|^2\, \sigma_s^2}{\sigma_w^2\,(\sigma_s^2 + \sigma_w^2)} + \ln \frac{\sigma_w^2}{\sigma_s^2 + \sigma_w^2}\right). \tag{2}$$

B. Bayesian approach - Shiryaev algorithm

The Bayesian approach [13] assumes that τ follows a specific prior distribution, and looks for the optimal algorithm that minimizes the average detection delay A_{DDl} under a constraint on the false alarm probability P_{FA}. The false alarm probability is defined as

$$P_{FA} = P(T \leq \tau) = \sum_{k=1}^{\infty} p(1-p)^k P(T \leq k). \tag{3}$$

The average detection delay is defined as

$$A_{DDl} = \mathbb{E}(T - \tau \mid T > \tau) = \frac{\mathbb{E}(T-\tau)^+}{P(T > \tau)} \tag{4}$$

where $x^+ = max\{0, x\}$. The optimal algorithm is the Shiryaev algorithm. As expressed in [13], its stopping time is

$$T_{\text{Sh}} = \inf\{n \geq 1 : C_{n,p} \geq \alpha\}, \text{with } C_{n,p} = \sum_{k=1}^{n} \prod_{m=k}^{n} \frac{L_m}{1-p} \tag{5}$$

where α is selected in such a way that $P_{FA} \leq \omega$, $\omega \in]0, 1[$.

IV. Low-complexity GLR-Shiryaev algorithm

In this paragraph, the new LC GLR-Shiryaev algorithm is introduced. It is obtained by applying the GLR approach to Shiryaev algorithm, for situations where σ_s^2 is unknown to the SU. Similarly to the GLR-CUSUM, the statistic $C_{n,p}$ has to be maximized over the set of possible values of σ_s^2. The GLR-Shiryaev statistic is given by

$$\begin{aligned} \tilde{C}_{n,p} &= \sup_{\sigma_s^2} C_{n,p}(\sigma_s^2) \\ &= \sup_{\sigma_s^2} \sum_{k=1}^{n} \prod_{m=k}^{n} \left(\frac{1}{1-p}\right) \frac{\sigma_w^2}{\sigma_s^2 + \sigma_w^2} \exp\left(\frac{|y(m)|^2\, \sigma_s^2}{\sigma_w^2\,(\sigma_s^2 + \sigma_w^2)}\right) \\ &= \max_{\sigma_s^2 \geq 0} \sum_{k=1}^{n} \frac{1}{(1-p)^{n-k+1}} \left(\left(\frac{\sigma_w^2}{\sigma_s^2 + \sigma_w^2}\right)^{n-k+1} \right. \\ &\qquad \left. \exp\left(\frac{\sigma_s^2}{\sigma_w^2\,(\sigma_s^2 + \sigma_w^2)} \sum_{m=k}^{n} |y(m)|^2\right)\right). \end{aligned}$$

The metric $C_{n,p}(\sigma_s^2)$ can be shown to be a non-convex function and its derivative is a nonlinear function whose zeros are hard to find analytically. To relax the optimization problem, we make the following approximation. Instead of directly maximizing the sum $C_{n,p}$, each term of the sum is maximized independently, making the problem easier to solve. In other words, rather than computing the σ_s^2 that maximizes $C_{n,p}$, we compute for each term of $C_{n,p}$, the $\sigma_{s,k}^2$ that maximizes this term. The new decision statistic is thus

$$\tilde{C}_{n,p} \simeq \sum_{k=1}^{n} \frac{1}{(1-p)^{n-k+1}} \sup_{\sigma_{s,k}^2} \left(\left(\frac{\sigma_w^2}{\sigma_{s,k}^2 + \sigma_w^2} \right)^{n-k+1} \exp\left(\frac{\sigma_{s,k}^2}{\sigma_w^2 \left(\sigma_{s,k}^2 + \sigma_w^2\right)} \sum_{m=k}^{n} |y(m)|^2 \right) \right). \tag{6}$$

Looking at one term at a time, the logarithm of the k^{th} term is redefined as $g_k(\sigma_{s,k}^2)$, and the optimization problem that needs to be solved is the following

$$\max_{\sigma_{s,k}^2 \geq 0} \quad g_k(\sigma_{s,k}^2) = (n-k+1)\ln\left(\frac{\sigma_w^2}{\sigma_{s,k}^2+\sigma_w^2}\right) + \frac{\sigma_{s,k}^2}{\sigma_w^2\left(\sigma_{s,k}^2+\sigma_w^2\right)} \sum_{m=k}^{n} |y(m)|^2. \tag{7}$$

The KKT conditions for this problem are given by

$$\begin{cases} \frac{1}{(\sigma_{s,k}^2+\sigma_w^2)^2} \sum_{m=k}^{n} |y(m)|^2 - \frac{n-k+1}{\sigma_{s,k}^2+\sigma_w^2} + \lambda = 0 \\ \sigma_{s,k}^2 \geq 0 \\ \lambda \geq 0 \\ \lambda = 0 \text{ or } \sigma_{s,k}^2 = 0 \end{cases}$$

where λ is the Lagrange multiplier. The solutions can be split in two cases:

- **Case 1:** $\sigma_{s,k}^2 > 0 \implies \lambda = 0, \sigma_{s,k}^2 = \frac{\sum_{m=k}^{n} |y(m)|^2}{n-k+1} - \sigma_w^2$, which is obtained if $\sum_{m=k}^{n} |y(m)|^2 \geq (n-k+1)\sigma_w^2$.
- **Case 2:** $\sigma_{s,k}^2 = 0 \implies \lambda = \frac{(n-k+1)\sigma_w^2 - \sum_{m=k}^{n} |y(m)|^2}{\sigma_w^4}$, which is obtained for the reversed inequality, and thus satisfies $\lambda \geq 0$.

Consequently, the solution can be compactly written as $\tilde{\sigma}_{s,k}^2 = \left\{ \frac{\sum_{m=k}^{n} |y(m)|^2}{n-k+1} - \sigma_w^2 \right\}^+$.
Inserting it in (6), the decision statistic of the *low-complexity GLR-Shiryaev* algorithm becomes

$$\tilde{C}_{n,p} \simeq \sum_{k=1}^{n} \frac{1}{(1-p)^{n-k+1}} \left(\left(\frac{\sigma_w^2}{\tilde{\sigma}_{s,k}^2 + \sigma_w^2} \right)^{n-k+1} \exp\left(\frac{\tilde{\sigma}_{s,k}^2}{\sigma_w^2 \left(\tilde{\sigma}_{s,k}^2 + \sigma_w^2\right)} \sum_{m=k}^{n} |y(m)|^2 \right) \right). \tag{8}$$

In summary, at each time sample n, the SU just has to compute the decision statistic $\tilde{C}_{n,p}$ in (8) and compare it with the threshold α, to decide of the PU appearance.

V. Numerical Results

In order to assess the precision of the approximation (6), done in deriving LC GLR-Shiryaev, the initial optimization problem could be solved numerically by using a grid search over a finite set of values for σ_s^2. An equivalent method for doing that is the M-Shiryaev algorithm presented in [6]. M-Shiryaev consists in running multiple (M) parallel Shiryaev procedures for each of the possible values of σ_s^2 and deciding the PU appearance when any one of the procedures stops [7].

A. Simulation setup

It is assumed that the SU has a limited amount of time, defined by the number N of measurement samples, for secondary communication and τ is geometrically generated with parameter p. As mentioned in section II, the value of p could be, in practice, inferred from the SU speed, and other characteristics of the environment. It should also be mentioned that it depends on the sampling frequency of the SU. In a practical mobile CR scenario, the SU is expected to be able to transmit for some reasonable amount of time, so the value of p is rather small.

B. Global Penalty

In CR, it is desirable to reduce the interference with the PU as much as possible but also to maximize the spectrum usage. For this reason, we define a new criterion that tries to take both effects into account. The criterion also has to take into account for how long the interference or spectrum waste is present with respect to the duration of the communication. For this reason, the average detection delay, representing the interference duration, is first redefined as

$$A_{DD} = \mathbb{E}\Big(\min\{T, N\} - \min\{\tau, N\} \Big)^+. \tag{9}$$

This varies from the usual average detection delay definition of the literature (4) in two ways. First, it is no longer conditional to the correct detection event $\{T > \tau\}$, but the average is done over all cases (correct detection and false alarm) since the cases with $\{T < \tau\}$ must be counted as creating no interference penalty. Moreover, it also handles the cases where there is no detection until the end of the frame ($T = N$), and the cases where the PU does not appear during the frame ($\tau > N$). Similarly, the *average false alarm duration* (A_{FAD}) is defined as the average time between a false alarm decision and the changepoint

$$A_{FAD} = \mathbb{E}\Big(\min\{\tau, N\} - \min\{T, N\} \Big)^+. \tag{10}$$

This metric represents the average time of spectrum waste during the time frame of communication. The penalties induced by interference and waste of spectral opportunity are respectively denoted by P_{int} and P_{wso}. Finally, the overall ***global penalty*** (G_P) induced by the detection algorithm during the time frame N can be defined as

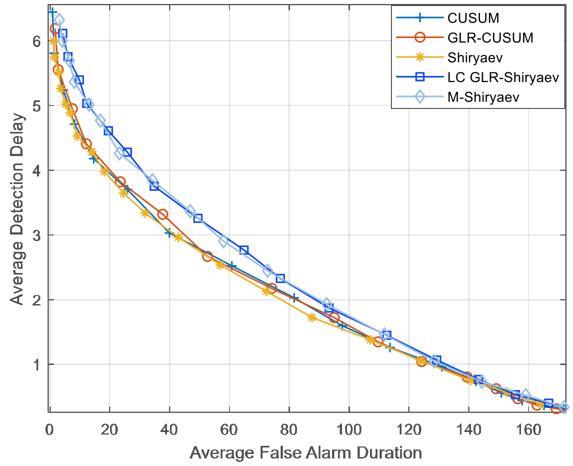
$$G_P = P_{int}\ A_{DD} + P_{wso}\ A_{FAD}. \tag{11}$$

C. Simulations results & Discussion

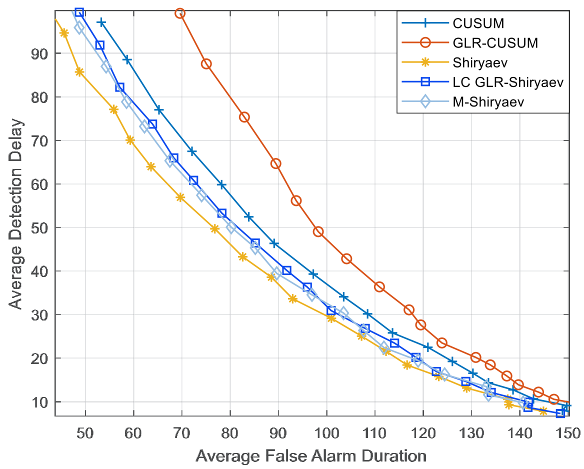
The simulation parameters are given the following values: $p \in \{0.005, 0.05\}$, received SNR $= \frac{\sigma_s^2}{\sigma_w^2} \in \{-10, 0, 5\}\,dB$ when the PU is present, SU's communication duration $N = 1000$ samples, penalties[1] $P_{int} = 0.5$ and $P_{wso} = 0.14$.

In order to concurrently visualize the interference and the spectral opportunity waste induced by each algorithm, the A_{DD} vs A_{FAD} curve, obtained by varying the threshold, is represented in Fig. 1. It can be viewed as a kind of complementary receiver operating characteristics (ROC) curve for changepoint spectrum sensing algorithms. Shiryaev algorithm has the best performance at every SNR value. CUSUM and GLR-CUSUM have comparable performance with Shiryaev

[1] To compute the penalties in this paper, we make a simple assumption of a multi-band transmission (20 sub-bands at least) with uniform power allocation and compute the loss in the total channel rate when the SU fails in detecting the PU activity on a sub-band (interference) and when it releases a sub-band before the PU appearance (spectrum waste).

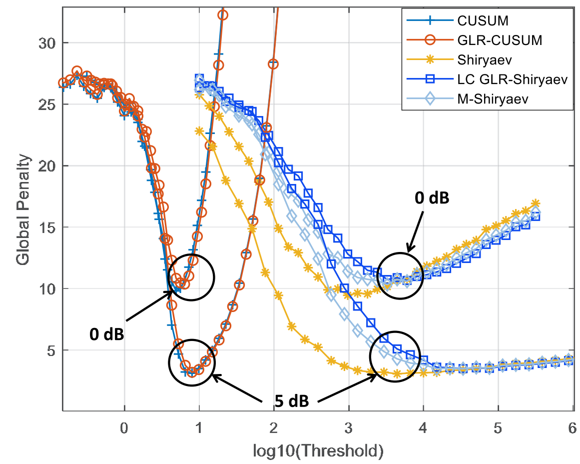


(a) p = 0.005, SNR = 5 dB.

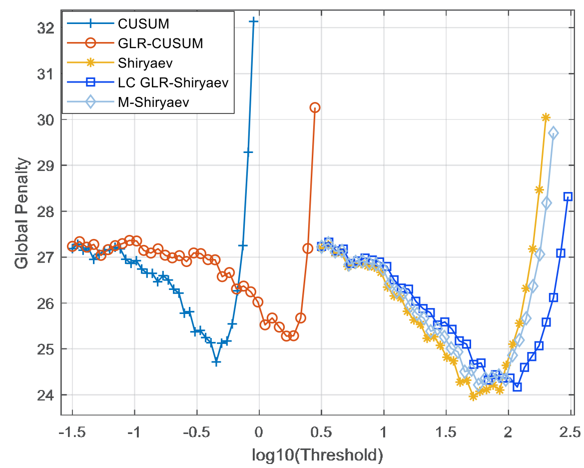


(b) p = 0.005, SNR = -10 dB.

Fig. 1: A_{DD} vs A_{FAD} for p = 0.005.



(a) p = 0.005, SNR = 5 dB & 0 dB.



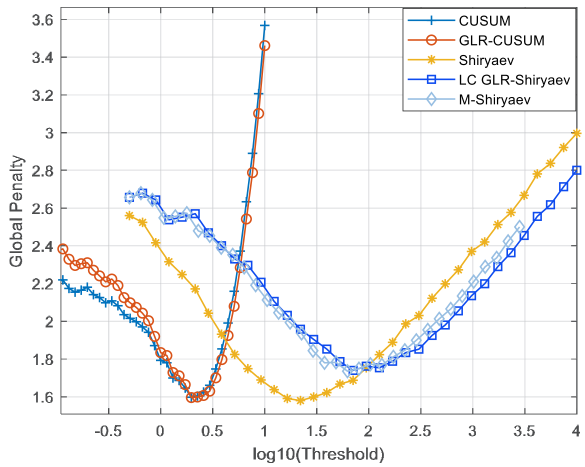
(b) p = 0.005, SNR = -10 dB.

Fig. 2: G_P vs $log_{10}(\alpha)$ for p = 0.005.

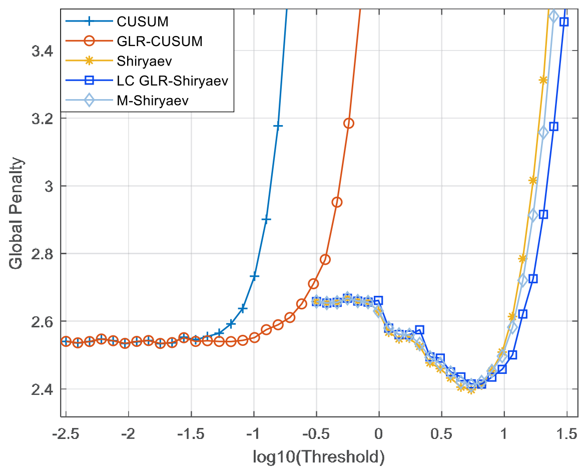
at high SNR values, whereas LC GLR-Shiryaev and M-Shiryaev outperform them at low SNR values. As it can be seen, LC GLR-Shiryaev and M-Shiryaev have comparable performance in every cases, showing that the approximation (6) is acceptable. It can be shown with further simulations that when a narrower range of possible values of σ_s^2 is chosen around the true value, the M-Shiryaev algorithm approaches the optimal Shiryaev algorithm. However this requires an accurate *a priori* knowledge of the SNR which is usually not available. LC GLR-Shiryaev has the advantage of not requiring any *a priori* knowledge as it uses an analytical form. It also does not require to compute multiple parallel procedures.

The choice of the threshold α is an important issue and choosing a specific value will put the algorithm at a specific operating point on the A_{DD} vs A_{FAD} curve. It is desirable to choose a threshold that gives a good compromise between A_{DD} and A_{FAD}. We suggest to fix this threshold with the help of the G_P vs $log_{10}(\alpha)$ curves plotted in Fig. 2 and Fig. 3. First, it can be seen that, for every algorithm, a minimum global penalty exists, and the corresponding threshold depends on the SNR and the value of p. By comparing Fig. 2 (a) and Fig. 2 (b), the same tendency as before can be noticed: LC GLR-Shiryaev and M-Shiryaev yield lower minimum global penalties than CUSUM and GLR-CUSUM at low SNR, for $p = 0.005$. The same trend is observed when p is increased tenfold ($p = 0.05$) in Fig. 3. Moreover, looking at Fig. 2, it can be noticed that CUSUM and GLR-CUSUM are more sensitive to the threshold variation, i.e., small variations of the threshold around its optimal value can result in high variations of the global penalty.

These results show that Shiryaev-based algorithms work better than CUSUM-based algorithms at low SNR. This can be explained by the fact that at low SNR, the PU's signal becomes more difficult to distinguish from noise and the prior knowledge of the PU appearance can help increasing the detection performance. These results are consistent with the literature on changepoint detection theory, since a low value of the Kullback-Leibler information corresponds to a low value of the SNR in a CR context [4], [13]. This advantage of Shiryaev-based algorithms over CUSUM-based algorithms in low SNR ranges is valuable in practical mobile CR. In fact, multipath fading and shadowing can happen very frequently and the SU



(a) p = 0.05, SNR = 5 dB.



(b) p = 0.05, SNR = -10 dB.

Fig. 3: G_P vs $log_{10}(\alpha)$ for p = 0.05.

could be required to detect primary signals with very low SNR (low as -20 dB, according to IEEE 802.22 wireless regional area network standard) [14].

D. Optimal threshold setting

In the literature, the optimal threshold is classically chosen so that the false alarm probability is below a given value. In this work, we suggest to choose the optimal threshold as the one that minimizes the global penalty. We have noticed that this optimal threshold depends on p and the SNR, but is approximately independent of the time frame duration N. If the SNR is known, the optimal thresholds can be pre-computed offline and recovered from a table during the actual sensing. Based on simulations, the optimal thresholds have been computed for a set of given SNR values. The results are presented in TABLE I below. When the SNR is unknown, the SNR could be estimated first or a fixed threshold, independent from the SNR, could be used.

VI. Conclusion

In this work, Bayesian changepoint detection is applied to mobile cognitive radio. A low-complexity algorithm, based on the Shiryaev algorithm and using the GLR test, is introduced. It shows a reduced complexity, when compared with the real GLR-based Shiryaev which needs to solve a complex optimization problem at each time sample. Moreover it gives better performance than the GLR-based CUSUM in low SNR situations, thanks to the prior knowledge of the PU appearance. We assume that the a priori probability of spectrum change could be inferred from mobility parameters but do not consider a specific model relating them in this paper. Such a model may be derived either mathematically, using stochastic geometry, or experimentally, by means of field measurements. It thus may be subject for future research. Collaborative schemes of the proposed LC GLR-Shiryaev are other interesting perspectives to investigate.

TABLE I: Optimal thresholds computed through simulations

SNR (dB) / Algorithm	-10	-5	0	5
		p=0.005		
CUS	0.4516	1.8248	4.8819	7.9771
GLR-CUS	1.6521	2.9753	6.3437	7.9771
Shi	52.2811	171.9072	966.7053	4.4367e+03
LC GLR-Shi	117.8190	258.0862	6.0174e+03	3.3839e+04
M-Shi	57.4652	258.0862	3.2712e+03	3.3839e+04
		p=0.05		
CUS	0.0179	0.1317	0.5987	1.9684
GLR-CUS	0.0179	0.1838	0.6471	1.9684
Shi	5.4419	6.0174	6.7289	22.0537
LC GLR-Shi	6.5786	9.3193	10.3830	70.7107
M-Shi	5.4419	5.3150	9.2552	62.3929

References

[1] G. Lorden, "Procedures for Reacting to a Change in Distribution," *The Annals of Mathematical Statistics*, vol. 42, pp. 1897–1908, jul 1971.

[2] G. V. Moustakides, "Optimal stopping times for detecting changes in distributions," *Annals of Statistics*, vol. 14(4), pp. 1379–1387, 1986.

[3] A. G. Tartakovsky and V. V. Veeravalli, "General asymptotic Bayesian theory of quickest change detection," *Theory of Probability and its Applications*, vol. 49, pp. 458–497, oct 2005.

[4] A. G. Tartakovsky and G. V. Moustakides, "State-of-the-art in bayesian changepoint detection," *Sequential Analysis*, vol. 29, pp. 125–145, apr 2010.

[5] A. Mariani, A. Giorgetti, and M. Chiani, "Effects of noise power estimation on energy detection for cognitive radio applications," *IEEE Transactions on Communications*, vol. 59, pp. 3410–3420, December 2011.

[6] J. Geng and L. Lai, "Bayesian quickest detection with unknown post-change parameter," in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, vol. 2016-May, pp. 4169–4173, may 2016.

[7] J. Geng, E. Bayraktar, and L. Lai, "Multi-Chart Detection Procedure for Bayesian Quickest Change-Point Detection with Unknown Post-Change Parameters," tech. rep., 2017.

[8] L. Lai, Y. Fan, and H. V. Poor, "Quickest detection in cognitive radio: A sequential change detection framework," in *GLOBECOM - IEEE Global Telecommunications Conference*, pp. 2957–2961, 2008.

[9] Q. Zhao and J. Ye, "Quickest detection in multiple on-off processes," *IEEE Transactions on Signal Processing*, vol. 58, pp. 5994–6006, dec 2010.

[10] A. Tajer and J. Heydari, "Quickest Wideband Spectrum Sensing over Correlated Channels," *IEEE Transactions on Communications*, vol. 63, pp. 3082–3091, sep 2015.

[11] E. Hanafi, P. A. Martin, P. J. Smith, and A. J. Coulson, "Quickest spectrum sensing with multiple antennas in rician and correlated rayleigh channels," *IEEE Communications Letters*, vol. 18, no. 8, pp. 1455–1458, 2014.

[12] T.-C. Hsu, T.-Y. Wang, and Y.-W. P. Hong, "Collaborative Change Detection for Efficient Spectrum Sensing in Cognitive Radio Networks," in *2010 IEEE 71st Vehicular Technology Conference*, pp. 1–5, jul 2010.

[13] A. Tartakovsky, I. Nikiforov, and M. Basseville, *Sequential Analysis*. Chapman and Hall/CRC, 1st ed., aug 2014.

[14] S. Atapattu, C. Tellambura, H. Jiang, and N. Rajatheva, "Unified Analysis of Low-SNR Energy Detection and Threshold Selection," *IEEE Transactions on Vehicular Technology*, vol. 64, pp. 5006–5019, nov 2015.