Sequential Likelihood Ratio Test for Cognitive Radios

Julien Renard, *Student Member, IEEE*, Lutz Lampe, *Senior Member, IEEE,* and Francois Horlin, *Member, IEEE*

***Abstract*—This paper presents a sequential likelihood ratio test (SLRT) detector for spectrum sensing scenarios in cognitive radios (CRs). Similar to other CR detectors, we exploit the structure of the sample covariance matrix of the received signal to achieve detection with minimal information regarding the signal. Unlike the majority of covariance-based CR detectors, the SLRT is a sequential detector that allows for smaller detection delays, which is a prized asset in CR systems. Using methods borrowed from the theory of continuous-time diffusion processes, we derive the statistical properties of the SLRT detector and compare it with an eigenvalue-based sequential detector which has been presented in previous work for CR systems. The comparison also includes detection scenarios with non-Gaussian noise to illustrate the robustness of the proposed detector in these situations.**

***Index Terms*—Sequential Log-Likelihood Ratio, Spectrum Sensing, Cognitive Radios, Bayesian detection**

I. Introduction

Spectrum sensing for cognitive radios (CRs) has been extensively studied for the past decade and several detectors have emerged as a result, see e.g. [1], [2], [3]. Energy detection, cyclic-feature detection and covariance-based detection are the most commonly cited types of spectrum sensing strategies for CRs with many research articles dedicated to their performance in typical CR detection scenarios: blind detection, low signal-to-noise ratio (SNR), noise variance uncertainties, non-Gaussian noise, etc. (see e.g. [4], [5], [6], [7], [8], [9]). Emphasis is usually put on the reliability of the detector and several strategies have been proposed to improve it, using for instance collaborative sensing schemes [10], [3] and nonparametric detectors [11], [12], [13]. However, detection delays are just as important as reliability in CR detectors in order to avoid causing harmful interferences to primary users and must fall within the specifications of the established standards [14], [15]. While the majority of the aforementioned detectors have been derived as batch detectors, sequential detectors offer significant advantages with respect to detection delays and constitute the main subject of this paper.

Batch processing algorithms assume that a given number of samples are available at the detector. These are used to reach a decision regarding the presence or absence of a communication signal (hypotheses H_1 and H_0, respectively) and are discarded afterwards. The detector always assumes that only one hypothesis is valid over the entire observation window (i.e. all samples belong to H_0 or H_1) and often tries to optimize detection performance for such scenarios, using for instance the Neyman-Pearson criterion. As pointed out in [16], such detectors are not optimal if one considers that the signal can appear or disappear at any random time (i.e. change-point detection scenarios). On the other hand, sequential detectors can minimize detection delays for such scenarios and outperform batch detectors [17]. The tracking property of a sequential detector also helps to reliably detect signals in changing environments plagued with noise and signal parameters uncertainties [18], a common assumption for CR detection scenarios. Moreover, CR systems need to continuously monitor their environment in order to detect opportunities for their own transmissions as well as to avoid interferences with other transmitters making sequential detectors a logical choice when they can be implemented[1].

The theory for sequential detection can be traced

This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC).

L. Lampe is with the Department of Electrical and Computer Engineering (ECE) at the University of British Columbia (UBC), Vancouver, Canada (email: Lampe@ece.ubc.ca).

F. Horlin is with the Department of "Optique, Photonique, Electromagnetisme, Radio-communications et Acoustique" (OPERA) at the Universite Libre de Bruxelles (ULB), Brussels, Belgium (email: fhorlin@ulb.ac.be).

J. Renard is with both the ECE and OPERA Departments at UBC and ULB (email: jrenard@ece.ubc.ca).

[1]As mentioned in [17], a drawback of sequential detectors comes from the mathematical tractability of their statistical distribution. Several assumptions about the received signal are needed to derive a useful analytical expression for the detector statistic.

back to the 1960's, and several sequential detectors have already been developed for CR systems. For instance, a Bayesian detector exploiting the structure of the covariance matrix to differentiate uncorrelated Gaussian noise from man-made signals has been presented in [18]. The system periodically updates prior distributions associated with the covariance matrix models used. More recently, subspace tracking techniques [19] have been used in [20] to derive a sequential largest-eigenvalue detector that also circumvents the need to explicitly compute covariance matrix estimates or their eigenvalues. The authors provide analytical results regarding the statistical distribution of the detector under additive white Gaussian noise (AWGN) conditions. This detector is a promising candidate for covariance-based sequential CR detectors and will be used as the reference for our own sequential likelihood ratio test (SLRT) detector.

In the context of SLRT detectors applied to CRs, Lai et al. consider in [21] change-point detection methods that aim to detect a transition in the underlying statistical distribution (specifically its energy) of the received samples as quickly as possible. They focus on the cumulative sum (CUSUM) algorithm, related to Wald's sequential probability ratio test (SPRT) [22], known to minimize the worst mean detection delay when all the parameters of the statistical distributions before and after the change are known. Additionally, they consider a rank-based likelihood ratio test as a robust detector when parameters such as the noise variance and signal power are unknown. The detector is however devised for single-antenna systems and suffers from a high implementation complexity. Another detector based on the CUSUM test has been derived in [23].

Traditional generalized-likelihood ratio (GLR) CUSUM algorithms [24] can be used when unknown parameters are present which can be estimated incrementally. These algorithms reuse past samples with the current estimates of the unknown parameters to minimize detection delays. This optimality requires to sacrifice the recursive nature of the CUSUM algorithm and also requires a large memory to store past samples. The detector considered in [25] combines the properties of a GLR test, namely sequential parameter estimation, with a parallel CUSUM algorithm, in order to overcome parameter uncertainties while keeping the iterative property of the CUSUM algorithm. Unlike most SLRT detectors, the unknown parameters are defined over a discrete space and the detector alternates between steps during which the parameter space is redefined and estimation steps during which candidates from the parameter space are discarded. This requires to run multiple (i.e. one for each parameter times the size of the parameter candidates set) CUSUM tests in parallel.

In [26], the authors derive an SLRT detector based on an auto-regressive moving-average (ARMA) signal model, applied to CR systems. While the SLRT detector proposed in our paper is based on a similar procedure, the underlying signal assumptions are different.

Regarding the general theory of sequential statistical tests, Lai [24] and Kailath et al. [27, p. 2252] both present extensive reviews of SLRT and other sequential detection schemes as well as the statistical tools needed to study them.

In this paper, we propose an SLRT similar to the one developed in [26], but do not assume an ARMA model for the received signal, nor that the noise variance is known. Furthermore, the detector exploits the structure of the sample covariance matrix of the received signal in order to achieve detection, as opposed to the signal energy level as used in [26] (and [21]). We also provide analytical results that offer more insights into the choice of detector parameters and its performance. Unlike [18], we do not need strong Gaussian assumptions regarding the distribution of the received signal. The proposed detector is a sequential variant of the batch detector proposed in [11] and retains its constant false-alarm rate (CFAR) property with respect to the noise variance. In comparison to [25], the SLRT detector studied in our paper possesses a simpler mode of operation. The proposed detector updates its parameter estimation at every step and does not require multiple CUSUM algorithms operating in parallel or the need to periodically redefine the parameter space. In summary, we propose a computationally simple and robust SLRT detector exploiting the correlation of the received signal and derive the necessary tools to parametrize the detector and study its performance in a typical CR spectrum sensing scenario. Numerical results show the accuracy of the statistical model derived to study the detector while a comparison with the sequential covariance-based detector of [20] reveals the competitiveness of the proposed SLRT detector and highlights its robustness to impulsive noise[2].

The remainder of this paper is organized as follows. In Section II we introduce the received signal model and some notations used in this paper. Section III presents the proposed detector statistic, whose properties are studied in detail in Section IV. Numerical results allowing a comparison with the theoretical results of Section IV are presented in Section V, along with a comparison of the

[2]We do not compare the proposed SLRT detector to [26] since the latter is not robust to noise variance uncertainties and would therefore be subjected to the SNR wall phenomenon [28].

proposed detector with the sequential detector developed in [20]. Section VI concludes the paper.

II. System Model

In this section, we present our detection scenario and introduce some notations that will be used throughout this paper.

A. Detection strategy

Let us first introduce the operating principle of an SLRT detector. Using a random variable, whose samples are received sequentially and whose distribution changes with hypothesis H_0 (signal absent) or H_1 (signal present), we compute the likelihood ratio, or log-likelihood ratio (LLR) statistic for this random variable, which is itself a random variable. Inherent to the properties of the LLR, the mean value decreases (increases) over time when hypothesis H_0 (H_1) is valid. As a result, the LLR statistic "drifts" over time, with a positive or negative drift depending on which hypothesis is true. One can therefore set lower and upper detection thresholds and infer which hypothesis is true based on the behavior of the LLR statistic. Once a decision is made, the statistic is reset to a point in between the detection thresholds and the cycle starts anew.

One then has to decide which random variable (i.e. metric) is suitable as a starting point for the SLRT detector. In blind detection scenarios, where the receiver possesses little information about the signal to be detected (typical for CR systems), second-order statistical properties of the received signal are by far the most used metrics to differentiate signal from noise (e.g. energy detector, covariance-based detectors, cyclic-feature detectors). While some statistics pertaining to covariance-based detection are often based on the likelihood ratio test (LRT) optimality criterion [29], [9], others are *ad hoc* statistics based on the assumption that the noise is either spatially uncorrelated or temporally white [30], [11]. While ad hoc statistics do not possess any optimality property, they may possess the advantage of simpler signal model, making the detector more robust to uncertainties or deviations from the model. The proposed SLRT detector uses the ad hoc statistic developed in [11], which only requires the noise to be spatially uncorrelated, as the time-variant random variable for which the LLR is derived.

B. Received signal

Since we are trying to develop a detector suitable for CR systems, we will focus on low SNR scenarios. Additionally, we will consider a multiantenna system suitable for a detection schemes based on the spatial covariance matrix of the received signal, similar to [20]. We note that the proposed detector would also be suitable for a single-antenna system working on the temporal covariance of the received signal.

We consider a multiantenna ($q \in \{1, \ldots, 10\}$ antennas) system receiving a baseband discrete-time multivariate signal $\boldsymbol{x}(n)$ (of size $1 \times q$ $\forall n$) and attempt to discriminate between the following hypotheses:

$$\begin{cases} H_0 : \boldsymbol{x}(n) = \boldsymbol{v}(n) \\ H_1 : \boldsymbol{x}(n) = \boldsymbol{s}(n) + \boldsymbol{v}(n) \end{cases} \tag{1}$$

where $\boldsymbol{s}(n)$ is the communication signal to be detected and $\boldsymbol{v}(n)$ represents an i.i.d. process of noise and interference (not necessarily Gaussian). All signals are assumed to be zero mean and if needed, we may subtract the sample mean from the received sample vectors. We regroup the received sample vectors in non-overlapping windows of N samples and denote by $\boldsymbol{X}_k$ the $(N \times q)$ matrix composed of the sample vectors received within the observation window k ($k = 0, 1, \ldots$). We will sometime use the terms k^{th} iteration or value at time k to refer to parameters or variables whose discrete-time index is k. We assume that the input signals are wide-sense stationary (WSS) over each observation window.

As previously stated, we plan on exploiting the structure of the covariance matrix of the received signal ($\boldsymbol{\Sigma}_x$) to differentiate the signal from the noise. In this paper we will use the spatial covariance matrix of the signal, but the theory will apply to any covariance matrix[3] (temporal, spatial or a mix of both), as long as it remains diagonal under H_0. Moreover, we do not suppose that the receiver is calibrated, allowing for different noise variance at the antennas ($\boldsymbol{\Sigma}_v$ is diagonal but not necessarily a scaled identity matrix). Upon collection of N sample vectors, the sample covariance matrix can be computed as

$$\hat{\boldsymbol{\Sigma}}_k = \frac{1}{N} \boldsymbol{X}_k^H \boldsymbol{X}_k, \tag{2}$$

where $(.)^H$ denotes the Hermitian operator.

In the same fashion as the procedure developed by Tugnait [11] for his covariance-based detector, we are going to normalize the entries of $\hat{\boldsymbol{\Sigma}}_k$ in order to remove dependencies on the noise variance. We first form the matrix $\boldsymbol{\Gamma} = diag(\hat{\boldsymbol{\Sigma}}_x)$, composed of the main diagonal entries of $\hat{\boldsymbol{\Sigma}}_k$, and use it to create the matrix

$$\tilde{\boldsymbol{\Sigma}}_k = \sqrt{N} \boldsymbol{\Gamma}^{-1/2} \hat{\boldsymbol{\Sigma}}_k (\boldsymbol{\Gamma}^{-1/2})^H, \tag{3}$$

[3] If the covariance matrix has a Toeplitz structure the detector would work with the normalized auto-correlation vector of the received signal.

whose off-diagonal entries have a unit variance. We then regroup the lower-diagonal entries of $\tilde{\boldsymbol{\Sigma}}_k$ in the vector $\boldsymbol{r}_k = [\tilde{\boldsymbol{\Sigma}}_k(2,1), \tilde{\boldsymbol{\Sigma}}_k(3,1), \ldots, \tilde{\boldsymbol{\Sigma}}_k(q,q-1)]$. The vector $\boldsymbol{r}_k$ is of size $(p \times 1)$, where $p = q(q-1)/2$. Owing to the central-limit theorem (CLT) for large enough values of N, the vector $\boldsymbol{r}_k$ is approximately distributed as

$$\begin{aligned} H_0 &: \boldsymbol{r}_k \sim \mathcal{N}_p(\boldsymbol{0}, \boldsymbol{\mathcal{I}}_p) \\ H_1 &: \boldsymbol{r}_k \sim \mathcal{N}_p(\boldsymbol{\mu_r}, \boldsymbol{\mathcal{I}}_p), \end{aligned} \tag{4}$$

where $\boldsymbol{\mathcal{I}}_p$ is an identity matrix of size p and $\boldsymbol{\mu_r}$ is the mean value vector of $\boldsymbol{r}$ [11]. Note that under hypothesis H_1, the elements of the vector $\boldsymbol{r}_k$ are in fact correlated. However, as pointed out in [11], the correlation vanishes when the SNR, defined as the ratio of the signal power σ_s^2 over the noise power σ_n^2, tends toward zero. We are mostly concerned with low SNR spectrum sensing scenarios and will therefore neglect this correlation.

Looking at equation (4), it is apparent that we can use statistical inference on the mean value of $\boldsymbol{r}_k$ to detect the presence of the signal. Additionally, we notice that the distribution of $\boldsymbol{r}_k$ does not depend on the noise variance, which will allow our detector to be CFAR with respect to this nuisance parameter (a very desirable property for CR detectors).

Since we will often need to refer to mean value or variance of variables, while specifying the variable considered, the hypothesis and the time index, we will adopt the notation ${}_i(.)_{X_k}$, where $i \in \{0,1\}$ is the hypothesis (H_0 or H_1), X is the variable considered and $k \in \mathbb{N}_0$ is the time index. When there is no assumption regarding the hypothesis or when the context makes it clear which hypothesis is considered, we will omit the left subscript.

III. Sequential Likelihood Ratio Test (SLRT)

We will now develop the SLRT statistic based on $\boldsymbol{r}_k$ and then discuss pertinent properties of the LLR update of its statistic.

A. SLRT statistic

We denote by $\zeta_{\boldsymbol{r}_k}(H_0)$ and $\zeta_{\boldsymbol{r}_k}(H_1)$ the marginal likelihoods of the vector $\boldsymbol{r}_k$ under hypothesis H_0 and H_1, respectively. Under hypothesis H_0, we know that $\boldsymbol{r}_k \sim \mathcal{N}_p(0, \boldsymbol{\mathcal{I}}_p)$, therefore $\zeta_{\boldsymbol{r}_k}(H_0) = \phi_p(\boldsymbol{0}, \boldsymbol{\mathcal{I}}_p)$, where $\phi_p(.,.)$ denotes the p-variate normal density function. On the other hand, $\zeta_{\boldsymbol{r}_k}(H_1)$ depends on the unknown parameter $\boldsymbol{\mu_r}$ and to address this issue, we now turn to a Bayesian framework, such as the one used in [18]. We will assign a prior distribution to the parameter $\boldsymbol{\mu_r}$ and, upon reception of a vector $\boldsymbol{r}_k$, compute its posterior distribution. The posterior distribution will then become the prior distribution for the next iteration.

Since the distribution of $\boldsymbol{r}_k$ is approximately Gaussian with a known variance, we can assign a conjugate prior distribution [31] to the mean parameter: $\boldsymbol{\mu_r} \sim \mathcal{N}_p(\boldsymbol{\mu}_k, \sigma_k^2 \boldsymbol{\mathcal{I}}_p)$, where the variables $\boldsymbol{\mu}_k$ and σ_k^2 are called the hyperparameters. Upon reception of the vector $\boldsymbol{r}_{k+1}$, the hyperparameters are updated as ([31, p. 318])

$$\begin{aligned} \boldsymbol{\mu}_{k+1} &= \frac{\boldsymbol{\mu}_k}{1+\sigma_k^2} + \frac{\sigma_k^2}{1+\sigma_k^2}\boldsymbol{r}_{k+1} \\ \sigma_{k+1}^2 &= \frac{\sigma_k^2}{\sigma_k^2+1}. \end{aligned} \tag{5}$$

Using such a conjugate prior allows us to easily compute $\zeta_{\boldsymbol{r}_k}(H_1)$ as

$$\zeta_{\boldsymbol{r}_k}(H_1) = \phi_p(\boldsymbol{\mu}_{k-1}, (1+\sigma_{k-1}^2)\boldsymbol{\mathcal{I}}_p). \tag{6}$$

Let us define the binary random variable $u_k \in \{0,1\}$ indicating the presence ($u_k = 1$) or absence ($u_k = 0$) of a communication signal at time k. Given prior probabilities on u_k, the posterior probability $\Pr(u_k = 1) \triangleq {}_1P_{u_k}$ is given by

$${}_1P_{u_k} = \frac{\zeta_{\boldsymbol{r}_k}(H_1)\,{}_1P_{u_{k-1}}}{\zeta_{\boldsymbol{r}_k}(H_1)\,{}_1P_{u_{k-1}} + \zeta_{\boldsymbol{r}_k}(H_0)\,{}_0P_{u_{k-1}}}. \tag{7}$$

Rearranging the terms of equation (7), we obtain

$$2\ln(\frac{{}_1P_{u_{k+1}}}{{}_0P_{u_{k+1}}}) = \mathcal{L}_k + 2\ln(\frac{{}_1P_{u_k}}{{}_0P_{u_k}}), \tag{8}$$

where $\ln(.)$ is the natural logarithm and $\mathcal{L}_k$ is the LLR of $\boldsymbol{r}_k$ explicitly given by

$$\mathcal{L}_k = \|\boldsymbol{r}_k\|^2 - \frac{\|\boldsymbol{r}_k - \boldsymbol{\mu}_{k-1}\|^2}{1+\sigma_{k-1}^2} - p\ln(1+\sigma_{k-1}^2). \tag{9}$$

Defining the new variable $Z_k \triangleq 2\ln(\frac{{}_1P_{u_k}}{{}_0P_{u_k}})$, we finally obtain the SLRT statistic

$$Z_{k+1} = Z_k + \mathcal{L}_k. \tag{10}$$

The proposed SLRT detector based on (10) works as follows. Starting from an initial value Z_0, the variable Z_k is continuously updated. Under hypothesis H_0, the mean value of the log-likelihood ratio $\mu_{\mathcal{L}_k}$ is negative and Z_k tends toward $-\infty$. Whenever it reaches a lower threshold A, we infer that hypothesis H_0 is valid and the value Z_k is reset to a value a ($a \geq A$ and typically $Z_0 = a$). Under hypothesis H_1, $\mu_{\mathcal{L}_k} \geq 0$ and the drift of the variable Z_k is positive. When it reaches an upper threshold B, we infer that a signal has been detected and its value is set to b ($b \leq B$). This is illustrated in Figure 1, which shows the evolution of the statistic Z_k over time. It depicts the statistic in the H_0 stationary regime,

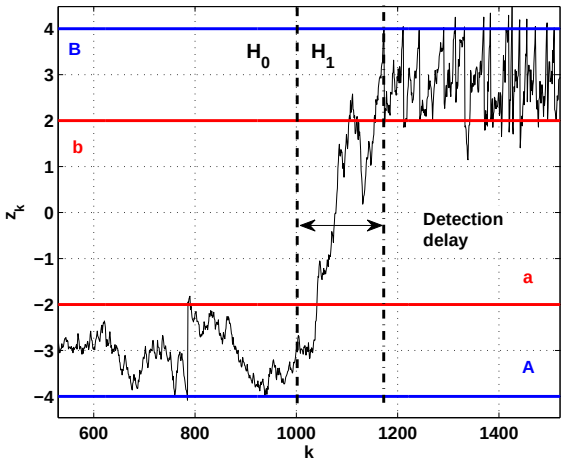

Fig. 1. Illustration of the SLRT statistic Z_k over time.

the $H_0 \to H_1$ transition and the H_1 stationary regime.

In the limit $a \to A$, and $b \to B$ the boundaries A and B become reflecting screens as termed in [32]. This limit is particularly important since it satisfies some optimality criterion for the detector. However, it also corresponds to a case of "degenerate analysis" [32] since several uncertainties of the type $\frac{0}{0}$ appear in equations describing the statistical properties of Z_k. Therefore, we will develop the generic model, which includes parameters a and b, and take the limit $a = A$ and $b = B$ when needed.

Later on, we will need a definition of the signal "strength", that appropriately reflects the influence of the signal on the statistic Z_k. We will use the following metric:

$$SNR_L = \frac{\|\boldsymbol{\mu_r}\|^2}{{}_0\sigma_{\boldsymbol{r}}^2} = \sum_{i=1}^{p} \boldsymbol{\mu_r}^2(i), \tag{11}$$

where $\boldsymbol{\mu_r}(i)$ is the ith component of the vector $\boldsymbol{\mu_r}$.

B. LLR

Before we move on to the statistical properties of the variable Z_k, we first discuss the properties of the LLR in (9).

Starting from deterministic initial values $\boldsymbol{\mu}_0$ and σ_0^2, the hyperparameters (5) will converge over time (i.e. increasing k) toward $\boldsymbol{\mu_r}$ and 0, respectively. If we suppose that the received signal is at first only composed of noise samples (hypothesis H_0), it implies that:

- The likelihoods under hypotheses H_0 and H_1 become identical: $\lim_{k\to\infty} \zeta_{\boldsymbol{r}_k}(H_1) = \zeta_{\boldsymbol{r}_k}(H_0)$, forcing $\mu_{\mathcal{L}_k}$ toward 0.
- The SLRT detector will take an increasingly longer time (eventually infinite) to react to a change in the distribution of $\boldsymbol{r}_k$ (for instance if a signal suddenly appears) since the influence of new samples $\boldsymbol{r}_k$ on the value of $\boldsymbol{\mu}_k$ becomes vanishingly small as σ_k^2 tends toward 0 (see (5)).

These issues come from the infinite memory of the system through the hyperparameters (5) and can be avoided using a "back-to-the-prior" strategy [33], [18] on the hyperparameter σ_k^2 after each iteration (5) via

$$\sigma_{k+1}^2 =: (1-\kappa)\sigma_{k+1}^2 + \kappa\sigma_0^2, \tag{12}$$

where $\kappa \in (0,1)$. This strategy only needs to be applied when we believe that hypothesis H_0 is true. Under hypothesis H_1, $\zeta_{\boldsymbol{r}_k}(H_1) > \zeta_{\boldsymbol{r}_k}(H_0)$ and therefore $\mu_{\mathcal{L}_k} > 0$ (i.e. there is no "stagnation" of the LLR near 0).

The choice of κ and σ_0^2 influences the mean value and variance of the LLR. In order to choose appropriate values for these two parameters, let us now focus on the asymptotic regime under H_0 (when a large number of samples belonging to H_0 have been received) and the transition $H_0 \to H_1$.

In the asymptotic H_0 regime, σ_k^2 tends toward σ_∞^2 which satisfies the equation

$$\sigma_\infty^2 = (1-\kappa)\frac{\sigma_\infty^2}{1+\sigma_\infty^2} + \kappa\sigma_0^2, \tag{13}$$

whose solution is given by

$$\sigma_\infty^2 = \frac{1}{2}(\kappa(\sigma_0^2 - 1) + \sqrt{\kappa^2(1-\sigma_0^2)^2 + 4\kappa\sigma_0^2}). \tag{14}$$

Additionally, the hyperparameter $\boldsymbol{\mu}_k$ tends toward a random variable $\boldsymbol{\mu}_\infty \sim \mathcal{N}_p(\mathbf{0}, \frac{\sigma_\infty^2}{2+\sigma_\infty^2}\boldsymbol{\mathcal{I}}_p)$.

If σ_∞^2 is different from zero, the statistic Z_k has a negative drift under H_0, forcing it to periodically reach the lower threshold A. This is beneficial to avoid false-alarms by driving Z_k away from the upper threshold. Moreover, the variance of the estimator $\boldsymbol{\mu}_k$ is different from zero, allowing the detector to react to a change of hypothesis. On the other hand, when a signal appears, it can only be detected if $\mu_{\mathcal{L}_k} > 0$. This indicates that if the limit σ_∞^2 is too large (for a given value of the SNR_L), the LLR will maintain a negative drift under H_1 and the SLRT will not be able to detect the signal[4]. Hence, we need to find the upper bound on σ_∞^2.

Starting from the asymptotic H_0 regime and receiving samples belonging to the hypothesis H_1, the upper bound

[4]This is an approximation since the probability that Z_k reaches the upper threshold is not equal to zero. However, detection events become comparable to false-alarms (long delays and short duration) indicating that the detector effectively fails to reliably detect the signal.

on σ_∞^2 is the value for which the drift is asymptotically equal to zero ($_1\mu_{\mathcal{L}_\infty} = 0$). Using equation (9) with the asymptotic values of the variables gives us

$$_1\mu_{\mathcal{L}_\infty} = \sum_{i=1}^{p}(1 + \boldsymbol{\mu}_{\boldsymbol{r}}^2(i)) - \frac{2p}{(2 + {}_0\sigma_\infty^2)} - p\ln(1 + {}_0\sigma_\infty^2). \tag{15}$$

Notice that the hyperparameter σ_k^2 takes on its H_0 limit $_0\sigma_\infty^2$, despite the fact that H_1 is the valid hypothesis, since we consider that the detector never manages to detect the signal.

The upper bound $\bar{\sigma}_\infty^2$ is the root of the equation

$$\begin{aligned} 0 &= p(1 - \frac{2}{2 + \bar{\sigma}_\infty^2} - \ln(1 + \bar{\sigma}_\infty^2)) + SNR_L \\ &\simeq p(1 - \frac{2}{2 + \bar{\sigma}_\infty^2} - \bar{\sigma}_\infty^2) + SNR_L. \end{aligned} \tag{16}$$

where the first-order Taylor series approximation $\ln(1 + \bar{\sigma}_\infty^2) \simeq \bar{\sigma}_\infty^2$ is valid provided that $\bar{\sigma}_\infty^2 \ll 1$. While equation (16) can be solved, we can reduce it further to a linear equation at the cost of one additional approximation ($\frac{1+\sigma_\infty^2}{2+\sigma_\infty^2} \simeq \frac{1}{2}$) which results in

$$SNR_L - \frac{p}{2}\bar{\sigma}_\infty^2 + \simeq 0, \tag{17}$$

whose root is simply $\bar{\sigma}_\infty^2 \simeq \frac{2SNR_L}{p}$.

Equations (14) and (17) can be used to choose the parameters κ and σ_0^2 as a function of the SNR_L (i.e. assumed minimum value of the SNR_L for the signal to be detected) via the variable σ_∞^2.

Since the limits $\sigma_\infty^2 \to 0$ and $\sigma_\infty^2 \to \bar{\sigma}_\infty^2$ result in a detection time that tends toward infinity, it is reasonable to expect that there exists a value of $\sigma_\infty^2 \in (0, \bar{\sigma}_\infty^2)$ that minimizes the detection delay. This will be discussed further in Section IV-D.

In order to clarify the procedure used by the SLRT detector, Algorithm 1 summarizes the steps involved in the algorithm and the main equations involved.

Algorithm 1 SLRT

Require: SNR_L, A, B, p, N
- $\bar{\sigma}_\infty^2 \Leftarrow SNR_L$ ▷ (17)
- $\sigma_\infty^2 \Leftarrow \nu\bar{\sigma}_\infty^2 \quad \nu \in [0, 1]$
- $\sigma_0^2, \kappa \Leftarrow \sigma_\infty^2$ ▷ (14)
- A, B
- $Z_0 \Leftarrow A$, $H_0 \Leftarrow$ true, $k \Leftarrow 0$
- **loop**
  - $\boldsymbol{r}_k \Leftarrow \boldsymbol{X}_k$ ▷ (2), (3)
  - $\mathcal{L}_k \Leftarrow \boldsymbol{r}_k$ ▷ (9)
  - $Z_{k+1} \Leftarrow Z_k + \mathcal{L}_k$
  - $\{\boldsymbol{\mu}_{k+1}, \sigma_{k+1}^2\} \Leftarrow \{\boldsymbol{\mu}_k, \sigma_k^2\}$ ▷ (5)
  - **if** $Z_{k+1} \leq A$ **then**
    - $Z_{k+1} \Leftarrow A$
    - $H_0 \Leftarrow$ true
    - $\sigma_{k+1}^2 \Leftarrow (1 - \kappa)\sigma_{k+1}^2 + \kappa\sigma_0^2$ ▷ (12)
  - **else if** $Z_{k+1} \geq B$ **then**
    - $Z_{k+1} \Leftarrow B$
    - $H_1 \Leftarrow$ true
  - **end if**
  - $k \Leftarrow k + 1$
- **end loop**

IV. Statistical properties of the SLRT

This section focuses on the statistical properties of the SLRT statistic Z_k. Specifically, we will introduce a statistical model for Z_k based on diffusion processes, which will then allow us to set the detector thresholds A and B, and find analytical expressions for the SLRT detection delays as a function of the signal parameters.

In batch detection scenarios, the detector accumulates a fixed amount of samples before computing its statistic. Statistical properties of the detector can therefore be determined as a function of the number of samples received during the observation window. For instance, the probability of false-alarm (P_{fa}) is typically defined as a the probability that the detector statistic exceeds a threshold when computed over a batch of samples, under hypothesis H_0.

On the other hand, sequential detectors continuously update their statistic instead of working with a batch of samples and a signal may appear or disappear at any time. A sequential detector can attain any probability of detection for a fixed probability of false-alarm by using as many samples as needed. As a result, change-point detection scenarios, where the statistical distribution of the received samples changes at some random time (e.g. $H_0 \to H_1$ or vice versa), use different criteria to assess the performance of the detector [17]. Let us call λ_{fa} the process which records over time the delay between false-alarms (i.e. the delay until Z_k reaches boundary B again under hypothesis H_0). It is of interest to compute the mean delay between false-alarms $E[\lambda_{fa}]$. Similarly, we define λ_d as the process which records detection delays over time (i.e. delay before the detector registers a change in the signal underlying distribution after a change-point occurred). The main performance criterion of the SLRT detector will be the mean detection delay ($E[\lambda_d]$). Obviously, it is of interest to minimize $E[\lambda_d]$ while maximizing $E[\lambda_{fa}]$.

In order to compute $E[\lambda_{fa}]$, one needs to know the statistical characteristics of Z_k in the stationary (asymptotic) H_0 regime. On the other hand, $E[\lambda_d]$ depends on the transient distribution of Z_k after a change-point occurred. To find these distributions using equation (10) and the distribution of the LLR proves intractable due to the non-linear behavior of Z_k once it reaches the thresholds A or B. To circumvent these difficulties, we will approximate the SLRT statistic by a continuous-time diffusion process.

A. Diffusion process model

A diffusion model is a continuous-time stochastic process which has the Markov property. A discrete-time random walk such as Z_k can be approximated by a diffusion process Z_t [34] according to Donsker's theorem [35], [36], allowing the use of all the statistical tools available to characterize theses processes. Unfortunately, this approximation requires the successive increments of the stochastic process to be i.i.d.. Under the stationary H_0 regime, the log-likelihood ratios increments $\mathcal{L}_k$ are identically distributed but not independent, as can be seen from equations (5) and (9). The correlation coefficient $\rho_{\mathcal{L}}(\tau) = \frac{E[\mathcal{L}_k \mathcal{L}_{k-\tau}]}{\sigma^2_{\mathcal{L}_\infty}}$ for LLR that are τ samples apart is a function of σ_0^2 and in the limit ($\sigma_0^2 \to 0$) tends toward zero. For typical values of $\sigma_0^2 \in [10^{-4}, 10^{-3}]$, the correlation coefficient is small ($\rho_{\mathcal{L}}(\tau) \sim 10^{-4}$) and we will neglect it. We later show the impact of this correlation on the accuracy of the analytical model derived for Z_k.

Additionally, the continuous-time approximation does not take into account the overshoot Δ_t that occurs when the process reaches a boundary. However, Lorden's inequality [37, p. 160] $E[\Delta_t] \leq \frac{E[\mathcal{L}^2_\infty]}{\mu_{\mathcal{L}_\infty}}$ gives us a bound on the overshoot mean value.

The diffusion process model that we will use to characterize the statistic Z_k in the stationary H_0 regime is defined by the differential equation

$$dZ_t = \mu_Z dt + \sigma_Z^2 dB_t, \tag{18}$$

where B_t is a Wiener process (Brownian motion). The parameters μ_Z and σ_Z^2 are respectively the drift and diffusion coefficients of the process Z_t (the subscript t indicates that we are now working with a continuous-time process). Note that since we approximate the random walk (10) by a diffusion process, we have $\mu_Z = \mu_{\mathcal{L}_k}$ and $\sigma_Z^2 = \sigma^2_{\mathcal{L}_k}$.

Additionally, we define the parameter $\varphi = \frac{2\mu_Z}{\sigma_Z^2}$, which will be used extensively while developing the statistical properties of the process Z_t. In the stationary H_0 regime we have

$$\varphi = \frac{1 - 1/\omega(1 + \frac{\sigma^2_\infty}{2+\sigma^2_\infty}) - \ln(\omega)}{1 - 2/\omega + 1/\omega^2(1 + \frac{\sigma^2_\infty}{2+\sigma^2_\infty})^2}, \tag{19}$$

where we introduced the parameter $\omega = 1 + \sigma^2_\infty$ to simplify notations. Provided that $\sigma^2_\infty \ll 1$, φ is closely approximated by $-\frac{1}{2}$ in the H_0 regime, but we will develop the theory using the general φ parameter.

In his paper [16], Shiryaev considers a multistage cyclic-return Wald system using a very similar diffusion model, but for which the parameter $\varphi = 1$, leading to some simplifications. We refer the reader to the excellent book by Karlin and Taylor [34] as well as the aforementioned work of Shiryaev for most of the theory that we will now apply to identify the statistical properties of Z_t.

B. Mean false-alarm delay

In order to find the mean delay between false-alarms, we need to compute the probability that the process reaches the boundary B under H_0, as well as the average time that it takes to reach either boundaries (A or B) from a starting point x.

We define as τ_1 the first time the process reaches a boundary A or B from a starting point x at $t = 0$. Let $\alpha(x) = \Pr(Z_{\tau_1} = B | Z_0 = x, H_0)$ be the probability that the process first exits via boundary B under hypothesis H_0. It can be shown that the probability $\alpha(x)$ is a solution of the Backward Kolmogorov equation [34]

$$\frac{1}{2}\sigma_Z^2 \alpha^{''}(x) + \mu_Z \alpha^{'}(x) = 0, \tag{20}$$

with boundary conditions $\alpha(B) = 1$ and $\alpha(A) = 0$. The solution is given by

$$\alpha(x) = \frac{e^{\varphi(A+B-x)} - e^{\varphi B}}{e^{\varphi A} - e^{\varphi B}}. \tag{21}$$

Similarly, we define

$$\beta(x) = \Pr(Z_{\tau_1} = A | Z_0 = x, H_0) = \frac{e^{\varphi A} - e^{\varphi(A+B-x)}}{e^{\varphi A} - e^{\varphi B}}. \tag{22}$$

We denote by $M_0(x) = E[\tau_1 | Z_0 = x, H_0]$ the average time taken by the process to reach either boundary A or B given a starting point $A \leq x \leq B$. It is a solution of the following differential equation [34, p. 193]

$$\frac{1}{2}\sigma_Z^2 M_0^{''}(x) + \mu_Z M_0^{'}(x) = -1, \tag{23}$$

given by

$$M_0(x) = \frac{1}{\mu_Z}[(B - A)\alpha(x) + (A - x)] \tag{24}$$

The mean false-alarm delay $E[\lambda_{fa}] \triangleq T_0(a)$ (i.e. the average time for the process to reach boundary B given that the process starts at a under H_0) is then given by the law of total probabilities:

$$\begin{aligned} T_0(a) &= \alpha(a)M_0(a) + (1-\alpha(a))[M_0(a) + T_0(a)] \\ &= \frac{M_0(a)}{\alpha(a)}. \end{aligned} \tag{25}$$

In the limit $a \to A$ and $b \to B$, the mean false-alarm delay becomes equal to

$$T_0(A) = \frac{1}{\mu_Z}\left[(B-A) + \frac{1}{\varphi}(e^{\varphi(A-B)} - 1)\right]. \tag{26}$$

From a practical point of view, it is more interesting to obtain A and B for a desired $T_0(A)$. Assuming without loss of generality that $A = -B$ and using the approximation $\varphi \simeq -1/2$, this requires us to solve the exponential-linear equation

$$T_0(A) = \frac{1}{\mu_Z}\left[2B - 2(e^B - 1)\right]. \tag{27}$$

Using the method derived in [38], we first rewrite (27) in the form $a^x = b(x+c)$:

$$e^B = B + (1 - \frac{T_0(A)\mu_Z}{2}), \tag{28}$$

whose solution is given by

$$B = -W_i(-e^{-c}) - c, \tag{29}$$

where $W_i(x)$ is the i^{th} branch of the Lambert W-function. The branch to be used depends on the value of c.

C. Stationary distribution

We will now derive the stationary distribution $\Psi_Z(x)$ of the process Z_t under H_0, which will be needed to compute the SLRT mean detection delay $E[\lambda_d]$. The method closely follows the theory presented in [34] to compute the distribution of instantaneous return processes.

In order to obtain the distribution, we will need the Green function $G(x_0, \xi)$ associated with the process. It corresponds to the average time the process spends in an infinitesimal interval $\{\xi, \xi + d\xi\}$ ($\xi \in \{A, B\}$), before exiting via either A or B. x_0 denotes the starting point at $t = 0$. The Green function for the process Z_t is given by

$$\begin{aligned} G(x,\xi) &= \begin{cases} \dfrac{2\varphi(e^{\varphi(A+B-x)} - e^{\varphi B})(e^{\varphi(\xi-B)} - 1)}{(e^{\varphi A} - e^{\varphi B})\sigma_Z^2} \\ \dfrac{2\varphi(e^{\varphi A} - e^{\varphi(A+B-x)})(1 - e^{\varphi(\xi-A)})}{(e^{\varphi A} - e^{\varphi B})\sigma_Z^2} \end{cases} \\ &= \begin{cases} G^+(x,\xi) & A \le x \le \xi \le B \\ G^-(x,\xi) & A \le \xi \le x \le B \end{cases} \end{aligned} \tag{30}$$

The process Z_t is a form of instantaneous return process whereby the process restarts at a point a or b whenever it reaches the boundary A or B (this defines a cycle). Therefore, in its stationary regime, the process is composed of multiple i.i.d. cycles of random duration. It has been proven [34, p. 261] that the stationary distribution of an instantaneous return process that starts at a point x_0 takes on the form

$$\Psi(x_0, x) = \frac{G(x_0, x)}{\int_A^B G(x_0, \xi)d\xi}. \tag{31}$$

In our case, the starting points are either a or b depending on which boundary is reached on the previous cycle and therefore there is a probability associated with them. If we call $\Pi(a)$ the probability that the process restarts at a, we have the following equation:

$$\begin{aligned} \Pi(a) &= \Pi(a)\beta(a) + (1 - \Pi(a))\beta(b) \\ &= \frac{\beta(b)}{1 + \beta(b) - \beta(a)}. \end{aligned} \tag{32}$$

In the same fashion,

$$\Pi(b) = \frac{\alpha(a)}{1 + \alpha(a) - \alpha(b)}. \tag{33}$$

Using equations (31), (32) and (33), we derive the stationary distribution

$$\begin{aligned} \Psi_Z(x) &= \frac{\Pi(a)G(a,x) + \Pi(b)G(b,x)}{\int_A^B \Pi(a)G(a,\xi) + \Pi(b)G(b,\xi)d\xi} \\ &= \frac{\beta(b)G(a,x) + \alpha(a)G(b,x)}{\int_A^B \beta(b)G(a,\xi) + \alpha(a)G(b,\xi)d\xi}. \end{aligned} \tag{34}$$

In the limit $a \to A$ and $b \to B$, the distribution becomes

$$\Psi_Z(x) = \frac{\varphi e^{\varphi x}}{e^{\varphi B} - e^{\varphi A}}. \tag{35}$$

Its mean value is equal to

$$\mu_\Psi = \frac{1}{e^{\varphi B} - e^{\varphi A}}\left[e^{\varphi B}(B - 1/\varphi) - e^{\varphi A}(A - 1/\varphi)\right]. \tag{36}$$

D. Mean detection delay

In his work [32], Shiryaev considers an instantaneous return (to zero) diffusion process with fixed (and known) drift parameters under hypotheses H_0 and H_1. This allows him to derive the mean detection delay $E[\lambda_d] \triangleq R_Z$

$$R_Z = \int_A^B T_1(x)\Psi_Z(x)dx, \tag{37}$$

where $T_1(x)$ is the mean delay to reach boundary B, starting from x, under hypothesis H_1. Moreover, Shiryaev provides a proof that, for large mean false-alarm delay $T_0(0)$, the minimum detection delay occurs when $A = 0$ (i.e. the lower boundary coincides with the starting point) and this detection delay is proportional to $\ln(T_0(0))$. He also conjectures that the condition $A = 0$ is likely true for all values of $T_0(0)$ (not just large ones) and these results are further analyzed in [39]. Intuitively, restarting the process on the boundary itself is appealing. When a change of distribution occurs, the trajectory of the process becomes longer if the process is allowed to venture below its starting point, leading to a longer detection delay. This is also similar to the CUSUM detection rule, which only considers increments of the process in the direction of the upper boundary. We have not been able to provide a mathematical proof showing that the detection delay reaches a minimum under the limit $a, b \to A, B$ but numerical results indicate that Shiryaev's conjecture holds true.

We will now derive an approximation for R_Z. The process Z_t is slightly different than the one considered by Shiryaev, since its drift and diffusion (variance) during the transition $H_0 \to H_1$ are time-dependent. This prevents us from deriving the function $T_1(x)$ using the same approach as for equation (25). Instead we will use a general form of Wald's identity, stating that the expected value of a sum of M random variables (M being also a random variable) is equal to:

$$E[\sum_{k=1}^{M} X_k] = E[\sum_{k=1}^{M} \mu_{X_k}] \tag{38}$$

If M is the stopping time corresponding to the process reaching the boundary B starting from a point x, then we can write

$$T_1(x) = \arg\min_M \{|\sum_{k=1}^{M} \mu_{\mathcal{L}_k} - (B - x)|\}. \tag{39}$$

Note that due to the existing correlation between successive $\mathcal{L}_k$, equation (39) is only an approximation[5].

Using equations (9), (11) and considering that at $k = 0$ the process is in its stationary H_0 regime, we obtain the expression for $\mu_{\mathcal{L}_k}$ in the transient $H_0 \to H_1$ regime as

$$\mu_{\mathcal{L}_k} = SNR_L \left(\frac{\omega^{2k-1} - 1}{\omega^{2k-1}}\right) + p\left(\frac{\sigma_\infty^2}{2 + \sigma_\infty^2} - \ln(\omega)\right), \tag{40}$$

where ω has been defined after equation (19). Using equation (17) as an approximation for $\bar{\sigma}_\infty^2$ and choosing $\sigma_\infty^2 = \nu\bar{\sigma}_\infty^2,\ \nu \in [0, 1]$, we can simplify equation (40) to

$$\mu_{\mathcal{L}_k} \simeq SNR_L \left((1 - \nu) - \frac{1}{\omega^{2k-1}}\right). \tag{41}$$

The cumulative sum of $\mu_{\mathcal{L}_k}$ is given by

$$\chi_k \simeq SNR_L \left[k(1 - \nu) - \left(\frac{1}{\omega^{2k-1}}\right)\left(\frac{\omega^{2k} - 1}{\omega^2 - 1}\right)\right], \tag{42}$$

which is a convex function that exhibits a linear asymptotic ($k \to \infty$) behavior. Using (39), we can see that $T_1(x)$ corresponds to the index k that minimizes

$$T_1(x) = \arg\min_k \{|\chi_k - (B - x)|\}. \tag{43}$$

In order to alleviate the need to solve (43) for all $x \in [A, B]$, we will use a first-order Taylor series approximation for $T_1(x)$. Replacing $T_1(x)$ in (37) by its Taylor series around an arbitrary point x_0, we obtain the series

$$R_Z = \sum_{i=1}^{\infty} \frac{1}{i!} \left.\frac{d^i T_1(x)}{dx^i}\right|_{x_0} \int_A^B (x - x_0)^i \Psi_Z(x)dx, \tag{44}$$

Taking the first-order series around $x_0 = \mu_\Psi$, we obtain

$$R_Z = T_1(\mu_\Psi) + \mathcal{O}(\sigma_\Psi^2), \tag{45}$$

indicating that the mean detection delay is approximately given by

$$R_Z = \arg\min_k \{|\chi_k - (B - \mu_\Psi)|\}. \tag{46}$$

Developing the terms of (46) using (42) and rearranging the terms in the same fashion as (27) (see also [38]), we

[5]It is worth mentioning that an analytical approximation for the probability that a process with time-dependent drift and diffusion remains below a given threshold over an interval $[0, T]$ is given in [40]. This approximation can be used to derive $T_1(x)$, but numerical results suggest that the approximation is not as good as (39).

obtain

$$\underbrace{(\omega^{-2})}_{a}^{R_Z} \simeq \underbrace{\frac{\omega^2-1}{\omega}}_{b}(\nu-1) \times \left[R_Z + \underbrace{\frac{B-\mu_\Psi}{SNR_L(1-\nu)} + \frac{\omega}{(\omega^2-1)(\nu-1)}}_{c}\right] \tag{47}$$

whose solution is

$$R_Z \simeq \frac{-1}{\ln(a)} W_i\left(\frac{-\ln(a)}{ba^c}\right) - c. \tag{48}$$

The parameter ν influences R_Z. As we previously mentioned, since the values 0 and 1 approximately correspond to vertical asymptotes for the function $R_Z(\nu)$ (i.e. the mean delay tends toward infinity), it is logical to suspect that there is a value ν_{opt} that minimizes $R_Z(\nu)$. Unfortunately, finding it analytically would require to find the root of the derivative of (48), which also includes terms such as (29). This is mathematically very complicated (possibly intractable) and therefore of little practical value. We will therefore only show numerical results pertaining to ν_{opt}.

E. Influence of the sample size N on the mean detection delay

The sample size N used to compute sequential estimations of the covariance matrix (3) has so far been left as an arbitrary value. We will now investigate its effect on R_Z.

It is obvious that under H_1, a larger N increases the value of $\boldsymbol{\mu_r}$. This in turn increases the drift of Z_t toward the upper boundary B and one would expect that the SLRT will then require fewer samples $\boldsymbol{r}_k$ (i.e. fewer steps) to detect the presence of a signal. However, a larger N also increases the delay that occurs between successive samples $\boldsymbol{r}_k$ and it is not obvious if the overall effect is a reduction in the detection delay.

In the following we will again consider that $\sigma^2_\infty \ll 1$ and therefore $\varphi = -1/2, \forall N$.

Let us consider two sample sizes N and $\tilde{N}$ and call their ratio ($\frac{\tilde{N}}{N} = \gamma$). To provide a fair comparison of the mean detection delay, we first need to make sure the overall mean false-alarm delay remains constant ($NT_0(A) = \tilde{N}\tilde{T}_0(A)$). Therefore, $\tilde{T}_0(A) = 1/\gamma T_0(A)$. Moreover, $S\tilde{N}R_L = \sum_{i=1}^{p} \tilde{\boldsymbol{\mu}}^2_{\tilde{\boldsymbol{r}}}(i) = \sum_{i=1}^{p} (\sqrt{\gamma}\boldsymbol{\mu_r}(i))^2 = \gamma SNR_L$. Using equations (17), (40) and once again choosing $\sigma^2_\infty = \nu\bar{\sigma}^2_\infty$, we see that ${}_0\tilde{\mu}_{\mathcal{L}_\infty} = \nu S\tilde{N}R_L = \gamma\,{}_0\mu_{\mathcal{L}_\infty}$. This leads to the identity

$$\tilde{T}_0(A)\,{}_0\tilde{\mu}_{\mathcal{L}_\infty} = T_0(A)\,{}_0\mu_{\mathcal{L}_\infty}. \tag{49}$$

Using equation (26) we deduce that the boundaries A and B remain constant (i.e. $\tilde{A} = A$). Rewriting the term $B - \mu_\Psi$ in equation (46) using (36), we obtain

$$B - \mu_\Psi = \frac{1}{\varphi} - 2A\left(\frac{2 - e^{2\varphi A}}{1 - e^{2\varphi A}}\right), \tag{50}$$

which is independent of N.

From equation (50) we get the following identity:

$$\tilde{\chi}_{\tilde{t}} = B - \mu_\Psi = \chi_t, \tag{51}$$

where the discrete-time parameter k has been replaced by its continuous-time counterpart t to avoid round-off errors. Expanding (51) using (42), we see that equality between the left and right sides of the equation implies $\tilde{\omega}^{\tilde{t}} = \omega^t$. Thereby,

$$\tilde{t}\ln(1+\gamma\sigma^2_\infty) = t\ln(1+\sigma^2_\infty). \tag{52}$$

Removing the logarithm using a first-order Taylor approximation, we finally get $\tilde{t} = 1/\gamma t$. As a result, the total mean detection delay using $\tilde{N}$ can be written as

$$\tilde{N}\tilde{t} = \tilde{N}\frac{1}{\gamma}t = Nt, \tag{53}$$

indicating that the mean detection delay of the SLRT is approximately independent of N. This result is intuitively appealing since it implies that the detector requires the same amount of information (i.e. total number of received samples) to detect a signal. Moreover, it allows us to choose a small N, which will ensure a minimum detection delay when the actual SNR is higher than expected. It is however important to remember that a key assumption for this theoretical result resides in the use of the CLT to consider that $\boldsymbol{r}_k$ follows a Gaussian distribution, preventing N from being arbitrary small. A smaller N also implies that the SLRT needs to refresh its statistic more often, which may impose additional hardware constraints. As a general rule, the lower the SNR, the larger N can be without detrimental effects on the detection delay.

V. NUMERICAL RESULTS

In this section we will compare the SLRT detector to the sequential eigenvalue-based detector [20] and through the use of Monte Carlo (MC) simulations assess the quality of the theoretical results derived so far (theoretical results are labeled "Th" in the figures). The thresholds A and B are mirrored with respect to the origin ($A = -B$) and are set to achieve a specified

TABLE I
$T_0(A)$ RELATIVE ERROR BETWEEN THEORETICAL VALUES AND MONTE CARLO ESTIMATIONS.

$T_0(A)$	5×10^3	10^4	2×10^4	3×10^4
$T_0(A)$ MC	5933	11315	22109	33537
error (%)	18.6	13.1	10.5	11.6

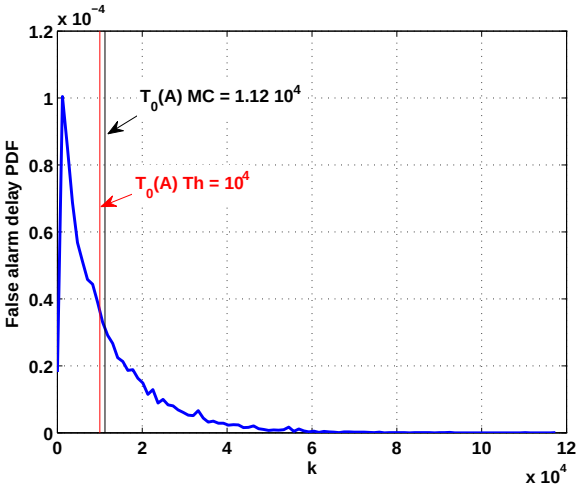

Fig. 2. False-alarm delay statistical distribution and $T_0(A)$. $N = 50$, $p = 10$, $\kappa = 10^{-3}$, $\sigma_0^2 = 2.4 \times 10^{-4}$, $\nu = 20\%$.

mean false-alarm delay $T_0(A)$ (measured with respect to the time index k). We further set $a = A$ and $b = B$ to minimize detection delays. Unless otherwise specified, detection or false-alarm delays are expressed using the "time" index k instead of a physical time value (e.g. seconds) since our results are independent of the sampling frequency used by the detector.

Figure 2 shows the probability density function (PDF) of the delay between successive false-alarms. Parameters κ and σ_0^2 have been set using equations (14) and (17) to achieve optimal detection at a target SNR_L of -15 dB. Thresholds have been set to achieve $T_0(A) = 10^4$, however the mean delay is approximately 12% larger due to the correlation between successive LLR_k. The actual delay is always larger than the theoretical one since the correlation coefficient is negative. Interestingly, the relative error does not remain constant and instead decreases as $T_0(A)$ increases, as can be seen in Table I. While the minimum error is not known, it seems that one should expect an average of 10% deviations from the theoretical results for values of $T_0(A)$ above 10^4 which, while not ideal, seems acceptable to set the detector thresholds.

Next, we compare in Figure 3 the stationary distribution of Ψ_Z computed using (34) with that obtained via

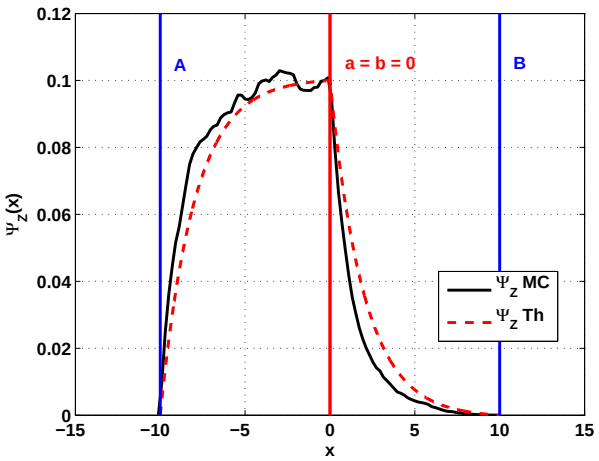

Fig. 3. Stationary distribution of Z_k under H_0 for the case $a = b = 0$ (multicycle Wald detector).$N = 50$, $p = 10$, $\sigma_0^2 = 10^{-4}$, $\kappa = 10^{-3}$.

Monte Carlo simulations. The theoretical result captures the overall shape of the distribution relatively well, but the impact of the correlation is noticeable from the difference between the curves. Fortunately, for the SLRT detector studied in this paper, we are mostly interested in the mean value of the distribution, used to compute the mean detection delay, rather than the exact shape of the distribution.

To study the accuracy of the expression for the mean detection delay in (48) we consider a multivariate signal whose covariance matrix is of rank one, embedded in AWGN. We note that while the detector performance depends on the specific signal model applied, the quality of the theoretical results considered here is not affected. The theoretical and simulated results are plotted in Figure 4. We observe that our solution for R_Z offers a close approximation in the typical SNR range considered for this study.

Figure 5 shows the theoretical mean detection delay (48) as a function of the parameter ν, used to minimize the detection delay of the SLRT as a function of the SNR_L. The minimum delay appears around $\nu = 20\%$ and seems to deviate from this value by only a few percents for a wide range of SNR and mean false-alarm delays. Additionally, small deviations from the optimal value do not drastically affect R_Z, prompting us to use 20% as the default value. It also indicates that small deviations of the SNR do not drastically affect the performance of the SLRT parametrized for a specific SNR.

In Figure 6 we compare the SLRT detector to the sequential eigenvalue-based detector presented in [20].

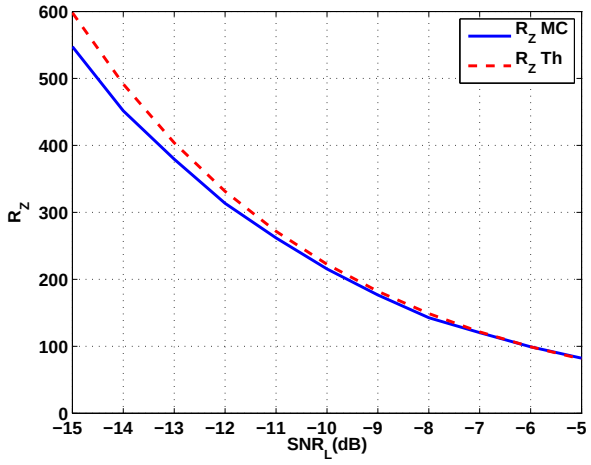

Fig. 4. Mean detection delay R_Z as a function of SNR_L. $T_0(A) = 10^4$, $N = 50$, $p = 10$, $\kappa = 10^{-3}$, $\nu = 20\%$.

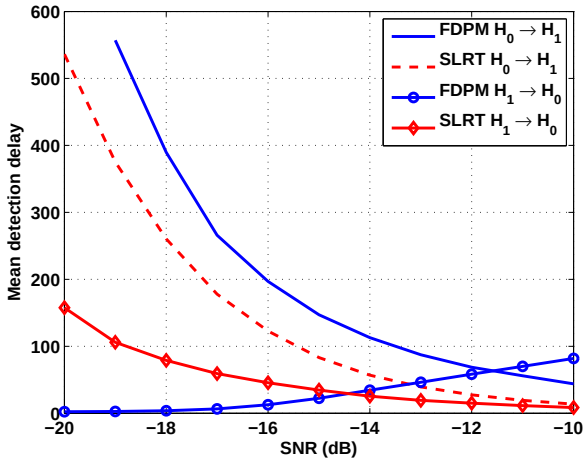

Fig. 6. Mean detection delays for the SLRT and FDPM detectors. $T_0(A) = 10^4$, $\nu = 20\%$, $N = 50$, $p = 10$, $\kappa = 10^{-3}$.

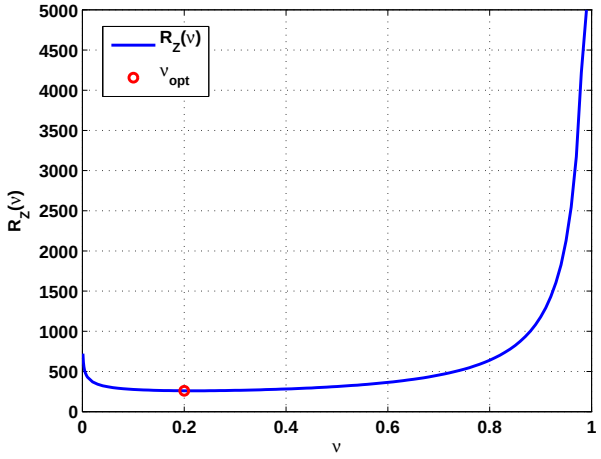

Fig. 5. Mean detection delay R_Z as a function of ν. $T_0(A) = 10^4$, $p = 10$, $\kappa = 10^{-3}$, $SNR_L = -10$dB.

This detector is based on the fast-data projection method (FDPM) and is therefore referred to as such in the following. We again consider a multivariate signal, with a rank-one covariance matrix, embedded in AWGN. This covariance matrix structure corresponds to the ideal scenario for which the scaled-largest eigenvalue (SLE) detector (the batch equivalent of the FDPM detector studied here) has been developed, as a generalized likelihood ratio test (GLRT). As we will see, even though the signal model tends to favor the FDPM detector, the SLRT detector still outperforms it.

Both detectors are set to achieve a mean false-alarm delay of 10^4 using equation (29) for the SLRT. We then compare the mean detection delays ($H_0 \rightarrow H_1$ and $H_1 \rightarrow H_0$), computed as the first-time passage over the detection thresholds after a change in the underlying signal distribution. The SLRT is consistently faster at detecting the presence of a signal, starting from the stationary H_0 regime. Looking at the detection delays corresponding to the disappearance of the communication signal ($H_1 \rightarrow H_0$), one may think that the FDPM detector outperforms the SLRT detector since the detection delay is very small (close to 0 at -20dB), but this actually reflects the SNR limit of the FDPM detector. As the SNR decreases, the mean value of the FDPM statistic under H_1 tends toward its H_0 value. To reliably detect the signal, the variance of the FDPM statistic needs to be reduced accordingly (by increasing N) otherwise the detector cannot distinguish between hypotheses H_0 and H_1 anymore. This is illustrated on Figure 7 which shows the probability of misdetection P_{md} (probability that the statistic remains under the detection threshold) of the SLRT and FDPM detectors under H_1. At low SNR, the FDPM statistic remains below the detection threshold most of the time (e.g. $P_{md} \sim 95\%$ at -20dB). Furthermore, the FDPM $H_1 \rightarrow H_0$ detection delay increases with the SNR since the statistic is not limited by an upper threshold: the mean value of the FDPM statistic increases with the SNR and the detector needs more iterations after the signal disappears to bring the statistic below the detection threshold. Overall, the SLRT clearly outperforms the FDPM detector for the investigated detection scenarios.

Another advantage of the SLRT detector over the FDPM detector is its robustness to the noise distribution.

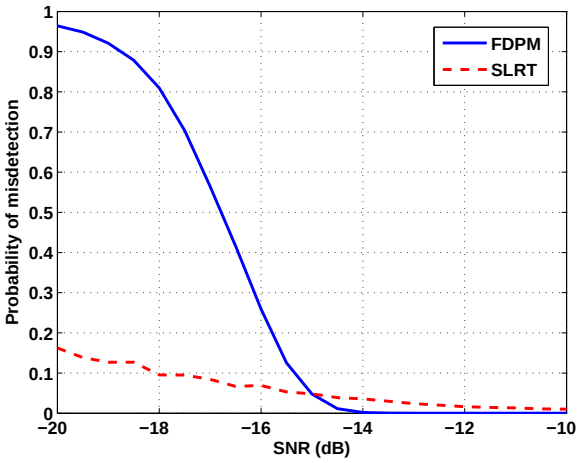

Fig. 7. Probability of misdetection as a function of the SNR. $T_0(A) = 10^4$, $\nu = 20\%$, $p = 10$, $\kappa = 10^{-3}$, $N = 50$.

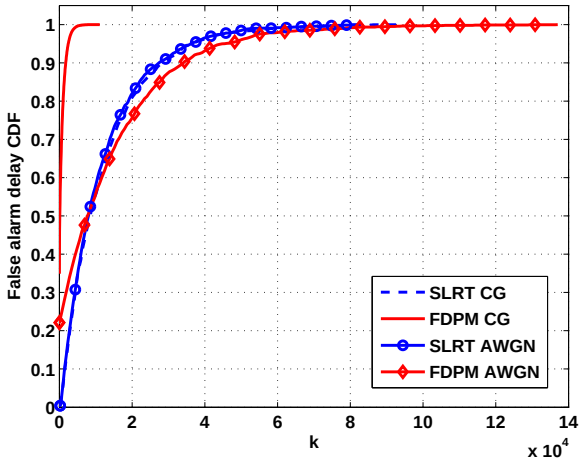

Fig. 8. λ_{fa} cumulative distribution function (CDF) under AWGN and CG noise conditions for the SLRT and FDPM detectors. $T_0(A) = 10^4$, $N = 50$, $p = 10$, $\kappa = 10^{-3}$, $\sigma_0^2 = 10^{-4}$, $p_{CG} = 5\%$, $R = 10$.

The AWGN assumption is a key requirement to use the formulas developed in [20] to set the detection threshold of the FDPM detector. Instead, let us use a contaminated Gaussian (CG) distribution $f_{\boldsymbol{v}} = (1 - p_{CG})\mathcal{N}_q(\mathbf{0}, \boldsymbol{\mathcal{I}}_q) + p_{CG}\mathcal{N}_q(\mathbf{0}, R\,\boldsymbol{\mathcal{I}}_q)$, which corresponds to a two-component Gaussian mixture distribution where the impulsive (i.e. heavy-tail) part of the noise possesses a variance R times higher than the background noise and a small probability of occurrence p_{CG}.

In the following example, we consider a moderately impulsive noise ($R = 10$, $p_{CG} = 0.05$) and consider the effect on the detectors mean false-alarm delay distribution. The impact on the FDPM detector, displayed in Figure 8, is obvious. Under AWGN conditions, both detectors are set to achieve $T_0(A) = 10^4$. If we then change the noise distribution, the FDPM detector shows a mean false-alarm delay $E[\lambda_{fa}]$ that is much lower than the specified value. On the other hand, the SLRT detector remains unaffected (both curves are superimposed).

Finally, to illustrate the advantages of sequential detectors over batch detectors, we compare the SLRT detector to its batch variant developed by Tugnait [11]. We receive a digital video broadcasting terrestrial mode (DVB-T) signal, whose orthogonal frequency division multiplexing (OFDM) frame structure is specified by the European Telecommunications Standards Institute (ETSI) [41]. The signal passes through a multipath Rayleigh fading channel defined by the International Telecommunications Union (ITU) ITU-R, channel B pedestrian model [42] (Appendix 7.D). The batch detector (labeled "Tugnait" in the figures) is parametrized to achieve a probability of detection (P_d) of 95% and a probability of false-alarm (P_{fa}) of 1% at an SNR of -10 dB. The SLRT detector mean false-alarm delay $T_0(A)$ is set to match the batch detector P_{fa} in the sense that the SLRT detector spends on average 1% of the time in a false-alarm mode. The SLRT detector is also configured to achieve optimal detection at an SNR of -10 dB. Keeping these parameters constant, we analyze the detectors detection performance when the SNR varies from -15 dB to -5 dB. Of interest are the detection delays and probabilities of misdetection, illustrated in Figures 9 and 10, respectively.

The parameters N_{SLRT} and N_{Tugnait} correspond to the sample window sizes used by the sequential and batch detectors respectively. N_{Tugnait} is chosen to achieve the desired 95% probability of detection at an SNR of -10 dB, while N_{SLRT} is chosen based on the considerations discussed in Section IV-E. As we can see, the mean detection delay of the SLRT detector is consistently lower than that of the batch detector, even at the target SNR of -10 dB where the SLRT detector is twice as fast to detect the signal. The interquartile range allows to visualize the dispersion of the detection delays at each SNR. Not surprisingly the range decreases as the SNR increases but, for the batch detector, it reaches a limit due to the "dead time" of the detector. This "dead time" reflects the inability for a batch detector to react to the appearance of a signal before the start of a new observation window. On the other hand, a sequential detector works with considerably smaller observation windows ($N_{\text{SLRT}} \ll N_{\text{Tugnait}}$) and can therefore bring the interquartile range to much smaller values.

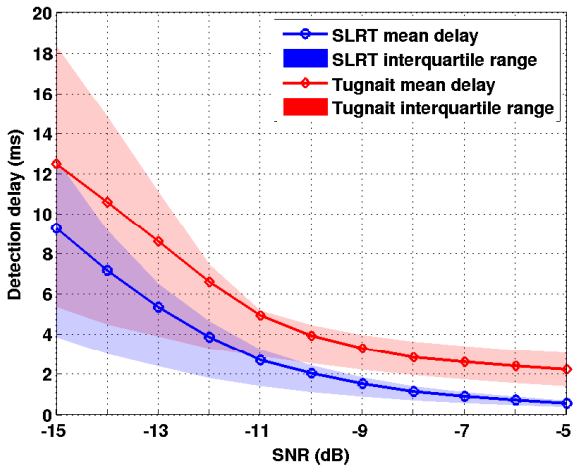

Fig. 9. SLRT and Tugnait detectors mean detection delays as a function of the SNR. $N_{\text{Tugnait}} = 3 \times 10^4$. SLRT parameters: $N_{\text{SLRT}} = 1024$, $p = 12$, $T_0(A) = 5 \times 10^2$, $\nu = 20\%$, $\kappa = 10^{-3}$, $\sigma_0^2 = 0.96$.

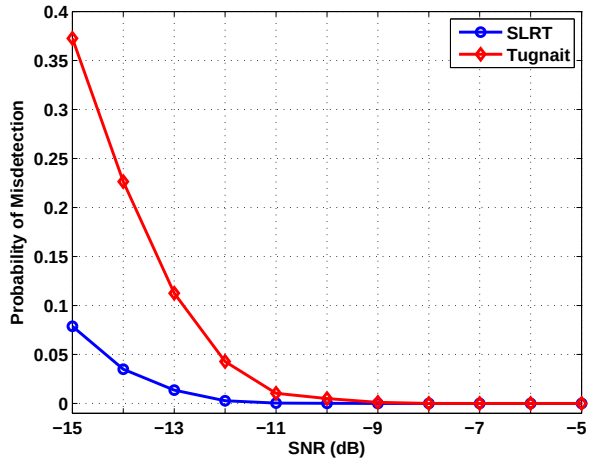

Fig. 10. SLRT and Tugnait detectors probability of misdetection as a function of the SNR. $N_{\text{Tugnait}} = 3 \times 10^4$. SLRT parameters: $N_{\text{SLRT}} = 1024$, $p = 12$, $T_0(A) = 5 \times 10^2$, $\nu = 20\%$, $\kappa = 10^{-3}$, $\sigma_0^2 = 0.96$.

Complementary to Figure 9, Figure 10 shows the probability of misdetection for both detectors. We consider that the detectors fail to detect the signal if they cannot detect it before a maximum delay set to 10 observation windows of sample size N_{Tugnait}. As we can see, the SLRT detector clearly outperforms the batch detector when the actual SNR is lower than expected. For instance, at an SNR of -15 dB, the probability of misdetection for the batch detector rises to 37%, as opposed to 8% for the SLRT detector. This indicates that the SLRT detector is much more robust to SNR uncertainties, both in terms of detection delays and probability of misdetection, than its batch counterpart.

VI. Conclusion

In this paper we have proposed a new SLRT detector exploiting the correlation of a multivariate communication signal and considered its use for CR spectrum sensing scenarios. We derived analytical results regarding the statistical properties of the test, needed to parametrize the detector and evaluate its performance, and compared the detector to a sequential eigenvalue-based detector and a batch detector designed to operate under the same conditions. The comparison revealed the competitiveness of the SLRT as well as its robustness to impulsive noise.

References

[1] E. Axell, G. Leus, E. Larsson, and H. Poor, "Spectrum Sensing for Cognitive Radio: State-of-the-Art and Recent Advances," *IEEE Signal Processing Mag.*, vol. 29, pp. 101–116, May 2012.

[2] J. Lunden, V. Koivunen, and H. V. Poor, "Spectrum Exploration and Exploitation for Cognitive Radio: Recent Advances," *IEEE Signal Processing Mag.*, vol. 32, pp. 123–140, May 2015.

[3] Z. Quan, S. Cui, H. V. Poor, and A. H. Sayed, "Collaborative wideband sensing for cognitive radios," *IEEE Trans. Signal Processing*, vol. 25, pp. 60–73, Nov. 2008.

[4] J. Lunden, S. Kassam, and V. Koivunen, "Robust Nonparametric Cyclic Correlation-Based Spectrum Sensing for Cognitive Radio," *IEEE Trans. Signal Processing*, vol. 58, pp. 38 –52, Jan. 2010.

[5] F. Bhatti, G. Rowe, K. Sowerby, and C. da Silva, "Blind Signal Detection Using a Linear Antenna Array: An Experimental Approach," *IEEE Trans. Veh. Technol.*, vol. 63, pp. 1135–1145, Mar. 2014.

[6] F. Penna and S. Staczak, "Decentralized Eigenvalue Algorithms for Distributed Signal Detection in Wireless Networks," *IEEE Trans. Signal Processing*, vol. 63, pp. 427–440, Jan. 2015.

[7] K. Hassan, R. Gautier, I. Dayoub, M. Berbineau, and E. Radoi, "Multiple-Antenna-Based Blind Spectrum Sensing in the Presence of Impulsive Noise," *IEEE Trans. Veh. Technol.*, vol. 63, pp. 2248–2257, June 2014.

[8] F. Moghimi, A. Nasri, and R. Schober, "Adaptive Lp Norm Spectrum Sensing for Cognitive Radio Networks," *IEEE Trans. Commun.*, vol. 59, pp. 1934–1945, July 2011.

[9] E. Axell and E. Larsson, "A unified framework for GLRT-based spectrum sensing of signals with covariance matrices with known eigenvalue multiplicities," in *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 2956–2959, May 2011.

[10] J. Lunden, V. Koivunen, A. Huttunen, and H. Poor, "Collaborative Cyclostationary Spectrum Sensing for Cognitive Radio Systems," *IEEE Trans. Signal Processing*, vol. 57, pp. 4182 –4195, Nov. 2009.

[11] J. Tugnait, "On autocorrelation-based multiantenna spectrum sensing for cognitive radios in unknown noise," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 2944–2947, May 2011.

[12] J. Font-Segura, J. Riba, J. Villares, and G. Vazquez, "Quadratic sphericity test for blind detection over time-varying frequency-selective fading channels," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 4708–4712, May 2013.

[13] Y. Zeng, Y.-C. Liang, and R. Zhang, "Blindly Combined Energy Detection for Spectrum Sensing in Cognitive Radio," *IEEE Signal Processing Lett.*, vol. 15, pp. 649–652, 2008.
[14] FCC, "Notice of Proposed Rule Making and Order (FCC 03-322): Facilitating Opportunities for Flexible, Efficient, and Reliable Spectrum Use Employing Cognitive Radio Technologies," ET Docket No. 03-108, Federal Communications Commission, Dec. 2003.
[15] ECC REPORT 159, "Technical and Operational Requirements for the Possible Operation of Cognitive Radio Systems in the White Spaces of the Frequency Band 470-790 MHz," Tech. Rep. 159, CEPT, Cardiff, Wales, Jan. 2011. [online, March 2016] http://www.erodocdb.dk.
[16] A. N. Shiryaev, "Quickest Detection Problems: Fifty Years Later," *Sequential Analysis*, vol. 29, no. 4, pp. 345–385, 2010.
[17] V. H. Poor, *An Introduction to Signal Detection and Estimation*. Springer, 1994.
[18] J. Manco-Vásquez, M. Lázaro-Gredilla, D. Ramírez, J. Vía, and I. Santamaría, "A Bayesian approach for adaptive multiantenna sensing in cognitive radio networks," *Elsevier Signal Processing*, vol. 96, Part B, p. 228240, Mar. 2014.
[19] X. Doukopoulos and G. Moustakides, "Fast and Stable Subspace Tracking," *IEEE Trans. Signal Processing*, vol. 56, pp. 1452–1465, Apr. 2008.
[20] C. G. Tsinos and K. Berberidis, "Decentralized Adaptive Eigenvalue-Based Spectrum Sensing for Multiantenna Cognitive Radio Systems," *IEEE Trans. Wireless Commun.*, vol. 14, pp. 1703–1715, Mar. 2015.
[21] L. Lai, Y. Fan, and H. Poor, "Quickest Detection in Cognitive Radio: A Sequential Change Detection Framework," in *IEEE GLOBECOM 2008*, pp. 1–5, Nov. 2008.
[22] A. Wald and J. Wolfowitz, "Optimum Character of the Sequential Probability Ratio Test," *The Annals of Mathematical Statistics*, vol. 19, no. 3, pp. 326–339, 1948.
[23] F. Lin, R. C. Qiu, and J. P. Browning, "Spectrum Sensing With Small-Sized Data Sets in Cognitive Radio: Algorithms and Analysis," *IEEE Trans. Veh. Technol.*, vol. 64, pp. 77–87, Jan. 2015.
[24] T. L. Lai, "Sequential analysis: Some classical problems and new challenges," *Statistica Sinica*, vol. 11, pp. 303–408, Apr. 2001.
[25] H. Li, C. Li, and H. Dai, "Quickest spectrum sensing in cognitive radio," in *CISS 2008. 42nd Annual Conference on Information Sciences and Systems.*, pp. 203–208, Mar. 2008.
[26] W.-H. Chung, "Sequential Likelihood Ratio Test under Incomplete Signal Model for Spectrum Sensing," *IEEE Trans. Wireless Commun.*, vol. 12, pp. 494–503, Feb. 2013.
[27] T. Kailath and H. Poor, "Detection of stochastic processes," *IRE Trans. Inform. Theory*, vol. 44, pp. 2230–2231, Oct. 1998.
[28] R. Tandra and A. Sahai, "SNR Walls for Signal Detection," *IEEE J. Select. Topics. Signal Processing*, vol. 2, pp. 4 –17, Feb. 2008.
[29] R. Lopez-Valcarce, G. Vazquez-Vilar, and J. Sala, "Multiantenna spectrum sensing for Cognitive Radio: overcoming noise uncertainty," in *2nd International Workshop on Cognitive Information Processing (CIP)*, pp. 310–315, June 2010.
[30] Y. Zeng and Y.-C. Liang, "Spectrum-Sensing Algorithms for Cognitive Radio Based on Statistical Covariances," *IEEE Trans. Veh. Technol.*, vol. 58, pp. 1804–1815, May 2009.
[31] S. Kay, *Fundamentals of Statistical Signal Processing: Vol I Estimation Theory*. Prentice-Hall PTR, 1998.
[32] A. N. Shiryaev, "On the Detection of Disorder in a Manufacturing Process. I.," *Theory of Probability & Its Applications*, vol. 8, no. 3, pp. 247–265, 1963.
[33] S. Van Vaerenbergh, M. Lazaro-Gredilla, and I. Santamaria, "Kernel Recursive Least-Squares Tracker for Time-Varying Regression," *IEEE Trans. Neural Networks*, vol. 23, pp. 1313–1326, Aug. 2012.
[34] S. Karlin and H. Taylor, *A Second Course in Stochastic Processes*. Academic Press, 1981.
[35] P. W. Glynn, "Stochastic Models, Chapter 4: Diffusion approximations," *Handbooks in operations research and management science*, vol. 2, pp. 145–198, 1990.
[36] A. Klenke, *Probability theory: a comprehensive course*. Springer, 2008.
[37] S. Asmussen, *Applied probability and queues*. Applications of mathematics, Springer, 2003.
[38] D. Kalman, "A Generalized Logarithm for Exponential-Linear Equations," *The College Mathematics Journal*, vol. 32, Jan. 2001.
[39] T. T. Chien, *An adaptive technique for a redundant-sensor navigation system*. PhD thesis, Massachusetts Institute of Technology. Dept. of Aeronautics and Astronautics, July 1972.
[40] T. Guillaume, "On the Probability of Hitting a Constant or a Time-Dependent Boundary for a Geometric Brownian Motion with Time-Dependent Coefficients," *Hikari, Applied Mathematical Sciences*, vol. 8, no. 20, pp. 989 – 1009, 2014.
[41] ETSI, "ETSI EN 300 744 V1.6.2, Digital Video Broadcasting (DVB); Framing structure, channel coding and modulation for digital terrestrial television," tech. rep., European Telecommunications Standards Institute (ETSI), Oct. 2015. [online, oct 2015] http://www.etsi.org.
[42] A. Molisch, *Wireless Communications*. J. Wiley & Sons, 2005.

Julien Renard (J'10) received the M.Sc. (Eng.) degree with High Distinction in electrical engineering from the Applied Science Faculty of the Universite Libre de Bruxelles (ULB), Brussels, Belgium, in 2009.

He is currently pursuing a D.Sc (eng.) degree, organized as a joint degree between the ULB and the University of British Columbia, Vancouver, Canada.

Lutz Lampe (M02SM08) received the Dipl.-Ing. and Dr.-Ing. degrees in electrical engineering from the University of Erlangen, Erlangen, Germany, in 1998 and 2002, respectively. Since 2003, he has been with the Department of Electrical and Computer Engineering, The University of British Columbia, Vancouver, BC, Canada, where he is a Full Professor. His research interests are broadly in theory and application of wireless, optical wireless, and power line communications. Dr. Lampe was the General (Co-)Chair for 2005 ISPLC and 2009 IEEE ICUWB and the General Co-Chair for the 2013 IEEE SmartGridComm. He is currently an Associate Editor of the IEEE WIRELESS COMMUNICATIONS LETTERS and the IEEE COMMUNICATIONS SURVEYS AND TUTORIALS and has served as an Associate Editor and a Guest Editor of several IEEE transactions and journals. He was a (co-)recipient of a number of best paper awards, including awards at the 2006 IEEE International Conference on Ultra-Wideband, the 2010 IEEE International Communications Conference, and the 2011 IEEE International Conference on Power Line Communications.

Francois Horlin received the electrical engineering degree and the Ph.D. degree from the Universit catholique de Louvain (UCL), Louvain-la-Neuve, Belgium, in 1998 and 2002 respectively. During his studies, he specialized in the field of digital signal processing for communications. He led the project aiming at developing a fourth-generation wireless communication system in collaboration with Samsung Korea. Since 2014 Francois Horlin is full professor at the Universite libre de Bruxelles (ULB), Brussels, Belgium. He is currently giving three lectures in the field of digital telecommunications and is advisor of 6 Ph.D. students (plus 13 already defended Ph.D. theses). He chairs currently the IEEE signal processing chapter of the Benelux.