

Scaled Largest Eigenvalue Detection for Stationary Time-Series

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

Abstract—This paper studies the performance of the Scaled Largest Eigenvalue (SLE) detector used for the detection of stationary time-series. We focus on a single-antenna setup and a blind detection scenario (neither the signal covariance, nor the noise variance are known). The SLE detector has received much attention in the context of Cognitive Radios (CR) due to its simplicity, good performance and robustness to noise level uncertainties. Specifically, our goal is to analyze the detector based on the statistic $\Gamma = \frac{\lambda_1}{\sum_{i=1}^p \lambda_i}$, where $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$ represent the ordered eigenvalues of the sample covariance matrix. We derive a large-sample-size closed-form approximation for the test statistic which allows us to derive its statistical distribution and set up the detector to achieve the required Probability of False-Alarm (P_{fa}) and Probability of Detection (P_d). We also study the robustness of the detector in the presence of noise uncertainty and impulsive-noise and investigate the benefits of the spatial sign filter for such scenarios.

Index Terms—Eigenvalue Detection, Cognitive Radios, Spatial Sign Function, Matrix Perturbation Theory

I. INTRODUCTION

Among the various blind spectrum sensing techniques for Cognitive Radios (CR) proposed in the literature [1], eigenvalue detection has received its share of interest due to its simplicity, good performance in low Signal-to-Noise Ratio (SNR) environments, and its robustness to SNR wall phenomena. While several algorithms exploiting the structure of the sample covariance matrix

Copyright (c) 2015 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

This work was supported by the “Fonds pour la formation a la Recherche dans l’Industrie et dans l’Agriculture” (FRIA), Belgium and 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).

through its eigenvalues do exist [2], [3], [4], [5], we focus our attention on the Scaled Largest Eigenvalue (SLE) detector whose test statistic is given by the ratio of the largest eigenvalue to the trace (or equivalently, the sum of its eigenvalues) of the sample covariance matrix $\hat{\Sigma}_x$.

When the received signal is a sequence of i.i.d multivariate Gaussian vectors and the signal to be detected possesses a rank-one covariance matrix, this detector has been derived as a sufficient statistic for the Generalized Log-Likelihood Ratio Test (GLRT) [2]. It has also been shown to maximize the SNR among all the statistics that linearly combine the received samples [6]. Moreover, its statistical distribution is not affected by the noise variance making it a Constant False-Alarm Rate (CFAR) detector. CFAR detectors have sparked considerable interest due to their adequacy in low SNR environments, which is a typical scenario for CR systems.

Despite its apparent simplicity, studies of the SLE detector have been hampered by the complexity of its statistical distribution. Deriving the distribution of a ratio of eigenvalues is not a trivial problem. To the authors’ knowledge, analytical expressions for the Probability Density Function (PDF) of the SLE detector have been derived for the following scenarios:

- The sample covariance matrix follows a Wishart distribution, and its size p tends toward infinity (the ratio of number of samples to the size of the matrix $\beta = \frac{p}{N}$ being constant), or is large enough to use the asymptotic distribution (Tracy-Widom) as an approximation [7].
- The sample covariance matrix follows a Wishart distribution, and its size is small enough ($p \leq 5$) to use the closed-form expression derived in [8].

However, the sample covariance matrix $\hat{\Sigma}_x$ will only follow a Wishart distribution if the received samples are i.i.d. vectors of size p following a multivariate normal distribution $\mathcal{N}_p(0, \Sigma_x)$. Any deviation from this assumption leaves us without any expression for the statistical distribution of the detector and therefore poses a problem to adjust its parameters. In many situations of

practical importance the aforementioned hypothesis can be violated, for instance in the presence of impulsive noise, temporal correlation between the received samples or the addition of a non-linear filter to improve the detector performance. Moreover, if the detector exploits the temporal covariance of a stationary time-series, the covariance matrix becomes a Toeplitz matrix and therefore cannot follow a Wishart distribution.

In this paper, we set out to investigate the statistical distribution of the SLE detector working with temporal covariance matrices under the assumption that the received signals are wide-sense stationary. We consider a single-antenna system in blind-detection scenario where neither the covariance matrix of the signal nor the noise variance are known. In such a blind-detection scenario, tests such as the Ljung-Box test [9] or some variation can be used to test for the presence of correlation among samples of the time-series. However, we will show that eigenvalue-based detectors can be more efficient for such scenarios (even though they are not GLRT anymore).

Our main goal is to derive a closed-form approximation for the statistical distribution of $\Gamma = \frac{\lambda_1}{\sum_{i=1}^p \lambda_i}$, the statistic of the SLE detector. This approximation will provide some insight regarding the influence of the various parameters of the signal on the detector performance and allow us to compute the detector threshold without the use of Monte Carlo (MC) simulations. We will address the issue of robustness in the presence of noise model uncertainty, such as deviations from the Average White Gaussian Noise (AWGN) assumption commonly used and see how the SLE detector copes with such non-ideal conditions. Subsequently, we will study the benefits of the spatial sign function non-linearity, applied to the received samples, to mitigate the impact of such uncertainties and use our model to seek out the statistical distribution of the SLE detector when combined with the spatial sign function.

Additionally, as a by-product of our statistical model, we will introduce a new, computationally simpler detector called the Scaled Largest Circulant Eigenvalue (SLCE) detector and compare it to the SLE detector.

The remainder of this paper is organized as follows: In Section II we introduce the signal models and hypotheses used to derive our main result. Section III is used to approximate the statistical distribution of λ_1 , the largest eigenvalue of the sample covariance matrix, which is needed in Section IV, where we move on to the statistical distribution of the SLE detector statistic Γ . Section V introduces the spatial sign function and the modifications imparted on the theoretical results of the previous section. Numerical results are presented in Section VI,

where we also analyze the limitations of our approximations. Furthermore, we will compare the SLE detector to other eigenvalue detectors and the Ljung-Box detector. We conclude with a summary of our contributions in Section VII.

II. SYSTEM MODEL

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

A. Received signal

As is typically done in the context of spectrum sensing, we consider a baseband discrete-time signal $x(n)$ at the receiver and set out to discriminate between the following hypotheses:

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

where $s(n)$ is the communication signal to be detected and $v(n)$ represents an i.i.d. process of noise and interference. All signals are assumed to be zero mean and if needed, we may subtract the sample mean from the received sample vectors. We shall refer to the hypothesis H_0 as the null hypothesis and the hypothesis H_1 as the alternative hypothesis. We assume that the input signals are Wide-Sense Stationary (WSS) over the observation window made of N samples. As a result, the unknown temporal covariance matrix of the signal $x(n)$, denoted as Σ_x , has a Toeplitz structure and in order to proceed with the detection of the signal $s(n)$, we need an estimate $\hat{\Sigma}_x$ of Σ_x .

B. Covariance matrix estimate

Constraining a sample covariance matrix to have a Toeplitz structure is a difficult problem, that often involves Maximum Likelihood (ML) estimation¹ and iterative algorithms (see [11], [12] and the references therein). Such approaches make it difficult to derive the statistical distribution of the estimated covariance matrix and for this reason, we will instead use more naive (but tractable) estimation techniques². More specifically, we may use either of the following two approximations. We

¹Please note that the usual ML sample covariance matrix estimate for multivariate Gaussian signals [10], typically used in multiantenna scenarios, does not apply here since we consider a time-series scenario and do not assume that the received samples follow a Gaussian distribution.

²As a result, please note that even if the time-series samples follow a Gaussian distribution, neither of the following estimators for Σ_x are ML.

can first compute the vector $\hat{\boldsymbol{\zeta}} = [\hat{\zeta}_0, \hat{\zeta}_1, \dots, \hat{\zeta}_{p-1}]$ that contains the sample estimates of the auto-correlation at time lags $\{0, 1, 2, \dots, p-1\}$:

$$\hat{\zeta}_i = \frac{1}{N-i+1} \sum_{n=0}^{N-i+1} x(n)x^*(n+i), \quad (2)$$

where $(\cdot)^*$ denotes the complex conjugate operator, and build the Hermitian Toeplitz matrix $\hat{\boldsymbol{\Sigma}}_x = \text{Toep}(\hat{\boldsymbol{\zeta}})$. This matrix has a Toeplitz structure, but is not guaranteed to be positive-definite. Alternatively, we can construct the matrix

$$\mathbf{X} = \begin{pmatrix} x_0 & x_1 & \dots & x_{p-1} \\ x_1 & x_2 & \dots & x_p \\ \vdots & \vdots & \dots & \vdots \\ x_{N-p} & x_{N-p+2} & \dots & x_{N-1} \end{pmatrix}, \quad (3)$$

using overlapping sample vectors and build the sample covariance matrix $\hat{\boldsymbol{\Sigma}}_x = \frac{1}{N-p+1} \mathbf{X}^H \mathbf{X}$, where $(\cdot)^H$ is the conjugate transpose operator. The matrix $\hat{\boldsymbol{\Sigma}}_x$ defined this way is a true covariance matrix in the sense that it is Hermitian and positive-definite, but is not Toeplitz. However, when the sample size N becomes large with respect to the size p of the matrix, the difference between the two approximations becomes negligible. From now on, we will assume that the ratio $\beta = \frac{N}{p}$ is large enough (e.g. $\beta = 100$) to consider that $\hat{\boldsymbol{\Sigma}}_x$ is both Toeplitz and positive-definite.

C. SLE detector statistic

Under hypothesis H_0 , the received samples are assumed to be i.i.d³ and the true covariance matrix $\boldsymbol{\Sigma}_x$ is a scaled identity matrix. Intuitively, under the hypothesis H_1 , the signal $s(n)$ adds correlation to the received signal $x(n)$ which causes the off-diagonal elements of $\boldsymbol{\Sigma}_x$ to be non-zero. The SLE detector uses the eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$ of $\hat{\boldsymbol{\Sigma}}_x$ to measure these changes by computing the statistic

$$\Gamma = \frac{\lambda_1}{\sum_{i=1}^p \lambda_i}. \quad (4)$$

Under hypothesis H_1 , the largest eigenvalue λ_1 increases, indicating the presence of a communication signal. λ_1 is divided by the sum of all the eigenvalues in order to make the statistic Γ independent from any scaling effect (i.e. multiplying $x(n)$ by a constant does not affect Γ), which implies that the detector does not depend on the variance of the noise and is thereby CFAR.

For finite-size matrices, the statistical distribution of $\hat{\boldsymbol{\Sigma}}_x$ is paramount to evaluating the distribution of the

³Any correlation arising from the presence of a filter at the receiver can be canceled using the procedure described in [13].

SLE statistic Γ . If $\hat{\boldsymbol{\Sigma}}_x$ is Toeplitz, its entries are entirely determined by the vector $\hat{\boldsymbol{\zeta}}$, whose statistical distribution is linked to the input signal $x(n)$ via equation (2). If we suppose that the noise samples are i.i.d complex Gaussian distributed, $v(n) \sim \mathcal{CN}(0, \sigma_n^2)$, the vector $\hat{\boldsymbol{\zeta}}$ is composed of quadratic forms of Gaussian random variables, which can be approximated using Chi-squared distributions as shown in [14]. However, in order to extend the theory derived in this paper to non-Gaussian signals, we will instead assume that the number of samples N is large enough so that the Central-Limit Theorem (CLT) can be used to model the statistical distribution of $\hat{\boldsymbol{\zeta}}$ using a multivariate complex Gaussian distribution.

III. STATISTICAL DISTRIBUTION OF λ_1

As can be seen from equation (4), the SLE detector statistic depends on the largest eigenvalue λ_1 of $\hat{\boldsymbol{\Sigma}}_x$. In order to derive the statistical distribution of Γ , we first need to derive the distribution of λ_1 , which is the main objective of this section.

Despite being heavily structured, there is no general solution to the eigenvalue problem of random Toeplitz matrices such as $\hat{\boldsymbol{\Sigma}}_x$. Therefore, we seek an approximation for the statistical distribution of λ_1 , which we will construct in two main steps: first we approximate $\hat{\boldsymbol{\Sigma}}_x$ by a circulant matrix \mathbf{C} and derive the statistical distribution of its largest eigenvalue. Second, using the results of step one and the matrix perturbation theory developed in [15], [16], [17], we propose an approximation for the statistical distribution of λ_1 .

A. Circulant approximation for Toeplitz matrices

Circulant matrices are a special type of Toeplitz matrices for which the last element of a given row becomes the first element of the next row via a cyclic shift:

$$\begin{pmatrix} a & b & c & d \\ d & a & b & c \\ c & d & a & b \\ b & c & d & a \end{pmatrix} \quad (5)$$

Circulant matrices benefit from many interesting properties related to their eigenvalues and eigenvectors. For instance, their eigenvectors correspond to the columns of a Discrete Fourier Transform (DFT) matrix:

$$\mathbf{D}_{i,j} = \left(\frac{1}{\sqrt{p}} e^{-j \frac{2\pi}{p} (i \times j)} \right) \{i, j\} \in \{0, 1, \dots, p-1\}, \quad (6)$$

where the indices $\{i, j\}$ denote the i^{th} and j^{th} row and column respectively. As a result, the eigenvalues of a

circulant matrix C can be obtained as

$$\Lambda = D^H C D, \quad (7)$$

where Λ is a diagonal matrix of the eigenvalues.

It is well known that banded Toeplitz matrices can be asymptotically (i.e. when their size increases to infinity) approximated by circulant matrices, in the sense that their eigenvalue spectrum behave similarly [18], [19]. Unfortunately, the matrix $\hat{\Sigma}_x$ is not banded and we seek an approximation for its individual eigenvalues, instead of its spectrum. Nevertheless, approximating $\hat{\Sigma}_x$ by a circulant matrix C will give us a starting point to use the matrix perturbation theory later on.

In order to construct C , we use the approximation developed in [20]:

$$\begin{aligned} C &= D^H \text{diag}(D \hat{\Sigma}_x D^H) D = D^H \text{diag}(\mathbf{R}) D \\ &= D^H \text{diag}(\mathbf{r}) D, \end{aligned} \quad (8)$$

where $\text{diag}(\mathbf{R})$ is a diagonal matrix built from the main diagonal of the matrix $\mathbf{R} = D \hat{\Sigma}_x D^H$ and $\text{diag}(\mathbf{r})$ is a diagonal matrix built from the vector \mathbf{r} , which contains the main diagonal elements of \mathbf{R} . Since the DFT matrix is unitary, the matrix \mathbf{R} possesses the same eigenvalues as $\hat{\Sigma}_x$ and $\text{diag}(\mathbf{R})$ turns out to be a diagonal matrix whose elements $\mathbf{r} = [r_1, r_2, \dots, r_p]$ are the eigenvalues of C .

We see that approximating $\hat{\Sigma}_x$ by C is equivalent to approximating the eigenvalues of \mathbf{R} by \mathbf{r} . Gershgorin circles [21] can be used to bound the error of such approximations (replacing the eigenvalues of a matrix by its diagonal elements). Unfortunately, the bounds are much larger than the spread of the statistical distribution of the eigenvalues, indicating that the approximation is too crude to directly derive the distribution of Γ . Nevertheless, we will make use of this approximation later on and therefore need the statistical distribution of the largest eigenvalue of the matrix C (which is equivalent to $\max(\mathbf{r})$)⁴.

B. Joint statistical distribution of the eigenvalues of a circulant matrix

We will now derive the joint statistical distribution of the eigenvalues of the matrix C . In order to do that, we will start from the distribution of the estimated covariance vector $\hat{\zeta}$ and track down its transformation going from $\hat{\zeta}$ to \mathbf{r} .

As we explained in Section II-C, we assume that the vector $\hat{\zeta}$, whose elements form the matrix $\hat{\Sigma}_x$, is distributed as a multivariate complex Gaussian random vari-

⁴Note that since we are working with covariance matrices, which are positive semi-definite, all the eigenvalues are real and non-negative.

able. Moreover, equation (8) indicates that the elements of \mathbf{r} can be obtained from the real and imaginary parts of $\hat{\zeta}$ using a linear transformation. More specifically, let us define the matrix

$$T = \begin{bmatrix} D_1 T_1 D_1^H & \dots & D_1 T_p D_1^H & D_1 \tilde{T}_1 D_1^H & \dots \\ D_2 T_1 D_2^H & \dots & D_2 T_p D_2^H & D_2 \tilde{T}_1 D_2^H & \dots \\ \vdots & & & & \\ D_p T_1 D_p^H & \dots & D_p T_p D_p^H & D_p \tilde{T}_1 D_p^H & \dots \end{bmatrix}$$

where D_i is the i^{th} row of the DFT matrix D , T_i is a $p \times p$ symmetric Toeplitz matrix whose elements are different from zero only on the upper and lower i^{th} diagonals, where they are equal to one, and \tilde{T}_i is a $p \times p$ Hermitian Toeplitz matrix whose elements are different from zero only on the upper and lower i^{th} diagonals, where they are equal to $-j$ and j respectively (defined as $j = \sqrt{-1}$). The first diagonal is defined as the main diagonal. Then the vector \mathbf{r} is equal to

$$\mathbf{r} = T \times \begin{bmatrix} \Re(\hat{\zeta}) \\ \Im(\hat{\zeta}) \end{bmatrix}, \quad (9)$$

where the operators $\Re(\cdot)$ and $\Im(\cdot)$ extract the real and imaginary parts of their argument.

As a result, the statistical distribution of \mathbf{r} is given by $\mathbf{r} \sim \mathcal{N}(T\boldsymbol{\mu}_{\hat{\zeta}}, T\boldsymbol{\Sigma}_{\hat{\zeta}}T^T) = \mathcal{N}(\boldsymbol{\mu}_{\mathbf{r}}, \boldsymbol{\Sigma}_{\mathbf{r}})$, where $\boldsymbol{\mu}_{\hat{\zeta}}$ and $\boldsymbol{\Sigma}_{\hat{\zeta}}$ are the mean vector and covariance matrix of $[\Re(\hat{\zeta}) \ \Im(\hat{\zeta})]^T$, respectively.

C. Statistical distribution of the largest eigenvalues of a circulant matrix

Our interest lies with the largest eigenvalue of C or equivalently, the maximum element of \mathbf{r} .

If we suppose that the maximum element of \mathbf{r} is r_1 , then its statistical distribution can be derived from \mathbf{r} using appropriate constraints on dummy variables. We define the vector $\mathbf{u} = [u_2, u_3, \dots, u_p]$ whose elements are $u_i = r_i - r_1$ and the variable $v = r_1$:

$$\begin{pmatrix} u_2 \\ u_3 \\ \vdots \\ u_p \\ v \end{pmatrix} = \underbrace{\begin{pmatrix} -1 & 1 & 0 & \dots & 0 \\ -1 & 0 & 1 & \dots & 0 \\ -1 & 0 & 0 & \dots & 0 \\ \vdots & & \ddots & & \vdots \\ 1 & 0 & 0 & \dots & 0 \end{pmatrix}}_{\mathbf{B}} \begin{pmatrix} r_1 \\ r_2 \\ r_3 \\ \vdots \\ r_p \end{pmatrix} \quad (10)$$

The statistical distribution of r_1 given that $r_1 > r_n, \forall n > 2$ is therefore the distribution of v under the constraints that $\mathbf{u} < \mathbf{0}$ (where $\mathbf{0}$ is a null vector of size $(p-1)$).

Since any element of \mathbf{r} can be the maximum, we need to repeat this process for all the elements of \mathbf{r} (by successively assigning $v = r_2, v = r_3, \dots$) and combine

the corresponding conditional distributions. However, under hypothesis H_0 , we suppose that the noise samples are uncorrelated, making the covariance matrix Σ_x a scaled identity matrix. Upon transformation via (9), Σ_r becomes a circulant matrix, which is invariant over cyclic permutations of the elements r . Therefore, the distribution of $\max(r)$ is simply equal to the conditional distribution of r_1 . As a result, the distribution of λ_{C_1} , the largest eigenvalue of C , is given by

$$F_{\lambda_{C_1}}(x) = p \Pr(v \leq x, u_2 \leq 0, \dots, u_p \leq 0) \quad (11)$$

which shows that the variable λ_{C_1} possesses (under H_0) a Selection Normal distribution $SLCT - N_{1,p-1}(\mathbf{B}\mu_r, \mathbf{B}\Sigma_r\mathbf{B}^T, \mathbf{0})$, whose density function is given in Appendix A. Under hypothesis H_1 the covariance matrix Σ_r is not circulant anymore and the distribution of λ_{C_1} becomes a mixture of Selection Normal distributions whose elements can be computed using (10) after a suitable permutation of the elements of r .

D. λ_1 as a function of λ_{C_1}

In our second main step, we aim at closing the gap from approximating the sample covariance matrix $\hat{\Sigma}_x$ by the circulant matrix C . Since we are only interested in the eigenvalues of the aforementioned matrices, we will instead work with the matrices \mathbf{R} and $\text{diag}(\mathbf{R})$ which have the same eigenvalues as $\hat{\Sigma}_x$ and C , respectively. We will use the matrix perturbation theory derived in [15], [16], [17] and presented in Appendix B⁵, which allows to derive the eigenvalues of the matrix \mathbf{R} as a function of the eigenvalues of $\text{diag}(\mathbf{R})$ and the perturbation matrix $\mathbf{E} = \mathbf{R} - \text{diag}(\mathbf{R})$.

We will need the following submatrices.

- \mathbf{R}_{11} is the maximum element of the main diagonal of \mathbf{R} .
- \mathbf{R}_{12} and \mathbf{R}_{21} are the row and column vectors, respectively, corresponding to the off-diagonal elements of the row and column that contain \mathbf{R}_{11} .
- \mathbf{R}_{22} is the submatrix made of the elements of \mathbf{R} , excluding \mathbf{R}_{11} , \mathbf{R}_{12} and \mathbf{R}_{21} .

Using equation (33) we can express λ_1 as a function of λ_{C_1} ,

$$\lambda_1 = \lambda_{C_1} + \mathbf{R}_{12}\mathbf{P}, \quad (12)$$

where the vector \mathbf{P} is defined in equation (34). Unfortunately, (12) does not directly lead to a useful analytical

⁵The theory presented in Appendix B is derived for a general case. Some simplifications occur in this section due to our particular choice of matrices.

expression for λ_1 due to the term $\mathbf{R}_{12}\mathbf{P}$ ⁶. The difficulty inherent to the statistical distribution of $\mathbf{R}_{12}\mathbf{P}$ combined with its small variance⁷ (compared to λ_{C_1}) motivates the approximation $\mathbf{R}_{12}\mathbf{P} \sim E[\mathbf{R}_{12}\mathbf{P}]$, whereby we replace the random variable by its expectation $E[\cdot]$. As a result, we consider that the *distribution* of λ_1 is identical to that of λ_{C_1} apart from the offset $\Delta \equiv E[\mathbf{R}_{12}\mathbf{P}] = E[\lambda_1] - E[\lambda_{C_1}]$. We will demonstrate through numerical results that this approximation is valid for the upper tail of the distribution. Please note that we only approximate the *distribution* $F_{\lambda_1}(x)$ by $F_{\lambda_{C_1}}(x - \Delta)$. We cannot approximate the eigenvalues directly ($\lambda_1 \neq \lambda_{C_1} + \Delta$).

E. Distribution offset (Δ)

In order to compute the offset Δ , we will start from equation (36). Combining both parts, we get the quadratic equation

$$\lambda_1(\lambda_1 - \lambda_{C_1}) - \mathbf{R}_{12}\mathbf{R}_{22}\mathbf{P} - \|\mathbf{R}_{12}\|_2^2 = 0 \quad (13)$$

The term $\mathbf{R}_{12}\mathbf{R}_{22}\mathbf{P}$ poses a problem since it depends on the unknown vector \mathbf{P} . We will therefore apply two successive approximations to eliminate it: first, when the parameter N is large enough, the submatrix \mathbf{R}_{22} tends toward the identity matrix and the variable $\mathbf{R}_{12}\mathbf{R}_{22}\mathbf{P}$ can be approximated by $\mathbf{R}_{12}\mathbf{P}$. Second, we replace the random variable $\mathbf{R}_{12}\mathbf{P}$ by its expectation Δ . Applying this to (13) we obtain

$$\lambda_1(\lambda_1 - \lambda_{C_1}) - \Delta - \|\mathbf{R}_{12}\|_2^2 + \xi = 0, \quad (14)$$

where ξ is an error term proportional to the standard deviation of $\mathbf{R}_{12}\mathbf{P}$. The solution of this equation is given by

$$\lambda_1 = \frac{1}{2}\lambda_{C_1} + \frac{1}{2}\sqrt{\lambda_{C_1}^2 + 4(\Delta + \|\mathbf{R}_{12}\|_2^2 + \xi)} \quad (15)$$

We then apply the expectation operator $E[\cdot]$, which leads to the expression:

$$2E[\lambda_1] = E[\lambda_{C_1}] + \sqrt{E[\lambda_{C_1}^2] + 4\Delta + 4E[\|\mathbf{R}_{12}\|_2^2] + \xi_2}, \quad (16)$$

where the term ξ_2 accounts for ξ and the approximation error due to bringing the expectation operator inside the square root. Using the equality $E[\lambda_1] = E[\lambda_{C_1}] + \Delta$, we obtain a quadratic equation for Δ , whose solution

⁶Analyzing \mathbf{P} using equation (35) shows that its statistical distribution involves a multidimensional integral over a fairly complex path, making it intractable.

⁷If one looks at the explicit terms of the k^{th} iteration of equation (35) and neglects the remainder terms that still depend on \mathbf{P} , we can see that the variance of $\mathbf{R}_{12}\mathbf{P}$ has the same order of magnitude as $\|\mathbf{R}_{12}\|_2$ which, provided that the ratio β is large enough, is much smaller than λ_{C_1} .

(neglecting the error term) is given by

$$\Delta \simeq \frac{1}{2} \left((1 - \mu_{\lambda_{C_1}}) + \sqrt{(1 - \mu_{\lambda_{C_1}})^2 + \sigma_{\lambda_{C_1}}^2 + 4E[\|\mathbf{R}_{12}\|_2^2]} \right) F_{\Gamma}(x) = \Pr\left(\frac{\lambda_1}{\sum_{i=1}^p \lambda_i} \leq x\right) \quad (17)$$

where $\mu_{\lambda_{C_1}}$ and $\sigma_{\lambda_{C_1}}^2$ are respectively the mean and variance of λ_{C_1} . We will use equation (17) as an approximation for Δ . This approximation only depends on parameters that belong to the matrix \mathbf{C} or the entries of \mathbf{R} and therefore can be obtained analytically⁸. We will discuss in Section VI the errors resulting from the use of our approximation (17) (i.e. neglecting ξ_2) in deriving the statistical distribution of the SLE detector.

F. Summary

Let us write down the key results of this section:

- We built a circulant matrix (\mathbf{C}) to approximate a Toeplitz matrix ($\hat{\Sigma}_x$).
- We derived the statistical distribution of the largest eigenvalue λ_{C_1} of \mathbf{C} .
- Using the matrix perturbation theory, we obtained equation (12) relating the largest eigenvalue λ_1 of $\hat{\Sigma}_x$ to λ_{C_1} .
- We simplified equation (12) to remove the unknown variables, thereby postulating that the distribution of λ_1 can be approximated by the distribution of λ_{C_1} , up to an offset Δ .
- We derived an approximation for the offset Δ , which can be computed analytically.

By combining these results, we have modeled the statistical distribution of λ_1 as $F_{\lambda_1}(x - \Delta) = SLCT - N_{1,p-1}(\mathbf{B}\mu_r, \mathbf{B}\Sigma_r\mathbf{B}^T, \mathbf{0})$, indicating that the largest eigenvalue of $\hat{\Sigma}_x$ follows a Selection Normal distribution.

IV. STATISTICAL DISTRIBUTION OF Γ

We now move on to the next and main result of this paper, the statistical distribution of Γ . Using the result of the previous section, the statistical distribution $F_{\Gamma}(x)$

⁸For instance, the vector \mathbf{R}_{12} is obtained using a selection mechanism identical to (10): the position of the maximum element of $diag(\mathbf{R})$ dictates which off-diagonal elements are used to form \mathbf{R}_{12} . Thereby, \mathbf{R}_{12} follows a Selection Normal distribution whose parameters can be easily calculated. The Euclidean norm $\|\mathbf{R}_{12}\|_2$ is a quadratic form of a selection normal distribution and its variance can be calculated using the moment generating function of \mathbf{R}_{12} (28), or by direct integration using the density function (27).

can be approximated as

$$\begin{aligned} & \Pr\left(\frac{\lambda_1}{\sum_{i=1}^p \lambda_i} \leq x\right) \\ & \simeq \Pr(\lambda_{C_1} + \Delta \leq x \sum_{i=1}^p \lambda_{C_i}) \\ & = \Pr(\lambda_{C_1}(1-x) + \Delta - x \sum_{i=2}^p \lambda_{C_i} \leq 0) \\ & = \Pr(\max(\mathbf{r})(1-x) + \Delta - x \sum_{i=2}^p r_i \leq 0). \end{aligned} \quad (18)$$

Please note that $\sum_{i=1}^p \lambda_i = \sum_{i=1}^p \lambda_{C_i}$ due to the trace invariance property of our approximation (8). In the same fashion as equation (10), we define the random variables v and $\mathbf{u} = [u_2, \dots, u_p]^T$ through

$$\begin{pmatrix} u_2 \\ u_3 \\ \vdots \\ u_p \\ v \end{pmatrix} = \underbrace{\begin{pmatrix} -1 & 1 & 0 & \dots & 0 \\ -1 & 0 & 1 & \dots & 0 \\ -1 & 0 & 0 & \dots & 0 \\ \vdots & & \ddots & & \vdots \\ 1-x & -x & -x & \dots & -x \end{pmatrix}}_{\mathbf{B}_{\Gamma}} \begin{pmatrix} r_1 \\ r_2 \\ r_3 \\ \vdots \\ r_p \end{pmatrix} + \underbrace{\begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ \Delta \end{pmatrix}}_{\mathbf{\Delta}} \quad (19)$$

This allows us to rewrite equation (18) as

$$\begin{aligned} F_{\Gamma}(x) & \simeq p \Pr(v \leq 0, u_2 \leq 0, \dots, u_p \leq 0) \\ & = p \Phi_p(\mathbf{0}, \mathbf{B}_{\Gamma}\mu_r + \mathbf{\Delta}, \mathbf{B}_{\Gamma}\Sigma_r\mathbf{B}_{\Gamma}^T), \end{aligned} \quad (20)$$

where Φ_p is the standard Gaussian cumulative distribution function. This shows that the distribution $F_{\Gamma}(x)$ can be computed using a multivariate normal distribution. This result will allow us to set the SLE detector parameters to obtain a desired Probability of False-Alarm (P_{fa}) and to analyze its probability of detection without using MC simulations, for various noise and signal distributions.

A. SLCE detector

Interestingly, the theory developed so far can be used to obtain the distribution of a statistic similar to the SLE. In particular, we consider the statistic $\Gamma_C = \frac{\lambda_{C_1}}{\sum_{i=1}^p \lambda_{C_i}}$ derived using the eigenvalues of \mathbf{C} . We call the detector based on Γ_C the Scaled-Largest Circulant Eigenvalue (SLCE) detector. This detector is of interest since its statistic does not depend on the cumbersome approximation of the offset Δ . It is easy to see that its distribution is simply approximated by

$$F_{\Gamma_C}(x) \simeq p \Phi_p(\mathbf{0}, \mathbf{B}_{\Gamma}\mu_r, \mathbf{B}_{\Gamma}\Sigma_r\mathbf{B}_{\Gamma}^T). \quad (21)$$

Moreover, when the number of samples N is large

enough, the covariance matrix $\mathbf{B}_\Gamma \Sigma_r \mathbf{B}_\Gamma^T$ remains approximately constant over the domain of x , which shows that the two distributions $F_\Gamma(x)$ and $F_{\Gamma_C}(x)$ are identical, up to a shift in mean $\Delta_\Gamma \simeq \frac{\Delta}{p}$. Provided that the shift Δ_Γ remains constant between hypotheses H_0 and H_1 , the SLE and SLCE detectors have the same performance, but with the added benefit for the SLCE of a simpler distribution model.

V. SPATIAL SIGN FILTERING

To highlight the versatility of the presented methodology we consider its applicability to robust CR detectors. A detector is called robust if a slight modification of the noise distribution induces only a slight modification of the detector statistical distribution, thereby allowing us to maintain its threshold. For instance, CFAR detectors are robust to a change in the noise variance. As we previously mentioned, the SLE detector has been almost exclusively studied under the assumption that the noise follows a Gaussian distribution and is temporally uncorrelated [22], [8], [2]. As we will see, the SLE detector statistic is dramatically affected by any deviation from these assumptions. For instance, in the presence of impulsive noise, the sample covariance matrix does not follow a Wishart distribution anymore and the distribution of Γ is unknown, making it difficult to use the detector. Impulsive noise distributions occur frequently both in indoor and outdoor environments (see [23] and the references therein).

It has been shown on multiple occasions [23], [24], [25] that filtering the input signal using a non-linear sigmoid function, which constrains the signal between a lower and upper bound, improves the robustness of the detector with regards to impulsive noise by attenuating the influence of high amplitude samples. One such function is the spatial sign function

$$S(x) = \frac{x}{|x|}, \quad (22)$$

which extracts the phase of the received signal. We therefore wish to obtain the statistical distribution of the SLE detector when the received samples are first filtered by the spatial sign function.

Provided that the CLT remains valid, the vector $\hat{\zeta}$ (2) is still distributed as a multivariate normal random variable. The spatial sign function only affects its mean $\boldsymbol{\mu}_\zeta$ and covariance matrix Σ_ζ . Consequently, all the results regarding the statistical distribution of λ_1 and λ_{C1} remain valid. However, the main diagonal of the covariance matrix $\hat{\Sigma}_x$ becomes a constant (equal to 1) and the statistic Γ becomes identical to λ_1 , up to the rescaling factor $\frac{1}{p}$. In addition, the covariance matrix

$\hat{\Sigma}_x$ becomes singular, which subsequently affects the covariance matrix $\mathbf{B} \Sigma_r \mathbf{B}^T$ of the SLCT-N distribution (cf. equation (10)). However, as explained in [26], the consequences of this singularity are only of computational nature and the distribution $F_{\lambda_1}(x)$ still follows a SLCT-N distribution. As a result, our model can be used to approximate the distribution of the SLE detector when the received signal is first modified by the spatial sign function, which we will use to analyze the performance of the modified detector in the presence of impulsive noise.

VI. NUMERICAL RESULTS

In this section we present numerical results pertaining to the analytical results obtained so far. First, we will illustrate to what extent our approximations and models are valid by comparing the analytical results with Monte Carlo (MC) simulations. Second, we will compare the performances of the SLE and SLCE detectors for various parameters, noise distributions and signal distributions.

To clearly identify the underlying methods for the different results, the statistics based on our model incorporate a hat notation ($\hat{\cdot}$) ($\hat{\lambda}_{C1}$, $\hat{\lambda}_1$, etc), while the equivalent MC estimations do not. Furthermore, when presenting the validity of our model, we will also show results incorporating a MC estimation of the offset Δ . The statistics based on that estimation feature a \sim sign on top. For instance, we call $F_{\lambda_1}^\sim(x)$ the approximate distribution for λ_1 based on the analytical distribution of λ_{C1} and MC estimation of Δ . To simplify notation in the figures, we may write Γ , $\hat{\Gamma}$ and $\tilde{\Gamma}$ to designate the various *distributions* estimated using MC simulations, the theory previously developed, or a combination of both.

In order to test the robustness of the detectors with respect to deviations from the noise model assumptions, we will also show results incorporating non-Gaussian heavy-tailed noise, such Contaminated Gaussian (CG) noise, whose PDF is given by:

$$f_{CG}(x) = p_1 \mathcal{N}(x, 0, \sigma^2) + (1-p_1) \mathcal{N}(x, 0, r\sigma^2). \quad (23)$$

This is a two-component Gaussian mixture distribution. The background noise component has a high probability of occurrence (typically $p_1 = 95\%$) while the second component, controlling the impulsive part of the noise, is characterized by a low probability of occurrence ($1-p_1$) but possesses a variance that is r times higher than that of the background noise. Typical parameters used in this paper will be ($p_1 = 0.95, r = 100$).

A. λ_{C1} and λ_1 PDF approximations

Figure 1 shows the PDF of λ_{C1} and its approximation $\hat{\lambda}_{C1}$ based on the SLCT-N distribution. The

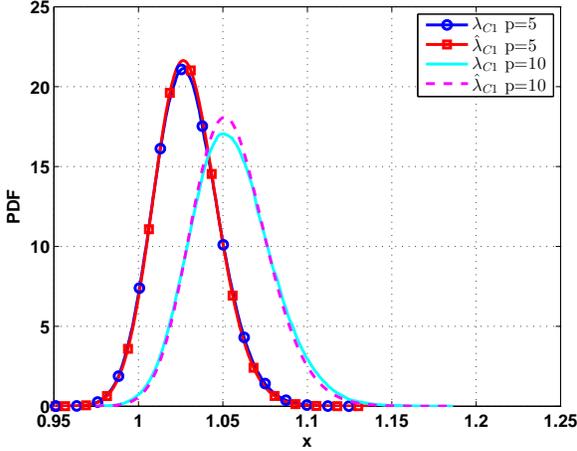


Fig. 1. PDF of λ_{C1} , obtained via Monte Carlo simulations, and its SLCT-N approximation $\hat{\lambda}_{C1}$. The input signal is AWGN noise, $N = 5000$.

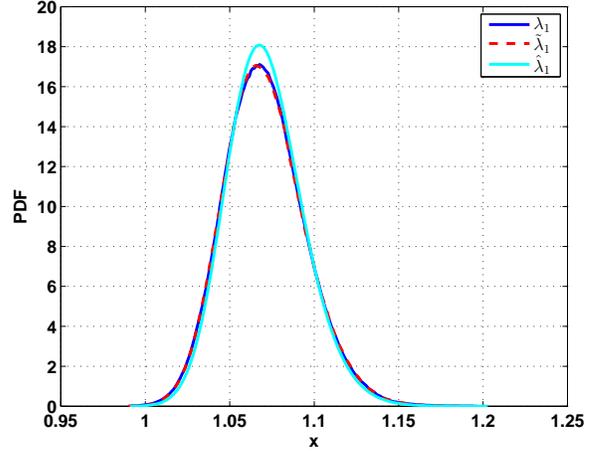


Fig. 2. PDF of λ_1 , $\tilde{\lambda}_1$ and $\hat{\lambda}_1$, the approximations $\hat{\lambda}_{C1}$ and $\hat{\Delta}$ are derived using the SLCT-N distribution and equ (17), respectively. The input signal is AWGN noise, $N = 5000$, and $p = 10$. $\Delta = 0.0154$, $\hat{\Delta} = 0.0164$

SLCT-N model is based on the assumption that the sample autocorrelation vector $\hat{\zeta}$ follows a multivariate normal distribution. This is an approximation based on the CLT and, as p increases, the overall error in the statistical model for $\hat{\zeta}$ increases, which in turn degrades the goodness-of-fit of the SLCT-N model. Increasing the number of samples available or using a Gaussian mixture model for the distribution of $\hat{\zeta}$ improves the goodness-of-fit. Nevertheless, we are mostly concerned with the upper tail behavior of our approximation, for which the error remains manageable as we will see by comparing deviations from the target P_{fa} for the SLE and SLCE detectors. We decided to use that metric as a way to illustrate the suitability of our model rather than, say, the Kullback-Leibler divergence, since it is more relevant for the parametrization of the detectors.

Figure 2 shows the PDF of λ_1 and models $\tilde{\lambda}_1$ and $\hat{\lambda}_1$. The PDF of $\tilde{\lambda}_1$ ($F_{\lambda_{C1}}(x)$ shifted by Δ) shows a remarkable match with the PDF of λ_1 , as does the resulting Cumulative Distribution Function (CDF) for $\tilde{\Gamma}$ shown in Figure 3. Despite being a mostly empirical observation, it indicates that a model accurately describing the PDF of λ_{C1} and the offset Δ can be used to model the distribution of λ_1 . The SLCT-N distribution approximation for $\hat{\lambda}_{C1}$ and our approximation $\hat{\Delta}$ attempt to fulfill this role. As can be seen on Figure 2, the approximation is valid for the upper quartile of the distribution, which in turn provides a good fit for the distribution function of $\hat{\Gamma}$ as shown in Figure 3.

The capabilities and limitations of our model are highlighted in Figure 4, where we show the deviations

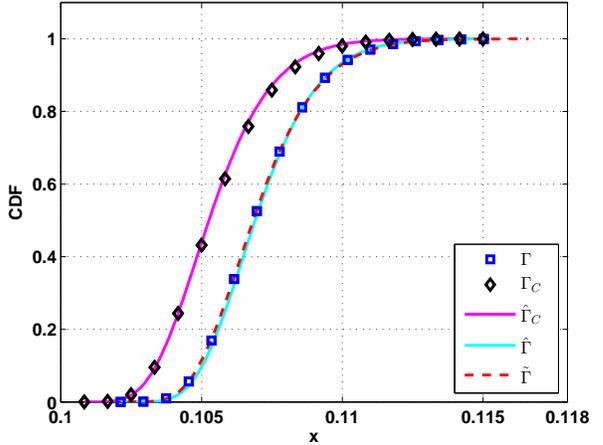


Fig. 3. CDF of the SLE and SLCE detectors statistics. The input signal is AWGN noise, $N = 5000$, and $p = 10$. $\Delta = 0.0154$, $\hat{\Delta} = 0.0164$

from the target P_{fa} caused by our model. For instance, if the desired P_{fa} is 0.01, simulations show that the real P_{fa} will have an offset of about 15%, bringing its value to 0.0115. The relative error goes down when the target P_{fa} increases and the error appears to be reasonable for typical P_{fa} values such as 0.1. In the particular case of Figure 4, the distribution $\tilde{\Gamma}$ is slightly worse than $\hat{\Gamma}$ in the range of interest ($0.01 \leq P_{fa} \leq 0.5$), but this is not generally true, as can be seen in Table I, which shows the P_{fa} relative error for various values of the parameters N and p . These errors are caused by deviation from the

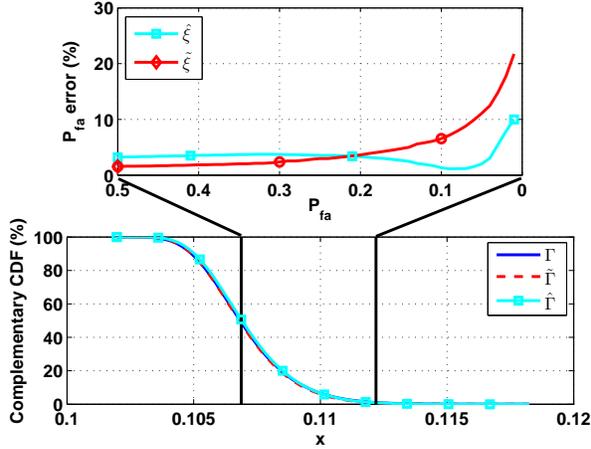


Fig. 4. Absolute value of the P_{fa} relative error for the SLE detector parametrized using the statistics $\hat{\Gamma}$ and $\tilde{\Gamma}$. The error is expressed in % compared to the P_{fa} obtained using MC simulations (i.e. the statistic Γ): $\tilde{\xi} = 100 \frac{|F_{\tilde{\Gamma}}(x) - F_{\Gamma}(x)|}{1 - F_{\tilde{\Gamma}}(x)}$. The input signal is AWGN noise, $N = 5000$, and $p = 10$.

multivariate Gaussian distribution of the vector $\hat{\zeta}$. The pdf of $\hat{\zeta}$ relies on the CLT and therefore the model is not accurate for small values of N . Moreover, for a fixed N (and thereby constant deviation from the Gaussian model for every element $\hat{\zeta}_i$), increasing the parameter p increases the dimension of $\hat{\zeta}$ and the overall error caused by the CLT approximation. The error caused by our model also appears in Figure 5, which shows the probability of detection of the SLE detector in the presence of impulsive noise. As the SNR decreases the probability of detection becomes equivalent to the P_{fa} and we can readily see that the real P_{fa} slightly differs from the target P_{fa} .

B. SLCE Detector and spatial sign filtering

The SLCE detector statistic does not rely on an estimation of the offset Δ , thereby removing a source of error in the model, which should reduce any deviation from the desired P_{fa} . For instance, Figure 6 shows that the P_{fa} for the SLCE detector is equal to: 0.108 instead of 0.13 for the SLE detector (as shown on Figure 5). Additionally, as we previously mentioned in Section IV, the distribution Γ and Γ_C are almost identical, apart from the offset Δ_{Γ} and the performance of both detectors should therefore be similar. This is illustrated in Figure 7, which shows that the SLCE detector has almost identical performance compared to the SLE detector, making it an interesting alternative to the SLE detector.

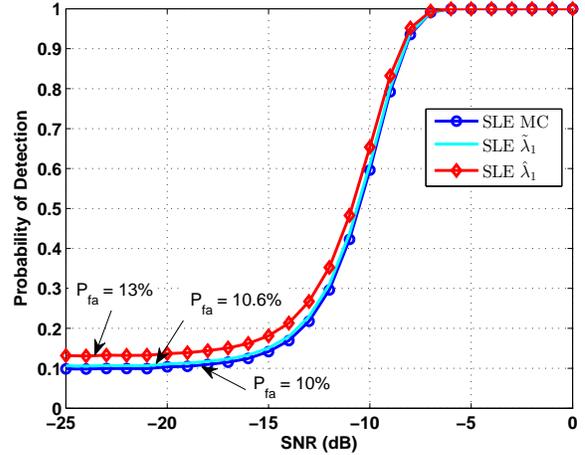


Fig. 5. Probability of detection for the SLE detector based on the statistics Γ , $\hat{\Gamma}$ and $\tilde{\Gamma}$. Target $P_{fa} = 10\%$, $N = 5000$, and $p = 10$. The input signal is a 64 Quadrature Amplitude Modulated (QAM) signal filtered by a root-raised cosine filter and embedded in CG noise (0.95, 100). The oversampling factor is set to 2.

The spatial sign non-linearity makes the SLE and SLCE detectors impervious to changes in amplitude of the received signal, thereby increasing their robustness to non-Gaussian noise signals. As shown in Figure 7, both detectors greatly benefit from the presence of the spatial sign filtering in the presence of impulsive noise, with an increase of about 6 dB in SNR⁹ for an identical probability of detection. Moreover, an added benefit of the spatial sign filter comes from the hardware simplification that occurs by working with a signal of fixed amplitude. Since our analytical model only requires the CLT approximation to be valid for the vector $\hat{\zeta}$, we can use our model to derive the SLE statistic distribution even when the input signal is modified by the spatial sign function (cf. Figure 6). This contrasts with the usual Wishart-based analysis, which cannot be used when the input signal deviates from the AWGN assumption.

C. Spatial Covariance matrices

While the theory developed in Sections III and IV originates from temporal covariance matrices of wide-sense stationary signals, the SLE detectors is often used in multi-antenna systems that work with spatial covariance matrices, exploiting the correlation between antennas. Therefore, given the importance of this fairly practical scenario, we present a set of results regarding

⁹The SNR for CG noise is defined using the overall variance of the noise, including the impulsive part.

TABLE I
 P_{fa} RELATIVE ERROR (%) FOR THE STATISTICS $\tilde{\Gamma}$, $\hat{\Gamma}$ AND Γ_C FOR VARIOUS VALUES OF THE PARAMETERS N AND p . THE TARGET P_{fa} IS 0.1. AWGN.

	$N = 500$		$N = 1000$			$N = 5000$			
	$p = 5$	$p = 10$	$p = 5$	$p = 10$	$p = 15$	$p = 5$	$p = 10$	$p = 15$	$p = 20$
$\tilde{\Gamma}$	18.3	31	7.1	21.8	45.7	0.1	6.9	13.8	22.8
$\hat{\Gamma}$	19.2	62.5	20.5	29.6	56.0	18.8	0.4	21.0	38.3
$\hat{\Gamma}_C$	17.4	38.8	10.4	29.1	48.9	3.9	9.3	19.2	27.2

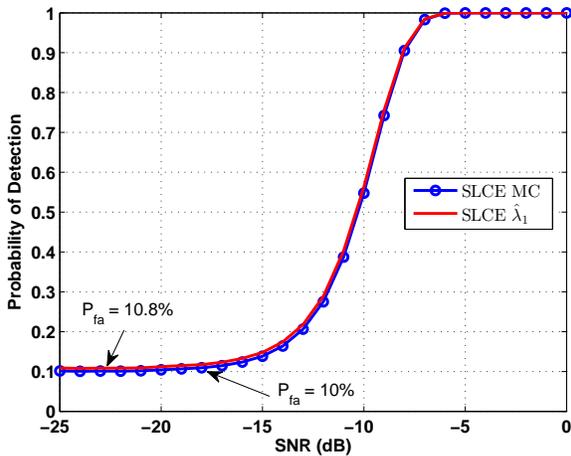


Fig. 6. Probability of detection for the SLCE detector, parametrized using MC simulations and the SLCT-N approximation. Target $P_{fa} = 10\%$, $N = 5000$, and $p = 10$, 64QAM signal embedded in CG noise (0.95, 100).

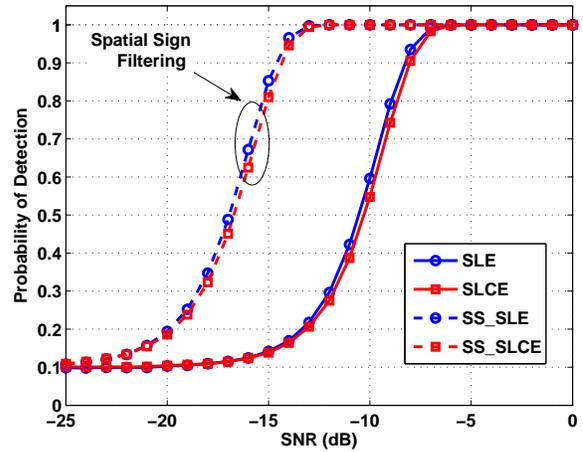


Fig. 7. Probability of detection for the SLE and SLCE detectors, in the presence of impulsive noise, with and without spatial sign filtering (Prefix "SS" in the legend) $N = 5000$, and $p = 10$, CG noise (0.95, 100).

the distribution of the SLE and SLCE detectors exploiting spatial covariance matrices¹⁰.

Spatial covariance matrices do not have a Toeplitz structure. Nevertheless, under the null hypothesis, both the spatial and temporal sample covariance matrices take on the form of perturbed identity matrices (provided that the noise power is identical at all antennas), and it is reasonable to expect that the theory derived for temporal covariance matrices could be reused for spatial covariance matrices (only for the purpose of tuning the detector for the null hypothesis).

Figure 8 shows the CDF of the SLE and SLCE statistics for several noise distributions. For AWGN noise, the empirical (MC) distribution is closely approximated

¹⁰Please note that, despite their empirical nature, we will denote by "theoretical results" the ones obtained using the theory developed for the temporal covariance matrices, as opposed to the "empirical results" obtained through MC simulations.

by the theoretical result of equation (20). Next, we add random noise variance fluctuations at the antennas: the noise variance at each antenna follows a uniform distribution centered around unity with a maximum deviation of 1 dB from unity. When the noise variance is not identical across all antennas, we can readily see that the SLE distribution is drastically different from the AWGN scenario, indicating that the detector is not robust with respect to uncalibrated noise and that the theory based on Wishart distribution cannot be used to parametrize the detector in such conditions. Unfortunately, our model also fails to approximate the empirical distribution. However, when the received signal is first modified by the spatial sign function, the distribution of the SLE becomes similar to the AWGN scenario and the theoretical CDF

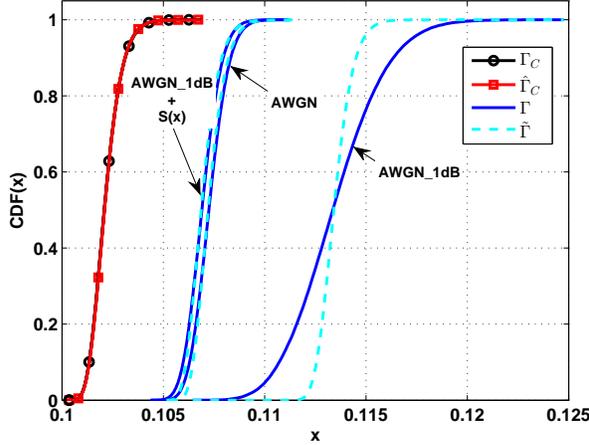


Fig. 8. Theoretical and empirical CDF of the SLE and SLCE detectors computed for spatial covariance matrices, in the presence of AWGN, AWGN uncalibrated among the different antennas (1 dB uncertainty, denoted as *AWGN_1dB*) and *AWGN_1dB* modified by the spatial sign function (*AWGN_1dB* + $S(x)$). The CDF of Γ_C and $\hat{\Gamma}_C$, remain identical for all scenarios and the curves are superimposed (leftmost curve). $p = 10$, $N = 5000$

matches the empirical one¹¹. The SLCE statistic remains unchanged for all scenarios, indicating an inherent robustness to the noise distribution/calibration. Indeed, the lack of noise variance calibration only results in a scaling effect for the entries of the vector \mathbf{r} , which disappears in the SLCE statistic. The theoretical approximation offers a good fit, making it possible to use the SLCE detector with spatial covariance matrices, even when subjected to noise variance uncertainties.

D. Probability of Detection

The analysis leading to expression (20) has been focused on the statistical distribution of Γ under hypothesis H_0 , in order to set a detector threshold matching a target P_{fa} . It is however entirely possible to use the methodology to derive the distribution under the alternative hypothesis H_1 , provided that the presence of a communication signal only modifies the parameters (i.e. $\mu_{\hat{\xi}}$ and $\Sigma_{\hat{\xi}}$) of the estimated covariance vector $\hat{\xi}$ (2) (the joint multivariate Gaussian assumption has to remain valid). For instance, Figure 9 illustrates the CDF of the SLE and SLCE detectors for an Orthogonal Frequency-Division Multiplexing (OFDM) signal¹² embedded in

¹¹The offset Δ is derived using MC simulations as its approximation (17) is only valid for Toeplitz matrices. We did not derive an approximation suitable for spatial covariance matrices.

¹²We could have used other signals such as the QAM signals used earlier.

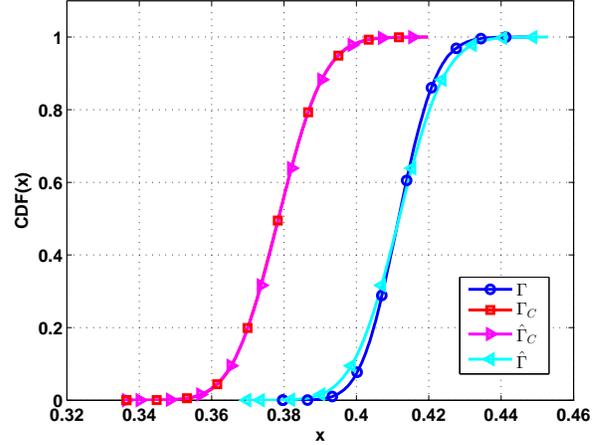


Fig. 9. Theoretical and empirical CDF of the SLE and SLCE detectors computed under hypothesis H_1 . The received signal consist of AWGN noise and an OFDM signal (SNR= -5 dB), filtered using the spatial sign function. $p = 10$, $N = 5000$

AWGN noise. We can see that the CDF approximation remains valid for both the SLCE and SLE statistics.

E. Comparison with the Ljung-Box detector and sphericity test

In testing for the presence of autocorrelation of a Gaussian time-series, whenever the noise variance or signal covariance matrix is known, it is possible to derive a GLRT statistic [27]. When neither the signal covariance matrix nor the noise variance are known, computing a GLRT statistic is impractical due to the amount of unknown parameters and in any case requires a Gaussian distribution for the received signal. For such detection scenarios, the Ljung-Box detector [9], [28] is often used even though it is not optimal. In its simplest form, the Ljung-Box detector sums the square of the serial correlation coefficients normalized by the variance of the signal:

$$\Gamma_{LB} = N(N+2) \sum_{i=1}^p \frac{1}{N-i} \left(\frac{\hat{\zeta}_i}{\hat{\zeta}_0} \right)^2 \quad (24)$$

While the SLE detector is a GLRT for the detection of a rank-one multivariate Gaussian signal, it also maximizes the SNR among all statistics that linearly combine the received signal samples [6], motivating its use in single-antenna detectors. Therefore, it is interesting to compare it to the Ljung-Box detector.

Another eigenvalue detector that is worth comparing to the SLE is the sphericity, also known as the Spherical Test Method (STM detector). The sphericity test is a

GLRT for the detection of multivariate Gaussian signals with positive-definite covariance matrices (hence full-rank) in the presence of i.i.d Gaussian noise [2], [5]. Its statistic is given by

$$\Gamma_{sph} = \frac{\frac{1}{p} \text{tr}(\hat{\Sigma}_x)}{\det(\hat{\Sigma}_x)^{1/p}}, \quad (25)$$

where $\text{tr}(\mathbf{A})$ and $\det(\mathbf{A})$ are the trace and determinant of the matrix \mathbf{A} , respectively. In a blind-detection scenario, where the structure of Σ_x is unknown, one may expect for the sphericity test to outperform the SLE detector. However, the covariance matrix $\hat{\Sigma}_x$ tends toward a rank-one matrix as the oversampling rate of the received signal increases and warrants a comparison between the two detectors for small to moderate values of the oversampling factor. The oversampling factor has been set to 2 for all previous results to guarantee some amount of correlation, irrespective of the signal or channel parameters.

Figures 10 and 11 show the Receiver Operating Characteristic (ROC) curves for the SLE, Sphericity and Ljung-Box detectors for different signals and size of the covariance matrix. For the typical lag values used in the paper ($p \in \{5, 20\}$) and a small oversampling factor (2, 3), the SLE and sphericity detector have similar performances. When the oversampling factor increases (Figure 11), the SLE detector outperforms the Sphericity detector. However, when the oversampling increases, the overall SNR decreases due to the added out-of-band noise collected, resulting in a degradation of performance for the three detectors. As a result, an oversampling factor as small as possible is always desirable and when no oversampling is needed (e.g. an OFDM signal with guard bands), the Sphericity test outperforms the SLE detector.

VII. CONCLUSION

In this paper, we have been concerned with the analysis of the SLE detector statistic Γ . The objective was to develop an analytical model for its statistical distribution when the received signal is wide-sense stationary. We considered small matrix sizes ($5 \leq p \leq 20$) and large sample sizes ($500 \leq N \leq 5000$). The model we developed relies on a CLT assumption and therefore requires enough samples to be used. However, the CLT assumption also makes it applicable to a wide range of noise and signal distributions, unlike previous models found in the literature. Furthermore, it allowed us to consider non-linear transformations of the signals, (e.g. spatial sign filtering) to improve the detector robustness to impulsive noise. We also presented a new detector, the SLCE detector, based on the largest eigenvalue of

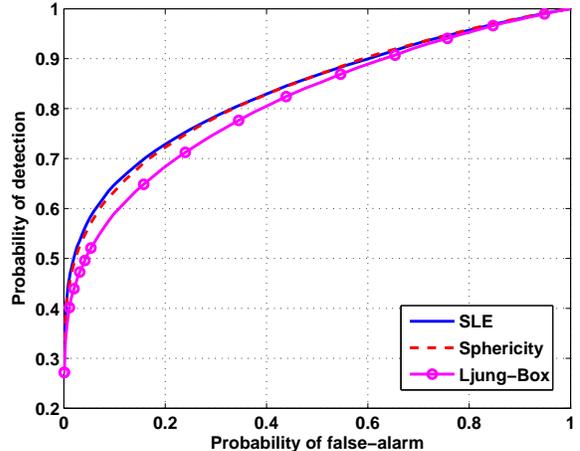


Fig. 10. Receiver Operating Characteristic (ROC) curves for the SLE, Sphericity and Ljung-Box detectors. 64 Quadrature Amplitude Modulated (QAM) signal in AWGN. $SNR = -10$ dB, $p = 10$, $N = 5000$, oversampling factor = 3

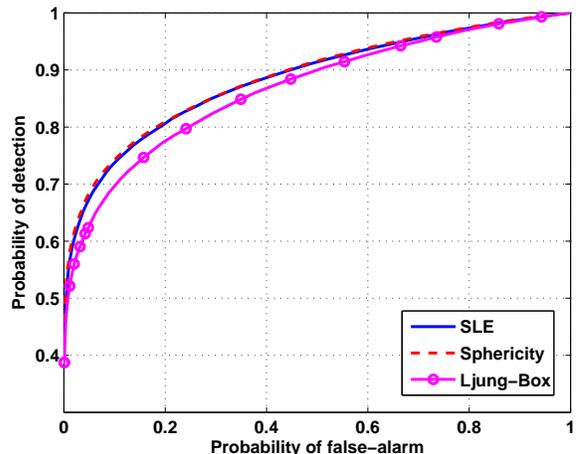


Fig. 11. ROC curves for the SLE, Sphericity and Ljung-Box detectors. OFDM signal in AWGN. $SNR = -15$ dB, $p = 20$, $N = 5000$, oversampling factor = 2

a circulant matrix approximating the sample covariance matrix. Our motivation regarding the SLCE detector resides in its competitiveness with the SLE detector as well as the simplified closed-form solution for its statistical distribution.

APPENDIX

A. Selection Normal Distribution

We introduce here the selection normal distribution (also called SUN distribution), as presented in [26], [29].

Given two multivariate normal random variables \mathbf{U} and \mathbf{V} , i.e.

$$\begin{pmatrix} \mathbf{U} \\ \mathbf{V} \end{pmatrix} \sim \mathcal{N}_{q+p} \left(\boldsymbol{\xi} = \begin{bmatrix} \boldsymbol{\xi}_U \\ \boldsymbol{\xi}_V \end{bmatrix}, \boldsymbol{\Omega} = \begin{bmatrix} \boldsymbol{\Omega}_U & \Delta^T \\ \Delta & \boldsymbol{\Omega}_V \end{bmatrix} \right), \quad (26)$$

the selection normal distribution $SLCT - N_{p,q}(C, \boldsymbol{\xi}, \boldsymbol{\Omega})$ corresponds to the distribution of the random variable \mathbf{V} when the variable \mathbf{U} is constrained to a domain C : $\Pr(\mathbf{V} \leq \mathbf{x} | \mathbf{U} \in C)$. Its density function is given by

$$f_{\mathbf{V}}(\mathbf{x}) = \phi_p(\mathbf{x}, \boldsymbol{\xi}_V, \boldsymbol{\Omega}_V) \times \frac{\Phi_q(C, \Delta^T \boldsymbol{\Omega}_V^{-1}(\mathbf{x} - \boldsymbol{\xi}_V) + \boldsymbol{\xi}_U, \boldsymbol{\Omega}_U - \Delta^T \boldsymbol{\Omega}_V^{-1} \Delta)}{\Phi_q(C, \boldsymbol{\xi}_U, \boldsymbol{\Omega}_U)}, \quad (27)$$

where ϕ_p and Φ_p represent the p-variate Gaussian density and distribution functions, respectively.

The moment generating function is given by

$$M_{\mathbf{V}}(\mathbf{x}) = e^{\{\mathbf{x}^T \boldsymbol{\xi}_V + \frac{1}{2} \mathbf{x}^T \boldsymbol{\Omega}_V \mathbf{x}\}} \frac{\Phi_q(C, \Delta^T \mathbf{x} + \boldsymbol{\xi}_U, \boldsymbol{\Omega}_U)}{\Phi_q(C, \boldsymbol{\xi}_U, \boldsymbol{\Omega}_U)}. \quad (28)$$

B. Matrix perturbation theory

It is well known that eigenvalues of a matrix are a continuous function of the entries of that matrix. Given a $p \times p$ matrix \mathbf{A} with known eigenvalues and eigenvectors, the matrix perturbation theory presented in [15], [16], [17] allows us to derive the eigenvalues of a matrix $\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{E}$, where \mathbf{E} is called the perturbation matrix. In the following, \mathbf{A} and \mathbf{E} are Hermitian. The notations are consistent with the ones used in [15].

We call \mathbf{L}_1 the diagonal matrix formed by $k \leq p$ eigenvalues of \mathbf{A} and \mathbf{L}_2 the diagonal matrix formed by the remaining eigenvalues of \mathbf{A} . \mathbf{X}_1 and \mathbf{X}_2 are matrices made of the eigenvectors corresponding to the eigenvalues of \mathbf{L}_1 and \mathbf{L}_2 , respectively. We can then write

$$(\mathbf{X}_1 \quad \mathbf{X}_2)^H \mathbf{A} (\mathbf{X}_1 \quad \mathbf{X}_2) = \begin{pmatrix} \mathbf{L}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{L}_2 \end{pmatrix}. \quad (29)$$

Given a perturbation \mathbf{E} , we have

$$(\mathbf{X}_1 \quad \mathbf{X}_2)^H \mathbf{E} (\mathbf{X}_1 \quad \mathbf{X}_2) = \begin{pmatrix} \mathbf{E}_{11} & \mathbf{E}_{12} \\ \mathbf{E}_{21} & \mathbf{E}_{22} \end{pmatrix}. \quad (30)$$

We define the following parameters:

$$\begin{aligned} \tilde{\gamma} &= \|\mathbf{E}_{21}\|_F \\ \tilde{\delta} &= sep_F(\mathbf{L}_1, \mathbf{L}_2) - \|\mathbf{E}_{11}\|_F - \|\mathbf{E}_{22}\|_F \\ &= \min |\mathcal{L}(\mathbf{L}_1) - \mathcal{L}(\mathbf{L}_2)| - \|\mathbf{E}_{11}\|_F - \|\mathbf{E}_{22}\|_F, \end{aligned} \quad (31)$$

where $\|\cdot\|_F$ is the Frobenius norm and $\mathcal{L}(\mathbf{A})$ is the set of the eigenvalues of \mathbf{A} . The separation function

$sep_F(\cdot, \cdot)$ is defined in [15] (p. 231) and is equivalent to $\min |\mathcal{L}(\cdot) - \mathcal{L}(\cdot)|$ only for Hermitian matrices.

As was demonstrated in [15], if $\frac{\tilde{\gamma}}{\tilde{\delta}} < \frac{1}{2}$ then there is a unique matrix \mathbf{P} , solution of the nonlinear equation

$$\mathbf{P}(\mathbf{L}_1 + \mathbf{E}_{11}) - (\mathbf{L}_2 + \mathbf{E}_{22})\mathbf{P} = \mathbf{E}_{21} - \mathbf{P}\mathbf{E}_{12}\mathbf{P}, \quad (32)$$

such that

$$\tilde{\mathbf{L}}_1 = \mathbf{L}_1 + \mathbf{E}_{11} + \mathbf{E}_{12}\mathbf{P}. \quad (33)$$

The eigenvalues of $\tilde{\mathbf{L}}_1$ are the eigenvalues of the perturbed matrix $\tilde{\mathbf{A}}$ corresponding to the eigenvalues \mathbf{L}_1 , in the sense that $\mathcal{L}(\tilde{\mathbf{L}}_1) \rightarrow \mathcal{L}(\mathbf{L}_1)$ if $\mathbf{E} \rightarrow \mathbf{0}$. In this paper, our interest lies with $\tilde{\mathbf{L}}_1$ and equation (33) provides the link we need between the known eigenvalues \mathbf{L}_1 and the unknown eigenvalues $\tilde{\mathbf{L}}_1$. We are only interested in the largest eigenvalue of $\tilde{\mathbf{A}}$ and as a result, $k = 1$, \mathbf{L}_1 and \mathbf{E}_{11} are scalars and the matrix \mathbf{P} is a vector¹³ of size $(p-1)$. This leads to the following simplification of equation (32):

$$\mathbf{P} = \mathbf{B}[\mathbf{E}_{21} - \mathbf{E}_{11}\mathbf{P} + \mathbf{E}_{22}\mathbf{P} - \mathbf{P}\mathbf{E}_{12}\mathbf{P}], \quad (34)$$

where $\mathbf{B} = (\mathbf{L}_1 \mathbf{I}_{p-1} - \mathbf{L}_2)^{-1}$.

Provided that $\frac{\tilde{\gamma}}{\tilde{\delta}} < \frac{1}{2}$, equation (34) can be solved iteratively:

$$\begin{cases} \mathbf{P}_0 &= \mathbf{0} \\ \mathbf{P}_{k+1} &= \mathbf{B}[\mathbf{E}_{21} - \mathbf{E}_{11}\mathbf{P}_k + \mathbf{E}_{22}\mathbf{P}_k - \mathbf{P}_k\mathbf{E}_{12}\mathbf{P}_k]. \end{cases} \quad (35)$$

Additionally, if $\mathbf{A} = \text{diag}(\tilde{\mathbf{A}})$, \mathbf{P} can be linked directly to $\tilde{\mathbf{X}}_1$ (the eigenvector of $\tilde{\mathbf{L}}_1$). We denote the position of the maximum element of $\text{diag}(\mathbf{A})$ with the subscript M and \tilde{x}_M is the M^{th} element of the vector $\tilde{\mathbf{X}}_1 = [\tilde{x}_1 \quad \dots \quad \tilde{x}_p]$. Then, using basic properties of eigenvalues, we can readily obtain the following system of equations (we suppose without loss of generality that $M = 1$):

$$\begin{aligned} \tilde{\mathbf{A}}\tilde{\mathbf{X}}_1 &= \tilde{\mathbf{L}}_1\tilde{\mathbf{X}}_1 \\ \tilde{\mathbf{A}}\left(\frac{1}{\tilde{x}_M}\right)\tilde{\mathbf{X}}_1 &= \tilde{\mathbf{L}}_1\left(\frac{1}{\tilde{x}_M}\right)\tilde{\mathbf{X}}_1 \\ \begin{pmatrix} \mathbf{L}_1 & \mathbf{E}_{12} \\ \mathbf{E}_{21} & \mathbf{L}_2 + \mathbf{E}_{22} \end{pmatrix} \begin{bmatrix} 1 \\ \mathbf{P}_1 \\ \vdots \\ \mathbf{P}_{p-1} \end{bmatrix} &= \tilde{\mathbf{L}}_1 \begin{bmatrix} 1 \\ \mathbf{P} \end{bmatrix} \\ \begin{cases} \mathbf{L}_1 + \mathbf{E}_{12}\mathbf{P} &= \tilde{\mathbf{L}}_1 \\ \mathbf{E}_{21} + (\mathbf{L}_2 + \mathbf{E}_{22})\mathbf{P} &= \tilde{\mathbf{L}}_1\mathbf{P} \end{cases} \end{aligned} \quad (36)$$

This link between $\tilde{\mathbf{X}}_1$ and \mathbf{P} proves very useful in computer simulations to assess the convergence of equation (35).

¹³We will nonetheless keep the matrix notation to avoid an unnecessary change of notation.

Applying the theory to the matrices defined in equation (8), we see that $\text{diag}(\mathbf{R})$ is our original matrix \mathbf{A} and the perturbation matrix is defined as $\mathbf{E} = \mathbf{R} - \text{diag}(\mathbf{R})$. Since our unperturbed matrix is diagonal, its eigenvectors are unit vectors $\mathbf{e}_i = [0 \dots 1 \dots 0]$.

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 Magazine*, vol. 29, pp. 101–116, May 2012.
- [2] 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.
- [3] D. Ramirez, G. Vazquez-Vilar, R. Lopez-Valcarce, J. Via, and I. Santamaria, "Multiantenna detection under noise uncertainty and primary user's spatial structure," in *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 2948–2951, May 2011.
- [4] G. Vazquez-Vilar, R. Lopez-Valcarce, and J. Sala, "Multi-antenna Spectrum Sensing Exploiting Spectral a priori Information," *IEEE Transactions on Wireless Communications*, vol. 10, pp. 4345–4355, Dec. 2011.
- [5] 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.
- [6] 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.
- [7] P. Bianchi, M. Debbah, M. Maida, and J. Najim, "Performance of statistical tests for single-source detection using random matrix theory," *IEEE Transactions on Information Theory*, vol. 57, pp. 2400–2419, April 2011.
- [8] L. Wei, O. Tirkkonen, P. Dharmawansa, and M. McKay, "On the exact distribution of the scaled largest eigenvalue," in *2012 IEEE International Conference on Communications (ICC)*, pp. 2422–2426, 2012.
- [9] G. M. Ljung and G. E. P. Box, "On a measure of lack of fit in time series models," *Biometrika*, vol. 65, no. 2, pp. 297–303, 1978.
- [10] T. Anderson, *An Introduction to Multivariate Statistical Analysis, 3rd Edition*. Wiley series in probability and mathematical statistics, 2003.
- [11] P. Wirfalt and M. Jansson, "On kronecker and linearly structured covariance matrix estimation," *IEEE Transactions on Signal Processing*, vol. 62, pp. 1536–1547, Mar. 2014.
- [12] H. Li, P. Stoica, and J. Li, "Computationally efficient maximum likelihood estimation of structured covariance matrices," *IEEE Trans. Signal Processing*, vol. 47, pp. 1314–1323, May 1999.
- [13] Y. Zeng and Y.-C. Liang, "Spectrum-Sensing Algorithms for Cognitive Radio Based on Statistical Covariances," *IEEE Transactions on Vehicular Technology*, vol. 58, pp. 1804–1815, May 2009.
- [14] A. M. Mathai and S. B. Provost, *Quadratic Forms in Random Variables, Theory and Applications*. Marcel Dekker, New York, 1992.
- [15] G. W. Stewart and J. Sun, *Matrix Perturbation Theory*. Academic Press, 1990.
- [16] B. Ames, "Taylor series expansions of the spectrum of a symmetric matrix," Master's thesis, University of Guelph, Apr. 2007.
- [17] B. P. Ames and H. S. Sendov, "Asymptotic expansions of the ordered spectrum of symmetric matrices," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 72, pp. 4288 – 4297, June 2010.
- [18] R. M. Gray, "Toeplitz and circulant matrices: A review," *Foundations and Trends in Communications and Information Theory*, vol. 2, no. 3, pp. 155–239, 2006.
- [19] R. Gray, "On the asymptotic eigenvalue distribution of toeplitz matrices," *IEEE Trans. Inform. Theory*, vol. 18, no. 6, pp. 725–730, 1972.
- [20] J. Pearl, "On coding and filtering stationary signals by discrete fourier transforms (corresp.)," *IEEE Transactions on Information Theory*, vol. 19, pp. 229–232, Mar. 1973.
- [21] J. Kierzkowski and A. Smoktunowicz, "Block normal matrices and gershgorin-type discs," *Electronic Journal of Linear Algebra*, vol. 22, pp. 1053–1063, 2011.
- [22] A. Kortum, M. Sellathurai, T. Ratnarajah, and C. Zhong, "Distribution of the ratio of the largest eigenvalue to the trace of complex wishart matrices," *IEEE Trans. Signal Processing*, vol. 60, no. 10, pp. 5527–5532, 2012.
- [23] 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.
- [24] G. Jacovitti and A. Neri, "Estimation of the autocorrelation function of complex Gaussian stationary processes by amplitude clipped signals," *IEEE Trans. Inform. Theory*, vol. 40, pp. 239 –245, Jan. 1994.
- [25] T. Biedka, L. Mili, and J. Reed, "Robust estimation of cyclic correlation in contaminated Gaussian noise," *29th Asilomar Conference on Signals, Systems and Computers*, p. 511, 1995.
- [26] R. B. Arellano-Valle, M. D. Branco, and M. G. Genton, "A unified view on skewed distributions arising from selections," *Canadian Journal of Statistics*, vol. 34, no. 4, pp. 581–601, 2006.
- [27] S. Kay, *Fundamentals of Statistical Signal Processing: Vol II Detection Theory*. Prentice-Hall PTR, 1998.
- [28] G. M. Goerg, "Testing for white noise against locally stationary alternatives," *Statistical Analysis and Data Mining*, vol. 5, pp. 478 – 492, 2012.
- [29] R. B. Arellano-Valle and A. Azzalini, "On the unification of families of skew-normal distributions," *Scandinavian Journal of Statistics*, vol. 33, no. 3, pp. 561–574, 2006.



Julien Renard (J'10) received the M.Sc. (Eng.) degree with High Distinction in electrical engineering from the Applied Science Faculty of the Université 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 SmartGrid-Comm. 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. In 2002, he joined the Inter-university Micro-Electronics Center (IMEC), Leuven, Belgium. He led the project aiming at developing a fourth-generation wireless communication system in

collaboration with Samsung Korea. In 2007, Francois Horlin became associate professor at the Universit libre de Bruxelles (ULB), Brussels, Belgium. Since 2014, he is full professor. 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 is author of a book, author of a book chapter, co-author of two patents, author or co-author of more than 180 publications in well-recognized journals and conferences. He chairs currently the IEEE signal processing chapter of the Benelux.