Detection of Spatially Correlated Gaussian Time Series

David Ramírez, *Student Member, IEEE*, Javier Vía, *Member, IEEE*, Ignacio Santamaría, *Senior Member, IEEE*, and Louis L. Scharf, *Life Fellow, IEEE*

***Abstract*—This work addresses the problem of deciding whether a set of realizations of a vector-valued time series with unknown temporal correlation are spatially correlated or not. For wide sense stationary (WSS) Gaussian processes, this is a problem of deciding between two different power spectral density matrices, one of them diagonal. Specifically, we show that for arbitrary Gaussian processes (not necessarily WSS) the generalized likelihood ratio test (GLRT) is given by the quotient between the determinant of the sample space–time covariance matrix and the determinant of its block-diagonal version. Furthermore, for WSS processes, we present an asymptotic frequency-domain approximation of the GLRT which is given by a function of the Hadamard ratio (quotient between the determinant of a matrix and the product of the elements of the main diagonal) of the estimated power spectral density matrix. The Hadamard ratio is known to be the GLRT detector for vector-valued random variables and, therefore, what this paper shows is how frequency-dependent Hadamard ratios must be merged into a single test statistic when the vector-valued random variable is replaced by a vector-valued time series with temporal correlation. For bivariate time series, the derived frequency domain detector can be rewritten as a function of the well-known magnitude squared coherence (MSC) spectrum, which suggests a straightforward extension of the MSC spectrum to the general case of multivariate time series. Finally, the performance of the proposed method is illustrated by means of simulations.**

***Index Terms*—Coherence spectrum, generalized likelihood ratio test (GLRT), Hadamard ratio, multiple-channel signal detection, power spectral density matrix.**

I. Introduction

THE multiple-channel signal detection problem appears in many applications, such as sensor networks [1], cooperative networks with multiple relays using the amplify-and-forward (AF) scheme [2]–[4], or radar detection with multiple antennas [5]. It is also an important problem in cognitive radio, when a secondary (non-licensed) user equipped with multiple antennas needs to sense the channel to detect whether a specific frequency sub-band is occupied or not [6]–[9]. To solve the multiple-channel signal detection problem in its most general formulation, we should exploit the fact that, under the null hypothesis, the signal is spatially uncorrelated. In the case of complex circular Gaussian processes, the general detection problem consists in deciding whether the space–time covariance matrix (a matrix which contains all space–time second order statistics of the vector-valued time series) is block diagonal (null cross second order statistics) or not. However, in the particular case of wide sense stationary processes, the problem can be seen as that of deciding between two different power spectral density (PSD) matrices,[1] one of them (representing the null hypothesis) diagonal.

The problem of multiple-channel signal detection has been addressed in [10] and [11] for vector-valued random variables, where the authors proposed a new measure called the generalized coherence (GC). The derivation of the GC is based on a geometrical interpretation of the correlation coefficient between two random variables, which can readily be generalized to random vectors. In [12] and [13], Leshem and van der Veen have derived the generalized likelihood ratio test (GLRT) for detecting the presence of an unknown white Gaussian signal acquired by a set of sensors. In fact, for Gaussian random vectors the GC and the GLRT result in the same detector, which is given by the Hadamard ratio of the estimated covariance matrix, i.e., by the quotient between its determinant and the product of its diagonal elements.

In this paper, we extend these works to the case of time-correlated signals, and derive the GLRT in the time and frequency domains. Interestingly, for wide sense stationary (WSS) processes we show that the frequency-domain detector asymptotically converges to the integrated logarithm of the frequency dependent Hadamard ratios, which nicely extends the results in [10]–[13] to the case of time-correlated signals.

In the case of bivariate signals, the proposed test is just a function of the estimated magnitude squared coherence (MSC) spectrum [14]. This suggests that the proposed frequency-domain detector can be seen as an extension of the MSC spectrum for more than two signals. Moreover, it admits a straightforward information-theoretic interpretation as the measure of the mutual information among more than two time-series. Finally, the proposed detector is compared with the GC by means of some numerical simulations and its application to cognitive radio is presented. As expected, exploiting the time structure of the spatially distributed signals notably improves the receiver operating characteristic (ROC) curve of the detector.

Manuscript received August 10, 2009; accepted May 27, 2010. Date of publication June 21, 2010; date of current version September 15, 2010. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Erik G. Larsson. This work was supported by the Spanish Government, Ministerio de Ciencia e Innovación (MICINN), under project MultiMIMO (TEC2007-68020-C04-02), project COMONSENS (CSD2008-00010, CONSOLIDER-INGENIO 2010 Program) and FPU Grant AP2006-2965.

D. Ramírez, J. Vía, and I. Santamaría are with the Department of Communications Engineering, University of Cantabria, 39005 Santander, Spain (e-mail: ramirezgd@gtas.dicom.unican.es; jvia@gtas.dicom.unican.es; nacho@gtas.dicom.unican.es).

L. L. Scharf is with the Departments of Electrical and Computer Engineering and Statistics, Colorado State University, Ft. Collins, CO 80523 USA (e-mail: scharf@engr.colostate.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2010.2053360

[1]The PSD matrix of a vector-valued time series is a matrix which contains all pairwise cross-spectra between each component of the vector-valued time series.

The paper is organized as follows. Section II presents the problem of multiple-channel signal detection. The GLRT in the time domain is obtained in Section III. A frequency domain representation of the GLRT and an approximation for WSS processes are presented in Section IV. Section V shows the relationship among the frequency domain detector, the coherence spectrum, the mutual information and the latent signal model, and also provides a practical approximation of the detector in the low-correlation regime. Finally, the performance of the proposed detector is illustrated by means of numerical simulations in Section VI, and the main conclusions are summarized in Section VII.

II. Preliminaries

Notation

In this paper, we use bold-faced upper case letters to denote matrices, with elements $[\mathbf{X}]_{i,k}$; bold-faced lower case letters for column vectors, and light-face lower case letters for scalar quantities. The superscripts $(\cdot)^T$ and $(\cdot)^H$ denote transpose and Hermitian, respectively. The superscript $\hat{(\cdot)}$ will denote estimated matrices, vectors or scalars. The determinant, trace and Frobenius norm of a matrix $\mathbf{A}$ will be denoted, respectively, as $\det(\mathbf{A})$, $\mathrm{tr}(\mathbf{A})$ and $||\mathbf{A}||_F$. The notation $\mathbf{A} \in \mathbb{C}^{M\times N}$ $\left(\mathbf{A} \in \mathbb{R}^{M\times N}\right)$ will be used to denote that $\mathbf{A}$ is a complex (real) matrix of dimension $M \times N$. For vectors, the notation $\mathbf{x} \in \mathbb{C}^M$ $\left(\mathbf{x} \in \mathbb{R}^M\right)$ denotes that $\mathbf{x}$ is a complex (real) vector of dimension M. $\mathbf{x} \sim \mathcal{CN}(\boldsymbol{\mu}, \mathbf{R})$ indicates that $\mathbf{x}$ is a complex circular Gaussian random vector of mean $\boldsymbol{\mu}$ and covariance matrix $\mathbf{R}$. The expectation operator will be denoted as $E[\cdot]$, $\mathrm{vec}\,(\mathbf{A})$ is the column-wise vectorization of $\mathbf{A}$, $\otimes$ is the Kronecker product and $*$ denotes the convolution operator. $\mathbf{I}_L$ is the identity matrix of size $L \times L$ and $\mathbf{0}_L$ denotes the zero vector or the zero matrix (depending on the context) of sizes $L \times 1$ and $L \times L$, respectively. Finally, $\mathbf{A}^{1/2}$ $(\mathbf{A}^{-1/2})$ is the Hermitian square root matrix of the Hermitian matrix $\mathbf{A}$ $(\mathbf{A}^{-1})$ and $\mathrm{diag}\,(\mathbf{A}_1, \ldots, \mathbf{A}_N)$ is the operator that forms a block-diagonal matrix from the matrices $\mathbf{A}_1, \ldots, \mathbf{A}_N$.

A. Problem Formulation

In this paper, we address the problem of testing for the covariance structure of the vector-valued time series $\{\mathbf{x}[n],\ n = 0, \pm 1, \ldots\}$, where $\mathbf{x}[n] = [x_1[n], \ldots, x_L[n]]^T$ is a vector of measurements at time n, or equivalently, $\{x_i[n]\}$ is the time series at sensor i. In order to proceed, we need the probability distribution of $\{\mathbf{x}[n]\}$, which we take to be circular complex Gaussian. We shall proceed by constructing the data matrix $\mathbf{X}$

$$\mathbf{X} = \begin{bmatrix} x_1[0] & x_1[1] & \ldots & x_1[N-1] \\ x_2[0] & x_2[1] & \ldots & x_2[N-1] \\ \vdots & \vdots & \ddots & \vdots \\ x_L[0] & x_L[1] & \ldots & x_L[N-1] \end{bmatrix} = \begin{bmatrix} \mathbf{x}_1^T \\ \mathbf{x}_2^T \\ \vdots \\ \mathbf{x}_L^T \end{bmatrix}$$

where the ith row, $\mathbf{x}_i^T = [x_i[0], x_i[1], \ldots, x_i[N-1]]$, contains N-samples of the ith time series $\{x_i[n]\}$, and the kth column is the kth sample of the vector-valued time series $\{\mathbf{x}[n]\}$. The vector $\mathbf{z} = \mathrm{vec}\left(\mathbf{X}^T\right)$ stacks the columns of $\mathbf{X}^T$, and its covariance matrix is

$$\mathbf{R} = E\left[\mathbf{z}\mathbf{z}^H\right] = \begin{bmatrix} \mathbf{R}_{11} & \mathbf{R}_{21}^H & \ldots & \mathbf{R}_{L1}^H \\ \mathbf{R}_{21} & \mathbf{R}_{22} & \ldots & \mathbf{R}_{L2}^H \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{R}_{L1} & \mathbf{R}_{L2} & \ldots & \mathbf{R}_{LL} \end{bmatrix}$$

where $\mathbf{R} \in \mathbb{C}^{LN\times LN}$ and the covariance matrix $\mathbf{R}_{ik} = E\left[\mathbf{x}_i\mathbf{x}_k^H\right],\ 1 \le i, k \le L$ captures all space–time second-order information within and without the random vectors $\{\mathbf{x}_i\}_{i=1}^L$.

We consider the two following hypotheses:

$$\mathcal{H}_0 : \mathbf{z} \sim \mathcal{CN}\left(\mathbf{0}_{LN}, \mathbf{R}_0\right)$$
$$\mathcal{H}_1 : \mathbf{z} \sim \mathcal{CN}\left(\mathbf{0}_{LN}, \mathbf{R}_1\right)$$

where $\mathbf{R}_1 \in \mathfrak{R}_1$, $\mathbf{R}_0 \in \mathfrak{R}_0$ are two unknown covariance matrices, $\mathfrak{R}_1$ is the set of covariance matrices with no particular temporal or spatial structure (i.e., they are only constrained to be positive definite) and $\mathfrak{R}_0$ is the set of block-diagonal covariance matrices, i.e., $\mathbf{R}_{ik} = \mathbf{0}_N,\ i \neq k$. Therefore, under the null hypothesis, the spatially uncorrelated vector-valued time series may be temporally correlated.

III. Derivation of the GLRT

The first result we shall discuss is an extension to multivariate time series of a standard result in multivariate normal theory, wherein a GLRT is used to test model $\mathcal{H}_0$ versus $\mathcal{H}_1$. To this end, we shall assume an experiment producing M independent copies of the data matrix $\mathbf{X}$, or equivalently its vectorized version $\mathbf{z} = \mathrm{vec}\left(\mathbf{X}^T\right)$ (see Fig. 1). The joint probability density function (pdf) for these measurements is the product of the pdfs for $\mathbf{z}[m] \sim \mathcal{CN}\left(\mathbf{0}_{LN}, \mathbf{R}\right)$, and is given by

$$\begin{aligned} p\left(\mathbf{z}[0], \ldots, \mathbf{z}[M-1]; \mathbf{R}\right) &= \prod_{m=0}^{M-1} p\left(\mathbf{z}[m]; \mathbf{R}\right) \\ &= \frac{1}{\pi^{LNM}\det\left(\mathbf{R}\right)^M} \exp\left\{-\sum_{m=0}^{M-1} \mathbf{z}^H[m]\mathbf{R}^{-1}\mathbf{z}[m]\right\} \\ &= \frac{1}{\pi^{LNM}\det\left(\mathbf{R}\right)^M} \exp\left\{-M\mathrm{tr}\left(\mathbf{R}^{-1}\hat{\mathbf{R}}\right)\right\}. \end{aligned}$$

Here, $\hat{\mathbf{R}}$ is the sample covariance matrix

$$\begin{aligned} \hat{\mathbf{R}} &= \frac{1}{M}\sum_{m=0}^{M-1} \mathbf{z}[m]\mathbf{z}^H[m] \\ &= \begin{bmatrix} \hat{\mathbf{R}}_{11} & \hat{\mathbf{R}}_{21}^H & \ldots & \hat{\mathbf{R}}_{L1}^H \\ \hat{\mathbf{R}}_{21} & \hat{\mathbf{R}}_{22} & \ldots & \hat{\mathbf{R}}_{L2}^H \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\mathbf{R}}_{L1} & \hat{\mathbf{R}}_{L2} & \ldots & \hat{\mathbf{R}}_{LL} \end{bmatrix}, \end{aligned}$$

and $\hat{\mathbf{R}}_{ik} \in \mathbb{C}^{N\times N}$ is the $\{i, k\}$th block of $\hat{\mathbf{R}}$, which is the sample cross-covariance matrix between the N-sample windows of the ith and kth time series.

To solve our hypothesis testing problem, we will use the generalized likelihood ratio test (GLRT). Although it is known that the GLRT is not optimal in the Neyman-Pearson sense, it provides good performance [15]; and as we will see, it results in a

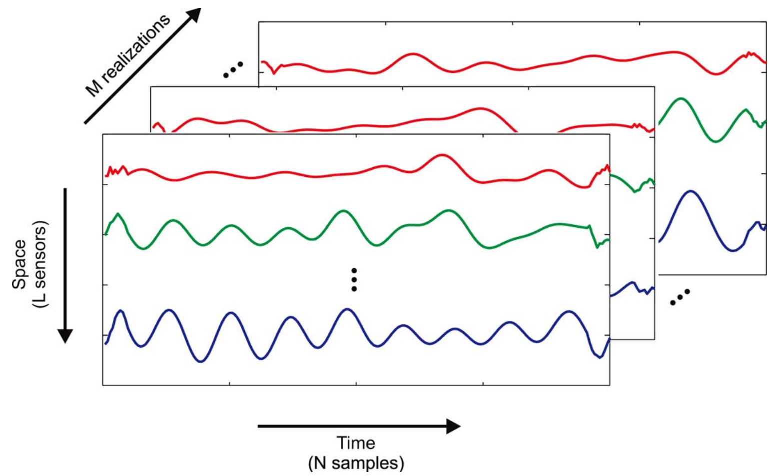

Fig. 1. Observations consist of M space–time snapshots of dimensions $L \times N$, each one representing the signals acquired by M sensors during N time instants.

simple detector with several interesting properties. The GLRT for $\mathcal{H}_0 : \mathbf{R} \in \mathfrak{R}_0$ versus $\mathcal{H}_1 : \mathbf{R} \in \mathfrak{R}_1$ is based on the generalized likelihood ratio (GLR) [15]

$$\lambda = \frac{\max\limits_{\mathbf{R}\in\mathfrak{R}_0} p\left(\mathbf{z}[0], \ldots, \mathbf{z}[M-1]; \mathbf{R}\right)}{\max\limits_{\mathbf{R}\in\mathfrak{R}_1} p\left(\mathbf{z}[0], \ldots, \mathbf{z}[M-1]; \mathbf{R}\right)} = \det\left(\hat{\mathbf{R}}_0^{-1}\hat{\mathbf{R}}_1\right)^M \exp\left\{-M\mathrm{tr}\left[\left(\hat{\mathbf{R}}_0^{-1} - \hat{\mathbf{R}}_1^{-1}\right)\hat{\mathbf{R}}\right]\right\}$$

where $\hat{\mathbf{R}}_0$ and $\hat{\mathbf{R}}_1$ are the maximum likelihood estimates of $\mathbf{R}$ under hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$, respectively.

As previously pointed out, under $\mathcal{H}_0$ the covariance matrix $\mathbf{R}$ is block-diagonal, therefore, $\mathfrak{R}_0$ is the set of matrices $\mathfrak{R}_0 = \{\mathbf{R} \mid \mathbf{R} = \mathrm{diag}\,(\mathbf{R}_{11}, \ldots, \mathbf{R}_{LL})\}$, with the only constraint that $\mathbf{R}_{ii}$ is Hermitian positive definite. That is, we force spatial uncorrelatedness but do not force any temporal structure.

Now, we will obtain the ML estimates of the covariance matrices under both hypotheses, for which we need to assume $M \geq N$. Taking this into account, it is easy to show that the ML estimate of $\mathbf{R}_0$ is $\hat{\mathbf{R}}_0 = \mathrm{diag}\left(\hat{\mathbf{R}}_{11}, \ldots, \hat{\mathbf{R}}_{LL}\right)$, [16]. As previously pointed out, we take $\mathfrak{R}_1$ to be the set of matrices $\mathbf{R}$, with no temporal or spatial structure imposed, with the only constraint being that $\mathbf{R}$ is an Hermitian positive definite matrix. Then, the ML estimate of $\mathbf{R}_1$ is given by $\hat{\mathbf{R}}_1 = \hat{\mathbf{R}}$, [16] and the GLRT becomes

$$\lambda^{1/M} = \det\left(\hat{\mathbf{R}}_0^{-1}\hat{\mathbf{R}}_1\right) = \frac{\det\left(\hat{\mathbf{R}}\right)}{\prod\limits_{i=1}^{L}\det\left(\hat{\mathbf{R}}_{ii}\right)} = \det\left(\hat{\mathbf{D}}^{-1/2}\hat{\mathbf{R}}\hat{\mathbf{D}}^{-1/2}\right) = \det(\hat{\mathbf{C}}) \quad (1)$$

where $\hat{\mathbf{D}} = \mathrm{diag}\left(\hat{\mathbf{R}}_{11}, \ldots, \hat{\mathbf{R}}_{LL}\right)$ and the matrix $\hat{\mathbf{C}} = \hat{\mathbf{D}}^{-1/2}\hat{\mathbf{R}}\hat{\mathbf{D}}^{-1/2}$ is a coherence matrix, sometimes also called a signal-to-noise ratio matrix when $\mathcal{H}_0$ may be considered as a noise-only hypothesis. The GLRT in (1) is a special case of a general result in [15] and, interestingly, it is a *generalized* Hadamard ratio.[2] Interestingly, it has been recently shown in [9] that the GLRT given by (1) is also related with the geodesic distance between $\hat{\mathbf{R}}_0$ and $\hat{\mathbf{R}}_1$ on the manifold of positive definite matrices.

Finally, we present the following property of the statistic.

Lemma 1: The GLRT in (1) is invariant to linear transformations of each time-series $x_i[n]$, which includes as a particular case any independent arbitrary scaling or filtering.

Proof: Defining the transformed time series as $\mathbf{y}_i = \mathbf{A}_i\mathbf{x}_i$, where $\mathbf{A}_i$ is any invertible matrix, it is easy to obtain $\hat{\mathbf{R}}_0^{(y)} = \mathbf{A}\hat{\mathbf{R}}_0^{(x)}\mathbf{A}^H$ and $\hat{\mathbf{R}}_1^{(y)} = \mathbf{A}\hat{\mathbf{R}}_1^{(x)}\mathbf{A}^H$, where $\mathbf{A} = \mathrm{diag}\,(\mathbf{A}_1, \ldots, \mathbf{A}_L)$ and $\hat{\mathbf{R}}_i^{(y)}$ and $\hat{\mathbf{R}}_i^{(x)}$ are, respectively, the sample covariance matrices of the transformed signals and the original ones. The proof concludes by substituting these matrices into (1). □

IV. GLRT in the Frequency Domain

In this section, we derive a frequency-domain version of the GLRT by exploiting its invariance to linear transformations. We also present an approximation of the proposed detector which, for WSS processes, asymptotically converges to the frequency-domain version of the GLRT.

We shall start by considering the signals $\mathbf{y}_i = \mathbf{F}_N\mathbf{x}_i$, where $\mathbf{F}_N$ is the $N \times N$ Fourier matrix with entries given by $[\mathbf{F}_N]_{i,k} = 1/\sqrt{N}e^{-j2\pi ik/N}$. Then, taking into account Lemma 1, the GLRT is equal for both sets of signals, $\{\mathbf{x}_1, \mathbf{x}_2, \ldots\}$ and $\{\mathbf{y}_1, \mathbf{y}_2, \ldots\}$, and is given by

$$\lambda^{1/M} = \det\left(\left(\mathbf{F}_N \otimes \mathbf{I}_L\right)\hat{\mathbf{C}}\left(\mathbf{F}_N \otimes \mathbf{I}_L\right)^H\right)$$

where $\hat{\mathbf{C}}$ is the estimated coherence matrix defined in the previous section. Now, introducing a simple permutation of the rows and columns of the matrix inside the determinant, the GLRT can be rewritten as

$$\lambda^{1/M} = \det(\tilde{\mathbf{C}})$$

where[3]

$$\tilde{\mathbf{C}} = \begin{bmatrix} \hat{\mathbf{C}}\left(e^{j\theta_0}\right) & \cdots & \hat{\mathbf{C}}\left(e^{j\theta_0}, e^{j\theta_{N-1}}\right) \\ \vdots & \ddots & \vdots \\ \hat{\mathbf{C}}\left(e^{j\theta_0}, e^{j\theta_{N-1}}\right) & \cdots & \hat{\mathbf{C}}\left(e^{j\theta_{N-1}}\right) \end{bmatrix}$$

[2]We use the term generalized to point out that the denominator is given by the determinant of the block-diagonal (instead of the diagonal) version of R.

[3]For notational simplicity we will use $\hat{\mathbf{C}}\left(e^{j\theta_i}\right)$ as a shorthand for $\hat{\mathbf{C}}\left(e^{j\theta_i}, e^{j\theta_i}\right)$.

is the global coherence matrix in the frequency domain, $\theta_i = 2\pi i/N$. The elements of the coherence matrices in the frequency domain are given by

$$[\hat{\mathbf{C}}\left(e^{j\theta_i}, e^{j\theta_k}\right)]_{l,m} = \mathbf{f}^H\left(e^{j\theta_i}\right)\hat{\mathbf{R}}_{ll}^{-1/2}\hat{\mathbf{R}}_{lm}\hat{\mathbf{R}}_{mm}^{-1/2}\mathbf{f}\left(e^{j\theta_k}\right)$$

and $\mathbf{f}\left(e^{j\theta}\right)$ is the Fourier vector at angular frequency θ.

We shall continue by decomposing $\tilde{\mathbf{C}}$ as

$$\tilde{\mathbf{C}} = \tilde{\mathbf{C}}_{\mathrm{WSS}}^{1/2}\tilde{\mathbf{C}}_{\mathrm{NS}}\tilde{\mathbf{C}}_{\mathrm{WSS}}^{1/2}$$

where $\tilde{\mathbf{C}}_{\mathrm{WSS}} = \mathrm{diag}\left(\hat{\mathbf{C}}\left(e^{j\theta_0}\right), \ldots, \hat{\mathbf{C}}\left(e^{j\theta_{N-1}}\right)\right)$ and $\tilde{\mathbf{C}}_{\mathrm{NS}} = \tilde{\mathbf{C}}_{\mathrm{WSS}}^{-1/2}\tilde{\mathbf{C}}\tilde{\mathbf{C}}_{\mathrm{WSS}}^{-1/2}$. Thus, taking the logarithm, the GLRT can be rewritten as

$$\frac{1}{M}\log\lambda = \log\lambda_{\mathrm{WSS}} + \log\lambda_{\mathrm{NS}}$$

where $\lambda_{\mathrm{WSS}} = \det(\tilde{\mathbf{C}}_{\mathrm{WSS}})$ and $\lambda_{\mathrm{NS}} = \det(\tilde{\mathbf{C}}_{\mathrm{NS}})$.

Now, let us consider the case in which the time series are jointly WSS. In this situation the covariance matrices are Toeplitz and, therefore, the elements of the matrix $\hat{\mathbf{C}}\left(e^{j\theta}\right)$ can be seen as estimates of the pairwise coherence spectra. Moreover, we have the asymptotic result

$$\frac{1}{N}\log\lambda_{\mathrm{WSS}} \underset{M,N\to\infty}{\longrightarrow} \int_{-\pi}^{\pi}\log\det\left(\hat{\mathbf{C}}\left(e^{j\theta}\right)\right)\frac{d\theta}{2\pi} = \int_{-\pi}^{\pi}\log\left[\frac{\det\left(\hat{\mathbf{S}}\left(e^{j\theta}\right)\right)}{\prod_{i=1}^{L}\hat{S}_{ii}\left(e^{j\theta}\right)}\right]\frac{d\theta}{2\pi}$$

where $[\hat{\mathbf{S}}\left(e^{j\theta}\right)]_{l,m} = \mathbf{f}^H\left(e^{j\theta}\right)\hat{\mathbf{R}}_{lm}\mathbf{f}\left(e^{j\theta}\right)$ is a quadratic estimator (averaged over M realizations) of the power spectral density matrix [17], [18], and $\hat{S}_{ii}\left(e^{j\theta}\right)$ are the elements of the main diagonal of $\hat{\mathbf{S}}\left(e^{j\theta}\right)$. In addition, the term $\log\lambda_{\mathrm{NS}}$ will approach zero,[4] which allows us to see it as a measure of the contribution to the test statistic of the *non-stationary part* of the time series. Therefore, the limiting form (L fixed and $M, N \to \infty$) of the GLRT statistic is

$$l = \lambda^{1/NM} \underset{M,N\to\infty}{\longrightarrow} \exp\left\{\int_{-\pi}^{\pi}\log\left[\frac{\det\left(\hat{\mathbf{S}}\left(e^{j\theta}\right)\right)}{\prod_{i=1}^{L}\hat{S}_{ii}\left(e^{j\theta}\right)}\right]\frac{d\theta}{2\pi}\right\}. \quad (2)$$

Here, we must note that this asymptotic version of the GLRT in the frequency domain is not the true GLRT for WSS processes. The reason is that the ML estimates of the covariance matrices should take into account their Toeplitz structure, which is a problem that, to the best of our knowledge, does not have a closed form solution [21], [22]. However, as we will see in the simulations section, the finite version of (2) presents better performance than (1). In addition, it is computationally efficient and has very nice properties, as we will see in the next section.

[4]These two results are easily proven taking into account the Szegö's Theorem for sequences of Toeplitz matrices [19] and its extension to Block-Toeplitz matrices, see for instance [20].

V. Properties and Further Discussion

In this section we show some interesting relationships between the frequency-domain approximation of the GLRT, given by (2), and other statistical measures such as the coherence spectrum or the mutual information among L WSS Gaussian processes. In addition, we present an approximation of (2) in the low correlation regime (i.e., when the pairwise cross-spectra are low in comparison to the power spectral densities), which can be useful to avoid some of the difficulties posed by the estimation of (2). Finally, we discuss the relationship between the problem addressed in this paper and the spatially reduced-rank signal-plus-noise model. All the results of this section specifically consider WSS processes.

A. Relationship With the Magnitude Squared Coherence Spectrum

Let us start by particularizing the approximation of the GLRT in (2) for $L = 2$ processes

$$l = \exp\left\{\int_{-\pi}^{\pi}\log\left(\frac{\hat{S}_{11}(e^{j\theta})\hat{S}_{22}(e^{j\theta}) - \left|\hat{S}_{12}(e^{j\theta})\right|^2}{\hat{S}_{11}(e^{j\theta})\hat{S}_{22}(e^{j\theta})}\right)\frac{d\theta}{2\pi}\right\}$$
$$= \exp\left\{\int_{-\pi}^{\pi}\log\left(1 - \left|\hat{\gamma}_{12}(e^{j\theta})\right|^2\right)\frac{d\theta}{2\pi}\right\}$$

where $\left|\hat{\gamma}_{12}(e^{j\theta})\right|^2 = \left|\hat{S}_{12}(e^{j\theta})\right|^2 / \left(\hat{S}_{11}(e^{j\theta})\hat{S}_{22}(e^{j\theta})\right)$ is an estimate of the MSC spectrum. Therefore, for this particular case the term inside the logarithm in (2) is a simple function of the MSC, which is a frequency-dependent measure of the linear relationship between two processes. Taking this into account, we propose the following generalization of the MSC spectrum for $L \geq 2$ random processes

$$\left|\gamma(e^{j\theta})\right|^2 = 1 - \frac{\det\left(\mathbf{S}(e^{j\theta})\right)}{\prod_{i=1}^{L}S_{ii}(e^{j\theta})} \quad (3)$$

where $\mathbf{S}(e^{j\theta})$ is the theoretical power spectral density matrix of the vector-valued time series. It is easy to show that this generalization has the following interesting properties:

1) Property 1: The generalized MSC in (3) is bounded between 0 and 1.

Proof: This is a direct consequence of the fact that the Hadamard ratio is always bounded between 0 and 1, which results from a majorization result that shows that the product of eigenvalues of $\mathbf{S}(e^{j\theta})$ is less than or equal to the product of its diagonal elements [23]. □

2) Property 2: The generalized MSC in (3) attains its maximum when the L time series admit a low-rank representation, i.e., when they can be represented as a linear combination of $L' < L$ processes.

Proof: The low-rank vector-valued time series admit the following representation

$$\mathbf{x}[n] = (\mathbf{H} * \mathbf{s})[n]$$

where $(\mathbf{H} * \mathbf{s})[n]$ denotes the convolution operation between $\mathbf{H}[n]$ and $\mathbf{s}[n]$, $\mathbf{H}[n] \in \mathbb{C}^{L\times L'}$ is a filtering matrix and $\mathbf{s}[n] \in$

$\mathbb{C}^{L'\times 1}$ is a time series of dimension L' whose matrix-valued covariance sequence is given by $E\left[\mathbf{s}[n+m]\mathbf{s}^H[n]\right] = \mathbf{I}_{L'}\delta[m]$. The power spectral density matrix of $\mathbf{x}[n]$ is given by

$$\mathbf{S}\left(e^{j\theta}\right) = \mathbf{H}\left(e^{j\theta}\right)\mathbf{H}^H\left(e^{j\theta}\right).$$

Obviously, $\mathbf{S}\left(e^{j\theta}\right)$ is rank-deficient ($L' < L$) and, therefore, $\det\left(\mathbf{S}\left(e^{j\theta}\right)\right) = 0$ which yields $\left|\gamma(e^{j\theta})\right|^2 = 1$. □

3) Property 3: The generalized MSC in (3) is invariant to independent linear filtering of the signals. This means that if we consider the following filtered signals

$$\mathbf{y}[n] = (\mathbf{D} * \mathbf{x})[n]$$

where $\mathbf{D}[n]$ is a diagonal matrix containing the impulse responses of L stable discrete-time filters along its main diagonal. Moreover, the Fourier transforms of the filters, $\mathbf{D}\left(e^{j\theta}\right) = \mathrm{diag}\left(D_1\left(e^{j\theta}\right), \ldots, D_L\left(e^{j\theta}\right)\right)$, must satisfy $D_i\left(e^{j\theta}\right) \neq 0$. Then $\left|\gamma(e^{j\theta})\right|^2$ is equal for the signals $\mathbf{y}[n]$ and $\mathbf{x}[n]$.

Proof: The power spectral density matrix of the filtered time series can be rewritten as

$$\mathbf{S}_y\left(e^{j\theta}\right) = \mathbf{D}\left(e^{j\theta}\right)\mathbf{S}_x\left(e^{j\theta}\right)\mathbf{D}^H\left(e^{j\theta}\right);$$

therefore, the determinant of $\mathbf{S}_y\left(e^{j\theta}\right)$ is given by

$$\det\left(\mathbf{S}_y\left(e^{j\theta}\right)\right) = \det\left(\mathbf{S}_x\left(e^{j\theta}\right)\right)\prod_{i=1}^{L}\left|D_i\left(e^{j\theta}\right)\right|^2.$$

Finally, substituting $\det\left(\mathbf{S}_y\left(e^{j\theta}\right)\right)$ into the definition, we obtain

$$\begin{aligned}\left|\gamma_y(e^{j\theta})\right|^2 &= 1 - \frac{\det\left(\mathbf{S}_y(e^{j\theta})\right)}{\prod_{i=1}^{L} S_{ii}^{(y)}(e^{j\theta})} \\ &= 1 - \frac{\det\left(\mathbf{S}_x\left(e^{j\theta}\right)\right)\prod_{i=1}^{L}\left|D_i\left(e^{j\theta}\right)\right|^2}{\prod_{i=1}^{L}\left|D_i\left(e^{j\theta}\right)\right|^2 S_{ii}^{(x)}(e^{j\theta})} \\ &= \left|\gamma_x(e^{j\theta})\right|^2\end{aligned}$$

i.e., the proposed generalization of the coherence spectrum of the original and filtered signals are identical, which concludes the proof. □

B. Relationship With the Mutual Information

The mutual information among L stochastic processes is defined as [24]

$$\begin{aligned}I(x_1[n], \ldots, x_L[n]) = &\sum_{i=1}^{L}\lim_{N\to\infty}\frac{1}{N}H([x_i[0], \ldots, x_i[N-1]) \\ &- \lim_{N\to\infty}\frac{1}{N}H(\mathbf{x}[0], \ldots, \mathbf{x}[N-1])\end{aligned}$$

where the two terms in the right-hand side of the above equation are, respectively, the marginal and joint entropy rates of the random processes $\{x_i[n], i = 1, \ldots, L\}$, [24]. For the case of L complex circular univariate Gaussian jointly wide sense stationary processes, the mutual information can be expressed in terms of the pairwise cross-power spectra of the processes similar to [25] and [26] for $L = 2$,[5]

$$I(x_1[n], \ldots, x_L[n]) = -\int_{-\pi}^{\pi}\log\left(\frac{\det\left(\mathbf{S}(e^{j\theta})\right)}{\prod_{i=1}^{L} S_{ii}(e^{j\theta})}\right)\frac{d\theta}{2\pi}.$$

Therefore, the proposed test statistic in the frequency domain given by (2) is an estimate of the mutual information among L Gaussian processes, i.e., $\hat{I}(x_1[n], \ldots, x_L[n]) = -\log l$.

C. Low Correlation Regime Approximation

In this subsection, following the ideas of [12], [13], we present an approximation of (2) in the low correlation regime. This is an interesting scenario in cognitive radio [6], where the signal-to-noise ratio is usually very low, which is equivalent to a very low correlation among signals. Let us consider again the coherence matrix

$$\begin{aligned}\hat{\mathbf{C}}(e^{j\theta}) &= \hat{\mathbf{D}}^{-1/2}(e^{j\theta})\hat{\mathbf{S}}(e^{j\theta})\hat{\mathbf{D}}^{-1/2}(e^{j\theta}) \\ &= \begin{bmatrix} 1 & \hat{C}_{21}^*(e^{j\theta}) & \cdots & \hat{C}_{L1}^*(e^{j\theta}) \\ \hat{C}_{21}(e^{j\theta}) & 1 & \cdots & \hat{C}_{L2}^*(e^{j\theta}) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{C}_{L1}(e^{j\theta}) & \hat{C}_{L2}(e^{j\theta}) & \cdots & 1 \end{bmatrix}\end{aligned}$$

where $\hat{\mathbf{D}}(e^{j\theta}) = \mathrm{diag}\left(\hat{S}_{11}(e^{j\theta}), \hat{S}_{22}(e^{j\theta}), \ldots, \hat{S}_{LL}(e^{j\theta})\right)$, and

$$\hat{C}_{ik}(e^{j\theta}) = \frac{\hat{S}_{ik}(e^{j\theta})}{\sqrt{\hat{S}_{ii}(e^{j\theta})\hat{S}_{kk}(e^{j\theta})}}, \quad i, k = 1, \ldots, L$$

is an estimate of the complex coherence spectrum [14] between the ith and kth signals. Thus, (2) can be alternatively written as

$$l = \exp\left\{\int_{-\pi}^{\pi}\log\det(\hat{\mathbf{C}}(e^{j\theta}))\frac{d\theta}{2\pi}\right\}. \tag{4}$$

In the low correlation regime $\left|\hat{C}_{ik}(e^{j\theta})\right| \approx 0$, $i \neq k$, i.e., $\hat{\mathbf{C}}(e^{j\theta})$ is approximately equal to the identity matrix. Therefore, its eigenvalues may be approximated as

$$\lambda_i(e^{j\theta}) = 1 + \epsilon_i(e^{j\theta})$$

where $|\epsilon_i(e^{j\theta})| \ll 1$ and $\sum_{i=1}^{L}\epsilon_i(e^{j\theta}) = 0$, since $\mathrm{tr}\left(\hat{\mathbf{C}}(e^{j\theta})\right) = L$. Then, the logdet of the pairwise coherence matrix is

$$\log\det\left(\hat{\mathbf{C}}(e^{j\theta})\right) = \sum_{i=1}^{L}\log\lambda_i(e^{j\theta}) = \sum_{i=1}^{L}\log\left[1 + \epsilon_i(e^{j\theta})\right]$$

[5]This result is a direct application of the Szegö's theorem for sequences of Toeplitz matrices [19], [27] and its extension to sequences of Block-Toeplitz matrices [20].

which, using the Taylor series expansion of $\log(1+x)$ up to the second order, can be approximated by

$$\begin{aligned}\sum_{i=1}^{L}\log\left[1+\epsilon_i(e^{j\theta})\right] &\approx \sum_{i=1}^{L}\left[\epsilon_i(e^{j\theta})-\frac{1}{2}\epsilon_i^2(e^{j\theta})\right]\\ &= -\frac{1}{2}\sum_{i=1}^{L}\epsilon_i^2(e^{j\theta})\\ &= -\frac{1}{2}\left(\left\|\hat{\mathbf{C}}(e^{j\theta})\right\|_F^2 - L\right).\end{aligned}$$

Finally, the test (2) can be approximated by

$$l \approx \exp\left\{-\frac{1}{2}\int_{-\pi}^{\pi}\left\|\hat{\mathbf{C}}(e^{j\theta})\right\|_F^2\frac{d\theta}{2\pi}+\frac{L}{2}\right\} \tag{5}$$

which generalizes the result of [12], [13] to vector-valued time series. The Frobenius norm in (5) is, in general, easier to estimate than the determinant in (2). Thus, in addition to being a good approximation of (4) in the low correlation regime (or, equivalently, in the low signal-to-noise ratio regime), the Frobenius norm approximation (5) is a more robust test statistic than the logdet detector (2) when the number of available samples is small and, in consequence, $\hat{\mathbf{C}}(e^{j\theta})$ is poorly estimated.

Finally, we must point out that other approximations are also valid, depending on the number of dominant eigenvalues $\lambda_i(e^{j\theta})$. Some interesting cases are those based on the largest and the smallest eigenvalue [28], [29]. However, they will not be considered in this work.

D. Latent Signal Model and Structured Matrices

A particular problem that fits within the general signal model presented in Section II-B is the following binary hypotheses testing problem, which is commonly encountered in multi-sensor array processing:

$$\begin{aligned}\mathcal{H}_1 &: \mathbf{x}[n] = (\mathbf{H}*\mathbf{s})[n] + \mathbf{w}[n]\\ \mathcal{H}_0 &: \mathbf{x}[n] = \mathbf{w}[n]\end{aligned}$$

where $\mathbf{H}[n] \in \mathbb{C}^{L\times P}$ is an unknown and deterministic multiple-input multiple-output channel, $\mathbf{s}[n] \in \mathbb{C}^P$ is an spatially and temporally uncorrelated signal transmitted at the nth time instant, $\mathbf{w}[n]$ is the additive spatially uncorrelated noise (though it might be correlated in time) and P is usually smaller than or equal to L, i.e., $P \leq L$. In signal processing and communications this is the conventional signal-plus-noise model, whereas in other fields such as statistics or econometrics it is referred to as the latent signal model.

It is important to remark that during our derivation of the test we did not impose any constraint on the estimate of the covariance matrix (apart from its block-diagonal structure under $\mathcal{H}_0$). However, if known, the test should exploit the structure induced by the spatially rank-reduced model. Here, we analyze the effect of the order P on the ML estimates. To this end, we rewrite the estimated power spectral density matrix (under hypothesis $\mathcal{H}_1$) as

$$\begin{aligned}\hat{\mathbf{S}}(e^{j\theta}) &= \hat{\mathbf{H}}(e^{j\theta})\hat{\mathbf{H}}^H(e^{j\theta}) + \hat{\mathbf{S}}_w(e^{j\theta})\\ &= \mathbf{U}(e^{j\theta})\mathbf{\Sigma}^2(e^{j\theta})\mathbf{U}^H(e^{j\theta}) + \hat{\mathbf{S}}_w(e^{j\theta})\end{aligned} \tag{6}$$

where $\hat{\mathbf{H}}(e^{j\theta}) = \mathbf{U}(e^{j\theta})\mathbf{\Sigma}(e^{j\theta})\mathbf{V}^H(e^{j\theta})$ is the compact singular value decomposition of the estimated Fourier transform of the channel, $\hat{\mathbf{S}}_w(e^{j\theta})$ is the estimated power spectral density matrix of the noise and, without loss of generality, we assume a spatially and temporally white excitation of the matrix-valued filter $\mathbf{H}[n]$, i.e., $E\left[\mathbf{s}[n+m]\mathbf{s}^H[m]\right] = \mathbf{I}_P\delta[n]$. In general, the sample covariance matrix in (6) can not be matched by the spatially low-rank model. To be more precise, we use dimension counting arguments to derive a *necessary* condition on the value of P that allows us to determine the cases in which the detector is given by (2). To find these values of P, we have to count the number of equations and the number of independent parameters (or degrees of freedom).

The number of equations is easily found by taking into account the structure of the sample covariance matrix, which is Hermitian. So, we have $L(L-1)/2$ complex equations above the diagonal plus L real equations in the diagonal, which sums (real equations) to

$$2\frac{L(L-1)}{2} + L = L^2.$$

The determination of the degrees of freedom is more involved [30]: there are P non-zero singular values in the decomposition of $\hat{\mathbf{H}}(e^{j\theta})$ and L real elements in the diagonal matrix $\hat{\mathbf{S}}_w(e^{j\theta})$, whereas, the number of real parameters in $\mathbf{U}(e^{j\theta})$ is $2LP$. However, not all the parameters in $\mathbf{U}(e^{j\theta})$ are independent. The eigenvectors should have unit norm, which reduces the degrees of freedom in $2P$, and should be mutually orthogonal which reduces the number of independent parameters in $2\sum_{i=1}^{P-1} i = 2P(P-1)/2$. Summarizing, the number of free parameters is

$$P + L + 2LP - 2P - 2\frac{P(P-1)}{2} = L + 2LP - P^2.$$

Thus, we can easily see that if the following condition is satisfied

$$L + 2LP - P^2 \geq L^2 \Rightarrow P \geq L - \sqrt{L} \tag{7}$$

then we have at least as many degrees of freedom as equations, and therefore there might exist a solution[6] $\mathbf{U}(e^{j\theta})$, $\mathbf{\Sigma}^2(e^{j\theta})$, $\hat{\mathbf{S}}_w(e^{j\theta})$ exactly satisfying (6), which implies that the GLRT is given by (2). Finally, let us mention that for white Gaussian processes this rank-P model has been analyzed in [13], obtaining an equivalent result.

VI. Numerical Results

A. Non-Stationary Processes

In this subsection, we evaluate the performance of the proposed detectors, that is, the GLRT given by (1) and the frequency domain detector given by (2). The observations are generated as follows:

$$\begin{aligned}\mathcal{H}_1 &: \mathbf{x}[n] = v[n]\left(\sum_{\tau=0}^{T-1}\mathbf{H}[\tau]\mathbf{s}[n-\tau] + \mathbf{w}[n]\right)\\ \mathcal{H}_0 &: \mathbf{x}[n] = v[n]\mathbf{w}[n]\end{aligned}$$

where $v[n] = 1 + \alpha\cos(2\pi\nu n)$, $\mathbf{s}[n] \in \mathbb{C}^P$ is a spatially and temporally white process distributed as $\mathbf{s}[n] \sim \mathcal{CN}(0, \mathbf{I}_L/L)$

[6]The solution must satisfy that $[\mathbf{\Sigma}^2(e^{j\theta})]_{i,i},\ i = 1,\ldots,P$, and $[\hat{\mathbf{S}}_w(e^{j\theta})]_{i,i},\ i = 1,\ldots,L$, are real and positive.

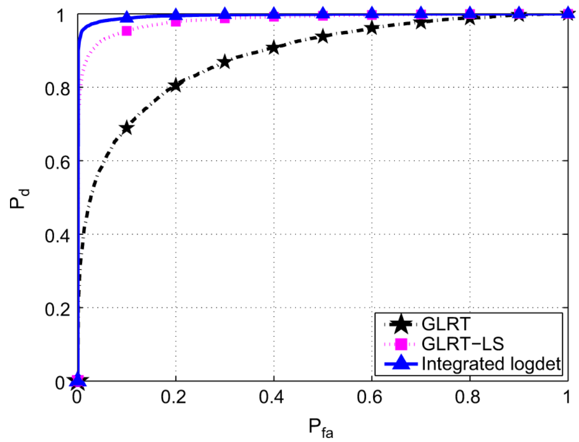

Fig. 2. ROC for $L = 3$, SNR = - 5 dB, $N = 48$, $M = 150$, $\alpha = 0.9$, $\nu = 10/N$.

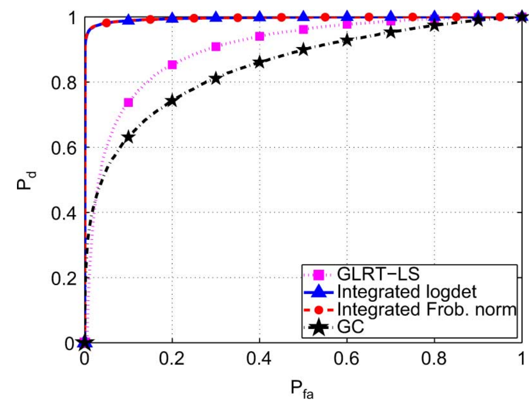

Fig. 3. ROC for $L = 3$, SNR = - 8 dB, $N = 48$, $M = 120$.

and $\mathbf{H}[n] \in \mathbb{C}^{L\times P}$. Notice that the observations correspond to a non-stationary signal where the *degree* of non-stationarity varies with α and ν. Moreover, we consider $P = L$ and, therefore, it is not possible to find a reduced-rank representation which could improve the performance of the detector.

Each coefficient of the channel matrix is generated as follows:

$$[\mathbf{H}[n]]_{i,k} \sim \mathcal{CN}\left(0, \frac{1}{T}\right), \quad \begin{cases} n = 0, \dots, T-1 \\ i = 1, \dots, L \\ k = 1, \dots, P \end{cases} \tag{8}$$

and the additive noises at each sensor $w_i[n]$ are independent zero-mean and complex white Gaussian processes with unknown but common variance σ^2, colored through unknown finite impulse response (FIR) filters of T_w random coefficients generated as

$$a_i[n] \sim \mathcal{CN}\left(0, \frac{1}{T_w}\right), \quad n = 0, \dots, T_w - 1;\ i = 1, \dots, L.$$

The signal-to-noise ratio (SNR) for each sensor is defined as SNR(dB) $= 10\log_{10}(1/\sigma^2)$, and we have considered $T = 10$ and $T_w = 4$.

In this first example, we compare the receiver operating characteristic (ROC) curves of the following detectors:

- the GLRT in the time domain given by (1) (denoted as GLRT in the figures);
- the frequency-domain approximation of the GLRT given by (2) (denoted as Integrated logdet);
- the time domain detector which imposes a Toeplitz structure on the covariance matrix using the least squares (LS) estimator [22], i.e., averaging the sample covariance matrices along diagonals (denoted as GLRT-LS).

1) Example 1: In this example, we have considered $L = 3$, SNR = - 5 dB, $N = 48$, $M = 150$, $\alpha = 0.9$ and $\nu = 10/N$. The results are shown in Fig. 2, where we can see that the GLRT detector presents poor results and it is clearly outperformed by the frequency-domain detector even for these non-stationary signals. The main reason for this poor performance is that the determinant estimates in the GLRT are very sensitive to the finite sample effect due to the large size of the estimated covariance matrices. The GLRT-LS detector presents better performance than the GLRT (although worse than the integrated logdet) because it is based on a simpler model, with a reduced number of parameters (the correlation values) to be estimated.

B. Stationary Processes

In this subsection, we consider the case of wide sense stationary signals; analyze in more depth the frequency domain detector (2) and its approximation (5); and compare them with the generalized coherence detector [10]. The observations are generated as in the previous subsection but, in order to obtain a stationary signal, we have selected $\alpha = 0$.

The ROC curves of the following detectors are compared:

- the frequency-domain approximation of the GLRT (denoted as Integrated logdet);
- the detector based on the Frobenius norm of the power spectral density matrix (denoted as Integrated Frob. norm);
- the generalized coherence detector proposed by Cochran [10] (denoted as GC in the figures), which anticipates the detector of [12] and [13] and is given by

$$1 - \frac{\det(\hat{\mathbf{R}}[0])}{[\hat{\mathbf{R}}[0]]_{11} \dots [\hat{\mathbf{R}}[0]]_{LL}} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \nu$$

  where

$$\hat{\mathbf{R}}[0] = \frac{1}{NM} \sum_{m=0}^{M-1} \mathbf{X}[m]\mathbf{X}^H[m];$$

- the time domain detector which imposes a Toeplitz structure to the covariance matrix using the least squares (LS) estimator [22] (denoted as GLRT-LS).

1) Example 2: In this example, we have considered $L = 3$, SNR = - 8 dB, $N = 48$ and $M = 120$. The results are shown in Fig. 3, where we can see that the proposed frequency domain detector and its approximation provide the best results. This example also serves to validate the Frobenius norm approximation of the optimal logdet detector for this low correlation scenario (i.e., low SNR). Obviously, the GC performs poorly because it was designed for temporally white processes, and never intended for correlated time series. Finally, the difference in performance between the logdet detector and GLRT-LS detector is greater than in the previous example, mainly, due to the smaller number of realizations M.

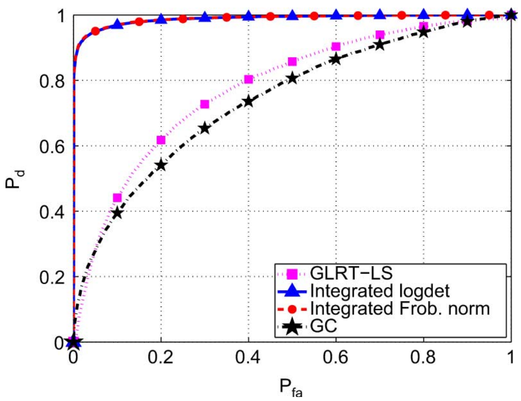

Fig. 4. ROC for $L = 5$, SNR = - 13 dB, $N = 48$, $M = 350$.

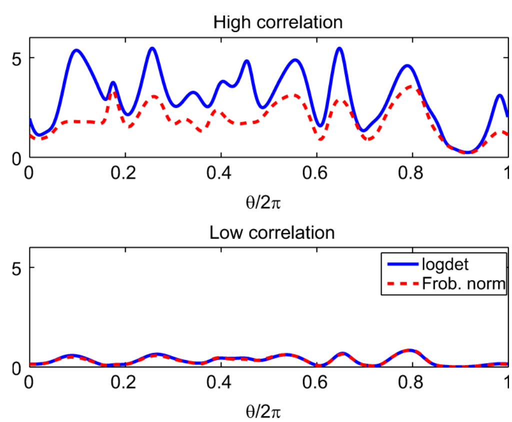

Fig. 5. Theoretical values of the logdet and Frobenius norm of the pairwise coherence spectra matrix under $\mathcal{H}_1$.

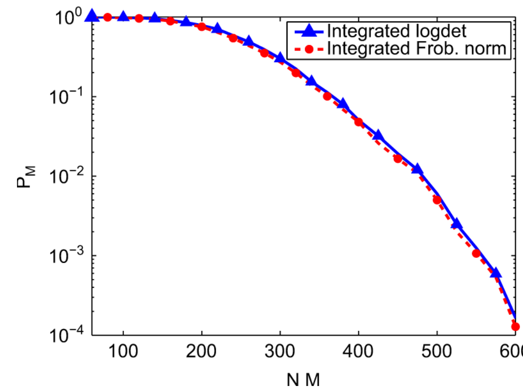

Fig. 6. Probability of missed detection for $p_{\mathrm{fa}} = 10^{-4}$ as a function of NM for the low correlation scenario.

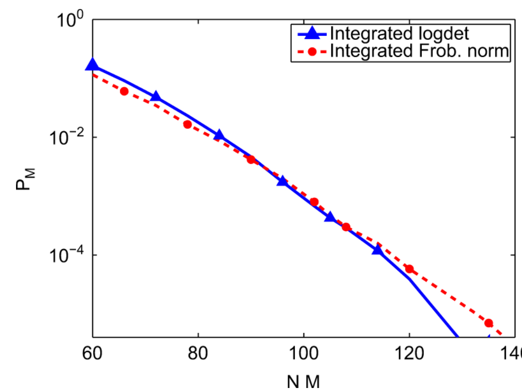

Fig. 7. Probability of missed detection for $p_{\mathrm{fa}} = 10^{-4}$ as a function of NM for the high correlation scenario.

2) Example 3: In the third example, the parameters are $L = 5$, SNR = - 13 dB, $N = 48$, $M = 350$. Fig. 4 shows the results for this example, from which similar conclusions are drawn.

3) Example 4: This example illustrates the robustness of the Frobenius norm based detector to finite size effects. In particular, we compare the performance of the detectors of (2) and (5) as a function of NM in the low and high correlation cases for a fixed channel and noise spectral densities, i.e., $a_i[n]$ and $[\mathbf{H}[n]]_{i,k}$ are fixed. Concretely, the probability of missed detection (p_M) for a fixed false alarm probability $(p_{\mathrm{fa}} = 10^{-4})$ is compared as a function of the number of available samples in two different scenarios, highly correlated signals (SNR = 20 dB) and low correlated signals (SNR = 5 dB). Fig. 5 shows the theoretical values of the proposed generalization of the magnitude squared coherence spectrum and its approximation based on the Frobenius norm, where we can see that they are approximately equal for low correlated signals and different for highly correlated signals. The number of signals for this example is $L = 3$. As can be seen in Fig. 6, the performance of both detectors is the same in the low correlation regime. However, in the high correlation regime (Fig. 7), the performance of the logdet detector is worse for small NM, and it is obviously better for a sufficiently large number of samples.

4) Example 5: This example illustrates the effect of the choice of M and N. The value of N determines the spectral resolution (bias) whereas M determines the quality of the estimate, i.e., its variance. Therefore, for a fixed value of MN there is a bias-variance trade-off. This is a classical problem in the field of spectral estimation and statistics, and therefore, we will not analyze its effects on the estimate of $\log\det(\hat{\mathbf{C}}(e^{j\theta}))$. Instead, we present some simulations showing how it affects the performance of the detector. Fig. 8 shows the probability of missed detection of the logdet detector ($p_{\mathrm{fa}} = 10^{-5}$) for two different values of N (two different spectral resolutions) for the high correlated case of Example 4. In this figure, we can see that for a small number of samples, i.e., NM small, it is advisable to sacrifice some spectral resolution in order to reduce the variance of the estimate. On the other hand, when the number of realizations increases, the spectral resolution becomes more important.

C. Application to Cognitive Radio

In this subsection, we present the application of the proposed detector to cognitive radio (CR) [6]. CR is a new paradigm in communications in which the users make an opportunistic access to the wireless channel when it is free. The basic idea is that there are some primary (or licensed) users which have assigned a frequency band, and there are some secondary users who can access that frequency band if no primary user is transmitting. Therefore, any CR system must rely on a spectrum sensing device (see [7] and references therein for a description of previously proposed detectors). If the primary users and the CR node are equipped with multiple antennas, and making the assumption that the noise processes at different antennas are uncorrelated, the detection problem in CR is equivalent to the hypothesis test described in Section V-D. Then, defining P as the number of antennas of the primary user, L as the number of antennas at the CR node and assuming that (7) is satisfied and that

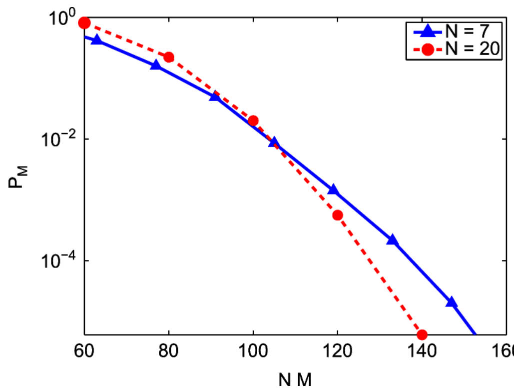

Fig. 8. Probability of missed detection for $p_{\rm fa} = 10^{-5}$ as a function of NM for high correlation scenario and two different resolutions.

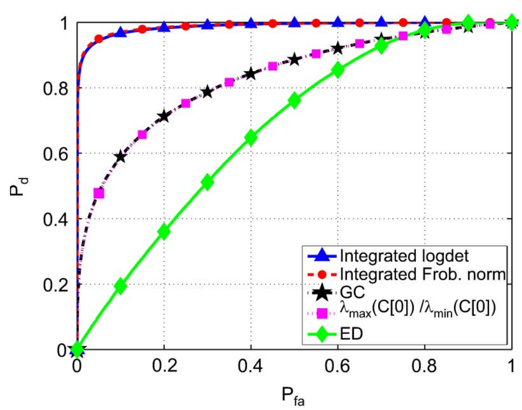

Fig. 9. ROC for $L = 3$, SNR = 0 dB, $N = 100$, $M = 10$.

the pdf of the primary signal is Gaussian, the detector given by (2) and also its approximation in (5) can be directly applied to detection of primary users in CR [31].

1) Example 6: In this final example, we present some simulation results to illustrate the application of the proposed detector in CR. In addition, we have compared its performance with that of the following CR detectors:

- the generalized coherence detector (denoted as GC);
- a modification of the detector proposed in [32] to handle noises with different powers at each antenna; the detector is based on the ratio of largest to smallest eigenvalues of the $L \times L$ spatial coherence matrix in the time domain
$$\hat{\mathbf{C}}[0] = \hat{\mathbf{D}}^{-1/2}[0]\hat{\mathbf{R}}[0]\hat{\mathbf{D}}^{-1/2}[0]$$
and $\hat{\mathbf{D}}[0]$ is a diagonal matrix formed from the main diagonal of $\hat{\mathbf{R}}[0]$. This detector will be denoted as $\lambda_{\max}(\hat{\mathbf{C}}[0])/\lambda_{\min}(\hat{\mathbf{C}}[0])$;
- the energy detector (denoted as ED) using LN samples per realization (the total number of samples is therefore MLN).

For the simulation, we have used OFDM-modulated DVB-T signals[7] with a bandwidth of 7.61 MHz. We have considered a 3×3 spatially uncorrelated frequency-selective Rayleigh fading channel with unit power and an exponential power delay profile with delay spread of 0.779 μs [33] ($L = P = 3$). The additive noises at each antenna are generated by filtering independent zero-mean and complex white Gaussian processes with common variance σ^2 with FIR filters with 4 i.i.d. random taps distributed as $a_i[n] \sim \mathcal{CN}(0, 1/4)$, $n = 0, \ldots, 3$; $i = 1, \ldots, L$, and the common SNR for all antennas is SNR(dB) $= 10\log_{10}(1/\sigma^2) = 0$ dB. Finally, for the detection process, there are available $M = 10$ realizations of length $N = 100$. Fig. 9 shows the results for this example, where we can see that the best results are obtained by the proposed detector.

[7] 8K mode, 64-QAM, guard interval $1/4$ and inner code rate $2/3$.

VII. Conclusion

In this work, we have derived the generalized likelihood ratio test (GLRT) for deciding whether L complex circular Gaussian signals with unknown arbitrary covariance matrices are spatially correlated or not. This is an interesting problem since it appears in a wide variety of applications, such as detection in sensor networks, or in multiple-input multiple-output (MIMO) radar. The most interesting findings are provided by the GLRT in the frequency domain, since it is closely related to other statistical measures such as the coherence spectrum or the mutual information. Specifically, we present an interesting frequency domain approximation of the GLRT given by the integral of the logarithm of the Hadamard ratio of the estimated cross power spectral density matrix. In the case of $L = 2$ signals, this test is given by a function of the well-known magnitude squared coherence (MSC) spectrum. This fact has prompted us to propose a generalization of the MSC spectrum for more than two signals which is essentially defined as the determinant of a matrix containing all the pairwise complex coherence spectra. This generalizes Cochran's multi-channel coherence from L random variables to L time series. In addition, we have presented an approximation of the integrated logdet for low SNR scenarios, which provides good results and is robust under small sets of data. The derivation of detectors for more structured detection problems (e.g., the reduced-rank latent signal model) is an interesting future research line.

References

[1] F. Zhao and L. Guibas*, Wireless Sensor Networks: An Information Processing Approach*. Amsterdam, The Netherlands: Elsevier, 2004.

[2] *Cooperation in Wireless Networks: Principles and Applications*, F. H. P. Fitzek and M. D. Katz, Eds. New York: Springer, 2006.

[3] N. Laneman, D. Tse, and G. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[4] H. Mheidat, M. Uysa, and N. Al-Dhahir, "Equalization techniques for distributed space–time block codes with amplify-and-forward relaying," *IEEE Trans. Signal Process.*, vol. 55, no. 5, pp. 1839–1852, May 2007.

[5] J. Li and P. Stoica*, MIMO Radar Signal Processing*. New York: Wiley-IEEE Press, 2008.

[6] *Cognitive Wireless Communication Networks*, E. Hossain and V. K. Bhargava, Eds. New York: Springer, 2007.

[7] J. Ma, G. Y. Li, and B. Hwang, "Signal processing in cognitive radio," *Proc. IEEE*, vol. 97, no. 5, pp. 805–823, May 2009.

[8] R. López-Valcarce and G. Vazquez-Vilar, "Multiantenna detection of multicarrier primary signals exploiting spectral *a priori* information," presented at the CrownCom, Hannover, Germany, Jun. 2009.

[9] A. Pérez-Neira, M. A. Lagunas, M. A. Rojas, and P. Stoica, "Correlation matching approach for spectrum sensing in open spectrum communications," *IEEE Trans. Signal Process.*, vol. 57, no. 12, pp. 4823–4836, Dec. 2009.

[10] D. Cochran, H. Gish, and D. Sinno, "A geometric approach to multiple-channel signal detection," *IEEE Trans. Signal Process.*, vol. 43, no. 9, pp. 2049–2057, Sep. 1995.

[11] A. Clausen and D. Cochran, "Non-parametric multiple channel detection in deep ocean noise," in *Proc. Conf. Rec. 31st Asilomar Conf. Signal, Systems, Computers*, Oct. 1997, vol. 1, pp. 850–854.

[12] A. Leshem and A.-J. van der Veen, "Multichannel detection of Gaussian signals with uncalibrated receivers," *IEEE Signal Process. Lett.*, vol. 8, no. 4, pp. 120–122, Apr. 2001.

[13] A. Leshem and A. J. van der Veen, "Multichannel detection and spatial signature estimation with uncalibrated receivers," in *Proc. 11th IEEE Workshop Stat. Signal Process.*, Aug. 6–8, 2001, pp. 190–193.

[14] P. Stoica and R. Moses, *Spectral Analysis of Signals*. Englewood Cliffs, NJ: Prentice-Hall, 2005.

[15] K. V. Mardia, J. T. Kent, and J. M. Bibby, *Multivariate Analysis*. New York: Academic, 1979.

[16] J. R. Magnus and H. Neudecker, *Matrix Differential Calculus With Applications in Statistics and Econometrics*, 2nd ed. New York: Wiley, 1999.

[17] D. J. Thomson, "Spectrum estimation and harmonic analysis," *Proc. IEEE*, vol. 70, pp. 1055–1096, Sep. 1982.

[18] C. T. Mullis and L. L. Scharf, "Quadratic estimators of the power spectrum," in *Advances in Spectrum Analysis and Array Processing*, S. Haykin, Ed. Englewood Cliffs, NJ: Prentice-Hall, 1991, vol. I, ch. 1, pp. 1–57.

[19] U. Grenander and G. Szegö, *Toeplitz Forms and Their Applications*. Berkeley, CA: Univ. of California Press, 1958.

[20] J. Gutierrez-Gutierrez and P. M. Crespo, "Asymptotically equivalent sequences of matrices and Hermitian block Toeplitz matrices with continuous symbols: Applications to MIMO systems," *IEEE Trans. Inf. Theory*, vol. 54, no. 12, pp. 5671–5680, Dec. 2008.

[21] J. P. Burg, D. G. Luenberger, L. Wenger, and D. L. Wenger, "Estimation of structured covariance matrices," *Proc. IEEE*, vol. 70, pp. 963–974, 1982.

[22] L. L. Scharf, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*. Reading, MA: Addison-Wesley, 1991.

[23] E. Jorswieck and H. Boche, "Majorization and matrix-monotone functions in wireless communications," *Foundations and Trends in Communications and Information Theory*, vol. 3, no. 6, Jun. 2007.

[24] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. New York: Wiley-Interscience, 2006.

[25] L. Scharf and J. Thomas, "Wiener filters in canonical coordinates for transform coding, filtering and quantizing," *IEEE Trans. Signal Process.*, vol. 46, no. 3, pp. 647–654, Mar. 1998.

[26] L. L. Scharf and C. T. Mullis, "Canonical coordinates and the geometry of inference, rate and capacity," *IEEE Trans. Signal Process.*, vol. 48, no. 3, pp. 824–831, Mar. 2000.

[27] R. M. Gray, "Toeplitz and circulant matrices: A review," *Found. Trends Commun. Inf. Theory*, vol. 2, no. 3, 2006.

[28] D. Ramírez, J. Vía, and I. Santamaría, "A generalization of the magnitude squared coherence spectrum for more than two signals: Definition, properties and estimation," in *Proc. IEEE Int. Conf Acoust., Speech, Signal Process. (ICASSP)*, Apr. 2008.

[29] D. Ramírez, J. Vía, and I. Santamaría, "Multiple-channel signal detection using the generalized coherence spectrum," in *Proc. IAPR Work. Cognitive Inf. Process.*, Jun. 2008.

[30] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-33, no. 2, pp. 387–387, Apr. 1985.

[31] D. Ramírez, J. Vía, I. Santamaría, R. López-Valcarce, and L. L. Scharf, "Multiantenna spectrum sensing: Detection of spatial correlation among time-series with unknown spectra," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Mar. 2010.

[32] Y. Zeng and Y.-C. Liang, "Eigenvalue-based spectrum sensing algorithms for cognitive radio," *IEEE Trans. Commun.*, vol. 57, no. 6, pp. 1784–1793, Jun. 2009.

[33] M. Falli, "Digital land mobile radio communications—Cost 207: Final report," Luxembourg, 1989.

David Ramírez (S'07) received the Telecommunication Engineer degree from the University of Cantabria, Spain, in 2006.

Since 2006 he has been working towards the Ph.D. degree at the Communications Engineering Department, University of Cantabria, Spain, under the supervision of I. Santamaría and J. Vía. He has been a visiting researcher, under the supervision of Prof. P. J. Schreier at the University of Newcastle, Australia. His current research interests include signal processing for wireless communications, MIMO Systems, MIMO testbeds, cognitive radio, and multivariate statistical analysis. He has been involved in several national and international research projects on these topics.

Javier Vía (M'08) received the Telecommunication Engineer degree and the Ph.D. degree in electrical engineering from the University of Cantabria, Spain, in 2002 and 2007, respectively.

In 2002, he joined the Department of Communications Engineering, University of Cantabria, Spain, where he is currently Assistant Professor. He has spent visiting periods at the Smart Antennas Research Group of Stanford University, Stanford, CA, and at the Department of Electronics and Computer Engineering (Hong Kong University of Science and Technology). He has actively participated in several European and Spanish research projects. His current research interests include blind channel estimation and equalization in wireless communication systems, multivariate statistical analysis, quaternion signal processing, and kernel methods.

Ignacio Santamaría (M'96–SM'05) received the Telecommunication Engineer degree and the Ph.D. degree in electrical engineering from the Universidad Politécnica de Madrid (UPM), Spain, in 1991 and 1995, respectively.

In 1992, he joined the Departamento de Ingeniería de Comunicaciones, Universidad de Cantabria, Spain, where he is currently Full Professor. He has been a visiting researcher at the Computational NeuroEngineering Laboratory of the University of Florida and at the Wireless Networking and Communications Group of the University of Texas at Austin. He has more than 100 publications in refereed journals and international conference papers. His current research interests include signal processing algorithms for wireless communication systems, MIMO systems, multivariate statistical techniques, and machine learning theories. He has been involved in several national and international research projects on these topics.

Dr. Santamaría is currently serving as a member of the Machine Learning for Signal Processing Technical Committee of the IEEE Signal Processing Society.

Louis L. Scharf (S'67–M'69–SM'77–F'86–LF'07) received the Ph.D. degree from the University of Washington, Seattle.

From 1971 to 1982, he served as Professor of electrical engineering and statistics at Colorado State University (CSU), Ft. Collins. From 1982 to 1985, he was Professor and Chairman of electrical and computer engineering at the University of Rhode Island, Kingston. From 1985 to 2000, he was Professor of electrical and computer engineering at the University of Colorado, Boulder. In January 2001, he rejoined CSU as Professor of electrical and computer engineering and statistics. He has held several visiting positions here and abroad, including the Ecole Superieure d'electricité, Gif-sur-Yvette, France; Ecole Nationale Superieure des Télécommunications, Paris, France; EURECOM, Nice, France; the University of La Plata, La Plata, Argentina; Duke University, Durham, NC; the University of Wisconsin, Madison; and the University of Tromsø, Tromsø, Norway. His interests are in statistical signal processing, as it applies to adaptive radar, sonar, and wireless communication. His most important contributed to date are to invariance theories for detection and estimation; matched and adaptive subspace detectors and estimators for radar, sonar, and data communication; and canonical decompositions for reduced dimensional filtering and quantizing. His current interests are in rapidly adaptive receiver design for space–time and frequency-time signal processing in the radar/sonar and wireless communication channels.

Prof. Scharf was Technical Program Chair for the 1980 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Denver, CO; Tutorials Chair for ICASSP 2001, Salt Lake City, UT; and Technical Program Chair for the Asilomar Conference on Signals, Systems, and Computers 2002. He is past-Chair of the Fellow Committee for the IEEE Signal Processing Society and serves on it Technical Committee for Sensor Arrays and Multichannel Signal Processing. He has received numerous awards for his research contributions to statistical signal processing, including a College Research Award, an IEEE Distinguished Lectureship, an IEEE Third Millennium Medal, and the Technical Achievement and Society Awards from the IEEE Signal Processing Society.