Detection of Rank-P Signals in Cognitive Radio Networks With Uncalibrated Multiple Antennas

David Ramírez, *Student Member, IEEE*, Gonzalo Vazquez-Vilar, *Member, IEEE*, Roberto López-Valcarce, *Member, IEEE*, Javier Vía, *Member, IEEE*, and Ignacio Santamaría, *Senior Member, IEEE*

***Abstract*—Spectrum sensing is a key component of the cognitive radio paradigm. Primary signals are typically detected with uncalibrated receivers at signal-to-noise ratios (SNRs) well below decodability levels. Multiantenna detectors exploit spatial independence of receiver thermal noise to boost detection performance and robustness. We study the problem of detecting a Gaussian signal with rank-P unknown spatial covariance matrix in spatially uncorrelated Gaussian noise with unknown covariance using multiple antennas. The generalized likelihood ratio test (GLRT) is derived for two scenarios. In the first one, the noises at all antennas are assumed to have the same (unknown) variance, whereas in the second, a generic diagonal noise covariance matrix is allowed in order to accommodate calibration uncertainties in the different antenna frontends. In the latter case, the GLRT statistic must be obtained numerically, for which an efficient method is presented. Furthermore, for asymptotically low SNR, it is shown that the GLRT does admit a closed form, and the resulting detector performs well in practice. Extensions are presented in order to account for unknown temporal correlation in both signal and noise, as well as frequency-selective channels.**

***Index Terms*—Cognitive radio, generalized likelihood ratio test (GLRT), macimum likelihood (ML) estimation, spectrum sensing.**

I. Introduction

Cognitive Radio (CR) has the potential to improve wireless spectrum usage and alleviate the apparent scarcity of spectral resources as seen today [1], [2]. The key idea behind CR is to allow opportunistic access to temporally and/or geographically unused licensed bands. Thus, spectrum sensing constitutes a key component in CR, in order to identify vacant channels and avoid interference to rightful license owners [3].

The wireless medium makes reliable detection of these users a very challenging task: Due to fading and shadowing phenomena, the received primary signal may be very weak, resulting in very low signal-to-noise ratio (SNR) operation conditions [4]. Any structure in the primary signal, such as the presence of pilots or cyclostationary features, could, in principle, be exploited for detection purposes. However, most such approaches require some level of synchronization with the primary signal, which cannot be guaranteed in very low SNRs [4]. In order to avoid these drawbacks, asynchronous detectors can be considered. The simpler asynchronous detectors, including the popular energy detector, require knowledge of the noise variance in order to compute the detection threshold. Any uncertainty regarding this parameter translates in severe performance degradation, so that the detection/false alarm requirements may not be satisfied [5].

This serious drawback motivates the search for asynchronous detectors robust to noise uncertainty, one possibility being the use of multiple-antenna sensors. Several authors have explored this strategy in order to enhance detection performance in the context of CR systems. Assuming a temporally white Gaussian model for both signal and noise, spatially white noise with the same unknown variance across antennas, and an unknown rank-1 spatial covariance matrix for the signal, a generalized likelihood ratio test (GLRT) was proposed in [6]. Other ad hoc detectors that have been proposed under this model include those in [7]–[9]. A GLRT framework has also been applied in [10] assuming certain prior information of the unknown parameters.

However, in practical scenarios, the spatial rank of the received signals may be larger than one. This is the case, for example, if multiple independent users (e.g., from adjacent cells) simultaneously access the same frequency channel. Alternatively, many state-of-the-art communication standards consider the simultaneous transmission of different data streams through multiple antennas to achieve multiplexing gain and/or the use of space-time codes to enhance spatial diversity. For these systems, the signal received at the multiantenna sensor will exhibit a spatial rank equal to the number of independent streams or the spatial size of the code, respectively. Examples range from broadcasting standards, such as the European DVB-T2 [11] which considers two-antenna space-time Alamouti codes, to point-to-multipoint standards, such as IEEE 802.11n [12],

Manuscript received September 13, 2010; revised February 10, 2011 and April 15, 2011; accepted April 15, 2011. Date of publication April 25, 2011; date of current version July 13, 2011. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Yimin D. Zhang. This work was supported by the Spanish Government's Ministerio de Ciencia e Innovación (MICINN) and the European Regional Development Fund (ERDF), under projects MULTIMIMO (TEC2007-68020-C04-02), COSIMA (TEC2010-19545-C04-03), DYNACS (TEC2010-21245-C02-02/TCM), SPROACTIVE (TEC2007-68094-C02-01/TCM), COMONSENS (CONSOLIDER-INGENIO 2010 CSD2008-00010), and FPU Grant AP2006-2965. This work was presented in part at the Second IAPR International Workshop on Cognitive Information Processing (CIP), Elba Island, Italy, June 2010, and at the Sixth IEEE Sensor Array and Multichannel Signal Processing, Jerusalem, Israel, October 2010.

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).

G. Vazquez-Vilar and R. López-Valcarce are with the Department of Signal Theory and Communications, University of Vigo, 36310 Vigo, Spain (e-mail: gvazquez@gts.tsc.uvigo.es; valcarce@gts.tsc.uvigo.es).

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.2011.2146779

IEEE 802.16 [13], or LTE [14], which support up to four transmit antennas. Hence, it is of interest to develop detectors for signals with spatial rank $P > 1$.

In this work, we focus our attention to the GLRT detection of vector-valued rank-P signals when the noise covariance is assumed unknown. In particular, the contributions of this paper are as follows.

1) We derive the GLRT detector of vector-valued rank-P signals for independent and identically distributed (i.i.d.) noises at each of the components when both signal and noise are assumed temporally white.
2) We formulate the GLRT detector for a similar scenario in which the noise components present different (unknown) variances. The GLRT for this case requires solving a non-convex optimization problem. We propose an efficient numerical method based on an alternating minimization approach to compute the exact GLRT statistic for noises with different variances and for arbitrary SNR. Additionally we show that this detector admits a closed form expression in the asymptotic low SNR regime.
3) The proposed GLRT detectors are rederived for signals with unknown PSDs, extending [15]–[17] to rank-P signals. This is of special interest in applications with frequency selective channels and/or temporally colored noise.

Our results are related to previous works. When the signal covariance matrix is unstructured, and the noise assumed i.i.d., the GLRT is the well known *test for sphericity* [18], which was applied to CR in [19] and [20]. As previously pointed out, for $P = 1$ and i.i.d. noises the GLRT is derived in [6] and its application to CR was presented in [21] and [22]. In [19] and [20], the authors derived the GLRT for rank-$P > 1$ under the assumption of i.i.d. noises with known variance. Furthermore, in [19], there is also an heuristic approach to handle the case of unknown σ^2, where the noise variance is estimated as the smallest eigenvalue of the sample covariance matrix.

In [23], the GLRT was derived for the case of an unstructured signal covariance matrix for non-i.i.d. noise. This detector was later applied to array signal processing in [24] and [25]. Other detectors which can handle different (unknown) noise variances have been proposed in [26]–[28]. However, they either assume rank-1 primary signals or unstructured signals.

Notation: We use light-face letters for scalars, and bold-face uppercase and lowercase letters for matrices and vectors, respectively. The elements of matrix $\mathbf{A}$ and vector $\mathbf{x}$ are denoted by $[\mathbf{A}]_{i,j}$ and x_i respectively. Calligraphic uppercase letters denote block-Toeplitz matrices. $\text{diag}(\mathbf{x})$ is a diagonal matrix with the elements of vector $\mathbf{x}$ on its diagonal. Table I summarizes other nonstandard notation.

TABLE I
NOTATION USED IN THE PAPER

$\hat{(\cdot)}$	Estimated matrices, vectors or scalars
$\det(\mathbf{A})$, $\text{tr}(\mathbf{A})$	Determinant and trace of $\mathbf{A}$
$\text{vec}(\mathbf{A})$	Column-wise vectorization of $\mathbf{A}$
$\mathbf{0}_L$	Zero $L \times 1$ vector or $L \times L$ matrix
$\mathbf{a}_k$	k-th column of matrix $\mathbf{A}$
$E[\cdot]$	Expectation operator
$\mathcal{F}(\cdot)$	Discrete-time Fourier transform
$\mathbf{x} \sim \mathcal{CN}(\boldsymbol{\mu}, \mathbf{R})$	Complex circular Gaussian random vector of mean $\boldsymbol{\mu}$ and covariance matrix $\mathbf{R}$
$\odot$	Hadamard product
$(h * s)[n]$	Convolution operation between $h[n]$ and $s[n]$
$\delta[m]$	Discrete delta impulse

II. PROBLEM FORMULATION

Consider a spectrum monitor equipped with L antennas which is to sense a given frequency channel. The received signals are downconverted and sampled at the Nyquist rate. No synchronization with any potentially present primary signal is assumed. Primary transmission, if present, is known to have spatial rank P, and a frequency-flat channel is assumed. Thus, for a single observation $\mathbf{x} \in \mathbb{C}^L$, the hypothesis testing problem can be written as

$$\begin{aligned} \mathcal{H}_1 &: \mathbf{x} = \mathbf{Hs} + \mathbf{v}, \\ \mathcal{H}_0 &: \mathbf{x} = \mathbf{v}, \end{aligned} \tag{1}$$

where $\mathbf{s} \in \mathbb{C}^P$ is the primary signal, $\mathbf{H} \in \mathbb{C}^{L \times P}$ is the unknown multiple-input multiple-output (MIMO) channel between the primary user and the spectrum sensor, and $\mathbf{v} \in \mathbb{C}^L$ is the additive noise, which is assumed to be zero-mean circular complex Gaussian spatially uncorrelated.

We model $\mathbf{s}$ as zero-mean circular complex Gaussian, which is particularly accurate if the primary transmitter uses orthogonal frequency division multiplexing (OFDM). Even if this is not the case, the Gaussian model leads to tractable analysis and useful detectors. It is assumed that $\mathbf{s}$ is spatially white and power-normalized, as any spatial correlation and scaling of the primary signal can be absorbed in the channel matrix $\mathbf{H}$. For the time being, we will assume that the primary signal and the noise are temporally white.[1] Taking this into account, the (spatial) covariance matrices of the primary signal and noise are given by

$$E\left[\mathbf{s}\mathbf{s}^H\right] = \mathbf{I}_P, \quad E\left[\mathbf{v}\mathbf{v}^H\right] = \mathbf{\Sigma}^2, \tag{2}$$

where $\mathbf{I}_P$ is the identity matrix of size $L \times L$ and $\mathbf{\Sigma}^2$ is an unknown diagonal covariance matrix. The detection problem in (1) amounts to testing between two different structures for the covariance of the vector-valued random variable $\mathbf{x} \sim \mathcal{CN}(\mathbf{0}_L, \mathbf{R})$:

$$\begin{aligned} \mathcal{H}_1 &: \mathbf{R} = \mathbf{H}\mathbf{H}^H + \mathbf{\Sigma}^2, \\ \mathcal{H}_0 &: \mathbf{R} = \mathbf{\Sigma}^2. \end{aligned} \tag{3}$$

That is, under $\mathcal{H}_0$ the covariance matrix $\mathbf{R}$ is diagonal whereas under $\mathcal{H}_1$ it is a rank-P matrix plus a diagonal one. We shall assume that $\mathbf{H}$ has full rank.

[1]These results will be extended in Section V to the case in which noise and primary signals are time series with *unknown* temporal structure.

III. Derivation of the GLRT for i.i.d. Noises

As a first step, we derive a detector for the simpler case of i.i.d. noises, i.e., $\mathbf{\Sigma}^2 = \sigma^2\mathbf{I}$, which amounts to saying that all the L analog frontends are perfectly calibrated. As there are unknown parameters under both hypotheses, the Neyman–Pearson detector is not implementable for this composite test. Therefore, we adopt a GLRT approach, since it usually results in simple detectors with good performance [29].

We shall consider $M \geq L$ snapshots $\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}$. Assuming that the channel remains constant during the sensing period, these can be regarded as i.i.d. realizations of $\mathbf{x} \sim \mathcal{CN}(\mathbf{0}_L, \mathbf{R})$. The likelihood is given by the product of the individual pdfs, i.e.,

$$p\left(\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}; \mathbf{R}\right) = \frac{1}{\pi^{LM} \det\left(\mathbf{R}\right)^M} \exp\left\{-M \mathrm{tr}\left(\hat{\mathbf{R}}\mathbf{R}^{-1}\right)\right\} \quad (4)$$

where $\hat{\mathbf{R}} = \frac{1}{M}\sum_{m=0}^{M-1} \mathbf{x}_m \mathbf{x}_m^H$ is the sample covariance matrix. The GLRT for $\mathcal{H}_0 : \mathbf{R} = \sigma^2\mathbf{I}$ versus $\mathcal{H}_1 : \mathbf{R} = \mathbf{H}\mathbf{H}^H + \sigma^2\mathbf{I}$ is based on the generalized likelihood ratio $\mathscr{L}$ [29]

$$\mathscr{L} = \frac{\max_{\sigma^2} p\left(\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}; \sigma^2\right)}{\max_{\mathbf{H},\sigma^2} p\left(\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}; \mathbf{H}, \sigma^2\right)} \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\gtrless}} \eta, \quad (5)$$

with η a threshold. First, the maximum likelihood (ML) estimate of the noise variance under $\mathcal{H}_0$ is given by

$$\hat{\sigma}^2 = \frac{1}{L}\mathrm{tr}\left(\hat{\mathbf{R}}\right). \quad (6)$$

In order to obtain the ML estimates under $\mathcal{H}_1$, we consider two cases depending on the rank P.

Lemma 1: If $P \geq L-1$, the ML estimates of $\mathbf{H}$ and σ^2 satisfy $\hat{\mathbf{H}}\hat{\mathbf{H}}^H + \hat{\sigma}^2\mathbf{I} = \hat{\mathbf{R}}$.

Proof: For $P \geq L-1$, $\mathbf{R} = \mathbf{H}\mathbf{H}^H + \sigma^2\mathbf{I}$ has no additional structure besides being positive definite Hermitian. In that case, the log-likelihood is maximized for $\mathbf{R} = \hat{\mathbf{R}}$, as shown in [30]. ■

Thus, for $P \geq L-1$, the GLRT is the well-known *Sphericity test* [18]

$$\log \mathscr{L} = ML \log\left[\frac{\det^{\frac{1}{L}}\left(\hat{\mathbf{R}}\right)}{\frac{1}{L}\mathrm{trace}\left(\hat{\mathbf{R}}\right)}\right]. \quad (7)$$

When $P < L-1$, the low-rank structure of the primary signal can be used to further improve the detection. In that case, to obtain the ML estimates under $\mathcal{H}_1$, let $\mathbf{H}\mathbf{H}^H = \mathbf{U}\mathbf{\Psi}^2\mathbf{U}^H$ be an eigenvalue decomposition (EVD) of $\mathbf{H}\mathbf{H}^H$, with

$$\mathbf{\Psi}^2 = \mathrm{diag}(\psi_1^2, \psi_2^2, \ldots, \psi_P^2, 0, 0, \ldots, 0), \quad (8)$$

with $\psi_1 \geq \psi_2 \geq \cdots \geq \psi_P$.

Lemma 2: Let $\hat{\mathbf{R}} = \mathbf{W}\mathrm{diag}\left(\lambda_1, \ldots, \lambda_L\right)\mathbf{W}^H$ be an EV decomposition of the sample covariance matrix, with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_L$. For $P < L-1$, the ML estimates of $\mathbf{U}$, $\mathbf{\Psi}^2$ and σ^2 under $\mathcal{H}_1$ are given by

$$\hat{\mathbf{U}} = \mathbf{W}, \quad \hat{\sigma}^2 = \frac{1}{L-P}\sum_{k=P+1}^{L} \lambda_k \quad (9)$$

$$\hat{\psi}_i^2 = \lambda_i - \hat{\sigma}^2, \quad i = 1, \ldots, P. \quad (10)$$

Proof: This result was proved by Anderson in [31]. ■

Taking into account (6) and Lemma 2, the log-GLRT for $P < L-1$ is given, after some straightforward manipulations, by

$$\log \mathscr{L} = ML \log\left[\frac{\left(\prod_{i=1}^{L} \lambda_i\right)^{\frac{1}{L}}}{\frac{1}{L}\sum_{i=1}^{L} \lambda_i}\right] - M(L-P)\log\left[\frac{\left(\prod_{i=P+1}^{L} \lambda_i\right)^{\frac{1}{(L-P)}}}{\frac{1}{L-P}\sum_{i=P+1}^{L} \lambda_i}\right] \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\gtrless}} \eta. \quad (11)$$

Note that the logarithmic terms in (11) are functions of the ratio between the geometric and arithmetic means of all eigenvalues and the $L-P$ smallest eigenvalues of $\hat{\mathbf{R}}$, respectively. The first term is the statistic of the sphericity test (7), whereas the second term can be seen as a test for the sphericity of the noise subspace, or as a *reference* for sphericity due to finite sample size effects (since as $M \to \infty$, then $\hat{\mathbf{R}} \to \mathbf{R}$ and thus $\lambda_i \to \sigma^2$ for $i = P+1, \ldots, L$, so that the second term in (11) goes to zero). Thus, the log-GLRT may be seen as a *sphericity ratio* (quotient between the sphericity statistics of the sample covariance matrix and its noise subspace).

Remark 1: The GLRT in (11) generalizes the results in [6], [21], and [22] obtained for the special case of $P = 1$.

IV. Derivation of the GLRT for Non-i.i.d. Noises

In this section, we derive the GLRT for the more involved model of non-i.i.d. noises. In this case, the only constraint on $\mathbf{\Sigma}^2$ is being diagonal with positive entries. Let us start by the ML estimate of $\mathbf{\Sigma}^2$ under $\mathcal{H}_0$, which is given by [24], [25]

$$\hat{\mathbf{\Sigma}}^2 = \mathrm{diag}\left([\hat{\mathbf{R}}]_{1,1}, \ldots, [\hat{\mathbf{R}}]_{L,L}\right) = \hat{\mathbf{D}}. \quad (12)$$

Similar to the case of i.i.d. noises, we study first the effect of the signal rank P on the ML estimate of the covariance matrix.

Lemma 3: If $P \geq L - \sqrt{L}$, the ML estimates of $\mathbf{H}$ and $\mathbf{\Sigma}^2$ under $\mathcal{H}_1$ satisfy $\hat{\mathbf{H}}\hat{\mathbf{H}}^H + \hat{\mathbf{\Sigma}}^2 = \hat{\mathbf{R}}$.

Proof: The proof can be found in [16], [24]. It hinges on the fact that if $P \geq L - \sqrt{L}$, then $\mathbf{H}\mathbf{H}^H + \mathbf{\Sigma}^2$ has no further structure beyond being positive definite Hermitian. ■

Hence, for $P \geq L - \sqrt{L}$, the GLRT is given by the *Hadamard ratio* of the sample covariance matrix [23]–[25]:

$$\mathscr{L} = \frac{\det\left(\hat{\mathbf{R}}\right)}{\prod_{i=1}^{L} [\hat{\mathbf{R}}]_{i,i}}. \quad (13)$$

If $P < L - \sqrt{L}$, the low-rank structure of the primary signal can be further exploited. In order to simplify the derivation of the ML estimates under $\mathcal{H}_1$, let $\hat{\mathbf{R}}_{\mathbf{\Sigma}} = \mathbf{\Sigma}^{-1}\hat{\mathbf{R}}\mathbf{\Sigma}^{-1}$ (the *whitened* sample covariance matrix) and $\mathbf{H}_{\boldsymbol{\Sigma}} = \boldsymbol{\Sigma}^{-1}\mathbf{H}$. We can rewrite the log-likelihood as

$$\begin{aligned}\log p\left(\mathbf{x}_0,\ldots,\mathbf{x}_{M-1};\mathbf{H}_{\boldsymbol{\Sigma}},\boldsymbol{\Sigma}^2\right) &= -LM\log\pi - M\log\det\left(\mathbf{H}_{\boldsymbol{\Sigma}}\mathbf{H}_{\boldsymbol{\Sigma}}^H+\mathbf{I}\right)\\ &\quad - M\log\det\left(\boldsymbol{\Sigma}^2\right)\\ &\quad - M\mathrm{tr}\left[\hat{\mathbf{R}}_{\boldsymbol{\Sigma}}\left(\mathbf{H}_{\boldsymbol{\Sigma}}\mathbf{H}_{\boldsymbol{\Sigma}}^H+\mathbf{I}\right)^{-1}\right]. \end{aligned} \tag{14}$$

Let $\mathbf{H}_{\boldsymbol{\Sigma}}\mathbf{H}_{\boldsymbol{\Sigma}}^H = \mathbf{G}\boldsymbol{\Phi}^2\mathbf{G}^H$ be the EVD of $\mathbf{H}_{\boldsymbol{\Sigma}}\mathbf{H}_{\boldsymbol{\Sigma}}^H$. The ML estimates of $\mathbf{G}$ and $\boldsymbol{\Phi}^2$ are given next.

Lemma 4: Let $\hat{\mathbf{R}}_{\boldsymbol{\Sigma}} = \mathbf{Q}\mathrm{diag}\left(\gamma_1,\ldots,\gamma_L\right)\mathbf{Q}^H$ be the EVD of $\hat{\mathbf{R}}_{\boldsymbol{\Sigma}}$, with $\gamma_1 \geq \cdots \geq \gamma_L$. The ML estimates of $\mathbf{G}$ and $\boldsymbol{\Phi}^2 = \mathrm{diag}\left(\phi_1,\ldots,\phi_L\right)$ (which are functions of $\boldsymbol{\Sigma}^2$) are

$$\hat{\mathbf{G}} = \mathbf{Q}, \tag{15}$$

$$\hat{\phi}_i^2 = \begin{cases}\gamma_i - 1, & i = 1,\ldots P,\\ 0, & i = P+1,\ldots,L.\end{cases} \tag{16}$$

Proof: Once $\hat{\mathbf{R}}$ and $\mathbf{H}$ have been prewhitened, the problem reduces to the i.i.d. case and, therefore, the proof follows the same lines as those in [31]. ■

Finally, replacing the ML estimate of $\mathbf{H}_{\boldsymbol{\Sigma}}\mathbf{H}_{\boldsymbol{\Sigma}}^H$ into (14), we obtain

$$\begin{aligned}\log p\left(\mathbf{x}_0,\ldots,\mathbf{x}_{M-1};\boldsymbol{\Sigma}^2\right) &= -LM\log\pi - MP\\ &\quad -M\log\det\left(\hat{\mathbf{R}}\right) - M\sum_{i=P+1}^{L}\left[\gamma_i - \log\gamma_i\right]. \end{aligned}\tag{17}$$

As previously mentioned, for unstructured covariance matrices, i.e., $P \geq L - \sqrt{L}$, there does exist a closed-form GLRT given by (13). However, to the best of our knowledge, the maximization of (17) with respect to $\boldsymbol{\Sigma}^2$ does not admit a closed-form solution if $P < L - \sqrt{L}$. We present two different approaches: an alternating optimization scheme and a closed-form GLRT detector obtained in the limit of asymptotically small SNR.

A. Alternating Optimization

The ML estimation problem in (14) can be written as

$$\begin{aligned}\underset{H_{\Sigma},\boldsymbol{\Sigma}}{\text{minimize}}\quad & \mathrm{tr}\left(\hat{\mathbf{R}}\boldsymbol{\Sigma}^{-1}\mathbf{R}_{\boldsymbol{\Sigma}}^{-1}\boldsymbol{\Sigma}^{-1}\right)\\ & - \log\det\left(\boldsymbol{\Sigma}^{-2}\right) + \log\det\mathbf{R}_{\boldsymbol{\Sigma}},\\ \text{subject to}\quad & \mathbf{R}_{\boldsymbol{\Sigma}} = \mathbf{I}_L + \mathbf{H}_{\boldsymbol{\Sigma}}\mathbf{H}_{\boldsymbol{\Sigma}}^H,\\ & [\boldsymbol{\Sigma}]_{i,i} \geq 0. \end{aligned}\tag{18}$$

While this optimization problem is nonconvex, it is possible to partition the free variables in two different sets to obtain an alternating optimization scheme. Then, we will alternatively perform the minimization over each set of parameters while the remaining ones are held fixed. Since at each step the value of the cost function can only decrease the method is guaranteed to converge to a (local) minimum.

From (18), we note that the individual minimization with respect to $\boldsymbol{\Sigma}$ (considering $\mathbf{H}_{\boldsymbol{\Sigma}}$ fixed) and with respect to $\mathbf{H}_{\boldsymbol{\Sigma}}$ (considering $\boldsymbol{\Sigma}$ fixed) can be easily written as convex problems individually and, therefore, they can be efficiently solved.

Algorithm 1: Iterative Estimation of $\mathbf{H}_{\boldsymbol{\Sigma}}$ and $\boldsymbol{\Sigma}$ Via Alternating Optimization

Input: Starting point $\boldsymbol{\alpha}_{(0)}$ and $\hat{\mathbf{R}}$.
Output: ML estimates of $\mathbf{H}_{\boldsymbol{\Sigma}}$ and $\boldsymbol{\Sigma}$.
Initialize: $n = 0$;
repeat
  Compute $\boldsymbol{\Sigma}_{(n)}^{-1} = \mathrm{diag}(\boldsymbol{\alpha}_{(n)})$;
  Obtain $\hat{\mathbf{R}}_{\boldsymbol{\Sigma}}^{(n+1)} = \boldsymbol{\Sigma}_{(n)}^{-1}\hat{\mathbf{R}}\boldsymbol{\Sigma}_{(n)}^{-1}$ and its EVD;
  Compute $\mathbf{H}_{\boldsymbol{\Sigma}}^{(n+1)}$ from **(19)** (fixed $\boldsymbol{\Sigma}_{(n)}^{-1}$);
  Solve **(21)** to obtain $\boldsymbol{\alpha}_{(n+1)}$ (fixed $\mathbf{H}_{\boldsymbol{\Sigma}}^{(n+1)}$);
  Update $n = n + 1$;
until Convergence;

1) Minimization With Respect to $\mathbf{H}_{\boldsymbol{\Sigma}}$*:* For fixed $\boldsymbol{\Sigma}$, the optimal $\mathbf{H}_{\boldsymbol{\Sigma}}$ minimizing (18) is (up to a right multiplication by a unitary matrix) given by Lemma 4, that is

$$\hat{\mathbf{H}}_{\boldsymbol{\Sigma}} = \left[\mathbf{q}_1 \quad \cdots \quad \mathbf{q}_P\right]\left(\mathrm{diag}\left(\gamma_1,\ldots,\gamma_P\right) - \mathbf{I}_P\right)^{\frac{1}{2}}. \tag{19}$$

2) Minimization With Respect to $\boldsymbol{\Sigma}$*:* For fixed $\mathbf{H}_{\boldsymbol{\Sigma}}$ the minimization problem in (18) reduces to

$$\begin{aligned}\underset{\boldsymbol{\Sigma}}{\text{minimize}}\quad & \mathrm{tr}\left(\hat{\mathbf{R}}\boldsymbol{\Sigma}^{-1}\mathbf{R}_{\boldsymbol{\Sigma}}^{-1}\boldsymbol{\Sigma}^{-1}\right) - \log\det\left(\boldsymbol{\Sigma}^{-2}\right),\\ \text{subject to}\quad & [\boldsymbol{\Sigma}]_{i,i} \geq 0. \end{aligned}\tag{20}$$

Defining the vector $\boldsymbol{\alpha} = \left[[\boldsymbol{\Sigma}^{-1}]_{1,1},\ldots,[\boldsymbol{\Sigma}^{-1}]_{L,L}\right]^T$, the trace term in (20) can be reorganized to obtain an equivalent minimization problem given by

$$\begin{aligned}\underset{\boldsymbol{\alpha}}{\text{minimize}}\quad & \boldsymbol{\alpha}^T(\hat{\mathbf{R}}^T \odot \mathbf{R}_{\boldsymbol{\Sigma}}^{-1})\boldsymbol{\alpha} - \sum_{i=1}^{L}\log\alpha_i^2,\\ \text{subject to}\quad & \alpha_i \geq 0. \end{aligned}\tag{21}$$

Note that, given the trace term in (20), the matrix $\hat{\mathbf{R}}^T \odot \mathbf{R}_{\boldsymbol{\Sigma}}^{-1}$ is positive semidefinite. Hence, the problem (21) is convex with respect to the parameter vector $\boldsymbol{\alpha}$ and, therefore, it can be efficiently solved using any convex optimization solver.

The proposed alternating minimization algorithm is summarized in Algorithm 1. While the alternating minimization approach does not guarantee that the global maximizer of the log-likelihood is found, in the numerical experiments conducted this detector shows good performance.

B. Low SNR Approximation of the GLRT

The usefulness of the detector given in Algorithm 1 in practical settings may be hindered by its complexity. In this context, simpler closed-form detectors become of practical interest. Now, we derive a closed-form expression for the GLRT in the low SNR regime, of particular interest in CR applications. As the SNR goes to zero, the covariance matrix will become close to diagonal, and thus it is possible to approximate the ML estimate of $\boldsymbol{\Sigma}^2$ as $\hat{\boldsymbol{\Sigma}}^2 \approx \hat{\mathbf{D}}$ defined in (12). Substituting this back into (17), we obtain the final compressed log-likelihood:

$$\begin{aligned}\log p\left(\mathbf{x}_0,\ldots,\mathbf{x}_{M-1}\right) &= -LM\log\pi - MP\\ &\quad -M\log\det\left(\hat{\mathbf{R}}\right) - M\sum_{i=P+1}^{L}\left[\beta_i - \log\beta_i\right]. \end{aligned}\tag{22}$$

where β_i is the ith largest eigenvalue of the sample spatial coherence matrix $\hat{\mathbf{C}} = \hat{\mathbf{D}}^{-1/2}\hat{\mathbf{R}}\hat{\mathbf{D}}^{-1/2}$. Then, the asymptotic log-GLRT is

$$\log \mathscr{L} \approx M \sum_{i=1}^{P} [\log \beta_i - \beta_i] + MP \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\gtrless}} \eta. \tag{23}$$

Alternatively, (23) can be rewritten as

$$\log \mathscr{L} \approx MP + M \log \prod_{i=1}^{P} \beta_i e^{-\beta_i} \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\gtrless}} \eta, \tag{24}$$

and, thus, the test statistic is seen to be given by the product of the P largest eigenvalues of $\hat{\mathbf{C}}$, each equalized by an exponential term. Note that $\beta e^{-\beta}$ is maximum at $\beta = 1$. Hence, the statistic $\prod_{i=1}^{P} \beta_i e^{-\beta_i}$ measures, in some sense, how far the vector of the P largest eigenvalues $[\beta_1 \cdots \beta_P]$ is from the vector of all ones. Note that (23) yields a closed-form test, in contrast with the iterative scheme presented in the previous section.

V. Extension to Time Series With Temporal Structure

We extend now the detectors in Sections III and IV in order to deal with frequency-selective channels, as well as *unknown* temporal correlation in signals and noise.

A. Problem Formulation

The detection problem can be expressed now as

$$\begin{aligned} \mathcal{H}_1 &: \mathbf{x}[n] = (\mathbf{H} * \mathbf{s})[n] + \mathbf{v}[n], \quad n = 0, \ldots, N-1, \\ \mathcal{H}_0 &: \mathbf{x}[n] = \mathbf{v}[n], \quad n = 0, \ldots, N-1, \end{aligned} \tag{25}$$

where $\mathbf{s}[n] \in \mathbb{C}^P$ is the wide sense stationary (WSS) zero-mean circular complex Gaussian primary signal; $\mathbf{H}[n] \in \mathbb{C}^{L \times P}$ is the frequency-selective MIMO channel between the primary user and the spectrum monitor; and $\mathbf{v}[n] \in \mathbb{C}^L$ is the additive noise vector, which is assumed to be WSS zero-mean circular complex Gaussian and spatially uncorrelated, i.e., $E[v_i[n]v_k^*[m]] = 0$ for $i \neq k$ and $\forall n, m$. No assumptions are made about the *temporal* correlation of the primary signal or the noise processes. Note, however, that any spatial and temporal correlation present in the signal can be absorbed in the unknown channel without altering the model. Therefore, the matrix-valued covariance function of the primary signal and the noise are given by

$$E\left[\mathbf{s}[n]\mathbf{s}^H[n-m]\right] = \mathbf{I}\delta[m] \tag{26}$$

$$E\left[\mathbf{v}[n]\mathbf{v}^H[n-m]\right] = \mathbf{\Sigma}^2[m] \tag{27}$$

where $\mathbf{\Sigma}^2[m]$ is a diagonal matrix for all values of m.

Let us introduce the data matrix

$$\mathbf{X} = [\mathbf{x}[0] \quad \mathbf{x}[1] \quad \ldots \quad \mathbf{x}[N-1]] \in \mathbb{C}^{L \times N}, \tag{28}$$

where the ith row contains N samples of the time series $\{x_i[n]\}$ at the ith antenna, and the nth column is the nth sample of the vector-valued time series. The vector $\mathbf{z} = \mathrm{vec}(\mathbf{X})$ stacks the columns of $\mathbf{X}$, and in view of the WSS assumption, its block-Toeplitz covariance matrix $\mathcal{R} \in \mathbb{C}^{LN \times LN}$ is given by

$$\mathcal{R} = \begin{bmatrix} \mathbf{R}[0] & \mathbf{R}[-1] & \cdots & \mathbf{R}[-N+1] \\ \mathbf{R}[1] & \mathbf{R}[0] & \cdots & \mathbf{R}[-N+2] \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{R}[N-1] & \mathbf{R}[N-2] & \cdots & \mathbf{R}[0] \end{bmatrix}, \tag{29}$$

where $\mathbf{R}[m] = E\left[\mathbf{x}[n]\mathbf{x}^H[n-m]\right]$ is a matrix-valued covariance function. Therefore, under the Gaussian assumption, the hypothesis testing problem becomes

$$\begin{aligned} \mathcal{H}_1 &: \mathbf{z} \sim \mathcal{CN}(\mathbf{0}_{LN}, \mathcal{R}_1), \\ \mathcal{H}_0 &: \mathbf{z} \sim \mathcal{CN}(\mathbf{0}_{LN}, \mathcal{R}_0). \end{aligned} \tag{30}$$

We are testing two different block-Toeplitz matrices where each block has a different structure under each hypothesis. Under $\mathcal{H}_0$ each block $\mathbf{R}[m] = \mathbf{\Sigma}^2[m]$ is diagonal, whereas, under $\mathcal{H}_1$, it is given by

$$\mathbf{R}[m] = \sum_k \mathbf{H}[k]\mathbf{H}^H[k-m] + \mathbf{\Sigma}^2[m]. \tag{31}$$

B. Asymptotic Log-likelihood

The structure of $\mathcal{R}_1$ induced by the rank-P primary signal, along with the block-Toeplitz structure, prevents the ML estimation of $\mathcal{R}_1$ in closed-form, even in the case of i.i.d. noises. In fact, the ML estimation of Toeplitz covariance matrices is known to be a nonconvex problem with no closed-form solution [32]. To overcome this limitation, we introduce Theorem 1, which states the convergence (in the mean square sense) between the log-likelihood and its asymptotic version. This theorem allows us to work with the log-likelihood in the frequency domain which, as we will see, simplifies the derivation of the GLRT (actually, the asymptotic GLRT). The asymptotic log-likelihood is now a function of the estimated and theoretical power spectral density (PSD) matrices instead of being a function of the estimated and theoretical covariance matrices. Additionally, the proposed asymptotic log-likelihood is an extension of Whittle's likelihood [33], [34] to multivariate Gaussian processes.

Let us introduce some definitions before presenting the theorem. Consider an experiment producing M $(M \geq L)$ independent realizations[2] of the data vector $\mathbf{z}$. Then, its log-likelihood is given by

$$\begin{aligned} \log p(\mathbf{z}_0, \ldots, \mathbf{z}_{M-1}; \mathcal{R}) &= -LNM \log \pi \\ &\quad - M \log \det(\mathcal{R}) - M \mathrm{tr}\left(\hat{\mathbf{R}}\mathcal{R}^{-1}\right), \end{aligned} \tag{32}$$

and the asymptotic $(N \longrightarrow \infty)$ log-likelihood is

$$\begin{aligned} &\log p\left(\mathbf{z}_0, \ldots, \mathbf{z}_{M-1}; \mathbf{S}\left(e^{j\theta}\right)\right) \\ &= -LNM \log \pi - NM \int_{-\pi}^{\pi} \log \det\left(\mathbf{S}\left(e^{j\theta}\right)\right) \frac{d\theta}{2\pi} \\ &\quad - NM \int_{-\pi}^{\pi} \mathrm{tr}\left(\hat{\mathbf{S}}\left(e^{j\theta}\right)\mathbf{S}^{-1}\left(e^{j\theta}\right)\right) \frac{d\theta}{2\pi}, \end{aligned} \tag{33}$$

[2]In this case, the *snapshots* are matrix-valued.

where $\mathcal{R}$ is the theoretical block-Toeplitz covariance matrix, $\mathbf{S}\left(e^{j\theta}\right) = \mathcal{F}\left(\mathbf{R}[m]\right)$ is the theoretical PSD matrix, and their sample estimates are $\hat{\mathbf{R}} = \frac{1}{M}\sum_{i=0}^{M-1}\mathbf{z}_i\mathbf{z}_i^H$ and $\hat{\mathbf{S}}\left(e^{j\theta}\right) = \frac{1}{M}\sum_{i=0}^{M-1}\mathbf{x}_i\left(e^{j\theta}\right)\mathbf{x}_i^H\left(e^{j\theta}\right)$, with $\mathbf{x}_i\left(e^{j\theta}\right) = \frac{1}{\sqrt{N}}\sum_{n=0}^{N-1}\mathbf{x}_i[n]\,e^{-j\theta n}$.

Theorem 1: As $N\rightarrow\infty$, the asymptotic log-likelihood converges in the mean square sense to the true log-likelihood:

$$\lim_{N\rightarrow\infty} E\left[\left|\frac{1}{N}\left[\log p\left(\mathbf{z}_0,\ldots,\mathbf{z}_{M-1};\mathcal{R}\right) - \log p\left(\mathbf{z}_0,\ldots,\mathbf{z}_{M-1};\mathbf{S}\left(e^{j\theta}\right)\right)\right]\right|^2\right] = 0. \quad (34)$$

Proof: The proof can be found in the Appendix. ■

As a direct consequence of Theorem 1, the hypothesis test asymptotically becomes

$$\mathcal{H}_1 : \mathbf{x}\left(e^{j\theta}\right) \sim \mathcal{CN}\left(\mathbf{0},\mathbf{S}_1\left(e^{j\theta}\right)\right), \quad (35)$$

$$\mathcal{H}_0 : \mathbf{x}\left(e^{j\theta}\right) \sim \mathcal{CN}\left(\mathbf{0},\mathbf{S}_0\left(e^{j\theta}\right)\right), \quad (36)$$

where $\mathbf{S}_1\left(e^{j\theta}\right) = \mathbf{H}\left(e^{j\theta}\right)\mathbf{H}^H\left(e^{j\theta}\right) + \mathbf{\Sigma}^2\left(e^{j\theta}\right)$, $\mathbf{S}_0\left(e^{j\theta}\right) = \mathbf{\Sigma}^2\left(e^{j\theta}\right)$, $\mathbf{H}\left(e^{j\theta}\right)$ is the Fourier transform of the MIMO channel and $\mathbf{\Sigma}^2\left(e^{j\theta}\right)$ is a diagonal matrix which contains the PSD of the noises. Therefore, under $\mathcal{H}_0$ the PSD matrix is diagonal, whereas under $\mathcal{H}_1$ it is the sum of a rank-P matrix plus a diagonal one.

C. Derivation of the GLRT

Without imposing any temporal structure to the time series, the ML estimation of the unknown parameters can be carried out on a frequency-by-frequency basis. Thus, we can directly apply the results from Sections III and IV. First, the log-GLRT for the case of noises with equal PSDs, i.e., $\mathbf{\Sigma}^2(e^{j\theta}) = S_v(e^{j\theta})\mathbf{I}$ and assuming $P < L-1$, is given by

$$\log\mathscr{L} = NML\int_{-\pi}^{\pi}\log\left[\frac{\left(\prod_{i=1}^{L}\lambda_i\left(e^{j\theta}\right)\right)^{\frac{1}{L}}}{\frac{1}{L}\sum_{i=1}^{L}\lambda_i\left(e^{j\theta}\right)}\right]\frac{d\theta}{2\pi} - NM(L-P)\int_{-\pi}^{\pi}\log\left[\frac{\left(\prod_{i=P+1}^{L}\lambda_i\left(e^{j\theta}\right)\right)^{\frac{1}{(L-P)}}}{\frac{1}{L-P}\sum_{i=P+1}^{L}\lambda_i\left(e^{j\theta}\right)}\right]\frac{d\theta}{2\pi} \quad (37)$$

where $\lambda_i\left(e^{j\theta}\right)$ is the ith largest eigenvalue of $\hat{\mathbf{S}}\left(e^{j\theta}\right)$. The asymptotic log-GLRT for time series with unknown temporal structure is the integral of the frequency-wise GLRT statistic for white vector-valued time-series, derived in (11).

For the case of noises with different PSDs along the antennas, and assuming $P < L-\sqrt{L}$, the log-GLRT in the low SNR region[3] is approximately given by

$$\log\mathscr{L} \approx NM\int_{-\pi}^{\pi}\sum_{i=1}^{P}\log\beta_i\left(e^{j\theta}\right)\frac{d\theta}{2\pi} - NM\int_{-\pi}^{\pi}\sum_{i=1}^{P}\beta_i\left(e^{j\theta}\right)\frac{d\theta}{2\pi} + NMP \quad (38)$$

[3]For other SNR regimes, it would be possible to apply, on a frequency-by-frequency basis, the alternating minimization algorithm presented in Section IV-A.

where β_i is the ith largest eigenvalue of the coherence matrix

$$\hat{\mathbf{C}}\left(e^{j\theta}\right) = \hat{\mathbf{D}}^{-1/2}\left(e^{j\theta}\right)\hat{\mathbf{S}}\left(e^{j\theta}\right)\hat{\mathbf{D}}^{-1/2}\left(e^{j\theta}\right) \quad (39)$$

and

$$\hat{\mathbf{D}}\left(e^{j\theta}\right) = \operatorname{diag}\left([\hat{\mathbf{S}}\left(e^{j\theta}\right)]_{1,1},\ldots,[\hat{\mathbf{S}}\left(e^{j\theta}\right)]_{L,L}\right). \quad (40)$$

The asymptotic log-GLRT is, again, the integral of the GLRT for vector-valued random variables.

VI. Numerical Results

In this section we evaluate the performance of the proposed algorithms under different scenarios by means of Monte Carlo simulations. First, we consider frequency-flat channels and temporally white signals and noises. Unless otherwise specified, the noise level at each antenna is fixed for each experiment, and for each Monte Carlo realization the entries of the channel matrix $\mathbf{H}$ are independently drawn from a Gaussian distribution (thus obtaining a Rayleigh fading scenario) and scaled so that the SNR is constant during the experiment

$$\text{SNR (dB)} = 10\log_{10}\frac{\text{tr}(\mathbf{H}\mathbf{H}^H)}{\text{tr}(\mathbf{\Sigma}^2)}. \quad (41)$$

We evaluate two detectors derived under the i.i.d. noise assumption: the proposed GLRT statistic in (11) denoted here by *i.i.d.-GLRT*, and the sphericity test or GLRT for nonstructured primary signals [18] (denoted as *Sphericity*). In addition, three detectors derived for uncalibrated receivers ($\mathbf{\Sigma}^2$ diagonal with positive entries) are also evaluated: the proposed alternating optimization scheme from Algorithm 1 denoted here as *alterntng-GLRT*,[4] the asymptotic closed-form detector in (23) (*asympt-GLRT*), the Hadamard ratio test [23] or GLRT for unstructured primary signals (13) (Hadamard). Additionally, we also include two heuristic detectors for comparison: the detector based on statistical covariances [28, Alg. 1] (*Covariance*) and that of (33) in [19] (Lim *et al.*).

A. Detection Performance for Rank-P Primary Signals

First we compare the performance of the different schemes in terms of the spatial rank of the signal. Fig. 1 shows the missed detection probability in a scenario with $\text{SNR} = -6$ dB, $L = 6$ antennas for primary signals with rank $P = 1,\ldots,5$, for i.i.d. and non-i.i.d. noises. Note that, for increasing P the structure present in the covariance matrix decreases. This effect translates into the performance degradation for all the detectors under study.

From Fig. 1(a) and (b), we can see that both for i.i.d. and non-i.i.d. noises the proposed GLRT detectors are the best performing detectors for arbitrary values of P. While the GLRTs for $P = 1$ present poor performance if the actual rank of the signal is larger than one, Sphericity and Hadamard ratio tests

[4]Given the observed convergence properties, the iterations are stopped when the cost improvement between iterations is less than 10^{-5} with a maximum of 100 allowed iterations. As starting point we use an estimate given by the scaled low SNR asymptotic solution $\boldsymbol{\alpha}_{(0)} = \sqrt{\frac{L}{(L-P)}}\,[\,[\hat{\mathbf{D}}]_{1,1}^{-1},\ldots,[\hat{\mathbf{D}}]_{L,L}^{-1}\,]^T$.

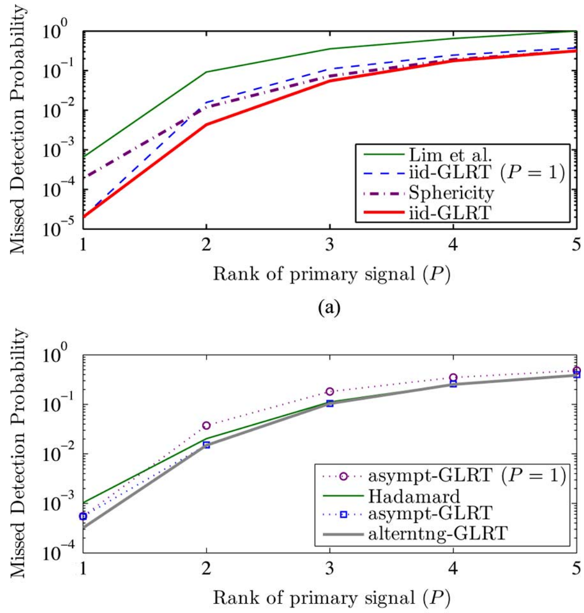

Fig. 1. Missed detection probability versus P: (a) i.i.d. noise and (b) non-i.i.d. noise.

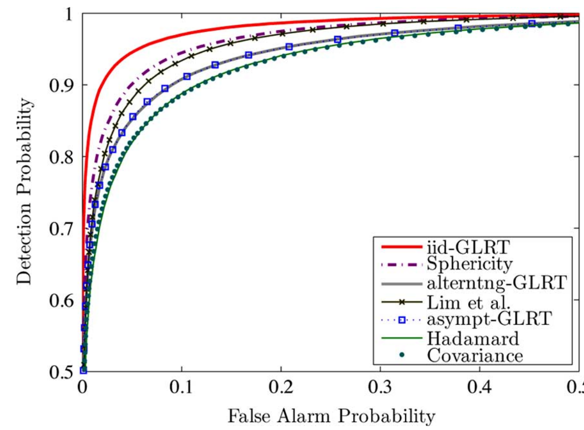

Fig. 2. ROC for the different detectors ($\mathrm{SNR} = -8$ dB, $L = 4$ antennas, $M = 128$ samples) without noise power mismatch (noise level at each antenna equal to 0 dB).

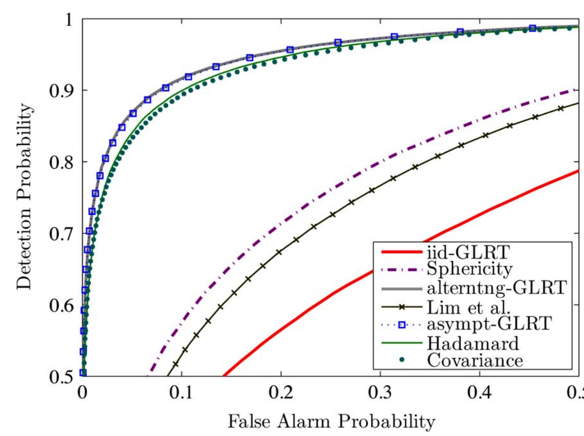

Fig. 3. ROC for the different detectors ($\mathrm{SNR} = -8$ dB, $L = 4$ antennas, $M = 128$ samples) with noise power mismatch (noise powers at each antenna equal to 0, -1, 1.5, and -0.5 dB, respectively).

(which do not assume any structure on the primary signal) degrade for strong structure, i.e., small P. It is interesting to note however that as the rank of the signal grows (for $P \geq 4$) the Sphericity and Hadamard ratio tests offer similar performance to that of the rank-based detectors at a lower computational cost. Regarding the heuristic detectors, the covariance based detector (*Covariance*) presents virtually the same performance as the Hadamard ratio test and it was not included in the plot for clarity. On the other hand, the detector by Lim *et al.* [19] can be seen to present a poor performance for all values of P, which can be attributed to the heuristic estimation of the noise variance.

Finally, it is interesting to note that for $P > 1$, the advantage of the iterative scheme *alterntng-GLRT* over the asymptotic GLRT decreases. This can be explained from the fact that, as the total SNR is divided among a growing number of dimensions, the effective SNR per dimension decreases and one gets closer to the asymptotic regime for which *asympt-GLRT* was derived.

B. Noise Mismatch Effect on the Detection Performance

We now investigate the effect of a noise level mismatch at the different antennas on the different detectors. To focus on this effect we fix $P = 1$. Fig. 2 shows the corresponding receiver operating characteristic (ROC) curves in a scenario with i.i.d. noises. Note that the *i.i.d.-GLRT* test, corresponding to the GLRT under this model (rank-1 signals and i.i.d. noises), yields the best detection performance, whereas the detectors designed for nonuniform noise variances suffer a noticeable penalty. From the detectors designed for uncalibrated receivers, it is seen that the GLRT based schemes, both asymptotic and iterative, behave similarly and outperform the Hadamard ratio detector. The heuristic detector based on statistical covariances [28] presents almost the same performance as the Hadamard ratio test, while the detector by Lim *et al.* suffers a penalty compared to the GLRT for the same model.

In Fig. 3, we represent the same scenario as in Fig. 2 except for set of noise variances at each of the antennas, now given by 0, -1, 1.5, and -0.5 dB. Note that the behavior of the detectors designed for uncalibrated receivers has not changed with respect to that in Fig. 2; however, the detectors based on the i.i.d. noise assumption suffer an important performance degradation.

C. Asymptotic GLRT Performance for Finite SNR Values

Although the asymptotic GLRT detector (*asympt-GLRT*) given by (23) is appealing due to its computational simplicity, it is not clear how much can be gained when the iterative scheme (*alterntng-GLRT*) is used in order to implement the exact GLRT. Fig. 4 shows the missed detection probability of the detectors versus the SNR in a scenario similar to the previous subsection ($P = 1$, $L = 4$, $M = 128$, different noise levels at each of the antennas fixed to 0, -1, 1.5, and -0.5 dB, respectively). The probability of false alarm is fixed to $P_{\mathrm{FA}} = 0.01$ and 0.1. In Fig. 4, it is seen that for very low SNR values the asymptotic detector presents the same performance as the alternating minimization scheme. However, as the SNR increases, the GLRT outperforms the detector derived for asymptotically low SNR, as it could be expected. Note, however, that the performance loss of the asymptotic detector

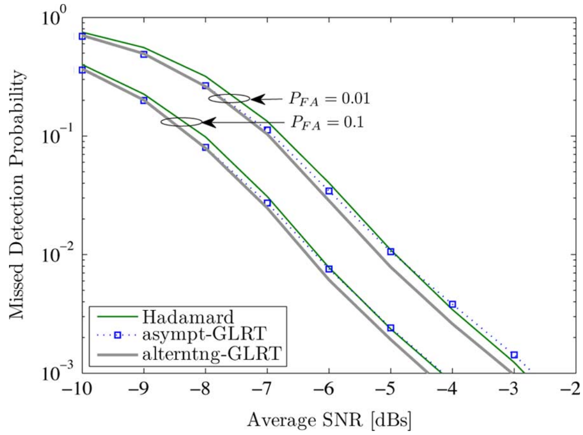

Fig. 4. Missed detection probability versus SNR for different detectors. Same scenario as in Fig. 3.

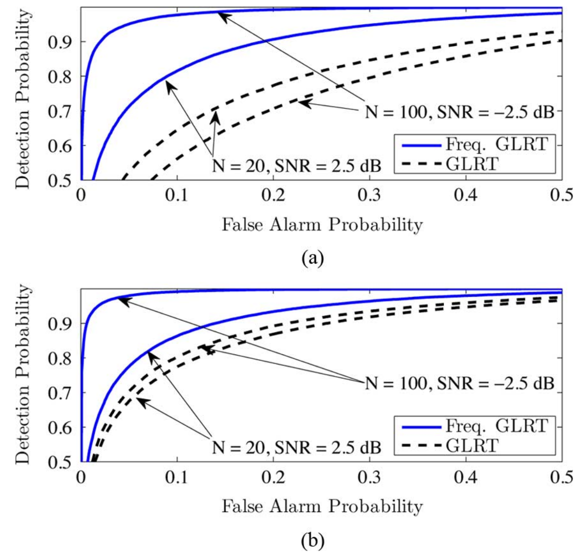

Fig. 5. Performance comparison of the frequency-domain GLRT against the GLRT: (a) i.i.d. noise and (b) non-i.i.d. noise.

is rather small, and therefore it offers a good tradeoff between performance and complexity.

D. Detection Performance in Environments With Unknown Temporal Structure

Now, we evaluate the performance of the proposed detectors in environments with unknown temporal structure. We consider an scenario with primary signals of rank $P = 2$, a receiver with $L = 5$ antennas, which captures $M = 5$ realizations of length $N = 100$ or $N = 20$ for the detection process. The transmitted signals use OFDM modulation, have a bandwidth of 7.61 MHz, and undergo propagation through a 5×2 uncorrelated frequency-selective channel with exponential power delay profile and delay spread 0.779 μs [35], which is fixed through the experiment. At the receiver, temporally white noises are added and the signals are sampled at 16 MHz. The SNR, defined in a frequency selective environment as

$$\text{SNR (dB)} = 10\log_{10}\frac{\int_{-\pi}^{\pi}\operatorname{tr}\left(\mathbf{H}\left(e^{j\theta}\right)\mathbf{H}^H\left(e^{j\theta}\right)\right)\frac{d\theta}{2\pi}}{\int_{-\pi}^{\pi}\operatorname{tr}\left(\boldsymbol{\Sigma}^2\left(e^{j\theta}\right)\right)\frac{d\theta}{2\pi}}, \quad (42)$$

is given in the figures.

First, we assume i.i.d. noises. The ROC curves for the GLRT [given by (11)] and the frequency-domain GLRT for i.i.d. noises [which assumes frequency-selective channels and is given by (37)] are shown in Fig. 5(a). The advantage of exploiting the temporal structure of the time series is clear. However, as can be seen in the figure, this improvement becomes smaller when the number of available samples decrease. Similar conclusions can be drawn from Fig. 5(b), which shows the performance of the two GLRT detectors for uncalibrated receivers, given by (23) and (38). In this experiment the noise variance values at the different antennas were drawn from a uniform distribution in linear scale to obtain mismatches no larger than 7.5 dB.

VII. Conclusion

We have derived the GLRTs for the problem of detecting vector-valued rank-P signals when the noise covariance is assumed unknown. In particular, when the noise at each of the components is assumed i.i.d. and for uncorrelated noises non necessarily i.i.d.. As it turns out, detectors derived under the assumption of i.i.d. noises are not robust to a mismatch in noise levels across the antennas. Although the GLRT for the more challenging non-i.i.d. noise case does not admit a closed form, one can resort to numerical optimization techniques. A closed-form low SNR approximation of the GLRT has also been presented, providing a tradeoff between complexity and performance. These detectors include as particular cases several previous schemes derived either for $P = 1$ or for large P. These results have also been extended to a model considering vector-valued time series with unknown PSDs, of interest in applications with frequency selective channels and/or temporally colored noise.

All of the detectors considered here assume knowledge of the signal rank P. While this may be reasonable in some contexts, for example if the space-time coding scheme used by primary transmitters is known, there are scenarios in which P is unknown, for example if it is related to the number of primary users simultaneously transmitting. Future research should consider estimation of P [36] and primary signal detection jointly.

Appendix
Proof of Theorem 1

In this Appendix, we prove the mean square convergence of the log-likelihoods, which can be seen as an extension of Whittle's likelihood [33], [34] to multivariate processes. We shall start by the conditions under which the theorem holds:

c.1 The block-Toeplitz matrices are generated by continuous symbols, i.e., each matrix block is given by the Fourier coefficients of an $L \times L$ continuous matrix function.

c.2 The power spectral matrices are positive definite.

In order to proceed, let us rewrite (32) as follows:

$$\log p\left(\mathbf{z}_0,\ldots,\mathbf{z}_{M-1};\mathcal{R}\right) = -LNM\log\pi - M\log\det\left(\mathcal{R}\right) - \sum_{m=0}^{M-1}\mathbf{z}_m^H\mathcal{R}^{-1}\mathbf{z}_m, \quad (43)$$

where $\mathcal{R}$ is given by (29). Additionally, let us introduce the block-Toeplitz matrix $\tilde{\mathcal{R}}$ which is given by

$$\tilde{\mathcal{R}} = \begin{bmatrix} \tilde{\mathbf{R}}[0] & \tilde{\mathbf{R}}[-1] & \cdots & \tilde{\mathbf{R}}[-N+1] \\ \tilde{\mathbf{R}}[1] & \tilde{\mathbf{R}}[0] & \cdots & \tilde{\mathbf{R}}[-N+2] \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{\mathbf{R}}[N-1] & \tilde{\mathbf{R}}[N-2] & \cdots & \tilde{\mathbf{R}}[0] \end{bmatrix}, \quad (44)$$

where $\tilde{\mathbf{R}}[m] = \mathcal{F}^{-1}\left[\mathbf{S}^{-1}\left(e^{j\theta}\right)\right]$; and the sum of quadratic forms of the matrix $\tilde{\mathcal{R}}$

$$\sum_{m=0}^{M-1} \mathbf{z}_m^H \tilde{\mathcal{R}} \mathbf{z}_m = NM \int_{-\pi}^{\pi} \operatorname{tr}\left(\hat{\mathbf{S}}\left(e^{j\theta}\right)\mathbf{S}^{-1}\left(e^{j\theta}\right)\right)\frac{d\theta}{2\pi}. \quad (45)$$

The mean-square convergence of a random variable is defined as follows (see, e.g., [37]):

$$\lim_{N\to\infty} E\left[\left|\frac{1}{N}\left[\log p\left(\mathbf{z}_0,\ldots,\mathbf{z}_{M-1};\mathcal{R}\right) - \log p\left(\mathbf{z}_0,\ldots,\mathbf{z}_{M-1};\mathbf{S}\left(e^{j\theta}\right)\right)\right]\right|^2\right] = 0. \quad (46)$$

Substituting (32) and (43) into the left-hand side of (46), one has

$$\begin{aligned} &\lim_{N\to\infty} E\left[\left|\frac{1}{N}\left[-M\log\det\left(\mathcal{R}\right) - \sum_{m=0}^{M-1}\mathbf{z}_m^H\mathcal{R}^{-1}\mathbf{z}_m + NM\int_{-\pi}^{\pi}\log\det\left(\mathbf{S}\left(e^{j\theta}\right)\right)\frac{d\theta}{2\pi} - NM\int_{-\pi}^{\pi}\operatorname{tr}\left(\hat{\mathbf{S}}\left(e^{j\theta}\right)\mathbf{S}^{-1}\left(e^{j\theta}\right)\right)\frac{d\theta}{2\pi}\right]\right|^2\right] \\ &\overset{(a)}{\leq} \lim_{N\to\infty} 2M\left|\frac{1}{N}\log\det\left(\mathcal{R}\right) - \int_{-\pi}^{\pi}\log\det\left(\mathbf{S}\left(e^{j\theta}\right)\right)\frac{d\theta}{2\pi}\right|^2 \\ &\quad + \lim_{N\to\infty} 2E\left[\left|\frac{1}{N}\sum_{m=0}^{M-1}\mathbf{z}_m^H\mathcal{R}^{-1}\mathbf{z}_m - M\int_{-\pi}^{\pi}\operatorname{tr}\left(\hat{\mathbf{S}}\left(e^{j\theta}\right)\mathbf{S}^{-1}\left(e^{j\theta}\right)\right)\frac{d\theta}{2\pi}\right|^2\right], \quad (47) \end{aligned}$$

where (a) follows from [37, Th. 8, p. 287]. We shall proceed by splitting the proof in two parts, one for each term in the right-hand side of (47). It is easy to show that under c.1 and c.2, the following holds (see [38, Th. 6])

$$\lim_{N\to\infty}\frac{1}{N}\log\det\left(\mathcal{R}\right) - \int_{-\pi}^{\pi}\log\det\left(\mathbf{S}\left(e^{j\theta}\right)\right)\frac{d\theta}{2\pi} = 0. \quad (48)$$

Then, taking into account that

$$\lim_{N\to\infty} a_N = 0 \quad \Rightarrow \quad \lim_{N\to\infty} |a_N|^2 = 0 \quad (49)$$

where a_N is any sequence of real numbers, it follows that

$$\lim_{N\to\infty} 2M\left|\frac{1}{N}\log\det\left(\mathcal{R}\right) - \int_{-\pi}^{\pi}\log\det\left(\mathbf{S}\left(e^{j\theta}\right)\right)\frac{d\theta}{2\pi}\right|^2 = 0. \quad (50)$$

Now, taking into account (45), it is readily shown that

$$\begin{aligned} &\lim_{N\to\infty} E\left[\left|\frac{1}{N}\sum_{m=0}^{M-1}\mathbf{z}_m^H\mathcal{R}^{-1}\mathbf{z}_m - M\int_{-\pi}^{\pi}\operatorname{tr}\left(\hat{\mathbf{S}}\left(e^{j\theta}\right)\mathbf{S}^{-1}\left(e^{j\theta}\right)\right)\frac{d\theta}{2\pi}\right|^2\right] \\ &= \lim_{N\to\infty} E\left[\left|\frac{1}{N}\sum_{m=0}^{M-1}\mathbf{z}_m^H\left(\mathcal{R}^{-1} - \tilde{\mathcal{R}}\right)\mathbf{z}_m\right|^2\right]. \quad (51) \end{aligned}$$

We now prove that (50) converges to zero. First, we must note that it is a mean of the square of a sum of quadratic forms. Therefore, taking into account that the quadratic forms are uncorrelated and the formulas for the mean and variance of a quadratic form of a multivariate normal distribution [39], it follows that

$$\begin{aligned} &\lim_{N\to\infty} E\left[\left|\frac{1}{N}\sum_{m=0}^{M-1}\mathbf{z}_m^H\left(\mathcal{R}^{-1} - \tilde{\mathcal{R}}\right)\mathbf{z}_m\right|^2\right] \\ &= \lim_{N\to\infty}\frac{1}{N^2}\left\{M^2\left(\operatorname{tr}\left[\mathcal{R}\left(\mathcal{R}^{-1} - \tilde{\mathcal{R}}\right)\right]\right)^2 + 2M\operatorname{tr}\left[\mathcal{R}\left(\mathcal{R}^{-1} - \tilde{\mathcal{R}}\right)\mathcal{R}\left(\mathcal{R}^{-1} - \tilde{\mathcal{R}}\right)\right]\right\} = 0 \quad (52) \end{aligned}$$

where we have applied [38, Th. 3 and Th. 5] and (49). Therefore, (51) converges to zero, i.e.,

$$\lim_{N\to\infty} E\left[\left|\frac{1}{N}\sum_{m=0}^{M-1}\mathbf{z}_m^H\mathcal{R}^{-1}\mathbf{z}_m - M\int_{-\pi}^{\pi}\operatorname{tr}\left(\hat{\mathbf{S}}\left(e^{j\theta}\right)\mathbf{S}^{-1}\left(e^{j\theta}\right)\right)\frac{d\theta}{2\pi}\right|^2\right] = 0. \quad (53)$$

Finally, the proof follows from (50) and (53). ■

References

[1] J. Mitola and G. Q. Maguire, Jr., "Cognitive radio: Making software radios more personal," *IEEE Pers. Commun.*, vol. 6, pp. 13–18, Aug. 1999.

[2] J. M. Peha, "Sharing spectrum through spectrum policy reform and cognitive radio," *Proc. IEEE*, vol. 97, no. 4, pp. 708–719, Apr. 2009.

[3] I. Akyildiz, W.-Y. Lee, M. Vuran, and S. Mohanty, "A survey on spectrum management in cognitive radio networks," *IEEE Commun. Mag.*, vol. 46, no. 4, pp. 40–48, Apr. 2008.

[4] D. Cabric, "Addressing the feasibility of cognitive radios," *IEEE Signal Process. Mag.*, vol. 25, no. 6, pp. 85–93, Nov. 2008.

[5] R. Tandra and A. Sahai, "SNR walls for signal detection," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 4–17, Feb. 2008.

[6] O. Besson, S. Kraut, and L. L. Scharf, "Detection of an unknown rank-one component in white noise," *IEEE Trans. Signal Process.*, vol. 54, no. 7, pp. 2835–2839, Jul. 2006.

[7] Y. Zeng, Y.-C. Liang, and R. Zhang, "Blindly combined energy detection for spectrum sensing in cognitive radio," *IEEE Signal Process. Lett.*, vol. 15, pp. 649–652, 2008.

[8] M. Alamgir, M. Faulkner, J. Gao, and P. Conder, "Signal detection for cognitive radio using multiple antennas," in *Proc. IEEE Int. Symp. Wireless Commun. Syst.*, Reykjavic, Iceland, Oct. 2008.

[9] 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.

[10] J. Font-Segura and X. Wang, “GLRT-based spectrum sensing for cognitive radio with prior information,” *IEEE Trans. Commun.*, vol. 58, no. 7, pp. 2137–2146, Jul. 2010.
[11] *ETSI, Digital Video Broadcasting (DVB); Frame Structure Channel Coding and Modulation for a Second Generation Digital Terrestrial Television Broadcasting System (DVB-T2)*, ETSI EN 302 755, 2009.
[12] *IEEE Computer Society, Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. Amendment 5: Enhancements for Higher Throughput*, IEEE Standard 802.11n-2009. Part 11, 2009.
[13] IEEE Computer Society and the IEEE Microwave Theory and Techniques Society, IEEE Standard 802.16-2009. Part 16: Air Interface for Broadband Wireless Access Systems 2009.
[14] 3GPP, LTE Release 9. 2009.
[15] D. Ramírez, J. Vía, and I. Santamaría, “Multiantenna spectrum sensing: The case of wideband rank-one primary signals,” in *Proc. IEEE Sensor Array Multichannel Signal Process. Workshop*, Israel, Oct. 2010.
[16] D. Ramírez, J. Vía, I. Santamaría, and L. L. Scharf, “Detection of spatially correlated Gaussian time series,” *IEEE Trans. Signal Process*, vol. 58, no. 10, pp. 5006–5015, Oct. 2010.
[17] D. Ramirez, 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.*, Dallas, TX, Mar. 2010.
[18] J. Mauchly, “Significance test for sphericity of a normal n-variate distribution,” *Ann. Math. Statist.*, vol. 11, pp. 204–209, 1940.
[19] T. J. Lim, R. Zhang, Y.-C. Liang, and Y. Zeng, “GLRT-based spectrum sensing for cognitive radio,” in *Proc. Global Commun. Conf.*, New Orleans, LA, 2008.
[20] R. Zhang, T. J. Lim, Y.-C. Liang, and Y. Zeng, “Multi-antenna based spectrum sensing for cognitive radios: A GLRT approach,” *IEEE Trans. Commun.*, vol. 58, no. 1, pp. 84–88, Jan. 2010.
[21] A. Taherpour, M. Nasiri-Kenari, and S. Gazor, “Multiple antenna spectrum sensing in cognitive radios,” *IEEE Trans. Wireless Commun.*, vol. 9, no. 2, pp. 814–823, Feb. 2010.
[22] P. Wang, J. Fang, N. Han, and H. Li, “Multiantenna-assisted spectrum sensing for cognitive radio,” *IEEE Trans. Veh. Technol.*, vol. 59, no. 4, pp. 1791–1800, May 2010.
[23] S. Wilks, “On the independence of k sets of normally distributed statistical variables,” *Econometrica*, vol. 3, pp. 309–325, 1935.
[24] 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.
[25] 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.
[26] A.-J. Boonstra and A.-J. van der Veen, “Gain calibration methods for radio telescope arrays,” *IEEE Trans. Signal Process.*, vol. 51, no. 1, pp. 25–38, Jan. 2003.
[27] R. López-Valcarce, G. Vazquez-Vilar, and J. Sala, “Multiantenna spectrum sensing for cognitive radio: Overcoming noise uncertainty,” presented at the Int. Work. Cognitive Inf. Process., Elba Island, Italy, Jun. 2010.
[28] Y. Zeng and Y.-C. Liang, “Spectrum-sensing algorithms for cognitive radio based on statistical covariances,” *IEEE Trans. Veh. Technol.*, vol. 58, no. 4, pp. 1804–1815, May 2009.
[29] K. V. Mardia, J. T. Kent, and J. M. Bibby, *Multivariate Analysis*. New York: Academic, 1979.
[30] J. R. Magnus and H. Neudecker, *Matrix Differential Calculus With Applications in Statistics and Econometrics*, 2nd ed. New York: Wiley, 1999.
[31] T. W. Anderson, “Asymptotic theory for principal component analysis,” *Ann. Math. Stat.*, vol. 34, no. 1, pp. 122–148, Mar. 1963.
[32] J. P. Burg, D. G. Luenberger, and D. L. Wenger, “Estimation of structured covariance matrices,” *Proc. IEEE*, vol. 70, no. 9, pp. 963–974, 1982.
[33] P. Whittle, “Gaussian estimation in stationary time series,” *Bull. Inst. Int. Statist.*, vol. 39, pp. 105–129, 1962.
[34] P. Whittle, “Estimation and information in time series analysis,” *Skand. Aktuar.*, vol. 35, pp. 48–60, 1952.
[35] M. Failli, “Digital land mobile radio communications,” Cost Action 207 Final Rep. Luxemburg, Luxemburg, 1989.
[36] M. Chiani and M. Win, “Estimating the number of signals observed by multiple sensors,” presented at the Int. Work. Cognitive Inf. Process., Elba Island, Italy, Jun. 2010.
[37] G. Grimmett and D. Stirzaker, *Probability and Random Processes*. Oxford, U.K.: Oxford Science, 1992.
[38] 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.
[39] K. Dzhaparidze, *Parameter Estimation and Hypothesis Testing in Spectral Analysis of Stationary Time Series*. New York: Springer-Verlag, 1986.

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.

Gonzalo Vazquez-Vilar (S’04–M’08) received the Telecommunications Engineer degree from the Universidad de Vigo, Vigo, Spain, in 2004 and the M.S. degree in electrical engineering from Stanford University, Stanford, CA, in 2008.

During 2005 and 2006, he was respectively with the University of Applied Sciences of Saarbruecken and with Siemens A.G., both in Saarbruecken, Germany. Currently he is working towards the Ph.D. degree at the University of Vigo on the area of signal processing applied to communications. His research focuses on wireless MIMO systems, cognitive radio, and interference management.

Roberto López-Valcarce (M’01) received the undergraduate diploma in telecommunication engineering from the Universidad de Vigo, Vigo, Spain, in 1995, and the M.S. and Ph.D. degrees in electrical engineering from the University of Iowa, Iowa City, in 1998 and 2000, respectively.

During 1995, he was a Project Engineer with Intelsis. He was a Postdoctoral Fellow of the Spanish Ministry of Science and Technology from 2001 to 2006. During that period, he was with the Signal Theory and Communications Department, Universidad de Vigo, where he is currently an Associate Professor. His main research interests lie in the area of adaptive signal processing, digital communications, and cognitive radio. He holds two European patents in the field of adaptive processing applied to communications systems.

Dr. López-Valcarce was a recipient of a 2005 Best Paper Award of the IEEE Signal Processing Society. Since 2008 he serves as an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING.

Javier Vía (S'04–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, and at the Department of Electronics and Computer Engineering (Hong Kong University of Science and Technology). His current research interests include blind channel estimation and equalization in wireless communication systems, multivariate statistical analysis, quaternion signal processing, and kernel methods.

Dr. Vía has actively participated in several European and Spanish research projects.

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 (University of Florida), and at the Wireless Networking and Communications Group (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.