Multiantenna GLR detection of rank-one signals with known power spectrum in white noise with unknown spatial correlation

Josep Sala†, *Senior Member, IEEE,* Gonzalo Vazquez-Vilar‡, *Student Member, IEEE,* Roberto López-Valcarce‡, *Member, IEEE,*

†Dept. of Signal Theory and Communications, Technical University of Catalonia, 08034 Barcelona, SPAIN
E-mail: josep.sala@upc.edu
‡Dept. of Signal Theory and Communications, University of Vigo, 36310 Vigo, SPAIN
E-mail: {gvazquez,valcarce}@gts.uvigo.es
submitted (minors after AQ) to: IEEE Transactions on Signal Processing
EDICS: SPC-DETC (Detection, Estimation and Demodulation), SSP-DETC (Detection)

Abstract—Multiple-antenna detection of a Gaussian signal with spatial rank one in temporally white Gaussian noise with arbitrary and unknown spatial covariance is considered. This is motivated by spectrum sensing problems in the context of Dynamic Spectrum Access in which several secondary networks coexist but do not cooperate, creating a background of spatially correlated broadband interference. When the temporal correlation of the signal of interest is assumed known up to a scale factor, the corresponding Generalized Likelihood Ratio Test is shown to yield a scalar optimization problem. Closed-form expressions of the test are obtained for the general signal spectrum case in the low signal-to-noise ratio (SNR) regime, as well as for signals with binary-valued power spectrum in arbitrary SNR. The two resulting detectors turn out to be equivalent. An asymptotic approximation to the test distribution for the low-SNR regime is derived, closely matching empirical results from spectrum sensing simulation experiments.

Index Terms—GLR test, detection, multiantenna array, correlated noise, noise uncertainty, cognitive radio, spectrum sensing, spectral flatness measure, Capon beamformer

I. INTRODUCTION

Array processing for signal detection and parameter estimation has been a topic of deep research interest for decades. Many techniques have been proposed under the assumption that the noise is Gaussian and spatially uncorrelated (or has a known spatial structure that allows prewhitening) [1]–[5]. However, in many practical cases the noise field may present unknown spatial correlation, due to co-channel interference in communications applications, and to clutter and jamming in radar array processing [6]–[10]. Our main motivation stems from the problem of spectrum sensing for Dynamic Spectrum Access (DSA) in licensed bands [11], [12]. DSA aims at a more efficient usage of spectrum by allowing unlicensed (or *secondary*) users to opportunistically access channels in those bands, as long as interference to licensed (or *primary*) users is avoided; this can be achieved by dynamically tuning to different carrier frequencies to sense the radio environment for unused channels. Spectrum sensing schemes must be robust to wireless propagation phenomena such as large- and small-scale fading as well as to interference. The first two issues can be dealt with by resorting to cooperative detection strategies [13] and multiantenna detectors [14]–[17], respectively.

Multiantenna spectrum sensing methods usually exploit spatial correlation of the signal, assuming a spatially uncorrelated noise field. Whereas this assumption may be adequate in early stages of DSA-based schemes, in which a single secondary network coexists with the primary system, this is not necessarily so when several independent secondary systems access the same frequency band. Unless strict coordination among such systems is imposed, no quiet periods will be available in order to sense primary activity in a given channel. Hence, such activity will have to be detected in the presence of a background noise consisting not only of thermal noise, but also the aggregate interference from the rest of secondary networks. The physical layer of secondary systems is likely to make use of channel fragmentation, aggregation, and/or bonding in order to enhance spectrum utilization (such is the case for e.g. the IEEE 802.22 standard for opportunistic access to TV white spaces [18]), so that the background interference will likely be broadband (with respect to the bandwidth of a primary channel), and potentially correlated in space, due to the well-localized origin of secondary transmissions. Other potential sources of broadband interference of course exist, e.g. the emissions of Broadband over Power Line (BPL) systems in the HF and VHF bands [19]. After bandlimiting by the channel selection filter of the receiver, the contribution due to the broadband interference will tend to be Gaussian distributed, and will be modeled as temporally uncorrelated in this work. More sophisticated noise models, e.g. autoregressive in time and/or space [20], may be of interest in certain scenarios but fall out of the scope of this paper.

On the other hand, it may well be possible for the spectrum sensor to have knowledge about the (normalized) Power Spectrum Density (PSD) of primary signals, as the emission masks of many licensed services (e.g., broadcast and cellular networks) are in the public domain. Knowledge of the signal

PSD has been exploited in [21] assuming a single antenna and known noise variance; and in [22], [23] for multiple antennas with spatially uncorrelated noise with known common variance, and respectively known and unknown channel gains.

Motivated by the above considerations, we derive the Generalized Likelihood Ratio (GLR) detector and its Receiver Operating Characteristic (ROC) for a random signal with spatial rank one and known PSD in temporally white Gaussian noise of arbitrary spatial covariance. The nuisance parameters are the unknown antenna gains for the signal of interest and the unknown spatial covariance of the noise. No structure is assumed on the latter, thus avoiding artificial constraints on the unknown spatial rank of the interference, for the sake of robustness and detector complexity. With this model, intuition suggests that detection should be feasible as long as the temporal correlation of the signal differs from that of the noise process, i.e., the signal is not temporally white. Indeed, the performance of the GLR detector will be shown to be dependent on the spectral flatness measure at the output of the minimum variance beamformer.

Following common practice, the Gaussian model is adopted for the primary signal. The resulting model is tractable and leads to useful detectors even under different signal distributions. In addition, if the primary system uses multicarrier modulation, the Gaussian assumption becomes quite accurate for the number of subcarriers usually found in practice [24].

Previous works in the literature on signal detection are related to the scheme presented in this paper. The single-antenna scenario considering real- rather than complex-valued data in the proposed signal model has been treated in [30] using the Expectation-Maximization (EM) algorithm. On the other hand, the problem of signal detection with multiple antennas and spatially correlated noise has been addressed by the Generalized Multivariate Analysis of Variance (GMANOVA) approach [25] when the signal is deterministic dependent on unknown parameters. Adaptive subspace detectors can be used if the signal has low spatial rank and signal-free training data are available for estimating the unknown noise covariance; see [26] for the case of deterministic signals, and [27] for Gaussian signal components. Other approaches for the Gaussian signal model exploit rotational invariance of the noise field (if it exists) [28] or assume a parametric model of the noise [20]. When the spatial correlation of noise is unstructured, and lacking training data, as in our model, it may be possible to exploit the *temporal* correlation properties of signal and noise as in [29], where it is assumed that noise has a much shorter temporal correlation length than that of the signals. However, none of these works derives the GLR test under the proposed signal model. Moreover, the detector herein described can be validated with the result in [30] in the single antenna setting, leading to a simpler univariate optimization scheme, and can significantly outperform other detectors which apply under the same model, as shown by simulation results for that from [29].

The paper is structured as follows. The signal model is given in Sec. II, together with a summary of the main results to be developed in the sequel. The derivation of the exact GLR detector is carried out in Sec. III. The resulting scheme involves the optimization of a data-dependent term with respect to a scalar variable, which cannot be solved in closed form in general. Nevertheless, in Sec. IV two important particular instances are considered that will result in a closed-form scheme, namely the low-SNR and binary-valued signal PSD cases. The asymptotic performance of the GLR detector is analytically derived in Sec. V. Numerical results are given in Sec. VI, and Sec. VII presents some closing remarks.

Notation: lower- and upper-case boldface symbols denote vectors and matrices, respectively. The k-th unit vector is denoted by $\boldsymbol{e}_k$. The trace, determinant, transpose, conjugate, conjugate transpose (Hermitian), adjoint, and Moore-Penrose pseudoinverse of $\boldsymbol{A}$ are denoted by $\operatorname{tr}\boldsymbol{A}$, $\det\boldsymbol{A}$, $\boldsymbol{A}^T$, $\boldsymbol{A}^*$, $\boldsymbol{A}^H$, $\operatorname{adj}(\boldsymbol{A})$ and $\boldsymbol{A}^\dagger$ respectively. $\operatorname{diag}(\boldsymbol{A})$ is a diagonal matrix with diagonal equal to that of $\boldsymbol{A}$. The Kronecker product of $\boldsymbol{A}$ and $\boldsymbol{B}$ is denoted by $\boldsymbol{A}\otimes\boldsymbol{B}$. The column-wise vectorization of $\boldsymbol{A}$ is denoted by $\operatorname{vec}(\boldsymbol{A})$. For Hermitian $\boldsymbol{A}$, its largest (resp. smallest) eigenvalue is denoted by $\lambda_{\max}[\boldsymbol{A}]$ (resp. $\lambda_{\min}[\boldsymbol{A}]$); if $\boldsymbol{A}$ is positive (semi)definite, $\boldsymbol{A}^{1/2}$ denotes its unique Hermitian square root. The natural (base e) logarithm is denoted by $\log$. Finally, $\boldsymbol{x}\sim\mathcal{CN}(\boldsymbol{\mu},\boldsymbol{P})$ indicates that $\boldsymbol{x}$ is circularly complex Gaussian with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{P}$.

II. PROBLEM FORMULATION

A. Signal Model

The general signal model considered in this paper is as follows:

$$\boldsymbol{y}_n = \boldsymbol{h}s_n + \boldsymbol{z}_n \in \mathbb{C}^M, \qquad n \in \{1, \cdots, N\}, \qquad (1)$$

where $\boldsymbol{y}_n$ represents a snapshot (sample of sensor outputs at time n) from an array of M sensors, $\boldsymbol{h}\in\mathbb{C}^M$ is the unknown spatial signature vector, and $\{s_n\}\in\mathbb{C}$, $\{\boldsymbol{z}_n\}\in\mathbb{C}^M$ are independent zero-mean complex circular Gaussian processes. The signal vector $\boldsymbol{s}\doteq[\begin{array}{ccc} s_1 & \cdots & s_N \end{array}]^T\in\mathbb{C}^N$ is zero mean with known temporal covariance $\boldsymbol{C}\doteq\mathbb{E}[\boldsymbol{s}\boldsymbol{s}^H]$, i.e., $\boldsymbol{s}\sim\mathcal{CN}(\boldsymbol{0},\boldsymbol{C})$. The process $\{s_n\}$ is assumed wide-sense stationary with PSD $S_{ss}(e^{j\omega})$; thus, $\boldsymbol{C}$ is Hermitian Toeplitz. In addition, we assume that $\mathbb{E}[|s_n|^2]=1$ without loss of generality (since any scaling factor can be absorbed into $\boldsymbol{h}$), so that $\boldsymbol{C}$ has an all-ones diagonal. On the other hand, the process $\{\boldsymbol{z}_n\}$ models the background noise and interference. It is assumed to be temporally white[1], but spatially correlated with *unknown and unstructured* spatial covariance matrix $\boldsymbol{\Sigma}$: $\mathbb{E}[\boldsymbol{z}_k\boldsymbol{z}_l^H]=\boldsymbol{\Sigma}$ if $k=l$, and zero otherwise. Therefore, defining the $MN\times 1$ data and noise vectors respectively as $\boldsymbol{y}\doteq[\begin{array}{ccc}\boldsymbol{y}_1^T & \cdots & \boldsymbol{y}_N^T\end{array}]^T$ and $\boldsymbol{z}\doteq[\begin{array}{ccc}\boldsymbol{z}_1^T & \cdots & \boldsymbol{z}_N^T\end{array}]^T$, one can rewrite (1) as

$$\boldsymbol{y} = \boldsymbol{s}\otimes\boldsymbol{h} + \boldsymbol{z}, \qquad (2)$$

which is Gaussian with covariance $\boldsymbol{R}\doteq\mathbb{E}[\boldsymbol{y}\boldsymbol{y}^H]=\boldsymbol{C}\otimes\boldsymbol{h}\boldsymbol{h}^H+\boldsymbol{I}_N\otimes\boldsymbol{\Sigma}$.

The problem considered is the detection of the presence of signal $\boldsymbol{s}$ given the data vector $\boldsymbol{y}$. Thus, the corresponding

---

[1]If $\boldsymbol{z}_n$ is temporally nonwhite but with *known* temporal correlation, the proposed model can be obtained after a temporal prewhitening step.

hypothesis test is given by

$$\mathcal{H}_0 \ : \ \boldsymbol{y} \ \sim \ \mathcal{CN}(\boldsymbol{0}, \boldsymbol{R}_0), \quad \boldsymbol{R}_0 = \boldsymbol{I}_N \otimes \boldsymbol{\Sigma}, \tag{3}$$

$$\mathcal{H}_1 \ : \ \boldsymbol{y} \ \sim \ \mathcal{CN}(\boldsymbol{0}, \boldsymbol{R}_1), \quad \boldsymbol{R}_1 = \boldsymbol{C} \otimes \boldsymbol{h}\boldsymbol{h}^H + \boldsymbol{I}_N \otimes \boldsymbol{\Sigma}$$
$$\boldsymbol{h} \neq \boldsymbol{0}, \tag{4}$$

where $\boldsymbol{\Sigma}$ is an unknown, unstructured, Hermitian positive definite matrix under both hypotheses.

Observe from (3)-(4) that a necessary condition for the problem to be well-posed is that $\boldsymbol{C} \neq \boldsymbol{I}_N$. Otherwise, the covariance matrix under $\mathcal{H}_1$ could be written as $\boldsymbol{I}_N \otimes \tilde{\boldsymbol{\Sigma}}$ with $\tilde{\boldsymbol{\Sigma}} = \boldsymbol{h}\boldsymbol{h}^H + \boldsymbol{\Sigma}$; the fact that both $\boldsymbol{\Sigma}$ and $\tilde{\boldsymbol{\Sigma}}$ are unknown, together with the lack of structure in these matrices, yields the same admissible set of covariance matrices under both hypotheses, which then become indistinguishable.

In order to cope with the unknown parameters $\boldsymbol{h}$ and $\boldsymbol{\Sigma}$, a sensible approach is the Generalized Likelihood Ratio (GLR) test [31], in which these parameters are substituted by their Maximum Likelihood (ML) estimates under each hypothesis:

$$T = \frac{\max_{\boldsymbol{\Sigma},\boldsymbol{h}} f(\boldsymbol{y} \,|\, \boldsymbol{\Sigma}, \boldsymbol{h})}{\max_{\boldsymbol{\Sigma}} f(\boldsymbol{y} \,|\, \boldsymbol{\Sigma}, \boldsymbol{0})} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \gamma, \tag{5}$$

where the probability density function (p.d.f.) f is given by

$$f(\boldsymbol{y} \,|\, \boldsymbol{\Sigma}, \boldsymbol{h}) = \frac{1}{\pi^{MN} \det \boldsymbol{R}} \exp\{-\boldsymbol{y}^H \boldsymbol{R}^{-1} \boldsymbol{y}\}. \tag{6}$$

B. Summary of Main Results

Before delving into the derivation of the GLR test (5), we summarize and discuss our main results. To this end, let us introduce the data matrix $\boldsymbol{Y} \in \mathbb{C}^{M \times N}$ and its (economy-size) singular value decomposition (SVD):

$$\boldsymbol{Y} \doteq [\ \boldsymbol{y}_1 \ \cdots \ \boldsymbol{y}_N \ ] = \boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^H, \tag{7}$$

where $\boldsymbol{U} \in \mathbb{C}^{M \times M}$ is unitary, $\boldsymbol{S} \in \mathbb{C}^{M \times M}$ is positive definite diagonal, and $\boldsymbol{V} \in \mathbb{C}^{N \times M}$ is semi-unitary, i.e., $\boldsymbol{V}^H \boldsymbol{V} = \boldsymbol{I}_M$. Note that $\boldsymbol{Y}$ is related to the data vector $\boldsymbol{y} \in \mathbb{C}^{MN}$ by $\boldsymbol{y} = \text{vec}(\boldsymbol{Y})$. Then we have the following.

Theorem 1: The GLR test (5) can be written as follows, with γ a suitable threshold:

$$T = \max_{\rho \geq 0} t(\rho) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \gamma, \quad t(\rho) \doteq \frac{\lambda_{\min}^{-N} \left[\boldsymbol{V}^H (\boldsymbol{I}_N + \rho \boldsymbol{C}^*)^{-1} \boldsymbol{V} \right]}{\det(\boldsymbol{I}_N + \rho \boldsymbol{C}^*)}, \tag{8}$$

which involves a scalar optimization problem in the variable ρ.

Note that the dependence of (8) with the data is only through the semi-unitary matrix $\boldsymbol{V}$, which is related to the temporal dimension of the data. This is due to the lack of knowledge about the spatial covariance of the noise-plus-interference process $\boldsymbol{z}_n$. In addition, it is readily checked that if $\boldsymbol{C} = \boldsymbol{I}_N$ then the statistic in (8) becomes $T = 1$ independently of the data, reflecting again the fact that white signals are not detectable under this model.

The parameter ρ featuring in (8) is related to the Signal-to-Noise Ratio (SNR), as will be shown in the sequel. To the best of our knowledge, there is no closed-form solution for the general maximization problem (8). Nevertheless, the following result applies in the asymptotic regime of low SNR.

Lemma 1: For vanishingly small SNR, the GLR test (8) becomes equivalent to the following test:

$$T' \doteq \lambda_{\max}[\boldsymbol{V}^H \boldsymbol{C}^* \boldsymbol{V}] \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \gamma'. \tag{9}$$

In order to illustrate the meaning of (9), let $\boldsymbol{C} = \boldsymbol{F}\boldsymbol{\Lambda}\boldsymbol{F}^H$ with $\boldsymbol{\Lambda} = \text{diag}(\lambda_0, \cdots, \lambda_{N-1})$ be an eigenvalue decomposition (EVD) of $\boldsymbol{C}$. It is well known [31] that as $N \to \infty$ the eigenvalues of $\boldsymbol{C}$ approach $\lambda_k \to S_{ss}(e^{j\frac{2\pi k}{N}})$, $0 \leq k \leq N-1$, whereas the matrix of eigenvectors $\boldsymbol{F}$ approach the $N \times N$ orthonormal Inverse Discrete Fourier Transform (IDFT) matrix (this can be formally justified in terms of the asymptotic equivalence between sequences of matrices and the asymptotic eigenvalue distribution of circulant and Toeplitz matrices [32]). Let us now write $\boldsymbol{V}^* = [\ \boldsymbol{v}_1 \ \cdots \ \boldsymbol{v}_M \ ]$. The (i, j) element of $\boldsymbol{V}^H \boldsymbol{C}^* \boldsymbol{V}$ is therefore $\boldsymbol{v}_j^H \boldsymbol{C} \boldsymbol{v}_i = (\boldsymbol{F}^H \boldsymbol{v}_j)^H \boldsymbol{\Lambda}(\boldsymbol{F}^H \boldsymbol{v}_i)$, which for $N \to \infty$ can be thought of as a spectrally weighted frequency-domain crosscorrelation between the outputs of the i-th and j-th orthonormalized data streams (since $\boldsymbol{F}^H \boldsymbol{v}_i$ approaches the N-point DFT of $\boldsymbol{v}_i$). The spectral weights are given by the eigenvalues of $\boldsymbol{C}$, i.e., the sampled PSD of the signal process. Thus, the test statistic T' in (9) is the largest eigenvalue of this spectrally weighted frequency-domain correlation matrix. The discussion above also shows that an approximation to $\boldsymbol{V}^H \boldsymbol{C}^* \boldsymbol{V}$ for N large can be efficiently computed by means of the FFT operation, similarly to [21], with no significant performance loss even for moderate values of N.

If $M = 1$, i.e., only one antenna is available, then it is readily checked that

$$T' = \frac{\boldsymbol{y}^H \boldsymbol{C} \boldsymbol{y}}{\boldsymbol{y}^H \boldsymbol{y}}, \tag{10}$$

which is the ratio of the spectrally weighted energy to total energy, and can be seen as a generalization to the unknown noise power case of the spectral correlation-based detector from [21].

The following result gives the *exact* expression of the GLR test in closed form for a particular family of temporal covariance matrices $\boldsymbol{C}$.

Lemma 2: Assume that $\{0, \lambda\}$ are the only eigenvalues of $\boldsymbol{C}$, with $\lambda > 0$. Then the GLR test (8) is equivalent to the closed-form test (9), independently of the SNR value.

Thus, for the case of a process $\{s_n\}$ with a flat bandpass PSD, the asymptotic (for low SNR) GLR test is also the corresponding GLR test for all SNR values. This suggests that the loss incurred by the detector (9) with other types of signal spectra at moderate SNR may be small. Numerical results will attest to this remark.

To close this section, we note that it is possible to analytically derive the asymptotic distribution of the test statistic T under both hypotheses. The corresponding expressions will be presented in Sec. V.

III. DERIVATION OF THE GLR TEST

In this section we obtain the required ML estimates for the derivation of the GLR detector (5).

A. Preliminaries

Under $\mathcal{H}_0$, (6) can be written as

$$f(\boldsymbol{y} \,|\, \boldsymbol{\Sigma}, \boldsymbol{0}) = \left[\frac{\exp\{-\operatorname{tr}(\boldsymbol{\Sigma}^{-1}\hat{\boldsymbol{R}})\}}{\pi^M \det \boldsymbol{\Sigma}} \right]^N, \qquad (11)$$

where we have introduced the sample covariance matrix

$$\hat{\boldsymbol{R}} \doteq \frac{1}{N} \boldsymbol{Y}\boldsymbol{Y}^H. \qquad (12)$$

It is well known [5] that (11) is maximized for $\hat{\boldsymbol{\Sigma}}_0 = \hat{\boldsymbol{R}}$, yielding

$$\max_{\boldsymbol{\Sigma}} f(\boldsymbol{y} \,|\, \boldsymbol{\Sigma}, \boldsymbol{0}) = [(\pi e)^M \det \hat{\boldsymbol{R}}]^{-N}. \qquad (13)$$

Obtaining the ML estimates under $\mathcal{H}_1$ is more involved. Let us begin by introducing

$$\boldsymbol{u}_\| \doteq \boldsymbol{\Sigma}^{-1}\boldsymbol{h}, \qquad \rho \doteq \boldsymbol{h}^H \boldsymbol{\Sigma}^{-1}\boldsymbol{h}, \qquad \boldsymbol{G}(\rho) \doteq (\boldsymbol{I}_N + \rho\boldsymbol{C})^{-1}. \qquad (14)$$

Vector $\boldsymbol{u}_\|$ is the *Capon beamformer* [33], whereas ρ is the *maximum* SNR that can be obtained at the output of a linear combiner $x_n = \boldsymbol{w}^H \boldsymbol{y}_n$, attained precisely at $\boldsymbol{w} = \boldsymbol{u}_\|$. Note that $\boldsymbol{G}(\rho)$ and $\boldsymbol{C}$ share the same eigenvectors, and therefore they commute. It will be also convenient to introduce the unit norm vectors

$$\bar{\boldsymbol{h}} \doteq \frac{\boldsymbol{h}}{\sqrt{\boldsymbol{h}^H\boldsymbol{h}}}, \qquad \bar{\boldsymbol{u}}_\| \doteq \frac{\boldsymbol{u}_\|}{\sqrt{\boldsymbol{u}_\|^H\boldsymbol{u}_\|}}. \qquad (15)$$

Now we need expressions for the determinant and inverse of $\boldsymbol{R}_1 = \boldsymbol{C} \otimes \boldsymbol{h}\boldsymbol{h}^H + \boldsymbol{I}_N \otimes \boldsymbol{\Sigma}$, featuring in the p.d.f. (6). Applying Sylvester's determinant theorem [37] and the properties of the Kronecker product,

$$\det \boldsymbol{R}_1 = (\det \boldsymbol{\Sigma})^N \det(\boldsymbol{I}_N + \rho\boldsymbol{C}). \qquad (16)$$

Regarding the inverse of $\boldsymbol{R}_1$, it is shown in Appendix A that

$$\boldsymbol{R}_1^{-1} = \boldsymbol{I}_N \otimes \boldsymbol{\Sigma}^{-1} - \boldsymbol{C}\boldsymbol{G}(\rho) \otimes \boldsymbol{u}_\|\boldsymbol{u}_\|^H. \qquad (17)$$

With this, the quadratic form $\boldsymbol{y}^H\boldsymbol{R}_1^{-1}\boldsymbol{y}$ can be written as (see Appendix B):

$$\boldsymbol{y}^H\boldsymbol{R}_1^{-1}\boldsymbol{y} = \operatorname{tr}[\boldsymbol{W}_\perp \boldsymbol{Y}\boldsymbol{Y}^H] + \operatorname{tr}[\boldsymbol{W}_\| \boldsymbol{Y}\boldsymbol{G}^*(\rho)\boldsymbol{Y}^H], \qquad (18)$$

where the matrices $\boldsymbol{W}_\|$, $\boldsymbol{W}_\perp$ are defined as

$$\boldsymbol{W}_\| \doteq \rho^{-1}\boldsymbol{u}_\|\boldsymbol{u}_\|^H \qquad (19)$$

$$= \boldsymbol{\Sigma}^{-1}\left[\rho^{-1}\boldsymbol{h}\boldsymbol{u}_\|^H\right], \qquad (20)$$

$$\boldsymbol{W}_\perp \doteq \boldsymbol{\Sigma}^{-1} - \rho^{-1}\boldsymbol{u}_\|\boldsymbol{u}_\|^H \qquad (21)$$

$$= \boldsymbol{\Sigma}^{-1}\left[\boldsymbol{I}_M - \rho^{-1}\boldsymbol{h}\boldsymbol{u}_\|^H\right]. \qquad (22)$$

Note that $\boldsymbol{\Sigma}^{-1} = \boldsymbol{W}_\| + \boldsymbol{W}_\perp$, and that $\boldsymbol{W}_\|$ has rank one and is positive semidefinite. In addition, $\boldsymbol{W}_\perp\boldsymbol{h} = \boldsymbol{0}$, and thus the rank of $\boldsymbol{W}_\perp$ is at most $M - 1$. The geometric interpretation of these matrices is as follows. For any $\boldsymbol{x} \in \mathbb{C}^M$, write $\boldsymbol{x} = c\boldsymbol{h} + \tilde{\boldsymbol{x}}$ for some $c \in \mathbb{C}$, $\tilde{\boldsymbol{x}} \in \mathbb{C}^M$ with $\tilde{\boldsymbol{x}}^H\boldsymbol{u}_\| = 0$. Note that as long as $\rho = \boldsymbol{u}_\|^H\boldsymbol{h} \neq 0$, this decomposition exists and is unique. (This condition means that the Capon beamformer is not orthogonal to the spatial signature, i.e., it does not block

the signal at its output completely. It is satisfied for full rank $\boldsymbol{\Sigma}$, as $\rho = \boldsymbol{h}^H\boldsymbol{\Sigma}^{-1}\boldsymbol{h} > 0$). Then

$$\left[\rho^{-1}\boldsymbol{h}\boldsymbol{u}_\|^H\right]\boldsymbol{x} = c\boldsymbol{h}, \qquad \left[\boldsymbol{I}_M - \rho^{-1}\boldsymbol{h}\boldsymbol{u}_\|^H\right]\boldsymbol{x} = \tilde{\boldsymbol{x}}, \quad (23)$$

i.e., $\boldsymbol{\Pi} \doteq \boldsymbol{I}_M - \rho^{-1}\boldsymbol{h}\boldsymbol{u}_\|^H$ is the oblique projector along $\boldsymbol{h}$ onto the subspace orthogonal to $\boldsymbol{u}_\|$.

Note that $\boldsymbol{x}^H\boldsymbol{\Sigma}^{-1}\boldsymbol{x} = |c|^2\boldsymbol{h}^H\boldsymbol{\Sigma}^{-1}\boldsymbol{h} + \tilde{\boldsymbol{x}}^H\boldsymbol{W}_\perp\tilde{\boldsymbol{x}}$, which is positive for $\boldsymbol{x} \neq 0$. Then for $c = 0$, one has that $\boldsymbol{x} = \tilde{\boldsymbol{x}}$ is in a subspace of dimension $M - 1$, and $\boldsymbol{x}^H\boldsymbol{\Sigma}^{-1}\boldsymbol{x} = \tilde{\boldsymbol{x}}^H\boldsymbol{W}_\perp\tilde{\boldsymbol{x}} > 0$ for $\tilde{\boldsymbol{x}} \neq \boldsymbol{0}$, showing that $\boldsymbol{W}_\perp$ is positive semidefinite of rank $M - 1$.

For completeness, we provide an expression for $\boldsymbol{\Sigma}$ in terms of these matrices, given by

$$\boldsymbol{\Sigma} = \boldsymbol{\Pi}\boldsymbol{W}_\perp^\dagger\boldsymbol{\Pi}^H + \rho^{-1}\boldsymbol{h}\boldsymbol{h}^H, \qquad (24)$$

which is proved in Remark 2 (Appendix D).

The procedure whereby two different orthogonal spaces are considered for the signal and the noise plus interference subspaces in the spatial domain for this non-collaborative spectrum sensing scenario constitutes a recurrent and successful approach in multiantenna detection [30] [34], array signal processing [35] as well as in spectral estimation contexts [36]. The use of suitable parameter transformations in detection schemes has also been addressed in [30].

B. A convenient change of variables

The decomposition of the unstructured inverse spatial co-variance $\boldsymbol{\Sigma}^{-1} = \boldsymbol{W}_\| + \boldsymbol{W}_\perp$ into two structured positive semidefinite matrices will be useful in order to obtain the ML estimates. To this end, we introduce a change of variables in order to replace the original set of unstructured unknown parameters $\Omega \doteq \{\boldsymbol{h}, \boldsymbol{\Sigma}\}$ by another set Ω' based on $\boldsymbol{W}_\|$ and $\boldsymbol{W}_\perp$. In order to account for the structure of these matrices, let us introduce the following (economy-size) EVDs:

$$\boldsymbol{W}_\| = \gamma_\|\bar{\boldsymbol{u}}_\|\bar{\boldsymbol{u}}_\|^H, \qquad \boldsymbol{W}_\perp = \boldsymbol{U}_\perp\boldsymbol{\Gamma}_\perp\boldsymbol{U}_\perp^H, \qquad (25)$$

where $\gamma_\| \in \mathbb{R}$ is positive, $\bar{\boldsymbol{u}}_\| \in \mathbb{C}^M$ has unit norm (note from (19) that $\bar{\boldsymbol{u}}_\|$ is the unit-norm Capon beamformer and $\gamma_\|$ the inverse of the noise power at its output), $\boldsymbol{\Gamma}_\perp \in \mathbb{R}^{(M-1)\times(M-1)}$ is positive definite diagonal, and $\boldsymbol{U}_\perp \in \mathbb{C}^{M\times(M-1)}$ has orthonormal columns. In addition, we require that (i) $\boldsymbol{U}_\perp^H\bar{\boldsymbol{h}} = \boldsymbol{0}$ (to ensure that $\boldsymbol{W}_\perp\bar{\boldsymbol{h}} = \boldsymbol{0}$); and (ii) $\bar{\boldsymbol{u}}_\|^H\bar{\boldsymbol{h}} \neq 0$ (to ensure that $\boldsymbol{W}_\| + \boldsymbol{W}_\perp$ has full rank). Note that $[\boldsymbol{U}_\perp \ \bar{\boldsymbol{h}}] \in \mathbb{C}^{M\times M}$ is unitary.

With these, the new parameter space we consider is given by

$$\Omega' \doteq \{\rho, \bar{\boldsymbol{h}}, \gamma_\|, \bar{\boldsymbol{u}}_\|, \boldsymbol{\Gamma}_\perp, \boldsymbol{U}_\perp\}. \qquad (26)$$

It is readily checked that under the standing assumptions $\boldsymbol{h} \neq \boldsymbol{0}$, $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}^H$ full rank, there is a one-to-one correspondence[2] between the parameter spaces Ω and Ω', which is summarized in Table I. Hence, the maximization of the likelihood function can be carried out over either Ω or Ω' with identical result.

[2]Apart from irrelevant rotational ambiguities in $\bar{\boldsymbol{h}}$ and $\bar{\boldsymbol{u}}_\|$, and rotational/ordering ambiguities in the EVD of $\boldsymbol{W}_\perp$.

In order to write the likelihood function in terms of the new parameters, note from (16) and (18) that

$$
\begin{aligned}
& -\log f(\boldsymbol{y} \,|\, \boldsymbol{\Sigma}, \boldsymbol{h}) \\
={} & MN \log \pi + N \log \det \boldsymbol{\Sigma} - \log \det \boldsymbol{G}^*(\rho) \\
+{} & N \operatorname{tr}[\boldsymbol{W}_\perp \hat{\boldsymbol{R}}] + \operatorname{tr}[\boldsymbol{W}_\| \boldsymbol{Y} \boldsymbol{G}^*(\rho) \boldsymbol{Y}^H].
\end{aligned} \tag{27}
$$

The following result links the determinant of the unstructured spatial covariance matrix with the new parameters; see Appendix C for the proof.

Lemma 3: Under the parameterization (25)-(26), it holds that $\det(\boldsymbol{\Sigma}^{-1}) = \gamma_\| \left|\bar{\boldsymbol{h}}^H \bar{\boldsymbol{u}}_\|\right|^2 \det(\boldsymbol{\Gamma}_\perp)$.

Then (27) becomes

$$
\begin{aligned}
& -\log f(\boldsymbol{y} \,|\, \boldsymbol{\Sigma}, \boldsymbol{h}) \\
={} & MN \log \pi - N \log \gamma_\| - N \log \left|\bar{\boldsymbol{h}}^H \bar{\boldsymbol{u}}_\|\right|^2 \\
-{} & N \log \det \boldsymbol{\Gamma}_\perp - \log \det \boldsymbol{G}^*(\rho) \\
+{} & N \operatorname{tr}[\boldsymbol{U}_\perp \boldsymbol{\Gamma}_\perp \boldsymbol{U}_\perp^H \hat{\boldsymbol{R}}] + \gamma_\| \bar{\boldsymbol{u}}_\|^H \boldsymbol{Y} \boldsymbol{G}^*(\rho) \boldsymbol{Y}^H \bar{\boldsymbol{u}}_\|.
\end{aligned} \tag{28}
$$

C. ML estimation under $\mathcal{H}_1$

We proceed now to minimize (28) with respect to the parameters in (26). The optimum values of $\gamma_\|$ and $\boldsymbol{\Gamma}_\perp$ are readily found:

$$
\hat{\gamma}_\| = \frac{N}{\bar{\boldsymbol{u}}_\|^H \boldsymbol{Y} \boldsymbol{G}^*(\rho) \boldsymbol{Y}^H \bar{\boldsymbol{u}}_\|}, \qquad \hat{\boldsymbol{\Gamma}}_\perp = \left[\operatorname{diag}(\boldsymbol{U}_\perp^H \hat{\boldsymbol{R}} \boldsymbol{U}_\perp)\right]^{-1}. \tag{29}
$$

Substituting (29) back in (28), we obtain[3]

$$
\begin{aligned}
& -\log f \\
={} & N\left(M + \log \frac{\pi^M}{N}\right) - N \log \left|\bar{\boldsymbol{h}}^H \bar{\boldsymbol{u}}_\|\right|^2 - \log \det \boldsymbol{G}^*(\rho) \\
+{} & N \log \det\left[\operatorname{diag}(\boldsymbol{U}_\perp^H \hat{\boldsymbol{R}} \boldsymbol{U}_\perp)\right] \\
+{} & N \log \left(\bar{\boldsymbol{u}}_\|^H \boldsymbol{Y} \boldsymbol{G}^*(\rho) \boldsymbol{Y}^H \bar{\boldsymbol{u}}_\|\right).
\end{aligned} \tag{30}
$$

The optimum value of $\boldsymbol{U}_\perp$ is provided by the following result, whose proof is given in Appendix D:

Lemma 4: Let us consider the following cost: $J(\boldsymbol{U}_\perp) \doteq \det\left[\operatorname{diag}(\boldsymbol{U}_\perp^H \hat{\boldsymbol{R}} \boldsymbol{U}_\perp)\right]$. Then the solution of

$$
\min_{\boldsymbol{U}_\perp} J(\boldsymbol{U}_\perp) \quad \text{subject to} \quad \boldsymbol{U}_\perp^H \boldsymbol{U}_\perp = \boldsymbol{I}_{M-1}, \quad \boldsymbol{U}_\perp^H \bar{\boldsymbol{h}} = \boldsymbol{0} \tag{31}
$$

is given by $J_{\min} = \bar{\boldsymbol{h}}^H \hat{\boldsymbol{R}}^{-1} \bar{\boldsymbol{h}} \cdot \det \hat{\boldsymbol{R}}$, and is attained when $\boldsymbol{U}_\perp$ is an orthonormal basis for the range space of $(\boldsymbol{I}_M - \bar{\boldsymbol{h}}\bar{\boldsymbol{h}}^H) \hat{\boldsymbol{R}} (\boldsymbol{I}_M - \bar{\boldsymbol{h}}\bar{\boldsymbol{h}}^H)$.

Substituting the value of $J_{\min}$ into (30), one has

$$
\begin{aligned}
& -\log f \\
={} & N\left(M + \log \frac{\pi^M}{N} + \log \det \hat{\boldsymbol{R}}\right) - N \log \left|\bar{\boldsymbol{h}}^H \bar{\boldsymbol{u}}_\|\right|^2 \\
-{} & \log \det \boldsymbol{G}^*(\rho) \\
+{} & N \log(\bar{\boldsymbol{h}}^H \hat{\boldsymbol{R}}^{-1} \bar{\boldsymbol{h}}) + N \log \left(\bar{\boldsymbol{u}}_\|^H \boldsymbol{Y} \boldsymbol{G}^*(\rho) \boldsymbol{Y}^H \bar{\boldsymbol{u}}_\|\right)
\end{aligned} \tag{32}
$$

[3]For the sake of clarity and with some abuse of the notation, we denote the value of $f(\boldsymbol{y} \,|\, \boldsymbol{\Sigma}, \boldsymbol{h})$ obtained after maximization w.r.t. a parameter in Ω' simply by f.

Now, minimizing (32) with respect to $\bar{\boldsymbol{h}}$ and $\bar{\boldsymbol{u}}_\|$ (unit-norm signature vector and Capon beamformer respectively) amounts to minimizing

$$
\frac{(\bar{\boldsymbol{h}}^H \hat{\boldsymbol{R}}^{-1} \bar{\boldsymbol{h}}) \cdot (\bar{\boldsymbol{u}}_\|^H \boldsymbol{Y} \boldsymbol{G}^*(\rho) \boldsymbol{Y}^H \bar{\boldsymbol{u}}_\|)}{\left|\bar{\boldsymbol{h}}^H \bar{\boldsymbol{u}}_\|\right|^2}, \tag{33}
$$

subject to $\bar{\boldsymbol{h}}^H \bar{\boldsymbol{u}}_\| \neq 0$. To this end, the following result will be applied; see Appendix E for the proof.

Lemma 5: Let $\boldsymbol{A}_1$, $\boldsymbol{A}_2$ be two $M \times M$ positive definite Hermitian matrices. The minimum value of

$$
F(\boldsymbol{u}_1, \boldsymbol{u}_2) \doteq \frac{(\boldsymbol{u}_1^H \boldsymbol{A}_1 \boldsymbol{u}_1) \cdot (\boldsymbol{u}_2^H \boldsymbol{A}_2 \boldsymbol{u}_2)}{|\boldsymbol{u}_1^H \boldsymbol{u}_2|^2} \tag{34}
$$

subject to $\boldsymbol{u}_1^H \boldsymbol{u}_2 \neq 0$ is given by $F_{\min} = \lambda_{\min}(\boldsymbol{A}_1 \boldsymbol{A}_2) = \lambda_{\min}(\boldsymbol{A}_2 \boldsymbol{A}_1)$, and is attained if $\boldsymbol{u}_1$ (resp. $\boldsymbol{u}_2$) is a minimum right eigenvector[4] of $\boldsymbol{A}_2 \boldsymbol{A}_1$ (resp. $\boldsymbol{A}_1 \boldsymbol{A}_2$).

Therefore, from the SVD (7), the minimum value of (33) is found to be $\lambda_{\min}[\boldsymbol{Y} \boldsymbol{G}^*(\rho) \boldsymbol{Y}^H \hat{\boldsymbol{R}}^{-1}] = N \lambda_{\min}[\boldsymbol{V}^H \boldsymbol{G}^*(\rho) \boldsymbol{V}]$. Using $\boldsymbol{Y}^\dagger = \boldsymbol{V} \boldsymbol{S}^{-1} \boldsymbol{U}^H$, the corresponding ML estimates $\hat{\bar{\boldsymbol{h}}}$ and $\hat{\bar{\boldsymbol{u}}}_\|$ are seen to be the unit-norm minimum eigenvectors of the matrices $\boldsymbol{Y} \boldsymbol{G}^*(\rho) \boldsymbol{Y}^\dagger$ and $(\boldsymbol{Y}^H)^\dagger \boldsymbol{G}^*(\rho) \boldsymbol{Y}^H$, respectively.

Substitution of these optimum values into (32) finally yields

$$
\max_{\boldsymbol{\Sigma}, \boldsymbol{h}} f(\boldsymbol{y} \,|\, \boldsymbol{\Sigma}, \boldsymbol{h}) \tag{35}
$$

$$
= [(\pi e)^M \det \hat{\boldsymbol{R}}]^{-N} \cdot \max_{\rho \geq 0} \left\{ \left(\frac{\lambda_{\min}^N[\boldsymbol{V}^H \boldsymbol{G}^*(\rho) \boldsymbol{V}]}{\det \boldsymbol{G}^*(\rho)}\right)^{-1} \right\},
$$

which, together with (13), proves that the GLR test is indeed (8) as in Theorem 1.

The ML estimates of the parameters are summarized in Table I. The expression for the ML estimate $\hat{\boldsymbol{W}}_\perp = \hat{\boldsymbol{U}}_\perp \hat{\boldsymbol{\Gamma}}_\perp \hat{\boldsymbol{U}}_\perp^H$ given in Table I is justified in Remark 1 of Appendix D while that of $\boldsymbol{\Sigma}$ stems from (24).

As previously mentioned, there is no closed-form expression for the ML estimate of ρ, which has to be computed by numerical means, e.g. a gradient search as in Sec. VI. This raises the issue of local optima, about which the following can be said. First, experimental evidence suggests that the likelihood function is unimodal in ρ. Second, the asymptotic (as $N \to \infty$) likelihood function does turn out to be unimodal; see Remark 3 in Appendix F. And finally, for certain kinds of temporal covariance matrices unimodality can be analytically established, as shown in the next section.

IV. GLR TEST FOR LOW-SNR AND BINARY-VALUED EIGENSPECTRUM CASES

In this section we first explore the behavior of the GLR test (8) for asymptotically low SNR, with arbitrary (but non-white) temporal correlation of the signal of interest. After this, we particularize the GLR test (8) to the case in which the eigenvalues of the temporal covariance matrix $\boldsymbol{C}$ reduce to $\{0, \lambda\}$ with $\lambda > 0$ (recall that these eigenvalues approach the samples of the signal PSD as $N \to \infty$; thus, this situation includes the case of a PSD occupying only a fraction of the Nyquist bandwidth, in which it is constant), showing

[4]In the sequel we shall refer to *right eigenvectors* simply as *eigenvectors*.

Transformed parameters: $\Omega' = \{\rho, \boldsymbol{h}, \gamma_\parallel, \bar{\boldsymbol{u}}_\parallel, \boldsymbol{\Gamma}_\perp, \boldsymbol{U}_\perp\}$	ML estimates of transformed parameters
$\rho = \boldsymbol{h}^H \boldsymbol{\Sigma}^{-1} \boldsymbol{h}$	$\hat{\rho} = \arg\min_{\rho \geq 0} \left\{ \frac{\lambda_{\min}^N \left[(\boldsymbol{Y}\boldsymbol{G}^*(\rho)\boldsymbol{Y}^H)(\boldsymbol{Y}\boldsymbol{Y}^H)^{-1} \right]}{\det \boldsymbol{G}^*(\rho)} \right\}$
$\bar{\boldsymbol{h}} = \boldsymbol{h}/\sqrt{\boldsymbol{h}^H \boldsymbol{h}}$	$\hat{\bar{\boldsymbol{h}}} =$ unit-norm minimum eigenvector of $\boldsymbol{Y}\boldsymbol{G}^*(\hat{\rho})\boldsymbol{Y}^\dagger$
$\bar{\boldsymbol{u}}_\parallel = (\boldsymbol{\Sigma}^{-1}\boldsymbol{h})/\sqrt{\boldsymbol{h}^H \boldsymbol{\Sigma}^{-2}\boldsymbol{h}}$	$\hat{\bar{\boldsymbol{u}}}_\parallel =$ unit-norm minimum eigenvector of $(\boldsymbol{Y}^H)^\dagger \boldsymbol{G}^*(\hat{\rho}) \boldsymbol{Y}^H$
$\gamma_\parallel = (\boldsymbol{h}^H \boldsymbol{\Sigma}^{-2}\boldsymbol{h})/(\boldsymbol{h}^H \boldsymbol{\Sigma}^{-1}\boldsymbol{h})$	$\hat{\gamma}_\parallel = \left(\frac{1}{N}\hat{\bar{\boldsymbol{u}}}_\parallel^H \boldsymbol{Y}\boldsymbol{G}^*(\hat{\rho})\boldsymbol{Y}^H \hat{\bar{\boldsymbol{u}}}_\parallel \right)^{-1}$
$\boldsymbol{U}_\perp \boldsymbol{\Gamma}_\perp \boldsymbol{U}_\perp^H = \boldsymbol{\Sigma}^{-1} - (\boldsymbol{h}^H \boldsymbol{\Sigma}^{-1}\boldsymbol{h})^{-1}(\boldsymbol{\Sigma}^{-1}\boldsymbol{h})(\boldsymbol{\Sigma}^{-1}\boldsymbol{h})^H$	$\hat{\boldsymbol{U}}_\perp \hat{\boldsymbol{\Gamma}}_\perp \hat{\boldsymbol{U}}_\perp^H = \left[(\boldsymbol{I}_M - \hat{\bar{\boldsymbol{h}}}\hat{\bar{\boldsymbol{h}}}^H)\hat{\boldsymbol{R}}(\boldsymbol{I}_M - \hat{\bar{\boldsymbol{h}}}\hat{\bar{\boldsymbol{h}}}^H) \right]^\dagger$
Original parameters: $\Omega = \{\boldsymbol{\Sigma}, \boldsymbol{h}\}$	ML estimates of original parameters
$\boldsymbol{h} = (\sqrt{\rho/\gamma_\parallel}/|\bar{\boldsymbol{u}}_\parallel^H \bar{\boldsymbol{h}}|) \cdot \bar{\boldsymbol{h}}$	$\hat{\boldsymbol{h}} = \left(\sqrt{\hat{\rho}/\hat{\gamma}_\parallel}/|\hat{\bar{\boldsymbol{u}}}_\parallel^H \hat{\bar{\boldsymbol{h}}}| \right) \cdot \hat{\bar{\boldsymbol{h}}}$
$\boldsymbol{\Sigma}^{-1} = \boldsymbol{U}_\perp \boldsymbol{\Gamma}_\perp \boldsymbol{U}_\perp^H + \gamma_\parallel \bar{\boldsymbol{u}}_\parallel \bar{\boldsymbol{u}}_\parallel^H$	$\widehat{\boldsymbol{\Sigma}^{-1}} = \hat{\boldsymbol{U}}_\perp \hat{\boldsymbol{\Gamma}}_\perp \hat{\boldsymbol{U}}_\perp^H + \hat{\gamma}_\parallel \hat{\bar{\boldsymbol{u}}}_\parallel \hat{\bar{\boldsymbol{u}}}_\parallel^H$
$\boldsymbol{\Sigma} = \boldsymbol{\Pi}(\boldsymbol{U}_\perp \boldsymbol{\Gamma}_\perp^{-1} \boldsymbol{U}_\perp^H)\boldsymbol{\Pi}^H + (\gamma_\parallel |\bar{\boldsymbol{u}}_\parallel^H \bar{\boldsymbol{h}}|^2)^{-1}\bar{\boldsymbol{h}}\bar{\boldsymbol{h}}^H$	$\hat{\boldsymbol{\Sigma}} = \hat{\boldsymbol{\Pi}}\hat{\boldsymbol{R}}\hat{\boldsymbol{\Pi}}^H + \hat{\rho}^{-1}\hat{\boldsymbol{h}}\hat{\boldsymbol{h}}^H$
$\boldsymbol{\Pi} \doteq \boldsymbol{I}_M - (\bar{\boldsymbol{u}}_\parallel^H \bar{\boldsymbol{h}})^{-1}\bar{\boldsymbol{h}}\bar{\boldsymbol{u}}_\parallel^H$	$\hat{\boldsymbol{\Pi}} = \boldsymbol{I}_M - (\hat{\bar{\boldsymbol{u}}}_\parallel^H \hat{\bar{\boldsymbol{h}}})^{-1}\hat{\bar{\boldsymbol{h}}}\hat{\bar{\boldsymbol{u}}}_\parallel^H$

TABLE I
TRUE PARAMETERS (LEFT) AND ML ESTIMATES (RIGHT) UNDER $\mathcal{H}_1$, WITH $\hat{\boldsymbol{R}} = \frac{1}{N}\boldsymbol{Y}\boldsymbol{Y}^H$, $\boldsymbol{G}(\rho) \doteq (\boldsymbol{I}_N + \rho\boldsymbol{C})^{-1}$.

that a closed-form expression for the detector exists in this case, valid for all SNR values. Remarkably, the closed-form detectors obtained in both cases are the same.

A. GLR test for asymptotically low SNR

From (8), the GLR statistic is $T = \max_{\rho \geq 0} t(\rho)$. For small ρ, the following first-order approximations

$$(\boldsymbol{I}_N + \rho\boldsymbol{C}^*)^{-1} \approx \boldsymbol{I}_N - \rho\boldsymbol{C}^*, \qquad \det(\boldsymbol{I}_N + \rho\boldsymbol{C}^*) \approx 1 + \rho\operatorname{tr}\boldsymbol{C}^* \tag{36}$$

hold. Therefore,

$$t(\rho) \approx \frac{\lambda_{\min}^{-N}[\boldsymbol{V}^H(\boldsymbol{I}_N - \rho\boldsymbol{C}^*)\boldsymbol{V}]}{1 + \rho\operatorname{tr}\boldsymbol{C}^*} = \frac{(1 - \rho\lambda_{\max}[\boldsymbol{V}^H\boldsymbol{C}^*\boldsymbol{V}])^{-N}}{1 + \rho\operatorname{tr}\boldsymbol{C}^*}. \tag{37}$$

Taking logarithms and using the approximation $\log(1+x) \approx x$ for small $|x|$, one has

$$\log t(\rho) \approx \rho N \left(\lambda_{\max}[\boldsymbol{V}^H\boldsymbol{C}^*\boldsymbol{V}] - \frac{1}{N}\operatorname{tr}\boldsymbol{C}^* \right). \tag{38}$$

Suppose now that ρ were known; in that case, the GLR statistic is directly $t(\rho)$, and (38) shows that, in low SNR, the GLR test is equivalent to comparing $\lambda_{\max}[\boldsymbol{V}^H\boldsymbol{C}^*\boldsymbol{V}]$ against a threshold. This amounts to saying that for vanishingly small SNR, knowledge of ρ becomes irrelevant, and the GLR test can be rephrased as in (9) since $\operatorname{tr}\boldsymbol{C}^*$ is a constant independent of the data, thus establishing Lemma 1.

B. Binary-valued eigenspectrum

We now set to prove Lemma 2. Consider the EVD $\boldsymbol{C} = \boldsymbol{F}\boldsymbol{\Lambda}\boldsymbol{F}^H$, and assume that the only eigenvalues of $\boldsymbol{C}$ are λ and zero, with multiplicities L and $N - L$ respectively. Since it is assumed that $\boldsymbol{C}$ has ones on its diagonal, it follows that $\operatorname{tr}\boldsymbol{C} = L\lambda = N$. Letting $b \doteq L/N < 1$ denote the fraction of nonzero eigenvalues, then $\lambda = b^{-1}$. The GLR statistic T from (8) can be written as

$$T = \max_{\rho \geq 0} \left\{ \frac{\lambda_{\min}^{-N}\left[(\boldsymbol{F}^*\boldsymbol{V})^H(\boldsymbol{I}_N + \rho\boldsymbol{\Lambda})^{-1}\boldsymbol{F}^*\boldsymbol{V} \right]}{\det(\boldsymbol{I}_N + \rho\boldsymbol{\Lambda})} \right\} \tag{39}$$

It is readily checked that $(\boldsymbol{I}_N + \rho\boldsymbol{\Lambda})^{-1} = \boldsymbol{I}_N - \frac{\rho}{1 + b^{-1}\rho}\boldsymbol{\Lambda}$ and $\det(\boldsymbol{I}_N + \rho\boldsymbol{\Lambda}) = (1 + b^{-1}\rho)^{bN}$. Therefore,

$$\begin{aligned} T^{1/N} & \tag{40} \\ &= \max_{\rho \geq 0} \left\{ \frac{\lambda_{\min}^{-1}\left[(\boldsymbol{F}^*\boldsymbol{V})^H(\boldsymbol{I}_N - \frac{\rho}{1 + b^{-1}\rho}\boldsymbol{\Lambda})\boldsymbol{F}^*\boldsymbol{V} \right]}{(1 + b^{-1}\rho)^b} \right\} \\ &= \left\{ \min_{\rho \geq 0} \left[\left(1 - \frac{\rho\bar{\lambda}}{1 + b^{-1}\rho} \right)(1 + b^{-1}\rho)^b \right] \right\}^{-1}, \quad \tag{41} \end{aligned}$$

where $\bar{\lambda} \doteq \lambda_{\max}[\boldsymbol{V}^H\boldsymbol{C}^*\boldsymbol{V}]$. It is straightforward to show that the minimum in (41) is attained at

$$\hat{\rho} = \max\left\{ 0, \frac{\bar{\lambda} - 1}{b^{-1} - \bar{\lambda}} \right\}, \qquad \text{if } 0 \leq \bar{\lambda} \leq b^{-1}, \tag{42}$$

whereas no minimum exists if $\bar{\lambda} > b^{-1}$. However, this case need not be considered: since $\boldsymbol{V}^H\boldsymbol{V} = \boldsymbol{I}_M$, by the Poincaré separation theorem [38] one has $\lambda_{\max}[\boldsymbol{V}^H\boldsymbol{C}^*\boldsymbol{V}] \leq \lambda_{\max}[\boldsymbol{C}^*] = b^{-1}$. Thus, (41) becomes

$$T^{1/N} = \begin{cases} 1, & \text{if } 0 \leq \bar{\lambda} \leq 1, \\ \bar{\lambda}^{-b}\left(\frac{1-b}{1-b\bar{\lambda}} \right)^{1-b}, & \text{if } 1 < \bar{\lambda} \leq b^{-1}, \end{cases} \tag{43}$$

which is a non-decreasing function of $\bar{\lambda} \in [0, b^{-1}]$. Hence, the GLR test for binary-valued eigenspectrum is equivalent to the test (9), as was to be shown. It has a nice geometrical interpretation: if we collect in $\bar{\boldsymbol{F}}$ the columns of $\boldsymbol{F}$ corresponding to the nonzero eigenvalue $\lambda = b^{-1}$, then $\boldsymbol{C} = b^{-1}\bar{\boldsymbol{F}}\bar{\boldsymbol{F}}^H$, and

$$\begin{aligned} \lambda_{\max}[\boldsymbol{V}^H\boldsymbol{C}^*\boldsymbol{V}] &= b^{-1}\lambda_{\max}[(\bar{\boldsymbol{F}}^H\boldsymbol{V}^*)^H(\bar{\boldsymbol{F}}^H\boldsymbol{V}^*)] \\ &= b^{-1}\cos^2\theta, \tag{44} \end{aligned}$$

where θ is the *minimum principal angle* [39] between the subspaces spanned by $\boldsymbol{V}^*$ and $\bar{\boldsymbol{F}}$. The GLR statistic is thus a measure of the largest achievable projection of a vector in the range space of $\boldsymbol{V}^*$ (data subspace) onto the range space of $\bar{\boldsymbol{F}}$ (reference subspace). Note that $\cos^2\theta > b$ in order to have $T > 1$.

V. PERFORMANCE ANALYSIS IN LOW SNR

The following result, whose proof is given in Appendix F, provides the asymptotic distributions of the GLR statistic under each hypothesis, assuming that the SNR is small. Thus, it enables us to compute the probabilities of detection P_D and false alarm P_{FA} for a given threshold.

Theorem 2: As $N \to \infty$ and for sufficiently small SNR, the GLR statistic (8) is asymptotically distributed as

$$2 \log T \sim \begin{cases} \chi^2_{2M-1}, & \text{under } \mathcal{H}_0, \\ \chi'^2_{2M-1}(\alpha_{\mathrm{sph}}(\rho)), & \text{under } \mathcal{H}_1, \end{cases} \quad (45)$$

$$\alpha_{\mathrm{sph}}(\rho) \doteq 2N \log \frac{\frac{1}{N} \mathrm{tr}(\boldsymbol{I}_N + \rho \boldsymbol{C})}{[\det(\boldsymbol{I}_N + \rho \boldsymbol{C})]^{1/N}}, \quad (46)$$

where $\rho \doteq \boldsymbol{h}^H \boldsymbol{\Sigma}^{-1} \boldsymbol{h}$, and χ^2_d, $\chi'^2_d(\alpha)$ denote respectively central and non-central chi-square distributions with d degrees of freedom and non-centrality parameter α.

The argument of the log in (46) is the ratio of the arithmetic mean (AM) to geometric mean (GM) of the eigenvalues of $\boldsymbol{I}_N + \rho \boldsymbol{C}$, which is the temporal covariance matrix at the output of the Capon beamformer. By the AM-GM inequality, this ratio is no less than one, and it is equal to one iff all eigenvalues are equal, i.e., iff $\boldsymbol{C} = \boldsymbol{I}_N$. In that case $\alpha_{\mathrm{sph}}(\rho) = 0$ for all ρ and the asymptotic distributions under $\mathcal{H}_1$ and $\mathcal{H}_0$ coincide, which is consistent with the fact that white signals are not detectable under this signal model. Applying Szegö's theorem [40], the asymptotic value of the AM-GM ratio can be written in terms of the signal PSD:

$$\lim_{N \to \infty} \frac{\frac{1}{N} \mathrm{tr}(\boldsymbol{I}_N + \rho \boldsymbol{C})}{[\det(\boldsymbol{I}_N + \rho \boldsymbol{C})]^{1/N}} \quad (47)$$
$$= \frac{\frac{1}{2\pi} \int_{-\pi}^{\pi} [1 + \rho S_{ss}(e^{j\omega})] \mathrm{d}\omega}{\exp\left\{\frac{1}{2\pi} \int_{-\pi}^{\pi} \log[1 + \rho S_{ss}(e^{j\omega})] \mathrm{d}\omega\right\}},$$

which is the inverse of the Spectral Flatness Measure (SFM) [41] associated with the power spectrum $1 + \rho S_{ss}(e^{j\omega})$. Its minimum value is 1 for $\rho = 0$, increasing monotonically with ρ toward the inverse of the SFM associated to $S_{ss}(e^{j\omega})$ [42].

We could ask for the signal eigenvalue distribution maximizing the performance of the GLR detector at a given SNR. In view of Theorem 2, this amounts to maximizing $\alpha_{\mathrm{sph}}(\rho)$. Recalling that $\mathrm{tr}\, \boldsymbol{C} = N$ due to signal power normalization, and with $\{\lambda_0, \ldots, \lambda_{N-1}\}$ the eigenvalues of $\boldsymbol{C}$, the problem becomes

$$\min_{\{\lambda_n\}} \prod_{n=0}^{N-1} (1 + \rho \lambda_n) \quad (48)$$
$$\text{subject to} \quad \sum_{n=0}^{N-1} \lambda_n = N, \quad \lambda_i \geq 0, \ i = 0, \ldots, N-1,$$

whose solutions are $\lambda_j = N$ for some $j \in \{0, 1, \ldots, N-1\}$ and $\lambda_n = 0$ for $n \neq j$ (independently of ρ), as can be easily shown following the same reasoning as in [21, Sec. IV]. For $N \to \infty$, this implies that the optimum PSD concentrates all its power at a single frequency. This is not surprising, as this kind of peaky spectra minimize the SFM and are easier to detect in the presence of noise.
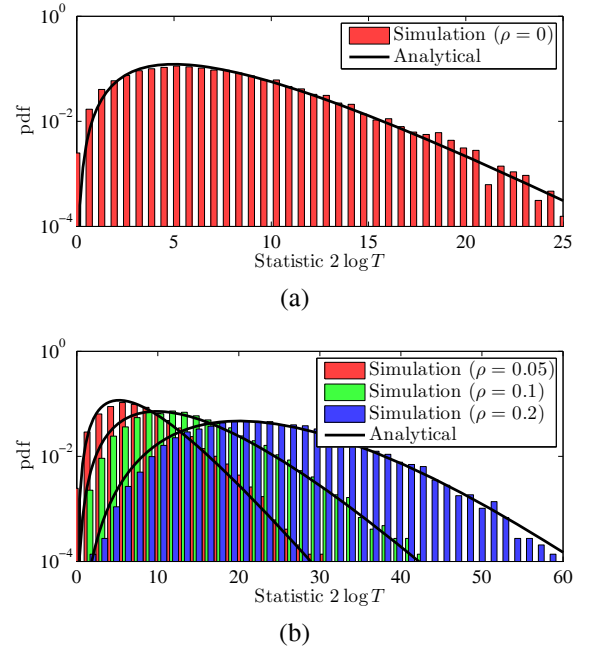

Fig. 1. Distribution of the statistic $2 \log T$ for $M = 4$ and $N = 128$. (a) Under $\mathcal{H}_0$ (b) Under $\mathcal{H}_1$.

Finally, note from Theorem 2 that, as desired, the asymptotic performance of the GLR detector does not depend on the specific spatial correlation profile of the noise, but only on the operational SNR ρ.

VI. NUMERICAL RESULTS

We proceed to examine the performance of the proposed detectors via Monte Carlo simulations and check the accuracy of the analytical approximations. In all experiments, the signature vector $\boldsymbol{h}$ and the noise spatial covariance matrix $\boldsymbol{\Sigma} = \boldsymbol{H} \boldsymbol{H}^H$ are randomly generated at each Monte Carlo run, with $\boldsymbol{h}$ scaled in order to obtain a given SNR value $\rho = \boldsymbol{h}^H \boldsymbol{\Sigma}^{-1} \boldsymbol{h}$. The entries of $\boldsymbol{h}$ and $\boldsymbol{H} \in \mathbb{C}^{M \times M}$ are independent zero-mean circular Gaussian. We refer to the two proposed detectors as:

1) *Iterative GLRT*: the exact GLR test from (8), in which the optimization w.r.t. ρ is performed using a gradient descent algorithm detailed in Appendix G.
2) *Asymptotic GLRT*: the closed-form detector from (9), which has been shown to coincide with the GLR detector for vanishing SNR or for a binary-valued eigenspectrum (PSD) of the primary signal.

For each signal type used in the simulations, the autocorrelation of the signal is estimated from a register of 10^6 noise-free samples, and the Toeplitz matrix $\boldsymbol{C}$ is built from these estimates.

In order to check the accuracy of the theoretical distributions of the GLR statistic given in Theorem 2, these are compared in Fig. 1 against the empirical histograms of the *iterative GLRT* scheme, for a setting with $M = 4$ antennas and $N = 128$. The PSD of the signal is constant within its support, which equals half the Nyquist bandwidth. A very good agreement is observed, even for this moderate value of the sample size N.

To establish suitable benchmarks for the proposed schemes, we additionally consider the three following detectors, all of which incorporate knowledge of C.

1) A direct generalization of the single-antenna detector of Quan *et al.* [21], in which the following statistic is compared against a threshold:

$$T_{\text{Quan et al.}} = \text{tr}(\boldsymbol{Y}\boldsymbol{C}\boldsymbol{Y}^H). \qquad (49)$$

2) An *energy ratio* (ER) detector in which (49) is normalized by the total observed energy, in order to cope with the noise uncertainty issue. The corresponding statistic is

$$T_{\text{ER}} = \frac{\text{tr}(\boldsymbol{Y}\boldsymbol{C}\boldsymbol{Y}^H)}{\text{tr}(\boldsymbol{Y}\boldsymbol{Y}^H)}, \qquad (50)$$

which is a direct generalization of (10) to the multi-antenna setting.

3) The "rule T" detector proposed by Stoica & Cedervall in [29], which exploits that the temporal correlation length of the noise is much shorter than that of the signal of interest. Its specification is too lengthy and the reader is referred to [29] for more information. Using the notation from [29], our implementation assumes $\hat{n} = 1$ (signal spatial rank), $M = 16$ (truncation point) and $K = 1$ (number of correlation lags).

A. Detection performance in the low SNR regime

The empirical ROC curves of the different detectors considered are shown in Fig. 2, together with the analytical approximation from Theorem 2, for a setting with $M = 4$, $N = 512$ and $\rho = 0.2$. Two different kinds of signals are considered. The first one is an OFDM-modulated digital TV baseband signal with a bandwidth of 3.805 MHz, sampled at 16 Msps and quantized to 9-bit precision. The signal samples are approximately Gaussian distributed, with a PSD that is almost constant within its support (about 48% of the Nyquist bandwidth). The second signal uses a 16-QAM constellation and square-root raised cosine pulses with roll-off factor 1, and it is sampled at twice its baud rate with random timing offset. In this case, the corresponding samples do not follow a Gaussian distribution. In accordance with the signal model, frequency-flat channels are assumed; the impact of frequency selectivity will be considered in Sec. VI-C.

As can be seen in Fig. 2, the proposed schemes significantly outperform the detectors of Quan *et al.* and Stoica & Cedervall, as well as the Energy Ratio detector. The SNR is sufficiently low for the asymptotic GLRT to achieve the same performance as the iterative GLRT detector. The agreement of the empirical and theoretical ROC curves for the GLR test is again quite good. Also note that the GLR detectors perform noticeably better with the OFDM signal than with the QAM signal. This is explained by the shape of the respective PSDs: in this setting, the inverse SFM given by (47) equals 1.01525 and 1.00705 respectively for the OFDM and QAM signals.

B. Detection performance vs. SNR

The exact GLR detector (8) and its asymptotic version (9) were tested with the same OFDM and QAM signals of the
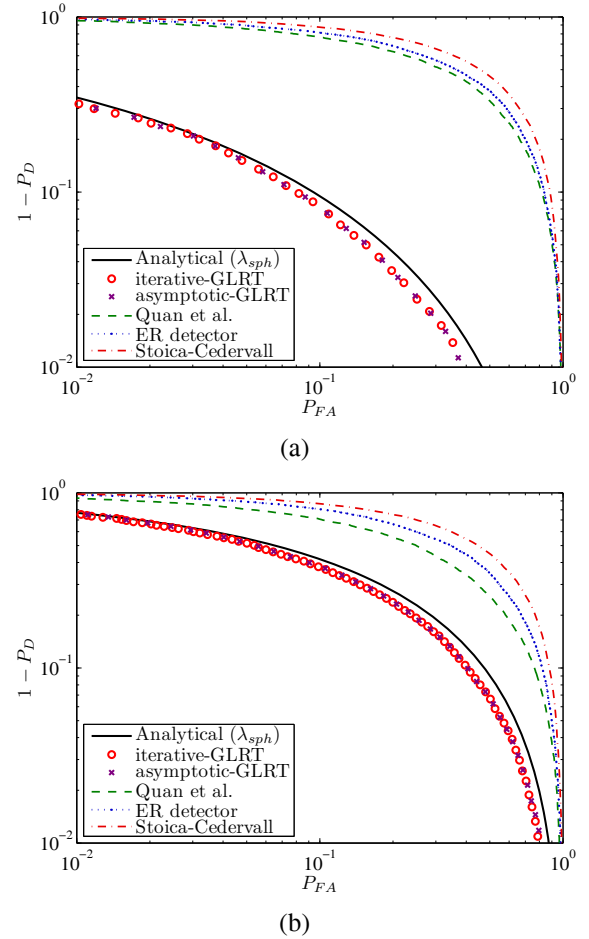

(a)

(b)

Fig. 2. ROC curves for $M = 4$, $N = 512$ and $\rho = 0.2$: (a) OFDM and (b) square root raised cosine signals.
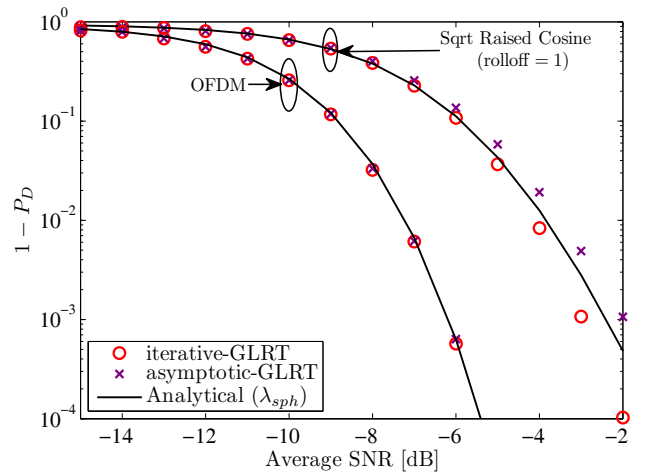

Fig. 3. Missed detection probability vs. SNR for fixed $P_{\text{FA}} = 0.05$, $M = 4$ and $N = 128$.

previous section, but now varying the operational SNR in order to check the performance loss incurred by the asymptotic GLR scheme. Fig. 3 shows the probability of missed detection in terms of the average SNR per antenna, ρ/M, for $M = 4$, $N = 128$ and fixed $P_{\text{FA}} = 0.05$. Notice the good match between
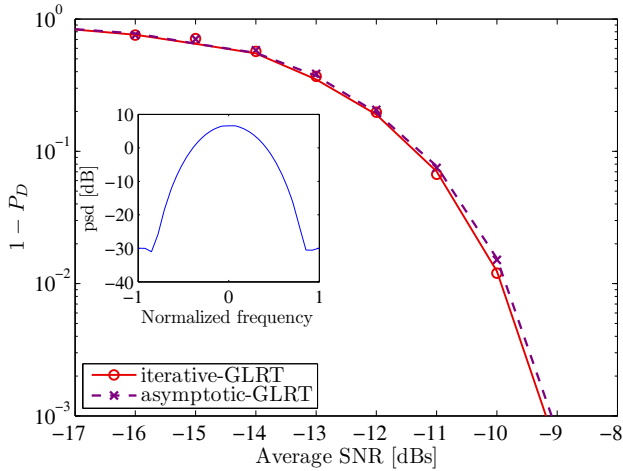
Fig. 4. Detection performance vs. SNR for fixed $P_{FA} = 0.05$, with $M = 4$ and $N = 256$. Lines: frequency-flat channels. Markers: Frequency-selective WINNER II channel model.

the empirical and analytical (asymptotic) results over the SNR range considered, which is not limited to just the very low SNR region. The better results observed for the OFDM signal are again explained by the fact that the corresponding PSD is "less flat". Also note that the two versions of the GLR detector yield the same result with the OFDM signal; this is as expected, since the PSD of the OFDM signal is very close to the binary-valued eigenspectrum case for which both detectors are equivalent (cf. Sec IV-B). On the other hand, for the raised cosine PSD the asymptotic form of the GLR detector presents a performance loss that increases with the SNR. Hence, with non-binary valued power spectra, this detector offers a tradeoff between complexity and performance.

C. Effect of frequency selective channels

To close this section, we consider a realistic scenario based on the parameters of the GSM system. The separation between GSM channels is 200 kHz, although in a given geographical area two adjacent channels cannot be active in order to avoid interference. Thus, the effective channel separation is 400 kHz, which is the same as the approximate bandwidth of the GSM signal, based on Gaussian Minimum Shift Keying (GMSK) [43]. We thus synthesized samples of a baseband GMSK waveform based on GSM parameters at a sampling rate of 400 ksps.

Frequency-selective channels were generated according to the WINNER Phase II model [44] with Profile C1 (suburban), and Non-line-of-sight (NLOS). The central frequency is 1.8 GHz and the channel bandwidth is 400 kHz. Transmitter and receiver are randomly distributed on a 10 km-side square. For the receiver we take a linear array with $M = 4$ antennas with inter-element separation of 10 cm. This channel is rescaled to obtain a fixed operational SNR and is applied to the GSM waveform. Fig. 4 shows the probability of detection of the two proposed schemes vs. average per antenna SNR for fixed $P_{FA} = 0.05$. The inset also shows the PSD of the GMSK waveform. It is observed that both the exact

and asymptotic versions of the GLR detector are robust to frequency selectivity effects, and in fact the results obtained virtually match those of a frequency-flat scenario.

We note that some small degree of non-whiteness in the combined spectral content of noise plus interfering broadband sources should be expected within the detector's bandwidth due to random frequency selectivity affecting the secondary users (assuming a uniform PSD at the transmitter over the detector's bandwidth). Nonetheless, for the detection of a comparatively narrow-band primary signal in the simulated WINNER Phase II channel model, this effect (fluctuations in the tenths of dB range over a span of 400 kHz) is less critical than the mitigation of spatially correlated interference. We should remark as reported in [45], that under all real-world conditions, unavoidable model uncertainties appear in terms of an SNR wall that establishes a trade-off between primary user's SNR and detector robustness. That is, below a minimum SNR, *any* detector ceases to operate reliably regardless of the duration of the observation window. Although worth mentioning, a detailed statistical analysis of these degrading effects[5] falls out of the scope of this paper and has not been included for reasons of space.

VII. CONCLUSIONS

The multiantenna GLR detector for a Gaussian signal of known (up to a scaling) PSD and spatial rank one in temporally white Gaussian noise-plus-interference of arbitrary and unknown spatial correlation has been derived in terms of a univariate optimization problem over an SNR parameter. An analytical expression for the statistical distributions of the test in the low SNR regime has been obtained, resulting in close agreement with empirical results for realistic scenarios featuring practical (non-Gaussian) communication signals. This is attributed to the fact that, as the detector statistic is based on sample correlations, the underlying distribution of individual samples becomes asymptotically irrelevant for large data records. The spectral flatness measure at the output of the minimum variance beamformer was found to determine the performance of the GLR detector, such that less spectrally flat signals enjoy improved detectability. At the other extreme, temporally white signals are not detectable under the model considered. The asymptotic version of the GLR detector for low SNR was also derived, resulting in a closed-form test with much lower computational cost to which the exact GLR detector also reduces if the signal PSD is binary-valued.

APPENDIX

A. Proof of (17)

Letting $q \doteq \Sigma^{-1/2} h$, one has

$$
\begin{aligned}
\boldsymbol{R}_1^{-1} \qquad & (51) \\
= \quad & [\boldsymbol{C} \otimes \boldsymbol{h}\boldsymbol{h}^H + \boldsymbol{I}_N \otimes \boldsymbol{\Sigma}]^{-1} \\
= \quad & (\boldsymbol{I}_N \otimes \boldsymbol{\Sigma}^{-1/2})[\boldsymbol{C} \otimes \boldsymbol{q}\boldsymbol{q}^H + \boldsymbol{I}_{MN}]^{-1}(\boldsymbol{I}_N \otimes \boldsymbol{\Sigma}^{-1/2}).
\end{aligned}
$$

---

[5]For example, detection within the steep spectral roll-off region of the broadband interferer(s).

Recalling the definition (14) of $G(\rho)$, we now claim that

$$[C \otimes qq^H + I_{MN}]^{-1} = I_{MN} - CG(\rho) \otimes qq^H, \quad (52)$$

which can be checked by directly multiplying the right-hand side of (52) by $C \otimes qq^H + I_{MN}$ to see that it yields the identity matrix (using that C and $G(\rho)$ share the same eigenvectors). Substituting (52) back into (51) and noting from (14) that $\Sigma^{-1/2}q = u_\parallel$, the desired result is obtained.

B. Proof of (18)

From (17), and using the property $\text{tr}(A_1^T A_2^H A_3 A_4) = \text{vec}(A_2)^H (A_1 \otimes A_3)\text{vec}(A_4)$ [47], one has

$$y^H R_1^{-1} y \quad (53)$$
$$= \text{vec}(Y)^H (I_N \otimes \Sigma^{-1} - CG(\rho) \otimes u_\parallel u_\parallel^H)\text{vec}(Y)$$
$$= \text{tr}(\Sigma^{-1} YY^H) - \text{tr}(u_\parallel u_\parallel^H Y (CG(\rho))^T Y^H). \quad (54)$$

Applying now the identity $CG(\rho) = \rho^{-1}[I_N - G(\rho)]$, which follows from the definition (14) of $G(\rho)$, noting that $G^T(\rho) = G^*(\rho)$, and then rearranging terms yields the desired result.

C. Proof of Lemma 3

First, note that $W_\perp = [U_\perp \ \bar{h}]\tilde{\Gamma}_\perp [U_\perp \ \bar{h}]^H$ constitutes a full EVD of $W_\perp$, where

$$\tilde{\Gamma}_\perp \doteq \begin{bmatrix} \Gamma_\perp & \\ & 0 \end{bmatrix}. \quad (55)$$

Then the determinant of Σ^{-1} can be written as

$$\det(\Sigma^{-1})$$
$$= \det(W_\parallel + W_\perp)$$
$$= \det\left(\gamma_\parallel \bar{u}_\parallel \bar{u}_\parallel^H + [U_\perp \ \bar{h}]\tilde{\Gamma}_\perp [U_\perp \ \bar{h}]^H\right) \quad (56)$$
$$= \det\left(\gamma_\parallel [U_\perp \ \bar{h}]^H \bar{u}_\parallel \bar{u}_\parallel^H [U_\perp \ \bar{h}] + \tilde{\Gamma}_\perp\right) \quad (57)$$
$$= \det(\tilde{\Gamma}_\perp) + \gamma_\parallel \bar{u}_\parallel^H [U_\perp \ \bar{h}] \text{adj}(\tilde{\Gamma}_\perp)[U_\perp \ \bar{h}]^H \bar{u}_\parallel \quad (58)$$

where in the last step we used the fact that $\det(A + pq^H) = \det A + q^H \text{adj}(A)p$ [46]. Note now from (55) that $\det(\tilde{\Gamma}_\perp) = 0$, whereas $\text{adj}(\tilde{\Gamma}_\perp) = \det(\Gamma_\perp)e_M e_M^H$. The desired result then follows.

D. Proof of Lemma 4

Let $Y_\perp \doteq (I_M - \bar{h}\bar{h}^H)Y$ be the orthogonal projection of the data matrix onto the subspace orthogonal to h. Let $\hat{R}_\perp \doteq \frac{1}{N} Y_\perp Y_\perp^H$ be the corresponding sample covariance matrix, with (economy-size) EVD $\hat{R}_\perp = QDQ^H$ (i.e., $D \in \mathbb{R}^{(M-1)\times(M-1)}$ is positive diagonal and $Q \in \mathbb{C}^{M\times(M-1)}$ has orthonormal columns). Note that since $U_\perp^H \bar{h} = 0$, then $U_\perp^H \hat{R} U_\perp = U_\perp^H \hat{R}_\perp U_\perp$. Therefore, by virtue of Hadamard's inequality [46], the cost J satisfies

$$J(U_\perp) = \det\left[\text{diag}(U_\perp^H QDQ^H U_\perp)\right]$$
$$\geq \det(U_\perp^H QDQ^H U_\perp), \quad (59)$$

with equality holding in (59) iff the matrix between parentheses is diagonal. This happens if $U_\perp = Q$, which is feasible

since its columns are orthonormal and $Q^H \bar{h} = 0$. Thus $J_{\min} = \det D$.

Let now $\tilde{Q} \doteq [Q \ \bar{h}]$, which is unitary. Write the inverse of $\tilde{Q}^H \hat{R}\tilde{Q}$ in terms of the adjoint matrix as

$$(\tilde{Q}^H \hat{R}\tilde{Q})^{-1} = \frac{\text{adj}(\tilde{Q}^H \hat{R}\tilde{Q})}{\det(\tilde{Q}^H \hat{R}\tilde{Q})}, \quad (60)$$

which can be rewritten as

$$\tilde{Q}^H \hat{R}^{-1}\tilde{Q} = \frac{\text{adj}(\tilde{Q}^H \hat{R}\tilde{Q})}{\det \hat{R}}. \quad (61)$$

The (M, M) element of this matrix is therefore

$$\bar{h}^H \hat{R}^{-1}\bar{h} = \frac{e_M^H \text{adj}(\tilde{Q}^H \hat{R}\tilde{Q})e_M}{\det \hat{R}} = \frac{\det(Q^H \hat{R}Q)}{\det \hat{R}}. \quad (62)$$

Finally, observe that $D = Q^H \hat{R}_\perp Q = Q^H \hat{R}Q$ because $Q^H \bar{h} = 0$. This, together with (62), yields the desired result.

Remark 1: Note from (29) that the ML estimate of $\Gamma_\perp$ is given by

$$\hat{\Gamma}_\perp = \left[\text{diag}\left(\hat{U}_\perp^H \hat{R}\hat{U}_\perp\right)\right]^{-1} = D^{-1}, \quad (63)$$

since the ML estimate of $U_\perp$ is $\hat{U}_\perp = Q$. Hence, the ML estimate of $W_\perp = U_\perp \Gamma_\perp U_\perp^H$ turns out to be $\hat{W}_\perp = \hat{U}_\perp \hat{\Gamma}_\perp \hat{U}_\perp^H = QD^{-1}Q^H = \hat{R}_\perp^\dagger$.

Remark 2: We note that the Moore-Penrose pseudoinverse $W_\perp^\dagger = U_\perp \Gamma_\perp^{-1} U_\perp^H$ fulfils $W_\perp^\dagger \bar{h} = 0$, with $W_\perp W_\perp^\dagger$ the projector onto the subspace orthogonal to $\bar{h}$ and $W_\perp W_\perp^\dagger + \bar{h}\bar{h}^H = I_M$. Then, to verify (24), we premultiply by Σ^{-1}: $\Sigma^{-1}(\Pi W_\perp^\dagger \Pi^H + \rho^{-1}hh^H) = W_\perp W_\perp^\dagger \Pi^H + I_M - \Pi^H = I_M - \bar{h}\bar{h}^H \Pi^H$. As $\Pi\bar{h} = 0$, the previous result is I_M, verifying that (24) is a valid expression for Σ.

E. Proof of Lemma 5

Since $F(u_1, u_2)$ is invariant to scalings in u_1 and u_2, its minimization is equivalent to the problem

$$\min_{u_1, u_2} (u_1^H A_1 u_1)(u_2^H A_2 u_2) \quad \text{subject to} \quad |u_1^H u_2|^2 = c^2, \quad (64)$$

with $c^2 > 0$ an arbitrary positive constant. Thus, consider the Langrangian

$$\mathcal{L} = (u_1^H A_1 u_1)(u_2^H A_2 u_2) - \chi(|u_1^H u_2|^2 - c^2), \quad (65)$$

with χ the Lagrange multiplier. Equating the gradient of $\mathcal{L}$ w.r.t. u_1, u_2 to zero, we respectively obtain

$$(u_2^H A_2 u_2)A_1 u_1 = \chi(u_2^H u_1)u_2, \quad (66)$$
$$(u_1^H A_1 u_1)A_2 u_2 = \chi(u_1^H u_2)u_1. \quad (67)$$

To discard the possibility of having $u_1^H u_2 = 0$, note that this would mean that the left-hand sides of (66)-(67) are both zero, which can only happen if $u_1 = u_2 = 0$ since A_1, A_2 are positive definite. Thus, assuming $u_1 \neq 0, u_2 \neq 0$, solving for χ in (66)-(67) yields

$$\chi = \frac{(u_1^H A_1 u_1)(u_2^H A_2 u_2)}{|u_1^H u_2|^2}, \quad (68)$$

that is, the Lagrange multiplier takes the value of the attained cost. Substituting now the value of $\boldsymbol{u}_2$ from (66) and of χ from (68) into (67), one has $\boldsymbol{A}_2\boldsymbol{A}_1\boldsymbol{u}_1 = \chi\boldsymbol{u}_1$. Similarly, it also holds that $\boldsymbol{A}_1\boldsymbol{A}_2\boldsymbol{u}_2 = \chi\boldsymbol{u}_2$. Thus χ is an eigenvalue of $\boldsymbol{A}_2\boldsymbol{A}_1$ (resp. $\boldsymbol{A}_1\boldsymbol{A}_2$) with associated eigenvector $\boldsymbol{u}_1$ (resp. $\boldsymbol{u}_2$). Hence the cost is minimized when χ is the *smallest* eigenvalue of these matrices. Equivalently, $\boldsymbol{u}_1$ (resp. $\boldsymbol{u}_2$) is a *generalized* eigenvector [39] of the pair $(\boldsymbol{A}_1, \boldsymbol{A}_2^{-1})$ [resp. $(\boldsymbol{A}_2, \boldsymbol{A}_1^{-1})$].

F. Proof of Theorem 2

The asymptotic distributions of the GLR statistic T_{GLR} in the weak signal regime are given in [31, Sec. 6.5] for a test of the form

$$\mathcal{H}_0 : \boldsymbol{\theta}_r = \boldsymbol{\theta}_{r_0}, \boldsymbol{\theta}_s \qquad \mathcal{H}_1 : \boldsymbol{\theta}_r \neq \boldsymbol{\theta}_{r_0}, \boldsymbol{\theta}_s, \qquad (69)$$

where $\boldsymbol{\theta}_r \in \mathbb{R}^r$, $\boldsymbol{\theta}_s \in \mathbb{R}^s$. One wishes to test whether $\boldsymbol{\theta}_r = \boldsymbol{\theta}_{r_0}$ as opposed to $\boldsymbol{\theta}_r \neq \boldsymbol{\theta}_{r_0}$, and $\boldsymbol{\theta}_s$ is a vector of nuisance parameters which are unknown (but the same) under either hypothesis. The distributions are

$$2\log T_{\mathrm{GLR}} \sim \begin{cases} \chi_r^2, & \text{under } \mathcal{H}_0, \\ \chi_r'^2(\alpha), & \text{under } \mathcal{H}_1, \end{cases} \qquad (70)$$

where the non-centrality parameter α is given in [31, eq. (6.24)]. In our case, the nuisance parameters are given by the elements of the noise spatial covariance $\boldsymbol{\Sigma}$. On the other hand, the parameter to be tested is $\boldsymbol{h} = \boldsymbol{0}$ versus $\boldsymbol{h} \neq \boldsymbol{0}$. However, since the p.d.f. depends on $\boldsymbol{h}$ only through $\boldsymbol{h}\boldsymbol{h}^H$, which is invariant to multiplication of $\boldsymbol{h}$ by a unit-magnitude complex scalar, we can fix the imaginary part of the last element of $\boldsymbol{h}$ (say) to zero. Thus the vector $\boldsymbol{\theta}_r$ comprises the real parts of the elements of $\boldsymbol{h}$, plus the imaginary parts of the elements of $\boldsymbol{h}$ save the last one. The size of $\boldsymbol{\theta}_r$ is therefore $r = 2M - 1$.

Direct application of [31, eq. (6.24)] to determine α is involved for the model at hand. We propose an alternative approach based on the following result, whose proof will be presented in turn.

Lemma 6: Consider the GLR statistic T from (8). Then one has

$$\lim_{N \to \infty} \mathbb{E}[|T - \tilde{T}(\rho_0)|^2] = 0, \qquad (71)$$

where $\rho_0 = \boldsymbol{h}^H\boldsymbol{\Sigma}^{-1}\boldsymbol{h} = \boldsymbol{u}_\parallel^H\boldsymbol{h}$ is the true value of the SNR, and

$$\tilde{T}(\rho) \doteq \left[\frac{\frac{1}{N}\mathrm{tr}(\boldsymbol{I}_N + \rho\boldsymbol{C})}{[\det(\boldsymbol{I}_N + \rho\boldsymbol{C})]^{1/N}}\right]^N. \qquad (72)$$

From Lemma 6, it follows that

$$\lim_{N \to \infty} \mathbb{E}[2\log T] = \lim_{N \to \infty} 2\log\tilde{T}(\rho_0) \qquad (73)$$

$$= \lim_{N \to \infty} 2N\log\frac{\frac{1}{N}\mathrm{tr}(\boldsymbol{I}_N + \rho_0\boldsymbol{C})}{[\det(\boldsymbol{I}_N + \rho_0\boldsymbol{C})]^{1/N}}.$$

On the other hand, note that if $x \sim \chi_r'^2(\alpha)$, then $\mathbb{E}[x] = r + \alpha$. Hence,

$$\lim_{N \to \infty} \mathbb{E}[2\log T] = (2M - 1) + \lim_{N \to \infty} \alpha. \qquad (74)$$

It follows from (73)-(74) that for sufficiently large N, $\alpha \gg 2M - 1$ and one may approximate $\alpha \approx 2\log\tilde{T}(\rho_0) = \alpha_{\mathrm{sph}}(\rho_0)$.

Proof of Lemma 6: Write the data matrix as $\boldsymbol{Y} = \boldsymbol{h}\boldsymbol{s}^T + \boldsymbol{\Sigma}^{1/2}\boldsymbol{W}$, where the entries of $\boldsymbol{W} \doteq \boldsymbol{\Sigma}^{-1/2}[\boldsymbol{z}_1\ \boldsymbol{z}_2\ \cdots\ \boldsymbol{z}_N]$ are independent zero-mean complex circular Gaussian random variables with unit variance. It follows that for $\boldsymbol{A} \in \mathbb{C}^{N \times N}$,

$$\mathbb{E}[\boldsymbol{Y}\boldsymbol{A}\boldsymbol{Y}^H] = \mathrm{tr}(\boldsymbol{C}^*\boldsymbol{A})\boldsymbol{h}\boldsymbol{h}^H + \mathrm{tr}(\boldsymbol{A})\boldsymbol{\Sigma}. \qquad (75)$$

Now note that $\frac{1}{N}\boldsymbol{Y}\boldsymbol{A}\boldsymbol{Y}^H$ is a consistent estimator of its mean for the matrices $\boldsymbol{A}$ in this paper, and hence

$$\frac{1}{N}\boldsymbol{Y}\boldsymbol{A}\boldsymbol{Y}^H \overset{\mathrm{var}}{=} \frac{\mathrm{tr}(\boldsymbol{C}^*\boldsymbol{A})}{N}\boldsymbol{h}\boldsymbol{h}^H + \frac{\mathrm{tr}(\boldsymbol{A})}{N}\boldsymbol{\Sigma}, \qquad (76)$$

where $\overset{\mathrm{var}}{=}$ denotes stochastic convergence in variance [48], i.e., $a_N \overset{\mathrm{var}}{=} b_N$ iff $\lim_{N \to \infty} \mathbb{E}[|a_N - b_N|^2] = 0$, applying componentwise for matrices. Let us rewrite the GLR statistic T as

$$T = \max_{\rho \geq 0} \frac{\lambda_{\min}^{-N}\left[(\frac{1}{N}\boldsymbol{Y}\boldsymbol{Y}^H)^{-1}(\frac{1}{N}\boldsymbol{Y}\boldsymbol{G}^*(\rho)\boldsymbol{Y}^H)\right]}{[\det\boldsymbol{G}^*(\rho)]^{-1}}. \qquad (77)$$

Then, taking into account that $\mathrm{tr}\,\boldsymbol{C}^* = \mathrm{tr}\,\boldsymbol{I}_N = N$, the asymptotic behavior of T can be found:

$$T \overset{\mathrm{var}}{=} \qquad (78)$$

$$\max_{\rho \geq 0} \frac{\lambda_{\min}^{-N}\left[\frac{1}{N}\mathrm{tr}\,\boldsymbol{G}^*(\rho)\cdot(\boldsymbol{\Sigma}+\boldsymbol{h}\boldsymbol{h}^H)^{-1}(\boldsymbol{\Sigma}+a(\rho)\boldsymbol{h}\boldsymbol{h}^H)\right]}{[\det\boldsymbol{G}^*(\rho)]^{-1}} =$$

$$\max_{\rho \geq 0} \left[\frac{\frac{1}{N}\mathrm{tr}\,\boldsymbol{G}^*(\rho)}{[\det\boldsymbol{G}^*(\rho)]^{1/N}}\cdot\lambda_{\min}\left[\boldsymbol{I}_M - \frac{1-a(\rho)}{1+\rho_0}\boldsymbol{u}_\parallel\boldsymbol{h}^H\right]\right]^{-N} \qquad (79)$$

where

$$a(\rho) \doteq \frac{\mathrm{tr}(\boldsymbol{C}^*\boldsymbol{G}^*(\rho))}{\mathrm{tr}\,\boldsymbol{G}^*(\rho)}. \qquad (80)$$

In order to obtain (79), the Matrix Inversion Lemma has been applied. Denote now the matrix in brackets in (79) by $\boldsymbol{B}$. Then $\boldsymbol{B}$ has an eigenvalue equal to one with multiplicity $M - 1$, since $\boldsymbol{B}\boldsymbol{v} = \boldsymbol{v}$ for any $\boldsymbol{v}$ such that $\boldsymbol{v}^H\boldsymbol{h} = 0$. The remaining eigenvalue is associated to the eigenvector $\boldsymbol{u}_\parallel$ and is given by

$$\lambda = \frac{1 + \rho_0 a(\rho)}{1 + \rho_0}. \qquad (81)$$

We claim that $a(\rho) \leq 1$ for all $\rho \geq 0$, so that λ in (81) is less than one and hence it is the minimum eigenvalue of $\boldsymbol{B}$. First, note that $a(0) = \mathrm{tr}\,\boldsymbol{C}^*/\mathrm{tr}\,\boldsymbol{I}_N = 1$. On the other hand, for $\rho > 0$, since $\boldsymbol{C}$ and $\boldsymbol{G}(\rho)$ share the same eigenvectors, we can write $a(\rho)$ in (80) in terms of the eigenvalues $\{\lambda_n\}_{n=0}^{N-1}$ of $\boldsymbol{C}$ as

$$a(\rho) = \frac{\sum_{n=0}^{N-1}\frac{\lambda_n}{1+\rho\lambda_n}}{\sum_{n=0}^{N-1}\frac{1}{1+\rho\lambda_n}} \qquad (82)$$

$$= \frac{1}{\rho}\left[\frac{N}{\sum_{n=0}^{N-1}\frac{1}{1+\rho\lambda_n}} - 1\right], \qquad (83)$$

where (83) follows from the fact that $N = \sum_n\frac{1}{1+\rho\lambda_n} + \rho\sum_n\frac{\lambda_n}{1+\rho\lambda_n}$. Note now that for positive numbers $\{x_n\}_{n=0}^{N-1}$, the Cauchy-Schwarz inequality can be invoked to show that $\left(\frac{1}{N}\sum_n x_n\right)\left(\frac{1}{N}\sum_n\frac{1}{x_n}\right) \geq 1$. Hence, taking $x_n = 1 + \rho\lambda_n$, and noting that $\sum_n\lambda_n = N$, one has $\frac{(1+\rho)}{N}\sum_n\frac{1}{1+\rho\lambda_n} \geq 1$.

Applying this to (83) yields $a(\rho) \leq 1$, as desired. Therefore, from (79), one has

$$T \stackrel{\text{var}}{=} \left[\min_{\rho \geq 0} \left[\frac{\frac{1}{N} \operatorname{tr} \boldsymbol{G}^*(\rho)}{\det^{1/N} \boldsymbol{G}^*(\rho)} \cdot \frac{1 + \rho_0 a(\rho)}{1 + \rho_0} \right] \right]^{-N}. \tag{84}$$

Note from (80) that $1 + \rho_0 a(\rho)$ can be written as

$$1 + \rho_0 a(\rho) = \frac{\operatorname{tr}[(\boldsymbol{I}_N + \rho_0 \boldsymbol{C}^*) \boldsymbol{G}^*(\rho)]}{\operatorname{tr} \boldsymbol{G}^*(\rho)}. \tag{85}$$

Substituting (85) in (84),

$$T \stackrel{\text{var}}{=} \left(\frac{1}{N(1 + \rho_0)} \min_{\rho \geq 0} H(\rho) \right)^{-N}, \tag{86}$$

where we have introduced

$$H(\rho) \doteq \frac{\operatorname{tr}[(\boldsymbol{I}_N + \rho_0 \boldsymbol{C}^*) \boldsymbol{G}^*(\rho)]}{\det^{1/N} \boldsymbol{G}^*(\rho)}. \tag{87}$$

Since $\boldsymbol{G}^*(\rho) = (\boldsymbol{I}_N + \rho \boldsymbol{C}^*)^{-1}$, applying Jensen's inequality one has

$$\log H(\rho)$$
$$= \log \sum_{n=0}^{N-1} \frac{1 + \rho_0 \lambda_n}{1 + \rho \lambda_n} + \frac{1}{N} \sum_{n=0}^{N-1} \log(1 + \rho \lambda_n) \tag{88}$$
$$\geq \log N + \frac{1}{N} \sum_{n=0}^{N-1} \log \frac{1 + \rho_0 \lambda_n}{1 + \rho \lambda_n} + \frac{1}{N} \sum_{n=0}^{N-1} \log(1 + \rho \lambda_n) \tag{89}$$
$$= \log N + \frac{1}{N} \sum_{n=0}^{N-1} \log(1 + \rho_0 \lambda_n) \tag{90}$$
$$= \log H(\rho_0), \tag{91}$$

showing that the minimum is attained at $\rho = \rho_0$. Hence,

$$T \stackrel{\text{var}}{=} \left(\frac{\left[\prod_{n=0}^{N-1} (1 + \rho_0 \lambda_n) \right]^{1/N}}{1 + \rho_0} \right)^{-N} = \tilde{T}(\rho_0), \tag{92}$$

as was to be shown.

Remark 3: Asymptotic unimodality in ρ. In addition to being minimized at $\rho = \rho_0$, the function $H(\rho)$ in (87) has no other local minimum in $\rho \geq 0$ provided that the detectability condition $\boldsymbol{C} \neq \boldsymbol{I}_N$ holds, as we show next. Differentiating (88),

$$\frac{\mathrm{d} \log H(\rho)}{\mathrm{d} \rho} = \frac{1}{N} \sum_{n=0}^{N-1} \frac{\lambda_n}{1 + \rho \lambda_n}$$
$$- \left[\sum_{n=0}^{N-1} \frac{1 + \rho_0 \lambda_n}{1 + \rho \lambda_n} \right]^{-1} \sum_{n=0}^{N-1} \frac{\lambda_n(1 + \rho_0 \lambda_n)}{(1 + \rho \lambda_n)^2}. \tag{93}$$

Let now

$$\delta \doteq \rho_0 - \rho, \qquad x_n \doteq \frac{\lambda_n}{1 + \rho \lambda_n} \quad \Rightarrow \quad \frac{1 + \rho_0 \lambda_n}{1 + \rho \lambda_n} = 1 + \delta x_n, \tag{94}$$

so that (93) becomes

$$\frac{\mathrm{d} \log H(\rho)}{\mathrm{d} \rho} \tag{95}$$
$$= \frac{1}{N} \sum_n x_n - \left[\sum_n (1 + \delta x_n) \right]^{-1} \sum_n x_n(1 + \delta x_n)$$
$$= - \underbrace{\left[\sum_n (1 + \delta x_n) \right]^{-1}}_{\doteq c_1} \underbrace{\left[\sum_n x_n^2 - \frac{1}{N} \left(\sum_n x_n \right)^2 \right]}_{\doteq c_2} \cdot \delta. \tag{96}$$

Note from (94) that $c_1 > 0$ for all $\rho \geq 0$. On the other hand, from the Cauchy-Schwarz inequality one has $c_2 \geq 0$, with equality iff all x_n are equal. But this would imply that all eigenvalues of $\boldsymbol{C}$ are equal, i.e., $\boldsymbol{C} = \boldsymbol{I}_N$; thus, $c_2 > 0$. We conclude that the derivative (96) is negative if $\delta > 0$ and positive if $\delta < 0$, which shows that $H(\rho)$ has a unique minimum at $\rho = \rho_0$.

G. Gradient descent computations

The goal is to maximize $t(\rho)$ in (8) over $\rho \geq 0$. Equivalently, we can minimize $t_0(\rho) \doteq t^{-1/N}(\rho)$ by means of an iterative gradient descent of the form $\rho_{k+1} = \rho_k - \mu_k t_0'(\rho_k)$, with $\mu_k > 0$ a suitable stepsize sequence. In the simulations, we initialized $\rho_0 = 1$, $\mu_0 = 100$. The stepsize is reduced as per $\mu_k = 0.25 \mu_{k-1}$ whenever there is a sign change in the descent direction. The stopping criterion adopted is $|t_0'(\rho_k)| < 10^{-5}$ with a maximum of 100 iterations.

The required derivative is obtained as follows. Denote the minimum unit-norm eigenvector of $\boldsymbol{V}^H (\boldsymbol{I}_N + \rho \boldsymbol{C}^*)^{-1} \boldsymbol{V}$ by $\boldsymbol{u}_0(\rho)$. Then we can write

$$\lambda_0(\rho) \doteq \lambda_{\min}[\boldsymbol{V}^H (\boldsymbol{I}_N + \rho \boldsymbol{C}^*)^{-1} \boldsymbol{V}] \tag{97}$$
$$= \boldsymbol{u}_0^H(\rho) \left[\boldsymbol{V}^H (\boldsymbol{I}_N + \rho \boldsymbol{C}^*)^{-1} \boldsymbol{V} \right] \boldsymbol{u}_0(\rho), \tag{98}$$

and therefore, using the expressions for eigenvalues derivatives in [47], one has

$$t_0'(\rho)$$
$$= [\det(\boldsymbol{G}^*(\rho))]^{-1/N} \cdot \frac{\partial}{\partial \rho} \left(\boldsymbol{u}_0^H(\rho) \left[\boldsymbol{V}^H \boldsymbol{G}^*(\rho) \boldsymbol{V} \right] \boldsymbol{u}_0(\rho) \right)$$
$$+ \lambda_0(\rho) \cdot \frac{\partial}{\partial \rho} [\det(\boldsymbol{G}^*(\rho))]^{-1/N}$$
$$= - [\det(\boldsymbol{G}^*(\rho))]^{-1/N} \cdot$$
$$\cdot \left(\boldsymbol{u}_0^H(\rho) \left[\boldsymbol{V}^H \boldsymbol{G}^*(\rho) \boldsymbol{C}^* \boldsymbol{G}^*(\rho) \boldsymbol{V} \right] \boldsymbol{u}_0(\rho) \right)$$
$$+ \lambda_0(\rho) \cdot [\det(\boldsymbol{G}^*(\rho))]^{-1/N} \frac{1}{N} \operatorname{tr} \left(\boldsymbol{G}^*(\rho) \boldsymbol{C}^* \right)$$
$$= \left[-\boldsymbol{u}_0^H(\rho) \boldsymbol{A}(\rho) \boldsymbol{u}_0(\rho) + \frac{1}{N} \lambda_0(\rho) \cdot \operatorname{tr} \left(\boldsymbol{G}^*(\rho) \boldsymbol{C}^* \right) \right] \cdot$$
$$\cdot [\det(\boldsymbol{G}^*(\rho))]^{-1/N},$$

where $\boldsymbol{A}(\rho) \doteq \boldsymbol{V}^H \boldsymbol{G}^*(\rho) \boldsymbol{C}^* \boldsymbol{G}^*(\rho) \boldsymbol{V}$. Using the EVD $\boldsymbol{C} = \boldsymbol{F} \boldsymbol{\Lambda} \boldsymbol{F}^H$ and the fact that $\boldsymbol{F}^H$ approaches the orthonormal DFT matrix for large N, one can precompute $\boldsymbol{F}^H \boldsymbol{V}^*$ efficiently by applying the FFT to the columns of $\boldsymbol{V}^*$. The elements of $\boldsymbol{A}(\rho)$ are then obtained as weighted crosscorrelations between

the columns of $\boldsymbol{F}^H \boldsymbol{V}^*$, where the weights are of the form $\lambda_n/(1+\rho\lambda_n)^2$. A similar approach can be used to compute $\boldsymbol{V}^H \boldsymbol{G}^*(\rho)\boldsymbol{V}$ efficiently.

REFERENCES

[1] M. Wax, T. Kailath, "Detection of signals by information-theoretic criteria," *IEEE Trans. Acoust., Speech, Signal Process.,* vol. ASSP-33, pp. 387–392, Apr. 1985.

[2] M. Viberg, B. Ottersten, T. Kailath, "Detection and estimation in sensor arrays using weighted subspace fitting," *IEEE Trans. Signal Process.,* vol. 39, pp. 2436–2449, Nov. 1991.

[3] L. L. Scharf, B. Friedlander, "Matched subspace detectors," *IEEE Trans. Signal Process.,* vol. 42, pp. 2146–2157, Aug. 1994.

[4] O. Besson, S. Kraut, L. L. Scharf, "Detection of an unknown rank-one component in white noise," *IEEE Trans. Signal Process.,* vol. 54, pp. 2835–2839, Jul. 2006.

[5] A. Leshem, A.-J. van der Veen, "Multichannel detection of Gaussian signals with uncalibrated receivers," *IEEE Signal Process. Lett.,* vol. 8, pp. 120–122, Apr. 2001.

[6] S. A. Vorobyov, A. Gershman, Kon Max Wong, "Maximum likelihood direction-of-arrival estimation in unknown noise fields using sparse sensor arrays," *IEEE Trans. Signal Process.,* vol. 53, pp. 34–43, Jan. 2005.

[7] M. Agrawal, S. Prasad, "A modified likelihood function approach to DOA estimation in the presence of unknown spatially correlated Gaussian noise using a uniform linear array," *IEEE Trans. Signal Process.,* vol. 48, pp. 2743–2749, Oct. 2000.

[8] Yuehua Wu, Kwok-Wai Tam, Fu Li, "Determination of number of sources with multiple arrays in correlated noise fields," *IEEE Trans. Signal Process.,* vol. 50, pp. 1257–1260, Jun. 2002.

[9] P. Stoica, Kon Max Wong, Qiang Wu, "On a nonparametric detection method for array signal processing in correlated noise fields," *IEEE Trans. Signal Process.,* vol. 44, pp. 1030–1032, Apr. 1996.

[10] M. Viberg, P. Stoica, B. Ottersten, "Maximum likelihood array processing in spatially correlated noise fields using parameterized signals," *IEEE Trans. Signal Process.,* vol. 45, pp. 996–1004, Apr. 1997.

[11] I. Akyildiz, Won-Yeol Lee, M. Vuran, S. Mohanty, "A survey on spectrum management in cognitive radio networks," *IEEE Commun. Mag.,* vol. 46, pp. 40–48, Apr. 2008.

[12] J. M. Peha, "Sharing spectrum through spectrum policy reform and cognitive radio," *Proc. IEEE,* vol. 97, pp. 708–719, Apr. 2009.

[13] J. Unnikrishnan, V. Veeravalli, "Cooperative sensing for primary detection in cognitive radio," *IEEE J. Sel. Topics Signal Process.,* vol. 2, pp. 18–27, Feb. 2008.

[14] A. Taherpour, M. Nasiri-Kenari, S. Gazor, "Multiple antenna spectrum sensing in cognitive radios," *IEEE Trans. Wireless Commun.,* vol. 9, pp. 814–823, Feb. 2010.

[15] Rui Zhang, Teng Lim, Ying-Chang Liang, Yonghong Zeng, "Multi-antenna based spectrum sensing for cognitive radios: A GLRT approach," *IEEE Trans. Commun.,* vol. 58, pp. 84–88, Jan. 2010.

[16] Pu Wang, Jun Fang, Ning Han, Hongbin Li, "Multiantenna-assisted spectrum sensing for cognitive radio," *IEEE Trans. Veh. Technol.,* vol. 59, pp. 1791–1800, May 2010.

[17] D. Ramírez, G. Vazquez-Vilar, R. López-Valcarce, J. Vía, I. Santamaría, "Detection of rank-P signals in cognitive radio networks with uncalibrated multiple antennas," *IEEE Trans. Signal Process.,* vol. 59, pp. 3764–3774, Aug. 2011.

[18] IEEE 802.22 WRAN Working Group, *IEEE 802.22-2011 Standard for Cognitive Wireless Regional Area Networks (RAN) for Operation in TV Bands.* Piscataway, NJ: IEEE Standards Association, Jul. 2011.

[19] National Telecommunications and Information Administration, *Report 04-413: Potential interference from broadband over power line (BPL) systems to Federal Government radiocommunications at 1.7-80 MHz.* Washington, D.C.: NTIA, Apr. 2004.

[20] J. P. Le Cadre, "Parametric methods for spatial signal processing in the presence of unknown colored noise fields," *IEEE Trans. Acoust., Speech, Signal Process.,* vol. ASSP-37, pp. 965–983, Jul. 1989.

[21] Zhi Quan, Wenyi Zhang, S. J. Shellhammer, A. H. Sayed, "Optimal spectral feature detection for spectrum sensing at very low SNR," *IEEE Trans. Commun.,* vol. 59, pp. 201–212, Jan. 2011.

[22] Wenyi Zhang, H. Poor, Zhi Quan, "Frequency-domain correlation: an asymptotically optimum approximation of quadratic likelihood ratio detectors," *IEEE Trans. Signal Process.,* vol. 58, pp. 969–979, Mar. 2010.

[23] G. Vazquez-Vilar, R. López-Valcace, J. Sala, "Multiantenna spectrum sensing exploiting spectral a priori information," *IEEE Trans. Wireless Comm.,,* Vol.10, No.12, pp. 4345–4355, Dec. 2011.

[24] J. Tellado, *Multicarrier modulation with low PAR: Applications to DSL and wireless.* Norwell, MA: Kluwer Academic, 2000.

[25] A. Dogandžić, A. Nehorai, "Generalized multivariate analysis of variance - A unified framework for signal processing in correlated noise," *IEEE Signal Process. Mag.,* vol. 20, pp. 39–54, Sep. 2003.

[26] S. Kraut, L. L. Scharf, L. T. McWhorter, "Adaptive subspace detectors," *IEEE Trans. Signal Process.,* vol. 49, pp. 1–16, Jan. 2001.

[27] Yuanwei Jin, B. Friedlander, "A CFAR adaptive subspace detector for second-order Gaussian signals," *IEEE Trans. Signal Process.,* vol. 53, pp. 871–884, Mar. 2005.

[28] A. Paulraj, T. Kailath, "Eigenstructure methods for direction-of-arrival estimation in the presence of unknown noise fields," *IEEE Trans. Acoust., Speech, Signal Process.,* vol. ASSP-34, pp. 13–20, Jan. 1986.

[29] P. Stoica, M. Cedervall, "Detection tests for array processing in unknown correlated noise fields," *IEEE Trans. Signal Process.,* vol. 45, pp. 2351–2362, Sep. 1997.

[30] B. C. Levy, *Principles of Signal Detection and Parameter Estimation*, Springer, 2010.

[31] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory.* Englewood Cliffs, NJ: Prentice-Hall, 1998.

[32] R. M. Gray, *Toeplitz and Circulant Matrices: A Review.* NY: Hanover/Now, 2006.

[33] J. Capon, "High-resolution frequency-wavenumber spectrum analysis," *Proc. IEEE,* vol. 57, pp. 1408–1418, Aug. 1969.

[34] A. J. Laub, *Matrix Analysis for Scientists and Engineers*, Soc. for Industrial and Applied Math., Philadelphia, 2005.

[35] H. L. van Trees, *Optimum Array Processing (Detection, Estimation and Modulation Theory)*, Wiley, 2002.

[36] S. M. Kay, *Modern Spectral Estimation: Theory and Applications*, Prentice-Hall, 1987.

[37] D. E. Bernstein, *Matrix Mathematics: Theory, Facts and Formulas, 2nd ed.* Princeton University Press, 2009.

[38] R. A. Horn, C. R. Johnson, *Matrix Analysis.* Cambridge University Press, 1986.

[39] G. H. Golub, C. F. van Loan, *Matrix Computations, 2nd ed.* John Hopkins University Press, 1989.

[40] U. Grenander, G. Szegö, *Toeplitz Forms and Their Applications.* New York, Chelsea Pub. Co., 1984.

[41] A. J. Gray, J. Markel, "A spectral flatness measure for studying the autocorrelation method of linear prediction of speech analysis," *IEEE Trans. Acoust., Speech, Signal Process.,* vol. ASSP-22, pp. 207–217, Jun. 1974

[42] J. P. Dugre, L. L. Scharf, A. Beex, "A note on the measurement of spectral flatness and the calculation of prediction error variances," *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.,* 1980, vol. 5, pp. 607–611.

[43] EN 300 910, "Digital cellular telecommunications system (Phase 2+); Radio transmission and reception (GSM 05.05 version 8.5.1 Release 1999)," ETSI, Nov. 2000.

[44] L. Hentilä, P. Kyösti, M. Käske, M. Narandzic, M. Alatossava, "Matlab implementation of the WINNER Phase II Channel Model v1.1", Online: https://www.ist-winner.org/phase_2_model.html, 2007.

[45] R. Tandra, A. Sahai, "SNR Walls for Signal Detection", *IEEE Journal of Selected Topics in Signal Processing*, Vol.2, No.1, pp. 3-17, Feb. 2008.

[46] P. Lancaster, M. Tismenetsky, *The Theory of Matrices, 2nd ed.* San Diego, CA: Academic Press, 1985.

[47] J. R.Magnus, H. Neudecker, *Matrix Differential Calculus with Applications to Statistics and Econometrics, revised ed.* Chichester, England: Wiley, 1999.

[48] R. J. Serfling, *Approximation Theorems of Mathematical Statistics.* Chichester, England: Wiley, 1980.