MULTIANTENNA GLR DETECTION OF A GAUSSIAN SIGNAL IN SPATIALLY UNCORRELATED NOISE

Josep Sala

Dpt. of Signal Theory and Communications, Technical University of Catalonia (UPC)
c/Jordi Girona 1-3, Campus Nord UPC, D5-114b, 08034 Barcelona, SPAIN
Phone:+34-93-4015894, Fax:+34-93-4016447, E-mail: josep.sala@upc.edu

ABSTRACT

The exact all-SNR Generalized Log-likelihood Ratio Test (GLRT) of a Gaussian rank-one signal impinging on an M-antenna receiver in unknown spatially uncorrelated white noise is derived and compared with some previous low-SNR approximations to this problem. It is confirmed that the coherence matrix used in the mentioned approximations is a sufficient statistic for the GLRT over the complete SNR range. In contrast, this exact, albeit more complex test, requires the unconstrained optimization of a highly nonlinear M-variate complex function which can be addressed using classical optimization techniques. A maximum eigenvalue problem combined with a univariate optimization problem serves the purpose of initialization. Results validate the yet unproven close to optimum performance of previous detectors under the rank-one signal model for the tested SNR's.

1. INTRODUCTION

Spectrum sensing constitutes an important tool in applications ranging from Cognitive Radio to radio-astronomy, where statistical tests are required for assessing the presence or absence of a signal in a given data record. The Generalized Log-Likelihood Ratio Test (GLRT) has been successfully applied to a wide range of scenarios: given a parametrized data model, it evaluates the ratio of probability densities between the two hypotheses (presence vs. absence), each of them maximized with respect to its corresponding parameter set. Therefore, the resulting performance depends on the closeness of the assumed model to the actual data generation process. Robustness to noise uncertainty is an important feature required of this type of detectors, but, except for the simplest cases, it usually leads to complex optimization problems when formulated in terms of the GLRT. In some cases, the operating conditions justify some assumptions: it is typical to consider the low-SNR regime, whereby reasonable approximations improve mathematical tractability at some unknown though reduced performance loss. This paper addresses instead the exact solution of a multiantenna detection problem valid for the all-SNR regime: we consider the detection of a low-bandwidth signal by an uncalibrated receiver modeled with an unknown diagonal noise spatial correlation matrix $\mathbf{\Sigma}^2$. This signal model has been previously considered in [1] for spatial detection schemes and in radio-astronomy contexts [2] for array calibration. The field of Factor Analysis [3] has also addressed this type of problems, although the solution herein presented for the rank-one (single factor) model is, to the author's knowledge, hitherto unknown. Although this is not an exhaustive list, previous work on array processing for spatially uncorrelated noise may also be found in [5],[6],[7],[8],[9]. This paper considers the detection of a low-bandwidth signal where a flat-fading channel can be reasonably assumed. In the absence of interference, the spatial correlation matrix is expressed as an unknown diagonal (noise) with an unknown rank-one perturbation (signal): $\mathbf{R} = \mathbf{\Sigma}^2 + \sigma^2\boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{H}}$.

2. SIGNAL MODEL

We consider a passband signal impinging on an M-antenna receiver, with $y_i(t) = s_i(t) + w_i(t)$ the passband signal at carrier frequency f_c picked up by the i-th antenna, and $s_i(t)$ and $w_i(t)$ the signal of interest and Additive White Gaussian Noise (AWGN), respectively. After down-conversion and sampling at $f_s = 1/T_s$, the complex baseband signal samples $x_i(nT_s)$ at $t_n = nT_s$ are stored into vector $\mathbf{x}_n$, of components $[\mathbf{x}_n]_i = x_i(nT_s)$. As in [1], both signal and noise are considered spectrally white: the received signal vector is expressed as $\mathbf{x}_n = s_n \cdot \boldsymbol{\alpha} + \boldsymbol{\eta}_n$, with s_n a stationary Gaussian discrete white process, $\boldsymbol{\alpha}$ the signal steering vector and $\boldsymbol{\eta}_n$ a stationary discrete additive white Gaussian noise (AWGN) vector process with diagonal spatial correlation matrix $\mathbf{\Sigma}^2 = \mathbb{E}[\boldsymbol{\eta}_n\boldsymbol{\eta}_n^{\mathrm{H}}]$, with $(\cdot)^{\mathrm{H}}$ the conjugate transpose operation. The following detection problem is stated,

$$\mathcal{H}_0 \quad : \quad \mathbf{x}_n \overset{\mathrm{sp}}{\sim} \mathcal{N}(\mathbf{0}, \mathbf{R}_0 = \mathbf{\Sigma}_0^2) \tag{1}$$

$$\mathcal{H}_1 \quad : \quad \mathbf{x}_n \overset{\mathrm{sp}}{\sim} \mathcal{N}(\mathbf{0}, \mathbf{R}_1 = \mathbf{\Sigma}_1^2 + \sigma_s^2 \cdot \boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{H}}) \tag{2}$$

with $\overset{\mathrm{sp}}{\sim}$ denoting 'spatially distributed as' , $\mathcal{N}(\cdot,\cdot)$ the normal distribution with specified mean and covariance matrix, $\mathbf{\Sigma}_0^2$ and $\mathbf{\Sigma}_1^2$ the unknown noise correlation matrices under hypotheses $\mathcal{H}_0, \mathcal{H}_1$, respectively, $\sigma_s^2, \boldsymbol{\alpha}$ the unknown signal power and steering (channel) vector under $\mathcal{H}_1$, respectively. Note that there exists an ambiguous (irrelevant) scale factor in the product $\sigma_s^2 \cdot \boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{H}} = (\sigma_s^2/a^2)(\boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{H}} \cdot a^2)$: in the following we set $\sigma_s^2 = 1$. The data matrix $\mathbf{X} = [\mathbf{x}_1^{\mathrm{H}}; \cdots; \mathbf{x}_N^{\mathrm{H}}]$, with ; the vector stacking operation, incorporates the signal $\mathbf{x}_i$ received at the i-th antenna. Letting $\mathbf{R}_e = \mathbb{E}[\mathbf{x}_n\mathbf{x}_n^{\mathrm{H}}]$ denote the spatial correlation matrix, the probability density function (p.d.f.) is expressed for either hypothesis as,

$$p(\mathbf{X}, \mathbf{R}_e) \quad = \quad \frac{1}{(\pi \det \mathbf{R}_e)^N} \exp\left(-N\mathrm{tr}[\mathbf{R}_e^{-1}\widehat{\mathbf{R}}]\right) \tag{3}$$

with $\mathrm{tr}[\cdot]$ the matrix trace, N the number of samples per antenna and $\widehat{\mathbf{R}} = \mathbf{X}^{\mathrm{H}}\mathbf{X}/N$ the sample correlation matrix.

3. DERIVATION OF TEST

The GLRT under the hypotheses $\mathcal{H}_1, \mathcal{H}_0$ is defined in terms of the Maximum Likelihood (ML) estimates of the p.d.f. parameters of the respective hypothesis,

$$\Lambda_{\mathrm{GLRT}} \quad = \quad \log \frac{\sup_{\mathbf{\Sigma}_1^2, \boldsymbol{\alpha}} p(\mathbf{X}|\mathbf{\Sigma}_1^2, \boldsymbol{\alpha})}{\sup_{\mathbf{\Sigma}_0^2} p(\mathbf{X}|\mathbf{\Sigma}_0^2)} \tag{4}$$

$$\quad = \quad \Lambda_{\mathcal{H}_1}\left(\widehat{\mathbf{\Sigma}}_1^2, \widehat{\boldsymbol{\alpha}}\right) - \Lambda_{\mathcal{H}_0}\left(\widehat{\mathbf{\Sigma}}_0^2\right) \tag{5}$$

for $\widehat{\mathbf{\Sigma}}_1^2, \widehat{\boldsymbol{\alpha}}$ and $\widehat{\mathbf{\Sigma}}_0^2$ the corresponding ML estimates and $\Lambda = \log[p]$ the associated log-likelihood. The value of this and the other tests considered in this paper is compared with a threshold γ_{th} to decide either hypothesis. The procedure for solving the non-trivial optimization problem under $\mathcal{H}_1$ in (5) is based on establishing a sequence of transformations on the initial parameter space $\Theta_1 = \{\boldsymbol{\Sigma}_1^2, \boldsymbol{\alpha}\}$ so that partial optimizations can be successively carried out on the new parameters, as described in the two following subsections.

3.1 GLRT derivation: stage 1

The initial few lines of our derivation (this subsection) are based on the initial conventions of the approximate low-SNR algorithm developed in [1] and have been incorporated for completeness of the exposition. The following subsection (stage 2) contains the novel contribution of this paper. Under $\mathcal{H}_0$, the optimization of the loglikelihood $\Lambda_{\mathcal{H}_0}$ is not difficult and was shown to be [1],

$$\begin{aligned} \Lambda_{\mathcal{H}_0}\left(\widehat{\boldsymbol{\Sigma}}_0^2\right) &= \log p(\mathbf{X}|\widehat{\boldsymbol{\Sigma}}_0^2 = \widehat{\mathbf{D}}) && (6) \\ &= -N(\log \pi + M + \log \det \widehat{\mathbf{D}}) && (7) \end{aligned}$$

with $\widehat{\mathbf{D}} = \operatorname{diag}[\widehat{\mathbf{R}}]$. Under $\mathcal{H}_1$ and using Sylvester's determinant property [1], we have for $\mathbf{R}_1$ in (2) that: $\det \mathbf{R}_1 = (1 + \boldsymbol{\alpha}^{\mathrm{H}}\boldsymbol{\Sigma}_1^{-2}\boldsymbol{\alpha}) \det \boldsymbol{\Sigma}_1^2$. From now on, we set $\boldsymbol{\Sigma}^2 = \boldsymbol{\Sigma}_1^2$ and $\rho = \boldsymbol{\alpha}^{\mathrm{H}}\boldsymbol{\Sigma}^{-2}\boldsymbol{\alpha}$. Using the Matrix Inversion Lemma for $\mathbf{R}_1^{-1}$, the log-likelihood becomes,

$$\begin{aligned} &\Lambda_{\mathcal{H}_1}(\boldsymbol{\Sigma}^2, \boldsymbol{\alpha}) = \log p(\mathbf{X}|\boldsymbol{\Sigma}^2, \boldsymbol{\alpha}) = && (8) \\ &= -\log \pi^N - N \log \det \boldsymbol{\Sigma}^2 - N \log(1+\rho) - \\ &- N \operatorname{tr}\left[\left(\boldsymbol{\Sigma}^{-2} - \frac{\boldsymbol{\Sigma}^{-2}\boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{H}}\boldsymbol{\Sigma}^{-2}}{1+\rho}\right)\widehat{\mathbf{R}}\right] \end{aligned}$$

From [1], $\mathbf{g} = \boldsymbol{\Sigma}^{-2}\boldsymbol{\alpha}/\sqrt{\rho}$, constrained by construction to $\mathbf{g}^{\mathrm{H}}\boldsymbol{\Sigma}^2\mathbf{g} = 1$. We define the cost function $J_1(\boldsymbol{\Sigma}, \mathbf{g}, \rho) = -\Lambda_{\mathcal{H}_1}(\boldsymbol{\Sigma}^2, \boldsymbol{\alpha})/N - \log \pi$, which has to be minimized in terms of the specified parameters,

$$\begin{aligned} J_1(\boldsymbol{\Sigma}, \mathbf{g}, \rho) &= \log \det \boldsymbol{\Sigma}^2 + \operatorname{tr}[\boldsymbol{\Sigma}^{-2}\widehat{\mathbf{R}}] + \log(1+\rho) \\ &- \frac{\rho}{1+\rho}\mathbf{g}^{\mathrm{H}}\widehat{\mathbf{R}}\mathbf{g} && (9) \\ &\text{s.t.} \quad \mathbf{g}^{\mathrm{H}}\boldsymbol{\Sigma}^2\mathbf{g} = 1 && (10) \end{aligned}$$

Henceforth, the derivation differs from [1].

3.2 GLRT derivation: stage 2

We introduce the unitary vector $\mathbf{e}_g = \mathbf{g}/||\mathbf{g}||_2$ and denote $\gamma = ||\mathbf{g}||_2^2$, with $||\cdot||_2$ the Euclidean norm. Then, $\mathbf{g}^{\mathrm{H}}\boldsymbol{\Sigma}^2\mathbf{g} = \gamma \cdot \mathbf{e}_g^{\mathrm{H}}\boldsymbol{\Sigma}^2\mathbf{e}_g = 1$, so that $\gamma = (\mathbf{e}_g^{\mathrm{H}}\boldsymbol{\Sigma}^2\mathbf{e}_g)^{-1}$. The cost function is now expressed as $J_2(\boldsymbol{\Sigma}, \mathbf{e}_g, \rho) = J_1(\boldsymbol{\Sigma}, \mathbf{g}, \rho)$, with the new parameters in the transformed set $\Theta_2 = \{\boldsymbol{\Sigma}, \mathbf{e}_g, \rho\}$ now mutually independent (although constrained: $\boldsymbol{\Sigma} \geq \mathbf{0}, ||\mathbf{e}_g||_2 = 1, \rho \geq 0$). Minimizing $J_2(\boldsymbol{\Sigma}, \mathbf{e}_g, \rho)$ with respect to ρ, we set $\nabla_{\rho=\widehat{\rho}} J_2(\boldsymbol{\Sigma}, \mathbf{e}_g, \rho) = 0$ and,

$$\frac{1}{1+\widehat{\rho}} - \frac{\mathbf{g}^{\mathrm{H}}\widehat{\mathbf{R}}\mathbf{g}}{(1+\widehat{\rho})^2} = 0 \qquad (11)$$

$$\Rightarrow \quad 1 + \widehat{\rho} = \gamma \cdot \mathbf{e}_g^{\mathrm{H}}\widehat{\mathbf{R}}\mathbf{e}_g = \frac{\mathbf{e}_g^{\mathrm{H}}\widehat{\mathbf{R}}\mathbf{e}_g}{\mathbf{e}_g^{\mathrm{H}}\boldsymbol{\Sigma}^2\mathbf{e}_g} \qquad (12)$$

Hence,

$$\begin{aligned} J_2(\boldsymbol{\Sigma}, \mathbf{e}_g, \widehat{\rho}) &= 1 + \log \det \boldsymbol{\Sigma}^2 + \operatorname{tr}[\boldsymbol{\Sigma}^{-2}\widehat{\mathbf{R}}] + && (13) \\ &+ \log \frac{\mathbf{e}_g^{\mathrm{H}}\widehat{\mathbf{R}}\mathbf{e}_g}{\mathbf{e}_g^{\mathrm{H}}\boldsymbol{\Sigma}^2\mathbf{e}_g} - \frac{\mathbf{e}_g^{\mathrm{H}}\widehat{\mathbf{R}}\mathbf{e}_g}{\mathbf{e}_g^{\mathrm{H}}\boldsymbol{\Sigma}^2\mathbf{e}_g} && (14) \end{aligned}$$

We define the new diagonal matrix $\boldsymbol{\Gamma}$, with $\widehat{\mathbf{D}} = \boldsymbol{\Sigma}^2 + \boldsymbol{\Gamma}$ (it will be shown later that $\boldsymbol{\Gamma} \geq \mathbf{0}$). Therefore,

$$\begin{aligned} \mathbf{e}_g^{\mathrm{H}}\boldsymbol{\Sigma}^2\mathbf{e}_g &= \mathbf{e}_g^{\mathrm{H}}\widehat{\mathbf{D}}\mathbf{e}_g - \mathbf{e}_g^{\mathrm{H}}\boldsymbol{\Gamma}\mathbf{e}_g && (15) \\ &= \mathbf{e}_g^{\mathrm{H}}\widehat{\mathbf{D}}\mathbf{e}_g \cdot \left(1 - \frac{\mathbf{e}_g^{\mathrm{H}}\boldsymbol{\Gamma}\mathbf{e}_g}{\mathbf{e}_g^{\mathrm{H}}\widehat{\mathbf{D}}\mathbf{e}_g}\right) && (16) \end{aligned}$$

We define the unitary vector $\mathbf{e}_d = \widehat{\mathbf{D}}^{1/2}\mathbf{e}_g/||\widehat{\mathbf{D}}^{1/2}\mathbf{e}_g||_2$ and the coherence matrix $\mathbf{C} = \widehat{\mathbf{D}}^{-1/2}\widehat{\mathbf{R}}\widehat{\mathbf{D}}^{-1/2}$. Hence, incorporating $\mathbf{e}_d$ into (16) and combining it with the ratio of quadratic forms in (14), yields,

$$\begin{aligned} \frac{\mathbf{e}_g^{\mathrm{H}}\widehat{\mathbf{R}}\mathbf{e}_g}{\mathbf{e}_g^{\mathrm{H}}\boldsymbol{\Sigma}^2\mathbf{e}_g} &= \frac{\mathbf{e}_d^{\mathrm{H}}\mathbf{C}\mathbf{e}_d}{\mathbf{e}_d^{\mathrm{H}}\mathbf{e}_d} \cdot \left(1 - \frac{\mathbf{e}_d^{\mathrm{H}}\boldsymbol{\Gamma}\widehat{\mathbf{D}}^{-1}\mathbf{e}_d}{\mathbf{e}_d^{\mathrm{H}}\mathbf{e}_d}\right)^{-1} && (17) \\ &= \mathbf{e}_d^{\mathrm{H}}\mathbf{C}\mathbf{e}_d \cdot (1 - \mathbf{e}_d^{\mathrm{H}}\mathbf{Q}\mathbf{e}_d)^{-1} && (18) \end{aligned}$$

as $\mathbf{e}_d^{\mathrm{H}}\mathbf{e}_d = 1$ and where we have defined the new diagonal matrix $\mathbf{Q} = \boldsymbol{\Gamma}\widehat{\mathbf{D}}^{-1} = \mathbf{I} - \boldsymbol{\Sigma}^2\widehat{\mathbf{D}}^{-1}$. Now, the full matrix $\mathbf{C}$ is known but the diagonal matrix $\mathbf{Q}$ is unknown as it depends on $\boldsymbol{\Sigma}^2$. Therefore, we can perform an additional parameter transformation based on $\Theta_3 = \{\mathbf{Q}, \mathbf{e}_d, \widehat{\rho}\}$ instead of on Θ_2 and operate with the modified cost function $J_3(\mathbf{Q}, \mathbf{e}_d, \widehat{\rho}) = J_2(\boldsymbol{\Sigma}, \mathbf{e}_g, \widehat{\rho})$. As $\operatorname{tr}[\boldsymbol{\Sigma}^{-2}\widehat{\mathbf{R}}] = \operatorname{tr}[\boldsymbol{\Sigma}^{-2}\widehat{\mathbf{D}}]$, we may write,

$$\begin{aligned} J_3(\mathbf{Q}, \mathbf{e}_d, \widehat{\rho}) &= 1 + \log \det \widehat{\mathbf{D}} + \log \det(\mathbf{I} - \mathbf{Q}) + \\ &+ \operatorname{tr}[(\mathbf{I} - \mathbf{Q})^{-1}] + && (19) \\ &+ \log \frac{\mathbf{e}_d^{\mathrm{H}}\mathbf{C}\mathbf{e}_d}{1 - \mathbf{e}_d^{\mathrm{H}}\mathbf{Q}\mathbf{e}_d} - \frac{\mathbf{e}_d^{\mathrm{H}}\mathbf{C}\mathbf{e}_d}{1 - \mathbf{e}_d^{\mathrm{H}}\mathbf{Q}\mathbf{e}_d} \end{aligned}$$

We now minimize the cost function $J_3(\mathbf{Q}, \mathbf{e}_d, \widehat{\rho})$ with respect to $\mathbf{Q}$: we compute the derivatives $\nabla_{q_k = \widehat{q}_k} J_3(\mathbf{Q}, \mathbf{e}_d, \widehat{\rho}) = 0$, with $\widehat{\mathbf{Q}} = \operatorname{diag}[\widehat{q}_1, \cdots, \widehat{q}_M]$,

$$\begin{aligned} &\nabla_{q_k = \widehat{q}_k} J_3(\mathbf{Q}, \mathbf{e}_d, \widehat{\rho}) && (20) \\ &= -\frac{1}{1 - \widehat{q}_k} + \frac{1}{(1 - \widehat{q}_k)^2} + \\ &+ \frac{|e_{d,k}|^2}{1 - \mathbf{e}_d^{\mathrm{H}}\widehat{\mathbf{Q}}\mathbf{e}_d} - \frac{|e_{d,k}|^2 \cdot \mathbf{e}_d^{\mathrm{H}}\mathbf{C}\mathbf{e}_d}{(1 - \mathbf{e}_d^{\mathrm{H}}\widehat{\mathbf{Q}}\mathbf{e}_d)^2} = 0 \end{aligned}$$

with $e_{d,k}$ the k-th component of $\mathbf{e}_d$. Defining $P_s = P_s(\widehat{\mathbf{Q}}, \mathbf{e}_d) = \mathbf{e}_d^{\mathrm{H}}\mathbf{C}\mathbf{e}_d/(1 - \mathbf{e}_d^{\mathrm{H}}\widehat{\mathbf{Q}}\mathbf{e}_d)$, we get,

$$\begin{aligned} \frac{1}{(1 - \widehat{q}_k)^2} - \frac{1}{1 - \widehat{q}_k} &= \frac{\widehat{q}_k}{(1 - \widehat{q}_k)^2} && (21) \\ &= \frac{|e_{d,k}|^2}{\mathbf{e}_d^{\mathrm{H}}\mathbf{C}\mathbf{e}_d}(P_s^2 - P_s) && (22) \end{aligned}$$

From (22), either all $\widehat{q}_k$ are positive and $P_s \geq 1$ or all $\widehat{q}_k$ are negative and $P_s \leq 1$. Note also that $P_s \geq 0$ by construction as a quotient of positive quadratic forms: see (18), where $P_s = \mathbf{e}_g^{\mathrm{H}}\widehat{\mathbf{R}}\mathbf{e}_g/\mathbf{e}_g^{\mathrm{H}}\boldsymbol{\Sigma}^2\mathbf{e}_g$. Hence, $1 \geq \mathbf{e}_d^{\mathrm{H}}\widehat{\mathbf{Q}}\mathbf{e}_d$, which requires $\widehat{\mathbf{D}} \geq \boldsymbol{\Gamma}$, and necessarily $0 \leq \widehat{q}_k \leq 1, P_s \geq 1$. From (22), adding over k and using the unitary character of $\mathbf{e}_d$, we can define $\widehat{\xi} = \xi(\widehat{\mathbf{Q}})$ (a compression of $\widehat{\mathbf{Q}}$) as,

$$\begin{aligned} \widehat{\xi} &= \sum_{k=1}^{M}\left(\frac{1}{(1 - \widehat{q}_k)^2} - \frac{1}{1 - \widehat{q}_k}\right) && (23) \\ &= \frac{P_s^2 - P_s}{\mathbf{e}_d^{\mathrm{H}}\mathbf{C}\mathbf{e}_d} \geq 0 && (24) \end{aligned}$$

Additionally, multiplying (22) by $1-\widehat{q}_k$ and adding over k,

$$\sum_{k=1}^{M}\left(\frac{1}{1-\widehat{q}_k}-1\right) \tag{25}$$

$$= \frac{P_s^2-P_s}{\mathbf{e}_d^{\mathrm{H}}\mathbf{C}\mathbf{e}_d}\cdot\sum_{k=1}^{M}(1-\widehat{q}_k)|e_{d,k}|^2 \tag{26}$$

$$= \frac{P_s^2-P_s}{\mathbf{e}_d^{\mathrm{H}}\mathbf{C}\mathbf{e}_d}\cdot(1-\mathbf{e}_d^{\mathrm{H}}\widehat{\mathbf{Q}}\mathbf{e}_d) \quad = \quad \frac{P_s^2-P_s}{P_s} = P_s - 1$$

Combining (7) and (19) the normalized log-GLRT $\Lambda^1_{\mathrm{GLRT}}$ can be defined in terms of J_3 as,

$$\Lambda^1_{\mathrm{GLRT}} = \frac{\Lambda_{\mathcal{H}_1}-\Lambda_{\mathcal{H}_0}}{N} \tag{27}$$

$$= M + \log\det\widehat{\mathbf{D}} - J_3(\widehat{\mathbf{Q}},\mathbf{e}_d,\widehat{\rho}) \tag{28}$$

which yields,

$$\Lambda^1_{\mathrm{GLRT}} = (M-1)+\log\det[(\mathbf{I}-\widehat{\mathbf{Q}})^{-1}] - \tag{29}$$

$$- \mathrm{tr}[(\mathbf{I}-\widehat{\mathbf{Q}})^{-1}] + P_s - \log P_s \tag{30}$$

We note that $\widehat{\mathbf{Q}}$ depends on $\mathbf{e}_d$: $\widehat{\mathbf{Q}} = \widehat{\mathbf{Q}}(\mathbf{e}_d)$ in terms of the system of nonlinear equations in (22). Additionally, we will be able optimize with respect to the compressed parameter $\widehat{\xi} \geq 0$ in (24), so that the constraint $\widehat{\xi} = \xi(\widehat{\mathbf{Q}})$ in (24) is enforced. From (22), we can solve for $(1-\widehat{q}_k)^{-1}$ and keep the only valid solution to the second-degree equation consistent with $0 \leq \widehat{q}_k \leq 1$ and $\widehat{\xi} \geq 0$. Using the expression for $\widehat{\xi}$ in (24), we get,

$$\frac{1}{1-\widehat{q}_k} = \frac{1}{2}\left(1+\sqrt{1+4\widehat{\xi}\cdot|e_{d,k}|^2}\right) \tag{31}$$

We can also solve for P_s in the second-degree equation in (24). The only solution that guarantees $P_s \geq 1$ is,

$$P_s = \frac{1}{2}\left(1+\sqrt{1+4\widehat{\xi}\cdot\mathbf{e}_d^{\mathrm{H}}\mathbf{C}\mathbf{e}_d}\right) \tag{32}$$

where $(1-\widehat{q}_k)^{-1}$ in (31) and P_s in (32) are expressed in terms of the same function. These two expressions can be readily substituted into $\Lambda^1_{\mathrm{GLRT}}$ in (30) so that the compressed log-GLRT is expressed in terms of parameters $\widehat{\xi} \geq 0$ and unitary $\mathbf{e}_d$. For simplification, we note that the following monotone increasing non-convex function $\mathrm{g}(\tau)$ can be defined,

$$\mathrm{g}(\tau) = -\log\frac{1+\sqrt{1+\tau}}{2} + \frac{1+\sqrt{1+\tau}}{2} \tag{33}$$

with $\mathrm{g}(\tau) \geq 1$. Therefore,

$$\Lambda^1_{\mathrm{GLRT}}(\widehat{\xi},\mathbf{e}_d) \tag{34}$$

$$= (M-1)+\mathrm{g}(4\widehat{\xi}\cdot\mathbf{e}_d^{\mathrm{H}}\mathbf{C}\mathbf{e}_d) - \sum_{k=1}^{M}\mathrm{g}(4\widehat{\xi}\cdot|e_{d,k}|^2)$$

If we define the unconstrained vector $\mathbf{v}_d = (4\widehat{\xi})^{1/2}\cdot\mathbf{e}_d$, with $v_{d,k}$ its k-th component, the log-GLRT can be expressed alternatively as,

$$\Lambda^1_{\mathrm{GLRT}}(\mathbf{v}_d) \tag{35}$$

$$= (M-1)+\mathrm{g}(\mathbf{v}_d^{\mathrm{H}}\mathbf{C}\mathbf{v}_d) - \sum_{k=1}^{M}\mathrm{g}(|v_{d,k}|^2)$$

and the following test can be finally defined,

$$\mathrm{T}_1 = \max_{\mathbf{v}_d\in\mathbb{C}^M}\left[(M-1)+\mathrm{g}(\mathbf{v}_d^{\mathrm{H}}\mathbf{C}\mathbf{v}_d) - \sum_{k=1}^{M}\mathrm{g}(|v_{d,k}|^2)\right] \tag{36}$$

Hence, classical unconstrained optimization methods can be applied for obtaining the final value of the test. An initial estimate must be provided for the first iteration. The optimization algorithm is described in section 4. We note that from (24) and (22), $\widehat{\xi}$ can be computed from $\mathbf{v}_d$ as follows, which is consistent with the definition of $\mathbf{v}_d$: $\mathbf{v}_d = (4\widehat{\xi})^{1/2}\cdot\mathbf{e}_d$,

$$\widehat{\xi} = \sum_{k=1}^{M}\frac{q_k}{(1-q_k)^2} = \frac{1}{4}\mathbf{v}_d^{\mathrm{H}}\mathbf{v}_d \tag{37}$$

3.3 Other detectors

This section briefly describes two detectors that will be used for comnparison in the simulations section. We note that reference [1] considered the following constrained optimization algorithm as a low-SNR approximation to the GLRT,

$$\mathrm{T}_2 = \max_{\mathbf{v}'_d:||\mathbf{v}'_d||_2=1}\left[\mathbf{v}'^{\mathrm{H}}_d\mathbf{C}\mathbf{v}'_d\right] \tag{38}$$

$$= \lambda_{\max}[\mathbf{C}] \tag{39}$$

with a stationary point solution determined by the typical eigenvalue/eigenvector relationship,

$$\mathbf{C}\mathbf{v}'_d = \lambda\mathbf{v}'_d \quad , \quad \lambda = \frac{\mathbf{v}'^{\mathrm{H}}_d\mathbf{C}\mathbf{v}'_d}{\mathbf{v}'^{\mathrm{H}}_d\mathbf{v}'_d} \tag{40}$$

In comparing with (36), we observe the following facts,

1. equation (38) is constrained by $||\mathbf{v}'_d||_2 = 1$. Hence, it is not sensitive to scale (Euclidean norm of $\mathbf{v}'_d$). On the contrary, equation (36) is sensitive to scale. This is due to the fact that (38) was derived under low-SNR assumptions while (36) remains valid over the whole SNR range. This scale sensitiveness is thus directly related to SNR, as shown in section (5).
2. as $\mathrm{g}(\tau)$ is a monotone increasing function, equation (38) maximizes in fact the term $\mathrm{g}(\mathbf{v}'^{\mathrm{H}}_d\mathbf{C}\mathbf{v}_d)$ on the constraint $||\mathbf{v}'_d||_2 = 1$. The final term $\sum_{k=1}^{M}\mathrm{g}(|v_{d,k}|^2)$ has been thus disregarded.

Yet another point of similarity will be established in section 7 in terms of the stationary point solutions of T_1 and T_2.

A second detector was analyzed in [4], which operates on the Euclidean norm of the off-diagonal components of $\mathbf{C}$. As the diagonal components of $\mathbf{C}$ are unity, it can also be expressed in terms of the Frobenius norm of $\mathbf{C}$,

$$\mathrm{T}_3 = ||\mathbf{C}||_F \tag{41}$$

Although this detector was derived for a more general signal model, it has also been included in the comparative analysis.

4. OPTIMIZATION ALGORITHM

The detector T_1 is, potentially, a multimodal function in terms of parameters $\widehat{\xi},\mathbf{e}_d$ or of $\mathbf{v}_d$. Therefore, global optimization rests on the ability to provide an initial guess of $\mathbf{v}_d$ as close as possible to the global maximum. Following [1], we choose the maximum eigenvector of $\mathbf{C}$ for the first iteration. This is consistent with the presence of the first term $\mathrm{g}(4\widehat{\xi}\cdot\mathbf{e}_d^{\mathrm{H}}\mathbf{C}\mathbf{e}_d)$ in (34), as $\mathrm{g}(\cdot)$ is monotone increasing. Instead, the second term $\sum_{k=1}^{M}\mathrm{g}(4\widehat{\xi}\cdot|e_{d,k}|^2)$ is sensitive to the squared components $|e_{d,k}|^2$ being as uniform as possible as can be seen considering its minimization under the constraint $\mathbf{e}_d^{\mathrm{H}}\mathbf{e}_d = 1$. The following algorithm is proposed,

1. Compute $\mathbf{e}_d^0 = \mathbf{e}_{\max}[\mathbf{C}]$, with $\mathbf{e}_d^{0,\mathrm{H}}\mathbf{C}\mathbf{e}_d^0 = \lambda_{\max}[\mathbf{C}]$.
2. Solve the single-variate optimization problem: $\widehat{\xi}^0 = \operatorname{argmax}_{\widehat{\xi}\geq 0}\left[\Lambda^1_{\mathrm{GLRT}}(\widehat{\xi},\mathbf{e}_d^0)\right]$.
3. Set $\mathbf{v}_d^0 = (4\widehat{\xi}^0)^{1/2}\cdot\mathbf{e}_d^0$.
4. Apply some iterative unconstrained optimization scheme to $\Lambda^1_{\mathrm{GLRT}}(\mathbf{v}_d)$ in (35) using the initial guess $\mathbf{v}_d^0$.

5. SNR ESTIMATION

We establish the relationship between the GLRT and the SNR estimates at each antenna, where the per-antenna SNR is defined from the model of the signal and noise power as $\mathrm{snr}_k = [\mathbf{\Gamma}\mathbf{\Sigma}^{-2}]_{k,k}$. From the previous definition of $\mathbf{Q}$: $\mathbf{Q} = \mathbf{\Gamma}\widehat{\mathbf{D}}^{-1} = \mathbf{\Gamma}(\mathbf{\Gamma}+\mathbf{\Sigma}^2)^{-1}$, we get $\widehat{\mathrm{snr}}_k = [\widehat{\mathbf{Q}}(\mathbf{I}-\widehat{\mathbf{Q}})^{-1}]_{k,k} = \widehat{q}_k/(1-\widehat{q}_k)$. Now, from (25), we may establish for P_s that,

$$P_s = 1+\sum_{k=1}^{M}\frac{\widehat{q}_k}{1-\widehat{q}_k} = 1+\sum_{k=1}^{M}\widehat{\mathrm{snr}}_k \qquad (42)$$

Substituting the identity (42) into the log-GLRT expression in (30) and expressing $\widehat{q}_k$ as $\widehat{q}_k = \widehat{\mathrm{snr}}_k/(1+\widehat{\mathrm{snr}}_k)$, it is not difficult to show that the log-GLRT can be expressed in terms of the optimum per-antenna SNR estimates as,

$$\begin{aligned}\Lambda^1_{\mathrm{GLRT}} &= (M-1)-\sum_{k=1}^{M}\log(1-\widehat{q}_k)-\sum_{k=1}^{M}\frac{1}{1-\widehat{q}_k}+\\ &+ P_s-\log P_s \qquad (43)\\ &= \log\frac{\prod_{k=1}^{M}(1+\widehat{\mathrm{snr}}_k)}{1+\sum_{k=1}^{M}\widehat{\mathrm{snr}}_k} \geq 0 \qquad (44)\end{aligned}$$

It should be remarked that this expression does not constitute an equivalent criterion, but is simply a relationship between the value of the log-GLRT resulting from the optimization procedure and the corresponding SNR estimates at the same maximizing values of the parameter set $\Theta = \{\widehat{\xi},\mathbf{e}_d\}$. Finally, from (24), we may also express $\widehat{\xi}$ in terms of the per-antenna SNR estimates as,

$$\widehat{\xi} = \sum_{k=1}^{M}\widehat{\mathrm{snr}}_k\cdot(1+\widehat{\mathrm{snr}}_k) \qquad (45)$$

6. LOW-SNR ANALYSIS OF T_1

From (45), we take $\widehat{\xi}\to 0^+$, so that the low-SNR analysis reduces to examining the first non-zero term of the Taylor series of (36) about $\widehat{\xi}=0$, where $\mathbf{v}_d = (4\widehat{\xi})^{1/2}\mathbf{e}_d$. The Taylor series of $\mathrm{g}(\tau)$ yields $\mathrm{g}(\tau) = 1+\tau^2/16+o(\tau^2)$ and, retaining the second order term, the following detector results,

$$\mathrm{T}_4 = \max_{\mathbf{e}_d:||\mathbf{e}_d||_2=1}\left[(\mathbf{e}_d^{\mathrm{H}}\mathbf{C}\mathbf{e}_d)^2-\sum_{k=1}^{M}|e_{d,k}|^4\right] \qquad (46)$$

where the scale factor $4\widehat{\xi}$ becomes an irrelevant common factor and can be dispensed with for the optimization. We note the difference with T_2 in (38): the second term in T_4 does not appear. This is due to the fact that, although T_2 was derived considering an intuitive low-SNR approximation, an explicit Taylor series expansion was not available at the moment. Additionally, we note that this low-SNR detector should also be implemented iteratively as an explicit expression cannot be derived for the optimization problem in (46).

7. STATIONARY POINT ANALYSIS

We examine the equations for the stationary point of detector T_1 as opposed to detector T_2. The following Lagrangian is constructed, where, the comparison requires that the unitary vector $\mathbf{e}_d$ (included as a norm constraint to the Lagrangian) and the scale parameter ξ be treated separately,

$$\mathcal{L} = -\mathrm{g}(4\xi\mathbf{e}_d^{\mathrm{H}}\mathbf{C}\mathbf{e}_d)+\sum_k \mathrm{g}(4\xi|e_{d,k}|^2)-\mu_1(\mathbf{e}_d^{\mathrm{H}}\mathbf{e}_d-1) \qquad (47)$$

The stationary points are obtained from $\nabla_{\mathbf{e}_d^{\mathrm{H}}}\mathcal{L} = \mathbf{0}$,

$$\mathbf{0} = -\mathrm{g}'(4\xi\mathbf{e}_d^{\mathrm{H}}\mathbf{C}\mathbf{e}_d)4\xi\mathbf{C}\mathbf{e}_d+\mathbf{G}'4\xi\mathbf{e}_d-\mu_1\mathbf{e}_d \qquad (48)$$

with $\mathrm{g}'(\tau) = (\mathrm{d}/\mathrm{d}\tau)\mathrm{g}(\tau)$ and $\mathbf{G}'$ a diagonal matrix of components $[\mathbf{G}']_{k,k} = \mathrm{g}'(4\xi|e_{d,k}|^2)$. Pre-multiplying by $\mathbf{e}_d^{\mathrm{H}}$,

$$\mu_1 = 4\xi(-\mathrm{g}'(4\xi\mathbf{e}_d^{\mathrm{H}}\mathbf{C}\mathbf{e}_d)\mathbf{e}_d^{\mathrm{H}}\mathbf{C}\mathbf{e}_d+\mathbf{e}_d^{\mathrm{H}}\mathbf{G}'\mathbf{e}_d) \qquad (49)$$

Substituting μ_1 into (48),

$$\mathbf{C}\mathbf{e}_d = (\mathbf{e}_d^{\mathrm{H}}\mathbf{C}\mathbf{e}_d)\mathbf{e}_d+\frac{1}{\mathrm{g}'_\xi}(\mathbf{I}-\mathbf{e}_d\mathbf{e}_d^{\mathrm{H}})\mathbf{G}'\mathbf{e}_d \qquad (50)$$

where $\mathrm{g}'_\xi = \mathrm{g}'(4\xi\mathbf{e}_d^{\mathrm{H}}\mathbf{C}\mathbf{e}_d)$ has been defined. Comparing with T_2 (38), which fulfils the stationary point equation $\mathbf{C}\mathbf{e}_d = \lambda\mathbf{e}_d = (\mathbf{e}_d^{\mathrm{H}}\mathbf{C}\mathbf{e}_d)\mathbf{e}_d$, a second scale-dependent (ξ) correction term has appeared. Now, differentiating with respect to ξ,

$$\begin{aligned}-\mathrm{g}'_\xi\cdot 4\mathbf{e}_d^{\mathrm{H}}\mathbf{C}\mathbf{e}_d+\sum_k \mathrm{g}'(4\xi|e_{d,k}|^2)4|e_{d,k}|^2 &= 0 \qquad (51)\\ -\mathrm{g}'_\xi\cdot 4\mathbf{e}_d^{\mathrm{H}}\mathbf{C}\mathbf{e}_d+\mathbf{e}_d^{\mathrm{H}}\mathbf{G}'\mathbf{e}_d &= 0 \qquad (52)\end{aligned}$$

which provides an additional relationship to the non-linear system associated with the stationary point. Hence, we get $1/\mathrm{g}'_\xi = \mathbf{e}_d^{\mathrm{H}}\mathbf{C}\mathbf{e}_d/\mathbf{e}_d^{\mathrm{H}}\mathbf{G}'\mathbf{e}_d$ and, substituting into (50), we get the following expression in terms of generalized eigenvalues,

$$\mathbf{C}\mathbf{e}_d = \frac{\mathbf{e}_d^{\mathrm{H}}\mathbf{C}\mathbf{e}_d}{\mathbf{e}_d^{\mathrm{H}}\mathbf{G}'\mathbf{e}_d}\cdot\mathbf{G}'\mathbf{e}_d \qquad (53)$$

$$\mathbf{e}_d = \mathbf{e}[\mathbf{C},\mathbf{G}'] \quad , \quad \frac{\mathbf{e}_d^{\mathrm{H}}\mathbf{C}\mathbf{e}_d}{\mathbf{e}_d^{\mathrm{H}}\mathbf{G}'\mathbf{e}_d} = \lambda[\mathbf{C},\mathbf{G}'] \qquad (54)$$

This is a nonlinearly modified eigenvector equation, as $\mathbf{G}' = \mathbf{G}'(\mathbf{e}_d)$. The stationary point is found instead by direct optimization of T_1 in (36) via a gradient descent algorithm.

8. SIMULATIONS

Simulations have been carried out for the signal model in (2). All figures display, in logarithmic axes, the probability of missed detection (vertical axis) versus the probability of false alarm (horizontal axis). The following four detectors have been considered:

1. Detector T_1 (this paper) in equation (36), exact all-SNR algorithm for the rank-1 signal model.
2. Detector T_2 in equation (38), approximate low-SNR algorithm derived in [1] for the rank-1 signal model.
3. Detector T_3 in equation (41), approximate low-SNR algorithm derived in [4] for the rank-r signal model, with r unknown.
4. Detector T_4 in equation (46), asymptotic low-SNR algorithm for the rank-1 signal model.

The convergence rate of the iterative detectors T_1 and T_4 depends on the SNR. We have verified that, when optimization is implemented with a gradient descent algorithm, slower

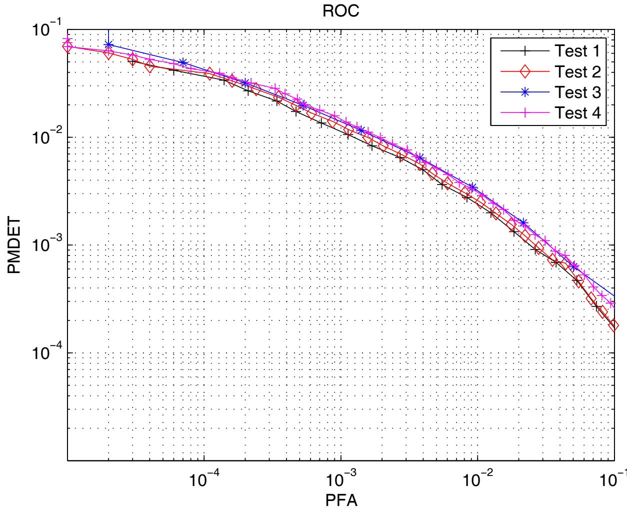

Figure 1: ROC: $M = 4$; $N = 60$ samples/antenna. Per-antenna SNR's (dB): $[-1; -3; -8; -9]$. Relative per-antenna noise power (dB): $[-1.5; -2; -1; 0]$.

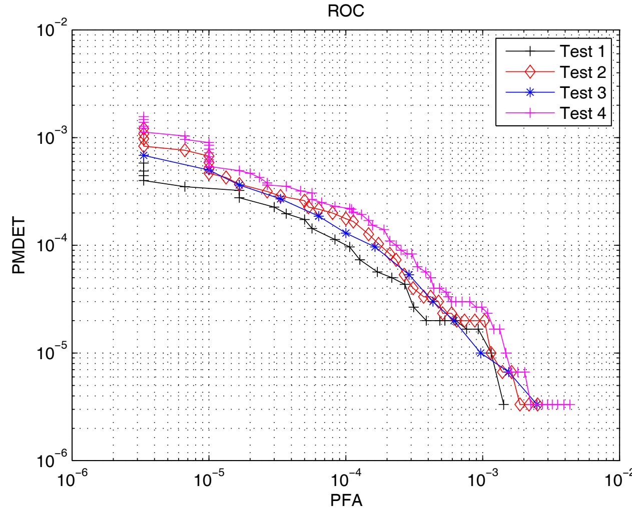

Figure 2: ROC: $M = 4$; $N = 60$ samples/antenna. Per-antenna SNR's (dB): $[1; -1; -6; -7]$. Relative per-antenna noise power (dB): $[-1.5; -2; -1; 0]$.

convergence occurs with increasingly higher SNR (the maximum number of iterations has been limited to $N_{\mathrm{it}} = 80$). Different useful signal powers are considered for each antenna as indicated in the corresponding figure caption in terms of their per-antenna SNR. The noise power at each antenna relative to that antenna with maximum noise power is also indicated. Figure 1 illustrates for a 4-antenna scenario that at low-SNR and when not many samples are available, the performance of all four algorithms is practically the same: detector T_1 and its corresponding low-SNR approximation (detector T_4) show very similar performance, as well as detectors T_2 and T_3. Figure 2 illustrates the same scenario under a uniform 2 dB increment in the per-sensor SNR, where a slight dominance of detector T_1 over the other detectors is observed. MonteCarlo runs of $1 \cdot 10^5$ (fig.1) and $3 \cdot 10^5$ (fig.2) iterations have been performed, so that probabilities below 10^{-4} may be subject to some statistical noise.

9. CONCLUSIONS

We have transformed the GLR optimization problem for the Gaussian rank-one signal model exposed in the introduction, which was initially expressed in terms of M unknown complex variables (steering vector) and M unknown real variables (per-antenna noise power profile), into a concentrated GLR problem. The new GLR problem (detector T_1), which, is *exact* over the complete range of SNR, constitutes an unconstrained M-variate nonlinear optimization problem in complex variables, thus reducing the initial dimensionality of the parameter space. The mathematical expression of detector T_1 has been compared with the closed-form (but approximate) algorithm in [1] (detector T_2). Experimental performance evaluation has validated the close-to-optimum performance of detectors T_2 and T_3 in [4] against the exact detector T_1 (and its low-SNR approximation T_4) for the tested SNR's. As far as has been possible to measure due to the low probabilities involved, small improvements in T_1's performance over the other detectors have been observed when the per-antenna SNR increases (although still below the 0-dB reference). In terms of complexity, detector T_1 does not need eigenvalue evaluation as detector T_2 but it is iterative and the number of iterations required to converge to a local maximum increases with SNR (a maximum of 80 iterations has been considered) and involves square-root and logarithm computations. Thus is the most complex of all four considered. In contrast, detector T_3 is the simplest as it only requires evaluation of a Frobenius norm. The theoretical objective of the paper has been the comparison between the variational expressions of T_1 in (36) and T_2 in (38).

REFERENCES

[1] Lopez-Valcarce, R.; Vazquez-Vilar, G.; Sala, J.; "Multi-antenna spectrum sensing for Cognitive Radio: overcoming noise uncertainty"; 2nd Intl. Workshop on Cognitive Information Processing (CIP 2010), Elba Island, Italy.

[2] Boonstra, A.J.; Van der Ween, A.J.; "Gain decomposition methods for radio-telescope arrays"; IEEE Workshop on Statistical Signal Processing.

[3] Mulaik, Stanley A.; "Foundations of Factor Analysis"; Chapman and Hall, 2010 (2nd edition).

[4] Leshem, A.; van der Veen, A.-J.; "Multichannel detection of Gaussian signals with uncalibrated receivers"; Signal Process. Letters, IEEE, Vol. 8, 4, 2001, 120 - 122.

[5] Pesavento, M.; Gershman, A.B.; "Array processing in the presence of unknown nonuniform sensor noise: a maximum likelihood direction finding algorithm and cramer-rao bounds"; Stat. Signal and Array Processing, 2000. Procs. of the 10th IEEE Workshop on, 2000, 78 - 82.

[6] Ramirez, D.; Vazquez V., G.; Lopez-V., R.; Via, J.; Santamaria, I.; "Detection of rank-P signals in cognitive radio networks with uncalibrated multiple antennas", Signal Processing, IEEE Trans. on, 2011 (accepted).

[7] Besson, O.; Kraut, S.; Scharf, L.L.; "Detection of an unknown rank-one component in white noise"; Signal Processing, IEEE Trans. on; 2006, Vol. 54, 7, 2835-2839.

[8] Taherpour, A.; Nasiri-Kenari, M.; Gazor, S.; "Multiple antenna spectrum sensing incognitive radios"; Wireless Communications, IEEE Trans. on; Vol. 9, 2, 814-823.

[9] Wang, P.; Fang, J.; Han, N.; Li, H.; "Multiantenna-assisted spectrum sensing for cognitive radio""; Vehicular Technology, IEEE Trans. on; Vol. 59, 4, 1791-1800.