RADIO TRANSIENT DETECTION IN RADIO ASTRONOMICAL ARRAYS

Avishay Antman and Amir Leshem
Faculty of Engineering, Bar-Ilan University, Israel

Abstract—Celestial transient radio sources have attracted considerable scientific interest recently, but their investigation is hampered by the fact that they cannot be effectively detected by commonly used radio astronomy signal processing techniques. One significant obstacle to observing radio transients is intermittent terrestrial radio frequency interference, which can appear as a transient signal. In this paper we present a generalized likelihood ratio test detector for near field sources, which is often terrestrial interferences, for which no prior knowledge about the steering vector is assumed. The proposed detector has the desirable property of a constant false alarm rate. Computer simulation results suggest that it outperforms the detector used by the Low Frequency Array (LOFAR) radio telescope, and that it can also be harnessed for the detection of far-field signals.

I. INTRODUCTION

A short, non-periodic celestial radio signal was detected for the first time in 2007 [2]. Since then, roughly 100 other such fast radio bursts (FRB) have been detected [3]. The study of FRBs and their origin is of considerable scientific interest because they are considered to be a product of exotic high energy astronomical events that are not yet understood [4]. It is likely that only a small fraction of FRBs incidents on Earth are detected, since current radio telescopes and imaging methods are not adapted for a large field of view, high sensitivity to transients, or full-time coverage. Enormous efforts are being made to develop new telescopes, processing hardware and algorithms that will provide better temporal, spatial and frequency coverage for the detection and accurate analysis of transient astronomical signals [5].

Like the detectors in [6] and [7], the detectors developed in this work are also applicable to cognitive radio, passive radar, passive acoustic sensors and any other multi-sensor receiving system that can benefit from detecting short, spatially localized Gaussian signals that are embedded in colored Gaussian noise.

The sensor (specifically, a radio telescope designed for the search of astronomical transients) is assumed to generate a sequence of sample covariance matrices (SCM), where each SCM corresponds to a distinct set of vector samples collected during a time frame of constant length [8] [9]. The SCM integration interval is short enough so that the statistical properties of all the received signals, except for transients, can be assumed to remain constant during two consecutive time frames, and the bandwidth is large enough so that a large number of vector samples can be collected during each interval (on the order of $10^4$). The signal of interest is expected to be a Gaussian point source that appears and disappears during one integration interval. We further assume that all the signals, both static and transient, are zero-mean Gaussian, and that the occurrence of transients is rare, so that there may be at most one transient point source in a single covariance matrix.

The model here makes the necessary assumptions regarding the signal: it is rank-1 and Gaussian, and the noise is Gaussian. The data-structure fed to the detector, formulated as a sequence of sample covariance matrices, corresponds to the data-structure streamed in the transient detection pipeline of "Low Frequency Array" (LOFAR). The resulting generalized likelihood ratio test (GLRT) is shown to possess several desirable properties, such as a constant false alarm rate (CFAR), sensitivity comparable to detectors that assume a structured signal, robustness to the number of samples and array calibration errors, and low computational complexity.

The work presented below focuses on detecting a spatially correlated Gaussian signal using a given pair of consecutive SCMs. The problem of detecting spatially correlated signals using one or more covariance matrices has been addressed in works on various statistical signal and noise models. In [10], [11], [6], [7] and [12] it is assumed that the signal is Gaussian and that the noise is spatially uncorrelated, or that its spatial covariance is very small compared to the covariance of the desired signal and that the signal is Gaussian with an unknown steering vector. In [13], [14], [15] and [16] GLRT detectors have been developed for the detection of unknown deterministic signals that are restricted to some given subspace, and embedded in spatially correlated Gaussian noise. In [13] the noise covariance matrix is assumed to be known up to some scaling factor, in [14] the signal is modeled as a linear combination of signals with known steering vectors and time signatures, in [15] the signal is a linear combination of a small set of parameterized functions of time and in [16] the availability of a secondary dataset that contains only background signal is assumed. Our approach is a two-dataset extension of these models to a Gaussian rank-1 signal with a steering vector that is not restricted to any predetermined structure. In recent papers [17], [18] detectors of Gaussian signals with known steering vector have been developed. Like [16] and this work, they use two sets of observations: a primary dataset in which the signal is sought and a secondary dataset containing only noise. Related approaches in which two or more datasets are used to detect a Gaussian signal can be found in [19], [20]. In these works the signal is assumed to be simultaneously collected by different channels, which is often the case in passive radar or passive sensing of communication.
signals.

All the detection methods above are based on some assumed knowledge of the signal or noise parameters or their structure. This dependence on prior knowledge makes them sensitive to model errors such as interference, multipath, array calibration errors or signals with unexpected structures. The detector proposed in this work does not rely on any assumptions related to the signal or the noise, except that they are Gaussian. Therefore it is particularly appropriate for detecting radio astronomical transients - a detection scenario prone to uncertainties relating to the receiving array, the background signals and noise, and the direction of arrival of the transient signal. Furthermore, as shown below, the transient detector is also a GLRT detector corresponding to a model of a deterministic unknown signal with unknown unstructured steering vector, embedded in correlative noise, and can be used in conjunction with a detector of signals with structured steering vector to enhance its computability.

The remainder of this paper is organized as follows: Section II details the notation. The problem of detecting Gaussian signals in correlated noise and its corresponding detector and estimators are presented in Section III, along with a review of related detectors derived in previous works. The mathematical derivation of the detector and the estimators is provided in Section IV. Extensions of existing detectors to models with unknown steering vectors, their computational complexity and the benefits of combining different types of detectors are discussed in Section V. Numerical simulations appear in Section VI.

II. NOTATION

In what follows, vectors are denoted by bold-face lower-case letters, matrices by bold-face upper-case letters and matrix pairs by calligraphic upper-case letters. $\mathbb{R}$ is the set of real numbers, and $\mathbb{R}_+$ and $\mathbb{R}_{++}$ are the sets of non-negative and positive numbers, respectively. $\mathbb{C}$ is the set of complex numbers, and $\mathbb{C}^q$ is the unit sphere of $\mathbb{C}^q$, restricted to vectors whose first non-zero element is real and positive (so that no two linearly dependent vectors are contained in $\mathbb{C}^q$). The notations $|A|$, $\text{tr}(A)$, $\text{vec}(A)$, $A^T$, $A^H$ and $A^{-H}$ denote the determinant, trace, vectorization, transpose, conjugate transpose and inverse conjugate transpose of a matrix $A$, respectively. The $(i,j)$th element of a matrix $A$ is denoted by $A_{ij}$ and the $i$th element of a vector $x$ is denoted by $x(i)$. The notation $A = \text{diag}(a_1,\ldots,a_q)$ indicates that $A_{ij} = a_{\delta_{ij}}$ where $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ if $i \neq j$. The q-dimensional identity matrix is denoted by $I_q$. The vector in $\mathbb{R}^q$ whose elements are all 1 is denoted by $1_q$, $\lambda_{\text{max}}(A)$ is the largest eigenvalue of $A$ with corresponding right eigenvector $v_{\text{max}}(A) \in \mathbb{C}^q$. We write $x \sim \mathcal{CN}_q(\mu, \Sigma)$ if $x$ is a q-dimensional complex normal vector with mean $\mu$ and covariance $\Sigma$. If $x_1,\ldots,x_n$ is a sequence of i.i.d. random vectors with $x_i \sim \mathcal{CN}_q(0, \Sigma)$ for $i = 1,\ldots,n$ and $R = \frac{1}{n}\sum_{i=1}^n x_i x_i^H$, then $R$ is said to be a q-dimensional random Wishart matrix with scaling matrix $\Sigma$ and $n$ degrees of freedom; this distribution is denoted by $\mathcal{W}_q(\Sigma, n)$. If $\Sigma = I_q$ then $R$ is said to be a white Wishart matrix. We define the following sets of $q \times q$ complex matrices: $\mathbb{S}^q$ is the set of Hermitian matrices, $\mathbb{S}_+^q$ is the set of positive definite (PD) matrices, $\mathbb{F}^q$ is the set of invertible matrices, $\mathbb{D}^q_\Sigma$ is the set of diagonal PD matrices, $\mathbb{U}^q_{++}$ is the set of unitary matrices, $\mathbb{P}^q$ is the set of permutation matrices, $\mathbb{U}^q_{\Sigma}$ is the set of unitary matrices that are also diagonal. The $i$th vector of the natural basis of $\mathbb{C}^q$ is denoted by $e_i$. Its elements satisfy $e_i(j) = \delta_{ij}$. For a real variable $x$ we define $(x)_{+} = \max\{x,0\}$. Defining the function $m : \mathbb{C}^q \times \mathbb{S}_+^q \times \mathbb{S}_+^q \to \mathbb{R}_{++}$ and $h : \mathbb{R} \to \mathbb{R}_{+}$,

\[
m(x,B,A) = \frac{x^H B^{-1} A B^{-1} x}{x^H B^{-1} x} \quad (1)
\]

\[
h(x) = \begin{cases} 0 & x < 1 \\ -\log x + x - 1 & x \geq 1 \end{cases} \quad (2)
\]

If $T_1(Y)$ and $T_2(Y)$ are two real functions of an observation set used as test statistics, then the notation $T_1(Y) \overset{\Delta}{=} T_2(Y)$ indicates the equivalence of their corresponding detectors, which occurs when $T_1(Y)$ and $T_2(Y)$ are increasing functions of each other.

III. PROBLEM FORMULATION AND MAIN RESULTS

This section overviews previous detectors and formulates the main results of this work and relating them to existing detectors. III-A briefly presents different signal models related to the one considered here. The unconditional signal model and the corresponding maximum likelihood (ML) estimators and GLRT detector are described in III-B. Detectors derived for related models are discussed in III-C.

A. Overview of signal models

The following assumption applies to all the models presented in this section. An array of $q$ antennas receives a dataset comprised of $N$ $q$-dimensional complex vectors. Under $H_0$ the vectors consist of zero-mean i.i.d. Gaussian background signals and noise whose combined covariance matrix will be referred to as the background covariance matrix (BCM) and denoted by $R_s \in \mathbb{S}_+^q$. Under $H_1$ they contain the signal of interest in addition to the noise. This dataset is referred to as the “primary dataset” and is denoted by the matrix $X = (x_1,\ldots,x_N)$ where each column $x_i \in \mathbb{C}^q$ is a single sample. Some models assume that a secondary dataset, comprised of $N$ samples of noise is available. In this case, the noise in the secondary dataset is zero-mean i.i.d. with the same covariance as the noise in the primary dataset. The secondary dataset is denoted by $Y = (y_1,\ldots,y_N)$. In all models $x_i = \alpha_i a_s + n_i$ where $a_s \in \mathbb{C}^q$ is the signal steering vector and $n_i \sim \mathcal{CN}_q(0,R_s)$ is the additive noise. In deterministic models, $\alpha_1,\ldots,\alpha_N$ is a set of deterministic unknown scalars and in the Gaussian models $\alpha_i \sim \mathcal{CN}_1(0,\sigma^2_s)$ where $\sigma^2_s$ is the signal power. We define the primary SCM and secondary SCM by

\[
\hat{R}_p = N^{-1}XX^H, \quad \hat{R}_s = N^{-1}YY^H \quad (3)
\]

Denote the observation set by $Y = (\hat{R}_p, \hat{R}_s)$. In a Gaussian model, the SCMs $\hat{R}_p$ and $\hat{R}_s$ are a sufficient statistic for estimating the unknown parameters, and therefore contain all
the information that can be used when deriving the estimators and the test statistic.

We consider two types of signal steering vector search domains: a known array manifold, denoted by \( S \subset C^q \), and a general vector in \( C^q \). The former is used when working with a calibrated array and a far field source. The latter is used when calibration is inaccurate or the sources are in the near field. These two cases are referred to as “structured” and “unstructured”, respectively. A detailed analysis of the near field versus the far field assumption is provided in [21], [22].

Models can be categorized by the assumed type of signal (Gaussian or deterministic unknown), whether \( R_s \) is known or not, and the knowledge available on \( a_\ast \) - known, structured or unstructured.

A deterministic model is used when there is no prior knowledge on the distribution of the signal amplitude. Such is the case, for example, when trying to detect a burst of communication signal with unknown power and modulation. It is generally preferable to use a parametric probabilistic model when possible, since it reduces the number of unknown parameters and leads to likelihood functions that can model the mechanism that generated the observations better.

The noise covariance is considered known in the case where the channel gains are calibrated and there are no external noise sources, or there are stationary noise sources and the available noise-only dataset is much larger than the dataset in which the signal is being sought. This type of scenario occurs in the context of astronomical transient detection when searching for signals whose duration is far shorter than the time interval in which the array phase response to a stationary source in the near field versus the far field assumption is provided in [21], [22].

The main results of this work are presented in the following theorem.

**Theorem 1.** Under the unconditional model given in (4), the ML estimator of the BCM under \( H_0 \) is

\[
\hat{M} \equiv \frac{1}{2}(\hat{R}_s + \hat{R}_p) .
\]

Under \( H_1 \), the ML estimators of the signal parameters and the BCM are given by:

\[
\hat{a}(Y) = v_{\text{max}} \left( \hat{R}_p \hat{R}_s^{-1} \right)
\]

\[
\hat{\sigma}^2(Y) = \lambda_{\text{max}} \left( \hat{R}_p \hat{R}_s^{-1} \right) - 1
\]

\[
\hat{R}(Y) = \frac{1}{2} \left( \hat{R}_s + \hat{R}_p - \hat{\sigma}^2 \hat{a} \hat{a}^H \right)
\]

and the GLRT test statistic is equivalent to

\[
T_{\text{GLRT}}(Y) = \lambda_{\text{max}} \left( \hat{R}_p \hat{R}_s^{-1} \right)
\]

This theorem is proved in section IV.

Note that this model has an inherent uncertainty since a complex factor can be exchanged between \( \hat{a} \) and \( \hat{\sigma}^2 \). As is customary in array processing problems, we resolve this by requiring that \( \hat{a} \in C^q \), i.e. \( ||\hat{a}|| = 1 \) and its first non-zero coordinate is real and positive.

Note that the estimated BCM under \( H_1 \) is the average of the two SCMs, after the rank-1 component attributed to the transient signal, namely \( \hat{\sigma}^2 \hat{a} \hat{a}^H \), is subtracted from \( \hat{R}_p \). Furthermore, \( \hat{R}(Y) \in S^q_{++} \) as will be observed in (37), where it is represented as a diagonal PD matrix multiplied from both sides by an invertible matrix and its Hermitian conjugate.

**Sufficient statistic and CFAR:** Let \( L \) be the Cholesky decomposition of \( R_s \), constrained such that \( R_s = LL^H \) and \( \|L^{-1}a_\ast\|^{2}L^{-1}a_\ast = e_1 \). We define the observed SCMs, whitened w.r.t. the true, unknown \( R_s \) as \( A \equiv L^{-1}R_pL^{-H} \), \( B \equiv L^{-1}R_sL^{-H} \). Under \( H_0 \), they are statistically independent random white Wishart matrices, and under \( H_1 \) we have
that \( A \sim W_q(I_q + \sigma_s^2e_ie_i^H, N) \) where \( \sigma_s^2 \) denotes the signal to background ratio (SBR), and is defined as follows:

\[
\sigma_s^2 \equiv \sigma_s^2a_s^Ha_s^{-1}a_s \quad \text{(SBR)} .
\] (10)

The unconditional test statistic can be written as \( T_{\theta}^{(Y)} = \lambda_{\text{max}}(AB^{-1}) \). Under \( H_0 \), it is distributed as the largest generalized eigenvalue of two independent white Wishart matrices, and is hence independent of all the unknown parameters, a property known as the constant false alarm rate (CFAR). Under \( H_1 \), the distribution of \( T_{\theta}^{(Y)} \) depends solely on the SBR; therefore, it is a sufficient statistic for the unconditional model.

### C. Related models

In this section we review GLRT detectors obtained for signal models similar to the unstructured Gaussian model described in subsection III-B. We begin this review with two detectors designed for deterministic unknown signals. In [14] the authors consider a general signal model that can be specialized to a model similar to ours, in which two datasets of \( N \) observations can be distributed as follows

\[
y_i \sim CN_q(0, R_s) \quad x_i \sim \begin{cases} \mathcal{CN}_q(0, R_s) & H_0 \text{ (deterministic)} \\ \mathcal{CN}_q(\alpha_i a_s, R_s) & H_1 \end{cases} ,
\] (11)

where the noise covariance \( R_s \in S^n_+ \) and scaling factors \( \alpha_1, ..., \alpha_N \in \mathbb{C} \) are deterministic unknown, \( a_s \in \mathbb{C}^q \) is known and \( i = 1, ..., N \), as is the case for all the distribution laws in this subsection. The generalized likelihood ratio (GLR) for this model was presented by Besson et al. [17] as the ratio of the minimum variance distortionless responses (MVDR) [23] evaluated at the known source direction using \( \hat{R}_s \) and \( (R_s + \hat{R}_p) \)

\[
T_{\theta}^{\text{MVDR}}(Y) = \frac{a_s^H \hat{R}_s^{-1}a_s}{a_s^H (R_s + \hat{R}_p)^{-1}a_s} .\] (12)

The matched subspace detector for noncoherent detection developed in [13] was designed for a single vector sample which under \( H_0 \) contains only complex zero mean Gaussian noise with a known covariance matrix. Under \( H_1 \) it also contains a deterministic unknown signal that lies in a known subspace. It is straightforward to extend this detector to a multisample model, as given in (11) with both \( a_s \) and \( R_s \) assumed to be known, which obviates the need for the secondary dataset. The detector obtained by this extension is referred to as the “clairvoyant deterministic detector” and its test statistic is given by

\[
T_{\theta}^{\text{1-step}}(Y) = Nm(a_s, R_s, \hat{R}_p)
\] (13)

where \( m(\ldots) \) is defined in (1).

In [17] a GLRT detector for zero-mean Gaussian signals with known \( a_s \) was developed. The detector uses two datasets distributed according to (4) where \( R_s \) and \( \sigma_s^2 \) are unknown. The GLR corresponding to this detector, which is referred to as the “one-step GLRT” and is denoted by \( T_{\theta}^{\text{1-step}}(Y) \), is the product of (12) and another term whose computation involves optimizing a scalar function of the observed SCMs and \( a_s \).

Since the GLRT in [17] is computationally intensive, Besson et al. developed a simpler approach, referred to as the “two-step GLRT”. It is based on a multiple sample extension to the single sample adaptive matched filter in [24] which relies on the unrealistic assumption that the SCM of the secondary dataset is the actual noise covariance matrix (\( \hat{R}_s \rightarrow R_s \)). This leads to the following distributions

\[
y_i \sim \mathcal{CN}_q(0, R_s) \quad x_i \sim \begin{cases} \mathcal{CN}_q(0, \hat{R}_s) & H_0 \\ \mathcal{CN}_q(0, \hat{R}_s + \sigma_s^2a_s a_s^H) & H_1 \end{cases} .
\] (14)

Whitening \( x_i \) w.r.t. \( \hat{R}_s \) reduces the two-step model to a simple Gauss-Gauss detection problem, where the signal is Gaussian with known steering vector \((\hat{R}_s)^{-1/2}a_s\) and the noise is Gaussian with covariance \( \hat{R}_q \). The GLR test statistic for this model is

\[
T_{\theta}^{\text{2-step}}(Y) = Nh \left( m(a_s, \hat{R}_s, \hat{R}_p) \right)
\] (15)

where \( m(\ldots) \) and \( h(x) \) are defined in (1) and (2), respectively.

It has a computational advantage over the one-step detector since it does not require an optimization process, at the cost of reduced \( p_d \). It is shown in [17] that for large \( N \) the loss of \( p_d \) is not significant. A GLRT detector for Gaussian signals with a known steering vector and a known BCM is obtained by substituting the known \( R_s \) for \( \hat{R}_s \) in \( T_{\theta}^{\text{2-step}}(Y) \), resulting in a test statistic equivalent to (13).

A model in which two datasets are collected by separate channels which receive a common zero-mean Gaussian signal under \( H_1 \), is considered in [19] for rank-1 signals and in [25] for subspace signals. Let \( x_s[n] \) and \( x_r[n] \) be the \( q \)-vectors sampled by two channels, and \( R \) be the covariance of the two vectors stacked into \( x[n] = (x_s[n]^T \ x_r[n]^T)^T \in \mathbb{C}^{2q} \). The two hypotheses can be formulated through the structure of \( R \) as follows

\[
R = \begin{pmatrix} \Sigma_{ss} & \Sigma_{sr} & 0 \\ \Sigma_{sr}^T & H_r H_r^H & H_r H_r^H \\ 0 & H_r H_r^H & \Sigma_{rr} + H_r H_r^H \end{pmatrix}
\] (16)

where \( \Sigma_{ss}, \Sigma_{sr} \in S^q_+ \) are the noise covariance matrices which are unknown unstructured PD matrices under the most general model considered, and \( H_s, H_r \in \mathbb{C}^{M \times q} \) represent the \( M \times q \) channels from the transmitters to the two receiving arrays. In contrast to our model, in [25] it is assumed that the correlation between samples from different datasets (given by \( H_r H_r^H \)) is non-zero, which enables detection of a common signal without any prior knowledge of the noise or signal statistics. Under our assumptions the detector of [25] completely fails, since it has no reference signal available (\( H_r = 0 \)).

A GLRT detector for the case where the signal is Gaussian with an unknown steering vector \( a_s \) and a power \( \sigma_s^2 \), using a single dataset, was developed in [26] and in [12], under the
assumption that the noise is uncorrelated with the constant unknown power $\sigma_n^2$. Formally

$$x_i \sim \begin{cases} \mathcal{CN}_q(0, \sigma_n^2 I_q) & H_0 \\ \mathcal{CN}_q(0, \sigma_n^2 I_q + \sigma_s^2 a_i a_i^H) & H_1 \end{cases} \quad (17)$$

The test statistic of the detector corresponding to (17) is

$$T_{\hat{R}_n} = \frac{\lambda_{\max}(\hat{R}_p)}{\text{tr}(\hat{R}_p)} \cdot (18)$$

It can be used to detect near-field interference in radio telescopes when the elements of the array are calibrated (at least w.r.t. gain) and the static celestial and terrestrial signal sources are much weaker than the thermal noise, whose power is unknown. In [11] the GLRT was derived for an arbitrary unknown diagonal covariance matrix.

Finally we present the detector implemented in the main stage of the LOFAR transient detection pipeline [27], which is based on subtracting consecutive dirty images. A dirty image is generated from a given SCM $\bar{R}$ as follows: Let $a(\phi) \in \mathcal{N}$ be the steering vector that corresponds to direction $\phi$. The brightness of the dirty image in direction $\phi$ is given by $a(\phi)^H \bar{R} a(\phi)$. The test statistic of the dirty image subtraction detector is

$$T_{\text{DI}}(\hat{y}) = \max_{a \in \mathcal{N}} a^H (\hat{R}_p - \hat{R}_n) a \cdot (19)$$

The actual processing pipeline of LOFAR is more complicated than the simple detection formula that was presented here, but the other processing stages that were omitted are not related to the topic discussed in this paper.

IV. PROOF OF THE MAIN THEOREM

In this section we derive the results presented in Theorem 1. We begin by introducing the generalized eigen-decomposition of matrix pairs in IV-A. In IV-B the log-likelihood functions are developed and maximized under the assumption that the BCM is given, obtaining ML estimators of the source steering vector and power that depend on the BCM. In subsection IV-C we derive the BCM ML estimators under the two hypotheses. The BCM estimator under $H_1$ is then used in IV-D to obtain ML estimators of the transient parameters; i.e., we estimate $\hat{a}(R)$ and then $\hat{\sigma}^2(\hat{a}(R), R)$. Then we substitute $\hat{a}(R)$ and $\hat{\sigma}^2(\hat{a}(R), R)$ into the log-likelihood and optimize over $R$ to compute $\hat{R}(\hat{y})$. This is used to derive the GLRT in IV-E.

The derivation of the detector is relatively complicated, since the maximization of the likelihood function under $H_1$ should be performed w.r.t. three parameters simultaneously, where the first is a matrix (BCM), the second is a vector (steering vector) and the third is a scalar (power). Moreover, the likelihood as a function of the BCM is not concave, which significantly complicates the derivation and proofs.

A. Generalized eigen-decomposition

For a given ordered pair of PD matrices $P = (A, B) \in S^q_+ \times S^q_+$, there exists a triple of matrices $(V_P, U_P \in F^q, \Lambda_P \in D^q_+)$ which satisfies the following set of equations:

$$B = V_P V_P^H$$
$$A = V_P \Lambda_P V_P^H$$
$$U_P = V_P^H \cdot (20)$$

The matrices $U_P$, $V_P$ and $\Lambda_P$ are referred to as the generalized-eigen-matrix, the inverse generalized-eigen-matrix, and the joint eigen-matrix of $P$ respectively. The columns of $U_P$ and $V_P$ denoted by $\{u_{p,1}, \ldots, u_{p,q}\}$ and $\{v_{p,1}, \ldots, v_{p,q}\}$, are the generalized eigenvectors and inverse eigenvectors, respectively, and the joint eigenvalue that is associated with $u_{p,i}$ and $v_{p,i}$ is the $i$th diagonal entry of $\Lambda_P$, which is denoted by $\lambda_{p,i}$. It follows from (22) that $U_P^H v_{p,i} = V_P^H u_{p,i} = e_i$ for all $i = 1, \ldots, q$, hence $v_{p}^H, u_{p} \sim \delta_{ij}$ and

$$A u_{p,i} = \lambda_{p,i} u_{p,i}, B v_{p,i} = \lambda_{p,i} v_{p,i}$$
$$A^{-1} v_{p,i} = \lambda_{p,i}^{-1} B^{-1} v_{p,i} = \lambda_{p,i}^{-1} u_{p,i} \cdot (23)$$

The largest joint eigenvalue of $P$, its associated generalized eigenvector and inverse eigenvector are denoted by $\lambda_P$, $\bar{u}_P$ and $\bar{v}_P$, respectively. Multiplying (24) by $A$ or by $V_P^H$ we obtain the following identities

$$v_{\max} (A B^{-1}) = \frac{v_P}{\|v_P\|} \cdot (25)$$
$$\lambda_{\max} (A B^{-1}) = \lambda_P \cdot (26)$$
$$\bar{v}_P B^{-1} \bar{v}_P^H = 1 \cdot (27)$$

The set $(V_P, U_P, \Lambda_P)$ corresponding to a given pair of PD matrices $P = (A, B)$ can be obtained as follows. Use the Cholesky decomposition to write $B = C C^H$ with $C \in F^q$ as a consequence of $B \in S^q_+$. Let $T$ be a $q$-dimensional unitary matrix whose columns are the eigenvectors of the matrix $D \equiv (C^{-1} A C^{-H}) \in S^q_+$ and $\Lambda$ a diagonal matrix whose diagonal elements are the eigenvalues of $D$ ordered to correspond to the eigenvector ordering in $T$, so that $D = T \Lambda T^H$. The matrices $V_P$ and $\Lambda_P$ can be identified as $C T$ and $\Lambda$, respectively. For $U_P$ we use (22). It is shown that the generalized eigen-decomposition presented in (20) and (21) exists for any pair of PD matrices.

B. Maximization of the log-likelihood functions for a given $R$

We now derive the log-likelihood functions of the unknown parameters under the two hypotheses for the measured $\hat{R}_n$ and $\hat{R}_p$. It is useful to recall that the log-probability-density of a measured SCM $\hat{C}$ computed from $N$ i.i.d realizations of a $q$-dimensional zero mean complex normal random vector with covariance $C$, is

$$\log p(\hat{C}; C) = -N \left( q \log \pi + \log |C| + \text{tr} \left( C^{-1} \hat{C} \right) \right) \cdot (28)$$
In what follows, we omit the constant terms $N\log\pi$ from the log-likelihood expressions, since they play no role in their maximizations or in the log-likelihood ratio.

Because the two sets of observations are statistically independent under both hypotheses, the two log-likelihood functions can be separated into the sum of the two parts, each corresponding to the log-likelihood of a single observed SCM. Under $H_0$, $\hat{R}$ is the covariance of all the samples in $\bar{R}$ and $\hat{R}_p$; hence, from (28) we have that the $H_0$ log-likelihood is

$$\ell_0(\gamma; \hat{R}) = -2\log|\hat{R}| + \text{tr} \left( \hat{R}^{-1} \left( \bar{R}_a + \hat{R}_p \right) \right).$$  \hspace{1cm} (29)

For the log-likelihood function under $H_1$ we subtract the log-likelihood of $\hat{R}$, given by $\log|\hat{R}| + \text{tr} \left( \hat{R}^{-1} \bar{R}_a \right)$ from the log of the $H_1$ likelihood that was derived in Appendix B of [17]

$$\ell_1(\gamma; \hat{R}, a, \sigma^2) = -2\log|\hat{R}| - \log(1 + \sigma^2a^H\hat{R}^{-1}a)$$

$$- \frac{\sigma^2a^H\hat{R}^{-1}\bar{R}_a\sigma^2}{1 + \sigma^2a^H\hat{R}^{-1}a} - \left( \log|\hat{R}| + \text{tr}\left( \hat{R}^{-1} \bar{R}_a \right) \right)$$

$$= \ell_0(\gamma; \hat{R}) - \log(1 + \sigma^2) + \frac{\sigma^2m(a, R, \hat{R}_p)}{1 + \sigma^2}$$  \hspace{1cm} (30)

where we define $\sigma^2 \equiv \sigma^2a^H\hat{R}^{-1}a$ and $m(\cdots)$ is defined in (1).

Next, we derive estimators of the signal parameters and the maximal $H_1$ log-likelihood as functions of the BCM. The log-likelihood maximized w.r.t. $\sigma^2$, as obtained in [17], is

$$\max_{\sigma^2 \geq 0} \ell_1(\gamma; \hat{R}, a, \sigma^2) = \ell_0(\gamma; \hat{R}) + h \left( m(a, R, \hat{R}_p) \right)$$  \hspace{1cm} (31)

where $h(x)$ is defined in (2). Using the derivation of the maximum of $m(a, R, \hat{R}_p)$ from appendix A and the fact that $h(x)$ is an increasing function we have that the log-likelihood maximized w.r.t. $a$ is

$$\ell_1(\gamma; \hat{R}) \equiv \max_{a \in C^n} \max_{\sigma^2 \geq 0} \ell_1(\gamma; \hat{R}, a, \sigma^2) = \ell_0(\gamma; \hat{R}) + h(\hat{\lambda}_W)$$  \hspace{1cm} (32)

where we define the matrix pair $W \equiv (\hat{R}_p, R)$, this maximum is attained at the estimator of the steering vector, which at this stage is expressed as a function of another still unknown estimator $\hat{R}$

$$\hat{a}(\gamma, \hat{R}) = \arg \max_{a \in C^n} \left\{ \max_{\sigma^2 \geq 0} \hat{\ell}_1(\gamma; \hat{R}, a, \sigma^2) \right\} = \frac{\psi_W}{\|\psi_W\|}$$  \hspace{1cm} (33)

The power estimator as a function of $\hat{R}$ is obtained by plugging $\hat{a}(\gamma, \hat{R})$ back into the log-likelihood (30) and maximizing the result w.r.t. $\sigma^2$

$$\hat{\sigma}^2(\gamma, \hat{R}) = \arg \max_{\sigma^2 \geq 0} \left\{ \frac{\sigma^2\hat{\lambda}_W}{\|\psi_W\|^2 + \sigma^2} - \log \left( 1 + \frac{\sigma^2}{\|\psi_W\|^2} \right) \right\}$$

$$= \|\psi_W\|^2 \hat{\lambda}_W^{-1} \hat{\lambda}_W - 1 \right)^+$$  \hspace{1cm} (34)

where we used $\hat{a}(\gamma, R)^H \hat{R}^{-1} \hat{a}(\gamma, R) = \|\psi_W\|^{-2}$ which follows from (27).

### C. BCM Estimators

We observe that $\ell_0(\gamma; \hat{R})$, given in (29), is the standard log-likelihood of a set of observed zero-mean Gaussian vectors whose SCM is $M$. Anderson [28] proved that this function is maximized w.r.t. the covariance matrix by the observed SCM; hence, as stated by Theorem 1, $M$ is the ML estimator of the BCM under $H_0$, and the maximal log-likelihood is

$$\ell_0(\gamma; \hat{R}) \equiv \ell_0(\gamma; M) = -2\log|M| - 2q.$$  \hspace{1cm} (35)

We now prove that $\hat{R}(\gamma)$ given in (8) is indeed the ML estimator of the BCM under $H_1$ for the full unstructured model.

**Proof Outline:** In the following proof we first introduce a new function $s : S^+_n \rightarrow \mathbb{R}$ which attains its maximum at the inverse of the maximum of $\ell_1(\gamma; \hat{R})$, and is simpler to maximize. By diagonalizing its arguments and analyzing its local behavior on an arbitrary straight line passing through $S^+_n$, we show that local maximum points of $s$ can only be found in a subset of $S^+_p$ given by a linear map of the set of PD diagonal matrices (which has considerably fewer dimensions than $S^+_n$). The next step is to further narrow down the set of candidates for maximum points of $s$ into a set that contains a single matrix. Finally, by showing that $s$ must have a maximum point, it is established that the only candidate left is indeed the ML estimator of the BCM.

The detailed proof is deferred to Appendix B.

### D. Signal Parameter Estimators

In this subsection we prove that the steering vector and power estimators given in Theorem 1 are obtained by substituting $\hat{R}(\gamma)$ into the dependent estimators given in (33) and (34). First, we use (25) and (27) to rewrite the estimators given in (6) and (7) as follows

$$\hat{\sigma}^2(\gamma) = \frac{\|\psi_\gamma\|^2}{\|\psi_\gamma\|^2} \left( \hat{\lambda}_Y - 1 \right), \quad \hat{a}(\gamma) = \frac{\psi_\gamma}{\|\psi_\gamma\|}$$  \hspace{1cm} (36)

Using these expressions, $\psi_\gamma = V_\gamma e_1$ and the generalized eigen-decomposition of $(\hat{R}_p, \hat{R})$, we write (8) as

$$\hat{R}(\gamma) = V_\gamma \Lambda_\gamma V_\gamma^H = V_\gamma V_\gamma^H$$  \hspace{1cm} (37)

where $\hat{W} \equiv (\hat{R}_p, \hat{R}(\gamma))$ and

$$\Lambda_\gamma = \frac{1}{2} \left( I_q + \Lambda_Y - (\hat{\lambda}_Y - 1) e_1 e_1^H \right)$$

$$= \text{diag} \left( 1, \frac{\lambda_{y,2} + 1}{2}, ..., \frac{\lambda_{y,q} + 1}{2} \right)$$  \hspace{1cm} (38)

It follows from (37) that $V_\gamma V_\gamma^H = V_\gamma \Lambda_\gamma V_\gamma^H$ where $\Lambda_\gamma \in \mathbb{D}_{++}^q$. Combining this with the decomposition $\hat{R}_p = V_\gamma \Lambda_\gamma V_\gamma^H$ we obtain

$$\Lambda_{\hat{W}} = \Lambda_\gamma^{-1} \Lambda_Y = \text{diag} \left( \hat{\lambda}_Y, \frac{2\lambda_{y,2}}{1 + \lambda_{y,2}}, ..., \frac{2\lambda_{y,q}}{1 + \lambda_{y,q}} \right).$$  \hspace{1cm} (39)

Because $\hat{\lambda}_Y > 1$ and $\lambda_{y,i} \geq \lambda_{y,1}$ for any $i$, it is easy to show that the largest joint eigenvalue of $\hat{W}$ is

$$\hat{\lambda}_{\hat{W}} = \lambda_{\hat{W},1} = \lambda_Y.$$  \hspace{1cm} (40)
and hence \( \psi_{\psi} = \psi_Y \). These results, when substituted into the general expressions (33) and (34), yield the estimators in (36) which are equal to the estimators given in Theorem 1 (using the fact that \( \lambda_Y > 1 \) to omit the \((-)\) from the power estimator).

E. Test Statistic

For the test statistic, we first develop the trace and log-determinant terms of (29) with \( \hat{R}(Y) \rightarrow R_\psi \) by substituting (37), (69) and \( M = \frac{1}{2}V_Y(I_q + \Lambda_Y)\bar{V}^H_Y \) and using the properties of the trace and determinant

\[
\text{tr} \left( \hat{R}(Y)^{-1} \left( \hat{R} + \hat{p} \right) \right) = \text{tr} \left( \Lambda_p^{-1}(I_q + \Lambda_Y) \right) = \lambda_Y + 2\eta - 1
\]

\[
\log |\hat{R}(Y)| = \log |\hat{M}| - \log \left| \frac{1}{2}(I_q + \Lambda_Y)\Lambda_p^{-1} \right|
\]

\[
= \log |\hat{M}| + \log \frac{2}{1 + \lambda_Y}.
\]

These results are now used to express the difference between the \( H_0 \) log-likelihood function (29) evaluated at \( \hat{R}(Y) \) and its maximal value (35)

\[
\ell_0(Y; \hat{R}(Y)) - \ell_0^*(Y) = -2\log \frac{2}{1 + \lambda_Y} - \lambda_Y + 1 \quad (41)
\]

It follows from (32) (with \( \mathcal{W} = \hat{\mathcal{W}} \)) together with (41) and (40) that the GLR corresponding to the unconditional model is

\[
\ell_1^*(Y; \hat{R}(Y)) - \ell_0^*(Y) = \left( \ell_0(Y; \hat{R}(Y)) + h(\lambda_Y) \right) - \ell_0^*(Y)
\]

\[
= -2\log \frac{2}{1 + \lambda_Y} - \lambda_Y + 1 + h(\lambda_Y) = 2\log \frac{1 + \lambda_Y}{2\sqrt{\lambda_Y}}. \quad (42)
\]

When \( \lambda_Y \geq 1 \) the GLR is a strictly increasing function of \( \lambda_Y \), which is equal to the test statistic of the unconditional model given in (9), as shown in (26); thus, the proof of Theorem 1 is complete.

V. Extension of Existing Detectors to an Unknown Steering Vector

We now discuss the applicability of the detectors in [11], [17], [14] to our problem. Since these detectors were derived for a known steering vector, the generalization of these detectors relies on optimization over the complete field of view, whether it is the array manifold \( S \) or \( C^N \).

All the detectors presented here that assume knowledge of \( a_* \) can also be used to detect signals whose steering vector is unknown by using the maximum of the GLR over the domain in which the steering vector is expected to be as the new test statistic. Extensions of a detector assuming known steering vector, say \( T_{i|S}(\cdot) \), to a model with a structured and unstructured unknown steering vector, which are denoted by \( T_{i|S}(\cdot) \) and \( T(\cdot) \), respectively, are computed as follows

\[
T_{i|S}(\cdot) \equiv \max_{a_i \in S} T_{i|a_i}(\cdot) \quad \text{(structured)} \quad (43)
\]

\[
T(\cdot) \equiv \max_{a_i \in C^N} T_{i|a_i}(\cdot) \quad \text{(unstructured)} \quad (44)
\]

Note that the extension of the one-step model to the unstructured steering vector settings results in the unconditional model, which implies that \( T_{i|0}(Y) = \max_{a_i \in C^N} T_{i|a_i}(Y) \). However, it is simpler to derive (9) using our model as in Section IV.

Combining the test statistic from Theorem 1 and the equivalences proved in Appendix A we conclude that if no parameter is given, then \( T_{i|0}(Y) = \lambda_Y \) serves as a GLRT test statistic for the Gaussian (4), simplified Gaussian (14) and deterministic (11) signal models. This is a surprising and important observation. While knowledge of the array manifold differentiates the various models in the literature. Not knowing the directional vector, which is realistic in many cases, makes all these detectors equivalent in the unstructured case. In contrast, using the structural array data results in different detectors.

Thus below, we propose a general two step approach which is instrumental in the rapid removal of terrestrial radio-frequency interference (RFI) from FRB searches.

Similarly, the extension of the detectors that assume knowledge of \( R_\psi \) (13) to the unstructured case is given by \( T_{i|R_\psi}(R_q) = \lambda_{max} \hat{R}_p R_{\psi}^{-1} \).

Table I lists the relationships between detectors in the cases of a known and unknown steering vector.

A. Computational Complexity

The computation of the largest joint eigenvalue of the pair \( (\hat{R}_p, \hat{R}_s) \), required for the unconditional detector, can be performed by whitening \( \hat{R}_p \) w.r.t. \( \hat{R}_s \) and finding the largest eigenvalue of the resulting Hermitian matrix using standard eigenvalue methods [29]. This scheme involves the following steps: (1) compute the SCMs \( R_s, R_p \) (2) compute the Cholesky decomposition \( J \) of \( R_s \), (3) invert \( J \), (4) compute \( \hat{R}_p = J^{-1} \hat{R}_p J^{-H} \) and (5) compute the largest eigenvalue of \( \hat{R}_p \). The computation of the standard largest eigenvalue test statistic only requires steps (1) (for a single SCM) and (5), and its computational complexity is \( q^2(2N - 1) + O(q^3) \) [12]. To these we add the addition of another SCM \( q^2(2N - 1) \) operations), the computation of \( J \) (\( O(q^3) \) operations), inversion of \( J \) (\( O(q^2) \) operations) and two matrix multiplication for \( J^{-1} \hat{R}_p J^{-H} \) (\( O(q^3) \) operations). In total, the computational complexity of the unstructured Gaussian detector is:

\[
\text{complexity} (T_{i|0}(Y)) = 2q^2(2N - 1) + O(q^3). \quad (45)
\]

Next, we derive the computational complexity of the 2-step structured detector, since it is the simpler of the two Gaussian structured detectors \( T_{i|S}(Y) \) and \( T_{i|S}(Y) \). Suppose we define a grid of \( L \) directions that spans the area of interest and corresponds to a set of steering vectors \( \{ a_{11}, ..., a_{1L} \} \subset S \). The detector computes \( T_{i|S}(Y) \) from (15) for \( i = 1, ..., L \) and then uses the maximal value as the test statistic. Three computations are performed once and used by all the sampled directions: (1) computation of \( \hat{R}_p \) and \( \hat{R}_s \) (2) inversion of \( \hat{R}_s \) and (3) the matrix multiplication \( \hat{R}_s^{-1} \hat{R}_p \hat{R}_s^{-1} \). These computations sum to \( 2q^2(2N - 1 + 1) + 2q^3 \) operations. Then for every \( a_i \) we compute \( (a_i^H \hat{R}_s^{-1} \hat{R}_p \hat{R}_s^{-1} a_i) \) and divide the result.
by \((a^H R_s^{-1} a_i)\) to obtain \(m(a_i, R_s, R_p)\). The computation of each of the terms requires \((q^2 + q)\) operations. By summing them all, we have that the computational complexity of the structured two-step detector is

\[
\text{complexity} \left( T^{2\text{-step}}_{|S}(Y) \right) = L (q^2 + q + 1) \\
+ 2q^2 (2N - 1) + O(q^3) .
\]

If \(L\) is large compared to \(N\), which is typically the case when covering the sky with narrow beams, the structured detector is more computationally expensive than the unstructured detector.

**B. Cascade detector**

A structured and an unstructured detector for Gaussian signals with an unknown steering vector were introduced in subsection III-A. The structured detectors provide directional selectivity at the cost of higher computational complexity, compared to the unstructured detector, as shown in the subsection above and demonstrated by simulation in VI-E. Applying a structured detector in cascade with an unstructured one combines the advantages of the two types. In this detection scheme, the unconditional detector \(T_{|\emptyset}(Y)\) is applied first in order to detect any suspicious signal, and then uses the steering vector estimation to only pass signals that might come from a celestial source to the next stage. In the second stage, the structured detector \(T^{1\text{-step}}_{|S}(Y)\) or \(T^{2\text{-step}}_{|S}(Y)\) is used to confirm or reject the first stage detections. This way, terrestrial interferences are filtered out from the detection process, and the computational demands are relaxed compared to a detection scheme that passes all the data to the structured detector.

We provide a qualitative discussion of the performance of the cascade detector. Let \(T_i(Y), C_i, \gamma_i\) for \(i = 1, 2\), denote the test statistic, computational complexity and detection threshold of the two detection stages, where \(i = 1\) is associated with the unstructured stage and \(i = 2\) with the structured stage. Assuming that \(H_1\) is a rare event, the complexity of the cascade compared to a single stage structured detector is given by

\[
\frac{C_{12}}{C_2} = \frac{C_1}{C_2} + p^{(1)}_m (\gamma_1) .
\]

The test statistics \(T_1(Y)\) and \(T_2(Y)\) satisfy (43) and (44) for some test statistic \(T_{|\emptyset}(Y)\) and array manifold \(S \subset C^q\); hence \(T_1(Y) \geq T_2(Y)\) for any \(Y\). Assume \(\gamma_2\) is calibrated to a value at which the structured detector provides satisfactory performance as a single stage detector. In this case the \(p_{fa}, p_{d}\) and complexity of the two stage detector depend on \(\gamma_1\). If \(\gamma_1 = \gamma_2\), the cascade detector has the same detection performance as the structured detector, and reduced complexity if \(p_{fa}^{(1)}(\gamma_2) < 1 - C_1/C_2\). By raising \(\gamma_1\) above \(\gamma_2\) the complexity can be further reduced at the expense of lower performance. The simulation shows that computation speed can be significantly improved without considerable performance degradation.

**VI. SIMULATIONS**

We now present a full sky simulation of the proposed detector \(T_{|\emptyset}(Y)\) and the detectors presented in Subsection III-C or their extensions as given in Section V. For that purpose we used the 3C catalog of sources [30] as a model for the sky. We generated a transient with varying power and direction whose incoming signal is superimposed on the signal received from the stationary sources and the receiver noise.

The key simulation parameters are as follows. We simulated the signals received by an array with \(q = 12\) elements arranged in a non-uniform (to avoid directional ambiguity) \(3 \times 4\) grid, with East-West span of 16 meter and South-North span of 19 meters. The center frequency was 50MHz. The BCM used for the simulation is given by

\[
R_s = I_q + \sum_{m=1}^{50} \sigma_m^2 a_m a_m^H
\]

where \(a_1, \ldots, a_{50}\) are the steering vectors corresponding to the 50 brightest sources that appeared at 6AM on December 12th 2013 in the sky above the LOFAR, computed for the simulated array at a frequency of 50MHz. Their corresponding intensities are given by \(\sigma_1^2, \ldots, \sigma_{50}^2\), normalized such that \(\max_m \sigma_m^2 = 20\), which is a typical SNR (13dB) for a bright source at this frequency. For the structured detectors, the array manifold covers a square patch in the sky, sampled at \(50 \times 50\) points. The simulated transient SBR, defined in (10), ranges from -15dB to 22dB. Performance was evaluated for two locations of the transient: one in a quiet area of the sky, and the other near a strong stationary source. These locations are referred to as the “dark” and “bright” points, respectively. If one transient is placed at the bright point and another at the dark point, the bright point transient must be 6dB brighter than the dark point transient (in terms of absolute power \(\sigma^2\)) to have the same SBR. Using the model given in (4), a set of

---

**TABLE I**

**SUMMARY OF GLRT DETECTORS**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>signal type (G=Gaussion)</td>
<td>known (R_s)</td>
<td>(R_s \in S^I)</td>
<td>(R_s \approx I_s)</td>
<td>(R_s \approx I_s)</td>
<td>(R_s \approx I_s)</td>
</tr>
<tr>
<td>uses secondary dataset</td>
<td>X</td>
<td>X</td>
<td>V</td>
<td>V</td>
<td>X</td>
</tr>
<tr>
<td>BCM assumption</td>
<td>X</td>
<td>X</td>
<td>V</td>
<td>V</td>
<td>X</td>
</tr>
<tr>
<td>test statistic for (a_s \in C^q)</td>
<td>(13)</td>
<td>(12)</td>
<td>(12)</td>
<td>(12)</td>
<td>(12)</td>
</tr>
</tbody>
</table>

---
5,000 pairs of sample matrices $\mathbf{X} = (x_1 \cdots x_N)$ and $\mathbf{Y} = (y_1 \cdots y_N)$ were generated to simulate $H_0$ and a set of 1,000 datasets were generated to simulate every SBR value under $H_1$.

The extension of the one-step detector to the structured model was not included in the simulation because its computation was too demanding for the machine that was used. Its performance when $\alpha_*$ is known can be found in [17].

A. Full sky simulation

The probabilities of detection of the simulated detectors as a function of SBR are shown in Figure 1 for $N = 10^4$ and in Figure 2 for $N = 30$. Once the SBR is fixed, the location of the transient only affects the detectors that assume white noise, namely $T_{R_{\infty}I} (\mathbf{R}_p)$ and $T_{S}^{\text{MVDR}}(\mathbf{Y})$, given in (18) and (19). For the tested scenario with large datasets ($N = 10^4$), the unstructured clairvoyant $T_{R_{\infty}I} (\mathbf{R}_p)$ and two structured detectors $T_{I_{S}^{\text{2-step}}}(\mathbf{Y})$ and $T_{S}^{\text{MVDR}}(\mathbf{Y})$ performed identically. The unconditional detector was about 1dB less sensitive and the dirty image detector was less sensitive than the structured detectors by 8-15dB, depending on the position of the transient relative to the background sources. The detector $T_{R_{\infty}I} (\mathbf{R}_p)$ that assumed white noise was far worse than all the rest and its performance also depended on the position of the transient.

Figure 2 plots the simulation results when the number of samples in each dataset was reduced to 30. The sensitivity gap between the unconditional detector and the structured detectors remained about the same as it was in the scenario where $N = 10^4$. The clairvoyant unstructured detector, that is not sensitive to BCM estimation errors outperformed the rest of the simulated detectors.

These simulation findings call for two comments. The first is that the simulated transient steering vector was contained in the array manifold. Normally, there is some mismatch due to array calibration errors which can significantly reduce the probability of detection of structured detectors, as shown in [17] for detectors that assume knowledge of $\alpha_*$. This mismatch does not affect the unconditional detector since it assumed no array manifold. Second, the detection threshold was set to achieve $p_{fa} = 10^{-2}$ by all the detectors in the simulated scenario. In practice, the threshold of non-CFAR detectors is set such that the desired $p_{fa}$ is obtained in a worst case scenario so that the $p_{fa}$ in all other scenarios is lower. This lower value is what would be used in a more realistic simulation, that computes the threshold based on the worst case scenario. Lower $p_{fa}$ values for non-CFAR detectors are expected to result from this type of simulation. As in the case of steering vector mismatch, CFAR detectors such as the unconditional detector are not affected by this consideration.

B. Dependence on the number of samples

The dependence on the number of samples in each dataset was evaluated for three detectors: the unconditional (9), the unstructured detector that knows $\mathbf{R}_c$ whose test statistic is computed as $T_{I_{S}^{\text{2-step}}}(\mathbf{Y})$ with $\mathbf{R}_c$ substituted by $\mathbf{R}_c$ (15), and the structured MVDR ratio detector (12) whose performance is somewhere between the 2-step and the 1-step detectors.

We simulated a range of scenarios where the total number of samples (denoted by $M$ in Figure 3) was constant and the partition of the samples between the primary and secondary datasets varied. Except for the number of samples, the simulated scenario was the same as described in Subsection VI-A.

Figure 3 shows that the performance of the detectors based on two datasets, namely $T_{S}^{\text{2-step}}(\mathbf{Y})$ and $T_{S}^{\text{MVDR}}(\mathbf{Y})$ performed worse when given a small secondary dataset (the right side of the graph), than with a small primary dataset. The optimal point was when the secondary dataset was slightly larger than the primary dataset. The unstructured and structured detectors performed about the same when the unstructured detector was given 1.5 times more samples. This factor cannot be taken as an inherent property of the detectors, since it depends on the scenario parameters. The clairvoyant unstructured detector improved as the size of the primary dataset increased, as expected (since it makes no use of the secondary dataset). It should be noted that the unconditional detector was derived for a model of equally sized datasets. By further derivation it could be adapted to a scenario as simulated here.
The choice of \( p \) at a small range of SBRs, for the cost of losing a detection with a probability of 4% or 15% can significantly improve computation speed at the tolerable cost of losing a detection with a probability of 4% or 15%. The maximum loss is incurred when \( \lambda = 0.7 \) in the structured stage. The reduction in the probability of detection can be significant for certain SBR values. Its importance depends on prior knowledge of the FRB’s. However since these are rare events, losing 15% of them can be very significant. If this is the case, it is preferable to set the threshold of the first stage such that \( p_{fa}^{(1)}>0.025 \). The choice of \( p_{fa}^{(1)} \) only affects computation speed and not the overall probability of false alarm (which mainly depends on \( p_{fa}^{(2)} \)).

C. Cascade detector

We evaluated the effects of using the cascade detection scheme described in Subsection V-B, compared to a single stage structured detector. For the first stage we used the unconditional detector \( T_\theta(\mathcal{Y}) \), and the MVDR ratio detector \( T_{MVDR}^{(S)}(\mathcal{Y}) \) for the second stage. The scenario corresponded to the one described in the beginning of this section. The computational complexity reduction was linear to the probability of false alarm in the first stage, denoted by \( p_{fa}^{(1)} \) in (47), where the detection performance degradation was expressed as the difference between the \( p_{fa} \) of the structured detector and the \( p_{fa} \) of the cascade detector. The simulation results are shown in Figure 4 for \( N = 1,000 \) and a probability of false alarm in the second stage of \( p_{fa}^{(2)} = 10^{-3} \). The upper graph depicts the \( p_{fa} \) curves for the second stage when used alone, and the \( p_{fa} \) for the cascade detector for two values of \( p_{fa} \) in the first stage, as denoted by \( p_{fa}^{(1)} \). The differences between the probabilities of detection of the single stage and the cascade detectors are shown in the lower graph. By taking a smaller \( N \), the curves shifted to the right but the probabilities did not change. It is evident from Figure 4 that if the computation of \( T_\theta(\mathcal{Y}) \) is faster than the structured detector, which was the case in this simulation, the use of the cascade detector can significantly improve computation speed at the tolerable cost of losing a detection with a probability of 4% or 15% at a small range of SBRs, for the \( p_{fa}^{(1)} \) values used in the simulation (0.1 and 0.025). The maximum loss is incurred when \( p_{fa} = 0.7 \) in the structured stage. The reduction in the probability of detection can be significant for certain SBR values. Its importance depends on prior knowledge of the FRB’s. However since these are rare events, losing 15% of them can be very significant. If this is the case, it is preferable to set the threshold of the first stage such that \( p_{fa}^{(1)}>0.025 \). The choice of \( p_{fa}^{(1)} \) only affects computation speed and not the overall probability of false alarm (which mainly depends on \( p_{fa}^{(2)} \)).

D. Non-stationary background

All the non-clairvoyant detectors are based on the assumption that the noise statistics remain unchanged between the secondary and primary sampling intervals. The robustness of the unconditional detector and the two structured detectors to violations of this assumption was examined in the next simulation. We generated \( \hat{R}_s \) and \( \hat{R}_b \) based on the static source position at two time instances separated by a few seconds, during which the entries of the steering vectors associated with the sources shifted by up to 0.1 [radian] (depending on the time difference, East-West baseline and wavelength). We measured the sensitivity of the simulated detectors to a set of time differences between the first and second intervals, ranging from 0 to 10 minutes. Sensitivity was defined as the minimal SBR required to obtain \( p_{fa} = 0.9 \) with \( p_{fa} = 0.01 \). To make the results more general, below we plot the sensitivity curves as a function of the phase shift measured between the farthest elements on the East-West axis for a source positioned directly above the array (for time intervals much shorter than a day we used the approximation \( \varphi \approx \omega \delta t D/\lambda \), where \( \varphi \) is the phase shift, \( \omega \) is the angular velocity of Earth, \( \delta t \) is the time difference, \( D \) is the maximal East-West baseline and \( \lambda \) is the wavelength). Figure 5 shows that the sensitivity of the unconditional and structured MVDR ratio detectors degraded similarly as the time difference became larger. It is also evident that the dirty image detector was worse even in the most favorable conditions.

These results illustrate the considerations involved in designing transient detectors. On one hand it is desirable to use as many noise samples as possible to reduce the estimation error. On the other, if the noise collecting interval is too long, the covariance of the noise in the primary sampling interval drifts...
away from the noise covariance measured in the secondary interval, which increases the probability of false detection.

E. Runtime

The following runtimes were measured using a standard i7 PC running MATLAB R2019b. Note that they do not include the computation time of $\hat{R}$, and $R_{ps}$, since they are common to all the detectors. The computation for the unstructured detectors: $T_{10}(Y)$, $T_{R,R_p}(R_{ps})$ and $T_{R,R_{ps}}(R_{ps})$, took 100 – 220 µs for a single dataset with $q = 12$. The computation time for the structured detectors over a grid of 50 × 50 = 2,500 steering vectors with $q = 12$ was as follows: $T_{MVDR}(Y) \approx 4$ ms, $T_{MVDR}(Y) \approx 4.1$ ms, $T_{R,R_p}(Y) \approx 2.6$ ms, $T_{R,R_p}(Y) \approx 636$ ms, where for $T_{MVDR}(Y)$ we used a grid of 20 points to achieve the optimization required for the computation of each pixel. Computation of a detector that knows $a_c$ would take as long as the computation of a single pixel with the corresponding structured detector.

This shows that the computation of the structured detectors can be significantly longer than the unstructured detectors, which makes the cascade detector an attractive solution.

VII. CONCLUSION

We developed a closed form GLRT detector for radio transient detection in radio astronomical arrays that can be utilized in several detection schemes. It was also proven to be a GLRT detector for deterministic unknown signals from unknown directions in correlated Gaussian noise with unknown statistics. We showed that it naturally extends various known steering vector techniques and that all these techniques yield the same GLRT when we include optimization over the unknown directional vector. This makes it possible to generalize state of the-art detectors to a general two step detector. The simulation results demonstrated the superiority of the proposed detectors over standard radio transient detectors used in existing radio telescopes. These simulations also showed that if array calibration is assumed, the two-step and MVDR ratio structured detectors perform equally well with better sensitivity compared to the unconditional detector, as could be expected. We expect these results to have an impact on the design of future fast radio burst detectors in radio telescopes such as the Square Kilometer Array.

APPENDIX A

EQUIVALENCE OF DETECTORS

We prove that the test statistics (12) and (15), when maximized w.r.t. $a$ over $C^q$ become equivalent to the largest generalized eigenvalue $\lambda_Y$ of the pair $Y = (R_p, R_{s})$.

First, we find the maximum of $m(a, B, A)$ for a matrix pair $P = (A, B) \in S^+ \times S^+_q$, using the identities $B^{-1} = U_P U_P^H$ and $\Lambda_P = U_P U_P^H$ that were introduced in subsection IV-A

$$\max_{a \in C^q} m(a, B, A) = \max_{a \in C^q} \left( U_P^H a \right)^H U_P^H A U_P \left( U_P^H a \right)$$

$$= \lambda_{\max} \left( U_P^H A U_P \right) = \lambda_{\max} (\Lambda_P) = \bar{\lambda}_P ,$$

Hence it is proved that

$$T_{2-step}(Y) \equiv \max_{a \in C^q} T_{2-step}(a) \triangleq \bar{\lambda}_Y .$$

We also note that the maximum of $m(a, B, A)$ is attained at $a = v_P/\|v_P\|$, since

$$m \left( v_P/\|v_P\|, B, A \right) = \frac{v_P^H B^{-1} A B^{-1} v_P}{v_P^H B^{-1} v_P} = \frac{u_P^H A u_P}{v_P^H u_P} = \bar{\lambda}_P .$$

Next, we derive the test statistic of the MVDR ratio detector for the case where the signal steering vector is unknown

$$T_{MVDR}(Y) \equiv \max_{a \in C^q} T_{MVDR}(Y)$$

$$= \max_{a \in C^q} \left( U_Y^H a \right)^H \left( U_Y^H a \right)$$

$$= \max_{a \in C^q} a^H (I_q + A_Y)^{-1} a = 1 + \lambda_Y$$

hence $T_{MVDR}(Y) \triangleq \bar{\lambda}_Y$.

APPENDIX B

DERIVATION OF THE ML ESTIMATOR OF THE BCM

For the following derivation let $s : S^+_q \rightarrow \mathbb{R}$ be defined by:

$$s(Z) \equiv \ell_1^*(Y; Z^{-1}) = 2 \log |Z| - 2 \text{tr} \left( Z \hat{M} \right) + h \left( \lambda_{\max}(Z, \hat{R}_2^{-1}) \right)$$

where $\ell_1^*(\cdot)$ is defined in (32) and $b(\cdot)$ is defined in (2). $s(\cdot)$ has a maximum point at the inverse of the ML estimator of the BCM

$$\hat{R}(Y) = \arg \max_{R \in S^+_q} \ell_1^*(Y; R) = \left( \arg \max_{Z \in S^+_q} s(Z) \right)^{-1}$$

(54)
By searching for local maximum points, we examine the local behavior of \(s(Z + \epsilon X)\) around \(\epsilon \approx 0\), where \(Z \in S_q^+\) is the tested matrix and \(X \in S^q\) is the direction of perturbation.

To simplify the parameterization of the problem, we define
\[
A_Y \equiv U_Y A_Y^{-\frac{1}{2}}
\]
Let \(\Gamma \in D_q^+\) and \(T \in U^q\). Define
\[
Z(\Gamma, T) = A_Y T^T T^H A_Y^H
\]
Using the same representation with \(\Gamma\) replaced by a non-diagonal matrix \(\Delta \in S^q\), define
\[
X(\Delta, T) = A_Y T \Delta T^H A_Y^H
\]
Note that any pair \(Z \in S_q^+, X \in S^q\) can be represented as \(Z(\Gamma, T), X(\Delta, T)\). Perturbing \(s(\cdot)\) at \(Z(\Gamma, T)\) by \(\epsilon X(\Delta, T)\), we obtain a representation of \(s(\cdot)\) restricted to a line passing through \(S^q_+\), as a function of a single real variable \(\epsilon\), which has \(\Delta, \Gamma, T\) as parameters.

\[
s_0(\epsilon; \Delta, \Gamma, T) \equiv s(Z(\Gamma, T) + \epsilon X(\Delta, T))
\]
\[
= 2\log |\Gamma + \epsilon \Delta| - 2\log |\tilde{R}_{\epsilon}| + h(\lambda_{\text{max}}(\Gamma + \epsilon \Delta)) - \text{tr} \left( T(\Gamma + \epsilon \Delta) T^H (I_q + A_Y^{-1}) \right)
\]
Note that we only consider the open interval of \(\epsilon\) where \((Z(\Gamma, T) + \epsilon X(\Delta, T)) \in S_q^+\).

**Lemma 2.** The function \(s_0(\epsilon; \Delta, \Gamma, T)\) defined in (58) is right differentiable w.r.t. \(\epsilon\) over its domain for any \((\Delta, \Gamma, T)\).

The proof of this lemma is deferred to Appendix C.

Given the fact that \(s_0(\epsilon; \Delta, \Gamma, T)\) is always right-differentiable w.r.t. \(\epsilon\), a sufficient condition for \(Z(\Gamma, T)\) to be a local maximum point of \(s(Z)\) is that for any \(\Delta \in S^q\):
\[
\partial_+ s_0(\Delta, \Gamma, T) \equiv \lim_{\epsilon \to 0^+} \frac{s_0(\epsilon) - s_0(0)}{\epsilon} \leq 0
\]
implies that \(s(Z(\Gamma, T))\) decreases when adding to \(Z(\Gamma, T)\) any incremental matrix of the form \(\epsilon X\). We denote the set of points satisfying this condition by \(A\):
\[
A \equiv \{Z(\Gamma, T) : \partial_+ s_0(\Delta, \Gamma, T) \leq 0 \forall \Delta \in S^q\} \subseteq S_q^+
\]
Define a subset of unitary matrices
\[
U_q^U = \{P \Phi : P \in P^q, \Phi \in U_D\}
\]
In the following lemma we prove that every \(Z \in A\), when represented as \(Z(\Gamma, T)\), has \(T \in U_q^U\).

**Lemma 3.** Let \(s_0(\epsilon; \Delta, \Gamma, T)\) and \(\partial_+ s_0(\Delta, \Gamma, T)\) be defined by (58) and (59). If \(T \in U_q^U, U_q^U\) and the joint eigenvalues of \(\mathcal{Y}\) are distinct, then there exists some \(\Delta \in S^q\) such that \(\partial_+ s_0(\Delta, \Gamma, T) > 0\).

The proof of this lemma is deferred to Appendix D. We conclude that if the joint eigenvalues of \(\mathcal{Y}\) are distinct then
\[
A \subseteq \tilde{A} \equiv \{Z(\Gamma, T) : \Gamma \in D_q^+, T \in U_q^U\}
\]

By definition of \(U_q^U\),
\[
\tilde{A} = \{Z(\Gamma, P \Phi) : \Gamma \in D_q^+, P \in P^q, \Phi \in U_D\}
\]
\[
= \{A_Y P \Gamma P^H A_Y^H : \Gamma \in D_q^+, P \in P^q\}
\]
\[
= \{Z(\Gamma, I_q) : \Gamma \in D_q^+\}
\]
where for the first transition we used \(P \Gamma P^H = \Gamma\), for the second transition we used the fact that \(P \Gamma P^H = \Gamma\) whose diagonal entries are permuted according to \(P\); hence \(\{P \Gamma P^H : \Gamma \in D_q^+, P \in P^q\} = D_q^+\).

If the domain of \(s(\cdot)\) is restricted to \(\tilde{A}\) then (53) can be rewritten as the following function of \(\Gamma\) (the following expression can also be obtained by substituting \(\epsilon = 0\) and \(T = I_q\) into (58))
\[
s(\Gamma) \equiv s(Z(\Gamma, I_q))
\]
\[
= -2\log |\tilde{R}_{\Gamma}^{-1}| - \text{tr} (\Gamma (A_Y^{-1} + I_q)) + h \left( \max_i \Gamma_{ii} \right)
\]

**Lemma 4.** For any \(\mathcal{Y} \in S_q^+ \times S^q_q\) with distinct joint eigenvalues and \(\lambda_{\mathcal{Y},1} = \cdots = \lambda_{\mathcal{Y},q} \geq 1\), the function \(s(\cdot) : D_q^+ \rightarrow \mathbb{R}\), defined in (64) has a unique maximum point at:
\[
\Gamma^* \equiv \text{diag} \left( \lambda_{\mathcal{Y},1}, \frac{2\lambda_{\mathcal{Y},2}}{1 + \lambda_{\mathcal{Y},2}}, \cdots, \frac{2\lambda_{\mathcal{Y},q}}{1 + \lambda_{\mathcal{Y},q}} \right)
\]

This lemma is proved in Appendix E. We see that when its domain is restricted to \(\tilde{A}\), \(s(Z)\) has a single maximum point at \(Z(\Gamma^*, I_q)\). This result eliminates all the other members of \(\tilde{A}\) from being maximum points of \(s(Z)\). If \(s(Z)\) has a maximum point, it must be attained at \(Z(\Gamma^*, I_q)\) and only there, since it is the only candidate left. To prove that \(s(Z)\) has a maximum point, we present a sufficient condition for its existence in Lemma 5 and show in Lemma 6 that this condition is satisfied by \(s(Z)\). These lemmas are proven in Appendices F and G, respectively.

**Lemma 5.** Define the function \(g : S_q^+ \rightarrow \mathbb{R}\) by
\[
g(Z) = \max \{\lambda_{\text{max}}(Z), \lambda_{\text{min}}^{-1}(Z)\}
\]
and define a “divergent matrix sequence” as a sequence of PD matrices \(\{Z_i\}\) such that \(g(Z_i) \rightarrow \infty\). If \(f : S_q^+ \rightarrow \mathbb{R}\) is a continuous function such that \(f(Z_i) \rightarrow (-\infty)\) for every \(\{Z_i\}\) that is a divergent matrix sequence, then \(f\) has a maximum point.

**Lemma 6.** Let \(s, g : S_q^+ \rightarrow \mathbb{R}\) be defined by (53) and (66), respectively. If \(\{Z_i \in S_q^+\}\) is a sequence of matrices such that \(g(Z_i) \rightarrow \infty\) then \(s(Z_i) \rightarrow (-\infty)\).

We see from (54) that \(\tilde{R}(\mathcal{Y})\) is the inverse of \(Z(\Gamma^*, I_q)\), which is now proved to be the maximizer of \(s(Z)\). Using (56) with \(\Gamma = \Gamma^*\) and \(T = I_q\) together with (55) and (22) we have
\[
\tilde{R}(\mathcal{Y}) = Z(\Gamma^*, I_q)^{-1} = V_\mathcal{Y} A_\mathcal{Y} \Gamma^{*-1} V_\mathcal{Y}^H
\]
Inverting \(\Gamma^*\) from (65) and multiplying the result by \(A_\mathcal{Y}\) (both multipliers are diagonal)
\[
A_\mathcal{Y} \Gamma^{*-1} = \frac{1}{2}(A_\mathcal{Y} + I_q) - \frac{1}{2}(\lambda_{\mathcal{Y}} - 1)e_1e_1^H
\]
From the generalized eigen-decomposition of \((\hat{R}_p, \hat{R}_s)\) we have
\[
\hat{R}_s + \hat{R}_p = V_Y(I_q + \Lambda_Y) V_Y^H
\] (69)

Using the identities (25)-(27) the steering vector and power estimators from (6) and (7) can be expressed as 
\[\hat{\delta} = \|V_Y\|^{-1} V_Y \hat{e}_1\] and 
\[\hat{\sigma}^2 = \|V_Y\|^2 (\lambda_Y - 1).\] Combining these representations with the identity \(V_Y = V_Y \hat{e}_1\) we have
\[
\hat{\sigma}^2 \hat{\delta}^H = (\lambda_Y - 1) V_Y \hat{e}_1 \hat{e}_1^H V_Y^H
\] (70)
The expression in (8) is obtained by plugging (68), (69) and (70) into (67).

**APPENDIX C**

**PROOF OF LEMMA 2**

To prove Lemma 2 we note that the function \(s_0(\epsilon; \Delta, \Gamma, T)\), given in (58), depends on \(\epsilon\) through the sum of three terms: 
\[\log(\cdots), h(\max(\cdots))\] and \(\max(\cdots)\) which are identified as the log-determinant, largest eigenvalue and trace terms. We show that each is right differentiable w.r.t. \(\epsilon\):  

1) Substituting \(X = \Gamma + \epsilon \Delta\) and \(\partial X = (de) \Delta\) into equation (43) from [31], we obtain
\[
\frac{d \log |\Gamma + \epsilon \Delta|}{d \epsilon} = \text{tr} \left( (\Gamma + \epsilon \Delta)^{-1} \Delta \right)
\] (71)
This derivative is well-defined over the set of \(\epsilon\) for which \((\Gamma + \epsilon \Delta)\) is invertible.

2) For a given \(\Gamma \in \mathbb{S}^d_{+}, \Delta \in \mathbb{S}^q\) and \(\epsilon \in \mathbb{C}^q\), define the open interval \(E = \{ \epsilon : \Gamma + \epsilon \Delta \in \mathbb{S}^d_{+} \}\), the function \(f_x(\epsilon) \equiv x^H(\Gamma + \epsilon \Delta)x\) which is differentiable w.r.t. \(\epsilon\) and \(f(\epsilon) \equiv \lambda_{\max}(\Gamma + \epsilon \Delta)\). It can easily be shown that since \(f(\epsilon) = \max_{x \in \mathbb{C}^q} f_x(\epsilon)\) for every \(\epsilon \in E\), then \(\partial \epsilon f(\epsilon) \leq \max_{x \in \mathbb{C}^q} |df_x(\epsilon)/d\epsilon|;\) hence \(f(\epsilon)\) is right-differentiable at any \(\epsilon \in E\).

3) The term \(\text{cTR} (\hat{T} \hat{\Delta} \hat{T}^H (I_q + \Lambda_Y^{-1}))\) is linear w.r.t. \(\epsilon\), and hence differentiable.

**APPENDIX D**

**PROOF OF LEMMA 3**

To prove Lemma 3 we need the following lemma, which will be proved in subsection D-A:

**Lemma 7.** Let \(T \in \mathbb{U}^q \setminus \mathbb{U}_q^1\). Define the sets of vectors \(\mathcal{V} \equiv \{ t_{i,j} \in \mathbb{C}^p : t_{i,j}(k) = T_{ki} T_{kj}, i,j \in \{1, \ldots, q\} \text{ and } T \equiv T \setminus \{ t_{i,i}, i \in \{1, \ldots, q\} \}\). If \(x \in \mathbb{C}^q\) has \(q\) distinct entries then there exists at least one member of \(\mathcal{V}\) that is not orthogonal to \(x\).

We now use Lemma 7 to prove Lemma 3. The function \(s_o(\epsilon; \Delta, \Gamma, T)\) is expressed below as the sum of a differentiable part and a second term which is possibly non-differentiable at \(\epsilon = 0\) (but right-differentiable there), which are denoted by \(s_0(\epsilon; \Delta, \Gamma, T)\) and \(s_0(\epsilon; \Delta, \Gamma, T)\), respectively, and defined as follows
\[
s_0(\epsilon; \Delta, \Gamma, T) = 2 \log |\Gamma + \epsilon \Delta| - 2 \log |\hat{R}_p| - \text{tr} \left( T (\Gamma + \epsilon \Delta) \hat{T}^H (I_q + \Lambda_Y^{-1}) \right)
\] (72)
\[
s_0(\epsilon; \Delta, \Gamma, T) = h(\lambda_{\max}(\Gamma + \epsilon \Delta))
\] (73)
\[
s_0(\epsilon; \Delta, \Gamma, T) = s_0(\epsilon; \Delta, \Gamma, T) + s_0(\epsilon; \Delta, \Gamma, T)
\] (74)
Suppose we choose \(\Delta\) whose entries are given by
\[
\Delta_{ij} = z \delta_{i,0} \delta_{j,0} + z^* \delta_{i,0} \delta_{j,0}
\] (75)
Then \(z \neq 0\) and \(i_0 < j_0\). It will be shown that for any \(\Delta\) of this form, the right-derivative of \(s_0(\epsilon; \Delta, \Gamma, T)\) at \(\epsilon = 0\) is non-negative, and that there must exist a choice of \((i_0, j_0)\) and \(z\) such that the right derivative of \(s_0(\epsilon; \Delta, \Gamma, T)\) is positive.

For the derivation of \(\partial \epsilon s_0(\Delta, \Gamma, T)\), we consider the behavior of the eigenvalues of \((\Gamma + \epsilon \Delta)\), which are continuous functions of \(\epsilon\), that are denoted by \(\gamma_1(\cdot), \ldots, \gamma_q(\cdot)\) and ordered such that \(\gamma_1(0) = \gamma_l \equiv \Gamma_{l,l}\) for all \(l\). Due to the diagonality of \(\Gamma\) and the sparse structure of \(\Delta\), when \(\epsilon \neq 0\), \(\gamma_i(\epsilon) = \gamma_i\) for all \(\epsilon\). The two remaining eigenvalues, \(\gamma_{i_0}(\epsilon)\) and \(\gamma_{j_0}(\epsilon)\) can be identified as the solutions of the following determinant equation
\[
\begin{vmatrix}
\gamma_{i_0} - z & 0 \\
0 & \gamma_{j_0}
\end{vmatrix}
= 0
\] (76)
Assuming that the larger of the two solutions of (76) is assigned to \(\gamma_{i_0}(\epsilon)\), then
\[
\gamma_{i_0}(\epsilon) = \left( 1 + \frac{1}{2} (\gamma_{i_0} + \gamma_{j_0} + \sqrt{(\gamma_{i_0} - \gamma_{j_0})^2 + 4 \epsilon^2 |z|^2}) \right)
\] (77)
The right derivative of \(\gamma_{i_0}(\epsilon)\) at \(\epsilon = 0\), denoted by \(\partial \epsilon \gamma_{i_0}(0)\), is \(|z|\) if \(\gamma_{i_0} = \gamma_{j_0}\) and zero otherwise. The right-derivative of the largest eigenvalue of \((\Gamma + \epsilon \Delta)\) at \(\epsilon = 0\) is \(\partial \epsilon \gamma_{i_0}(0)\) when \(\gamma_{i_0}(0) = \gamma_{i_0}\) is max, \(\gamma_i\) and zero otherwise (since then, when \(\epsilon\) is in the vicinity of zero the maximum eigenvalue of \((\Lambda + \epsilon \Delta)\) is independent of \(\epsilon\); hence it is always non-negative. Given this non-negativity and because \(h(\cdot)\) is a non-decreasing function, for any \(\Gamma\) and any choice of \(i_0, j_0\) and \(z\):
\[
\partial \epsilon s_0(\Delta, \Gamma, T) \geq 0
\] (78)
According to (71), the derivative of the log-determinant term of \(s_0(\epsilon; \Delta, \Gamma, T)\) is zero at \(\epsilon = 0\) since the diagonal of \(\Delta\) is zero. The right-derivative of \(s_0(\epsilon; \Delta, \Gamma, T)\) at \(\epsilon = 0\) is
\[
\partial \epsilon \hat{s}_0(\Delta, \Gamma, T) = \frac{d \hat{s}_0(\epsilon; \Delta, \Gamma, T)}{d \epsilon} = \text{tr} \left( T \hat{\Delta} T^H (I_q + \Lambda_Y^{-1}) \right)
\] (79)
where \(e_{ij}(k) = 1 + \lambda_{ij}^{-1}\) and \(t_{i,j}(k) = T_{ki} T_{kj}\). It is given that \(T \in \mathbb{U}^q \setminus \mathbb{U}_q^1\) and that the \(q\) joint eigenvalues of \(\Sigma\) are distinct, which implies that \(e_{ij}\) has distinct entries; hence according to Lemma 7, there must exist an index pair \((i_* ,j_*)\) such that \(e_{ij}(0) \neq 0\) and \(i_* < j_*\). Setting \(\Delta\) according to (75) with \((i_0, j_0) = (i_*, j_*)\) and \(z = e_{ij}(0)\), we obtain from (79) that
\[
\partial \epsilon \hat{s}_0(\Delta, \Gamma, T) = 2 |t_{i_0, j_0} e_{ij}(0) |^2 > 0
\] (80)
Adding the inequalities (78) and (80) we have from (74) that for any \( \Gamma \in D^q_{+} \) and \( T \in U^n \backslash U^q \) there must exist \( \Delta \in S^q \) for which \( \partial_{q+}(\Delta, \Gamma, T) > 0 \).

A. Proof of Lemma 7

To prove Lemma 7 we need the following lemma which will be proved in subsection D-B:

**Lemma 8.** If \( T \in U^n \backslash U^q \), \( x = (x_1, \ldots, x_q)^T \in \mathbb{R}^q \) is a vector with distinct entries (\( x_i \neq x_j \) for any \( i \neq j \)) and \( A \) is such that \( A_{ij} = |T_{ij}|^2 \) for all \( i, j = 1, \ldots, q \) then \( \|Ax\| < \|x\| \).

We now use Lemma 8 to prove Lemma 7. Define the following scalar function of the given unitary matrix \( T \) and a vector \( x \): \( \sigma(T, x) \equiv \sum_{i=1}^{q} \sum_{j \neq i} x_i x_i^H \). A statement equivalent to Lemma 7 is that for all \( T \in U^n \backslash U^q \) and \( x \) with distinct entries, \( \sigma(T, x) \neq 0 \).

We begin by representing \( \sigma(T, x) \) using a simple quadratic form. Based on \( T \), we define a set of \( q \) vectors \( \{t_i \in C^q : t_i(j) = T_{ij}\} \), a matrix \( A \equiv (t_{i1}, \ldots, t_{iq}) \) and two sets of \( q \) matrices \( \{Y(i) = \text{diag}(T_{i1}, \ldots, T_{iq})\} \) and \( \{S(i) \in C^{q \times (q - 1)} \} \) where the \( i \)th matrix of this set is given by

\[
S(i) = \begin{pmatrix} t_{i1} & \cdots & t_{i,i-1} & t_{i,i+1} & \cdots & t_{iq} \end{pmatrix} \tag{81}
\]

Using the following identities

\[
\sum_{i=1}^{q} t_{i,ij} t_{i,j}^H = \sum_{i=1}^{q} Y(i) \sum_{i=1}^{q} t_{i,j} t_{i,j}^H Y(i)^H = I_q \tag{82}
\]

\[
\sum_{i=1}^{q} S(i) S(i)^H = \sum_{i=1}^{q} \left( \sum_{i=1}^{q} t_{i,ij} t_{i,j}^H - t_{i,i} t_{i,i}^H \right) = I_q - AA^H \tag{83}
\]

we express \( \sigma(T, x) \) as

\[
\sigma(T, x) = \sum_{i=1}^{q} x_i^H S(i) S(i)^H x = x_i^H (I_q - AA^H) x \tag{84}
\]

According to Lemma 8, since \( A_{ij} = |T_{ij}|^2 \) for all \( i, j \), if \( T \notin U^q \) then the right-hand expression in (84), which is equal to \( \sigma(T, x) \), cannot be zero for \( x \) with distinct entries.

B. Proof of Lemma 8

Assume that \( x = (x_1, \ldots, x_q)^T \in C^q \) is a vector with distinct entries. The lemma is proved by contradiction, we show that \( \|Ax\| = \|x\| \) implies \( T \notin U^q \).

By the definition of \( A \): \( A_{ij} \geq 0 \) and \( A_{1q} = A_{1q}^T = 1_q \), hence \( A \) is a doubly stochastic matrix, and by \( A_{1q} A_{1q}^T = A_{1q}^T 1_q = 1_q \) it is proved that \( B \equiv A^T A \) is a symmetric doubly stochastic matrix. Given the fact that \( A \) is stochastic, it follows from the Gershgorin circle theorem [32, p. 320] that \( \rho(A) \leq 1 \) where \( \rho(A) \equiv \max \{ |\lambda_1(A)|, \ldots, |\lambda_q(A)| \} \) is the spectral radius of \( A \). Because \( A_{1q} = 1_q \), we have that \( \lambda_{\max}(A) = \rho(A) = 1 \). The same is true for \( B \); i.e., \( \lambda_{\max}(B) = \rho(B) = 1 \).

We now prove by induction that \( Bx = x \) implies \( B = I_q \).

In the \( n \)th step of the proof, it is shown that the \( n \)th column and row of \( B \) are \( e_n \) and \( e_n^T \), respectively. It is assumed, without loss of generality that \( x_1 > \cdots > x_q \).

Base case: For the first element of \( Bx \) we can write

\[
e_{1}^H Bx = \sum_{j=1}^{q} B_{1,j} x_j = x_1 = \max x_i \tag{85}
\]

This equation represents \( x_1 \) as a convex combination of all the elements of \( x \), which is bounded by the extreme values of \( x \). Equation (85) is satisfied only if \( B_{1,j} = \delta_{1j} \). Since \( B \) is symmetric, the first column is also described by \( B_{1,j} = \delta_{1j} \) and the induction statement is proved for \( n = 1 \).

Inductive step: Suppose that for some \( n \geq 1 \), \( B_{n+1,1} = \cdots = B_{n+1,n} = 0 \). The \((n + 1)\)th element of \( Bx \) can be written as

\[
e_{n+1}^H Bx = \sum_{j=1}^{q} B_{n+1,j} x_j = \sum_{j=n+1}^{q} B_{n+1,j} x_j = x_{n+1} = \max\{x_{n+1}, \ldots, x_q\} \tag{86}
\]

The same logic that was used for the base case, we conclude that \( B_{n+1,j} = \delta_{n+1,j} \), and thus we prove that if \( Bx = x \) then for any \( n \), \( B_{n,j} = \delta_{n,j} \) or simply \( B = I_q \).

It is implied by \( B = A^T A = I_q \) that \( A \) is an orthogonal matrix, hence for every \( i: \sum_{j=1}^{q} (A_{ij})^2 = \sum_{j=1}^{q} (A_{ji})^2 = \sum_{j=1}^{q} A_{ij} = 1_j \), which is satisfied if and only if \( A \in P^n \) (recalling that the entries of \( A \) are non-negative by definition). \( A \) can be a permutation matrix if and only if \( T = A \Phi \) where \( \Phi \in U^n_{+} \); hence by definition of \( U^n_{+} \), \( \|Ax\| = \|x\| \) \( \Rightarrow Bx = x \) \( \Rightarrow B = I_q \) \( \Rightarrow A \in P^n \Rightarrow T \in U^n_{+} \), so that \( T \in U^n \backslash U^q \Rightarrow \|Ax\| < \|x\| \).

**APPENDIX E**

**PROOF OF LEMMA 4**

We define a set of functions \( \{f^{(i)} : D^q_{+} \rightarrow \mathbb{R}, i = 1, \ldots, q\} \)

\[
f^{(i)}(\Gamma) = 2 \log |\Gamma| - \text{tr} \left( \Gamma (I_q + A_{\gamma}^{-1}) \right) + h(\gamma_i) \hspace{1em} i = 1, \ldots, q \tag{87}
\]

where we denote \( \gamma_i \equiv \Gamma_{ii} \), and discard the dependence on \( \gamma \), to conform with the notation used in subsection IV-C. It follows from the fact that \( h(\cdot) \) is a non-decreasing function that \( h(\max \gamma_i) = \max h(\gamma_i) \) so that \( \hat{s}^{(i)}(\Gamma) \) from (64) can be written as:

\[
\hat{s}^{(i)}(\Gamma) = \max_i f^{(i)}(\Gamma) - 2 \log |R_{\gamma}|. \tag{88}
\]

If \( \Gamma^{(i)} \) is the unique maximum of \( f^{(i)}(\Gamma) \) and the set of maximal values \( \{f^{(i)}(\Gamma^{*}) \}, i = 1, \ldots, q \) has a unique maximum whose index is denoted by \( i \), then \( \Gamma^{(i)} \) is the unique maximum of \( \hat{s}^{(i)}(\Gamma) \).

We decompose each \( f^{(i)}(\Gamma) \) into the sum of \( q \) functions, each depending on a single diagonal element of \( \Gamma \) and a single \( \lambda_{\gamma,i} \)

\[
f^{(i)}(\Gamma) = \sum_{j=1}^{q} f^{(i,j)}(\gamma_j; \lambda_{\gamma,j}) \tag{89}
\]

where \( \bigwedge \) and \( \vee \) denotes the logic “and” and “or”, respectively. Using straightforward algebra, we derive the unique maximum
point of each \( f^{(i,j)}(\gamma; \lambda_{Y,j}) \) w.r.t. \( \gamma \) (using the fact that \( \lambda_{Y,j} > 0 \)):

\[
\gamma^*_j = \arg \max_{\gamma > 0} f^{(i,j)}(\gamma; \lambda_{Y,j}) = \frac{\lambda_{Y,j}}{1 + \lambda_{Y,j}} \quad (i = j) \land (\lambda_{Y,j} \geq 1) \\
\frac{2\lambda_{Y,j}}{1 + \lambda_{Y,j}} - 2 \quad (i \neq j) \lor (\lambda_{Y,j} < 1)
\]

with the maximal value of \( f^{(i,j)}(\gamma; \lambda_{Y,j}) \) given by

\[
f^{(i,j)}(\gamma) = \frac{2\lambda_{Y,j}}{1 + \lambda_{Y,j}} - 2 \quad (i \neq j) \lor (\lambda_{Y,j} < 1)
\]

The sum in (88) is maximized when each of the summands is maximized; hence the \( j \)th diagonal element of the unique maximizer of \( f^{(i,j)}(\Gamma) \) is \( \Gamma^j_{jj} = \gamma^*_j \). The corresponding maximal value is

\[
f^*_j = \max_{\Gamma \in \Delta_i} f^{(i)}(\Gamma) = \sum_{j=1}^q f^{(i,j)}(\gamma^*_j) = \frac{2q\lambda_{Y,j}}{1 + \lambda_{Y,j}} - 2q \quad (\lambda_{Y,j} > 0)
\]

It follows from the fact that the additional term in (92) that depends on \( \lambda_{Y,j} \) is an increasing function and strictly increasing when \( \lambda_{Y,j} > 1 \) and \( \lambda_{Y,j} \equiv \max_j \lambda_{Y,j} > 1 \) that \( f^{(i)} > f^{(j)} \) for all \( j > 1 \); hence \( i = 1 \) which implies \( \Gamma^*_i = \Gamma^1 \). The \( j \)th diagonal element of \( \Gamma^*_i \), as given in (65) is obtained by plugging \( i = 1 \) into \( \gamma^*_i \) from (90).

**APPENDIX F**

**PROOF OF LEMMA 5**

Let \( f : S^q_+ \to \mathbb{R} \) be continuous with \( f(Z_i) \to -\infty \) for any divergent matrix sequence \( \{Z_i\} \). Define a sequence of closed bounded sets of PD matrices \( G_i = \{Z \in S^q_+ : i - 1 < g(Z) \leq i\} \) for \( i = 1, 2, \ldots, \infty \). According to the extreme value theorem, because \( f \) is continuous, for every \( i \) there exists a matrix \( Z_i \in G_i \) such that \( f_i \equiv f(Z_i) = \max_{Z \in G_i} f(Z) \). By definition of \( G_i \), \( Z_i \) is a divergent matrix sequence; hence by definition of \( f_i \), \( f_i \to -\infty \) and therefore the sequence \( \{f_i\} \) has a maximal value \( f_* = \max_i f_i = \max_{Z \in G_i} f(Z) \). Because the infinite union of \( G_i \) is \( S^q_+ \), \( f_* \) is the global maximum of \( f \).

**APPENDIX G**

**PROOF OF LEMMA 6**

The following lemma is essential for the proof of lemma 6:

**Lemma 9.** Let \( \{Z_i \in S^q_+\} \), \( \{\Gamma_i \in \Delta_i^q\} \), \( \{T_i \in U^q\} \) and \( \mathcal{Y} \in S^q_+ \times S^q_+ \) be such that for every \( i \), \( T_i = Z(\Gamma_i, T_i) \), where the representation \( Z(\cdot, \cdot) \) is given in (56). A sufficient and necessary condition for \( g(Z_i) \to -\infty \) is that \( g(\Gamma_i) \to -\infty \).

**Proof:** Using the inequalities \( \lambda_{\max}(ABA^H) \leq \lambda_{\max}(B)\lambda_{\max}(A^H) \) and \( \lambda_{\min}(ABA^H) \geq \lambda_{\min}(B)\lambda_{\min}(A^H) \) which hold for any \( B \in S^q_+ \) and \( A \in \mathbb{F}^q \), the ratio of the smallest and the largest eigenvalues of \( Z_i \) to the corresponding eigenvalues of \( \Gamma_i \) can be bounded as follows

\[
\frac{1}{\lambda_{\min}(A^H)\lambda_{\max}(A^H)} \leq \frac{\lambda_{\max}(Z_i)}{\lambda_{\min}(\Gamma_i)} \leq \frac{1}{\lambda_{\min}(A^H)\lambda_{\max}(A^H)}
\]

It is implied from these bounds that \( \lambda_{\min}(Z_i) \to \infty \) if and only if \( \lambda_{\max}(\Gamma_i) \to \infty \) and that \( \lambda_{\min}(Z_i) \to 0 \) if and only if \( \lambda_{\min}(\Gamma_i) \to 0 \); hence \( g(Z_i) \to -\infty \) if and only if \( g(\Gamma_i) \to -\infty \).

Using Lemma 9 we now prove Lemma 6. Substituting \( \epsilon = 0 \) into (58) and omitting the constant term \( -2 \log |R_\epsilon| - 1 \) we obtain \( s(Z) \) in terms of the diagonal entries of \( \Gamma \), denoted by \( \gamma_1 \geq \cdots \geq \gamma_q \), of \( T \) and of \( A^Y \):

\[
s(Z(\Gamma, T)) = 2\log |\Gamma| + \gamma_1 - \log \gamma_1 - \text{tr} (\Gamma + TTT^H A^H_Y^{-1}) = \log \gamma_1 + \sum_{i=2}^q (2\log \gamma_i - \gamma_i) - \text{tr} (TTT^H A^H_Y^{-1})
\]

Defining for \( x > 0 \) the functions \( r_1(x) \equiv 2\log x - x \) and \( r_2(x) \equiv \log x - \lambda_{Y,j} x \) with upper bounds given by \( r_1(2) \equiv r_2(1) \) and \( r_1(1) \equiv r_2(2) \) with \( r_1(1) \equiv r_2(2) \). The terms of (93) have the following upper bounds:

\[
\sum_{i=2}^q r_1(\gamma_i) \leq r_1(\gamma_1) + (q-2)r_1 \leq (q-1)r_1
\]

\[
\log \gamma_i - \text{tr} (TTT^H A^H_Y^{-1}) \leq r_2(\gamma_i) \leq r_2(1)
\]

where (95) follows from:

\[
\text{tr} (TTT^H A^H_Y^{-1}) \geq \gamma_1 \text{tr}(\Gamma_Y)
\]

where we define the constants \( C_2 \equiv (q-2)\gamma_1 + 2 \gamma_1 \). Let \( \{Z_i \in S^q_+\}, \{\Gamma_i \in \text{diag}(\gamma_1, \ldots, \gamma_q) \in \Delta_i^q\}, \{T_i \in U^q\} \) be matrix sequences such that for every \( i \), \( Z_i = Z(\Gamma_i, T_i) \), \( \gamma_1, \ldots, \gamma_q \geq \gamma_1 = \gamma_0 \), and \( g(Z_i) \to \infty \). According to Lemma 9, it is implied by \( g(Z_i) \to \infty \) that 

\[
\lambda_{\min}(\Gamma_i, Y_i) \to \infty \quad \text{and} \quad \min(\gamma_i, 1) \to \infty.
\]

**REFERENCES**


