

This is a postprint version of the following published document:

Horstmann, Stefanie; Ramírez, David; Schreier, Peter J. (2020). Two-Channel Passive Detection of Cyclostationary Signals. *IEEE Transactions on Signal Processing*, v. 68, pp.: 2340 - 2355.

DOI: <https://doi.org/10.1109/TSP.2020.2981767>

©2020 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

# Two-Channel Passive Detection of Cyclostationary Signals

Stefanie Horstmann, *Student Member, IEEE*, David Ramírez, *Senior Member, IEEE*, and Peter J. Schreier, *Senior Member, IEEE*

**Abstract**—This paper considers passive detection of a cyclostationary signal in two multiple-input multiple-output (MIMO) channels. The passive detection system consists of an illuminator of opportunity (IO), a reference array, and a surveillance array, each equipped with multiple antennas. As common transmission signals of the IO are cyclostationary, our goal is to detect the presence of cyclostationarity at the surveillance array, given observations from both channels. To this end, we analyze the existence of optimal invariant tests, and we derive an alternative and more insightful expression for a previously proposed generalized likelihood ratio test (GLRT). Since we show that neither the uniformly most powerful invariant test (UMPIT) nor the locally most powerful invariant test (LMPIT) exist, we propose an LMPIT-inspired detector that is given by a function of the cyclic cross-power spectral density. We show that the LMPIT-inspired detector outperforms the GLRT, and both detectors outperform state-of-the-art techniques.

**Index Terms**—Cyclostationarity, generalized likelihood ratio test (GLRT), locally most powerful invariant test (LMPIT), multiple-input multiple-output (MIMO) passive detection

## I. INTRODUCTION

**I**N this work, we consider a multiple-input multiple-output (MIMO) passive bistatic radar system. Such systems are of special interest as they are simple, cheap, and undetectable because the transmitter is not part of the system [1]. A passive bistatic radar system consists of one receiver and one non-cooperative transmitter, which is referred to as an illuminator of opportunity (IO). The passive radar receives a direct-path signal, which is a noisy version of the transmitted signals from the IO, and a target-path signal, which is the echo from the target if it is present, or only noise, otherwise. In order to separate these two signals at the receiver, either directional antennas [2], digital beamforming [3], [4] or both could be employed. The target-path signal may also be corrupted with direct-path and clutter components. Given a strong direct-path signal in the reference channel, techniques to cancel these kinds of interferences are presented in, e.g., [5], [6]. Typically,

S. Horstmann and P. J. Schreier were supported by the German Research Foundation (DFG) under grant SCHR 1384/6-1. The work of D. Ramírez was supported by the Ministerio de Ciencia, Innovación y Universidades under grant TEC2017-92552-EXP (aMBITION), by the Ministerio de Ciencia, Innovación y Universidades, jointly with the European Commission (ERDF), under grant TEC2017-86921-C2-2-R (CAIMAN), by The Comunidad de Madrid under grant Y2018/TCS-4705 (PRACTICO-CM), and by the German Research Foundation (DFG) under grant RA 2662/2-1.

S. Horstmann and P. J. Schreier are with the Signal and System Theory Group, University of Paderborn, 33098 Paderborn, Germany (e-mail: stefanie.horstmann@sst.upb.de, peter.schreier@sst.upb.de)

D. Ramírez is with the Department of Signal Theory and Communications, Universidad Carlos III de Madrid, Spain and the Gregorio Marañón Health Research Institute, Madrid, Spain (e-mail: david.ramirez@uc3m.es)

the IO is a commercial radio or TV broadcast system, or it could be a space-based source such as communication or navigation satellites [7]–[9].

Various techniques have been derived to detect the presence of the target echo at the surveillance channel (SC) assuming that the transmission signal is temporally white. The most common approach is based on cross-correlating the signals at SC and reference channel (RC), e.g., [5], [10]–[14]. Although this resembles the matched filter, it is suboptimal due to noise at the RC [14]. In [15]–[18] generalized likelihood ratio tests (GLRT) were derived for the case of unknown stochastic waveforms and for various assumptions on the signal and noise models. Reference [15] considered the detection of a rank-one signal received by a multiantenna array, whereas [16] generalized these results to a rank- $p$  signal. These detectors assume that the noise has an arbitrary spatial correlation. The GLRT for spatially white noise with the same variance at SC and RC was derived in [17]. Finally, [18] extended the results to the detection of a rank- $p$  signal in white noise with different variances at SC and RC and spatially uncorrelated noise with arbitrary variances. The GLRTs for the case of unknown deterministic waveforms in temporally and spatially white noise were presented in [19] and [20], where [19] assumed unknown and [20] assumed known noise variance. For the same problem, an approximate Bayesian test was derived in [19] and the exact Bayesian test was presented in [21]. The work in [22] proposed an ad-hoc detector based on the generalized coherence [23].

However, all these aforementioned detectors do not exploit the fact that digital communication signals transmitted by potential IOs are cyclostationary [24]. For single array detection this property was exploited in [25], [26], which derived locally optimum tests for a known signal waveform and different assumptions on the noise. In [25] temporally and spatially white Gaussian noise was considered, whereas [26] considered non-Gaussian noise. The GLRT and locally most powerful invariant test (LMPIT) for detecting an unknown cyclostationary signal with a single array in temporally and spatially correlated noise was derived in [27] and specialized to various noise structures in [28], [29].

## A. Contributions

In this work, we solve the two-channel passive detection problem by exploiting cyclostationarity. This aims at detecting the presence of cyclostationarity at the SC given the additional reference channel. We evaluate the performance of the proposed detectors with Monte Carlo simulations and show that

they outperform existing tests. The main contributions can be summarized as follows:

- 1) We derive an alternative and more insightful expression for the GLRT, which we have previously proposed in [30]. Deriving the GLRT requires the maximum likelihood (ML) estimates of the covariance matrices, which have a block-Toeplitz structure. Since there exists no closed-form solution for (block) Toeplitz covariance matrices, we use an asymptotic result from [27], which allows us to obtain approximate closed-form ML estimates of the covariance matrices under both hypotheses. Moreover, we show that the distribution under the null hypothesis can be asymptotically approximated by the distribution of the product of independent beta random variables.
- 2) We examine the existence of the uniformly most powerful invariant test (UMPIT) and the LMPIT. In order to do so, we exploit Wijsman's theorem [31], which avoids the necessity of deriving the maximal invariant statistic and its distribution under both hypotheses. This approach has already been applied in, e.g., [27], [29]. We show that neither UMPIT nor the LMPIT exist. However, based on these derivations we are able to propose an LMPIT-inspired detector.
- 3) We provide an interpretation of the two proposed test statistics. The generalized likelihood ratio (GLR) is a function of a coherence matrix (that accounts for the spectral correlation at the SC) and a function of the cross-coherence matrix (that accounts for the cross-spectral correlation between the SC and the RC). Our proposed LMPIT-inspired detector only depends on the latter.

### B. Outline

The detection problem is formulated in Section II followed by the derivation of the GLRT in Section III. In Section IV we examine the existence of the LMPIT, and in Section V we provide an interpretation of the statistics. In Section VI we propose the LMPIT-inspired detector. Finally, the performance of the GLRT and the LMPIT-inspired tests is numerically evaluated with Monte Carlo simulations in Section VII.

### C. Notation

In this paper  $\mathbf{A} \in \mathbb{C}^{M \times N}$  denotes a complex-valued matrix of dimension  $M \times N$ ,  $\mathbf{u} \in \mathbb{C}^M$  denotes a complex-valued vector of dimension  $M$ , and  $(\cdot)^T$  and  $(\cdot)^H$  denote the transpose and Hermitian transpose, respectively. Light-face lower case letters indicate scalars. Furthermore, the trace, determinant, and Frobenius norm of a matrix are denoted by  $\text{tr}(\cdot)$ ,  $\det(\cdot)$ , and  $\|\cdot\|^2$ , respectively. The operator  $\text{vec}(\mathbf{A})$  takes the column-wise vectorization of matrix  $\mathbf{A}$  and  $\text{diag}_M(\mathbf{A})$  is the block-diagonal matrix with block size  $M$  obtained from the  $M \times M$  main diagonal blocks of  $\mathbf{A}$ . The square root matrix is denoted by  $\mathbf{A}^{1/2}$  and the identity matrix of dimension  $N \times N$  by  $\mathbf{I}_N$ . The set of block-diagonal Hermitian matrices of size  $N \times N$  with block size  $M \times M$  is written as  $\mathbb{S}_M^N$  and similarly the set of Hermitian block-Toeplitz matrices of size

$N \times N$  with block size  $M \times M$  as  $\mathbb{T}_M^N$ . The subscript  $k$  in  $\mathbf{B}_k$  denotes the  $k$ th block on the main diagonal of  $\mathbf{B}$  of the appropriate dimensions. Moreover, the superscripts  $(i, j)$  indicate the  $(i, j)$ th possibly matrix-valued element in  $\mathbf{B}_k$ . The corresponding dimensions are given in the context. We denote the Kronecker product of two matrices by  $\otimes$ ,  $*$  denotes the convolution, and  $\propto$  indicates equality up to data-independent positive multiplicative and additive terms. Finally,  $\mathbf{x} \sim \mathcal{CN}(\boldsymbol{\mu}, \mathbf{R})$  stands for a proper complex Gaussian-distributed vector  $\mathbf{x}$  with mean  $\boldsymbol{\mu}$  and covariance matrix  $\mathbf{R}$ .

## II. PROBLEM FORMULATION

We consider a passive bistatic radar setup, in which there are an RC and an SC. Without loss of generality, we assume that each array is equipped with  $L$  antennas.<sup>1</sup> Furthermore, we assume that the IO is equipped with  $L_I$  antennas, and a noisy version of its transmitted signal is received at the RC. The cancellation of interference and clutter in the RC has been considered in e.g. [32], [12]. If there is a target present, the echo of the transmitted signal is observed at the SC. If there is no target present, only noise is received at the surveillance array. Hence, we assume that there is no clutter, interference, or direct-path signal present in the SC, which is achieved by either physical shielding [33] or cancellation by signal processing techniques presented in e.g. [4]–[6], [34]. The complete cancellation of direct-path interference in the SC is, admittedly, an idealized assumption as was pointed out in [4] and the works in [14], [35], [36] have considered the direct-path interference in their signal models. Furthermore, we restrict our attention to the true velocity of the target corresponding to a Doppler shift, which allows us to assume that the target echo observed at the SC is synchronized to the reference signal [18], [20], [36]. The time-delay of the target echo is inherently accounted for in the frequency-selective channel, which we assume in our signal model in the following paragraph. Moreover, considering that direct-path interference has zero Doppler-shift as opposed to the target path signal, it can be filtered [37]. Thus, taking into account the aforementioned assumptions, the passive radar system considered in this paper is illustrated in Figure 1 and the detection problem can be formulated as

$$\begin{aligned} \mathcal{H}_0 : & \begin{cases} \mathbf{u}_s[n] = \mathbf{v}_s[n], \\ \mathbf{u}_r[n] = \mathbf{H}_r[n] * \mathbf{s}[n] + \mathbf{v}_r[n], \end{cases} \\ \mathcal{H}_1 : & \begin{cases} \mathbf{u}_s[n] = \mathbf{H}_s[n] * \mathbf{s}[n] + \mathbf{v}_s[n], \\ \mathbf{u}_r[n] = \mathbf{H}_r[n] * \mathbf{s}[n] + \mathbf{v}_r[n], \end{cases} \end{aligned} \quad (1)$$

for  $n = 0, \dots, NP - 1$  and where  $\mathbf{H}_s[n] \in \mathbb{C}^{L \times L_I}$  and  $\mathbf{H}_r[n] \in \mathbb{C}^{L \times L_I}$  represent the time-invariant frequency-selective channels from the IO to the reference and surveillance arrays, respectively. The additive noise terms  $\mathbf{v}_s[n] \in \mathbb{C}^L$  and  $\mathbf{v}_r[n] \in \mathbb{C}^L$  are assumed to be wide-sense stationary (WSS) with arbitrary temporal and spatial correlation, but they are assumed to be uncorrelated between reference and surveillance arrays. The signal  $\mathbf{s}[n] \in \mathbb{C}^{L_I}$  transmitted by

<sup>1</sup>Note that the derivations can easily be generalized to different numbers of antennas at both arrays.

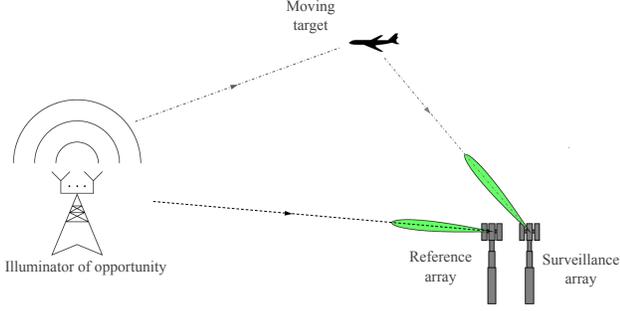


Fig. 1. MIMO passive bistatic radar system that consists of an IO, a reference, and a surveillance array. The reference array receives the direct-path signal from the IO illustrated by the black dashed line and in the presence of a moving target the surveillance array receives the target-path signal, which is depicted by the gray dashed dotted line.

the IO is assumed to be a discrete-time zero-mean second-order cyclostationary (CS) signal with cycle period  $P$ , i.e., its matrix-valued covariance sequence  $\mathbf{R}_{s_s}[n, m] = \mathbb{E}[s[n]\mathbf{s}^H[n-m]] = \mathbf{R}_{s_s}[n+P, m]$  is periodic in  $n$  with period  $P$ . Since the transmitted signal  $s[n]$  is CS, the signal  $\mathbf{u}_r[n] \in \mathbb{C}^L$  received at the reference array is a multivariate CS process with cycle period  $P$  under both hypotheses, whereas the signal  $\mathbf{u}_s[n] \in \mathbb{C}^L$  received at the surveillance array is WSS under  $\mathcal{H}_0$  and CS with cycle period  $P$  under  $\mathcal{H}_1$ . As the cycle period is related to signal features such as carrier frequency, symbol rate, or, for instance, cyclic prefix length, which are known by the standards used by the IO, we can assume that the cycle period  $P$  is known a priori. If this is not the case, the cycle period may be estimated with techniques presented in, e.g., [38]–[40]. Moreover, we assume that  $L_I \geq L$ , which implies that the cyclic (cross) power spectral densities (PSD) of  $\mathbf{H}_s[n] * s[n]$  and  $\mathbf{H}_r[n] * s[n]$  have full rank  $L$ . We make this assumption because the low-rank case would impose additional structure that is not considered in this work.

In order to formulate the hypothesis test, let us consider the vector-valued process

$$\mathbf{x}[n] = [\mathbf{u}^T[nP] \cdots \mathbf{u}^T[(n+1)P-1]]^T \in \mathbb{C}^{LP}, \quad (2)$$

which is WSS if the  $L$ -variate process  $\mathbf{u}[n] \in \mathbb{C}^L$  is CS with cycle period  $P$  [41]. This implies that its matrix-valued covariance function  $\mathbf{R}_{x_x}[n, m] = \mathbb{E}[\mathbf{x}[n]\mathbf{x}^H[n-m]] = \mathbf{R}_{x_x}[m]$  only depends on the time-shift. Moreover, the stack of  $N$  observations  $\mathbf{w} = [\mathbf{x}^T[0] \cdots \mathbf{x}^T[N-1]]^T \in \mathbb{C}^{LNP}$  has a block-Toeplitz structured covariance matrix with block size  $LP$ :

$$\begin{aligned} \mathbf{R}_{w_w} &= \mathbb{E}[\mathbf{w}\mathbf{w}^H] \\ &= \begin{bmatrix} \mathbf{R}_{x_x}[0] & \cdots & \mathbf{R}_{x_x}[N-1] \\ \vdots & \ddots & \vdots \\ \mathbf{R}_{x_x}^H[N-1] & \cdots & \mathbf{R}_{x_x}[0] \end{bmatrix} \in \mathbb{T}_{LP}^{LNP}. \end{aligned} \quad (3)$$

Exploiting the latter considerations we observe that the stack of  $NP$  samples of  $\mathbf{u}_r[n]$

$$\mathbf{w}_r = [\mathbf{u}_r^T[0] \cdots \mathbf{u}_r^T[NP-1]]^T \in \mathbb{C}^{LNP}, \quad (4)$$

has covariance matrix  $\mathbf{R}_r = \mathbb{E}[\mathbf{w}_r\mathbf{w}_r^H]$ , which is a block-Toeplitz matrix with block size  $LP$  under both hypotheses since the signal  $\mathbf{u}_r[n]$  is CS with cycle period  $P$  regardless of the hypothesis. On the other hand, the stack of observations of  $\mathbf{u}_s[n]$

$$\mathbf{w}_s = [\mathbf{u}_s^T[0] \cdots \mathbf{u}_s^T[NP-1]]^T \in \mathbb{C}^{LNP}, \quad (5)$$

has a block-Toeplitz structured covariance matrix  $\mathbf{R}_s^{(0)} = \mathbb{E}[\mathbf{w}_s\mathbf{w}_s^H|\mathcal{H}_0]$  with block size  $L$  under the null hypothesis, where  $\mathbf{u}_s[n] \in \mathbb{C}^L$  is WSS, and covariance matrix  $\mathbf{R}_s^{(1)} = \mathbb{E}[\mathbf{w}_s\mathbf{w}_s^H|\mathcal{H}_1]$ , which is block-Toeplitz with block size  $LP$  under the alternative, where  $\mathbf{u}_s[n]$  is CS with cycle period  $P$ . Moreover, we stack the observations from SC and RC into one long vector

$$\mathbf{w} = [\mathbf{w}_s^T \quad \mathbf{w}_r^T]^T \in \mathbb{C}^{2LNP}. \quad (6)$$

Now let us investigate the structure of the covariance matrix of  $\mathbf{w}$  under both hypotheses. Since the vectors  $\mathbf{w}_s$  and  $\mathbf{w}_r$  are uncorrelated under the null hypothesis, the covariance matrix will simply be a  $2 \times 2$  block-diagonal matrix wherein the covariance matrices of  $\mathbf{w}_s$  and  $\mathbf{w}_r$  are the first and second blocks on the main diagonal, respectively,

$$\mathbf{R}_0 = \mathbb{E}[\mathbf{w}\mathbf{w}^H|\mathcal{H}_0] = \begin{bmatrix} \mathbf{R}_s^{(0)} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_r \end{bmatrix}. \quad (7)$$

The covariance matrix of  $\mathbf{w}$  under the alternative becomes more involved as  $\mathbf{w}_s$  and  $\mathbf{w}_r$  are correlated:

$$\mathbf{R}_1 = \mathbb{E}[\mathbf{w}\mathbf{w}^H|\mathcal{H}_1] = \begin{bmatrix} \mathbf{R}_s^{(1)} & \mathbf{R}_{sr} \\ \mathbf{R}_{rs} & \mathbf{R}_r \end{bmatrix}, \quad (8)$$

where  $\mathbf{R}_{sr} = \mathbf{R}_{rs}^H = \mathbb{E}[\mathbf{w}_s\mathbf{w}_r^H|\mathcal{H}_1]$  is the cross-covariance matrix of  $\mathbf{w}_s$  and  $\mathbf{w}_r$ , which is a block-Toeplitz matrix with block size  $LP$  since the matrix-valued cross-covariance sequence of  $\mathbf{u}_s[n]$  and  $\mathbf{u}_r[n]$  is also periodic with period  $P$ . Thus, all of the matrices  $\mathbf{R}_s^{(1)}$ ,  $\mathbf{R}_r$ , and  $\mathbf{R}_{sr}$  are block-Toeplitz matrices with block size  $LP$ . Assuming that  $\mathbf{u}_s[n]$  and  $\mathbf{u}_r[n]$  are zero-mean proper complex Gaussian random processes, we can formulate the hypothesis test as

$$\begin{aligned} \mathcal{H}_0 : \mathbf{w} &\sim \mathcal{CN}_{2LNP}(\mathbf{0}, \mathbf{R}_0), \\ \mathcal{H}_1 : \mathbf{w} &\sim \mathcal{CN}_{2LNP}(\mathbf{0}, \mathbf{R}_1). \end{aligned} \quad (9)$$

As  $\mathbf{R}_0$  and  $\mathbf{R}_1$  are unknown, (9) is a composite hypothesis test, which is typically approached by a GLRT, a UMPIT, or an LMPIT. The block-Toeplitz structure of the covariance matrices precludes the derivation of the aforementioned detectors. This is because there is no closed-form for the ML estimate of block-Toeplitz covariance matrices and they do not have the necessary invariances for the existence of the UMPIT or the LMPIT. To overcome this issue, we follow an approach similar to [27], where it is shown that we can asymptotically ( $N \rightarrow \infty$ ) approximate a block-Toeplitz covariance matrix by a block-circulant matrix and the likelihood with the block-circulant matrix converges to that with the block-Toeplitz matrix.

Before proceeding we should note that a (block) circulant matrix can be diagonalized by the DFT. To this end let us consider the following linear transformation of  $\mathbf{w}_\bullet$ ,

$$\mathbf{z}_\bullet = (\mathbf{L}_{NP,N} \otimes \mathbf{I}_L)(\mathbf{F}_{NP} \otimes \mathbf{I}_L)^H \mathbf{w}_\bullet, \quad (10)$$

where  $\clubsuit \in \{s, r\}$ ,  $\mathbf{L}_{NP,N}$  is the commutation matrix,<sup>2</sup> and  $\mathbf{F}_{NP}$  is the DFT matrix of size  $NP$ . Hence,  $\mathbf{z}_\clubsuit$  contains a specific reordering of the frequencies in  $\mathbf{w}_\clubsuit$ . In order to give an insight into the reordering let us first partition  $\mathbf{z}_s$  into  $N$  blocks  $\mathbf{x}[n] \in \mathbb{C}^{LP}$  and  $\mathbf{z}_r$  into  $N$  blocks  $\mathbf{y}[n] \in \mathbb{C}^{LP}$  for  $n = 0, \dots, N-1$ . Furthermore, the DFT of length  $NP$  of  $\mathbf{u}_\clubsuit[n]$  is defined as

$$\mathbf{u}_\clubsuit(\theta_k) = \sum_{n=0}^{NP-1} \mathbf{u}_\clubsuit[n] e^{-j\theta_k n}, \quad (11)$$

with  $\theta_k = \frac{2\pi}{NP}k$ . Then the  $n$ th block of size  $LP$  of  $\mathbf{z}_s$  is given by

$$\mathbf{x}[n] = [\mathbf{u}_s^T(\theta_n) \mathbf{u}_s^T(\theta_{N+n}) \cdots \mathbf{u}_s^T(\theta_{(P-1)N+n})]^T \quad (12)$$

and similarly for  $\mathbf{z}_r$ :

$$\mathbf{y}[n] = [\mathbf{u}_r^T(\theta_n) \mathbf{u}_r^T(\theta_{N+n}) \cdots \mathbf{u}_r^T(\theta_{(P-1)N+n})]^T. \quad (13)$$

Hence, each of the blocks  $\mathbf{x}[n]$  and  $\mathbf{y}[n]$  contains  $P$  frequencies separated by multiples of the fundamental cycle frequency  $\frac{2\pi}{P}$ . Recall that frequency components of a CS process separated by multiples of a cycle frequency may be correlated [42].

Let us now investigate the (cross) covariance matrices of  $\mathbf{z}_s$  and  $\mathbf{z}_r$ . Under the null hypothesis we obtain

$$\mathbf{S}_0 = \mathbb{E}[\mathbf{z}\mathbf{z}^H | \mathcal{H}_0] = \begin{bmatrix} \mathbf{S}_s^{(0)} & \mathbf{0} \\ \mathbf{0} & \mathbf{S}_r \end{bmatrix}, \quad (14)$$

where the off-diagonal blocks are zero since observations at SC and RC are uncorrelated, and the covariance matrices  $\mathbf{S}_s^{(0)} = \mathbb{E}[\mathbf{z}_s \mathbf{z}_s^H | \mathcal{H}_0] \in \mathbb{S}_L^{LNP}$  and  $\mathbf{S}_r = \mathbb{E}[\mathbf{z}_r \mathbf{z}_r^H] \in \mathbb{S}_{LP}^{LNP}$  are block-diagonal matrices with block size  $L$  and  $LP$ , respectively, since the covariance matrices of  $\mathbf{w}_s$  and  $\mathbf{w}_r$  are asymptotically block-circulant and diagonalized by the linear transformation in (10). Note that  $\mathbf{S}_r$  is block-diagonal with block size  $LP$  regardless of the hypothesis. Under  $\mathcal{H}_1$ , the covariance matrix is given by

$$\mathbf{S}_1 = \mathbb{E}[\mathbf{z}\mathbf{z}^H | \mathcal{H}_1] = \begin{bmatrix} \mathbf{S}_s^{(1)} & \mathbf{S}_{sr} \\ \mathbf{S}_{rs} & \mathbf{S}_r \end{bmatrix}, \quad (15)$$

where  $\mathbf{S}_s^{(1)} = \mathbb{E}[\mathbf{z}_s \mathbf{z}_s^H | \mathcal{H}_1]$  and  $\mathbf{S}_{sr} = \mathbf{S}_{rs}^H = \mathbb{E}[\mathbf{z}_s \mathbf{z}_r^H | \mathcal{H}_1]$  are block-diagonal with block size  $LP$ . Hence, each of the four blocks in  $\mathbf{S}_1$  is now given by a block-diagonal matrix with block size  $LP$ . Finally, the hypotheses can be formulated as

$$\begin{aligned} \mathcal{H}_0 : \mathbf{z} &\sim \mathcal{CN}_{2LNP}(\mathbf{0}, \mathbf{S}_0), \\ \mathcal{H}_1 : \mathbf{z} &\sim \mathcal{CN}_{2LNP}(\mathbf{0}, \mathbf{S}_1). \end{aligned} \quad (16)$$

### III. DERIVATION OF THE GLRT

The GLR is given by

$$\mathcal{G} = \frac{p(\mathbf{z}_0, \dots, \mathbf{z}_{M-1}; \hat{\mathbf{S}}_0)}{p(\mathbf{z}_0, \dots, \mathbf{z}_{M-1}; \hat{\mathbf{S}}_1)}, \quad (17)$$

where  $\mathbf{z}_0, \dots, \mathbf{z}_{M-1}$  denote  $M$  independent and identically distributed (i.i.d.) realizations<sup>3</sup> of  $\mathbf{z}$  and  $\hat{\mathbf{S}}_0$  and  $\hat{\mathbf{S}}_1$  denote the ML estimates of  $\mathbf{S}_0$  and  $\mathbf{S}_1$ , respectively. Under the Gaussian assumption the likelihoods are given by

$$\begin{aligned} p(\mathbf{z}_0, \dots, \mathbf{z}_{M-1}; \hat{\mathbf{S}}_j) &= \frac{1}{\pi^{2LNP} \det(\hat{\mathbf{S}}_j)^M} \\ &\times \exp \left\{ -M \operatorname{tr}(\mathbf{Q} \hat{\mathbf{S}}_j^{-1}) \right\}, \end{aligned} \quad (18)$$

where  $\mathbf{Q} = \frac{1}{M} \sum_{m=0}^{M-1} \mathbf{z}_m \mathbf{z}_m^H = \begin{bmatrix} \mathbf{Q}_s & \mathbf{Q}_{sr} \\ \mathbf{Q}_{rs} & \mathbf{Q}_r \end{bmatrix}$  is the sample covariance matrix of  $\mathbf{z}$  and  $j \in \{0, 1\}$  indicates whether it is the ML estimate under  $\mathcal{H}_0$  or  $\mathcal{H}_1$ .

In the following we will derive the GLRT, which requires the ML estimation of the covariance matrices under both hypotheses. Although this is straightforward under the null hypothesis as it requires the ML estimation of a block-diagonal matrix, it demands a suitable permutation under  $\mathcal{H}_1$  to obtain another block-diagonal covariance matrix that is easy to estimate.

**Theorem 1.** *The GLR (17) is given by*

$$\mathcal{G}^{\frac{1}{M}} = \prod_{k=1}^N \det(\mathbf{D}_k - \mathbf{C}_k \mathbf{C}_k^H), \quad (19)$$

where  $\mathbf{D}_k$  is the  $k$ th  $LP \times LP$  block of

$$\mathbf{D} = \operatorname{diag}_L(\mathbf{Q}_s)^{-1/2} \operatorname{diag}_{LP}(\mathbf{Q}_s) \operatorname{diag}_L(\mathbf{Q}_s)^{-1/2}, \quad (20)$$

and  $\mathbf{C}_k$  the  $k$ th  $LP \times LP$  block of

$$\mathbf{C} = \operatorname{diag}_L(\mathbf{Q}_s)^{-1/2} \operatorname{diag}_{LP}(\mathbf{Q}_{sr}) \operatorname{diag}_{LP}(\mathbf{Q}_r)^{-1/2}. \quad (21)$$

*Proof.* See Appendix A.  $\square$

As can be observed, the GLR consists of two parts. The first one is the coherence matrix  $\mathbf{D}$ , which accounts for the spectral correlation present at the SC. The second part is the cross-coherence matrix  $\mathbf{C}$ , which captures the cross-correlation between SC and RC, i.e., it accounts for the inherent cross-correlation and also for cross-spectral correlation induced by the presence of cyclostationarity.

Note that there are also the Rao and Wald tests, which could be applied to our problem. Asymptotically, these tests have the same performance as the GLRT [43] but in the finite sample case their performance depends on the specific underlying model as was pointed out for a different problem in, e.g., [44], [45].

#### Threshold selection and null distribution

In order to apply the proposed detector, it is necessary to determine a threshold that assures a given probability of false alarm. To this end we propose two alternatives. The first one considers the invariances of the tests. We observe that  $\mathbf{z}_s$  can be multiplied by any non-singular block-diagonal matrix with block size  $L$  and  $\mathbf{z}_r$  with any non-singular block-diagonal

<sup>2</sup>The commutation matrix fulfills the following equation:  $\operatorname{vec}(\mathbf{A}) = \mathbf{L}_{MN,N} \operatorname{vec}(\mathbf{A}^T)$  for an  $M \times N$  matrix  $\mathbf{A}$ .

<sup>3</sup>In practice i.i.d. observations are rarely available. This may be addressed by dividing a long observation into  $M$  windows and treating them as if they were i.i.d.

matrix with block size  $LP$  without changing the structure of  $\mathbf{S}_0$  and  $\mathbf{S}_1$ , i.e., the test is invariant to the noise PSD in the SC and signal-plus-noise PSD in the RC. In the time-domain this corresponds to a circular convolution of  $\mathbf{u}_s[n]$  with an arbitrary  $L$ -variate sequence and a circular convolution of the stack of  $P$  observations of  $\mathbf{u}_r[n]$  with an arbitrary  $LP$ -variate sequence, which is asymptotically equivalent to (MIMO) linear filtering. These invariances allow us to assume, without loss of generality, that under  $\mathcal{H}_0$   $\mathbf{z} \stackrel{N \rightarrow \infty}{\sim} \mathcal{CN}(\mathbf{0}, \mathbf{I}_{2LPN})$ . Hence, numerical simulations with a temporally and spatially white process can be used to obtain the threshold under the null hypothesis for any arbitrary process.

The second approach decomposes the GLR, similar to [46], such that its distribution is asymptotically equivalent to a product of independent beta random variables.

**Proposition 1.** *Under the null hypothesis the likelihood ratio (19) is distributed as*

$$\mathcal{G}^{\frac{1}{M}} \stackrel{\mathcal{D}}{=} \prod_{n=1}^N \prod_{l=1}^L V_{l0} \prod_{p=1}^{P-1} U_{lp} V_{lp}, \quad (22)$$

where  $U_{lp} \sim \text{Beta}(\alpha_{lp}, \alpha_p)$  and  $V_{lp} \sim \text{Beta}(\beta_{lp}, \beta)$  with

$$\alpha_{lp} = M - (Lp + l - 1), \quad (23)$$

$$\alpha_p = Lp, \quad (24)$$

$$\beta_{lp} = M - (LP + Lp + l - 1), \quad (25)$$

$$\beta = LP. \quad (26)$$

*Proof.* Please refer to Appendix B.  $\square$

Since both approaches only hold asymptotically, the finite-sample size effects will be studied in Section VII.

#### IV. DERIVATION OF OPTIMAL INVARIANT TESTS

In this section we study the existence of invariant tests. In particular, we first consider the UMPIT, which is the optimal detector among those that are invariant. Moreover, we also consider the LMPIT, which is optimal only for close hypotheses. In order to derive the UMPIT or the LMPIT there are several steps that need to be accomplished [47]: (i) determine the group of invariant transformations, (ii) identify the maximal invariant statistic, (iii) determine the distribution of the maximal invariant under both hypotheses, and (iv) obtain the likelihood ratio of the densities. If this ratio (or a monotone transformation thereof) does not depend on the unknown parameters, it would yield the UMPIT. Although there are some scenarios in which the maximal invariant statistic and its distributions can be established, e.g. [44], [48], in general this can be a tedious approach. In order to avoid these involved tasks, we will make use of Wijsman's theorem, which allows us to directly compute the ratio of maximal invariants [31]. In the derivation, we will show that neither the UMPIT nor the LMPIT exist for the given hypothesis test. The first step of this proof is to identify the invariances of the hypothesis test as they are required in Wijsman's theorem.

Considering only linear operations, which will maintain Gaussianity, we may first observe that we can multiply  $\mathbf{z}_s$  by any non-singular block-diagonal matrix with block size

$L$  and  $\mathbf{z}_r$  with any non-singular block-diagonal matrix with block size  $LP$  without changing the structure of  $\mathbf{S}_0$  and  $\mathbf{S}_1$ . Secondly, we can permute the blocks  $\mathbf{x}[n]$  in  $\mathbf{z}_s$  arbitrarily, provided that we apply the same permutation to the blocks  $\mathbf{y}[n]$  in  $\mathbf{z}_r$ . This corresponds to a reordering of the blocks that contain  $P$  frequencies separated by multiples of  $\frac{2\pi}{P}$ . Moreover, we may arbitrarily permute these  $P$  frequencies within each block  $\mathbf{x}[n]$  and  $\mathbf{y}[n]$  for every  $n = 0, \dots, N-1$ . Hence, the invariance group can be formulated as

$$\mathcal{G} = \{g : \mathbf{z} \rightarrow g(\mathbf{z}) = \Psi \mathbf{z}\}, \quad (27)$$

where  $\Psi = \begin{bmatrix} \mathbf{P}_s \mathbf{G} & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_r \mathbf{H} \end{bmatrix}$  with

$$\mathbf{P}_\star = \left( \sum_{k=1}^N \epsilon_k \epsilon_k^T \otimes \mathbf{V}_\star^{(k)} \otimes \mathbf{I}_L \right) (\mathbf{U} \otimes \mathbf{I}_{LP}), \quad (28)$$

$\epsilon_k$  is the  $k$ th column of  $\mathbf{I}_N$ ,  $\mathbf{V}_\star^{(k)} \in \mathbb{V}$  denotes a  $P \times P$  permutation matrix, and  $\mathbf{U} \in \mathbb{U}$  is a permutation matrix of size  $N \times N$ .  $\mathbb{V}$  and  $\mathbb{U}$  denote the corresponding sets of  $P$ - and  $N$ -dimensional permutation matrices, respectively. Furthermore,  $\mathbf{G} \in \mathbb{G}$  and  $\mathbf{H} \in \mathbb{H}$ , where  $\mathbb{G}$  is the set of nonsingular block-diagonal matrices with block size  $L$  and  $\mathbb{H}$  denotes the set of nonsingular block-diagonal matrices with block size  $LP$ . In (28), the left parenthesized expression performs the permutation within the blocks  $\mathbf{x}[n]$  or  $\mathbf{y}[n]$ , respectively, and the right parenthesized expression applies the same permutation to the blocks  $\mathbf{x}[n]$  and the blocks  $\mathbf{y}[n]$ .

Now we will use Wijsman's theorem [31] to obtain the ratio of the maximal invariant densities under the two hypotheses, which is given by

$$\mathcal{L} = \frac{\int_{\mathcal{G}} p(g(\mathbf{z}); \mathcal{H}_1) |\det(\mathbf{J}_g)| dg}{\int_{\mathcal{G}} p(g(\mathbf{z}); \mathcal{H}_0) |\det(\mathbf{J}_g)| dg}, \quad (29)$$

where  $\mathcal{G}$  denotes the group of invariant transformations, which we identified for the given problem in the previous paragraph, the transformation  $g(\cdot) \in \mathcal{G}$ ,  $p(\mathbf{z}; \mathcal{H}_i)$  is the probability density function of  $\mathbf{z}$  under hypothesis  $\mathcal{H}_i$ ,  $\mathbf{J}_g$  denotes the Jacobian matrix of the transformation  $g(\cdot)$ , and finally  $dg$  denotes the invariant group measure, which in our case is the usual Lebesgue measure.

For the problem considered in this paper, Wijsman's theorem states that the ratio of the distributions of the maximal invariant statistic is given by (30) at the top of the next page, where  $\sum_{\mathbb{V}_0^N, \mathbb{V}_1^N, \mathbb{U}} = \sum_{\mathbb{V}_0^{(1)}} \cdots \sum_{\mathbb{V}_0^{(N)}} \sum_{\mathbb{V}_1^{(1)}} \cdots \sum_{\mathbb{V}_1^{(N)}} \sum_{\mathbb{U}}$ , and  $d\mathbf{G}$  and  $d\mathbf{H}$  are the invariant measures on the sets  $\mathbb{G}$  and  $\mathbb{H}$ , respectively. If the ratio did not depend on unknown parameters, the UMPIT would exist. However, it will turn out by further simplifying (30) that the UMPIT does not exist for this problem.

**Lemma 1.** *The ratio (30) can be simplified as*

$$\mathcal{L} \propto \sum_{\mathbb{V}_0^N, \mathbb{V}_1^N, \mathbb{U}} \int_{\mathbb{G}} \int_{\mathbb{H}} \beta(\mathbf{G}) \beta(\mathbf{H}) e^{-M[\alpha_1(\mathbf{G}) + \alpha_2(\mathbf{G}, \mathbf{H})]} d\mathbf{G} d\mathbf{H}, \quad (31)$$

$$\mathcal{L} = \frac{\sum_{\mathbf{V}_0^N, \mathbf{V}_1^N, \mathbf{U}} \int_{\mathbf{G}} \int_{\mathbf{H}} \det(\mathbf{S}_1)^{-M} |\det(\mathbf{G})|^{2M} |\det(\mathbf{H})|^{2M} \exp \left\{ -M \text{tr} \left( \Psi \mathbf{Q} \Psi^H \mathbf{S}_1^{-1} \right) \right\} d\mathbf{G} d\mathbf{H}}{\sum_{\mathbf{V}_0^N, \mathbf{V}_1^N, \mathbf{U}} \int_{\mathbf{G}} \int_{\mathbf{H}} \det(\mathbf{S}_0)^{-M} |\det(\mathbf{G})|^{2M} |\det(\mathbf{H})|^{2M} \exp \left\{ -M \text{tr} \left( \Psi \mathbf{Q} \Psi^H \mathbf{S}_0^{-1} \right) \right\} d\mathbf{G} d\mathbf{H}} \quad (30)$$

with

$$\beta(\mathbf{A}) = |\det(\mathbf{A})| e^{-M \text{tr}(\mathbf{A} \mathbf{A}^H)}, \quad (32)$$

$$\alpha_1(\mathbf{G}) = \sum_{k=1}^N \sum_{\substack{m,n=1 \\ m \neq n}}^P \text{tr} \left( \Gamma_k^{(m,n)} \mathbf{G}_k^{(n,n)} \mathbf{D}_k^{(n,m)} \mathbf{G}_k^{(m,m)H} \right), \quad (33)$$

$$\Gamma = \mathbf{P}_s^T \text{diag}_L(\boldsymbol{\Sigma}_1)^{-\frac{1}{2}} \boldsymbol{\Sigma}_1 \text{diag}_L(\boldsymbol{\Sigma}_1)^{-\frac{1}{2}} \mathbf{P}_s, \quad (34)$$

$$\alpha_2(\mathbf{G}, \mathbf{H}) = \sum_{k=1}^N \text{tr} \left( \Lambda_k \mathbf{G}_k \mathbf{C}_k \mathbf{H}_k^H \right), \quad (35)$$

$$\Lambda = \mathbf{P}_r^T \boldsymbol{\Sigma}_2^{-\frac{1}{2}} \boldsymbol{\Sigma}_{21} \text{diag}_L(\boldsymbol{\Sigma}_1)^{-\frac{1}{2}} \mathbf{P}_r, \quad (36)$$

$$\text{where } \mathbf{S}_1^{-1} = \boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_2 \end{bmatrix}.$$

*Proof.* See Appendix C.  $\square$

We should note that both  $\Gamma$  and  $\Lambda$  depend on unknown parameters in  $\boldsymbol{\Sigma}$ . For this reason we can conclude that the UMPIT does not exist. However, we may focus on the case of close hypotheses to examine the existence of an LMPIT. In our scenario the hypotheses are close if the SNR at the SC is very low. In this case the cross-correlation between SC and RC is close to zero, i.e.,  $\mathbf{S}_{sr} \approx \mathbf{0}$ , and at the SC the covariance matrix  $\mathbf{S}_s$  is close to block-diagonal with block size  $L$ . For this reason it follows that  $\boldsymbol{\Sigma}_{12} \approx \mathbf{0}$ , and  $\boldsymbol{\Sigma}_1$  is also close to block-diagonal with block size  $L$ . Therefore, both  $\alpha_1(\mathbf{G}) \approx 0$  and  $\alpha_2(\mathbf{G}, \mathbf{H}) \approx 0$ , and we may use a second-order Taylor series approximation to approximate the exponential in (31) around  $\alpha_1(\mathbf{G}) + \alpha_2(\mathbf{G}, \mathbf{H}) = 0$  as

$$e^{-M(\alpha_1(\mathbf{G}) + \alpha_2(\mathbf{G}, \mathbf{H}))} \approx 1 - M(\alpha_1(\mathbf{G}) + \alpha_2(\mathbf{G}, \mathbf{H})) + \frac{M^2}{2} [\alpha_1^2(\mathbf{G}) + 2\alpha_1(\mathbf{G})\alpha_2(\mathbf{G}, \mathbf{H}) + \alpha_2^2(\mathbf{G}, \mathbf{H})]. \quad (37)$$

Thus, (31) can be approximated as

$$\mathcal{L} \propto \mathcal{L}_1 + \mathcal{L}_2 + \mathcal{L}_3 + \mathcal{L}_4 + \mathcal{L}_5, \quad (38)$$

where

$$\mathcal{L}_1 = -M \sum_{\mathbf{V}_0^N, \mathbf{V}_1^N, \mathbf{U}} \int_{\mathbf{G}} \beta(\mathbf{G}) \alpha_1(\mathbf{G}) d\mathbf{G} \int_{\mathbf{H}} \beta(\mathbf{H}) d\mathbf{H}, \quad (39)$$

$$\mathcal{L}_2 = -M \sum_{\mathbf{V}_0^N, \mathbf{V}_1^N, \mathbf{U}} \int_{\mathbf{G}} \int_{\mathbf{H}} \beta(\mathbf{G}) \beta(\mathbf{H}) \alpha_2(\mathbf{G}, \mathbf{H}) d\mathbf{G} d\mathbf{H}, \quad (40)$$

$$\mathcal{L}_3 = \frac{M^2}{2} \sum_{\mathbf{V}_0^N, \mathbf{V}_1^N, \mathbf{U}} \int_{\mathbf{G}} \beta(\mathbf{G}) \alpha_1^2(\mathbf{G}) d\mathbf{G} \int_{\mathbf{H}} \beta(\mathbf{H}) d\mathbf{H}, \quad (41)$$

$$\mathcal{L}_4 = \frac{M^2}{2} \sum_{\mathbf{V}_0^N, \mathbf{V}_1^N, \mathbf{U}} \int_{\mathbf{G}} \int_{\mathbf{H}} \beta(\mathbf{G}) \beta(\mathbf{H}) \alpha_1(\mathbf{G}) \alpha_2(\mathbf{G}, \mathbf{H}) d\mathbf{G} d\mathbf{H}, \quad (42)$$

$$\mathcal{L}_5 = \frac{M^2}{2} \sum_{\mathbf{V}_0^N, \mathbf{V}_1^N, \mathbf{U}} \int_{\mathbf{G}} \int_{\mathbf{H}} \beta(\mathbf{G}) \beta(\mathbf{H}) \alpha_2^2(\mathbf{G}, \mathbf{H}) d\mathbf{G} d\mathbf{H}. \quad (43)$$

**Lemma 2.** *The following terms are zero:*

$$\mathcal{L}_1 = 0, \quad (44)$$

$$\mathcal{L}_2 = 0, \quad (45)$$

$$\mathcal{L}_4 = 0. \quad (46)$$

*Proof.* Let us first focus on  $\mathcal{L}_1$ , which is given by

$$\begin{aligned} \mathcal{L}_1 &\propto \sum_{\mathbf{V}_0^N, \mathbf{V}_1^N, \mathbf{U}} \int_{\mathbf{G}} \int_{\mathbf{H}} \beta(\mathbf{G}) \beta(\mathbf{H}) \\ &\times \sum_{k=1}^N \sum_{\substack{m,n=1 \\ m \neq n}}^P \text{tr} \left( \Gamma_k^{(m,n)} \mathbf{G}_k^{(n,n)} \mathbf{D}_k^{(n,m)} \mathbf{G}_k^{(m,m)H} \right) d\mathbf{G} d\mathbf{H}. \end{aligned} \quad (47)$$

Applying the change of variables  $\mathbf{G}_k^{(n,n)} \rightarrow -\mathbf{G}_k^{(n,n)}$  and it can be seen that the integrals need to be equal to their opposites, i.e., they are zero. In a similar fashion, it can be shown that the terms  $\mathcal{L}_2$  and  $\mathcal{L}_4$  are zero.  $\square$

Finally, the quadratic terms in  $\alpha_1(\mathbf{G})$  and  $\alpha_2(\mathbf{G}, \mathbf{H})$  remain in (38). In the following theorem we will show that these terms can be expressed as functions of the (cross) coherence matrices (20) and (21).

**Theorem 2.** *The ratio of the distribution of the maximal invariant statistic in (30) is*

$$\mathcal{L} \propto \mathcal{L}_S + \gamma \mathcal{L}_{SR}, \quad (48)$$

where

$$\mathcal{L}_S = \sum_{k=1}^N \|\mathbf{D}_k\|^2 \quad (49)$$

and

$$\mathcal{L}_{SR} = \sum_{k=1}^N \|\mathbf{C}_k\|^2 \quad (50)$$

with  $\mathbf{D}$  and  $\mathbf{C}$  given by (20) and (21), respectively. The parameter  $\gamma$  is a constant that depends on unknown parameters but is independent of the observations.

*Proof.* See Appendix D.  $\square$

Since  $\mathcal{L}$  still depends on unknown parameters via the constant  $\gamma$ , we can conclude that the LMPIT does not exist. We should note that the term  $\mathcal{L}_S$  is the LMPIT for the single array CS detection problem [27]. After giving an interpretation of both  $\mathcal{L}_S$  and  $\mathcal{L}_{SR}$  in the following section, we will study

the influence of the two terms on the detection performance as a function of  $\gamma$  in Section VI, which will finally show that an LMPIT-inspired detector can be suggested.

## V. INTERPRETATION OF THE TEST STATISTICS

As can be seen in (19) and (48), both the GLRT and the ratio of the distribution of the maximal invariant statistics are functions of the sample coherence matrix  $\mathbf{D}$  and the sample cross-coherence matrix  $\mathbf{C}$  given in (20) and (21), respectively. Similarly to [27], we will provide an interpretation of these statistics. Recall that the cyclic (cross) PSD at cycle frequency  $\frac{2\pi}{P}l$  is given by [49]

$$\mathbf{\Pi}_{\clubsuit\heartsuit}^{(l)}(\theta) d\theta = \mathbb{E} \left[ d\boldsymbol{\xi}_{\clubsuit}(\theta) d\boldsymbol{\xi}_{\heartsuit}^H \left( \theta - \frac{2\pi}{P}l \right) \right], \quad (51)$$

where  $\clubsuit, \heartsuit \in \{s, r\}$  and  $d\boldsymbol{\xi}_{\clubsuit}(\theta) \in \mathbb{C}^L$  denotes the increment of a spectral process  $\boldsymbol{\xi}_{\clubsuit}(\theta)$  that generates the time series

$$\mathbf{u}_{\clubsuit}[n] = \int_{-\pi}^{\pi} e^{j\theta n} d\boldsymbol{\xi}_{\clubsuit}(\theta). \quad (52)$$

Furthermore, the cyclic (cross) PSD and the bi-frequency spectrum are related by [49]

$$\mathbf{S}_{\clubsuit\heartsuit}(\theta_i, \theta_j) = \sum_l \mathbf{\Pi}_{\clubsuit\heartsuit}^{(l)}(\theta_j) \delta \left( \theta_i - \theta_j - \frac{2\pi}{P}l \right) \in \mathbb{C}^{L \times L}. \quad (53)$$

Note that the line for  $l = 0$  is the stationary manifold, which contains the usual PSD. Moreover, the support of  $\mathbf{S}_{\clubsuit\heartsuit}(\theta_i, \theta_j)$  may only contain frequencies separated by multiples of the fundamental cycle frequency  $\frac{2\pi}{P}$ , i.e.,  $\theta_i - \theta_j = \frac{2\pi}{P}l$ , for a CS process with cycle period  $P$ . As we have already mentioned in Section II, these possibly non-zero components are contained in the  $LP \times LP$  blocks on the main diagonal of  $\mathbf{S}_s^{(1)}$ ,  $\mathbf{S}_r$ , and  $\mathbf{S}_{sr}$  in (15). For instance, the  $(i, j)$ th  $L \times L$  sized block of the  $k$ th diagonal block of  $\mathbf{S}_{sr}$  is given by

$$[\mathbf{S}_{sr}]_k^{(i,j)} = \mathbf{S}_{sr}(\theta_{iN+k}, \theta_{jN+k}) = \mathbf{\Pi}_{sr}^{(i-j)}(\theta_{jN+k}) \in \mathbb{C}^{L \times L}, \quad (54)$$

where  $\theta_l = \frac{2\pi l}{NP}$ ,  $k = 0, \dots, N-1$ ,  $i, j = 0, \dots, P-1$ , and similarly for  $\mathbf{S}_s^{(1)}$  and  $\mathbf{S}_r$ . Accordingly, the ML estimates of the covariance matrices contain samples of cyclic (cross) PSDs. Comparing (53) and (54) shows that the  $L \times L$  diagonal blocks for  $i = j$  correspond to the (cross) PSD on the stationary manifold for frequency  $\theta_{jN+k}$ , and the off-diagonal blocks for  $i \neq j$  correspond to the cyclic (cross) PSD at frequency  $\theta_{jN+k}$  and at cycle frequency  $\frac{2\pi(i-j)}{P}$ .

The latter considerations allow us to rewrite the (cross) coherence matrices  $\mathbf{D}$  and  $\mathbf{C}$  as functions of the cyclic (cross) PSDs.

**Proposition 2.** *The  $L \times L$  blocks in the (cross) coherence matrices  $\mathbf{D}$  and  $\mathbf{C}$  can be expressed by the samples of the cyclic (cross) PSDs as*

$$\mathbf{D}^{(q)}(\theta_{jN+k}) = \left[ \mathbf{\Pi}_{ss}^{(0)} \left( \theta_{jN+k} + \frac{2\pi}{P}q \right) \right]^{-\frac{1}{2}} \times \mathbf{\Pi}_{ss}^{(q)}(\theta_{jN+k}) \left[ \mathbf{\Pi}_{ss}^{(0)}(\theta_{jN+k}) \right]^{-\frac{1}{2}}, \quad (55)$$

and

$$\mathbf{C}^{(q)}(\theta_{jN+k}) = \left[ \mathbf{\Pi}_{ss}^{(0)} \left( \theta_{jN+k} + \frac{2\pi}{P}q \right) \right]^{-\frac{1}{2}} \times \sum_{m=-j}^{P-1-j} \mathbf{\Pi}_{sr}^{(q-m)} \left( \theta_{mN+k} + \frac{2\pi}{P}j \right) \left[ \mathbf{\Pi}_{rr}^{(m)}(\theta_{jN+k}) \right]^{-\frac{1}{2}}, \quad (56)$$

for  $j = 0, \dots, P-1$ ,  $q = -j, \dots, P-1-j$ , and  $k = 0, \dots, N-1$ .

*Proof.* See Appendix E.  $\square$

As can be seen, the coherence matrix  $\mathbf{D}$  contains the cyclic PSD of the SC signal  $\mathbf{\Pi}_{ss}^{(q)}(\theta_{jN+k})$  for  $q \neq 0$  normalized by the PSD, which lives on the stationary manifold. The cross-coherence matrix  $\mathbf{C}$ , on the other hand, contains the cross-cyclic PSD between SC and RC,  $\mathbf{\Pi}_{sr}^{(q-m)}(\theta_{mN+k} + \frac{2\pi}{P}j)$ , normalized by  $\mathbf{\Pi}_{rr}^{(m)}(\theta_{jN+k})$  and  $\mathbf{\Pi}_{ss}^{(0)}(\theta_{jN+k} + \frac{2\pi}{P}q)$  and sums it over  $m = -j, \dots, P-1-j$ . Note that the main diagonal blocks of  $\mathbf{C}$  given by (57), at the top of the next page, do not only account for the cyclic components but also for the usual cross-coherence between the WSS components at frequency  $\theta_{jN+k}$  given by the first term in the equation.

In a nutshell, the coherence matrix  $\mathbf{D}$  accounts for the spectral correlation at the SC, whereas the cross-coherence matrix  $\mathbf{C}$  accounts for the cross-spectral correlation between SC and RC. Furthermore, comparing the GLRT  $\mathcal{G}$  in (19) and the ratio  $\mathcal{L}$  in (48), it can be observed that the GLRT inherently merges the information provided by the presence of cyclostationarity at the SC via  $\mathbf{D}$  and the correlation of SC and RC present in  $\mathbf{C}$ , whereas in  $\mathcal{L}$  these terms are connected by the unknown parameter  $\gamma$  in (48). Moreover, another difference is the way the spectral correlation is measured in the two tests. The GLRT employs the determinant, whereas the ratio of the distribution of maximal invariants uses the Frobenius norm.

## VI. LMPIT-INSPIRED DETECTOR

Since no LMPIT exists, we now analyze the influence of  $\gamma$ , i.e., the influence of the individual terms  $\mathcal{L}_S$  and  $\mathcal{L}_{SR}$  in (48) on the detection performance. As mentioned before, on the one hand the term  $\mathcal{L}_S$  is the LMPIT for CS detection at a single array (the SC). Specifically, it measures the strength of the cyclic components relative to the stationary components. On the other hand,  $\mathcal{L}_{SR}$  measures the strength of cross-spectral correlation between SC and RC, i.e., it accounts for the inherent correlation between SC and RC and also for the spectral correlation induced by cyclostationarity. For this reason it is expected that  $\mathcal{L}_{SR}$  will have a bigger influence on the detection performance than  $\mathcal{L}_S$  provided that the signals are not too weak.

Since the theoretical distribution of (48) is very difficult to obtain, we used Monte Carlo simulations to study the influence of  $\gamma$ . In order to do so, we used the signal model to be described in Section VII to generate realizations under  $\mathcal{H}_0$  and  $\mathcal{H}_1$ . For a given set of values for  $\gamma$  we obtained the probability of detection  $p_d$  based on the statistic  $\mathcal{L}$  for a fixed probability of false alarm. Although we use  $\mathcal{L}$  throughout this section as a

$$\mathbf{C}^{(0)}(\theta_{jN+k}) = \left[ \mathbf{\Pi}_{ss}^{(0)}(\theta_{jN+k}) \right]^{-\frac{1}{2}} \mathbf{\Pi}_{sr}^{(0)}(\theta_{jN+k}) \left[ \mathbf{\Pi}_{rr}^{(0)}(\theta_{jN+k}) \right]^{-\frac{1}{2}} + \left[ \mathbf{\Pi}_{ss}^{(0)}(\theta_{jN+k}) \right]^{-\frac{1}{2}} \sum_{m=-j}^{P-1-j} \mathbf{\Pi}_{sr}^{(-m)}(\theta_{mN+k} + \frac{2\pi}{P}j) \left[ \mathbf{\Pi}_{rr}^{(m)}(\theta_{jN+k}) \right]^{-\frac{1}{2}} \quad (57)$$

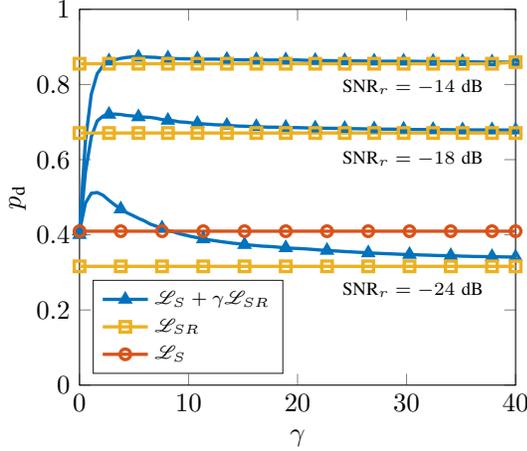


Fig. 2. Probability of detection as a function of  $\gamma$  based on different detection statistics for an experiment with the following parameters:  $P = 4$ ,  $N = 128$ ,  $M = 20$ ,  $L = L_I = 2$ , a rectangular pulse,  $\text{SNR}_s = -18$  dB,  $\text{SNR}_r = \{-14, -18, -24\}$  dB, and  $p_{fa} = 0.01$ .

benchmark for detectors based on the two individual terms  $\mathcal{L}_S$  and  $\mathcal{L}_{SR}$ , note that in practice this is not possible as it depends on unknown parameters in  $\gamma$ . Additionally, we obtained the probability of detection based on using either  $\mathcal{L}_S$  or  $\mathcal{L}_{SR}$  individually.

The impact of the parameter  $\gamma$  on detection probability is shown in Figure 2. For a fixed  $\text{SNR}_s = -18$  dB we obtained the detection probabilities for the detectors based on  $\mathcal{L}_S$  and  $\mathcal{L}_{SR}$  for three different values of the SNR at the RC, which are  $-14$  dB,  $-18$  dB, and  $-24$  dB, and also the detection probability of  $\mathcal{L}_S$ , which is independent of  $\text{SNR}_r$ . It can be observed that for this scenario a reasonable performance is only reached for  $\text{SNR}_r = -14$  dB. Moreover, the probability of detection based on  $\mathcal{L}_{SR}$  almost overlaps with that based on the optimal statistic  $\mathcal{L}$ . For lower  $\text{SNR}_r$  the correlation between signals at SC and RC is getting weaker and the right choice of  $\gamma$  becomes more critical for the best performance. At  $\text{SNR}_r = -24$  dB, we observe that a detector based on  $\mathcal{L}_S$  outperforms a detector based on  $\mathcal{L}_{SR}$ , i.e., better performance is obtained by simply detecting the presence of cyclostationarity at the SC. If the optimal  $\gamma$  were known, the performance of  $\mathcal{L}$  could be reached. However, it should be noted that for such a low  $\text{SNR}_r$  even the optimal detector would not provide satisfactory performance.

As the SNR at the reference array is typically not less than the SNR at the surveillance array, we compare the probabilities of detection for equal SNRs at SC and RC for detectors based on  $\mathcal{L}$  (where  $\gamma$  has been determined by a brute-force search to maximize the probability of detection),  $\mathcal{L}_{SR}$ , and  $\mathcal{L}_S$  in

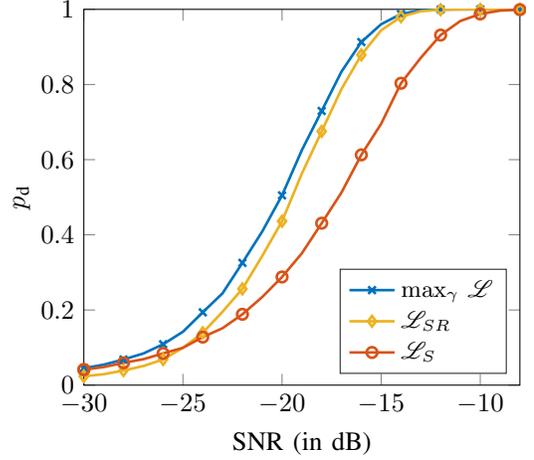


Fig. 3. Probability of detection vs. SNR for various detectors, where  $\text{SNR}_s = \text{SNR}_r = \text{SNR}$  for an experiment with the following parameters:  $P = 4$ ,  $N = 128$ ,  $M = 20$ ,  $L = L_I = 2$ , a rectangular pulse, and  $p_{fa} = 0.01$ .

Figure 3. It can be observed that although there is a gap between the optimal  $p_d$  and the  $p_d$  of  $\mathcal{L}_{SR}$ , it is comparatively small and it decreases as the SNR increases. Moreover, the gap between  $\mathcal{L}_{SR}$  and  $\mathcal{L}_S$  decreases with decreasing SNR, which we expect because the lower the SNR, the more beneficial the CS detection at the SC only. For different scenarios where we vary, for instance,  $M$  or  $N$ , we have also observed (the results are not reproduced here) that the performance of  $\mathcal{L}_{SR}$  is close to  $\mathcal{L}$  with the optimal  $\gamma$  obtained by brute-force search (which is not possible in practice).

Based on these considerations we propose

$$\mathcal{L}_{SR} = \sum_{k=1}^N \|\mathbf{C}_k\|^2, \quad (58)$$

with  $\mathbf{C}$  defined in (21), as an LMPIT-inspired detector. In the following section, we will present further numerical results that show that such an LMPIT-inspired detector outperforms the state-of-the-art.

In order to determine a threshold that assures a given probability of false alarm, we utilize again the invariances of the test, specifically, its asymptotic invariance to linear filtering. Similar to the GLRT statistic, we assume, without loss of generality, that under  $\mathcal{H}_0$   $\mathbf{z} \stackrel{N \rightarrow \infty}{\sim} \mathcal{CN}(\mathbf{0}, \mathbf{I}_{2LPN})$ . For this reason we can use numerical simulations with a white process to obtain the threshold under the null hypothesis for any arbitrary noise. Note that the threshold selection is (asymptotically) invariant to the signal-plus-noise PSD at the RC. In the next section we investigate the accuracy of the distribution for different sample sizes.

## VII. NUMERICAL RESULTS

In this section we evaluate the performance of the GLRT and the LMPIT-inspired test using Monte Carlo simulations.<sup>4</sup> According to our model in (1) we generate the CS signal  $\mathbf{s}[n]$  as a QPSK-signal with either a raised-cosine pulse with roll-off factor  $\rho$  or a rectangular pulse. The number of samples per symbol is equal to the cycle period  $P$ . Furthermore, the frequency-selective channels  $\mathbf{H}_s[n]$  and  $\mathbf{H}_r[n]$  are both Rayleigh-fading channels with a delay spread of 10 times the symbol duration and an exponential power delay profile. In each Monte Carlo simulation we draw new realizations of the channels. The independent noises between SC and RC are both colored Gaussian generated with a moving average filter of order 20 and correlated among antennas. This correlation is generated by multiplying the noise realizations with a random matrix with elements drawn from unit complex normals. Moreover, we define the SNRs at SC and RC as

$$\text{SNR}_\star = 10 \log_{10} \left( \frac{\text{tr}(\hat{\mathbf{R}}_\star)}{\text{tr}(\hat{\mathbf{V}}_\star)} \right), \quad (59)$$

where  $\star \in \{s, r\}$  and

$$\hat{\mathbf{R}}_\star = \frac{1}{MNP} \sum_{n=0}^{MNP-1} (\mathbf{H}_\star[n] * \mathbf{s}[n]) (\mathbf{H}_\star[n] * \mathbf{s}[n])^H \in \mathbb{C}^{L \times L} \quad (60)$$

$$\hat{\mathbf{V}}_\star = \frac{1}{MNP} \sum_{n=0}^{MNP-1} \mathbf{v}_\star[n] \mathbf{v}_\star^H[n] \in \mathbb{C}^{L \times L}. \quad (61)$$

Furthermore, we compare the proposed detectors with the following benchmark techniques: The first one is the correlated subspace detector proposed in [18], which employs the following test statistic

$$\mathcal{H} = \prod_{i=1}^{\min(L_I, L)} \frac{1}{1 - k_i^2} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta, \quad (62)$$

where  $k_i$  is the  $i$ th sample canonical correlation between the SC and the RC. The second competitor is the multiantenna extension of the popular cross-correlation detector [14], [18] that uses the statistic

$$\mathcal{C} = |\text{tr}(\mathbf{R}_{sr}^H \mathbf{R}_{sr})| \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta, \quad (63)$$

where

$$\mathbf{R}_{sr} = \frac{1}{MNP} \sum_{n=0}^{MNP-1} \mathbf{u}_s[n] \mathbf{u}_r^H[n], \quad (64)$$

denotes the sample cross covariance matrix of SC and RC. It should be noted that the cross-correlation detector does not require any prior knowledge, whereas the correlated subspace detector needs to know the number of antennas  $L_I$  at the IO, and our proposed techniques also need to know the cycle period  $P$ . Generally, both  $P$  and  $L_I$  could be estimated or they may be known from the standards used by the IO.

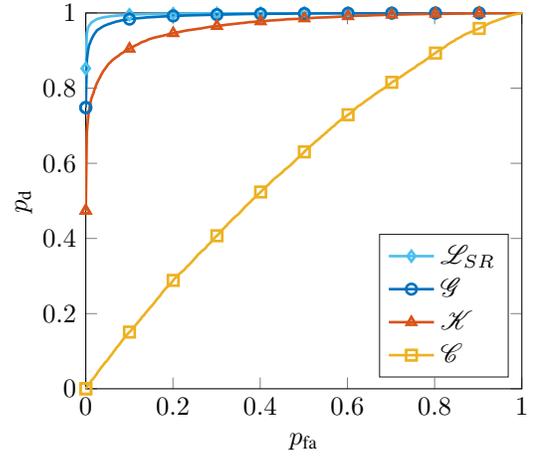


Fig. 4. ROC curves in a scenario with  $P = 2$ ,  $N = 64$ ,  $L = L_I = 4$ ,  $M = 20$ , a rectangular pulse,  $\text{SNR}_s = -15$  dB and  $\text{SNR}_r = -5$  dB.

To evaluate the performance of the proposed detectors, we first choose a scenario with  $P = 2$ ,  $N = 64$ ,  $L = L_I = 4$ ,  $M = 20$ , and a rectangular pulse, Figure 4 shows the receiver operating characteristic (ROC) for  $\text{SNR}_s = -15$  dB at the SC and  $\text{SNR}_r = -5$  dB at the RC. As can be seen, the proposed detectors outperform the competing techniques. We observe that the LMPIT-inspired detector performs better than the GLRT, while the cross-correlation detector performs little better than chance.

Figure 5 depicts the probability of detection versus the  $\text{SNR}_s$  for  $\text{SNR}_r = 0$  in the top plot and  $\text{SNR}_r = -5$  in the bottom plot. The remaining parameters are chosen as  $P = 4$ ,  $N = 128$ ,  $L = L_I = 2$ ,  $M = 20$ , a rectangular pulse, and  $p_{fa} = 0.01$ .<sup>5</sup> Again we can observe that the proposed detectors outperform the competing techniques. In the  $\text{SNR}_s$  range of practical interest, the performance of the LMPIT-inspired test is better than that of the GLRT. It is also shown that the performance drop due to decreasing  $\text{SNR}_r$  is smallest for the LMPIT-inspired test and the GLRT whereas it is largest for the cross-correlation detector.

For another scenario with  $P = 3$ ,  $N = 128$ ,  $L = L_I = 2$ ,  $M = 20$  we study the influence of the pulse shape, i.e. the amount of cyclostationarity present in the signal. A signal with raised-cosine pulse with  $\rho > 0$  has a non-zero cyclic PSD only for the cycle frequency  $\pm 2\pi/P$  and on the stationary manifold (for  $\rho = 0$  it is only non-zero on the stationary manifold), whereas the PSD of a rectangular pulse shaped signal is non-zero for all harmonics of the cycle frequency [49]. Figure 6 shows the ROC for an  $\text{SNR}_s = -15$  dB at the SC and  $\text{SNR}_r = -15$  dB at the RC for  $\rho = \{0, 0.5, 1\}$  and a rectangular pulse shape. As can be seen, detection performance increases with the amount of cyclostationarity present. Specifically, we can observe best performance for the rectangular pulse and worst performance for  $\rho = 0$ . Note that the detection performance does not drop to zero for  $\rho = 0$  as both proposed detectors also account for the usual cross-coherence between RC and

<sup>4</sup>Matlab code is available for download from: <https://github.com/SSTGroup/Cyclostationary-Signal-Processing>

<sup>5</sup>Note that in a passive radar scenario  $p_{fa}$  would be a few orders of magnitude smaller.

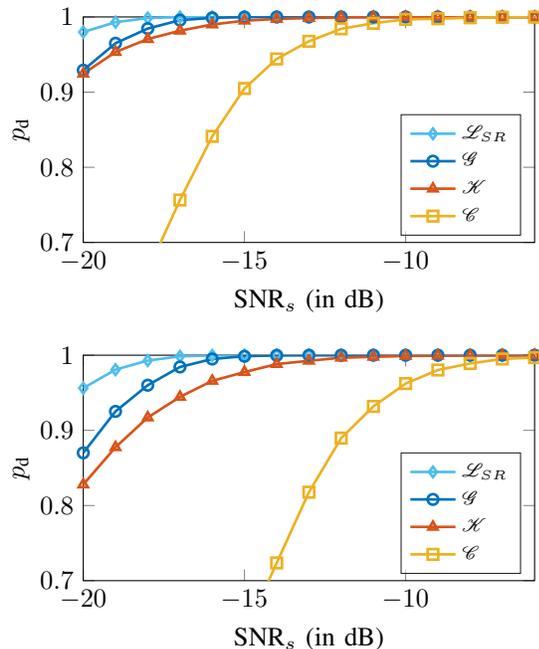


Fig. 5. Probability of detection vs.  $\text{SNR}_s$ , where the top plot shows the results for  $\text{SNR}_r = 0$  dB and the bottom plot for  $\text{SNR}_r = -5$  dB for the following remaining parameters  $P = 4$ ,  $N = 128$ ,  $L = L_I = 2$ ,  $M = 20$ , a rectangular pulse, and  $p_{fa} = 0.01$ .

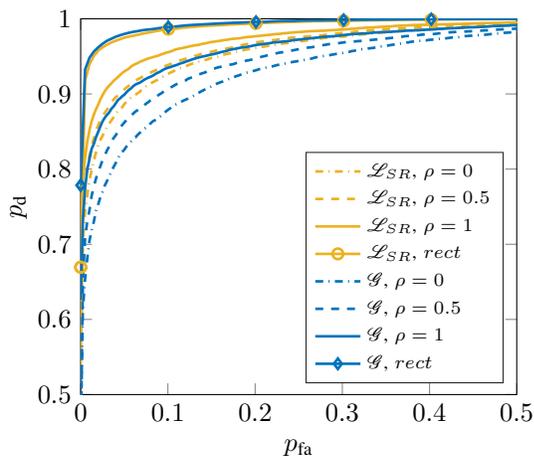


Fig. 6. ROC curves for roll-off factors  $\rho = \{0, 0.5, 1\}$  and a rectangular pulse shaping in a scenario with  $P = 3$ ,  $N = 128$ ,  $L = L_I = 2$ ,  $M = 20$ ,  $\rho = 0.9$ , and  $\text{SNR}_s = \text{SNR}_r = -15$  dB.

SC components on the stationary manifold as can be seen in equations (56) and (57).

Now we will investigate the influence of the particular choice of  $N$  and  $M$  on the detection performance. We should note that  $N$  influences the spectral resolution, i.e., the bias of the estimates, and  $M$  determines the variance of the estimates. Hence, the choice of  $N$  and  $M$  is a bias-variance trade-off, which was already studied in [50] for a related problem. Figure 7 shows the probability of detection versus the total number of samples  $NM$  for the GLRT and the LMPIT-inspired detector for two different choices of  $N$ , namely,  $N = 16$  and  $N = 128$ . If the total number of samples  $NM$  is rather small, we should

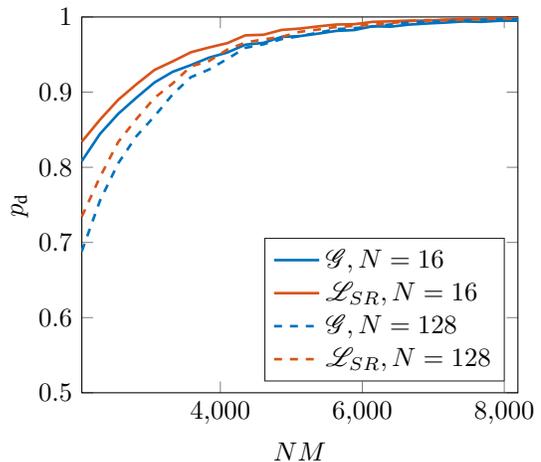


Fig. 7. Probability of detection for  $N = 16$  and  $N = 128$  for different number of samples for  $P = 2$ ,  $L = L_I = 2$ , a rectangular pulse,  $\text{SNR}_s = -18$  dB and  $\text{SNR}_r = -12$  dB.

sacrifice spectral resolution by choosing a smaller  $N$  in order to decrease the variance of the estimate with a larger  $M$ . On the other hand, if a larger number of samples is available, we may choose a larger  $N$  to increase the spectral resolution.

Finally, we examine the accuracy of the distributions under the null hypothesis obtained for the GLRT and the LMPIT-inspired detector. The top plots in Figures 8 and 9 show the distribution of the logarithm of the product of beta random variables and compare it to (i) the distributions obtained numerically with white noise realizations and (ii) the distribution obtained under  $\mathcal{H}_0$ , for  $N = 32$  (Figure 8) and  $N = 128$  (Figure 9). As can be observed, the GLRs for white noise and the product of beta random variables are an accurate match independently of  $N$ , which is not to our surprise since the product of beta random variables is derived for white noise. Either distribution is a reasonably good, albeit not perfect, match for the actual distribution under  $H_0$  for  $N = 32$ , and a very good match for  $N = 128$ . Similar observations can be made for distributions under the null hypothesis of the LMPIT-inspired, which are shown in the bottom plots in Figures 8 and 9.

## VIII. CONCLUSION

In this paper we derived the GLRT for a two-channel passive detection problem by exploiting cyclostationarity. We also examined the existence of optimal invariant tests for this problem. As it turned out that neither the UMPIT nor the LMPIT exists, we proposed an LMPIT-inspired detector. Both detectors, GLRT and LMPIT-inspired, are functions of a cyclic cross-coherence function, but only the GLRT accounts for the cyclic coherence at the SC.

Possible future extensions of our work might consider the case of unknown cycle frequencies as well as almost-cyclostationary signals. Further extensions could be to remove the idealized assumptions of complete cancellation of direct-path interference and clutter at the SC and multipath propagation and clutter at the RC.

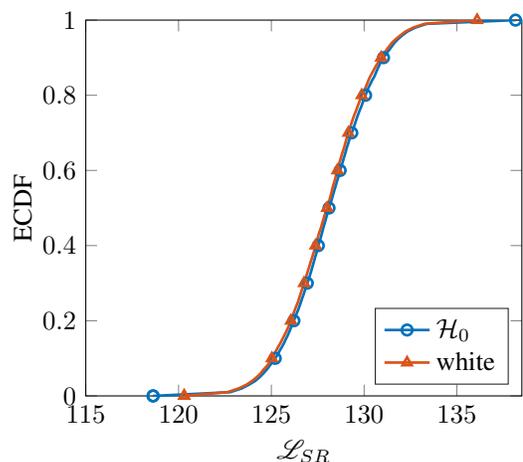
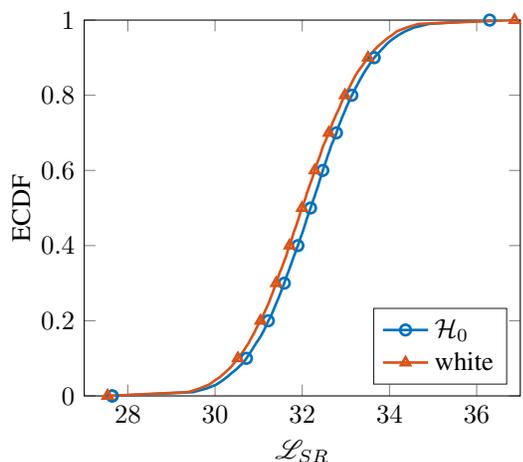
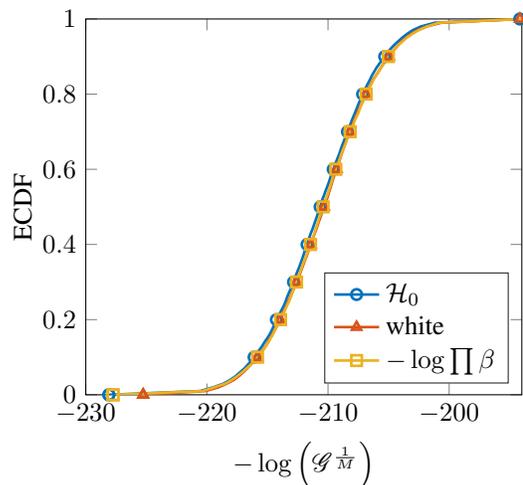
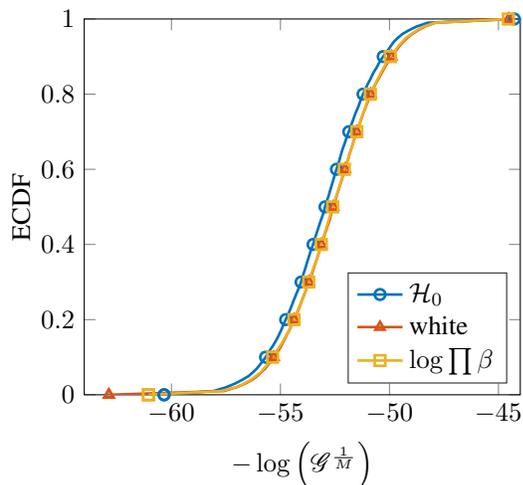


Fig. 8. ECDF of the test statistics under  $\mathcal{H}_0$  and for white noise for the GLRT (top) and the LMPIT-inspired test (bottom) for a scenario with  $P = 2$ ,  $N = 32$ ,  $M = 16$ ,  $L = L_I = 2$ . The top figure also displays the approximation as a product of beta random variables.

Fig. 9. Same as Figure 8, except for  $P = 2$ ,  $N = 128$ ,  $M = 16$ ,  $L = L_I = 2$ .

#### APPENDIX A PROOF OF THEOREM 1

The ML estimate  $\hat{\mathbf{S}}_0$  can be easily found considering the block-diagonal structure of the covariance matrix under  $\mathcal{H}_0$ . With results from complex-valued matrix differentiation [51], the ML estimate is given by

$$\hat{\mathbf{S}}_0 = \begin{bmatrix} \text{diag}_L(\mathbf{Q}_s) & \mathbf{0} \\ \mathbf{0} & \text{diag}_{LP}(\mathbf{Q}_r) \end{bmatrix}. \quad (65)$$

In order to find the ML estimate under  $\mathcal{H}_1$  we note that the permutation of the elements in  $\mathbf{w}$ , given by

$$\tilde{\mathbf{w}} = \mathbf{T}\mathbf{w}, \quad (66)$$

where  $\mathbf{T} = (\mathbf{L}_{2NP, NP} \otimes \mathbf{I}_L)$ , yields a block-Toeplitz structured covariance matrix of  $\tilde{\mathbf{w}}$  with block size  $2LP$ . This is easily shown by noticing that  $\tilde{\mathbf{w}}$  contains the samples  $\mathbf{u}_s[n]$  and  $\mathbf{u}_r[n]$  in alternating order and considering that  $[\mathbf{u}_s[n]^T \ \mathbf{u}_r[n]^T]^T \in \mathbb{C}^{2L}$  is a  $2L$ -variate CS process with cycle period  $P$ . Again this block-Toeplitz covariance matrix

can be approximated by a block-circulant matrix, and the latter can be block-diagonalized by the transformation

$$\tilde{\mathbf{z}} = (\mathbf{L}_{NP, N} \otimes \mathbf{I}_{2L})(\mathbf{F}_{NP} \otimes \mathbf{I}_{2L})^H \tilde{\mathbf{w}}, \quad (67)$$

i.e., the covariance matrix  $\tilde{\mathbf{S}}_1 = \mathbb{E}[\tilde{\mathbf{z}}\tilde{\mathbf{z}}^H | \mathcal{H}_1] \in \mathbb{S}_{2LP}^{2LNP}$  is block-diagonal with block size  $2LP$ . Exploiting properties of the Kronecker product we can rewrite (67) as

$$\tilde{\mathbf{z}} = [(\mathbf{L}_{NP, N} \mathbf{F}_{NP}^H \otimes \mathbf{I}_2) \otimes \mathbf{I}_L] \tilde{\mathbf{w}}. \quad (68)$$

Considering (10), the linear transformation of  $\mathbf{w}$  is given by

$$\mathbf{z} = [(\mathbf{I}_2 \otimes \mathbf{L}_{NP, N} \mathbf{F}_{NP}^H) \otimes \mathbf{I}_L] \mathbf{w}. \quad (69)$$

It can be observed that (68) and (69) are equal up to the commutation of the Kronecker product inside the parentheses. We should further notice that the matrix  $\mathbf{T}$  commutes with that product. After putting these pieces together,  $\tilde{\mathbf{z}}$  and  $\mathbf{z}$  are also related by the linear transformation  $\mathbf{T}$  as

$$\tilde{\mathbf{z}} = \mathbf{T}\mathbf{z}. \quad (70)$$

$$\mathcal{G}_M^{\frac{1}{M}} = \frac{\det(\hat{\mathbf{S}}_1)}{\det(\hat{\mathbf{S}}_0)} = \frac{\det(\text{diag}_{LP}(\mathbf{Q}_s) - \text{diag}_{LP}(\mathbf{Q}_{sr}) \text{diag}_{LP}(\mathbf{Q}_r)^{-1} \text{diag}_{LP}(\mathbf{Q}_{rs}))}{\det(\text{diag}_L(\mathbf{Q}_s))} = \prod_{k=1}^N \det(\mathbf{D}_k - \mathbf{C}_k \mathbf{C}_k^H) \quad (74)$$

Hence, similar to  $\tilde{\mathbf{w}}$ ,  $\tilde{\mathbf{z}}$  contains  $\mathbf{u}_s(\theta_n)$  and  $\mathbf{u}_r(\theta_n)$  in alternating order. As  $\hat{\mathbf{S}}_1$  is block-diagonal, we can easily find its ML estimate as

$$\hat{\mathbf{S}}_1 = \text{diag}_{2LP}(\tilde{\mathbf{Q}}), \quad (71)$$

where  $\tilde{\mathbf{Q}} = \mathbf{T} \mathbf{Q} \mathbf{T}^T$ . After exploiting the invariance of the ML estimate [47], we find

$$\hat{\mathbf{S}}_1 = \mathbf{T}^T \hat{\mathbf{S}}_1 \mathbf{T} = \mathbf{T}^T \text{diag}_{2LP}(\tilde{\mathbf{Q}}) \mathbf{T}. \quad (72)$$

In order to express this as a function of the sample covariance matrix  $\mathbf{Q}$ , let us study the effect of the permutation. The  $(k, l)$ th  $L \times L$  block of  $\hat{\mathbf{S}}_1$  with  $k = mNP + i$  and  $l = nNP + j$  for  $m, n = 0, 1$  and  $i, j = 0, \dots, NP - 1$  is shifted to the  $(k', l')$ th entry in  $\hat{\mathbf{S}}_1$  with  $k' = 2i + m$  and  $l' = 2j + n$ . Applying the permutation to every element, (72) can be expressed as a function of  $\mathbf{Q}$  as

$$\hat{\mathbf{S}}_1 = \begin{bmatrix} \text{diag}_{LP}(\mathbf{Q}_s) & \text{diag}_{LP}(\mathbf{Q}_{sr}) \\ \text{diag}_{LP}(\mathbf{Q}_{rs}) & \text{diag}_{LP}(\mathbf{Q}_r) \end{bmatrix}. \quad (73)$$

Now we plug the ML estimates (65) and (73) into the likelihood ratio (17) to obtain (74) as shown at the top of this page, where  $\mathbf{D}$  and  $\mathbf{C}$  are given by (20) and (21), respectively. In this expression, we exploited the fact that the determinant of a block-diagonal matrix is equal to the product of the determinants of the single blocks, and the expression for the determinant of a  $2 \times 2$  block matrix with invertible blocks.

## APPENDIX B PROOF OF PROPOSITION 1

Let us define the matrices  $\mathbf{X}[n]$  and  $\mathbf{Y}[n]$  as the concatenation of all realizations  $M$  of  $\mathbf{x}[n]$  and  $\mathbf{y}[n]$  given by equations (12) and (13) as

$$\mathbf{X}[n] = [\mathbf{x}_1[n] \cdots \mathbf{x}_M[n]] \in \mathbb{C}^{LP \times M}, \quad (75)$$

$$\mathbf{Y}[n] = [\mathbf{y}_1[n] \cdots \mathbf{y}_M[n]] \in \mathbb{C}^{LP \times M}. \quad (76)$$

Moreover, we define

$$\mathbf{U}_p[n] = [\mathbf{u}_{s1}(\theta_{pN+n}) \cdots \mathbf{u}_{sM}(\theta_{pN+n})] \in \mathbb{C}^{L \times M}, \quad (77)$$

$$\mathbf{V}_p[n] = [\mathbf{u}_{r1}(\theta_{pN+n}) \cdots \mathbf{u}_{rM}(\theta_{pN+n})] \in \mathbb{C}^{L \times M}, \quad (78)$$

for  $p = 0, \dots, P - 1$ , and the  $l$ th rows for  $l = 1, \dots, L$  of these matrices are referred to as  $\mathbf{u}_p^{(l)}[n] \in \mathbb{C}^{1 \times M}$  and  $\mathbf{v}_p^{(l)}[n] \in \mathbb{C}^{1 \times M}$ , respectively. Now  $\mathbf{X}[n]$  is partitioned as

$$\mathbf{X}[n] = \begin{bmatrix} \mathbf{X}_p[n] \\ \mathbf{X}_p^{(l)}[n] \\ \mathbf{u}_p^{(l)}[n] \\ \vdots \end{bmatrix}, \quad (79)$$

where

$$\mathbf{X}_p[n] = [\mathbf{U}_0^T[n] \cdots \mathbf{U}_{p-1}^T[n]]^T \in \mathbb{C}^{LP \times M}, \quad (80)$$

$$\mathbf{X}_p^{(l)}[n] = [\mathbf{u}_p^{(1)T}[n] \cdots \mathbf{u}_p^{(l-1)T}[n]]^T \in \mathbb{C}^{(l-1) \times M}. \quad (81)$$

Equivalently, we can partition  $\mathbf{Y}[n]$  into  $\mathbf{Y}_p[n]^{LP \times M}$ ,  $\mathbf{Y}_p^{(l)}[n] \in \mathbb{C}^{(l-1) \times M}$ , and  $\mathbf{v}_p^{(l)}[n] \in \mathbb{C}^{1 \times M}$ , and define

$$\bar{\mathbf{Y}}_p^{(l)}[n] = \begin{bmatrix} \mathbf{Y}_p[n] \\ \mathbf{Y}_p^{(l)}[n] \end{bmatrix}. \quad (82)$$

Let us now provide some generic projection matrices, which will be used subsequently:

$$\mathbf{P}_A^\perp = \mathbf{I} - \mathbf{A}^H [\mathbf{A} \mathbf{A}^H]^{-1} \mathbf{A}, \quad (83)$$

$$\mathbf{P}_{\mathbf{A}\mathbf{B}}^\perp = \mathbf{I} - [\mathbf{A}^H \ \mathbf{B}^H] \left( \begin{bmatrix} \mathbf{A} \\ \mathbf{B} \end{bmatrix} [\mathbf{A}^H \ \mathbf{B}^H] \right)^{-1} \begin{bmatrix} \mathbf{A} \\ \mathbf{B} \end{bmatrix}, \quad (84)$$

$$\mathbf{P}_{\mathbf{B}\mathbf{P}_A^\perp} = \mathbf{P}_A^\perp \mathbf{B}^H (\mathbf{B} \mathbf{P}_A^\perp \mathbf{B}^H)^{-1} \mathbf{B} \mathbf{P}_A^\perp, \quad (85)$$

for some matrices  $\mathbf{A}$  and  $\mathbf{B}$  of suitable dimensions. With these considerations we can rewrite the likelihood ratio given in (74) in terms of the Gram matrices as<sup>6</sup>

$$\frac{\det(\hat{\mathbf{S}}_1)}{\det(\hat{\mathbf{S}}_0)} = \frac{\prod_{n=1}^N \det(\mathbf{Z}\mathbf{Z}^H)}{\prod_{n=1}^N \prod_{p=0}^{P-1} \det(\mathbf{U}_p \mathbf{U}_p^H) \det(\mathbf{Y}\mathbf{Y}^H)}, \quad (86)$$

where  $\mathbf{Z} = [\mathbf{X}^T \ \mathbf{Y}^T]^T$  and in the numerator we have exploited the permutation invariance of the determinant. As shown in [47], we can decompose the determinants into products of scalars to obtain

$$\det(\mathbf{U}_p \mathbf{U}_p^H) = \prod_{l=1}^L \mathbf{u}_p^{(l)} \mathbf{P}_{\mathbf{X}_p^{(l)}}^\perp \mathbf{u}_p^{(l)H} \quad (87)$$

$$\det(\mathbf{Y}\mathbf{Y}^H) = \prod_{p=0}^{P-1} \prod_{l=1}^L \mathbf{v}_p^{(l)} \mathbf{P}_{\bar{\mathbf{Y}}_p^{(l)}}^\perp \mathbf{v}_p^{(l)H} \quad (88)$$

$$\det(\mathbf{Z}\mathbf{Z}^H) = \prod_{p=0}^{P-1} \prod_{l=1}^L \mathbf{u}_p^{(l)} \mathbf{P}_{\mathbf{X}_p \mathbf{X}_p^{(l)}}^\perp \mathbf{u}_p^{(l)H} \mathbf{v}_p^{(l)} \mathbf{P}_{\mathbf{X} \bar{\mathbf{Y}}_p^{(l)}}^\perp \mathbf{v}_p^{(l)H}. \quad (89)$$

We can further decompose these matrices as [46]

$$\mathbf{P}_{\mathbf{X}_p^{(l)}}^\perp = \mathbf{P}_{\mathbf{X}_p \mathbf{X}_p^{(l)}}^\perp + \mathbf{P}_{\mathbf{X}_p} \mathbf{P}_{\mathbf{X}_p^{(l)}}^\perp, \quad (90)$$

$$\mathbf{P}_{\bar{\mathbf{Y}}_p^{(l)}}^\perp = \mathbf{P}_{\mathbf{X} \bar{\mathbf{Y}}_p^{(l)}}^\perp + \mathbf{P}_{\mathbf{X} \mathbf{P}_{\bar{\mathbf{Y}}_p^{(l)}}^\perp}. \quad (91)$$

<sup>6</sup>We drop the index  $[n]$  for notational convenience.

Finally, we can plug in (87), (88), and (89) into (86) to obtain

$$\mathcal{G}_M^{\frac{1}{M}} = \prod_{n=1}^N \prod_{p=0}^{P-1} \prod_{l=1}^L \frac{\mathbf{u}_p^{(l)} \mathbf{P}_{\mathbf{X}_p \mathbf{X}_p}^\perp \mathbf{u}_p^{(l)H}}{\mathbf{u}_p^{(l)} \left( \mathbf{P}_{\mathbf{X}_p \mathbf{X}_p}^\perp + \mathbf{P}_{\mathbf{X}_p \mathbf{P}_{\mathbf{X}_p}^\perp} \right) \mathbf{u}_p^{(l)H}} \times \frac{\mathbf{v}_p^{(l)} \mathbf{P}_{\mathbf{X} \bar{\mathbf{Y}}_p}^\perp \mathbf{v}_p^{(l)H}}{\mathbf{v}_p^{(l)} \left( \mathbf{P}_{\mathbf{X} \bar{\mathbf{Y}}_p}^\perp + \mathbf{P}_{\mathbf{X} \mathbf{P}_{\bar{\mathbf{Y}}_p}^\perp} \right) \mathbf{v}_p^{(l)H}}. \quad (92)$$

In order to characterize the distribution of the last expression under the null hypothesis, we will again exploit the invariances of the likelihood ratio. Under  $\mathcal{H}_0$  the vector  $\mathbf{z}$  can always be pre-whitened and without loss of generality, we can assume that  $\mathbf{z} \stackrel{N \rightarrow \infty}{\sim} \mathcal{CN}(\mathbf{0}, \mathbf{I}_{2LPN})$  or, since we consider  $M$  i.i.d. observations,  $\mathbf{u}_p^{(l)} \stackrel{N \rightarrow \infty}{\sim} \mathcal{CN}(\mathbf{0}, \mathbf{I}_M)$  and  $\mathbf{v}_p^{(l)} \stackrel{N \rightarrow \infty}{\sim} \mathcal{CN}(\mathbf{0}, \mathbf{I}_M)$ . Therefore, each of the quadratic terms in (92) is chi-squared distributed with degrees of freedom equal to two times the rank of the projection matrices involved in each of the quadratic terms. Specifically, the ranks of these matrices are given by

$$\alpha_{lp} = \text{rank}(\mathbf{P}_{\mathbf{X}_p \mathbf{X}_p}^\perp) = M - (Lp + l - 1), \quad (93)$$

$$\alpha_p = \text{rank}(\mathbf{P}_{\mathbf{X}_p \mathbf{P}_{\mathbf{X}_p}^\perp}) = Lp, \quad (94)$$

$$\beta_{lp} = \text{rank}(\mathbf{P}_{\mathbf{X} \bar{\mathbf{Y}}_p}^\perp) = M - (LP + Lp + l - 1), \quad (95)$$

$$\beta = \text{rank}(\mathbf{P}_{\mathbf{X} \mathbf{P}_{\bar{\mathbf{Y}}_p}^\perp}) = LP. \quad (96)$$

Finally, considering that the ratio  $\frac{A}{A+B} \sim \text{Beta}(\gamma_1/2, \gamma_2/2)$  if  $A \sim \chi_{\gamma_1}^2$  and  $B \sim \chi_{\gamma_2}^2$ , the proof follows.

#### APPENDIX C PROOF OF LEMMA 1

First we observe that the terms  $\det(\mathbf{S}_0)^{-M}$  and  $\det(\mathbf{S}_1)^{-M}$  in (30) neither depend on the observations nor the invariances. Hence, they can be discarded in the ratio. Secondly, let us focus on the denominator of the ratio, specifically, on the exponential term. Taking into account that  $\Psi$  and  $\mathbf{S}_0$  are block-diagonal with two blocks, the denominator can be simplified as

$$\text{tr}(\Psi \mathbf{Q} \Psi^H \mathbf{S}_0^{-1}) = \text{tr}(\mathbf{P}_s \mathbf{G} \mathbf{Q}_s \mathbf{G}^H \mathbf{P}_s^T \mathbf{S}_s^{-1}) + \text{tr}(\mathbf{P}_r \mathbf{H} \mathbf{Q}_r \mathbf{H}^H \mathbf{P}_r^T \mathbf{S}_r^{-1}). \quad (97)$$

Applying the change of variables  $\mathbf{G} \rightarrow \mathbf{G} \mathbf{B}_s^{-\frac{1}{2}}$  and  $\mathbf{H} \rightarrow \mathbf{H} \mathbf{B}_r^{-\frac{1}{2}}$ , where  $\mathbf{B}_s = \text{diag}_L(\mathbf{Q}_s)$  and  $\mathbf{B}_r = \text{diag}_{LP}(\mathbf{Q}_r)$ , yields

$$\text{tr}(\Psi \mathbf{Q} \Psi^H \mathbf{S}_0^{-1}) = \text{tr}(\mathbf{P}_s \mathbf{G} \mathbf{B}_s^{-\frac{1}{2}} \mathbf{Q}_s \mathbf{B}_s^{-\frac{1}{2}} \mathbf{G}^H \mathbf{P}_s^T \mathbf{S}_s^{-1}) + \text{tr}(\mathbf{P}_r \mathbf{H} \mathbf{B}_r^{-\frac{1}{2}} \mathbf{Q}_r \mathbf{B}_r^{-\frac{1}{2}} \mathbf{H}^H \mathbf{P}_r^T \mathbf{S}_r^{-1}). \quad (98)$$

We should note that the data-dependent terms  $\mathbf{B}_s^{-\frac{1}{2}} \mathbf{Q}_s \mathbf{B}_s^{-\frac{1}{2}}$  and  $\mathbf{B}_r^{-\frac{1}{2}} \mathbf{Q}_r \mathbf{B}_r^{-\frac{1}{2}}$  are whitened on their main diagonal blocks, i.e., these are given by  $\mathbf{I}_L$  and  $\mathbf{I}_{LP}$ , respectively. These whitened main diagonal blocks are the only blocks of those matrices involved in the trace operations since the other

matrices are block-diagonal, in the first trace operator with block size  $L \times L$  and in the second trace operator with block size  $LP \times LP$ . For this reason the denominator can be discarded in the ratio (30).

Hence,  $\mathcal{L}$  simplifies as

$$\mathcal{L} \propto \sum_{\mathbf{v}_0^N, \mathbf{v}_1^N, \mathbf{U}} \int_{\mathbb{G}} \int_{\mathbb{H}} |\det(\mathbf{G})|^{2M} |\det(\mathbf{H})|^{2M} \times \exp\left\{-M \text{tr}(\Psi \mathbf{Q} \Psi^H \mathbf{S}_1^{-1})\right\} d\mathbf{G} d\mathbf{H}. \quad (99)$$

In order to further reduce this expression, we first consider the structure of the inverse covariance matrix  $\mathbf{S}_1^{-1}$ . Let us define the matrices  $\Sigma = \mathbf{S}_1^{-1}$  and  $\bar{\Psi} = \Psi \mathbf{Q} \Psi^H$ , where these matrices can be partitioned as  $\Sigma = \begin{bmatrix} \Sigma_1 & \Sigma_{12} \\ \Sigma_{21} & \Sigma_2 \end{bmatrix}$  and  $\bar{\Psi} = \begin{bmatrix} \bar{\Psi}_1 & \bar{\Psi}_{12} \\ \bar{\Psi}_{21} & \bar{\Psi}_2 \end{bmatrix}$ . Note that each of the four  $LNP$ -sized blocks in  $\Sigma$  are block-diagonal with block size  $LP$ . Furthermore, after another change of variables  $\mathbf{G} \rightarrow \mathbf{G} \mathbf{B}_s^{-\frac{1}{2}}$  and  $\mathbf{H} \rightarrow \mathbf{H} \mathbf{B}_r^{-\frac{1}{2}}$ , the blocks of  $\bar{\Psi}$  are given by  $\bar{\Psi}_1 = \mathbf{P}_s \mathbf{G} \mathbf{B}_s^{-\frac{1}{2}} \mathbf{Q}_s \mathbf{B}_s^{-\frac{1}{2}} \mathbf{G}^H \mathbf{P}_s^H$ ,  $\bar{\Psi}_2 = \mathbf{P}_r \mathbf{H} \mathbf{B}_r^{-\frac{1}{2}} \mathbf{Q}_r \mathbf{B}_r^{-\frac{1}{2}} \mathbf{H}^H \mathbf{P}_r^H$  and  $\bar{\Psi}_{12} = \bar{\Psi}_{21} = \mathbf{P}_s \mathbf{G} \mathbf{B}_s^{-\frac{1}{2}} \mathbf{Q}_{sr} \mathbf{B}_r^{-\frac{1}{2}} \mathbf{H}^H \mathbf{P}_r^H$ . Finally, putting things together, (99) becomes

$$\mathcal{L} \propto \sum_{\mathbf{v}_0^N, \mathbf{v}_1^N, \mathbf{U}} \int_{\mathbb{G}} \int_{\mathbb{H}} |\det(\mathbf{G})|^{2M} |\det(\mathbf{H})|^{2M} \times e^{-M \text{tr}(\bar{\Psi}_1 \Sigma_1)} e^{-2M \text{tr}(\bar{\Psi}_{12} \Sigma_{21})} e^{-M \text{tr}(\bar{\Psi}_2 \Sigma_2)} d\mathbf{G} d\mathbf{H}. \quad (100)$$

In order to further disentangle this expression, we consider the traces in the exponential terms individually. Introducing the change of variables  $\mathbf{G} \rightarrow \mathbf{P}_s^T \text{diag}_L(\Sigma_1)^{-\frac{1}{2}} \mathbf{P}_s \mathbf{G}$  and considering that the trace is given by the sum of the diagonal elements,  $\text{tr}(\bar{\Psi}_1 \Sigma_1)$  simplifies to

$$\text{tr}(\bar{\Psi}_1 \Sigma_1) = \text{tr}(\mathbf{G} \mathbf{G}^H) + \sum_{k=1}^N \sum_{\substack{m,n=1 \\ m \neq n}}^P \text{tr}(\Gamma_k^{(m,n)} \mathbf{G}_k^{(n,n)} \mathbf{D}_k^{(n,m)} \mathbf{G}_k^{(m,m)H}), \quad (101)$$

where we also considered that both  $\Gamma$  and  $\mathbf{D}$ , given by (34) and (20), respectively, are whitened on their  $L \times L$  main diagonal and  $\mathbf{G} \in \mathbb{G}$ . It should be noted that (101) depends on unknown parameters through  $\Gamma$ .

The second exponential term in (100) can be reduced by introducing the change of variables  $\mathbf{H} \rightarrow \mathbf{P}_r^T \Sigma_2^{-\frac{1}{2}} \mathbf{P}_r \mathbf{H}$  as

$$\text{tr}(\bar{\Psi}_2 \Sigma_2) = \sum_{k=1}^N \text{tr}(\Lambda_k \mathbf{G}_k \mathbf{C}_k \mathbf{H}_k^H), \quad (102)$$

where  $\Lambda$  is given by (36).

Finally, by plugging in the previous change of variables, the last exponential term in (100) becomes

$$\text{tr}(\bar{\Psi}_2 \Sigma_2) = \text{tr}(\Sigma_2^{-\frac{1}{2}} \mathbf{P}_r \mathbf{H} \mathbf{B}_r^{-\frac{1}{2}} \mathbf{Q}_r \mathbf{B}_r^{-\frac{1}{2}} \mathbf{H}^H \mathbf{P}_r^T \Sigma_2^{-\frac{1}{2}} \Sigma_2) = \text{tr}(\mathbf{H} \mathbf{H}^H), \quad (103)$$

$$\begin{aligned} \mathcal{L}_3 &= \frac{M^2}{2} \sum_{\mathbf{V}_0^N, \mathbf{V}_1^N, \mathbf{U}} \int_{\mathbf{G}} \beta(\mathbf{G}) \left[ \sum_{k=1}^N \sum_{\substack{m,n=1 \\ m \neq n}}^P \text{tr} \left( \mathbf{\Gamma}_k^{(m,n)} \mathbf{G}_k^{(n,n)} \mathbf{D}_k^{(n,m)} \mathbf{G}_k^{(m,m)H} \right) \right]^2 d\mathbf{G} \int_{\mathbf{H}} \beta(\mathbf{H}) d\mathbf{H} \\ &\propto \sum_{\mathbf{V}_0^N, \mathbf{V}_1^N, \mathbf{U}} \int_{\mathbf{G}} \beta(\mathbf{G}) \sum_{k=1}^N \sum_{\substack{m,n=1 \\ m \neq n}}^P \text{tr}^2 \left( \mathbf{\Gamma}_k^{(m,n)} \mathbf{G}_k^{(n,n)} \mathbf{D}_k^{(n,m)} \mathbf{G}_k^{(m,m)H} \right) d\mathbf{G} \end{aligned} \quad (104)$$

$$\begin{aligned} \mathcal{L}_5 &= \frac{M^2}{2} \sum_{\mathbf{V}_0^N, \mathbf{V}_1^N, \mathbf{U}} \int_{\mathbf{G}} \int_{\mathbf{H}} \beta(\mathbf{G}) \beta(\mathbf{H}) \left[ \sum_{i,j,l=1}^P \text{tr} \left( \mathbf{\Lambda}_k^{(i,j)} \mathbf{G}_k^{(j,j)} \mathbf{C}_k^{(j,l)} \mathbf{H}_k^{(l,i)H} \right) \right]^2 d\mathbf{G} d\mathbf{H} \\ &\propto \sum_{\mathbf{V}_0^N, \mathbf{V}_1^N, \mathbf{U}} \int_{\mathbf{G}} \int_{\mathbf{H}} \beta(\mathbf{G}) \beta(\mathbf{H}) \sum_{i,j,l=1}^P \text{tr}^2 \left( \mathbf{\Lambda}_k^{(i,j)} \mathbf{G}_k^{(j,j)} \mathbf{C}_k^{(j,l)} \mathbf{H}_k^{(l,i)H} \right) d\mathbf{G} d\mathbf{H} \end{aligned} \quad (106)$$

where the last simplification follows from the fact that  $\mathbf{H}$  is block-diagonal with block size  $LP$  and  $\mathbf{B}_r^{-\frac{1}{2}} \mathbf{Q}_r \mathbf{B}_r^{-\frac{1}{2}}$  is white on its  $LP$ -sized main diagonal blocks. The proof follows by plugging (101), (102), and (103) into (100).

#### APPENDIX D PROOF OF THEOREM 2

Let us first focus on (41), which can be simplified in (104) at the top of this page, where the integrals involving the cross-terms of the square, i.e., those elements of the sum that are not multiplied by themselves, are zero since they are equal to their opposites as can be seen by applying the change of variables  $\mathbf{G}_k^{(n,n)} \rightarrow -\mathbf{G}_k^{(n,n)}$ . Now (104) became the same expression as in Appendix C in [27] and we can simplify it in the same way to obtain

$$\mathcal{L}_3 \propto \mathcal{L}_S = \sum_{k=1}^N \|\mathbf{D}_k\|^2, \quad (105)$$

where  $\mathbf{D}$  is given by (20). Secondly, we can reduce (43) in (106) at the top of this page, where the cross-terms of the square cancel out by another change of variables  $\mathbf{G}_k^{(j,j)} \rightarrow -\mathbf{G}_k^{(j,j)}$  or  $\mathbf{H}_k^{(l,i)} \rightarrow -\mathbf{H}_k^{(l,i)}$ , respectively. Finally, following similar steps as in Appendix C of [27], we obtain

$$\mathcal{L}_5 \propto \mathcal{L}_{SR} = \sum_{k=1}^N \|\mathbf{C}_k\|^2. \quad (107)$$

It should be noted that  $\mathcal{L}_3$  and  $\mathcal{L}_5$  are equal to  $\mathcal{L}_S$  and  $\mathcal{L}_{SR}$  up to constant terms that depend on data-independent but unknown values in  $\mathbf{\Gamma}$  and  $\mathbf{\Lambda}$ . These constant terms are taken into account via one constant  $\gamma$ , which allows us to express (38) as

$$\mathcal{L} \propto \mathcal{L}_S + \gamma \mathcal{L}_{SR}. \quad (108)$$

#### APPENDIX E PROOF OF PROPOSITION 2

Recall that the sample coherence matrix  $\mathbf{D}$  is given by

$$\mathbf{D} = \text{diag}_L(\mathbf{Q}_s)^{-1/2} \text{diag}_{LP}(\mathbf{Q}_s) \text{diag}_L(\mathbf{Q}_s)^{-1/2}, \quad (109)$$

and considering the block-diagonal structure of the matrices, its  $(i, j)$ th  $L \times L$  element in the  $k$ th  $LP \times LP$  block is given by

$$\mathbf{D}_k^{(i,j)} = \left( \mathbf{Q}_{sk}^{(i,i)} \right)^{-1/2} \mathbf{Q}_{sk}^{(i,j)} \left( \mathbf{Q}_{sk}^{(j,j)} \right)^{-1/2}, \quad (110)$$

for  $k = 0, \dots, N-1$  and  $i, j = 0, \dots, P-1$ . These elements of the sample covariance matrix can again be expressed as samples of the cyclic PSDs similar to (54). Hence,  $\mathbf{D}_k^{(i,j)}$  can be written as a function of frequency  $\theta_{iN+k}$  as

$$\begin{aligned} \mathbf{D}_k^{(i,j)} &= \mathbf{D}^{(i-j)}(\theta_{jN+k}) = \left[ \mathbf{\Pi}_{ss}^{(0)} \left( \theta_{jN+k} + \frac{2\pi}{P}q \right) \right]^{-\frac{1}{2}} \\ &\quad \times \mathbf{\Pi}_{ss}^{(i-j)}(\theta_{jN+k}) \left[ \mathbf{\Pi}_{ss}^{(0)}(\theta_{jN+k}) \right]^{-\frac{1}{2}}, \end{aligned} \quad (111)$$

and the proof follows with  $q = i - j$ .

Similarly the  $(i, j)$ th  $L \times L$  element in the  $k$ th  $LP \times LP$  block of the sample cross-coherence matrix

$$\mathbf{C} = \text{diag}_L(\mathbf{Q}_s)^{-1/2} \text{diag}_{LP}(\mathbf{Q}_{sr}) \text{diag}_{LP}(\mathbf{Q}_r)^{-1/2}. \quad (112)$$

can be written as

$$\mathbf{C}_k^{(i,j)} = \left( \mathbf{Q}_{sk}^{(i,i)} \right)^{-1/2} \sum_{l=0}^{P-1} \mathbf{Q}_{srk}^{(i,l)} \left( \mathbf{Q}_{rk}^{(l,j)} \right)^{-1/2}, \quad (113)$$

since the  $k$ th diagonal blocks of  $\mathbf{Q}_{sr}$  and  $\mathbf{Q}_r$  are both full matrices whereas the  $k$ th block of  $\mathbf{Q}_s$  is block-diagonal with block size  $L$ . Again  $\mathbf{C}_k^{(i,j)}$  can be expressed as a function of  $\theta_{jN+k}$  as

$$\begin{aligned} \mathbf{C}_k^{(i,j)} &= \mathbf{C}_k^{(i-j)}(\theta_{jN+k}) = \left[ \mathbf{\Pi}_{ss}^{(0)}(\theta_{jN+k} + \frac{2\pi}{P}q) \right]^{-\frac{1}{2}} \\ &\quad \times \sum_{l=0}^{P-1} \mathbf{\Pi}_{sr}^{(i-l)}(\theta_{lN+k}) \left[ \mathbf{\Pi}_{rr}^{(l-j)}(\theta_{jN+k}) \right]^{-\frac{1}{2}}, \end{aligned} \quad (114)$$

and with  $q = i - j$  the proof follows.

#### REFERENCES

- [1] H. D. Griffiths and C. J. Baker, "Passive coherent location radar systems. Part 1: Performance prediction," *IEE Proc. Radar Sonar Navig.*, vol. 152, no. 3, pp. 153–159, June 2005.
- [2] D. O'Hagan, F. Colone, C. Baker, and H. Griffiths, "Passive bistatic radar (PBR) demonstrator," in *Proc. IET Int. Radar Syst. Conf.*, Edinburgh, UK, 2007.
- [3] R. Zemmari, U. Nickel, and W.-D. Wirth, "GSM passive radar for medium range surveillance," in *Proc. Eur. Radar Conf. (EuRAD)*, Rome, Italy, 2009, pp. 49–52.
- [4] R. Tao, H. Wu, and T. Shan, "Direct-path suppression by spatial filtering in digital television terrestrial broadcasting-based passive radar," *IET Radar, Sonar, Navigation*, vol. 4, no. 6, pp. 791–805, Sept. 2010.
- [5] F. Colone, D. O'Hagan, P. Lombardo, and C. Baker, "A Multistage Processing Algorithm for Disturbance Removal and Target Detection in Passive Bistatic Radar," *IEEE Trans. on Aerospace, Electr. Syst.*, vol. 45, no. 2, pp. 698–722, Apr. 2009.

- [6] J. E. Palmer and S. J. Searle, "Evaluation of adaptive filter algorithms for clutter cancellation in passive bistatic radar," in *IEEE Radar Conference*, Atlanta, GA, USA, 2012, pp. 493–498.
- [7] H. Griffiths and N. Long, "Television-based bistatic radar," *IEE Proc. Radar Sonar Navig.*, vol. 133, no. 7, pp. 649–657, 1986.
- [8] H. Griffiths, A. Garnett, C. Baker, and S. Keaveney, "Bistatic radar using satellite-borne illuminator of opportunity," in *Proc. RADAR*, Edinburgh, UK, 1992, pp. 276–279.
- [9] M. Glende, J. Heckenbach, H. Kuschel, S. Müller, J. Schell, and C. Schumacher, "Experimental passive radar systems using digital illuminators (DAB/DVB-T)," in *Proc. Int. Radar Symp.(IRS)*, Cologne, Germany, 2007.
- [10] P. Howland, D. Maksimiuk, and G. Reitsma, "FM radio based bistatic radar," *IEE Proc. Radar Sonar Navig.*, vol. 152, no. 3, pp. 107–115, June 2005.
- [11] K. Kulpa and Z. Czekala, "Masking effect and its removal in PCL radar," *IEE Proc. Radar Sonar Navig.*, vol. 152, no. 3, pp. 174–178, June 2005.
- [12] F. Colone, R. Cardinali, P. Lombardo, O. Crognale, A. Cosmi, A. Lauri, and T. Bucciarelli, "Space-time constant modulus algorithm for multipath removal on the reference signal exploited by passive bistatic radar," *IET Radar, Sonar, Navigation*, vol. 3, no. 3, pp. 253–264, 2009.
- [13] G. Smith, K. Chetty, C. Baker, and K. Woodbridge, "Extended time processing for passive bistatic radar," *IET Radar, Sonar, Navigation*, vol. 7, no. 9, pp. 1012–1018, Dec. 2013.
- [14] J. Liu, H. Li, and B. Himed, "On the performance of the cross-correlation detector for passive radar applications," *Signal Process.*, vol. 113, pp. 32–37, 2015.
- [15] I. Santamaría, L. L. Scharf, D. Cochran, and J. Vía, "Passive detection of rank-one signals with a multiantenna reference channel," in *Proc. 24th European Signal Proc. Conf. (EUSIPCO)*, Budapest, Hungary, 2016, pp. 140–144.
- [16] H. Wang, Y. Wang, L. L. Scharf, and I. Santamaría, "Canonical correlations for target detection in a passive radar network," in *Proc. 50th Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, USA, 2016, pp. 1159 – 1163.
- [17] I. Santamaría, J. Vía, L. L. Scharf, and Y. Wang, "A GLRT approach for detecting correlated signals in white noise in two MIMO channels," in *Proc. 25th European Signal Proc. Conf. (EUSIPCO)*, Kos, Greece, 2017, pp. 1395–1399.
- [18] I. Santamaría, L. L. Scharf, J. Vía, H. Wang, and Y. Wang, "Passive detection of correlated subspace signals in two MIMO channels," *IEEE Trans. on Signal Process.*, vol. 65, no. 20, pp. 5266–5280, Oct. 2017.
- [19] S. D. Howard and S. Sirianunpiboon, "Passive radar detection using multiple transmitters," in *Proc. 47th Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, USA, Nov. 2013, pp. 945–948.
- [20] D. E. Hack, L. K. Patton, and B. Himed, "Detection in passive MIMO radar networks," *IEEE Trans. on Signal Process.*, vol. 62, no. 11, pp. 780–785, June 2014.
- [21] S. D. Howard, S. Sirianunpiboon, and D. Cochran, "An exact bayesian detector for multistatic passive radar," in *Proc. 50th Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, USA, Nov. 2016, pp. 1077–1080.
- [22] K. S. Bialkowski, I. V. L. Clarkson, and S. D. Howard, "Generalized canonical correlation for passive multistatic radar detection," in *2011 IEEE Stat. Sig. Proc. Workshop (SSP)*, Nice, France, 2011, pp. 417–420.
- [23] D. Cochran, H. Gish, and D. Sinno, "A geometric approach to multiple-channel signal detection," *IEEE Trans. on Signal Process.*, vol. 43, no. 9, pp. 2049–2057, Sept. 1995.
- [24] W. A. Gardner, W. A. Brown, and C. Chen, "Spectral correlation of modulated signals: Part II - digital modulation," *IEEE Trans. on Comms.*, vol. Com-35, no. 6, pp. 595–601, June 1987.
- [25] W. A. Gardner and C. M. Spooner, "Detection and Source Location of Weak Cyclostationary Signals: Simplifications of the Maximum-Likelihood Receiver," *IEEE Trans. on Comms.*, vol. 41, no. 6, pp. 905–916, June 1993.
- [26] G. Gelli, L. Izzo, and L. Paura, "Cyclostationarity-Based Signal Detection and Source Location in Non-Gaussian noise," *IEEE Trans. on Comms.*, vol. 44, no. 3, pp. 368–376, Mar. 1996.
- [27] D. Ramírez, P. J. Schreier, J. Vía, I. Santamaría, and L. L. Scharf, "Detection of multivariate cyclostationarity," *IEEE Trans. on Signal Process.*, vol. 63, no. 20, pp. 5395–5408, Oct. 2015.
- [28] A. Pries, D. Ramírez, and P. J. Schreier, "Detection of cyclostationarity in the presence of temporal or spatial structure with applications to cognitive radio," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Proc. (ICASSP)*, Shanghai, China, 2016, pp. 4249–4253.
- [29] —, "LMPIT-inspired tests for detecting a cyclostationary signal in noise with spatio-temporal structure," *IEEE Trans. on Wireless Comms.*, vol. 17, no. 9, pp. 6321–6334, Sept. 2018.
- [30] S. Horstmann, D. Ramírez, and P. J. Schreier, "Two-channel passive detection exploiting cyclostationarity," in *Proc. 27th European Signal Proc. Conf. (EUSIPCO)*, A Coruña, Spain, 2019.
- [31] R. A. Wijsman, "Cross-sections of orbits and their application to densities of maximal invariants," in *Proc. Fifth Berkeley Symp. Math. Stat. Prob.*, vol. 1, 1967, pp. 389–400.
- [32] R. Cardinali, F. Colone, P. Lombardo, O. Crognale, A. Cosmi, and A. Lauri, "Multipath cancellation on reference antenna for passive radar which exploits FM transmission," *IET Conference Publications*, no. 530 CP, 2007.
- [33] H. Griffiths and C. Baker, *An Introduction to Passive Radar*. Artech House, 2017.
- [34] F. Colone, R. Cardinali, and P. Lombardo, "Cancellation of clutter and multipath in passive radar using a sequential approach," *IEEE National Radar Conference - Proceedings*, pp. 393–399, 2006.
- [35] A. Zaimbashi, M. Derakhtian, and A. Sheikhi, "GLRT-Based CFAR detection in passive bistatic radar," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 49, no. 1, pp. 134–159, 2013.
- [36] H. Zhao, J. Liu, Z. Zhang, H. Liu, and S. Zhou, "Linear fusion for target detection in passive multistatic radar," *Signal Process.*, vol. 130, pp. 175–182, 2017.
- [37] S. Searle and S. D. Howard, "Clutter cancellation in passive radar as a dual basis projection," in *Proc. 53rd Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, USA, Nov. 2019.
- [38] D. Ramírez, P. J. Schreier, J. Vía, I. Santamaría, and L. L. Scharf, "A regularized maximum likelihood estimator for the period of a cyclostationary process," in *Proc. 48th Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, USA, Nov. 2014, pp. 1972–1976.
- [39] A. V. Dandawaté and G. B. Giannakis, "Statistical tests for presence of cyclostationarity," *IEEE Trans. on Signal Process.*, vol. 42, no. 9, 1994.
- [40] J. Wang, T. Chen, and B. Huang, "Cyclo-period estimation for discrete-time cyclo-stationary signals," *IEEE Trans. on Signal Process.*, vol. 54, no. 1, pp. 83–94, Jan. 2006.
- [41] E. D. Gladyshev, "Periodically correlated random sequences," *Soviet Math. Dokl.*, vol. 2, pp. 385–388, 1961.
- [42] W. A. Gardner, A. Napolitano, and L. Paura, "Cyclostationarity: Half a century of research," *Signal Process.*, vol. 86, no. 4, pp. 639–697, 2006.
- [43] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume 2: Detection Theory*. Upper Saddle River, NJ, US: Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1998.
- [44] D. Ciuonzo, A. De Maio, and D. Orlando, "A unifying framework for adaptive radar detection in homogeneous plus structured interference-Part I: On the maximal invariant statistic," *IEEE Trans. on Signal Process.*, vol. 64, no. 11, pp. 2894–2906, June 2016.
- [45] W. Liu, W. Xie, J. Liu, and Y. Wang, "Adaptive double subspace signal detection in Gaussian background Part I: Homogeneous environments," *IEEE Trans. on Signal Process.*, vol. 62, no. 9, pp. 2345–2357, May 2014.
- [46] N. H. Klausner, M. R. Azimi-Sadjadi, and L. L. Scharf, "Detection of spatially correlated time series from a network of sensor arrays," *IEEE Trans. on Signal Process.*, vol. 62, no. 6, pp. 1396–1407, Mar. 2014.
- [47] L. L. Scharf, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*. Addison - Wesley, 1991.
- [48] D. Ciuonzo, V. Carotenuto, and A. De Maio, "On multiple covariance equality testing with application to SAR change detection," *IEEE Trans. on Signal Process.*, vol. 65, no. 19, pp. 5078–5091, Oct. 2017.
- [49] P. J. Schreier and L. L. Scharf, *Statistical Signal Processing of Complex-Valued Data: The Theory of Improper and Noncircular Signals*. Cambridge University Press, 2010.
- [50] D. Ramírez, J. Vía, I. Santamaría, and L. L. Scharf, "Detection of spatially correlated Gaussian time series," *IEEE Trans. on Signal Process.*, vol. 58, no. 10, pp. 5006–5015, Oct. 2010.
- [51] A. Hjørungnes, *Complex-Valued Matrix Derivatives: With Applications in Signal Processing and Communications*. Cambridge University Press, 2011.