

Testing Equality of Multiple Power Spectral Density Matrices

David Ramírez, *Senior Member, IEEE*, Daniel Romero, *Member, IEEE*, Javier Vía, *Senior Member, IEEE*, Roberto López-Valcarce, *Member, IEEE*, and Ignacio Santamaría, *Senior Member, IEEE*

Abstract—This work studies the existence of optimal invariant detectors for determining whether P multivariate processes have the same power spectral density (PSD). This problem finds application in multiple fields, including physical layer security and cognitive radio. For Gaussian observations, we prove that the optimal invariant detector, i.e., the uniformly most powerful invariant test (UMPIT), does not exist. Additionally, we consider the challenging case of close hypotheses, where we study the existence of the locally most powerful invariant test (LMPIT). The LMPIT is obtained in closed form only for univariate signals. In the multivariate case, it is shown that the LMPIT does not exist. However, the corresponding proof naturally suggests one LMPIT-inspired detector, which outperforms previously proposed detectors.

Index Terms—Generalized likelihood ratio test (GLRT), locally most powerful invariant test (LMPIT), power spectral density (PSD), Toeplitz matrix, uniformly most powerful invariant test (UMPIT).

I. INTRODUCTION

This work studies the problem of determining whether P Gaussian multivariate time series possess the same (possibly matrix-valued) power spectral density (PSD) at every frequency. This interesting problem has many applications, such as comparison of gas pipes [1], analysis of hormonal times series [2], earthquake-explosion discrimination [3], light-intensity emission stability determination [4], physical-layer security [5], and spectrum sensing [6].

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

D. Romero is with the Department of Information and Communication Technology, University of Agder, Grimstad, Norway (e-mail: daniel.romero@uia.no).

J. Vía and I. Santamaría are with the Department of Communications Engineering, University of Cantabria, Santander, Spain (e-mail: {javier.via,i.santamaria}@unican.es).

R. López-Valcarce is with the Department of Signal Theory and Communications, University of Vigo, Vigo, Spain (e-mail: valcarce@gts.uvigo.es).

This work was partly supported by the Spanish MINECO grants COMONSENS Network (TEC2015-69648-REDC) and KERMES Network (TEC2016-81900-REDT/AEI); by the Spanish MINECO and the European Commission (ERDF) grants ADVENTURE (TEC2015-69868-C2-1-R), WINTER (TEC2016-76409-C2-2-R), CARMEN (TEC2016-75067-C4-4-R) and CAIMAN (TEC2017-86921-C2-1-R and TEC2017-86921-C2-2-R); by the Comunidad de Madrid grant CASI-CAM-CM (S2013/ICE-2845); by the Xunta de Galicia and ERDF grants GRC2013/009, R2014/037 and ED431G/04 (Agrupación Estratégica Consolidada de Galicia accreditation 2016-2019); by the SODERCAN and ERDF grant CAIMAN (12.JU01.64661); and by the Research Council of Norway grant FRIPRO TOPPFORSK (250910/F20). This paper was presented in part at the 2018 IEEE International Conference on Acoustics, Speech and Signal Processing.

The first work to consider this problem was developed by Coates and Diggle [1]. This work proposed tests, for univariate and real-valued time series, based on the ratio of periodograms. First, they presented non-parametric tests based on the comparison of the maximum and minimum of the log-ratio of periodograms over all frequencies. Moreover, assuming a parametric (quadratic) model for the log-ratio of the PSDs, they developed a generalized likelihood ratio test (GLRT). These detectors are further studied in [7]. Following similar ideas to [1], [7], the work in [2] proposed a graphical procedure, which resulted in another non-parametric test. The authors of [4] also considered detectors based on the ratio of periodograms for a problem with several time series and are based on a semi-parametric log-linear model for the ratio of PSDs.

A different kind of detector is presented in [8], where the GLRT was derived without imposing any parametric model. In particular, they computed the GLRT for testing whether the PSD of two real multivariate time series are equal at a given frequency. The extension to complex time series is considered in [9], where the information from all frequencies is fused into a single statistic. An alternative way of fusing the information at all frequencies is derived in [10], but the proposed detector is not a GLRT anymore. All aforementioned detectors were developed in the frequency domain. However, there are also other works that propose time-domain detectors; see, for instance, [11] and references therein.

The problem considered in this work is to decide whether at each frequency all P PSD matrices are identical or not. Thus, the tests for equality of several PSDs may be addressed as an extension of the classical problem of testing the homogeneity of covariance matrices [12]. This allows us to use the statistics for homogeneity at each frequency, which need to be fused into one statistic afterwards by combining the different frequencies. Depending on the chosen combination rule, different detectors (with different performance) may be derived.

Every proposed test so far is either based on ad-hoc principles, or on the GLRT, which is optimal in the asymptotic regime [13]. However, their behavior for finite data records is unknown. That is, for finite data records, they may very well be suboptimal. Actually, to the best of the authors' knowledge, neither the uniformly most powerful invariant test (UMPIT) nor the locally most powerful invariant test (LMPIT) have been studied for the considered problem, with the exception of our previous conference paper [14], which considers the particular case of $P = 2$ processes. The conventional approach to derive these optimal invariant detectors is based on obtaining the

ratio of the densities of the maximal invariant statistic under each hypothesis [15]. The derivation of these distributions is for most problems a very complicated task, if possible at all, which in many cases precludes the derivation of optimal invariant detectors. Instead of pursuing this conventional approach, we may invoke Wijsman's theorem [16], [17], as we did in our previous works [18], [19]. This theorem allows us to derive the UMPIT or the LMPIT, if they exist, without identifying the maximal invariant statistic and, more importantly, without computing its distributions. Exploiting Wijsman's theorem, and assuming Gaussian-distributed data, this work proves that the UMPIT does not exist for testing equality of the PSD matrices of $P (\geq 2)$ processes. Moreover, focusing on the case of close hypotheses (similar PSDs), we prove that the LMPIT only exists in the case of univariate processes, for which we find a closed-form expression; whereas it does not exist for the general case of multivariate time series. However, the non-existence proof of the LMPIT in the multivariate case suggests one LMPIT-inspired detector, which turns out to outperform previously proposed schemes.

The paper is organized as follows. Section II presents the mathematical formulation of the problem. The proof of the non-existence of the UMPIT and the LMPIT for the general case is presented in Section III, whereas Section IV derives the LMPIT for univariate processes. Due to the non-existence of the LMPIT in the general case, we present a LMPIT-inspired detector in Section V. The performance of the proposed detector is illustrated by means of numerical simulations in Section VI, and Section VII summarizes the main conclusions of this work.

A. Notation

In this paper, matrices are denoted by bold-faced upper case letters; column vectors are denoted by bold-faced lower case letters, and light-face lower case letters correspond to scalar quantities. The superscripts $(\cdot)^T$ and $(\cdot)^H$ denote transpose and Hermitian, respectively. A complex (real) matrix of dimension $M \times N$ is denoted by $\mathbf{A} \in \mathbb{C}^{M \times N}$ ($\mathbf{A} \in \mathbb{R}^{M \times N}$) and $\mathbf{x} \in \mathbb{C}^M$ ($\mathbf{x} \in \mathbb{R}^M$) denotes that \mathbf{x} is a complex (real) vector of dimension M . The absolute value of the complex number x is denoted as $|x|$, and the determinant, trace and Frobenius norm of a matrix \mathbf{A} will be denoted, respectively, as $\det(\mathbf{A})$, $\text{tr}(\mathbf{A})$ and $\|\mathbf{A}\|_F$. The Kronecker product between two matrices is denoted by \otimes , \mathbf{I}_L is the identity matrix of size $L \times L$, $\mathbf{0}$ denotes the zero matrix of the appropriate dimensions, and \mathbf{F}_N is the Fourier matrix of size N . We use $\mathbf{A}^{-1/2}$ to denote the Hermitian square root matrix of the Hermitian matrix \mathbf{A}^{-1} (the inverse of \mathbf{A}), the operator $\text{diag}_L(\mathbf{A})$ constructs a block diagonal matrix from the $L \times L$ blocks in the diagonal of \mathbf{A} , whereas $\text{diag}(\mathbf{A})$ yields a vector with the elements of the main diagonal of \mathbf{A} , and $\mathbf{x} = \text{vec}(\mathbf{X})$ constructs a vector by stacking the columns of \mathbf{X} . $\mathbf{x} \sim \mathcal{CN}(\boldsymbol{\mu}, \mathbf{R})$ indicates that \mathbf{x} is a proper complex circular Gaussian random vector of mean $\boldsymbol{\mu}$ and covariance matrix \mathbf{R} and $E[\cdot]$ represents the expectation operator. Finally, \propto stands for equality up to additive and positive multiplicative constant (not depending on data) terms.

II. PROBLEM FORMULATION

We are given N samples of P time series that are L variate, $\mathbf{x}_i[n] \in \mathbb{C}^L, i = 1, \dots, P$, and $n = 0, \dots, N - 1$. These samples are realizations of zero-mean proper Gaussian processes that are independent and wide-sense stationary. The problem studied in this work is to determine whether *all* these P processes have the same power spectral density (PSD) matrix at *every* frequency or, alternatively, the same matrix-valued covariance function. To mathematically formulate the detection problem, it is necessary to introduce the vector $\mathbf{y}_i = [\mathbf{x}_i^T[0] \dots \mathbf{x}_i^T[N-1]]^T \in \mathbb{C}^{NL}$, which stacks the N samples of the i th time series, and $\mathbf{y} = [\mathbf{y}_1^T \dots \mathbf{y}_P^T]^T \in \mathbb{C}^{PNL}$. Thus, the problem can be cast as the following binary hypothesis test:

$$\begin{aligned} \mathcal{H}_1 : \mathbf{y} &\sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_{\mathcal{H}_1}), \\ \mathcal{H}_0 : \mathbf{y} &\sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_{\mathcal{H}_0}), \end{aligned} \quad (1)$$

where $\mathcal{CN}(\mathbf{0}, \mathbf{R}_{\mathcal{H}_i})$ denotes a zero-mean circular complex Gaussian distribution with covariance matrix $\mathbf{R}_{\mathcal{H}_i}$. Taking the independence into account, the covariance matrices are

$$\mathbf{R}_{\mathcal{H}_1} = \begin{bmatrix} \mathbf{R}_1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_2 & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{R}_P \end{bmatrix}, \quad (2)$$

and

$$\mathbf{R}_{\mathcal{H}_0} = \begin{bmatrix} \mathbf{R}_0 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_0 & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{R}_0 \end{bmatrix} = \mathbf{I}_P \otimes \mathbf{R}_0, \quad (3)$$

under \mathcal{H}_1 and \mathcal{H}_0 , respectively, with

$$\mathbf{R}_i = E[\mathbf{y}_i \mathbf{y}_i^H] = \begin{bmatrix} \mathbf{M}_i[0] & \dots & \mathbf{M}_i[-N+1] \\ \vdots & \ddots & \vdots \\ \mathbf{M}_i[N-1] & \dots & \mathbf{M}_i[0] \end{bmatrix}, \quad (4)$$

being a block-Toeplitz covariance matrix built from the (unknown) covariance sequence of $\mathbf{x}_i[n]$, $\mathbf{M}_i[m] = E[\mathbf{x}_i[n] \mathbf{x}_i^H[n-m]]$. Moreover, under \mathcal{H}_1 , we do not assume any further particular structure and, for notational simplicity, we have defined the common covariance sequence under \mathcal{H}_0 as $\mathbf{M}_0[m]$.

Dealing with block-Toeplitz covariance matrices is challenging, since they typically prevent the derivation of closed-form detectors, as our previous works in [19]–[21] show. These works also presented a solution to overcome this problem, which is based on an asymptotic (as $N \rightarrow \infty$) approximation of the likelihood. Concretely, this approach performs a block-circulant approximation of \mathbf{R}_i , and results in convergence in the mean-square sense of the asymptotic likelihood to the likelihood [21].

Assume now that M , with $M \geq L$,¹ independent and identically distributed (i.i.d.) realizations of \mathbf{y} , say

¹This reasonable assumption is necessary for the derivation of the GLRT. However, for the study of the existence of optimal invariant detectors, only $PM \geq L$ is required.

$\mathbf{y}^{(0)}, \dots, \mathbf{y}^{(M-1)}$, are available. Then, to obtain the asymptotic likelihood, we must define the transformation $\mathbf{z} = [\mathbf{z}_1^T \cdots \mathbf{z}_P^T]^T$, where

$$\mathbf{z}_i = (\mathbf{F}_N^H \otimes \mathbf{I}_L) \mathbf{y}_i = [\mathbf{z}_i^T[0] \cdots \mathbf{z}_i^T[N-1]]^T, \quad (5)$$

with $\mathbf{z}_i[k]$ being the discrete Fourier transform (DFT) of $\mathbf{x}_i[n]$ at frequency $\theta_k = 2\pi k/N, k = 0, \dots, N-1$. Exploiting this transformation, the asymptotic approximation of the likelihood is [19], [20]

$$p(\mathbf{z}^{(0)}, \dots, \mathbf{z}^{(M-1)}; \mathbf{S}_{\mathcal{H}_i}) = \frac{1}{\pi^{PNLM} \det(\mathbf{S}_{\mathcal{H}_i})^M} \exp \left\{ -M \operatorname{tr} \left(\mathbf{S}_{\mathcal{H}_i}^{-1} \hat{\mathbf{S}} \right) \right\}, \quad (6)$$

where the $PNL \times PNL$ sample covariance matrix of the transformed observations is

$$\hat{\mathbf{S}} = \frac{1}{M} \sum_{m=0}^{M-1} \mathbf{z}^{(m)} \mathbf{z}^{(m)H}, \quad (7)$$

and the covariance matrices under both hypotheses are

$$\mathbf{S}_{\mathcal{H}_1} = \begin{bmatrix} \mathbf{S}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{S}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{S}_P \end{bmatrix}, \quad (8)$$

and

$$\mathbf{S}_{\mathcal{H}_0} = \mathbf{I}_P \otimes \mathbf{S}_0. \quad (9)$$

Moreover, \mathbf{S}_i is an $NL \times NL$ block-diagonal matrix whose $L \times L$ blocks $\mathbf{S}_{i,1}, \dots, \mathbf{S}_{i,N}$ are given by the power spectral density matrix, i.e.,

$$\mathbf{S}_{i,k+1} = \mathbf{S}_i(e^{j\theta_k}) = \sum_{n=0}^{N-1} \mathbf{M}_i[n] e^{-j\theta_k n}, \quad (10)$$

with $i = 1, \dots, P$, and $k = 0, \dots, N-1$. Finally, using the asymptotic likelihood, the detection problem in (1) is asymptotically equivalent to

$$\begin{aligned} \mathcal{H}_1 : \mathbf{z}^{(m)} &\sim \mathcal{CN}(\mathbf{0}, \mathbf{S}_{\mathcal{H}_1}), \quad m = 0, \dots, M-1, \\ \mathcal{H}_0 : \mathbf{z}^{(m)} &\sim \mathcal{CN}(\mathbf{0}, \mathbf{S}_{\mathcal{H}_0}), \quad m = 0, \dots, M-1. \end{aligned} \quad (11)$$

That is, we are testing two different covariance matrices with known structure but unknown values.

Although our formulation assumes Gaussian data, as well as the availability of M realizations each of length N , this is not very restrictive in practice. First, the Gaussianity assumption can be dropped since the transformed observations, $\mathbf{z}_i[k]$, are samples of the DFT and it is well known, see [22], that under some mild conditions the DFT of large data records yields $\mathbf{z}_i[k]$ that are independent and Gaussian distributed with covariance matrix $\mathbf{S}_i(e^{j\theta_k})$. Moreover, if only $M = 1$ realization is available, it is possible to split this single realization into M windows, but keeping in mind that the realizations may no longer be i.i.d., as the samples in different windows may be dependent. This resembles the Welch method for PSD estimation. Moreover, further exploiting on this idea, it would be possible to increase the number of realizations allowing some overlap among different windows. However, the study of the side effects (due to a higher correlation

among windows) will not be analyzed in this work. Finally, the case of different numbers M_i of realizations for each process would require a different treatment, which would be equivalent to the introduction of further structure (subsets of identical covariance matrices) under the alternative hypothesis. Although this is a very interesting case, which is currently under consideration, the modification of the problem invariances introduces an additional complexity that is beyond the scope of this paper.

A. The generalized likelihood ratio test

The typical approach to solve detection problems with unknown parameters, as (11), is based on the GLRT. Actually, the works in [8], [9], [11] derived the GLRT for this problem under different assumptions, but they only studied the case of $P = 2$ time series. The GLRT in [11] was derived for univariate real time series, which was extended to multivariate real signals in [8] and multivariate complex signals in [9]. Here, we present the (straightforward) extension of these GLRTs to $P (\geq 2)$ complex and multivariate processes. Concretely, the GLRT is given by

$$\log \mathcal{G} \propto \sum_{k=0}^{N-1} \sum_{i=1}^P \log \left[\frac{\det(\hat{\mathbf{S}}_i(e^{j\theta_k}))}{\det\left(\frac{1}{P} \sum_{p=1}^P \hat{\mathbf{S}}_p(e^{j\theta_k})\right)} \right], \quad (12)$$

where

$$\hat{\mathbf{S}}_i(e^{j\theta_k}) = \frac{1}{M} \sum_{m=0}^{M-1} \mathbf{z}_i^{(m)}[k] \mathbf{z}_i^{(m)H}[k], \quad (13)$$

is the sample PSD matrix at frequency θ_k , i.e., an $L \times L$ block of $\hat{\mathbf{S}}$. Alternatively, the GLRT may be rewritten as

$$\log \mathcal{G} \propto \sum_{k=0}^{N-1} \sum_{i=1}^P \log \det \left(\hat{\mathbf{C}}_i(e^{j\theta_k}) \right), \quad (14)$$

where the frequency coherence is

$$\begin{aligned} \hat{\mathbf{C}}_i(e^{j\theta}) &= \\ &\left(\frac{1}{P} \sum_{p=1}^P \hat{\mathbf{S}}_p(e^{j\theta}) \right)^{-1/2} \hat{\mathbf{S}}_i(e^{j\theta}) \left(\frac{1}{P} \sum_{p=1}^P \hat{\mathbf{S}}_p(e^{j\theta}) \right)^{-1/2}. \end{aligned} \quad (15)$$

Finally, it is important to address how the threshold must be chosen. On one hand, we could use Wilks' theorem [13], which states that the GLRT is asymptotically (as $M \rightarrow \infty$) distributed as

$$-2M \sum_{k=0}^{N-1} \sum_{i=1}^P \log \det \left(\hat{\mathbf{C}}_i(e^{j\theta_k}) \right) \sim \chi_{(P-1)NL}^2, \quad (16)$$

that is, it is distributed as a Chi-squared distribution with $(P-1)NL^2$ degrees of freedom. On the other, and for the finite case, we could take into account that the detector is invariant to MIMO filtering, which allows us to consider that under \mathcal{H}_0 the PSDs are $\mathbf{S}_i(e^{j\theta}) = \mathbf{I}_L, \forall \theta$ and $i = 1, \dots, P$. Thus, under this assumption we could obtain the thresholds using Monte

Carlo simulations, which will depend on P, L and M , but will be independent of the specific values of $\mathbf{S}_i(e^{j\theta})$.
 should still be valid for any $\mathbf{S}_i(e^{j\theta})$.

III. ON THE EXISTENCE OF OPTIMAL DETECTORS

Since GLRTs are not necessarily optimal for finite data records [23], the goal of this section is to study the existence of optimal invariant detectors for the hypothesis test in (11). In particular, we will show that neither the uniformly most powerful invariant test (UMPIT) nor the locally most powerful invariant test (LMPIT) exist in the general case.

To derive invariant detectors, such as the UMPIT or the LMPIT, we must first identify the problem invariances [15]. Specifically, we must define the group of invariant transformations, which is composed only of linear transformations since Gaussianity must be preserved. Among those linear transformations, it is clear that applying the same invertible multiple-input-multiple-output (MIMO) filtering to all time series does not modify the structure of the hypotheses. That is, if the PSDs are equal, the same MIMO filtering yields also identical PSDs and if they are different, they will stay different. In particular, this transformation is $\tilde{\mathbf{x}}_i[n] = (\mathbf{H} * \mathbf{x}_i)[n]$, where $\mathbf{H}[n] \in \mathbb{C}^{L \times L}$ is a filtering matrix common to all processes and $*$ denotes convolution, which when applied to \mathbf{z}_i becomes

$$\tilde{\mathbf{z}}_i = \mathbf{G}\mathbf{z}_i, \quad (17)$$

where \mathbf{G} is a block-diagonal matrix with invertible $L \times L$ blocks. Additionally, we may label the processes in any arbitrary order, which may be even done on a frequency-by-frequency basis. The last invariance consists of a frequency reordering. That is, we may permute the frequencies, i.e., permute $\mathbf{z}_i[k]$, provided that the same permutation is applied to all processes. Then, the group of invariant transformations for the hypothesis test in (11) is

$$\mathcal{G} = \{\mathbf{g} : \mathbf{z} \mapsto \mathbf{g}(\mathbf{z}) = \tilde{\mathbf{G}}\mathbf{z}\}, \quad (18)$$

where

$$\tilde{\mathbf{G}} = (\mathbf{I}_P \otimes \mathbf{G}) \left(\sum_{k=1}^N \mathbf{P}_k^T \otimes \mathbf{e}_k \mathbf{e}_k^T \otimes \mathbf{I}_L \right) (\mathbf{I}_P \otimes \mathbf{T} \otimes \mathbf{I}_L), \quad (19)$$

with \mathbf{e}_k being the k th column of \mathbf{I}_N , $\mathbf{P}_k \in \mathbb{R}^{P \times P}$ and $\mathbf{T} \in \mathbb{R}^{N \times N}$. Moreover, the matrix $\mathbf{I}_P \otimes \mathbf{T} \otimes \mathbf{I}_L$ applies a frequency reordering (permutation) to every process, \mathbf{G} is a block-diagonal matrix with $L \times L$ invertible blocks \mathbf{G}_k , and \mathbf{P}_k is a matrix that permutes the k th frequency of all processes, i.e., permutes the position of $\mathbf{z}_1[k-1], \dots, \mathbf{z}_P[k-1]$ in \mathbf{z} . Then, $\mathbf{P}_k \in \mathbb{P}_k$, $\mathbf{T} \in \mathbb{T}$, and $\mathbf{G}_k \in \mathbb{G}$, where \mathbb{P}_k and \mathbb{T} are the set of permutation matrices formed by \mathbf{P}_k and \mathbf{T} , respectively, and \mathbb{G} is the set of $L \times L$ invertible matrices. In Appendix I, we prove that $\left(\sum_{k=1}^N \mathbf{P}_k^T \otimes \mathbf{e}_k \mathbf{e}_k^T \otimes \mathbf{I}_L \right) \mathbf{z}$ corresponds to the relabeling of the processes at each frequency.

Equipped with the transformation group \mathcal{G} , it is possible to study the existence of the UMPIT. The typical approach [15] involves finding the maximal invariant statistic and computing its densities under both hypotheses. An alternative to this process, which is usually very involved or even intractable, is based on Wijsman's theorem [16], [17], and allows us to

derive the UMPIT, if it exists, without finding the maximal invariant statistic nor its distributions. This theorem directly gives the ratio of the distributions of the maximal invariant statistic as follows

$$\mathcal{L} = \frac{\sum_{\mathbb{T}, \mathbb{P}_1, \dots, \mathbb{P}_N} \int_{\mathbb{G}^N} |\det(\mathbf{G})|^{2MP} \exp \left\{ -M \operatorname{tr} \left(\mathbf{S}_{\mathcal{H}_1}^{-1} \tilde{\mathbf{G}} \hat{\mathbf{S}} \tilde{\mathbf{G}}^H \right) \right\} d\mathbf{G}}{\sum_{\mathbb{T}, \mathbb{P}_1, \dots, \mathbb{P}_N} \int_{\mathbb{G}^N} |\det(\mathbf{G})|^{2MP} \exp \left\{ -M \operatorname{tr} \left(\mathbf{S}_{\mathcal{H}_0}^{-1} \tilde{\mathbf{G}} \hat{\mathbf{S}} \tilde{\mathbf{G}}^H \right) \right\} d\mathbf{G}}, \quad (20)$$

where $\mathbb{G}^N = \mathbb{G} \times \dots \times \mathbb{G}$ and $d\mathbf{G}$ is an invariant measure on the set \mathbb{G}^N and the sum over \mathbb{P}_k represents the sum over all permutations matrices in the set \mathbb{P}_k . If the ratio \mathcal{L} , or a monotone transformation thereof, did not depend on unknown parameters, it would yield the UMPIT. When such dependence is present, the UMPIT does not exist, and in that case it is sensible to consider the case of close hypotheses to study the existence of the LMPIT.

In the following, we will simplify the ratio \mathcal{L} , show that the UMPIT does not exist, and study the case of close hypotheses. The next lemma presents the first simplification of \mathcal{L} .

Lemma 1: The ratio \mathcal{L} in (20) can be written as

$$\mathcal{L} \propto \sum_{\mathbb{T}} \sum_{\mathbb{P}_1, \dots, \mathbb{P}_N} \int_{\mathbb{G}^N} \exp(-M\alpha(\mathbf{G})) \prod_{l=1}^N \beta(\mathbf{G}_l) d\mathbf{G}_l, \quad (21)$$

where

$$\alpha(\mathbf{G}) = \sum_{k=1}^N \sum_{i=1}^P \operatorname{tr} \left(\mathbf{W}_{i,k} \mathbf{G}_k \hat{\mathbf{C}}_{\pi_k[i], \Pi[k]} \mathbf{G}_k^H \right), \quad (22)$$

the matrix $\hat{\mathbf{C}}_{\pi_k[i], \Pi[k]}$ is a permutation of the sample coherence matrix

$$\hat{\mathbf{C}}_{i,k} = \left[\frac{1}{P} \sum_{p=1}^P \hat{\mathbf{S}}_{p,k} \right]^{-1/2} \hat{\mathbf{S}}_{i,k} \left[\frac{1}{P} \sum_{p=1}^P \hat{\mathbf{S}}_{p,k} \right]^{-1/2}, \quad (23)$$

with

$$\hat{\mathbf{S}}_{i,k+1} = \frac{1}{M} \sum_{m=0}^{M-1} \mathbf{z}_i^{(m)}[k] \mathbf{z}_i^{(m)H}[k], \quad (24)$$

being the sample estimate of the PSD matrix of $\{\mathbf{x}_i[n]\}_{n=0}^{N-1}$ at frequency $2\pi k/N$. Moreover, the scalar $\beta(\mathbf{G}_l)$ is given by

$$\beta(\mathbf{G}_l) = |\det(\mathbf{G}_l)|^{2MP} \exp \left\{ -PM \operatorname{tr} \left(\mathbf{G}_l^H \mathbf{G}_l \right) \right\}, \quad (25)$$

and the matrix $\mathbf{W}_{i,k}$ is

$$\mathbf{W}_{i,k} = \left(\frac{1}{P} \sum_{p=1}^P \mathbf{S}_{p,k} \right)^{-1/2} \mathbf{S}_{i,k}^{-1} \left(\frac{1}{P} \sum_{p=1}^P \mathbf{S}_{p,k} \right)^{-1/2} - \mathbf{I}_L. \quad (26)$$

Proof: The proof is presented in Appendix II. \square

As can be seen in (21), the ratio \mathcal{L} depends on the matrices $\mathbf{W}_{i,k}$, which are unknown, proving that the UMPIT does not exist. Hence, as previously mentioned, we focus hereafter on the case of close hypotheses, i.e., the PSD matrices are very similar. In this case, $\mathbf{S}_1 \approx \dots \approx \mathbf{S}_P$, which yields $\mathbf{W}_{i,k} \approx \mathbf{0}$.

Under this assumption, the exponent term in (21) becomes small and allows us to perform a Taylor series expansion around $\alpha(\mathbf{G}) = 0$ as follows

$$\exp(-M\alpha(\mathbf{G})) \approx \frac{1}{2} (2 - 2M\alpha(\mathbf{G}) + M^2\alpha^2(\mathbf{G})), \quad (27)$$

which yields

$$\mathcal{L} \propto \mathcal{L}_l + \mathcal{L}_q, \quad (28)$$

where the linear and quadratic terms are respectively given by

$$\mathcal{L}_l = -2M \sum_{\mathbf{T}} \sum_{\mathbb{P}_1, \dots, \mathbb{P}_N} \int_{\mathbb{G}^N} \alpha(\mathbf{G}) \prod_{l=1}^N \beta(\mathbf{G}_l) d\mathbf{G}_l, \quad (29)$$

and

$$\mathcal{L}_q \propto M^2 \sum_{\mathbf{T}} \sum_{\mathbb{P}_1, \dots, \mathbb{P}_N} \int_{\mathbb{G}^N} \alpha^2(\mathbf{G}) \prod_{l=1}^N \beta(\mathbf{G}_l) d\mathbf{G}_l. \quad (30)$$

Next, we analyze the linear term \mathcal{L}_l .

Lemma 2: The linear term is zero, i.e.,

$$\mathcal{L}_l = 0. \quad (31)$$

Proof: The proof can be found in Appendix III. \square

Since the linear term is zero, only the quadratic term has to be taken into account. The final expression is provided in the following theorem.

Theorem 1: The ratio of the distributions of the maximal invariant statistic is given by

$$\mathcal{L} \propto \sum_{k=1}^N \sum_{i=1}^P \|\hat{\mathbf{C}}_{i,k}\|_F^2 + \beta \sum_{k=1}^N \sum_{i=1}^P \text{tr}^2(\hat{\mathbf{C}}_{i,k}), \quad (32)$$

where β is a data-independent function of the matrices $\mathbf{W}_{i,k}$, which are unknown.

Proof: See Appendix IV. \square

As Theorem 1 shows, the ratio \mathcal{L} depends on unknown parameters, which are summarized in β . Hence, the LMPIT for testing the equality of PSD matrices at all frequencies does not exist in the general case. An exception is examined in the following section. Moreover, an LMPIT-inspired detector is also presented in Section V, and its performance is analyzed using computer simulations.

One final comment is in order. Since the ratio \mathcal{L} is given by a linear combination (with unknown weights) of the Frobenius norm and the trace of $\hat{\mathbf{C}}_{i,k}$, the optimal detector would be a function of the eigenvalues of $\hat{\mathbf{C}}_{i,k}$. This makes sense as the distributions of these eigenvalues are not modified by any of the invariances.

IV. THE LMPIT FOR UNIVARIATE TIME SERIES

The case of univariate time series, $L = 1$, is interesting since the LMPIT does exist, as shown in the next corollary.

Corollary 1: For $L = 1$ the ratio in (32) reduces to

$$\mathcal{L} \propto \sum_{k=1}^N \sum_{i=1}^P |\hat{C}_{i,k}|^2, \quad (33)$$

which is therefore the LMPIT.

Proof: In the univariate case, the coherence matrices $\hat{\mathbf{C}}_{i,k}$ become scalar, that is, $\hat{\mathbf{C}}_{i,k} = \hat{C}_{i,k}$, and consequently

$$\|\hat{\mathbf{C}}_{i,k}\|_F^2 = \text{tr}^2(\hat{C}_{i,k}) = |\hat{C}_{i,k}|^2, \quad (34)$$

which yields

$$\mathcal{L} \propto (1 + \beta) \sum_{k=1}^N \sum_{i=1}^P |\hat{C}_{i,k}|^2 \propto \sum_{k=1}^N \sum_{i=1}^P |\hat{C}_{i,k}|^2. \quad (35)$$

\square

Interestingly, using (10) and the definition of $\hat{C}_{i,k}$ in (23), the LMPIT in Corollary 1 may be rewritten in a more insightful form as

$$\mathcal{L} \propto \sum_{k=0}^{N-1} \frac{\sum_{i=1}^P \hat{S}_i^2(e^{j\theta_k})}{\left[\sum_{i=1}^P \hat{S}_i(e^{j\theta_k}) \right]^2}, \quad (36)$$

or asymptotically (as $N \rightarrow \infty$)

$$\mathcal{L} \propto \int_{-\pi}^{\pi} \frac{\sum_{i=1}^P \hat{S}_i^2(e^{j\theta})}{\left[\sum_{i=1}^P \hat{S}_i(e^{j\theta}) \right]^2} \frac{d\theta}{2\pi}. \quad (37)$$

Thus, for $L = 1$, the LMPIT is given by the integral of the sum of the squares of the PSD estimates normalized by the square of their sum.

V. AN LMPIT-INSPIRED DETECTOR

Since the LMPIT does not exist in the multivariate case ($L > 1$), we present here an LMPIT-inspired detector. In particular, we could use each of the terms in (32) as test statistics, which are

$$\mathcal{L}_F = \sum_{k=1}^N \sum_{i=1}^P \|\hat{\mathbf{C}}_{i,k}\|_F^2 = \sum_{k=0}^{N-1} \sum_{i=1}^P \|\hat{\mathbf{C}}_i(e^{j\theta_k})\|_F^2, \quad (38)$$

and

$$\mathcal{L}_T = \sum_{k=1}^N \sum_{i=1}^P \text{tr}^2(\hat{\mathbf{C}}_{i,k}) = \sum_{k=0}^{N-1} \sum_{i=1}^P \text{tr}^2(\hat{\mathbf{C}}_i(e^{j\theta_k})), \quad (39)$$

where the frequency coherence, $\hat{\mathbf{C}}_i(e^{j\theta})$, was defined in (15). Note that in the univariate case ($L = 1$), both (38) and (39) reduce to the true LMPIT (33). However, for multivariate processes we only propose \mathcal{L}_F as a detector and discard \mathcal{L}_T . To understand why, let us analyze both. Considering the case where all PSDs are identical (\mathcal{H}_0), the coherence matrices are $\hat{\mathbf{C}}_i(e^{j\theta}) \approx \mathbf{I}_L, \forall \theta$. Actually, for a large number of realizations, $M \rightarrow \infty$, they converge to $\hat{\mathbf{C}}_i(e^{j\theta}) = \mathbf{I}_L, \forall \theta$. Hence, to distinguish between both hypotheses, the detectors must measure how different $\hat{\mathbf{C}}_i(e^{j\theta})$ is from \mathbf{I}_L . This is exactly what the statistics \mathcal{L}_F and \mathcal{L}_T do. The only difference resides in the way they quantify this difference: while \mathcal{L}_F uses the Frobenius norm, \mathcal{L}_T uses the trace. Since the Frobenius norm exploits information provided by the cross-spectral densities of

each multivariate time series, information which is neglected by the trace operator, one would expect \mathcal{L}_F to outperform \mathcal{L}_T .

Finally, as in Section IV, it is possible to write the asymptotic versions of (38) as

$$\mathcal{L}_F = \int_{-\pi}^{\pi} \sum_{i=1}^P \|\hat{\mathbf{C}}_i(e^{j\theta})\|_F^2 \frac{d\theta}{2\pi}. \quad (40)$$

A. Threshold selection

In this section, we study the threshold selection problem for the LMPIT-inspired detector \mathcal{L}_F . Similarly to the approaches described for the GLRT, we could consider that the PSDs are $\mathbf{S}_i(e^{j\theta}) = \mathbf{I}_L, \forall \theta$ and $i = 1, \dots, P$, and use Monte Carlo simulations to obtain the thresholds. Of course, due to the invariance to MIMO filtering, these thresholds should be valid for other PSDs. The second approach is also based on Wilks' theorem, but it cannot be directly applied. Concretely, we will follow along the lines in [18]. First, for close hypotheses, the GLRT may be approximated by

$$\begin{aligned} \sum_{k=0}^{N-1} \sum_{i=1}^P \log \det \left(\hat{\mathbf{C}}_i(e^{j\theta_k}) \right) &= \sum_{k=0}^{N-1} \sum_{i=1}^P \sum_{s=1}^L \log \left(1 + \epsilon_{i,s}(e^{j\theta_k}) \right) \\ &\approx \sum_{k=0}^{N-1} \sum_{i=1}^P \sum_{s=1}^L \left(\epsilon_{i,s}(e^{j\theta_k}) - \frac{\epsilon_{i,s}^2(e^{j\theta_k})}{2} \right), \end{aligned} \quad (41)$$

where $1 + \epsilon_{i,s}(e^{j\theta_k})$ is the s th eigenvalue of $\hat{\mathbf{C}}_i(e^{j\theta_k})$ and $|\epsilon_{i,s}(e^{j\theta_k})| \ll 1$. After some straightforward manipulations, the above approximation becomes

$$\begin{aligned} \sum_{k=0}^{N-1} \sum_{i=1}^P \log \det \left(\hat{\mathbf{C}}_i(e^{j\theta_k}) \right) &\approx \frac{NPL}{2} + \\ &\sum_{k=0}^{N-1} \sum_{i=1}^P \sum_{s=1}^L \left(2\epsilon_{i,s}(e^{j\theta_k}) - \frac{(1 + \epsilon_{i,s}(e^{j\theta_k}))^2}{2} \right), \end{aligned} \quad (42)$$

Now, using (60), we get

$$\sum_{i=1}^P \sum_{s=1}^L \epsilon_{i,s}(e^{j\theta_k}) = 0, \quad (43)$$

which yields

$$\begin{aligned} -2M \sum_{k=0}^{N-1} \sum_{i=1}^P \log \det \left(\hat{\mathbf{C}}_i(e^{j\theta_k}) \right) \\ \approx M \sum_{k=0}^{N-1} \sum_{i=1}^P \|\hat{\mathbf{C}}_i(e^{j\theta_k})\|_F^2 - MNPL. \end{aligned} \quad (44)$$

Hence, we obtain the following asymptotic distribution of \mathcal{L}_F

$$(M\mathcal{L}_F - MNPL) \sim \chi_{(P-1)NL}^2. \quad (45)$$

VI. NUMERICAL RESULTS

This section studies the performance of the proposed detectors using Monte Carlo simulations, and compare it with that

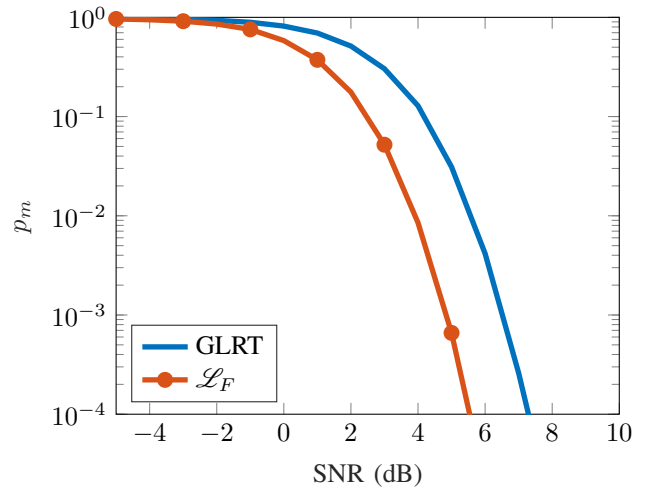


Fig. 1: Probability of missed detection for Experiment 1: $P = 3$, $L = 3$, $Q = 1$, $T = 20$, $\rho = 0.75$, $\Delta_h = 0.1$, $M = 4$ and $N = 128$

of the GLRT. The performance evaluation is carried out in a communication setup, where the signals are generated as

$$\mathbf{x}_i[n] = \sum_{\tau=0}^{T-1} \mathbf{H}_i[\tau] \mathbf{s}_i[n - \tau] + \mathbf{v}_i[n], \quad i = 1, \dots, P,$$

which corresponds to a MIMO channel with finite impulse response. In this expression, the transmitted signals $\mathbf{s}_i[n] \in \mathbb{C}^Q$ are independent multivariate processes whose entries are independent QPSK symbols with unit energy, the noise vectors $\mathbf{v}_i[n] \in \mathbb{C}^L$ are independent with variance σ^2 , and spatially and temporally white. The channel $\mathbf{H}_1[n]$ is a Rayleigh MIMO channel with unit energy,² spatially uncorrelated, and with exponential power delay profile of parameter ρ . Finally, $\mathbf{H}_i[n] = \sqrt{1 - \Delta_h} \mathbf{H}_1[n] + \sqrt{\Delta_h} \mathbf{E}_i[n]$, $i = 2, \dots, P$, with $\mathbf{E}_i[n]$ possessing the same statistical properties as $\mathbf{H}_1[n]$ and being independent. Under this model, $\Delta_h = 0$ corresponds to signals having the same PSD, that is \mathcal{H}_0 , and $0 < \Delta_h \leq 1$ measures how far the hypotheses are, along with the signal-to-noise ratio, which is defined as

$$\text{SNR (dB)} = 10 \log \left(\frac{1}{\sigma^2} \right). \quad (46)$$

Experiment 1: In this first experiment, we have considered $P = 3$ processes of dimension $L = 3$, $Q = 1$ signals are transmitted through MIMO channels with $T = 20$ taps, the parameter of the exponential power delay profile is $\rho = 0.75$, and $\Delta_h = 0.1$ under \mathcal{H}_1 . Moreover, to carry out the detection, a realization of 512 samples is available, which is then divided into $M = 4$ windows of length $N = 128$. The probability of missed detection for a fixed probability of false alarm $p_{fa} = 10^{-2}$ and for a varying SNR is depicted in Fig. 1. As this figure shows, the proposed LMPIT-inspired detector, \mathcal{L}_F , outperforms the GLRT, with an approximate gain of 2 dB.

²Each element of $\mathbf{H}_1[n]$ follows a proper complex Gaussian distribution with zero mean and unit variance.

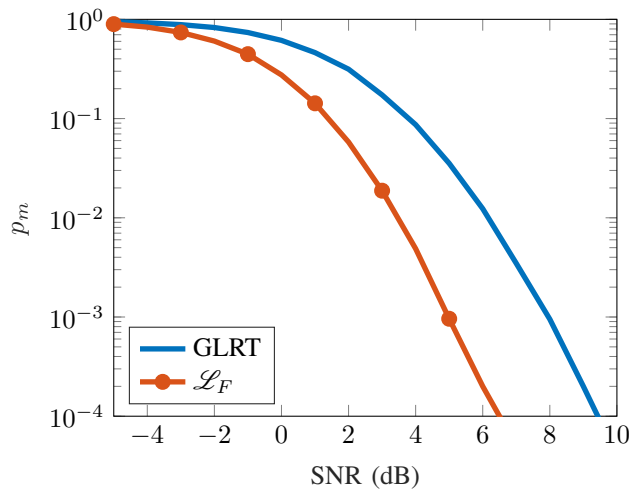


Fig. 2: Probability of missed detection for Experiment 2: $P = 3$, $L = 3$, $Q = 3$, $T = 20$, $\rho = 0.75$, $\Delta_h = 0.1$, $M = 4$ and $N = 128$

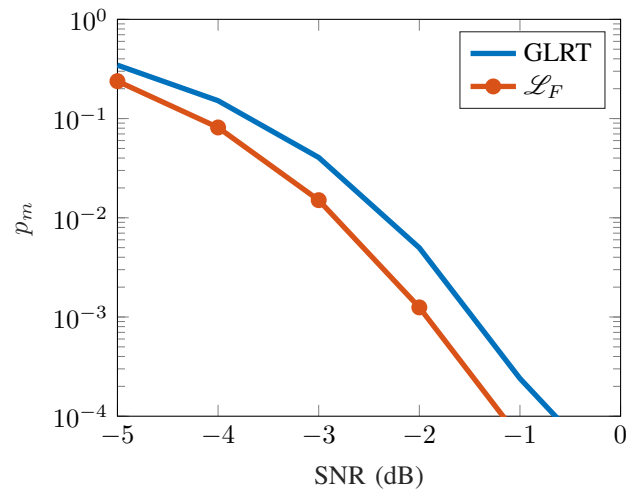


Fig. 3: Probability of missed detection for Experiment 3: $P = 3$, $L = 5$, $Q = 5$, $T = 20$, $\rho = 0.75$, $\Delta_h = 0.1$, $M = 8$ and $N = 128$

Experiment 2: The setup of this experiment is equivalent to that of Experiment 1, with the exception that $Q = 3$ signals are transmitted. The results for this scenario are presented in Fig. 2, where similar conclusions can be drawn. In this scenario, which could be expected to be a bit more favorable for the GLRT, since the true PSD matrices are full rank, the detector \mathcal{L}_F still outperforms the GLRT.

Experiment 3: The third experiment considers $P = 3$ processes of dimension $L = 5$, and $Q = 5$ signals are transmitted. The channel parameters remain the same as in the two previous experiments. However, in this case, 1024 samples are acquired, which are divided into $M = 8$ windows of length $N = 128$. In this case, the difference between the GLRT and the Frobenius norm detector is reduced as Fig. 3 shows. This was expected since, in this case, there is a larger M , which reduces the accuracy of the second-order Taylor expansion in (27), i.e., the assumption of close hypotheses begins to not hold true.

Experiment 4: The fourth experiment also analyzes the effect of the distance between the hypotheses. In particular, we have obtained the probability of missed detection ($p_{fa} = 10^{-2}$) for a varying Δ_h and two different SNRs, 0 and 5 dB. In this experiment, there are $P = 2$ processes of dimension $L = 3$, and $Q = 3$ signals are transmitted. The channel parameters remain the same as in the previous experiments with the exception of Δ_h , and 512 samples are obtained, which are divided into $M = 4$ realizations of length $N = 128$. The results for this experiment are shown in Fig. 4, which demonstrates that the close hypotheses assumption holds for many setups, where the detector \mathcal{L}_F still outperforms the GLRT.

Experiment 5: In the fifth experiment, we evaluate the effect of N and M . Specifically, we have considered a scenario with the same parameters of Experiment 1, with the exception of the SNR, which is fixed to 3 dB, the length of the long realization and how it is divided. Concretely, we have swept N between 32 and 256 in steps of 16 samples, and considered

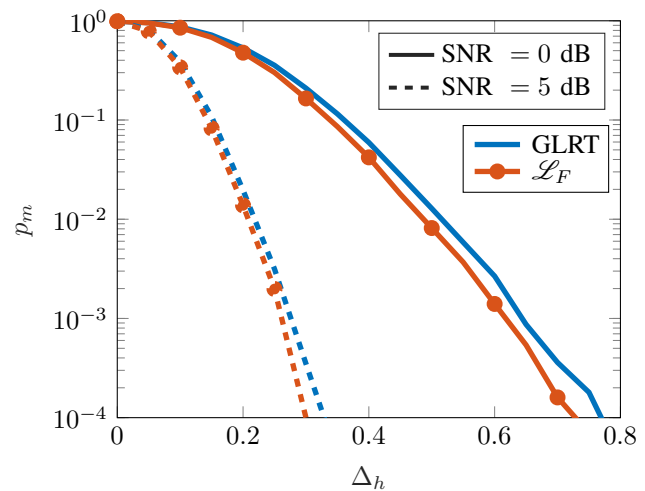


Fig. 4: Probability of missed detection for Experiment 4: $P = 2$, $L = 3$, $Q = 3$, $T = 20$, $\rho = 0.75$, $M = 4$ and $N = 128$

$M = 4$ and $M = 8$. Thus, a long realization of the appropriate length is generated and divided according to the values of N and M , that is, for each point of the curves the total number of samples may be different. Figure 5 shows the probability of missed detection for $p_{fa} = 10^{-2}$, which shows that in this setup it is more convenient to reduce the variance of the estimator (increase M) at the expense of a lower resolution (small N). This can be seen if we compare the probability of missed detection for $N = 32$ and $M = 8$ with that of $N = 64$ and $M = 4$. Both have same number of samples $MN = 256$, but p_m is smaller for $M = 8$, i.e., for \mathcal{L}_F we have $p_m = 0.1346$ against $p_m = 0.3454$. Finally, note that this analysis affects both, the GLRT and the proposed detector.

Experiment 6: This experiment evaluates the performance of the detectors in a larger-scale scenario. Specifically, we have considered $P = 2$ processes of dimension $L = 30$, $Q = 3$, the MIMO channels have $T = 20$ taps, $\rho = 0.75$ and $\text{SNR} = -8$ dB, $\Delta_h = 0.1$ under \mathcal{H}_1 , and a long realization

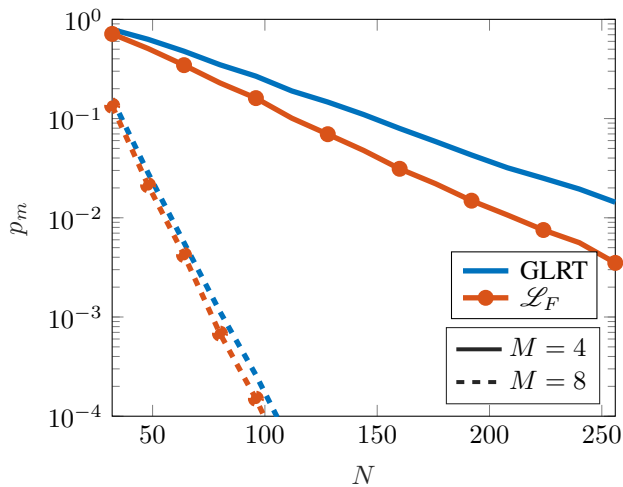


Fig. 5: Probability of missed detection for Experiment 5: $P = 3$, $L = 3$, $Q = 1$, $T = 20$, $\rho = 0.75$, $\Delta_h = 0.1$, and $\text{SNR} = 3$ dB

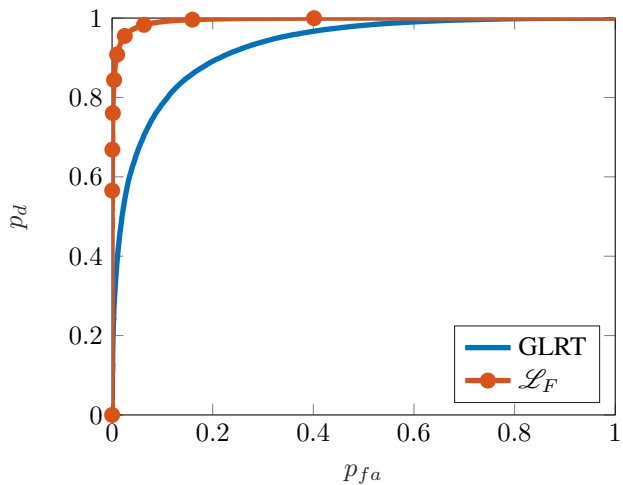


Fig. 6: ROC curves for Experiment 6: $P = 2$, $L = 30$, $Q = 3$, $T = 20$, $\rho = 0.75$, $\Delta_h = 0.1$, $\text{SNR} = -8$ dB, $M = 30$ and $N = 128$

of 3840 samples is generated, which is then divided into $M = 30$ windows of length $N = 128$. The receiver operating characteristic (ROC) curve for both detectors is depicted in Figure 6, which shows the better performance of the proposed detector even in this large-scale scenario.

Experiment 7: In this experiment, we evaluate the accuracy of Wilks' approximations for both, GLRT and the LMPIT-inspired detector. Fig. 7 shows the accuracy of these approximations for a scenario with $P = 3$ processes of dimension $L = 4$, and $Q = 4$ signals are transmitted. The channel parameters are those of Experiment 1 with an $\text{SNR} = 5$ dB. We have considered a long realization (of the appropriate length) that is divided into $M = 10, 20, \dots, 100$ windows of length $N = 64$. As can be seen in the figure, the χ^2 distribution approximates much better the distribution of \mathcal{L}_F than that of the GLRT. That is, it seems to converge faster for the Frobenius norm based detector than for the GLRT.

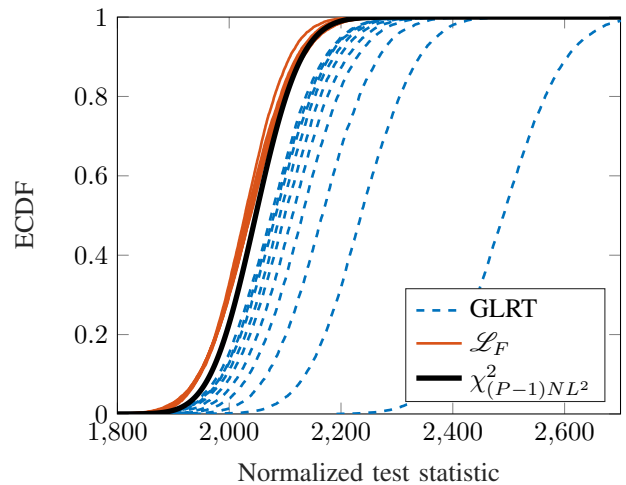


Fig. 7: Empirical cumulative distribution functions (ECDF) for Experiment 7: $P = 3$, $L = 4$, $Q = 4$, $T = 20$, $\rho = 0.75$, $M = 10, 20, \dots, 100$, $N = 64$, and $\text{SNR} = 5$ dB

VII. CONCLUSIONS

This work has studied the existence of optimal invariant detectors for testing whether P multivariate time series have the same power spectral density (PSD) matrix at every frequency. Specifically, our derivation shows that the uniformly most powerful invariant test (UMPIT) for this problem does not exist. Considering close hypotheses (i.e., similar PSD matrices), we obtained the locally most powerful invariant test (LMPIT) for the case of univariate time series, and showed that the LMPIT does not exist in the multivariate case. Nevertheless, our derivation suggested one LMPIT-inspired detector for multivariate time series, which outperforms previously proposed detectors as shown by extensive numerical examples.

APPENDIX I

DERIVATION OF THE RELABELING MATRICES

Let us start by rewriting \mathbf{z} as $\mathbf{z} = \text{vec}(\mathbf{Z})$, where

$$\mathbf{Z} = \begin{bmatrix} \mathbf{z}_1[0] & \cdots & \mathbf{z}_P[0] \\ \vdots & \ddots & \vdots \\ \mathbf{z}_1[N-1] & \cdots & \mathbf{z}_P[N-1] \end{bmatrix} = \sum_{k=1}^N \mathbf{e}_k \otimes [\mathbf{z}_1[k-1] \cdots \mathbf{z}_P[k-1]]. \quad (47)$$

Relabeling the processes at the k th frequency corresponds to permuting the columns of $[\mathbf{z}_1[k-1] \cdots \mathbf{z}_P[k-1]]$, that is, $[\mathbf{z}_1[k-1] \cdots \mathbf{z}_P[k-1]] \mathbf{P}_k$, where \mathbf{P}_k is an arbitrary $P \times P$ permutation matrix. It is important to note that the permutation matrix \mathbf{P}_k depends on the frequency since we may apply different relabelings at each frequency. The observations after the N (possibly) different relabelings are given by

$$\tilde{\mathbf{Z}} = \sum_{k=1}^N \mathbf{e}_k \otimes ([\mathbf{z}_1[k-1] \cdots \mathbf{z}_P[k-1]] \mathbf{P}_k) = \sum_{k=1}^N (\mathbf{e}_k \otimes [\mathbf{z}_1[k-1] \cdots \mathbf{z}_P[k-1]]) \mathbf{P}_k. \quad (48)$$

Taking now into account that

$$[\mathbf{z}_1[k-1] \ \cdots \ \mathbf{z}_P[k-1]] = (\mathbf{e}_k^T \otimes \mathbf{I}_L) \mathbf{Z}, \quad (49)$$

it is easy to show that

$$\begin{aligned} \mathbf{e}_k \otimes [(\mathbf{e}_k^T \otimes \mathbf{I}_L) \mathbf{Z}] &= (\mathbf{e}_k \otimes \mathbf{e}_k^T \otimes \mathbf{I}_L) \mathbf{Z} \\ &= (\mathbf{e}_k \mathbf{e}_k^T \otimes \mathbf{I}_L) \mathbf{Z}, \end{aligned} \quad (50)$$

which yields

$$\tilde{\mathbf{Z}} = \sum_{k=1}^N (\mathbf{e}_k \mathbf{e}_k^T \otimes \mathbf{I}_L) \mathbf{Z} \mathbf{P}_k. \quad (51)$$

Finally, vectorizing $\tilde{\mathbf{Z}}$ yields

$$\tilde{\mathbf{z}} = \text{vec}(\tilde{\mathbf{Z}}) = \left(\sum_{k=1}^N \mathbf{P}_k^T \otimes \mathbf{e}_k \mathbf{e}_k^T \otimes \mathbf{I}_L \right) \mathbf{z}. \quad (52)$$

APPENDIX II PROOF OF LEMMA 1

Since $\tilde{\mathbf{G}}$ and $\mathbf{S}_{\mathcal{H}_i}$ are block-diagonal matrices with block size L , we may substitute $\hat{\mathbf{S}}$ in (20) by $\text{diag}_L(\hat{\mathbf{S}})$ without modifying the integrals. Before proceeding, we must define the block-diagonal matrix

$$\begin{aligned} \hat{\mathbf{S}}_\pi &= \sum_{k,l=1}^N (\mathbf{P}_k^T \otimes \mathbf{e}_k \mathbf{e}_k^T \otimes \mathbf{I}_L) (\mathbf{I}_P \otimes \mathbf{T} \otimes \mathbf{I}_L) \text{diag}_L(\hat{\mathbf{S}}) \times \\ &\quad (\mathbf{I}_P \otimes \mathbf{T}^T \otimes \mathbf{I}_L) (\mathbf{P}_l \otimes \mathbf{e}_l \mathbf{e}_l^T \otimes \mathbf{I}_L), \end{aligned} \quad (53)$$

and, taking into account the effect of the permutations, it becomes

$$\hat{\mathbf{S}}_\pi = \begin{bmatrix} \hat{\mathbf{S}}_{\pi_1[1],\Pi[1]} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \hat{\mathbf{S}}_{\pi_2[1],\Pi[2]} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \hat{\mathbf{S}}_{\pi_N[P],\Pi[N]} \end{bmatrix}, \quad (54)$$

where $\pi_k[\cdot]$, $k = 1, \dots, N$, are possibly different permutations of size P and $\Pi[\cdot]$ is a permutation of size N . That is, $\pi_k[\cdot]$ is the permutation associated to \mathbf{P}_k , $\mathbf{P}_k \rightarrow \pi_k[\cdot]$, and $\Pi[\cdot]$ is the permutation associated to \mathbf{T} , $\mathbf{T} \rightarrow \Pi[\cdot]$. Applying now the change of variables $\mathbf{G}_{\Pi[k]} \rightarrow \mathbf{G}_{\Pi[k]} \left[1/P \sum_{i=1}^P \hat{\mathbf{S}}_{i,\Pi[k]} \right]^{-1/2}$ to the integrals in the numerator and denominator of (20), the ratio \mathcal{L} becomes

$$\mathcal{L} = \frac{\sum_{\mathbf{T}} \sum_{\mathbf{P}_1, \dots, \mathbf{P}_N} \int_{\mathbf{G}^N} |\det(\mathbf{G})|^{2MP} \exp(-M\gamma_1) d\mathbf{G}}{\sum_{\mathbf{T}} \sum_{\mathbf{P}_1, \dots, \mathbf{P}_N} \int_{\mathbf{G}^N} |\det(\mathbf{G})|^{2MP} \exp(-M\gamma_2) d\mathbf{G}}, \quad (55)$$

where

$$\gamma_1 = \text{tr} \left(\mathbf{S}_{\mathcal{H}_1}^{-1} (\mathbf{I}_P \otimes \mathbf{G}) \hat{\mathbf{C}}_\pi (\mathbf{I}_P \otimes \mathbf{G}^H) \right), \quad (56)$$

and

$$\gamma_2 = \text{tr} \left((\mathbf{I}_P \otimes \mathbf{G}^H \mathbf{S}_0^{-1} \mathbf{G}) \hat{\mathbf{C}}_\pi \right), \quad (57)$$

with

$$\hat{\mathbf{C}}_\pi = \begin{bmatrix} \hat{\mathbf{C}}_{\pi_1[1],\Pi[1]} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \hat{\mathbf{C}}_{\pi_2[1],\Pi[2]} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \hat{\mathbf{C}}_{\pi_N[P],\Pi[N]} \end{bmatrix}, \quad (58)$$

where we have used (9). We continue by applying the change of variables $\mathbf{G}_k \rightarrow \mathbf{S}_{0,k}^{1/2} \mathbf{G}_k$ to the integral in the denominator, which simplifies γ_2 to

$$\begin{aligned} \gamma_2 &= \text{tr} \left((\mathbf{I}_P \otimes \mathbf{G}^H \mathbf{G}) \hat{\mathbf{C}}_\pi \right) \\ &= \sum_{k=1}^N \text{tr} \left(\mathbf{G}_k^H \mathbf{G}_k \sum_{i=1}^P \hat{\mathbf{C}}_{\pi_k[i],\Pi[k]} \right) = P \sum_{k=1}^N \text{tr} \left(\mathbf{G}_k^H \mathbf{G}_k \right), \end{aligned} \quad (59)$$

where we have taken into account that

$$\sum_{i=1}^P \hat{\mathbf{C}}_{i,k} = P \mathbf{I}_L. \quad (60)$$

Since this exponent does not depend on the observations, the denominator may be discarded from the ratio. Let us now consider the change of variables $\mathbf{G} \rightarrow (1/P \sum_i \mathbf{S}_i^{-1})^{-1/2} \mathbf{G}$, which yields

$$\gamma_1 = \text{tr} \left[(\mathbf{W} + \mathbf{I}) (\mathbf{I}_P \otimes \mathbf{G}) \hat{\mathbf{C}}_\pi (\mathbf{I}_P \otimes \mathbf{G}^H) \right], \quad (61)$$

where

$$\mathbf{W} = \begin{bmatrix} \mathbf{W}_{1,1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{W}_{1,2} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{W}_{P,N} \end{bmatrix}, \quad (62)$$

with $\mathbf{W}_{i,k}$ defined in (26). The proof concludes by taking (60) again into account and the block-diagonal structure of all matrices. \square

APPENDIX III PROOF OF LEMMA 2

The linear term \mathcal{L}_l can be rewritten as

$$\begin{aligned} \mathcal{L}_l &= -2M\psi^{N-1} \sum_{\mathbf{T}} \sum_{k=1}^N \sum_{\mathbf{P}_1, \dots, \mathbf{P}_N} \sum_{i=1}^P \\ &\quad \int_{\mathbf{G}} \beta(\mathbf{G}_k) \text{tr} \left(\mathbf{W}_{i,k} \mathbf{G}_k \hat{\mathbf{C}}_{\pi_k[i],\Pi[k]} \mathbf{G}_k^H \right) d\mathbf{G}_k, \end{aligned} \quad (63)$$

where

$$\psi = \int_{\mathbf{G}} \beta(\mathbf{G}_l) d\mathbf{G}_l, \quad (64)$$

does not depend on l . Now, performing the sum over $\mathbb{P}_1, \dots, \mathbb{P}_N$, \mathcal{L}_1 becomes

$$\begin{aligned} \mathcal{L}_1 &= -2[(P-1)!]^N M \psi^{N-1} \sum_{\mathbb{T}} \sum_{k=1}^N \sum_{p_k=1}^P \sum_{i=1}^P \\ &\quad \int_{\mathbb{G}} \beta(\mathbf{G}_k) \text{tr} \left(\mathbf{W}_{i,k} \mathbf{G}_k \hat{\mathbf{C}}_{p_k, \Pi[k]} \mathbf{G}_k^H \right) d\mathbf{G}_k \\ &= -2[(P-1)!]^N M \psi^{N-1} \sum_{\mathbb{T}} \sum_{k=1}^N \sum_{p_k=1}^P \\ &\quad \int_{\mathbb{G}} \beta(\mathbf{G}_k) \text{tr} \left[\left(\sum_{i=1}^P \mathbf{W}_{i,k} \right) \mathbf{G}_k \hat{\mathbf{C}}_{p_k, \Pi[k]} \mathbf{G}_k^H \right] d\mathbf{G}_k, \end{aligned} \quad (65)$$

where for notational convenience

$$\sum_{k=1}^N \sum_{p_k=1}^P = \sum_{k=1}^N \sum_{p_1=1}^P \cdots \sum_{p_N=1}^P, \quad (66)$$

and we have taken into account that

$$\sum_{\mathbb{P}} f(\pi[i]) = (P-1)! \sum_{p=1}^P f(p), \quad (67)$$

for permutations of size P and any arbitrary function $f(\cdot)$. The proof follows from

$$\sum_{i=1}^P \mathbf{W}_{i,k} = \mathbf{0}, \quad \forall k. \quad (68)$$

□

APPENDIX IV PROOF OF THEOREM 1

Exploiting the results of Lemma 2, the ratio of distributions of the maximal invariant statistic simplifies to

$$\mathcal{L} \propto \sum_{\mathbb{T}} \sum_{\mathbb{P}_1, \dots, \mathbb{P}_N} \int_{\mathbb{G}^N} \alpha^2(\mathbf{G}) \prod_{l=1}^N \beta(\mathbf{G}_l) d\mathbf{G}_l, \quad (69)$$

and expanding $\alpha^2(\mathbf{G})$ yields

$$\mathcal{L} \propto \mathcal{L}_1 + \mathcal{L}_2 + \mathcal{L}_3 + \mathcal{L}_4 \quad (70)$$

where the expressions of the terms in the right-hand side of (70) are shown at the top of the next page. After summing over $\mathbb{P}_1, \dots, \mathbb{P}_N$ and \mathbb{T} , \mathcal{L}_1 becomes

$$\mathcal{L}_1 \propto \sum_{k,l=1}^N \sum_{p_k=1}^P \sum_{i=1}^P \int_{\mathbb{G}} \beta(\mathbf{G}_k) \text{tr}^2 \left(\mathbf{W}_{i,k} \mathbf{G}_k \hat{\mathbf{C}}_{p_k, l} \mathbf{G}_k^H \right) d\mathbf{G}_k. \quad (75)$$

Let us now consider \mathcal{L}_2 , which after summing over $\mathbb{P}_1, \dots, \mathbb{P}_N$, becomes

$$\begin{aligned} \mathcal{L}_2 &\propto \sum_{\mathbb{T}} \sum_{\substack{k,l=1 \\ k \neq l}}^N \sum_{p_k=1}^P \sum_{q_l=1}^P \sum_{i=1}^P \\ &\quad \left[\int_{\mathbb{G}} \beta(\mathbf{G}_k) \text{tr} \left(\mathbf{W}_{i,k} \mathbf{G}_k \hat{\mathbf{C}}_{p_k, \Pi[k]} \mathbf{G}_k^H \right) d\mathbf{G}_k \right] \\ &\quad \times \left[\int_{\mathbb{G}} \beta(\mathbf{G}_l) \text{tr} \left(\mathbf{W}_{i,l} \mathbf{G}_l \hat{\mathbf{C}}_{q_l, \Pi[l]} \mathbf{G}_l^H \right) d\mathbf{G}_l \right]. \end{aligned} \quad (76)$$

Carrying out the summations in p_k and q_l , and taking (60) into account, the term \mathcal{L}_2 simplifies to

$$\begin{aligned} \mathcal{L}_2 &\propto \sum_{\mathbb{T}} \sum_{\substack{k,l=1 \\ k \neq l}}^N \sum_{i=1}^P \left[\int_{\mathbb{G}} \beta(\mathbf{G}_k) \text{tr} \left(\mathbf{W}_{i,k} \mathbf{G}_k \mathbf{G}_k^H \right) d\mathbf{G}_k \right] \\ &\quad \times \left[\int_{\mathbb{G}} \beta(\mathbf{G}_l) \text{tr} \left(\mathbf{W}_{i,l} \mathbf{G}_l \mathbf{G}_l^H \right) d\mathbf{G}_l \right], \end{aligned} \quad (77)$$

which does not depend on the observations and therefore can be discarded. We shall now focus on \mathcal{L}_3 , which can be rewritten as

$$\begin{aligned} \mathcal{L}_3 &\propto \sum_{\mathbb{T}} \sum_{k=1}^N \sum_{p_k=1}^P \sum_{\substack{q_k=1 \\ p_k \neq q_k}}^P \sum_{\substack{i,j=1 \\ i \neq j}}^P \\ &\quad \int_{\mathbb{G}} \beta(\mathbf{G}_k) \text{tr} \left(\mathbf{W}_{i,k} \mathbf{G}_k \hat{\mathbf{C}}_{p_k, \Pi[k]} \mathbf{G}_k^H \right) \\ &\quad \times \text{tr} \left(\mathbf{W}_{j,k} \mathbf{G}_k \hat{\mathbf{C}}_{q_k, \Pi[k]} \mathbf{G}_k^H \right) d\mathbf{G}_k, \end{aligned} \quad (78)$$

since

$$\sum_{\mathbb{P}} f(\pi[i]) g(\pi[j]) = (P-2)! \sum_{\substack{p,q=1 \\ p \neq q}}^P f(p) g(q), \quad (79)$$

for permutations of size P , $i \neq j$, and two arbitrary functions $f(\cdot)$ and $g(\cdot)$. The expression for \mathcal{L}_3 simplifies to

$$\begin{aligned} \mathcal{L}_3 &\propto \sum_{\mathbb{T}} \sum_{k=1}^N \sum_{p_k=1}^P \sum_{\substack{i,j=1 \\ i \neq j}}^P \int_{\mathbb{G}} \beta(\mathbf{G}_k) \text{tr} \left(\mathbf{W}_{i,k} \mathbf{G}_k \hat{\mathbf{C}}_{p_k, \Pi[k]} \mathbf{G}_k^H \right) \\ &\quad \times \text{tr} \left(\mathbf{W}_{j,k} \mathbf{G}_k \left(\mathbf{P}\mathbf{I} - \hat{\mathbf{C}}_{p_k, \Pi[k]} \right) \mathbf{G}_k^H \right) d\mathbf{G}_k, \end{aligned} \quad (80)$$

after summing in q_k and taking into account that

$$\sum_{\substack{q_k=1 \\ p_k \neq q_k}}^P \hat{\mathbf{C}}_{q_k, \Pi[k]} = \mathbf{P}\mathbf{I}_L - \hat{\mathbf{C}}_{p_k, \Pi[k]}. \quad (81)$$

Finally, expanding the second trace and summing over p_k and j yields

$$\mathcal{L}_3 \propto \sum_{k,l=1}^N \sum_{p_k=1}^P \sum_{i=1}^P \int_{\mathbb{G}} \beta(\mathbf{G}_k) \text{tr}^2 \left(\mathbf{W}_{i,k} \mathbf{G}_k \hat{\mathbf{C}}_{p_k, l} \mathbf{G}_k^H \right) d\mathbf{G}_k, \quad (82)$$

which is identical (up to a multiplicative and additive constants) to \mathcal{L}_1 . Summing over $\mathbb{P}_1, \dots, \mathbb{P}_N$, the term \mathcal{L}_4 becomes

$$\begin{aligned} \mathcal{L}_4 &\propto \sum_{\mathbb{T}} \sum_{\substack{k,l=1 \\ k \neq l}}^N \sum_{p_k=1}^P \sum_{q_l=1}^P \sum_{\substack{i,j=1 \\ i \neq j}}^P \\ &\quad \left[\int_{\mathbb{G}} \beta(\mathbf{G}_k) \text{tr} \left(\mathbf{W}_{i,k} \mathbf{G}_k \hat{\mathbf{C}}_{p_k, \Pi[k]} \mathbf{G}_k^H \right) d\mathbf{G}_k \right] \\ &\quad \times \left[\int_{\mathbb{G}} \beta(\mathbf{G}_l) \text{tr} \left(\mathbf{W}_{j,l} \mathbf{G}_l \hat{\mathbf{C}}_{q_l, \Pi[l]} \mathbf{G}_l^H \right) d\mathbf{G}_l \right], \end{aligned} \quad (83)$$

$$\mathcal{L}_1 = \psi^{N-1} \sum_{\mathbb{T}} \sum_{k=1}^N \sum_{\mathbb{P}_1, \dots, \mathbb{P}_N} \sum_{i=1}^P \int_{\mathbb{G}} \beta(\mathbf{G}_k) \text{tr}^2 \left(\mathbf{W}_{i,k} \mathbf{G}_k \hat{\mathbf{C}}_{\pi_k[i], \Pi[k]} \mathbf{G}_k^H \right) d\mathbf{G}_k, \quad (71)$$

$$\begin{aligned} \mathcal{L}_2 = \psi^{N-2} \sum_{\mathbb{T}} \sum_{\substack{k,l=1 \\ k \neq l}}^N \sum_{\mathbb{P}_1, \dots, \mathbb{P}_N} \sum_{i=1}^P \left[\int_{\mathbb{G}} \beta(\mathbf{G}_k) \text{tr} \left(\mathbf{W}_{i,k} \mathbf{G}_k \hat{\mathbf{C}}_{\pi_k[i], \Pi[k]} \mathbf{G}_k^H \right) d\mathbf{G}_k \right] \\ \times \left[\int_{\mathbb{G}} \beta(\mathbf{G}_l) \text{tr} \left(\mathbf{W}_{i,l} \mathbf{G}_l \hat{\mathbf{C}}_{\pi_l[i], \Pi[l]} \mathbf{G}_l^H \right) d\mathbf{G}_l \right], \quad (72) \end{aligned}$$

$$\mathcal{L}_3 = \psi^{N-1} \sum_{\mathbb{T}} \sum_{k=1}^N \sum_{\substack{\mathbb{P}_1, \dots, \mathbb{P}_N \\ i,j=1 \\ i \neq j}} \int_{\mathbb{G}} \beta(\mathbf{G}_k) \text{tr} \left(\mathbf{W}_{i,k} \mathbf{G}_k \hat{\mathbf{C}}_{\pi_k[i], \Pi[k]} \mathbf{G}_k^H \right) \text{tr} \left(\mathbf{W}_{j,k} \mathbf{G}_k \hat{\mathbf{C}}_{\pi_k[j], \Pi[k]} \mathbf{G}_k^H \right) d\mathbf{G}_k, \quad (73)$$

$$\begin{aligned} \mathcal{L}_4 = \psi^{N-2} \sum_{\mathbb{T}} \sum_{\substack{k,l=1 \\ k \neq l}}^N \sum_{\substack{\mathbb{P}_1, \dots, \mathbb{P}_N \\ i,j=1 \\ i \neq j}} \left[\int_{\mathbb{G}} \beta(\mathbf{G}_k) \text{tr} \left(\mathbf{W}_{i,k} \mathbf{G}_k \hat{\mathbf{C}}_{\pi_k[i], \Pi[k]} \mathbf{G}_k^H \right) d\mathbf{G}_k \right] \\ \times \left[\int_{\mathbb{G}} \beta(\mathbf{G}_l) \text{tr} \left(\mathbf{W}_{j,l} \mathbf{G}_l \hat{\mathbf{C}}_{\pi_l[j], \Pi[l]} \mathbf{G}_l^H \right) d\mathbf{G}_l \right]. \quad (74) \end{aligned}$$

which reduces to

$$\begin{aligned} \mathcal{L}_4 \propto \sum_{\mathbb{T}} \sum_{\substack{k,l=1 \\ k \neq l}}^N \sum_{\substack{i,j=1 \\ i \neq j}}^P \left[\int_{\mathbb{G}} \beta(\mathbf{G}_k) \text{tr} \left(\mathbf{W}_{i,k} \mathbf{G}_k \mathbf{G}_k^H \right) d\mathbf{G}_k \right] \\ \times \left[\int_{\mathbb{G}} \beta(\mathbf{G}_l) \text{tr} \left(\mathbf{W}_{j,l} \mathbf{G}_l \mathbf{G}_l^H \right) d\mathbf{G}_l \right], \quad (84) \end{aligned}$$

after summing over p_k and q_l , and taking (60) into account. It is clear that \mathcal{L}_4 does not depend on the observations and therefore can be discarded. Then, combining all \mathcal{L}_i , we obtain

$$\mathcal{L} \propto \sum_{k,l=1}^N \sum_{p_k=1}^P \sum_{i=1}^P \int_{\mathbb{G}} \beta(\mathbf{G}_k) \text{tr}^2 \left(\mathbf{W}_{i,k} \mathbf{G}_k \hat{\mathbf{C}}_{p_k,l} \mathbf{G}_k^H \right) d\mathbf{G}_k. \quad (85)$$

Applying the change of variables $\mathbf{G}_k \rightarrow \mathbf{V}_{i,k} \mathbf{G}_k \mathbf{U}_{p_k,l}^H$, with $\hat{\mathbf{C}}_{p_k,l} = \mathbf{U}_{p_k,l} \boldsymbol{\Sigma}_{p_k,l} \mathbf{U}_{p_k,l}^H$ and $\mathbf{W}_{i,k} = \mathbf{V}_{i,k} \boldsymbol{\Lambda}_{i,k} \mathbf{V}_{i,k}^H$ being the eigenvalue decompositions of $\hat{\mathbf{C}}_{p_k,l}$ and $\mathbf{W}_{i,k}$, yields

$$\mathcal{L} \propto \sum_{k,l=1}^N \sum_{p_k=1}^P \sum_{i=1}^P \int_{\mathbb{G}} \beta(\mathbf{G}_k) \text{tr}^2 \left(\mathbf{G}_k^H \boldsymbol{\Lambda}_{i,k} \mathbf{G}_k \boldsymbol{\Sigma}_{p_k,l} \right) d\mathbf{G}_k. \quad (86)$$

Now, expressing the trace as

$$\text{tr} \left(\mathbf{G}_k^H \boldsymbol{\Lambda}_{i,k} \mathbf{G}_k \boldsymbol{\Sigma}_{p_k,l} \right) = \boldsymbol{\sigma}_{p_k,l}^T \tilde{\mathbf{G}}_k \boldsymbol{\lambda}_{i,k}, \quad (87)$$

where $\boldsymbol{\sigma}_{p_k,l} = \text{diag}(\boldsymbol{\Sigma}_{p_k,l})$, $\boldsymbol{\lambda}_{i,k} = \text{diag}(\boldsymbol{\Lambda}_{i,k})$, and $[\tilde{\mathbf{G}}_k]_{m,n} = |[\mathbf{G}_k]_{m,n}|^2$, the ratio becomes

$$\mathcal{L} \propto \sum_{k,l=1}^N \sum_{p_k=1}^P \sum_{i=1}^P \boldsymbol{\sigma}_{p_k,l}^T \mathbf{E}_{i,k} \boldsymbol{\sigma}_{p_k,l}, \quad (88)$$

where

$$\mathbf{E}_{i,k} = \int_{\mathbb{G}} \beta(\mathbf{G}_k) \tilde{\mathbf{G}}_k \boldsymbol{\lambda}_{i,k} \boldsymbol{\lambda}_{i,k}^T \tilde{\mathbf{G}}_k^T d\mathbf{G}_k. \quad (89)$$

The quadratic form $\boldsymbol{\sigma}_{p_k,l}^T \mathbf{E}_{i,k} \boldsymbol{\sigma}_{p_k,l}$ is invariant to permutations³ of the elements of $\boldsymbol{\sigma}_{p_k,l}$, then $\mathbf{E}_{i,k}$ must be of the form

$$\mathbf{E}_{i,k} = \bar{e}_{i,k} \mathbf{I} + \tilde{e}_{i,k} \mathbf{1}\mathbf{1}^T, \quad (90)$$

yielding

$$\begin{aligned} \mathcal{L} \propto \sum_{k,l=1}^N \sum_{p_k=1}^P \sum_{i=1}^P \boldsymbol{\sigma}_{p_k,l}^T (\bar{e}_{i,k} \mathbf{I} + \tilde{e}_{i,k} \mathbf{1}\mathbf{1}^T) \boldsymbol{\sigma}_{p_k,l} \\ = \sum_{k,l=1}^N \sum_{p_k=1}^P \sum_{i=1}^P \left(\bar{e}_{i,k} \|\hat{\mathbf{C}}_{p_k,l}\|^2 + \tilde{e}_{i,k} \text{tr}^2(\hat{\mathbf{C}}_{p_k,l}) \right) \\ = \sum_{k,l=1}^N \sum_{p_k=1}^P \left(\bar{e}_k \|\hat{\mathbf{C}}_{p_k,l}\|^2 + \tilde{e}_k \text{tr}^2(\hat{\mathbf{C}}_{p_k,l}) \right), \quad (91) \end{aligned}$$

where

$$\bar{e}_k = \sum_{i=1}^P \bar{e}_{i,k}, \quad \tilde{e}_k = \sum_{i=1}^P \tilde{e}_{i,k}. \quad (92)$$

The proof concludes by rewriting \mathcal{L} as

$$\mathcal{L} \propto \bar{e} \sum_{l=1}^N \sum_{p_k=1}^P \|\hat{\mathbf{C}}_{p_k,l}\|^2 + \tilde{e} \sum_{l=1}^N \sum_{p_k=1}^P \text{tr}^2(\hat{\mathbf{C}}_{p_k,l}), \quad (93)$$

where

$$\bar{e} = \sum_{k=1}^N \bar{e}_k, \quad \tilde{e} = \sum_{k=1}^N \tilde{e}_k, \quad (94)$$

dividing (93) by \bar{e} and defining $\beta = \tilde{e}/\bar{e}$. \square

REFERENCES

- [1] D. S. Coates and P. J. Diggle, "Tests for comparing two estimated spectral densities," *Journal of Time Series Analysis*, vol. 7, no. 1, pp. 7–20, Jan. 1986.

³This permutation is an element of the set \mathbb{G} , i.e., particular case of the multiplication by \mathbf{G}_k .

- [2] P. J. Diggle and N. I. Fisher, "Nonparametric comparison of cumulative periodograms," *Journal Royal Stat. Soc. Series C (App. Stat.)*, vol. 40, no. 3, pp. 423–434, 1991.
- [3] Y. Kakizawa, R. H. Shumway, and M. Taniguchi, "Discrimination and clustering for multivariate time series," *Journal Ame. Stat. Assoc.*, vol. 93, no. 441, pp. 328–340, Mar. 1998.
- [4] K. Fokianos and A. Savvides, "On comparing several spectral densities," *Technometrics*, vol. 50, no. 3, pp. 317–331, 2008.
- [5] J. K. Tugnait, "Wireless user authentication via comparison of power spectral densities," *IEEE J. Sel. Areas Comm.*, vol. 31, no. 9, pp. 1791–1802, 2013.
- [6] A. Tani and R. Fantacci, "A low-complexity cyclostationary-based spectrum sensing for UWB and WiMAX coexistence with noise uncertainty," *IEEE Trans. Vehicular Tech.*, vol. 59, no. 6, pp. 2940–2950, 2010.
- [7] P. J. Diggle, *Time Series*. Oxford University Press, 1990.
- [8] N. Ravishanker, J. R. M. Hosking, and J. Mukhopadhyay, "Spectrum-based comparison of stationary multivariate time series," *Meth. Comp. App. Prob.*, vol. 12, no. 4, pp. 749–762, 2010.
- [9] J. K. Tugnait, "Comparing multivariate complex random signals: Algorithm, performance analysis and application," *IEEE Trans. Signal Process.*, vol. 64, no. 4, 2016.
- [10] C. Jentsch and M. Pauly, "Testing equality of spectral densities using randomization techniques," *Bernoulli*, vol. 21, no. 2, pp. 697–739, May 2015.
- [11] R. Lund, H. Bassily, and B. Vidakovic, "Testing equality of stationary autocovariances," *Journal of Time Series Analysis*, vol. 30, no. 3, pp. 332–348, 2009.
- [12] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*. New York: Wiley, 1958.
- [13] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*. Prentice Hall, 1998, vol. II.
- [14] D. Ramírez, D. Romero, J. Vía, R. López-Valcarce, and I. Santamaría, "Locally optimal invariant detector for testing equality of two power spectral densities," in *IEEE Int. Conf. on Acoustics, Speech and Signal Process.*, Calgary, Canada, Apr. 2018.
- [15] L. L. Scharf, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*. Addison - Wesley, 1991.
- [16] 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.
- [17] J. R. Gabriel and S. M. Kay, "Use of Wijsman's theorem for the ratio of maximal invariant densities in signal detection applications," in *Asilomar Conf. Signals, Systems, and Computers*, vol. 1, Nov. 2002, pp. 756 – 762.
- [18] D. Ramírez, J. Vía, I. Santamaría, and L. L. Scharf, "Locally most powerful invariant tests for correlation and sphericity of Gaussian vectors," *IEEE Trans. Inf. Theory*, vol. 59, no. 4, pp. 2128–2141, Apr. 2013.
- [19] D. Ramírez, P. J. Schreier, J. Vía, I. Santamaría, and L. L. Scharf, "Detection of multivariate cyclostationarity," *IEEE Trans. Signal Process.*, vol. 63, no. 20, pp. 5395–5408, Oct. 2015.
- [20] D. Ramírez, J. Vía, I. Santamaría, and L. L. Scharf, "Detection of spatially correlated Gaussian time series," *IEEE Trans. Signal Process.*, vol. 58, no. 10, pp. 5006–5015, Oct. 2010.
- [21] D. Ramírez, G. Vázquez-Vilar, R. López-Valcarce, J. Vía, and I. Santamaría, "Detection of rank- P signals in cognitive radio networks with uncalibrated multiple antennas," *IEEE Trans. Signal Process.*, vol. 59, no. 8, pp. 3764–3774, Aug. 2011.
- [22] D. R. Brillinger, *Time Series: Data Analysis and Theory*. McGraw-Hill, 1981.
- [23] K. V. Mardia, J. T. Kent, and J. M. Bibby, *Multivariate Analysis*. New York: Academic, 1979.