UNIVERSITÄT PADERBORN
Die Universität der Informationsgesellschaft

FAKULTÄT FÜR
ELEKTROTECHNIK,
INFORMATIK UND
MATHEMATIK

Detection of (Almost) Cyclostationary Signals in Single and Multiple Channels

Von der Fakultät für Elektrotechnik, Informatik und Mathematik
der Universität Paderborn

zur Erlangung des akademischen Grades

Doktor der Ingenieurwissenschaften (Dr.-Ing.)

genehmigte Dissertation

von

M. Sc. Stefanie Horstmann

Erster Gutachter:	Prof. Dr. Peter J. Schreier
Zweiter Gutachter:	Prof. Dr. David Ramírez

Tag der mündlichen Prüfung: 11.01.2024

Paderborn 2024

Diss. EIM-E/371

Abstract

The detection of cyclostationary signals is of interest in various fields of science and engineering such as meteorology and climatology, economics, mechanics, and communications. In this thesis cyclostationarity is encountered as an inherent property of communication signals, specifically in the context of passive radar and cognitive radio.

The key of passive radar is that the system uses a non-cooperative transmitter, also known as illuminator of opportunity (IO). However, without control of the transmitted signal, the detection task of a passive radar is more challenging than in active radar. For this reason, in this thesis it is exploited that an IO typically emits signals that are cyclostationary.

The goal is to detect the presence of cyclostationarity at a surveillance array given observations from two arrays, the surveillance array and a reference array. For this two-channel passive detection problem the asymptotic generalized likelihood ratio test (GLRT) is derived. Furthermore, the existence of optimal invariant tests is studied and it is shown that neither the uniformly most powerful invariant test (UMPIT) nor the locally most powerful invariant test (LMPIT) exist for this problem. Based on an insightful interpretation of the test statistics and Monte Carlo simulations, it is possible to propose an LMPIT-inspired detector. Additionally, GLRTs are derived for different assumptions on the noise's spatio-temporal structure and its influence on the detection performance is investigated. Furthermore, the two-channel problem is generalized to multiple surveillance and reference channels for which an asymptotic GLRT is derived and the existence of UMPIT and LMPIT is also analyzed. In order to select the thresholds of the GLRT statistics to guarantee a constant false alarm rate, it is shown that under the null hypothesis the GLRT statistic is asymptotically distributed as the product of independent Beta random variables.

In addition to the two and multi-channel detection problems, a generalized version of the single array detection problem, encountered in applications such as cognitive radio, is also considered. Specifically, a detector for the presence of an almost-cyclostationary signal in wide-sense stationary noise is derived that simultaneously provides an estimate of its cycle period. This problem is treated by combining a resampling approach that allows for the application of a previously proposed GLRT with a multiple hypothesis test.

For all problems that are considered in this thesis extensive Monte Carlo simulations are performed. It is demonstrated that the derived detectors outperform comparable state-of-the-art competing techniques.

Zusammenfassung der Dissertation

Die Detektion zyklostationärer Signale ist in vielen Feldern der Wissenschaft und den Ingenieurswissenschaften von Interesse, darunter Meteorologie und Klimatologie, Ökonomie, Maschinenbau und Kommunikationstechnik. In dieser Arbeit begegnen wir Zyklostationarität als inärente Eigenschaft von Kommunikationssignalen im Kontext von passivem Radar und Cognitive Radio.

Die Kernidee passiven Radars ist das Ausnutzen von nicht-kooperativen Sendern, die als "illuminator of opportunity" (IO) bezeichnet werden. Die Detektion ohne Einflussnahme auf das Sendesignal ist allerdings deutlich herausfordernder. Aufgrunddessen wird in dieser Arbeit ausgenutzt, dass das Sendesignal des IO zyklostationär ist.

Ziel ist die Detektion eines zyklostationären Signals an einem Überwachungs-Array unter Berücksichtigung der Beobachtungen von dem Überwachungs-Array und einem Referenz-Array. Für dieses zwei-kanalige passive Detektionsproblem wird ein asymptotischer "generalized likelihood ratio test" (GLRT) hergeleitet. Darüber hinaus wird die Existenz optimaler invarianter Tests untersucht wobei gezeigt wird, dass weder der "uniformly most powerful invariant test" (UMPIT) noch der "locally most powerful invariant test" (LMPIT) für dieses Problem existieren. Basierend auf der aufschlussreichen Interpretation der Test Statistik und Monte Carlo Simulationen, kann ein vom LMPIT inspirierter Detektor vorgeschlagen werden.

Außerdem werden die GLRTs für verschiedene Annahmen der räumlichen und zeitlichen Struktur des Rauschens und deren Einfluss auf die Detektionsleistung untersucht. Zusätzlich wird das zwei-kanalige Detektionsproblem verallgemeinert für eine beliebige Anzahl der Überwachungs- und Referenz-Arrays. Für dieses Problem wird der asymptotische GLRT hergeleitet und die Existenz des UMPIT und LMPIT analysiert. Um die Schwellwerte der GLRT Statistiken für eine konstante Fehlalarmrate zu wählen, wird gezeigt, dass die GLRT Statistik unter der Null Hypothese asymptotisch äquivalent zur Verteilung eines Produkts unabhängiger Beta Zufallsvariablen ist.

Über das zwei-kanalige und mehr-kanalige Problem hinaus, wird das ein-kanalige Problem, welches im Zusammenhang mit Cognitive Radio Anwendungen, relevant ist, betrachtet. Insbesondere wird ein Detektor für die Detektion fast-zyklostationärer Signale in im weitesten Sinne stationärem Rauschen hergeleitet, das gleichzeitig die Periode des fast-zyklostationären Signals schätzt. Das Problem wird durch die Kombination eines Resampling Ansatzes, der dafür sorgt, dass man einen zuvor hergeleiteten GLRT anwenden kann, und eines Mehrfachen Hypothesentests gelöst.

Für alle Problemstellungen, die in dieser Arbeit untersucht sind, ist mit Monte Carlo Simulationen gezeigt, dass die hergeleiteten Detektoren vergleichbare Techniken übertreffen.

Acknowledgements

First of all I would like to thank my supervisor Prof. Dr. Peter Schreier. I am genuinely thankful that he gave me the opportunity to work as a Ph.D. student and research associate in the SST group to follow my research interests. I am glad to have had his support and guidance throughout my academic career from my Bachelor's degree, through my Master's degree until finalizing this thesis. I am also grateful that, in the meantime, he also provided me access to my post-academic career. I really appreciate all the technical and non-technical things I learned from him.

This thesis and my publications would not have been possible without Prof. Dr. David Ramírez. I am particularly grateful for his support and patience. David has always been available for discussions, to challenge me, and to evaluate and discuss my results. I am truly thankful for the countless things that I learned from him.

I am especially thankful to Peter and David who have both kept pushing me in the past three years to finalize this thesis. Especially Peter's pragmatism combined with David's constant help and reminders have enabled me to finish this work.

I would also like to thank all my former SST group colleagues, especially, Dr. Tanuj Hasija, Dr. Mohammad Soleymani, and Dr. Timothy Marrinan for their time and help during many fruitful discussions. I would also like to thank all my metamorphosis colleagues for their support during my last steps of this work.

Moreover, I am grateful that I had the opportunity to meet Prof. Dr. Louis Scharf. Every occasion to meet him and talk to him, resulted in lots of feedback and new input for my research.

Finally, I would like to thank my family and friends for their support and for providing me refuge and distraction of academics whenever needed.

Stefanie Horstmann
Paderborn, April 2023

Contents

Acronyms and notation

Acronyms

CS	cyclostationary
CR	cognitive radio
DPI	direct-path interference
ECDF	empirical cumulative distribution function
GLRT	generalized likelihood ratio test
IO	illuminator of opportunity
LMPIT	locally most powerful invariant test
MIMO	multiple-input multiple-output
PBR	passive bistatic radar
PSD	power spectral density
p.s.d.	positive semi-definite
QPSK	quadrature phase shift keying
RC	reference channel
ROC	receiver operating characteristic
RP	random processes
SC	surveillance channel
SNR	signal-to-noise ratio
UMPIT	uniformly most powerful invariant test
WSS	wide-sense stationary

Vector and matrix notation

x	scalar
$\mathbf{x} \in \mathbb{C}^M$	denotes a complex-valued vector of dimension M
$\mathbf{A} \in \mathbb{C}^{M \times N}$	complex-valued matrix of dimension $M \times N$
$\mathbf{A}^{1/2}$	square root matrix
$\mathbf{I}_N$	identity matrix of dimension $N \times N$
ϵ_n	nth column of identity matrix
$\mathbb{S}_M^N$	set of block-diagonal matrices of size $N \times N$ with block size $M \times M$
$\mathbb{T}_M^N$	set of block-Toeplitz matrices of size $N \times N$ with block size $M \times M$
$\mathbb{B}_K\mathbb{T}_M^N$	set of Hermitian $K \times K$ block matrices, where each block is a block-Toeplitz matrix of size $N \times N$ with block size $M \times M$
$\mathbf{L}_{NP,N}$	commutation matrix that fulfills the following equation: $\text{vec}(\mathbf{A}) = \mathbf{L}_{MN,N}\,\text{vec}\left(\mathbf{A}^T\right)$ for an $M \times N$ matrix $\mathbf{A}$

Mathematical operations

$(\cdot)^T$	transpose
$(\cdot)^H$	Hermitian transpose
$\mathrm{tr}(\cdot)$	matrix trace
$\det(\cdot)$	matrix determinant
$\|\| \cdot \|\|^2$	Frobenius norm
$\mathrm{vec}(\mathbf{A})$	column-wise vectorization of matrix $\mathbf{A}$
$\mathrm{diag}_M(\mathbf{A})$	block-diagonal matrix with block size M obtained from the $M \times M$ main diagonal blocks of $\mathbf{A}$
$\mathrm{block}_{L,M}\,\mathrm{diag}_N\,(\mathbf{A})$	$L \times M$ block matrix obtained from the matrix $\mathbf{A}$, where each block is a square block-diagonal matrix with block size N obtained from the respective block in $\mathbf{A}$, where $\mathbf{A}$ is of suitable dimension
$\mathbf{B}_k$	kth block on the main diagonal of $\mathbf{B}$ of the appropriate dimension
$\mathbf{B}_k^{(i,j)}$	indicate the (i, j)th possibly matrix-valued element in $\mathbf{B}_k$
$\otimes$	Kronecker product
$*$	convolution
$\propto$	equality up to data-independent positive multiplicative and additive terms
$\mathbf{x} \sim \mathcal{CN}(\boldsymbol{\mu}, \mathbf{R})$	proper complex Gaussian-distributed vector $\mathbf{x}$ with mean $\boldsymbol{\mu}$ and covariance matrix $\mathbf{R}$

1 Introduction

1.1 Motivation and related work

The detection of cyclostationary signals is of interest in several fields of science and engineering. Its relevance is described in the following section. The applications of cyclostationary signal detection in cognitive radio and passive radar are introduced in Sections 1.1.2 and 1.1.3.

1.1.1 Relevance of cyclostationarity detection

In various fields of science and engineering we encounter time-varying physical phenomena that are subject to random variations. In order to statistically process those phenomena they are modeled as random processes (RP). RP that reveal periodically or *almost* periodically varying statistics are refered to as *cyclostationary* (CS) or *almost-cyclostationary* (ACS) processes, which are specific classes of non-stationary RP. Yet another important class of RP of practical relevance are *wide-sense stationary* (WSS) RP, which are commonly used to model phenomena with time-invariant first and second-order moments. Early work on CS processes has been published in the Russian literature in 1959 [1, 2]. In 1961 Gladyshev showed that CS RP can be stationarized [3], which is an important result that has been exploited by [4], which is the basis of this thesis. Further contributions are most notably published by Hurd and Gardner, e.g. [5, 6] and later by Izzo and Napolitano, e.g. [7, 8]. Generally, there is an abundance of publications in CS literature, e.g. refer to [9–11].

The theory of cyclostationarity has been successfully applied in the context of estimation and detection in a variety of different problems. For instance, in oceanography such as passive sonar, CS properties are used to model underwater acoustic signals that enable estimation of propeller shaft rates [12]. In the context of meteorology, where periodicity may be induced by seasonal variations, it is applied in time series prediction with applications in meteorology [13] and for modeling time series in climatology, e.g. carbon dioxide concentrations [14]. In geography CS process are used for modeling seismic records [15]. Another interesting application area is economics, where, for instance, financial data (e.g. electric prices) may also be subject to seasonal changes and reveal periodic statistics. In [16] the authors applied CS theory for forecasting business cycles or to model seasonal decisions of consumers [16]. Additionally, in astronomy, CS detectors have

been investigated to detect celestial objects [17] and the CS signal properties have been exploited for satellite communications [18, 19]. In mechanics CS signals occur due to gear or propeller rotation or in [20] the authors exploit cyclostationarity for fault rolling-element bearing detection by analyzing bearing vibration signals. Finally, in communications, where cyclostationarity is induced by operations such as modulation, multiplexing, and sampling, see e.g. [10], the CS properties are exploited, for instance, for signal modeling (see e.g. [21, 22]), interference cancellation (see e.g. [23, 24]), equalization (see e.g. [25, 26]), channel estimation (see e.g. [27, 28]), filtering (see e.g. [24]), antenna array processing, i.e., beamforming (see e.g. [29, 30]), direction and time difference of arrival estimation (see e.g. [31, 32]). More recently applications of the CS theory include topics such as spectrum sensing for cognitive radio or compressive sensing [11].

Before providing some more insight into the applications considered in this thesis, let us first dive into the basic detection problem that will also serve as the basis of the contributions provided hereafter. The problem can be formulated as the following hypothesis test:

$$\begin{aligned} \mathcal{H}_0 &: \quad \text{measurement is WSS,} \\ \mathcal{H}_1 &: \quad \text{measurement is CS with given cycle period,} \end{aligned} \tag{1.1}$$

where under the null hypothesis $\mathcal{H}_0$ the observations are WSS, e.g., noise, and under the alternative $\mathcal{H}_1$ the observations are CS with given cycle period, e.g., CS signal plus noise. In order to distinguish a WSS process from a CS process the following inherent properties of CS signals (but not WSS signals) are taken advantage of in the reported detectors in the literature[1]: First property is that the CS signal is correlated with its frequency-shifted version by the amount of the cycle frequency as opposed to stationary signals, see e.g. [24, 33]. Second, the covariance function of a CS signal is not only a function of the time-shift but it is a periodic function in absolute time. Under certain conditions the periodicity allows a Fourier-series expansion and the coefficients are referred to as cyclic covariance functions. Those are non-zero for cycle frequencies unequal to zero for CS signals but zero for stationary signals. Equivalently, this can be expressed in the frequency domain as a non-zero cyclic power spectral density for non-zero cycle frequencies. These properties were exploited in e.g. [34–36]. Finally, CS signals can be represented in the bifrequency spectrum. The bifrequency spectrum is the frequency representation of a non-stationary signal obtained from both the time-shift and the absolute time. In this representation a CS signal reveals non-zero contribution on lines besides the stationary manifold depending on the cycle period as opposed to WSS signals that solely have support on the stationary manifold [16, 37]. The aforementioned detectors are designed to exploit the respective properties that allow for distinction between CS and WSS signals, however, they have not been designed based on solid statistical principles. Yet in [4] the authors derived an asymptotic *generalized likelihood ratio test* (GLRT) and later extended this work in [38] and also derived an asymptotic *locally most powerful invariant test* (LMPIT). The test statistics of

[1] Please refer to Section 2.1.2 for the definitions of the CS signal properties listed in the following paragraph.

both GLRT and LMPIT are functions of the spectral coherence matrix, which can be interpreted in terms of all three of the aforementioned CS properties [38]. These two detectors are the basis for the work presented in this thesis. Hence, the hypothesis test in (1.1) will be encountered throughout this work in a modified version for the application in a cognitive radio scenario with unknown cycle period and in an extended version for a passive radar application. In the two subsequent sections more insights into the two applications, i.e., cognitive radio and passive radar, will be presented.

1.1.2 Application in cognitive radio

Let us consider the application of CS in the context of cognitive radio (CR). CR is a communication paradigm, which aims to improve the utilization of the wireless spectrum by increasing the spectral efficiency in wireless systems. The concept of CR was first proposed by Mitola in 1999 [39]. An increasing number of wireless systems results in a limited wireless spectrum, which gave rise to the idea of CR. The main idea is to use licensed frequency bands by unlicensed users. This, however, presumes that unlicensed (cognitive or secondary) users do not degrade the performance of the licensed (primary) users. To this end, novel spectrum allocation policies and the utilization of advanced radio and signal processing technologies are required [40].

Goldsmith [40] defines cognitive radio as follows: *A cognitive radio is a wireless communication system that intelligently utilizes any available side information about the a) activity, b) channel conditions, c) codebooks, or d) messages of other nodes with which it shares the spectrum.*. Hence, the cognitive user must be aware of its environment in order to exploit different available side information. Depending on the assumptions made and side information available, it is possible to cluster different CR technologies into three main paradigms: *underlay*, *overlay*, and *interweave* cognitive radio [40]. Those terms reflect how the cognitive user interacts with the non-cognitive users, i.e., either the signals are underlayed, overlayed, or interweaved with those of the licensed users. In the underlay scenario it is assumed that the cognitive users have knowledge about their interference with each primary user and there is a limit set to the amount of acceptable interference caused, i.e., it must always be below a threshold. In the overlay paradigm, the cognitive user requires knowledge about the licensed user's codebooks and messages, which allows for a cancellation of the interference caused at both cognitive and noncognitive users' receiver. Finally, in the interweave scenario the cognitive user opportunistically accesses the licensed frequency band by prior sensing for space-time-frequency gaps (spectrum holes) in the spectrum. Once such gaps are found the cognitive user may communicate over those frequencies. A constant monitoring of the radio spectrum ensures that the interference to licensed users is minimal [40].

Here we will focus on the interweave CR approach. Specifically, we concentrate on *spectrum sensing*, i.e., the determination of availability of the spectrum. In its simplest form it can be boiled down to the detection of primary signals in noise, which can again be formulated by a hypothesis

test with null hypothesis being noise only and alternative being signal embedded in noise [41,42]. A straightforward approach to solve this problem is *energy detection* that estimates the received energy over a certain period of time and compares it against a threshold value. This approach, however, does not exploit any properties of the underlying communication signal that can highly improve the detection performance. Especially, in environments with low signal-to-noise ratios it is essential to make use of further information about or features of the primary user signals. Such distinctive features of communication signals include, for instance, the cyclic prefix in OFDM modulation or pilot sequences multiplexed into the signal [42]. As it was indicated before, these distinctive signal features give rise to cyclostationary statistical properties, which have been exploited to derive relevant detectors, e.g. in [4,36].

Typically, the previously mentioned detectors assume prior knowledge of the cycle period, e.g. [4,35,36,38]. However, in practice the period might be unknown or not known exactly due to clock or oscillator errors, which decreases the performance of the detectors [43]. Moreover, the cycle period itself is an interesting signal parameter. For instance, it relates to the symbol rate and carrier frequency [9,10].

For this reason, we consider the more general case of almost-cyclostationary detection since in practice sampling of a CS time series will turn into a discrete-time ACS signal [8]. This will generally be the case if the cycle period of the continuous-time signal is unknown [44]. This problem results in the following modified detection problem:

$$\begin{aligned} \mathcal{H}_0 :&\quad \text{measurement is WSS,} \\ \mathcal{H}_1 :&\quad \text{measurement is ACS with unknown cycle period.} \end{aligned} \tag{1.2}$$

In this thesis a joint detector and cycle period estimator is derived to approach (1.2). To this end, a resampling approach is combined with a multiple hypothesis test. By minimizing the family-wise error rate and simultaneously comparing the test statistic from [45] of the resampled signal to a threshold determined based on order-statistics, it was possible to show that the proposed combined detector outperforms the comparable state-of-the-art detectors.

1.1.3 Application in passive radar

A *passive radar* system consists of one or more receivers and exploits one or more *non-cooperative* transmitters, which are also referred to as *illuminators of opportunity* (IO). Exploiting non-cooperative transmitters is the key idea of a passive radar. Where an active radar has full control of the signal used to sense targets, a passive radar utilizes other transmitters in its surrounding. Typically, IOs are commercial radio or TV broadcast system, or space-based sources such as communication or navigation satellites [46–48]. The lack of a dedicated transmitter as part of the system itself makes the passive radar especially interesting as it is cheap, simple and undetectable [49]. However,

it comes at a cost since utilizing an IO does not only provide advantages: it makes the detection problem more challenging as well. Before pointing out the challenges let us briefly introduce different topologies of passive radar systems.
The simplest setup is the passive bistatic radar (PBR) that consists of one receiver placed remotely from one IO. The passive multistatic radar consists of multiple transmitters and multiple distributed receivers in which multiple bistatic measurements are combined [50]. In [51] the system consisting of multiple transmitters and multiple receivers is referred to as a passive MIMO radar (PMR), which does not only consider the individual bistatic pairs but which exploits all transmitter-receiver combinations.

Let us consider the PBR system for now. The measurements received are typically separated into two channels. One is the *direct-path* signal from IO to receiver, which will serve as a reference signal. For this reason it is commonly called *reference channel* (RC). This reference signal will always be a *noisy version* of the signal transmitted by the IO. Moreover, it may be corrupted by clutter and interference. The second channel is composed of the *target-path* signal, which is a delayed version of the signal transmitted by the IO, reflected by the target, and noise. We refer to this channel as *surveillance channel* (SC). This channel, however, may be corrupted by the direct-path interference of the IO and again it may also be corrupted by clutter and interference in general. In order to handle all these kinds of interferences at the receiver, various algorithms have been derived to handle these problems: The cancellation of interference and clutter in the RC has been considered in e.g. [52], [53]. 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 observed at the surveillance array. Clutter, interference, or direct-path signal in the target-path signal can be prevented by either physical shielding [50] or cancelled by signal processing techniques presented in e.g. [54–57]. The complete cancellation of direct-path interference in the SC is, admittedly, an idealized assumption as was pointed out in [54] and the works in [58–60] have considered the direct-path interference in their signal models.

Assuming that all these kinds of interferences have been canceled, we can consider the system illustrated in Figure 1.1, which is a simple MIMO PBR system. The goal is to detect the presence of a target echo at the SC given measurements at SC and RC. Various techniques have been derived to approach this problem. The most common and intuitive approach for single-input single-out (SISO) signals is based on cross-correlating the signals at SC and reference channel (RC), e.g., [53, 57, 58, 61–63]. Although this resembles the matched filter, it is suboptimal due to noise at the RC [58]. In [64–67] GLRT were derived for the case of unknown stochastic waveforms and for various assumptions on the signal and noise models for MIMO signals. The GLRTs for the case of unknown deterministic waveforms for different assumptions on signal and noise models were presented in [51, 68]. Bayesian tests were derived in [68, 69]. The work in [70] proposed an ad-hoc detector based on the generalized coherence [71].

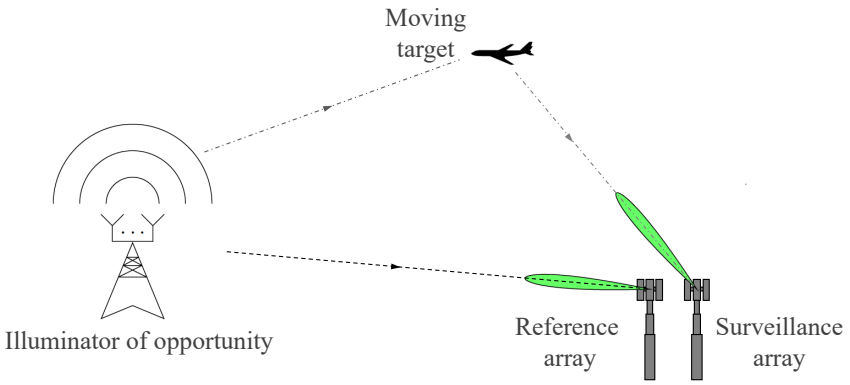

Figure 1.1: MIMO passive bistatic radar system that consists of an IO, a reference, and a surveillance array. The reference array receives the direct-path signal from the IO illustrated by the black dashed line and in the presence of a moving target the surveillance array receives the target-path signal, which is depicted by the gray dashed dotted line.

In this thesis, the CS property of the IO signal will be exploited to approach the detection problem. Generally, this may be regarded as a generalization of the single array detection problem, which was considered for the application in spectrum sensing for cognitive radio with additional reference arrays, which is further generalized to the case of multiple surveillance and multiple reference arrays. The GLRTs are derived for different variations of the setup and it is shown that neither the uniformly most powerful invariant test (UMPIT) nor an LMPIT exist for the problem.

1.2 Contributions

The main goal of this thesis is to derive detectors for cyclostationarity in one or multiple surveillance channels given one or multiple reference channels based on solid statistical principles.
First, a generalized version of the single array detection problem is considered with the goal to derive an ACS detector for the case of unknown cycle period. This problem is formulated as a multiple hypothesis test combined with a resampling stage that allows to jointly detect a signal and to estimate its cycle period. By minimizing the family-wise error rate and simultaneously comparing the test statistic from [45] of the resampled signal to a threshold determined based on order-statistics and the fact that the statistic is asymptotically distributed as a product of independent Beta random variables under the null hypothesis, it was possible to show that the proposed combined detector outperforms comparable state-of-the-art detectors.

Furthermore, an asymptotic GLRT for the problem of MIMO two-channel detection is derived. To this end, the maximum likelihood (ML) estimates of the covariance matrices, which have a block-Toeplitz structure, must be determined. Since there exists no closed-form solution for the MLE of (block) Toeplitz covariance matrices, an asymptotic result from [38] is used, which allows to obtain approximate closed-form ML estimates of the covariance matrices under both hypotheses.

In addition, the existence of the uniformly most powerful invariant test (UMPIT) and the LMPIT for the same detection problem is examined. To this end, Wijsman's theorem [72] is exploited, which avoids the necessity of deriving the maximal invariant statistic and its distribution under both hypotheses. It only requires to identify all groups of invariant transformations. It has been shown that neither the UMPIT nor the LMPIT exist for the problem since the ratio of the distributions of the maximal invariant statistic still depends on unknown parameters. Nevertheless, the interpretation of this ratio (depending on an unknown parameter) together with extensive results of Monte Carlo simulations, provides a rationale for proposing an LMPIT-inspired detector with the potential to outperform the GLRT. It is shown that the GLRT inherently merges the information provided by 1) the presence of cyclostationarity at the SC via a coherence matrix computed only based on observations at the SC and 2) the correlation of SC and RC present in a cross-coherence matrix that combines the cross-correlation with the cross-spectral correlation between SC and RC. In the ratio of the distributions of the maximal invariant statistic, however, these terms are connected by an unknown parameter. It is demonstrated that depending on the SNR at both channels it may be sufficient to include only the information provided by the cross-coherence matrix into the test statistic, which is the proposed LMPIT-inspired detector. It has shown to have the capability to outperform the proposed GLRT and both proposed detectors outperform comparable state-of-the-art detectors.

The two-channel detection problem is also generalized to the case of multiple surveillance and multiple reference arrays. Similar to the two-channel case the GLRT combines the spectral correlations of all SC as well as the cross correlation and cross-spectral correlations of all RC and SC. It is shown that neither UMPIT nor LMPIT exist for this problem. It could be considered, however, to propose another LMPIT-inspired test similar to the two-channel case.

The aforementioned derivations considered the scenario of noise that is spatially and temporally correlated. Two-channel GLRTs for noise with spatio-temporal structure have been derived, which yield three additional test statistics for the following kinds of noise models: 1) temporally white and spatially uncorrelated, 2) temporally white and spatially correlated, and 3) temporally colored and spatially uncorrelated. This has been published in [73], where it is concluded that, depending on the model, it is essential to account for appropriate temporal structure of the noise when selecting a test statistic.

In order to apply all the aforementioned test statistics it is necessary to select an appropriate threshold. For the GLRT statistics, a representation of the null distribution that can be used to determine a threshold for a fixed probability of false alarm is derived based on the work in [74]. Specifically, by considering the invariant transformation that allows for whitening the observations, the distribution of the proposed GLRT statistics is shown to be a product of independent Beta distributed random variables, where the degrees of freedom of the Beta random variables depend on the number of surveillance and reference arrays, the number of antennas per array, the assumptions on noise, the cycle period, and the total number of samples.

1.3 Outline

This thesis is structured as follows: In the following chapter, basic concepts of random processes including wide-sense stationary and CS processes are outlined. Additionally, a brief overview of relevant concepts of detection theory is presented. The single array detection problem of almost-cyclostationarity with unknown cycle period is presented in Chapter 4. Thereafter, in Chapter 5, the GLRT for the two-channel cyclostationarity detection is derived. This is complemented by examining the existence of UMPIT and LMPIT for the same problem and followed by proposing an LMPIT-inspired test. In Chapter 6 the GLRT for the two-channel problem under different assumptions on the spatio-temporal structure of the noise is derived. Finally, Chapter 7 generalizes the two-channel detection problem to multiple surveillance and reference channels and also investigates the existence of optimal invariant tests, where again it turns out that they do not exist.

2 Preliminaries

In this chapter we briefly review basic statistical signal processing concepts used within this thesis. In the first section, we introduce random processes with focus on wide-sense stationary (WSS) random processes (RP) and cyclostationary RP. Afterwards, we will concisely introduce the most relevant tools of hypothesis testing, which will be exploited in this work. In particular, we will present optimal and optimal invariant tests, generalized likelihood ratio tests and multiple hypothesis tests. The information provided in this chapter is mainly based on [75–79].

2.1 Random processes

Random processes are a statistical tool to model random and time-varying phenomena, which arise naturally but also in man-made contexts. Naturally occuring RPs include quantities such as temperature or air pressure whereas man-made RP can originate in technical contexts such as speech or communication signal processing but also in economics. In communication systems, for instance, randomness is induced by fluctuating currents and voltages due to noise (background noise, i.e., impinging noise and interfering signals). Also the information-carrying signals themselves are modeled as random signals varying over time [77]. Within this thesis we deal with the second-order theory of random processes, i.e., we deal with first and second-order moments. In theory those are expected values. In practice, however, we have to estimate the quantities of interest, which is oftentimes performed by time averaging instead of taking ensemble averages.

Before we get to the first and second-order moments of interest, we introduce complex-valued random vectors. Then we extend this concept to the statistical characterization of continuous-time RPs. Finally, we consider discrete-time processes. The information in the following sections is based on [75, 77] and the notation is predominantly consistent with [75].

We denote a complex-valued random vector by $\mathbf{x} : \Omega \to \mathbb{C}^L$, which we define as $\mathbf{x} = \mathbf{u} + j\mathbf{v}$ with $\mathbf{u} : \Omega \to \mathbb{R}^L$ and $\mathbf{v} : \Omega \to \mathbb{R}^L$, where Ω denotes the sample space. Now, we can define the *joint cumulative probability* of $\mathbf{u}$ and $\mathbf{v}$ as

$$P(\mathbf{x}) \triangleq P(\mathbf{u}_0 + j\mathbf{v}_0) = \text{Prob}(\mathbf{u} \leq \mathbf{u}_0, \mathbf{v} \leq \mathbf{v}_0) \tag{2.1}$$

and the *joint probability density function*

$$p(\mathbf{x}) \triangleq p(\mathbf{u} + j\mathbf{v}) = \frac{\partial^2}{\partial \mathbf{u} \partial \mathbf{v}} P(\mathbf{u}, \mathbf{v}). \tag{2.2}$$

If $\mathbf{u}$ and $\mathbf{v}$ depend on time t, $\mathbf{x}(t) = \mathbf{u}(t) + j\mathbf{v}(t)$ is a continuous-time *random process*, which at a given point in time t_0 reduces to a random variable, i.e., $\mathbf{x}(t_0)$. The RP $\mathbf{x}(t)$ is described by an ensemble of sample functions $\{\mathbf{x}_n(t)\}$, which are realizations of the RP.

Typically, we focus on moment functions that summarize the most important statistical characteristics. Let us introduce the first and second-order moments for complex-valued RP. The *mean function* of the random process is given by

$$\boldsymbol{\mu}_x(t) = \mathrm{E}\left[\mathbf{x}(t)\right] \in \mathbb{C}^L, \tag{2.3}$$

where $\mathrm{E}\left[\cdot\right]$ denotes the expectation operator defined as

$$\mathrm{E}\left[f(\mathbf{x}(t))\right] = \iint f\left(\mathbf{u}(t) + j\mathbf{v}(t)\right) p(\mathbf{u}(t) + j\mathbf{v}(t))\, d\mathbf{u} d\mathbf{v} \tag{2.4}$$

The matrix-valued *autocorrelation function* is defined as

$$\mathbf{M}_{\mathbf{xx}}(t, \tau) = \mathrm{E}\left[\mathbf{x}(t)\mathbf{x}^H(t-\tau)\right] \in \mathbb{C}^{L \times L} \tag{2.5}$$

and the *complementary* matrix-valued autocorrelation function

$$\tilde{\mathbf{M}}_{\mathbf{xx}}(t, \tau) = \mathrm{E}\left[\mathbf{x}(t)\mathbf{x}^T(t-\tau)\right] \in \mathbb{C}^{L \times L}. \tag{2.6}$$

Similarly, we can define the matrix-valued *covariance function* and *complementary covariance* function as

$$\mathbf{R}_{\mathbf{xx}}(t, \tau) = \mathrm{E}\left[\left(\mathbf{x}(t) - \boldsymbol{\mu}(t)\right)\left(\mathbf{x}(t-\tau) - \boldsymbol{\mu}(t-\tau)\right)^H\right] \in \mathbb{C}^{L \times L} \tag{2.7}$$

and

$$\tilde{\mathbf{R}}_{\mathbf{xx}}(t, \tau) = \mathrm{E}\left[\left(\mathbf{x}(t) - \boldsymbol{\mu}(t)\right)\left(\mathbf{x}(t-\tau) - \boldsymbol{\mu}(t-\tau)\right)^T\right] \in \mathbb{C}^{L \times L}, \tag{2.8}$$

respectively. An important property of complex-valued random vectors is the case in which a random vector $\mathbf{x}$ is called *proper*. This is the case when the complementary covariance function is identical zero $\tilde{\mathbf{R}}_{\mathbf{xx}}(t, \tau) = \mathbf{0}$. If the complementary covariance matrix is non-zero, $\mathbf{x}$ is called *improper* [75]. Note that for zero-mean signals the (auto-) correlation function and covariance function are identical. For the sake of simplicity, from here on we will consider zero-mean signals throughout the thesis. In order to completely characterize a random process all n order statistics have to be specified. Certain classes of RP, e.g. Gaussian RP, can be completely characterized by their first and second-order statistics.

RPs that are characterized for at least one n order time-varying joint probability density functions and corresponding time-varying moment functions are generally referred to as *non-stationary* RP. However, there exist some classes of RP of great practical interest that reveal certain properties regarding the first and second-order moments that we will introduce in the following two sections.

2.1.1 Wide-sense stationary random processes

Wide-sense stationary random processes are an important class of RPs. They find their application, for instance, in communications for modeling signals and noise. Thermal noise, for instance, depends on temperature, which is a time-invariant effect resulting in time-invariant statistical parameters [77]. An RP is defined to be WSS if the mean function is time-invariant and the covariance function is shift-invariant. Hence, the matrix-valued covariance function only depends on the time-shift τ

$$\mathbf{R}_{\mathbf{xx}}(\tau) = \mathrm{E}[\mathbf{x}(t)\,\mathbf{x}^{H}(t-\tau)], \tag{2.9}$$

whereas all higher-order moments may be time-variant. The complementary covariance function is defined equivalently to (2.7).

So far we have only considered continuous-time RP. Similarly, we can introduce the discrete-time RP $\mathbf{x}[n]$ with $n \in \mathbb{Z}$. The matrix-valued covariance function is defined as

$$\mathbf{R}_{\mathbf{xx}}[n,m] = \mathrm{E}[\mathbf{x}[n]\mathbf{x}^{H}[n-m]] \in \mathbb{C}^{L\times L}, \tag{2.10}$$

with time instant n and time-shift m. In the case of a WSS RP, the covariance function only depends on the time-shift m, i.e. $\mathbf{R}_{\mathbf{xx}}[n,m] = \mathbf{R}_{\mathbf{xx}}[m]$. Again the complementary covariance function is defined accordingly. Furthermore, the covariance function allows a frequency domain interpretation of the signal $\mathbf{x}[n]$. First note that the Fourier transform of a WSS process does not exist since it has infinite energy. However, a spectral representation of the process can be provided considering that it has a finite power on average. Although, we cannot Fourier transform $\mathbf{x}[n]$ itself, we can, however, Fourier transform the covariance function with respect to τ in the continuous-time case and with respect to m in the discrete-time case. This transformation is referred to as the *power spectral density* (PSD) with $\theta \in (-\pi, \pi)$

$$\mathbf{\Pi}(\theta) = \sum_{n=-\infty}^{\infty} \mathbf{R}_{\mathbf{xx}}[m] e^{-j\theta n}. \tag{2.11}$$

Complementary, we can express the process as follows

$$\mathbf{x}[n] = \int_{-\pi}^{\pi} e^{j\theta n} d\boldsymbol{\xi}(\theta), \tag{2.12}$$

where $\boldsymbol{\xi}(\theta)$ is a spectral process and $d\boldsymbol{\xi}(\theta)$ denotes its orthogonal increments. This expression allows for expressing the PSD as the second-order moment of $\boldsymbol{\xi}(\theta)$, i.e.

$$\boldsymbol{\Pi}(\theta)\, d\theta = \mathrm{E}\left[d\boldsymbol{\xi}(\theta)\, d\boldsymbol{\xi}^{H}(\theta)\right]. \tag{2.13}$$

This representation of a spectral process $\boldsymbol{\xi}(\theta)$ is again of interest when discussing another class of random processes namely cyclostationary random processes and allows for nice interpretations as we will see in later chapters.

Let us now consider the case in which we collect N observations of a zero-mean vector-valued WSS process $\mathbf{u}[n] \in \mathbb{C}^L$. It is convenient to stack these observations in one vector in order to investigate its covariances, i.e.

$$\mathbf{x} = \begin{bmatrix}\mathbf{u}[0] & \mathbf{u}[1] & \cdots & \mathbf{u}[N-1]\end{bmatrix}^{\mathrm{T}}. \tag{2.14}$$

The covariance matrix of $\mathbf{x}$ is given

$$\mathbf{R} = \mathrm{E}\left[\mathbf{x}\mathbf{x}^{H}\right] = \begin{bmatrix} \mathrm{E}\left[\mathbf{u}[0]\mathbf{u}^{H}[0]\right] & \mathrm{E}\left[\mathbf{u}[0]\mathbf{u}^{H}[1]\right] & \cdots & \mathrm{E}\left[\mathbf{u}[0]\mathbf{u}^{H}[N-1]\right] \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{E}\left[\mathbf{u}[N-1]\mathbf{u}^{H}[0]\right] & \mathrm{E}\left[\mathbf{u}[N-1]\mathbf{u}^{H}[1]\right] & \cdots & \mathrm{E}\left[\mathbf{u}[N-1]\mathbf{u}^{H}[N-1]\right] \end{bmatrix}. \tag{2.15}$$

Considering that a WSS process $\mathbf{u}[n]$ depends on the time-shift only, we can rewrite the covariance matrix as follows

$$\mathbf{R} = \begin{bmatrix} \mathbf{R}_{\mathbf{uu}}[0] & \mathbf{R}_{\mathbf{uu}}[-1] & \mathbf{R}_{\mathbf{uu}}[-2] & \cdots & \mathbf{R}_{\mathbf{uu}}[-(N-1)] \\ \mathbf{R}_{\mathbf{uu}}[1] & \mathbf{R}_{\mathbf{uu}}[0] & \mathbf{R}_{\mathbf{uu}}[-1] & \cdots & \mathbf{R}_{\mathbf{uu}}[-(N-2)] \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{R}_{\mathbf{uu}}[N-1] & \mathbf{R}_{\mathbf{uu}}[N-2] & \mathbf{R}_{\mathbf{uu}}[N-3] & \cdots & \mathbf{R}_{\mathbf{uu}}[0] \end{bmatrix}. \tag{2.16}$$

We can observe that there are identical $L \times L$ sized blocks $\mathbf{R}_{\mathbf{uu}}[m]$ on each diagonal of $\mathbf{R}$, i.e., $\mathbf{R}$ is a block-*Toeplitz* matrix with block size L. Note that $\mathbf{R}$ itself is a Hermitian positive definite matrix, the $L \times L$ blocks on the off-diagonals do not reveal any further structure but being positive semidefinite, and the main diagonal blocks are Hermitian and positive semidefinite.

2.1.2 Cyclostationary RP

In this section we introduce cyclostationary random processes, which are of key interest within this work. CS processes are a special class of non-stationary RPs that reveal periodicity. As already mentioned in the introduction, CS RP are used to model periodic phenomena that arise in many different fields such as climatology, mechanics, astronomy and communications [10]. These processes consist of random data with periodically varying statistical properties, however, the

processes themselves are not necessarily periodic functions of time. In engineering fields, such as communications, periodicity is induced by operations such as sampling, modulation, multiplexing and coding operations. In mechanics it originates, for instance, due to gear rotation. The information provided in this section mainly follows the works by Gardner and Schreier [10, 75, 77].

Within this thesis we only deal with second-order cyclostationarity, also referred to as wide-sense cyclostationarity. In contrast to WSS RP, cyclostationary RP have periodic first and second-order moments, i.e., mean function and correlation functions, vary periodically over time. Again without loss of generality, we deal with zero-mean processes. Hence, a zero-mean wide-sense CS RP $x(t) \in \mathbb{C}$ with *cycle period* T_0 is characterized by a periodic second-order moment, i.e. its covariance function is periodic:

$$r_{xx}(t,\tau) = \mathrm{E}[x(t)x^*(t-\tau)] = \mathrm{E}[x(t+T_0)x^*(t+T_0-\tau)] = r_{xx}(t+T_0,\tau). \tag{2.17}$$

The periodic covariance function $r_{xx}(t,\tau)$ can be expanded into a Fourier series assuming that the expansion converges [77]:

$$r_{xx}(t,\tau) = \sum_{n=-\infty}^{\infty} R_{xx}(n/T_0,\tau)\mathrm{e}^{j2\pi(\frac{n}{T_0})t}.$$

Here, $R_{xx}(n/T_0,\tau)$ are the Fourier coefficients obtained as

$$R_{xx}(n/T_0,\tau) = \frac{1}{T_0}\int_{-T_0/2}^{T_0/2} r_{xx}(t,\tau)\mathrm{e}^{-j2\pi(\frac{n}{T_0})t}\mathrm{d}t, \tag{2.18}$$

which we refer to as *cyclic autocorrelation functions* with *cycle frequencies* $\{\frac{n}{T_0}\}$ for $n \in \mathbb{Z}$.

We will also consider a more general class of CS RPs, which are referred to as *almost* cyclostationary processes. The previous paragraph considered a single cycle period T_0. If there are multiple periodicities, the cycle period of that process is their least common multiple. However, if the individual periodicities are incommensurate, the underlying process is almost CS (ACS). Since CS processes are a special case of ACS processes, we deal with ACS processes for the remainder of this section instead. To this end, let us assume that the covariance function $r_{xx}(t,\tau)$ is an almost periodic function in t. A function $x(t)$ is said to be almost periodic in the sense of Bohr [10, 80] if it is the limit of trigonometric polynomials, i.e.,

$$x(t) = \sum_{\alpha\in\mathcal{A}} x^\alpha e^{j2\pi\alpha t}, \tag{2.19}$$

where $\mathcal{A}$ is a countable set of possibly incommensurate frequencies and

$$x^\alpha = \lim_{T\to\infty}\frac{1}{T}\int_{-\frac{T}{2}}^{\frac{T}{2}} x(t)e^{-j2\pi\alpha t}\mathrm{d}t. \tag{2.20}$$

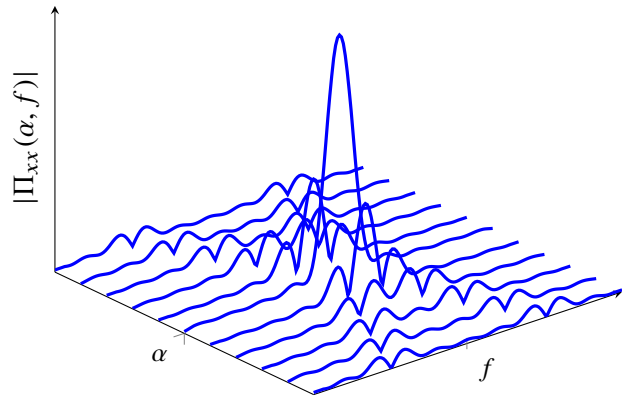

Figure 2.1: Magnitude of cyclic PSD of a rectangular shaped quadrature phase shift keying signal for different frequencies f and cycle frequencies α.

This allows us to expand the covariance function of an ACS signal $x(t)$ as follows

$$r_{xx}(t,\tau) = \sum_{\alpha\in\mathcal{A}} R_{xx}(\alpha,\tau)e^{j2\pi\alpha t}, \tag{2.21}$$

where

$$R_{xx}(\alpha,\tau) = \lim_{T\to\infty}\frac{1}{T}\int_{-\frac{T}{2}}^{\frac{T}{2}} r_{xx}(t,\tau)e^{-j2\pi\alpha t}\mathrm{d}t, \tag{2.22}$$

is the cyclic autocorrelation function at cycle frequency $\alpha \in \mathcal{A}$. It is easy to see that CS processes are a special case of ACS processes if $\mathcal{A} \equiv \{\frac{n}{T_0}\}$ for $n \in \mathbb{Z}$.

We can rewrite (2.22) by plugging in the definition of $r_{xx}(t,\tau) = \mathrm{E}[x(t)x^*(t-\tau)]$ and observe that an ACS signal $x(t)$ is correlated with its frequency-shifted version $x(t)e^{j2\pi\alpha t}$ when $\alpha \in \mathcal{A}$, i.e., CS signals exhibit correlation between some frequency-shifted versions of the process. Since the cyclic autocorrelation functions are Fourier coefficients, they allow for a frequency domain interpretation and the set of cycle frequencies α for which $R_{xx}(\alpha,\tau) \not\equiv 0$ is sometimes also referred to as *cycle spectrum*. We can further characterize the ACS process in the frequency domain by Fourier transforming $R_{xx}(\alpha,\tau)$ with respect to τ

$$\Pi_{xx}(\alpha, f) = \int_{-\infty}^{\infty} R_{xx}(\alpha,\tau)\mathrm{e}^{-j2\pi f\tau}\mathrm{d}\tau, \tag{2.23}$$

which we refer to as the *cyclic power spectral density*. Note that for $\alpha = 0$ it reduces to the usual PSD as introduced in (2.11). We illustrate an exemplary cyclic PSD quadrature phase shift keying (QPSK) signal in Figure 2.1. It can be observed that for a countable set of cycle frequencies α, the cyclic PSD is non-zero. The cyclic PSD for $\alpha = 0$ is identical to the usual PSD. Additionally, in Figure 2.2 we show a four-corners diagram that illustrates the relation between the functions of

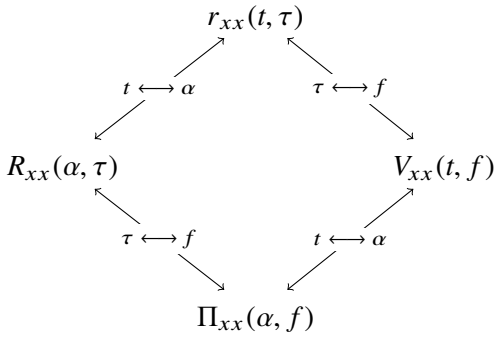

Figure 2.2: Four-corners diagram illustrating the relation between the autocorrelation function $r_{xx}(t,\tau)$, cyclic autocorrelation function $R_{xx}(\alpha,\tau)$, cyclic power spectral density $\Pi_{xx}(\alpha,f)$ and Rihaczek time-frequency representation $V_{xx}(t,f)$.

interest, i.e.,the autocorrelation function $r_{xx}(t,\tau)$, cyclic autocorrelation function $R_{xx}(\alpha,\tau)$ and the cyclic power spectral density $\Pi_{xx}(\alpha,f)$. For the sake of completeness, we also display the so-called *Rihaczek time-frequency representation* $V_{xx}(t,f)$ [75], which we will not further consider here.

As we have seen in the previous section, we can complementarily express the cyclic PSD as the second-order moment of the stochastic process

$$x(t) = \int_{-\infty}^{\infty} e^{jft} \mathrm{d}\xi(f), \tag{2.24}$$

as follows

$$\Pi_{xx}(\alpha,f)\mathrm{d}f = \mathrm{E}\left[\xi(f)\xi^*(f-\alpha)\right]. \tag{2.25}$$

Hence, the cyclic spectrum is a measure of the amount of correlation between $x(t)$ and its frequency shifted version $x(t)e^{j2\pi\alpha t}$. More generally, we can define the *Loève spectrum* or *bifrequency spectrum* for non-stationary processes $\mathrm{d}\xi(f)$ as

$$S(f_1,f_2)\mathrm{d}f_1\mathrm{d}f_2 = \mathrm{E}\left[\mathrm{d}\xi(f_1)\mathrm{d}\xi^*(f_2)\right]. \tag{2.26}$$

For a CS process the bifrequency spectrum can be expressed as a function of the cyclic PSD as follows

$$S(f_1,f_2) = \sum_{\alpha\in\mathcal{A}} \Pi_{xx}(\alpha,f_1)\delta(f_2-f_1-\alpha), \tag{2.27}$$

i.e. the support of the spectral correlation in the bifrequency spectrum consists of δ-ridges with unit slope and its density is given by the cyclic PSD. Moreover, for $\alpha = 0$ the cyclic PSD reduces to the usual PSD, which lies on the line with unit slope through the origin, which is also referred to as *stationary manifold*.

Anything what was presented in this section so far can be extended to the discrete-time case. Specifically, for a zero-mean CS process $x[n]$ with cycle period $P \in \mathbb{N}$ the covariance function is given by the following periodic function

$$r_{xx}[n, m] = \mathrm{E}[x[n]x^*[n-m]] = r_{xx}[n+P, m]. \tag{2.28}$$

Here we want to point out an important property of CS processes, namely that they can be stationarized. This was first shown by Gladyshev [3]. Specifically, we consider a polyphase decomposition of $x[n]$, or in other words, the set of sub-sampled time-series

$$x_i[n] = x[nP + i], \tag{2.29}$$

with $i = 0, \ldots, P-1$, are jointly WSS [3]. This property is easily derived by computing the cross-covariance function of $x_i[n]$ and $x_j[n]$

$$r_{X_i X_j}[n, m] = \mathrm{E}[x_i[n+m]x_j^*[n]] = E[x[(n+m)P + i]x^*[nP + j]], \tag{2.30}$$

which can be further simplified as

$$r_{X_i X_j}[n, m] = \mathrm{E}[x[nP + mP + i]x^*[nP + j]] = E[x[m + i]x^*[j]], \tag{2.31}$$

by exploiting the fact that the covariance function of a CS signal is periodic in P. Apparently, (2.31) is independent of the instant n and thus WSS.[1] Let us now arrange every P sub-samples of $x[n]$ into a vector

$$\mathbf{x}[n] = \begin{bmatrix} x[nP] & x[kP+1] & \ldots & x[(n+1)P-1] \end{bmatrix}^{\mathrm{T}}, \tag{2.32}$$

which is a vector-valued WSS process, i.e. its covariance sequence is shift-invariant. Finally, we can collect N samples of $\mathbf{x}[n]$ in the vector

$$\mathbf{y} = [\mathbf{x}[0], \cdots, \mathbf{x}[N-1]]^T. \tag{2.33}$$

The covariance matrix of $\mathbf{y}$ reveals the same structure as we have observed in the previous section, i.e., the covariance matrix of $\mathbf{y}$ is a block-Toeplitz structured matrix with block size LP.

[1] This can be similary shown for continous-time processes. A CS process with period T_0 can be stationarized by introducing a random time-shift. Let us consider the CS signal $x(t)$ with cycle period T_0. It has been shown in [5] that the process $y(t+\theta)$, where θ is a random variable uniformly distributed on the interval $[0, T_0)$ and which is statistically independent of $x(t)$, is a WSS process.

2.2 Detection

Detection theory, also referred to as hypothesis testing, is a fundamental technique to select one of a finite number of models for given measurements. Typical fields in signal processing for detection are, for instance, signal detection in radar and sonar, decoding symbols in communications, and speech recognition. As already mentioned in the introduction, in this thesis we will apply detection theory to differentiate between the presence of a signal and noise in the context of cognitive radio and passive radar. Here, we give a brief overview about the concepts that will be applied in the main part of this work. First, we present the Neyman-Pearson lemma that is an optimal test in some sense and secondly, the generalized likelihood ratio test is explained. The generalized likelihood ratio test is a practical approach to design a hypothesis test when the Neyman-Pearson test cannot be applied. Moreover, we give a brief insight into optimal and optimal invariant tests. The information provided in this section is primarily based on [76, 78, 79] and the notation is predominantly kept consistent with the work in [76].

2.2.1 Optimal and optimal invariant hypothesis tests

Let us consider a binary hypothesis test of

$$\begin{aligned} \mathcal{H}_0 : &\quad \boldsymbol{\theta} \in \Theta_0, \\ \mathcal{H}_1 : &\quad \boldsymbol{\theta} \in \Theta_1. \end{aligned} \tag{2.34}$$

Given a vector of measurements $\mathbf{x}$ with pdf $p(\mathbf{x}; \boldsymbol{\theta}_i)$ with $\boldsymbol{\theta}_i \in \Theta_i$ for $i \in \{0, 1\}$ the goal is to derive a test statistic, or detector, that takes the form

$$\phi(\mathbf{x}) = \begin{cases} 0 \sim \mathcal{H}_0, \ \mathbf{x} \in \mathcal{R}_0, \\ 1 \sim \mathcal{H}_1, \ \mathbf{x} \in \mathcal{R}_1, \end{cases} \tag{2.35}$$

where $\mathcal{R}_0$ and $\mathcal{R}_1$ denote the decision regions for $\mathcal{H}_0$ and $\mathcal{H}_1$, respectively. Before we present the test statistic that fulfills certain optimality constraints, we consider two different types of errors that occur in hypothesis testing. The first kind of error, the Type I error, or *false alarm*, is the probability of rejecting $\mathcal{H}_0$ although the measurement $\mathbf{x}$ is drawn from $\mathcal{H}_0$, i.e., the probability of false alarm, also referred to as size, p_{fa} is given by

$$p_{\text{fa}} = \int_{\mathbf{x} \in \mathcal{R}_1} p(\mathbf{x}; \boldsymbol{\theta}_0) d\mathbf{x}. \tag{2.36}$$

The Type II error is the probability of *miss detection*, this is, we fail to reject $\mathcal{H}_0$ although the measurement $\mathbf{x}$ is drawn from $\mathcal{H}_1$. Typically, the goal is to minimize this type of error, or equivalently, to maximize the detection probability, or power, which is given by

$$p_{\mathrm{d}} = \int_{\mathbf{x}\in\mathcal{R}_1} p(\mathbf{x};\boldsymbol{\theta}_1)d\mathbf{x}. \tag{2.37}$$

Neyman and Pearson [81] introduced the likelihood ratio test, which is the most powerful test of size p_{fa}, i.e., it maximizes the probability of detection given a probability of false alarm if a *simple* hypothesis $\mathcal{H}_0$ is tested versus a simple hypothesis $\mathcal{H}_1$. A simple hypothesis test refers to the case in which $\boldsymbol{\theta} = \boldsymbol{\theta}_i$ as opposed to being greater than or less than $\boldsymbol{\theta}_i$ as it is the case for *composite* hypotheses. The likelihood ratio is given by

$$l(\mathbf{x}) = \frac{p(\mathbf{x};\boldsymbol{\theta}_0)}{p(\mathbf{x};\boldsymbol{\theta}_1)} \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\gtrless}} \eta, \tag{2.38}$$

where η is chosen to meet the constraint on the probability of false alarm

$$p_{\mathrm{fa}} = \int_{\mathbf{x}:\, l(\mathbf{x})<\eta} p(\mathbf{x};\boldsymbol{\theta}_0)d\mathbf{x}. \tag{2.39}$$

Often it is possible to factorize the pdf $p(\mathbf{x};\boldsymbol{\theta})$ according to the Neyman-Fisher factorization theorem as follows [79]

$$p(\mathbf{x};\boldsymbol{\theta}) = f(t(\mathbf{x});\boldsymbol{\theta})g(\mathbf{x}). \tag{2.40}$$

If this factorization exists, then $t(\mathbf{x})$ is a *sufficient statistic*, i.e., the likelihood ratio simplifies to

$$l(\mathbf{x}) = \frac{f(t(\mathbf{x});\boldsymbol{\theta}_0)}{f(t(\mathbf{x});\boldsymbol{\theta}_1)} \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\gtrless}} \eta. \tag{2.41}$$

As we can see the statistic (2.41) only depends on data $\mathbf{x}$ through $t(\mathbf{x})$. Consequently, such a *scalar-valued* sufficient statistic contains all the relevant information of $\boldsymbol{\theta}$ that allows us to distinguish the hypotheses [79].

The Neyman-Pearson likelihood ratio test is the most powerful test if we deal with simple hypotheses. However, in practice we often deal with composite hypothesis tests, i.e., the pdf under $\mathcal{H}_0$ is not completely characterized by a single parameter but multiple parameters. Still it is desirable to derive an optimal test that provides maximum power, or probability of detection, for a given size, i.e., a given probability of false alarm. These tests are referred to as *uniformly most powerful tests* (UMP) tests. Their general definition is as follows [76]: A test $\phi(\mathbf{x})$ is UMP with a given p_{fa} if its power is uniformly greater than the power of any other test $\phi'(\mathbf{x})$ whose size is less than or equal to p_{fa}.

Now the Karlin-Rubin Theorem [76] states that if the likelihood ratio $l(x) = \dfrac{p(x;\theta_1)}{p(x;\theta_0)}$ with scalar measurement x is a monotone-likelihood ratio, i.e., a non-decreasing function of x, then the threshold test

$$\phi(x) = \begin{cases} 1, & x > x_0, \\ \gamma, & x = x_0, \\ 0, & x < x_0, \end{cases} \tag{2.42}$$

is UMP for the test $\mathcal{H}_0 : \ \theta \leq \theta_0$ vs. $\mathcal{H}_1 : \ \theta > \theta_0$. However, UMP tests exist for only very rare cases and, generally, only for one-sided hypothesis tests. For this reason the concept of *invariance* is used in order to derive optimal tests for a broader class of composite hypotheses.

A hypothesis test is invariant to a transformation group $\mathcal{G}$ if $\phi(g(\mathbf{x})) = \phi(\mathbf{x})$ for $g(\cdot) \in \mathcal{G}$ [78]. Moreover, a statistic $M(\mathbf{x})$ is said to be *maximally invariant* [76] if it is 1) invariant, i.e., $M(g(\mathbf{x})) = M(\mathbf{x})$ for all $g \in \mathcal{G}$, and 2) maximal, i.e. $M(\mathbf{x}_1) = M(\mathbf{x}_2)$ implies $\mathbf{x}_2 = g(\mathbf{x}_1)$ for some $g \in \mathcal{G}$. Making use of the concepts of invariance and sufficiency, measurements may be compressed into a low dimensional test statistic for which a UMP test exists. Although this may seem rather theoretical, in Chapter 5 and Chapter 7 we show that the invariant transformations arise naturally to the problem. Now a uniformly most powerful *invariant* test (UMPIT) can be defined as follows [76]: An invariant test $\phi(\mathbf{x}) = \phi(M(\mathbf{x}))$ is UMP with given p_{fa} if its power is uniformly greater than the power of any other test $\phi'(\mathbf{x})$ that is also invariant to $\mathcal{G}$ and whose size is less than or equal to p_{fa}. In order to prove that an invariant test $\phi(M(\mathbf{x}))$ is UMPI, it has to be established that the likelihood ratio is a monotonically increasing function.

In [76] the author also summarizes the procedure to establish UMPI tests in the following seven steps

1. define the hypothesis testing problem: $\mathcal{H}_0 : \ \boldsymbol{\theta} \in \Theta_0$ vs. $\mathcal{H}_1 : \ \boldsymbol{\theta} \in \Theta_1$
2. replace the measurements $\mathbf{x}$ by the sufficient statistic $t(\mathbf{x})$ if possible
3. reformulate the hypothesis test for the sufficient statistic $t(\mathbf{x})$
4. find a meaningful group of invariant transformations $\mathcal{G}$
5. find a maximal invariant statistic $M(t(\mathbf{x}))$
6. construct the likelihood ratio, i.e., derive the densities of the maximal invariant statistic under both hypotheses and form their ratio
7. check if the likelihood ratio of the maximal invariant statistic is monotonically increasing. If that is the case, the test is UMPIT for testing $\mathcal{H}_0 : \ \theta \leq \theta_0$ vs. $\mathcal{H}_1 : \ \theta > \theta_0$.

Although in some cases invariance and sufficiency allows to get rid of unknown nuisance parameters that finally yield a UMPIT, in practice this only happens in rare cases. Still it might be possible to find an optimal test that is UMP within a certain range of the parameter, i.e., for close hypotheses. Based on a Taylor series approximation of the ratio of the maximal invariants, it may be possible to get rid of the dependence on unknown parameters. These tests are referred to as locally most powerful invariant tests (LMPIT) [78].

In order to derive an LMPIT it is still necessary to complete steps 1. to 6. of finding a UMPIT, which may be a challenging task. Especially, deriving the maximal invariant statistic may be difficult and establishing its distribution under both hypotheses is even more involved since it can be a complicated function of the measurements. However, Wijsman [72] has proven that the ratio of densities of the maximal invariant is given by

$$\mathscr{L} = \frac{\int_{\mathcal{G}} p(g(\mathbf{x}); \mathcal{H}_1) |\det(\mathbf{J}_g)| dg}{\int_{\mathcal{G}} p(g(\mathbf{x}); \mathcal{H}_0) |\det(\mathbf{J}_g)| dg}, \tag{2.43}$$

where $\mathbf{J}_g$ denotes the Jacobian matrix of the transformation $g(\cdot)$, and finally dg denotes the invariant group measure, which in usual signal processing applications is the Lebesgue measure, which allows us to evaluate the resulting integrals [82]. The concept behind integrating over the invariance group is averaging out the unknown parameters in the ratio. This concept was first used by Stein in [83], who proposed to average over a group to obtain the density of a maximal invariant statistic. Based on this ratio it can be established whether a UMPIT exists or if it is necessary to settle for an LMPIT by expanding the ratio into a Taylor series focusing on close hypotheses, which exists only if the final statistic does not depend on unknown parameters.

2.2.2 Generalized Likelihood Ratio Test

A different concept to deal with composite hypotheses are generalized likelihood ratio tests (GLRTs), which replace the unknown parameters in the likelihood ratio by their MLEs [79] rather than averaging them out as it is shown in Wijsman's theorem (2.43). The GLRT is given by

$$\mathscr{G} = \frac{p(\mathbf{x}; \hat{\boldsymbol{\theta}}_0)}{p(\mathbf{x}; \hat{\boldsymbol{\theta}}_1)} \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\gtrless}} \eta, \tag{2.44}$$

where $\hat{\boldsymbol{\theta}}_i$ for $i = 0, 1$ denote the ML estimates of $\boldsymbol{\theta}_i$. The test does not guarantee optimality for a finite number of measurements, however, it is asymptotically UMPI [82]. In practice it has been shown that it often performs well even if there are only finite data records available.

2.2.3 Multiple hypotheses tests

So far we have only considered binary hypothesis tests. However, comparing multiple (null) hypotheses[2], such as $\mathcal{H}_1$ vs. $\mathcal{H}_2$ vs. $\cdots$ vs. $\mathcal{H}_K$, is of practical interest as well. In order to draw inference in a multiple hypothesis test, we have to concern about error control. Different from dealing with binary tests there are several types of error control that may be of interest. Traditionally, the so-called *familywise error rate* (FWER), which is the error of having at least one false rejection, i.e.,

$$\text{FWER} = P\left(\text{reject any } \mathcal{H}_i\right), \tag{2.45}$$

is controlled at level α, i.e., FWER $\leq \alpha$. There have been various procedures proposed that control the FWER such as Bonferroni's correction [78] or Holm's sequential test procedure [84], which we will briefly describe at the end of this section.

Typically, procedures that control either the FWER or the FDR are based on p-values. Formally, a p-value is defined as

$$p = p(x) = \inf\{\alpha : \ x \in S_\alpha\}, \tag{2.46}$$

where S_α denotes the rejection region. In other words the p-value is the smallest significance level at which a hypothesis is rejected. For instance, for the case of a right-sided hypothesis test the p-value can be expressed as

$$p = p(x) = 1 - F_X(x), \tag{2.47}$$

where $F_X(x)$ is the cumulative distribution function of X under hypothesis $\mathcal{H}$. As can be seen, a p-value indicates how strongly a measurement x contradicts the underlying hypothesis, i.e., the smaller the p-value the less likely the hypothesis.

Suppose we want to test the hypotheses $\mathcal{H}_1$ vs. $\mathcal{H}_2$ vs. $\cdots$ vs. $\mathcal{H}_K$ and the goal is to control the FWER at level α. Exemplarily, we will present Holm's sequential test [84]. Firstly, the p-values $p_1, \ldots, p_K$ for all K hypotheses are computed and ordered as $p_{(1)} \leq p_{(2)} \leq \cdots p_{(K)}$ associated to hypotheses $\mathcal{H}_{(1)}, \mathcal{H}_{(2)}, \ldots, \mathcal{H}_{(K)}$. Then the test procedure sequentially tests for $i = 1, \ldots, K$ whether $p_{(i)} \geq \alpha/(K - i + 1)$, if that holds, we fail to reject $\mathcal{H}_{(1)}, \ldots, \mathcal{H}_{(K-i+1)}$. If the condition does not hold, we reject $\mathcal{H}_{(1)}, \ldots, \mathcal{H}_{(i)}$, increment i, and continue until we fail to reject the hypotheses.

[2] These kinds of test may also be viewed as a set of binary tests with null hypotheses $\mathcal{H}_i$.

3 Stochastic representation of the generalized likelihood ratio test statistics under the null hypothesis

In this chapter we derive stochastic representations of a specific class of GLRT statistics under the null hypothesis. Specifically, we consider a general class of problems encountered in various fields of application, which boil down to test different block-diagonal structures of the covariance matrix under both hypotheses. In this chapter, we do not consider a particular application of the derived test statistics, nevertheless, they provide a basis for stochastic representations we will encounter throughout the thesis and we will make use of the results and derivations made in this chapter.

3.1 Introduction

In many fields of science such as engineering, medicine, and economics, it is a common challenge to test either for independence between multiple random vectors, their equivalence, or a combination thereof. The typical approach to solve these detection problems is the likelihood ratio test that, assuming Gaussianity of the data, results in tests for the structure of the covariance matrices. Hence, under the null hypothesis the covariance has a particular block-diagonal structure and under the alternative, it has either no further structure or it is block-diagonal itself, where the blocks do not reveal further structure except being positive semi-definite (p.s.d.). These tests are composite hypothesis tests, which are typically approached by generalized likelihood ratio tests (GLRTs).

There is an abundance of literature in which the GLRTs for problems that reveal various kinds of covariance matrix structures are derived. The test statistics of these tests are ratios of determinants of functions of the sample covariance matrix, which are referred to as generalized Hadamard ratios. In [85] the author derived the test for independence of univariate real-valued Gaussian observations in the context of economics, psychology, and anthropology. This test examines whether the sample covariance matrix is diagonal or whether it is a full p.s.d. matrix. In [86, 87] the authors derived the GLRT for a similar problem in the context of radio astronomy, where the independence of complex-valued Gaussian signals with uncalibrated and nonidentical sensors is established.

This test was generalized to the vector-valued case in [74, 88] with applications to multi-channel signal detection in networks of sensor arrays, radar detection with multiple antennas, or cognitive radio. The authors derived the GLRT to detect spatially correlated complex Gaussian vectors from multiple antennas for arbitrary temporal structure. This aims at testing whether the sample covariance matrix has block-diagonal structure or whether it is only p.s.d., i.e., whether it is a test for independence of blocks.

Furthermore, in [89, 90] the authors adressed the problem of testing for equivalence of covariance matrices.This is the test of having a diagonal covariance matrix with identical elements on its main diagonal versus a full p.s.d. covariance matrix. This problem is referred to as sphericity test [89] and was generalized to the block-sphericity test in [90, 91], where the goal is to examine whether the main diagonal blocks of the covariance matrix are identical.

In the context of the detection of cyclostationarity with applications in cognitive radio and passive radar, the aforementioned tests of covariance matrix structures also arise. We will consider these applications in the subsequent chapters, where we will encounter the following matrix structures (asymptotically): The covariance matrix under the alternative is no longer a p.s.d. matrix without further structure but it is also a block-diagonal matrix, where each block is p.s.d. without further structure. Under the null hypothesis, the structure of the test depends on the spatio-temporal structure of the noise and results in either tests for block-independence [4] or block-sphericity tests [92]. These tests are generalized to two-channel problems in [73, 93], that makes the structure of the covariance matrices more complex as we derive in Chapters 5 and 6. In Chapter 7, we further generalize this problem to multiple channels, which also increases the variety of matrix structure under the null and the alternative hypotheses.

All of the aforementioned GLRTs have in common that their test statistics are ratios of determinants of main-diagonal blocks of the sample covariance matrix or functions thereof. These ratios are referred to as Hadamard or generalized Hadamard ratios [88]. Since the application of these tests requires to set a threshold that fixes the probability of false alarm at a desired level, we need to know the distribution of the test statistic under the null hypothesis. Unfortunately, the exact distributions of these generalized Hadamard ratios can only be established for some special cases as shown in [90, 94], which is a tedious process. For this reason, it is common to resort on asymptotic distributions that are available in closed-form such as the well known Wilks' Theorem [95]. However, we do not always operate in asymptotic regimes if there are insufficiently many observations available.

To overcome this problem, a stochastic representation for various kinds of ratios can be derived. For instance, in [74] the authors derived the stochastic representation of the GLR under the null hypothesis for the detection of spatially correlated time series. In [90] the distribution was derived for the (block-) sphericity tests. The authors in [96] generalized the model of the covariance under null hypothesis for a combination of independence and sphericity tests. Additionally, they derive the

stochastic representation for this general GLRT under the null hypothesis. In [97, 98] the stochastic representation for testing different block-diagonal structures in a single channel has been proposed and it has been derived for the two-channel case in [99]. In [100] the authors also derive the null distributions for various coherence statistics.

In this chapter we further generalize the problem based on the work in [96] by additionally imposing block-diagonal structure to the covariance matrix under the alternative, which allows for an application in [4, 73, 92, 93]. All of these stochastic representations of Hadamard ratios under the null have in common that they are products of independent Beta random variables, which do not depend on data but simply on the parameters of the setup. Hence, these distributions can be obtained offline and allow for computing the thresholds to set the probability of false alarm.

3.1.1 Outline

First of all, we formulate the detection problem in a general fashion in Section 3.2. Second, we derive the GLRT in Section 3.3. This is divided into two different cases. The first one considers the case of independence and/or sphericity among block matrices without further structure, whereas the second case further specializes the test of independence and/or sphericity among block matrices that are themselves diagonal with distinct elements. In Section 3.4 we derive the stochastic representation of the GLR under the null hypothesis for the two cases mentioned in the preceding lines. Finally, in Section 3.5 we demonstrate that this representation shows agreement with the distribution of the test statistic under the null with numerical simulations.

3.2 Problem formulation

The problem considered in this chapter is testing for different block-diagonal structures of covariance matrices of zero-mean proper complex Gaussian random variables. Specifically, given a random vector $\mathbf{x} \in \mathbb{C}^N$ the test for the structure of its covariance matrix is the following

$$\begin{aligned} \mathcal{H}_0 &: \mathbf{x} \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_0), \\ \mathcal{H}_1 &: \mathbf{x} \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_1), \end{aligned} \tag{3.1}$$

where $\mathbf{R}_0 \in \mathbb{C}^{N \times N}$ is a block-diagonal matrix with the following structure

$$\mathbf{R}_0 = \mathrm{E}[\mathbf{x}\mathbf{x}^H | \mathcal{H}_0] = \begin{bmatrix} \mathbf{I}_{\delta_1} \otimes \mathbf{R}_0^{(1)} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{\delta_2} \otimes \mathbf{R}_0^{(2)} & \mathbf{0} & \vdots \\ \vdots & & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{I}_{\delta_K} \otimes \mathbf{R}_0^{(K)} \end{bmatrix}, \tag{3.2}$$

where $\mathbf{R}_0^{(k)} \in \mathbb{C}^{\kappa_k \times \kappa_k}$ for $k = 1, \ldots, K$ with $\sum_{k=1}^{K} \delta_k \kappa_k = N$ and $\delta_k, \kappa_k > 0$. Moreover, we consider two different structures for $\mathbf{R}_0^{(k)}$. The first one is $\mathbf{R}_0^{(k)}$ being a p.s.d. matrix without further structure and the second one being $\mathbf{R}_0^{(k)}$ to be a diagonal matrix with positive diagonal elements itself, i.e.,

Case 1) $\mathbf{R}_0^{(k)}$ is a positive semi-definite matrix without further structure,

Case 2) $\mathbf{R}_0^{(k)}$ is a diagonal matrix itself.

Note that it is possible to combine Cases 1) and 2) to more sophisticated structures under the null hypothesis. However, since the combination does not provide additional value here and for the sake of simplicity of the notation, we treat the cases separately.

The covariance matrix $\mathbf{R}_1 \in \mathbb{C}^{N \times N}$ under the alternative is given by

$$\mathbf{R}_1 = \mathrm{E}[\mathbf{x}\mathbf{x}^H | \mathcal{H}_1] = \begin{bmatrix} \mathbf{R}_1^{(1)} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_1^{(2)} & \mathbf{0} & \vdots \\ \vdots & & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{R}_1^{(L)} \end{bmatrix}, \tag{3.3}$$

where $\mathbf{R}_1^{(l)} \in \mathbb{C}^{\lambda_l \times \lambda_l}$ for $l = 1, \ldots, L$ with $\sum_{l=1}^{L} \lambda_l = N$ and $L > 0$ is a positive semi-definite matrix without further structure. Moreover, we consider the block sizes of $\mathbf{R}_0$ and $\mathbf{R}_1$ to be related as $\lambda_l = \sum_{k=n_{l-1}+1}^{n_l} \delta_k \kappa_k$ with $n_0 = 1$, where n_l is the number of blocks under the null that are tested against a single block l without further structure under the alternative. In other words, the block size of block l under the alternative is always a multiple of block sizes under the null. Since this is the case throughout the thesis and in order to avoid even more complicated expressions in this chapter, we do not consider these cases within this chapter.

In Figure 3.1 we illustrate the structure of the covariance matrix under $\mathcal{H}_0$ for Case 1) as described in Section 3.5.

3.3 Derivation of the GLR

A common approach to deal with composite hypothesis tests is the GLRT, which we will employ for this problem as well. To this end, unknown parameters in the likelihood ratio are replaced by their MLEs, which, in our case, are the MLEs of the covariance matrices $\mathbf{R}_0$ and $\mathbf{R}_1$. Hence, the GLR is given by

$$\mathscr{G} = \frac{p(\mathbf{x}_0, \cdots, \mathbf{x}_{M-1}; \hat{\mathbf{R}}_0)}{p(\mathbf{x}_0, \cdots, \mathbf{x}_{M-1}; \hat{\mathbf{R}}_1)}, \tag{3.4}$$

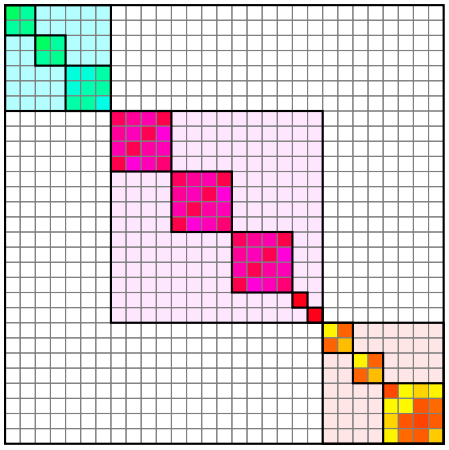

Figure 3.1: This figure illustrates the structure of the covariance matrices for Case 1) for the following parameters: $L = 3$ with $\lambda_1 = 7$, $\lambda_2 = 14$, and $\lambda_3 = 8$, i.e., $N = 29$ under the alternative displayed by the transparent colors. Under the null we have $K = 6$ diagonal blocks with $\delta_k = \{2, 1, 3, 2, 2, 1\}$ and $\kappa_k = \{2, 3, 4, 1, 2, 4\}$ for $k = 1, \ldots, 6$ illustrated by the opaque colors.

where $\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}$ denote M independent and identically distributed (i.i.d.) realizations of $\mathbf{x}$, and $\hat{\mathbf{R}}_0$ and $\hat{\mathbf{R}}_1$ denote the ML estimates of $\mathbf{R}_0$ and $\mathbf{R}_1$, respectively. Since $\mathbf{x}$ is zero-mean proper complex Gaussian, the likelihoods are given by

$$p(\mathbf{x}_0, \cdots, \mathbf{x}_{M-1}; \hat{\mathbf{R}}_j) = \frac{1}{\pi^{NM} \det\left(\hat{\mathbf{R}}_j\right)^M} \exp\left\{-M \operatorname{tr}\left(\mathbf{Q}\hat{\mathbf{R}}_j^{-1}\right)\right\}, \tag{3.5}$$

with $j \in \{0, 1\}$ indicating the MLE $\hat{\mathbf{R}}_0$ under $\mathcal{H}_0$ or the MLE $\hat{\mathbf{R}}_1$ under $\mathcal{H}_1$. The matrix $\mathbf{Q}$ denotes the sample covariance matrix of $\mathbf{x}$ given by

$$\mathbf{Q} = \frac{1}{M}\mathbb{X}\mathbb{X}^H, \tag{3.6}$$

where $\mathbb{X} = [\mathbf{x}_1 \cdots \mathbf{x}_M] \in \mathbb{C}^{N \times M}$ is the data matrix. Note that we assume that $M \geq \max\{\kappa_1, \ldots, \kappa_K, \lambda_1, \ldots, \lambda_L\}$ such that each block on the main diagonal is full-rank[1]. Based on the likelihood (3.5), we can derive the MLEs of the covariance matrices. To this end, we can partition the data matrix $\mathbb{X}$ in two different ways as follows

$$\mathbb{X} = \begin{bmatrix} \mathbf{X}_1 \\ \vdots \\ \mathbf{X}_K \end{bmatrix} = \begin{bmatrix} \mathbf{Y}_1 \\ \vdots \\ \mathbf{Y}_L \end{bmatrix}, \tag{3.7}$$

[1] If we dropped this assumption on M, the covariance matrices would reveal further structure that we do not consider in this problem.

where $\mathbf{X}_k \in \mathbb{C}^{\delta_k \kappa_k \times M}$ for $k = 1, \ldots, K$ and $\mathbf{Y}_l \in \mathbb{C}^{\lambda_l \times M}$ for $l = 1, \ldots, L$ are related to the sizes under $\mathcal{H}_0$ and $\mathcal{H}_1$, respectively. We can further decompose $\mathbf{X}_k$ as follows

$$\mathbf{X}_k = \begin{bmatrix} \mathbf{X}_k^{(1)} \\ \vdots \\ \mathbf{X}_k^{(\delta_k)} \end{bmatrix}, \tag{3.8}$$

and

$$\mathbf{X}_k^{(j)} = \begin{bmatrix} \mathbf{x}_{k\cdot 1}^{(j)} \\ \vdots \\ \mathbf{x}_{k\cdot \kappa_k}^{(j)} \end{bmatrix} \in \mathbb{C}^{\kappa_k \times M}, \tag{3.9}$$

for $j = 1, \ldots, \delta_k$.

The MLEs of $\mathbf{R}_0$ and $\mathbf{R}_1$ are straightforward to derive considering their block-diagonal structure and using results from complex-valued matrix differentiation [75, 101]. Hence, they are given by

$$\hat{\mathbf{R}}_0 = \begin{bmatrix} \mathbf{I}_{\delta_1} \otimes \bar{\mathbf{Q}}_{0,1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{\delta_2} \otimes \bar{\mathbf{Q}}_{0,2} & \mathbf{0} & \vdots \\ \vdots & & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{I}_{\delta_K} \otimes \bar{\mathbf{Q}}_{0,K} \end{bmatrix}, \tag{3.10}$$

where $\bar{\mathbf{Q}}_{0,k}$ is given by

Case 1) $\bar{\mathbf{Q}}_{0,k} = \frac{1}{\delta_k} \sum_{j=1}^{\delta_k} \mathbf{Q}_{0,k}^{(j)}$, for the case in which $\mathbf{R}_0^{(k)}$ for $k = 1, \ldots, K$ are p.s.d. matrices without further structure,

Case 2) $\bar{\mathbf{Q}}_{0,k} = \frac{1}{\delta_k} \sum_{j=1}^{\delta_k} \operatorname{diag}\left\{\mathbf{Q}_{0,k}^{(j)}\right\}$, for the case in which $\mathbf{R}_0^{(k)}$ for $k = 1, \ldots, K$ are diagonal matrices,

with $\mathbf{Q}_{0,k}^{(j)} = 1/M \mathbf{X}_k^{(j)} \left(\mathbf{X}_k^{(j)}\right)^H$ and

$$\hat{\mathbf{R}}_1 = \begin{bmatrix} \mathbf{Q}_{1,1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_{1,2} & \mathbf{0} & \vdots \\ \vdots & & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{Q}_{1,L} \end{bmatrix}, \tag{3.11}$$

with $\mathbf{Q}_{1,l} = 1/M \mathbf{Y}_l \left(\mathbf{Y}_l\right)^H$ for $l = 1, \ldots, L$.

Now we can plug the MLEs (3.10) and (3.11) into the likelihood ratio (3.4) to obtain the simplified expressions for both cases. The GLR for Case 1) is given by

$$\mathscr{G}_{1)}^{1/M} = \frac{\det\left(\hat{\mathbf{R}}_1\right)}{\det\left(\hat{\mathbf{R}}_0\right)} = \frac{\prod_{l=1}^{L}\det\left(\mathbf{Q}_{1,l}\right)}{\prod_{k=1}^{K}\det\left(\bar{\mathbf{Q}}_{0,k}\right)^{\delta_k}}. \tag{3.12}$$

and we further simplify it for Case 2), i.e., we consider a diagonal covariance matrix under the null

$$\mathscr{G}_{2)}^{1/M} = \frac{\prod_{l=1}^{L}\det\left(\hat{\mathbf{R}}_1^{(l)}\right)}{\prod_{k=1}^{K}\prod_{m=1}^{\kappa_k}\bar{q}_{k,mm}^{\delta_k}}, \tag{3.13}$$

where $\bar{q}_{k,mm}$ denotes the mth element on the diagonal of $\bar{\mathbf{Q}}_{0,k}$.

Having derived the test statistics we still lack the distributions for (3.12) and (3.13) given that the null hypothesis is true. The distribution allows for fixing the probabilty of false alarm at a desired level. In the literature exact distributions are tediously derived for special combinations of the parameters only in, e.g., [90, 94]. Instead we derive a stochastic representation for the general case considered here in the subsequent sections. We will show that we can factorize the GLRs into products of statistically independent scalars, where each of the scalar variables follows a Beta distribution.

3.4 Stochastic representation of the GLR under $\mathcal{H}_0$

In order to derive the distribution of the GLR, we exploit that $\lambda_l = \sum_{k=n_{l-1}+1}^{n_l}\delta_k\kappa_k$ to factorize the GLR in (3.12) into L terms as follows

$$\mathscr{G}_{i)}^{1/M} = \frac{\det\left(\mathbf{Q}_{1,1}\right)}{\prod_{k=1}^{n_1}\det\left(\bar{\mathbf{Q}}_{0,k}\right)^{\delta_k}} \times \cdots \times \frac{\det\left(\mathbf{Q}_{1,L}\right)}{\prod_{k=n_{L-1}+1}^{n_L}\det\left(\bar{\mathbf{Q}}_{0,k}\right)^{\delta_k}} = \prod_{l=1}^{L}\Lambda_l^{i)}, \tag{3.14}$$

where $i \in \{1, 2\}$ and

$$\Lambda_l^{1)} = \frac{\det\left(\mathbf{Q}_{1,l}\right)}{\prod_{k=n_{l-1}+1}^{n_l}\det\left(\bar{\mathbf{Q}}_{0,k}\right)^{\delta_k}} \tag{3.15}$$

$$\Lambda_l^{2)} = \frac{\det\left(\mathbf{Q}_{1,l}\right)}{\prod_{k=1}^{K}\prod_{m=1}^{\kappa_k}\bar{q}_{k,mm}^{\delta_k}}. \tag{3.16}$$

In order to derive the stochastic representation for each factor $\Lambda_l^{i)}$, we expand the ratios (3.15) and (3.16) such that they can be split into two statistically independent terms (under the null hypothesis). Let us first consider $\Lambda_l^{1)}$, which we expand by $\prod_{k=n_{l-1}+1}^{n_l} \prod_{j=1}^{\delta_k} \det\left(\mathbf{Q}_{0,k}^{(j)}\right)$ to factorize it into the following terms

$$\Lambda_l^{1)} = \underbrace{\frac{\det\left(\mathbf{Q}_{1,l}\right)}{\prod_{k=n_{l-1}+1}^{n_l} \prod_{j=1}^{\delta_k} \det\left(\mathbf{Q}_{0,k}^{(j)}\right)}}_{\Lambda_{l,\mathrm{I}}^{1)}} \times \underbrace{\frac{\prod_{k=n_{l-1}+1}^{n_l} \prod_{j=1}^{\delta_k} \det\left(\mathbf{Q}_{0,k}^{(j)}\right)}{\prod_{k=n_{l-1}+1}^{n_l} \det\left(\bar{\mathbf{Q}}_{0,k}\right)^{\delta_k}}}_{\Lambda_{l,\mathrm{II}}^{1)}}. \tag{3.17}$$

Hence, $\Lambda_l^{1)}$ is split into one term that is the block-independence test and second term that is the block-sphericity test, which are independent given that the null hypothesis is true. The authors in [100] show in Appendix H.3 that $\Lambda_{l,\mathrm{I}}^{1)}$ and $\Lambda_{l,\mathrm{II}}^{1)}$ are statistically independent by exploiting Basu's theorem [102] similarly as it has been done in Appendix A of [103].

Similar to the preceding paragraph we can split $\Lambda_l^{2)}$ into two statistically independent terms under the null hypothesis. To this end, we should first recall that for Case 2) $\mathbf{R}_0^{(k)}$ is a diagonal matrix with positive elements. Hence, we rewrite the denominator in (3.16) as $\prod_{k=n_{l-1}+1}^{n_l} \prod_{j=1}^{\delta_k} \prod_{m=1}^{\kappa_k} q_{k,mm}^{(j)}$ and factorize it as follows

$$\Lambda_l^{2)} = \underbrace{\frac{\det\left(\mathbf{Q}_{1,l}\right)}{\prod_{k=n_{l-1}+1}^{n_l} \prod_{j=1}^{\delta_k} \prod_{m=1}^{\kappa_k} q_{k,mm}^{(j)}}}_{\Lambda_{l,\mathrm{I}}^{2)}} \times \underbrace{\frac{\prod_{k=n_{l-1}+1}^{n_l} \prod_{j=1}^{\delta_k} \prod_{m=1}^{\kappa_k} q_{k,mm}^{(j)}}{\prod_{k=n_{l-1}+1}^{n_l} \prod_{m=1}^{\kappa_k} \bar{q}_{k,mm}^{\delta_k}}}_{\Lambda_{l,\mathrm{II}}^{2)}}. \tag{3.18}$$

Again the results of Appendix H.3 in [100] can be applied to conclude that the terms $\Lambda_{l,\mathrm{I}}^{2)}$ and $\Lambda_{l,\mathrm{II}}^{2)}$ are statistically independent.

Let us briefly outline the next steps. First, we will introduce the concept of Gram determinants as they are exploited throughout this chapter. Afterwards, we propose the distributions of the two independent terms $\Lambda_{l,\mathrm{I}}^{1)}$ and $\Lambda_{l,\mathrm{II}}^{1)}$, which will be combined in the distribution of $\Lambda_l^{1)}$ eventually. Finally, we will also derive the distributions of $\Lambda_{l,\mathrm{I}}^{2)}$ and $\Lambda_{l,\mathrm{II}}^{2)}$ and combine them to obtain the distribution of $\Lambda_l^{2)}$.

Distributions for Case 1)

Before we derive the distributions of $\Lambda_{l,\mathrm{I}}^{1)}$ and $\Lambda_{l,\mathrm{II}}^{1)}$, we first introduce the concept of Gram determinants [76]. Exemplarily, we consider the numerator $\det\left(\mathbf{Q}_{1,l}\right)$ of $\Lambda_l^{i)}$. The matrix $\mathbf{Q}_{1,l} = 1/M \mathbf{Y}_l \left(\mathbf{Y}_l\right)^H$ for $l = 1, \ldots, L$ is a Gram matrix, which contains the inner products between all combinations of rows of $\mathbf{Y}_l$. In [76] the author presents a procedure to sequentially test each row

in $\mathbf{Y}_l$ to check whether it is independent to its preceding vectors. Here we will use this concept to rewrite the determinants of Gram matrices as a product of scalars similar to the authors in [74], which is presented in the following paragraph, where we apply the notation from [76].

Firstly, we will rewrite the data matrix $\mathbf{Y}_l$ in terms of $\mathbf{X}_k$. By noting that $\lambda_l = \sum_{k=n_{l-1}+1}^{n_l} \delta_k \kappa_k$ it follows

$$\mathbf{Y}_l = \begin{bmatrix} \mathbf{X}_{n_{l-1}+1} \\ \vdots \\ \mathbf{X}_{n_l} \end{bmatrix}. \tag{3.19}$$

Secondly, we partition the Gram matrix $\mathbf{Q}_{1,l}$ into submatrices and we denote the $i \times i$ north-west block of $\mathbf{Q}_{1,l}$ by $\mathbf{Q}_{1,l}^{(i)}$. Moreover, we can partition the matrix $\mathbf{Q}_{1,l}^{(i)}$ as follows

$$\mathbf{Q}_{1,l}^{(i)} = \begin{bmatrix} \mathbf{Q}_{1,l}^{(i-1)} & \left(\mathbf{q}_{1,l}^{(i)}\right)^H \\ \mathbf{q}_{1,l}^{(i)} & q_{ii}^{(l)} \end{bmatrix}, \tag{3.20}$$

where $\mathbf{q}_{1,l}^{(i)} \in \mathbb{C}^{1\times(i-1)}$ and $q_{ii}^{(l)} \in \mathbb{R}$ is a scalar. Given this notation, the Gram determinant can be written as a product of scalar values as follows

$$\det\left(\mathbf{Q}_{1,l}\right) = \prod_{i=1}^{\lambda_l} \left\{ q_{ii}^{(l)} - \mathbf{q}_{1,l}^{(i)} \left(\mathbf{Q}_{1,l}^{(i-1)}\right)^{-1} \left(\mathbf{q}_{1,l}^{(i)}\right)^H \right\}, \tag{3.21}$$

where the terms $q_{ii}^{(l)} - \mathbf{q}_{1,l}^{(i)} \left(\mathbf{Q}_{1,l}^{(i-1)}\right)^{-1} \left(\mathbf{q}_{1,l}^{(i)}\right)^H$ are referred to as *Schur complements* [76]. Products of this style will accompany us throughout the derivations. Given this as a basis, we propose the distributions of Λ_l separately for Cases 1) and 2) in the upcoming paragraphs, which we will denote by $\Lambda_l^{1)}$ and $\Lambda_l^{2)}$, respectively.

In the following proposition we present the distribution of the first factor that is $\Lambda_{l,\mathrm{I}}^{1)}$.

Proposition 3.4.1. *The distribution of* $\Lambda_{l,\mathrm{I}}^{1)}$ *in* (3.17) *is given by a product of independent Beta random variables that is*

$$\Lambda_{l,\mathrm{I}}^{1)} \overset{\mathcal{D}}{=} \prod_{k=n_{l-1}+1}^{n_l} \prod_{j=1}^{\delta_k} \prod_{m=1}^{\kappa_k} V_{k,m,j}, \tag{3.22}$$

where

$$V_{k,m,j} \sim Beta\left\{ M - \left(\sum_{p=n_{l-1}+1}^{k-1} \delta_p \kappa_p + (j-1)\,\kappa_k + m - 1 \right), \sum_{p=n_{l-1}+1}^{k-1} \delta_p \kappa_p + (j-1)\,\kappa_k \right\}. \tag{3.23}$$

Proof. Let us begin the proof by replacing the determinants in the numerator and denominator of

$$\Lambda_{l,\mathrm{I}}^{1)} = \frac{\det\left(\mathbf{Q}_{1,l}\right)}{\prod_{k=n_{l-1}+1}^{n_l} \prod_{j=1}^{\delta_k} \det\left(\mathbf{Q}_{0,k}^{(j)}\right)} \tag{3.24}$$

by their corresponding product of Schur complements from (3.21). To this end, we first rewrite the numerator as

$$\det\left(\mathbf{Q}_{1,l}\right) = \det\left(\frac{1}{M}\mathbf{Y}_l\mathbf{Y}_l^H\right) = \prod_{i=1}^{\lambda_l}\left\{q_{ii}^{(l)} - \mathbf{q}_{1,l}^{(i)}\left(\mathbf{Q}_{1,l}^{(i-1)}\right)^{-1}\left(\mathbf{q}_{1,l}^{(i)}\right)^H\right\}. \tag{3.25}$$

Recall that this product is a sequential test for linear independence of the vectors building the Gram matrix to its preceding ones and that $\mathbf{Y}_l$ can be expressed in terms of $\mathbf{X}_k$ with $\lambda_l = \sum_{k=n_{l-1}+1}^{n_l}\delta_k\kappa_k$ by

$$\mathbf{Y}_l = \begin{bmatrix}\mathbf{X}_{n_{l-1}+1}\\ \vdots \\ \mathbf{X}_{n_l}\end{bmatrix}. \tag{3.26}$$

Hence, to get a better insight into this expression and its distribution, we will replace the components of the Gram matrices by the corresponding elements of the data matrix as follows

$$\begin{aligned}&\prod_{i=1}^{\lambda_l}\left\{q_{ii}^{(l)} - \mathbf{q}_{1,l}^{(i)}\left(\mathbf{Q}_{1,l}^{(i-1)}\right)^{-1}\left(\mathbf{q}_{1,l}^{(i)}\right)^H\right\} = \\ &\frac{1}{M^{\lambda_l}}\prod_{k=n_{l-1}+1}^{n_l}\prod_{j=1}^{\delta_k}\prod_{m=1}^{\kappa_k}\left\{\mathbf{x}_{k,m}^{(j)}\left(\mathbf{x}_{k,m}^{(j)}\right)^H - \mathbf{x}_{k,m}^{(j)}\left(\mathbf{W}_{k,m}^{(j)}\right)^H\left[\mathbf{W}_{k,m}^{(j)}\left(\mathbf{W}_{k,m}^{(j)}\right)^H\right]^{-1}\mathbf{W}_{k,m}^{(j)}\left(\mathbf{x}_{k,m}^{(j)}\right)^H\right\},\end{aligned} \tag{3.27}$$

where $\mathbf{x}_{k,m}^{(j)} \in \mathbb{C}^{1\times M}$ is the mth element in $\mathbf{X}_k^{(j)}$ from (3.9) and $\mathbf{W}_{k,m}^{(j)} \in \mathbb{C}^{\left[\sum_{p=n_{l-1}+1}^{k-1}\delta_p\kappa_p + (j-1)\kappa_k + m - 1\right]\times M}$ is the matrix that contains all vectors preceding $\mathbf{x}_{k,m}^{(j)}$ in (3.26)

$$\mathbf{W}_{k,m}^{(j)} = \begin{bmatrix}\mathbf{X}_{n_{l-1}+1}\\ \vdots \\ \mathbf{X}_{k-1}\\ \mathbf{X}_k^{(1)}\\ \vdots \\ \mathbf{X}_k^{(j-1)}\end{bmatrix}\Bigg\}\,\mathbf{U}_k^{(j-1)} \\ \begin{bmatrix}\mathbf{x}_{k,1}^{(j)}\\ \vdots \\ \mathbf{x}_{k,m-1}^{(j)}\end{bmatrix}\Bigg\}\,\mathbf{V}_{k,m}^{(j)}. \tag{3.28}$$

Note, that the matrices $\mathbf{U}_k^{(j-1)}$ and $\mathbf{V}_{k,m}^{(j)}$ will be used subsequently to handle the denominator of (3.25).

Now we can factor out $\mathbf{x}_{k,m}^{(j)}$ to obtain

$$\prod_{k=n_{l-1}+1}^{n_l} \prod_{j=1}^{\delta_k} \prod_{m=1}^{\kappa_k} \mathbf{x}_{k,m}^{(j)} \left(\mathbf{x}_{k,m}^{(j)}\right)^H - \mathbf{x}_{k,m}^{(j)} \left(\mathbf{W}_{k,m}^{(j)}\right)^H \left[\mathbf{W}_{k,m}^{(j)} \left(\mathbf{W}_{k,m}^{(j)}\right)^H\right]^{-1} \mathbf{W}_{k,m}^{(j)} \left(\mathbf{x}_{k,m}^{(j)}\right)^H = \prod_{k=n_{l-1}+1}^{n_l} \prod_{j=1}^{\delta_k} \prod_{m=1}^{\kappa_k} \mathbf{x}_{k,m}^{(j)} \underbrace{\left[\mathbf{I}_M - \left(\mathbf{W}_{k,m}^{(j)}\right)^H \left[\mathbf{W}_{k,m}^{(j)} \left(\mathbf{W}_{k,m}^{(j)}\right)^H\right]^{-1} \mathbf{W}_{k,m}^{(j)}\right]}_{\mathbf{P}^{\perp}_{\mathbf{W}_{k,m}^{(j)}}} \left(\mathbf{x}_{k,m}^{(j)}\right)^H, \quad (3.29)$$

where $\mathbf{P}^{\perp}_{\mathbf{W}_{k,m}^{(j)}}$ denotes the projection onto the subspace orthogonal to $\mathbf{W}_{k,m}^{(j)}$. This expression allows us to easily determine its distribution.

In [104] it is shown that a quadratic form of two Gaussians is chi-squared distributed. This result was generalized in [105, 106] to proper complex Gaussians and will be exploited throughout the remainder of this proof. Moreover, under the null hypothesis we can pre-whiten the covariance matrices since the test statistic is invariant to multiplications with matrices of the structure of $\mathbf{R}_0$. Hence, without loss of generality it holds that $\mathbf{x} \sim \mathcal{CN}(0, \mathbf{I}_N)$. Furthermore, there are M i.i.d. samples of $\mathbf{x}$ available, for this reason, we can equivalently state that $\left(\mathbf{x}_{k,m}^{(j)}\right)^H \sim \mathcal{CN}(0, \mathbf{I}_M)$. According to [104–106], the quadratic term $2\mathbf{x}_{k,m}^{(j)} \mathbf{P}^{\perp}_{\mathbf{W}_{k,m}^{(j)}} \left(\mathbf{x}_{k,m}^{(j)}\right)^H$ is chi-squared distributed with degrees of freedom equal to $2 \operatorname{rank}\left(\mathbf{P}^{\perp}_{\mathbf{W}_{k,m}^{(j)}}\right)$, i.e.,

$$2\mathbf{x}_{k,m}^{(j)} \mathbf{P}^{\perp}_{\mathbf{W}_{k,m}^{(j)}} \left(\mathbf{x}_{k,m}^{(j)}\right)^H \sim \chi^2\left(2\left[M - \left(\sum_{p=n_{l-1}+1}^{k-1} \delta_p \kappa_p + (j-1)\kappa_k + m - 1\right)\right]\right). \quad (3.30)$$

Similarly, we can decompose the denominator of $\Lambda_{l,\mathrm{I}}^{1)}$ in (3.24) as follows

$$\prod_{k=n_{l-1}+1}^{n_l} \prod_{j=1}^{\delta_k} \frac{1}{M^{\kappa_k}} \det\left(\mathbf{Q}_{0,k}^{(j)}\right) = \frac{1}{M^{\lambda_l}} \prod_{k=n_{l-1}+1}^{n_l} \prod_{j=1}^{\delta_k} \prod_{m=1}^{\kappa_k} \mathbf{x}_{k,m}^{(j)} \underbrace{\left[\mathbf{I}_M - \left(\mathbf{V}_{k,m}^{(j)}\right)^H \left[\mathbf{V}_{k,m}^{(j)} \left(\mathbf{V}_{k,m}^{(j)}\right)^H\right]^{-1} \mathbf{V}_{k,m}^{(j)}\right]}_{\mathbf{P}^{\perp}_{\mathbf{V}_{k,m}^{(j)}}} \left(\mathbf{x}_{k,m}^{(j)}\right)^H, \quad (3.31)$$

where we recall that $\mathbf{Q}_{0,k}^{(j)} = 1/M \mathbf{X}_k^{(j)} \left(\mathbf{X}_k^{(j)}\right)^H$ and $\mathbf{V}_{k,m}^{(j)}$ is given in (3.28) and contains all rows of $\mathbf{X}_k^{(j)}$ preceding $\mathbf{x}_{k,m}^{(j)}$.

Now we can rewrite $\Lambda_{l,\mathrm{I}}^{1)}$ in (3.17) as a function of $\mathbf{x}_{k,m}^{(j)}$

$$\Lambda_{l,\mathrm{I}}^{1)} = \prod_{k=n_{l-1}+1}^{n_l} \prod_{j=1}^{\delta_k} \prod_{m=1}^{\kappa_k} \frac{\mathbf{x}_{k,m}^{(j)} \mathbf{P}^{\perp}_{\mathbf{W}_{k,m}^{(j)}} \left(\mathbf{x}_{k,m}^{(j)}\right)^H}{\mathbf{x}_{k,m}^{(j)} \mathbf{P}^{\perp}_{\mathbf{V}_{k,m}^{(j)}} \left(\mathbf{x}_{k,m}^{(j)}\right)^H}. \tag{3.32}$$

It should be noted that the ratios are independent for all k, j, and m due to the nature of Gram determinants that sequentially test for orthogonality. Similar to the distribution of the numerator in (3.30), the distribution of $2\mathbf{x}_{k,m}^{(j)} \mathbf{P}^{\perp}_{\mathbf{V}_{k,m}^{(j)}} \left(\mathbf{x}_{k,m}^{(j)}\right)^H$ is given by a chi-squared distribution as well. However, to determine the distribution of the fraction in (3.32), we will split $\mathbf{P}^{\perp}_{\mathbf{V}_{k,m}^{(j)}}$ into two orthogonal projections [74]

$$\mathbf{P}^{\perp}_{\mathbf{V}_{k,m}^{(j)}} = \mathbf{P}^{\perp}_{\mathbf{W}_{k,m}^{(j)}} + \mathbf{P}_{\mathbf{P}^{\perp}_{\mathbf{V}_{k,m}^{(j)}} \mathbf{U}_k^{(j-1)}}, \tag{3.33}$$

where $\mathbf{U}_k^{(j-1)}$ is given by (3.28) and $P_{\mathbf{P}^{\perp}_{\mathbf{V}_{k,m}^{(j)}} \mathbf{U}_k^{(j-1)}}$ is the projection onto the subspace $\mathbf{P}^{\perp}_{\mathbf{V}_{k,m}^{(j)}} \mathbf{U}_k^{(j-1)}$, which itself is the projection of $\mathbf{U}_k^{(j-1)}$ onto the orthogonal subspace of $\mathbf{V}_{k,m}^{(j)}$. Note that $\mathbf{P}_{\mathbf{P}^{\perp}_{\mathbf{V}_{k,m}^{(j)}} \mathbf{U}_k^{(j-1)}}$ is orthogonal to the orthogonal subspace of $\mathbf{W}_{k,m}^{(j)}$ since $\mathbf{W}_{k,m}^{(j)}$ is the stack of $\mathbf{U}_k^{(j-1)}$ and $\mathbf{V}_{k,m}^{(j)}$. Furthermore, we should note that the sum in the denominator is equal to the numerator plus $\mathbf{x}_{k,m}^{(j)} \mathbf{P}_{\mathbf{P}^{\perp}_{\mathbf{V}_{k,m}^{(j)}} \mathbf{U}_k^{(j-1)}} \left(\mathbf{x}_{k,m}^{(j)}\right)^H$. The latter term is distributed as

$$2\mathbf{x}_{k,m}^{(j)} \mathbf{P}_{\mathbf{P}^{\perp}_{\mathbf{V}_{k,m}^{(j)}} \mathbf{U}_k^{(j-1)}} \left(\mathbf{x}_{k,m}^{(j)}\right)^H \sim \chi^2\left(2\left[\sum_{p=n_{l-1}+1}^{k-1} \delta_p \kappa_p + (j-1)\,\kappa_k\right]\right). \tag{3.34}$$

Hence, the fraction is of the form $\frac{a}{a+b}$, where $a \sim \chi^2(d_a)$ and $b \sim \chi^2(d_b)$ are independent chi-squared distributions with degrees of freedom d_a and d_b, respectively. In, e.g., [107] it is shown that the ratio is distributed as

$$\frac{a}{a+b} \sim \mathrm{Beta}\left\{1/2d_a, 1/2d_b\right\}. \tag{3.35}$$

This concludes the proof.

□

Having derived the distribution of $\Lambda_{l,\mathrm{I}}^{1)}$ in (3.17), we will now provide the distribution of the second term $\Lambda_{l,\mathrm{II}}^{1)}$. Note that $\Lambda_{l,\mathrm{II}}^{1)}$ is the statistic to test for the equivalence of covariance matrices, i.e., the block-sphericity test [90].

Proposition 3.4.2. *The distribution of* $\Lambda_{l,\mathrm{II}}^{1)}$ *in* (3.17) *is given by a product of independent Beta random variables, specifically,*

$$\Lambda_{l,\mathrm{II}}^{1)} \overset{\mathcal{D}}{=} \prod_{k=n_{l-1}+1}^{n_l} \delta_k^{\delta_k \kappa_k} \prod_{j=2}^{\delta_k} \left\{ \prod_{m=1}^{\kappa_k} \left(1 - W_{m,j-1}\right) \left(W_{m,j-1}\right)^{j-1} \right\} \left\{ \prod_{m=2}^{\kappa_k} \left(U_{m,j}\right)^{j} \right\}, \tag{3.36}$$

where

$$W_{m,j-1} \sim \mathit{Beta}\left\{M(j-1) - (m-1), M - (m-1)\right\}, \tag{3.37}$$

$$U_{m,j} \sim \mathit{Beta}\left\{Mj - 2(m-1), m-1\right\}, \tag{3.38}$$

are independent for all k*,* m*, and* j*.*

Proof. In order to proof this proposition, we basically follow Chapter 10.4 in [90], where the author derived the GLRT for the matrix sphericity test and showed that its stochastic representation under the null is given by a product of independent beta random variables. Here we follow those lines in order to provide the complete picture for our more general test. From (3.17) $\Lambda_{l,\mathrm{II}}^{1)}$ is given by

$$\Lambda_{l,\mathrm{II}}^{1)} = \prod_{k=n_{l-1}+1}^{n_l} \frac{\prod_{j=1}^{\delta_k} \det\left(\mathbf{Q}_{0,k}^{(j)}\right)}{\det\left(\bar{\mathbf{Q}}_{0,k}\right)^{\delta_k}} = \prod_{k=n_{l-1}+1}^{n_l} \delta_k^{\delta_k \kappa_k} \frac{\prod_{j=1}^{\delta_k} \det\left(\mathbf{Q}_{0,k}^{(j)}\right)}{\det\left(\mathbf{Q}_{0,k}^{(1)} + \cdots + \mathbf{Q}_{0,k}^{(\delta_k)}\right)^{\delta_k}}. \tag{3.39}$$

Note that $\Lambda_{l,\mathrm{II}}^{1)} = 1$ if $\delta_k = 1$, i.e., the case, where the blocks $\mathbf{Q}_{0,k}^{(j)}$ are not repeated on the main diagonal. We expand the fraction as follows

$$\Lambda_{l,\mathrm{II}}^{1)} = \prod_{k=n_{l-1}+1}^{n_l} \delta_k^{\delta_k \kappa_k} \frac{\det\left(\mathbf{Q}_{0,k}^{(1)}\right)}{\det\left(\mathbf{Q}_{0,k}^{(1)}\right)} \times \frac{\det\left(\mathbf{Q}_{0,k}^{(1)}\right) \det\left(\mathbf{Q}_{0,k}^{(2)}\right)}{\det\left(\mathbf{Q}_{0,k}^{(1)} + \mathbf{Q}_{0,k}^{(2)}\right)^2} \times \cdots \tag{3.40}$$

$$\times \frac{\det\left(\mathbf{Q}_{0,k}^{(1)} + \cdots + \mathbf{Q}_{0,k}^{(\delta_k - 1)}\right)^{\delta_k - 1} \det\left(\mathbf{Q}_{0,k}^{(\delta_k)}\right)}{\det\left(\mathbf{Q}_{0,k}^{(1)} + \cdots + \mathbf{Q}_{0,k}^{(\delta_k)}\right)^{\delta_k}} \tag{3.41}$$

$$= \prod_{k=n_{l-1}+1}^{n_l} \delta_k^{\delta_k \kappa_k} \prod_{j=2}^{\delta_k} \frac{\det\left(\mathbf{Q}_{0,k}^{(1)} + \cdots + \mathbf{Q}_{0,k}^{(j-1)}\right)^{j-1} \det\left(\mathbf{Q}_{0,k}^{(j)}\right)}{\det\left(\mathbf{Q}_{0,k}^{(1)} + \cdots + \mathbf{Q}_{0,k}^{(j)}\right)^{j}}. \tag{3.42}$$

Note that these ratios are independent for all $j = 2, \ldots, \delta_k$, which was shown in Theorem 10.4.1. in [90]. Similar to the proof of Proposition 3.4.1, we make use of Gram determinants. Furthermore, to keep readability, we denote the sum $\mathbf{Q}_{0,k}^{(1)} + \cdots + \mathbf{Q}_{0,k}^{(j-1)}$ as follows

$$\mathbf{A}_k^{(j-1)} = \mathbf{Q}_{0,k}^{(1)} + \cdots + \mathbf{Q}_{0,k}^{(j-1)} \tag{3.43}$$

and rewrite (3.39) as

$$\Lambda_{l,\mathrm{II}}^{1)} = \prod_{k=n_{l-1}+1}^{n_l} \delta_k^{\delta_k \kappa_k} \prod_{j=2}^{\delta_k} \frac{\det\left(\mathbf{A}_k^{(j-1)}\right)^{j-1} \det\left(\mathbf{Q}_{0,k}^{(j)}\right)}{\det\left(\mathbf{A}_k^{(j-1)} + \mathbf{Q}_{0,k}^{(j)}\right)^{j}}. \tag{3.44}$$

Now we replace the determinants by products of Schur complements as we have seen in (3.21)

$$\Lambda_{l,\mathrm{II}}^{1)} = \prod_{k=n_{l-1}+1}^{n_l} \delta_k^{\delta_k \kappa_k} \prod_{j=2}^{\delta_k} \prod_{m=2}^{\kappa_k} \Delta_0^{(k,j,m)} \tag{3.45}$$

where

$$\Delta_0^{(k,j,m)} = \frac{\left(\alpha_{k,mm}^{(j-1)}\right)^{j-1} \beta_{k\cdot mm}^{(j)}}{\left(a_{k\cdot mm}^{(j-1)} + q_{k,mm}^{(j)} - \left(\mathbf{a}_{k,m}^{(j-1)} + \mathbf{q}_{k,m}^{(j)}\right)\left(\mathbf{A}_{k,m}^{(j-1)} + \mathbf{Q}_{k,m}^{(j)}\right)^{-1}\left(\mathbf{a}_{k,m}^{(j-1)} + \mathbf{q}_{k,m}^{(j)}\right)^{H}\right)^{j}} \tag{3.46}$$

with

$$\alpha_{k,mm}^{(j-1)} = a_{k\cdot mm}^{(j-1)} - \mathbf{a}_{k,m}^{(j-1)}\left(\mathbf{A}_{k,m-1}^{(j-1)}\right)^{-1}\left(\mathbf{a}_{k,m}^{(j-1)}\right)^{H}, \tag{3.47}$$

$$\beta_{k\cdot mm}^{(j)} = q_{k,mm}^{(j)} - \mathbf{q}_{k,m}^{(j)}\left(\mathbf{Q}_{k,m-1}^{(j)}\right)^{-1}\left(\mathbf{q}_{k,m}^{(j)}\right)^{H}, \tag{3.48}$$

and $a_{k,mm}^{(j-1)}, \mathbf{a}_{k,m}^{(j-1)}, \mathbf{A}_{k\cdot m-1}^{(j-1)}, q_{k,mm}^{(j)}, \mathbf{q}_{k,m}^{(j)}$, and $\mathbf{Q}_{k,m-1}^{(j)}$ coming from partitioning $\mathbf{A}_k^{(j-1)}$ and $\mathbf{Q}_{0,k}^{(j)}$ equivalently to $\mathbf{Q}_i^{(l)}$ in (3.20). Note that for $m = 1$ it holds that $\Delta_0^{(k,j,m)} = 1$, i.e., for the case that $\kappa_k = 1$, we can neglect the term $\Lambda_{l,\mathrm{II}}^{1)}$ in (3.17) as it is 1.

Now we expand $\Delta_0^{(k,j,m)}$ by $\left(\alpha_{k,mm}^{(j-1)} + \beta_{k\cdot mm}^{(j)}\right)^{j}$ in order to divide it into two terms $\Delta_1^{(k,j,m)}$ and $\Delta_2^{(k,j,m)}$ as follows

$$\Delta_0^{(k,j,m)} = \Delta_1^{(k,j,m)} \times \Delta_2^{(k,j,m)}, \tag{3.49}$$

where

$$\Delta_1^{(k,j,m)} = \frac{\left(\alpha_{k,mm}^{(j-1)}\right)^{j-1} \beta_{k\cdot mm}^{(j)}}{\left(\alpha_{k,mm}^{(j-1)} + \beta_{k\cdot mm}^{(j)}\right)^{j}}, \tag{3.50}$$

$$\Delta_2^{(k,j,m)} = \left(\frac{\alpha_{k,mm}^{(j-1)} + \beta_{k\cdot mm}^{(j)}}{a_{k\cdot mm}^{(j-1)} + q_{k,mm}^{(j)} - \left(\mathbf{a}_{k,m}^{(j-1)} + \mathbf{q}_{k,m}^{(j)}\right)\left(\mathbf{A}_{k,m}^{(j-1)} + \mathbf{Q}_{k,m}^{(j)}\right)^{-1}\left(\mathbf{a}_{k,m}^{(j-1)} + \mathbf{q}_{k,m}^{(j)}\right)^{H}}\right)^{j}. \tag{3.51}$$

For each of the terms $\Delta_1^{(k,j,m)}$ and $\Delta_2^{(k,j,m)}$ we will determine their distribution separately.

Let us first focus on $\Delta_1^{(k,j,m)}$, which we factorize as

$$\Delta_1^{(k,j,m)} = \left(\frac{\alpha_{k,mm}^{(j-1)}}{\alpha_{k,mm}^{(j-1)} + \beta_{k\cdot mm}^{(j)}} \right)^{j-1} \times \frac{\beta_{k\cdot mm}^{(j)}}{\alpha_{k,mm}^{(j-1)} + \beta_{k\cdot mm}^{(j)}} \tag{3.52}$$

$$= \left(\frac{\alpha_{k,mm}^{(j-1)}}{\alpha_{k,mm}^{(j-1)} + \beta_{k\cdot mm}^{(j)}} \right)^{j-1} \left(1 - \frac{\alpha_{k,mm}^{(j-1)}}{\alpha_{k,mm}^{(j-1)} + \beta_{k\cdot mm}^{(j)}} \right). \tag{3.53}$$

In order to find the distribution of $\dfrac{\alpha_{k,mm}^{(j-1)}}{\alpha_{k,mm}^{(j-1)} + \beta_{k\cdot mm}^{(j)}}$, we identify the distributions of $\alpha_{k,mm}^{(j-1)}$ and $\beta_{k\cdot mm}^{(j)}$. First, $\alpha_{k,mm}^{(j-1)}$ is given by the Schur complement (3.47). Since $\mathbf{A}_k^{(j-1)}$ is the sum of $j-1$ Gram matrices each having a complex Wishart distribution with M degrees of freedom, $\mathbf{A}_k^{(j-1)}$ itself is also complex Wishart distributed with $M(j-1)$ degrees of freedom. Theorem 1 of Chapter 4 in [108], where the authors derived the distribution of $\alpha_{k,mm}^{(j-1)}$ when $\mathbf{A}_k^{(j-1)}$ is a Wishart matrix, can be applied. The extension of that theorem to the proper complex case is straightforward. Hence, the distribution is given by

$$2\alpha_{k,mm}^{(j-1)} \sim \chi^2 \left(2 \left[M(j-1) - (m-1) \right] \right). \tag{3.54}$$

Equivalently, we can derive the distribution of $\beta_{k\cdot mm}^{(j)}$, which is given by

$$2\beta_{k\cdot mm}^{(j)} \sim \chi^2 \left(2 \left[M - (m-1) \right] \right). \tag{3.55}$$

Finally, it is easy to see that the terms $\alpha_{k,mm}^{(j-1)}$ and $\beta_{k\cdot mm}^{(j)}$ are independent. Hence, similar to the proof of Proposition 3.4.1, we can find the following

$$\Delta_1^{(k,j,m)} \overset{\mathcal{D}}{=} \left(W_{m,j-1} \right)^{j-1} \left(1 - W_{m,j-1} \right) \tag{3.56}$$

with

$$W_{m,j-1} \sim \text{Beta} \left\{ M(j-1) - (m-1), M - (m-1) \right\}. \tag{3.57}$$

Now we will derive the distribution of $\Delta_2^{(k,j,m)}$ in (3.49)

$$\Delta_2^{(k,j,m)} = \left(\frac{\alpha_{k,mm}^{(j-1)} + \beta_{k\cdot mm}^{(j)}}{a_{k\cdot mm}^{(j-1)} + q_{k,mm}^{(j)} - \left(\mathbf{a}_{k,m}^{(j-1)} + \mathbf{q}_{k,m}^{(j)} \right) \left(\mathbf{A}_{k,m}^{(j-1)} + \mathbf{Q}_{k,m-1}^{(j)} \right)^{-1} \left(\mathbf{a}_{k,m}^{(j-1)} + \mathbf{q}_{k,m}^{(j)} \right)^H} \right)^j. \tag{3.58}$$

To this end, we will continue following the lines from [90]. Firstly, we will expand the denominator such that the term in the numerator $\alpha_{k,mm}^{(j-1)} + \beta_{k\cdot mm}^{(j)}$ also occurs in the denominator. This aims in getting another fraction of the form (3.35). For now we rewrite the denominator of $\Delta_2^{(k,j,m)}$ as follows

$$a_{k\cdot mm}^{(j-1)} + q_{k,mm}^{(j)} - \left(\mathbf{a}_{k,m}^{(j-1)} + \mathbf{q}_{k,m}^{(j)}\right)\left(\mathbf{A}_{k,m}^{(j-1)} + \mathbf{Q}_{k,m}^{(j)}\right)^{-1}\left(\mathbf{a}_{k,m}^{(j-1)} + \mathbf{q}_{k,m}^{(j)}\right)^{H} = \alpha_{k,mm}^{(j-1)} + \beta_{k,m}^{(j)} + \delta^{(k,j,m)}, \tag{3.59}$$

where we expanded the term with $\mathbf{a}_k^{(j-1)}\left(\mathbf{A}_k^{(j-1)}\right)^{-1}\left(\mathbf{a}_k^{(j-1)}\right)^{H}$ and $\mathbf{q}_{0,k}^{(j)}\left(\mathbf{Q}_{0,k}^{(j)}\right)^{-1}\left(\mathbf{q}_{0,k}^{(j)}\right)^{H}$ such that

$$\begin{aligned}\delta^{(k,j,m)} = {} & \mathbf{a}_k^{(j-1)}\left(\mathbf{A}_k^{(j-1)}\right)^{-1}\left(\mathbf{a}_k^{(j-1)}\right)^{H} + \mathbf{q}_{0,k}^{(j)}\left(\mathbf{Q}_{0,k}^{(j)}\right)^{-1}\left(\mathbf{q}_{0,k}^{(j)}\right)^{H} \\ & - \left(\mathbf{a}_{k,m}^{(j-1)} + \mathbf{q}_{k,m}^{(j)}\right)\left(\mathbf{A}_{k,m}^{(j-1)} + \mathbf{Q}_{k,m}^{(j)}\right)^{-1}\left(\mathbf{a}_{k,m}^{(j-1)} + \mathbf{q}_{k,m}^{(j)}\right)^{H}. \end{aligned} \tag{3.60}$$

In (3.54) and (3.55) we figured that the distributions of $\alpha_{k,mm}^{(j-1)}$ and $\beta_{k,m}^{(j)}$ are both chi-squared and so is the sum $\alpha_{k,mm}^{(j-1)} + \beta_{k,m}^{(j)}$ with degrees of freedom equal to the sum of the single ones, i.e.,

$$2\left(\alpha_{k,mm}^{(j-1)} + \beta_{k,m}^{(j)}\right) \sim \chi^2\left(2\left[Mj - 2(m-1)\right]\right). \tag{3.61}$$

However, the distribution of $\delta^{(k,j,m)}$ is not easy to identify at the first glance. For this reason, we use Lemma 10.4.1 in [90], which states

Lemma 3.4.1.

$$\begin{aligned}\delta^{(k,j,m)} = {} & \left[\mathbf{a}_k^{(j-1)}\left(\mathbf{A}_k^{(j-1)}\right)^{-1} - \mathbf{q}_{0,k}^{(j)}\left(\mathbf{Q}_{0,k}^{(j)}\right)^{-1}\right] \\ & \left[\left(\mathbf{A}_{k,m}^{(j-1)}\right)^{-1} + \left(\mathbf{Q}_{k,m}^{(j)}\right)^{-1}\right]^{-1}\left[\mathbf{a}_k^{(j-1)}\left(\mathbf{A}_k^{(j-1)}\right)^{-1} - \mathbf{q}_{0,k}^{(j)}\left(\mathbf{Q}_{0,k}^{(j)}\right)^{-1}\right]^{H}. \end{aligned} \tag{3.62}$$

Proof. Refer to Lemma 10.4.1 in [90]. □

This Lemma allows for easily coming up with the distribution of $\delta^{(k,j,m)}$. Specifically, we first consider the terms $\mathbf{a}_k^{(j-1)}\left(\mathbf{A}_k^{(j-1)}\right)^{-1}$ and $\mathbf{q}_{0,k}^{(j)}\left(\mathbf{Q}_{0,k}^{(j)}\right)^{-1}$. In Theorem 2 in [108] it is shown that these vectors are Gaussian distributed with covariance matrices $\left(\mathbf{A}_k^{(j-1)}\right)^{-1}$ and $\left(\mathbf{Q}_{0,k}^{(j)}\right)^{-1}$, respectively. Although, in [108] the author focused on the case of real-valued variables, it can be directly extended to the proper complex Gaussian case. Furthermore, the difference of both expressions is still zero-mean proper complex Gaussian with covariance matrix $\left(\mathbf{A}_k^{(j-1)}\right)^{-1} + \left(\mathbf{Q}_{0,k}^{(j)}\right)^{-1}$ considering that $\mathbf{a}_k^{(j-1)}\left(\mathbf{A}_k^{(j-1)}\right)^{-1}$ and $\mathbf{q}_{0,k}^{(j)}\left(\mathbf{Q}_{0,k}^{(j)}\right)^{-1}$ are independent. Hence, it turns out that $\delta^{(k,j,m)}$ is another quadratic term of two complex Gaussians scaled by the inverse of its covariance matrix. Now it is straightforward to obtain the distribution of $\delta^{(k,j,m)}$, which is given by another chi-squared distribution

$$2\delta^{(k,j,m)} \sim \chi^2\left(2\left[m-1\right]\right). \tag{3.63}$$

Noting that $\delta^{(k,j,m)}$ is independent of $\alpha_{k,mm}^{(j-1)} + \beta_{k,m}^{(j)}$, we apply once again (3.35) to find the distribution of $\Delta_2^{(k,j,m)}$, which is given by

$$\Delta_2^{(k,j,m)} \overset{\mathcal{D}}{=} \left(U_{m,j}\right)^j \tag{3.64}$$

with

$$U_{m,j} \sim \text{Beta}\left\{Mj - 2(m-1), m-1\right\}, \tag{3.65}$$

which concludes the proof.

□

Finally, everything is set up to obtain the stochastic representations of the GLR for Case 1) in the following theorem.

Theorem 3.4.1. *The GLR for Case 1) given in* (3.12) *has the following distribution under the null hypothesis*

$$\begin{aligned}\mathcal{G}^{1/M} \overset{\mathcal{D}}{=} &\prod_{l=1}^{L}\left\{\prod_{k=n_{l-1}+1}^{n_l}\prod_{j=1}^{\delta_k}\prod_{m=1}^{\kappa_k} V_{k,m,j}\right\} \\ &\times\left\{\prod_{k=n_{l-1}+1}^{n_l}\delta_k^{\delta_k\kappa_k}\prod_{j=2}^{\delta_k}\left\{\prod_{m=1}^{\kappa_k}\left(1-W_{m,j-1}\right)\left(W_{m,j-1}\right)^{j-1}\right\}\left\{\prod_{m=2}^{\kappa_k}\left(U_{m,j}\right)^j\right\}\right\},\end{aligned} \tag{3.66}$$

where all V, W, and U are independent Beta distributed random variables. Specifically, they are given by

$$V_{k,m,j} \sim Beta\left\{M - \left(\sum_{p=n_{l-1}+1}^{k-1}\delta_p\kappa_p + (j-1)\kappa_k + m - 1\right), \sum_{p=n_{l-1}+1}^{k-1}\delta_p\kappa_p + (j-1)\kappa_k\right\}, \tag{3.67}$$

$$W_{m,j-1} \sim Beta\left\{M(j-1) - (m-1), M - (m-1)\right\}, \tag{3.68}$$

$$U_{m,j} \sim Beta\left\{Mj - 2(m-1), m-1\right\}. \tag{3.69}$$

Proof. This theorem is a direct consequence of Propositions 3.4.1 and 3.4.2, the independence of $\Lambda_{l,\mathrm{I}}^{1)}$ and $\Lambda_{l,\mathrm{II}}^{1)}$, and also the independence of Λ_l for $l = 1, \ldots, L$ in (3.14). □

Distributions for Case 2)

Here we present the distributions of $\Lambda_{l,\mathrm{I}}^{2)}$ and $\Lambda_{l,\mathrm{II}}^{2)}$ in the following two propositions.

Proposition 3.4.3. *The distribution of* $\Lambda_{l,\mathrm{I}}^{2)}$ *in* (3.18) *is given by a product of independent Beta random variables that is*

$$\Lambda_{l,\mathrm{I}}^{2)} \overset{\mathcal{D}}{=} \prod_{k=n_{l-1}+1}^{n_l} \prod_{j=1}^{\delta_k} \prod_{m=1}^{\kappa_k} Y_{k,m,j}, \tag{3.70}$$

where

$$Y_{k,m,j} \sim Beta\left\{ M - \left(\sum_{p=n_{l-1}+1}^{k-1} \delta_p \kappa_p + (j-1)\,\kappa_k + m - 1 \right), \sum_{p=n_{l-1}+1}^{k-1} \delta_p \kappa_p + (j-1)\,\kappa_k + m - 1 \right\}. \tag{3.71}$$

Proof. The proof is similar to that of Proposition 3.4.1. The numerator is identical to that in $\Lambda_{l,\mathrm{I}}^{1)}$ and can be factored into

$$\det\left(\mathbf{Q}_{1,l}\right) = \prod_{k=n_{l-1}+1}^{n_l} \prod_{j=1}^{\delta_k} \prod_{m=1}^{\kappa_k} \mathbf{x}_{k,m}^{(j)} \mathbf{P}_{\mathbf{W}_{k,m}^{(j)}}^{\perp} \left(\mathbf{x}_{k,m}^{(j)}\right)^H \tag{3.72}$$

and its distribution is given by (3.30). The denominator of $\Lambda_{l,\mathrm{I}}^{2)}$ in (3.18) can easily be rewritten as a function of quadratic terms of $\mathbf{x}_{k,m}^{(j)}$ since $q_{k,mm}^{(j)} = \mathbf{x}_{k,m}^{(j)} \left(\mathbf{x}_{k,m}^{(j)}\right)^H$, as follows

$$\prod_{k=n_{l-1}+1}^{n_l} \prod_{j=1}^{\delta_k} \prod_{m=1}^{\kappa_k} q_{k,mm}^{(j)} = \prod_{k=n_{l-1}+1}^{n_l} \prod_{j=1}^{\delta_k} \prod_{m=1}^{\kappa_k} \mathbf{x}_{k,m}^{(j)} \left(\mathbf{x}_{k,m}^{(j)}\right)^H \tag{3.73}$$

which split into two orthogonal terms

$$\prod_{k=n_{l-1}+1}^{n_l} \prod_{j=1}^{\delta_k} \prod_{m=1}^{\kappa_k} \mathbf{x}_{k,m}^{(j)} \left(\mathbf{x}_{k,m}^{(j)}\right)^H = \prod_{k=n_{l-1}+1}^{n_l} \prod_{j=1}^{\delta_k} \prod_{m=1}^{\kappa_k} \mathbf{x}_{k,m}^{(j)} \left(\mathbf{P}_{\mathbf{W}_{k,m}^{(j)}}^{\perp} + \mathbf{P}_{\mathbf{W}_{k,m}^{(j)}} \right) \left(\mathbf{x}_{k,m}^{(j)}\right)^H. \tag{3.74}$$

The distribution of $\mathbf{x}_{k,m}^{(j)} \mathbf{P}_{\mathbf{W}_{k,m}^{(j)}} \left(\mathbf{x}_{k,m}^{(j)}\right)^H$ can be easily identified as

$$2\mathbf{x}_{k,m}^{(j)} \mathbf{P}_{\mathbf{W}_{k,m}^{(j)}} \left(\mathbf{x}_{k,m}^{(j)}\right)^H \sim \chi^2\left(2 \left[\sum_{p=n_{l-1}+1}^{k-1} \delta_p \kappa_p + (j-1)\,\kappa_k + m - 1 \right] \right). \tag{3.75}$$

Finally, we exploit again (3.35) and the proof follows.

□

To conclude the derivation of the stochastic representation of $\Lambda_l^{2)}$ for Case 2), we propose the distribution of $\Lambda_{l,\mathrm{II}}^{2)}$ in the following proposition.

Proposition 3.4.4. *The distribution of* $\Lambda_{l,\mathrm{II}}^{2)}$ *in* (3.17) *is given by a product of independent Beta random variables that is*

$$\Lambda_{l,\mathrm{II}}^{2)} \overset{\mathcal{D}}{=} \prod_{k=n_{l-1}+1}^{n_l} \delta_k^{\delta_k \kappa_k} \prod_{j=2}^{\delta_k} \left\{ \prod_{m=1}^{\kappa_k} \left(1 - Z_{m,j-1}\right) \left(Z_{m,j-1}\right)^{j-1} \right\}, \tag{3.76}$$

where

$$Z_{m,j-1} \sim Beta\left\{M(j-1), M\right\}, \tag{3.77}$$

are independent for all k*,* m*, and* j*.*

Proof. First, we rewrite $\Lambda_{l,\mathrm{II}}^{2)}$ similarly to $\Lambda_{l,\mathrm{II}}^{1)}$ in (3.44) as follows

$$\Lambda_{l,\mathrm{II}}^{2)} = \prod_{k=n_{l-1}+1}^{n_l} \delta_k^{\delta_k \kappa_k} \prod_{j=2}^{\delta_k} \frac{\left(a_k^{(j-1)}\right)^{j-1} \left(q_k^{(j)}\right)}{\left(a_k^{(j-1)} + q_k^{(j)}\right)^{j}}, \tag{3.78}$$

with $a_k^{(j-1)}$ and $q_k^{(j)}$ as before in the proof of Proposition 3.4.2. Now this expression is equivalent to $\Delta_1^{(k,j,m)}$ in (3.50) and the distribution of $a_k^{(j-1)}$ and $q_k^{(j)}$ is given by

$$2a_k^{(j-1)} \sim \chi^2\left(2M\left(j-1\right)\right), \tag{3.79}$$

$$2q_k^{(j)} \sim \chi^2\left(2M\right). \tag{3.80}$$

Once again we use (3.35) and the proof follows. □

Finally, we put the pieces together in the following theorem, where we obtain the distribution of $\mathscr{G}_{2)}^{1/M}$.

Theorem 3.4.2. *The GLR for Case 2) given in* (3.13) *has the following distribution under the null hypothesis*

$$\mathscr{G}^{1/M} \overset{\mathcal{D}}{=} \prod_{l=1}^{L} \left\{ \prod_{k=n_{l-1}+1}^{n_l} \prod_{j=1}^{\delta_k} \prod_{m=1}^{\kappa_k} Y_{k,m,j} \right\} \left\{ \prod_{k=n_{l-1}+1}^{n_l} \delta_k^{\delta_k \kappa_k} \prod_{j=2}^{\delta_k} \prod_{m=1}^{\kappa_k} \left(1 - Z_{m,j-1}\right) \left(Z_{m,j-1}\right)^{j-1} \right\}, \tag{3.81}$$

where all Y *and* Z *are independent beta distributed random variables. Specifically, they are given by*

$$Y_{k,m,j} \sim Beta\left\{ M - \left(\sum_{p=n_{l-1}+1}^{k-1} \delta_p \kappa_p + (j-1)\kappa_k + m - 1 \right), \sum_{p=n_{l-1}+1}^{k-1} \delta_p \kappa_p + (j-1)\kappa_k + m - 1 \right\}, \tag{3.82}$$

$$Z_{m,j-1} \sim Beta\left\{M(j-1), M\right\}. \tag{3.83}$$

Proof. This theorem is a direct consequence of Propositions 3.4.3 and 3.4.4, the independence of $\Lambda_{l,\mathrm{I}}^{2)}$ and $\Lambda_{l,\mathrm{II}}^{2)}$, and also the independence of Λ_l for $l = 1, \ldots, L$ in (3.14). □

Now that we have proposed the stochastic representations of the GLR under the null hypothesis for Cases 1) and 2), we will evaluate them numerically in the following section.

3.5 Numerical Evaluation

In order to verify the distributions of the GLR under $\mathcal{H}_0$ derived in the previous section, we run Monte Carlo simulations. First, we generate complex Gaussian data according to our null hypothesis and compute the GLRs given by (3.12) for Case 1) and (3.13) for Case 2). Second, we draw independent Beta random variables and form the products according to Theorems 3.4.1 and 3.4.2. Finally, we will compare the agreement of both distributions based on their empirical cumulative distribution functions (ECDF). We consider two different sets of parameters (scenarios) and for each we generate data for Cases 1) and 2).

First, we consider Scenario 1) with the following parameters. We choose $\mathbf{R}_1$ with $L = 3$ blocks on its main diagonal with dimensions $\lambda_1 = 7$, $\lambda_2 = 14$, and $\lambda_3 = 8$, i.e., $N = 29$. Furthermore, $\mathbf{R}_0$ is composed of $K = 6$ diagonal blocks with $\delta_k = \{2, 1, 3, 2, 2, 1\}$ and $\kappa_k = \{2, 3, 4, 1, 2, 4\}$ for $k = 1, \ldots, 6$. Given λ_l, δ_k, κ_k it follows that $n_1 = n_2 = n_3 = 2$. Finally, we consider $M = 15$ realizations. Once again note that we consider the distribution under $\mathcal{H}_0$.

In Figures 3.2 and 3.3 we illustrate the ECDFs of 10^6 Monte Carlo simulations with the setup described in the preceding paragraph for Cases 1) and 2), respectively. In blue we illustrate the ECDF of product of beta distributed random variables and the red graph corresponds to the ECDF of the GLRs. As can be seen the graphs for both distributions exactly overlap which shows that the distribution of the GLR under the null hypothesis is identical to the distribution of the product of independent Beta random variables.

For Scenario 2), we choose $\mathbf{R}_1$ with $L = 2$ blocks of dimensions $\lambda_1 = 15$ and $\lambda_2 = 8$ on its main diagonal. Hence, we consider $N = 23$ samples and additionally $M = 20$ realizations. $\mathbf{R}_0$ is composed of $K = 5$ diagonal blocks with $\delta_k = \{3, 2, 1, 2, 2\}$ and $\kappa_k = \{2, 3, 3, 1, 3\}$ for $k = 1, \ldots, 5$, which results in $n_1 = 3$ and $n_2 = 2$. Again we compare the distributions of the product of independent Beta random variables with the GLRs for Cases 1) and 2) via Monte Carlo simulations. The results are shown in Figures 3.4 and 3.5. Similar to Scenario 1), we can again observe that the distributions of the products of independent Beta random variables agree very well with the distribution of the GLRs. Hence, we have demonstrated that the distribution of the product of independent Beta random variables can be used to obtain the threshold that allows us to set the probability of false alarm when using the GLRT.

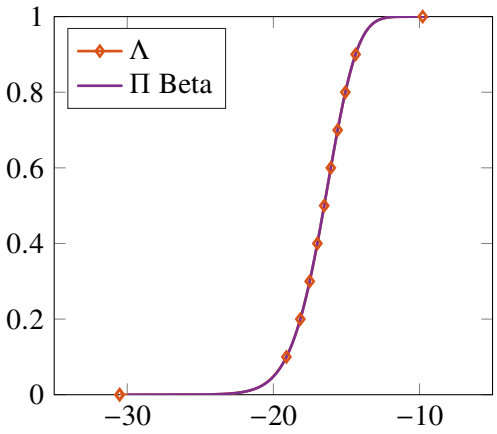

Figure 3.2: ECDF illustrating the agreement of distributions of the product of Beta random variables and the GLR for Case 1) under Scenario 1).

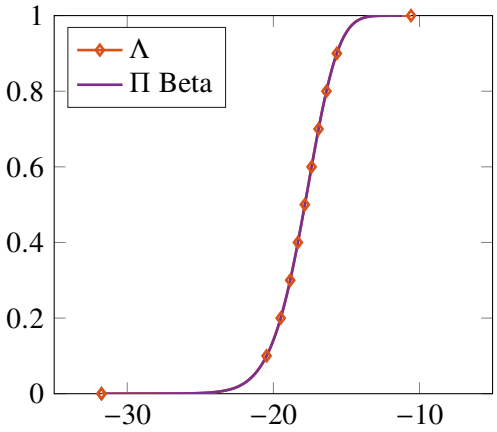

Figure 3.3: ECDF illustrating the agreement of distributions of the product of Beta random variables and the GLR for Case 2) under Scenario 1).

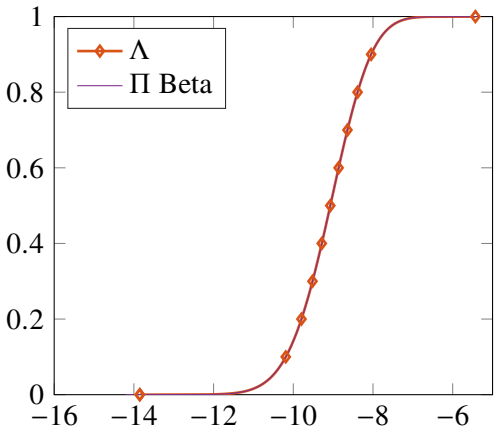

Figure 3.4: ECDF illustrating the agreement of distributions of the product of Beta random variables and the GLR for Case 1) under Scenario 2).

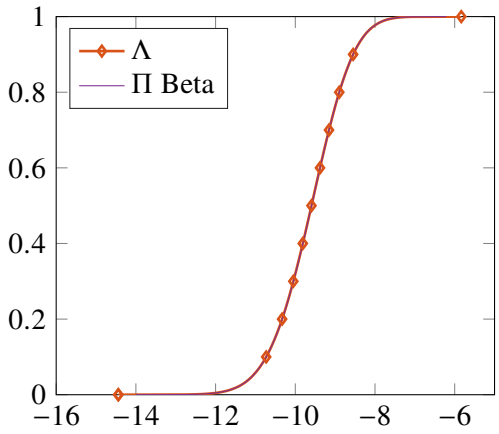

Figure 3.5: ECDF illustrating the agreement of distributions of the product of Beta random variables and the GLR for Case 2) under Scenario 2).

3.6 Conclusion

In this chapter we have derived a mathematical expression for the distribution of a general class of GLRTs under the null hypothesis. These GLRTs aim in testing their covariance structures, specifically, we test for different block-diagonal structures of covariance matrices. It is shown that the stochastic representations of such ratios under the null hypothesis are distributed as the product of independent Beta random variables. Although this is not a closed-form expression it still allows us to fix the probability of false alarm for these GLRTs without having to rely on asymptotic distributions to find the threshold. This result will be applied in the subsequent chapters for the derived GLRT statistics.

4 Joint detection of almost-cyclostationary signals and estimation of their cycle period

In this chapter we consider the problem of detecting almost-cyclostationary signals in a single array. Specifically, we consider the case of a sampled continuous-time CS signal that becomes discrete-time ACS after sampling. Moreover, we assume that the cycle period of the continuous-time signal is unknown. In order to approach this problem, we formulate it as a multiple hypothesis test combined with a resampling stage. Together it allows us to 1) to detect the presence of an ACS signal and 2) infer the cycle period. Monte Carlo simulations have been performed in order to demonstrate the detection performance and to compare with state-of-the-art detectors[1].

4.1 Introduction

Continuous-time signals are typically sampled before further processing. A sampled discrete-time signal is almost-cyclostationary (ACS) if the sampling interval is not a sub-multiple of the cycle period. This is generally the case if the cycle period of the continuous-time signal is unknown [8,44]. In this chapter we propose a detector for discrete-time second-order ACS processes at a single L dimensional array.

As we have previously outlined in the introduction of this thesis, the detection of ACS signals is of interest in many fields. For instance, in communications it is of great importance in spectrum sensing for cognitive radio, and in mechanics it can be used for fault gear or bearing diagnostics [9, 10]. Generally, there are different kinds of CS detectors. Some test for non-zero cyclic autocorrelation function, such as the detector in [35], which extends [34] to multivariate signals. Others [36, 109] test for correlation between the process and its frequency-shifted version. A third class of ACS detectors determines whether the Loève spectrum has support on lines parallel to the stationary manifold [37]. The GLRT and LMPIT for the detection of multivariate comprises an interpretation

[1] This chapter is based on the papers "Detection of Almost-Cyclostationarity : An Approach Based on a Multiple Hypothesis Test, S. Horstmann, D. Ramírez, P.J. Schreier, *Proc. Asilomar Conference on Signals, Systems and Computers, 2017*" and "Joint Detection of Almost-Cyclostationary Signals and Estimation of Their Cycle Period, S. Horstmann, D. Ramírez, P.J. Schreier, *IEEE Signal Processing Letters, 2018*"

of all these three features of CS processes. However, these state-of-the-art (A)CS detectors assume prior knowledge of the cycle period, which, in practice might be unknown or not known exactly due to clock or oscillator errors, which decreases the performance of the detectors [43]. Furthermore, the cycle period itself is an interesting signal parameter. For instance, in communications it relates to the symbol rate and carrier frequency [9, 10]. For this reason it is desirable to use a detector that jointly detects ACS signals and estimates their cycle period.

To this end we propose a technique that detects the presence of ACS signals with unknown cycle period in wide-sense stationary (WSS) noise and simultaneously estimate their cycle period. The idea behind our approach is as follows. A multiple hypothesis test is employed to determine the unknown integer part, and a resampling stage deals with the unknown fractional part. For each potential integer part the signal is resampled at potential fractional parts such that the optimal resampling rate yields a CS signal with cycle period equal to the candidate integer part. This allows us to apply the GLRT derived in [38] for each candidate integer part. Subsequently, the multiple hypothesis test is used to determine the unknown integer part. To reach a decision in the overall test WSS vs. ACS signals we use Holm's sequentially rejective test [84], which controls the probability of false alarm of the overall test and simultaneously provides an estimate of the integer part of the cycle period. This result combined with the corresponding resampling rate yields an estimate of the cycle period.

In order to achieve good detection and estimation performance, the grid of candidate resampling rates has to be fine enough, which increases the computational complexity of the technique. For this reason we propose a filter bank structure to efficiently resample the signal at many different rates.

4.2 Problem Formulation

We consider a continuous-time zero-mean multivariate process $\mathbf{u}(t) \in \mathbb{C}^L$ that is second-order cyclostationary with cycle period T_0, or, equivalently, cyclostationary with fundamental cycle frequency $\alpha_0 = 1/T_0$. Hence, it has a periodic matrix-valued covariance function with period T_0

$$\mathbf{R}(t, \tau) = \mathrm{E}\left[\mathbf{u}(t)\mathbf{u}^{\mathrm{H}}(t-\tau)\right] = \mathbf{R}(t + T_0, \tau) \in \mathbb{C}^{L\times L}. \tag{4.1}$$

Once the signal $\mathbf{u}(t)$ is sampled with sampling interval $T_s < T_0$, the discrete-time signal $\mathbf{u}[n]$ is ACS with fundamental cycle frequency $\tilde{\alpha}_0 = T_s/T_0$ [8]. Now we define a cycle period P as the reciprocal of $\tilde{\alpha}_0$, i.e. $P = T_0/T_s$, which is in general a real number rather than an integer. Assuming that $T_0 \geq 3/2T_s$, we can divide the cycle period P into an integer and fractional part as follows

$$P = P_{\mathrm{int}} + \epsilon, \tag{4.2}$$

where $P_{\text{int}} = 2, 3, \ldots$ and $\epsilon \in [-0.5, 0.5)$. Note that if we dropped the assumption $T_0 \geq 3/2T_s$ and allowed for $T_0 < 3/2T_s$, the range of integer part of the cycle period would also need to cover the case $P_{\text{int}} = 1$, which however corresponds to a wide-sense stationary signals, which, by design, we cannot differentiate from noise.

Now our aim is to solve the following hypothesis test

$$\begin{aligned} \mathcal{H} &: \mathbf{u}[n] \text{ is WSS,} \\ \mathcal{A} &: \mathbf{u}[n] \text{ is ACS.} \end{aligned} \tag{4.3}$$

Additionally, if the presence of an ACS signal is detected, then we simultaneously estimate the cycle period P. We assume that $\mathbf{u}[n]$ is proper complex Gaussian and that we observe M i.i.d. realizations of $\mathbf{u}[n]$ of length N. We propose a detector to solve the hypothesis test (4.3) by combining a resampling stage, which enables to apply the GLRT proposed in [38], with a multiple hypothesis test. The resampling stage allows us to estimate the fractional part of the cycle period, and the multiple hypothesis test provides an estimate for the integer part if the presence of ACS signals was detected. In order to reliably estimate the fractional part of the cycle period, it is necessary to resample the signal at many different rates. Since this is a computationally complex process, we also propose a filter bank structure to decrease the computational costs.

4.3 Resampling Stage

In order to handle the unknown fractional part of the cycle period, let us assume first that the integer part of the cycle period P_{int} is known. Now our goal is to find the resampling rate Δ such that the resampled signal becomes CS with cycle period P_{int}, which allows us to apply the test CS vs. WSS proposed in [38]. Specifically, for a set of D candidate resampling rates

$$\Delta_d = \frac{P_{\text{int}}}{P_{\text{int}} + \epsilon_d}, \quad d = 1, \ldots, D, \tag{4.4}$$

where $\epsilon_d = -0.5 + (d-1)/D$, we obtain the resampled signal $\tilde{\mathbf{u}}_d[m]$ for which we compute the GLRT statistic proposed in [38]. Let us briefly outline the computation of the statistic:

$$\mathcal{G}(\Delta_d | P_{\text{int}}) = \prod_{k=1}^{N/P_{\text{int}}} \det\left(\hat{\mathbf{C}}_k\right), \tag{4.5}$$

where $\hat{\mathbf{C}}_k$ is the kth $LP_{\text{int}} \times LP_{\text{int}}$ diagonal block of the coherence matrix

$$\hat{\mathbf{C}} = \left[\text{diag}_L(\hat{\mathbf{S}})\right]^{-1/2} \text{diag}_{LP_{\text{int}}}(\hat{\mathbf{S}}) \left[\text{diag}_L(\hat{\mathbf{S}})\right]^{-1/2}. \tag{4.6}$$

Here, $\mathrm{diag}_L(\hat{\mathbf{S}})$ and $\mathrm{diag}_{LP_{\mathrm{int}}}(\hat{\mathbf{S}})$ denote block-diagonal matrices obtained from the $L \times L$ and $LP_{\mathrm{int}} \times LP_{\mathrm{int}}$ blocks on the diagonal of

$$\hat{\mathbf{S}} = \frac{1}{M} \sum_{i=1}^{M} \mathbf{z}_i \mathbf{z}_i^H, \tag{4.7}$$

where

$$\mathbf{z}_i = (\mathbf{L}_{N,N/P_{\mathrm{int}}} \otimes \mathbf{I}_L)(\mathbf{F}_N \otimes \mathbf{I}_L)^H \mathbf{y}_i, \tag{4.8}$$

$\mathbf{L}_{N,N/P_{\mathrm{int}}}$ is the commutation matrix, $\mathbf{F}_N$ is the DFT matrix, and $\mathbf{y}_i = [\tilde{\mathbf{u}}_i^T[0] \cdots \tilde{\mathbf{u}}_i^T[N-1]]^T$.

Subsequently, we find the maximum likelihood (ML) estimate of the resampling rate Δ_d or equivalently, due to their relation in (4.4), the ML estimate of the fractional part of the cycle period, by maximizing the likelihood under $\mathcal{A}$. Equivalently, we can minimize the GLR (4.5) for a given integer part P_{int}

$$\Delta_{\min} = \underset{\Delta_{d=1,\ldots,D}}{\arg\min}\, \mathcal{G}(\Delta_d | P_{\mathrm{int}}). \tag{4.9}$$

In order to make the resampling grid fine enough, D must be a large number, and hence the resampling of the signal at D different rates is computationally quite expensive. There are efficient strategies to convert the sample rate, but only for a single given Δ_d, see e.g. [110]. In the following, we propose a filter bank that is designed to deal with a set of D different resampling rates for each candidate integer P_{int}.

Sample Rate Conversion

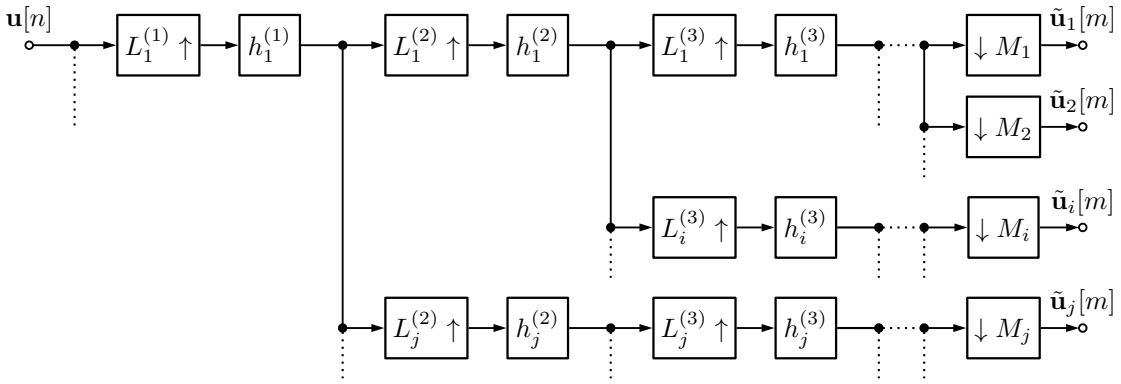

Figure 4.1: Proposed filter bank structure to convert the sampling rate $1/T_s$ to D different rates.

The sample rate conversion with the proposed filter bank structure is illustrated in Figure 4.1. Each resampling rate Δ_d for $d = 1, \ldots, D$ specifies an interpolation factor L_d and decimation factor M_d such that $\frac{L_d}{M_d} = \Delta_d$. Since the number of resampling rates D can be quite large, it is desirable to

exploit common interpolation stages among all upsampling rates L_d for $d = 1, \ldots, D$ in order to save as many computations as possible. To this end, each L_d is factorized into a product of prime numbers $L_d = \prod_{\lambda=1}^{\Lambda} L_d^{(\lambda)}$, where $L_d^{(1)} \geq L_d^{(2)} \geq \cdots \geq L_d^{(\Lambda)}$. Within the set of D upsampling rates, we identify common factors such as $\prod_{\kappa=1}^{K} L_i^{(\kappa)} = \prod_{\kappa=1}^{K} L_j^{(\kappa)}$ for $K < \Lambda$, $i \neq j$ and $i, j \in \{1, \ldots, D\}$. Hence, the signal is interpolated only once at those common stages. After the signal has been interpolated at all required rates, it is downsampled by the respective rates M_d. Each interpolation stage is implemented as a polyphase filter.

4.4 Multiple Hypothesis Test

The integer part of the cycle period in the detection problem can be handled by a multiple hypothesis test, which at the same time provides an estimate of the integer part of the cycle period P_{int}. To this end, the multiple hypothesis test is implemented as a set of binary tests with a common null hypothesis, i.e.

$$\mathcal{H} : \ \mathbf{u}[n] \text{ is WSS}, \tag{4.10}$$

versus the following set of alternatives

$$\begin{aligned} \mathcal{A}_1 : \ & \mathbf{u}[n] \text{ is ACS with } P_{\text{int}} = 2, \\ \mathcal{A}_2 : \ & \mathbf{u}[n] \text{ is ACS with } P_{\text{int}} = 3, \\ & \vdots \\ \mathcal{A}_K : \ & \mathbf{u}[n] \text{ is ACS with } P_{\text{int}} = P_{\max}, \end{aligned} \tag{4.11}$$

where $P_{\max} = K + 1$ is the largest integer cycle period under consideration. In the overall test WSS vs. ACS we reject $\mathcal{H}$ if it is rejected in at least one of the binary tests. The decision is reached by employing Holm's sequentially rejective test [84]. This test controls the familywise error rate (FWER), which is the probability of at least one false rejection and, therefore, in our case, it is identical to the probability of false alarm p_{fa} of the test in (4.3), i.e.

$$\begin{aligned} \text{FWER} &= P\left(\text{reject } \mathcal{H} \text{ in any test } \mathcal{H} \text{ vs. } \mathcal{A}_i | \mathcal{H}\right) \\ &\equiv p_{\text{fa}}. \end{aligned} \tag{4.12}$$

Following Holm's test procedure we reach a decision in the overall test by obtaining the minimum p-value p_γ, where γ indicates the index of the corresponding binary test, i.e.

$$\gamma = \underset{i=1,\ldots,K}{\arg\min}\, p_i, \tag{4.13}$$

which we compare to a threshold p_{fa}/K. Hence, if $p_\gamma \geq p_{\text{fa}}/K$, then we fail to reject $\mathcal{H}$, and if $p_\gamma < p_{\text{fa}}/K$, we reject $\mathcal{H}$ and the signal is said to be ACS. To estimate the p-values we exploit that the distribution of the GLR under $\mathcal{H}$ for a given Δ_d, $\mathcal{G}(\Delta_d|P_{\text{int}})$, can be approximated as the distribution of a product of independent Beta random variables as shown in the following proposition:

Proposition 4.4.1. *The distribution of the GLRT statistic under the null hypothesis is asymptotically given by a product of independent Beta random variables that is*

$$\mathcal{G}(\Delta_d|P_{int}) \overset{\mathcal{D}}{=} \prod_{l=1}^{N/P_{\text{int}}} \left\{ \prod_{k=1}^{P_{\text{int}}} \prod_{m=1}^{L} U_{k,m} \right\} \tag{4.14}$$

where all U are independent Beta distributed random variables. Specifically, they are given by

$$U_{k,m} \sim Beta\left\{M - ((k-1)L + m - 1), (k-1)L\right\}. \tag{4.15}$$

Proof. The proof directly follows from Theorem 3.4.1 for the following parameters: The total number of blocks under the alternative is given by $L = N/P_{\text{int}}$, where each block is of dimension $\lambda_l = LP_{\text{int}}\ \forall l$. Finally, each block under the alternative is tested against $n_l = P_{\text{int}}\ \forall l$ blocks under the null with $\kappa_k = L,\ k = 1, \ldots, P_{\text{int}}$, which concludes the proof. □

In order to obtain the distribution of the minimum value of the statistic $\mathcal{G}(\Delta_{\min}|P_{\text{int}})$, we simplify the problem by assuming independence among the GLRTs to apply results from order statistics [111]. Specifically, we approximate the cumulative distribution of $\mathcal{G}(\Delta_{\min}|P_{\text{int}})$ by the first-order statistic

$$F_1(\mathcal{G}(\Delta_{\min}|P_{\text{int}})) = 1 - [1 - F(\mathcal{G}(\Delta_{\min}|P_{\text{int}}))]^D. \tag{4.16}$$

This result allows us to estimate the p-value p_i for each binary hypothesis test $\mathcal{H}$ vs. $\mathcal{A}_i$ for $i = 1, \ldots, K$.

In Figure 4.2 we compare the ECDFs of the first-order statistic computed under the null hypothesis, shown by the dashed plot, with the approximate distribution obtained from the product of independent Beta distributed random variables, shown by the solid plot. It can be observed that the distributions agree quite well, even though the dashed plot is a little left of the blue one. This will result in slightly underestimated p-values, which will also be discussed in the following section. Further simulations (not shown here) confirm these findings.

At the same time we can obtain an estimate of the cycle period of the ACS process as follows. If the multiple hypothesis test rejects $\mathcal{H}$, the integer part P_{int} is estimated as

$$\hat{P}_{\text{int}} = \gamma + 1. \tag{4.17}$$

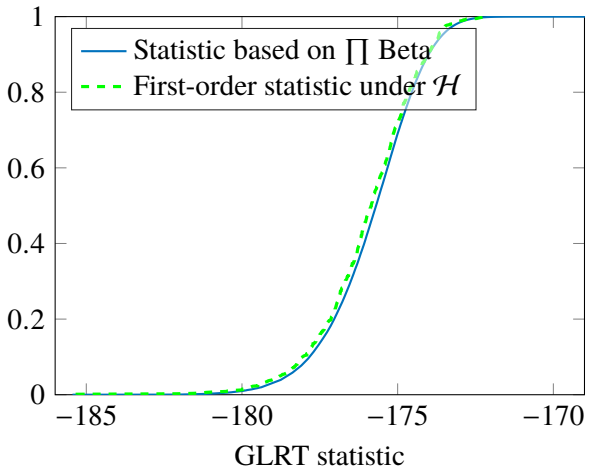

Figure 4.2: Comparison of ECDFs of the first-order statistic computed under the null hypothesis with the approximated distribution shown in blue for the following parameters: $P_{\text{int}} = 4$, $N = 768$, $L = 2$, $M = 25$ and $D = 20$

Furthermore, the fractional part of the cycle period is obtained from (4.4) as

$$\hat{\epsilon} = \hat{P}_{\text{int}}(1/\Delta_{\text{min}}^{\gamma} - 1), \tag{4.18}$$

where $\Delta_{\text{min}}^{\gamma}$ denotes the optimal resampling rate corresponding to hypothesis $\mathcal{A}_{\gamma}$. Hence, the estimate of the cycle period is given by

$$\hat{P} = \hat{P}_{\text{int}} + \hat{\epsilon}. \tag{4.19}$$

The complete test procedure is summarized in Algorithm 4.1.

4.5 Numerical Results

In this section we evaluate the performance of the proposed technique and compare it to the techniques presented in [35] and [36]. For the evaluation we use Monte Carlo simulations in a communications scenario:

$$\begin{aligned} \mathcal{H}: \ & \mathbf{u}[n] = \mathbf{w}[n], \\ \mathcal{A}: \ & \mathbf{u}[n] = \mathbf{H}[n] * \mathbf{s}[n] + \mathbf{w}[n], \end{aligned} \tag{4.20}$$

where $\mathbf{w}[n] \in \mathbb{C}^L$ is colored Gaussian noise generated with a moving average filter of order 20 and $\mathbf{H}[n] \in \mathbb{C}^{L \times L}$ is a Rayleigh fading channel with a delay spread of 10 times the symbol duration T_0 and a sampling frequency of $f_s = 1.2$ MHz. The ACS transmission signal $\mathbf{s}[n] \in \mathbb{C}^L$ is obtained by

Algorithm 4.1 Multiple-Hypothesis-Test Procedure

Input: Discrete-time process $\mathbf{u}[n]$, maximum integer part $P_{\max}$, number D of grid points of fractional parts, desired probability of false alarm α.

Resampling

$k \leftarrow 1$
for $k \leq P_{\max} - 1$ **do**
  $P_{\text{int}} = k + 1$
  $d \leftarrow 1$
  for $d \leq D$ **do**
    $\tilde{\mathbf{u}}[n] \leftarrow$ resample $\mathbf{u}[n]$ by Δ_d
    Compute $\mathcal{G}(\Delta_d)$ with (4.5)
    $d \leftarrow d + 1$
  end for
  Obtain $\Delta_{\min} = \underset{\Delta_{d=1,\ldots,D}}{\arg\min}\, \mathcal{G}(\Delta_d)$
  Determine p-value p_k of $\mathcal{G}(\Delta_{\min})$
  $k \leftarrow k + 1$
end for

Multiple Hypothesis Test

$\gamma = \underset{i=1,\ldots,K}{\arg\min}\, p_i$,
if $p_\gamma \geq p_{\text{fa}}/(P_{\max} - 1)$ **then**
  Fail to reject $\mathcal{H}$: $\mathbf{u}[n]$ is WSS
else
  Reject $\mathcal{H}$: $\mathbf{u}[n]$ is ACS with $\hat{P}_{\text{int}} = \gamma + 1$ and $\hat{\epsilon} = \hat{P}_{\text{int}}(1/\Delta_{\min}^{\gamma} - 1)$
end if

subsampling a long QPSK-signal with raised-cosine pulse shaping and roll-off factor 1. In order to obtain M realizations, which are required to obtain the test statistic, we generate one long sequence $\mathbf{u}[n]$ and cut that into M pieces.

We obtain the joint probability of detection and correctly estimating the cycle period

$$p_{\text{d}} = P(p_\gamma < p_{\text{fa}}/K \;\cap\; |\hat{P} - P| < 1/D \mid \mathcal{A}) \tag{4.21}$$

for a given probability of false alarm p_{fa}. The benchmark techniques [35] and [36] do not provide estimates of the cycle period — they rather need this knowledge a priori. For a fair comparison we obtain the test statistics of [35] and [36] for the same grid of cycle periods we use for our technique. Moreover, instead of performing a multiple hypothesis test, we simply use the maximum test statistic. If this is greater than the threshold used to solve the detection problem, the cycle period corresponding to this maximum test statistic is its estimate. Other cycle period estimation techniques such as [37, 112] cannot be used as comparisons since they are only cycle period estimators but not ACS signal detectors.

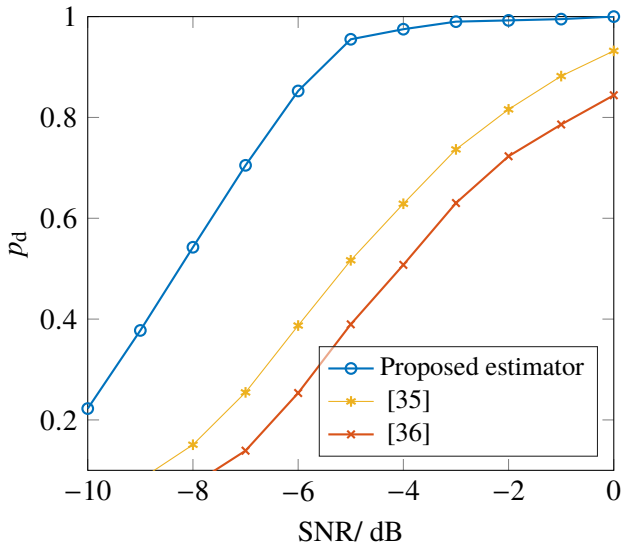

Figure 4.3: Probability of jointly detecting ACS signals and estimating P for $p_{fa} = 0.05$ for the following scenario: $P = 3.19$, $L = 2$ antennas, $N = 300$ samples, $M = 25$ realizations, $P_{max} = 10$, and $D = 100$

For the simulation we use the following parameters: A symbol duration of $T_0 = 2.6583\mu$s, which yields a cycle period of $P = 3.19$, $L = 2$ antennas, $N = 300$ samples per antenna, $M = 25$ snapshots, and a probability of false alarm fixed at $p_{fa} = 5\%$. In Figure 4.3 we can see p_d for a scenario with a resampling grid of size $D = 100$, where the true cycle period lies on the grid of candidates. As can be seen, the proposed technique substantially outperforms the two competitors. For instance, for an SNR of −3dB we observe a relative performance gain of 35% and 58% of our technique compared to [35] and [36], respectively. If we now choose a grid size $D = 110$, where the true cycle period is off-grid by $0.9 \cdot 10^{-3}$, we observe in Figure 4.4 that the performance of all three techniques decreases although the proposed method still significantly outperforms the competitors.

Let us now consider an appropriate choice of D. Generally, D depends on the frequency resolution of the estimates, which is in the order of $1/N$. Since the resolution of the grid of fractional parts is in the order of $1/D$, a reasonable choice for D would be $D \geq N$ as illustrated in Figure 4.5. It shows the joint probability of detection and correctly estimating the cycle period p_d as a function of grid size D of the grid of fractional parts for a cycle period of $P = 3.19$, $L = 2$, $N = 300$ samples, $M = 25$ snapshots, and a probability of false alarm fixed at $p_{fa} = 5\%$. Furthermore, the true cycle period lies off the grid of candidates. It can be observed that the performance increases until $D \geq N$, where it saturates since then the limiting factor is the frequency resolution determined by N and not the distance between grid points and fractional parts determined by D. However, the larger D the higher the computational costs. Therefore, the price to pay for good performance is computational complexity. Comparing the relative computation time of a MATLAB implementation with respect

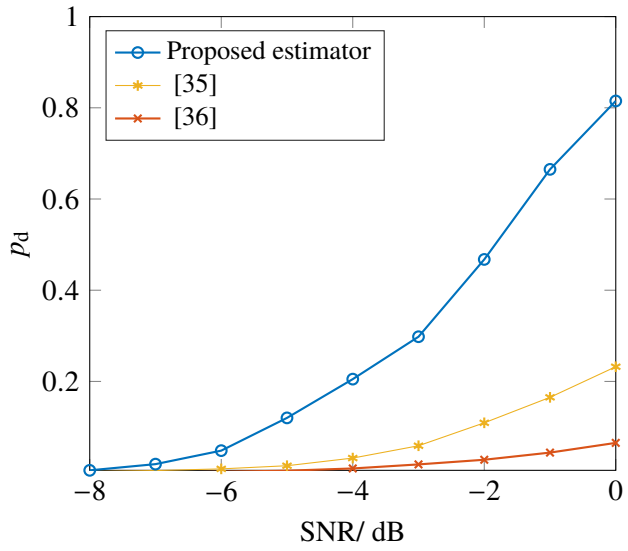

Figure 4.4: Probability of jointly detecting ACS signals and estimating P for $p_{fa} = 0.05$ for the following scenario: $P = 3.19$, $L = 2$ antennas, $N = 300$ samples, $M = 25$ realizations, $P_{max} = 10$, and $D = 110$

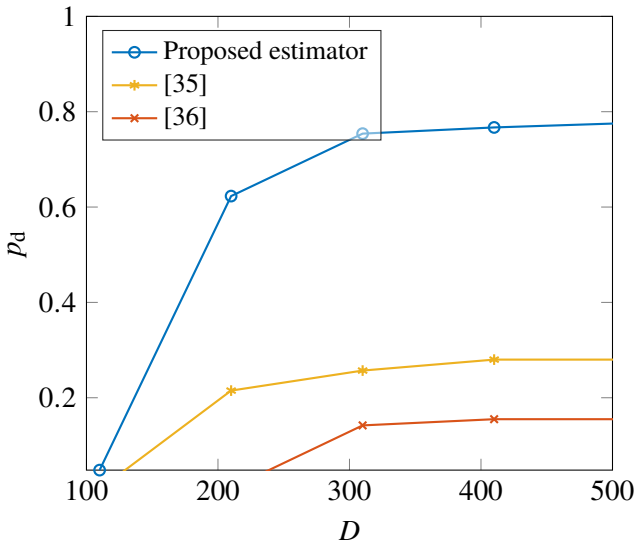

Figure 4.5: Probability of jointly detecting ACS signals and estimating P for $p_{fa} = 0.05$ as a function of D for the following scenario: $P = 3.19$, $L = 2$, $N = 300$ samples, $M = 25$ realizations, $P_{max} = 10$, and SNR $= -6$ dB

to the technique proposed in [35] reveals that our detector requires double the computation time, whereas [36] requires only one fifth of the time of [35]. The relative computational complexity of the techniques is independent of the grid size D.

4.6 Conclusion

In this chapter, we have presented a technique that jointly detects almost-cyclostationarity and estimates the cycle period. While our technique is more computationally complex than competing techniques, it also substantially outperforms them, which we have shown for the single array application for spectrum sensing in cognitive radio. Hence, increased performance comes at the price of computational complexity.

5 Two-channel passive detection

This chapter considers the two-channel passive detection of cyclostationarity. Specifically, based on two receiver arrays, one of which is a reference array, we consider the problem of detecting a Gaussian cyclostationary signal corrupted with temporally colored and spatially correlated noise. We derive an asymptotic GLRT and show that neither a UMPIT nor an LMPIT for close hypotheses exist. Furthermore, an insightful interpretation of the test statistic is provided, which leads to the proposition of an LMPIT-inspired test statistic. Monte Carlo simulations demonstrate that the proposed detectors outperform comparable state-of-the-art detectors[1].

5.1 Introduction

In this chapter we consider a 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 [49]. 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 [113], digital beamforming [54, 114] 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., [56, 57]. Typically, the IO is a commercial radio or TV broadcast system, or it could be a space-based source such as communication or navigation satellites [46–48].

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., [53, 57, 58, 61–63].

[1] This chapter is based on the papers "Two-channel passive detection exploiting cyclostationarity, S. Horstmann, D. Ramirez, P.J. Schreier, *Proc. 27th European Signal Proc. Conf. (EUSIPCO), 2019* " and "Two-Channel Passive Detection of Cyclostationary Signals, S. Horstmann, D. Ramirez, P.J. Schreier, *IEEE Transactions on Signal Processing, 2020*"

Although this resembles the matched filter, it is suboptimal due to noise at the RC [58]. In [64–67] 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 [64] considered the detection of a rank-one signal received by a multiantenna array, whereas [65] 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 [66]. Finally, [67] 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 [68] and [51], where [68] assumed unknown and [51] assumed known noise variance. For the same problem, an approximate Bayesian test was derived in [68] and the exact Bayesian test was presented in [69]. The work in [70] proposed an ad-hoc detector based on the generalized coherence [71].

However, all these aforementioned detectors do not exploit the fact that digital communication signals transmitted by potential IOs are cyclostationary [22]. For single array detection this property was exploited in [115, 116], which derived locally optimum tests for a known signal waveform and different assumptions on the noise. In [115] temporally and spatially white Gaussian noise was considered, whereas [116] 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 [38] and specialized to various noise structures in [92, 117].

5.1.1 Outline of this chapter

The detection problem is formulated in Section 5.2 followed by the derivation of the GLRT in Section 5.3. In Section 5.4 we examine the existence of the UMPIT and the LMPIT, and in Section 5.5 we provide an interpretation of the statistics. In Section 5.5 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 5.7. This chapter is concluded in Section 5.8.

5.2 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.[2] Furthermore, we assume that the IO is equipped with L_I antennas, and a noisy version of its transmitted signal is received at the

[2]Note that the derivations can easily be generalized to different numbers of antennas at both arrays.

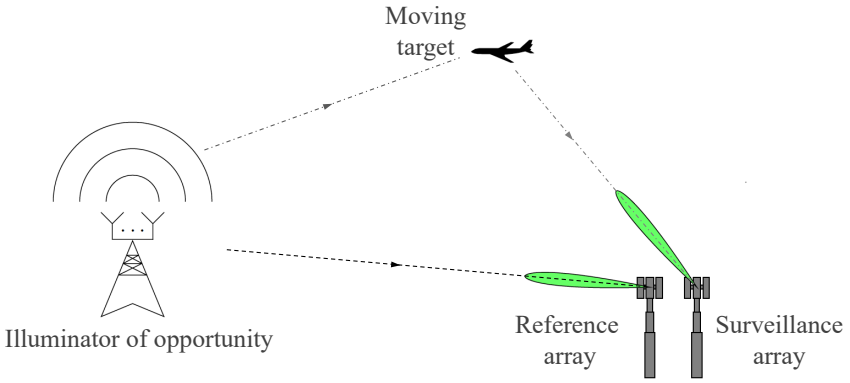

Figure 5.1: MIMO passive bistatic radar system that consists of an IO, a reference, and a surveillance array. The reference array receives the direct-path signal from the IO illustrated by the black dashed line and in the presence of a moving target the surveillance array receives the target-path signal, which is depicted by the gray dashed dotted line.

RC. The cancellation of interference and clutter in the RC has been considered in e.g. [52], [53]. 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 [50] or cancellation by signal processing techniques presented in e.g. [54–57]. The complete cancellation of direct-path interference in the SC is, admittedly, an idealized assumption as was pointed out in [54] and the works in [58–60] 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 [51, 60, 67]. 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 [118]. Thus, taking into account the aforementioned assumptions, the passive radar system considered in this chapter is illustrated in Figure 5.1. The detection problem can be formulated as

$$\begin{aligned} \mathcal{H}_0: &\begin{cases} \mathbf{u}_s[n] = \mathbf{v}_s[n], \\ \mathbf{u}_r[n] = \mathbf{H}_r[n] * \mathbf{s}[n] + \mathbf{v}_r[n], \end{cases} \\ \mathcal{H}_1: &\begin{cases} \mathbf{u}_s[n] = \mathbf{H}_s[n] * \mathbf{s}[n] + \mathbf{v}_s[n], \\ \mathbf{u}_r[n] = \mathbf{H}_r[n] * \mathbf{s}[n] + \mathbf{v}_r[n], \end{cases} \end{aligned} \tag{5.1}$$

for $n = 0, \ldots, NP-1$ 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 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 $\mathbf{R}_{ss}[n, m] = \mathrm{E}[\mathbf{s}[n]\mathbf{s}^H[n-m]] = \mathbf{R}_{ss}[n+P, m]$ is periodic in n with period P. Since the transmitted signal $\mathbf{s}[n]$ 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., [34, 112, 119]. Moreover, we assume that $L_I \geq L$, which implies that the cyclic (cross) power spectral densities (PSD) of $\mathbf{H}_s[n] * \mathbf{s}[n]$ and $\mathbf{H}_r[n] * \mathbf{s}[n]$ 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

$$\mathbf{x}[n] = \left[\mathbf{u}^T[nP] \ \cdots \ \mathbf{u}^T[(n+1)P-1]\right]^T \in \mathbb{C}^{LP}, \tag{5.2}$$

which is WSS if the L-variate process $\mathbf{u}[n] \in \mathbb{C}^L$ is CS with cycle period P [3]. This implies that its matrix-valued covariance function $\mathbf{R}_{xx}[n, m] = \mathrm{E}\left[\mathbf{x}[n]\mathbf{x}^H[n-m]\right] = \mathbf{R}_{xx}[m]$ only depends on the time-shift. Moreover, the stack of N observations $\mathbf{w} = \left[\mathbf{x}^T[0] \ \cdots \ \mathbf{x}^T[N-1]\right]^T \in \mathbb{C}^{LNP}$ has a Hermitian block-Toeplitz structured covariance matrix with block size LP:

$$\mathbf{R}_{ww} = \mathrm{E}[\mathbf{w}\mathbf{w}^H] = \begin{bmatrix} \mathbf{R}_{xx}[0] & \cdots & \mathbf{R}_{xx}[N-1] \\ \vdots & \ddots & \vdots \\ \mathbf{R}_{xx}^H[N-1] & \cdots & \mathbf{R}_{xx}[0] \end{bmatrix} \in \mathbb{T}_{LP}^{LNP}. \tag{5.3}$$

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

$$\mathbf{w}_r = \left[\mathbf{u}_r^T[0] \ \cdots \ \mathbf{u}_r^T[NP-1]\right]^T \in \mathbb{C}^{LNP}, \tag{5.4}$$

has covariance matrix $\mathbf{R}_r = \mathrm{E}[\mathbf{w}_r \mathbf{w}_r^H]$, 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 $\mathbf{u}_s[n]$

$$\mathbf{w}_s = \left[\mathbf{u}_s^T[0] \ \cdots \ \mathbf{u}_s^T[NP-1]\right]^T \in \mathbb{C}^{LNP}, \tag{5.5}$$

has a block-Toeplitz structured covariance matrix $\mathbf{R}_s^{(0)} = \mathrm{E}[\mathbf{w}_s \mathbf{w}_s^H | \mathcal{H}_0]$ with block size L under the null hypothesis, where $\mathbf{u}_s[n] \in \mathbb{C}^L$ is WSS, and covariance matrix $\mathbf{R}_s^{(1)} = \mathrm{E}[\mathbf{w}_s \mathbf{w}_s^H | \mathcal{H}_1]$, 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

$$\mathbf{w} = \begin{bmatrix} \mathbf{w}_s^T & \mathbf{w}_r^T \end{bmatrix}^T \in \mathbb{C}^{2LNP}. \tag{5.6}$$

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 2×2 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,

$$\mathbf{R}_0 = \mathrm{E}[\mathbf{w}\mathbf{w}^H | \mathcal{H}_0] = \begin{bmatrix} \mathbf{R}_s^{(0)} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_r \end{bmatrix}. \tag{5.7}$$

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

$$\mathbf{R}_1 = \mathrm{E}[\mathbf{w}\mathbf{w}^H | \mathcal{H}_1] = \begin{bmatrix} \mathbf{R}_s^{(1)} & \mathbf{R}_{sr} \\ \mathbf{R}_{rs} & \mathbf{R}_r \end{bmatrix}, \tag{5.8}$$

where $\mathbf{R}_{sr} = \mathbf{R}_{rs}^H = \mathrm{E}[\mathbf{w}_s \mathbf{w}_r^H | \mathcal{H}_1]$ 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

$$\begin{aligned} \mathcal{H}_0 &: \mathbf{w} \sim \mathcal{CN}_{2LNP}(\mathbf{0}, \mathbf{R}_0), \\ \mathcal{H}_1 &: \mathbf{w} \sim \mathcal{CN}_{2LNP}(\mathbf{0}, \mathbf{R}_1). \end{aligned} \tag{5.9}$$

As $\mathbf{R}_0$ and $\mathbf{R}_1$ are unknown, (5.9) 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 [38], where it is shown that we can asymptotically ($N \to \infty$) 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 $\mathbf{w}_\clubsuit$,

$$\mathbf{z}_\clubsuit = (\mathbf{L}_{NP,N} \otimes \mathbf{I}_L)(\mathbf{F}_{NP} \otimes \mathbf{I}_L)^H \mathbf{w}_\clubsuit, \tag{5.10}$$

where $\clubsuit \in \{s, r\}$, $\mathbf{F}_{NP}$ is the DFT matrix of size NP, and $\mathbf{L}_{NP,N}$ is the commutation matrix, which fulfills $\text{vec}(\mathbf{A}) = \mathbf{L}_{MN,N}\,\text{vec}\left(\mathbf{A}^T\right)$ for an $M \times N$ matrix $\mathbf{A}$. 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 $n = 0, \ldots, N-1$. Furthermore, the DFT of length NP of $\mathbf{u}_\clubsuit[n]$ is defined as

$$\mathbf{u}_\clubsuit(\theta_k) = \sum_{n=0}^{NP-1} \mathbf{u}_\clubsuit[n] e^{-j\theta_k n}, \tag{5.11}$$

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

$$\mathbf{x}[n] = \left[\mathbf{u}_s^T(\theta_n)\ \mathbf{u}_s^T(\theta_{N+n})\ \cdots\ \mathbf{u}_s^T(\theta_{(P-1)N+n})\right]^T \tag{5.12}$$

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

$$\mathbf{y}[n] = \left[\mathbf{u}_r^T(\theta_n)\ \mathbf{u}_r^T(\theta_{N+n})\ \cdots\ \mathbf{u}_r^T(\theta_{(P-1)N+n})\right]^T. \tag{5.13}$$

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 [10].

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

$$\mathbf{S}_0 = \text{E}[\mathbf{z}\mathbf{z}^H|\mathcal{H}_0] = \begin{bmatrix} \mathbf{S}_s^{(0)} & \mathbf{0} \\ \mathbf{0} & \mathbf{S}_r \end{bmatrix}, \tag{5.14}$$

where the off-diagonal blocks are zero since observations at SC and RC are uncorrelated, and the covariance matrices $\mathbf{S}_s^{(0)} = \text{E}[\mathbf{z}_s\mathbf{z}_s^H|\mathcal{H}_0] \in \mathbb{S}_L^{LNP}$ and $\mathbf{S}_r = \text{E}[\mathbf{z}_r\mathbf{z}_r^H] \in \mathbb{S}_{LP}^{LNP}$ are block-diagonal Hermitian 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 (5.10). 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

$$\mathbf{S}_1 = \mathrm{E}[\mathbf{z}\mathbf{z}^H|\mathcal{H}_1] = \begin{bmatrix} \mathbf{S}_s^{(1)} & \mathbf{S}_{sr} \\ \mathbf{S}_{rs} & \mathbf{S}_r \end{bmatrix}, \tag{5.15}$$

where $\mathbf{S}_s^{(1)} = \mathrm{E}[\mathbf{z}_s\mathbf{z}_s^H|\mathcal{H}_1]$ and $\mathbf{S}_{sr} = \mathbf{S}_{rs}^H = \mathrm{E}[\mathbf{z}_s\mathbf{z}_r^H|\mathcal{H}_1]$ 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

$$\begin{aligned} \mathcal{H}_0 &: \mathbf{z} \sim \mathcal{CN}_{2LNP}(\mathbf{0}, \mathbf{S}_0), \\ \mathcal{H}_1 &: \mathbf{z} \sim \mathcal{CN}_{2LNP}(\mathbf{0}, \mathbf{S}_1). \end{aligned} \tag{5.16}$$

5.3 Derivation of the GLRT

The GLR is given by

$$\mathscr{G} = \frac{p(\mathbf{z}_0, \cdots, \mathbf{z}_{M-1}; \hat{\mathbf{S}}_0)}{p(\mathbf{z}_0, \cdots, \mathbf{z}_{M-1}; \hat{\mathbf{S}}_1)}, \tag{5.17}$$

where $\mathbf{z}_0, \ldots, \mathbf{z}_{M-1}$ denote M independent and identically distributed (i.i.d.) realizations[3] 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

$$p(\mathbf{z}_0, \cdots, \mathbf{z}_{M-1}; \hat{\mathbf{S}}_j) = \frac{1}{\pi^{2LNPM} \det\left(\hat{\mathbf{S}}_j\right)^M} \times \exp\left\{-M \operatorname{tr}\left(\mathbf{Q}\hat{\mathbf{S}}_j^{-1}\right)\right\}, \tag{5.18}$$

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 5.3.1. *The GLR* (5.17) *is given by*

$$\mathscr{G}^{\frac{1}{M}} = \prod_{k=1}^{N} \det\left(\mathbf{D}_k - \mathbf{C}_k\mathbf{C}_k^H\right), \tag{5.19}$$

[3] 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.

where $\mathbf{D}_k$ *is the kth* $LP \times LP$ *block of*

$$\mathbf{D} = \mathrm{diag}_L\left(\mathbf{Q}_s\right)^{-1/2} \mathrm{diag}_{LP}(\mathbf{Q}_s)\, \mathrm{diag}_L\left(\mathbf{Q}_s\right)^{-1/2}, \tag{5.20}$$

and $\mathbf{C}_k$ *the kth* $LP \times LP$ *block of*

$$\mathbf{C} = \mathrm{diag}_L\left(\mathbf{Q}_s\right)^{-1/2} \mathrm{diag}_{LP}(\mathbf{Q}_{sr})\, \mathrm{diag}_{LP}\left(\mathbf{Q}_r\right)^{-1/2}. \tag{5.21}$$

Proof. 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 [120], the ML estimate is given by

$$\hat{\mathbf{S}}_0 = \begin{bmatrix} \mathrm{diag}_L(\mathbf{Q}_s) & \mathbf{0} \\ \mathbf{0} & \mathrm{diag}_{LP}(\mathbf{Q}_r) \end{bmatrix}. \tag{5.22}$$

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

$$\tilde{\mathbf{w}} = \mathbf{T}\mathbf{w}, \tag{5.23}$$

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 $\left[\mathbf{u}_s[n]^T \;\; \mathbf{u}_r[n]^T\right]^T \in \mathbb{C}^{2L}$ 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

$$\tilde{\mathbf{z}} = (\mathbf{L}_{NP,N} \otimes \mathbf{I}_{2L})(\mathbf{F}_{NP} \otimes \mathbf{I}_{2L})^H \tilde{\mathbf{w}}, \tag{5.24}$$

i.e., the covariance matrix $\tilde{\mathbf{S}}_1 = \mathrm{E}\left[\tilde{\mathbf{z}}\tilde{\mathbf{z}}^H | \mathcal{H}_1\right] \in \mathbb{S}_{2LP}^{2LNP}$ is Hermitian and block-diagonal with block size $2LP$. Exploiting properties of the Kronecker product we can rewrite (5.24) as

$$\tilde{\mathbf{z}} = \left[(\mathbf{L}_{NP,N}\mathbf{F}_{NP}^H \otimes \mathbf{I}_2) \otimes \mathbf{I}_L\right] \tilde{\mathbf{w}}. \tag{5.25}$$

Considering (5.10), the linear transformation of $\mathbf{w}$ is given by

$$\mathbf{z} = \left[(\mathbf{I}_2 \otimes \mathbf{L}_{NP,N}\mathbf{F}_{NP}^H) \otimes \mathbf{I}_L\right] \mathbf{w}. \tag{5.26}$$

It can be observed that (5.25) and (5.26) 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

$$\tilde{\mathbf{z}} = \mathbf{T}\mathbf{z}. \tag{5.27}$$

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 $\tilde{\mathbf{S}}_1$ is block-diagonal, we can easily find its ML estimate as

$$\hat{\tilde{\mathbf{S}}}_1 = \mathrm{diag}_{2LP}\left(\tilde{\mathbf{Q}}\right), \tag{5.28}$$

where $\tilde{\mathbf{Q}} = \mathbf{T}\mathbf{Q}\mathbf{T}^T$. After exploiting the invariance of the ML estimate [76], we find

$$\hat{\mathbf{S}}_1 = \mathbf{T}^T\hat{\tilde{\mathbf{S}}}_1\mathbf{T} = \mathbf{T}^T\,\mathrm{diag}_{2LP}\left(\tilde{\mathbf{Q}}\right)\mathbf{T}. \tag{5.29}$$

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 $\hat{\tilde{\mathbf{S}}}_1$ with $k = mNP + i$ and $l = nNP + j$ for $m, n = 0, 1$ and $i, j = 0, \ldots, NP-1$ 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, (5.29) can be expressed as a function of $\mathbf{Q}$ as

$$\hat{\mathbf{S}}_1 = \begin{bmatrix} \mathrm{diag}_{LP}(\mathbf{Q}_s) & \mathrm{diag}_{LP}(\mathbf{Q}_{sr}) \\ \mathrm{diag}_{LP}(\mathbf{Q}_{rs}) & \mathrm{diag}_{LP}(\mathbf{Q}_r) \end{bmatrix}. \tag{5.30}$$

Now we plug the ML estimates (5.22) and (5.30) into the likelihood ratio (5.17) to obtain the following expression:

$$\begin{aligned} \mathscr{G}^{\frac{1}{M}} &= \frac{\det\left(\hat{\mathbf{S}}_1\right)}{\det\left(\hat{\mathbf{S}}_0\right)} = \frac{\det\left(\mathrm{diag}_{LP}(\mathbf{Q}_s) - \mathrm{diag}_{LP}(\mathbf{Q}_{sr})\,\mathrm{diag}_{LP}(\mathbf{Q}_r)^{-1}\,\mathrm{diag}_{LP}(\mathbf{Q}_{rs})\right)}{\det\left(\mathrm{diag}_L(\mathbf{Q}_s)\right)} \\ &= \prod_{k=1}^{N}\det\left(\mathbf{D}_k - \mathbf{C}_k\mathbf{C}_k^H\right), \end{aligned} \tag{5.31}$$

where $\mathbf{D}$ and $\mathbf{C}$ are given by (5.20) and (5.21), 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 2×2 block matrix with invertible blocks. □

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 [79] 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., [121, 122].

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 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 $\mathcal{H}_0$ $\mathbf{z} \overset{N\to\infty}{\sim} \mathcal{CN}(\mathbf{0}, \mathbf{I}_{2LPN})$. 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 follows our work in Chapter 3 that allows us to decompose the distribution of the GLR under the null into a product of independent beta random variables as follows

Proposition 5.3.1. *The distribution of the GLRT statistic proposed in Theorem 5.3.1 under the null hypothesis is given by a product of independent Beta random variables that is*

$$\mathscr{G}^{1/M} \overset{\mathcal{D}}{=} \prod_{l=1}^{N} \left\{ \prod_{k=1}^{P} \prod_{m=1}^{L} U_{k,m} \right\} \left\{ \prod_{n=1}^{LP} V_n \right\} \tag{5.32}$$

where all U, V are independent Beta distributed random variables. Specifically, they are given by

$$U_{k,m} \sim Beta\left\{M - ((k-1)L + m - 1), (k-1)L\right\}, \tag{5.33}$$

$$V_n \sim Beta\left\{M - (LP + n - 1), LP\right\}. \tag{5.34}$$

Proof. The proof directly follows from Theorem 3.4.1 for the following parameters: The total number of blocks under the alternative is given by $L = N$, where each block is of dimension $\lambda_l = 2LP \ \forall l$. Now we can permute the blocks under the null such that each block under the alternative is tested against $n_l = P + 1 \ \forall l$ blocks under the null with $\kappa_k = \begin{cases} L, & k = 1, \ldots, P, \\ LP, & k = P + 1. \end{cases}$ □

Since both approaches only hold asymptotically, the finite-sample size effects will be studied in Section 5.7.

5.4 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 [76]: (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. [121, 123], 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 [72]. 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 $n = 0, \ldots, N-1$. Hence, the invariance group can be formulated as

$$\mathcal{G} = \{g : \mathbf{z} \rightarrow g(\mathbf{z}) = \mathbf{\Psi}\mathbf{z}\}, \tag{5.35}$$

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

$$\mathbf{P}_{\clubsuit} = \left(\sum_{k=1}^{N} \boldsymbol{\epsilon}_k \boldsymbol{\epsilon}_k^T \otimes \mathbf{V}_{\clubsuit}^{(k)} \otimes \mathbf{I}_L \right) (\mathbf{U} \otimes \mathbf{I}_{LP}), \tag{5.36}$$

$\boldsymbol{\epsilon}_k$ is the kth column of $\mathbf{I}_N$, $\mathbf{V}_{\clubsuit}^{(k)} \in \mathbb{V}$ 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 (5.36), 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 [72] to obtain the ratio of the maximal invariant densities under the two hypotheses, which is given by

$$\mathscr{L} = \frac{\int_{\mathcal{G}} p(g(\mathbf{z}); \mathcal{H}_1) |\det(\mathbf{J}_g)| dg}{\int_{\mathcal{G}} p(g(\mathbf{z}); \mathcal{H}_0) |\det(\mathbf{J}_g)| dg}, \tag{5.37}$$

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 $\mathcal{H}_i$. Moreover, $\mathbf{J}_g$ 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 work, Wijsman's theorem states that the ratio of the distributions of the maximal invariant statistic is given by

$$\mathscr{L} = \frac{\sum_{\mathbb{V}_0^N, \mathbb{V}_1^N, \mathbb{U}} \int_{\mathbb{G}} \int_{\mathbb{H}} \det(\mathbf{S}_1)^{-M} |\det(\mathbf{G})|^{2M} |\det(\mathbf{H})|^{2M} \exp\left\{-M \operatorname{tr}\left(\mathbf{\Psi Q \Psi}^H \mathbf{S}_1^{-1}\right)\right\} d\mathbf{G} d\mathbf{H}}{\sum_{\mathbb{V}_0^N, \mathbb{V}_1^N, \mathbb{U}} \int_{\mathbb{G}} \int_{\mathbb{H}} \det(\mathbf{S}_0)^{-M} |\det(\mathbf{G})|^{2M} |\det(\mathbf{H})|^{2M} \exp\left\{-M \operatorname{tr}\left(\mathbf{\Psi Q \Psi}^H \mathbf{S}_0^{-1}\right)\right\} d\mathbf{G} d\mathbf{H}}, \tag{5.38}$$

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 (5.38) that the UMPIT does not exist for this problem.

Lemma 5.4.1. *The ratio* (5.38) *can be simplified as*

$$\mathscr{L} \propto \sum_{\mathbb{V}_0^N, \mathbb{V}_1^N, \mathbb{U}} \int_{\mathbb{G}} \int_{\mathbb{H}} \beta(\mathbf{G}) \beta(\mathbf{H}) e^{-M[\alpha_1(\mathbf{G}) + \alpha_2(\mathbf{G}, \mathbf{H})]} d\mathbf{G} d\mathbf{H}, \tag{5.39}$$

with

$$\beta(\mathbf{A}) = |\det(\mathbf{A})| e^{-M \operatorname{tr}(\mathbf{A}\mathbf{A}^H)}, \tag{5.40}$$

$$\alpha_1(\mathbf{G}) = \sum_{k=1}^{N} \sum_{\substack{m,n=1 \\ m \neq n}}^{P} \operatorname{tr}\left(\mathbf{\Gamma}_k^{(m,n)} \mathbf{G}_k^{(n,n)} \mathbf{D}_k^{(n,m)} \mathbf{G}_k^{(m,m)H}\right), \tag{5.41}$$

$$\mathbf{\Gamma} = \mathbf{P}_s^T \operatorname{diag}_L\left(\mathbf{\Sigma}_1\right)^{-\frac{1}{2}} \mathbf{\Sigma}_1 \operatorname{diag}_L\left(\mathbf{\Sigma}_1\right)^{-\frac{1}{2}} \mathbf{P}_s, \tag{5.42}$$

$$\alpha_2(\mathbf{G}, \mathbf{H}) = \sum_{k=1}^{N} \operatorname{tr}\left(\mathbf{\Lambda}_k \mathbf{G}_k \mathbf{C}_k \mathbf{H}_k^H\right), \tag{5.43}$$

$$\mathbf{\Lambda} = \mathbf{P}_r^T \mathbf{\Sigma}_2^{-\frac{1}{2}} \mathbf{\Sigma}_{21} \operatorname{diag}_L\left(\mathbf{\Sigma}_1\right)^{-\frac{1}{2}} \mathbf{P}_r, \tag{5.44}$$

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

Proof. First we observe that the terms $\det(\mathbf{S}_0)^{-M}$ and $\det(\mathbf{S}_1)^{-M}$ in (5.38) 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 $\boldsymbol{\Psi}$ and $\mathbf{S}_0$ are block-diagonal with two blocks, the denominator can be simplified as

$$\operatorname{tr}\left(\boldsymbol{\Psi}\mathbf{Q}\boldsymbol{\Psi}^H\mathbf{S}_0^{-1}\right) = \operatorname{tr}\left(\mathbf{P}_s\mathbf{G}\mathbf{Q}_s\mathbf{G}^H\mathbf{P}_s^T\mathbf{S}_s^{-1}\right) + \operatorname{tr}\left(\mathbf{P}_r\mathbf{H}\mathbf{Q}_r\mathbf{H}^H\mathbf{P}_r^T\mathbf{S}_r^{-1}\right). \tag{5.45}$$

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 $\mathbf{B}_s = \operatorname{diag}_L(\mathbf{Q}_s)$ and $\mathbf{B}_r = \operatorname{diag}_{LP}(\mathbf{Q}_r)$, yields

$$\operatorname{tr}\left(\boldsymbol{\Psi}\mathbf{Q}\boldsymbol{\Psi}^H\mathbf{S}_0^{-1}\right) = \operatorname{tr}\left(\mathbf{P}_s\mathbf{G}\mathbf{B}_s^{-\frac{1}{2}}\mathbf{Q}_s\mathbf{B}_s^{-\frac{1}{2}}\mathbf{G}^H\mathbf{P}_s^T\mathbf{S}_s^{-1}\right) + \operatorname{tr}\left(\mathbf{P}_r\mathbf{H}\mathbf{B}_r^{-\frac{1}{2}}\mathbf{Q}_r\mathbf{B}_r^{-\frac{1}{2}}\mathbf{H}^H\mathbf{P}_r^T\mathbf{S}_r^{-1}\right). \tag{5.46}$$

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 (5.38).

Hence, $\mathscr{L}$ simplifies as

$$\mathscr{L} \propto \sum_{\mathbb{V}_0^N, \mathbb{V}_1^N, \mathbb{U}} \int_{\mathbb{G}} \int_{\mathbb{H}} |\det(\mathbf{G})|^{2M} |\det(\mathbf{H})|^{2M} \exp\left\{-M \operatorname{tr}\left(\boldsymbol{\Psi}\mathbf{Q}\boldsymbol{\Psi}^H\mathbf{S}_1^{-1}\right)\right\} d\mathbf{G}d\mathbf{H}. \tag{5.47}$$

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 $\boldsymbol{\Sigma} = \mathbf{S}_1^{-1}$ and $\bar{\boldsymbol{\Psi}} = \boldsymbol{\Psi}\mathbf{Q}\boldsymbol{\Psi}^H$, where these matrices can be partitioned as $\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_2 \end{bmatrix}$ and $\bar{\boldsymbol{\Psi}} = \begin{bmatrix} \bar{\boldsymbol{\Psi}}_1 & \bar{\boldsymbol{\Psi}}_{12} \\ \bar{\boldsymbol{\Psi}}_{21} & \bar{\boldsymbol{\Psi}}_2 \end{bmatrix}$. Note that each of the four LNP-sized blocks in $\boldsymbol{\Sigma}$ 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 $\bar{\boldsymbol{\Psi}}$ are given by $\bar{\boldsymbol{\Psi}}_1 = \mathbf{P}_s\mathbf{G}\mathbf{B}_s^{-\frac{1}{2}}\mathbf{Q}_s\mathbf{B}_s^{-\frac{1}{2}}\mathbf{G}^H\mathbf{P}_s^H$, $\bar{\boldsymbol{\Psi}}_2 = \mathbf{P}_r\mathbf{H}\mathbf{B}_r^{-\frac{1}{2}}\mathbf{Q}_r\mathbf{B}_r^{-\frac{1}{2}}\mathbf{H}^H\mathbf{P}_r^H$ and $\bar{\boldsymbol{\Psi}}_{12} = \bar{\boldsymbol{\Psi}}_{21}^H = \mathbf{P}_s\mathbf{G}\mathbf{B}_s^{-\frac{1}{2}}\mathbf{Q}_{sr}\mathbf{B}_r^{-\frac{1}{2}}\mathbf{H}^H\mathbf{P}_r^H$. Finally, putting things together, (5.47) becomes

$$\mathscr{L} \propto \sum_{\mathbb{V}_0^N, \mathbb{V}_1^N, \mathbb{U}} \int_{\mathbb{G}} \int_{\mathbb{H}} |\det(\mathbf{G})|^{2M} |\det(\mathbf{H})|^{2M} e^{-M \operatorname{tr}(\bar{\boldsymbol{\Psi}}_1\boldsymbol{\Sigma}_1)} e^{-2M \operatorname{tr}(\bar{\boldsymbol{\Psi}}_{12}\boldsymbol{\Sigma}_{21})} e^{-M \operatorname{tr}(\bar{\boldsymbol{\Psi}}_2\boldsymbol{\Sigma}_2)} d\mathbf{G}d\mathbf{H}. \tag{5.48}$$

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

$$\operatorname{tr}(\bar{\boldsymbol{\Psi}}_1 \boldsymbol{\Sigma}_1) = \operatorname{tr}\left(\mathbf{G}\mathbf{G}^H\right) + \sum_{k=1}^{N} \sum_{\substack{m,n=1 \\ m \neq n}}^{P} \operatorname{tr}\left(\boldsymbol{\Gamma}_k^{(m,n)} \mathbf{G}_k^{(n,n)} \mathbf{D}_k^{(n,m)} \mathbf{G}_k^{(m,m)H}\right), \tag{5.49}$$

where we also considered that both $\boldsymbol{\Gamma}$ and $\mathbf{D}$, given by (5.42) and (5.20), respectively, are whitened on their $L \times L$ main diagonal and $\mathbf{G} \in \mathbb{G}$. It should be noted that (5.49) depends on unknown parameters through $\boldsymbol{\Gamma}$.

The second exponential term in (5.48) can be reduced by introducing the change of variables $\mathbf{H} \to \mathbf{P}_r^T \boldsymbol{\Sigma}_2^{-\frac{1}{2}} \mathbf{P}_r \mathbf{H}$ as

$$\operatorname{tr}\left(\bar{\boldsymbol{\Psi}}_{12} \boldsymbol{\Sigma}_{21}\right) = \sum_{k=1}^{N} \operatorname{tr}\left(\boldsymbol{\Lambda}_k \mathbf{G}_k \mathbf{C}_k \mathbf{H}_k^H\right), \tag{5.50}$$

where $\boldsymbol{\Lambda}$ is given by (5.44).

Finally, by plugging in the previous change of variables, the last exponential term in (5.48) becomes

$$\begin{aligned} \operatorname{tr}\left(\bar{\boldsymbol{\Psi}}_2 \boldsymbol{\Sigma}_2\right) &= \operatorname{tr}\left(\boldsymbol{\Sigma}_2^{-\frac{1}{2}} \mathbf{P}_r \mathbf{H} \mathbf{B}_r^{-\frac{1}{2}} \mathbf{Q}_r \mathbf{B}_r^{-\frac{1}{2}} \mathbf{H}^H \mathbf{P}_r^T \boldsymbol{\Sigma}_2^{-\frac{1}{2}} \boldsymbol{\Sigma}_2\right) \\ &= \operatorname{tr}\left(\mathbf{H}\mathbf{H}^H\right), \end{aligned} \tag{5.51}$$

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 (5.49), (5.50), and (5.51) into (5.48). □

We should note that both $\boldsymbol{\Gamma}$ and $\boldsymbol{\Lambda}$ 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 (5.39) around $\alpha_1(\mathbf{G}) + \alpha_2(\mathbf{G}, \mathbf{H}) = 0$ as

$$\begin{aligned} e^{-M(\alpha_1(\mathbf{G})+\alpha_2(\mathbf{G},\mathbf{H}))} \approx 1 &- M(\alpha_1(\mathbf{G}) + \alpha_2(\mathbf{G}, \mathbf{H})) \\ &+ \frac{M^2}{2}\left[\alpha_1^2(\mathbf{G}) + 2\alpha_1(\mathbf{G})\alpha_2(\mathbf{G}, \mathbf{H}) + \alpha_2^2(\mathbf{G}, \mathbf{H})\right]. \end{aligned} \tag{5.52}$$

Thus, (5.39) can be approximated as

$$\mathscr{L} \propto \mathscr{L}_1 + \mathscr{L}_2 + \mathscr{L}_3 + \mathscr{L}_4 + \mathscr{L}_5, \tag{5.53}$$

where

$$\mathscr{L}_1 = -M \sum_{\mathbb{V}_0^N, \mathbb{V}_1^N, \mathbb{U}} \int_{\mathbb{G}} \beta(\mathbf{G}) \alpha_1(\mathbf{G}) d\mathbf{G} \int_{\mathbb{H}} \beta(\mathbf{H}) d\mathbf{H}, \tag{5.54}$$

$$\mathscr{L}_2 = -M \sum_{\mathbb{V}_0^N, \mathbb{V}_1^N, \mathbb{U}} \int_{\mathbb{G}} \int_{\mathbb{H}} \beta(\mathbf{G}) \beta(\mathbf{H}) \alpha_2(\mathbf{G}, \mathbf{H}) d\mathbf{G} d\mathbf{H}, \tag{5.55}$$

$$\mathscr{L}_3 = \frac{M^2}{2} \sum_{\mathbb{V}_0^N, \mathbb{V}_1^N, \mathbb{U}} \int_{\mathbb{G}} \beta(\mathbf{G}) \alpha_1^2(\mathbf{G}) d\mathbf{G} \int_{\mathbb{H}} \beta(\mathbf{H}) d\mathbf{H}, \tag{5.56}$$

$$\mathscr{L}_4 = \frac{M^2}{2} \sum_{\mathbb{V}_0^N, \mathbb{V}_1^N, \mathbb{U}} \int_{\mathbb{G}} \int_{\mathbb{H}} \beta(\mathbf{G}) \beta(\mathbf{H}) \alpha_1(\mathbf{G}) \alpha_2(\mathbf{G}, \mathbf{H}) d\mathbf{G} d\mathbf{H}, \tag{5.57}$$

$$\mathscr{L}_5 = \frac{M^2}{2} \sum_{\mathbb{V}_0^N, \mathbb{V}_1^N, \mathbb{U}} \int_{\mathbb{G}} \int_{\mathbb{H}} \beta(\mathbf{G}) \beta(\mathbf{H}) \alpha_2^2(\mathbf{G}, \mathbf{H}) d\mathbf{G} d\mathbf{H}. \tag{5.58}$$

Lemma 5.4.2. *The following terms are zero:*

$$\mathscr{L}_1 = 0, \tag{5.59}$$

$$\mathscr{L}_2 = 0, \tag{5.60}$$

$$\mathscr{L}_4 = 0. \tag{5.61}$$

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

$$\mathscr{L}_1 \propto \sum_{\mathbb{V}_0^N, \mathbb{V}_1^N, \mathbb{U}} \int_{\mathbb{G}} \int_{\mathbb{H}} \beta(\mathbf{G}) \beta(\mathbf{H}) \sum_{k=1}^{N} \sum_{\substack{m,n=1 \\ m \neq n}}^{P} \operatorname{tr}\left(\mathbf{\Gamma}_k^{(m,n)} \mathbf{G}_k^{(n,n)} \mathbf{D}_k^{(n,m)} \mathbf{G}_k^{(m,m)H} \right) d\mathbf{G} d\mathbf{H}. \tag{5.62}$$

Applying the change of variables $\mathbf{G}_k^{(n,n)} \to -\mathbf{G}_k^{(n,n)}$ 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 $\mathscr{L}_2$ and $\mathscr{L}_4$ are zero. □

Finally, the quadratic terms in $\alpha_1(\mathbf{G})$ and $\alpha_2(\mathbf{G}, \mathbf{H})$ remain in (5.53). In the following theorem we will show that these terms can be expressed as functions of the (cross) coherence matrices (5.20) and (5.21).

Theorem 5.4.1. *The ratio of the distribution of the maximal invariant statistic in* (5.38) *is*

$$\mathscr{L} \propto \mathscr{L}_S + \gamma \mathscr{L}_{SR}, \tag{5.63}$$

where

$$\mathscr{L}_S = \sum_{k=1}^{N} ||\mathbf{D}_k||^2 \tag{5.64}$$

and

$$\mathscr{L}_{SR} = \sum_{k=1}^{N} ||\mathbf{C}_k||^2 \tag{5.65}$$

with $\mathbf{D}$ *and* $\mathbf{C}$ *given by* (5.20) *and* (5.21)*, respectively. The parameter* γ *is a constant that depends on unknown parameters but is independent of the observations.*

Proof. Let us first focus on (5.56), which can be simplified as

$$\begin{aligned}\mathscr{L}_3 &= \tfrac{M^2}{2} \textstyle\sum_{\mathbb{V}_0^N,\mathbb{V}_1^N,\mathbb{U}} \int_{\mathbb{G}} \beta(\mathbf{G}) \left[\sum_{k=1}^{N} \sum_{\substack{m,n=1 \\ m\neq n}}^{P} \operatorname{tr}\left(\mathbf{\Gamma}_k^{(m,n)} \mathbf{G}_k^{(n,n)} \mathbf{D}_k^{(n,m)} \mathbf{G}_k^{(m,m)H}\right)\right]^2 d\mathbf{G} \int_{\mathbb{H}} \beta(\mathbf{H}) d\mathbf{H} \\ &\propto \textstyle\sum_{\mathbb{V}_0^N,\mathbb{V}_1^N,\mathbb{U}} \int_{\mathbb{G}} \beta(\mathbf{G}) \sum_{k=1}^{N} \sum_{\substack{m,n=1 \\ m\neq n}}^{P} \operatorname{tr}^2\left(\mathbf{\Gamma}_k^{(m,n)} \mathbf{G}_k^{(n,n)} \mathbf{D}_k^{(n,m)} \mathbf{G}_k^{(m,m)H}\right) d\mathbf{G}\end{aligned} \tag{5.66}$$

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 $\mathbf{G}_k^{(n,n)} \to -\mathbf{G}_k^{(n,n)}$. Now (5.66) becomes the same expression as in Appendix C in [38] and we can simplify it in the same way to obtain

$$\mathscr{L}_3 \propto \mathscr{L}_S = \sum_{k=1}^{N} ||\mathbf{D}_k||^2, \tag{5.67}$$

where $\mathbf{D}$ is given by (5.20). Secondly, we can reduce (5.58) as follows

$$\begin{aligned}\mathscr{L}_5 &= \tfrac{M^2}{2} \textstyle\sum_{\mathbb{V}_0^N,\mathbb{V}_1^N,\mathbb{U}} \int_{\mathbb{G}} \int_{\mathbb{H}} \beta(\mathbf{G})\beta(\mathbf{H}) \left[\sum_{i,j,l=1}^{P} \operatorname{tr}\left(\mathbf{\Lambda}_k^{(i,j)} \mathbf{G}_k^{(j,j)} \mathbf{C}_k^{(j,l)} \mathbf{H}_k^{(l,i)H}\right)\right]^2 d\mathbf{G} d\mathbf{H} \\ &\propto \textstyle\sum_{\mathbb{V}_0^N,\mathbb{V}_1^N,\mathbb{U}} \int_{\mathbb{G}} \int_{\mathbb{H}} \beta(\mathbf{G})\beta(\mathbf{H}) \sum_{i,j,l=1}^{P} \operatorname{tr}^2\left(\mathbf{\Lambda}_k^{(i,j)} \mathbf{G}_k^{(j,j)} \mathbf{C}_k^{(j,l)} \mathbf{H}_k^{(l,i)H}\right) d\mathbf{G} d\mathbf{H},\end{aligned} \tag{5.68}$$

where the cross-terms of the square cancel out by another change of variables $\mathbf{G}_k^{(j,j)} \to -\mathbf{G}_k^{(j,j)}$ or $\mathbf{H}_k^{(l,i)} \to -\mathbf{H}_k^{(l,i)}$, respectively. Finally, following similar steps as in Appendix C of [38], we obtain

$$\mathscr{L}_5 \propto \mathscr{L}_{SR} = \sum_{k=1}^{N} ||\mathbf{C}_k||^2. \tag{5.69}$$

It should be noted that $\mathscr{L}_3$ and $\mathscr{L}_5$ are equal to $\mathscr{L}_S$ and $\mathscr{L}_{SR}$ 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 γ, which allows us to express (5.53) as

$$\mathscr{L} \propto \mathscr{L}_S + \gamma \mathscr{L}_{SR}. \tag{5.70}$$

□

Since $\mathscr{L}$ still depends on unknown parameters via the constant γ, we can conclude that the LMPIT does not exist. We should note that the term $\mathscr{L}_S$ is the LMPIT for the single array CS detection problem [38]. After giving an interpretation of both $\mathscr{L}_S$ and $\mathscr{L}_{SR}$ in the following section, we will study the influence of the two terms on the detection performance as a function of γ in Section 5.6, which will finally show that an LMPIT-inspired detector can be suggested.

5.5 Interpretation of the test statistics

As can be seen in (5.19) and (5.63), 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 (5.20) and (5.21), respectively. Similarly to [38], 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 [75]

$$\mathbf{\Pi}_{\clubsuit\heartsuit}^{(l)}(\theta)\, d\theta = \mathrm{E}\left[d\boldsymbol{\xi}_{\clubsuit}(\theta)\, d\boldsymbol{\xi}_{\heartsuit}^{H}\left(\theta - \frac{2\pi}{P}l\right)\right], \tag{5.71}$$

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

$$\mathbf{u}_{\clubsuit}[n] = \int_{-\pi}^{\pi} e^{j\theta n} d\boldsymbol{\xi}_{\clubsuit}(\theta). \tag{5.72}$$

Furthermore, the cyclic (cross) PSD and the bi-frequency spectrum are related by [75]

$$\mathbf{S}_{\clubsuit\heartsuit}(\theta_i, \theta_j) = \sum_l \mathbf{\Pi}_{\clubsuit\heartsuit}^{(l)}(\theta_j)\delta\left(\theta_i - \theta_j - \frac{2\pi}{P}l\right) \in \mathbb{C}^{L\times L}. \tag{5.73}$$

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 5.2, 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 (5.15). For instance, the (i, j)th $L \times L$ sized block of the kth diagonal block of $\mathbf{S}_{sr}$ is given by

$$[\mathbf{S}_{sr}]_k^{(i,j)} = \mathbf{S}_{sr}(\theta_{iN+k}, \theta_{jN+k}) = \mathbf{\Pi}_{sr}^{(i-j)}(\theta_{jN+k}) \in \mathbb{C}^{L\times L}, \tag{5.74}$$

where $\theta_l = \frac{2\pi l}{NP}$, $k = 0, \ldots, N-1$, $i, j = 0, \ldots, P-1$, 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 (5.73) and (5.74) shows that the $L \times L$ diagonal blocks for $i = j$ correspond to the (cross) PSD on the stationary manifold for frequency θ_{jN+k}, and the off-diagonal blocks for $i \neq j$ correspond to the cyclic (cross) PSD at frequency θ_{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 5.5.1. *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*

$$\mathbf{D}^{(q)}(\theta_{jN+k}) = \left[\mathbf{\Pi}_{ss}^{(0)}\left(\theta_{jN+k} + \frac{2\pi}{P}q\right)\right]^{-\frac{1}{2}} \mathbf{\Pi}_{ss}^{(q)}(\theta_{jN+k}) \left[\mathbf{\Pi}_{ss}^{(0)}(\theta_{jN+k})\right]^{-\frac{1}{2}}, \tag{5.75}$$

and

$$\mathbf{C}^{(q)}(\theta_{jN+k}) = \left[\mathbf{\Pi}_{ss}^{(0)}\left(\theta_{jN+k} + \frac{2\pi}{P}q\right)\right]^{-\frac{1}{2}} \sum_{m=-j}^{P-1-j} \mathbf{\Pi}_{sr}^{(q-m)}\left(\theta_{mN+k} + \frac{2\pi}{P}j\right) \left[\mathbf{\Pi}_{rr}^{(m)}(\theta_{jN+k})\right]^{-\frac{1}{2}}, \tag{5.76}$$

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

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

$$\mathbf{D} = \operatorname{diag}_L(\mathbf{Q}_s)^{-1/2} \operatorname{diag}_{LP}(\mathbf{Q}_s) \operatorname{diag}_L(\mathbf{Q}_s)^{-1/2}, \tag{5.77}$$

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

$$\mathbf{D}_k^{(i,j)} = \left(\mathbf{Q}_{sk}^{(i,i)}\right)^{-1/2} \mathbf{Q}_{sk}^{(i,j)} \left(\mathbf{Q}_{sk}^{(j,j)}\right)^{-1/2}, \tag{5.78}$$

for $k = 0, \ldots, N-1$ and $i, j = 0, \ldots, P-1$. These elements of the sample covariance matrix can again be expressed as samples of the cyclic PSDs similar to (5.74). Hence, $\mathbf{D}_k^{(i,j)}$ can be written as a function of frequency θ_{iN+k} as

$$\mathbf{D}_k^{(i,j)} = \mathbf{D}^{(i-j)}(\theta_{jN+k}) = \left[\mathbf{\Pi}_{ss}^{(0)}\left(\theta_{jN+k} + \frac{2\pi}{P}q\right)\right]^{-\frac{1}{2}} \mathbf{\Pi}_{ss}^{(i-j)}(\theta_{jN+k}) \left[\mathbf{\Pi}_{ss}^{(0)}(\theta_{jN+k})\right]^{-\frac{1}{2}}, \tag{5.79}$$

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

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

$$\mathbf{C} = \operatorname{diag}_L(\mathbf{Q}_s)^{-1/2} \operatorname{diag}_{LP}(\mathbf{Q}_{sr}) \operatorname{diag}_{LP}(\mathbf{Q}_r)^{-1/2}. \tag{5.80}$$

can be written as

$$\mathbf{C}_k^{(i,j)} = \left(\mathbf{Q}_{sk}^{(i,i)}\right)^{-1/2} \sum_{l=0}^{P-1} \mathbf{Q}_{srk}^{(i,l)} \left(\mathbf{Q}_{rk}^{(l,j)}\right)^{-1/2}, \tag{5.81}$$

since the kth diagonal blocks of $\mathbf{Q}_{sr}$ and $\mathbf{Q}_r$ are both full matrices whereas the kth 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 θ_{jN+k} as

$$\mathbf{C}_k^{(i,j)} = \mathbf{C}_k^{(i-j)}(\theta_{jN+k}) = \left[\mathbf{\Pi}_{ss}^{(0)}(\theta_{jN+k} + \frac{2\pi}{P}q)\right]^{-\frac{1}{2}} \sum_{l=0}^{P-1} \mathbf{\Pi}_{sr}^{(i-l)}(\theta_{lN+k}) \left[\mathbf{\Pi}_{rr}^{(l-j)}(\theta_{jN+k})\right]^{-\frac{1}{2}}, \tag{5.82}$$

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

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-cylic 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 $m = -j, ..., P-1-j$. Note that the main diagonal blocks of $\mathbf{C}$ given by

$$\begin{aligned}\mathbf{C}^{(0)}(\theta_{jN+k}) = &\left[\mathbf{\Pi}_{ss}^{(0)}(\theta_{jN+k})\right]^{-\frac{1}{2}} \mathbf{\Pi}_{sr}^{(0)}(\theta_{jN+k}) \left[\mathbf{\Pi}_{rr}^{(0)}(\theta_{jN+k})\right]^{-\frac{1}{2}} \\ &+ \left[\mathbf{\Pi}_{ss}^{(0)}(\theta_{jN+k})\right]^{-\frac{1}{2}} \sum_{m=-j}^{P-1-j} \mathbf{\Pi}_{sr}^{(-m)}(\theta_{mN+k} + \frac{2\pi}{P}j) \left[\mathbf{\Pi}_{rr}^{(m)}(\theta_{jN+k})\right]^{-\frac{1}{2}},\end{aligned} \tag{5.83}$$

do not only account for the cyclic components but also for the usual cross-coherence between the WSS components at frequency θ_{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 $\mathscr{G}$ in (5.19) and the ratio $\mathscr{L}$ in (5.63), 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 $\mathscr{L}$ these terms are connected by the unknown parameter γ in (5.63). 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.

5.6 LMPIT-inspired detector

Since no LMPIT exists, we now analyze the influence of γ, i.e., the influence of the individual terms $\mathscr{L}_S$ and $\mathscr{L}_{SR}$ in (5.63) on the detection performance. As mentioned before, on the one hand the term $\mathscr{L}_S$ 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, $\mathscr{L}_{SR}$ measures the strength of cross-spectral correlation between SC and RC, i.e., it accounts for the inherent

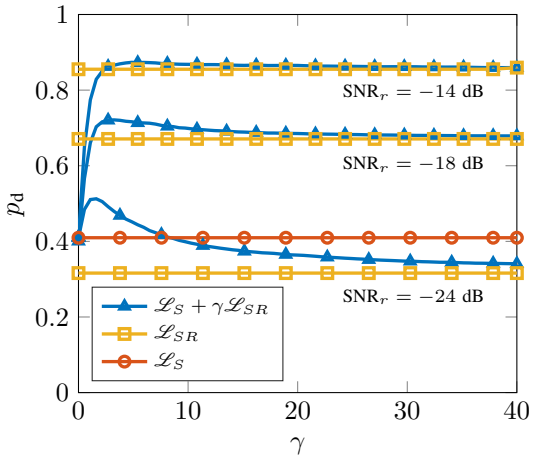

Figure 5.2: Probability of detection as a function of γ based on different detection statistics for an experiment with the following parameters: $P = 4$, $N = 128$, $M = 20$, $L = L_I = 2$, a rectangular pulse, $\text{SNR}_s = -18$ dB, $\text{SNR}_r = \{-14, -18, -24\}$ dB, and $p_{\text{fa}} = 0.01$.

correlation between SC and RC and also for the spectral correlation induced by cyclostationarity. For this reason it is expected that $\mathscr{L}_{SR}$ will have a bigger influence on the detection performance than $\mathscr{L}_S$ provided that the signals are not too weak.

Since the theoretical distribution of (5.63) is very difficult to obtain, we used Monte Carlo simulations to study the influence of γ. In order to do so, we used the signal model to be described in Section 5.7 to generate realizations under $\mathcal{H}_0$ and $\mathcal{H}_1$. For a given set of values for γ we obtained the probability of detection p_{d} based on the statistic $\mathscr{L}$ for a fixed probability of false alarm. Although we use $\mathscr{L}$ throughout this section as a benchmark for detectors based on the two individual terms $\mathscr{L}_S$ and $\mathscr{L}_{SR}$, note that in practice this is not possible as it depends on unknown parameters in γ. Additionally, we obtained the probability of detection based on using either $\mathscr{L}_S$ or $\mathscr{L}_{SR}$ individually.

The impact of the parameter γ on detection probability is shown in Figure 5.2. For a fixed $\text{SNR}_s = -18$ dB we obtained the detection probabilities for the detectors based on $\mathscr{L}_S$ and $\mathscr{L}_{SR}$ 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 $\mathscr{L}_S$, which is independent of 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 $\mathscr{L}_{SR}$ almost overlaps with that based on the optimal statistic $\mathscr{L}$. For lower SNR_r the correlation between signals at SC and RC is getting weaker and the right choice of γ becomes more critical for the best performance. At $\text{SNR}_r = -24$ dB, we observe that a detector based on $\mathscr{L}_S$ outperforms a detector based on $\mathscr{L}_{SR}$, i.e., better performance is obtained by simply detecting the presence of cyclostationarity at the SC. If the optimal γ were known, the performance of $\mathscr{L}$ could be reached. However, it should be noted that for such a low SNR_r even the optimal detector would not provide satisfactory performance.

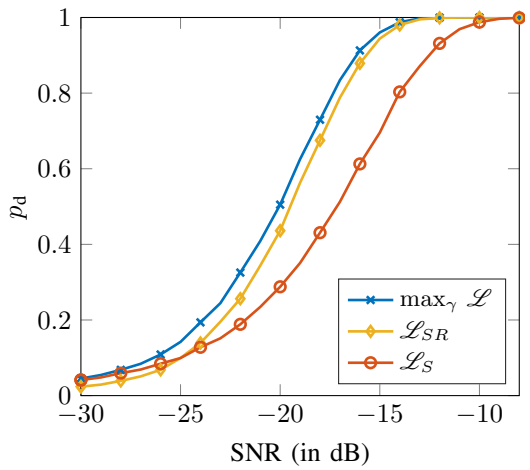

Figure 5.3: Probability of detection vs. SNR for various detectors, where $\mathrm{SNR}_s = \mathrm{SNR}_r = \mathrm{SNR}$ for an experiment with the following parameters: $P = 4$, $N = 128$, $M = 20$, $L = L_I = 2$, a rectangular pulse, and $p_{\mathrm{fa}} = 0.01$.

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 $\mathscr{L}$ (where γ has been determined by a brute-force search to maximize the probability of detection), $\mathscr{L}_{SR}$, and $\mathscr{L}_S$ in Figure 5.3. It can be observed that although there is a gap between the optimal p_d and the p_d of $\mathscr{L}_{SR}$, it is comparatively small and it decreases as the SNR increases. Moreover, the gap between $\mathscr{L}_{SR}$ and $\mathscr{L}_S$ 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 $\mathscr{L}_{SR}$ is close to $\mathscr{L}$ with the optimal γ obtained by brute-force search (which is not possible in practice).

Based on these considerations we propose

$$\mathscr{L}_{SR} = \sum_{k=1}^{N} ||\mathbf{C}_k||^2, \tag{5.84}$$

with $\mathbf{C}$ defined in (5.21), 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 $\mathcal{H}_0$ $\mathbf{z} \overset{N\to\infty}{\sim} \mathcal{CN}(\mathbf{0}, \mathbf{I}_{2LPN})$. 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.

5.7 Numerical results

In this section we evaluate the performance of the GLRT and the LMPIT-inspired test using Monte Carlo simulations. According to our model in (5.1) we generate the CS signal $\mathbf{s}[n]$ as a QPSK-signal with either a raised-cosine pulse with roll-off factor ρ 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

$$\mathrm{SNR}_{\clubsuit} = 10\log_{10}\left(\frac{\mathrm{tr}\left(\hat{\mathbf{R}}_{\clubsuit}\right)}{\mathrm{tr}\left(\hat{\mathbf{V}}_{\clubsuit}\right)}\right), \tag{5.85}$$

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

$$\hat{\mathbf{R}}_{\clubsuit} = \frac{1}{MNP}\sum_{n=0}^{MNP-1}\left(\mathbf{H}_{\clubsuit}[n] * \mathbf{s}[n]\right)\left(\mathbf{H}_{\clubsuit}[n] * \mathbf{s}[n]\right)^H \in \mathbb{C}^{L\times L} \tag{5.86}$$

$$\hat{\mathbf{V}}_{\clubsuit} = \frac{1}{MNP}\sum_{n=0}^{MNP-1}\mathbf{v}_{\clubsuit}[n]\mathbf{v}_{\clubsuit}^H[n] \in \mathbb{C}^{L\times L}. \tag{5.87}$$

Furthermore, we compare the proposed detectors with the following benchmark techniques: The first one is the correlated subspace detector proposed in [67], which employs the following test statistic

$$\mathscr{K} = \prod_{i=1}^{\min(L_I, L)}\frac{1}{1-k_i^2}\underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}}\eta, \tag{5.88}$$

where k_i is the ith sample canonical correlation between the SC and the RC. The second competitor is the multiantenna extension of the popular cross-correlation detector [58, 67] that uses the statistic

$$\mathscr{C} = |\,\mathrm{tr}\left(\mathbf{R}_{sr}^H\mathbf{R}_{sr}\right)| \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}}\eta, \tag{5.89}$$

where

$$\mathbf{R}_{sr} = \frac{1}{MNP} \sum_{n=0}^{MNP-1} \mathbf{u}_s[n]\mathbf{u}_r^H[n], \tag{5.90}$$

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.

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 5.4 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.

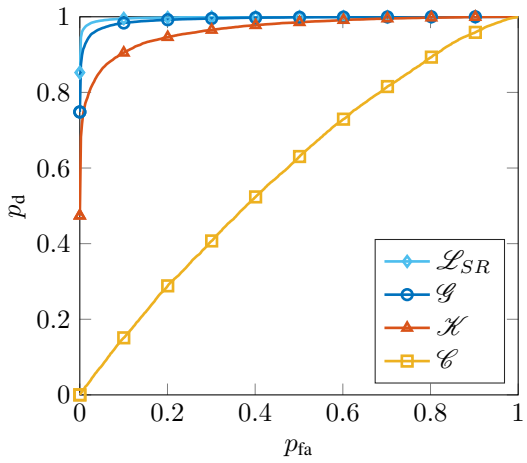

Figure 5.4: ROC curves in a scenario with $P = 2$, $N = 64$, $L = L_I = 4$, $M = 20$, a rectangular pulse, $\text{SNR}_s = -15$ dB and $\text{SNR}_r = -5$ dB.

Figure 5.5 depicts the probability of detection versus the 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 $p_{\text{fa}} = 0.01$.[4] Again we can observe that the proposed detectors outperform the competing techniques. In the 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 SNR_r is smallest for the LMPIT-inspired test and the GLRT whereas it is largest for the cross-correlation detector.

[4]Note that in a passive radar scenario p_{fa} would be a few orders of magnitude smaller.

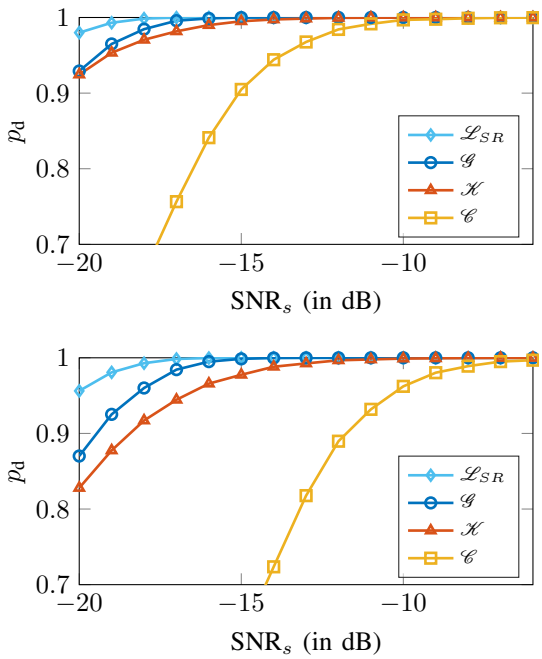

Figure 5.5: Probability of detection vs. SNR$_s$, where the top plot shows the results for SNR$_r$ = 0 dB and the bottom plot for SNR$_r$ = −5 dB for the following remaining parameters $P = 4$, $N = 128$, $L = L_I = 2$, $M = 20$, a rectangular pulse, and $p_{\text{fa}} = 0.01$.

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 [75]. Figure 5.6 shows the ROC for an SNR$_s$ = −15 dB at the SC and 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 SC components on the stationary manifold as can be seen in equations (5.76) and (5.83).

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 [88] for a related problem. Figure 5.7 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

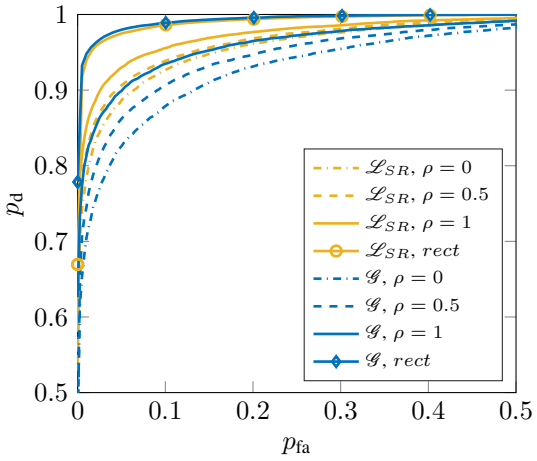

Figure 5.6: ROC curves for roll-off factors $\rho = \{0, 0.5, 1\}$ and a rectangular pulse shaping in a scenario with $P = 3$, $N = 128$, $L = L_I = 2$, $M = 20$, $\rho = 0.9$, and $\text{SNR}_s = \text{SNR}_r = -15$ dB.

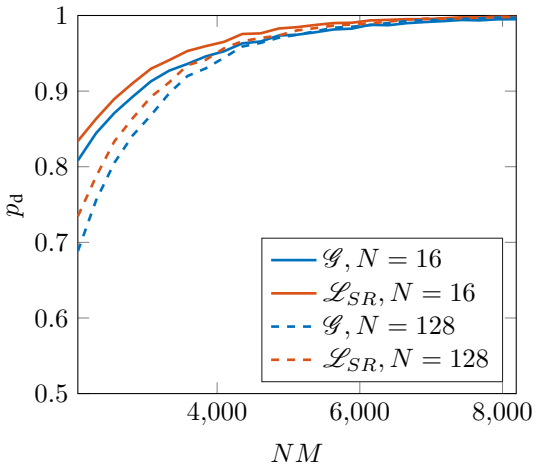

Figure 5.7: Probability of detection for $N = 16$ and $N = 128$ for different number of samples for $P = 2$, $L = L_I = 2$, a rectangular pulse, $\text{SNR}_s = -18$ dB and $\text{SNR}_r = -12$ dB.

rather small, we should 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 5.8 and 5.9 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 5.8) and $N = 128$ (Figure 5.9). 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

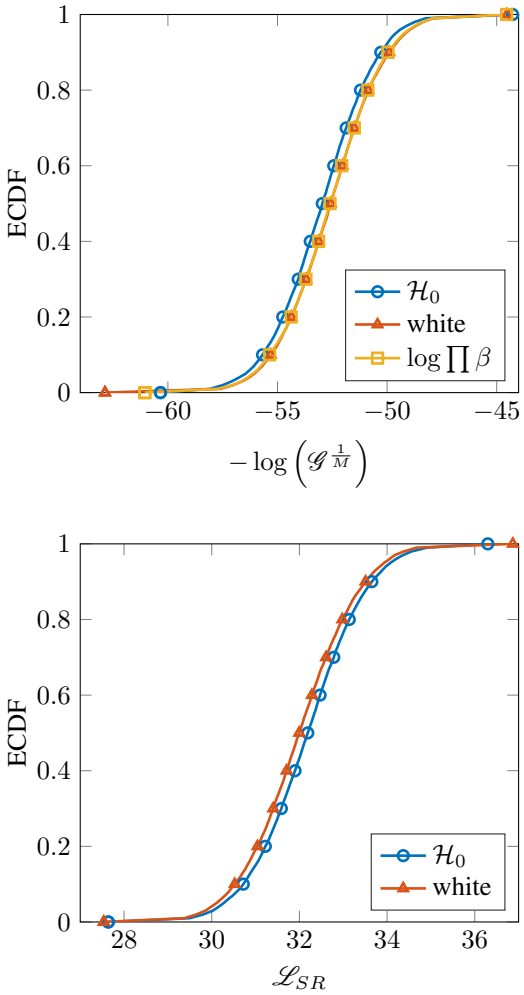

Figure 5.8: ECDF of the test statistics under $\mathcal{H}_0$ and for white noise for the GLRT (top) and the LMPIT-inspired test (bottom) for a scenario with $P = 2$, $N = 32$, $M = 16$, $L = L_I = 2$. The top figure also displays the approximation as a product of beta random variables.

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 5.8 and 5.9.

5.8 Conclusion

In this chapter 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 neither the UMPIT nor the LMPIT exists since the ratio of maximal invariant statistics depends on unknown parameters. Nevertheless, it is possible to propose an LMPIT-inspired detector based on this ratio. Both detectors, GLRT and LMPIT-inspired, are functions of a cyclic

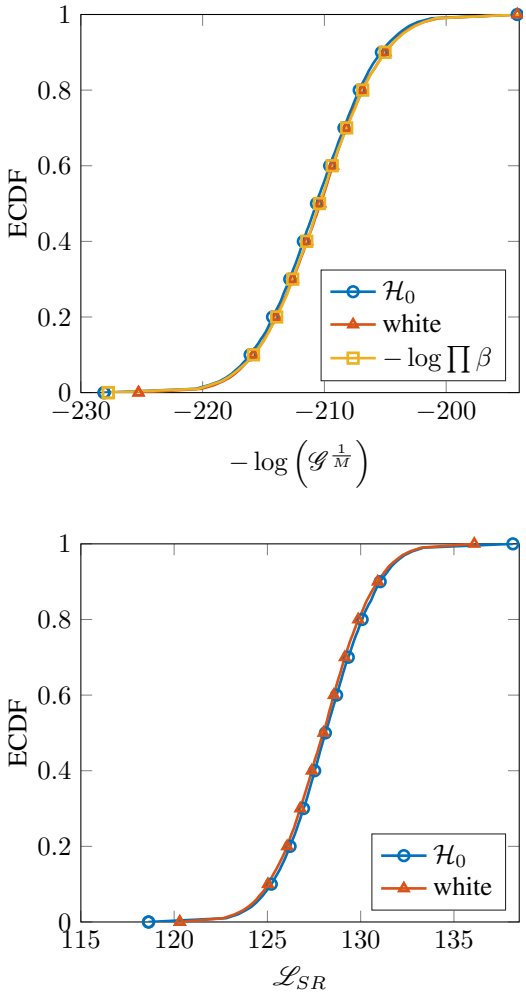

Figure 5.9: Same as Figure 5.8, except for $P = 2$, $N = 128$, $M = 16$, $L = L_I = 2$.

cross-coherence function, but only the GLRT accounts for the cyclic coherence at the SC. Finally, we derived the stochastic representations of the GLRT statistic under the null hypothesis that can be used to find a threshold to fix a desired probability of false alarm.

6 Two-channel passive detection of cyclostationary signals in noise with spatio-temporal structure

This chapter considers the two-channel passive detection of cyclostationarity in noise with spatio-temporal structure. In the previous chapter we considered the problem of two-channel passive detection of cyclostationary signals in temporally colored noise with spatial correlation. Here we specialize the noise model to different combinations of spatially correlated or uncorrelated noise that can be temporally white or colored. For the three combinations that have not been considered so far, we derive the GLRT and provide an insightful interpretation of the statistics. Moreover, we apply the results of Chapter 3 to propose stochastic representations for the three noise constellations under the null hypothesis. Monte Carlo simulations demonstrate that the proposed detectors outperform comparable state-of-the-art detectors and that it can be crucial to account for the appropriate noise model[1].

6.1 Introduction

The problem considered in this chapter is similar to the one presented in the previous one. We consider a passive bistatic radar system that consists of an illuminator of opportunity, a reference channel, and a surveillance channel. As previously mentioned, these systems are of special interest as they are cheap, simple and undetectable since the transmitter, which is the IO, is not part of the passive radar system itself [49]. Since, commonly, IOs are commercial video or audio broadcast systems or they could be space-based sources such as communication or navigation satellites [49], the transmitted signals are cyclostationary. Here we consider a MIMO system in which there are a reference antenna array that receives a noisy version of the direct-path signal transmitted by the IO, and a surveillance array. This array receives the target-path signal if a target is present, or noise only, otherwise. We assume that clutter or direct-path interference, which may corrupt the

[1] This chapter is based on the paper "Two-channel passive detection of cyclostationary signals in noise with spatio-temporal structure, S. Horstmann, D. Ramirez, P.J. Schreier, A. Pries *Proc. Asilomar Conference on Signals, Systems and Computers, 2019*"

signal received at the SC, is canceled by employing the techniques presented in, e.g., [54, 56, 57]. Furthermore, the SC and RC signals are obtained by spatial filtering [54, 114] or by employing directional antennas [113].

Commonly, this target detection problem is approached by cross-correlating SC and RC signals [49, 53, 63], which is, however, suboptimal due to noise at the RC [58]. In the past, various generalized likelihood ratio tests (GLRT) have been derived for the case of unknown deterministic waveforms in temporally and spatially white noise [51, 68]. Also, GLRTs considering the case of stochastic waveforms in white noise with various assumptions on the spatial correlation among each array and across reference and surveillance array have been derived in [64–67].

These aforementioned detectors assume that the signals are temporally white. However, digital communication signals transmitted by IOs are not temporally white but cyclostationary (CS) [22, 49]. We exploited this property in the previous chapter for which we derived the GLRT and proposed an LMPIT-inspired test. The noise considered in Chapter 5 is spatially correlated and temporally colored. Spatial correlation refers to the correlation at the individual arrays but not across reference and surveillance arrays. Assuming this kind of noise does not impose an additional structure to the detection problem that needs to be considered. For the single array case, detectors exploiting cyclostationarity and additional noise structure have been derived in [92, 117]. In [92] they derived the GLRTs and in [117] the same authors proposed an LMPIT-inspired test because no LMPITs exist for the problems. Considering that the problem does not have enough invariances for the single array case for the existence of an LMPIT, we do not show that no LMPIT (or UMPIT) exists for the problem considered in this chapter as this conclusion can already be made based on the observations in Chapter 5 and in [117].

In this chapter, we will see, however, that different noise models induce further structure that should be accounted for, or in other words, it should be exploited in the detector design. Specifically, if the antenna arrays are properly calibrated, the noise at each antenna can be considered temporally white and spatially uncorrelated. However, if the calibration fails, these assumptions may be violated. Hence, in this work, we consider noise with further structure: temporally white noise, which can be either spatially correlated or not, and temporally colored noise that is spatially uncorrelated.

We derive the GLRTs for these noise models and show that for all cases the GLRT can be factorized into the single channel GLRT [38, 117], which accounts for the presence of cyclostationarity at the SC, and a second term that accounts for the cross-cyclic correlations between observations at SC and RC. We show that exploiting cyclostationarity and, additionally, the structure of the noise increases the performance of the previously proposed detector for the general case of temporally colored and spatially correlated noise [73]. For the case that the noise structure is not known a priori, we may apply the tests proposed in [117] to specify the noise structure.

6.2 Problem formulation

As before, we consider a passive bistatic radar system, which consists of a reference array and a surveillance array. The received signal at the SC is denoted by $\mathbf{u}_s[n] \in \mathbb{C}^L$ and the received signal at the RC is given by $\mathbf{u}_r[n] \in \mathbb{C}^L$, where we assume again without loss of generality that there are L antennas at both arrays. As in Chapter 5, we consider the true pair of time-delay and Doppler shift, which allows us to synchronize the target-path signal with the direct-path signal, in our signal model, i.e., the test statistic derived in this chapter serves as an ambiguity score [67]. The problem considered here is therefore given by the two hypotheses

$$\begin{aligned} \mathcal{H}_0: \quad & \begin{cases} \mathbf{u}_s[n] = \mathbf{v}_s[n], \\ \mathbf{u}_r[n] = \mathbf{H}_r[n] * \mathbf{s}[n] + \mathbf{v}_r[n], \end{cases} \\ \mathcal{H}_1: \quad & \begin{cases} \mathbf{u}_s[n] = \mathbf{H}_s[n] * \mathbf{s}[n] + \mathbf{v}_s[n], \\ \mathbf{u}_r[n] = \mathbf{H}_r[n] * \mathbf{s}[n] + \mathbf{v}_r[n], \end{cases} \end{aligned} \tag{6.1}$$

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 frequency-selective channels between the IO and reference and surveillance arrays. The signal transmitted by the IO, which is equipped with L_I antennas, where $L_I \geq L$, is denoted by $\mathbf{s}[n] \in \mathbb{C}^{L_I}$. As in Chapter 5, $\mathbf{s}[n]$ is considered to be a discrete-time zero-mean second-order CS signal with known cycle period P. Finally, $\mathbf{v}_s[n] \in \mathbb{C}^L$ and $\mathbf{v}_r[n] \in \mathbb{C}^L$ are additive noise terms that are wide-sense stationary (WSS) and uncorrelated between RC and SC. However, within each array the noise may have further structure, which is either 1) *temporally white & spatially uncorrelated*, 2) *temporally white & spatially correlated*, or 3) *temporally colored & spatially uncorrelated*. In order to derive the GLRT we basically follow the derivations in the proof of Theorem 5.3.1 in Chapter 5 in which we have considered the case of temporally colored and spatially correlated noise.

Let us collect NP samples of each array into the vectors

$$\mathbf{y}_{\clubsuit} = \left[\mathbf{u}_{\clubsuit}^T[0], \ldots, \mathbf{u}_{\clubsuit}^T[NP-1]\right]^T, \tag{6.2}$$

for $\clubsuit \in \{s, r\}$. Before deriving the GLRT we will first investigate the structures of the (cross) covariance matrices $\mathbf{R}_{\clubsuit\heartsuit} = \mathrm{E}[\mathbf{y}_{\clubsuit}\mathbf{y}_{\heartsuit}^H] \in \mathbb{C}^{LNP \times LNP}$. Under both hypotheses the covariance matrix $\mathbf{R}_{rr}$ has the same structure. Considering that the signal $\mathbf{u}_r[n]$ is CS with cycle period P, we exploit a result from [3] that shows that the stack of P observations of a CS process is WSS, i.e. the

covariance matrix $\mathbf{R}_{rr}$ is a block-Toeplitz matrix with block size LP. Following the same argument it is easy to see that under $\mathcal{H}_1$ the matrices $\mathbf{R}_{ss}^{(1)}$ and $\mathbf{R}_{sr} = \mathbf{R}_{rs}^H$ are also block-Toeplitz matrices with block size LP.[2]

Considering the null hypothesis, we should first notice that the observations from SC and RC are uncorrelated since $\mathbf{v}_s[n]$ and $\mathbf{v}_r[n]$ are assumed to be uncorrelated. Secondly, the structure of $\mathbf{R}_{ss}^{(0)}$ depends on the specific noise model under consideration [92]. Specifically, it is given by

1) $\mathbf{R}_{ss}^{(0)} = \mathbf{I}_{NP} \otimes \mathbf{D}_L$, where $\mathbf{D}_L \in \mathbb{D}_L$ and $\mathbb{D}_L$ denotes the set of diagonal covariance matrices of dimension L,

2) $\mathbf{R}_{ss}^{(0)} = \mathbf{I}_{NP} \otimes \mathbf{A}_L$, where $\mathbf{A}_L \in \mathbb{A}_L$ and $\mathbb{A}_L$ denotes the set of positive semidefinite matrices of dimension L,

3) $\mathbf{R}_{ss}^{(0)}$ is block-Toeplitz with block size L and each block is diagonal.

Now we stack $\mathbf{y}_s$ and $\mathbf{y}_r$ into one long vector

$$\mathbf{y} = \left[\mathbf{y}_s^T, \mathbf{y}_r^T\right]^T . \tag{6.3}$$

Assuming $\mathbf{y}$ to be zero-mean proper complex Gaussian we can reformulate (6.1) as

$$\begin{aligned} \mathcal{H}_0 &: \mathbf{y} \sim \mathcal{CN}_{2LNP}(\mathbf{0}, \mathbf{R}_0), \\ \mathcal{H}_1 &: \mathbf{y} \sim \mathcal{CN}_{2LNP}(\mathbf{0}, \mathbf{R}_1), \end{aligned} \tag{6.4}$$

with

$$\mathbf{R}_0 = \mathrm{E}\left[\mathbf{y}\mathbf{y}^H | \mathcal{H}_0\right] = \begin{bmatrix} \mathbf{R}_{ss}^{(0)} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_{rr} \end{bmatrix}, \tag{6.5}$$

and

$$\mathbf{R}_1 = \mathrm{E}\left[\mathbf{y}\mathbf{y}^H | \mathcal{H}_1\right] = \begin{bmatrix} \mathbf{R}_{ss}^{(1)} & \mathbf{R}_{sr} \\ \mathbf{R}_{rs} & \mathbf{R}_{rr} \end{bmatrix}. \tag{6.6}$$

In order to derive the GLRT it would be necessary to obtain the maximum likelihood estimates (MLE) of $\mathbf{R}_0$ and $\mathbf{R}_1$. However, this would involve finding the MLE of a block-Toeplitz matrix for which there exists no closed-form solutions [124]. For this reason we make use of the approximation proposed in [38], where it is shown that the log-likelihood parameterized by a block-Toeplitz covariance matrix converge asymptotically ($N \to \infty$) to the log-likelihood parameterized by a block-circulant covariance matrix. Moreover, we exploit that a block-circulant covariance matrix can be block-diagonalized by the DFT matrix and a block-diagonal covariance matrix in turn has a closed-form MLE.

[2] Note that the superscripts (0) and (1) emphasize that it is the covariance matrix obtained under the null hypothesis or the alternative, respectively.

6.3 Derivation of the GLRT

In order to derive a closed-form expression for the (asymptotic) GLRT, we follow the approach in [38] and transform the samples from both arrays into the frequency domain as follows

$$\mathbf{z}_{\clubsuit} = (\mathbf{L}_{NP,N} \otimes \mathbf{I}_L)(\mathbf{F}_{NP} \otimes \mathbf{I}_L)^H \mathbf{y}_{\clubsuit} \in \mathbb{C}^{LNP}, \tag{6.7}$$

with commutation matrix $\mathbf{L}_{NP,N}$ and DFT matrix $\mathbf{F}_{NP}$ of size NP. Similar to the time domain we stack $\mathbf{z}_s$ and $\mathbf{z}_r$ into a vector

$$\mathbf{z} = [\mathbf{z}_s^T, \mathbf{z}_r^T]^T, \tag{6.8}$$

to reformulate the hypotheses (6.4) as

$$\begin{aligned} \mathcal{H}_0 &: \mathbf{z} \sim \mathcal{CN}_{2LNP}(\mathbf{0}, \mathbf{S}_0), \\ \mathcal{H}_1 &: \mathbf{z} \sim \mathcal{CN}_{2LNP}(\mathbf{0}, \mathbf{S}_1). \end{aligned} \tag{6.9}$$

Given $M \geq LP$ independent and identically distributed (i.i.d.) realizations of $\mathbf{z}$, the GLR is given by the ratio of the determinants of the MLEs of the covariance matrices [67]

$$\Lambda = \frac{p(\mathbf{z}_0, \cdots, \mathbf{z}_{M-1}; \hat{\mathbf{S}}_0)}{p(\mathbf{z}_0, \cdots, \mathbf{z}_{M-1}; \hat{\mathbf{S}}_1)} = \left(\frac{\det\left(\hat{\mathbf{S}}_1\right)}{\det\left(\hat{\mathbf{S}}_0\right)} \right)^M, \tag{6.10}$$

where $\hat{\mathbf{S}}_0$ and $\hat{\mathbf{S}}_1$ denote the MLEs of $\mathbf{S}_0$ and $\mathbf{S}_1$, respectively.

Similarly to the time-domain covariance matrices $\mathbf{R}_0$ and $\mathbf{R}_1$, we partition the frequency-domain covariance matrices into blocks as follows

$$\mathbf{S}_0 = \begin{bmatrix} \mathbf{S}_{ss}^{(0)} & \mathbf{0} \\ \mathbf{0} & \mathbf{S}_{rr} \end{bmatrix}, \tag{6.11}$$

and

$$\mathbf{S}_1 = \begin{bmatrix} \mathbf{S}_{ss}^{(1)} & \mathbf{S}_{sr} \\ \mathbf{S}_{rs} & \mathbf{S}_{rr} \end{bmatrix}, \tag{6.12}$$

where each block is given by the (cross) covariance matrix of $\mathbf{z}_s$ and $\mathbf{z}_r$, respectively. The permutation and transformation in (6.7) is designed such that the corresponding (cross) covariance matrices all have a block-diagonal structure. Specifically, $\mathbf{S}_{ss}^{(1)}$, $\mathbf{S}_{sr}$, and $\mathbf{S}_{rr}$ are block-diagonal with block size LP, and depending on the noise model $\mathbf{S}_{ss}^{(0)}$ is given by the following expressions as previously derived in [92]:

1) $\mathbf{S}_{ss}^{(0)} = \mathbf{I}_{NP} \otimes \tilde{\mathbf{D}}_L$, where $\tilde{\mathbf{D}}_L \in \mathbb{D}_L$,

2) $\mathbf{S}_{ss}^{(0)} = \mathbf{I}_{NP} \otimes \tilde{\mathbf{A}}_L$, where $\tilde{\mathbf{A}}_L \in \mathbb{A}_L$,

3) $\mathbf{S}_{ss}^{(0)}$ is diagonal.

6.3.1 MLEs of the covariance matrices $\mathbf{S}_0$ and $\mathbf{S}_1$

Given the sample covariance matrix of M i.i.d. samples of $\mathbf{z}$,

$$\mathbf{Q} = \frac{1}{M} \sum_{m=0}^{M-1} \mathbf{z}_m \mathbf{z}_m^H = \begin{bmatrix} \mathbf{Q}_{ss} & \mathbf{Q}_{sr} \\ \mathbf{Q}_{rs} & \mathbf{Q}_{rr} \end{bmatrix}, \tag{6.13}$$

the MLE of $\mathbf{S}_0$ may be easily obtained by considering its block-diagonal structure, i.e.

$$\hat{\mathbf{S}}_0 = \begin{bmatrix} \hat{\mathbf{S}}_{ss}^{(0)} & \mathbf{0} \\ \mathbf{0} & \operatorname{diag}_{LP}(\mathbf{Q}_{rr}) \end{bmatrix}, \tag{6.14}$$

where the noise structure-dependent MLEs, $\hat{\mathbf{S}}_{ss}^{(0)}$, were derived in [92] and are given by

1) $\hat{\mathbf{S}}_{ss}^{(0)} = \mathbf{I}_{NP} \otimes \left(\frac{1}{NP} \sum_{i=1}^{NP} \operatorname{diag}\left([\mathbf{Q}_{ss}]_i \right) \right)$,

2) $\hat{\mathbf{S}}_{ss}^{(0)} = \mathbf{I}_{NP} \otimes \left(\frac{1}{NP} \sum_{i=1}^{NP} [\mathbf{Q}_{ss}]_i \right)$,

3) $\hat{\mathbf{S}}_{ss}^{(0)} = \operatorname{diag}(\mathbf{Q}_{ss})$,

where $[\mathbf{Q}_{ss}]_i$ denotes the ith block of dimension L on the main diagonal of $\mathbf{Q}_{ss}$. The structure of $\mathbf{S}_1$ is given by a 2×2 block matrix, where each block itself is a block-diagonal matrix. Deriving the MLE of this matrix may seem to be more involved. However, it can be observed that it is always possible to permute the blocks in $\mathbf{S}_1$ such that another block-diagonal matrix with block size $2LP$ can be obtained. The MLE of this permuted matrix is again given by the block-diagonal matrix obtained from the corresponding blocks of the sample covariance matrix [73]. After applying the inverse permutation to the MLE, we end up with

$$\hat{\mathbf{S}}_1 = \begin{bmatrix} \operatorname{diag}_{LP}(\mathbf{Q}_{ss}) & \operatorname{diag}_{LP}(\mathbf{Q}_{sr}) \\ \operatorname{diag}_{LP}(\mathbf{Q}_{rs}) & \operatorname{diag}_{LP}(\mathbf{Q}_{rr}) \end{bmatrix}. \tag{6.15}$$

6.3.2 GLRT

Finally, we can plug in (6.14) and (6.15) into (6.10) to obtain

$$\Lambda^{\frac{1}{M}} = \det\left(\mathbf{D} - \mathbf{C}\mathbf{C}^H \right) \tag{6.16}$$

$$= \det(\mathbf{D}) \det\left(\mathbf{I} - \tilde{\mathbf{C}}\tilde{\mathbf{C}}^H \right), \tag{6.17}$$

where

$$\mathbf{D} = \left(\hat{\mathbf{S}}_{ss}^{(0)}\right)^{-1/2} \mathrm{diag}_{LP}\left(\mathbf{Q}_{ss}\right) \left(\hat{\mathbf{S}}_{ss}^{(0)}\right)^{-1/2}, \tag{6.18}$$

$$\mathbf{C} = \mathrm{diag}_{L}\left(\mathbf{Q}_{ss}\right)^{-1/2} \mathrm{diag}_{LP}\left(\mathbf{Q}_{sr}\right) \mathrm{diag}_{LP}\left(\mathbf{Q}_{rr}\right)^{-1/2}, \tag{6.19}$$

$$\tilde{\mathbf{C}} = \mathrm{diag}_{LP}\left(\mathbf{Q}_{ss}\right)^{-1/2} \mathrm{diag}_{LP}\left(\mathbf{Q}_{sr}\right) \mathrm{diag}_{LP}\left(\mathbf{Q}_{rr}\right)^{-1/2}, \tag{6.20}$$

and where we exploited the properties of the determinant of block matrices. Note that only the first term in (6.17) depends on the noise structure through $\hat{\mathbf{S}}_{ss}^{(0)}$. Furthermore, we observe that $\mathbf{D}$ is a coherence matrix that only depends on the correlations between observations at the SC, whereas $\tilde{\mathbf{C}}$ is a cross-coherence matrix that depends on the cross-correlations between SC and RC. After noticing that both $\mathbf{D}$ and $\tilde{\mathbf{C}}$ are block-diagonal matrices with block size LP, we may further simplify (6.17) as follows

$$\Lambda^{\frac{1}{M}} = \prod_{k=1}^{N} \prod_{l=1}^{LP} \lambda_k^{(l)} \left[1 - \left(\kappa_k^{(l)}\right)^2\right], \tag{6.21}$$

where $\lambda_k^{(l)}$ denotes the lth eigenvalue of the kth diagonal block of $\mathbf{D}$, which depends on the noise models 1)-3), and $\kappa_k^{(l)}$ denotes the lth eigenvalue of the kth block of $\tilde{\mathbf{C}}$. Finally, the GLRT is given by

$$\Lambda^{\frac{1}{M}} \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\gtrless}} \eta, \tag{6.22}$$

where η is selected to guarantee a given probability of false alarm.

It should be noted that $\lambda_k^{(l)}$ and $\kappa_k^{(l)}$ can be given an interpretation. Recall the linear transformation (6.7) that transforms the samples into the frequency domain and orders them such that every LP consecutive samples are those frequency components that are separated by multiples of the cycle frequency $2\pi/P$ [38]. These frequencies may be correlated if the signal is CS, and they are uncorrelated for a WSS process. It can be observed that $\lambda_k^{(l)}$ are the sample canonical correlations between frequency components of the SC observations separated by multiples of $2\pi/P$. Moreover, considering the products over these $\lambda_k^{(l)}$ individually, this product is the GLRT for detecting the presence of CS at a single channel (here the SC) [38, 92]. Similarly, $\kappa_k^{(l)}$ are the sample canonical correlations between frequency components of SC and RC observations.

6.3.3 Threshold selection and null distribution

In order to find the threshold for a given probability of false alarm p_{fa}, we propose two different approaches. The first exploits the invariances of the test statistic and the second uses the results of Chapter 3. To identify the invariances note that we can multiply $\mathbf{z}$ with any matrix of the structure of $\mathbf{S}_0$ without modifying the likelihood ratio, i.e., it is invariant to these transformations.

Hence, without loss of generality, we can assume that the observations under the null hypothesis are whitened. Thus, to find the threshold for any arbitrary process, we can run numerical simulations with temporally and spatially white processes.

For the second approach we use our results from Chapter 3 and propose the stochastic representations of the GLRTs for the three different cases in this paragraph. In order to apply the previous results, we have to split the GLRs for the temporally white cases once more into one factor being an independence test and a certain kind of sphericity test, which tests whether (block) diagonal elements of (block) matrices are identical. The latter term only consists of samples from the surveillance channel, which is not surprising considering that we can only test for temporally white observations based on the surveillance channel observations. First, we propose the stochastic representation for the temporally white and spatially uncorrelated case.

Proposition 6.3.1. *The distribution of the GLRT statistic* (6.17) *for* ***temporally white & spatially uncorrelated*** *noise (Case 1) under the null hypothesis is given by a product of independent Beta random variables that is*

$$\mathscr{G}^{1/M} \overset{\mathcal{D}}{=} \prod_{l=1}^{N} \left\{ \prod_{k=1}^{LP} U_k \right\} \left\{ \prod_{n=1}^{LP} V_n \right\} (NP)^{LNP} \prod_{j=2}^{NP} \left\{ \prod_{m=1}^{L} \left(1 - Z_{m,j-1}\right) \left(Z_{m,j-1}\right)^{j-1} \right\}, \tag{6.23}$$

where all U, V, Z are independent Beta distributed random variables. Specifically, they are given by

$$U_k \sim Beta\left\{M - ((k-1)LP), (k-1)LP\right\}, \tag{6.24}$$

$$V_n \sim Beta\left\{M - (LP + n - 1), LP\right\}. \tag{6.25}$$

$$Z_{m,j-1} \sim Beta\left\{M(j-1), M\right\}. \tag{6.26}$$

Proof. In order to exploit the results from Chapter 3 let us first rewrite the GLR (6.17). First we should note that the determinant is invariant to permutations and that $\hat{\mathbf{S}}_1$ can be permuted such that it is a block diagonal matrix with blocks of dimension $2LP$, which we will denote by $\mathrm{diag}_{2LP}\left(\tilde{\mathbf{Q}}\right)$. Furthermore, we expand the ratio by expanding with $\det\left(\mathrm{diag}\left(\mathbf{Q}_{ss}\right)\right)$. Now the ratio can be rewritten as follows

$$\mathscr{G}^{1/M} = \frac{\det\left(\hat{\mathbf{S}}_1\right)}{\det\left(\hat{\mathbf{S}}_{ss}^{(0)}\right) \det\left(\mathrm{diag}_{LP}\left(\mathbf{Q}_{rr}\right)\right)} \tag{6.27}$$

$$= \frac{\det\left(\mathrm{diag}_{2LP}\left(\tilde{\mathbf{Q}}\right)\right)}{\det\left(\mathrm{diag}\left(\mathbf{Q}_{ss}\right)\right) \det\left(\mathrm{diag}_{LP}\left(\mathbf{Q}_{rr}\right)\right)} \times \frac{\det\left(\mathrm{diag}\left(\mathbf{Q}_{ss}\right)\right)}{\det\left(\mathbf{I}_{NP} \otimes \left(\frac{1}{NP}\sum_{i=1}^{NP} \mathrm{diag}\left([\mathbf{Q}_{ss}]_i\right)\right)\right)}. \tag{6.28}$$

Similar to Appendix A in [103] it can be shown these two factors are statistically independent. Hence, we can come up with the distributions for each factor and can simply multiply those results. We can observe that the first term tests for block-diagonal matrices of different sizes and we can

directly apply Proposition 3.4.1 for the following parameters: The total number of blocks in the numerator is given by $L = N$, where each block is of dimension $\lambda_l = 2LP \; \forall l$. Now we can permute the blocks under the null such that each block under the alternative is tested against $n_l = LP + 1 \; \forall l$ blocks under the null with $\kappa_k = \begin{cases} 1, & k = 1, \ldots, LP, \\ LP, & k = LP + 1. \end{cases}$

In order to find the stochastic representation of the second term, we can directly apply Proposition 3.4.4. The test can be phrased as a test for equivalence of every L elements on the main diagonal elements of a matrix. The Proposition can be applied with the following parameters: The total number of elements under the alternative is given by LNP, where each block is of dimension $\lambda_l = 1$ for $l = 1, \ldots, LNP$. Under the null we have $\delta_k = NP$ and $\kappa_k = L$ for $k = 1, \ldots, NP$. □

In (6.28) we can observe that the GLR is divided into two independent tests. One of which is the test for independence of a set of scalars and vectors of dimension LP, which is identical to the GLR for the case of temporally colored and spatially correlated noise. The second term tests for the equivalence of sets of L scalars.

In the following proposition, we introduce the stochastic representation for the temporally white and spatially correlated noise case.

Proposition 6.3.2. *The distribution of the GLRT statistic* (6.17) *for* ***temporally white & spatially correlated*** *noise (Case 2) under the null hypothesis is given by a product of independent Beta random variables that is*

$$\mathscr{G}^{1/M} \overset{\mathcal{D}}{=} \prod_{l=1}^{N} \left\{ \prod_{k=1}^{P} \prod_{m=1}^{L} U_{k,m} \right\} \left\{ \prod_{n=1}^{LP} V_n \right\} \times (NP)^{LNP} \prod_{j=2}^{NP} \left\{ \prod_{m=1}^{L} \left(1 - Z_{m,j-1}\right) \left(Z_{m,j-1}\right)^{j-1} \right\} \left\{ \prod_{m=1}^{L} \left(W_{m,j}\right)^{j} \right\}, \quad (6.29)$$

where all U, V are independent Beta distributed random variables. Specifically, they are given by

$$U_{k,m} \sim Beta\left\{M - ((k-1)L + m - 1), (k-1)L\right\}, \quad (6.30)$$

$$V_n \sim Beta\left\{M - (LP + n - 1), LP\right\}, \quad (6.31)$$

$$Z_{m,j-1} \sim Beta\left\{M(j-1) - (m-1), M - (m-1)\right\}, \quad (6.32)$$

$$W_{m,j-1} \sim Beta\left\{Mj - 2(m-1), m - 1\right\}. \quad (6.33)$$

Proof. We first rewrite the GLR (6.17) by expanding it with $\det\left(\mathrm{diag}_L\left(\mathbf{Q}_{ss}\right)\right)$. We exploit, equivalently to the previous proof, that $\hat{\mathbf{S}}_1$ can be permuted such that it is a block diagonal matrix with blocks of dimension $2LP$ and its ML estimate is given by $\mathrm{diag}_{2LP}\left(\tilde{\mathbf{Q}}\right)$. We obtain the following expression for the GLR:

$$\mathscr{G}^{1/M} = \frac{\det\left(\hat{\mathbf{S}}_1\right)}{\det\left(\hat{\mathbf{S}}_{ss}^{(0)}\right)\det\left(\mathrm{diag}_{LP}\left(\mathbf{Q}_{rr}\right)\right)} \tag{6.34}$$

$$= \frac{\det\left(\mathrm{diag}_{2LP}\left(\tilde{\mathbf{Q}}\right)\right)}{\det\left(\mathrm{diag}_L\left(\mathbf{Q}_{ss}\right)\right)\det\left(\mathrm{diag}_{LP}\left(\mathbf{Q}_{rr}\right)\right)} \times \frac{\det\left(\mathrm{diag}_L\left(\mathbf{Q}_{ss}\right)\right)}{\det\left(\mathbf{I}_{NP}\otimes\left(\frac{1}{NP}\sum_{i=1}^{NP}\left[\mathbf{Q}_{ss}\right]_i\right)\right)}. \tag{6.35}$$

Again using the results of Appendix H.3 in [100], it can be shown that the two factors are independent. We can observe that the first term tests for block-diagonal matrices of different sizes and we can directly apply Proposition 3.4.1 for the following parameters: The total number of blocks under the alternative is given by $L = N$, where each block has dimension $\lambda_l = 2LP\ \forall l$. Since we can arbitrarily permute the elements in the denominator, the test can be formulated as follows. Each of the $2LP$-sized blocks is tested against LP one-dimensional elements, i.e., $\kappa_k = 1$ for $k = 1, \ldots, LNP$ and one block of dimension LP, i.e., $\kappa_k = LP$ for $k = LNP + 1, \ldots, LNP + N$.

In order to find the stochastic representation of the second term, we can apply Proposition 3.4.4. The test can be phrased as a test for equivalence of every L elements on the main diagonal elements of a matrix. The Proposition can be applied with the following parameters: The total number of elements under the alternative is given by $L = LNP$, where each block is of dimension $\lambda_l = 1$ for $l = 1, \ldots, LNP$. Under the null we have $\delta_k = NP$ and $\kappa_k = L$ for $k = 1, \ldots, NP$. □

In the proof of the proposition it can be observed that the GLR is once again splitted into two factors in (6.35). The first term is again an independence test. In this case it is the independence of a set of L and LP-sized vectors. The second term includes terms of the surveillance array only and tests the equivalence of block diagonal elements, i.e., it is again a special kind of sphericity test.

Finally, we present the stochastic representation for the case of temporally colored and spatially uncorrelated noise in the following proposition.

Proposition 6.3.3. *The distribution of the GLRT statistic* (6.17) *for* ***temporally colored & spatially uncorrelated*** *noise (Case 3) under the null hypothesis is given by a product of independent Beta random variables that is*

$$\mathscr{G}^{1/M} \overset{\mathcal{D}}{=} \prod_{l=1}^{N}\left\{\prod_{k=1}^{LP} U_k\right\}\left\{\prod_{n=1}^{LP} V_n\right\} \tag{6.36}$$

where all U, V are independent Beta distributed random variables. Specifically, they are given by

$$U_k \sim Beta\{M - ((k-1)LP), (k-1)LP\}, \tag{6.37}$$

$$V_n \sim Beta\{M - (LP + n - 1), LP\}. \tag{6.38}$$

Proof. The proof directly follows from Theorem 3.4.1 for the following parameters: The total number of blocks under the alternative is given by $L = N$, where each block is of dimension $\lambda_l = 2LP \ \forall l$. Now we can permute the blocks under the null such that each block under the alternative is tested against $n_l = LP + 1 \ \forall l$ blocks under the null with $\kappa_k = \begin{cases} 1, & k = 1, \ldots, LP, \\ LP, & k = LP + 1. \end{cases}$ Note that this expression is also identical to the first factor encountered in (6.28). □

The stochastic representations for the three considered problems can be used to determine a threshold to fix a desired probability of false alarm when applying the GLRT statistics.

6.4 Numerical results

We evaluate the performance of the GLRT for the different noise models using Monte Carlo simulations. According to our model in (6.1), we generate the CS signal $\mathbf{s}[n]$ as a QPSK-signal with rectangular pulse shaping. 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. The SNR at the surveillance array is set to $\mathrm{SNR}_s = -10$ dB and at the reference array it is set to $\mathrm{SNR}_s = -5$ dB. In each Monte Carlo simulation we draw new realizations of the channels. We consider simulation setups with the noise structures in Model 1) and Model 3), i.e., temporally white or colored, and spatially uncorrelated noise. Here, the noise terms are independent between SC and RC, white Gaussian, or colored Gaussian generated with a moving average filter of order 10 and uncorrelated among antennas. Moreover, in the figures presented in this section, 1) refers to the detector exploiting noise structure from Model 1, 2) and 3) refer to the detectors that exploits only temporal or spatial structure of the noise, respectively.

In Figure 6.1 we compare the receiver operating characteristic (ROC) curves of the three proposed detectors 1) - 3), which exploit various degrees of noise structure. Detector 2) only considers temporal structure, Detector 3) only considers spatial structure, whereas Detector 1) takes into account both. Since the noise in our example is temporally and spatially white, Detector 1) outperforms Detectors 2) and 3), but taking into account temporal structure is more advantageous.

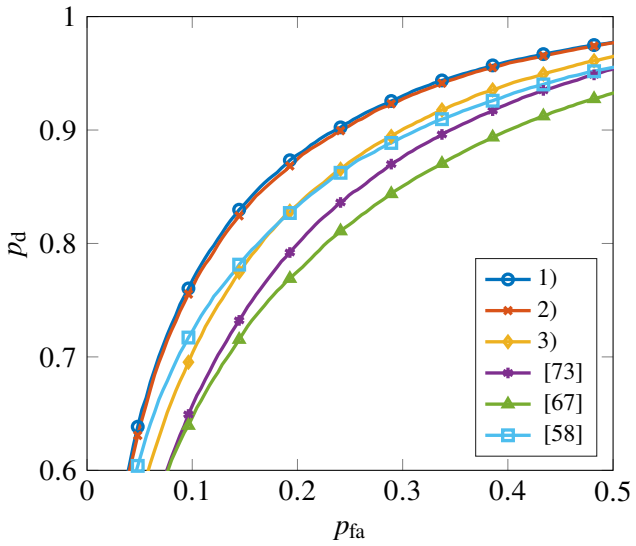

Figure 6.1: ROC for a scenario with temporally white and spatially uncorrelated noise (Model 1) for $P = 2$, $N = 16$, $L = L_I = 4$, $M = 32$, $\mathrm{SNR}_s = -10$ dB, and $\mathrm{SNR}_r = -5$ dB.

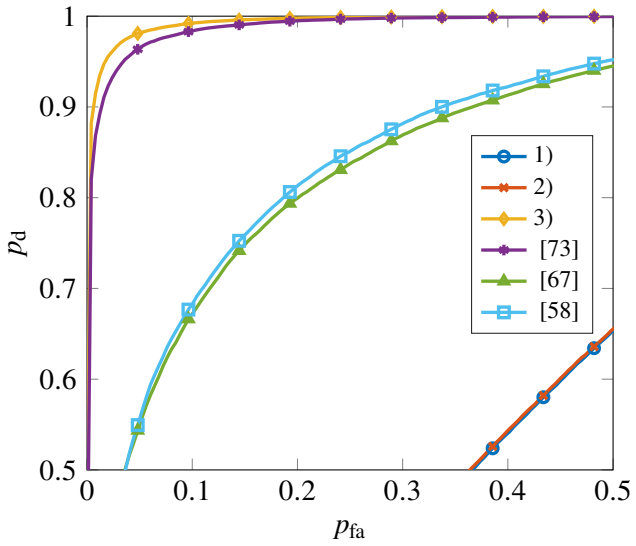

Figure 6.2: ROC for a scenario with temporally colored and spatially uncorrelated noise (Model 3) for $P = 2$, $N = 16$, $L = L_I = 4$, $M = 32$, $\mathrm{SNR}_s = -10$ dB, and $\mathrm{SNR}_r = -5$ dB.

Moreover, all detectors outperform competing approaches, namely, the GLRT from [73] that does not exploit the noise structure, the correlated subspace detector proposed in [67], and the cross-correlation detector [58].

Similar observations can be made in Figure 6.2, where we compare the ROCs of the various detectors for simulations under noise Model 3. It can be observed that Detector 3), which matches the model, performs best among the competing ones. Interestingly, applying the white noise detectors 1) and 2) decreases the performance substantially. The reason for this is that Detectors 1) and 2) underfit

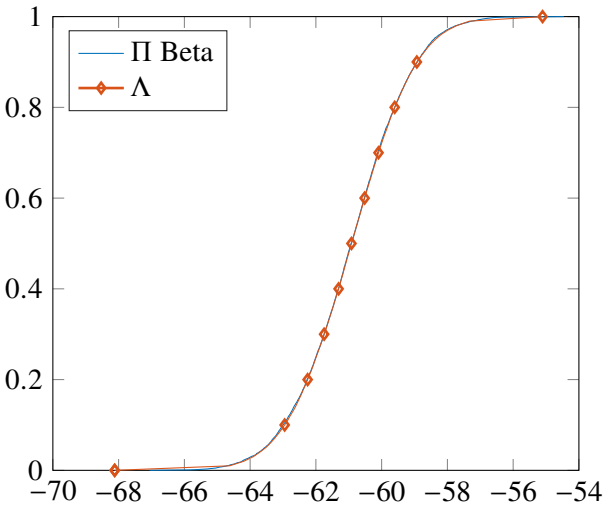

Figure 6.3: ECDF of the test statistics under $\mathcal{H}_0$ and the approximation as a product of beta random variables for a scenario with temporally white and spatially uncorrelated noise (Model 1) for $P = 2$, $N = 16$, $L = L_I = 4$, $M = 32$.

noise model 3), which has many more degrees of freedom as captured by 1) and 2). Hence, in general accounting for the temporal structure is essential. Note that the influence of the spatial structure is not as large as the related degrees of freedom depend on L, whereas the number of free parameters of the model corresponding to the temporal structure depends on N, which is generally much larger.

Finally, in Figure 6.3 we illustrate the ECDF of the GLRT under the null hypothesis for the case of temporally white and spatially uncorrelated noise (Model 1)) compared to the approximation with the product of independent Beta random variables as proposed in Section 6.3.3. As can be observed the approximation fits well to the distribution of the GLRT under the null hypothesis. This can be equivalently shown for Models 2) and 3).

6.5 Conclusion

In this chapter we have derived the GLRT statistics for the problem of two-channel passive detection of cyclostationary signals in the presence of three different noise models. It has been demonstrated that accounting for the proper noise model in the detector design is generally beneficial. Furthermore, it has turned out to be essential to incorporate the temporal structure of the noise for best detector performance. Similarly it should also be accounted for the spatial structure, however, its influence is less significant than the temporal structure since generally less degrees of freedom are determined

by the spatial structure compared to the number of degrees of freedom affected by the temporal structure. Finally, we have also proposed stochastic representations of all three GLRs under the null hypothesis that allow for fixing the probability of false alarm to a desired level.

7 Generalization to multistatic passive detection

In this chapter the two-channel detection problem is generalized to multiple channels. Specifically, we consider different numbers of surveillance and reference arrays for the detection of Gaussian cyclostationary signal in temporally colored and spatially correlated noise. Again, we derive the GLRT for this problem and investigate the existence of optimal invariant tests. The goal of this chapter is to analyze the influence of the number of surveillance and reference arrays on the detection performance. To this end, Monte Carlo simulations are performed to investigate the performance for various scenarios.

7.1 Introduction

We consider a MIMO passive multistatic radar system, that consists of one transmitter and multiple multichannel receivers. Other multistatic detectors may also consider multiple transmitters, here we consider a single illuminator of opportunity for the sake of simplicity. In the literature different multistatic detector systems have been reported: Hack et al. derive the GLRT for the problem considering unknown deterministic waveforms, where the author discusses benefits of centralized processing approaches and compares to similar detectors encountered in active MIMO radar and passive source localization problems [51]. In [70] the authors propose an ad-hoc detector based on the generalized coherence.

The previously proposed detectors do not exploit, however, that the signals transmitted by IOs are cyclostationary. Hence, we will derive the GLRT for the case of multiple receiver arrays. Yet in this work we limit the problem to a single IO.

7.1.1 Outline of this chapter

We will first formulate the detection problem in the following section. Afterwards we will derive the asymptotic GLRT for the problem in Section 7.3. This will be followed by analyzing the existence of optimal invariant detectors in Section 7.4. Note that the derivations made in Sections 7.3 and 7.4 mainly follow those from the previous chapter. In Section 7.5 we will present the numerical results

where we investigate the performance of the derived detectors with those derived in the previous chapter as well as comparable competing techniques. This chapter is concluded with a summary in Section 7.6.

7.2 Problem formulation

We consider a scenario with J surveillance antenna arrays and K reference antenna arrays. For the sake of simplicity, we assume again that all arrays are equipped with L antennas but the results presented here can be generalized to different numbers of antennas at each array. The IO is assumed to be equipped with L_I antennas. A noisy version of its transmission signal is received at the reference arrays. In the presence of a target, there will also be a noisy version of the IO signal observed at the surveillance arrays whereas there is only noise if no target echo is received. As already pointed out in detail in the previous chapter, we consider a simplified model in the sense that we assume that direct-path interference is canceled. The hypothesis test can be formulated as

$$\begin{aligned}
\mathcal{H}_0 : \quad & \begin{cases} \mathbf{u}_{s,1}[n] = \mathbf{v}_{s,1}[n], \\ \mathbf{u}_{s,2}[n] = \mathbf{v}_{s,2}[n], \\ \vdots \\ \mathbf{u}_{s,J}[n] = \mathbf{v}_{s,J}[n], \end{cases} \\
\mathcal{H}_1 : \quad & \begin{cases} \mathbf{u}_{s,1}[n] = \mathbf{H}_{s,1}[n] * \mathbf{s}[n] + \mathbf{v}_{s,1}[n], \\ \mathbf{u}_{s,2}[n] = \mathbf{H}_{s,2}[n] * \mathbf{s}[n] + \mathbf{v}_{s,2}[n], \\ \vdots \\ \mathbf{u}_{s,J}[n] = \mathbf{H}_{s,J}[n] * \mathbf{s}[n] + \mathbf{v}_{s,J}[n], \end{cases}
\end{aligned} \tag{7.1}$$

and reference signals $\mathbf{u}_{r,k}[n] = \mathbf{H}_{r,k}[n] * \mathbf{s}[n] + \mathbf{v}_{r,k}[n]$ for $k = 1, \ldots, K$ given both hypotheses. Furthermore, $\mathbf{H}_{s,j}[n] \in \mathbb{C}^{L \times L_I}$ and $\mathbf{H}_{r,k}[n] \in \mathbb{C}^{L \times L_I}$ represent the time-invariant frequency-selective channels from the IO to the reference and surveillance arrays, respectively. The additive noise terms $\mathbf{v}_{s,j}[n] \in \mathbb{C}^L$ for $j = 1, \ldots, J$ and $\mathbf{v}_{r,k}[n] \in \mathbb{C}^L$ for $k = 1, \ldots, K$ 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. As in Chapters 5 and 6, the IO transmission signal $\mathbf{s}[n] \in \mathbb{C}^{L_I}$ is assumed to be discrete-time zero-mean second-order CS with cycle period P. For this reason $\mathbf{u}_{r,k}[n] = \mathbf{H}_{r,k}[n] * \mathbf{s}[n] + \mathbf{v}_{r,k}[n]$ is also CS with cycle period P and so are $\mathbf{u}_{s,j}[n] = \mathbf{H}_{s,j}[n] * \mathbf{s}[n] + \mathbf{v}_{s,j}[n]$ under the alternative. Hence, our goal is to detect cyclostationarity at the surveillance arrays. Moreover, we assume that $L_I \geq L$, which implies that the cyclic (cross) power spectral densities (PSD) of $\mathbf{H}_{s,j}[n] * \mathbf{s}[n]$ and $\mathbf{H}_{r,k}[n] * \mathbf{s}[n]$ have full rank L.

	$\mathcal{H}_0$	$\mathcal{H}_1$
$\mathrm{E}\left[\mathbf{w}_{s,i}\mathbf{w}_{s,j}^H\right]$	$\begin{cases}\mathbb{T}_L^{LNP} \text{ for } i=j,\\ \mathbf{0} \text{ for } i\neq j,\end{cases}$	$\mathbb{T}_{LP}^{LNP}\ \forall\, i,j$
$\mathrm{E}\left[\mathbf{w}_{s,i}\mathbf{w}_{r,j}^H\right]$	$\mathbf{0}\ \forall\, i,j$	$\mathbb{T}_{LP}^{LNP}\ \forall\, i,j$
$\mathrm{E}\left[\mathbf{w}_{r,k}\mathbf{w}_{r,l}^H\right]$	$\mathbb{T}_{LP}^{LNP}\ \forall\, k,l$	

Table 7.1: Summary of (cross) covariance matrices of $\mathbf{w}_{s,j}$ and $\mathbf{w}_{r,k}$ under both hypotheses.

First, let us stack NP observations from each array into a vector

$$\mathbf{w}_{s,j} = \left[\mathbf{u}_{s,j}^T[0] \ \cdots \ \mathbf{u}_{s,j}^T[NP-1]\right]^T \in \mathbb{C}^{LNP}, \tag{7.2}$$

$$\mathbf{w}_{r,k} = \left[\mathbf{u}_{r,k}^T[0] \ \cdots \ \mathbf{u}_{r,k}^T[NP-1]\right]^T \in \mathbb{C}^{LNP}. \tag{7.3}$$

As exploited throughout this thesis, we will again make use of the fact that $\mathbf{w}_{s,j}$ and $\mathbf{w}_{r,k}$ are LP-variate WSS signals given that $\mathbf{u}_{s,j}[n]$ and $\mathbf{u}_{r,k}[n]$ are CS with cycle period P. In that case, their (cross) covariance matrix have block-Toeplitz structure with block size LP. Under the null hypothesis, i.e., when $\mathbf{u}_{s,j}[n]$ is an L-variate WSS signal rather than CS its (cross) covariance matrices have block-Toeplitz structure with block size L. A summary of these matrix structures is provided in Table 7.1. Similar to Section 5.2, we will next stack all observations from SC and RC into a long vector $\mathbf{w}$ as follows

$$\mathbf{w} = \left[\mathbf{w}_s^T \ \mathbf{w}_r^T\right]^T \in \mathbb{C}^{(J+K)LNP}, \tag{7.4}$$

where $\mathbf{w}_s = \left[\mathbf{w}_{s,1}^T \ \cdots \ \mathbf{w}_{s,J}^T\right]^T \in \mathbb{C}^{JLNP}$ and $\mathbf{w}_r = \left[\mathbf{w}_{r,1}^T \ \cdots \ \mathbf{w}_{r,K}^T\right]^T \in \mathbb{C}^{KLNP}$. Given the results from Table 7.1 it is now easy to determine the structure of the covariance matrix of $\mathbf{w}$ under both hypotheses. Under $\mathcal{H}_0$ the covariance matrix is given by

$$\mathbf{R}_0 = \mathrm{E}[\mathbf{w}\mathbf{w}^H|\mathcal{H}_0] = \begin{bmatrix} \mathbf{R}_s^{(0)} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_r \end{bmatrix}, \tag{7.5}$$

where $\mathbf{R}_s^{(0)}\,\mathrm{E}\left[\mathbf{w}_s\mathbf{w}_s^H|\mathcal{H}_0\right] \in \mathbb{T}_L^{JLNP}$ is a Hermitian block-Toeplitz matrix of dimension $JLNP \times JLNP$ with block size L and $\mathbf{R}_r = \mathrm{E}\left[\mathbf{w}_r\mathbf{w}_r^H\right] \in \mathbb{B}_K\mathbb{T}_{LP}^{LNP}$ is a $K \times K$ Hermitian block matrix, where each block is a block-Toeplitz matrix of dimension $LNP \times LNP$ with block size LP regardless of the hypothesis. Under the alternative we obtain the following structure

$$\mathbf{R}_1 = \mathrm{E}[\mathbf{w}\mathbf{w}^H|\mathcal{H}_1] = \begin{bmatrix} \mathbf{R}_s^{(1)} & \mathbf{R}_{sr} \\ \mathbf{R}_{rs} & \mathbf{R}_r \end{bmatrix}, \tag{7.6}$$

	$\mathcal{H}_0$	$\mathcal{H}_1$
$\mathrm{E}\left[\mathbf{z}_{s,i}\mathbf{z}_{s,j}^H\right]$	$\begin{cases}\mathbb{S}_L^{LNP} \text{ for } i=j,\\ \mathbf{0} \text{ for } i\neq j,\end{cases}$	$\mathbb{S}_{LP}^{LNP}\ \forall\, i,j$
$\mathrm{E}\left[\mathbf{z}_{s,i}\mathbf{z}_{r,j}^H\right]$	$\mathbf{0}\ \forall\, i,j$	$\mathbb{S}_{LP}^{LNP}\ \forall\, i,j$
$\mathrm{E}\left[\mathbf{z}_{r,k}\mathbf{z}_{r,l}^H\right]$	$\mathbb{S}_{LP}^{LNP}\ \forall\, k,l$	

Table 7.2: Summary of (cross) covariance matrices of $\mathbf{z}_{s,j}$ and $\mathbf{z}_{r,k}$ under both hypotheses.

where $\mathbf{R}_s^{(1)} = \mathrm{E}\left[\mathbf{w}_s\mathbf{w}_s^H|\mathcal{H}_1\right] \in \mathbb{B}_J\mathbb{T}_{LP}^{LNP}$ and $\mathbf{R}_{sr} = \mathbf{R}_{rs}^H = \mathrm{E}[\mathbf{w}_s\mathbf{w}_r^H|\mathcal{H}_1] \in \mathbb{B}_{J,K}\mathbb{T}_{LP}^{LNP}$. Hence, $\mathbf{R}_1$ is a $J+K$ block matrix, where each block is a block-Toeplitz matrix of dimension $LNP \times LNP$ with block size LP.

Assuming that $\mathbf{u}_{s,j}[n]$ for $j = 1,\ldots,J$ and $\mathbf{u}_{r,k}[n]$ for $k = 1,\ldots,K$ are zero-mean proper complex Gaussian random processes, we can formulate the hypothesis test as

$$\begin{aligned}\mathcal{H}_0 &: \mathbf{w} \sim \mathcal{CN}_{(J+K)LNP}(\mathbf{0}, \mathbf{R}_0),\\ \mathcal{H}_1 &: \mathbf{w} \sim \mathcal{CN}_{(J+K)LNP}(\mathbf{0}, \mathbf{R}_1).\end{aligned} \tag{7.7}$$

This is a composite hypothesis test since the two covariance matrices $\mathbf{R}_0$ and $\mathbf{R}_1$ are unknown. As it is already pointed out in the previous chapters, the (block) Toeplitz structure of the covariance matrices precludes the derivation of the aforementioned detectors since there is no closed-form ML estimate of the covariance matrices available. Hence, once more we exploit the results from [4] and approximate the block-Toeplitz matrices by block-circulant ones. In order to do so, we transform the observations into the frequency domain since the DFT matrix can be used to (block) diagonalize a (block) circulant matrix.

To this end, the observations are transformed in the following fashion:

$$\mathbf{z}_{s,j} = (\mathbf{L}_{NP,N} \otimes \mathbf{I}_L)(\mathbf{F}_{NP} \otimes \mathbf{I}_L)^H \mathbf{w}_{s,j}, \tag{7.8}$$

$$\mathbf{z}_{r,k} = (\mathbf{L}_{NP,N} \otimes \mathbf{I}_L)(\mathbf{F}_{NP} \otimes \mathbf{I}_L)^H \mathbf{w}_{r,k}. \tag{7.9}$$

Recall from equations (5.12) and (5.13) in Section 5.2 that $\mathbf{z}_{s,j}$ and $\mathbf{z}_{r,k}$ contain N blocks of P frequencies, which are separated by multiples of the cycle frequency $2\pi/P$. Given that the transformed signal is cyclostationary, these frequencies are correlated. Hence, under the null hypothesis the covariance matrices of $\mathbf{z}_{s,j}$ are block-diagonal with block size L, whereas the cross covariance matrices of $\mathbf{z}_{s,i}$ and $\mathbf{z}_{s,j}$ are zero. Under the alternative the (cross) covariance matrices of $\mathbf{z}_{s,i}$ and $\mathbf{z}_{s,j}$ are block-diagonal with block size LP and so are the (cross) covariance matrices of $\mathbf{z}_{r,k}$ and $\mathbf{z}_{r,l}$. The structures of the matrices are summarized in Table 7.2

Similar to the time-domain, we collect the frequency representations from all surveillance and all reference arrays in one long vector

$$\mathbf{z} = \left[\mathbf{z}_s^T \ \mathbf{z}_r^T\right]^T \in \mathbb{C}^{(J+K)LNP}, \tag{7.10}$$

where $\mathbf{z}_s = \left[\mathbf{z}_{s,1}^T \ \cdots \ \mathbf{z}_{s,J}^T\right]^T \in \mathbb{C}^{JLNP}$ and $\mathbf{z}_r = \left[\mathbf{z}_{r,1}^T \ \cdots \ \mathbf{z}_{r,K}^T\right]^T \in \mathbb{C}^{KLNP}$.

Given this as a basis, we can put all pieces together to establish the structure of the covariance matrix of $\mathbf{z}$ under both hypotheses. Under the null hypothesis we obtain

$$\mathbf{S}_0 = \mathrm{E}[\mathbf{z}\mathbf{z}^H|\mathcal{H}_0] = \begin{bmatrix} \mathbf{S}_s^{(0)} & \mathbf{0} \\ \mathbf{0} & \mathbf{S}_r \end{bmatrix}, \tag{7.11}$$

where the off-diagonal blocks are zero since observations at SC and RC are uncorrelated, i.e., $\mathbf{S}_s^{(0)} = \mathrm{E}\left[\mathbf{z}_s\mathbf{z}_s^H|\mathcal{H}_0\right] \in \mathbb{S}_L^{JLNP}$ is a block-diagonal Hermitian matrix with block size L and $\mathbf{S}_r = \mathrm{E}\left[\mathbf{z}_r\mathbf{z}_r^H\right] \in \mathbb{B}_K\mathbb{S}_{LP}^{LNP}$ is a $K \times K$ Hermitian block matrix, where each block of dimension $LNP \times LNP$ is a block-diagonal matrix with block size LP, which is identical under both hypotheses. Now, under $\mathcal{H}_1$, the covariance matrix is given by

$$\mathbf{S}_1 = \mathrm{E}[\mathbf{z}\mathbf{z}^H|\mathcal{H}_1] = \begin{bmatrix} \mathbf{S}_s^{(1)} & \mathbf{S}_{sr} \\ \mathbf{S}_{rs} & \mathbf{S}_r \end{bmatrix} \in \mathbb{B}_{(J+K)}\mathbb{S}_{LP}^{LNP}, \tag{7.12}$$

where $\mathbf{S}_s^{(1)} = \mathrm{E}\left[\mathbf{z}_s\mathbf{z}_s^H|\mathcal{H}_1\right] \in \mathbb{B}_J\mathbb{S}_{LP}^{LNP}$, $\mathbf{S}_{sr} = \mathrm{E}\left[\mathbf{z}_s\mathbf{z}_r^H|\mathcal{H}_1\right] \in \mathbb{B}_{J,K}\mathbb{S}_{LP}^{LNP}$, and $\mathbf{S}_r$ as before. Finally, we can reformulate the hypothesis test (7.7) asymptotically in the frequency domain as

$$\begin{aligned} \mathcal{H}_0 &: \mathbf{z} \sim \mathcal{CN}_{(J+K)LNP}(\mathbf{0}, \mathbf{S}_0), \\ \mathcal{H}_1 &: \mathbf{z} \sim \mathcal{CN}_{(J+K)LNP}(\mathbf{0}, \mathbf{S}_1). \end{aligned} \tag{7.13}$$

In the following Section we will derive the GLRT for this problem.

7.3 Derivation of the GLRT

The GLR is given by

$$\mathscr{G} = \frac{p(\mathbf{z}_0, \cdots, \mathbf{z}_{M-1}; \hat{\mathbf{S}}_0)}{p(\mathbf{z}_0, \cdots, \mathbf{z}_{M-1}; \hat{\mathbf{S}}_1)}, \tag{7.14}$$

where $\mathbf{z}_0, \ldots, \mathbf{z}_{M-1}$ denote M independent and identically distributed (i.i.d.) realizations 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

$$p(\mathbf{z}_0, \cdots, \mathbf{z}_{M-1}; \hat{\mathbf{S}}_i) = \frac{1}{\pi^{2(J+K)LNPM} \det\left(\hat{\mathbf{S}}_i\right)^M} \times \exp\left\{-M \operatorname{tr}\left(\mathbf{Q}\hat{\mathbf{S}}_i^{-1}\right)\right\}, \tag{7.15}$$

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

Theorem 7.3.1. *The GLR* (7.14) *is given by*

$$\mathscr{G}^{1/M} = \det\left(\mathbf{D} - \mathbf{C}\mathbf{C}^H\right), \tag{7.16}$$

$$= \det(\mathbf{D}) \det\left(\mathbf{I}_{KLNP} - \tilde{\mathbf{C}}^H \tilde{\mathbf{C}}\right), \tag{7.17}$$

where

$$\mathbf{D} = \left(\hat{\mathbf{S}}_s^{(0)}\right)^{-1/2} \hat{\mathbf{S}}_s^{(1)} \left(\hat{\mathbf{S}}_s^{(0)}\right)^{-1/2}, \tag{7.18}$$

$$\mathbf{C} = \left(\hat{\mathbf{S}}_s^{(0)}\right)^{-1/2} \hat{\mathbf{S}}_{sr} \left(\hat{\mathbf{S}}_r\right)^{-1/2}, \tag{7.19}$$

$$\tilde{\mathbf{C}} = \left(\hat{\mathbf{S}}_s^{(1)}\right)^{-1/2} \hat{\mathbf{S}}_{sr} \left(\hat{\mathbf{S}}_r\right)^{-1/2}, \tag{7.20}$$

with MLEs given by the following expressions

$$\hat{\mathbf{S}}_s^{(0)} = \operatorname{diag}_L(\mathbf{Q}_s), \tag{7.21}$$

$$\hat{\mathbf{S}}_s^{(1)} = \operatorname{block}_{J,J} \operatorname{diag}_{LP}(\mathbf{Q}_s), \tag{7.22}$$

$$\hat{\mathbf{S}}_{sr} = \operatorname{block}_{J,K} \operatorname{diag}_{LP}(\mathbf{Q}_s), \tag{7.23}$$

$$\hat{\mathbf{S}}_r = \operatorname{block}_{K,K} \operatorname{diag}_{LP}(\mathbf{Q}_s), \tag{7.24}$$

where the operation $\operatorname{block}_{L,M} \operatorname{diag}_N(\mathbf{A})$ *obtains an* $L \times M$ *block matrix from the matrix* $\mathbf{A}$*, where each block is a square block-diagonal matrix with block size* N *obtained from the respective block in* $\mathbf{A}$*, where* $\mathbf{A}$ *is of suitable dimension.*

Proof. Firstly, we derive the MLEs of the covariance matrices $\mathbf{S}_0$ and $\mathbf{S}_1$. In order to do so, we basically follow the lines of Theorem 5.3.1 in the previous chapter, where it is shown that the MLEs may be easily derived by noting that a permutation of the elements of $\mathbf{z}$ in a particular way leads to

block-diagonal matrices, which are easy to estimate. Specifically, we can reorder the elements of $\mathbf{w}_r$ such that the samples of each reference antenna array is stacked consecutively as follows

$$\tilde{\mathbf{w}}_r = \left[\mathbf{u}_{r,1}^T[0] \ \cdots \ \mathbf{u}_{r,K}^T[0], \mathbf{u}_{r,1}^T[1], \ \cdots, \ \mathbf{u}_{r,K}^T[1], \cdots, \ \mathbf{u}_{r,K}^T[NP-1]\right]^T. \tag{7.25}$$

Under the alternative hypothesis this permutation is required not only for the reference observations but also for the surveillance channel observations. Hence,

$$\tilde{\mathbf{w}}_r = (\mathbf{L}_{KNP,NP} \otimes \mathbf{I}_L)\mathbf{w}_r, \tag{7.26}$$

and similarly we permute the whole vector $\mathbf{w}$ under the alternative

$$\tilde{\mathbf{w}} = (\mathbf{L}_{(J+K)NP,NP} \otimes \mathbf{I}_L)\mathbf{w}. \tag{7.27}$$

The covariance matrix $\tilde{\mathbf{Q}}_r = \mathrm{E}[\tilde{\mathbf{w}}_r\tilde{\mathbf{w}}_r^H]$ is a block-diagonal matrix with block size KLP and similarly the covariance matrix $\tilde{\mathbf{Q}} = \mathrm{E}[\tilde{\mathbf{w}}\tilde{\mathbf{w}}^H]$ is a block diagonal matrix with block size $(J+K)LP$.

Similar to the two-channel case, we consider the following linear transformations

$$\tilde{\mathbf{z}}_r = \left[(\mathbf{L}_{NP,N}\mathbf{F}_{NP}^H \otimes \mathbf{I}_K) \otimes \mathbf{I}_L\right] \tilde{\mathbf{w}}_r, \tag{7.28}$$

$$\tilde{\mathbf{z}} = \left[(\mathbf{L}_{NP,N}\mathbf{F}_{NP}^H \otimes \mathbf{I}_{(J+K)}) \otimes \mathbf{I}_L\right] \tilde{\mathbf{w}}, \tag{7.29}$$

which lead to block-diagonal covariance matrices of $\tilde{\mathbf{z}}_r$ and $\tilde{\mathbf{z}}$ under both hypotheses. Next we exploit properties of the Kronecker product and equation (7.9) to obtain the following relation

$$\mathbf{z}_r = \left[(\mathbf{I}_K \otimes \mathbf{L}_{NP,N}\mathbf{F}_{NP}^H) \otimes \mathbf{I}_L\right] \mathbf{w}_r. \tag{7.30}$$

Similarly, we can express the linear transformation of the complete observation vector $\mathbf{w}$ as

$$\mathbf{z} = \left[(\mathbf{I}_{(J+K)} \otimes \mathbf{L}_{NP,N}\mathbf{F}_{NP}^H) \otimes \mathbf{I}_L\right] \mathbf{w}. \tag{7.31}$$

If we now plug back in equations (7.26) and (7.27), we can solve for $\mathbf{z}_r$ and $\mathbf{z}$. Putting all pieces together, we end up with the following linear transformations

$$\tilde{\mathbf{z}}_r = \underbrace{(\mathbf{L}_{KNP,NP} \otimes \mathbf{I}_L)}_{\mathbf{T}_r} \mathbf{z}_r, \tag{7.32}$$

$$\tilde{\mathbf{z}} = \underbrace{(\mathbf{L}_{(J+K)NP,NP} \otimes \mathbf{I}_L)}_{\mathbf{T}} \mathbf{z}, \tag{7.33}$$

where we exploited that $\mathbf{L}_{KNP,NP}$ and $\mathbf{L}_{(J+K)NP,NP}$ commute the Kronecker products $(\mathbf{I}_K \otimes \mathbf{L}_{NP,N}\mathbf{F}_{NP}^H)$ and $(\mathbf{I}_{(J+K)} \otimes \mathbf{L}_{NP,N}\mathbf{F}_{NP}^H)$, respectively. Comparing equations (7.32) and (7.33) with (7.26) and (7.27) shows that $\tilde{\mathbf{w}}_r$ and $\tilde{\mathbf{w}}$ are related by the same permutation matrices $\mathbf{T}_r$ and $\mathbf{T}$ as $\tilde{\mathbf{z}}_r$ and $\tilde{\mathbf{z}}$.

Given the permutations (7.32) and (7.33) we can now easily obtain the MLEs $\mathbf{S}_r$ and $\mathbf{S}_1$ by

$$\hat{\mathbf{S}}_r = \mathbf{T}_r^T \operatorname{diag}_{KLP}\left(\tilde{\mathbf{Q}}_r\right)\mathbf{T}_r, \tag{7.34}$$

$$\hat{\mathbf{S}}_1 = \mathbf{T}^T \operatorname{diag}_{(J+K)LP}\left(\tilde{\mathbf{Q}}\right)\mathbf{T}, \tag{7.35}$$

which is equivalent to

$$\hat{\mathbf{S}}_r = \operatorname{block}_{K,K} \operatorname{diag}_{LP}(\mathbf{Q}_r), \tag{7.36}$$

$$\hat{\mathbf{S}}_1 = \operatorname{block}_{(J+K),(J+K)} \operatorname{diag}_{LP}(\mathbf{Q}). \tag{7.37}$$

Finally, the MLE of $\mathbf{S}_0$ is given by

$$\hat{\mathbf{S}}_0 = \begin{bmatrix} \operatorname{diag}_L(\mathbf{Q}_s) & \mathbf{0} \\ \mathbf{0} & \hat{\mathbf{S}}_r \end{bmatrix}. \tag{7.38}$$

Now we plug the MLEs into the likelihood ratio (7.14) to obtain the following expressions

$$\mathscr{G}^{1/M} = \frac{\det\left(\hat{\mathbf{S}}_1\right)}{\det\left(\hat{\mathbf{S}}_0\right)}, \tag{7.39}$$

in which we can replace $\det\left(\hat{\mathbf{S}}_1\right)$ in two different ways noting that the determinant of a 2×2 block matrix can be formulated in two different ways

$$\det\left(\hat{\mathbf{S}}_1\right) = \det\left(\hat{\mathbf{S}}_r\right)\det\left(\hat{\mathbf{S}}_s^{(1)} - \hat{\mathbf{S}}_{rs}\hat{\mathbf{S}}_r^{-1}\hat{\mathbf{S}}_{sr}\right) = \det\left(\hat{\mathbf{S}}_s^{(1)}\right)\det\left(\hat{\mathbf{S}}_r - \hat{\mathbf{S}}_{sr}\left(\hat{\mathbf{S}}_s^{(1)}\right)^{-1}\hat{\mathbf{S}}_{rs}\right). \tag{7.40}$$

Considering that the denominator in (7.39) is given by $\hat{\mathbf{S}}_0 = \det\left(\hat{\mathbf{S}}_s^{(0)}\right)\det\left(\hat{\mathbf{S}}_r\right)$, we can obtain the following two expressions for (7.39)

$$\mathscr{G}^{1/M} = \det\left(\mathbf{D} - \mathbf{C}\mathbf{C}^H\right), \tag{7.41}$$

$$= \det(\mathbf{D})\det\left(\mathbf{I}_{KLNP} - \tilde{\mathbf{C}}^H\tilde{\mathbf{C}}\right), \tag{7.42}$$

where

$$\mathbf{D} = \left(\hat{\mathbf{S}}_s^{(0)}\right)^{-1/2} \hat{\mathbf{S}}_s^{(1)} \left(\hat{\mathbf{S}}_s^{(0)}\right)^{-1/2}, \tag{7.43}$$

$$\mathbf{C} = \left(\hat{\mathbf{S}}_s^{(0)}\right)^{-1/2} \hat{\mathbf{S}}_{sr} \left(\hat{\mathbf{S}}_r\right)^{-1/2}, \tag{7.44}$$

$$\tilde{\mathbf{C}} = \left(\hat{\mathbf{S}}_s^{(1)}\right)^{-1/2} \hat{\mathbf{S}}_{sr} \left(\hat{\mathbf{S}}_r\right)^{-1/2}. \tag{7.45}$$

□

7.3.1 Interpretation of the cross-coherence matrices

Similar to Section 5.5 we will provide a brief interpretation of the test statistics composed by the (cross) coherence matrices (7.18), (7.19), and (7.20) in this paragraph. In order to get an insight into the test statistics let us first consider the (cross) covariance matrices of each channel pair. To this end, we split the sample covariance matrix $\mathbf{Q}$ into $LNP \times LNP$ sized blocks that represent the sample covariance matrices of each array and channel combination. We will denote them as follows

$$\mathbf{Q}_s^{(i,j)} = \frac{1}{M} \sum_{m=0}^{M-1} \mathbf{z}_{s,i}^{(m)} \mathbf{z}_{s,j}^{(m)}, \tag{7.46}$$

$$\mathbf{Q}_{sr}^{(j,k)} = \frac{1}{M} \sum_{m=0}^{M-1} \mathbf{z}_{s,j}^{(m)} \mathbf{z}_{r,k}^{(m)}, \tag{7.47}$$

$$\mathbf{Q}_r^{(k,l)} = \frac{1}{M} \sum_{m=0}^{M-1} \mathbf{z}_{r,k}^{(m)} \mathbf{z}_{r,l}^{(m)}, \tag{7.48}$$

where $i, j = 1, \ldots, J$, $k, l = 1, \ldots, K$, and $\mathbf{z}_{\clubsuit,i}^{(m)}$ for $\clubsuit \in \{s, r\}$ denotes the mth observation at array $\clubsuit$ in channel i. Now we can rewrite each $LNP \times LNP$ sized block of the MLEs in (7.21) - (7.24) as

$$\hat{\mathbf{S}}_{s_i s_i}^{(0)} = \mathrm{diag}_L \left(\mathbf{Q}_s^{(i,i)}\right), \tag{7.49}$$

$$\hat{\mathbf{S}}_{s_i s_j}^{(1)} = \mathrm{diag}_{LP} \left(\mathbf{Q}_s^{(i,j)}\right), \tag{7.50}$$

$$\hat{\mathbf{S}}_{s_j r_k} = \mathrm{diag}_{LP} \left(\mathbf{Q}_{sr}^{(j,k)}\right), \tag{7.51}$$

$$\hat{\mathbf{S}}_{r_l r_l} = \mathrm{diag}_{LP} \left(\mathbf{Q}_r^{(l,k)}\right), \tag{7.52}$$

for $i, j = 1, \ldots, J$, $k, l = 1, \ldots, K$. Now we can exploit the results from Section 5.5 and explicitly Equation (5.74) in order to rewrite the blocks of the MLEs as functions of the cyclic (cross) PSDs. The (m, n)th $L \times L$ sized block of the qth diagonal block of $\hat{\mathbf{S}}^{(1)}_{s_i s_j}$, $\hat{\mathbf{S}}_{s_j r_k}$, and $\hat{\mathbf{S}}_{r_l r_l}$ are given by

$$\left[\hat{\mathbf{S}}^{(1)}_{s_i s_j}\right]^{(m,n)}_k = \hat{\mathbf{S}}^{(1)}_{s_i s_j}(\theta_{mN+q}, \theta_{nN+q}) = \mathbf{\Pi}^{(m-n)}_{s_i s_j}(\theta_{nN+q}) \in \mathbb{C}^{L\times L}, \tag{7.53}$$

$$\left[\hat{\mathbf{S}}_{s_j r_k}\right]^{(m,n)}_k = \hat{\mathbf{S}}_{s_j r_k}(\theta_{mN+q}, \theta_{nN+q}) = \mathbf{\Pi}^{(m-n)}_{s_j r_k}(\theta_{nN+q}) \in \mathbb{C}^{L\times L}, \tag{7.54}$$

$$\left[\hat{\mathbf{S}}_{r_l r_l}\right]^{(m,n)}_k = \hat{\mathbf{S}}_{r_l r_l}(\theta_{mN+q}, \theta_{nN+q}) = \mathbf{\Pi}^{(m-n)}_{r_l r_l}(\theta_{nN+q}) \in \mathbb{C}^{L\times L}, \tag{7.55}$$

where $\theta_l = \frac{2\pi l}{NP}$, $k = 0, \ldots, N-1$, $m, n = 0, \ldots, P-1$. The superscript of $\mathbf{\Pi}^{(m-n)}$ denotes the $m-n$ diagonal of the qth $LP \times LP$ block on the main diagonal of respective $LNP \times LNP$ sized block. Note that the main diagonal, i.e., $\mathbf{\Pi}^{(0)}$ denotes the usual (cross) PSD living on the stationary manifold. Hence,

$$\left[\hat{\mathbf{S}}^{(0)}_{s_i s_i}\right]^{(m,m)}_k = \hat{\mathbf{S}}^{(0)}_{s_i s_i}(\theta_{mN+q}, \theta_{mN+q}) = \mathbf{\Pi}^{(0)}_{s_i s_i}(\theta_{mN+q}) \in \mathbb{C}^{L\times L} \tag{7.56}$$

denotes the usual PSD at frequency θ_{mN+q} of channel i.

Accordingly, the coherence matrix $\mathbf{D}$ contains samples of the cyclic (cross) PSDs between surveillance arrays i and j for $i, j = 1, \ldots, J$ scaled by the usual PSD of arrays i and j, respectively. Each $LNP \times LNP$ sized block of the cross coherence matrix $\mathbf{C}$ can be interpreted similarly to the cross coherence matrix introduced in Chapter 5 as it contains the cyclic cross PSDs between all combinations of surveillance and reference arrays that are scaled by usual PSDs of the surveillance arrays and complicated functions of the cyclic cross PSDs across the combinations of reference arrays. Finally, the cross coherence matrix $\tilde{\mathbf{C}}$ is composed of the cyclic cross PSDs between all combinations of surveillance and reference arrays that are scaled by complicated functions of the cyclic cross PSDs of the surveillance arrays and of the reference arrays.

Hence, $\mathbf{D}$ accounts for the cross-spectral correlations across all combinations of surveillance arrays, whereas $\mathbf{C}$ and $\tilde{\mathbf{C}}$ account for the cross-spectral correlations among all combinations of surveillance and reference arrays. $\mathbf{C}$ and $\tilde{\mathbf{C}}$ differ in the way the cyclic cross PSDs are scaled.

Finally, the GLRT statistic combines the information about the presence of cyclostationarity among all surveillance arrays via the determinant of $\mathbf{D}$ with the determinant of $\mathbf{I}_{KLNP} - \tilde{\mathbf{C}}^H\tilde{\mathbf{C}}$. Where the latter accounts for the correlation between surveillance and reference arrays as well as the spectral correlation of those.

7.3.2 Threshold selection and null distribution

In order to apply the proposed GLR as a test statistic, we have to determine its distribution under the null hypothesis in order to fix the probability of false alarm.The GLR in (7.14) can be expressed as a ratio of determinants of the ML estimates of the covariance matrices $\hat{\mathbf{S}}_0$ and $\hat{\mathbf{S}}_1$ as shown in (7.39) in Section 7.3.1. Since a properly selected permutation of $\hat{\mathbf{S}}_0$ and $\hat{\mathbf{S}}_1$ makes them block-diagonal, we can apply our results from Chapter 3.

Proposition 7.3.1. *The distribution of the GLRT statistic proposed in Theorem 7.3.1 under the null hypothesis is given by a product of independent Beta random variables that is*

$$\mathscr{G}^{1/M} \overset{\mathcal{D}}{=} \prod_{l=1}^{N} \left\{ \prod_{k=1}^{JP} \prod_{m=1}^{L} U_{k,m} \right\} \left\{ \prod_{n=1}^{KLP} V_n \right\} \tag{7.57}$$

where all U, V are independent Beta distributed random variables. Specifically, they are given by

$$U_{k,m} \sim Beta\left\{M - ((k-1)L + m - 1), (k-1)L\right\}, \tag{7.58}$$

$$V_n \sim Beta\left\{M - (JLP + n - 1), JLP\right\}. \tag{7.59}$$

Proof. The proof directly follows from Theorem 3.4.1 for the following parameters: The total number of blocks under the alternative is given by $L = N$, where each block is of dimension $\lambda_l = (J+K)LP \; \forall l$. Now we can permute the blocks under the null such that each block under the alternative is tested against $n_l = JP + 1$ blocks under the null with $\kappa_k = \begin{cases} L, & k = 1, \ldots, JP, \\ KLP, & k = JP + 1. \end{cases}$ □

7.4 Derivation of optimal invariant tests

In order to examine the existence of UMPI and LMPI tests, we exploit Wijsman's theorem [72] once more. To this end, we have to identify the invariances of the hypothesis test similarly to Section 5.4 for the two-channel problem.

7.4.1 Invariances of the hypothesis test

The invariance group of the multistatic problem is given by

$$\mathcal{G} = \{g : \mathbf{z} \rightarrow g(\mathbf{z}) = \mathbf{\Psi}\mathbf{z}\}, \tag{7.60}$$

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

$$\mathbf{P}_{\clubsuit} = \left(\sum_{k=1}^{N} \boldsymbol{\epsilon}_k \boldsymbol{\epsilon}_k^T \otimes \mathbf{V}_{\clubsuit}^{(k)} \otimes \mathbf{I}_L \right) (\mathbf{I}_\kappa \otimes \mathbf{U} \otimes \mathbf{I}_{LP}) \, (\mathbf{W}_{\clubsuit} \otimes \mathbf{I}_{LNP}) \, . \tag{7.61}$$

The term in the first parentheses accounts for an arbitrary permutation within every LP block in $\mathbf{z}$, i.e., a permutation of those frequency components that are separated by multiples of $2\pi/P$. Hence, $\boldsymbol{\epsilon}_k$ is the kth column of $\mathbf{I}_N$, $\mathbf{V}_{\clubsuit}^{(k)} \in \mathbb{V}$ denotes a $P \times P$ permutation matrix. Moreover, the middle term in parentheses permutes the LP-sized blocks within every $\mathbf{z}_{s,j}$ and $\mathbf{z}_{r,k}$ in the same way. To this end, $\mathbf{U} \in \mathbb{U}$ is a permutation matrix of size $N \times N$ and $\kappa \in \{J, K\}$. The multistatic problem also allows for a permutation of the J and K channels without changing the structure of $\mathbf{S}_0$ and $\mathbf{S}_1$, i.e. $\mathbf{W}_{\clubsuit} \in \mathbb{W}$ denotes a $J \times J$ and $K \times K$ permutation matrix, respectively. $\mathbb{V}$, $\mathbb{U}$, and $\mathbb{W}$ denote the corresponding sets of P-, N-, and J-,K-dimensional permutation matrices, respectively. Finally, we should note that we can multiply $\mathbf{z}_s$ with any $\mathbf{G} \in \mathbb{G}$, where $\mathbb{G}$ is the set of nonsingular block-diagonal matrices with block size L and we can multiply $\mathbf{z}_r$ with any $\mathbf{H} \in \mathbb{B}_K\mathbb{H}$, where $\mathbb{B}_K\mathbb{H}$ denotes the set of nonsingular $K \times K$ block matrices, where each block is of dimension $LNP \times LNP$ and a block-diagonal matrix with block size LP itself. Note that we only consider linear operations in order to preserve Gaussianity of $\mathbf{z}$.

7.4.2 Ratio of maximal invariant densities

According to Wijsman's theorem the ratio of maximal invariant densities is given by

$$\mathscr{L} = \frac{\int_{\mathcal{G}} p(g(\mathbf{z}); \mathcal{H}_1) |\det(\mathbf{J}_g)| dg}{\int_{\mathcal{G}} p(g(\mathbf{z}); \mathcal{H}_0) |\det(\mathbf{J}_g)| dg}, \tag{7.62}$$

where $\mathcal{G}$ denotes the group of invariant transformations given by (7.60), 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 is the Lebesgue measure. Plugging in the likelihoods and the group of invariant transformations yields the following ratio

$$\mathscr{L} = \frac{\sum_{\mathbb{V}_0^N, \mathbb{V}_1^N, \mathbb{U}, \mathbb{W}} \int_{\mathbb{G}} \int_{\mathbb{H}} \det(\mathbf{S}_1)^{-M} |\det(\mathbf{G})|^{2M} |\det(\mathbf{H})|^{2M} \exp\left\{ -M \operatorname{tr}\left(\boldsymbol{\Psi}\mathbf{Q}\boldsymbol{\Psi}^H \mathbf{S}_1^{-1} \right) \right\} d\mathbf{G} d\mathbf{H}}{\sum_{\mathbb{V}_0^N, \mathbb{V}_1^N, \mathbb{U}, \mathbb{W}} \int_{\mathbb{G}} \int_{\mathbb{H}} \det(\mathbf{S}_0)^{-M} |\det(\mathbf{G})|^{2M} |\det(\mathbf{H})|^{2M} \exp\left\{ -M \operatorname{tr}\left(\boldsymbol{\Psi}\mathbf{Q}\boldsymbol{\Psi}^H \mathbf{S}_0^{-1} \right) \right\} d\mathbf{G} d\mathbf{H}}. \tag{7.63}$$

Although this ratio is a complicated function it can be simplified if we discard all those terms that do neither depend on the observations nor on the invariances as shown in the following lemma

Lemma 7.4.1. *Equation 7.63 can be simplified as*

$$\mathscr{L} \propto \sum_{\mathbb{V}_0^N, \mathbb{V}_1^N, \mathbb{U}, \mathbb{W}} \int_{\mathbb{G}} \int_{\mathbb{H}} \beta(\mathbf{G}) \gamma(\mathbf{H}) e^{-M[\alpha_1(\mathbf{G}) + \alpha_2(\mathbf{G}, \mathbf{H})]} d\mathbf{G} d\mathbf{H}, \tag{7.64}$$

with

$$\beta(\mathbf{G}) = |\det(\mathbf{G})|^{2M}, \tag{7.65}$$

$$\gamma(\mathbf{H}) = |\det(\mathbf{H})|^{2M} e^{-M \operatorname{tr}(\mathbf{H}\mathbf{H}^H)}, \tag{7.66}$$

$$\alpha_1(\mathbf{G}) = \sum_{j=1}^{JN} \sum_{k=1}^{J} \sum_{m,n=1}^{P} \operatorname{tr}\left(\mathbf{\Gamma}^{(m,n)}_{(j,f(j,k))} \mathbf{G}^{(n,n)}_{(f(j,k),f(j,k))} \mathbf{D}^{(n,m)}_{(f(j,k),j)} \mathbf{G}^{(m,m)H}_{(j,j)} \right) \tag{7.67}$$

$$\mathbf{\Gamma} = \mathbf{P}_s^T \operatorname{diag}_L\left(\mathbf{\Sigma}_1\right)^{-\frac{1}{2}} \mathbf{\Sigma}_1 \operatorname{diag}_L\left(\mathbf{\Sigma}_1\right)^{-\frac{1}{2}} \mathbf{P}_s, \tag{7.68}$$

$$\mathbf{D} = \left(\hat{\mathbf{S}}_s^{(0)}\right)^{-1/2} \hat{\mathbf{S}}_s^{(1)} \left(\hat{\mathbf{S}}_s^{(0)}\right)^{-1/2} = \left(\operatorname{diag}_L\left(\mathbf{Q}_s\right)\right)^{-1/2} \operatorname{block}_{J,J} \operatorname{diag}_{LP}\left(\mathbf{Q}_s\right) \left(\operatorname{diag}_L\left(\mathbf{Q}_s\right)\right)^{-1/2}, \tag{7.69}$$

$$\alpha_2(\mathbf{G}, \mathbf{H}) = \sum_{l=1}^{KN} \sum_{j=1}^{J} \sum_{k=1}^{K} \sum_{m,n,p=1}^{P} \operatorname{tr}\left(\mathbf{\Lambda}^{(m,n)}_{(l,f(l,j))} \mathbf{G}^{(n,n)}_{(f(l,j),f(l,j))} \mathbf{C}^{(n,p)}_{(f(l,j),f(l,k))} \mathbf{H}^{(p,m)H}_{(f(l,k),l)} \right), \tag{7.70}$$

$$\mathbf{\Lambda} = \mathbf{P}_r^T \mathbf{\Sigma}_2^{-\frac{1}{2}} \mathbf{\Sigma}_{21} \operatorname{diag}_L\left(\mathbf{\Sigma}_1\right)^{-\frac{1}{2}} \mathbf{P}_s, \tag{7.71}$$

$$\mathbf{C} = \left(\hat{\mathbf{S}}_s^{(0)}\right)^{-1/2} \hat{\mathbf{S}}_{sr} \left(\hat{\mathbf{S}}_r\right)^{-1/2} \tag{7.72}$$

$$= \left(\operatorname{diag}_L\left(\mathbf{Q}_s\right)\right)^{-1/2} \operatorname{block}_{K,K} \operatorname{diag}_{LP}\left(\mathbf{Q}_{sr}\right) \left(\operatorname{block}_{K,K} \operatorname{diag}_{LP}\left(\mathbf{Q}_r\right)\right)^{-1/2}, \tag{7.73}$$

where $f(j,k) = (k-1)N + \operatorname{mod}(j,N)$ *and* $\mathbf{S}_1^{-1} = \mathbf{\Sigma} = \begin{bmatrix} \mathbf{\Sigma}_1 & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_2 \end{bmatrix}$.

Proof. Starting from (7.63), we can observe that the terms $\det(\mathbf{S}_0)^{-M}$ and $\det(\mathbf{S}_1)^{-M}$ neither depend on the observations nor on the invariances. Hence, they can be discarded in the ratio. Before we further simplify the ratio, we will first present the following lemma:

Lemma 7.4.2. *Assume we have a matrix* $\mathbf{H} \in \mathbb{B}_K \mathbb{S}_{LP}^{LNP}$*, which is a* $K \times K$ *block matrix with each block being block-diagonal with block size* LP*. Then the structure of its inverse* $\mathbf{H}^{-1}$ *has the identical structure as* $\mathbf{H}$*, i.e.,* $\mathbf{H}^{-1} \in \mathbb{B}_K \mathbb{S}_{LP}^{LNP}$.

Proof. We will proof this lemma by induction. Let us start with $K = 1$, i.e. $\mathbf{H} \in \mathbb{S}_{LP}^{LNP}$ and clearly, its inverse is a block-diagonal matrix with block size LP as well, i.e., $\mathbf{H}^{-1} \in \mathbb{S}_{LP}^{LNP}$.

Now, let us assume that $K = n + 1$ and we partition the matrix $\mathbf{H} \in \mathbb{B}_{n+1}\mathbb{S}_{LP}^{LNP}$ as follows

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_{11} & \mathbf{H}_{12} \\ \mathbf{H}_{21} & \mathbf{H}_{22} \end{bmatrix}, \tag{7.74}$$

where $\mathbf{H}_{11} \in \mathbb{B}_n\mathbb{S}_{LP}^{LNP}$, $\mathbf{H}_{22} \in \mathbb{S}_{LP}^{LNP}$, $\mathbf{H}_{12}$ is an $n \times 1$ block matrix, where each block is in $\mathbb{S}_{LP}^{LNP}$, and similarly, $\mathbf{H}_{21}$ is an $1 \times n$ block matrix, where each block is in $\mathbb{S}_{LP}^{LNP}$. Moreover, by induction we presume that $\mathbf{H}_{11}^{-1} \in \mathbb{B}_n\mathbb{S}_{LP}^{LNP}$. Given the matrix inversion of a 2×2 block matrix the inverse of $\mathbf{H}$ is given by

$$\mathbf{H}^{-1} = \begin{bmatrix} \mathbf{H}_{11}^{-1} + \mathbf{H}_{11}^{-1}\mathbf{H}_{12}\left(\mathbf{H}_{22} - \mathbf{H}_{21}\mathbf{H}_{11}^{-1}\mathbf{H}_{12}\right)^{-1}\mathbf{H}_{21}\mathbf{H}_{11}^{-1} & -\mathbf{H}_{11}^{-1}\mathbf{H}_{12}\left(\mathbf{H}_{22} - \mathbf{H}_{21}\mathbf{H}_{11}^{-1}\mathbf{H}_{12}\right)^{-1} \\ -\left(\mathbf{H}_{22} - \mathbf{H}_{21}\mathbf{H}_{11}^{-1}\mathbf{H}_{12}\right)^{-1}\mathbf{H}_{21}\mathbf{H}_{11}^{-1} & \left(\mathbf{H}_{22} - \mathbf{H}_{21}\mathbf{H}_{11}^{-1}\mathbf{H}_{12}\right)^{-1} \end{bmatrix}. \tag{7.75}$$

In order to show that $\mathbf{H}^{-1} \in \mathbb{B}_{n+1}\mathbb{S}_{LP}^{LNP}$, we will establish that all four blocks in (7.75) have the identical structure as the blocks in (7.74). First, we consider the south-east block $\left(\mathbf{H}_{22} - \mathbf{H}_{21}\mathbf{H}_{11}^{-1}\mathbf{H}_{12}\right)^{-1}$. Let us have a look at the product $\mathbf{H}_{21}\mathbf{H}_{11}^{-1}$: Since both $\mathbf{H}_{21}$ and $\mathbf{H}_{11}^{-1}$ matrices consist of $n \times 1$ and $n \times n$ blocks of LNP-sized block-diagonal matrices, building the product leads to multiplying block-diagonal matrices with identical block sizes and summing those up. Hence, each block in the product will be block-diagonal with block size LP again, i.e., $\mathbf{H}_{21}\mathbf{H}_{11}^{-1}$ has the same structure as $\mathbf{H}_{21}$. Similarly, it is easy to see that multiplying $\mathbf{H}_{21}\mathbf{H}_{11}^{-1}$ with $\mathbf{H}_{12}$ ends up in a block-diagonal matrix. The difference of two block-diagonal matrices is still block-diagonal and so is its inverse. Hence, the south-east block of $\mathbf{H}^{-1}$ has the same structure as the south-east block of $\mathbf{H}$.

Second, given the previous paragraph, it is obvious that the north-east and south-east blocks are $n \times 1$ and $1 \times n$ block matrices consisting of block-diagonal matrices, respectively. Finally, for the north-west block we first consider the product $\mathbf{H}_{11}^{-1}\mathbf{H}_{12}\left(\mathbf{H}_{22} - \mathbf{H}_{21}\mathbf{H}_{11}^{-1}\mathbf{H}_{12}\right)^{-1}\mathbf{H}_{21}\mathbf{H}_{11}^{-1}$, which is, given the previous results, a product of an $n \times 1$ and an $1 \times n$ block matrix, which is an $n \times n$ block matrix of which every block is block-diagonal. Hence, the north-west block is the sum of two matrices in $\mathbb{B}_n\mathbb{S}_{LP}^{LNP}$, which is also in $\mathbb{B}_n\mathbb{S}_{LP}^{LNP}$. In summary, every block in $\mathbf{H}^{-1}$ has the same structure as the blocks in $\mathbf{H}$, which concludes the proof. □

Handling the denominator of (7.63)

Now we simplify the denominator of (7.63). Let us first focus on the exponential term. Taking into account that $\mathbf{\Psi}$ and $\mathbf{S}_0$ are block-diagonal matrices that consist of one block of dimension $JLNP \times JLNP$ and a second block of dimension $KLNP \times KLNP$, the trace operation in the exponential term can be rewritten as

$$\operatorname{tr}\left(\mathbf{\Psi}\mathbf{Q}\mathbf{\Psi}^H\mathbf{S}_0^{-1}\right) = \operatorname{tr}\left(\mathbf{P}_s\mathbf{G}\mathbf{Q}_s\mathbf{G}^H\mathbf{P}_s^T\left[\mathbf{S}_s^{(0)}\right]^{-1}\right) + \operatorname{tr}\left(\mathbf{P}_r\mathbf{H}\mathbf{Q}_r\mathbf{H}^H\mathbf{P}_r^T\mathbf{S}_r^{-1}\right). \tag{7.76}$$

Applying the change of variables $\mathbf{G} \rightarrow \mathbf{G}\mathbf{B}_s^{-\frac{1}{2}}$,where $\mathbf{B}_s = \text{diag}_L(\mathbf{Q}_s)$ and another change of variables $\mathbf{H} \rightarrow \mathbf{H}\mathbf{B}_r^{-\frac{1}{2}}$ with $\mathbf{B}_r = \text{block}_{K,K}\,\text{diag}_{LP}(\mathbf{Q}_r)$ and where we exploit Lemma 7.4.2, which assures that $\mathbf{B}_r^{-\frac{1}{2}} \in \mathbb{B}_K\mathbb{S}_{LP}^{LNP}$, yields

$$\text{tr}\left(\mathbf{\Psi}\mathbf{Q}\mathbf{\Psi}^H\mathbf{S}_0^{-1}\right) = \text{tr}\left(\mathbf{P}_s\mathbf{G}\mathbf{B}_s^{-\frac{1}{2}}\mathbf{Q}_s\mathbf{B}_s^{-\frac{1}{2}}\mathbf{G}^H\mathbf{P}_s^T\left[\mathbf{S}_s^{(0)}\right]^{-1}\right) + \text{tr}\left(\mathbf{P}_r\mathbf{H}\mathbf{B}_r^{-\frac{1}{2}}\mathbf{Q}_r\mathbf{B}_r^{-\frac{1}{2}}\mathbf{H}^H\mathbf{P}_r^T\mathbf{S}_r^{-1}\right). \quad (7.77)$$

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. Moreover, considering the LNP-sized off-diagonal blocks of $\mathbf{B}_r^{-\frac{1}{2}}\mathbf{Q}_r\mathbf{B}_r^{-\frac{1}{2}}$, we should note that the LP-sized elements on the main diagonals of those blocks are zero.

Let us consider the first trace operation in the sum (7.76). Since both $\mathbf{G}$ and $\mathbf{S}_s^{-1}$ are block-diagonal with block size L, only the whitened main diagonal blocks of $\mathbf{B}_s^{-\frac{1}{2}}\mathbf{Q}_s\mathbf{B}_s^{-\frac{1}{2}}$ are involved in the trace operation, i.e., there is no data-dependency in the first term. Note that $\mathbf{P}_s$ preserves the block-diagonal structure. In order to handle the second trace operation, we should recall that $\mathbf{H} \in \mathbb{B}_K\mathbb{S}_{LP}^{LNP}$ and by Lemma 7.4.2 $\mathbf{S}_r^{-1} \in \mathbb{B}_K\mathbb{S}_{LP}^{LNP}$. Since the $LP \times LP$ main diagonal blocks of $\mathbf{B}_r^{-\frac{1}{2}}\mathbf{Q}_r\mathbf{B}_r^{-\frac{1}{2}}$ are whitened and the LP-sized main diagonal blocks of the off-diagonals are zero, there is no data-dependent terms in the trace operation. Here $\mathbf{P}_r$ preserves the $\mathbb{B}_{J+K}\mathbb{S}_{LP}^{LNP}$ structure.

Hence, we can discard the denominator in the ratio (7.63). and $\mathscr{L}$ simplifies as

$$\mathscr{L} \propto \sum_{\mathbb{V}_0^N,\mathbb{V}_1^N,\mathbb{U},\mathbb{W}} \int_{\mathbb{G}}\int_{\mathbb{H}} |\det(\mathbf{G})|^{2M}|\det(\mathbf{H})|^{2M}\exp\left\{-M\text{tr}\left(\mathbf{\Psi}\mathbf{Q}\mathbf{\Psi}^H\mathbf{S}_1^{-1}\right)\right\} d\mathbf{G}d\mathbf{H}. \quad (7.78)$$

Further simplifying the ratio

Let us first consider the exponential term in (7.78). We define the matrices $\mathbf{\Sigma} = \mathbf{S}_1^{-1} \in \mathbb{B}_{J+K}\mathbb{S}_{LP}^{LNP}$, which keeps the structure of $\mathbf{S}_1$ by Lemma 7.4.2 and $\bar{\mathbf{\Psi}} = \mathbf{\Psi}\mathbf{Q}\mathbf{\Psi}^H \in \mathbb{B}_{J+K}\mathbb{S}_{LP}^{LNP}$. These matrices can be partitioned into $\mathbf{\Sigma} = \begin{bmatrix} \mathbf{\Sigma}_1 & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_2 \end{bmatrix}$ and $\bar{\mathbf{\Psi}} = \begin{bmatrix} \bar{\mathbf{\Psi}}_1 & \bar{\mathbf{\Psi}}_{12} \\ \bar{\mathbf{\Psi}}_{21} & \bar{\mathbf{\Psi}}_2 \end{bmatrix}$. 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 $\bar{\mathbf{\Psi}}$ are given by $\bar{\mathbf{\Psi}}_1 = \mathbf{P}_s\mathbf{G}\mathbf{B}_s^{-\frac{1}{2}}\mathbf{Q}_s\mathbf{B}_s^{-\frac{1}{2}}\mathbf{G}^H\mathbf{P}_s^H$, $\bar{\mathbf{\Psi}}_2 = \mathbf{P}_r\mathbf{H}\mathbf{B}_r^{-\frac{1}{2}}\mathbf{Q}_r\mathbf{B}_r^{-\frac{1}{2}}\mathbf{H}^H\mathbf{P}_r^H$ and $\bar{\mathbf{\Psi}}_{12} = \bar{\mathbf{\Psi}}_{21}^H = \mathbf{P}_s\mathbf{G}\mathbf{B}_s^{-\frac{1}{2}}\mathbf{Q}_{sr}\mathbf{B}_r^{-\frac{1}{2}}\mathbf{H}^H\mathbf{P}_r^H$. Finally, putting things together, (7.78) becomes

$$\mathscr{L} \propto \sum_{\mathbb{V}_0^N,\mathbb{V}_1^N,\mathbb{U}} \int_{\mathbb{G}}\int_{\mathbb{H}} |\det(\mathbf{G})|^{2M}|\det(\mathbf{H})|^{2M} e^{-M\,\text{tr}(\bar{\mathbf{\Psi}}_1\mathbf{\Sigma}_1)} e^{-2M\,\text{tr}(\bar{\mathbf{\Psi}}_{12}\mathbf{\Sigma}_{21})} e^{-M\,\text{tr}(\bar{\mathbf{\Psi}}_2\mathbf{\Sigma}_2)} d\mathbf{G}d\mathbf{H}. \quad (7.79)$$

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

$$\begin{aligned}\operatorname{tr}(\bar{\boldsymbol{\Psi}}_1 \boldsymbol{\Sigma}_1) &= \sum_{j=1}^{JN} \sum_{k=1}^{J} \operatorname{tr}\left(\boldsymbol{\Gamma}_{(j,f(j,k))} \mathbf{G}_{(f(j,k),f(j,k))} \mathbf{D}_{(f(j,k),j)} \mathbf{G}_{(j,j)}^H\right), \\ &= \sum_{j=1}^{JN} \sum_{k=1}^{J} \sum_{m,n=1}^{P} \operatorname{tr}\left(\boldsymbol{\Gamma}_{(j,f(j,k))}^{(m,n)} \mathbf{G}_{(f(j,k),f(j,k))}^{(n,n)} \mathbf{D}_{(f(j,k),j)}^{(n,m)} \mathbf{G}_{(j,j)}^{(m,m)H}\right) \end{aligned} \tag{7.80}$$

with $\boldsymbol{\Gamma}$ and $\mathbf{D}$, given by (7.68) and (7.69), respectively, and $f(j,k) = (k-1)N + \operatorname{mod}(j,N)$. Both matrices $\mathbf{D}$ and $\boldsymbol{\Gamma}$ are in $\mathbb{B}_J \mathbb{S}_{LP}^{LNP}$ and the index $f(j,k) = (k-1)N + \operatorname{mod}(j,N)$ selects the LP-sized blocks on the main diagonals of every LNP-sized block. It should be noted that (7.80) depends on unknown parameters through $\boldsymbol{\Gamma}$.

The second exponential term in (7.79) can be reduced by introducing another change of variables $\mathbf{H} \to \mathbf{P}_r^T \boldsymbol{\Sigma}_2^{-\frac{1}{2}} \mathbf{P}_r \mathbf{H}$ as

$$\begin{aligned}\operatorname{tr}\left(\bar{\boldsymbol{\Psi}}_{12} \boldsymbol{\Sigma}_{21}\right) &= \sum_{l=1}^{KN} \sum_{j=1}^{J} \sum_{k=1}^{K} \operatorname{tr}\left(\boldsymbol{\Lambda}_{(l,f(l,j))} \mathbf{G}_{(f(l,j),f(l,j))} \mathbf{C}_{(f(l,j),f(l,k))} \mathbf{H}_{(f(l,k),l)}^H\right), \\ &= \sum_{l=1}^{KN} \sum_{j=1}^{J} \sum_{k=1}^{K} \sum_{m,n,p=1}^{P} \operatorname{tr}\left(\boldsymbol{\Lambda}_{(l,f(l,j))}^{(m,n)} \mathbf{G}_{(f(l,j),f(l,j))}^{(n,n)} \mathbf{C}_{(f(l,j),f(l,k))}^{(n,p)} \mathbf{H}_{(f(l,k),l)}^{(p,m)H}\right), \end{aligned} \tag{7.81}$$

where $\boldsymbol{\Lambda}$ is given by (7.71).

Finally, by plugging in the previous change of variables, the last exponential term in (7.79) becomes

$$\begin{aligned}\operatorname{tr}\left(\bar{\boldsymbol{\Psi}}_2 \boldsymbol{\Sigma}_2\right) &= \operatorname{tr}\left(\boldsymbol{\Sigma}_2^{-\frac{1}{2}} \mathbf{P}_r \mathbf{H} \mathbf{B}_r^{-\frac{1}{2}} \mathbf{Q}_r \mathbf{B}_r^{-\frac{1}{2}} \mathbf{H}^H \mathbf{P}_r^T \boldsymbol{\Sigma}_2^{-\frac{1}{2}} \boldsymbol{\Sigma}_2\right) \\ &= \operatorname{tr}\left(\mathbf{H}\mathbf{H}^H\right), \end{aligned} \tag{7.82}$$

where the last simplification follows from the fact that $\mathbf{H}$ is a $K \times K$ block matrix with blocks that are 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 and zero on its LP-sized off-diagonal blocks of interest. The proof follows by plugging (7.80), (7.81), and (7.82) into (7.79).

□

We should note that both $\boldsymbol{\Gamma}$ and $\boldsymbol{\Lambda}$ 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 surveillance arrays are 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 $\mathbf{\Sigma}_{12} \approx \mathbf{0}$, and $\mathbf{\Sigma}_1$ is also close to being block-diagonal with block size L. Therefore, we can follow the same approach as in Section 5.4 since 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 (7.64) around $\alpha_1(\mathbf{G}) + \alpha_2(\mathbf{G}, \mathbf{H}) = 0$ as

$$e^{-M(\alpha_1(\mathbf{G})+\alpha_2(\mathbf{G},\mathbf{H}))} \approx 1 - M(\alpha_1(\mathbf{G}) + \alpha_2(\mathbf{G}, \mathbf{H})) + \frac{M^2}{2}\left[\alpha_1^2(\mathbf{G}) + 2\alpha_1(\mathbf{G})\alpha_2(\mathbf{G}, \mathbf{H}) + \alpha_2^2(\mathbf{G}, \mathbf{H})\right]. \quad (7.83)$$

Thus, (7.64) can be approximated as

$$\mathscr{L} \propto \mathscr{L}_1 + \mathscr{L}_2 + \mathscr{L}_3 + \mathscr{L}_4 + \mathscr{L}_5, \quad (7.84)$$

where

$$\mathscr{L}_1 = -M \sum_{\mathbb{V}_0^N, \mathbb{V}_1^N, \mathbb{U}, \mathbb{W}} \int_{\mathbb{G}} \beta(\mathbf{G})\alpha_1(\mathbf{G}) d\mathbf{G} \int_{\mathbb{H}} \gamma(\mathbf{H}) d\mathbf{H}, \quad (7.85)$$

$$\mathscr{L}_2 = -M \sum_{\mathbb{V}_0^N, \mathbb{V}_1^N, \mathbb{U}, \mathbb{W}} \int_{\mathbb{G}} \int_{\mathbb{H}} \beta(\mathbf{G})\gamma(\mathbf{H})\alpha_2(\mathbf{G}, \mathbf{H}) d\mathbf{G} d\mathbf{H}, \quad (7.86)$$

$$\mathscr{L}_3 = \frac{M^2}{2} \sum_{\mathbb{V}_0^N, \mathbb{V}_1^N, \mathbb{U}, \mathbb{W}} \int_{\mathbb{G}} \beta(\mathbf{G})\alpha_1^2(\mathbf{G}) d\mathbf{G} \int_{\mathbb{H}} \gamma(\mathbf{H}) d\mathbf{H}, \quad (7.87)$$

$$\mathscr{L}_4 = \frac{M^2}{2} \sum_{\mathbb{V}_0^N, \mathbb{V}_1^N, \mathbb{U}, \mathbb{W}} \int_{\mathbb{G}} \int_{\mathbb{H}} \beta(\mathbf{G})\gamma(\mathbf{H})\alpha_1(\mathbf{G})\alpha_2(\mathbf{G}, \mathbf{H}) d\mathbf{G} d\mathbf{H}, \quad (7.88)$$

$$\mathscr{L}_5 = \frac{M^2}{2} \sum_{\mathbb{V}_0^N, \mathbb{V}_1^N, \mathbb{U}, \mathbb{W}} \int_{\mathbb{G}} \int_{\mathbb{H}} \beta(\mathbf{G})\gamma(\mathbf{H})\alpha_2^2(\mathbf{G}, \mathbf{H}) d\mathbf{G} d\mathbf{H}. \quad (7.89)$$

In the following lemma, we show that (7.84) further simplifies.

Lemma 7.4.3. *The following terms are zero:*

$$\mathscr{L}_1 = 0, \quad (7.90)$$

$$\mathscr{L}_2 = 0, \quad (7.91)$$

$$\mathscr{L}_4 = 0. \quad (7.92)$$

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

$$\mathscr{L}_1 \propto \sum_{\mathbb{V}_0^N, \mathbb{V}_1^N, \mathbb{U}, \mathbb{W}} \int_{\mathbb{G}} \int_{\mathbb{H}} \beta(\mathbf{G})\gamma(\mathbf{H}) \times \sum_{j=1}^{KN} \sum_{k=1}^{K} \sum_{m,n=1}^{P} \operatorname{tr}\left(\mathbf{\Gamma}^{(m,n)}_{(j,f(j,k))} \mathbf{G}^{(n,n)}_{(f(j,k),f(j,k))} \mathbf{D}^{(n,m)}_{(f(j,k),j)} \mathbf{G}^{(m,m)H}_{(j,j)}\right) d\mathbf{G} d\mathbf{H}. \quad (7.93)$$

Applying the change of variables $\mathbf{G}^{(n,n)}_{(f(j,k),f(j,k))} \rightarrow -\mathbf{G}^{(n,n)}_{(f(j,k),f(j,k))}$ shows 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 $\mathscr{L}_2$ and $\mathscr{L}_4$ are zero. □

Hence, (7.84) simplifies to $\mathscr{L} \propto \mathscr{L}_3 + \mathscr{L}_5$. Moreover, we can further simplify the terms $\mathscr{L}_3$ and $\mathscr{L}_5$, which lead to the following theorem

Theorem 7.4.1. *The ratio of the distribution of the maximal invariant statistic in* (7.63) *is*

$$\mathscr{L} \propto \mathscr{L}_S + \gamma \mathscr{L}_{SR}, \quad (7.94)$$

where

$$\mathscr{L}_S = ||\mathbf{D}||^2 \quad (7.95)$$

and

$$\mathscr{L}_{SR} = ||\mathbf{C}||^2 \quad (7.96)$$

with $\mathbf{D}$ *and* $\mathbf{C}$ *given by* (7.69) *and* (7.73)*, respectively. The parameter* γ *is a constant that depends on unknown parameters through* $\mathbf{S}_1$ *but is independent of the observations.*

Proof.

$$\begin{aligned} \mathscr{L}_3 \quad &= \tfrac{M^2}{2} \textstyle\sum_{\mathbb{V}_0^N, \mathbb{V}_1^N, \mathbb{U}, \mathbb{W}} \int_{\mathbb{G}} \beta(\mathbf{G}) \left[\sum_{j=1}^{KN} \sum_{k=1}^{K} \sum_{m,n=1}^{P} \right. \\ &\quad \left. \times \operatorname{tr}\left(\mathbf{\Gamma}^{(m,n)}_{(j,f(j,k))} \mathbf{G}^{(n,n)}_{(f(j,k),f(j,k))} \mathbf{D}^{(n,m)}_{(f(j,k),j)} \mathbf{G}^{(m,m)H}_{(j,j)}\right)\right]^2 d\mathbf{G} \int_{\mathbb{H}} \gamma(\mathbf{H}) d\mathbf{H}, \end{aligned} \quad (7.97)$$

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 $\mathbf{G}^{(n,n)}_{(f(j,k),f(j,k))} \rightarrow -\mathbf{G}^{(n,n)}_{(f(j,k),f(j,k))}$ as before. Moreover, the last integral is just a constant that does not involve observations or any of the unknowns and for this reason it can be neglected. Now $\mathscr{L}_3$ simplifies as follows

$$\begin{aligned} \mathscr{L}_3 \quad &\propto \textstyle\sum_{\mathbb{V}_0^N, \mathbb{V}_1^N, \mathbb{U}, \mathbb{W}} \int_{\mathbb{G}} \beta(\mathbf{G}) \sum_{j=1}^{KN} \sum_{k=1}^{K} \sum_{m,n=1}^{P} \\ &\quad \times \operatorname{tr}^2\left(\mathbf{\Gamma}^{(m,n)}_{(j,f(j,k))} \mathbf{G}^{(n,n)}_{(f(j,k),f(j,k))} \mathbf{D}^{(n,m)}_{(f(j,k),j)} \mathbf{G}^{(m,m)H}_{(j,j)}\right) d\mathbf{G}. \end{aligned} \quad (7.98)$$

Now (7.98) became a similar expression as in Appendix C in [38] and we can simplify it similarly to obtain

$$\mathscr{L}_3 \propto \mathscr{L}_S = ||\mathbf{D}||^2, \tag{7.99}$$

where $\mathbf{D}$ is given by (7.69).

Secondly, we can reduce (7.89) as follows

$$\begin{aligned}\mathscr{L}_5 \quad &= \tfrac{M^2}{2} \sum_{\mathbb{V}_0^N, \mathbb{V}_1^N, \mathbb{U}, \mathbb{W}} \int_{\mathbb{G}} \int_{\mathbb{H}} \beta(\mathbf{G})\gamma(\mathbf{H}) \Big[\sum_{j=1}^{KN} \sum_{k,l=1}^{K} \sum_{m,n,p=1}^{P} \\ &\quad \times \operatorname{tr}\left(\mathbf{\Lambda}^{(m,n)}_{(j,f(j,k))} \mathbf{G}^{(n,n)}_{(f(j,k),f(j,k))} \mathbf{C}^{(n,p)}_{(f(j,k),f(j,l))} \mathbf{H}^{(p,m)H}_{(f(j,l),j)} \right) \Big]^2 d\mathbf{G} d\mathbf{H},\end{aligned} \tag{7.101}$$

where the cross-terms of the square cancel out by another change of variables of either $\mathbf{G}^{(n,n)}_{(f(j,k),f(j,k))} \to -\mathbf{G}^{(n,n)}_{(f(j,k),f(j,k))}$ or $\mathbf{H}^{(p,m)}_{(f(j,l),j)} \to -\mathbf{H}^{(p,m)}_{(f(j,l),j)}$. Then the $\mathscr{L}_5$ simplifies to

$$\begin{aligned}\mathscr{L}_5 \quad &\propto \sum_{\mathbb{V}_0^N, \mathbb{V}_1^N, \mathbb{U}, \mathbb{W}} \int_{\mathbb{G}} \int_{\mathbb{H}} \beta(\mathbf{G})\gamma(\mathbf{H}) \sum_{j=1}^{KN} \sum_{k,l=1}^{K} \sum_{m,n,p=1}^{P} \\ &\quad \times \operatorname{tr}^2 \left(\mathbf{\Lambda}^{(m,n)}_{(j,f(j,k))} \mathbf{G}^{(n,n)}_{(f(j,k),f(j,k))} \mathbf{C}^{(n,p)}_{(f(j,k),f(j,l))} \mathbf{H}^{(p,m)H}_{(f(j,l),j)} \right) d\mathbf{G} d\mathbf{H}.\end{aligned} \tag{7.102}$$

Finally, following similar steps as in Appendix C of [38], we obtain

$$\mathscr{L}_5 \propto \mathscr{L}_{SR} = ||\mathbf{C}||^2, \tag{7.103}$$

where $\mathbf{C}$ is given by (7.73). It should be noted that $\mathscr{L}_3$ and $\mathscr{L}_5$ are equal to $\mathscr{L}_S$ and $\mathscr{L}_{SR}$ 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 γ, which allows us to express (7.84) as

$$\mathscr{L} \propto \mathscr{L}_S + \gamma \mathscr{L}_{SR}. \tag{7.104}$$

□

From this theorem, we can conclude that the LMPIT for the multistatic problem does not exist since the parameter γ depends on unknown parameters. However, we can observe that similar to the derivations in Section 5.4 we end up with an expression that is a weighted sum of the Frobenius norms of the coherence matrix $\mathbf{D}$ and the Frobenius norm of the cross coherence matrix $\mathbf{C}$. Although the first term $\mathscr{L}_S$ resembles the single array LMPIT derived in [38], it should be noted that $\mathbf{D}$ consists of $J \times J$ blocks, where each block is block diagonal with block size LP. Hence, the Frobenius norm of this matrix is actually a sum of all J single arrays LMPITs, which correspond to the blocks on the main diagonal of $\mathbf{D}$, but it also accounts for all spectral correlations across the J arrays by the sum over the Frobenius norm of all non-zero off-diagonal blocks.

Similarly, the second term $\mathscr{L}_{SR}$ in (7.94) is the multi-array extension of (5.65) and accounts for all cross-spectral correlations of all combinations of surveillance and reference arrays. Hence, $\mathscr{L}_{SR}$ is the sum of all two-channel LMPIT-inspired detectors proposed in (5.84). Please note that the GLRT for the multistatic case cannot be divided into a function of all two-channel detectors since it involves the determinant of block matrices and is a complicated function of Schur complements.

Since the LMPIT does not exist, we analyze the performance of the individual terms $\mathscr{L}_S$ and $\mathscr{L}_{SR}$ as well as the influence of the unknown constant γ in order to investigate whether an LMPIT-inspired test can be proposed. To this end we run Monte Carlo simulations for different scenarios in the following section.

7.5 Numerical Results

In this section we evaluate the performance of the multistatic GLRT and of $\mathscr{L}_S$ and $\mathscr{L}_{SR}$ from (7.94) and the influence of the unknown parameter γ using Monte Carlo simulations. To this end, we simulate data according to the model (7.1). Specifically, for all J, K we generate $\mathbf{H}_{s,j}[n]$ and $\mathbf{H}_{r,k}[n]$ and $\mathbf{v}_{s,j}[n]$ and $\mathbf{v}_{r,k}[n]$ for $j = 1, \ldots, J$ and $k = 1, \ldots, K$ independently. Also note that we draw new realizations of all parameters in each Monte Carlo simulation.

Similar to Section 5.7 we define the SNR in our simulation as follows

$$\mathrm{SNR}_{\clubsuit} = 10 \log_{10} \left(\frac{\mathrm{tr}\left(\hat{\mathbf{R}}_{\clubsuit}\right)}{\mathrm{tr}\left(\hat{\mathbf{V}}_{\clubsuit}\right)} \right), \tag{7.105}$$

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

$$\hat{\mathbf{R}}_{\clubsuit} = \frac{1}{MNP} \sum_{n=0}^{MNP-1} \mathbf{h}_{\clubsuit}[n] \mathbf{h}_{\clubsuit}^{H}[n] \in \mathbb{C}^{L \times L}, \tag{7.106}$$

$$\hat{\mathbf{V}}_{\clubsuit} = \frac{1}{MNP} \sum_{n=0}^{MNP-1} \mathbf{v}_{\clubsuit}[n] \mathbf{v}_{\clubsuit}^{H}[n] \in \mathbb{C}^{L \times L}, \tag{7.107}$$

with

$$\mathbf{h}_s[n] = \left[\left(\mathbf{H}_{s,1}[n] * \mathbf{s}[n]\right)^T, \cdots, \left(\mathbf{H}_{s,J}[n] * \mathbf{s}[n]\right)^T \right]^T, \tag{7.108}$$

$$\mathbf{h}_r[n] = \left[\left(\mathbf{H}_{r,1}[n] * \mathbf{s}[n]\right)^T, \cdots, \left(\mathbf{H}_{r,K}[n] * \mathbf{s}[n]\right)^T \right]^T, \tag{7.109}$$

and similarly, $\mathbf{v}_s[n] = \left[\mathbf{v}_{s,1}[n]^T, \cdots, \mathbf{v}_{s,J}[n]^T\right]$ and $\mathbf{v}_r[n] = \left[\mathbf{v}_{r,1}[n]^T, \cdots, \mathbf{v}_{r,K}[n]^T\right]$. Hence, the SNR is defined as an average across all surveillance and reference arrays, respectively.

In order to evaluate the performance of the test statistics, we further generalize the detector proposed in [67] to the case of multiple surveillance and reference arrays as follows

$$\mathscr{K} = \frac{\det(\mathbf{R})}{\det\left(\mathrm{diag}_L\{\mathbf{R}_{ss}\}\right)\det(\mathbf{R}_{rr})} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta, \tag{7.110}$$

where

$$\mathbf{R} = \begin{bmatrix} \mathbf{R}_{ss} & \mathbf{R}_{sr} \\ \mathbf{R}_{sr}^H & \mathbf{R}_{rr} \end{bmatrix} \tag{7.111}$$

with

$$\mathbf{R}_{ss} = \frac{1}{MNP}\sum_{n=0}^{MNP-1} \mathbf{u}_s[n]\mathbf{u}_s^H[n], \tag{7.112}$$

$$\mathbf{R}_{sr} = \frac{1}{MNP}\sum_{n=0}^{MNP-1} \mathbf{u}_s[n]\mathbf{u}_r^H[n], \tag{7.113}$$

$$\mathbf{R}_{rr} = \frac{1}{MNP}\sum_{n=0}^{MNP-1} \mathbf{u}_r[n]\mathbf{u}_r^H[n]. \tag{7.114}$$

The second competitor is the multiarray extension of the popular cross-correlation detector [58, 67] that uses the statistic

$$\mathscr{C} = |\operatorname{tr}\left(\mathbf{R}_{sr}^H\mathbf{R}_{sr}\right)| \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta. \tag{7.115}$$

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.

In addition to the two benchmark techniques presented in the previous paragraph, we also compare to the two-channel detectors presented in Chapter 5. To this end, we treat the J surveillance arrays with L antennas each as one large array with JL antennas and similarly, the reference arrays are treated as another large array with KL antennas. This inherently imposes a different structure to the covariance matrix under the null hypothesis. Specifically, under the null it is presumed that there is correlation across the J surveillance and across the K reference arrays, although there is no correlation noise across those arrays present. This effect is similar to that of assuming spatial correlation although arrays are uncorrelated as described in Chapter 6.

Note that other test statistics either need additional information such as the noise variance [51], which we assume to be unknown, or they only account for rank-one signals [70], which would be an unfair comparison for the competing technique. For these reasons we use the generalized versions of [67] and [58] for comparison.

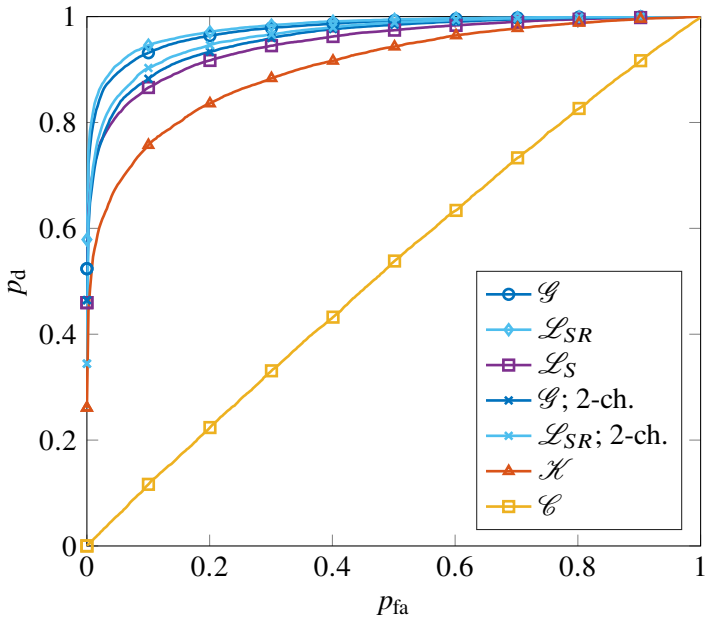

Figure 7.1: ROC curves in a scenario with $P = 2$, $N = 16$, $M = 64$, $L = L_I = 2$, a rectangular pulse, $\mathrm{SNR}_s = \mathrm{SNR}_r = -20dB$, and an equal number of surveillance and reference arrays $J = K = 5$.

In the first scenario that we investigate, we choose the following parameters: $P = 2$, $N = 16$, $M = 64$, $L = L_I = 2$, and a rectangular pulse. Moreover, we will vary the SNR and the number of surveillance arrays and reference arrays but we will keep the total number of arrays constant at 10.

In Figure 7.1 we investigate the ROC for $\mathrm{SNR}_s = \mathrm{SNR}_r = -20dB$ in the case of an equal number of arrays at surveillance and reference arrays, i.e., $J = K = 5$. We can observe that the GLRT $\mathscr{G}$ and $\mathscr{L}_{SR}$ perform best, where $\mathscr{L}_{SR}$ is slightly better than $\mathscr{G}$. The two-channel detectors from Chapter 5, i.e., the GLRT and the LMPIT-inspired test perform second best. They are followed by $\mathscr{L}_S$ and with a larger gap by $\mathscr{K}$, and the cross-correlation detector is almost the chance line. The reason for the poor performance of the cross-correlation detector is the low SNR regime as we have similarly seen in the previous chapter. Since the performance of this detector is particularly bad, we will neglect it in further figures as it does not provide further insight. This result resembles that of Chapter 5.7, where we have seen that the LMPIT-inspired test outperforms the GLRT although it is not an optimal test but the way the spectral correlation is measured seemed to be more beneficial to the detection performance compared to merging the spectral correlation at the SC only in an optimal way with the cross-spectral correlation between SC and RC. Here we can observe a similar effect. The gap between the multichannel detectors and the two-channel detectors can be explained by the fact that the two-channel detectors inherently presumes a different noise structure than the one present in the simulations. Specifically, the two-channel detectors overfit the underlying noise model as they presume more degrees of freedom than actually present.

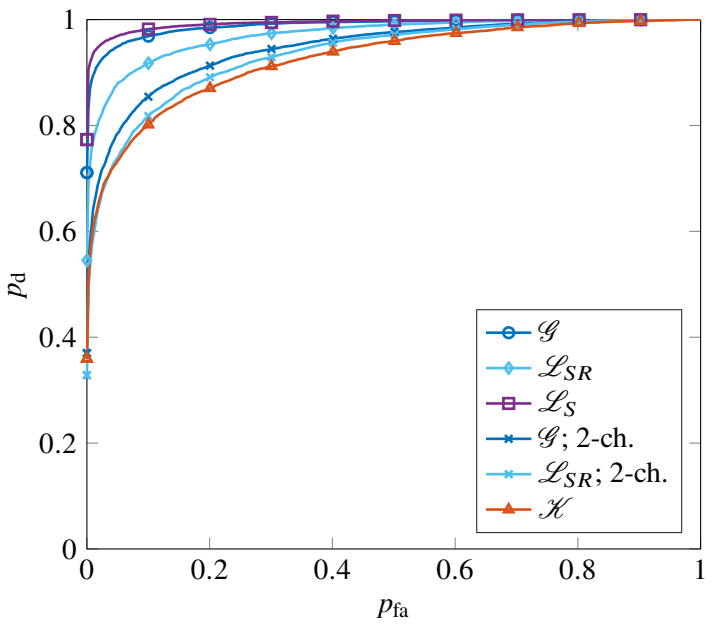

Figure 7.2: ROC curves in a scenario with $P = 2$, $N = 16$, $M = 64$, $L = L_I = 2$, a rectangular pulse, $\mathrm{SNR}_s = \mathrm{SNR}_r = -20dB$, $J = 8$ surveillance and $K = 2$ reference arrays.

Now we choose all parameters as before and keep the total number of arrays constant but increase the number of surveillance arrays to $J = 8$ and decrease the number of reference arrays to $K = 2$. Interestingly, the ROC in Figure 7.2 shows that the statistic $\mathscr{L}_S$ performs best among all competing detectors and is closely followed by the GLRT $\mathscr{G}$. At third place one can observe $\mathscr{L}_{SR}$, which is followed by the two-channel detectors. Note that the order of the two-channel GLRT and the two-channel LMPIT-inspired test has changed. This can be explained by the fact that the GLRT combines spectral correlations within the surveillance array with the cross-correlations between SC and RC. Since for this scenario we have more surveillance arrays than reference arrays it is more beneficial to include the inter-SC correlations rather than only accounting for the cross-spectral correlations only as it is done by the two-channel LMPIT-inspired test. This is also the reason for $\mathscr{L}_S$ outperforming $\mathscr{G}$ and $\mathscr{L}_{SR}$, i.e., it is more beneficial to only account for the inter-SC correlations if measured with the Frobenius norm rather than choosing the GLRT that combines the informations of all arrays but measures the correlations with the determinant. However, it is still superior to $\mathscr{L}_{SR}$, which only accounts for the cross-spectral correlations between SC and RC.

In a third scenario we still keep all parameters as before but now we choose less surveillance arrays ($J = 2$) but increase the number of reference arrays to $K = 8$. As illustrated in Figure 7.3 the multistatic and the two-channel detectors all perform almost equally, whereas the statistic $\mathscr{L}_S$ performs much worse compared to all detectors but $\mathscr{C}$, which is still the chance line. Since in this scenario we only have two surveillance arrays with $L = 2$ antennas each available, it is essential to incorporate the information provided by the cross-spectral correlations between SC and RC. The

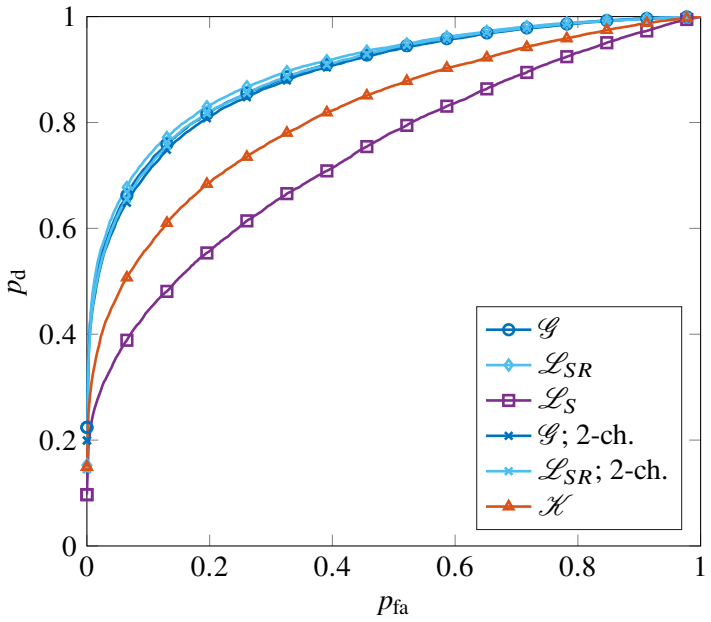

Figure 7.3: ROC curves in a scenario with $P = 2$, $N = 16$, $M = 64$, $L = L_I = 2$, a rectangular pulse, $\text{SNR}_s = \text{SNR}_r = -20dB$, $J = 2$ surveillance and $K = 8$ reference arrays.

reason for $\mathscr{G}$ and $\mathscr{L}_{SR}$ performing almost equally compared to their two-channel variants is that the influence of the noise model is for $J = 2$ with $L = 2$ almost negligible, i.e., the influence of a mismatched model is rather small for the given number of surveillance arrays.

Overall we can observe that for a given number of antennas arrays, it is more beneficial to have more surveillance than reference arrays, which is an intuitive result. Depending on the SNR and for a certain number of surveillance channels it appears to be beneficial to neglect the influence of the reference array and choosing $\mathscr{L}_S$, which only accounts for the inter-SC spectral correlations as a detector rather than the GLRT that fuses the information of both channels. This seems counterintuitive but considering that the GLRT and $\mathscr{L}_S$ determine the test statistics based on different functions and considering that the result is also dependent on the SNR at the reference channel, it is reasonable that one can find such a scenario. Nevertheless, for an equal number of surveillance and reference arrays it should be accounted for the cross-spectral correlations by $\mathscr{L}_{SR}$. In the following paragraph we study the influence of the unknown parameter γ from (7.94).

In Figures 7.4 to 7.6 we investigate the performance of the terms $\mathscr{L}_S$ and $\mathscr{L}_{SR}$ from (7.94) compared to its weighted sum $\mathscr{L}$ depending on the unknown γ, which would be the LMPIT if γ was known. To this end, we fix a set of values for γ and for each value we obtain p_d of $\mathscr{L}$ for a fixed SNR at the surveillance arrays $\text{SNR}_s = -20$ dB and for $\text{SNR}_r = \{-10, -15, -20\}$ dB. Those are displayed in blue. In yellow we illustrate $\mathscr{L}_{SR}$ and in red we illustrate $\mathscr{L}_S$. Both are independent of γ and, additionally, $\mathscr{L}_S$, which only accounts for the spectral correlations within and across the surveillance arrays, is independent of the SNR at the reference arrays.

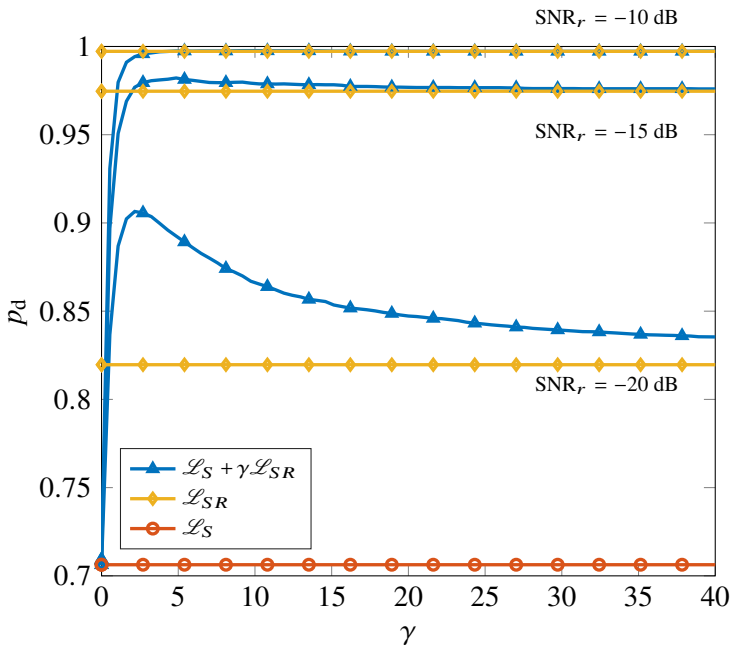

Figure 7.4: Probability of detection as a function of γ based on different detection statistics for an experiment with the following parameters: $P = 2$, $N = 16$, $M = 64$, $L = L_I = 2$, a rectangular pulse, $J = K = 5$ surveillance and reference arrays, and $\text{SNR}_s = -20$ dB, $\text{SNR}_r = \{-10, -15, -20\}$ dB, and $p_{\text{fa}} = 0.01$.

Generally, the results are similar to those we observed in Chapter 5.6. In all three figures we can observe that $\mathscr{L}_{SR}$ is close to the optimal detector $\mathscr{L}$ if the SNR_r is reasonably large, i.e., the cross-spectral correlations provide sufficiently many information for detection. The lower the SNR_r the more beneficial it is to also incorporate the spectral correlation within and across the surveillance arrays. This is especially well visible in Figures 7.4 and 7.5 comparing the yellow detector $\mathscr{L}_{SR}$ with the blue detector $\mathscr{L}$ for an $\text{SNR}_r = -20$ dB. In Figure 7.5 we can also observe that $\mathscr{L}_S$ performs better than $\mathscr{L}_{SR}$ for low $\text{SNR}_r = -20$ dB since the number of surveillance arrays is four times larger than the number of reference arrays. Hence, as long as the SNR at the reference array is too low it is more beneficial to account for the spectral correlations at the surveillance arrays only. In Figure 7.6, we investigate the case with four times as many reference as surveillance arrays. The cross-spectral correlations measured by $\mathscr{L}_{SR}$ perform almost as well as the optimal detector $\mathscr{L}$ even for low SNR_r. Obviously, they also significantly outperform $\mathscr{L}_S$. However, regarding the overall detection performance it is not too beneficial to increase the number of reference rather than the number of surveillance arrays.

In this paragraph we investigate the behavior of the GLRT statistic, the terms $\mathscr{L}_{SR}$ and $\mathscr{L}_S$ from (7.94) and the two-channel detectors from the previous chapter denoted by "2-ch" for different number of surveillance and reference arrays as a function of SNR_r for the following parameters $P = 2$, $N = 16$, $M = 64$, $L = L_I = 2$. In Figure 7.7 we show three different plots where the total

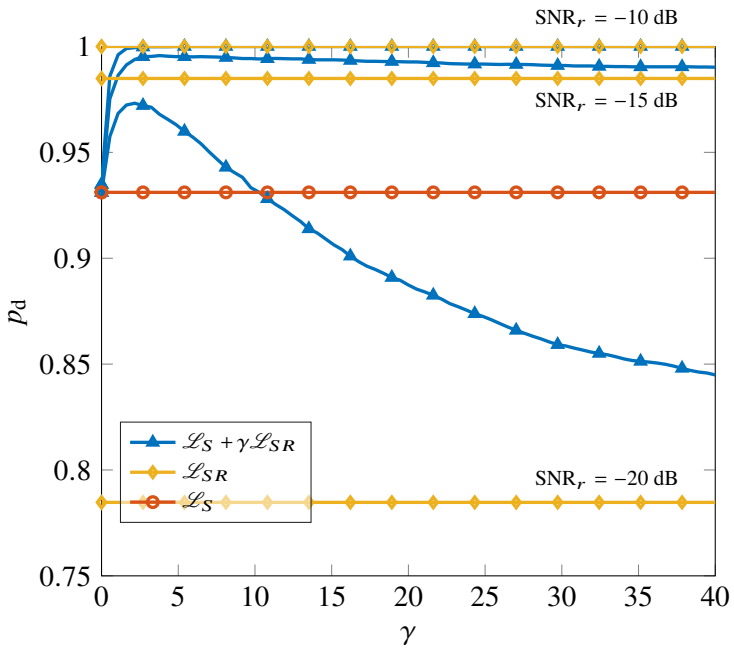

Figure 7.5: Probability of detection as a function of γ based on different detection statistics for an experiment with the following parameters: $P = 2$, $N = 16$, $M = 64$, $L = L_I = 2$, a rectangular pulse, $J = 8$ surveillance arrays and $K = 2$ reference arrays, and $\text{SNR}_s = -20$ dB, $\text{SNR}_r = \{-10, -15, -20\}$ dB, and $p_{\text{fa}} = 0.01$.

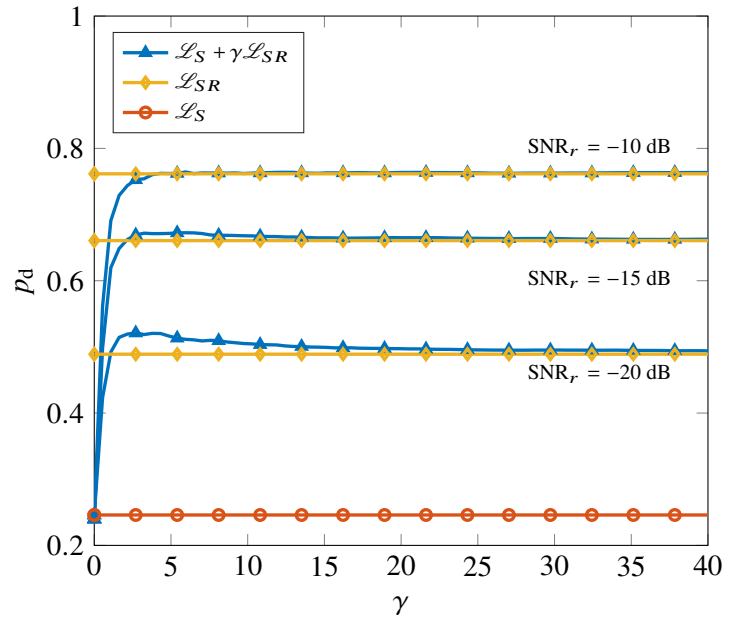

Figure 7.6: Probability of detection as a function of γ based on different detection statistics for an experiment with the following parameters: $P = 2$, $N = 16$, $M = 64$, $L = L_I = 2$, a rectangular pulse, $J = 2$ surveillance arrays and $K = 8$ reference arrays, and $\text{SNR}_s = -20$ dB, $\text{SNR}_r = \{-10, -15, -20\}$ dB, and $p_{\text{fa}} = 0.01$.

number of arrays is kept constant but the number of surveillance arrays varies from $J = 3$ in the top figure down to $J = 1$ in the bottom figure and the number of reference arrays from $K = 1$ in the top figure up to $K = 3$ in the bottom figure. In the top figure it can be observed that the GLRT and $\mathscr{L}_S$ perform equally well for low SNR. However, $\mathscr{L}_S$ is constant with respect to the SNR at the reference arrays, which can be observed in all three figures. The worst performance is given by the two-channel LMPIT-inspired test and $\mathscr{L}_{SR}$, whereas the two-channel GLRT performs medium best. This can be explained by the fact that only one reference array compared to three surveillance arrays is available. Hence, less information is contained in the cross-spectral coherences across surveillance and reference arrays and it is more beneficial to also include cross-spectral coherences within the three surveillance arrays.
The middle plot shows the case for an equal number of surveillance and reference arrays. Note that for low SNR at the reference array the GLRT performs best since it merges the information from both arrays including the inter-array cross-spectral correlations, whereas $\mathscr{L}_{SR}$ and $\mathscr{L}_S$ only exploit the cross-spectral correlations across surveillance and reference arrays and the inter-array cross spectral correlations of the surveillance array, respectively.
Finally, in the bottom plot only three statistics are shown for only one surveillance channel, the two-channel detector statistics are identical to the multistatic test statistics. For this specific setup, we can observe that GLRT and $\mathscr{L}_{SR}$ perform equally, whereas $\mathscr{L}_S$ has worse performance, which is not surprising since it is the single-array detector for cyclostationarity at the surveillance arrays only and no cross-spectral information across surveillance and reference arrays can be exploited.

In Figure 7.8 we consider the same setup as before but now we keep the SNR at the reference array constant at -12 dB and we vary the SNR at the surveillance array. We can observe that for $J = 3$ and $K = 1$ in the top figure and $J = 2$ and $K = 2$ the overall performance of $\mathscr{L}_{SR}$, the GLRT and the two-channel detectors is in a similar range. Where $\mathscr{L}_{SR}$ is a little more beneficial in the top figure, the gap to the other detectors shrinks for an equal number of reference and surveillance array. In all three constellations, the statistic $\mathscr{L}_S$ performs worse, which is not surprising considering that it only detects cyclostationarity based on the surveillance array observations. In the bottom plot we can observe that the GLRT and $\mathscr{L}_{SR}$ perform almost equally for the given constellation of $J = 1$ and $K = 3$, which can be explained by observing that due to only a single surveillance channel, the cross-spectral correlations across surveillance and reference channels is not too beneficial in $\mathscr{L}_{SR}$ and the additional information incorporated by the GLRT, i.e., the spectral correlation at the surveillance array only helps the GLRT but does not provide enough benefit to outperform $\mathscr{L}_{SR}$. Recall that $\mathscr{L}_{SR}$ and $\mathscr{L}_S$ use a different function (the Frobenius norm) to measure the coherence than the GLRT that uses the determinant, which was already shown in the previous chapter and also in [38] that measuring the coherence with the Frobenius norm seem to be more beneficial.

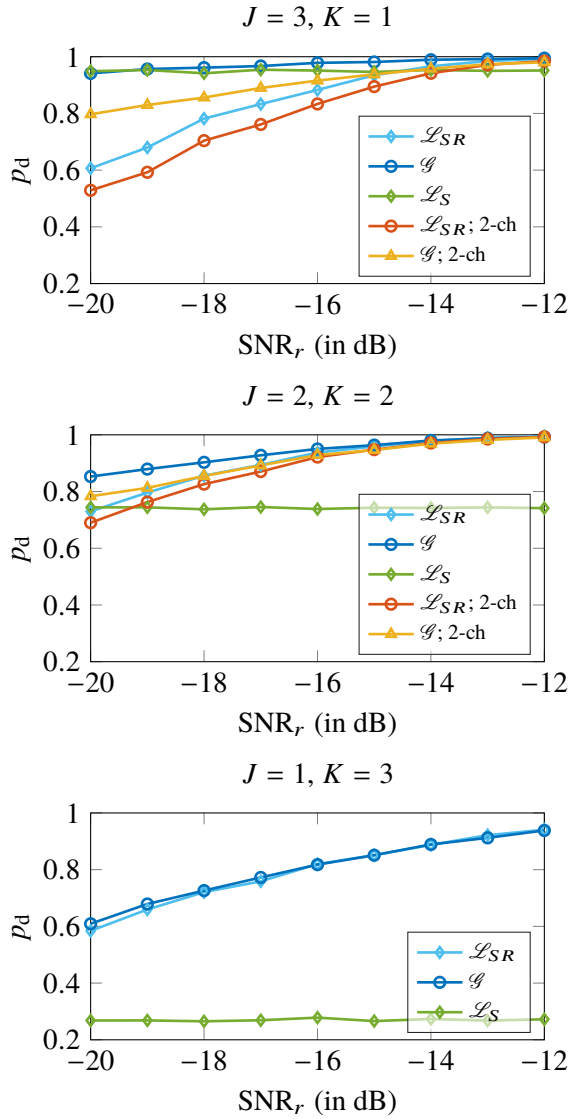

Figure 7.7: Probability of detection as a function of the SNR$_r$ and different constellation of numbers of surveillance and reference arrays for the following parameters $P = 2$, $N = 16$, $M = 64$, $L = L_I = 2$, and SNR$_s = -15$ dB.

In this section we have seen that depending on the number of surveillance and reference arrays it can be beneficial to employ $\mathcal{L}_{SR}$ as an "LMPIT-inspired" test, especially if the number of arrays are in a similar range. However, this observation cannot be generalized to arbitrary constellations and arbitrary SNR at the arrays as we have seen in the previous paragraphs, where applying the GLRT that inherently fuses the cross-spectral correlations inter-surveillance arrays with the cross-spectral correlation across surveillance and reference arrays.

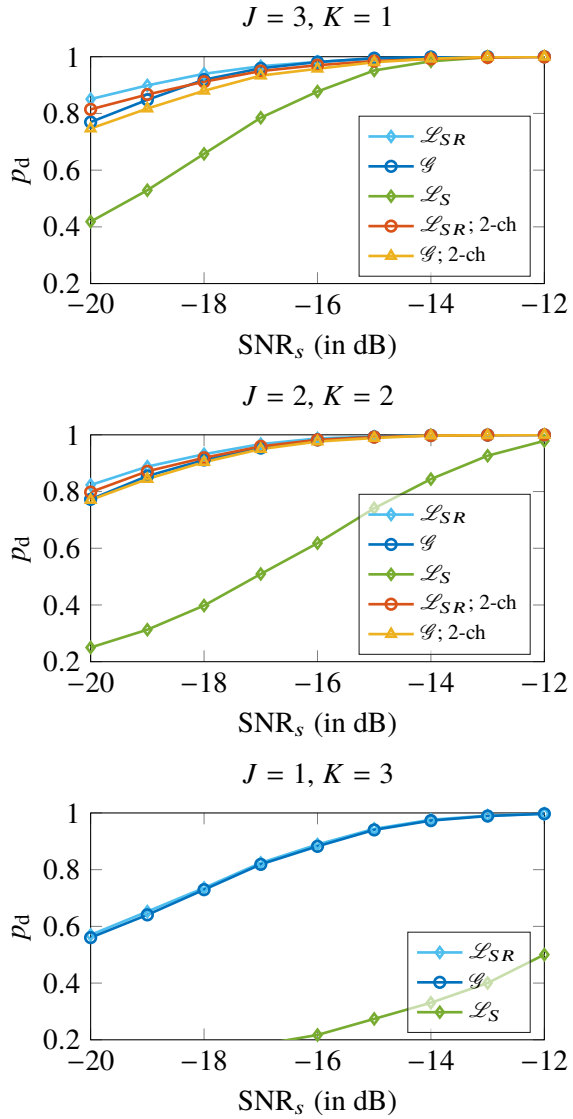

Figure 7.8: Probability of detection as a function of the SNR_s and different constellation of numbers of surveillance and reference arrays for the following parameters $P = 2$, $N = 16$, $M = 64$, $L = L_I = 2$, and $\text{SNR}_r = -12$ dB.

7.6 Conclusion

In this chapter we have derived the GLRT for the problem of detecting the presence of cyclostationarity at multiple surveillance channels given observations from the surveillance channels and multiple reference channels. This is a generalization of the two-channel passive detection problem that was considered in the previous chapter. We showed that the GLRT can be divided into one factor that measures the cross-spectral correlations within each and across all surveillance channels and a second factor that accounts for all cross-spectral correlations between all combinations of surveillance and reference arrays. Moreover, we have examined the existence of optimal invariant

tests for the same problem and have shown that neither the UMPIT nor the LMPIT exist for the problem. Nevertheless, the ratio of optimal invariant densities for close hypotheses is a weighted sum of two functions of cross-spectral coherence, where the first depends on the cross-spectral coherence across all surveillance arrays and the second depends on the cross-spectral coherence between all combinations of reference and surveillance arrays both measured by the Frobenius norm. However, the terms are connected by an unknown factor and therefore no LMPIT exists but still, we could investigate the detection performance of the two individual factors in order to analyze whether an LMPIT-inspired test can be proposed. Since the performance of the different statistics depend on the various involved parameters, it was not possible to consistently show that one of the factors would, in the majority of cases, be beneficial compared to, e.g., the GLRT for the same problem. For this reason, we did not propose an LMPIT-inspired test.

8 Summary

8.1 Conclusions

The results of this thesis are twofold. On the one hand, a detector has been derived for the problem of the detection of almost-cyclostationary signals in a single channel, which jointly provides an estimate of the cycle period. On the other hand, test statistics for passive detection of cyclostationary signals in one or multiple surveillance arrays given one or multiple reference arrays have been derived based on statistically established techniques. Specifically, an asymptotic GLRT for the problem of MIMO two-channel detection is derived. To this end, it was assumed that the received signals are zero-mean Gaussian and that the cycle period is known. The derivation of a GLRT requires to obtain the MLEs of unknown parameters, in the case of zero-mean Gaussian signals, covariance matrices. Due to the cyclostationarity the covariance matrices have a particular structure that is block-Toeplitz. However, there are no closed-form solutions for MLEs of block-Toeplitz matrices. For this reason, a previously proposed approach has been exploited that approximates the block-Toeplitz covariance matrix as a block circulant matrix in order to find a closed-form but asymptotic solution for the ML estimate. This concept has been applied throughout the thesis, not only for the two-channel detection case but also for its generalizations.

Furthermore, the existence of the UMPIT and the LMPIT for the same detection problems have been investigated. To this end, Wijsman's theorem is exploited, which avoids the necessity of deriving the maximal invariant statistic and its distribution under both hypotheses. Instead, all groups of invariant transformations are identified in order to come up with an expression for the ratio of the distributions of maximal invariant statistics. One can come rather quickly to the conclusion that a UMPIT does not exist for the studied detection problems. By considering the case of close hypotheses, i.e., a small cross-correlation between SC and RC, the existence of the LMPIT is analyzed. Although a simplified expression of the ratio of maximal invariant statistics can be found, it still depends on unknown parameters and it is concluded that an LMPIT does not exist either.

Nonetheless, an interpretation of both the GLRT and the ratio of the distribution of the maximal invariant statistics allows to propose another detector. It is shown that the GLRT inherently merges the information provided by 1) the presence of cyclostationarity at the SC via the coherence matrix computed based on observations at the SC only and 2) the cross-coherence matrix that

combines the usual cross-coherence with the cross-cyclic coherence between SC and RC. After performing extensive Monte Carlo simulations an LMPIT-inspired detector can be proposed that only incorporates the latter information. This detector has shown to have the capability to outperform the GLRT and both proposed detectors outperform comparable state-of-the-art detectors.

The two-channel detection problem has also been generalized to the case of multiple surveillance and multiple reference arrays. Similar to the two-channel case it is shown that the GLRT combines the cyclic (cross) coherence of all SC arrays as well as the cyclic (cross) coherences of all RC and SC in an optimal way. Moreover, it is shown that neither UMPIT nor LMPIT exist for this problem.

The previously mentioned derivations all consider the scenario of noise that reveals spatial and temporal correlation. Taking into account noise with spatio-temporal structure, the degrees of freedom of the detection problems are reduced. Therefore, three additional asymptotic GLRT statistics for the two-channel passive detection problem are derived for the following kinds of noise models: 1) temporally white and spatially uncorrelated, 2) temporally white and spatially correlated, and 3) temporally colored and spatially uncorrelated. It has been shown that, depending on the model, it can be crucial to account for the appropriate temporal structure of the noise when selecting a test statistic.

In order to apply all the aforementioned test statistics it is necessary to select an appropriate threshold. For the GLRT statistics, the stochastic representation of the null distribution that can be used to determine a threshold for a fixed probability of false alarm is derived. Specifically, by taking into account the invariant transformation that allows for whitening the observations, the distribution of the proposed GLRT statistics is shown to be a product of independent Beta distributed random variables, where the degrees of freedom of the Beta random variables depend on the number of surveillance and reference arrays, the number of antennas per array, the assumptions on noise, the cycle period, and the total number of samples.

Finally, a more generalized version of the single array detection problem has been considered. Specifically, it has been generalized to the detection of almost-cyclostationarity with unknown cycle period. This problem is formulated as a multiple hypothesis test combined with a resampling stage. To this end, the test statistic for the resampled signal is compared to a threshold, which is determined based on order-statistics and the fact that the statistic is distributed as a product of independent Beta random variables under the null hypothesis. It was possible to show that the proposed detector outperforms the comparable state-of-the-art detectors.

8.2 Outlook

The detectors derived in this thesis show the capability to improve detection performance in passive radar and cognitive radio scenarios by exploiting the (almost-) cyclostationary nature of communication signals. Here an overview of potential future directions are outlined.

One extension of the proposed two-channel and multichannel detectors is the consideration of unknown cycle periods and almost-cyclostationarity. As already pointed out in Chapter 4 sampling a continuous-time cyclostationary signal results in a discrete-time almost-cyclostationary signal. Similar to the technique derived in Chapter 4, it is possible to compute the derived GLRT statistics for a set of resampled signals and apply a multiple hypothesis test for the detection of ACS signals.

So far in this thesis attention is restricted to synchronized surveillance and reference signals with respect to the Doppler-shift, which itself is related to the target velocity. In practice this implies that the derived test statistics serve as ambiguity scores, i.e., it is scanned through candidate Doppler-shifts. Note that in our model the time-shift of the target echo may already be accounted for in the signal model through the frequency-selective channel. Another extension of the passive detection problem addresses assumptions made about the received signals. In practice it is required to account for direct-path interference in the surveillance channel since a complete cancellation cannot be guaranteed. However, dropping this assumption makes the signal at the surveillance channel not only cyclostationary, since the direct-path signal is cyclostationary, but also correlated with the reference channel under the null hypothesis. At the first glance this makes the problem much more complicated to solve. Yet one approach to consider is exploiting the Doppler-shift induced at the surveillance channel.

For a given cycle frequency and Doppler-shift, we can absorb all terms with different cycle frequency and Doppler-shift into the noise term. This also allows for the detection of multiple targets considering that they impose different Doppler-shifts. Similarly, direct-path-interference (DPI) and clutter can be absorbed into the noise term as they are related to zero Doppler-shifts, whereas all contributions made by targets have non-zero Doppler-shifts. Note that absorbing the DPI into the noise term, which is typically a strong signal compared to the signal of interest, implicates that the SNR in both channels has to be considered very small. Additionally, we have assumed so far that the true Doppler-shifts are among the candidates, i.e., the resolution in Doppler plane must be sufficiently large, and that the cycle period is an integer. Furthermore, this approach implies that all cyclostationary contributions that are absorbed into the noise term are treated as if they were wide-sense stationary for the given candidate cycle frequency.

Additionally, in this work we assumed so far that the number of transmit antennas is greater than or equal to the number of receive antennas. It would be beneficial to allow a large number of receive antennas as it would improve the signal separation performance. However, less transmit than receive

antennas induce a low-rank structure to the covariance matrices, which makes the derivations of the GLRT much more involved since in addition to the block-Toeplitz structure of the covariance matrices they are additionally of low-rank. In order to approach that problem, it could be possible to follow the lines in [125]. The derivation of the LMPIT for the same problem, however, is not straightforward.

Finally, more than a single illuminator of opportunity can be considered. Each receiver array observes a superposition of multiple direct-path signals and also multiple target-path signals at the reference and surveillance arrays, respectively. Assuming that each transmission signal is itself cyclostationary with a given cycle period, the superposition of the signals is also cyclostationary, where the cycle period is the least common multiple of all individual cycle periods. In this case, the problem can be approached by the statistics derived in this chapter, however, the cycle period may become rather large, which comes along with the necessity to increase the number of samples observed at each array.

List of Figures

List of Tables

List of publications

Journal articles

1. S. Horstmann, D. Ramírez, P.J. Schreier, "Joint Detection of Almost-Cyclostationary Signals and Estimation of Their Cycle Period", *IEEE Signal Processing Letters, 2018*
2. S. Horstmann, D. Ramírez, P.J. Schreier, "Two-Channel Passive Detection of Cyclostationary Signals", *IEEE Transactions on Signal Processing, 2020*

Conference contributions

1. S. Horstmann, D. Ramírez, P.J. Schreier, "Detection of Almost-Cyclostationarity : An Approach Based on a Multiple Hypothesis Test", *Proc. Asilomar Conference on Signals, Systems and Computers, 2017*
2. S. Horstmann, D. Ramírez, P.J. Schreier, "Two-channel passive detection exploiting cyclostationarity", *Proc. 27th European Signal Proc. Conf. (EUSIPCO), 2019*
3. S. Horstmann, D. Ramírez, P.J. Schreier, A. Pries "Two-channel passive detection of cyclostationary signals in noise with spatio-temporal structure", *Proc. Asilomar Conference on Signals, Systems and Computers, 2019*

Bibliography

[1] L. Gudzenko, "On periodic nonstationary processes," *Radio Eng. Electron. Phys. (in Russian)*, vol. 4, pp. 220–224, 1959.

[2] V. Lebedev, "On random processes having nonstationarity of periodic character," *Nauchn. Dokl. Vysshch. Shchk. Ser. Radiotekh. Elektron. (in Russian)*, vol. 2, pp. 32–34, 1959.

[3] E. D. Gladyshev, "Periodically correlated random sequences," *Soviet Math. Dokl.*, vol. 2, pp. 385–388, 1961.

[4] D. Ramírez, L. L. Scharf, J. Vía, I. Santamaría, and P. J. Schreier, "An asymptotic glrt for the detection of cyclostationary signals," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2014, pp. 3415–3419.

[5] H. L. Hurd, "Stationarizing properties of random shifts," *SIAM Journal on Applied Mathematics*, vol. 26, no. 1, pp. 203–212, 1974.

[6] W. A. Gardner and L. E. Franks, "Characterization of Cyclostationary Random Signal Processes," *IEEE Transactions on Information Theory*, vol. 21, no. 1, pp. 4–14, 1975.

[7] L. Izzo, A. Napolitano, and L. Paura, "Modified Cyclic Methods for Signal Selective TDOA Estimation," *IEEE Transactions on Signal Processing*, vol. 42, no. 11, pp. 3294–3298, 1994.

[8] L. Izzo and A. Napolitano, "Higher-order cyclostationarity properties of sampled time-series," *Signal Process.*, vol. 54, no. 3, pp. 303–307, 1996.

[9] E. Serpedin, F. Panduru, I. Sarı, and G. B. Giannakis, "Bibliography on cyclostationarity," *Signal Process.*, vol. 85, no. 12, pp. 2233–2303, 2005.

[10] W. A. Gardner, A. Napolitano, and L. Paura, "Cyclostationarity: Half a century of research," *Signal Process.*, vol. 86, no. 4, pp. 639–697, 2006.

[11] A. Napolitano, "Asymptotic normality of cyclic autocorrelation estimate with estimated cycle frequency," *Proc. 23rd European Signal Processing Conference (EUSIPCO)*, pp. 1481–1485, 2015.

[12] H. Amindavar and P. P. Moghaddam, "Estimation of propeller shaft rate and vessel classification in multipath environment," *Proceedings of the IEEE Sensor Array and Multichannel Signal Processing Workshop*, pp. 125–128, 2000.

[13] R. H. Jones and W. M. Brelsford, "Time Series with Periodic Structure," *Biometrika*, vol. 54, no. 3/4, pp. 403–408, 1967.

[14] R. Lund, H. Hurd, P. Bloomfield, and R. Smith, "Climatological time series with periodic correlation," *Journal of Climate*, vol. 8, no. 11, pp. 2787–2809, 1995.

[15] G. R. Dargahi-Noubary and A. G. Miamee, "Seismic waves and correlation autoregressive processes," *Mathematical Geology*, vol. 25, no. 6, pp. 671–688, 1993.

[16] E. Broszkiewicz-Suwaj, A. Makagon, R. Weron, and A. Wylomanska, "On detecting and modeling periodic correlation in financial data," *Physica A 336*, pp. 196–205, 2004.

[17] S. Bretteil and R. Weber, "Comparison between two cyclostationary detectors for RFI mitigation in radio astronomy," *European Signal Processing Conference (EUSIPCO)*, vol. 06-10-Sept, pp. 769–772, 2004.

[18] B. J. Currivan, "Satellite performance monitoring using digital measurement techniques," *American Institute of Aeronautics and Astronautics (AIAA)*, vol. 15, pp. 1156–1161, 1994.

[19] M. L. Boucheret, I. Mortensen, and H. Favaro, "Fast Convolution Filter Banks for Satellite Payloads with On-Board Processing," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 2, pp. 238–247, 1999.

[20] I. Antoniadis and G. Glossiotis, "Cyclostationary analysis of rolling-element bearing vibration signals," *Journal of Sound and Vibration*, vol. 248, pp. 829–845, 2001.

[21] R. A. Boyles and W. A. Gardner, "Cycloergodic Properties of Discrete-Parameter Nonstationary Stochastic Processes," *IEEE Transactions on Information Theory*, vol. 29, no. 1, pp. 105–114, 1983.

[22] 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.

[23] M. Abdulrahman and D. D. Falconer, "Cyclostationary Crosstalk Suppression by Decision Feedback Equalization on Digital Subscriber Loops," *IEEE Journal on Selected Areas in Communications*, vol. 10, no. 3, pp. 640–649, 1992.

[24] W. A. Gardner and S. Venkataraman, "Performance of optimum and adaptive frequency-shift filters for cochannel interference and fading," *Conference Record - Asilomar Conference on Circuits, Systems & Computers*, vol. 1, pp. 242–247, 1991.

[25] G. Gelli and F. Verde, "Blind LPTV joint equalization and interference suppression," *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, vol. 5, no. 5, pp. 2753–2756, 2000.

[26] G. Gelli, L. Paura, and F. Verde, "On the existence of FIR zero-forcing equalizers for nonredundantly precoded transmissions through FIR channels," *IEEE Signal Processing Letters*, vol. 12, no. 3, pp. 202–205, 2005.

[27] G. B. Giannakis, G. Zhou, and M. K. Tsatsanis, "On blind channel estimation with periodic misses and equalization of periodically varying channels," *Asilomar Conference on Signals, Systems and Computers*, 1992.

[28] J. K. Hwang and J. H. Chiou, "A New method for blind identification of multipath channel by exploiting signal cyclostationarity," *IEEE Signal Processing Workshop on Signal Processing Advances in Wireless Communications, SPAWC*, pp. 25–28, 1997.

[29] J. H. Lee and Y. T. Lee, "Robust adaptive array beamforming for cyclostationary signals under cycle frequency error," *IEEE Transactions on Antennas and Propagation*, vol. 47, no. 2, pp. 233–241, 1999.

[30] J. H. Lee and C. C. Huang, "Blind adaptive beamforming for cyclostationary signals: A subspace projection approach," *IEEE Antennas and Wireless Propagation Letters*, vol. 8, no. Lcmv, pp. 1406–1409, 2009.

[31] W. A. Gardner, "Simplification of MUSIC and ESPRIT by Exploitation of Cyclostationarity," *Proceedings of the IEEE*, vol. 76, no. 7, pp. 845–847, 1988.

[32] ——, "Identification of Systems with Cyclostationary Input and Correlated Input/Output Measurement Noise," *IEEE Transactions on Automatic Control*, vol. 35, no. 4, 1990.

[33] S. Enserink and D. Cochran, "On Detection of Cyclostationary Signals," *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, pp. 2004–2007, 1995.

[34] A. V. Dandawaté and G. B. Giannakis, "Statistical tests for presence of cyclostationarity," *IEEE Trans. on Signal Process.*, vol. 42, no. 9, 1994.

[35] J. Lundén, V. Koivunen, A. Huttunen, and H. V. Poor, "Collaborative cyclostationary spectrum sensing for cognitive radio systems," *IEEE Trans. Signal Process.*, vol. 57, no. 11, pp. 4182–4195, 2009.

[36] P. Urriza, E. Rebeiz, and D. Cabric, "Multiple antenna cyclostationary spectrum sensing based on the cyclic correlation significance test," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 11, pp. 2185–2195, 2013.

[37] H. Hurd and N. L. Gerr, "Graphical methods for determining the presence of periodic correlation," *J. Time Ser. Anal.*, vol. 12, no. 4, pp. 337–350, 1991.

[38] 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.

[39] J. Mitola and G. Q. Maguire, "Cognitive radio: making software radios more personal," *IEEE Personal Communications*, vol. 6, no. 4, pp. 13–18, 1999.

[40] A. Goldsmith, S. A. Jafar, I. Maric, and S. Srinivasa, "Breaking Spectrum Gridlock with Cognitive Radios : An Information Theoretic Perspective," *Proceedings of the IEEE*, vol. 97, no. 5, pp. 894–914, 2009.

[41] S. Haykin, D. J. Thomson, and J. H. Reed, "Spectrum sensing for cognitive radio," *Proceedings of the IEEE*, vol. 97, no. 5, pp. 849–877, 2009.

[42] E. Axell, G. Leus, E. G. Larsson, and H. V. Poor, "Spectrum Sensing for Cognitive Radio: State-of-the-Art and Recent Advances," *IEEE Signal Processing Magazine*, no. 3, pp. 101–116, 2012.

[43] Y. Zeng and Y.-C. Liang, "Robustness of the cyclostationary detection to cyclic frequency mismatch," *IEEE 21st Int. Symp. on Pers. Indoor and Mobile Radio Comm.*, pp. 2704–2709, 2010.

[44] G. Giannakis, "Cyclostationary Signal Analysis," in *The Digital Signal Processing Handbook*, V. K. Madisetti and D. B. Williams, Eds. Boca Raton: CRC Press LLC, 1999, ch. 17.

[45] D. Ramírez, J. Vía, I. Santamaría, and L. L. Scharf, "Locally most powerful invariant tests for correlation and sphericity of gaussian vectors," *IEEE Trans. on Inf. Theory*, vol. 59, no. 4, pp. 2128–2141, April 2013.

[46] H. Griffiths and N. Long, "Television-based bistatic radar," vol. 133, no. 7, pp. 649–657, 1986.

[47] 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.

[48] 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.

[49] 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.

[50] H. Griffiths and C. Baker, *An Introduction to Passive Radar*. Artech House, 2017.

[51] 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.

[52] 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," pp. 1–5, 2007.

[53] 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.

[54] 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.

[55] 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.

[56] 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.

[57] 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.

[58] 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.

[59] 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.

[60] 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.

[61] 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.

[62] 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.

[63] 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.

[64] 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.

[65] 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.

[66] 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.

[67] 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.

[68] 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.

[69] 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.

[70] 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.

[71] 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.

[72] 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.

[73] 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.

[74] 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.

[75] 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.

[76] L. L. Scharf, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*. Addison - Wesley, 1991.

[77] W. Gardner, "Statistical Spectral Analysis A Nonprobabilistic Theory," p. 99, 1986.

[78] E. Lehmann and J. P. Romano, *Testing Statistical Hypotheses*. Springer Series in Statistics, 2008.

[79] 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.

[80] H. Bohr, "Zur Theorie der Fast Periodischen Funktionen I." *Acta Math.*, vol. 45, pp. 29–127, 1925.

[81] J. Neyman and E. S. Pearson, "On the Problem of the Most Efficient Tests of Statistical Hypotheses," *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, vol. 231, pp. 289–337, 1933.

[82] J. R. Gabriel and S. M. Kay, "Use of Wijsman's theorem for the ratio of maximal invariant densities in signal detection applications," *Conference Record of the Asilomar Conference on Signals, Systems and Computers*, vol. 1, pp. 756–762, 2002.

[83] C. Stein, "Some problems in multivariate analysis part 1," *Technical Report No. 6*, 1956.

[84] S. Holm, "A simple sequentially rejective multiple test procedure," *Scand. J. Statist.*, vol. 6, no. 2, pp. 65–70, 1979.

[85] S. S. Wilks, "On the Independence of k Sets of Normally Distributed Statistical Variables," *Econometrica*, vol. 3, no. 3, p. 309, 1935.

[86] A. Leshem and A.-J. V. D. Veen, "Multichannel Detection of Gaussian Signals with Uncalibrated Receivers," *IEEE Signal Processing Letters*, vol. 8, no. 4, pp. 120–122, 2001.

[87] A. Leshem and A. J. Van der Veen, "Multichannel detection and spatial signature estimation with uncalibrated receivers," *Proc. 11th IEEE Workshop on Statistical Signal Processing*, no. 2, pp. 190–193, 2001.

[88] 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.

[89] J. Mauchly, "Significance Test for Sphericity of a Normal n-Variate Distribution," *The Annals of Mathematical Statistics*, vol. 11, no. 2, pp. 204–209, 1940.

[90] T. Anderson, *An Introduction to Multivariate Statistical Analysis*, 3rd ed. Hoboken, New Jersey: John Wiley & Sons, Inc., 2003.

[91] F. J. Marques and C. A. Coelho, "The multi-sample block-matrix sphericity test," *Proc. AIP Conference*, vol. 389, no. 1, 2011.

[92] 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.

[93] S. Horstmann, D. Ramirez, P. J. Schreier, and A. Pries, "Two-channel passive detection of cyclostationary signals in noise with spatio-temporal structure," in *Proc. 53rd Asilomar Conf. Signals, Syst., Comput.* Pacific Grove, CA, USA: IEEE, Nov. 2019, pp. 1333–1337.

[94] A. M. Mathai and R. K. Saxena, *Generalized Hypergeometric Functions with Applications in Statistics and Physical Sciences*. Berlin: Springer-Verlag, 1973, vol. 4, no. 2.

[95] S. Wilks, "The large-sample distribution of the likelihood ratio for testing composite hypotheses," *The Annals of Mathematical Statistics*, vol. 9, no. 1, pp. 60–62, 1938.

[96] B. R. Correia, C. A. Coelho, and F. J. Marques, "Likelihood ratio test for the hyper-block matrix sphericity covariance structure — Characterization of the exact distribution and development of near-exact distributions for the test statistic," *Revstat Statistical Journal*, vol. 16, no. 3, pp. 365–403, 2018. [Online]. Available: https://www.ine.pt/revstat/pdf/REVSTAT{_}v16-n3-6.pdf

[97] S. Horstmann, D. Ramírez, and P. J. Schreier, "Detection of almost-cyclostationarity : an approach based on a multiple hypothesis test," *Proc. Asilomar Conference on Signals, Systems and Computers*, 2017.

[98] S. Horstmann, D. Ramírez, and P. J. Schreier, "Joint detection of almost-cyclostationary signals and estimation of their cycle period," *IEEE Signal Processing Letters*, vol. 25, no. 11, pp. 1695–1699, Nov. 2018.

[99] S. Horstmann, D. Ramírez, and P. J. Schreier, "Two-Channel Passive Detection of Cyclostationary Signals," *IEEE Transactions on Signal Processing*, vol. 68, pp. 2340–2355, 2020.

[100] D. Ramírez, I. Santamaría, and L. L. Scharf, *Coherence: In Signal Processing and Machine Learning*, 1st ed. Springer International, 2023.

[101] J. Magnus and H. Neudecker, *Matrix Differential Calculus With Applications in Statistics and Econometrics*. New York: John Wiley & Sons, Inc., 1999.

[102] D. Basu, "On statistics independent of a complete sufficient statistic," *Indian Statistical Institute*, vol. 15, no. 4, pp. 377–380, 1955.

[103] S. Sedighi, A. Taherpour, J. Sala-Alvarez, and T. Khattab, "On the performance of hadamard ratio detector-based spectrum sensing for cognitive radios," *IEEE Transactions on Signal Processing*, vol. 63, no. 14, pp. 3809–3824, 2015.

[104] W. G. Cochran, "The distribution of quadratic forms The distribution of quadratic forms in a normal system," *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 30, pp. 178–191, 1934.

[105] N. R. Goodman, "Statistical Analysis Based on a certain Multivariate Complex Gaussian Distribution (an Introduction)," *The Annals of Mathematical Statistics*, 1963.

[106] C. Khatri, "Classical Statistical Analysis Based on a Certain Multivariate Complex Gaussian Distribution," *The Annals of Mathematical Statistics*, vol. 36, no. 1, pp. 98–114, 1965.

[107] R. W. Bailey, "Distributional Identities of Beta and Chi-Squared Variates : A Geometrical Interpretation," *The American Statistician*, vol. 46, no. 2, pp. 117–120, 1992.

[108] A. M. Kshirsagar, *Multivariate Analysis*. New York: Marcel Dekker, Inc., 1972.

[109] S. V. Schell and W. A. Gardner, "Detection of the number of cyclostationary signals in unknown interference and noise," *in Proc. Asilomar Conference on Signals, Systems and Computers*, pp. 3–7, 1990.

[110] T. Hentschel and G. Fettweis, "Sample rate conversion for software radio," *IEEE Communications Magazine*, no. August, pp. 142–150, 2000.

[111] N. Balakrishnan and A. Clifford Cohen, *Order Statistics and Inference*. Academic Press, INC., 1991.

[112] 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.

[113] 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.

[114] 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.

[115] 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.

[116] 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.

[117] A. Pries, D. Ramírez, and P. J. Schreier, "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.

[118] 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.

[119] 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.

[120] A. Hjørungnes, *Complex-Valued Matrix Derivatives: With Applications in Signal Processing and Communications*. Cambridge University Press, 2011.

[121] 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.

[122] 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.

[123] 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.

[124] J. P. Burg and D. G. Luenberger, "Robust estimation of structured covariance matrices," *IEEE Trans. on Signal Process.*, vol. 70, no. 9, pp. 963–974, Sept. 1982.

[125] D. Ramírez, G. Vazquez-Vilar, R. López-Valcarce, J. Vía, and I. Santamaría, "Detection of rank-P signals in cognitive radio networks with uncalibrated multiple antennas," *IEEE Trans. on Signal Process.*, vol. 59, no. 8, pp. 3764–3774, Aug. 2011.