UNIVERSIDAD DE CANTABRIA

DEPARTAMENTO DE
INGENIERÍA DE COMUNICACIONES

TESIS DOCTORAL

ESTIMA Y DETECCIÓN DE SERIES TEMPORALES MEDIANTE MÚLTIPLES SENSORES

Autor: David Ramírez García
Directores: Ignacio Santamaría Caballero
Javier Vía Rodríguez

2011

Estima y Detección de Series Temporales Mediante Múltiples Sensores

Tesis que se presenta para optar al título de

Doctor por la Universidad de Cantabria

Autor: David Ramírez García
Directores: Ignacio Santamaría Caballero
Javier Vía Rodríguez

Programa Oficial de Posgrado en Tecnologías de la Información y Comunicaciones en Sistemas de Telecomunicación

Grupo de Tratamiento Avanzado de Señal

Departamento de Ingeniería de Comunicaciones

Escuela Técnica Superior de Ingenieros Industriales y de Telecomunicación

Universidad de Cantabria

2011

© Copyright 2011 David Ramírez García

Tesis Doctoral:	Estima y Detección de Series Temporales Mediante Múltiples Sensores
Autor:	David Ramírez García
Directores:	Ignacio Santamaría Caballero Javier Vía Rodríguez

El tribunal nombrado para juzgar la tesis doctoral citada, compuesto por los señores

Presidente:

Vocales:

Secretario:

Acuerda otorgarle la calificación de

Santander, a de de 2011

Tamdiu discendum est, quamdiu vivas
(Debemos aprender mientras estemos vivos)

Seneca Philosophus
Epistulae morales

Afiliación

Grupo de Tratamiento Avanzado de Señal
Departamento de Ingeniería de Comunicaciones
Universidad de Cantabria

Este trabajo ha sido financiado por la beca FPU AP2006-2965 del Ministerio de Educación.

Agradecimientos

Escribir los agradecimientos de cualquier trabajo siempre es la parte más bonita. Por un lado significa que has acabado y, por otro, puedes agradecer a todos los que te han ayudado en su desarrollo. Además, en esta ocasión es más importante, si cabe, dado que finaliza una etapa que comenzó hace más de seis años, y comienza una nueva.

Para no ser menos que nadie, empezaré por mis directores: Ignacio Santamaría y Javier Vía. En un principio comenzó a dirigirme Nacho únicamente, pero pronto se incorporó Vía a la ardua tarea de dirigir a alguien como yo, con los constantes *no funciona*, *esto no vale pa na* o alguna cosa peor que no se puede escribir. Y ¿qué puedo decir de ellos? Pues sólo agradecerles todo el esfuerzo dedicado a este trabajo, que es tan mío como suyo. Y en un plano no académico, los habituales partidos de padel, las cenas GTAS, los cafes y las charlas hicieron más llevadero el trabajo, haciendo que estos años hayan pasado volando.

También merece una mención especial el Profesor Louis Scharf, culpable, en parte, de esta Tesis. Gracias a sus *Rejected* y *is not derived from first principles* contribuyó a definir el tema de este trabajo: la detección multicanal. Asimismo, quería agradecer al Profesor Peter Schreier su cálido recibimiento en la Universidad de Newcastle, Australia y la oportunidad que me brinda para continuar en el maravilloso mundo de la investigación en la Universidad de Paderborn, Alemania. Además, he tenido la suerte de colaborar con el Profesor Roberto López-Valcarce y con Gonzalo Vázquez-Vilar durante una mini-visita a Vigo.

Por otro lado, también quería agradecer a los miembros del GTAS, desde los profesores: los dos Jesuses y Luis, hasta postdocs y doctorandos: Chuspy, Foaud, Fredo, Gonzy, Steven y Víctor, incluyendo las últimas incorporaciones: Christian y Miguel. Todos ellos me han ayudado cuando han podido, y han conseguido que las horas de curro pasasen volando. Ahhh, se me olvidaban las viejas glorias: Bea, David FM, David JP, Gabri, Sara, Susana y alguno más que seguro se me pasa. Y no penséis que se me olvidan los no-GTAS: Jose, Mon y Óscar (electromagnético).

Mención aparte merecen todos los mitiquísimos tele[ñe]cos: Belén, Chus H., Chuspy (otra vez), Dar, Fonsín, Franuco, Gallosín, Lino, Lu, Moi, More, Nacho, Pablín, Quique, Rubio, Salas, Sara y Vali. Gracias a ellos, los cinco años de la carrera se hicieron cortísimos. Bien sea por hacer más amenas muchas de las asignaturas (cosa en ocasiones muy complicada), o por las aventuras que vivimos fuera de la Uni. Además, cada vez que nos juntamos, parece que no haya pasado el tiempo, pensando en nuevas liadas o recordando las antiguas.

Le debo mucho también a mis amigos. Los de toda la vida de Santander: Abraham, Álvaro, Fernando, Mario, Rey y los del Pueblo: Aingeru, Axel, Bruno, Demel, Felipe, Fer, Gabri, Javi, Jon, Ponga, Regi, Ruli y Ruth. Y no me quiero olvidar de otros que he conocido más tarde, y no por ello son menos importantes: Julia, Óscar, Oscarín y Vero. Todos ellos me han influido, esperemos que para bien, contribuyendo a ser quien soy.

Y, por último, a mi familia. Desde mis abuelos hasta mis padres, pasando por tíos y primos. Todos me han aguantado, cosa nada sencilla en muchas ocasiones. Además, quiero

recordar a uno de ellos que ya no está entre nosotros desde hace muchos años, cerca de quince, pero estoy seguro que ahora estaría muy orgulloso.

Seguro que se me queda gente por ahí, pero mi memoria ya no es lo que era (tanto pensar ...), por ello espero que no se lo tomen a mal. Y sin más, me despido. ¡Ha sido genial conoceros a todos! Nos vemos,

David

Santander, 13 de mayo de 2011

Resumen

El problema de detección de series temporales multivariadas surge en aplicaciones tan dispares como radar, sonar, ingeniería biomédica o comunicaciones. Aunque, sin lugar a dudas, dos de las aplicaciones más importantes en la actualidad son el sensado espectral multiantena para radio cognitiva (CR) y las redes de sensores. En el primer caso, el objetivo consiste en detectar *spectral holes*, es decir, bandas del espectro vacías, para permitir la transmisión oportunista de usuarios secundarios. En el segundo, a partir de las señales adquiridas por un conjunto de sensores, se quiere determinar qué modelo ha generado dichas observaciones.

Los problemas de detección multicanal se pueden resolver de diferentes maneras en función de la información *a priori* disponible. Por ejemplo, se puede disponer de información sobre las señales originales, también es posible diferenciar entre las hipótesis cuando éstas tienen diferentes estructuras temporales (espectros planos o coloreados, señales cicloestacionarias, etc.) o, dado que las señales son multivariadas, explotando diferencias en la estructura espacial de las mismas. En esta Tesis, se consideran los detectores basados en la estructura espacial, que es una característica especialmente interesante dado que permite desarrollar tests que necesitan muy poca información *a priori* sobre las señales.

El primer problema considerado es la detección de procesos contaminados por ruidos independientes e idénticamente distribuidos (IID). Bajo esta suposición, se obtiene el test basado en el cociente de verosimilitudes generalizado (GLRT) para señales con diferentes estructuras temporales y espaciales. En primer lugar, se deriva el GLRT para señales con densidades espectrales de potencia (PSD) planas y ninguna suposición adicional respecto a la estructura espacial. Posteriormente, se generaliza dicho modelo considerando señales con matriz de covarianza deficiente en rango. Estos dos modelos se extienden a procesos con PSD arbitrarias mediante la verosimilitud asintótica, evitando la estimación de máxima verosimilitud de matrices Toeplitz por bloques. Finalmente, se deriva el GLRT para señales no estacionarias.

El segundo problema consiste en la detección de series temporales contaminadas por ruidos independientes, pero no idénticamente distribuidos (no-IID). En este caso, también se deriva el GLRT para señales blancas sin estructura espacial, no obstante, el modelo de rango reducido conduce a un problema más complejo para ruidos no-IID. Debido a esto, se propone optimizar numéricamente la verosimilitud u obtener estimas aproximadas bajo la suposición de baja relación señal a ruido (SNR). Estos resultados se extienden a señales coloreadas temporalmente y, por último, se deriva el GLRT para señales no estacionarias.

El análisis de los tests derivados para señales con PSD arbitrarias contaminadas por ruidos no-IID muestra su relación con el espectro de coherencia (CS). Motivados por este hecho, en la última parte de la Tesis, se propone una generalización del espectro de coherencia (GCS) definida como una función Schur-convexa de los autovalores de la matriz compuesta por todos los CS por parejas. Además, se estudian sus principales propiedades y se proponen distintos estimadores del GCS. Por otro lado, se discute la relación del GCS con la información mutua de múltiples procesos y con los GLRT previamente derivados. Esta última relación permite

desarrollar detectores, con prestaciones diferentes al GLRT, basados en diferentes estimadores del GCS. Finalmente, el GCS es aplicado a un problema de ingeniería forestal.

En definitiva, en esta Tesis se aborda el problema de la detección de series temporales multivariadas bajo diferentes suposiciones acerca de su estructura espacio-temporal, y se derivan detectores basados en el cociente de verosimilitudes generalizado, cuyas prestaciones se evalúan mediante simulaciones de Monte Carlo.

Summary

The multiple-channel signal detection problem appears in many applications, such as radar and sonar, bioengineering, or communications. Nevertheless, two of the most important applications nowadays are multiantenna spectrum sensing and sensor networks. In the first case, we must detect spectral holes, that is, frequency sub-bands which are not occupied, to allow the opportunistic transmission of secondary users. In the second case, the goal consists in determining, from a set of observations, the model which better explains the data.

Multiple-channel detection problems may be solved following different approaches depending on the available information. For instance, the transmitted signals may be known in advance, or the different time structure under each hypothesis may be used (white vs. colored spectra, cyclostationary processes, etc.) or, in the case of multivariate time series, the spatial structure. In this Thesis, spatial-structure based detectors are considered, providing tests that use no or very limited *a priori* information about the signal to be detected.

Firstly, the detection of processes in independent and identically distributed (IID) noises is considered. Under this assumption, we derive the generalized likelihood ratio test (GLRT) for different time and spatial structures. The GLRT for time uncorrelated signals without further spatial structure is derived and, then, we tackle the detection of processes with rank-deficient covariance matrices. Both problems are generalized to time series with arbitrary spectra applying the asymptotic likelihood, which allows us to overcome the ML estimation of block-Toeplitz matrices. The last problem considered for IID noises is the detection of non-stationary time series.

Secondly, signals in independent but non identically distributed (non-IID) noises are considered. In this case, the GLRT for uncorrelated signals with unstructured covariance matrix is also obtained. However, for rank-deficient signal covariance matrices, the problem becomes more difficult, and we propose to numerically optimize the likelihood. Alternatively, we also derive an approximate solution for very low signal to noise ratio (SNR). These GLRT are generalized to signals with arbitrary power spectral densities (PSD) and, finally, we consider the detection problem for non-stationary processes.

The GLRT for bivariate signals with arbitrary PSD in non-IID noises reduces to a function of the coherence spectrum (CS). This prompts us to propose a generalization of the coherence spectrum (GCS), defined as a Schur-convex function of the eigenvalues of a matrix containing all the pairwise CS. Moreover, we discuss its properties and several GCS estimators are introduced. The relationship with the mutual information is presented and some GLRT are rewritten as functions of the GCS. This allows us to develop multiple-channel detectors based on different GCS estimators, which provides several alternatives to the GLRT. Finally, the GCS is applied to a forestry problem.

In summary, in this Thesis the multiple-channel signal detection problem is considered under several assumptions about the space-time structure of the signals, and detectors based on the generalized likelihood ratio are derived, whose performance is evaluated by means of Monte Carlo simulations.

Notación y Abreviaturas

Notación

a	Escalar
$\mathbf{a}$	Vector columna
$\mathbf{A}$	Matriz
$\mathcal{R}$	Matriz Toeplitz o Toeplitz por bloques
$\hat{(\cdot)}$	Estima de una matriz, vector o escalar
$\mathbf{I}$	Matriz identidad de las dimensiones adecuadas
$\mathbf{0}$	Matriz (o vector) de todo ceros de las dimensiones adecuadas
sup	Supremo de un conjunto
$\mathbb{S}_+$	Conjunto de matrices Hermíticas definidas positivas
$\mathbb{D}$	Conjunto de matrices diagonales
$\mathbb{D}_+$	Conjunto de matrices diagonales definidas positivas
$\mathbb{D}_{\mathrm{B}+}$	Conjunto de matrices diagonales por bloques (definidas positivas)
$\mathbf{x} \sim \mathcal{CN}(\boldsymbol{\mu}, \mathbf{R})$	$\mathbf{x}$ es un vector aleatorio complejo, circular y Gaussiano de media $\boldsymbol{\mu}$ y matriz de covarianza $\mathbf{R}$
$\mathbf{x} \overset{a}{\sim} \mathcal{N}(\boldsymbol{\mu}, \mathbf{R})$	$\mathbf{x}$ es un vector aleatorio con distribución asintótica Gaussiana de media $\boldsymbol{\mu}$ y matriz de covarianza $\mathbf{R}$
$\odot$	Producto de Hadamard (o elemento a elemento)
$\otimes$	Producto de Kronecker
$\mathcal{F}(\cdot)$	Transformada de Fourier en tiempo discreto
$\lVert\mathbf{A}\rVert_F$	Norma Frobenius de la matriz $\mathbf{A}$
$\det(\mathbf{A})$	Determinante de la matriz $\mathbf{A}$
$\mathrm{tr}(\mathbf{A})$	Traza de la matriz $\mathbf{A}$

Abreviaturas

ADC	Conversión analógico-digital (*Analog to digital conversion*)
ALS	Mínimos cuadrados alternados (*Alternating least squares*) [optimización]
AR	Autoregresivo (*Autoregressive*) [proceso estocástico]
ARMA	Autoregresivo media móvil (*Autoregressive moving average*) [proceso estocástico]
CR	Radio cognitiva (*Cognitive radio*)
CS	Espectro de coherencia (*Coherence spectrum*)
CCA	Análisis de correlaciones canónicas (*Canonical correlation analysis*)

DC	Clases de descomposición (*Decay classes*)
DL	Técnicas de Regularización (*Diagonal loading*)
DVB	Difusión de video digital (*Digital video broadcasting*)
ECG	Electrocardiográfica (*Electrocardiographic*) [Señal]
ED	Detector de energía (*Energy detector*)
EEG	Electroencefalográfica (*Electroencephalographic*) [Señal]
EV	Valores propios (*Eigenvalues*) [problema]
EVD	Descomposición en valores propios (*Eigenvalue decomposition*)
FIR	Respuesta al impulso finita (*Finite impulse response*) [filtro]
fMRI	Imagen por resonancia magnética funcional (*functional magnetic resonance imaging*)
GEV	Valores propios generalizados (*Generalized eigenvalues*) [problema]
GCS	Espectro de coherencia generalizado (*Generalized coherence spectrum*)
GLR	Cociente de verosimilitudes generalizado (*Generalized likelihood ratio*)
GLRT	Test basado en el cociente de verosimilitudes generalizado (*Generalized likelihood ratio test*)
IEEE	Instituto de ingenieros eléctricos y electrónicos (*Institute of electrical and electronics engineers*)
IID	Independientes e idénticamente distribuidas (*Independent and identically distributed*)
INMSPS	Ruidos idénticos, señales estacionarias con estructura espacial de rango P y canales planos *(IID noises, stationary signals with rank-P spatial structure and flat-fading channels*) [modelo]
INNSNS	Ruidos idénticos y señales no estacionarias sin estructura espacial *(IID noises, non stationary signals without spatial structure*) [modelo]
INSSNS	Ruidos idénticos y señales estacionarias sin estructura espacial *(IID noises, stationary signals without spatial structure*) [modelo]
INSSPS	Ruidos idénticos y señales estacionarias con estructura espacial de rango P *(IID noises, stationary signals with rank-P spatial structure*) [modelo]
INWSNS	Ruidos idénticos y señales estacionarias blancas sin estructura espacial *(IID noises, white stationary signals without spatial structure*) [modelo]
INWSPS	Ruidos idénticos y señales estacionarias blancas con estructura espacial de rango P *(IID noises, white stationary signals with rank-P spatial structure*) [modelo]
KKT	Karush-Kuhn-Tucker [condiciones]
LMP	Localmente más potente (*Locally most powerful*) [test]
LMPI	Localmente más potente invariante (*Locally most powerful invariant*) [test]
LS	Mínimos cuadrados (*Least squares*) [estimación]
LTE	Evolución a largo plazo (*Long term evolution*) [Estándar]
MA	Media móvil (*Moving average*) [proceso estocástico]
MAXVAR	Máxima varianza (*Maximum variance*) [generalización de CCA]
MF	Filtro adaptado (*Matched filter*)
MIMO	Múltiples entradas y salidas (*Multiple-input multiple-output*)
ML	Máxima verosimilitud (*Maximum likelihood*) [estimador]
MSE	Error cuadrático medio (*Mean square error*)
MUSIC	Clasificación múltiple de señales (*Multiple signal classification*)
MVDR	Mínima varianza sin distorsión (*Minimum variance distortionless response*) [filtro]

NNNSNS	Ruidos con distribuciones diferentes y señales no estacionarias sin estructura espacial *(Non-IID noises, non stationary signals without spatial structure)* [modelo]
NNSSNS	Ruidos con distribuciones diferentes y señales estacionarias sin estructura espacial *(Non-IID noises, stationary signals without spatial structure)* [modelo]
NNSSPS	Ruidos con distribuciones diferentes y señales estacionarias con estructura espacial de rango P *(Non-IID noises, stationary signals with rank-P spatial structure)* [modelo]
NNWSNS	Ruidos con distribuciones diferentes y señales estacionarias blancas sin estructura espacial *(Non-IID noises, white stationary signals without spatial structure)* [modelo]
NNWSPS	Ruidos con distribuciones diferentes y señales estacionarias blancas con estructura espacial de rango P *(Non-IID noises, white stationary signals with rank-P spatial structure)* [modelo]
NS	No estacionario (*Non stationary*)
OFDM	Multiplexación por división de frecuencias ortogonales (*Orthogonal frequency domain multiplexing*)
PCA	Análisis de componentes principales (*Principal component analysis*)
PDF	Función densidad de probabilidad (*Probability density function*)
PDP	Perfil de potencia de los retardos (*Power delay profile*)
PSD	Densidad espectral de potencia (*Power spectral density*)
ROC	Característica de operación del receptor (*Receiver operating characteristic*) [curva]
SISO	Una entrada y una salida (*Single-input single-output*)
SIMO	Una entrada y múltiples salidas (*Single-input multiple-output*)
SNR	Relación señal a ruido (*Signal to noise ratio*)
SVD	Descomposición en valores singulares (*singular value decomposition*)
UMP	Uniformemente más potente (*Uniformly most powerful*) [test]
UMPI	Uniformemente más potente invariante (*Uniformly most powerful invariant*) [test]
WSS	Estacionario en sentido amplio (*Wide sense stationary*) [proceso]

Índice general

Agradecimientos ix

Resumen xi

Summary xiii

Notación y Abreviaturas xv

Índice general xxi

1. Introducción **3**
1.1. Objetivos de la Tesis . . . 3
1.2. Organización y Contribuciones de la Tesis . . . 4

2. Detección en Series Temporales Multivariadas **7**
2.1. Introducción . . . 7
2.2. Modelo de Señal . . . 9
2.2.1. Ruidos con Distribuciones Idénticas . . . 10
2.2.2. Ruidos con Distribuciones Diferentes . . . 12
2.3. Técnicas de Detección . . . 13
2.3.1. Test Uniformemente más Potente (UMP) . . . 14
2.3.2. Test Localmente más Potente (LMP) . . . 15
2.3.3. Test Invariantes . . . 15
2.3.4. Test basado en el Cociente de Verosimilitudes Generalizado (GLRT) . . . 16
2.3.5. Tests de Rao y de Wald . . . 17
2.4. Aplicaciones de la Detección Multicanal . . . 18
2.4.1. Radio Cognitiva . . . 19
2.5. Conclusiones . . . 21

3. Detección Multicanal de Procesos Contaminados con Ruidos IID **23**
3.1. Detección Multicanal de Procesos Estacionarios Blancos . . . 23
3.1.1. Señales sin Estructura Espacial . . . 24
3.1.2. Señales con Estructura Espacial . . . 27
3.2. Detección Multicanal de Procesos Estacionarios con PSD Arbitrarias . . . 31
3.2.1. La Verosimilitud Asintótica . . . 32
3.2.2. Señales sin Estructura Espacial . . . 33
3.2.3. Señales con Estructura Espacial . . . 35
3.2.4. Señales Coloreadas en Canales Planos . . . 36
3.3. Detección Multicanal de Procesos No Estacionarios . . . 42
3.3.1. GLRT en el Dominio de la Frecuencia . . . 44
3.4. Resultados Numéricos . . . 45
3.4.1. Efecto del Rango P . . . 47
3.4.2. Detección de Series Temporales con PSD Arbitrarias . . . 48
3.4.3. Análisis del Detector en el Dominio de la Frecuencia . . . 51
3.4.4. Aplicación a Apuntamiento de Haz . . . 52

3.4.5. Detección de Series Temporales no Estacionarias . . . 54
3.4.6. Aplicación a Radio Cognitiva . . . 55
3.5. Conclusiones . . . 57

4. Detección Multicanal de Procesos Contaminados por Ruidos no-IID 59
4.1. Detección Multicanal de Procesos Estacionarios Blancos . . . 60
4.1.1. Señales sin Estructura Espacial . . . 60
4.1.2. Señales con Estructura Espacial . . . 62
4.1.3. Optimización Alternada . . . 64
4.1.4. Aproximación para Baja SNR . . . 65
4.2. Detección Multicanal de Procesos Estacionarios con PSD Arbitrarias . . . 66
4.2.1. Señales sin Estructura Espacial . . . 67
4.2.2. Señales con Estructura Espacial . . . 68
4.3. Detección de Correlación Espacial en Series Temporales No Estacionarias . . . 69
4.3.1. GLRT en el Dominio de la Frecuencia . . . 71
4.4. Resultados Numéricos . . . 71
4.4.1. Efecto del Rango P . . . 72
4.4.2. Comportamiento del GLRT para Baja SNR . . . 74
4.4.3. Impacto de la Distribución del Ruido . . . 74
4.4.4. Detección de Series Temporales con PSD Arbitrarias . . . 75
4.4.5. Determinante frente a la Norma Frobenius . . . 77
4.4.6. Señales No Estacionarias . . . 77
4.4.7. Aplicación a Radio Cognitiva . . . 79
4.5. Conclusiones . . . 79

5. Nuevas Medidas Multicanal de Correlación en el Dominio Frecuencial 81
5.1. El Espectro de Coherencia . . . 82
5.1.1. Definición y Propiedades . . . 82
5.1.2. Estimadores del Espectro de Coherencia . . . 83
5.2. El Espectro de Coherencia Generalizado . . . 86
5.2.1. Propiedades . . . 89
5.2.2. Estimación del GCS-MAX . . . 92
5.3. Relaciones entre el GCS y otras Medidas Estadísticas . . . 94
5.3.1. Relación con el GLRT de Señales Coloreadas Contaminadas por Ruidos no-IID 94
5.3.2. Relación con la Información Mutua de L Series Temporales . . . 95
5.4. Resultados Numéricos . . . 96
5.4.1. Comparación del GCS-MAX, GCS-MIN y GCS-DET . . . 97
5.4.2. Comparación de los Diferentes Estimadores del GCS-MAX . . . 98
5.4.3. El GCS-MAX en Problemas de Detección . . . 100
5.4.4. Una Aplicación del Espectro de Coherencia Generalizado en Ingeniería Forestal 103
5.5. Conclusiones . . . 103

6. Conclusiones y Líneas Futuras 107
6.1. Conclusiones . . . 107
6.2. Líneas Futuras de Investigación . . . 108
6.2.1. Extensiones para otros modelos . . . 108
6.2.2. Obtención del UMPIT y del LMPIT . . . 109
6.2.3. Evaluación de las prestaciones sobre una plataforma MIMO Hardware . . . 110
6.2.4. Otras aplicaciones del GCS . . . 110
6.2.5. Estimación multicanal . . . 111

A. Demostración del Teorema 3.7 113

B. Mayorización **115**
B.1. Definición . . . 115
B.2. Funciones Schur-convexas y Schur-cóncavas . . . 116
B.2.1. Condición de Schur . . . 117

C. Análisis de correlaciones canónicas **119**
C.1. Formulación Clásica . . . 119
C.1.1. Soluciones CCA Sucesivas . . . 120
C.2. Generalizaciones a Varios Conjuntos de Datos . . . 120
C.2.1. Máxima Varianza (MAXVAR) . . . 121
C.2.2. Mínima Varianza (MINVAR) . . . 121
C.2.3. Otras Generalizaciones . . . 122

D. Publicaciones **123**
D.1. Publicaciones derivadas de la Tesis . . . 123
D.1.1. Revistas Internacionales . . . 123
D.1.2. Congresos Internacionales . . . 123
D.2. Otras Publicaciones . . . 124
D.2.1. Revistas Internacionales . . . 124
D.2.2. Congresos Internacionales . . . 124

Índice de Figuras **128**

Índice de Tablas **129**

Índice de Algoritmos **131**

Bibliografía **133**

Capítulo 1

Introducción

1.1. Objetivos de la Tesis

Los problemas de detección y estimación de series temporales multicanal o multivariadas surgen en un gran número de aplicaciones. Se pueden encontrar, por ejemplo, en radar con múltiples antenas [Fishler et al., 2004a, Li y Stoica, 2008], sonar [Azimi-Sadjadi et al., 2000], bioingeniería o análisis de señales médicas [Luesma, 2006, Friston et al., 1994], radioastronomía [Leshem et al., 2000, Leshem y Van der Veen, 2001b, Leshem y Van der Veen, 2001a] y sensado espectral para radio cognitiva [Mitola y Maguire Jr., 1999, Hossain y Bhargava, 2007, Chen y Prasad, 2009].

Lógicamente, en función de la información disponible acerca de las señales a detectar es posible derivar distintos detectores. Por ejemplo, se puede disponer de información sobre las señales originales o explotar ciertas características como la cicloestacionariedad o diferencias en la estructura temporal de las observaciones bajo ambas hipótesis. El uso de múltiples sensores puede inducir también estructura espacial que puede ser empleada para derivar los detectores. Esta característica es especialmente útil ya que, como se verá, permite derivar detectores que necesitan muy poca información *a priori* sobre las señales. De hecho, bajo la suposición de ruidos independientes, los detectores propuestos en la Tesis sólo necesitan conocer el número de señales transmitidas. Por lo tanto, el objetivo principal de la Tesis consiste en el desarrollo de tests para problemas de detección multicanal bajo diferentes supuestos acerca de la estructura temporal y espacial de los procesos y diferentes suposiciones acerca de la distribución del ruido.

El primer caso considera ruidos independientes e idénticamente distribuidos (IID). Bajo esta suposición, se deriva, en primer lugar, los tests para señales estacionarias en sentido amplio (WSS) con densidades espectrales de potencia (PSD) planas, con y sin estructura espacial, respectivamente. En particular, consideramos detectores basados en el cociente de verosimilitudes generalizado (GLRT), dado que las hipótesis son compuestas. A continuación, los GLRT anteriores son generalizados a series temporales con PSD arbitrarias. Sin embargo, para derivar los tests es necesario obtener las estimas de máxima verosimilitud (ML) de matrices Toeplitz por bloques, que es un problema de optimización no convexo. Debido a esto, se propone usar la verosimilitud asintótica, dada en el dominio frecuencial. El último test bajo ruidos IID trata el caso de señales no estacionarias.

En segundo lugar, el problema anterior se generalizas considerando ruidos independientes, pero no idénticamente distribuidos (no-IID). Comenzamos por derivar los GLRT para series temporales con PSD planas, con y sin estructura espacial. Para señales sin estructura espacial se obtiene una fórmula cerrada para el GLRT, sin embargo, incorporar estructura espacial

al problema implica que no se pueda encontrar solución cerrada. En la Tesis se proponen dos maneras de abordar dicho problema: (i) una técnica de optimización numérica basada en minimización alternada y (ii) una aproximación para baja relación señal a ruido (SNR). Los detectores anteriores se extienden, una vez más, al caso de procesos con PSD arbitrarias. El análisis de los problemas de detección con ruidos no-IID finaliza derivando el GLRT para determinar si una serie temporal multivariada no estacionaria tiene sus componentes incorreladas espacialmente.

Para el caso de series temporales bidimensionales, algunos de los GLRT presentados previamente se reducen a funciones del espectro de coherencia (CS). Motivados por este hecho, se propone una generalización del espectro de coherencia (GCS) para más de dos series temporales, definida como una función de los autovalores de la matriz que contiene todos los CS por parejas. Concretamente, para obtener una generalización adecuada del CS, se consideran funciones Schur-convexas, acotadas entre 0 y 1, las cuales para dos series temporales se reducen a la definición clásica del espectro de coherencia. Asimismo, también se presentan sus propiedades y estimadores. Por otro lado, se discute la relación entre el GCS y la información mutua, así como, su relación con los GLRT de algunos modelos considerados previamente. La interpretación de los GLRT como funciones del GCS, permite derivar otros detectores, empleando distintos estimadores del GCS.

1.2. Organización y Contribuciones de la Tesis

La Tesis se organiza de la siguiente manera.

- El Capítulo 2 introduce y motiva el problema de la detección multicanal, planteado como un test de hipótesis compuestas, y presenta un resumen de la teoría de la detección. En concreto, se introducen brevemente los detectores para tests binarios de hipótesis compuestas, debido a que son los tests que aparecen en la Tesis. Posteriormente, se introduce el modelo de señal y las distintas particularizaciones consideradas en este trabajo. El capítulo concluye comentando posibles aplicaciones del problema de detección multicanal, haciendo especial énfasis en un nuevo paradigma de comunicaciones conocido como radio cognitiva.

- El Capítulo 3 aborda la detección de procesos estocásticos contaminados por ruidos independientes e idénticamente distribuidos (IID). En primer lugar se considera la detección de procesos estacionarios con densidades espectrales de potencia (PSD) planas sin estructura espacial adicional. Dicho test, conocido como test de esfericidad, fue derivado en [Mauchly, 1940] para señales reales y Gaussianas, siendo su cociente de verosimilitudes generalizado (GLRT) el cociente de las medias geométricas y aritméticas de la matriz de covarianza muestral. A continuación, se extiende dicho modelo al caso de matrices de covarianza compuestas por la suma de una matriz deficiente en rango más una versión escalada de la matriz identidad. Una vez más, el detector propuesto para este test de hipótesis es una función de los autovalores de la matriz de covarianza muestral. Por otro lado, también se considera la detección de procesos con PSD arbitrarias. La derivación del GLRT para este problema es bastante complicada, requiriendo la estima de máxima verosimilitud (ML) de matrices Toeplitz por bloques que, como se sabe, es un problema de optimización no convexo. Para solventar dicho problema se usa la verosimilitud asintótica, definida en el dominio frecuencial, y que converge en error cuadrático medio a la verosimilitud convencional. Haciendo uso de

la verosimilitud asintótica, se derivan los correspondientes GLRT. El capítulo concluye considerando el problema de detección para series temporales no estacionarias. En este caso, no es necesario recurrir a la verosimilitud asintótica debido a que las matrices de covarianza no son Toeplitz.

Los resultados presentados en este capítulo han dado lugar a las siguientes publicaciones:

- D. Ramírez, G. Vazquez-Vilar, R. López-Valcarce, J. Vía y I. Santamaría. "Detection of rank-P signals in cognitive radio networks with uncalibrated multiple antennas". *IEEE Transactions on Signal Processing (Aceptado)*, 2011.
- D. Ramírez, J. Vía, I. Santamaría y L. L. Scharf. "Multi-sensor beamsteering based on the asymptotic likelihood for colored signals". *IEEE Workshop on Statistical Signal Processing (SSP 2011)*, Niza, Francia, Junio 2011.
- D. Ramírez, J. Vía, I. Santamaría y L. L. Scharf. "Multiple-channel detection of a Gaussian time series over frequency-flat channels". *IEEE Int. Conf. on Acoust., Speech, and Signal Processing (ICASSP 2011)*, República Checa, Mayo 2011.
- D. Ramírez, J. Vía y I. Santamaría. "Multiantenna spectrum sensing: The case of wideband rank-one primary signals". *6th IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM 2010)*, Israel, Octubre 2010.

- El caso de ruidos independientes, pero no idénticamente distribuidos (no-IID), se considera en el Capítulo 4. Al igual que en el capítulo anterior se comienza considerando la detección de procesos estocásticos con PSD planas, cuyo GLRT fue derivado en [Wilks, 1935] y que viene dado por el cociente de Hadamard de la matriz de covarianza muestral. Este modelo se extiende al caso de señales con estructura espacial, demostrándose que no existe una solución cerrada para el GLRT. No obstante, se proponen dos alternativas. La primera de ellas consiste en la optimización numérica de la verosimilitud. Concretamente, se propone una técnica basada en minimización alternada, donde sendos problemas de optimización son convexos. Por otro lado, motivados por el alto coste computacional de la técnica de optimización numérica, se deriva una fórmula cerrada para el GLRT bajo la suposición de baja relación señal a ruido (SNR). A continuación, se considera la detección de series temporales con PSD arbitrarias usando la verosimilitud asintótica. Por último, se aborda el problema de detección de correlación espacial en series temporales no estacionarias.

  Los resultados presentados en este capítulo han dado lugar a las siguientes publicaciones:[1]

  - D. Ramírez, G. Vazquez-Vilar, R. López-Valcarce, J. Vía y I. Santamaría. "Detection of rank-P signals in cognitive radio networks with uncalibrated multiple antennas". *IEEE Transactions on Signal Processing (Aceptado)*, 2011.
  - D. Ramírez, J. Vía, I. Santamaría and L. L. Scharf. "Detection of spatially correlated Gaussian time series". *IEEE Transactions on Signal Processing*, vol.58, no.10, pp.5006-5015, Octubre 2010.
  - D. Ramírez, G. Vazquez-Vilar, R. López-Valcarce, J. Vía y I. Santamaría. "Multiantenna detection under noise uncertainty and primary users spatial structure". *IEEE Int. Conf. on Acoust., Speech, and Signal Processing (ICASSP 2011)*, República Checa, Mayo 2011.

[1] Nótese que algunas publicaciones han contribuido a varios capítulos.

- D. Ramírez, J. Vía y I. Santamaría. "Multiantenna spectrum sensing: The case of wideband rank-one primary signals". *6th IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM 2010)*, Israel, Octubre 2010.
- D. Ramírez, J. Vía, I. Santamaría, R. López-Valcarce y L. L. Scharf. "Multiantenna spectrum sensing: detection of spatial correlation among time-series with unknown spectra". *IEEE Int. Conf. on Acoust., Speech, and Signal Processing (ICASSP 2010)*, Dallas, EE.UU., Marzo 2010.

- Un análisis detallado de los GLRT derivados en el Capítulo 4 para series temporales bivariadas con PSD arbitrarias muestra que estos pueden reescribirse como funciones del espectro de coherencia (CS). Motivados por este hecho, en el Capítulo 5 se propone una definición del espectro de coherencia generalizado (GCS) basada en la matriz que contiene todos los CS por parejas. Además, se estudian sus propiedades más importantes, así como algunos estimadores del GCS. Por otro lado, esta medida puede relacionarse de manera directa con la información mutua de múltiples señales. Adicionalmente, es interesante comprobar que varios de los GLRT propuestos en capítulos anteriores pueden reescribirse como funciones del GCS. Por último, basándonos en dichas interpretaciones se proponen tests para el problema de detección multicanal usando los estimadores del GCS que presentan mejor comportamiento.

  Los resultados presentados en este capítulo han dado lugar a las siguientes publicaciones:

  - V. Pichler, M. Homolák, W. Skierucha, M. Pichlerová, D. Ramírez, J. Gregor y P. Jaloviar. "Variability of moisture in coarse woody debris from several ecologically important tree species of the temperate zone of Europe". *Ecohydrology (Aceptado)*, 2011.
  - D. Ramírez, J. Vía y I. Santamaría. "Multiple-channel signal detection using the generalized coherence spectrum". *1st IAPR Workshop on Cognitive Information Processing (CIP 2008)*, Santorini, Grecia, Junio 2008.
  - D. Ramírez, J. Vía y I. Santamaría. "A generalization of the magnitude squared coherence spectrum for more than two signals: definition, properties and estimation". *IEEE Int. Conf. on Acoust., Speech, and Signal Processing (ICASSP 2008)*, Las Vegas, EE.UU., Abril 2008.

- Por último, el Capítulo 6 presenta las principales conclusiones derivadas de la Tesis, así como posibles líneas futuras de investigación en este campo.

Capítulo 2

Detección en Series Temporales Multivariadas

Esta Tesis estudia el problema de detección en series temporales multivariadas o multicanal. Éste es un problema que aparece en multitud de aplicaciones, que van desde la bioingeniería hasta los sistemas radar con múltiples antenas, o los sistemas de comunicaciones inalámbricas más modernos basados en el paradigma conocido como radio cognitiva.

El uso de múltiples sensores en el proceso de detección añade un grado adicional de libertad (y de complejidad) al problema, la estructura espacial. La estructura espacial de las señales recibidas en los distintos sensores permite desarrollar detectores que, prácticamente, no necesitan información *a priori*, es decir, no necesitan conocer las formas de onda transmitidas, las densidades espectrales de potencia del ruido, etc. Esto es muy importante porque los detectores basados en la estructura espacial serán más robustos que los detectores que usen información *a priori*, ya que ésta puede no conocerse completamente, o conocerse con errores (de estimación, de transmisión, ...). Todos los tests desarrollados en la Tesis son, por tanto, tests sobre la estructura de la covarianza de la serie temporal multivariada. Además, para no restringir el uso de los detectores, tanto la estructura espacial como la temporal son desconocidas, obteniendo tests de hipótesis con parámetros desconocidos bajo ambas hipótesis.

Este capítulo está organizado de la siguiente manera: en la Sección 2.1 se plantea el problema de detección en series temporales multivariadas, motivando su importancia y analizando trabajos previos. A continuación, los diferentes modelos de señal empleados durante la Tesis se presentan en la Sección 2.2. Una breve introducción al problema de detección o test de hipótesis se realiza en la Sección 2.3, haciendo especial énfasis en las técnicas de detección con parámetros desconocidos. La Sección 2.4 presenta varias aplicaciones de la detección multicanal en diferentes ámbitos, presentando con mayor detalle el paradigma de comunicaciones conocido como radio cognitiva; y, por último, las principales conclusiones se presentan en la Sección 2.5.

2.1. Introducción

El problema de detección de series temporales multivariadas se puede formular de la siguiente manera: dadas M observaciones de longitud N de una serie temporal multivariada adquiridas por un conjunto de L sensores (ver Figura 2.1), $\{\mathbf{x}[n],\ n = 0, 1, \ldots, N-1\}$, el problema consiste en determinar qué modelo ha generado dichas observaciones. A lo largo

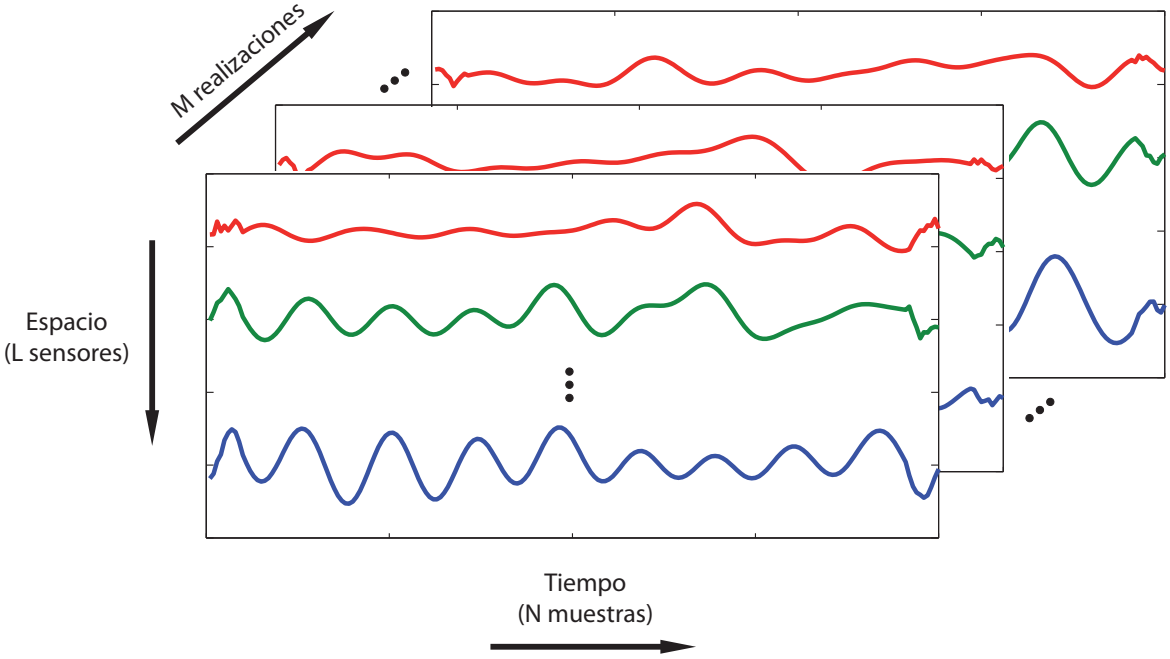

Figura 2.1: Las observaciones están compuestas por ventanas de longitud N de cada una de las L series temporales univariadas.

de la Tesis únicamente se considera el caso binario (dos hipótesis). Esto puede parecer excesivamente restrictivo, pero veremos que los tests de hipótesis binarios tienen aplicación en multitud de ámbitos.

El problema de la detección multicanal puede plantearse de diferentes maneras en función de la información disponible. Si bajo ambas hipótesis las formas de onda son conocidas, esta información se usará para decidir qué hipótesis ha generado las observaciones. Este es el típico problema que surge en comunicaciones digitales [Proakis, 1988], donde el detector óptimo está basado en correlar la señal recibida con cada una de las posibles formas de onda (filtro adaptado) y ver cuál de estas correlaciones proporciona mayor valor. En otras ocasiones, las señales son desconocidas, siendo necesario recurrir a otras propiedades de las señales para discernir entre las hipótesis.

Típicamente, la energía juega un papel importante en problemas de detección. Por ejemplo, esto sucede en el modelo de señal en ruido [Kay, 1998], donde en una hipótesis únicamente hay ruido y en la otra, las observaciones están compuestas por una señal (desconocida) contaminada por ruido aditivo. Para este modelo, como veremos, suponiendo señales y ruidos Gaussianos, el detector óptimo consiste en comparar la energía de las observaciones con un umbral [Urkowitz, 1967, Kay, 1998].

Otra característica útil es la diferente estructura temporal de las observaciones bajo cada hipótesis. Por ejemplo, podemos considerar el problema de determinar si las observaciones están o no correladas temporalmente, es decir, si su densidad espectral de potencia es plana. Si la señal es cicloestacionaria [Gardner et al., 2006], se puede usar esta característica para desarrollar detectores. Esta es una idea muy importante en el campo de las comunicaciones, ya que las señales suelen ser cicloestacionarias [Gardner et al., 2006], mientras que el ruido puede considerarse estacionario. Diferentes tests que explotan la cicloestacionariedad

en una serie temporal univariada pueden encontrarse en [Dandawaté y Giannakis, 1994]. En [Lundén et al., 2009], los citados detectores se extienden considerando múltiples sensores y, además, se proponen detectores energéticamente eficientes.

La última característica, en la que se centra esta Tesis, consiste en aprovechar las distintas estructuras espaciales de las observaciones bajo ambas hipótesis. Obviamente, la estructura espacial surge de manera natural cuando se emplean receptores multiantena o múltiples sensores distribuidos. Diferentes tests han sido propuestos en la literatura estadística para este tipo de problemas desde los años 30. En [Wilks, 1935], se deriva el test que permite determinar si un conjunto de L variables aleatorias Gaussianas son independientes o no. Por otro lado, en [Mauchly, 1940] se resuelve un problema de detección donde la matriz de covarianza bajo una de las hipótesis no tiene estructura espacial y bajo la hipótesis alternativa es conocida hasta un escalado. Asimismo, en [Anderson, 1958] se considera el test donde una de las matrices de covarianza es perfectamente conocida. Todos los detectores mencionados previamente suponen alguna de las matrices de covarianza teóricas desconocida, por lo tanto, las hipótesis son compuestas.

Además de ser ampliamente utilizada en estadística, la estructura espacial también se ha usado en diferentes problemas de conformación de haz o procesado en array [Bose y Steinhardt, 1995, Bose y Steinhardt, 1996, Stoica y Cedervall, 1997, Besson et al., 2006]. Por otro lado, los problemas de detección donde se comprueba la diferente estructura espacial de cada una de las hipótesis son una herramienta potente en sensado espectral para radio cognitiva [Mitola y Maguire Jr., 1999]. En la Sección 2.4 se presenta una revisión más detallada de posibles aplicaciones de la detección multicanal, con especial énfasis en radio cognitiva.

2.2. Modelo de Señal

Esta sección presenta el modelo de señal considerado durante la Tesis. Recordemos la Figura 2.1, donde se ilustran los principales parámetros de la señal adquirida. Concretamente, se dispone para la detección de M realizaciones de una ventana espacio-temporal, compuesta por N muestras de la serie temporal multivariada. Dichas señales, adquiridas por L sensores, se denotan por

$$\{\mathbf{x}[n],\ n = 0, \ldots, N-1\}$$

donde $\mathbf{x}[n] \in \mathbb{C}^L$ es la señal adquirida por los L sensores en el instante n-ésimo y $\{x_i[n],\ n = 0, \ldots, N-1\}$ es la serie temporal del i-ésimo sensor en los N instantes temporales. Además, consideramos señales con media nula, por lo tanto, los estadísticos de segundo orden de $\mathbf{x}[n]$ se resumen en la función de covarianza matricial, definida como

$$\mathbf{R}[n, n-m] = E\left[\mathbf{x}[n]\mathbf{x}^H[n-m]\right] \in \mathbb{C}^{L\times L},$$

donde, por el momento, no se ha supuesto procesos estacionarios en sentido amplio (WSS). Como se ha mencionado previamente, la mayor parte de la Tesis se centra en tests sobre la estructura de covarianza de la serie temporal $\mathbf{x}[n]$. En particular, los tests consisten en decidir si las componentes de $\mathbf{x}[n]$, $x_i[n]$, están correladas espacialmente o no; es decir, si la matriz $\mathbf{R}[n, n-m]$ es diagonal o no. Así, los tests considerados en la Tesis son de la forma

$$\begin{aligned}\mathcal{H}_1 &: E\left[\mathbf{x}[n]\mathbf{x}^H[n-m]\right] = \mathbf{R}_1[n, n-m],\\ \mathcal{H}_0 &: E\left[\mathbf{x}[n]\mathbf{x}^H[n-m]\right] = \mathbf{R}_0[n, n-m],\end{aligned}$$

donde $\mathbf{R}_1[n, n-m] \in \mathbb{S}_+$ es la función de covarianza bajo $\mathcal{H}_1$, $\mathbf{R}_0[n, n-m] \in \mathbb{D}_+$ es la función de covarianza bajo $\mathcal{H}_0$, $\mathbb{S}_+$ denota el conjunto de matrices Hermíticas definidas positivas (sin ninguna estructura adicional) y $\mathbb{D}_+$ es el conjunto de matrices diagonales definidas positivas. Éste es un modelo muy general que será necesario particularizar para derivar los diferentes tests. Así, nos centraremos en los modelos de mayor aplicación práctica, y que se presentan a continuación.

2.2.1. Ruidos con Distribuciones Idénticas

La primera característica empleada para clasificar los diferentes modelos es la distribución del ruido recibido por los diferentes sensores o antenas. En esta subsección se considera el caso donde los ruidos en cada sensor tienen la misma función de covarianza, es decir, son independientes e idénticamente distribuidos (IID). Con lo cual, la covarianza del ruido es

$$\mathbf{R}_0[n, n-m] = \begin{bmatrix} \sigma^2[n, n-m] & 0 & \cdots & 0 \\ 0 & \sigma^2[n, n-m] & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma^2[n, n-m] \end{bmatrix} = \sigma^2[n, n-m]\mathbf{I}.$$

Teniendo en cuenta las restricciones sobre la función de covarianza bajo $\mathcal{H}_0$, una serie de modelos con diferentes suposiciones sobre la estructura temporal y espacial bajo $\mathcal{H}_1$ se presentan a continuación.

Definición 2.1 (Modelo INNSNS: ruidos idénticos y señales no estacionarias sin estructura espacial)**.** *El test de hipótesis para este modelo viene dado por*

$$\begin{aligned} \mathcal{H}_1 &: E\left[\mathbf{x}[n]\mathbf{x}^H[n-m]\right] = \mathbf{R}_1[n, n-m], \\ \mathcal{H}_0 &: E\left[\mathbf{x}[n]\mathbf{x}^H[n-m]\right] = \sigma^2[n, n-m]\mathbf{I}, \end{aligned}$$

donde $\sigma^2[n, n-m] > 0$ y $\mathbf{R}_1[n, n-m] \in \mathbb{S}_+$, es una matriz de covarianza sin estructura espacial adicional.

Es decir, el test para el modelo INNSNS consiste en determinar si la serie temporal multivariada, $\mathbf{x}[n]$, tiene componentes espacialmente incorreladas o no, teniendo la misma función de covarianza en el caso de estar incorreladas.

En el modelo anterior se han supuesto señales no estacionarias. Sin embargo, en la mayor parte de los problemas considerados en la Tesis, se incorpora estructura temporal adicional suponiendo procesos estocásticos estacionarios en sentido amplio (WSS). Además, las señales WSS surgen en muchos problemas prácticos y, como veremos, permiten incorporar de manera sencilla estructura espacial. Para este tipo de series temporales, es bien conocido que la covarianza cumple la siguiente condición

$$\mathbf{R}_i[n, n-m] = \mathbf{R}_i[m], \quad i = 0, 1.$$

Definición 2.2 (Modelo INSSNS: ruidos idénticos y señales estacionarias sin estructura espacial)**.** *El test de hipótesis para este modelo es*

$$\begin{aligned} \mathcal{H}_1 &: E\left[\mathbf{x}[n]\mathbf{x}^H[n-m]\right] = \mathbf{R}_1[m], \\ \mathcal{H}_0 &: E\left[\mathbf{x}[n]\mathbf{x}^H[n-m]\right] = \sigma^2[m]\mathbf{I}, \end{aligned}$$

donde $\sigma^2[m] > 0$ y $\mathbf{R}_1[m] \in \mathbb{S}_+$. La única diferencia con el modelo INNSNS es la suposición de señales estacionarias.

Por otro lado, en ciertas aplicaciones, las señales pueden considerarse frecuencialmente planas, es decir, no presentan estructura temporal. Por ejemplo, algunos estándares de comunicación como el Bluetooth o el ZigBee transmiten con tasas binarias reducidas en bandas frecuenciales donde el ancho de banda de coherencia del canal es elevado, por lo tanto las señales no se distorsionan, obteniendo señales temporalmente incorreladas. Bajo esta suposición, el modelo INSSNS puede simplificarse de la siguiente manera.

Definición 2.3 (Modelo INWSNS: ruidos idénticos y señales estacionarias blancas sin estructura espacial)**.** *El test de hipótesis para este modelo viene dado por*

$$\begin{aligned}\mathcal{H}_1 &: E\left[\mathbf{x}[n]\mathbf{x}^H[n-m]\right] = \mathbf{R}_1\delta[m],\\ \mathcal{H}_0 &: E\left[\mathbf{x}[n]\mathbf{x}^H[n-m]\right] = \sigma^2\delta[m]\mathbf{I},\end{aligned}$$

donde $\delta[m]$ *es la función impulso discreto,* $\sigma^2 > 0$ *y* $\mathbf{R}_1 \in \mathbb{S}_+$.

De momento, la matriz $\mathbf{R}_1[m]$ no presenta ninguna estructura espacial adicional. Sin embargo, en muchos problemas de comunicaciones el modelo típico es el que se conoce como señal en ruido, o modelo de señal latente en otros campos, como la estadística o la econometría. Para este modelo, las observaciones bajo cada hipótesis se generan como

$$\begin{aligned}\mathcal{H}_1 &: \mathbf{x}[n] = (\mathbf{H} * \mathbf{s})[n] + \mathbf{v}[n],\\ \mathcal{H}_0 &: \mathbf{x}[n] = \mathbf{v}[n],\end{aligned}$$

donde $\mathbf{H}[n] \in \mathbb{C}^{L\times P}$ es el canal con múltiples entradas y salidas (MIMO), $\mathbf{s}[n] \in \mathbb{C}^P$ es la señal transmitida por P fuentes o una única fuente con P antenas, $*$ denota el operador convolución y $\mathbf{v}[n] \in \mathbb{C}^L$ es ruido aditivo. Para este modelo, se consideran nuevamente diferentes estructuras temporales para la señal y el ruido.

Definición 2.4 (Modelo INSSPS: ruidos idénticos y señales estacionarias con estructura espacial de rango P)**.** *El test de hipótesis para este modelo se puede formular como*

$$\begin{aligned}\mathcal{H}_1 &: E\left[\mathbf{x}[n]\mathbf{x}^H[n-m]\right] = \mathbf{H}[m] * \mathbf{R}_s[m] * \mathbf{H}^H[-m] + \sigma^2[m]\mathbf{I},\\ \mathcal{H}_0 &: E\left[\mathbf{x}[n]\mathbf{x}^H[n-m]\right] = \sigma^2[m]\mathbf{I},\end{aligned}$$

donde $\mathbf{R}_s[m] = E\left[\mathbf{s}[n]\mathbf{s}^H[n-m]\right] \in \mathbb{C}^{P\times P}$*, es la covarianza de la señal transmitida.*

También es posible considerar señales blancas y canales no selectivos en frecuencia, es decir, $\mathbf{R}_s[m] = \mathbf{R}_s\delta[m]$, con $\mathbf{R}_s = E\left[\mathbf{s}[n]\mathbf{s}^H[n]\right]$ y $\mathbf{H}[m] = \mathbf{H}\delta[m]$. Bajo estas suposiciones resulta el siguiente modelo.

Definición 2.5 (Modelo INWSPS: ruidos idénticos y señales estacionarias blancas con estructura espacial de rango P)**.** *El test de hipótesis para este modelo viene dado por*

$$\begin{aligned}\mathcal{H}_1 &: E\left[\mathbf{x}[n]\mathbf{x}^H[n-m]\right] = \mathbf{H}\mathbf{R}_s\mathbf{H}^H\delta[m] + \sigma^2\mathbf{I}\delta[m],\\ \mathcal{H}_0 &: E\left[\mathbf{x}[n]\mathbf{x}^H[n-m]\right] = \sigma^2\delta[m]\mathbf{I}.\end{aligned}$$

El último modelo considerado para ruidos IID es una combinación de los modelos INSSPS e INWSPS. Concretamente, los canales son planos en frecuencia, pero las señales presentan correlación temporal.

Definición 2.6 (Modelo INMSPS: ruidos idénticos, señales estacionarias con estructura espacial de rango P y canales planos)**.** *El test de hipótesis para este modelo es*

$$\begin{aligned}\mathcal{H}_1 &: E\left[\mathbf{x}[n]\mathbf{x}^H[n-m]\right] = \mathbf{H}\mathbf{R}_s[m]\mathbf{H}^H + \sigma^2[m]\mathbf{I},\\ \mathcal{H}_0 &: E\left[\mathbf{x}[n]\mathbf{x}^H[n-m]\right] = \sigma^2[m]\mathbf{I}.\end{aligned}$$

Finalmente, a modo de resumen, la Tabla 2.1 muestra todos los modelos considerados en la Tesis para ruidos IID.

	No estacionario	Estacionario		
		Blancas	Coloreadas	
Sin estructura espacial	INNSNS	INWSNS	INSSNS	
Estructura espacial de rango P	—	INWSPS	INMSPS	INSSPS

Tabla 2.1: Modelos considerados en la Tesis para ruidos IID.

2.2.2. Ruidos con Distribuciones Diferentes

Esta subsección presenta un modelo más general que el anterior considerando ruidos con distribuciones diferentes. Bajo esta suposición, la única restricción sobre la función de covarianza es

$$\mathbf{R}_0[n, n-m] = \begin{bmatrix} \sigma_1^2[n, n-m] & 0 & \cdots & 0 \\ 0 & \sigma_2^2[n, n-m] & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_L^2[n, n-m] \end{bmatrix} = \mathbf{\Sigma}^2[n, n-m],$$

con $\sigma_i^2[n, n-m] > 0$, es decir, $\mathbf{\Sigma}^2[n, n-m] \in \mathbb{D}_+$. A continuación, dada la similitud con el escenario de ruidos IID, únicamente se presentan los diferentes modelos, sin discutir cada modelo particular.

Definición 2.7 (Modelo NNNSNS: ruidos con distribuciones diferentes y señales no estacionarias sin estructura espacial)**.** *El test de hipótesis para este modelo viene dado por*

$$\begin{aligned} \mathcal{H}_1 &: E\left[\mathbf{x}[n]\mathbf{x}^H[n-m]\right] = \mathbf{R}_1[n, m], \\ \mathcal{H}_0 &: E\left[\mathbf{x}[n]\mathbf{x}^H[n-m]\right] = \mathbf{\Sigma}^2[n, m]. \end{aligned}$$

Definición 2.8 (Modelo NNSSNS: ruidos con distribuciones diferentes y señales estacionarias sin estructura espacial)**.** *El test de hipótesis para este modelo es*

$$\begin{aligned} \mathcal{H}_1 &: E\left[\mathbf{x}[n]\mathbf{x}^H[n-m]\right] = \mathbf{R}_1[m], \\ \mathcal{H}_0 &: E\left[\mathbf{x}[n]\mathbf{x}^H[n-m]\right] = \mathbf{\Sigma}^2[m]. \end{aligned}$$

Definición 2.9 (Modelo NNWSNS: ruidos con distribuciones diferentes y señales estacionarias blancas sin estructura espacial)**.** *El test de hipótesis para este modelo puede expresarse como*

$$\begin{aligned} \mathcal{H}_1 &: E\left[\mathbf{x}[n]\mathbf{x}^H[n-m]\right] = \mathbf{R}_1\delta[m], \\ \mathcal{H}_0 &: E\left[\mathbf{x}[n]\mathbf{x}^H[n-m]\right] = \mathbf{\Sigma}^2\delta[m]. \end{aligned}$$

Definición 2.10 (Modelo NNSSPS: ruidos con distribuciones diferentes y señales estacionarias con estructura espacial de rango P)**.** *El test de hipótesis para este modelo es*

$$\begin{aligned} \mathcal{H}_1 &: E\left[\mathbf{x}[n]\mathbf{x}^H[n-m]\right] = \mathbf{H}[m] * \mathbf{R}_s[m] * \mathbf{H}^H[-m] + \mathbf{\Sigma}^2[m], \\ \mathcal{H}_0 &: E\left[\mathbf{x}[n]\mathbf{x}^H[n-m]\right] = \mathbf{\Sigma}^2[m]. \end{aligned}$$

Definición 2.11 (Modelo NNWSPS: ruidos con distribuciones diferentes y señales estacionarias blancas con estructura espacial de rango P)**.** *El test de hipótesis para este modelo viene dado por*

$$\begin{aligned} \mathcal{H}_1 &: E\left[\mathbf{x}[n]\mathbf{x}^H[n-m]\right] = \mathbf{H}\mathbf{R}_s\mathbf{H}^H\delta[m] + \mathbf{\Sigma}^2\delta[m], \\ \mathcal{H}_0 &: E\left[\mathbf{x}[n]\mathbf{x}^H[n-m]\right] = \mathbf{\Sigma}^2\delta[m]. \end{aligned}$$

Finalmente, a modo de resumen, la Tabla 2.2 muestra todos los modelos para ruidos no-IID considerados en la Tesis.

	No estacionario	Estacionario		
		Blancas	Coloreadas	
Sin estructura espacial	NNNSNS	NNWSNS	NNSSNS	
Estructura espacial de rango P	—	NNWSPS	—	NNSSPS

Tabla 2.2: Modelos considerados en la Tesis para ruidos no-IID.

2.3. Técnicas de Detección

El problema de detección o test de hipótesis es un área muy amplia dentro de la estadística o del procesado (estadístico) de señales. En esta sección se presenta un resumen de las principales técnicas, haciendo mayor énfasis en las que encuentran aplicación en la Tesis. Para una revisión más detallada existen excelente libros, como por ejemplo [Ferguson, 1967, Lehmann, 1986, Scharf, 1991, Kay, 1998].

En primer lugar, existen dos grandes filosofías a la hora de desarrollar detectores: la filosofía clásica o frecuentista y la filosofía Bayesiana. La diferencia entre ambas consiste en la incorporación, o no, de información *a priori* sobre las hipótesis. Los problemas planteados en la Tesis no consideran información *a priori* y, por lo tanto, las técnicas clásicas son más apropiadas. Los problemas de detección también se pueden clasificar en función del número de hipótesis consideradas. De acuerdo a este criterio, los tests de hipótesis pueden ser binarios o múltiples, es decir, el test está compuesto por dos hipótesis o más. Como hemos visto en la sección anterior, todos los modelos considerados en la Tesis son binarios, y por ello, únicamente se revisan las técnicas de detección clásicas para tests de hipótesis binarios.

A continuación se define qué es una hipótesis simple/compuesta y un test óptimo, para presentar el detector más conocido, el detector de Neyman-Pearson [Neyman y Pearson, 1933, Scharf, 1991, Kay, 1998].

Definición 2.12 (Hipótesis simples y compuestas)**.** *Una hipótesis es simple cuando la verosimilitud de las medidas es perfectamente conocida. Por el contrario, una hipótesis es compuesta si la verosimilitud depende de un conjunto de parámetros desconocidos.*

Definición 2.13 (Test óptimo o más potente)**.** *Un test óptimo para un problema de detección con hipótesis simples, bajo la filosofía clásica, es aquel que maximiza la probabilidad de detección (o potencia en la literatura estadística) para una probabilidad de falsa alarma (o tamaño) fija, donde la probabilidad de detección se define como* $p_{\mathrm{D}} = p(\hat{\mathcal{H}}_1|\mathcal{H}_1)$ *y la probabilidad de falsa alarma es* $p_{\mathrm{FA}} = p(\hat{\mathcal{H}}_1|\mathcal{H}_0)$*, siendo* $p(\hat{\mathcal{H}}_i|\mathcal{H}_l)$ *la probabilidad de decidir la hipótesis* i*-ésima cuando la hipótesis activa era la* l*-ésima.*

Es fácil demostrar que el test óptimo para un problema con hipótesis simples es el detector de Neyman-Pearson, basado el cociente de verosimilitudes. Por lo tanto, el estadístico de este test viene dado por

$$T_{\mathrm{NP}}(\mathbf{x}) = \frac{p(\mathbf{x};\boldsymbol{\theta}_1)}{p(\mathbf{x};\boldsymbol{\theta}_0)},$$

donde $\mathbf{x}$ es el conjunto de observaciones disponibles para realizar el test, $p(\mathbf{x};\boldsymbol{\theta}_1)$ es la verosimilitud bajo $\mathcal{H}_1$, dependiente del conjunto de parámetros conocidos $\boldsymbol{\theta}_1$, y $p(\mathbf{x};\boldsymbol{\theta}_0)$ es la verosimilitud bajo $\mathcal{H}_0$, que depende del conjunto de parámetros conocidos $\boldsymbol{\theta}_0$. Finalmente, el

detector se obtiene comparando el estadístico con un umbral

$$\phi_{\mathrm{NP}}(\mathbf{x}) = \begin{cases} 1, & T_{\mathrm{NP}}(\mathbf{x}) \geq \eta, \\ 0, & T_{\mathrm{NP}}(\mathbf{x}) < \eta, \end{cases}$$

donde η se elige para obtener la probabilidad de falsa alarma deseada.

En la mayoría de problemas prácticos no es posible conocer completamente la verosimilitud de las medidas, y éstas suelen depender de unos parámetros desconocidos. Entonces, las hipótesis son compuestas y, por lo tanto, no es posible aplicar el detector de Neyman-Pearson. A continuación se presentan los diferentes criterios y técnicas más populares para tests de hipótesis compuestas.

2.3.1. Test Uniformemente más Potente (UMP)

Como se acaba de ver, al diseñar un detector para un test de hipótesis, se busca su optimalidad en el sentido de Neyman-Pearson. Esto es, que maximize la probabilidad de detección para una probabilidad de falsa alarma fija. Sin embargo, para aplicar este criterio de optimalidad a tests compuestos es necesario generalizarlo. Las verosimilitudes dependen ahora de un conjunto de parámetros $\boldsymbol{\theta}$ desconocidos y $\boldsymbol{\Theta}_1$ y $\boldsymbol{\Theta}_0$ denotan el conjunto de valores que pueden tomar los parámetros bajo $\mathcal{H}_1$ y $\mathcal{H}_0$, respectivamente. Teniendo esto en cuenta, la probabilidad de falsa alarma del test $\phi(\mathbf{x})$ se define como

$$\alpha = \sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}_0} E_{\boldsymbol{\theta}} \left[\phi \left(\mathbf{x} \right) \right],$$

es decir, la probabilidad de falsa alarma para un problema con hipótesis compuestas se define como el supremo sobre el conjunto $\boldsymbol{\Theta}_0$ de las probabilidades de falsa alarma cuando el test es simple. Con esta definición de la probabilidad de falsa alarma, diremos que un test para hipótesis compuestas, $\phi\left(\mathbf{x}\right)$, es un test uniformemente más potente de tamaño α, si para cualquier otro test $\phi'\left(\mathbf{x}\right)$ con $\sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}_0} E_{\boldsymbol{\theta}} \left[\phi' \left(\mathbf{x} \right) \right] \leq \alpha$, se cumple

$$E_{\boldsymbol{\theta}} \left[\phi \left(\mathbf{x} \right) \right] \geq E_{\boldsymbol{\theta}} \left[\phi' \left(\mathbf{x} \right) \right], \quad \forall \boldsymbol{\theta} \in \boldsymbol{\Theta}_1,$$

donde $p_{\mathrm{D}}(\boldsymbol{\theta}) = E_{\boldsymbol{\theta}} \left[\phi \left(\mathbf{x} \right) \right]$ es la probabilidad de detección del test para el conjunto de parámetros $\boldsymbol{\theta} \in \boldsymbol{\Theta}_1$.

A pesar de la optimalidad de los tests UMP, en un gran número de problemas no existen. Además, hay muy pocos resultados relacionados con la existencia y obtención del test UMP, y los que hay son muy concretos: parámetros escalares y un tipo de test especial. A continuación se presentan dos resultados para obtener tests UMP [Scharf, 1991].

Teorema 2.1 (Karlin-Rubin)**.** *Sea x una variable aleatoria escalar cuya verosimilitud está parametrizada por θ, un parámetro escalar. Supongamos que el cociente de verosimilitudes dado por*

$$\frac{p(\mathbf{x}; \theta_1)}{p(\mathbf{x}; \theta_0)},$$

es no decreciente para $\theta_1 > \theta_0$. Entonces, para el problema de detección[1]

$$\begin{aligned} \mathcal{H}_1 &: \theta > \theta_0, \\ \mathcal{H}_0 &: \theta \leq \theta_0, \end{aligned}$$

[1]Este tipo de tests se conoce como tests unilaterales (*one-sided tests*).

existe un test UMP de tamaño α, dado por

$$x \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta.$$

Teorema 2.2. *Sea* $\mathbf{x}$ *una variable aleatoria vectorial cuya verosimilitud está parametrizada por* θ*, un parámetro escalar, y además pertenece a la familia exponencial, es decir*

$$p(\mathbf{x};\theta) = c(\theta)h(\mathbf{x})\exp(\pi(\theta)t(\mathbf{x})).$$

Si el cociente de verosimilitudes es monótono en el estadístico suficiente, $t(\mathbf{x})$*, lo cual se cumple si* $\pi(\theta)$ *es no decreciente en* θ*, se puede demostrar que para el test de hipótesis unilateral*

$$\begin{aligned}\mathcal{H}_1 &: \theta > \theta_0,\\ \mathcal{H}_0 &: \theta \leq \theta_0,\end{aligned}$$

existe un test UMP.

2.3.2. Test Localmente más Potente (LMP)

Para una amplia mayoría de tests de hipótesis compuestas no es posible derivar el test UMP. En algunos casos, considerando hipótesis próximas, se define el test LMP como aquel que maximiza $p_{\mathrm{D}}(\boldsymbol{\theta})$ para una probabilidad de falsa alarma fija, donde los parámetros $\boldsymbol{\theta} \in \tilde{\boldsymbol{\Theta}}_1$, siendo $\tilde{\boldsymbol{\Theta}}_1 \in \boldsymbol{\Theta}_1$ una vecindad abierta de $\boldsymbol{\Theta}_0$. No obstante, los tests LMP no serán óptimos para hipótesis muy distanciadas. La obtención de tests LMP se basa en desarrollos en serie de Taylor y, al igual que los tests UMP, en general su existencia no está garantizada.

2.3.3. Test Invariantes

La principal idea que hay tras la invarianza es la siguiente: imaginemos, por ejemplo, que de todos los parámetros $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_p)$, algunos de ellos, $(\theta_r, \ldots, \theta_p)$, no dependen de la hipótesis. Dichos parámetros (*nuisance parameters*) no ayudan a distinguir entre ambas hipótesis y pueden impedir la obtención del test UMP. Entonces, parece razonable diseñar tests que eliminen los parámetros $(\theta_r, \ldots, \theta_p)$, siendo una posible manera considerar tests invariantes. Es decir, en problemas donde no existe un test UMP, imponiendo ciertas restricciones a la clase de detectores considerados, se puede encontrar, dentro de dicha clase restringida de detectores, un test UMP, es decir, se deriva un test UMP invariante (UMPI). Este concepto también puede aplicarse a tests LMP, obteniendo tests LMPI.

A continuación, se presenta brevemente la teoría de la invarianza. Un análisis exhaustivo de la invarianza en problemas de detección se puede encontrar en [Ferguson, 1967, Capítulo 5] y su aplicación a problemas de procesado de señal en [Scharf, 1991]. Comencemos por definir un problema de detección invariante. Sea $\mathbf{y} = g(\mathbf{x})$ una transformación de las observaciones obtenida del grupo de transformaciones G, distribuidas según $q(\mathbf{y};\boldsymbol{\theta})$. Si la transformación deja invariante la distribución de las observaciones y sólo transforma los parámetros mediante la función $\overline{g}(\cdot)$, cumpliéndose

$$q(\mathbf{y};\boldsymbol{\theta}) = p(\mathbf{y};\overline{g}(\boldsymbol{\theta})),$$

la familia de distribuciones $p(\mathbf{x};\boldsymbol{\theta})$ será invariante al grupo de transformaciones G. Asimismo, si se cumple $\overline{g}(\boldsymbol{\Theta}_i) = \boldsymbol{\Theta}_i$, el test de hipótesis

$$\begin{aligned}\mathcal{H}_1 &: \boldsymbol{\theta} \in \boldsymbol{\Theta}_1,\\ \mathcal{H}_0 &: \boldsymbol{\theta} \in \boldsymbol{\Theta}_0,\end{aligned}$$

es invariante a G. Por último, un test es invariante a G si

$$T(g(\mathbf{x})) = T(\mathbf{x}), \quad \forall g(\cdot) \in G.$$

Un concepto importante de los tests invariantes es el estadístico máximamente invariante (*maximal invariant statistic*), $M(\mathbf{x})$, definido a continuación.

Definición 2.14. *Un estadístico, $M(\mathbf{x})$, es máximamente invariante si cumple:*

- $M(g(\mathbf{x})) = M(\mathbf{x}), \quad \forall g(\cdot) \in G.$
- $M(\mathbf{x}_1) = M(\mathbf{x}_2)$, *implica* $\mathbf{x}_2 = g(\mathbf{x}_1)$ *para alguna transformación* $g(\cdot) \in G$.

La importancia del estadístico máximamente invariante radica en que cualquier test invariante se puede reescribir como una función de él, es decir,

$$T(\mathbf{x}) = T(M(\mathbf{x})).$$

Por lo tanto, es posible restringirse a funciones del estadístico máximamente invariante al derivar tests invariantes. Finalmente, en [Scharf, 1991], se propone obtener el test UMPI basándose en el cociente de verosimilitudes del estadístico máximamente invariante.

2.3.4. Test basado en el Cociente de Verosimilitudes Generalizado (GLRT)

Sin duda, uno de los detectores más conocidos y ampliamente utilizados para test binarios de hipótesis compuestas es el GLRT [Neyman y Pearson, 1928, Kay, 1998]. El GLRT para $\mathcal{H}_0 : \boldsymbol{\theta} \in \boldsymbol{\Theta}_0$ frente a $\mathcal{H}_1 : \boldsymbol{\theta} \in \boldsymbol{\Theta}_1$ se basa en el cociente de verosimilitudes generalizado (GLR), dado por [Mardia et al., 1979]

$$\mathscr{L} = \frac{\operatorname*{máx}_{\boldsymbol{\theta}\in\boldsymbol{\Theta}_0} p\left(\mathbf{x};\boldsymbol{\theta}\right)}{\operatorname*{máx}_{\boldsymbol{\theta}\in\boldsymbol{\Theta}_1} p\left(\mathbf{x};\boldsymbol{\theta}\right)},$$

donde $p\left(\mathbf{x};\boldsymbol{\theta}\right)$ es la verosimilitud (que depende de $\boldsymbol{\theta}$) de las medidas, y el test se obtiene comparando el GLR con un umbral

$$\mathscr{L} \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\gtrless}} \eta,$$

donde η se elige para obtener una determinada probabilidad de falsa alarma. Con lo cual, la idea del GLRT es muy sencilla y consiste en derivar el detector Neyman-Pearson sustituyendo los parámetros desconocidos por sus estimas de máxima verosimilitud (ML).

El uso de estimadores ML en el GLRT se debe a sus propiedades asintóticas. Para ver dichas propiedades, definamos primero el estimador ML como

$$\hat{\boldsymbol{\theta}}_{\mathrm{ML}} = \arg \operatorname*{máx}_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} p\left(\mathbf{x};\boldsymbol{\theta}\right),$$

donde $\boldsymbol{\Theta}$ es el conjunto de posibles valores de los parámetros. Si la función densidad de probabilidad (PDF) de $\mathbf{x}$ satisface unas condiciones de regularidad, definidas a continuación, y las estimas ML se obtienen igualando a cero la derivada de la verosimilitud, esto es,

$$\left.\frac{\partial \log p(\mathbf{x};\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}_{\mathrm{ML}}} = \mathbf{0},$$

se obtienen estimadores asintóticamente insesgados y eficientes, es decir, alcanzan la cota de Cramer-Rao [Kay, 1993]. Además el estimador es consistente y se distribuye (asintóticamente) de manera normal, esto es

$$\hat{\boldsymbol{\theta}}_{\mathrm{ML}} \overset{a}{\sim} \mathcal{CN}\left(\boldsymbol{\theta}, \mathbf{J}^{-1}\left(\boldsymbol{\theta}\right)\right),$$

donde $\overset{a}{\sim}$ indica distribución asintótica (en el número de observaciones) y $\mathbf{J}\left(\boldsymbol{\theta}\right)$ es la matriz de información de Fisher [Kay, 1993], definida como

$$\mathbf{J}\left(\boldsymbol{\theta}\right) = -E\left[\frac{\partial^2}{\partial \boldsymbol{\theta}^2} \log p\left(\mathbf{x}; \boldsymbol{\theta}\right)\right].$$

La demostración de esta propiedad se puede encontrar en [Kay, 1993]. Por otro lado, referenciamos al lector a [Dudewicz, 1976] para una demostración más rigurosa de la consistencia y a [Rao, 1988] para la demostración de Gaussianidad. Por último, las condiciones de regularidad referenciadas previamente son:

1. la primera y segunda derivadas de la log-verosimilitud están bien definidas,
2. se cumple
$$E\left[\frac{\partial \log p\left(\mathbf{x}; \boldsymbol{\theta}\right)}{\partial \boldsymbol{\theta}}\right] = \mathbf{0}.$$

Por último, se presentan algunas propiedades del GLRT. Comencemos por destacar que el GLRT no es óptimo en el sentido de Neyman-Pearson. Sin embargo, debido a las propiedades asintóticas de los estimadores ML, el GLRT proporciona buenos resultados cuando se dispone de un número suficiente de muestras [Kay, 1998, Mardia et al., 1979]. Además, en una amplia mayoría de problemas, el GLRT resulta sencillo de implementar como se verá en capítulos posteriores. Por último, otra propiedad asintótica del GLRT que justifica su popularidad es la siguiente.

Teorema 2.3 (Wilks)**.** *Si $p(\mathbf{x}; \boldsymbol{\theta})$ cumple las condiciones de regularidad y las estimas ML se obtienen igualando a cero la derivada de la verosimilitud, el log-GLRT bajo $\mathcal{H}_0$ tiene la siguiente distribución asintótica*

$$-2 \log \mathscr{L} \overset{a}{\sim} \chi^2_{p_1 - p_0},$$

donde χ^2_ν denota una variable aleatoria Chi-cuadrado centrada con ν grados de libertad, p_1 es el número de parámetros desconocidos bajo $\mathcal{H}_1$ y p_0 es el número de parámetros desconocidos bajo $\mathcal{H}_0$.

Por lo tanto, aplicando el Teorema de Wilks se pueden obtener los umbrales, para una probabilidad de falsa alarma fija, de manera teórica. Asimismo, en [Kay, 1998] se presenta una generalización del teorema de Wilks, que establece la distribución asintótica del GLRT bajo $\mathcal{H}_1$. No obstante, esta generalización es menos rigurosa que el Teorema de Wilks y no se considerará.

2.3.5. Tests de Rao y de Wald

A pesar de que el GLRT es uno de los tests más ampliamente utilizados en el campo del tratamiento estadístico de señales, como se ha señalado previamente no existen resultados generales sobre su optimalidad. Debido a esto, es interesante buscar tests alternativos que proporcionen mejores resultados o que sean más sencillos de derivar. En esta línea se han

propuesto los detectores de Wald [Wald, 1943] y Rao [Rao, 1948]. Estos detectores son más sencillos de implementar que el GLRT, siendo válidos únicamente para tests binarios en los que sólo una de las hipótesis es compuesta, es decir, $\mathcal{H}_0$ es una hipótesis simple.

El estadístico del test de Wald viene dado por

$$\mathscr{W} = \left(\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_0\right)^T \mathbf{J}\left(\hat{\boldsymbol{\theta}}_1\right)\left(\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_0\right)$$

donde $\hat{\boldsymbol{\theta}}_1$ es la estima ML de los parámetros bajo $\mathcal{H}_1$, $\boldsymbol{\theta}_0$ son los parámetros bajo $\mathcal{H}_0$ y $\mathbf{J}\left(\hat{\boldsymbol{\theta}}_1\right)$ es la matriz de información de Fisher evaluada en $\hat{\boldsymbol{\theta}}_1$. Por otro lado, el estadístico del test de Rao es

$$\mathscr{R} = \left.\frac{\partial \log p\left(\mathbf{x};\boldsymbol{\theta}\right)}{\partial \boldsymbol{\theta}}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}^{T} \mathbf{J}^{-1}\left(\boldsymbol{\theta}_0\right) \left.\frac{\partial \log p\left(\mathbf{x};\boldsymbol{\theta}\right)}{\partial \boldsymbol{\theta}}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}.$$

Como se puede ver, el test de Wald sólo necesita la estima ML bajo $\mathcal{H}_1$ y, por el contrario, el test de Rao no necesita ninguna estima ML. Además, es posible demostrar que los tests de Rao, de Wald y el GLRT son asintóticamente equivalentes [Kay, 1998]. Debido a estas propiedades, su aplicación en el campo del tratamiento estadístico de señales está creciendo en los últimos años [Baggenstoss y Kay, 1992, de Maio, 2004, de Maio, 2007]. Además, en [de Maio et al., 2010] se presenta un análisis de la relación entre los tests de Rao, de Wald y el GLRT para un modelo de señal particular.

2.4. Aplicaciones de la Detección Multicanal

Esta sección presenta un conjunto de aplicaciones de la detección multicanal con gran interés práctico. Un primer ejemplo consiste en sistemas radar equipados con múltiples antenas en recepción, y una o más antenas en transmisión [Fishler et al., 2004a, Li y Stoica, 2008], conocidos como sistemas MIMO-radar. Este tipo de sistemas radar ha ganado interés en los últimos años [Fishler et al., 2004a] debido a que proporcionan diversidad espacial [Fishler et al., 2004b, Fishler et al., 2006], de manera análoga a los sistemas MIMO en comunicaciones inalámbricas. Normalmente, en radar las formas de onda transmitidas son conocidas por los sensores en el momento de realizar el test de hipótesis. Sin embargo, los problemas considerados en la Tesis, no son conocidas. Si éste fuera el caso, todos los detectores desarrollados en esta Tesis podrían aplicarse en MIMO-radar.

La detección multicanal también es un problema importante en procesado de señal para sonar, siendo una posible aplicación la detección de minas subacuáticas u otros objetos [Azimi-Sadjadi et al., 2000]. Éste es un problema complicado debido a la no-repetibilidad, la variabilidad de la firma del blanco u objetivo (*target signature*), el ruido generado por el hombre, la reverberación, la gran variabilidad del canal y la falta de información *a priori*. Por esto, es necesario usar múltiples sensores para obtener más medidas sin aumentar el tiempo de sensado. Sin embargo, este entorno de propagación tan complejo ha obligado a emplear técnicas de detección diferentes a las que se proponen en la Tesis, basadas en el aprendizaje-máquina. Por ejemplo, en [Li et al., 2004] se comparan algunas de dichas técnicas: el vecino más próximo [Denoeux, 1995], redes neuronales probabilísticas [Specht, 1990], máquinas de vectores soporte [Vapnik, 1995, Schölkopf y Smola, 2001], etc. No obstante, cuando los ruidos en los sensores son independientes, los detectores propuestos en la Tesis encuentran aplicación en este problema.

Otro campo importante donde surgen problemas de detección multicanal es la bioingeniería o el análisis de datos médicos. El ejemplo más típico es el análisis de las señales

electroencefalográfica (EEG) y electrocardiográfica (ECG). Dichas señales capturan la actividad eléctrica en el cerebro y en el corazón, respectivamente; y son, por naturaleza, multicanal ya que se usan múltiples sensores (o derivaciones). En otras ocasiones se obtienen procesos multivariados fusionando diferentes tipos de señales: ECG, variabilidad del ritmo cardíaco y la variabilidad de la presión sanguínea [Luesma, 2006]. Por otro lado, un tipo de prueba no invasiva que ha ido aumentando su popularidad es la imagen por resonancia magnética funcional (fMRI) [Friston et al., 1994]. La fMRI mide la respuesta hemodinámica del cerebro, es decir, cambios en el flujo de sangre como consecuencia de la actividad cerebral. La naturaleza multicanal de este tipo de señales radica en la alta resolución espacial de la imágenes adquiridas y su variación temporal. Surge, entonces, una serie temporal multivariada considerando cada posición espacial como una componente de la serie.

En radioastronomía cada vez es mayor la contaminación que sufren las observaciones debido a sistemas de comunicaciones inalámbricas o sistemas de navegación por satélite. Por lo tanto, la detección de dichas señales interferentes permitirá su posterior cancelación. Además, dada la naturaleza de los sistemas de medida, la detección multicanal también es importante en este campo [Leshem et al., 2000, Leshem y Van der Veen, 2001a, Leshem y Van der Veen, 2001b].

Últimamente, en el campo de las comunicaciones ha surgido un nuevo paradigma conocido como radio cognitiva (CR) [Mitola y Maguire Jr., 1999, Hossain y Bhargava, 2007, Chen y Prasad, 2009] y dada la importancia en la Tesis del CR, en la Sección 2.4.1 se realiza una breve introducción al tema.

2.4.1. Radio Cognitiva

Actualmente, hay una aparente escasez de espectro radioeléctrico [FCC, 2002]. Sin embargo, dicha escasez está motivada por la estricta política de asignación frecuencial existente, que no permite el uso de las bandas frecuenciales asignadas, incluso estando vacías. La idea principal de CR se basa en el acceso oportunista, tanto temporal como espacialmente, a dichas bandas cuando no están siendo ocupadas [Akyildiz et al., 2008]. El sensado espectral es, por ello, un ingrediente clave de CR.

El canal inalámbrico, debido a los desvanecimientos de pequeña y gran escala, provoca que la detección de bandas no ocupadas sea una tarea realmente complicada, en especial cuando la relación señal a ruido (SNR) recibida es muy baja. Por este motivo, cualquier característica conocida acerca de la señal transmitida debe ser empleada en la detección. El modelo típico para la detección en CR es el de señal en ruido [Kay, 1998], es decir,

$$\begin{aligned} \mathcal{H}_1 &: x[n] = s[n] + v[n], \\ \mathcal{H}_0 &: x[n] = v[n], \end{aligned}$$

donde $x[n]$ es la señal recibida, $s[n]$ es la señal transmitida y $v[n]$ es el ruido aditivo en el sensor. Para este modelo, si el ruido es Gaussiano y la señal $s[n]$ conocida, el detector óptimo es el filtro adaptado [Kay, 1998], cuyo estadístico es

$$T_{\mathrm{MF}} = \sum_{n=0}^{N-1} s^*[n]x[n],$$

donde N es el número de muestras disponibles para detectar. Como se ha mencionado previamente, es necesario conocer la señal transmitida para implementar el detector, algo no muy realista en ciertos escenarios. Otro detector típico se basa en la cicloestacionariedad de

la señal [Gardner et al., 2006]. Esto es posible dado que las señales de comunicaciones suelen ser cicloestacionarias [Proakis, 1988, Gardner et al., 2006] y el ruido puede considerarse estacionario.

Las dos aproximaciones anteriores necesitan algún tipo de sincronización con la señal transmitida, que puede ser poco fiable en la región de baja SNR [Cabric, 2008]. Debido a esto, es importante desarrollar detectores asíncronos. El detector asíncrono más conocido es el detector de energía [Urkowitz, 1967, Kay, 1998], que además es el detector óptimo para el modelo de señal en ruido presentado anteriormente, si $s[n]$ es una señal Gaussiana desconocida sin correlación temporal. El test de Neyman-Pearson para este modelo es

$$T_{\mathrm{ED}} = \sum_{n=0}^{N-1} |x[n]|^2 .$$

El mayor problema del detector de energía, además de sus pobres prestaciones, consiste en la necesidad de conocer (o estimar) la varianza del ruido para el cálculo del umbral. Sin embargo, cualquier desviación respecto al nivel nominal de este valor conlleva severas degradaciones en las prestaciones del detector [Tandra y Sahai, 2008]. Esto puede llegar a provocar que, incluso teniendo infinitas observaciones, $N \to \infty$, haya un valor de SNR, conocido como muro de SNR (*SNR wall*), por debajo del cual sea imposible detectar [Tandra y Sahai, 2008].

Las limitaciones del detector de energía motivan la búsqueda de detectores asíncronos robustos a incertidumbres en la varianza del ruido. Una manera de solucionar este problema es el uso de detectores con múltiples antenas, lo que permite explotar la incorrelación espacial del ruido en el proceso de detección. El modelo básico es un caso particular del INWSPS o del NNWSPS presentados en esta Tesis

$$\begin{aligned} \mathcal{H}_1 &: \mathbf{x}[n] = \mathbf{h}s[n] + \mathbf{v}[n], \quad n = 0, \ldots, N-1, \\ \mathcal{H}_0 &: \mathbf{x}[n] = \mathbf{v}[n], \qquad\qquad\quad n = 0, \ldots, N-1, \end{aligned}$$

donde $\mathbf{x}[n] \in \mathbb{C}^L$ es la señal adquirida por las L antenas, $\mathbf{h} \in \mathbb{C}^L$ es el canal con una entrada y múltiples salidas (SIMO) entre el usuario primario y el sensor, $s[n]$ es la señal transmitida y $\mathbf{v}[n]$ es el ruido, que suponemos espacialmente incorrelado. Bajo estos supuestos, en [Besson et al., 2006, Taherpour et al., 2010, López-Valcarce et al., 2010, Wang et al., 2010, Zeng et al., 2008, Alamgir et al., 2008, Zeng y Liang, 2009] se proponen detectores para este modelo, basados en los autovalores de la matriz de covarianza muestral o de la matriz de coherencia.

Sin embargo, en escenarios realistas, puede haber múltiples señales transmitidas o puede que el transmisor sea multiantena. Surgen entonces los modelos INWSPS y NNWSPS de la Tesis

$$\begin{aligned} \mathcal{H}_1 &: \mathbf{x}[n] = \mathbf{H}\mathbf{s}[n] + \mathbf{v}[n], \quad n = 0, \ldots, N-1, \\ \mathcal{H}_0 &: \mathbf{x}[n] = \mathbf{v}[n], \qquad\qquad\quad n = 0, \ldots, N-1, \end{aligned}$$

donde $\mathbf{H} \in \mathbb{C}^{L \times P}$ es el canal MIMO y $\mathbf{s}[n] \in \mathbb{C}^P$ es la señal transmitida que, en este caso, es vectorial. Este modelo surge cuando múltiples usuarios comparten recursos (tiempo, espacio y frecuencia). También, este modelo aparece en los nuevos estándares de comunicación que consideran la transmisión de múltiples flujos de datos para obtener ganancia por multiplexado [Vucetic y Yuan, 2003] y/o usan códigos espacio-temporales para obtener ganancia por diversidad [Vucetic y Yuan, 2003]. Ejemplos de estos estándares incluyen los de radiodifusión, como el europeo DVB-T2 [ETSI, 2009], el cual considera la transmisión con el esquema de Alamouti [Alamouti, 1998], es decir,

usa 2 antenas en transmisión. Otros ejemplos son los estándares de transmisión punto a multipunto, como el IEEE 802.11n [IEEE Computer Society, 2009], el IEEE 802.16 [IEEE Computer Society and the IEEE Microwave Theory and Techniques Society, 2009] o el estándar de telefonía móvil de cuarta generación LTE [3GPP, 2009], que soportan hasta cuatro antenas transmisoras, es decir, $P = 4$.

Todos los modelos y, consecuentemente, todos los detectores considerados se basan en la suposición de señales temporalmente incorreladas y canales planos en frecuencia. Sin embargo, la reciente demanda de mayores tasas de transmisión en las comunicaciones inalámbricas actuales, requiere la transmisión de señales de banda ancha. En esta situación, los modelos INSSPS y NNSSPS, que vienen dados por

$$\begin{aligned} \mathcal{H}_1 &: \mathbf{x}[n] = (\mathbf{H} * \mathbf{s})[n] + \mathbf{v}[n], \quad && n = 0, \ldots, N-1, \\ \mathcal{H}_0 &: \mathbf{x}[n] = \mathbf{v}[n], && n = 0, \ldots, N-1, \end{aligned}$$

adquieren mayor importancia en CR. Por ejemplo, esto se ha considerado en los trabajos [Font-Segura et al., 2011b, Font-Segura et al., 2011a]

Finalmente, si se conocieran otras propiedades de la señal, además de su estructura espacial se podrían explotar en el proceso de detección multicanal, siendo un ejemplo de esto el trabajo considerado en [Romero y López-Valcarce, 2011b], donde se suponen señales con envolvente constante. Por otro lado, considerando sensores multiantena distribuidos se obtiene un modelo equivalente al anterior. No obstante, las técnicas de detección empleadas deben ser diferentes ya que no es posible transmitir toda la información sobre las señales al centro de fusión, por ello, en [Romero y López-Valcarce, 2011a] se proponen diferentes detectores combinando la información de los detectores individuales.

2.5. Conclusiones

Este capítulo nos ha permitido introducir el problema considerado en la Tesis: la detección multicanal. Se ha comenzado por motivar el problema, presentando un resumen de la información o propiedades de las señales útiles para resolverlo. Además, se ha introducido brevemente la teoría de la detección clásica, haciendo especial énfasis en los tests binarios de hipótesis compuestas y los detectores resultantes. Entre estos detectores, merece una mención especial el GLRT. La utilización del GLRT en la Tesis se debe a su derivación relativamente sencilla y a su buen comportamiento en todos los problemas considerados. Por otro lado, se han presentado posibles aplicaciones de la detección multicanal. Dichas aplicaciones varían desde la detección en MIMO-radar hasta la bioingeniería, pasando por sonar. Finalmente, se ha considerado con mayor detenimiento el problema de sensado espectral para radio cognitiva, por ser la principal aplicación de las técnicas de detección propuestas en esta Tesis, y dada su gran popularidad en el campo de las comunicaciones inalámbricas actuales.

Capítulo 3

Detección Multicanal de Procesos Contaminados con Ruidos IID

Este capítulo estudia el problema de detección multicanal en series temporales contaminadas con ruidos independientes e idénticamente distribuidos (IID). En primer lugar, se resuelven los diferentes problemas de detección para series temporales estacionarias con densidades espectrales de potencia (PSD) planas, es decir, procesos sin correlación temporal. Como veremos, estos tests de hipótesis se reducen a testear la estructura de diferentes matrices de covarianza. Los problemas de detección tienen hipótesis compuestas, y no es posible derivar el detector de Neyman-Pearson. Por lo tanto, se propone resolver dichos tests de hipótesis aplicando el detector basado en el cociente de verosimilitudes generalizado (GLRT), debido a su simplicidad y buenas prestaciones en la mayoría de problemas prácticos.

En segundo lugar, se considera el caso en el que los procesos estocásticos son estacionarios con PSD arbitrarias y desconocidas. En esta situación, las matrices de covarianza son Toeplitz por bloques y, por lo tanto, la estimación ML de dichas matrices es un problema de optimización no convexo. Por ello, no es posible derivar una fórmula cerrada para el GLRT. No obstante, si se supone un número de observaciones elevado, aplicando la verosimilitud asintótica (en el dominio frecuencial), es posible obtener fórmulas cerradas para el GLRT. Además, trabajando en el dominio frecuencial es sencillo explotar la estructura espacial.

El último escenario considera la detección de series temporales no estacionarias. En este caso, las matrices de covarianza no tienen estructura Toeplitz y, por ello, su estimación no se basa en la verosimilitud asintótica.

Este capítulo está organizado de la siguiente manera: la Sección 3.1 considera el problema de detección con procesos estacionarios blancos sin estructura espacial bajo $\mathcal{H}_1$, y posteriormente, se generaliza considerando modelos de rango reducido. La Sección 3.2 extiende dichos problemas, con y sin estructura espacial, a series temporales estacionarias con espectros arbitrarios. La Sección 3.3 aborda la detección de señales no estacionarias. Las prestaciones de todos los detectores, así como la comparación con otros detectores propuestos en la literatura, se analizan mediante simulaciones de Monte Carlo en la Sección 3.4. Por último, en la Sección 3.5 se resumen las principales conclusiones obtenidas.

3.1. Detección Multicanal de Procesos Estacionarios Blancos

En esta sección se aborda el problema de detección multicanal de procesos temporalmente blancos y contaminados con ruidos IID. Recordemos que para el test de hipótesis, se dispone

de M ventanas de longitud N de las señales adquiridas por L sensores (ver Figura 2.1), denotadas como

$$\{\mathbf{x}[n],\ n = 0, \ldots, N-1\},$$

donde $\mathbf{x}[n] \in \mathbb{C}^L$ son las medidas de todos los sensores en el instante n-ésimo.[1] Para derivar el GLRT es necesario conocer la función densidad de probabilidad (PDF) de $\mathbf{x}[n]$, que modelaremos como una variable aleatoria Gaussiana multidimensional, compleja y circular. No obstante, en el caso de señales no Gaussianas, esta suposición permite hacer desarrollos matemáticos de manera sencilla, obtener detectores útiles y fácilmente interpretables. Asimismo, para baja relación señal a ruido o constelaciones multinivel, la suposición de señales Gaussianas proporciona estimadores asintóticamente óptimos [Villares y Vázquez, 2007], y puede considerarse como la suposición más desfavorable [Stoica y Babu, 2011].

El problema de detección multicanal considerado en esta sección puede formularse como

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

donde $\mathbf{R}_i[m] = E[\mathbf{x}[n]\mathbf{x}^H[n-m]]$ es la función de covarianza bajo la i-ésima hipótesis, que se supone desconocida. Dada la estructura temporal, el problema (3.1) se puede simplificar de la siguiente manera

$$\begin{aligned} \mathcal{H}_1 &: \mathbf{x} \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{R}_1\right), \\ \mathcal{H}_0 &: \mathbf{x} \sim \mathcal{CN}\left(\mathbf{0}, \sigma^2\mathbf{I}\right), \end{aligned} \tag{3.2}$$

donde $\sigma^2 > 0$ y $\mathbf{R}_1 \in \mathbb{S}_+$, es decir, se elimina la dependencia temporal. En los siguientes apartados, se resuelve el test de hipótesis bajo diferentes suposiciones sobre la estructura espacial de $\mathbf{R}_1$.

3.1.1. Señales sin Estructura Espacial

En primer lugar, se considera el problema de detección dado por el modelo INWSNS, es decir, la matriz de covarianza bajo $\mathcal{H}_1$ no tiene estructura adicional. Este test, conocido como test de esfericidad, fue desarrollado en [Mauchly, 1940] para variables aleatorias multidimensionales reales y Gaussianas. No obstante, el problema resuelto en [Mauchly, 1940] es más general, considerando una matriz de covarianza bajo $\mathcal{H}_0$ genérica conocida hasta un factor de escala, siendo en nuestro caso la identidad. Sin embargo, las diferencias entre ambos casos son mínimas y bastaría con preblanquear las medidas por la matriz conocida para que ambos problemas sean equivalentes. Además, cuando la matriz desconocida es la identidad, se está testeando si las curvas de equiprobabilidad son esféricas o no, de ahí que se llame test de esfericidad.

Formulado de manera matemática, el problema de detección resultante es

$$\begin{aligned} \mathcal{H}_1 &: \mathbf{x} \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{R}_1\right), \\ \mathcal{H}_0 &: \mathbf{x} \sim \mathcal{CN}\left(\mathbf{0}, \sigma^2\mathbf{I}\right), \end{aligned} \tag{3.3}$$

donde $\sigma^2 > 0$ es un parámetro desconocido y $\mathbf{R}_1 \in \mathbb{S}_+$ es una matriz de covarianza también desconocida, cuya única restricción es ser definida positiva ($\mathbb{S}_+$ es el conjunto de matrices Hermíticas definidas positivas). Supongamos M ($M \geq L$) observaciones independientes e idénticamente distribuidas (IID) para derivar el test: $\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}$. En [Mauchly, 1940] se

[1] Nótese, además, que para series temporalmente blancas es indiferente tener M ventanas de longitud N, que NM ventanas de longitud 1. Por simplicidad notacional se considerarán M ventanas de longitud $N = 1$.

empleó el GLRT para este problema, basado en el cociente de verosimilitudes generalizado (GLR)

$$\mathscr{L} = \frac{\max\limits_{\sigma^2>0} p\left(\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}; \sigma^2\mathbf{I}\right)}{\max\limits_{\mathbf{R}_1\in\mathbb{S}_+} p\left(\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}; \mathbf{R}_1\right)},$$

donde $p\left(\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}; \mathbf{R}\right)$ es la verosimilitud de las medidas, $\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}$, que depende de la matriz de covarianza $\mathbf{R}$.

Dado que las medidas son IID y teniendo en cuenta su distribución marginal, $\mathbf{x}_m \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{R}\right)$, la verosimilitud viene dada por

$$p\left(\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}; \mathbf{R}\right) = \frac{1}{\pi^{LM}\det\left(\mathbf{R}\right)^M} \exp\left\{-M\mathrm{tr}\left(\mathbf{R}^{-1}\hat{\mathbf{R}}\right)\right\}, \tag{3.4}$$

donde $\hat{\mathbf{R}}$ es la matriz de covarianza muestral definida como

$$\hat{\mathbf{R}} = \frac{1}{M}\sum_{m=0}^{M-1} \mathbf{x}_m\mathbf{x}_m^H,$$

siendo $\det(\cdot)$ y $\mathrm{tr}(\cdot)$ el determinante y la traza de una matriz, respectivamente.

Para obtener el GLRT, como se ha descrito en el Capítulo 2, es necesario derivar las estimas ML de los parámetros desconocidos bajo ambas hipótesis. Dichas estimas se presentan en los siguientes lemas.

Lema 3.1. *La estima ML de σ^2 viene dada por*

$$\hat{\sigma}^2 = \frac{1}{L}\mathrm{tr}\left(\hat{\mathbf{R}}\right) = \frac{1}{L}\sum_{i=1}^{L} \lambda_i\left(\hat{\mathbf{R}}\right), \tag{3.5}$$

donde $\lambda_i\left(\hat{\mathbf{R}}\right)$ es el i-ésimo autovalor de $\hat{\mathbf{R}}$.

Demostración. Sustituyendo $\mathbf{R} = \sigma^2\mathbf{I}$ en (3.4), la verosimilitud es

$$p\left(\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}; \sigma^2\mathbf{I}\right) = \frac{1}{\pi^{LM}\sigma^{2LM}} \exp\left\{-\frac{M}{\sigma^2}\mathrm{tr}\left(\hat{\mathbf{R}}\right)\right\}. \tag{3.6}$$

La estima ML de σ^2 se obtendrá maximizando (3.6) sujeto a $\sigma^2 > 0$. Teniendo en cuenta la monotonicidad del logaritmo, el problema de optimización resultante es

$$\underset{\sigma}{\text{maximizar}} \quad -LM\log\pi - LM\log\sigma^2 - \frac{M}{\sigma^2}\mathrm{tr}\left(\hat{\mathbf{R}}\right), \tag{3.7}$$

$$\text{sujeto a} \quad \sigma^2 > 0. \tag{3.8}$$

Este problema de optimización puede resolverse mediante el método de los multiplicadores de Lagrange [Boyd y Vandenberghe, 2004]. Nótese, además, que la restricción dada por (3.8) puede omitirse, dado que está implícita en la función objetivo (3.7). Teniendo esto en cuenta, es sencillo resolver el problema de optimización. Para ello, se calcula la derivada de (3.7) respecto a σ y se iguala a 0, obteniéndose (3.5). □

En el lema anterior, se ha derivado la estima ML de σ^2 bajo $\mathcal{H}_0$, que puede interpretarse como la energía promedio de los L sensores. No obstante, el GLRT no necesita dicha estima, sino la verosimilitud comprimida, es decir, la verosimilitud donde todos los parámetros desconocidos han sido sustituidos por sus estimas ML. En este caso, sustituyendo (3.5) en (3.6) y tomando logaritmos, la log-verosimilitud comprimida bajo $\mathcal{H}_0$ es

$$\log p\left(\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}; \hat{\sigma}^2\mathbf{I}\right) = -LM\log\pi - LM - LM\log\left[\frac{1}{L}\mathrm{tr}\left(\hat{\mathbf{R}}\right)\right]. \tag{3.9}$$

Lema 3.2. *La estima ML de* $\mathbf{R}_1$ *viene dada por*

$$\hat{\mathbf{R}}_1 = \hat{\mathbf{R}}. \tag{3.10}$$

Demostración. La log-verosimilitud es

$$\log p\left(\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}; \mathbf{R}_1\right) = -LM\log\pi - M\log\det\left(\mathbf{R}_1\right) - M\mathrm{tr}\left(\mathbf{R}_1^{-1}\hat{\mathbf{R}}\right).$$

De manera análoga a la demostración bajo $\mathcal{H}_0$, omitiendo la restricción de ser definida positiva, tomando el gradiente respecto a $\mathbf{R}_1$ [Magnus y Neudecker, 1999] e igualando a $\mathbf{0}$ se obtiene (3.10). Además, es fácil comprobar que $\hat{\mathbf{R}}_1 = \hat{\mathbf{R}}$ es definida positiva. □

Bajo $\mathcal{H}_1$, se obtiene la siguiente log-verosimilitud comprimida

$$\log p\left(\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}; \hat{\mathbf{R}}_1\right) = -LM\log\pi - LM - M\log\det\left(\hat{\mathbf{R}}\right). \tag{3.11}$$

Finalmente, restando (3.9) y (3.11), el log-GLR[2] es

$$\begin{aligned}\log\mathscr{L} &= \log p\left(\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}; \hat{\sigma}^2\mathbf{I}\right) - \log p\left(\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}; \hat{\mathbf{R}}_1\right) \\ &= \log\left\{\frac{\left[\prod_{i=1}^{L}\lambda_i\left(\hat{\mathbf{R}}\right)\right]^{1/L}}{\frac{1}{L}\sum_{i=1}^{L}\lambda_i\left(\hat{\mathbf{R}}\right)}\right\}.\end{aligned} \tag{3.12}$$

Analicemos el GLRT derivado para el modelo INWSNS. En primer lugar, (3.12) es el cociente entre la media geométrica y la media aritmética de todos los autovalores de la matriz de covarianza muestral. Además, el GLRT es invariante a ciertas transformaciones, propiedad que se discute en el siguiente lema.

Lema 3.3. *El GLRT* (3.12) *es invariante a transformaciones* $\mathbf{y} = \mathbf{A}\mathbf{x}$*, con* $\mathbf{A}^H\mathbf{A} = |a|^2\mathbf{I}$ *y* $|a| \neq 0$.

Demostración. Es fácil demostrar que los autovalores de la matriz de covarianza muestral de $\mathbf{y}$ vienen dados por

$$\lambda_i^{(y)} = |a|^2\lambda_i^{(x)}, \quad i = 1, \ldots, L,$$

donde $\lambda_i^{(x)}$ y $\lambda_i^{(y)}$ son los autovalores de las matrices de covarianza muestrales de $\mathbf{x}$ e $\mathbf{y}$, respectivamente. Finalmente, la demostración concluye sustituyendo $\lambda_i^{(y)} = |a|^2\lambda_i^{(x)}$ en (3.12). □

[2]Por simplicidad notacional, en esta Tesis, se omiten factores de escala y términos constantes en algunas igualdades, dado que no modifican el comportamiento del test. No obstante, deben tenerse en cuenta cuando se aplique el teorema de Wilks.

El Lema 3.3 establece la invarianza del GLRT a multiplicaciones de las observaciones por matrices unitarias y escalados, siendo sencillo demostrar que el problema de detección dado por (3.3) también es invariante a este tipo de transformaciones. Esta es una propiedad muy interesante pues permite obtener, de manera numérica, la distribución del GLRT bajo $\mathcal{H}_0$, necesaria para determinar el umbral de detección. Para ello, mediante simulaciones, se obtiene la distribución del estadístico para $\sigma^2 = 1$ y, aplicando el Lema 3.3, dicha distribución es válida para cualquier otro valor de σ^2.

Por último, el GLRT (3.12) cumple las condiciones del teorema de Wilks, siendo su distribución asintótica χ^2_ν bajo $\mathcal{H}_0$, donde ν depende de la diferencia de grados de libertad entre las hipótesis. Bajo $\mathcal{H}_0$ sólo σ^2 es desconocido (un único parámetro desconocido). Por el contrario, bajo $\mathcal{H}_1$ la matriz de covarianza es desconocida, que posee L^2 parámetros reales libres, valor obtenido al sumar los grados de libertad correspondientes a los $L(L-1)/2$ elementos complejos por encima de la diagonal y a las L componentes reales en la diagonal. Finalmente, teniendo en cuenta el número de parámetros desconocidos bajo ambas hipótesis, el GLRT bajo $\mathcal{H}_0$ se distribuye como

$$-2\log\mathscr{L} \overset{a}{\sim} \chi^2_{L^2-1}.$$

3.1.2. Señales con Estructura Espacial

Esta subsección generaliza el problema anterior incorporando estructura espacial al problema de detección, que surge de manera natural, por ejemplo, en problemas de comunicaciones inalámbricas. Concretamente, se deriva el GLRT para el modelo INWSPS, cuyo test de hipótesis es

$$\begin{aligned}\mathcal{H}_1 &: \mathbf{x} \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{H}\mathbf{R}_s\mathbf{H}^H + \sigma^2\mathbf{I}\right),\\ \mathcal{H}_0 &: \mathbf{x} \sim \mathcal{CN}\left(\mathbf{0}, \sigma^2\mathbf{I}\right),\end{aligned}$$

donde $\mathbf{H} \in \mathbb{C}^{L\times P}$ es completa en rango de columnas, o de manera equivalente, todas sus columnas son linealmente independientes, $\mathbf{R}_s \in \mathbb{C}^{P\times P}$ es la matriz de covarianza de la señal transmitida y $\sigma^2 > 0$. Dado que $\mathbf{H}$ y $\mathbf{R}_s$ son desconocidas, existe una ambigüedad sobre ellas y no se pueden estimar unívocamente. Debido a esto, se supone $\mathbf{R}_s = \mathbf{I}$, quedando el test reducido a

$$\begin{aligned}\mathcal{H}_1 &: \mathbf{x} \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{H}\mathbf{H}^H + \sigma^2\mathbf{I}\right),\\ \mathcal{H}_0 &: \mathbf{x} \sim \mathcal{CN}\left(\mathbf{0}, \sigma^2\mathbf{I}\right).\end{aligned} \tag{3.13}$$

El test de hipótesis (3.13) considera una matriz de covarianza bajo $\mathcal{H}_1$ con estructura adicional (además de ser definida positiva), que será necesario imponer para derivar las estimas ML. Dicha estructura permitirá mejorar la estima de $\mathbf{R}_1$ y, por consiguiente, las prestaciones del detector.

Bajo $\mathcal{H}_0$, los modelos INWSPS e INWSNS (resuelto en el apartado anterior) son idénticos, por lo tanto, la estima de σ^2 bajo $\mathcal{H}_0$ viene dada por (3.5). Por otro lado, la estructura espacial inducida por este modelo depende de P. Para un valor de P pequeño, $\mathbf{H}\mathbf{H}^H$ es una matriz deficiente en rango y dicha estructura espacial se usará en la estimación ML. Por el contrario, para un valor de P elevado, $\mathbf{H}\mathbf{H}^H$ únicamente posee la estructura de una matriz definida positiva. El siguiente lema, establece para qué valores de P se pierde la estructura espacial adicional.

Lema 3.4. *Para $P \geq L-1$, la matriz $\mathbf{R}_1 = \mathbf{H}\mathbf{H}^H + \sigma^2\mathbf{I}$ es una matriz definida positiva sin otra estructura adicional.*

Demostración. La demostración de este lema se basa en contar el número de parámetros independientes de $\mathbf{R}_1$ y comprobar si es mayor o menor que el número de parámetros de

una matriz definida positiva. Para simplificar el cálculo reparametricemos $\mathbf{R}_1$. Comencemos calculando la descomposición en autovalores (EVD) de $\mathbf{H}\mathbf{H}^H$, dada por $\mathbf{H}\mathbf{H}^H = \mathbf{U}\mathbf{\Psi}^2\mathbf{U}^H$, donde $\mathbf{U} \in \mathbb{C}^{L\times P}$ es una matriz unitaria de autovectores y $\mathbf{\Psi}^2 = \operatorname{diag}\left(\psi_1^2, \ldots, \psi_P^2\right) \in \mathbb{C}^{P\times P}$ es una matriz diagonal de valores propios. $\mathbf{\Psi}^2$ tiene P elementos reales y $\mathbf{U}$ tiene $2LP$ elementos reales, sin embargo no todos son independientes. Cada columna tiene norma unidad, por lo tanto, hay P parámetros libres menos. Además, un cambio de fase común a todos los elementos de una columna de $\mathbf{U}$ no modifica $\mathbf{H}\mathbf{H}^H$, lo que resta otros P grados de libertad. Finalmente, las columnas de $\mathbf{U}$ deben ser mutuamente ortogonales, eliminando los siguiente parámetros libres

$$2\sum_{i=1}^{P-1} i = 2\frac{P(P-1)}{2}.$$

Teniendo en cuenta la varianza del ruido, σ^2, el número total de parámetros reales libres de $\mathbf{R}_1 = \mathbf{H}\mathbf{H}^H + \sigma^2\mathbf{I}$ es

$$\underbrace{P}_{\mathbf{\Psi}^2} + \underbrace{2LP - P - P - P(P-1)}_{\mathbf{U}} + \underbrace{1}_{\sigma^2} = 2LP - P^2 + 1. \tag{3.14}$$

Por otro lado, una matriz definida positiva Hermítica tiene $L(L-1)/2$ elementos complejos por encima de la diagonal y L componentes reales en la diagonal, es decir, sus grados de libertad son

$$2\frac{L(L-1)}{2} + L = L^2. \tag{3.15}$$

Comparando (3.15) y (3.14) obtenemos que si se cumple

$$2LP - P^2 + 1 \geq L^2 \Rightarrow P \geq L - 1,$$

el número de variables libres en $\mathbf{R}_1$ es mayor que el de una matriz definida positiva arbitraria, por lo tanto concluimos que la matriz de covarianza no tiene estructura adicional que se pueda explotar. □

Para $P \geq L-1$, la matriz de covarianza bajo $\mathcal{H}_1$ no tiene estructura espacial adicional y el problema de detección se reduce al presentado en la Sección 3.1.1. Por otro lado, para $P < L-1$ sí es posible explotar el rango reducido en la estimación ML de la matriz de covarianza. No obstante, la derivación de las estimas ML bajo $\mathcal{H}_1$ de $\mathbf{H}$ y σ^2 es un problema complicado que puede simplificarse usando la parametrización presentada en la demostración del Lema 3.4. Por lo tanto, el Lema 3.5 presenta, para $P < L-1$, las estimas ML de los autovalores y autovectores de $\mathbf{R}_1$.

Lema 3.5. *Teniendo en cuenta la EVD de la matriz de covarianza muestral,* $\hat{\mathbf{R}} = \mathbf{W}\operatorname{diag}(\lambda_1, \ldots, \lambda_L)\,\mathbf{W}^H$*, con* $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_L$*, las estimas ML de* $\mathbf{U}$*,* $\mathbf{\Psi}^2$ *y* σ^2 *vienen dadas por*

$$\begin{aligned}
\hat{\mathbf{U}} &= \mathbf{W}_{\mathrm{s}}, \\
\hat{\psi}_i^2 &= \lambda_i - \hat{\sigma}^2, \quad i = 1, \ldots, P, \\
\hat{\sigma}^2 &= \frac{1}{L-P}\sum_{k=P+1}^{L} \lambda_k,
\end{aligned}$$

donde $\mathbf{W}_s$ *son los autovectores asociados a los* P *mayores autovalores de* $\hat{\mathbf{R}}$*, es decir*

$$\mathbf{W} = [\ \overbrace{\mathbf{W}_\mathrm{s}}^{P\ columnas}\quad \overbrace{\mathbf{W}_\mathrm{n}}^{L-P\ columnas}\ ].$$

Demostración.[3] Aplicando la EVD a $\mathbf{R}_1$ se obtiene $\mathbf{R}_1 = \mathbf{U}_\mathrm{c}(\mathbf{\Psi}_\mathrm{c}^2 + \sigma^2\mathbf{I})\mathbf{U}_\mathrm{c}^H$, donde ahora $\mathbf{U}_\mathrm{c} \in \mathbb{C}^{L\times L}$ es una matriz unitaria de autovectores que incluye el subespacio de ruido y $\mathbf{\Psi}_\mathrm{c}^2 = \mathrm{diag}\left(\psi_1^2, \ldots, \psi_P^2, 0, \ldots, 0\right) \in \mathbb{C}^{L\times L}$ es una matriz diagonal de valores propios. Es decir, en lugar de usar la EVD reducida, se usa la EVD completa. Sustituyendo $\mathbf{R}_1$ por su EVD, la log-verosimilitud viene dada por

$$\begin{aligned}\log p\left(\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}; \mathbf{U}_\mathrm{c}, \mathbf{\Psi}_\mathrm{c}^2, \sigma^2\right) = &- LM\log\pi - M\log\det\left(\mathbf{\Psi}_\mathrm{c}^2 + \sigma^2\mathbf{I}\right)\\ &- M\mathrm{tr}\left[\left(\mathbf{\Psi}_\mathrm{c}^2 + \sigma^2\mathbf{I}\right)^{-1}\mathbf{U}_\mathrm{c}^H\hat{\mathbf{R}}\mathbf{U}_\mathrm{c}\right],\end{aligned}$$

y definiendo $\mathbf{A} = \mathbf{U}_\mathrm{c}^H\hat{\mathbf{R}}\mathbf{U}_\mathrm{c}$, puede reescribirse como

$$\begin{aligned}\log p\left(\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}; \mathbf{U}_\mathrm{c}, \mathbf{\Psi}_\mathrm{c}^2, \sigma^2\right) = &-LM\log\pi - M\sum_{i=1}^{P}\log\left(\psi_i^2 + \sigma^2\right) - M(L-P)\log\sigma^2\\ &- M\sum_{i=1}^{P}\frac{[\mathbf{A}]_{i,i}}{\psi_i^2 + \sigma^2} - \frac{M}{\sigma^2}\sum_{i=P+1}^{L}[\mathbf{A}]_{i,i}. \quad (3.16)\end{aligned}$$

Calculando derivadas e igualando a cero, es fácil demostrar que (3.16) se maximiza respecto a ψ_i^2 y σ^2 para

$$\hat{\sigma}^2 = \frac{1}{L-P}\sum_{i=P+1}^{L}[\mathbf{A}]_{i,i}, \quad (3.17)$$

$$\hat{\psi}_i^2 = [\mathbf{A}]_{i,i} - \hat{\sigma}^2, \quad i = 1, \ldots, P. \quad (3.18)$$

Sustituyendo ahora (3.17) y (3.18) en (3.16), la estimación ML se reduce a maximizar

$$\begin{aligned}\log p\left(\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}; \mathbf{U}_\mathrm{c}, \hat{\mathbf{\Psi}}_\mathrm{c}^2, \hat{\sigma}^2\right) = &- LM\left(\log\pi + 1\right) - M\sum_{i=1}^{P}\log[\mathbf{A}]_{i,i}\\ &- M(L-P)\log\left(\frac{1}{L-P}\sum_{i=P+1}^{L}[\mathbf{A}]_{i,i}\right), \quad (3.19)\end{aligned}$$

sujeto a $\mathbf{A} = \mathbf{U}_\mathrm{c}^H\hat{\mathbf{R}}\mathbf{U}_\mathrm{c}$ y $\mathbf{U}_\mathrm{c}^H\mathbf{U}_\mathrm{c} = \mathbf{I}_L$. Se puede comprobar que (3.19) es una función Schur-convexa [Marshall y Olkin, 1979, Jorswieck y Boche, 2007] de $[\mathbf{A}]_{i,i},\ i = 1, \ldots, L$. Entonces, (3.19) tiene una cota superior dada por el i-ésimo autovalor de $\mathbf{A}$ [Jorswieck y Boche, 2007], debido a que el vector de autovalores de una matriz mayoriza el vector compuesto por los elementos de la diagonal de dicha matriz. Dicha cota superior se alcanzará para $\hat{\mathbf{U}}_\mathrm{c} = \mathbf{W}$, obteniéndose una matriz $\mathbf{A}$ diagonal. □

[3] La demostración de este lema se puede encontrar en [Anderson, 1963], sin embargo, aquí vamos a presentar una demostración alternativa basada en mayorización.

Teniendo en cuenta las estimas ML del Lema 3.5, la log-verosimilitud comprimida bajo $\mathcal{H}_1$ viene dada por

$$\begin{aligned}\log p\left(\mathbf{x}_0,\ldots,\mathbf{x}_{M-1};\hat{\mathbf{H}},\hat{\sigma}^2\right) =& -LM - LM\log\pi - M\sum_{i=1}^{P}\log\lambda_i \\ & - M(L-P)\log\left(\frac{1}{L-P}\sum_{i=P+1}^{L}\lambda_i\right). \end{aligned} \tag{3.20}$$

Sustituyendo (3.20) y (3.9) en el log-GLRT, se obtiene

$$\begin{aligned}\log\mathscr{L} =& \log p\left(\mathbf{x}_0,\ldots,\mathbf{x}_{M-1};\hat{\sigma}^2\mathbf{I}\right) - \log p\left(\mathbf{x}_0,\ldots,\mathbf{x}_{M-1};\hat{\mathbf{H}},\hat{\sigma}^2\right) \\ =& \sum_{i=1}^{P}\log\lambda_i + (L-P)\log\left(\frac{1}{L-P}\sum_{i=P+1}^{L}\lambda_i\right) - L\log\left[\frac{1}{L}\mathrm{tr}\left(\hat{\mathbf{R}}\right)\right].\end{aligned}$$

Finalmente, sumando y restando $\log\left[\prod_{i=P+1}^{L}\lambda_i\right]$, el log-GLRT es

$$\log\mathscr{L} = L\log\left[\frac{\left(\prod_{i=1}^{L}\lambda_i\right)^{1/L}}{\frac{1}{L}\sum_{i=1}^{L}\lambda_i}\right] - (L-P)\log\left[\frac{\left(\prod_{i=P+1}^{L}\lambda_i\right)^{1/(L-P)}}{\frac{1}{L-P}\sum_{i=P+1}^{L}\lambda_i}\right]. \tag{3.21}$$

Los dos términos de (3.21) vienen dados, respectivamente, por el cociente entre la media geométrica y aritmética de todos los autovalores y de los $L-P$ menores autovalores de $\hat{\mathbf{R}}$. Por otro lado, el primer término de (3.21) es el estadístico del test de esfericidad (3.12), y el segundo puede interpretarse como el estadístico del test de esfericidad del subespacio de ruido, siendo el GLRT un cociente de esfericidades. Alternativamente, el segundo término puede interpretarse como una referencia de esfericidad debida a los efectos de disponer de un número finito de muestras (*finite sample size effects*). Esto es debido a que cuando el número de muestras tiende a infinito ($M\to\infty$), la matriz de covarianza muestral converge a la teórica ($\hat{\mathbf{R}}\to\mathbf{R}$). Por lo tanto, los $L-P$ autovalores más pequeños convergen a σ^2 ($\lambda_i\to\sigma^2$ para $i=P+1,\ldots,L$), y el segundo término de (3.21) tiende a 0.

En el siguiente lema se presenta una invarianza del GLRT que se acaba de obtener.

Lema 3.6. *El GLRT* (3.21) *es invariante a transformaciones* $\mathbf{y}=\mathbf{A}\mathbf{x}$, *con* $\mathbf{A}^H\mathbf{A}=|a|^2\mathbf{I}$ *y* $|a|\neq 0$.

Demostración. La demostración es equivalente a la del Lema 3.3. □

Como en la sección anterior, este lema permite obtener el umbral de detección para una probabilidad de falsa alarma dada, independientemente del valor de σ^2. Además, para este GLRT tiene mayor importancia, dado que no se cumplen las condiciones del Teorema de Wilks. En concreto, las estimas ML no se obtienen igualando a cero la log-verosimilitud. Más adelante, se presentarán unos resultados numéricos que ilustran este comportamiento.

Relación con Trabajos Previos

En este apartado se relaciona el GLRT dado por (3.21) con trabajos previos. Comencemos por el caso particular de $P = 1$. Para este modelo, cuyo test de hipótesis es

$$\begin{aligned}\mathcal{H}_1 &: \mathbf{x} \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{h}\mathbf{h}^H + \sigma^2\mathbf{I}\right),\\ \mathcal{H}_0 &: \mathbf{x} \sim \mathcal{CN}\left(\mathbf{0}, \sigma^2\mathbf{I}\right),\end{aligned}$$

se demostró en [Besson et al., 2006] que no existe un test UMPI para valores de $L > 2$ y se derivó el siguiente GLRT

$$\log \mathscr{L} = \log \left[\frac{\operatorname{tr}\left(\hat{\mathbf{R}}\right)}{\lambda_1}\right]. \tag{3.22}$$

Además, el GLRT (3.22) se ha aplicado a sensado espectral para radio cognitiva en [Taherpour et al., 2010, Wang et al., 2010].

Finalmente, el modelo INWSPS ha sido considerado también en [Lim et al., 2008] donde se deriva un detector de manera heurística. Concretamente, se obtiene el GLRT para el modelo INWSNS (sin estructura espacial), suponiendo conocida la varianza del ruido. Bajo estas suposiciones el GLRT es

$$\log \mathscr{L} = \sum_{i=1}^{\tilde{L}} \left(\frac{\lambda_i}{2\sigma^2} - 1\right) - \sum_{i=1}^{\tilde{L}} \log\left(\frac{\lambda_i}{2\sigma^2}\right),$$

donde $\tilde{L}$ es el mayor número natural tal que $\lambda_{\tilde{L}} > 2\sigma^2$. A partir del GLRT anterior, los autores incorporan estructura espacial y proponen estimar, bajo ambas hipótesis, la varianza de ruido como el menor autovalor de la matriz de covarianza muestral, es decir, $\hat{\sigma}^2 = \lambda_L$. Sustituyendo esta estima en las verosimilitudes, el detector propuesto en [Lim et al., 2008] para el modelo INWSPS es

$$T_{\text{SSE}} = \frac{1}{P}\sum_{i=1}^{P} \frac{\lambda_i}{\lambda_L} - \log\left(\prod_{i=1}^{P} \frac{\lambda_i}{\lambda_L}\right)^{1/P}. \tag{3.23}$$

Por último, nótese que el detector (3.23) no se corresponde con el GLRT para el modelo INWSPS.

3.2. Detección Multicanal de Procesos Estacionarios con PSD Arbitrarias

Esta sección extiende los detectores de la sección anterior considerando series temporales estacionarias con densidades espectrales de potencia (PSD) arbitrarias. En concreto, se deriva el GLRT para los modelos INSSNS e INSSPS, que tienen las siguientes funciones de covarianza

$$\begin{aligned}\mathcal{H}_1 &: E\left[\mathbf{x}[n]\mathbf{x}^H[n-m]\right] = \mathbf{R}_1[m],\\ \mathcal{H}_0 &: E\left[\mathbf{x}[n]\mathbf{x}^H[n-m]\right] = \sigma^2[m]\mathbf{I}.\end{aligned}$$

Para obtener el test, comencemos por construir la matriz de datos $\mathbf{X}$

$$\mathbf{X} = \begin{bmatrix} x_1[0] & x_1[1] & \dots & x_1[N-1] \\ x_2[0] & x_2[1] & \dots & x_2[N-1] \\ \vdots & \vdots & \ddots & \vdots \\ x_L[0] & x_L[1] & \dots & x_L[N-1] \end{bmatrix},$$

donde la i-ésima fila, $\mathbf{x}_i^T = [x_i[0], x_i[1], \ldots, x_i[N-1]]$, contiene N muestras de la i-ésima serie temporal $\{x_i[n]\}$, y la k-ésima columna es la k-ésima muestra de la serie temporal multivariada $\{\mathbf{x}[n]\}$. A continuación, se define el vector $\mathbf{z}$, que apila las columnas de $\mathbf{X}$, de la siguiente manera

$$\mathbf{z} = \mathrm{vec}(\mathbf{X}) = \begin{bmatrix}\mathbf{x}^T[0] & \mathbf{x}^T[1] & \ldots & \mathbf{x}^T[N-1]\end{bmatrix}^T.$$

La matriz de covarianza del vector aleatorio $\mathbf{z}$ captura toda la información estadística de segundo orden de la serie $\mathbf{x}[n]$, y viene dada por

$$\mathcal{R} = \begin{bmatrix} \mathbf{R}[0] & \mathbf{R}[-1] & \cdots & \mathbf{R}[-N+1] \\ \mathbf{R}[1] & \mathbf{R}[0] & \cdots & \mathbf{R}[-N+2] \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{R}[N-1] & \mathbf{R}[N-2] & \cdots & \mathbf{R}[0] \end{bmatrix} \in \mathbb{C}^{LN\times LN}, \tag{3.24}$$

donde $\mathbf{R}[m] = E\left[\mathbf{x}[n]\mathbf{x}^H[n-m]\right]$ es una secuencia de covarianza matricial y se han considerado señales estacionarias, resultando una matriz de covarianza, $\mathcal{R}$, Toeplitz por bloques. Por lo tanto, suponiendo gaussianidad, en esta sección se resuelve el siguiente problema de detección

$$\begin{aligned} \mathcal{H}_1 &: \mathbf{z} \sim \mathcal{CN}\left(\mathbf{0}_{LN}, \mathcal{R}_1\right), \\ \mathcal{H}_0 &: \mathbf{z} \sim \mathcal{CN}\left(\mathbf{0}_{LN}, \mathcal{R}_0\right), \end{aligned} \tag{3.25}$$

donde $\mathcal{R}_1$ y $\mathcal{R}_0$ son las matrices (Toeplitz por bloques) de covarianza bajo $\mathcal{H}_1$ y $\mathcal{H}_0$, respectivamente.

3.2.1. La Verosimilitud Asintótica

La derivación del GLRT para el test de hipótesis (3.25) requiere obtener estimas ML de matrices de covarianza Toeplitz por bloques. Sin embargo, la estimación ML de matrices Toeplitz, y por consiguiente Toeplitz por bloques, es un problema de optimización no convexo [Burg et al., 1982, Scharf, 1991]. Debido a esto, emplearemos la verosimilitud asintótica, en el dominio frecuencial, que converge en media cuadrática a la verosimilitud en el dominio del tiempo cuando el número de muestras, N, tiende a infinito.

A continuación, se presentan unas definiciones para, posteriormente, enunciar el teorema que establece la convergencia de las verosimilitudes. Considerando un experimento que produce M ($M \geq L$) realizaciones independientes de la matriz de datos o, de manera equivalente, de su vectorización, $\mathbf{z}_i = \begin{bmatrix}\mathbf{x}_i^T[0] & \ldots & \mathbf{x}_i^T[N-1]\end{bmatrix}^T$, $i = 0, \ldots, M-1$, la log-verosimilitud (en el dominio del tiempo) viene dada por

$$\log p\left(\mathbf{z}_0, \ldots, \mathbf{z}_{M-1}; \mathcal{R}\right) = -LNM\log\pi - M\log\det\left(\mathcal{R}\right) - M\mathrm{tr}\left(\mathcal{R}^{-1}\hat{\mathbf{R}}\right), \tag{3.26}$$

y la log-verosimilitud asintótica es

$$\begin{aligned} \log p\left(\mathbf{z}_0, \ldots, \mathbf{z}_{M-1}; \mathbf{S}\left(e^{j\theta}\right)\right) = &- LNM\log\pi - NM\int_{-\pi}^{\pi}\log\det\left(\mathbf{S}\left(e^{j\theta}\right)\right)\frac{d\theta}{2\pi} \\ &- NM\int_{-\pi}^{\pi}\mathrm{tr}\left(\mathbf{S}^{-1}\left(e^{j\theta}\right)\hat{\mathbf{S}}\left(e^{j\theta}\right)\right)\frac{d\theta}{2\pi}, \end{aligned} \tag{3.27}$$

donde $\mathcal{R}$ es la matriz teórica de covarianza con estructura Toeplitz por bloques, $\mathbf{S}\left(e^{j\theta}\right) = \mathcal{F}\left(\mathbf{R}[m]\right)$ es la matriz teórica de densidades espectrales de potencia y sus versiones muestrales

son

$$\hat{\mathbf{R}} = \frac{1}{M}\sum_{i=0}^{M-1} \mathbf{z}_i \mathbf{z}_i^H,$$

y

$$\hat{\mathbf{S}}\left(e^{j\theta}\right) = \frac{1}{M}\sum_{i=0}^{M-1} \mathbf{x}_i\left(e^{j\theta}\right)\mathbf{x}_i^H\left(e^{j\theta}\right),$$

respectivamente, con

$$\mathbf{x}_i\left(e^{j\theta}\right) = \frac{1}{\sqrt{N}}\sum_{n=0}^{N-1} \mathbf{x}_i[n]\, e^{-j\theta n}.$$

Por último, $\mathcal{F}(\cdot)$ denota la transformada de Fourier en tiempo discreto.

Teorema 3.7. *Bajo los siguientes supuestos:*

c.1 las matrices Toeplitz por bloques son generadas por símbolos continuos, es decir, cada bloque son los coeficientes de Fourier de una función matricial continua de dimensiones $L \times L$; y

c.2 las matrices de densidades espectrales de potencia son definidas positivas,

la log-verosimilitud asintótica converge en media cuadrática a la log-verosimilitud, esto es

$$\lim_{N\to\infty} E\left[\left|\frac{1}{N}\left[\log p\left(\mathbf{z}_0,\ldots,\mathbf{z}_{M-1};\mathcal{R}\right) - \log p\left(\mathbf{z}_0,\ldots,\mathbf{z}_{M-1};\mathbf{S}\left(e^{j\theta}\right)\right)\right]\right|^2\right] = 0. \tag{3.28}$$

Demostración. La demostración puede encontrarse en el Apéndice A. □

Como una consecuencia directa del Teorema 3.7, el test (3.25) puede reescribirse como

$$\begin{aligned}\mathcal{H}_1 &: \mathbf{x}\left(e^{j\theta}\right) \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{S}_1\left(e^{j\theta}\right)\right),\\ \mathcal{H}_0 &: \mathbf{x}\left(e^{j\theta}\right) \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{S}_0\left(e^{j\theta}\right)\right).\end{aligned}$$

De esta manera, testeamos la diferente estructura espacial, bajo cada hipótesis, de las matrices de densidades espectrales de potencia.

3.2.2. Señales sin Estructura Espacial

Esta subsección presenta el GLRT para el modelo INSSNS, es decir, ruidos IID y señales estacionarias con PSD arbitrarias sin estructura espacial. Bajo estos supuestos, aplicando la verosimilitud asintótica, el test de hipótesis considerado en este apartado es

$$\begin{aligned}\mathcal{H}_1 &: \mathbf{x}\left(e^{j\theta}\right) \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{S}_1\left(e^{j\theta}\right)\right),\\ \mathcal{H}_0 &: \mathbf{x}\left(e^{j\theta}\right) \sim \mathcal{CN}\left(\mathbf{0}, S_n\left(e^{j\theta}\right)\mathbf{I}\right),\end{aligned} \tag{3.29}$$

con $S_n\left(e^{j\theta}\right) = \mathcal{F}\left(r_n[m]\right) > 0,\ \forall\theta$, y $\mathbf{S}_1\left(e^{j\theta}\right) = \mathcal{F}\left(\mathbf{R}_1[m]\right) \in \mathbb{S}_+,\ \forall\theta$.

El GLR para este test de hipótesis es

$$\mathscr{L} = \frac{\displaystyle\max_{S_n\left(e^{j\theta}\right)>0} p\left(\mathbf{z}_0,\ldots,\mathbf{z}_{M-1};S_n\left(e^{j\theta}\right)\mathbf{I}\right)}{\displaystyle\max_{\mathbf{S}_1\left(e^{j\theta}\right)\in\mathbb{S}_+} p\left(\mathbf{z}_0,\ldots,\mathbf{z}_{M-1};\mathbf{S}_1\left(e^{j\theta}\right)\right)},$$

siendo necesario derivar las estimas ML de los parámetros desconocidos bajo cada hipótesis para obtenerlo.

Como se ha dicho previamente, se suponen señales con PSD arbitrarias, es decir, no consideramos modelos paramétricos como, por ejemplo, modelos autoregresivos (AR), de media móvil (MA) o autoregresivos de media móvil (ARMA). Teniendo esto en cuenta, no existe ningún conjunto de parámetros más reducido que las matrices de densidades espectrales de potencia en todas las frecuencias y, por ello, la estimación ML se puede hacer independientemente para cada frecuencia. Además, el test de hipótesis dado por (3.29) tiene la misma estructura espacial (a cada frecuencia) que el test de esfericidad (3.3). Por lo tanto, las estimas ML del test de esfericidad son válidas para este problema, sustituyendo las matrices de covarianza por las matrices de densidades espectrales de potencia. Resumiendo, la estima ML bajo $\mathcal{H}_0$ viene dada por

$$\hat{S}_{\mathrm{n}}\left(e^{j\theta}\right) = \frac{1}{L}\mathrm{tr}\left[\hat{\mathbf{S}}\left(e^{j\theta}\right)\right], \tag{3.30}$$

y bajo $\mathcal{H}_1$ es

$$\hat{\mathbf{S}}_1\left(e^{j\theta}\right) = \hat{\mathbf{S}}\left(e^{j\theta}\right),$$

obteniéndose el siguiente log-GLRT asintótico

$$\log\mathscr{L} = \int_{-\pi}^{\pi} \log\left\{\frac{\left[\prod_{i=1}^{L}\lambda_i\left(e^{j\theta}\right)\right]^{1/L}}{\frac{1}{L}\sum_{i=1}^{L}\lambda_i\left(e^{j\theta}\right)}\right\}\frac{d\theta}{2\pi}, \tag{3.31}$$

donde $\lambda_i\left(e^{j\theta}\right)$ es el i-ésimo autovalor de $\hat{\mathbf{S}}\left(e^{j\theta}\right)$. El GLRT (3.31) se puede interpretar como la integral del test de esfericidad, y además es invariante a transformaciones que permiten obtener el umbral de detección mediante simulaciones, independientemente del valor de la PSD del ruido. Dicha invarianza se presenta en el siguiente lema.

Lema 3.8. *El GLRT* (3.31) *es invariante a transformaciones* $\mathbf{y}[n] = (\mathbf{A} * \mathbf{x})[n]$, *con* $\mathbf{A}^H\left(e^{j\theta}\right)\mathbf{A}\left(e^{j\theta}\right) = \left|A\left(e^{j\theta}\right)\right|^2\mathbf{I}$, $\left|A\left(e^{j\theta}\right)\right| \neq 0, \forall\theta$, *donde* $\mathbf{A}\left(e^{j\theta}\right) = \mathcal{F}\left(\mathbf{A}[n]\right)$.

Demostración. Los autovalores de la matriz muestral de densidades espectrales de potencia de $\mathbf{y}[n]$ vienen dados por

$$\lambda_i^{(y)}\left(e^{j\theta}\right) = \left|A\left(e^{j\theta}\right)\right|^2\lambda_i^{(x)}\left(e^{j\theta}\right), \quad i = 1, \ldots, L,$$

donde $\lambda_i^{(x)}\left(e^{j\theta}\right)$ y $\lambda_i^{(y)}\left(e^{j\theta}\right)$ son los autovalores de las matrices PSD muestrales de $\mathbf{x}[n]$ e $\mathbf{y}[n]$, respectivamente. Finalmente, la demostración concluye sustituyendo $\lambda_i^{(y)}\left(e^{j\theta}\right) = \left|A\left(e^{j\theta}\right)\right|^2\lambda_i^{(x)}\left(e^{j\theta}\right)$ en (3.31). □

Para el log-GLRT (3.31) es posible obtener de manera analítica la distribución bajo ambas hipótesis. Teniendo en cuenta la independencia asintótica de las estimas $\hat{\mathbf{S}}\left(e^{j\theta}\right)$ en diferentes frecuencias [Brillinger, 1981] y aplicando el teorema del límite central [Papoulis y Pillai, 2002], el log-GLRT se distribuye, bajo ambas hipótesis, como

$$\log\mathscr{L} \overset{a}{\sim} \mathcal{N}\left(\mu, \sigma^2\right),$$

donde $\overset{a}{\sim}$ indica distribución asintótica en N, y μ y σ^2 son la media y la varianza, respectivamente. Dichos parámetros son, en general, muy difíciles de obtener, dado que son funciones complicadas de los autovalores de matrices Wishart. No obstante, como se ha mencionado previamente, el GLRT es la integral (suma) de los GLRT para señales blancas, y recordando que, asintóticamente bajo $\mathcal{H}_0$, se distribuyen de manera χ^2, es posible calcular la media y la varianza analíticamente bajo esta hipótesis. En particular, es la suma de N variables aleatorias $\chi^2_{L^2-1}$, por lo tanto, el log-GLRT se distribuye como

$$-2\log\mathscr{L} \overset{a}{\sim} \mathcal{N}\left(NL^2 - N, 2NL^2 - 2N\right),$$

donde ahora $\overset{a}{\sim}$ indica distribución asintótica en N y M. Finalmente, la media y la varianza bajo $\mathcal{H}_1$ se obtendrán mediante simulaciones, las cuales dependerán del valor concreto de $\mathbf{R}_1[m]$.

3.2.3. Señales con Estructura Espacial

A continuación, se consideran señales con estructura espacial con densidades espectrales de potencia arbitrarias y contaminadas por ruidos IID, es decir, el modelo INSSPS. Para este caso, el test de hipótesis se puede formular, en el dominio frecuencial, como

$$\begin{aligned}\mathcal{H}_1 &: \mathbf{x}\left(e^{j\theta}\right) \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{H}\left(e^{j\theta}\right)\mathbf{S}_s\left(e^{j\theta}\right)\mathbf{H}^H\left(e^{j\theta}\right) + S_n\left(e^{j\theta}\right)\mathbf{I}\right),\\ \mathcal{H}_0 &: \mathbf{x}\left(e^{j\theta}\right) \sim \mathcal{CN}\left(\mathbf{0}, S_n\left(e^{j\theta}\right)\mathbf{I}\right),\end{aligned}$$

con $\mathbf{H}\left(e^{j\theta}\right) = \mathcal{F}\left(\mathbf{H}[n]\right) \in \mathbb{C}^{L\times P}$ y $\mathbf{S}_s\left(e^{j\theta}\right) = \mathcal{F}\left(\mathbf{R}_s[m]\right) \in \mathbb{C}^{P\times P}$. Dado que ni el canal, ni los estadísticos de la señal son conocidos, es imposible estimar $\mathbf{H}\left(e^{j\theta}\right)$ y $\mathbf{S}_\mathrm{s}\left(e^{j\theta}\right)$ unívocamente, existiendo un problema de identificabilidad de ambos factores. Debido a esto, suponemos señales blancas espacial y temporalmente, es decir, la matriz de densidades espectrales de potencia de $\mathbf{s}[n]$ es $\mathbf{S}_s\left(e^{j\theta}\right) = \mathbf{I}$, $\forall\theta$, obteniéndose el siguiente test de hipótesis simplificado

$$\begin{aligned}\mathcal{H}_1 &: \mathbf{x}\left(e^{j\theta}\right) \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{H}\left(e^{j\theta}\right)\mathbf{H}^H\left(e^{j\theta}\right) + S_n\left(e^{j\theta}\right)\mathbf{I}\right),\\ \mathcal{H}_0 &: \mathbf{x}\left(e^{j\theta}\right) \sim \mathcal{CN}\left(\mathbf{0}, S_n\left(e^{j\theta}\right)\mathbf{I}\right).\end{aligned}$$

De manera análoga al apartado anterior, la estimación ML se puede realizar independientemente para cada frecuencia, obteniéndose un modelo equivalente (en el dominio frecuencial) al modelo INWSPS. Teniendo esto en cuenta, las estimas ML se presentan a continuación, y de manera equivalente al escenario con señales blancas, dependerán del valor concreto de P. Para $P \geq L-1$, el problema no tiene estructura de rango reducido, la estima ML de $\mathbf{S}_1\left(e^{j\theta}\right)$ viene dada por la matriz muestral de densidades espectrales de potencia, $\hat{\mathbf{S}}\left(e^{j\theta}\right)$, siendo el GLRT (3.31). No obstante, para $P < L-1$ es posible mejorar las estimas explotando la estructura espacial adicional. Bajo $\mathcal{H}_0$, la estima ML es

$$\hat{S}_n\left(e^{j\theta}\right) = \frac{1}{L}\mathrm{tr}\left[\hat{\mathbf{S}}\left(e^{j\theta}\right)\right].$$

Para obtener las estimas bajo $\mathcal{H}_1$, reescribamos $\mathbf{H}\left(e^{j\theta}\right)\mathbf{H}^H\left(e^{j\theta}\right) + S_n\left(e^{j\theta}\right)\mathbf{I}$ como

$$\mathbf{H}\left(e^{j\theta}\right)\mathbf{H}^H\left(e^{j\theta}\right) + S_n\left(e^{j\theta}\right)\mathbf{I} = \mathbf{U}_\mathrm{c}\left(e^{j\theta}\right)\left[\mathbf{\Psi}^2_\mathrm{c}\left(e^{j\theta}\right) + S_\mathrm{n}\left(e^{j\theta}\right)\mathbf{I}\right]\mathbf{U}^H_\mathrm{c}\left(e^{j\theta}\right),$$

donde $\mathbf{\Psi}^2_\mathrm{c}\left(e^{j\theta}\right) = \mathrm{diag}\left(\Psi^2_1\left(e^{j\theta}\right), \ldots, \Psi^2_P\left(e^{j\theta}\right), 0, \ldots, 0\right)$ es una matriz diagonal de autovalores y $\mathbf{U}_\mathrm{c}\left(e^{j\theta}\right)$ es una matriz de autovectores. Entonces, las estimas ML de $\mathbf{\Psi}^2_\mathrm{c}\left(e^{j\theta}\right)$, $\mathbf{U}_\mathrm{c}\left(e^{j\theta}\right)$

y $S_n\left(e^{j\theta}\right)$ son

$$\begin{aligned}
\hat{\mathbf{U}}_{\mathrm{c}}\left(e^{j\theta}\right) &= \mathbf{W}\left(e^{j\theta}\right), \\
\hat{\psi}_i^2\left(e^{j\theta}\right) &= \lambda_i\left(e^{j\theta}\right) - \hat{S}_n\left(e^{j\theta}\right), \quad i = 1, \ldots, P, \\
\hat{S}_n\left(e^{j\theta}\right) &= \frac{1}{L-P}\sum_{k=P+1}^{L} \lambda_k\left(e^{j\theta}\right),
\end{aligned}$$

donde $\mathbf{W}\left(e^{j\theta}\right)$ es una matriz que contiene todos los autovectores de $\hat{\mathbf{S}}\left(e^{j\theta}\right)$. Teniendo en cuenta las estimas ML bajo ambas hipótesis, el log-GLRT viene dado por

$$\begin{aligned}
\log \mathscr{L} = L \int_{-\pi}^{\pi} \log \left[\frac{\left(\prod_{i=1}^{L} \lambda_i\left(e^{j\theta}\right)\right)^{1/L}}{\frac{1}{L}\sum_{i=1}^{L} \lambda_i\left(e^{j\theta}\right)} \right] \frac{d\theta}{2\pi} \\
- (L-P) \int_{-\pi}^{\pi} \log \left[\frac{\left(\prod_{i=P+1}^{L} \lambda_i\left(e^{j\theta}\right)\right)^{1/(L-P)}}{\frac{1}{L-P}\sum_{i=P+1}^{L} \lambda_i\left(e^{j\theta}\right)} \right] \frac{d\theta}{2\pi}.
\end{aligned} \tag{3.32}$$

El log-GLRT asintótico se puede interpretar, otra vez, como la integral del GLRT para el modelo equivalente con PSD planas. Además, es invariante a un filtrado común de todas las componentes de la serie temporal multivariada, permitiendo obtener los umbrales de detección mediante simulaciones, independientemente del valor de la PSD del ruido.

Lema 3.9. *El GLRT* (3.32) *es invariante a transformaciones* $\mathbf{y}[n] = (\mathbf{A} * \mathbf{x})[n]$, *con* $\mathbf{A}^H\left(e^{j\theta}\right)\mathbf{A}\left(e^{j\theta}\right) = \left|A\left(e^{j\theta}\right)\right|^2 \mathbf{I}$, $\left|A\left(e^{j\theta}\right)\right| \neq 0, \forall\theta$, *siendo* $\mathbf{A}\left(e^{j\theta}\right) = \mathcal{F}\left(\mathbf{A}[n]\right)$.

Demostración. La demostración es equivalente a la del Lema 3.8. □

Por último, aplicando el mismo razonamiento que para el test de esfericidad en el dominio frecuencial, se puede derivar la distribución del log-GLRT, obteniendo

$$\log \mathscr{L} \sim \mathcal{N}\left(\mu, \sigma^2\right).$$

No obstante, en esta ocasión no es posible aplicar el teorema de Wilks y será necesario calcular la media y la varianza mediante simulaciones. Teniendo en cuenta el Lema 3.9, dichos parámetros no dependen de las PSD del ruido bajo $\mathcal{H}_0$, sin embargo, sí dependen de $\mathbf{H}[n]$ bajo $\mathcal{H}_1$.

3.2.4. Señales Coloreadas en Canales Planos

Esta subsección considera el modelo INMSPS: detección de señales temporalmente correladas en canales frecuencialmente planos contaminadas por ruidos IID. A pesar de no ser un modelo habitual, este tipo de señales aparece cuando se desea detectar una señal de comunicaciones sin realizar la sincronización temporal, es decir, cuando la detección se hace

justo después de la conversión analógico-digital (ADC). Otro ejemplo en el campo de las comunicaciones donde surge el modelo INMSPS es la detección de una señal que ha sido precodificada linealmente [Proakis, 1988, Lender, 1987, Tugnait y Gummadavelli, 1997] con un filtro desconocido. En esta situación, incluso si se realiza sincronización y muestreo al periodo de símbolo, se obtiene una señal temporalmente correlada sobre un canal frecuencialmente plano, si el ancho de banda de transmisión es reducido en comparación con el ancho de banda de coherencia del canal.

Para este problema de detección, el modelo de señal es

$$\begin{array}{ll} \mathcal{H}_1 : \mathbf{x}[n] = \mathbf{h}s[n] + \mathbf{v}[n], & n = 0, \ldots, N-1, \\ \mathcal{H}_0 : \mathbf{x}[n] = \mathbf{v}[n], & n = 0, \ldots, N-1, \end{array} \tag{3.33}$$

donde $\mathbf{x}[n] \in \mathbb{C}^L$ es un vector de medidas, $\mathbf{h} \in \mathbb{C}^L$ es un canal con una entrada y múltiples salidas (SIMO) plano en frecuencia,[4] $s[n]$ es un proceso estacionario en sentido amplio de media nula cuya función de covarianza es $r_{\mathrm{s}}[m] = E\left[s[n]s^*[n-m]\right]$ y $\mathbf{v}[n] \in \mathbb{C}^L$ es un vector de ruidos IID con función de covarianza matricial dada por $\mathbf{R}_{\mathrm{v}}[m] = E\left[\mathbf{v}[n]\mathbf{v}^H[n-m]\right] = r_{\mathrm{v}}[n]\mathbf{I}$. Al contrario que en modelos anteriores, donde la señal, el canal y el ruido son planos o selectivos en frecuencia, esta subsección considera señales y ruidos temporalmente correlados, pero canales planos. Aplicando la verosimilitud asintótica, el test (3.33) puede reescribirse como

$$\begin{array}{l} \mathcal{H}_1 : \mathbf{x}\left(e^{j\theta}\right) \sim \mathcal{CN}\left(0, \mathbf{S}_1\left(e^{j\theta}\right)\right), \\ \mathcal{H}_0 : \mathbf{x}\left(e^{j\theta}\right) \sim \mathcal{CN}\left(0, \mathbf{S}_0\left(e^{j\theta}\right)\right), \end{array} \tag{3.34}$$

donde $\mathbf{S}_1\left(e^{j\theta}\right) = \mathbf{h}S_{\mathrm{s}}\left(e^{j\theta}\right)\mathbf{h}^H + S_{\mathrm{v}}\left(e^{j\theta}\right)\mathbf{I}$ es la matriz de densidades espectrales de potencia bajo $\mathcal{H}_1$ y $\mathbf{S}_0\left(e^{j\theta}\right) = S_{\mathrm{v}}\left(e^{j\theta}\right)\mathbf{I}$ es dicha matriz bajo $\mathcal{H}_0$. $S_{\mathrm{s}}\left(e^{j\theta}\right)$ y $S_{\mathrm{v}}\left(e^{j\theta}\right)$ son, respectivamente, las transformadas de Fourier de $r_{\mathrm{s}}[m]$ y $r_{\mathrm{v}}[m]$.

El GLR para $\mathcal{H}_0 : \mathbf{S}_0\left(e^{j\theta}\right) = S_{\mathrm{v}}\left(e^{j\theta}\right)\mathbf{I}$, vs. $\mathcal{H}_1 : \mathbf{S}_1\left(e^{j\theta}\right) = \mathbf{h}S_{\mathrm{s}}\left(e^{j\theta}\right)\mathbf{h}^H + S_v\left(e^{j\theta}\right)\mathbf{I}$, viene dado por

$$\mathscr{L} = \frac{\underset{S_{\mathrm{v}}\left(e^{j\theta}\right)}{\text{máx}}\; p\left(\mathbf{z}_0, \ldots, \mathbf{z}_{M-1}; S_{\mathrm{v}}\left(e^{j\theta}\right)\right)}{\underset{\mathbf{h}, S_{\mathrm{s}}\left(e^{j\theta}\right), S_{\mathrm{v}}\left(e^{j\theta}\right)}{\text{máx}}\; p\left(\mathbf{z}_0, \ldots, \mathbf{z}_{M-1}; \mathbf{h}, S_{\mathrm{s}}\left(e^{j\theta}\right), S_{\mathrm{v}}\left(e^{j\theta}\right)\right)}. \tag{3.35}$$

La estima ML de la densidad espectral de potencia del ruido bajo $\mathcal{H}_0$ ha sido derivada anteriormente, obteniéndose

$$\hat{S}_{\mathrm{v}}\left(e^{j\theta}\right) = \frac{1}{L}\mathrm{tr}\left(\hat{\mathbf{S}}\left(e^{j\theta}\right)\right).$$

Para obtener las estimas ML bajo $\mathcal{H}_1$, se usa un procedimiento diferente al presentado en apartados anteriores. Comencemos por definir $\alpha\left(e^{j\theta}\right) = \mathbf{h}^H\hat{\mathbf{S}}\left(e^{j\theta}\right)\mathbf{h}$, que puede verse como una estima a la frecuencia θ de la energía total en el subespacio de señal. Aplicando ahora los lemas para la inversa de una matriz y para su determinante, dados por

$$\det\left(\mathbf{S}_1\left(e^{j\theta}\right)\right) = \left(1 + \frac{S_{\mathrm{s}}\left(e^{j\theta}\right)}{S_{\mathrm{v}}\left(e^{j\theta}\right)}\right) S_{\mathrm{v}}^L\left(e^{j\theta}\right),$$

$$\mathbf{S}_1^{-1}\left(e^{j\theta}\right) = \frac{1}{S_{\mathrm{v}}\left(e^{j\theta}\right)}\mathbf{I} - \frac{\mathbf{h}S_{\mathrm{s}}\left(e^{j\theta}\right)\mathbf{h}^H}{S_{\mathrm{v}}^2\left(e^{j\theta}\right) + S_{\mathrm{v}}\left(e^{j\theta}\right)S_{\mathrm{s}}\left(e^{j\theta}\right)},$$

[4] Sin pérdida de generalidad suponemos $\|\mathbf{h}\|^2 = 1$, debido a que cualquier factor de escala puede incorporarse a $s[n]$.

la log-verosimilitud bajo $\mathcal{H}_1$ es

$$\begin{aligned}
&\log p\left(\mathbf{z}_0,\ldots,\mathbf{z}_{M-1};\mathbf{h},S_{\mathrm{s}}\left(e^{j\theta}\right),S_{\mathrm{v}}\left(e^{j\theta}\right)\right)\\
&\quad=-LNM\log\pi-NM\int_{-\pi}^{\pi}\log\left(1+\frac{S_{\mathrm{s}}\left(e^{j\theta}\right)}{S_{\mathrm{v}}\left(e^{j\theta}\right)}\right)\frac{d\theta}{2\pi}-LNM\int_{-\pi}^{\pi}\log S_{\mathrm{v}}\left(e^{j\theta}\right)\frac{d\theta}{2\pi}\\
&\qquad-NM\int_{-\pi}^{\pi}\frac{1}{S_{\mathrm{v}}\left(e^{j\theta}\right)}\mathrm{tr}\left[\hat{\mathbf{S}}\left(e^{j\theta}\right)\right]\frac{d\theta}{2\pi}+NM\int_{-\pi}^{\pi}\frac{S_{\mathrm{s}}\left(e^{j\theta}\right)\alpha\left(e^{j\theta}\right)}{S_{\mathrm{v}}^2\left(e^{j\theta}\right)+S_{\mathrm{v}}\left(e^{j\theta}\right)S_{\mathrm{s}}\left(e^{j\theta}\right)}\frac{d\theta}{2\pi}.
\end{aligned}$$

Para obtener la estima ML de $S_{\mathrm{s}}\left(e^{j\theta}\right)$, suponiendo por el momento $S_{\mathrm{v}}\left(e^{j\theta}\right)$ conocido, debemos resolver el siguiente problema de optimización

$$\begin{aligned}
&\underset{S_{\mathrm{s}}\left(e^{j\theta}\right)}{\text{maximizar}} && \log p\left(\mathbf{z}_0,\ldots,\mathbf{z}_{M-1};\mathbf{h},S_{\mathrm{s}}\left(e^{j\theta}\right),S_{\mathrm{v}}\left(e^{j\theta}\right)\right),\\
&\text{sujeto a} && S_{\mathrm{s}}\left(e^{j\theta}\right)\geq 0,
\end{aligned}$$

cuya solución viene dada por

$$\hat{S}_{\mathrm{s}}\left(e^{j\theta}\right)=\left[\alpha\left(e^{j\theta}\right)-S_{\mathrm{v}}\left(e^{j\theta}\right)\right]^{+},$$

donde $[a]^{+}=\text{máx}\left(a,0\right)$.

Ahora se considerarán dos casos diferentes: (i) $\hat{S}_{\mathrm{s}}\left(e^{j\theta}\right)=0$ y (ii) $\hat{S}_{\mathrm{s}}\left(e^{j\theta}\right)>0$. El primero se reduce a $\mathcal{H}_0$, obteniendo la estima ML de $S_{\mathrm{v}}\left(e^{j\theta}\right)$ de apartados anteriores, dada por (3.30). En el segundo, la estima ML de $S_{\mathrm{v}}\left(e^{j\theta}\right)$ es

$$\hat{S}_{\mathrm{v}}\left(e^{j\theta}\right)=\frac{1}{L-1}\left(\mathrm{tr}\left[\hat{\mathbf{S}}\left(e^{j\theta}\right)\right]-\alpha\left(e^{j\theta}\right)\right),$$

que puede interpretarse como una estima de la energía normalizada (por dimensión) en el subespacio de ruido a la frecuencia θ.

Para obtener la log-verosimilitud comprimida, se define

$$\alpha_{\mathrm{tr}}(e^{j\theta})=\text{máx}\left(\alpha(e^{j\theta}),\frac{1}{L}\mathrm{tr}\left[\hat{\mathbf{S}}(e^{j\theta})\right]\right),$$

que es el máximo entre la energía en el subespacio de señal y la energía promedio por dimensión a la frecuencia θ. Sustituyendo las estimas ML de $S_{\mathrm{s}}\left(e^{j\theta}\right)$ y $S_{\mathrm{v}}\left(e^{j\theta}\right)$, la log-verosimilitud comprimida es

$$\begin{aligned}
\log p\left(\mathbf{z}_0,\ldots,\mathbf{z}_{M-1};\mathbf{h},\hat{S}_{\mathrm{s}}\left(e^{j\theta}\right),\hat{S}_{\mathrm{v}}\left(e^{j\theta}\right)\right)&=-NM\int_{-\pi}^{\pi}\log\alpha_{\mathrm{tr}}\left(e^{j\theta}\right)\frac{d\theta}{2\pi}\\
&\quad-(L-1)NM\int_{-\pi}^{\pi}\log\left[\frac{1}{L-1}\left(\mathrm{tr}\left[\hat{\mathbf{S}}\left(e^{j\theta}\right)\right]-\alpha_{\mathrm{tr}}\left(e^{j\theta}\right)\right)\right]\frac{d\theta}{2\pi}.
\end{aligned}\tag{3.36}$$

La estima ML de $\mathbf{h}$ se obtiene ahora como la solución al siguiente problema de optimización

$$\begin{aligned}
&\underset{\mathbf{h},\alpha_{\mathrm{tr}}\left(e^{j\theta}\right),\alpha\left(e^{j\theta}\right)}{\text{maximizar}} && \log p\left(\mathbf{z}_0,\ldots,\mathbf{z}_{M-1};\mathbf{h},\hat{S}_{\mathrm{s}}\left(e^{j\theta}\right),\hat{S}_{\mathrm{v}}\left(e^{j\theta}\right)\right)\\
&\text{sujeto a} && \alpha_{\mathrm{tr}}\left(e^{j\theta}\right)=\text{máx}\left(\alpha\left(e^{j\theta}\right),\tfrac{1}{L}\mathrm{tr}\left[\hat{\mathbf{S}}\left(e^{j\theta}\right)\right]\right),\\
& && \alpha\left(e^{j\theta}\right)=\mathbf{h}^H\hat{\mathbf{S}}\left(e^{j\theta}\right)\mathbf{h},\\
& && \mathbf{h}^H\mathbf{h}=1,
\end{aligned}\tag{3.37}$$

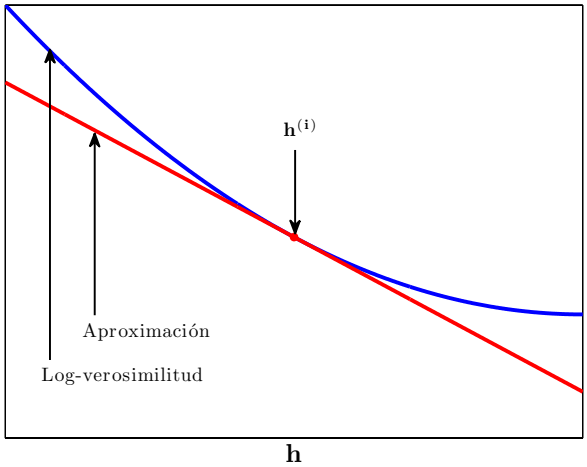

Figura 3.1: Ilustración gráfica de las condiciones para la convergencia de la técnica basada en las aproximaciones convexas sucesivas.

que es un problema difícil de resolver. Por un lado, las restricciones no son convexas, y por otro, la función (3.36) es convexa en $\alpha_{\text{tr}}\left(e^{j\theta}\right)$, y recordemos que maximizar una función convexa es un problema de optimización no-convexo [Boyd y Vandenberghe, 2004]. Resumiendo, el problema de optimización (3.37) es no-convexo, no existiendo soluciones eficientes, en general, para este tipo de problemas.

Para resolver (3.37) se propone usar una técnica denominada aproximaciones convexas sucesivas (o método de la condensación) [Marks y Wright, 1978, Chiang et al., 2007]. La idea básica de dicha técnica consiste en resolver una serie de problemas convexos, donde el problema original no-convexo es sustituido por una aproximación convexa. En [Marks y Wright, 1978] se demuestra que, cuando la aproximación cumple unas restricciones que enunciaremos a continuación, esta técnica converge a una solución que satisface las condiciones de Karush-Kuhn-Tucker (KKT), pudiendo ser, o no, un óptimo global. Las condiciones referenciadas anteriormente son:

1. $f\left(\mathbf{h}\right) \leq \log p\left(\mathbf{z}_0, \ldots, \mathbf{z}_{M-1}; \mathbf{h}, \hat{S}_{\text{s}}\left(e^{j\theta}\right), \hat{S}_{\text{v}}\left(e^{j\theta}\right)\right)$, $\forall \mathbf{h}$, donde $f\left(\mathbf{h}\right)$ es la aproximación.

2. $f\left(\mathbf{h}^{(i)}\right) = \log p\left(\mathbf{z}_0, \ldots, \mathbf{z}_{M-1}; \mathbf{h}^{(i)}, \hat{S}_{\text{s}}\left(e^{j\theta}\right), \hat{S}_{\text{v}}\left(e^{j\theta}\right)\right)$, donde $\mathbf{h}^{(i)}$ es la solución óptima del problema aproximado en la iteración anterior.

3. $\left.\frac{\partial f(\mathbf{h})}{\partial \mathbf{h}}\right|_{\mathbf{h}=\mathbf{h}^{(i)}} = \left.\frac{\partial \log p\left(\mathbf{z}_0,\ldots,\mathbf{z}_{M-1};\mathbf{h},\hat{S}_{\text{s}}\left(e^{j\theta}\right),\hat{S}_{\text{v}}\left(e^{j\theta}\right)\right)}{\partial \mathbf{h}}\right|_{\mathbf{h}=\mathbf{h}^{(i)}}$.

La Figura 3.1 muestra, de manera gráfica, las condiciones anteriores para un ejemplo muy sencillo. En primer lugar, la aproximación debe ser siempre inferior a la función original para cualquier valor de $\mathbf{h}$. Esta condición implica que el valor de la función objetivo evaluada en la solución del problema aproximado es siempre mayor que el valor de la función objetivo del problema aproximado. Además, la función original y su aproximación deben ser iguales evaluadas en la solución de la iteración previa. Estas dos condiciones garantizan el aumento, en cada iteración, del valor de la función objetivo. Por último, la función original y su aproximación tienen la misma derivada en la solución de la iteración anterior, garantizándose que

la solución del problema aproximado, cuando ha convergido, satisface las condiciones KKT del problema original.

Comencemos por encontrar una aproximación de la función objetivo que satisfaga las condiciones presentadas anteriormente. De entre todas las posibles aproximaciones se ha elegido el polinomio de Taylor de orden 1 en torno a $\mathbf{h}^{(i)}$, debido a su sencillez, y siendo fácil demostrar que cumple las condiciones anteriores. Es decir, la aproximación elegida viene dada por

$$f(\mathbf{h}) = a + \left[\int_{-\pi}^{\pi} \mathbf{b}^H\left(e^{j\theta}\right)\frac{d\theta}{2\pi}\right]\left(\mathbf{h}-\mathbf{h}^{(i)}\right),$$

donde $a = \log p\left(\mathbf{z}_0,\ldots,\mathbf{z}_{M-1};\mathbf{h}^{(i)},\hat{S}_{\mathrm{s}}\left(e^{j\theta}\right),\hat{S}_{\mathrm{v}}\left(e^{j\theta}\right)\right)$,

$$\mathbf{b}^H\left(e^{j\theta}\right) = \left.\frac{\partial \log p\left(\mathbf{z}_0,\ldots,\mathbf{z}_{M-1};\mathbf{h},\hat{S}_{\mathrm{s}}\left(e^{j\theta}\right),\hat{S}_{\mathrm{v}}\left(e^{j\theta}\right)\right)}{\partial \mathbf{h}}\right|_{\mathbf{h}=\mathbf{h}^{(i)}} = b\left(e^{j\theta}\right)\mathbf{h}^{(i)H}\hat{\mathbf{S}}\left(e^{j\theta}\right),$$

con $b\left(e^{j\theta}\right)$ dado por

$$b\left(e^{j\theta}\right) = \begin{cases} \dfrac{L-1}{\mathrm{tr}\left[\hat{\mathbf{S}}\left(e^{j\theta}\right)\right]-\alpha^{(i)}\left(e^{j\theta}\right)} - \dfrac{1}{\alpha^{(i)}\left(e^{j\theta}\right)}, & \text{if } \theta\in\Theta_+^{(i)}, \\ 0, & \text{if } \theta\in\Theta_-^{(i)}, \end{cases} \tag{3.38}$$

$\alpha^{(i)}\left(e^{j\theta}\right) = \mathbf{h}^{(i)H}\hat{\mathbf{S}}\left(e^{j\theta}\right)\mathbf{h}^{(i)}$ es la estima de la energía en el subespacio de señal en la i-ésima iteración, y $\Theta_+^{(i)}$ y $\Theta_-^{(i)}$ son los siguientes conjuntos de frecuencias

$$\Theta_+^{(i)} = \left\{\theta \mid \alpha^{(i)}\left(e^{j\theta}\right) \geq \frac{1}{L}\mathrm{tr}\left[\hat{\mathbf{S}}\left(e^{j\theta}\right)\right]\right\},$$
$$\Theta_-^{(i)} = \left\{\theta \mid \alpha^{(i)}\left(e^{j\theta}\right) < \frac{1}{L}\mathrm{tr}\left[\hat{\mathbf{S}}\left(e^{j\theta}\right)\right]\right\}.$$

Antes de continuar, se va a interpretar el valor de $b\left(e^{j\theta}\right)$. Para ello, es posible reescribir $b\left(e^{j\theta}\right)$ como

$$b\left(e^{j\theta}\right) = \left[\frac{\alpha^{(i)}\left(e^{j\theta}\right)}{\mathrm{SNR}^{(i)}\left(e^{j\theta}\right)}\right]^{-1},$$

donde

$$\mathrm{SNR}^{(i)}\left(e^{j\theta}\right) = \frac{\hat{S}_s^{(i)}\left(e^{j\theta}\right)}{\hat{S}_v^{(i)}\left(e^{j\theta}\right)},$$

es la SNR estimada en la i-ésima iteración a la frecuencia θ y $\hat{S}_s^{(i)}\left(e^{j\theta}\right)$ y $\hat{S}_v^{(i)}\left(e^{j\theta}\right)$ son las estimas en la i-ésima iteración de las densidades espectrales de la señal y el ruido, respectivamente. Por lo tanto, $b\left(e^{j\theta}\right)$ es inversamente proporcional a la energía total en el subespacio de señal, normalizado por la SNR, a la frecuencia correspondiente.

Teniendo en cuenta el desarrollo en serie de Taylor, el problema de optimización en cada iteración es

$$\begin{array}{ll} \underset{\mathbf{h}}{\text{maximizar}} & \mathbf{h}^{(i)H}\hat{\mathbf{R}}_{\mathrm{w}}\mathbf{h}, \\ \text{sujeto a} & \mathbf{h}^H\mathbf{h} = 1, \end{array} \tag{3.39}$$

3.1: Estima de $\mathbf{h}$ basada en aproximaciones convexas sucesivas.

Entradas: Punto inicial $\mathbf{h}^{(0)}$ y $\hat{\mathbf{S}}\left(e^{j\theta}\right)$
Salidas: Estima de $\mathbf{h}$ bajo $\mathcal{H}_1$
Inicializar: $i = 0$
repeat
  Calcular $b\left(e^{j\theta}\right)$ usando (3.38)
  Obtener $\hat{\mathbf{R}}_{\mathrm{w}} = \int_{-\pi}^{\pi} b\left(e^{j\theta}\right) \hat{\mathbf{S}}\left(e^{j\theta}\right) \frac{d\theta}{2\pi}$
  Estimar $\mathbf{h}^{(i)}$ a partir de (3.41)
  Actualizar $i = i + 1$
until Convergencia

donde

$$\hat{\mathbf{R}}_{\mathrm{w}} = \int_{-\pi}^{\pi} b\left(e^{j\theta}\right) \hat{\mathbf{S}}\left(e^{j\theta}\right) \frac{d\theta}{2\pi},$$

es una matriz de covarianza ponderada. Es decir, es la transformada de Fourier inversa (para $m = 0$) de la matriz de densidades espectrales de potencia estimada y ponderada a cada frecuencia por $b\left(e^{j\theta}\right)$. Analicemos dos casos límite: $\mathrm{SNR}^{(i)}\left(e^{j\theta}\right) = 0$ y $\mathrm{SNR}^{(i)}\left(e^{j\theta}\right) \gg 1$. En el primero, la matriz $\hat{\mathbf{R}}_{\mathrm{w}}$ viene dada por

$$\hat{\mathbf{R}}_{\mathrm{w}} = \int_{\Omega} b\left(e^{j\theta}\right) \hat{\mathbf{S}}\left(e^{j\theta}\right) \frac{d\theta}{2\pi}, \tag{3.40}$$

donde Ω es el siguiente conjunto $\left\{\theta \mid \mathrm{SNR}^{(i)}\left(e^{j\theta}\right) \neq 0\right\}$, por ello, las frecuencias con mucho ruido no contribuyen a la matriz $\hat{\mathbf{R}}_{\mathrm{w}}$. En el segundo, la matriz viene dada por (3.40), donde ahora Ω es el conjunto $\left\{\theta \mid \mathrm{SNR}^{(i)}\left(e^{j\theta}\right) \gg 1\right\}$, es decir, la matriz $\hat{\mathbf{R}}_{\mathrm{w}}$ se forma únicamente con las frecuencias sin ruido.

La solución del problema de optimización en (3.39) viene dada por

$$\mathbf{h}^{(i+1)} = \begin{cases} \dfrac{\hat{\mathbf{R}}_{\mathrm{w}}\mathbf{h}^{(i)}}{\left\|\hat{\mathbf{R}}_{\mathrm{w}}\mathbf{h}^{(i)}\right\|}, & \text{si } \|\hat{\mathbf{R}}_{\mathrm{w}}\mathbf{h}^{(i)}\| \neq 0, \\ \text{cualquier vector de norma unidad}, & \text{si } \|\hat{\mathbf{R}}_{\mathrm{w}}\mathbf{h}^{(i)}\| = 0. \end{cases} \tag{3.41}$$

Esta solución puede interpretarse como una iteración de un método de potencias [Golub y van Loan, 1996] para la extracción del autovector principal de $\hat{\mathbf{R}}_{\mathrm{w}}$. Dada la naturaleza iterativa del método propuesto, esta estima se vuelve a sustituir en la función de coste, y el procedimiento se repite hasta que converja. Básicamente, la solución propuesta para la estima del canal SIMO, presentada en el Algoritmo 3.1, es un método de potencias donde la matriz varía entre iteraciones. Finalmente, el log-GLRT se obtiene sustituyendo las estimas ML obtenidas en (3.35).

Una vez más, el GLRT presenta la siguiente invarianza.

Lema 3.10. *El GLRT iterativo derivado en esta sección es invariante a transformaciones* $\mathbf{y}[n] = (a * \mathbf{x})[n]$, *con* $\left|A\left(e^{j\theta}\right)\right| \neq 0$, *siendo* $A\left(e^{j\theta}\right) = \mathcal{F}\left(a[n]\right)$.

Demostración. La demostración es consecuencia directa de

$$\hat{\mathbf{S}}^{(y)}\left(e^{j\theta}\right) = \left|A\left(e^{j\theta}\right)\right|^2 \hat{\mathbf{S}}^{(x)}\left(e^{j\theta}\right),$$

donde $\hat{\mathbf{S}}^{(x)}\left(e^{j\theta}\right)$ y $\hat{\mathbf{S}}^{(x)}\left(e^{j\theta}\right)$ son las PSD muestrales de $\mathbf{x}[n]$ e $\mathbf{y}[n]$, respectivamente. □

3.3. Detección Multicanal de Procesos No Estacionarios

Esta sección considera la detección multicanal de procesos no estacionarios, sin estructura espacial y contaminados por ruidos IID (modelo INNSNS), siendo las matrices de covarianza para este modelo

$$\begin{aligned}\mathcal{H}_1 : E\left[\mathbf{x}[n]\mathbf{x}^H[n-m]\right] &= \mathbf{R}_1[n, n-m],\\ \mathcal{H}_0 : E\left[\mathbf{x}[n]\mathbf{x}^H[n-m]\right] &= \sigma^2[n, n-m]\mathbf{I}.\end{aligned}$$

Recordemos que la matriz de datos viene dada por

$$\mathbf{X} = \begin{bmatrix} x_1[0] & x_1[1] & \dots & x_1[N-1] \\ x_2[0] & x_2[1] & \dots & x_2[N-1] \\ \vdots & \vdots & \ddots & \vdots \\ x_L[0] & x_L[1] & \dots & x_L[N-1] \end{bmatrix} = \begin{bmatrix} \mathbf{x}_1^T \\ \mathbf{x}_2^T \\ \vdots \\ \mathbf{x}_L^T \end{bmatrix},$$

y, en esta ocasión, se define el vector $\mathbf{g} = \text{vec}\left(\mathbf{X}^T\right)$, que apila las columnas de $\mathbf{X}^T$, siendo su matriz de covarianza

$$\mathbf{R} = E\left[\mathbf{g}\mathbf{g}^H\right] = \begin{bmatrix} \mathbf{R}_{11} & \mathbf{R}_{21}^H & \dots & \mathbf{R}_{L1}^H \\ \mathbf{R}_{21} & \mathbf{R}_{22} & \dots & \mathbf{R}_{L2}^H \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{R}_{L1} & \mathbf{R}_{L2} & \dots & \mathbf{R}_{LL} \end{bmatrix} \in \mathbb{C}^{LN \times LN},$$

donde las matrices de covarianza cruzada $\mathbf{R}_{ik} = E\left[\mathbf{x}_i\mathbf{x}_k^H\right]$, $1 \leq i, k \leq L$, contienen toda la información espacio-temporal de segundo orden de los vectores $\{\mathbf{x}_i\}_{i=1}^L$. En esta sección, se consideran series temporales no estacionarias, obteniendo matrices de covarianza sin estructura Toeplitz. Además, tampoco se impone ninguna estructura temporal adicional, es decir, las matrices $\mathbf{R}_{ik}$ son sólo definidas positivas, $\mathbf{R}_{ik} \in \mathbb{S}_+$.

Teniendo en cuenta estas definiciones, se puede formular el problema de detección como

$$\begin{aligned}\mathcal{H}_1 &: \mathbf{g} \sim \mathcal{CN}\left(\mathbf{0}_{LN}, \mathbf{R}_1\right),\\ \mathcal{H}_0 &: \mathbf{g} \sim \mathcal{CN}\left(\mathbf{0}_{LN}, \mathbf{R}_0\right),\end{aligned}$$

donde $\mathbf{R}_1 \in \mathbb{S}_+$, $\mathbf{R}_0 \in \mathbb{D}_\text{B}$ son dos matrices de covarianza desconocidas. Además, debido a la suposición de ruidos IID, la matriz $\mathbf{R}_0$ posee la siguiente estructura

$$\mathbf{R}_0 = \begin{bmatrix} \mathbf{R}_{11} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_{11} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{R}_{11} \end{bmatrix} = \mathbf{I} \otimes \mathbf{R}_{11}.$$

Considerando M ($M \geq LN$) observaciones independientes e idénticamente distribuidas, $\mathbf{g}_1, \dots, \mathbf{g}_M$, de $\mathbf{g}$, la verosimilitud es

$$p\left(\mathbf{g}_0, \dots, \mathbf{g}_{M-1}; \mathbf{R}\right) = \frac{1}{\pi^{LNM}\det\left(\mathbf{R}\right)^M} \exp\left\{-M\text{tr}\left(\mathbf{R}^{-1}\hat{\mathbf{R}}\right)\right\},$$

donde, en esta ocasión, la matriz de covarianza muestral se define como

$$\hat{\mathbf{R}} = \frac{1}{M}\sum_{m=0}^{M-1} \mathbf{g}_m\mathbf{g}_m^H = \begin{bmatrix} \hat{\mathbf{R}}_{11} & \hat{\mathbf{R}}_{21}^H & \dots & \hat{\mathbf{R}}_{L1}^H \\ \hat{\mathbf{R}}_{21} & \hat{\mathbf{R}}_{22} & \dots & \hat{\mathbf{R}}_{L2}^H \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\mathbf{R}}_{L1} & \hat{\mathbf{R}}_{L2} & \dots & \hat{\mathbf{R}}_{LL} \end{bmatrix},$$

con $\hat{\mathbf{R}}_{ik} \in \mathbb{C}^{N\times N}$ el $\{i,k\}$-ésimo bloque de $\hat{\mathbf{R}}$, que es la covarianza muestral cruzada entre las ventanas de longitud N de la i-ésima y k-ésima series temporales. Con lo cual, el GLRT para $\mathcal{H}_0 : \mathbf{R} = \mathbf{I} \otimes \mathbf{R}_{11}, \mathbf{R}_{11} \in \mathbb{S}_+$ vs. $\mathcal{H}_1 : \mathbf{R} \in \mathbb{S}_+$ se basa en el siguiente GLR

$$\mathscr{L} = \frac{\operatorname*{máx}_{\mathbf{R}=\mathbf{I}\otimes\mathbf{R}_{11},\mathbf{R}_{11}\in\mathbb{S}_+} p\left(\mathbf{g}_0,\ldots,\mathbf{g}_{M-1};\mathbf{R}\right)}{\operatorname*{máx}_{\mathbf{R}\in\mathbb{S}_+} p\left(\mathbf{g}_0,\ldots,\mathbf{g}_{M-1};\mathbf{R}\right)},$$

siendo necesario derivar las estimas ML de $\mathbf{R}_0$ y $\mathbf{R}_1$, que se presentan en los siguientes lemas.

Lema 3.11. *La estima de la matriz de covarianza bajo $\mathcal{H}_0$ viene dada por*

$$\hat{\mathbf{R}}_0 = \mathbf{I} \otimes \left[\frac{1}{L}\sum_{i=1}^{L}\hat{\mathbf{R}}_{ii}\right]. \tag{3.42}$$

Demostración. Bajo $\mathcal{H}_0$, es posible reescribir la verosimilitud como

$$p\left(\mathbf{g}_0,\ldots,\mathbf{g}_{M-1};\mathbf{R}\right) = p\left(\mathbf{g}_0,\ldots,\mathbf{g}_{M-1};\mathbf{R}_{11}\right),$$

que es la verosimilitud de una variable Gaussiana con la siguiente matriz de covarianza muestral equivalente

$$\hat{\mathbf{R}}_{\text{eq}} = \frac{1}{L}\sum_{i=1}^{L}\hat{\mathbf{R}}_{ii}.$$

Por lo tanto, teniendo en cuenta que $\mathbf{R}_{11}$ no tiene ninguna estructura adicional (ni espacial ni temporal), su estima ML será la matriz de covarianza muestral equivalente [Magnus y Neudecker, 1999]. La demostración concluye sustituyendo la estima ML en la matriz $\mathbf{R}_0$. □

Lema 3.12. *La estima de la matriz de covarianza bajo $\mathcal{H}_1$ es*

$$\hat{\mathbf{R}}_1 = \hat{\mathbf{R}}. \tag{3.43}$$

Demostración. La demostración se puede encontrar en [Magnus y Neudecker, 1999]. □

Teniendo en cuenta (3.42) y (3.43), el GLRT es

$$\mathscr{L} = \frac{\det\left(\hat{\mathbf{R}}\right)}{\det\left(\frac{1}{L}\sum_{i=1}^{L}\hat{\mathbf{R}}_{ii}\right)^{L}}. \tag{3.44}$$

Este GLRT se puede interpretar como el test de esfericidad para señales no estacionarias, y podemos ver sus similitudes con los GLRT dados por (3.12) y (3.31). Por último, se presenta la invarianza del GLRT (3.44) a transformaciones lineales.

Lema 3.13. *El GLRT es invariante a transformaciones $\mathbf{g}^{(y)} = (\mathbf{Q}\otimes\mathbf{A})\,\mathbf{g}^{(x)}$, siendo $\mathbf{g}^{(x)}$ y $\mathbf{g}^{(y)}$ las vectorizaciones de las matrices de datos original y transformada, respectivamente, con $\mathbf{Q}^H\mathbf{Q} = \mathbf{I}$ y $\mathbf{A}$ una matriz invertible.*

Demostración. Es fácil demostrar que las matrices de covarianza muestrales de las señales $\mathbf{y}_i$ vienen dadas por $\hat{\mathbf{R}}_0^{(y)} = (\mathbf{Q}\otimes\mathbf{A})\,\hat{\mathbf{R}}_0^{(x)}\,(\mathbf{Q}\otimes\mathbf{A})^H$ y $\hat{\mathbf{R}}_1^{(y)} = (\mathbf{Q}\otimes\mathbf{A})\,\hat{\mathbf{R}}_1^{(x)}\,(\mathbf{Q}\otimes\mathbf{A})^H$. Sustituyendo estas matrices en (3.44) se obtiene el resultado deseado. □

3.3.1. GLRT en el Dominio de la Frecuencia

En esta subsección, aplicando la propiedad de invarianza dada por el Lema 3.13, se reescribe el GLRT en (3.44), presentando adicionalmente una aproximación en el dominio frecuencial.

Comencemos considerando las señales $\mathbf{g}^{(y)} = (\mathbf{I} \otimes \mathbf{F}_N)\,\mathbf{g}^{(x)}$, donde $\mathbf{F}_N$ es la matriz de Fourier de dimensiones $N \times N$. Teniendo en cuenta el Lema 3.13, el GLRT de las señales originales y transformadas es

$$\mathscr{L} = \frac{\det\left((\mathbf{I} \otimes \mathbf{F}_N)\,\hat{\mathbf{R}}\,(\mathbf{I} \otimes \mathbf{F}_N^H)\right)}{\det\left(\mathbf{F}_N \hat{\mathbf{R}}_{\text{eq}} \mathbf{F}_N^H\right)^L}. \tag{3.45}$$

Introduciendo una simple permutación de filas y columnas de las matrices dentro de los determinantes en (3.45), el GLRT puede reescribirse como

$$\mathscr{L} = \frac{\det\left(\tilde{\mathbf{R}}\right)}{\det\left(\tilde{\mathbf{R}}_{\text{eq}}\right)^L},$$

donde

$$\tilde{\mathbf{R}} = \begin{bmatrix} \hat{\mathbf{S}}\left(e^{j\theta_0}\right) & \cdots & \hat{\mathbf{S}}\left(e^{j\theta_0}, e^{j\theta_{N-1}}\right) \\ \vdots & \ddots & \vdots \\ \hat{\mathbf{S}}\left(e^{j\theta_0}, e^{j\theta_{N-1}}\right) & \cdots & \hat{\mathbf{S}}\left(e^{j\theta_{N-1}}\right) \end{bmatrix},$$

es la matriz de densidades espectrales de potencia global en el dominio de la frecuencia,

$$\tilde{\mathbf{R}}_{\text{eq}} = \begin{bmatrix} \hat{S}_{\text{eq}}\left(e^{j\theta_0}\right) & \cdots & \hat{S}_{\text{eq}}\left(e^{j\theta_0}, e^{j\theta_{N-1}}\right) \\ \vdots & \ddots & \vdots \\ \hat{S}_{\text{eq}}\left(e^{j\theta_0}, e^{j\theta_{N-1}}\right) & \cdots & \hat{S}_{\text{eq}}\left(e^{j\theta_{N-1}}\right) \end{bmatrix},$$

$\theta_i = 2\pi i/N$ y, por simplicidad notacional, se usa $\hat{\mathbf{S}}\left(e^{j\theta_i}\right)$ en lugar de $\hat{\mathbf{S}}\left(e^{j\theta_i}, e^{j\theta_i}\right)$. Los elementos de las matrices de densidades espectrales de potencia son

$$\left[\hat{\mathbf{S}}\left(e^{j\theta_i}, e^{j\theta_k}\right)\right]_{l,m} = \mathbf{f}^H\left(e^{j\theta_i}\right)\hat{\mathbf{R}}_{lm}\mathbf{f}\left(e^{j\theta_k}\right),$$

$$\hat{S}_{\text{eq}}\left(e^{j\theta_i}, e^{j\theta_k}\right) = \mathbf{f}^H\left(e^{j\theta_i}\right)\left(\frac{1}{L}\sum_{i=1}^{L}\hat{\mathbf{R}}_{ii}\right)\mathbf{f}\left(e^{j\theta_k}\right) = \frac{1}{L}\sum_{i=1}^{L}\hat{S}_{ii}\left(e^{j\theta_i}, e^{j\theta_k}\right),$$

y $\mathbf{f}\left(e^{j\theta}\right)$ denota el vector de Fourier a la frecuencia angular θ.

A continuación, factorizamos $\tilde{\mathbf{R}}$ y $\tilde{\mathbf{R}}_{\text{eq}}$ de la siguiente manera

$$\tilde{\mathbf{R}} = \tilde{\mathbf{R}}_{\text{WSS}}^{1/2}\tilde{\mathbf{R}}_{\text{NS}}\tilde{\mathbf{R}}_{\text{WSS}}^{1/2},$$

y

$$\tilde{\mathbf{R}}_{\text{eq}} = \tilde{\mathbf{R}}_{\text{eq, WSS}}^{1/2}\tilde{\mathbf{R}}_{\text{eq, NS}}\tilde{\mathbf{R}}_{\text{eq, WSS}}^{1/2},$$

donde

$$\tilde{\mathbf{R}}_{\text{WSS}} = \begin{bmatrix} \hat{\mathbf{S}}\left(e^{j\theta_0}\right) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \hat{\mathbf{S}}\left(e^{j\theta_1}\right) & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \hat{\mathbf{S}}\left(e^{j\theta_{N-1}}\right) \end{bmatrix},$$

$$\tilde{\mathbf{R}}_{\text{eq, WSS}} = \begin{bmatrix} \hat{S}_{\text{eq}}\left(e^{j\theta_0}\right) & 0 & \dots & 0 \\ 0 & \hat{S}_{\text{eq}}\left(e^{j\theta_1}\right) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \hat{S}_{\text{eq}}\left(e^{j\theta_{N-1}}\right) \end{bmatrix},$$

$\tilde{\mathbf{R}}_{\text{NS}} = \tilde{\mathbf{R}}_{\text{WSS}}^{-1/2}\tilde{\mathbf{R}}\tilde{\mathbf{R}}_{\text{WSS}}^{-1/2}$ y $\tilde{\mathbf{R}}_{\text{eq, NS}} = \tilde{\mathbf{R}}_{\text{eq, WSS}}^{-1/2}\tilde{\mathbf{R}}_{\text{eq}}\tilde{\mathbf{R}}_{\text{eq, WSS}}^{-1/2}$. Así, el logaritmo del GLRT se puede reescribir como

$$\log \mathscr{L} = \log \mathscr{L}_{\text{WSS}} + \log \mathscr{L}_{\text{NS}},$$

con

$$\mathscr{L}_{\text{NS}} = \frac{\det(\tilde{\mathbf{R}}_{\text{NS}})}{\det(\tilde{\mathbf{R}}_{\text{eq, NS}})^L},$$

y

$$\mathscr{L}_{\text{WSS}} = \frac{\det(\tilde{\mathbf{R}}_{\text{WSS}})}{\det(\tilde{\mathbf{R}}_{\text{eq, WSS}})^L}.$$

Consideremos ahora series temporales estacionarias en sentido amplio. En este caso las matrices de covarianza $\mathbf{R}_{lm}$ son Toeplitz y, por lo tanto, los elementos de la matriz $\hat{\mathbf{S}}\left(e^{j\theta}\right)$ pueden verse como estimas de las densidades espectrales potencia. Si el tamaño de la ventana tiende a infinito, $N \to \infty$, se puede aplicar el teorema de Szëgo [Grenander y Szegö, 1958, Gray, 2006] para obtener

$$\frac{1}{N}\log \mathscr{L} = \frac{1}{N}\log \mathscr{L}_{\text{WSS}} \underset{N\to\infty}{\longrightarrow} \int_{-\pi}^{\pi} \log\left\{\frac{\det\left[\hat{\mathbf{S}}\left(e^{j\theta}\right)\right]^{1/L}}{\frac{1}{L}\text{tr}\left[\hat{\mathbf{S}}\left(e^{j\theta}\right)\right]}\right\}, \tag{3.46}$$

donde $[\hat{\mathbf{S}}\left(e^{j\theta}\right)]_{l,m} = \mathbf{f}^H\left(e^{j\theta}\right)\hat{\mathbf{R}}_{lm}\mathbf{f}\left(e^{j\theta}\right)$ es un estimador cuadrático (promediado sobre M realizaciones) de la matriz de densidades espectrales de potencia [Thomson, 1982, Mullis y Scharf, 1991]. Por lo tanto, la aproximación del GLRT (3.46) converge al GLRT (3.31) para señales estacionarias. Adicionalmente, cabe destacar que el término $\log \mathscr{L}_{\text{NS}}$ converge a 0, por lo tanto, es posible interpretarlo como una medida de la contribución al estadístico de la parte no estacionaria de la serie temporal multivariada.

3.4. Resultados Numéricos

En esta sección, las prestaciones de los detectores propuestos se evalúan mediante simulaciones de Monte Carlo. Concretamente, nos centraremos en los siguientes detectores:

- Sphericity: GLRT para el modelo INWSNS, conocido como test de esfericidad, cuyo estadístico es

$$\log \mathscr{L} = \log\left\{\frac{\left[\prod_{i=1}^{L}\lambda_i\right]^{1/L}}{\frac{1}{L}\sum_{i=1}^{L}\lambda_i}\right\},$$

  donde λ_i son los autovalores de la matriz de covarianza muestral.

- IID-GLRT: GLRT para el modelo INWSPS, siendo su estadístico

$$\mathscr{L} = L \log \left[\frac{\left(\prod_{i=1}^{L} \lambda_i \right)^{1/L}}{\frac{1}{L} \sum_{i=1}^{L} \lambda_i} \right] - (L-P) \log \left[\frac{\left(\prod_{i=P+1}^{L} \lambda_i \right)^{1/(L-P)}}{\frac{1}{L-P} \sum_{i=P+1}^{L} \lambda_i} \right].$$

- IID-GLRT ($P = 1$): GLRT para el modelo INWSPS con $P = 1$, propuesto en [Besson et al., 2006], y basado en el siguiente estadístico

$$\log \mathscr{L} = \log \left[\frac{\operatorname{tr}\left(\hat{\mathbf{R}}\right)}{\lambda_1} \right].$$

- Freq. Sphericity: GLRT para el modelo INSSNS, cuyo estadístico es

$$\log \mathscr{L} = \int_{-\pi}^{\pi} \log \left\{ \frac{\left[\prod_{i=1}^{L} \lambda_i \left(e^{j\theta}\right) \right]^{1/L}}{\frac{1}{L} \sum_{i=1}^{L} \lambda_i \left(e^{j\theta}\right)} \right\} \frac{d\theta}{2\pi},$$

donde $\lambda_i \left(e^{j\theta}\right)$ son los autovalores de la matriz PSD muestral.

- Freq. IID-GLRT: GLRT para el modelo INSSPS, basado en el siguiente GLR

$$\log \mathscr{L} = L \int_{-\pi}^{\pi} \log \left[\frac{\left(\prod_{i=1}^{L} \lambda_i \left(e^{j\theta}\right) \right)^{1/L}}{\frac{1}{L} \sum_{i=1}^{L} \lambda_i \left(e^{j\theta}\right)} \right] \frac{d\theta}{2\pi}$$
$$- (L-P) \int_{-\pi}^{\pi} \log \left[\frac{\left(\prod_{i=P+1}^{L} \lambda_i \left(e^{j\theta}\right) \right)^{1/(L-P)}}{\frac{1}{L-P} \sum_{i=P+1}^{L} \lambda_i \left(e^{j\theta}\right)} \right] \frac{d\theta}{2\pi}.$$

- Freq. GLRT iterativo: GLRT para el modelo INWSPS basado en el Algoritmo 3.1.

- Sphericity Gen.: GLRT para el modelo INNSNS, que puede interpretarse como una generalización del test de esfericidad, siendo su estadístico

$$\mathscr{L} = \frac{\det\left(\hat{\mathbf{R}}\right)}{\det\left(\frac{1}{L} \sum_{i=1}^{L} \hat{\mathbf{R}}_{ii} \right)^{L}}.$$

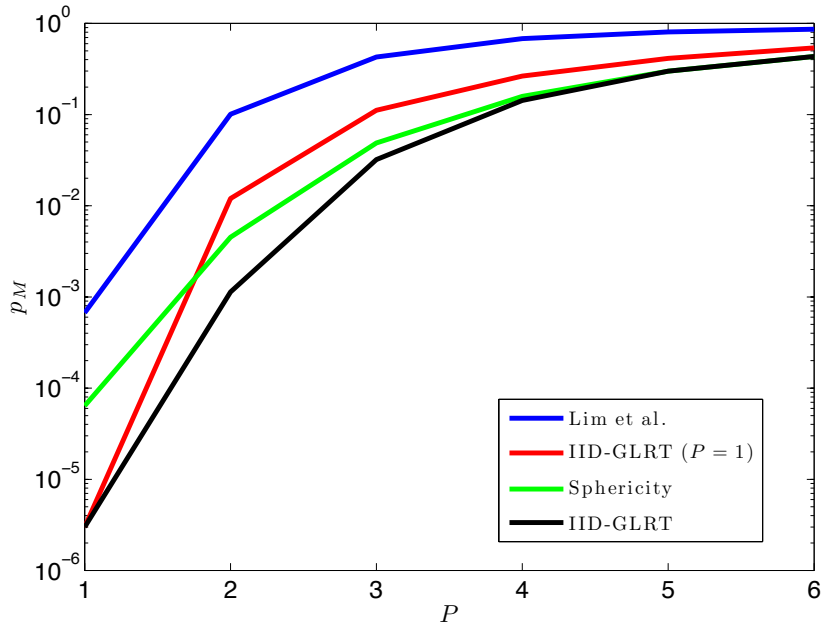

Figura 3.2: Probabilidad de pérdida frente a P para ruidos IID en el siguiente experimento: $L = 6$ sensores, SNR (dB) $= -2$ dB y $M = 64$ muestras.

- Lim et al.: Detector heurístico propuesto en [Lim et al., 2008], dado por

$$T_{\text{SSE}} = \frac{1}{P}\sum_{i=1}^{P}\frac{\lambda_i}{\lambda_L} - \log\left(\prod_{i=1}^{P}\frac{\lambda_i}{\lambda_L}\right)^{1/P}.$$

3.4.1. Efecto del Rango P

Este primer ejemplo analiza el comportamiento de los diferentes detectores para señales temporalmente blancas en función del rango P. Para ello, en la Figura 3.2 se muestra la probabilidad de pérdida (*missed detection probability*) para una probabilidad de falsa alarma fija, $p_{\text{FA}} = 0.001$, y distintos valores de P. Concretamente, se han considerado señales Gaussianas y canales Rayleigh normalizados para obtener una SNR instantánea, definida como

$$\text{SNR (dB)} = 10\log_{10}\frac{\text{tr}(\mathbf{H}\mathbf{H}^H)}{L\sigma^2} = -2 \text{ dB},$$

constante en todas las realizaciones de Monte Carlo. El número de sensores es $L = 6$ y el número de observaciones $M = 64$. Como se puede ver en la figura, las mejores prestaciones las obtiene el GLRT propuesto, dado por (3.21). El detector del test de esfericidad presenta peores prestaciones que el GLRT (3.21), salvo para P grandes, debido a que el GLRT es el único que explota completamente la estructura espacial del problema. De hecho, como se demostró en el Lema 3.5 y se puede ver en la Fig. 3.2, para $P = 5$ y $P = 6$, ambos detectores son equivalentes. Por otro lado, observando el GLRT que asume siempre $P = 1$, se ve la ganancia obtenida al explotar el rango P. Obviamente, para $P = 1$, las prestaciones son iguales, pero cuando aumenta P, las diferencias se incrementan. El detector de [Lim et al., 2008] presenta las peores prestaciones debido, principalmente, a la pobre estima del ruido, en particular

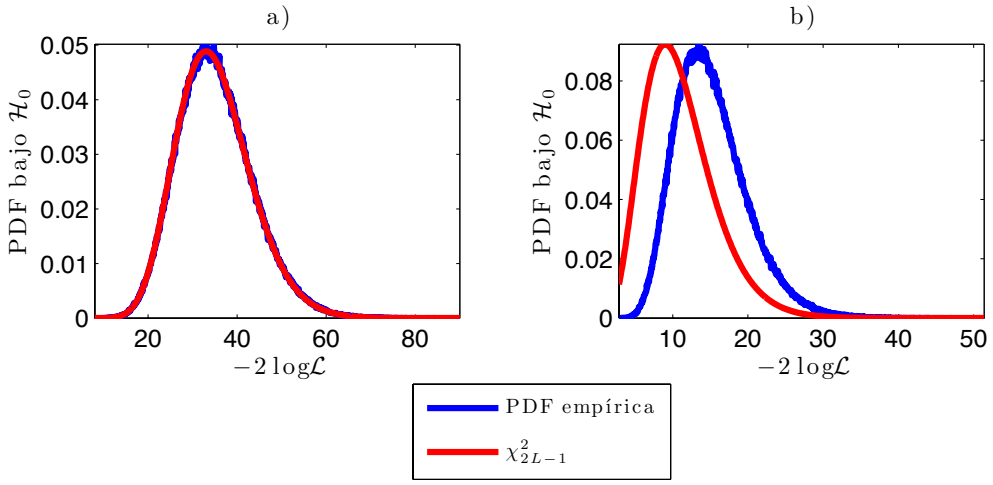

Figura 3.3: PDF estimada y aproximación χ^2 para: a) el test de esfericidad y b) el GLRT para señales blancas con rango P y ruidos IID.

bajo $\mathcal{H}_0$. Por otro lado, es importante destacar, que las prestaciones de todos los detectores empeoran cuando P crece. Una explicación posible es que al aumentar P, las matrices de covarianza pierden estructura que se emplea para mejorar su estima y, por consiguiente, empeoran las prestaciones de los detectores.

Para este mismo experimento, seleccionando $M = 1000$, la Figura 3.3 muestra el histograma bajo $\mathcal{H}_0$, para $P = 1$ y $P = 6$. En el caso de $P = 6$ (ver Figura 3.3a), el GLRT es el test de esfericidad, y la aproximación $\chi^2_{L^2-1}$ ajusta correctamente la PDF estimada. Por otro lado, para $P = 1$ (ver Figura 3.3b) la aproximación[5] $-2\log\mathscr{L} \overset{a}{\sim} \chi^2_{2L-1}$ no es una buena, dado que no se cumplen las condiciones del teorema de Wilks.

3.4.2. Detección de Series Temporales con PSD Arbitrarias

En esta sección se evalúa el comportamiento de los GLRT asintóticos, comparando sus prestaciones con las de los GLRT para procesos blancos. Concretamente, se consideran las siguientes señales[6]

$$\begin{aligned} \mathcal{H}_1: &\quad \mathbf{x}[n] = \sum_{\tau=0}^{T-1} \mathbf{H}[\tau]\mathbf{s}[n-\tau] + \mathbf{v}[n], \\ \mathcal{H}_0: &\quad \mathbf{x}[n] = \mathbf{v}[n], \end{aligned} \tag{3.47}$$

donde $\mathbf{s}[n]$ es la señal transmitida, considerada blanca espacial y temporalmente, $\mathbf{v}[n]$ es el ruido aditivo en los sensores donde sus componentes, $v_i[n]$, son procesos MA idénticos de orden q_v y potencia σ^2; y el canal $\mathbf{H}[\tau]$, compuesto por T coeficientes, es Rayleigh con perfil de potencia de los retardos (PDP) plano, que se normaliza para obtener una SNR instantánea, definida ahora como

$$\text{SNR (dB)} = 10\log_{10} \frac{\int_{-\pi}^{\pi} \operatorname{tr}\left(\mathbf{H}\left(e^{j\theta}\right)\mathbf{H}^H\left(e^{j\theta}\right)\right)\frac{d\theta}{2\pi}}{L\sigma^2}, \tag{3.48}$$

[5]La diferencia de grados de libertad entre $\mathcal{H}_1$ y $\mathcal{H}_0$ para $P = 1$ es $2L - 1$.

[6]Nótese que este modelo de señal incorpora al problema estructura temporal adicional. Por el momento se ignora dicha estructura, que será explotada en trabajos futuros.

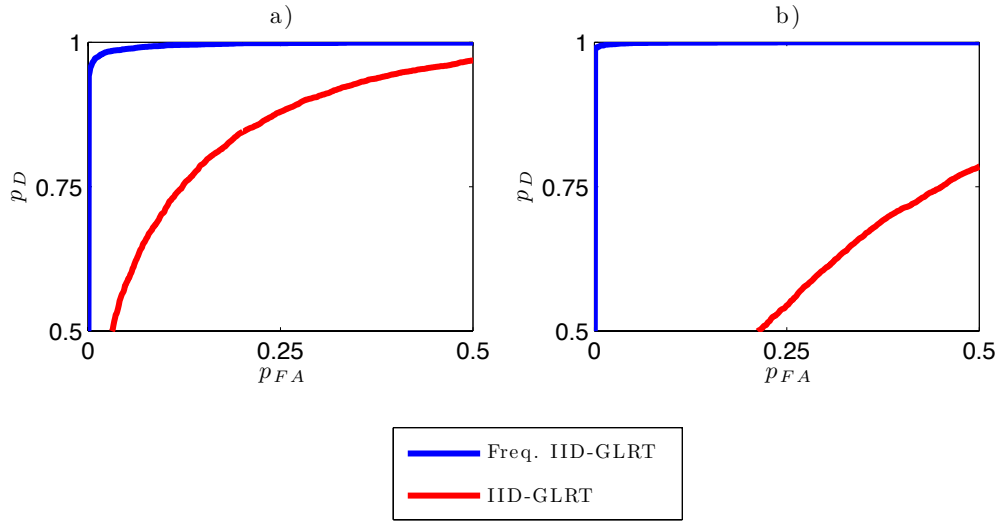

Figura 3.4: Curva ROC de los diferentes detectores para señales de rango P y ruidos IID, considerando dos selectividades frecuenciales distintas: a) $T = q_v = 4$ y b) $T = q_v = 8$. Los parámetros del experimento son: $L = 6$ sensores, $P = 1$, $M = 6$ realizaciones de longitud $N = 100$ y la SNR $= -10$ dB.

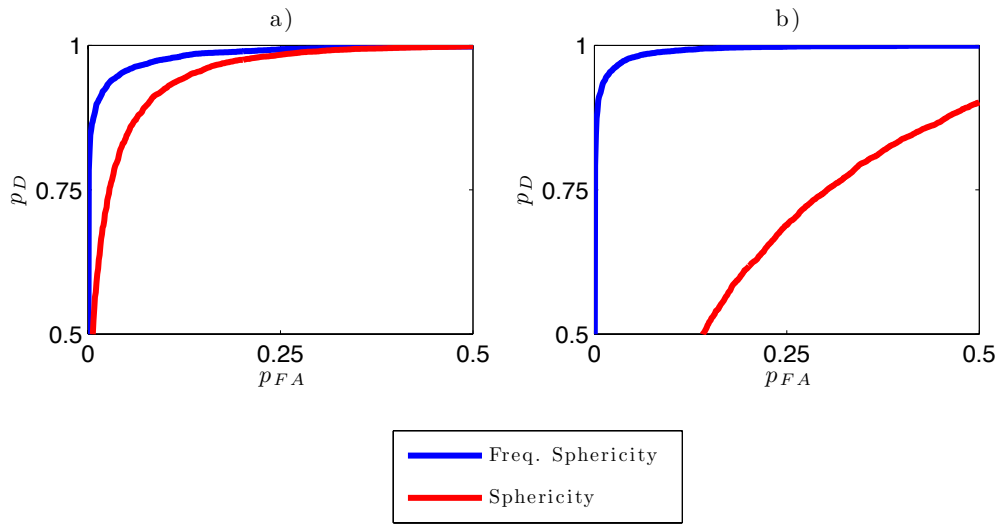

Figura 3.5: Curva ROC de los diferentes detectores para señales sin estructura espacial y ruidos IID, considerando dos selectividades frecuenciales distintas: a) $T = q_v = 4$ y b) $T = q_v = 8$. Mismos parámetros que en la Figura 3.4 salvo $P = 6$ y SNR $= -4$ dB.

constante en todas las realizaciones. Por otro lado, las integrales necesarias para los GLRT asintóticos se evalúan mediante la regla del trapecio compuesta, utilizando $K = 256$ frecuencias equiespaciadas entre $-\pi$ y π.

En el primer ejemplo se analiza el caso de señales con estructura espacial y ruidos IID en un experimento con los siguientes parámetros: $L = 6$, $P = 1$, $M = 6$ realizaciones de longitud $N = 100$, una SNR $= -10$ dB y dos selectividades frecuenciales diferentes $T = q_v = 4$ y

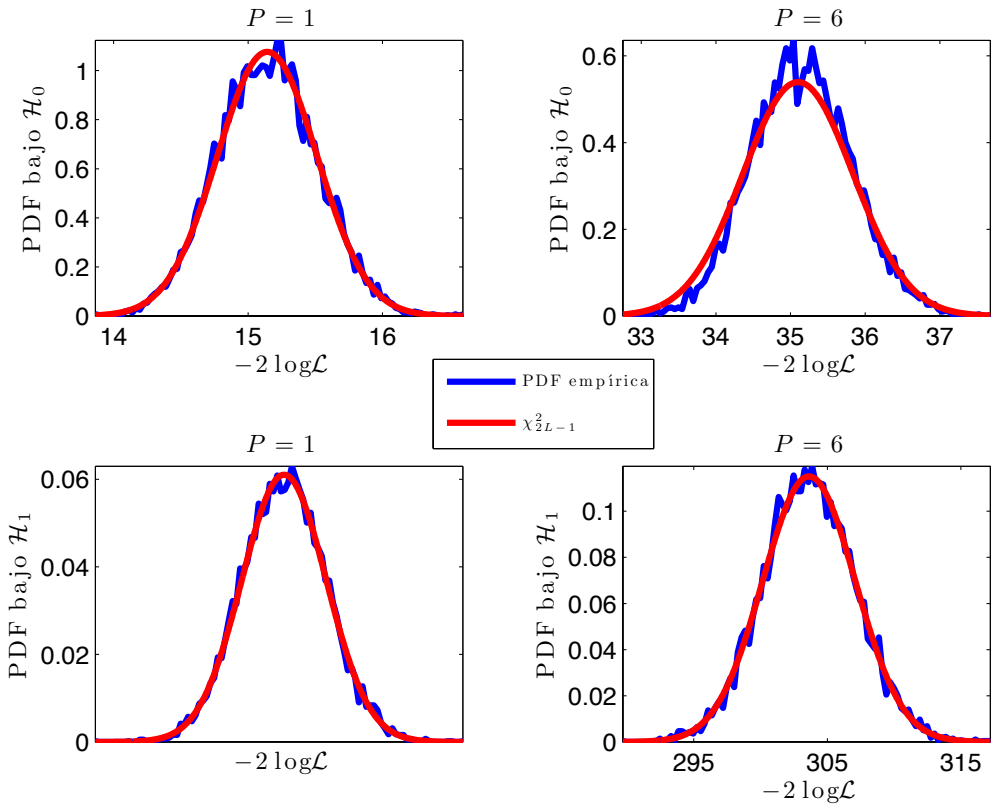

Figura 3.6: PDF estimada y aproximación Gaussiana de los GLRT asintóticos para ruidos IID en un experimento con $L = 6$, $M = 6$ y $N = 100$.

$T = q_v = 8$. Para este experimento, se comparan los GLRT dados por (3.21) y (3.32) que ignoran y tienen en cuenta, respectivamente, la estructura temporal. La Figura 3.4 muestra las curvas características de operación del receptor (ROC), donde se observa que el GLRT asintótico (3.32) presenta unas prestaciones muy superiores al GLRT para el modelo INWSPS, dado por (3.21). Dicha ganancia es más importante en escenarios con una alta selectividad frecuencial, alto T y/o q_v.

En el siguiente ejemplo, se considera un experimento similar al anterior con $P = 6$ y SNR $= -4$ dB. En esta ocasión, no existe ninguna estructura espacial adicional y, por lo tanto, el GLRT es (3.31). Concretamente, la Figura 3.5 muestra la curva ROC del GLRT (3.31) y la del GLRT (3.12), que no explota la estructura temporal, donde se pueden extraer conclusiones similares a las del experimento anterior.

Para estos dos ejemplos, considerando fijas las densidades espectrales de potencia y $M = 1000$, la Figura 3.6 muestra la PDF estimada y su aproximación Gaussiana bajo ambas hipótesis, demostrando el buen comportamiento de dicha aproximación. En este ejemplo, las medias y varianzas se han estimado a partir de las realizaciones, con la excepción de la distribución bajo $\mathcal{H}_0$ del ejemplo con $P = 6$, donde se han usado los valores teóricos derivados previamente.

Por último, el tercer experimento evalúa las prestaciones del GLRT para el modelo INMSPS. En particular, se evalúa el GLRT iterativo para señales de rango $P = 1$ y canales frecuencialmente planos, resumido en el Algoritmo 3.1; eligiendo para $\mathbf{h}^{(0)}$ el autovector

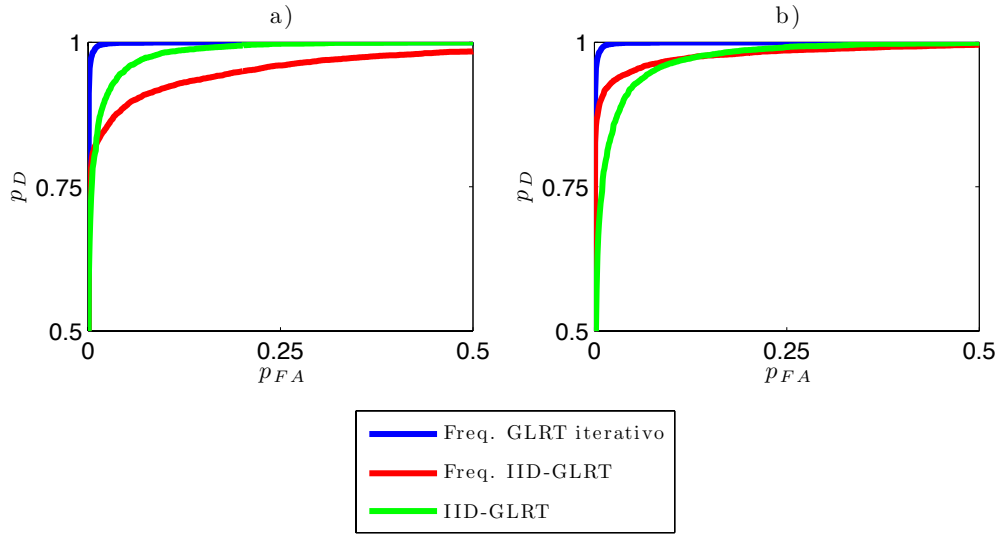

Figura 3.7: Curva ROC de los diferentes detectores para señales de rango $P = 1$, ruidos IID y canales frecuencialmente planos considerando: a) $q_s = q_v = 4$ y b) $q_s = q_v = 8$. Los parámetros del experimento son: $L = 6$ sensores, $M = 6$ realizaciones de longitud $N = 100$ y SNR $= -12$ dB.

principal de $\hat{\mathbf{R}}[0]$. Asimismo, se comparan sus prestaciones con las de los GLRT que suponen que todos los parámetros son planos o selectivos en frecuencia, dados por (3.21) y (3.32), respectivamente. Es importante señalar que únicamente el detector iterativo explota completamente la estructura espacio-temporal del problema.

Concretamente, se considera el siguiente experimento: $L = 6$ sensores, $M = 6$ realizaciones de longitud $N = 100$ y SNR $= -12$ dB, donde ahora la SNR está definida como

$$\mathrm{SNR} = 10 \log_{10} \left(\frac{E_{\mathrm{s}}}{L\sigma^2} \right),$$

siendo E_{s} la energía de la señal $s[n]$. Para este experimento, $s[n]$ es un proceso de media móvil (MA) con q_{s} coeficientes. La Figura 3.7 muestra las curvas ROC de los tres detectores mencionados previamente, donde se observa claramente las mejores prestaciones del GLRT iterativo. Por otro lado, las prestaciones de los GLRT (3.32) y (3.21) dependen de la selectividad frecuencial.

La Figura 3.8 muestra la convergencia de la log-verosimilitud bajo $\mathcal{H}_1$ del experimento anterior para 20 inicializaciones aleatorias (en azul) y la basada en el autovector principal de $\hat{\mathbf{R}}[0]$ (en rojo). Es importante señalar que todas las inicializaciones convergen, para este ejemplo, a una solución con el mismo valor de la log-verosimilitud, independientemente de la inicialización. También, la inicialización propuesta presenta una mejor convergencia. No obstante, independientemente de la inicialización, el algoritmo converge rápidamente.

3.4.3. Análisis del Detector en el Dominio de la Frecuencia

A continuación, se presentan unos resultados para ilustrar el comportamiento de los detectores frecuenciales en función de N y M, evaluando las prestaciones del GLRT asintótico (3.31). El valor de N determina la resolución espectral (el sesgo o *bias*) y el valor de M

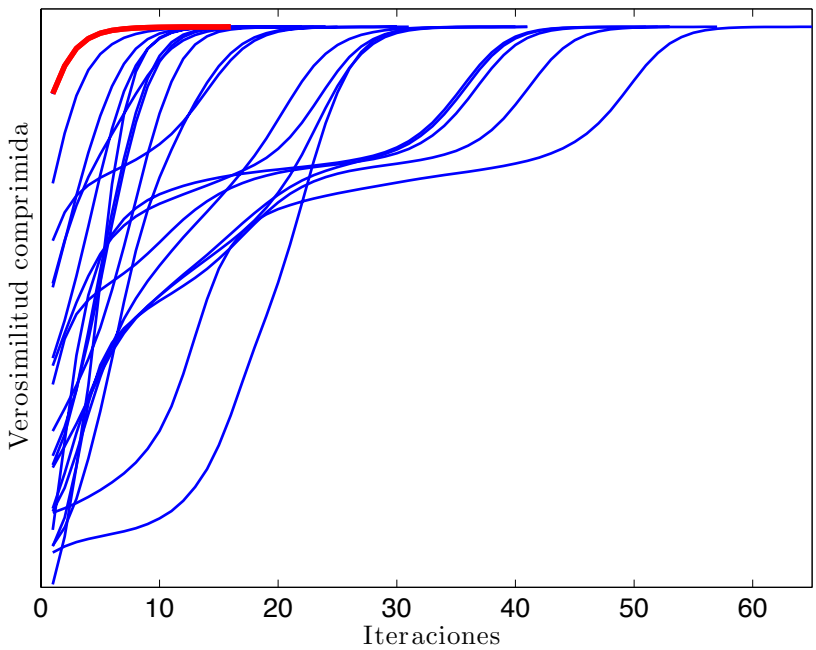

Figura 3.8: Curvas de convergencia del GLRT iterativo basado en las aproximaciones convexas sucesivas para diferentes inicializaciones.

está relacionado con la varianza, existiendo un compromiso entre sesgo y varianza para un valor prefijado de NM. Éste es un problema ampliamente estudiado en los campos de análisis espectral y estadística [Stoica y Moses, 2005], y por ello no se analizarán sus efectos en la estima de $\hat{\mathbf{S}}(e^{j\theta})$. Únicamente, se presenta una simulación para ilustrar sus efectos en el comportamiento del detector. Concretamente, se ha considerado un experimento con los siguientes parámetros: $L = 3$ sensores, $P = 3$, SNR $= 0$ dB y $T = q_v = 4$. La Figura 3.9 muestra la probabilidad de pérdida para una $p_{\mathrm{FA}} = 0.005$ y dos resoluciones diferentes, $N = 5$ y $N = 20$. En dicha figura se observa que para valores de NM pequeños se obtienen mejores resultados sacrificando resolución (N pequeño) en favor de disminuir la varianza (M grande). Por el contrario, cuando NM crece, conviene aumentar la resolución, es decir, aumentar N.

3.4.4. Aplicación a Apuntamiento de Haz

A pesar de no ser estrictamente un problema de detección, en este apartado se obtiene un patrón de apuntamiento de haz (*beamsteering*) basado en la estimación ML. La idea consiste en obtener el patrón con la respuesta de apuntamiento del array (*array's bearing response pattern*) a partir de un conjunto de medidas [Van Trees, 2002, Hua et al., 2004], usando la verosimilitud comprimida. Esta idea ha sido empleada previamente, por ejemplo en [Böhme, 1986, Bresler, 1988] y [Hua et al., 2004, Ch. 1], para señales temporalmente incorreladas. Por ello, aquí se generaliza considerando señales y ruidos temporalmente correlados. No obstante, existen algunas técnicas que consideran (y explotan) la correlación temporal de las señales, ver [Azimi-Sadjadi et al., 2008] y referencias ahí incluidas. Concretamente, en [Azimi-Sadjadi et al., 2008] se obtiene un *beamformer* banda estrecha de Capon a cada frecuencia y esta información se fusiona para obtener un conformador banda ancha mediante técnicas heurísticas, como por ejemplo, las medias aritmética, geométrica y armónica.

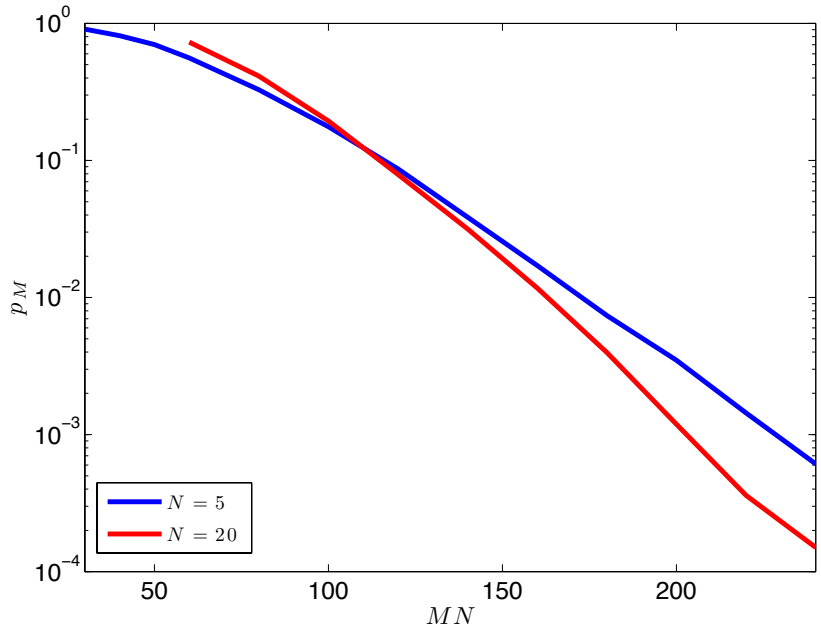

Figura 3.9: Probabilidad de pérdida para $p_{\text{FA}} = 0.005$ y dos resoluciones diferentes. Los parámetros del experimento son: $L = 3$ sensores, $P = 3$, SNR $= 0$ dB y $T = q_v = 4$.

El algoritmo de apuntamiento de haz se obtiene a partir de la verosimilitud comprimida del siguiente modelo

$$\mathbf{x}[n] = \mathbf{h}(\phi)\, s[n] + \mathbf{v}[n], \quad n = 0, \ldots, N-1,$$

donde, la única diferencia con el modelo (3.33) es el canal $\mathbf{h}$, siendo ahora

$$\mathbf{h}(\phi) = \frac{1}{\sqrt{L}} \begin{bmatrix} 1 & e^{j\phi} & \cdots & e^{j(L-1)\phi} \end{bmatrix}^T.$$

Usando las estimas ML bajo $\mathcal{H}_1$ del modelo INMSPS derivadas previamente, la log-verosimilitud comprimida es

$$\begin{aligned} \log p\left(\mathbf{z}_0, \ldots, \mathbf{z}_{M-1}; \mathbf{h}(\phi), \hat{S}_{\text{s}}\left(e^{j\theta}\right), \hat{S}_{\text{v}}\left(e^{j\theta}\right)\right) &= -NM \int_{-\pi}^{\pi} \log \alpha_{\text{tr}}\left(e^{j\theta}, \phi\right) \frac{d\theta}{2\pi} \\ &- (L-1)NM \int_{-\pi}^{\pi} \log \left[\frac{1}{L-1}\left(\text{tr}\left[\hat{\mathbf{S}}\left(e^{j\theta}\right)\right] - \alpha_{\text{tr}}\left(e^{j\theta}, \phi\right)\right)\right] \frac{d\theta}{2\pi}, \end{aligned} \tag{3.49}$$

donde

$$\alpha_{\text{tr}}\left(e^{j\theta}, \phi\right) = \text{máx}\left\{\alpha\left(e^{j\theta}, \phi\right), \frac{1}{L}\text{tr}\left[\hat{\mathbf{S}}_{\mathbf{x}}\left(e^{j\theta}\right)\right]\right\},$$

con $\alpha\left(e^{j\theta}, \phi\right) = \mathbf{h}^H(\phi)\hat{\mathbf{S}}_{\mathbf{x}}\left(e^{j\theta}\right)\mathbf{h}(\phi)$. Con lo cual, el algoritmo para apuntamiento de haz se basa en barrer en ϕ la log-verosimilitud comprimida (3.49).

Para evaluar las prestaciones de la técnica propuesta se propone aplicarla a estimación del ángulo de llegada. Para ello, se compara el MSE, definido como MSE (dB) $= 20 \log \left|\hat{\phi} - \phi\right|$, de los siguientes estimadores:

- Estimador basado en la verosimilitud asintótica comprimida dada por (3.49). La estima del ángulo se realiza mediante una búsqueda lineal (*line search*) en una rejilla de 1024 puntos.
- Estimador basado en la verosimilitud dada por [Bresler, 1988, Böhme, 1986]

$$\log p\left(\mathbf{z}_0,\ldots,\mathbf{z}_{M-1};\mathbf{h}\left(\phi\right)\right) = -NM\log\alpha_{\mathrm{tr}}\left(\phi\right) - (L-1)NM\log\left[\frac{1}{L-1}\left(\mathrm{tr}\left[\hat{\mathbf{R}}[0]\right]-\alpha_{\mathrm{tr}}\left(\phi\right)\right)\right], \quad (3.50)$$

donde

$$\alpha_{\mathrm{tr}}\left(\phi\right) = \text{máx}\left\{\alpha\left(\phi\right), \frac{1}{L}\mathrm{tr}\left[\hat{\mathbf{R}}[0]\right]\right\},$$

con $\alpha\left(\phi\right) = \mathbf{h}^H(\phi)\hat{\mathbf{R}}[0]\mathbf{h}(\phi)$. El ángulo de llegada se obtiene también mediante una búsqueda lineal.

- Estimador basado en una búsqueda lineal del patrón de apuntamiento de la técnica MUSIC. Dicho patrón se obtiene como

$$f_{\mathrm{MUSIC}}\left(\phi\right) = \frac{1}{\mathbf{h}^H\left(\phi\right)\mathbf{W}_{\mathrm{n}}\mathbf{W}_{\mathrm{n}}^H\mathbf{h}\left(\phi\right)},$$

donde $\mathbf{W}_{\mathrm{n}}$ es el subespacio de ruido de la matriz $\hat{\mathbf{R}}[0]$, es decir, los autovectores asociados a los $L-1$ autovalores más pequeños.

- Estimador basado en una búsqueda lineal del patrón de apuntamiento del *beamformer* de Capon de banda ancha [Azimi-Sadjadi et al., 2008]. El *beamformer* de Capon de banda ancha se basa en el diseño de banda estrecha mediante el criterio de Capon, y su combinación mediante técnicas heurísticas, como por ejemplo, las medias aritmética, geométrica y armónica. En concreto, el patrón de apuntamiento que mejores prestaciones presenta está basado en la media geométrica, dado por

$$f_{\mathrm{Capon}}\left(\phi\right) = \prod_{k=1}^{K}\frac{1}{\mathbf{h}^H\left(\phi\right)\hat{\mathbf{S}}^{-1}\left(e^{j\theta_k}\right)\mathbf{h}\left(\phi\right)}.$$

La Figura 3.10 muestra el MSE del siguiente experimento: $L=3$ sensores, $M=3$ realizaciones de longitud $N=128$ y $q_s=q_v=4$. En dicha figuras se observa una mejora de aproximadamente 2 dB respecto a la técnica basada en Capon y unos 4 dB respecto a las técnicas que ignoran el estructura temporal de las señales, es decir, las técnicas basadas en el MUSIC y en la verosimilitud comprimida (3.50).

3.4.5. Detección de Series Temporales no Estacionarias

Esta subsección estudia la detección de señales no estacionarias. El modelo de señal considerado es

$$\mathbf{y}[n] = \left(1+d_{\mathrm{NS}}\cos(2\pi\nu n)\right)\mathbf{x}[n],$$

donde $\mathbf{x}[n]$ está definida, bajo ambas hipótesis, en (3.47) y el grado de no estacionariedad varía con d_{NS} y ν. Para ello, se supone un experimento con: $L=3$ sensores y $P=3$ señales transmitidas, por lo tanto el GLRT es (3.44). En primer lugar, la Figura 3.11a, muestra las

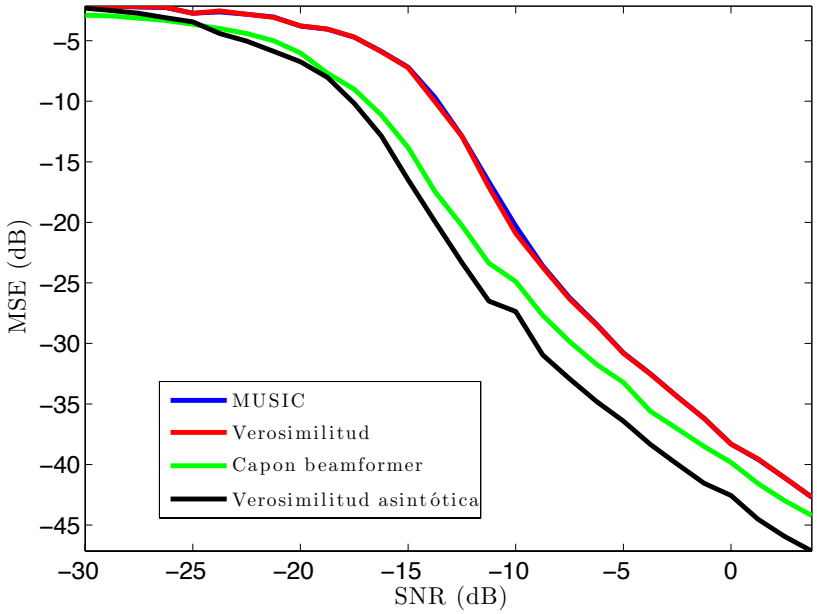

Figura 3.10: MSE de la estima del ángulo de llegada de las diferentes técnicas. Los parámetros del experimento son $L = 3$ sensores, $M = 3$ realizaciones de longitud $N = 128$ y $q_s = q_v = 4$.

curvas ROC del GLRT (3.44) y del test (3.31) para un ejemplo con los siguientes parámetros: $M = 50$ realizaciones de longitud $N = 4$, SNR $= 4$ dB, $T = q_v = 4$, $d_{\text{NS}} = 0.9$ y $\nu = 10/N$, donde se puede observar las superiores prestaciones del GLRT (3.44) comparadas con las del detector para series temporales estacionarias.

Por otro lado, la Figura 3.11b muestra las curvas ROC para otro experimento con $M = 100$ ventanas de longitud $N = 30$ y la SNR $= -7$ dB, los demás parámetros no cambian. En este caso, el detector (3.31) obtiene mejores resultados, incluso para señales no estacionarias. Este peor comportamiento se explica porque el GLRT en el dominio del tiempo necesita estimar una matriz de grandes dimensiones, $LN \times LN$, a partir de un conjunto reducido de observaciones. Sin embargo, para el detector en el dominio de la frecuencia, la matriz tiene dimensiones $L \times L$, es decir, N veces más pequeña.

3.4.6. Aplicación a Radio Cognitiva

Esta subsección presenta la aplicación de los detectores derivados en este capítulo a un problema de sensado espectral para radio cognitiva. Concretamente, se considera un usuario primario equipado con una antena ($P = 1$) transmitiendo una señal de televisión digital terrestre: señal OFDM con 8192 portadoras, ancho de banda 7.61 MHz, intervalo de guarda 1/4 y tasa del código interno 2/3. La señal se propaga en un entorno urbano (o rural): canales con desvanecimiento Rayleigh incorrelados espacialmente y PDP exponencial con una dispersión temporal 0.779 μsec (0.097 μsec) [M. Falli, e.d., 1989]. En el receptor del usuario secundario las señales son demoduladas y muestreadas asíncronamente a 16 MHz. La distribución de los ruidos no cambia respecto a apartados anteriores.

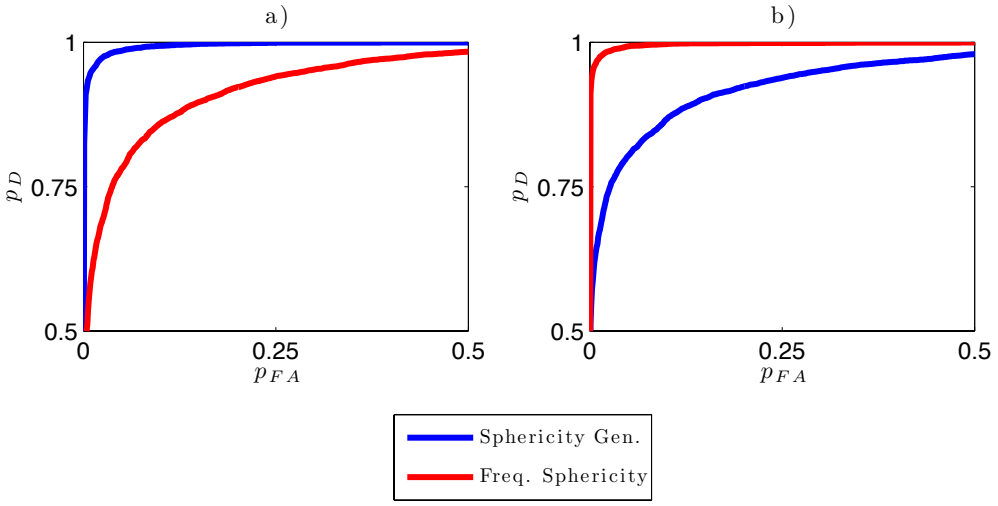

Figura 3.11: Curva ROC de los diferentes detectores para señales no estacionarias considerando diferentes tamaños de ventana: a) $N = 4$ y b) $N = 30$.

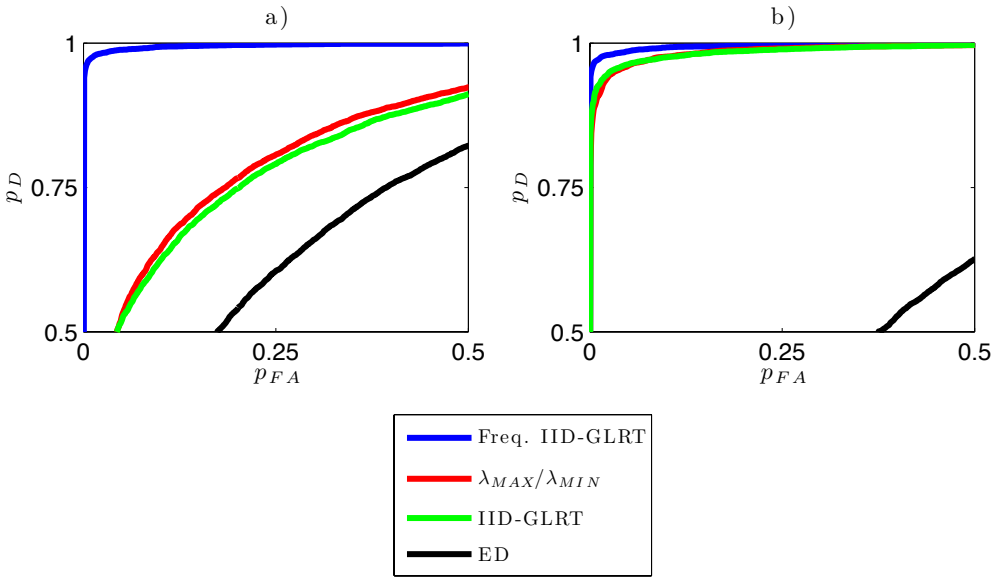

Figura 3.12: Curva ROC para los diferentes detectores en un ejemplo de radio cognitiva con ruidos IID en dos entornos de propagación: a) área urbana y b) área rural.

En concreto, la Figura 3.12 muestra la curvas ROC de los GLRT (3.21) y (3.32) para un ejemplo con los siguiente parámetros: $L = 3$ antenas, $P = 1$, SNR $= -8$ dB y $M = 10$ realizaciones de longitud $N = 100$. Asimismo, se presentan también las prestaciones de dos detectores típicos de radio cognitiva, el detector de energía (denotado como ED) y el detector propuesto en [Zeng y Liang, 2009] (denotado como $\lambda_{\text{MAX}}/\lambda_{\text{MIN}}$). Una vez más, incluso para señales OFDM, las mejores prestaciones se obtienen con el detector que explota completamente la estructura espacio-temporal. Asimismo, las prestaciones del detector de energía son pobres en este escenario y el comportamiento del detector (3.21) y el detector propuesto

en [Zeng y Liang, 2009] es bastante similar. Por último, la ventaja de explotar la estructura temporal se reduce para entornos poco selectivos en frecuencia (entorno rural).

3.5. Conclusiones

Este capítulo ha estudiado el problema de detección de series temporales multivariadas contaminadas por ruidos independientes e idénticamente distribuidos (IID). En concreto, se ha derivado el test basado en el cociente de verosimilitudes generalizado (GLRT) bajo diferentes suposiciones acerca de la estructura temporal y espacial de las señales.

El modelo más sencillo ha considerado señales y ruidos estacionarios en sentido amplio (WSS) con densidades espectrales de potencia (PSD) planas. Para este escenario, el GLRT es una función, dependiente de la estructura espacial concreta, de los autovalores de la matriz de covarianza muestral. Por ejemplo, para señales sin estructura espacial adicional, el GLRT es el cociente de las medias aritmética y geométrica de los autovalores. Por otro lado, los GLRT son invariantes a un escalado de la serie temporal, permitiendo obtener, mediante simulaciones, la distribución del estadístico bajo $\mathcal{H}_0$, independientemente de la SNR. Las prestaciones de los detectores se han evaluado mediante simulaciones de Monte Carlo, ilustrando el mejor comportamiento obtenido por explotar correctamente la estructura espacial.

Los detectores anteriores se han extendido considerando señales WSS con PSD arbitrarias. La derivación del GLRT para estos modelos requiere obtener las estimas de máxima verosimilitud (ML) de matrices Toeplitz por bloques, que es un problema de optimización no convexo. No obstante, aplicando la verosimilitud asintótica, en el dominio frecuencia, se ha evitado la estimación ML de matrices Toeplitz, facilitando la obtención de fórmulas cerradas para el GLRT. En todos los modelos considerados, el GLRT para procesos con PSD arbitrarias ha podido interpretarse como la integral del GLRT análogo para series con PSD planas. Adicionalmente, se ha considerado un modelo mixto, es decir, señales y ruidos coloreados en canales frecuencialmente planos. Para este problema, incluso aplicando la verosimilitud asintótica, no existe formulación cerrada para el GLRT, siendo necesario emplear técnicas numéricas. Además, al igual que para procesos blancos, los GLRT para procesos con PSD arbitrarias son invariantes a filtrados comunes, y por lo tanto, se puede obtener su distribución bajo $\mathcal{H}_0$, independientemente de la PSD del ruido. Una vez más, los resultados numéricos han mostrado el mejor comportamiento de los detectores que explotan completamente la estructura espacio-temporal.

Finalmente, se han considerado series temporales no estacionarias. Bajo esta suposición, las matrices de covarianza no poseen estructura Toeplitz por bloques, permitiendo la derivación del GLRT sin recurrir a la verosimilitud asintótica. Asimismo, se ha presentado una aproximación del GLRT, en el dominio frecuencial, aplicando la invarianza a transformaciones lineales. Los resultados numéricos ilustran el buen comportamiento del GLRT en el dominio del tiempo para ventanas cortas. No obstante, para ventanas largas las prestaciones del detector se degradan, presentando mejores prestaciones la aproximación propuesta en el dominio de la frecuencia.

Capítulo 4

Detección Multicanal de Procesos Contaminados por Ruidos no-IID

Este capítulo generaliza los detectores presentados en el Capítulo 3 considerando ruidos independientes con diferentes distribuciones (no-IID). De manera equivalente al capítulo anterior, se comienza por el caso más sencillo de series temporales con densidades espectrales de potencia (PSD) planas. Para el test de hipótesis sin estructura espacial bajo $\mathcal{H}_1$, el GLRT se desarrolló en los años 30 y viene dado por el cociente de Hadamard de la matriz de covarianza muestral. Por otro lado, también se considera el problema de detección con estructura espacial adicional, y en esta situación la matriz de covarianza bajo $\mathcal{H}_1$ viene dada por la suma de una matriz de rango reducido y una matriz diagonal. En el capítulo se demuestra que el problema de estimación de máxima verosimilitud bajo $\mathcal{H}_1$ es no convexo, siendo necesario recurrir a soluciones numéricas subóptimas o aproximaciones. En primer lugar, se considera la minimización alternada de la función de coste y, a continuación, motivado por el alto coste computacional de esta técnica, se propone una solución aproximada para escenarios de baja relación señal a ruido (SNR), de especial interés en radio cognitiva.

Los resultados obtenidos se extienden posteriormente a series temporales con PSD arbitrarias. Una vez más, la extensión se basa en la verosimilitud asintótica, y los GLRT derivados pueden interpretarse como la integral de los GLRT para procesos blancos. Asimismo, también se deriva el GLRT para la detección de correlación espacial en series temporales no estacionarias sin ninguna estructura espacial adicional, que puede considerarse un cociente de Hadamard generalizado.

Este capítulo está organizado de la siguiente manera: la Sección 4.1 considera la detección de procesos estacionarios blancos y sin estructura espacial bajo $\mathcal{H}_1$, contaminados con ruidos no-IID; y posteriormente, se generaliza a modelos de rango reducido. A continuación, la Sección 4.2 extiende dichos problemas, con y sin estructura espacial, a series temporales estacionarias con espectros arbitrarios, considerando únicamente la aproximación de baja SNR. La Sección 4.3 presenta el último escenario analizado, la detección de correlación espacial en señales no estacionarias. Las prestaciones de todos los detectores, así como la comparación con otros detectores, se analizan mediante simulaciones de Monte Carlo en la Sección 4.4. Por último, la Sección 4.5 presenta las principales conclusiones obtenidas.

4.1. Detección Multicanal de Procesos Estacionarios Blancos

Esta sección aborda el problema de detección multicanal de procesos temporalmente blancos contaminados con ruidos no-IID. Siguiendo un desarrollo similar al de la Sección 3.1, la familia genérica de tests de hipótesis considerada en esta sección es

$$\begin{aligned} \mathcal{H}_1 &: \mathbf{x} \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{R}_1\right), \\ \mathcal{H}_0 &: \mathbf{x} \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{\Sigma}^2\right), \end{aligned} \tag{4.1}$$

donde $\mathbf{\Sigma}^2 \in \mathbb{D}_+$ y $\mathbf{R}_1 \in \mathbb{S}_+$, siendo $\mathbb{D}_+$ el conjunto de matrices diagonales definidas positivas y $\mathbb{S}_+$ el conjunto de matrices definidas positivas. En particular, se deriva el GLRT bajo las mismas suposiciones que en el capítulo anterior respecto a la estructura espacial de las señales.

4.1.1. Señales sin Estructura Espacial

En este apartado, únicamente se supone que la matriz $\mathbf{R}_1$ es definida positiva, es decir, no posee ninguna estructura espacial adicional (modelo NNWSNS). Este test de hipótesis fue resuelto en [Wilks, 1935] para vectores aleatorios reales y Gaussianos, donde se testeaba la independencia de un conjunto de L variables aleatorias. Posteriormente, este test se aplicó a procesado en arrays y radioastronomía en [Leshem y Van der Veen, 2001a, Leshem y Van der Veen, 2001b].

El test para el modelo NNWSNS es

$$\begin{aligned} \mathcal{H}_1 &: \mathbf{x} \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{R}_1\right), \\ \mathcal{H}_0 &: \mathbf{x} \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{\Sigma}^2\right), \end{aligned}$$

donde $\mathbf{\Sigma}^2 \in \mathbb{D}_+$ es una matriz de covarianza diagonal y $\mathbf{R}_1 \in \mathbb{S}_+$. Suponiendo M ($M \geq L$) medidas independientes e idénticamente distribuidas (IID) disponibles para derivar el test: $\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}$, el GLRT se obtiene a partir del siguiente GLR

$$\mathscr{L} = \frac{\operatorname*{máx}_{\mathbf{\Sigma}^2 \in \mathbb{D}_+} p\left(\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}; \mathbf{\Sigma}^2\right)}{\operatorname*{máx}_{\mathbf{R}_1 \in \mathbb{S}_+} p\left(\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}; \mathbf{R}_1\right)}.$$

Para derivar el GLR es necesario obtener las estimas ML de $\mathbf{\Sigma}^2$ y $\mathbf{R}_1$, las cuales se presentan en los siguientes lemas.

Lema 4.1. *La estima ML de* $\mathbf{\Sigma}^2$ *es*

$$\hat{\mathbf{\Sigma}}^2 = \hat{\mathbf{D}} = \operatorname{diag}\left([\hat{\mathbf{R}}]_{1,1}, \ldots, [\hat{\mathbf{R}}]_{L,L}\right), \tag{4.2}$$

donde $\operatorname{diag}\left([\hat{\mathbf{R}}]_{1,1}, \ldots, [\hat{\mathbf{R}}]_{L,L}\right)$ *es una matriz diagonal formada por los elementos* $[\hat{\mathbf{R}}]_{i,i}$, $i = 1, \ldots, L$.

Demostración. Dada la estructura diagonal de $\mathbf{\Sigma}^2$, (3.4) puede reescribirse como el producto de las PDF marginales, siendo posible optimizar independientemente cada $[\mathbf{\Sigma}^2]_{i,i}$. Resumiendo, la i-ésima log-verosimilitud viene dada por

$$\log p\left(\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}; [\mathbf{\Sigma}^2]_{i,i}\right) = -M \log \pi - M \log\left([\mathbf{\Sigma}^2]_{i,i}\right) - M \frac{[\hat{\mathbf{R}}]_{i,i}}{[\mathbf{\Sigma}^2]_{i,i}},$$

y las estimas ML son $[\hat{\mathbf{\Sigma}}^2]_{i,i} = [\hat{\mathbf{R}}]_{i,i}$, $i = 1, \ldots, L$. □

Teniendo en cuenta la estima (4.2), la log-verosimilitud comprimida es

$$\log p\left(\mathbf{x}_0,\ldots,\mathbf{x}_{M-1};\hat{\mathbf{\Sigma}}^2\right) = -LM\log\pi - LM - M\log\det\left(\hat{\mathbf{D}}\right) \tag{4.3}$$

Bajo $\mathcal{H}_1$, la estima de la matriz de covarianza se obtuvo en la Sección 3.1.1 y viene dada por $\hat{\mathbf{R}}_1 = \hat{\mathbf{R}}$. Por lo tanto, teniendo en cuenta las verosimilitudes comprimidas, dadas por (4.3) y (3.11), el log-GLRT es

$$\log\mathscr{L} = \log p\left(\mathbf{x}_0,\ldots,\mathbf{x}_{M-1};\hat{\mathbf{\Sigma}}^2\right) - \log p\left(\mathbf{x}_0,\ldots,\mathbf{x}_{M-1};\hat{\mathbf{R}}_1\right) = \log\left\{\frac{\det\left(\hat{\mathbf{R}}\right)}{\prod_{i=1}^{L}\left[\hat{\mathbf{D}}\right]_{i,i}}\right\}. \tag{4.4}$$

El GLRT (4.4) es el cociente de Hadamard de la matriz de covarianza muestral, que es una medida de la relación lineal entre las componentes del vector $\mathbf{x}$ [Scharf y Thomas, 1998]. Finalmente, definiendo la matriz de coherencia

$$\hat{\mathbf{C}} = \hat{\mathbf{D}}^{-1/2}\hat{\mathbf{R}}\hat{\mathbf{D}}^{-1/2},$$

se puede reescribir (4.4) como

$$\log\mathscr{L} = \log\det\left(\hat{\mathbf{C}}\right). \tag{4.5}$$

Esta es una expresión interesante del GLRT que permite relacionarlo con el análisis de correlaciones canónicas[1] (CCA) [Hotelling, 1936] y su extensión a más de dos conjuntos de datos [Kettenring, 1971].

Adicionalmente, el GLRT (4.4) es invariante a escalados independientes de las series temporales, propiedad que permite obtener, mediante simulaciones, los umbrales de decisión, independientemente de la matriz $\mathbf{\Sigma}^2$.

Lema 4.2. *El GLRT* (4.4) *es invariante a transformaciones* $\mathbf{y} = \mathbf{A}\mathbf{x}$*, siendo* $\mathbf{A}$ *una matriz diagonal invertible.*

Demostración. Es fácil demostrar que las matrices de covarianza muestrales de las señales $\mathbf{y}$ vienen dadas por

$$\hat{\mathbf{R}}^{(y)} = \mathbf{A}\hat{\mathbf{R}}^{(x)}\mathbf{A}^H,$$
$$\hat{\mathbf{D}}^{(y)} = \mathbf{A}\hat{\mathbf{D}}^{(x)}\mathbf{A}^H,$$

y sustituyendo dichas matrices en (4.4) se obtiene el resultado deseado. □

Por otro lado, el GLRT (4.4) sí cumple las condiciones del teorema de Wilks, siendo su distribución asintótica χ^2_ν bajo $\mathcal{H}_0$, donde ν es la diferencia entre el número de grados de libertad de ambas hipótesis. Los grados de libertad bajo $\mathcal{H}_1$, L^2, se han obtenido previamente; por otro lado, bajo $\mathcal{H}_0$ sólo $\mathbf{\Sigma}^2$ es desconocido, siendo sus grados de libertad L. Resumiendo, el GLRT basado en el cociente de Hadamard se distribuye asintóticamente bajo $\mathcal{H}_0$ como

$$-2\log\mathscr{L} \overset{a}{\sim} \chi^2_{L^2-L}.$$

Una vez más, el teorema de Wilks permite obtener, de manera teórica, los umbrales cuando el número de observaciones es elevado.

[1] Un resumen de CCA se puede encontrar en el Apéndice C.

Aproximación de baja correlación

En este apartado, se presenta la aproximación de baja correlación espacial o SNR, derivada en [Leshem y Van der Veen, 2001a, Leshem y Van der Veen, 2001b], del GLRT (4.5). Para baja correlación, la matriz de coherencia se aproxima a la matriz identidad, es decir, $\mathbf{C} \approx \mathbf{I}$ y, por ello, sus autovalores se pueden expresar como

$$\beta_i = 1 + \epsilon_i, \quad i = 1, \ldots, L,$$

donde $|\epsilon_i| \ll 1$ y $\sum_{i=1}^{L} \epsilon_i = 0$, dado que $\operatorname{tr}\left(\hat{\mathbf{C}}\right) = L$. Sustituyendo β_i en (4.5), el GLRT es

$$\log \mathscr{L} = \log \det \left(\hat{\mathbf{C}}\right) = \sum_{i=1}^{L} \log \beta_i = \sum_{i=1}^{L} \log \left(1 + \epsilon_i\right).$$

Para obtener el GLRT para baja correlación, consideremos la aproximación de Taylor de orden 2, dada por $\log(1 + x) \approx x - x^2/2$ para $|x| \ll 1$, obteniendo el GLRT basado en la norma Frobenius de la matriz de coherencia

$$\log \mathscr{L} \approx - \left\|\hat{\mathbf{C}}\right\|_{\mathrm{F}}^2. \tag{4.6}$$

Es importante destacar, como se verá en la sección de resultados numéricos, que la norma Frobenius es menos sensible al número de observaciones disponibles, es decir, el detector basado en la norma Frobenius es más robusto que el detector basado en el determinante.

4.1.2. Señales con Estructura Espacial

Esta subsección generaliza el modelo con ruidos no-IID, incorporando estructura espacial al problema de detección. Por lo tanto, el test de hipótesis se corresponde con el modelo NNWSPS, dado por[2]

$$\begin{aligned} \mathcal{H}_1 &: \mathbf{x} \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{H}\mathbf{H}^H + \mathbf{\Sigma}^2\right), \\ \mathcal{H}_0 &: \mathbf{x} \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{\Sigma}^2\right), \end{aligned} \tag{4.7}$$

donde $\mathbf{H} \in \mathbb{C}^{L \times P}$ tiene rango de columnas completo.

Comencemos analizando para qué valores de P la matriz $\mathbf{R}_1 = \mathbf{H}\mathbf{H}^H + \mathbf{\Sigma}^2$ tiene estructura adicional que pueda explotarse para mejorar las estimas ML.

Lema 4.3. *Para $P \geq L - \sqrt{L}$, la matriz $\mathbf{R}_1 = \mathbf{H}\mathbf{H}^H + \mathbf{\Sigma}^2$ es una matriz definida positiva, sin otra estructura adicional.*

Demostración. La única diferencia respecto al caso de ruidos IID es la matriz $\mathbf{\Sigma}^2$. Por lo tanto, el número de grados de libertad de $\mathbf{R}_1$ es

$$\underbrace{2LP - P^2}_{\mathbf{H}\mathbf{H}^H} + \underbrace{L}_{\mathbf{\Sigma}^2} = 2LP - P^2 + L. \tag{4.8}$$

La demostración concluye comparando el número de parámetros libres de una matriz definida positiva, dado por (3.15), y (4.8), es decir,

$$2LP - P^2 + L \geq L^2 \Rightarrow P \geq L - \sqrt{L}.$$

□

[2] Sin pérdida de generalidad se ha supuesto $\mathbf{R}_s = \mathbf{I}$.

Para $P \geq L - \sqrt{L}$, el problema de detección es equivalente al presentado en la Sección 4.1.1, siendo su GLRT el cociente de Hadamard de la matriz de covarianza muestral. Por el contrario, para $P < L - \sqrt{L}$ es posible explotar la estructura de $\mathbf{R}_1$ para mejorar su estima.

La estima bajo $\mathcal{H}_0$ viene dada por (4.2) y las estimas bajo $\mathcal{H}_1$ se presentan en el Lema 4.4. Previamente, para simplificar la estimación ML de $\mathbf{H}$ y $\mathbf{\Sigma}^2$ se reparametriza el problema. Definiendo $\hat{\mathbf{R}}_{\mathbf{\Sigma}} = \mathbf{\Sigma}^{-1}\hat{\mathbf{R}}\mathbf{\Sigma}^{-1}$ y $\mathbf{H}_{\mathbf{\Sigma}} = \mathbf{\Sigma}^{-1}\mathbf{H}$ como la matriz de covarianza muestral blanqueada y el canal blanqueado, respectivamente, se puede reescribir la log-verosimilitud como

$$\begin{aligned}\log p\left(\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}; \mathbf{H}_{\mathbf{\Sigma}}, \mathbf{\Sigma}^2\right) = &- LM \log \pi - M \log\det\left(\mathbf{H}_{\mathbf{\Sigma}}\mathbf{H}_{\mathbf{\Sigma}}^H + \mathbf{I}\right) \\ &- M \log\det\left(\mathbf{\Sigma}^2\right) - M\mathrm{tr}\left[\left(\mathbf{H}_{\mathbf{\Sigma}}\mathbf{H}_{\mathbf{\Sigma}}^H + \mathbf{I}\right)^{-1}\hat{\mathbf{R}}_{\mathbf{\Sigma}}\right]. \end{aligned} \tag{4.9}$$

Por otro lado, derivar las estimas ML de $\mathbf{H}_{\mathbf{\Sigma}}$ y $\mathbf{\Sigma}$ es complicado, y por ello se propone obtener las estimas ML de los autovectores y autovalores de $\mathbf{H}_{\mathbf{\Sigma}}\mathbf{H}_{\mathbf{\Sigma}}^H$. En concreto, aplicando la EVD a $\mathbf{H}_{\mathbf{\Sigma}}\mathbf{H}_{\mathbf{\Sigma}}^H = \mathbf{G}_{\mathrm{c}}\mathbf{\Phi}_{\mathrm{c}}^2\mathbf{G}_{\mathrm{c}}^H$ se obtienen en el siguiente lema las estimas ML de $\mathbf{G}_{\mathrm{c}}$ y $\mathbf{\Phi}_{\mathrm{c}}^2$.

Lema 4.4. *Teniendo en cuenta la EVD de $\hat{\mathbf{R}}_{\mathbf{\Sigma}}$ dada por $\hat{\mathbf{R}}_{\mathbf{\Sigma}} = \mathbf{Q}\mathrm{diag}\left(\gamma_1, \ldots, \gamma_L\right)\mathbf{Q}^H$, con $\gamma_1 \geq \cdots \geq \gamma_L$. Las estimas ML de $\mathbf{G}_{\mathrm{c}}$ y $\mathbf{\Phi}_{\mathrm{c}}^2 = \mathrm{diag}\left(\phi_1, \ldots, \phi_L\right)$ (que dependen de $\mathbf{\Sigma}^2$) son*

$$\begin{aligned}\hat{\mathbf{G}}_{\mathrm{c}} &= \mathbf{Q}, \\ \hat{\phi}_i^2 &= \begin{cases} \gamma_i - 1, & i = 1, \ldots, P, \\ 0, & i = P+1, \ldots, L. \end{cases}\end{aligned}$$

Demostración. Una vez que $\hat{\mathbf{R}}_{\mathbf{\Sigma}}$ y $\mathbf{H}_{\mathbf{\Sigma}}$ han sido blanqueadas, la demostración de este lema sigue las mismas líneas que la del Lema 3.5. □

Sustituyendo las estimas ML del Lema 4.4 en (4.9), la verosimilitud se simplifica a

$$\begin{aligned}\log p\left(\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}; \hat{\mathbf{H}}_{\mathbf{\Sigma}}, \mathbf{\Sigma}^2\right) = &- LM \log \pi - M\sum_{i=1}^{P}\log\gamma_i - M\log\det\left(\mathbf{\Sigma}^2\right) \\ &- M\left(P + \sum_{i=P+1}^{L}\gamma_i\right).\end{aligned}$$

y sumando y restando $\log\det\left(\hat{\mathbf{R}}\right)$, la expresión anterior puede reescribirse como

$$\begin{aligned}\log p\left(\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}; \hat{\mathbf{H}}_{\mathbf{\Sigma}}, \mathbf{\Sigma}^2\right) = &- LM \log \pi - MP - M\log\det\left(\hat{\mathbf{R}}\right) \\ &- M\sum_{i=P+1}^{L}\left[\gamma_i - \log\gamma_i\right].\end{aligned} \tag{4.10}$$

Finalmente, para obtener el GLRT hay que maximizar en $\mathbf{\Sigma}^2$ la función objetivo (4.10), sin embargo, se puede demostrar que es un problema de optimización no-convexo [Boyd y Vandenberghe, 2004]; y tampoco es posible encontrar una solución de manera analítica, necesaria para derivar una expresión cerrada para el GLRT. Debido a esto, se propone usar técnicas de optimización no convexas, que no garantizan la obtención del óptimo global en tiempo polinómico.

4.1: Estima iterativa de $\mathbf{H}_{\boldsymbol{\Sigma}}$ y $\boldsymbol{\Sigma}$ vía optimización alternada.

Entradas: Punto inicial $\boldsymbol{\alpha}_{(0)}$ y $\hat{\mathbf{R}}$
Salidas: Estimas de $\mathbf{H}_{\boldsymbol{\Sigma}}$ y $\boldsymbol{\Sigma}$ bajo $\mathcal{H}_1$
Inicializar: $i = 0$
repeat
- Construir $\boldsymbol{\Sigma}_{(i)}^{-1} = \mathrm{diag}(\boldsymbol{\alpha}_{(i)})$
- Obtener $\hat{\mathbf{R}}_{\boldsymbol{\Sigma}}^{(i+1)} = \boldsymbol{\Sigma}_{(i)}^{-1}\hat{\mathbf{R}}\boldsymbol{\Sigma}_{(i)}^{-1}$ y su EVD
- Estimar $\mathbf{H}_{\boldsymbol{\Sigma}}^{(i+1)}$ a partir de (4.12) (para $\boldsymbol{\Sigma}_{(i)}^{-1}$ fijo)
- Resolver (4.13) para obtener $\boldsymbol{\alpha}_{(i+1)}$ (para $\mathbf{H}_{\boldsymbol{\Sigma}}^{(i+1)}$ fijo)
- Actualizar $i = i + 1$

until Convergencia

4.1.3. Optimización Alternada

Esta subsección presenta una solución subóptima para maximizar (4.10). Comencemos por reformular la maximización de (4.9) como el siguiente problema optimización

$$\begin{aligned} \underset{\mathbf{H}_{\boldsymbol{\Sigma}},\boldsymbol{\Sigma}}{\text{minimizar}} \quad & \mathrm{tr}\left(\hat{\mathbf{R}}\boldsymbol{\Sigma}^{-1}\mathbf{R}_{\boldsymbol{\Sigma}}^{-1}\boldsymbol{\Sigma}^{-1}\right) - \log\det\left(\boldsymbol{\Sigma}^{-2}\right) + \log\det\mathbf{R}_{\boldsymbol{\Sigma}}, \\ \text{sujeto a} \quad & \mathbf{R}_{\boldsymbol{\Sigma}} = \mathbf{I} + \mathbf{H}_{\boldsymbol{\Sigma}}\mathbf{H}_{\boldsymbol{\Sigma}}^{H}, \\ & \boldsymbol{\Sigma}^2 \in \mathbb{D}_+. \end{aligned} \tag{4.11}$$

El problema (4.11) es aún no convexo, pero esta formulación permite dividir las variables libres en dos conjuntos, obteniendo sendos problemas convexos para cada grupo de parámetros, si las restantes variables se consideran constantes. Es decir, se optimizará alternativamente cada conjunto de incógnitas, manteniendo el otro fijo, hasta que el algoritmo converja a un punto fijo. Es importante destacar que, debido a que en cada iteración la función de coste únicamente puede decrecer, se garantiza la convergencia del algoritmo a un mínimo, que puede ser local.

- Minimización respecto a $\mathbf{H}_{\boldsymbol{\Sigma}}$: Para $\boldsymbol{\Sigma}$ fija, la matriz $\mathbf{H}_{\boldsymbol{\Sigma}}$ que minimiza (4.11) viene dada por el Lema 4.4, es decir,

$$\hat{\mathbf{H}}_{\boldsymbol{\Sigma}} = \begin{bmatrix}\mathbf{q}_1 & \cdots & \mathbf{q}_P\end{bmatrix}\left(\mathrm{diag}\left(\gamma_1, \ldots, \gamma_P\right) - \mathbf{I}\right)^{1/2}, \tag{4.12}$$

  donde $\mathbf{q}_i$ es la i-ésima columna de $\mathbf{Q}$. Cabe mencionar que esta solución no es única y cualquier multiplicación de $\hat{\mathbf{H}}_{\boldsymbol{\Sigma}}$ por una matriz unitaria por la derecha es también una solución.

- Minimización respecto a $\boldsymbol{\Sigma}$: Para $\mathbf{H}_{\boldsymbol{\Sigma}}$ fija, (4.11) se simplifica a

$$\begin{aligned} \underset{\boldsymbol{\Sigma}}{\text{minimizar}} \quad & \mathrm{tr}\left(\hat{\mathbf{R}}\boldsymbol{\Sigma}^{-1}\mathbf{R}_{\boldsymbol{\Sigma}}^{-1}\boldsymbol{\Sigma}^{-1}\right) - \log\det\left(\boldsymbol{\Sigma}^{-2}\right), \\ \text{sujeto a} \quad & \boldsymbol{\Sigma} \in \mathbb{D}_+. \end{aligned}$$

  A continuación, es fácil manipular la traza para obtener

$$\mathrm{tr}\left(\hat{\mathbf{R}}\boldsymbol{\Sigma}^{-1}\mathbf{R}_{\boldsymbol{\Sigma}}^{-1}\boldsymbol{\Sigma}^{-1}\right) = \boldsymbol{\alpha}^T\left(\hat{\mathbf{R}}^T \odot \mathbf{R}_{\boldsymbol{\Sigma}}^{-1}\right)\boldsymbol{\alpha},$$

donde $\boldsymbol{\alpha} = \left[[\boldsymbol{\Sigma}^{-1}]_{1,1}, \ldots, [\boldsymbol{\Sigma}^{-1}]_{L,L}\right]^T$ y $\odot$ denota el producto de Hadamard, obteniéndose el siguiente problema de optimización

$$\begin{aligned} \underset{\boldsymbol{\alpha}}{\text{minimizar}} \quad & \boldsymbol{\alpha}^T \left(\hat{\mathbf{R}}^T \odot \mathbf{R}_{\boldsymbol{\Sigma}}^{-1}\right) \boldsymbol{\alpha} - \sum_{i=1}^{L} \log \alpha_i^2, \\ \text{sujeto a} \quad & \alpha_i \geq 0. \end{aligned} \tag{4.13}$$

Recordemos que $\operatorname{tr}\left(\hat{\mathbf{R}}\boldsymbol{\Sigma}^{-1}\mathbf{R}_{\boldsymbol{\Sigma}}^{-1}\boldsymbol{\Sigma}^{-1}\right)$ se obtuvo al reescribir una fórmula cuadrática de una matriz de covarianza definida positiva, por lo tanto, se cumple $\operatorname{tr}\left(\hat{\mathbf{R}}\boldsymbol{\Sigma}^{-1}\mathbf{R}_{\boldsymbol{\Sigma}}^{-1}\boldsymbol{\Sigma}^{-1}\right) > 0$ para cualquier matriz $\boldsymbol{\Sigma}^2 \in \mathbb{D}_+$. Esto implica que $\hat{\mathbf{R}}^T \odot \mathbf{R}_{\boldsymbol{\Sigma}}^{-1}$ es una matriz definida positiva y el problema de optimización (4.13) es convexo [Boyd y Vandenberghe, 2004], siendo posible entonces obtener el óptimo global (para $\mathbf{H}_{\boldsymbol{\Sigma}}$ fijo) de manera eficiente.

Los principales pasos de la técnica de optimización alternada se resumen en el Algoritmo 4.1. Finalmente, el GLRT se obtiene sustituyendo las estimas ML derivadas previamente en (4.9) y restando las log-verosimilitudes de ambas hipótesis.

4.1.4. Aproximación para Baja SNR

La complejidad computacional del detector presentado en el Algoritmo 4.1 puede reducir su utilidad en aplicaciones prácticas. Debido a esto, desarrollar detectores más simples y que admitan una fórmula cerrada es importante. A continuación, se deriva un GLRT aproximado que admite una expresión cerrada en el supuesto de baja SNR. En este contexto, la matriz de covarianza bajo $\mathcal{H}_1$ estará dominada por la matriz de covarianza del ruido, es decir, $\mathbf{R}_1 \approx \boldsymbol{\Sigma}^2$, y la estima ML de $\boldsymbol{\Sigma}^2$ es $\hat{\boldsymbol{\Sigma}}^2 \approx \hat{\mathbf{D}}$, donde $\hat{\mathbf{D}}$ está definida en (4.2). Por lo tanto, la verosimilitud comprimida es

$$\log p\left(\mathbf{x}_0, \ldots, \mathbf{x}_{M-1}, \hat{\mathbf{H}}, \hat{\boldsymbol{\Sigma}}^2\right) = -LM \log \pi - MP - M \log \det\left(\hat{\mathbf{R}}\right) - M \sum_{i=P+1}^{L} \left[\beta_i - \log \beta_i\right]. \tag{4.14}$$

donde β_i es el i-ésimo autovalor de la matriz de coherencia muestral $\hat{\mathbf{C}} = \hat{\mathbf{D}}^{-1/2}\hat{\mathbf{R}}\hat{\mathbf{D}}^{-1/2}$. Teniendo en cuenta (4.3) y (4.14), el log-GLRT se puede aproximar por

$$\log \mathscr{L} \approx \sum_{i=1}^{P} \left[\log \beta_i - \beta_i\right]. \tag{4.15}$$

Es decir, el GLRT para baja SNR es una función de los P mayores autovalores de la matriz de coherencia estimada. De manera alternativa, y para obtener otra interpretación del detector, reescribamos (4.15) como

$$\log \mathscr{L} \approx \log \prod_{i=1}^{P} \beta_i e^{-\beta_i},$$

donde se observa que el estadístico es el producto de los P mayores autovalores de $\hat{\mathbf{C}}$, cada uno ponderado por un término exponencial. Teniendo en cuenta que $\beta e^{-\beta}$ es máximo para $\beta = 1$, el log-GLRT puede interpretase como una medida de la distancia al vector de todo unos. Además, al igual que el cociente de Hadamard, el GLRT (4.15) es invariante a escalados independientes de cada serie temporal.

Lema 4.5. *El GLRT* (4.15) *es invariante a transformaciones* $\mathbf{y} = \mathbf{A}\mathbf{x}$*, siendo* $\mathbf{A}$ *una matriz diagonal invertible.*

Demostración. La demostración es equivalente a la del Lema 4.2. □

Finalmente, teniendo en cuenta que no todas la estimas ML del GLRT se han obtenido igualando a cero la verosimilitud, no se puede asegurar que la distribución asintótica del GLRT (4.15) sea χ^2 bajo la hipótesis nula. Por ello, su distribución se estudiará en la Sección 4.4 mediante simulaciones de Monte Carlo.

Relación con trabajos previos

En este apartado se relaciona el GLRT para el modelo NNWSPS derivado anteriormente con trabajos previos. En primer lugar, la aproximación del GLRT para baja SNR se obtuvo en [López-Valcarce et al., 2010] para el caso particular de $P = 1$. En dicha situación, el log-GLRT es aproximadamente

$$\log \mathscr{L} \approx -\beta_1, \tag{4.16}$$

es decir, el GLRT se obtiene como el mayor autovalor de la matriz de coherencia muestral.

Por otro lado, en [Leshem y Van der Veen, 2001a] también se considera la estimación ML de la matriz de covarianza $\mathbf{R} = \mathbf{H}\mathbf{H}^H + \mathbf{\Sigma}^2$. Sin embargo, en dicho trabajo no se obtienen las estimas ML de manera analítica, sino que se proponen diversas técnicas heurísticas para su estimación, en concreto:

- Una técnica de minimización alternada como la propuesta anteriormente. No obstante, los autores no presentan una solución para la estima ML de $\mathbf{\Sigma}^2$ y proponen emplear estimadores de mínimos cuadrados en lugar de estimadores ML.
- Técnicas clásicas de optimización numérica como, por ejemplo, el algoritmo de Gauss-Newton.
- Técnicas *ad-hoc* para valores de P concretos. Por ejemplo, para $P = 1$, se presentan diferentes algoritmos heurísticos en [Boonstra y Van der Veen, 2003].

4.2. Detección Multicanal de Procesos Estacionarios con PSD Arbitrarias

Esta sección extiende los modelos anteriores considerando series temporales con PSD arbitrarias. En particular, se resuelven los modelos NNSSNS y NNSSPS, donde las funciones de covarianza son

$$\begin{aligned} \mathcal{H}_1 &: E\left[\mathbf{x}[n]\mathbf{x}^H[n-m]\right] = \mathbf{R}_1[m], \\ \mathcal{H}_0 &: E\left[\mathbf{x}[n]\mathbf{x}^H[n-m]\right] = \mathbf{\Sigma}^2[m], \end{aligned}$$

con $\mathbf{R}_1[m] \in \mathbb{S}_+$, y $\mathbf{\Sigma}^2[m] \in \mathbb{D}_+$. Siguiendo el mismo razonamiento que en el capítulo anterior, la derivación del GLRT para este problema requiere las estimas de matrices Toeplitz por bloques, que, una vez más, se puede evitar aplicando la verosimilitud asintótica (3.27). Por lo tanto, los tests de hipótesis en el dominio frecuencial son

$$\begin{aligned} \mathcal{H}_1 &: \mathbf{x}\left(e^{j\theta}\right) \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{S}_1\left(e^{j\theta}\right)\right), \\ \mathcal{H}_0 &: \mathbf{x}\left(e^{j\theta}\right) \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{\Sigma}^2\left(e^{j\theta}\right)\right), \end{aligned} \tag{4.17}$$

donde $\mathbf{S}_1\left(e^{j\theta}\right) = \mathcal{F}\left(\mathbf{R}_1[m]\right) \in \mathbb{S}_+$, y $\mathbf{\Sigma}^2\left(e^{j\theta}\right) = \mathcal{F}\left(\mathbf{\Sigma}^2[m]\right) \in \mathbb{D}_+$.

4.2.1. Señales sin Estructura Espacial

Esta subsección aborda la derivación del GLRT para el test de hipótesis (4.17), sin incorporar estructura espacial adicional bajo $\mathcal{H}_1$. Por lo tanto, el GLR para el problema de detección viene dado por

$$\mathscr{L} = \frac{\underset{\mathbf{\Sigma}^2\left(e^{j\theta}\right)\in\mathbb{D}_+}{\text{máx}} \; p\left(\mathbf{z}_0,\ldots,\mathbf{z}_{M-1};\mathbf{\Sigma}^2\left(e^{j\theta}\right)\right)}{\underset{\mathbf{S}_1\left(e^{j\theta}\right)\in\mathbb{S}_+}{\text{máx}} \; p\left(\mathbf{z}_0,\ldots,\mathbf{z}_{M-1};\mathbf{S}_1\left(e^{j\theta}\right)\right)},$$

y, siguiendo un procedimiento al presentado en la Sección 3.2, se demuestra que las estimas ML son $\hat{\mathbf{S}}_1\left(e^{j\theta}\right) = \hat{\mathbf{S}}\left(e^{j\theta}\right)$, y

$$\hat{\mathbf{\Sigma}}^2\left(e^{j\theta}\right) = \hat{\mathbf{D}}\left(e^{j\theta}\right) = \text{diag}\left([\hat{\mathbf{S}}\left(e^{j\theta}\right)]_{1,1},\ldots,[\hat{\mathbf{S}}\left(e^{j\theta}\right)]_{L,L}\right).$$

bajo $\mathcal{H}_1$ y $\mathcal{H}_0$, respectivamente. Así, sustituyendo estas estimas en las verosimilitudes, el log-GLRT viene dado por

$$\log\mathscr{L} = \int_{-\pi}^{\pi}\log\left[\frac{\det\left(\hat{\mathbf{S}}\left(e^{j\theta}\right)\right)}{\prod_{i=1}^{L}[\hat{\mathbf{S}}\left(e^{j\theta}\right)]_{i,i}}\right]\frac{d\theta}{2\pi} = \int_{-\pi}^{\pi}\log\left[\det\left(\hat{\mathbf{C}}\left(e^{j\theta}\right)\right)\right]\frac{d\theta}{2\pi}, \tag{4.18}$$

donde la matriz de coherencia es

$$\hat{\mathbf{C}}\left(e^{j\theta}\right) = \hat{\mathbf{D}}^{-1/2}\left(e^{j\theta}\right)\hat{\mathbf{S}}\left(e^{j\theta}\right)\hat{\mathbf{D}}^{-1/2}\left(e^{j\theta}\right).$$

El test (4.18) se puede interpretar como la integral del test análogo para procesos blancos. Además, el GLRT (4.18) presenta invarianza a un tipo de transformaciones que permite obtener, mediante simulaciones, los umbrales de decisión independientemente de las PSD del ruido.

Lema 4.6. *El GLRT* (4.18) *es invariante a transformaciones* $\mathbf{y}[n] = (\mathbf{A} * \mathbf{x})[n]$*, donde* $\mathbf{A}[n]$ *es una matriz diagonal que satisface* $\left|[\mathbf{A}\left(e^{j\theta}\right)]_{ii}\right| \neq 0$*,* $\forall\theta$*, siendo* $\mathbf{A}\left(e^{j\theta}\right) = \mathcal{F}\left(\mathbf{A}[n]\right)$*.*

Demostración. La demostración es equivalente a la demostración del Lema 4.2. □

Finalmente, de manera similar a los GLRT derivados en el Capítulo 3 para procesos con PSD arbitrarias, el GLRT (4.18) se distribuye asintóticamente en N de manera Gaussiana bajo ambas hipótesis. Asimismo, bajo $\mathcal{H}_0$ su distribución asintótica en N y M es

$$-2\log\mathscr{L} \overset{a}{\sim} \mathcal{N}\left(NL^2 - NL, 2NL^2 - 2NL\right).$$

Aproximación de Baja SNR

De manera análoga a (4.6), se puede aproximar (4.18) en el caso de baja SNR como

$$\log\mathscr{L} \approx -\int_{-\pi}^{\pi}\left\|\hat{\mathbf{C}}\left(e^{j\theta}\right)\right\|_{\text{F}}^2\frac{d\theta}{2\pi}, \tag{4.19}$$

donde $\hat{\mathbf{C}}\left(e^{j\theta}\right)$ es la matriz de coherencia espectral, definida previamente.

4.2.2. Señales con Estructura Espacial

Esta subsección considera bajo $\mathcal{H}_1$ una matriz de densidades espectrales de potencia con estructura espacial adicional. Por lo tanto, suponiendo $\mathbf{S}_s\left(e^{j\theta}\right) = \mathbf{I}$, el test de hipótesis viene dado por

$$\begin{aligned} \mathcal{H}_1 &: \mathbf{x}\left(e^{j\theta}\right) \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{H}\left(e^{j\theta}\right)\mathbf{H}^H\left(e^{j\theta}\right) + \mathbf{\Sigma}^2\left(e^{j\theta}\right)\right), \\ \mathcal{H}_0 &: \mathbf{x}\left(e^{j\theta}\right) \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{\Sigma}^2\left(e^{j\theta}\right)\right), \end{aligned}$$

con $\mathbf{H}\left(e^{j\theta}\right) = \mathcal{F}\left(\mathbf{H}[n]\right) \in \mathbb{C}^{L\times P}$. De manera similar al caso de señales blancas, sólo existirá estructura espacial adicional bajo $\mathcal{H}_1$ para ciertos valores de P. Analizando el número de parámetros libres en $\mathbf{S}_1\left(e^{j\theta}\right)$, puede demostrarse que nuevamente para $P \geq L - \sqrt{L}$, dicha matriz no tiene estructura adicional y, por lo tanto, el GLRT es el derivado en la subsección anterior. Por el contrario, para $P < L - \sqrt{L}$, la estructura de rango reducido se puede explotar en la estimación ML.

Bajo $\mathcal{H}_0$, la estima ML se obtuvo en la subsección anterior y viene dada por $\hat{\mathbf{\Sigma}}\left(e^{j\theta}\right) = \hat{\mathbf{D}}\left(e^{j\theta}\right)$. Por el contrario, las estimas ML bajo $\mathcal{H}_1$ no tienen solución cerrada, como sucedía para series temporales estacionarias con PSD planas. Para este problema, también se puede aplicar el Algoritmo 4.1 frecuencia a frecuencia. No obstante, esta opción no se considera en esta Tesis, ya que su coste computacional sería excesivo; además, la aproximación de baja SNR presenta, como veremos, buenas prestaciones, incluso para SNR moderadas. Debido a esto, únicamente se considera la aproximación de baja SNR. Para obtener las estimas ML bajo $\mathcal{H}_1$, descompongamos $\mathbf{H}_{\mathbf{\Sigma}}\left(e^{j\theta}\right)\mathbf{H}_{\mathbf{\Sigma}}^H\left(e^{j\theta}\right)$ de la siguiente manera

$$\mathbf{H}_{\mathbf{\Sigma}}\left(e^{j\theta}\right)\mathbf{H}_{\mathbf{\Sigma}}^H\left(e^{j\theta}\right) = \mathbf{G}_{\mathrm{c}}\left(e^{j\theta}\right)\mathbf{\Phi}_{\mathrm{c}}^2\left(e^{j\theta}\right)\mathbf{G}_{\mathrm{c}}^H\left(e^{j\theta}\right),$$

donde $\mathbf{H}_{\mathbf{\Sigma}}\left(e^{j\theta}\right) = \mathbf{\Sigma}^{-1}\left(e^{j\theta}\right)\mathbf{H}\left(e^{j\theta}\right)$ y $\mathbf{\Phi}_{\mathrm{c}}^2\left(e^{j\theta}\right) = \mathrm{diag}\left(\phi_1^2\left(e^{j\theta}\right), \ldots, \phi_P^2\left(e^{j\theta}\right), 0, \ldots, 0\right)$. Con lo cual, las estimas ML vienen dadas por

$$\begin{aligned} &\hat{\mathbf{\Sigma}}^2\left(e^{j\theta}\right) \approx \hat{\mathbf{D}}\left(e^{j\theta}\right), \\ &\hat{\mathbf{G}}_{\mathrm{c}}\left(e^{j\theta}\right) = \mathbf{Q}\left(e^{j\theta}\right), \\ &\hat{\phi}_i^2\left(e^{j\theta}\right) = \begin{cases} \beta_i\left(e^{j\theta}\right) - 1, & i = 1, \ldots, P, \\ 0, & i = P+1, \ldots, L, \end{cases} \end{aligned}$$

donde $\beta_i\left(e^{j\theta}\right)$ es el i-ésimo autovalor de $\hat{\mathbf{C}}\left(e^{j\theta}\right)$ y $\mathbf{Q}\left(e^{j\theta}\right)$ su matriz de autovectores. Teniendo en cuenta las estimas ML presentadas anteriormente, el log-GLRT para baja SNR es

$$\log \mathscr{L} \approx \int_{-\pi}^{\pi} \sum_{i=1}^{P} \left[\log \beta_i\left(e^{j\theta}\right) - \beta_i\left(e^{j\theta}\right)\right] \frac{d\theta}{2\pi}. \tag{4.20}$$

Este GLRT se puede interpretar como la integral del GLRT análogo para series temporales blancas, dado por (4.15). Asimismo, el GLRT (4.20) presenta las mismas invarianzas que el GLRT (4.18).

Lema 4.7. *El GLRT* (4.20) *es invariante a transformaciones* $\mathbf{y}[n] = (\mathbf{A} * \mathbf{x})[n]$, *donde* $\mathbf{A}[n]$ *es una matriz diagonal que satisface* $\left|[\mathbf{A}\left(e^{j\theta}\right)]_{ii}\right| \neq 0$, *siendo* $\mathbf{A}\left(e^{j\theta}\right) = \mathcal{F}\left(\mathbf{A}[n]\right)$.

Demostración. La demostración es equivalente a la del Lema 4.2. □

El GLRT (4.20) también tiene una distribución normal bajo ambas hipótesis, siendo necesario obtener las medias y las varianzas mediante simulaciones. Asimismo, aplicando el Lema 4.7, dichos parámetros no dependen de las PSD de los ruidos.

4.3. Detección de Correlación Espacial en Series Temporales No Estacionarias

Esta sección considera el problema de detección de correlación espacial de la serie temporal multivariada, bajo la suposición de procesos no estacionarios. Por lo tanto, las funciones de covarianza corresponden al modelo NNNSNS

$$\begin{aligned}\mathcal{H}_1 : E\left[\mathbf{x}[n]\mathbf{x}^H[n-m]\right] &= \mathbf{R}_1[n, n-m],\\ \mathcal{H}_0 : E\left[\mathbf{x}[n]\mathbf{x}^H[n-m]\right] &= \mathbf{\Sigma}^2[n, n-m].\end{aligned}$$

Siguiendo un desarrollo análogo al de la Sección 3.3, el problema de detección se puede formular como

$$\begin{aligned}\mathcal{H}_1 : \mathbf{g} &\sim \mathcal{CN}\left(\mathbf{0}_{LN}, \mathbf{R}_1\right),\\ \mathcal{H}_0 : \mathbf{g} &\sim \mathcal{CN}\left(\mathbf{0}_{LN}, \mathbf{R}_0\right),\end{aligned}$$

donde $\mathbf{R}_1 \in \mathbb{S}_+$, $\mathbf{R}_0 \in \mathbb{D}_{\mathrm{B}+}$ son dos matrices de covarianza desconocidas, $\mathbb{S}_+$ es el conjunto de matrices de covarianza sin estructura espacial ni temporal adicional (sólo son definidas positivas) y $\mathbb{D}_{\mathrm{B}+}$ es el conjunto de matrices de covarianza diagonales por bloques, es decir, $\mathbf{R}_{ik} = \mathbf{0}_N,\ i \neq k$. Resumiendo, bajo la hipótesis nula las series temporales están espacialmente incorreladas, pero pueden tener correlación temporal.

El GLRT para $\mathcal{H}_0 : \mathbf{R} \in \mathbb{D}_{\mathrm{B}+}$ vs. $\mathcal{H}_1 : \mathbf{R} \in \mathbb{S}_+$ se basa en el GLR

$$\mathscr{L} = \frac{\operatorname*{máx}_{\mathbf{R}\in\mathbb{D}_{\mathrm{B}+}} p\left(\mathbf{g}_0, \ldots, \mathbf{g}_{M-1}; \mathbf{R}\right)}{\operatorname*{máx}_{\mathbf{R}\in\mathbb{S}_+} p\left(\mathbf{g}_0, \ldots, \mathbf{g}_{M-1}; \mathbf{R}\right)},$$

y, en los siguientes lemas, se obtienen las estimas ML de $\mathbf{R}_1$ y $\mathbf{R}_0$ necesarias para derivar el GLRT.

Lema 4.8. *La estima de la matriz de covarianza bajo $\mathcal{H}_0$ viene dada por*

$$\hat{\mathbf{R}}_0 = \begin{bmatrix} \hat{\mathbf{R}}_{11} & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \hat{\mathbf{R}}_{22} & \ldots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \ldots & \hat{\mathbf{R}}_{LL} \end{bmatrix}, \tag{4.21}$$

esto es, una matriz compuesta por los bloques de la diagonal de $\hat{\mathbf{R}}$.

Demostración. Bajo $\mathcal{H}_0$, dada la estructura diagonal por bloques, es posible factorizar la verosimilitud como

$$p\left(\mathbf{g}_0, \ldots, \mathbf{g}_{M-1}; \mathbf{R}\right) = \prod_{i=1}^{L} p\left([\mathbf{g}_0]_i, \ldots, [\mathbf{g}_{M-1}]_i; \mathbf{R}_{ii}\right),$$

donde $\mathbf{g}_l = \left[[\mathbf{g}_l]_1^T, \ldots, [\mathbf{g}_l]_L^T\right]^T$. Teniendo en cuenta esta descomposición, se demuestra que $p\left([\mathbf{g}_0]_i, \ldots, [\mathbf{g}_{M-1}]_i; \mathbf{R}_{ii}\right)$ puede ser optimizada independientemente para cada $\mathbf{R}_{ii}$. Por otro lado, dado que $\mathbf{R}_{ii}$ no tiene ninguna estructura adicional (ni espacial ni temporal), su estima ML será [Magnus y Neudecker, 1999]

$$\mathbf{R}_{ii} = \hat{\mathbf{R}}_{ii}, \qquad i = 1, \ldots, L.$$

La demostración concluye sustituyendo las estimas ML individuales en la matriz $\mathbf{R}_0$. □

Lema 4.9. *La estima de la matriz de covarianza bajo $\mathcal{H}_1$ viene dada por*

$$\hat{\mathbf{R}}_1 = \hat{\mathbf{R}}. \tag{4.22}$$

Demostración. La demostración se puede encontrar en [Magnus y Neudecker, 1999]. □

Teniendo en cuenta las estimas ML bajo ambas hipótesis, es decir (4.21) y (4.22), el GLRT es

$$\mathscr{L} = \frac{\det\left(\hat{\mathbf{R}}\right)}{\prod_{i=1}^{L} \det\left(\hat{\mathbf{R}}_{ii}\right)} = \det(\hat{\mathbf{C}}), \tag{4.23}$$

donde $\hat{\mathbf{D}} = \hat{\mathbf{R}}_0$ y $\hat{\mathbf{C}} = \hat{\mathbf{D}}^{-1/2}\hat{\mathbf{R}}\hat{\mathbf{D}}^{-1/2}$ es la matriz de coherencia o matriz de relación señal a ruido si se considera $\mathcal{H}_0$ como la hipótesis de sólo ruido. El GLRT (4.23) puede interpretarse como una generalización del cociente de Hadamard. Además, si tenemos en cuenta la expresión (4.5) para el GLRT del modelo NNWSNS, se puede ver que ambos detectores tienen la misma expresión, pero con diferentes definiciones para la matriz de coherencia. Finalmente, el GLRT (4.23) es invariante a transformaciones lineales de las series temporales.

Lema 4.10. *El GLRT* (4.23) *es invariante a transformaciones* $\mathbf{y}_i = \mathbf{A}_i\mathbf{x}_i,\ i = 1, \ldots, L,$ *suponiendo matrices* $\mathbf{A}_i,\ i = 1, \ldots, L,$ *invertibles.*

Demostración. Definiendo la matriz $\mathbf{A}$

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2 & \ldots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \ldots & \mathbf{A}_L \end{bmatrix},$$

es fácil demostrar que las matrices de covarianza muestrales de las señales $\mathbf{y}_i$ vienen dadas por $\hat{\mathbf{R}}_0^{(y)} = \mathbf{A}\hat{\mathbf{R}}_0^{(x)}\mathbf{A}^H$ y $\hat{\mathbf{R}}_1^{(y)} = \mathbf{A}\hat{\mathbf{R}}_1^{(x)}\mathbf{A}^H$. Sustituyendo estas matrices en (4.23) obtenemos el resultado deseado. □

Relación con Trabajos Previos

Recientemente, en [Pérez-Neira et al., 2009] se ha demostrado que el GLRT (4.23) está relacionado con la distancia geodésica entre las matrices $\hat{\mathbf{R}}_0$ y $\hat{\mathbf{R}}_1$ sobre la variedad analítica (*manifold*) de las matrices definidas positivas. Para ilustrar esta relación, definamos la distancia geodésica [Boutros et al., 2006] como la longitud del segmento

$$\breve{\mathbf{R}}(t) = \hat{\mathbf{R}}_0^{1/2} \exp\left[t\left(\hat{\mathbf{R}}_0^{-1/2}\log(\hat{\mathbf{R}}_1)\hat{\mathbf{R}}_0^{-1/2}\right)\right]\hat{\mathbf{R}}_0^{1/2},$$

donde $t \in [0, 1]$. En [Boutros et al., 2006], se demuestra que la distancia geodésica puede expresarse como $d^2_{\text{geo}} = \sum_{i=1}^{LN} \log^2 \beta_i$, donde β_i son los autovalores de $\hat{\mathbf{C}}$. Por otro lado, el logaritmo del GLRT en (4.23) es $\log \mathscr{L} = \sum_{i=1}^{LN} \log \beta_i$, demostrando que ambas medidas son funciones del logaritmo de los autovalores de la matriz de coherencia $\hat{\mathbf{C}}$.

4.3.1. GLRT en el Dominio de la Frecuencia

En esta subsección, usando la propiedad de invarianza a transformaciones lineales, se reescribe el GLRT (4.23) y se presenta una aproximación del mismo. Dada la similitud con la Sección 3.3.1, sólo se presenta el resultado, omitiendo su derivación.

El GLRT (4.23) puede descomponerse en un término correspondiente a la componente estacionaria de la serie y uno correspondiente a la componente no estacionaria, es decir,

$$\log \mathscr{L} = \log\det(\tilde{\mathbf{C}}_{\mathrm{NS}}) + \log\det(\tilde{\mathbf{C}}_{\mathrm{WSS}}),$$

donde $\tilde{\mathbf{C}}_{\mathrm{NS}} = \tilde{\mathbf{C}}_{\mathrm{WSS}}^{-1/2}\tilde{\mathbf{C}}\tilde{\mathbf{C}}_{\mathrm{WSS}}^{-1/2}$, $\tilde{\mathbf{C}}_{\mathrm{WSS}}$ es una matriz diagonal por bloques, dados por $\hat{\mathbf{C}}\left(e^{j\theta_i}\right)$, $i = 1, \ldots, N$, y

$$\tilde{\mathbf{C}} = \begin{bmatrix} \hat{\mathbf{C}}\left(e^{j\theta_0}\right) & \cdots & \hat{\mathbf{C}}\left(e^{j\theta_0}, e^{j\theta_{N-1}}\right) \\ \vdots & \ddots & \vdots \\ \hat{\mathbf{C}}\left(e^{j\theta_0}, e^{j\theta_{N-1}}\right) & \cdots & \hat{\mathbf{C}}\left(e^{j\theta_{N-1}}\right) \end{bmatrix},$$

siendo

$$\left[\hat{\mathbf{C}}\left(e^{j\theta_i}, e^{j\theta_k}\right)\right]_{l,m} = \mathbf{f}^H\left(e^{j\theta_i}\right)\hat{\mathbf{R}}_{ll}^{-1/2}\hat{\mathbf{R}}_{lm}\hat{\mathbf{R}}_{mm}^{-1/2}\mathbf{f}\left(e^{j\theta_k}\right).$$

Por otro lado, bajo el supuesto de estacionariedad, el log-GLRT se simplifica a

$$\frac{1}{N}\log\mathscr{L} = \frac{1}{N}\log\mathscr{L}_{\mathrm{WSS}} \underset{N\to\infty}{\longrightarrow} \int_{-\pi}^{\pi} \log\det\left(\hat{\mathbf{C}}\left(e^{j\theta}\right)\right)\frac{d\theta}{2\pi},$$

obteniéndose conclusiones similares a las presentadas en la Sección 3.3.1.

4.4. Resultados Numéricos

Esta sección presenta el comportamiento de los detectores propuestos mediante simulaciones de Monte Carlo. Concretamente, se evalúan las prestaciones de los siguientes detectores:

- Hadamard: GLRT para el modelo NNWSNS, basado en el cociente de Hadamard, cuyo estadístico es
$$\log\mathscr{L} = \log\det\left(\hat{\mathbf{C}}\right).$$
- Alt-GLRT: GLRT para el modelo NNWSPS basado en la técnica de minimización alternada del Algoritmo 4.1. La inicialización considerada en las simulaciones es $\hat{\mathbf{\Sigma}}_{(0)} = \hat{\mathbf{D}}$.
- Asin-GLRT: GLRT para el modelo NNWSPS basado en la aproximación de baja SNR, siendo su estadístico
$$\log\mathscr{L} \approx \sum_{i=1}^{P}\left[\log\beta_i - \beta_i\right],$$
donde β_i son los autovalores de la matriz de coherencia muestral.
- *Asin-GLRT ($P = 1$)* GLRT para el modelo NNWSPS con $P = 1$, propuesto en [López-Valcarce et al., 2010], y basado en el siguiente estadístico
$$\mathscr{L} \approx -\beta_1.$$

- Norma Frob.: Aproximación de baja correlación del GLRT (4.4), basado en la norma Frobenius
$$\log \mathscr{L} \approx -\left\|\hat{\mathbf{C}}\right\|_{\mathrm{F}}^{2}.$$
- Freq. Hadamard: GLRT para el modelo NNSSNS, dado por
$$\log \mathscr{L} = \int_{-\pi}^{\pi} \log \left[\det\left(\hat{\mathbf{C}}\left(e^{j\theta}\right)\right)\right] \frac{d\theta}{2\pi}.$$
- Freq. Asin-GLRT: GLRT para el modelo NNSSPS bajo la suposición de baja SNR, basado en el siguiente GLR aproximado
$$\log \mathscr{L} \approx \int_{-\pi}^{\pi} \sum_{i=1}^{P} \left[\log \beta_i\left(e^{j\theta}\right) - \beta_i\left(e^{j\theta}\right)\right] \frac{d\theta}{2\pi},$$
donde $\beta_i\left(e^{j\theta}\right)$ son los autovalores de $\hat{\mathbf{C}}\left(e^{j\theta}\right)$.
- Freq. Norma Frob.: Aproximación de baja correlación del GLRT (4.18) basado en la norma Frobenius
$$\log \mathscr{L} \approx -\int_{-\pi}^{\pi} \left\|\hat{\mathbf{C}}\left(e^{j\theta}\right)\right\|_{\mathrm{F}}^{2} \frac{d\theta}{2\pi}.$$
- Hadamard Gen.: GLRT para el modelo NNNSNS, considerado como una generalización del cociente de Hadamard, siendo su estadístico
$$\mathscr{L} = \frac{\det\left(\hat{\mathbf{R}}\right)}{\prod_{i=1}^{L} \det\left(\hat{\mathbf{R}}_{ii}\right)}.$$

4.4.1. Efecto del Rango P

El primer experimento considerado analiza el efecto de P en el modelo NNWSPS sobre los diferente detectores. De manera similar al capítulo anterior, se suponen señales y ruidos Gaussianos y canales Rayleigh normalizados para obtener una SNR instantánea constante, definida para ruidos no-IID como
$$\text{SNR (dB)} = 10 \log_{10} \frac{\operatorname{tr}(\mathbf{H}\mathbf{H}^{H})}{\operatorname{tr}(\mathbf{\Sigma}^{2})}.$$
En concreto, los parámetros del experimento son: $L = 6$ sensores, SNR $= -8$ dB, $M = 180$ observaciones y los niveles de ruido relativos en cada sensor son 0.5, -1, 0, 1, -0.5 y 0 dB. En la Figura 4.1 se dibuja la probabilidad de pérdida para una probabilidad de falsa alarma $p_{\mathrm{FA}} = 0.001$, mostrando, una vez más, las mejores prestaciones del detector que explota completamente la estructura espacial, siendo para este modelo el basado en la optimización alternada (Algoritmo 4.1). Asimismo, el cociente de Hadamard presenta unas prestaciones prácticamente equivalentes al GLRT para P grande, como se demostró en el Lema 4.4. Por otro lado, nótese que el GLRT para baja SNR (4.15) presenta prestaciones similares a las del detector basado en la optimización alternada, disminuyendo la diferencia entre ambos cuando P crece. La razón por la que ambos detectores presentan prestaciones parecidas es

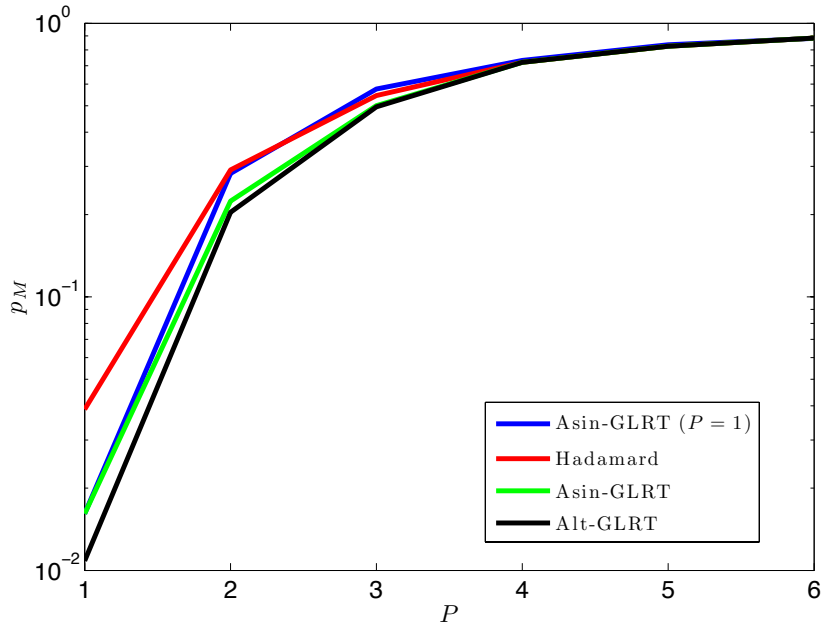

Figura 4.1: Probabilidad de pérdida frente a P con ruidos no-IID para el siguiente experimento: $L = 6$ sensores, SNR (dB) $= -8$ dB y $M = 180$ muestras. Los niveles de ruido relativos en cada antena son 0.5, -1, 0, 1, -0.5 y 0 dB.

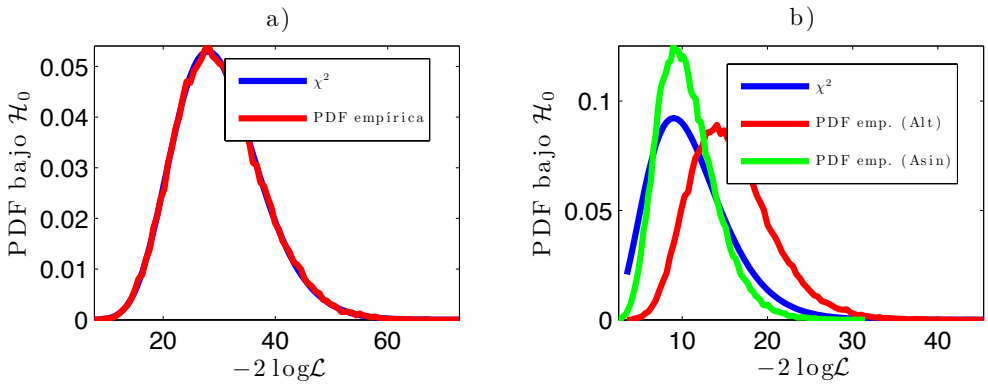

Figura 4.2: PDF estimada y aproximación χ^2 para: a) el test de Hadamard y b) el GLRT para señales blancas con rango P y ruidos no-IID.

la baja SNR del experimento. Además, cuando P crece, dicha SNR se reparte entre más dimensiones, por lo tanto la SNR efectiva, o por dimensión, se reduce, favoreciendo aún más la aproximación de baja SNR. Finalmente, al igual que en el caso de ruidos IID, el GLRT que siempre asume $P = 1$ presenta prestaciones pobres para otros valores de P.

Para este mismo experimento, seleccionando $M = 1000$ y SNR $= -2$ dB, se estima la PDF del estadístico bajo $\mathcal{H}_0$. Para el caso de $P = 6$ el GLRT es el cociente de Hadamard, pudiéndose observar en la Figura 4.2a el buen comportamiento de la aproximación $\chi^2_{L^2-L}$. Por el contrario, para $P = 1$ el GLRT no cumple las condiciones del teorema de Wilks, obteniendo

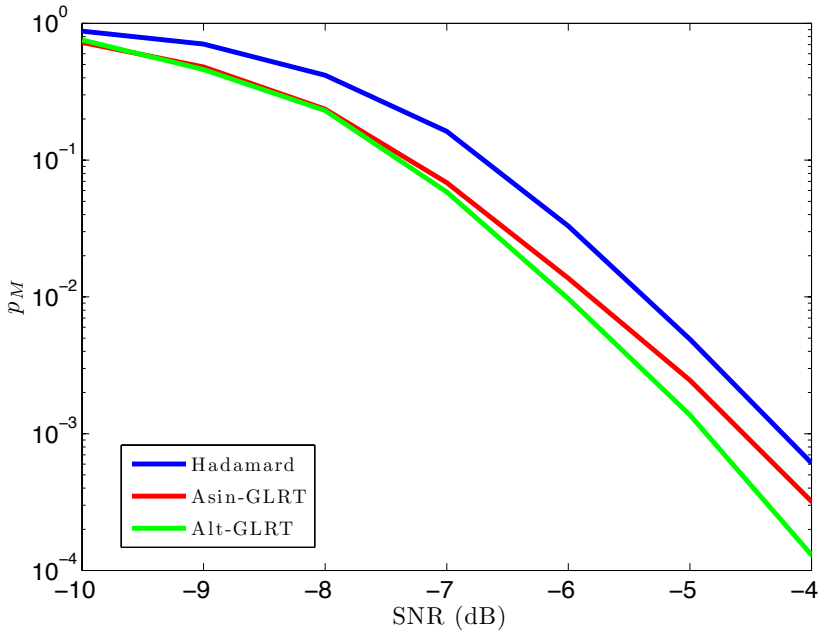

Figura 4.3: Probabilidad de pérdida frente a la SNR de los diferentes detectores. Los parámetros de este experimento son: $L = 6$ sensores o antenas, $P = 1$ y $M = 96$ y los mismos niveles de ruido relativos del apartado anterior.

una PDF estimada que no se puede ajustar por una χ^2_{2L-1}, como se puede observar en la Figura 4.2b.

4.4.2. Comportamiento del GLRT para Baja SNR

El experimento anterior ha ilustrado el comportamiento similar del GLRT para ruidos no-IID y rango P basado en la optimización alternada y su aproximación de baja SNR. Sin embargo, en dicho experimento la SNR considerada era pequeña. En esta subsección, se analizan las prestaciones del GLRT (4.15) en un rango de SNR moderado, y se estudiará si el mayor coste computacional del detector presentado en el Algoritmo 4.1 está justificado en un caso práctico. Para ello, en la Figura 4.3 se ha obtenido la probabilidad de pérdida para $P_{\text{FA}} = 0.001$ y diferentes valores de SNR. Los parámetros del experimento son los siguientes: $L = 6$ sensores o antenas, $P = 1$, $M = 96$ y los mismos niveles de ruido relativos del apartado anterior. La Figura 4.3 ilustra el mejor comportamiento del detector basado en la optimización alternada para todos los valores de SNR, lo cual, como ya se ha mencionado previamente, es esperable por ser el único detector que explota completamente la estructura del problema. Asimismo, podemos observar que el GLRT para baja SNR (4.15) se comporta de manera similar al detector basado en la minimización alternada para bajas SNR y empeora cuando la SNR aumenta. Finalmente, nótese que los detectores que explotan la estructura espacial se comportan, una vez más, mejor que el detector que ignora dicha estructura (Hadamard).

4.4.3. Impacto de la Distribución del Ruido

Esta subsección evalúa el impacto de asumir incorrectamente la distribución del ruido. Concretamente, se analiza el comportamiento de los GLRT derivados bajo ambas suposiciones

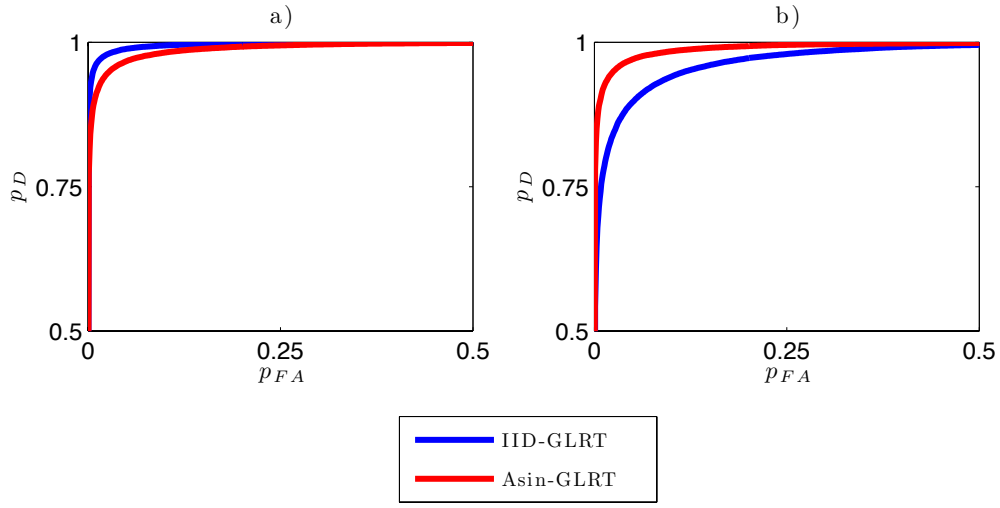

Figura 4.4: Curva ROC de los GLRT propuestos bajo: a) ruidos IID y b) ruidos no-IID. En este ejemplo los parámetros del experimento son: SNR $= -7$ dB, $L = 6$ sensores o antenas, $P = 2$ y $M = 128$.

respecto al ruido en escenarios con ruidos IID y no-IID. Para ello, se considera el siguiente experimento: SNR $= -7$ dB, $L = 6$ sensores o antenas, $P = 2$ y $M = 128$. La Figura 4.4 muestra las curvas ROC bajo ambos supuestos, donde se observa la pérdida de prestaciones de los GLRT cuando suponen incorrectamente la distribución del ruido. No obstante, es importante destacar la mayor pérdida de prestaciones del GLRT para ruidos IID cuando, en realidad, son no-IID (Fig. 4.4b) que en la situación opuesta (Fig. 4.4a). Finalmente, los resultados de esta subsección demuestran que, desde un punto de vista práctico, es preferible usar los detectores derivados bajo la suposición de ruidos no-IID.

4.4.4. Detección de Series Temporales con PSD Arbitrarias

Esta subsección analiza las prestaciones de los diferentes GLRT para series temporales con PSD arbitrarias, comparando su comportamiento con los GLRT para procesos blancos.

El primer ejemplo considera un experimento con señales generadas por el modelo (3.47), donde los ruidos son ahora procesos MA distintos y con diferentes potencias. Los parámetros del experimento considerado son: $L = 6$ sensores, $P = 1$, $M = 6$ realizaciones de longitud $N = 100$, $T = q_v = 8$, la SNR $= -10$ dB y los niveles de ruidos relativos son 0.5, -1, 0, 1, -0.5 y 0 dB. Para esta situación el GLRT viene dado por[3] (4.20), y la Figura 4.5 compara sus prestaciones con las del detector (4.15). Como se puede observar en la figura, las prestaciones del detector que explota la estructura temporal son mucho mejores. Asimismo, en la Figura 4.5 se muestran también las curvas ROC para un experimento con los mismos parámetros que el anterior, salvo $P = 6$ y SNR $= -3$ dB. En este ejemplo se comparan los detectores (4.4) y (4.18), obteniendo conclusiones similares.

La Figura 4.6 muestra para PSD fijas y $M = 1000$, la exactitud de la aproximación Gaussiana para los dos ejemplos anteriores. La media y la varianza se han obtenido mediante

[3]Las integrales se calculan mediante la regla del trapecio compuesta con $K = 256$ frecuencias equiespaciadas.

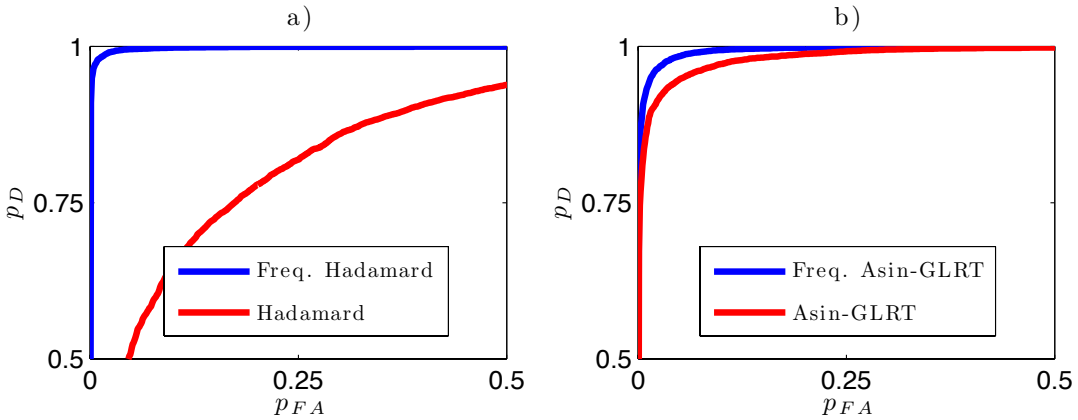

Figura 4.5: Curva ROC de los diferentes detectores para señales con PSD arbitrarias contaminadas con ruidos no-IID. Los parámetros del experimento son: $L = 6$ sensores, $M = 6$ realizaciones de longitud $N = 100$, $T = q_v = 8$ y : a) $P = 1$, SNR $= -10$ dB y b) $P = 6$, SNR $= -3$ dB

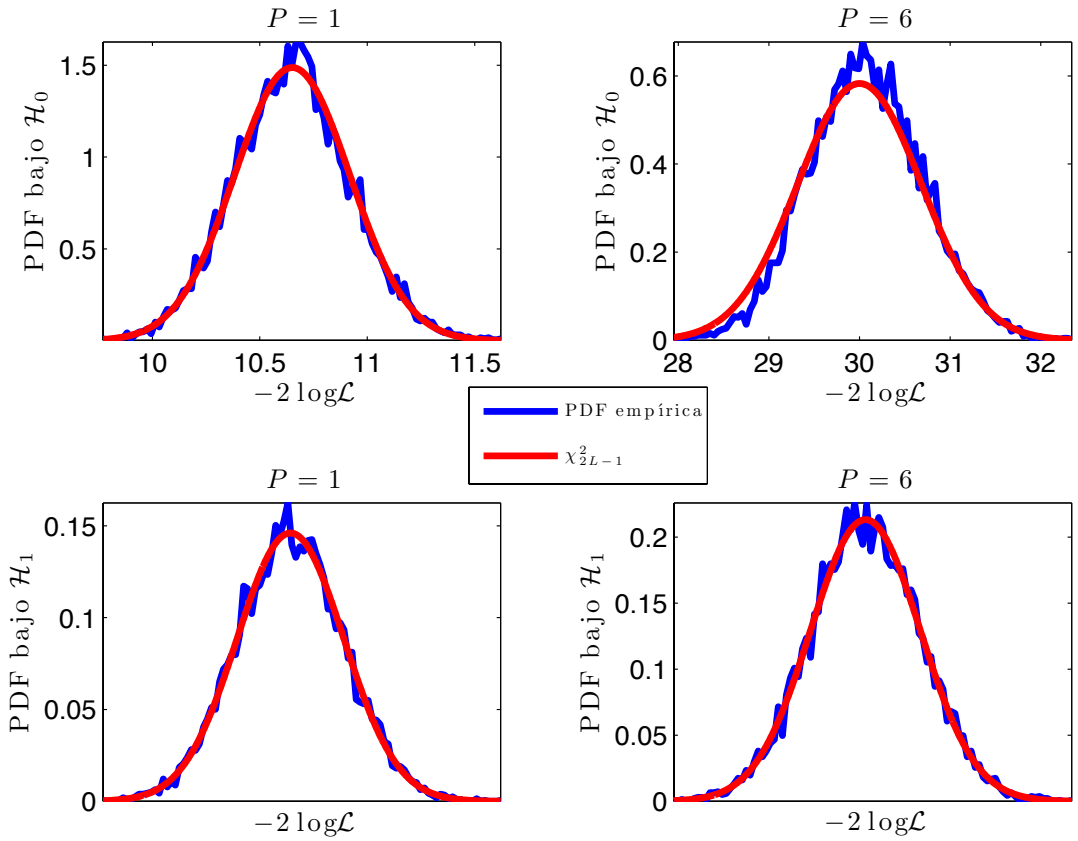

Figura 4.6: PDF estimada y aproximación Gaussiana de los GLRT asintóticos para ruidos no-IID en un experimento con $L = 6$, $M = 6$ y $N = 100$.

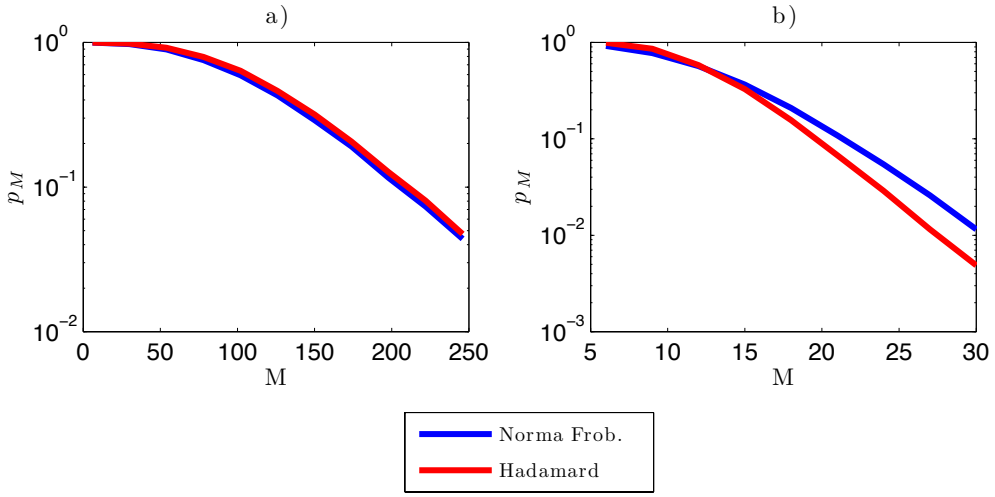

Figura 4.7: Probabilidad de pérdida, para $p_{\text{FA}} = 10^{-3}$, del detector basado en el cociente de Hadamard y su aproximación, para señales blancas con: a) SNR $= -5$ dB y SNR $= 5$ dB.

simulaciones, a excepción de la distribución bajo $\mathcal{H}_0$ del ejemplo con $P = 6$, donde se han usado los valores teóricos derivados previamente.

4.4.5. Determinante frente a la Norma Frobenius

Esta subsección analiza la aproximación de baja SNR de los detectores basados en los cocientes de Hadamard. En el primer ejemplo se consideran señales temporalmente blancas para dos correlaciones espaciales diferentes: baja correlación (SNR $= -5$ dB) y alta correlación (SNR $= 5$ dB). La Figura 4.7 muestra la probabilidad de pérdida, para $p_{\text{FA}} = 10^{-3}$, de un experimento con $L = 6$ y $P = 6$, en función de M. En la situación de baja correlación (Figura 4.7a), el detector basado en la norma Frobenius presenta unos resultados ligeramente mejores para todos los valores de M analizados. Sin embargo, para alta correlación (Figura 4.7b) el GLRT presenta mejor comportamiento para un número suficiente de observaciones, mientras que para un número de muestras reducido su aproximación basada en la norma Frobenius es mejor. Estos resultados demuestran que el detector basado en la norma Frobenius es más robusto cuando el número de observaciones disponible es pequeño.

Adicionalmente, la Figura 4.8 muestra la probabilidad de pérdida, para $p_{\text{FA}} = 10^{-4}$, de los detectores para procesos coloreados en un experimento con los siguientes parámetros: $L = 3$, $P = 3$, SNR $= -3$ dB o SNR $= 12$ dB y los niveles de ruido relativos son -2.5, 0 y 1.5. En este ejemplo, las conclusiones obtenidas son similares que para procesos blancos.

4.4.6. Señales No Estacionarias

Esta subsección analiza el GLRT para señales no estacionarias (4.23), comparando sus prestaciones con las de su aproximación. En este ejemplo, las observaciones se generan igual que en la Sección 3.4.5 con la excepción de los ruidos, que ahora son procesos MA distintos con diferentes potencias.

En la Figura 4.9a se pueden ver las curvas ROC de los diferentes detectores en un experimento con diferentes valores de N y M, para $L = 3$, $P = 3$, $T = q_v = 4$, $d_{\text{NS}} = 0.9$ y

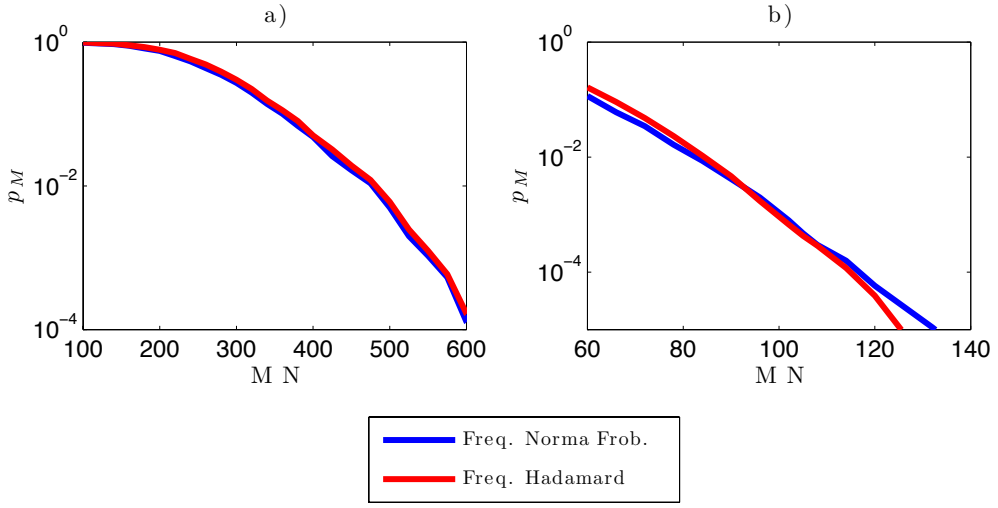

Figura 4.8: Probabilidad de pérdida, para $p_{\text{FA}} = 10^{-4}$, del detector basado en el cociente de Hadamard y su aproximación, para señales coloreadas con: a) SNR $= -5$ dB y SNR $= 5$ dB.

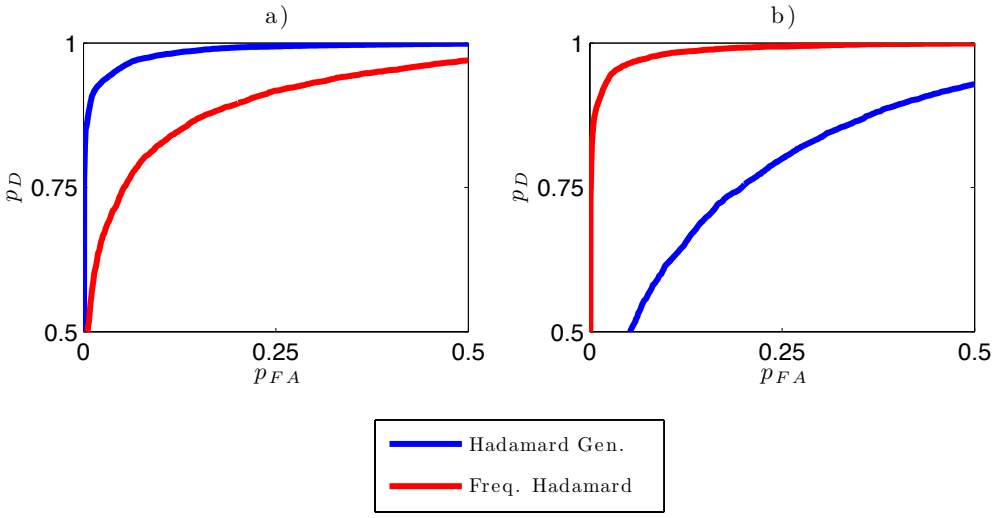

Figura 4.9: Curva ROC de los diferentes detectores para señales no estacionarias considerando diferentes tamaños de ventana: a) $N = 4$ y b) $N = 30$. En este ejemplo se han considerado los siguientes parámetros: $L = 3$ sensores, $T = q_v = 4$, $d_{\text{NS}} = 0.9$ y $\nu = 10/N$.

$\nu = 10/N$. En primer lugar se considera $N = 4$, $M = 50$ y SNR $= 0$ dB. Para este caso, el GLRT en el dominio del tiempo (4.23) presenta mejores prestaciones que su aproximación frecuencial. Sin embargo, para $N = 30$ ($M = 100$ y SNR $= -7$ dB) la aproximación frecuencial presenta mejor comportamiento que el GLRT (ver Figura 4.9b). Esto es debido a que el detector en el dominio de la frecuencia estima matrices de tamaño $L \times L$, por el contrario, el detector en el dominio del tiempo trabaja con matrices mucho mayores, $LN \times LN$.

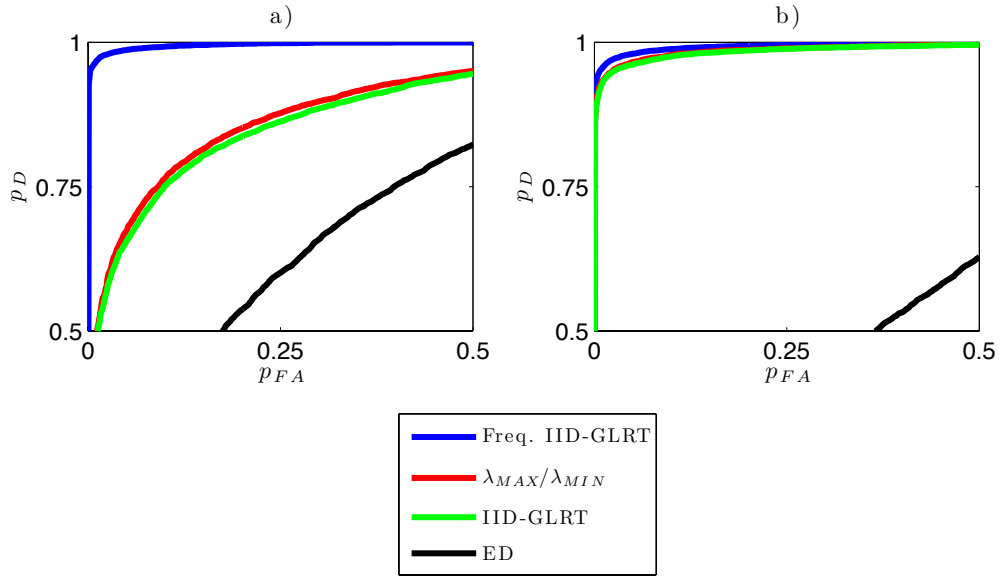

Figura 4.10: Curva ROC para los diferentes detectores para señales en un ejemplo de radio cognitiva con ruidos no-IID en dos entornos de propagación: a) área urbana y b) área rural.

4.4.7. Aplicación a Radio Cognitiva

Esta subsección presenta la aplicación de los detectores derivados en el capítulo a un problema de sensado espectral para radio cognitiva, considerando un escenario equivalente al presentado en la Sección 3.4.6, con la excepción de los ruidos que ahora son procesos MA distintos. En concreto, se considera un experimento con los siguientes parámetros: $L = 3$ antenas, $P = 1$, SNR $= -8$ dB, $M = 10$ realizaciones de longitud $N = 100$ y los siguientes niveles de potencia relativos: 0.5, -1 y 0 dB.

La Figura 4.10 muestra las curvas ROC de los GLRT para los modelos NNWSPS y NNSSPS. Asimismo, se presentan también las prestaciones del detector de energía (denotado como ED) y del detector propuesto en [Zeng y Liang, 2009], modificado para considerar ruidos no-IID[4] (denotado como $\lambda_{\text{MAX}}/\lambda_{\text{MIN}}$). Una vez más, las mejores prestaciones las obtiene el detector que explota completamente la estructura espacio-temporal.

4.5. Conclusiones

Este capítulo ha considerado el problema de detección de series temporales multivariadas contaminadas por ruidos independientes, pero no idénticamente distribuidos (no-IID). Para resolver los test de hipótesis se ha usado, una vez más, el detector basado en el cociente de verosimilitudes generalizado (GLRT).

En primer lugar se ha estudiado el problema con procesos estacionarios en sentido amplio (WSS) y densidades espectrales de potencia (PSD) planas, bajo dos suposiciones sobre su estructura espacial. Para señales sin estructura espacial adicional, se puede obtener una solución cerrada para el GLRT. Por el contrario, cuando la matriz de covarianza bajo $\mathcal{H}_1$ es la suma de una matriz deficiente en rango más una matriz diagonal, la estimación de máxima

[4]Para este caso, el detector se basa en el número de condición de la matriz de coherencia estimada.

verosimilitud (ML) de la matriz diagonal es un problema no convexo, que se ha resuelto de dos maneras diferentes. En primer lugar se ha propuesto una técnica de minimización alternada que, en cada paso, resuelve un problema convexo. No obstante, motivado por el alto coste computacional de la técnica iterativa, se ha derivado una solución aproximada para baja relación señal a ruido (SNR), con un coste computacional inferior. Las prestaciones de los citados detectores se evaluaron mediante simulaciones de Monte Carlo, pudiéndose comprobar la ventaja obtenida al explotar correctamente la estructura espacial. Asimismo, se ha comprobado que la aproximación de baja SNR presenta un comportamiento similar al de la técnica basada en la minimización alternada para SNR bajas y moderadas, comportándose ligeramente peor para SNR altas. Por último, los detectores derivados para ruidos no-IID son más robustos a suposiciones erróneas sobre la distribución del ruido, como demuestran las simulaciones.

Posteriormente, se han generalizado los resultados anteriores considerando series temporales con PSD arbitrarias. Una vez más, se ha aplicado la verosimilitud asintótica para derivar los GLRT de estos modelos, dados por la integral de los GLRT análogos para procesos blancos. Asimismo, los resultados numéricos han demostrado el mejor comportamiento de los GLRT que explotan la estructura espacial.

Finalmente, se han considerado series temporales no estacionarias sin estructura espacial adicional. Para este modelo, no es necesario recurrir a la verosimilitud asintótica, obteniéndose un GLRT que se puede considerar una generalización del cociente de Hadamard. Asimismo, se ha presentado una aproximación del GLRT en el dominio frecuencial y se han evaluado las prestaciones de ambos mediante simulaciones de Monte Carlo. En concreto, para tamaños de ventana pequeños, el GLRT presenta mejor comportamiento que la aproximación frecuencial. No obstante, cuando el tamaño de la ventana es elevado las prestaciones del GLRT son inferiores a las de la aproximación, debido a que el GLRT necesita estimar una matriz de mayores dimensiones, a partir de un conjunto reducido de observaciones.

Capítulo 5

Nuevas Medidas Multicanal de Correlación en el Dominio Frecuencial

En capítulos anteriores se ha estudiado el problema de detección multicanal de series temporales bajo diferentes suposiciones acerca de la distribución del ruido. Concretamente, para series temporales con densidades espectrales de potencia (PSD) arbitrarias, todos los GLRT derivados son funciones de los autovalores de la matriz de densidades espectral de potencia y de la matriz de espectros de coherencia estimados, para ruidos IID y no-IID, respectivamente. El caso de ruidos no-IID es especialmente interesante debido a que el GLRT para $L = 2$ series temporales puede reescribirse como una función del espectro de coherencia. Esto permite proponer generalizaciones basadas en funciones de los autovalores de la matriz de coherencia. Dichas funciones tienen que mantener el ordenamiento, es decir, ser funciones Schur-convexas del vector que contiene todos los autovalores de la matriz de coherencia.

Una vez definido el espectro de coherencia generalizado (o una nueva familia de funciones para generalizar el espectro de coherencia) se presentan y analizan sus propiedades. Asimismo, se proponen diferentes estimadores para el espectro de coherencia generalizado. Básicamente, las técnicas de estimación se agrupan en dos clases. La primera de ellas se basa en estimar los espectros de coherencia por parejas con técnicas propuestas en la literatura, y sustituir dichas estimas en la definición. La segunda, que como se verá presenta mejores prestaciones, se basa en el análisis de correlaciones canónicas (CCA). Finalmente, aplicando el estimador basado en CCA, se propone un nuevo detector para el problema de detección multicanal con ruidos no-IID, que presenta un comportamiento diferente del GLRT.

Este capítulo está organizado de la siguiente manera: la Sección 5.1 presenta el espectro de coherencia clásico, así como sus propiedades y estimadores. A continuación, en la Sección 5.2 se define el espectro de coherencia generalizado (GCS), se presentan sus propiedades y se proponen estimadores. Posteriormente, la Sección 5.3 discute la relación del GCS con la información mutua de L series temporales y se presenta su relación con los GLRT derivados en capítulos anteriores. Por último, en la Sección 5.4 se presentan resultados numéricos para ilustrar el comportamiento de los diferentes GCS, analizar las prestaciones de los distintos estimadores y mostrar los efectos resultantes de emplear mejores estimadores del GCS en el problema de detección multicanal. Finalmente, se presentan las principales conclusiones obtenidas en la Sección 5.5.

5.1. El Espectro de Coherencia

El espectro de coherencia (CS) es una medida frecuencial de la relación estadística lineal entre dos procesos estocásticos [Brillinger, 1981, Stoica y Moses, 2005], que suele ser considerado como el coeficiente de correlación en el dominio de la frecuencia. Por otro lado, para procesos Gaussianos, el espectro de coherencia proporciona una medida de la información mutua [Scharf y Thomas, 1998]. A continuación se presenta el espectro de coherencia y sus propiedades, así como algunos estimadores.

5.1.1. Definición y Propiedades

Consideremos dos procesos estocásticos estacionarios en sentido amplio y de media nula: $x_1[n]$ y $x_2[n]$ y se define su covarianza cruzada ($r_{12}[m]$) y autocovarianzas ($r_{11}[m]$ y $r_{22}[m]$) de la siguiente manera

$$r_{ij}[m] = E\left[x_i[n]x_j^*[n-m]\right], \quad i,j = 1,2.$$

El espectro de potencia cruzado y las densidades espectrales de potencia vienen dados por la transformada de Fourier de la covarianza cruzada y las autocovarianzas, respectivamente, es decir

$$S_{ij}(e^{j\theta}) = \mathcal{F}\left(r_{ij}[m]\right), \quad i,j = 1,2.$$

Teniendo en cuenta las definiciones anteriores, ahora se definen el espectro de coherencia complejo y el espectro de coherencia.

Definición 5.1. *El espectro de coherencia complejo se define como el siguiente cociente*

$$C_{12}(e^{j\theta}) = \frac{S_{12}(e^{j\theta})}{\sqrt{S_{11}(e^{j\theta})S_{22}(e^{j\theta})}}. \tag{5.1}$$

Definición 5.2. *El espectro de coherencia es el valor absoluto del espectro de coherencia complejo*

$$\gamma(e^{j\theta}) = \left|C_{12}(e^{j\theta})\right|. \tag{5.2}$$

Ambas definiciones ilustran su estrecha relación con el coeficiente de correlación de dos variables aleatorias. De hecho, el espectro de coherencia es una medida de la relación lineal frecuencial entre las señales. Para demostrar esto, definamos la señal de error de la siguiente manera

$$e[n] = x_2[n] - \left(h * x_1\right)[n],$$

y su potencia como

$$P_{\mathrm{e}} = E\left[|e[n]|^2\right] = \int_{-\pi}^{\pi} S_{\mathrm{e}}(e^{j\theta})\frac{d\theta}{2\pi}.$$

Por lo tanto, minimizando la potencia del error respecto a $h[n]$, se obtiene el siguiente filtro

$$h[n] = \mathcal{F}^{-1}\left(\frac{S_{21}(e^{j\theta})}{S_{11}(e^{j\theta})}\right),$$

y la potencia del error residual es

$$P_{\mathrm{residual}} = \int_{-\pi}^{\pi}\left(1 - |C_{12}(e^{j\theta})|^2\right)S_{22}(e^{j\theta})\frac{d\theta}{2\pi}.$$

La ecuación anterior demuestra que para valores del espectro de coherencia próximos a 1, ambas señales se consideran linealmente dependientes dado que la potencia del error residual es próxima a 0 y, al contrario, si el espectro de coherencia es 0, no existe relación lineal entre ambas. Aunque no se puede asegurar nada sobre la posible existencia de relaciones no lineales.

Para finalizar este apartado, se presentan dos propiedades del espectro de coherencia, que se generalizarán (y demostrarán), posteriormente, para el GCS.

Propiedad 5.1. *El espectro de coherencia toma valores entre* 0 *y* 1*:*

$$0 \leq \gamma(e^{j\theta}) \leq 1.$$

Propiedad 5.2. *El espectro de coherencia es invariante a filtrados independientes. Definamos las señales filtradas* $y_i[n]$

$$y_i[n] = (a_i * x_i)[n], \quad i = 1, 2,$$

donde $a_i[n]$ $(i = 1, 2)$ *son dos filtros estables cualesquiera. Entonces el espectro de coherencia de las señales* $y_i[n]$ *coincide con el espectro de las señales* $x_i[n]$*:*

$$\gamma^{(\mathrm{y})}(e^{j\theta}) = \gamma^{(\mathrm{x})}(e^{j\theta}).$$

Aplicando la Propiedad 5.2, es sencillo demostrar que el espectro de coherencia de las señales blanqueadas es igual a su densidad de potencia cruzada, es decir,

$$\tilde{S}_{12}(e^{j\theta}) = C_{12}(e^{j\theta}),$$

donde $\tilde{S}_{12}(e^{j\theta})$ denota la densidad de potencia cruzada de las señales blanqueadas, $\tilde{x}_i[n]$, con

$$\tilde{x}_i[n] = \left(x_i * h_i^{(\mathrm{w})}\right)[n], \quad i = 1, 2,$$

donde $h_i^{(\mathrm{w})}[n]$ es un filtro dado por

$$h_i^{(\mathrm{w})}[n] = \mathcal{F}^{-1}\left(\frac{1}{\sqrt{S_{ii}(e^{j\theta})}}\right).$$

5.1.2. Estimadores del Espectro de Coherencia

En esta subsección se considera la estimación del CS. En primer lugar, se demuestra que la estima basada en el periodograma no es adecuada. Después se introducen dos alternativas al periodograma y, finalmente, se presenta una técnica basada en el análisis de correlaciones canónicas (CCA) para estimar el CS.

Estimador basado en el Periodograma

Este apartado presenta una estima del CS basada en el periodograma, viéndose posteriormente que no es un estimador adecuado del CS. Para ello, se define la señal bidimensional $\mathbf{x}[n] = [x_1[n], x_2[n]]^T$ y se calcula la estima basada en el periodograma de la matriz de densidades espectrales de la siguiente manera

$$\hat{\mathbf{S}}(e^{j\theta}) = \mathbf{x}(e^{j\theta})\mathbf{x}^H(e^{j\theta}) = \begin{bmatrix} \hat{S}_{11}(e^{j\theta}) & \hat{S}_{12}(e^{j\theta}) \\ \hat{S}_{12}^*(e^{j\theta}) & \hat{S}_{22}(e^{j\theta}) \end{bmatrix},$$

donde

$$\mathbf{x}(e^{j\theta}) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \mathbf{x}[n] e^{-j\theta n}.$$

Es fácil ver que la matriz $\hat{\mathbf{S}}(e^{j\theta})$ tiene rango uno y, por lo tanto, su determinante vale 0, es decir,

$$\det\left(\hat{\mathbf{S}}(e^{j\theta})\right) = \hat{S}_{11}(e^{j\theta})\hat{S}_{22}(e^{j\theta}) - \left|\hat{S}_{12}(e^{j\theta})\right|^2 = 0. \tag{5.3}$$

Finalmente, definiendo la estima del espectro de coherencia (al cuadrado) como

$$\hat{\gamma}^2(e^{j\theta}) = \frac{\left|\hat{S}_{12}(e^{j\theta})\right|^2}{\hat{S}_{11}(e^{j\theta})\hat{S}_{22}(e^{j\theta})}$$

y teniendo en cuenta (5.3), se demuestra $\hat{\gamma}(e^{j\theta}) = 1$, $\forall\theta$.

Estimador basado en el Método de Welch

En este apartado, se presenta un estimador basado en el método de Welch [Stoica y Moses, 2005] que fue propuesto en [Carter et al., 1973]. Este método soluciona el problema anterior mediante un enventanado y segmentado de las señales. Comencemos por dividir la señal i-ésima en Q segmentos, que pueden o no solapar entre sí. Dichos segmentos, enventanados por $w[n]$, se denotan de la siguiente manera

$$x_i^{(q)}[n] = w[n]x[n+qR], \quad n = 0, \ldots, N-1, \; q = 0, \ldots, Q-1.$$

y las estimas de las densidades espectrales de potencia y de la densidad de potencia cruzada son

$$\hat{S}_{ii}(e^{j\theta}) = \sum_{q=0}^{Q-1} \left|X_i^{(q)}(e^{j\theta})\right|^2, \quad i = 1, 2,$$

$$\hat{S}_{12}(e^{j\theta}) = \sum_{q=0}^{Q-1} X_1^{(q)}(e^{j\theta}) X_2^{(q)*}(e^{j\theta}),$$

donde $X_i^{(q)}(e^{j\theta})$ es la transformada de Fourier del segmento q-ésimo de la señal i-ésima. Finalmente, teniendo en cuenta las estimas anteriores y las definiciones, el espectro de coherencia complejo y el espectro de coherencia se estiman como

$$\hat{C}_{12}(e^{j\theta}) = \frac{\hat{S}_{12}(e^{j\theta})}{\sqrt{\hat{S}_{11}(e^{j\theta})\hat{S}_{22}(e^{j\theta})}},$$

$$\hat{\gamma}(e^{j\theta}) = \left|\hat{C}_{12}(e^{j\theta})\right|.$$

Nótese que la selección del número de segmentos, el % de solape y el tipo ventana, afectarán al rendimiento del estimador. Ventanas más suaves en los bordes mejoran el nivel de los lóbulos secundarios, pero disminuyen su resolución (aumentan la anchura del lóbulo principal), y al contrario. Sin embargo, este es un tema muy estudiado en el análisis espectral clásico [Stoica y Moses, 2005] y, por lo tanto, no se tratará en este trabajo.

Estimador basado en el MVDR

El filtro de mínima varianza sin distorsión (MVDR) o filtro de Capon [Capon, 1969], ha sido usado ampliamente en el campo de conformación de haz (*beamforming*) y del análisis espectral. Sin embargo, no se empleó para la estima del CS hasta ser propuesto en [Benesty et al., 2006]. Esta técnica se basa en una aproximación al problema mediante bancos de filtros, uno por cada frecuencia donde se estima el CS. Cada filtro se diseña para minimizar la energía a su salida, manteniendo su respuesta inalterada en esa frecuencia. Por lo tanto, el filtro se obtiene resolviendo el siguiente problema de optimización

$$\begin{array}{ll} \underset{\mathbf{h}_i(e^{j\theta})}{\text{minimizar}} & P_i(e^{j\theta}), \\ \text{sujeto a} & \mathbf{f}^H(e^{j\theta})\mathbf{h}_i(e^{j\theta}) = \sum_{n=0}^{N-1} h_i[n, e^{j\theta}]e^{-j\theta n} = 1, \end{array}$$

donde $\mathbf{h}_i(e^{j\theta}) = [h_i[0, e^{j\theta}), \ldots, h_i[N-1, e^{j\theta})]^T$ es un vector que contiene los N coeficientes del filtro a la frecuencia θ, $\mathbf{f}(e^{j\theta}) = 1/L\left[1, e^{j\theta}, \ldots, e^{j\theta(N-1)}\right]^T$, es el vector de Fourier de longitud N a la frecuencia θ, y $P_i(e^{j\theta})$ es la potencia a la salida del filtro, dada por la siguiente forma cuadrática

$$P_i(e^{j\theta}) = \mathbf{h}_i^H(e^{j\theta}) \overbrace{\begin{bmatrix} r_{ii}[0] & r_{ii}[1] & \cdots & r_{ii}[N-1] \\ r_{ii}[-1] & r_{ii}[0] & \cdots & r_{ii}[N-2] \\ \vdots & \vdots & \ddots & \vdots \\ r_{ii}[-N+1] & r_{ii}[-N+2] & \cdots & r_{ii}[0] \end{bmatrix}}^{\mathbf{R}_{ii}^T} \mathbf{h}_i(e^{j\theta}), \quad i = 1, 2.$$

Resolviendo el problema de optimización mediante la técnica de los multiplicadores de Lagrange [Boyd y Vandenberghe, 2004], los filtros son

$$\mathbf{h}_i(e^{j\theta}) = \frac{\mathbf{R}_{ii}^{-T}\mathbf{f}(e^{j\theta})}{\mathbf{f}^H(e^{j\theta})\mathbf{R}_{ii}^{-T}\mathbf{f}(e^{j\theta})}. \tag{5.4}$$

Por último, la densidad espectral de potencia y la densidad de potencia cruzada vienen dadas por

$$\begin{aligned} S_{ii}(e^{j\theta}) &= \mathbf{h}_i^H(e^{j\theta})\mathbf{R}_{ii}^T\mathbf{h}_i(e^{j\theta}), \quad i = 1, 2 \\ S_{12}(e^{j\theta}) &= \mathbf{h}_1^H(e^{j\theta})\mathbf{R}_{12}^T\mathbf{h}_2(e^{j\theta}), \end{aligned}$$

donde $\mathbf{R}_{12}$ se define de manera análoga a $\mathbf{R}_{ii}$, y el espectro de coherencia es

$$\gamma(e^{j\theta}) = \frac{\left|\mathbf{f}^H(e^{j\theta})\mathbf{R}_{11}^{-T}\mathbf{R}_{12}^T\mathbf{R}_{22}^{-T}\mathbf{f}(e^{j\theta})\right|}{\sqrt{\mathbf{f}^H(e^{j\theta})\mathbf{R}_{11}^{-T}\mathbf{f}(e^{j\theta})\mathbf{f}^H(e^{j\theta})\mathbf{R}_{22}^{-T}\mathbf{f}(e^{j\theta})}}.$$

En el desarrollo anterior, se ha supuesto que las matrices de covarianza son conocidas sin error. Sin embargo, esto no será cierto en general, al ser estimadas a partir de un conjunto finito de observaciones.

Estimador basado en CCA

Este apartado presenta un estimador del CS basado en CCA [Santamaría y Vía, 2007]. Comencemos por reescribir el espectro de coherencia complejo de la siguiente manera [Scharf y Thomas, 1998]

$$C_{12}(e^{j\theta}) = \mathbf{f}^H(e^{j\theta})\mathbf{R}_{11}^{-1/2}\mathbf{R}_{12}\mathbf{R}_{22}^{-1/2}\mathbf{f}(e^{j\theta}) = \mathbf{f}^H(e^{j\theta})\mathbf{C}_{12}\mathbf{f}(e^{j\theta}),$$

donde $\mathbf{C}_{12} = \mathbf{R}_{11}^{-1/2}\mathbf{R}_{12}\mathbf{R}_{22}^{-1/2}$ es una matriz de coherencia, $\mathbf{f}(e^{j\theta})$ es el vector de Fourier de tamaño infinito y $\mathbf{R}_{ij}$ es la matriz Toeplitz de covarianza de tamaño infinito dada por

$$\mathbf{R}_{ij} = \begin{bmatrix} r_{ij}[0] & r_{ij}[-1] & r_{ij}[-2] & \cdots \\ r_{ij}[1] & r_{ij}[0] & r_{ij}[-1] & \cdots \\ r_{ij}[2] & r_{ij}[1] & r_{ij}[0] & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \quad i,j = 1,2.$$

Adicionalmente se considera la descomposición en valores singulares (SVD) [Golub y van Loan, 1996] de la matriz de coherencia de dimensiones finitas y estimada $\hat{\mathbf{C}}_{12} = \hat{\mathbf{R}}_{11}^{-1/2}\hat{\mathbf{R}}_{12}\hat{\mathbf{R}}_{22}^{-1/2} = \hat{\mathbf{F}}_1\hat{\mathbf{\Lambda}}\hat{\mathbf{F}}_2^H$.

La idea básica del estimador consiste en eliminar el subespacio asociado con las correlaciones canónicas (ver apéndice C) de menor valor, las cuales se ven más afectadas por errores de estimación. De esta manera, este estimador es similar a otras propuestas basadas en las coordenadas canónicas, usadas previamente en problemas de codificación, filtrado, etc. [Scharf y Thomas, 1998, Scharf y Mullis, 2000, Hua et al., 2001, Schreier y Scharf, 2006]. Concretamente, se propone usar una versión de rango reducido de la matriz de coherencia estimada $\breve{\mathbf{C}}_{12} = \breve{\mathbf{F}}_1\breve{\mathbf{\Lambda}}\breve{\mathbf{F}}_2^H$, donde $\breve{\mathbf{F}}_i (i = 1,2)$ contiene los vectores singulares asociados a los κ mayores valores singulares, $\breve{\mathbf{\Lambda}}$ es una matriz diagonal que contiene los κ mayores valores singulares de $\hat{\mathbf{\Lambda}}$, obteniéndose la siguiente estima del CS

$$\hat{\gamma}(e^{j\theta}) = \mathbf{f}^H(e^{j\theta})\breve{\mathbf{C}}_{12}\mathbf{f}(e^{j\theta}).$$

La selección del valor κ es un problema muy complicado que no se tratará teóricamente. No obstante, en la Sección 5.4 se da alguna idea sobre cómo seleccionarlo.

Finalmente, analizando la estima del CS en términos de banco de filtros, es sencillo obtener los siguientes filtros equivalentes para la estima basada en CCA $\mathbf{h}_i(e^{j\theta}) = \hat{\mathbf{R}}_{ii}^{-1/2}\breve{\mathbf{F}}_i\breve{\mathbf{F}}_i^H\hat{\mathbf{R}}_{ii}^{1/2}\mathbf{f}(e^{j\theta})$. Esto demuestra que, a diferencia del estimador basado en el MVDR, los filtros para la técnica de CCA no sólo dependen de la autocorrelación de las señales, sino también de su correlación cruzada.

5.2. El Espectro de Coherencia Generalizado

A pesar de no existir una definición comúnmente aceptada para la generalización del espectro de coherencia, algunos investigadores han propuesto diferentes técnicas para analizar la correlación entre $L \geq 2$ series temporales. Por un lado, tenemos la coherencia parcial o coherencia parcial dirigida [Sameshima y Baccalá, 1999, Winterhalder et al., 2005], la cual selecciona dos series temporales de entre las L, substrae la influencia lineal del resto y computa el espectro de coherencia clásico entre ellas. Por otro lado, las L series temporales se pueden agrupar en dos subconjuntos y tratar dichos conjuntos como dos series temporales multidimensionales. Este idea permite aplicar una versión frecuencial del análisis de correlaciones

canónicas (CCA) propuesto en [Brillinger, 1981]. Sin embargo, ambas propuestas dependen de cómo se agrupen las señales, es decir, son medidas no simétricas.

Para evitar estos problemas, basándonos en GLRT derivados en capítulos anteriores, se propone una nueva medida frecuencial de correlación para múltiples series temporales, definiendo el espectro de coherencia generalizado (GCS) como una función Schur-convexa [Marshall y Olkin, 1979, Jorswieck y Boche, 2007, Schreier, 2008] del vector de autovalores de una matriz que contiene todos los espectros de coherencia complejos a una frecuencia dada. Consideremos L procesos estocásticos estacionarios en sentido amplio y de media nula: $x_1[n], \ldots, x_L[n]$ y recordemos la definición del espectro de coherencia complejo entre las señales $x_i[n]$ y $x_j[n]$ de la Sección 5.1

$$C_{ij}(e^{j\theta}) = \frac{S_{ij}(e^{j\theta})}{\sqrt{S_{ii}(e^{j\theta})S_{jj}(e^{j\theta})}}, \quad i,j = 1, \ldots, L.$$

Para extender la idea del espectro de coherencia al caso de más de dos series temporales, se define la matriz $\mathbf{C}(e^{j\theta}) \in \mathbb{C}^{L\times L}$, que contiene todos los espectros de coherencia complejos por parejas, de la siguiente manera

$$\mathbf{C}(e^{j\theta}) = \begin{bmatrix} 1 & C_{12}(e^{j\theta}) & \ldots & C_{1L}(e^{j\theta}) \\ C_{21}(e^{j\theta}) & 1 & \ldots & C_{2L}(e^{j\theta}) \\ \vdots & \vdots & \ddots & \vdots \\ C_{L1}(e^{j\theta}) & C_{L2}(e^{j\theta}) & \ldots & 1 \end{bmatrix},$$

la cual, se puede reescribir como $\mathbf{C}(e^{j\theta}) = \mathbf{D}^{-1/2}(e^{j\theta})\mathbf{S}(e^{j\theta})\mathbf{D}^{-1/2}(e^{j\theta})$, donde

$$\mathbf{S}(e^{j\theta}) = \begin{bmatrix} S_{11}(e^{j\theta}) & S_{12}(e^{j\theta}) & \ldots & S_{1L}(e^{j\theta}) \\ S_{21}(e^{j\theta}) & S_{22}(e^{j\theta}) & \ldots & S_{2L}(e^{j\theta}) \\ \vdots & \vdots & \ddots & \vdots \\ S_{L1}(e^{j\theta}) & S_{L2}(e^{j\theta}) & \ldots & S_{LL}(e^{j\theta}) \end{bmatrix},$$

y $\mathbf{D}(e^{j\theta})$ es una matriz diagonal con los siguientes elementos $[\mathbf{D}(e^{j\theta})]_{i,i} = S_{ii}(e^{j\theta})$, $(i = 1, \ldots, L)$. Adicionalmente se define el vector $\boldsymbol{\beta}(e^{j\theta}) = \left[\beta_1(e^{j\theta}), \ldots, \beta_L(e^{j\theta})\right]^T$, que contiene, en orden decreciente, todos los autovalores del siguiente problema de autovalores

$$\mathbf{C}(e^{j\theta})\mathbf{v}(e^{j\theta}) = \beta(e^{j\theta})\mathbf{v}(e^{j\theta}). \tag{5.5}$$

o del problema de autovalores generalizado dado por

$$\mathbf{S}(e^{j\theta})\tilde{\mathbf{v}}(e^{j\theta}) = \beta(e^{j\theta})\mathbf{D}(e^{j\theta})\tilde{\mathbf{v}}(e^{j\theta}), \tag{5.6}$$

donde $\tilde{\mathbf{v}}(e^{j\theta}) = \mathbf{D}^{-1/2}(e^{j\theta})\mathbf{v}(e^{j\theta})$ es el autovector generalizado.

Definición 5.3. *El espectro de coherencia generalizado (GCS) se define como*

$$\gamma(e^{j\theta}) = f\left(\boldsymbol{\beta}(e^{j\theta})\right),$$

donde $f(\boldsymbol{\beta}(e^{j\theta}))$ es una función que debe cumplir las siguiente condiciones:

c.1) $f\left(\boldsymbol{\beta}(e^{j\theta})\right)$ es una función escalar acotada del vector $\boldsymbol{\beta}(e^{j\theta})$ en el intervalo $[0,1]$, es decir,

$$f : \mathbb{R}^{L\times 1} \to \mathcal{A},$$

donde $\mathcal{A} = \{x \in \mathbb{R} \mid 0 \leq x \leq 1\}$.

c.2) Para $L = 2$ procesos estocásticos, $f\left(\boldsymbol{\beta}(e^{j\theta})\right)$ debe ser el CS.

c.3) $f\left(\boldsymbol{\beta}(e^{j\theta})\right)$ debe ser una función Schur-convexa.[1]

La elección de la función $f(\cdot)$ determina la medida de coherencia concreta y, por consiguiente, la información extraída de los L procesos. En particular, se proponen las siguientes tres opciones para $f(\cdot)$:

Definición 5.4. *El GCS-MAX se define como*

$$\gamma_{\mathrm{MAX}}(e^{j\theta}) = \frac{1}{L-1}\left(\beta_{\mathrm{MAX}}(e^{j\theta}) - 1\right),$$

donde $\beta_{\mathrm{MAX}}(e^{j\theta}) = \beta_1(e^{j\theta})$ es el mayor autovalor de (5.5).

En la Sección 5.2.1 se demuestra que esta medida alcanza su máximo cuando todas las señales están perfectamente correladas por parejas.

Definición 5.5. *El GCS-MIN se define como*

$$\gamma_{\mathrm{MIN}}(e^{j\theta}) = 1 - \beta_{\mathrm{MIN}}(e^{j\theta}),$$

donde $\beta_{\mathrm{MIN}}(e^{j\theta}) = \beta_L(e^{j\theta})$ es el menor autovalor de (5.5).

Usando el menor autovalor, la medida alcanza su máximo cuando al menos dos de las L señales están perfectamente correladas. O, más formalmente, cuando cualquier señal se puede expresar como una combinación lineal de las $L-1$ restantes.

Definición 5.6. *El GCS-DET se define como*

$$\gamma_{\mathrm{DET}}(e^{j\theta}) = \sqrt{1 - \prod_{i=1}^{L} \beta_i(e^{j\theta})} = \sqrt{1 - \det(\mathbf{C}(e^{j\theta}))}.$$

El GCS-DET también presenta su máximo cuando al menos dos de los L procesos están perfectamente correlados. Sin embargo, su comportamiento es, en general, diferente al del GCS-MIN. Por otro lado, la definición del GCS-DET se puede reescribir como

$$\gamma^2_{\mathrm{DET}}(e^{j\theta}) = 1 - \frac{\det(\mathbf{S}_x(e^{j\theta}))}{\prod_{i=1}^{L} S_{ii}(e^{j\theta})},$$

donde se observa su relación con el cociente de Hadamard, el cual es una medida de dependencia lineal como se demuestra en [Scharf y Thomas, 1998].

Las tres funciones $f(\cdot)$ propuestas cumplen la condición c.2, es decir, en el caso de 2 señales el GCS coincide con el CS. Esto se demuestra en la Sección 5.2.1 para el GCS-MAX, aunque es sencillo demostrarlo para el resto de definiciones. Además, se puede comprobar que todas las particularizaciones son funciones Schur-convexas [Jorswieck y Boche, 2007].

[1] Esta propiedad se establece debido a que las funciones Schur-convexas mantienen el ordenamiento de los vectores, ver Apéndice B o [Marshall y Olkin, 1979, Jorswieck y Boche, 2007].

Resumiendo, toda la información necesaria para extender el CS a más de dos series temporales la proporcionan los autovalores de la matriz $\mathbf{C}(e^{j\theta})$. Diferentes funciones de los autovalores proporcionan diferentes medidas, con comportamientos diferentes; y todas ellas con utilidad en diferentes problemas. A continuación se presenta un ejemplo ilustrativo del comportamiento del GCS: consideremos cuatro señales con las siguientes matrices $\mathbf{C}(e^{j\theta})$ a las frecuencias θ_1, θ_2 y θ_3

$$\mathbf{C}(e^{j\theta_1}) = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}, \quad \mathbf{C}(e^{j\theta_2}) = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{C}(e^{j\theta_3}) = \begin{bmatrix} 1 & 0.25 & 0.5 & 0.25 \\ 0.25 & 1 & 0.25 & 0.5 \\ 0.5 & 0.25 & 1 & 0.25 \\ 0.25 & 0.5 & 0.25 & 1 \end{bmatrix}.$$

Dichas matrices indican que las señales están perfectamente correladas a la frecuencia θ_1, a la frecuencia θ_2 sólo las dos primeras presentan correlación, mientras que a la frecuencia θ_3 la interpretación de la matriz es más complicada. Para este ejemplo, se muestran los valores del GCS-MAX, GCS-MIN y GCS-DET en la Tabla 5.1. En dicha tabla se ve que el GCS-MAX es la única definición que alcanza su máximo sólo cuando todas las señales están perfectamente correladas. Por otro lado, el GCS-MIN y el GCS-DET proporcionan el mismo valor a la frecuencia θ_2 y valores diferentes a θ_3.

Frecuencia	GCS-MAX	GCS-MIN	GCS-DET
θ_1	1	1	1
θ_2	1/3	1	1
θ_3	1/3	0.5	$\sqrt{2}/2$

Tabla 5.1: GCS-MAX, GCS-MIN y GCS-DET para el ejemplo ilustrativo.

A pesar del interés de todas las definiciones, a continuación, se presenta únicamente el análisis del GCS-MAX, pudiéndose extender algunos resultados al resto de particularizaciones del GCS de manera sencilla.

5.2.1. Propiedades

Esta sección presenta las propiedades del GCS-MAX.[2]

Propiedad 5.3. *El GCS-MAX está acotado entre* 0 *y* 1*, es decir,*

$$0 \leq \gamma_{MAX}(e^{j\theta}) \leq 1.$$

Demostración. Teniendo en cuenta $\operatorname{tr}\left(\mathbf{C}(e^{j\theta})\right) = \sum_{i=1}^{L} \beta_i(e^{j\theta}) = L$, donde $\beta_i(e^{j\theta}) \geq 0$ son los autovalores, es fácil demostrar que el máximo valor de $\beta_1(e^{j\theta})$ es L cuando $\beta_i(e^{j\theta}) = 0$, $i = 2, \ldots, L$; mientras que su mínimo valor es 1 cuando todos los autovalores son iguales $\beta_i(e^{j\theta}) = 1$, $i = 1, \ldots, L$. Entonces, $1 \leq \lambda_{\mathrm{MAX}}(e^{j\theta}) \leq L$, lo cual implica $0 \leq \gamma_{\mathrm{MAX}}(e^{j\theta}) \leq 1$. □

Propiedad 5.4. *En el caso de $L = 2$ señales, el GCS-MAX se reduce a la definición del CS estándar.*

[2]Las Propiedades 5.3, 5.4 y 5.5 son válidas para cualquier definición del GCS, incluyendo el GCS-MIN y el GCS-DET.

Demostración. Para $L = 2$, la matriz $\mathbf{C}(e^{j\theta})$ es

$$\mathbf{C}(e^{j\theta}) = \begin{bmatrix} 1 & C_{12}(e^{j\theta}) \\ C_{12}^*(e^{j\theta}) & 1 \end{bmatrix},$$

siendo su mayor autovalor $\beta_{\mathrm{MAX}}(e^{j\theta}) = 1 + |C_{12}(e^{j\theta})|$. Por lo tanto, el GCS-MAX es $\gamma_{\mathrm{MAX}}(e^{j\theta}) = (\beta_{\mathrm{MAX}}(e^{j\theta}) - 1) = |C_{12}(e^{j\theta})|$. □

Propiedad 5.5. *Consideremos las siguientes señales*

$$y_i[n] = (x_i * a_i)[n], \quad i = 1, \ldots, L,$$

donde $a_i[n]$ es la respuesta impulsional de unos filtros estables arbitrarios con respuesta frecuencial $A_i(e^{j\theta}) = |A_i(e^{j\theta})|e^{j\phi_i(e^{j\theta})}$, que satisface $|A_i(e^{j\theta})| \neq 0$, $\forall\theta \in [0, 2\pi)$. Entonces, el GCS-MAX de las señales $x_i[n]$ e $y_i[n]$ son idénticos,

$$\gamma_{\mathrm{MAX}}^{(\mathrm{y})}(e^{j\theta}) = \gamma_{\mathrm{MAX}}^{(\mathrm{x})}(e^{j\theta}),$$

donde $\gamma_{\mathrm{MAX}}^{(\mathrm{x})}(e^{j\theta})$ y $\gamma_{\mathrm{MAX}}^{(\mathrm{y})}(e^{j\theta})$ son el GCS-MAX de las señales originales $x_i[n]$ y de las filtradas $y_i[n]$, respectivamente.

Demostración. La covarianza cruzada entre la señales filtradas i y l es

$$r_{il}^{(\mathrm{y})}[m] = E\left[y_i[n]y_l^*[n-m]\right] = \sum_{k=-\infty}^{\infty}\sum_{s=-\infty}^{\infty} a_i[k]a_l[s]r_{il}^{(\mathrm{x})}[m+s-k],$$

donde $r_{il}^{(\mathrm{y})}[m]$ y $r_{il}^{(\mathrm{x})}[m]$ son, respectivamente, las secuencias de covarianza cruzada entre las señales i y l filtradas y originales. Tomando la transformada de Fourier de $r_{il}^{(\mathrm{y})}[m]$ se obtiene

$$S_{il}^{(\mathrm{y})}(e^{j\theta}) = \left(\sum_{m=-\infty}^{\infty} a_i[m]e^{-j\theta m}\right)\left(\sum_{k=-\infty}^{\infty} a_l[k]e^{-j\theta k}\right)^*\left(\sum_{s=-\infty}^{\infty} r_{il}^{(\mathrm{x})}[k]e^{-j\theta k}\right)$$

que finalmente se puede reescribir como $S_{il}^{(\mathrm{y})}(e^{j\theta}) = A_i(e^{j\theta})A_l^*(e^{j\theta})S_{il}^{(\mathrm{x})}(e^{j\theta})$. Por lo tanto, el espectro de coherencia complejo es $C_{il}^{(\mathrm{y})}(e^{j\theta}) = e^{j\phi_i(e^{j\theta})}e^{-j\phi_j(e^{j\theta})}C_{il}^{(\mathrm{x})}(e^{j\theta})$, y $\mathbf{C}_{\mathrm{y}}(e^{j\theta})$ viene dada por

$$\mathbf{C}_{\mathrm{y}}(e^{j\theta}) = \mathbf{Q}(e^{j\theta})\mathbf{C}_{\mathrm{x}}(e^{j\theta})\mathbf{Q}^H(e^{j\theta}), \tag{5.7}$$

donde $\mathbf{Q}(e^{j\theta}) = \mathrm{diag}\left(e^{j\phi_1(e^{j\theta})}, \ldots, e^{j\phi_M(e^{j\theta})}\right)$ es una matriz unitaria. Entonces, el problema de autovalores para las señales filtradas es

$$\mathbf{C}_{\mathrm{y}}(e^{j\theta})\mathbf{v}_{\mathrm{y}}(e^{j\theta}) = \beta_{\mathrm{y}}(e^{j\theta})\mathbf{v}_{\mathrm{y}}(e^{j\theta}),$$

y, teniendo en cuenta (5.7), obtenemos $\mathbf{v}_{\mathrm{x}}(e^{j\theta}) = \mathbf{Q}^H(e^{j\theta})\mathbf{v}_{\mathrm{y}}(e^{j\theta})$ y $\beta_{\mathrm{y}}(e^{j\theta}) = \beta_{\mathrm{x}}(e^{j\theta})$. Esto significa que los autovalores no cambian debido a un filtrado de las señales si los filtros son invertibles. Por otro lado, los autovectores sólo varían en una rotación. Dado que el GCS-MAX depende del mayor autovalor y no de los autovectores, los espectros de coherencia generalizados de las señales originales y de las señales filtradas son iguales. □

Propiedad 5.6. *El GCS-MAX es máximo cuando las L series temporales están perfectamente correladas por parejas a una determinada frecuencia y mínimo cuando están incorreladas.*

Demostración. Para señales perfectamente correladas por parejas, la matriz $\mathbf{C}(e^{j\theta})$ es una matriz de rango uno, siendo su mayor autovalor $\beta_{\text{MAX}}(e^{j\theta}) = L$. Por otro lado, en el caso de señales incorreladas, $\mathbf{C}(e^{j\theta}) = \mathbf{I}$ y $\beta_{\text{MAX}}(e^{j\theta}) = 1$. □

Propiedad 5.7. *La contribución de la i-ésima señal al GCS-MAX en la frecuencia θ se puede medir con el i-ésimo elemento, $v_i(e^{j\theta})$, del autovector $\mathbf{v}(e^{j\theta})$ asociado al mayor autovalor de $\mathbf{C}(e^{j\theta})$.*

Demostración. Comencemos por reescribir el mayor autovalor de $\mathbf{C}(e^{j\theta})$ como

$$\beta_{\text{MAX}}(e^{j\theta}) = \mathbf{v}^H(e^{j\theta})\mathbf{C}(e^{j\theta})\mathbf{v}(e^{j\theta}) = \sum_{i=1}^{L}\sum_{l=1}^{L} v_i^*(e^{j\theta})v_l(e^{j\theta})\tilde{S}_{il}(e^{j\theta}), \tag{5.8}$$

donde se ha tenido en cuenta que el espectro de coherencia complejo coincide con el espectro cruzado de las señales blanqueadas, $\tilde{x}_i[n]$, y $\mathbf{v}(e^{j\theta})$ es el autovector asociado a $\beta_{\text{MAX}}(e^{j\theta})$. Teniendo en cuenta la definición del espectro cruzado, se puede reescribir (5.8) como

$$\beta_{\text{MAX}}(e^{j\theta}) = \sum_{m=-\infty}^{\infty} E\left[\sum_{i=1}^{L}\sum_{l=1}^{L} v_i^*(e^{j\theta})v_l(e^{j\theta})\tilde{x}_i[n]\tilde{x}_l^*[n-m]\right] e^{-j\theta m},$$

y definiendo[3] $z[n, e^{j\theta}) = \sum_{i=1}^{M} v_i^*(e^{j\theta})\tilde{x}_i[n]$, $\beta_{\text{MAX}}(e^{j\theta})$ es

$$\beta_{\text{MAX}}(e^{j\theta}) = \sum_{m=-\infty}^{\infty} E\left[z[n, e^{j\theta})z^*[n-m, e^{j\theta})\right] e^{-j\theta m} = \mathcal{F}\left(r_z[m, e^{j\theta})\right).$$

Evidentemente, $\beta_{\text{MAX}}(e^{j\theta})$ es la densidad espectral de potencia de la señal $z[n, e^{j\theta})$, la cual es una suma ponderada de las señales blanqueadas $\tilde{x}_i[n]$, y los pesos vienen dados por los elementos del autovector $\mathbf{v}(e^{j\theta})$. La prueba concluye recordando que el GCS-MAX es una transformación afín de $\beta_{\text{MAX}}(e^{j\theta})$. □

Clarifiquemos esta propiedad con el siguiente ejemplo. Se considera un experimento con $L = 3$ procesos estocásticos, dos de ellos perfectamente correlados a la frecuencia θ, mientras que el tercero está incorrelado con ambos. Para este ejemplo, la matriz $\mathbf{C}(e^{j\theta})$ es

$$\mathbf{C}(e^{j\theta}) = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

el GCS-MAX a la frecuencia θ vale $\gamma(e^{j\theta}) = 1/2$, y el autovector correspondiente es

$$\mathbf{v}(e^{j\theta}) = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \end{bmatrix}^T,$$

indicando, claramente, que el tercer proceso no contribuye al GCS-MAX, mientras que los dos primeros contribuyen por igual.

[3] Aquí, se introduce la notación $z[n, e^{j\theta})$ para destacar que la señal es también función de la frecuencia en la que estamos evaluando el GCS-MAX.

5.2.2. Estimación del GCS-MAX

Esta sección presenta dos técnicas para estimar el GCS-MAX. La primera técnica, que es muy sencilla, puede ser aplicada a cualquier definición del GCS. Dicha técnica está basada en la estima de los $L(L-1)/2$ espectros de coherencia complejos por parejas diferentes, mediante cualquier estimador presentado en la Sección 5.1.2. Con estas estimas se forma la matriz $\hat{\mathbf{C}}(e^{j\theta})$

$$\hat{\mathbf{C}}(e^{j\theta}) = \begin{bmatrix} 1 & \hat{C}_{12}(e^{j\theta}) & \dots & \hat{C}_{1L}(e^{j\theta}) \\ \hat{C}^*_{12}(e^{j\theta}) & 1 & \dots & \hat{C}_{2L}(e^{j\theta}) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{C}^*_{1L}(e^{j\theta}) & \hat{C}^*_{2L}(e^{j\theta}) & \dots & 1 \end{bmatrix},$$

obteniendo directamente la estima a partir de la definición.

Como alternativa a esta técnica, en la siguiente subsección se presenta un estimador basado en CCA-MAXVAR, siendo específico para el GCS-MAX.

Estima del GCS-MAX mediante CCA-MAXVAR

En primer lugar se analiza el caso asintótico para ver la relación entre el GCS-MAX y la generalización de máxima varianza (MAXVAR) de CCA (ver Apéndice C y [Kettenring, 1971]), considerando, a continuación, el caso práctico. Comencemos por expresar el espectro cruzado entre la señal i-ésima y la l-ésima de la siguiente manera [Papoulis y Pillai, 2002]

$$S_{il}(e^{j\theta}) = \mathbf{f}^H(e^{j\theta})\mathbf{R}_{il}\mathbf{f}(e^{j\theta}),$$

donde $\mathbf{R}_{il}$ es la matriz Toeplitz infinita de correlación cruzada

$$\mathbf{R}_{il} = \begin{bmatrix} r_{il}[0] & r_{il}[-1] & r_{il}[-2] & \cdots \\ r_{il}[1] & r_{il}[0] & r_{il}[-1] & \cdots \\ r_{il}[2] & r_{il}[1] & r_{il}[0] & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

y $\mathbf{f}(e^{j\theta})$ es el vector de Fourier de longitud infinita a la frecuencia θ. De manera similar, las matrices $\mathbf{S}(e^{j\theta})$ y $\mathbf{D}(e^{j\theta})$ se pueden reescribir como

$$\mathbf{S}(e^{j\theta}) = \mathbf{F}^H(e^{j\theta}) \overbrace{\begin{bmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} & \dots & \mathbf{R}_{1L} \\ \mathbf{R}_{21} & \mathbf{R}_{22} & \dots & \mathbf{R}_{2L} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{R}_{L1} & \mathbf{R}_{L2} & \dots & \mathbf{R}_{LL} \end{bmatrix}}^{\mathbf{R}} \mathbf{F}(e^{j\theta}), \tag{5.9}$$

$$\mathbf{D}(e^{j\theta}) = \mathbf{F}^H(e^{j\theta}) \underbrace{\begin{bmatrix} \mathbf{R}_{11} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_{22} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{R}_{LL} \end{bmatrix}}_{\mathbf{D}} \mathbf{F}(e^{j\theta}), \tag{5.10}$$

donde $\mathbf{F}(e^{j\theta}) = \mathbf{I} \otimes \mathbf{f}(e^{j\theta})$ y $\otimes$ denota el producto de Kronecker. Teniendo en cuenta que los autovectores de las matrices $\mathbf{R}_{il}$ se aproximan asintóticamente a los de Fourier

[Grenander y Szegö, 1958, Scharf y Thomas, 1998], es sencillo demostrar que si $\beta\left(e^{j\theta_p}\right)$ y $\mathbf{v}\left(e^{j\theta_p}\right)$ son, respectivamente, un autovalor y su correspondiente autovector a la frecuencia θ_p de la matriz $\mathbf{C}\left(e^{j\theta_p}\right)$, entonces $\beta\left(e^{j\theta_p}\right)$ también es un autovalor de la matriz $\mathbf{D}^{-1/2}\mathbf{R}\mathbf{D}^{-1/2}$ con autovector $\mathbf{w}^{(p)}$, es decir,

$$\mathbf{D}^{-1/2}\mathbf{R}\mathbf{D}^{-1/2}\mathbf{w}^{(p)} = \beta\left(e^{j\theta_p}\right)\mathbf{w}^{(p)}, \tag{5.11}$$

donde $\mathbf{R}$ y $\mathbf{D}$ se han definido en (5.9) y (5.10), respectivamente, y $\mathbf{w}^{(p)} = \mathbf{F}\left(e^{j\theta_p}\right)\mathbf{v}\left(e^{j\theta_p}\right) = \mathbf{v}\left(e^{j\theta_p}\right)\otimes\mathbf{f}\left(e^{j\theta_p}\right)$. El problema EV (5.11) es la formulación clásica de la generalización de máxima varianza (MAXVAR) de CCA [Kettenring, 1971, Vía et al., 2007] que puede reescribirse como el siguiente problema GEV

$$\mathbf{R}\tilde{\mathbf{w}}^{(p)} = \beta\left(e^{j\theta_p}\right)\mathbf{D}\tilde{\mathbf{w}}^{(p)},$$

donde $\tilde{\mathbf{w}}^{(p)} = \mathbf{D}^{-1/2}\mathbf{w}^{(p)}$. La Ecuación (5.11) demuestra que en el caso asintótico, matrices de correlación cruzada infinitas, el GCS-MAX puede obtenerse directamente de los autovalores de $\mathbf{D}^{-1/2}\mathbf{R}\mathbf{D}^{-1/2}$. Sin embargo, en la práctica se trabaja con matrices de covarianza cruzada de tamaño finito $(L \times L)$, estimadas a partir de un conjunto finito de observaciones. Esto provoca diferencias entre los autovectores teóricos y estimados, $\hat{\mathbf{w}}^{(p)} = \left[\hat{\mathbf{w}}_1^{(p)T}, \ldots, \hat{\mathbf{w}}_L^{(p)T}\right]^T$, que se obtienen al resolver

$$\hat{\mathbf{D}}^{-1/2}\hat{\mathbf{R}}\hat{\mathbf{D}}^{-1/2}\hat{\mathbf{w}}^{(p)} = \hat{\beta}^{(p)}\hat{\mathbf{w}}^{(p)}, \quad p = 1, \ldots, LN, \tag{5.12}$$

donde $\hat{\mathbf{D}} \in \mathbb{C}^{LN\times LN}$ y $\hat{\mathbf{R}} \in \mathbb{C}^{LN\times LN}$ son las versiones finitas y estimadas de $\mathbf{D}$ y $\mathbf{R}$. Nótese que para un problema EV de dimensiones infinitas (5.11) existe un autovector $\mathbf{w}^{(p)}$ por cada frecuencia θ_p. Por el contrario, para problemas de dimensión finita (5.12), cada uno de los LN autovectores estimados proporciona información sobre todas las frecuencias. Entonces, $\hat{\mathbf{w}}^{(p)}$ puede interpretarse como un filtro banda estrecha centrado en θ_p, cuyo ancho de banda tiende a 0 cuando N tiende a infinito. Sólo queda una pregunta: ¿Se deben usar todos los autovectores o un subconjunto de ellos para estimar el GCS-MAX? Al igual que ocurre con el estimador del CS basado en CCA, puede argumentarse que autovectores asociados a menores correlaciones canónicas están más afectados por errores en la estimación. Debido a esto, proponemos usar una versión de rango reducido de la matriz $\hat{\mathbf{D}}^{-1/2}\hat{\mathbf{R}}\hat{\mathbf{D}}^{-1/2}$, de manera similar a las técnicas de [Scharf y Thomas, 1998, Scharf y Mullis, 2000, Hua et al., 2001, Schreier y Scharf, 2006, Santamaría y Vía, 2007]. Concretamente, la estima del GCS-MAX propuesta viene dada por una suma ponderada del módulo al cuadrado de la transformada de Fourier de los κ vectores canónicos principales

$$\hat{\gamma}_{MAX}(e^{j\theta}) = \frac{1}{L-1}\sum_{p=1}^{\kappa}\sum_{k=1}^{L}\left(\hat{\beta}^{(p)} - 1\right)\left|\mathbf{f}^H(e^{j\theta})\hat{\mathbf{w}}_k^{(p)}\right|^2,$$

donde $\mathbf{f}(e^{j\theta})$ es el vector de Fourier de longitud N, κ es el rango seleccionado, y $\hat{\beta}^{(p)}$, $p = 1, \ldots, \kappa$, son los κ mayores autovalores de (5.12). De manera similar a la estima del CS basada en CCA, la selección óptima de κ mediante un análisis teórico es muy complicada, y sólo se proponen algunas ideas para seleccionarlo en la sección de resultados.

5.3. Relaciones entre el GCS y otras Medidas Estadísticas

Esta sección discute la relación entre el GCS y otras medidas estadísticas. Por un lado, se presenta la relación entre el GCS y el GLRT para ruidos no-IID. Por otro, se relaciona el GCS con una medida ampliamente utilizada en la teoría de la información, la información mutua.

5.3.1. Relación con el GLRT de Señales Coloreadas Contaminadas por Ruidos no-IID

El GLRT para series temporales con PSD arbitrarias sin estructura espacial contaminadas por ruidos no-IID se puede expresar como la integral de una función del GCS-DET, lo cual se demuestra a continuación. El GLRT para este test de hipótesis viene dado por (4.18), pudiéndose reescribir como

$$\log \mathscr{L} = \int_{-\pi}^{\pi} \log\det\left(\hat{\mathbf{C}}\left(e^{j\theta}\right)\right)\frac{d\theta}{2\pi},$$

y aplicando la Definición 5.6 el GLRT es

$$\log \mathscr{L} = \int_{-\pi}^{\pi} \log\left(1-\hat{\gamma}^2_{DET}\left(e^{j\theta}\right)\right)\frac{d\theta}{2\pi},$$

demostrando la relación entre el GCS-DET y el GLRT.

Por otro lado, bajo la suposición de baja SNR y $N \to \infty$, el GLRT para testear la correlación espacial de una señal coloreada temporalmente de rango P en ruidos no-IID es

$$\log \mathscr{L} \approx \int_{-\pi}^{\pi} \sum_{i=1}^{P}\left[\log\beta_i\left(e^{j\theta}\right)-\beta_i\left(e^{j\theta}\right)\right]\frac{d\theta}{2\pi}.$$

Haciendo uso del Apéndice B o de [Jorswieck y Boche, 2007, Proposition 2.7] es fácil ver que el integrando del log-GLRT es una función Schur-cóncava en $\beta_i\left(e^{j\theta}\right),\ i=1,\ldots,L$. Con lo cual, el GLRT puede interpretarse como la integral de un GCS definido, hasta factores de normalización para que esté acotado entre 0 y 1, de la siguiente manera

$$\gamma_{\text{GLRT}}\left(e^{j\theta}\right) = \sum_{i=1}^{P}\left[\beta_i\left(e^{j\theta}\right)-\log\beta_i\left(e^{j\theta}\right)\right].$$

Para ilustrar mejor esto, se particulariza este resultado para $P=1$. En este caso el GLRT viene dado por

$$\log \mathscr{L} \approx \int_{-\pi}^{\pi}\left[\log\beta_1\left(e^{j\theta}\right)-\beta_1\left(e^{j\theta}\right)\right]\frac{d\theta}{2\pi}.$$

que, teniendo en cuenta la definición del GCS-MAX, se puede reescribir como

$$\log \mathscr{L} \approx \int_{-\pi}^{\pi}\left[\log\left[(L-1)\hat{\gamma}_{\text{MAX}}\left(e^{j\theta}\right)+1\right]-(L-1)\hat{\gamma}_{\text{MAX}}\left(e^{j\theta}\right)\right]\frac{d\theta}{2\pi}. \tag{5.13}$$

Por lo tanto, interpretando los GLRT como la integral de un GCS, permite derivar otros detectores empleando diferentes estimadores del GCS.

A continuación, se deriva un detector que fue presentado de manera heurística en [Ramírez et al., 2008]; evaluando, posteriormente, sus prestaciones en la sección de resultados. Para ello, teniendo en cuenta que $(L-1)\hat{\gamma}_{MAX}\left(e^{j\theta}\right) \ll 1$, se puede simplificar (5.13) a

$$\log \mathscr{L} \approx -\int_{-\pi}^{\pi} \hat{\gamma}_{\mathrm{MAX}}\left(e^{j\theta}\right) \frac{d\theta}{2\pi}, \tag{5.14}$$

es decir, el GLRT es una función de la integral del GCS-MAX. Finalmente, usando la estima basada en CCA del GCS-MAX, se obtiene

$$\log \mathscr{L} \approx -\int_{-\pi}^{\pi} \hat{\gamma}_{\mathrm{MAX}}\left(e^{j\theta}\right) \frac{d\theta}{2\pi} = -\frac{1}{L-1}\sum_{i=1}^{\kappa}\sum_{k=1}^{L}(\hat{\beta}_i - 1)\int_{-\pi}^{\pi}\left|\mathbf{f}^H\left(e^{j\theta}\right)\hat{\mathbf{w}}_k^{(i)}\right|^2 \frac{d\theta}{2\pi},$$

y, aplicando el Teorema de Parseval, se simplifica a

$$\log \mathscr{L} \approx -\frac{1}{L-1}\sum_{i=1}^{\kappa}(\hat{\beta}_i - 1),$$

donde se ha tenido en cuenta que los autovectores $\hat{\mathbf{w}}^{(i)} = \left[\hat{\mathbf{w}}_1^{(i)T}, \ldots, \hat{\mathbf{w}}_L^{(i)T}\right]^T$ tienen norma unidad. Como se puede ver, el detector se ha simplificado a la suma de los κ mayores autovalores del problema de autovalores (5.11).

5.3.2. Relación con la Información Mutua de L Series Temporales

A pesar de no existir una única definición de la información mutua para más de dos variables aleatorias o procesos estocásticos, sí hay una ampliamente aceptada. Dicha definición está basada en las tasas de entropía (*entropy rates*) y viene dada por [Cover y Thomas, 2006]

$$I(x_1[n], \ldots, x_L[n]) = \sum_{i=1}^{L} \lim_{N\to\infty} \frac{1}{N} H(x_i[0], \ldots, x_i[N-1]) - \lim_{N\to\infty} \frac{1}{N} H(\mathbf{x}[0], \ldots, \mathbf{x}[N-1]), \tag{5.15}$$

donde $\lim_{N\to\infty} H(x_i[0], \ldots, x_i[N-1])/N$ es la tasa de entropía marginal de la serie temporal i-ésima y $\lim_{N\to\infty} H(\mathbf{x}[0], \ldots, \mathbf{x}[N-1])/N$ es la tasa de entropía conjunta del proceso multivariado. A continuación se calculan ambas tasas de entropía. Para la entropía marginal, la demostración se puede encontrar, por ejemplo, en [Gray, 2006]; pero aquí se repite por completitud. En primer lugar se obtiene $H(x_i[0], \ldots, x_i[N-1])$, siendo su valor para procesos estocásticos Gaussianos, complejos y circulares, el siguiente

$$H(x_i[0], \ldots, x_i[N-1]) = N\log(\pi \mathrm{e}) + \log\det\mathcal{R},$$

donde $\mathcal{R} \in \mathbb{C}^{N\times N}$ es la matriz de covarianza del vector $[x_i[0], \ldots, x_i[N-1]]^T$. Dicha matriz de covarianza, para procesos estocásticos estacionarios en sentido amplio, es Toeplitz y viene dada por

$$\mathcal{R} = \begin{bmatrix} r_i[0] & r_i[-1] & \cdots & r_i[-N+1] \\ r_i[1] & r_i[0] & \cdots & r_i[-N+2] \\ \vdots & \vdots & \ddots & \vdots \\ r_i[N-1] & r_i[N-2] & \cdots & r_i[0] \end{bmatrix},$$

donde la secuencia de covarianza es $r_i[m] = E\left[x_i[n]x_i^*[n-m]\right]$. Por lo tanto, la tasa de entropía marginal es

$$\lim_{N\to\infty}\frac{1}{N}H(x_i[0],\ldots,x_i[N-1]) = \log(\pi\mathrm{e}) + \lim_{N\to\infty}\frac{1}{N}\log\det\mathcal{R},$$

y aplicando el Teorema de Szegö para secuencias de matrices Toeplitz, concretamente [Gray, 2006, Theorem 13], puede reescribirse como

$$\lim_{N\to\infty}\frac{1}{N}H(x_i[0],\ldots,x_i[N-1]) = \log(\pi\mathrm{e}) + \int_{-\pi}^{\pi}\log\left[S_i\left(e^{j\theta}\right)\right]\frac{d\theta}{2\pi}, \tag{5.16}$$

donde $S_i\left(e^{j\theta}\right) = \mathcal{F}\left(r_i[m]\right)$. La entropía conjunta de N observaciones de la serie temporal multivariada viene dada por

$$H(\mathbf{x}[0],\ldots,\mathbf{x}[N-1]) = NL\log(\pi\mathrm{e}) + \log\det\mathcal{R},$$

donde la matriz de covarianza es Toeplitz por bloques

$$\mathcal{R} = \begin{bmatrix} \mathbf{R}[0] & \mathbf{R}[-1] & \cdots & \mathbf{R}[-N+1] \\ \mathbf{R}[1] & \mathbf{R}[0] & \cdots & \mathbf{R}[-N+2] \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{R}[N-1] & \mathbf{R}[N-2] & \cdots & \mathbf{R}[0] \end{bmatrix} \in \mathbb{C}^{LN\times LN},$$

con $\mathbf{R}[m] = E\left[\mathbf{x}[n]\mathbf{x}^H[n-m]\right]$, una secuencia de covarianza matricial. La tasa de entropía conjunta es

$$\lim_{N\to\infty}\frac{1}{N}H(\mathbf{x}[0],\ldots,\mathbf{x}[N-1]) = L\log(\pi\mathrm{e}) + \lim_{N\to\infty}\frac{1}{N}\log\det\mathcal{R},$$

y, aplicando [Gutiérrez-Gutiérrez y Crespo, 2008, Theorem 6], se obtiene

$$\lim_{N\to\infty}\frac{1}{N}H(\mathbf{x}[0],\ldots,\mathbf{x}[N-1]) = L\log(\pi\mathrm{e}) + \int_{-\pi}^{\pi}\log\det\left[\mathbf{S}\left(e^{j\theta}\right)\right]\frac{d\theta}{2\pi}, \tag{5.17}$$

donde $\mathbf{S}\left(e^{j\theta}\right) = \mathcal{F}\left(\mathbf{R}[m]\right)$. Finalmente, sustituyendo (5.16) y (5.17) en (5.15), la información mutua viene dada por

$$I(x_1[n],\ldots,x_L[n]) = -\int_{-\pi}^{\pi}\log\left(\frac{\det\left(\mathbf{S}(e^{j\theta})\right)}{\prod_{i=1}^{L}S_i(e^{j\theta})}\right)\frac{d\theta}{2\pi} = -\int_{-\pi}^{\pi}\log\left[\det\left(\mathbf{C}(e^{j\theta})\right)\right]\frac{d\theta}{2\pi},$$

que, teniendo en cuenta la Definición 5.6, se puede expresar como una función del GCS-DET

$$I(x_1[n],\ldots,x_L[n]) = -\int_{-\pi}^{\pi}\log\left(1-\gamma^2_{\mathrm{DET}}(e^{j\theta})\right)\frac{d\theta}{2\pi}.$$

Finalmente, esta interpretación demuestra que el GLRT del modelo NNSSNS, dado por (4.18), es una estima de la información mutua. Por lo tanto, el GLRT es también un test para la independencia de las L series temporales univariadas.

5.4. Resultados Numéricos

Esta sección ilustra las diferencias entre las tres definiciones del GCS propuestas mediante simulaciones. Por otro lado, también compara el rendimiento de los diferentes estimadores del GCS-MAX y, por último, analiza las prestaciones de la estima del GCS-MAX basada en CCA en problemas de detección multicanal.

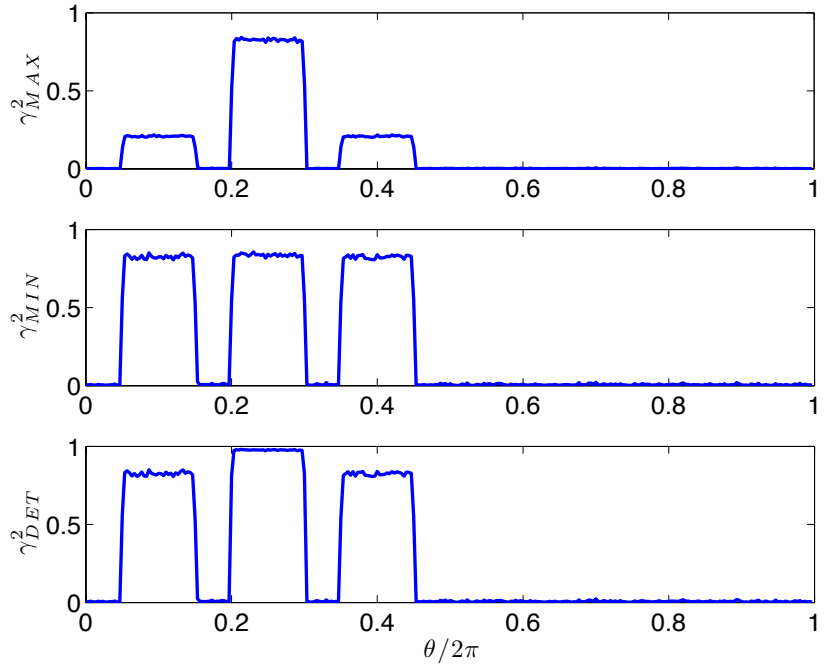

Figura 5.1: Comparativa del GCS-MAX, GCS-MIN y GCS-DET.

5.4.1. Comparación del GCS-MAX, GCS-MIN y GCS-DET

Este primer ejemplo estudia el comportamiento de las tres definiciones del GCS propuestas (MAX, MIN y DET). Para que la comparación sea justa, se emplea el estimador basado en la técnica de Welch [Carter et al., 1973], usando una ventana de Hanning para todas las definiciones. Además, el número de observaciones considerado es elevado, $N = 2^{17}$, al igual que la longitud de la ventana, 600, (con un solape del 50 %) para considerar un caso próximo al teórico, y ver exclusivamente la diferencia entre las definiciones. Para cada definición, el GCS se ha evaluado en $K = 300$ frecuencias normalizadas entre 0 y 1.

Las señales consideradas en este ejemplo son:

$$\begin{bmatrix} x_1[n] \\ x_2[n] \\ x_3[n] \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} s_1[n] \\ s_2[n] \\ s_3[n] \end{bmatrix} + \begin{bmatrix} v_1[n] \\ v_2[n] \\ v_3[n] \end{bmatrix}$$

donde $s_1[n]$, $s_2[n]$, $s_3[n]$ son tres procesos Gaussianos complejos de media nula, potencia unitaria, ancho de banda $BW = 0.1$, y frecuencias[4] centrales 0.1, 0.25 y 0.4, respectivamente. Por otro lado, $v_k[n]$ son procesos independientes e idénticamente distribuidos (IID), que se distribuyen de acuerdo a $\mathcal{CN}(0, 1)$. Es importante señalar que, en la primera banda sólo están correladas $x_1[n]$ y $x_2[n]$, en la segunda banda todas las señales están correladas; y en la tercera banda únicamente están correladas $x_2[n]$ y $x_3[n]$.

La Figura 5.1 muestra las tres estimas del GCS: GCS-MAX, GCS-MIN y GCS-DET. En dicha figura se puede ver que el GCS-MIN y el GCS-DET presentan, aproximadamente, el mismo valor, independientemente del número de señales correladas, aunque con pequeñas diferencias entre ambos. Sin embargo, el GCS-MAX depende fuertemente del número de señales

[4] Las frecuencias se consideran normalizadas, es decir, $\nu = \theta/2\pi$.

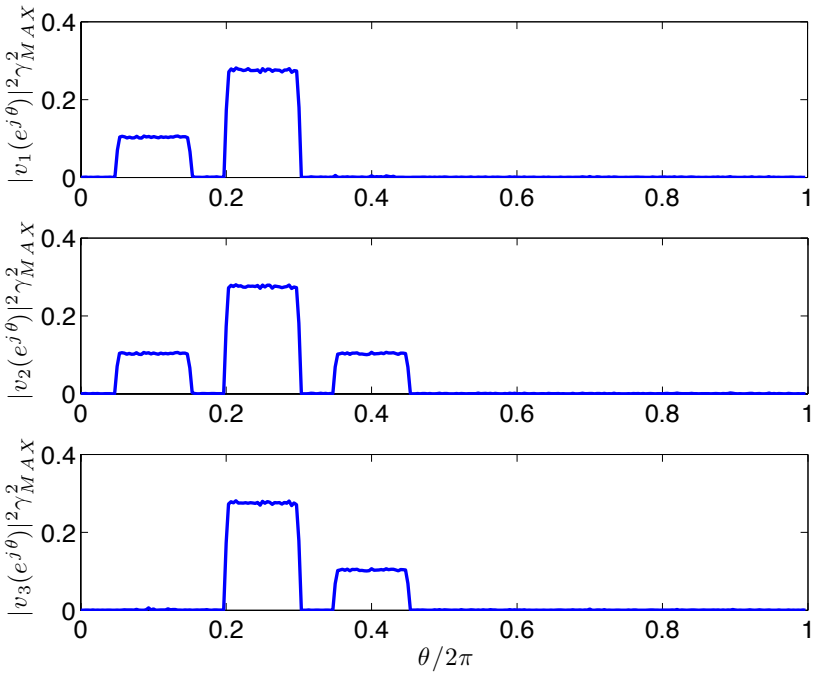

Figura 5.2: Valor de $|v_i(e^{j\theta})|^2\gamma^2_{MAX}(e^{j\theta})$.

correladas en cada banda. De hecho, y como demuestra la Propiedad 5.7, la contribución de cada señal al GCS-MAX se mide con los elementos del autovector asociado. Esto se muestra en la Figura 5.2, donde los valores de las componentes de los autovectores indican claramente este hecho. Concretamente, se observa que en la primera banda sólo $x_1[n]$ y $x_2[n]$ están correladas, lo mismo que ocurre en la tercera banda con $x_2[n]$ y $x_3[n]$. Mientras que en la segunda banda todas las componentes son iguales, dado que todas las señales están correladas.

5.4.2. Comparación de los Diferentes Estimadores del GCS-MAX

A continuación se comparan los estimadores del GCS-MAX basados en Welch, en el MVDR y en CCA mediante dos ejemplos de simulación. En ambas simulaciones se han considerado $N = 1024$ observaciones de $L = 3$ señales diferentes, y el GCS-MAX se ha evaluado en $K = 200$ frecuencias equiespaciadas. Hemos seleccionado 100 como tamaño de las matrices de covarianza cruzada y longitud de los segmentos o filtros, y una ventana de Hanning con un solape del 50 % para la técnica de Welch; el rango elegido es $\kappa = 5$ para la técnica basada en CCA.

En el primer ejemplo se han considerado las siguiente señales

$$x_k[n] = v_k[n] + \sum_{i=1}^{N_f} \frac{1}{\sqrt{2}} e^{j\left(2\pi\nu_k^{(i)} n + \phi_k\right)}, \quad k = 1, \ldots, 3,$$

donde $v_k[n]$ son procesos Gaussianos complejos independientes, de media nula y varianza unidad, el conjunto de frecuencias elegido se muestra en la Tabla 5.2 con $N_f = 5$, siendo las fases ϕ_k variables aleatorias uniformes entre 0 y 2π. Es importante destacar que todas las frecuencias excepto $\nu = 0.152$ coinciden con las frecuencias donde se evalúa el GCS. El

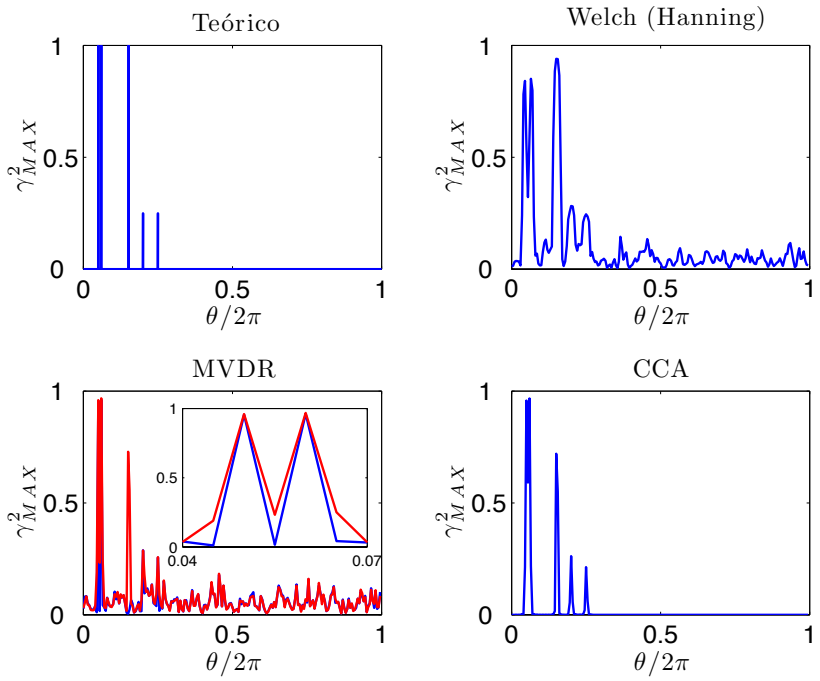

Figura 5.3: Estima del GCS-MAX para el primer ejemplo. Para la estima MVDR se muestra en azul la técnica sin DL y en rojo con DL.

Primer ejemplo	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$
$\nu_1^{(i)}$	0.05	0.06	0.152	0.20	0.25
$\nu_2^{(i)}$	0.05	0.06	0.152	0.20	0.35
$\nu_3^{(i)}$	0.05	0.06	0.152	0.25	0.4

Tabla 5.2: Frecuencias de las exponenciales complejas del primer ejemplo.

GCS-MAX estimado (al cuadrado) se muestra en la Figura 5.3, donde se observan los mejores resultados obtenidos por la técnica basada en CCA, la cual tiene una resolución elevada y, lo más importante, elimina correlaciones espurias que surgen por trabajar con matrices de covarianza estimadas. Asimismo, es importante destacar que la técnica basada en el MVDR posee mayor resolución, sin embargo, su comportamiento se ve severamente degradado para la frecuencia $\nu = 0.152$. Este problema puede resolverse en la práctica mediante técnicas de regularización (*diagonal loading* o DL). Pero el incremento de robustez provoca una pérdida de resolución, como se observa en el *zoom* de la figura. Este problema es equivalente al problema de mal apuntamiento (*mismatch*) en *beamforming* que ha sido estudiado extensamente en [Cox, 1973, Li y Stoica, 2005].

En el segundo ejemplo, las tres señales se generan de acuerdo al modelo

$$x_k[n] = s[n] + v_k[n], \quad k = 1, \ldots, 3,$$

donde $v_k[n]$ son procesos IID distribuidos como $\mathcal{CN}(0,1)$, y la señal común $s[n]$ es un proceso Gaussiano complejo de banda estrecha, media nula, potencia unitaria y banda de paso entre

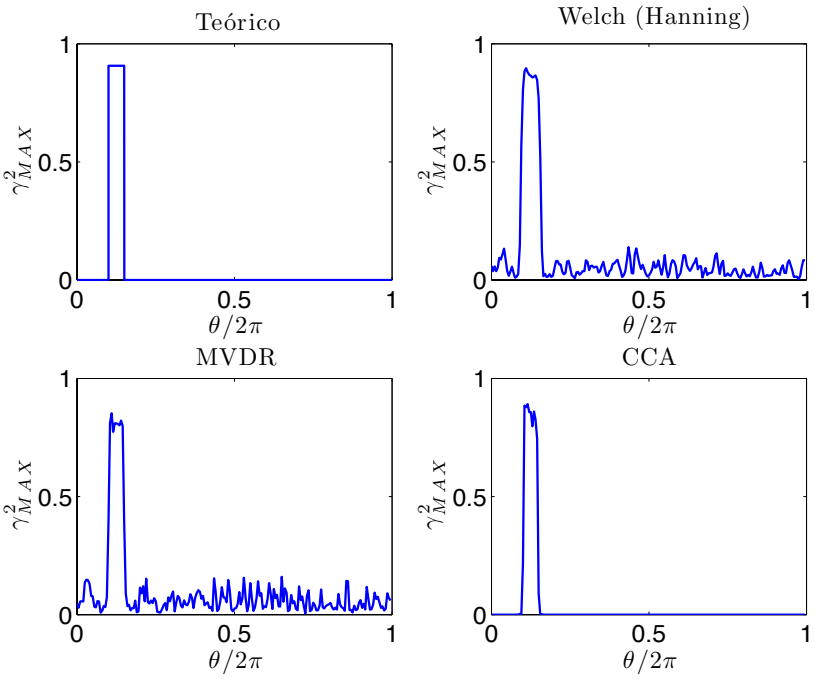

Figura 5.4: Estima del GCS-MAX para el segundo ejemplo.

0.1 y 0.15. Los resultados se muestran en la Figura 5.4 donde, una vez más, se observan las mejores prestaciones de la técnica basada en CCA.

Por último, la Figura 5.5 muestra los autovalores del problema GEV (5.12) obtenidos en los dos ejemplos anteriores. En ambos casos, el número de autovalores dominantes es aproximadamente $\kappa = 5$, justificando la elección del orden realizada previamente para la técnica basada en CCA. Como se observó en [Santamaría y Vía, 2007], el número de autovalores relevantes para la reducción de orden de la técnica de CCA está relacionado con el producto del ancho de banda por el tamaño de la ventana, o con el número de exponenciales complejas en el caso de procesos armónicos; influyendo también el número de observaciones disponibles, N. En la práctica, la elección se basa en el perfil de autovalores (Figura 5.5). Sin embargo, la elección de κ no es crítica para la estima del GCS-MAX. De hecho, como se puede ver en la Figura 5.6, no hay demasiadas diferencias al estimar el GCS-MAX con $\kappa = 5$ o $\kappa = 6$, lo cual es bastante importante dado que la estima del orden es un problema complicado.

5.4.3. El GCS-MAX en Problemas de Detección

Esta sección, usando la interpretación de los GLRT como GCS, evalúa las prestaciones de los dos estimadores del GCS-MAX en problemas de detección multicanal. En el primer ejemplo se ha considerado un experimento con las señales generadas por el modelo (3.47) y los siguientes parámetros: $L = 5$ sensores, $P = 1$, $M = 10$ realizaciones de longitud $N = 20$, SNR $= -8$ dB y $T = q_v = 8$. No obstante, en este ejemplo, la señal $s[n]$ es banda estrecha con su energía concentrada entre 0.1 y 0.15. Las curvas ROC de este experimento se muestran en la Figura 5.7a donde se han comparado las prestaciones de los siguientes detectores:

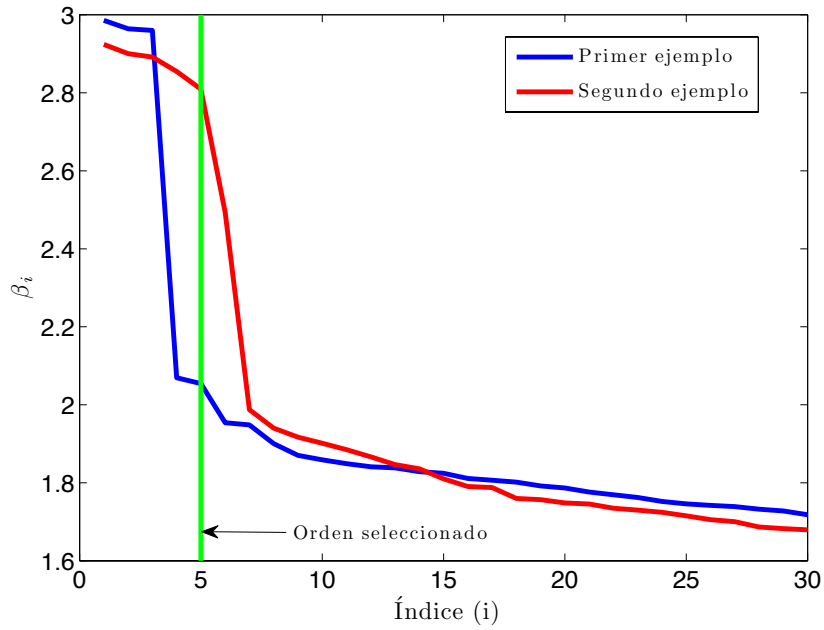

Figura 5.5: Principales autovalores de los ejemplos 1 y 2.

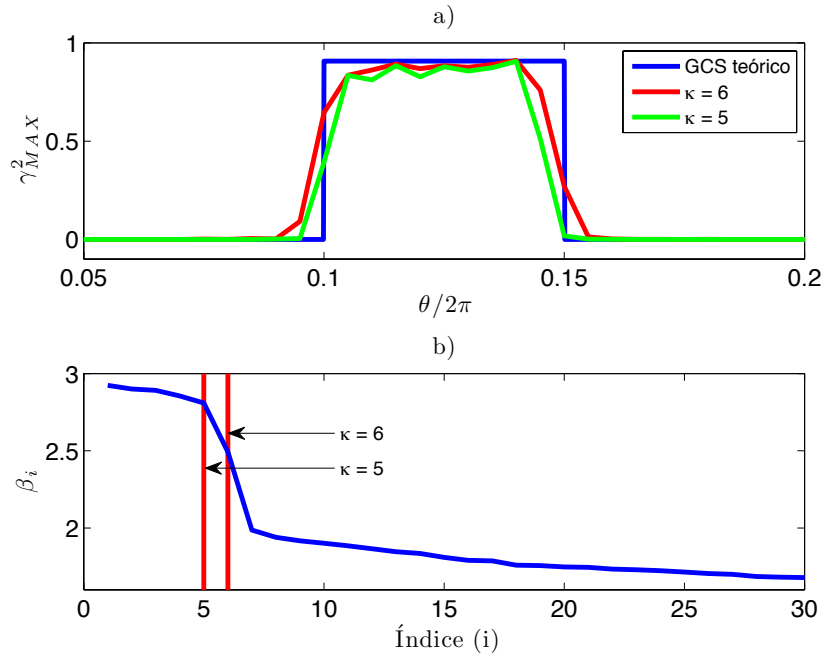

Figura 5.6: a) Estima del GCS-MAX para diferentes ordenes κ; y b) autovalores del problema CCA.

- Freq. Asin-GLRT: GLRT para el modelo NNSSPS bajo la suposición de baja SNR, basado en el siguiente GLR

$$\log \mathscr{L} \approx \int_{-\pi}^{\pi} \left[\log \beta_1 \left(e^{j\theta} \right) - \beta_1 \left(e^{j\theta} \right) \right] \frac{d\theta}{2\pi},$$

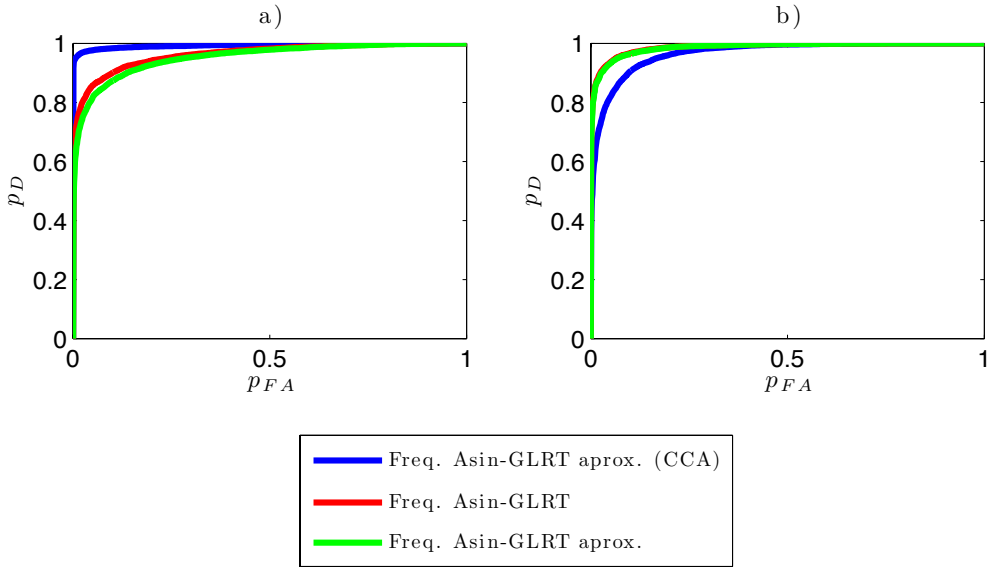

Figura 5.7: Curva ROC de los diferentes detectores para un experimento con: a) señales banda estrecha y b) señales banda ancha. Los parámetros de este ejemplo son: $L = 5$ sensores, $P = 1$, $M = 10$ realizaciones de longitud $N = 100$, SNR $= -8$ dB, $T = q_v = 8$ y $\kappa = 1$.

donde $\beta_1\left(e^{j\theta}\right)$ es el mayor autovalor de la matriz de coherencia muestral.

- Freq. Asin-GLRT aprox.: Aproximación del GLRT anterior, cuyo estadístico es

$$\log \mathscr{L} \approx -\int_{-\pi}^{\pi} \hat{\gamma}_{\text{MAX}}\left(e^{j\theta}\right) \frac{d\theta}{2\pi},$$

donde $\hat{\gamma}_{\text{MAX}}\left(e^{j\theta}\right)$ es el GCS-MAX estimado mediante la técnica de Welch.

- Freq. Asin-GLRT aprox. (CCA): Aproximación del GLRT para el modelo NNSSPS cuando el GCS-MAX se estima mediante la técnica basada en CCA. El estadístico de este test es

$$\log \mathscr{L} \approx -\frac{1}{L-1} \sum_{i=1}^{\kappa} (\hat{\beta}_i - 1),$$

donde $\hat{\beta}_i$ son los autovalores en (5.12). Concretamente, se ha seleccionado $\kappa = 1$; por ello, el test está basado aproximadamente en la frecuencia más correlada.

Para este ejemplo, las mejores prestaciones las obtiene el test basado en CCA, debido a que explota implícitamente la estructura temporal de $s[n]$. Por otro lado, el GLRT y su aproximación presentan prestaciones muy similares.

No obstante, en este mismo escenario, se ha considerado también evaluar las prestaciones de los detectores anteriores cuando la señal $s[n]$ es banda ancha. En este ejemplo, se observa en la Figura 5.7b la pérdida de prestaciones del detector basado en CCA y el mejor comportamiento del GLRT y su aproximación.

5.4.4. Una Aplicación del Espectro de Coherencia Generalizado en Ingeniería Forestal

Esta sección muestra una posible aplicación del espectro de coherencia generalizado. Concretamente, el GCS se usa en el análisis de la variabilidad de la humedad en desechos de madera de diferentes especies de árboles en la zona templada de Europa. Este análisis forma parte de un estudio más amplio que estudia la variabilidad de la humedad y la dinámica estacional mediante diferentes técnicas. La mayor parte de dicho trabajo se encuentra fuera del alcance de la Tesis, por ello, sólo se presentan las señales utilizadas, se obtiene el GCS asociado a dichas señales y se enuncian las principales conclusiones del estudio.

Las medidas se tomaron por un reflectómetro en el dominio del tiempo con una periodicidad de dos semanas en diferentes bosques de los Cárpatos. En concreto, se hicieron medidas de la humedad en el corazón (o duramen) interno y externo y en la albura del tronco de un árbol para diferentes clases de descomposición (DC) y tipos de árboles. Las clases de descomposición se clasifican de la siguiente manera [Korpel, 1989]

- DC1: árboles caídos recientemente, cubiertos por corteza y madera intactas, que permiten una fácil identificación de la especie.
- DC2: árboles con la corteza despedazada o desprendida y madera podrida, cuyas especies aún pueden determinarse.
- DC3: árboles que apenas puede reconocerse su especie, debido al avanzado estado de descomposición.

Las medidas constan de 11 muestras por señal (humedad en el corazón interno y externo y en la albura) para las tres clases de descomposición y cuatro especies de árboles, que son

- Roble Turco o Cabelludo,
- Abeto Plateado o Común,
- Haya Europea,
- Picea de Noruega o Abeto Rojo.

Es importante destacar que el número de observaciones es muy reducido y, por ello, la estima del GCS no será muy precisa. No obstante, es bastante laborioso obtener un conjunto mayor de medidas y se seleccionarán los parámetros del estimador del GCS para obtener la mejor estima posible. En concreto, en la Figura 5.8 se muestra el GCS-MAX estimado mediante la técnica de Welch con ventanas rectangulares de longitud 7 y un solapamiento entre ellas del 50 %. Estas gráficas revelan que las tendencias estacionales de la humedad se componen por oscilaciones de la humedad entre las tres capas del tronco donde se han realizado las medidas. La coherencia a bajas frecuencias se debe al secado sufrido por la madera debido a las variaciones de la temperatura a lo largo del año, y a altas frecuencias es debida a la influencia puntual de la lluvia.

5.5. Conclusiones

El análisis de los detectores derivados bajo la suposición de ruidos no-IID para series temporales bivariadas ha demostrado su relación con el espectro de coherencia (CS). Motivado

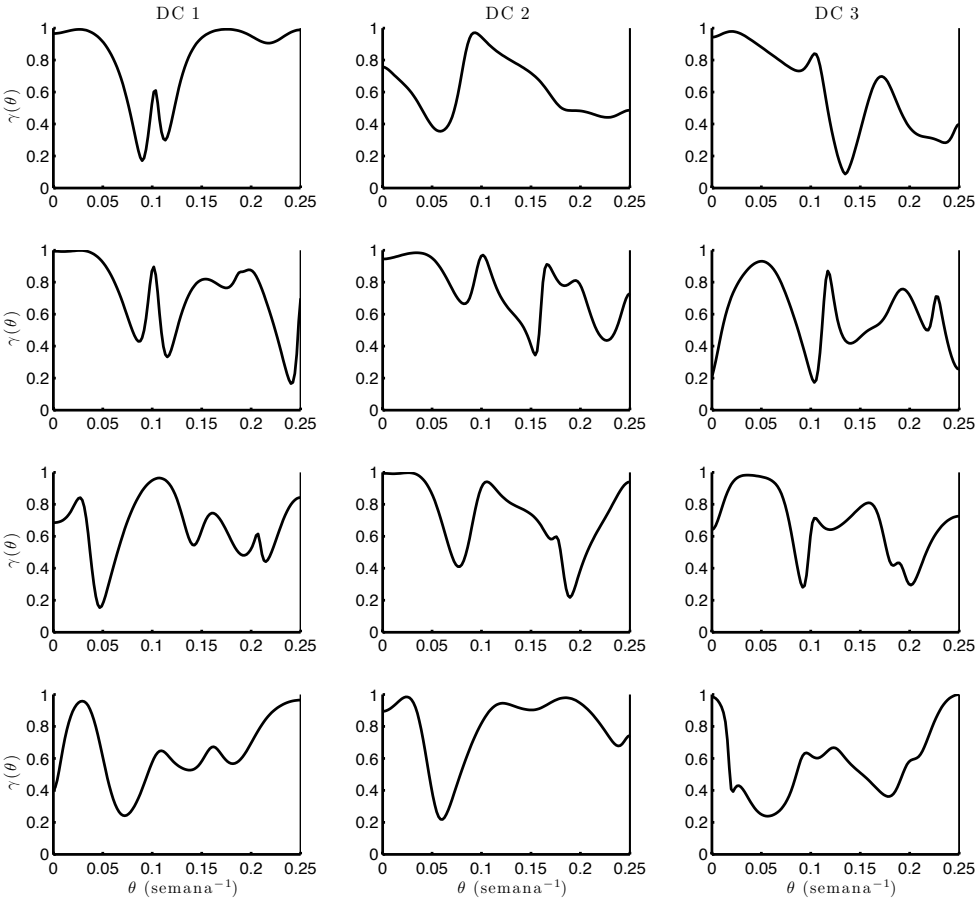

Figura 5.8: GCS para las diferentes especies y clases de descomposición. En cada fila se muestra el GCS para una especie de árbol (Roble Turco, Abeto Plateado, Haya Europea y Picea de Noruega) y cada columna corresponde a una clase de descomposición.

por este hecho, en este capítulo se ha propuesto una generalización del espectro de coherencia (GCS) a múltiples procesos. Concretamente, la generalización propuesta es una función de los autovalores de la matriz que contiene todos los espectros de coherencia por parejas. No obstante, como se ha visto, dicha función debe cumplir una serie de restricciones: tiene que ser Schur-convexa (para mantener la idea intuitiva de correlación), simplificarse al CS para dos series temporales y, además, debe estar acotada entre 0 y 1. Asimismo, se han discutido sus propiedades y se han propuesto dos clases de estimadores. La primera de ellas se basa en estimar los CS mediante técnicas convencionales, y aplicar directamente la definición. Para evitar los problemas presentados por esta técnica, se ha propuesto un estimador basado en el análisis de correlaciones canónicas (CCA), pudiéndose obtener estimadores de rango reducido de manera sencilla en el dominio de las correlaciones canónicas. Las prestaciones de ambos estimadores se han evaluado mediante simulaciones, que ilustran el mejor comportamiento del estimador basado en CCA, debido a la reducción del rango.

Por otro lado, los GLRT derivados bajo la suposición de ruidos no-IID admiten una formulación en términos del GCS, permitiendo aplicar el estimador basado en CCA a problemas de detección y explotar sus ventajas. Esto se ha comprobado mediante simulaciones que ilustran las mejores prestaciones del detector basado en CCA en problemas donde la señal transmitida es banda estrecha. Adicionalmente, la información mutua se puede expresar en términos del GCS-DET, al igual que el GLRT para el modelo NNSSNS, lo que ha permitido interpretar dicho GLRT como una medida de independencia de la serie temporal.

Capítulo 6

Conclusiones y Líneas Futuras

6.1. Conclusiones

Esta Tesis ha considerado el problema de detección de series temporales multivariadas o multicanal contaminadas por ruido. Este es un problema que surge en multitud de aplicaciones que van desde el radar o sonar hasta problemas de ingeniería biomédica o comunicaciones, destacando en este último campo el sensado espectral para radio cognitiva y las redes de sensores. En la mayoría de los casos, este problema de detección resulta en un test de hipótesis compuestas, dado que las verosimilitudes dependen de parámetros desconocidos.

El primer caso considerado es el de ruidos independientes e idénticamente distribuidos (IID). Para este escenario, se ha derivado el test basado en el cociente de verosimilitudes generalizado (GLRT) bajo diferentes estructuras temporales y espaciales. Se ha comenzado por considerar series temporales estacionarias en sentido amplio (WSS) con densidades espectrales de potencia (PSD) planas, con y sin estructura espacial. El GLRT para señales sin estructura espacial ya había sido derivado previamente, pero el GLRT para señales con estructura espacial es novedoso. En ambos casos, el GLRT es una función de los autovalores de la matriz de covarianza muestral. A continuación se han extendido los resultados anteriores, considerando series temporales estacionarias con PSD arbitrarias. La obtención del GLRT para este modelo requiere la estimación de máxima verosimilitud (ML) de matrices de covarianza Toeplitz por bloques, que es un problema no convexo. Debido a esto, se empleó la verosimilitud asintótica, en el dominio frecuencial, y que converge en error cuadrático medio a la verosimilitud convencional. Los GLRT para procesos con PSD arbitrarias, derivados usando la verosimilitud asintótica, se han podido interpretar como la integral del GLRT para los modelos con PSD planas y la misma estructura espacial. Asimismo, también se ha considerado un modelo diferente, donde una señal coloreada se propaga por un canal plano. Para este modelo no existe solución cerrada para el GLRT, siendo necesario recurrir a soluciones numéricas. En concreto, se propuso una técnica que puede interpretarse como un método de potencias para la obtención de un autovector, en el que la matriz varía entre iteraciones. El último modelo resuelto considera señales no estacionarias y, en esta ocasión, no es necesario recurrir a la verosimilitud asintótica dado que las matrices de covarianza no son Toeplitz. Todos estos GLRT han sido evaluados por medio de simulaciones numéricas, mostrando la gran mejora que se produce al explotar completamente la estructura espacio-temporal de las señales.

La segunda parte de la Tesis considera la detección de series temporales contaminadas por ruidos independientes, pero no idénticamente distribuidos (no-IID). Al igual que en el caso anterior, primero se han considerados series temporales con PSD planas, con y sin

estructura espacial. Para el caso de estructura espacial adicional, no es posible derivar una fórmula cerrada para el GLRT. Debido a esto, se han propuesto dos soluciones: (i) optimizar numéricamente la verosimilitud mediante una minimización alternada, y (ii) considerar una relación señal a ruido (SNR) baja, siendo posible obtener, en este caso, una solución cerrada para el GLRT. Posteriormente, se extendieron estos GLRT, haciendo uso de la verosimilitud asintótica, a series temporales con PSD arbitrarias y, por último, también se consideraron series temporales no estacionarias. Las prestaciones de los GLRT se evaluaron por medio de simulaciones numéricas demostrando, una vez más, la importancia de explotar completamente la estructura espacio-temporal. Además se ha mostrado que los GLRT para ruidos no-IID son más robustos que los tests para ruidos IID y, por lo tanto, desde un punto de vista práctico, es preferible usarlos.

Particularizando a series temporales bivariadas los GLRT derivados para series temporales con PSD arbitrarias y ruidos no-IID, se demostró que éstos son funciones del espectro de coherencia. Motivados por este hecho, hemos propuesto una generalización del espectro de coherencia (GCS) a más de dos señales. Concretamente, se ha definido el GCS como una función Schur-convexa de los autovalores de una matriz compuesta por todos los espectros de coherencia por parejas. Además, se derivaron propiedades y se desarrollaron estimadores del GCS. Por otro lado, se relacionó el GCS con la información mutua de más de dos procesos estocásticos y con los GLRT para ruidos no-IID. Esta relación ha permitido obtener tests para el problema de detección multicanal usando diferentes estimadores del GCS, que presentan comportamientos distintos al GLRT. El comportamiento del GCS, así como, las prestaciones de los diferentes estimadores se evaluaron por simulación, mostrando las mejores prestaciones del estimador del GCS basado en el análisis de correlaciones canónicas (CCA), que presenta una alta resolución y elimina las correlaciones espurias. Finalmente, la utilidad del GCS se ha demostrado en un problema de ingeniería forestal.

6.2. Líneas Futuras de Investigación

6.2.1. Extensiones para otros modelos

Esta Tesis ha derivado el test basado en el cociente de verosimilitudes generalizado (GLRT) para la detección de series temporales multivariadas bajo diferentes suposiciones acerca de sus estructuras espaciales y temporales. No obstante, hay ciertos casos con gran interés en el campo de las comunicaciones no considerados en este trabajo. Concretamente, no se han analizado problemas donde las densidades espectrales de potencia (PSD) de las series temporales admiten un modelo paramétrico, como por ejemplo: modelos autoregresivos (AR), de media móvil (MA) y autoregresivos de media móvil.

Un ejemplo típico en comunicaciones inalámbricas es el siguiente. Consideremos un conjunto de L sensores que adquieren la señal transmitida por un transmisor con P antenas, o P transmisores con una antena compartiendo recursos. Suponiendo una señal transmitida banda ancha que atraviesa un canal de múltiples entradas y salidas (MIMO) selectivo en frecuencia [Paulraj et al., 2003], el modelo de señal es

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

donde $\mathbf{H}[\tau] \in \mathbb{C}^{L\times P}$ es la respuesta impulsional del canal MIMO, $\mathbf{s}[n] \in \mathbb{C}^{P}$ es la señal transmitida y $\mathbf{v}[n] \in \mathbb{C}^{L}$ es un ruido aditivo. Aplicando, una vez más, la verosimilitud asintó-

tica, las estimas de máxima verosimilitud (ML) de los parámetros desconocidos se obtienen resolviendo el siguiente problema de optimización

$$\begin{aligned}
\underset{\substack{\mathbf{H}[0],\ldots,\mathbf{H}[T-1]\\ \mathbf{A}_v[0],\ldots,\mathbf{A}_v[q_v-1]}}{\text{minimizar}} \quad & \int_{-\pi}^{\pi} \log\det\left(\mathbf{S}\left(e^{j\theta}\right)\right)\frac{d\theta}{2\pi} + \int_{-\pi}^{\pi} \operatorname{tr}\left(\mathbf{S}^{-1}\left(e^{j\theta}\right)\hat{\mathbf{S}}\left(e^{j\theta}\right)\right)\frac{d\theta}{2\pi} \\
\text{sujeto a} \quad & \mathbf{S}\left(e^{j\theta}\right) = \mathbf{H}\left(e^{j\theta}\right)\mathbf{H}^H\left(e^{j\theta}\right) + \mathbf{\Sigma}^2\left(e^{j\theta}\right), \\
& \mathbf{H}\left(e^{j\theta}\right) = \sum_{n=0}^{T-1}\mathbf{H}[n]e^{-j\theta n}, \\
& \mathbf{\Sigma}^2\left(e^{j\theta}\right) = \left|\sum_{n=0}^{q_v-1}\mathbf{A}_v[n]e^{-j\theta n}\right|^2, \\
& \mathbf{A}_v[n] \in \mathbb{D},
\end{aligned}$$

donde $\mathbb{D}$ es el conjunto de matrices diagonales, se ha supuesto una señal blanca espacial y temporalmente, $\mathbf{R}_s\left(e^{j\theta}\right) = \mathbf{I}$, y ruidos independientes para cada antena, modelados como procesos MA; no obstante, la formulación para procesos ARMA o AR es equivalente. Nótese que, a diferencia de los modelos considerados en la Tesis, no es posible hacer la estimación independientemente para cada frecuencia. Por ello, (6.1) es un problema de optimización no convexo muy complicado, siendo necesario recurrir a aproximaciones para resolverlo, de manera similar a la derivación del GLRT para el modelo INMSPS.

6.2.2. Obtención del UMPIT y del LMPIT

Todos los tests derivados en la Tesis están basados en el GLRT que, como ya se mencionó, es un detector muy popular para tests de hipótesis compuestas y, además, suele resultar en detectores con fácil interpretación y sencillos de implementar. Sin embargo, existen multitud de detectores para tests con hipótesis compuestas (ver Capítulo 2) que pueden ser considerados para resolver los problemas de detección presentados en la Tesis. Entre ellos, el test uniformemente más potente e invariante (UMPIT) y el test localmente más potente e invariante merecen una mención especial.

La aproximación clásica para derivar el UMPIT es la siguiente [Scharf, 1991]:

1. Identificar las invarianzas del problema y describirlas como un grupo de transformaciones.

2. Elegir el estadístico máximamente invariante.

3. Derivar la distribución del estadístico máximamente invariante bajo ambas hipótesis.

4. Obtener el detector basado en el cociente de verosimilitudes del estadístico máximamente invariante.

Si el cociente de verosimilitudes (o una función de él que no varíe el estadístico) no depende de los parámetros desconocidos se habrá derivado el UMPIT. Por el contrario, si depende de parámetros desconocidos, aún es posible derivar el LMPIT. Para ello, se realiza una aproximación en serie de Taylor del cociente de verosimilitudes y, si no depende de los parámetros, se habrá obtenido el LMPIT.

La mayor complicación que presenta el procedimiento anterior es la derivación de las distribuciones del estadístico máximamente invariante siendo, en general, un problema extremadamente complicado. No obstante, existe una manera alternativa para derivar el UMPIT,

basada en el teorema de Wijsman [Wijsman, 1967]; que a pesar de su potencial no es ampliamente conocido en el campo del procesado de señal [Gabriel y Kay, 2002]. Concretamente, el teorema de Wijsman permite obtener el cociente de verosimilitudes del estadístico máximamente invariante sin derivar las distribuciones. A continuación, se enuncia el teorema de Wijsman.

Teorema 6.1 (Wijsman)**.** *Bajo una serie de condiciones [Wijsman, 1985], el cociente de verosimilitudes del estadístico máximamente invariante viene dado por*

$$\frac{\int_G p\left(g(\mathbf{x}); \boldsymbol{\theta}_1\right) \xi_0(g) d\nu_l(g)}{\int_G p\left(g(\mathbf{x}); \boldsymbol{\theta}_0\right) \xi_0(g) d\nu_l(g)} \tag{6.1}$$

donde $g(\cdot)$ es un elemento del grupo de transformaciones G, $\xi_0(g) = \det^{-1}\left(J_{g^{-1}}(\mathbf{x})\right)$ es el inverso del Jacobiano [Papoulis y Pillai, 2002] y $\nu_l(g)$ denota una medida invariante por la izquierda para el grupo G.

Resumiendo, (6.1) permite derivar el cociente de verosimilitudes del estadístico máximamente invariante, necesario para obtener el UMPIT o el LMPIT, sin calcular las distribuciones del estadístico.

6.2.3. Evaluación de las prestaciones sobre una plataforma MIMO Hardware

Esta Tesis ha considerado, únicamente, un estudio teórico del problema de detección multicanal. No obstante, es importante la evaluación de las prestaciones de los detectores presentados en sistemas reales. Para ello, a partir de la experiencia adquirida en [Ramírez et al., 2006, García-Naya et al., 2007, Ramírez et al., 2008, González et al., 2011] se evaluarán las prestaciones de los diferentes tests usando una plataforma MIMO Hardware.

En concreto, la plataforma disponible en el grupo de tratamiento avanzado de señal (GTAS) consta de tres nodos Lyrtech de altas prestaciones equipados con cuatro antenas y capacidades de comunicación semi-dúplex. Asimismo, en [Vielva et al., 2010, Gutiérrez et al., 2011] se ha desarrollado un software para el control remoto de dicho hardware, permitiendo su manejo de manera sencilla y centralizada en un único PC. Por otro lado, dicha plataforma va a ser ampliada con la adquisición de cinco nodos USRP2 con dos antenas y prestaciones algo inferiores a los anteriores.

Con el hardware disponible, descrito anteriormente, se propone el siguiente escenario:

- Una red primaria compuesta por los tres nodos de altas prestaciones,
- y una red secundaria formada por los cinco nodos de bajas prestaciones.

Para este escenario se medirán las prestaciones de las diferentes técnicas presentadas y se discutirán las ventajas o inconvenientes de todos los modelos teóricos considerados (procesos blancos o coloreados, estructura espacial de rango reducido, ruidos con diferentes distribuciones). Por lo tanto, se comprobará, en un escenario práctico, si el coste computacional de los GLRT para modelos más complicados está justificado en estos casos.

6.2.4. Otras aplicaciones del GCS

En la Tesis se ha considerado la aplicación del GCS a un problema de ingeniería forestal y a radio cognitiva. No obstante, existen multitud de aplicaciones donde es necesario comprimir la información correspondiente a la estructura espacial en una única medida.

Por ejemplo, en [Luesma, 2006] se aplica el espectro de coherencia a un problema de bioingeniería, analizando las diferentes coherencias por parejas entre las señales correspondientes a:

- la variabilidad del ritmo cardíaco,
- la respiración,
- y la variabilidad de la presión sanguínea.

Concretamente, a partir de los tres CS se extraen diversas conclusiones. No obstante, creemos que el análisis del GCS en este problema puede proporcionar un punto de vista interesante.

6.2.5. Estimación multicanal

Finalmente, como en muchos problemas de procesado de señal, la detección constituye únicamente un primer paso, siendo necesario estimar posteriormente la señales transmitidas o algunos parámetros que las caracterizan. Muchos problemas de detección ya han sido considerados, de manera indirecta, en al Tesis al derivar el GLRT. No obstante, la estimación multicanal es un campo muy amplio y dichos trabajos constituyen únicamente una pequeña parte del mismo.

Por ejemplo, en [Washizawa et al., 2010] se considera la estimación de una señal común observada mediante sensores con diferentes ganancias y ruidos con diferentes energías. Este problema se corresponde perfectamente con el modelos NNWSPS considerado en la Tesis, para el cual se obtuvieron, de manera iterativa, las estimas ML de las ganancias y de las potencias de los ruidos. Por lo tanto, una linea futura en este ámbito consiste en evaluar las prestaciones en problemas de estimación de los estimadores ML propuestos y su comparación con otras técnicas propuestas en la literatura. Asimismo, será necesario derivar estimadores de la señal. Por otro lado, en [Schizas et al., 2007, Cui et al., 2007, Xiao et al., 2008] se considera el problema de la estimación multi-sensor distribuida.

Apéndice A

Demostración del Teorema 3.7

Este apéndice demuestra la convergencia en media cuadrática de las log-verosimilitudes, que puede interpretarse como una extensión de verosimilitud de Whittle [Whittle, 1952, Whittle, 1962] a procesos multivariados.

La log-verosimilitud (3.26) puede reescribirse como

$$\log p\left(\mathbf{z}_0, \cdots, \mathbf{z}_{M-1}; \mathcal{R}\right) = -LNM \log \pi - M \log \det\left(\mathcal{R}\right) - \sum_{m=0}^{M-1} \mathbf{z}_m^H \mathcal{R}^{-1} \mathbf{z}_m, \tag{A.1}$$

donde $\mathcal{R}$ está definida en (3.24). Adicionalmente, se define la matriz Toeplitz por bloques $\tilde{\mathcal{R}}$, dada por

$$\tilde{\mathcal{R}} = \begin{bmatrix} \tilde{\mathbf{R}}[0] & \tilde{\mathbf{R}}[-1] & \cdots & \tilde{\mathbf{R}}[-N+1] \\ \tilde{\mathbf{R}}[1] & \tilde{\mathbf{R}}[0] & \cdots & \tilde{\mathbf{R}}[-N+2] \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{\mathbf{R}}[N-1] & \tilde{\mathbf{R}}[N-2] & \cdots & \tilde{\mathbf{R}}[0] \end{bmatrix}, \tag{A.2}$$

donde $\tilde{\mathbf{R}}[n] = \mathcal{F}^{-1}\left[\mathbf{S}^{-1}\left(e^{j\theta}\right)\right]$; y la suma de formas cuadráticas de $\tilde{\mathcal{R}}$

$$\sum_{m=0}^{M-1} \mathbf{z}_m^H \tilde{\mathcal{R}} \mathbf{z}_m = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \sum_{n'=0}^{N-1} \mathbf{x}_m^H[n] \tilde{\mathbf{R}}\left[n - n'\right] \mathbf{x}_m[n'],$$

que, teniendo en cuenta la definición de $\tilde{\mathbf{R}}[n] = \mathcal{F}^{-1}\left[\mathbf{S}^{-1}\left(e^{j\theta}\right)\right]$, es

$$\begin{aligned} \sum_{m=0}^{M-1} \mathbf{z}_m^H \tilde{\mathcal{R}} \mathbf{z}_m &= \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \sum_{n'=0}^{N-1} \int_{-\pi}^{\pi} \mathbf{x}_m^H[n] \mathbf{S}^{-1}\left(e^{j\theta}\right) \mathbf{x}_m[n'] e^{j(n-n')\theta} \frac{d\theta}{2\pi} \\ &= NM \int_{-\pi}^{\pi} \operatorname{tr}\left(\mathbf{S}^{-1}\left(e^{j\theta}\right) \hat{\mathbf{S}}\left(e^{j\theta}\right)\right) \frac{d\theta}{2\pi}. \end{aligned} \tag{A.3}$$

Por otro lado, la convergencia en media cuadrática de una variable se define [Grimmett y Stirzaker, 1992]

$$\lim_{N\to\infty} E\left[\left|\frac{1}{N}\left[\log p\left(\mathbf{z}_0, \ldots, \mathbf{z}_{M-1}; \mathcal{R}\right) - \log p\left(\mathbf{z}_0, \ldots, \mathbf{z}_{M-1}; \mathbf{S}\left(e^{j\theta}\right)\right)\right]\right|^2\right] = 0, \tag{A.4}$$

y sustituyendo (3.27) y (A.1) en (A.4), se obtiene

$$\begin{aligned}\lim_{N\to\infty} E\left[\left|\frac{1}{N}\left[-M\log\det(\mathcal{R})-\sum_{m=0}^{M-1}\mathbf{z}_m^H\mathcal{R}^{-1}\mathbf{z}_m+NM\int_{-\pi}^{\pi}\log\det\left(\mathbf{S}\left(e^{j\theta}\right)\right)\frac{d\theta}{2\pi}\right.\right.\right.\\ \left.\left.\left.-NM\int_{-\pi}^{\pi}\operatorname{tr}\left(\mathbf{S}^{-1}\left(e^{j\theta}\right)\hat{\mathbf{S}}\left(e^{j\theta}\right)\right)\frac{d\theta}{2\pi}\right]\right|^2\right]\\ \overset{(a)}{\leq}\lim_{N\to\infty}2M\left|\frac{1}{N}\log\det(\mathcal{R})-\int_{-\pi}^{\pi}\log\det\left(\mathbf{S}\left(e^{j\theta}\right)\right)\frac{d\theta}{2\pi}\right|^2\\ +\lim_{N\to\infty}2E\left[\left|\frac{1}{N}\sum_{m=0}^{M-1}\mathbf{z}_m^H\mathcal{R}^{-1}\mathbf{z}_m-M\int_{-\pi}^{\pi}\operatorname{tr}\left(\mathbf{S}^{-1}\left(e^{j\theta}\right)\hat{\mathbf{S}}\left(e^{j\theta}\right)\right)\frac{d\theta}{2\pi}\right|^2\right] \quad \text{(A.5)}\end{aligned}$$

donde (a) está demostrado en [Grimmett y Stirzaker, 1992, Th.8, p. 287]. La demostración se divide en dos partes, una para cada término en la parte derecha de la ecuación (A.5). Aplicando [Gutiérrez-Gutiérrez y Crespo, 2008, Th. 6] se obtiene

$$\lim_{N\to\infty}\frac{1}{N}\log\det(\mathcal{R})-\int_{-\pi}^{\pi}\log\det\left(\mathbf{S}\left(e^{j\theta}\right)\right)\frac{d\theta}{2\pi}=0,$$

y, teniendo en cuenta que

$$\lim_{N\to\infty}a_N=0 \quad\Rightarrow\quad \lim_{N\to\infty}|a_N|^2=0,$$

donde a_N es cualquier secuencia de números reales, se puede demostrar

$$\lim_{N\to\infty}2M\left|\frac{1}{N}\log\det(\mathcal{R})-\int_{-\pi}^{\pi}\log\det\left(\mathbf{S}\left(e^{j\theta}\right)\right)\frac{d\theta}{2\pi}\right|^2=0. \quad \text{(A.6)}$$

Aplicando ahora (A.3), se obtiene

$$\begin{aligned}\lim_{N\to\infty}E\left[\left|\frac{1}{N}\sum_{m=0}^{M-1}\mathbf{z}_m^H\mathcal{R}^{-1}\mathbf{z}_m-M\int_{-\pi}^{\pi}\operatorname{tr}\left(\mathbf{S}^{-1}\left(e^{j\theta}\right)\hat{\mathbf{S}}\left(e^{j\theta}\right)\right)\frac{d\theta}{2\pi}\right|^2\right]\\ =\lim_{N\to\infty}E\left[\left|\frac{1}{N}\sum_{m=0}^{M-1}\mathbf{z}_m^H\left(\mathcal{R}^{-1}-\tilde{\mathcal{R}}\right)\mathbf{z}_m\right|^2\right]. \quad \text{(A.7)}\end{aligned}$$

A continuación se demuestra que (A.7) converge a cero. Primero, nótese que es la esperanza matemática de la suma de formas cuadráticas, y teniendo en cuenta que las formas cuadráticas están incorreladas y la media y la varianza derivadas en [Dzhaparidze, 1986], (A.7) es

$$\begin{aligned}\lim_{N\to\infty}E\left[\left|\frac{1}{N}\sum_{m=0}^{M-1}\mathbf{z}_m^H\left(\mathcal{R}^{-1}-\tilde{\mathcal{R}}\right)\mathbf{z}_m\right|^2\right]=\\ \lim_{N\to\infty}\frac{1}{N^2}\left\{M^2\left(\operatorname{tr}\left[\mathcal{R}\left(\mathcal{R}^{-1}-\tilde{\mathcal{R}}\right)\right]\right)^2+2M\operatorname{tr}\left[\mathcal{R}\left(\mathcal{R}^{-1}-\tilde{\mathcal{R}}\right)\mathcal{R}\left(\mathcal{R}^{-1}-\tilde{\mathcal{R}}\right)\right]\right\}=0, \quad \text{(A.8)}\end{aligned}$$

donde se ha aplicado [Gutiérrez-Gutiérrez y Crespo, 2008, Th. 3 y Th. 5]. La demostración concluye teniendo en cuenta (A.6) y (A.8). □

Apéndice B
Mayorización

Este apéndice presenta una breve introducción a la teoría de la mayorización [Marshall y Olkin, 1979, Jorswieck y Boche, 2007].

B.1. Definición

Es importante destacar que la mayorización es un ordenamiento parcial sobre vectores. Esto es importante, dado que no es posible ordenar cualquier conjunto de vectores. A continuación se presenta su definición:

Definición B.1 (Mayorización)**.** *Dados dos vectores,* $\mathbf{x}, \mathbf{y} \in \mathcal{R}^{L\times 1}$*, con sus componentes ordenadas en orden decreciente, es decir,*

$$x_1 \geq x_2 \geq \cdots \geq x_L \geq 0,$$
$$y_1 \geq y_2 \geq \cdots \geq y_L \geq 0,$$

se dice que $\mathbf{x}$ *mayoriza a* $\mathbf{y}$ *(o* $\mathbf{y}$ *es mayorizado por* $\mathbf{x}$*) y se denota por*

$$\mathbf{x} \succeq \mathbf{y},$$

si

$$\sum_{k=1}^{K} x_k \geq \sum_{k=1}^{K} y_k, \quad K = 1, \ldots, L-1,$$

y

$$\sum_{k=1}^{L} x_k = \sum_{k=1}^{L} y_k.$$

La interpretación de la mayorización es la siguiente: $\mathbf{x}$ mayoriza a $\mathbf{y}$ si las componentes de $\mathbf{x}$ están menos dispersas que las de $\mathbf{y}$. A continuación se presentan dos ejemplos para ilustrar esta definición. En el primero se consideran los siguientes vectores

$$\mathbf{x} = [2.2, 1, 0.9, 0.6, 0.3]^T,$$
$$\mathbf{y} = [2.2, 1, 0.6, 0.6, 0.6]^T,$$

y es fácil comprobar que $\mathbf{x} \succeq \mathbf{y}$. Sin embargo, considerando

$$\mathbf{x} = [0.6, 0.25, 0.15]^T,$$
$$\mathbf{y} = [0.55, 0.4, 0.05]^T,$$

no se puede establecer que $\mathbf{x}$ mayoriza a $\mathbf{y}$, ni al contrario. Esto es debido a que, como se ha dicho antes, la mayorización sólo proporciona un ordenamiento parcial. A continuación se presenta un teorema que asegura la mayorización entre dos vectores.

Teorema B.1. *Una condición necesaria y suficiente para que* $\mathbf{x} \succeq \mathbf{y}$ *es que exista una matriz doblemente estocástica* $\mathbf{P}$*, tal que*

$$\mathbf{y} = \mathbf{P}\mathbf{x},$$

donde una matriz doblemente estocástica es aquella que cumple

$$\begin{aligned} [\mathbf{P}]_{i,j} &\geq 0, \quad i,j = 1,\ldots,L, \\ \sum_{i=1}^{L} [\mathbf{P}]_{i,j} &= 1, \quad j = 1,\ldots,L, \\ \sum_{j=1}^{L} [\mathbf{P}]_{i,j} &= 1, \quad i = 1,\ldots,L, \end{aligned}$$

es decir, sus filas y columnas pueden interpretarse como probabilidades.

Finalmente, se introducen dos vectores de gran interés en mayorización

$$\begin{aligned} \mathbf{1}_L &= [1,1,\ldots,1]^T, \\ \mathbf{x} &= [L,0,\ldots,0]^T. \end{aligned}$$

Este interés se debe a que el vector $\mathbf{1}_L$ es mayorizado por cualquier otro vector, y el vector $\mathbf{x}$ mayoriza a cualquier vector.

B.2. Funciones Schur-convexas y Schur-cóncavas

En esta sección se introduce un conjunto muy particular de funciones, las funciones Schur-convexas y Schur-cóncavas:

Definición B.2 (Función Schur-convexa)**.** *Una función* $f(\cdot)$ *real definida en* $A \subset \mathcal{R}^{L\times 1}$ *es Schur-convexa si cumple*

$$f(\mathbf{x}) \geq f(\mathbf{y}),$$

para

$$\mathbf{x} \succeq \mathbf{y}.$$

Definición B.3 (Función Schur-cóncava)**.** *Una función* $f(\cdot)$ *real definida en* $A \subset \mathcal{R}^{L\times 1}$ *es Schur-cóncava si cumple*

$$f(\mathbf{x}) \leq f(\mathbf{y}),$$

para

$$\mathbf{x} \succeq \mathbf{y}.$$

Las funciones Schur-convexas poseen una propiedad interesante para medir correlaciones, y es que mantienen el ordenamiento. Es evidente, que si el vector de autovalores de una matriz de covarianza mayoriza al vector de autovalores de otra matriz

$$\boldsymbol{\lambda}_a \succeq \boldsymbol{\lambda}_b,$$

la primera matriz corresponde a procesos más correlados o linealmente dependientes [Jorswieck y Boche, 2007]. Sin embargo, el vector de autovalores tiene L componentes que proporcionan información. Entonces, para compactar dicha información en una única medida, lo más apropiado es aplicar una función que preserve dicho ordenamiento, es decir, una función Schur-convexa.

B.2.1. Condición de Schur

Sea $\mathcal{I} \subset \mathbb{R}$ un intervalo abierto y $f : \mathcal{I}^L \to \mathbb{R}$ una función diferenciable continuamente. Las condiciones necesarias y suficientes para que f sea Schur-convexa en $\mathcal{I}^L$ son

1. f es simétrica en $\mathcal{I}^L$,
2. y se cumple
$$(x_i - x_j)\left(\frac{\partial f(\mathbf{x})}{\partial x_i} - \frac{\partial f(\mathbf{x})}{\partial x_j}\right) \geq 0, \qquad 1 \leq i, j \leq L,$$

donde una función es simétrica si el argumento se puede permutar sin variar el valor de la función.

Una condición derivada a partir de la condición de Schur es la siguiente. Sea $\mathcal{I} \subset \mathbb{R}$ un intervalo abierto y $g : \mathcal{I} \to \mathbb{R}$ una función convexa (cóncava) doblemente diferenciable. Entonces
$$f(\mathbf{x}) = \sum_{i=1}^{L} g(x_i)$$
es Schur-convexa (cóncava) en $\mathcal{I}^L$.

Apéndice C

Análisis de correlaciones canónicas

Este apéndice presenta un breve resumen del análisis de correlaciones canónicas (CCA), comenzando por su formulación más sencilla, consistente en la maximización de la correlación entre proyecciones de dos conjuntos de datos reales [Hotelling, 1936]. A partir de esta formulación se presenta la extensión para la obtención de varias soluciones CCA y se introducen diversas generalizaciones del problema a varios conjuntos de datos, las cuales tuvieron su origen en los trabajos de Kettenring [Kettenring, 1971] y Gifi [Gifi, 1990].

C.1. Formulación Clásica

Dadas dos matrices de datos $\mathbf{X}_1 \in \mathbb{R}^{N\times m_1}$ y $\mathbf{X}_2 \in \mathbb{R}^{N\times m_2}$ completas en rango de columnas, el análisis de correlaciones canónicas (CCA) se puede definir como el problema de encontrar dos vectores canónicos: $\mathbf{h}_1$ de dimensiones $m_1 \times 1$ y $\mathbf{h}_2$ de dimensiones $m_2 \times 1$, tales que se maximice la correlación entre las variables canónicas $\mathbf{z}_1 = \mathbf{X}_1\mathbf{h}_1$ y $\mathbf{z}_2 = \mathbf{X}_2\mathbf{h}_2$, es decir,

$$\underset{\mathbf{h}_1,\mathbf{h}_2}{\arg\ \text{máx}}\ \rho = \frac{\mathbf{z}_1^T\mathbf{z}_2}{\|\mathbf{z}_1\|\|\mathbf{z}_2\|} = \frac{\mathbf{h}_1^T\mathbf{R}_{12}\mathbf{h}_2}{\sqrt{\mathbf{h}_1^T\mathbf{R}_{11}\mathbf{h}_1\mathbf{h}_2^T\mathbf{R}_{22}\mathbf{h}_2}}, \tag{C.1}$$

donde $\mathbf{R}_{kl} = \mathbf{X}_k^T\mathbf{X}_l$ es una estima de la matriz de covarianza cruzada. El problema (C.1) es equivalente a maximizar la correlación canónica

$$\rho = \mathbf{h}_1^T\mathbf{R}_{12}\mathbf{h}_2, \tag{C.2}$$

sujeta a la restricción

$$\mathbf{h}_1^T\mathbf{R}_{11}\mathbf{h}_1 = \mathbf{h}_2^T\mathbf{R}_{22}\mathbf{h}_2 = 1, \tag{C.3}$$

la cual, en el caso de dos conjuntos de datos, se puede sustituir por la restricción equivalente

$$\frac{\mathbf{h}_1^T\mathbf{R}_{11}\mathbf{h}_1 + \mathbf{h}_2^T\mathbf{R}_{22}\mathbf{h}_2}{2} = 1. \tag{C.4}$$

De esta forma, maximizar (C.2) equivale a minimizar la siguiente función de coste, basada en distancias,

$$J = \frac{1}{2}\|\mathbf{X}_1\mathbf{h}_1 - \mathbf{X}_2\mathbf{h}_2\|^2 = \frac{\|\mathbf{z}_1\|^2 + \|\mathbf{z}_2\|^2}{2} - \rho,$$

la cual, sujeta a (C.3) o (C.4), implica $J = 1 - \rho$, permitiendo obtener una generalización de manera sencilla a múltiples conjuntos de datos complejos. La solución a este problema viene

dada por el autovector principal del siguiente problema de autovalores generalizado (GEV) [Borga, 1998]

$$\begin{bmatrix} \mathbf{0} & \mathbf{R}_{12} \\ \mathbf{R}_{21} & \mathbf{0} \end{bmatrix} \mathbf{h} = \rho \begin{bmatrix} \mathbf{R}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_{22} \end{bmatrix} \mathbf{h},$$

donde ρ es la correlación canónica y $\mathbf{h} = [\mathbf{h}_1^T, \mathbf{h}_2^T]^T$ es el autovector. Finalmente, resulta sencillo comprobar que, en este caso, un cambio en el signo de $\mathbf{h}_1$ o $\mathbf{h}_2$ implica un cambio en el signo de ρ. Por lo tanto, si $\mathbf{h} = [\mathbf{h}_1^T, \mathbf{h}_2^T]^T$ es el autovector correspondiente al máximo ρ, $\mathbf{h} = [\mathbf{h}_1^T, -\mathbf{h}_2^T]^T$ es el autovector asociado al menor autovalor.

Por otro lado, las restricciones (C.3) se pueden reescribir como

$$\mathbf{f}_1^T \mathbf{f}_1 = \mathbf{f}_2^T \mathbf{f}_2 = 1,$$

donde $\mathbf{f}_i = \mathbf{R}_{ii}^{-1/2} \mathbf{h}_i,\ i = 1, 2$. Por lo tanto, en términos de $\mathbf{f}_i$, la correlación canónica viene dada por

$$\rho = \mathbf{f}_1^T \mathbf{R}_{11}^{-1/2} \mathbf{R}_{12} \mathbf{R}_{22}^{-1/2} \mathbf{f}_2 = \mathbf{f}_1^T \mathbf{C}_{12} \mathbf{f}_2,$$

y la solución que maximiza ρ son los vectores singulares asociados al mayor valor singular de la matriz de coherencia $\mathbf{C}_{12}$.

C.1.1. Soluciones CCA Sucesivas

Con objeto de encontrar soluciones CCA adicionales, se resuelven diferentes problemas de optimización de manera sucesiva. Denotando los i-ésimos vectores, variables, y correlaciones canónicas como $\mathbf{h}_1^{(i)}$, $\mathbf{h}_2^{(i)}$, $\mathbf{z}_1^{(i)} = \mathbf{X}_1 \mathbf{h}_1^{(i)}$, $\mathbf{z}_2^{(i)} = \mathbf{X}_2 \mathbf{h}_2^{(i)}$ y $\rho^{(i)}$, respectivamente; se define $\mathbf{z}^{(i)} = \frac{1}{2}\left(\mathbf{z}_1^{(i)} + \mathbf{z}_2^{(i)}\right)$ e impone la siguiente condición de ortogonalidad para $i \neq j$

$$\mathbf{z}^{(i)T} \mathbf{z}^{(j)} = 0, \tag{C.5}$$

la cual, en el caso de dos conjuntos de datos también implica

$$\mathbf{z}_k^{(i)T} \mathbf{z}_l^{(j)} = 0, \qquad k, l = 1, 2. \tag{C.6}$$

La equivalencia entre (C.5) y (C.6) se puede demostrar fácilmente resolviendo el problema de optimización con restricciones por el método de los multiplicadores de Lagrange. Finalmente, se obtiene nueva solución CCA resolviendo el siguiente problema GEV

$$\begin{bmatrix} \mathbf{0} & \mathbf{R}_{12} \\ \mathbf{R}_{21} & \mathbf{0} \end{bmatrix} \mathbf{h}^{(i)} = \rho^{(i)} \begin{bmatrix} \mathbf{R}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_{22} \end{bmatrix} \mathbf{h}^{(i)},$$

donde $\rho^{(i)}$ es la i-ésima correlación canónica y $\mathbf{h}^{(i)} = [\mathbf{h}_1^{(i)T}, \mathbf{h}_2^{(i)T}]^T$ es el autovector asociado.

C.2. Generalizaciones a Varios Conjuntos de Datos

En este apartado se resumen las distintas generalizaciones de CCA a varios conjuntos de datos propuestas en [Kettenring, 1971].

C.2.1. Máxima Varianza (MAXVAR)

La generalización CCA de máxima varianza (MAXVAR) fue propuesta por Horst en 1961 [Horst, 1961a]. Dados L conjuntos de datos $\mathbf{X}_k \in \mathbb{C}^{N\times m_k}$ $(k = 1, \ldots, L)$, en [Kettenring, 1971] se define la generalización MAXVAR como el problema de encontrar, de manera sucesiva, un conjunto de vectores $\mathbf{f}_k^{(i)}$ y las correspondientes proyecciones $\mathbf{y}_k^{(i)} = \mathbf{X}_k\mathbf{f}_k^{(i)}$, que admitan la mejor representación PCA uni-dimensional posible $\mathbf{z}^{(i)}$, todo ello sujeto a restricciones de energía unidad $\|\mathbf{y}_k^{(i)}\| = 1$, y a restricciones de ortogonalidad tales como[1] $\mathbf{z}^{(i)H}\mathbf{z}^{(j)} = 0$ para $j = 1, \ldots, i-1$. De esta forma, la función de coste a minimizar, sujeta a $\|\mathbf{a}^{(i)}\|^2 = L$, es

$$J_{\mathrm{PCA}}(\mathbf{f}^{(i)}) = \min_{\mathbf{z}^{(i)},\mathbf{a}^{(i)}} \frac{1}{L}\sum_{k=1}^{L}\left\|\mathbf{z}^{(i)} - a_k^{(i)}\mathbf{y}_k^{(i)}\right\|^2, \tag{C.7}$$

donde $\mathbf{a}^{(i)} = [a_1^{(i)}, \ldots, a_L^{(i)}]^T$ es el vector que proporciona los pesos para la mejor combinación lineal de las salidas y $\mathbf{f}^{(i)} = [\mathbf{f}_1^{(i)T}, \ldots, \mathbf{f}_L^{(i)T}]^T$. Tomando la derivada de (C.7) respecto a $\mathbf{z}^{(i)}$ e igualando a cero se obtiene

$$\mathbf{z}^{(i)} = \frac{1}{L}\mathbf{Y}^{(i)}\mathbf{a}^{(i)}, \tag{C.8}$$

donde $\mathbf{Y}^{(i)} = [\mathbf{y}_1^{(i)} \cdots \mathbf{y}_L^{(i)}]$. Finalmente, sustituyendo (C.8) en (C.7), la función de coste resultante es

$$J_{\mathrm{PCA}}(\mathbf{f}^{(i)}) = 1 - \frac{\mathbf{a}^{(i)H}\mathbf{Y}^{(i)H}\mathbf{Y}^{(i)}\mathbf{a}^{(i)}}{L^2} = 1 - \beta^{(i)},$$

cuya solución es el autovector principal del siguiente problema GEV [Vía et al., 2007]

$$\frac{1}{L}\begin{bmatrix}\mathbf{R}_{11} & \mathbf{R}_{12} & \ldots & \mathbf{R}_{1L}\\ \mathbf{R}_{21} & \mathbf{R}_{22} & \ldots & \mathbf{R}_{2L}\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{R}_{L1} & \mathbf{R}_{L2} & \ldots & \mathbf{R}_{LL}\end{bmatrix}\mathbf{h}^{(i)} = \beta^{(i)}\begin{bmatrix}\mathbf{R}_{11} & \mathbf{0} & \ldots & \mathbf{0}\\ \mathbf{0} & \mathbf{R}_{22} & \ldots & \mathbf{0}\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{0} & \mathbf{0} & \ldots & \mathbf{R}_{LL}\end{bmatrix}\mathbf{h}^{(i)}, \tag{C.9}$$

donde

$$\beta^{(i)} = \frac{1 + (L-1)\rho^{(i)}}{L},$$

siendo $\rho^{(i)}$ la correlación canónica.

C.2.2. Mínima Varianza (MINVAR)

De forma análoga a MAXVAR, en [Kettenring, 1971] se define la generalización CCA de mínima varianza (MINVAR) como el problema de encontrar sucesivamente un conjunto de vectores $\mathbf{f}_k^{(i)}$, y las correspondientes proyecciones $\mathbf{y}_k^{(i)} = \mathbf{X}_k\mathbf{f}_k^{(i)}$, que admitan la mejor representación PCA $(L-1)$-dimensional posible, todo ello sujeto a las restricciones $\|\mathbf{y}_k^{(i)}\| = 1$ y $\mathbf{z}^{(i)H}\mathbf{z}^{(j)} = 0$ para $j = 1, \ldots, i-1$, donde $\mathbf{Y}^{(i)} = [\mathbf{y}_1^{(i)} \cdots \mathbf{y}_L^{(i)}]$ y

$$\mathbf{z}^{(i)} = \frac{1}{L}\mathbf{Y}^{(i)}\mathbf{a}^{(i)},$$

[1] La obtención de soluciones CCA sucesivas puede basarse también en otras restricciones de ortogonalidad, por ejemplo $\mathbf{y}_k^{(i)H}\mathbf{y}_k^{(j)} = 0$ para $i \neq j$.

es la proyección de $\mathbf{Y}^{(i)}$ en el subespacio uni-dimensional complementario, es decir, $\mathbf{z}^{(i)}$ es el error residual de la representación PCA $(L-1)$-dimensional de $\mathbf{Y}^{(i)}$. De esta forma, la función de coste es

$$J_{\mathrm{PCA}}(\mathbf{f}^{(i)}) = \left\|\mathbf{z}^{(i)}\right\|^2 = \frac{\mathbf{a}^{(i)H}\mathbf{Y}^{(i)H}\mathbf{Y}^{(i)}\mathbf{a}^{(i)}}{L^2} = \beta^{(i)},$$

cuya solución viene dada por el autovector asociado al menor autovalor de (C.9).

C.2.3. Otras Generalizaciones

A continuación se enumeran otras generalizaciones de CCA al caso de más de dos conjuntos de datos.

- El Sistema de Gifi: En [Gifi, 1990], el autor propone una generalización de los problemas MAXVAR y MINVAR (un buen resumen se puede consultar en [Michailidis y de Leeuw, 1998]) consistente en encontrar, sucesivamente, el conjunto de proyecciones que admita la mejor aproximación P-dimensional posible. Obviamente, en los casos particulares $P = 1$ y $P = L - 1$, el sistema de Gifi se reduce a las generalizaciones MAXVAR y MINVAR, respectivamente. En un caso general, la solución del sistema de Gifi se obtiene mediante un algoritmo de mínimos cuadrados alternados (ALS) [Michailidis y de Leeuw, 1998].

- Suma de Correlaciones (SUMCOR): El método de suma de correlaciones fue introducido por Horst en 1961 [Horst, 1961b]. Esta generalización obtiene los vectores $\mathbf{f}_k$ y las proyecciones $\mathbf{y}_k = \mathbf{X}_k\mathbf{f}_k$, con $\|\mathbf{y}_k\| = 1$, que minimizan la siguiente función de coste

$$J_{SC}(\mathbf{f}) = \min_{\mathbf{z}} \frac{1}{L}\sum_{k=1}^{L} \|\mathbf{z} - a_k\mathbf{y}_k\|^2$$

  donde, a diferencia de (C.7), a_k son parámetros del modelo fijados *a priori*. De esta manera, SUMCOR se puede interpretar como un caso particular de MAXVAR cuando los pesos a_k son fijos, o en otras palabras, MAXVAR es una generalización de SUMCOR donde también se optimizan los pesos a_k. Finalmente, en ambos casos la función objetivo es $\mathbf{a}^T\mathbf{Y}^T\mathbf{Y}\mathbf{a}$, es decir, una suma ponderada de las correlaciones entre las proyecciones $\mathbf{y}_k$. En particular, para $\mathbf{a} = [1, \ldots, 1]^T$, el problema se reduce a encontrar las proyecciones $\mathbf{y}_k$ que proporcionan una mayor suma de las correlaciones.

- Suma de Correlaciones al Cuadrado (SSQCOR): Esta generalización fue propuesta por Kettenring [Kettenring, 1971] para combinar las ventajas de los métodos de máxima y mínima varianza. El método maximiza los mayores autovalores de la matriz $\mathbf{Y}^{(i)T}\mathbf{Y}^{(i)}/L$ (análogamente a MAXVAR) a la vez que minimiza los autovalores más pequeños (de forma análoga a MINVAR). El criterio empleado para conseguir este objetivo es maximizar la suma de los cuadrados de los autovalores de $\mathbf{Y}^{(i)T}\mathbf{Y}^{(i)}/L$.

- Varianza Generalizada (GENVAR): Esta técnica fue propuesta en [Steel, 1951], y se puede interpretar como una alternativa al método SSQCOR. En esta ocasión, el criterio consiste en minimizar el producto de los autovalores de $\mathbf{Y}^{(i)T}\mathbf{Y}^{(i)}/L$.

Apéndice D

Publicaciones

D.1. Publicaciones derivadas de la Tesis

El desarrollo de esta Tesis ha dado lugar a las siguientes publicaciones:

D.1.1. Revistas Internacionales

1. V. Pichler, M. Homolák, W. Skierucha, M. Pichlerová, D. Ramírez, J. Gregor y P. Jaloviar. "Variability of moisture in coarse woody debris from several ecologically important tree species of the temperate zone of Europe". *Ecohydrology (Aceptado)*, 2011.

2. D. Ramírez, G. Vazquez-Vilar, R. López-Valcarce, J. Vía y I. Santamaría. "Detection of rank-P signals in cognitive radio networks with uncalibrated multiple antennas". *IEEE Transactions on Signal Processing (Aceptado)*, 2011.

3. D. Ramírez, J. Vía, I. Santamaría and L. L. Scharf. "Detection of spatially correlated Gaussian time series". *IEEE Transactions on Signal Processing*, vol.58, no.10, pp.5006-5015, Octubre 2010.

D.1.2. Congresos Internacionales

4. D. Ramírez, J. Vía, I. Santamaría y L. L. Scharf. "Multi-sensor beamsteering based on the asymptotic likelihood for colored signals". *IEEE Workshop on Statistical Signal Processing (SSP 2011)*, Niza, Francia, Junio 2011.

5. D. Ramírez, G. Vazquez-Vilar, R. López-Valcarce, J. Vía y I. Santamaría. "Multiantenna detection under noise uncertainty and primary users spatial structure". *IEEE Int. Conf. on Acoust., Speech, and Signal Processing (ICASSP 2011)*, República Checa, Mayo 2011.

6. D. Ramírez, J. Vía, I. Santamaría y L. L. Scharf. "Multiple-channel detection of a Gaussian time series over frequency-flat channels". *IEEE Int. Conf. on Acoust., Speech, and Signal Processing (ICASSP 2011)*, República Checa, Mayo 2011.

7. D. Ramírez, J. Vía y I. Santamaría. "Multiantenna spectrum sensing: The case of wideband rank-one primary signals". *6th IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM 2010)*, Israel, Octubre 2010.

8. D. Ramírez, J. Vía, I. Santamaría, R. López-Valcarce y L. L. Scharf. "Multiantenna spectrum sensing: detection of spatial correlation among time-series with unknown spectra". *IEEE Int. Conf. on Acoust., Speech, and Signal Processing (ICASSP 2010)*, Dallas, EE.UU., Marzo 2010.

9. D. Ramírez, J. Vía y I. Santamaría. "Multiple-channel signal detection using the generalized coherence spectrum". *1st IAPR Workshop on Cognitive Information Processing (CIP 2008)*, Santorini, Grecia, Junio 2008.

10. D. Ramírez, J. Vía y I. Santamaría. "A generalization of the magnitude squared coherence spectrum for more than two signals: definition, properties and estimation". *IEEE Int. Conf. on Acoust., Speech, and Signal Processing (ICASSP 2008)*, Las Vegas, EE.UU., Abril 2008.

D.2. Otras Publicaciones

Durante el desarrollo de esta Tesis, se ha trabajado adicionalmente en otros temas de investigación no relacionados con la temática de la Tesis:

D.2.1. Revistas Internacionales

11. J. Gutiérrez, Ó. González, J. Pérez, D. Ramírez, L. Vielva, J. Ibáñez y I. Santamaría. "Frequency-domain methodology for measuring MIMO channels using a generic test bed". *IEEE Transactions on Instrumentation and Measurement*, vol.60, no.3, pp.827-838, Marzo 2011.

12. J. Vía, D. Ramírez y I. Santamaría. "Properness and widely linear processing of quaternion random vectors". *IEEE Transactions on Information Theory*, vol.56, no.7, pp.3502-3515, Julio 2010.

13. D. Ramírez, I. Santamaría, J. Pérez, J. Vía, J. A. García-Naya, T. M. Fernández-Caramés, H. Pérez-Iglesias, M. González López, L. Castedo y J. M. Torres-Royo. "A comparative study of STBC transmissions at 2.4 GHz over indoor channels using a 2×2 MIMO testbed". *Wireless Communications and Mobile Computing*, vol.8, no.9, pp.1149-1164, Noviembre 2008.

14. D. Ramírez y I. Santamaría. "Regularised approach to detection of constant modulus signals in MIMO channels". *Electronic Letters*, vol.42, no.3, pp.184-186, Febrero 2006.

D.2.2. Congresos Internacionales

15. Ó. González, D. Ramírez, I. Santamaría, J. A. García-Naya y L. Castedo. "Experimental validation of interference alignment techniques using a multiuser MIMO testbed". *International ITG Workshop on Smart Antennas (WSA 2011)*, Aachen, Alemania, Febrero 2011.

16. J. Vía, D. Ramírez, I. Santamaría y L. Vielva. "Improperness measures for quaternion random vectors". *IEEE International Workshop on Machine Learning for Signal Processing (MLSP 2010)*, Kittilä, Finlandia, Agosto 2010.

17. J. Vía, D. Ramírez, I. Santamaría y L. Vielva. "Widely and semi-widely linear processing of quaternion vectors". *IEEE Int. Conf. on Acoust., Speech, and Signal Processing (ICASSP 2010)*, Dallas, EE.UU., Marzo 2010.

18. D. Ramírez, J. Vía y I. Santamaría. "Coherent fusion of information for optimal detection in sensor networks". *IEEE Workshop on Statistical Signal Processing (SSP 09)*, Cardiff, Reino Unido, Septiembre 2009.

19. D. Ramírez, J. Vía, I. Santamaría y P. Crespo. "Entropy and Kullback-Leibler divergence estimation based on Szegö's theorem". *17th European Signal Processing Conference (EUSIPCO 2009)*, Glasgow, Escocia, Agosto 2009.

20. I. Santamaría, V. Elvira, J. Vía, D. Ramírez, J. Pérez, J. Ibáñez, R. Eickhoff y F. Ellinger. "Optimal MIMO transmission schemes with adaptive antenna combining in the RF path". *16th European Signal Processing Conference (EUSIPCO 2008)*, Lausanne, Suiza, Agosto 2008.

21. J. A. García-Naya, T. M. Fernández-Caramés, H. J. Pérez-Iglesias, M. González-López, L. Castedo, D. Ramírez, I. Santamaría, J. Pérez, J. Vía y J. M. Torres-Royo. "Performance of STBC transmissions with real data". *16th IST Mobile & Wireless Communications Summit*, Budapest, Hungría, Julio 2007.

22. D. Ramírez, I. Santamaría, J. Pérez, J. Vía, A. Tazón, J. A. Garcia-Naya, T. M. Fernández-Caramés, M. González-López, H. J. Pérez Iglesias y L. Castedo. "A flexible testbed for the rapid prototyping of MIMO baseband modules", *3rd International Symposium on Wireless Communication Systems (ISWCS 2006)*, Valencia, España, Septiembre 2006.

23. J. Vía, I. Santamaría, J. Pérez y D. Ramírez. "Blind decoding of MISO-OSTBC systems based on principal component analysis". *IEEE Int. Conf. on Acoust., Speech, and Signal Processing (ICASSP 2006)*, Toulouse, Francia, Mayo 2006.

Índice de Figuras

2.1. Las observaciones están compuestas por ventanas de longitud N de cada una de las L series temporales univariadas. 8

3.1. Ilustración gráfica de las condiciones para la convergencia de la técnica basada en las aproximaciones convexas sucesivas. 39
3.2. Probabilidad de pérdida frente a P para ruidos IID. 47
3.3. PDF estimada y aproximación χ^2. 48
3.4. Curva ROC de los diferentes detectores para señales de rango P y ruidos IID. 49
3.5. Curva ROC de los diferentes detectores para señales sin estructura espacial y ruidos IID. 49
3.6. PDF estimada y aproximación Gaussiana de los GLRT asintóticos para ruidos IID. 50
3.7. Curva ROC de los diferentes detectores para señales de rango $P = 1$, ruidos IID y canales frecuencialmente planos. 51
3.8. Curvas de convergencia del GLRT iterativo basado en las aproximaciones convexas sucesivas. 52
3.9. Probabilidad de pérdida de los GLRT asintóticos considerando dos resoluciones diferentes. 53
3.10. MSE de la estima del ángulo de llegada de las diferentes técnicas. 55
3.11. Curva ROC de los diferentes detectores para señales no estacionarias. 56
3.12. Curva ROC para los diferentes detectores en un ejemplo de radio cognitiva con ruidos IID. 56

4.1. Probabilidad de pérdida frente a P con ruidos no-IID. 73
4.2. PDF estimada y aproximación χ^2. 73
4.3. Probabilidad de pérdida frente a la SNR de los diferentes detectores. 74
4.4. Curva ROC de los GLRT propuestos considerando ruidos IID y ruidos no-IID. 75
4.5. Curva ROC de los diferentes detectores para señales con PSD arbitrarias contaminadas con ruidos no-IID. 76
4.6. PDF estimada y aproximación Gaussiana de los GLRT asintóticos para ruidos no-IID. 76
4.7. Probabilidad de pérdida del detector basado en el cociente de Hadamard y su aproximación para señales blancas. 77
4.8. Probabilidad de pérdida del detector basado en el cociente de Hadamard y su aproximación para señales coloreadas. 78
4.9. Curva ROC de los diferentes detectores para señales no estacionarias. 78
4.10. Curva ROC para los diferentes detectores para señales en un ejemplo de radio cognitiva con ruidos no-IID. 79

5.1. Comparativa del GCS-MAX, GCS-MIN y GCS-DET. 97
5.2. Valor de las componentes del autovector principal. 98
5.3. Estima del GCS-MAX para señales armónicas. 99
5.4. Estima del GCS-MAX para señales banda estrecha. 100
5.5. Principales autovalores de la matriz de coherencia. 101
5.6. Influencia del orden en la estima del GCS-MAX. 101
5.7. Curva ROC de los diferentes detectores para señales banda estrecha y banda ancha. 102
5.8. GCS para las diferentes especies y clases de descomposición. 104

Índice de Tablas

2.1. Modelos considerados en la Tesis para ruidos IID. 12
2.2. Modelos considerados en la Tesis para ruidos no-IID. 13

5.1. GCS-MAX, GCS-MIN y GCS-DET en un ejemplo ilustrativo. 89
5.2. Frecuencias de las exponenciales complejas. 99

Índice de Algoritmos

3.1. Estima del canal SIMO basada en aproximaciones convexas sucesivas. 41

4.1. Estima iterativa de $\mathbf{H}_{\boldsymbol{\Sigma}}$ y $\boldsymbol{\Sigma}$ vía optimización alternada. 64

Bibliografía

[3GPP, 2009] 3GPP. "LTE Release 9". *3GPP*, 2009.

[Akyildiz et al., 2008] I. Akyildiz, W.-Y. Lee, M. Vuran, y S. Mohanty. "A survey on spectrum management in cognitive radio networks". *IEEE Comm. Magazine*, volumen 46, nº 4, páginas 40–48, 2008.

[Alamgir et al., 2008] M. Alamgir, M. Faulkner, J. Gao, y P. Conder. "Signal detection for cognitive radio using multiple antennas". En *IEEE Int. Symp. Wireless Comm. Systems.* 2008.

[Alamouti, 1998] S. Alamouti. "A simple transmit diversity technique for wireless communications". *IEEE J. Sel. Areas Comm.*, volumen 45, nº 9, páginas 1451–1458, 1998.

[Anderson, 1958] T. W. Anderson. *An Introduction to Multivariate Statistical Analysis.* Wiley, New York, 1958.

[Anderson, 1963] T. W. Anderson. "Asymptotic theory for principal component analysis". *The Annals of Mathematical Statistics*, volumen 34, nº 1, páginas 122–148, 1963.

[Azimi-Sadjadi et al., 2008] M. R. Azimi-Sadjadi, N. Roseveare, y A. Pezeshki. "Wideband DOA estimation algorithms for multiple moving sensors using unattended acoustic sensors". *IEEE Trans. Aero. Elect. Systems*, volumen 44, nº 4, páginas 1585–1599, 2008.

[Azimi-Sadjadi et al., 2000] M. R. Azimi-Sadjadi, D. Yao, Q. Huang, y G. J. Dobeck. "Underwater target classification using wavelet packets and neural networks". *IEEE Trans. Neural Networks*, volumen 11, páginas 784–794, 2000.

[Baggenstoss y Kay, 1992] P. M. Baggenstoss y S. M. Kay. "An adaptive detector for deterministic signals in noise of unknown spectra using the Rao test". *IEEE Trans. Signal Process.*, volumen 40, nº 6, páginas 1460–1468, 1992.

[Benesty et al., 2006] J. Benesty, J. Chen, y Y. Huang. "Estimation of the coherence function with the MVDR approach". En *Proc. IEEE Int. Conf. Acoust., Speech and Signal Process. (ICASSP).* 2006.

[Besson et al., 2006] O. Besson, S. Kraut, y L. L. Scharf. "Detection of an unknown rank-one component in white noise". *IEEE Trans. Signal Process.*, volumen 54, nº 7, páginas 2835–2839, 2006.

[Böhme, 1986] J. F. Böhme. "Estimation of spectral parameters of correlated signals in wavefields". *Signal Process.*, volumen 11, páginas 329–337, 1986.

[Boonstra y Van der Veen, 2003] A.-J. Boonstra y A.-J. Van der Veen. "Gain calibration methods for radio telescope arrays". *IEEE Trans. Signal Process.*, volumen 51, nº 1, páginas 25–38, 2003.

[Borga, 1998] M. Borga. *Learning Multidimensional Signal Processing.* Tesis Doctoral, Linköping University, Linköping, Sweden, 1998. Dissertation No 531.

[Bose y Steinhardt, 1995] S. Bose y A. O. Steinhardt. "A maximal invariant framework for adaptive detection with structured and unstructured covariance matrices". *IEEE Trans. Signal Process.*, volumen 43, nº 9, páginas 2164–2175, 1995.

[Bose y Steinhardt, 1996] S. Bose y A. O. Steinhardt. "Adaptive array detection of uncertain rank one waveforms". *IEEE Trans. Signal Process.*, volumen 44, nº 11, páginas 2801–2809, 1996.

[Boutros et al., 2006] J. J. Boutros, F. Kharrat, y H. Randriambololona. "A classification of multiple antenna channels". En *Proc. Int. Zurich Seminar Communications (IZS).* 2006.

[Boyd y Vandenberghe, 2004] S. Boyd y L. Vandenberghe. *Convex Optimization.* Cambridge University Press, 2004.

[Bresler, 1988] Y. Bresler. "Maximum likelihood estimation of a linearly structured covariance with application to antenna array processing". En *Proc. 4th Spectrum Estimation Workshop.* 1988.

[Brillinger, 1981] D. R. Brillinger. *Time Series: Data Analysis and Theory.* McGraw-Hill, 1981.

[Burg et al., 1982] J. P. Burg, D. G. Luenberger, y D. L. Wenger. "Estimation of structured covariance matrices". *Proc. IEEE*, volumen 70, nº 9, páginas 963–974, 1982.

[Cabric, 2008] D. Cabric. "Addressing the feasibility of cognitive radios". *IEEE Signal Process. Magazine*, volumen 25, nº 6, páginas 85–93, 2008.

[Capon, 1969] J. Capon. "High resolution frequency-wavenumber spectrum analysis". *Proc. IEEE*, volumen 57, páginas 1408–1418, 1969.

[Carter et al., 1973] G. Carter, C. Knapp, y A. Nuttall. "Estimation of the magnitude-squared coherence function via overlapped Fast Fourier Transform processing". *IEEE Trans. Audio Electroacoust.*, volumen 21, nº 4, páginas 331–344, 1973.

[Chen y Prasad, 2009] K.-C. Chen y R. Prasad. *Cognitive Radio Networks.* Wiley, 2009.

[Chiang et al., 2007] M. Chiang, C. W. Tan, D. P. Palomar, D. O'Neill, y D. Julian. "Power control by geometric programming". *IEEE Trans. Wireless Comm.*, volumen 6, nº 7, páginas 2640–2651, 2007.

[Cover y Thomas, 2006] T. M. Cover y J. A. Thomas. *Elements of Information Theory.* Wiley-Interscience, 2006.

[Cox, 1973] H. Cox. "Resolving power and sensitivity to mismatch of optimum array processors". *Journal of the Acoustic Society of America*, volumen 54, nº 3, páginas 771–785, 1973.

[Cui et al., 2007] S. Cui, J.-J. Xiao, A. J. Goldsmith, Z.-Q. Luo, y H. V. Poor. "Estimation diversity and energy efficiency in distributed sensing". *IEEE Trans. Signal Process.*, volumen 55, nº 9, páginas 4683–4695, 2007.

[Dandawaté y Giannakis, 1994] A. V. Dandawaté y G. B. Giannakis. "Statistical tests for presence of cyclostationarity". *IEEE Trans. Signal Process.*, volumen 42, nº 9, páginas 2355–2369, 1994.

[de Maio, 2004] A. de Maio. "A new derivation of the adaptive matched filter". *IEEE Signal Process. Lett.*, volumen 11, nº 10, páginas 792–793, 2004.

[de Maio, 2007] A. de Maio. "Rao test for adaptive detection in Gaussian interference with unknown covariance matrix". *IEEE Trans. Signal Process.*, volumen 55, nº 7, páginas 3577–3584, 2007.

[de Maio et al., 2010] A. de Maio, S. M. Kay, y A. Farina. "On the invariance, coincidence, and statistical equivalence of the GLRT, Rao test, and Wald test". *IEEE Trans. Signal Process.*, volumen 58, nº 4, páginas 1967–1979, 2010.

[Denoeux, 1995] T. Denoeux. "A k-nearest neighbor classification rule based on Dempster-Shafer theory". *IEEE Trans. Syst., Man, Cybern.*, volumen 25, páginas 804–813, 1995.

[Dudewicz, 1976] E. J. Dudewicz. *Introduction to probability and Statistics.* Holt, Rinehart, and Winston, New York, 1976.

[Dzhaparidze, 1986] K. Dzhaparidze. *Parameter Estimation and Hypothesis Testing in Spectral Analysis of Stationary Time Series.* Springer-Verlag, 1986.

[ETSI, 2009] ETSI. "Digital Video Broadcasting (DVB); Frame structure, channel coding and modulation for a second generation digital terrestrial television broadcasting system (DVB-T2)". *ETSI EN 302 755*, 2009.

[FCC, 2002] FCC. "Spectrum policy task force report". En *Proc. of the Federal Communications Commission.* Washington DC, USA, 2002.

[Ferguson, 1967] T. S. Ferguson. *Mathematical statistics.* New York: Academic Press, 1967.

[Fishler et al., 2004a] E. Fishler, A. Haimovich, R. Blum, D. Chizhik, L. Cimini, y R. Valenzuela. "MIMO radar: An idea whose time has come". En *Proc. IEEE Radar Conf.*, páginas 71–78. 2004.

[Fishler et al., 2004b] E. Fishler, A. Haimovich, R. Blum, L. Cimini, D. Chizhik, y R. Valenzuela. "Performance of MIMO radar systems: Advantages of angular diversity". En *Proc. 38th Asilomar Conf. Signals, Systems and Computers*, páginas 305–309. Pacific Grove, CA, 2004.

[Fishler et al., 2006] E. Fishler, A. Haimovich, R. Blum, L. Cimini, D. Chizhik, y R. Valenzuela. "Spatial diversity in radars – models and detection performance". *IEEE Trans. Signal Process.*, volumen 54, nº 3, páginas 823–838, 2006.

[Font-Segura et al., 2011a] J. Font-Segura, G. Vázquez, y J. Riba. "Compressed correlation-matching for spectrum sensing in sparse wideband regimes". En *Proc. IEEE Int. Conf. Comm.* Kyoto, Japan, 2011.

[Font-Segura et al., 2011b] J. Font-Segura, G. Vázquez, y J. Riba. "Multi-frequency GLRT spectrum sensing for wideband cognitive radio". En *Proc. IEEE Int. Conf. Comm.* Kyoto, Japan, 2011.

[Friston et al., 1994] K. Friston, P. Jezzard, y R. Turner. "Analysis of functional MRI time-series". *Hum. Brain Mapp.*, volumen 1, nº 1, páginas 153–171, 1994.

[Gabriel y Kay, 2002] J. Gabriel y S. Kay. "Use of Wijsman's theorem for the ratio of maximal invariant densities in signal detection applications". En *Asilomar Conf. Signals, Systems and Computers*, volumen 1, páginas 756 – 762. 2002.

[García-Naya et al., 2007] J. A. García-Naya, T. M. Fernández-Caramés, H. Pérez-Iglesias, M. González-López, L. Castedo, D. Ramírez, I. Santamaría, J. Pérez, J. Vía, y A. Tazón. "Performance of STBC transmissions with real data". En *16th Mobile & Wireless Comm. Summit.* 2007.

[Gardner et al., 2006] W. A. Gardner, A. Napolitano, y L. Paura. "Cyclostationarity: Half a century of research". *Signal Process.*, volumen 86, nº 4, páginas 639–697, 2006.

[Gifi, 1990] A. Gifi. *Nonlinear Multivariate Analysis.* Chichester: Wiley, 1990.

[Golub y van Loan, 1996] G. H. Golub y C. F. van Loan. *Matrix Computations.* Johns Hopkins University Press, 1996.

[González et al., 2011] Ó. González, D. Ramírez, I. Santamaría, J. A. García-Naya, y L. Castedo. "Experimental validation of interference alignment techniques using a multiuser MIMO testbed". En *International ITG Workshop on Smart Antennas (WSA 2011).* Aachen, Germany, 2011.

[Gray, 2006] R. M. Gray. *Toeplitz and Circulant Matrices: A Review.* Foundations and Trends in Communications and Information Theory, 2006.

[Grenander y Szegö, 1958] U. Grenander y G. Szegö. *Toeplitz Forms and Their Applications.* Univ. Calif. Press, Berkeley, 1958.

[Grimmett y Stirzaker, 1992] G. Grimmett y D. Stirzaker. *Probability and Random Processes.* Oxford Science Publications, 1992.

[Gutiérrez et al., 2011] J. Gutiérrez, Ó. González, J. Pérez, D. Ramírez, L. Vielva, J. Ibáñez, y I. Santamaría. "Frequency-domain methodology for measuring MIMO channels using a generic test bed". *IEEE Trans. on Instrumentation and Measurement*, volumen 60, nº 3, páginas 827–838, 2011.

[Gutiérrez-Gutiérrez y Crespo, 2008] J. Gutiérrez-Gutiérrez y P. M. Crespo. "Asymptotically equivalent sequences of matrices and Hermitian block Toeplitz matrices with continuous symbols: Applications to MIMO systems". *IEEE Trans. Inf. Theory*, volumen 54, nº 12, páginas 5671–5680, 2008.

[Horst, 1961a] P. Horst. "Generalized canonical correlations and their applications to experimental data". *J. Clin. Psychol. (Monograph supplement.)*, volumen 14, páginas 331–347, 1961.

[Horst, 1961b] P. Horst. "Relations among m sets of measures". *Psychometrica*, volumen 26, nº 2, páginas 129–149, 1961.

[Hossain y Bhargava, 2007] E. Hossain y V. K. Bhargava, editores. *Cognitive Wireless Communication Networks*. Springer, 2007.

[Hotelling, 1936] H. Hotelling. "Relations between two sets of variates". *Biometrika*, volumen 28, páginas 321–377, 1936.

[Hua et al., 2004] Y. Hua, A. B. Gershman, y Q. Cheng, editores. *High-Resolution and Robust Signal Processing*. Marcel Dekker, 2004.

[Hua et al., 2001] Y. Hua, M. Nikpour, y P. Stoica. "Optimal reduced-rank estimation and filtering". *IEEE Trans. Signal Process.*, volumen 49, nº 3, páginas 457–469, 2001.

[IEEE Computer Society, 2009] IEEE Computer Society. "IEEE Standard 802.11n-2009. Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. Amendment 5: Enhancements for Higher Throughput". 2009.

[IEEE Computer Society and the IEEE Microwave Theory and Techniques Society, 2009] IEEE Computer Society and the IEEE Microwave Theory and Techniques Society. "IEEE Standard 802.16-2009. Part 16: Air Interface for Broadband Wireless Access Systems". 2009.

[Jorswieck y Boche, 2007] E. Jorswieck y H. Boche. *Majorization and Matrix-Monotone Functions in Wireless Communications*. Foundations and Trends in Communications and Information Theory, 2007.

[Kay, 1993] S. M. Kay. *Fundamentals of Statistical Signal Processing: Estimation Theory*, volumen I. Prentice Hall, 1993.

[Kay, 1998] S. M. Kay. *Fundamentals of Statistical Signal Processing: Detection Theory*, volumen II. Prentice Hall, 1998.

[Kettenring, 1971] J. R. Kettenring. "Canonical analysis of several sets of variables". *Biometrika*, volumen 58, nº 3, páginas 433–451, 1971.

[Korpel, 1989] S. Korpel. *Pralesy Slovenska (The natural forests of Slovakia)*. Veda Bratislava, 1989.

[Lehmann, 1986] E. L. Lehmann. *Testing Statistical Hypotheses*. Springer texts in Statistics, 1986.

[Lender, 1987] A. Lender. "Correlative (partial response) techniques and applications to digital radio systems". En K. Feher, editor, *Digital Communications: Microwave Application*. Prentice-Hall, Englewood Cliffs, NJ, 1987.

[Leshem y Van der Veen, 2001a] A. Leshem y A.-J. Van der Veen. "Multichannel detection and spatial signature estimation with uncalibrated receivers". En *Proc. 11th IEEE Work. Stat. Signal Process.*, páginas 190–193. 2001.

[Leshem y Van der Veen, 2001b] A. Leshem y A.-J. Van der Veen. "Multichannel detection of Gaussian signals with uncalibrated receivers". *IEEE Signal Process. Lett.*, volumen 8, nº 4, páginas 120–122, 2001.

[Leshem et al., 2000] A. Leshem, A.-J. Van der Veen, y A.-J. Boonstra. "Multichannel interference mitigation techniques in radio astronomy". *The Astrophysical Journal Supplement Series*, volumen 131, páginas 355–373, 2000.

[Li et al., 2004] D. Li, M. R. Azimi-Sadjadi, y M. Robinson. "Comparison of different classification algorithms for underwater target discrimination". *IEEE Trans. Neural Networks*, volumen 15, nº 1, páginas 189–194, 2004.

[Li y Stoica, 2005] J. Li y P. Stoica. *Robust Adaptive Beamforming.* Wiley, New York, 2005.

[Li y Stoica, 2008] J. Li y P. Stoica. *MIMO Radar Signal Processing.* Wiley-IEEE Press, 2008.

[Lim et al., 2008] T. J. Lim, R. Zhang, Y.-C. Liang, y Y. Zeng. "GLRT-based spectrum sensing for cognitive radio". En *Proc. Global Comm. Conf.* 2008.

[López-Valcarce et al., 2010] R. López-Valcarce, G. Vazquez-Vilar, y J. Sala. "Multiantenna spectrum sensing for cognitive radio: overcoming noise uncertainty". En *Proc. Int. Work. Cognitive Inf. Process.* Elba, Italy, 2010.

[Luesma, 2006] R. B. Luesma. *Análisis de la señal electrocardiográfica durante prueba de esfuerzo y su aplicación en el diagnostico de cardiopatías.* Tesis Doctoral, Universidad de Zaragoza, 2006.

[Lundén et al., 2009] J. Lundén, V. Koivunen, A. Huttunen, y H. V. Poor. "Collaborative cyclostationary spectrum sensing for cognitive radio systems". *IEEE Trans. Signal Process.*, volumen 57, nº 11, páginas 4182–4195, 2009.

[M. Falli, e.d., 1989] M. Falli, e.d. "Digital land mobile radio communications - cost 207: Final report". Luxembourg, 1989.

[Magnus y Neudecker, 1999] J. R. Magnus y H. Neudecker. *Matrix Differential Calculus with Applications in Statistics and Econometrics.* John Wiley & Sons, 1999.

[Mardia et al., 1979] K. V. Mardia, J. T. Kent, y J. M. Bibby. *Multivariate Analysis.* New York: Academic, 1979.

[Marks y Wright, 1978] B. R. Marks y G. P. Wright. "A general inner approximation algorithm for nonconvex mathematical programs". *Operations Research*, volumen 26, nº 4, páginas 681–683, 1978.

[Marshall y Olkin, 1979] A. W. Marshall y I. Olkin. *Inequalities: Theory of Majorization and Its Application.* Mathematics in Science and Engineering, Academic Press, Inc. (London) Ltd., 1979.

[Mauchly, 1940] J. Mauchly. "Significance test for sphericity of a normal n-variate distribution". *Ann. Math. Statist.*, volumen 11, páginas 204–209, 1940.

[Michailidis y de Leeuw, 1998] G. Michailidis y J. de Leeuw. "The Gifi system for descriptive multivariate analysis". *Statistical Science*, volumen 13, páginas 307–336, 1998.

[Mitola y Maguire Jr., 1999] J. Mitola y G. Q. Maguire Jr. "Cognitive radio: Making software radios more personal". *IEEE Pers. Comm.*, volumen 6, páginas 13–18, 1999.

[Mullis y Scharf, 1991] C. T. Mullis y L. L. Scharf. "Quadratic estimators of the power spectrum". En S. Haykin, editor, *Advances in Spectrum Analysis and Array Processing: Volume I*, capítulo 1, páginas 1–57. Prentice Hall, 1991.

[Neyman y Pearson, 1928] J. Neyman y E. S. Pearson. "On the use and interpretation of certain test criteria for purpose of statistical inference". *Biometrika*, volumen 20, páginas 175–240, 1928.

[Neyman y Pearson, 1933] J. Neyman y E. S. Pearson. "On the problem of the most efficient tests of statistical hypotheses". *Philosophical Transactions of the Royal Society of London*, volumen 231, páginas 289–337, 1933.

[Papoulis y Pillai, 2002] A. Papoulis y S. Pillai. *Probability, Random Variables and Stochastic Processes.* Mc Graw Hill, 2002.

[Paulraj et al., 2003] A. Paulraj, R. Nabar, y D. Gore. *Introduction to Space-Time Wireless Communications.* Cambridge University Press, 2003.

[Pérez-Neira et al., 2009] A. Pérez-Neira, M. A. Lagunas, M. A. Rojas, y P. Stoica. "Correlation matching approach for spectrum sensing in open spectrum communications". *IEEE Trans. Signal Process.*, volumen 57, nº 12, páginas 4823–4836, 2009.

[Proakis, 1988] J. Proakis. *Digital Communications.* Prentice Hall, NJ, USA, 1988.

[Ramírez et al., 2008] D. Ramírez, I. Santamaría, J. Pérez, J. Vía, J. A. García-Naya, T. M. Fernández-Caramés, H. Pérez-Iglesias, M. González-López, L. Castedo, y J. M. Torres-Royo. "A comparative study of STBC transmissions at 2.4 GHz over indoor channels using a 2×2 MIMO testbed". *Wireless Comm. and Mobile Comput.*, volumen 8, nº 9, páginas 1149–1164, 2008.

[Ramírez et al., 2006] D. Ramírez, I. Santamaría, J. Pérez, J. Vía, A. Tazón, J. A. García-Naya, T. M. Fernández-Caramés, H. Pérez-Iglesias, M. González-López, y L. Castedo. "A flexible testbed for the rapid prototyping of MIMO baseband modules". En *3rd Int. Symp. on Wireless Comm. Systems.* Valencia, Spain, 2006.

[Ramírez et al., 2008] D. Ramírez, J. Vía, y I. Santamaría. "Multiple-channel signal detection using the generalized coherence spectrum". En *Proc. Int. Work. Cognitive Inf. Process.* 2008.

[Rao, 1948] C. R. Rao. "Large sample tests of statistical hypothesis concerning several parameters with applications to problems of estimation". *Proc. Cambridge Philosophical Soc.*, volumen 44, páginas 50–57, 1948.

[Rao, 1988] C. R. Rao. *Statistical inference and its applications.* Prentice Hall, Englewood Cliffs, New York, 1988.

[Romero y López-Valcarce, 2011a] D. Romero y R. López-Valcarce. "Distributed spectrum sensing with multiantenna sensors under calibration errors". En *Proc. IEEE Int. Work. Signal Process. Advances in Wireless Comm.* San Francisco, USA, 2011.

[Romero y López-Valcarce, 2011b] D. Romero y R. López-Valcarce. "Multiantenna detection of constant-envelope signals in noise of unknown variance". En *Proc. IEEE Int. Work. Signal Process. Advances in Wireless Comm.* San Francisco, USA, 2011.

[Sameshima y Baccalá, 1999] K. Sameshima y L. A. Baccalá. "Using partial directed coherence to describe neuronal ensemble interactions". *Journal of Neoroscience Methods*, volumen 94, nº 1, páginas 93–103, 1999.

[Santamaría y Vía, 2007] I. Santamaría y J. Vía. "Estimation of the magnitude squared coherence spectrum based on reduced-rank canonical coordinates". En *Proc. IEEE Int. Conf. Acoust., Speech and Signal Process. (ICASSP)*. 2007.

[Scharf y Thomas, 1998] L. Scharf y J. Thomas. "Wiener filters in canonical coordinates for transform coding, filtering and quantizing". *IEEE Trans. Signal Process.*, volumen 46, nº 3, páginas 647–654, 1998.

[Scharf, 1991] L. L. Scharf. *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis.* Addison - Wesley, 1991.

[Scharf y Mullis, 2000] L. L. Scharf y C. T. Mullis. "Canonical coordinates and the geometry of inference, rate and capacity". *IEEE Trans. Signal Process.*, volumen 48, nº 3, páginas 824–831, 2000.

[Schizas et al., 2007] I. D. Schizas, G. B. Giannakis, y Z.-Q. Luo. "Distributed estimation using reduced-rank dimensionality sensor observations". *IEEE Trans. Signal Process.*, volumen 55, nº 8, páginas 4284–4299, 2007.

[Schölkopf y Smola, 2001] B. Schölkopf y A. J. Smola. *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond.* MIT Press, Cambridge, MA, USA, 2001.

[Schreier, 2008] P. J. Schreier. "A unifying discussion of correlation analysis for complex random vectors". *IEEE Trans. Signal Process.*, volumen 56, nº 4, páginas 1327–1336, 2008.

[Schreier y Scharf, 2006] P. J. Schreier y L. L. Scharf. "Canonical coordinates for transform coding of noisy sources". *IEEE Trans. Signal Process.*, volumen 54, nº 1, páginas 235–243, 2006.

[Specht, 1990] D. F. Specht. "Probabilistic neural networks". *Neural Networks*, volumen 3, páginas 109–118, 1990.

[Steel, 1951] R. Steel. "Minimum generalized variance for a set of linear functions". *Ann. Math. Statist.*, volumen 22, páginas 456–460, 1951.

[Stoica y Babu, 2011] P. Stoica y P. Babu. "The Gaussian data assumption leads to the largest Cramér-Rao bound". *IEEE Signal Process. Magazine*, volumen 28, nº 3, páginas 132–133, 2011.

[Stoica y Cedervall, 1997] P. Stoica y M. Cedervall. "Detection tests for array processing in unknown correlated noise fields". *IEEE Trans. Signal Process.*, volumen 45, nº 9, páginas 2351–2362, 1997.

[Stoica y Moses, 2005] P. Stoica y R. Moses. *Spectral Analysis of Signals.* Prentice Hall, 2005.

[Taherpour et al., 2010] A. Taherpour, M. Nasiri-Kenari, y S. Gazor. "Multiple antenna spectrum sensing in cognitive radios". *IEEE Trans. Wireless Comm.*, volumen 9, nº 2, páginas 814–823, 2010.

[Tandra y Sahai, 2008] R. Tandra y A. Sahai. "SNR walls for signal detection". *IEEE J. Sel. Topics Signal Process.*, volumen 2, nº 1, páginas 4–17, 2008.

[Thomson, 1982] D. J. Thomson. "Spectrum estimation and harmonic analysis". *Proc. IEEE*, volumen 70, nº 9, páginas 1055–1096, 1982.

[Tugnait y Gummadavelli, 1997] J. K. Tugnait y U. Gummadavelli. "Blind equalization and channel estimation with partial response input signals". *IEEE Trans. Comm.*, volumen 45, nº 9, páginas 1025–1031, 1997.

[Urkowitz, 1967] H. Urkowitz. "Energy detection of unknown deterministic signals". *Proc. IEEE*, volumen 55, nº 4, páginas 523–531, 1967.

[Van Trees, 2002] H. L. Van Trees. *Detection, estimation and modulation theory: Optimum array processing (part IV).* John Wiley & Sons, 2002.

[Vapnik, 1995] V. Vapnik. *The Nature of Statistical Learning Theory.* Springer-Verlag, NY, USA, 1995.

[Vía et al., 2007] J. Vía, I. Santamaría, y J. Pérez. "A learning algorithm for adaptive canonical correlation analysis of several data sets". *Neural Networks*, volumen 20, nº 1, páginas 139–152, 2007.

[Vielva et al., 2010] L. Vielva, J. Vía, J. Gutiérrez, Ó. González, J. Ibáñez, y I. Santamaría. "Building a web platform for learning advanced digital communications using a MIMO testbed". En *IEEE Int. Conf. on Acoustics, Speech and Signal Process.* Dallas, USA, 2010.

[Villares y Vázquez, 2007] J. Villares y G. Vázquez. "The Gaussian assumption in second-order estimation problems in digital communications". *IEEE Trans. Signal Process.*, volumen 55, nº 10, páginas 4994–5002, 2007.

[Vucetic y Yuan, 2003] B. Vucetic y J. Yuan. *Space-Time Coding.* Wiley, 2003.

[Wald, 1943] A. Wald. "Tests of statistical hypothesis concerning several parameters when the number of observations is large". *Trans. American Math. Soc.*, volumen 54, páginas 426–482, 1943.

[Wang et al., 2010] P. Wang, J. Fang, N. Han, y H. Li. "Multiantenna-assisted spectrum sensing for cognitive radio". *IEEE Trans. Vehicular Tech.*, volumen 59, nº 4, páginas 1791–1800, 2010.

[Washizawa et al., 2010] Y. Washizawa, Y. Yamashita, T. Tanaka, y A. Cichocki. "Blind extraction of global signal from multi-channel noisy observations". *IEEE Trans. Neural*, volumen 21, nº 9, páginas 1742–1481, 2010.

[Whittle, 1952] P. Whittle. "Estimation and information in time series analysis". *Skand. Aktuar.*, volumen 35, páginas 48–60, 1952.

[Whittle, 1962] P. Whittle. "Gaussian estimation in stationary time series". *Bull. Inst. Int. Statist.*, volumen 39, páginas 105–129, 1962.

[Wijsman, 1967] R. A. Wijsman. "Cross-sections of orbits and their application to densities of maximal invariants". En *Proc. Fifth Berkeley Symp. on Math. Stat. and Prob.*, volumen 1, páginas 389–400. 1967.

[Wijsman, 1985] R. A. Wijsman. "Proper action in steps, with application to density ratios of maximal invariants". *The Annals of Statistics*, volumen 10, páginas 955–961, 1985.

[Wilks, 1935] S. Wilks. "On the independence of k sets of normally distributed statistical variables". *Econometrica*, volumen 3, páginas 309–325, 1935.

[Winterhalder et al., 2005] M. Winterhalder, B. Schelter, W. Hesse, K. Schwab, L. Leistritz, D. Klan, R. Bauer, J. Timmer, y H. Witte. "Comparison of linear signal processing techniques to infer directed interactions in multivariate neural systems". *Signal Process.*, volumen 85, nº 11, páginas 2137–2160, 2005.

[Xiao et al., 2008] J.-J. Xiao, S. Cui, Z.-Q. Luo, y A. J. Goldsmith. "Linear coherent decentralized estimation". *IEEE Trans. Signal Process.*, volumen 56, nº 2, páginas 757–770, 2008.

[Zeng y Liang, 2009] Y. Zeng y Y.-C. Liang. "Eigenvalue-based spectrum sensing algorithms for cognitive radio". *IEEE Trans. Comm.*, volumen 57, nº 6, páginas 1784–1793, 2009.

[Zeng et al., 2008] Y. Zeng, Y.-C. Liang, y R. Zhang. "Blindly combined energy detection for spectrum sensing in cognitive radio". *IEEE Signal Process. Lett.*, volumen 15, páginas 649–652, 2008.