TECHNICAL UNIVERSITY OF CATALONIA (UPC)
DEPARTMENT OF SIGNAL THEORY AND COMMUNICATIONS

WIDEBAND COGNITIVE RADIO: MONITORING, DETECTION AND SPARSE NOISE SUBSPACE COMMUNICATION

PH.D. DISSERTATION BY
Josep Font-Segura

ADVISOR
Prof. Gregori Vázquez Grau

BARCELONA, MAY 2014

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Escola de Doctorat

Acta de qualificació de tesi doctoral

Curs acadèmic:

Nom i cognoms

Programa de doctorat

Unitat estructural responsable del programa

Resolució del Tribunal

Reunit el Tribunal designat a l'efecte, el doctorand / la doctoranda exposa el tema de la seva tesi doctoral titulada

_____

_____.

Acabada la lectura i després de donar resposta a les qüestions formulades pels membres titulars del tribunal, aquest atorga la qualificació:

☐ NO APTE ☐ APROVAT ☐ NOTABLE ☐ EXCEL·LENT

(Nom, cognoms i signatura) President/a	(Nom, cognoms i signatura) Secretari/ària	
(Nom, cognoms i signatura) Vocal	(Nom, cognoms i signatura) Vocal	(Nom, cognoms i signatura) Vocal

_____, _____ d'/de _____ de _____

El resultat de l'escrutini dels vots emesos pels membres titulars del tribunal, efectuat per l'Escola de Doctorat, a instància de la Comissió de Doctorat de la UPC, atorga la MENCIÓ CUM LAUDE:

☐ SÍ ☐ NO

(Nom, cognoms i signatura) President de la Comissió Permanent de l'Escola de Doctorat	(Nom, cognoms i signatura) Secretària de la Comissió Permanent de l'Escola de Doctorat

Barcelona, _____ d'/de _____ de _____

A l'Alba i als meus pares.

Wideband Cognitive Radio: Monitoring, Detection and Sparse Noise Subspace Communication

ABSTRACT

We are surrounded by electronic devices that take advantage of wireless technologies, from our computer mice, which require little amounts of information, to our cellphones, which demand increasingly higher data rates. Until today, the coexistence of such a variety of services has been guaranteed by a fixed assignment of spectrum resources by regulatory agencies. This has resulted into a blind alley, as current wireless spectrum has become an expensive and a scarce resource. However, recent measurements in dense areas paint a very different picture: there is an actual underutilization of the spectrum by legacy systems.

Cognitive radio exhibits a tremendous promise for increasing the spectral efficiency for future wireless systems. Ideally, new secondary users would have a perfect panorama of the spectrum usage, and would opportunistically communicate over the available resources without degrading the primary systems. Yet in practice, monitoring the spectrum resources, detecting available resources for opportunistic communication, and transmitting over the resources are hard tasks. This thesis addresses the tasks of monitoring, detecting and transmitting, in challenging scenarios including wideband signals, nonuniform sampling, inaccurate side information, and frequency-selective fading channels.

In the first task of monitoring the spectrum resources, this thesis derives the periodogram and Capon spectral estimates in nonuniform sampling exploiting a correlation-matching fitting from linearly projected data. It is shown that nonuniform sampling incurs the phenomenon of noise enhancement, which is circumvented by the proposed spectral estimates by implementing a denoising process, and further theoretically characterized in Bernoulli nonuniform sampling by establishing equivalence between nonuniform sampling and signal-to-noise ratio (SNR).

In the second task of detecting the available resources, this thesis considers the problems of multi-frequency signal detection, asymptotic performance, and cyclostationary signal detection. In multi-frequency signal detection, a unified framework based on the generalized likelihood ratio test (GLRT) is derived by considering different degrees of side information and performing maximum likelihood (ML) and correlation-matching estimation over the unknown parameters in uniform and nonuniform sampling, respectively. The asymptotic performance of signal detection is considered from two perspectives: the Stein's lemma, which allows discovering the influence of the main parameters on the error exponents in the error probabilities; and the asymptotic statistical characterization of the GLRT in Bernoulli nonuniform sampling, which allows the derivation of sampling walls in noise uncertainty, i.e., sampling densities below which the target detection probabilities

cannot be guaranteed. Finally, this thesis exploits the cyclostationarity properties of primary signals by deriving the quadratic sphericity test (QST), which is the ratio between the squared mean and the arithmetic mean of the eigenvalues of the autocorrelation matrix of the observations; and the optimal GLRT in a parameterized model of the frequency-selective channel, which exploits the low rank structure of small spectral covariance matrices.

In the last task of transmitting over the available resources, a cyclostationary secondary waveform scheme is first proposed to mitigate the interference that an active cognitive radio may cause to an inactive cognitive radio that performs spectrum sensing, by projecting the oversampled observations into a reduced subspace. Second, this thesis derives and statistically characterizes the sphericity minimum description length (MDL) for estimating the primary signal subspace. And third, this thesis finally considers the minimum norm waveform optimization problem with imperfect side information, whose benefits are those of linear predictors: flat frequency response and rotationally invariance.

Ràdio Cognitiva de Banda Ampla: Monitoratge, Detecció i Comunicació Dispersa de Subespai de Soroll

RESUM

Estem envoltats de dispositius electrònics que utilitzen tecnologia sense fils, des del ratolí de l'ordinador que requereix petites quantitats d'informació, fins als nostres telèfons mòbil que demanen cada vegada més velocitat de dades. Fins avui, la coexistència de tants serveis ha estat garantida per una assignació fixa dels recursos freqüencials per part de les agències de regulació. Això ens ha portat a un atzucac, ja que l'espectre actual ha esdevingut un recurs car i escàs. Tanmateix, mesures recents dibuixen una situació molt diferent: de fet hi ha una utilització molt baixa de l'espectre per part dels sistemes amb llicència.

La tecnologia de ràdio cognitiva promet millorar l'eficiència espectral dels futurs sistemes de comunicació sense fils. En teoria, un usuari secundari coneix perfectament la utilització de l'espectre, i és capaç de transmetre de manera oportuna sense degradar els sistemes primaris. A la pràctica, però, monitoritzar els recursos freqüencials, detectar-los i transmetre-hi són tasques difícils. Aquesta tesi tracta aquestes tres tasques en escenaris complicats com senyals de banda ampla, mostreig no uniforme, informació lateral imprecisa i canals selectius en freqüència.

En la primera tasca de monitoritzar els recursos freqüencials, aquesta tesi desenvolupa els estimadors espectrals de periodograma i Capon en mostreig no uniforme a partir d'un ajust per correlació de les observacions linealment projectades. Es demostra que el mostreig no uniforme genera el fenomen d'increment de soroll, el qual és solucionat pels estimadors espectrals proposats, i a més a més és caracteritzat teòricament pel cas de Bernoulli, establint una equivalència entre el mostreig no uniforme i la relació senyal soroll (SNR).

En la segona tasca de detectar els recursos disponibles, la tesi considera els problemes de detecció de senyals multifreqüència, avaluació de les prestacions asimptòtiques, i detecció de senyals cicloestacionàries. En detecció multifreqüència, es proposa una formulació unificada basada en el test generalitzat de màxima versemblança (GLRT), considerant diferents graus d'informació lateral, i efectuant estimació de màxima versemblança (ML) i d'ajust per correlació dels paràmetres desconeguts en mostreig uniforme i mostreig no uniforme, respectivament. Les prestacions asimptòtiques dels detectors són avaluades des de dues perspectives: el lema d'Stein, que permet descobrir la influència dels diferents paràmetres sobre els exponents de les probabilitats d'error; i la caracterització estadística asimptòtica del GLRT en mostreig no uniforme de Bernoulli, que permet derivar les parets de mostreig en incertesa de soroll, és a dir, aquelles densitats de mostreig per sota de les quals les probabilitats de detecció objectiu no són garantides. Finalment, la tesi explota les propietats cicloestacionàries dels senyals primaris: es deriva el test d'esfericitat quadràtica

(QST), que és la divisió entre la mitjana quadràtica i la mitjana aritmètica dels autovalors de la matriu de correlació de les observacions; i també es deriva el GLRT en un model parametritzat del canal selectiu en freqüència, que explota l'estructura rang deficient de petites matrius de covariància espectral.

En l'última tasca de transmetre en els recursos disponibles, es proposa en primer lloc un esquema de forma d'ona cicloestacionària per reduir la interferència que un usuari cognitiu pot causar a un altre usuari cognitiu inactiu que fa sensat de l'espectre, projectant les observacions sobremostrejades en un subespai reduït. En segon lloc, aquesta tesi deriva i caracteritza estadísticament la llargària mínima de descripció (MDL) d'esfericitat per estimar el subespai de senyal primària. I en tercer lloc, la tesi considera el problema d'optimització de forma d'ona de norma mínima amb informació lateral imperfecta, els beneficis del qual són els dels predictors lineals: resposta freqüencial plana i invariància a la rotació.

Agraïments

Escriure aquestes línies significa novament tancar una etapa important de la meva vida. Però com deia l'escriptor, el viatge no s'acaba mai, només els viatgers s'acaben, i fins i tot ells poden perdurar en forma de memòria, de record, de relat. Heu estat molts els viatgers que m'heu acompanyat durant aquests últims anys, que no tan sols conclouen en aquesta tesi, sinó que també m'han fet créixer com a persona.

Gràcies Gregori. Gràcies per tenir sempre els braços oberts i ensenyar-me a estimar el món de la recerca: ets una persona excepcional, en tots els sentits, i m'has ajudat a superar molts aspectes, tan professionals com de la vida. Aquesta tesi tampoc hauria estat possible sense la particular visió i les genials idees del Jaume, moltíssimes gràcies. Voldria també agrair al Xavi i al Francesc les bones estones de discussions tècniques, els bons dinars, i les experiències en els viatges de congrés. La Xell, el Josep, el Paco i el David també es mereixen el meu agraïment per haver tingut sempre la porta oberta.

Unes paraules d'agraïment per als companys de despatx, als que ja han marxat i als que encara queden, i en especial a la Míriam, a l'Adri, a l'Eva, al Màrius i a la Sandra, pels bons moments fent un cafè, o senzillament pels bon dia del matí.

Més sincers agraïments per als meus amics Joan, Mònica, Maria, Ton, Jordi, Aleix i Marcial per haver-me ofert la vostra complicitat i les vostres rialles constantment durant aquests anys: tenir-vos al meu voltant m'ha ajudat a arribar fins aquí indubtablement.

Finalment, voldria donar gràcies a la meva família, sobretot als meus pares i a la meva germana, per entendre les meves dèries i fer-me saber que sempre esteu al meu costat. Alba, des del primer moment tu has estat la peça clau en aquest viatge: gràcies per la teva perseverança, per la teva intensitat de viure i per compartir-ho amb mi, gràcies de tot cor.

– Josep

Aquesta tesi ha estat parcialment finançada per l'Agència de Gestió d'Ajust Universitaris i de Recerca (AGAUR) de la Generalitat de Catalunya mitjançant el programa d'ajuts destinats a universitats i centres de recerca per a la contractació de personal investigador novell (FI-DGR 2010).

List of Acronyms

AGC Analog-to-digital conversion.

AGM Arithmetic-to-geometric mean.

APSK Amplitude phase-shift keying.

BER Bit error rate.

BPSK Binary phase-shift keying.

CDF Cumulative density function.

CDMA Code division multiple-access.

CSI Channel state information.

DPC Dirty paper coding.

DSA Dynamic spectrum access.

DS-CDMA Direct-sequence CDMA.

DSSS Direct-sequence spread-spectrum.

DVB-T Terrestrial digital video broadcasting.

EVD Eigenvalue decomposition.

FCC Federal Communications Commission.

FD Feature detector.

FDM Frequency-division multiplexing.

FDMA Frequency-division multiple-access.

FH-CDMA Frequency-hopping CDMA.

FSSS Frequency-hopping spread-spectrum.

GLR Generalized likelihood ratio.

GLRT Generalized likelihood ratio test.

GSM Global system for mobile communications.

ISI Inter symbol interference.

KLD Kullback-Leibler divergence.

LDPC Low-density parity-check.

LMPIT Locally most powerful invariant test.

LoS Line of sight.

LSE Least squares estimation.

LRT Likelihood ratio test.

MAC Multiple-access channel.

MAP Maximum a posteriori.

MCSCMD Multi-cycle spectral correlation magnitude detector.

MDL Minimum description length.

MIMO Multiple-input multiple-output.

ML Maximum likelihood.

MME Maximum to minimum eigenvalue.

NME Normalized maximum eigenvalue.

OFDM Orthogonal frequency-division multiplexing.

OFDMA Orthogonal frequency-division multiple-access.

OSTBC Orthogonal space-time block coding.

OSTC Orthogonal space-time coding.

PAM Pulse amplitude modulation.

PDF Probability density function.

PER Packet error rate.

PSD Power spectral density.

QAM Quadrature amplitude modulation.

QoS Quality of service.

QST Quadratic sphericity test.

RIP Restricted isometry property.

ROC Receiver operating characteristics.

SDMA Space division multiple-access.

SDR Software-defined-radio.

SINR Signal-to-interference-plus-noise ratio.

SNR Signal-to-noise ratio.

SRRC Squared-root raised cosine.

TDMA Time-division multiple-access.

UWB Ultra wideband.

WLAN Wireless local area network.

WRAN Wireless regional area network.

WSS Wide-sense stationary.

Notation

$x(t)$	A random process.
$X(\omega)$	The Fourier transform of $x(t)$.
$\mathbf{x}$	A column vector.
$\mathbf{X}$	A matrix.
$\phi_x(\omega)$	The PSD of $x(t)$.
$\mathbf{I}$	The identity matrix.
$\mathbf{0}$	All-zeroes matrix or vector with appropriate dimensions.
$\mathrm{tr}(\mathbf{X})$	The trace of $\mathbf{X}$.
$\det(\mathbf{X})$	The determinant of $\mathbf{X}$.
$\mathrm{diag}(\mathbf{x})$	A matrix with the elements of $\mathbf{x}$ on the main diagonal.
$\mathrm{diag}(\mathbf{X})$	A vector with the elements of the main diagonal of $\mathbf{X}$.
$\|\mathbf{X}\|_F$	The Frobenius norm of $\mathbf{X}$.
x^+	The operator $\max(0, x)$.
$\mathbb{R}_+$	The set of positive real numbers.
$\mathbb{E}$	The expectation operator.
$\mathbb{B}$	The set of bandlimited signals.
$\mathcal{N}(\mathbf{m}, \mathbf{R})$	Gaussian distribution with mean $\mathbf{m}$ and covariance matrix $\mathbf{R}$ of appropriate dimensions.
$\mathcal{CN}(\mathbf{m}, \mathbf{R})$	Complex Gaussian distribution with mean $\mathbf{m}$ and covariance matrix $\mathbf{R}$ of appropriate dimensions.
$\mathcal{X}_r^2(\lambda)$	Chi-squared distribution of r degrees of freedom and λ as non-centrality parameter.

Contents

Abstract iii

Resum v

Agraïments vii

List of Acronyms viii

Notation xi

1 Introduction 1
- 1.1 Scope . . . 1
- 1.2 Publications . . . 2
- 1.3 Contributions and Organization . . . 4

2 Signal Processing for Wideband Cognitive Radio 8
- 2.1 Cognitive Radio . . . 8
- 2.2 Nonuniform Sampling . . . 14
- 2.3 Spectrum Sensing . . . 19
- 2.4 Dynamic Spectrum Access . . . 24

3 Spectral Analysis in Nonuniform Sampling 30
- 3.1 Introduction . . . 30
- 3.2 Nonuniform Correlation-Matching . . . 32
- 3.3 Bernoulli Nonuniform Sampling . . . 37
- 3.4 Numerical Results . . . 44
- 3.5 Conclusions . . . 48

4 Multi-Frequency Primary Signal Detection 50
- 4.1 Introduction . . . 50
- 4.2 Maximum Likelihood Estimates in Uniform Sampling . . . 53
- 4.3 Correlation-Matching Estimates in Nonuniform Sampling . . . 64

4.4 Numerical Results . . . 67
4.5 Conclusions . . . 77

5 Asymptotic Performance in Primary Signal Detection 79
5.1 Introduction . . . 79
5.2 Stein's Lemma . . . 81
5.3 Nonuniform Sampling Walls . . . 96
5.4 Numerical Results . . . 108
5.5 Conclusions . . . 117

6 Cyclostationary Primary Signal Detection 118
6.1 Introduction . . . 118
6.2 Quadratic Sphericity Test . . . 120
6.3 Frequency-Domain Detection . . . 122
6.4 Numerical Results . . . 126
6.5 Conclusions . . . 130

7 Noise Subspace Communication 132
7.1 Introduction . . . 132
7.2 Cognitive Cyclostationary Waveform . . . 135
7.3 Sphericity Minimum Description Length . . . 139
7.4 Minimum Norm Waveform Optimization . . . 144
7.5 Numerical Results . . . 152
7.6 Conclusions . . . 165

8 Conclusions and Future Work 167
8.1 Conclusions . . . 167
8.2 Future Work . . . 170

A Proofs of Chapter 3 174

B Proofs of Chapter 4 176

C Proofs of Chapter 5 183

D Proofs of Chapter 6 190

E Proofs of Chapter 7 195

References 226

Chapter 1
Introduction

1.1 Scope

It is well known that every single device that transmits wirelessly must satisfy tight regulations, not only to confine interference, but also because wireless spectrum is an expensive scarce resource. This scarcity has become more notorious due to the recent increase of wireless systems which demand higher data rates. However, recent measurements show an opposite situation: many licensed services transmit only sporadically, giving an overall spectrum occupancy of roughly 35%, even in most crowded areas. It is precisely this inefficiency in the use of the frequency resources which has motivated the current research in cognitive radio technology. Ideally, cognitive radios have a perfect picture of the spectrum usage in a given place and a given time, being able to smartly adapt the transmission scheme to perform opportunistic communication.

Despite cognitive radio technology holds a big promise in the proliferation of new wireless services, there still exist many technical challenges which are addressed in this thesis. More precisely, this thesis has focussed on a variety of statistical signal processing problems involved in

1. monitoring the spectral resources,
2. detecting available resources for opportunistic communication, and
3. transmitting over the available resources without or causing little interference to the licensed services,

in several challenging scenarios, such as

1. wideband signals,
2. nonuniform sampling,
3. inaccurate side information, and
4. frequency-selective fading channels.

Wideband signals are costly to monitor due to the cumbersome required sampling rates, but also difficult to detect due to the low transmission power and spectrum shape. Nonuniform sampling has been proposed to reduce the sampling rate, however monitoring and detecting nonuniformly sampled wideband signals are not easy tasks, as random sampling affects the statistics of the signal. Also, a good calibration of the receiver noise and legacy signal parameters has a fundamental role in monitoring and detecting wideband signals, as uncertainties can lead to detection walls beyond which the detection fails to be robust. Another issue involved in the detection is the interference caused by an active cognitive radios, which can lead to licensed signal missed detections by other silent cognitive radios which are far from the licensed system but close to the transmitting cognitive radio. Furthermore, the frequency-selective fading nature of the channels between the licensed services and the cognitive radios play an important part, as the statistical properties of the signals involved in detecting the available resources and transmitting over the available resources are degraded. A last example of research challenge consists of the design of cognitive radio waveforms in a decentralized network, where each cognitive radio has different versions of the spectrum occupancy due to the different propagation channels between the legacy system and the cognitive radio network.

This thesis has tackled the aforementioned problems whose research has lead to the publications detailed in Section 1.2, and to the core contributions reported in Section 1.3, which summarize the core structure of the thesis.

1.2 Publications

The following journal publications have been produced as a result of this thesis research.

[J1] **J. Font-Segura**, G. Vázquez and J. Riba, "Robust minimum norm waveform optimization for wideband cognitive radio communications", *IEEE Transactions on Communications*, 2014, (to submit).

[J2] **J. Font-Segura**, G. Vázquez and J. Riba, "Sphericity minimum description length", *IEEE Signal Processing Letters*, 2014, (to submit).

[J3] **J. Font-Segura**, G. Vázquez and J. Riba, "Single and Multi-Frequency Wideband Spectrum Sensing with Side Information ", *IET Signal Processing*, 2014 (in press).

[J4] J. Riba, **J. Font-Segura**, J. Villares and G. Vázquez, "Frequency-domain GLR detection of a second-order cyclostationary signal over dading channels", *IEEE Transactions on Signal Processing*, Vol. 62, No. 8, pp. 1899–1912, April 2014.

[J5] **J. Font-Segura**, G. Vázquez and J. Riba, "Nonuniform sampling walls in wideband signal detection", *IEEE Transactions on Signal Processing*, Vol. 62, No. 1, pp. 44–55, January 2014.

[J6] **J. Font-Segura** and X. Wang, "GLRT-based Spectrum Sensing for Cognitive Radio with Prior Information", *IEEE Transactions on Communications*, Vol. 58, No. 7, pp. 2137–2146, July 2010.

as well as the following publications in conference proceedings

[C1] **J. Font-Segura**, J. Riba, J. Villares and G. Vázquez, "Frequency-domain GLR detection of cyclostationary signals in frequency-selective channels", *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Florence, Italy, May 2014.

[C2] **J. Font-Segura**, G. Vázquez and J. Riba, "Exploiting cognitive cyclostationary noise subspace for noncognitive spectrum sensing", *IEEE International Workshop on Signal Processing Advances for Wireless Communications (SPAWC)*, Darmstadt, Germany, June 2013.

[C3] **J. Font-Segura**, G. Vázquez and J. Riba, "Sampling walls in signal detection of Bernoulli nonuniformly sampled signals", *IEEE International Conference on Communications (ICC)*, Budapest, Hungary, June 2013.

[C4] **J. Font-Segura**, J. Riba, J. Villares and G. Vázquez, "Quadratic sphericity test for blind detection over time-varying frequency-selective fading channels", *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Vancouver, Canada, May 2013.

[C5] **J. Font-Segura**, G. Vázquez and J. Riba, "Asymptotic error exponents in energy-detector and estimator-correlator signal detection", *IEEE International Conference on Communications (ICC)*, Ottawa, Canada, June 2012.

[C6] **J. Font-Segura**, G. Vázquez and J. Riba, "Novel periodogram and Capon spectral analysis based on nonuniform sampling", *IEEE Global Communications Conference (GLOBECOM)*, Houston, TX, December 2011.

[C7] **J. Font-Segura**, G. Vázquez and J. Riba, "Noise enhancement and SNR equivalence in Bernoulli nonuniform sampling", *International Conference on Cognitive Radio and Advanced Spectrum Management (CogART)*, Barcelona, Spain, October 2011.

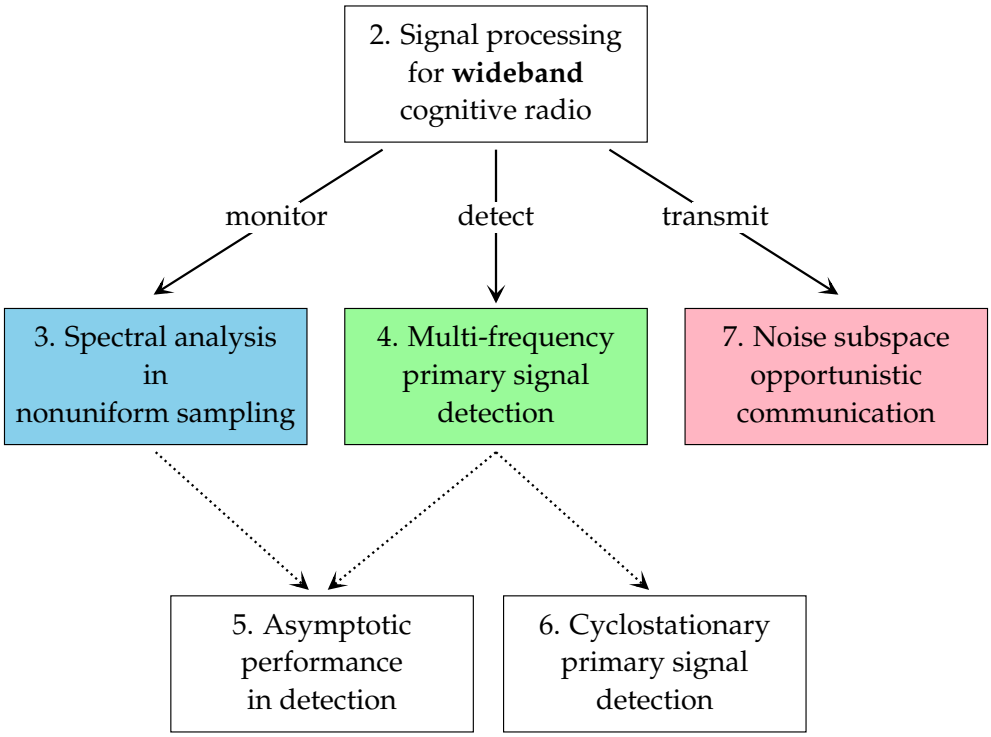

Figure 1.1: Thesis structure.

[C8] **J. Font-Segura**, G. Vázquez and J. Riba, "Multi-frequency GLRT spectrum sensing for wideband cognitive radio", *IEEE International Conference on Communications (ICC)*, Kyoto, Japan, June 2011.

[C9] **J. Font-Segura**, G. Vázquez and J. Riba, "Compressed correlation-Matching for spectrum sensing in sparse wideband regimes", *IEEE International Conference on Communications (ICC)*, Kyoto, Japan, June 2011.

1.3 Contributions and Organization

The contributions of this thesis are structured into six parts, as depicted in Figure 1.1, which correspond to Chapters 2–7. In the sequel, the contributions and the publications derived from each chapter are detailed.

Chapter 2 introduces the reader to the cognitive radio technology and the signal processing research challenges in two fundamental issues of cognitive radio communication: spectrum sensing and dynamic access, i.e., how the cognitive radios obtain information on the spectrum usage and how the cognitive radios actually transmit over the available resources, respectively. This Chapter further reports the most relevant background bibliography on state-of-the-art works related to the specific signal processing problems of nonuniform sampling, signal detection, and dynamic spectrum access.

The first problem of monitoring the spectral resources by means of spectral analysis in nonuniform sampling is addressed in Chapter 3. Nonuniform sampling is motivated by the fact that the spectrum usage is sparse, which allows the application of the foreseen theory of compressed sensing to further reduce the sampling rate. However, by simply exploring classical spectral estimation over nonuniformly sampled data, it is observed that nonuniform sampling incurs a phenomenon of noise enhancement. In order to circumvent this phenomenon, the periodogram and Capon spectral estimates based on the correlation-matching metric are designed to match the spectral contribution of the legacy signal and remove the spectral contribution due to nonuniform sampling. Chapter 3 further provides a theoretical framework to quantize and predict the noise enhancement effect in nonuniform sampling by resorting to the simple but utile Bernoulli nonuniform sampling. In particular, a signal-to-noise ratio (SNR) equivalence is derived to predict the conditions in which any signal processing problem will operate. The SNR equivalence is extended to matrix formulation, which stands for an important tool in the subsequent chapters. The contributions of Chapter 3 have derived to the publications [C6] and [C7].

The second problem of detecting a signal immersed in additive noise is addressed in Chapter 4. Statistical signal detection is one of the most fundamental problems for decision making in a wide range of applications, from radar to biomedicine. The performance of signal detectors is clearly improved as more side information on the noise and signal statistics is exploited. However, inaccurate knowledge on these parameters incurs severe degradation in the detector's performance, which can lead to nonrobust behaviors such as SNR walls due to noise uncertainty. For that reason, Chapter 4 addresses optimal detection in the Neyman-Pearson sense by means of the generalized likelihood ratio test (GLRT), which natively incorporates joint parameter estimation for inaccurate model parameters. Several scenarios of partial side information are addressed, and a low complexity detector for multi-frequency systems is proposed. Chapter 4 is divided into two parts: the first part addresses the estimation involved in the GLRT by maximum likelihood (ML) estimation in uniform sampling, whereas the second part addresses the estimation from a correlation-matching view point from nonuniform sampling in a similar way that it has been proposed in Chapter 3. The publications [C8], [C9], [J3] and [J6] have been produced as a consequence of the research presented in this chapter.

It is well known that the statistical characterization of signal detectors is an important issue to predict and benchmark the detectors' performance within the detection conditions, i.e., the SNR, the sensing time, the channel diversity, the legacy signal occupancy, or the nonuniform sampling rate. However, the statistical characterization is, in general, a cumbersome problem which has no direct solution. In contrary, Chapter 5 adopts the asymptotic performance of the GLRT detector by means of two tools. On the one hand, the Stein's lemma is a fundamental result that relates the rate in which the error probabilities decay within the number of observations, with the Kullback-Leibler divergence (KLD)

between the probability density functions (PDFs) of the hypotheses. More specifically, the Stein's lemma is applied to the energy detector and the estimator-correlator in a wide variety of scenarios including channel diversity and Bernoulli nonuniform sampling. The derived expressions show that the error exponents of the false-alarm and missed-detection probabilities depend on the observation size and the second-order statistics of the problem, leading to the fact that the observation size scales as the inverse of a monotonically increasing function of the SNR. On the other hand, the asymptotic performance of the signal detector in Bernoulli nonuniform sampling in the presence of noise uncertainty is addressed in the second part of Chapter 5 by approximating the PDF of the detector by Gaussian distributions. As a consequence, this thesis shows the existence of sampling walls, i.e., the sampling density below which the target error probabilities, i.e., the missed-detection and false-alarm probabilities, cannot be guaranteed at a given SNR regardless the number of acquired samples. The contributions of Chapter 5 have produced the publications [J5], [C3] and [C5].

An alternative approach to fight against noise uncertainty is cyclostationary signal detection, as cyclostationary processes exhibit a form of frequency diversity which can be exploited by cognitive radios to perform spectrum sensing invariant to noise variance. However, these detectors are highly sensitive to frequency-selective fading channels. Chapter 6 addresses the design of cyclostationary detectors in the presence of unknown channel gains. In particular, a first invariant cyclostationarity detector is proposed by deriving a blind detector based on the squared mean to arithmetic mean of the eigenvalues of the autocorrelation matrix of the observations, denoted as the quadratic sphericity test (QST). The main advantage of the QST is that it shows invariance with respect to both the noise variance and the channel gains. Chapter 6 further exploits cyclostationarity by adopting a rank-1 frequency-domain representation of digital waveforms. This allows the derivation of an optimal frequency-domain GLRT which addresses ML estimation of a parametric channel model based on the coherence bandwidth. The proposed detector outperforms classical spectral correlation magnitude detectors by exploiting the rank-1 structure of small spectral covariance matrices. The publications [J4], [C1] and [C4] have been derived from the research presented in this chapter.

The third problem of cognitive radio waveform optimization is reported in Chapter 7, which is organized into three parts. Firstly, the cognitive radio transmission scheme is design so that the interference generated to other cognitive radios which are performing spectrum sensing is minimized. If the cognitive radio network is localized in a small area, the transmitting users may act as strong interference to the inactive users. By operating directly on the oversampled cyclostationary signal subspace, the inactive cognitive radios are able to exploit the cyclostationarity in order to identify a noise subspace where legacy signal detection can be done free of interference. A further analysis on the deflection as performance indicator reveals that the performance of the detector is proportional to the

roll-off factor of the squared-root raised cosine (SRRC) pulse employed by the cognitive radio transmitters, and proportional to the squared of the SNR. Chapter 7 also considers the optimization of waveforms that lie in the noise subspace of the legacy signal second-order statistics. This requires the previous determination of the legacy signal subspace dimension from local observations, i.e., solving a model order selection problem. Therefore, the second part of Chapter 7 exploits the minimum description length (MDL) statistic with unknown noise variance and unknown low-rank signal correlation matrix. This leads to the formulation of the GLRT, which has been recently reported as the sphericity test for low-rank Gaussian signals. Because the GLRT asymptotically follows a non-central Chi-squared distribution, the statistical characterization of the proposed sphericity MDL, together with the MDL with known noise variance and imperfect known noise variance is provided. The last research of this thesis seeks cognitive radio waveforms that satisfy the favorable properties of white frequency response, invariance to rotation, and that lie on the noise subspace of the second-order statistics of the legacy signal. It is shown that linear predictors that arise from minimum norm filtering exhibit these properties. Therefore, the resulting waveforms implement a form of null space communication similar to that in multiple antennas, but will show enhanced stability and low complexity benefits of temporal correlation matrices. In front of statistical channel state information (CSI) mismatching, a robust waveform optimization problem is addressed, based on the worst-case performance scenario, which regularizes the objective function by adding a penalty term equivalent to the squared value of the spectrum. Both the nonrobust and robust waveform optimization problems are solved iteratively, and show efficient implementation in the frequency-domain. Chapter 7 has produced the publications reported in [J1], [J2] and [C2].

Finally, Chapter 8 summarizes the main conclusions obtained in the thesis, as well as brief description of future lines of research.

Chapter 2

Signal Processing for Wideband Cognitive Radio

This Chapter is structured as follows. Section 2.1 motivates the cognitive radio technology and outlines the main signal processing research challenges. The problems of nonuniform sampling, spectrum sensing, and dynamic spectrum access (DSA), along with state-of-the-art literature, are detailed in Sections 2.2, 2.3 and 2.4, respectively.

2.1 Cognitive Radio

2.1.1 Motivation

Today's wireless communications are characterized by a fixed spectrum resource assignment, i.e., the spectrum is regulated by the governmental agencies and is assigned to the license holders for a long term in a large geographical region. However, a large portion of this assigned spectrum is used only sporadically by the primary services, who concentrate the communication over certain portions of the spectrum while a significant amount of the spectrum remains underutilized. According to recent studies by the Federal Communications Commission (FCC), the utilization of the available spectrum can be as low as 15% in determined geographical areas [FCC03].

Specifically out of this premise was born the new paradigm of *cognitive radio*. Cognitive radios utilize advanced signal processing along with novel DSA policies to support new wireless users who opportunistically communicate in the existing congested spectrum without degrading established users. For that purpose, a cognitive radio is an adap-

tive wireless communications system that takes advantage of all type of side information on the network, e.g., users' activity, channel conditions, modulations and codes format, or even the information message sent by other users who share the spectrum [GJMS09]. Depending on the degree of knowledge on this side information, cognitive radio systems are based on underlying, overlaying or interweaving the new users' signals with those of existing users [Kod05]. In detail:

1. The *underlay* cognitive radios operate only if the interference caused to the non-cognitive services is below a given threshold. This approach assumes that the secondary users have perfect knowledge of the interference caused by their transmitters to the primary services. One possible approach for transmitting signals causing little interference is to use a large bandwidth over which the cognitive radio signal is spread below the noise level. This is the basis idea of spread spectrum and ultra wideband (UWB) communications. Underlay cognitive radio is the most common scheme when coexisting with licensed primary services, e.g., UWB underlays many licensed bands.

2. In *overlay* cognitive radio, the secondary users can transmit at any power and simultaneously with the primary users by using sophisticated techniques like dirty paper coding (DPC) or relays. For this to be possible, the cognitive transmitters need to know the primary codebooks and messages, which may be impractical in many situations.

3. Thirdly, the idea of opportunistic communication is exploited in *interweave* cognitive radios, which indeed is the original motivation of cognitive radio [MM99]. The spectrum gaps or holes can be used sporadically by secondary users for their communication in given time and geographical location conditions. The primal requirement of interweave cognitive radio is that the secondary users must be aware of the spectral activity of the primary systems.

Even though the concept of cognitive radio, under the umbrella of software-defined-radio (SDR) and DSA, was proposed in the early 2000s, it is still a state-of-the-art topic that receives much attention in the research society. For a detailed collection of recent work on advanced signal processing for cognitive radio networks the reader is referred to [ZKL$^+$11, LCL$^+$11a, LCL$^+$11b], whereas the practical implementation and application of the cognitive radio technology, standardization issues and testbeds is reported in, e.g., [WGC11, FHMI11, PND$^+$11, NIK12].

This thesis focusses on the interweave concept of cognitive radio and its signal processing and communications challenges. Specifically, two of the main functionalities indispensable for the cognitive radio technology to proliferate are investigated: spectrum sensing and dynamic access. On the one hand, *spectrum sensing* refers to the signal processing

detection techniques that reliably identify the radio resources unused by the primary systems in order to cause little interference. On the other hand, *dynamic access* consists of the enabling communication strategies that capture the best multiple access technique in order to meet the user communications requirements. The envisioned open fields on spectrum sensing and dynamic access are discussed in Section 2.1.2 and Section 2.1.3, respectively.

2.1.2 Challenges in Spectrum Sensing

For the last decades, cognitive radio has been one of the major focus of academic research, e.g., [Hay05, MLJ09, WL11, ALLP12], as well as application initiatives such as the IEEE 802.22 standard on wireless regional area network (WRAN) [CCBSS05]. Yet, there are many open research challenges which are identify in the following for both spectrum sensing and dynamic access. Spectrum awareness is an important prerequisite in the envisioned applications of wireless cognitive radio networks. Creating an interference map of the operational spatial region plays a fundamental role in enabling spatial frequency reuse and allowing DSA. Spectrum awareness can be obtained through

1. Databases (e.g., geolocation, modulation and propagation models),
2. Sensing, or
3. Beacons

This thesis focusses on sensing for spectrum awareness, as it constitutes a broader solution and requires much less infrastructure. Therefore, designing fast and reliable spectrum sensing techniques based on cognitive radios local observations is a challenging task with many open research topics.

1. *Sensing duration and sensing rate*.

   The time needed by the spectrum sensing algorithm to detect the presence of primary user activity is probably the most primordial requirement of cognitive radio communications, especially when transmitting because the primary users can claim their spectrum resources at any time. In order to avoid interference from and to the primary users, the secondary users must identify the presence of primary activity as fast as possible and immediately vacate the used band. The main factor that determines a tradeoff between sensing time and detection performance is the selection of the parameters to be sensed.

   On the other hand, the sensing rate, i.e., how often the cognitive radios perform spectrum sensing, is another important design parameter. The optimum value of the sensing rate depends on the technical capabilities of the cognitive radio and the statistical properties of the primary services [JJL+07]. For instance, if the status of the

primary users is known to change slowly in time, spectrum sensing can be performed in a more relaxed manner, whereas a more constant spectrum sensing is required in public safety bands to immediately vacate the band to prevent any interference. Selecting the optimum sensing rate is a novel research problem, and the goal is to maximize the cognitive radio channel capacity while maintain the interference to primary users constraint. Many approaches can be proposed to solve these problems. For instance, one could use the guard interval in orthogonal frequency-division multiplexing (OFDM) systems to perform sensing, or focusing only in the changing parts of the spectrum.

Another problem associated to sensing time is that spectrum sensing cannot be performed on spectrum resources over which the cognitive radio users are communicating. To mitigate this problem, frequency-hopping has arisen as a practical solution, as reported in [HWA+07].

2. *Detection of spread-spreading primary users*.

   Typically, primary system either consists of a fixed narrowband or spread spectrum services. The two major approaches of spread spectrum communications are the frequency-hopping spread-spectrum (FSSS) and the direct-sequence spread-spectrum (DSSS). In both situations, detecting such systems is difficult as the power of the primary users is distributed over a wide frequency range, even though the actual information bandwidth is much narrower [CMB04]. One possible solution is to uncover the hopping pattern or achieve perfect synchronization with the primary users, however such detection in the code-domain is not straightforward.

3. *Wideband sensing*.

   Very high sampling rates are required by conventional spectral estimation methods which have to operate at or above the Nyquist rate when the frequency range to be sensed is very large. Meanwhile, the stringent timing requirements for monitoring the dynamically changing spectrum only allow for a limited number of measurements to be collected for sensing, which makes it challenging to reliably perform high-resolution signal reconstruction. Furthermore, the wideband regime is characterized by close to zero spectral efficiency and signal-to-noise ratio (SNR) close the minimum required for reliable communication [Ver02]. Therefore, spectrum sensing becomes indeed especially challenging in the wideband regime.

4. *Cooperative sensing*.

   Cooperative sensing becomes a practical methodology for increasing the reliability of the cognitive radio network spectrum sensing performance and addresses some problems of individual spectrum sensing such as the hidden primary user problem

or punctual deep fading. However, sharing the information among cognitive radios and combining the measurements is a challenging task.

On the one hand, the shared information can be based on soft-decisions or hard-decisions. Soft-decisions outperform hard-decisions in terms of probability of missed-detection; however they require a larger overhead for transmitting such information.

On the other hand, there are many approaches for making the final decision. In a distributed network, each node can base its decision on the local measurements of the neighborhood and pass the information in an intelligent manner (e.g., belief propagation). In a centralized network, a decision center collects all the measurements/decisions from the cognitive radios and performs the final decision employing optimal combination (e.g., the likelihood ratio test (LRT)), or alternatively simpler techniques such as equal gain-combining, selection combining, or switch and stay combining.

5. *Complexity*.

   Finally, complexity is one of the major factors affecting the implementation of practical sensing methods. To detect a signal at very low SNR and in a harsh environment is not a simple task.

   The research objectives of this thesis focus on statistical-based detection methods. The major advantage of such spectrum sensing detectors is their little dependence on signal and channel knowledge, as well as relatively low complexity. A promising research line is then the design of statistical-based spectrum sensing algorithms that exploit the signal features. At last, detecting the presence of a primary user signal is only the basic task of sensing. For a radio with high level of cognition, further information such as the modulation format employed by the primary service may be exploited. Therefore, signal identification turns to be an advanced task of sensing.

6. *Training*.

   Training is the task of guiding a cognitive radio engine through the process of learning a desired system's behavior and capabilities. The training speed and expected performance during this task are of paramount importance to the system's operation, especially when the system is facing new conditions [VB12].

2.1.3 Challenges in Dynamic Wideband Access

All the cognitive nodes willing to transmit or receive information perform spectrum sensing to acquire a map on the situation of the available spectrum prior to opportunistic communication. Recent theoretical work (see [JS07] and the references therein for an information theory formulation of cognitive radio channels) recognizes that the spectral environment is distributed and dynamic. The distributed access occurs due to the fact that the

primary activity detected by a cognitive radio differs from that detected around the other cognitive radios.

Multiple-access techniques are in charge of regulating the access of different users to a common access radio resource, which usually consists of a dedicated frequency bandwidth, but not limited to other natures such as time, code, or space [Pro00]. The common resource in which a set of M users wish to communicate is referred as multiple-access channel (MAC) [Cov72]. By its nature, the MAC is interference limited in dense networks. This fact, though, becomes a challenge for cognitive radio MATs, when contrasted to the dedicated multiple-access techniques, i.e., classical techniques where each user or service has a dedicated (not shared) radio resource. A typical example for such techniques is the time-division multiple-access (TDMA) employed in the global system for mobile communications (GSM) network. When considering the MAC, many information theoretical network issues arise. How do the various senders cooperate with each other to send information to the receivers? What rates of communication are simultaneously achievable? What limitations does interference among the senders and to the primary services put on the total rate of communications?

There are still many open lines of investigation in order to answer all these questions. In the field of DSA [ZS07], the following open challenges are identified.

1. *Sharing dimension*.

   In general, spectrum sharing can be done at any dimension of the spectrum space, i.e., frequency, code, space, and time. Selecting the sharing dimension in cognitive radios is a design parameter that depends on the primary user activity.

   In frequency-division multiple-access (FDMA) and code division multiple-access (CDMA), the main challenge is to perform efficient channel allocation. More advanced signal processing techniques are required in space multiple-access as power control is necessary to avoid interference to the primary systems. Finally, time access is based on the concept of time leases.

2. *Opportunistic Access.*

   Contrary to open spectrum access, i.e., a cognitive radio accessing unlicensed bands, opportunistic DSA is regulated by the priority established between primary and secondary users.

   The primary task in opportunistic access is channel selection. The choice of the frequency bands that can be used for opportunistic communication depends on the amount of primary systems occupancy information available at the cognitive radio receiver.

   (a) *Narrowband Spectrum Access*. In Narrowband DSA, the cognitive radios can transmit and receive the information on a predetermined or dynamically chosen

frequency band, depending on the primary user occupancy. This gives rise to two techniques: frequency-hopping and frequency-tracking. On the one hand, in frequency-hopping spectrum access, the cognitive transmitter and receiver simultaneously hop across multiple frequencies according to a predetermined sequence. The primal challenge of the frequency-hopping approach is that the transmitter and the receiver need always to be synchronized. On the other hand, frequency-tracking consists of choosing one of the free frequency bands for communication. In this setting, the receiver chooses the expected frequency band based on local observations so as to maximize the probability of coincidence with the transmitter.

(b) *Wideband Spectrum Access*. In a wideband cognitive radio system, the transmitter and receiver can scan the spectral activity in all the frequency bands. Then, the cognitive radios can span the information using a spanning code over multiple frequency slots that are potentially unused by the primary systems. Unlike the narrowband spectrum access schemes, the secondary users need to have information of all the bands prior to every transmission. The basic challenge in this scheme then goes to the spectrum sensing problem.

3. *Centralized or Distributed Spectrum Access*.

   An important aspect of a dynamic spectrum access is whether decisions on the access are taken based on complete or partial information about the current utilization of the spectrum. In centralized cognitive radio networks, a central device is responsible for gathering the spectrum utilization from all the cognitive radios and performing spectrum allocation and controlling based on the global information on the network. Although the performance of centralized spectrum access is optimal due to the acquisition of network information, reporting to a central device may not be practical in many situations due to communication overhead. Alternatively, in distributed approaches (e.g., [TPD+10]) the spectrum access is controlled by each individual cognitive radio based on its own or common policies imposed by the MAT, and neighborhood information. Distributed spectrum access reduces the reporting overhead, but achieves suboptimal performance due to the lack of complete information on the network.

2.2 Nonuniform Sampling

Conventional approaches to sampling signals follow the well-known Nyquist theorem, in which the sampling rate must be at least twice the maximum frequency present in the information signal. However, the new theory of compressed sensing asserts that one can recover certain signals from few samples than the Nyquist rate. Compressed sensing has

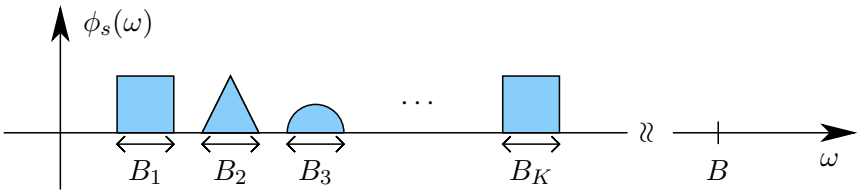

Figure 2.1: Multi-frequency primary signal allows nonuniform sampling.

arisen from the mathematical society and has been applied to many engineering fields. It relays on two fundamental characteristics: the sparsity of the signal to be sampled, and the incoherence between the sampling method and the underlying domain in which the signal is sparse. The concept of sparsity expresses the idea that the information rate behind a signal can be much smaller than the one suggested by its bandwidth (less degrees of freedom compared to its length), when the signal to be sampled has low occupancy in the sensed bandwidth, e.g., as depicted in Figure 2.1.

In the following, the mathematical fundamentals of the classical compressed sensing theory are reported, as well as a discussion on the state-of-the-art works which are based or relate to compressed sensing.

2.2.1 Classical Compressed Sensing Theory

The compressed sensing theory has been motivated by E. J. Candès in the last years [CT05, CRT06, CW08]. Consider a continuous-time random process $s(t)$, and an arbitrary N-size sampled vector $\mathbf{s} \doteq (s(t_1), \ldots, s(t_N))^T$, where the sampling intervals t_n accomplish the Nyquist theorem. It is said that $\mathbf{x}$ is *sparse* in the $\mathbf{\Phi}$ domain, if its representation onto the basis $\mathbf{\Phi}$ has only S non-zero coefficients, with $S \ll N$. The compressed sensing theory claims that $\mathbf{s}$ can be perfectly recovered from the compressed signal $\mathbf{x} = \mathbf{\Psi}(\mathbf{s} + \mathbf{w})$, where $\mathbf{w}$ is the additive Gaussian noise with double-sided spectral density $N_0/2$, and $\mathbf{\Psi}$ is a $K \times N$ nonuniform sampling matrix with enough incoherence with $\mathbf{\Phi}$. The sparsest solution of the coefficients of $\mathbf{s}$ onto $\mathbf{\Phi}$ is given by the ℓ_0-norm constrained minimization problem

$$\hat{\boldsymbol{\alpha}} = \arg\min_{\boldsymbol{\alpha}} \|\boldsymbol{\alpha}\|_{\ell_0} \tag{2.1}$$

subject to $\mathbf{y} = \mathbf{\Psi}\mathbf{\Phi}\boldsymbol{\alpha}$, where $\|\boldsymbol{\alpha}\|_{\ell_0}$ denotes the ℓ_0-norm of $\boldsymbol{\alpha}$, which is equal to the number of non-zero elements in $\boldsymbol{\alpha}$, i.e., $\|\boldsymbol{\alpha}\|_{\ell_0} \doteq \lim_{d\to 0} \sum_{n=1}^{N} |\alpha_n|^d$. Solving (2.1) is a NP-hard problem [CT05] and very sensitive to the noise. Hence, several alternatives to (2.1) have been proposed in the literature. Basis pursuit [CRT06] is one of the most used algorithms as replaces the ℓ_0-norm in (2.1) by the ℓ_1-norm, with a relaxed second-order constraint

$$\hat{\boldsymbol{\alpha}} = \arg\min_{\boldsymbol{\alpha}} \|\boldsymbol{\alpha}\|_{\ell_1}, \tag{2.2}$$

subject to $\|\mathbf{x} - \mathbf{\Psi\Phi\alpha}\|^2 \leq \epsilon$. The main advantage of basis pursuit is that it leads to linear programming, so (2.2) can efficiently solved in polynomial time. However, due to the performance degradation in noisy environments, several improvements on the basis pursuit algorithm have been considered, e.g., employing the gradient projection [FNW07], or the interior-point [KKL$^+$07] techniques. Alternative options encompass greedy versions of basis pursuit, like the matching pursuit variants [TG07b], which improve the rapidness in solving (2.2), but may lead to poor estimation performance. Despite the noise sensitivity of the original ℓ_0-norm minimization, it has been shown that it gives the highest possibility of sparse recovery with very few measurements, which motivates the use of approximations to the ℓ_0-norm function to directly solve (2.1). Among them, the smoothed ℓ_0 cost function [MBZJ09] approximation is a fast algorithm, and gives significantly improved performance in scenarios with low SNR regimes. The main idea behind this approach is to approximate the nonlinear function (2.1) by a suitable continuous function, and take advantage of linearity. The weighted ℓ_1-minimization for nonuniform sampling has been further addressed in [KXAH11]. The construction of deterministic sensing matrices satisfying the statistical restricted isometry property (RIP) has been reported in [CHJ10].

The main drawback of the aforementioned works is that they focus on solving the compressed sensing problem in a very general fashion, without reaching engineering-like solutions. Furthermore, most of the reported algorithms focus on the reconstruction of the information at the signal level. However, as shown in this thesis, the reconstruction of the information at the signal level is not necessary for the purpose of primary signal detection or design of secondary transmission waveforms.

2.2.2 Rate of Innovation

Parallel to the Candès work, M. Vetterli *et al.* [VMB02] presented the concept of *rate of innovation* of signals as a generalization to the concept of bandwidth. The rate of innovation, defined as

$$\rho \doteq \lim_{T\to\infty} \frac{1}{T}\tau_X(T), \tag{2.3}$$

where $\tau_X(T)$ counts the number of degrees of freedom of X in an interval T, expresses the level of innovation per unit of time. The idea behind it is that if a signal has a finite rate of innovation ρ, one hopes to be able to measure and reconstruct the signal by taking only ρ samples per unit of time. There is a clear analogy to compressed sensing, as ρ may measure the sparsity level of X. However, this generalization allows to consider signals which, even though they are not band limited, they can be sampled uniformly at or above the rate of innovation using appropriate kernels.

From a spectral analysis view point, P. Stoica *et al.* provide a nonuniform sampling approach in [SLH09] and show that nonuniform sampling does not suffer from tome drawbacks present in traditional uniform sampling. The connection to compressed sensing is

feasible due the nonuniform sampling matrix $\mathbf{\Psi}$. The multichannel sampling problem at the rate of innovation has been recently addressed by [GTE11].

2.2.3 Multicoset Sampling

A third line of investigation related to compressed sensing has been pursued by G. Eldar *et al.* [ME09, EKB10, ME10, GE11, DE11, BHME12]. The work presented in [ME09, ME10] establishes a framework related to compressed sensing, called multi-coset sampling for multi-band signals with sparse support over the sensed band, that establishes the general conditions for perfect reconstruction. This theory is based on universal sampling patters which are a particularization of the nonuniform sampling matrix $\mathbf{\Psi}$ and obey a predetermined structure so that the uniqueness of the solution is preserved. Under some circumstances, if the universal multi-coset pattern takes p samples out of L, it is shown that

$$\frac{1}{T_S} = \frac{p}{LT} \geq nNB, \tag{2.4}$$

where T_S is the average sampling rate, $1/T$ is the Nyquist rate, and NB is the Landau rate for N bands whose width do not exceed B. If the spectrum support is known, $n = 1$, and the sampling rate must be above the Landau rate. On the other hand, if the spectrum support is known, there is a penalty of $n = 2$ in the minimum required sampling rate. Even the authors provide a reconstruction algorithm based on finite-size observations, numerical algorithms are needed as the reconstruction is based on the classical compressed sensing theory (2.1). The theory presented in [ME09] is applied to wideband analog signals with unknown spectral support in [ME10]. The idea of multi-coset sampling implementation with filter banks has been further reported in [MEDS11, MEE11] for the problem of analog-to-digital conversion (AGC).

Sparsity can be naturally found in many signals, and the reconstruction of a compressed sparse signal relies on the incoherence between the nonuniform sampling and the domain of sparsity. In [GE11], the authors claim that reconstruction is possible even when the sparsity base is not known, i.e., blind compressed sensing of *sparke* signals. Uniqueness is ensured for a constrained family of sparsity basis, whose restrictions and reconstruction algorithms are discussed in the work by [GE11].

Another problem of interest is that of detecting the activity of primary systems employing burst transmissions. Settled in the time-domain, the sparsity does not occur symbol-by-symbol, rather in a block-sparsity fashion. The work in [EKB10] conditions the compressed sensing problem setting for block-sparse signals, and show that block-sparsity can yield to better reconstruction results than treating the signal as being sparse in the conventional sense.

An outstanding review on the compressed sensing theory has recently been reported in [DE11], where the main theme is to exploit signal and measurement structure in com-

pressed sensing. In other words, the main focus of [DE11] is to establish the link between theory and practice, i.e., to bring the maths to the hardware. Further in [BHME12], the performance bounds for estimating the finite rate of innovation of signals are established. This provides a useful metric for designing sampling criteria.

2.2.4 Advances in Nonuniform Sampling

A connection to random coding has been established in the recent work by [RG10]. Random coding [VO79] is an underlying theory developed by R. G. Gallager [Gal68] for the classical and non-classical (i.e., quantum) channel coding theorem, which presumes of a known a priori distribution of the input alphabet to provide an upper bound on the error probability. The authors in [RG10] assume that there is some prior distribution on the location of the zero elements of the sparse signal, e.g., which frequency bands are more likely to be empty, and derive information theoretical bounds on the entropy of the non-zero elements of the sparse signal to show that the sampling density can be further reduced.

Other connections between nonuniform sampling and other signal processing and information theory topics include kernel-induced sampling [TIM10], the sparse RLS algorihm [BKT10], regularized sampling for multiband signals [Sel10], co-prime sampling [VP11], segmented compressed sensing [TV11], ower spectrum blind sampling via periodic sampling and simple least squares reconstruction [LA11], sparse residuals [KG11], semi-parametric methods such as SPICE [SBL11], grid solution problems, [SB12b], or 1-bit compressed sensing [LWYB11]. Finally, a reconciliation of compressed sensing systems with spectrally sparse continuous-time signals is reported in [LDT12], which establishes a good frequency-domain analysis and retaliations to the known problems of windowing and aliasing.

2.2.5 Fundamental Limits

Lastly, a large bunch of research has been generated with the aim of theoretically characterize the process of nonuniform sampling, as well as determine the fundamental limits associated to sampling.

In [MM10a] and [MM10b], the sampling process has been studied from the view point of systems theory, providing the main concepts and tools, as well as noncasual solutions. More information theoretical limits related to Shannon are established in the works by [AT10] and [ASZ10]. For instance, [CEG13] explores two fundamental questions at the intersection of sampling theory and information theory: how channel capacity is affected by sampling below the channel's Nyquist rate, and what sub-Nyquist sampling strategy should be employed to maximize capacity.

An study of compressed sensing as an estimation procedure has been conducted in [JBHHE11], whereas an analysis of the support recovery (only the position of the non-zero

entries) by casting the problem as a MAC has been addressed in [JKR11]. Further Cramér-Rao bounds in noisy compressed sensing [BKT09]. In [OSW12], the authors establish the universal sampling sets for the reconstruction of non uniformly sampled signals by means of interpolation when the signals belong to the set of bandlimited signals $\mathbb{B}$. Also, the nonuniform sampling problem has been analyzed from a rate-distortion theory, such as [WV12].

Finally, it is worth mentioning the work [DBWB10] on statistical signal processing (detection and estimation) with compressed measurements. The idea is naturally based on random projections which behave as near-optimal measurements. However, despite the intense focus of the community on signal recovery, many signal processing problems (e.g., detection, classification, estimation or filtering) do not require full signal recovery.

2.3 Spectrum Sensing

The primary function of a cognitive radio receiver is to reliably identify available spectrum resources temporally unused by primary users in a geographical area. This awareness can be obtained through a database, by using beacons, or by local spectrum sensing at the cognitive radios [GS08a]. This thesis focusses on the spectrum sensing problem performed at the cognitive radio receivers as it constitutes a broader solution and has less infrastructure requirements. In the sequel, it is assumed that the cognitive radio receiver has a dataset of N observations of the wideband signal $x(t)$, $\mathbf{X} \doteq (\mathbf{x}_1, \ldots, \mathbf{x}_N)$, and has to decide between hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$ whether a primary signal is present in the sensed resource or not, respectively.

2.3.1 Classical Spectrum Sensing Methods

The *energy detector* [Poo94] is the most common way of spectrum sensing because of its low computational and implementation complexities. In addition, it is the most generic detector as the cognitive radio receivers do not need to know any further properties on the primary user signals. The primary user signal is detected by comparing the output of the energy detector, $T(\mathbf{X}) = \sum_n \|\mathbf{x}_n\|^2 \geq \tau$ with a threshold τ that depends on the noise floor. The main challenges behind the energy detector are the threshold computation with inaccurate noise variance, and its poor performance in low SNR regimes [Tan05]. However, the simplicity of the detector makes its theoretical analysis very clear, and its performance can be accurately characterized.

A second method of spectrum sensing for cognitive radio are the *waveform-based* detectors. Known patterns are commonly employed in wireless systems for synchronization purposes. In the presence of a known pattern pilot signal in the primary services transmission, the cognitive radio can correlate the detected signal with a known copy of itself [Tan05], i.e., $T(\mathbf{X}) = \sum_n \mathbf{x}_n^H \mathbf{r}_n \geq \tau$, where $\mathbf{r}_n$ is the pilot sequence. It is shown that the

waveform-based detector outperforms the energy detector in both reliability and convergence time. However, including pattern preambles may not be feasible in some primary services, or it might derive to storage problems when sensing a very wide band.

Cyclostationarity feature detection has been also applied in cognitive radio spectrum sensing [SND08, HTR09]. This method exploits the cyclostationarity of the primary signals, as periodicity occurs in the signal or in its statistics like the mean of autocorrelation. Even in low SNR regimes, the cyclostationarity-based detectors can differentiate the primary signal from the noise. This is a result of the fact that the thermal noise is a wide-sense stationary (WSS) with no correlation, while the modulated signals show correlation. The detection is then performed by detecting periodic peaks on the cyclic spectral density function, defined as the Fourier transform of the cyclic autocorrelation $R_x(\tau) \doteq \mathbb{E}[x(t+\tau)x^*(t-\tau)e^{\jmath\omega t}]$. Cyclic frequencies can be assumed to be known or they can be extracted from the actual observations, but inaccurate values on them may cause severe degradation on the performance of the detector. For instance, second-order cyclostationarity is exploited in [AHDVP12]. A comparison of feature-based sensing methods is addressed in [CWZ11].

Match-filtering is known as the optimum method for detecting the primary signal when the transmitted signal is perfectly known [Pro00]. The main advantage of match-filters is that compared to the aforementioned methods, it achieves high probability of detecting $\mathcal{H}_1$ in a very short time. However, matched-filtering requires the cognitive radios to demodulate the received signals. In other words, it requires perfect knowledge on the primary signal deterministic characteristics such as the bandwidth, operating frequency, or modulation type and order. Moreover, its complexity increases as it needs radio-frequency receivers for each signal types. Alternatively, data-independent *filter-banks* detectors have been recently applied to cognitive radio [HRS08] as they provide an integrated tool for wireless communications with less complexity than match-filtering. However, these detectors are not robust to interference. A generalization of filter-banks design has been deployed in [PNLRS09] in the framework of correlation-matching [Gar86]. The correlation-matching scheme presented in [PNLRS09] provides power level estimation and frequency location for a candidate correlation matrix to be suited with the sample covariance matrix. If the proper candidate matrix is selected, the resulting detector becomes robust to interferences and significantly outperforms other feature-based detectors.

Finally, Bayesian spectrum sensing has also gained recent attention, e.g., [JVLN$^+$11, LS12, ZKLZ13]. In [JVLN$^+$11], the location of the secondary users is employed as side information. In [LS12], a Bayesian detector is proposed to minimize the error probability via the Chernoff bound. In [ZKLZ13], the fact that the primary system is highly likely idle is further exploited.

2.3.2 GLRT-based Spectrum Sensing

The detectors discussed in Section 2.3.1 relay on the fact that the cognitive radio receiver has perfect knowledge on some deterministic parameters of the noise and primary signal, such as the noise variance or the signal features, e.g., the cyclic frequency.

This thesis investigates the generalized likelihood ratio test (GLRT) based spectrum sensing detectors for cognitive radio for two reasons. Firstly, the GLRT is the optimal detector in the Neyman-Pearson sense [Kay98a] because it maximizes the probability of detection, $\mathbb{P}(\mathcal{H}_1|\mathcal{H}_1)$, in front of a fixed false alarm probability, $\mathbb{P}(\mathcal{H}_1|\mathcal{H}_0)$. And secondly, GLRT detectors become a solution to inaccurate parameter models because they natively incorporate a joint parameter estimation framework:

$$L(\mathbf{X}, \Theta) \doteq \frac{\pi(\hat{\Xi}_1)p(\mathbf{X}|\hat{\Xi}_1, \Theta_1, \mathcal{H}_1)}{\pi(\hat{\Xi}_0)p(\mathbf{X}|\hat{\Xi}_0, \Theta_0, \mathcal{H}_0)} \geq \gamma \tag{2.5}$$

where $\hat{\Xi}_1$ and $\hat{\Xi}_0$ are the maximum a posteriori (MAP) estimates of the unknown model parameters, i.e.,

$$\hat{\Xi}_0 = \arg\max_{\Xi} \pi(\Xi)p(\mathbf{X}|\Xi, \Theta_0, \mathcal{H}_0), \tag{2.6}$$

and

$$\hat{\Xi}_1 = \arg\max_{\Xi} \pi(\Xi)p(\mathbf{X}|\Xi, \Theta_1, \mathcal{H}_1), \tag{2.7}$$

respectively. The parameters Θ_0 and Θ_1 indicate the side information available in the problem. If the prior distribution of the unknowns, $\pi(\Xi_1)$ and $\pi(\Xi_0)$, is not available, $\hat{\Xi}_1$ and $\hat{\Xi}_0$ become the maximum likelihood (ML) estimates. The optimality of the GLRT in the finite size case has been recently proved for the joint detection and estimation problem by G. Moustakides *et. al.* in [MJTW12]. Furthermore, S. Kay has shown the optimality of the Neyman-Pearson test in statistical signal detection and its relation to signal estimation and the Fisher information matrix [Kay12]. The work done in [FSW10] studies the effect of side information on the noise, channel and signal statistics over the performance of the GLRT. The authors also provide extensions to multiple-input multiple-output (MIMO) channels, fast and slow-fading Rayleigh channels, and orthogonal frequency-division multiple-access (OFDMA) schemes. The work concludes that under the white-noise assumption, estimating the noise variance incurs notable performance gain if the true value is inaccurate. This result suggests that the second-order statistics of the noise and signal are a sufficient statistic for the detection problem. In fact, spectrum sensing detectors exclusively based on statistical covariances to exploit spectral features have been reported in the literature, e.g., [ZL09, QZSS11, YLPC11]. On the other hand, in [GA11] and [GA12] it is shown that implementing the detection procedure in the frequency-domain offers fewer false alarms than its implementation in the time-domain, while offering the possibility of making use of well-known signal processing tools such as spectral analysis. Finally, GLRT-based spectrum sensing algorithms are also developed in [HLC10, BNS11, PBD11a, AL11] for OFDM

cognitive radio networks.

2.3.3 Blind Spectrum Sensing

Numerous spectrum sensing algorithms are blind signal detectors, i.e., are designed to exclusively exploit the characteristics of the primary signal (energy, stationarity, cyclostationarity, pilots, spatial rank, etc). However, the statistics of the communication channel between the primary transmitter and the secondary receiver may also be employed to improve detection. The viability of blind spectrum sensing (i.e., only relying on the eigenvalue decomposition (EVD) distribution of the sample covariance matrix GLRT) is addressed in [ZQG11] with hardware implementations. The correlation between sub-bands occupancy is exploited in to fight against the propagation channel [HC11].

2.3.4 Multiantenna Spectrum Sensing

It is clear that another manner of improving the detection performance is to exploit spatial correlation by the use of multiple antennas at the cognitive radio receivers.

In [LZLZ08], the multiantenna spectrum sensing problem is first formulated for the wideband low SNR regime and for unknown noise covariance matrix. A rank-1 GLRT by blindly learning the leading eigenvector of the observations has been reported by [ZQ13]. The multiantenna spectrum sensing problem with unknown primary signal correlation matrix is addressed in [RVVLV$^+$11]. Spectral a priori information is exploited for the multiantenna spectrum sensing problem by [VVLVS11]. Totally blind spectrum sensing (i.e., unknown primary signaling, uknown noise variance and unknown channel coefficients) using antenna arrays and path correlation is addressed in [OSNP11]. By exploiting the empirical characteristic function from the observations, the multiantenna spectrum sensing problem is addressed by [SWZZ12] when no information is required on the primary signal nor the noise. Space-time domain spectrum sensing is addressed in [SAVVLV12] by considering known power spectral density (PSD) and unknown noise spatial correlation. Finally, the locally most powerful invariant test (LMPIT) has been proposed in [RVSS13] which shows that under close hypotheses (e.g., low SNR), the additional rank structure of the primary covariance matrix is actually irrelevant for optimal detection.

2.3.5 Sensing Time

Another bunch of spectrum sensing literature has focussed on minimizing the sensing time, which is a crucial parameter in cognitive radio as outlined above.

In [PLGZ11], a power control of the secondary users is proposed which allows the maximization of the secondary system throughput and optimize the sensing time at the same time. The sensing time is also optimized in [MMS12] together with power control in secondary multiantenna communications. A similar approach has been taken in [LR12],

where the sensing time is optimized along with the sensing threshold. Sensing time is also discussed in [LLL12], where the joint sensing block length and detection threshold optimization to maximize the throughput of a cognitive radio network is investigated. Furthermore, in [SN11] the sensing time and the power allocation problem are jointly considered in the case of multiband signals.

Finally, the concept of *quickest detection* is presented in [ZP13] on opportunistic detection based on finite-sample-size setting, where the detection rule possible makes an early decision on the $\mathcal{H}_1$ hypothesis, while always deferring to $\mathcal{H}_0$ until collecting all the samples. The asymptotic behavior is analyzed by means of a Chernoff-Stein lemma, which is the ratio between the Kullback-Leibler divergence (KLD) and the Chernoff information of the two hypotheses.

2.3.6 Compressed Spectrum Sensing

Tian *et. al.* [TG07a] study the problem of collaborative distributed spectrum sensing in wideband communications, and provide an extension to spectrum estimation in [Tia08]. The common factor of these works is that the authors employ compressed sensing and spectrum sensing independently. On the one hand, each cognitive radio reconstructs the received wideband signal from the compressed local observations using the basis pursuit method, and on the other hand, an energy detector with majority vote decision is performed in the fusion center of the cooperative cognitive radio network. Similarly, the authors in [CZL09] address the problem of spectrum sensing in wideband cognitive radio scenarios in a two-step approach. Firstly, the basis-pusuit on the Fourier transform of the observations is employed to recover the received spectrum. Secondly, the band location estimation problem is solved using a wavelet-based edge detector. Recent spectrum sensing techniques for wideband cognitive radio based on the detection of the cyclic feature from sub-Nyquist samples have been proposed in [TTS12].

Alternatively, the work done by Leus *et. al.* [WPPL09, PWPL09] on compressed spectrum sensing is focused on the reconstruction of the signal PSD of the primary user signals motivated by the sparsity present on the autocorrelation signal. The authors solve the distributed collaborative sensing problem by considering a fusion center which, after collecting the autocorrelation signals from the cognitive radios, employs the simultaneous orthogonal matching pursuit algorithm to recover the observations PSD. Also, the authors in [BG10] investigate the problem of cooperative spectrum sensing in cognitive radio networks that show sparsity in two domains: frequency and space. Based on local energy detectors, the cooperative scheme aims to determine the locations of the cognitive transmitters as well as the frequency bands unused by the users for opportunistic frequency reuse. Obtaining an interference map in space and frequency is enforced by statistical models for the location of the cognitive transmitters [GTS07]. The compressed wideband power spectral estimation problem has been addressed in by Leus [AL12].

Finally, in a recent work by [PHB12], the sensing time in which the channel is sensed has nonuniform duration, which has some benefits in terms of maximizing the secondary throughput.

2.3.7 Cooperative Spectrum Sensing

Even though this thesis is not focussed on cooperative spectrum sensing, in the following works that relate to the topics addressed in this thesis are listed. The combination of cooperative spectrum sensing and nonuniform sampling has been reported in [MYL+11], where sparse observations are employed in the detection procedure. A good reference in cooperative spectrum sensing when reporting imperfect channel state information (CSI) is the work by [CLKP12] and the references therein, where the comparison between hard and soft decisions is addressed. In the context of multiantenna, the generalized energy detector in cooperative sensing is proposed in [SBM12]. Finally, it is worth noting that in cooperative spectrum sensing a secondary user may misbehave in order to artificially increase or reduce the throughput of the cognitive network by sending wrong sensing information. In [PSDC12], a statistical attack model is adopted by characterizing the malicious nodes with a probability of attack.

2.3.8 Adaptive Sensing

Lastly, adaptive sensing has also been proposed by the literature in order to exploit the structure of the primary signal by adapting the sensing procedure to those *directions* to maximize the detection. In [LSI12], the combined problem of adaptive sensing and resource allocation with imperfect spectrum sensing is addressed. Similarly, adaptive sensing in congested bands is reported in [TCW12]. Moreover, the problem of compressed spectrum sensing from an ℓ_2-norm perspective and adaptive fashion is reported in [SCN12].

Finally, in [ACCD13], the authors show that, despite the folk belief that adaptive techniques always improve the estimation error compared to nonadaptive counterparts, the advantages offered by clever adaptive strategies and sophisticated estimation procedures over classical compressed ℓ_1-norm schemes are minimal.

2.4 Dynamic Spectrum Access

The second primary issue in cognitive radio networks after spectrum sensing is the access strategy that integrates the opportunity exploitation. The main challenge behind DSA for cognitive radio is the established hierarchical structure in which the secondary system is designed such that no or only insignificant interference is generated toward the primary system.

2.4.1 Multiple-Access Techniques

As a wideband wireless system, the cognitive radio multiple-access technique must optimize the utilization of the available spectrum, and guarantee the maximum number of reliable links. Conversely, wideband communications are characterized by arbitrarily large delays, low SNR and close to zero spectral efficiency [Ver02]. The spectrum efficiency in classical point-to-point communication channels, defined as the information rate that can be transmitted over a given bandwidth [BB99], is a suitable metric for benchmarking among different multiple-access techniques. The spectrum efficiency in wireless networks has recently been casted as the transmission capacity [WAJ10]. The transmission capacity is defined as the number of successful transmissions taking place in the network per geographical unit area, subject to a constraint on outage probability. The outage capacity was first presented for both the broadcast channel and the MAC [LG01, LG05]. The importance of this novel concept lies largely in that it can be exactly derived in some important cases, or tightly bounded in many others, hence providing a useful tool of comparison. In general, one distinguishes between the following fundamental multiple-access schemes: TDMA, FDMA, CDMA, space division multiple-access (SDMA), and random access. Spread-spectrum techniques [Pur87] have been adopted in many interference-challenged wireless communication systems because of its efficiency in which several terminals transmit over the same frequency bandwidth, without requiring planned infrastructure (ad hoc). Consequently, frequency-hopping CDMA (FH-CDMA) and direct-sequence CDMA (DS-CDMA) are potential dynamic access strategies for cognitive radios.

2.4.2 Transmission Capacity and Outage Capacity

With no delay constraint, the classical information theoretical MAC capacity is the relevant performance indicator. This applies, for example, to variable-rate systems. On the other hand, most of today's communication systems carry services for which constant-rate and delay-limited transmission should be considered. In this case, the outage probability is the appropriate capacity metric as it is able to provide the probability that any quality of service (QoS) performance metric such as mutual information, SNR, transmitted power, or bit error rate (BER) are bellow or above the resource limit, that is, target transmission rate, target SNR, maximum power or target BER [BPS98, CTB99, OSW94, KS95, CA07].

The transmission capacity framework focuses on the statistics of the received signal-to-interference-plus-noise ratio (SINR) in the MAC. They key underlying this mathematical concept is the use of spatial models for the location of the M terminals on a plane, whose channel response is a function of their relative distances. Spatial models have been used in wireless communications since the late 1970's [KS78, MW78], in which the received SINR determines the conditions of transmission success. Moreover, mathematical formulation for spatial models consisting of stochastic geometry has been shown to be well suited for

a wide range of problems within wireless communications [BB09].

The transmission capacity metric was first introduced by the authors in [WYAdV05]. It is defined as the spatial intensity of attempted transmissions associated to an outage probability of the SINR. The transmission capacity metric has been employed in a wide variety of wireless network design and performance analysis problems, including spread spectrum [WYdVA04], interference-cancellation [WAYdV07], cognitive radio [JAW08, CA09], scheduling and power control [WAJ07, JWA00], and the use of multiple antennas for beamforming or orthogonal space-time block coding (OSTBC) [HAW08]. For a general survey and unification of a number of recent contributions that have collectively developed the transmission capacity metric, see [WAJ10]. One of the most relevant conclusions of the transmission capacity work is that for CDMA multiple-access techniques, the ratio between the transmission capacity of FH-CDMA and DS-CDMA only depends on the spreading factor K and the path-loss parameter n by [AWH07]

$$\frac{c_{FH}}{c_{DS}} = K^{1-2/n}. \tag{2.8}$$

In other words, if $n > 2$, FH-CMDA outperforms DS-CDMA for fixed system conditions by a factor of, e.g., $\sqrt{K}$ in indoor environment with $n = 4$. Recently, the concept of transmission capacity has been extended to spectrum sharing scenarios with the combination of interference alignment [LAH13]. The potentials of interference alignment in cognitive radio networks are reported in Section 8.2.4 of Chapter 8.

Lastly, the work by Menon *et al.* [MBR05], also based on the outage probability and exponential models [GTS07], addresses an assessment on the performance on interference avoidance and interference averaging multiple-access techniques. The authors show that the interference-avoidance techniques dramatically reduce the interference seen at the primary services. Therefore, underlay schemes such as FH-CDMA should also incorporate interference-avoidance. Moreover, interference-avoidance underlay cognitive radios result in lower outage probabilities as compared to interference-avoidance overlay cognitive radios, with pronounced benefits as the transmission bandwidth is increased.

2.4.3 Linear Prediction

Not surprisingly, FH-CDMA has already been adopted in many research lines in the field of cognitive radio DSA by L. Tong *et al.* in [GSTS09]. In this work, the authors investigate spectrum access policies of cognitive radios in a wireless local area network (WLAN) scenario. The main idea is to sense and predict interference patterns so as to adapt the spectrum access accordingly.

In the line of primary user activity prediction, the works [ZTSC07, GTS08, ZGTS08] discuss the cognitive medium access problem via stochastic modeling of the primary service. Recognizing the apparent spectrum sensing constraints, it is feasible to assume that

a secondary user may not be able to perform full spectrum sensing or may not be willing to monitor the spectrum when it has no data to transmit. Contrarily, a secondary user can choose to sense a subset of the possible channels. The work in [ZTSC07] integrates spectrum sensing and spectrum access by adopting a decision-theoretic approach that casts the design of the spectrum access in the framework of partially observable Markov decision process [Alt99]. This formulation leads to optimal policies for spectrum sensing and access, and a systematic tradeoff between performance and complexity. Specifically, the traffic statistics of the primary system follow a discrete-time Markov process with 2^K states, being K the number of channels, each state of the type $(S_1(t), \ldots, S_K(t)) \in \mathbb{B}^K$. The hardware and energy limitation on the spectrum sensing capabilities of the secondary users is translated to a maximum number of K_1 sensing and access K_2 channels. The protocol is designed under this constraints, and the spectrum and access decisions are made to maximize the throughput of the secondary user while limiting the interference to the primary system. The authors in [GTS08] further discuss the spectrum access problem in cognitive radio networks with partial and fully-observable radio resources ($K_1 = K_2 = K$). The opportunistic spectrum access via periodic channel sensing is discussed in [ZGTS08].

2.4.4 Resource Allocation

The secondary transmitter addresses the communication problem of resource allocation based on the sensing information on the primary users. A resource is understood as any dimension in which a secondary signal can be confined for communication. The problem of power allocation in cognitive radio networks has been addressed in the literature in the derivation of the capacity of spectrum sharing channels, including the additive white Gaussian noise channel [Gas07] and fading channels [KLN$^+$09, MA09]. A major amount of work on power allocation for OFDM spectrum sharing channels has been reported in the last years [BHB08, BBKB11, SB12a, CWC11, BHB11, HBHB09, KGLZ10, CT12, EAQS12, KBB12, WGZ13].

The capacity of the secondary user is maximized in [BHB08] when the average power interference introduced to the primary user remains within a tolerable range. The capacity of the secondary user is also maximized in the context of relay networks with joint relay and power allocation [BBKB11, SB12a], joint subcarrier and power allocation [CWC11]. Perfect CSI of the secondary-to-primary channel and perfect knowledge on the spectral occupancy of the primary signal is a general assumption in [BHB08, BBKB11, SB12a, CWC11]. This knowledge, which requires feedback from and to the primary users, is not a practical assumption and, as a consequence, the power interference constraint needs to be redefined. The authors in [BHB08] extend their work to an statistical formulation of the interference constraint in the sense that the secondary user guarantees that the probability of the interference remaining within a tolerable range is lower-bounded [BHB11].

In parallel, a rate loss formulation of the interference has been proposed on the capacity of the secondary user due to the presence of the primary user in probability [HBHB09]. In parallel, a rate loss formulation of the interference has been also proposed but as a constraint on the primary user capacity due to the secondary transmissions [KGLZ10].

Finally, it is worth mentioning a few works that have addressed the difficult problem of establishing the fundamental capacity limits of cognitive radio networks. Among them, [TFL11] derives the network capacity regions by finding the optimal spectrum management policy, in [RTD12] the inner and outer bounds for the Gaussian cognitive interference channel are invetigared, the capacity of overlay cognitive radio with partial cognitition is derived in [CSVH12], the capacity of the interference channel with causal and noncausal feedback at the cognitive transmitted is addressed in [MAA12], the capacity limits of multiuser multiantenna cognitive networks has been established by [LN12], whereas the capacity of the Z interference channel in the overlay cognitive radio paradigm has been addressed in [LMGSS13], where the cognitive transmitter has information on both the secondary and primary messages, and the interference is caused by the primary transmitted to the secondary transmitter.

2.4.5 Imperfect Sensing and Channel State Information

All the aforementioned works establish the fundamental limits of cognitive radio networks, and show that both spectrum sensing and CSI are crucial factors in the proliferation of this technology. For instance, the works [BHB08, BBKB11, SB12a, CWC11, BHB11, HBHB09, KGLZ10] assume perfect spectrum sensing, i.e., perfect knowledge on the presence of the primary users. However, imperfect spectrum sensing has severe consequences on the performance of the secondary link, as reported in [CT12] and [KBB12]. In this sense, joint spectrum sensing and power control for secondary communication has been investigated in the literature, as e.g., in [AS12], [Wan11] and [WT11].

On the other hand, a control feedback channel is also proposed for DSA in cognitive radio networks, e.g., as reported in the works by [EGR11] and [HLD11].

Finally, in order to diminish the dependence on instantaneous CSI of the primary system, several works in the literature propose dynamic access strategies that exhibit robustness in front of CSI. As an example, the work in [LZY11] proposes a dynamic access scheme based on only quality information of the primary signal. Similarly, several degrees of CSI of the secondary-to-primary channel are considered in the problem of multiantenna transceiver optimization [LCHK11] and multiantenna selection [HSTM11]. In the same way, the works [AKSL11] and [BFT11] assume imperfect CSI in the problems of spectrum management and power control, respectively.

2.4.6 Spectral Precoding and Beamforming

A last approach to the DSA in cognitive radio networks involves the concepts of spectral precoding and beamforming applied to the secondary signal.

In the context of OFDM, the interference caused to the primary system by the out-of-band interference is a bottleneck issue. With frequency notching, the approach presented in [ZLS13] improves the bandwidth efficiency and suprresses the overall radiation without sacrificing BER performance of the secondary users. Similarly, the adaptation of the precoded cycling prefix in OFDM-based communications is investigated in [WLL11].

The extension to MIMO-OFDM is considered in [PBD11b] and [SSR13], where orthogonal space-time coding (OSTC) in the linear precoding and beamforming versions is considered, respectively.

Finally, in the more challenging context of single antenna cognitive radio networks where space coding is not applicable (although easily extendable), time-domain waveform adaptation and more sophisticated frequency-domain adaptation are required. On the one hand, in [TLL11], opportunistic waveform adaptation is proposed by taking advantage of the sparse properties of wideband signals. On the other hand, [CKCD13] exploits the Vandermonde subspace for frequency multiplexing by making use of a precoder based on the Vandermonde structure. By doing so, the proposed technique is able to provide secondary communication while keeping the primary system free of interference.

Chapter 3

Spectral Analysis in Nonuniform Sampling

3.1 Introduction

Given the importance of second-order statistics in signal processing, this Chapter focusses on the analysis and estimation of the second-order parameters, namely the correlation, the spectrum and the signal-to-noise ratio (SNR) of the primary signal in Gaussian noise from nonuniform sampling.

3.1.1 Spectral Analysis based on Correlation-Matching

Illustrated in Figure 3.1, spectral analysis is the signal processing problem that consists of recovering the frequency content in a sampled signal, and is a relevant problem that occurs in many applications ranging from detecting economic cycles to radar imaging. The classical approaches to spectral analysis are based on uniformly sampled observations, e.g., the periodogram [Kay98b], not assuming any spectral structure of the signal. Another class of spectral estimators which are based on adaptive filter-bank methods, e.g., the Capon [Cap69], show outstanding properties and are able to provide better spectral resolution that the periodogram in small number of uniformly sampled observations.

Scanning the whole spectrum is a challenging task due to the power consumption involved in the analog-to-digital conversion (AGC). Adopted alternatives include sequen-

$$S + W \longrightarrow \boxed{\text{Nonuniform Sampling}} \longrightarrow X \longrightarrow \boxed{\text{Spectral Analysis}} \longrightarrow \hat{\phi}_s(\omega)$$

Figure 3.1: Spectrum analysis in nonuniform sampling problem.

tial sensing and filter bank techniques. However, the former incurs significant delay in the sensing process, whereas the latter burdens implementation efficiency. To overcome the aforementioned limitations, nonuniform sampling has gained recent attention as a machinery to exploit the sparsity in the spectrum of the primary signal [Lan67]. In the recent years, this concept has derived to parallel works such as compressed sensing [CW08], nonuniform sampling [SLH09], multicoset sampling [ME09, ME10], kernel-induced sampling [TIM10], and co-prime sampling [VP11], among others.

In this thesis, the problem of spectral analysis based on the periodogram and Capon methods from nonuniformly sampled noisy observations of signals with sparse spectrum is addressed. The nonuniform sampling is formulated as a linear projection matrix which may adopt any structure and, hence, covers many of the aforementioned sampling strategies. Preliminary results of spectral analysis over nonuniform samples show that sampling below the Nyquist rate exhibits the phenomenon of noise enhancement/folding. To overcome this problem, a rank-1 matrix-level correlation-matching based on nonuniform sampling is presented. The potentials of correlation-matching applied to spectral analysis have been recently addressed by [PNLRS09]. Furthermore, correlation-matching behaves as an approximation to maximum likelihood at the low SNR regime with asymptotically large observations [Por08]. The derived correlation-matching estimate permits to formulate the nonuniform periodogram and nonuniform Capon estimates as a particular case when the correlation model is the outer product of the linear filter applied for spectral analysis. It is shown that the nonuniform correlation-matching implements a denoising process to fight against the noise enhancement phenomenon inherent in nonuniform sampling.

3.1.2 Noise Enhancement in Bernoulli Nonuniform Sampling

In order to further comprehend the noise enhancement phenomenon observed in the nonuniform correlation-matching problem, a theoretical sound analysis is established by means of a simple Bernoulli sampling. The Bernoulli distribution, initially advocated to solve reconstruction problems in the 1980s [KM82], has been investigated in the literature for sequential sampling [MSW10] and sparse signal reconstructions [DT10]. The advantages of Bernoulli nonuniform sampling are that it satisfies the restricted isometry property (RIP) with overwhelming probability [CT06] on the one hand, and that is exhibits good tradeoff properties between complexity and performance on the other hand.

More specifically, the statistical treatment of sampling shows that nonuniformly sampling below the Nyquist rate produces in average the effects of noise enhancement and power loss on the second-order statistics of the signal (both correlation and spectrum), *even in noise free scenarios*. The main focus is to obtain an equivalence in terms of SNR to model the noise enhancement and power loss effects. The Bernoulli nonuniform sampling is further extended to matrix formulation. This allows the derivation of the noise enhancement and SNR equivalences in matrix forms, which are important parameters involved in

the signal detection problems addressed in the following Chapters.

As related work, in [DLTB12], the authors show that in compressed sensing systems that aim to reduce the number of measurements are sensitive to signal noise, exhibiting a 3 dB SNR loss per octave of subsampling, which parallels the classic noise folding phenomenon.

3.1.3 Chapter Organization

This Chapter is structured as follows. The nonuniform correlation-matching problem and the derivation of the nonuniform periodogram and Capon estimates is addressed in Section 3.2. The noise enhancement and SNR equivalence in Bernoulli nonuniform sampling are established in Section 3.3. Numerical results are provided in Section 3.4 in order to assess the behavior of the noise enhancement effect from both the nonuniform correlation-matching and the Bernoulli nonuniform sampling views, as well as to illustrate the performance of the proposed nonuniform Capon and periodogram spectrum estimates. Finally, Section 3.5 outlines the main conclusions obtained.

3.2 Nonuniform Correlation-Matching

3.2.1 Problem Statement

The nonuniform correlation-matching problem is considered over the second-order statistics of a signal $s(t)$ with sparse spectrum, which may represent, but not limited to, a multiband signal consisting of the superposition of the primary services in a cognitive radio network. The sensed signal is of the form

$$x(t) = s(t) + w(t), \tag{3.1}$$

where $w(t)$ is the double-sided complex zero-mean additive white Gaussian noise with spectral density $N_0/2$. The N-size uniformly sampled signal and noise processes are defined as $\mathbf{s}_m \doteq [s(t_1^m), \ldots, s(t_N^m)]^T$ and $\mathbf{w}_m \doteq [w(t_1^m), \ldots, w(t_N^m)]^T$, respectively, where the intervals satisfy piecewise stacking, i.e., $t_n^m = (mN + n)/f_s$, and f_s is the Nyquist rate. Let $\phi_s(\omega)$ be the spectrum of $s(t)$. In this setting, the sensing is performed over the frequency range $\omega \in [0, B)$. As a result, the samples consists of a data record of size $N \times M$ given by $\mathbf{S} \doteq (\mathbf{s}_1, \ldots, \mathbf{s}_M)$, with $\mathbf{W}$ similarly defined. As the spectral analysis bandwidth increases $B \to \infty$, an impractical large number of samples would be required. Nonetheless, the motivation of nonuniform sampling arises when the frequency contribution of $\phi_s(\omega)$ is sparse. In such a case, the average sampling density may be decreased much bellow the Nyquist rate. In the sequel, it is stated that $s(t)$ has a sparse spectrum if the support of $\phi_s(\omega)$ is small compared to the sensed bandwidth, i.e., if $B_s \ll B$, where B_s is the Lebesgue

measure of the frequency set $\{\omega \in [0, B) \ : \ \phi_s(\omega) > 0\}$. In other words, the spectral analysis problem of the signal $s(t)$ from the a lower dimensional discrete-time sampled signal is investigated. The nonuniformly sampled signal is defined as $\mathbf{x}_m \doteq [x(t_1^m), \ldots, x(t_K^m)]^T$, with $K < N$ and the sampling intervals satisfying nonuniform sampling and piece-wise stacking. The nonuniform sampling process may be cast as a linear projection of $\mathbf{x}_m$ onto a lower dimensional space through $K \times N$ projection matrices $\mathbf{\Psi}_m$, i.e.,

$$\mathbf{x}_m = \mathbf{\Psi}_m(\mathbf{s}_m + \mathbf{w}_m), \tag{3.2}$$

for $1 \leq m \leq M$. In what follows, $\mathbf{\Psi}_m$ is a pinning matrix that randomly selects K samples of $\mathbf{s}_m + \mathbf{w}_m$, and it is given by randomly selecting K rows of the identity matrix $\mathbf{I}_N$. In average, the sampling rates of $\mathbf{x}_m$ and $\mathbf{s}_m + \mathbf{w}_m$ are related through the sampling density defined as

$$\kappa \doteq \frac{K}{N}. \tag{3.3}$$

The occupancy, an important parameter used throughout this thesis, is defined as

$$\kappa_0 \doteq \frac{B_s}{B}. \tag{3.4}$$

As the spectrum information of $s(t)$ is contained in the second-order statistics, a rank-1 correlation-matching approach based on nonuniform sampling below the Nyquist rate is proposed, and shown that it allows to formulate the nonuniform periodogram and Capon estimates as particular cases.

3.2.2 Correlation-Matching Approach

The reconstruction of the second-order statistics of $s(t) + w(t)$ from the nonuniformly sampled observations $\mathbf{X} \doteq (\mathbf{x}_1, \ldots, \mathbf{x}_M)$ is discussed. Let Θ denote the set of parameters that uniquely parametrize the spectrum of the received signal. The correlation-matching approach is based on a second-order fitting between the observations and the correlation model $\mathcal{R}(\Theta)$. In many practical situations, the prior knowledge on the noise and signal statistics enables signal processing algorithms for the estimation problem. Moreover, because nonuniform sampling projects the received observations onto a lower dimensional space, a phenomenon of noise enhancement is observed at the output of the reconstructed signal. With the purpose of diminishing this effect, the correlation model $\mathcal{R}(\Theta)$ is

$$\mathcal{R}[P(\omega), \sigma^2] = P(\omega)\mathbf{R}_0(\omega) + \sigma^2\mathbf{I}_N, \tag{3.5}$$

where

$$\mathbf{R}_0(\omega) = \mathbf{v}(\omega)\mathbf{v}^H(\omega) \tag{3.6}$$

denotes the rank-1 correlation matrix of the spectral template $\mathbf{v}(\omega)$, normalized with the condition $\mathrm{tr}(\mathbf{R}_0) = 1$, and $P(\omega)$ and σ^2 are the signal and noise levels at the sensed frequency, respectively. The spectral template has the role of bandpass filter in spectral analysis. The notation is simplified to P and $\mathbf{R}_0$ as the frequency argument is clear from context. The nominal frequency-continuous SNR is further defined as

$$\mathrm{SNR}(\omega) \doteq \frac{P(\omega)}{\sigma^2}. \tag{3.7}$$

As for spectral analysis it is clear that $\mathrm{rank}(\mathbf{R}_0) = 1$, in what follows a second-order fitting of the nonuniformly sampled observations to the correlation model (3.5) is proposed. That is, consider the correlation model $\mathcal{R}(P, \sigma^2) \doteq P\mathbf{R}_0 + \sigma^2\mathbf{I}_N$. The signal and noise power levels that minimize the nonuniform correlation-matching (3.10) with nonuniformly sampled observations $\mathbf{X}$ are given by the solution to the optimization problem

$$(\hat{P}, \hat{\sigma}^2) = \arg\min_{P,\sigma^2} \mathcal{M}\left[\mathbf{X}, \mathcal{R}(P, \sigma^2)\right], \tag{3.8}$$

where

$$\mathcal{M}\left[\mathbf{X}, \mathcal{R}(P, \sigma^2)\right] = \frac{1}{M}\sum_m \|\mathbf{x}_m\mathbf{x}_m^H - \mathbf{\Psi}_m\mathcal{R}(P, \sigma^2)\mathbf{\Psi}_m^H\|^2. \tag{3.9}$$

The solution to the former optimization problem is derived in Appendix A.1, where it is shown that the correlation-matching estimates of the signal power and noise variance are given by

$$\hat{P} = \frac{\mathrm{tr}\left[\mathbf{B}\left(\mathbf{R}_0 - \frac{1}{K}\mathrm{tr}(\overline{\mathbf{R}}_0)\mathbf{I}_N\right)\right]^+}{\mathrm{tr}(\overline{\mathbf{R}_0^2}) - \frac{1}{K}\mathrm{tr}^2(\overline{\mathbf{R}}_0)}, \tag{3.10}$$

and

$$\hat{\sigma}^2 = \frac{\mathrm{tr}\left[\mathbf{B}\left(\frac{1}{K}\mathrm{tr}(\overline{\mathbf{R}_0^2})\mathbf{I}_N - \frac{1}{K}\mathrm{tr}(\overline{\mathbf{R}}_0)\mathbf{R}_0\right)\right]^+}{\mathrm{tr}(\overline{\mathbf{R}_0^2}) - \frac{1}{K}\mathrm{tr}^2(\overline{\mathbf{R}}_0)}, \tag{3.11}$$

respectively, where the matrices involved read

$$\mathbf{B} \doteq \frac{1}{M}\sum_{m=1}^{M} \mathbf{\Psi}_m^H\mathbf{x}_m\mathbf{x}_m^H\mathbf{\Psi}_m, \tag{3.12}$$

$$\overline{\mathbf{R}}_0 \doteq \frac{1}{M}\sum_{m=1}^{M} \mathbf{\Psi}_m\mathbf{R}_0\mathbf{\Psi}_m^H, \tag{3.13}$$

and

$$\overline{\mathbf{R}_0^2} \doteq \frac{1}{M}\sum_{m=1}^{M} \left(\mathbf{\Psi}_m\mathbf{R}_0\mathbf{\Psi}_m^H\right)^2, \tag{3.14}$$

It is worth noting that $\mathbf{B}$ can be interpreted as the reconstructed sample covariance matrix, whereas $\overline{\mathbf{R}}_0$ is the average nonuniformly sampled version of the normalized signal corre-

lation matrix, and $\overline{\mathbf{R}_0^2}$ is the average second-order nonuniformly sampled version of the normalized signal correlation matrix.

The spectral analysis is focused on the signal of interest, i.e., on $s(t)$. Hence, the expression of the noise level estimate is ignored in the sequel, because it is implicit as a denoising process in (3.10). From (3.10), it is interpreted that the spectrum estimation is the result of the projection of the reconstructed second-order statistics of the observations, i.e., $\mathbf{B}$, onto the diagonal off-loaded correlation matrix of the spectral template, which is given by

$$\mathbf{R}_{\mathrm{D}} = \mathbf{R}_0 - \frac{1}{K}\mathrm{tr}(\overline{\mathbf{R}}_0)\mathbf{I}_N. \tag{3.15}$$

On the one hand, the first part of $\mathbf{R}_{\mathrm{D}}$ is the classical spectrum estimation filtering, because $\mathrm{tr}(\mathbf{B}\mathbf{R}_0) = \mathbf{v}(\omega)\mathbf{B}\mathbf{v}^H(\omega)$. On the other hand, the second part of $\mathbf{R}_{\mathrm{D}}$ is in charge of subtracting the part of the reconstructed second-order statistics of the observations that are considered as white noise according to the signal model (3.2). Hence, the estimate (3.10) performs denoising. The denominator of (3.10) is a normalizing term that depends on the nonuniformly sampled squared norm of the spectral template, i.e., it can be shown that $\mathrm{tr}(\overline{\mathbf{R}_0^2}) = \|\mathbf{v}^2(\omega)\|_\kappa^2$, and $\mathrm{tr}(\overline{\mathbf{R}}_0) = \|\ \mathbf{v}(\omega)\|_\kappa^2$, with

$$\|\mathbf{x}\|_\kappa^2 \doteq \frac{1}{M}\sum_{m=1}^{M}\|\mathbf{\Psi}_m\mathbf{x}\|^2. \tag{3.16}$$

For M sufficiently large, it follows $\|\mathbf{v}^2(\omega)\|_\kappa^2 \to \kappa^2\|\mathbf{v}(\omega)\|^4$ and $\|\mathbf{v}(\omega)\|_\kappa^2 \to \kappa\|\mathbf{v}(\omega)\|^2$. Therefore, $\mathrm{tr}(\overline{\mathbf{R}_0^2}) - \frac{1}{K}\mathrm{tr}^2(\overline{\mathbf{R}}_0)$ is asymptotically equivalent to $\left(1-\frac{1}{K}\right)\kappa^2\|\mathbf{v}(\omega)\|^2$, from which it is deduced that the positivity of the denominator of (3.10) is guaranteed for $K > 1$.

Next, the particularization of (3.10) to the problem of spectral analysis of signals with sparse spectrum based on the periodogram and Capon techniques is discussed. In the sequel, the N-points normalized frequency vector is defined as

$$\mathbf{e}(\omega) \doteq \frac{1}{\sqrt{N}}[1,\ e^{j\omega},\ldots,\ e^{j(N-1)\omega}]^H. \tag{3.17}$$

3.2.3 Nonuniform Periodogram

The nonuniform periodogram has been addressed in [SLH09], and the following can be regarded as an extension for general projection matrices and noisy observations. The periodogram of $\phi_s(\omega)$ based on the nonuniformly sampled observations $\mathbf{X}$ is given by letting $\mathbf{R}_0 \doteq \mathbf{e}(\omega)\mathbf{e}^H(\omega)$ in (3.10), i.e., the *nonuniform periodogram* is given by

$$\boxed{\hat{\phi}_s^{\mathrm{P}}(\omega) = \frac{N^2}{K(K-1)}\left[\mathbf{e}^H(\omega)\mathbf{B}\mathbf{e}(\omega) - \frac{1}{N}\mathrm{tr}(\mathbf{B})\right]^+,} \tag{3.18}$$

where $\mathbf{B}$ is the reconstructed sample covariance matrix given in (3.12) It is noticed that P in $\mathcal{R}(P,\sigma^2)$ reflects the periodogram of $s(t)$ as a result of the denoising process, i.e., $\hat{P} = \hat{\phi}_s^{\mathrm{P}}(\omega)$. Noting that in this particular case $\mathrm{tr}(\overline{\mathbf{R}}_0) = K/N$ and $\mathrm{tr}(\overline{\mathbf{R}_0^2}) = K^2/N^2$, the numerator and denominator of (3.10) particularize to $\mathrm{tr}\left[\mathbf{B}\mathbf{e}(\omega)\mathbf{e}^H(\omega) - \frac{1}{N}\mathbf{I}_N\right]$, and $K^2/N^2 - K/N^2$, respectively.

As it can be appreciated, $\mathbf{B}$ acts as the reconstructed correlation matrix of the observations. Hence, (3.18) is a modified periodogram which subtracts the white part of the observations (denoising) and has an scaling factor that depends on the sampling density.

For sake of comparison, the periodogram when omitting the denoising term in (3.5), i.e., with $\mathcal{R}(P) = P\mathbf{R}_0$, is also included. Following the same procedure, the derivative with respect to the signal level P gives the equation $\mathrm{tr}(\mathbf{B}\mathbf{R}_0) - P\mathrm{tr}(\overline{\mathbf{R}_0^2}) = 0$, which leads to the noisy correlation-matching estimate

$$\tilde{P} = \frac{\mathrm{tr}(\mathbf{B}\mathbf{R}_0)}{\mathrm{tr}(\overline{\mathbf{R}_0^2})}, \tag{3.19}$$

For $\kappa = 1$, (3.19) reduces to the correlation-matching estimate of uniformly sampled data. The noisy periodogram is obtained as a particularization of (3.19) when $\mathbf{R}_0 = \mathbf{e}(\omega)\mathbf{e}^H(\omega)$. Noting that $\mathrm{tr}(\overline{\mathbf{R}_0^2}) = K^2/N^2 = \kappa^2$, it rapidly follows that the *nonuniform noisy periodogram* reads

$$\hat{\phi}_s^{\mathrm{NP}}(\omega) = \frac{1}{\kappa^2}\mathbf{e}^H(\omega)\mathbf{B}\mathbf{e}(\omega). \tag{3.20}$$

For uniform sampling, i.e., $\kappa = 1$, the classical uniform periodogram is obtained [Kay98b].

3.2.4 Nonuniform Capon

The Capon estimate generalizes the periodogram concept by designing the frequency template $\mathbf{v}(\omega)$ under the minimum variance or minimum leakage criterion [Cap69]. Consider (3.10) with $\mathbf{R}_0 \doteq \mathbf{v}(\omega)\mathbf{v}^H(\omega)$, with the additional constraint $\mathbf{v}^H(\omega)\mathbf{e}(\omega) = 1$. Given that the denominator of (3.10) depends on the template $\mathbf{v}(\omega)$, the asymptotic behavior of the Capon template is considered. As $N \to \infty$, the Capon template has infinite degrees of freedom, and $\|\mathbf{v}(\omega)\|^2$ is a constant value for all $\mathbf{X}$. Therefore, the Capon template is designed to minimize the first part of the numerator of (3.10), i.e.,

$$\hat{\mathbf{v}}(\omega) = \arg\min_{\mathbf{v}(\omega)} \mathrm{tr}\left[\mathbf{B}\mathbf{v}(\omega)\mathbf{v}^H(\omega)\right] \quad \text{subject to} \quad \mathbf{v}^H(\omega)\mathbf{e}(\omega) = 1. \tag{3.21}$$

The solution of the former constrained optimization problem is obtained by taking the derivative of the Lagrangian $\mathrm{tr}\left[\mathbf{B}\mathbf{v}(\omega)\mathbf{v}^H(\omega)\right] + \varepsilon\left[1 - \mathbf{v}^H(\omega)\mathbf{e}(\omega)\right]$ and setting it to zero. This leads to

$$\hat{\mathbf{v}}(\omega) = \frac{\mathbf{B}^\dagger\mathbf{e}(\omega)}{\mathbf{e}^H(\omega)\mathbf{B}^\dagger\mathbf{e}(\omega)}, \tag{3.22}$$

where $\mathbf{B}^\dagger$ denotes the Moore-Penrose pseudo-inverse of $\mathbf{B}$. It is further noted that, because $\|\mathbf{v}(\omega)\|^2 = 1$ asymptotically as $N \to \infty$, (3.23) is a normalized Capon estimate. Therefore, when M is sufficiently large, if follows that $\mathrm{tr}(\overline{\mathbf{R}}_0) = \frac{K}{N}\|\mathbf{v}(\omega)\|^2$ and $\mathrm{tr}(\overline{\mathbf{R}_0^2}) = \frac{K^2}{N^2}\|\mathbf{v}(\omega)\|^4$, or, equivalently, $\mathrm{tr}(\overline{\mathbf{R}}_0) = K/N$ and $\mathrm{tr}(\overline{\mathbf{R}_0^2}) = K^2/N^2$. Finally, plugging $\hat{\mathbf{v}}(\omega)$ into (3.10) and after some mathematical manipulations, the *nonuniform Capon* is given by

$$\boxed{\hat{\phi}_s^{\mathrm{C}}(\omega) = \frac{N^2}{K(K-1)}\left[\frac{1}{\mathbf{e}^H(\omega)\mathbf{B}^\dagger\mathbf{e}(\omega)} - \frac{1}{N}\mathrm{tr}(\mathbf{B})\right]^+,} \tag{3.23}$$

where $\mathbf{B}$ is the reconstructed sample covariance matrix (3.12), and $\dagger$ denotes the Moore-Penrose pseudoinverse.

Comparing (3.23) to (3.18), it is observed that the same scaling and denoising process are applied to the spectral estimates. It is worth highlighting that the higher the sampling rate reduction, the sparser is the reconstructed correlation matrix $\mathbf{B}$ and therefore a larger scaling is required to preserve the power level. Similarly to the periodogram, the *noisy Capon* estimate is derived for comparison reasons from the noisy correlation-matching estimate (3.19). Letting $\mathbf{R}_0 = \mathbf{v}(\omega)\mathbf{v}^H(\omega)$, the Capon template is designed to minimize the numerator of (3.19) with the additional constraint $\mathbf{v}^H(\omega)\mathbf{e}(\omega) = 1$. This leads to the Capon filter (3.22). Noting, again, the asymptotic property $\mathrm{tr}(\overline{\mathbf{R}_0^2}) = \kappa^2$, from (3.19) and (3.22) it follows that the *nonuniform noisy Capon* is given by

$$\hat{\phi}_s^{\mathrm{NC}}(\omega) = \frac{1}{\kappa^2}\frac{1}{\mathbf{e}^H(\omega)\mathbf{B}^\dagger\mathbf{e}(\omega)}, \tag{3.24}$$

which for $\kappa = 1$ reduces to the classical uniform Capon [Cap69].

3.3 Bernoulli Nonuniform Sampling

3.3.1 Noise Enhancement in Noiseless Observations

In order to derive a theoretical basis for the noise enhancement phenomenon, an analogy between the Gaussian-Bernoulli distribution and nonuniform sampling is established as follows. Let s be a zero-mean Gaussian process that represents the uniform Nyquist sampled signal. Instead of acquiring all the samples of s the problem of sensing is proposed from a smaller number of samples represented by the Gaussian-Bernoulli process, i.e.,

$$x = \psi \cdot s, \tag{3.25}$$

where ψ has a Bernoulli distribution independent from s with parameter κ, leading to $\mathbb{P}(\psi = 1) = \kappa$ and $\mathbb{P}(\psi = 0) = 1 - \kappa$. Therefore, κ is defined as the sampling density and denotes the amount of samples of s that will be represented in x. Because $\mathbb{E}(\psi) = \kappa$, it rapidly follows that in average, the equivalent sampling rate of x is κ times the Nyquist

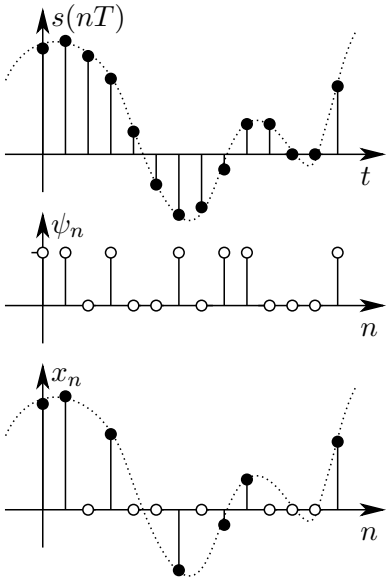

Figure 3.2: Example of Bernoulli nonuniform sampling.

rate. An illustrative example of (3.25) is given in Figure 3.2.

The spectrum recovery problem from nonuniform sampling consists of estimating the autocorrelation or the spectrum of s from the observation process x with complete or partial information on the Bernoulli sampling pattern ψ. In this Section, the problem is addressed in average with statistical knowledge on s and ψ. Given the statistical independence between the signal s and the sampling pattern ψ, it follows that the autocorrelation of x is given by the product of autocorrelations, i.e., $r_x[m] \doteq \mathbb{E}\left[x_n x_{n-m}^*\right] = r_s[m] \cdot r_\psi[m]$, where $r_s[m] \doteq \mathbb{E}\left(s_n s_{n-m}^*\right)$ and $r_\psi[m] \doteq \mathbb{E}\left(\psi_n \psi_{n-m}\right)$. As ψ has Bernoulli distribution, its autocorrelation function is given by the expression $r_\psi[m] = \kappa\delta[m] + \kappa^2 \sum_{l \neq m} \delta[m-l]$. A more convenient way to write $r_\psi[m]$ is

$$r_\psi[m] = \kappa(1-\kappa)\delta[m] + \kappa^2 \sum_{l=-\infty}^{\infty} \delta[m-l]. \tag{3.26}$$

The advantage of (3.26) is the consistency with Nyquist rate uniform sampling when setting the Bernoulli parameter to $\kappa = 1$. In such a case, $r_\psi[m] = \sum_{l=-\infty}^{+\infty} \delta[m-l]$ becomes a uniformly-spaced impulse train and, hence, $r_x[m] = r_s[m]$. In the frequency-domain, the spectrum of x is related to that of ψ and s by the continuous-frequency convolution $\phi_x(\omega) = \frac{1}{2\pi}\phi_s(\omega) * \phi_\psi(\omega)$. After some mathematical operations, it follows that the spectrum of the Gaussian-Bernoulli process x, $\phi_x(\omega)$, is given by

$$\phi_x(\omega) = \kappa(1-\kappa)P + \kappa^2 \sum_{k=-\infty}^{\infty} \phi_s\left(\omega - 2\pi k\right), \tag{3.27}$$

where P is the power of s. From the former equation it is deduced that nonuniform sampling suffers from self-interference (first summand) and attenuated periodic replicas (second summand), *even for noiseless observations*. Further, noting the periodicity of the Fourier transform, without loss of generality, the spectrum of x reads

$$\phi_x(\omega) = \underbrace{\kappa(1-\kappa)P}_{\text{self-interference}} + \underbrace{\kappa^2\phi_s(\omega)}_{\text{signal}}. \tag{3.28}$$

A metric that evaluates the conditions under which the spectrum recovery problem is conceived is the signal-to-interference-plus-noise ratio (SINR). From (3.28), it is deduced that in the noise free scenario,

$$\text{SINR}(\kappa) = \frac{\kappa}{1-\kappa} = \sum_{n=1}^{\infty} \kappa^n. \tag{3.29}$$

Since the sampling density is bounded by $0 \leq \kappa \leq 1$, the SINR (3.29) is bounded by $0 \leq \text{SINR}(\kappa) < \infty$ accordingly. As expected, the maximum SINR is obtained when $\kappa = 1$, whereas the worst scenario follows for $\kappa = 0$, i.e., no samples are taken. It is worth noting that the SINR is not linearly related to κ in general, but only when the sampling density is small the SINR (3.29) approximates to $\text{SINR}(k) \approx \kappa$ as $\kappa \to 0.$, i.e., the conditions vary in proportion to the sampling density.

3.3.2 SNR Equivalence in Noisy Observations

In a realistic scenario, the sampled signal is contaminated by additive Gaussian noise. Because the nonuniform sampling of the noise is also modeled as a Gaussian-Bernoulli process, it follows that the noisy observations are drawn from a Gaussian-Bernoulli process given by

$$x = \psi \cdot (s + w), \tag{3.30}$$

where the noise process is distributed as $w \sim \mathcal{N}(0, \sigma^2)$, being σ^2 its variance. As a consequence, the autocorrelation function of the observations is given by the sum of terms $r_x[m] = r_s[m] \cdot r_\psi[m] + r_w[m] \cdot r_\psi[m]$. Analogous to the developments in Section 3.3.1, from (3.28) it follows that

$$\phi_x(\omega) = \underbrace{\kappa(1-\kappa)P}_{\text{self-interference}} + \underbrace{\kappa\sigma^2}_{\text{noise}} + \underbrace{\kappa^2\phi_s(\omega)}_{\text{signal}}. \tag{3.31}$$

It is appreciated that the noise contribution is a white spectrum of variance σ^2 scaled by the sampling density κ. This is a consequence of $\phi_w(\omega)$ being white. The SNR conditions before nonuniform sampling are defined by the metric $\text{SNR} \doteq P/\sigma^2$. As for the noiseless scenario, in this setting the SINR is a convenient metric to evaluate the conditions of the recovery problem. According to (3.31), the SINR after nonuniform sampling is given, after

some mathematical operations, by

$$\text{SINR}(\kappa) = \frac{\kappa}{1 - \kappa + 1/\text{SNR}}. \tag{3.32}$$

Similarly, the fact that the sampling density satisfies $0 \leq \kappa \leq 1$ allows to establish that the SINR (3.32) for a fixed SNR is bounded above by the SNR itself, i.e., it satisfies $0 \leq \text{SINR} \leq \text{SNR}$. The expression (3.32) allows the following interpretations in the asymptotic cases of low SNR regime ($\text{SNR} \to 0$) and high SNR regime ($\text{SNR} \to \infty$). On the one hand, in the low SNR regime, the SINR behaves linearly with κ and SNR, as (3.32) may be re-written as

$$\text{SINR}(\kappa) \approx \kappa \cdot \text{SNR} \quad \text{as} \quad \text{SNR} \to 0. \tag{3.33}$$

On the other hand, the SINR (3.32) must show consistency with its noiseless counterpart (3.29) in the high SNR regime. Therefore, the SINR for a fixed κ is also bounded above by

$$\text{SINR}(\kappa) \leq \frac{\kappa}{1 - \kappa} \quad \text{as} \quad \text{SNR} \to \infty, \tag{3.34}$$

which states that even in the high SNR regime, the effect of noise enhancement introduced by nonuniform sampling yields the SINR (3.29).

In general, as nonuniform sampling incurs self-interference, relating any variation of κ to an equivalent variation on the SNR becomes a useful metric. More precisely, let $\text{SINR}(\kappa_1, \text{SNR})$ and $\text{SINR}(\kappa_2, \text{SNR})$ be two SINR conditions with sampling densities κ_1 and κ_2, respectively, and $\kappa_1 \geq \kappa_2$, at fixed SNR conditions. By a simple inspection of the definition of (3.32), it follows that the equality $\text{SINR}(\kappa_1, \text{SNR}_e) = \text{SINR}(\kappa_2, \text{SNR})$ occurs when

$$\text{SNR}_e = \frac{\kappa_2 \text{SNR}}{\kappa_1 + (\kappa_1 - \kappa_2)\text{SNR}}. \tag{3.35}$$

From (3.35), it is observed that decreasing the sampling density from κ_1 to κ_2 is equivalent to the original sampling density κ_1 with a lower SNR given by (3.35), as $\text{SNR}_e \leq \text{SNR}$. In the low SNR regime, this equivalence is linearly proportional to the ratio of densities κ_2/κ_1, because

$$\text{SNR}_e \approx \frac{\kappa_2}{\kappa_1}\text{SNR} \quad \text{as} \quad \text{SNR} \to 0. \tag{3.36}$$

For instance, reducing the sampling density by a factor of 2 (i.e., $\kappa_1 = 2\kappa_2$) incurs an equivalent penalty of 3 dB in SNR. On the other hand, in the high SNR regime,

$$\text{SNR}_e \approx \frac{\kappa_2}{\kappa_1 - \kappa_2} \quad \text{as} \quad \text{SNR} \to \infty, \tag{3.37}$$

which behaves independent on the original SNR as expected.

By setting $\kappa_1 = 1$ and $\kappa_2 = \kappa$, the SNR penalty incurred by any sampling density when compared to uniform Nyquist sampling is reduced to $\text{SNR}_e = \text{SINR}(\kappa, \text{SNR})$, i.e., the *SNR*

equivalence is given by

$$\boxed{\text{SNR}_e = \frac{\kappa}{1-\kappa+1/\text{SNR}}.} \tag{3.38}$$

In other words, the SINR (3.32) is the equivalent SNR decrease incurred by Bernoulli nonuniform sampling with respect to uniform sampling. Finally, as the performance of signal processing algorithms is mostly evaluated across SNR, due to the SNR equivalence (3.35), the effect of nonuniform sampling is translated to an SNR-dependent shift of the performance curves. For that purpose, the SNR gap metric is introduced as the ratio between uniform sampling SNR and nonuniform sampling SNR as

$$\Gamma(\kappa) \doteq \frac{\text{SNR}}{\text{SINR}(\kappa)} = \frac{(1-\kappa)\text{SNR}+1}{\kappa}. \tag{3.39}$$

From (3.39), the following SNR asymptotic interpretations apply. In the low SNR regime, the SNR gap becomes independent on the SNR and saturates at

$$\Gamma(\kappa) \approx \frac{1}{\kappa} \quad \text{as} \quad \text{SNR} \to 0. \tag{3.40}$$

In the high SNR regime, the SNR gap introduced by nonuniform sampling is directly proportional to SNR because

$$\Gamma(\kappa) \approx \frac{1-\kappa}{\kappa}\text{SNR} \quad \text{as} \quad \text{SNR} \to \infty. \tag{3.41}$$

This means that in logarithmic scale, all the SNR gaps behave with slope equal to 1 and with a shift with respect to the uniform sampling of

$$\text{SNR}_{\text{shift}} = 10\log_{10}[(1-\kappa)/\kappa] \text{ dB}. \tag{3.42}$$

3.3.3 Finite-Size Vector Signal Processing

In a practical signal processing problem, such as estimation or detection, a finite-size treatment of the observations is required. For that purpose, in the sequel, the Gaussian-Bernoulli process formulation is presented in matrix notation. Let x_n be the n-th observation of the Gaussian-Bernoulli process. The following discrete-time vectors, $\mathbf{s} \doteq (s_0, \ldots, s_{N-1})^T$ and $\boldsymbol{\psi} \doteq (\psi_0, \ldots, \psi_{N-1})^T$, are defined. As a consequence, the Gaussian-Bernoulli process $\mathbf{x} \doteq (x_0, \ldots, x_{N-1})^T$ is related to $\mathbf{s}$ through

$$\mathbf{x} = \boldsymbol{\Psi}\mathbf{s}, \tag{3.43}$$

where $\boldsymbol{\Psi} \doteq \text{diag}(\boldsymbol{\psi})$. Clearly, (3.43) is of the from of a nonuniform sampling problem, where the sampling matrix $\boldsymbol{\Psi}$ is diagonal. The matrix notation is useful because temporal correlation rises in the form of cross-correlation between samples, as the Gaussian part of

the Gaussian-Bernoulli process, i.e., $\mathbf{s}$, is distributed according to $\mathcal{N}(\mathbf{0}, \mathbf{R}_s)$, where $\mathbf{R}_s$ is the correlation matrix of $\mathbf{s}$. It is a well-known statistical signal processing result that any correlation matrix has, asymptotically as $N \to \infty$, the following eigenvalue decomposition (EVD) [Kay98a]

$$\mathbf{R}_s \simeq \mathbf{F}\mathbf{\Lambda}_s\mathbf{F}^H, \tag{3.44}$$

where the equivalence is understood in the weak norm sense. In (3.44), the diagonal matrix $\mathbf{\Lambda}_s \doteq \mathrm{diag}[\phi_s(\omega_0), \ldots, \phi_s(\omega_{N-1})]$ contains the power spectral density (PSD) of s at the $\omega_n = 2\pi n/N$ frequencies, and $\mathbf{F} = (\mathbf{f}_0, \ldots, \mathbf{f}_{N-1})$ is the unitary Fourier matrix that contains as columns $(\mathbf{f}_n)_m = \frac{1}{\sqrt{2}} e^{\jmath m\omega_n}$ with $0 \leq m, n \leq N-1$. From (3.43), the correlation matrix of the Gaussian-Bernoulli observations $\mathbf{x}$ depends on the realizations of the sampling patters $\mathbf{\Psi}$. As in Section 3.3.1, the problem is addressed in average with statistical knowledge on $\mathbf{s}$ and $\mathbf{\Psi}$. Therefore, the correlation matrix of the Gaussian-Bernoulli process must obey the following decomposition according to (3.44)

$$\mathbf{R}_x \simeq \mathbb{E}_{\mathbf{\Psi}}\left(\mathbf{\Psi}\mathbf{F}\mathbf{\Lambda}_s\mathbf{F}^H\mathbf{\Psi}\right) \simeq \mathbf{F}\mathbf{\Lambda}_x\mathbf{F}^H. \tag{3.45}$$

The expectation on $\mathbf{\Psi}$ over $\mathbf{\Psi}\mathbf{F}\mathbf{\Lambda}_s\mathbf{F}^H\mathbf{\Psi}$ requires an element-wise evaluation. Noting that the columns of the Fourier matrix $\mathbf{F}$ are mutually orthonormal, i.e., $\mathbf{f}_i^H\mathbf{f}_j = \delta_{ij}$, it is convenient to express the EVD of $\mathbf{R}_s$ as $\mathbf{R}_s \simeq \sum_{n=0}^{N-1} \lambda_{sn}\mathbf{f}_n\mathbf{f}_n^H$, where $\lambda_{sn} \doteq [\mathbf{\Lambda}_s]_{nn}$. Similarly, $\mathbf{R}_x \simeq \sum_{n=0}^{N-1} \lambda_{xn}\mathbf{f}_n\mathbf{f}_n^H$, where the eigenvalues λ_{xn} are defined analogously to λ_{sn}. From (3.45), the (i,j)-th element of $\mathbf{R}_x$ is therefore given by $(\mathbf{R}_x)_{ij} = \mathbb{E}_{\mathbf{\Psi}}\left(\sum_{n=0}^{N-1} \lambda_{sn} f_{in} f_{jn}^* \Psi_i \Psi_j\right)$. Similarly, $(\mathbf{R}_s)_{ij} = \sum_{n=0}^{N-1} \lambda_{sn} f_{in} f_{jn}^*$. From this point, as the elements of the sampling pattern are drawn from a Bernoulli distribution with parameter κ, it follows that the elements of the correlation matrix of the Gaussian-Bernoulli process $\mathbf{x}$ and the elements of the correlation matrix of the signal are related through

$$(\mathbf{R}_x)_{ij} = (\mathbf{R}_s)_{ij} \cdot \begin{cases} \kappa & i = j \\ \kappa^2 & i \neq j \end{cases}. \tag{3.46}$$

As a stationary process, the diagonal elements are all equal to the arithmetic mean of the eigenvalues, or, equivalently, it is true that $(\mathbf{R}_s)_{ii} = \frac{1}{N}\mathrm{tr}(\mathbf{\Lambda}_s) = P$. As a consequence, matrices $\mathbf{R}_x$ and $\mathbf{R}_s$ can be related through the following equation

$$\mathbf{R}_x \simeq \kappa^2\mathbf{R}_s + \kappa(1-\kappa)P\mathbf{I}, \tag{3.47}$$

where $\mathbf{I}$ is the identity matrix. Finally, noting that $\mathbf{R}_s \simeq \mathbf{F}\mathbf{\Lambda}_s\mathbf{F}^H$, $\mathrm{tr}(\mathbf{R}_s) = \mathrm{tr}(\mathbf{\Lambda}_s)$ and $\mathbf{I} = \mathbf{F}\mathbf{F}^H$, the Fourier matrices $\mathbf{F}$ and $\mathbf{F}^H$ are taken as common left and right factors, respectively, in (3.47). Equating to $\mathbf{F}\mathbf{\Lambda}_x\mathbf{F}^H$, it follows that, necessarily, the eigenvalues of

$\mathbf{R}_x$ and $\mathbf{R}_s$ are related, asymptotically as $N \to \infty$, by

$$\mathbf{\Lambda}_x = \kappa^2 \mathbf{\Lambda}_s + \kappa(1-\kappa)P\mathbf{I}. \tag{3.48}$$

Equations (3.47) and (3.48) are nothing but the extension of (3.28) to matrix formulation. Both (3.47) and (3.48) share the property that the part of the useful signal has been attenuated by a factor of the squared value of the sampling density, and a noise contribution has been introduced in proportion to the variance of the sampled signal. This noise enhancement has the following interpretations. From (3.47), it is seen that the off-diagonal entries of original correlation matrix $\mathbf{R}_s$ have been diminished by a factor of κ^2, whereas the elements in the diagonal only suffer from a factor of κ, as $[\mathbf{R}_x]_{nn} = \kappa^2[\mathbf{R}_s]_{nn} + \kappa(1-\kappa)P = \kappa P$. Therefore, the effect of nonuniform sampling is translated to an attenuation of the nonzero lags correlation coefficients. The interpretations on (3.48) are analogous to (3.28) as the eigenvalues of the correlation matrices contain the PSD of the processes.

The matrix notation introduced in this Section and the noise enhancement effect on the correlation matrix (3.47) are used in the sequel with the aim of establishing an SNR equivalence formulation in signal detection. In particular, throughout the following Chapters, it is shown that the SNR matrix $\mathbf{SNR} \doteq \mathbf{R}_w^{-1}\mathbf{R}_s$ is the main statistical information involved in second-order signal detection. Therefore, the characterization of the SNR equivalence in matrix form is of interest. By considering the estimator-correlator [Kay98a] as reference, it follows that the detection of s in (3.43) with additive white noise with variance σ^2 involves the test

$$T_\kappa(\mathbf{x}) = \mathbf{x}^H \mathbf{SNR}_\kappa (\mathbf{I} + \mathbf{SNR}_\kappa)^{-1}\mathbf{x} \geq \tau, \tag{3.49}$$

where τ is the detection threshold, and $\mathbf{SNR}_\kappa$ is the SNR matrix, which, under the AWGN model is given as

$$\mathbf{SNR}_\kappa \doteq \frac{\mathbf{R}_x}{\sigma^2} = \frac{\kappa\mathbf{R}_s + (1-\kappa)P\mathbf{I}}{\sigma^2}, \tag{3.50}$$

where the equivalence (3.47) has been employed. A simple comparison to the estimator-correlator in uniform sampling (i.e., $T_1(\mathbf{x}) = \mathbf{x}^H \mathbf{SNR}(\mathbf{I} + \mathbf{SNR})^{-1}\mathbf{x}$) where $\mathbf{SNR} = \frac{1}{\sigma^2}\mathbf{R}_s$ derives to the following *SNR matrix equivalence*:

$$\boxed{\mathbf{SNR}_\kappa = \kappa\mathbf{SNR} + (1-\kappa)\mathrm{SNR}\mathbf{I}.} \tag{3.51}$$

Interestingly, the SNR equivalence in signal detection for white noise is analogous to the noise enhancement produced in (3.47). In other words, it is seen that the SNR matrix employed in the estimator-correlator in nonuniform sampling is equal to the SNR matrix in the estimator-correlator in uniform sampling with an equivalent SNR matrix given by (3.51), i.e., a linear combination of the true second-order conditions **SNR** and a contribution proportional to the average SNR with white statistics. When the sampling density is arbitrarily small, $\mathbf{SNR}_\kappa \approx \mathrm{SNR}\mathbf{I}$ and the estimator-correlator reduces (neglecting data

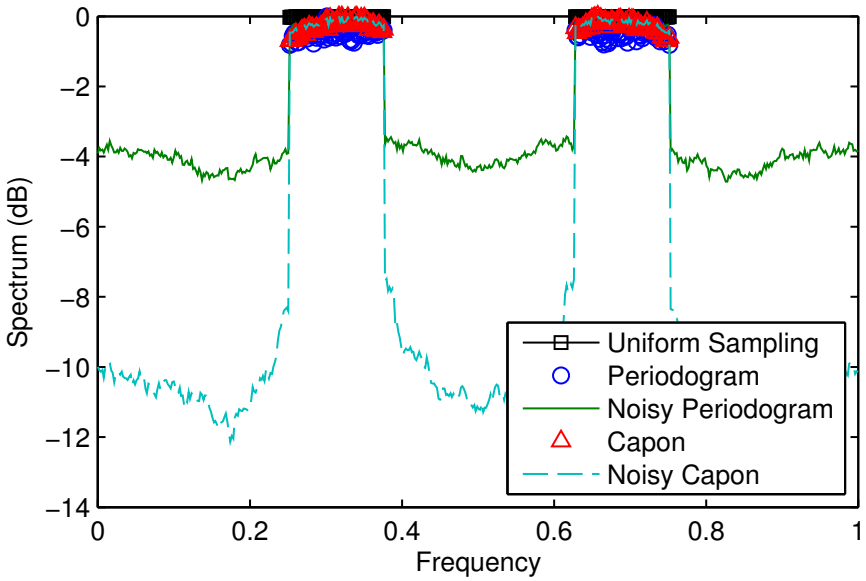

Figure 3.3: Spectrum of a primary signal with occupancy of 1/8 at average SNR of 15 dB and sampling density of 1/4.

independent terms) to $T_\kappa(\mathbf{x}) \approx \mathbf{x}^H\mathbf{x}$, i.e., the energy detector as the time correlation in $\mathbf{R}_s$ is not further exploited. Finally, it is noted that $\mathrm{tr}(\mathbf{SNR}_\kappa) = \mathrm{tr}(\mathbf{SNR})$ for all κ. This means that the comparison developed in this Section is fair in terms of energy (e.g., the total amount of samples is fixed to N) so that the penalties incurred by nonuniform sampling arise from the nature of the sampling and not by the fact of having a smaller density of samples.

3.4 Numerical Results

In this Section, numerical results are provided to show the existence of the noise enhancement phenomenon in nonuniform sampling, to illustrate the performance of the proposed nonuniform periodogram and Capon spectrum estimates, as well as to assess the SNR equivalences established in the Bernoulli nonuniform sampling framework. In the sequel, $s(t)$ is modeled as the equivalent baseband primary signal based on the terrestrial digital video broadcasting (DVB-T) standard [ETS04], and $w(t)$ is modeled as a zero-mean Gaussian noise.

3.4.1 Nonuniform Periodogram and Nonuniform Capon

The performance of the spectral analysis methods presented in Section 3.2, i.e., (3.18), (3.20), (3.23) and (3.24), is presented by means of simulation results. The size of the uniformly sampled observations is $N = 64$. In order to strictly focus on the performance behavior due to sampling rate reduction and remove the effect of insufficient data records,

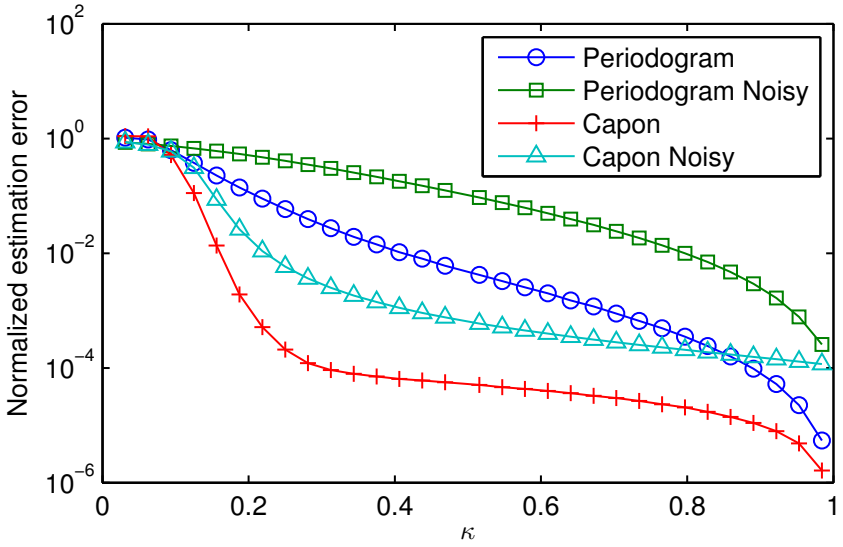

Figure 3.4: Normalized estimation error of the periodogram and Capon estimates versus the sampling density of a primary signal with occupancy of 1/8 at average SNR of 15 dB.

the size of the observations $\mathbf{X}$ is fixed to $2N$ by setting $M(\kappa) = 2N\kappa^{-1}$. The normalized estimation error is defined as

$$\epsilon^2[\phi_s(\omega), \hat{\phi}(\omega)] \doteq \int_B \mathbb{E}\left[\frac{|\phi_s(\omega) - \hat{\phi}(\omega)|^2}{|\phi_s(\omega)|^2}\right] d\omega, \tag{3.52}$$

where $\hat{\phi}(\omega)$ is the spectrum estimate, and $\phi_s(\omega)$ is the spectrum of the uniformly sampled $s(t)$ at Nyquist rate.

An example of spectral analysis with occupancy $\kappa_0 = 1/8$ (e.g., two active channels out of eight in this example), an SNR of 15 dB, and sampling density $\kappa = 1/4$ is depicted in Figure 3.3. The resulting spectrum estimates are averaged according to 1,000 Monte Carlo runs. Firstly, it is observed that the noisy periodogram and noisy Capon, even in the high SNR regime, introduce noise enhancement because of the projection implicit in nonuniform sampling. The denoising process in (3.18) reduces the noise level by approximately 60 dB, whereas the Capon estimate is able to completely subtract the noise floor.

The behavior of the normalized estimation error (3.52) versus the sampling density is illustrated in Figure 3.4, when the occupancy is $\kappa_0 = 1/8$. As it can be appreciated, both Capon and noisy Capon estimates provide better performance when compared to the periodogram counterparts for a wide range of κ. The gain introduced by the denoising process is outlined when comparing the nonuniform spectral estimates to their noisy counterparts. In the case of the Capon estimate, a change of performance behavior is observed around a sampling density of $\kappa = 1/4$.

Figure 3.5 plots the normalized estimation error versus average SNR. In the low SNR regimes, the noisy periodogram and noisy Capon are not able to provide good spectral res-

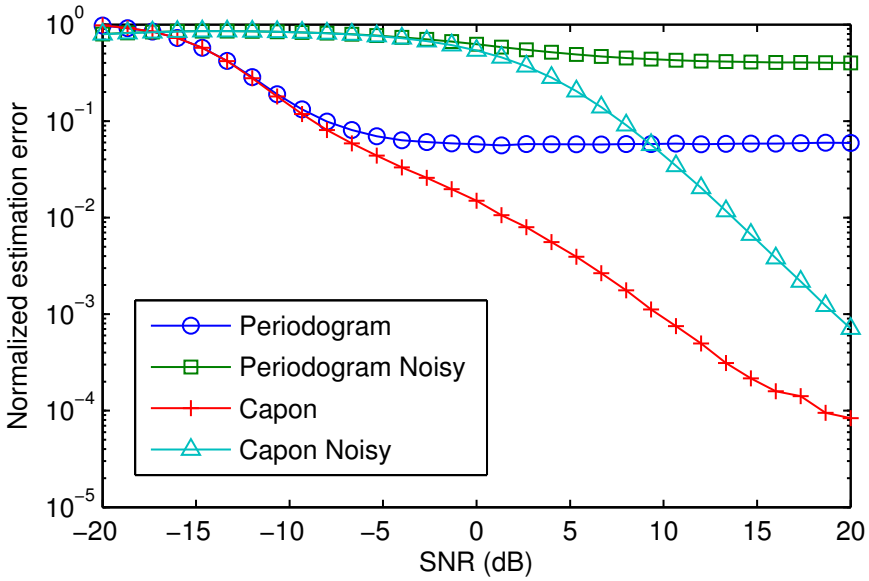

Figure 3.5: Normalized estimation error of the periodogram and Capon estimates versus average SNR of a primary signal with occupancy of 1/8 and sampling density of 1/4.

olution because of the noise floor; whereas the nonuniform periodogram and Capon are capable of distinguishing the primary signal. In the high SNR regimes, the two versions of the periodogram saturate because the spectral pattern is not adaptive. Hence, it is identified that the nonuniform Capon estimate (3.23) might provide outstanding performance for detection purposes for a wide range of SNRs.

Finally, Figure 3.6 illustrates the performance of the spectrum estimates when, for a fixed sampling density $\kappa = 1/4$, the occupancy of the primary signal κ_0 increases from $\kappa_0 = 0$ (no occupancy) to $\kappa_0 = 1$ (100% occupied). It is observed that the normalized estimation error increases with the occupancy, as the dimensionality of the lower space, i.e., K, is not sufficient for high resolution spectral analysis when the sparsity of the spectrum is low. However, it is concluded that the Capon estimate (3.23) behaves as a high resolution spectral method, e.g., $\epsilon^2 \leq 0.01$, as far as $\kappa < \kappa_0$.

3.4.2 SNR Equivalence

The periodogram of two DVB-T signals immersed in Gaussian noise is depicted in Figure 3.7, for $N = 512$. The solid black curve depicts $\phi_x(\omega)$ for an SNR of 10 dB. After Bernoulli nonuniform sampling with a sampling density of $\kappa = 1/4$, the resulting spectrum $\phi_x(\omega)$, which is represented by the dashed red line, suffers from noise enhancement and signal attenuation, according to the expression obtained in (3.31). By employing the SNR equivalence established in (3.32), it is confirmed that $\phi_x(\omega)$ with an SNR of 10 dB and $\phi_x(\omega)$ with an equivalent SNR of

$$\text{SNR}_e = \text{SINR}(1/4) = -1 \text{ dB} \tag{3.53}$$

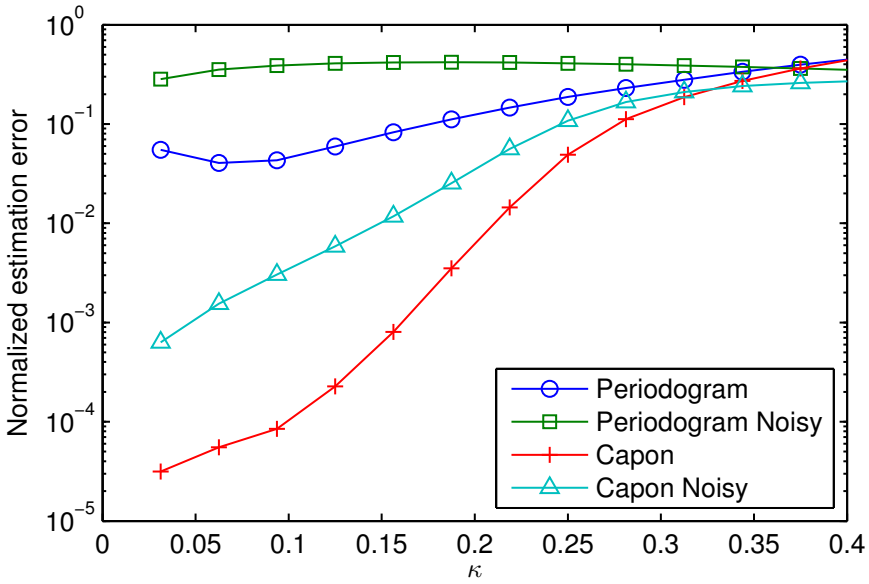

Figure 3.6: Normalized estimation error of the periodogram and Capon estimates versus primary signal occupancy level at average SNR of 15 dB and sampling density of 1/4.

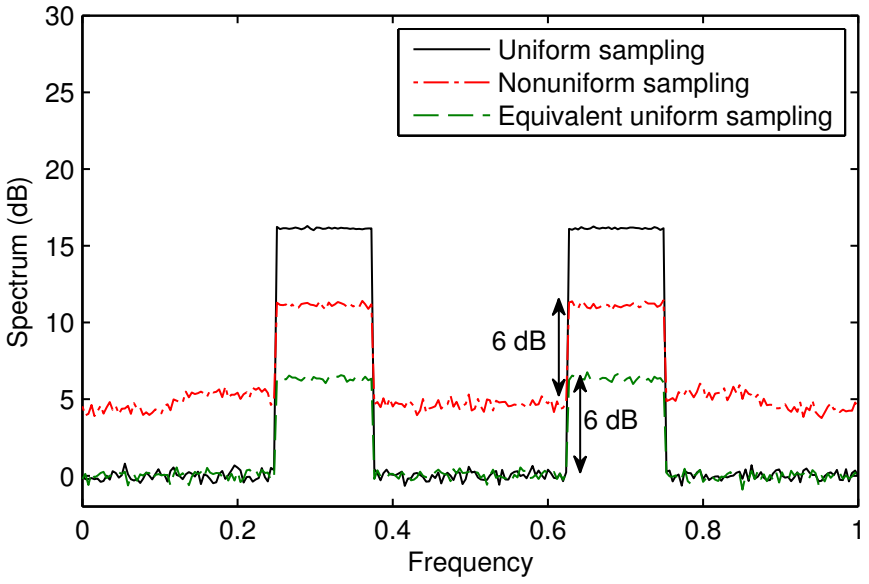

Figure 3.7: Spectrum of uniform sampling, nonuniform sampling with sampling density of 1/4, and equivalent uniform sampling, at average SNR of 10 dB.

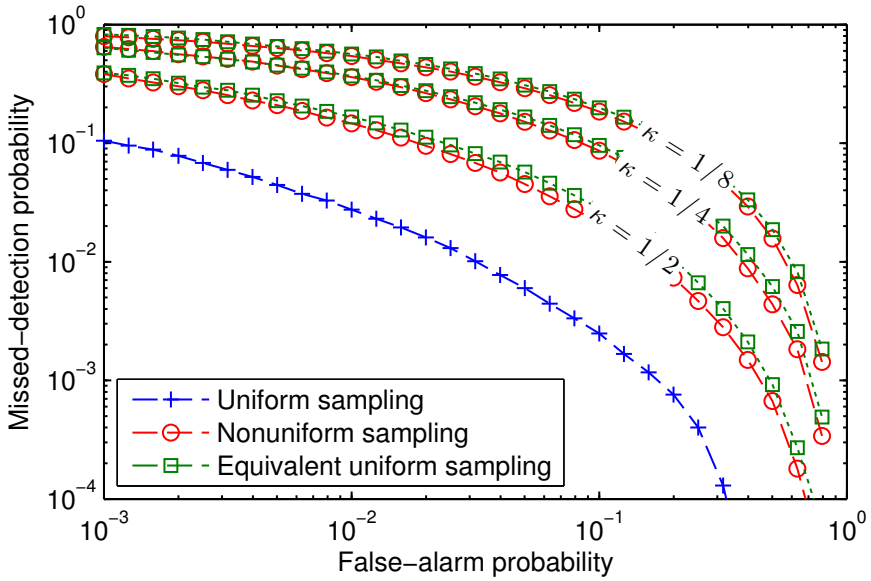

Figure 3.8: ROC of uniform sampling, nonuniform sampling, and equivalent uniform sampling at SNR of -12.5 dB.

show the same spectral SNR. Therefore, the penalty of nonuniform sampling with $\kappa = 1/4$ equals to an SNR penalty of 11 dB.

Finally, the receiver operating characteristics (ROC) of the estimator-correlator (3.49) shown in Figure 3.8 is evaluated. The performance curves are obtained in several sampling conditions for an average SNR of -12.5 dB when detecting a DVB signal immersed in Gaussian noise with $N = 32$ samples and an occupancy of $\kappa_0 = 1/8$. The degradation incurred by nonuniform sampling is observed (dashed lines). The dotted lines depict the performance of the equivalent uniform estimator-detector associated to the given sampling densities, i.e., $T_1(\mathbf{x})$ is evaluated under the SNR matrix conditions (3.51). From Figure 3.8 it follows that the equivalence between SNR and noise enhancement introduced in this thesis accurately models the effect of nonuniform sampling. As a result, the performance of any spectrum sensing detector based on nonuniformly sampled data can be forecasted beforehand by applying an SNR shift according to (3.51).

3.5 Conclusions

In this Chapter, the existence of the noise enhancement phenomenon has been shown by addressing the spectral analysis problem under both empirical and theoretical frameworks.

On the one hand, the problem of spectral analysis employing the periodogram and Capon estimates from nonuniformly sampled data has been addressed. The nonuniform sampling has been modeled as a linear projection onto a lower dimensional space. The correlation-matching metric has been proposed in this setting as a rank-1 fitting. This for-

mulation has allowed the derivation of the nonuniform periodogram and Capon estimates as particular cases. It is concluded that the nonuniform Capon, which performs denoising and adaptive pattern, provides good resolution for spectral analysis.

On the other hand, an equivalence between the noise enhancement produced by Bernoulli nonuniform sampling and SNR has been introduced. The equivalence is firstly established in terms of signal correlation and signal spectral density, and is further extended to matrix formulation and spectrum sensing detection. Numerical results on spectral analysis and spectrum sensing show that the degradation incurred by nonuniform sampling can be easily translated to a penalty in SNR.

Chapter 4

Multi-Frequency Primary Signal Detection

4.1 Introduction

Statistical signal detection is one of the most fundamental signal processing problems for decision making in radar, communications, biomedicine, and so on [Kay98a]. In this Chapter, the problem of detecting the presence of a primary signal over a set of frequency resources is addressed. As illustrated in Figure 4.1, the objective is the design of test statistics $T_q(X)$ that, based on the nonuniformly sampled observations X, provide optimal detection probability at each of the q-th frequency resource.

4.1.1 Generalized Likelihood Ratio Test

This thesis addresses the optimal spectrum sensing detection based on the complete or partial side information on the signal and noise statistics. The primary function of interweave cognitive radios [GJMS09] is to reliably identify the available spectrum resources temporally unused by primary users. This awareness can be obtained through a database, using beacons, or by local spectrum sensing [ALLP12]. This thesis focuses on spectrum sensing performed at the cognitive radio receivers as it constitutes a broader solution and has less infrastructure requirements. The energy detector, cyclostationarity feature detection, and match-filtering are the most commonly employed techniques for spectrum sensing. How-

$$\sum_{q=1}^{Q} S_q + W \longrightarrow \boxed{\text{Nonuniform Sampling}} \longrightarrow X \longrightarrow \boxed{\text{Signal Detection}} \longrightarrow T_q(X)$$

Figure 4.1: Multi-Frequency primary signal detection problem.

ever, the performance of such detectors is severely degraded when the side information on the signal and noise features is incomplete, e.g., the cyclic frequencies. Spectrum sensing detectors based on the generalized likelihood ratio test (GLRT) have received recent attention as the GLRT statistic is optimal in the Neyman-Pearson sense [Kay12] and natively incorporates joint parameter maximum likelihood (ML) estimation for inaccurate model parameters [MJTW12]. The effect of side information on the signal and noise statistics has been reported in [FSW10] in realistic scenarios.

Wideband spectrum sensing has gained recent attention [SNWC13]. It is recognized that ML estimation in wideband cognitive radios is especially challenging because wideband regimes are characterized by close to zero spectral efficiency and low signal-to-noise ratio (low SNR) [Ver02]. Furthermore, the method of ML, despite its theoretical appeal, is often difficult to implement, and analytical solutions are not available in many circumstances [Por08]. However, the tractability of the ML formulation is shown in asymptotically low SNR regimes, and it is identified that the second-order statistics of the observations are sufficient statistics for the spectrum detection problem when the noise variance is high. In this thesis, the optimal ML estimates for GLRT spectrum sensing detection are derived in the single-frequency scenario, and show through low SNR approximations that any GLRT spectrum sensing detector exclusively depends on a kernel operator and the sample covariance matrix of the observations, asymptotically as the SNR tends to zero. Further, it is extended to cognitive radio networks operating over primary systems that employ multi-frequency communications, such as the terrestrial digital video broadcasting (DVB-T). Simulation results assess the performance comparison of the derived algorithms.

A numerous amount of spectrum sensing detectors based on the GLRT have been reported in the last years. It is a well-known signal processing result that in the low SNR regime, the second-order statistics of the signal and noise involved in the detection are sufficient statistics for detection [Kay98a]. More precisely, the cross-correlation between signal and noise can be exploited in the frequency, spatial, or temporal domains. Firstly, frequency correlation detectors are intended to uncover the spectral coherence of cyclostationary processes (see [GNP06] and the references thereby). However, these detectors are very sensitive to the signal and noise features, and require in general high computational complexity. A spectral feature based detector has been reported in [QZSS11]. Secondly, a bunch of signal detectors that exploit spatial correlation have been proposed in the recent literature [ZLLZ10, WFHL10, TNKG10, RVS10, VVLVS11, RVVLV$^+$11, SWZZ12, SAVVLV12, STM13, RVSS13, ZQ13, ZKLZ13]. Addressing the GLRT problem in the multiple antenna framework has allowed the formulation of well-known Gaussian detectors such as the arithmetic-to-geometric mean (AGM) detector [ZLLZ10], rank-1 detectors [RVS10, SAVVLV12, ZQ13], the rank-P detector [RVVLV$^+$11], and the locally most powerful invariant test (LMPIT)[RVSS13]. In the case of fading, the algorithms reported in [WFHL10,

TNKG10, SWZZ12] address blind detection with unknown channel parameters, whereas the works by [VVLVS11, STM13, ZKLZ13] investigate the use of prior information in the detection process. The common frawmework of [ZLLZ10, WFHL10, TNKG10, RVS10, VVLVS11, RVVLV$^+$11, SWZZ12, SAVVLV12, STM13, RVSS13, ZQ13, ZKLZ13] is that the detectors are based on the eigenvalue decomposition (EVD) of the spatial correlation matrix. Therefore, these algorithms are only valid when the secondary users are equipped with multiple antennas. Thirdly, the exploitation of temporal correlation has gained little attention, even though temporal correlation matrices exhibit good properties of stability and computational complexity. Related work involving GLRT detection include [NPI10, YLPC11, BFPRI12]. In [NPI10], the sensitivity of oversampled temporal correlation based detection to frequency offsets is investigated. In order to further improve the efficiency, [YLPC11] reports a detector based on the Cholesky factorization of the sample correlation matrix, whereas the work by [BFPRI12] is concerned in performing signal detection while communicating at the same time. These works does not consider side information on the primary signal correlation.

In this thesis, the temporal correlation of the primary users is exploited given the Toeplitz structure of the correlation matrices when one single antenna is employed at the cognitive receiver. An $N \times N$ Hermitian Toeplitz matrix has only N degrees of freedom. Therefore, the algorithms derived in this Chapter will have good properties of stability and computational complexity. It is noted that the detection of a stationary process with a single antenna constitutes a well-defined fundamental problem by itself, which takes advantage of the second-order statistics distinctness between hypotheses. Theoretical interpretations and numerical results show the tradeoff between detection performance and the degree of side information on the most informative statistics for detection, i.e., the modulation format and spectrum distribution of the primary users.

4.1.2 Nonuniform Correlation Matching

Due to the low occupancy of many communication systems, it is recognized that primary signals are sparse in the spectrum domain, which allows the application of the foreseen theory of compressed sensing [CW08] to further reduce the sampling rate required for spectrum sensing detection. A comparison addressed by [SLH09] shows that nonuniform sampling does not suffer from some drawbacks present in traditional uniform Nyquist rate sampling.

As stated above, ML estimation is an important method used in a wide range of statistical signal processing problems [Kay98b], which is often difficult to implement, even more in the nonuniform sampling scenario. As an alternative, this thesis proposes the application of the correlation-matching technique to spectrum sensing of wideband signals. Correlation-matching is a least-squares fitting of second-order statistics, and behaves as an approximation to ML at the low SNR regime with asymptotic large data records [Por08].

More specifically, the nonuniform sampling version of the correlation-matching approach for spectrum reconstruction of sparse wideband signals is discussed. The potentials of correlation-matching as a matrix-level fitting has been recently addressed by [PNLRS09], and a reformulation is carried out under the nonuniform sampling scenario for asymptotically sparse wideband regimes. A general closed-form estimate of the signal power level based on nonuniformly sampled observations. This unified formulation allows physical interpretation of the problem, and straightforward application to the problem of GLRT spectrum sensing in multi-frequency cognitive radios.

4.1.3 Chapter Organization

This Chapter is organized as follows. The single and multi-frequency GLRT detection based on ML estimation with side information is addressed in Section 4.2. The extension to nonuniform sampling with nonuniform correlation-matching estimation is provided in Section 4.3. Numerical performance evaluation of the detection performance of the GLRT detectors derived in this Chapter is reported in Section 4.4. Section 4.5 concludes the Chapter.

4.2 Maximum Likelihood Estimates in Uniform Sampling

4.2.1 Problem Statement

The spectrum sensing problem for wideband cognitive radio networks is considered, consisting of a set of secondary users, each equipped with one receiving antenna with the purpose of individually monitoring the activity of the primary system, denoted by the wide-sense stationary (WSS) signal $s(t)$, which accounts for the superposition of the primary services over the sensed spectrum of bandwidth B. The sensed signal at the local cognitive radio is

$$x(t) = s(t) + w(t) \tag{4.1}$$

where $w(t)$ is the double-sided complex zero-mean additive white Gaussian noise with spectral density $N_0/2$. In this Section, a block processing of the signal is considered. On the one hand, the N-dimensional discrete-time received signal is defined as

$$\mathbf{x}[m] \doteq [x(t_1^m), \ldots, x(t_N^m)]^T, \tag{4.2}$$

where the sampling instants satisfy the Nyquist condition, i.e, $t_{n+1}^m - t_n^m \leq \frac{1}{f_s}$ for $1 \leq m \leq M$, where f_s is the Nyquist rate. Here, Nyquist rate uniform sampling and piecewise stacking are assumed, for which $t_n^m = t_0 + (mN + n)/f_s$. On the other hand, each cognitive radio acquires M blocks given by $\mathbf{X} \doteq (\mathbf{x}[1], \ldots, \mathbf{x}[M])$. Therefore, the block size N is a side information parameter large enough to cope with the temporal correlation

of the primary signal, and M is a factor to improve the performance of the detectors by averaging independent blocks. Similarly, $\mathbf{S}$ and $\mathbf{W}$ are defined as the uniformed-sampled piece-wise stacked versions of $x(t)$ and $w(t)$, respectively.

Let $\mathcal{H}_0$ be the hypothesis representing the event such that the primary users are not transmitting over the sensed band. In such a case, the local observations during the sensing interval will consist of only noise, i.e., $\mathbf{X} = \mathbf{W}$. On the other hand, let $\mathcal{H}_1$ denote the event in which there is activity of the primary systems during the sensing interval, i.e., $\mathbf{X} = \mathbf{S} + \mathbf{W}$. The spectrum sensing problem may be therefore cast as the binary hypothesis testing problem

$$\begin{aligned} \mathcal{H}_0 &: \mathbf{X} = \mathbf{W} \\ \mathcal{H}_1 &: \mathbf{X} = \mathbf{S} + \mathbf{W}. \end{aligned} \tag{4.3}$$

In (4.3), the column-entries of the signal and noise observations are complex multivariate zero-mean Gaussian distributed with correlation matrices $\mathbf{R}_s = P\mathbf{R}_0$, and $\mathbf{R}_w = \sigma^2\mathbf{I}$, respectively, with $\mathbf{R}_0$ the normalized signal correlation matrix such that $\mathrm{tr}(\mathbf{R}_0) = N$. The detection SNR is further defined as

$$\text{SNR} \doteq \frac{P}{\sigma^2}. \tag{4.4}$$

In the problem at hand, it is a valid assumption that both noise and signal are normally distributed. While facilitating the analysis, this is reasonable because usually there is no line of sight (LoS) path between the cognitive radio and the primary user. Hence, the resulting signal is the superposition of no-LoS signals and approximates Gaussian distribution as the number of observations is sufficiently large, according to the central limit theorem. Moreover, it has recently proved that Gaussian ML estimation provides, asymptotically as SNR $\to 0$, the optimum second-order estimator [VV07]. As the size of the spectrum portion taken into consideration increases, the relative overall occupancy of primary systems becomes low, which motivates the derivation of robust spectrum sensing detectors focused on low SNR regimes. Therefore, it is assumed that asymptotically SNR $\to 0$, where SNR is the nominal SNR at detection.

This work is interested in detecting the presence of the signal $\mathbf{S}$ based on the local observations $\mathbf{X}$ in (4.3). It is known that the GLRT is asymptotically optimal in the Neyman-Pearson sense, i.e., to maximize the detection probability, $\mathbb{P}(\mathcal{H}_1|\mathcal{H}_1)$, for a given false-alarm probability level, $\mathbb{P}(\mathcal{H}_1|\mathcal{H}_0) \leq \epsilon_0$, when the number of observations tends to infinity [Kay98a]. Recently, the finite-sample optimality of GLRT has been established in [Mou09]. Let Ξ_0 and Ξ_1 denote the unknown model parameters under $\mathcal{H}_0$ and $\mathcal{H}_1$, respectively. The optimal test in the Neyman-Pearson sense for deciding between hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$ is given by

$$L(\mathbf{X}, \Theta) \doteq \frac{p(\mathbf{X}|\hat{\Xi}_1, \Theta_1, \mathcal{H}_1)}{p(\mathbf{X}|\hat{\Xi}_0, \Theta_0, \mathcal{H}_0)} \geq \gamma, \tag{4.5}$$

where the ML estimates of the unknown parameters are given by

$$\hat{\Xi}_0 = \arg\max_{\Xi} p(\mathbf{X}|\Xi, \Theta_0, \mathcal{H}_0), \tag{4.6}$$

and

$$\hat{\Xi}_1 = \arg\max_{\Xi} p(\mathbf{X}|\Xi, \Theta_1, \mathcal{H}_1), \tag{4.7}$$

respectively. In (4.5), Θ denotes the set of side information parameters, and $\Theta_1, \Theta_0 \subseteq \Theta$. The threshold γ sets the decision level for which the test $L(\mathbf{X}, \Theta) \geq \gamma$ decides for $\mathcal{H}_1$, and for $\mathcal{H}_0$ otherwise, and is selected to satisfy the false-alarm level $\mathbb{P}(\mathcal{H}_1|\mathcal{H}_0) = \mathbb{P}(L(\mathbf{X}, \Theta) \geq \gamma|\mathcal{H}_0) = \epsilon_0$. Despite its theoretical appeal, the ML estimation and GLRT detection in (4.5) are often difficult to implement, and analytical solutions are not available in many circumstances [Por08]. However, it is shown that GLRT spectrum sensing detection for wideband cognitive radio is encompassed in a low-complex unified framework, stated in the following Theorem.

Theorem 4.1 *Consider the wideband cognitive radio spectrum sensing problem* (4.3)*. Under the Gaussian assumption, any GLRT spectrum sensing detector* (4.5) *is, asymptotically as SNR* $\to 0$ *and* $N \to \infty$*, given by*

$$\boxed{T(\mathbf{X}, \Theta) \doteq \int_B \mathcal{K}(\omega, \Theta) P(\omega) d\omega \geq \tau,} \tag{4.8}$$

where $\mathbf{X}$ *is the asymptotic data record,* Θ *is the set of side information parameters,* $\mathcal{K}(\omega, \Theta)$ *is a kernel associated to each detector,* $P(\omega)$ *is the continuous-frequency periodogram of* $\mathbf{X}$*, and* τ *is the detection threshold. Furthermore, any GLRT spectrum sensing detector only depends, asymptotically as SNR* $\to 0$*, on the second-order statistics of the observations.*

Proof 4.1 *See Appendix B.2.*

From (4.8), it is deduced that GLRT spectrum sensing detection for wideband cognitive radio is strictly based on the second-order statistics of the observations shaped by a kernel that highlights the signal and noise features which are relevant for detection. Moreover, it is shown that frequency-domain asymptotic kernels derived from GLRT spectrum sensing detectors as SNR $\to$ 0 have a common inner structure that depends on the signal and interference plus noise statistics. Let $\phi_s(\omega)$ and $\phi_\nu(\omega)$ denote the power spectral density (PSD) of the signal to be detected and the interference plus noise, respectively. It is shown in Appendix B.2 that the detection kernel in (4.8) has an internal structure given by

$$\boxed{\mathcal{K}_0\left[\phi_s(\omega), \phi_\nu(\omega)\right] \doteq \frac{1}{\phi_\nu(\omega)} \frac{\phi_s(\omega)}{\phi_s(\omega) + \phi_\nu(\omega)},} \tag{4.9}$$

which will be used in subsequent sections to provide a unified perspective of different approaches to the spectrum sensing problem. In the high SNR regime, the $\mathcal{K}_0\left[\phi_s(\omega), \phi_\nu(\omega)\right]$

approaches to $\frac{1}{\phi_\nu(\omega)}$ and the detector asymptotically behaves as the energy detector with noise PSD given by $\phi_\nu(\omega)$. In such a case, the detector focuses on the parts of the spectrum that are less affected by the noise. On the other hand, it is shown that the kernel becomes proportional to $\frac{\phi_s(\omega)}{\phi_\nu^2(\omega)}$ as the interference plus noise increases, and the detector performs the spectral correlation between the periodogram of the observations and the signal PSD weighted by the inverse of the squared spectrum of the interference plus noise. This result corresponds to the locally optimum detector for the cognitive radio problem (4.3), obtained through expanding the optimum quadratic statistic in the low SNR limit [Poo94]. When the signal and interference plus noise statistics are not perfectly characterized, the expression of $\mathcal{K}_0\left[\phi_s(\omega), \phi_\nu(\omega)\right]$ depends on the ML estimates of the unknown parameters. Therefore, the performance of the GLRT spectrum sensing detectors is related to the derivation of the ML estimates in (4.5). Even though parameter estimation is especially challenging in wideband regimes, in what follows, it is shown that analytical solutions are obtained for asymptotic SNR $\to 0$ for many cases of interest.

4.2.2 Single-Frequency Detection with Known Noise Variance

In the following, the optimal single-frequency GLRT detectors are derived based on the assumption that noise variance present at the cognitive radio receiver is known. This is a valid assumption in most cognitive radio networks, as the control layers set predetermined silent periods devoted to threshold computation.

Estimator-Correlator

The optimal test with known statistical parameters is the *estimator-correlator*, given by [Kay98a, Eq. (5.16)]

$$T_1(\mathbf{X}|\mathbf{R}_s, \sigma^2) = \frac{1}{\sigma^2}\mathrm{tr}\left(\mathbf{R}_s(\mathbf{R}_s + \sigma^2\mathbf{I})^{-1}\hat{\mathbf{R}}\right) \geq \tau_1, \tag{4.10}$$

where $\hat{\mathbf{R}}$ stands for the sample covariance matrix, i.e.,

$$\hat{\mathbf{R}} \doteq \frac{1}{M}\sum_m \mathbf{x}_m\mathbf{x}_m^H. \tag{4.11}$$

The detector (4.10) is a classical detection result, which correlates the observations with the output of the Wiener filter or minimum mean square error (MMSE) estimate of the signal, i.e., with the term $\mathbf{R}_s(\mathbf{R}_s + \sigma^2\mathbf{I})^{-1}\mathbf{x}_m$. The frequency-domain asymptotic interpretation of the estimator-correlator has recently been studied in [ZPQ10], from which it is identified that the kernel associated to the estimator-correlator is given by

$$\mathcal{K}_1(\omega, \phi_s, \sigma^2) = \mathcal{K}_0\left[\phi_s(\omega), \sigma^2\right]. \tag{4.12}$$

It is noted that (4.10) and its associated kernel symbolize an upper-bound on the sensing performance of GLRT spectrum sensing detectors, and therefore provide a fundamental limit useful for performance assessments.

Signal Level Detector

Now, the known primary signal level assumption is relaxed. This assumption represents a more realistic scenario, as the power at which the primary signal reaches the cognitive radio receiver is an unknown parameter. Yet, the modulation format employed by the primary services is a side information that can easily be incorporated in the cognitive radio detectors. Under the white noise assumption, the presence of correlation in the observations can be further exploited by the spectrum sensing algorithm to detect the presence of this signal.

As proven in Appendix B.3, for a given normalized signal correlation matrix $\mathbf{R}_0$, the optimal GLRT spectrum sensing detector in the wideband regime with known noise variance, denoted as the *signal level detector*, is given by

$$T_2(\mathbf{X}|\mathbf{R}_0,\sigma^2) = \frac{1}{\sigma^2}\mathrm{tr}\left(\hat{P}\mathbf{R}_0(\hat{P}\mathbf{R}_0+\sigma^2\mathbf{I})^{-1}\hat{\mathbf{R}}\right) \geq \tau_2, \tag{4.13}$$

where the ML estimate of the signal level is given by

$$\hat{P} = \left(\frac{\mathrm{tr}(\mathbf{R}_0\hat{\mathbf{R}}) - \sigma^2 N}{\mathrm{tr}(\mathbf{R}_0^2)}\right)^+. \tag{4.14}$$

The evaluation and comparison of (4.13) in front of (4.10) uncovers the incurred degradation of the spectrum sensing detector when the signal level is unknown by the cognitive radio receiver, as the frequency-domain asymptotic kernel associated to (4.13) is $\mathcal{K}_2(\omega,\phi_0,\sigma^2) = \mathcal{K}_0\left[\hat{P}\phi_0(\omega),\sigma^2\right]$. It is seen that the asymptotic kernel is based on the ML estimate of the signal level P. From the frequency-domain asymptotic interpretation of (4.14), it is deduced that kernel exploits the correlated structure of $\mathbf{R}_0$ and the side information on the noise variance to recover P. It is emphasized that the correlated structure of $\mathbf{R}_0$ is only required for unknown noise variance (c.f. Section 4.2.3), whereas the estimate (4.14) is still valid even for white signal, i.e., $\mathbf{R}_0$ a diagonal matrix, when the noise variance is known.

Toeplitz Detector

Next, the spectrum sensing detection problem is discussed when the primary signal correlation matrix is unknown. According to the GLRT formulation, under $\mathcal{H}_1$ an estimate of $\mathbf{R}_s$ based on the local observations is required to perform the detection. Because $\mathbf{R}_s$ represents the correlation of a stationary signal, it is the solution to ML estimate problem

with the additional constraints $\mathbf{R}_s \succeq \mathbf{0}$ and complex Hermitian Toeplitz structure. Any complex Hermitian Toeplitz matrix of order N is represented by a unique vector of length N containing, e.g., the element of the first row of the matrix. A direct consequence of this property is that $\mathbf{R}_s$ has only N degrees of freedom and, then, it can uniquely be represented by the N first correlation lags, $(r_s[0], \ldots, r_s[N-1])$. However, for the problem at hand, the following decomposition for complex Hermitian Toeplitz matrices is proposed

$$\mathbf{R}[\boldsymbol{\beta}] = \beta_0 \mathbf{T}_0 + \sum_{n=1}^{N-1} \left(\beta_n \mathbf{T}_n^T + \beta_n^* \mathbf{T}_n \right), \tag{4.15}$$

where $\boldsymbol{\beta} \doteq (\beta_0, \ldots, \beta_{N-1})$ are the N coefficients that uniquely represent the matrix $\mathbf{R}$ onto the orthogonal basis $\mathcal{T} \doteq \{\mathbf{T}_n\}$ of the space of complex Hermitian Toeplitz matrices. It is noted that $\mathrm{tr}(\mathbf{T}_i \mathbf{T}_j^T) \propto \delta_{ij}$, where δ_{ij} is the Kronecker delta.

For a given Toeplitz orthogonal basis $\mathcal{T} = \{\mathbf{T}_n\}$, the optimal GLRT spectrum sensing detector in the wideband regime with known noise variance, denoted as *Toeplitz detector*, is derived in Appendix B.4 and is given by

$$T_3(\mathbf{X}|\mathcal{T}, \sigma^2) = \frac{1}{\sigma^2} \mathrm{tr}\left(\mathbf{R}_s[\hat{\boldsymbol{\beta}}] (\mathbf{R}_s[\hat{\boldsymbol{\beta}}] + \sigma^2 \mathbf{I})^{-1} \hat{\mathbf{R}} \right) \geq \tau_3, \tag{4.16}$$

where the coefficients of $\hat{\mathbf{R}}_s$ onto $\mathcal{T}$ are given by

$$\hat{\beta}_0 = \left(\frac{1}{N} \mathrm{tr}(\hat{\mathbf{R}}) - \sigma^2 \right)^+ \tag{4.17a}$$

$$\hat{\beta}_n = \frac{\mathrm{tr}(\mathbf{T}_n \hat{\mathbf{R}})}{\mathrm{tr}(\mathbf{T}_n \mathbf{T}_n^T)}, \quad 1 \leq n \leq N-1. \tag{4.17b}$$

It is noticed that when employing diagonal matrices, i.e., $\mathbf{T}_0 = \mathbf{I}$, and $\mathbf{T}_n$ all-zeros matrix with an all-ones semi-diagonal n-positions above the main diagonal, the coefficients β_n have the physical meaning of the correlation lags, i.e., $\hat{\beta}_n = \hat{R}_s[n]$, for $0 \leq n \leq N-1$. Therefore, the computation of the zero-lag β_0 in (4.17a) is based on the detected energy and the side information on the noise variance. Conversely, because $\mathrm{tr}(\mathbf{T}_n) = 0$ for $n > 0$, each coefficient β_n in (4.17b) uncovers the stationary part of the received observations, i.e., takes into account the off-diagonal information contained in $\hat{\mathbf{R}}$. The kernel associated to (4.16) is given by

$$\mathcal{K}_3(\omega, \mathcal{T}, \sigma^2) = \mathcal{K}_0 \left[\left(\hat{\beta}_0 + \sum_{n=1}^{N-1} \mathrm{Re}(\hat{\beta}_n) \phi_n(\omega) \right)^+, \sigma^2 \right], \tag{4.18}$$

where $\phi_n(\omega)$ is spectral density associated to the correlation matrices $(\mathbf{T}_n + \mathbf{T}_n^T)$, and $\mathrm{Re}(z)$ takes the real part of $z \in \mathbb{C}$. The GLRT spectrum sensing detector (4.16) exploits the side information on the noise variance to perform optimal matching between the observations

and the frequency patterns ϕ_n. When noting that $\sigma^2 + \hat{\beta}_0$ is the energy detector, it is appreciated that the detection takes advantage of the frequency variations of the periodogram. Hence, it is expected to achieve performance gain with respect to the energy detector.

4.2.3 Single-Frequency Detection with Unknown Noise Variance

Many detectors, including the energy detector, assume exact side information on the noise variance to properly perform the detection. Yet in practice, a mismatch on the noise variance may significantly degrade the sensing performance of the detectors, as reported in [TS05, TS08]. For this purpose, the single-frequency GLRT spectrum sensing detectors derived in Section 4.2.2 are extended for unknown noise variance. Under $\mathcal{H}_0$, the ML estimate of σ^2 is given by [Kay98a, Eq. (9.10)]

$$\hat{\sigma}_0^2 = \frac{1}{N}\mathrm{tr}(\hat{\mathbf{R}}). \tag{4.19}$$

Noise Level Detector

It is first assumed perfect side information on the signal correlation matrix $\mathbf{R}_s$. Even though this scenario is not realistic, it provides a relaxed upper-bound on the sensing performance with unknown noise variance. The ML estimate of σ^2 under $\mathcal{H}_1$ is derived in [Kay98a] for the low SNR regime, and is given by

$$\hat{\sigma}_1^2 = \frac{1}{N}\left[\mathrm{tr}\left(\hat{\mathbf{R}} - \mathbf{R}_s\right)\right]^+. \tag{4.20}$$

For a given signal correlation matrix $\mathbf{R}_s$, Particularizing (B.5) by (4.20) and $\hat{\mathbf{R}}_s = \mathbf{R}_s$, it follows that the optimal GLRT spectrum sensing detector in the wideband regime with unknown noise variance, denoted as *noise level detector*, is given by

$$T_4(\mathbf{X}|\mathbf{R}_s) = \frac{1}{\hat{\sigma}_1^2}\mathrm{tr}\left(\mathbf{R}_s(\hat{\sigma}_1^2\mathbf{I} + \mathbf{R}_s)^{-1}\hat{\mathbf{R}}\right) \geq \tau_4, \tag{4.21}$$

where the ML estimate of the noise variance under $\mathcal{H}_1$ is given by (4.20). The side information on $\mathbf{R}_s$ is twofold. On the one hand, recalling that the kernel associated to (4.21) is

$$\mathcal{K}_4(\omega, \phi_s) = \mathcal{K}_0\left[\phi_s(\omega), \hat{\sigma}_1^2\right], \tag{4.22}$$

it is noted that the side information pattern for detecting primary systems is supplied by $\phi_s(\omega)$. On the other hand, the side information on the energy of $\phi_s(\omega)$ diminishes the problem of noise mismatching, as the estimate of the noise variance is based on both the received observations and $\phi_s(\omega)$.

Signal and Noise Levels Detector

Next, the spectrum sensing problem is considered when the cognitive radio receivers have perfect side information on the normalized signal correlation matrix, $\mathbf{R}_0$. Hence, based on the structure of $\mathbf{R}_0$, the optimal GLRT spectrum sensing detector aims at recover both the signal level and noise variance. For a given normalized signal correlation matrix $\mathbf{R}_0$, in Appendix B.5 it is proved that the optimal GLRT spectrum sensing detector in the wide-band regime with unknown noise variance, denoted as the *signal and noise levels detector*, is given by

$$T_5(\mathbf{X}|\mathbf{R}_0) = \frac{1}{\hat{\sigma}_1^2}\mathrm{tr}\left(\hat{P}\mathbf{R}_0(\hat{\sigma}_1^2\mathbf{I} + \hat{P}\mathbf{R}_0)^{-1}\hat{\mathbf{R}}\right) \geq \tau_5, \tag{4.23}$$

where the ML estimates of the signal level and noise variance are given by

$$\hat{P} = \left(\frac{\mathrm{tr}\left(\hat{\mathbf{R}}(\mathbf{R}_0 - \mathbf{I})\right)}{\mathrm{tr}(\mathbf{R}_0^2) - N}\right)^+ \tag{4.24a}$$

$$\hat{\sigma}_1^2 = \left(\frac{1}{N}\mathrm{tr}(\hat{\mathbf{R}}) - \hat{P}\right)^+. \tag{4.24b}$$

The ML estimates of P and σ^2 under $\mathcal{H}_1$ provide further insight on the sensing performance of the test statistic (4.23). It is first observed that (4.24a) is proportional to $\mathrm{tr}(\hat{\mathbf{R}}\mathbf{R}_\mathrm{D})$, where $\mathbf{R}_\mathrm{D}$ is the diagonal off-loaded correlation matrix of the primary signal, defined as $\mathbf{R}_\mathrm{D} \doteq \mathbf{R}_0 - \mathbf{I}$. Under the stationarity assumption, it is outlined that (4.24a) is an energy detector that takes into account solely the statistics of the received observations that are not affected by the noise, i.e., the presence of non-zero correlation lags. This is because the diagonal entries of $\mathbf{R}_\mathrm{D}$ are all null. Conversely, if the primary signal is cyclostationary, the main diagonal of $\mathbf{R}_0$ is not uniform, and (4.24a) evaluates the variability around the mean of the main diagonal. In conclusion, the ML estimate of the signal level is a linear combination of a measure of the instantaneous energetic variability and a measure of the degree of autocorrelation present in the non-zero lags. Similar interpretations can be obtained in view of the asymptotic associated kernel

$$\mathcal{K}_5(\omega, \phi_0) = \mathcal{K}_0\left[\hat{P}\phi_0(\omega), \hat{\sigma}_1^2\right] \tag{4.25}$$

in the frequency-domain, because the ML estimate (4.24a) is strictly based on the variations of the normalized pattern $\phi_0(\omega)$ around its spectral mean.

Toeplitz and Noise Level Detector

Finally, the optimal spectrum sensing detector is discussed when the signal correlation matrix is fully unknown to the cognitive radio receivers with the additional constraint of complex Hermitian Toeplitz structure. For a given Toeplitz orthogonal basis $\mathcal{T}$, the optimal

GLRT spectrum sensing detector in the wideband regime with unknown noise variance, i.e., the *Toeplitz and noise level detector*, is given by

$$T_6(\mathbf{X}|\mathcal{T}) = \frac{1}{\hat{\sigma}_1^2 + \hat{\beta}_0} \mathrm{tr}\left(\mathbf{R}_s[\hat{\boldsymbol{\beta}}_1](\hat{\sigma}_1^2\mathbf{I} + \mathbf{R}_s[\hat{\boldsymbol{\beta}}_1])^{-1}\hat{\mathbf{R}}\right) \geq \tau_6, \tag{4.26}$$

where $\boldsymbol{\beta}_1 \doteq (0, \beta_1, \ldots, \beta_{N-1})$, and the ML estimates of the noise variance and the coefficients of $\hat{\mathbf{R}}_s$ onto $\mathcal{T}$ are given by

$$\hat{\beta}_0 + \hat{\sigma}_1^2 = \frac{1}{N}\mathrm{tr}(\hat{\mathbf{R}}) \tag{4.27a}$$

$$\hat{\beta}_n = \frac{\mathrm{tr}(\mathbf{T}_n\hat{\mathbf{R}})}{\mathrm{tr}(\mathbf{T}_n\mathbf{T}_n)}, \quad 1 \leq n \leq N-1. \tag{4.27b}$$

The details are reported in Appendix B.6. The frequency-domain asymptotic interpretation of (4.26) can be outlined from its associated kernel, which is given by

$$\mathcal{K}_6(\omega, \mathcal{T}) = \mathcal{K}_0\left[\left(\sum_{n=1}^{N-1} \mathrm{Re}(\hat{\beta}_n)\phi_n(\omega)\right)^+, \hat{\beta}_0 + \hat{\sigma}_1^2\right]. \tag{4.28}$$

When releasing the structure of $\mathbf{R}_0$, it is observed that the test statistic (4.26) cannot separate the signal and noise energy contributions, because (4.27a) is treated as interference plus noise. As a consequence, the sensing performance depends on the variability of the periodogram, i.e., the contributions of the patterns $\phi_n(\omega)$, for $n \geq 1$, with respect to the total received energy. It is further noted that indoor environments are characterized by a high degree of diversity, due to multi-path propagation. Therefore, detection is expected due to temporal correlation, although the white part of the signal is not recovered.

4.2.4 Multi-Frequency Spectrum Sensing

Next, the spectrum sensing problem for cognitive radio systems is considered where the primary system employ frequency-division multiplexing (FDM) with predetermined channelization. The sensed multi-frequency system is characterized by Q adjacent channels, and hence the sampled baseband observations admit the multi-frequency structure

$$\mathbf{x}[m] = \sum_{q=1}^{Q} \mathbf{s}_q[m] + \mathbf{w}[m], \tag{4.29}$$

for $1 \leq m \leq M$, where $\mathbf{s}_q$ denotes the received signal located at the q-th channel, and $\mathbf{w}$ is the complex zero-mean additive white Gaussian noise with variance σ^2. The statistics of $\mathbf{s}_q$ are modeled as complex zero-mean Gaussian with correlation matrix $P_q\mathbf{R}_q$, where P_q stands for the received power level on the q-th channel, and $\mathbf{R}_q$ is the normalized Toeplitz correlation matrix of the primary signal on the q-th channel, with $\mathrm{tr}(\mathbf{R}_q) = N$. If the sensed

multi-frequency system employs homogeneous services, the signal statistics across channels further accomplish $\mathbf{R}_q = \mathbf{R}_0 \odot (\mathbf{e}(\omega_q)\mathbf{e}^H(\omega_q))$, where $\mathbf{R}_0$ is the baseband basic modulation format, and $\mathbf{e}(\omega)$ is the frequency vector at ω, i.e., $\mathbf{e}^H(\omega) \doteq \left[1\ e^{\jmath\omega}\ \ldots\ e^{\jmath\omega(N-1)}\right]$. Let $\mathbf{R}_s[\mathbf{P}] \doteq \sum_{q=1}^{Q} P_q\mathbf{R}_q$. For notation purposes, $\mathcal{M} = \{\mathbf{R}_q\}$ is defined as the set of normalized multi-frequency correlation matrices, and $\mathbf{P} \doteq (P_1, \ldots, P_Q)^T$.

The spectrum sensing problem for cognitive radio systems employing multi-frequency signals consists in detecting the presence of primary user activity at each channel, $1 \leq q \leq Q$. The corresponding joint multiple-hypotheses test [Kay98a] evaluates the $2^Q - 1$ possible combinatorial occupancies. Hence, the complexity of multiple-hypotheses testing grows exponentially with the number of channels and becomes an impractical approach due to the limited processing capabilities of the wideband cognitive radios. It is noticed that spectrum sensing on the q-th channel is governed by the signal level P_q. Therefore, (4.29) can be treated as a nuisance-hypotheses testing problem. Let $\mathcal{H}_{1,q}$ and $\mathcal{H}_{0,q}$ denote the hypotheses representing the primary system transmitting or not transmitting over the q-th channel, respectively. Let $\mathbf{P}_{\bar{q}} \doteq (P_1, \ldots, P_{q-1}, 0, P_{q+1}, \ldots, P_Q)^T$, for $1 \leq q \leq Q$. The nuisance-hypotheses testing problem at the q-th channel is then given by

$$L_q(\mathbf{X}, \Theta) = \frac{p(\mathbf{X}|\hat{\mathbf{P}}, \Theta_1, \mathcal{H}_{1,q})}{p(\mathbf{X}|\hat{\mathbf{P}}_{\bar{q}}, \Theta_0, \mathcal{H}_{0,q})} \geq \gamma_q, \tag{4.30}$$

for $1 \leq q \leq Q$. The complexity has been reduced to a set of Q binary tests.

The optimal GLRT spectrum sensing detectors for the multi-frequency model (4.29) with known and unknown noise variance are discussed in the sequel. It is assumed that the frequencies ω_q and the baseband modulation format $\mathbf{R}_0$ are perfectly known.

Multi-Frequency Detector

Assume that each cognitive radio device has perfect side information on the noise variance σ^2, as well as the primary system multi-frequency structure $\mathcal{M}$. It is shown in Appendix B.7 that for a given multi-frequency system $\mathcal{M}$, the optimal GLRT spectrum sensing detector at the q-th frequency in the wideband regime with known noise variance, i.e., the *multi-frequency detector*, is given by

$$T_{7,q}(\mathbf{X}|\mathcal{M}, \sigma^2) = \mathrm{tr}\left(\hat{P}_q\boldsymbol{\Xi}_q^{-1}\mathbf{R}_q(\hat{P}_q\mathbf{R}_q + \boldsymbol{\Xi}_q)^{-1}\hat{\mathbf{R}}\right) \geq \tau_{7,q}, \tag{4.31}$$

where $\boldsymbol{\Xi}_q \doteq \sum_{l\neq q} \hat{P}_l\mathbf{R}_l + \sigma^2\mathbf{I}$, and the ML estimates of the signal levels are given by

$$\begin{pmatrix} \mathrm{tr}(\mathbf{R}_1^2) & \ldots & \mathrm{tr}(\mathbf{R}_1\mathbf{R}_Q) \\ \vdots & & \vdots \\ \mathrm{tr}(\mathbf{R}_Q\mathbf{R}_1) & \ldots & \mathrm{tr}(\mathbf{R}_Q^2) \end{pmatrix} \times \begin{pmatrix} P_1 \\ \vdots \\ P_Q \end{pmatrix} = \begin{pmatrix} \mathrm{tr}(\mathbf{R}_1\hat{\mathbf{R}}) \\ \vdots \\ \mathrm{tr}(\mathbf{R}_Q\hat{\mathbf{R}}) \end{pmatrix} - \sigma^2 N\mathbf{1}, \tag{4.32}$$

where $\mathbf{1}$ is the all-ones column vector. It is noted that when the number of available samples is small, the orthogonality between channels is not preserved and, in general, the system of equations (4.32) is coupled because $\mathrm{tr}(\mathbf{R}_q\mathbf{R}_l) \neq 0$, for $l \neq q$. However, for large data records, the system matrix in (4.32) becomes nearly diagonal, and the ML estimates at each channel are independent on the other channels, giving $\mathrm{tr}(\mathbf{R}_q^2)\hat{P}_q = \mathrm{tr}(\mathbf{R}_q\hat{\mathbf{R}}) - \sigma^2 N$. In both cases, the coefficients $\mathrm{tr}(\mathbf{R}_q\mathbf{R}_l)$ in (4.32) can be computed off-line. The frequency-domain asymptotic kernel for the spectrum sensing detection at the q-th channel is given by

$$\mathcal{K}_{7,q}(\omega, \mathcal{M}, \sigma^2) = \mathcal{K}_0 \left[\hat{P}_q\phi_q(\omega), \sum_{l\neq q} \hat{P}_l\phi_l(\omega) + \sigma^2 \right]. \tag{4.33}$$

It is seen that the detector employs the relative occupancy on the remaining frequencies as interference for sensing the q-th channel. As expected, the performance of (4.31) is affected by the signal-to-interference-plus-noise ratio (SINR) based on the cross-correlation that arises from the adjacent channels. Finally, it is worth highlighting that (4.31) and (4.32) are a generalization of the wideband signal level detector (4.13) and the estimate (4.14).

Multi-Frequency and Noise Level Detector

Finally, the detection of multi-frequency systems with unknown noise variance is considered. For a given multi-frequency system $\mathcal{M}$, the optimal GLRT spectrum sensing detector at the q-th channel in the wideband regime with unknown noise variance is derived in Appendix B.8 and denoted as the *multi-frequency and noise level detector*, and is given by

$$T_{8,q}(\mathbf{X}|\mathcal{M}) = \mathrm{tr}\left(\mathbf{\Xi}_q^{-1}\hat{P}_q\mathbf{R}_q(\hat{P}_q\mathbf{R}_q + \mathbf{\Xi}_q)^{-1}\hat{\mathbf{R}}\right) \geq \tau_{8,q}, \tag{4.34}$$

where $\mathbf{\Xi}_q \doteq \sum_{l\neq q}\hat{P}_l\mathbf{R}_l + \hat{\sigma}_1^2\mathbf{I}$, and the ML estimates of the signal levels and noise variance are given by

$$\begin{pmatrix} \mathrm{tr}(\mathbf{R}_1^2) & \dots & \mathrm{tr}(\mathbf{R}_1\mathbf{R}_Q) & N \\ \vdots & & \vdots & \vdots \\ \mathrm{tr}(\mathbf{R}_Q\mathbf{R}_1) & \dots & \mathrm{tr}(\mathbf{R}_Q^2) & N \\ N & \vdots & N & N \end{pmatrix} \times \begin{pmatrix} P_1 \\ \vdots \\ P_Q \\ \sigma_1^2 \end{pmatrix} = \begin{pmatrix} \mathrm{tr}(\mathbf{R}_1\hat{\mathbf{R}}) \\ \vdots \\ \mathrm{tr}(\mathbf{R}_Q\hat{\mathbf{R}}) \\ \mathrm{tr}(\hat{\mathbf{R}}) \end{pmatrix}. \tag{4.35}$$

The main advantage of the test statistic (4.30) is that all the information on the sensed bandwidth is exploited for joint detection and estimation at a given band. Whereas filter-bank based detectors may suffer from adjacent channel leakage, the nuisance parameter formulation allows the detector to take advantage of the multi-frequency structure $\mathcal{M}$ for estimating both signal levels and noise variance. The frequency-domain asymptotic inter-

pretation of the associated kernel

$$\mathcal{K}_{8,q}(\omega, \mathcal{M}) = \mathcal{K}_0 \left[\hat{P}_q \phi_q(\omega), \sum_{l \neq q} \hat{P}_l \phi_l(\omega) + \hat{\sigma}_1^2 \right] \tag{4.36}$$

shows that the ML estimate of the noise variance flourishes together with the remaining bands as interference. It is also noted that for $Q = 1$, (4.34) and (4.35) reduce to (4.23) and (4.24), respectively.

4.3 Correlation-Matching Estimates in Nonuniform Sampling

4.3.1 Problem Formulation

Due to the difficulty of ML estimation in nonuniform sampling, the primary signal detection problem is addressed by correlation-matching estimation. Analogously to the signal model described in Section 3.2, the observation set satisfies the following nonuniform sampling model

$$\mathbf{x}_m = \mathbf{\Psi}_m(\mathbf{s}_m + \mathbf{w}_m). \tag{4.37}$$

where $\mathbf{\Psi}_m$ is a pinning matrix that randomly selects K samples of $\mathbf{x}_m$, and it is given by randomly selecting K rows of $\mathbf{I}_N$, with $K < N$. For signal detection, the correlation model $\mathcal{R}(\Theta)$ is now given by

$$\mathcal{R}(P, \sigma^2) = P\mathbf{R}_0 + \sigma^2 \mathbf{I}_N, \tag{4.38}$$

where $\mathbf{R}_0$ denotes the normalized signal correlation matrix, e.g., with $\mathrm{tr}(\mathbf{R}_0) = N$, and P and σ^2 denote the signal and noise power levels over B, respectively. Notice that now $\mathbf{R}_0$ has the meaning of the primary signal correlation matrix, not the rank-1 matrix result of the outer product of a frequency template. For that reason, in this Section the correlation-matching metric is reformulated as

$$\mathcal{M}\left[\mathbf{X}, \mathcal{R}(P, \sigma^2)\right] \doteq \|\mathbf{X}\mathbf{X}^H - \sum_m \mathbf{\Psi}_m \mathcal{R}(P, \sigma^2) \mathbf{\Psi}_m^H\|^2, \tag{4.39}$$

i.e., the matching is extended to full-rank matrices. While having all the advantages than that of (3.8) in Chapter 3, the solution to the problem

$$\hat{P}, \hat{\sigma}^2 = \arg\min_{P, \sigma^2} \mathcal{M}\left[\mathbf{X}, \mathcal{R}(P, \sigma^2)\right] \tag{4.40}$$

has a different structure. More concretely, as derived in Appendix B.9, it is given by

$$\boxed{\hat{P} = \frac{\mathrm{tr}\left[\hat{\mathbf{R}}\left(\overline{\mathbf{R}}_0 - \frac{1}{K}\mathrm{tr}(\overline{\mathbf{R}}_0)\mathbf{I}_K\right)\right]^+}{\mathrm{tr}(\overline{\mathbf{R}}_0^2) - \frac{1}{K}\mathrm{tr}^2(\overline{\mathbf{R}}_0),}} \tag{4.41}$$

where

$$\hat{\mathbf{R}} \doteq \frac{1}{M}\mathbf{X}\mathbf{X}^H \tag{4.42}$$

is the sample covariance matrix of the nonuniformly sampled observations,

$$\overline{\mathbf{R}}_0 \doteq \frac{1}{M}\sum_{m=1}^{M} \mathbf{\Psi}_m \mathbf{R}_0 \mathbf{\Psi}_m^H \tag{4.43}$$

is the nonuniformly sampled version of the normalized signal correlation matrix, and $x^+ \doteq \max(0, x)$.

The physical interpretation of (4.41) can be discussed from the nonuniformly sampled statistics $\hat{\mathbf{R}}$, and $\overline{\mathbf{R}}_0$. It can be appreciated that the expression of the numerator of $\hat{P}$ can be rewritten as the scalar product $\mathrm{tr}(\hat{\mathbf{R}}\mathbf{R}_\mathrm{D})$, the $K \times K$ *diagonal off-loaded signal correlation matrix* has been defined as

$$\mathbf{R}_\mathrm{D} \doteq \overline{\mathbf{R}}_0 - \frac{1}{K}\mathrm{tr}(\overline{\mathbf{R}}_0)\mathbf{I}_K. \tag{4.44}$$

The first term of $\mathbf{R}_\mathrm{D}$ performs the projection of the sample covariance matrix $\hat{\mathbf{R}}$ onto the nonuniform sampling normalized signal space defined by $\overline{\mathbf{R}}_0$. The second term is in charge of subtracting the part of the reconstructed observations that are considered as noise according to the signal model (4.37). Because the spectrum of $s(t)$ is strongly sparse, the average energy is measured in the nonuniform sampling domain, i.e., over $\overline{\mathbf{R}}_0$, rather than directly on $\mathbf{R}_0$. If $s(t)$ is a stationary process, it is noticed that $\mathbf{R}_\mathrm{D} = \overline{\mathbf{R}}_0 - \mathbf{I}_K$, because $\overline{\mathbf{R}}_0$ has an all-ones main diagonal. The projection of $\mathbf{B}$ is then performed onto the components of $\overline{\mathbf{R}}_0$ that are not affected by the noise, i.e., the non-zero correlation lags contained out of the main diagonal. Conversely, if $s(t)$ is cyclostationary, the main diagonal of $\overline{\mathbf{R}}_0$ is not uniform, and $\mathbf{R}_\mathrm{D}$ evaluates the energetic variability around the average energy, in addition to the correlation of the non-zero lags.

Moreover, the denominator of (4.41) reflects the distinctness between $\overline{\mathbf{R}}_0$ and $\bar{\mathbf{I}}_N$, as it becomes subjected to the Cauchy-Schwarz inequality. Let $\langle \overline{\mathbf{A}}, \overline{\mathbf{B}} \rangle$ be the inner product between $\overline{\mathbf{A}}$ and $\overline{\mathbf{B}}$, defined as $\langle \overline{\mathbf{A}}, \overline{\mathbf{B}} \rangle \doteq \mathrm{tr}\left[\left(\frac{1}{M}\sum_m \mathbf{\Psi}_m \mathbf{A} \mathbf{\Psi}_m^H\right)\left(\frac{1}{M}\sum_l \mathbf{\Psi}_l \mathbf{B} \mathbf{\Psi}_l^H\right)\right]$. Further consider the squared norm of $\overline{\mathbf{A}}$ as $\|\overline{\mathbf{A}}\|^2 \doteq \langle \overline{\mathbf{A}}, \overline{\mathbf{A}} \rangle$. By writing K times the denominator of (4.41), each term can be written in function of inner products and norms, i.e., $K\mathrm{tr}(\overline{\mathbf{R}}_0^2) = \|\overline{\mathbf{R}}_0^2\|^2 \|\bar{\mathbf{I}}_N^2\|^2$, and $\mathrm{tr}^2(\overline{\mathbf{R}}_0) = \langle \overline{\mathbf{R}}_0^2, \bar{\mathbf{I}}_N^2 \rangle^2$. After taking the squared root of both terms, it is shown the non-negativity of the denominator, because it holds that

$$\|\overline{\mathbf{R}}_0^2\| \|\bar{\mathbf{I}}_N^2\| \geq \langle \overline{\mathbf{R}}_0^2, \bar{\mathbf{I}}_N^2 \rangle. \tag{4.45}$$

It is noticed that the more distinct are the second-order statistics of the signal and the noise, the more robust is the estimate (4.41). The equality in (4.45) only holds when the signal and noise statistics are linearly dependent, i.e., both white, or when $K = 1$, for which in the sequel it is assumed $K > 1$.

4.3.2 Application to Multi-Frequency Spectrum Sensing

It is recognized that wideband cognitive radio networks are characterized by a sparse primary systems' spectrum, which motivates the application of the nonuniform correlation-matching scheme to spectrum sensing. Now, $s(t)$ is proposed in (4.37) as a multi-frequency system given by

$$s(t) \doteq \sum_{q=1}^{Q} s_q(t), \tag{4.46}$$

i.e., the addition of primary signals all with the same spectral pattern and occupying Q adjacent bands. This model actually corresponds to many real primary systems such as the DVB-T standard. The correlation model $\mathcal{R}(\Theta)$ for multi-frequency systems is hereby given by

$$\mathcal{R}(P_1, \ldots, P_Q, \sigma^2) \doteq \sum_{q=1}^{Q} P_q \mathbf{R}_q + \sigma^2 \mathbf{I}_N, \tag{4.47}$$

where $\mathcal{W} \doteq \{\mathbf{R}_q\}$ is the set of normalized correlation matrices of uniform Nyquist rate sampled versions of $s_q(t)$, with $\mathrm{tr}(\mathbf{R}_q) = N$, and P_q denote the power levels. ML estimation, which constitutes a natural subprocess within the spectrum sensing under the GLRT perspective, offers robustness in front of inaccurate model parameters. In the sequel, the nonuniform correlation-matching approach as an alternative to ML in the low SNR regime is discussed, as well as its application to spectrum sensing.

For given $\mathcal{W}$, the nonuniform sampling GLRT spectrum sensing of multi-frequency systems involves the estimation of the signal and noise power levels from the nonuniformly sampled observations $\mathbf{X}$ [Kay98a]. The following result generalizes (4.41) to multi-frequency signals. By considering the the multi-frequency model (4.47), in Appendix B.10 it is shown that the solution of the signal and noise contributions under the the metric (4.39) is given by the solution of the system of equations

$$\begin{pmatrix} \mathrm{tr}(\overline{\mathbf{R}}_1^2) & \ldots & \mathrm{tr}(\overline{\mathbf{R}}_1 \overline{\mathbf{R}}_Q) & \mathrm{tr}(\overline{\mathbf{R}}_1) \\ \vdots & \ddots & \vdots & \vdots \\ \mathrm{tr}(\overline{\mathbf{R}}_Q \overline{\mathbf{R}}_1) & \ldots & \mathrm{tr}(\overline{\mathbf{R}}_Q^2) & \mathrm{tr}(\overline{\mathbf{R}}_Q) \\ \mathrm{tr}(\overline{\mathbf{R}}_1) & \ldots & \mathrm{tr}(\overline{\mathbf{R}}_Q) & K \end{pmatrix} \times \begin{pmatrix} \hat{P}_1 \\ \vdots \\ \hat{P}_Q \\ \hat{\sigma}^2 \end{pmatrix} \begin{pmatrix} \mathrm{tr}(\hat{\mathbf{R}} \overline{\mathbf{R}}_1) \\ \vdots \\ \mathrm{tr}(\hat{\mathbf{R}} \overline{\mathbf{R}}_Q) \\ \mathrm{tr}(\hat{\mathbf{R}}) \end{pmatrix}, \tag{4.48}$$

where

$$\overline{\mathbf{R}}_q \doteq \frac{1}{M} \sum_{m=1}^{M} \mathbf{\Psi}_m \mathbf{R}_q \mathbf{\Psi}_m^H \tag{4.49}$$

is the nonuniform sampling version of the normalized signal correlation matrix at the q-th band. The detection of the signal levels $(P_1, \ldots, P_Q)$ is based on the nonuniformly sampled observations $\mathbf{X}$, which contain spectral information of the sensed bandwidth. Contrarily to conventional filter-bank techniques, the nonuniform correlation-matching approach is

based on the system matrix (4.48) which reflects the cross-correlation between different bands given that orthogonality is not preserved after compressed sensing.

4.3.3 Nonuniform Sampling Multi-Frequency Spectrum Sensing

By taking the optimal multi-frequency GLRT (MF-GLRT) for asymptotic wideband regime, i.e.,

$$T_q\mathbf{X}|\mathcal{W}) = \mathrm{tr}\left[\boldsymbol{\Xi}_q^{-1}\hat{P}_q\overline{\mathbf{R}}_q\left(\boldsymbol{\Xi}_q + \hat{P}_q\overline{\mathbf{R}}_q\right)^{-1}\hat{\mathbf{R}}\right] \geq \tau_q, \tag{4.50}$$

where $\boldsymbol{\Xi}_q \doteq \sum_{l\neq q}\hat{P}_l\overline{\mathbf{R}}_l + \hat{\sigma}^2\mathbf{I}_K$ represents the equivalent noise-plus-interferences covariance when sensing the q-th band, $(\hat{P}_1,\ldots,\hat{P}_Q,\hat{\sigma}^2)$ are the estimates of the multi-frequency signal and noise levels, given by (4.48), and $\hat{\mathbf{R}} \doteq \frac{1}{M}\mathbf{X}\mathbf{X}^H$ is the sample covariance matrix of the nonuniformly sampled observations. In (4.50), the threshold τ_q is set to satisfy the false-alarm constraint

$$\mathbb{P}\left[T_q(\mathbf{X}|\mathcal{W}) \geq \tau_q|\mathcal{H}_0\right] = \epsilon_1^q \tag{4.51}$$

which, without loss of generality, it is set to $\epsilon_1^q = \epsilon_1$ for $q = 1,\ldots,Q$. While according to the GLRT theory $(\hat{P}_1,\ldots,\hat{P}_Q,\hat{\sigma}^2)$ are the ML estimates of the signal and noise power levels, (4.48) are alternatively considered as a valid approximation to ML in the low SNR regime [Por08]. It is seen that the MF-GLRT (4.50) takes advantage of the occupancy on the remaining frequencies as equivalent noise when sensing the q-th band. Therefore, the performance of (4.50) is affected by the SINR based on the cross-correlation that arises from the adjacent bands, and the signal and noise power level reconstruction in (4.48).

4.4 Numerical Results

4.4.1 Maximum Likelihood Estimates

The sensing performance is assessed by means of experimental simulations modeling practical wideband cognitive radio scenarios. The GLRT spectrum sensing detectors derived in this work, which are summarized in Table 4.1, are evaluated along with the energy detector [Poo94], and the AGM detector [ZLLZ10]. The AGM detector is based on the white noise assumption with unknown noise variance and signal correlation matrix. Even though it was originally motivated by the use of multiple antennas, the AGM statistic is applied on $\hat{\mathbf{R}}$ because it is a good measure of the spread of the set of eigenvalues that arise temporal correlation.

First, the cognitive radio model (4.3) consisting of a secondary user equipped with a single sensing antenna is simulated, which takes $N = 32$ samples of the wideband signal $x(t)$ with a sampling depth of $M = 2N$. In this scenario, the relative occupancy of the primary signal $s(t)$ is 25% of the sensing band, i.e., the spectral support of $\phi_s(\omega)$ is $\frac{1}{4}B$. Figure 4.2 depicts the receiver operating characteristics (ROC) of the wideband GLRT spectrum

Table 4.1: Summary of GLRT Spectrum Sensing Detectors' Kernels and ML Estimates

Detector	Test Statistic	Equation No.	ML Estimates
Estimator-Correlator	$T_1(\mathbf{X}\|\mathbf{R}_s, \sigma^2)$	(4.10)	-
Signal Level	$T_2(\mathbf{X}\|\mathbf{R}_0, \sigma^2)$	(4.13)	(4.14)
Toeplitz	$T_3(\mathbf{X}\|\mathcal{T}, \sigma^2)$	(4.16)	(4.17)
Noise Level	$T_4(\mathbf{X}\|\mathbf{R}_s)$	(4.21)	(4.20)
Signal and Noise Levels	$T_5(\mathbf{X}\|\mathbf{R}_0)$	(4.23)	(4.24)
Toeplitz and Noise Level	$T_6(\mathbf{X}\|\mathcal{T})$	(4.26)	(4.27)
Multi-Frequency	$T_{7,q}(\mathbf{X}\|\mathcal{M}, \sigma^2)$	(4.31)	(4.32)
Multi-Frequency Noise	$T_{8,q}(\mathbf{X}\|\mathcal{M})$	(4.34)	(4.35)

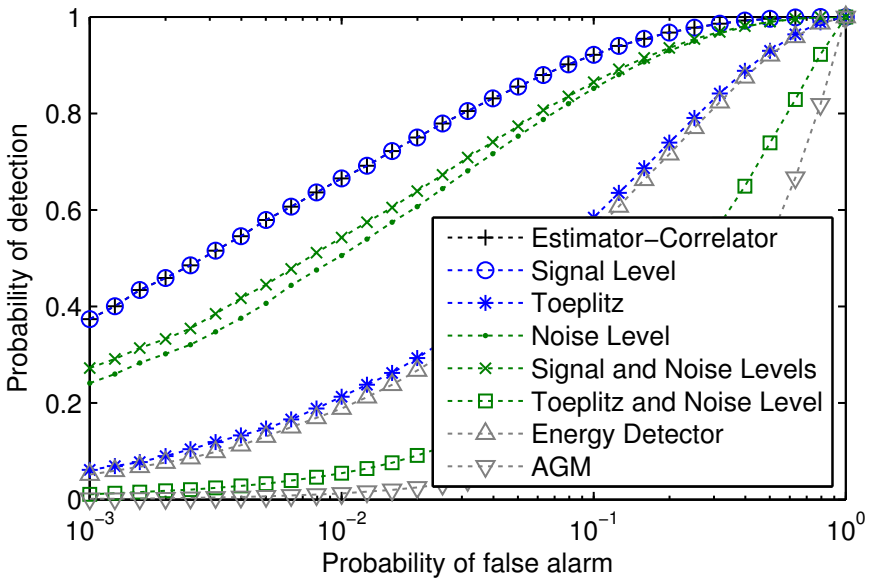

Figure 4.2: ROC of the wideband spectrum sensing detectors at SNR of -15 dB.

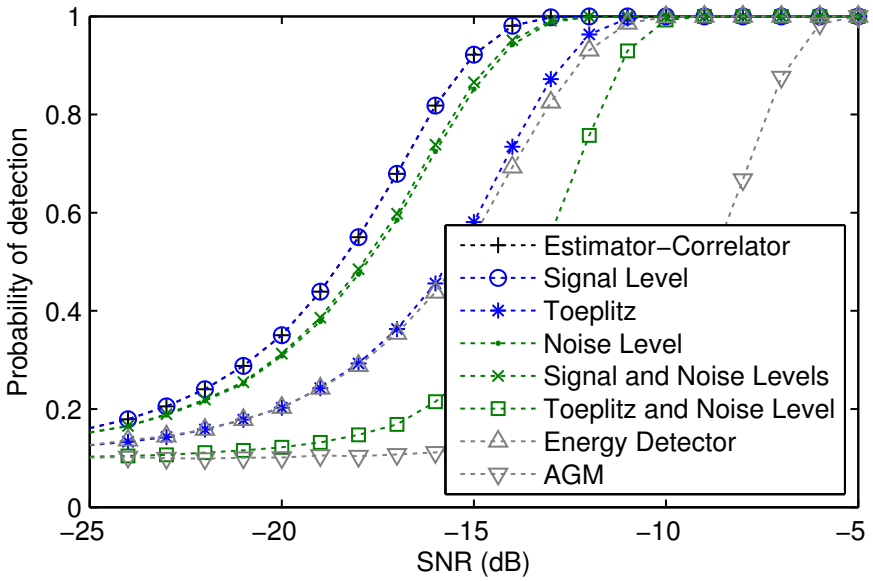

Figure 4.3: detection probability of the wideband spectrum sensing detectors versus average SNR with false-alarm level of 0.1.

sensing detectors derived in Sections 4.2.2 and 4.2.3, along with the energy detector and the AGM detector when the noise variance is perfectly calibrated by the secondary device, at a nominal SNR $= -15$ dB. It is noticed that the detectors with side information on the second-order statistics of the primary systems signal, i.e., $\mathbf{R}_s$ or $\mathbf{R}_0$, provide sensing performance close the estimator-correlator upper-bound. Whereas estimating the signal level incurs no performance loss, the accuracy in estimating the noise variance is reflected in maximum loss of 0.1 points in detection probability along a wide range of false-alarm probability constraints. It is further observed that the Toeplitz detector shows a slight gain with respect to the energy detector, because it exploits the correlation lags present in the off-diagonal of the sample correlation matrix, in addition to the energy computation of the main diagonal. Finally, as expected, the AGM detector's ROC curve is below that of the Toeplitz and noise variance detector, since the test statistic is less informative, e.g., does not exploit the Toeplitz structure of $\mathbf{R}_s$. Figure 4.3 plots the detection probability of the wideband GLRT spectrum sensing detectors versus average nominal SNR, at a fixed false-alarm level of 0.1. For a detection probability requirement of 0.9, it seen that the sensitivity of the detectors with known signal statistics is approximately 3 dB above the Toeplitz and energy detectors, and up to 10 dB with respect to the AGM detector.

In this work, the noise uncertainty is modeled as the uniform distribution $\sigma^2/\delta \leq \hat{\sigma}^2 \leq \delta\sigma^2$, where $\hat{\sigma}^2$ is the prior information on the noise variance and $u = 10\log_{10}\delta$ is the uncertainty level in dB. In the sequel, $u = 0.1$ dB. The ROC at SNR $= -15$ dB, and the detection probability with false-alarm level $\epsilon_0 = 0.1$ of the GLRT spectrum sensing detectors are drawn in Figure 4.4 and Figure 4.5, respectively. A simple comparison between Figure 4.2 and Figure 4.4, and Figure 4.3 and 4.5, respectively, outlines that the test statistics

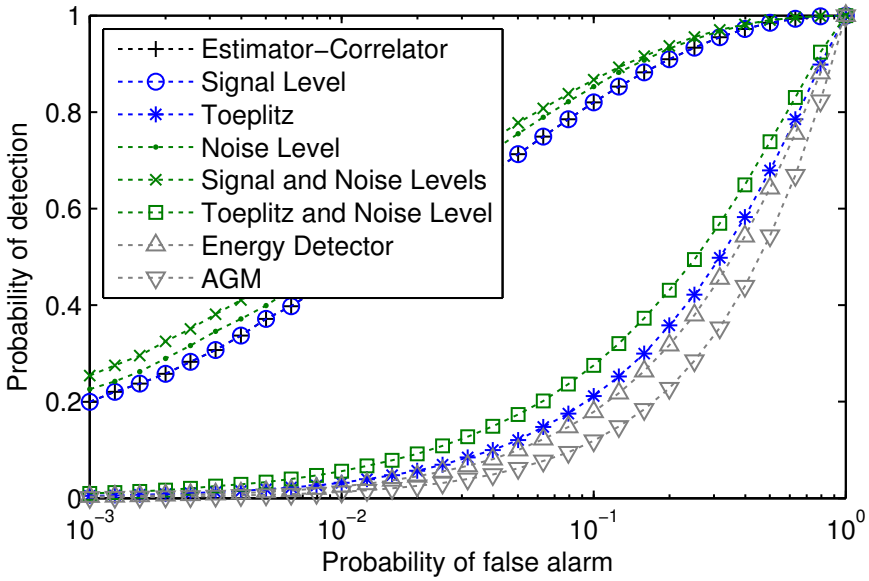

Figure 4.4: ROC of the wideband spectrum sensing detectors at SNR of -15 dB with noise uncertainty of 0.1 dB.

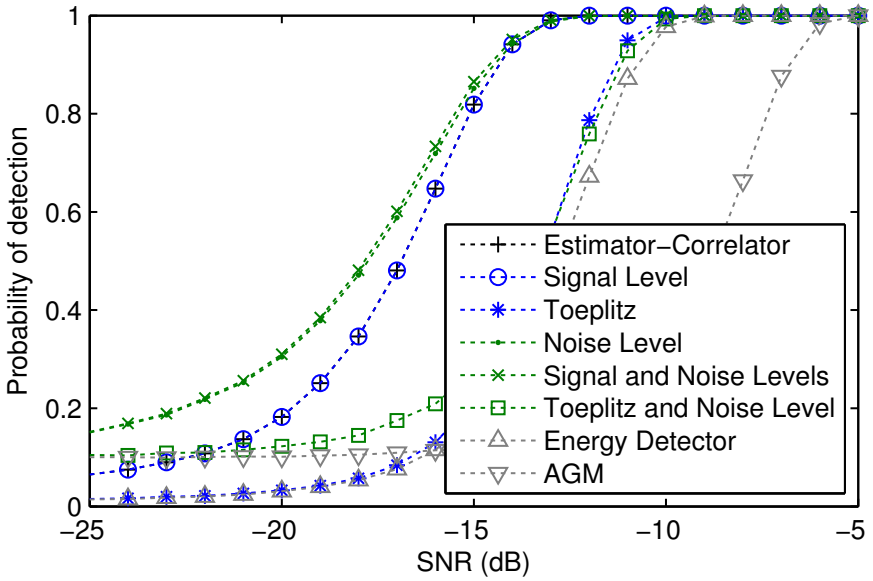

Figure 4.5: detection probability of the wideband spectrum sensing detectors versus average SNR with false-alarm level of 0.1 and noise uncertainty of 0.1 dB.

that assume perfect noise variance side information suffer from performance degradation, whereas the rest of the detectors which incorporate noise variance ML estimation remain unaltered. This is a well-known result [TS08], and becomes the main motivation for the spectrum sensing detectors derived in Section 4.2.3. As appreciated in both figures, the performance curves of the estimator-correlator and the signal level detectors are displaced below the signal level and noise variance, and noise variance detectors. Likewise, the Toeplitz and energy detectors incur a penalty of approximately 2-3 dB of SNR sensibility, as seen by comparing Figure 4.3 and Figure 4.5.

The ROC interpretation of the wideband GLRT spectrum sensing detectors is now studied with fixed SNR $= -15$ dB, in two opposite conditions of primary systems' relative occupancy. As depicted in Figure 4.6, the relative occupancies of 12.5%, and 62.5% have been considered. By comparing both situations, it can be concluded that for a fixed noise and signal powers, spectrum sensing is a more challenging task when the primary signal is more spread over the sensed bandwidth. This effect is more remarkable for the detectors that exploit temporal correlation (e.g., the signal level, noise level, signal and noise levels, and toeplitz and noise level detectors) because temporal correlation decays with frequency occupancy.

Finally, the performance of the GLRT spectrum sensing detectors is evaluated for multi-frequency systems. A cognitive radio network with primary systems based on the DVB-T standard in the 2k-mode is considered in an example with $Q = 8$ channels when sensing an arbitrary channel. The secondary users take $N = 32$ samples with a sampling depth of $M = 2N$. For comparison reasons, the nuisance estimator-correlator, i.e., the test statistic (4.31) with perfect side information on $(P_1, \ldots, P_Q)$, is also added. The SNR at the q-th channel is defined as $\text{SNR}_q \doteq \frac{P_q}{\sigma^2}$, with $\sum_{q=1}^{Q} \text{SNR}_q = \text{SNR}$. On the one hand, the sensing performance of the multi-frequency GLRT spectrum sensing detectors derived in Section 4.2.4 are analyzed as follows. The ROC at $\text{SNR}_q = -15$ dB is analyzed in Figure 4.7, whereas the detection probability versus average SNR is shown in Figure 4.8. It can be highlighted that, analogous to the wideband detectors, the side information on the normalized correlation matrices $\mathbf{R}_q$ is the most informative statistic on the primary signal as the multi-frequency detector (4.31) incurs roughly no sensing loss in estimating the signal levels when the noise variance is perfectly known. However, the degradation of the multi-frequency detector due to noise variance estimation can be clearly appreciated in both figures. Whereas in terms of sensitivity the performance loss is roughly only 1-2 dB in SNR, the ROC for very restrictive false-alarm levels incurs a large penalty. The main reason for this last appreciation is that the false-alarm probability depends on the remaining frequency contributions, whose estimates become highly sensitive to the noise variance computation. It is further shown in Figure 4.9 how the kernels actuate over the periodogram of the observations in the multi-frequency detector when sensing the 3rd band with $P_1 = P_3 = P_7$, $P_2 = P_4 = P_6 = P_8 = 0$ and P_5 with slightly more power As it can

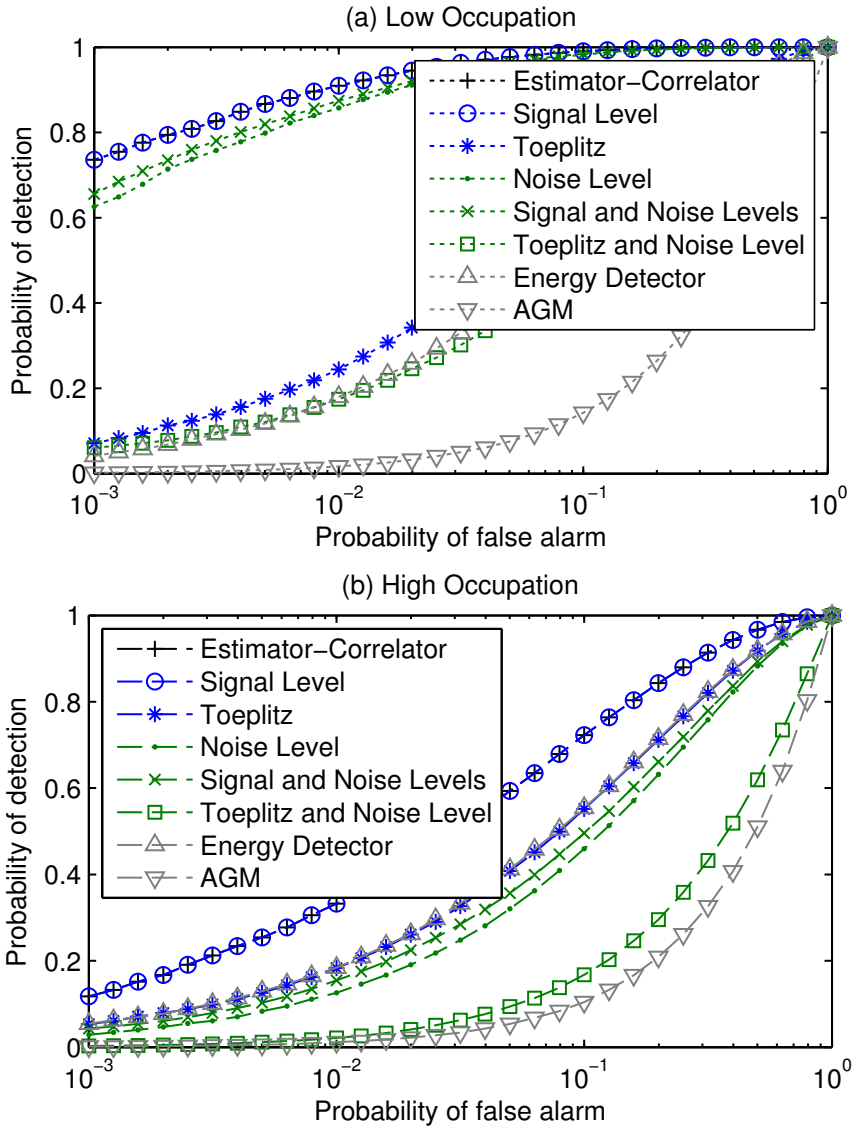

Figure 4.6: ROC of the wideband spectrum sensing detectors at SNR of -15 dB with low (12.5%) and high (62.5%) primary systems relative occupancy.

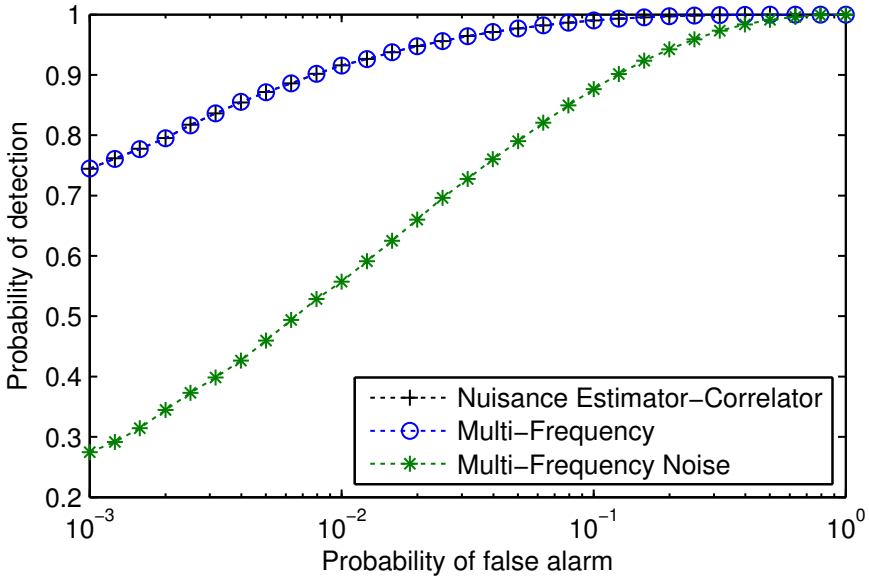

Figure 4.7: ROC of the multi-frequency spectrum sensing detectors at SNR per carrier of -15 dB.

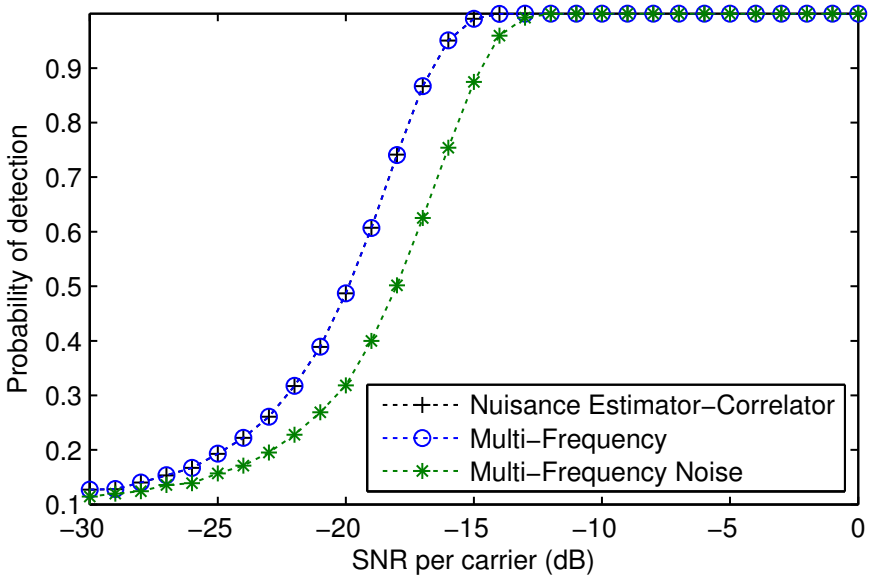

Figure 4.8: detection probability of the multi-frequency spectrum sensing detectors versus average SNR per carrier with false-alarm level of 0.1.

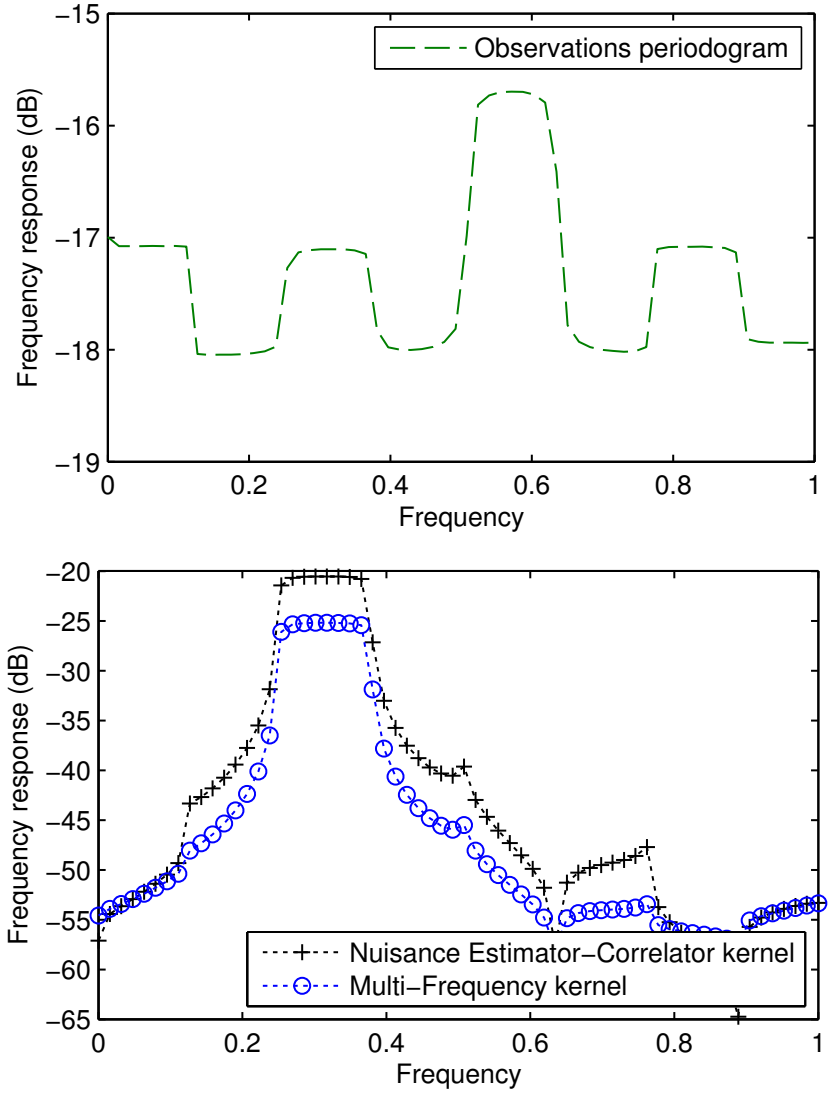

Figure 4.9: Frequency-domain interpretation of the kernels in multi-frequency spectrum sensing detectors at average SNR per carrier of -15 dB.

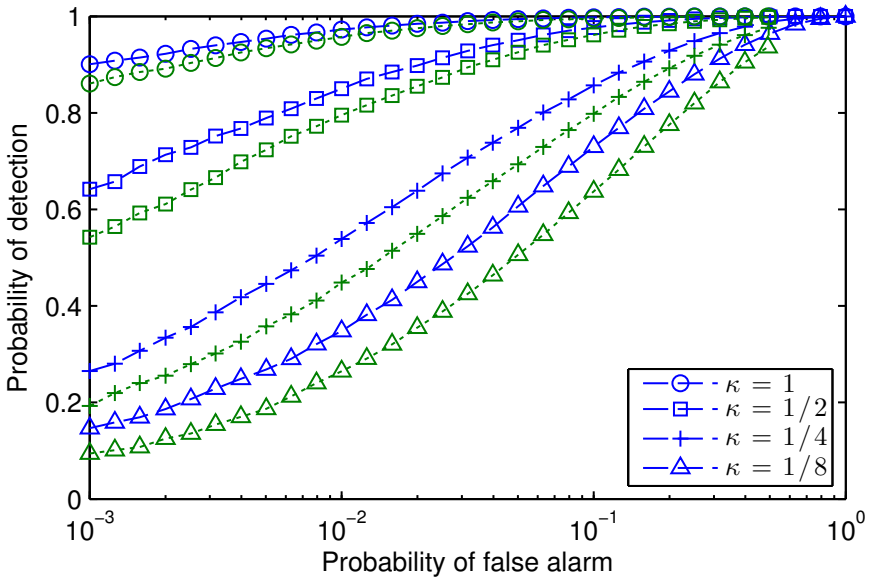

Figure 4.10: ROC of the estimator-correlator (dashed blue lines) and the MF-GLRT (dotted green lines) in a wideband scenario with occupancy of 1/8 and SNR of -12.5 dB.

be appreciated, the kernel $\mathcal{K}_{7,q}(\omega, \mathcal{M})$ at $q = 3$ keeps the spectral shape of the estimator-correlator detector with a small shift and scaling. Both kernels show how they are affected by the spectral information outside the sensing frequency, because they are incorporated in $\phi_\nu(\omega)$ of $\mathcal{K}_{7,q}(\omega, \mathcal{M})$, along with the additive thermal noise. Hence, as an example, it is observed the effort of the kernels in diminishing the contribution of the signals around ω_5 and ω_7 while augmenting the focus on the detected channel, i.e., on ω_3.

4.4.2 Nonuniform Correlation-Matching Estimates

Finally, simulation results are provided to support the novel nonuniform correlation-matching approach in terms of spectrum sensing detection in multi-frequency cognitive radio. Specifically, a cognitive radio network with primary systems based on the DVB-T standard in the 2k-mode with $Q = 8$ bands is considered, when sensing an arbitrary band. The size of the observation before nonuniform sampling is set to $N = 32$ samples. In order to strictly focus on the performance behavior due to nonuniform sampling and remove the effect of insufficient data records, $M(\kappa) = 2N\kappa^{-1}$, so that the size of the nonuniformly sampled observations database $\mathbf{X}$ is $2N$ for any sampling density. In other words, small κ, the cognitive radio takes samples from a larger period of time to preserve the total number of available samples. The average occupancy of the system is $\kappa_0 = 1/8$. For comparison purposes, the estimator-correlator detector [Kay98a], i.e., the test statistic (4.50) with perfectly known signal and noise power levels $(P_1, \ldots, P_Q, \sigma^2)$, is also included.

Firstly, the ROC of the MF-GLRT (4.50) and the estimator-correlator are presented in Figure 4.10 for several sampling densities at an average SNR of SNR $= -12.5$ dB. It is ob-

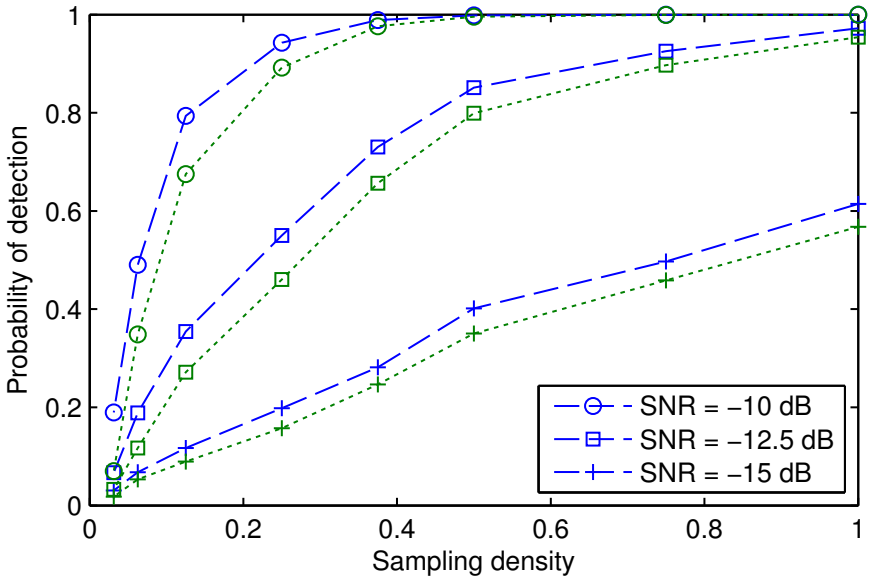

Figure 4.11: detection probability performance of the estimator-correlator (dashed blue lines) and the MF-GLRT (dotted green lines) versus sampling density, with occupancy of 1/8 and false-alarm level of 0.01.

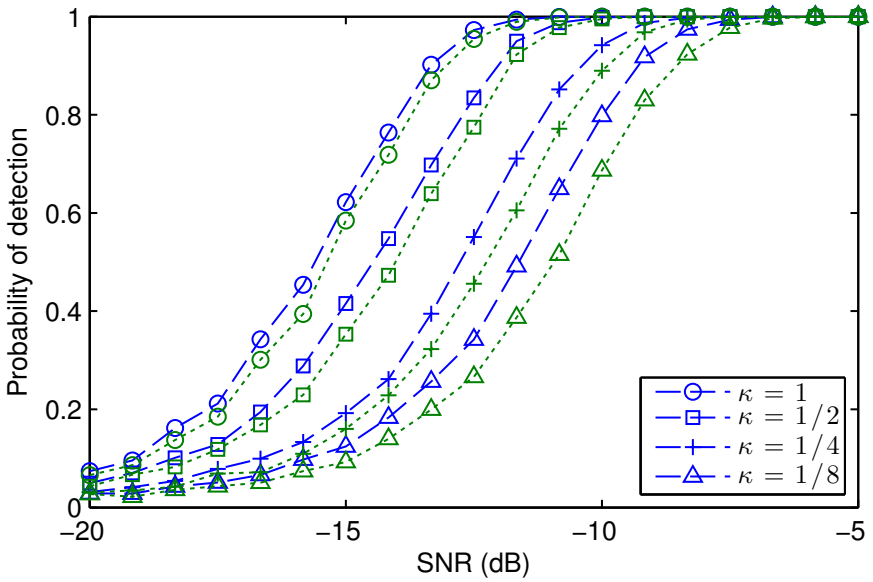

Figure 4.12: detection probability performance of the estimator-correlator (dashed blue lines) and the MF-GLRT (dotted green lines) versus average SNR, with occupancy of 1/8, and false-alarm level of 0.01.

served that the degradation in detection probability incurred by the estimates of the signal and noise power levels (4.48) is small when compared to the performance bound established by the estimator-correlator, for a wide range of false-alarm levels. As the sampling density decreases (smaller κ), the ROC curves of both detectors suffer from degradation in terms of detection probability, mainly for very restrictive false-alarm levels.

Secondly, Figure 4.11 plots the behavior of the MF-GLRT (4.50) and the estimator-correlator versus κ, in terms of detection probability for several average SNR values. It is deduced that for a fixed detection probability target, lower sampling densities (smaller κ) are allowed as the average SNR increases. As an example, for an average SNR of SNR ≥ -10 dB, it is argued that for a false-alarm level of $\epsilon_0 = 0.01$, the MF-GLRT offers outstanding detection probability for a wide range of sampling densities , i.e., it provides probabilities of detection near to 1 for $\kappa \geq \kappa_0$.

Finally, Figure 4.12 depicts the detection probability versus average SNR curves of the MF-GLRT (4.50) and the estimator-correlator for several sampling densities, for a false-alarm level of $\epsilon_0 = 0.01$. While supporting the former interpretations, it can be appreciated from Figure 4.10 and Figure 4.12 that the performance loss incurred by (4.48) increases with smaller κ, as the coupling between adjacent bands of the system matrix in (4.48) is larger as κ becomes smaller. Moreover, it is seen that the performance loss incurred by the MF-GLRT (4.50) is at most 1 dB in SNR when compared to the estimator-correlator, even for a sampling density of the order of the primary system occupancy, i.e., for $\kappa = \kappa_0$.

4.5 Conclusions

In this Chapter, the problem of spectrum sensing in wideband cognitive radios has been investigated. Under the low SNR assumption, a unified framework based on the frequency-domain asymptotic interpretation of the optimal GLRT spectrum sensing detectors has been derived. This unified framework uniquely consists of a kernel inherent to the detector, and the periodogram of the observations. The corresponding kernels for a variety of scenarios of practical interest by obtaining closed-form ML estimates of the unknown parameters has been further obtained, including the signal level and noise variance, and multi-frequency systems. Theoretical interpretations and simulation results show that the primary signal's second-order statistics constitute the most informative statistics for detection. It has been observed that while noise variance estimation guarantees robustness in front of noise uncertainty, the detection kernel takes advantage of the spectral information contained over all the sensed bandwidth.

Furthermore, a novel nonuniform correlation-matching approach and its application to spectrum sensing of wideband sparse signals has been investigated. The closed-form expression of the signal power level estimate has provided physical insight on the problem. The optimal MF-GLRT based on the nonuniform correlation-matching estimates of

multi-frequency signal and noise power levels has been derived. Numerical results have assessed the performance of the proposed technique, which remains tight to the estimator-correlator bound for a wide range of sampling densitiy, SNR, and false-alarm level.

Chapter 5

Asymptotic Performance in Primary Signal Detection

5.1 Introduction

The theoretical characterization of the primary signal detection performance is an important problem in order to predict and benchmark the behavior of the detectors. In this Chapter, the theoretical derivation of the false-alarm and missed-detection probabilities is addressed in order to predict their behavior within the detection parameters, e.g., the signal-to-noise ratio (SNR) the observation size, the correlation level of the primary signal, the noise uncertainty, and the sampling density.

5.1.1 Stein's Lemma

Statistical signal detection is evaluated by the associated error probabilities, i.e., the false-alarm and the missed-detection probabilities. Finding closed-form expressions of the error probabilities is impractical in many situations because the statistical characterization of the detection test is of high complexity. Nonetheless, the asymptotic properties of the statistics are sometimes useful to characterize the behavior and to obtain performance bounds. The Stein's lemma [CT91] is a fundamental result that provides the asymptotic behavior

$$S + W \longrightarrow \boxed{\text{Nonuniform Sampling}} \longrightarrow X \longrightarrow \boxed{\text{Signal Detection}} \longrightarrow \lim_{N\to\infty} T(X)$$

Figure 5.1: Asymptotic performance in primary signal detection problem.

of the error probabilities associated to detection problems when the number of observations grows to infinity. More specifically, it shows that the error probabilities decay exponentially with the number of observations [Tun05], hence playing the role of error exponent as in coding theory. The potentials of the Stein's lemma and the Kullback-Leibler divergence (KLD) as a measure of distance have been recently explored in a wide variety of information theory and communication problems, from multiple-input multiple-output (MIMO) radar [TLWP09] to sensor networks [MT08]. In the field of cognitive radio, the Stein's lemma has been employed in asymptotic performance analysis in collaborative spectrum sensing [RACV11] and quadratic likelihood detection [ZLLZ10]. Also, the lemma has been recently applied in state-of-the-art mathematical tools such as random matrix theory [BDMN11], or sparse principal components analysis [DJSS10].

In this thesis, the error exponents for the energy detector and the estimator-correlator are derived. Both the energy detector and the estimator-correlator are optimal likelihood ratio tests (LRTs) under the Neyman-Pearson criterion and have a central role in spectrum sensing for cognitive radio [GS08a]. It is shown that the error exponents of the false-alarm and missed-detection probabilities in detection depend on the observation size and the second-order statistics of the problem, i.e., the SNR. The main focus is to establish an asymptotic relation between the observation size and the SNR for a fixed error probability. In general, it is shown that the observation size scales as the inverse of a monotonically increasing function of the SNR.

5.1.2 Sampling Walls in Nonuniform Sampling

This thesis shows the existence of sampling walls in detection of wideband signals from Bernoulli nonuniform sampling in the presence of noise uncertainty. A sampling wall is defined as the sampling density below which the target error probabilities, i.e., the missed-detection and false-alarm probabilities, cannot be guaranteed at a given SNR regardless the number of acquired samples. The Bernoulli nonuniform sampling is adopted because it exhibits good tradeoff properties between complexity and performance. As shown in Chapter 3, the Bernoulli nonuniform sampling suffers from noise enhancement, which translates into a whitening effect in the correlation of the primary signal. Contrarily to the existing literature, the signal detection problem is addressed without having to reconstruct neither the signal nor its spectrum. More specifically, the optimal low SNR detector is formulated as a generalized likelihood ratio test (GLRT) to exploit the available side information of the problem, i.e., the noise variance, the sampling density and the primary signal autocorrelation. As shown in Chapter 4, the optimal GLRT signal detector in the low SNR regime strictly depends on the normalized correlation matrix of the primary signal, which by the Bernoulli nonuniform sampling equivalence depends on the sampling density.

Therefore, by deriving the asymptotic performance of the GLRT in the presence of noise

uncertainty, explicit expressions for sampling walls are obtained as a function of the primary signal occupancy, the SNR and the noise uncertainty. Furthermore, a novel noise uncertainty outage model is introduced, and SNR walls are derived as a function of the sampling density and the primary signal occupancy. In the end, it is shown that, when the SNR is comprised within the *sampling walls region*, closed-form expressions of *sampling walls*, denoted as κ_{wall}, are derived. Hence, defining the sampling density κ as the rate between Bernoulli nonuniform sampling and Nyquist rate uniform sampling, it is shown that signal detection is not feasible in the sense that the target missed-detection and false-alarm probabilities cannot be guaranteed regardless the number of samples as far as $\kappa < \kappa_{\text{wall}}$.

5.1.3 Chapter Organization

The rest of the Chapter is organized as follows. In Section 5.2, a first approach to asymptotic performance analysis is addressed with the Stein's lemma. One of the most important results derived in this thesis is presented in Section 5.3: nonuniform sampling walls. Numerical results are reported in Section 5.4, and Section 5.5 concludes the Chapter.

5.2 Stein's Lemma

In this Section, it is shown that the observation size scales as the inverse of a monotonically increasing function of the SNR, i.e.,

$$N \propto \frac{1}{f(\text{SNR})}. \tag{5.1}$$

The main contribution is the derivation of simple, closed-form expressions of $f(\text{SNR})$ that permit the evaluation of the main factors yielded in the signal detection task. In what follows, Section 5.2.1 describes the mathematical tools required for the derivation of $f(\text{SNR})$, whereas the specific application of the lemma to the energy detector, estimator-correlator, nonuniform sampling, diversity and coherent detection are carried out in Sections 5.2.2–5.2.6, respectively.

5.2.1 The Stein's Lemma

In a general framework of binary hypothesis testing problems based on block signal processing, the set of i.i.d. vector observations $\mathbf{X} \doteq (\mathbf{x}_1, \ldots, \mathbf{x}_M)$ is distributed according to the probability density function (PDF) $p(\mathbf{x})$. Each vector $\mathbf{x}_m$ is composed by N consecutive samples, being N the observation size large enough to arise cross-correlation between samples and guarantee independence between samples within two separate vectors. Despite the total available amount of samples is MN, in this Section it is assumed that N is a design parameter, while M can be arbitrarily large. The PDF $p(\mathbf{x})$ takes one of the

following forms under the two hypotheses

$$\mathcal{H}_0 \quad : \quad p(\mathbf{x}) = p_0(\mathbf{x}) \tag{5.2a}$$

$$\mathcal{H}_1 \quad : \quad p(\mathbf{x}) = p_1(\mathbf{x}). \tag{5.2b}$$

The detection problem consists of designing a decision function or test $T(\mathbf{x})$ whose output implies accepting either $\mathcal{H}_0$ or $\mathcal{H}_1$, as a function of the observations and the statistics of the problem. The test is specified by the detection sets $\mathcal{T}_0$ and its complementary $\mathcal{T}_1$ over which $\mathcal{H}_0$ or $\mathcal{H}_1$ is decided, respectively. Given the boundary which defines the aforementioned sets, the error probabilities associated to the test $T(\mathbf{x})$ are defined as

$$P_{\text{FA}} \quad = \quad \mathbb{P}\left[T(\mathbf{X}) \in \mathcal{T}_1 \mid \mathcal{H}_0\right] \tag{5.3a}$$

$$P_{\text{MD}} \quad = \quad \mathbb{P}\left[T(\mathbf{X}) \in \mathcal{T}_0 \mid \mathcal{H}_1\right]. \tag{5.3b}$$

In communications problems, the hypothesis $\mathcal{H}_0$ denotes the only noise situation (i.e., $\mathbf{x} = \mathbf{w}$), whereas $\mathcal{H}_1$ denotes the signal plus noise situation (i.e., $\mathbf{x} = \mathbf{s} + \mathbf{w}$). Therefore, P_{FA} is called the false-alarm probability, and P_{MD} the missed-detection probability. In cognitive radio, P_{FA} is important to guarantee opportunistic communication for the secondary users, while P_{MD} protects the primary systems. According to the Neyman-Pearson criterion, the LRT defined as

$$L(\mathbf{X}) = \frac{p_1(\mathbf{X})}{p_0(\mathbf{X})} \geq \gamma \tag{5.4}$$

provides optimal error probabilities pair $(P_{\text{FA}}, P_{\text{MD}})$ in the sense that for one fixed error probability (e.g., the false-alarm probability), any other test will provide a higher probability pair (e.g., higher missed-detection probability). The LRT (5.4) makes decisions by comparing to a threshold γ, which defines the decision sets as $\mathcal{T}_0 = \{L(\mathbf{X}) < \gamma\}$, and $\mathcal{T}_1 = \{L(\mathbf{X}) \geq \gamma\}$.

The Stein's lemma may be expressed in two versions, which are summarized in the following lemmas.

Lemma 5.1 (False-alarm probability Stein's lemma) *Consider the binary hypothesis testing problem* (5.2) *and the likelihood ratio test* (5.4)*. For a fixed missed-detection probability $P_{\text{MD}} \leq \epsilon_1$, the false-alarm probability asymptotically behaves as*

$$\lim_{M \to \infty} \frac{1}{M} \log P_{\text{FA}} = -\mathbb{D}(\mathcal{H}_1 \| \mathcal{H}_0) \tag{5.5}$$

where $\mathbb{D}(\mathcal{H}_1 \| \mathcal{H}_0)$ evaluates the KLD given by

$$\mathbb{D}(\mathcal{H}_1 \| \mathcal{H}_0) = \int p_1(\mathbf{x}) \log \frac{p_1(\mathbf{x})}{p_0(\mathbf{x})} d\mathbf{x}. \tag{5.6}$$

Lemma 5.2 (Missed-detection probability Stein's lemma) *Consider the binary hypothesis test-*

ing problem (5.2) *and the likelihood ratio test* (5.4)*. For a fixed false-alarm probability* $P_{\mathrm{FA}} \leq \epsilon_0$*, the missed-detection probability asymptotically behaves as*

$$\lim_{M\to\infty} \frac{1}{M} \log P_{\mathrm{MD}} = -\mathbb{D}(\mathcal{H}_0 \| \mathcal{H}_1). \tag{5.7}$$

A direct consequence of Lemmas 1 and 2 is that both error probabilities decay, as M grows to infinity, exponentially with respect to each associated KLD, i.e.,

$$P_{\mathrm{FA}} \approx u(M) e^{-M\mathbb{D}(\mathcal{H}_1 \| \mathcal{H}_0)} \tag{5.8a}$$

$$P_{\mathrm{MD}} \approx v(M) e^{-M\mathbb{D}(\mathcal{H}_0 \| \mathcal{H}_1)}, \tag{5.8b}$$

where $u(M)$ and $v(M)$ are slow-varying functions compared to the exponential, such that

$$\lim_{M\to\infty} \frac{1}{M} \log u(M) = \lim_{M\to\infty} \frac{1}{M} \log v(M) = 0. \tag{5.9}$$

Therefore, given a number of observations, the detection performance exclusively depends on the KLD between hypotheses, which in the sequel it is shown that it is related to the observation size N and the second-order statistics of the problem. Let $p_0(\mathbf{x}) = \mathcal{CN}(\mathbf{0}, \mathbf{R}_0)$ and $p_1(\mathbf{x}) = \mathcal{CN}(\mathbf{0}, \mathbf{R}_1)$. The KLD (5.6) is given by

$$\mathbb{D}(\mathcal{H}_1 \| \mathcal{H}_0) = \mathrm{tr}(\mathbf{R}_0^{-1}\mathbf{R}_1) + (\mathbf{m}_1 - \mathbf{m}_0)^H \mathbf{R}_0^{-1}(\mathbf{m}_1 - \mathbf{m}_0) - \log \frac{\det(\mathbf{R}_1)}{\det(\mathbf{R}_0)} - N. \tag{5.10}$$

From (5.10) it follows that under the Gaussian assumption the pseudo-distance between two processes depends on four terms: the ratio between the second-order statistics, a distance measures between means, the log-radio of determinants of the covariance matrices, and a penalty term of dimensionality.

5.2.2 Energy Detector

Though its simplicity, the energy detector is a low-complexity and well-studied test that has been adopted in recent standards [CCBSS05] as a fast-sensing algorithm. The IEEE 802.22 standard defines the sensing requirements for detecting TV white spaces for wireless regional area network (WRAN) devices opportunities. The spectrum sensing defined in the standard is based on two stages: fast and fine sensing. The energy detector is employed in the fast sensing stage as a coarse detector, whereas a more sophisticated detector is used in the fine sensing stage when the fast sensing stage detects the presence of the signal. Hence, the energy detector is still an important statistical test for practical engineering problems as it allows simple formulations to obtain insights on the required observation size and SNR.

The energy detector is the optimal test in the Neyman-Pearson criterion when the pri-

mary signal and noise have Gaussian white statistics, and is given by [Kay98a]

$$T_{\mathrm{ED}}(\mathbf{x}) = \mathrm{tr}(\mathbf{X}\mathbf{X}^H) \geq \tau. \tag{5.11}$$

Under this assumption, the statistics of the observations are

$$p_0(\mathbf{x}) = \mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I}) \tag{5.12a}$$
$$p_1(\mathbf{x}) = \mathcal{CN}(\mathbf{0}, P\mathbf{I} + \sigma^2\mathbf{I}), \tag{5.12b}$$

where P and σ^2 are the signal and noise powers, respectively. As a consequence, the KLD that define the error exponents in P_{FA} and P_{MD} for the energy detector are

$$\mathbb{D}(\mathcal{H}_1\|\mathcal{H}_0) = N\left[\mathrm{SNR} - \log(1+\mathrm{SNR})\right] \tag{5.13a}$$
$$\mathbb{D}(\mathcal{H}_0\|\mathcal{H}_1) = N\left[\log(1+\mathrm{SNR}) - \frac{\mathrm{SNR}}{1+\mathrm{SNR}}\right], \tag{5.13b}$$

respectively, where $\mathrm{SNR} \doteq P/\sigma^2$. Due to the white statistics under both hypotheses, the KLD (5.6) equals N times the KLD of each individual observation. Hence, the error exponents of the energy detector (5.11) grow linearly with the observation size N in the same way it linearly grows with the number of observations M, in the sense that the equivalent total number of available samples is MN. In other words, no correlation needs to be exploited.

The slope of the error probabilities is given by functions of the SNR, which from (5.13), are given by $\mathrm{SNR} - \log(1+\mathrm{SNR})$ and $\log(1+\mathrm{SNR}) - \mathrm{SNR}(1+\mathrm{SNR})^{-1}$. The non-negativity of the error exponents is guaranteed provided that the former functions are nonnegative for all $\mathrm{SNR} \geq 0$. Additionally, they are monotonically increasing functions in SNR, which ensures that in the limit of the high SNR regime the detection is error-free. Another interpretation of (5.13) is how SNR and the observation size scale to preserve the error probabilities in the asymptotic cases of low SNR (i.e., $\mathrm{SNR} \to 0$) and high SNR (i.e., $\mathrm{SNR} \to \infty$) regimes.

False-alarm probability

The error exponent associated to the false-alarm probability admits the following approximations. In the low SNR regime, it can be approximated by the second degree polynomial $\log(1+\mathrm{SNR}) \approx \mathrm{SNR} - \frac{1}{2}\mathrm{SNR}^2$. Therefore, (5.13a) approximates by

$$\mathbb{D}(\mathcal{H}_1\|\mathcal{H}_0) \approx \frac{1}{2}N \cdot \mathrm{SNR}^2. \tag{5.14}$$

This means that the required observation size is inversely proportional to the squared value of the SNR, i.e.,

$$N \propto \frac{1}{\text{SNR}^2} \tag{5.15}$$

to preserve a target false-alarm probability ϵ_0. Contrarily, the approximation $\text{SNR} - \log(1 + \text{SNR}) \approx \text{SNR}$ is valid in the high SNR regime. In that case, the error exponent may be approximated by

$$\mathbb{D}(\mathcal{H}_1 \| \mathcal{H}_0) \approx N \cdot \text{SNR}, \tag{5.16}$$

for which the required observation size is related to the SNR by

$$N \propto \frac{1}{\text{SNR}}. \tag{5.17}$$

This concludes that the energy detector is more sensitive to a change in channel conditions (i.e., SNR) when operating in the low SNR regime rather than in the high SNR regime.

Missed-detection probability

Regarding the error exponent associated to the missed-detection probability, the following two approximations $\log(1 + \text{SNR}) \approx \text{SNR} - \frac{1}{2}\text{SNR}^2$ and $\text{SNR}(1 + \text{SNR})^{-1} \approx \text{SNR} - \text{SNR}^2$ apply in the low SNR regime. Hence, (5.15) holds for P_{MD}, as $\mathbb{D}(\mathcal{H}_0 \| \mathcal{H}_1) \approx \frac{1}{2} N \text{SNR}^2$ as well. In the high SNR regime, the behavior of the error exponent is more conservative with the SNR, because by the approximation $\log(1 + \text{SNR}) - \text{SNR}(1 + \text{SNR})^{-1} \approx \log(\text{SNR})$ as $\text{SNR} \to \infty$ it follows that

$$\mathbb{D}(\mathcal{H}_0 \| \mathcal{H}_1) \approx N \cdot \log(\text{SNR}). \tag{5.18}$$

In other words, the missed-detection probability is more restrictive in the observation size, in the sense that is inversely proportional to the logarithm of the SNR.

$$N \propto \frac{1}{\log(\text{SNR})}. \tag{5.19}$$

As a common factor in, the observation size is always inversely proportional to a monotonically increasing function of the SNR, as claimed by (5.1), which has the closed-form expressions in (5.13)–(5.19).

5.2.3 Estimator-Correlator

The detection of a signal in Gaussian noise describes many real engineering situations, including the energy detector discussed above. In the problem in hand, it is also a valid assumption that the signal to be detected has Gaussian distribution. While it facilitates the analysis, it is reasonable in signal detection problems in low SNR regimes as the Gaussian distribution provides optimum second-order treatment [VV07], and acts as a worst-case

distribution.

Claimed by the KLD, the performance of the detection in terms of error probabilities depends on the distinctness between the two hypotheses (5.2). For zero-mean Gaussian signals, this distinctness is reflected by the cross-correlation between hypotheses. The correlation can be found in the time, frequency, or space domains. In what follows, the temporal correlation of the signal to be detected is exploited when both noise and signal correlation matrices are known.

In this setting, the observations are distributed according to

$$p_0(\mathbf{x}) = \mathcal{CN}(\mathbf{0}, \mathbf{R}_w) \tag{5.20a}$$

$$p_1(\mathbf{x}) = \mathcal{CN}(\mathbf{0}, \mathbf{R}_s + \mathbf{R}_w), \tag{5.20b}$$

under $\mathcal{H}_0$ and $\mathcal{H}_1$, respectively; where $\mathbf{R}_s$ and $\mathbf{R}_w$ are the correlation matrices of the signal and noise, respectively, defined as $\mathbf{R}_s \doteq \mathbb{E}[\mathbf{s}\mathbf{s}^H]$ and $\mathbf{R}_w \doteq \mathbb{E}[\mathbf{w}\mathbf{w}^H]$. The optimal test in the Neyman-Pearson criterion is the estimator-correlator, given by [Kay98a]

$$T_{\mathrm{EC}}(\mathbf{x}) = \mathrm{tr}\left[\mathbf{R}_w^{-1}\mathbf{R}_s \left(\mathbf{R}_s + \mathbf{R}_w\right)^{-1} \mathbf{X}\mathbf{X}^H\right] \geq \tau. \tag{5.21}$$

From (5.10), it follows that the KLD that define the error exponents in P_{FA} and P_{MD} for the estimator-correlator are

$$\mathbb{D}(\mathcal{H}_1 \| \mathcal{H}_0) = \mathrm{tr}(\mathbf{SNR}) - \log\det\left(\mathbf{I} + \mathbf{SNR}\right) \tag{5.22a}$$

$$\mathbb{D}(\mathcal{H}_0 \| \mathcal{H}_1) = \log\det\left(\mathbf{I} + \mathbf{SNR}\right) - \mathrm{tr}\left[\mathbf{SNR}\left(\mathbf{I} + \mathbf{SNR}\right)^{-1}\right], \tag{5.22b}$$

respectively, where the Woodbury matrix identity has been used. The SNR matrix has been defined as

$$\mathbf{SNR} \doteq \mathbf{R}_w^{-1}\mathbf{R}_s. \tag{5.23}$$

As it can be appreciated, the KLD (5.13) and (5.22) share the property that they exclusively depend on the ratio of signal and noise second-order statistics, as well as the observation size N. For the estimator-correlator, however, the effect of temporal correlation and the effect of the observation size are both inherent in the structure of the SNR matrix **SNR**.

For clarity and comparison purposes, the expressions of the error exponents for asymptotically large observation size, i.e., $N \to \infty$, as also provided. It is a well-known result in statistical signal processing that any autocorrelation matrix asymptotically follows the eigenvalue decomposition (EVD) $\mathbf{R} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H$, where the eigenvectors in $\mathbf{U}$ are related to the Fourier matrix and independent of the process, and the eigenvalues in $\mathbf{\Lambda}$ are equal to the power spectral density (PSD) at the $f_n \doteq n/N$ frequency [Kay98a]. Therefore, **SNR** is asymptotically equivalent to

$$\mathbf{SNR} \to \mathbf{U}\mathbf{\Lambda}_w^{-1}\mathbf{\Lambda}_s\mathbf{U}^H, \tag{5.24}$$

in a weak norm sense. That is, its eigenvalues reflect the SNR at the f_n frequency, denoted by $[\mathbf{\Lambda}_w^{-1}\mathbf{\Lambda}_s]_{nn} = \mathrm{SNR}_n$. The expressions in (5.22) and the result (5.24) allow the following interpretations.

False-alarm probability

For (5.22a), by applying the following Taylor series approximation $\log\det(\mathbf{I} + \mathbf{SNR}) \approx \mathrm{tr}(\mathbf{SNR}) - \frac{1}{2}\mathrm{tr}(\mathbf{SNR}^2)$ when $\mathrm{SNR} \to 0$, it follows that in the low SNR regime,

$$\mathbb{D}(\mathcal{H}_1\|\mathcal{H}_0) \approx \frac{1}{2}\mathrm{tr}(\mathbf{SNR}^2) \to \frac{1}{2}N \cdot \mathrm{SNR}_q^2, \tag{5.25}$$

where the quadratic mean of the SNR has been defined as

$$\mathrm{SNR}_q \doteq \sqrt{\frac{1}{N}\sum_n \mathrm{SNR}_n^2} \tag{5.26}$$

That is, the error exponent associated to P_{FA} is proportional to the squared value of the quadratic mean of the SNR. Furthermore, the observation size scales as

$$N \propto \frac{1}{\mathrm{SNR}_q^2}, \tag{5.27}$$

which shows consistency with (5.15) for white PSD. Contrarily, at the high SNR regime, the KLD can be approximated as

$$\mathbb{D}(\mathcal{H}_1\|\mathcal{H}_0) \approx \mathrm{tr}(\mathbf{SNR}) \to N \cdot \mathrm{SNR}_a, \tag{5.28}$$

i.e., N times the arithmetic mean of the SNR defined as

$$\mathrm{SNR}_a \doteq \frac{1}{N}\sum_n \mathrm{SNR}_n. \tag{5.29}$$

Therefore, at the high SNR regime, the P_{FA} error exponent of the estimator-correlator scales linearly with the observation size and arithmetic mean of SNR

$$N \propto \frac{1}{\mathrm{SNR}_a}. \tag{5.30}$$

Missed-detection probability

Regarding (5.22b), on the one hand, the same low SNR approximation aforementioned can be applied to $\mathbb{D}(\mathcal{H}_0\|\mathcal{H}_1)$, together with the Taylor series approximation of the second term $\mathrm{tr}[\mathbf{SNR}(\mathbf{I} + \mathbf{SNR})^{-1}] \approx \mathrm{tr}(\mathbf{SNR}) - \mathrm{tr}(\mathbf{SNR}^2)$ as $\mathrm{SNR} \to 0$. As a consequence, this leads to the same approximation (5.25), i.e., $\mathbb{D}(\mathcal{H}_0\|\mathcal{H}_1) \approx \frac{1}{2}\mathrm{tr}(\mathbf{SNR}^2)$ and (5.27). This shows that both error probabilities, P_{FA} and P_{MD}, show the same behavior in the low SNR regime in

proportion to the squared of the second-order statistics. On the other hand, as SNR $\to \infty$, the term $\mathrm{tr}\left[\mathbf{SNR}\left(\mathbf{I}+\mathbf{SNR}\right)^{-1}\right] \approx \mathrm{tr}(\mathbf{I})$ can be neglected in front of the first term $\log\det(\mathbf{I}+\mathbf{SNR}) \approx \log\det(\mathbf{SNR})$. Therefore,

$$\mathbb{D}(\mathcal{H}_0\|\mathcal{H}_1) \approx \log\det(\mathbf{SNR}) \doteq N \cdot \log(\mathrm{SNR}_g), \tag{5.31}$$

which coincides with the asymptotic behavior of the channel capacity formula. The geometric mean of the SNR is defined as

$$\mathrm{SNR}_g \doteq \sqrt[n]{\prod_n \mathrm{SNR}_n}. \tag{5.32}$$

This concludes that the error exponent of the missed-detection probability scales with the logarithm of the geometric mean of the SNR, whereas the required observation size becomes proportional to

$$N \propto \frac{1}{\log(\mathrm{SNR}_g)}. \tag{5.33}$$

It is noticed that (5.33) reduces to (5.19) for white statistics.

In conclusion, it has been shown that signal detection based on the estimator-correlator asymptotically (as M and N go to infinity) depends on the quadratic mean of the SNR profile in the low SNR regime (i.e., wideband signals), and on the arithmetic and geometric means in the high SNR regime (i.e., narrowband signals). In general, it has been shown that the required observation size scales as inversely proportional to a function of the SNR (5.1), whose closed-form expressions have been derived.

Frequency-Domain Interpretation

The asymptotic behavior of the error exponents of the estimator-correlator may be expressed in the frequency-domain as a function of the power density functions of the signal and noise processes, denoted by $\phi_s(\omega)$ and $\phi_w(\omega)$. As $N \to \infty$, it follows that (5.22) may be written as

$$\mathbb{D}(\mathcal{H}_1\|\mathcal{H}_0) = \frac{N}{2\pi}\left[\int_0^{2\pi}\frac{\phi_s(\omega)}{\phi_w(\omega)}d\omega - \int_0^{2\pi}\log\frac{\phi_s(\omega)+\phi_w(\omega)}{\phi_w(\omega)}d\omega\right] \tag{5.34a}$$

$$\mathbb{D}(\mathcal{H}_0\|\mathcal{H}_1) = \frac{N}{2\pi}\left[\int_0^{2\pi}\log\frac{\phi_s(\omega)+\phi_w(\omega)}{\phi_w(\omega)}d\omega - \int_0^{2\pi}\frac{\phi_s(\omega)}{\phi_s(\omega)+\phi_w(\omega)}d\omega\right]. \tag{5.34b}$$

The expressions (5.34) reflect that the error exponents linearly scale with N and with a slope that depends on the spectral coherence and spectral SNR of the signal and noise processes of the problem averaged over the whole sensed bandwidth. From the expressions in (5.34), the following considerations can be made. As $N \to \infty$, the error exponents in P_{FA} and P_{MD} scale proportionally to the observation size N, multiplied by the continuous-

frequency average infinitesimal error exponents over frequency. As a result, both exponents depend on the spectral SNR $\phi_s(\omega)/\phi_w(\omega)$ in terms of capacity $\log(1+\phi_s(\omega)/\phi_w(\omega))$ and spectral coherence.

5.2.4 Nonuniform Sampling Signal Detection

As the number of observations increases, the Szëgo's Theorem [Kay98a] applies and it rapidly follows that, in the particular case of Bernoulli nonuniform sampling, the correlation equivalence (3.47) derived in Section 3.3 translates into

$$\mathbf{R}_0 = \kappa^2\mathbf{R}_w + \kappa(1-\kappa)\sigma^2\mathbf{I}, \tag{5.35a}$$

$$\mathbf{R}_1 = \kappa^2\mathbf{R}_s + \kappa^2\mathbf{R}_w + \kappa(1-\kappa)P\mathbf{I} + \kappa(1-\kappa)\sigma^2\mathbf{I} \tag{5.35b}$$

under $\mathcal{H}_0$ and $\mathcal{H}_1$, respectively. In this formulation, the signal power and noise variance are defined as $P = \frac{1}{N}\mathrm{tr}(\mathbf{R}_s)$ and $\sigma^2 = \frac{1}{N}\mathrm{tr}(\mathbf{R}_w)$, respectively, which is consistent with $r_s[0]$ and $r_w[0]$ in the stationary case. From this point, it is seen that if the signal and noise processes are i.d.d., the equivalent covariance matrices at both hypotheses are the ones for the uniform sampling case with a common scaling factor of κ, i.e., $\mathbf{R}_0 = \kappa\sigma^2\mathbf{I}$ and $\mathbf{R}_1 = \kappa(P+\sigma^2)\mathbf{I}$. As the error exponents exclusively depend on **SNR** under the zero-mean Gaussian assumption, the pseudo-distances will not change. This means that for white spectra, the effect of nonuniform sampling is only appreciated with the effective number of samples K, hence having the same asymptotic behavior.

For the general case, it is obtained that the pseudo-distances are exactly as those in the uniformly sampled estimator-correlator (5.22) with equivalent SNR matrix

$$\mathbf{SNR}_\kappa = \left[\kappa\mathbf{R}_w + (1-\kappa)\sigma^2\mathbf{I}\right]^{-1}\left[\kappa\mathbf{R}_s + (1-\kappa)P\mathbf{I}\right]. \tag{5.36}$$

As expected, $\mathbf{SNR}_1 = \mathbf{SNR}$. An interesting problem is relating the general expression of $\mathbf{SNR}_\kappa$ to that of **SNR** as a function of κ. This will allow a derivation of a relationship between SNR, sampling density and number of samples and, maybe, establish SNR walls. By employing the Woodbury matrix identity and after some mathematical manipulations, it follows that the noise autocorrelation matrix can be expressed as follows

$$\mathbf{SNR}_\kappa = \left(\mathbf{I} + \frac{1-\kappa}{\kappa}\mathbf{\Sigma}^{-1}\right)^{-1}\left(\mathbf{SNR} + \mathrm{SNR}\frac{1-\kappa}{\kappa}\mathbf{\Sigma}^{-1}\right), \tag{5.37}$$

where $\mathbf{\Sigma}$ is the normalized noise correlation matrix and the SNR in this scenario is the ratio of signal and noise variances, i.e., $\mathrm{SNR} \doteq \frac{P}{\sigma^2}$. Note that $\mathrm{SNR} = \mathrm{SNR}_a$ if at least one of the processes (noise or signal) is white, which is consistent with the definitions of SNR and SNR_a for the energy detector and estimator-correlator above.

In the important case of white noise, $\mathbf{\Sigma} = \mathbf{I}$, (5.37) reduces to (3.51), which is recalled in

this Chapter:

$$\mathbf{SNR}_{\kappa} = \kappa\mathbf{SNR} + (1-\kappa)\mathrm{SNR}\mathbf{I}. \tag{5.38}$$

In the sequel, the asymptotic behavior of the error exponents for the estimator-correlator in Bernoulli nonuniform sampling is analyzed in the slow SNR and high SNR regimes.

False-alarm probability

The error exponents in P_{FA} for the white noise case are given by (5.22a) and (5.22b) with the equivalent SNR matrix (5.38). Therefore, in the low SNR regime, the following approximation applies

$$\mathbb{D}(\mathcal{H}_0\|\mathcal{H}_1) \approx \frac{1}{2}\kappa^2\mathrm{tr}(\mathbf{SNR}^2) + \frac{1}{2}\mathrm{SNR}^2N(1-\kappa^2). \tag{5.39}$$

The former equation allows the following considerations caused by nonuniform sampling with respect to the uniformly sampled estimator-correlator. Interestingly, the error exponent in this setting is bounded by the extremes cases of $\kappa = 1$, i.e., uniform sampling, and $\kappa = 0$, i.e., no samples. For $\kappa = 1$, the error exponent reduces to $\mathbb{D}(\mathcal{H}_1\|\mathcal{H}_0) \approx \frac{1}{2}\mathrm{tr}(\mathbf{SNR}^2)$, i.e., the error exponent of the estimator-correlator (5.25). On the other hand, for $\kappa = 0$, it is obtained that the error exponent equals $\mathbb{D}(\mathcal{H}_0\|\mathcal{H}_1) \approx \frac{1}{2}\mathrm{SNR}^2N$, which is indeed that of the energy detector (5.14). In conclusion, the error exponent of the false-alarm probability in the low SNR regime in Bernoulli nonuniform sampling is bounded by the best scenario (uniform sampling estimator-correlator) and the worst scenario (energy detector). Also, it follows that the energy detector provides a lower bound in performance of signal detection in nonuniform sampling, as it behaves in the same way as if asymptotically no samples are taken.

In the high SNR regime, the error exponent associated to P_{FA} equals

$$\mathbb{D}(\mathcal{H}_0\|\mathcal{H}_1) = N\mathrm{SNR}, \tag{5.40}$$

that is, the effect of nonuniform sampling is not experienced when SNR is relatively high.

Missed-detection probability

For the error exponent in P_{MD}, the low SNR regime is equivalent to P_{FA}, whereas in the high SNR regime it follows from (5.31) that $\mathbb{D}(\mathcal{H}_0\|\mathcal{H}_1)$ approximates by

$$\log\det(\mathbf{SNR}_{\kappa}) = \log\det\left[\kappa\mathbf{SNR} + (1-\kappa)\mathrm{SNR}\mathbf{I}\right]. \tag{5.41}$$

For $\kappa = 0$, the former expression reduces to $N\log(\mathrm{SNR})$, i.e., (5.18), whereas for $\kappa = 1$ goes back to the uniformly sampled estimator-correlator (5.31).

An interesting problem consists in finding the optimum value of κ that maximizes the error exponents in the error probabilities. For that purpose, in the following the KLD in

the case of white noise are first analyzed. However, as a constrained optimization problem with $0 \leq \kappa \leq 1$, it follows that in any scenario the maximum error exponent is always achieved by $\kappa = 1$, i.e., uniform sampling.

Relation to the Sparsity Level

For the white noise case, let $\mathbf{c} \doteq [\mathrm{SNR}_1, \ldots, \mathrm{SNR}_N]^T$ be the vector containing the eigenvalues of the SNR matrix $\mathbf{SNR}$. In an Q-sparse environment, the vector $\mathbf{c}$ contains only Q nonzero elements. The sparsity level is defined as as

$$\kappa_0 \doteq \frac{1}{N}\|\mathbf{c}\|_{\ell_0}, \tag{5.42a}$$

where the ℓ_0-norm counts the number of nonzero elements of $\mathbf{c}$, i.e., $\|\mathbf{c}\|_{\ell_0} = Q$. The following lower-bounds on the operations on the SNR matrix $\mathbf{SNR}$ are noted. First, for comparison purposes, a fixed average SNR is considered, i.e.,

$$\frac{1}{N}\mathrm{tr}(\mathbf{SNR}) = \frac{1}{N}\sum_{n=1}^{N}\mathrm{SNR}_n = \mathrm{SNR}_a \tag{5.43}$$

for any sparsity pattern of the SNR. As a consequence, for large enough N, it is shown in Appendix C.1 that

$$\mathrm{tr}(\mathbf{SNR}^2) \geq \frac{1}{\kappa_0}\mathrm{tr}^2(\mathbf{SNR}). \tag{5.44}$$

Employing (5.44), it follows that the KLD (5.39) associated to both the false-alarm and missed-detection error exponents in the low SNR regime are given by

$$\mathbb{D}(\mathcal{H}_0\|\mathcal{H}_1) = \mathbb{D}(\mathcal{H}_1\|\mathcal{H}_0) \geq \frac{1}{2}\mathrm{tr}^2(\mathbf{SNR}) \cdot \underbrace{\frac{\kappa_0 + (1-\kappa_0)\kappa^2}{\kappa_0}}_{\doteq \rho}, \tag{5.45}$$

i.e., the KLD in nonuniform sampling is that of uniform sampling, i.e., (5.25), scaled by a factor ρ which is a function of the sampling density and the sparsity level, i.e., the primary signal occupancy.

5.2.5 Diversity in Signal Detection

Diversity has been introduced in communication systems to efficiently combat fading. Based on the premise that the SNR in fading channels is a random variable, transmitting the same signal through L separate fading channels will, with high probability, produce a reliable signal at the output of the channel. Many techniques have been advocated to generate diversity, such as space diversity through multiple antennas, polarization diversity though cross-polarized antennas, frequency diversity through different frequency carriers,

and time diversity through coding and interleaving.

In a real scenario of spectrum sensing for cognitive radio, the diversity is obtained by the secondary user through the use of multiple antennas (all the work by [ZLLZ10, WFHL10, TNKG10, RVS10, RVVLV$^+$11]) or time diversity (if the channel between the primary user and the secondary user has fast fading). In multi-frequency systems, frequency diversity is exploited [FSVR11]. Slow and fast-fading with unknown parameters and prior information is addressed in [FSW10].

In the following, it is assumed that the channel response between the primary users and the secondary users is known by the latter. This knowledge can be obtained by exploiting the transmission system employed by the primary users (e.g., orthogonal frequency-division multiplexing (OFDM)) when this one is broadcasting, or extracting information from the pilot sequences in a communication systems. Furthermore, coherent signal detection provides performance upper bounds on the best achievable error exponent, and the behavior on the diversity L.

A general formulation to include both time correlation and diversity is the following. Consider the reception of L independent signals of the form

$$\mathbf{x}_l[m] = \mathbf{s}_l[m] + \mathbf{w}_l[m], \tag{5.46}$$

where $\mathbf{s}_l[m]$ and $\mathbf{w}_l[m]$ contain N consecutive signal and noise samples at the l-th path, respectively, for the m-th observation, with $l = 1, \ldots, L$ and $m = 1, \ldots, M$. In a diversity scenario, the received signals $\mathbf{s}_l[m]$ are the different signatures of the primary user signal, $\mathbf{u}[m]$, multiplied by the channel gain $h_l[m]$, i.e., $\mathbf{s}_l[m] = h_l[m]\mathbf{u}[m]$. In a slow-fading scenario, $h_l[m] \approx h_l$. The joint time-diversity formulation is encompassed by stacking the L vectors $\mathbf{x}_l[m]$ in a larger vector of dimension NL denoted by $\mathbf{x}[m] = (\mathbf{x}_1[m]^T, \ldots, \mathbf{x}_L[m]^T)^T$. By defining the channel vector as $\mathbf{h} = (h_1, \ldots, h_L)^T$, it follows that

$$\mathbf{x}[m] = \mathbf{h} \otimes \mathbf{s}[m] + \mathbf{w}, \tag{5.47}$$

where $\mathbf{s}[m]$ and $\mathbf{w}[m]$ have been similarly defined. As a result, the following PDF hold for the spectrum sensing problem

$$p_0(\mathbf{x}) = \mathcal{CN}(\mathbf{0}, \mathbf{R}_w) \tag{5.48a}$$

$$p_1(\mathbf{x}) = \mathcal{CN}(\mathbf{0}, \mathbf{h}\mathbf{h}^H \otimes \mathbf{R}_u + \mathbf{R}_w), \tag{5.48b}$$

where $\mathbf{R}_w = \mathbb{E}(\mathbf{w}[m]\mathbf{w}^H[m])$ is the $NL \times NL$ noise correlation matrix, and on the other hand $\mathbf{R}_u = \mathbb{E}(\mathbf{u}[m]\mathbf{u}^H[m])$ is the $N \times N$ primary signal correlation matrix. The Kronecker product $\mathbf{h}\mathbf{h}^H \otimes \mathbf{R}_u$ reflects the diversity in the $\mathbf{h}\mathbf{h}^H$ term and the time correlation in the $\mathbf{R}_u$ term. The KLD associated to the error probabilities for the detection problem (5.48) are given by the error exponents of the estimator-correlator (5.22) with a new associated

$NL \times NL$ SNR matrix given by

$$\mathbf{SNR} = \mathbf{R}_w^{-1} \left(\mathbf{h}\mathbf{h}^H \otimes \mathbf{R}_u \right). \tag{5.49}$$

Obtaining insights on (5.49) is difficult, as it jointly addresses the diversity and the time correlation. In the special case of i.i.d. white noise ($\mathbf{R}_w = \sigma^2 \mathbf{I}$ and defining $\mathbf{R}_u = P\boldsymbol{\Sigma}_u$ and $\text{SNR} = P/\sigma^2$), the following low SNR and high SNR approximations hold for KLD.

False-alarm probability

By taking the asymptotic expression of the error exponent associated to the false-alarm probability in the low SNR regime (5.25) with $\mathbf{SNR} = \text{SNR}(\mathbf{h}\mathbf{h}^H \otimes \boldsymbol{\Sigma}_u)$ it rapidly follows that in a diversity scheme,

$$\mathbb{D}(\mathcal{H}_1 \| \mathcal{H}_0) \approx \frac{1}{2} \text{SNR}^2 \text{tr}(\boldsymbol{\Sigma}_u^2) \|\mathbf{h}\|^4, \tag{5.50}$$

where the property $\text{tr}(\mathbf{A} \otimes \mathbf{B}) = \text{tr}(\mathbf{A}) \cdot \text{tr}(\mathbf{B})$ has been used. The error exponent evolves with the squared value of the SNR, the squared value of the second-order statistics of the primary signal, and the squared value of the squared norm of the diversity channel. For $L = 1$ and $h_1 = 1$, (5.50) reduces to (5.25). On the other hand, the high SNR behavior of the KLD is given by the trace of the SNR matrix **SNR**, i.e.,

$$\mathbb{D}(\mathcal{H}_1 \| \mathcal{H}_0) \approx N \cdot \text{SNR} \cdot \|\mathbf{h}\|^2. \tag{5.51}$$

The linear dependence with N and SNR is clear from (5.51). However, how the number of paths in the diversity channel affects the false-alarm probability is hidden inside the statistics of the $\|\mathbf{h}\|^2$ term. By modeling the channel gains as independent complex Gaussian variables with unit gain, it is shown in Appendix C.2 that the average error exponent (5.51) in Rayleigh fading is given by

$$\mathbb{E}[\mathbb{D}(\mathcal{H}_1 \| \mathcal{H}_0)] = LN \cdot \text{SNR}. \tag{5.52}$$

Therefore, in average, the number of paths in the diversity channel L plays the same role as the observation size in the time-domain N. As a result, the false-alarm probability P_{FA} scales proportional to

$$\log P_{\text{FA}} \propto -LN \log \text{SNR}, \tag{5.53}$$

i.e., at high SNR the false-alarm probability decreases as the LN-th power of the SNR, which corresponds to a slope of $-LN$ in the logarithmic-logarithmic scale.

Missed-detection probability

Regarding the error exponent of the missed-detection probability, from (5.31) it is deduced that in the high SNR regime,

$$\mathbb{D}(\mathcal{H}_0\|\mathcal{H}_1) \approx N \cdot \log\left(\text{SNR} \cdot \|\mathbf{h}\|^2 \cdot \sqrt[N]{\det(\mathbf{\Sigma}_u)}\right). \tag{5.54}$$

In (5.54), the time correlation is evaluated by the geometric mean of the eigenvalues of the primary signal correlation matrix $\mathbf{\Sigma}_u$ as $\sqrt[N]{\det(\mathbf{\Sigma}_u)}$. The contribution of the SNR and the diversity is, as expected, as the product $\text{SNR} \cdot \|\mathbf{h}\|^2$. This means that, contrary to the false-alarm probability, the only parameter that affects to the slope of the probability is N, as is a common factor in the error exponent. That is,

$$P_{\text{MD}} \propto \frac{1}{\left(\text{SNR} \cdot \|\mathbf{h}\|^2 \cdot \sqrt[N]{\det(\mathbf{\Sigma}_u)}\right)^N}. \tag{5.55}$$

Evaluating the average probability with respect to the fading is difficult, in the following study the particularization to $N = 1$ is addressed. From a wider perspective, however, it is seen that the effective SNR grows linearly with the number of diversity paths, as $\|\mathbf{h}\|^2$ is the sum of L squared magnitudes. This means that L does not change the slope of P_{MD} in the logarithmic scale, rather than produces a shift in the SNR.

5.2.6 Coherent Signal Detection

To be more precise with the diversity, in the sequel the observation size N is set to one. Under this assumption, the signal model (5.46) reduces to the scalar expression $x_l[m] = h_l[m]u[m] + w_l[m]$. This expression can be further formulated in matrix notation in the case of slow-fading as

$$\mathbf{x}[m] = \mathbf{h}u[m] + \mathbf{w}[m], \tag{5.56}$$

where now $\mathbf{w}[m] = (w_1[m], \ldots, w_L[m])^T$ and $\mathbf{x}[m] = (x_1[m], \ldots, x_L[m])^T$. As a result, the binary hypotheses problem (5.48) reduces to

$$p_0(\mathbf{x}) = \mathcal{CN}(\mathbf{0}, \mathbf{R}_w) \tag{5.57a}$$

$$p_1(\mathbf{x}) = \mathcal{CN}(\mathbf{0}, P\mathbf{h}\mathbf{h}^H + \mathbf{R}_w), \tag{5.57b}$$

where P is the variance of $u[m]$ and $\mathbf{R}_w = \mathbb{E}(\mathbf{w}[m]\mathbf{w}^H[m])$ is the $L \times L$ noise correlation matrix which exploits the noise correlation among the diversity paths. The KLD associated to the error probabilities for the detection problem (5.57) are given by the error exponents of the estimator-corrector (5.22) with an associated $L \times L$ SNR matrix given by

$$\mathbf{SNR} = \mathbf{R}_w^{-1} P\mathbf{h}\mathbf{h}^H. \tag{5.58}$$

In the special case of i.i.d. noises with equal variance σ^2, the former SNR matrix reduces to $\mathbf{SNR} = \mathrm{SNR} \cdot \mathbf{h}\mathbf{h}^H$, i.e., a matrix with rank equal to one. This derives to the following low SNR and high SNR approximations of the KLD.

False-alarm probability

Either from (5.25) with an equivalent **SNR** as (5.58), or from (5.50) with $N = 1$, it follows that in the low SNR regime the error exponent associated to the false-alarm probability in this scenario becomes

$$\mathbb{D}(\mathcal{H}_1 \| \mathcal{H}_0) \approx \frac{1}{2} \mathrm{SNR}^2 \|\mathbf{h}\|^4. \tag{5.59}$$

Mimicking the Rayleigh scenario as in the former developments, the average error exponent with respect to the fading is derived in Appendix C.3 and is given by

$$\mathbb{E}[\mathbb{D}(\mathcal{H}_1 \| \mathcal{H}_0)] = \frac{1}{2}(L+1) \cdot L \cdot \mathrm{SNR}^2. \tag{5.60}$$

Interestingly, in the low SNR regime the the false-alarm probability is related to L by its squared value. If L is large enough, then the false-alarm probability scales as

$$\log P_{\mathrm{FA}} \propto -2L^2 \cdot \log(\mathrm{SNR}), \tag{5.61}$$

i.e., in a log-log scale, the false-alarm probability is linear with the SNR with slope $-2L^2$. Finally, in the high SNR regime, the following approximation holds

$$\mathbb{D}(\mathcal{H}_1 \| \mathcal{H}_0) \approx \mathrm{SNR} \|\mathbf{h}\|^2. \tag{5.62}$$

In Rayleigh fading, the average of (5.62) is further given by

$$\mathbb{E}[\mathbb{D}(\mathcal{H}_1 \| \mathcal{H}_0)] = L \cdot \mathrm{SNR}. \tag{5.63}$$

Therefore, P_{FA} decays with slope L in the logarithmic scale with SNR. Presumably, this does not happen with the missed-detection probability, which is addressed in the following analysis.

Missed-detection probability

In the low SNR, the KLD of P_{MD} is the same as the KLD of P_{FA}. Hence, the high SNR regime is considered here. From (5.54) with $N = 1$, it follows that

$$\mathbb{D}(\mathcal{H}_0 \| \mathcal{H}_1) \approx \log\left(\mathrm{SNR} \|\mathbf{h}\|^2\right). \tag{5.64}$$

As a direct consequence, the missed-detection probability scales as the inverse of the effective SNR $\mathrm{SNR} \cdot \|\mathbf{h}\|^2$, i.e.,

$$P_{\mathrm{MD}} \propto \frac{1}{\mathrm{SNR} \cdot \|\mathbf{h}\|^2}. \tag{5.65}$$

As a concluding remark, the average missed-detection probability with Rayleigh fading is evaluated in Appendix C.4 and is given as

$$\mathbb{E}[P_{\mathrm{MD}}] = \frac{1}{(L-1)\mathrm{SNR}} \tag{5.66a}$$

for $L > 1$. This is a surprising result which states that, contrary to the false-alarm probability, the missed-detection probability asymptotically as $M \to \infty$ and $\mathrm{SNR} \to \infty$ has an effective SNR given by $(L-1)\mathrm{SNR}$. Therefore, the diversity does not introduce a change in slope in the error probability, rather than only a shift in SNR.

5.3 Nonuniform Sampling Walls

5.3.1 Problem Formulation

This Section considers the problem of detecting the presence of a primary signal in the additive noise model. Let $s(t)$ and $w(t)$ denote the wide-sense stationary (WSS) signal and noise complex analytic processes, respectively. Within the band of interest B, it is assumed that the additive noise is a zero-mean white Gaussian process with known PSD equal to N_0, whereas the signal has known spectral support and unknown power P[1]. The spectral support of $s(t)$ is parameterized by the occupancy κ_0 comprised within $0 < \kappa_0 \leq 1$, which is defined as the ratio between the ℓ_1-norm of the spectral support and B. The detector must infer between the hypothesis $\mathcal{H}_0$ in which the observation process has only noise, and the hypotheses $\mathcal{H}_1$ in which the observation process has both signal and noise.

The detection problem is cast as an hypotheses testing problem based on a set of $N = TB$ nonuniform samples acquired during a sensing period of T seconds. The aim of the hypotheses testing problem is to exploit the side information available in the problem, i.e., the noise variance, the sampling density and the primary signal autocorrelation. To the end of the Chapter, the wideband regime $B \gg 1$ and asymptotically large number of samples $N \gg 1$ are considered.

The sampling pattern ψ_n, for $0 \leq n < N$, is distributed according to a Bernoulli distri-

[1]Because this work focuses on detecting whether the band of interest is totally or partially occupied, a general formulation on $s(t)$ is adopted based on the normalized correlation matrix. This model encompasses practical scenarios such as multiple primary signals $s(t) = \sum_j s_j(t)$, or the exploitation of channel state information (CSI) $s(t) = \sum_j h_j(t) * s_j(t)$.

bution, independent from the signal process, with parameter κ such that

$$\mathbb{P}\left[\psi_n = \frac{1}{\sqrt{\kappa}}\right] = \kappa, \tag{5.67a}$$

$$\mathbb{P}\left[\psi_n = 0\right] = 1 - \kappa, \tag{5.67b}$$

where $0 < \kappa \leq 1$. Therefore, as shown in Chapter 3, the second-order statistics of the nonuniformly sampled vector[2]

$$\mathbf{x} = \mathbf{\Psi}\mathbf{s}, \tag{5.68}$$

exhibit the following equivalence

$$\mathbf{R}_x = \kappa\mathbf{R}_s + (1-\kappa)P\mathbf{I}_N, \tag{5.69}$$

where $\mathbf{R}_x \doteq \mathbb{E}\left[\mathbf{x}\mathbf{x}^H\right]$, $\mathbf{R}_s \doteq \mathbb{E}\left[\mathbf{s}\mathbf{s}^H\right]$, $\mathbf{x} \doteq (x_0, \ldots, x_{N-1})^T$, $\mathbf{s} \doteq (s_0, \ldots, s_{N-1})^T$, $\mathbf{\Psi} \doteq \mathrm{diag}(\psi_0, \ldots, \psi_{N-1})$, and

$$x_n = \psi_n \cdot s_n, \tag{5.70}$$

where s_n are uniformly samples at a rate B. Note that P is the power of $s(t)$. It is interesting that, even in the noise free case, the correlation of a Bernoulli nonuniform sampling signal suffers from a self interference term or noise enhancement term which is proportional to the product between $(1-\kappa)$ and the signal power. This self interference has white statistics, i.e., it only affects to the main diagonal of $\mathbf{R}_s$. It is also important to point out that the signal power is preserved, since $\mathrm{tr}(\mathbf{R}_x) = \mathrm{tr}(\mathbf{R}_s)$ for any value of κ. As a consequence, the nonzero correlation lags suffer from an attenuation factor of κ, decreasing the overall correlation of the Bernoulli nonuniform sampling signal. This is not surprising, as the nonzero samples in $\mathbf{x}$ become more separate in average as κ decreases.

The detector acquires samples during a sensing time of duration T. Because this Chapter focuses on the asymptotic performance of exploiting the side information of the problem, i.e., the noise variance, the sampling density and the primary signal autocorrelation, the sensing time is scaled in order to satisfy a constant average nonzero samples in the observation process equal to M. This parameter will be referred to as the number of effective samples, or simply the number of samples. This is straightforwardly accomplished by setting the sensing time T to satisfy the relation $T\kappa B = M$, where M is a constant independent on κ. Under these assumptions, the signal detection problem reads

$$\mathcal{H}_0 \;:\; \mathbf{x} = \mathbf{\Psi}\mathbf{w}, \tag{5.71a}$$

$$\mathcal{H}_1 \;:\; \mathbf{x} = \mathbf{\Psi}\left(\mathbf{s} + \mathbf{w}\right), \tag{5.71b}$$

[2]Even though (5.68) reminds the compressed sensing formulation, this work adopts a different approach by considering zeros in the non-sampled positions. This allows to tackle the problem as a product of random discrete-time sequences and conduct a second-order statistical analysis.

where the observation size is the nearest integer N to satisfy $M = \kappa N$. Signal detection in wideband regimes is a challenging problem due to the hard low SNR and sparsity conditions. Let

$$\text{SNR} \doteq \frac{P}{N_0 B}. \tag{5.72}$$

Notice that the SNR of the problem is preserved in Bernoulli nonuniform sampling, as the power is preserved for both the signal and noise processes. Clearly, for a fixed signal power, it follows that when the sensing bandwidth increases as $B \gg 1$, the SNR asymptotically approaches to zero. In statistical signal detection, the signal detection problem described in (5.71) is equivalent to consider

$$\mathcal{H}_0 \quad : \quad \mathbf{x} \sim g_0(\mathbf{x}), \tag{5.73a}$$

$$\mathcal{H}_1 \quad : \quad \mathbf{x} \sim g_1(\mathbf{x}), \tag{5.73b}$$

where $g_0(\mathbf{x})$ and $g_1(\mathbf{x})$ denote the Gaussian distributions of the observations under $\mathcal{H}_0$ and $\mathcal{H}_1$, respectively[3].

Let $\mathcal{CN}(\boldsymbol{\mu}, \mathbf{R})$ denote the N-dimensional circular symmetric complex Gaussian distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\mathbf{R}$. Then, $g_0(\mathbf{x}) = \mathcal{CN}(\mathbf{0}, \mathbf{R}_0)$ and $g_1(\mathbf{x}) = \mathcal{CN}(\mathbf{0}, \mathbf{R}_1)$, where the correlation matrices are given, from (5.69) and (5.71), as

$$\mathcal{H}_0 \quad : \quad \mathbf{R}_0 = \sigma^2 \mathbf{I}_N, \tag{5.74a}$$

$$\mathcal{H}_1 \quad : \quad \mathbf{R}_1 = \sigma^2 \mathbf{I}_N + P\boldsymbol{\Sigma}_\kappa, \tag{5.74b}$$

where $\sigma^2 = N_0 W$ is the total noise power, and $\boldsymbol{\Sigma}_\kappa$ is the normalized correlation matrix of $\boldsymbol{\Psi}\mathbf{s}$, given from (5.69) as

$$\boldsymbol{\Sigma}_\kappa \doteq \kappa \boldsymbol{\Sigma}_s + (1 - \kappa)\mathbf{I}_N, \tag{5.75}$$

being $\boldsymbol{\Sigma}_s$ the normalized correlation matrix of $\mathbf{s}$, related to P and $\mathbf{R}_s$ as $\mathbf{R}_s = P\boldsymbol{\Sigma}_s$. The definition of the normalized correlation matrices $\boldsymbol{\Sigma}_\kappa$ and $\boldsymbol{\Sigma}_s$ has been done in analogy to the noise correlation matrix, i.e., all satisfy $\text{tr}(\boldsymbol{\Sigma}_\kappa) = \text{tr}(\boldsymbol{\Sigma}_s) = \text{tr}(\mathbf{I}_N) = N$. The matrix $\boldsymbol{\Sigma}_\kappa$ is a function of the sampling density and the normalized correlation of $s(t)$. In a cognitive radio scenario the detector may have knowledge on the noise power σ^2 and on the normalized statistics of the signal. However, the knowledge on the signal power is rather impractical. For this reason, this work considers signal detection in the general case of unknown P and studies the effect of uncertainty around the knowledge of σ^2.

[3]While under $\mathcal{H}_0$ the Gaussian assumption on the noise process is widely accepted, in general the distribution of the observations under $\mathcal{H}_1$ remains unknown. In this thesis, the properness of the Gaussian assumption of $\mathbf{s}$ is twofold. Firstly, it has been proved that adopting the Gaussian distribution for the signal statistic is valid in the low SNR regime [VV07]. And secondly, the Gaussian distribution is the least favorable distribution in signal detection for a given signal power, as claimed in Appendix C.5. Furthermore, in practical scenarios such as multiple primary signals $s(t) = \sum_j s_j(t)$, the Gaussian assumption holds for each $s_j(t)$ and, hence, for $s(t)$.

5.3.2 Optimal Detector

In this Section, the formulation of the optimal test statistic to solve the signal detection problem (5.71) in the Neyman-Pearson sense is discussed. The optimal test statistic that exploits the available side information of the problem, i.e, the noise variance, the sampling density and the primary signal autocorrelation, and addresses unknown parameters (P) is given by the GLRT. The GLRT consists of the ratio of the PDF of the observations where the unknown parameter is substituted by its maximum likelihood (ML) estimate. In the sequel, the ML estimation of the signal power is first addressed for the case of $B \gg 1$. The expression of the ML estimate is then used to formulate the GLRT. The asymptotic performance of the GLRT is further analyzed for the asymptotic cases of large data records, i.e., when the observation size $N \gg 1$.

The ML estimate of the signal power, derived in Appendix C.6, is given as $B \gg 1$ by

$$\hat{P} = \frac{\mathbf{x}^H \boldsymbol{\Sigma}_\kappa \mathbf{x} - N\sigma^2}{\mathrm{tr}(\boldsymbol{\Sigma}_\kappa^2)}. \tag{5.76}$$

The estimate (5.76) admits the following two interpretations. At first, the term $\mathbf{x}^H \boldsymbol{\Sigma}_\kappa \mathbf{x}$ implements a generalized energy detector in the norm defined by the Bernoulli nonuniform sampling normalized correlation matrix. As the power in the observations includes the noise contribution, the latter is compensated with the term $N\sigma^2$. The power estimation is finally obtained after a normalization factor given by the trace of the squared value of the Bernoulli nonuniform sampling normalized correlation matrix. A second interpretation is understood when arranging the numerator of (5.76) as $\mathrm{tr}[\boldsymbol{\Sigma}_\kappa(\mathbf{x}\mathbf{x}^H - \sigma^2 \mathbf{I}_N)]$, which implements the scalar product between the Bernoulli nonuniform sampling normalized correlation matrix and the matrix $(\mathbf{x}\mathbf{x}^H - \sigma^2 \mathbf{I}_N)$. The latter evaluates solely the contribution of the signal in the second-order statistic of the observations, because the rank-1 matrix $\mathbf{x}\mathbf{x}^H$ behaves as an estimation of the correlation matrix and the noise contribution is balanced by subtracting its correlation matrix, i.e., $\sigma^2 \mathbf{I}_N$.

The GLRT is given by the ratio

$$L(\mathbf{x}) = \frac{\mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I}_N + \hat{P}\boldsymbol{\Sigma}_\kappa)}{\mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I}_N)} \geq \gamma, \tag{5.77}$$

where $\hat{P}$ is given by (5.76), and the threshold γ defines the regions for which $\mathcal{H}_0$ $(L(\mathbf{x}) < \gamma)$ or $\mathcal{H}_1$ $(L(\mathbf{x}) \geq \gamma)$ are decided. The threshold can be set analytically when the test statistic (5.77) admits perfect statistical characterization, which is not the case in most situations except in some special assumptions such as the asymptotic scenarios considered in this work with $B \gg 1$ and $N \gg 1$. Under these premises, the GLRT signal detector is derived in Appendix C.7, and is given as $B \gg 1$ by

$$T(\mathbf{x}) = \mathbf{x}^H \boldsymbol{\Sigma}_\kappa \mathbf{x} \geq \tau, \tag{5.78}$$

where the new threshold can be related to the original threshold by the expression

$$\tau = \sigma^2 \left[N + \sqrt{\mathrm{tr}(\mathbf{\Sigma}_\kappa^2) \log(\gamma)} \right]. \tag{5.79}$$

Interestingly, the expression of the GLRT signal detector (5.78) is simply given by the part of the ML estimate of the signal power which involves the observations. It is also worth noting that, in general, the non-negativity of the numerator of (5.76) is not guaranteed, and therefore (5.76) is not properly an estimate of the signal power. If the constraint $P \in \mathbb{R}_+$ is included as a complimentary Lagrange multiplier in the optimization problem involved in the derivation of (5.76), it will follow that the non-negativity must be forced by $\max(0, \hat{P})$. However, this remains unnecessary when the purpose of (5.76) is signal detection, because the mathematical operations which relate (5.76) and (5.78) incur in a shift and scaling in the computation of the threshold, independent on the observations.

Recall that in this work the GLRT signal detector (5.78) is derived under the wideband assumption $B \gg 1$. Since the wideband regime is equivalent to the low SNR regime, (5.78) puts in remembrance classical results [Kay98a] and several works in the recent literature. One recent method for spectrum sensing in open spectrum communications based on the correlation matching approach has been reported in [PNLRS09]. Although [PNLRS09] adopts a different technique, it is worth noting that (5.78) acts as a correlation matching estimate when the candidate matrix is the Bernoulli nonuniform sampling normalized correlation matrix. This is not surprising, as the correlation matching behaves as an equivalence to ML in the low SNR regime [Por08]. Similarly, the recent work [QZSS11] has shown that exploiting spectral features, i.e., a second-order statistic of the primary signal, is optimal for asymptotically low SNR regimes.

5.3.3 Asymptotic Statistical Characterization

The statistical characterization of GLRTs is usually a hard problem. The GLRT signal detector (5.78) is a quadratic form on the observation vector $\mathbf{x}$. The statistics of a quadratic form is known if and only if the matrix involved in the quadratic form has properties of symmetry and idempotence within the covariance matrix of the observations [GM69]. That is not the case, as in general $\mathbf{\Sigma}_\kappa$ is not idempotent, i.e., $\mathbf{\Sigma}_\kappa^2 \neq \mathbf{\Sigma}_\kappa$, nor it is idempotent within any correlation matrix under the two hypotheses (5.74), i.e., $\mathbf{\Sigma}_\kappa \mathbf{R}_0 \mathbf{\Sigma}_\kappa \neq \mathbf{\Sigma}_\kappa$ and $\mathbf{\Sigma}_\kappa \mathbf{R}_1 \mathbf{\Sigma}_\kappa \neq \mathbf{\Sigma}_\kappa$. Alternatively, the quadratic form (5.78) has the same distribution as the following weighted sum of independent Chi-square random variables with two degrees of freedom [Gra76]

$$T(\mathbf{x}) \sim \sum_{n=1}^{r} \frac{\lambda_n(\mathbf{R}_x \mathbf{\Sigma}_\kappa) w_n}{2}, \tag{5.80}$$

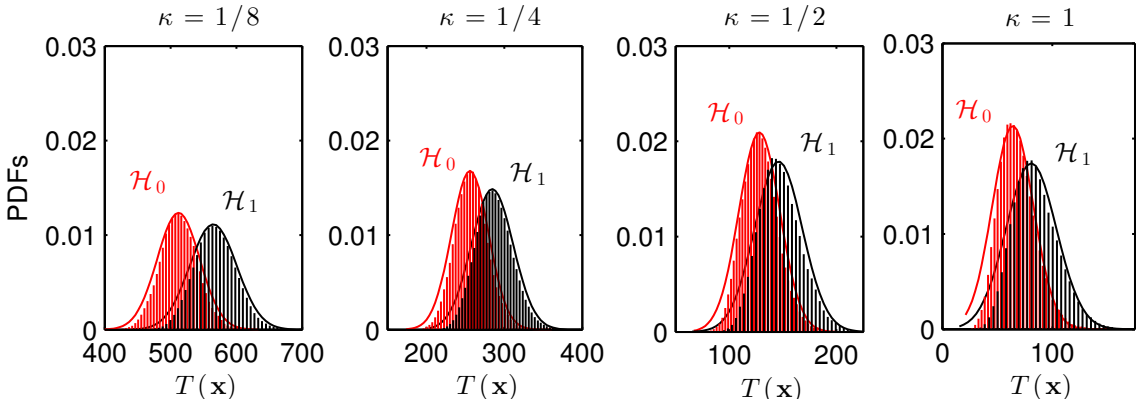

Figure 5.2: Theoretical (lines) and empirical (stems) PDF of the GLRT signal detector under both hypotheses, for several sampling densities of a randomly generated signal with squared correlation coefficient of 8/3, observation size of 64 samples, and SNR of -10 dB.

where $r = \text{rank}(\mathbf{\Sigma}_\kappa)$, $w_n \sim \mathcal{X}_2^2$ and the weights $\lambda_n(\mathbf{R}_x\mathbf{\Sigma}_\kappa)$ denote the n-th largest eigenvalue of the quadratic form, given by

$$\mathcal{H}_0 \;:\; \lambda_n(\mathbf{R}_0\mathbf{\Sigma}_\kappa) = \sigma^2\lambda_n(\mathbf{\Sigma}_\kappa) \tag{5.81a}$$

$$\mathcal{H}_1 \;:\; \lambda_n(\mathbf{R}_1\mathbf{\Sigma}_\kappa) = \sigma^2\lambda_n(\mathbf{\Sigma}_\kappa) + P\lambda_n^2(\mathbf{\Sigma}_\kappa), \tag{5.81b}$$

where $\lambda_n(\mathbf{\Sigma}_\kappa)$ the n-th largest eigenvalue of $\mathbf{\Sigma}_\kappa$. As $N \gg 1$, the former weighted sum of Chi-square random variables can be approximated by Gaussian distributions [AS64] whose means and variance are derived in Appendix C.8. As a result, the GLRT signal detector (5.78) is asymptotically distributed as

$$\mathcal{H}_0 \;:\; T(\mathbf{x}) \sim \mathcal{N}\left(N\sigma^2, 2N\sigma^4\rho\right) \tag{5.82a}$$

$$\mathcal{H}_1 \;:\; T(\mathbf{x}) \sim \mathcal{N}\left(N\sigma^2(1+\text{SNR}\rho), 2\sigma^4\text{tr}(\mathbf{\Sigma}_\kappa^2(\mathbf{I}_N + \text{SNR}\mathbf{\Sigma}_\kappa)^2)\right), \tag{5.82b}$$

where $\mathcal{N}(\mu, \sigma^2)$ denotes the continuous Gaussian probability distribution with mean μ and variance σ^2. The properness of the Gaussian distribution for $T(\mathbf{x})$ is corroborated in Figure 5.2, which shows the theoretical (5.82) and empirical PDF in four sampling density scenarios. As it can be appreciated, the Gaussian distribution is more pessimistic than the true distribution. Hence, the Gaussian assumption on the distribution of the GLRT signal detector (5.78) behaves again as a worst case formulation. Also, it is worth noting that as the sampling density diminishes, the theoretical distribution behaves as better approximation because the Bernoulli nonuniform sampling observations become more uncorrelated. In (5.82), the squared correlation coefficient has been introduced as

$$\rho \doteq \frac{1}{N}\sum_{n=1}^{r}\lambda_n^2(\mathbf{\Sigma}_\kappa). \tag{5.83}$$

This parameter, which can be rewritten as $\rho = (1/N)\mathrm{tr}(\mathbf{\Sigma}_\kappa^2)$, is a measure of dependence, and plays a relevant role in the rest of the Chapter. The parameter ρ strongly depends on the sampling density κ of Bernoulli nonuniform sampling, as well as the structure of the correlation matrix $\mathbf{\Sigma}_s$. By employing the Cauchy-Schwarz inequality $\sum_{n=1}^{r} \lambda_n^2(\mathbf{\Sigma}_\kappa) \geq (1/N)\left(\sum_{n=1}^{r} \lambda_n(\mathbf{\Sigma}_\kappa)\right)^2$ and recalling the normalization $\sum_{n=1}^{r} \lambda_n(\mathbf{\Sigma}_\kappa) = N$, it follows that the squared correlation coefficient (5.83) is lower bounded by $\rho \geq 1$, where the equality $\rho = 1$ is achieved for $\mathbf{\Sigma}_\kappa = \mathbf{I}_N$. It is easy to see from the Bernoulli nonuniform sampling correlation equivalence (5.75) that this situation is asymptotically true as $\kappa \ll 1$. On the other extreme, since ρ computes the sum of the squared value of the eigenvalues of $\mathbf{\Sigma}_\kappa$, the maximum of ρ is achieved when $\mathbf{\Sigma}_\kappa$ has a unique eigenvalue equal to N. This corresponds to the uniform sampling ($\kappa = 1$) fully correlated case (e.g., deterministic signals employed as pilots) in which $\mathbf{\Sigma}_\kappa$ is a rank-1. In such a scenario, $\rho = N$. Therefore, the squared correlation coefficient is comprised within the interval $1 \leq \rho \leq N$.

5.3.4 Frequency-Domain Interpretation

In [ZPQ10], it has been shown that quadratic detectors can be approximated by the correlation between the periodogram of the observations and an spectral mask corresponding to the N-points sampled vector of the Fourier transform of the matrix involved in the quadratic detection. In fact, this approximation behaves asymptotically as $N \gg 1$ equivalently to the optimum quadratic detector, because the matrix involved in the quadratic detection is equivalent to a circulant matrix whose eigenvalues are given by the spectrum evaluated at the n-th frequency bins and whose eigenvectors are from the Fourier matrix [Gra06]. In this work, the GLRT signal detector (5.78) is asymptotically equivalent to the frequency-domain test

$$T(\mathbf{x}) \simeq \frac{N}{2\pi} \int_{-\pi}^{\pi} \phi_\kappa(\omega) P(\omega) d\omega, \tag{5.84}$$

In (5.84), $\phi_\kappa(\omega)$ is the normalized spectrum of the Bernoulli nonuniform sampling signal (i.e, the frequency-domain version of $\mathbf{\Sigma}_\kappa$) normalized to the unit power as

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \phi_\kappa(\omega) d\omega = 1, \tag{5.85}$$

and $P(\omega)$ is the periodogram of the observations, i.e.,

$$P(\omega) = \frac{1}{N} \left| \mathbf{e}^H(\omega)\mathbf{x} \right|^2, \tag{5.86}$$

where $\mathbf{e}(\omega)^H = [1\ e^{-\jmath\omega}\ \ldots\ e^{-\jmath(N-1)\omega}]$.

Analogously, the squared correlation coefficient ρ defined in (5.83) admits the following

asymptotic frequency-domain implementation

$$\rho \simeq \frac{1}{2\pi} \int_{-\pi}^{\pi} \phi_{\kappa}^{2}(\omega) d\omega. \tag{5.87}$$

Because $\phi_{\kappa}(\omega)$ is normalized, the squaring operation in (5.87) attenuates the areas in which the spectrum is low and preserves the areas in which the spectrum is high. Therefore, the frequency-domain squared correlation coefficient is a measure of the concentration of the spectrum in frequency. In Bernoulli nonuniform sampling, the spectrum of the observations suffers from the self interference term or noise enhancement term analogous to the normalized correlation (5.75). The normalized spectra are then related by the linear combination

$$\phi_{\kappa}(\omega) = \kappa \phi_{s}(\omega) + (1 - \kappa), \tag{5.88}$$

being $\phi_s(\omega)$ the normalized spectrum of $s(t)$. Restricting the study to flat spectrum signals[4] with unknown spectral support and known occupancy κ_0, it rapidly follows that the normalized spectrum of $s(t)$ is of the form $\phi_s(\omega) = 1/\kappa_0$ when ω lies in the spectrum support, and $\phi_s(\omega) = 0$ otherwise. This formulation allows to parameterize the frequency-domain squared correlation coefficient as a function of the sampling density and the occupancy by

$$\rho \simeq \frac{\kappa_0 + (1 - \kappa_0)\kappa^2}{\kappa_0}, \tag{5.89}$$

which is in fact the same factor ρ that affects the KLD in the Stein's lemma in (5.45). The former expression will essentially describe the behavior of the asymptotic performance of the GLRT signal detector (5.78) in terms of the relation between κ and κ_0. To this point, the coherence between (5.83) and (5.89) is complemented in the following extreme cases. When the occupancy approaches to one, $\rho \simeq 1$ regardless the sampling density κ. This corresponds to the case in which the Bernoulli nonuniform sampling signal has white statistics, i.e., $\mathbf{\Sigma}_{\kappa} = \mathbf{I}_N$, and the GLRT signal detector only exploits the zero lag correlation. Finally, if the sampling density approaches to one (uniform sampling), the remaining terms simplify to $\rho \simeq 1/\kappa_0$ which increases inversely proportional to the occupancy. Therefore, the frequency-domain squared correlation coefficient is comprised within the interval $1 \leq \rho \leq 1/\kappa_0$, provided that $N \gg 1/\kappa_0$.

[4]The properness of adopting the flat spectrum assumption is threefold. Firstly, it allows to obtain clear insights as the frequency-domain squared correlation coefficient is characterized by a single parameter of the primary signal, i.e, the occupancy κ_0. Secondly, the flat spectrum assumption holds in many practical primary systems, such as terrestrial digital video broadcasting (DVB-T) based OFDM. And thirdly, for a given occupancy, the flatness assumption acts as a worst case scenario, as any signal with equal power and non-flat spectrum will exhibit larger correlation coefficient. This claim is proved in Appendix C.9.

5.3.5 Asymptotic Performance

The performance of the GLRT signal detector (5.78) is evaluated by means of the error probability pair $(P_{\mathrm{FA}}, P_{\mathrm{MD}})$, namely the false-alarm and missed-detection probabilities, defined as $P_{\mathrm{FA}} \doteq \mathbb{P}\left[T(\mathbf{x}) \geq \tau \mid \mathcal{H}_0\right]$ and $P_{\mathrm{MD}} \doteq \mathbb{P}\left[T(\mathbf{x}) < \tau \mid \mathcal{H}_1\right]$, respectively. From the asymptotic PDF (5.82), in Appendix C.10 it is shown that the minimum SNR for achieving a target error probability pair (ϵ_0, ϵ_1) scales with the observation size and the squared correlation coefficient as

$$\mathrm{SNR}_{\min} \propto \frac{1}{\sqrt{N\rho}}. \tag{5.90}$$

The rate at which $\mathrm{SNR}_{\min}$ decreases is strongly related to the correlation level present in the Bernoulli nonuniform sampling signal. Restricting to uniform sampling ($\kappa = 1$ and $N = M$), the squared correlation coefficient (5.83) evaluates the correlation level of $s(t)$. Therefore, the scaling of $\mathrm{SNR}_{\min}$ has the following extreme values. On the one hand, the worst scenario is given by the lower bound $\rho = 1$, which represents a totally uncorrelated signal. In such a case, $\mathrm{SNR}_{\min} \propto \sqrt{1/M}$, which corresponds to the scaling of the SNR in the energy detector [Kay98a]. This is a known detection result in digital communications, as the error probability in noncoherent communications (e.g., detection or synchronization of unknown waveforms) scales as the inverse of the squared root of the sensing time. On the other hand, the best scenario is given by the upper bound $\rho = M$, which represents a fully correlated signal. In such a case, it follows that $\mathrm{SNR}_{\min} \propto 1/M$, i.e., the minimum SNR scales inversely proportional to the observation size. This is also a known result in digital communications, as the error probability in coherent communications (e.g., detection or synchronization of known waveforms) scales with the inverse of the sensing time [Kay98a].

For arbitrary sampling density $0 < \kappa \leq 1$, one can evaluate the evolution of $\mathrm{SNR}_{\min}$ by making use of the frequency-domain interpretation of the squared correlation coefficient (5.87). After some mathematical manipulations and recalling that N is the nearest integer that satisfies $M = \kappa N$, it follows that the minimum SNR scales as

$$\mathrm{SNR}_{\min} \propto \sqrt{\frac{\kappa\kappa_0}{\kappa_0 + (1-\kappa_0)\kappa^2}} \cdot \frac{1}{\sqrt{M}}. \tag{5.91}$$

The importance of the former expression is given by the fact that it evaluates the rate at which the minimum SNR required for signal detection in Bernoulli nonuniform sampling scales with the sampling density and the occupancy. In Figure 5.3, the normalized minimum SNR is plotted against the sampling density for several occupancies. It is appreciated that if the occupancy is below $\kappa_0 < 1/2$, there exists a sampling density given by

$$\kappa_{\min} = \sqrt{\frac{\kappa_0}{1-\kappa_0}}, \tag{5.92}$$

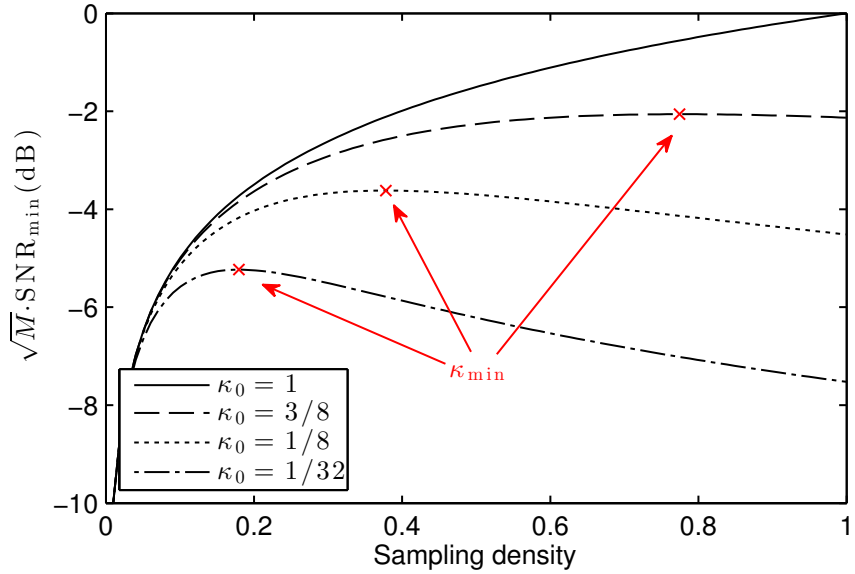

Figure 5.3: Evolution of the minimum SNR as a function of the sampling density and the occupancy, and location of the sampling densities that achieve worst case minimum SNR.

above which the GLRT signal detector is able to exploit the remaining correlation in the Bernoulli nonuniform sampling signal. Hence, the minimum SNR is monotonically decreasing (improves) with the sampling density (higher rates) only if $\kappa > \kappa_{\min}$. On the other hand, for sampling densities $\kappa < \kappa_{\min}$, the white self interference incurred in Bernoulli nonuniform sampling degrades the signal autocorrelation in such a level that the GLRT signal detector essentially resorts to energy detection. Therefore, in this range the minimum SNR decreases with a rate of $\sqrt{\kappa}$ as $\kappa \ll 1$, because the sensing time is increased with a factor of κ. In conclusion, the minimum sampling density $\kappa_{\min}$ acts as a worst case scenario on the minimum SNR in Bernoulli nonuniform sampling signal detection. Interestingly, both the location of the minimum sampling density and the value of the worst case minimum SNR (red cross markers in Figure 5.3) increase with the occupancy. This corroborates that smaller sampling rates can be attained when the occupancy of the signal is small compared to B. In the limit, as $\kappa_0 \gg 1$, the minimum sampling density can be approximated by $\kappa_{\min} \approx \sqrt{\kappa_0}$. In such a case, the Bernoulli nonuniform sampling sampling density must be higher than the squared root of the occupancy in order to take advantage of the side information of the problem, i.e., the noise variance, the sampling density and the primary signal autocorrelation.

A final highlight is that an alternative indicator of the asymptotic performance of the GLRT signal detector is the normalized minimum sensing time $T_{\min}$. This parameter behaves as a measure of complexity involved in the detection process, as the complexity of a quadratic form with Toeplitz matrices is of the order $\mathcal{O}(2N)$. From (5.90) and after some mathematical manipulations, it follows that the normalized minimum sensing time that achieves a target error probability pair scales with the SNR and the squared correlation

coefficient as

$$T_{\min} \cdot B \propto \frac{1}{\mathrm{SNR}^2 \rho}, \tag{5.93}$$

i.e., it scales inversely to the squared value of the SNR and with the inverse of the squared correlation coefficient.

5.3.6 Sampling Walls in Noise Uncertainty

This Section addresses the performance of (5.78) in the presence of noise uncertainty. Because the noise variance cannot be known with infinite precision in real physical systems, a small error model is adopted as noise uncertainty. This model imposes fundamental limitations on the detection performance, even though when the noise level in silent times is estimated in a time duration that do not scale with the sensing time [MGC11]. These limitations are presented in terms of SNR walls [TS08]. In this thesis, these detection walls are extended to sampling walls, i.e., sampling densities below which the detection is not possible, even for infinite sensing time.

Let $\hat{\sigma}^2$ be a random variable denoting the prior information on the noise variance. In the small error model, it is assumed to be uniformly distributed within the range

$$\sigma^2/\delta \leq \hat{\sigma}^2 \leq \delta\sigma^2. \tag{5.94}$$

In the sequel, this will be referred as having a noise uncertainty of $u = 10\log_{10}\delta$ dB. From the asymptotic PDF (5.82), the error probability pair (ϵ_0, ϵ_1) cannot be guaranteed because the uncertainty on the noise variance translates to a wrong setting of the threshold (5.79).

In this work, the threshold is set to guarantee a given outage on the false-alarm and missed-detection probabilities. In other words, the value of τ delivers a false-alarm P_{FA} and a missed-detection P_{MD} such that $\mathbb{P}[P_{\mathrm{FA}} \geq \epsilon_0] < \epsilon$ and $\mathbb{P}[P_{\mathrm{MD}} \geq \epsilon_1] < \epsilon$, $0 \leq \epsilon < 1$, where the probability is evaluated in the randomness of σ^2. If $\epsilon = 0$, then the target error probability pair is satisfied for any range of $\hat{\sigma}^2$ (worst case approach). In Appendix C.11, it is shown that the minimum SNR for achieving a target error probability pair (ϵ_0, ϵ_1) with outage probability ϵ scales with the observation size and the squared correlation coefficient as

$$\mathrm{SNR}_{\min} \propto \frac{1}{\sqrt{N\rho}} + \frac{U-1}{\rho}, \tag{5.95}$$

where U a parameter that relates the uncertainty level and the outage probability as

$$U = \frac{\delta^2(1-\epsilon)+\epsilon}{\delta^2\epsilon + 1 - \epsilon}. \tag{5.96}$$

Notice that for $\epsilon = 0$, U becomes the peak-to-peak ratio of the uncertainty model. By comparing (5.95) to (5.90), it is easy to see that the noise uncertainty incurs in a penalty in terms of SNR which is inversely proportional to the squared correlation coefficient and

Region	Interval
Unfeasibility Region	$0 \leq \text{SNR} < \kappa_0(U-1)$
Sampling Walls Region	$\kappa_0(U-1) < \text{SNR} < U-1$
Wallfree Region	$\text{SNR} > U-1$

Table 5.1: SNR Regions for Sampling Walls.

that it is always nonnegative, as $U \geq 1$. This term is referred as the SNR wall associated to the GLRT signal detector (5.78), defined as

$$\text{SNR}_{\text{wall}} \doteq \lim_{N\to\infty} \text{SNR}_{\text{min}} = \frac{U-1}{\rho}. \tag{5.97}$$

In the worst case of uncorrelated signal ($\rho = 1$), the SNR wall is linear with the parameter (5.96) as $\text{SNR}_{\text{wall}} = U-1$. This is the classical energy detector SNR wall [TS08]. In the most favorable scenario of fully correlated signal, i.e., when $\rho = N$, the SNR wall improves as it scales inversely proportional to the observation size.

The SNR wall (5.97) admits the following frequency-domain interpretation by plugging the expression of the squared correlation coefficient that relates the Bernoulli nonuniform sampling sampling density and the occupancy. This leads to

$$\text{SNR}_{\text{wall}} = \frac{\kappa_0}{\kappa_0 + (1-\kappa_0)\kappa^2} \cdot (U-1). \tag{5.98}$$

For Nyquist uniform sampling, i.e, $\kappa = 1$, the SNR wall is directly related to the occupancy as $\text{SNR}_{\text{wall}} = \kappa_0(U-1)$. This means that the SNR wall is smaller when the primary signal has low occupancy, because the signal detector preserves the ability of distinguishing a narrowband signal immersed in white noise, regardless the exact level. On the contrary, for a very low sampling density, i.e., κ approaching zero, the SNR wall approaches the energy detector: $\text{SNR}_{\text{wall}} = U-1$. The two aforementioned SNR walls, which correspond to Nyquist uniform sampling and energy detector, define the SNR region in which sampling walls are experienced. In Table 5.1, the three following SNR regions are introduced:

- First, the *unfeasibility region* is the SNR region characterized by the inability of the GLRT signal detector (5.78) to guarantee the target error probabilities, even for an infinite number of samples.

- Second, the *wallfree region* is lower bounded by the SNR wall of the energy detector. In this region, the nonexistence of sampling walls is guaranteed regardless the combination of sampling density and primary signal occupancy.

- And third, the *sampling walls region* is the SNR region comprised within the SNR walls $\kappa_0(U-1)$ and $U-1$. The existence of sampling walls inside this region is

consequence from the feasibility condition $\text{SNR} > \text{SNR}_{\text{wall}}$ in (5.98).

After some mathematical manipulations, the feasibility condition (5.98) is translated to the feasibility condition

$$\kappa > \kappa_{\text{wall}}, \tag{5.99}$$

where the sampling wall κ_{wall} is given as a function of the noise uncertainty, the primary signal occupancy and the SNR by the following expression

$$\boxed{\kappa_{\text{wall}} = \sqrt{\frac{\kappa_0(U - 1 - \text{SNR})}{(1 - \kappa_0)\text{SNR}}}.} \tag{5.100}$$

The sampling wall (5.100) is a fundamental limit, below which detection is not feasible, i.e., the target error probabilities cannot be guaranteed. From (5.100), it is seen that the sampling wall is proportional to the occupancy. Therefore and as expected, smaller sampling densities can be achieved in low occupancy scenarios.

Finally, an alternative representation of sampling walls is reported through the normalized minimum sensing time that achieves the target error probabilities, $T_{\min}$. By rewriting (5.95), the normalized minimum sensing time scales as

$$T_{\min} \cdot B \propto \frac{1}{\rho\left(\text{SNR} - \text{SNR}_{\text{wall}}\right)^2}, \tag{5.101}$$

i.e., inversely proportional to the squared correlation coefficient and with asymptotes located at the SNR walls and sampling walls.

5.4 Numerical Results

5.4.1 Stein's Lemma

Simulation results are provided to assess the behavior of the error probabilities of the energy detector (5.11) and the estimator-correlator (5.21) in low SNR conditions. More specifically, the linear behavior of the logarithm of the error probabilities as well as the tightness of the Stein's lemma are evaluated. In the sequel, the number of observations M vary from 1 to 1024 and are equally spaced in a base-2 logarithmic scale.

The Stein's lemmas on the false-alarm probability and missed-detection probability for the energy detector (5.11) are evaluated in Figures 5.4 and 5.5, respectively, for $N = 2$. As it can be appreciated in both figures, the error probabilities in signal detection obey a linear scaling with the number of observations as $M \to \infty$. Furthermore, the rate at which the error probabilities diminish with M, i.e., the slope associated to the error exponents, is asymptotically given by the KLD (5.13) which depend on the SNR of the problem, as well as the observation size N. As the scenario conditions are in the low SNR regime, the

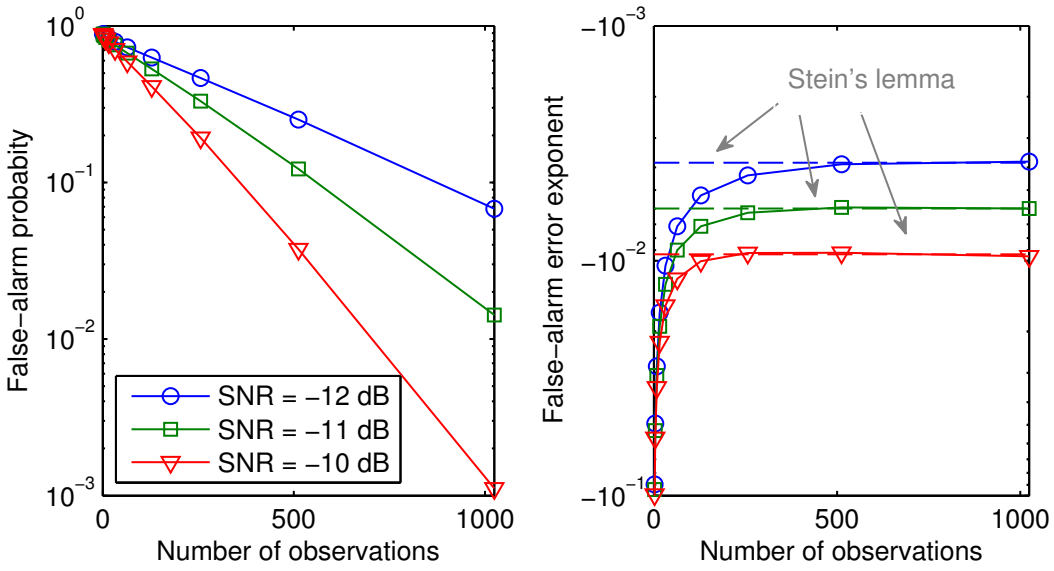

Figure 5.4: False-alarm probability versus observation size and simulated false-alarm error exponent versus the number of observations compared to the theoretical error exponents for the energy detector.

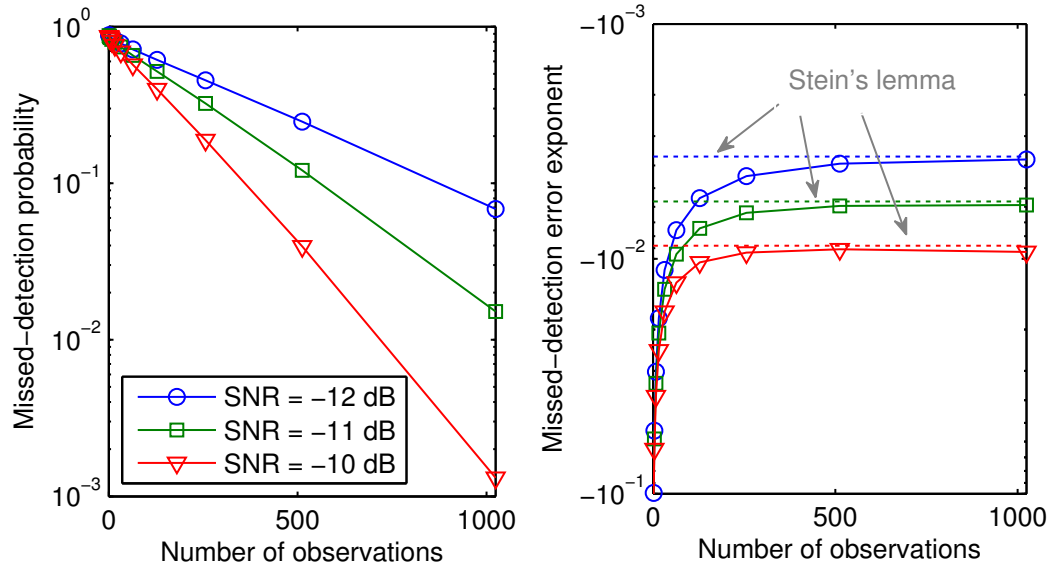

Figure 5.5: Missed-detection probability versus observation size and simulated missed-detection error exponent versus the number of observations compared to the theoretical error exponents for the energy detector.

asymptotic error exponents depend on the squared value of the SNR. Therefore, the three error exponents in the right hand of the Figures are equally spaced, as the SNR points are equally spaced as well in the logarithmic scale. Also, because of the low SNR regime, both false-alarm and missed-detection probabilities show very similar performances in terms of achievable error exponents. This corroborates the fact that the KLD (5.22) have the same

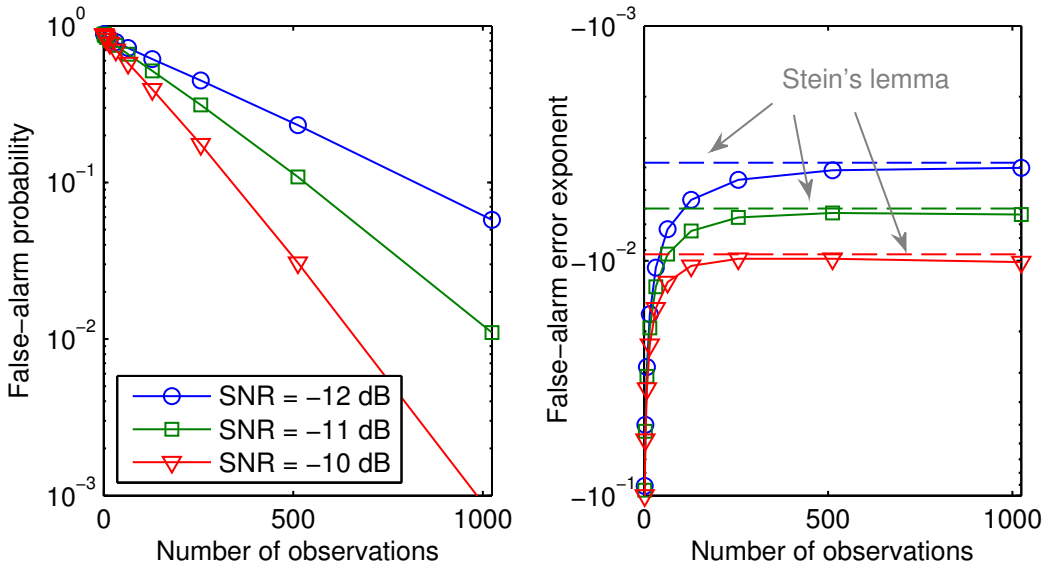

Figure 5.6: False-alarm probability versus number of observations and simulated false-alarm error exponents versus number of observations compared to the theoretical error exponents for the estimator-correlator.

approximation in the low SNR regime.

Secondly, the asymptotic performance of the estimator-correlator (5.21) is evaluated in terms of Monte Carlo trials on the error probabilities in several low SNR conditions. In this example, the following simple correlation matrix $\mathbf{R}_s = \text{SNR}[1\ -0.5; -0.5\ 1]$ has been employed, and $\mathbf{R}_w = \mathbf{I}$. The false-alarm probability and its associated error exponent is depicted in Figure 5.6, whereas the missed-detection probability and its associated error exponent is depicted in Figure 5.7. Under the same conditions of SNR, observation size, and number of samples, the estimator-correlator performs slightly better than the energy detector as it is able to exploit the correlation of the signal in $\mathbf{R}_s$. This is seen in the asymptotic error exponents (5.22) which, for white Gaussian noise, reduce to $\mathbf{SNR} = \frac{1}{\sigma^2}\mathbf{R}_s$. Finally, as expected, for arbitrarily large M, e.g., for $M \geq 100$, the linear behavior of the logarithm of the error probabilities with M is observed, as well as the tightness of the Stein's lemma for $M \geq 500$.

Thirdly, the behavior of the Stein's lemma in Bernoulli nonuniform sampling is illustrated in Figure 5.8 and Figure 5.9 for an SNR of -10 dB. In both Figures, where the error probabilities and their associated simulated and theoretical error exponents are depicted for several sampling densities, it can be appreciated how the Stein's lemma in this scenario accurately predicts the slope of the error probabilities as M grows. In particular, the error exponents match the KLD of the estimator-correlator (5.22) with the equivalent SNR matrix (5.37) in Bernoulli nonuniform sampling. Further as expected, the error exponents suffer from degradation as the sampling density is decreased due to the noise enhancement phenomenon.

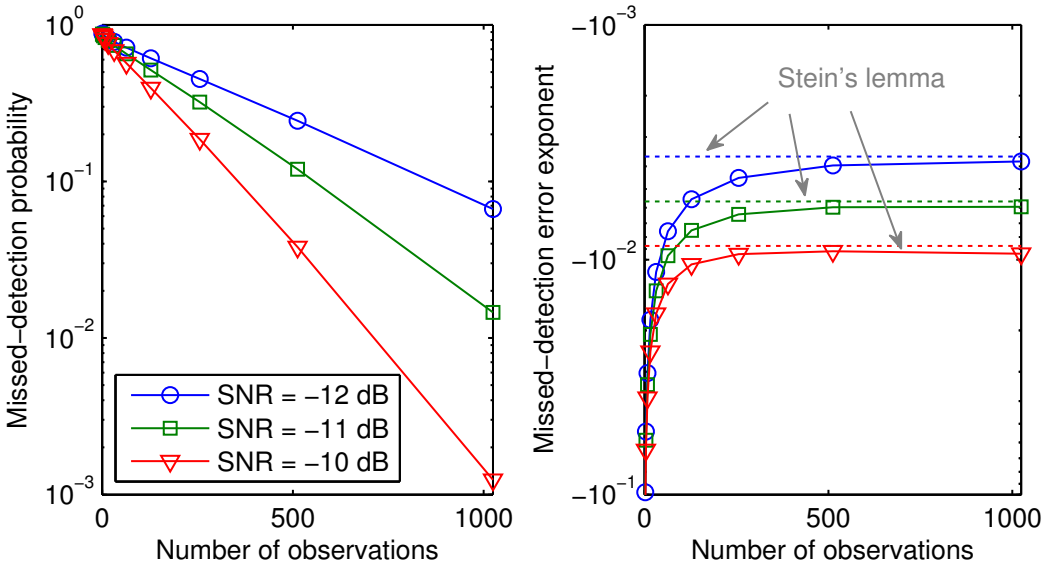

Figure 5.7: Missed-detection probability versus observation size and simulated missed-detection error exponent versus the number of observations compared to the theoretical error exponents for the estimator-correlator.

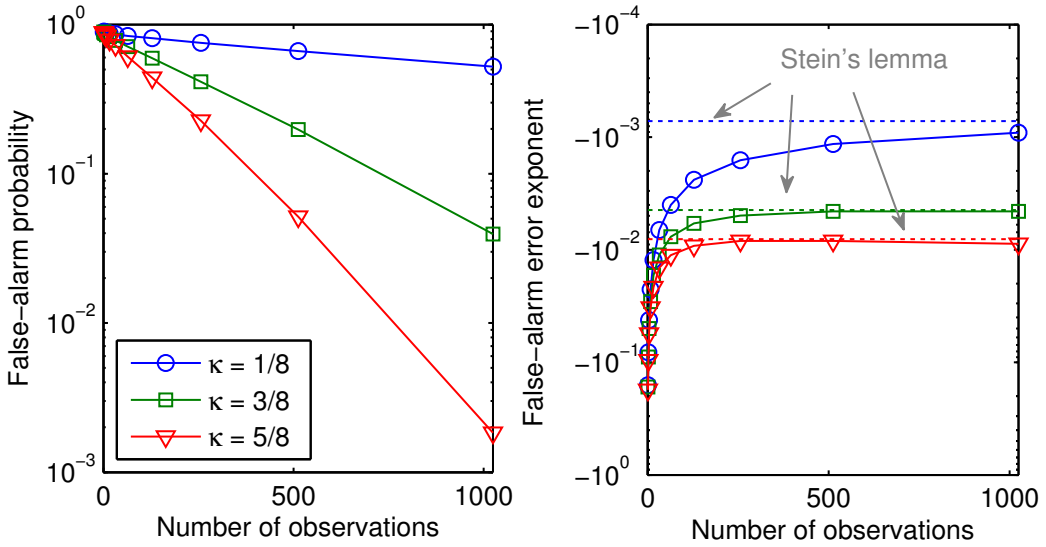

Figure 5.8: False-alarm probability versus number of observations and simulated false-alarm error exponents versus number of observations compared to the theoretical error exponents for the estimator-correlator in Bernoulli nonuniform sampling.

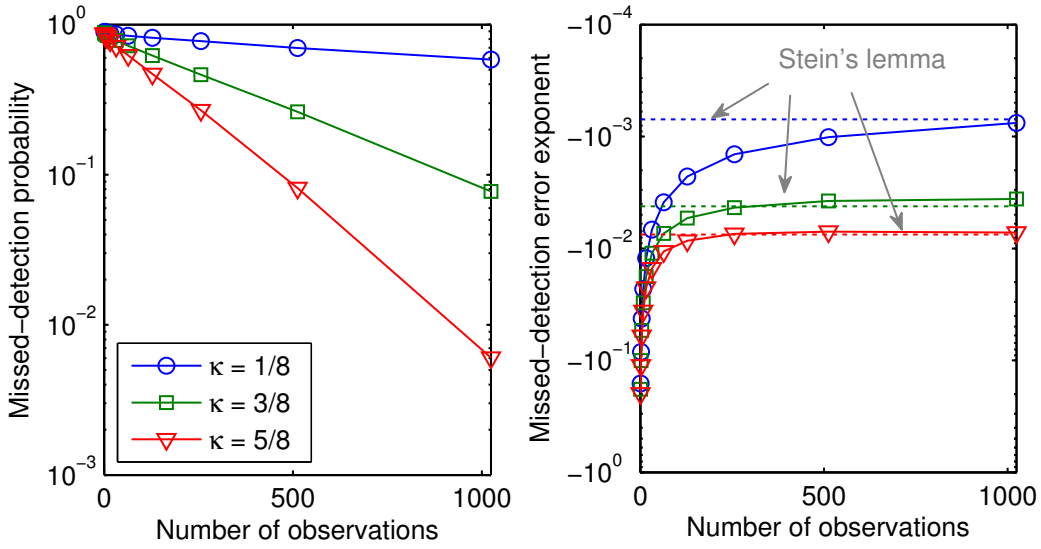

Figure 5.9: Missed-detection probability versus observation size and simulated missed-detection error exponent versus the number of observations compared to the theoretical error exponents for the estimator-correlator in Bernoulli nonuniform sampling.

Finally, the false-alarm probability and its associated error exponent of the estimator-correlator in coherent signal detection with $N = 1$ is depicted in Figure 5.10, whereas the missed-detection probability and its associated error exponent is plotted in Figure 5.11. In both figures, several diversity levels L are considered, at an average SNR of -10 dB. In this particular scenario, the Stein's lemma has been applied on the KLD of the estimator-correlator, i.e, (5.22), with the SNR matrix (5.58), where the diversity channel $\mathbf{h}$ has been generated as i.i.d. complex Gaussian with unit variance. Apart from noticing that in this scenario the Stein's lemma provides a tight bound on the error exponents as well, it can be appreciated how the diversity incurs a high improvement in the slope of the error exponents, as predicted in Section 5.2.6.

5.4.2 Sampling Walls

In this Section, the performance of the GLRT signal detector (5.78) and the behavior of the fundamental limits derived in this Chapter are empirically evaluated. In what follows, the primary signal is generated according to the DVB-T standard in the 2k-mode with $Q = 8$ bands, which exhibits a flat spectrum in the spectral support and allows concrete configurations of the occupancy.

The performance of the GLRT signal detector (5.78) in Bernoulli nonuniform sampling is evaluated in terms of the false-alarm and missed-detection probabilities in the following scenarios of occupancies and noise uncertainties. In the following, the number of samples has been fixed to $M = 64$ samples. First, Figure 5.12 plots the receiver operating characteristics (ROC) at two values of SNR for several occupancies when the sampling density

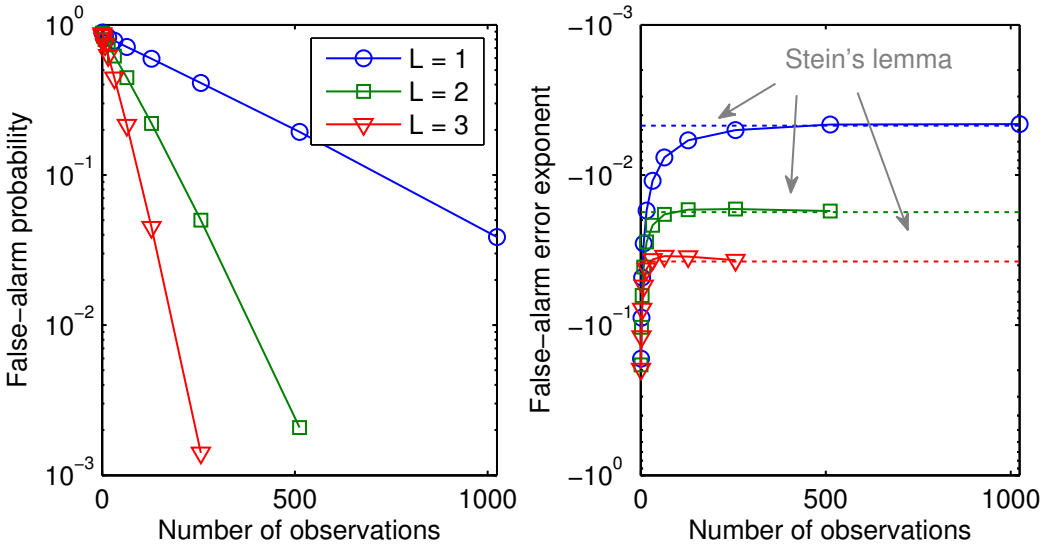

Figure 5.10: False-alarm probability versus number of observations and simulated false-alarm error exponents versus number of observations compared to the theoretical error exponents for the estimator-correlator in coherent signal detection.

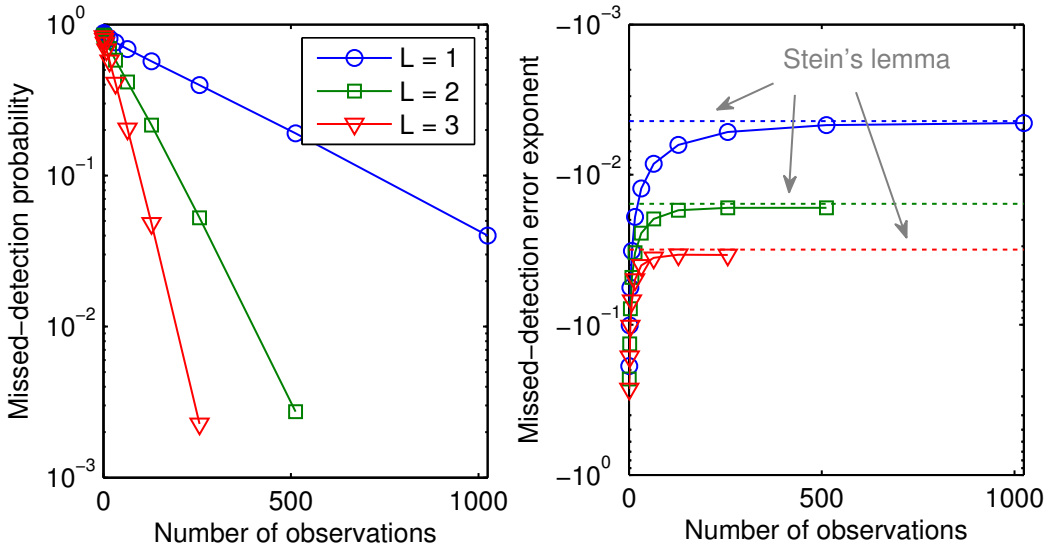

Figure 5.11: Missed-detection probability versus observation size and simulated missed-detection error exponent versus the number of observations compared to the theoretical error exponents for the estimator-correlator in coherent signal detection.

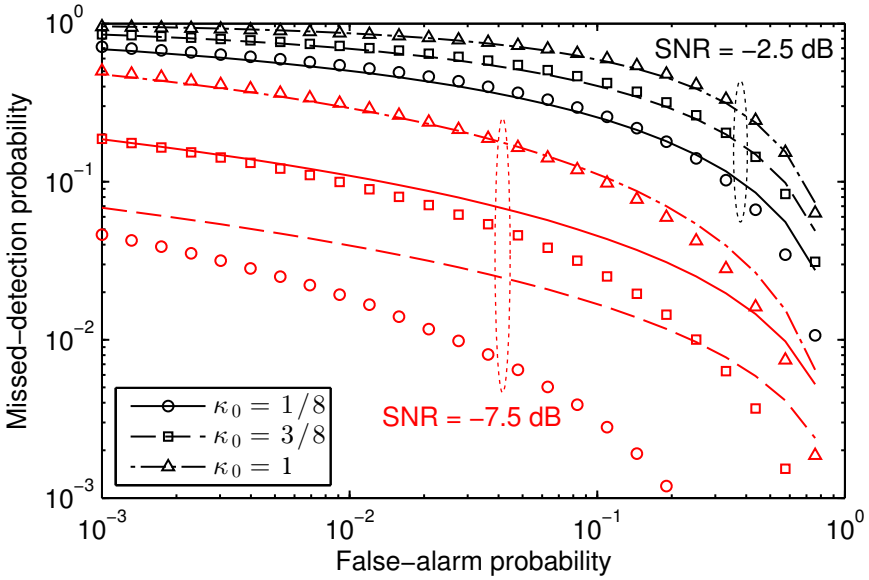

Figure 5.12: Theoretical (lines) and empirical (markers) ROC for several occupancies at the minimum sampling density and two values of SNR.

of Bernoulli nonuniform sampling has been set to $\kappa = \kappa_{\min}$, according to the expression (5.92). Markers denote the empirical probabilities obtained through 100,000 random realizations of the Bernoulli nonuniform sampling with occupancy κ_0. Because the sampling density has been set to the minimum sampling density, the ROC depicted in Figure 5.12 are fundamentally worst case scenarios. In other words, the maximum penalty incurred in Bernoulli nonuniform sampling is evaluated. At an SNR $= -7.5$ dB, it can be appreciated that the theoretical false-alarm and missed-detection probability pair given in (C.20) and (C.22) in Appendix C.10 behave as good approximations of the empirical probabilities. However, as the SNR increases, it is seen that the theoretical expressions based on the Gaussian assumption act as a worst case scenario, being an upper bound on the empirical missed-detection probability for any false-alarm probability. In both SNR ranges, the performance of the GLRT signal detector improves with smaller occupancies. For instance, for a false-alarm probability of 10% and SNR $= -2.5$ dB, a signal with an occupancy of 12.5% has an improvement in terms of missed-detection probability of almost two orders of magnitude with respect to the energy detection of white signals.

Analogous results are obtained in the evolution of the complimentary missed-detection probability, as plotted in Figure 5.13 against the SNR for a fixed false-alarm probability of 1% for several occupancies when the sampling density of Bernoulli nonuniform sampling has been also set to $\kappa = \kappa_{\min}$. Markers denote the empirical probability obtained through 100,000 random realization of the Bernoulli nonuniform sampling and a flat spectrum signal with occupancy κ_0. In this plot, it is noted that for a fixed missed-detection probability level, the gain attained with Bernoulli nonuniform sampling in low occupancy scenarios is of several dBs of SNR. Further, the penalty incurred by a noise uncertainty of $u = 0.5$ dB

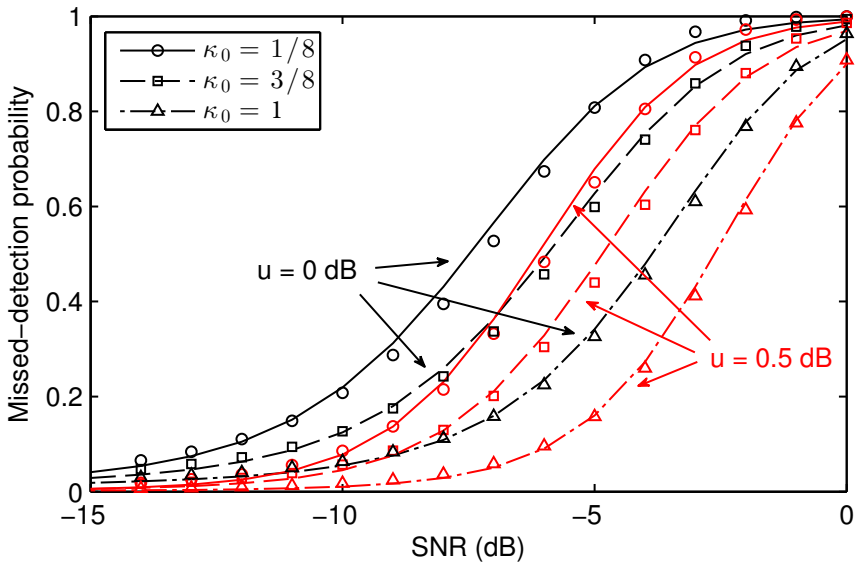

Figure 5.13: Theoretical (lines) and empirical (markers) complimentary missed-detection probability versus SNR for several occupancy at the minimum sampling density, with (red) and without (black) noise uncertainty.

with a probability outage of $\epsilon = 0.1$ is of the same order regardless the occupancy of the signal, and implies a shift in the missed-detection probability of about 1-2 dB.

In this thesis, two fundamental detection walls have been derived under noise uncertainty, namely the SNR walls and the sampling walls. The theoretical SNR walls and sampling walls are visually depicted as asymptotes in Figure 5.14 and 5.15, respectively, both plotting the evolution of the normalized minimum sensing time (5.101) for a target error probability pair $\epsilon_0 = \epsilon_1 = 0.01$. In both figures, markers denote the empirical sensing times obtained through 100,000 random realizations of the Bernoulli nonuniform sampling signal.

The SNR walls are illustrated in Figure 5.14 when plotting the normalized minimum sensing time versus the SNR for several sampling densities for an occupancy of 37.5%. When there is no noise uncertainty ($u = 0$ dB), the scaling of $T_{\min}$ with the SNR is linear in the logarithmic scale, according to (5.93). In other words, for every increase of 5 dB in SNR, the target error probability pair can be achieved 10 times faster. Further, sampling at a lower density also requires sensing during a larger period of time. Finally, when the noise variance has an uncertainty of $u = 0.5$ dB with a probability outage of $\epsilon = 0.1$, the SNR walls (5.98) appear as asymptotes on the normalized minimum sensing time, as if $\text{SNR} \leq \text{SNR}_{\text{wall}}$, the false-alarm and missed-detection target error probability pair cannot be achieved regardless the normalized minimum sensing time devoted to detection. Also, as it can be appreciated in the expression of the SNR walls, the location of the walls moves toward higher SNRs as the sampling density is smaller. The empirical results validate both the slope of the minimum sensing time with the SNR and the location of the SNR walls for

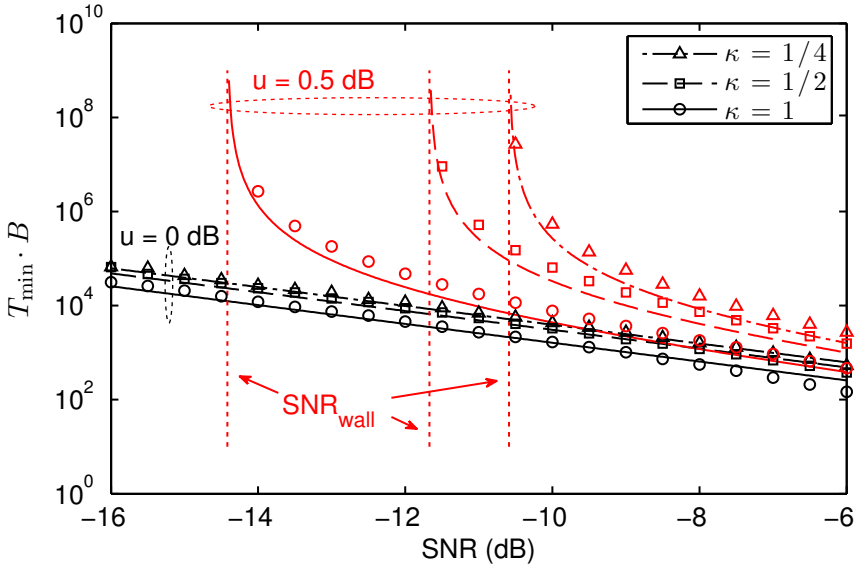

Figure 5.14: Theoretical (lines) and empirical (markers) evolution of the normalized minimum sensing time as a function of the SNR for several sampling densities, several noise uncertainties, and occupancy of 3/8.

the sampling densities considered.

On the other hand, the sampling walls are illustrated in Figure 5.15 when plotting the normalized minimum sensing time versus the sampling density for several occupancies and for a fixed SNR of -11.5 dB. For a given occupancy, the normalized minimum sensing time is monotonically increasing when moving toward smaller sampling densities. Since the SNR conditions are inside the sampling wall region, the sampling walls (5.100) appear as asymptotes in the normalized minimum sensing time. For instance, for an occupancy of $\kappa_0 = 1/8$ and a noise uncertainty of $u = 0.5$ dB, the SNR wall in the energy detector and the SNR wall in Nyquist uniform sampling are given by $U - 1 = -10.15$ dB and $\kappa_0(U-1) = -19.18$ dB, respectively. Hence, for SNR $= -11.5$ dB, the inequality

$$\kappa_0(U-1) < \text{SNR} < U - 1 \tag{5.102}$$

is satisfied. The location of the specific sampling wall is given, from (5.100), by $\kappa_{\text{wall}} = 0.2275$, which is slightly below the minimum sampling density $\kappa_{\min} = \sqrt{1/7} = 0.3780$. In other words, sampling below $\kappa_{\min}$ makes impossible signal detection based on the signal autocorrelation, while sampling below κ_{wall} makes the target error probability pair unattainable. Finally, the empirical markers confirm the existence and location of sampling walls for the occupancies considered, as well as the behavior of the minimum sensing time with the sampling density of Bernoulli nonuniform sampling.

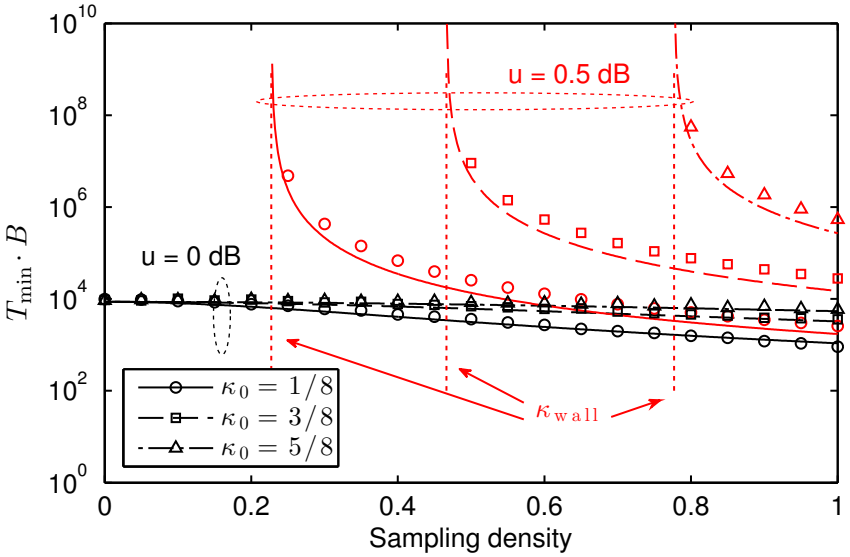

Figure 5.15: Theoretical (lines) and empirical (markers) evolution of the normalized minimum sensing time as a function of the sampling density for several occupancies and noise uncertainties, at SNR of -11.5 dB.

5.5 Conclusions

In this Chapter, the asymptotic behavior of the error probabilities in signal detection has been addressed by means of the Stein's lemma in several scenarios including the energy detector, the estimator-correlator, nonuniform sampling, diversity and coherent detection. It has been shown that the error exponents in the energy detector and the estimator-correlator depend on the observation size and the SNR profile of the problem. Closed-form expressions of the asymptotic relation between observation size and SNR have been obtained. Simulation results have been reported to assess the linear behavior of the logarithm of the error probabilities, as well as to evaluate the tightness of the lemma.

Furthermore, closed-form expressions of sampling walls in Bernoulli nonuniform sampling have been derived in the presence of noise uncertainty. The optimal low SNR GLRT signal detector has been derived. The asymptotic statistical characterization of the test has provided theoretical insights into the minimum SNR, false-alarm probability, missed-detection probability, SNR walls and sampling walls as functions of the primary signal occupancy and sampling density. Numerical results have been reported to assess the validity of the derived theoretical expressions.

Chapter 6

Cyclostationary Primary Signal Detection

6.1 Introduction

Cyclostationary signal detection is a problem of interest in cognitive radio networks, as cyclostationary processes exhibit a form of frequency diversity which can be exploited by secondary users to perform spectrum sensing. Even though their robustness in front of noise uncertainty, cyclostationary-based detectors are highly sensible to the frequency-selective nature of the primary-to-secondary propagation channel. For this reason, this Chapter focusses on the design of cyclostationary detectors under the assumptions of unknown white noise, unknown timing parameter, and unknown channel gains.

6.1.1 Quadratic Sphericity Test

A first invariant cyclostationarity detector is proposed by deriving a blind detector based on the squared mean to arithmetic mean ratio of the eigenvalues of the autocorrelation matrix of the observations, denote as the quadratic sphericity test (QST). First, a test statistic which is optimal in terms of a correlation-matching of the second-order statistics of the problem is derived. This test statistics further shows invariance with respect to the noise power and the channel gain. The QST is derived by optimizing the correlation-matching cost function with respect to the unknown channel fading coefficients.

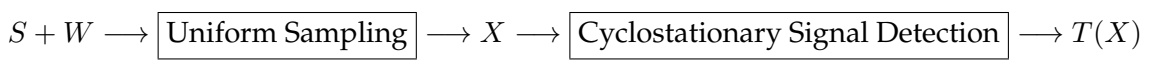

Figure 6.1: Cyclostationary primary signal detection problem.

Prior work related to the QST is detailed as follows. Spectrum sensing detectors based on the autocorrelation matrix of the observations are reported in [PNLRS09, ZL09, QZSS11]. These works exploit the structure of the correlation matrix of the transmitted primary signal, but fading is not considered. Spectrum sensing detectors based on the eigenvalues of the sample covariance matrix include the maximum to minimum eigenvalue (MME) [ZcL09] and the normalized maximum eigenvalue (NME) [ZLZ08]. The latter is considered in the numerical results for comparison. There is also a large body of spectrum sensing detectors taking into account unknown parameters, e.g., [ZLLZ10, TNKG10, RVVLV$^+$11, VVLVS11, SAVVLV12]. However, these works exploit the spatial correlation that arises from the use of multiple antennas. In this thesis, a single antenna receiver in time-varying frequency-selective channels is rather considered. Finally, the works [MMA11, FSW10] consider signal detection in fading channels, but considering prior knowledge on the statistics of the channel.

6.1.2 Frequency-Domain Detection

The frequency diversity exhibited by cyclostationary signals is further exploited in Section 6.3. A novel rank-1 frequency-domain representation of a digital waveform is proposed to address the generalized likelihood ratio (GLR) detection of a cyclostationary signal with unknown white noise. With the aim of avoiding the well-known sensitivity of cyclostationary-based detectors to frequency-selective fading channels, a parametric channel model based on the coherence bandwidth is adopted and incorporated in the generalized likelihood ratio test (GLRT). The proposed detector outperforms the classical spectral correlation magnitude detectors by exploiting the rank-1 structure of small spectral covariance matrices.

The adoption of the primary user model is a relevant issue, and may include features such as the transmitted power spectral density (PSD) [QZSS11], cyclostationarity, modulation type, etc. [ALLP12]. On the other hand, the frequency-selective and time-varying nature of the wireless propagation channel can severely affect the detectors' performance [TS07]. The use of multiple antennas has received attention in order to improve performance [BKS06, ZLLZ10, RVVLV$^+$11, RVSS13]. The multiple antenna formulation leads to low rank spatial covariance matrices. This fact is exploited to improve, for instance, the noise floor estimation, as one has access to a noise subspace which is free of the signal component. In the single antenna case, the specification of a model for the temporal autocorrelation function of primary user and noise components is required. By resorting to asymptotic properties of Toeplitz matrices, [ZPQ10] shows that detection can be formulated from the signal periodogram, smoothed by an appropriately selected spectral mask. This yields to significant saving on computational complexity, which is specially interesting in order to make cognitive radios feasible.

This work formulates the single-antenna detection problem of a second-order cyclosta-

tionary signal by taking benefit from the mature and recent detection theory advances in array processing. This provides an original formulation of cyclostationary detection based on a vectorial frequency-domain processing that leads to a signal model which exhibits low rank structure. Particularly, the rank-1 structure for pulse-shaped digital modulations considered in our formulation plays the equivalent role of the spatial signature typically found in the array processing field.

Cyclostationarity feature detection had originally emerged as a tool to relax the assumptions on the noise statistics [Gar86], and it has hence gained recent attention in the field of spectrum sensing for weak signals [GS93, FSRVV13] and unknown noise statistics [HT13, RVSS13, FSRVV13] to achieve robustness to signal-to-noise ratio (SNR) walls [TS05]. In particular, the scheme reported in [GS93] and method based on [RVVLV$^+$11, FSRVV13] are considered as benchmarking. The frequency-domain formulation adopted in this thesis was introduced by the authors in [RVV10] in the SNR estimation problem.

6.1.3 Chapter Organization

The rest of the Chapter is organized as follows. The QST is proposed in Section 6.2. The frequency-domain GLR detection is addressed in Section 6.3. Section 6.4 reports numerical results, and Section 6.5 provides the concluding remarks.

6.2 Quadratic Sphericity Test

6.2.1 Problem Formulation

The detection of a primary wide-sense stationary (WSS) signal is cast as the following binary hypotheses testing problem:

$$\mathcal{H}_0 \quad : \quad \mathbf{x} = \mathbf{w}, \tag{6.1a}$$

$$\mathcal{H}_1 \quad : \quad \mathbf{x} = \mathbf{s} + \mathbf{w} \tag{6.1b}$$

where $\mathbf{x} = [x(0), \ldots, x(N-1)]^T$ collects N consecutive samples as observations, $\mathbf{w}$ denotes the noise samples of variance σ^2, $\mathbf{s}$ denotes the signal samples with power P. The signal-to-noise ratio (SNR) is defined as the ratio $\text{SNR} = P/\sigma^2$.

In this Section, the signal $\mathbf{s}$ is the output of a time-varying linear system given as[1]

$$\mathbf{s} = \mathbf{A_H}\mathbf{u}, \tag{6.2}$$

where $\mathbf{s}$ is a column vector containing a sufficiently large number of K digitally modulated symbols, and $\mathbf{A_H}$ models the time-varying frequency-selective nature of the fading

[1]For simplicity it is assumed that the convolution matrix $\mathbf{A_H}$ models the pulse shaping at transmission, the fading nature of the channel, and the matched-filtering at the receiver.

channel as

$$\mathbf{A_H} = \begin{bmatrix} h_0(1) & 0 & \cdots & 0 \\ \vdots & h_1(1) & & \vdots \\ h_0(L) & \vdots & \ddots & 0 \\ 0 & h_1(L) & & h_{K-1}(1) \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & h_{K-1}(L) \end{bmatrix}. \tag{6.3}$$

It is worth noting that the model adopted in (6.3) illustrates how each of the transmitted symbol $s(k)$ is propagated through the channel with fading coefficients $\mathbf{h}_k$. For sake of notation, the channel matrix is defined as

$$\mathbf{H} = [\mathbf{h}_0, \ldots, \mathbf{h}_{K-1}] \tag{6.4}$$

where each column contains the channel impulse response at time index k, i.e.,

$$\mathbf{h}_k = [h_k(1), \ldots, h_k(L)]^T. \tag{6.5}$$

6.2.2 Blind Detector for Time-Varying Frequency-Selective Channels

In this Section, the following test statistic is adopted, which shows invariance with respect to the time varying frequency-selective channel gain and the noise variance, given by

$$T(\mathbf{x}) = \frac{1}{\mathbf{x}^H\mathbf{x}} \max_{\mathbf{H}} \frac{\mathrm{tr}(\mathbf{H}\mathbf{H}^H\hat{\mathbf{R}})}{\|\mathbf{H}\mathbf{H}^H\|_F} \geq \tau, \tag{6.6}$$

where

$$\hat{\mathbf{R}} = \frac{1}{K}\sum_{k=0}^{K-1} \mathbf{x}_k\mathbf{x}_k^H \tag{6.7}$$

is the $L \times L$ autocorrelation matrix of the observations, τ is the detection threshold, and

$$\mathbf{x}_k = [x(k), \ldots, x(k+L-1)]^T. \tag{6.8}$$

The correlation-matching proof of (6.6) is reported in Appendix D.1. In general, the solution to the maximization problem (6.6) leads to the following blind detector as a function of the eigenvalues of the autocorrelation matrix of the observations, as elaborated in Appendix D.2:

$$\boxed{T_{\mathrm{QST}}(\mathbf{x}) = \frac{\sqrt{\sum_{i=1}^{L}\lambda_i^2(\hat{\mathbf{R}})}}{\sum_{i=1}^{L}\lambda_i(\hat{\mathbf{R}})} \geq \tau.} \tag{6.9}$$

Note that the numerator and denominator of (6.9) can be computed as the squared root of $\mathrm{tr}(\hat{\mathbf{R}}^2)$ and as $\mathrm{tr}(\hat{\mathbf{R}})$ with reduced complexity, respectively. This detector is denoted as the

QST, as it evaluates the squared mean of the eigenvalues of the autocorrelation matrix of the observations divided by the arithmetic mean of the eigenvalues. Therefore, from the structure of the autocorrelation matrix (6.7), it is observed that the QST is a measure of dispersion of the eigenvalues in the L-dimensional space that arises from the structure of the fading channel $\mathbf{H}$. Under $\mathcal{H}_0$, the QST will observe a flat distribution of the eigenvalues, whereas under $\mathcal{H}_1$ the distribution of the eigenvalues will depend on the rank of $\mathbf{H}$. Finally, because slow-fading scenarios have lower variability, the eigenvalues of $\hat{\mathbf{R}}$ will exhibit large dispersion as some of them will approach zero, whereas in fast-fading scenarios the dispersion will be lower.

6.3 Frequency-Domain Detection

6.3.1 Problem Formulation

This Section addresses the detection of a digital pulse-shaped modulated signal as

$$\mathcal{H}_0 \quad : \quad x(t) = w(t) \tag{6.10a}$$

$$\mathcal{H}_1 \quad : \quad x(t) = s(t) + w(t), \tag{6.10b}$$

where $w(t)$ is a WSS circular noise of PSD $N_0/2$ inside the band of interest, and $s(t)$ is given by

$$s(t) = \sqrt{\gamma T} h(t) * \sum_n a[n] p(t - nT). \tag{6.11}$$

In (6.11), $\gamma > 0$ is the signal strength, T is the symbol interval, $h(t)$ is the complex propagation channel, $a[n]$ is the unit variance symbol sequence, and $p(t)$ is the unit energy modulation pulse. This signal model is valid for a wide class of digital modulations such as pulse amplitude modulation (PAM), quadrature amplitude modulation (QAM) and amplitude phase-shift keying (APSK). The two-sided bandwidth of $s(t)$ is

$$B = \frac{1+\alpha}{T}, \tag{6.12}$$

where $0 < \alpha < 1$ is the excess bandwidth or roll-off parameter.

6.3.2 Cyclostationary Signal Background

This section summarizes the main concepts behind cyclostationarity. The reader is referred to [SD98, SS10, Gar86, SPSG05, GNP06] for more details on this subject. The cyclic spectral density (CSD) of a process $s(t)$ is defined as the cross spectral density

$$\phi_s^\alpha(f) \doteq \lim_{T_0 \to \infty} E\left[S_{T_0}\left(f + \frac{\alpha}{2}\right) S_{T_0}^*\left(f - \frac{\alpha}{2}\right) \right], \tag{6.13}$$

where $S_{T_0}(f)$ is the normalized finite-size Fourier transform of a T_0-size realization of the process. Assuming stationary zero-mean uncorrelated symbols, the CSD of (6.11) is given as

$$\phi_s^{l/T}(f) = \gamma G\left(f + \frac{l}{2T}\right) G^*\left(f - \frac{l}{2T}\right), \tag{6.14}$$

for $l \in \mathbb{Z}$ and 0 for $l \notin \mathbb{Z}$, and $G(f)$ is

$$G(f) = H(f)P(f) \tag{6.15}$$

i.e., the product between the channel frequency response and the modulation pulse Fourier transform. The factorization (6.14) holds even in the presence of a frequency-selective channel, as it is consequence of the pulse-shaped modulation structure of (6.11).

6.3.3 Asymptotic Frequency-Domain Signal Detection

Now, consider the Fourier transform of M blocks of length T of $x(t)$, namely, $X_{MT}(f)$. In vector notation, the frequency-domain observation vector is defined as

$$\mathbf{x}(\upsilon) \doteq X_{MT}(\mathbf{f}(\upsilon)), \tag{6.16}$$

where $\mathbf{f}(\upsilon)$ scans the sensing interval $(-B/2, B/2)$ with small sampling intervals of $L \doteq \lceil BT \rceil$ samples through the auxiliary variable $-1/(2T) \leq \upsilon \leq 1/(2T)$, i.e.,

$$\mathbf{f}(\upsilon) = \upsilon\mathbf{1} + \underbrace{\frac{1}{T}\begin{pmatrix} (L-1)/2 \\ \vdots \\ -(L-1)/2 \end{pmatrix}}_{\doteq \mathbf{z}}. \tag{6.17}$$

Under this notation, a consequence of (6.14) and the stationarity of the noise is that the second-order statistics of (6.16) obey

$$\mathbb{E}\left[\mathbf{x}(\upsilon+\delta)\mathbf{x}^H(\upsilon)\right] \to \begin{cases} \mathbf{\Phi}_x(\upsilon) & \delta = 0 \\ \mathbf{0} & \delta \neq 0 \end{cases}, \tag{6.18}$$

where the entries of the $L \times L$ spectral matrix $\mathbf{\Phi}_x$ are related to the CSD of $x(t)$ by

$$[\mathbf{\Phi}_x]_{i,j\in\{1,\ldots,\mathrm{L}\}} = \phi_x^{f_i(\upsilon)_i - f_j(\upsilon)}\left(\frac{f_i(\upsilon) + f_j(\upsilon)}{2}\right), \tag{6.19}$$

where $f_i(\upsilon)$ is the i-th element of $\mathbf{f}(\upsilon)$. By introducing a vectorial spectral process to treat cyclostationary signals, the incorrelation property (6.18) is hereby unveiled. In fact, this constitutes and extension of the frequency-domain treatment of WSS signals to cyclosta-

tionary signals, just by extending the dimension of the frequency-domain process.

It is a well-known result that for a fixed value of f, $S_{T_0}(f)$ has asymptotic normality as $T_0 \to \infty$ [Bri81]. As a result of, the log-likelihood function of the frequency-domain observation vector (6.16) can be written as

$$L(\mathbf{x}) = -\int_{1/T} \ln\det \mathbf{\Phi}_x(\upsilon) + \mathbf{x}^H(\upsilon)\mathbf{\Phi}_x^{-1}(\upsilon)\mathbf{x}(\upsilon)d\upsilon. \tag{6.20}$$

An advantage of this formulation is that the matrix involved in (6.20) exhibits low dimensionality L, independent from the data size. Hence, mathematical and computational simplicity will be attained compared with the time-domain approach which typically involves the estimation of cumbersome covariance matrices, e.g., as in [DG94]. Using (6.14), the spectral matrices under both hypotheses read

$$\mathcal{H}_0 \quad : \quad \mathbf{\Phi}_x(\upsilon) = N_0\mathbf{I}, \tag{6.21a}$$

$$\mathcal{H}_1 \quad : \quad \mathbf{\Phi}_x(\upsilon) = N_0\mathbf{I} + \gamma\mathbf{g}(\upsilon)\mathbf{g}^H(\upsilon), \tag{6.21b}$$

where $\mathbf{g}(\upsilon) = G(\mathbf{f}_\upsilon) \odot e^{j2\pi\varepsilon\mathbf{z}}$, being $0 \leq \varepsilon \leq 1$ the time-delay or timing parameter, and $\mathbf{z}$ has been defined in (6.17).

It is noted that as a result of the frequency-domain formulation of the likelihood function, the general problem posed in (6.10) is translated to the specific problem (6.21). Here, the important feature of (6.21) is that the signal component has rank-1 structure, while the noise component is full-rank. This low rank nature of the signal to be detected is what provides the basis of the detection of cyclostationary signals developed in the sequel.

6.3.4 GLR Detection over Frequency-Selective Fading Channels

It is known that in the presence of unknown frequency-selective channels, the performance of cyclostationary-based detection methods is severely downed [TS07]. This is due to the fact that the channel can attenuate or even kill the specific cyclostationary feature which is exploited for detection. However, as stated in the former section, pulse-shaped modulated signals preserve the factorization property (6.14) even in frequency-selective channels. This means that the rank-1 structure is still present in the received waveform as indicated in (6.21), but with a modified pulse shaped according to (6.15). Under the asymptotic frequency-domain formulation, this is modeled as

$$\mathbf{g}(\upsilon) = \sqrt{g(\upsilon)}\tilde{\mathbf{g}}(\upsilon), \tag{6.22}$$

i.e., $g(\upsilon) \doteq \|\mathbf{g}(\upsilon)\|$ models the frequency-dependent arbitrary gain, $\tilde{\mathbf{g}}(\upsilon)$ is a unit-norm vector. If the coherence bandwidth of the channel is known in the form $B_c = 1/(KT)$, a

K-parameterization of (6.22) is proposed as

$$\mathbf{g}(\upsilon) = \sqrt{g_k}\tilde{\mathbf{g}}_k, \tag{6.23}$$

for $\upsilon \in I_k$, where

$$\mathcal{W}_k = \left(\frac{-1/2 + (k-1)/K}{T}, \frac{-1/2 + k/K}{T} \right), \tag{6.24}$$

for $k = 1, \ldots, K$. Note that the key assumption in the previous model is that every entry of the frequency vector (6.23) remains constant for values of υ which are within the coherence bandwidth. Under this model, the eigenvalue decomposition (EVD) of the spectral coherence matrices under $\mathcal{H}_1$ within the range (6.24) is

$$\mathbf{\Phi}_x(\upsilon) = \underbrace{N_0\mathbf{I} + g_k\tilde{\mathbf{g}}_k\tilde{\mathbf{g}}_k^H}_{\doteq \mathbf{U}_k\mathbf{\Lambda}_k\mathbf{U}_k^H} \tag{6.25}$$

where the eigenvector matrix $\mathbf{U}_k$ has been defined as $\mathbf{U}_k = \left[\tilde{\mathbf{g}}_k, \mathbf{U}_k^{\perp}\right]$, and $\mathbf{\Lambda}_k$ is given as

$$\mathbf{\Lambda}_k = \begin{bmatrix} N_k & \mathbf{0} \\ \mathbf{0} & N_0\mathbf{I}_{L-1} \end{bmatrix}. \tag{6.26}$$

Because this Section addresses GLR detection with unknown white noise PSD and unknown frequency-selective channel, the set of nuisance parameters for the log-likelihood function (6.20) are

$$N_0, \quad N_k = N_0 + \gamma_k \quad \text{and} \quad \tilde{\mathbf{g}}_k, \tag{6.27}$$

for $k = 1, \ldots, K$. The Neyman-Pearson optimal frequency-domain GLR detector of the second-order cyclostationary signal (6.11) over a frequency-selective channel is derived in Appendix D.3 and is given by

$$\boxed{T(\mathbf{x}) = \frac{1}{\left(\prod_{k=1}^{K}\lambda_k\right)^{1/K}\left(1 - \sum_{k=1}^{K}\lambda_k\right)^{\alpha}} \geq \tau,} \tag{6.28}$$

where λ_k is given as $\lambda_k = \lambda_{\max}\{\mathbf{B}_k\}$, where the matrices $\mathbf{B}_k$ are given as

$$\mathbf{B}_k = \frac{1}{\hat{P}_x}\int_{\mathcal{W}_k} \mathbf{x}(\upsilon)\mathbf{x}^H(\upsilon)d\upsilon. \tag{6.29}$$

and $\hat{P}_x$ is an estimate of the total received signal power.

Here, the matrices $\mathbf{B}_k$ in (6.29) are short-band estimates of the spectral covariance matrix of the received signal normalized by the total received signal power. The entries of these matrices can be seen as samples of the frequency-smoothed cyclic periodogram of the signal. The interpretation of the detector (6.28) goes as follows. While the first term of

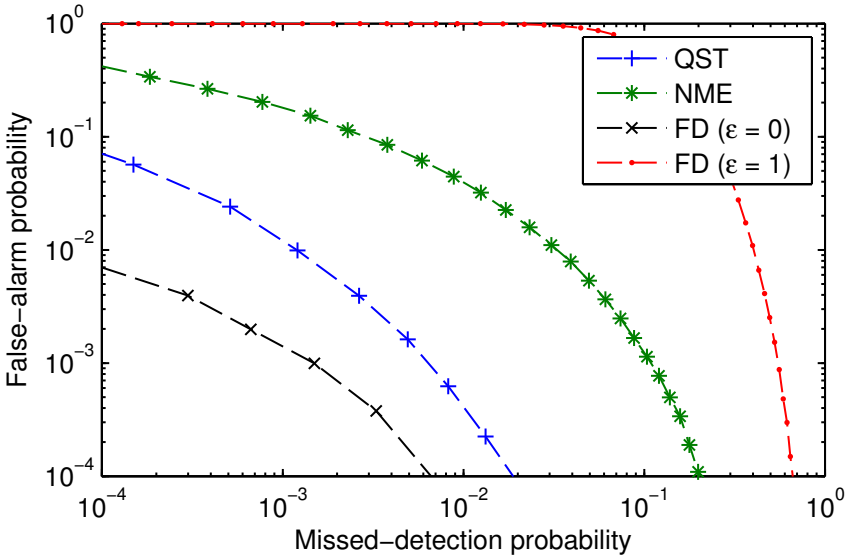

Figure 6.2: ROC of the QST, DF, and NME in a slow-fading scenario.

the denominator implements a classical measure of inter-band sphericity (a term that refers to flatness), the second term of the denominator is in charge of measuring the intra-band sphericity, i.e., the lack of spectral correlation. The second term has no influence for $\alpha = 0$, whereas its relative importance increases within the excess band. As a whole, the value of the detector increases when high spectral correlation, i.e., high eigenvalue dispersion in $\mathbf{B}_k$, is assessed, and/or when the the measured spectrum is far from white.

6.4 Numerical Results

6.4.1 Quadratic Sphericity Test

In this Section, numerical results are provided to assess the performance of the QST (6.9) in terms of the receiver operating characteristics (ROC) curve (missed-detection probability P_{MD} versus false-alarm probability P_{FA}) in several scenarios. As a benchmarking, the following additional signal detectors are similarly considered. On the one hand, the feature detector (FD) which is given by the invariant detector (6.6) with side information of the channel coefficients $\hat{\mathbf{H}}$, i.e.,

$$T_{\mathrm{FD}}(\mathbf{x}) = \frac{1}{\mathbf{x}^H\mathbf{x}} \frac{\mathrm{tr}(\hat{\mathbf{H}}\hat{\mathbf{H}}^H\hat{\mathbf{R}})}{\|\hat{\mathbf{H}}\hat{\mathbf{H}}^H\|_F} \geq \tau, \tag{6.30}$$

with

$$\hat{\mathbf{H}} = \sqrt{\frac{1}{1+\epsilon^2}}\mathbf{H} + \sqrt{\frac{\epsilon^2}{1+\epsilon^2}}\mathbf{\Delta}, \tag{6.31}$$

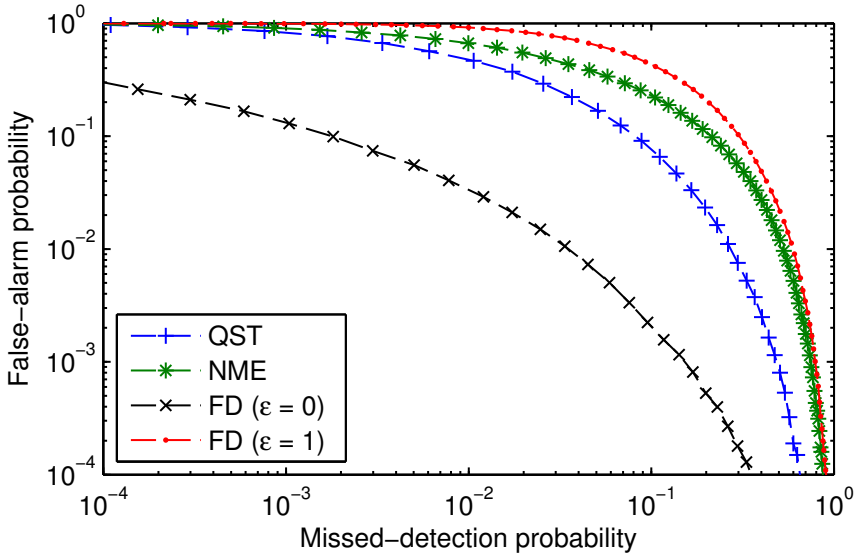

Figure 6.3: ROC of the QST, DF, and NME in a fast-fading scenario.

where $\mathbf{H}$ is the true channel matrix, and $\mathbf{\Delta}$ is uniformly distributed within the range $\|\mathbf{\Delta}\| \leq 1$ to model the error incurred in the estimation of the channel in pilot-based cognitive radio networks. Remarkably, both the perfectly known case $\epsilon = 0$, and the FD with channel uncertainty $\epsilon = 1$ are considered. On the other hand, the NME , given by [ZLZ08]

$$T_{\text{NME}}(\mathbf{x}) = \frac{\lambda_1(\hat{\mathbf{R}})}{\sum_{i=1}^{L} \lambda_i(\hat{\mathbf{R}})} \geq \tau, \tag{6.32}$$

is included as well.

In particular, the simulations contemplate the transmission of K QPSK symbols over a fading channel $\mathbf{H}$ containing K realizations of the channel coefficients according to the block fading model [BPS98], i.e., throughout K channel realizations, it is assumed that the channel response $\mathbf{h}$ remains constant through T realizations and then changes independently to another value. In the sequel, the slow fading scenario ($T \gg K$), and the fast fading scenario ($1 \leq T \ll K$) are considered. The local receiver operates at a SNR of 5 dB and acquires $N = 256$ samples. A channel of length $L = 12$ unit variance i.i.d. lags is employed.

The ROC of the QST (6.9), the FD (6.30) with $\epsilon = 0$ and $\epsilon = 1$, as well as the NME (6.32) in a slow-fading and fast-fading ($T = 5$) scenarios are depicted in Figure 6.2 and Figure 6.3, respectively. As it can be appreciated, the FD with channel errors incurs a severe performance loss compared to the FD without errors in both scenarios due to the mismatching between the true channel and the available side information. On the other hand, both QST and NME outperform the FD with errors. In general, it is seen how the proposed QST has a performance gain with respect to the NME , as the NME is only concerned in extracting a channel feature corresponding to the largest eigenvalue. Finally, whereas detecting the

channel structure in fast-fading scenarios is a challenging task for both QST and NME , in the slow-fading scenario the QST provides good performance characteristics and incurs a degradation relative to the FD with perfectly known channel smaller than the NME .

6.4.2 Frequency-Domain Detection

The performance of the proposed detector (6.28), referred as GLRT-K, is evaluated by Monte Carlo simulations using 100,000 realizations per scenario. The transmission is modeled by M QPSK symbols using root raised cosine (RRC) pulses. A tapped delay line channel model [BPS98] with exponentially decaying power profile is considered with mean delay-spread proportional to $1/B_c$ symbols, such that is remains constant in each realization of MT seconds. In all Figures, $M = 256$, $K = 8$ and $\alpha = 0.8$.

For comparison, the following detectors are evaluated in all the scenarios. First, the well-known multi-cycle spectral correlation magnitude detector (MCSCMD-K) proposed by Gardner [GS93] with $\ell_0 = 1$. Second, the GLR detector not exploiting the frequency-selective nature of the fading channel, i.e., considering $\mathbf{g}(\upsilon) = \sqrt{g}\tilde{\mathbf{g}}(\upsilon)$ instead of (6.22). This yields

$$T_{\mathrm{GLRT}-1}(\mathbf{x}) = \frac{1}{\lambda(1-\lambda)^{\alpha}} \geq \tau, \tag{6.33}$$

with $\lambda = \hat{P}_s/\hat{P}_T$ [RFSVV14]. It is noted that this corresponds to the GLRT-1, i.e., (6.28) particularized for $K = 1$. This detector is considered to illustrate the lack of robustness in presence of a frequency-selective channel. Third, the LMPIT-K, i.e., the Frobenius norm of a normalized version of the autocorrelation matrix

$$T_{\mathrm{LMPIT}-K}(\mathbf{x}) = \sum_{k=1}^{K} \|\mathbf{B}_k\|_F^2 \geq \tau. \tag{6.34}$$

This detector has been proposed [LdV01, FSRVV13] and proved [RVSS13] to be optimal for close hypotheses, i.e., low SNR. This is the main reference for comparison, as it assumes identical prior knowledge.

Figure 6.4 depicts the complementary ROC when the channel has a mean delay-spread of $\tau = 1$ symbol and illustrates the high sensitivity of GLRT-1 in front of an unknown channel response. In contrast, the performance of the proposed GLRT-K detector exhibits a robust behavior in detection performance in front of unknown frequency-selective channels. The missed-detection probability of the detectors versus the SNR is shown in Figure 6.5, where the thresholds have been set to satisfy a false-alarm level of 0.05. It is rapidly observed that the error exponent associated to this error probability, i.e., the scaling of the detection performance with the SNR conditions, is significantly improved by the proposed GLRT-K. Figure 6.6 illustrates the main advantage of the proposed detector by showing the missed-detection probability versus the mean delay-spread, with a false-alarm level of 0.05. It is rapidly appreciated that small values of delay-spread cause a significant degrada-

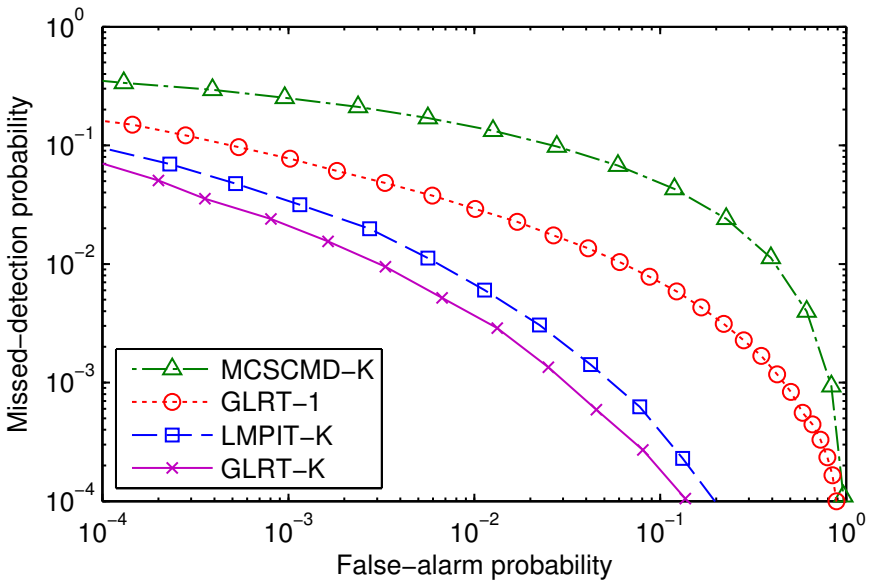

Figure 6.4: Complementary ROC with a delay-spread of 1 symbol and SNR of -0.5 dB.

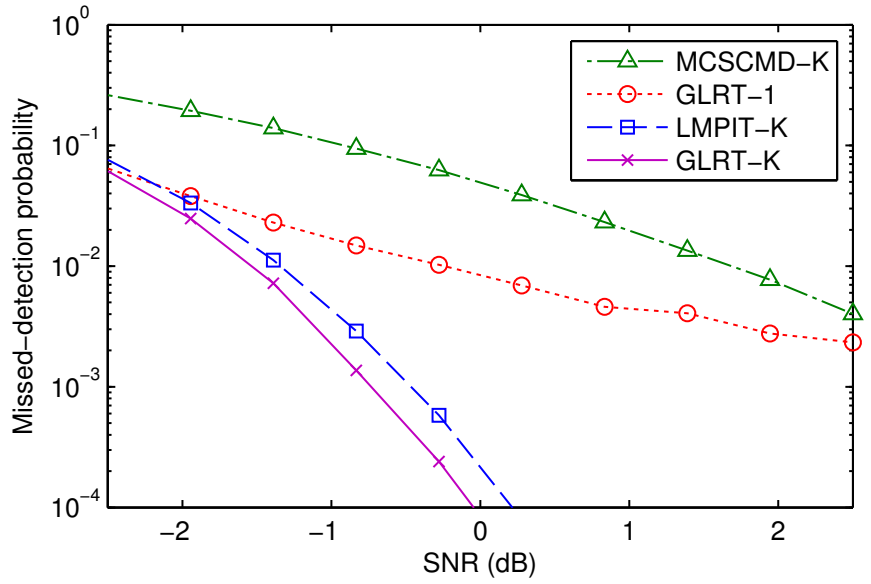

Figure 6.5: Missed-detection probability versus SNR for a false-alarm level of 0.05 with a delay-spread of 1 symbol.

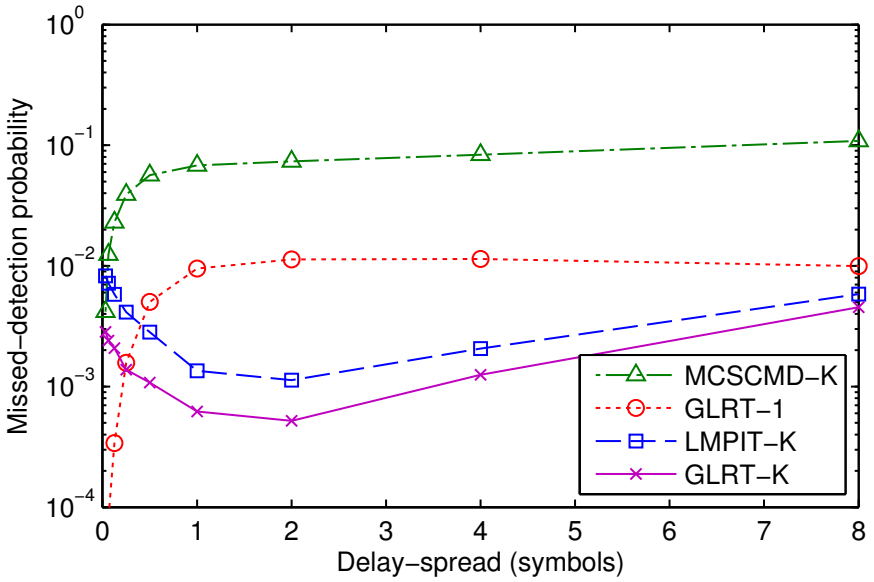

Figure 6.6: Missed-detection probability versus delay-spread for a false-alarm level of 0.05 with and SNR of -0.5 dB.

tion in the detectors' performance. In contrast, the GLRT-K tolerates higher delay-spread and shows a degradation only when the delay-spread is of the same order as the number of sub-bands, i.e., K. Furthermore, the performance of the GLRT-K improves for small values of the delay-spread. This is due to the fact that the GLRT-K is partly sensitive to the lack of sphericity measured by the geometrical mean term in the expression (6.28). Hence, in contrast to other cyclostationary detectors, the proposed detector does not ignore the stationary spectral component of the signal, which is less flat in the presence of a frequency-selective channel. Finally, compared to the benchmarking LMPIT-K, the proposed GLRT-K provides a systematic gain in the simulated scenarios, whereas the performance of the LMPIT-K approaches that of GLRT-K for asymptotically low SNR, as expected and confirmed in Figure 6.6.

6.5 Conclusions

In this Chapter, the QST for blind detection of WSS signals in time-varying frequency-selective fading channels has been introduced. By deriving a correlation-matching optimal invariant test statistic, it has been shown that the QST is the solution in the blind case. The QST evaluates the ratio between the squared mean and the arithmetic mean of the eigenvalues of the autocorrelation matrix. Numerical results have been reported to assess and benchmark the performance of the QST in slow and fast-fading scenarios. In slow-fading, the proposed blind QST exhibits good performance and incurs smaller degradation compared to the NME .

On the other hand, the GLR detection of a cyclostationary signal is addressed by ex-

ploiting the rank-1 structure of small spectral covariance matrices in a novel frequency-domain representation of digital waveforms. By incorporating a K parameterized frequency-selective channel model based on the coherence bandwidth, the proposed GLRT-K is derived. Numerical results have assessed the detection performance and the robustness in front of frequency-selective channels.

Chapter 7

Noise Subspace Communication

7.1 Introduction

This last Chapter focusses on the design of the transmission scheme of the secondary users. The goal is to design the secondary transmitted signal Y from the local noisy observations X of the primary signal S, as shown in Figure 7.1.

7.1.1 Cyclostationary Secondary Transmission

Firstly, the design of the secondary transmitted signal is addressed with the objective of mitigating the interference caused to other secondary users, while keeping orthogonality to the primary users. More specifically, the problem of spectrum sensing performed by a silent non intended cognitive receiver in the presence of an active cognitive transmitter is considered.

Because the silent receiver does not decode the information, the transmitted signal acts as strong interference to the spectrum sensing process. It is known that cyclic feature detectors show robustness in front of noise uncertainty and have gained much recent attention [SND08, LKHP09, LKK10]. However, acquiring the cyclic statistics requires knowledge on the primary cyclic frequencies. Here, a primary spectrum sensing detector which exploits the statistical features of cyclostationarity is developed, albeit from a different context that

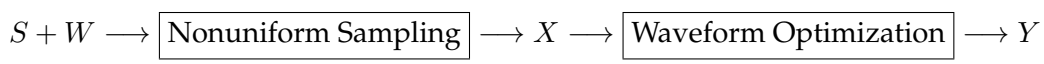

Figure 7.1: Waveform optimization problem.

the aforementioned methods. Specifically, it is considered that spectrum sensing is performed by a silent cognitive user in a cognitive radio network when a cognitive transmitter is active. A silent cognitive user is referred to as a user who is not transmitting nor receiving information from the active transmitter, but it is synchronized with the cognitive radio network. If the cognitive radio network is localized in a small area, the transmitting user may act as strong interference to the non intended receivers, making spectrum sensing a further challenging problem. By operating directly on the oversampled cyclostationary signal at the receive antenna, the silent cognitive user is able to exploit the cyclostationarity in order to identify a noise subspace where primary spectrum sensing can be performed free of interference [RVV10]. Because the key signal property to be exploited is related with the cyclic spectral density of the cognitive signal, the noise subspace detector is formulated rather in the frequency-domain. In order to evaluate the performance of the proposed detector, the deflection [Pic95] is used as a performance indicator. Although the formulation of the problem is general, the derived results are particularized to the realistic scenario in which the primary signal has flat spectrum within the detection interval, e.g., a terrestrial digital video broadcasting (DVB-T), and in which the active cognitive transmitter employs a squared-root raised cosine (SRRC) pulse. The deflection reveals that the performance of the noise subspace detector is proportional to the roll-off factor of the SRRC pulse, and proportional to the squared of the signal-to-noise ratio (SNR), defined as the ratio between the primary signal average power and the average noise power at the silent cognitive antenna. This results demonstrates that the proposed detector operates in the free-interference regime, whereas any of the spectrum sensing methods mentioned above would see a strong interference.

It is worth noting that the spectrum sensing problem formulated in the presence of other non-primary interference has received little attention in the literature. For further information on this topic, the reader is referred to the work [MH13].

7.1.2 Minimum Description Length

One important problem when designing secondary waveforms is the determination of the noise subspace dimension, i.e., the amount of available resources for secondary communication orthogonal to the primary system. This problem is cast as a model order selection problem.

The minimum description length (MDL) is the most commonly adopted model order selection statsitic [SS04]. The model order selection problem consists of determining the dimension of the parameter vector of the data model, which has many applications in a wide range of signal processing problems. This work focusses on the model order selection of the dimension of the second-order statistics of a signal. This particular scenario has practical applications in the cognitive radio context [MM99], where the dimension of the unused resources by the primary users has to be determined [HTR09].

More precisely, the model order selection problem is formulated for a zero-mean Gaussian signal with unknown low rank correlation matrix immersed in zero-mean white Gaussian noise with unknown noise variance. Because a likelihood ratio is involved in the MDL statistic, the signal correlation matrix and the noise variance are incorporated as nuisance parameters. This leads to the formulation of the generalized likelihood ratio (GLR) statistic, which has been recently reported as the sphericity test for low rank Gaussian signals [RVVLV$^+$11]. Because the GLR asymptotically follows a non-central Chi-squared distribution, the statistical characterization of the proposed sphericity MDL is further addressed in this Chapter.

The problem of model order selection of Gaussian random variables with unknown parameters has been addressed in [SL07, GS08b, BGH09] employing least squares estimation (LSE), instead of maximum likelihood (ML) estimation.

7.1.3 Rotationally Invariant Minimum Norm Waveform Optimization

The last research of this thesis seeks secondary waveforms that satisfy the favorable properties of white frequency response, invariance to rotation, and that rely only on the second-order statistics of the secondary-to-primary interference channel. Signals with white spectrum exhibit good properties in terms of entropy and residual correlation, whereas invariance to rotation allows noncoherent detection. The benefits of second-order channel state information (CSI) are straightforward: it represents a slower time-varying statistic compared to first-order CSI, and it further exhibits quantization advantages.

Under these requirements, it is shown that linear predictors that arise from minimum norm filtering exhibit the white frequency response and rotation invariance rotation. The resulting waveforms implement a form of null space communication similar to that in multiple antennas (e.g., see [YHR$^+$09]), but will show enhanced stability and low complexity benefits as it is based on the temporal correlation rather than the spatial correlation. In front of statistical CSI mismatching, a robust waveform optimization problem is also addressed. The robust waveform optimization is based on the worst-case performance scenario initially motivated by Gershman *et. al* in [SGLW03]. The worst-case performance scenario regularizes the objective function by adding a penalty term which is equivalent to the squared value of the spectrum of the optimization waveform, which is a well-known result in spectral estimation theory. Both the nonrobust and robust waveform optimization problems are solved in an iterative fashion, and show efficient implementation in the asymptotic frequency-domain.

7.1.4 Chapter Organization

The rest of the Chapter is organized as follows. The design of cognitive cyclostationary waveforms is addressed in Section 7.2. The obtention of second-order statistics of the

primary system by means of the MDL problem is addressed in Section 7.3. Section 7.4 presents the minimum norm waveforms based on the second-order statistics of the primary system, in both nonrobust and robust versions. Numerical results are reported in Section 7.5, whereas the conclusions are outlined in Section 7.6.

7.2 Cognitive Cyclostationary Waveform

7.2.1 Noise Subspace Induced By Cylostationarity

This Section considers a cognitive radio network consisting of $Q = 3$ cognitive users: an active cognitive transmitter, an active cognitive intended received, and a silent cognitive unintended receiver. The spectrum sensing problem will be addressed in the AWGN channel at the antenna of the silent cognitive unintended receiver. Let $y(t)$, $w(t)$ and $s(t)$ denote the active cognitive transmitted signal, the AWGN and the primary signal to be detected, respectively. The cognitive signal admits the following linear modulation form:

$$y(t) = \sum_{k} a[k]g(t - \varepsilon - kT), \tag{7.1}$$

being $a[k]$ the complex zero-mean uncorrelated symbols, ϵ the timing parameter, T the symbol period and $g(t)$ the unit-energy pulse shape. Because the silent user belongs to the same cognitive radio network, it has perfect knowledge on the pulse shape and timing parameter ε involved in (7.1). The frequency response of $g(t)$ is given by $G(f)$, with spectral support

$$\mathcal{W}_G = \left(-\frac{1+\alpha}{2T}, \frac{1+\alpha}{2T} \right), \tag{7.2}$$

where α is the roll-off factor that determines the excess of bandwidth with respect to the Nyquist bandwidth $1/T$.

Because this Section is motivated to exploit the statistical features of cyclostationarity available in $y(t)$, the MT-size Fourier transform of the complex received signal is considered as

$$X(f) = S(f) + W(f) + Y(f), \tag{7.3}$$

where $Y(f)$, $W(f)$ and $S(f)$ are the finite-size Fourier transforms of the cognitive signal, noise and primary signal components, respectively, and M is the signal block length. In order to display the inherent second-order periodicity of $y(t)$, the following normalized two-dimensional frequency-domain observation vector is adopted [RVV10]

$$\mathbf{o}_x(\nu) \doteq \frac{1}{\sqrt{MT}} \begin{bmatrix} X(\nu_+) \\ X(\nu_-) \end{bmatrix}, \tag{7.4}$$

where the frequency pair $\nu_{\pm}$, defined through the auxiliary variable ν as

$$\nu_{\pm} \doteq \nu \pm \frac{1}{2T}, \tag{7.5}$$

indicates the correlation of the spectral components of $y(t)$, which are separated by the fundamental cycle frequency $1/T$ [Gar90]. In the cognitive radio network context addressed in this Section,

$$\mathbf{o}_x(\nu) = \mathbf{o}_y(\nu) + \mathbf{o}_w(\nu) + \mathbf{o}_s(\nu), \tag{7.6}$$

where $\mathbf{o}_y(\nu)$, $\mathbf{o}_w(\nu)$ and $\mathbf{o}_s(\nu)$ have been similarly defined as (7.4) for the cognitive signal, noise and primary signal components, respectively. Assuming uncorrelated components[1], the second-order statistics of the observation vector $\mathbf{\Phi}_x(\nu) \doteq \mathbb{E}\left\{\mathbf{o}_x(\nu)\mathbf{o}_x^H(\nu)\right\}$ may be written as

$$\mathbf{\Phi}_x(\nu) = \mathbf{\Phi}_y(\nu) + \mathbf{\Phi}_w(\nu) + \mathbf{\Phi}_s(\nu), \tag{7.7}$$

where $\mathbf{\Phi}_y(\nu)$, $\mathbf{\Phi}_w(\nu)$ and $\mathbf{\Phi}_s(\nu)$ are the spectral coherence matrices of the cognitive signal, noise and primary signal components, respectively. The structure of the former matrices is the enabling key factor of the proposed detector.

On the one hand, because $y(t)$ is a pulse amplitude modulation (PAM), its spectral coherence matrix will exhibit a low rank structure. For signals having an excess of bandwidth $\alpha \leq 1$ and with an associated pulse shape fulfilling the Nyquist criterion of free inter symbol interference (ISI)

$$\frac{1}{T}\sum_{k=-\infty}^{\infty}\left|G\left(f+\frac{k}{T}\right)\right|^2 \tag{7.8}$$

it can be shown that the spectral coherence matrix of the cognitive signal is a rank-one matrix which can be expressed as the self outer product

$$\mathbf{\Phi}_y(\nu) = P\mathbf{g}(\nu)\mathbf{g}^H(\nu), \tag{7.9}$$

where P is the cognitive signal power, and the unitary vector $\mathbf{g}(\nu)$ is related to the frequency response of the pulse shape, the timing parameter and the symbol period through

$$\mathbf{g}(\nu) = \frac{1}{\sqrt{T}}\begin{bmatrix} G\left(\nu_+\right)e^{-j\pi\varepsilon/T} \\ G\left(\nu_-\right)e^{j\pi\varepsilon/T} \end{bmatrix}. \tag{7.10}$$

For a more detailed understanding of the cyclostationary concepts involved in the derivation of (7.9), c.f. [RVV10, Equations (8)–(18)] and the references therein. On the other hand,

[1]This assumption is only valid if the cognitive signal is generated independently on the statistics of the primary signal, e.g., if spectral shaping is not employed.

because the noise and primary signal components are wide-sense stationary (WSS)[2], their spectral coherence matrices can be expressed as

$$\mathbf{\Phi}_w(\nu) = N_0\mathbf{I}, \tag{7.11}$$

and

$$\mathbf{\Phi}_s(\nu) = \begin{bmatrix} \phi_s(\nu_+) & 0 \\ 0 & \phi_s(\nu_-) \end{bmatrix}, \tag{7.12}$$

respectively, where N_0 is the power spectral density (PSD) of the AWGN.

7.2.2 Primary Spectrum Sensing

The detection of the primary signal is addressed from the second-order statistics of the observation vector exploiting the noise subspace induced by the cyclostationary cognitive signal. Therefore, a signal detector which jointly operates with the transmission and reception of cognitive signals is proposed, without resorting to silent periods to perform detection. In particular, the signal detection problem is cast as the following binary hypotheses testing problem

$$\mathcal{H}_0 \;:\; \mathbf{\Phi}_x(\nu) = \mathbf{\Phi}_y(\nu) + \mathbf{\Phi}_w(\nu) \tag{7.13a}$$

$$\mathcal{H}_1 \;:\; \mathbf{\Phi}_x(\nu) = \mathbf{\Phi}_y(\nu) + \mathbf{\Phi}_w(\nu) + \mathbf{\Phi}_s(\nu), \tag{7.13b}$$

where the spectral coherence matrices are given by (7.9), (7.11) and (7.12) derived above.

The matrix $\mathbf{\Phi}_y(\nu)$ is seen as interference in the spectrum sensing process and it is present in both hypotheses in (7.13). Because the cognitive signal space is spanned by the unitary vector $\mathbf{g}(\nu)$, the proposed detector projects the observation vector onto the space orthogonal to the signal space, namely the noise subspace, by means of the orthogonal unitary vector that spans the noise subspace. That is, the *noise subspace detector* is given by

$$\boxed{T_{\mathrm{NS}}(\mathbf{x}) = \frac{1}{|\mathcal{W}|} \cdot \int_{\mathcal{W}} \left|\mathbf{g}_\perp^H(\nu)\mathbf{o}_x(\nu)\right|^2 d\nu \geq \tau,} \tag{7.14}$$

where the integration is done over a frequency region $\mathcal{W}$, τ is the detection threshold for which $\mathcal{H}_1$ is decided in front of $\mathcal{H}_0$, and $\mathbf{g}_\perp(\nu)$ is the unitary vector that spans the orthogonal space to $\mathbf{g}(\nu)$, and is related to the frequency response of the pulse shape, the timing parameter and the symbol period as

$$\mathbf{g}_\perp(\nu) = \frac{1}{\sqrt{T}} \begin{bmatrix} G^*(\nu_-)\, e^{-j\pi\varepsilon/T} \\ -G^*(\nu_+)\, e^{j\pi\varepsilon/T} \end{bmatrix}. \tag{7.15}$$

[2]Even though the primary signal may exhibit cyclostationarity, it is assumed that it is not exploited at the cognitive receiver due to the lack of synchronism.

Because the spectral support of the cognitive signal is $\mathcal{W}_G$ given in (7.2), the detection interval $\mathcal{W}$ is restricted by the possible values that the frequency variable ν can take:

$$\mathcal{W} = \left(-\frac{\alpha}{2T}, \frac{\alpha}{2T}\right). \tag{7.16}$$

Therefore, the detection interval will be given by the spectral support of $\mathcal{W}$, i.e.,

$$|\mathcal{W}| = \frac{\alpha}{T}. \tag{7.17}$$

In the following, the statistical properties of $T_{\mathrm{NS}}(\mathbf{y})$ are derived in order to illustrate how the projection onto the noise subspace is able to cancel out the cyclostationary cognitive component from the observation vector, as well as to demonstrate that the conditions in which the detector operates are related to the SNR, defined as the ratio between the primary signal power and the noise power, instead of signal-to-interference-plus-noise ratio (SINR).

7.2.3 Deflection Performance

In order to evaluate the performance of the proposed detector, as well as to investigate the parameters of the cognitive and primary signals involved in the detection process, this Section makes use of the deflection [Pic95] as a performance indicator of the test $T_{\mathrm{NS}}(\mathbf{y})$. The motivation of employing deflection is twofold. On the one hand, it is a *one number* instead of a curve, hence it provides simpler and more concise interpretations. And on the other hand, maximizing the deflection becomes asymptotically equivalent to optimizing the performance calculated in terms of the ROC. The deflection associated to a detector L is given by

$$\mathbb{D}(T) = \frac{\left(\mathbb{E}_1[T] - \mathbb{E}_0[T]\right)^2}{\mathbb{V}_0[T]}, \tag{7.18}$$

where $\mathbb{E}_0$ and $\mathbb{E}_1$ evaluate the expectation under $\mathcal{H}_0$ and $\mathcal{H}_1$, respectively; and $\mathbb{V}_0$ evaluates the variance under $\mathcal{H}_0$. In Appendix E.1 it is shown that the deflection of the noise subspace detector is given by

$$\mathbb{D}\left(T_{\mathrm{NS}}\right) = \frac{1}{N_0^2|\mathcal{W}|}\left(\int_{\mathcal{W}} \mathbf{g}_{\perp}^H(\nu)\mathbf{\Phi}_s(\nu)\mathbf{g}_{\perp}(\nu)d\nu\right)^2. \tag{7.19}$$

Because (7.19) can be regarded as the inner product between two matrices, in Appendix E.2 the Cauchy-Schwarz inequality is applied to obtain the following upper bound

$$\mathbb{D}\left(T_{\mathrm{NS}}\right) \leq \frac{1}{N_0^2}\int_{\mathcal{W}} \mathrm{tr}\left(\mathbf{\Phi}_s^2(\nu)\right) d\nu, \tag{7.20}$$

i.e., the deflection is related to the expression $\mathrm{tr}\left(\mathbf{\Phi}_s^2(\nu)\right)$, which is a fourth-order measure of the primary signal.

Finally, it is worth noting that the upper bound is achieved when the primary signal has flat spectrum, i.e., if

$$\mathbf{\Phi}_s = N_1 \mathbf{I}. \tag{7.21}$$

In such a case, the deflection reduces to the expression

$$\mathbb{D}\{T_{\mathrm{NS}}\} = \frac{\alpha}{T} \cdot \mathrm{SNR}^2, \tag{7.22}$$

where (7.17) has been used, and the SNR has been defined as the ratio between the PSD of the primary signal and noise components, i.e.,

$$\mathrm{SNR} = \frac{N_1}{N_0}. \tag{7.23}$$

On the one hand, the deflection is proportional the squared value of the SNR, therefore the proposed noise subspace detector operates at an SNR conditions free from any interference caused by the cognitive signal present in the cognitive radio network. On the other hand, it rapidly follows that the detection performance increases with the roll-off factor. Even though higher roll-off factors incur lower spectral efficiency due to the excess of bandwidth, higher values of α are also preferred in non data aided systems for timing synchronization purposes [RSV01].

7.3 Sphericity Minimum Description Length

7.3.1 Problem Formulation

The problem of determining the dimension of the primary signal subspace spanned by the eigenvectors of the correlation matrix is cast as follows. Each of the secondary users acquire a data set of LM samples in the silent periods in which the secondary system is inactive. These observations are vectorized in L vectors $\mathbf{x}_l$ of dimension M which follow the discrete-time model

$$\mathbf{x}_l = \mathbf{s}_l + \mathbf{w}_l, \tag{7.24}$$

for $1 \leq l \leq L$, where each realization $\mathbf{s}_l$ and $\mathbf{w}_l$ represent the primary signal and the noise components at secondary user with power P and variance σ^2, respectively. The average SNR in this problem is defined as

$$\mathrm{SNR} \doteq \frac{NP}{\sigma^2}, \tag{7.25}$$

where N is the true dimension of the primary signal. Typically, the primary signal subspace estimation problem involves low SNR regimes, as the secondary system may be far away from the primary system. In the assumption of white noise with unknown noise

variance[3], the primary signal subspace estimation problem requires estimating the following parameters: the noise variance, the dimension of the primary signal subspace, and the eigenvectors associated to the primary signal subspace. Because the estimation of the dimension of the primary signal subspace is a model order selection problem, the formulation of the following MDL statistic is required [Kay98a]

$$\text{MDL}(n) = -\log \text{GLR}(n) + nM\log(L), \tag{7.26}$$

where $\text{GLR}(n)$ is the GLR which jointly estimates the unknown parameters, i.e., the noise variance and the eigenvectors associated to the primary signal space. That is,

$$\text{GLR}(n) = \frac{\max_{\sigma^2,\mathbf{R}_s} p(\mathbf{x}_1,\ldots,\mathbf{x}_L|\mathcal{H}_n)}{\max_{\sigma^2} p(\mathbf{x}_1,\ldots,\mathbf{x}_L|\mathcal{H}_0)}, \tag{7.27}$$

where $\mathcal{H}_n$ and $\mathcal{H}_0$ denote the hypotheses that the primary signal has dimension n or zero, respectively, in (7.24). Solving (7.27) for $\mathbf{R}_s$ and σ^2 for a given primary signal subspace size n derives to the rank-n sphericity test [RVVLV+11]. As a result, for $n < M-1$, the MDL for the primary signal subspace estimation problem is given by the expression

$$\boxed{\text{MDL}(n) = LM\log\frac{\prod_{m=1}^{M}\lambda_m^{1/M}}{\frac{1}{M}\sum_{m=1}^{M}\lambda_m} - L(M-n)\log\frac{\prod_{m=n+1}^{M}\lambda_m^{1/(M-n)}}{\frac{1}{M-n}\sum_{m=n+1}^{M}\lambda_m} + nM\log(L),} \tag{7.28}$$

where $\lambda_1,\ldots,\lambda_M$ are the eigenvalues of the sample covariance matrix

$$\hat{\mathbf{R}} = \frac{1}{L}\sum_{l=1}^{L}\mathbf{x}_l\mathbf{x}_l^H. \tag{7.29}$$

As noted in [RVVLV+11], for $n \geq M-1$ the low rank structure of the signal correlation matrix cannot be exploited, and hence (7.28) particularizes to the arithmetic-to-geometric mean (AGM) detector, i.e.,

$$\text{MDL}(n) = LM\log\frac{\prod_{m=1}^{M}\lambda_m^{1/M}}{\frac{1}{M}\sum_{m=1}^{M},\lambda_m} + nM\log(L). \tag{7.30}$$

From (7.28) and (7.30), the estimation of the primary signal subspace dimension reads

$$\hat{N} = \arg\min_n \text{MDL}(n), \tag{7.31}$$

[3]As in signal detection, the phenomenon of noise uncertainty is an important problem in noise subspace estimation when the noise variance is inaccurately known a priori by the secondary user. Typically the noise variance is considered as nuisance parameter, hence providing robustness to noise uncertainty.

where the search is in the set $n = 0, \ldots, M$.

7.3.2 Statistical Characterization

The statistical characterization of the primary signal subspace estimation problem resorts to identify the statistical properties of the MDL function in (7.31). Because the MDL(n) in (7.26) implements a GLR statistic for $n > 0$, its distribution is asymptotically given by a non-central Chi-square distribution, i.e.,

$$nM \log(L) - \text{MDL}(n) \sim \mathcal{X}^2_{r_n}(\mu_n), \tag{7.32}$$

where r_n are the degrees of freedom and μ_n is the non-centrality parameter. Because the noise variance is a nuisance parameter and MDL(n) involves the estimation of an Toeplitz Hermitian complex matrix of rank n, it follows that the degrees of freedom r_n are given as

$$r_n = M^2 - (M - n - 1)(M - n) - 1. \tag{7.33}$$

On the other hand, the non-centrality parameter μ_n depends on the occupancy of the primary signal, i.e., the true hypotheses $\mathcal{H}_N$. If the primary signal is not present (i.e., if $\mathcal{H}_0$ is true), the distribution is non-central and $\mu_n = 0$ for all $1 \leq n \leq M$. This task under $\mathcal{H}_N$ is a more difficult problem, as it involves the computation of the non-centrality parameter μ_n. Making use of the Jensen's inequality, a good approximation holds for $n \geq N$:

$$\mu_n + r_n \approx L \log \frac{\left(1 + \dfrac{\text{SNR}}{M}\right)^M}{\left(1 + \dfrac{\text{SNR}}{N}\right)^N}. \tag{7.34}$$

It is worth noting that for $n \geq N$, i.e., when the MDL(n) statistic has reached the true primary signal dimension N, the non-centrality parameter does not improve with n. This is due to the fact that the second part of the MDL(n) in (7.28) uncovers no structure in the noise subspace. Even tough, as the $nM \log(L)$ penalty term in (7.26) continues to increase with n, the MDL will "prefer" $n = N$ in front of higher dimensions. Finally, for $n < N$, the non-centrality parameter is affected by a negative term, i.e.,

$$\mu_n + r_n \approx L \log \frac{\left(1 + \dfrac{\text{SNR}}{M}\right)^M}{\left(1 + \dfrac{\text{SNR}}{N}\right)^N} - L \log \frac{\left(1 + \dfrac{N-n}{M-n}\dfrac{\text{SNR}}{N}\right)^{M-n}}{\left(1 + \dfrac{\text{SNR}}{N}\right)^{N-n}}, \tag{7.35}$$

because the second part of the MDL(n) in (7.28) now evaluates a portion of the signal subspace of dimension $N - n$. A last note on MDL is that by construction, MDL$(0) = 0$ in

a deterministic fashion. A sketch of the proof of the non-centrality parameters (7.34) and (7.35) is provided in Appendix E.3.

An important application of the statistical characterization of the primary signal subspace estimation problem is the computation of the error probability. In particular, the statistical characterization of the sphericity MDL allows to determine the probability of selecting n as the dimension of the primary signal subspace is employed, i.e.,

$$P_n \doteq \mathbb{P}\left[\arg\min_n \text{MDL}(n) = n\right], \tag{7.36}$$

for $n = 0, \ldots, M$. In Appendix E.4 it is shown that this probability is given by

$$\begin{aligned} P_n = \prod_{m=0}^{n-1} \text{CDF}_{-\mathcal{X}^2_{r_n - r_m}(\mu_n - \mu_m)}\left((m-n)M\log(L)\right) \\ \times \prod_{m=n+1}^{M} \text{CDF}_{\mathcal{X}^2_{r_m - r_n}(\mu_m - \mu_n)}\left((m-n)M\log(L)\right), \end{aligned} \tag{7.37a}$$

where r_n and μ_n are given in (7.33) and (7.34)–(7.35), respectively, under either $\mathcal{H}_0$ or $\mathcal{H}_N$. The error probability is defined as the complementary probability to the probability of detection, i.e,

$$P_e \doteq 1 - P_N, \tag{7.38}$$

being P_N the probability of detection given by (7.37) for $n = N$ under $\mathcal{H}_N$. As it is appreciated from (7.37), the pairs (n, m) that exhibit smaller $\mu_n - \mu_m$ will contribute to the error probability.

It is worth noting that in cognitive radio an incorrect primary signal subspace detection will produce a different effect, depending if either the estimated dimension is smaller or larger than the true dimension. If $\hat{N} < N$, the secondary user will underestimate the primary signal subspace, hence will cause interference, whereas if $\hat{N} > N$, the secondary user will overestimate the primary signal subspace, hence losing opportunity. However, this Section focusses on the error probability as a single parameter measure of the behavior of the MDL, for sake of clarity.

7.3.3 Problem Formulation with Known Noise Variance

Assuming that the noise variance is known at the receiver, the MDL problem now involves the following GLR, which only accounts for the estimation of the primary signal correlation matrix, i.e.,

$$\text{GLR}(n, \sigma^2) = \frac{\max_{\mathbf{R}_s} p(\mathbf{x}_1, \ldots, \mathbf{x}_L | \mathcal{H}_n)}{p(\mathbf{x}_1, \ldots, \mathbf{x}_L | \mathcal{H}_0)}. \tag{7.39}$$

Similarly to the unknown noise variance scenario, the GLR statistic (7.39) is asymptotically distributed as

$$nM\log(L) - \text{MDL}(n,\sigma^2) \sim \mathcal{X}^2_{r_n}(\mu_n). \tag{7.40}$$

In Appendix E.5 it is shown that, after ML estimating $\mathbf{R}_s$ and compressing the GLR statsitic, the MDL is given by

$$\text{MDL}(n,\sigma^2) = LM\log\left(\prod_{m=1}^{n}\left(\frac{\lambda_m}{\sigma^2}\right)^{1/M}\right) - LM\left(\frac{1}{M}\sum_{m=1}^{n}\left(\frac{\lambda_m}{\sigma^2}-1\right)\right) + nM\log(L), \tag{7.41}$$

for $n \leq M$. In view of (7.40), the MDL with known noise variance has the same r_n parameter as the MDL with unknown noise variance (as the degrees of freedom in $\mathcal{H}_1$ minus the degrees of freedom in $\mathcal{H}_0$ are preserved), whereas the non-centrality parameter can be approximated by making use of the Jensen's inequality analogously to the developments in Appendix E.3, i.e., by

$$\mu_n + r_n \approx Ln\left(\frac{\text{SNR}}{N} - \log\left(1+\frac{\text{SNR}}{N}\right)\right) \tag{7.42}$$

if $n \leq N$, and

$$\mu_n + r_n \approx LN\left(\frac{\text{SNR}}{N} - \log\left(1+\frac{\text{SNR}}{N}\right)\right) \tag{7.43}$$

if $n > N$. If there is a mismatching between the *prior* σ^2 and the *true* σ^2, then the ratios λ_m/σ^2 in (7.30) will produce random shifts in (7.42)–(7.43) in average as

$$\mathbb{E}\left[\frac{\lambda_m}{\hat{\sigma}^2}\right] = \frac{P+\sigma^2}{\hat{\sigma}^2} \neq 1 + \frac{\text{SNR}}{N}. \tag{7.44}$$

The noise uncertainty in the model order selection problem requires a different analysis than that of signal detection. In particular, the SNR shift due to noise variance mismatching will affect in the same way to all the MDLs, and therefore the key point in the analysis is the evaluation of the parameters in P_n (7.37). As an example, the worst-case for a uniformly distributed noise variance with uncertainty $\delta \geq 1$, i.e.,

$$\sigma^2 \sim U\left(\frac{\hat{\sigma}^2}{\delta}, \delta\hat{\sigma}^2\right), \tag{7.45}$$

corresponds to the case $\hat{\sigma}^2 = \delta\sigma^2$ (i.e., the receiver has a prior noise variance which is δ times the true noise variance). In such a case, for instance,

$$\mu_n - \mu_m \approx L(n-m)\left(\frac{1}{\delta}\frac{\text{SNR}}{N} + \frac{1-\delta}{\delta} - \log\left(\frac{1}{\delta}\left(\frac{\text{SNR}}{N}+1\right)\right)\right), \tag{7.46}$$

which decreases with δ. The specific behavior of the MDL with known noise variance with respect to the noise uncertainty δ and the error probability is numerically evaluated in Section 7.5.2.

7.4 Minimum Norm Waveform Optimization

7.4.1 Problem Formulation and Requirements

The problem of coexistence of a secondary system with a primary system is addressed in this Section. The secondary system must opportunistically transmit over the unused resources of the primary system, causing no or little interference to the primary system. Hence, the coexistence problem is cast as a waveform optimization problem in which the each waveform os the set of M-samples discrete-time complex secondary waveforms

$$\mathbf{A} = (\mathbf{a}_1, \ldots, \mathbf{a}_K), \tag{7.47}$$

is designed to minimize the average interference caused to the primary system, i.e.,

$$\min_{\mathbf{a}_k} \xi(\mathbf{a}_k), \tag{7.48}$$

iteratively for $k = 1, \ldots, K$. The average interference (7.48) induced by the transmission of the waveforms $\mathbf{a}_k$ is related to the second-order statistics of the primary system, $\mathbf{R}$ as

$$\xi(\mathbf{a}_k) = \mathbf{a}_k^H \mathbf{R} \mathbf{a}_k. \tag{7.49}$$

This means that $\mathbf{R}$ is the minimum required CSI for the waveform optimization problem (7.48). Hence, the secondary waveforms designed in this Section will depend on the second-order statistics of the interference channel (i.e., the eigenvectors and the eigenvalues of $\mathbf{R}$), rather than the complete CSI contained in the primary-to-secondary channel at signal level. The advantages of second-order statistical CSI are twofold. On the one hand, $\mathbf{R}$ represents a slower time-varying statistic compared to first-order (instantaneous) CSI, which provides a more realistic feedback scenario. And on the other hand, $\mathbf{R}$ is quantization friendly due to its Hermitian and positive definiteness properties [SMPI10]. Asymptotically, the minimization of the objective function (7.49) has the meaning of *maximum spectrum mismatch*, as

$$\xi(\mathbf{a}_k) = \mathbf{a}_k^H \mathbf{R} \mathbf{a}_k = \mathrm{tr}\left(\mathbf{R} \mathbf{a}_k \mathbf{a}_k^H\right) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \phi(\omega) |A_k(\omega)|^2 d\omega, \tag{7.50}$$

where $A_k(\omega)$ is the frequency response of the waveform $\mathbf{a}_k$, and $\phi(\omega)$ is the Fourier transform of the second-order statistics of the primary system. Therefore, minimizing (7.50) involves finding the secondary waveforms that exhibit higher mismatch in spectrum.

Because opportunistic communication investigates low primary system occupancy scenarios, the low rank structure of $\mathbf{R}$ is exploited. In particular, the dimension of the primary system is defined as

$$N = \mathrm{rank}(\mathbf{R}), \tag{7.51}$$

i.e., the first N eigenvectors of $\mathbf{R}$ span the *primary subspace*. It is precisely through these eigenvectors that the secondary waveforms will produce interference. As far as $N < M$, the aim of the secondary system is to direct the secondary waveforms into the *noise subspace*, i.e., the orthogonal subspace to the primary subspace, spanned by the $K = M - N$ remaining eigenvectors of $\mathbf{R}$. In the following, the primary occupancy is defined as

$$\kappa_0 \doteq \frac{N}{M} < 1. \tag{7.52}$$

It is worth noting that CSI on the secondary-to-primary communication channel is not sufficient to exploit the unused resources by the primary system, as the activity of the primary users is not reflected in this particular channel. To this end, it is assumed that the interference channel is the combined effect of the true secondary-to-primary communication channel plus a spectral mask that models the spectrum holes of the primary system, i.e., the Fourier transform of the interference channel response has zero spectral contribution in the frequency resources not used by the primary system.

7.4.2 Nonrobust Waveform Optimization with Perfect CSI

In the case of perfect CSI, the optimization problem (7.49) reads

$$\min_{\mathbf{a}_k} \mathbf{a}_k^H \mathbf{R} \mathbf{a}_k \quad \text{subject to} \quad \mathbf{a}_k^H \mathbf{e}_k = 1, \tag{7.53}$$

for $k = 1, \ldots, K$, and where the additional constraint $\mathbf{a}_k^H \mathbf{e}_k = 1$ has been added to avoid the trivial solution. The solution of the optimization problem (7.53) is derived in Appendix E.7 and is given by

$$\mathbf{a}_k = \frac{\mathbf{P}_{\mathbf{U}_k}^{\perp} \mathbf{e}_k}{\sqrt{\alpha_k}}, \tag{7.54}$$

where α_k is a scaling factor to guarantee that $\mathbf{a}_k$ is unit-norm, and

$$\mathbf{P}_{\mathbf{U}_k}^{\perp} \doteq \mathbf{I} - \mathbf{U}_k \mathbf{U}_k^H, \tag{7.55a}$$

$$\mathbf{U}_1 \doteq \mathbf{U}, \tag{7.55b}$$

$$\mathbf{R} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^H, \tag{7.55c}$$

$$\mathbf{U}_{k+1} \doteq \left[\mathbf{U}_k \quad \mathbf{a}_k\right], \tag{7.55d}$$

$$\mathbf{e}_k \doteq [\underbrace{0 \cdots 0}_{M-k} \; 1 \; \underbrace{0 \cdots 0}_{k-1}]^T. \tag{7.55e}$$

The solution of the waveform optimization problem is presented as an iterative algorithm, which has implementation advantages as the secondary system is able to sequentially obtain new waveforms until the full noise subspace is exploited, or until the quality of service (QoS) required (i.e., the size of the physical layer alphabet) is achieved.

From (7.54), it is seen that the secondary waveforms implement an orthogonal projector to the primary subspace, i.e., (7.55), which depends on the eigenvectors of $\mathbf{R}$ and is independent on the eigenvalues of $\mathbf{R}$. This means that to achieve complete orthogonality, the information contained in the eigenvalues, i.e., the strength of each *eigenchannel* is not necessary. As a result, when perfect second-order statistical CSI (P-CSI) on the interference channel is available, the average interference is completely nulled, i.e.,

$$\xi_{\text{P-CSI}}(\mathbf{A}) = 0. \tag{7.56}$$

Interestingly, the interference minimization problem (7.53) becomes equivalent to the following minimum norm problem with orthogonality constraint

$$\min_{\mathbf{a}_k} \mathbf{a}_k^H \mathbf{a}_k \quad \text{subject to} \quad \mathbf{a}_k^H \mathbf{e}_k = 1 \text{ and } \mathbf{U}^H \mathbf{a}_k = \mathbf{0}, \tag{7.57}$$

as proved in Appendix E.8. This means that (7.54) will exhibit all the good properties of the classical linear predictor [TKK82]. In particular, the polynomial zeroes of the secondary waveforms will have an asymptotic distribution which is uniform on a circle of radius less than unity, as advocated in [Pak87]. A direct consequence of this result it that the proposed waveforms satisfy the desired property of white frequency response in the noise subspace, a preferable condition in cognitive radio communication.

Furthermore, one key property of minimum norm waveforms is that it relies on the orthogonal projector (7.55), which exhibits invariance to rotations within the primary subspace. This property is suitable in cognitive radio networks, as several secondary users may obtain rotated versions of the second-order statistical CSI (e.g., by observing the primary signal or exploiting available pilots). Modeling the rotation by a rotation matrix $\mathbf{\Psi}_i$[4], each of the versions available at the secondary users can be modeled as

$$\mathbf{V}_i = \mathbf{U}\mathbf{\Psi}_i. \tag{7.58}$$

Distances are preserved in the primary subspace, therefore the proposed minimum norm waveforms are rotationally invariant as $\mathbf{P}_{\mathbf{V}_i} = \mathbf{P}_{\mathbf{U}}$ for any $\mathbf{\Psi}_i$. However, if the rotation is not within the primary signal subspace, i.e., of the type $\mathbf{\Psi}_i\mathbf{U}$, then mismatching between known subspaces may arise. This motivates the design of robust waveforms, as addressed in the following.

[4]Note that a rotation matrix satisfies $\mathbf{\Psi}\mathbf{\Psi}^H = \mathbf{I}$ and $\det(\mathbf{\Psi}) = \pm 1$.

7.4.3 Robust Waveform Optimization with Imperfect Channel State Information

In practical cognitive radio networks, the errors in obtaining the second-order statistical CSI of the interference channel could translate to an inaccurate knowledge on $\mathbf{R}$, as it has been studied in the former Section 7.3 in the MDL problem. If the *priorly* known second-order statistical CSI is denoted as $\hat{\mathbf{R}}$, the existence of the following mismatching in the interference channel

$$\mathbf{R} = \hat{\mathbf{R}} + \mathbf{\Delta}, \tag{7.59}$$

where $\mathbf{\Delta}$ is the semi-definite positive second-order mismatching. The existence of $\mathbf{\Delta}$ could lead to increasing interference to the primary system, as well as performance loss in the secondary system, as the knowledge on the primary subspace may differ from the actual primary subspace, as well as among secondary users. For this reason, the robust waveform optimization problem is addressed. In contrast to (7.47), this section optimizes the following waveforms

$$\tilde{\mathbf{A}} = (\tilde{\mathbf{a}}_1, \ldots, \tilde{\mathbf{a}}_K) \tag{7.60}$$

to minimize (7.48) taking into account the mismatching (7.59).

Particularly, the mismatching (7.59) is bounded by [Höl89]

$$\|\mathbf{\Delta}\|_\infty \leq \epsilon, \tag{7.61}$$

where the parameter $\epsilon > 0$ takes the meaning of uncertainty level, and $\|\mathbf{\Delta}\|_\infty$ evaluates the infinite norm of $\mathbf{\Delta}$, i.e., its maximum eigenvalue. Hence, addressing robust waveform optimization involves solving the following problem

$$\min_{\tilde{\mathbf{a}}_k} \tilde{\mathbf{a}}_k^H \left(\hat{\mathbf{R}} + \mathbf{\Delta}\right) \tilde{\mathbf{a}}_k \quad \text{for any} \quad \|\mathbf{\Delta}\|_\infty \leq \epsilon, \tag{7.62}$$

instead of that of (7.49). Contrarily to the P-CSI scenario, the waveforms cannot meet the orthogonality equality for any possible mismatching $\mathbf{\Delta}$ in the imperfect second-order statistical CSI (I-CSI) scenario. In this sense, the worst-case performance optimization, motivated by [SGLW03] in the beamforming problem, is addressed. More specifically, finding the worst-case subspace mismatching $\mathbf{\Delta}$ given an uncertainty level ϵ will produce a set of waveforms which will exhibit better interference performance for any other realization of $\mathbf{\Delta}$. The worst-case performance optimization is written as

$$\max_{\mathbf{\Delta}} \tilde{\mathbf{a}}_k^H \left(\hat{\mathbf{R}} + \mathbf{\Delta}\right) \tilde{\mathbf{a}}_k \;\; \text{subject to} \;\; \|\mathbf{\Delta}\|_\infty \leq \epsilon. \tag{7.63}$$

Solving (7.63) leads (see Appendix E.9 for the details) to

$$\mathbf{\Delta} = \epsilon \tilde{\mathbf{a}}_k \tilde{\mathbf{a}}_k^H \tag{7.64}$$

In other words, the worst-case subspace mismatching scenario with uncertainty ϵ is such mismatching that concentrates all the uncertainty in a subspace isotropically spanned by the direction to be minimized. This is indeed a very pessimistic scenario, as in general mismatching could expand or shrink the primary subspace up to a certain measure. After plugging (7.64) into the objective function (7.62), the objective function becomes

$$\xi(\tilde{\mathbf{A}}) = \tilde{\mathbf{a}}_k^H \left(\hat{\mathbf{R}} + \epsilon \mathbf{I}\right) \tilde{\mathbf{a}}_k \tag{7.65}$$

i.e., the worst-case performance scenario involves a diagonal loading into the Gramm matrix of the interference channel. Hence, together with the nontriviality constraint, the robust waveform optimization problem reads

$$\min_{\tilde{\mathbf{a}}_k} \tilde{\mathbf{a}}_k^H \left(\hat{\mathbf{R}} + \epsilon \mathbf{I}\right) \tilde{\mathbf{a}}_k \quad \text{subject to} \quad \tilde{\mathbf{a}}_k^H \mathbf{e}_k = 1, \tag{7.66}$$

for $k = 1, \ldots, K$.

The solution to the former problem is derived in Appendix E.10 and is given by the expression

$$\tilde{\mathbf{a}}_k = \frac{(\mathbf{I} - \hat{\mathbf{U}}_k \mathbf{\Phi}_k \hat{\mathbf{U}}_k^H)\mathbf{e}_k}{\sqrt{\tilde{\alpha}_k}} \tag{7.67}$$

where $\tilde{\alpha}_k$ is a factor that guarantees unit-norm waveforms, and where

$$\begin{aligned}
\mathbf{\Phi}_1 &= \begin{bmatrix} \frac{\hat{\lambda}_1}{\hat{\lambda}_1+\epsilon} & & 0 \\ & \ddots & \\ 0 & & \frac{\hat{\lambda}_N}{\hat{\lambda}_N+\epsilon} \end{bmatrix}, && \text{(7.68a)} \\
\hat{\mathbf{U}}_1 &= \hat{\mathbf{U}}, && \text{(7.68b)} \\
\hat{\mathbf{R}} &= \hat{\mathbf{U}} \hat{\mathbf{\Lambda}} \hat{\mathbf{U}}^H, && \text{(7.68c)} \\
\mathbf{\Phi}_{k+1} &= \begin{bmatrix} \mathbf{\Phi}_k & 0 \\ 0 & 1 \end{bmatrix}, && \text{(7.68d)} \\
\hat{\mathbf{U}}_{k+1} &= \begin{bmatrix} \hat{\mathbf{U}}_k \ \tilde{\mathbf{a}}_k \end{bmatrix} && \text{(7.68e)} \\
\mathbf{e}_k &\doteq [\underbrace{0 \cdots 0}_{M-k} \ 1 \ \underbrace{0 \cdots 0}_{k-1}]^T. && \text{(7.68f)}
\end{aligned}$$

In (7.68), $\hat{\lambda}_n$ denotes the n-th eigenvalue of $\hat{\mathbf{R}}$. From (7.67), it can be appreciated that the robust waveforms implement a *soft* orthogonal projector of the form $(\mathbf{I} - \hat{\mathbf{U}}_k \mathbf{\Phi}_k \hat{\mathbf{U}}_k^H)$, where the eigenvectors of the priorly known Gram matrix are combined by means of the diagonal elements of $\mathbf{\Phi}_k$, which take into account the tradeoff between the uncertainty level ϵ and the magnitude of each eigenchannel, $\hat{\lambda}_k$. In other words, it implements a measure of the

quality level of the CSI, as each of the elements of the diagonal of $\mathbf{\Phi}_k$ can be seen as

$$\frac{\hat{\lambda}_n}{\hat{\lambda}_n + \epsilon} = \frac{\text{SNR}_n}{1 + \text{SNR}_n}, \tag{7.69}$$

where the SNR_n has been defined as $\text{SNR}_n \doteq \hat{\lambda}_n / \epsilon$.

If the quality of the CSI is asymptotically high, it follows that $\epsilon \to 0$, $\text{SNR}_n \to \infty$, $\mathbf{\Phi}_k \to \mathbf{I}$, $(\mathbf{I} - \hat{\mathbf{U}}_k \mathbf{\Phi}_k \hat{\mathbf{U}}_k^H) \to \mathbf{P}_{\mathbf{U}_k}$ in (7.55), and therefore that

$$\tilde{\mathbf{a}}_k \to \mathbf{a}_k \tag{7.70}$$

in (7.54) for $k = 1, \ldots, K$, which is consistent with the nonrobust solution. On the other hand, if the quality of the CSI is asymptotically poor, it follows that $\epsilon \to \infty$, $\text{SNR}_n \to 0$, $\mathbf{\Phi}_1 \to \mathbf{0}$, and hence that

$$\tilde{\mathbf{a}}_k \to \mathbf{e}_k \tag{7.71}$$

for $k = 1, \ldots, K$. In general, the robust solution is partially using the information contained in $\mathbf{R}$ and the *pre-agreed* structure inherent in the nontriviality constraint $\tilde{\mathbf{a}}_k^H \mathbf{e}_k = 1$.

7.4.4 Efficient Frequency-Domain Implementation

In order to investigate the frequency components exploited by the secondary waveforms in opportunistic communication, the asymptotic behavior of (7.54) and (7.67) is addressed in the following for asymptotically large M. A direct consequence of the asymptotic behavior is that the implementation of the minimum norm waveforms can be efficiently done as a function of the frequency bins (library), instead of computing the eigenvalue decomposition (EVD) of an autocorrelation matrix.

Nonrobust Waveforms Asymptotic Frequency Response

As $M \to \infty$, the eigenvectors of the second-order statsitical CSI of the interference channel, i.e., the Gram matrix $\mathbf{R}$, approach to the frequency bins with non-zero spectral contribution [Gra06]. In this setting, the orthogonal projector (7.55) for $k = 1$ can be written as a function of the frequency bins

$$\mathbf{P}_{\mathbf{U}}^{\perp} = \sum_{n \in \mathcal{W}} \mathbf{f}_n \mathbf{f}_n^H, \tag{7.72}$$

where $\mathcal{W}$ indexes the frequency bins not occupied by the primary subspace, i.e., the noise subspace, and where

$$\mathbf{f}_n = \frac{1}{\sqrt{M}} \begin{bmatrix} 1 \\ \omega_n \\ \vdots \\ \omega_n^{M-1} \end{bmatrix}, \tag{7.73}$$

being ω_n the primitive n-th root of unity. After some mathematical manipulations, the first waveform can be expressed as

$$\mathbf{a}_1 \propto \sum_{n \in \mathcal{W}} \mathbf{f}_n, \tag{7.74}$$

i.e., it implements a uniform linear combination of the frequency bins located in the noise subspace. Hence, it is straightforward to show that the first waveform exhibits the maximally flat frequency response in the noise subspace. As stated above, this is a known property of minimum norm filtering with orthogonality constraints with respect to a certain positive measure on the unit circle.

Taking into account the iterative procedure for the subsequent waveforms $\mathbf{a}_2, \ldots, \mathbf{a}_K$, it follows that the asymptotic frequency response will strictly depend on the frequency bins located in the noise subspace, i.e.,

$$\mathbf{a}_k \propto \sum_{n \in \mathcal{W}} \beta_{kn} \mathbf{f}_n \tag{7.75}$$

where the set of parameters β_{kn} depend on the particular configuration of $\mathcal{W}$. In other words and given the orthogonality constraints with respect to the former waveforms, the rest of the waveforms will also exhibit the mostly flatter frequency response.

The same interpretations follow from the Z-transform of (7.74) and, in general, (7.75), given by

$$A_1(z) \propto \sum_{n \in \mathcal{W}} \frac{1 - z^{-M}}{1 - \omega_n z^{-1}}, \tag{7.76}$$

and

$$A_k(z) \propto \sum_{n \in \mathcal{W}} \beta_{kn} \frac{1 - z^{-M}}{1 - \omega_n z^{-1}}, \tag{7.77}$$

respectively. In any case, it is easy to observe that each of the frequencies occupied by the primary subspace are actually zeros of the secondary waveforms, i.e.,

$$A_k(z \to \omega_n) = 0, \tag{7.78}$$

for $k = 1, \ldots, K$ and $n \in \mathcal{P}$, being $\mathcal{P}$ the complementary set to $\mathcal{W}$, i.e., the set that indexes the frequency components of the primary subspace. In general, the weights in (7.77) outline that a secondary system based on minimum norm waveforms exhibit a multi-carrier code division multiple-access (CDMA) strategy.

A final note is that $|\mathcal{P}| = N$ and $|\mathcal{W}| = K$, meaning that κ_0 in (7.52) asymptotically acquires the meaning of spectrum occupancy, i.e., the primary system spectrum support to the total available bandwidth ratio.

Robust Waveforms Asymptotic Frequency Response

In order to study the asymptotic frequency response of the robust waveforms (7.67), the complete spectral decomposition of $\hat{\mathbf{R}}$ is required, i.e,

$$\hat{\mathbf{R}} = \sum_{n\in\mathcal{P}} \hat{g}_n \mathbf{f}_n \mathbf{f}_n^H, \tag{7.79}$$

where $\hat{g}_n$ is the squared value of the spectral contribution of the interference channel at the n-th frequency bin, i.e.,

$$\hat{g}_n \doteq |G(f_n)|^2, \tag{7.80}$$

where $G(f)$ is the frequency response of secondary-to-primary interference channel. Employing (7.79), the first robust waveform can be expressed as

$$\tilde{\mathbf{a}}_1 \propto \left(\mathbf{e}_1 - \frac{1}{\sqrt{M}} \sum_{n\in\mathcal{P}} \frac{\hat{g}_n}{\hat{g}_n + \epsilon} \mathbf{f}_n \right), \tag{7.81}$$

which implements a negative linear combination of he frequency bins located in the primary subspace, with an associated weights that depend on both $\hat{g}_n$ and ϵ. In particular, those frequency bins with high

$$\text{SNR}_n \doteq \frac{\hat{g}_n}{\epsilon} \tag{7.82}$$

will have a large contribution in the negative linear combination. Because of the iterative procedure for the subsequent waveforms $\tilde{\mathbf{a}}_2, \ldots, \tilde{\mathbf{a}}_K$, in general

$$\tilde{\mathbf{a}}_k \propto \left(\mathbf{e}_k - \frac{1}{\sqrt{M}} \sum_{n\in\mathcal{P}} \tilde{\beta}_{kn} \mathbf{f}_n \right). \tag{7.83}$$

It is straightforward to appreciate that complete orthogonality to the prior CSI $\hat{\mathbf{U}}$ is not achieved in the robust waveforms, as the orthogonal projector in the numerator of (7.67) is implementing a soft version taking into account the quality of the CSI, i.e., SNR_n.

Similar interpretations follow from the Z-transform of (7.81) and (7.83), given as

$$\tilde{A}_1(z) \propto \left(1 - \sum_{n\in\mathcal{P}} \frac{\hat{g}_n}{\hat{g}_n + \epsilon} \frac{1 - z^{-M}}{1 - \hat{\omega}_n z^{-1}} \right), \tag{7.84}$$

and

$$\tilde{A}_k(z) \propto \left(1 - \sum_{n\in\mathcal{P}} \tilde{\beta}_{kn} \frac{1 - z^{-M}}{1 - \hat{\omega}_n z^{-1}} \right), \tag{7.85}$$

respectively. Only when $\beta_{kn} = 1$, i.e., when the quality of the CSI in the n-th frequency bin is high, complete nulling $\tilde{W}_k(z \to \omega_n) = 0$ will occur. For other values of β_{kn}, the position of the zeroes in $\tilde{W}_k(z)$ is shifted to gain robustness in front of CSI uncertainty.

Residual Interference

As the robust formulation is motivated to reduce the residual interference that the nonrobust waveforms may produce due to imperfect CSI, this section evaluates the theoretical basis of this residual interference. Under the mismatching model (7.59), the waveforms (7.54) based on the CSI $\hat{\mathbf{R}}$ have the following normalized residual interference

$$\xi_{\text{I-CSI}}(\mathbf{a}_k) = \mathbf{a}_k^H \boldsymbol{\Delta} \mathbf{a}_k. \tag{7.86}$$

This is direct consequence of the orthogonality constraint between $\mathbf{a}_1, \ldots, \mathbf{a}_K$ and the eigenvectors of $\hat{\mathbf{R}}$. If $\boldsymbol{\Delta}$ is a random error matrix constrained by (7.61), it is clear to see that the residual interference associated to the nonrobust waveforms is bounded by the worst-case scenario, i.e., (7.64), which provides an upper bound on (7.86) as

$$\xi_{\text{I-CSI}}(\mathbf{a}_k) \leq \epsilon. \tag{7.87}$$

On the other hand, the residual interference associated to the robust waveforms is given by the general expression

$$\xi_{\text{I-CSI}}(\tilde{\mathbf{a}}_k) = \tilde{\mathbf{a}}_k^H \mathbf{R} \tilde{\mathbf{a}}_k. \tag{7.88}$$

Focussing on the first waveform, in Appendix E.11 it is shown that $\xi(\tilde{\mathbf{a}}_1)$ is upper bounded by

$$\xi_{\text{I-CSI}}(\tilde{\mathbf{a}}_1) \leq \epsilon \cdot \underbrace{\frac{1 - \dfrac{1}{M} \displaystyle\sum_{n \in \mathcal{P}} \frac{\text{SNR}_n}{\text{SNR}_n + 1}}{1 - \kappa_0 + \dfrac{1}{M} \displaystyle\sum_{n \in \mathcal{P}} \left(\frac{1}{1 + \text{SNR}_n} \right)^2}}_{\doteq 1/\Gamma_1(\text{SNR}_1, \ldots, \text{SNR}_N)}, \tag{7.89}$$

where $\Gamma_1(\text{SNR}_1, \ldots, \text{SNR}_N)$ has been defined as the interference gain. As expected, for $\text{SNR}_n \to \infty$, there is no gain between the robust and the nonrobust waveform, i.e.,

$$\Gamma_1(\infty) = 1. \tag{7.90}$$

7.5 Numerical Results

In this Section, the performance of the proposed secondary communication strategies, as well as the primary signal identification problem, is evaluated by means of numerical results.

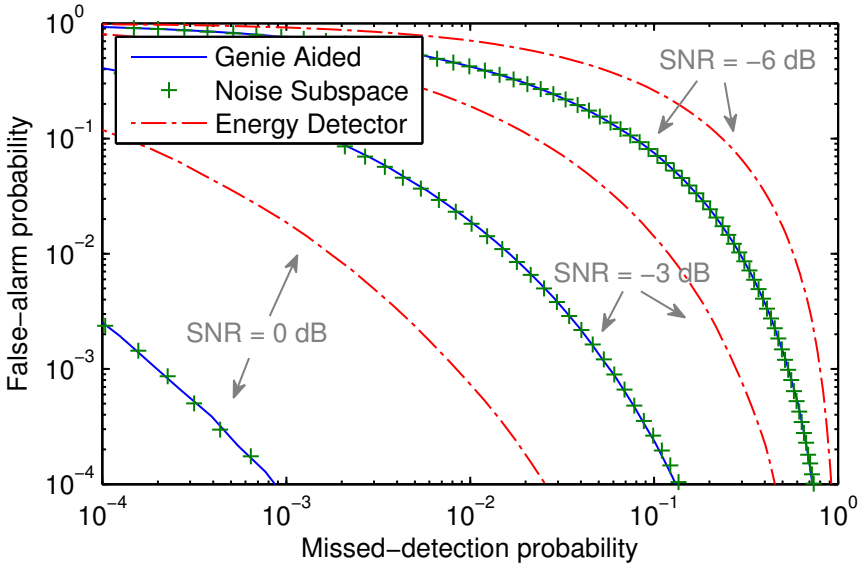

Figure 7.2: Complementary ROC of the genie aided, noise subspace and energy detector methods for several average SNR conditions with 256 samples and roll-off factor of 0.8.

7.5.1 Interference-Free Primary Signal Detection

In the following, numerical results are shown to assess the performance of the noise subspace detector. For sake of practical convenience, the operations involved in (7.14) have been implemented in the time-domain, as in view of the unitary property of the Fourier transform, the results presented in this work will hold for large enough signal block length. In the following simulations, $M = 256$, the noise and primary signal are white Gaussian distributed, the cognitive signal is a binary phase-shift keying (BPSK), with $G(f)$ a SRRC pulse with roll-off factor α, and the received signal is sampled at two samples per symbol. With the aim of benchmarking the following detectors are included in the simulations: a genie-aided detector which is able to subtract the $x(t)$ contribution from the observations and performs energy detection in the interval $\mathcal{W}$, as well as the energy detector which performs energy detection in the interval $\mathcal{W}$ directly on the observations.

Figure 7.2 depicts the receiver operating characteristics (ROC) of the genie aided, noise subspace and energy detector for several average SNR conditions and a roll-off factor of $\alpha = 0.8$. The performance of the proposed noise subspace detector reaches that of genie aided in all SNR conditions, whereas the energy detector incurs a severe performance loss due to the interference caused by the cognitive signal.

The same conclusions are obtained in Figure 7.3, which plots the ROC of the aforementioned detectors at the average SNR of -6 dB, for several values of α. It is worth noting that the ability of the noise subsapce detector to perform spectrum sensing free of interference is preserved for a wide range of roll-off factors.

In both figures, it is appreciated that the performance of the proposed detector in-

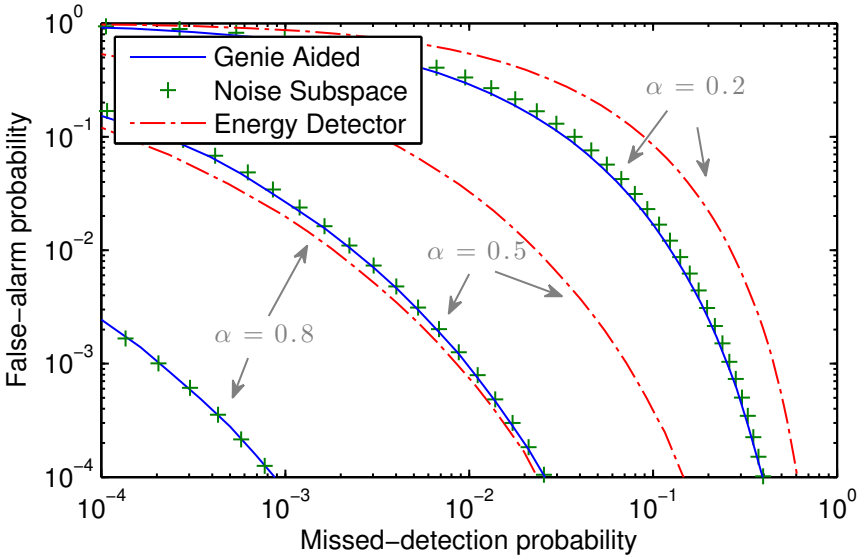

Figure 7.3: Complementary ROC of the genie aided, noise subspace and energy detector methods for several roll-off factors with 256 samples and SNR of 0 dB.

creases with α and the SNR, as it is forecast by the deflection (7.22) and corroborated by Figure 7.4, which illustrates the deflection of the genie aided, noise subspace and energy detector detectors versus the roll-off factor for several average SNR conditions. On the one hand, it can be noted how numeric simulations match the theoretical results as, e.g., for an SNR of 0 dB, $\alpha = 1$ and $|\mathcal{W}| = \alpha M = 256$, the deflection of the proposed detector equals $\mathbb{D}\{T_{\mathrm{NS}}\} = 256 \times 1^2 = 24$ dB. Furthermore, it is seen that a change in SNR conditions of 3 dB translates to a shift of 6 dB in deflection, as expected. On the other hand, in the energy detector case, it follows that the deflection is approximately 3.5 dB below that of the noise subspace and genie aided detectors, because the energy detector operates under SINR conditions, i.e.,

$$\mathbb{D}(L_{\mathrm{ED}}) \propto |\mathcal{W}| \cdot \mathrm{SINR}^2, \tag{7.91}$$

where the SINR is defined as

$$\mathrm{SINR} \doteq \frac{N_1}{N_0 + P/2}. \tag{7.92}$$

The degradation of the energy detector increases in high interference scenarios. In this simulation, $N_0 = P = 1$. Because, $\mathrm{SINR} = \mathrm{SNR} - 10\log_{10}(1 + P/2N_0)$ dB, the energy detector incurs a systematic loss with respect to the noise subspace and the genie aided of 1.75 dB, which translates into 3.5 dB in deflection, as prognosticated by Figure 7.4. Broadly speaking, the performance of the genie aided, noise subspace and energy detector accomplish the following condition:

$$\mathbb{D}(T_{\mathrm{ED}}) < \mathbb{D}(T_{\mathrm{NS}}) \approx \mathbb{D}(T_{\mathrm{GA}}). \tag{7.93}$$

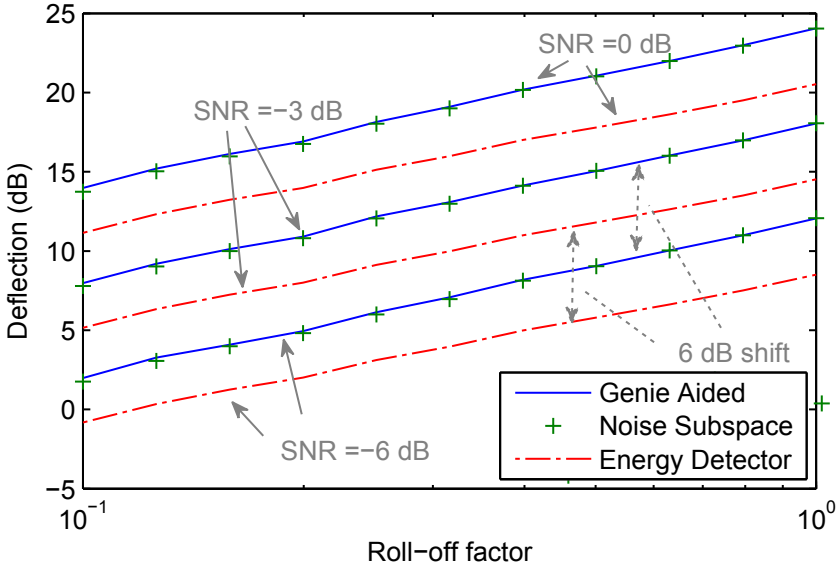

Figure 7.4: Evolution of the deflection of the gene aided, noise subspace and energy detector methods with the roll-off factor for several SNR conditions.

7.5.2 Minimum Description Length

The probability density function (PDF) of the proposed sphericity MDL is depicted in Figure 7.5 under the true hypotheses of $\mathcal{H}_0$ and $\mathcal{H}_3$, for $L = 1,000$ and SNR $= 10$ dB, along with the PDF of the MDL with known noise variance with perfect prior information on σ^2, and the PDF of the MDL with known noise variance with imperfect prior information on σ^2 with uncertainty level of 0.5 dB, in Figures 7.6 and 7.7, respectively. In all Figures it is appreciated that the theoretical statistical characterization of the MDLs is accurate. Under $\mathcal{H}_3$, the MDL for $n < 3$ are disjoint with those for $n > 3$ due to incompleteness of the GLR(n) for $n < 3$. On the other hand, since the sphericity is complete for $n \geq 3$, it is the penalty term in (7.26) that plays a role in the model order selection. It is worthy to point out that the noise uncertainty plays a fundamental role under $\mathcal{H}_0$, as seen in Figure 7.7.

The error probability is depicted in Figure 7.8 and Figure 7.9 versus SNR and L, respectively, under $\mathcal{H}_3$. It is appreciated that the theoretical statistical characterization forecasts an accurate behavior of the error probability as a function of the main parameters involved in the model order selection problem, namely, N, L and the SNR. In both figures, it is seen that for a fixed pair (L, SNR), the error probability degrades as the signal subspace dimension N increases and approaches to the full dimension. This is due to the fact that the sphericity ratios involved in the sphericity MDL (e.g., the ratio inside the logarithm in (7.34)) are closer to one. Compared to the MDL with known noise variance, in all scenarios the proposed sphericity MDL suffers from a degradation in either SNR or sensing time (required number of observations). However, the proposed sphericity MDL shows robustness in front of inaccurate knowledge on the noise variance, as the MDL with known noise

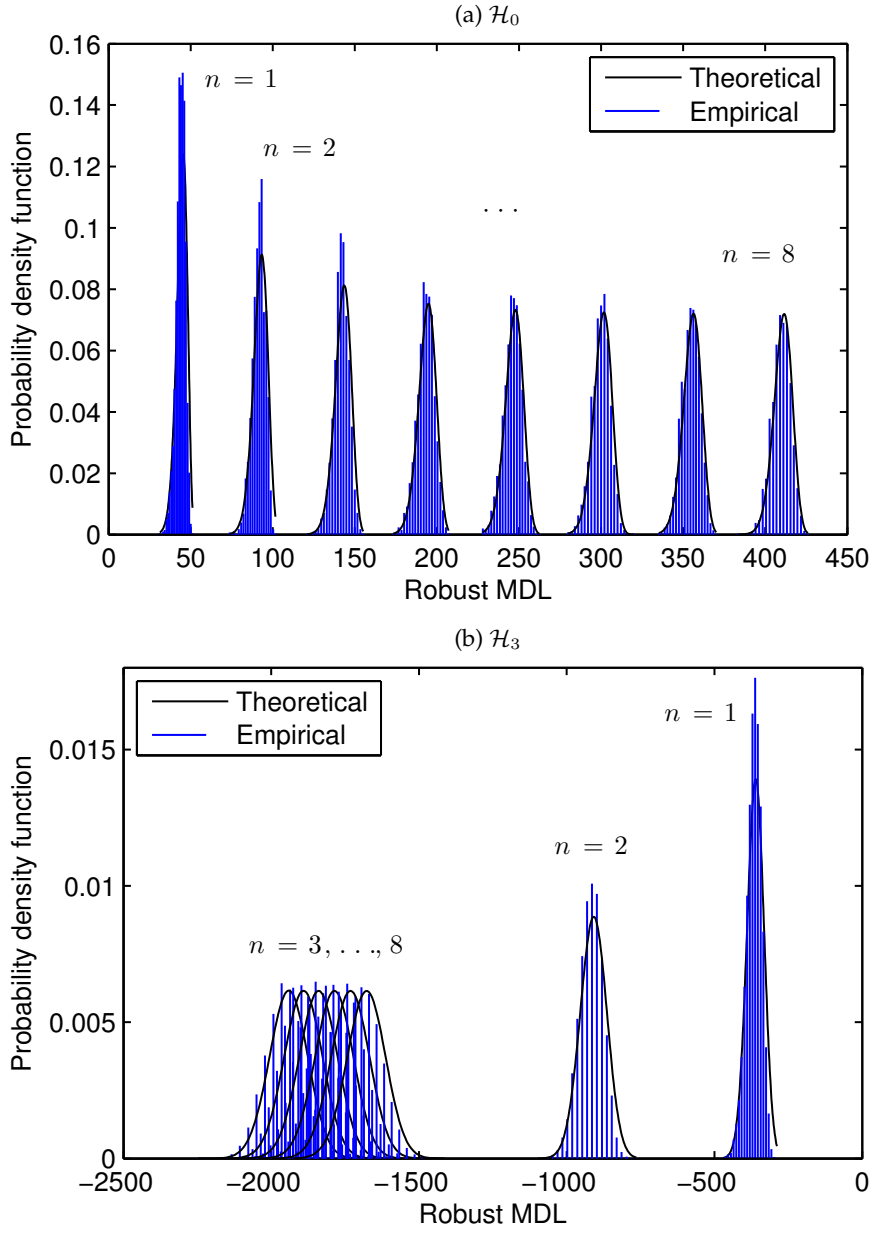

Figure 7.5: PDF of the robust MDL under the true hypotheses of (a) no primary signal, and (b) a primary signal with dimension 3.

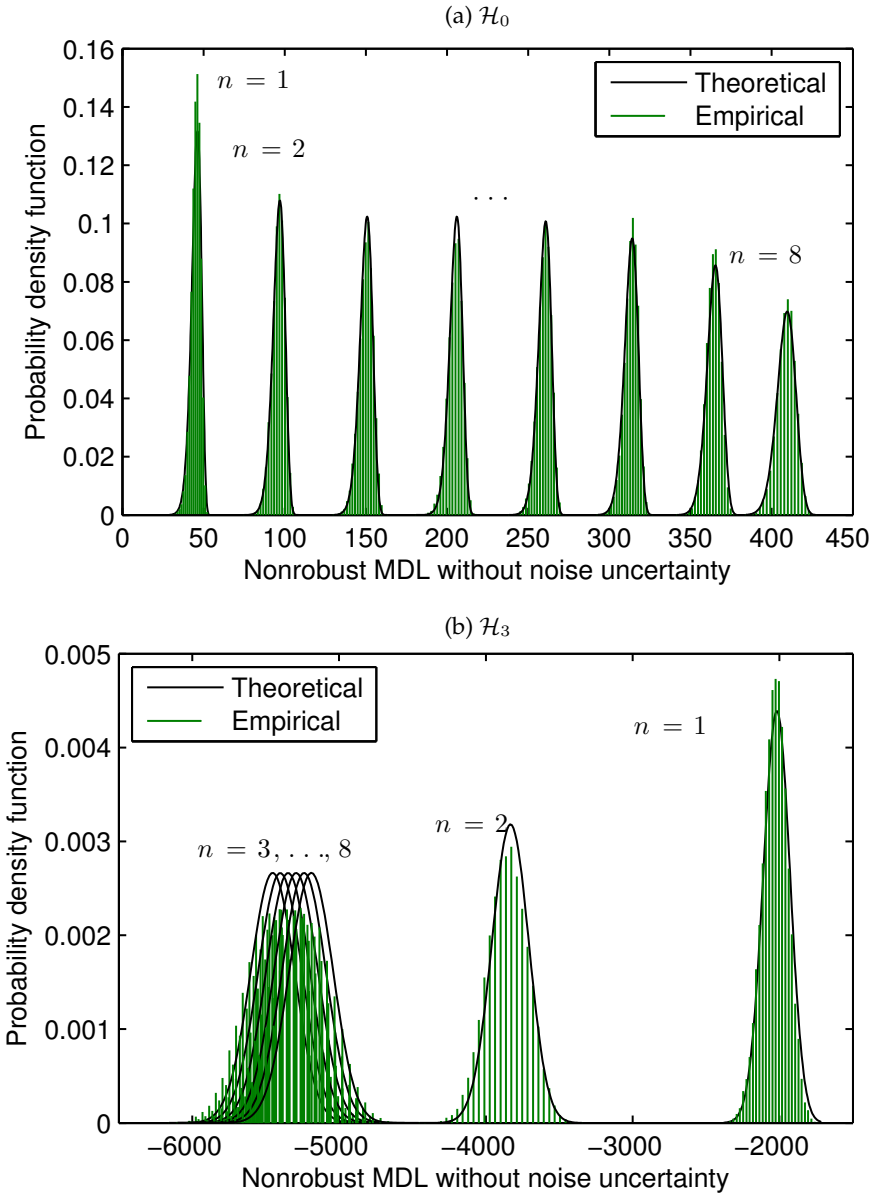

Figure 7.6: PDF of the nonrobust MDL without noise uncertainty under the true hypotheses of (a) no primary signal, and (b) a primary signal with dimension 3.

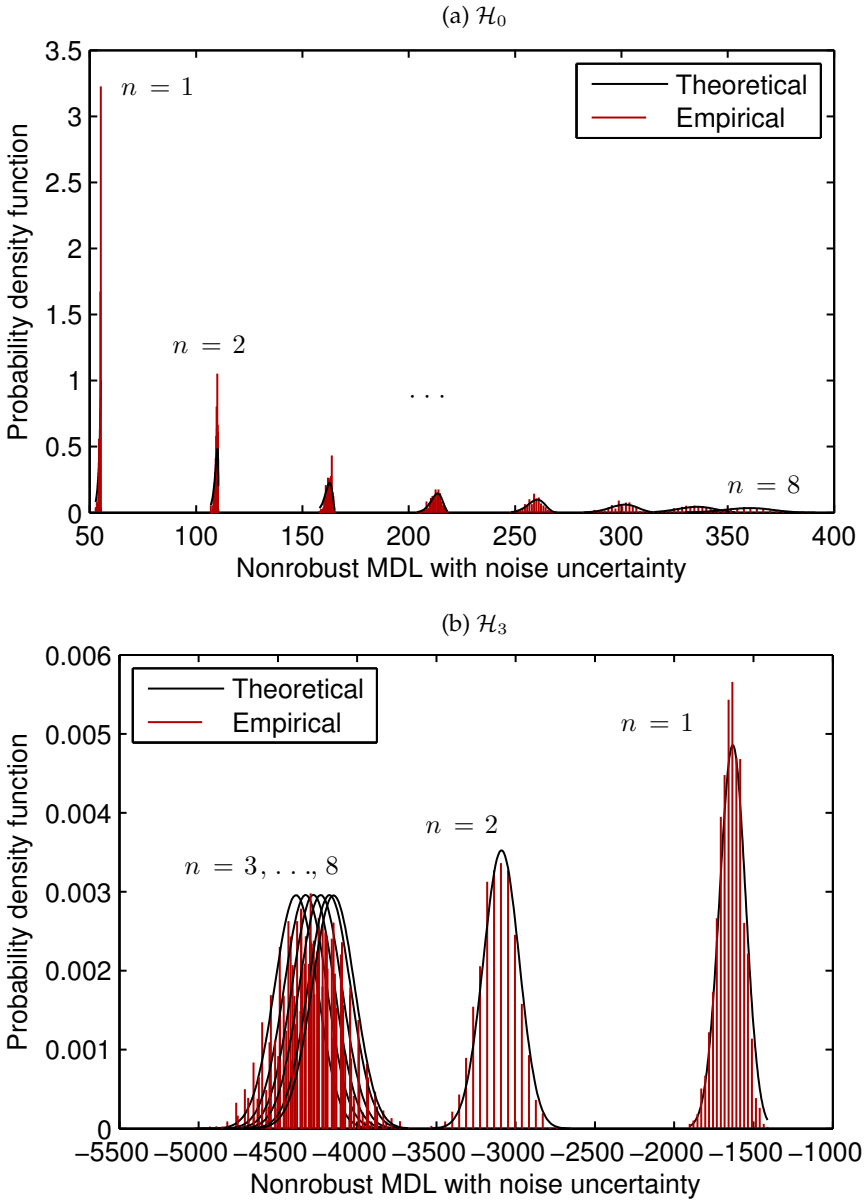

Figure 7.7: PDF of the nonrobust MDL with noise uncertainty of 0.5 dB under the true hypotheses of (a) no primary signal, and (b) a primary signal with dimension 3.

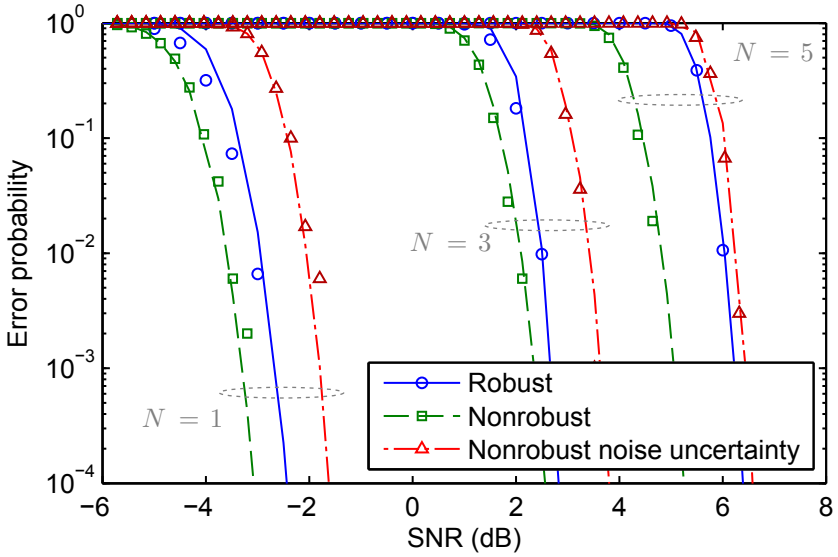

Figure 7.8: Theoretical (lines) and empirical (markers) error probability versus SNR of the robust, nonrobust, and nonrobust with noise uncertainty of 0.5 dB MDLs, with an observation size of 1,000 samples.

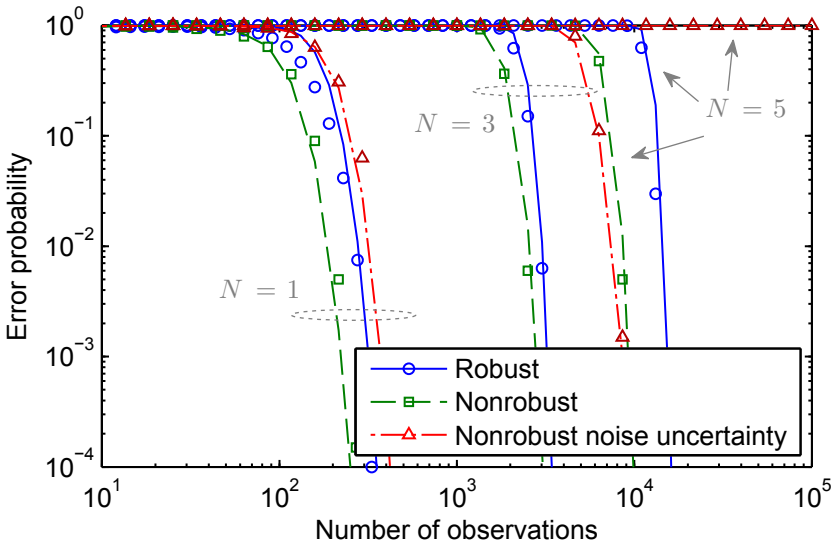

Figure 7.9: Theoretical (lines) and empirical (markers) error probability versus number of observations of the robust, nonrobust, and nonrobust with noise uncertainty of 0.5 dB MDLs, with an SNR of 0 dB.

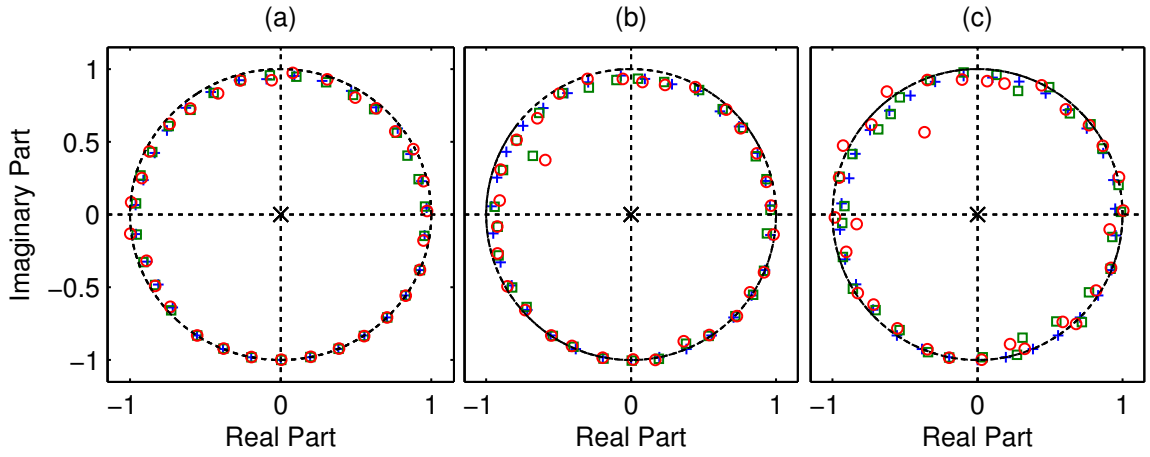

Figure 7.10: Zero-pole distribution of the genie aided (blue cross), nonrobust (red circle) and robust (green square) minimum norm waveforms of length 32 samples, for a primary system with occupancy of 3/8, with uncertainty levels of (a) 0, (b) 1/2, and (c) 1.

variance under noise uncertainty of 0.5 dB shows a severer degradation to the point that, e.g., at an SNR of 0 dB the error free model order selection for $N = 5$ is unfeasible.

7.5.3 Performance of Minimum Norm Waveform in Noncoherent Communication

The theoretical assessments derived in Section 7.4 are evaluated in numerical simulations. For sake of simplicity, in the sequel the primary system is modeled as a set of 8 DVB-T channels, and a flat secondary-to-primary channel is considered. The nonrobust (7.54) and robust (7.67) waveforms are considered for secondary transmissions, and a genie aided waveform consisting of (7.54) with perfect CSI is employed as benchmark. In the imperfect CSI scenarios, i.e., when $\epsilon > 0$, a random error matrix $\mathbf{\Delta}$ is generated such that the norm constraint (7.61) and the positive semi-definite structure of $\hat{\mathbf{R}}$ are satisfied.

As waveform optimization in cognitive radio networks involves spectral shaping, the frequency response of the genie aided, robust, and nonrobust waveforms is computed, when the primary system consists of 3 occupied channels out of the 8 available for opportunistic communication, i.e., a primary occupancy of $\kappa_0 = 3/8$.

Frequency-Response

The zero-pole distribution of the genie aided, nonrobust, and robust waveforms of length $M = 32$ samples is depicted in Figure 7.10, for three values of uncertainty $\epsilon = 1/10$, $\epsilon = 1/2$ and $\epsilon = 1$. In the three scenarios, it can be appreciated how the proposed waveforms exhibit the zero distribution characteristic of minimum norm filtering, i.e., the zeroes are distributed uniformly in a circle of radius less than unity. The orthogonality with the primary system met by the genie aided waveforms is not satisfied in imperfect CSI, as the

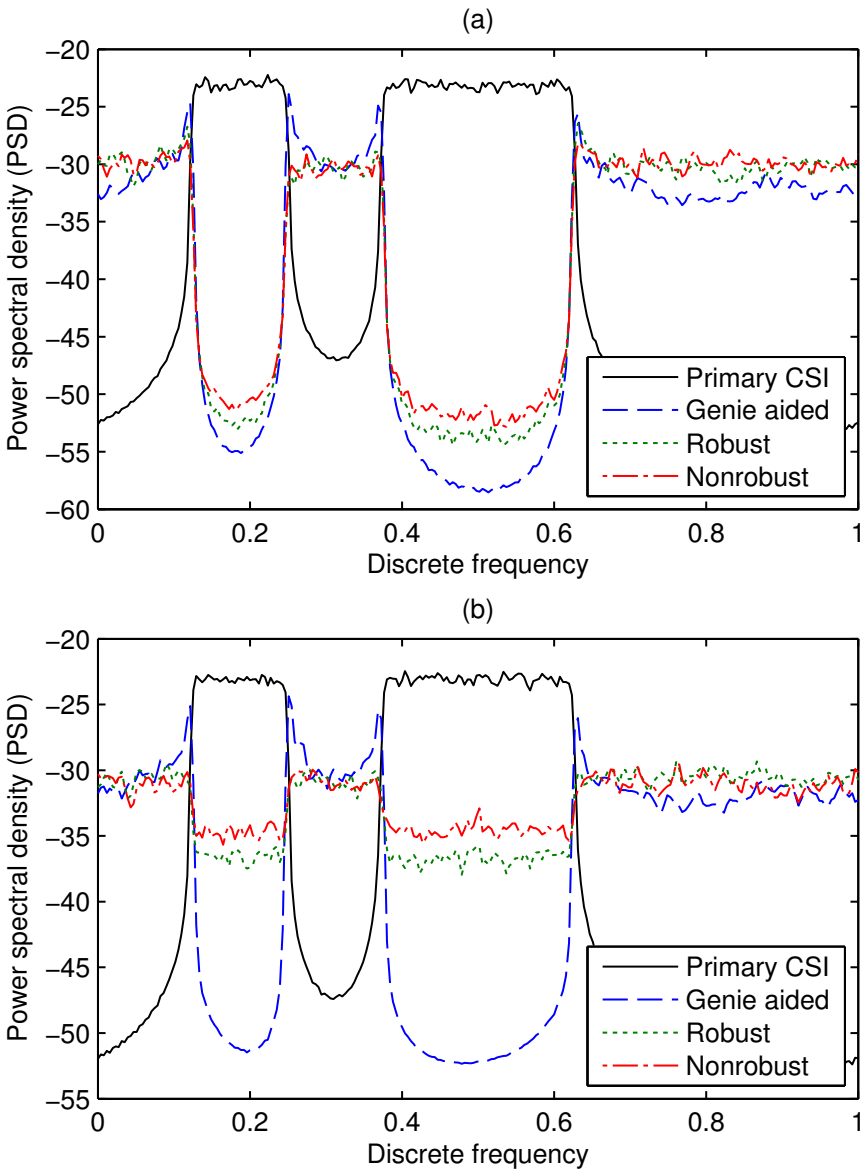

Figure 7.11: PSD of the genie aided, nonrobust and robust minimum norm waveforms of length 256 samples for a primary system with occupancy of 3/8, with uncertainty levels of (a) 1/2, and (b) 1.

zeros of the waveforms are not located in the primary subspace, i.e., the set of frequencies occupied by the primary system. This fact can be appreciated for $\epsilon = 1$ in the range, e.g., $-\pi/2 \leq \omega \leq 0$.

The comparison between the nonrobust and the robust counterparts is more clear in the frequency response. In particular, Figure 7.11 plots the PSD of the genie aided, robust and nonrobust waveforms for uncertainty levels of $\epsilon = 1/2$ and $\epsilon = 1$. In both cases, it can be appreciated that the proposed waveforms take advantage of the white spaces, and distribute the power over the noise subspace, i.e., the frequency components unused by the primary system, achieving maximally flat frequency response. Compared to the genie aided waveform, the proposed nonrobust waveform exhibits a noise floor which increases with the uncertainty level ϵ, as the orthogonality is not met. On the other hand, the robust waveform is able to diminish the floor effect by implementing a soft projection that takes into account the uncertainty in $\mathbf{R}$, i.e., it is more conservative on the information of the primary subspace.

Performance in Noncoherent Communication

This Section discusses the performance of the waveforms (7.47) and (7.60) over the secondary channel. More specifically, a point-to-point noncoherent communication channel is considered such that the channel remains invariant within the transmission of each waveform, i.e., during M symbols. The transmission of L waveforms is considered. The signal at the secondary receiver is given by $\mathbf{Z} = \mathbf{HY} + \mathbf{W}$, where the secondary channel matrix is

$$\mathbf{H} = \text{diag}(h_1, \ldots, h_L), \tag{7.94}$$

and the transmitted symbols $\mathbf{Y}$, the received symbols $\mathbf{Z}$ and the noise $\mathbf{W}$ are defined as

$$\mathbf{Y} = [\mathbf{y}_1, \ldots, \mathbf{y}_L], \tag{7.95a}$$

$$\mathbf{Z} = [\mathbf{z}_1, \ldots, \mathbf{z}_L], \tag{7.95b}$$

$$\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_L], \tag{7.95c}$$

respectively. For each transmission, the secondary transmitter selects with equal probability a waveform from $\mathbf{A}$, i.e, $\mathbf{y}_l = \sqrt{P}\mathbf{a}_k$, where P is the transmitted power. Noncoherent communication are adopted in hostile environments where the channel conditions change from transmission to transmission, as reflected by (7.94), so channel equalization is not practical. As the channel gain is unknown at each transmission, the noncoherent detection involves [BB99]

$$\hat{\mathbf{a}} = \max_{\mathbf{a} \in \mathbf{A}} |\mathbf{a}^H \mathbf{z}|^2. \tag{7.96}$$

In this setting, the noncoherent channel (7.94) is drawn from a zero-mean unit-norm normal distribution at each symbol transmission over a coherence of $M = 8$ samples.

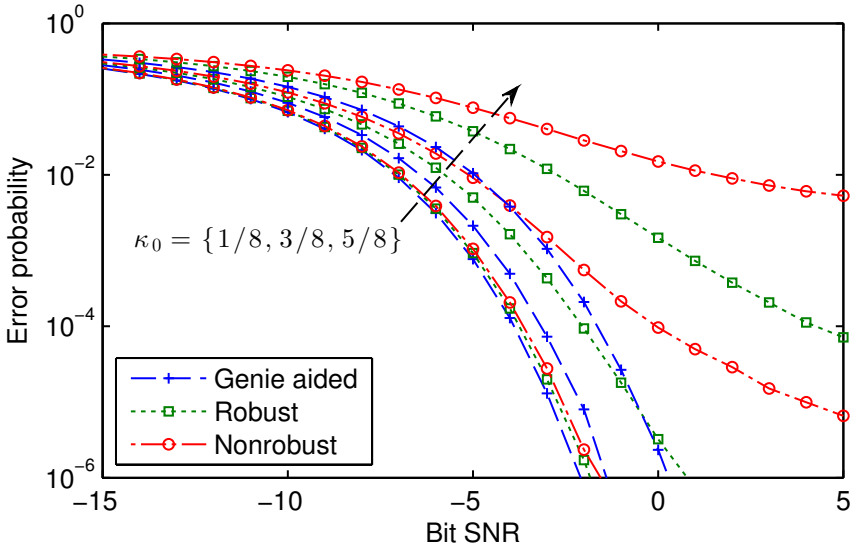

Figure 7.12: BER versus bit SNR per Hz of the genie aided, nonrobust and robust minimum norm waveforms of length 8 samples with uncertainty level of 1/2, with occupancies of 1/8, 3/8, and 5/8.

In Figure 7.12, the bit error rate (BER) versus the average bit SNR of the noncoherent channel with several primary occupancies is depicted, for an uncertainty level of $\epsilon = 1/2$. The curves show better error probability vs per bit SNR tradeoff in low occupancy scenario, as the cognitive radio users are able to exploit a larger noise subspace (larger K) for opportunistic communication. Furthermore, the degradation incurred by imperfect CSI is reflected as a error floor in the error probability, i.e., as SNR $\to \infty$ the error exponent of BER tends to zero. This degradation is further increased with κ_0, as the scenario is more restrictive. Compared to the nonrobust counterpart, the proposed robust waveforms achieve significant gain in imperfect CSI, as the error floor is displaced several orders of magnitude.

To observe the effect of ϵ over the error floor Figure 7.13 depicts the BER versus per bit SNR when the primary system has an occupancy of $\kappa_0 = 3/8$ with uncertainty levels of $\epsilon = 0$, $\epsilon = 1/2$, and $\epsilon = 1$. It is seen that for moderate uncertainty on the CSI, the proposed robust waveform exhibits low error exponent degradation compared to the genie aided waveform. However, in severely imperfect CSI, both nonrobust and robust waveform saturate at high SNR, as the mismatching between the CSI among secondary users is too severe.

The contour plot of the BER versus the uncertainty level of the secondary transmitter, ϵ_1, and the uncertainty level of the secondary receiver, ϵ_2, at SNR of -1.5 dB is plotted in Figure 7.15. It can be observed that for a wide range of uncertainty pair (ϵ_1, ϵ_2) the gain in error probability of the robust waveform with respect to the nonrobust waveform is of 1 order of magnitude. Furthermore, a symmetric behavior is appreciated, in the sense that an increment in either the uncertainty level of the transmitter or the uncertainty level of

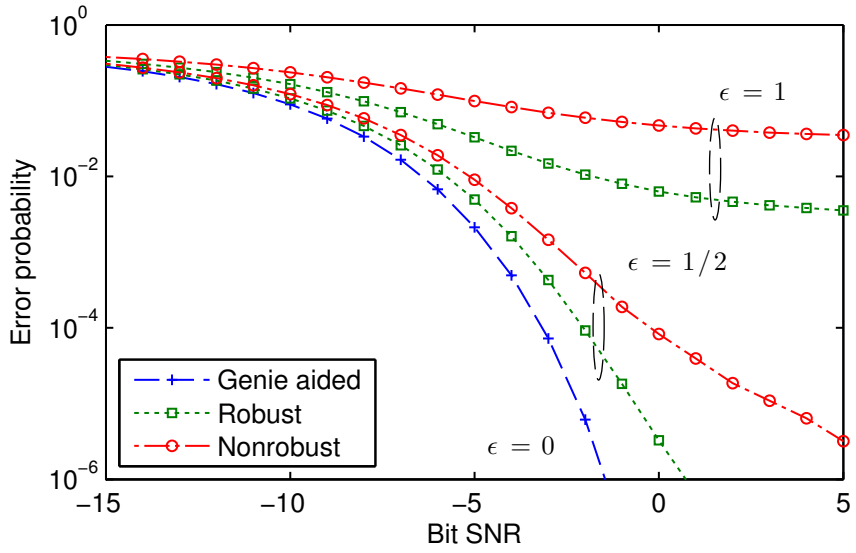

Figure 7.13: BER versus bit SNR per Hz of the genie aided, nonrobust and robust minimum norm waveforms of length 8 samples with occupancy of 3/8, with uncertainty levels of 0, 1/2, and 1.

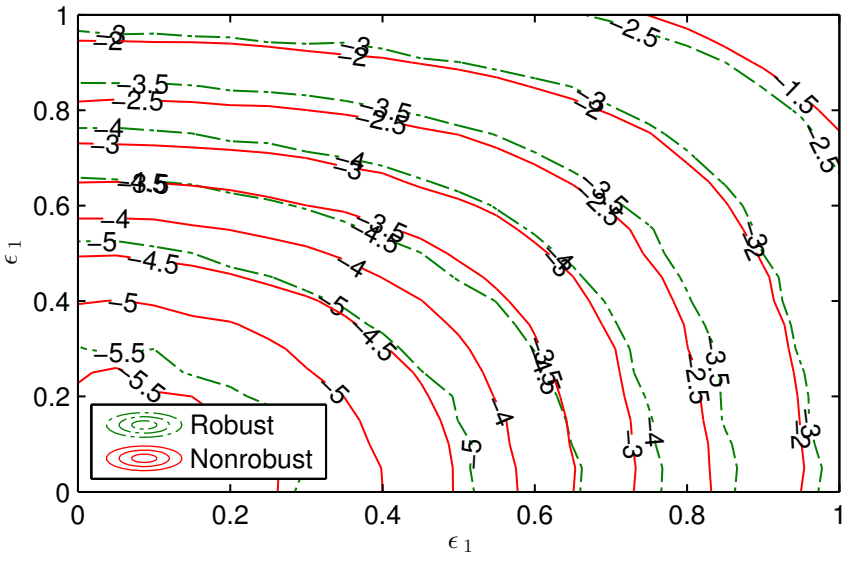

Figure 7.14: Contour plot of BER versus the uncertainty levels of user 1 and user 2 at bit SNR of -1.5 dB per Hz and primary occupancy of 3/8.

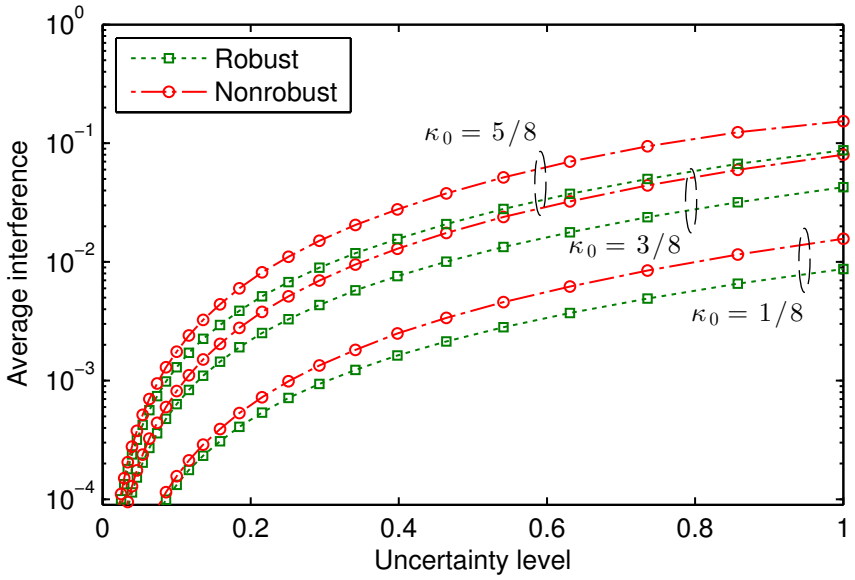

Figure 7.15: Average interference versus uncertainty level of the nonrobust and robust minimum norm waveforms of length 8 samples, with occupancies 1/8, 3/8, and 5/8.

the receiver acts as bootle neck regardless the CSI quality of the other.

Finally, the average interference generated by a secondary signal employing the nonrobust and nonrobust waveforms is examined in Figure 7.14 versus the uncertainty level ϵ for primary occupancies of $\kappa_0 = 1/8$, $\kappa_0 = 3/8$, and $\kappa_0 = 5/8$. The interference has been normalized so that it admits the meaning of portion of secondary transmitted power that actually interferes to the primary system. For fixed uncertainty level ϵ, the interference caused to the primary system is larger in more crowded scenarios. Additionally, the robust waveforms offer an improvement throughout a wide range of CSI uncertainty.

7.6 Conclusions

This Chapter has proposed a primary signal detector that exploits the noise subspace induced by the cyclostationarity exhibited by the cognitive signal. This detector allows performing spectrum sensing free of the interference caused by other cognitive users transmitting in the network. It has been shown that the deflection associated to the proposed detector is proportional to the roll-off factor of the cognitive signal and proportional to the squared value of the SNR, instead of the SINR. Numerical simulations corroborate the theoretical results, and show, for flat primary spectrum, detection performance is practically that of the genie aided detector.

Secondly, the sphericity MDL statistic with unknown signal correlation matrix and unknown noise variance has been derived for the Gaussian model order selection problem with application to cognitive radio. The asymptotic statistical characterization of the sphericity MDL has been derived, which allows to predict the performance of model order

selection as a function of the main parameters of the system, i.e,. the SNR, the observation size and the primary signal dimension. Simulation results have been provided to assess the performance of the proposed sphericity MDL in comparison to the MDL with known noise variance.

Finally, a minimum norm based approach for waveform optimization to achieve orthogonality to the primary signal subspace has been presented. This solution has the important properties of rotational invariance and maximally flat frequency response. Furthermore, in order to cope with the possible mismatching among the subspaces estimated by each cognitive radio user, a robust version of the minimum norm waveform has been proposed. Simulation results on both the frequency properties of the waveforms as well as its performance in secondary communication have been provided.

Chapter 8
Conclusions and Future Work

8.1 Conclusions

This thesis has considered three signal processing problems in the context of the cognitive radio technology: the sensing of the activity of the primary system, the detection of available resources, and the waveform optimization for opportunistic communication. The three signal processing problems have been posed under a common denominator: wideband regimes. As a consequence and as shown throughout the thesis, the proposed algorithms are focalized in low signal-to-noise ratio (SNR) scenarios, exploit the sufficient temporal second-order statistics of the primary signal and noise components, and take advantage of the potentials of nonuniform sampling. The objective of this thesis has been twofold: not only to derive practical, low complexity signal processing tools, but also to establish fundamental, theoretical limits in asymptotic regimes. Both perspectives have been contrasted with Monte Carlo numerical simulations.

As recalled by Figure 1.1 of Chapter 1, this thesis has been organized in six main parts, whose conclusions are summarized as follows.

The signal processing research challenges for spectrum sensing and dynamic access for wideband cognitive radio communication have been outlined in Chapter 2. The design of low complexity sensing and transmitting schemes that exhibit robustness to model inaccuracies has been identified as a major objective in the spreading of cognitive radio technology. State-of-the-art nonuniform sampling algorithms have been revised while noting that signal detection does not require actual data reconstruction. The recent advances in signal detection have also been outlined, with special highlights on the sensitivity of signal

detectors to signal, noise and channel model inaccuracies. Thirdly, several approaches to opportunistic communication have been reviewed, keeping in mind the degree of channel state information (CSI) and spectrum sensing errors as crucial factors in the design of transmitting waveforms.

The first problem of spectral analysis in nonuniform sampling has been addressed in Chapter 3. It has been shown that nonuniform sampling incurs the phenomenon of noise enhancement and power loss in the second-order statistics of the sampled signal. This phenomenon has been shown from experimentation in the design of the periodogram and Capon spectral estimates, and from the asymptotic theoretical analysis of Bernoulli nonuniform sampling. The proposed correlation-matching based spectral estimates, based on a rank-1 fitting from linearly projected data, circumvent the noise enhancement problem by implementing a denoising process in the spectral estimation. On the other hand, an equivalence between the noise enhancement and SNR has been introduced, which relates the degradation incurred by nonuniform sampling with the sampling density. Hence, this has allowed to predict the conditions in which many signal detection problems will operate under nonuniform sampling.

Chapter 4 has addressed the problem of signal detection in multi-frequency systems with priorly known normalized correlation matrices and by computing the estimates of the power profiles by maximum likelihood (ML) estimation and correlation-matching estimation for uniform and nonuniform sampling, respectively. On the one hand, under the low SNR assumption, a unified framework based on the frequency-domain asymptotic interpretation of the optimal generalized likelihood ratio test (GLRT) detectors has been derived, which consists of a kernel inherent to the detector and the periodogram of the observations. The detectors and associated kernels corresponding to a wide variety of situations involving several degrees of knowledge on the signal and noise statistics have been addressed, including the important case of multi-frequency systems. On the other hand, a correlation-matching approach for parameter estimation has allowed the derivation of the GLRT in nonuniform sampling for multi-frequency systems. The performance of the resulting detector remains tight to the estimator-correlator for a wide range of sampling densities and SNRs.

The asymptotic performance of signal detection has been addressed from two perspectives in Chapter 5. On the one hand, the Stein's lemma has been applied to the energy detector and the estimator-correlator in several scenarios, ranging from nonuniform sampling to noncoherent channels. The Stein's lemma allows to discover the main parameters that shape the error exponents in the error probabilities, i.e., the SNR, the number of observations, the primary signal occupancy, the sampling density, the diversity order. On the other hand, the asymptotic statistical characterization of the GLRT with unknown signal power in Bernoulli nonuniform sampling has been further derived. This has allowed the derivation of closed-form solution of sampling walls in noise uncertainty, specifically

as functions of the primary signal occupancy and the sampling density. Sampling walls state a sampling density limit bellow which the target detection probabilities cannot be guaranteed, regardless the number of observations.

In Chapter 6, the cyclostationarity properties of primary signals have been exploited to improve performance in the primary signal detection problem. More precisely, in the first part, by deriving a correlation-matching optimal invariant test statistic, the quadratic sphericity test (QST) has been obtained for the blind scenario. The QST evaluates the ratio between the squared mean and the arithmetic mean of the eigenvalues of the autocorrelation matrix. In the second part, the GLRT of a cyclostationary signal has been addressed by exploiting the rank-1 structure of small spectral covariance matrices in a frequency-domain representation of digital waveforms. By incorporating a parameterization of the frequency-selective channel, the proposed detector shows robustness in front of fading channels and outperforms the classical cyclostationarity-based detectors.

Finally, the last problem of waveform optimization has been reported in Chapter 7. In particular, the design of a secondary transmitting scheme has been addressed from two perspectives. First, a cyclostationary secondary waveform scheme has been proposed in order to mitigate the interference that active secondary users may cause to inactive secondary users who are performing spectrum sensing. By oversampling the received observations, the inactive cognitive radios are able to project the observations to a reduced noise subspace which is free from interference. It has been shown that the performance of the proposed detector is proportional to the roll-off factor of the cognitive transmitting signal, and proportional to the squared value of the SNR, instead of the signal-to-interference-plus-noise ratio (SINR). The second approach has been the design of maximally flat, invariant secondary waveforms that lie in the primary noise subspace. Beforehand, the sphericity minimum description length (MDL) statistic with unknown noise variance has been addressed to determine the dimension and directions of the primary signal subspace. The asymptotic characterization of the sphericity MDL has allowed to predict the performance of the model order selection as a function of the main parameters, i.e., the SNR, the observation size and the primary signal occupancy. When the primary noise subspace has been identified, a minimum norm based approach for waveform optimization has been presented. The advantage of minimum norm filtering is that it exhibits the properties of linear predictors: maximally flat frequency response and invariance. The proposed minimum norm waveform optimization has been further addressed in the important case of imperfect CSI in order to cope with the possible mismatching among the subspaces acquired by each cognitive user. The proposed waveforms exhibit low complexity implementation in the frequency-domain, and show robustness in front of imperfect CSI, as advocated by the numerical results on both the performance of the secondary channel and the interference level caused to the legacy system.

8.2 Future Work

After completing this thesis, the following future lines of research have been identified.

8.2.1 Sampling Pattern Optimization

This thesis has considered the Bernoulli nonuniform sampling, as the Gaussian-Bernoulli distribution which results from the formulation as product of sequences allows a neat analysis of the second-order statistics of the sampling. A possible future line of investigation would consist of generalizing the sampling patter, i.e., the sequence ψ_n, to adaptively optimize a performance metric such as the SNR or the Kullback-Leibler divergence (KLD) in signal detection, e.g.,

$$\psi_n = \arg\max \mathbb{D}(\mathcal{H}_1||\mathcal{H}_0). \tag{8.1}$$

This would allow to design nonuniform sampling patterns that exhibit matching and are adapted to the primary signal, noise and channel statistical parameters, while at the same time enhancing the detection performance, or reducing the actual sampling density for a given target performance.

8.2.2 Capacity Tradeoff with Imperfect Sensing

In the ideal scenario of perfect spectrum sensing, the cognitive radios are able to fully take advantage of the available resources to perform opportunistic communication, while keeping the legacy system free of interference. However, in practice, spectrum sensing techniques incur errors, namely false-alarms and missed-detections, with arbitrary small probabilities related to the signal detection conditions, i.e., the SNR, the observation size, the legacy system occupancy and the nonuniform sampling density. This nonzero error probabilities are translated into a degradation in both the secondary and primary channel capacities. Hence, the establishment of a tradeoff between the primary user capacity, the secondary user capacity, and the spectrum sensing performance could be explored. In a first attempt by resorting to the flat fading channel case, an scalar parameterization of the tradeoff would consist of determining an upper-bound on the product between capacities as a function of the channel and spectrum sensing parameters, i.e.,

$$C_{\mathrm{P}} \cdot C_{\mathrm{S}} \leq f(\mathrm{SNR}, N, P_{\mathrm{FA}}, P_{\mathrm{MD}}, \kappa, \kappa_0, ...). \tag{8.2}$$

The characterization of such a function would allow to establish fundamental limits and observe how the detection parameters affect the primary and secondary links actual quality.

8.2.3 Cutoff-Rate Waveform Optimization

Resource allocation is an important problem in underlay cognitive radios in order to take advantage of the underutilized resources by the licensed primary users. In the underlay paradigm, the secondary transmissions may cause little interference to the primary users, hence utilizing a larger amount of resources. With the use of side information provided by the spectrum sensing part, a possible line of investigation consists of studying the problem of power allocation in coded cognitive radio communications from a maximum cutoff-rate perspective, i.e.,

$$P_1, \ldots, P_L = \arg\max R_0(\text{SNR}, P_{\text{MD}}, P_{\text{FA}}, \kappa, \kappa_0, ...), \tag{8.3}$$

where R_0 stands for the cutoff-rate. The cutoff-rate was originally proposed as a sensitive criterion for the design of block and convolutional coded communications, as with low-density parity-check (LDPC) codes and iterative decoding the cutoff-rate was considered as an upper bound for the rates where feasible codes can be achieved. However, although current codes achieve rates higher than the cutoff-rate, this parameter still maintains the importance as an indicator of the error exponent behavior. The motivation of using the cutoff-rate instead of capacity rises from the fact that since the capacity only gives a potential maximum rate, probably the more convenient form of the coding theorem is the evaluation of the packet error rate (PER) as a function of the block length and the information rate [Sha59]. Particularly, Gallager showed that the average error probability of a family of randomly generated codes decreases exponentially with the block length according to an error exponent, as far as the code rate is below the channel capacity [Gal68]. The main motivation behind the minimization of the PER is that this parameter is a realistic indicator of the quality of the communication link in numerous communications systems. Specifically, it is directly related to the real throughput as the PER potentially minimizes the total amount of packets required in a transmission.

8.2.4 Interference Alignment

Along with fading, interference is one of the main concerns in digital communications. While fading has been the central phenomenon in point-to-point communications, the performance in current communication networks is often limited by the interference produced among users. Therefore, managing the interference is an important factor in cognitive radio communications. Basically, there are three conventional approaches in treating the interference in communication networks: decoding the interference, orthogonalization, and consider the interference as noise. The first approach consists in decoding the interference caused by the interferer, to further subtract the interferer contribution to the received signal and finally decode the useful signal. The second approach tries to avoid interference by having signal and interference to be orthogonal in some domain, e.g., time, frequency,

space, or code. Lastly, the third scheme treats the interference as noise, which is a useful approach when the interference is weak.

The three aforementioned interference management approaches share the property that the sum-rate capacity of the network is independent on the number of users, hence limited by interference. Surprisingly, the theory behind interference alignment states that the sum-rate capacity of a time-varying interference network with limited resources can be increased linearly with the number of users Interference alignment was independently introduced by several authors, e.g., [JS07, JS08, MAMK08, CJ08, CJ09, SCKP10]. Interference alignment permits a transmitter to partially or completely align its interference within the unused dimensions of the channel, given by the degrees of the interference channel. In other words, it refers to the construction of signals (codes) such that the resulting transmitted signal lies in a subspace which is orthogonal to the one spanned by the signal of interest at each receiver. In a frequency selective channel with K users, up to $K/2$ degrees of freedom are achievable when using interference alignment as a transmitting technique, and zero-forcing as the receiving technique.

In dense networks, white spaces might not be enough persistent to guarantee the quality of service (QoS) of the secondary users. In this scenario, secondary users are unable to transmit without harming the primary users by means of interference. To circumvent this problem, a possible line of investigation would consists of proposing interference alignment techniques to operate in the primary signal noise subspace and reveal how the degrees of freedom behave within the cognitive radio parameters, i.e.,

$$d \leq f(K, \kappa_0, \kappa, ...). \tag{8.4}$$

8.2.5 Noncoherent Communication and Grassmann Manifolds

In the last part of this thesis, the concepts of primary signal subspace and noise subspace have been used in several problems. Grassmann manifolds are mathematical tools obtained by identifying orthogonal matrices whose columns span the same subspace. In other words, a point in the Grassmann manifold is a representative of the signal subspace or the noise subspace, which represents the available subspaces for underlay communication. Since a point in the Grassmann manifold is a linear subspace, the choice of this matrix is not unique, and understanding the geometry of Grassmann manifolds allows the formulation of important mathematical instruments in optimization problems (such as detection and estimation) such as the geodesics distance, parallel translation, gradients and Hessians in Grassmann manifolds.

In many communication systems, the receiver or both the receiver and the transmitter make use of some a priori knowledge on the CSI to achieve reliable communication. However, accurate CSI can only be acquired through a sufficient amount of training or pilot data. Generally, data-aided schemes show suitable performance in slow time-varying

channels, and have become no longer efficient under fast time-varying channels of current wireless communication systems. In other words, with shorter coherence times, CSI acquisition means larger overheads and significant estimation errors. To improve the spectral efficiency in such scenarios, noncoherent communications including both noncoherent data detection and blind channel estimation have gained recent attention. Noncoherent communication is appealing as only a few pilot data or no pilot is used for data detection and channel estimation, and therefore have a big promise for deployment under fast time-varying channels.

As a result of the two aforementioned tools, a future line of investigation would consist of applying the Grassmann manifolds theory to noncoherent communication in the primary noise subspace, i.e., the design of secondary waveforms to maximize the distance between the signal and noise subspaces, e.g.,

$$\mathbf{y}_1, \ldots, \mathbf{y}_K = \arg\max d(\mathbf{U}_s, \mathbf{U}_w), \tag{8.5}$$

where $d(\mathbf{U}_s, \mathbf{U}_w)$ evaluates the (geodesic) distance between the signal and noise subspaces.

Appendix A

Proofs of Chapter 3

A.1 Proof of (3.10) and (3.11)

Noting the convexity of the problem (3.8), the derivative is taken with respect to the signal level and set it to zero, which leads to the equation

$$\frac{1}{M}\sum_{m=1}^{M}\mathrm{tr}\left[\left(\mathbf{x}_m\mathbf{x}_m^H-\mathbf{\Psi}_m\left(P\mathbf{R}_0+\sigma^2\mathbf{I}_N\right)\mathbf{\Psi}_m^H\right)\mathbf{\Psi}_m\mathbf{R}_0\mathbf{\Psi}_m^H\right]=0. \tag{A.1}$$

Similarly, after taking the derivative with respect to the noise power level and equal it to zero, it is obtained that

$$\frac{1}{M}\sum_{m=1}^{M}\mathrm{tr}\left[\left(\mathbf{x}_m\mathbf{x}_m^H-\mathbf{\Psi}_m\left(P\mathbf{R}_0+\sigma^2\mathbf{I}_N\right)\mathbf{\Psi}_m^H\right)\mathbf{I}_K\right]=0, \tag{A.2}$$

where the property $\mathbf{\Psi}_m\mathbf{\Psi}_m^H=\mathbf{I}_K$ has been used. Clearly, the derivatives outline the coupling between the signal and noise statistics, which is based on the projection of the local observations onto the normalized correlation matrices. Defining the matrices $\mathbf{B}$, $\overline{\mathbf{R}}_0$, and $\overline{\mathbf{R}_0^2}$ as

$$\mathbf{B}\doteq\frac{1}{M}\sum_{m=1}^{M}\mathbf{\Psi}_m^H\mathbf{x}_m\mathbf{x}_m^H\mathbf{\Psi}_m, \tag{A.3}$$

$$\overline{\mathbf{R}}_0\doteq\frac{1}{M}\sum_{m=1}^{M}\mathbf{\Psi}_m\mathbf{R}_0\mathbf{\Psi}_m^H, \tag{A.4}$$

and

$$\overline{\mathbf{R}_0^2} \doteq \frac{1}{M}\sum_{m=1}^{M}\left(\mathbf{\Psi}_m\mathbf{R}_0\mathbf{\Psi}_m^H\right)^2, \tag{A.5}$$

the derivatives can be written as $\mathrm{tr}(\mathbf{B}\mathbf{R}_0) - P\mathrm{tr}(\overline{\mathbf{R}_0^2}) - \sigma^2\mathrm{tr}(\overline{\mathbf{R}}_0) = 0$, and $\mathrm{tr}(\mathbf{B}) - P\mathrm{tr}(\overline{\mathbf{R}}_0) - \sigma^2 K = 0$, respectively. Finally, after some mathematical manipulations, it follows that the correlation-matching estimates $\hat{P}$ and $\hat{\sigma}^2$ are given, respectively, by (3.10) and (3.11), as wanted to show.

Appendix B

Proofs of Chapter 4

B.1 Low Signal-to-Noise Ratio Approximation

$$\left(\mathbf{R}_s + \sigma^2\mathbf{I}\right)^{-1} \approx \frac{1}{\sigma^2}\left(\mathbf{I} - \frac{1}{\sigma^2}\mathbf{R}_s\right). \tag{B.1}$$

For any stationary signal, the correlation matrix can be expressed as the product of the signal variance by a normalized correlation matrix, i.e., $\mathbf{R}_s \doteq P\mathbf{R}_0$, where $P \geq 0$ and $\mathrm{diag}(\mathbf{R}_0) = N\mathbf{1}$. By taking σ^2 as a common factor, $\left(\mathbf{R}_s + \sigma^2\mathbf{I}\right)^{-1}$ can be expressed as

$$\left(\mathbf{R}_s + \sigma^2\mathbf{I}\right)^{-1} = \frac{1}{\sigma^2}\left(\mathbf{I} + \mathrm{SNR}\mathbf{R}_0\right)^{-1} \doteq \mathbf{R}(\mathrm{SNR}), \tag{B.2}$$

where the parameter $\mathrm{SNR} \doteq \frac{P}{\sigma^2}$ reflects the signal-to-noise ratio (SNR) of the problem. For asymptotic $\mathrm{SNR} \to 0$, the following matrix Taylor expansion $\mathbf{R}(\mathrm{SNR}_0) = \mathbf{R}(\mathrm{SNR}_0) + \frac{\partial}{\partial \mathrm{SNR}}\mathbf{R}(\mathrm{SNR}_0)(\mathrm{SNR} - \mathrm{SNR}_0) + \mathcal{O}(\mathrm{SNR}^2)$, as $\mathrm{SNR}_0 \to 0$, is considered. The low SNR approximation is formally defined as the truncation of the Taylor series at the squared order of the SNR and higher. By making use of the matrix derivation properties, it is obtained that $\frac{\partial}{\partial \mathrm{SNR}}\mathbf{R}(\mathrm{SNR}_0) = -\frac{1}{\sigma^2}\left(\mathbf{I} + \mathrm{SNR}_0\right)^{-1}\mathbf{R}_0\left(\mathbf{I} + \mathrm{SNR}_0\right)^{-1} = -\frac{1}{\sigma^2}\mathbf{R}_0$ as $\mathrm{SNR}_0 \to 0$. Then, as $\mathbf{R}(\mathrm{SNR}_0) = \frac{1}{\sigma^2}$, it follows that $\mathbf{R}(\mathrm{SNR}_0) \approx \frac{1}{\sigma^2}\mathbf{I} - \frac{1}{\sigma^2}\mathbf{R}_0\mathrm{SNR}$, which after some simple manipulations concludes the proof.

B.2 Proof of Theorem 4.1

It is first proved that, for asymptotically SNR $\to 0$, the generalized likelihood ratio test (GLRT) (4.5) is of the form $T(\mathbf{X},\Theta) = \mathrm{tr}(\mathbf{K}\hat{\mathbf{R}})$. Consider the GLRT (4.5) for the spectrum sensing problem (4.3) under the Gaussian assumption, i.e.,

$$L(\mathbf{X}) = \frac{p(\mathbf{X}|\hat{\mathbf{R}}_s,\hat{\sigma}_1^2,\mathcal{H}_1)}{p(\mathbf{X}|\hat{\sigma}_0^2,\mathcal{H}_0)} = \frac{\det(\hat{\sigma}_0^2\mathbf{I})^M}{\det(\hat{\mathbf{R}}_s+\hat{\sigma}_1^2\mathbf{I})^M} \times \frac{\exp\mathrm{tr}(-\mathbf{X}^H(\hat{\mathbf{R}}_s+\hat{\sigma}_1^2\mathbf{I})^{-1}\mathbf{X})}{\exp\mathrm{tr}(-\mathbf{X}^H(\hat{\sigma}_0^2\mathbf{I})^{-1}\mathbf{X})}. \tag{B.3}$$

Taking the logarithm and defining $\hat{\mathbf{R}} \doteq \frac{1}{M}\mathbf{X}\mathbf{X}^H$, after grouping terms it is obtained

$$\frac{1}{M}\log L(\mathbf{X}) = \mathrm{tr}\left(\left(\frac{1}{\hat{\sigma}_0^2}\mathbf{I} - (\hat{\mathbf{R}}_s+\hat{\sigma}_1^2\mathbf{I})^{-1}\right)\hat{\mathbf{R}}\right) + \log\left(\frac{\hat{\sigma}_0^2}{\hat{\sigma}_1^2}\right) - \log\det\left(\mathbf{I}+\frac{1}{\hat{\sigma}_1^2}\mathbf{R}_s\right). \tag{B.4}$$

On the one hand, the approximations $\hat{\sigma}_0^2 \approx \hat{\sigma}_1^2$ and $\log\det\left(\mathbf{I}+\frac{1}{\hat{\sigma}_1^2}\hat{\mathbf{R}}_s\right) \approx \mathrm{SNR}\mathrm{tr}(\hat{\mathbf{R}}_s) \approx 0$ are considered asymptotically as SNR $\to 0$. On the other hand, by recalling the matrix inversion lemma $(\mathbf{A}+\mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{DA}^{-1}\mathbf{B}+\mathbf{C}^{-1})^{-1}\mathbf{DA}^{-1}$ and setting $\mathbf{A} = \hat{\sigma}_1^2\mathbf{I}$, $\mathbf{B} = \hat{\mathbf{R}}_s$ and $\mathbf{C} = \mathbf{D} = \mathbf{I}$ in $(\hat{\mathbf{R}}_s+\hat{\sigma}_1^2\mathbf{I})^{-1}$, it is obtained that $(\hat{\mathbf{R}}_s+\hat{\sigma}_1^2\mathbf{I})^{-1} = \frac{1}{\hat{\sigma}_1^2}\mathbf{I} - \frac{1}{\hat{\sigma}_1^4}\hat{\mathbf{R}}_s(\frac{1}{\hat{\sigma}_1^2}\hat{\mathbf{R}}_s+\mathbf{I})^{-1} = \frac{1}{\hat{\sigma}_1^2}\mathbf{I} - \frac{1}{\hat{\sigma}_1^2}\hat{\mathbf{R}}_s(\hat{\mathbf{R}}_s+\hat{\sigma}_1^2\mathbf{I})^{-1}$. Applying this last result to (B.4),

$$\frac{1}{M}\log L(\mathbf{X}) \approx \mathrm{tr}\left(\frac{1}{\hat{\sigma}_1^2}\hat{\mathbf{R}}_s(\hat{\sigma}_1^2\mathbf{I}+\hat{\mathbf{R}}_s)^{-1}\hat{\mathbf{R}}\right) \doteq \mathrm{tr}\left(\mathbf{K}\hat{\mathbf{R}}\right). \tag{B.5}$$

For the second part of the proof, it is recalled that for large data records (i.e., as $N \to \infty$) in (4.3), it is established in [Kay98a, Ch. 5, Sec. 5] though approximating the probability density function (PDF) of $\mathbf{X}$ that log-likelihood decision statistic (B.5) can be approximated as (4.8), where $\mathcal{K}(\omega,\Theta)$ is the asymptotic continuous-frequency transform of the second-order statistic $\mathbf{K}$ given by (4.9), and $P(\omega)$ is the continuous-frequency periodogram of $\mathbf{X}$.

B.3 Proof of (4.13)

The ML estimate of P under $\mathcal{H}_1$ for the spectrum sensing problem (4.3) is given by $\hat{P} = \arg\min_P \log\det(P\mathbf{R}_0+\sigma^2\mathbf{I}) + \mathrm{tr}\left((P\mathbf{R}_0+\sigma^2\mathbf{I})^{-1}\hat{\mathbf{R}}\right)$ subject to $P \geq 0$. By making use of the following properties $\frac{\partial}{\partial x}\mathbf{A}^{-1} = -\mathbf{A}^{-1}\frac{\partial}{\partial x}\mathbf{A}\mathbf{A}^{-1}$ and $\frac{\partial}{\partial x}\log\det(\mathbf{A}) = \mathrm{tr}\left(\mathbf{A}^{-1}\frac{\partial}{\partial x}\mathbf{A}\right)$, the derivative with respect to P is taken and set to zero. This provides that the maximum likelihood (ML) estimate of P is the solution to the equation

$$\mathrm{tr}\left((P\mathbf{R}_0+\sigma^2\mathbf{I})^{-1}\mathbf{R}_0\right) - \mathrm{tr}\left((P\mathbf{R}_0+\sigma^2\mathbf{I})^{-1}\mathbf{R}_0(P\mathbf{R}_0+\sigma^2\mathbf{I})^{-1}\hat{\mathbf{R}}\right) = 0, \tag{B.6}$$

which has no closed-form solution. Yet, by further making use of Approximation (B.1), it follows that $\mathrm{tr}\left((P\mathbf{R}_0+\sigma^2\mathbf{I})^{-1}\mathbf{R}_0\right) \approx \frac{N}{\sigma^2} - \frac{P}{\sigma^4}\mathrm{tr}(\mathbf{R}_0^2)$, and

$$\mathrm{tr}\left((P\mathbf{R}_0+\sigma^2\mathbf{I})^{-1}\mathbf{R}_0(P\mathbf{R}_0+\sigma^2\mathbf{I})^{-1}\hat{\mathbf{R}}\right) \approx \frac{1}{\sigma^4}\mathrm{tr}(\mathbf{R}_0\hat{\mathbf{R}}) - \frac{2P}{\sigma^6}\mathrm{tr}(\mathbf{R}_0^2\hat{\mathbf{R}}), \tag{B.7}$$

as the ratio $\frac{P}{\sigma^2} \to 0$. Applying the former approximations, it is obtained that the derivative is equivalent to the equation $P\left(\frac{2}{\sigma^6}\mathrm{tr}(\mathbf{R}_0^2\hat{\mathbf{R}}) - \frac{P}{\sigma^4}\mathrm{tr}(\mathbf{R}_0^2)\right) = \frac{1}{\sigma^4}\mathrm{tr}(\mathbf{R}_0\hat{\mathbf{R}}) - \frac{N}{\sigma^2}$. After noting that $\mathrm{tr}(2\hat{\mathbf{R}}-\sigma^2\mathbf{I}) \approx \sigma^2\mathbf{I}$ and multiplying both sides of the equation by σ^4, it is obtained that the ML estimate of P reads

$$\hat{P} = \frac{\mathrm{tr}(\mathbf{R}_0\hat{\mathbf{R}}) - \sigma^2 N}{\mathrm{tr}(\mathbf{R}_0^2)}. \tag{B.8}$$

After applying the non-negativity restriction on $\hat{P}$, $(\hat{P})^+$, it is obtained (4.14). The expression of the GLRT spectrum sensing detectors (4.5) with known noise variance is given by (B.5) with $\hat{\sigma}_1^2 = \sigma^2$. For the particular case of unknown signal level, the optimal GLRT spectrum sensing detector obtained by further substituting $\hat{\mathbf{R}}_s = \hat{P}\mathbf{R}_0$, which concludes the proof.

B.4 Proof of (4.16)

Let $\mathbf{R}_s[\boldsymbol{\beta}]$ be the decomposition of $\mathbf{R}_s$ given by (4.15). The ML estimate of the coefficients $\boldsymbol{\beta} \doteq (\hat{\beta}_0,\ldots,\hat{\beta}_{N-1})$ onto $\mathcal{T}$ is given by

$$\hat{\boldsymbol{\beta}} = \arg\min_{\boldsymbol{\beta}} \log\det\left(\mathbf{R}_s[\boldsymbol{\beta}]+\sigma^2\mathbf{I}\right) + \mathrm{tr}\left(\left(\mathbf{R}_s[\boldsymbol{\beta}]+\sigma^2\mathbf{I}\right)^{-1}\hat{\mathbf{R}}\right), \tag{B.9}$$

with the additional constraint $\beta_0 \in \mathbb{R}_+$. By making use of the derivative properties, the derivative of the objective with respect to $\beta_{n'}^*$ is taken, which leads to the equation $\mathrm{tr}\left(\left(\mathbf{R}_s[\boldsymbol{\beta}]+\sigma^2\mathbf{I}\right)^{-1}\mathbf{T}_{n'}\right) - \mathrm{tr}\left(\left(\mathbf{R}_s[\boldsymbol{\beta}]+\sigma^2\mathbf{I}\right)^{-1}\mathbf{T}_{n'}\left(\mathbf{R}_s[\boldsymbol{\beta}]+\sigma^2\mathbf{I}\right)^{-1}\hat{\mathbf{R}}\right) = 0$, for $0 \le n \le N-1$. Solving for $\boldsymbol{\beta}$ requires the use of iterative algorithms which do not provide interpretation insight on the solution. However, in the wideband regime, the use of Approximation (B.1) allows to to approximate the following terms $\mathrm{tr}\left(\left(\mathbf{R}_s[\boldsymbol{\beta}]+\sigma^2\mathbf{I}\right)^{-1}\mathbf{T}_{n'}\right) \approx \frac{1}{\sigma^2}\mathrm{tr}(\mathbf{T}_{n'}) - \frac{1}{\sigma^4}\mathrm{tr}\left(\mathbf{R}_s[\boldsymbol{\beta}]\mathbf{T}_{n'}\right)$, and

$$\mathrm{tr}\left(\left(\mathbf{R}_s[\boldsymbol{\beta}]+\sigma^2\mathbf{I}\right)^{-1}\mathbf{T}_{n'}\left(\mathbf{R}_s[\boldsymbol{\beta}]+\sigma^2\mathbf{I}\right)^{-1}\hat{\mathbf{R}}\right) \approx \frac{1}{\sigma^4}\mathrm{tr}(\mathbf{T}_{n'}\hat{\mathbf{R}}) - \frac{2}{\sigma^6}\mathrm{tr}\left(\mathbf{R}_s[\boldsymbol{\beta}]\mathbf{T}_{n'}\hat{\mathbf{R}}\right), \tag{B.10}$$

as $\frac{\beta_n}{\sigma^2} \to 0$. Applying the former approximations, it is obtained that the ML estimate of $\mathbf{R}_s$ accomplishes $\frac{2}{\sigma^6}\mathrm{tr}\left(\mathbf{R}_s[\boldsymbol{\beta}]\mathbf{T}_{n'}\hat{\mathbf{R}}\right) - \frac{1}{\sigma^4}\mathrm{tr}\left(\mathbf{R}_s[\boldsymbol{\beta}]\mathbf{T}_{n'}\right) = \frac{1}{\sigma^4}\mathrm{tr}(\mathbf{T}_{n'}\hat{\mathbf{R}}) - \frac{1}{\sigma^2}\mathrm{tr}(\mathbf{T}_{n'})$. Noting that $\mathrm{tr}(2\hat{\mathbf{R}}-\sigma^2\mathbf{I}) \approx \sigma^2\mathbf{I}$, it reduces, after multiplying both sides by σ^2, to $\mathrm{tr}(\mathbf{R}_s[\boldsymbol{\beta}]\mathbf{T}_{n'}) = \mathrm{tr}(\mathbf{T}_{n'}\hat{\mathbf{R}}) - \sigma^2\mathrm{tr}(\mathbf{T}_{n'})$. Now, making use of the orthogonality of $\mathcal{T}$, it is noted that the term

$\mathrm{tr}(\mathbf{R}_s[\boldsymbol{\beta}]\mathbf{T}_{n'})$, using the decomposition (4.15), equals to $\beta_n \mathrm{tr}(\mathbf{T}_n^T\mathbf{T}_n)$, i.e., only the term $n = n'$ survives. Finally, $\beta_n \mathrm{tr}(\mathbf{T}_n^T\mathbf{T}_n) = \mathrm{tr}(\mathbf{T}_n\hat{\mathbf{R}}) - \sigma^2\mathrm{tr}(\mathbf{T}_n)$, which for $n = 0$ as $\mathbf{T}_0 = \mathbf{I}$ gives β_0 in (4.17a), and for $n \geq 1$ as $\mathrm{tr}(\mathbf{T}_n) = 0$ gives β_n in (4.17b). Finally, taking (B.5) with $\hat{\sigma}_1^2 = \sigma^2$ and $\hat{\mathbf{R}}_s = \mathbf{R}_s[\hat{\boldsymbol{\beta}}]$, proves (4.16).

B.5 Proof of (4.23)

The ML estimates of P and σ^2 are given by the convex optimization problem

$$\hat{P}, \hat{\sigma}_1^2 = \arg\min_{P,\sigma^2} \log\det(P\mathbf{R}_0 + \sigma^2\mathbf{I}) + \mathrm{tr}\left((P\mathbf{R}_0 + \sigma^2\mathbf{I})^{-1}\hat{\mathbf{R}}\right). \tag{B.11}$$

The solution for P is given by (B.8) for $\sigma^2 = \hat{\sigma}_1^2$. On the other hand, similarly the derivative with respect to σ^2 are taken and set it zero, which gives

$$\mathrm{tr}\left((P\mathbf{R}_0 + \sigma^2\mathbf{I})^{-1}\right) - \mathrm{tr}\left((P\mathbf{R}_0 + \sigma^2\mathbf{I})^{-2}\hat{\mathbf{R}}\right) = 0. \tag{B.12}$$

To solve for the noise variance, the Approximation (B.1) leads to $\mathrm{tr}\left((P\mathbf{R}_0 + \sigma^2\mathbf{I})^{-1}\right) \approx \frac{N}{\sigma^2} - \frac{P}{\sigma^4}N$, and $\mathrm{tr}\left((P\mathbf{R}_0 + \sigma^2\mathbf{I})^{-2}\hat{\mathbf{R}}\right) \approx \frac{1}{\sigma^4}\mathrm{tr}(\hat{\mathbf{R}}) - \frac{2P}{\sigma^6}\mathrm{tr}(\mathbf{R}_0\hat{\mathbf{R}})$. As a result, the derivative becomes equivalent to $P\left(\frac{2}{\sigma^6}\mathrm{tr}(\mathbf{R}_0\hat{\mathbf{R}}) - \frac{N}{\sigma^4}\right) = \frac{1}{\sigma^4}\mathrm{tr}(\hat{\mathbf{R}}) - \frac{N}{\sigma^2}$. After applying the low SNR approximation $\mathrm{tr}(2\hat{\mathbf{R}} - \sigma^2\mathbf{I}) \approx \sigma^2\mathbf{I}$ and rearranging (B.8), it is obtained a system of equations formed by $P\mathrm{tr}(\mathbf{R}_0^2) + \sigma^2 N = \mathrm{tr}(\mathbf{R}_0\hat{\mathbf{R}})$, and $PN + \sigma^2 N = \mathrm{tr}(\hat{\mathbf{R}})$. After some mathematical manipulations, it is obtained that the solutions are given by (4.24). Finally, taking (B.5) and placing $\hat{\mathbf{R}}_s = \hat{P}\mathbf{R}_0$, it proves (4.23).

B.6 Proof of (4.26)

The ML estimate of the coefficients $\boldsymbol{\beta} \doteq (\beta_0, \ldots, \beta_{N-1})$ and the noise variance are given by the optimization problem

$$\hat{\boldsymbol{\beta}}, \hat{\sigma}_1^2 = \arg\min_{\boldsymbol{\beta},\sigma^2} \log\det\left(\mathbf{R}_s[\boldsymbol{\beta}] + \sigma^2\mathbf{I}\right) + \mathrm{tr}\left(\left(\mathbf{R}_s[\boldsymbol{\beta}] + \sigma^2\mathbf{I}\right)^{-1}\hat{\mathbf{R}}\right), \tag{B.13}$$

with the additional constraint $\beta_0 \in \mathbb{R}_+$. As the problem is convex on $\boldsymbol{\beta}$ and σ^2, the derivative of the objective with respect to β_n^* and σ^2 is taken. On the one hand, the derivative with respect to β_n^* has been derived in Appendix B.4 and, after some mathematical manipulations and making use of Approximation (B.1) and the orthogonality of $\mathcal{T}$, reduces to $\beta_n \mathrm{tr}(\mathbf{T}_n^T\mathbf{T}_n) + \sigma^2\mathrm{tr}(\mathbf{T}_n) = \mathrm{tr}(\mathbf{T}_n\hat{\mathbf{R}})$, for $0 \leq n \leq N-1$. Hence, for $n \geq 1$, (4.27b) is proved because $\mathrm{tr}(\mathbf{T}_n) = 0$. For $n = 0$, $\beta_0 + \sigma^2 = \frac{1}{N}\mathrm{tr}(\hat{\mathbf{R}})$, together with the derivative with respect to σ^2, i.e., $\mathrm{tr}\left(\left(\mathbf{R}_s[\boldsymbol{\beta}] + \sigma^2\mathbf{I}\right)^{-1}\right) - \mathrm{tr}\left(\left(\mathbf{R}_s[\boldsymbol{\beta}] + \sigma^2\mathbf{I}\right)^{-2}\hat{\mathbf{R}}\right) = 0$. After applying Approximation (B.1), it is obtained the following approximations $\mathrm{tr}\left(\left(\mathbf{R}_s[\boldsymbol{\beta}] + \sigma^2\mathbf{I}\right)^{-1}\right) \approx$

$\frac{1}{\sigma^2}N - \frac{1}{\sigma^4}\mathrm{tr}\left(\mathbf{R}_s[\boldsymbol{\beta}]\right)$, and

$$\mathrm{tr}\left(\left(\mathbf{R}_s[\boldsymbol{\beta}]+\sigma^2\mathbf{I}\right)^{-1}\mathbf{T}_{n'}\left(\mathbf{R}_s[\boldsymbol{\beta}]+\sigma^2\mathbf{I}\right)^{-1}\hat{\mathbf{R}}\right) \approx \frac{1}{\sigma^4}\mathrm{tr}(\hat{\mathbf{R}}) - \frac{2}{\sigma^6}\mathrm{tr}\left(\mathbf{R}_s[\boldsymbol{\beta}]\hat{\mathbf{R}}\right). \tag{B.14}$$

Finally, noting that at the low SNR regime $\mathrm{tr}(2\hat{\mathbf{R}} - \sigma^2\mathbf{I}) \approx \sigma^2\mathbf{I}$, it is obtained $\mathrm{tr}(\mathbf{R}_s[\boldsymbol{\beta}]) + \sigma^2 N = \mathrm{tr}(\hat{\mathbf{R}})$. As $\mathrm{tr}(\mathbf{R}_s[\boldsymbol{\beta}]) = \beta_0 N$, it is obtained, again, $\beta_0 + \sigma^2 = \frac{1}{N}\mathrm{tr}(\hat{\mathbf{R}})$. Hence, employing the matrix decomposition $\mathcal{T}$, it is noted that both β_0 and σ^2 account for the white component of the received observations, and the ML estimates cannot be found further than $\hat{\beta}_0 + \hat{\sigma}_1^2 = \frac{1}{N}\mathrm{tr}(\hat{\mathbf{R}})$, proving (4.27b). Finally, (B.5) is taken with $\mathbf{R}_s = \mathbf{R}_s[\hat{\boldsymbol{\beta}}]$ to compute the GLRT. As the white component of the signal, β_0, cannot be recovered in a white noise environment, it is considered as noise, and systematically add it to $\hat{\sigma}_1^2$. Then, by defining $\boldsymbol{\beta}_1 \doteq (0, \beta_1, \ldots, \beta_{N-1})$, it proves (4.26).

B.7 Proof of (4.31)

The ML estimate of the power levels $\mathbf{P} \doteq (P_1, \ldots, P_Q)$ for the multi-frequency model (4.29) reads

$$\hat{\mathbf{P}} = \arg\min_{\mathbf{P}} \log\det\left(\sum_{q=1}^{Q} P_q\mathbf{R}_q + \sigma^2\mathbf{I}\right) + \mathrm{tr}\left(\left(\sum_{q=1}^{Q} P_q\mathbf{R}_q + \sigma^2\mathbf{I}\right)^{-1}\hat{\mathbf{R}}\right), \tag{B.15}$$

with the additional constraint $\mathbf{P} \succeq \mathbf{0}$. The problem is convex on $\boldsymbol{P}$ and the derivative of the objective with respect to P_l is taken and equaled to zero, which gives the following equation

$$\mathrm{tr}\left(\left(\mathbf{R}_s[\mathbf{P}]+\sigma^2\mathbf{I}\right)^{-1}\mathbf{R}_q\right) - \mathrm{tr}\left(\left(\mathbf{R}_s[\mathbf{P}]+\sigma^2\mathbf{I}\right)^{-1}\mathbf{R}_l\left(\mathbf{R}_s[\mathbf{P}]+\sigma^2\mathbf{I}\right)^{-1}\hat{\mathbf{R}}\right) = 0. \tag{B.16}$$

With analogous developments as Appendix B.4 with $\mathbf{R}_s[\boldsymbol{\beta}]$, the Approximation (B.1) is used to approximate the former equation by $\mathrm{tr}(\mathbf{R}_s[\mathbf{P}]\mathbf{R}_l) = \mathrm{tr}(\mathbf{R}_l\hat{\mathbf{R}}) - \sigma^2\mathrm{tr}(\mathbf{R}_l)$. By placing the Q equations in a matrix form, this proves (4.32). The expression of the multi-frequency nuisance GLRT (4.30) with known noise variance is given by (B.5) with the following modification. By letting $\sum_{l\neq q} P_l\mathbf{R}_l$ be a nuisance parameter, it behaves as known part of the noise when sensing at the q-th channel. By defining $\boldsymbol{\Xi}_q \doteq \sum_{l\neq q}\hat{P}_l\mathbf{R}_l + \sigma^2\mathbf{I}$, it proves (4.31).

B.8 Proof of (4.34)

Consider the multi-frequency model (4.29). The ML estimates of the power levels $\boldsymbol{P}$ and the noise variance under $\mathcal{H}_1$, $\hat{\sigma}_1^2$, are given by the optimization problem

$$\hat{\mathbf{P}}, \hat{\sigma}_1^2 = \arg\min_{\boldsymbol{P}} \log\det\left(\sum_{q=1}^{Q} P_q \mathbf{R}_q + \sigma^2 \mathbf{I}\right) + \mathrm{tr}\left(\left(\sum_{q=1}^{Q} P_q \mathbf{R}_q + \sigma^2 \mathbf{I}\right)^{-1} \hat{\mathbf{R}}\right). \tag{B.17}$$

Under the low SNR assumption, the derivative with respect to the power levels is given, similarly to Appendix B.7, by the set of equations $\mathrm{tr}(\mathbf{R}_s[\mathbf{P}]\mathbf{R}_l) + \sigma^2 \mathrm{tr}(\mathbf{R}_l) = \mathrm{tr}(\mathbf{R}_l \hat{\mathbf{R}})$. To the K former equations, the derivative with respect to the noise variance is added, which after analogous manipulations, leads to $\mathrm{tr}(\mathbf{R}_s[\mathbf{P}]) + \sigma^2 N = \mathrm{tr}(\hat{\mathbf{R}})$. The $Q+1$ equations in vectorial notation is then given by (4.35). The expression of the multi-frequency nuisance GLRT (4.30) with unknown noise variance is given by (B.5) with following straightforward modifications. By letting $\sum_{l \neq q} P_l \mathbf{R}_l$ be a nuisance parameter, it systematically behaves as known part of the noise when sensing at the q-th channel. By defining $\boldsymbol{\Xi}_q \doteq \sum_{l \neq q} \hat{P}_l \mathbf{R}_l + \hat{\sigma}_1^2 \mathbf{I}$, this concludes the proof.

B.9 Proof of (4.41)

Consider the correlation model $\mathcal{R}(P, \sigma^2) \doteq P\mathbf{R}_0 + \sigma^2 \mathbf{I}_N$. The signal and noise power levels that minimize the nonuniform correlation-matching (4.40) with nonuniformly sampled observations $\mathbf{X}$ are the solution of the optimization problem

$$(\hat{P}, \hat{\sigma}^2) = \arg\min_{P, \sigma^2} \mathcal{M}\left[\mathbf{X}, \mathcal{R}(P, \sigma^2)\right]. \tag{B.18}$$

Noting the convexity of the problem, the derivative with respect to the signal level is taken and set to zero, which leads to the equation

$$\mathrm{tr}\left[\left(\mathbf{X}\mathbf{X}^H - \sum_m \boldsymbol{\Psi}_m \left(P\mathbf{R}_0 + \sigma^2 \mathbf{I}_N\right) \boldsymbol{\Psi}_m^H\right) \sum_m \boldsymbol{\Psi}_m \mathbf{R}_0 \boldsymbol{\Psi}_m^H\right] = 0. \tag{B.19}$$

Similarly, after taking the derivative with respect to the noise power level and equal it to zero, it is obtained that

$$\mathrm{tr}\left[\left(\mathbf{X}\mathbf{X}^H - \sum_m \boldsymbol{\Psi}_m \left(P\mathbf{R}_0 + \sigma^2 \mathbf{I}_N\right) \boldsymbol{\Psi}_m^H\right) \sum_m \boldsymbol{\Psi}_m \mathbf{I}_N \boldsymbol{\Psi}_m^H\right] = 0. \tag{B.20}$$

Clearly, the derivatives outline the coupling between the signal and noise statistics, which is based on the projection of the local observations onto the normalized correlation matrices. By dividing each term in both equations by M, the following matrices $\hat{\mathbf{R}} \doteq \frac{1}{M}\mathbf{X}\mathbf{X}^H$

and $\overline{\mathbf{R}}_0 \doteq \frac{1}{M}\sum_m \mathbf{\Psi}_m \mathbf{R}_0 \mathbf{\Psi}_m^H$ are defined. The derivatives can be then written as $\mathrm{tr}(\hat{\mathbf{R}}\overline{\mathbf{R}}_0) - P\mathrm{tr}(\overline{\mathbf{R}}_0^2) - \sigma^2\mathrm{tr}(\overline{\mathbf{R}}_0) = 0$, and $\mathrm{tr}(\hat{\mathbf{R}}) - P\mathrm{tr}(\overline{\mathbf{R}}_0) - \sigma^2 K = 0$, respectively, where the property $\mathbf{\Psi}_m \mathbf{\Psi}_m^H = \mathbf{I}_K$ was employed. After some mathematical manipulations, it can be shown that the system of equations formed by the former derivatives has as solution (4.41).

B.10 Proof of (4.48)

Consider the multi-frequency correlation model given by $\mathcal{R}(P_1, \ldots, P_Q, \sigma^2) \doteq \sum_{q=1}^{Q} P_q \mathbf{R}_q + \sigma^2 \mathbf{I}_N$. The multi-frequency signal and noise power levels that minimize the nonuniform correlation-matching (4.40) with nonuniformly sampled observations $\mathbf{X}$ are the solution to the optimization problem

$$(\hat{P}_1, \ldots, \hat{P}_Q, \hat{\sigma}^2) = \arg \min_{P_1,\ldots,P_Q,\sigma^2} \mathcal{M}\left[\mathbf{X}, \mathcal{R}(P_1, \ldots, P_Q, \sigma^2)\right]. \tag{B.21}$$

It is noticed that the problem is convex on the multi-frequency signal power levels. By taking the derivative of $\mathcal{M}\left[\mathbf{X}, \mathcal{R}(P_1, \ldots, P_Q, \sigma^2)\right]$ with respect to the signal level at the l-th band and set it to zero, it leads to the equation

$$\mathrm{tr}\left[\left(\mathbf{X}\mathbf{X}^H - \sum_m \mathbf{\Psi}_m \left(\sum_q P\mathbf{R}_q + \sigma^2 \mathbf{I}_N\right) \mathbf{\Psi}_m^H\right) \sum_m \mathbf{\Psi}_m \mathbf{R}_l \mathbf{\Psi}_m^H\right] = 0. \tag{B.22}$$

Similarly, by the derivative with respect to the noise power level and equal it to zero, it leads to the equation

$$\mathrm{tr}\left[\left(\mathbf{X}\mathbf{X}^H - \sum_m \mathbf{\Psi}_m \left(\sum_q P\mathbf{R}_q + \sigma^2 \mathbf{I}_N\right) \mathbf{\Psi}_m^H\right) \sum_m \mathbf{\Psi}_m \mathbf{I}_N \mathbf{\Psi}_m^H\right] = 0. \tag{B.23}$$

It is noted that now the derivatives also outline the coupling among the statistics of the remaining bands and the noise. By dividing each term of both equations by M and further defining the matrices $\hat{\mathbf{R}} \doteq \frac{1}{M}\mathbf{X}\mathbf{X}^H$ and $\overline{\mathbf{R}}_l \doteq \frac{1}{M}\sum_m \mathbf{\Psi}_m \mathbf{R}_l \mathbf{\Psi}_m^H$ for $1 \leq l \leq Q$, the derivatives can be written as $\mathrm{tr}(\hat{\mathbf{R}}\overline{\mathbf{R}}_l) - \sum_q P_q \mathrm{tr}(\overline{\mathbf{R}}_q \overline{\mathbf{R}}_l) - \sigma^2 \mathrm{tr}(\overline{\mathbf{R}}_l) = 0$, and $\mathrm{tr}(\hat{\mathbf{R}}) - \sum_{q=1}^{Q} P_q \mathrm{tr}(\overline{\mathbf{R}}_q) - \sigma^2 K = 0$, respectively, where the property $\mathbf{\Psi}_m \mathbf{\Psi}_m^H = \mathbf{I}_K$ has been used. After some mathematical manipulations, the $Q + 1$ equations are placed in a system of equations, which proves (4.48).

Appendix C
Proofs of Chapter 5

C.1 Proof of (5.44)

Let $\mathcal{W} \doteq \{n : \text{SNR}_n > 0\}$ be the support of the SNR. Then, clearly, $\sum_{n\in\mathcal{W}} \text{SNR}_n = N{\cdot}\text{SNR}_a$. On the other hand, from (5.39) it is seen that the error exponent for the false-alarm probability depends on the squared value of the signal-to-noise ratio (SNR) matrix as $\text{tr}(\mathbf{SNR}^2)$. Hence,

$$\text{tr}(\mathbf{SNR}^2) = \sum_{n=1}^{N} \text{SNR}_n^2 = \sum_{n\in\mathcal{W}} \text{SNR}_n^2 \geq \frac{1}{|\mathcal{W}|}\left(\sum_{n\in\mathcal{W}} \text{SNR}_n\right)^2 = \frac{1}{Q}(N{\cdot}\text{SNR}_a)^2 = \frac{1}{\kappa_0}\text{tr}^2(\mathbf{SNR}), \tag{C.1}$$

as wanted to show.

C.2 Proof of (5.52)

By modeling the channel gains as independent complex Gaussian variables with unit gain, i.e., $h_l \sim \mathcal{CN}(0,1)$, it follows that the squared norm of the channel vector $\mathbf{h}$ is the sum of squares of $2L$ independent real Gaussian variables, as each term $|h_l|^2$ is the sum of the squares of the real and imaginary parts of h_l. Therefore, $\xi \doteq \|\mathbf{h}\|^2$ is Chi-squared distributed with $2L$ degrees of freedom, i.e.,

$$f(\xi) = \frac{1}{(L-1)!}\xi^{L-1}e^{-\xi}, \tag{C.2}$$

for $\xi \geq 0$. The former results permits to obtain the average error exponent (5.51) in Rayleigh fading with explicit calculations given by

$$\begin{aligned}
\mathbb{E}_\xi[\mathbb{D}(\mathcal{H}_1\|\mathcal{H}_0)] &= \int_0^\infty \mathbb{D}(\mathcal{H}_1\|\mathcal{H}_0) f(\xi) d\xi && \text{(C.3a)}\\
&= N \cdot \text{SNR} \int_0^\infty \xi \cdot \frac{1}{(L-1)!} \xi^{L-1} e^{-\xi} d\xi && \text{(C.3b)}\\
&= \frac{N \cdot \text{SNR}}{(L-1)!} \int_0^\infty \xi^L e^{-\xi} d\xi && \text{(C.3c)}\\
&= \frac{N \cdot \text{SNR}}{(L-1)!} L! && \text{(C.3d)}\\
&= LN \cdot \text{SNR}, && \text{(C.3e)}
\end{aligned}$$

which concludes the proof.

C.3 Proof of (5.60)

Being $\xi = \|\mathbf{h}\|^2$ Chi-squared distributed with $2L$ degrees of freedom, the average error exponent with respect to fading can be found as

$$\begin{aligned}
\mathbb{E}_\xi[\mathbb{D}(\mathcal{H}_1\|\mathcal{H}_0)] &= \int_0^\infty \mathbb{D}(\mathcal{H}_1\|\mathcal{H}_0) f(\xi) d\xi && \text{(C.4a)}\\
&= \frac{1}{2} \cdot \text{SNR}^2 \int_0^\infty \xi^2 \cdot \frac{1}{(L-1)!} \xi^{L-1} e^{-\xi} d\xi && \text{(C.4b)}\\
&= \frac{\text{SNR}^2}{2(L-1)!} \int_0^\infty \xi^{L+1} e^{-\xi} d\xi && \text{(C.4c)}\\
&= \frac{\text{SNR}^2}{2(L-1)!} (L+1)! && \text{(C.4d)}\\
&= \frac{1}{2}(L+1) \cdot L \cdot \text{SNR}^2, && \text{(C.4e)}
\end{aligned}$$

as wanted to show.

C.4 Proof of (5.66)

By considering Rayleigh channel with unit gain, it rapidly follows that in average the missed-detection probability is given by

$$\begin{aligned}
\mathbb{E}_{\xi}[P_{\mathrm{MD}}] &= \int_0^{\infty} P_{\mathrm{MD}} f(\xi) d\xi && \text{(C.5a)}\\
&= \frac{1}{\mathrm{SNR}} \int_0^{\infty} \frac{1}{\xi} \cdot \frac{1}{(L-1)!} \xi^{L-1} e^{-\xi} d\xi && \text{(C.5b)}\\
&= \frac{1}{\mathrm{SNR}(L-1)!} \int_0^{\infty} \xi^{L-2} e^{-\xi} d\xi && \text{(C.5c)}\\
&= \frac{1}{\mathrm{SNR}(L-1)!} (L-2)! && \text{(C.5d)}\\
&= \frac{1}{(L-1)\mathrm{SNR}}. && \text{(C.5e)}
\end{aligned}$$

C.5 Gaussian Fitting as Worst Case Scenario

This Appendix proves the claim that the Gaussian distribution acts as a worst case scenario in signal detection. This affirmation is supported by the Stein's lemma [CT91], which states that the error exponents in signal detection are proportional to the Kullback-Leibler divergence (KLD) between the probability density function (PDF) of the observations under $\mathcal{H}_0$ and $\mathcal{H}_1$, i.e., $\mathbb{D}(g_1||g_0)$. The KLD is related to the entropy as

$$\mathbb{D}(g_1||g_0) = \mathbb{H}(g_1, g_0) - \mathbb{H}(g_1), \tag{C.6}$$

where $\mathbb{H}(g_1, g_0)$ is the cross entropy between g_1 and g_0, and $\mathbb{H}(g_1)$ is the entropy of g_1. Let p_1 be an arbitrary non Gaussian distribution with the same variance as g_1. Then, the following inequalities on the entropy $\mathbb{H}(g_1) \geq \mathbb{H}(p_1)$, and cross entropy $\mathbb{H}(p_1, g_0) \geq \mathbb{H}(g_1, g_0)$ hold. Therefore, the KLD between an arbitrary equal power signal distribution p_1 and g_0 is lower bounded by the KLD between g_1 and g_0, i.e.,

$$\mathbb{D}(p_1||g_0) \geq \mathbb{D}(g_1||g_0). \tag{C.7}$$

As a result, the performance of the signal detectors derived in this thesis are provided in the worst case scenario with respect to any other real distribution of the primary signal.

C.6 Proof of (5.76)

The maximum likelihood (ML) estimate of the signal power is given by the convex optimization problem

$$\hat{P} = \arg\max_{P} \mathcal{CN}(\mathbf{0}, \mathbf{R}_1), \tag{C.8}$$

where $\mathbf{R}_1 = \sigma^2\mathbf{I}_N + P\boldsymbol{\Sigma}_\kappa$. By equaling the derivative with respect to P to zero, and making use of the matrix derivatives properties

$$\partial/\partial x \mathbf{A}^{-1} = -\mathbf{A}^{-1}(\partial/\partial x \mathbf{A})\mathbf{A}^{-1}, \tag{C.9}$$

and

$$\partial/\partial x \log\det(\mathbf{A}) = \mathrm{tr}(\mathbf{A}^{-1}\partial/\partial x \mathbf{A}), \tag{C.10}$$

it follows that the derivative is equivalent to the equation $\mathrm{tr}\left(\mathbf{R}_1^{-1}\boldsymbol{\Sigma}_\kappa\right) = \mathbf{x}^H\mathbf{R}_1^{-1}\boldsymbol{\Sigma}_\kappa\mathbf{R}_1^{-1}\mathbf{x}$. The solution of P has unfortunately no closed form expression, and numerical algorithms must be used. However, as $W \gg 1$, SNR $\ll 1$, and the following low SNR approximation holds $(\mathbf{I}_N + \mathrm{SNR}\mathbf{A})^{-1} \approx \mathbf{I}_N - \mathrm{SNR}\mathbf{A}$. Applying this result into $\mathbf{R}_1$ and neglecting all the terms containing $\mathcal{O}(\sigma^{2p})$ with $p > 2$, it follows that $P\mathrm{tr}(\boldsymbol{\Sigma}_\kappa^2) = \mathbf{x}^H\boldsymbol{\Sigma}_\kappa\mathbf{x} - \sigma^2\mathrm{tr}(\mathbf{I}_N)$. From the former expression, obtaining (5.76) is straightforward.

C.7 Proof of (5.78)

Consider the log likelihood ratio test

$$l(\mathbf{x}) \doteq \log L(\mathbf{x}), \tag{C.11}$$

which is equivalent as far as the logarithm is a monotonically increasing function. From (5.77), it follows that the log likelihood ratio test is given by the expression

$$l(\mathbf{x}) = \mathbf{x}^H\left(\frac{1}{\sigma^2}\mathbf{I}_N - \mathbf{R}_1^{-1}\right)\mathbf{x} - \log\det\left(\mathbf{I}_N + \frac{P}{\sigma^2}\boldsymbol{\Sigma}_\kappa\right). \tag{C.12}$$

For consistency with the derivation of the ML estimate of the signal power, in the sequel the following low SNR approximation will be considered by neglecting all the terms containing $\mathcal{O}(\sigma^{2p})$ with $p > 2$: $(\mathbf{I}_N + \mathrm{SNR}\mathbf{A})^{-1} \approx \mathbf{I}_N - \mathrm{SNR}\mathbf{A}$ and $\log\det(\mathbf{I}_N + \mathrm{SNR}\mathbf{A}) \approx \mathrm{SNR}\mathrm{tr}(\mathbf{A})$. Therefore, $l(\mathbf{x})$ can be simplified and rearranged to

$$l(\mathbf{x}) = \frac{P}{\sigma^2}\left(\mathbf{x}^H\boldsymbol{\Sigma}_\kappa\mathbf{x} - \sigma^2 N\right) \geq \tau, \tag{C.13}$$

where $\tau = \log\gamma$. Interestingly, $l(\mathbf{x})$ is proportional to the product between an estimate of the SNR and a term which is equal to the numerator of the signal power estimate (5.76). As a consequence, $l(\mathbf{x})$ is nonnegative. By plugging (5.76) into $l(\mathbf{x})$, taking the squared root and moving all the data independent terms into the threshold, the generalized likelihood ratio test (GLRT) signal detector is finally given by (5.78), where its associated threshold is given by (5.79).

C.8 Proof of (5.82)

From the statistically equivalent test (5.80) and (5.81), it follows that the means and variances of $T(\mathbf{x})$ under $\mathcal{H}_0$ and $\mathcal{H}_1$ are given, respectively, by

$$\mathbb{E}_0\left[T(\mathbf{x})\right] = N\sigma^2 \tag{C.14a}$$

$$\mathbb{E}_1\left[T(\mathbf{x})\right] = N\sigma^2 + P\sum_{n=1}^{r}\lambda_n^2(\mathbf{\Sigma}_\kappa), \tag{C.14b}$$

$$\mathbb{V}_0\left[T(\mathbf{x})\right] = 2\sigma^4\sum_{n=1}^{r}\lambda_n^2(\mathbf{\Sigma}_\kappa), \tag{C.14c}$$

$$\mathbb{V}_1\left[T(\mathbf{x})\right] = 2\sum_{n=1}^{r}\lambda_n^2(\mathbf{\Sigma}_\kappa)\left[\sigma^2 + P\lambda_n(\mathbf{\Sigma}_\kappa)\right]^2. \tag{C.14d}$$

By noting that $\sum_{n=1}^{r}\lambda_n^p(\mathbf{\Sigma}_\kappa) = \mathrm{tr}(\mathbf{\Sigma}_\kappa^p)$ for any power $p \geq 1$, it follows that the means and variances are finally given by

$$\mathbb{E}_0\left[T(\mathbf{x})\right] = N\sigma^2, \tag{C.15a}$$

$$\mathbb{E}_1\left[T(\mathbf{x})\right] = N\sigma^2 + NP\rho, \tag{C.15b}$$

$$\mathbb{V}_0\left[T(\mathbf{x})\right] = 2N\sigma^4\rho, \tag{C.15c}$$

$$\mathbb{V}_1\left[T(\mathbf{x})\right] = 2\mathrm{tr}(\mathbf{\Sigma}_\kappa^2(\sigma^2\mathbf{I}_N + P\mathbf{\Sigma}_\kappa)^2) \tag{C.15d}$$

By taking N and σ^2 as common factors, the asymptotic PDF of $T(\mathbf{x})$ are hence given by (5.82), under $\mathcal{H}_0$ and $\mathcal{H}_1$.

C.9 Spectrum Flatness as Worst Case Scenario

This Appendix proves the claim that any signal with non-flat spectrum will exhibit a squared correlation coefficient larger than that of a flat spectrum signal, in the same conditions of power and occupancy. For sake of simplicity, a sampling density of $\kappa = 1$ is considered, as the following results hold for any κ. Let $\phi(\omega)$ be the normalized spectrum of the primary signal $S(t)$, accomplishing the normalization

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\phi(\omega)d\omega = 1 \tag{C.16}$$

such that the power is normalized. If the occupancy of $S(t)$ is κ_0, the spectral support of $\phi(\omega)$, denoted as $\mathcal{W}_s$, has cardinality $|\mathcal{W}_s| = 2\pi\kappa_0$. Let $\nu(\omega)$ be the spectral mask of occupancy κ_0 given by

$$\nu(\omega) = \begin{cases} 1 & \text{if } \omega \in \mathcal{W}_s \\ 0 & \text{elsewhere.} \end{cases} \tag{C.17}$$

Therefore, because $\phi(\omega) = \phi(\omega)\nu(\omega)$ it follows that

$$1 = \left(\int_{-\pi}^{\pi} \phi(\omega) \cdot \frac{\nu(\omega)}{2\pi} d\omega \right)^2 \tag{C.18a}$$

$$\leq \int_{-\pi}^{\pi} \phi^2(\omega) d\omega \cdot \int_{-\pi}^{\pi} \left(\frac{\nu(\omega)}{2\pi} \right)^2 d\omega \tag{C.18b}$$

$$= \underbrace{\frac{1}{2\pi} \int_{-\pi}^{\pi} \phi^2(\omega) d\omega}_{\rho} \cdot \underbrace{\frac{1}{2\pi} \int_{-\pi}^{\pi} \nu^2(\omega) d\omega}_{\doteq W_\nu} \tag{C.18c}$$

where the Cauchy-Schwarz inequality was employed, and W_ν is the equivalent bandwidth of the spectral mask $\nu(\omega)$. Therefore, the squared correlation coefficient is lower bounded by the inverse of the equivalent bandwidth, i.e.,

$$\rho \geq \frac{1}{W_\nu}. \tag{C.19}$$

The lower bound is achieved with equality if and only if $\phi(\omega) \propto \nu(\omega)$, i.e., if $S(t)$ has flat spectrum. Therefore, any other spectrum shape will exhibit a larger squared correlation coefficient, as claimed.

C.10 Proof of (5.90)

Consider the PDF of $T(\mathbf{x})$ given in (5.82). Firstly, the false-alarm probability is evaluated by means of the right tail probability of a Gaussian distribution, i.e.,

$$P_{\mathrm{FA}} = Q\left(\frac{\tau/\sigma^2 - N}{\sqrt{2N\rho}} \right), \tag{C.20}$$

where $Q(x) = 1/\sqrt{2\pi} \int_x^\infty \exp(-t^2/2) dt$. From the former expression, the threshold τ is set to satisfy the target false-alarm probability constraint $P_{\mathrm{FA}} = \epsilon_0$ as

$$\tau = \sigma^2 \left[N + \sqrt{2N\rho} Q^{-1}(\epsilon_0) \right]. \tag{C.21}$$

Further, the missed-detection probability is given, employing the former expression of the threshold, by

$$P_{\mathrm{MD}} = Q\left(\frac{N\mathrm{SNR}\rho - \sqrt{2N\rho} Q^{-1}(\epsilon_0)}{\sqrt{2\mathrm{tr}[\mathbf{\Sigma}_\kappa^2 (\mathbf{I}_N + \mathrm{SNR}\mathbf{\Sigma}_\kappa)^2]}} \right). \tag{C.22}$$

Solving for the minimum SNR that achieves the missed-detection probability constraint is a difficult task from (C.22). However, for small SNR the approximation $\mathrm{tr}[\mathbf{\Sigma}_\kappa^2 (\mathbf{I}_N + \mathrm{SNR}\mathbf{\Sigma}_\kappa)^2] \approx \mathrm{tr}(\mathbf{\Sigma}_\kappa^2) = N\rho$ holds. Hence, applying $P_{\mathrm{MD}} = \epsilon_1$ with the former approxima-

tion, it follows that the minimum SNR is given by the expression

$$\text{SNR}_{\min} = \frac{\sqrt{2}\left[Q^{-1}(\epsilon_0) + Q^{-1}(\epsilon_1)\right]}{\sqrt{N\rho}}. \tag{C.23}$$

As the numerator of $\text{SNR}_{\min}$ is a constant which depends on the design parameters (ϵ_0, ϵ_1), it is straightforward to see that the minimum SNR scales as (5.90).

C.11 Proof of (5.95)

Consider $\hat{\sigma}^2$ as a realization of the prior information on the noise variance. From the expression of the false-alarm probability (C.20), it is seen that the noise variance and the threshold affect in proportion τ/σ^2. Therefore, the uncertainty on the noise variance translates to a wrong setting of the threshold (C.21). By adopting the ϵ-outage probability on the false-alarm probability $\mathbb{P}[P_{\text{FA}} \geq \epsilon_0] < \epsilon$, the threshold is given, after some mathematical manipulations, as $\hat{\tau} = \frac{\hat{\sigma}^2}{1/\delta+\epsilon(\delta-1/\delta)}\left[N + \sqrt{2N\rho}Q^{-1}(\epsilon_0)\right]$. Also, because $Q(x)$ is a monotonically decreasing function in its argument, the value of $\hat{\sigma}^2$ that evaluates the ϵ-outage probability on the missed-detection probability $\mathbb{P}[P_{\text{MD}} \geq \epsilon_1] < \epsilon$ is given by $\hat{\sigma}^2 = [\delta - \epsilon(\delta - 1/\delta)]\,\sigma^2$. Therefore, the threshold in the ϵ-outage uncertainty model is given, after some mathematical manipulations, by

$$\hat{\tau} = U\tau, \tag{C.24}$$

where U has been defined as in (5.96) and τ is given in (C.21). From (C.24), it follows that the missed-detection probability is now given by

$$P_{\text{MD}} = Q\left(\frac{N(1-U) + N\text{SNR}\rho - \sqrt{2N\rho}UQ^{-1}(\epsilon_0)}{\sqrt{2\text{tr}[\mathbf{\Sigma}_\kappa^2(\mathbf{I}_N + \text{SNR}\mathbf{\Sigma}_\kappa)^2]}}\right). \tag{C.25}$$

Finally, from the condition $P_{\text{MD}} = \epsilon_1$ and making the analogous developments as in Appendix C.10, the minimum SNR solves for

$$\text{SNR}_{\min} = \frac{\sqrt{2}\left[UQ^{-1}(\epsilon_0) + Q^{-1}(\epsilon_1)\right] + \sqrt{N/\rho}(U-1)}{\sqrt{N\rho}}, \tag{C.26}$$

which proves the scaling (5.95).

Appendix D

Proofs of Chapter 6

D.1 Proof of (6.6)

The optimality is proved in the correlation-matching sense of the blind detector for time-varying frequency-selective channels (6.6). Specifically, the negative correlation-matching between the normalized autocorrelation of the data and a scaled version of the desired signal autocorrelation is adopted, i.e.,

$$T(\mathbf{x}) = \max_{\substack{\mathbf{R}_x \in \Omega \\ 0<\gamma<1}} -\left\| \frac{\mathbf{x}\mathbf{x}^H}{\mathbf{x}^H\mathbf{x}} - \gamma \mathbf{R}_x \right\|_\eta^2, \tag{D.1}$$

where γ is the coherence factor defined as

$$\gamma \approx \frac{\text{SNR}}{\text{SNR}+1} \tag{D.2}$$

and it conveniently scales the signal autocorrelation matrix $\mathbf{R}_x$ of the signal samples (6.2), i.e.,

$$\mathbf{R}_x = \mathbf{A_H} \mathbf{A_H}^H. \tag{D.3}$$

In (D.1), the maximization of the autocorrelation matrix is done over the space of all Hermitian matrices Ω, and the correlation-matching is implemented through a quadratic norm

η. Expanding the norm in (D.1) it follows that

$$-\left\|\frac{\mathbf{x}\mathbf{x}^H}{\mathbf{x}^H\mathbf{x}}-\gamma\mathbf{R}_x\right\|_\eta^2=-1+2\gamma\frac{\mathbf{x}^H\mathbf{R}_x\mathbf{x}}{\mathbf{x}^H\mathbf{x}}-\gamma^2\left\|\mathbf{R}_x\right\|_\eta^2. \tag{D.4}$$

The maximization with respect to the coherence factor γ leads to the following expression

$$\hat{\gamma}=\frac{\mathbf{x}^H\mathbf{R}_x\mathbf{x}}{(\mathbf{x}^H\mathbf{x})\left\|\mathbf{R}_x\right\|_\eta^2}. \tag{D.5}$$

By substituting (D.5) into (D.4), (D.1) simplifies to

$$-\left\|\frac{\mathbf{x}\mathbf{x}^H}{\mathbf{x}^H\mathbf{x}}-\hat{\gamma}\mathbf{R}_x\right\|_\eta^2=-1+\frac{\left(\mathbf{x}^H\mathbf{R}_x\mathbf{x}\right)^2}{\left(\mathbf{x}^H\mathbf{x}\right)^2\left\|\mathbf{R}_x\right\|_\eta^2}. \tag{D.6}$$

Finally, removing the additive constants and taking the squared root it follows that the maximization problem (D.1) is equivalent to

$$T(\mathbf{x})=\frac{1}{\mathbf{x}^H\mathbf{x}}\max_{\mathbf{R}_x\in\Omega}\frac{\mathbf{x}^H\mathbf{R}_x\mathbf{x}}{\left\|\mathbf{R}_x\right\|_\eta}. \tag{D.7}$$

To this point, an optimal test from a correlation-matching perspective general for any received signal model $\mathbf{x}$ has been derived. For fading channels, the received signal is given by (6.2), and its autocorrelation matrix as (D.3). Therefore, the numerator of (D.7) can be expressed as

$$\mathbf{x}^H\mathbf{R}_x\mathbf{x}=\sum_{k=0}^{K-1}\mathbf{h}_k^H\mathbf{x}_k\mathbf{x}_k^H\mathbf{h}_k, \tag{D.8}$$

where $\mathbf{h}_k$ and $\mathbf{x}_k$ have been previously defined in (6.5) and (6.8). The structure of (D.8) makes the optimization with respect to the channel coefficients a difficult problem. To simplify it, the study resorts to large data records. Reformulating (D.8) with the trace operator and dividing it by K it is obtained

$$\mathrm{tr}\left(\frac{1}{K}\sum_{k=0}^{K-1}\mathbf{h}_k\mathbf{h}_k^H\mathbf{x}_k\mathbf{x}_k^H\right). \tag{D.9}$$

If K is large, in virtue of the law of large numbers, or equivalently the ergodicity property, the sample mean approaches to the statistical mean. In addition, as the channel realization are statistically independent on the signal plus noise realizations, the statistical mean can be further factorized. As a result of both assessments,

$$\mathrm{tr}\left(\frac{1}{K}\sum_{k=0}^{K-1}\mathbf{h}_k\mathbf{h}_k^H\mathbf{x}_k\mathbf{x}_k^H\right)\simeq\mathrm{tr}\left(\frac{1}{K}\sum_{k=0}^{K-1}\mathbf{h}_k\mathbf{h}_k^H\times\frac{1}{K}\sum_{k=0}^{K-1}\mathbf{x}_k\mathbf{x}_k^H\right). \tag{D.10}$$

Hence, removing constants the numerator yields to

$$\mathbf{x}^H \mathbf{R}_x \mathbf{x} \simeq \mathrm{tr}(\mathbf{H}\mathbf{H}^H \hat{\mathbf{R}}), \tag{D.11}$$

where $\hat{\mathbf{R}}$ is defined from (D.10) as in (6.7). On the other hand, the denominator norm is adopted as the following:

$$\|\mathbf{R}_x\|_\eta = \|\mathbf{H}\mathbf{H}^H\|_F, \tag{D.12}$$

where $\|\cdot\|_F$ is the Frobenius norm, which preserves the invariance property of the detector. Finally, as both (D.11) and (D.12) strictly depend on the channel matrix $\mathbf{H}$, the maximization with respect to $\mathbf{R}_x$ is equivalent to the maximization with respect to $\mathbf{H}$. As a result, the test (D.7) rewrites

$$T(\mathbf{x}) = \frac{1}{\mathbf{x}^H \mathbf{x}} \max_{\mathbf{H}} \frac{\mathrm{tr}(\mathbf{H}\mathbf{H}^H \hat{\mathbf{R}})}{\|\mathbf{H}\mathbf{H}^H\|_F}, \tag{D.13}$$

as wanted to show.

D.2 Proof of (6.9)

From the invariant detector (D.13), the Hermitian matrix $\mathbf{H}\mathbf{H}^H$ is expressed as the linear combination

$$\mathbf{H}\mathbf{H}^H = \sum_{i=1}^{L} \mu_i \cdot \mathbf{v}_i \mathbf{v}_i^H, \tag{D.14}$$

where μ_i and $\mathbf{v}_i$ denote the corresponding eigenvalues and eigenvectors, respectively. Therefore, the optimization problem (D.7) can be equivalently reformulated as maximizing the numerator

$$T(\mathbf{x}) = \frac{1}{\mathbf{x}^H \mathbf{x}} \max_{\{\mu_i, \mathbf{v}_i\}} \sum_{i=1}^{L} \mu_i \mathbf{v}_i^H \hat{\mathbf{R}} \mathbf{v}_i, \tag{D.15}$$

subject to the denominator constraint $\|\mathbf{H}\mathbf{H}^H\|_F = \sum_{i=1}^{L} \mu_i^2 = 1$, and $\mathbf{v}_i^H \mathbf{v}_j = \delta_{ij}$ for $i, j = 1, \ldots, L$. First, the developments solve for the eigenvectors from the Lagrangian

$$L(\mathbf{v}_1, \ldots, \mathbf{v}_L) = \sum_{i=1}^{L} \mu_i \mathbf{v}_i^H \hat{\mathbf{R}} \mathbf{v}_i + \sum_{i=1}^{L} \rho_i (1 - \mathbf{v}_i^H \mathbf{v}_i), \tag{D.16}$$

where $\rho_1, \ldots, \rho_L$ are the Lagrange multipliers. By computing the derivative and setting it to zero, it yields to $\hat{\mathbf{R}}\mathbf{v}_i = (\rho_i/\mu_i)\mathbf{v}_i$. That is, the eigenvectors of $\mathbf{H}\mathbf{H}^H$ are given by

$$\mathbf{v}_i = \mathbf{u}_i, \tag{D.17}$$

where $\mathbf{u}_i$ denote the eigenvectors of $\hat{\mathbf{R}}$, and ρ_i/μ_i are the eigenvalues of $\hat{\mathbf{R}}$, denoted by λ_i. Notice that the orthonormality constraint on the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_L$ is automatically

fulfilled as the eigenvectors of $\hat{\mathbf{R}}$ are orthonormal. Next, solving for the eigenvalues as $L(\mu_1,\ldots,\mu_L) = \sum_{i=1}^{L} \mu_i\lambda_i + \mu(1-\sum_{i=1}^{L}\mu_i^2)$ involves computing the derivative and setting it to zero. This leads to

$$\mu_i = \frac{\lambda_i}{\sqrt{\sum_{j=1}^{L}\lambda_j^2}}. \tag{D.18}$$

Finally, asymptotically for large K, the $\mathbf{x}^H\mathbf{x}$ term can be approximated by

$$\mathbf{x}^H\mathbf{x} \simeq \frac{N}{L}\mathrm{tr}(\hat{\mathbf{R}}) = \frac{N}{L}\sum_{i=1}^{L}\lambda_i. \tag{D.19}$$

Using (D.17), (D.18) and (D.19) into (D.15) and removing constants, (6.9) is obtained.

D.3 Proof of (6.28)

The derivation of (6.28) involves the maximum likelihood (ML) estimation of the nuisance parameters, i.e., the maximization of $\ln \mathrm{GLR}(\mathbf{x}) \doteq L(\mathbf{x}|\mathcal{H}_1) - L(\mathbf{x}|\mathcal{H}_0)$ with respect to (6.27). Splitting the integral of the likelihood function in sub-intervals, it reads

$$L(\mathbf{x}|\mathcal{H}_1) = -\frac{|\mathcal{W}_s|}{K}\sum_k \ln N_k - \sum_k \frac{1}{N_k}\hat{P}_k - |\mathcal{W}_n|\ln N_0 - \frac{1}{N_0}\sum_k \hat{\sigma}_k^2, \tag{D.20}$$

where the signal and power estimates are given as

$$\hat{P}_k = \int_{\mathcal{W}_k} \left|\mathbf{h}_k^H(v)\mathbf{x}(v)\right|^2 dv = \mathbf{h}_k^H\hat{\mathbf{B}}_k\mathbf{h}_k \tag{D.21a}$$

$$\hat{\sigma}_k^2 = \int_{\mathcal{W}_k} \left|\mathbf{u}_k^{\perp H}(v)\mathbf{x}(v)\right|^2 dv. \tag{D.21b}$$

Under $\mathcal{H}_0$, the ML estimate of the noise floor is given by

$$\hat{N}_0^0 = \frac{1}{|\mathcal{W}_s| + |\mathcal{W}_n|}\left(\hat{P} + \hat{\sigma}^2\right) \tag{D.22a}$$

$$= \frac{1}{B}\int_{-B/2}^{B/2} |X(f)|^2 df \tag{D.22b}$$

$$\doteq \frac{1}{B}\hat{P}_x, \tag{D.22c}$$

where

$$\hat{P} = \sum_k \hat{P}_k, \tag{D.23}$$

and

$$\hat{\sigma}^2 = \sum_k \hat{\sigma}_k^2. \tag{D.24}$$

Concerning $\mathcal{H}_1$, the ML estimates of $\{N_k\}$ and N_0 are given as

$$\hat{N}_k = \frac{1}{|\mathcal{W}_s|/K}\hat{P}_k \tag{D.25}$$

and

$$\hat{N}_0^1 = \frac{1}{|\mathcal{W}_n|}\hat{\sigma}^2 = \frac{(1+\alpha)\hat{N}_0^0 - \frac{1}{K}\sum_k \hat{N}_k}{\alpha} \tag{D.26}$$

Substituting these estimates to the likelihood function and forming the generalized likelihood ratio test (GLRT) yields

$$\max_{\{\mathbf{h}_k\}} \frac{1}{\left(\prod_{k=1}^K \lambda_k\right)^{1/K}\left(1-\sum_{k=1}^K \lambda_k\right)^{\alpha}} \geq \tau, \tag{D.27}$$

with

$$\lambda_k = \frac{\mathbf{h}_k^H \mathbf{B}_k \mathbf{h}_k}{\hat{P}_x} \tag{D.28}$$

Maximizing with respect to the $\mathbf{h}_k$ for $k = 1, \ldots, K$, under the constraint $\|\mathbf{h}_k\|^2 = 1$ also for $k = 1, \ldots, \mathcal{K}_0\left[\phi_s(\omega), \phi_\nu(\omega)\right]$, yields (6.28).

Appendix

E Proofs of Chapter 7

E.1 Proof of (7.19)

Given (7.14), the components involved in the deflection are as follows. Under $\mathcal{H}_0$, the expected value of the test is given by only the noise component due to the orthogonality $\mathbf{g}_{\perp}^{H}(\nu)\mathbf{g}(\nu) = 0$, i.e.,

$$\mathbb{E}_0[T_{\mathrm{NS}}] = N_0. \tag{E.1}$$

Under $\mathcal{H}_1$, the expected value of the test will include the part of the noncongitive signal that remains after the orthogonal filtering:

$$\mathbb{E}_1[T_{\mathrm{NS}}] = N_0 + \frac{1}{|\mathcal{W}|}\int_{\mathcal{W}} \mathbf{g}_{\perp}^{H}(\nu)\mathbf{\Phi}_s(\nu)\mathbf{g}_{\perp}(\nu)d\nu. \tag{E.2}$$

Finally under $\mathcal{H}_0$, by making use of the following asymptotic property [RVV10]:

$$\frac{1}{MT}\mathbb{E}\left[S(\nu_1)S^*(\nu_2)\right] = \phi_s^{\nu_1-\nu_2}\left(\frac{\nu_1+\nu_2}{2}\right), \tag{E.3}$$

the variance of the test is given by

$$\mathbb{V}_0[T_{\mathrm{NS}}] = \frac{1}{|\mathcal{W}|^2}\int_{\nu}\int_{\eta} \mathbb{E}_0\left[\mathbf{g}_{\perp}^{H}(\nu)\mathbf{o}_x(\nu)\mathbf{o}_x^{H}(\nu)\mathbf{g}_{\perp}(\nu)\mathbf{g}_{\perp}^{H}(\eta)\mathbf{o}_x(\eta)\mathbf{o}_x^{H}(\eta)\mathbf{g}_{\perp}(\eta)\right] d\nu d\eta - \mathbb{E}_0^2[T_{\mathrm{NS}}] \tag{E.4}$$

Expanding terms,

$$\mathbb{V}_0[T_{\mathrm{NS}}] = \frac{1}{|\mathcal{W}|^2}\int_\nu\int_\eta \mathbf{g}_\perp^H(\nu)\left[\mathbf{\Phi}_y(\nu)+\mathbf{\Phi}_w(\nu)\right]\mathbf{g}_\perp(\nu)\times \tag{E.5a}$$

$$\mathbf{g}_\perp^H(\eta)\left[\mathbf{\Phi}_y(\eta)+\mathbf{\Phi}_w(\eta)\right]\mathbf{g}_\perp(\eta)d\nu d\eta + \tag{E.5b}$$

$$\frac{1}{|\mathcal{W}|^2}\int_\nu\int_\eta \mathbf{g}_\perp^H(\nu)\left[\mathbf{\Phi}_y(\nu,\eta)+\mathbf{\Phi}_w(\nu,\eta)\right]\mathbf{g}_\perp(\eta)\times \tag{E.5c}$$

$$\mathbf{g}_\perp^H(\eta)\left[\mathbf{\Phi}_y(\eta,\nu)+\mathbf{\Phi}_w(\eta,\nu)\right]\mathbf{g}_\perp(\nu)d\nu d\eta - N_0^2 \tag{E.5d}$$

$$= \frac{1}{|\mathcal{W}|^2}\int_\nu\int_\eta \mathbf{g}_\perp^H(\nu)\left[\mathbf{\Phi}_y(\nu,\eta)+\mathbf{\Phi}_w(\nu,\eta)\right]\mathbf{g}_\perp(\eta)\times \tag{E.5e}$$

$$\mathbf{g}_\perp^H(\eta)\left[\mathbf{\Phi}_y(\eta,\nu)+\mathbf{\Phi}_w(\eta,\nu)\right]\mathbf{g}_\perp(\nu)d\nu d\eta, \tag{E.5f}$$

where the cross-spectral coherence matrix is given by

$$\mathbf{\Phi}_y(\nu,\eta) = \begin{bmatrix} \phi_y^{\nu-\eta}\left(\frac{\nu+\eta+1/T}{2}\right) & \phi_y^{\nu-\eta+1/T}\left(\frac{\nu+\eta}{2}\right) \\ \phi_y^{\nu-\eta-1/T}\left(\frac{\nu+\eta}{2}\right) & \phi_y^{\nu-\eta}\left(\frac{\nu+\eta-1/T}{2}\right) \end{bmatrix}. \tag{E.6}$$

Since $y(t)$ is wide-sense cyclostationary with cycle frequencies that are multiples of the symbol rate, $\phi_y^\beta(f)$ is different from zero only for values of β of $-1/T$, 0, and $1/T$. Because the range of the integration is restricted to $\mathcal{W}$, the double integration is nonzero only if $\nu=\eta$. In such a case, $\mathbf{\Phi}_y(\nu,\eta)=\mathbf{\Phi}_y(\nu)\delta(\nu-\eta)$, and $\mathbf{\Phi}_w(\nu,\eta)=\mathbf{\Phi}_w(\nu)\delta(\nu-\eta)$, where $\delta(\nu-\eta)$ is the Dirac delta. Because $\mathbf{g}_\perp^H(\nu)\mathbf{\Phi}_y(\nu)\mathbf{g}_\perp(\nu)=0$, the computation of the variance reduces to the following integral

$$\mathbb{V}_0[T_{\mathrm{NS}}] = \frac{N_0^2}{|\mathcal{W}|^2}\int_{\nu\in\mathcal{W}}\int_{\eta\in\mathcal{W}} \|\mathbf{g}_\perp(\nu)\|^4\delta(\nu-\eta)d\nu d\eta \tag{E.7}$$

Because $\|\mathbf{g}_\perp(\nu)\|^2=1$, the variance under $\mathcal{H}_0$ rapidly reduces to

$$\mathbb{V}_0\left[T_{\mathrm{NS}}\right] = \frac{N_0^2}{|\mathcal{W}|}. \tag{E.8}$$

As a consequence, it is observed that the proposed detector shows consistency within the integration interval under hypothesis $\mathcal{H}_0$. Finally, plugging the derived statistics (E.1), (E.2), and (E.8) into (7.18), it is obtained that the deflection is given by (7.19).

E.2 Proof of (7.20)

In this Appendix, an upper bound on the deflection (7.19) will be obtained in order to learn how the deflection of the noise subspace detector scales with the problem parameters. Firstly, it is noted that the deflection is implementing an inner product map in the

Hermitian matrix space $\mathbb{M}$ defined as

$$\langle \cdot, \cdot \rangle \; : \; \mathbb{M} \times \mathbb{M} \;\; \mapsto \;\; \mathbb{R}_+ \tag{E.9a}$$

$$\langle \mathbf{X}_1(\nu), \mathbf{X}_2(\nu) \rangle \;\; \mapsto \;\; \int_{\mathcal{W}} \mathrm{tr}\left(\mathbf{X}_1(\nu) \cdot \mathbf{X}_2(\nu)\right) d\nu, \tag{E.9b}$$

it follows from the Cauchy-Schwarz inequality over this inner product, i.e.,

$$|\langle \mathbf{X}_1(\nu), \mathbf{X}_2(\nu) \rangle|^2 \leq \|\mathbf{X}_1(\nu)\| \cdot \|\mathbf{X}_2(\nu)\|, \tag{E.10}$$

that

$$\mathbb{D}(T_{\mathrm{NS}}) = \frac{1}{N_0^2 |\mathcal{W}|} \left(\int_{\mathcal{W}} \mathbf{g}_\perp^H(\nu) \mathbf{\Phi}_s(\nu) \mathbf{g}_\perp(\nu) d\nu \right)^2 \tag{E.11a}$$

$$= \frac{1}{N_0^2 |\mathcal{W}|} \left(\int_{\mathcal{W}} \mathrm{tr}\left(\mathbf{\Phi}_s(\nu) \mathbf{g}_\perp(\nu) \mathbf{g}_\perp^H(\nu) \right) d\nu \right)^2 \tag{E.11b}$$

$$\leq \frac{1}{N_0^2 |\mathcal{W}|} \int_{\mathcal{W}} \mathrm{tr}\left(\mathbf{\Phi}_s^2(\nu) \right) d\nu \cdot \int_{\mathcal{W}} \|\mathbf{g}_\perp(\nu)\|^4 d\nu \tag{E.11c}$$

$$= \frac{1}{N_0^2} \int_{\mathcal{W}} \mathrm{tr}\left(\mathbf{\Phi}_s^2(\nu) \right) d\nu, \tag{E.11d}$$

as wanted to show.

E.3 Proof of (7.34) and (7.35)

In this Appendix, the expected value of the MDL(n) statistic (7.28) is addressed. In particular, as

$$\log \mathrm{GLR}(n) \sim \mathcal{X}^2_{r_n}(\mu_n), \tag{E.12}$$

the following relationship

$$\mu_n + r_n = \mathbb{E}\left[\log \mathrm{GLR}(n)\right] \tag{E.13}$$

is employed to compute μ_n under $\mathcal{H}_N$. For sake of simplicity, a white primary signal is considered, whose power is uniformly distributed in the N dimensions of the signal subspace, i.e., the expected eigenvalues of the sample covariance matrix (7.29) are

$$\mathbb{E}[\lambda_n] = P + \sigma^2 \quad \text{for} \quad 1 \leq n \leq N \tag{E.14a}$$

$$\mathbb{E}[\lambda_n] = \sigma^2 \quad \text{for} \quad N+1 \leq n \leq M, \tag{E.14b}$$

where P and σ^2 are related by the average signal-to-noise ratio (SNR) through

$$\mathrm{SNR} = \frac{NP}{\sigma^2}. \tag{E.15}$$

If $n \geq N$, it then follows that the second term of MDL(n) is

$$\mathbb{E}\left[L(M-n)\log\frac{\prod_{m=n+1}^{M}\lambda_m^{1/(M-n)}}{\frac{1}{M-n}\sum_{m=n+1}^{M}\lambda_m}\right] \geq L(M-n)\log\mathbb{E}\left[\frac{\prod_{m=n+1}^{M}\lambda_m^{1/(M-n)}}{\frac{1}{M-n}\sum_{m=n+1}^{M}\lambda_m}\right] = 0, \quad \text{(E.16)}$$

whereas the first term is

$$\mathbb{E}\left[LM\log\frac{\prod_{m=1}^{M}\lambda_m^{1/M}}{\frac{1}{M}\sum_{m=1}^{M}\lambda_m}\right] \geq LM\log\mathbb{E}\left[\frac{\prod_{m=1}^{M}\lambda_m^{1/M}}{\frac{1}{M}\sum_{m=1}^{M}\lambda_m}\right], \quad \text{(E.17)}$$

which employing $\mathbb{E}[\lambda_n]$ it proves (7.34).

If $n < N$, then the second term of the MDL(n) is non-zero. Hence, after some mathematical manipulations on (E.16) and together with (E.17), (7.35) is obtained.

E.4 Proof of (7.37)

The auxiliary random variable D_{mn} is defined as the difference between the following GLRs

$$D_{mn} \doteq \log \text{GLR}(m) - \log \text{GLR}(n). \quad \text{(E.18)}$$

Employing (7.26),

$$D_{mn} = (m-n)M\log(L) - \text{MDL}(m) + \text{MDL}(n). \quad \text{(E.19)}$$

By noting that

$$P_n = \prod_{m\neq n}\mathbb{P}\left[\text{MDL}(n) < \text{MDL}(m)\right] \quad \text{(E.20a)}$$

$$= \prod_{m\neq n}\mathbb{P}\left[\text{MDL}(n) - \text{MDL}(m) < 0\right] \quad \text{(E.20b)}$$

the probability P_n is related to D_{mn} as

$$P_n = \prod_{m\neq n}\mathbb{P}\left[D_{mn} < (m-n)M\log(L)\right]. \quad \text{(E.21)}$$

Further, as the generalized likelihood ratio (GLR) statistic is asymptotically non-central Chi-squared distributed, the difference is also non-central Chi-squared distributed as

$$D_{mn} \sim \mathcal{X}^2_{r_m - r_n}(\mu_m - \mu_n), \quad \text{(E.22)}$$

if $m > n$ and

$$-D_{mn} \sim \mathcal{X}^2_{r_m - r_n}(\mu_m - \mu_n), \quad \text{(E.23)}$$

if $m < n$. As a consequence, the probability P_n can be computed by means of the cumulative density function (CDF) of D_{mn} as

$$P_n = \prod_{m<n} \text{CDF}_{D_{mn}} \left((m-n) M \log(L) \right) \tag{E.24}$$

which further employing (E.22) and (E.22) proves (7.37).

Note that the CDF of a non-central Chi-squared distribution is given by

$$\text{CDF}_{\mathcal{X}^2_r(\mu)}(x) = e^{-\mu/2} \sum_{j=0}^{\infty} \frac{(\mu/2)^j}{j!} Q(x, r+2j), \tag{E.25}$$

where $Q(x, r)$ is the CDF of a central Chi-squared distribution with r degrees of freedom. The CDF of the negative non-central Chi-squared distribution is given by symmetry as

$$\text{CDF}_{-\mathcal{X}^2_r(\mu)}(x) = 1 - \text{CDF}_{\mathcal{X}^2_r(\mu)}(x). \tag{E.26}$$

E.5 Proof of (7.41)

For the minimum description length (MDL) with known noise variance, the maximum likelihood (ML) estimation of the eigenvalues and eigenvectors of $\mathbf{R}_s$ are given as

$$\hat{\mathbf{u}}_m(\mathbf{R}_s + \sigma^2 \mathbf{I}) = \mathbf{u}_m(\hat{\mathbf{R}}), \tag{E.27}$$

and

$$\hat{\lambda}_m(\mathbf{R}_s + \sigma^2 \mathbf{I}) = \lambda_m(\hat{\mathbf{R}}) \tag{E.28}$$

respectively, for $m = 1, \ldots, n$. Therefore, the denominator and numerator of the GLR (7.39) reads

$$\log p(\mathbf{x}|\mathcal{H}_0) = -ML \log \pi - ML \log \sigma^2 - \frac{L}{\sigma^2} \sum_{m=1}^{M} \lambda_m, \tag{E.29}$$

and

$$\log p(\mathbf{x}|\mathcal{H}_n) = -ML \log \pi - ML \prod_{m=1}^{n} \lambda_m^{1/M} - ML \log \sigma^{2^{(M-n)/M}} - Ln - \frac{L}{\sigma^2} \sum_{m=n+1}^{M} \lambda_m, \tag{E.30}$$

respectively. Joining both terms,

$$\log \text{GLR}(n, \sigma^2) = \frac{L}{\sigma^2} \sum_{m=1}^{n} \lambda_m - ML \log \prod_{m=1}^{n} \lambda_m^{1/M} + Ln \log \sigma^2 - Ln. \tag{E.31}$$

After rearranging terms, the MDL with known noise variance is given as in (7.41).

E.6 Proof of (7.42) and (7.43)

Employing the relationships (E.13) and (E.14) of Appendix E.3, if $n \leq N$ it is noted that

$$\mathbb{E}\left(LM\left(\frac{1}{M}\sum_{m=1}^{n}\left(\frac{\lambda_m}{\sigma^2}-1\right)\right)\right) = LM\left(\frac{1}{M}\sum_{m=1}^{n}\frac{P}{\sigma^2}\right), \tag{E.32}$$

and

$$\mathbb{E}\left(LM\log\left(\prod_{m=1}^{n}\left(\frac{\lambda_m}{\sigma^2}\right)^{1/M}\right)\right) \geq LM\log\left(\prod_{m=1}^{n}\left(1+\frac{P}{\sigma^2}\right)^{1/M}\right). \tag{E.33}$$

Employing the SNR definition (E.15) and joining terms, it proves (7.42). For $n > N$, it rapidly follows that the terms for $m > N$ are zero in the summation (E.32), whereas the terms for $m > N$ in the product (E.33) are one (hence zero in the logarithm). Therefore, for $n > N$,

$$\mathbb{E}\left(LM\left(\frac{1}{M}\sum_{m=1}^{n}\left(\frac{\lambda_m}{\sigma^2}-1\right)\right)\right) = LM\left(\frac{1}{M}\sum_{m=1}^{N}\frac{P}{\sigma^2}\right), \tag{E.34}$$

and

$$\mathbb{E}\left(LM\log\left(\prod_{m=1}^{n}\left(\frac{\lambda_m}{\sigma^2}\right)^{1/M}\right)\right) \geq LM\log\left(\prod_{m=1}^{N}\left(1+\frac{P}{\sigma^2}\right)^{1/M}\right), \tag{E.35}$$

which employing (E.15) proves (7.43).

E.7 Proof of (7.54)

Consider the optimization problem (7.53) for $k = 1$. The Lagrangian associated to this problem is given by

$$\mathcal{L}(\mathbf{a}_1) = \mathbf{a}_1^H\mathbf{R}\mathbf{a}_1 + \lambda(1-\mathbf{a}_1^H\mathbf{e}_1), \tag{E.36}$$

where λ is the Lagrange multiplier for the nontriviality constraint. By taking the derivative of (E.36) with respect to $\mathbf{a}_1^H$ and equating to zero, i.e.,

$$\frac{\partial}{\partial\mathbf{a}_1^H}\mathcal{L}(\mathbf{a}_1) = \mathbf{R}\mathbf{a}_1 - \lambda\mathbf{e}_1 = 0 \tag{E.37}$$

it follows that $\mathbf{a}_1$ obeys

$$\mathbf{R}\mathbf{a}_1 = \lambda\mathbf{e}_1. \tag{E.38}$$

Because $\mathbf{R}$ is rank deficient, the inverse is implemented by means of the following limit

$$\mathbf{a}_1 = \lim_{\delta\to 0}\lambda\left(\mathbf{R}+\delta\mathbf{I}\right)^{-1}\mathbf{e}_1. \tag{E.39}$$

Applying the nontriviality constraint $\mathbf{a}_1^H \mathbf{e}_1 = 1$ to solve for λ and plugging the result into (E.39), it reads

$$\mathbf{a}_1 = \frac{\lim_{\delta \to 0} (\mathbf{R} + \delta \mathbf{I})^{-1} \mathbf{e}_1}{\mathbf{e}_1^T \lim_{\delta \to 0} (\mathbf{R} + \delta \mathbf{I})^{-1} \mathbf{e}_1} \tag{E.40a}$$

$$= \frac{\left(\mathbf{I} - \mathbf{U}\mathbf{U}^H\right) \mathbf{e}_1}{\mathbf{e}_1^T \underbrace{\left(\mathbf{I} - \mathbf{U}\mathbf{U}^H\right)}_{\doteq \mathbf{P}_{\mathbf{U}}^{\perp}} \mathbf{e}_1}, \tag{E.40b}$$

where the inversion lemma[1] on the term $(\mathbf{R} + \delta \mathbf{I})$ has been applied, i.e.,

$$(\mathbf{R} + \delta \mathbf{I})^{-1} = \left(\mathbf{U}\mathbf{\Lambda}\mathbf{U}^H + \delta \mathbf{I}\right)^{-1} \tag{E.41a}$$

$$= \frac{1}{\delta} - \frac{1}{\delta^2} \mathbf{U} \left(\mathbf{\Lambda}^{-1} + \frac{1}{\delta} \mathbf{U}^H \mathbf{U}\right)^{-1} \mathbf{U}^H \tag{E.41b}$$

$$= \frac{1}{\delta} \left(\mathbf{I} - \mathbf{U} \left(\delta \mathbf{\Lambda}^{-1} + \mathbf{I}\right)^{-1} \mathbf{U}^H\right). \tag{E.41c}$$

From (E.40), it is observed that $\mathbf{a}_1$ depends on the orthogonal projector defined in (7.55) for $k = 1$. In order to obtain the subsequent waveforms, each waveform must be normalized to unit-norm. Therefore,

$$\mathbf{a}_1 = \frac{\mathbf{P}_{\mathbf{U}}^{\perp} \mathbf{e}_1}{\underbrace{\sqrt{\mathbf{e}_1^T \mathbf{P}_{\mathbf{U}}^{\perp} \mathbf{e}_1}}_{\doteq \sqrt{\alpha_1}}}, \tag{E.42}$$

where α_1 is the corresponding scaling factor.

On the other hand, the subsequent waveforms can be derived by incorporating the previous waveforms into the Gram matrix, i.e., as

$$\mathbf{R}_{k+1} = \mathbf{R}_k + \mathbf{a}_k \mathbf{a}_k^H. \tag{E.43}$$

By doing so, it rapidly follows that

$$\mathbf{U}_{k+1} = \left[\mathbf{U}_k \quad \mathbf{a}_k\right]. \tag{E.44}$$

Therefore, the solution the subsequent waveforms in (7.53) becomes

$$\mathbf{a}_k = \frac{\mathbf{P}_{\mathbf{U}_k}^{\perp} \mathbf{e}_k}{\underbrace{\sqrt{\mathbf{e}_k^T \mathbf{P}_{\mathbf{U}_k}^{\perp} \mathbf{e}_k}}_{\doteq \sqrt{\alpha_k}}}, \tag{E.45}$$

where α_k are the corresponding scaling factors, as wanted to show.

[1] $(\mathbf{A} + \mathbf{U}\mathbf{C}\mathbf{V})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{U}\left(\mathbf{C}^{-1} + \mathbf{V}\mathbf{A}^{-1}\mathbf{U}\right)^{-1}\mathbf{V}\mathbf{A}^{-1}$

E.8 Equivalence between (7.53) and (7.57)

In order to show the equivalence between the minimum interference problem (7.53) and the minimum norm with orthogonality constraint problem (7.57), this Appendix proves that the solution to (7.57) equals the solution (7.54).

The Lagrangian associated to the first waveform of (7.57) is

$$\mathcal{L}(\mathbf{a}_1) = \mathbf{a}_1^H \mathbf{a}_1 + \lambda_0(1 - \mathbf{a}_1^H \mathbf{e}_1) + (\mathbf{0}^T - \mathbf{a}_1^H \mathbf{U})\boldsymbol{\lambda}, \tag{E.46}$$

where λ_0 and $\boldsymbol{\lambda}$ are the scalar and vector Lagrange multipliers for the nontriviality and orthogonality constraints, respectively. Taking the derivative with respect to $\mathbf{a}_1^H$ leads to

$$\frac{\partial}{\partial \mathbf{a}_1^H}\mathcal{L}(\mathbf{a}_1) = \mathbf{a}_1 - \lambda_0 \mathbf{e}_1 - \mathbf{U}\boldsymbol{\lambda} = 0, \tag{E.47}$$

which leads to

$$\mathbf{a}_1 = \lambda_0 \mathbf{e}_1 + \mathbf{U}\boldsymbol{\lambda}. \tag{E.48}$$

Applying the second constraint of (7.57) leads to

$$\boldsymbol{\lambda} = -\lambda_0 \mathbf{U}^H \mathbf{e}_1, \tag{E.49}$$

which combined with (E.48) leads to

$$\mathbf{a}_1 = \lambda_0 \left(\mathbf{I} - \mathbf{U}\mathbf{U}^H\right) \mathbf{e}_1, \tag{E.50}$$

which becomes equivalent to (7.54) as wanted to show.

E.9 Proof of (7.64)

In order to address the optimization problem (7.63), the eigenvalue decomposition (EVD) of the semi-definite positive mismatching matrix is defined as

$$\boldsymbol{\Delta} = \sum_{i=1}^{D} \lambda_i(\boldsymbol{\Delta}) \mathbf{u}_i(\boldsymbol{\Delta}) \mathbf{u}_i^H(\boldsymbol{\Delta}). \tag{E.51}$$

Hence, (7.63) is now given as

$$\max_{\boldsymbol{\Delta}} \sum_{i=1}^{D} \lambda_i(\boldsymbol{\Delta}) \mathbf{u}_i^H(\boldsymbol{\Delta}) \tilde{\mathbf{a}}_k \cdot \tilde{\mathbf{a}}_k^H \mathbf{u}_i(\boldsymbol{\Delta}), \tag{E.52}$$

subject to $\lambda_1(\boldsymbol{\Delta}) \leq \epsilon$, where λ_1 is the maximum eigenvalue. Clearly, the maximization requires that $\boldsymbol{\Delta}$ takes the eigenvalues of the quadratic form $\tilde{\mathbf{a}}_k\tilde{\mathbf{a}}_k^H$, i.e.,

$$\mathbf{u}_1(\boldsymbol{\Delta}) = \mathbf{a}_k. \tag{E.53}$$

As a result, the maximum is achieved for $\lambda_1(\boldsymbol{\Delta}) = \epsilon$, and hence it proves (7.64).

E.10 Proof of (7.67)

The Lagrangian associated to (7.66) for $k = 1$ is given by

$$\mathcal{L}(\tilde{\mathbf{a}}_1) = \tilde{\mathbf{a}}_1^H \left(\hat{\mathbf{R}} + \epsilon\mathbf{I}\right) \tilde{\mathbf{a}}_1 + \lambda(1 - \tilde{\mathbf{a}}_1^H\mathbf{e}_1), \tag{E.54}$$

where λ is the Lagrange multiplier for the nontriviality constraint. Taking the derivative of (E.54) with respect to $\tilde{\mathbf{a}}_1^H$ and equating to zero, i.e.,

$$\frac{\partial}{\partial\tilde{\mathbf{a}}_1^H}\mathcal{L}(\tilde{\mathbf{a}}_1) = \left(\hat{\mathbf{R}} + \epsilon\mathbf{I}\right) \tilde{\mathbf{a}}_1 - \lambda\mathbf{e}_1 = 0, \tag{E.55}$$

it follows that

$$\begin{aligned}
\mathbf{a}_1 &= \frac{\left(\hat{\mathbf{R}} + \epsilon\mathbf{I}\right)^{-1}\mathbf{e}_1}{\mathbf{e}_1^T\left(\hat{\mathbf{R}} + \epsilon\mathbf{I}\right)^{-1}\mathbf{e}_1} && \text{(E.56a)}\\
&= \frac{\left(\mathbf{I} - \hat{\mathbf{U}}\boldsymbol{\Phi}_1\hat{\mathbf{U}}^H\right)\mathbf{e}_1}{\mathbf{e}_1^T\left(\mathbf{I} - \hat{\mathbf{U}}\boldsymbol{\Phi}_1\hat{\mathbf{U}}^H\right)\mathbf{e}_1}. && \text{(E.56b)}
\end{aligned}$$

To obtain (E.56), the inversion lemma on the $\left(\hat{\mathbf{R}} + \epsilon\mathbf{I}\right)$ term has been applied, i.e.,

$$\begin{aligned}
\left(\hat{\mathbf{R}} + \epsilon\mathbf{I}\right)^{-1} &= \left(\hat{\mathbf{U}}\hat{\boldsymbol{\Lambda}}\hat{\mathbf{U}}^H + \epsilon\mathbf{I}\right)^{-1} && \text{(E.57a)}\\
&= \frac{1}{\epsilon} - \frac{1}{\epsilon^2}\hat{\mathbf{U}}\left(\hat{\boldsymbol{\Lambda}}^{-1} + \frac{1}{\epsilon}\hat{\mathbf{U}}^H\hat{\mathbf{U}}\right)^{-1}\hat{\mathbf{U}}^H && \text{(E.57b)}\\
&= \frac{1}{\epsilon}\left(\mathbf{I} - \hat{\mathbf{U}}\underbrace{\left(\epsilon\hat{\boldsymbol{\Lambda}}^{-1} + \mathbf{I}\right)^{-1}}_{\doteq\boldsymbol{\Phi}_1}\hat{\mathbf{U}}\right), && \text{(E.57c)}
\end{aligned}$$

where the diagonal matrix $\mathbf{\Phi}_1$ has been defined as

$$\mathbf{\Phi}_1 = \begin{bmatrix} \frac{\hat{\lambda}_1}{\hat{\lambda}_1+\epsilon} & & 0 \\ & \ddots & \\ 0 & & \frac{\hat{\lambda}_N}{\hat{\lambda}_N+\epsilon} \end{bmatrix}. \tag{E.58}$$

Hence, after normalization, the first robust waveform reads

$$\tilde{\mathbf{a}}_1 = \frac{\left(\mathbf{I} - \hat{\mathbf{U}}\mathbf{\Phi}_1\hat{\mathbf{U}}^H\right)\mathbf{e}_1}{\underbrace{\sqrt{\mathbf{e}_1^T\left(\mathbf{I} - \hat{\mathbf{U}}\mathbf{\Phi}_1\hat{\mathbf{U}}^H\right)^2\mathbf{e}_1}}_{\doteq\sqrt{\tilde{\alpha}_1}}}, \tag{E.59}$$

where $\tilde{\alpha}_1$ is the corresponding scaling factor.

By mimicking the iterative procedure proposed in (E.43), i.e.,

$$\hat{\mathbf{R}}_{k+1} = \hat{\mathbf{R}}_k + \hat{\mathbf{a}}_k\hat{\mathbf{a}}_k^H, \tag{E.60a}$$

$$\hat{\mathbf{U}}_{k+1} = \begin{bmatrix} \hat{\mathbf{U}}_k & \hat{\mathbf{a}}_k \end{bmatrix}, \tag{E.60b}$$

and

$$\mathbf{\Phi}_{k+1} = \begin{bmatrix} \mathbf{\Phi}_k & 0 \\ 0 & 1 \end{bmatrix}, \tag{E.61}$$

it follows that the k-th robust waveform is given by

$$\tilde{\mathbf{a}}_k = \frac{\left(\mathbf{I} - \hat{\mathbf{U}}_k\mathbf{\Phi}_k\hat{\mathbf{U}}_k^H\right)\mathbf{e}_k}{\underbrace{\sqrt{\mathbf{e}_k^T\left(\mathbf{I} - \hat{\mathbf{U}}_k\mathbf{\Phi}_k\hat{\mathbf{U}}_k^H\right)^2\mathbf{e}_k}}_{\doteq\sqrt{\tilde{\alpha}_k}}} \tag{E.62}$$

as wanted to show.

E.11 Proof of (7.89)

Noticing that the first robust waveform admits the structure (E.56), the individual interference associated to the first robust waveform is given by

$$\begin{aligned}
\xi(\tilde{\mathbf{a}}_1) &= \frac{\mathbf{e}_1^T \left(\hat{\mathbf{R}} + \epsilon \mathbf{I}\right)^{-1} \boldsymbol{\Delta} \left(\hat{\mathbf{R}} + \epsilon \mathbf{I}\right)^{-1} \mathbf{e}_1}{\mathbf{e}_1^T \left(\hat{\mathbf{R}} + \epsilon \mathbf{I}\right)^{-2} \mathbf{e}_1} && \text{(E.63a)} \\
&\leq \frac{\mathbf{e}_1^T \left(\hat{\mathbf{R}} + \epsilon \mathbf{I}\right)^{-1} \mathbf{e}_1}{\mathbf{e}_1^T \left(\hat{\mathbf{R}} + \epsilon \mathbf{I}\right)^{-2} \mathbf{e}_1} && \text{(E.63b)} \\
&= \epsilon \cdot \frac{\mathbf{e}_1^T \left(\mathbf{I} - \mathbf{U}\boldsymbol{\Phi}_1\mathbf{U}^H\right) \mathbf{e}_1}{\mathbf{e}_1^T \left(\mathbf{I} - \mathbf{U}\boldsymbol{\Phi}_1\mathbf{U}^H\right)^2 \mathbf{e}_1}, && \text{(E.63c)}
\end{aligned}$$

where to obtain (E.63) the worst-case performance scenario has been considered, and to obtain (E.63) the inversion lemma has been applied. Expanding using (7.79) and (7.80), it leads to (7.89).

References

[ACCD13] E. Arias-Castro, E. J. Candès, and M. A. Davenport, "On the fundamental limits of adaptive sensing," *IEEE Transactions on Information Theory*, vol. 59, no. 1, pp. 472–481, Jan. 2013.

[AHDVP12] A. Al-Habashna, O. A. Dobre, R. Venkatesan, and D. C. Popescu, "Second-order cyclostationarity of mobile WiMAX and LTE OFDM signals and application to spectrum awareness in cognitive radio systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 6, no. 1, pp. 26–42, Feb. 2012.

[AKSL11] T. Al-Khasib, M. B. Shenouda, and L. Lampe, "Dynamic spectrum management for multiple-antenna cognitive radio systems: Designs with imperfect CSI," *IEEE Transactions on Wireless Communications*, vol. 10, no. 9, pp. 2850–2859, Sep. 2011.

[AL11] E. Axell and E. G. Larsson, "Optimal and sub-optimal spectrum sensing of ofdm signals in known and unknown noise variance," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 2, Feb. 2011.

[AL12] D. Ariananda and G. Leus, "Compressive wideband power spectrum estimation," *IEEE Transactions on Signal Processing*, vol. 60, no. 9, pp. 4775–4789, Sep. 2012.

[ALLP12] E. Axell, G. Leus, E. Larsson, and H. Poor, "Spectrum sensing for cognitive radio : State-of-the-art and recent advances," *IEEE Signal Processing Magazine*, vol. 29, no. 3, pp. 101–116, May 2012.

[Alt99] E. Altman, *Constrained Markov Decision Process*. Chapman & Hall/CRC, 1999.

[AS64] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Washington D. C.: National Bureau of Standards Applied Mathematics Series 55, 1964.

[AS12] S. M. Almalfouh and G. L. Stuber, "Joint spectrum-sensing design and power control in cognitive radio networks: A stochastic approach," *IEEE Transactions on Wireless Communications*, vol. 11, no. 12, pp. 4372–4380, Dec. 2012.

[ASZ10] S. Aeron, V. Saligrama, and M. Zhao, "Information theoretic bounds for compressed sensing," *IEEE Transactions on Information Theory*, vol. 56, no. 10, pp. 5111–5130, Oct. 2010.

[AT10] M. Akcakaya and V. Tarokh, "Shannon-theoretic limits on noisy compressive sampling," *IEEE Transactions on Information Theory*, vol. 56, no. 1, pp. 492–504, Jan. 2010.

[AWH07] J. G. Andrews, S. Weber, and M. Haenggi, "Ad hoc networks: to spread or not to spread," *IEEE Communications Magazine*, vol. 45, no. 12, pp. 84–91, 2007.

[BB99] S. Benedetto and E. Biglieri, *Principles of Digital Communications*, ser. Information Technology: Transmission, Processing, and Storage. New York: Kluwer Academic / Plenum Publishers, 1999.

[BB09] F. Bacceli and B. Blaszczyszyn, *Stochastic geometry and wireless networks*, ser. Foundations and Trends in Networking. Now Publishers, Inc., 2009, vol. 3-4.

[BBKB11] D. Bharadia, G. Bansal, P. Kaligineedi, and V. K. Bhargava, "Relay and power allocation schemes for OFDM-based cognitive radio systems," *IEEE Transactions on Wireless Communications*, vol. 10, no. 9, pp. 2812–2817, Sep. 2011.

[BDMN11] P. Bianchi, M. Debbah, M. Maida, and J. Najim, "Performance of statistical tests for single-source detection using random matrix theory," *IEEE Transactions on Information Theory*, vol. 57, no. 4, pp. 2400–2419, Apr. 2011.

[BFPRI12] S. W. Boyd, J. M. Frye, M. B. Pursley, and T. C. Royster IV, "Spectrum monitoring during reception in dynamic spectrum access cognitive radio networks," *IEEE Transactions on Communications*, vol. 60, no. 2, pp. 547–558, Feb. 2012.

[BFT11] A. Bagayoko, I. Fijalkow, and P. Tortelier, "Power control of spectrum-sharing in fading environment with partial channel state information," *IEEE Transactions on Signal Processing*, vol. 59, no. 5, pp. 2244–2256, May 2011.

[BG10] J. A. Bazerque and G. B. Giannakis, "Distributed spectrum sensing for cognitive radio networks by exploiting sparsity," *IEEE Transactions on Signal Processing*, vol. 58, no. 3, pp. 1847–1862, Mar. 2010.

[BGH09] Y. Baraud, C. Giraud, and S. Huet, "Gaussian model selection with an unknown variance," *The Annals of Statistics*, vol. 37, no. 2, pp. 630–672, 2009.

[BHB08] G. Bansal, M. Hossain, and V. Bhargava, "Optimal and suboptimal power allocation schemes for OFDM-based cognitive radio systems," *IEEE Transactions on Wireless Communications*, vol. 7, no. 11, pp. 4710–4718, Nov. 2008.

[BHB11] G. Bansal, M. J. Hossain, and V. K. Bhargava, "Adaptive power loading for OFDM-based cognitive radio systems with statistical interference constraint," *IEEE Transactions on Wireless Communications*, vol. 10, no. 9, pp. 2786–2791, Sep. 2011.

[BHME12] Z. Ben-Haim, T. Michaeli, and Y. C. Eldar, "Performance bounds and design criteria for estimating finite rate of innovation signals," *IEEE Transactions on Information Theory*, vol. 58, no. 8, pp. 4993–5015, Aug. 2012.

[BKS06] O. Besson, S. Kraut, and L. L. Scharf, "Detection of an unknown rank- one component in white noise," *IEEE Transactions on Signal Processing*, vol. 54, no. 7, pp. 2835–2839, Jul. 2006.

[BKT09] B. Babadi, N. Kalouptsidis, and V. Tarokh, "Asymptotic achievability of the cramér-rao bound for noisy compressive sampling," *IEEE Transactions on Signal Processing*, vol. 57, no. 3, pp. 1233–1236, Mar. 2009.

[BKT10] ——, "SPARLS: The sparse RLS algorithm," *IEEE Transactions on Signal Processing*, vol. 58, no. 8, pp. 4013–4025, Aug. 2010.

[BNS11] S. Bokharaiee, H. H. Nguyen, and E. Shwedyk, "Blind spectrum sensing for OFDM-based cognitive radio systems," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 3, pp. 858–871, Mar. 2011.

[BPS98] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: information-theoretic and communications aspects," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2619–2692, Oct. 1998.

[Bri81] D. R. Brillinger, *Time Series: Data Analysis and Theory*. San Francisco: Holden-Day, 1981, (Section 4.4: asymptotic distribution of the finite Fourier transform).

[CA07] W. Choi and J. Andrews, "Spatial multiplexing in cellular MIMO-CDMA systems with linear receivers: Outage probability and capacity," *IEEE Transactions on Wireless Communications*, vol. 6, no. 7, pp. 2612–2621, Jul. 2007.

[CA09] V. Chandrasekhar and J. Andrews, "Spectrum allocation in tiered cellular networks," *IEEE Transactions on Communications*, vol. 57, no. 10, pp. 3059 – 3068, Oct. 2009.

[Cap69] J. Capon, "High-resolution frequency-wavenumber spectrum analysis," *Proceedings of the IEEE*, vol. 57, no. 8, pp. 1408–1418, Aug. 1969.

[CCBSS05] C. Cordeiro, K. Challapali, D. Birru, and N. Sai Shankar, "IEEE 802.22: the first worldwide wireless standard based on cognitive radios," in *IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN)*, Nov. 2005.

[CEG13] Y. Chen, Y. Eldar, and A. Goldsmith, "Shannon meets Nyquist: Capacity of sampled Gaussian channels," *IEEE Transactions on Information Theory*, vol. 59, no. 8, pp. 4889–4914, 2013.

[CHJ10] R. Calderbank, S. Howard, and S. Jafarpour, "Construction of a large class of deterministic sensing matrices that satisfy a statistical isometry property," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 358–374, Apr. 2010.

[CJ08] V. Cadambe and S. Jafar, "Interference alignment and degrees of freedom of the K-user interference channel," *IEEE Transactions on Information Theory*, vol. 54, no. 8, pp. 3425–3441, Aug. 2008.

[CJ09] ——, "Interference alignment and the degrees of freedom of wireless networks," *IEEE Transactions on Information Theory*, vol. 55, no. 9, pp. 3893–3908, Sep. 2009.

[CKCD13] L. S. Cardoso, M. Kobayashi, F. R. P. Cavalcanti, and M. Debbah, "Vandermonde-subspace frequency division multiplexing for two-tiered cognitive radio networks," *IEEE Transactions on Communications*, vol. 61, no. 6, pp. 2212–2220, Jun. 2013.

[CLKP12] S. Chaudhari, J. Lunden, V. Koivunen, and H. V. Poor, "Cooperative sensing with imperfect reporting channels: Hard decisions or soft decisions?" *IEEE Transactions on Signal Processing*, vol. 60, no. 1, pp. 18–28, Jan. 2012.

[CMB04] D. Cabric, S. Mishra, and R. Brodersen, "Implementation issues in spectrum sensing for cognitive radios," in *Asilomar Conference on Signals, Systems and Computers*, Nov. 2004.

[Cov72] T. M. Cover, "Broadcast channels," *IEEE Transactions on Information Theory*, vol. 18, no. 1, pp. 2–14, Jan. 1972.

[CRT06] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.

[CSVH12] G. Chung, S. Sridharan, S. Vishwanath, and C. S. Hwang, "On the capacity of overlay cognitive radios with partial cognition," *IEEE Transactions on Information Theory*, vol. 58, no. 5, pp. 2935–2949, May 2012.

[CT91] T. Cover and J. A. Thomas, *Elements of Information Theory*. John Wiley and Sons, Inc., 1991.

[CT05] E. J. Candès and T. Tao, "Decoding with linear programming," *IEEE Transactions on Information Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.

[CT06] ——, "Near-optimal signal recovery from random projections: Universal encoding strategies?" *IEEE Transactions on Information Theory*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006.

[CT12] Y. Chen and Z. Tang, "Effect of spectrum sensing errors on the performance of OFDM-based cognitive radio transmission," *IEEE Transactions on Wireless Communications*, vol. 11, no. 6, pp. 2342–2350, Jun. 2012.

[CTB99] G. Caire, G. Taricco, and E. Biglieri, "Optimum power control over fading channels," *IEEE Transactions on Information Theory*, vol. 45, no. 5, pp. 1468–1489, Jul. 1999.

[CW08] E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Processing Magazine*, pp. 21–30, Mar. 2008.

[CWC11] C.-H. Chen, C.-L. Wang, and C.-T. Chen, "A resource allocation scheme for cooperative multiuser OFDM-based cognitive radio systems," *IEEE Transactions on Communications*, vol. 59, no. 11, pp. 3204–3215, Nov. 2011.

[CWZ11] Y. Chen, C. Wang, and B. Zhao, "Performance comparison of feature-based detectors for spectrum sensing in the presence of primary user traffic," *IEEE Signal Processing Letters*, vol. 18, no. 5, pp. 291–294, May 2011.

[CZL09] X. Chen, L. Zhao, and J. Li, "A modified spectrum sensing method for wideband cognitive radio based on compressive sensing," in *International Conference on Communications and Networking (ChinaCOM)*, Aug. 2009.

[DBWB10] M. Davenport, P. Boufounos, M. Wakin, and R. Baraniuk, "Signal processing with compressive measurements," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 445–460, Apr. 2010.

[DE11] M. Duarte and Y. Eldar, "Structured compressed sensing: From theory to applications," *IEEE Transactions on Signal Processing*, vol. 59, no. 9, pp. 4053–4085, Sep. 2011.

[DG94] A. V. Dandawate and G. B. Giannakis, "Statistical tests for presence of cyclostationarity," *IEEE Transactions on Signal Processing*, vol. 42, no. 9, pp. 2355–2369, Sep. 1994.

[DJSS10] E. Diederichs, A. Juditsky, V. Spokoiny, and C. Schutte, "Sparse non-Gaussian component analysis," *IEEE Transactions on Information Theory*, vol. 56, no. 6, pp. 3033–3047, Jun. 2010.

[DLTB12] M. Davenport, J. Laska, J. Treichler, and R. Baraniuk, "The pros and cons of compressive sensing for wideband signal acquisition: Noise folding versus dynamic range," *IEEE Transactions on Signal Processing*, vol. 60, no. 9, pp. 4628–4642, Sep. 2012.

[DT10] N. Dobigeon and J.-Y. Tourneret, "Bayesian orthogonal component analysis for sparse representation," *IEEE Transactions on Signal Processing*, vol. 58, no. 5, pp. 2675–2685, May 2010.

[EAQS12] S. Ekin, M. Abdallah, K. Qaraqe, and E. Serpedin, "Random subcarrier allocation in OFDM-based cognitive radio networks," *IEEE Transactions on Signal Processing*, vol. 60, no. 9, pp. 4758–4774, Sep. 2012.

[EGR11] K. Eswaran, M. Gastpar, and K. Ramchandran, "Cognitive radio through primary control feedback," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 2, pp. 384–393, Feb. 2011.

[EKB10] Y. Eldar, P. Kuppinger, and H. Bolcskei, "Block-sparse signals: Uncertainty relations and efficient recovery," *IEEE Transactions on Signal Processing*, vol. 58, no. 6, pp. 3042–3054, Jun. 2010.

[ETS04] ETSI, "Digital Video Broadcasting (DVB); framing structure, channel coding and modulation for digital terrestrial television (DVB-T)," European Telecommunications Standards Institute (ETSI), Tech. Rep. EN 300 744 V1.5.1, Jun. 2004.

[FCC03] FCC, "Notice of proposed rule making and order," Federal Communications Commission (FCC), ET Docket No. 03-222, Dec. 2003.

[FHMI11] S. Filin, H. Harada, H. Murakami, and K. Ishizu, "International standardization of cognitive radio systems," *IEEE Communications Magazine*, vol. 49, no. 3, pp. 82–89, Mar. 2011.

[FNW07] M. Figueiredo, R. Nowak, and S. Wright, "Gradient projection for sparse reconstruction: application to compressed sensing and other inverse problems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 4, pp. 586–597, Jul. 2007.

[FSRVV13] J. Font-Segura, J. Riba, J. Villares, and G. Vázquez, "Quadratic sphericity test for blind detection over time-varying frequency-selective channels," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2013.

[FSVR11] J. Font-Segura, G. Vázquez, and J. Riba, "Multi-frequency GLRT spectrum sensing for wideband cognitive radio," in *IEEE International Conference on Communications (ICC)*, Jun. 2011.

[FSW10] J. Font-Segura and X. Wang, "GLRT-based spectrum sensing for cognitive radio with prior information," *IEEE Transactions on Communications*, vol. 58, no. 7, pp. 2137–2146, Jul. 2010.

[GA11] E. Gismalla and E. Alsusa, "Performance analysis of the periodogram-based energy-detector in fading channels," *IEEE Transactions on Signal Processing*, vol. 59, no. 8, pp. 3712–3721, Aug. 2011.

[GA12] ——, "On the performance of energy detection using Bartlett's estimate for spectrum sensing in cognitive radio systems," *IEEE Transactions on Signal Processing*, vol. 60, no. 7, pp. 3394–3404, Jul. 2012.

[Gal68] R. G. Gallager, *Information Theory and Reliable Communications*. New York: Wiley, 1968.

[Gar86] W. A. Gardner, *Statistical Spectral Analysis: A Nonprobabilistic Theory*. Prentice-Hall, 1986.

[Gar90] ——, *Introduction to random processes, with applications to signals and systems*, 2nd ed. New York: McGraw-Hill, 1990.

[Gas07] M. Gastpar, "On capacity under receive and spatial spectrum-sharing constraints," *IEEE Transactions on Information Theory*, vol. 53, no. 2, pp. 471–487, Feb. 2007.

[GE11] S. Gleichman and Y. C. Eldar, "Blind compressed sensing," *IEEE Transactions on Information Theory*, vol. 57, no. 10, pp. 6958–6975, Oct. 2011.

[GJMS09] A. Goldsmith, S. A. Jafar, I. Maric, and S. Srinivasa, "Breaking spectrum gridlock with cognitive radios: An information theoretic perspective," *Proceedings of the IEEE*, vol. 97, no. 5, pp. 894–914, May 2009.

[GM69] F. A. Graybill and G. Milliken, "Quadratic forms and idempotent matrices with random elements," *The Annals of Mathematical Statistics*, vol. 40, no. 4, pp. 1430–1438, 1969.

[GNP06] W. A. Gardner, A. Napolitano, and L. Paura, "Cyclostationarity: Half a century of research," *Signal Processing*, vol. 86, no. 4, pp. 639–697, Apr. 2006.

[Gra76] F. A. Graybill, *Theory and Applications of the Linear Model*. Cary, NC: SAS Institute, Inc., 1976.

[Gra06] R. M. Gray, *Toeplitz and Circulant Matrices: A Review*. New York: Hanover/Now, 2006.

[GS93] W. A. Gardner and C. M. Spooner, "Detection and source location of weak cyclostationary signals: simplifications of the maximum-likelihood receiver," *IEEE Transactions on Communications*, vol. 41, no. 6, pp. 905–916, Jun. 1993.

[GS08a] A. Ghasemi and E. S. Sousa, "Spectrum sensing in cognitive radio networks: requirements, challenges and design trade-offs," *IEEE Communications Magazine*, vol. 46, no. 4, pp. 32–39, Apr. 2008.

[GS08b] E. Gudmundson and P. Stoica, "On denoising via penalized least-squares rules," in *IEEE Int. Conf. Acoustics, Speech and Signal Process. (ICASSP)*, Apr. 2008.

[GSTS09] S. Geirhofer, J. Z. Sun, L. Tong, and B. M. Sadler, "Cognitive frequency hopping based on interference prediction: theory and experimental results," *Mobile Computing and Communications Review*, vol. 13, no. 2, pp. 49–61, Mar. 2009.

[GTE11] K. Gedalyahu, R. Tur, and Y. C. Eldar, "Multichannel sampling of pulse streams at the rate of innovation," *IEEE Transactions on Signal Processing*, vol. 59, no. 4, pp. 1491–1504, Apr. 2011.

[GTS07] S. Geirhofer, L. Tong, and B. M. Sadler, "Measurement-based models for cognitive medium access in WLAN bands," Cornell University, Adaptive Communications and Signal Processing Group (ACSP), Technical Report ACSP-TR-02-07-02, Feb. 2007, http://acsp.ece.cornell.edu/papers/ACSP-TR-02-07-02.pdf.

[GTS08] S. Geirhofer, L. Tong, and B. Sadler, "Cognitive medium access: Constraining interference based on experimental models," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 1, pp. 95–105, Jan. 2008.

[HAW08] A. Hunter, J. Andrews, and S. Weber, "The transmission capacity of ad hoc networks with spatial diversity," *IEEE Transactions on Wireless Communications*, vol. 7, no. 12, pp. 5058–5071, Dec. 2008.

[Hay05] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE Transactions on Communications*, vol. 23, no. 2, pp. 201–220, Feb. 2005.

[HBHB09] Z. Hasan, G. Bansal, E. Hossain, and V. Bhargava, "Energy-efficient power allocation in OFDM-based cognitive radio systems: A risk-return model," *IEEE Transactions on Wireless Communications*, vol. 8, no. 12, pp. 6078–6088, Dec. 2009.

[HC11] K. Hossain and B. Champagne, "Wideband spectrum sensing for cognitive radios with correlated subband occupancy," *IEEE Signal Processing Letters*, vol. 18, no. 1, pp. 35–38, Jan. 2011.

[HLC10] C.-H. Hwang, G.-L. Lai, and S.-C. Chen, "Spectrum sensing in wideband OFDM cognitive radios," *IEEE Transactions on Signal Processing*, vol. 58, no. 2, pp. 709–719, Feb. 2010.

[HLD11] S. Huang, X. Liu, and Z. Ding, "Decentralized cognitive radio control based on inference from primary link control information," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 2, pp. 394–406, Feb. 2011.

[HRS08] M. Hernandez, M. Rojas, and P. Stoica, "New spectral estimation based on filterbank for spectrum sensing," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Apr. 2008.

[HSTM11] M. Hanif, P. Smith, D. Taylor, and P. Martin, "MIMO cognitive radios with antenna selection," *IEEE Transactions on Wireless Communications*, vol. 10, no. 11, pp. 3688–3699, Nov. 2011.

[HT13] G. Huang and J. K. Tugnait, "On cyclostationarity based spectrum sensing under uncertain Gaussian noise," *IEEE Transactions on Signal Processing*, vol. 61, no. 8, pp. 2042–2054, Apr. 2013.

[HTR09] S. Haykin, D. J. Thomson, and J. H. Reed, "Spectrum sensing for cognitive radio," *Proceedings of the IEEE*, vol. 97, no. 5, pp. 849–877, May 2009.

[HWA+07] W. Hu, D. Willkomm, M. Abusubaih, J. Gross, G. Vlantis, M. Gerla, and A. Wolisz, "Dynamic frequency hopping communities for efficient IEEE 802.22 operation," *IEEE Communications Magazine*, vol. 45, no. 5, pp. 80–87, May 2007.

[Höl89] O. Hölder, "Über einen Mittelwertsatz," *Göttingen Nachrichten*, vol. 7, pp. 38–47, 1889.

[JAW08] N. Jindal, J. G. Andrews, and S. Weber, "Bandwidth partitioning in decentralized wireless networks," *IEEE Transactions on Wireless Communications*, vol. 7, no. 12, pp. 5408–5419, Dec. 2008.

[JBHHE11] A. Jung, Z. Ben-Haim, F. Hlawatsch, and Y. C. Eldar, "Unbiased estimation of a sparse vector in white Gaussian noise," *IEEE Transactions on Information Theory*, vol. 57, no. 12, pp. 7856–7876, Dec. 2011.

[JJL+07] S. D. Jones, E. Jung, X. Liu, N. Merheb, and I.-J. Wang, "Characterization of spectrum activities in the U. S. public safety band for opportunistic spectrum access," in *IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN)*, Apr. 2007.

[JKR11] Y. Jin, Y.-H. Kim, and B. D. Rao, "Limits on support recovery of sparse signals via multiple-access communication techniques," *IEEE Transactions on Information Theory*, vol. 57, no. 12, pp. 7877–7892, Dec. 2011.

[JS07] S. Jafar and S. Srinivasa, "Capacity limits of cognitive radio with distributed and dynamic spectral activity," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 3, pp. 529–537, Apr. 2007.

[JS08] S. Jafar and S. S. (Shitz), "Degrees of freedom region of the MIMO X channel," *IEEE Transactions on Information Theory*, vol. 54, no. 1, pp. 151–170, Jan. 2008.

[JVLN+11] P. Jia, M. Vu, T. Le-Ngoc, S.-C. Hong, and V. Tarokh, "Capacity- and Bayesian-based cognitive sensing with location side information," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 2, pp. 276–289, Feb. 2011.

[JWA00] N. Jindal, S. Weber, and J. G. Andrews, "Fractional power control for decentralized wireless networks," *IEEE Transactions on Wireless Communications*, vol. 7, no. 12, pp. 5482–5492, Dec. 200.

[Kay98a] S. M. Kay, *Fundamentals of Statistical Signal Processing*. Upper Saddle River, NJ: Prentice Hall, 1998, vol. 2 (Detection Theory).

[Kay98b] ——, *Fundamentals of Statistical Signal Processing*. Upper Saddle River, NJ: Prentice Hall, 1998, vol. 1 (Estimation Theory).

[Kay12] S. Kay, "A new proof of the Neyman-Pearson theorem using the EEF and the vindication of Sir R. Fisher," *IEEE Signal Processing Letters*, vol. 19, no. 8, pp. 451–454, Aug. 2012.

[KBB12] P. Kaligineedi, G. Bansal, and V. K. Bhargava, "Power loading algorithms for OFDM-based cognitive radio systems with imperfect sensing," *IEEE Transactions on Wireless Communications*, vol. 11, no. 12, pp. 4225–4230, Dec. 2012.

[KG11] V. Kekatos and G. B. Giannakis, "From sparse signals to sparse residuals for robust sensing," *IEEE Transactions on Signal Processing*, vol. 59, no. 7, pp. 3355–3368, Jul. 2011.

[KGLZ10] X. Kang, H. Garg, Y.-C. Liang, and R. Zhang, "Optimal power allocation for OFDM-based cognitive radio with new primary transmission protection criteria," *IEEE Transactions on Wireless Communications*, vol. 9, no. 6, pp. 2066–2075, Jun. 2010.

[KKL+07] S. Kim, K. Koh, M. Lustig, S. Boyd, and D. Gorinevsky, "An interior-point method for large-scale ℓ_1-regularized least squares," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 4, pp. 606–617, Dec. 2007.

[KLN+09] X. Kang, Y.-C. Liang, A. Nallanathan, H. Garg, and R. Zhang, "Optimal power allocation for fading channels in cognitive radio networks: Ergodic capacity and outage capacity," *IEEE Transactions on Wireless Communications*, vol. 8, no. 2, pp. 940–950, Feb. 2009.

[KM82] J. Kormylo and J. Mendel, "Maximum likelihood detection and estimation of Bernoulli-Gaussian processes," *IEEE Transactions on Information Theory*, vol. 28, no. 3, pp. 482–488, May 1982.

[Kod05] P. J. Kodolzy, "Cognitive radio fundamentals," in *Software Defined Radios (SDR) Forum*, Apr. 2005.

[KS78] L. Kleinrock and J. A. Silvester, "Optimum transmission radii in packet radio networks or why six is a magic number," in *National Telecommunications Conference*, Dec. 1978.

[KS95] G. Kaplan and S. Shamai, "Error probabilities for the block-fading gaussian channel," *Arch. Electronik Ubertragung*, vol. 49, no. 4, pp. 192–205, 1995.

[KXAH11] M. A. Khajehnejad, W. Xu, A. S. Avestimehr, and B. Hassibi, "Analyzing weighted ℓ_1 minimization for sparse recovery with nonuniform sparse models," *IEEE Transactions on Signal Processing*, vol. 59, no. 5, pp. 1985–2001, May 2011.

[LA11] G. Leus and D. D. Ariananda, "Power spectrum blind sampling," *IEEE Signal Processing Letters*, vol. 18, no. 8, pp. 443–446, Aug. 2011.

[LAH13] J. Lee, J. G. Andrews, and D. Hong, "Spectrum-sharing transmission capacity with interference cancellation," *IEEE Transactions on Communications*, vol. 61, no. 1, pp. 76–86, Jan. 2013.

[Lan67] H. J. Landau, "Necessary density conditions for sampling and interpolation of certain entire functions," *Acta Mathematica*, vol. 117, no. 1, pp. 37–52, 1967.

[LCHK11] K. Lee, C.-B. Chae, R. Heath, and J. Kang, "MIMO transceiver designs for spatial sensing in cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 10, no. 11, pp. 3570–3576, Nov. 2011.

[LCL+11a] Y.-C. Liang, K.-C. Chen, G. Y. Li, P. Mähönen, and D. Niyato, "Guest editorial advances in cognitive radio networking and communications (I)," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 2, pp. 273–275, Feb. 2011.

[LCL+11b] ——, "Guest editorial advances in cognitive radio networking and communications (II)," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 4, pp. 673–675, Apr. 2011.

[LDT12] M. A. Lexa, M. E. Davies, and J. S. Thompson, "Reconciling compressive sampling systems for spectrally sparse continuous-time signals," *IEEE Transactions on Signal Processing*, vol. 60, no. 1, pp. 155–171, Jan. 2012.

[LdV01] A. Leshem and A.-J. V. der Veen, "Multichannel detection of Gaussian signals with uncalibrated receivers," *IEEE Signal Processing Letters*, vol. 8, no. 4, pp. 120–122, Apr. 2001.

[LG01] L. Li and A. Goldsmith, "Capacity and optimal resource allocation for fading broadcast channels. II. Outage capacity," *IEEE Transactions on Information Theory*, vol. 47, no. 3, pp. 1103–1127, Mar. 2001.

[LG05] ——, "Outage capacities and optimal power allocation for fading multiple-access channels," *IEEE Transactions on Information Theory*, vol. 51, no. 4, pp. 1326–1347, Apr. 2005.

[LKHP09] J. Lunden, V. Koivunen, A. Huttunen, and H. V. Poor, "Collaborative cyclostationary spectrum sensing for cognitive radio systems," *IEEE Transactions on Signal Processing*, vol. 57, no. 11, pp. 4182–4195, Nov. 2009.

[LKK10] J. Lunden, S. Kassam, and V. Koivunen, "Robust nonparametric cyclic correlation-based spectrum sensing for cognitive radio," *IEEE Transactions on Signal Processing*, vol. 58, no. 1, pp. 38–52, Jan. 2010.

[LLL12] L. Lu, G. Y. Li, and S. Li, "Optimum periodic spectrum sensing for CR networks," *IEEE Communications Letters*, vol. 16, no. 12, pp. 1–4, Dec. 2012.

[LMGSS13] N. Liu, I. Maric, A. J. Goldsmith, and S. Shamai (Shitz), "Capacity bounds and exact results for the cognitive Z-interference channel," *IEEE Transactions on Information Theory*, vol. 59, no. 2, pp. 886–893, Feb. 2013.

[LN12] Y. Li and A. Nosratinia, "Capacity limits of multiuser multiantenna cognitive networks," *IEEE Transactions on Information Theory*, vol. 58, no. 7, pp. 4493–4508, Jul. 2012.

[LR12] L. Luo and S. Roy, "Efficient spectrum sensing for cognitive radio networks via joint optimization of sensing threshold and duration," *IEEE Transactions on Communications*, vol. 60, no. 10, pp. 2851–2860, Oct. 2012.

[LS12] Y. Lee and Y. Sung, "Generalized Chernoff information for mismatched Bayesian detection and its application to energy detection," *IEEE Signal Processing Letters*, vol. 19, no. 11, pp. 753–756, Nov. 2012.

[LSI12] H.-J. Lim, D.-Y. Seol, and G.-H. Im, "Joint sensing adaptation and resource allocation for cognitive radio with imperfect sensing," *IEEE Transactions on Communications*, vol. 60, no. 4, pp. 1091–1100, Apr. 2012.

[LWYB11] J. N. Laska, Z. Wen, W. Yin, and R. G. Baraniuk, "Trust, but verify: Fast and accurate signal recovery from 1-bit compressive measurements," *IEEE Transactions on Signal Processing*, vol. 59, no. 11, pp. 5289–5301, Nov. 2011.

[LZLZ08] T. J. Lim, R. Zhang, Y. C. Liang, and Y. Zeng, "GLRT-based spectrum sensing for cognitive radio," in *IEEE Global Communications Conference (GLOBECOM)*, Dec. 2008.

[LZY11] J. C. F. Li, W. Zhang, and J. Yuan, "Opportunistic spectrum sharing in cognitive radio networks based on primary limited feedback," *IEEE Transactions on Communications*, vol. 59, no. 12, pp. 3272–3277, Dec. 2011.

[MA09] L. Musavian and S. Aissa, "Capacity and power allocation for spectrum-sharing communications in fading channels," *IEEE Transactions on Wireless Communications*, vol. 8, no. 1, pp. 148–156, Jan. 2009.

[MAA12] M. Mirmohseni, B. Akhbari, and M. R. Aref, "On the capacity of interference channel with causal and noncausal generalized feedback at the cognitive transmitter," *IEEE Transactions on Information Theory*, vol. 58, no. 5, pp. 2813–2837, May 2012.

[MAMK08] M. Maddah-Ali, A. Motahari, and A. Khandani, "Communication over MIMO X channels: Interference alignment, decomposition, and performance analysis," *IEEE Transactions on Information Theory*, vol. 54, no. 8, pp. 3457–3470, Aug. 2008.

[MBR05] R. Menon, R. M. Buehrer, and J. H. Reed, "Outage probability based comparison of underlay and overlay spectrum sharing techniques," in *IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN)*, Nov. 2005.

[MBZJ09] H. Mohimani, M. Babaie-Zadeh, and C. Jutten, "A fast approach for overcomplete sparse decomposition based on smoothed ℓ_0-norm," *IEEE Transactions on Signal Processing*, vol. 57, no. 1, pp. 289–301, Jan. 2009.

[ME09] M. Mishali and Y. C. Eldar, "Blind multiband signal reconstruction: Compressed sensing for analog signals," *IEEE Transactions on Signal Processing*, vol. 57, no. 3, pp. 993–1009, Mar. 2009.

[ME10] M. Mishali and Y. Eldar, "From theory to practice: Sub-Nyquist sampling of sparse wideband analog signals," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 375–391, Apr. 2010.

[MEDS11] M. Mishali, Y. Eldar, O. Dounaevsky, and E. Shoshan, "Xampling: Analog to digital at sub-Nyquist rates," *IET Circuits, Devices Systems*, vol. 5, no. 1, pp. 8–20, Jan. 2011.

[MEE11] M. Mishali, Y. Eldar, and A. Elron, "Xampling: Signal acquisition and processing in union of subspaces," *IEEE Transactions on Signal Processing*, vol. 59, no. 10, pp. 4719–4734, Oct. 2011.

[MGC11] A. Mariani, A. Giorgetti, and M. Chiani, "Effects of noise power estimation on energy detection for cognitive radio applications," *IEEE Transactions on Communications*, vol. 59, no. 12, pp. 3410–3420, Dec. 2011.

[MH13] A. Makarfi and K. Hamdi, "Interference analysis of energy detection for spectrum sensing," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 6, pp. 2570–2578, Jul. 2013.

[MJTW12] G. Moustakides, G. Jajamovich, A. Tajer, and X. Wang, "Joint detection and estimation: Optimum tests and applications," *IEEE Transactions on Information Theory*, vol. 58, no. 7, pp. 4215–4229, Jul. 2012.

[MLJ09] J. Ma, G. Li, and B. H. Juang, "Signal processing in cognitive radio," *Proceedings of the IEEE*, vol. 97, no. 5, pp. 805–823, May 2009.

[MM99] J. Mitola and G. Q. Maguire, "Cognitive radio: making software radios more personal," *IEEE Personal Communications Magazine*, vol. 6, no. 4, pp. 13–18, Aug. 1999.

[MM10a] G. Meinsma and L. Mirkin, "Sampling from a system-theoretic viewpoint: Part I—concepts and tools," *IEEE Transactions on Signal Processing*, vol. 58, no. 7, pp. 3578–3590, Jul. 2010.

[MM10b] ——, "Sampling from a system-theoretic viewpoint: Part II—noncausal solutions," *IEEE Transactions on Signal Processing*, vol. 58, no. 7, pp. 3591–3606, Jul. 2010.

[MMA11] M. Mohammadkarimi, B. Mahboobi, and M. Ardebilipour, "Optimal spectrum sensing in fast fading Rayleigh channel for cognitive radio," *IEEE Communications Letters*, vol. 15, no. 10, pp. 1032–1034, Oct. 2011.

[MMS12] F. Moghimi, R. K. Mallik, and R. Schober, "Sensing time and power optimization in MIMO cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 11, no. 9, pp. 3398–3408, Sep. 2012.

[Mou09] G. V. Moustakides, "Finite Sample Size Optimality of GLR Tests," *ArXiv e-prints*, Mar. 2009, http://arxiv.org/abs/0903.3795.

[MSW10] D. Malioutov, S. Sanghavi, and A. Willsky, "Sequential compressed sensing," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 435–444, Apr. 2010.

[MT08] S. Misra and L. Tong, "Error exponents for the detection of Gauss-Markov signals using randomly spaced sensors," *IEEE Transactions on Signal Processing*, vol. 56, no. 8, pp. 3385–3396, Aug. 2008.

[MW78] S. Musa and W. Wasylkiwskyj, "Co-channel interference of spread spectrum systems in a multiple user environment," *IEEE Transactions on Communications*, vol. 26, no. 10, pp. 1405–1413, Oct. 1978.

[MYL+11] J. J. Meng, W. Yin, H. Li, E. Hossain, and Z. Han, "Collaborative spectrum sensing from sparse observations in cognitive radio networks," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 2, pp. 327–337, Feb. 2011.

[NIK12] M. Nekovee, T. Irnich, and J. Karlsson, "Worldwide trends in regulation of secondary access to white spaces using cognitive radio," *IEEE Wireless Communications Magazine*, vol. 19, no. 4, pp. 32–40, Aug. 2012.

[NPI10] M. Naraghi-Pour and T. Ikuma, "Autocorrelation-based spectrum sensing for cognitive radios," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 2, pp. 718–733, Feb. 2010.

[OSNP11] M. Orooji, R. Soosahabi, and M. Naraghi-Pour, "Blind spectrum sensing using antenna arrays and path correlation," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 8, pp. 3758–3767, Oct. 2011.

[OSW94] L. Ozarow, S. Shamai, and A. D. Wyner, "Information theoretic considerations for cellular mobile radio," *IEEE Transactions on Vehicular Technology*, vol. 43, no. 2, pp. 359–378, May 1994.

[OSW12] B. Osgood, A. Siripuram, and W. Wu, "Discrete sampling and interpolation: Universal sampling sets for discrete bandlimited spaces," *IEEE Transactions on Information Theory*, vol. 58, no. 7, pp. 4176–4200, Jul. 2012.

[Pak87] L. Pakula, "Asymptotic zero distribution of orthogonal polynomials in sinusoidal frequency estimation," *IEEE Transactions on Information Theory*, vol. 33, no. 4, pp. 569–576, Jul. 1987.

[PBD11a] A. Punchihewa, V. Bhargava, and C. Despins, "Blind estimation of OFDM parameters in cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 10, no. 3, pp. 733–738, Mar. 2011.

[PBD11b] ——, "Linear precoding for orthogonal space-time block coded MIMO-OFDM cognitive radio," *IEEE Transactions on Communications*, vol. 59, no. 3, pp. 767–779, Mar. 2011.

[PHB12] P. Paysarvi-Hoseini and N. C. Beaulieu, "On the benefits of multichannel/wideband spectrum sensing with non-uniform channel sensing durations for cognitive radio networks," *IEEE Transactions on Communications*, vol. 60, no. 9, pp. 2434–2443, Sep. 2012.

[Pic95] B. Picinbono, "On deflection as a performance criterion in detection," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 31, no. 3, pp. 1072–1081, Jul. 1995.

[PLGZ11] E. C. Y. Peh, Y.-C. Liang, Y. L. Guan, and Y. Zeng, "Power control in cognitive radios under cooperative and non-cooperative spectrum sensing," *IEEE Transactions on Wireless Communications*, vol. 10, no. 12, pp. 4238–4248, Dec. 2011.

[PND+11] P. Pawelczak, K. Nolan, L. Doyle, S. W. Oh, and D. Cabric, "Cognitive radio: Ten years of experimentation and development," *IEEE Communications Magazine*, vol. 49, no. 3, pp. 90–100, Mar. 2011.

[PNLRS09] A. I. Pérez-Neira, M. A. Lagunas, M. A. Rojas, and P. Stoica, "Correlation matching approach for spectrum sensing in open spectrum communications," *IEEE Transactions on Signal Processing*, vol. 57, no. 12, pp. 4823–4836, Dec. 2009.

[Poo94] H. V. Poor, *An Introduction to Signal Detection and Estimation*, 2nd ed. New York: Springer-Verlag, 1994.

[Por08] B. Porat, *Digital Processing of Random Signals. Theory and Methods*. Dover Publications, 2008.

[Pro00] J. G. Proakis, *Digital Communications*, 4th ed., ser. Science/Engineering/Math. McGraw-Hill, 2000.

[PSDC12] F. Penna, Y. Sun, L. Dolecek, and D. Cabric, "Detecting and counteracting statistical attacks in cooperative spectrum sensing," *IEEE Transactions on Signal Processing*, vol. 60, no. 4, pp. 1806–1822, Apr. 2012.

[Pur87] M. B. Pursley, "The role of spread spectrum in packet radio networks," *Proceedings of the IEEE*, vol. 75, no. 1, pp. 116–134, Jan. 1987.

[PWPL09] Y. L. Polo, Y. Wang, A. Pandharipande, and G. Leus, "Compressive wide-band spectrum sensing," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Apr. 2009.

[QZSS11] Z. Quan, W. Zhang, S. J. Shellhammer, and A. H. Sayed, "Optimal spectral feature detection for spectrum sensing at very low-SNR," *IEEE Transactions on Communications*, vol. 59, no. 1, pp. 201–212, Jan. 2011.

[RACV11] A. Rawat, P. Anand, H. Chen, and P. Varshney, "Collaborative spectrum sensing in the presence of Byzantine attacks in cognitive radio networks," *IEEE Transactions on Signal Processing*, vol. 59, no. 2, pp. 774–786, Feb. 2011.

[RFSVV14] J. Riba, J. Font-Segura, J. Villares, and G. Vázquez, "Frequency-domain GLR detection of a second-order cyclostationary signal over fading channels," *IEEE Transactions on Signal Processing*, vol. 62, no. 8, pp. 1899–1912, Apr. 2014.

[RG10] G. Reeves and M. Gastpar, "'Compressed' compressed sensing," Jan. 2010, http://arxiv.org/abs/1001.4295.

[RSV01] J. Riba, J. Sala, and G. Vázquez, "Conditional maximum likelihood timing recovery: estimators and bounds," *IEEE Transactions on Signal Processing*, vol. 49, no. 4, pp. 835–850, Apr. 2001.

[RTD12] S. Rini, D. Tuninetti, and N. Devroye, "Inner and outer bounds for the Gaussian cognitive interference channel and new capacity results," *IEEE Transactions on Information Theory*, vol. 58, no. 2, pp. 820–848, Feb. 2012.

[RVS10] D. Ramírez, J. Vía, and I. Santamaría, "Multiantenna spectrum sensing: The case of wideband rank-one primary signals," in *IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*, Oct. 2010.

[RVSS13] D. Ramírez, J. Vía, I. Santamaría, and L. Scharf, "Locally most powerful invariant tests for correlation and sphericity of gaussian vectors," *IEEE Transactions on Information Theory*, vol. 59, no. 4, pp. 2128–2141, 2013.

[RVV10] J. Riba, J. Villares, and G. Vázquez, "A nondata-aided SNR estimation technique for multilevel modulations exploiting signal cyclostationarity," *IEEE Transactions on Signal Processing*, vol. 58, no. 11, pp. 5767–5778, Nov. 2010.

[RVVLV+11] D. Ramírez, G. Vázquez-Vilar, R. López-Valcarce, J. Vía, and I. Santamaría, "Detection of rank-P signals in cognitive radio networks with uncalibrated multiple antennas," *IEEE Transactions on Signal Processing*, vol. 59, no. 8, pp. 3764–3774, Aug. 2011.

[SAVVLV12] J. Sala-Alvarez, G. Vázquez-Vilar, and R. López-Valcarce, "Multiantenna GLR detection of rank-one signals with known power spectrum in white noise with unknown spatial correlation," *IEEE Transactions on Signal Processing*, vol. 60, no. 6, pp. 3065–3078, Jun. 2012.

[SB12a] M. Shaat and F. Bader, "Asymptotically optimal resource allocation in OFDM-based cognitive networks with multiple relays," *Wireless Communications, IEEE Transactions on*, vol. 11, no. 3, pp. 892–897, Mar. 2012.

[SB12b] P. Stoica and P. Babu, "Sparse estimation of spectral lines: Grid selection problems and their solutions," *IEEE Transactions on Signal Processing*, vol. 60, no. 2, pp. 962–967, Feb. 2012.

[SBL11] P. Stoica, P. Babu, and J. Li, "New method of sparse parameter estimation in separable models and its use for spectral analysis of irregularly sampled data," *IEEE Transactions on Signal Processing*, vol. 59, no. 1, pp. 35–47, Jan. 2011.

[SBM12] A. Singh, M. R. Bhatnagar, and R. K. Mallik, "Cooperative spectrum sensing in multiple antenna based cognitive radio network using an improved energy-detector," *IEEE Communications Letters*, vol. 16, no. 1, pp. 64–67, Jan. 2012.

[SCKP10] X. Shang, B. Chen, G. Kramer, and H. Poor, "Capacity regions and sum-rate capacities of vector Gaussian interference channels," *IEEE Transactions on Information Theory*, vol. 56, no. 10, pp. 5030–5044, Oct. 2010.

[SCN12] H. Sun, W.-Y. Chiu, and A. Nallanathan, "Adaptive compressive spectrum sensing for wideband cognitive radios," *IEEE Communications Letters*, vol. 16, no. 11, pp. 1812–1815, Nov. 2012.

[SD98] B. M. Sadler and A. V. Dandawate, "Nonparametric estimation of the cyclic cross spectrum," *IEEE Transactions on Information Theory*, vol. 44, no. 1, pp. 351–358, Jan. 1998.

[Sel10] J. Selva, "Regularized sampling of multiband signals," *IEEE Transactions on Signal Processing*, vol. 58, no. 11, pp. 5624–5638, Nov. 2010.

[SGLW03] S. Shahbazpanahi, A. Gershman, Z.-Q. Luo, and K. M. Wong, "Robust adaptive beamforming for general-rank signal models," *IEEE Transactions on Signal Processing*, vol. 51, no. 9, pp. 2257–2269, Sep. 2003.

[Sha59] C. E. Shannon, "Probability of error for optimal codes in a Gaussian channel," *Bell Systems Technical Journal*, vol. 38, pp. 611–656, May 1959.

[SL07] Y. Selen and E. Larsson, "Empirical Bayes linear regression with unknown model order," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Apr. 2007.

[SLH09] P. Stoica, J. Li, and H. He, "Spectral analysis of nonuniformly sampled data: A new approach versus the periodogram," *IEEE Transactions on Signal Processing*, vol. 57, no. 3, pp. 843–858, Mar. 2009.

[SMPI10] D. Sacristan-Murga and A. Pascual-Iserte, "Differential feedback of MIMO channel Gram matrices based on geodesic curves," *IEEE Transactions on Wireless Communications*, vol. 9, no. 12, pp. 3714–3727, Dec. 2010.

[SN11] S. Stotas and A. Nallanathan, "Optimal sensing time and power allocation in multiband cognitive radio networks," *IEEE Transactions on Communications*, vol. 59, no. 1, pp. 226–235, Jan. 2011.

[SND08] P. Sutton, K. Nolan, and L. Doyle, "Cyclostationary signatures in practical cognitive radio applications," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 1, pp. 13–24, Jan. 2008.

[SNWC13] H. Sun, A. Nallanathan, C.-X. Wang, and Y. Chen, "Wideband spectrum sensing for cognitive radio networks: a survey," *IEEE Wireless Communications Magazine*, vol. 20, no. 2, pp. 74–81, 2013.

[SPSG05] E. Serpedin, F. Panduru, I. Sari, and G. B. Giannakis, "Bibliography on cyclostationarity," *Signal Processing*, vol. 85, no. 12, pp. 2233–2303, 2005.

[SS04] P. Stoica and Y. Selen, "Model-order selection: a review of information criterion rules," *IEEE Signal Processing Magazine*, vol. 21, no. 4, pp. 36–47, 2004.

[SS10] P. J. Schreier and L. L. Scharf, *Statistical Signal Processing of Complex-Valued Data. The Theory of Improper and Noncircular Signals*. Cambridge University Press, 2010.

[SSR13] E. Stathakis, M. Skoglund, and L. K. Rasmussen, "On beamforming and orthogonal space-time coding in cognitive networks with partial CSI," *IEEE Transactions on Communications*, vol. 61, no. 3, pp. 961–972, 2013.

[STM13] S. Sedighi, A. Taherpour, and S. Monfared, "Bayesian generalised likelihood ratio test-based multiple antenna spectrum sensing for cognitive radios," *IET Commun.*, vol. 7, no. 18, pp. 2151–2165, 2013.

[SWZZ12] L. Shen, H. Wang, W. Zhang, and Z. Zhao, "Multiple antennas assisted blind spectrum sensing in cognitive radio channels," *IEEE Communications Letters*, vol. 16, no. 1, pp. 92–94, Jan. 2012.

[Tan05] H. Tang, "Some physical layer issues of wide-band cognitive radio systems," in *IEEE International Symposium on New Frontiers on Dynamic Spectrum Access (DySPAN)*, Nov. 2005.

[TCW12] A. Tajer, R. M. Castro, and X. Wang, "Adaptive sensing of congested spectrum bands," *IEEE Transactions on Information Theory*, vol. 58, no. 9, pp. 6110–6125, Sep. 2012.

[TFL11] C. W. Tan, S. Friedland, and S. H. Low, "Spectrum management in multiuser cognitive wireless networks: Optimality and algorithm," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 2, pp. 421–430, Feb. 2011.

[TG07a] Z. Tian and G. B. Giannakis, "Compressed sensing for wideband cognitive radios," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Apr. 2007.

[TG07b] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.

[Tia08] Z. Tian, "Compressed wideband sensing in cooperative cognitive radio networks," in *IEEE Global Communications Conference (GLOBECOM)*, Dec. 2008.

[TIM10] A. Tanaka, H. Imai, and M. Miyakoshi, "Kernel-induced sampling theorem," *IEEE Transactions on Signal Processing*, vol. 58, no. 7, pp. 3569–3577, Jul. 2010.

[TKK82] D. Tufts, R. Kumaresan, and I. Kirsteins, "Data adaptive signal estimation by singular value decomposition of a data matrix," *Proceedings of the IEEE*, vol. 70, no. 6, pp. 684–685, Jun. 1982.

[TLL11] Z. Tian, G. Leus, and V. Lottici, "Joint dynamic resource allocation and waveform adaptation for cognitive networks," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 2, pp. 443–454, Feb. 2011.

[TLWP09] J. Tang, N. Li, Y. Wu, and Y. Peng, "On detection performance of MIMO radar: A relative entropy-based study," *IEEE Signal Processing Letters*, vol. 16, no. 3, pp. 184–187, Mar. 2009.

[TNKG10] A. Taherpour, M. Nasiri-Kenari, and S. Gazor, "Multiple antenna spectrum sensing in cognitive radios," *IEEE Transactions on Wireless Communications*, vol. 9, no. 2, pp. 814–823, Feb. 2010.

[TPD+10] M. Timmers, S. Pollin, A. Dejonghe, L. Van der Perre, and F. Catthoor, "A distributed multichannel MAC protocol for multihop cognitive radio networks," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 1, pp. 446–459, Jan. 2010.

[TS05] R. Tandra and A. Sahai, "Fundamental limits on detection in low SNR under noise uncertainty," in *International Conference on Wireless Networks, Communication and Mobile Computing*, Jun. 2005.

[TS07] ——, "SNR walls for feature detectors," in *IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN)*, Apr. 2007.

[TS08] ——, "SNR walls for signal detection," *IEEE Journal of Selected Topics in Signal Processing*, vol. 2, no. 1, pp. 4–17, Feb. 2008.

[TTS12] Z. Tian, Y. Tafesse, and B. M. Sadler, "Cyclic feature detection with sub-Nyquist sampling for wideband spectrum sensing," *IEEE Journal of Selected Topics in Signal Processing*, vol. 6, no. 1, pp. 58–69, Feb. 2012.

[Tun05] E. Tuncel, "On error exponents in hypothesis testing," *IEEE Transactions on Information Theory*, vol. 51, no. 8, pp. 2945–2950, Aug. 2005.

[TV11] O. Taheri and S. A. Vorobyov, "Segmented compressed sampling for analog-to-information conversion: Method and performance analysis," *IEEE Transactions on Signal Processing*, vol. 59, no. 2, pp. 554–572, Feb. 2011.

[VB12] H. I. Volos and R. M. Buehrer, "Cognitive radio engine training," *IEEE Transactions on Wireless Communications*, vol. 11, no. 11, pp. 3878–3889, Nov. 2012.

[Ver02] S. Verdú, "Spectral efficiency in the wideband regime," *IEEE Transactions on Information Theory*, vol. 48, no. 6, pp. 1319–1343, Jun. 2002.

[VMB02] M. Vetterli, P. Marziliano, and T. Blu, "Sampling signals with finite rate of innovation," *IEEE Transactions on Signal Processing*, vol. 50, no. 6, pp. 1417–1428, Jun. 2002.

[VO79] A. J. Viterbi and J. K. Omura, *Principles of Digital Communication and Coding*. McGraw-Hill Companies, 1979.

[VP11] P. P. Vaidyanathan and P. Pal, "Sparse sensing with co-pprime samplers and arrays," *IEEE Transactions on Signal Processing*, vol. 59, no. 2, pp. 573–586, Feb. 2011.

[VV07] J. Villares and G. Vázquez, "The Gaussian assumption in second-order estimation problems in digital communications," *IEEE Transactions on Signal Processing*, vol. 55, no. 10, pp. 4994–5002, Oct. 2007.

[VVLVS11] G. Vázquez-Vilar, R. López-Valcarce, and J. Sala, "Multiantenna spectrum sensing exploiting spectral a priori information," *IEEE Transactions on Wireless Communications*, vol. 10, no. 12, pp. 4345–4355, Dec. 2011.

[WAJ07] W. Weber, J. Andrews, and N. Jindal, "The effect of fading, channel inversion, and threshold scheduling on ad hoc networks," *IEEE Transactions on Information Theory*, vol. 53, no. 11, pp. 4127–4149, Nov. 2007.

[WAJ10] S. Weber, J. Andrews, and N. Jindal, "An overview of the transmission capacity of wireless networks," *IEEE Transactions on Communications*, vol. 58, no. 12, pp. 3593–3604, Dec. 2010.

[Wan11] X. Wang, "Joint sensing-channel selection and power control for cognitive radios," *IEEE Transactions on Wireless Communications*, vol. 10, no. 3, pp. 958–967, Mar. 2011.

[WAYdV07] S. Weber, J. Andrews, X. Yang, and G. de Veciana, "Transmission capacity of wireless ad hoc networks with successive interference cancellation," *IEEE Transactions on Information Theory*, vol. 53, no. 8, pp. 2799–2814, Aug. 2007.

[WFHL10] P. Wang, J. Fang, N. Han, and H. Li, "Multiantenna-assisted spectrum sensing for cognitive radio," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 4, pp. 1791–1800, May 2010.

[WGC11] J. Wang, M. Ghosh, and K. Challapali, "Emerging cognitive radio applications: A survey," *IEEE Communications Magazine*, vol. 49, no. 3, pp. 74–81, Mar. 2011.

[WGZ13] S. Wang, M. Ge, and W. Zhao, "Energy-efficient resource allocation for ofdm-based cognitive radio networks," *IEEE Transactions on Communications*, vol. 61, no. 8, pp. 3181–3191, Aug. 2013.

[WL11] B. Wang and K. Liu, "Advances in cognitive radio networks: A survey," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 1, pp. 5–23, Feb. 2011.

[WLL11] X. Wang, H. Li, and H. Lin, "A new adaptive OFDM system with precoded cyclic prefix for dynamic cognitive radio communications," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 2, pp. 431–442, Feb. 2011.

[WPPL09] Y. Wang, A. Pandharipande, Y. L. Polo, and G. Leus, "Distributed compressive wide-band spectrum sensing," in *Information Theory and Applications Workshop (ITA)*, Feb. 2009.

[WT11] Y. Wu and D. Tsang, "Energy-efficient spectrum sensing and transmission for cognitive radio system," *IEEE Communications Letters*, vol. 15, no. 5, pp. 545–547, May 2011.

[WV12] C. Weidmann and M. Vetterli, "Rate distortion behavior of sparse sources," *IEEE Transactions on Information Theory*, vol. 58, no. 8, pp. 4969–4992, Aug. 2012.

[WYAdV05] S. Weber, X. Yang, J. Andrews, and G. de Veciana, "Transmission capacity of wireless ad hoc networks with outage constraints," *IEEE Transactions on Information Theory*, vol. 51, no. 12, pp. 4091–4102, Dec. 2005.

[WYdVA04] S. Weber, X. Yang, G. de Veciana, and J. G. Andrews, "Transmission capacity of CDMA ad hoc networks," in *International Symposium on Spread Spectrum Techniques and Applications (ISSSTA)*, Aug. 2004.

[YHR+09] H. Yi, H. Hu, Y. Rui, K. Guo, and J. Zhang, "Null space-based precoding scheme for secondary transmission in a cognitive radio MIMO system using second-order statistics," in *IEEE Int. Conf. on Commun. (ICC)*, Jun. 2009.

[YLPC11] X. Yang, K. Lei, S. Peng, and X. Cao, "Blind detection for primary user based on the sample covariance matrix in cognitive radio," *IEEE Communications Letters*, vol. 15, no. 1, pp. 40–42, Jan. 2011.

[ZcL09] Y. Zeng and Y. chang Liang, "Eigenvalue-based spectrum sensing algorithms for cognitive radio," *IEEE Transactions on Communications*, vol. 57, no. 6, pp. 1784–1793, Jun. 2009.

[ZGTS08] Q. Zhao, S. Geirhofer, L. Tong, and B. Sadler, "Opportunistic spectrum access via periodic channel sensing," *IEEE Transactions on Signal Processing*, vol. 56, no. 2, pp. 785–796, Feb. 2008.

[ZKL+11] Q. Zhang, S. Kota, V. Lau, W. Su, and A. Kwasinski, "Introduction to the issue on cooperative communication and signal processing in cognitive radio systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 1, pp. 1–4, Feb. 2011.

[ZKLZ13] S. Zheng, P.-Y. Kam, Y.-C. Liang, and Y. Zeng, "Spectrum sensing for digital primary signals in cognitive radio: A Bayesian approach for maximizing spectrum utilization," *IEEE Transactions on Wireless Communications*, vol. 12, no. 4, pp. 1774–1782, Apr. 2013.

[ZL09] Y. Zeng and Y.-C. Liang, "Spectrum-sensing algorithms for cognitive radio based on statistical covariances," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 4, pp. 1804–1815, May 2009.

[ZLLZ10] R. Zhang, T. Lim, Y.-C. Liang, and Y. Zeng, "Multi-antenna based spectrum sensing for cognitive radios: A GLRT approach," *IEEE Transactions on Communications*, vol. 58, no. 1, pp. 84–88, Jan. 2010.

[ZLS13] X. Zhou, G. Y. Li, and G. Sun, "Multiuser spectral precoding for OFDM-based cognitive radio systems," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 3, pp. 345–352, Mar. 2013.

[ZLZ08] Y. Zeng, Y. C. Liang, and R. Zhang, "Blindly combined energy detection for spectrum sensing in cognitive radio," *IEEE Signal Processing Letters*, pp. 649–652, Oct. 2008.

[ZP13] W. Zhang and H. V. Poor, "Opportunistic detection under a fixed-sample-size setting," *IEEE Transactions on Information Theory*, vol. 59, no. 2, pp. 1107–1114, Feb. 2013.

[ZPQ10] W. Zhang, H. Poor, and Z. Quan, "Frequency-domain correlation: An asymptotically optimum approximation of quadratic likelihood ratio detectors," *IEEE Transactions on Signal Processing*, vol. 58, no. 3, pp. 969–979, Mar. 2010.

[ZQ13] P. Zhang and R. Qiu, "GLRT-based spectrum sensing with blindly learned feature under rank-1 assumption," *IEEE Transactions on Communications*, vol. 61, no. 1, pp. 87–96, Jan. 2013.

[ZQG11] P. Zhang, R. Qiu, and N. Guo, "Demonstration of spectrum sensing with blindly learned features," *IEEE Communications Letters*, vol. 15, no. 5, pp. 548–550, May 2011.

[ZS07] Q. Zhao and B. M. Sadler, "A survey of dynamic spectrum access," *IEEE Signal Processing Magazine*, vol. 24, no. 3, pp. 79–89, May 2007.

[ZTSC07] Q. Zhao, L. Tong, A. Swami, and Y. Chen, "Decentralized cognitive MAC for opportunistic spectrum access in ad hoc networks: A POMDP framework," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 3, pp. 589–600, Apr. 2007.