Collaborative Spectrum Sensing in Cognitive Radio Networks

Hongjian Sun

A thesis submitted for the degree of Doctor of Philosophy.
The University of Edinburgh.
January 2011

Abstract

The radio frequency (RF) spectrum is a scarce natural resource, currently regulated by government agencies. With the explosive emergence of wireless applications, the demands for the RF spectrum are constantly increasing. On the other hand, it has been reported that localised temporal and geographic spectrum utilisation efficiency is extremely low. Cognitive radio is an innovative technology designed to improve spectrum utilisation by exploiting those spectrum opportunities. This ability is dependent upon spectrum sensing, which is one of most critical components in a cognitive radio system. A significant challenge is to sense the whole RF spectrum at a particular physical location in a short observation time. Otherwise, performance degrades with longer observation times since the lagging response to spectrum holes implies low spectrum utilisation efficiency. Hence, developing an efficient wideband spectrum sensing technique is prime important.

In this thesis, a multirate asynchronous sub-Nyquist sampling (MASS) system that employs multiple low-rate analog-to-digital converters (ADCs) is developed that implements wideband spectrum sensing. The key features of the MASS system are, 1) low implementation complexity, 2) energy-efficiency for sharing spectrum sensing data, and 3) robustness against the lack of time synchronisation. The conditions under which recovery of the full spectrum is unique are presented using compressive sensing (CS) analysis. The MASS system is applied to both centralised and distributed cognitive radio networks. When the spectra of the cognitive radio nodes have a common spectral support, using one low-rate ADC in each cognitive radio node can successfully recover the full spectrum. This is obtained by applying a hybrid matching pursuit (HMP) algorithm - a synthesis of distributed compressive sensing simultaneous orthogonal matching pursuit (DCS-SOMP) and compressive sampling matching pursuit (CoSaMP). Moreover, a multirate spectrum detection (MSD) system is introduced to detect the primary users from a small number of measurements without ever reconstructing the full spectrum. To achieve a better detection performance, a data fusion strategy is developed for combining sensing data from all cognitive radio nodes. Theoretical bounds on detection performance are derived for distributed cognitive radio nodes suffering from additive white Gaussian noise (AWGN), Rayleigh fading, and log-normal fading channels.

In conclusion, MASS and MSD both have a low implementation complexity, high energy efficiency, good data compression capability, and are applicable to distributed cognitive radio networks.

Declaration of originality

I hereby declare that the work recorded in this thesis and the thesis itself was composed and originated entirely by myself in the School of Engineering at The University of Edinburgh as a candidate for the Doctorate of Philosophy.

Hongjian Sun
January, 2011

Acknowledgements

Foremost, I am heartily thankful to my principal supervisor, Dr. David I. Laurenson, whose warm encouragement, thoughtful guidance, and immense knowledge I will never forget. This thesis would not have been possible without his guidance, insight, patience, and enthusiasm. Dr. David I. Laurenson has been my source of inspiration when I hurdle all the obstacles throughout my research work. I could not have imagined having a better supervisor for my PhD study.

I would like to express my sincere gratitude to Dr. John S. Thompson for his associate supervision. Dr. Thompson has offered much advice and insight throughout my research work on cognitive radio, and opportunities for part-time jobs and attending conferences. I would also like to thank Dr. Thompson for offering a position of research consultant to ETRI in 2008.

I am grateful to Dr. Cheng-Xiang Wang, my associate supervisor in School of Engineering and Physical Sciences in Heriot-Watt University, for his valuable advice, friendly support, and detailed but important comments on my papers. Besides my supervisors, I owe my greatest gratitude to Prof. Mike E. Davies, who gave me lots of constructive comments and suggestions regarding compressive sensing technologies.

I also wish to thank, offer my regards and blessings to IDCOM members for creating an enthusiastic and dynamic research environment, especially to those who supported me in any respect during the completion of the research work. My special thanks to Dr. Mehrdad Yaghoobi and Dr. Gabriel Rilling for their friendly help on mathematics and programming.

The most special thanks I accord my beloved wife, Jing Jiang, for her sweet love, selfless support, and never-ending encouragement over the passed nine years. Without her, it would have been much harder for me to finish my study. With her, I believe I am the luckiest and happiest man in the world. My deepest debt of gratitude is to my parents, Shan Sun and Lijun Chen, for their constant source of support and boundless love.

Finally, I would like to acknowledge the support from the Wolfson Microelectronics Scholarship and the Scottish Funding Council for the Joint Research Institute with Heriot-Watt University which is a part of the Edinburgh Research Partnership in Engineering and Mathematics.

Contents

Declaration of originality . . . iii
Acknowledgements . . . iv
Contents . . . v
List of figures . . . vii
List of tables . . . x
Acronyms and abbreviations . . . xi
Nomenclature . . . xiii

1 Introduction **1**
1.1 Motivation . . . 1
1.2 Objectives and Contributions . . . 3
1.2.1 Objectives . . . 3
1.2.2 Key Contributions . . . 4
1.3 Thesis Outline . . . 4

2 Background **6**
2.1 Cognitive Radio . . . 6
2.1.1 Cognitive Radio Functionalities . . . 7
2.1.2 Potential Applications . . . 10
2.2 Spectrum Sensing Techniques . . . 11
2.2.1 Traditional Spectrum Sensing . . . 12
2.2.2 Cooperative Spectrum Sensing . . . 16
2.2.3 Wideband Spectrum Sensing . . . 19

3 Narrowband Collaborative Spectrum Sensing **30**
3.1 System Description . . . 31
3.2 Spectrum Sensing over A Single Fading Channel . . . 32
3.2.1 Nakagami-m Fading Channel . . . 33
3.2.2 Rician Fading Channel . . . 35
3.2.3 Slow Fading Channel . . . 37
3.3 Data Fusion Based Collaborative Spectrum Sensing . . . 38
3.3.1 Square-Law Combining . . . 39
3.3.2 Square-Law Selection . . . 41
3.3.3 Maximum Ratio Combining . . . 43
3.3.4 Selection Combining . . . 44
3.4 Simulation Results . . . 45
3.5 Conclusions . . . 53

4 Distributed Wideband Spectrum Sensing **56**
4.1 Preliminaries of Compressive Sensing . . . 57
4.2 Multirate Asynchronous Sub-Nyquist Sampling . . . 59
4.2.1 System Design . . . 59
4.2.2 Relate Sub-Nyquist DFT Spectrum to Nyquist DFT Spectrum . . . 61

4.2.3 Effect of Sub-Nyquist Sampling . . . 63
4.2.4 Recovery of the Full Spectrum via Multirate Sampling . . . 65
4.2.5 Practical Implementation Issues . . . 66
4.3 Extension to Distributed Wideband Spectrum Sensing . . . 67
4.4 Simulation Results . . . 71
4.5 Discussion . . . 75
4.5.1 Relationship with Multirate Sampling . . . 75
4.5.2 Connection with Multicoset Sampling . . . 76
4.5.3 Comparison with CS-based Models . . . 76
4.6 Conclusions . . . 77

5 Distributed Wideband Spectrum Detection 78
5.1 Preliminary . . . 79
5.2 Multirate Spectrum Detection . . . 81
5.2.1 System Model . . . 81
5.2.2 Effect of Sub-Nyquist Sampling in A Single Cognitive Radio . . . 83
5.2.3 Data Fusion Strategy among Multiple Cognitive Radios . . . 84
5.3 Performance Analysis . . . 86
5.3.1 Nonfading AWGN Channels . . . 86
5.3.2 Rayleigh Fading Channels . . . 87
5.3.3 Slow Fading Channels . . . 89
5.4 Simulations . . . 91
5.5 Conclusions . . . 98

6 Conclusions and Future Work 101
6.1 Conclusions . . . 101
6.2 Limitation of Work . . . 103
6.3 Future Work . . . 103

A Proof of Proposition 4.1 105

B Proof of Proposition 4.2 108

C Proof of Proposition 4.3 110

D Original publications 112
D.1 Journal papers . . . 112
D.2 Conference papers . . . 112

References 143

List of figures

1.1 Spectrum occupancy measurement results averaged over six locations. 2

2.1 Illustration of spectrum holes and the concept of dynamic spectrum access. . . 7
2.2 The cognitive capability of cognitive radio enabled by a basic cognitive cycle. . 8
2.3 Protocol stack in the cognitive radio networks. 9
2.4 Narrowband spectrum sensing algorithms. 12
2.5 Transmitter detection problems: (a) hidden terminal problem, (b) shadowing uncertainty. 13
2.6 Block diagrams for traditional spectrum sensing algorithms: (a) matched filter, (b) time domain energy detection, (c) frequency domain energy detection, and (d) cyclostationary detection. 14
2.7 Schematic illustration of cooperative spectrum sensing scheme that each cognitive radio transmits its individual observation or decision via control channels to a fusion center, which makes final decision on the spectral occupancy status. 17
2.8 Demonstration of the wideband spectrum of interest. The PSD is smooth within each subband, and exhibits discontinuities and irregularities with the adjacent subbands. 21
2.9 Block diagram of the wavelet transform based wideband spectrum sensing technique. 21
2.10 The graphic illustration of using filter bank on the wideband spectrum. 23
2.11 The demodulation of received wideband signal in the i-th subcarrier. 24
2.12 Block diagram of the compressive sensing based detection. 26
2.13 Block diagram for the analog-to-information converter. 26
2.14 Block diagram for the modulated wideband converter. 28
2.15 Multirate sampling system implemented by electro-optical devices. 29

3.1 Block diagram of the energy detector. 31
3.2 Complementary ROC curves of energy detection over a Nakagami-m fading channel, with (a) $m = 1$, and (b) $m = 2$. This figure is produced by changing the detection threshold λ from 0.1 to 1000. 47
3.3 Complementary ROC curves of energy detection over a Rician fading channel, with (a) $K = 1$, and (b) $K = 3$. This figure is produced by changing the detection threshold λ from 0.1 to 1000. 48
3.4 The truncation error after N iterations on index n, when the time bandwidth product $u = 1$, the local-mean SNR $\overline{\gamma} = 10$ dB, and the average probability of false alarm $\overline{P_f} = 0.01$, with (a) Nakagami-m fading, and (b) Rician fading. . . 49
3.5 Complementary ROC curves of energy detection over a slow fading channel with (a) the shadow standard deviation $\sigma = 4$ dB, and (b) the area-mean SNR $\ddot{\overline{\gamma}} = 10$ dB, compared with theoretical result in (3.31). 50
3.6 Complementary ROC curves of energy detection under MRC and SLC scheme, when the Nakagami fading factor $m = 3$, and the time bandwidth product $u = 1$. 51

3.7 Complementary ROC curves of energy detection under SC and SLS scheme, when the Nakagami fading factor $m = 3$, and the time bandwidth product $u = 1$. 52
3.8 Comparison of MRC, SLC, SC, and SLS, when the local-mean SNR $\overline{\gamma} =$ 15 dB, the number of collaborative cognitive radios $v = 3$, and the time bandwidth product $u = 1$. 53
3.9 Comparison of MRC, SLC, SC, and SLS, when the local-mean SNR $\overline{\gamma} =$ 15 dB, the number of collaborative cognitive radios $v = 3$, and the Nakagami fading factor $m = 3$. 54
3.10 Comparison of MRC, SLC, SC, and SLS, when the time bandwidth product $u = 1$, the number of collaborative cognitive radios $v = 3$, and the Nakagami fading factor $m = 3$. 55

4.1 The schematic illustration of the multirate asynchronous sub-Nyquist sampling system in one cognitive radio node. The wideband filters are altered to have the same bandwidth of W. 59
4.2 Interpretation of the relationship between the non-aliased spectrum, $\vec{X}$, and the sub-Nyquist rate spectrum, $\vec{Y}$, when the Fourier spectrum of $x(t)$ is denoted by $X(f)$. 61
4.3 Simulated probability of no signal overlap at a frequency $f = m\Delta f + \varsigma$ in the sub-Nyquist rate spectrum $\vec{Y}$, compared to the theoretical result in (4.15). The number of samples at the sub-Nyquist sampling rate is $M^i = \sqrt{N} = 3000$, where $N = 9 \times 10^6$ denotes the number of samples at the Nyquist sampling rate. 64
4.4 The schematic illustration of the distributed MASS system in cognitive radio networks. Each cognitive radio is only required to be equipped with one wideband filter, one low-rate ADC, and an FFT. The bandwidth of the wideband filters should be altered to have the same bandwidth of W. 68
4.5 Implementation of the distributed MASS system in cognitive radio networks. Each ADC senses the wideband spectrum at different sub-Nyquist sampling rate, and then sends the spectrum sensing data through a control channel to a fusion centre, where the full spectrum is reconstructed. 69
4.6 Illustration of the spectral correlation in distributed cognitive radio nodes. . . . 70
4.7 Comparison of the non-aliased DFT spectrum with the recovered spectrum, when the wideband signal has 5 subbands with bandwidth $0.3 \sim 30$ MHz, and SNR$=$ 15 dB. There are $v = 10$ ADCs, which experience non-fading AWGN channels. The total number of measurements is $23.85\%N$, rather than N when sampled at the Nyquist rate. 73
4.8 Influence of the number of sub-bands, and the compression ratio on the detection performance of the MASS system, when the signals from PUs have 10, 30, and 50 sub-bands with bandwidth $0.1 \sim 5$ MHz, and experience non-fading AWGN channels with SNR$=$ 15 dB. 74
4.9 Comparison of spectrum recovery performance for synchronous ADCs and asynchronous ADCs, with $N_b = 30$ and $N_b = 50$. The asynchronous ADCs have time offsets in range of $0 \sim 0.8$ μs, with a total observation time of 2 μs. 75
4.10 Performance of MASS over AWGN, Rayleigh, and slow fading channels, with the number of sub-bands $N_b = 30$. In this experiment, the average SNR is 15 dB, and the standard deviation of log-normal fading is 5 dB. 76

5.1 Schematic illustration of wideband spectrum estimation and detection. The Welch's method is employed for calculating the PSD. 79
5.2 Schematic illustration of the multirate spectrum detection system in a distributed cognitive radio network. Wideband filters prior to ADCs are altered to have the same bandwidth of W, and the wideband signal is sampled at different sub-Nyquist rates in different cognitive radio nodes. 81
5.3 Comparison of the original PSD in (a) with the MSD PSD in (b) over i.i.d. Rayleigh fading channels. In the experiment, the compression ratio is $M/N = 49.67\%$, with 75 ADCs at the average sampling rate of 132.46 MHz, and the SNR averaged over fading channels is 10 dB. 92
5.4 Comparison of the original PSD in (a) with the MSD PSD in (b) over i.i.d. slow fading channels with the standard deviation $\sigma = 5$ dB. In the experiment, the compression ratio is $M/N = 49.67\%$, with 75 ADCs at the average sampling rate of 132.46 MHz, and the SNR averaged over fading channels is 10 dB. . . . 93
5.5 Comparisons of empirical results and theoretical results for the probabilities of detection and false alarm over (a) AWGN, (b) Rayleigh, and (c) slow fading channels. The theoretical bounds are predicted in equations (5.26)-(5.27), (5.33)-(5.34), and (5.44)-(5.45). 94
5.6 ROC curves of MSD over different kinds of fading channels when the compression ratio $M/N = 49.2\%$ and the averaging times $J = 50$. The wideband signal is observed by 22 cognitive radios at different sampling rates (the average sampling rate is 448.68 MHz). The shadowing standard deviation is 5 dB. 95
5.7 ROC curves of MSD over slow fading channels with the compression ratio $M/N = 49.2\%$ and $J = 50$, when the shadowing standard deviation, σ, and the average SNR alter. The wideband signal is sampled at different sampling rates by 22 ADCs with the average sampling rate of 448.68 MHz. 96
5.8 Complementary ROC curves of the proposed MSD system over slow fading channels with SNR$= 5$ dB, $\sigma = 5$ dB and $J = 50$, when the compression ratio (M/N) varies. 97
5.9 ROC curves of the proposed MSD system over slow fading channels with $\sigma = 5$ dB, $M/N = 68.2\%$, and $J = 50$, when the SNR and the number of subbands change. 98
5.10 ROC curves of the MSD system over slow fading channels with SNR$= 5$ dB, $\sigma = 5$ dB, and the compression ratio $M/N = 68.2\%$, when the averaging times, J, and the number of subbands alter. 99
5.11 For fixed parameters: compression ratio $M/N = 49.67\%$, SNR$= 5$ dB, and $J = 50$, the tradeoff between the number of ADCs and the average sampling rate, i.e. 75 ADCs with the average sampling rate 132.46 MHz, 40 ADCs with the average sampling rate 247.45 MHz, and 22 ADCs with the average sampling rate 448.68 MHz. 100

List of tables

2.1 Summary of advantages and disadvantages of narrowband spectrum sensing algorithms. 15
2.2 Summary of advantages and disadvantages of wideband spectrum sensing algorithms. 20

4.1 Hybrid matching pursuit algorithm for distributed MASS system 71
4.2 List of simulation parameters setup for Figure 4.7, Figure 4.8, Figure 4.9, and Figure 4.10. 72

Acronyms and abbreviations

ADC	Analog-to-digital converter
AIC	Analog-to-information converter
AWGN	Additive white Gaussian noise
BP	Basis pursuit
BPF	Bandpass filter
CoSaMP	Compressive sampling matching pursuit
CS	Compressive sensing
CSF	Cyclic spectrum function
CSI	Channel state information
CWT	Continuous wavelet transform
DCS-SOMP	Distributed compressive sensing simultaneous orthogonal matching pursuit
DFT	Discrete Fourier transform
DTFT	Discrete-time Fourier transform
DoF	Degree of freedom
EGC	Equal gain combining
FC	Fusion centre
FCC	Federal communications commission
FFT	Fast Fourier transform
GSM	Global system for mobile
HMP	Hybrid matching pursuit
i.i.d.	Independent and identically distributed
JSM-2	Second joint sparsity model
LOS	Line of sight
LPBP	Linear programming basis pursuit
MASS	Multirate asynchronous sub-Nyquist sampling
MRS	Multirate sampling
MRC	Maximum ratio combining
MSD	Multirate spectrum detection
MWC	Modulated wideband converter

Ofcom	Office of communication
OMP	Orthogonal matching pursuit
OPG	Optical pulse generator
PDF	Probability density function
PSD	Power spectrum density
PU	Primary user
QoS	Quality of service
RF	Radio frequency
RIP	Restricted isometry property
ROC	Receiver operating characteristic
SC	Selection combining
SLC	Square-law combining
SLS	Square-law selection
SMRS	Synchronous multirate sampling
SNR	Signal-to-noise ratio
SPTF	Spectrum policy task force
TNBF	Tunable narrowband bandpass filter
TV	Television
UWB	Ultra-wideband
WLAN	Wireless local area network
WRAN	Wireless regional area network
WSS	Wide-sense stationary
xG	Next generation communication network

Nomenclature

$(a)_n$	Pochhammer symbol
$\mathbf{B}$	Off-diagonal matrix
B_l	Bandwidth of l-th subband
C	Sampling pattern
C_a^b	Binomial coefficient
$\mathrm{CWT}x$	Continuous wavelet transform of x
d	Dyadic scale
$\mathrm{E}(x)$	Expectation of x
E_l^i	Energy of the l-th subband in cognitive radio node i
e^x	Exponential function
$\mathbf{F}$	Fourier matrix
$f(x)$	Probability density function of x
f^i	Sub-Nyquist sampling rate at node i
f_l	Carrier frequency of l-th subband
f_s	Sampling rate greater than Nyquist rate
$\mathbf{G}$	Gram matrix
$G_N(f)$	Frequency response of a Nyquist (N) filter
$g_N(n)$	Temporal impulse response of a Nyquist (N) filter
H_0	Hypothesis that the signal is absent
H_1	Hypothesis that the signal is present
$H(f)$	Frequency response of a filter
$h(t)$	Complex channel gain between PU and cognitive radio
h^i	Complex channel gain between PU and cognitive radio node i
$\mathbf{I}$	Identity matrix
$I_a(x)$	a^{th} order modified Bessel function of the first kind
J	Number of segments
j	Imaginary unit that $j^2 = -1$ in Chapter 2
K	Rician fading shape parameter
$K_a(x)$	a^{th} order modified Bessel function of the second kind

k	Sparsity of a signal
L	Number of samples in a sampling block
M	Total number of measurements
$\overline{M}$	Average number of measurements
M^i	Number of measurements in cognitive radio node i
m	Nakagami-m fading factor in Chapter III
N	Number of samples when sampled at rate of f_s
	Number of computed terms in Chapter III
N_b	Number of subbands
$\mathcal{O}(x)$	Of the order of x
$\Pr(x)$	Probability of event x occurring
P_d	Probability of detection
P_f	Probability of false alarm
$p_c(t)$	Output of a pseudorandom number generator
$Q_u(a,x)$	Generalized Marcum Q-function
$\mathbf{R}$	Correlation matrix
R	Window hop size
$R_i(\mathbf{G})$	Deleted absolute row sums of matrix $\mathbf{G}$
$S(f)$	Power spectrum density
$S(f,a)$	Cyclic spectrum
$s(t)$	Transmit signal from PU
T	Observation time
$T(n)$	Truncation error after n computed terms
t	Time
t^i	Time offset of sample i from the origin
U	Universal set for all frequency bins in full spectrum
u	Time bandwidth product
v	Number of collaborative cognitive radio nodes
W	Signal bandwidth
$w_T(t)$	Tapered window
$X(f)$	Fourier transform of signal $x(t)$
$x(t)$	Time-domain signal
Y	Test statistic

$z(t)$	Additive white Gaussian noise
$\Gamma(x)$	Gamma function
$\Gamma(a, x)$	Upper incomplete gamma function
$\Gamma'(a, x)$	Lower incomplete gamma function
γ	Signal-to-noise ratio
$\overline{\gamma}$	Local-mean Signal-to-noise ratio
$\dddot{\gamma}$	Area-mean Signal-to-noise ratio
Δf	Frequency resolution
Δt^i	Sampling interval at cognitive radio node i
Δ_l	Time offset of the l-th subband
$\delta(t)$	Dirac delta function
δ^2	Time-domain noise variance
δ_n^2	Frequency-domain noise variance for full spectrum
δ_s^2	Noise variance in the scaled sub-Nyquist DFT spectrum
$\epsilon(x, \nu, z)$	Defined function in (3.46)
ε_i	i-th chipping sequence
η	Wald shape parameter
θ	Wald distribution parameter
λ	Detection threshold
λ_i	The i-th eigenvalue of a matrix
μ	Mutual coherence of a matrix
ξ	Constant has value of $10/\ln(10)$
ϖ	Alternating rate of a pseudorandom number generator
ρ	Recovery error tolerance
ϱ_k	Restricted isometry constant
σ	Standard deviation
$\mathbf{\Phi}$	Measurement matrix
$\mathbf{\Phi_\Omega}$	Sub-matrix of $\mathbf{\Phi}$ by choosing columns in support Ω
$\Phi(a, b; x)$	Confluent Hypergeometric function
$\vec{\phi}_j$	j-th column of measurement matrix $\mathbf{\Phi}$
$\phi_{i,j}$	Element of measurement matrix $\mathbf{\Phi}$
$\varphi(f)$	Wavelet smoothing function
$\varphi_d(f)$	Dilation of $\phi(f)$ with a dyadic scale of d

χ^2_a	Central chi-square distribution with a DoF
$\chi^2_a(b)$	Non-central chi-square distribution with a DoF, and non-central parameter b
$\Psi_{y_i y_i}(z)$	z-transform of correlation coefficient $\psi_{y_i y_i}(k)$
$\psi_{xx}(k)$	Correlation coefficients of x
Ω	Support
$\overline{\Omega}$	Relative complement of Ω in universal set U
${}_pF_q(a_1, \cdots, a_p;$ $b_1, \cdots, b_q; z)$	Generalized hypergeometric series
$\|\cdot\|_i$	ℓ_i norm
$\|\cdot\|$	Absolute operation
$\lceil\cdot\rceil$	The smallest integer not less than the argument
$\lfloor\cdot\rfloor$	The largest integer not greater than the argument
$\|\cdot\|_{\text{mod}}$	Modulo operation
$*$	Convolution operation
$< a, b >$	Inner product
arg min $(\cdot)$	Minimising argument
Var$(\cdot)$	Variance
$(\cdot)!$	Factorial of a positive integer
$\vec{x}$, $\overrightarrow{X}$	Vector
$\hat{\phi}$, $\widehat{\mathbf{\Phi}}$	Vector or matrix normalised with ℓ_2 norm
$[\cdot]^T$	Transposition operation
$[\cdot]^H$	Hermitian transpose operation
$[\cdot]^*$	Complex conjugate operation
$[\cdot]^\dagger$	Pseudoinverse operation
$[\cdot]^{-1}$	Inverse operation

Chapter 1
Introduction

This thesis addresses the issues of spectrum sensing in cognitive radio networks. Due to current spectral underutilisation, an innovative technology, cognitive radio, has been designed to exploit spectrum holes. Spectrum sensing is the critical component upon which the full operation of cognitive radio relies. Wideband spectrum sensing is the key technology that enables the efficient operation of both the primary user and the cognitive radio networks. However, wideband spectrum sensing systems are difficult to design, due to either high implementation complexity or high energy consumption from high-rate analog-to-digital converter (ADC). This thesis will present two sub-Nyquist sampling systems that implement wideband spectrum sensing by multiple low-rate ADCs. Both systems have low implementation complexity, low energy consumption, and are suitable for distributed cognitive radio networks.

1.1 Motivation

The radio frequency (RF) spectrum is a limited natural resource managed by government regulators, such as the office of communications (Ofcom [1]) in the United Kingdom, and the federal communications commission (FCC [2]) in the United States. Under current policy, all frequency bands are exclusively assigned to wireless networks on a long term basis for large geographical regions, and each system has to operate within a particular band. With the increasing emergence of new wireless products and the explosive development of mobile internet applications, the demands on RF spectrum have been constantly increasing. In recent years, it has become evident that there will not be enough spectrum exclusively available for all wireless systems currently under development. Interestingly, the spectrum policy task force (SPTF) within the FCC has reported that localised temporal and geographic spectrum utilisation efficiency ranges from 15% to 85% [3]. In another experiment as shown in Figure 1.1, the maximal occupancy of the spectrum from 30 MHz to 3 GHz (in New York city) has been reported to be only 13.1%, with average occupancy (over six locations) of 5.2% [4]. Spectral underutilisation can be solved by allowing a secondary user to access a licensed band when the primary user

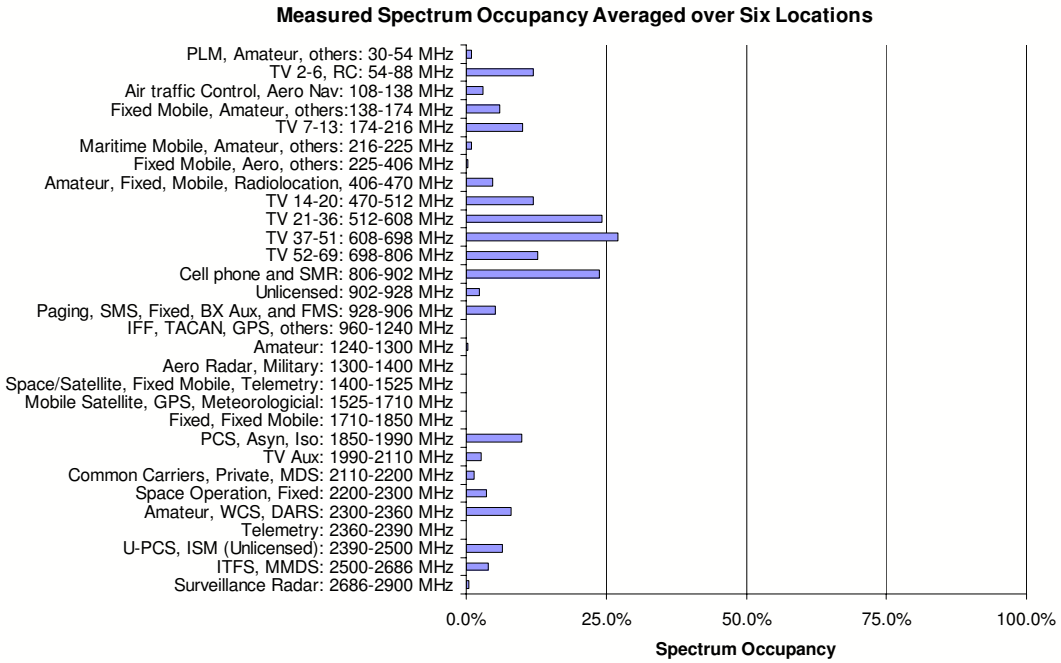

Figure 1.1: *Spectrum occupancy measurement results averaged over six locations [4].*

(PU) is absent [3]. Cognitive radio [5] has emerged as one of the most promising candidates for realising this [3, 6–11].

As an innovative technology, cognitive radio is designed to exploit spectrum opportunities by means of sensing and adapting to the environment. A crucial requirement of cognitive radios is that they must rapidly fill in spectrum holes without posing harmful interference on the PUs. This task is dependent upon the function of spectrum sensing, which is one of the critical components in a cognitive radio system. Due to effects of multipath and shadowing, a cognitive radio user cannot distinguish between a deeply faded band and an idle one. In order to mitigate these effects, cognitive radio users often collaborate for spectrum sensing [12–17]. Different collaborative strategies will result in distinct performance. On the other hand, there is a significant challenge in sensing the whole of the spectrum at a particular physical location in a short observation time. Otherwise, the performance of the cognitive radio system degrades due to a lagging response to spectrum holes. Hence, wideband spectrum sensing is of prime importance to ensure efficient operation of both the primary and the secondary (cognitive radio) networks.

Many extensive studies have been carried out to develop efficient and reliable spectrum sensing methods. Despite numerous spectrum sensing algorithms being reported in the literature [7, 12–29], few of them are effective for wideband spectrum sensing due to energy and hardware

constraints. To the best of the author's knowledge, only five approaches have appeared in the literature which offer the possibility of implementing wideband spectrum sensing. In the first model, a wavelet transform approach [24] is used to detect the PUs by searching discontinuities and irregularities in the power spectrum density (PSD) of the received signal. However, a high sampling rate ADC is required and the energy cost of that ADC will be prohibitive. The second model is known as filter bank detection [25, 26]. This model was developed based on the assumption that multicarrier communications are used in cognitive radio networks. A pair of matched root-Nyquist filters are employed in the primary transmitter and the cognitive radio receiver, requiring a large number of RF components [27, 28]. Moreover, in practice, the filter information at the PUs are usually unknown in the cognitive radio networks. Either multicoset sampling [30–32] or multirate sampling [33, 34] techniques can be used to reduce the high sampling rate, then the signal is recovered from measurements. However, to design a specific sampling pattern, requirements on both time synchronisation and devices are stringent. Finally, compressive sensing (CS) based methods were suggested in [22, 27–29, 35] to use under-sampled measurements for reconstructing the full spectrum. Then spectrum sensing is performed on the reconstructed spectrum. Despite their low sampling rate, the performance of CS based methods are dependent upon storage and transmission of measurement matrix.

1.2 Objectives and Contributions

1.2.1 Objectives

The aim of this thesis is to study the performance of collaborative spectrum sensing algorithms, and develop efficient wideband spectrum sensing techniques that can be used in distributed cognitive radio networks over fading channels. More specifically, the study has the following objectives:

- To analyse the detection performance of different collaborative spectrum sensing algorithms over fading channels.
- To develop wideband spectrum sensing techniques that are applicable to distributed cognitive radio networks with low implementation complexity, high energy efficiency, and good data compression capability.

1.2.2 Key Contributions

The main contributions of this thesis are summarised as follows:

- Easily computed expressions for the average probability of detection are derived for a cognitive radio using energy detection but suffering from Nakagami-m, and Rician fading channels. For a slow log-normal fading channel, an approximation of the average probability of detection is given, by using the Wald distribution to replace the log-normal distribution.
- The detection performance of energy detection using different collaborative strategies is derived. Specifically, maximum ratio combining (MRC), selective combining (SC), square-law combining (SLC), and square-law selection (SLS) approaches are analysed and compared under different constraints.
- A multirate asynchronous sub-Nyquist sampling (MASS) system that employs multiple low-rate ADCs is developed that implements wideband spectrum sensing using a few measurements. The MASS system not only has a low implementation complexity, but also is energy-efficient for sharing spectrum sensing data. The conditions under which recovery of the full spectrum is unique are presented using CS techniques. A trade-off is made between the number of sampling channels and the probability of successful spectrum recovery. The MASS system is applied to distributed cognitive radio networks. When the spectra of the cognitive radio nodes have a common spectral support, using one low-rate ADC in each cognitive radio node can successfully recover the full spectrum.
- A multirate spectrum detection (MSD) system is developed to detect the PUs from a small number of measurements without reconstructing the full spectrum. In order to achieve a better detection performance, a data fusion strategy is obtained for combining spectrum sensing data. The theoretical bounds on detection performance of MSD are derived for distributed cognitive radio nodes when they are suffering from independent and identically distributed (i.i.d.) additive white Gaussian noise (AWGN), Rayleigh fading, and log-normal fading channels.

1.3 Thesis Outline

The remainder of this thesis is organized as follows:

Chapter 2

This chapter first presents the structure, functionalities, and potential applications of cognitive radios. Some traditional spectrum sensing algorithms in the literature are then introduced, and their advantages and disadvantages are summarised. Finally, some related approaches for collaborative spectrum sensing and wideband spectrum sensing are discussed.

Chapter 3

This chapter derives some easily computed expressions for the average probabilities of false alarm and detection when the cognitive radio is using energy detection method but suffering from a fading channel. For cognitive radios designed to collaborate using MRC, SC, SLC, or SLS over i.i.d. Nakagami-m fading channels, the average probabilities of false alarm and detection are given.

Chapter 4

This chapter introduces a MASS system used to implement wideband spectrum sensing. Using CS theory, the sufficient conditions of full spectrum recovery are derived. The practical implementation issues, for example, the effect of noise and model mismatch, and the tradeoff between the number of sampling channels and the probability of successful spectrum recovery, are discussed. When the MASS system is applied to distributed cognitive radio networks, a hybrid matching pursuit (HMP) algorithm is proposed for reconstructing the full spectrum using fewer measurements.

Chapter 5

This chapter proposes a MSD system for implementing wideband spectrum detection without reconstructing the full spectrum. The effect of sub-Nyquist sampling in a single cognitive radio node is considered, followed by designing the data fusion rule for multiple cognitive radio nodes. The detection performance of MSD over fading channels is analysed, and some theoretical bounds on the detection performance of the proposed system are derived.

Chapter 2
Background

As an intelligent radio, cognitive radio is the key technology that provides the capability to use the RF spectrum in a dynamic manner [3, 5–11]. A crucial requirement of cognitive radios is that they must rapidly fill in spectrum holes without causing harmful interference to the PUs [7]. This task is fulfilled by the function of spectrum sensing. However, there are two challenges in spectrum sensing. One of them is that due to effect of multipath and shadowing, the sensing result of a single cognitive radio user is not reliable. Thus, collaborative spectrum sensing techniques are often used to combat the effect of fading. Another significant challenge is sensing the whole of the spectrum at a particular physical location in a short observation time. Hence, wideband spectrum sensing is of prime importance to ensure efficient operation of both the primary and the secondary (cognitive radio) networks.

In this chapter, Section 2.1 presents the structure and the functionalities of cognitive radios, and some potential applications of cognitive radio technology are introduced. A literature review of spectrum sensing algorithms is presented in Section 2.2. This section first discusses traditional spectrum sensing techniques, i.e. matched filter, energy detection, and cyclostationary detection, in Section 2.2.1. Cooperative spectrum sensing strategies, i.e. data fusion or decision fusion, are then analysed in Section 2.2.2. The more challenging techniques for wideband spectrum sensing, e.g. wavelet detection, filter bank detection, and compressive sensing based detection, are presented in Section 2.2.3.

2.1 Cognitive Radio

"Cognitive radio is viewed as a novel approach for improving the utilisation of a precious natural resource: the radio electromagnetic spectrum." **-S. Haykin** [6]

For the purpose of improving the spectrum utilisation efficiency and providing high bandwidth to mobile users, the next generation communication networks (xG) [3] program was developed to implement spectrum policy intelligent radios, also known as cognitive radios [5], by dynamic

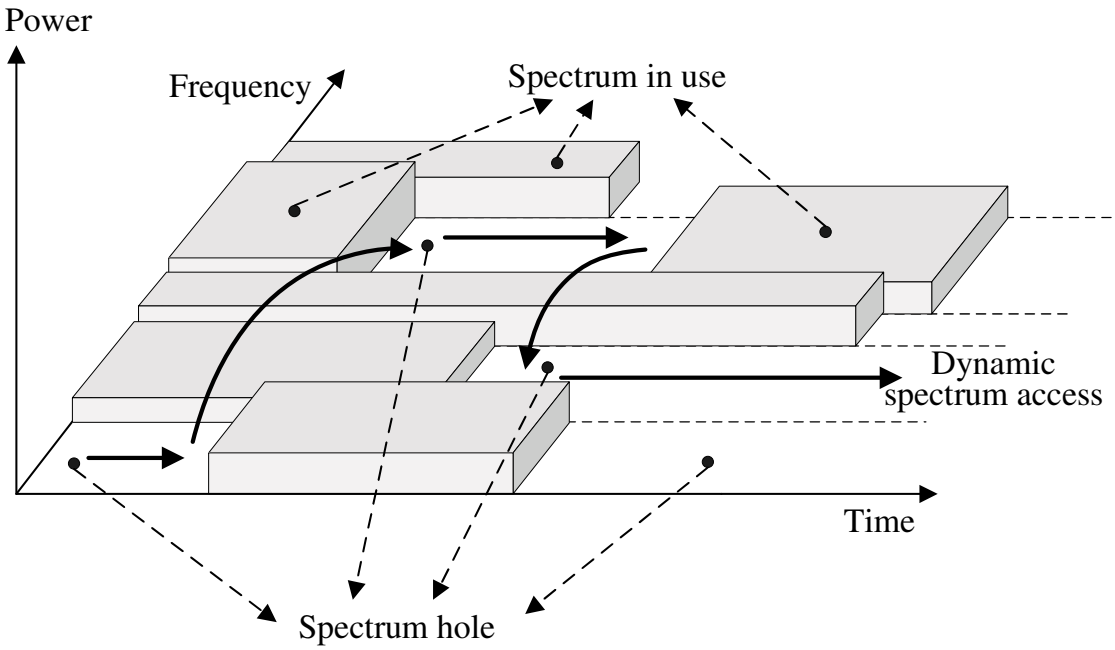

Figure 2.1: *Illustration of spectrum holes and the concept of dynamic spectrum access [38].*

spectrum access techniques as shown in Figure 2.1. Furthermore, the IEEE has organised a new working group, known as the wireless regional area network (WRAN, IEEE 802.22 [36]), for using cognitive radio techniques to allow sharing of geographically unused television (TV) spectrum on a non-interfering basis [9, 20, 37].

2.1.1 Cognitive Radio Functionalities

The term *cognitive radio* was first coined by Mitola in [5] and has the following formal definition as [6]:

"Cognitive radio is an intelligent wireless communication system that is aware of its surrounding environment (i.e. outside world), and uses the methodology of understanding-by-building to learn from the environment and adapt its internal states to statistical variations in the incoming RF stimuli by making corresponding changes in certain operating parameters (e.g., transmit-power, carrier-frequency, and modulation strategy) in real-time, with two primary objectives in mind:

- *highly reliable communications whenever and wherever needed;*
- *efficient utilisation of the radio spectrum. "***-S. Haykin**

From the definition, the two main characteristics of cognitive radio can be summarised as cogni-

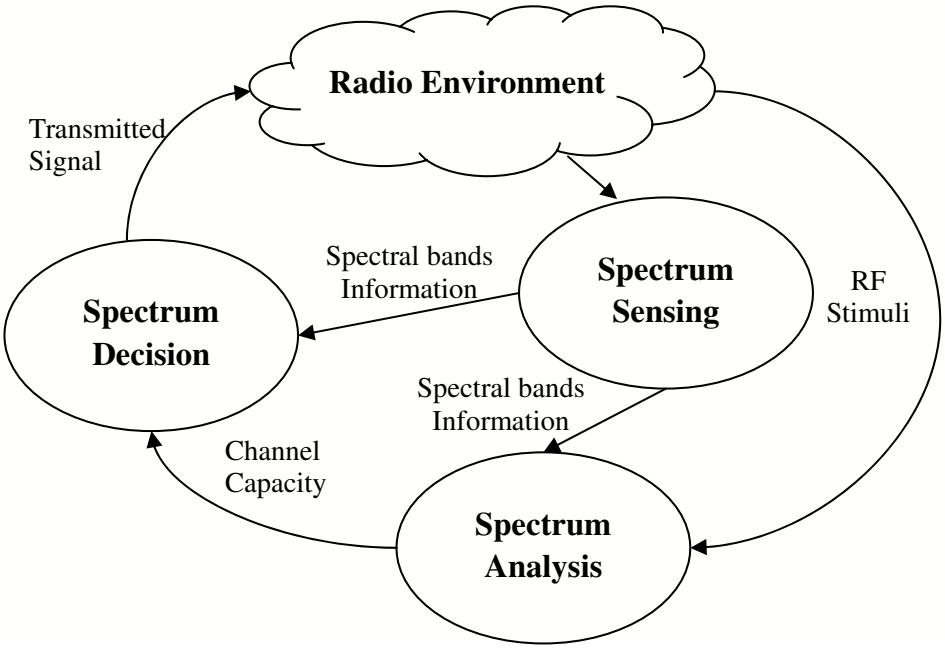

Figure 2.2: *The cognitive capability of cognitive radio enabled by a basic cognitive cycle [3].*

tive capability, and reconfigurability [3]. The former one enables the cognitive radio to interact with its environment in a real-time manner, and intelligently determine appropriate communication parameters based on quality of service (QoS) requirements. These tasks can be implemented by a basic cognitive cycle: spectrum sensing, spectrum analysis, and spectrum decision as shown in Figure 2.2 [3].

- *Spectrum sensing*: Either by cooperating or not, the cognitive radio nodes regularly monitor the RF environment. To improve the spectral usage efficiency, cognitive radio nodes should not only find spectrum holes by sensing some particular spectrum, but also monitor the whole spectral band.

- *Spectrum analysis*: The characteristics of the spectral bands that are sensed through spectrum sensing are estimated. The estimation results, e.g., capacity, and reliability, will be delivered to the spectrum decision step.

- *Spectrum decision*: According to the spectrum characteristics analysed above, an appropriate spectral band will be chosen for a particular cognitive radio node. Then the cognitive radio determines new configuration parameters, e.g., data rate, transmission mode, and bandwidth of the transmission.

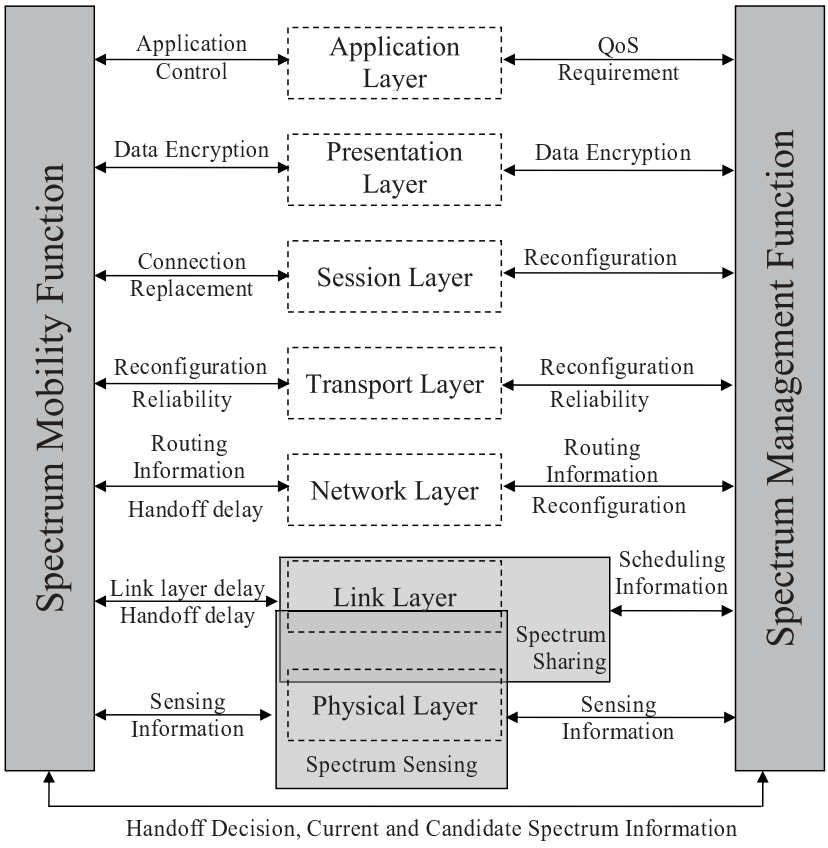

Figure 2.3: *Protocol stack in the cognitive radio networks [3,38].*

Another key feature of cognitive radio is reconfigurability. In order to adapt to RF environment, cognitive radio should change its operational parameters [3]:

- *Operating frequency*: cognitive radio is capable of changing its operating frequency in order to avoid the PU or to share spectrum with other users.
- *Modulation scheme*: cognitive radio should adaptively reconfigure the modulation scheme, according to the user requirements and the channel conditions.
- *Transmission power*: Within the power constraints, transmission power can be reconfigured in order to mitigate interference or improve spectral efficiency.
- *Communication technology*: cognitive radio can also be used to provide interoperability among different communication systems by changing modulation scheme etc.

Figure 2.3 shows the protocol stack in a cognitive radio network [3]. Spectrum sensing is the foundation of all other cognitive radio functions. Based on the results of spectrum sensing, the

function of spectrum sharing allocates spectrum holes considering both fairness and QoS requirements. Thus it requires coordination and reconfiguration among cognitive radio terminals. On the other hand, the functions of spectrum mobility and spectrum management require interactions with all other layers for exchanging information about QoS requirements, application control, routing, reconfiguration, and scheduling. The functionalities of the cognitive radio can be summarised as [3, 38]:

- *Spectrum sensing*: The aim of the spectrum sensing is to detect spectrum holes, and monitor the activity of the PUs. When the PUs reappear, the cognitive radio should release the spectrum without posing harmful interference on the PUs.
- *Spectrum sharing*: While there are several coexisting cognitive radio users, sharing the spectrum while considering both fairness and spectral efficiency is very important. These sharing strategies should depend not only on the spectrum availability, but also on the users' QoS requirements.
- *Spectrum management*: The goal of spectrum management is to provide flexible, fair, and efficient usage of the radio resource. According to the result of the spectrum analysis, spectrum management can improve the spectral utilisation efficiency by providing appropriate spectrum holes to the cognitive radio users, as well as considering their QoS requirements.
- *Spectrum mobility*: When the current operating frequency becomes unavailable during communication due to either changes over time, or movements of the cognitive radio users, the system needs to switch to other bands in a seamless manner.

2.1.2 Potential Applications

Because cognitive radio is aware of the RF environment and is capable of adapting its transmission parameters to the RF spectrum environment, cognitive radios and the concepts of cognitive radio can be applied to a variety of wireless communication environments, especially in commercial and military applications. A few of applications are listed below:

- *Coexistence of wireless technologies [8]*: Cognitive radio techniques were primarily considered for reusing the spectrum that is currently allocated to the TV service. WRAN users can take advantage of broadband data delivery by the opportunistic usage of the

underutilised spectrum. Additionally, the dynamic spectrum access techniques will play an important role in full interoperability and coexistence among diverse technologies for wireless networks. For example, cognitive radio concepts can be used to optimise and manage the spectrum when the wireless local area network (WLAN) and the Bluetooth devices coexist.

- *Military networks [8,39]*: In military communications, bandwidth is often at a premium. By using cognitive radio concepts, military radios can not only achieve substantial spectral efficiency on a noninterfering basis, but also reduce implementation complexity for defining the spectrum allocation for each user. Furthermore, military radios can obtain benefits from the opportunistic spectrum access function supported by the cognitive radio [8]. For example, the military radios can adapt their transmission parameters to use Global System for Mobile (GSM) bands, or other commercial bands when their original frequencies are jammed. The mechanism of spectrum management can help the military radios achieve information superiority on the battlefield. Furthermore, from the soldiers' perspective, cognitive radio can help the soldiers to reach an objective through its situational awareness.

- *Heterogeneous wireless networks [8,40]*: From a user's point of view, a cognitive radio device can dynamically discover information about access networks, e.g. WiFi and GSM, and makes decisions on which access network is most suitable for its requirements and preferences. Then the cognitive radio device will reconfigure itself to connect to the best access network. When the environmental conditions change, the cognitive radio device can adapt to these changes. The information as seen by the cognitive radio user is as transparent as possible to changes in the communication environment.

2.2 Spectrum Sensing Techniques

As PU systems have higher priority than secondary users for using the allocated frequencies, cognitive radios should either avoid interference to PUs, or keep the interference level lower than a threshold. To exploit spectrum opportunities, cognitive radio must detect spectrum holes. Most of the functions in the cognitive radio rely on spectrum sensing for implementing its environmental awareness. Narrowband spectrum sensing algorithms can be classified as shown in Figure 2.4. The most efficient way to sense spectrum holes is to detect active

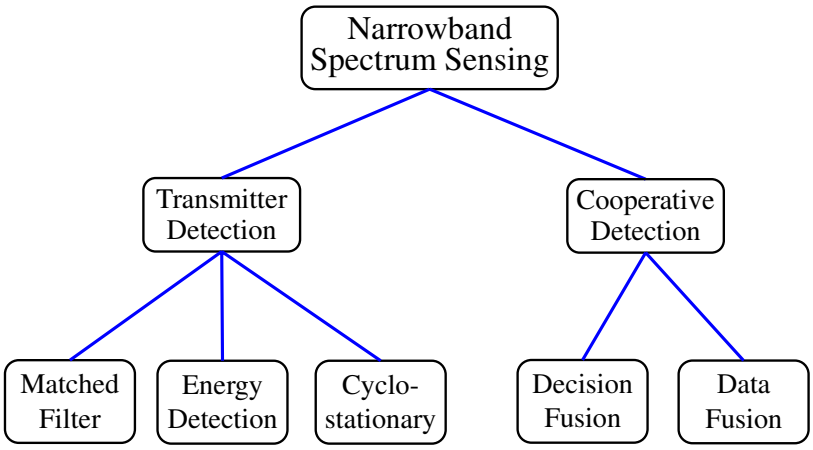

Figure 2.4: *Narrowband spectrum sensing algorithms [3, 7].*

primary transceivers in the vicinity of the cognitive radios [6]. However, as some primary receivers are passive, such as TVs, some are difficult to detect in practice. Three commonly used techniques for detecting the primary transmitters are matched filtering [23, 41], energy detection [10, 42–44], and cyclostationary detection [17, 21, 37, 45–49].

As a primary receiver may be passive, the cognitive radio may not be able to avoid generating interference to the primary users when the primary transmitter is out of the cognitive radio's detectable range. This problem is referred to as the hidden terminal problem as shown in Figure 2.5(a). Additionally, because of shadowing as illustrated in Figure 2.5(b), a cognitive radio user cannot distinguish between a deeply faded band and an idle one. To address these issues, cooperative spectrum sensing can be used to mitigate the effect of shadowing and the primary receiver location uncertainty [7, 8].

2.2.1 Traditional Spectrum Sensing

In this section, three typical spectrum sensing algorithms will be discussed. The implementation of these algorithms requires different conditions, and their detection performance are correspondingly distinguished. The advantages and disadvantages of these algorithms are summarised in Table 2.1.

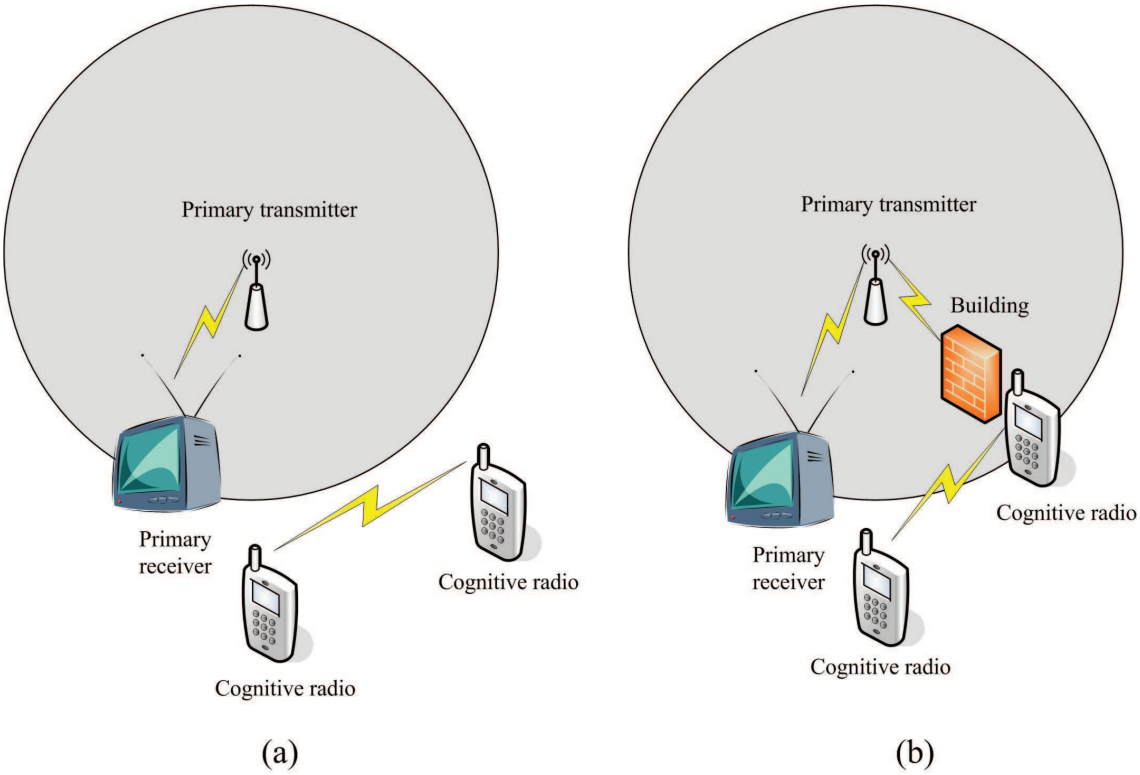

Figure 2.5: *Transmitter detection problems: (a) hidden terminal problem, (b) shadowing uncertainty [6].*

2.2.1.1 Matched filter

A block diagram of a matched filter is shown in Figure 2.6(a). The matched filter method is an optimal approach for spectrum sensing in the sense that it maximises the signal-to-noise ratio (SNR) in the presence of additive noise [23, 41]. Another advantage of the matched filter method is that it requires less observation time since the high processing gain can be achieved by coherent detection. For example, to meet a given probability of detection, only $\mathcal{O}(1/\text{SNR})$ samples are required [10]. This advantage is achieved by correlating the received signal with a template to detect the presence of a known signal in the received signal. However, it relies on prior knowledge of the PU, such as modulation type, and packet format, and requires the cognitive radio to be equipped with carrier synchronisation and timing devices. With more types of PUs, the implementation complexity grows making the matched filter impractical.

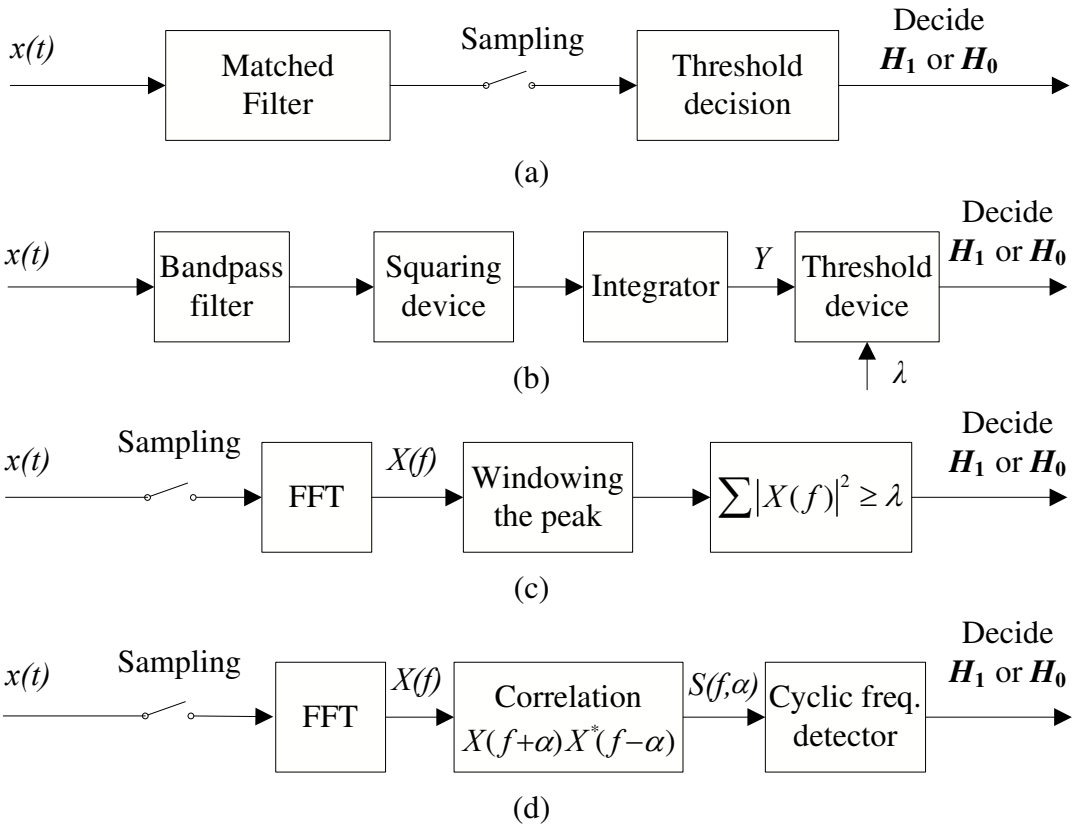

Figure 2.6: *Block diagrams for traditional spectrum sensing algorithms: (a) matched filter, (b) time domain energy detection, (c) frequency domain energy detection, and (d) cyclostationary detection [8].*

2.2.1.2 Energy detection

If the information about the PU is unknown in the cognitive radio, a commonly used method for detecting the PUs is energy detection (also known as radiometry) [50]. Energy detection is a non-coherent detection method that avoids the need for complicated receivers required by a matched filter. An energy detector can be implemented in both the time and the frequency domain. For time domain energy detection as shown in Figure 2.6(b), a bandpass filter (BPF) is applied to select a centre frequency and bandwidth of interest [7,20,21]. Then the energy of the received signal is measured by a magnitude squaring device, with an integrator to control the observation time. Finally, the energy of the received signal will be compared with a predetermined threshold to decide whether the PU is present or not. However, to sense a wide spectrum span, sweeping the BPF will result in a long measurement time. As shown in Figure 2.6(c), in the frequency domain, the energy detector can be implemented similarly to a spectrum analyser with a fast Fourier transform (FFT) [7,10,21]. Specifically, the received signal is sampled at or

Sensing algorithm	Advantages	Disadvantages
Matched filter	Optimal performance Low computational cost	Require prior information of the primary user
Energy detection	Do not require prior information Low computational cost	Poor performance for low SNR Cannot differentiate users
Cyclostationary	Valid in slow SNR region Robust against interference	Require partial prior information High computational cost

Table 2.1: *Summary of advantages and disadvantages of narrowband spectrum sensing algorithms [8,9].*

above the Nyquist rate over a time window. Then the PSD is computed using an FFT. The FFT is employed to analyze a wide frequency span in a short observation time, rather than sweeping the BPF in Figure 2.6(b). Finally, the PSD will be compared with a threshold, λ, to decide whether the corresponding frequency is occupied or not.

The advantages of energy detection are that prior knowledge of the PUs is not required, and both the implementation and the computational complexity are generally low. In addition, a short observation time is required, for example, $\mathcal{O}(1/\text{SNR}^2)$ samples are required to satisfy a given probability of detection [10]. Although energy detection has a low implementation complexity, it has some drawbacks. A major drawback is that it has poor detection performance under low SNR scenarios as it is a non-coherent detection scheme. Another drawback is that it cannot differentiate between the signal from a PU and the interference from other cognitive radios, thus, it cannot take advantage of adaptive signal processing, such as interference cancellation. Furthermore, noise level uncertainty can lead to further performance loss. These disadvantages can be overcome by using two-stage spectrum sensing technique, i.e. coarse spectrum sensing and fine spectrum sensing [21]. Coarse spectrum sensing can be implemented by energy detection or wideband spectrum analysing techniques. The aim of coarse spectrum sensing is to quickly scan the wideband spectrum and identify some possible spectrum holes in a short observation time. By contrast, fine spectrum sensing further investigates and analyses these suspected frequencies. More sophisticated detection techniques can be used at this stage, such as cyclostationary detection described below.

2.2.1.3 Cyclostationary detection

A block diagram of cyclostationary detection is shown in Figure 2.6(d). Cyclostationary detection is a method for detecting the PUs by exploiting the cyclostationary features in the modulated signals [23, 37]. In most cases, the received signals in cognitive radios are modulated signals, which in general exhibit built-in-periodicity within the training sequence or cyclic prefixes. This periodicity is generated by the primary transmitter so that the primary receiver can use it for parameter estimation, such as channel estimation, and pulse timing [21,23]. The cyclic correlation function, also called cyclic spectrum function (CSF), is used for detecting signals with a particular modulation type in the presence of noise. This is because noise is usually wide-sense stationary (WSS) without correlation, by contrast, modulated signals are cyclostationary with spectral correlation. Furthermore, since different modulated signals will exhibit different characteristics, cyclostationary detection can be used for distinguishing between different types of transmitted signals, noise, and interference in low SNR environments. One of the drawbacks of cyclostationary detection is that it still requires partial information of the PU. Another drawback is that the computational cost is high as the CSF is a two-dimensional function dependent on frequency and cyclic frequency [8, 9].

2.2.2 Cooperative Spectrum Sensing

Cooperative spectrum sensing can not only decrease the probabilities of false alarm and missed detection, but can also mitigate the hidden terminal problem. Thus, multiple cognitive radios are often required to collaborate for spectrum sensing. In a centralized sensing setup as shown in Figure 2.7, each cognitive radio observes the RF spectrum individually, and forwards its measured/processed data, or decisions to a fusion center (FC) via a common control channel. The common control channel is responsible for transferring sensing and control information between all cognitive radios and the FC. The FC then fuses all sensing data or decisions, identifies the available spectrum, and broadcasts the spectrum information to all nearby cognitive radios via a control channel. Generally, cooperative spectrum sensing algorithms can be categorised into data fusion and decision fusion, depending upon which type of sensing data is transmitted to the FC [7, 8, 37].

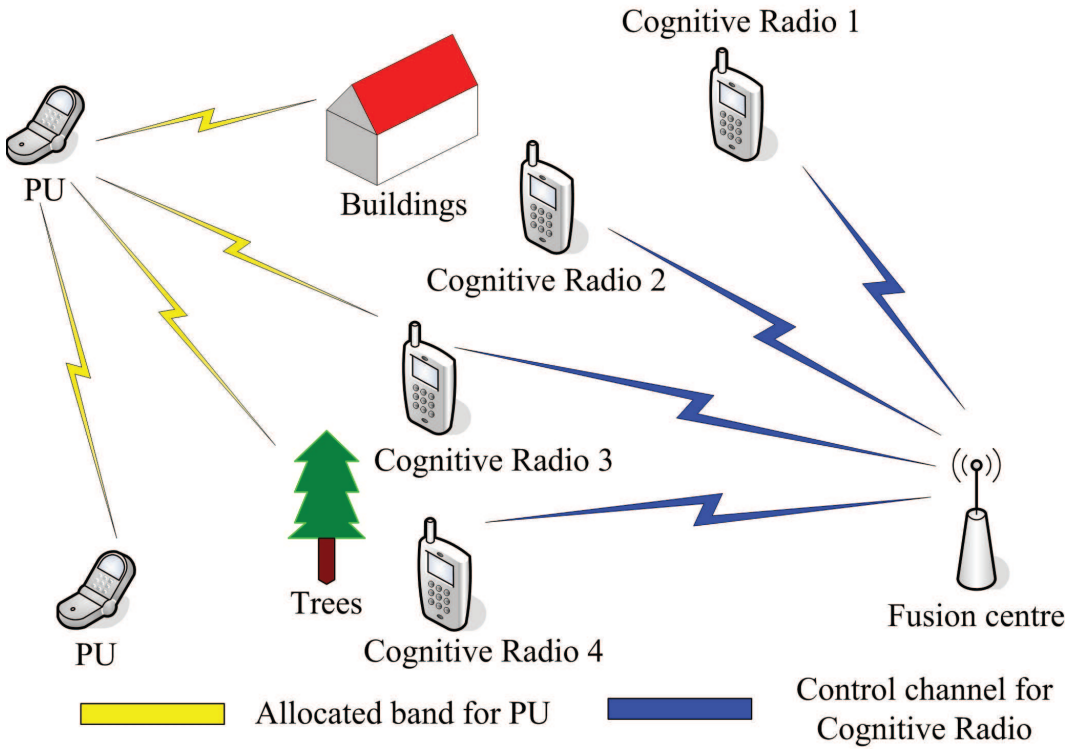

Figure 2.7: *Schematic illustration of cooperative spectrum sensing scheme that each cognitive radio transmits its individual observation or decision via control channels to a fusion center, which makes final decision on the spectral occupancy status.*

2.2.2.1 Data Fusion

If raw data from all cognitive radios are collected at the FC, some advanced signal processing techniques can be applied. This scheme is thus similar to a multi-antenna sensing case. Furthermore, if the channel state information (CSI) between the PUs and the cognitive radios are perfectly known, the optimal combining strategy, MRC [51], can be used for achieving the highest output SNR. In the case where partial CSI is available, other combining techniques can be employed, such as SC [51]. However, the communication burden of transmitting raw data is significant. Thus, it is preferred to send processed or compressed data to the FC for the purpose of saving transmission resources.

When an energy detector is deployed in each cognitive radio, energy vectors from all cognitive radios can be sent to the FC. In such a scenario, the FC can employ SLC, or SLS [52, 53]. The transmission bandwidth for sharing the data is half of the raw data case as the energy vectors are real, instead of complex. In addition, CSI is not required when using SLC or SLS. Alternatively, the measurements can be further compressed via source coding techniques, and then sent to the

FC. It is evident that sending compressed data will save transmission bandwidth and energy, but, it requires more computational resources in both the cognitive radios and the FC.

2.2.2.2 Decision Fusion

In a decision fusion scheme, every cognitive radio performs local spectrum measurements independently, and makes a binary decision on whether a PU is present or not. All decisions from the cognitive radios are then forwarded to the FC, where the decisions are fused and a final decision is made on the spectrum occupancy status. In general, three decision fusion rules can be adopted as below [37]:

- *Logical OR rule*: In this rule, the FC gives decision "1" (PU present) if any one of decisions from the cognitive radios is "1". Thus, using this rule, the probability of false alarm (when PU is absent, cognitive radios think that PU is using that band) will increase as shown in (2.1). Meanwhile, the probability of missed detection (when PU is present, cognitive radios sense that PU is not using this band) is reduced. Since cognitive radio occupying a frequency band used by the PU may interfere with the PU, the risk of cognitive radios causing interference to the PU is minimised using the logical OR rule. If these decisions from the $v(v \geq 1)$ cognitive radios are independent, the probabilities of false alarm and detection can be given by:

$$P_f = 1 - \prod_{i=1}^{v}(1 - P_f^i), \tag{2.1}$$

$$P_d = 1 - \prod_{i=1}^{v}(1 - P_d^i), \tag{2.2}$$

  where P_f^i and P_d^i denote the probabilities of false alarm and detection in the cognitive radio node i, respectively.

- *Logical AND rule*: In this rule, the FC decides "1" if and only if all decisions from the cognitive radios are "1". Hence, using this rule, the probability of false alarm is minimised, but the risk of causing interference will increase. Similarly, if these decisions from the v cognitive radios are independent, the probabilities of false alarm and detection can be given by:

$$P_f = \prod_{i=1}^{v} P_f^i, \tag{2.3}$$

$$P_d = \prod_{i=1}^{v} P_d^i. \tag{2.4}$$

- *c out of v voting*: The FC decides "1" if and only if c or more than c cognitive radios decide "1", where $c \in [1, v]$. If these decisions from the v cognitive radios are independent, the probabilities of false alarm and detection can be given as:

$$P_f = \sum_{i=0}^{v-c} \binom{v}{c+i} (1 - P_f^i)^{v-c-i} (P_f^i)^{c+i}, \tag{2.5}$$

$$P_d = \sum_{i=0}^{v-c} \binom{v}{c+i} (1 - P_d^i)^{v-c-i} (P_d^i)^{c+i}. \tag{2.6}$$

  Obviously, this fusion rule includes the logical OR rule ($c = 1$), and the logical AND rule ($c = v$) as special cases.

2.2.2.3 Data Fusion Versus Decision Fusion

For both fusion schemes, the measured data or decisions need to be transmitted through a control channel. In practice, the bandwidth of the control channel alters in cognitive radio networks, and the information exchanges may be unreliable. If the control channel is rented, the data transmission could be expensive. In comparison to data fusion, the one-bit decision fusion scheme needs a very low bandwidth for sharing the spectrum sensing data, as well as saving transmission energy. However, most decision fusion rules assume that the decisions from all cognitive radios are independent, which in practice may not occur. Any decision fusion scheme is open to similar abuse unless proper security is incorporated.

2.2.3 Wideband Spectrum Sensing

From the discussion in Section 2.1, the average spectrum occupancy is around 5%. Under such a circumstance, the cognitive radio can easily find spectrum holes by using a tunable narrowband bandpass filter (TNBF) [54] to search one narrowband portion of the spectrum at a given time. Traditional spectrum sensing algorithms can then be used for searching spectrum holes. Due to the explosive development of wireless products, the average spectrum occupancy will increase. A wideband spectrum sensing structure should be adopted to search multiple bands at a time [24]. In practice, wideband spectrum sensing systems are difficult to design,

Sensing algorithm	Advantages	Disadvantages
Wavelet detection	Flexibility in adapting to dynamic spectrum	Requires high sampling rate ADC High energy consumption
Filter bank detection	Low sampling rate High spectral dynamic range	Large implementation complexity Not flexible as filters are preset
Multicoset sampling	Low sampling rate Less measurements	Requires accurate time offsets Requires too many sampling channels
CS based detection	Low sampling rate Less processed data	High implementation complexity Matrix storage & transmission
Multirate sampling	Low sampling rate Less sampling channels	Stringent requirements on devices Non-aliased in at least one channel

Table 2.2: *Summary of advantages and disadvantages of wideband spectrum sensing algorithms [7, 27–29, 33–35].*

due to either high implementation complexity or high financial/energy costs [8]. The literature of wideband spectrum sensing is still in its early stages; five types of models are commonly discussed. They are: 1) wavelet detection [24], 2) filter bank detection [25, 26], 3) multicoset sampling based detection [30–32], 4) CS based methods [27, 28], and 5) multirate sampling based detection [33, 34, 55]. The advantages and disadvantages of these algorithms are summarised in Table 2.2.

2.2.3.1 Wavelet Detection

In [24], Tian and Giannakis proposed a wavelet-based wideband sensing approach. It provides an advantage of flexibility in adapting to a dynamic wideband spectrum. In their approach, the PSD of the wideband spectrum is modelled as a train of consecutive frequency subbands, where the PSD is smooth within each subband but exhibits discontinuities and irregularities on the border of two neighboring subbands as shown in Figure 2.8. The architecture of the wavelet transform based wideband spectrum sensing is illustrated in Figure 2.9. The wavelet transform of the wideband PSD is used to locate the singularities of the PSD.

Let $\varphi(f)$ be a wavelet smoothing function, the dilation of $\varphi(f)$ is given by,

$$\varphi_d(f) = \frac{1}{d}\varphi\left(\frac{f}{d}\right), \tag{2.7}$$

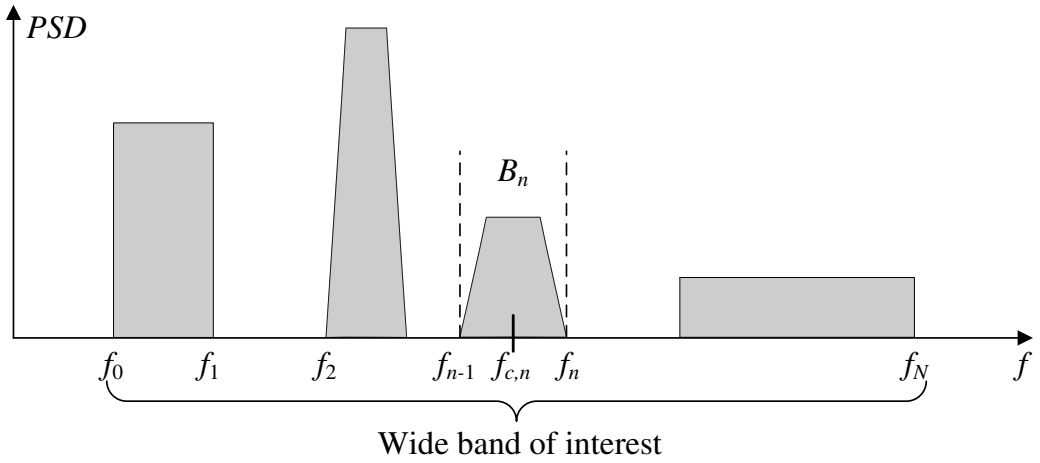

Figure 2.8: *Demonstration of the wideband spectrum of interest. The PSD is smooth within each subband, and exhibits discontinuities and irregularities with the adjacent subbands [24, 27].*

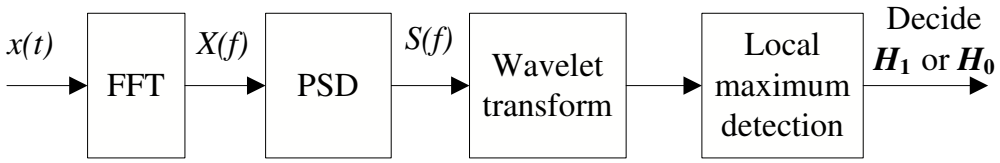

Figure 2.9: *Block diagram of the wavelet transform based wideband spectrum sensing technique [8].*

where d is a dyadic scale that can take values that are powers of 2, i.e. $d = 2^j$. The continuous wavelet transform (CWT) of the PSD is given by [24],

$$\text{CWT}\{S(f)\} = S(f) * \varphi_d(f), \tag{2.8}$$

where "$*$" denotes the convolution and $S(f)$ is the PSD of the received signal.

Then the first and second derivative of the CWT$\{S(f)\}$ are used to locate the irregularities and discontinuities in the wideband PSD. Specifically, the boundaries of each subbands are located by using the local maxima of the first derivative of CWT$\{S(f)\}$, and locations of the subbands are finally tracked by finding zero crossings in the second derivative of CWT$\{S(f)\}$.

By controlling the wavelet smoothing function, the wavelet transform based wideband spectrum sensing approach has flexibility in adapting to the dynamic spectrum. However, characterising the wideband spectrum will require a high sampling rate ADC (a signal with a bandwidth of

W must be sampled at or above $2W$ rate), due to the Nyquist sampling theorem [56], and the energy cost of that ADC will be prohibitive.

2.2.3.2 Filter Bank Detection

In [25], Farhang-Boroujeny presented a filter bank method for wideband spectrum sensing in cognitive radio systems. In his approach, multicarrier communications are assumed to be used in cognitive radio networks. It is assumed that a pair of matched root-Nyquist filters are employed in the primary transmitter and the cognitive radio receiver, respectively. The wideband spectrum is considered as the output of a bank of prototype filters (with different shifted central frequencies). As shown in Figure 2.10, the baseband spectrum can be directly estimated by using a prototype filter, and other bands can be obtained through modulating the prototype filter. In each subcarrier, the corresponding portion of the spectrum for the wideband signal is downconverted to baseband, then lowpass filtered as depicted in Figure 2.11. The PSD of the output signal, $S_{y_i y_i}(f)$, can be written as [25],

$$S_{y_i y_i}(f) = S_{xx}(f + f_i)|H(e^{2\pi jf})|^2 \approx S_{xx}(f_i)|H(e^{2\pi jf})|^2, \tag{2.9}$$

where $S_{xx}(f_i)$ denotes the PSD of the received signal $x(t)$ in the subband i, and $H(z)$ is assumed to be narrowband and is designed as a root Nyquist (N) filter.

The approximate result in (2.9) can be rewritten in terms of the z-transform variable z as [25],

$$\Psi_{y_i y_i}(z) = S_{xx}(f_i)H(z)H(z^{-1}) = S_{xx}(f_i)G_N(z), \tag{2.10}$$

where $G_N(z) = H(z)H(z^{-1})$ is called the Nyquist (N) filter. In the time domain, after normalising, $G_N(z)$ satisfies [25],

$$g_N(n) = \begin{cases} 1, & n = 0 \\ 0, & n = mN, m \neq 0, \end{cases} \tag{2.11}$$

where N is the maximum number of subcarriers in the filter bank.

Let $\psi_{y_i y_i}(k)$ represent the correlation coefficients of $y_i(n)$ when performing an inverse z-transform on $\Psi_{y_i y_i}(z)$, the correlation matrix of the measured vectors, $\mathbf{R}_{y_i y_i}$, can be given

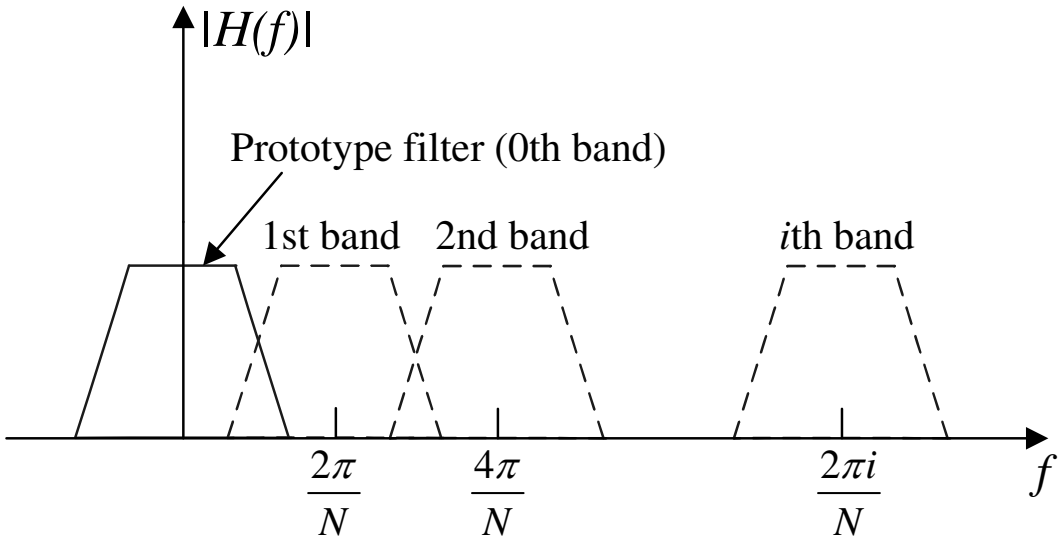

Figure 2.10: *The graphic illustration of using filter bank on the wideband spectrum [25].*

by [25],

$$\mathbf{R}_{y_i y_i} = S_{xx}(f_i)\mathbf{A}, \tag{2.12}$$

where the matrix $\mathbf{A}$ is a Toeplitz matrix, and each element of $\mathbf{A}$ is from the Nyquist (N) sequence, i.e. $g_N(n)$.

It can be shown that $S_{xx}(f_i)$ follows central or non-central chi-square distribution. Then the degree of freedom estimation is critical for the hypothesis test on $S_{xx}(f_i)$. The eigenvalue decomposition is performed on matrix $\mathbf{A}$, and the resultant eigenvalues, $\lambda_0, \lambda_1, \cdots$, are used to measure the degree of freedom in $S_{xx}(f_i)$. Then the estimated $S_{xx}(f_i)$ can be obtained by using observation vectors, eigenvalues, and degree of freedom as shown in equation (34) in [25].

As a result, compared with multitaper method [18], the filter bank method can obtain a lower variance when the PSD is low, because of its better response of the prototype filter. Regardless of the high spectral dynamic range of the filter bank approach, its computational speed is slower than that of the multitaper method [45]. Besides, the implementation of the filter bank approach requires a large number of RF components for sensing a wideband spectrum [27, 28]. For example, in order to sense wideband spectrum with a bandwidth of 10 GHz in a time, it needs at least 200 RF front-end components when the lowpass filter has the bandwidth of 100 MHz (100 ADCs and 100 filters as 10 GHz$/0.1$ GHz $= 100$). The range of filters, and the number of the narrow bands are always preset, thus, the filter bank model is not flexible. Furthermore, the filters at the primary transmitter and the cognitive radio are assumed to be a pair of matched

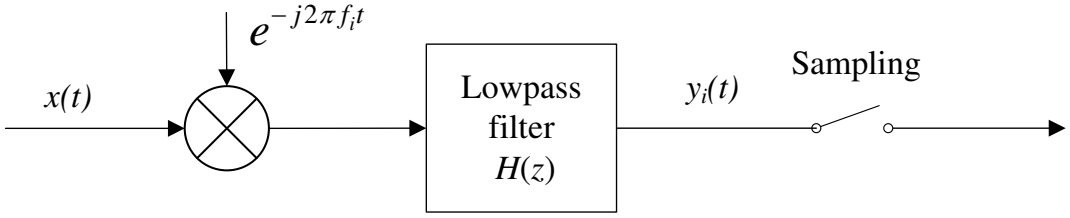

Figure 2.11: *The demodulation of received wideband signal in the i-th subcarrier [25].*

root-Nyquist filters [25], however, it is not practical to obtain the filter information of the PUs in the cognitive radio networks.

2.2.3.3 Multicoset Sampling based Detection

A potential method for implementing wideband spectrum sensing using a sub-Nyquist sampling rate is multicoset sampling [30–32]. Multicoset sampling is a selection of some samples from a uniform grid, which can be obtained when uniformly sampling signal, $x(t)$, at a rate of f_s (greater than the Nyquist rate). The uniform grid is then divided into blocks of L consecutive samples, and in each block $v(v < L)$ samples are retained while the rest of samples, i.e. $L - v$ samples, are skipped. A constant set C that describes the indexes of these v samples in each block is called a sampling pattern as [32],

$$C = \{t^i\}_{i=1}^{v}, \quad 0 \leq t^1 < t^2 < \cdots < t^v \leq L - 1. \tag{2.13}$$

The multicoset sampling can be implemented by using v sampling channels with sampling rate of $\frac{f_s}{L}$, where the i-th sampling channel is offset by $\frac{t^i}{f_s}$ from the origin as below [32],

$$x^i[n] = \begin{cases} x(\frac{n}{f_s}), & n = mL + t^i,\ m \in \mathbb{Z} \\ 0, & \text{otherwise}. \end{cases} \tag{2.14}$$

The discrete-time Fourier transform (DTFT) of the samples can be linked to the unknown Fourier transform of signal $x(t)$ by,

$$\overrightarrow{Y}(f) = \mathbf{\Phi}\overrightarrow{X}(f), \tag{2.15}$$

where $\overrightarrow{Y}(f)$ denotes a vector of DTFT of these measurements from v channels, $\overrightarrow{X}(f)$ is a vector of the Fourier transform of $x(t)$, and $\mathbf{\Phi}$ is the measurement matrix whose elements are

determined by the sampling pattern C. The problem of wideband spectrum sensing is thus equivalent to recovering $\overrightarrow{X}(f)$ from $\overrightarrow{Y}(f)$. In order to get a unique solution from (2.15), every set of v columns of $\mathbf{\Phi}$ should be linearly independent. However, searching for this sampling pattern is a combinatorial problem. In [57,58], some sampling patterns are proved to be valid for reconstruction. The advantage of multicoset sampling is that the sampling rate in each channel is L times lower than the Nyquist rate. Moreover, the number of measurements is $\frac{v}{L}$ lower than the Nyquist sampling case. One drawback of the multicoset sampling is that accurate time offsets between sampling channels are required to satisfy a specific sampling pattern. Another one is that the number of sampling channels should be sufficiently high [32].

2.2.3.4 Compressive Sensing based Detection

Tian and Giannakis were the first to exploit the sparsity of radio signals by introducing CS theory [59] to realise wideband spectrum sensing [27, 60, 61]. The technique takes advantage of using fewer samples closer to the information rate, rather than the inverse of the bandwidth, to perform wideband spectrum sensing [27]. After reconstruction of the wideband spectrum, wavelet-based edge detection was used to detect the wideband spectrum as shown in Figure 2.12.

Let $x(t)$ represent a wideband signal received at the cognitive radio. If $x(t)$ is sampled at the Nyquist sampling rate, the sequence vector, i.e. $\vec{x}$ ($\vec{x} \in \mathbb{C}^N$), will be obtained. The Fourier transform of the sequence, $\overrightarrow{X} = \mathbf{F}\vec{x}$, will therefore be alias-free, where $\mathbf{F}$ denotes the Fourier matrix. When the spectrum, $\overrightarrow{X}$, is k-sparse ($k \ll N$), which means k out of N values in $\overrightarrow{X}$ are not neglectable, $x(t)$ can be sampled at a sub-Nyquist rate while its spectrum can be reconstructed with a high probability. The sub-sampled/compressed signal, $\vec{y} \in \mathbb{C}^M$ ($k < M \ll N$), is linked to the Nyquist sequence $\vec{x}$ by [27],

$$\vec{y} = \mathbf{\Phi}\vec{x}, \tag{2.16}$$

where $\mathbf{\Phi} \in \mathbb{C}^{M \times N}$ is the measurement matrix, which is a selection matrix that randomly chooses M columns of the size-N identity matrix. Namely, $N - M$ samples out of N samples are skipped. The relationship between the spectrum $\overrightarrow{X}$ and the compressed sequence $\vec{y}$ is given by [27],

$$\vec{y} = \mathbf{\Phi}\mathbf{F}^{-1}\overrightarrow{X}, \tag{2.17}$$

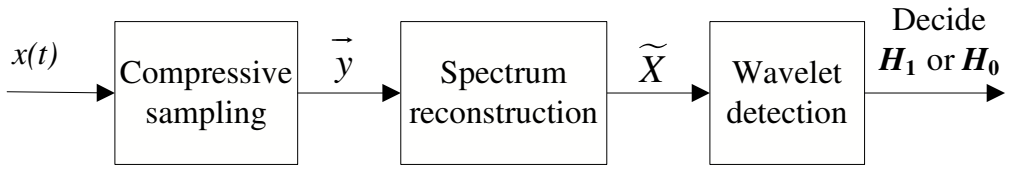

Figure 2.12: *Block diagram of the compressive sensing based detection.*

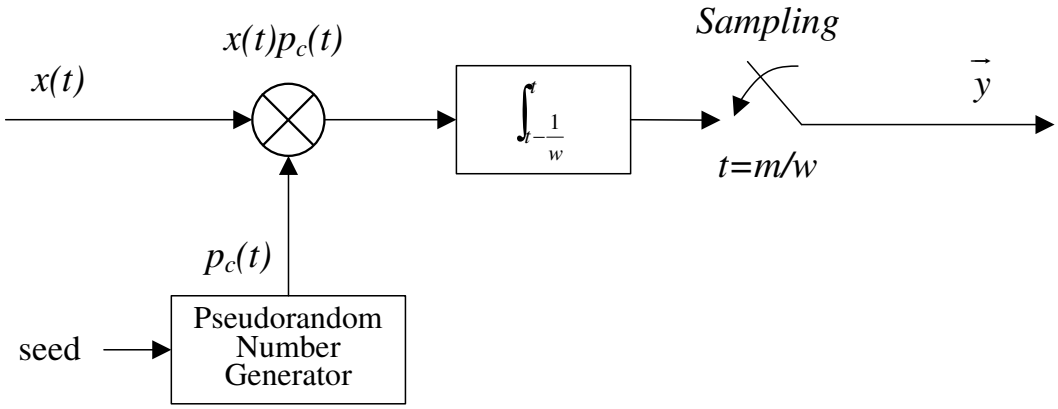

Figure 2.13: *Block diagram for the analog-to-information converter [69]. The received signal, $x(t)$, is randomly demodulated by a pseudorandom chipping sequence, integrated by an accumulator, and sampled at a sub-Nyquist rate.*

where $\mathbf{F^{-1}}$ denotes the inverse Fourier matrix.

Approximating $\overrightarrow{X}$ from $\vec{y}$ in (2.17) is a linear inverse problem and is NP-hard. The basis pursuit (BP) [62] algorithm can be used to solve $\overrightarrow{X}$ by linear programming [27]:

$$\widetilde{X} = \arg\ \min \|\overrightarrow{X}\|_1, \quad \text{s. t.} \quad \vec{y} = \mathbf{\Phi F^{-1}} \overrightarrow{X}. \tag{2.18}$$

After reconstructing the full spectrum $\overrightarrow{X}$, the PSD is calculated using $\widetilde{X}$. Then the wavelet detection approach in Section 2.2.3.1 can be used to analyse the edges in the PSD. Although less measurements are used for characterising the wideband spectrum, the requirement of high sampling rate on ADC is not relaxed. By contrast, in [28, 29, 63], Polo *et al.* suggested using an analog-to-information converter (AIC) [64–68] model (also known as random demodulator, [69]) for compressing the wideband signal in the analog domain. The block diagram of AIC is given in Figure 2.13.

A pseudorandom number generator is used to produce a discrete-time sequence $\varepsilon_0, \varepsilon_1, \cdots,$

called a chipping sequence, the number of which takes values of ± 1 with equal probability. The waveform should randomly alternate at or above the Nyquist rate, i.e. $\varpi \geq 2W$, where W is the bandwidth of signal. The output of the pseudorandom number generator, i.e. $p_c(t)$, is employed to demodulate a continuous-time input $x(t)$ by a mixer. Then an accumulator sums the demodulated signal for $1/w$ seconds, and the filtered signal is sampled at a sub-Nyquist rate of w. This sampling approach is called integrate-and-dump sampling since the accumulator is reset after each sample is taken. The samples acquired by the AIC, $\vec{y} \in \mathbb{C}^w$, can be related to the received signal, $\vec{x} \in \mathbb{C}^{\varpi}$, by,

$$\vec{y} = \mathbf{\Phi}\vec{x}, \tag{2.19}$$

where $\mathbf{\Phi} \in \mathbb{C}^{w \times \varpi}$ is the measurement matrix describing the overall action of the AIC system on the input signal $\vec{x}$. The signal $\vec{x}$ can be identified by solving the convex optimization problem,

$$\widetilde{x} = \text{arg } \min \|\vec{x}\|_1, \quad \text{s. t.} \quad \vec{y} = \mathbf{\Phi}\vec{x}, \tag{2.20}$$

by BP or other greedy pursuit algorithms. The PSD of the wideband spectrum can be estimated using the recovered signal $\widetilde{x}$, followed by a hypothesis test on the PSD. Alternatively, the PSD can be directly recovered from the measurements using CS algorithms [28]. Although the AIC bypasses the requirement for a high sampling rate ADC, it leads to a high computational complexity as the huge-scale of the measurement matrix [35]. Furthermore, it has been identified that the AIC model can easily be influenced by design imperfections or model mismatches [35].

In [22, 35], Mishali and Eldar proposed a parallel implementation of the AIC model, called modulated wideband converter (MWC), as shown in Figure 2.14. The key difference is that in each channel the accumulator for integrate-and-dump sampling is replaced by a general low-pass filter. One of the benefits of introducing parallel structure is that the dimension of the measurement matrix is reduced making the reconstruction easier. Another benefit is that it provides robustness to noise and model mismatch. On the other hand, the implementation complexity increases as multiple sampling channels are involved. An implementation issue of using MWC is that the storage and transmission of the measurement matrix must be considered when it is used in a distributed cognitive radio network under a data fusion collaborative scheme.

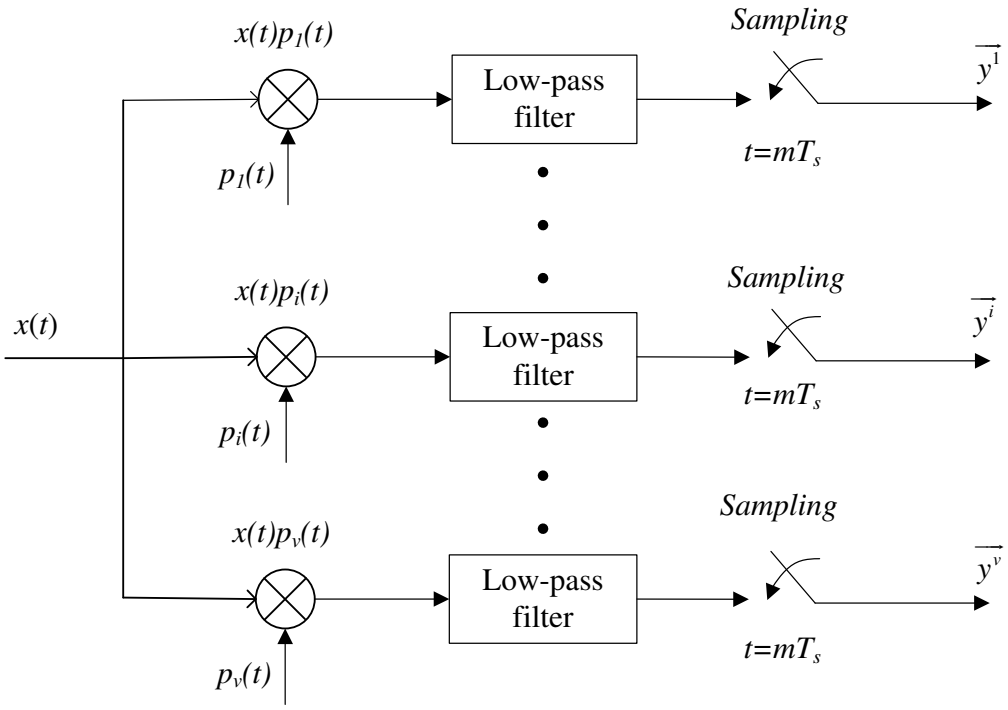

Figure 2.14: *Block diagram for the modulated wideband converter [35]. In each channel, the received signal is demodulated by a pseudorandom sequence, filtered by a low-pass filter, and sampled at a sub-Nyquist rate* $\frac{1}{T_s}$.

2.2.3.5 Multirate Sampling

An alternative model for compressing the wideband spectrum in the analog domain is a multi-rate sampling system as shown in Figure 2.15. Asynchronous multirate sampling (MRS) and synchronous multirate sampling (SMRS) were used for reconstructing sparse multiband signals in [33] and [34], respectively. In addition, MRS has been successfully implemented in experiments using an electro-optical system with three sampling channels as described in [55]. Both systems employ three optical pulsed sources that operate at different rates and at different wavelengths. The received signal is modulated with optical pulses, which provided by an optical pulse generator (OPG), in each channel. In order to reconstruct a wideband signal with an 18 GHz bandwidth, the modulated pulses are amplified, and sampled by an ADC at a rate of 4 GHz in each channel.

In [33], the sampling channels of the MRS can be implemented separately without synchronisation. However, reconstruction of the spectrum requires that each frequency of the signal must be non-aliased in at least one of the sampling channels. In [34] SMRS reconstructs the spectrum from linear equations, which relate the Fourier transform of the signal to the Fourier

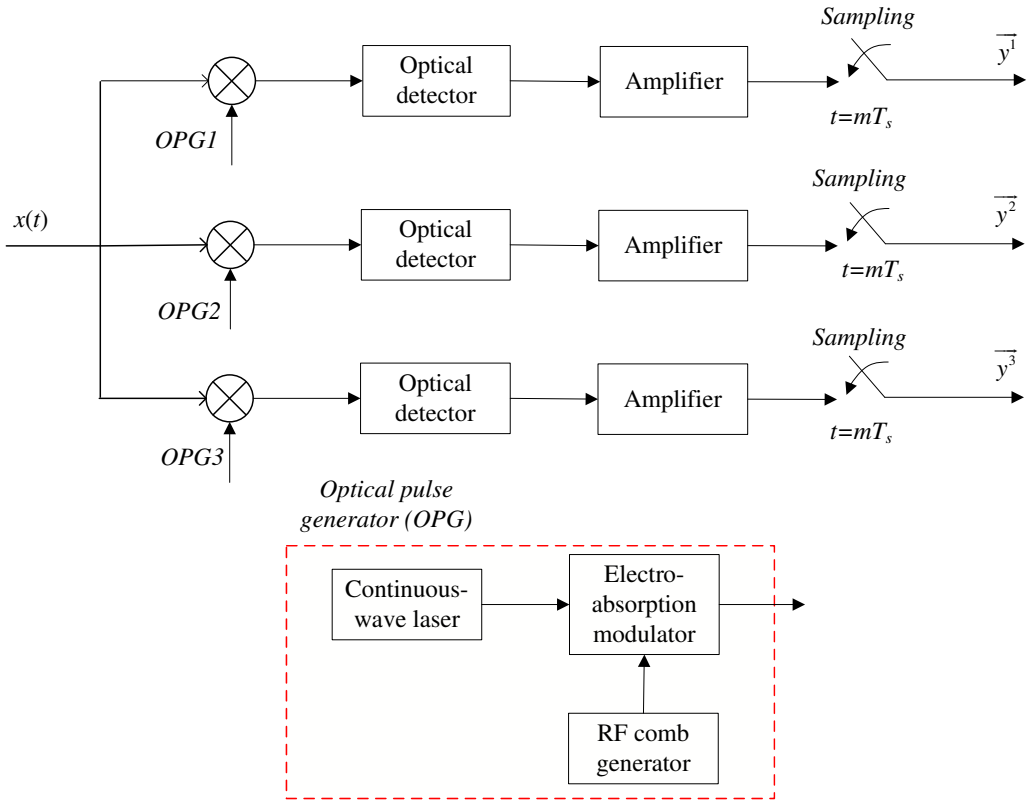

Figure 2.15: *Multirate sampling system implemented by electro-optical devices [55]. In each channel, the received signal is modulated by a train of short optical pulses. The modulated signal is then detected by an optical detector, amplified, and sampled by a low-rate ADC.*

transform of its samples. Using CS theory, sufficient conditions for perfectly reconstructing the spectrum are obtained; $v \geq 2k$ (the Fourier transform of the signal is k-sparse) sampling channels are required. In order to reconstruct the spectrum using MRS with fewer sampling channels, the spectrum to be recovered should possess certain properties, e.g., minimal bands, and uniqueness. Nonetheless, the spectral components from PUs may not possess these properties. Obviously, even though the multirate sampling system has broad application, there is a long way to go to implement it in a cognitive radio network because of its stringent requirements on both optical devices and the number of sampling channels.

Chapter 3
Narrowband Collaborative Spectrum Sensing

Energy detection is commonly used for spectrum sensing in a cognitive radio network, because it has a low implementation complexity and does not require CSI. In practical applications, fading occurs because of multipath propagation and shadowing. Thus it is important for cognitive radios to dynamically balance the probability of missed detection against the probability of false alarm. A computationally inexpensive means of calculating these is advantageous due to restricted computational resources of cognitive radios. On the other hand, to combat the effect of fading, multiple cognitive radios are often designed to collaborate in spectrum sensing. Considering the amount of CSI available at the cognitive radios, there are several different collaborative strategies, for example, MRC, SC, SLC, and SLS. It is noteworthy that different diversity reception schemes will result in distinct performance. Hence, a comparison of these diversity reception schemes is of great significance.

This chapter analyses the detection performance of energy detection over fading channels, and the contributions of this chapter are summarised as follow:

- For the Nakagami-m fading channel, a rapidly converging representation for the average probability of detection is obtained for any value of $m \in [1/2, \infty)$. This saves computational resources in cognitive radios.
- For the Rician fading channel, an easily computed expression for the average probability of detection is derived, which is applicable for any time bandwidth product of the test statistic.
- For a slow log-normal fading channel, an approximation of the average probability of detection is given, by using the Wald distribution to replace the log-normal distribution.
- The detection performance of energy detection using different collaborative strategies, i.e. MRC, SC, SLC, and SLS, are analysed.

The remainder of this chapter is organized as follows. Section 3.1 describes the energy detection method. When the signals from PUs experience fading, Section 3.2 derives some easily computed expressions for the average probabilities of false alarm and detection. For cognitive radios that are designed to collaborate using MRC, SC, SLC, or SLS over i.i.d. Nakagami-m fading channels, the average probabilities of false alarm and detection are derived in Section 3.3. Then simulation results are given in Section 3.4, followed by conclusions in Section 3.5.

3.1 System Description

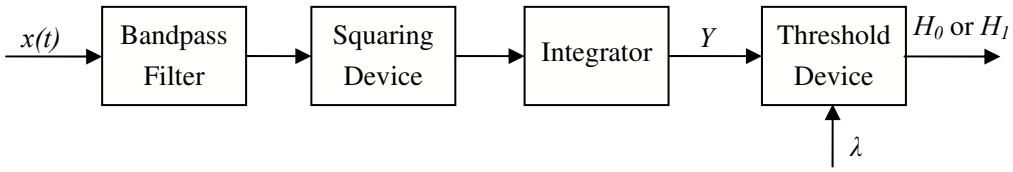

Figure 3.1: *Block diagram of the energy detector.*

A block diagram of an energy detector is shown in Figure 3.1. The received signal, $x(t)$, is filtered by a BPF, followed by a magnitude squaring device for measuring received energy, and an integrator that controls the observation interval, T. In order to decide whether the signal is present or not, the output of the integrator, Y, will act as a test statistic, and will be compared with a predetermined threshold, λ. The binary signal detection problem can be formulated as hypothesis test with H_0 (signal not present) or H_1 (signal present),

$$\begin{aligned} H_0: \quad x(t) &= z(t), \\ H_1: \quad x(t) &= h(t)\ s(t) + z(t), \end{aligned} \tag{3.1}$$

where $h(t)$ denotes the complex channel gain between the PU and the cognitive radio, $s(t)$ denotes the bandlimited signal coming from the PUs of unknown modulation format, and $z(t)$ is AWGN.

Following [43], the test statistic, Y, can be described as,

$$Y \sim \begin{cases} \chi^2_{2u}, & H_0 \\ \chi^2_{2u}(2\gamma), & H_1 \end{cases} \tag{3.2}$$

where $Y \sim \chi$ means Y follows the distribution of χ, γ denotes the signal-to-noise ratio (SNR)

at the cognitive radio, and χ^2_{2u} and $\chi^2_{2u}(2\gamma)$ denote the central and non-central chi-square distributions, respectively. Both distributions have the same degree of freedom (DoF), $2u$ (u is the time bandwidth product), and the latter one has a non-central parameter 2γ. The time bandwidth product, i.e. $u = TW$, denotes that signals of duration T have most of their energy within the frequency band $[-W/2, W/2]$. The value of u can be either integer or non-integer. The probability density function (PDF) of Y is given as [52],

$$f_Y(y) = \begin{cases} \frac{1}{2^u \cdot \Gamma(u)} \cdot y^{u-1} \cdot e^{-\frac{y}{2}}, & H_0 \\ \frac{1}{2} \cdot (\frac{y}{2\gamma})^{\frac{u-1}{2}} \cdot e^{-\frac{2\gamma+y}{2}} \cdot I_{u-1}(\sqrt{2\gamma y}), & H_1 \end{cases} \tag{3.3}$$

where $\Gamma(a)$ is the gamma function and $I_v(a)$ is the v-th order modified Bessel function of the first kind.

For a non-fading AWGN channel, the probabilities of false alarm and detection are given in [52] as below,

$$P_f = P_r(Y > \lambda | H_0) = \frac{\Gamma(u, \lambda/2)}{\Gamma(u)}, \tag{3.4}$$

$$P_d = P_r\left(Y > \lambda | H_1\right) = Q_u(\sqrt{2\gamma}, \sqrt{\lambda}), \tag{3.5}$$

where $\Gamma(a, x)$ denotes the upper incomplete gamma function given by $\Gamma(a, x) = \int_x^\infty t^{a-1} e^{-t} dt$, and $Q_u(a, x)$ denotes the generalised Marum Q-function given by,

$$Q_u(a, x) = \frac{1}{a^{u-1}} \int_x^\infty t^u e^{-\frac{a^2+t^2}{2}} I_{u-1}(at) dt. \tag{3.6}$$

The computation of the integral representation of P_d in (3.5) has a high complexity. Thus, in the first half of the chapter, a series form representation for the generalised Marcum Q-function in (27) of [70] is used, and P_d can be represented as,

$$P_d(\gamma, \lambda) = e^{-\frac{\lambda}{2}} \sum_{i=0}^{u-1} \frac{\left(\frac{\lambda}{2}\right)^i}{i!} + e^{-\frac{\lambda}{2}} \sum_{n=u}^{\infty} \frac{\left(\frac{\lambda}{2}\right)^n}{n!} \left(1 - e^{-\gamma} \sum_{k=0}^{n-u} \frac{\gamma^k}{k!}\right). \tag{3.7}$$

3.2 Spectrum Sensing over A Single Fading Channel

In a fading channel, the average probability of false alarm, $\overline{P_f}$, will not change [52]. In contrast, when the channel gain, $h(t)$, varies, the average probability of detection, $\overline{P_d}$, can be calculated

by averaging P_d in (3.5) over all SNR values as,

$$\overline{P_d} = \int_0^\infty P_d(\gamma, \lambda) f(\gamma) d\gamma = \int_0^\infty Q_u(\sqrt{2\gamma}, \sqrt{\lambda}) f(\gamma) d\gamma, \tag{3.8}$$

where $f(\gamma)$ denotes the PDF of the SNR in a fading channel.

3.2.1 Nakagami-m Fading Channel

In a Nakagami-m fading channel, the SNR of the signal is distributed according to a Gamma distribution as [51],

$$f(\gamma) = \frac{m^m (\gamma)^{m-1}}{(\overline{\gamma})^m \Gamma(m)} e^{-\frac{m\gamma}{\overline{\gamma}}}, \quad \gamma > 0, \tag{3.9}$$

where $\overline{\gamma}$ denote the local-mean SNR (SNR averaged over a few tens of wavelengths), and m is the Nakagami-m fading factor ($m \in [1/2,\ \infty)$).

The average probability of detection can be obtained by the following process: substituting (3.7) and (3.9) into (3.8), the average probability of detection, $\overline{P_{d,\text{Na}}}$, can be calculated as,

$$\overline{P_{d,\text{Na}}} = e^{-\frac{\lambda}{2}} \sum_{i=0}^{u-1} \frac{\left(\frac{\lambda}{2}\right)^i}{i!} + e^{-\frac{\lambda}{2}} \sum_{n=u}^{\infty} \frac{\left(\frac{\lambda}{2}\right)^n}{n!} \left(1 - \frac{m^m}{\Gamma(m)(\overline{\gamma})^m} \sum_{k=0}^{n-u} \frac{\int_0^\infty (\gamma)^{k+m-1} e^{-\frac{m+\overline{\gamma}}{\overline{\gamma}}\gamma} d\gamma}{k!} \right). \tag{3.10}$$

Using (3.351-3) in [71] for calculating the integral in (3.10), the result is obtained as,

$$\overline{P_{d,\text{Na}}} = e^{-\frac{\lambda}{2}} \sum_{i=0}^{u-1} \frac{\left(\frac{\lambda}{2}\right)^i}{i!} + e^{-\frac{\lambda}{2}} \sum_{n=u}^{\infty} \frac{\left(\frac{\lambda}{2}\right)^n}{n!} \left(1 - \left(\frac{m}{m+\overline{\gamma}}\right)^m \sum_{k=0}^{n-u} \frac{(m+k-1)!}{\Gamma(m)k!} \left(\frac{\overline{\gamma}}{m+\overline{\gamma}}\right)^k \right). \tag{3.11}$$

It can be shown that the above representation converges to 1 when the parameter $\overline{\gamma}$ goes to infinity for a constant m and any positive λ. As the detection threshold λ approaches to infinity, the average probability of detection converges to zero when $\overline{\gamma}$ is a constant. If the fading factor m goes to infinity (non-fading case), (3.11) converges to (3.7).

Since the above form contains infinite sums, the truncation error, $T_{\text{Nak}}(N)$, by truncating (3.11) after N iterations on index n, must be considered. As the number of computed terms, i.e. N,

varies, the truncation error is bounded by,

$$T_{\text{Nak}}(N) = e^{-\frac{\lambda}{2}} \sum_{n=u+N+1}^{\infty} \frac{\left(\frac{\lambda}{2}\right)^n}{n!} \left(1 - \left(\frac{m}{m+\overline{\gamma}}\right)^m \sum_{k=0}^{n-u} \frac{(m+k-1)!}{\Gamma(m)k!} \left(\frac{\overline{\gamma}}{m+\overline{\gamma}}\right)^k\right), \tag{3.12}$$

$$\leq e^{-\frac{\lambda}{2}} \sum_{n=u+N+1}^{\infty} \frac{\left(\frac{\lambda}{2}\right)^n}{n!} \left(1 - \left(\frac{m}{m+\overline{\gamma}}\right)^m \sum_{k=0}^{N+1} \frac{(m+k-1)!}{\Gamma(m)k!} \left(\frac{\overline{\gamma}}{m+\overline{\gamma}}\right)^k\right), \tag{3.13}$$

$$= \left(1 - e^{-\frac{\lambda}{2}} \sum_{n=0}^{u+N} \frac{\left(\frac{\lambda}{2}\right)^n}{n!}\right) \left(1 - \left(\frac{m}{m+\overline{\gamma}}\right)^m \sum_{k=0}^{N+1} \frac{(m+k-1)!}{\Gamma(m)k!} \left(\frac{\overline{\gamma}}{m+\overline{\gamma}}\right)^k\right), \tag{3.14}$$

where (3.13) holds because the term $1 - a^m \sum_{k=0}^{n-u} g(k)$ in (3.12) is monotonically decreasing with respect to n, and replacing n on index k by the smallest value $u+N+1$ will lead to an upper bound.

In order to obtain a specific accuracy when calculating $\overline{P_{d,\text{Na}}}$, the problem of "how many calculated terms are required in (3.11)?" is often of concern. Equation (3.14) can be used to determine the required N given the calculation accuracy.

Since the expression for the average probability of detection of (6) in [72] also contains infinite summations, to compare which converges faster, the result of (6) in [72] is rewritten by exchanging n and k as below,

$$\widetilde{\overline{P_{d,\text{Na}}}} = e^{-\frac{\lambda}{2}} \left(\frac{m}{m+\overline{\gamma}}\right)^m \sum_{k=0}^{\infty} \left(\frac{\overline{\gamma}}{m+\overline{\gamma}}\right)^k \frac{(m+k-1)!}{\Gamma(m)k!} \sum_{n=0}^{k+u-1} \frac{\left(\frac{\lambda}{2}\right)^n}{n!}. \tag{3.15}$$

If (3.15) were to be truncated after N iterations on k, then the truncation error of (3.15) is given by (7) in [72] as,

$$E_{\text{Nak}}(N) = \left(\frac{m}{m+\overline{\gamma}}\right)^m \left({}_1F_0\left(m;;\frac{\overline{\gamma}}{m+\overline{\gamma}}\right) - \sum_{k=0}^{N} \left(\frac{\overline{\gamma}}{m+\overline{\gamma}}\right)^k \frac{(m+k-1)!}{\Gamma(m)k!}\right), \tag{3.16}$$

$$= \left(\frac{m}{m+\overline{\gamma}}\right)^m \sum_{k=N+1}^{\infty} \left(\frac{\overline{\gamma}}{m+\overline{\gamma}}\right)^k \frac{(m+k-1)!}{\Gamma(m)k!}, \tag{3.17}$$

where ${}_1F_0(\alpha;;\beta)$ denotes a generalized hypergeometric series defined by,

$${}_pF_q(a_1, a_2, \cdots, a_p; b_1, b_2, \cdots, b_q; z) = \sum_{n=0}^{\infty} \frac{(a_1)_n (a_2)_n \cdots (a_p)_n}{(b_1)_n (b_2)_n \cdots (b_q)_n} \frac{z^n}{n!}, \tag{3.18}$$

where $(a)_n = \frac{\Gamma(a+n)}{\Gamma(a)}$ denotes the Pochhammer Symbol.

For comparison, (3.14) is rewritten as,

$$T_{\text{Nak}}(N) \leq \left(1 - e^{-\frac{\lambda}{2}} \sum_{n=0}^{u+N} \frac{\left(\frac{\lambda}{2}\right)^n}{n!}\right) \left(\frac{m}{m+\overline{\gamma}}\right)^m \sum_{k=N+2}^{\infty} \left(\frac{\overline{\gamma}}{m+\overline{\gamma}}\right)^k \frac{(m+k-1)!}{\Gamma(m)k!}. \tag{3.19}$$

Using (3.17) and (3.19), the relation of T_{Nak} and E_{Nak} is obtained as,

$$\frac{E_{\text{Nak}} - T_{\text{Nak}}}{\left(\frac{m}{m+\overline{\gamma}}\right)^m} \geq e^{-\frac{\lambda}{2}} \sum_{k=N+2}^{\infty} \left(\frac{\overline{\gamma}}{m+\overline{\gamma}}\right)^k \frac{(m+k-1)!}{\Gamma(m)k!} \sum_{n=0}^{u+N} \frac{\left(\frac{\lambda}{2}\right)^n}{n!} + \left(\frac{\overline{\gamma}}{m+\overline{\gamma}}\right)^{N+1} \frac{(m+N)!}{\Gamma(m)(N+1)!}. \tag{3.20}$$

Comparison of $E_{\text{Nak}}(N)$ and $T_{\text{Nak}}(N)$ shows that the proposed expression in (3.11) converges to the exact value more quickly than the existing expression in (3.15), since $E_{\text{Nak}}(N) - T_{\text{Nak}}(N) > 0$ holds true for any value of N when the local-mean SNR, $\overline{\gamma}$, is positive. The comparison of these expressions are shown in Figure 3.4.

3.2.2 Rician Fading Channel

The Rician model is often used to describe propagation paths which contain a strong dominant line of sight (LOS) component and several weaker scattered components. It includes both the nonfading AWGN channel and the Rayleigh fading channel as two special cases. The PDF of the SNR in a Rician fading channel is given by [51],

$$f(\gamma) = \frac{(K+1)e^{-K}}{\overline{\gamma}} e^{-\frac{(K+1)\gamma}{\overline{\gamma}}} I_0\left(2\sqrt{\frac{K(K+1)\gamma}{\overline{\gamma}}}\right), \quad \gamma > 0, \tag{3.21}$$

where K is the shape parameter.

The average probability of detection over Rician fading channel can be obtained by the following process: substituting (3.7) and (3.21) into (3.8), the average probability of detection can be given by,

$$\begin{aligned} \overline{P_{d,\text{Ri}}} \;=\; & e^{-\frac{\lambda}{2}} \sum_{i=0}^{u-1} \frac{\left(\frac{\lambda}{2}\right)^i}{i!} + e^{-\frac{\lambda}{2}} \sum_{n=u}^{\infty} \frac{\left(\frac{\lambda}{2}\right)^n}{n!} \Bigg(1 - \frac{(K+1)e^{-K}}{\overline{\gamma}} \\ & \times \sum_{k=0}^{n-u} \frac{\int_0^{\infty} (\gamma)^k e^{-\frac{K+\overline{\gamma}+1}{\overline{\gamma}}\gamma} I_0\left(2\sqrt{\frac{K(K+1)\gamma}{\overline{\gamma}}}\right) d\gamma}{k!}\Bigg). \end{aligned} \tag{3.22}$$

Using (6.643-2) in [71] for calculating the integral, with the aid of the identity for the Whittaker function in (9.220-2) of [71], the result of $\overline{P_{d,\mathrm{Ri}}}$ is obtained as,

$$\begin{aligned}\overline{P_{d,\mathrm{Ri}}} \;=\; & e^{-\frac{\lambda}{2}}\sum_{i=0}^{u-1}\frac{\left(\frac{\lambda}{2}\right)^i}{i!}+e^{-\frac{\lambda}{2}}\sum_{n=u}^{\infty}\frac{\left(\frac{\lambda}{2}\right)^n}{n!}\Bigg(1-\frac{(K+1)e^{-K}}{K+\overline{\gamma}+1}\\ & \times\sum_{k=0}^{n-u}\left(\frac{\overline{\gamma}}{K+\overline{\gamma}+1}\right)^k\Phi\left(k+1,1;\frac{K(K+1)}{K+\overline{\gamma}+1}\right)\Bigg). \qquad (3.23)\end{aligned}$$

where $\Phi(\alpha,\gamma;z)$ denotes the Confluent Hypergeometric function, given by,

$$\Phi(\alpha,\gamma;z)=1+\frac{\alpha}{\gamma}\frac{z}{1!}+\frac{\alpha(\alpha+1)}{\gamma(\gamma+1)}\frac{z^2}{2!}+\cdots\quad. \qquad (3.24)$$

In contrast to the result in [52], (3.23) is applicable for any value of the time bandwidth product. Since the time bandwidth product, u, describes the number of independent samples of the signal in the observation time T, the result in (3.23) is more flexible, and can be used to analyse the performance of energy detection, when the data is sampled at any sampling rate over any observation time of T.

The infinite sums are also involved in (3.23), thus it is necessary to analyse the truncation error when truncating the infinite sum to a finite one for calculation. As the number of computed terms, N, varies, the truncation error, $T_{\mathrm{Ric}}(N)$, is bounded by,

$$\begin{aligned}T_{\mathrm{Ric}}(N) \;=\; & e^{-\frac{\lambda}{2}}\sum_{n=u+N+1}^{\infty}\frac{\left(\frac{\lambda}{2}\right)^n}{n!}\Bigg(1-\frac{(K+1)e^{-K}}{K+\overline{\gamma}+1}\\ & \times\sum_{k=0}^{n-u}\left(\frac{\overline{\gamma}}{K+\overline{\gamma}+1}\right)^k\Phi\left(k+1,1;\frac{K(K+1)}{K+\overline{\gamma}+1}\right)\Bigg), \qquad (3.25)\\ \leq\; & \left(1-e^{-\frac{\lambda}{2}}\sum_{n=0}^{u+N}\frac{\left(\frac{\lambda}{2}\right)^n}{n!}\right)\Bigg(1-\frac{(K+1)e^{-K}}{K+\overline{\gamma}+1}\\ & \times\sum_{k=0}^{N+1}\left(\frac{\overline{\gamma}}{K+\overline{\gamma}+1}\right)^k\Phi\left(k+1,1;\frac{K(K+1)}{K+\overline{\gamma}+1}\right)\Bigg). \qquad (3.26)\end{aligned}$$

Equation (3.26) can be used to determine how many calculated terms, N, is required to obtain a specific accuracy when calculating $\overline{P_{d,\mathrm{Ri}}}$ using (3.23).

3.2.3 Slow Fading Channel

The strength of the received signals in cognitive radios is also affected by shadowing from buildings, hills, and other objects. In such a scenario, empirical measurements showed that the received power fluctuates with a log-normal distribution about the area-mean power for various outdoor and indoor environments [20]. In a slow fading channel, the PDF of the SNR is given by [51],

$$f(\gamma) = \frac{\xi}{\sqrt{2\pi}\sigma\gamma} \exp\left(-\frac{(10\log_{10}\gamma - \ddot{\gamma})^2}{2\sigma^2}\right), \quad \gamma > 0, \tag{3.27}$$

where $\xi = 10/\ln(10)$, $\ddot{\gamma}$ (dB) denotes the area-mean SNR (SNR averaged over an area with a radius of tens or hundreds of metres), and σ (dB) denotes the standard deviation of $10\log_{10}\gamma$. There exists no closed-form expression for the average probability of detection on substituting (3.7) and (3.27) into (3.8). However, a tractable expression can be found if the log-normal distribution is approximated by the Wald distribution (also known as the inverse Gaussian distribution) [73, 74], whose PDF is given by,

$$f(\gamma) = \sqrt{\frac{\eta}{2\pi}}\gamma^{-3/2} \exp\left(-\frac{\eta(\gamma-\theta)^2}{2\theta^2\gamma}\right), \quad \gamma > 0, \tag{3.28}$$

where $\theta = \mathrm{E}(\gamma)$ denotes the expectation of γ, and η is the shape parameter. The variance of γ is $\frac{\theta^3}{\eta}$, i.e. $\mathrm{Var}(\gamma) = \frac{\theta^3}{\eta}$. In order to approximate the log-normal distribution, the method of moments is used to obtain the parameters η, θ with $\ddot{\gamma}$, σ as below,

$$\begin{aligned} \theta &= \exp\left(\frac{\ddot{\gamma}}{\xi} + \frac{\sigma^2}{2\xi^2}\right), \\ \eta &= \frac{\theta}{\exp(\frac{\sigma^2}{\xi^2}) - 1}. \end{aligned} \tag{3.29}$$

The average probability of detection over slow fading channel can be obtained by the following process: substituting (3.7) and (3.28) into (3.8), the average probability of detection is given by,

$$\overline{P_{d,\mathrm{Sha}}} = e^{-\frac{\lambda}{2}} \sum_{i=0}^{u-1} \frac{\left(\frac{\lambda}{2}\right)^i}{i!} + e^{-\frac{\lambda}{2}} \sum_{n=u}^{\infty} \frac{\left(\frac{\lambda}{2}\right)^n}{n!} \left(1 - \sqrt{\frac{\eta}{2\pi}} e^{\frac{\eta}{\theta}} \sum_{k=0}^{n-u} \frac{\int_0^\infty (\gamma)^{k-\frac{3}{2}} e^{(-\frac{\eta/2}{\gamma} - \frac{2\theta^2+\eta}{2\theta^2}\gamma)} d\gamma}{k!}\right). \tag{3.30}$$

Using (3.471-9) in [71] for calculating the integral, the result of $\overline{P_{d,\text{Sha}}}$ is obtained as,

$$\overline{P_{d,\text{Sha}}} = \sum_{i=0}^{u-1} \frac{\left(\frac{\lambda}{2}\right)^i e^{-\frac{\lambda}{2}}}{i!} + \sum_{n=u}^{\infty} \frac{\left(\frac{\lambda}{2}\right)^n e^{-\frac{\lambda}{2}}}{n!} \left(1 - \sqrt{\frac{\eta}{2\pi}} e^{\frac{\eta}{\theta}} \sum_{k=0}^{n-u} \frac{\left(\sqrt{\frac{\eta\theta^2}{2\theta^2+\eta}}\right)^{k-\frac{1}{2}} \mathrm{K}_{k-\frac{1}{2}}\left(\frac{\sqrt{\eta(2\theta^2+\eta)}}{\theta}\right)}{k!} \right). \tag{3.31}$$

where $\mathrm{K}_{k-\frac{1}{2}}(a)$ denotes the modified Bessel function of the second kind with order $k-\frac{1}{2}$.

Likewise, the truncation error, $T_{\text{Sha}}(N)$, can be bounded by,

$$T_{\text{Sha}}(N) = \sum_{n=u+N+1}^{\infty} \frac{e^{-\frac{\lambda}{2}}\left(\frac{\lambda}{2}\right)^n}{n!} \left(1 - \sqrt{\frac{\eta}{2\pi}} e^{\frac{\eta}{\theta}} \sum_{k=0}^{n-u} \frac{\left(\sqrt{\frac{\eta\theta^2}{2\theta^2+\eta}}\right)^{k-\frac{1}{2}} \mathrm{K}_{k-\frac{1}{2}}\left(\frac{\sqrt{\eta(2\theta^2+\eta)}}{\theta}\right)}{k!} \right) \tag{3.32}$$

$$\leq \sum_{n=u+N+1}^{\infty} \frac{e^{-\frac{\lambda}{2}}\left(\frac{\lambda}{2}\right)^n}{n!} \left(1 - \sqrt{\frac{\eta}{2\pi}} e^{\frac{\eta}{\theta}} \sum_{k=0}^{N+1} \frac{\left(\sqrt{\frac{\eta\theta^2}{2\theta^2+\eta}}\right)^{k-\frac{1}{2}} \mathrm{K}_{k-\frac{1}{2}}\left(\frac{\sqrt{\eta(2\theta^2+\eta)}}{\theta}\right)}{k!} \right) \tag{3.33}$$

$$= \left(1 - \sum_{n=0}^{u+N} \frac{e^{-\frac{\lambda}{2}}\left(\frac{\lambda}{2}\right)^n}{n!} \right) \left(1 - \sqrt{\frac{\eta}{2\pi}} e^{\frac{\eta}{\theta}} \sum_{k=0}^{N+1} \frac{\left(\sqrt{\frac{\eta\theta^2}{2\theta^2+\eta}}\right)^{k-\frac{1}{2}} \mathrm{K}_{k-\frac{1}{2}}\left(\frac{\sqrt{\eta(2\theta^2+\eta)}}{\theta}\right)}{k!} \right). \tag{3.34}$$

3.3 Data Fusion Based Collaborative Spectrum Sensing

This section analyses the spectrum sensing performance using different data fusion approaches over multiple i.i.d. Nakagami-m fading channels, or slow fading channels. For convenience of derivation, the representation of the generalised Marcum Q-function in (3) of [75] is used in this section, and P_d is represented as,

$$P_d(\gamma, \lambda) = 1 - \sum_{n=0}^{\infty} \frac{\Gamma'(n+u, \frac{\lambda}{2})}{\Gamma(n+u)n!} (\gamma)^n e^{-\gamma}, \tag{3.35}$$

where $\Gamma'(a, x)$ denotes the lower incomplete gamma function, i.e. $\Gamma'(a, x) = \int_0^x t^{a-1} e^{-t} dt$.

3.3.1 Square-Law Combining

Using the SLC scheme, the squared and integrated energy vectors, $Y^1, Y^2, \cdots, Y^v$, from v distributed cognitive radios are gathered at a FC, where the test statistic, $Y_{\text{slc}} = \sum_{i=1}^{v} Y^i$ is formed [53]. Thus, under the H_0 hypothesis in (3.36), if these v fading channels are i.i.d., and all branches have the same noise variance, the test statistic, Y_{slc}, follows a central chi-square distribution with a $2vu$ DoF. On the other hand, under the H_1 hypothesis, it follows a non-central chi-square distribution with a $2vu$ DoF and non-central parameter of γ_{slc} as,

$$Y_{\text{slc}} \sim \begin{cases} \chi^2_{2vu}, & H_0, \\ \chi^2_{2vu}(2\gamma_{\text{slc}}), & H_1, \end{cases} \tag{3.36}$$

where $\gamma_{\text{slc}} = \sum_{i=1}^{v} \gamma^i$, and γ^i is the SNR in cognitive radio node i.

In the case of non-fading AWGN channels, the probabilities of false alarm and detection under the SLC diversity reception scheme can be obtained by substituting the DoF to (3.4) and (3.5) as,

$$P_f^{\text{slc}} = \frac{\Gamma(vu, \lambda/2)}{\Gamma(vu)}, \tag{3.37}$$

$$\begin{aligned} P_d^{\text{slc}} &= Q_{vu}(\sqrt{2\gamma_{\text{slc}}}, \sqrt{\lambda}), \\ &= 1 - \sum_{n=0}^{\infty} \frac{\Gamma'(n+vu, \frac{\lambda}{2})}{\Gamma(n+vu)n!} (\gamma_{\text{slc}})^n e^{-\gamma_{\text{slc}}}. \end{aligned} \tag{3.38}$$

When the signal experiences fading over v channels, the average probability of false alarm will not change, and the average probability of detection can be evaluated by averaging P_d^{slc} in (3.38) over the combined SNR distribution as,

$$\overline{P_d^{\text{slc}}} = \int_0^{\infty} P_d^{\text{slc}}(\gamma_{\text{slc}}, \lambda) f(\gamma_{\text{slc}}) d\gamma_{\text{slc}}. \tag{3.39}$$

When v Nakagami-m fading channels are i.i.d., the PDF of the SLC output SNR, i.e. $f(\gamma_{\text{slc}})$, is given by [51],

$$f(\gamma_{\text{slc}}) = \frac{m^{vm} (\gamma_{\text{slc}})^{vm-1}}{(\overline{\gamma})^{vm} \Gamma(vm)} e^{-\frac{m}{\overline{\gamma}} \gamma_{\text{slc}}}, \quad \gamma_{\text{slc}} > 0. \tag{3.40}$$

The average probability of detection using SLC can be obtained by the following process:

substituting (3.38) and (3.40) into (3.39), the average probability of detection is given by,

$$\overline{P_d^{\text{slc}}}=1-\frac{m^{vm}}{\Gamma(vm)(\overline{\gamma})^{vm}}\sum_{n=0}^{\infty}\frac{\Gamma'(n+vu,\frac{\lambda}{2})}{\Gamma(n+vu)n!}\int_0^{\infty}(\gamma)^{n+vm-1}e^{-\frac{m+\overline{\gamma}}{\overline{\gamma}}\gamma}d\gamma. \tag{3.41}$$

Using (3.351-3) in [71] for calculating the integral in (3.41), the result of $\overline{P_d^{\text{slc}}}$ can be obtained as,

$$\overline{P_d^{\text{slc}}}=1-\left(\frac{m}{m+\overline{\gamma}}\right)^{vm}\sum_{n=0}^{\infty}\frac{\Gamma'(n+vu,\frac{\lambda}{2})(vm)_n}{\Gamma(n+vu)n!}\left(\frac{\overline{\gamma}}{m+\overline{\gamma}}\right)^{n}. \tag{3.42}$$

Since the above form contains infinite sums, the truncation error is considered. The truncation error, $T_{\text{slc}}(N)$, is introduced by truncating (3.42) after N iterations on index n. As N varies, the truncation error is bounded by,

$$T_{\text{slc}}(N)=\left(\frac{m}{m+\overline{\gamma}}\right)^{vm}\sum_{n=N+1}^{\infty}\frac{\Gamma'(n+vu,\frac{\lambda}{2})(vm)_n}{\Gamma(n+vu)n!}\left(\frac{\overline{\gamma}}{m+\overline{\gamma}}\right)^{n}, \tag{3.43}$$

$$\leq\left(\frac{m}{m+\overline{\gamma}}\right)^{vm}\sum_{n=N+1}^{\infty}\frac{(vm)_n}{n!}\left(\frac{\overline{\gamma}}{m+\overline{\gamma}}\right)^{n}\frac{\Gamma'(vu+N+1,\frac{\lambda}{2})}{\Gamma(vu+N+1)}, \tag{3.44}$$

$$=(1-\epsilon(v\overline{\gamma},vm,N))\frac{\Gamma'(vu+N+1,\frac{\lambda}{2})}{\Gamma(vu+N+1)}, \tag{3.45}$$

where the inequality holds true due to that the function $\frac{\Gamma'(a,x)}{\Gamma(a)}$ is monotonically decreasing with respect to a, and $\epsilon(x,\nu,z)$ is defined by,

$$\epsilon(x,\nu,z)\triangleq\left(\frac{\nu}{\nu+x}\right)^{\nu}\sum_{n=0}^{z}\frac{(\nu)_n}{n!}\left(\frac{x}{\nu+x}\right)^{n}. \tag{3.46}$$

In slow fading channels, the PDF of the SNR in the node i, i.e. $f(\gamma^i)$, can be approximated by a Wald distribution. When all fading channels are stationary and i.i.d., the condition $\frac{\eta^i}{(\theta^i)^2}=\frac{\text{E}(\gamma^i)}{\text{Var}(\gamma^i)}=b$ (constant) can be satisfied. Thus, the combined SNR under the SLC scheme, γ_{slc}, will also follow the Wald distribution [76]. The PDF of γ_{slc} can be easily obtained by replacing each η with $v\eta$, each θ with $v\theta$, and each γ with γ_{slc} in (3.28). Using a similar method to that of

the single slow fading channel, the average probability of detection can be calculated as below,

$$\overline{P_{d,\text{slow}}^{\text{slc}}} = e^{-\frac{\lambda}{2}} \sum_{i=0}^{vu-1} \frac{\left(\frac{\lambda}{2}\right)^i}{i!} + e^{-\frac{\lambda}{2}} \sum_{n=vu}^{\infty} \frac{\left(\frac{\lambda}{2}\right)^n}{n!} \times \left(1 - \sqrt{\frac{v\eta}{2\pi}} e^{\frac{\eta}{\theta}} \sum_{k=0}^{n-vu} \frac{\left(\sqrt{\frac{\eta\theta^2 v^2}{2v\theta^2+\eta}}\right)^{k-\frac{1}{2}} \mathrm{K}_{k-\frac{1}{2}}\left(\frac{\sqrt{\eta(2v\theta^2+\eta)}}{\theta}\right)}{k!} \right). \quad (3.47)$$

The above result can also be obtained by replacing each η with $v\eta$, each θ with $v\theta$, and each u with vu in (3.31).

3.3.2 Square-Law Selection

Using the SLS scheme, the FC only selects the branch with the largest energy, i.e. $Y_{\text{sls}} = \max(Y^1, Y^2, \cdots, Y^v)$. In the case of non-fading AWGN channels, the probabilities of false alarm and detection under the SLS diversity reception scheme is given by [53],

$$P_f^{\text{sls}} = 1 - \left(1 - \frac{\Gamma(u, \lambda/2)}{\Gamma(u)}\right)^v, \quad (3.48)$$

$$P_d^{\text{sls}} = 1 - \prod_{i=1}^{v} \left(1 - Q_u(\sqrt{2\gamma^i}, \sqrt{\lambda})\right), \quad (3.49)$$

where the noise variance is assumed to be 1. When those v fading channels are i.i.d., the average probability of detection can be evaluated by averaging P_d^{sls} in (3.49) over all possible SNRs as,

$$\overline{P_d^{\text{sls}}} = \int_0^{\infty} P_d^{\text{sls}}(\gamma^i, \lambda) f(\gamma^i) d\gamma^i, \quad (3.50)$$

where $f(\gamma^i)$ is given by,

$$f(\gamma^i) = \frac{m^m (\gamma^i)^{m-1}}{(\overline{\gamma^i})^m \Gamma(m)} e^{-\frac{m}{\overline{\gamma^i}}\gamma^i}, \quad \gamma^i > 0, \quad (3.51)$$

where $\overline{\gamma^i}$ denotes the local-mean SNR in the i-th node.

By substituting (3.49) and (3.51) into (3.50), the expression for the average probability of de-

tection can be evaluated. Since the channels are i.i.d., the result is obtained as,

$$\overline{P_d^{\text{sls}}} = 1 - \prod_{i=1}^{v} \left(\frac{m}{m + \overline{\overline{\gamma^i}}} \right)^m \sum_{n=0}^{\infty} \frac{\Gamma'(n+u, \frac{\lambda}{2})(m)_n}{\Gamma(n+u)n!} \left(\frac{\overline{\gamma^i}}{m + \overline{\overline{\gamma^i}}} \right)^n. \tag{3.52}$$

In comparison with the results in [77], the results in (3.42) and (3.52) have lower computational complexity when u and m are integer multiples of $\frac{1}{24}$. This is because gamma function can be evaluated quickly using arithmetic-geometric mean iterations with computational complexity of $\mathcal{O}(\log bM(b))$ [78], where b denotes the number of digits of precision at which the function is to be evaluated, and $M(b)$ stands for the complexity of the chosen multiplication algorithm. By contrast, the computational complexity of confluent Hypergeometric function in [77] is $\mathcal{O}((\log b)^2 M(b))$ [79]. On the other hand, even though the proposed expressions contain infinite sums as well, it can be shown that they converge to the exact value very quickly.

In order to derive the truncation error, assume,

$$\zeta = \left(\frac{m}{m+\overline{\gamma}} \right)^m \sum_{n=0}^{N} \frac{\Gamma'(n+u, \frac{\lambda}{2})(m)_n}{\Gamma(n+u)n!} \left(\frac{\overline{\gamma}}{m+\overline{\gamma}} \right)^n, \tag{3.53}$$

and

$$\vartheta = \left(\frac{m}{m+\overline{\gamma}} \right)^m \sum_{n=N+1}^{\infty} \frac{\Gamma'(n+u, \frac{\lambda}{2})(m)_n}{\Gamma(n+u)n!} \left(\frac{\overline{\gamma}}{m+\overline{\gamma}} \right)^n. \tag{3.54}$$

Thus, the truncation error, $T_{\text{sls}}(N)$, in (3.52) can be represented by,

$$T_{\text{sls}}(N) = (\zeta + \vartheta)^v - \zeta^v = \sum_{l=1}^{v} \binom{v}{l} \zeta^{v-l} \vartheta^l. \tag{3.55}$$

Regarding (3.53) and (3.54), the following inequalities hold true,

$$\zeta \leq \epsilon(\overline{\gamma}, m, N) \frac{\Gamma'(u, \frac{\lambda}{2})}{\Gamma(u)}, \tag{3.56}$$

and

$$\vartheta \leq (1 - \epsilon(\overline{\gamma}, m, N)) \frac{\Gamma'(u+N+1, \frac{\lambda}{2})}{\Gamma(u+N+1)}. \tag{3.57}$$

The upper bound of $T_{\text{sls}}(N)$ can be determined by substituting (3.56) and (3.57) into (3.55).

3.3.3 Maximum Ratio Combining

Under the MRC scheme, multiple cognitive radios directly amplify and forward the received signals, rather than the energy, to the FC, where the data from multiple cognitive radios are combined by an MRC combiner. Then an energy detector measures the output of the MRC combiner. Therefore, the test statistic, Y_{mrc}, can be modeled by central and non-central chi-square distributed random variables as,

$$Y_{\text{mrc}} \sim \begin{cases} \chi^2_{2u}, & H_0 \\ \chi^2_{2u}(2\gamma_{\text{mrc}}), & H_1 \end{cases} \tag{3.58}$$

where $\gamma_{\text{mrc}} = \sum_{i=1}^{v} \gamma^i$ denotes the instantaneous SNR at the output of the MRC combiner [51].

Over AWGN channels, the probabilities of false alarm and detection under an MRC diversity scheme can be given by,

$$P_f^{\text{mrc}} = \frac{\Gamma(u, \lambda/2)}{\Gamma(u)}, \tag{3.59}$$

$$\begin{aligned} P_d^{\text{mrc}} &= Q_u(\sqrt{2\gamma_{\text{mrc}}}, \sqrt{\lambda}), \\ &= 1 - \sum_{n=0}^{\infty} \frac{\Gamma'(n+u, \frac{\lambda}{2})}{\Gamma(n+u)n!} (\gamma_{\text{mrc}})^n e^{-\gamma_{\text{mrc}}}. \end{aligned} \tag{3.60}$$

The PDF of the SNR, $f(\gamma_{\text{mrc}})$, at the output of the MRC combiner is given by,

$$f(\gamma_{\text{mrc}}) = \frac{m^{vm} (\gamma_{\text{mrc}})^{vm-1}}{(\overline{\gamma})^{vm} \Gamma(vm)} e^{-\frac{m}{\overline{\gamma}} \gamma_{\text{mrc}}}, \quad \gamma_{\text{mrc}} > 0. \tag{3.61}$$

Averaging (3.60) over (3.61), the average probability of detection is given by,

$$\overline{P_d^{\text{mrc}}} = 1 - \frac{m^{vm}}{\Gamma(vm)(\overline{\gamma})^{vm}} \sum_{n=0}^{\infty} \frac{\Gamma'(n+u, \frac{\lambda}{2})}{\Gamma(n+u)n!} \int_0^{\infty} (\gamma)^{n+vm-1} e^{-\frac{m+\overline{\gamma}}{\overline{\gamma}} \gamma} d\gamma. \tag{3.62}$$

Using (3.351-3) in [71] for calculating the integral in (3.62), the result of $\overline{P_d^{\text{mrc}}}$ is obtained as,

$$\overline{P_d^{\text{mrc}}} = 1 - \left(\frac{m}{m + \overline{\gamma}} \right)^{vm} \sum_{n=0}^{\infty} \frac{\Gamma'(n+u, \frac{\lambda}{2})(vm)_n}{\Gamma(n+u)n!} \left(\frac{\overline{\gamma}}{m + \overline{\gamma}} \right)^n. \tag{3.63}$$

The corresponding truncation error is bounded by,

$$T_{\text{mrc}}(N) = \left(\frac{m}{m+\overline{\gamma}}\right)^{vm} \sum_{n=N+1}^{\infty} \frac{\Gamma'(n+u, \frac{\lambda}{2})(vm)_n}{\Gamma(n+u)n!} \left(\frac{\overline{\gamma}}{m+\overline{\gamma}}\right)^{n}, \tag{3.64}$$

$$\leq \left(\frac{m}{m+\overline{\gamma}}\right)^{vm} \sum_{n=N+1}^{\infty} \frac{(vm)_n}{n!} \left(\frac{\overline{\gamma}}{m+\overline{\gamma}}\right)^{n} \frac{\Gamma'(u+N+1, \frac{\lambda}{2})}{\Gamma(u+N+1)}, \tag{3.65}$$

$$= (1-\epsilon(v\overline{\gamma}, vm, N)) \frac{\Gamma'(u+N+1, \frac{\lambda}{2})}{\Gamma(u+N+1)}. \tag{3.66}$$

3.3.4 Selection Combining

Under SC, the energy detector measures the output of the SC combiner. The test statistic, Y_{sc}, can be modeled by central and non-central chi-square distributed random variables as,

$$Y_{\text{sc}} \sim \begin{cases} \chi^2_{2u}, & H_0 \\ \chi^2_{2u}(2\gamma_{\text{sc}}), & H_1 \end{cases} \tag{3.67}$$

where $\gamma_{\text{sc}} = \max(\gamma^1, \gamma^2, \cdots, \gamma^v)$ denotes the instantaneous SNR at the output of the SC combiner [51]. In other words, rather than processing all fading branches, SC processes only one of the diversity branches with the highest SNR. Over AWGN channels, the probabilities of false alarm and detection under the SC diversity scheme can be given by,

$$P_f^{\text{sc}} = \frac{\Gamma(u, \lambda/2)}{\Gamma(u)}, \tag{3.68}$$

$$\begin{aligned} P_d^{\text{sc}} &= Q_u(\sqrt{2\gamma_{\text{sc}}}, \sqrt{\lambda}), \\ &= 1 - \sum_{n=0}^{\infty} \frac{\Gamma'(n+u, \frac{\lambda}{2})}{\Gamma(n+u)n!} (\gamma_{\text{sc}})^n e^{-\gamma_{\text{sc}}}. \end{aligned} \tag{3.69}$$

Over multiple i.i.d. Nakagami-m fading channels, if the fading factor m is restricted to integer values, the PDF of γ_{sc} can be obtained from the Appendix of [80] as,

$$f(\gamma_{\text{sc}}) = \frac{v}{\Gamma(m)} \sum_{l=0}^{v-1} (-1)^l \binom{v-1}{l} \sum_{k=0}^{l(m-1)} b_k^l \left(\frac{m}{\overline{\gamma}}\right)^{m+k} (\gamma_{\text{sc}})^{m+k-1} e^{-\frac{m(l+1)}{\overline{\gamma}}\gamma_{\text{sc}}}, \tag{3.70}$$

where b_k^l can be recursively computed as,

$$b_0^l = 1, \quad b_1^l = l, \quad b_{(m-1)l}^l = \frac{1}{\Gamma(m)^l},$$
$$b_k^l = \frac{1}{k}\sum_{j=1}^{J} b_{k-j}^l \frac{j(l+1)-k}{j!}, \tag{3.71}$$

with $J = \min(k, m-1)$, and $k \in [2, (m-1)l-1]$.

The average probability of detection can be evaluated by averaging P_d^{sc} in (3.69) over the SNR distribution in (3.70) as,

$$\overline{P_d^{\text{sc}}} = \int_0^\infty P_d^{\text{sc}}(\gamma_{\text{sc}}, \lambda) f(\gamma_{\text{sc}}) d\gamma_{\text{sc}}. \tag{3.72}$$

Substituting (3.69) and (3.70) into (3.72), with manipulation, the average probability of detection is given by,

$$\overline{P_d^{\text{sc}}} = 1 - \frac{v}{\Gamma(m)}\sum_{l=0}^{v-1}(-1)^l \binom{v-1}{l} \sum_{k=0}^{l(m-1)} b_k^l \left(\frac{m}{\overline{\gamma}}\right)^{m+k} \sum_{n=0}^{\infty} \frac{\Gamma'(n+u, \frac{\lambda}{2})}{\Gamma(n+u)n!} \times \int_0^\infty (\gamma)^{n+m+k-1} e^{-\frac{m(l+1)+\overline{\gamma}}{\overline{\gamma}}\gamma} d\gamma. \tag{3.73}$$

Using (3.351-3) in [71] for calculating the integral in (3.73), the result can be obtained as,

$$\overline{P_d^{\text{sc}}} = 1 - v\sum_{l=0}^{v-1}(-1)^l \binom{v-1}{l} \sum_{k=0}^{l(m-1)} b_k^l \left(\frac{m}{\overline{\gamma}}\right)^{m+k} \times \sum_{n=0}^{\infty} \frac{\Gamma'(n+u, \frac{\lambda}{2})(m)_{n+k}}{\Gamma(n+u)n!} \left(\frac{\overline{\gamma}}{m(l+1)+\overline{\gamma}}\right)^{n+m+k}. \tag{3.74}$$

3.4 Simulation Results

Receiver operating characteristic (ROC) analysis has been widely used in the signal detection theory. It is an ideal technique to quantify the tradeoff between the probability of detection and the probability of false alarm. In cognitive radio networks, the probabilities of false alarm and detection have different implications on the performance of spectrum sensing. Specifically, the probability of false alarm, P_f, is related with the spectral utilisation efficiency, where a larger P_f will trigger a lower spectral utilisation efficiency. On the other hand, the probability of missed detection, $1 - P_d$, measures the risk of the cognitive radios causing interference to

the PUs when the cognitive radios are filling in the spectrum holes. Thus, the ROC analysis is selected as the metric for evaluating the performance of spectrum sensing over fading channels. In this section, the complementary ROC curves (P_f vs $1 - P_d$) are used to show the detection performance of energy detection over various fading channels. The following parameters are considered,

- The local-mean SNR $\overline{\gamma}$ or the area-mean SNR $\ddot{\overline{\gamma}}$.
- The time bandwidth product in the test statistic, u.
- The shape parameters, e.g., m if Nakagami-m fading, K if Rician fading.
- The number of collaborative cognitive radio nodes, v.

Spectrum sensing over a single fading channel

Figure 3.2 presents the detection performance of energy detection over a Nakagami-m fading channel with $m = 1$ and $m = 2$. The Nakagami-m fading factor, m (also called shape parameter), indicates the severity of fading and the quality of the channel. The severity of fading reduces when the fading factor, m, increases. The Nakagami-m fading model includes the Rayleigh fading ($m = 1$) and the one-sided Gaussian fading ($m = \frac{1}{2}$) as special cases. In the limit as $m \to \infty$, it converges to a non-fading channel. In both figures, it is shown that, the higher the SNR is, the better a detection performance one can achieve. In addition, a smaller u will result in a smaller probability of missed detection for a fixed probability of false alarm. This is because that the parameter u is directly related to the DoF of the chi-square in (3.2), and the DoF implies how many independent standard normal random variables in the chi-square distribution. It is also evident that there exists performance improvement from $m = 1$ to $m = 2$, as in the latter case, the quality of the channel is better than that of the former case.

Figure 3.3 depicts the complementary ROC curves of the energy detector over a Rician fading channel with $K = 1$ in (a), and $K = 3$ in (b). The shape parameter, K, is defined as the ratio of signal power in dominant component over the (local-mean) scattered power. When the LOS component disappears, i.e. $K = 0$, Rician fading reduces to Rayleigh fading. In Figure 3.3, a similar phenomenon can be found that the energy detection has a better detection performance when the SNR increases, or the DoF, $2u$, decreases. For example, in the left hand figure, when $P_f = 1\%$, $u = 1$, and the SNR increases from 10 dB to 15 dB, the probability of missed detection declines from 30% to 10%. Comparing the left hand figure with the right hand one,

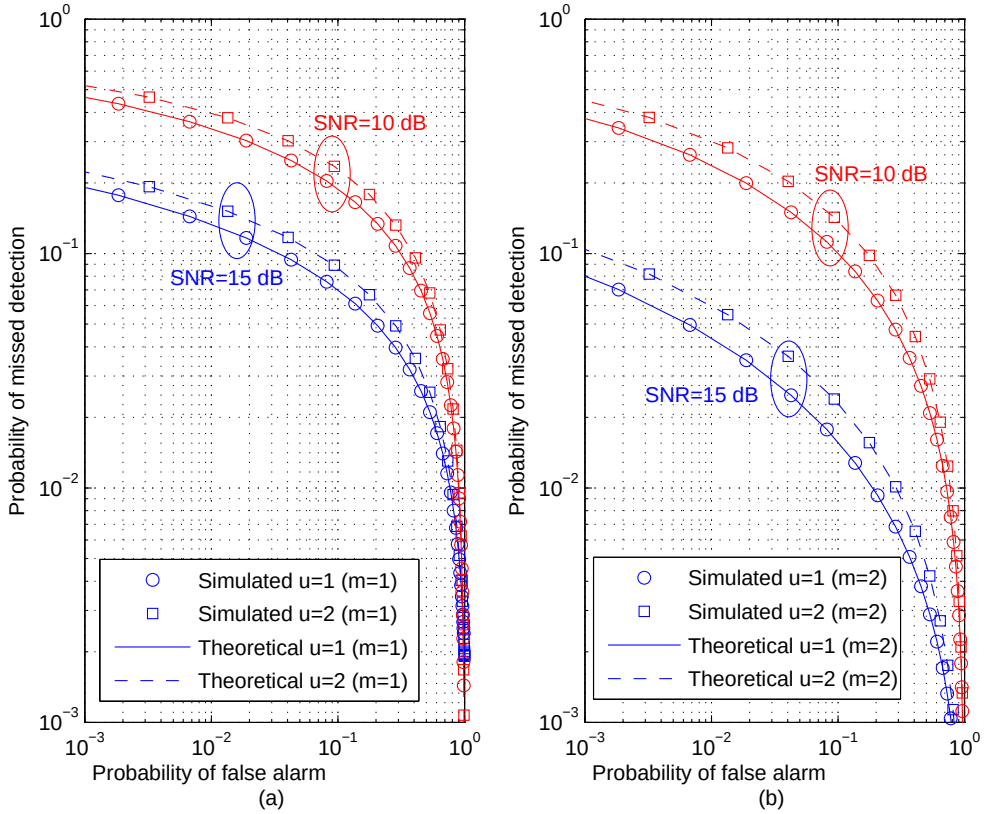

Figure 3.2: *Complementary ROC curves of energy detection over a Nakagami-m fading channel, with (a) $m = 1$, and (b) $m = 2$. This figure is produced by changing the detection threshold λ from 0.1 to 1000.*

it shows that when the shape parameter K reduces from 3 to 1, the performance of energy detection degrades due to the power of the dominant component decreasing.

Figure 3.4(a) compares the truncation error, (3.14), of the proposed expression in (3.11) with the truncation error, (3.16), of the existing expression in (3.15) for a Nakagami-m fading channel. The truncation error (3.26) for a Rician fading channel is depicted in Figure 3.4(b). Both figures reveal that using $N = 30$ computed terms, the expressions in (3.11) and (3.23) converge rapidly to their final values, with the corresponding truncation error achieving double-precision accuracy. By contrast, the truncation error $E_{\text{Nak}}(N)$ in (3.16) decreases more slowly, compared with the truncation error $T_{\text{Nak}}(N)$ in (3.14). For example, as shown in Figure 3.4(a), when $N = 10$, the truncation error $T_{\text{Nak}}(10)$ in (3.14) is approximately 10^{-2}. In contrast, to obtain $E_{\text{Nak}} = 10^{-2}$, more than 50 computed terms are required in (3.15). In other words, the proposed expression in (3.11) converges more quickly to its final value than the expression in (3.15).

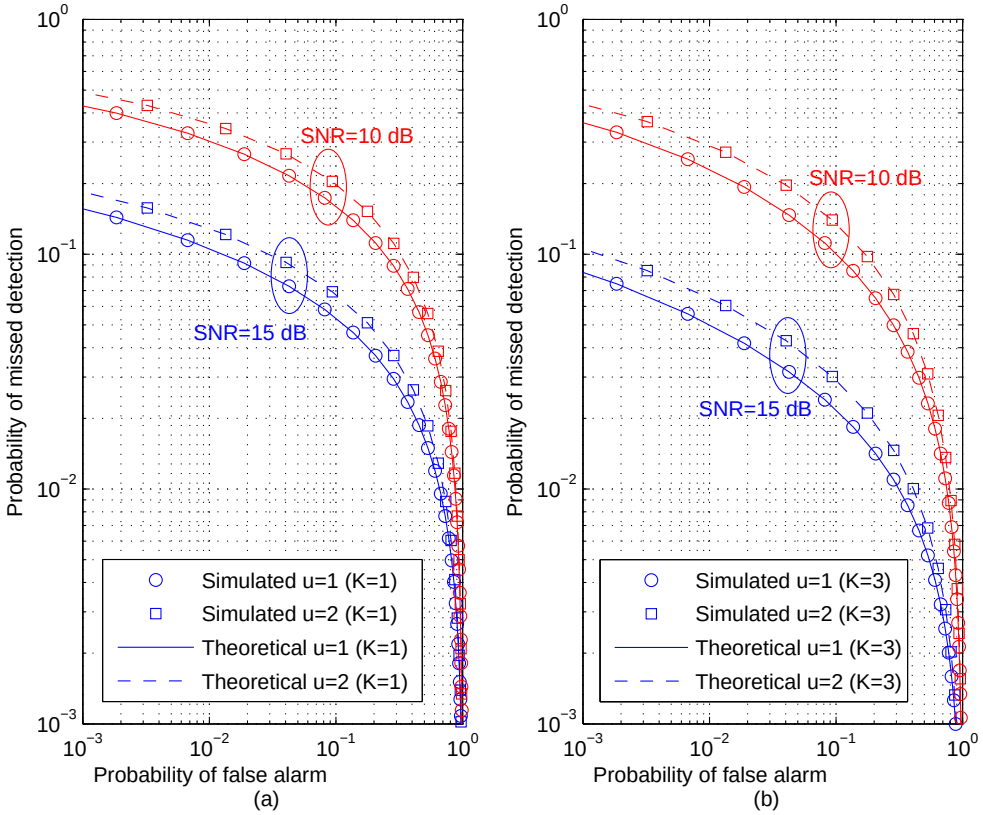

Figure 3.3: *Complementary ROC curves of energy detection over a Rician fading channel, with (a) $K = 1$, and (b) $K = 3$. This figure is produced by changing the detection threshold λ from 0.1 to 1000.*

Figure 3.4 shows that a larger shape parameter, e.g., m in (a), or K in (b), will lead to a slightly faster convergence.

Figure 3.5 demonstrates the performance of energy detection over a slow fading channel with a shadow standard deviation, $\sigma = 4$ dB in (a), and an area-mean SNR, $\ddot{\bar{\gamma}} = 10$ dB, in (b). Since the Wald distribution is used to approximate the log-normal distribution for deriving the average probability of detection, the accuracy of the proposed expression should be verified. Figure 3.5 compares the theoretical result in (3.31) with simulated results. It demonstrates that the theoretical results closely fit the experimental results, and the accuracy of the proposed expression varies under different conditions. As shown in the left hand figure, when the area-mean SNR, $\ddot{\bar{\gamma}}$, increases from 5 dB to 10 dB, the approximation error grows slightly. In addition, it is found that when the average probability of false alarm decreases, the approximation error gradually increases. This phenomenon stems from the long right tail of the long-normal dis-

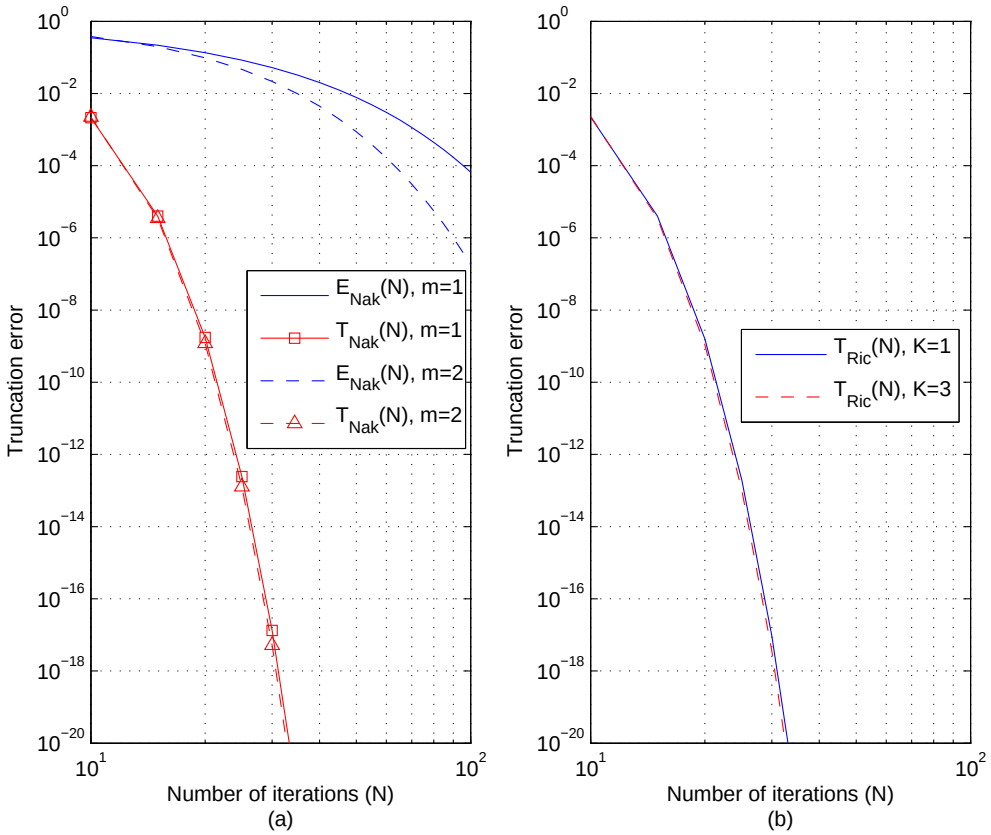

Figure 3.4: *The truncation error after N iterations on index n, when the time bandwidth product $u = 1$, the local-mean SNR $\overline{\gamma} = 10$ dB, and the average probability of false alarm $\overline{P_f} = 0.01$, with (a) Nakagami-m fading, and (b) Rician fading.*

tribution, which is difficult to match. In the right hand figure, the influence of the right tail of the long-normal distribution is further investigated. It illustrates that a larger shadow standard deviation (equivalent to a longer right tail in the log-normal distribution), σ, will cause a worse detection performance, and a larger mismatch. For example, the case of $\sigma = 5$ dB has a larger probability of missed detection, and a larger approximation error than that of $\sigma = 3$ dB.

Collaborative spectrum sensing over fading channels

In the following simulations, the detection performance of energy detection using collaborative approaches, i.e. MRC, SC, SLC, and SLS, will be compared. In Figure 3.6, it is manifest that both MRC and SLC have diversity gains for spectrum sensing, compared with the no-diversity case. In the case of $v = 3$ and SNR$= 20$ dB, the gain is several orders of magnitude on index $1 - P_d$. It is found that the diversity gain decreases when either the number of cognitive radio

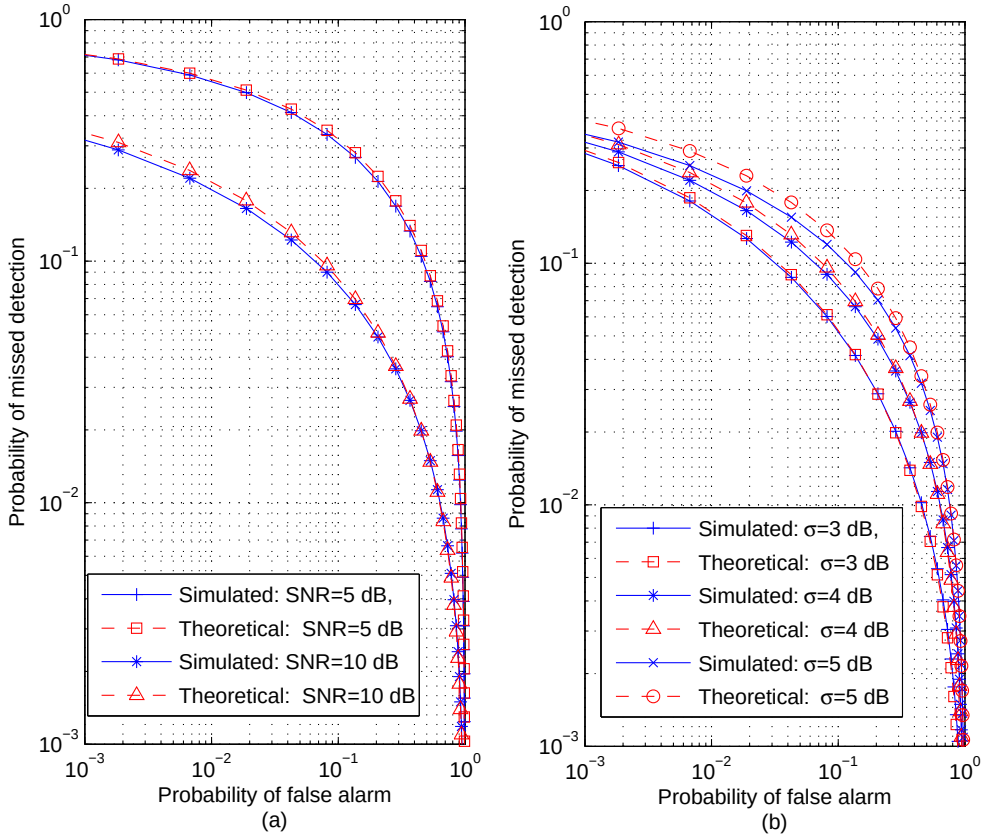

Figure 3.5: *Complementary ROC curves of energy detection over a slow fading channel with (a) the shadow standard deviation* $\sigma = 4$ dB*, and (b) the area-mean SNR* $\ddot{\bar{\gamma}} = 10$ dB*, compared with theoretical result in (3.31).*

nodes, v, or the SNR reduces. Comparing the detection performance of energy detection using MRC with that using SLC, it reveals that, with the aid of full CSI, MRC outperforms SLC. When SNR$=20$ dB and $v = 3$, the gain of using MRC over SLC is approximately one order of magnitude on index $1 - P_d$. But, the gain becomes less when the SNR or the number of cognitive radio nodes go down.

In Figure 3.7, a similar phenomenon can be found that either a higher SNR or a larger v will result in a better detection performance, and both SC and SLS have gains compared with the no-diversity case. However, the gain becomes less when the probability of false alarm decreases. For example, considering the SLS with $v = 3$ and SNR$=20$ dB, the gain is approximately seven orders of magnitude on index $1 - P_d$ when $P_f = 10\%$, and then becomes five orders of magnitude on index $1 - P_d$ when $P_f = 0.01\%$. It is more obvious for the case of SNR$=10$ dB since the diversity gain is very small when $P_f = 0.01 \sim 0.1\%$. Figure 3.7 also shows

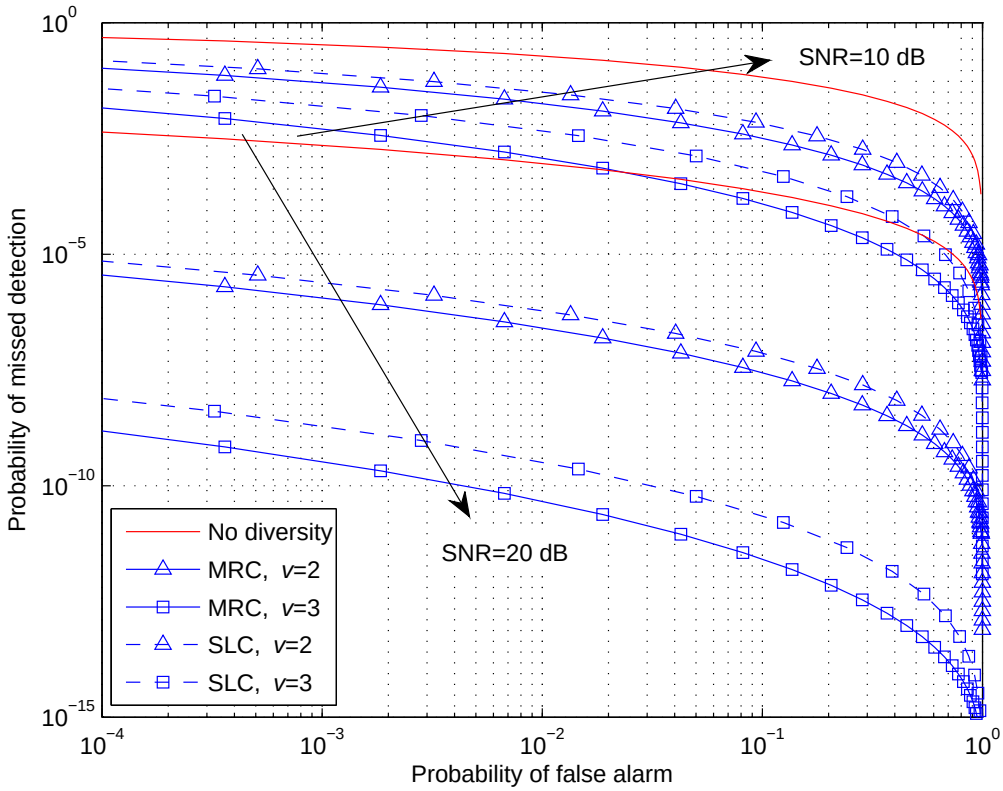

Figure 3.6: *Complementary ROC curves of energy detection under MRC and SLC scheme, when the Nakagami fading factor* $m = 3$*, and the time bandwidth product* $u = 1$.

that the SLS scheme is superior to the SC scheme. The reason behind this phenomenon is that, in comparison with choosing the branch with the highest SNR, selecting the branch with the largest energy to do the hypothesis test is more straightforward for improving detection performance.

Figure 3.8 describes the influence of the quality of channel on the performance of energy detection when using different collaborative strategies. It shows that using MRC one can obtain the best detection performance when the severity of fading changes. When $m = 3$, MRC outperforms SLC, SLS, and SC by roughly one, two, and three orders of magnitude on index $1 - P_d$, respectively. On the other hand, for Rayleigh fading channels ($m = 1$), the gain of MRC is less. The SLC is found to be the second best out of those four strategies. Comparing the scheme of MRC and SLC, it can be seen that MRC not only requires full CSI, but also needs double the transmission bandwidth of SLC. This is because when using SLC, the data to be transmitted is real, rather than complex data when using MRC. Since the full CSI is difficult to obtain, and

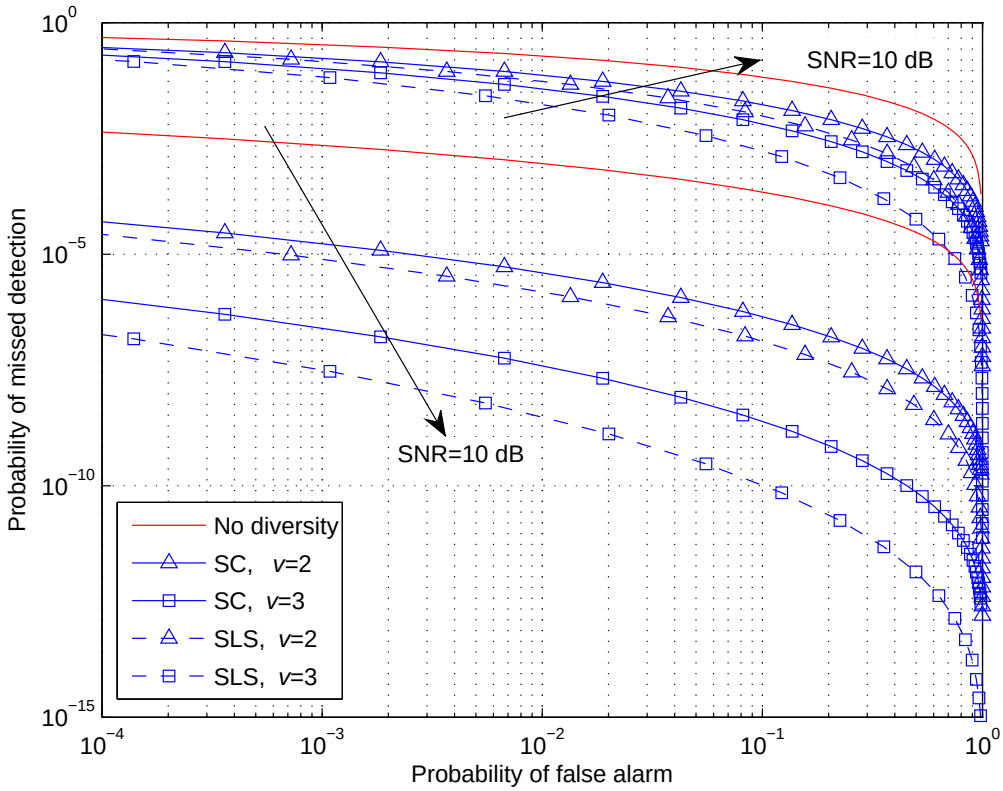

Figure 3.7: *Complementary ROC curves of energy detection under SC and SLS scheme, when the Nakagami fading factor $m = 3$, and the time bandwidth product $u = 1$.*

data transmission is expensive in cognitive radio networks, SLC is more attractive than MRC. Figure 3.8 also illustrates that when $m = 3$ the detection performance is much better than that over Rayleigh fading channels ($m = 1$). This is because the severity of fading for the former case is less than that of the latter case.

Figure 3.9 compares the performance of different collaborative strategies when the time bandwidth product, u, alters. It shows that MRC outperforms the other collaborative approaches for both $u = 1$ and $u = 2$. In comparison to SLC, the gain of using MRC is approximately one order of magnitude in $1 - P_d$, similar to the gain of SLC over SLS, and the gain of SLS over SC. Additionally, Figure 3.9 depicts that for a fixed probability of false alarm, the probability of missed detection decreases when u reduces. This is because the time bandwidth product, u, gives the number of independent standard normal random variables in the chi-square distribution, and more independent variables will lead to a poorer detection performance.

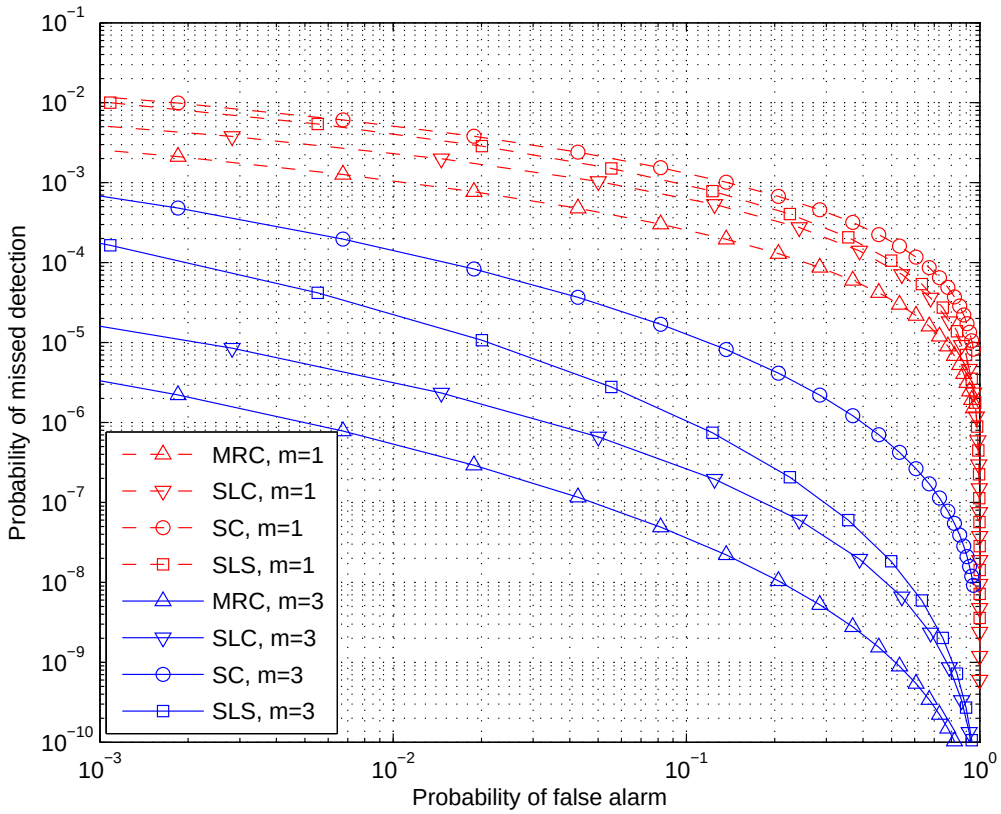

Figure 3.8: *Comparison of MRC, SLC, SC, and SLS, when the local-mean SNR* $\overline{\gamma} = 15$ dB, *the number of collaborative cognitive radios* $v = 3$, *and the time bandwidth product* $u = 1$.

Figure 3.10 presents the comparison of the performance of energy detection using different collaborative strategies when the local-mean SNR, $\overline{\gamma}$, changes. It is evident that SLC, SLS and SC are always inferior to MRC. The benefit of using MRC diminishes when the local-mean SNR goes down. It is also noticeable that SLS has a better detection performance than SC, however, the gain decreases when the probability of false alarm decreases. For example, when the local-mean SNR, $\overline{\gamma} = 5$ dB, and $P_f = 0.1\%$, the gain of SLS over SC is zero.

3.5 Conclusions

This chapter has derived rapidly converging expressions for the average probability of detection when using energy detection over a single Nakagami-m, or Rician fading channel. Even though infinite sums are involved, simulation results have shown that the proposed expressions

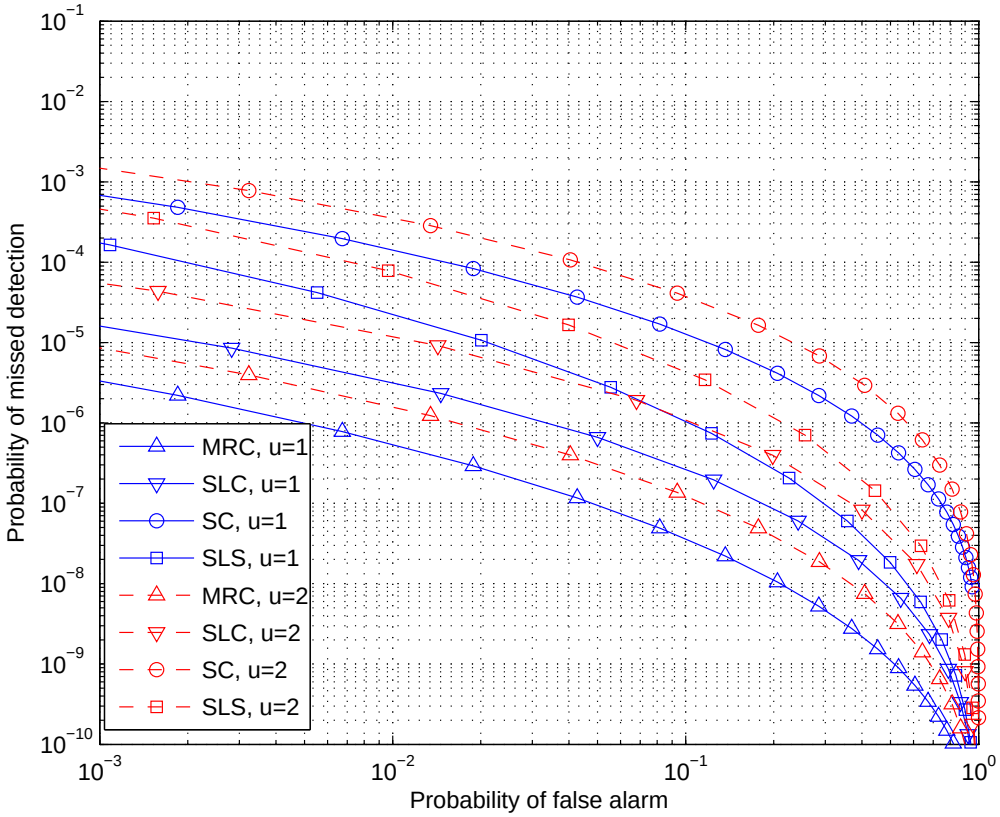

Figure 3.9: *Comparison of MRC, SLC, SC, and SLS, when the local-mean SNR* $\overline{\gamma} = 15$ dB, *the number of collaborative cognitive radios* $v = 3$, *and the Nakagami fading factor* $m = 3$.

converge rapidly. Additionally, a computationally tractable expression for the average probability of detection over a slow fading channel has been obtained, by replacing the log-normal distribution with the Wald distribution. It has been verified that the proposed expression closely fits the simulation results. Due to the effect of the long right tail in the log-normal distribution, the mismatch becomes larger when the shadow standard deviation increases.

Moreover, since collaborative spectrum sensing is a crucial means for combating fading, this chapter has studied the performance of spectrum sensing using different collaborative strategies, i.e. MRC, SC, SLC, and SLS. With perfect CSI, MRC gives an upper bound of the detection performance. In the case where CSI is not available, SLC is a good choice as it has a better detection performance than SLS. Furthermore, it is found that MRC and SC require double the transmission bandwidth compared with SLC and SLS, because they must forward complex data to the FC while the latter two only transmit real data (energy vectors). Hence, where there is a

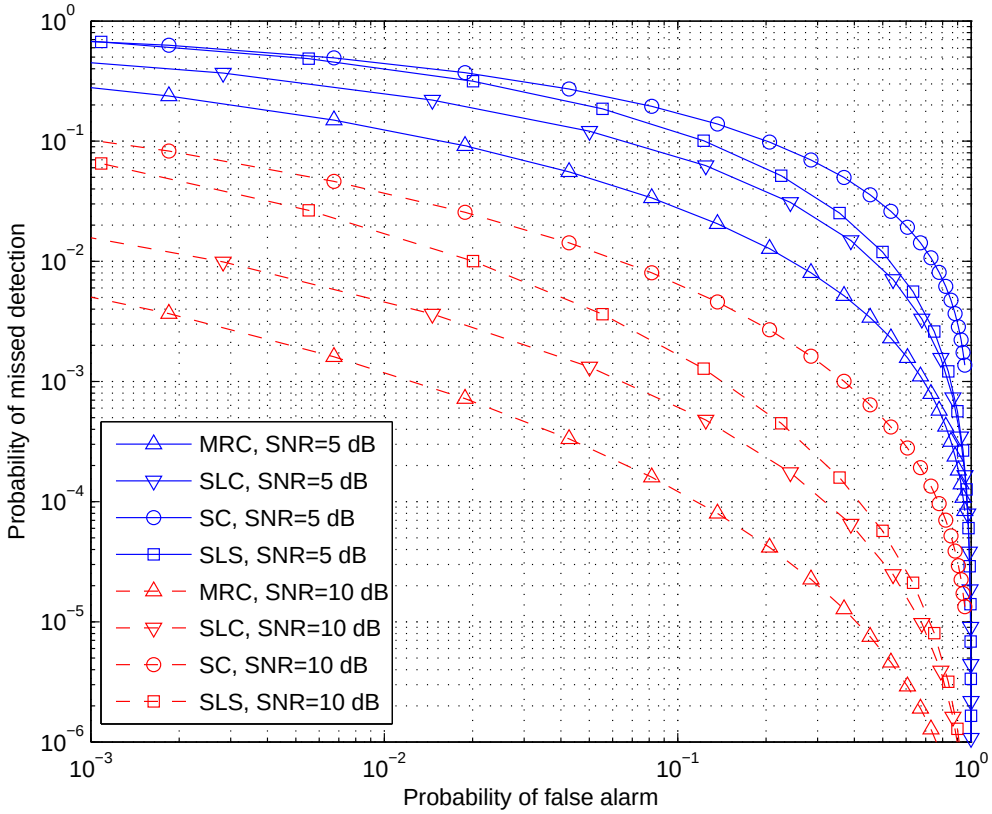

Figure 3.10: *Comparison of MRC, SLC, SC, and SLS, when the time bandwidth product* $u = 1$*, the number of collaborative cognitive radios* $v = 3$*, and the Nakagami fading factor* $m = 3$*.*

restriction on the bandwidth of the control channel, SLC is more attractive.

Chapter 4
Distributed Wideband Spectrum Sensing

Wideband spectrum sensing is becoming increasingly important to cognitive radio systems as a means of identifying spectrum holes or characterising interference. Meanwhile, due to the effects of multipath fading and shadowing, a single cognitive radio cannot distinguish between a deep fade and an idle band. In such a scenario, distributed spectrum sensing has been widely considered for combating fading or shadowing [21, 81]. A data fusion based approach is preferred for distributed spectrum sensing as it offers much better detection performance than decision fusion approaches [82]. Nonetheless, to jointly perform distributed and wideband spectrum sensing using conventional distributed spectrum sensing techniques, the transmission of raw data through a control channel is very expensive. Moreover, the high sampling rates are problematic when the distributed cognitive radios are battery powered.

This chapter proposes a novel wideband spectrum sensing model for distributed cognitive radio networks, with the following contributions:

- A multirate asynchronous sub-Nyquist sampling (MASS) system that employs multiple low-rate ADCs is developed to implement wideband spectrum sensing. The key features of the MASS system are, 1) low implementation complexity, 2) applicability to distributed cognitive radio networks, 3) energy-efficiency for sharing spectrum sensing data, and 4) robustness against the lack of time synchronisation.
- The conditions under which the recovery of the full spectrum is unique are presented by using CS analysis. A trade-off is made between the number of sampling channels and the probability of successful spectrum recovery. In addition, the effects of noise and the model mismatching are considered.
- The proposed MASS system is applied to distributed cognitive radio networks. When the spectra of the cognitive radio nodes have a common spectral support, using one low-rate ADC in each cognitive radio node can also successfully recover the full spectrum.

This is obtained by applying a hybrid matching pursuit (HMP) algorithm, a synthesis of the distributed compressive sensing simultaneous orthogonal matching pursuit (DCS-SOMP) and compressive sampling matching pursuit (CoSaMP).

This chapter is organized as follows. Section 4.1 briefly introduces techniques in CS theory. Section 4.2 proposes a MASS system. Using CS theory, the sufficient conditions for full spectrum recovery are derived. The practical implementation issues, for example, the effects of noise and model mismatch, are discussed in Section 4.2.5. Simulation results are presented in Section 4.4, followed by discussions and conclusions in Section 4.5 and Section 4.6, respectively.

4.1 Preliminaries of Compressive Sensing

CS theory [59] indicates that a signal, $\vec{x} \in \mathbb{C}^N$, which is k-sparse, can be exactly recovered from M $(M \ll N)$ linear projections/measurements $\vec{y} \in \mathbb{C}^M (\vec{y} = \mathbf{\Phi}\vec{x})$, where $\mathbf{\Phi} \in \mathbb{C}^{M \times N}$ is the measurement matrix. Note that k-sparse means that the k largest values of $\vec{x}$ are not negligible. Indeed, CS theory states that if the geometry of the sparse signals is preserved during measurements/sampling, the sparse signals can be recovered by some algorithm from only a few measurements. The performance of recovery is determined by three factors, namely, the sparsity, k, of the signal $\vec{x}$, the properties of the measurement matrix $\mathbf{\Phi}$, and the recovery algorithm.

One important property of the measurement matrix is the restricted isometry property (RIP) [83]. A measurement matrix $\mathbf{\Phi}$ has the RIP with parameter ($k, \varrho_k \in (0,1), N$) if,

$$(1 - \varrho_k)\|\vec{x}\|_2^2 \leq \|\mathbf{\Phi}\vec{x}\|_2^2 \leq (1 + \varrho_k)\|\vec{x}\|_2^2, \quad \forall \vec{x} \in \Omega. \tag{4.1}$$

where ϱ_k denotes the restricted isometry constant and the indices of k-nonzero components in $\vec{x}$ are assumed to be included by a support, i.e. Ω.

Calculating the RIP of a particular measurement matrix is practically impossible as it is NP-hard, but a lower bound can be obtained by calculating its mutual coherence. The mutual coherence is a computationally tractable metric for evaluating the suitability of the chosen measurement matrix.

Definition 4.1 [84]: Let $\mathbf{\Phi}$ be expressed as $\mathbf{\Phi} = [\vec{\phi}_1\ \vec{\phi}_2\ ...\ \vec{\phi}_N]$, where $\vec{\phi}_j$ denotes the j^{th} column of the matrix $\mathbf{\Phi}$. Then the mutual coherence, μ, of the matrix $\mathbf{\Phi}$ is given by,

$$\mu = \max_{j \neq h \in [1,N]} | < \hat{\phi}_j, \hat{\phi}_h > |, \tag{4.2}$$

where $\hat{\phi}_j = \frac{\vec{\phi}_j}{\|\vec{\phi}_j\|_2}$ denotes the ℓ_2 normalised column.

The aim is to keep μ, or ϱ_k to a minimum to allow linear projections to be inverted in a stable manner. Donoho and Elad [85, 86] have proved that a small mutual coherence μ can guarantee the performance of the recovery as below.

Theorem 4.1 [85, 86]: Assume that a signal $\vec{x}$ is k-sparse. When the mutual coherence μ of the measurement matrix $\mathbf{\Phi}$ satisfies,

$$\mu < \frac{1}{2k-1}, \tag{4.3}$$

BP [62] or orthogonal matching pursuit (OMP) [87] can be used to find the sparsest solution of the k-sparse signal $\vec{x} \in \mathbb{C}^N$ from measurements $\vec{y} \in \mathbb{C}^M$.

In practice, the measurements will be contaminated with noise. The observation is $\vec{y} = \mathbf{\Phi}\vec{x} + \vec{z}$, where $\vec{z}$ denotes noise vector. Linear programming BP (LPBP) [88] was proposed to solve $\vec{x}$ by,

$$\min \|\vec{x}\|_1, \quad \text{subject to:} \quad \|\vec{y} - \mathbf{\Phi}\vec{x}\|_2 \leq \rho \tag{4.4}$$

where ρ is greater than or equal to the noise level, i.e. $\rho \geq \|\vec{z}\|_2$.

The reconstruction can achieve good performance when $\mathbf{\Phi}$ has the RIP with small ϱ_{2k} as shown in Theorem 4.2 below.

Theorem 4.2 [88]: When the noise level $\|\vec{z}\|_2 \leq \rho$, and measurement matrix $\mathbf{\Phi}$ has the RIP with $\varrho_{2k} < \sqrt{2} - 1$, using (4.4) the recovered result $\tilde{x}$ will satisfy,

$$\|\tilde{x} - \vec{x}\|_2 \leq \frac{A}{\sqrt{k}} \|\vec{x} - \vec{x}^{\text{opt}}\|_1 + B\rho \tag{4.5}$$

where constants A and B are acceptably small, and $\vec{x}^{\text{opt}}$ denotes the optimal k-sparse approximation of $\vec{x}$.

In a distributed network, the received signals in different sensors are not only individually sparse, but also jointly sparse (have nonzero entries at the same locations). A separate recovery

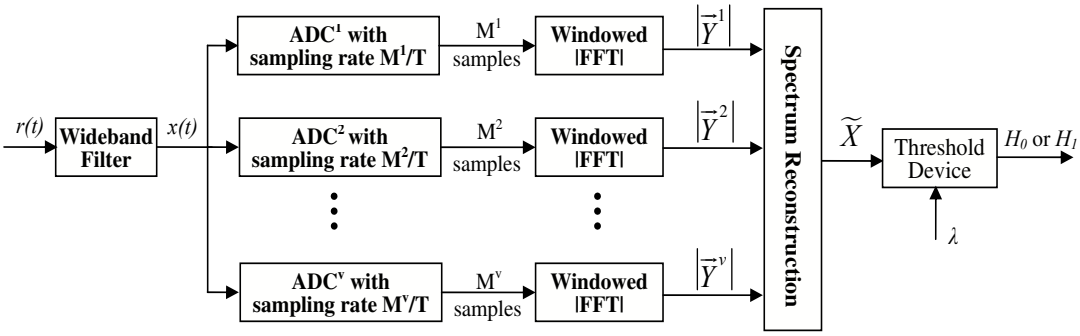

Figure 4.1: *The schematic illustration of the multirate asynchronous sub-Nyquist sampling system in one cognitive radio node. The wideband filters are altered to have the same bandwidth of W.*

strategy (each sensor recovers the signal individually) will require more measurements. One of the most important reasons is that it neglects the correlations of signals between sensors. The DCS-SOMP algorithm was presented in [89] for reconstructing the joint sparse signals with fewest measurements by a boost-and-recover approach.

4.2 Multirate Asynchronous Sub-Nyquist Sampling

This section will now present a MASS system to sense the wideband spectrum using multiple low-rate ADCs in Section 4.2.1. For simplicity, this section firstly consider the case that one cognitive radio node is equipped with parallel ADCs, which uniformly sample the wideband signal at different sub-Nyquist rates in the noiseless case. Section 4.2.2 will concentrate on exploring linear projection when performing sub-Nyquist sampling. After that the effect of sub-Nyquist sampling will be considered in Section 4.2.3. Using CS theory, the sufficient conditions for reconstructing the sparsest representation of the wideband spectrum are derived in Section 4.2.4. Some practical implementation issues are discussed in Section 4.2.5.

4.2.1 System Design

Partially motivated by MRS in [33], the system will use a multirate sampling scheme as shown in Figure 4.1. Instead of electro-optical devices, low-rate ADCs are employed in MASS. Since the average spectral occupancy is very low, it is assumed that the non-aliased discrete Fourier transform (DFT) spectrum (full spectrum, when the sampling rate is not less than the Nyquist

rate) is k-sparse, and consists of several subband signals with different unknown bandwidths. The k-sparse means that only k components in the full spectrum are non-negligible. In a carrier sense, it means that there are only a few active carriers even though most of them are allocated to different wireless systems. The sparsity level, k, of the non-aliased DFT spectrum can be obtained from initialisation, for example, coarse spectrum scanning [23, 90, 91], and will not be addressed here. The wideband filter prior to the ADCs removes only frequencies outside the spectrum of interest, and is altered to have the largest spectral estimation bandwidth, W. It is assumed that there are v ADCs that sample the wideband signal $x(t)$ at different rates, $\frac{M^1}{T}, \frac{M^2}{T}, ..., \frac{M^v}{T}$, over the same observation time of T. In addition, during the observation time T, the observed spectra are assumed to be wide-sense stationary (WSS). Note that no anti-aliasing filter is used prior to the ADCs, thus aliasing occurs. The length of samples in their corresponding channels will be $M^1, M^2, ..., M^v$ ($M^{i\in[1,v]} \sim \mathcal{O}(\sqrt{N})$. A tapered window, such as the Hamming window, is used to combat the effect of leakage, and keep the sparsity level, k, of the non-aliased DFT spectrum as small as possible. The spectral observations are obtained by applying an FFT to the windowed samples in each channel. The magnitude vectors of the sub-Nyquist rate spectra, $|\overrightarrow{Y}^1|, |\overrightarrow{Y}^2|, ..., |\overrightarrow{Y}^v|$ ($|\overrightarrow{Y}^i| \in \mathbb{R}^{M^i}$, $i \in [1, v]$), are used to form a concatenated equation. After that, the measurement matrix $\mathbf{\Phi}$ is constructed by only using M^i ($\forall\ i \in [1, v]$) and N. Then the non-aliased spectrum is recovered using a CS algorithm, e.g., BP, OMP, followed by spectrum detection on the reconstructed spectrum $\widetilde{X}$.

MASS has several advantages for application in cognitive radio networks, including,

1. Wideband spectrum sensing is implemented with sub-Nyquist sampling, which relaxes the stringent requirements on ADCs. Each ADC samples the wideband signal at rate of $\frac{M^{i\in[1,v]}}{T} \sim \mathcal{O}(\sqrt{\frac{2W}{T}})$, rather than the Nyquist sampling rate of $2W$.

2. The low-rate ADCs behave as acquisition devices as well as spectrum compression devices.

3. The compression/measurement matrix used in MASS is deterministic, and can be easily constructed once $M^{i\in[1,v]}$ and N are known. Thus, the transmission and storage of a measurement matrix is unnecessary.

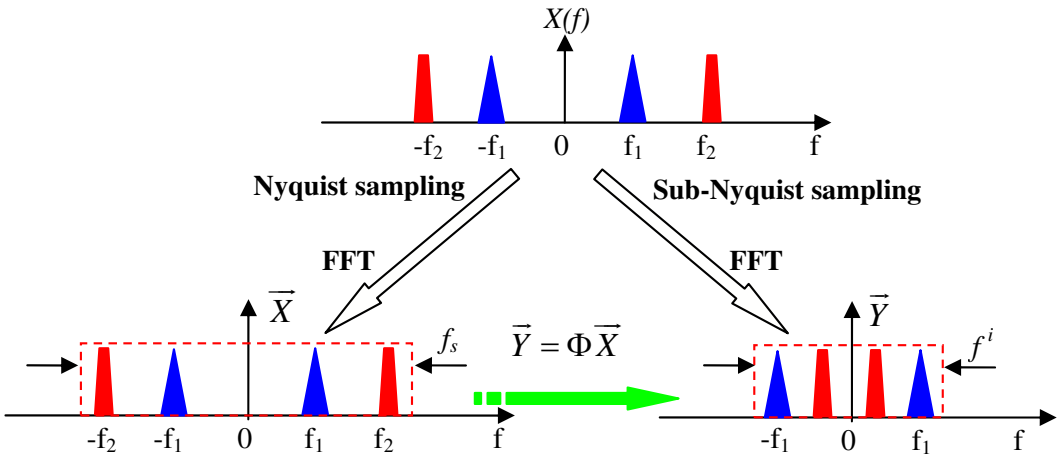

Figure 4.2: *Interpretation of the relationship between the non-aliased spectrum, $\vec{X}$, and the sub-Nyquist rate spectrum, $\vec{Y}$, when the Fourier spectrum of $x(t)$ is denoted by $X(f)$.*

4.2.2 Relate Sub-Nyquist DFT Spectrum to Nyquist DFT Spectrum

A sub-Nyquist rate spectrum (aliased spectrum) can be viewed as a linear projection from the non-aliased spectrum as shown in Figure 4.2.

$x(t)$ represents the output signal of the wideband filter in Figure 4.1, with a bandwidth of W. In the short observation time of T, the continuous signal $x(t)$ is sampled at a sub-Nyquist rate of $f^i = \frac{1}{\Delta t^i} = \frac{M^i}{T} < 2W$ in the i-th channel (different channels sample at different rates). After a tapered window, the sampled signal in the i-th channel, $y^i(t)$, can be represented by [92],

$$y^i(t) = \sum_{l=-\infty}^{\infty} x(t)\delta(t - l\Delta t^i)w_T(t), \tag{4.6}$$

where $w_T(t)$ is a tapered window, Δt^i is the sampling interval in the sampling channel i, and $\delta(t)$ is a Dirac delta function.

The Fourier transform of the sampled signal in the i-th sampling channel is given by [92],

$$Y^i(f) = f^i \sum_{l=-\infty}^{\infty} \int_{-\infty}^{\infty} X(\omega + lf^i)W_T(f - \omega)d\omega. \tag{4.7}$$

where $W_T(f)$ denotes the Fourier transform of the tapered window, and $X(f)$ denotes the Fourier transform of $x(t)$.

If $x(t)$ is sampled at or above the Nyquist rate, i.e. $f_s = \frac{N}{T} \geq 2W$, over an observation time of T, the Fourier transform of the sampled signal can be represented by,

$$X^i(f) = f_s \int_{-\infty}^{\infty} X(\omega) W_T(f-\omega) d\omega, \quad \forall |f| < f_s/2, \tag{4.8}$$

where $X^i(f)$ is a non-aliased full spectrum in the sampling channel i. Replacing f by $f + lf^i$, and ω by $\omega + lf^i$ in (4.8), and then substituting it into (4.7), the relationship between $Y^i(f)$ and $X^i(f)$ can be obtained as,

$$Y^i(f) = \frac{f^i}{f_s} \sum_{l=-\infty}^{\infty} X^i(f + lf^i), \quad \forall |f + lf^i| < f_s/2. \tag{4.9}$$

Since the observation time T in both cases are the same, the same frequency resolution applies to these two cases, i.e. $\Delta f = \frac{f_s}{N} = \frac{f^i}{M^i}$, where N, and M^i are integers and denote the number of samples at sampling rates f_s, and f^i, respectively. By defining an integer $m \in [-\lfloor \frac{M^i}{2} \rfloor, \lfloor \frac{M^i}{2} \rfloor]$, and a scalar $\varsigma \in [0, \Delta f)$, such that $f = m\Delta f + \varsigma$, (4.9) can be rewritten as,

$$Y^i(m\Delta f + \varsigma) = \frac{M^i}{N} \sum_{l=-\infty}^{\infty} X^i\big((m + lM^i)\Delta f + \varsigma\big) \tag{4.10}$$

$$= \frac{M^i}{N} \sum_{n=-\lfloor N/2 \rfloor}^{\lfloor N/2 \rfloor} X^i\left(n\Delta f + \varsigma\right) \sum_{l=-\infty}^{\infty} \delta[n - (m + lM^i)], \tag{4.11}$$

where $\lfloor a \rfloor$ denotes the floor function, and gives the largest integer not greater than a.

Sampling the Fourier transform spectrum in (4.11) at rate of $\frac{1}{\Delta f}$, the DFT spectrum is obtained as,

$$Y^i[m] = \frac{M^i}{N} \sum_{l=-\infty}^{\infty} X^i[m + lM^i] = \frac{M^i}{N} \sum_{n=-\lfloor N/2 \rfloor}^{\lfloor N/2 \rfloor} X^i[n] \sum_{l=-\infty}^{\infty} \delta[n - (m + lM^i)], \tag{4.12}$$

where $Y^i[m]$ denotes the sub-Nyquist rate DFT spectrum in the i-th channel, and $X^i[n]$ denotes the non-aliased DFT spectrum. In matrix form this becomes,

$$\overrightarrow{Y}^i = \mathbf{\Phi^i} \overrightarrow{X}^i, \tag{4.13}$$

where $\overrightarrow{Y}^i$ denotes the sub-Nyquist rate DFT spectrum vector in the i-th channel, $\overrightarrow{X}^i$ is the non-

aliased DFT spectrum vector, and the element of the linear projection operator $\mathbf{\Phi^i} \in \mathbb{R}^{M^i \times N}$ ($M^i < N$) can be represented as,

$$\phi_{m+\lfloor M^i/2 \rfloor+1, n+\lfloor N/2 \rfloor+1} = \frac{M^i}{N} \sum_{l=-\infty}^{\infty} \delta\big[n - (m + lM^i)\big], \tag{4.14}$$

where $m \in \mathbb{Z} \cap [-M^i/2, M^i/2)$, and $n \in \mathbb{Z} \cap [-N/2, N/2)$.

It is easy to see that in each column of $\mathbf{\Phi^i}$, there is only one non-zero element with value of $\frac{M^i}{N}$. In each row of $\mathbf{\Phi^i}$, there exists at most $\lceil \frac{N}{M^i} \rceil$ (ceil function gives the smallest integer not less than $\frac{N}{M^i}$) non-zero elements, which is also called the undersampling factor.

4.2.3 Effect of Sub-Nyquist Sampling

Sampling a signal at a sub-Nyquist sampling rate generates two issues. First, the exact location of the signals for those who have frequencies larger than the sub-Nyquist sampling rate is lost. Second, there is a risk of overlap, i.e. different frequencies are down-converted to the same frequency in the baseband. This is crucial, because it could lead to signal cancellation, and hence missed signal detection in the cognitive radio application. However, under certain assumptions, signal overlap has a very small probability of occurring. They are 1) the non-aliased DFT spectrum is k-sparse, 2) the number of subsamples in the i-th channel, $M^i \sim \mathcal{O}(\sqrt{N})$, for simplicity, let $M^i = \sqrt{N}$, and 3) $k \ll N$.

Assuming those k spectral components are i.i.d. over the frequency bins of $0, 1, \cdots, N-1$, the probability of one element in the non-aliased DFT spectrum being non-zero is $P = \Pr(X[n] \neq 0) = k/N$. Let $q[m]$ denote the number of signals overlapped at $Y^i[m]$, using (4.12) the probability of no signal overlap is given by,

$$\begin{aligned}\Pr(q[m] < 2) &= \Pr(q[m] = 0) + \Pr(q[m] = 1) = (1-P)^{\lceil \frac{N}{M^i} \rceil} + \binom{\lceil \frac{N}{M^i} \rceil}{1} P(1-P)^{\lceil \frac{N}{M^i} \rceil - 1} \\ &\approx (\frac{N-k}{N})^{\frac{N}{M^i}} + \frac{k}{M^i}(\frac{N-k}{N})^{\frac{N-M^i}{M^i}} = \frac{(\frac{N-k}{N})^{\sqrt{N}}(N-k+k\sqrt{N})}{N-k}.\end{aligned} \tag{4.15}$$

As shown in Figure 4.3, when $N = 9 \times 10^6$, and $M^i = \sqrt{N}$, the probability of no signal overlap at $Y^i[m]$ will be $\Pr(q[m] < 2) = 95 \sim 100\%$ for any $k \leq 1000 = M^i/3$, and $\Pr(q[m] < 2) \geq 99\%$ for any $k \leq 400$.

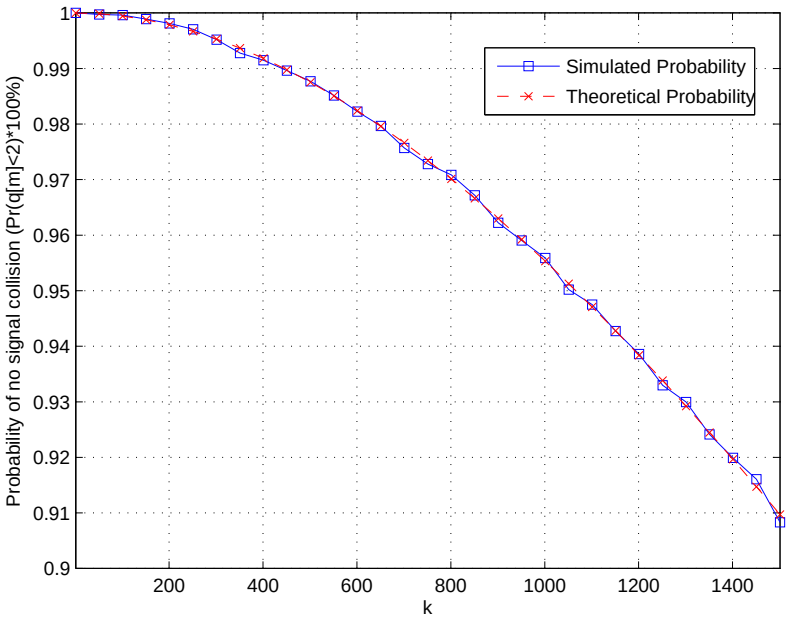

Figure 4.3: *Simulated probability of no signal overlap at a frequency $f = m\Delta f + \varsigma$ in the sub-Nyquist rate spectrum $\vec{Y}$, compared to the theoretical result in (4.15). The number of samples at the sub-Nyquist sampling rate is $M^i = \sqrt{N} = 3000$, where $N = 9 \times 10^6$ denotes the number of samples at the Nyquist sampling rate.*

Thus, under assumptions mentioned above, with the aid of (4.13), the following equation holds,

$$|\overrightarrow{Y}^i| = |\mathbf{\Phi^i}\overrightarrow{X}^i| \simeq \mathbf{\Phi^i}|\overrightarrow{X}^i|. \tag{4.16}$$

Equation (4.16) holds true because when $Y^i[m]$ is the projection of a single signal, the following equation holds,

$$|Y^i[m]| = \left|\frac{M^i}{N}\sum_l X^i[m + lM^i]\right| = \frac{M^i}{N}\left|X^i[m + lM^i]\right| = \mathbf{\Phi^i}\left|X^i[m + lM^i]\right|. \tag{4.17}$$

In the rare case where signal overlap occurs, i.e. $0 \leq |\overrightarrow{Y}^i| \leq \mathbf{\Phi^i}|\overrightarrow{X}^i|$, the spectral content of the full spectrum can still be tracked with the aid of other sampling channels [33]. This is because the blind spot of one ADC can be illuminated by other ADCs working at different sampling rates.

4.2.4 Recovery of the Full Spectrum via Multirate Sampling

This section will now introduce a method for reconstructing the full spectrum through multirate sampling. Since all ADCs are observing the same magnitude of the spectrum, i.e. $\overrightarrow{X} \triangleq |\overrightarrow{X}^1| = |\overrightarrow{X}^2| = \cdots = |\overrightarrow{X}^v|$, a concatenated equation relating $\overrightarrow{X} \in \mathbb{R}^N$ to $\overrightarrow{Y} \in \mathbb{R}^M$ can be formed as below,

$$\overrightarrow{Y} \triangleq \begin{pmatrix} |\overrightarrow{Y}^1| \\ |\overrightarrow{Y}^2| \\ \vdots \\ |\overrightarrow{Y}^v| \end{pmatrix} \simeq \mathbf{\Phi}\overrightarrow{X} = \begin{pmatrix} \mathbf{\Phi^1} \\ \mathbf{\Phi^2} \\ \vdots \\ \mathbf{\Phi^v} \end{pmatrix} \overrightarrow{X} \tag{4.18}$$

where $\mathbf{\Phi^1}, \mathbf{\Phi^2}, ..., \mathbf{\Phi^v}$ are disjoint submatrices of $\mathbf{\Phi}$, $\mathbf{\Phi^i} \in \mathbb{R}^{M^i \times N}$ is the measurement matrix of ADC i, and $M \triangleq \sum_{i=1}^{v} M^i$. The time offset between ADCs is required to be sufficiently small in the sense that the observed spectra are quasi-stationary. Thus, it will not influence the magnitude of the non-aliased spectrum, i.e. (4.18) holds true for asynchronous ADCs.

When certain conditions are satisfied as Proposition 4.1 below, the mutual coherence of the concatenated measurement matrix $\mathbf{\Phi}$ will be determined by the number of samples in each channel, and the number of channels v.

Proposition 4.1: When v ADCs observe the spectral magnitude, i.e. $\overrightarrow{X}$, in the same observation time of T, generating v measurement vectors, $|\overrightarrow{Y}^1|, |\overrightarrow{Y}^2|, \cdots, |\overrightarrow{Y}^v|$, whose length are different primes, $M^1, M^2, \cdots, M^v$, which satisfy,

$$M^l M^z > N, \quad \forall l, z \in [1, v], l \neq z, \tag{4.19}$$

then the mutual coherence μ of the measurement matrix $\mathbf{\Phi}$ is determined by,

$$\mu = \max_{j \neq h} | < \hat{\phi}_j, \hat{\phi}_h > | = \frac{\max_{i \in [1,v]} (M^i)^2}{\sum_{l=1}^{v} (M^l)^2}. \tag{4.20}$$

The proof of Proposition 4.1 is given in Appendix A.

From the discussion in Section 4.1, the full spectrum magnitude vector, i.e. $\overrightarrow{X}$, can be reconstructed when the mutual coherence of the measurement matrix is less than $\frac{1}{2k-1}$. Using (4.20),

with manipulation, this condition can be written as,

$$\sum_{l=1}^{v} \left(\frac{M^l}{\max_{i\in[1,v]} M^i} \right)^2 > 2k - 1. \tag{4.21}$$

Equation (4.19) illustrates that the number of samples in each ADC is of the order of $\sqrt{N}$. When $M^i \sim \mathcal{O}(\sqrt{N})$, it shows in (4.21) that at least $2k$ channels are required to recover the non-aliased spectrum $\overrightarrow{X}$. This is because $v \geq \sum_{l=1}^{v} \left(\frac{M^l}{\max_{i\in[1,v]} M^i} \right)^2$ holds true for all $v \in \mathbb{Z}^+$. Considering both (4.19) and (4.21), it is found that MASS needs the total number of observations to be $M = \sum_{i=1}^{v} M^i \sim \mathcal{O}(k\sqrt{N})$.

4.2.5 Practical Implementation Issues

This section will analyse influence from several factors when using MASS in cognitive radio networks, i.e. average sampling rate, number of sampling channels, noise level, and model mismatching. As discussed in Section 4.2.4, in order to recover the full spectrum with high probability, $2k$ sampling channels are required. For any number of sampling channels, the probability of successful spectrum recovery is determined by the average sampling rate.

Proposition 4.2: If a wideband signal $x(t)$, whose full spectrum is k-sparse, is sampled at different sampling rates, and the length of observations are different prime numbers, $M^1, M^2, ..., M^v$ which satisfy,

$$M^l M^z > N, \quad \forall l, z \in [1, v], l \neq z, \tag{4.22}$$

$$\frac{\sum_{l=1}^{v} (M^l)^2}{\sum_{l=1}^{v} M^l} > \frac{2k-1}{\epsilon}, \quad \forall v \in [2, \infty), \tag{4.23}$$

recovery algorithms such as BP, and OMP, have a probability of at least $(1 - \epsilon)$ of successfully reconstructing the full spectrum magnitude vector $\overrightarrow{X}$.

The proof of Proposition 4.2 is given in Appendix B.

In Proposition 4.2, it is found that the probability of successful spectrum recovery, i.e. $1 - \epsilon$, increases when the average sampling rate grows. Adding more sampling channels with a higher sampling rate will lead to a larger probability of successful spectrum recovery as the term $\frac{\sum_{l=1}^{v} (M^l)^2}{\sum_{l=1}^{v} M^l}$ increases. This is because $\frac{a^2}{a} \leq \frac{a^2+b^2}{a+b}$ when $0 < a \leq b$.

When noise is present, similarly to (4.18), a concatenated equation can be formed as,

$$\overrightarrow{Y} = \begin{pmatrix} |\overrightarrow{Y}^1| \\ |\overrightarrow{Y}^2| \\ \vdots \\ |\overrightarrow{Y}^v| \end{pmatrix} \simeq \mathbf{\Phi}\overrightarrow{X} + \vec{z} = \begin{pmatrix} \mathbf{\Phi^1} \\ \mathbf{\Phi^2} \\ \vdots \\ \mathbf{\Phi^v} \end{pmatrix} \overrightarrow{X} + \vec{z}. \tag{4.24}$$

The performance of spectrum recovery is related to the noise level and the model mismatching as follows:

Proposition 4.3: In the case of $\|\vec{z}\|_2 \leq \rho$, using LPBP for recovering $\overrightarrow{X}$ from (4.24), the recovered spectrum $\widetilde{X}$ satisfies,

$$\|\widetilde{X} - \overrightarrow{X}\|_2 \leq \frac{A}{\sqrt{k}}\|\overrightarrow{X} - \overrightarrow{X}^{\text{opt}}\|_1 + B\rho \tag{4.25}$$

if the number of samples $(M^1, M^2, ...M^v)$ in v channels satisfy:

1. $M^1, M^2, ...M^v$ are different prime numbers,
2. $M^l M^z > N, \quad \forall l, z \in [1, v], l \neq z$,
3. $\sum_{l=1}^{v}(\frac{M^l}{\max_{i\in[1,v]} M^i})^2 > \frac{2k-1}{\sqrt{2}-1}$,

where $\overrightarrow{X}^{\text{opt}}$ denotes the optimal k-sparse approximation of $\overrightarrow{X}$.

The proof of Proposition 4.3 is given in Appendix C.

It should be emphasised that Proposition 4.3 not only considers the effect of noise, but also shows the influence of model mismatching in the first expression. It illustrates that the performance of MCSS is stable when model mismatching occurs. This is because the model mismatch has a very small influence to the recovery error, i.e. $\|\widetilde{X} - \overrightarrow{X}\|_2$, due to the small coefficient $\frac{A}{\sqrt{k}}$.

4.3 Extension to Distributed Wideband Spectrum Sensing

In wireless communication scenarios, some cognitive radio users may suffer from fading either due to multipath propagation, or due to shadowing. In that case, the result of spectrum sensing

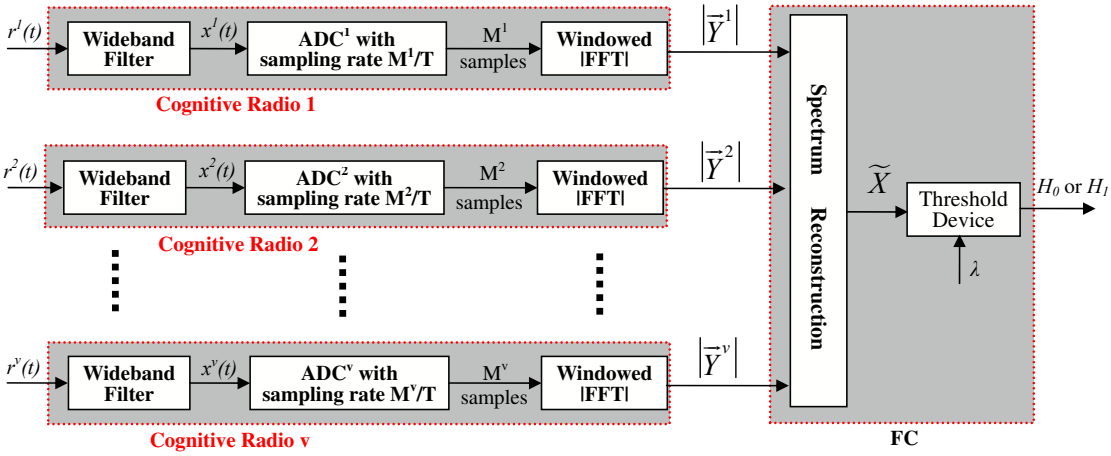

Figure 4.4: *The schematic illustration of the distributed MASS system in cognitive radio networks. Each cognitive radio is only required to be equipped with one wideband filter, one low-rate ADC, and an FFT. The bandwidth of the wideband filters should be altered to have the same bandwidth of W.*

from a single cognitive radio node is not reliable. A cooperative detection strategy offers a good solution as it minimises the effect of severe fading and achieves diversity gain [21]. If every cognitive radio forwards its measured or processed data to the FC, which makes a final decision based on the collected data, the cooperative scheme is often called data fusion based distributed spectrum sensing. If each cognitive radio uses multiple ADCs to perform sub-Nyquist sampling, the transmission of the measurements may be very expensive in distributed cognitive radio networks.

To minimise the required transmission bandwidth, this section proposes to extend the application of the MASS system from the case of parallel ADCs in one cognitive radio node to single ADCs in multiple cognitive radio nodes as shown in Figure 4.4. It is assumed that cognitive radios are situated in different locations. Each cognitive radio performs spectrum sensing (without decision) independently. The radios are sufficiently far apart that the spectrum sensing data can be assumed to be i.i.d. The FC then gathers these spectrum sensing data for making decision. The implementation of the distributed MASS system in cognitive radio networks is illustrated in Figure 4.5. Note that, as in the analysis of Section 4.4, the time offset between ADCs will not change the magnitude of the full spectrum, thus asynchronous ADCs have the same performance as synchronous ADCs. The wideband filters prior to the ADCs are assumed to have the same bandwidth of W. After measurements in each cognitive radios, the magnitude of the sub-Nyquist rate spectra are transmitted to the FC, where the full spectrum is

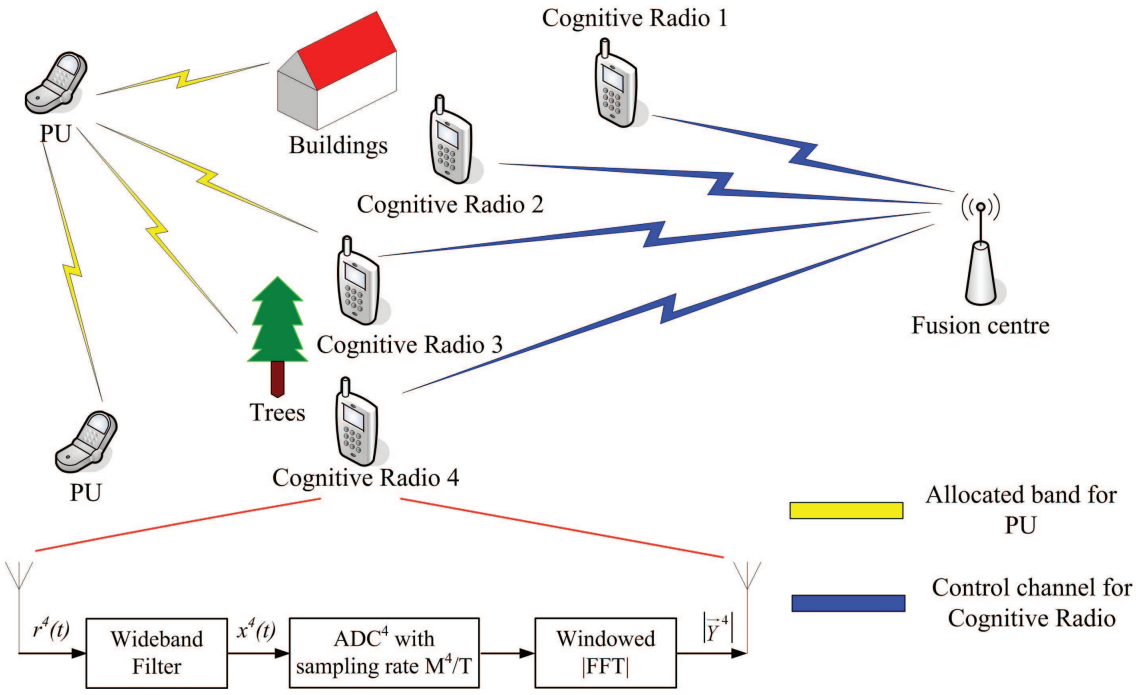

Figure 4.5: *Implementation of the distributed MASS system in cognitive radio networks. Each ADC senses the wideband spectrum at different sub-Nyquist sampling rate, and then sends the spectrum sensing data through a control channel to a fusion centre, where the full spectrum is reconstructed.*

reconstructed.

The advantage of the distributed MASS system is that only one low-rate ADC is required in each cognitive radio node, which not only simplifies the system structure at each cognitive radio node, but also decreases the bandwidth required for sharing spectrum sensing data. Section 4.2.4 shows that the total number of observations to be transmitted is of the order of $k\sqrt{N}$. Even though this is more than conventional CS, i.e. $M \sim \mathcal{O}(k\log(N))$, MASS is more amenable to implementation in a distributed cognitive radio network as neither a compression device nor a measurement matrix generator are required. Moreover, as the measurement matrix used in MASS is deterministic, then the transmission and storage of the measurement matrix is unnecessary. Another important advantage of the distributed system is the diversity gain. As cognitive radios are assumed to be in different locations, some cognitive radios may identify bands which are deeply faded (due to the shadowing effect) at a particular cognitive radio. Nevertheless, the disadvantage is that (4.18) no longer holds because of the influence of fading. Thus, conventional CS algorithms, such as BP and OMP, cannot be applied. In fading channels, the power of the signals coming from PUs are attenuated. Namely, the received signals at

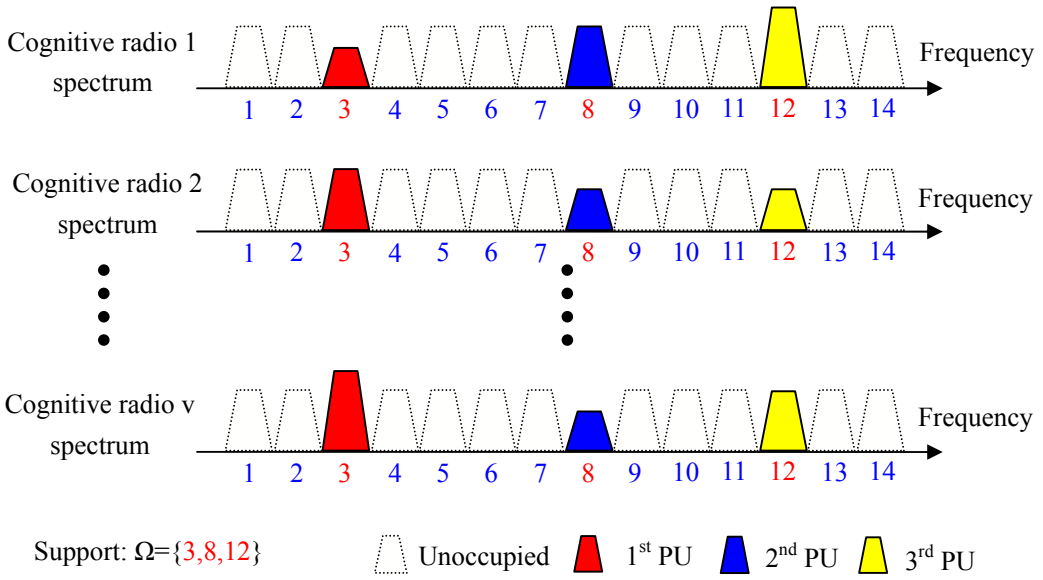

Figure 4.6: *Illustration of the spectral correlation in distributed cognitive radio nodes.*

distributed cognitive radio nodes, i.e. $r^1(t), r^2(t), \cdots r^v(t)$, may be different, and the spectra viewed by distributed cognitive radios therefore are often distinct. As illustrated in Figure 4.6, the spectra over the distributed cognitive radio nodes can be modelled by the second joint sparsity model (JSM-2) in [89]. Specifically, the spectra over cognitive radios have a common spectral support, Ω (the set of frequencies occupied), but with different amplitudes.

In order to exploit the joint sparse property over sensor nodes, DCS-SOMP was proposed in [89]. It has been observed that this algorithm requires fewer measurements when compared with a separate recovery approach. However, the drawback of DCS-SOMP is its calculation time as it only selects the maximum support in each iteration. Unlike the conventional greedy algorithms, CoSaMP [93] accelerates the calculation by identifying many possible solutions in each iteration. Hence, a hybrid matching pursuit algorithm is applied to the MASS system as shown in Table 4.1, by synthesizing DCS-SOMP and CoSaMP. In each iteration, the common support is boosted by summing up the correlation vectors from multiple channels, which will make the features easy to identify even if fading exists in some channels. Then multiple indices are selected in each iteration by choosing the top-$2k$ indices of the $2k$-largest values in the combined correlation vector, and merged with the previously computed support. After that, the full spectrum is recovered by least squares, where $\mathbf{\Phi}_{\mathbf{\Omega}}^{\dagger}$ denotes the pseudoinverse of $\mathbf{\Phi}_{\mathbf{\Omega}}$. Since the matrix $\mathbf{\Phi}_{\mathbf{\Omega}}$ is always well conditioned, the calculation of the pseudoinverse $\mathbf{\Phi}_{\mathbf{\Omega}}^{\dagger}$ can be im-

Input: measurement matrix $\mathbf{\Phi}$, spectral observation vector $\overrightarrow{Y}$, sparsity level k, and noise tolerance level ρ.
1. *Initialise:* Residual $\vec{r}_0 = \overrightarrow{Y}$, approximation $\overrightarrow{X}_0 = 0$, support $\Omega_0 = [\,]$, $j = 0$,
2. *While* halting criterion is false *do*
1). Form residual correlation vector by, $\overrightarrow{R}^i = (\mathbf{\Phi^i})^T \vec{r}_j^{\,i}$ for $i \in [1, v]$,
b). Find spectral support Ξ_j by boosting $\Xi_j = \text{Supp}_{2k}(\sum_{i=1}^{v} \overrightarrow{R}^i)$,
c). Merge support with previously computed support by $\Omega_j = \Omega_{j-1} \bigcup \Xi_j$,
d). Approximate the non-aliased spectrum by least squares $\overrightarrow{X}_j = \mathbf{\Phi_\Omega}^\dagger \overrightarrow{Y}$ for column index belong to Ω_j, other columns set to 0,
e). Prune locations of support by $\Theta = \text{Supp}_k(\overrightarrow{X}_j)$,
f). Update individual residual using a part of $\overrightarrow{X}_j$ with support Θ as $\vec{r}_j^{\,i} = \lvert\overrightarrow{Y}^i\rvert - \mathbf{\Phi^i}\overrightarrow{X}_{j\Theta}$ for $i \in [1, v]$,
g). $j = j + 1$.
3. *Halting criterion:* $\lVert\vec{r}^{\,i}\rVert_2 \leq \rho$
Output: A k-sparse approximation of the non-aliased spectrum, $\widetilde{X} = \overrightarrow{X}_{j\Theta}$.

Table 4.1: *Hybrid matching pursuit algorithm for distributed MASS system*

plemented quickly using an iterative method, such as the conjugate gradient method [94]. The support is pruned to have k-largest values, followed by updating the residuals. The algorithm will be halted when the ℓ_2 norm of the residual is not larger than the noise tolerance level ρ.

4.4 Simulation Results

In the experiments, it is assumed that each cognitive radio is equipped with a single low-rate ADC, and there are v cognitive radio nodes allocated in the same cluster. In the i-th cognitive radio node, the wideband signal $x^i(t)$, which is defined below, is sampled by a low-rate ADC over an observation time, T.

$$x^i(t) = \sum_{l=1}^{N_b} \sqrt{E_l^i} B_l \cdot \text{sinc}(B_l(t - \Delta_l)) \cdot \cos(2\pi f_l(t - \Delta_l)), \tag{4.26}$$

where Δ_l denotes the time offset of the signal, set to be $T/2$, and E_l^i denotes the energy of the l-th subband viewed by the i-th cognitive radio. Since the effect of fading, E_l^i varies subject to

the property of the fading channel. The wideband signal $x^i(t)$ consists of N_b subbands, whose bandwidth is B_l, and carrier frequency is f_l. The values of simulation parameters are specified in Table 4.2. As described in Fig. 4.4, after FFT analysis in each node, the sub-Nyquist rate spectral observations $|\overrightarrow{Y}^1|, |\overrightarrow{Y}^2|, ..., |\overrightarrow{Y}^v|$, are gathered at the FC. In the FC, the full spectrum is reconstructed using HMP. The spectral occupancy status is decided based on a hypothesis test on the reconstructed spectrum. The compression ratio is defined as $\frac{\sum_{i=1}^{v} M^i}{N}$.

	Figure 4.7	Figure 4.8	Figure 4.9	Figure 4.10
Observation time T	0.4 μs	2 μs	2 μs	2 μs
Wideband bandwidth W	5 GHz	20 GHz	20 GHz	20 GHz
Length of Nyquist samples N	4,000	80,000	80,000	80,000
No. of subbands N_b	5	10, 30, 50	30, 50	30
Bandwidth of subband (MHz)	$0.3 \sim 30$	$0.1 \sim 5$	$0.1 \sim 5$	$0.1 \sim 5$
No. of ADCs v	10	1$\sim$150	1$\sim$150	1$\sim$150
Sampling rates of ADCs (MHz)	182.5 $\sim$ 282.5	189.5$\sim$711.5	189.5$\sim$711.5	189.5$\sim$590.5
Compression ratio	23.85%	$0.47 \sim 163.45\%$	$0.47 \sim 163.45\%$	$0.47 \sim 163.45\%$

Table 4.2: *List of simulation parameters setup for Figure 4.7, Figure 4.8, Figure 4.9, and Figure 4.10.*

Figure 4.7 demonstrates the performance of spectrum recovery in the MASS system using HMP. It shows that by using $v = 10$ ADCs, the full spectrum, which consists of 5 subbands with bandwidth $0.3 \sim 30$ MHz, can be successfully reconstructed. It should be emphasised that, instead of working at or above 10 GHz, these ADCs are working at sampling rate of $182.5 \sim 282.5$ MHz (ADCs with different sampling rates), and the total number of measurements is $23.85\%N$ (N is the number of samples if $x^i(t)$ is sampled at Nyquist rate in one cognitive radio node). Obiviously, the sampling rate for characterising the wideband spectrum in the MASS system is much lower than the case of single node Nyquist sampling. Besides, in comparison to the filter bank method, which requires at least N measurements for wideband spectrum sensing,

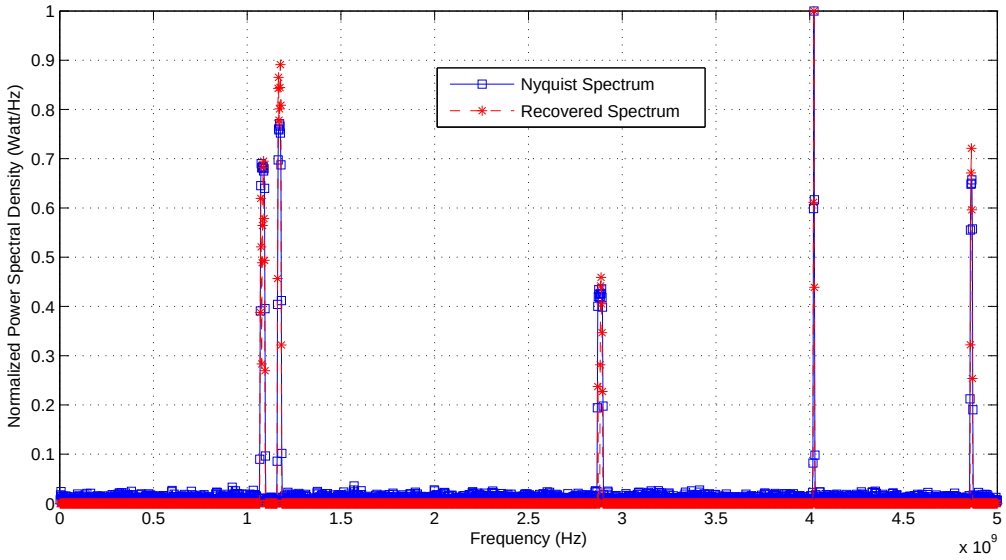

Figure 4.7: *Comparison of the non-aliased DFT spectrum with the recovered spectrum, when the wideband signal has* 5 *subbands with bandwidth* $0.3 \sim 30$ MHz*, and SNR*$= 15$ dB*. There are* $v = 10$ *ADCs, which experience non-fading AWGN channels. The total number of measurements is* $23.85\%N$*, rather than* N *when sampled at the Nyquist rate.*

the proposed MASS system needs fewer measurements. Thus, if the spectrum sensing data is shared in distributed cognitive radio networks, MASS consumes less transmission power.

Figure 4.8 depicts the influence of the number of subbands and the compression ratio on the detection performance of MASS. It illustrates that the fewer subbands there are, the better the detection performance one can achieve. If a cognitive radio system has constraint on the probability of false alarm, e.g., $P_f \leq 10\%$, the minimum number of measurements are $0.1N$, $0.3N$, and $0.55N$ for $N_b = 10$, $N_b = 30$, and $N_b = 50$, respectively. On the other hand, the probability of detection is more important, it is found that to obtain $P_d \geq 90\%$, at least $0.2N$ measurements are required to sense the wideband spectrum with up to 50 subbands. In addition, one can notice that a higher compression ratio, propotional to the number of measurements, will result in a smaller probability of false alarm and a larger probability of detection.

Figure 4.9 shows the effect of imperfect synchronisation between ADCs. Compared with a reference clock, the asynchronous ADCs have time offsets in range of $0 \sim 0.8$ μs, while the total observation time is 2 μs. It is evident that the detection performance of the asynchronous ADCs is roughly the same as that of the synchronous ADCs. Fig. 4.9 also illustrates that with

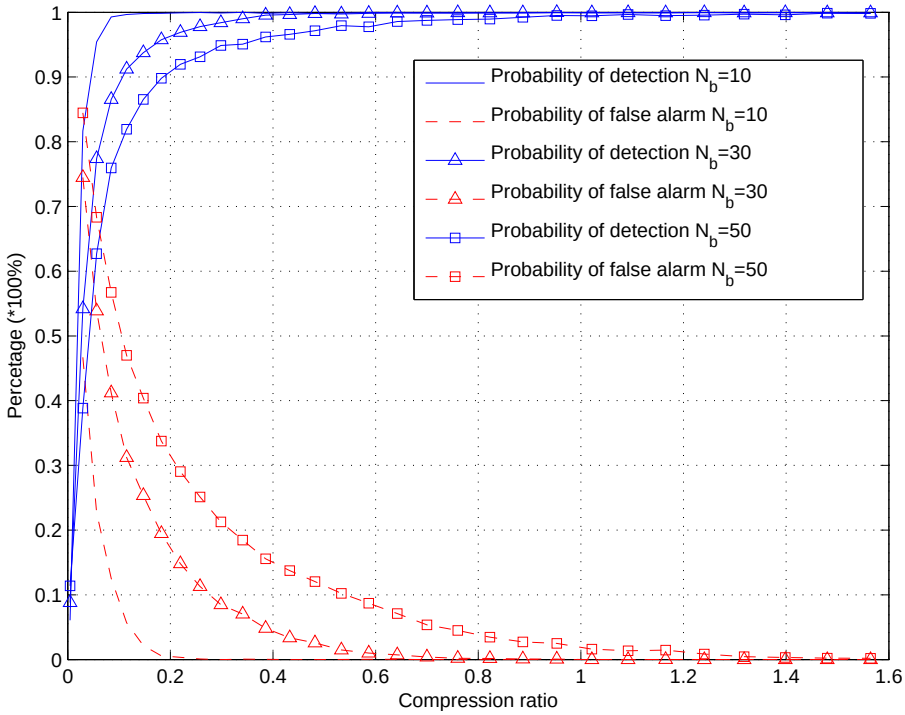

Figure 4.8: *Influence of the number of sub-bands, and the compression ratio on the detection performance of the MASS system, when the signals from PUs have* 10*,* 30*, and* 50 *sub-bands with bandwidth* $0.1 \sim 5$ MHz*, and experience non-fading AWGN channels with SNR*$= 15$ dB.

more ADCs in collaboration, a better spectrum sensing performance can be achieved. This is because with more sampling channels, a higher probability of successful spectrum recovery can be obtained as discussed in Section 4.2.5.

In order to quantify the detection performance of the proposed system over fast/slow fading channels, Figure 4.10 compares their detection performance with that over AWGN channels. It shows that, compared with AWGN case, more cognitive radios need to collaborate in order to combat the effect of fading. Specifically, to obtain $P_f = 10\%$, 40 cognitive radios are required to collaborate over AWGN channels, 50 cognitive radios are required for Rayleigh fading channels, and 60 cognitive radios are necessary for slow fading channels. Note that 60 ADCs in MASS can be translated to a compression ratio of 42.32%, which means that only $0.4232N$ measurements are needed to transmit to the FC through the control channel, rather than at least $2N$ (two users in collaboration) for a conventional collaborative spectrum sensing cognitive radio network.

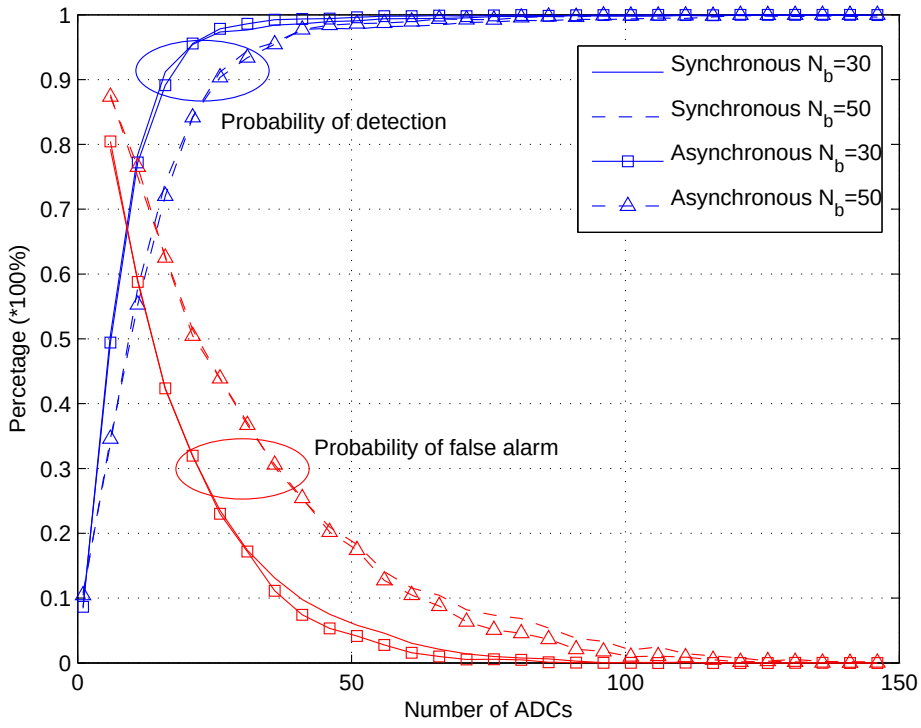

Figure 4.9: *Comparison of spectrum recovery performance for synchronous ADCs and asynchronous ADCs, with $N_b = 30$ and $N_b = 50$. The asynchronous ADCs have time offsets in range of $0 \sim 0.8$ μs, with a total observation time of 2 μs.*

4.5 Discussion

This section describes the connections between MASS and other undersampling approaches, which could be used for wideband spectrum sensing in distributed cognitive radio networks.

4.5.1 Relationship with Multirate Sampling

The sampling pattern of MASS can be implemented by both SMRS and MRS, by controlling the number of samples being different primes in each channel over an observation time T. However, it is difficult to meet the sufficient conditions for the MRS either in a single cognitive radio node or distributed cognitive radio nodes, because of its stringent requirements on electro-optical hardware and synchronisation.

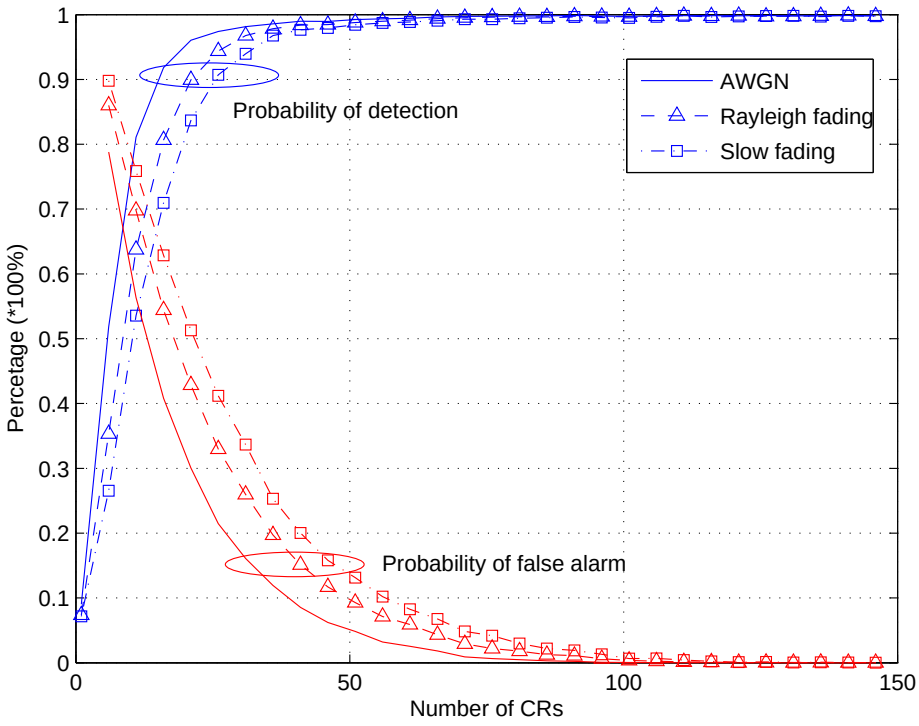

Figure 4.10: *Performance of MASS over AWGN, Rayleigh, and slow fading channels, with the number of sub-bands* $N_b = 30$*. In this experiment, the average SNR is* 15 dB*, and the standard deviation of log-normal fading is* 5 dB.

4.5.2 Connection with Multicoset Sampling

Some of the undersampling techniques are related to multicoset sampling [31]. In a multicoset scheme, v out of L ($v \leq L$) cosets of samples are chosen to reconstruct the signal. It can be implemented by using multiple sampling channels, which are offset by an integral multiple of a constant time. In order to reconstruct the signal with a high success rate, the number of sampling channels must be sufficiently high [32]. Note that even though the sampling pattern of MASS can also be implemented using a multicoset scheme, it is impractical to implement it this way in a distributed cognitive radio network as the synchronisation requirements will be extremely high.

4.5.3 Comparison with CS-based Models

Most CS based approaches, for example, [29, 35, 69], require Pseudo-random sequence generator as the measurement matrix generator. In order to use CS based approaches for wideband

spectrum sensing in a distributed cognitive radio network, the storage and transmission of the measurement matrix must be addressed. To exploit space diversity, each cognitive radio node should have a separate compression device, transmit both compressed data and their measurement matrix to the FC. This not only increases the implementation complexity of cognitive radio nodes, but also burdens the control channel. In contrast, when using MASS for wideband spectrum sensing, the measurement matrix is deterministic, thus the problem of storage and transmission of the measurement matrix is solved. Moreover, no separate device is required to generate the measurement matrix.

4.6 Conclusions

This chapter has presented a distributed wideband spectrum sensing model, MASS, which employs several low-rate ADCs to sample the wideband signal at different sub-Nyquist rates. The MASS system could be applied either in the scenario of multiple ADCs in one sensor or in the case of single ADCs in multiple sensors. The sufficient conditions have been derived to uniquely recover the full spectrum by the multirate asynchronous sub-Nyquist sampling. When the spectra of the cognitive radio nodes are jointly sparse in the distributed cognitive radio network, the HMP algorithm has been applied to recover the common Nyquist spectrum by the boost-and-recover strategy. If MASS is used in a distributed manner, it has very low implementation complexity in each cognitive radio node. Moreover, MASS is energy-efficient as it requires less transmission bandwidth than conventional collaborative spectrum sensing.

Simulation results have shown that MASS can recover the wideband Nyquist spectrum successfully. Besides, MASS has been shown to have a very robust property against the lack of synchronisation between ADCs or cognitive radios. When using the HMP algorithm, the MASS system is applicable to the distributed cognitive radio network. It has been shown that only a few more cognitive radios are needed to collaborate to combat the effect of fading.

Chapter 5
Distributed Wideband Spectrum Detection

In previous chapter, the spectral occupancy status can be obtained by a binary hypothesis test on the reconstructed full spectrum. However, the purpose of wideband spectrum sensing is to detect the presence or absence of the PUs, therefore, the task of wideband spectrum sensing does not necessarily require a full reconstruction of the spectrum, but only requires the estimation of relevant test statistics. It is possible to directly extract these test statistics from a small number of projections without ever reconstructing the full spectrum. This chapter proposes a novel model for directly detecting the PUs from a small number of measurements without reconstructing the wideband spectrum, called wideband spectrum detection, with the following contributions:

- A multirate spectrum detection (MSD) system is introduced to detect the PUs from a small number of measurements without reconstructing the full spectrum. This system has a low implementation complexity, and a low transmission cost for sharing spectrum sensing data.
- A data fusion strategy is developed to achieve a better detection performance. Compared with MASS, MSD has a low computational complexity at the FC because the full spectrum is not reconstructed.
- The detection performance of the proposed MSD system is analysed in the presence of AWGN, Rayleigh, and slow fading channels.

This chapter is organized as follows. A standard spectrum estimation method for wideband spectrum sensing is introduced in Section 5.1. In Section 5.2, MSD is proposed for wideband spectrum detection. The effect of sub-Nyquist sampling in a single cognitive radio node is considered, followed by the data fusion rule for multiple cognitive radio nodes. In Section 5.3, the detection performance of the MSD system over fading channels is analysed, and some

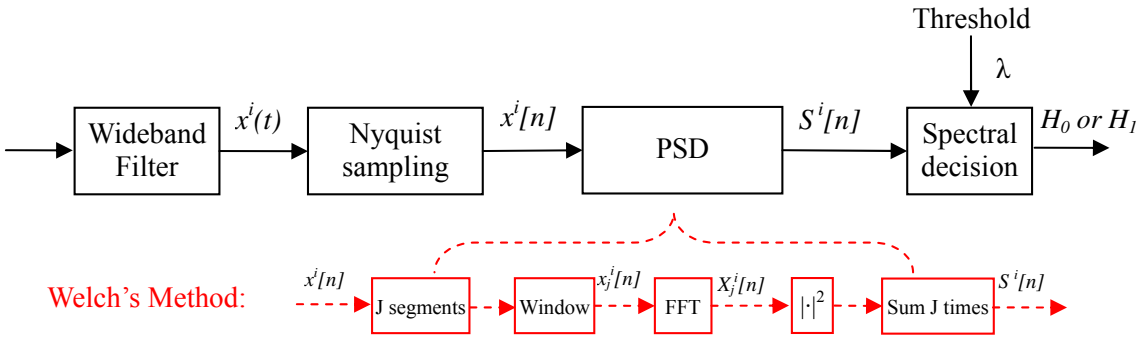

Figure 5.1: *Schematic illustration of wideband spectrum estimation and detection. The Welch's method is employed for calculating the PSD.*

theoretical bounds on the detection performance of the proposed system are derived. Simulation results are presented in Section 5.4, with conclusions given in Section 5.5.

5.1 Preliminary

Spectrum sensing refers to performing measurements on the RF spectrum and forming decisions related to spectrum usage based upon the measured data. Three steps are involved in forming a standard wideband spectrum estimation and detection as shown in Figure 5.1. They are 1) sample the received signal at or above the Nyquist rate over a time window of T_{all} with N_{T} samples, 2) compute PSD, and 3) compare it with a predetermined threshold, λ, to decide whether the corresponding frequency is occupied or not. After sampling, the received sequence at cognitive radio node i, $x^i[n]$ ($n = 0, 1, \cdots, N_{\text{T}} - 1$), can be modelled by a hypothesis test with H_0 (signal not present) or H_1 (signal present),

$$x^i[n] = \begin{cases} z[n], & H_0 \\ h^i\ s[n] + z[n], & H_1 \end{cases} \tag{5.1}$$

where h^i stands for the complex channel gain between the PU and the cognitive radio node i, $s[n]$ denotes transmit signal from the PU, and $z[n]$ is AWGN. For simplicity, the noise variance in each cognitive radio node is assumed to be the same, i.e. $z[n] \sim \mathcal{N}(0, \delta^2)$, where $\mathcal{N}(0, \delta^2)$ denotes a normal distribution with zero mean and variance of δ^2.

Using Welch's method [95], the received sequence is divided into J overlapping segments. The

j-th data segment, $x_j^i[n]$, can be represented as,

$$x_j^i[n] = w_T[n]x^i[n + jR], \quad n \in [0, N),\ j \in [0, J), \tag{5.2}$$

where $w_T[n]$ denotes the window, J is the number of segments, N is the number of samples when sampled at the Nyquist rate, and $R \leq N$ is the window hop size. Without loss of generality, let $R = N$, thus $N_{\mathrm{T}} = JN$. According to (5.1), the non-aliased DFT spectrum (full spectrum) in the j-th window, $X_j^i[n]$, follows a complex normal distribution given by,

$$X_j^i[n] \sim \begin{cases} \mathcal{CN}(0, \delta_n^2), & H_0 \\ \mathcal{CN}(s^i, \delta_n^2), & H_1, \end{cases} \tag{5.3}$$

where s^i denotes the received spectral component from the PU at the cognitive radio node i, δ_n^2 is the noise variance in the non-aliased DFT spectrum, and $\mathcal{CN}(s^i, \delta_n^2)$ denotes a complex normal distribution with mean of s^i and variance of δ_n^2. When $w_T[n]$ is a rectangular window, the noise variance in the full spectrum can be linked to the noise variance δ^2 in (5.1) by $\delta_n^2 = \frac{N}{2}\delta^2$, because the DFT operation is a linear operation. On the other hand, when different window function is used, the noise variance, δ_n^2, will be different. The output noise variance can be determined by both input noise variance and equivalent noise bandwidth of the window function [96, 97].

The scaled PSD observed by the cognitive radio i, i.e. $S^i[n]$, is calculated by,

$$S^i[n] = \sum_{j=1}^{J} |X_j^i[n]|^2, \quad n \in [0, N). \tag{5.4}$$

By (5.3) and (5.4), the distribution of the PSD can be modelled by central and non-central chi-square distributed random variables as,

$$S^i[n] \sim \begin{cases} \chi_{2J}^2, & H_0 \\ \chi_{2J}^2(2J\gamma^i), & H_1. \end{cases} \tag{5.5}$$

where γ^i denotes the SNR at the i-th cognitive radio node.

If the PSD, $S^i[n]$, is compared with a predetermined threshold, λ, to decide whether the PU is present or not, the probabilities of false alarm and detection can be obtained by using (5.5)

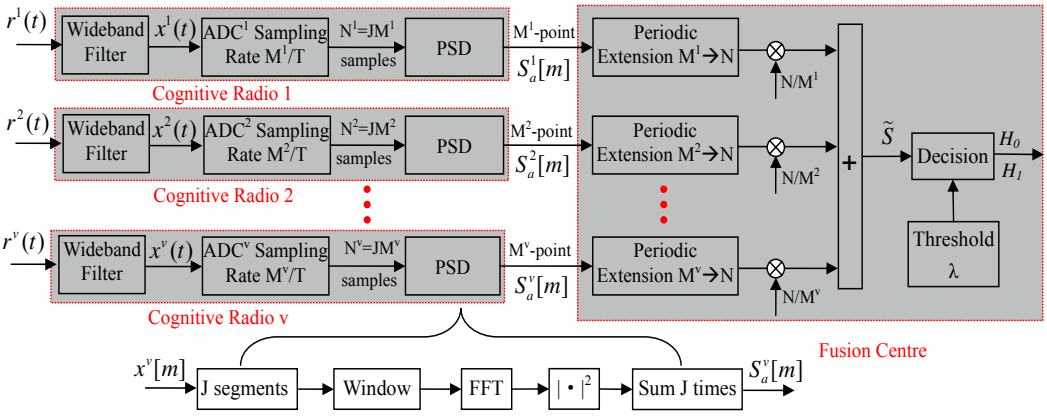

Figure 5.2: *Schematic illustration of the multirate spectrum detection system in a distributed cognitive radio network. Wideband filters prior to ADCs are altered to have the same bandwidth of W, and the wideband signal is sampled at different sub-Nyquist rates in different cognitive radio nodes.*

as [52],

$$P_f = \Pr(S^i > \lambda | H_0) = \frac{\Gamma(J, \frac{\lambda}{2\delta_n^2})}{\Gamma(J)}, \tag{5.6}$$

$$P_d = \Pr(S^i > \lambda | H_1) = Q_J\left(\sqrt{2J\gamma^i}, \sqrt{\frac{\lambda}{\delta_n^2}}\right). \tag{5.7}$$

where $\Gamma(a)$ is the gamma function, $\Gamma(a,x)$ is the upper incomplete gamma function, and $Q_J(a,x)$ denotes the generalised Marum Q-function.

5.2 Multirate Spectrum Detection

In practice, it is difficult to realize the standard spectrum estimation for wideband spectrum detection, because it requires a high speed ADC for acquiring the sequence $x^i[n]$. In order to implement distributed wideband spectrum sensing efficiently, a MSD system will be presented. The system is described in Section 5.2.1, and the effect of sub-Nyquist sampling is considered in Section 5.2.2, followed by the study of the data fusion rule in Section 5.2.3.

5.2.1 System Model

The MSD has a similar structure to the standard wideband spectrum estimation, however, Nyquist sampling is replaced by multirate sub-Nyquist sampling in parallel channels. Before the system description, there are several assumptions as follows. Since wideband sig-

nals are often sparse in the frequency domain [27], it is assumed that the full spectrum vector, $\overrightarrow{X}_j^i \in \mathbb{C}^N$, is k-sparse ($k \ll N$). Namely, only the largest k out of N components in $\overrightarrow{X}_j^i$ are non-neglectable. In addition, as the spectra viewed by distributed cognitive radios are often correlated when they are allocated in the same cluster, it is assumed that the spectra at distributed cognitive radios have a common spectral support (a set of active carrier locations in the frequency domain that is occupied by PUs) as illustrated in Figure. 4.6. Without loss of generality, let Ω represent the spectral support given by,

$$\Omega = \{n_1, n_2, \cdots, n_k\} \subset \{0, 1, \cdots, N-1\}. \tag{5.8}$$

In the MSD system, as shown in Figure 5.2, each cognitive radio is equipped with one wideband filter, one low-rate ADC, and an FFT for estimating the PSD. There are v distributed cognitive radio nodes that collaborate for wideband spectrum sensing. The wideband filters prior to the ADCs remove frequencies outside the spectrum of interest, and are altered to have the same bandwidth, W. Then the wideband signal viewed by the i-th cognitive radio, i.e. $x^i(t)$, is sampled by a single ADC with a different sub-Nyquist rate of M^i/T ($M^i < N$). Over an observation time of $T_{\text{all}} = JT$, the cognitive radio node i will obtain N_i-point samples, where $N_i = JM^i$. The aliased PSD in the i-th cognitive radio, i.e. $S_a^i[m]$, is calculated by,

$$S_a^i[m] = \sum_{j=1}^{J} \left|Y_j^i[m]\right|^2, \quad m \in [0, M^i), \tag{5.9}$$

where $Y_j^i[m]$ denotes the sub-Nyquist DFT spectrum of the j-th segment in the cognitive radio node i.

After that, the aliased PSD in each cognitive radio is transmitted to the FC, which forms the MSD PSD, i.e. $\widetilde{\mathrm{S}}[n]$, by,

$$\widetilde{\mathrm{S}}[n] = \sum_{i=1}^{v} \frac{N}{M^i} S_a^i[m+lM^i], \quad n \in [0, N), m \in [0, M^i), \tag{5.10}$$

where l is an integer and $l \in [0, N/M^i - 1]$. The MSD PSD, $\widetilde{\mathrm{S}}[n]$, will serve as a final test statistic, and will then be compared with a predetermined threshold, λ, to decide whether the PU is present or not.

5.2.2 Effect of Sub-Nyquist Sampling in A Single Cognitive Radio

Sub-Nyquist sampling will cause different spectral components to become indistinguishable, and change the distribution of the PSD. In the following, the change of spectral support will be considered, and the distribution of the MSD PSD will be modelled. From the discussion in Section 4.2, the relationship between the sub-Nyquist rate DFT spectrum and the full spectrum can be represented by,

$$Y_j^i[m] = \frac{M^i}{N} \sum_{l=-\infty}^{\infty} X_j^i[m + lM^i], \tag{5.11}$$

where $Y_j^i[m]$ and $X_j^i[n]$ denote the sub-Nyquist rate DFT spectrum and the full spectrum of the j-th segment in the cognitive radio node i, respectively. In matrix form this becomes,

$$\overrightarrow{Y}_j^i = \mathbf{\Phi^i} \overrightarrow{X}_j^i, \tag{5.12}$$

where $\overrightarrow{Y}_j^i$ and $\overrightarrow{X}_j^i$ denote the sub-Nyquist DFT spectrum vector and the full spectrum vector for the j-th segment in the cognitive radio node i, respectively, and $\mathbf{\Phi^i}$ is determined by (4.14).

Let Ω_i' represent the set of frequency bins in $\overrightarrow{Y}_j^i$ that corresponds to PUs,

$$\Omega_i' = \{m_1, m_2, \cdots, m_k\} \subset \{0, 1, \cdots, M^i - 1\}. \tag{5.13}$$

According to (5.11), Ω_i' and Ω can be related by,

$$m_j = |n_j|_{\bmod (M^i)}, \quad n_j \in \Omega, m_j \in \Omega_i', j \in [1, k], \tag{5.14}$$

where the subscript of m is named following the subscript of n. Evidently, when $n_j > M^i$, the exact location of n_j is lost in $\overrightarrow{Y}_j^i$.

Another risk caused by sub-Nyquist sampling is the signal overlap in $Y_j^i[m]$. It has been shown in Section 4.2.5 that when $M^i \sim \mathcal{O}(\sqrt{N})$ and $k \ll N$, the probability of signal overlap is very small. When only a single signal appears in the frequency bin m, the following equation holds from (5.11),

$$Y_j^i[m] = \frac{M^i}{N} X_j^i[m + lM^i], \tag{5.15}$$

where l is a particular integer that belongs to $[0, N/M^i - 1]$ but is unknown.

In such a scenario, with the aid of (5.3) and (5.15), the scaled sub-Nyquist rate DFT spectrum

can be approximately modelled by,

$$\sqrt{\frac{N}{M^i}}Y_j^i[m] \sim \begin{cases} \mathcal{CN}(0,\delta_s^2), & m \notin \Omega_i' \\ \mathcal{CN}(\sqrt{\frac{M^i}{N}}s^i,\delta_s^2), & m \in \Omega_i' \end{cases} \tag{5.16}$$

where δ_s^2 denotes the noise variance in the scaled sub-Nyquist rate DFT spectrum, and δ_s^2 can be approximately determined by using (5.3) and (5.11) as,

$$\delta_s^2 = \left\lceil \frac{N}{M^i} \right\rceil \left(\sqrt{\frac{M^i}{N}}\right)^2 \delta_n^2 \approx \delta_n^2, \tag{5.17}$$

where $\lceil \frac{N}{M^i} \rceil$ is the number of summations in (5.11).

Finally, according to (5.9) and (5.16), the scaled PSD in the cognitive radio node i can be modelled by,

$$\frac{N}{M^i}S_a^i[m] \sim \begin{cases} \chi_{2J}^2, & m \notin \Omega_i', \\ \chi_{2J}^2\left(2J\frac{M^i}{N}\gamma^i[m]\right), & m \in \Omega_i'. \end{cases} \tag{5.18}$$

5.2.3 Data Fusion Strategy among Multiple Cognitive Radios

The aliased PSD in distributed cognitive radio nodes will then be transmitted to and gathered at the FC, where the MSD PSD, $\widetilde{\mathrm{S}}[n]$, is calculated by using (5.10). Due to the periodic extension of $\overrightarrow{\mathrm{S}}_a^i$ ($\overrightarrow{\mathrm{S}}_a^i \in \mathrm{R}^{M^i}$) in (5.10), there will be more spectral components in the extended PSD than that in the original PSD, $\overrightarrow{\mathrm{S}}^i$ ($\overrightarrow{\mathrm{S}}^i \in \mathrm{R}^N$). For analysis convenience, let Ω_{M}^i represent the set of mirrored frequencies in Ω_{M}^i, and Ω_{U}^i represent the set of unaffected and unoccupied frequencies as,

$$\Omega_{\mathrm{M}}^i \triangleq \{n | n = m + lM^i, \ \ m \in \Omega_i', n \notin \Omega\}, \tag{5.19}$$

$$\Omega_{\mathrm{U}}^i \triangleq \{n | n = m + lM^i, \ \ m \notin \Omega_i', n \notin \Omega\}. \tag{5.20}$$

Using (5.10) and (5.18), the MSD PSD, $\widetilde{\mathrm{S}}[n]$, can be described by,

$$\widetilde{\mathrm{S}}[n] \sim \begin{cases} \chi_{2Jv}^2, & n \in \Omega_{\mathrm{U}} \\ \chi_{2Jv}^2\left(\frac{2J}{N}\sum\limits_{\substack{j=1 \\ i_j \in [1,v]}}^{j=p} M^{i_j}\gamma^{i_j}[n]\right), & n \in \Omega_{\mathrm{M}} \\ \chi_{2Jv}^2\left(\frac{2J}{N}\sum\limits_{i=1}^{i=v} M^i\gamma^i[n]\right), & n \in \Omega, \end{cases} \tag{5.21}$$

where $\Omega_{\mathrm{U}} = \Omega_{\mathrm{U}}^1 \cap \Omega_{\mathrm{U}}^2 \cdots \cap \Omega_{\mathrm{U}}^v$, $\Omega_{\mathrm{M}} = \Omega_{\mathrm{M}}^1 \cup \Omega_{\mathrm{M}}^2 \cdots \cup \Omega_{\mathrm{M}}^v$, and p denotes the number of cognitive radios, which have mirrored frequencies in the location of $n \in \Omega_{\mathrm{M}}$. Note that p could approach v, in which case, it is difficult to distinguish between $\widetilde{\mathrm{S}}[n](n \in \Omega_{\mathrm{M}})$ and $\widetilde{\mathrm{S}}[n](n \in \Omega)$. The aim of designing a data fusion strategy is to keep p to a minimum to achieve a best detection performance. First consider the case of $k = 1$, namely, only one location $n_1 \in \Omega$ in the full spectrum is occupied by PUs, then Proposition 5.1 will hold.

Proposition 5.1: If the length of segments in multiple cognitive radios, $M^1, M^2, ..., M^v$, are different prime numbers, and satisfy,

$$M^i M^j > N, \quad \forall\, i \neq j \in [1, v], \tag{5.22}$$

then two or more cognitive radios cannot have mirrored frequencies in the same location of $g \in \Omega_{\mathrm{M}}$.

Proof: Without loss of generality, the length of segment at the cognitive radio node i and j are assumed to be M^i and M^j, respectively. According to (5.14) and (5.19), the mirrored locations from the $n_1 \in \Omega$ are given by,

$$\begin{aligned} g_i &= |n_1|_{\mathrm{mod}\ (M^i)} + lM^i = n_1 - hM^i + lM^i, \quad h \neq l, \\ g_j &= |n_1|_{\mathrm{mod}\ (M^j)} + \check{l}M^j = n_1 - \check{h}M^j + \check{l}M^j, \quad \check{h} \neq \check{l}, \end{aligned} \tag{5.23}$$

where integers h and $\check{h}$ are from the operation of modulo, and $l - h \in [-\lceil \frac{N}{M^i} \rceil + 1, \lceil \frac{N}{M^i} \rceil - 1]$, $\check{l} - \check{h} \in [-\lceil \frac{N}{M^j} \rceil + 1, \lceil \frac{N}{M^j} \rceil - 1]$.

To avoid $g_i = g_j$, which is equivalent to avoiding $(l - h)M^i = (\check{l} - \check{h})M^j$, simply assume that M^i and M^j are different primes, and $\max(|l - h|) < M^j$, i.e. $\lceil \frac{N}{M^i} \rceil - 1 < M^j$. The condition $M^i M^j > N$ will satisfy this. Furthermore, if this holds for two cognitive radios, the case for more than two cognitive radios also holds. □

When $k = 1$ and the conditions in Proposition 5.1 are satisfied, based on the result of Proposition 5.1, the parameter p in (5.21) is found to be bounded by 1. On the other hand, when $k \geq 2$, the parameter p in (5.21) will be bounded by k. This is because only one cognitive radio node can map the frequency $n_j \in \Omega$ to the mirrored frequency $g \in \Omega_{\mathrm{M}}$, and the number of components in Ω is k.

Given $M^i M^j = b > N$ (b is constant), $M^i + M^j$ can be minimised when they are neigh-

bouring primes. Hence, to use the fewest measurements in MSD, the length of segments in multiple cognitive radios, $M^1, M^2, ..., M^v$, should be v neighbouring prime numbers. In such a scenario, the following approximations can be made,

$$\frac{2J\sum_{i=1}^{v} M^i\gamma^i}{N} \simeq \frac{2J\overline{M}}{N}\sum_{i=1}^{v}\gamma^i = \psi\gamma_{\mathrm{v}}, \tag{5.24}$$

$$\frac{2J\sum_{\substack{j=1\\ i_j\in[1,v]}}^{j=k} M^{i_j}\gamma^{i_j}}{N} \simeq \frac{2J\overline{M}}{N}\sum_{\substack{j=1\\ i_j\in[1,v]}}^{j=k}\gamma^{i_j} = \psi\gamma_{\mathrm{k}}, \tag{5.25}$$

where $\overline{M}$ denotes the average number of measurements over multiple cognitive radio nodes, $\gamma_{\mathrm{v}} \triangleq \sum_{i=1}^{v}\gamma^i$, $\gamma_{\mathrm{k}} \triangleq \sum_{\substack{j=1\\ i_j\in[1,v]}}^{j=k}\gamma^{i_j}$, and $\psi = \frac{2J\overline{M}}{N}$.

5.3 Performance Analysis

The MSD PSD, $\widetilde{\mathrm{S}}[n]$, will be compared with a predetermined threshold to determine whether the PU is absent or not. As described in (5.21), the noncentral parameters in the noncentral chi-square distributions contain the SNRs, thus both the probability of detection and the probability of false alarm in MSD will be influenced by the quality of the channels from PUs to cognitive radios.

5.3.1 Nonfading AWGN Channels

When the channels between PUs and cognitive radios can be modelled by nonfading AWGN models, the detection performance of MSD is given as follow.

Theorem 5.1: When all conditions in Proposition 5.1 are satisfied, the probabilities of false alarm and detection have bounds given by,

$$\frac{\Gamma(Jv, \frac{\lambda}{2\delta_s^2})}{\Gamma(Jv)} \leq P_f \leq Q_{Jv}\left(\sqrt{\psi\gamma_{\mathrm{k}}}, \sqrt{\frac{\lambda}{\delta_s^2}}\right), \tag{5.26}$$

$$P_d \geq Q_{Jv}\left(\sqrt{\psi\gamma_{\mathrm{v}}}, \sqrt{\frac{\lambda}{\delta_s^2}}\right). \tag{5.27}$$

Proof: Due to the mirrored spectral components, the following inequality can be obtained,

$$\Pr(\widetilde{\mathrm{S}}[n] > \lambda | n \in \Omega_{\mathrm{U}}) \leq \Pr(\widetilde{\mathrm{S}}[n] > \lambda | n \in \Omega_{\mathrm{M}}). \tag{5.28}$$

If let $U \triangleq \{0, 1, \cdots, N-1\}$ denote universal set for all frequency bins, $\overline{\Omega} \triangleq U - \Omega$ represent relative complement of Ω in U, then $\Omega_{\mathrm{U}} \bigcup \Omega_{\mathrm{M}} = \overline{\Omega}$. Thus, by (5.28), the following inequality holds,

$$\Pr(\widetilde{\mathrm{S}}[n] > \lambda | n \in \Omega_{\mathrm{U}}) \leq \Pr(\widetilde{\mathrm{S}}[n] > \lambda | n \in \overline{\Omega}) \leq \Pr(\widetilde{\mathrm{S}}[n] > \lambda | n \in \Omega_{\mathrm{M}}). \tag{5.29}$$

When conditions in Proposition 5.1 are satisfied, p will be bounded by k. With the aid of (5.21) and (5.29), (5.26) can be obtained. In the cases where one spectral component maps to another spectral component, the probability of detection will increase, and becomes larger than that predicted in (5.21). Thus, (5.27) holds. □

Theorem 5.1 shows that the number of distributed cognitive radios, v, should be larger than k. Additionally, more cognitive radios in collaboration will lead to a better detection performance, because with more cognitive radios, it is easier to distinguish between $\widetilde{\mathrm{S}}[n](n \in \Omega_{\mathrm{M}})$ and $\widetilde{\mathrm{S}}[n](n \in \Omega)$.

5.3.2 Rayleigh Fading Channels

When the signals from PUs experience i.i.d. Rayleigh fading, the PDF of the SNR at cognitive radio node i follows an exponential distribution,

$$f(\gamma^i) = \frac{1}{\overline{\gamma}} e^{-\frac{\gamma^i}{\overline{\gamma}}}, \quad \gamma^i > 0, \tag{5.30}$$

where $\overline{\gamma}$ denotes the local-mean SNR and γ^i is the SNR at cognitive radio node i.

Since cognitive radio nodes are distributed, the fading channels are assumed to be i.i.d.. The PDFs of $\gamma_{\mathrm{k}} = \sum_{\substack{j=1 \\ i_j \in [1,v]}}^{j=k} \gamma^{i_j}$ in (5.25) and $\gamma_{\mathrm{v}} = \sum_{i=1}^{v} \gamma^i$ in (5.24) will therefore follow Gamma distributions given by,

$$f(\gamma_{\mathrm{k}}) = \frac{(\gamma_{\mathrm{k}})^{k-1}}{(\overline{\gamma})^k \Gamma(k)} e^{-\frac{\gamma_{\mathrm{k}}}{\overline{\gamma}}}, \quad \gamma_{\mathrm{k}} \geq 0, \tag{5.31}$$

$$f(\gamma_{\mathrm{v}}) = \frac{(\gamma_{\mathrm{v}})^{v-1}}{(\overline{\gamma})^v \Gamma(v)} e^{-\frac{\gamma_{\mathrm{v}}}{\overline{\gamma}}}, \quad \gamma_{\mathrm{v}} \geq 0. \tag{5.32}$$

In the MSD system, the average probabilities of false alarm and detection can be calculated by averaging P_f in (5.26) and P_d in (5.27) over all possible SNRs.

Theorem 5.2: For the proposed MSD system over i.i.d. Rayleigh fading channels, the average probabilities of false alarm ($\overline{P_{f,\mathrm{R}}}$) and detection ($\overline{P_{d,\mathrm{R}}}$) will be bounded by,

$$\frac{\Gamma(Jv, \frac{\lambda}{2\delta_s^2})}{\Gamma(Jv)} \leq \overline{P_{f,\mathrm{R}}} \leq \Theta(k, Jv, \psi, \overline{\gamma}, \lambda, \delta_s^2), \tag{5.33}$$

$$\overline{P_{d,\mathrm{R}}} \geq \Theta(v, Jv, \psi, \overline{\gamma}, \lambda, \delta_s^2), \tag{5.34}$$

where $\Theta(x, Jv, \psi, \overline{\gamma}, \lambda, \delta_s^2)$ is defined by,

$$\Theta = \left(1 + \frac{\psi\overline{\gamma}}{2}\right)^{-x} \sum_{n=0}^{\infty} \mathrm{C}_{n+x-1}^{n} \left(\frac{\psi\overline{\gamma}}{\psi\overline{\gamma}+2}\right)^{n} \frac{\Gamma\left(n+Jv, \frac{\lambda}{2\delta_s^2}\right)}{\Gamma\left(n+Jv\right)},$$

where C_a^b denotes binomial coefficient given by $\mathrm{C}_a^b = \frac{b!}{a!(b-a)!}$.

Proof: In Rayleigh fading channels, the lower bound of the average probability of false alarm will remain the same as that in (5.26) as it is independent of the SNR, and the upper bound of the average probability of false alarm, $\overline{P_{f,\mathrm{R}}}^{\mathrm{up}}$, can be evaluated by,

$$\overline{P_{f,\mathrm{R}}}^{\mathrm{up}} = \int_0^{\infty} Q_{Jv}\left(\sqrt{\psi\gamma_{\mathrm{k}}}, \sqrt{\frac{\lambda}{\delta_s^2}}\right) \frac{(\gamma_{\mathrm{k}})^{k-1}}{(\overline{\gamma})^k \Gamma(k)} e^{-\frac{\gamma_{\mathrm{k}}}{\overline{\gamma}}} d\gamma_{\mathrm{k}}. \tag{5.35}$$

Using (4.74) in [51] and (8.352-2) in [71], the upper bound of P_f in (5.26) can be rewritten as,

$$P_f^{\mathrm{up}} = Q_{Jv}\left(\sqrt{\psi\gamma_{\mathrm{k}}}, \sqrt{\frac{\lambda}{\delta_s^2}}\right) = \sum_{n=0}^{\infty} \frac{\left(\frac{\psi\gamma_{\mathrm{k}}}{2}\right)^n e^{-\frac{\psi\gamma_{\mathrm{k}}}{2}}}{n!} \frac{\Gamma(n+Jv, \frac{\lambda}{2\delta_s^2})}{\Gamma(n+Jv)}. \tag{5.36}$$

Substituting (5.36) into (5.35), $\overline{P_{f,\mathrm{R}}}^{\mathrm{up}}$ can be written as,

$$\overline{P_{f,\mathrm{R}}}^{\mathrm{up}} = \frac{1}{(\overline{\gamma})^k} \sum_{n=0}^{\infty} \frac{\left(\frac{\psi}{2}\right)^n \Gamma(n+Jv, \frac{\lambda}{2\delta_s^2})}{n!(k-1)!\Gamma(n+Jv)} \int_0^{\infty} (\gamma_{\mathrm{k}})^{n+k-1} e^{-\frac{\psi\gamma_{\mathrm{k}}}{2} - \frac{\gamma_{\mathrm{k}}}{\overline{\gamma}}} d\gamma_{\mathrm{k}}. \tag{5.37}$$

Using (3.351-3) in [71] for calculating the integral, with manipulation, $\overline{P_{f,\mathrm{R}}}^{\mathrm{up}}$ is obtained as,

$$\overline{P_{f,\mathrm{R}}}^{\mathrm{up}} = \left(1 + \frac{\psi\overline{\gamma}}{2}\right)^{-k} \sum_{n=0}^{\infty} \mathrm{C}_{n+k-1}^{n} \left(\frac{\psi\overline{\gamma}}{\psi\overline{\gamma}+2}\right)^{n} \frac{\Gamma\left(n+Jv, \frac{\lambda}{2\delta_s^2}\right)}{\Gamma\left(n+Jv\right)}. \tag{5.38}$$

The lower bound of the average probability of detection can be approximated similarly. □

5.3.3 Slow Fading Channels

Over slow fading channels, the PDF of the SNR at the i-th cognitive radio node, $f(\gamma^i)$, is given by,

$$f(\gamma^i) = \frac{\xi}{\sqrt{2\pi}\sigma^i\gamma^i} \exp\left(-\frac{\left(10\log_{10}(\gamma^i) - \ddot{\bar{\gamma}}^i\right)^2}{2(\sigma^i)^2}\right), \quad \gamma^i > 0, \tag{5.39}$$

where $\xi = 10/\ln(10)$, $\ddot{\bar{\gamma}}^i$ (dB) denotes the area-mean SNR, and σ^i (dB) denotes the shadowing standard deviation of $10\log_{10}\gamma^i$ at the cognitive radio node i. Note that the PDF in (5.39) can be closely approximated by a Wald distribution as [73],

$$f(\gamma^i) = \sqrt{\frac{\eta^i}{2\pi}}(\gamma^i)^{-3/2} \exp\left(-\frac{\eta^i(\gamma^i - \theta^i)^2}{2(\theta^i)^2\gamma^i}\right), \quad \gamma^i > 0, \tag{5.40}$$

where $\theta^i = \mathrm{E}(\gamma^i)$ denotes the expectation of γ^i, and η^i is the shape parameter for the cognitive radio node i. By the method of moments, the parameters η^i, θ^i and $\ddot{\bar{\gamma}}^i$, σ^i are related as below,

$$\begin{aligned} \theta^i &= \exp\left(\frac{\ddot{\bar{\gamma}}^i}{\xi} + \frac{(\sigma^i)^2}{2\xi^2}\right), \\ \eta^i &= \frac{\theta^i}{\exp(\frac{(\sigma^i)^2}{\xi^2}) - 1}. \end{aligned} \tag{5.41}$$

Since the cognitive radio nodes are distributed and the observation time is small, all slow fading channels are assumed to be quasi stationary and i.i.d.. The condition $\frac{\eta^i}{(\theta^i)^2} = \frac{\mathrm{E}(\gamma^i)}{\mathrm{Var}(\gamma^i)} = b$ (constant) can be satisfied. Thus, γ_k and γ_v will also follow the Wald distribution [76]. The PDFs of γ_k and γ_v are given by,

$$f(\gamma_\mathrm{k}) = \sqrt{\frac{k\eta}{2\pi}}(\gamma_\mathrm{k})^{-3/2} \exp\left(-\frac{\eta(\gamma_\mathrm{k} - k\theta)^2}{2k\theta^2\gamma_\mathrm{k}}\right), \quad \gamma_\mathrm{k} > 0, \tag{5.42}$$

$$f(\gamma_\mathrm{v}) = \sqrt{\frac{v\eta}{2\pi}}(\gamma_\mathrm{v})^{-3/2} \exp\left(-\frac{\eta(\gamma_\mathrm{v} - v\theta)^2}{2v\theta^2\gamma_\mathrm{v}}\right), \quad \gamma_\mathrm{v} > 0, \tag{5.43}$$

where η and θ denote the average of η^i and θ^i, respectively.

The average probabilities of false alarm and detection can be calculated by averaging P_f in

(5.26) and P_d in (5.27) over the PDFs of the SNRs.

Theorem 5.3: For the MSD system over i.i.d. shadowing fading channels, the average probabilities of false alarm ($\overline{P_{f,\mathrm{S}}}$) and detection ($\overline{P_{d,\mathrm{S}}}$) will be bounded by,

$$\frac{\Gamma(Jv, \frac{\lambda}{2\delta_s^2})}{\Gamma(Jv)} \leq \overline{P_{f,\mathrm{S}}} \leq \Lambda(k, Jv, \psi, \lambda, \theta, \eta, \delta_s^2), \tag{5.44}$$

$$\overline{P_{d,\mathrm{S}}} \geq \Lambda(v, Jv, \psi, \lambda, \theta, \eta, \delta_s^2), \tag{5.45}$$

where $\Lambda(x, Jv, \psi, \lambda, \theta, \eta, \delta_s^2)$ is defined by,

$$\Lambda = \sqrt{\frac{2x\eta}{\pi}} e^{\frac{\eta}{\theta}} \sum_{n=0}^{\infty} \frac{\left(\frac{\psi}{2}\right)^n \Gamma\left(n+Jv, \frac{\lambda}{2\delta_s^2}\right)}{n!\Gamma\left(n+Jv\right)} \left(\sqrt{\frac{x^2\eta\theta^2}{x\psi\theta^2+\eta}}\right)^{n-\frac{1}{2}} \mathrm{K}_{n-\frac{1}{2}}\left(\sqrt{\frac{\eta(x\psi\theta^2+\eta)}{\theta^2}}\right), \tag{5.46}$$

where $\mathrm{K}_{n-\frac{1}{2}}(a)$ denotes the modified Bessel function of the second kind with order $n-\frac{1}{2}$.

Proof: In shadowing fading channels, the lower bound of $\overline{P_{f,\mathrm{S}}}$ in (5.44) will not change. By (5.42), the upper bound of the average probability of false alarm, i.e. $\overline{P_{f,\mathrm{S}}}^{\mathrm{up}}$, can be evaluated by,

$$\overline{P_{f,\mathrm{S}}}^{\mathrm{up}} = \int_0^{\infty} Q_{Jv}\left(\sqrt{\psi\gamma_{\mathrm{k}}}, \sqrt{\frac{\lambda}{\delta_s^2}}\right) \sqrt{\frac{k\eta}{2\pi}} (\gamma_{\mathrm{k}})^{-3/2} \exp\left(-\frac{\eta(\gamma_{\mathrm{k}}-k\theta)^2}{2k\theta^2\gamma_{\mathrm{k}}}\right) d\gamma_{\mathrm{k}}. \tag{5.47}$$

Substituting (5.36) into (5.47), $\overline{P_{f,\mathrm{S}}}^{\mathrm{up}}$ is calculated by,

$$\overline{P_{f,\mathrm{S}}}^{\mathrm{up}} = \sqrt{\frac{k\eta}{2\pi}} \sum_{n=0}^{\infty} \frac{\left(\frac{\psi}{2}\right)^n \Gamma\left(n+Jv, \frac{\lambda}{2\delta_s^2}\right)}{n!\Gamma\left(n+Jv\right)} \int_0^{\infty} (\gamma_{\mathrm{k}})^{n-\frac{3}{2}} \exp\left(-\frac{k\psi\theta^2+\eta}{2k\theta^2}\gamma_{\mathrm{k}} - \frac{k\eta}{2\gamma_{\mathrm{k}}} + \frac{\eta}{\theta}\right) d\gamma_{\mathrm{k}}. \tag{5.48}$$

Using (3.471-9) in [71] for calculating the integral in (5.48), with manipulation, $\overline{P_{f,\mathrm{S}}}^{\mathrm{up}}$ can be obtained as,

$$\overline{P_{f,\mathrm{S}}}^{\mathrm{up}} = \sqrt{\frac{2k\eta}{\pi}} e^{\frac{\eta}{\theta}} \sum_{n=0}^{\infty} \frac{\left(\frac{\psi}{2}\right)^n \Gamma\left(n+Jv, \frac{\lambda}{2\delta_s^2}\right)}{n!\Gamma(n+Jv)} \left(\sqrt{\frac{k^2\eta\theta^2}{k\psi\theta^2+\eta}}\right)^{n-\frac{1}{2}} \mathrm{K}_{n-\frac{1}{2}}\left(\sqrt{\frac{\eta(k\psi\theta^2+\eta)}{\theta^2}}\right). \tag{5.49}$$

Likewise, the lower bound of the average probability of detection can be approximated. Hence, Theorem 5.3 follows. □

5.4 Simulations

In the simulations, distributed cognitive radio nodes are assumed to have the setup as illustrated in Figure 5.2. The wideband signal $x^i(t)$ viewed by the i-th cognitive radio is assumed to be,

$$x^i(t) = \sum_{l=1}^{N_b} \sqrt{E_l^i} B_l \cdot \text{sinc}\left(B_l(t-\Delta)\right) \cdot \cos\left(2\pi f_l(t-\Delta)\right) + z(t), \tag{5.50}$$

where sinc(x) is defined by $\text{sinc}(x) = \frac{\sin(\pi x)}{\pi x}$, Δ denotes the time offset, set to be $T/2$, and E_l^i is the received power at cognitive radio node i and varies subject to the fading channel. Without any loss of generality, the average received power of all occupied subbands is normalized to be 1, and the noise is assumed to be AWGN, i.e. $z(t) \sim \mathcal{N}(0, \delta^2 = 1)$. The wideband signal $x^i(t)$ (10 GHz bandwidth) consists of $N_b = 6$ non-overlapping subbands, whose bandwidth varies in the range of $B_l = 1 \sim 10$ MHz, with the carrier frequency of $f_l = B_l \sim (W - B_l)$. The wideband signal is observed in a total time of 4 μs at different sub-Nyquist rates, and then the measurement sequence is divided into $J = 20$ or $J = 50$ segments. As the sidelobe of Hanning window decays rapidly, it is used for each data segments. Since the signal has the bandwidth of $W = 10$ GHz, if it were sampled at the Nyquist rate, i.e. $f_s = 20$ GHz, the length of each segment would be $N = 80,000$.

The aliased PSD, $S_a^i[m]$, is individually estimated at a different sub-Nyquist rate in each cognitive radio. The aliased PSD from all distributed cognitive radios is gathered at the FC, with total number of measurements $0.4967N$ (N denotes the number of samples when sampled at the Nyquist rate). The MSD PSD is finally calculated by using (5.10), and is compared with a predetermined threshold. The MSD PSD over i.i.d. Rayleigh and slow fading channels are illustrated in Figure 5.3 and Figure 5.4, respectively. It can be seen that the proposed system cannot achieve the benchmark system performance (obtained if Nyquist sampled). The observed SNR in the MSD PSD is lower than that in the Nyquist case. However, it is worthwhile to emphasize that these active frequencies can still be identified over various fading channels, when the compression ratio is approximately 50%. Furthermore, the average sampling rate in MSD is only 132.46 MHz, rather than 20 GHz in the Nyquist sampling case. Because the spectrum sensing data transmitted from cognitive radios to the FC is the PSD, MSD does not require perfect time synchronisation between distributed cognitive radios. In comparison to CS based detection methods, the transmission and storage of measurement matrix is unnecessary in the MSD system.

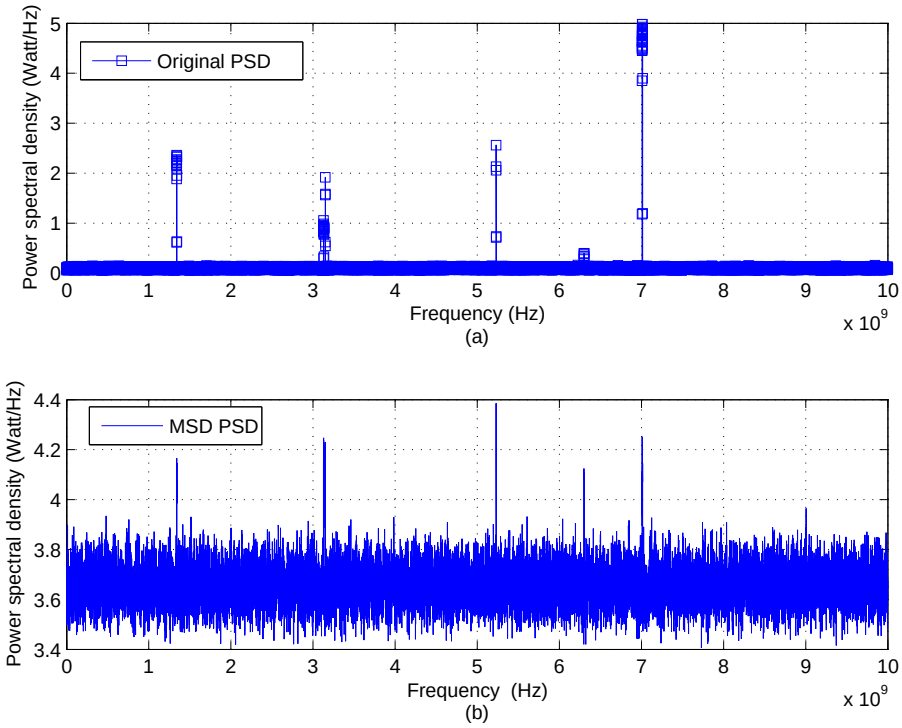

Figure 5.3: *Comparison of the original PSD in (a) with the MSD PSD in (b) over i.i.d. Rayleigh fading channels. In the experiment, the compression ratio is* $M/N = 49.67\%$*, with* 75 *ADCs at the average sampling rate of* 132.46 MHz*, and the SNR averaged over fading channels is* 10 dB.

Figure 5.5 compares the simulation results with the theoretical results predicted in (5.26)-(5.27), (5.33)-(5.34), and (5.44)-(5.45). These curves are obtained by using Monte Carlo method with 100,000 trials. It is evident that the empirical probabilities of false alarm in all figures are close to the lower bounds, but far away from the upper bounds. This is due to the fact that the assumption (all k components in the full spectrum have been mirrored to the same location when the wideband signal is sub-Nyquist sampled) for deriving the upper bound has a very low probability of occurring. Figure 5.5 also illustrates that the lower bound of the probability of detection can successfully predict the trend of empirical results. Comparing the fading cases with the non-fading case, it is found that in fading channels the probability of detection declines more slowly than that in AWGN channels. This is more obvious for the case of cognitive radios over slow fading channels as shown in Figure 5.5-(c). In contrast, the effect of fading has little or no influence on the probability of false alarm. This is because the mirrored frequencies will seldom apprear in the same location, and therefore validates the assumption in Section 5.2.2

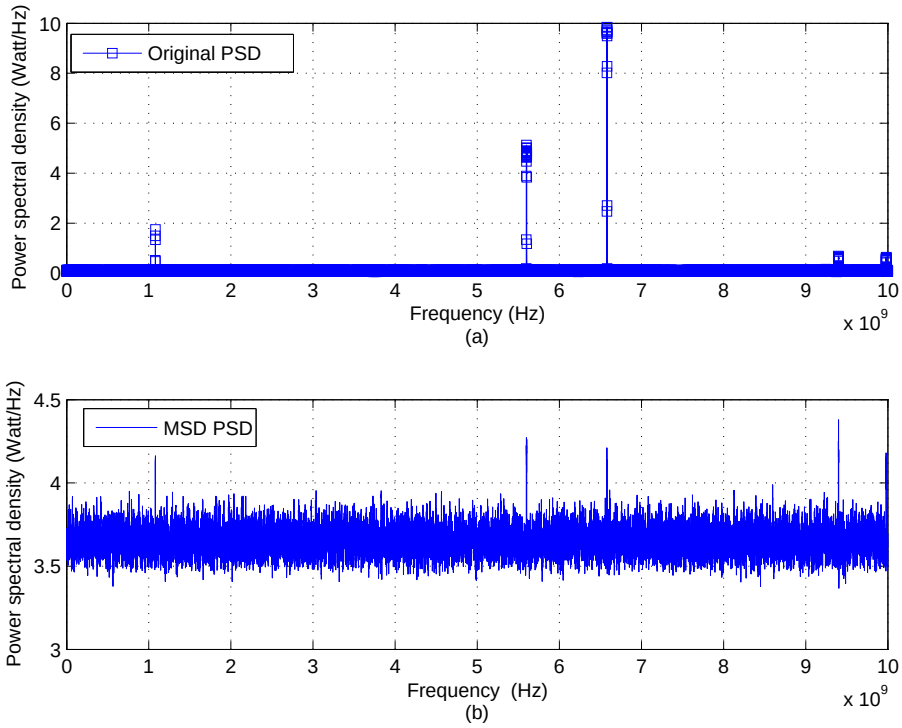

Figure 5.4: *Comparison of the original PSD in (a) with the MSD PSD in (b) over i.i.d. slow fading channels with the standard deviation* $\sigma = 5$ dB. *In the experiment, the compression ratio is* $M/N = 49.67\%$, *with* 75 *ADCs at the average sampling rate of* 132.46 MHz, *and the SNR averaged over fading channels is* 10 dB.

that the probability of signal overlap is very small.

In Figure 5.6, the ROC curves, i.e. P_f vs P_d, are used to to quantify the detection performance. Figure 5.6 demonstrates the detection performance of the proposed MSD system when cognitive radios are subject to non-fading AWGN, Rayleigh, and slow fading channels under low SNR scenarios. When SNR$= 0$ dB, the performance of MSD over fading channels perform roughly the same as that over non-fading AWGN channels. This is because the strength of signal is mostly annihilated by the noise. By contrast, the detection performance of MSD over AWGN channels outperform that over fading channels when SNR$= 5$ dB. In addition, it is found that the performance of MSD over slow fading channels is the poorest, in comparison to the case of MSD over AWGN, or Rayleigh fading channels. Nonetheless, even over slow fading channels, MSD has a probability of nearly 90% of detecting the presence of PUs when the probability of false alarm is 10%, by just using $0.492N$ measurements. On the contrary, wavelet detection and filter bank detection methods must use at least N measurements. Another

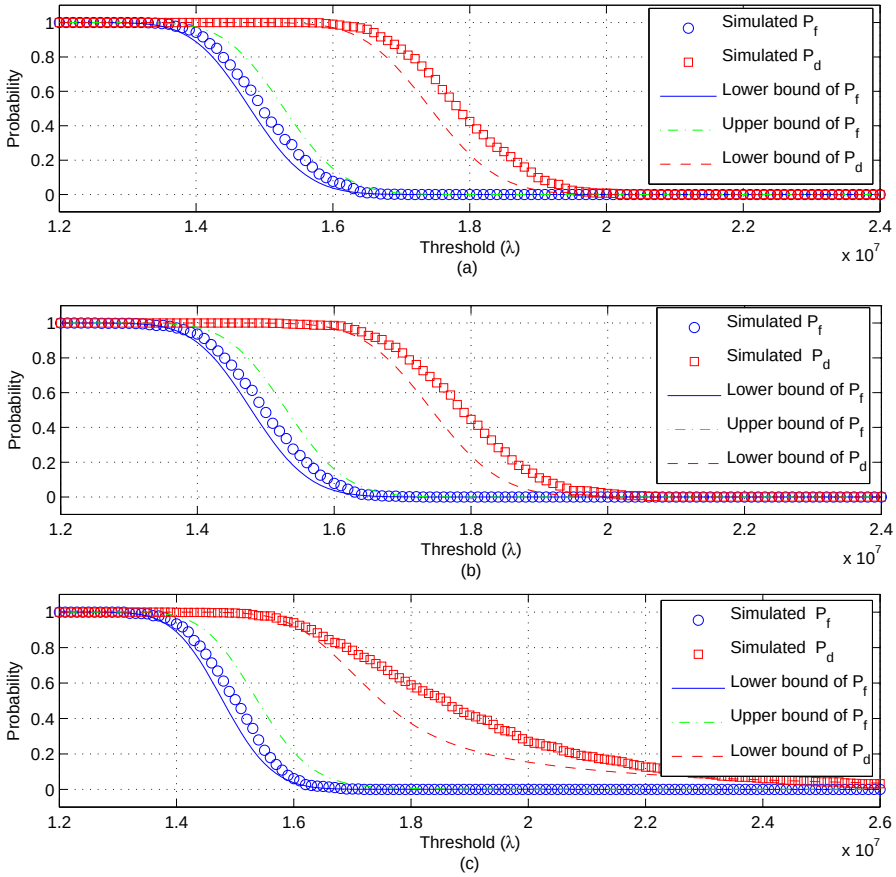

Figure 5.5: *Comparisons of empirical results and theoretical results for the probabilities of detection and false alarm over (a) AWGN, (b) Rayleigh, and (c) slow fading channels. The theoretical bounds are predicted in equations (5.26)-(5.27), (5.33)-(5.34), and (5.44)-(5.45).*

advantage of MSD is that the average sampling rate in MSD is only 448.68 MHz, instead of 20 GHz in wavelet detection method.

Figure 5.7 depicts the influence of the depth of shadowing when the MSD system is working over i.i.d. slow fading channels. It can be found that the larger shadowing standard deviation will lead to a worse detection performance for the MSD system. For example, when the average probability of false alarm is 10% and SNR$=5$ dB, the average probability of detection drops from 90% to 80% when σ rises from 5 dB to 7 dB. This is because a larger σ is equivalent to a longer tail in the log-normal distribution, thus making the detection difficult. The performance difference between $\sigma = 5$ dB and $\sigma = 7$ dB is small when the SNR decreases to 0 dB, because in that case the power level of the signals is similar to that of the noise.

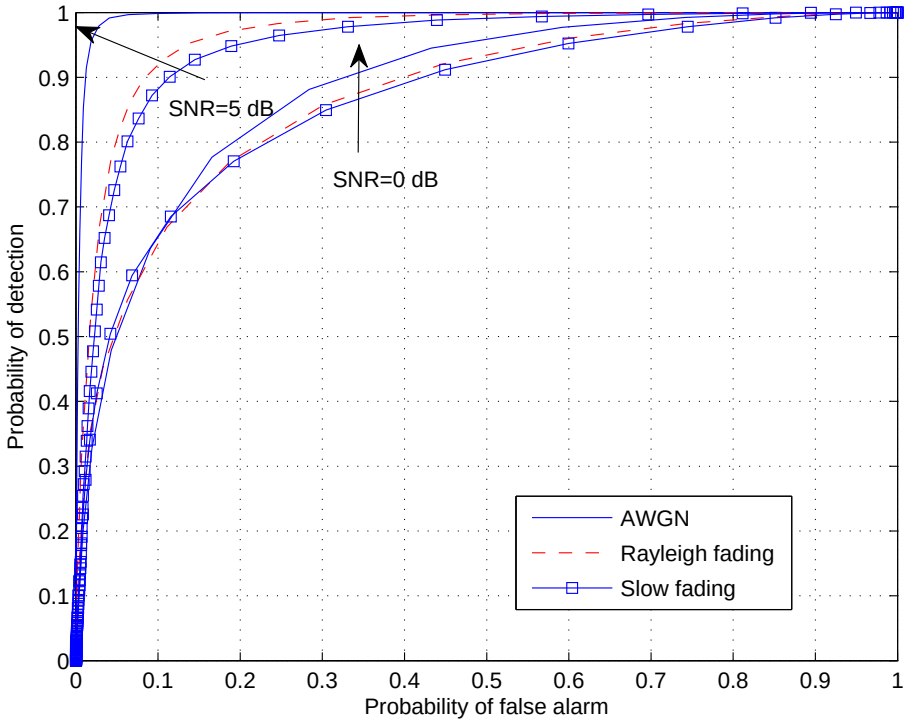

Figure 5.6: *ROC curves of MSD over different kinds of fading channels when the compression ratio $M/N = 49.2\%$ and the averaging times $J = 50$. The wideband signal is observed by 22 cognitive radios at different sampling rates (the average sampling rate is 448.68 MHz). The shadowing standard deviation is 5* dB.

Figure 5.8 shows the influence of the compression ratio (proportional to the number of distributed cognitive radios in collaboration) on the detection performance of MSD over slow fading channels, by using Complementary ROC curves, i.e. P_f vs $1 - P_d$. It is obvious that with increasing compression ratio, a better detection performance can be achieved. Considering the probability of false alarm $P_f = 10\%$, the average probability of missed detection reduces from 7% to 2% when the compression ratio increases from 49.2% to 150%. The reason is that adding more sampling channels (equivalent to more cognitive radios) makes it easier to distinguish the occupied frequencies from the mirrored frequencies as shown in (5.21). Thus, with increasing spatial diversity, the effect of fading can be mitigated in the MSD system.

The assumption validating the MSD system is that the spectrum is k-sparse. To investigate the influence of k, the ROC curves of MSD over slow fading channels are shown in Figure 5.9 with different values of k (proportional to the number of subbands). It shows that with the same SNR the detection performance degrades when the number of subbands increases. For

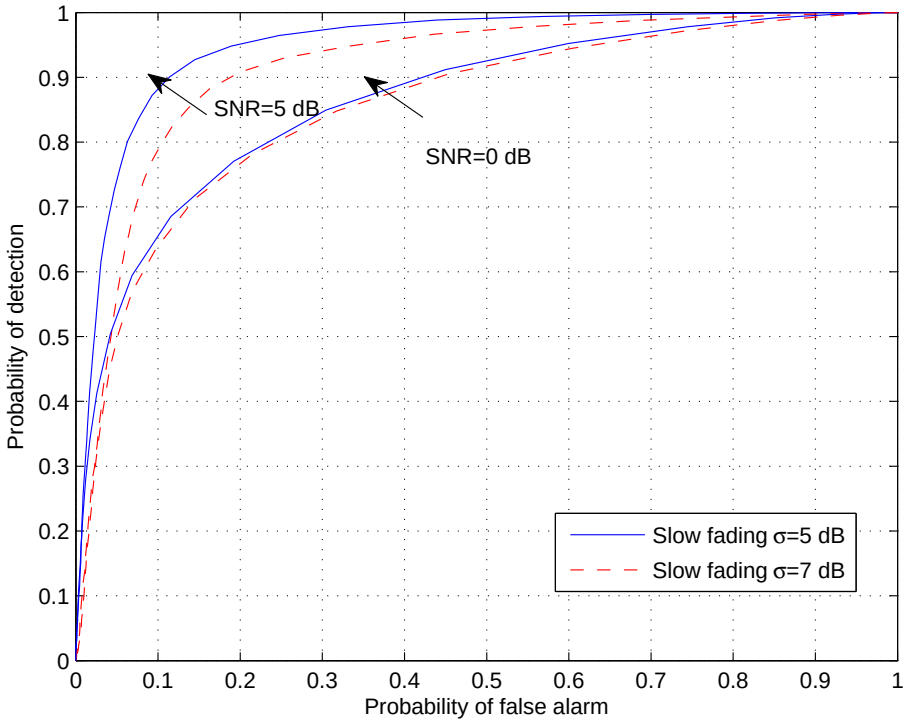

Figure 5.7: *ROC curves of MSD over slow fading channels with the compression ratio* $M/N = 49.2\%$ *and* $J = 50$*, when the shadowing standard deviation,* σ*, and the average SNR alter. The wideband signal is sampled at different sampling rates by* 22 *ADCs with the average sampling rate of* 448.68 MHz*.*

example, for a fixed probability of false alarm, i.e. $P_f = 10\%$, the probability of detection drops from 73% to 70% for SNR$= 0$ dB, and declines from 95% to 91% for SNR$= 5$ dB when the number of subbands increases from 6 to 12. The performance degradation of the MSD system stems from two reasons. One of them is that when k increases, the probability of signal overlap becomes larger when the wideband signal is sub-Nyquist sampled. The signal overlap may then lead to the missed detection of the PUs. The another reason is that, for a fixed number of sampling channels (or a fixed number of cognitive radio nodes in collaboration), increasing k makes it more difficult to distinguish the occupied frequencies from the mirrored frequencies as described in Section 5.3.1.

Figure 5.10 reveals the influence of sampling duration (number of segments), J, on the MSD system performance over slow fading channels. The ROC curves show that the performance of MSD becomes better when J increases or the number of subbands decreases. This is because averaging will reduce the noise variance in the PSD, thus makes the detection of the PUs easier.

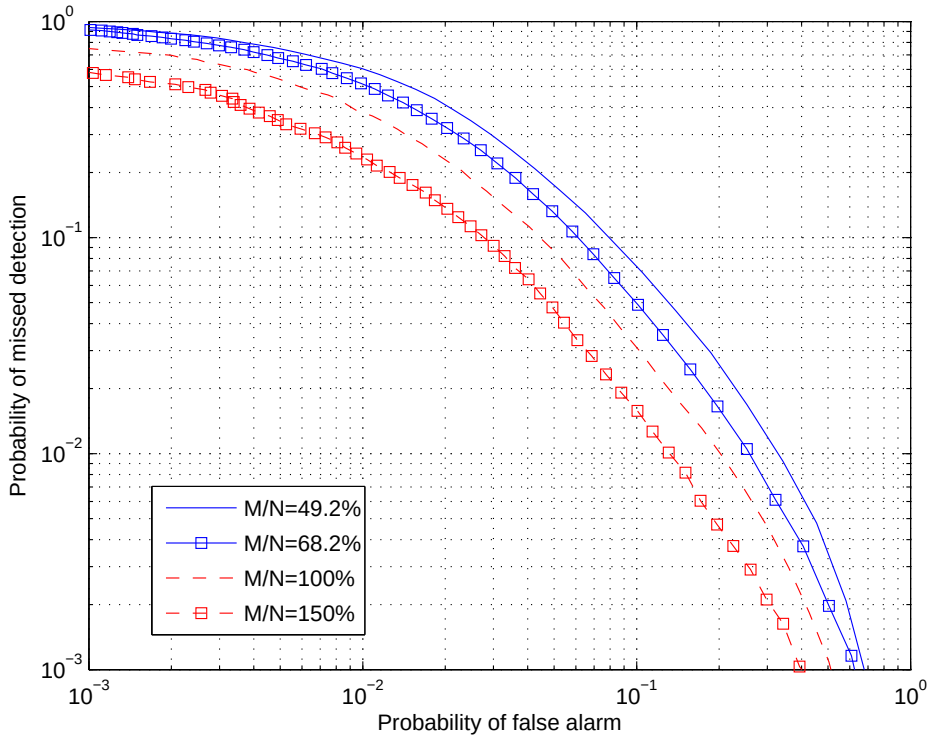

Figure 5.8: *Complementary ROC curves of the proposed MSD system over slow fading channels with SNR*$= 5$ dB, $\sigma = 5$ dB *and* $J = 50$, *when the compression ratio* (M/N) *varies.*

In practical environments, distributed spectrum sensing may face the problem of insufficient collaborative nodes. On the other hand, distributed cognitive radio nodes may be limited by the maximum sampling speed due to ADC technology or energy consumption considerations. In such scenarios, the tradeoff between the number of cognitive radio nodes and the average sampling rate across cognitive radio networks is critical. Figure 5.11 investigates the performance tradeoff with the fixed compression ratio when the MSD system is working over slow fading channels. It shows that the case of using less cognitive radios but with higher sampling rate outperforms that of using more cognitive radios with lower sampling rate. For example, when the probability of false alarm is 10%, the probability of detection is 92% for the case of 22 ADCs (with the average sampling rate 448.68 MHz), rather than 70% when using 75 cognitive radio nodes (with the average sampling rate 132.46 MHz). This seems in contradictory to (5.21) as the compression ratio doesn't change, thus the noncentral parameter will not change. However, it should be emphasized that a low sampling rate will lead to a larger probability of signal overlap, in which case the average probability of detection will decrease.

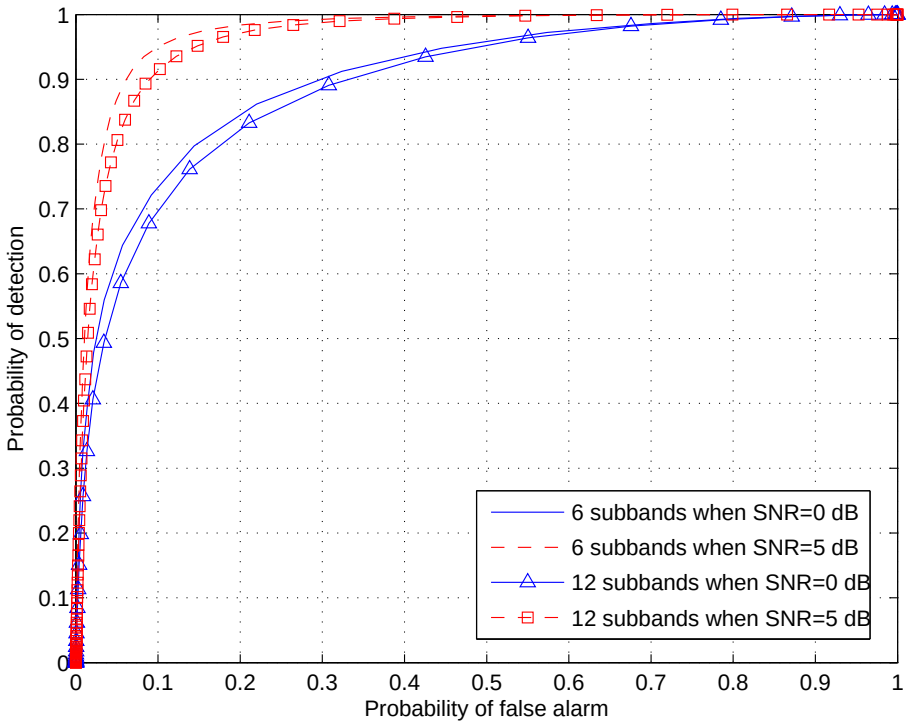

Figure 5.9: *ROC curves of the proposed MSD system over slow fading channels with* $\sigma = 5$ dB, $M/N = 68.2\%$, *and* $J = 50$, *when the SNR and the number of subbands change.*

5.5 Conclusions

In this chapter, a multirate spectrum detection system, called MSD, has been presented for implementing wideband spectrum detection in a distributed cognitive radio network. The detection performance of MSD over AWGN, Rayleigh, or slow fading channels has been derived. The MSD system has several attractive features of being used for wideband spectrum sensing in distributed cognitive radio networks. Firstly, the implementation complexity of the MSD system is low, thus each distributed cognitive radio only requires one low-rate ADC. To sense a wideband spectrum with 10 GHz bandwidth, the average sampling rate of ADCs is 448.68 MHz, opposed to the Nyquist rate of 20 GHz. Secondly, the MSD system is energy-efficient because only the sub-Nyquist PSD are processed and transmitted to the FC, which not only saves processing energy, but also saves transmission energy. Moreover, compared with MASS, or other CS based approaches, MSD can detect the full spectrum without reconstructing it. Therefore, the proposed MSD system can reduce the network overhead as it has a lower com-

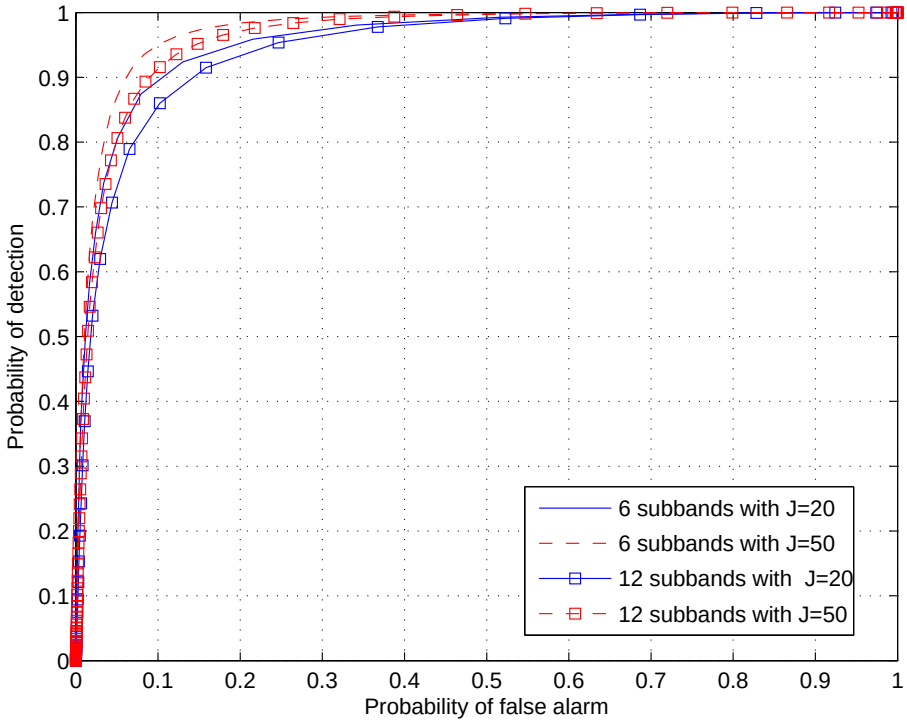

Figure 5.10: *ROC curves of the MSD system over slow fading channels with SNR*$= 5$ dB, $\sigma = 5$ dB, *and the compression ratio* $M/N = 68.2\%$, *when the averaging times,* J, *and the number of subbands alter.*

putational complexity in the FC.

Simulation results have shown that when the compression ratio is approximately 50%, MSD can achieve a good detection performance over slow fading channels. The detection performance of MSD becomes better when the number of distributed cognitive radio nodes increases or the number of PUs decreases. When there are limitations on the number of distributed cognitive radio nodes, or the maximum sampling rate in each cognitive radio, a tradeoff can be made to adapt the practical environment by changing the sampling rates.

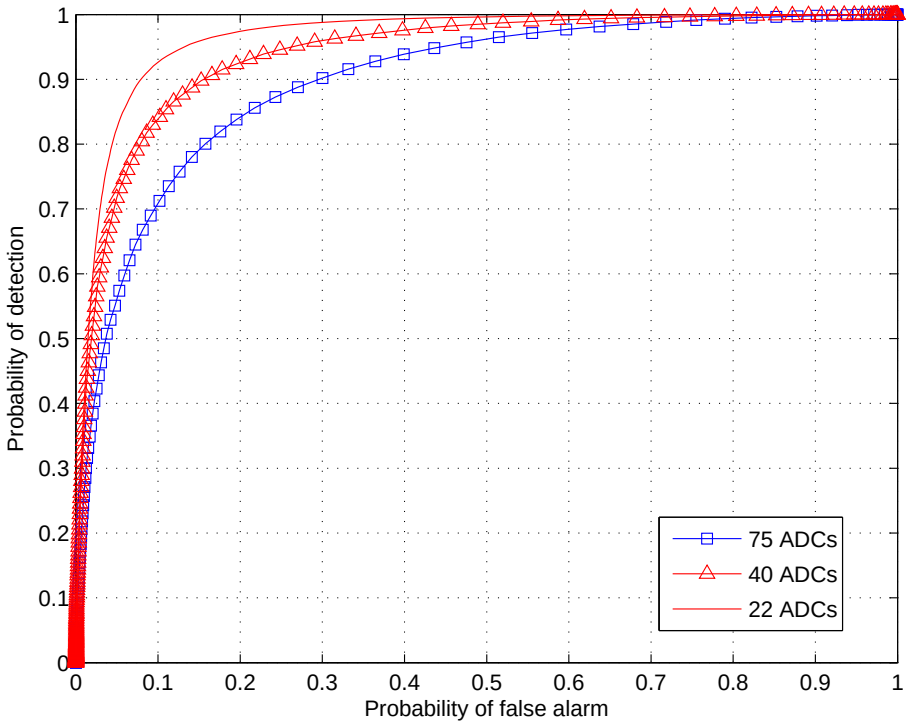

Figure 5.11: *For fixed parameters: compression ratio* $M/N = 49.67\%$*, SNR*$= 5$ dB*, and* $J = 50$*, the tradeoff between the number of ADCs and the average sampling rate, i.e.* 75 *ADCs with the average sampling rate* 132.46 MHz*,* 40 *ADCs with the average sampling rate* 247.45 MHz*, and* 22 *ADCs with the average sampling rate* 448.68 MHz*.*

Chapter 6
Conclusions and Future Work

This thesis has contributed to the performance evaluation of collaborative spectrum sensing algorithms as well as the development of wideband spectrum sensing techniques in centralised cognitive radio networks. In this concluding chapter, a summary of the key contributions from different chapters will be given in Section 6.1. Some limitations of work are discussed in Section 6.2. Several suggestions for future research areas are presented in Section 6.3.

6.1 Conclusions

Due to the low implementation complexity, energy detection is commonly used for spectrum sensing in a cognitive radio network. In fading environments, it is crucial for cognitive radios to dynamically balance the probability of missed detection against the probability of false alarm. As cognitive radios are always restricted by limited computational resources, fast converging expressions for these probabilities are of advantage. In Chapter 3, easily computed expressions for the average probabilities of detection and false alarm over a single Nakagami-m, or Rician fading channel have been derived. Despite infinite sums being involved, simulation results have shown that the proposed expressions converge more quickly than the methods in the literature. In the case of log-normal fading, a computationally tractable expression for the average probability of detection has been given, by approximating the log-normal distribution with the Wald distribution. It has been testified that the proposed expression closely fits the simulation results. On the other hand, as collaborative spectrum sensing approaches are commonly used for combating fading and improving detection performance, the performance of collaborative spectrum sensing algorithms, i.e. MRC, SC, SLC, and SLS, has been analysed. Among these algorithms, the MRC based collaborative spectrum sensing approach offers an upper bound of the detection performance, but at the expense of highest transmission bandwidth. When the bandwidth of the control channel is limited, the SLC based collaborative spectrum sensing scheme is more attractive as it not only saves 50% transmission bandwidth, but also does not require CSI.

In Chapter 4, a distributed wideband spectrum sensing system, called MASS, has been presented for implementing wideband spectrum sensing. It employs several low-rate ADCs to sample the wideband signal at different sub-Nyquist rates, instead of a single high speed ADC. One of the key advantages of the MASS is that the full wideband spectrum can be reconstructed from a few measurements. Sufficient conditions have been derived to uniquely recover the full spectrum by a multirate asynchronous sub-Nyquist sampling. Another advantage of MASS is that it can be applied either in the scenario of multiple ADCs in one sensor or in the case of single ADCs in multiple sensors. In the latter case, the implementation complexity of the MASS is extremely low as only one low-rate ADC is required in each cognitive radio node. The HMP algorithm has been applied to recover the common Nyquist spectrum by a boost-and-recover strategy. Compared with other wideband spectrum sensing techniques, MASS is energy-efficient as it requires less transmission bandwidth and fewer compression devices. MASS has been proved to be very robust against model mismatches and time synchronisation failures. The performance of MASS has been investigated and verified by simulation results. It has been demonstrated that MASS can recover the wideband Nyquist spectrum successfully when the compression ratio is approximately 50%. For the purpose of combating the effect of fading, only a few more cognitive radios are needed to collaborate in the MASS system.

In order to reduce the computational complexity in the FC, a multirate spectrum detection model has been proposed that implements wideband spectrum detection in distributed cognitive radio networks in Chapter 5. When the distributed cognitive radios suffer from AWGN, Rayleigh fading, or log-normal fading channels, the detection performance of MSD has been derived. The implementation complexity of the MSD system is as low as that of the MASS system in the sense that each distributed cognitive radio still requires one low-rate ADC and the sampling rate of that ADC is similar. The MSD system is also energy-efficient because only the sub-Nyquist power spectra are processed and transmitted to the FC, which not only saves processing energy, but also saves transmission energy. Compared with MASS, or other CS based approaches, the main advantage of MSD is that it can detect the PUs without ever reconstructing the full spectrum. Empirical results have shown that when the compression ratio is approximately 50%, MSD can achieve a good detection performance over i.i.d. Rayleigh or log-normal fading channels. The detection performance of MSD becomes better when the number of distributed cognitive radio nodes increases or the number of PUs decreases. More importantly, MSD has been shown to be very flexible that the number of sampling channels and the sampling rate can be balanced to adapt the practical environment.

6.2 Limitation of Work

In Chapter 3, the assumption of using energy detection is that the noise variance can be accurately estimated. However, in practice, this is difficult. The noise uncertainty will degrade the system performance even if collaborative spectrum sensing strategies are adopted. Besides, using energy detection, the signals from the PUs and interferences from other cognitive radios cannot be clearly differentiated. To further exploit the spectral scarcity, feature-based spectrum sensing techniques can be employed for fine spectrum sensing, such as cyclostationary detection. In both Chapter 4 and Chapter 5, the imperfect design of the wideband filters will influence the probability of successful full spectrum reconstruction. For example, a narrower bandwidth than the required one can make the frequencies in the sub-Nyquist spectrum congested, therefore more risk of being overlapped. In such a scenario, the performance degrades due to the missed detection of PUs.

6.3 Future Work

There are a number of research areas which have not been addressed in this thesis. Some of them merit much more work at a future stage as listed below:

- In Chapter 3, comparisons have been made between MRC, SC, SLC, and SLS strategies when the infrastructure of the cognitive radio network is centralised. For a comprehensive study of the collaborative spectrum sensing algorithms, more collaborative protocols should be involved for comparisons, e.g., decision fusion strategies, and equal gain combining (EGC) [98, 99] based collaborative spectrum sensing scheme.

- The performance of collaborative spectrum sensing algorithms over a composite fading channel deserves investigations, e.g., composite multipath/shadowed fading. There are various combinations suggested in the literature for obtaining the composite distribution, e.g., Suzuki distribution [100, 101], and K distribution [102, 103].

- In Chapter 4, to further reduce the transmission and the infrastructure costs, a multi-hop [60, 61, 104] implementation of MASS is attractive. In such a scenario, the spectrum sensing data will be transmitted via different spectrum holes, instead of common control channels. Thus, the spectra viewed by distributed cognitive radios are different at some

frequencies. This spectral difference will make the spectrum reconstruction become difficult. New spectrum reconstruction algorithms can be developed to address these issues.

- In Chapter 5, MSD can be implemented in a decision fusion manner that each cognitive radio decides the spectral occupancy status based on individual observations, and then transmits their decisions to the FC. This strategy will not only reduce sampling rate, but also significantly save transmission bandwidth. Furthermore, a sequential detection [105, 106] approach can be used to give a quickest detection result.

- In both Chapter 4 and Chapter 5, in order to let ADCs observe the signals of the same bandwidth, the wideband filters prior to ADCs are assumed to have the same bandwidth. In these cases, the spectra between ADCs are similar, and the sampling rates are different. It has been found that the observed SNR decreases when the bandwidth of the wideband filter increases. To improve the observed SNR in each cognitive radio, it is interesting to develop a system with different medium-band filters and different sub-Nyquist sampling rates under the condition that the full spectrum is still recoverable.

Appendix A
Proof of Proposition 4.1

An example, $M^1 = 3$, $M^2 = 5$, and $N = 9$, is used to demonstrate the structure of the measurement matrix. The measurement matrix $\mathbf{\Phi}$ can be constructed using (4.14) and (4.18), then the ℓ_2 column-wise normalised matrix $\hat{\mathbf{\Phi}}$ is,

$$\hat{\mathbf{\Phi}} = \begin{bmatrix} 0.5145 & 0 & 0 & 0.5145 & 0 & 0 & 0.5145 & 0 & 0 \\ 0 & 0.5145 & 0 & 0 & 0.5145 & 0 & 0 & 0.5145 & 0 \\ 0 & 0 & 0.5145 & 0 & 0 & 0.5145 & 0 & 0 & 0.5145 \\ 0 & 0 & 0.8575 & 0 & 0 & 0 & 0 & 0.8575 & 0 \\ 0 & 0 & 0 & 0.8575 & 0 & 0 & 0 & 0 & 0.8575 \\ 0 & 0 & 0 & 0 & 0.8575 & 0 & 0 & 0 & 0 \\ 0.8575 & 0 & 0 & 0 & 0 & 0.8575 & 0 & 0 & 0 \\ 0 & 0.8575 & 0 & 0 & 0 & 0 & 0.8575 & 0 & 0 \end{bmatrix}.$$

Using (4.14), the j-th normalised column $\hat{\phi}_j[m]$ can be expressed as,

$$\begin{aligned} \hat{\phi}_j[m] = & \frac{M^l}{\sqrt{\sum_{k=1}^{v}(M^k)^2}} \delta\Big[m - \Big|j - \Big\lfloor \frac{N}{2} \Big\rfloor + \Big\lfloor \frac{M^l}{2} \Big\rfloor \Big|_{\text{mod}^\star(M^l)}\Big], \\ & \forall j \in [1, N],\ l = \left\{ l | m \in \left[\sum_{k=0}^{l-1} M^k + 1, \sum_{k=0}^{l} M^k \right] \right\} \end{aligned} \tag{A.1}$$

where $M^0 = 0$ and $|\bullet|_{\text{mod}^\star(\oplus)}$ is a modified modulo operation in the sense of using $\oplus$ to replace all zeros in the results of the standard modulo operation.

Let us define the event m_l as there exists a value of $m \in [\sum_{k=0}^{l-1} M^k + 1, \sum_{k=0}^{l} M^k]$ such that $\hat{\phi}_j[m] = \hat{\phi}_h[m] > 0$, where $j \neq h \in [1, N]$. Then the average probability of the event m_l occurring is $Pr(m_l) = \frac{\mathrm{C}_{M^l}^1 \mathrm{C}_{D_l}^2}{\mathrm{C}_N^2} = \frac{D_l - 1}{N-1} = \frac{N/M^l - 1}{N-1}$, where $D_l = N/M^l$ is the undersampling factor in the l-th sampling channel. Obviously, any single event m_l ($l \in [1, v]$) could occur when $M^l < N$ because $Pr(m_l) > 0$. In the case of only one event m_i occurring, the mutual

coherence of the matrix $\mathbf{\Phi}$ can be calculated by using (4.2) and (A.1) as,

$$\mu = \max_{j \neq h} | < \hat{\phi}_j, \hat{\phi}_h > | = \max_{i \in [1,v]} \left| \frac{M^i}{\sqrt{\sum_{l=1}^{v} (M^l)^2}} \cdot \frac{M^i}{\sqrt{\sum_{l=1}^{v} (M^l)^2}} \right| = \frac{\max_{i \in [1,v]} (M^i)^2}{\sum_{l=1}^{v} (M^l)^2}. \tag{A.2}$$

Then next target is to make the probability of two events occurring to be zero. By the definition of the event m_l, it satisfies $|j - \lfloor \frac{N}{2} \rfloor + \lfloor \frac{M^l}{2} \rfloor|_{\text{mod}^\star(M^l)} = |h - \lfloor \frac{N}{2} \rfloor + \lfloor \frac{M^l}{2} \rfloor|_{\text{mod}^\star(M^l)}$, where $j \neq h \in [1, N]$. It is equivalent to $|h - j| = w_l M^l$ for $w_l \in \mathbb{Z} \cap [1, \lceil \frac{N}{M^l} \rceil - 1]$. Similarly, the event m_z happens when $|h - j| = w_z M^z$ for $w_z \in \mathbb{Z} \cap [1, \lceil \frac{N}{M^z} \rceil - 1]$. Both of them happen when

$$w_z M^z = w_l M^l, \quad w_l \in \mathbb{Z} \cap \left[1, \left\lceil \frac{N}{M^l} \right\rceil - 1\right], w_z \in \mathbb{Z} \cap \left[1, \left\lceil \frac{N}{M^z} \right\rceil - 1\right]. \tag{A.3}$$

The probability $Pr(m_z m_l)$ is the probability of both the event m_z and the event m_l satisfying equation (A.3). This is equivalent to being given M^l and M^z, and then finding how many pairs of numbers w_l and w_z out of the available number of $(\lceil \frac{N}{M^l} \rceil - 1)$ or $(\lceil \frac{N}{M^z} \rceil - 1)$ are required for equation (A.3) to hold. Without loss of generality, considering the possibility for w_l, the probability $Pr(m_z m_l)$ is given by,

$$Pr(m_z m_l) = \frac{\left\lfloor \frac{\lceil \frac{N}{M^l} \rceil - 1}{\ddot{w}_l} \right\rfloor}{\lceil \frac{N}{M^l} \rceil - 1}, \tag{A.4}$$

because of w_l should be integer multiples of $\ddot{w}_l$ as $w_l = w_z \frac{M^z}{M^l} = \ddot{w}_l \frac{w_z}{w'_z}$, where $\ddot{w}_l$ and $\ddot{w}_z$ are from simplest fraction form of,

$$\frac{w_l}{w_z} = \frac{M^z}{M^l} = \frac{a\ddot{w}_l}{a\ddot{w}_z} = \frac{\ddot{w}_l}{\ddot{w}_z}, \tag{A.5}$$

where a denotes an integer.

If assume that M^l and M^z are different prime numbers, then $\ddot{w}_l = M^z$. In order to let $Pr(m_z m_l) = 0$, which is equivalent to $M^z = \ddot{w}_l > \lceil \frac{N}{M^l} \rceil - 1$, M^l and M^z should satisfy,

$$M^l M^z > N, \quad \forall l, z \in [1, v], l \neq z \tag{A.6}$$

because $M^z > \frac{N}{M^l} \geq \lceil \frac{N}{M^l} \rceil - 1$.

In summary, in order to prevent any two events (m_l and m_z, $l \neq z$) from happening together, M^l and M^z should at least satisfy equation (A.6). Furthermore, if the probability that m_l and m_z occur together is zero, then the probability that more than two events (m_l, m_z...) occur together is zero too. It is equivalent to the condition that the maximum correlation of different columns in $\mathbf{\Phi}$ exists when only one event happens as,

$$\mu = \max_{j \neq h} | < \hat{\phi}_j, \hat{\phi}_h > | = \frac{\max_{i \in [1,v]} (M^i)^2}{\sum_{l=1}^{v} (M^l)^2}. \tag{A.7}$$

Hence, Proposition 4.1 is proved. □

Appendix B
Proof of Proposition 4.2

Based on Doob's maximal inequality [107], the following inequality holds,

$$Pr(\mu > x) = Pr\left(\max_{j \neq h} | < \hat{\phi}_j, \hat{\phi}_h > | > x\right) \leq \frac{\mathrm{E}(< \hat{\phi}_j, \hat{\phi}_h >)}{x}, \tag{B.1}$$

where $\mathrm{E}(a)$ denotes the expectation of a.

Given the proof for Proposition 4.1 in Appendix A, when the conditions in (4.22) are satisfied, there exists at most one value of $m \in [1, M]$ such that $\hat{\phi}_j[m] = \hat{\phi}_h[m] > 0$. Then the expected value of $< \hat{\phi}_j, \hat{\phi}_h >$ becomes,

$$\mathrm{E}(< \hat{\phi}_j, \hat{\phi}_h >) = \frac{(M^1)^2}{\sum_{l=1}^{v}(M^l)^2} Pr(m_1) + \frac{(M^2)^2}{\sum_{l=1}^{v}(M^l)^2} Pr(m_2) + \cdots + \frac{(M^v)^2}{\sum_{l=1}^{v}(M^l)^2} Pr(m_v), \tag{B.2}$$

where $Pr(m_l)$ denotes the average probability of the event m_l occuring as defined in Appendix A.

Substituting $Pr(m_l) = \frac{N/M^l - 1}{N-1} < 1/M^l$ into (B.2), the following inequality holds,

$$\mathrm{E}(< \hat{\phi}_j, \hat{\phi}_h >) < \frac{\sum_{l=1}^{v} M^l}{\sum_{l=1}^{v}(M^l)^2}. \tag{B.3}$$

Replacing x in (B.1) by $x = \frac{1}{2k-1}$, and substituting (B.3) into (B.1), the following inequalities can be obtained,

$$Pr\left(\mu > \frac{1}{2k-1}\right) \leq (2k-1)\mathrm{E}(< \hat{\phi}_j, \hat{\phi}_h >) < \frac{\sum_{l=1}^{v} M^l}{\sum_{l=1}^{v}(M^l)^2}(2k-1). \tag{B.4}$$

If the final term is less than ϵ, that is,

$$\frac{\sum_{l=1}^{v}(M^l)^2}{\sum_{l=1}^{v} M^l} > \frac{2k-1}{\epsilon}, \tag{B.5}$$

the following equation holds true,

$$Pr\left(\mu < \frac{1}{2k-1}\right) > 1 - \epsilon. \tag{B.6}$$

On applying Theorem 4.1, the proof of Proposition 4.2 is completed. □

Appendix C
Proof of Proposition 4.3

Inspired by [108], the Gram matrix of the measurement matrix will be used for the proof. Let Ω be the set of indices of $2k$-nonzero components in spectrum $\overrightarrow{X}$. The sub-matrix $\mathbf{\Phi_\Omega}$ can be obtained by selecting the columns, whose indices are in Ω, in the matrix $\widehat{\mathbf{\Phi}}$ (the ℓ_2 normalised $\mathbf{\Phi}$). The Gram matrix of $\mathbf{\Phi_\Omega}$ is expressed as,

$$\mathbf{G_\Omega} = \mathbf{\Phi_\Omega^H \Phi_\Omega}, \tag{C.1}$$

where $\mathbf{\Phi^H}$ denotes Hermitian transpose of $\mathbf{\Phi}$. Then $\mathbf{G_\Omega}$ is a symmetric and nonnegative definite matrix and can be written as,

$$\mathbf{G_\Omega} = \mathbf{I_{2k\times 2k}} + \mathbf{B_{2k\times 2k}}. \tag{C.2}$$

where $\mathbf{I_{2k\times 2k}}$ is an identity matrix, and $\mathbf{B_{2k\times 2k}}$ has values only in the off-diagonal entries.

The maximal entries $B^\star$ in $\mathbf{B_{2k\times 2k}}$ will be,

$$B^\star = \max_{j\neq h} | < \hat{\phi}_j, \hat{\phi}_h > | = \frac{\max_{i\in[1,v]}(M^i)^2}{\sum_{l=1}^{v}(M^l)^2}, \quad \forall j,h \in [1,2k], \tag{C.3}$$

The Gersgorin discs theorem [109] states that all the eigenvalues λ_i of $\mathbf{G_\Omega}$ are located in the union of $2k$ discs, then,

$$|\lambda_i - I_{ii}| \leq R_i(\mathbf{G_\Omega}) \leq B^\star(2k-1), \quad \forall i \in [1,2k]. \tag{C.4}$$

where $R_i(\mathbf{G_\Omega})$ denotes Deleted absolute row sums of $\mathbf{G_\Omega}$.

Thus, using the Triangle inequality, the eigenvalue of $\mathbf{G_\Omega}$, i.e. λ_i, has the bounds,

$$|I_{ii}| - |\lambda_i - I_{ii}| \leq |\lambda_i| \leq |I_{ii}| + |\lambda_i - I_{ii}|, \quad \forall i \in [1,2k]. \tag{C.5}$$

Substituting (C.4) into above equation, we end up with,

$$1 - B^{\star}(2k-1) \leq |\lambda_i| \leq 1 + B^{\star}(2k-1), \quad \forall i \in [1, 2k]. \tag{C.6}$$

Simply assume that $\varrho_{2k} = B^{\star}(2k-1)$, all the eigenvalues of $\mathbf{G}_{\mathbf{\Omega}}$ are in $[1-\varrho_{2k}, 1+\varrho_{2k}]$. Hence, apply Theorem 4.2, the Proposition 4.3 follows. □

Appendix D
Original publications

D.1 Journal papers

- H. Sun, D. I. Laurenson, and C.-X. Wang, "Computationally Tractable Models of Energy Detection Performance over Slow Fading Channels", *IEEE Communications Letters*, vol. 14, no. 10, pp. 924-926, October 2010.

- H. Sun, D. I. Laurenson, J. S. Thompson, M. E. Davies, M. Yaghoobi, and C.-X. Wang, "Distributed Wideband Spectrum Sensing by Multirate Asynchronous Sub-Nyquist Sampling", *IEEE Transactions on Wireless Communications*, submitted in June 2010.

- H. Sun, D. I. Laurenson, and C.-X. Wang, "Energy Detection of Unknown Signals over Fast Fading Channels in Cognitive Radios", *IEEE Communications Letters*, submitted in July 2010, revised in November 2010.

D.2 Conference papers

- H. Sun, D. I. Laurenson, J. S. Thompson, and C.-X. Wang, "Performance Analysis of Diversity Reception for Spectrum Sensing in Cognitive Radio Networks", *IEEE International Conference on Communications (ICC 2011)*, submitted in September 2010.

- H. Sun, D. I. Laurenson, J. Jiang, and C.-X. Wang, "A Novel Distributed Wideband Spectrum Sensing System for Cognitive Radio Networks", *IEEE International Conference on Communications (ICC 2011)*, submitted in September 2010.

- H. Sun, D. I. Laurenson, and J. S. Thompson, "Cooperative Compressive Spectrum Sensing by Sub-Nyquist Sampling", in *Proc. IEEE First UK-India International Workshop on Cognitive Wireless Systems (UKIWCWS 2009)*, IIT Delhi, India, pp. 1-6, December 2009.

- H. Sun, D. I. Laurenson, J. S. Thompson, and C.-X. Wang, “A Novel Centralized Network for Sensing Spectrum in Cognitive Radio”, in *Proc. IEEE International Conference on Communications (ICC 2008)*, Beijing, China, pp. 4186-4190, May 2008.

Computationally Tractable Model of Energy Detection Performance over Slow Fading Channels

Hongjian Sun, *Student Member, IEEE,* David I. Laurenson, *Member, IEEE,*
and Cheng-Xiang Wang, *Senior Member, IEEE*

***Abstract*—Energy detection (ED) has been widely used for detecting unknown deterministic signals in many wireless communication applications, e.g., cognitive radio, and ultra-wideband (UWB). However, the performance analysis of ED over slow fading channels is cumbersome, because it is difficult to derive closed-form expressions for the average probability of detection involving the generalised Marcum Q-function and the log-normal distribution. In this letter, we derive an approximation of the average probability of detection over a slow fading channel by replacing the log-normal distribution with a Wald distribution. In addition, we analyze the detection performance of the ED using a square-law combining scheme over multiple independent and identically distributed slow fading channels.**

***Index Terms*—Energy detection, slow fading, cognitive radio, UWB, Wald distribution.**

I. Introduction

SINCE Urkowitz's seminal paper [1], energy detection (ED, also known as radiometry) has been widely used for detecting unknown deterministic signals in many applications. For example, [2] employed ED for detecting the presence of ultra-wideband (UWB) signals, and [3] used ED to study the effect of the collaboration among cognitive radios. ED is commonly used not only due to its low computational and implementation complexity, but also because it does not need any prior knowledge of signals.

A non-fading additive white Gaussian noise (AWGN) channel is usually assumed when studying the performance of ED. However, in wireless communication applications, fading occurs because of multipath propagation and shadowing. The detection performance of ED over a variety of fading channels has gained interest recently in [4]–[6]. In [4], Digham *et al.* evaluated the performance of ED over Nakagami-m fading channels and Rician fading channels. The influence of the slow fading on ED was numerically studied in [7], but without a closed-form expression. The challenge of deriving closed-form expressions for the average probability of detection stems from the fact that it involves both the generalised Marcum Q-function and the log-normal distribution.

Manuscript received June 1, 2010. The associate editor coordinating the review of this letter and approving it for publication was H.-H. Chen.

H. Sun and D. I. Laurenson are with the Institute for Digital Communication and with the Joint Research Institute for Signal and Image Processing, University of Edinburgh, Kings Buildings, Mayfield Road, Edinburgh, EH9 3JL, UK (e-mail:mrhjsun@hotmail.com, dave.laurenson@ed.ac.uk).

C.-X. Wang is with the Joint Research Institute for Signal and Image Processing, School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh, EH14 4AS, UK (e-mail: cheng-xiang.wang@hw.ac.uk).

The authors acknowledge the support of Wolfson Microelectronics Scholarship, and the Scottish Funding Council for the Joint Research Institute with the Heriot-Watt University which is a part of the Edinburgh Research Partnership.

Digital Object Identifier 10.1109/LCOMM.2010.10.100934

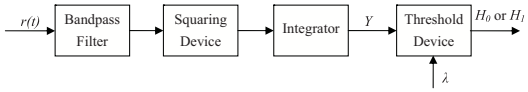

Fig. 1. Block diagram of the energy detector.

The contribution in this study is twofold. First, we give an approximation of the average probability of detection for the slow fading channel, by using the Wald distribution to replace the log-normal distribution. Second, we analyze the detection performance of the ED using a square-law combining (SLC) scheme over multiple independent and identically distributed (i.i.d.) slow fading channels, and derive a computationally tractable expression.

This paper is organized as follows. We briefly introduce the system model in Section II. In Section III, we derive an expression for the average probability of detection over a single slow fading channel. The detection performance of the ED using SLC over multiple slow fading channels is given in Section IV. Simulation results are presented in Section V, with conclusions given in Section VI.

II. System Model

A block diagram of an energy detector is shown in Fig. 1. The received signal, $r(t)$, is filtered by a bandpass filter (BPF), followed by a squaring device for measuring received energy, and an integrator that controls the observation interval, T. In order to decide whether the signal is present or not, the output of the integrator, Y, will act as a test statistic, and will be compared with a predetermined threshold, λ. The binary signal detection problem can be formulated as hypothesis test with H_0 (signal not present) or H_1 (signal present),

$$\begin{aligned} H_0: &\quad r(t) = n(t), \\ H_1: &\quad r(t) = h(t)\, s(t) + n(t), \end{aligned} \tag{1}$$

where $h(t)$ denotes the complex channel gain between the transmitter and the receiver, $s(t)$ denotes the bandlimited signal coming from the transmitter of unknown modulation format, and $n(t)$ is AWGN.

Following [1], the test statistic, Y, has the following distribution,

$$Y \sim \begin{cases} \chi^2_{2u}, & H_0 \\ \chi^2_{2u}(2\gamma), & H_1 \end{cases} \tag{2}$$

where "$\sim$" means "distributed as", γ is the signal-to-noise ratio (SNR) at the receiver, and χ^2_{2u} and $\chi^2_{2u}(2\gamma)$ denote the central and non-central chi-square distributions, respectively. Both distributions have the same degree of freedom (DoF),

$2u$ (u is the time bandwidth product), and the latter one has a non-central parameter 2γ.

The probability density function (PDF) of Y is given as [4]

$$f_Y(y) = \begin{cases} \frac{1}{2^u \cdot \Gamma(u)} \cdot y^{u-1} \cdot e^{-\frac{y}{2}}, & H_0 \\ \frac{1}{2} \cdot (\frac{y}{2\gamma})^{\frac{u-1}{2}} \cdot e^{-\frac{2\gamma+y}{2}} \cdot I_{u-1}(\sqrt{2\gamma y}), & H_1 \end{cases} \quad (3)$$

where $\Gamma(a)$ is the gamma function and $I_v(a)$ is the v^{th} order modified Bessel function of the first kind.

For a non-fading AWGN channel, the probabilities of false alarm and detection are given in [4] as below

$$P_{fa} = P_r(Y > \lambda | H_0) = \frac{\Gamma(u, \lambda/2)}{\Gamma(u)}, \quad (4)$$

$$P_d = P_r\left(Y > \lambda | H_1\right) = Q_u(\sqrt{2\gamma}, \sqrt{\lambda}), \quad (5)$$

where $\Gamma(a, x)$ denotes the incomplete gamma function given by $\Gamma(a, x) = \int_x^\infty t^{a-1} e^{-t} dt$, and $Q_u(a, x)$ denotes the generalised Marum Q-function given by

$$Q_u(a, x) = \frac{1}{a^{u-1}} \int_x^\infty t^u e^{-\frac{a^2+t^2}{2}} I_{u-1}(at) dt. \quad (6)$$

Another form of the generalised Marcum Q-function is given in (4.74) of [8] as

$$Q_u(\sqrt{2\gamma}, \sqrt{\lambda}) = \sum_{n=0}^{\infty} e^{-\gamma} \frac{\gamma^n}{n!} \sum_{l=0}^{n+u-1} e^{-\frac{\lambda}{2}} \frac{\left(\frac{\lambda}{2}\right)^l}{l!}. \quad (7)$$

With the aid of (8.352-2) in [9], we can rewrite above equation as

$$Q_u(\sqrt{2\gamma}, \sqrt{\lambda}) = \sum_{n=0}^{\infty} \frac{\gamma^n}{n!} \frac{\Gamma(n+u, \frac{\lambda}{2})}{\Gamma(n+u)} e^{-\gamma}. \quad (8)$$

III. Local Energy Detection in A Slow Fading Channel

When experiencing a fading channel, P_{fa} in (4) will remain the same, since it is independent of the SNR. On the other hand, when the channel gain, $h(t)$, varies, the average probability of detection can be calculated by averaging P_d in (5) over the SNR distribution as

$$\overline{P_d} = \int_0^\infty P_d(\gamma, \lambda) f(\gamma) d\gamma = \int_0^\infty Q_u(\sqrt{2\gamma}, \sqrt{\lambda}) f(\gamma) d\gamma, \quad (9)$$

where $f(\gamma)$ denotes the PDF of the SNR in a fading channel.

In terrestrial land-mobile wireless communication systems, the received SNR may be affected by the effect of shadowing due to objects obstructing the propagation path [8]. Empirical measurements showed that the received power fluctuates with a log-normal distribution about the area-mean power for various outdoor and indoor environments [7], [10]. The PDF of the SNR is given by [8]

$$f(\gamma) = \frac{\xi}{\sqrt{2\pi}\sigma\gamma} \exp\left(-\frac{(10\log_{10}\gamma - \mu)^2}{2\sigma^2}\right), \quad \gamma > 0, \quad (10)$$

where $\xi = 10/\ln(10)$, μ (dB) denotes the area-mean SNR, and σ (dB) denotes the standard deviation of $10\log_{10}\gamma$. To the best of our knowledge, there exists no closed-form expression for the average probability of detection when we substitute (8) and (10) into (9). The log-normal distribution can be closely approximated by the Wald distribution (also known as the inverse Gaussian distribution) [11], [12], whose PDF is given by

$$f(\gamma) = \sqrt{\frac{\eta}{2\pi}} \gamma^{-3/2} \exp\left(-\frac{\eta(\gamma-\theta)^2}{2\theta^2\gamma}\right), \quad \gamma > 0, \quad (11)$$

where $\theta = \mathrm{E}(\gamma)$ denotes the expectation of γ, and η is the shape parameter. The variance of γ is $\frac{\theta^3}{\eta}$, i.e., $\mathrm{Var}(\gamma) = \frac{\theta^3}{\eta}$. We propose to use the Wald distribution to approximate the log-normal distribution. In order to do so, by the method of moments we relate parameters η, θ with μ, σ as below

$$\begin{aligned} \theta &= \exp\left(\frac{\mu}{\xi} + \frac{\sigma^2}{2\xi^2}\right), \\ \eta &= \frac{\theta}{\exp(\frac{\sigma^2}{\xi^2}) - 1}. \end{aligned} \quad (12)$$

Substituting (8) and (11) into (9), with manipulation we obtain

$$\begin{aligned} \overline{P_{d,Sha}} &= \sqrt{\frac{\eta}{2\pi}} e^{\frac{\eta}{\theta}} \sum_{n=0}^{\infty} \frac{\Gamma(n+u, \frac{\lambda}{2})}{\Gamma(n+u) n!} \\ &\quad \times \int_0^\infty \gamma^{n-\frac{3}{2}} e^{(-\frac{\eta/2}{\gamma} - \frac{2\theta^2+\eta}{2\theta^2}\gamma)} d\gamma. \end{aligned} \quad (13)$$

Using (3.471-9) in [9] for calculating the integral, we obtain

$$\begin{aligned} \overline{P_{d,Sha}} &= \sqrt{\frac{2\eta}{\pi}} e^{\frac{\eta}{\theta}} \sum_{n=0}^{\infty} \frac{\Gamma(n+u, \frac{\lambda}{2})}{\Gamma(n+u) n!} \\ &\quad \times \left(\sqrt{\frac{\eta\theta^2}{2\theta^2+\eta}}\right)^{n-\frac{1}{2}} \mathrm{K}_{n-\frac{1}{2}}\left(\frac{\sqrt{\eta(2\theta^2+\eta)}}{\theta}\right), \end{aligned} \quad (14)$$

where $\mathrm{K}_{n-\frac{1}{2}}(a)$ denotes the modified Bessel function of the second kind with order $n - \frac{1}{2}$.

The truncation error, T_e, will be involved when using finite summations, N, to replace infinite summations in (14). Because $\frac{\Gamma(n,b)}{\Gamma(n)} = e^{-b} \sum_{i=0}^{n-1} \frac{b^i}{i!}$ can be viewed as the cumulative distribution function for a Poisson random variable of $X \sim \mathrm{Poi}(b)$, which results in $\frac{\Gamma(n,b)}{\Gamma(n)} \leq 1$, the truncation error is bounded by

$$\begin{aligned} T_e &\leq \sqrt{\frac{2\eta}{\pi}} e^{\frac{\eta}{\theta}} \sum_{n=N+1}^{\infty} \frac{\left(\sqrt{\frac{\eta\theta^2}{2\theta^2+\eta}}\right)^{n-\frac{1}{2}}}{n!} \mathrm{K}_{n-\frac{1}{2}}\left(\frac{\sqrt{\eta(2\theta^2+\eta)}}{\theta}\right), \\ &= 1 - \sqrt{\frac{2\eta}{\pi}} e^{\frac{\eta}{\theta}} \sum_{n=0}^{N} \frac{\left(\sqrt{\frac{\eta\theta^2}{2\theta^2+\eta}}\right)^{n-\frac{1}{2}}}{n!} \mathrm{K}_{n-\frac{1}{2}}\left(\frac{\sqrt{\eta(2\theta^2+\eta)}}{\theta}\right). \end{aligned} \quad (15)$$

IV. Energy Detection over Slow Fading Channels

The detection result of a single receiver may not be sufficiently reliable, which might be due to either the effect of fading or a low SNR. In such a scenario, diversity schemes are often employed because they can combat the severe fading [8]. Using SLC, the energy vectors, $Y_1, Y_2, \cdots, Y_L$, from L distributed receivers are gathered at a fusion centre (FC), where the test statistic, $Y_{slc} = \sum_{i=1}^{L} Y_i$ is formed [13]. When these L fading channels are i.i.d., and all branches have the same noise variance, the fused energy, Y_{slc}, has the following distribution,

$$Y_{slc} \sim \begin{cases} \chi^2_{2Lu}(2\gamma_{slc}), & H_1 \\ \chi^2_{2Lu}, & H_0, \end{cases} \quad (16)$$

where $\gamma_{slc} = \sum_{i=1}^{L} \gamma_i$ is given by [13].

In the case of non-fading AWGN channels, the probabilities of false alarm and detection under a SLC scheme can be given as below

$$P'_{fa} = \frac{\Gamma(Lu, \lambda/2)}{\Gamma(Lu)}, \tag{17}$$

$$\begin{aligned} P'_d &= Q_{Lu}(\sqrt{2\gamma_{slc}}, \sqrt{\lambda}), \\ &= \sum_{n=0}^{\infty} \frac{\gamma_{slc}^n}{n!} \frac{\Gamma(n+Lu, \frac{\lambda}{2})}{\Gamma(n+Lu)} e^{-\gamma_{slc}}. \end{aligned} \tag{18}$$

When the signal experiences fading over L channels, the average probability of false alarm will remain the same as (17), and the average probability of detection can be evaluated by averaging P'_d over the SNR distribution as

$$\begin{aligned} \overline{P'_d} &= \int_0^{\infty} P'_d(\gamma_{slc}, \lambda) f(\gamma_{slc}) d\gamma_{slc}, \\ &= \int_0^{\infty} Q_{Lu}(\sqrt{2\gamma_{slc}}, \sqrt{\lambda}) f(\gamma_{slc}) d\gamma_{slc}. \end{aligned} \tag{19}$$

In slow fading channels, the PDF of the SNR in the node i, γ_i, can be approximated by a Wald distribution. When all fading channels are stationary and i.i.d., the condition $\frac{\eta_i}{\theta_i^2} = \frac{\mathrm{E}(\gamma_i)}{\mathrm{Var}(\gamma_i)} = b$ (constant) can be satisfied. Thus, the combined SNR under the SLC scheme, γ_{slc}, will also follow the Wald distribution [14]. The PDF of γ_{slc} can be easily obtained by replacing each η with $L\eta$, each θ with $L\theta$, and each γ with γ_{slc} in (11). Using a similar method to that of the single slow fading channel, we can obtain the average probability of detection as below

$$\begin{aligned} \overline{P'_{d,Sha}} = &\sqrt{\frac{2L\eta}{\pi}} e^{\frac{\eta}{\theta}} \sum_{n=0}^{\infty} \frac{\Gamma(n+Lu, \frac{\lambda}{2})}{\Gamma(n+Lu) n!} \\ &\times \left(\sqrt{\frac{\eta\theta^2 L^2}{2L\theta^2+\eta}} \right)^{n-\frac{1}{2}} \mathrm{K}_{n-\frac{1}{2}} \left(\frac{\sqrt{\eta(2L\theta^2+\eta)}}{\theta} \right). \end{aligned} \tag{20}$$

The above result can also be obtained by replacing each η with $L\eta$, each θ with $L\theta$, and each u with Lu in (14).

V. Simulation results

Receiver operating characteristic (ROC) analysis has been widely used in the signal detection theory. It is an ideal technique to quantify the tradeoff between the probability of detection and the probability of false alarm. In the simulation, we use complementary ROC curves (P_{fa} vs $1 - P_d$) to show the detection performance of ED over slow fading channels. As we used the Wald distribution to approximate the log-normal distribution for deriving the average probability of detection, we compare the theoretical result in (14) with simulated result in Fig. 2. From both figures, we can find that the theoretical results closely fit the experimental results. In addition, we can see that, when the average probability of false alarm decreases, the approximation error slightly increases. This phenomenon may stem from the long right tail of the long-normal distribution, which is difficult to match. As shown in the right-hand figure, the mismatch becomes larger when the shadow standard deviation, σ, becomes larger (equivalent to a longer right tail in the log-normal distribution).

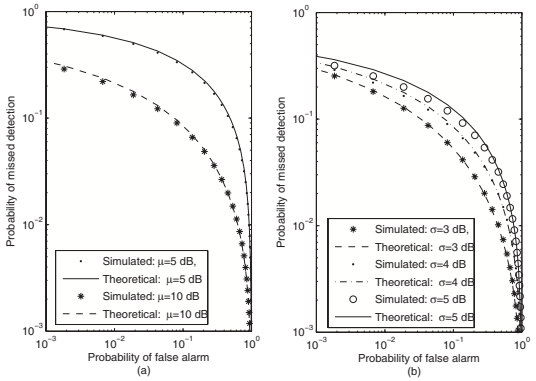

Fig. 2. Complementary ROC curves of energy detection over a slow fading channel with (a) the shadow standard deviation $\sigma = 4$ dB, and (b) the area-mean SNR $\mu = 10$ dB, compared with theoretical result in (14).

VI. Conclusions

In this study, we have obtained a computationally tractable expression for the average probability of detection over a slow fading channel, by using the Wald distribution to replace the log-normal distribution. Using SLC, we have studied the detection performance of ED over i.i.d. slow fading channels. It has been shown that the theoretical expression closely match the experimental results. Since the effect of the long right tail in the log-normal distribution, the mismatch becomes larger when the shadow standard deviation increases.

References

[1] H. Urkowitz, "Energy detection of unknown deterministic signals," *Proc. IEEE*, vol. 55, no. 4, pp. 523–531, Apr. 1967.

[2] C. Steiner and A. Wittneben, "Low complexity location fingerprinting with generalized UWB energy detection receivers," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1756–1767, Mar. 2010.

[3] C.-X. Wang, H.-H. Chen, X. Hong, and M. Guizani, "Cognitive radio network management: tuning in to real-time conditions," in *IEEE Veh. Technol. Mag.*, vol. 3, no. 1, pp. 28–35, Mar. 2008.

[4] F. F. Digham, M.-S. Alouini, and M. K. Simon, "On the energy detection of unkown signals over fading channels," in *Proc. IEEE ICC*, May 2003, pp. 3575–3579.

[5] V. Kostylev, "Energy detection of a signal with random amplitude," in *Proc. IEEE ICC*, 2002, pp. 1606–1610.

[6] K. Ruttik, K. Koufos, and R. Jantti, "Detection of unknown signals in a fading environment," *IEEE Commun. Lett.*, vol. 13, no. 7, pp. 498–500, July 2009.

[7] A. Ghasemi and E. Sousa, "Collaborative spectrum sensing for opportunistic access in fading environments," in *Proc. IEEE DySPAN*, Nov. 2005, pp. 131–136.

[8] M. K. Simon and M.-S. Alouini, *Digital Communication over Fading Channels*, 2nd edition. John Wiley & Sons, Inc., 2004.

[9] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 5th edition, A. Jeffrey, editor. Academic Press, Inc., 1994.

[10] H. Suzuki, "A statistical model for urban radio propagation," *IEEE Trans. Commun.*, vol. 25, no. 7, pp. 673–680, July 1977.

[11] A. H. Marcus, "Power sum distributions: an easier approach using the Wald distribution," *J. American Statistical Assoc.*, vol. 71, pp. 237–238, 1976.

[12] Karmeshu and R. Agrawal, "On efficacy of Rayleigh-inverse Gaussian distribution over K-distribution for wireless fading channels," *Wireless Commun. and Mobile Computing*, vol. 7, no. 1, pp. 1–7, 2007.

[13] F. F. Digham, M.-S. Alouini, and M. K. Simon, "On the energy detection of unknown signals over fading channels," *IEEE Trans. Commun.*, vol. 55, no. 1, pp. 21–24, Jan. 2007.

[14] R. S. Chhikara and J. L. Folks, *The Inverse Gaussian Distribution: Theory, Methodology, and Applications*. Marcel Dekker Inc., 1989.

Distributed Wideband Spectrum Sensing by Multirate Asynchronous Sub-Nyquist Sampling

Hongjian Sun, *Student Member, IEEE,* David I. Laurenson, *Member, IEEE,* John S. Thompson, *Member, IEEE,* Mike E. Davies, *Member, IEEE,* Mehrdad Yaghoobi, *Member, IEEE,* Cheng-Xiang Wang, *Senior Member, IEEE,*

Abstract

Wideband spectrum sensing is becoming increasingly important to modern electronic systems, e.g., cognitive radio (CR), for rapidly identifying spectrum holes or characterizing interference. On the other hand, distributed spectrum sensing has been widely suggested for combating the detrimental effects of fading. However, there is a significant challenge in sensing the wideband spectrum in a distributed CR network. Because both the high sampling rates and data transmission are problematic for battery powered CRs. In order to implement wideband spectrum sensing efficiently, we present a multirate asynchronous sub-Nyquist sampling (MASS) model. When the MASS is applied to a single CR node, sufficient conditions for uniquely reconstructing the full spectrum using basis pursuit, or orthogonal matching pursuit algorithms are derived. We also apply the proposed system to distributed CR networks. When the spectra observed by CR nodes have a common spectral support, using one lowrate ADC in each CR node can also successfully recover the full spectrum, by applying a matching pursuit algorithm.

Index Terms

Cognitive radio, Spectrum sensing, Compressed sensing, Multirate sampling.

H. Sun, D. I. Laurenson, J. S. Thompson, M. E. Davies and M. Yaghoobi are with the Institute for Digital Communication and with the Joint Research Institute for Signal and Image Processing, University of Edinburgh, Kings Buildings, Mayfield Road, Edinburgh, EH9 3JL, UK (Email: {H.Sun, Dave.Laurenson, John.Thompson, Mike.Davies}@ed.ac.uk, Yaghoobi@ieee.org).

C.-X. Wang is with the Joint Research Institute for Signal and Image Processing, School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh, EH14 4AS, UK (Email: cheng-xiang.wang@hw.ac.uk)

I. INTRODUCTION

A. Cognitive radio networks

The existing fixed frequency allocation of radio spectrum typically results in significant underutilisation of the available frequencies [1]. For example, the maximal occupancy of the spectrum from 30 MHz to 3 GHz (in New York city) has been reported to be only 13.1%, with average occupancy (over six locations) of 5.2% [2]. The spectral underutilisation can be solved by allowing a secondary user to access a licensed band when the primary user (PU) is absent. Cognitive radio (CR) [3] is widely agreed to be the most promising method for exploiting RF spectral scarcity [4]. A crucial requirement of CRs is that they must rapidly fill in spectrum holes without posing harmful interference to the PUs. This task is done by the function of spectrum sensing, which is one of the critical techniques in a CR system. However, there is a significant challenge in sensing the whole of the spectrum at a particular physical location in a short observation time. (Performance degrades with longer observation times since the lagging response implies low spectrum utilisation efficiency). Thus, wideband spectrum sensing is of prime importance to ensure efficient operation of both the primary and the secondary (CR) networks.

B. Wideband spectrum sensing

In order to exploit wideband spectrum, a bank of narrowband filters is often utilised to reduce the wideband spectrum sensing problem to a multiple narrowband one [5]–[7]. Nevertheless, the implementation of a bank of narrowband filters requires a large number of RF components for sensing a wideband spectrum [8]. Besides, it is not flexible to use a bank of narrowband filters because the range of filters, and the number of the narrow bands are always preset. A wavelet-based sensing approach [9] provides advantages of both low implementation cost and flexibility in adapting to a dynamic wideband spectrum. However, characterising the wide bandwidth will require a high sampling rate analog-to-digital converter (ADC), due to the Nyquist sampling theorem, and the financial cost of that ADC will be prohibitive. Hence, it motivates the research of using undersampling techniques for the wideband spectrum sensing.

Some of the undersampling techniques are related to multicoset sampling [10]. In a multicoset scheme, M out of L ($M \leq L$) cosets of samples are chosen to reconstruct the signal. It can be implemented by using multiple sampling channels, which are offset by an integral multiple of a constant time. In order to reconstruct the signal with a high success rate, the number of sampling channels must be sufficiently high [11]. Compressed sensing (CS) is an emerging sampling theory that allows undersampling the signal. Tian and Giannakis [8] were the first to introduce CS to CR using an edge detection based approach. It has been observed that this is very sensitive to low SNRs, which commonly exist in the wireless environment. An analog information converter (AIC) structure was proposed in [12] for compressing signals in the analog domain, and the AIC based compressed wideband spectrum sensing approach was studied in [13]. However, it has been identified that the AIC model will be easily influenced by design imperfections or model mismatches [14].

C. Distributed wideband spectrum sensing

On the other hand, due to the effect of multipath fading or shadowing, a single CR cannot distinguish between a deep fade and an idle band. In such scenarios, distributed spectrum sensing has been widely considered for combating fading or shadowing [15], [16]. A data fusion based approach is preferred for distributed spectrum sensing as it offers much better detection performance than the decision fusion approaches [17]. Nonetheless, using conventional distributed spectrum techniques, the transmission of raw data through a control channel is very expensive for the wideband spectrum sensing. Moreover, both the high sampling rates and the raw data transmission are problematic when the distributed CRs are battery powered.

A candidate model for distributed wideband spectrum sensing is a multirate sampling system. Asynchronous multirate sampling (MRS) and synchronous multirate sampling (SMRS) schemes were studied for reconstructing sparse multiband signals in [18] and [19], respectively. In addition, MRS has been successfully implemented in experiments using a novel electro-optical system with three sampling channels as described in [20]. Both systems employ three optical pulsed sources that operate at different rates and at different wavelengths. In order to

reconstruct a wideband signal with an 18 GHz bandwidth, the modulated pulses are amplified, and sampled by an ADC at a rate of 4 GHz in each channel. In [18], the sampling channels of the MRS can be implemented separately without synchronisation. However, reconstruction of the spectrum requires that each frequency of the signal must be non-aliased in at least one of the sampling channels. In [19] the SMRS reconstructs the spectrum from linear equations, which relate the Fourier transform of the signal to the Fourier transform of its samples. Using CS theory, sufficient conditions for perfectly reconstructing the spectrum are obtained; $v \geq 2k$ (the Fourier transform of the signal is k-sparse) sampling channels are required. It is difficult to implement such sufficient conditions either in a single CR node or distributed CR nodes, because of its stringent requirements on electro-optical hardware and synchronisation. For reconstructing the spectrum using fewer sampling channels, the spectrum to be recovered should possess certain properties, e.g., minimal bands, and uniqueness. However, the spectral components from PUs may not possess those properties.

D. Contributions

The contribution in this paper is threefold. First, we propose a multirate asynchronous sub-Nyquist sampling (MASS) system that employs multiple lowrate ADCs to implement wideband spectrum sensing. The key features of the MASS system are, 1) low implementation complexity, 2) applicability to distributed CR networks, 3) energy-efficiency for sharing spectrum sensing data, and 4) robustness against the lack of time synchronisation. Note that even though the sampling pattern of MASS can also be implemented using a multicoset scheme, it is not practical to use this technique because of its stringent synchronisation requirement. In order to guarantee a perfect spectrum recovery, the total number of observations from all ADCs is of the order of $2k\sqrt{N}$, instead of N (the number of samples required when sampling at the Nyquist rate). We should emphasise that even though we use more samples than the ideal CS-based method, which only requires $\mathrm{O}(k\log(N))$ samples, our system is more readily implemented in a distributed CR network than the CS-based model. We then present the conditions under which the recovery of the full spectrum is unique by using CS techniques. A trade-off is made between the number

of sampling channels and the probability of successful spectrum recovery. Finally we apply our model to distributed CR networks. When the spectra of the CR nodes have a common spectral support, using one lowrate ADC in each CR node can also successfully recover the full spectrum. This is obtained by applying a hybrid matching pursuit (HMP) algorithm by synthesizing the distributed compressed sensing simultaneous orthogonal matching pursuit (DCS-SOMP) and compressive sampling matching pursuit (CoSaMP).

E. Outline and Notations

In the following, we briefly introduce CS theory in Section II. In Section III we propose a MASS model. Using CS theory, the sufficient conditions of full spectrum recovery are derived. Simulation results are presented in Section IV, followed by conclusions in Section V. We note that, throughout this paper, we adopt the following notations. Scalars are denoted by italic font lowercase letters and constants by uppercase italic letters. Vectors and matrices are denoted by lowercase boldface and uppercase boldface letters, respectively.

II. Related Background on Compressed Sensing

CS theory [21] indicates that a signal, $\mathbf{x} \in \mathbb{C}^N$, which is k-sparse, can be exactly recovered from M $(M \ll N)$ linear projections $\mathbf{y} \in \mathbb{C}^M (\mathbf{y} = \mathbf{\Phi x})$, where $\mathbf{\Phi} \in \mathbb{C}^{M \times N}$ is the measurement matrix. By k-sparse, we mean that the k largest values of $\mathbf{x}$ are not negligible. The performance of recovery is determined by three factors, namely, the sparsity, k, of the signal $\mathbf{x}$, the properties of the measurement matrix $\mathbf{\Phi}$, and the recovery algorithm. The mutual coherence is a computationally tractable metric for evaluating the suitability of the chosen measurement matrix.

Definition 1 [22]: Let $\mathbf{\Phi}$ be expressed as $\mathbf{\Phi} = [\phi_\mathbf{1}\ \phi_\mathbf{2}\ ...\ \phi_\mathbf{N}]$, where $\phi_\mathbf{j}$ denotes the j-th column of the matrix $\mathbf{\Phi}$. Then the mutual coherence, μ, of the matrix $\mathbf{\Phi}$ is given by,

$$\mu = \max_{j \neq h \in [1,N]} | < \hat{\phi}_\mathbf{j}, \hat{\phi}_\mathbf{h} > |, \tag{1}$$

where $\hat{\phi}_\mathbf{j} = \frac{\phi_\mathbf{j}}{\|\phi_\mathbf{j}\|_2}$ denotes the ℓ_2 normalised column.

The aim is to keep μ to a minimum to allow linear projections to be inverted in a stable manner. Donoho and Elad have proved that a small mutual coherence μ can guarantee the performance

of the recovery as below.

Theorem 1 [23], [24]: Assume that a signal $\mathbf{x}$ is k-sparse, when the mutual coherence μ of the measurement matrix $\mathbf{\Phi}$ satisfies,

$$\mu < \frac{1}{2k-1}, \tag{2}$$

we can use techniques such as basis pursuit (BP) [25] or orthogonal matching pursuit (OMP) [26] to find the sparsest solution of the k-sparse signal $\mathbf{x} \in \mathbb{C}^N$ from measurements $\mathbf{y} \in \mathbb{C}^M$.

In a distributed network, the received signals in different sensors are not only individually sparse, but also jointly sparse (have nonzero entries at the same locations). A separate recovery strategy (each sensor recovers the signal individually) will require more measurements as it neglects the correlations of signal among sensors. The DCS-SOMP algorithm was presented in [27] for reconstructing the joint sparse signals with fewest measurements by a boost-and-recover approach.

III. Multirate Asynchronous Sub-Nyquist Sampling

We will now present a MASS system to sense the wideband spectrum using multiple lowrate ADCs in subsection III-A. For simplicity, we firstly consider the case that one CR node is equipped with parallel ADCs, which uniformly sample the wideband signal at different sub-Nyquist rates in the noiseless case. In subsection III-B, we will concentrate on exploring what kind of linear projection exists when performing sub-Nyquist sampling. The effect of sub-Nyquist sampling will then be considered in subsection III-C. Using CS theory, we will give sufficient conditions for reconstructing the full spectrum in subsection III-D. Finally the trade-off between the number of ADCs and the probability of successful spectrum recovery will be discussed in subsection III-E. In subsection III-F, this will be extended to apply in a distributed CR network.

A. System design

Partially motivated by MRS in [18], our system will use a multirate sampling scheme as shown in Fig. 1. Instead of electro-optical devices, lowrate ADCs are employed in MASS. Since the average spectral occupancy is very low, we assume that the non-aliased discrete Fourier transform

(DFT) spectrum (full spectrum, when the sampling rate is not less than the Nyquist sampling rate) is k-sparse, and consists of several subband signals with different unknown bandwidths. By k-sparse, we mean that only k components in the full spectrum are non-negligible. The sparsity level, k, of the non-aliased DFT spectrum can be obtained from initialisation, for example, coarse spectrum scanning, and will not be addressed here. The wideband filter prior to the ADCs removes only frequencies outside the spectrum of interest, and is altered to have the largest spectral estimation bandwidth, W. We assume that there are v ADCs that sample the wideband signal $x(t)$ at different rates, $\frac{M_1}{T}, \frac{M_2}{T}, ..., \frac{M_v}{T}$, over the same observation time of T. We note that no anti-aliasing filter is used prior to the ADCs, thus aliasing occurs. The length of samples in their corresponding channels will be $M_1, M_2, ..., M_v$ ($M_{i\in[1,v]} \sim \mathrm{O}(\sqrt{N})$ in (17)). A tapered window, such as the Hamming window, is used to combat the effect of leakage, and keep the sparsity level, k, of the non-aliased DFT spectrum as small as possible. The spectral observations are obtained by performing a windowed fast Fourier transform (FFT) of the samples in each channel. The magnitude vectors of the sub-Nyquist rate spectra, $\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_v$ ($\mathbf{y}_i \in \mathbb{R}^{M_i}$, $i \in [1, v]$), are used to form a concatenated equation as shown in (14). After that, the measurement matrix $\mathbf{\Phi}$ is constructed by only using M_i ($\forall\ i \in [1, v]$) and N in (10) and (14). Then we recover the non-aliased spectrum $|\mathbf{x_n}|$ using a CS algorithm, e.g., BP, OMP, followed by spectrum detection on the reconstructed spectrum $|\mathbf{x_n}|$.

MASS has several advantages for application in CR networks, including,

1) Wideband spectrum sensing is implemented with sub-Nyquist sampling, which relaxes the stringent requirements on ADCs.
2) The lowrate ADCs behave as acquisition devices as well as spectrum compression devices.
3) The compression/measurement matrix we used is deterministic, and can be easily constructed once we know $M_{i\in[1,v]}$ and N.

B. Relating the sub-Nyquist rate DFT spectrum to the non-aliased DFT spectrum

We can view the sub-Nyquist rate spectrum (aliased spectrum) as a linear projection from the non-aliased spectrum as shown in Fig. 2.

Let $x(t)$ represent the output signal of the wideband filter in Fig. 1, with a bandwidth of W. In the short observation time of T, we sample the continuous signal $x(t)$ at a sampling rate of $f_{s'} = \frac{M_i}{T} < 2W$ in the i-th channel (different sampling rates at different channels). After a tapered window, the sampled signal in the i-th channel, $x_{si}(t)$, can be represented by,

$$x_{si}(t) = \sum_{l=-\infty}^{\infty} x(t)\delta(t - l\Delta t_i)w_T(t), \tag{3}$$

where $w_T(t)$ is a tapered window, Δt_i is the sampling interval in the sampling channel i, and $\delta(t)$ is a Dirac delta function.

The Fourier transform of the sampled signal is given by,

$$X_{si}(f) = f_{s'} \sum_{l=-\infty}^{\infty} \int_{-\infty}^{\infty} X(\tau + lf_{s'})W_T(f - \tau)d\tau. \tag{4}$$

where $W_T(f)$ denotes the Fourier transform of the tapered window, $f_{s'} = 1/\Delta t$, and $X(f)$ denotes the Fourier transform of $x(t)$.

If $x(t)$ is sampled at a rate of $f_s = \frac{N}{T} > 2W$ over an observation time of T, the Fourier transform of the sampled signal can be represented by,

$$X_n(f) = f_s \int_{-\infty}^{\infty} X(\tau)W_T(f - \tau)d\tau, \quad \forall |f| < f_s/2, \tag{5}$$

where $X_n(f)$ is a non-aliased full spectrum. Replacing f by $f + lf_{s'}$, and τ by $\tau + lf_{s'}$ in (5), and then substituting it into (4), we obtain the relation between $X_{si}(f)$ and $X_n(f)$ as,

$$X_{si}(f) = \frac{f_{s'}}{f_s} \sum_{l=-\infty}^{\infty} X_n(f + lf_{s'}), \quad \forall |f + lf_{s'}| < f_s/2, \tag{6}$$

where $X_{si}(f)$ denotes the sub-Nyquist rate spectrum in the i-th sampling channel.

Since the observation time T in both cases are the same, the same frequency resolution applies to these two cases, i.e., $\Delta f = \frac{f_s}{N} = \frac{f_{s'}}{M_i}$, where N, and M_i are integers and denote the number of samples at sampling rates f_s, and $f_{s'}$, respectively. By defining an integer m, and a scalar $\xi \in [0, \Delta f)$, such that $f = m\Delta f + \xi$, we can rewrite (6) as,

$$X_{si}(m\Delta f+\xi) = \frac{M_i}{N} \sum_{l=-\infty}^{\infty} X_n((m+lM_i)\Delta f+\xi) = \frac{M_i}{N} \sum_{n=-\lfloor N/2 \rfloor}^{\lfloor N/2 \rfloor} X_n\left(n\Delta f+\xi\right) \sum_{l=-\infty}^{\infty} \delta[n-(m+lM_i)], \tag{7}$$

where $\lfloor a \rfloor$ denotes the floor function, and gives the largest integer not greater than a.

Sampling the Fourier transform spectrum in (7) at rate of $\frac{1}{\Delta f}$, we obtain the DFT spectrum,

$$X_{si}[m] = \frac{M_i}{N} \sum_{l=-\infty}^{\infty} X_n[m + lM_i] = \frac{M_i}{N} \sum_{n=-\lfloor N/2 \rfloor}^{\lfloor N/2 \rfloor} X_n[n] \sum_{l=-\infty}^{\infty} \delta[n - (m + lM_i)], \tag{8}$$

where $X_{si}[m]$ denotes the sub-Nyquist rate DFT spectrum in the i-th channel, and $X_n[n]$ denotes the non-aliased DFT spectrum. In matrix form this becomes,

$$\mathbf{x_{si}} = \mathbf{\Phi_i x_n}, \tag{9}$$

where $\mathbf{x_{si}}$ denotes the sub-Nyquist rate DFT spectrum vector in the i-th channel, $\mathbf{x_n}$ is the non-aliased DFT spectrum vector, and the element of the linear projection operator $\mathbf{\Phi_i} \in \mathbb{R}^{M_i \times N}$ ($M_i < N$) can be represented as,

$$\phi_{m+\lfloor M_i/2 \rfloor + 1, n+\lfloor N/2 \rfloor + 1} = \frac{M_i}{N} \sum_{l=-\infty}^{\infty} \delta\big[n - (m + lM_i)\big], \tag{10}$$

where $m \in \mathbb{Z} \cap [-M_i/2, M_i/2)$, and $n \in \mathbb{Z} \cap [-N/2, N/2)$.

It is easy to see that in each column of $\mathbf{\Phi_i}$, there is only one non-zero element with value of $\frac{M_i}{N}$. And in each row of $\mathbf{\Phi_i}$, there exists at most $\lceil \frac{N}{M_i} \rceil$ (ceil function gives the smallest integer not less than $\frac{N}{M_i}$) non-zero elements, which is also called the undersampling factor.

C. Effect of sub-Nyquist sampling

Sampling a signal at a sub-Nyquist sampling rate generates two issues. First, the exact location of the signals for those who have frequencies larger than the sub-Nyquist sampling rate is lost. Second, there is a risk of overlap, i.e., different frequencies are down-converted to the same frequency in the baseband. This is crucial, because it could lead to signal cancellation, and hence missed signal detection in the CR application. However, under certain assumptions, signal overlap has a very small probability of occurring. They are 1) the non-aliased DFT spectrum is k-sparse, 2) the number of subsamples in the i-th channel, $M_i \sim \mathrm{O}(\sqrt{N})$ (in (17), for simplicity, we use $M_i = \sqrt{N}$), and 3) $k \ll N$.

Assuming those k spectral components are independent and identically distributed (i.i.d.) over the frequency bins of $0, 1, \cdots, N-1$, the probability of one element in the non-aliased DFT spectrum being non-zero is $P = Pr(X_n[n] \neq 0) = k/N$. If let $q[m]$ denote the number of

signals overlapped at $X_{si}[m]$, using (8) the probability of no signal overlap is given by,

$$Pr(q[m]<2) = Pr(q[m]=0)+Pr(q[m]=1) = (1-P)^{\lceil \frac{N}{M_i} \rceil} + \binom{\lceil \frac{N}{M_i} \rceil}{1} P(1-P)^{\lceil \frac{N}{M_i} \rceil - 1}$$
$$\approx (\frac{N-k}{N})^{\frac{N}{M_i}} + \frac{k}{M_i}(\frac{N-k}{N})^{\frac{N-M_i}{M_i}} = \frac{(\frac{N-k}{N})^{\sqrt{N}}(N-k+k\sqrt{N})}{N-k}. \quad (11)$$

It is not difficult to test that, for any fixed k, (11) will converge to one when N increases to infinity. As shown in Fig. 3, when $N = 9 \times 10^6$, and $M_i = \sqrt{N}$, the probability of no signal overlap at $X_{si}[m]$ will be $Pr(q[m] < 2) = 95 \sim 100\%$ for any $k \leq 1000 = M_i/3$, and $Pr(q[m] < 2) \geq 99\%$ for any $k \leq 400$. Thus, under assumptions mentioned above, using (9) we can write,

$$\mathbf{y_i} = |\mathbf{x_{si}}| = |\mathbf{\Phi_i x_n}| \simeq \mathbf{\Phi_i}|\mathbf{x_n}|. \quad (12)$$

Equation (12) holds true because when $X_{si}[m]$ is the projection of a single signal, we have,

$$|X_{si}[m]| = \left| \frac{M_i}{N} \sum_l X_n[m + lM_i] \right| = \frac{M_i}{N} |X_n[m + lM_i]| = \mathbf{\Phi_i} |X_n[m + lM_i]|. \quad (13)$$

In the rare case where signal overlap occurs, i.e. $0 \leq |\mathbf{x_{si}}| \leq \mathbf{\Phi_i}|\mathbf{x_n}|$, we could still track the spectral content of $|\mathbf{x_n}|$ with the aid of other sampling channels [18]. This is because the blind spot of one ADC can be illuminated by other ADCs working at different sampling rates.

D. Recovery of the full spectrum via multirate sampling

We will now introduce a method for reconstructing the full spectrum through multirate sampling. Since we are observing the same magnitude of the spectrum, $|\mathbf{x_n}|$, in all channels, we can form a concatenated equation relating $\mathbf{x} = |\mathbf{x_n}| \in \mathbb{R}^N$ to $\mathbf{y} = |\mathbf{x_s}| \in \mathbb{R}^M$ as below,

$$\mathbf{y} = \begin{pmatrix} \mathbf{y_1} \\ \mathbf{y_2} \\ \vdots \\ \mathbf{y_v} \end{pmatrix} = \begin{pmatrix} |\mathbf{x_{s1}}| \\ |\mathbf{x_{s2}}| \\ \vdots \\ |\mathbf{x_{sv}}| \end{pmatrix} \simeq \mathbf{\Phi x} = \begin{pmatrix} \mathbf{\Phi_1} \\ \mathbf{\Phi_2} \\ \vdots \\ \mathbf{\Phi_v} \end{pmatrix} |\mathbf{x_n}| \quad (14)$$

where $\mathbf{\Phi_1}, \mathbf{\Phi_2}, ..., \mathbf{\Phi_v}$ are disjoint submatrices of $\mathbf{\Phi}$, $\mathbf{\Phi_i} \in \mathbb{R}^{M_i \times N}$ is the measurement matrix of ADC i, and $M = \sum_{i=1}^{v} M_i$. Since the time offset between ADCs will not influence the magnitude of the non-aliased spectrum, $|\mathbf{x_n}|$, (14) holds true for asynchronous ADCs.

When certain conditions are satisfied as Lemma 1 below, the mutual coherence of the concatenated measurement matrix $\mathbf{\Phi}$ will be determined by the number of samples in each channel, and the number of channels v.

Lemma 1 [28]: When v ADCs observe the spectrum, $|\mathbf{x_n}|$, in the same observation time, generating v measurement vectors, $\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_v$, whose length are different primes, $M_1, M_2, \cdots, M_v$, which satisfy,

$$M_l M_z > N, \quad \forall l, z \in [1, v], l \neq z, \tag{15}$$

then the mutual coherence μ of the measurement matrix $\mathbf{\Phi}$ is determined by,

$$\mu = \max_{j \neq h} | < \hat{\phi}_{\mathbf{j}}, \hat{\phi}_{\mathbf{h}} > | = \frac{\max_{i \in [1,v]} M_i^2}{\sum_{l=1}^{v} M_l^2}. \tag{16}$$

The proof of Lemma 1 is given in Appendix A.

Theorem 2 (Sufficient conditions): For the above system, if,

$$M_l M_z > N, \quad \forall l, z \in [1, v], l \neq z, \tag{17}$$

$$\sum_{l=1}^{v} \left(\frac{M_l}{\max_{i \in [1,v]} M_i} \right)^2 > 2k - 1, \tag{18}$$

we can reconstruct the full magnitude spectrum $|\mathbf{x_n}|$ without any aliasing.

Proof: Theorem 2 follows from the results of the Lemma 1 and Theorem 1.

From (17), we can find that the number of samples in each ADC is of the order of $\sqrt{N}$. When $M_i \sim \mathrm{O}(\sqrt{N})$, from (18), we can see that at least $2k$ channels are required to guarantee the recovery of the non-aliased spectrum $|\mathbf{x_n}|$. This is because $v \geq \sum_{l=1}^{v} \left(\frac{M_l}{\max_{i \in [1,v]} M_i} \right)^2$ holds true for all $v \in \mathbb{Z}^+$. Considering both (17) and (18), we note that MASS needs the total number of observations to be $\sum_{i=1}^{v} M_i \sim \mathrm{O}(2k\sqrt{N})$. If we define the compression ratio as $\frac{\sum_{i=1}^{v} M_i}{N}$, we can see from Table I that the compression ratio decreases when the length of Nyquist samples N increases, but it increases if the sparsity degree k increases. For example, for $v = 4$, the compression ratio descends from 16.4% to 1.33% when N increases from 1024 to 102400.

E. Trade-off number of sampling channels with recovery probability

In order to reduce the requirement for a large number of ADCs, the strict recovery criteria can be relaxed.

Theorem 3: If we sample a wideband signal $x(t)$, whose full spectrum is k-sparse, and obtain multirate spectral observations, $\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_v$, whose length are prime numbers of $M_1, M_2, ..., M_v$ which satisfy,

$$M_l M_z > N, \quad \forall l, z \in [1, v], l \neq z, \tag{19}$$

$$\frac{\sum_{l=1}^{v} M_l^2}{\sum_{l=1}^{v} M_l} > \frac{2k-1}{\epsilon}, \quad \forall v \in [2, \infty), \tag{20}$$

recovery algorithms such as BP, and OMP, have a probability of at least $(1-\epsilon)$ to reconstruct the non-aliased magnitude spectrum $|\mathbf{x_n}|$.

The proof of Theorem 3 is presented in Appendix B.

Compared with Theorem 2, we can see that when the condition $\max_{l \in [1,v]} M_l^2 > \frac{\sum_{l=1}^{v} M_l}{\epsilon}$ is satisfied, the recovery conditions can be relaxed. It is quite useful when we want to use fewer sampling channels to approximate the full spectrum with a probability of at least $(1-\epsilon)$.

F. Extension to distributed wideband spectrum sensing

In wireless communication scenarios, some CR users may suffer the effect of fading either due to multipath propagation, or due to shadowing. In that case, the result of spectrum sensing from a single CR node is not reliable. A cooperative detection strategy offers a good solution as it minimises the effect of severe fading and achieves diversity gain [16]. If every CR forwards its measured or processed data to a fusion center (FC), which makes a final decision based on collected data, the cooperative scheme is often called data fusion based distributed spectrum sensing. If each CR uses multiple ADCs to do sub-Nyquist sampling, the transmission of the measurements may be very expensive in distributed CR networks.

To minimise the required transmission bandwidth, we propose to extend the application of the MASS model from the case of parallel ADCs in one CR node to serial ADCs in multiple CR nodes as shown in Fig. 4. Note that, as in the analysis of subsection III-D, the time offset between ADCs will not change the magnitude of the non-aliased spectrum, thus asynchronous CRs have the same performance as synchronous CRs. The wideband filters prior to the ADCs are altered to have the same bandwidth of W. After measurements in each CRs, the magnitudes of the sub-Nyquist rate spectra are transmitted to the FC, where the full spectrum is reconstructed.

The advantage of the distributed MASS model is that only one lowrate ADC is required in each CR node, which not only simplifies the system structure at each CR node, but also decreases the bandwidth required for sharing spectrum sensing data. As the analysis in the subsection III-D, the total number of observations to be transmitted is of the order of $2k\sqrt{N}$. Even though this is more than conventional CS, i.e., $M \sim \mathrm{O}(k\log(N))$, MASS is more amenable to implementation in a distributed CR network as neither a compression device nor a measurement matrix generator are required. Moreover, the measurement matrix we used is deterministic, then the transmission and storage of the measurement matrix is unnecessary. Nevertheless, the disadvantage is that (14) no longer holds because of the influence of fading. Thus, conventional CS algorithms, such as BP and OMP, cannot be applied. In fading channels, the power of the signals coming from PUs are attenuated. Namely, the received signals at distributed CR nodes, i.e., $u^{(1)}(t), u^{(2)}(t), \cdots u^{(v)}(t)$, may be different, and the spectra viewed by distributed CRs therefore are often distinct. As illustrated in Fig. 5, the spectra over the distributed CR nodes can be modelled by the second joint sparsity model (JSM-2) in [27]. Specifically, the spectra over CRs have a common spectral support, Ω (the set of frequencies occupied), but with different amplitudes.

In order to exploit the joint sparse property over sensor nodes, DCS-SOMP was proposed in [27]. It has been observed that this algorithm requires fewer measurements when compared with a separate recovery approach. However, the drawback of DCS-SOMP is its calculation time as it only selects the maximum support in each iteration. Unlike the conventional greedy algorithms, CoSaMP [29] accelerates the calculation by identifying many possible solutions in each iteration. Hence, we propose to apply a hybrid matching pursuit algorithm to the MASS model as shown in Table II, by synthesizing DCS-SOMP and CoSaMP. In each iteration, we boost the common support by summing up the correlation vectors from multiple channels, which will make the features easy to identify even if fading exists in some channels. We select multiple indices in each iteration by choosing the top-$2k$ indices of the $2k$-largest values in the combined correlation vector, and merge with the previously computed support. After that, the non-aliased full spectrum is recovered by least squares, where $\mathbf{\Phi}_{\mathbf{\Omega}}^{\dagger}$ denotes the pseudoinverse of $\mathbf{\Phi}_{\mathbf{\Omega}}$, and $\mathbf{\Phi}_{\mathbf{\Omega}}$ is the sub-matrix obtained by only selecting the columns, whose indices are in Ω, in the matrix

$\mathbf{\Phi}$. Since the matrix $\mathbf{\Phi_\Omega}$ is always well conditioned, the calculation of the pseudoinverse $\mathbf{\Phi_\Omega^\dagger}$ can be implemented quickly using an iterative method, such as the conjugate gradient method in [30]. We then prune the support to have k-largest values, followed by updating the residuals. The algorithm will be halted when the ℓ_2 norm of the residual is not larger than the noise tolerance level ϱ.

IV. Simulation results

In the experiments, we consider that each CR is equipped with a single lowrate ADC, and there are v CR nodes allocated in the same cluster. In the i-th CR node, the wideband signal $x^{(i)}(t)$, which is defined below, is sampled by a lowrate ADC over an observation time, T.

$$x^{(i)}(t) = \sum_{l=1}^{N_b} \sqrt{E_l^i B_l} \cdot \text{sinc}(B_l(t - \Delta_l)) \cdot \cos(2\pi f_l(t - \Delta_l)), \tag{21}$$

where Δ_l denotes the time offset of the signal, set to be $T/2$, and E_l^i denotes the energy of the l-th subband viewed by the i-th CR. Since the effect of fading, E_l^i varies subject to the property of the fading channel. The wideband signal $x^{(i)}(t)$ consists of N_b subbands, whose bandwidth is B_l, and carrier frequency is f_l. The values of simulation parameters are specified in Table III. As described in Fig. 4, after FFT analysis in each node, the sub-Nyquist rate spectral observations $\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_v$, are gathered at the FC. In the FC, the full spectrum is reconstructed using HMP. The spectral occupancy status is decided based on a hypothesis test on the reconstructed spectrum.

In Fig. 6, we demonstrate the performance of spectrum recovery in MASS using HMP. We can see that by using $v = 10$ ADCs, the non-aliased spectrum, which consists of 5 subbands with bandwidth $0.3 \sim 30$ MHz, can be successfully reconstructed. We should emphasize that instead of working at or above 10 GHz, these ADCs are working at sampling rate of $182.5 \sim 282.5$ MHz, and the total number of measurements is $23.85\%N$ (N is the number of samples if $x^{(i)}(t)$ is Nyquist sampled in a CR). In Fig. 7, we find that the fewer subbands there are, the better the detection performance we can achieve. If a CR system has constraint on the probability of false alarm, e.g., $P_{fa} \leq 10\%$, the minimum number of measurements are $0.1N$, $0.3N$, and $0.55N$ for $N_b = 10$, $N_b = 30$, and $N_b = 50$, respectively. On the other hand, if one lays emphasis

on the probability of detection, we can find that to obtain $P_d \geq 90\%$, we need at least $0.2N$ measurements to sense the wideband spectrum with up to 50 subbands. In addition, one can notice that the higher compression ratio will result in a smaller probability of false alarm and a larger probability of detection.

In Fig. 8, we consider the effect of imperfect synchronisation among ADCs. Compared with a reference clock, the asynchronous ADCs have time offset in range of $0 \sim 0.8$ μs, while the total observation time is 2 μs. We can see that the detection performance of the asynchronous ADCs is roughly the same as that of the synchronous ADCs. Fig. 8 also illustrates that with more ADCs in collaboration, we can achieve a better spectrum sensing performance. This is because with more sampling channels, we can obtain higher probability of successful spectrum recovery as the discussion in the subsection III-E.

In order to quantify the detection performance of the distributed MASS system over fast/slow fading channels, we compare their detection performance with that over AWGN channels in Fig. 9. It shows that, compared with AWGN case, more CRs need to collaborate in order to combat the effect of fading. Specifically, to obtain $P_{fa} = 10\%$, 40 CRs are required to collaborate over AWGN channels, 50 CRs are required for Rayleigh fading channels, and 60 CRs are necessary for slow fading channels. Note that 60 ADCs in MASS can be converted to the compression ratio of 42.32%, which means that we only need to transmit $0.4232N$ measurements to the FC through the control channel, rather than $60N$ for a conventional distributed spectrum sensing approach.

V. Conclusions

In this paper, we have presented a distributed wideband spectrum sensing model, MASS, which employs several lowrate ADCs to sample the wideband signal at sub-Nyquist rates. The MASS system could be applied either in the scenario of multiple ADCs in one sensor or in the case of serial ADCs in multiple sensors. We have presented the sufficient conditions to uniquely recover the full spectrum by using CS theory. When the spectra of the CR nodes are jointly sparse, we have applied the MASS to the distributed CR network. Using one lowrate ADC in

each CR node, we can still recover the full spectrum by using the HMP algorithm.

Simulation results have shown that our approach can recover the wideband non-aliased spectrum successfully. Besides, MASS has been shown to be very robust against the lack of time synchronisation between ADCs. Moreover, when the MASS model is applied to the distributed CR network, we only need a few more CRs to collaborate for combating the effect of fading. To further save the transmission costs in transmitting local information to the FC, a multi-hop implementation of the MASS is attractive. We will investigate the performance of the MASS using multi-hop scheme in the near future.

APPENDIX A
PROOF OF LEMMA 1

We use an example, $M_1 = 3$, $M_2 = 5$, and $N = 9$. The measurement matrix $\mathbf{\Phi}$ can be constructed using (10) and (14), then the ℓ_2 column-wise normalised matrix $\hat{\mathbf{\Phi}}$ is

$$\hat{\mathbf{\Phi}} = \begin{bmatrix} 0.5145 & 0 & 0 & 0.5145 & 0 & 0 & 0.5145 & 0 & 0 \\ 0 & 0.5145 & 0 & 0 & 0.5145 & 0 & 0 & 0.5145 & 0 \\ 0 & 0 & 0.5145 & 0 & 0 & 0.5145 & 0 & 0 & 0.5145 \\ 0 & 0 & 0.8575 & 0 & 0 & 0 & 0 & 0.8575 & 0 \\ 0 & 0 & 0 & 0.8575 & 0 & 0 & 0 & 0 & 0.8575 \\ 0 & 0 & 0 & 0 & 0.8575 & 0 & 0 & 0 & 0 \\ 0.8575 & 0 & 0 & 0 & 0 & 0.8575 & 0 & 0 & 0 \\ 0 & 0.8575 & 0 & 0 & 0 & 0 & 0.8575 & 0 & 0 \end{bmatrix}.$$

Using (10), the j-th normalised column $\hat{\phi}_j[m]$ can be expressed as

$$\hat{\phi}_j[m] = \frac{M_l}{\sqrt{\sum_{k=1}^{v} M_k^2}} \delta\Big[m - \Big|j - \Big\lfloor \frac{N}{2} \Big\rfloor + \Big\lfloor \frac{M_l}{2} \Big\rfloor \Big|_{\text{mod}^\star(M_l)}\Big],$$
$$\forall j \in [1, N],\ l = \left\{ l | m \in \left[\sum_{k=0}^{l-1} M_k + 1, \sum_{k=0}^{l} M_k \right] \right\} \tag{22}$$

where $M_0 = 0$ and $|\bullet|_{\text{mod}^\star(\oplus)}$ is a modified modulo operation in the sense of using $\oplus$ to replace all zeros in the results of the standard modulo operation.

Let us define the event m_l as there exists a value of $m \in [\sum_{k=0}^{l-1} M_k + 1, \sum_{k=0}^{l} M_k]$ such that $\hat{\phi}_j[m] = \hat{\phi}_h[m] > 0$, where $j \neq h \in [1, N]$. Then the average probability of the event m_l

occurring is $Pr(m_l) = \frac{\mathrm{C}_{M_l}^1 \mathrm{C}_{D_l}^2}{\mathrm{C}_N^2} = \frac{D_l - 1}{N-1} = \frac{N/M_l - 1}{N-1}$, where $D_l = N/M_l$ is the undersampling factor in the l-th sampling channel. Obviously, any single event m_l ($l \in [1, v]$) could occur when $M_l < N$ because $Pr(m_l) > 0$. In case of only one event m_i occurring, the mutual coherence of the matrix $\mathbf{\Phi}$ can be calculated by using (1) and (22) as

$$\mu = \max_{j \neq h} | < \hat{\phi}_{\mathbf{j}}, \hat{\phi}_{\mathbf{h}} > | = \max_{i \in [1,v]} \left| \frac{M_i}{\sqrt{\sum_{l=1}^{v} M_l^2}} \cdot \frac{M_i}{\sqrt{\sum_{l=1}^{v} M_l^2}} \right| = \frac{\max_{i \in [1,v]} M_i^2}{\sum_{l=1}^{v} M_l^2}. \tag{23}$$

Then next target is to make the probability of two events occurring to be zero. By the definition of the event m_l, it satisfies $|j - \lfloor \frac{N}{2} \rfloor + \lfloor \frac{M_l}{2} \rfloor|_{\mathrm{mod}^\star(M_l)} = |h - \lfloor \frac{N}{2} \rfloor + \lfloor \frac{M_l}{2} \rfloor|_{\mathrm{mod}^\star(M_l)}$, where $j \neq h \in [1, N]$. It is equivalent to $|h - j| = w_l M_l$ for $w_l \in \mathbb{Z} \cap [1, \lceil \frac{N}{M_l} \rceil - 1]$. Similarly, the event m_z happens when $|h - j| = w_z M_z$ for $w_z \in \mathbb{Z} \cap [1, \lceil \frac{N}{M_z} \rceil - 1]$. Both of them happen when

$$w_z M_z = w_l M_l, \quad w_l \in \mathbb{Z} \cap \left[1, \left\lceil \frac{N}{M_l} \right\rceil - 1\right], w_z \in \mathbb{Z} \cap \left[1, \left\lceil \frac{N}{M_z} \right\rceil - 1\right]. \tag{24}$$

The probability $Pr(m_z m_l)$ is the probability of both the event m_z and the event m_l satisfying equation (24). This is equivalent to being given M_l and M_z, and then finding how many pairs of numbers w_l and w_z out of the available number of $(\lceil \frac{N}{M_l} \rceil - 1)$ or $(\lceil \frac{N}{M_z} \rceil - 1)$ are required for equation (24) to hold. Without lost of generality, we consider the possibility for w_l and get

$$Pr(m_z m_l) = \frac{\left\lfloor \frac{\left\lceil \frac{N}{M_l} \right\rceil - 1}{w_l^{'}} \right\rfloor}{\left\lceil \frac{N}{M_l} \right\rceil - 1}, \tag{25}$$

because of w_l should be integer multiples of w_l' as $w_l = w_z \frac{M_z}{M_l} = w_l' \frac{w_z}{w_z'}$, where $w_l^{'}$ and $w_z^{'}$ are from simplest fraction form of

$$\frac{w_l}{w_z} = \frac{M_z}{M_l} = \frac{a w_l^{'}}{a w_z^{'}} = \frac{w_l^{'}}{w_z^{'}}, \tag{26}$$

where a denotes an integer.

If we assume that M_l and M_z are different prime numbers, then $w_l^{'} = M_z$. In order to let $Pr(m_z m_l) = 0$, which is equivalent to $M_z = w_l^{'} > \lceil \frac{N}{M_l} \rceil - 1$, M_l and M_z should satisfy

$$M_l M_z > N, \quad \forall l, z \in [1, v], l \neq z \tag{27}$$

because $M_z > \frac{N}{M_l} \geq \lceil \frac{N}{M_l} \rceil - 1$.

In summary, in order to prevent any two events (m_l and m_z, $l \neq z$) from happening together, M_l and M_z should at least satisfy equation (27). Furthermore, if the probability that m_l and m_z occur together is zero, then the probability that more than two events (m_l, m_z...) occur together is zero too. It is equivalent to the condition that the maximum correlation of different columns in $\mathbf{\Phi}$ exists when only one event happens as

$$\mu = \max_{j \neq h} | < \hat{\phi}_{\mathbf{j}}, \hat{\phi}_{\mathbf{h}} > | = \frac{\max_{i \in [1,v]} M_i^2}{\sum_{l=1}^{v} M_l^2}. \tag{28}$$

Hence, Lemma 1 is proved.

APPENDIX B
PROOF OF THEOREM 3

From Doob's maximal inequality [31], we obtain

$$Pr(\mu > x) = Pr\left(\max_{j \neq h} | < \hat{\phi}_{\mathbf{j}}, \hat{\phi}_{\mathbf{h}} > | > x\right) \leq \frac{\mathrm{E}(< \hat{\phi}_{\mathbf{j}}, \hat{\phi}_{\mathbf{h}} >)}{x}, \tag{29}$$

where $\mathrm{E}(a)$ denotes the expectation of a.

Assume that we obtain the prime number of samples and the condition of equation (19) is satisfied, there exists at most one value of $m \in [1, M]$ that makes $\hat{\phi}_j[m] = \hat{\phi}_h[m] > 0$. Then the expected value of $< \hat{\phi}_{\mathbf{j}}, \hat{\phi}_{\mathbf{h}} >$ becomes

$$\mathrm{E}(< \hat{\phi}_{\mathbf{j}}, \hat{\phi}_{\mathbf{h}} >) = \frac{M_1^2}{\sum_{l=1}^{v} M_l^2} Pr(m_1) + \frac{M_2^2}{\sum_{l=1}^{v} M_l^2} Pr(m_2) + \cdots + \frac{M_v^2}{\sum_{l=1}^{v} M_l^2} Pr(m_v). \tag{30}$$

Because $Pr(m_l) = \frac{N/M_l - 1}{N-1} < 1/M_l$, substituting it into (30) we obtain

$$\mathrm{E}(< \hat{\phi}_{\mathbf{j}}, \hat{\phi}_{\mathbf{h}} >) < \frac{\sum_{l=1}^{v} M_l}{\sum_{l=1}^{v} M_l^2}. \tag{31}$$

Replacing x in (29) by $x = \frac{1}{2k-1}$, and substituting (31) into (29) we obtain

$$Pr\left(\mu > \frac{1}{2k-1}\right) \leq (2k-1)\mathrm{E}(< \hat{\phi}_{\mathbf{j}}, \hat{\phi}_{\mathbf{h}} >) < \frac{\sum_{l=1}^{v} M_l}{\sum_{l=1}^{v} M_l^2}(2k-1). \tag{32}$$

If we define the final term to be less than ϵ, that is

$$\frac{\sum_{l=1}^{v} M_l^2}{\sum_{l=1}^{v} M_l} > \frac{2k-1}{\epsilon}, \tag{33}$$

the following equation holds true

$$Pr\left(\mu < \frac{1}{2k-1}\right) > 1 - \epsilon. \tag{34}$$

On applying Theorem 1, we finish the proof of Theorem 3.

ACKNOWLEDGEMENT

The authors would like to acknowledge the support from the Wolfson Microelectronics Scholarship and the Scottish Funding Council for the Joint Research Institute with Heriot-Watt University which is a part of the Edinburgh Research Partnership in Engineering and Mathematics.

REFERENCES

[1] FCC, "Spectrum Policy Task Force," 11/2002, ET Docket 02-135.

[2] M. A. McHenry, "NSF spectrum occupancy measurements project summary," Shared spectrum company, Tech. Rep., Aug. 2005.

[3] J. Mitola, "Cognitive radio: An integrated agent architecture for software defined radio," Ph.D. dissertation, Dept. of Teleinformatics, Royal Institute of Technology Stockholm, Sweden, 8 May, 2000.

[4] C.-X. Wang, X. Hong, H.-H. Chen, and J. Thompson, "On capacity of cognitive radio networks with average interference power constraints," *IEEE Trans. Wireless Communications*, vol. 8, no. 4, pp. 1620 –1625, April 2009.

[5] B. Farhang-Boroujeny and R. Kempter, "Multicarrier communication techniques for spectrum sensing and communication in cognitive radios," *IEEE Communications Magazine*, vol. 46, no. 4, pp. 80 –85, April 2008.

[6] B. Farhang-Boroujeny, "Filter bank spectrum sensing for cognitive radios," *IEEE Trans. Signal Processing*, vol. 56, no. 5, pp. 1801 –1811, May 2008.

[7] Y. Pei, Y.-C. Liang, K. Teh, and K. Li, "How much time is needed for wideband spectrum sensing?" *IEEE Trans. Wireless Communications*, vol. 8, no. 11, pp. 5466 –5471, Nov. 2009.

[8] Z. Tian and G. B. Giannakis, "Compressed sensing for wideband cognitive radios," in *Proc. IEEE ICASSP*, 2007, pp. 1357–1360.

[9] ——, "A wavelet approach to wideband spectrum sensing for cognitive radios," in *Proc. 1st CrownCom*, 8–10 June 2006, pp. 1–5.

[10] P. Feng and Y. Bresler, "Spectrum-blind minimum-rate sampling and reconstruction of multiband signals," in *Proc. IEEE ICASSP*, vol. 3, 1996, pp. 1688 –1691.

[11] M. Mishali and Y. C. Eldar, "Blind multiband signal reconstruction: Compressed sensing for analog signals," *IEEE Trans. Signal Processing*, vol. 57, no. 3, pp. 993–1009, March 2009.

[12] J. Tropp, J. Laska, M. Duarte, J. Romberg, and R. Baraniuk, "Beyond Nyquist: Efficient sampling of sparse bandlimited signals," *IEEE Trans. Information Theory*, vol. 56, no. 1, pp. 520 –544, Jan. 2010.

[13] Y. Wang, A. Pandharipande, Y. L. Polo, and G. Leus, "Distributed compressive wide-band spectrum sensing," in *Proc. Information Theory and Applications Workshop*, 8–13 Feb. 2009, pp. 178–183.

[14] 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, April 2010.

[15] P. Kaligineedi, M. Khabbazian, and V. Bhargava, "Secure cooperative sensing techniques for cognitive radio systems," in *Proc. IEEE ICC*, 2008, pp. 3406 –3410.

[16] C.-X. Wang, H.-H. Chen, X. Hong, and M. Guizani, "Cognitive radio network management: tuning in to real-time conditions," *IEEE Vehicular Technology Magazine*, vol. 3, no. 1, pp. 28–35, March 2008.

[17] E. Visotsky, S. Kuffner, and R. Peterson, "On collaborative detection of TV transmissions in support of dynamic spectrum sharing," in *Proc. IEEE DySPAN*, 2005, pp. 338 –345.

[18] A. Rosenthal, A. Linden, and M. Horowitz, "Multi-rate asynchronous sampling of sparse multi-band signals," *J. Opt. Soc. Am.*, vol. 25, no. 9, pp. 2320–2330, 2008.

[19] M. Fleyer, A. Linden, M. Horowitz, and A. Rosenthal, "Multirate synchronous sampling of sparse multiband signals," *IEEE Trans. Signal Processing*, vol. 58, no. 3, pp. 1144–1156, 2010.

[20] A. Feldster, Y. Shapira, M. Horowitz, A. Rosenthal, S. Zach, and L. Singer, "Optical under-sampling and reconstruction of several bandwidth-limited signals," *Journal of Lightwave Technology*, vol. 27, no. 8, pp. 1027 –1033, April 2009.

[21] D. Donoho, "Compressed sensing," *IEEE Trans. Information Theory*, vol. 52, no. 4, pp. 1289 – 1306, April 2006.

[22] D. L. Donoho and X. Huo, "Uncertainty principles and ideal atomic decomposition," *IEEE Trans. Information Theory*, vol. 47, no. 7, pp. 2845–2862, Nov. 2001.

[23] D. L. Donoho and M. Elad, "Optimally sparse representation in general (nonorthogonal) dictionaries via L1 minimization," in *Proc. Nat. Acad. Sci. USA*, vol. 100, no. 5, 2003.

[24] J. A. Tropp, "Greed is good: algorithmic results for sparse approximation," *IEEE Trans. Information Theory*, vol. 50, no. 10, pp. 2231–2242, Oct. 2004.

[25] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Review*, vol. 43, no. 1, pp. 129–159, 2001. [Online]. Available: http://www.jstor.org/stable/3649687

[26] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Information Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.

[27] M. F. Duarte, S. Sarvotham, D. Baron, M. B. Wakin, and R. G. Baraniuk, "Distributed compressed sensing of jointly sparse signals," in *39th ASILOMAR*, 2005, pp. 1537–1541.

[28] H. Sun, D. Laurenson, and J. Thompson, "Cooperative compressive spectrum sensing by sub-Nyquist sampling," in *Proc. UKIWCWS, IIT Delhi, India*, 2009.

[29] D. Needell and J. Tropp, "CoSamp: Iterative signal recovery from incomplete and inaccurate samples," *Applied and Computational Harmonic Analysis*, vol. 26, no. 3, pp. 301 – 321, 2009.

[30] Y. Saad, *Iterative methods for sparse linear systems*, 2nd ed. Society for Industrial Mathematics, 2003.

[31] J. Doob, *Stochastic processes*. Chapman & Hall, 1953.

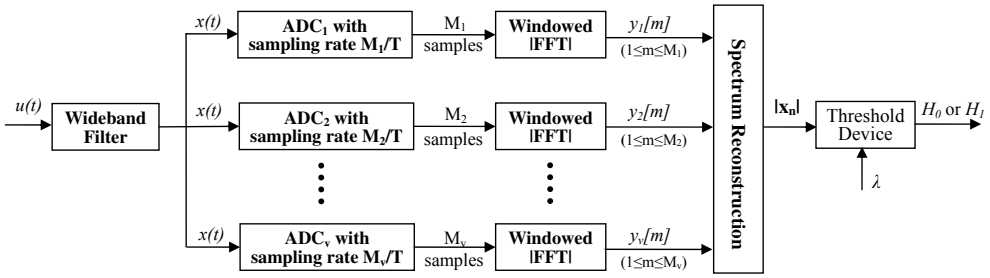

Fig. 1. The schematic illustration of the multirate asynchronous sub-Nyquist sampling system in one CR node. The wideband filter is altered to have a bandwidth of W.

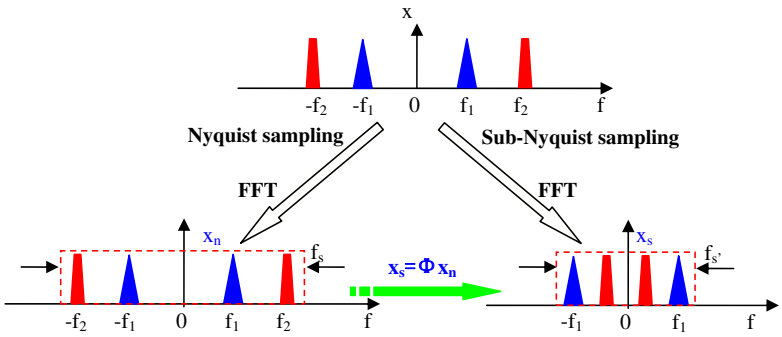

Fig. 2. Interpretation of the relationship between the non-aliased spectrum, $\mathbf{x_n}$, and the sub-Nyquist rate spectrum, $\mathbf{x_s}$, when the Fourier spectrum of $x(t)$ is denoted by $\mathbf{x}$.

TABLE I
SETS OF PRIMES SATISFYING CONDITIONS IN THEOREM 2

Sparsity (k)	Samples (N)	Set of primes (M_l)	Channels (v)	Compression ratio ($\frac{\sum_{l=1}^{v} M_l}{N} \times 100\%$)
1	9	3,5	2	88.9%
1	64	11,13	2	37.5%
1	1024	37,41	2	7.62%
2	1024	37,41,43,47	4	16.4%
1	10240	101,103	2	1.99%
2	10240	101,103,107,109	4	4.1%
3	10240	101, 103, 107, 109, 113, 127, 131	7	7.72%
1	102400	331,337	2	0.65%
2	102400	331,337,347,349	4	1.33%
3	102400	331,337, 347, 349, 353,359	6	2.02%

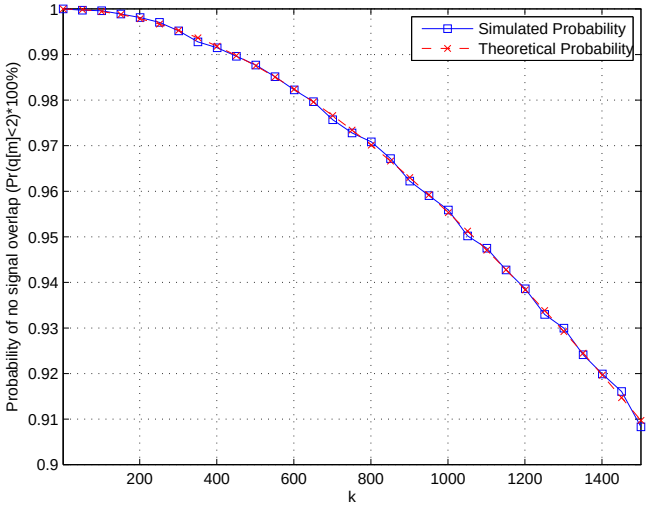

Fig. 3. Simulated probability of no signal overlap at a frequency $f = m\Delta f + \xi$ in the sub-Nyquist rate spectrum $\mathbf{x_s}$, compared to the theoretical result in (11). The number of samples at the sub-Nyquist sampling rate is $M_i = \sqrt{N} = 3000$, where $N = 9 \times 10^6$ denotes the number of samples at the Nyquist sampling rate.

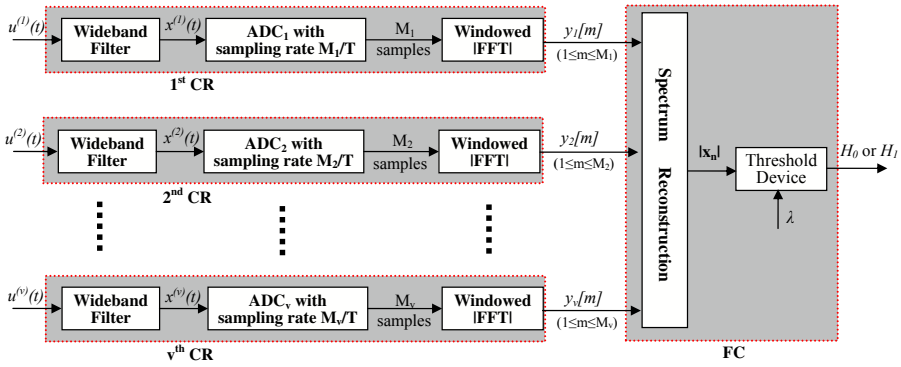

Fig. 4. The schematic illustration of the distributed MASS system.

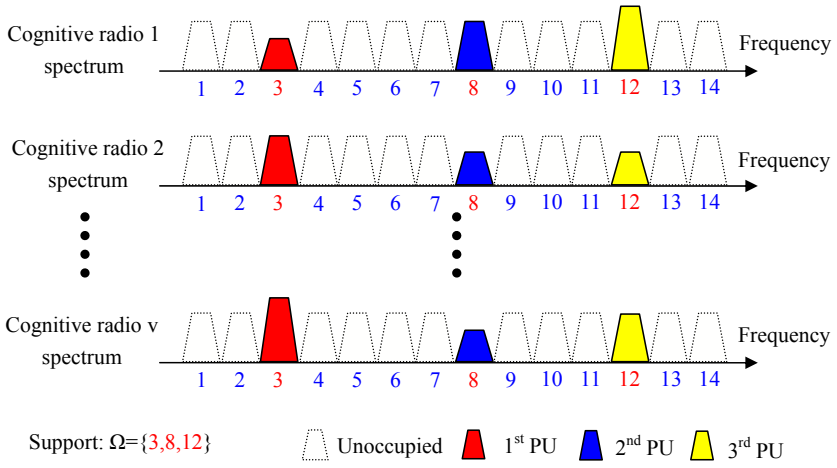

Fig. 5. Illustration of the spectral correlation in distributed CR nodes.

TABLE II
HYBRID MATCHING PURSUIT ALGORITHM FOR DISTRIBUTED MASS MODEL

Input: measurement matrix $\mathbf{\Phi}$, observation vector $\mathbf{y}$,
sparsity level k, and noise tolerance level ϱ.

1. *Initialise:*
Residual $\mathbf{r^0} = \mathbf{y}$, approximation $\mathbf{x^0} = \mathbf{0}$, support $\Omega^0 = []$, i$= 0$,
2. *While* halting criterion is false *do*
1). Form residual correlation vectors individually,
$\mathbf{c_j} = \mathbf{\Phi_j}^T \mathbf{r_j^i}$ for $j \in [1, v]$,
b). Find spectral support $\mathbf{s^i}$ by boosting
$\mathbf{s^i} = \text{Supp}_{2k}(\sum_{j=1}^{v} \mathbf{c_j})$,
c). Merge support with previously computed support by
$\Omega^i = \Omega^{i-1} \bigcup \mathbf{s^i}$,
d). Approximate the non-aliased spectrum by least squares
$\mathbf{x_\Omega^i} = \mathbf{\Phi_\Omega}^\dagger \mathbf{y}, \quad \mathbf{x_{\Omega^C}^i} = 0$,
e). Prune locations of support by
$\mathbf{p} = \text{Supp}_k(\mathbf{x^i})$,
f). Update individual residual by
$\mathbf{r_j^i} = \mathbf{y_j} - \mathbf{\Phi_j x_p^i}$ for $j \in [1, v]$,
g). i=i+1.
3. *Halting criterion:* $\|\mathbf{r^i}\|_2 \leq \varrho$

Output: A k-sparse approximation of the non-aliased spectrum, $|\mathbf{x_n}| = \mathbf{x_p^i}$.

TABLE III
LIST OF SIMULATION PARAMETERS SETUP FOR FIG. 6, FIG. 7, FIG. 8, AND FIG. 9.

	Figure 6	Figure 7	Figure 8	Figure 9
Observation time T	0.4 μs	2 μs	2 μs	2 μs
Wideband bandwidth W	5 GHz	20 GHz	20 GHz	20 GHz
Length of Nyquist samples N	4,000	80,000	80,000	80,000
No. of subbands N_b	5	10, 30, 50	30, 50	30
Bandwidth of subbands (MHz)	$0.3 \sim 30$	$0.1 \sim 5$	$0.1 \sim 5$	$0.1 \sim 5$
No. of ADCs v	10	1$\sim$150	1$\sim$150	1$\sim$150
Sampling rates of ADCs (MHz)	182.5 $\sim$ 282.5	189.5$\sim$711.5	189.5$\sim$711.5	189.5$\sim$590.5
Compression ratio	23.85%	$0.47 \sim 163.45\%$	$0.47 \sim 163.45\%$	$0.47 \sim 163.45\%$

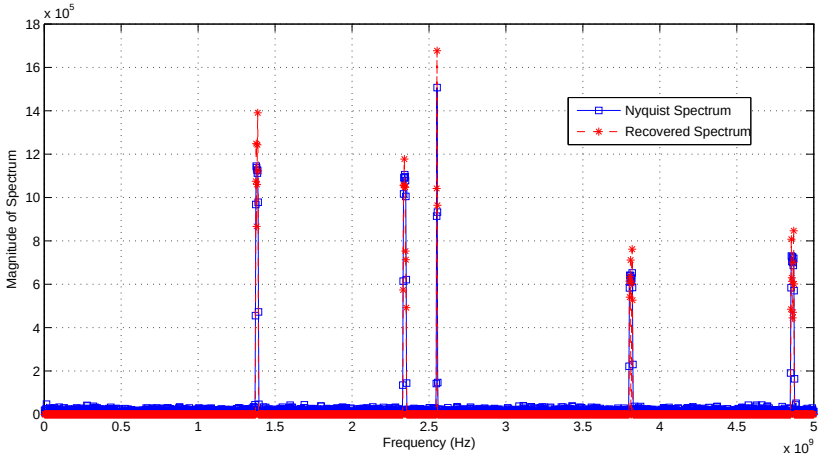

Fig. 6. Comparison of the non-aliased DFT spectrum with the recovered spectrum, when the wideband signal has 5 sub-bands with bandwidth $0.3 \sim 30$ MHz, and SNR$=15$dB. There are $v = 10$ ADCs, which experience non-fading AWGN channels.

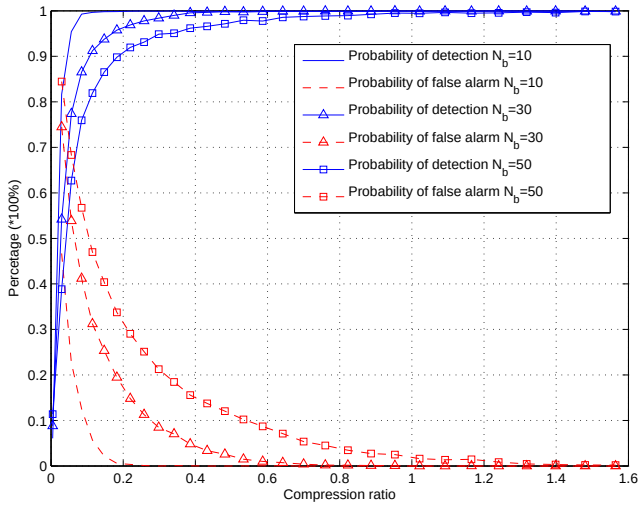

Fig. 7. Influence of the number of sub-bands, and the compression ratio on the detection performance, when the wideband signal has 10, 30, and 50 sub-bands with bandwidth $0.1 \sim 5$ MHz, and the ADCs experience non-fading AWGN channels.

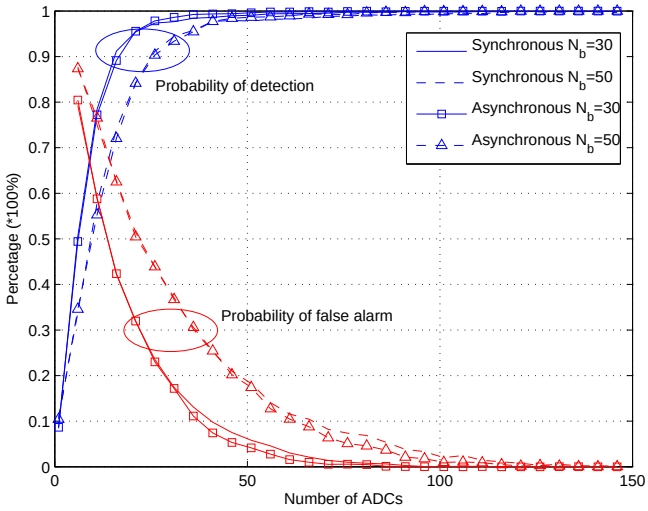

Fig. 8. Comparison of spectrum recovery performance for syn chronous ADCs and asynchronous ADCs, with $N_b = 30$ and $N_b = 50$. The asynchronous ADCs have time offsets in range of $0 \sim 0.8$ μs, with a total observation time of 2μs.

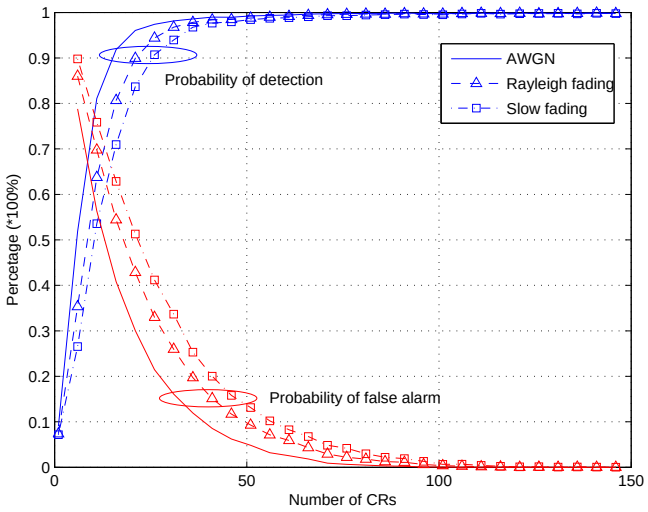

Fig. 9. Performance of MASS over AWGN, Rayleigh, and slow fading channels, with the number of sub-bands $N_b = 30$. In this experiment, the local-mean SNR is 15 dB, and the standard deviation of lognormal fading is 5 dB.

References

[1] Office of Communications, URL: www.ofcom.org.uk.

[2] Federal Communications Commission, URL: www.fcc.gov.

[3] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty, "NeXt generation/dynamic spectrum access/cognitive radio wireless networks: a survey," *Computer Networks*, vol. 50, no. 13, pp. 2127–2159, 2006.

[4] M. A. McHenry, "NSF spectrum occupancy measurements project summary," tech. rep., Shared spectrum company, Aug. 2005.

[5] J. Mitola, *Cognitive Radio: An Integrated Agent Architecture for Software Defined Radio*. PhD thesis, Dept. of Teleinformatics, Royal Institute of Technology Stockholm, Sweden, 8 May, 2000.

[6] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 23, pp. 201 – 220, Feb. 2005.

[7] W. Zhang and K. B. Letaief, "Cooperative communications for cognitive radio networks," *Proceedings of the IEEE*, vol. 97, pp. 878–893, May 2009.

[8] H. Ekram and B. V. K, *Cognitive Wireless Communications Networks*. Springer Publication, 2007.

[9] E. Hossain, D. Niyato, and Z. Han, *Dynamic Spectrum Access and Management in Cognitive Radio Networks*. Cambridge University Press, July 2009.

[10] D. Cabric, S. M. Mishra, and R. W. Brodersen, "Implementation issues in spectrum sensing for cognitive radios," in *Proceedings of Asilomar Conference on Signals, Systems, and Computers*, vol. 1, pp. 772–776, 2004.

[11] C.-X. Wang, X. Hong, H.-H. Chen, and J. Thompson, "On capacity of cognitive radio networks with average interference power constraints," *IEEE Transactions on Wireless Communications*, vol. 8, pp. 1620 –1625, Apr. 2009.

[12] W. Zhang and K. Letaief, "Cooperative spectrum sensing with transmit and relay diversity in cognitive radio networks - [transaction letters]," *IEEE Transactions on. Wireless Communications*, vol. 7, pp. 4761 –4766, Dec. 2008.

[13] B. Wang, K. Ray Liu, and T. Clancy, "Evolutionary cooperative spectrum sensing game: how to collaborate?," *IEEE Transactions on Communications*, vol. 58, pp. 890 –900, Mar. 2010.

[14] D. Duan, L. Yang, and J. Principe, "Cooperative diversity of spectrum sensing for cognitive radio systems," *IEEE Transactions on Signal Processing*, vol. 58, pp. 3218 –3227, June 2010.

[15] Z. Li, F. Yu, and M. Huang, "A distributed consensus-based cooperative spectrum-sensing scheme in cognitive radios," *IEEE Transactions on Vehicular Technology*, vol. 59, pp. 383 –393, Jan. 2010.

[16] G. Ganesan and Y. Li, "Cooperative spectrum sensing in cognitive radio, part I: Two user networks," *IEEE Transactions on Wireless Communications*, vol. 6, pp. 2204 –2213, June 2007.

[17] J. Lunden, V. Koivunen, A. Huttunen, and H. Poor, "Collaborative cyclostationary spectrum sensing for cognitive radio systems," *IEEE Transactions on Signal Processing*, vol. 57, pp. 4182 –4195, Nov. 2009.

[18] S. Haykin, D. Thomson, and J. Reed, "Spectrum sensing for cognitive radio," *Proceedings of the IEEE*, vol. 97, pp. 849 –877, May 2009.

[19] W.-Y. Lee and I. Akyildiz, "Optimal spectrum sensing framework for cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 7, pp. 3845 –3857, Oct. 2008.

[20] A. Ghasemi and E. Sousa, "Collaborative spectrum sensing for opportunistic access in fading environments," in *Proceedings of IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN 2005)*, pp. 131 –136, Nov. 2005.

[21] C.-X. Wang, H.-H. Chen, X. Hong, and M. Guizani, "Cognitive radio network management: tuning in to real-time conditions," in *IEEE Vehicular Technology Magazine*, vol. 3, pp. 28–35, Mar. 2008.

[22] M. Mishali and Y. C. Eldar, "Wideband spectrum sensing at sub-Nyquist rates." arXiv:1009.1305v1 [cs.AR], Sept. 2010.

[23] T. Yucek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *IEEE Communications Surveys Tutorials*, vol. 11, no. 1, pp. 116 –130, 2009.

[24] Z. Tian and G. B. Giannakis, "A wavelet approach to wideband spectrum sensing for cognitive radios," in *Proceedings of International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CrownCom 2006)*, pp. 1–5, June 2006.

[25] B. Farhang-Boroujeny, "Filter bank spectrum sensing for cognitive radios," *IEEE Transactions on Signal Processing*, vol. 56, pp. 1801 –1811, May 2008.

[26] B. Farhang-Boroujeny and R. Kempter, "Multicarrier communication techniques for spectrum sensing and communication in cognitive radios," *IEEE Communications Magazine*, vol. 46, pp. 80 –85, Apr. 2008.

[27] Z. Tian and G. B. Giannakis, "Compressed sensing for wideband cognitive radios," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2007)*, pp. 1357–1360, Apr. 2007.

[28] Y. L. Polo, Y. Wang, A. Pandharipande, and G. Leus, "Compressive wide-band spectrum sensing," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2009)*, pp. 2337–2340, Apr. 2009.

[29] Y. Wang, A. Pandharipande, Y. L. Polo, and G. Leus, "Distributed compressive wideband spectrum sensing," in *Proceedings of Information Theory and Applications Workshop*, pp. 178–183, Feb. 2009.

[30] Y. Bresler and P. Feng, "Spectrum-blind minimum-rate sampling and reconstruction of 2-D multiband signals," in *Proceedings of International Conference on Image Processing*, vol. 1, pp. 701 –704, Sep. 1996.

[31] P. Feng and Y. Bresler, "Spectrum-blind minimum-rate sampling and reconstruction of multiband signals," in *Proceedings of IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP 1996)*, vol. 3, pp. 1688 –1691, May 1996.

[32] M. Mishali and Y. C. Eldar, "Blind multiband signal reconstruction: Compressed sensing for analog signals," *IEEE Transactions on Signal Processing*, vol. 57, pp. 993–1009, Mar. 2009.

[33] A. Rosenthal, A. Linden, and M. Horowitz, "Multi-rate asynchronous sampling of sparse multi-band signals," *Journal of the Optical Society of America A*, vol. 25, no. 9, pp. 2320–2330, 2008.

[34] M. Fleyer, A. Linden, M. Horowitz, and A. Rosenthal, "Multirate synchronous sampling of sparse multiband signals," *IEEE Transactions on Signal Processing*, vol. 58, no. 3, pp. 1144–1156, 2010.

[35] 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, pp. 375 –391, Apr. 2010.

[36] IEEE 802.22, URL: www.ieee802.org/22/.

[37] Y. Zeng, Y.-C. Liang, A. T. Hoang, and R. Zhang, "A review on spectrum sensing for cognitive radio: Challenges and solutions," *EURASIP Journal on Advances in Signal Processing*, vol. 2010, pp. 1–15, 2010.

[38] I. Akyildiz, W.-Y. Lee, M. Vuran, and S. Mohanty, "A survey on spectrum management in cognitive radio networks," *IEEE Communications Magazine*, vol. 46, pp. 40 –48, Apr. 2008.

[39] H. Arslan, ed., *Cognitive Radio, Software Defined Radio, and Adaptive Wireless Systems*. Springer Publication, 2007.

[40] K.-C. Chen and R. Prasad, *Cognitive Radio Networks*. Wiley, June 22, 2009.

[41] J. G. Proakis, *Digital Communications*. McGraw-Hill, 4 ed., 2001.

[42] A. D'Amico, U. Mengali, and E. Arias-de Reyna, "Energy-detection UWB receivers with multiple energy measurements," *IEEE Transactions on Wireless Communications*, vol. 6, pp. 2652 –2659, July 2007.

[43] H. Urkowitz, "Energy detection of unknown deterministic signals," in *Proceedings of the IEEE*, vol. 55, pp. 523–531, Apr. 1967.

[44] V. Kostylev, "Energy detection of a signal with random amplitude," in *Proceedings of IEEE International Conference on Communications (ICC 2002)*, pp. 1606 – 1610, 2002.

[45] D.Ariananda, M. Lakshmanan, and H. Nikoo, "A survey on spectrum sensing techniques for cognitive radio," in *Proceedings of IEEE International Workshop on Cognitive Radio and Advanced Spectrum Management (CogART 2009)*, pp. 74 –79, May 2009.

[46] A. Tani and R. Fantacci, "A low-complexity cyclostationary-based spectrum sensing for UWB and WiMAX coexistence with noise uncertainty," *IEEE Transactions on Vehicular Technology*, vol. 59, pp. 2940 –2950, July 2010.

[47] K.-L. Du and W. H. Mow, "Affordable cyclostationarity-based spectrum sensing for cognitive radio with smart antennas," *IEEE Transactions on Vehicular Technology*, vol. 59, pp. 1877 –1886, May 2010.

[48] A. Perez-Neira, M. Lagunas, M. Rojas, and P. Stoica, "Correlation matching approach for spectrum sensing in open spectrum communications," *IEEE Transactions on Signal Processing*, vol. 57, pp. 4823 –4836, Dec. 2009.

[49] M. Derakhshani, M. Nasiri-Kenari, and T. Le-Ngoc, "Cooperative cyclostationary spectrum sensing in cognitive radios at low snr regimes," in *Proceedings of IEEE International Conference on Communications (ICC 2010)*, pp. 1 –5, May 2010.

[50] S. Atapattu, C. Tellambura, and H. Jiang, "Analysis of area under the ROC curve of energy detection," *IEEE Transactions on Wireless Communications*, vol. 9, pp. 1216 –1225, Mar. 2010.

[51] M. K. Simon and M.-S. Alouini, *Digital communication over fading channels*. John Wiley & Sons, Inc., 2 ed., Dec. 2004.

[52] F. F. Digham, M.-S. Alouini, and M. K. Simon, "On the energy detection of unkown signals over fading channels," in *Proceedings of IEEE International Conference on Communications (ICC 2003)*, pp. 3575–3579, May 2003.

[53] F. F. Digham, M.-S. Alouini, and M. K. Simon, "On the energy detection of unknown signals over fading channels," *IEEE Transactions on Communications*, vol. 55, pp. 21 –24, Jan. 2007.

[54] H. Harada, "A feasibility study on software defined cognitive radio equipment," in *Proceedings of IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN 2008)*, pp. 1 –12, Oct. 2008.

[55] A. Feldster, Y. Shapira, M. Horowitz, A. Rosenthal, S. Zach, and L. Singer, "Optical under-sampling and reconstruction of several bandwidth-limited signals," *Journal of Lightwave Technology*, vol. 27, pp. 1027 –1033, Apr. 2009.

[56] C. E. Shannon, "Communication in the presence of noise," *Proceedings of the IEEE*, vol. 86, pp. 447–457, Feb. 1998.

[57] R. Venkataramani and Y. Bresler, "Perfect reconstruction formulas and bounds on aliasing error in sub-Nyquist nonuniform sampling of multiband signals," *IEEE Transactions on Information Theory*, vol. 46, pp. 2173 –2183, Sep. 2000.

[58] T. Tao, "An uncertainty principle for cyclic groups of prime order," *Mathematical Research Letters*, vol. 12, pp. 121–127, 2005.

[59] D. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, pp. 1289 – 1306, Apr. 2006.

[60] F. Zeng, Z. Tian, and C. Li, "Distributed compressive wideband spectrum sensing in cooperative multi-hop cognitive networks," in *Proceedings of IEEE International Conference on Communications (ICC 2010)*, pp. 1 –5, May 2010.

[61] F. Zeng, C. Li, and Z. Tian, "Distributed compressive spectrum sensing in cooperative multihop cognitive networks," *IEEE Journal of Selected Topics in Signal Processing*, no. 99, pp. 1 –12, 2010.

[62] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Review*, vol. 43, no. 1, pp. 129–159, 2001.

[63] J. Liang, Y. Liu, W. Zhang, Y. Xu, X. Gan, and X. Wang, "Joint compressive sensing in wideband cognitive networks," in *Proceedings of IEEE Wireless Communications And Networking Conference (WCNC 2007)*, pp. 1 –5, Apr. 2010.

[64] S. Kirolos, J. Laska, M. Wakin, M. Duarte, D. Baron, T. Ragheb, Y. Massoud, and R. Baraniuk, "Analog-to-information conversion via random demodulation," in *Proceedings of IEEE Dallas/CAS Workshop on Design, Applications, Integration and Software*, pp. 71–74, Oct. 2006.

[65] S. Kirolos, T. Ragheb, J. Laska, M. E. Duarte, Y. Massoud, and R. G. Baraniuk, "Practical issues in implementing analog-to-information converters," in *Proceedings of 6th International Workshop on System-on-Chip for Real-Time Applications*, pp. 141–146, Dec. 2006.

[66] J. Laska, S. Kirolos, Y. Massoud, R. Baraniuk, A. Gilbert, M. Iwen, and M. Strauss, "Random sampling for analog-to-information conversion of wideband signals," in *Proceedings of the IEEE Dallas Circuits and Systems Workshop (DCAS 2006)*, pp. 119–122, Oct. 2006.

[67] J. N. Laska, S. Kirolos, M. F. Duarte, T. S. Ragheb, R. G. Baraniuk, and Y. Massoud, "Theory and implementation of an analog-to-information converter using random demodulation," in *Proceedings of IEEE International Symposium on Circuits and Systems (ISCAS 2008)*, pp. 1959–1962, May 2007.

[68] T. Ragheb, S. Kirolos, J. Laska, A. Gilbert, M. Strauss, R. Baraniuk, and Y. Massoud, "Implementation models for analog-to-information conversion via random sampling," in *Proceedings of 50th Midwest Symposium on Circuits and Systems*, pp. 325–328, Aug. 2007.

[69] J. A. Tropp, J. N. Laska, M. F. Duarte, J. K. Romberg, and R. Baraniuk, "Beyond Nyquist: Efficient sampling of sparse bandlimited signals," *IEEE Transactions on Information Theory*, vol. 56, pp. 520–544, Jan. 2010.

[70] D. Shnidman, "Efficient evaluation of probabilities of detection and the generalized Q-function," *IEEE Transactions on Information Theory*, vol. 22, pp. 746 – 751, Nov. 1976.

[71] I. S. Gradshteyn and I. M. Ryzhik, *Table of integrals, series, and products*. Academic Press, Inc., 5 ed., 1994.

[72] S. Herath, N. Rajatheva, and C. Tellambura, "On the energy detection of unknown deterministic signal over Nakagami channels with selection combining," in *Proceedings of Canadian Conference on Electrical and Computer Engineering (CCECE 2009)*, pp. 745 –749, May 2009.

[73] A. H. Marcus, "Power sum distributions: an easier approach using the Wald distribution," *Journal of American statistical association*, vol. 71, pp. 237–238, 1976.

[74] Karmeshu and R. Agrawal, "On efficacy of Rayleigh-inverse Gaussian distribution over K-distribution for wireless fading channels," *Journal of Wireless Communications and Mobile Computing*, vol. 7, no. 1, pp. 1–7, 2007.

[75] L. Brennan and I. Reed, "A recursive method of computing the Q function," *IEEE Transactions on Information Theory*, vol. 11, pp. 312 – 313, Apr. 1965.

[76] R. S. Chhikara and J. L. Folks, *The inverse Gaussian distribution: theory, methodology, and applications*. Marcel Dekker Inc., 1989.

[77] S. P. Herath and N. Rajatheva, "Analysis of diversity combining in energy detection for cognitive radio over Nakagami channels," in *IEEE International Conference on Communications CD-ROM (ICC 2009)*, vol. 32, pp. 2913–2917, 2009.

[78] J. M. Borwein and I. J. Zucker, "Fast evaluation of the Gamma function for small rational fractions using complete elliptic integrals of the first kind," *IMA Journal of Numerical Analysis*, vol. 12, no. 4, pp. 519–526, 1992.

[79] J. Borwein and P. Borwein, *Pi and the AGM: A Study in Analytic Number Theory and Computational Complexity*. Wiley-Interscience, 1998.

[80] G. Fedele, "N-branch diversity reception of M-ary DPSK signals in slow and nonselective Nakagami fading," *European Transactions on Telecommunications*, vol. 7, no. 2, pp. 119–123, 1996.

[81] P. Kaligineedi, M. Khabbazian, and V. Bhargava, "Secure cooperative sensing techniques for cognitive radio systems," in *Proceedings of IEEE International Conference on Communications (ICC 2008)*, pp. 3406 –3410, 2008.

[82] E. Visotsky, S. Kuffner, and R. Peterson, "On collaborative detection of TV transmissions in support of dynamic spectrum sharing," in *Proceedings of IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN 2005)*, pp. 338 –345, Nov. 2005.

[83] E. J. Candes and T. Tao, "Decoding by linear programming," *IEEE Transactions on Information Theory*, vol. 51, no. 12, pp. 4203–4215, 2005.

[84] D. L. Donoho and X. Huo, "Uncertainty principles and ideal atomic decomposition," *IEEE Transactions on Information Theory*, vol. 47, pp. 2845–2862, Nov. 2001.

[85] D. L. Donoho and M. Elad, "Optimally sparse representation in general (nonorthogonal) dictionaries via L1 minimization," in *Proceedings of the National Academy of Sciences*, vol. 100, pp. 2197–2202, Mar. 2003.

[86] J. A. Tropp, "Greed is good: algorithmic results for sparse approximation," *IEEE Transactions on Information Theory*, vol. 50, pp. 2231–2242, Oct. 2004.

[87] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Transactions on Information Theory*, vol. 53, pp. 4655–4666, Dec. 2007.

[88] E. Candes, "The restricted isometry property and its implications for compressed sensing," *Comptes Rendus Mathematique*, vol. 346, pp. 589–592, May 2008.

[89] M. F. Duarte, S. Sarvotham, D. Baron, M. B. Wakin, and R. G. Baraniuk, "Distributed compressed sensing of jointly sparse signals," in *Proceedings of 39th ASILOMAR Conference on Signals, Systems and Computation*, pp. 1537–1541, 2005.

[90] Y. Youn, H. Jeon, J. H. Choi, and H. Lee, "Fast spectrum sensing algorithm for 802.22 wran systems," in *Proceedings of IEEE International Symposium on Communications and Information Technologies (ISCIT 2006)*, pp. 960 –964, Sep. 2006.

[91] Y. Hur, J. Park, K. Kim, J. Lee, K. Lim, C.-H. Lee, H. Kim, and J. Laskar, "A cognitive radio (CR) testbed system employing a wideband multi-resolution spectrum sensing (MRSS) technique," in *Proceedings of IEEE Vehicular Technology Conference Fall (VTC 2006)*, pp. 1 –5, Sep. 2006.

[92] P. G. Bernard Mulgrew and J. Thompson, *Digital siganl processing concepts and applications*. Palgrave Macmillan, 2 ed., 2003.

[93] D. Needell and J. Tropp, "Cosamp: Iterative signal recovery from incomplete and inaccurate samples," *Journal of Applied and Computational Harmonic Analysis*, vol. 26, no. 3, pp. 301 – 321, 2009.

[94] Y. Saad, *Iterative methods for sparse linear systems*. Society for Industrial Mathematics, 2 ed., 2003.

[95] P. Welch, "The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms," *IEEE Transactions on Audio and Electroacoustics*, vol. 15, pp. 70 – 73, June 1967.

[96] F. Harris, "On the use of windows for harmonic analysis with the discrete Fourier transform," *Proceedings of the IEEE*, vol. 66, pp. 51 – 83, Jan. 1978.

[97] A. Nuttall, "Some windows with very good sidelobe behavior," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 29, pp. 84 – 91, Feb. 1981.

[98] S. Ikki and M. Ahmed, "Performance of cooperative diversity using equal gain combining (EGC) over Nakagami-m fading channels," *IEEE Transactions on Wireless Communications*, vol. 8, pp. 557 –562, Feb. 2009.

[99] G. Karagiannidis, D. Zogas, N. Sagias, S. Kotsopoulos, and G. Tombras, "Equal-gain and maximal-ratio combining over nonidentical Weibull fading channels," *IEEE Transactions on Wireless Communications*, vol. 4, pp. 841 – 846, May 2005.

[100] V. Aalo, G. Efthymoglou, T. Piboongungon, and C. Iskander, "Performance of diversity receivers in generalised gamma fading channels," *IET Communications*, vol. 1, pp. 341 –347, June 2007.

[101] A. Coulson, A. Williamson, and R. Vaughan, "A statistical basis for lognormal shadowing effects in multipath fading channels," *IEEE Transactions on Communications*, vol. 46, pp. 494 –502, Apr. 1998.

[102] P. Bithas, N. Sagias, and P. Mathiopoulos, "The bivariate generalized-k distribution and its application to diversity receivers," *IEEE Transactions on Communications*, vol. 57, pp. 2655 –2662, Sep. 2009.

[103] S. Al-Ahmadi and H. Yanikomeroglu, "On the approximation of the generalized-k distribution by a gamma distribution for modeling composite fading channels," *IEEE Transactions on Wireless Communications*, vol. 9, pp. 706 –713, Feb. 2010.

[104] X. Liu, K. Ravindran, and D. Loguinov, "A stochastic foundation of available bandwidth estimation: Multi-hop analysis," *IEEE/ACM Transactions on Networking*, vol. 16, pp. 130 –143, Feb. 2008.

[105] S.-J. Kim and G. B. Giannakis, "Sequential and cooperative sensing for multi-channel cognitive radios," *IEEE Transactions on Signal Processing*, vol. 58, pp. 4239 –4253, Aug. 2010.

[106] K. W. Choi, W. S. Jeon, and D. G. Jeong, "Sequential detection of cyclostationary signal for cognitive radio systems," *IEEE Transactions on Wireless Communications*, vol. 8, pp. 4480 –4485, Sep. 2009.

[107] J. Doob, *Stochastic processes*. Chapman & Hall, 1953.

[108] R. A. DeVore, "Deterministic constructions of compressed sensing matrices," *Journal of Complexity*, vol. 23, no. 4-6, pp. 918 – 925, 2007.

[109] R. A. Horn and C. R. Johnson, *Matrix analysis*. Cambridge University Press, 1985.