7

Primary User Detection in Multi-Antenna Cognitive Radio

Oscar Filio[1], Serguei Primak[1] and Valeri Kontorovich[2]
[1]*The University of Western Ontario*
[2]*Centre of Research and Advanced Studies (CINVESTAV-IPN)*
[1]*Canada*
[2]*Mexico*

1. Introduction

It is well known in the wireless telecommunications field that the most valuable resource available is the electromagnetic radio spectrum. Being a natural resource, it is obviously finite and has to be utilized in a rational fashion. Nevertheless the demand increase on wireless devices and services such as voice, short messages, Web, high-speed multimedia, as well as high quality of service (QoS) applications has led to a saturation of the currently available spectrum. On the other hand, it has be found that some of the major licensed bands like the ones used for television broadcasting are severely underutilized Federal Communications Commission (November) which at the end of the day results in a significant spectrum wastage. For this means, it is important to come up with a new paradigm that allows us to take advantage of the unused spectrum. Cognitive radio has risen as a solution to overcome the spectrum underutilization problem Mitola & Maguire (1999),Haykin (2005). The main idea under cognitive radio systems is to allow unlicensed users or cognitive users (those who have not paid for utilizing the electromagnetic spectrum), under certain circumstances, to transmit within a licensed band. In order to perform this task, cognitive users need to continuously monitor the spectrum activities and find a suitable spectrum band that allows them to:

- Transmit without or with the minimum amount of interference to the licensed or primary users.
- Achieve some minimum QoS required for their specific application.
- Share the spectrum with other cognitive users.

Therefore, it is easy to observe that spectrum sensing is the very task upon which the entire operation of cognitive radio rests Haykin et al. (2009). It is of extreme importance for the system to be able to detect the so-called spectrum holes (underutilized subbands of the radio spectrum). This is why in this chapter we focus all our attention to analyze some important aspects of spectrum sensing in cognitive radio, and particularly the case when it is performed using multiple antennas.
In order to take advantage of the cognitive radio features it is important to find which parts of the electromagnetic spectrum are unused at certain moment. These portions are also

called *spectrum holes* or *white spaces*. If these bands are further used by a licensed user the cognitive radio device has the alternative of either moving to another spectrum hoe or staying in the same band but altering its transmission power lever or modulation scheme in order to avoid the interference. Hence it is clear that an important requirement of any cognitive radio network is the ability to sense such spectrum holes. As the most recent literature suggests right now Akyildiz et al. (2008),Haykin et al. (2009), the most efficient way to detect spectrum holes is to detect the primary users that are receiving data within the communication range of a cognitive radio user. This approach is called *transmitter detection* which is based on the detection of the weak signal from a primary transmitter through the local observations of cognitive users. The hypotheses cab be defined as

$$x(t) = \begin{cases} \mathcal{H}_0 : & n(t) \\ \mathcal{H}_1 : & hs(t) + n(t) \end{cases}, \tag{1}$$

where $x(t)$ is the signal received by the cognitive user, $s(t)$ is the transmitted signal of the primary user, $n(t)$ is the AWGN and h is the amplitude gain of the channel. $\mathcal{H}_0$ is a null hypothesis, which states that there is no licensed user signal in a certain spectrum band. On the other hand, $\mathcal{H}_1$ is an alternative hypothesis, which states that there exist some licensed user signal. Three very famous models exist in order to implement transmitter detection according to the hypotheses model Poor & Hadjiliadis (2008). These are the matched filter detection, the energy detection and the cyclostationary feature detection.

1.1 Matched filter detection

When the information about the primary user signal is known to the cognitive user, the optimal detector in stationary Gaussian noise is the matched filter since it maximizes the received signal to noise ratio (SNR). While the main advantage of the matched filter is that it requires less time to achieve high processing gain due to coherency, it requires a priori knowledge of the primary user signal such as the modulation type and order, the pulse shape and the packet format. So that, if this information is not accurate, then the matched filter performs poorly. However, since most wireless networks systems have pilot, preambles, synchronization word or spreading codes, these can be used for coherent detection,

1.2 Energy detection

If the receiver cannot gather sufficient information about the primary user signal, for example, if the power of the random Gaussian noise is only known to the receiver, the optimal detector is an energy detector. In order to measure the energy of the received signal, the output signal of bandpass filter with bandwidth W is squared and integrated over the observation interval T. Finally, the output of the integrator Y, is compared with some threshold λ to decide whether a licensed user is present or not. Nevertheless, the performance of the energy detector is very susceptible to uncertainty in noise power. Hence, in order to solve this problem, a pilot tone from the primary transmitter can be used to help improve the accuracy of the energy detector. Another shortcoming is that the energy detector cannot differentiate signal types but can only determine the presence of the signal. Thus the energy detectors is prone to the false detection triggered by the unintended signals.

1.3 Cyclostationary feature detection

An alternative detection method is the cyclostationary feature detection. Modulated signals are in general couple with sine wave carriers, pulse trains, repeating spreading, hopping sequences or cyclic prefixes, which result in built-in periodicity Kontorovich et al. (2010). These modulated signals are characterized as cyclostationary since their mean and autocorrelation exhibit periodicity. These features are detected by analyzing a spectral correlation function. The main advantage of the spectral correlation function is that it differentiates the noise energy from modulated signal energy, which is a result of the fact that the noise is a wide-sense stationary signal with no correlation, while modulated signals are cyclostationary with spectral correlation due to the embedded redundancy of signal periodicity. Therefore, a cyclostationary feature detector can perform better than the energy detector in discriminating against noise due to its robustness to the uncertainty in noise power. Nonetheless, it is computationally complex and requires significantly long observation time.
Most of the previously mentioned techniques are investigated for a single sensor albeit some use of multiple sensors is suggested in (Zhang et al., 2010). In the latter, the authors consider a single sensor scenario equipped with multiple antennas and derived its performance in assumption of correlated antennas and constant channel. Also, most of these studies are focused on investigating the performance of particular schemes in ideal environments such as independent antennas in cooperative scenario or in uniform scattering. However, such consideration eliminate impact of real environment and its variation, while it is shown in many publications and realistic measurements that such environments change frequently, especially in highly build areas. Understanding how particular radio environment affects performance of cognitive radio sensing abilities is, therefore, and important issue to consider. Furthermore, it is well known (Haghighi et al., 2010) that the distribution of angle of arrival (AoA), itself defined by scattering environment (Haghighi et al., 2010), affects both temporal and spatial correlation of signals in antenna arrays. For these reasons in the first part of the chapter we utilize a simple but generic model of AoA distribution, suggested in (Abdi & Kaveh, 2002), to describe impact of scattering on statistical properties of received signals. Later the concept of Stochastic Degrees of Freedom (SDoF) is incorporated in order to obtain approximate expressions for the probability of miss detection in terms of number of antennas, scattering parameters and number of observations. Following, the trade-off between the number of antennas and required observation duration in correlated fading environments is investigated. It is shown that at low SNR it is more convenient having just a single antenna and many time samples so the noise suppression performs better. On the contrary, at high SNR, since the noise is suppressed relatively quickly is better to have more antennas in order to mitigate fading. Now, most of the existing spectrum schemes are based on fixed sample size detectors, which means that their sensing time is preset and fixed. Hence, in the second part of the chapter, we present some results based on the work of A. Wald (Wald, 2004) which showed that a detector based on sequential detection requires less average sensing time than a fixed size detector. We show that in general, it is possible to achieve the same performance that other fixed sample based techniques offer but using as low as half of the samples in average in the low signal to noise ratio regime. Afterwards, the impact of non-coherent detection is assessed when detecting signals using sequential analysis. We finished using sequential analysis as a new approach of cooperative approach for sensing. We call this an optimal fusion rule for distributed Wald detectors and a evaluate its performance. The last section of the chapter is devoted to conclusion remarks.

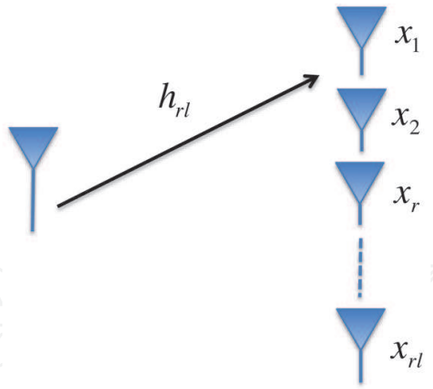


Fig. 1. System Model

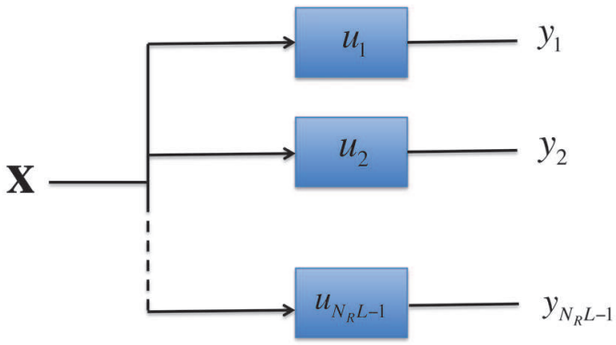


Fig. 2. Filtered Observations

2. Impact of scattering environment in spectrum sensing in multi-antenna cognitive radio systems

2.1 Signal model

Let us consider a primary transmitter which transmits some pilot signal s over L symbols in order to sound the primary channel. CR can sense the same signal using N_R receiving antennas. The received signal matrix $\mathbf{X}$ of size $N_R \times L$ can be written in terms of the $N_R \times L$ complex channel matrix $\mathbf{H} = \{h_{rl}\}$ and the noise matrix $\mathbf{W}$ of the same size as

$$\mathbf{X} = \mathbf{H}s + \mathbf{W}, \tag{2}$$

Here $\mathbf{W}$ is a zero mean Gaussian matrix of covariance $\sigma_n \mathbf{I}$ and $\mathbf{H}$ is a zero mean Gaussian matrix with covariance matrix $\mathbf{R_H}$ respectively. Element h_{rl} is the channel transfer coefficient from the transmitter to r-th antenna measured at l-th pilot symbol. Using vectorization operation ((van Trees, 2001)) , one can rewrite (2) as

$$\mathbf{x} = \mathbf{h}s + \mathbf{w}, \tag{3}$$

where $\mathbf{x} = \text{vec}\,\mathbf{X}$, $\mathbf{h} = \text{vec}\,\mathbf{H}$ and $\mathbf{w} = \text{vec}\,\mathbf{W}$[1]. Therefore, the detection problem is to distinguish between the hypotheses

$$\begin{array}{ll} \mathcal{H}_0 : x[n] = & w[n] \qquad n = 0, 1, \ldots, N_R L - 1 \\ \mathcal{H}_1 : x[n] = h[n]s + w[n] & n = 0, 1, \ldots, N_R L - 1 \end{array}. \tag{4}$$

[1] The vec$(\cdot)$ operator is defined as the $N_R L \times 1$ vector formed by stacking the columns of the $N_R \times L$ matrix i.e. $\text{vec}\,\mathbf{H} = [\mathbf{h}_1' \mathbf{h}_2' \ldots \mathbf{h}_L']'$

The sufficient statistic in this case is given by (van Trees, 2001), (Kay, 1998)

$$\mathcal{T} = \mathbf{x}^H \mathbf{Q} \mathbf{x} = |s|^2 \mathbf{x}^H \mathbf{R_h} \left[|s|^2 \mathbf{R_h} + \sigma_n^2 \mathbf{I} \right]^{-1} \mathbf{x}, \tag{5}$$

where $\mathbf{R_h} = \mathcal{E}\left\{ \mathbf{h}\mathbf{h}^H \right\}$ is the correlation matrix of the channel vector $\mathbf{h} = \operatorname{vec} \mathbf{H}$. This correlation matrix reflects both spatial correlation between different antennas and the time-varying nature of the channel. Let $\mathbf{R_h} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H$ be eigendecomposition of the correlation matrix $\mathbf{R_h}$. In this case, the statistic $\mathbf{T}$ could be recast in terms of the elements of the eigenvalues λ_i of the matrix $\mathbf{\Lambda}$ and filtered observations $\mathbf{y} = \mathbf{U^H x}$:

$$\mathcal{T} = \mathbf{y}^H \mathbf{\Lambda} \left[\mathbf{\Lambda} + \sigma_n^2 \mathbf{I} \right]^{-1} \mathbf{y} = \sum_{k=1}^{N_R L} \frac{\lambda_k^2}{\lambda_k^2 + \sigma_n^2} |y_k|^2, \tag{6}$$

which is analogous to equation (5.9) in (van Trees, 2001). Elements y_k of the vector $\mathbf{y}$ could be considered as filtered version of the received signal $\mathbf{x}$ with a set of orthogonal filters $\mathbf{u}_k$ (columns of the matrix $\mathbf{U}$), *i.e.* could be considered as multitaper analysis (Thomson, 1982). Linear filtering preserve Gaussian nature of the received signals, therefore, the distribution of $\mathcal{T}$ could be described by generalized χ^2 distribution[2] (Andronov & Fink, 1971):

$$p(x) = \sum_{k=1}^{N_R L} \alpha_k \exp(-x/2\lambda_k), \tag{7}$$

and

$$\alpha_k^{-1} = 2\lambda_k \prod_{l=1, l\neq k}^{N_R L} \left(1 - \frac{\lambda_l}{\lambda_k} \right). \tag{8}$$

Theoretically, equation (7) could be used to set up the detection threshold γ. However, it is difficult to use it for analytical investigation. Therefore, we would consider a few particular cases of the channel when the structure of the correlation matrix could be greatly simplified to reveal its effect on the detection performance.

2.2 Performance of estimator-correlator for PU detection

2.2.1 Constant independent channels

In this case the full covariance matrix $\mathbf{R_h} = \sigma_h^2 \mathbf{O}_L \otimes \mathbf{I}_{N_R}$ is a Kronecker product of $N_R \times N_R$ identity correlation matrix $\mathbf{I}_{N_R}$ and $\mathbf{O}_L = \mathbf{1}\mathbf{1}^H$ is a $L \times L$ matrix consisting of ones. Therefore, there are N_R eigenvalues λ_k, $k = 1, \cdots N_R$ equal to L. The k-th orthogonal filter $\mathbf{u}_k$ is the averaging operator applied to the data collected from the k-th antenna. Thus, the decision statistic is just

$$\mathcal{T}_{CI} = \sum_{k=1}^{N_R} \left| \sum_{l=1}^{L} x_{kl} \right|^2 = \sum_{k=1}^{N_R} P_k, \tag{9}$$

where

$$P_k = \left| \sum_{l=1}^{L} x_{il} \right|^2. \tag{10}$$

[2] Assuming that all eigenvalues λ_k of $\mathbf{R_h}$ are different.

In absence of the signal, samples x_{kl} are drawn from an i.i.d. complex Gaussian random variable with zero mean and variance σ_n^2. Therefore, the distribution of P_k is exponential, with the mean value $L\sigma_n^2$

$$p(P) = \frac{1}{L\sigma_n^2} \exp\left(-\frac{P}{L\sigma_n^2}\right), \tag{11}$$

and the distribution of $\mathcal{T}$ is just central χ^2 distribution with N_R degrees of freedom

$$p_{CI}(\mathcal{T}|\mathcal{H}_0) = \frac{1}{\Gamma(N_R)} \frac{\mathcal{T}^{N_R-1}}{(L\sigma_n^2)^{N_R}} \exp\left(-\frac{\mathcal{T}}{L\sigma_n^2}\right). \tag{12}$$

If γ is a detection threshold for the statistic $\mathcal{T}$ then the probability P_{FA} of the false alarm is

$$P_{FA} = \int_\gamma^\infty p_{CI}(\mathcal{T}|\mathcal{H}_0)d\mathcal{T} = \frac{\Gamma\left[N_R, \gamma/L\sigma_n^2\right]}{\Gamma(N_R)}, \tag{13}$$

or

$$\gamma_{CI} = L\sigma_n^2 \Gamma^{-1}\left[N_R, P_{FA}\Gamma(N_R)\right], \tag{14}$$

where $\Gamma^{-1}\left[N_R, \Gamma(N_R, x)\right] = x$. If the signal is present, *i.e.* if the hypothesis $\mathcal{H}_1$ is valid, then the signal y_i is a zero mean with the variance $\sigma^2 = L^2|s|^2\sigma_h^2 + L\sigma_n^2$. As the result, the distribution of the test statistic $\mathcal{T}$ under the hypothesis $\mathcal{H}_1$ is given by the central χ^2 distribution with N_R degree of freedom and the probability of the detection is just

$$P_D = \int_\gamma^\infty p(\mathcal{T}|\mathcal{H}_1) = \frac{1}{\Gamma(N_R)} \Gamma\left(N_R, \frac{\gamma}{\sigma^2}\right) = \frac{1}{\Gamma(N_R)} \Gamma\left(N_R, \frac{1}{1+L\bar{\mu}} \Gamma^{-1}\left[N_R, P_{FA}\Gamma(N_R)\right]\right), \tag{15}$$

where

$$\bar{\mu} = |s|^2 \frac{\sigma_h^2}{\sigma_n^2}, \tag{16}$$

is the average SNR per symbol. Performance curves for this case could be found in (van Trees, 2001).

It could be seen from both (9) and (15) that under the stated channel model, the improvement in performance P_D comes either through reduction of noise through accumulation of signal in each of the antennas (*i.e.* increase in the effective SNR) or through exploitation of diversity provided by N_R antennas Kang et al. (2010). Thus, increasing number of antennas leads to a faster detection.

2.2.2 Spatially correlated block fading (constant spatially correlated channel)

Now let us assume that the values of the channel remain constant over L symbols but the values of the channel coefficients for different antennas are correlated. In other words we will assume that $\mathbf{R_h} = \sigma_h^2 \mathbf{O}_L \otimes \mathbf{R}_s$ where $\mathbf{R}_s$ is the spatial correlation matrix between antennas. Let $\mathbf{R}_s = \mathbf{U}_s \mathbf{\Lambda}_s \mathbf{U}_s^H$ be spectral decomposition of $\mathbf{R}_s$. Then the test statistic $\mathcal{T}$ could be expressed, according to equation (6), as

$$\mathcal{T}_{CC} = \sum_{k=1}^{N_R} \frac{|s|^2\sigma_h^2\lambda_k}{|s|^2\sigma_h^2\lambda_k + \sigma_n^2} |y_k|^2 = \sum_{k=1}^{N_R} \frac{\bar{\mu}\lambda_k}{\bar{\mu}\lambda_k + 1} |y_k|^2, \tag{17}$$

where σ_h^2 is the variance of the channel per antenna. The eigenvalues λ_k of $\mathbf{R}_s$ reflect time accumulation of SNR in each "virtual branch" of the equivalent filtered value y_k. In general,

all the eigenvalues are different so one should utilize equation (7). While these calculations are relatively easy to implement numerically, it gives little insight into the effect of correlation on the performance of the detector.
Under certain scattering conditions (Haghighi et al., 2010), the eigenvalues of the matrix $\mathbf{R}_s$ are either all close to some constant $\lambda > 1$ or close to zero. If there is $N_{eq} < N_R$ non-zero eigenvalues, their values are equal to $\lambda_k = N_R/N_{eq}$ to preserve trace, and the rest $N_R - N_{eq}$ are equal to zero. In this case, the test statistic $\mathcal{T}_{CC}$ could be further simplified to

$$\mathcal{T}_{CC}(N_{eq}) = \sum_{k=1}^{N_{eq}} |y_k|^2, \tag{18}$$

where the index k corresponds to non-zero eigenvalues. Thus, the problem is equivalent to one considered in Section 2.2.1 with N_{eq} independent antennas and the expression for the threshold γ_{CC} and the probability of detection are given by

$$\alpha\gamma_{CC} = \sigma_n^2 \Gamma^{-1}\left[N_{eq}, P_{FA}\Gamma(N_{eq})\right], \tag{19}$$

where $0 < \alpha < 1$ performs as a corrector variable.
The effect of correlation between branches has dual effect on performance of the system. On one side, the number N_{eq} of equivalent independent branches is reduced, comparing to the number of antennas N_R, therefore reducing diversity. However, increased correlation results into additional accumulation of SNR (or, equivalently, additional noise reduction through averaging) by factor of $N_R/N_{eq} \geq 1$. Therefore

$$P_D = \int_{\gamma}^{\infty} p(\mathcal{T}|\mathcal{H}_1)d\mathcal{T} = \frac{1}{\Gamma(N_{eq})}\Gamma\left(N_{eq}, \frac{\alpha L\gamma_{CC}}{\sigma^2}\right) = \frac{1}{\Gamma(N_{eq})}\Gamma\left(N_{eq}, \frac{1}{1+LN_R\bar{\mu}/N_{eq}}\Gamma^{-1}\left[N_{eq}, P_{FA}\Gamma(N_{eq})\right]\right). \tag{20}$$

2.2.3 Independent channel with temporal correlation

In the case of independent antennas but temporally correlated fading, the full correlation matrix can be represented as $\mathbf{R}_h = \mathbf{R}_T \otimes \mathbf{I}_L$ where $\mathbf{R}_T = \mathbf{U}_T^H \mathbf{\Lambda}_T \mathbf{U}$ is the eigen decomposition temporal correlation matrix of an individual channel Paulraj et al. (2003). The decision statistic can now be represented as

$$\mathcal{T}_{ICC} = \sum_{k=1}^{N_R} \mathbf{x}_k^H \mathbf{R}_T \left(\mathbf{R}_T + \frac{1}{\bar{\mu}}\mathbf{I}_L\right)^{-1} \mathbf{x}_k = \sum_{k=1}^{N_R} \mathcal{T}_k, \tag{21}$$

where $\mathbf{x}_k$ is $1 \times L$ time sample received by the k-th antenna. Therefore, each antenna signal is processes separately and the results are added afterwards.
Taking advantage of eigendecomposition of the correlation matrix $\mathbf{R}_T$ calculation of decision statistic $\mathcal{T}_k$ can be recast as a multitaper analysis

$$\mathcal{T}_k = \mathbf{y}_k \mathbf{\Lambda}_k \left(\mathbf{\Lambda}_k + \frac{1}{\bar{\mu}}\mathbf{I}_L\right)^{-1} \mathbf{y}_k = \sum_{l=1}^{L} \frac{\lambda_l}{\lambda_l + 1/\bar{\mu}} |y_{kl}|^2. \tag{22}$$

Once again, we can utilize approximation of the correlation matrix by one with constant or zero eigenvalues as in Section 2.2.2. In this case there will be

$$L_{eq} = \frac{(\mathrm{tr}\,\mathbf{R}_T)^2}{\mathrm{tr}\,\mathbf{R}_T\mathbf{R}_T^H} \tag{23}$$

eigenvalues of size L/L_{eq} and the rest are zeros. Therefore, there is $N_R L_e q$ terms in the sum (21) each contributing

$$\frac{L/L_{eq}}{L/L_{eq} + 1/\bar{\mu}} = \frac{\bar{\mu}L + L_{eq}}{\bar{\mu}L}, \tag{24}$$

into the variance of $\mathcal{T}_{ICC}$. Corresponding equations for choosing the threshold become

$$\gamma_{CC} = L\sigma_n^2 \Gamma^{-1}\left[N_R L_{eq}, P_{FA}\Gamma(N_R L_{eq})\right], \tag{25}$$

$$P_D = \int_\gamma^\infty p(\mathcal{T}|\mathcal{H}_1)d\mathcal{T} = \frac{1}{\Gamma(N_{eq})}\,\Gamma\left(N_{eq}, \frac{\gamma}{\sigma^2}\right) = \frac{1}{\Gamma(N_{eq})}\,\Gamma\left(N_{eq}, \frac{1}{1 + LN_R\bar{\mu}/N_{eq}}\,\Gamma^{-1}\left[N_{eq}, P_{FA}\Gamma(N_{eq})\right]\right). \tag{26}$$

2.2.4 Channel with separable spatial and temporal correlation

The correlation matrix of the channel with separable temporal and spatial correlation has the correlation matrix of the form $\mathbf{R_h} = \mathbf{R}_T \otimes \mathbf{R}_s$. Correlation in both coordinates reduces total number of degrees of freedom from $N_R L$ to $N_{eq}L_{eq} \leq N_R L$ The loss of degrees of freedom is offset by accumulation of SNR due to averaging over the correlated samples. The equivalent increase in the average SNR is $N_R L/N_{eq}L_{eq}$. Thus, the problem is equivalent to detection using

$$K_{eq} = N_{eq}L_{eq} = \frac{(\mathrm{tr}\mathbf{R}_s)^2}{||\mathbf{R}_s||_F^2}\,\frac{(\mathrm{tr}\mathbf{R}_T)^2}{||\mathbf{R}_T||_F^2}, \tag{27}$$

independent samples in the noise with the average SNR

$$\bar{\mu}_{eq} = \frac{N_R L}{N_{eq}L_{eq}}\bar{\mu}. \tag{28}$$

The sufficient statistics in the case of the channel with separable spatial and temporal correlation could be easily obtained from the general expression (5) and (6). In fact, using Kronecker structure of $\mathbf{R_h}$ one obtains

$$\mathcal{T} = \sum_{k=1}^{K_{eq}} |z_k|^2. \tag{29}$$

2.3 Examples and simulation

2.3.1 Correlation models

While the Jakes correlation function $J_0(2\pi f_D \tau)$ is almost universally used in standards on wireless channels (Editors, 2006), realistic environment is much more complicated. A few other models could be found in the literature, some chosen for their simplicity, some are based on the measurements. In most cases we are able to calculate N_{eq} analytically, as shown below.

1. **Sinc type correlation** If scattering environment is formed by a single remote cluster (as it is shown in (Haghighi et al., 2010)), then the spatial covariance function $\mathbf{R}_s(d)$ as a function of electric distance between antennas d is given by

$$\mathbf{R}_s(d) = \exp\left(j2\pi d \sin\phi_0\right) \operatorname{sinc}\left(\Delta\phi d \cos\phi_0\right), \tag{30}$$

where ϕ_0 is the central angle of arrival, $\Delta\phi$ is the angular spread. This correlation matrix has approximately $\lfloor 2\Delta\phi \cos\phi_0 N + 1 \rfloor$ eigenvalues approximately equal eigenvalues with the rest close to zero (Slepian, 1978).

2. **Nearest neighbour correlation** Neglecting correlation between any two antennas which are not neighbours one obtains the following form of the correlation matrix $\mathbf{R}_s$

$$\mathbf{R}_s = \left\{r_{ij}\right\} = \begin{cases} 1 & \text{if } i = j \\ \rho & \text{if } i = j+1 \\ \rho^* & \text{if } i = j-1 \\ 0 & \text{if } |i-j| > 1 \end{cases}, \tag{31}$$

where ρ is the correlation coefficient. The eigenvalues of (31) are well know (Kotz & Adams, 1964)

$$\lambda_k = 1 - 2|\rho| \cos\frac{k\pi}{N+1}, \quad 1 \le k \le N. \tag{32}$$

The equivalent number of independent virtual antennas is given by

$$N_{eq} = \frac{N^2}{N + 2(N-1)|\rho|^2} = \frac{N}{1 + 2|\rho|^2\left(1 - 1/N\right)}. \tag{33}$$

3. **Exponential correlation**

$$\mathbf{R}_s = \left\{r_{ij}\right\} = \left\{|\rho|^{i-j}\right\}. \tag{34}$$

Eigenvalues of this matrix are well known (34) (Kotz & Adams, 1964)

$$\lambda_k = \frac{1 - |\rho|^2}{1 + 2|\rho|\cos\psi_k + |\rho|^2}, \tag{35}$$

where ψ_k are roots of the following equation

$$\frac{\sin(N+1)\psi - 2|\rho|\psi\sin N + |\rho|^2\sin(N-1)\psi}{\sin\psi} = 0. \tag{36}$$

4. **Temporal correlation model for nonisotropic scattering** Considering the extended case of the Clarke's temporal correlation model for the case of nonisotropic scattering around the user, we have the temporal correlation function as (Abdi & Kaveh, 2002):

$$\mathbf{R}_s(\tau) = \frac{I_0\left(\sqrt{\kappa^2 - 4\pi^2 f_D^2\tau^2 + j4\pi\kappa\cos(\theta) f_D\tau}\right)}{I_0(\kappa)}, \tag{37}$$

where $\kappa \ge 0$ controls the width of angle of arrival (AoA), f_d is the Doppler shift, and $\theta \in [-\pi, \pi)$ is the mean direction of AoA seen by the user; $I_0(\cdot)$ stands for the zeroth-order modified Bessel function.

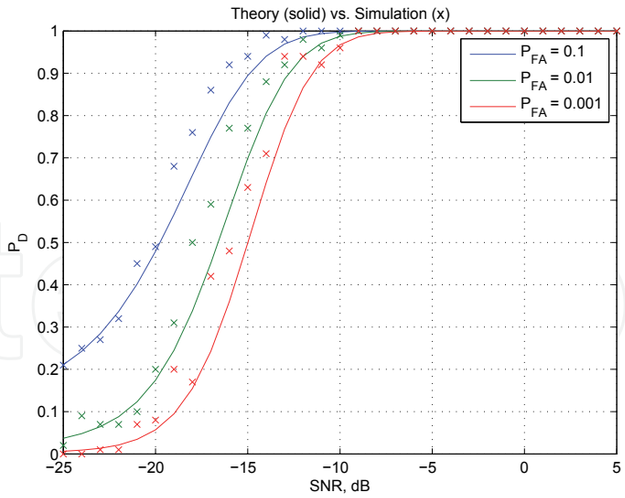


Fig. 3. ROC approximation vs. simulation results ($\alpha = 0.8$) . Solid lines - theory, x-lines - simulation.

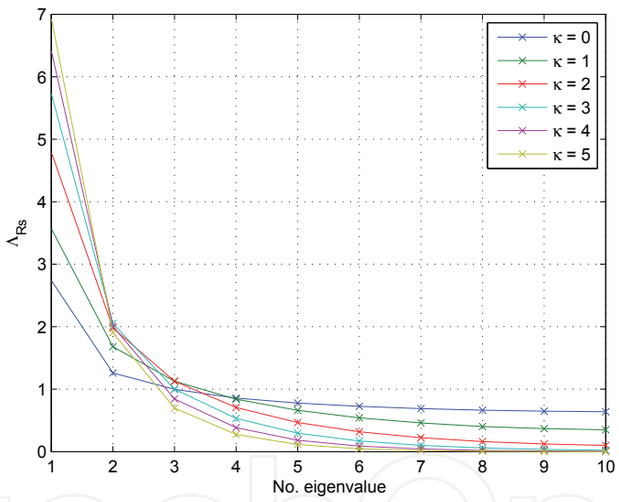


Fig. 4. Eigenvalues behavior of R_s temporal correlation matrix for nonisotropic scattering ($N = 10, \mu = 0$ and $f_d = 50Hz$)

Figure 4 shows the eigenvalues behaviour for different values of the κ factor. Notice that for $\kappa = 0$ (isotropic scattering) the values of the eigenvalues are spread in an almost equally and proportional fashion among all of them. As κ tends to infinity (extremely nonisotropic scattering), we obtain $N - 1$ zero eigenvalues and one eigenvalue with value N. In other words, as κ increases, the number of "significant" eigenvalues decreases and hence so the value of N_{eq} as shown in Figure 6.

2.4 Simulation procedure

In order to perform the simulations which verified these results, the hypothesis in eq. (4) was formed by giving the channel matrix $\mathbf{H}$ the desired correlation characteristics as shown in

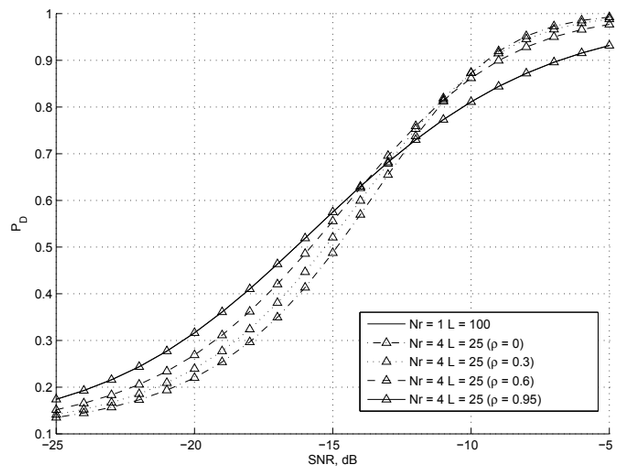


Fig. 5. Effect of correlation between antennas in the probability of detection.

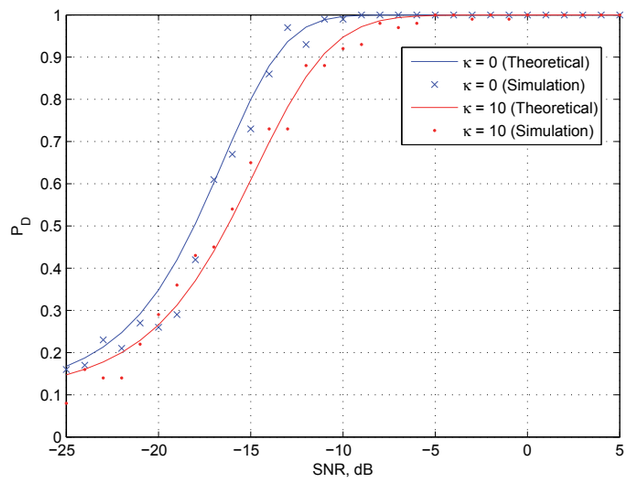


Fig. 6. ROC approximation for the estimator correlator considering the temporal correlation for isotropic and nonisotropic scattering ($\kappa = 0$ and $\kappa = 10$ respectively).

section 2.3.1. Therefore, the vectorization operations are performed and after evaluating the respective statistical tests, Monte Carlo method is utilized.

2.5 Space-time processing trade-off

It is common to assume that increasing number of antennas improves performance of detection algorithms due to increased degree of diversity. Such proposition is correct when the number of time samples remains the same. However, in cognitive networks it is desired to reduce decision time as much as possible, sometimes by introducing some added complexity in the form of additional number of antennas. The goal of this section is to show how one can trade speed of making decision with a number of antennas available for signal reception.
It can be seen from equation (18) that the processing of the signal consists of two separated procedures: averaging in time and accounting for diversity and suppressing noise in spatial

diversity brunches. Depending on amount of noise (SNR) and fading (fading figure (Simon & Alouini, 2000)) one of these two technique brings more benefit to the net procedure. For relatively low levels of SNR noise suppression is a dominant task, therefore it is more advantageous to have a single antenna and as many time samples as possible. On a contrary, if SNR is somewhat higher, noise is sufficiently suppressed even by short time averaging and suppressing fading through diversity combining is more beneficial. This can be seen from Fig. 5. This graph shows performance different configurations of the receiver in such a way that product $N_R L$ remains constant. The same Figure shows effect correlation, and thus, the scattering environment, plays on quality of reception. For very strong correlations $\rho \approx 1$ and $N_{eq} \approx 1$. Therefore all samples collected are used to reduce noise. Such scheme performs the best at low SNR. However, when $\rho = 0$ and $N_{eq} = N_R$ the gain from the diversity is highest and the scheme performs best for higher SNR. Intermediate case allows for smooth transition between these two regions. It also can be seen, that only in the case of $\rho = 1$ there is an equivalent trade off between number of antennas and time samples: *i.e.* the performance depends only on $Q = N_R L$ regardless how the product is divided between N_R and L. However, lesser correlation result in unequal trade-off with gain or loss defined by SNR and amount of correlation.

3. Sequential analysis for multi-antenna cognitive radio

3.1 Sequential analysis overview

The sequential analysis and the sequential probability ratio test (SPRT) introduced by A. Wald in 1943 (Wald, 2004) have proved to be highly effective in taking decisions between two known hypothesis $(\mathcal{H}_0, \mathcal{H}_1)$. Moreover, as it is shown in (Wald, 2004), the sequential probability ratio test frequently results in a saving of about 50 percent in the average number of observations in comparison to other well known detection techniques such as the Neyman-Pearson (NP) decision test which is based on fixed number of observations. Unlike NP test, which utilizes the logarithm of the Likelihood Ratio (log-LR) and compares it with a single predefined threshold γ, the sequential probability ratio test compares the log-LR with two thresholds $\mathcal{A}$ and $\mathcal{B}$ which are obtained in terms of the target probability of false alarm (P_{FA}) and probability of misdetection (P_{MD}) (or the complementary probability of detection $P_D = 1 - P_{MD}$) (Wald, 2004), (Middleton, 1960). Furthermore, in contrast to fixed decision time of NP test, the duration of testing of sequential analysis is a random variable.
The thresholds $\mathcal{A}$ and $\mathcal{B}$ are approximated as (Wald, 2004):

$$\mathcal{A} = \ln \frac{1 - P_{MD}}{P_{FA}}, \ \mathcal{B} = \ln \frac{P_{MD}}{1 - P_{FA}} \tag{38}$$

The test procedure consists of sequential accumulating of m samples and calculating the cumulative sum of the m-th log-LR as

$$\mathcal{B} < \sum_{i=1}^{m} \Lambda_i < \mathcal{A}, \tag{39}$$

where Λ_i is a single log-likelihood ratio sample.
If Eq. (39) is satisfied, the experiment is continued by taking an additional sample increasing m by 1. However, if

$$\sum_{i=1}^{m} \Lambda_i \geq \mathcal{A}, \tag{40}$$

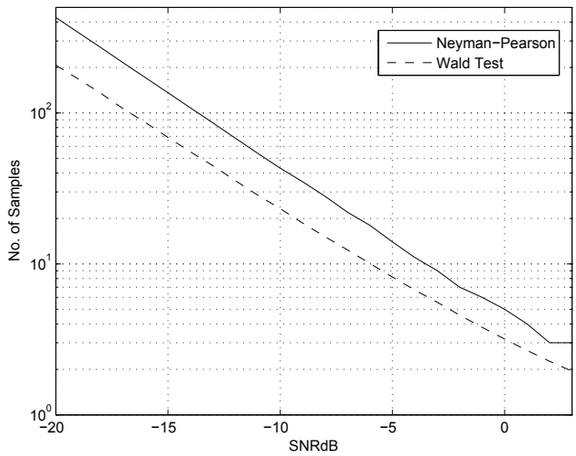


Fig. 7. Comparison of Neyman-Pearson Test and Sequential Probability Ratio Test ($P_{FA} = 0.1$,$P_D = 0.9$).

the process is terminated with the acceptance of $\mathcal{H}_1$. Similarly

$$\sum_{i=1}^{m} \Lambda_i \leq \mathcal{B}, \tag{41}$$

leads to termination with the acceptance of $\mathcal{H}_0$.

3.2 Wald test for complex random variable

Let us consider testing zero mean hypothesis in complex AWGN

$$\begin{array}{l} \mathcal{H}_0 : z_i = x_i + jy_i = w_i \\ \mathcal{H}_1 : z_i = \quad m + w_i \end{array}. \tag{42}$$

Here $m = m_I + jm_Q = \mu \exp(j\phi_m) \neq 0$ is complex non zero mean and w_i is i.i.d. complex zero mean Gaussian process of variance σ^2.
A single sample log-likelihood ratio Λ_i is given by

$$\Lambda_i = \ln \frac{p_1(z_i; \mathcal{H}_1)}{p_0(z_i; \mathcal{H}_0)} = \frac{2\mu \left(x_i \cos \phi_m + y_i \sin \phi_m \right) - \mu^2}{\sigma^2}. \tag{43}$$

After N steps of the sequential test the cumulative log-likelihood Λ ratio becomes

$$\Lambda = \sum_{n=1}^{N} \Lambda_n = \frac{2\mu}{\sigma^2} \mathcal{T}_N - \frac{N\mu^2}{\sigma^2}, \tag{44}$$

where

$$\mathcal{T}_N = \cos \phi_m \sum_{n=1}^{N} x_n + \sin \phi_m \sum_{n=1}^{N} y_n. \tag{45}$$

The rest of the test follows procedure of Section 3.1.

Figure 7 shows the performance comparison in number of samples needed between Wald Test and Neyman-Pearson Test. Notice that the latter needs in general almost twice the number of samples in order to detect the presence of the signal.
It follows from (45) that the sufficient statistic in the case of complex observation is

$$\mathcal{T} = \sum_{n=1}^{N} \Re\{x \exp(-j\phi_m)\}. \tag{46}$$

In other words, processing of the received signal is implemented in two stages: first, the data is unitary rotated by the angle ϕ_m to align the mean along real axis; then the real part of data is analyzed using the same procedure as purely real data.

3.2.1 Average number of samples

We have defined the sequential test procedure in eq.(39). Using this we can call

$$\Lambda = \sum_{i=1}^{m} \Lambda_i = \ln \frac{p(z_1, \ldots, z_m|\mathcal{H}_1)}{p(z_1, \ldots, z_m|\mathcal{H}_0)}, \tag{47}$$

where the random variable m stands for the required number of samples needed to terminate the test. As stated in Wald (2004) it is possible to neglect the excess on threshold $\mathcal{A}$ and $\mathcal{B}$, hence the the random variable can have the four possible combinations of terminations and hypotheses such as

$$\Lambda = \begin{cases} P_{FA}\mathcal{A} \text{ if } \mathcal{H}_0 \text{ is true} \\ P_D\mathcal{A} \text{ if } \mathcal{H}_1 \text{ is true} \\ (1 - P_{FA})\mathcal{B} \text{ if } \mathcal{H}_0 \text{ is true} \\ P_M\mathcal{B} \text{ if } \mathcal{H}_1 \text{ is true} \end{cases}. \tag{48}$$

Following same reasoning we can get the conditional expectation for the random variable Λ as

$$\bar{\Lambda} = \begin{cases} P_{FA}\mathcal{A} + (1 - P_{FA})\mathcal{B} & \text{if } \mathcal{H}_0 \text{ is true} \\ P_D\mathcal{A} + P_M\mathcal{B} & \text{if } \mathcal{H}_1 \text{ is true} \end{cases}. \tag{49}$$

It is possible now to obtain the average number of samples (decision time) for accepting one of the two hypothesis $(\mathcal{H}_0, \mathcal{H}_1)$ as:

$$\bar{n}(\mathcal{H}_0) = \frac{P_{FA}\mathcal{A} + (1 - P_{FA}\mathcal{B})}{\bar{\Lambda}(\mathcal{H}_0)}, \tag{50}$$

$$\bar{n}(\mathcal{H}_1) = \frac{P_D\mathcal{A} + (1 - P_D\mathcal{B})}{\bar{\Lambda}(\mathcal{H}_1)}, \tag{51}$$

where the term $\bar{\Lambda}(\mathcal{H}_0)$ can be calculated as

$$\bar{\Lambda}(\mathcal{H}_0) = \frac{\sum_{n=1}^{N} \Lambda_n}{N}, \tag{52}$$

if no signal is present. The term $\bar{\Lambda}(\mathcal{H}_1)$ can be calculated analogously assuming there is a signal present as follows

$$\bar{\Lambda}(\mathcal{H}_1) = \frac{\sum_{n=1}^{N} \Lambda_n}{N}. \tag{53}$$

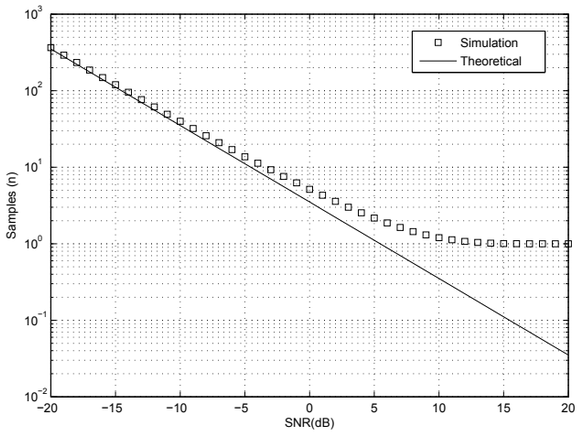


Fig. 8. Average Number of Samples for Detection in Sequential Analysis

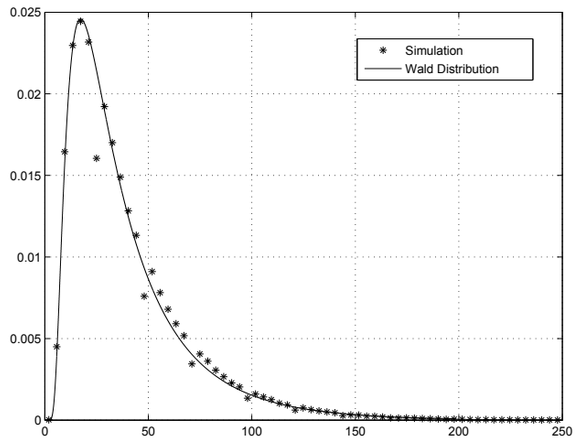


Fig. 9. Approximation of decision time using Wald Distribution.

Fig. 8 shows the average number of samples needed to achieve a $P_D = 0.9$ for different SNR. The deviation at high SNR's is due to the same effect already explained before. In practice for very high SNR's only one sample is enough to detect the presence of primary users.

3.2.2 Decision time distribution

as it was described earlier, decision time when using sequential analysis for detection is a random variable. Hence it can be described by its PDF. Although an exact shape for this PDF is not known in general, a very good approximation is available (specially for the low SNR regimen) called Wald distribution or inverse Gaussian distribution defined as

$$f(x) = \frac{\lambda}{2\pi x^3} \exp \frac{-\lambda(x-\mu)^2}{2\mu^2 x} \qquad x > 0, \tag{54}$$

where μ is the mean and $\lambda > 0$ is the shape parameter. In figure 9 it is shown Wald's distribution in order to approximate the decision time for a $P_D = 0.9$.

Though application of sequential analysis with high reliability of the hypothesis testing (P_{FA}, $P_M \to 0$) can provide an effective censoring of the information sent to other SU or FC together with reduction of the sampling size at SU.

3.3 Sequential probability ratio test for partially coherent channel

Let us consider detection of a signal in a channel a with partially known phase. The received signal can be modeled as

$$z_i = m \exp(j\Delta) + w_i, \tag{55}$$

where $m = m_I + jm_Q = \mu \exp(j\phi_m)$ is a deterministic and known complex constant, w_i is complex zero mean Gaussian noise with variance σ^2. Random variable Δ represents uncertainty in measuring the phase of the carrier. Its distribution can be described by PDF $p_\Delta(\Delta)$. In the following we assume that the phase uncertainty is described by the Von Mises (or Tikhonov) PDF (von Mises, 1964)

$$p_\Delta(\delta) = \frac{\exp\left[\kappa \cos(\Delta - \Delta_0)\right]}{2\pi I_0(\kappa)}. \tag{56}$$

Parameter Δ_0 represents bias in the determination of the carrier's phase, while κ represents quality of measurements. A few particular cases could be obtained from (56) by proper choice of parameters

1. Perfect phase recovery (coherent detection): $\kappa = \infty$, $\Delta_0 = 0$, and, thus, $p_\Delta(\Delta) = \delta(\Delta)$
2. No phase recovery (non-coherent detection): $\kappa = 0$ and, $p_\Delta(\Delta) = 1/2\pi$
3. Constant bias: $\kappa = \infty$, $\Delta_0 \neq 0$, $p_\Delta(\Delta) = \delta(\Delta - \Delta_0)$

We will derive the general expression first and then investigate particular cases to isolate effects of the parameters on performance of SPRT.

3.3.1 Average likelihood ratio

For a single observation z_i probability densities $p_1(z_i)$ and $p_0(z_i)$ corresponding to each of the hypothesis $\mathcal{H}_1$ and $\mathcal{H}_0$ are given by

$$p_1(z_i) = C \exp\left[-\frac{(x_i - \mu\cos(\phi_m + \Delta))^2}{\sigma^2}\right] \exp\left[-\frac{(y_i - \mu\sin(\phi_m + \Delta))^2}{\sigma^2}\right], \tag{57}$$

and

$$p_0(z_i) = C \exp\left[-\frac{x_i^2 + y_i^2}{\sigma^2}\right], \tag{58}$$

respectively Filio, Primak & Kontorovich (2011). For a given Δ, the likelihood ratio L_i could be easily calculated to be

$$L_i = \frac{p_1(z_i)}{p_0(z_i)} = \exp\left[\frac{2\mu\left(x_i\cos(\phi_m + \Delta) + y_i\sin(\phi_m + \Delta)\right) - \mu^2}{\sigma^2}\right]. \tag{59}$$

The conditional (on Δ) likelihood ratio $L(N|\Delta)$, considered over N observation is then just a product of likelihoods of individual observations, therefore

$$L(N|\Delta) = \prod_{n=1}^{N} \frac{p_1(z_n)}{p_0(z_n)} = \exp\left[\frac{2\mu \sum_{n=1}^{N} (x_n \cos(\phi_m + \Delta) + y_n \sin(\phi_m + \Delta)) - N\mu^2}{\sigma^2}\right]$$
$$= \exp\left[\frac{2\mu \mathcal{T}(N,\Delta)}{\sigma^2}\right] \exp\left[-\frac{N\mu^2}{\sigma^2}\right], \quad (60)$$

where

$$\mathcal{T}(N,\Delta) = \cos(\phi_m + \Delta) \sum_{n=1}^{N} x_n + \sin(\phi_m + \Delta) \sum_{n=1}^{N} y_n. \quad (61)$$

Let us introduce a new variables, $X(N)$, $Y(N)$, $Z(N)$ and $\Psi(N)$, defined by

$$X(N) = Z(N) \cos \Psi(N) = \sum_{n=1}^{N} x_n \quad (62)$$

$$Y(N) = Z(N) \sin \Psi(N) = \sum_{n=1}^{N} y_n \quad (63)$$

Using this notation equation (61) can now be rewritten as

$$\mathcal{T}(N,\Delta) = Z(N) \cos\left[\phi_m + \Delta - \Psi(N)\right]. \quad (64)$$

The average likelihood (Middleton, 1960) $L(N)$ could now be obtained by averaging (60) over the distribution of $p_\Delta(\Delta)$ to produce

$$\bar{L}(N) = \exp\left[-\frac{N\mu^2}{\sigma^2}\right] \int_{-\pi}^{\pi} \exp\left[\frac{2\mu Z(N) \cos\left[\phi_m + \Delta - \Psi(N)\right]}{\sigma^2}\right] p_\Delta(\Delta) d\Delta. \quad (65)$$

In turn, this expression could be further specialized if $p_\Delta(\Delta)$ is given by equation (56)

$$\bar{L}(N) = \exp\left[-\frac{N\mu^2}{\sigma^2}\right] \frac{1}{I_0(\kappa)} I_0\left[\sqrt{\frac{4\mu^2 Z^2(N)}{\sigma^4} + \frac{4\mu Z(N)\kappa}{\sigma^2} \cos\left[\phi_m - \Psi(N) - \Delta_0\right] + \kappa^2}\right]. \quad (66)$$

It can be seen from (66) that it is reduced to (45) if $\Delta_0 = 0$ and $\kappa = \infty$. Furthermore the deterministic phase bias Δ_0 could be easily eliminated from consideration by considering $\tilde{z}_i = z_i \exp[-j(\phi_m + \Delta_0)]$ instead of z_i. Therefore, equation (66) could be simplified to

$$\bar{L}(N) = \exp\left[-\frac{N\mu^2}{\sigma^2}\right] \frac{1}{I_0(\kappa)} I_0\left[\frac{1}{\sigma^2}\sqrt{4\mu^2 Y^2(N) + \left[2\mu X(N) + \kappa\sigma^2\right]^2}\right]. \quad (67)$$

In the case of non-coherent detection, *i.e.* in the case $\kappa = 0$, expression (67) assumes a well known form

$$\bar{L}(N) = \exp\left[-\frac{N\mu^2}{\sigma^2}\right] I_0\left[\frac{2\mu Z(N)}{\sigma^2}\right]. \quad (68)$$

Formation of the likelihood ratio could be considered as a two step process. As the first step, inphase and quadrature components independently accumulated to lessen the effect of AWGN. At the second step, values of $X(N)$ and $Y(N)$ must be combined in a fashion depending on available information. In the case of coherent reception it is know a priori that

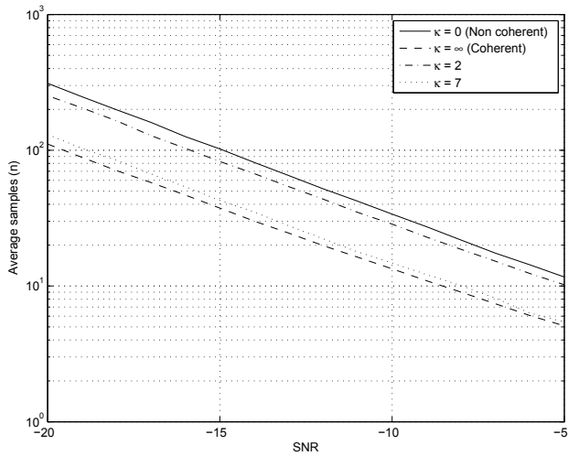


Fig. 10. Impact of coherency on the average number of samples.

the quadrature component $Y(N)$ contains only noise and it is ignored in the likelihood ratio. On the contrary, in the case of non-coherent reception one cannot distinguish between the in-phase and quadrature components and their powers are equally combined to for $Z(N)$. In the intermediate case both components are combined according to (67) with more and more emphasis put on in-phase component $X(N)$ as coherency increases with increase of κ.
In figure 10 we present the impact of non-coherent detection in the number of samples needed in order to detect a signal with respect to some P_D target. Notice that the main repercussion of non-coherence detection is the increase of samples to nearly twice in comparison to the coherent detector. In this terms, the non-coherent Wald sequential test procedure can be thought as having the same efficiency (in terms of number of samples) as the coherent NP-test.

3.4 Optimal fusion rule for distributed Wald detectors

This Section generalized results of Chair & Varshney (1986) to the case of distributed detection using Wal sequential analysis test. We assume that there are M sensors, making individual detection according to the Wald algorithm. Once a decision is made at an individual sensor the decision is send in the binary form to the Fusion Centre for further combining with other decisions. We assume that the value $u = -1$ is assigned if the hypothesis $\mathcal{H}_0$ is accepted, $u = 1$ if the hypothesis $\mathcal{H}_1$ is accepted and $u = 0$ if no decision has been made yet. Only $u = \pm 1$ are communicated to the Fusion Centre.
Since each node utilizes the Wald sequential detection Wald (2004), the decision is made at a random moments of time. Therefore, at any given moment of time t there is a random number $L(t) \leq M$ of decisions which are available at FC as can be seen in figure 11. Probability distribution of making decision could be approximated either by the two parametric Wald distribution Wald (2004), or by three parametric generalized inverse Gaussian distribution Jørgensen (1982) (as seen in fig.9). Parameters of those distributions could be found through moment/cumulant fitting, using expressions derived in Wald (2004), Filio, Kontorovich & Primak (2011). Following Chair & Varshney (1986) we would treat this problem as a two-hypothesis detection problem with individual detector decision being the observation. For a given number $L = L(t)$ of decisions made by the time t, the optimum decision rule is equivalent to the following likelihood ratio test

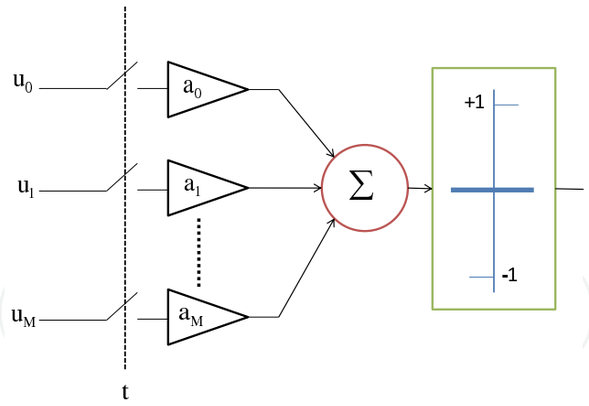


Fig. 11. System model of data fusion system

$$\frac{P\left(u_1,u_2,\cdots,u_L|L,\mathcal{H}_1\right)P(L|\mathcal{H}_1)}{P\left(u_1,u_2,\cdots,u_L|L,\mathcal{H}_0\right)P(L|\mathcal{H}_0)}\underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}}\frac{P_0(C_{10}-C_{00})}{P_1(C_{01}-C_{11})}. \tag{69}$$

Here $P(L|\mathcal{H}_0)$ is the probability of making exactly L decisions assuming that $\mathcal{H}_1$ is true and u_l is the decision made by l-th sensor.
Furthermore, assuming the minimum probability of error criteria, *i.e.* by setting $C_{00}=C_{11}=0$ and $C_{01}=C_{10}=1$, introducing notation $\mathbf{u}_L=\{u_1,u_2,\cdots,u_L\}$ and using the Bayes rule one can recast equation (69) as

$$\frac{P(\mathcal{H}_1|\mathbf{u}_L,\mathbf{l})P(\mathbf{l}|\mathcal{H}_1)}{P(\mathcal{H}_0|\mathbf{u}_L,\mathbf{l})P(\mathbf{l}|\mathcal{H}_0)}\underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}}1, \tag{70}$$

or, after taking natural log of both side

$$\ln\frac{P(\mathcal{H}_1|\mathbf{u}_L,\mathbf{l})}{P(\mathcal{H}_0|\mathbf{u}_L,\mathbf{l})}+\ln\frac{P(\mathbf{l}|\mathcal{H}_1)}{P(\mathbf{l}|\mathcal{H}_0)}\underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}}0, \tag{71}$$

where $\mathbf{l}$ is a vector which represents which sensors have made decisions.
Once again following Chair & Varshney (1986), one can calculate probabilities $P(\mathcal{H}_1|\mathbf{u}_L,\mathbf{l})$ and $P(\mathcal{H}_0|\mathbf{u}_L,\mathbf{l})$ as follows. In the case of the hypothesis $\mathcal{H}_1$ one can write

$$P(\mathcal{H}_1|\mathbf{u}_L,\mathbf{l})=\frac{P(\mathcal{H}_1,\mathbf{u}_L|\mathbf{l})}{P(\mathbf{u}_L|\mathbf{l})}=$$
$$\frac{P_1}{P(\mathbf{u}_L|\mathbf{l})}\prod_{S_+}P(u_l=+1|\mathcal{H}_1)\prod_{S_-}P(u_l=-1|\mathcal{H}_1)=\frac{P_1}{P(\mathbf{u}_L|\mathbf{l})}\prod_{S_+}(1-P_{M,l})\prod_{S_-}P_{M,l}, \tag{72}$$

where S_+ is the set of all i such that $u_i=+1$ and S_- is the set of all i such that $u_i=-1$.
Similarly, in the case of the hypothesis $\mathcal{H}_0$ one obtains

$$P(\mathcal{H}_0|\mathbf{u}_L,\mathbf{l})=\frac{P(\mathcal{H}_0,\mathbf{u}_L|\mathbf{l})}{P(\mathbf{u}_L|\mathbf{l})}=\frac{P_0}{P(\mathbf{u}_L|\mathbf{l})}\prod_{S_-}(1-P_{F,l})\prod_{S_+}P_{F,l}. \tag{73}$$

Finally, using equations (72) and (73) one obtains the following expression for the conditional log-likelihood

$$\ln\frac{P(\mathcal{H}_1|\mathbf{u}_L,\mathbf{l})}{P(\mathcal{H}_0|\mathbf{u}_L,\mathbf{l})}=\ln\frac{P_1}{P_0}+\sum_{S_+}\ln\frac{1-P_{M,l}}{P_{F,l}}+\sum_{S_-}\ln\frac{P_{M,l}}{1-P_{F,l}}. \tag{74}$$

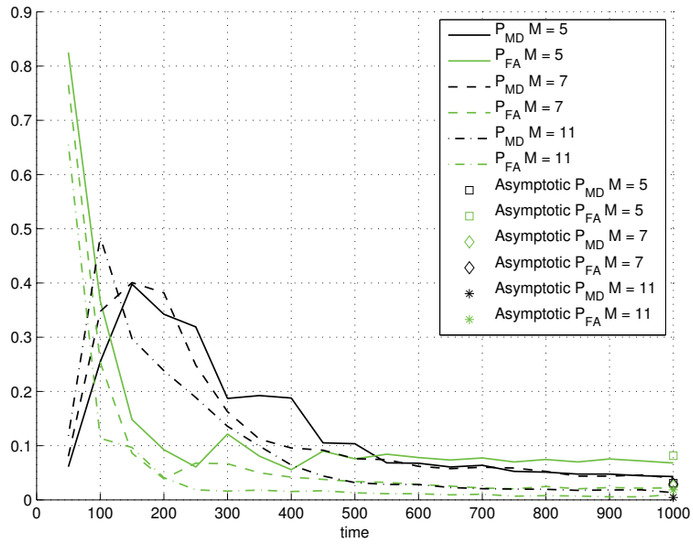


Fig. 12. Data fusion scheme considering sequential analysis decision from each sensor $(P_{M,l} = 0.3, P_{F,l} = 0.1)$.

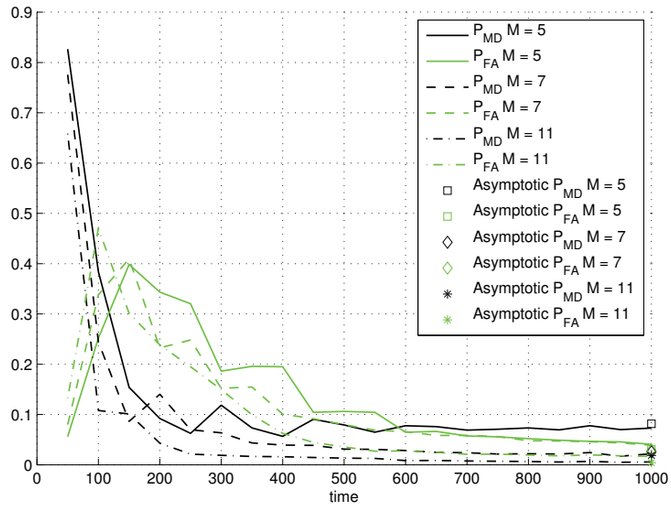


Fig. 13. Data fusion scheme considering sequential analysis decision from each sensor $(P_{M,l} = 0.1, P_{F,l} = 0.3)$.

In order to evaluate the second term in the sum (71) let us first consider an arbitrary node $1 \leq k \leq N$. Distribution $p_{T,k}(\tau)$ of the decision time at such a node is assumed to be known. Therefore, probability $P_{D,k}(t|\mathcal{H}_i)$ that the decision is made by the time $t = mT_s$ given that the hypothesis is true is given by

$$P_{D,k}(t|\mathcal{H}_i) = \int_0^{mT_S} p_k(\tau|\mathcal{H}_i)d\tau. \tag{75}$$

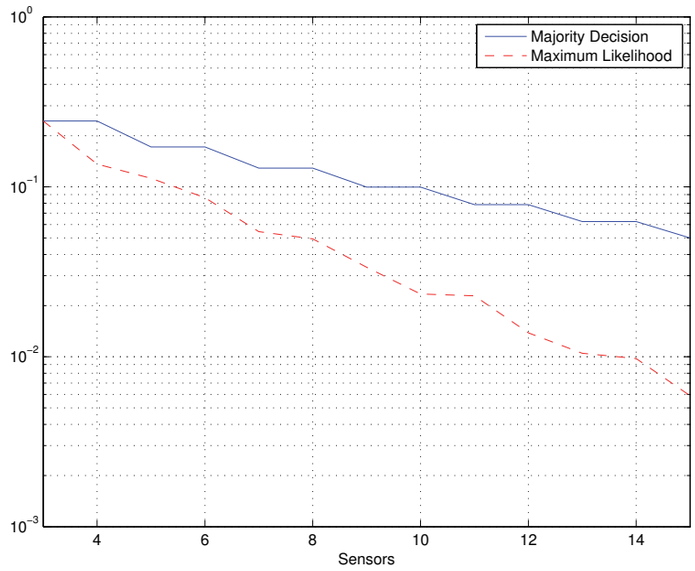


Fig. 14. Total Error Criteria

The probability that no decision has been made by the time t is then simply $1 - P_{D,k}(t|\mathcal{H}_i)$. As it has been mentioned earlier, parameters of this distribution could be defined as in Filio, Kontorovich & Primak (2011). Therefore , the second term in the equation (71).

$$\ln \frac{P(\mathbf{l}|\mathcal{H}_1)}{P(\mathbf{l}|\mathcal{H}_0)} = \sum_{l=1}^{L} \ln \frac{P_{D,l}(t|\mathcal{H}_1)}{P_{D,l}(t|\mathcal{H}_0)} + \sum_{l=L+1}^{M} \ln \frac{1 - P_{D,l}(t|\mathcal{H}_1)}{1 - P_{D,l}(t|\mathcal{H}_0)}. \tag{76}$$

Finally, the fusion rule in the case of nodes making decision according to Wald's criteria could be written as

$$f(\mathbf{u}) = \begin{cases} 1 & \text{if } a_0 + \sum_{l=1}^{L} a_l u_l > 0 \\ -1 & \text{otherwise} \end{cases} \tag{77}$$

where

$$a_0 = \ln \frac{P_1}{P_0} + \sum_{l=1}^{L} \ln \frac{P_{D,l}(t|\mathcal{H}_1)}{P_{D,l}(t|\mathcal{H}_0)} + \sum_{l=L+1}^{M} \ln \frac{1 - P_{D,l}(t|\mathcal{H}_1)}{1 - P_{D,l}(t|\mathcal{H}_0)}, \tag{78}$$

$$a_l = \ln \frac{1 - P_{M,l}}{P_{F,l}} \quad \text{if } u_l = 1, \tag{79}$$

$$a_l = \ln \frac{1 - P_{F,l}}{P_{M,l}} \quad \text{if } u_l = -1. \tag{80}$$

Thus, the combining rule is similar to that suggested in Chair & Varshney (1986), however, with some significant differences in the term of a_0. In figures 12 and 13 it is shown the performance of the data fusion scheme considering that each one of the sensors takes a decision based on the sequential detection criteria. In these figures it is plotted the probability of missdetection (P_{MD}) and the probability of false alarm (P_{FA}) versus the time in which the fusion centre gathers the decisions taken from the local observers. Notice that for $t \to \infty$ all graphs converge to the data fusion rule of Chair & Varshney (1986) for which the theoretical

expression for probability of missdetection and false alarm is derived in appendix 5. It is obvious that for small values of time, the fusion centre has less information (since not all the detectors might achieve a decision by then) and the final decision it takes is way less accurate that for large values of time. Nevertheless, in some practical systems it would be impossible to wait that long for getting the decision from the fusion centre, so we can use these results as a trade-off between the performance on the detection and the time of decision Gosan et al. (2010). Next thing it is possible to observe is the impact that $P_{D,l}$ and $P_{F,l}$ has on the performance of the data fusion detector.
Notice that for very small values of false alarm probability, $a_l \approx -\ln P_{M,l}$ if $u = -1$ in eq. (78) which means that hypothesis $\mathcal{H}_0$ is always less weighted in equation (77) or in other words, the fusion centre "trusts" more in those sensors who decide that $\mathcal{H}_1$ is true. For very small values of missdetection probability a similar thing occurs but in this case the hypothesis $\mathcal{H}_0$ is more weighted in the final sum in equation (77). A special case occurs when $P_{D,l} = P_{F,l}$, $P_0 = P_1$ and $t \to \infty$. In this situation, the scheme converts into the more simple majority decision approach, which just sums all u_l and compares with zero. Even though its simplicity, the maximum likelihood approach performs better than the majority decision scheme in the minimum probability of error criteria as can be seen in figure 14 Filio et al. (2010). The perceptive reader must have noticed by now that there might be some confusion at the fusion centre when there exists an even number of sensors and there is a tie in the decision. This can be settled by considering the a priori probabilities P_0 and P_1 which are inherent to the system.

4. Conclusion

In the first part of this chapter, we have investigated the impact the scattering environment on the performance of primary user detection in multiantennae confinguration. An approximate expressions for the probability of missed detection in function of the number of antennas, parameters of the scattering environment and number of observations. It is shown that at low SNR it is more beneficial to utilize just a single antenna and large number of time samples. This allowes for a better noise suppression via time averaging. On the contrary, at high SNR, it is more beneficial to have more antennas in order to mitigate fading which is a dominant cause of errors in detection in weak noise. It was also shown that for a very strong correlation, *i.e.* $\rho \approx 1$, the equivalent number of antennas is almost unity, $N_{eq} \approx 1$. Therefore this scheme can be usefully applied in a low SNR situations assuming enough time samples are obtained. As $\rho \approx 0$, the diversity gain is increased, therefore making it suitable for higher SNR situations.
The second part of the chapter was devoted to application of the sequential analysis technique to a chive a faster spectrum sensing in cognitive radio networks. It was shown that using the sequential probability ratio test it is possible to detect the presence of a primary user almost twice as fast as other fixed sample approaches such as Neyman Pearson detectors. This can be achieved when dealing in the low SNR region which is a quite often operating characteristic in real life. The effect of error in estimation phase of the carrier on the duration of the sequential analysis has also been investigated. It was shown that the impact of non-coherent detection in sensing the presence of primary users using sequential analysis is the increase of almost twice the samples needed in comparison to a coherent detection approach. Afterwards we derived an optimal fusion rule using detectors that use sequential analysis for taking decisions. We assessed the performance of the system in terms of the time that it takes to gather the decision from all detectors. It was shown that for faster decision, the fusion centre does not consider the opinions of all sensors and hence the performance gets reduced. On the other hand, as

the time of decision increases the performance is better but the system experiment a higher latency.

5. Appendix

Performance derivation of data fusion rule

Let us introduce the following notations

$$a_i = \begin{cases} \ln \frac{1-P_{MD}}{P_{FA}} & \text{if } u_i > 0 \\ \ln \frac{1-P_{FA}}{P_{MD}} = -\ln \frac{P_{MD}}{1-P_{FA}} & \text{if } u_i < 0 \end{cases}, \tag{81}$$

and

$$\xi_i = \begin{cases} a_i = \ln \frac{1-P_{MD}}{P_{FA}} & \text{if } u_i > 0 \\ b_i = -a_i = -\ln \frac{1-P_{FA}}{P_{MD}} & \text{if } u_i < 0 \end{cases}. \tag{82}$$

Consider a $\mathcal{T}$ (test statistic) given by eq. (77)

$$\mathcal{T} = a_0 + \sum_{k=1}^{K} a_k u_k = a_0 + \sum_{k=1}^{K} \xi_k |u_k| = a_0 + \sum_{k=1}^{K} \xi_k. \tag{83}$$

Here ξ_k could be considered as a random variable with PDF

$$\begin{aligned} P_{\xi}(x) = & \quad P_+ \delta(x-a) + P_- \delta x - b \\ = & \; P\delta(x-a) + (1-P)\delta(x-b), \end{aligned} \tag{84}$$

where

$$\begin{aligned} a = & \;\; a_i \;\; = \ln \frac{1-P_{MD}}{P_{FA}} \\ b = & -a_i = -\ln \frac{1-P_{FA}}{P_{MD}} \end{aligned}. \tag{85}$$

and P_+ is probability of $u = +1$ decision, equal to

$$\begin{aligned} P_+ = & \quad p(\mathcal{H}_1)(1-P_{MD}) + p(\mathcal{H}_0)P_{FA} \\ = & \; p(\mathcal{H}_1)(1-P_{MD}) + [1-p(\mathcal{H}_1)]\, P_{FA}, \end{aligned} \tag{86}$$

The corresponding characteristic function of ξ is then given by

$$\Theta_{\xi}(s) = P_+ e^{-sa} + P_- e^{-sb}, \tag{87}$$

and the characteristic function of $\mathcal{T}$ could be evaluated as

$$\begin{aligned} \Theta_{\mathcal{T}} = \Theta_{\xi}^{K} e^{-sa_0} = & \quad \left[P_+ e^{-sa} + (1-P_+)e^{-sb}\right]^K e^{-sa_0} \\ = & \; \textstyle\sum_{k=0}^{K} \binom{K}{k} P_+^k (1-P_+)^{K-k} e^{-s[ka+(K-k)b+a_0]} \end{aligned}. \tag{88}$$

Equivalently, the PDF is given by

$$P_{\mathcal{T}}(x) = \sum_{k=0}^{K} \binom{K}{k} P_+^k (1-P_+)^{K-k} \delta[x - (ka + (K-k)b + a_0)]. \tag{89}$$

If $k = 0$ then

$$ka + (K-k)b - a_0 = Kb + a_0$$

$$= -K \ln \frac{1 - P_{FA}}{P_{MD}} + \ln \frac{P(\mathcal{H}_1)}{1 - P(\mathcal{H}_1)}. \tag{90}$$

If $P(\mathcal{H}_1) \approx 1$ such that

$$\frac{P(\mathcal{H}_1)}{1 - P(\mathcal{H}_1)} > \left(\frac{1 - P_{FA}}{P_{MD}} \right)^K, \tag{91}$$

then the FC makes only $\mathcal{H}_1$ decisions i.e.

$$P_{MD} = 0, \qquad P_{FA} = P(\mathcal{H}_0) = 1 - p(\mathcal{H}_1).$$

If (91) is not satisfied then there is $k_{\max} > 0$ such that

$$k_{\max} a + (K - k_{\max}) b + a_0 < 0, \tag{92}$$

and

$$(k_{\max} + 1) a + (K - k_{\max} - 1) b + a_0 > 0. \tag{93}$$

In this case the scheme suggested in Chair & Varshney (1986) is equivalent to $(k_{\max} + 1)$ out of K scheme (this is assuming that are statistically equivalent).
Let $\mathcal{H}_1$ be true. Then the target is missed if there are no more than $k_{\max}$ positive decisions, or, equivalently, no less than $K - k_{\max}$ negative decisions. The probability of miss detection at FC is then given by

$$P_{MD_F} = \sum_{k=0}^{k_{\max}} \binom{K}{k} P_{MD}^k (1 - P_{MD})^{K-k}. \tag{94}$$

To more decisions $\mathcal{H}_1$ there should be at least $k_{\max} + 1$ partial 1. If $\mathcal{H}_0$ is true, the probability of false alarm at the fusion center is then:

$$\begin{array}{l} \mathcal{H}_1 : P_{MD_F} = \sum_{k=0}^{k_{\max}} \binom{K}{k} (1 - P_{MD})^k P_{MD}^{K-k} \\ \mathcal{H}_0 : P_{FA_F} = \sum_{k=k_{\max}}^{K} \binom{K}{k} (P_{FA})^k (1 - P_{FA})^{K-k}. \end{array} \tag{95}$$

6. References

Abdi, A. & Kaveh, M. (2002). A space-time correlation model for multielement antenna systems in mobile fading channels, 20(3): 550–560.

Akyildiz, I., Lee, W., Vuran, M. & Mohanty, S. (2008). A Survey on Spectrum Management in Cognitive Radio Networks, 46(4): 40–48.

Almalfouh, S. & Stuber, G. (2010). Interference-aware power allocation in cognitive radio networks with imperfect spectrum sensing, *Proc. of ICC 2010*, pp. 1 –6.

Andronov, I. & Fink, L. (1971). *Peredacha diskretnykh soobshchenii po parallel'nym kanalam*, Sov. Radio, Moscow.

Chair, Z. & Varshney, P. (1986). Optimal data fusion in multiple sensor detection systems, 22(1): 98 –101.

Chenhui, H., Xinbing, W., Zichao, Y., Jianfeng, Z., Youyun, X. & Xinbo, G. (2010). A geometry study on the capacity of wireless networks via percolation, 58(10): 2916 – 2925.

Editors, S. (2006). UMTS: Spatial channel model for MIMO simulations, *Tech. report 25.996*.

Federal Communications Commission (November). Spectrum policy task force, *Technical report*, FCC.

Filio, O., Kontorovich, V. & Primak, S. (2011). Characteristics of sequencial detection in cognitive radio networks, *Proc. ICACT 2011*, Phoenix Park, Korea.

Filio, O., Kontorovich, V., Primak, S. & Ramos-Alarcon, F. (2010). Collaborative spectrum sensing for cognitive radio: Diversity combining approach, *Wireless Sensor Network* 3(1): 24–37.

Filio, O., Primak, S. & Kontorovich, V. (2011). On performance of wald test in partially coherent channels, *Proc. ICCIT 2011*, Aqaba, Jordan.

Gosan, N., Jemin, L., Hano, W., Sungtae, K., Sooyong, C. & Daesik, H. (2010). Throughput analysis and optimization of sensing-based cognitive radio systems with markovian traffic, 59(8): 4163–4169.

Haghighi, S. J., Primak, S., Kontorovich, V. & Sejdic, E. (2010). *Mobile and Wireless Communications Physical layer development and implementatiom*, InTech Publishing, chapter Wireless Communications and Multitaper Analysis: Applications to Channel Modelling and Estimation.

Haykin, S. (2005). Cognitive radio: brain-empowered wireless communications, 23(2): 201–220.

Haykin, S., Thomson, D. & Reed, J. (2009). Spectrum sensing for cognitive radio, *Proceedings of the IEEE* 97(5): 849 –877.

Jørgensen, B. (1982). *Statistical Properties of the Generalized Inverse Gaussian Distribution. Lecture Notes in Statistics. Vol. 9*, Springer-Verlag, New York, Berlin.

Kang, H., Song, I., Yoon, S. & Kim, Y. (2010). A class of spectrum-sensing schemes for cognitive radio under impulsive noise circumstances: Structure and performance in nonfading and fading environments, 59(9): 4322 – 4339.

Kay, S. (1998). *Statistical Signal Processing: Detection Theory*, Prentice Hall PTR, Upper Saddle River, NJ.

Kontorovich, V., Ramos-Alarcón, F., Filio, O. & Primak, S. (2010). Cyclostationary spectrum sensing for cognitive radio and multiantenna systems, *Wireless Communications and Signal Processing (WCSP), 2010 International Conference on*, pp. 1 –6.

Kotz, S. & Adams, J. W. (1964). Distribution of sum of identically distributed exponentially correlated gamma-variables, *The Annals of Mathematical Statistics* 35(1): 277–283.

Li, H. & Han, Z. (2010a). Catch me if you can: An abnormality detection approach for collaborative spectrum sensing in cognitive radio networks, 9(11): 3554 – 3565.

Li, H. & Han, Z. (2010b). Dogfight in spectrum: Combating primary user emulation attacks in cognitive radio systems, part i: Known channel statistics, 9(11): 3566 – 3577.

Lin, S.-C., Lee, C.-P. & Su, H.-J. (2010). Cognitive Radio with Partial Channel State Information at the Transmitter, 9(11): 3402–3413.

Middleton (1960). *Introduction to Statistical Communications Theory*, first edn, McGraw-Hill, New York.

Mitola, J. & Maguire, G. (1999). Cognitive radio: making software radios more personal, 6(4): 13–18.

Mitran, P., Le, L. B. & Rosenberg, C. (2010). Queue-aware resource allocation for downlink ofdma cognitive radio networks, 9(10): 3100 – 3111.

Paulraj, A., Nabar, R. & Gore, D. (2003). *Introduction to Space-Time Wireless Communications*, Cambridge University Press, Cambridge, UK.

Poor, H. & Hadjiliadis, O. (2008). *Quickest detection*, Cambridge University Press.

Rui, Z. (2010). On active learning and supervised transmission of spectrum sharing based cognitive radios by exploiting hidden primary radio feedback, 58(10): 2960 – 2970.

Shengli, X., Yi, L., Yan, Z. & Rong, Y. (2010). A parallel cooperative spectrum sensing in cognitive radio networks, 59(8): 4079 – 4092.

Simon, M. & Alouini, M.-S. (2000). *Digital Communication over Fading Channels:A Unified Approach to Performance Analysis*, John Wiley & Sons, New York.

Slepian, D. (1978). Prolate spheroidal wave functions, Fourier analysis, and uncertainty. V-The discrete case, *Bell System Technical Journal* 57: 1371–1430.

Song, M.-G., Kim, D. & Im, G.-H. (2010). Recursive channel estimation method for ofdm-based cooperative systems, 14(11): 1029 – 1031.

Sun, H., Laurenson, D. & Wang, C.-X. (2010). Computationally tractable model of energy detection performance over slow fading channels, 14(10): 924 –926.

Tang, P. K. & Han, C. (2010). On the modeling and performance of three opportunistic spectrum access schemes, 59(8): 4070 – 4078.

Thomson, D. (1982). Spectral estimation and harmonic analysis, 70(9): 1055–1096.

van Trees, H. (2001). *Detection, Estimation, and Modulation Theory: Part I*, first edn, John Wiley & Sons, New York.

von Mises, R. (1964). *Mathematical theory of probability and statistics*, Academic Press, New York.

Wald, A. (2004). *Sequential Analysis*, 6th edn, Dover Phoenix, Mineola, NY.

Wang, R. & Meixia, T. (2010). Blind spectrum sensing by information theoretic criteria for cognitive radios, 59(8): 3806 – 3817.

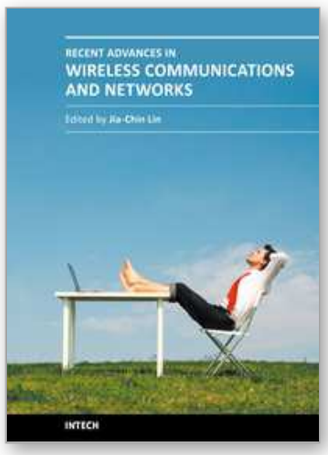
Wenyi, Z. (2010a). Spectrum shaping: A new perspective on cognitive radioŮpart ii: Coexistence with uncoded legacy transmission, 58(10): 2971 – 2983.

Wenyi, Z. M. (2010b). Spectrum shaping: a new perspective on cognitive radio-part i: coexistence with coded legacy transmission, 58(6): 1857 – 1867.

Yücek, T. & Arslam, H. (2009). A survey of spectrum sensing algorithms for cognitive radio applications, *IEEE Comm. Surveys & Tutorials* 11(1): 116–130.

Zhang, R., Lim, T. J., Liang, Y.-C. & Zeng, Y. (2010). Multi-antenna based spectrum sensing for cognitive radios: A GLRT approach, 58(1): 84–88.

Zou, Q., Zheng, S. & Sayed, A. (2010). Cooperative sensing via sequential detection, 58(12): 6266 –6283.



Recent Advances in Wireless Communications and Networks
Edited by Prof. Jia-Chin Lin

ISBN 978-953-307-274-6
Hard cover, 454 pages
Publisher InTech
Published online 23, August, 2011
Published in print edition August, 2011

This book focuses on the current hottest issues from the lowest layers to the upper layers of wireless communication networks and provides â€œreal-timeâ€ research progress on these issues. The authors have made every effort to systematically organize the information on these topics to make it easily accessible to readers of any level. This book also maintains the balance between current research results and their theoretical support. In this book, a variety of novel techniques in wireless communications and networks are investigated. The authors attempt to present these topics in detail. Insightful and reader-friendly descriptions are presented to nourish readers of any level, from practicing and knowledgeable communication engineers to beginning or professional researchers. All interested readers can easily find noteworthy materials in much greater detail than in previous publications and in the references cited in these chapters.

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen