Cluster-Based Differential Energy Detection for Spectrum Sensing in Multi-Carrier Systems

Parisa Cheraghi, *Student Member, IEEE*, Yi Ma, *Senior Member, IEEE*, Rahim Tafazolli, *Senior Member, IEEE*, and Zhengwei Lu, *Student Member, IEEE*

Abstract—This paper presents a novel differential energy detection scheme for multi-carrier systems, which can form fast and reliable decision of spectrum availability even in very low signal-to-noise ratio (SNR) environment. For example, the proposed scheme can reach 90% in probability of detection (PD) and 10% in probability of false alarm (PFA) for the SNRs as low as -21 dB, while the observation length is equivalent to 2 multi-carrier symbol duration. The underlying initiative of the proposed scheme is applying order statistics on the clustered differential energy-spectral-density (ESD) in order to exploit the channel frequency diversity inherent in high data-rate communications. Specifically, to enjoy a good frequency diversity, the clustering operation is utilized to group uncorrelated subcarriers, while, the differential operation applied onto each cluster can effectively remove the noise floor and consequently overcome the impact of noise uncertainty while exploiting the frequency diversity. More importantly, the proposed scheme is designed to allow robustness in terms of both, time and frequency offsets. In order to analytically evaluate the proposed scheme, PFA and PD for Rayleigh fading channel are derived. The closed-form expressions show a clear relationship between the sensing performance and the cluster size, which is an indicator of the diversity gain. Moreover, we are able to observe up to 10 dB gain in the performance compared to the state-of-the-art spectrum sensing schemes.

Index Terms—Differential, energy detection, low signal-to-noise ratio (SNR), multi-carrier, spectrum sensing.

I. INTRODUCTION

OPPORTUNISTIC SPECTRUM ACCESS (OSA) [1], first coined by Mitola *et al.* [2] under the term "spectrum pooling" in cognitive radio terminology, promises tremendous gain in improving spectral efficiency. The main objective of OSA is to offer the ability of identifying and exploiting the under-utilized spectrum in an instantaneous manner in a wireless device, without any user intrusion. This allows the wireless devices to rapidly change their modulation scheme and communication protocol so as to better and more efficient communication. The initial requirement of any OSA device is to determine the spectrum availability. There are three possible solutions for monitoring the spectrum availability proposed

in the literature: 1) through an ubiquitous connection to the database, 2) a dedicated standardised channel to broadcast a beacon signal, 3) spectrum sensing [3]. Recently, Federal Communication Committee (FCC) [4] considered database connection for inclusion in the IEEE 802.22 standard [5]. However, it has been shown in [6] that the geo-location database solution might incur additional costs, e.g., signalling overhead, scheduling complexity and database maintenance costs. Nevertheless, spectrum sensing, thanks to its relatively low infrastructure cost, still receives more and more attention. Therefore, proposed by the Office of Communications (Ofcom) [7] a complementary application of both spectrum sensing and database connection can provide a practical solution for enabling spectrum availability monitoring.

Spectrum sensing is a traditional topic in the scope of signal processing for mobile communications. It is quite mature yet for carrier sense multiple access (CSMA) [8] based random access networks, i.e., "listen before talk" communication protocols. New challenges and problems arise for spectrum sensing in flexible networks, e.g., cognitive radio, where it is required to meet the following three requirements.

- **Fast spectrum sensing.** Since it is not possible to transmit and sense simultaneously at a particular frequency band, sensing has to be interleaved with data transmission. Hence, the required observation time (or window) should be as short as possible in order to maximise the overall throughput.
- **High accuracy.** The spectrum sensing device identifies vacant spectrum bands by detecting presence of primary signal, i.e., by performing a binary hypothesis test. With noise and fading available in any communication system, sensing errors are inevitable. For example, false alarm occurs when an idle channel is detected as busy, and miss detection occurs when an occupied channel is declared as idle. In the occasion of a false alarm a transmission opportunity is overlooked, resulting in waste of the spectrum, while miss detection leads to collision with primary users and hence, interference.
- **Low complexity.** The computational complexity of the sensing device should be kept as low as possible in order to reduce the signal processing time, device energy consumption as well as the infrastructure cost.

In addition, it is becoming increasingly demanding for delivering reliable spectrum sensing in very low signal-to-noise ratio (SNR) range, such that heavily shadowed signals can be identified.

A. Related Work

Spectrum sensing has been quite well investigated in the literature, for which there are many approaches reported. In general the existing spectrum sensing approaches can be divided into three main categories:

- **Exploiting energy difference.** The most well-known spectrum sensing approach developed under this category is the energy detection [9]. The energy detection is recognized as a blind sensing scheme with advantages such as low complexity and low latency. However, it is very sensitive to the noise uncertainty such that its performance is limited by the SNR wall [10]. In the last four decades since the publication of [9], many solutions have been developed to make energy detection more robust in terms of SNR wall (e.g., [11]-[13]), yet the noise uncertainty problem in spectrum sensing approaches based on the energy difference still exists.
- **Exploiting stationarity difference.** The initial works of spectrum sensing through stationarity difference can be traced back to work of Dandawate *et al.* in [14], where second order cyclostionarity is employed. The cyclostationarity based scheme can trade latency with high sensing reliability. It is less sensitive to the noise uncertainty, provided the knowledge of the signals cyclic frequency [15]. The wavelet scheme is able to perform wide-band sensing with the aid of edge detection [16]. It is particularly useful for fast coarse spectrum sensing based on a number of non-stationary samples, hence, making use of the signals non-stationarity features. Fine spectrum sensing is further required in order to determine the vacancy of specified frequency bands. In addition, the Wigner-Ville based spectrum sensing [17] derives a greyscale image of the time-frequency description of the received signal through the Wigner-Ville transform and, similar to wavelet based detection, with the aid of edge detection is able to detect occupied frequency bands. Matched-filtering pilot based detection, given the knowledge of pilot symbols, reasonably good timing and frequency synchronizations, exploits the cyclostationary property of the pilot symbols to deliver fast and reliable sensing [18]. Furthermore, the eigenvalue-based detection scheme exploits orthogonality between the signal subspace and noise subspace using second order stationarity features to offer highly reliable spectrum sensing [19]. However, it often needs the support of multiple antennas, and the subspace decomposition costs cubic complexity.
- **Exploiting the distribution difference.** Given that in almost all communication system models, noise is assumed to be additive, white and Gaussian, one can determine the vacancy of a particular frequency band by observing the difference of the received signals distribution and that of the additive white Gaussian noise (AWGN). An example of such approaches would be the kurtosis-type scheme, which exploits the non-Gaussianity of communication signals [20] [21]. This scheme features excellent accuracy at the price of large latency, due to higher-order statistics. A critical point is that the sensing performance degrades

significantly when signals are approximately Gaussian. Entropy-based spectrum sensing can be thought of as an approach which also benefits from this property [22]. The probability space is partitioned into fixed dimensions and the Shannon entropy is employed as the information measure of the received signal, acting as the test statistic.

A complete survey on the existing spectrum sensing techniques can be found in [23]-[25].

It is realized that existing schemes can hardly meet the requirements of a fast and accurate spectrum sensing particularly in very low SNR range, (considering that the target SNR for a reliable spectrum sensing sensitivity is about -20 dB [5]) without introducing high complexity to the system. This observation motivates us to develop a new local spectrum sensing scheme, which can significantly improve the state-of-the-arts and provides a practical solution.

B. Contribution

The key contribution presented in this paper is a novel spectrum sensing scheme namely, cluster-based differential energy detection. It has several distinctive features including low latency, high accuracy reasonable computational complexity, as well as robustness to very low SNR. For example, the proposed scheme can reach 90% in probability of detection (PD) and 10% in probability of false alarm (PFA) for an SNR as low as -21 dB, while the observation window is equivalent to 2 multi-carrier symbol duration. The proposed scheme at this stage is specially designed for sensing multi-carrier sources but we would argue that most of the current and future mobile networks are multi-carrier based systems, and thus it has a wide implication for practical applications.

The key idea of the proposed scheme is to exploit the channel frequency diversity inherent in high data-rate communications using the clustered differential ordered energy spectral density (ESD). Specifically, after the ESD computation, the clustering operation is utilized to group uncorrelated subcarriers based on the coherence bandwidth to enjoy a good frequency diversity. The knowledge of coherence bandwidth does not need to be very accurate (in this paper we employ the reciprocal of the maximal channel delay). Furthermore, making use of order statistics of the estimated ESD, we further increase the reliability of the sensing algorithm.

In order to exploit the second order moment diversity of the observed signal, a differential operation is performed on the rank ordered ESD. When the channel is frequency selective and the noise is white, the differential process can effectively remove the noise floor resulting in elimination of the noise uncertainty impact which is the main factor making energy detection reluctant [10]. At the final stage of the proposed scheme, the differential rank ordered ESD within different clusters are linearly combined in order to further reduce the effect of impulse/spike noise. Binary hypothesis testing is then applied on either the maximum or the extremal quotient (maximum-to-minimum ratio) depending on the wireless channel characteristics of the sensed environment. More importantly, the proposed spectrum sensing scheme is designed to allow robustness in terms of both, time and frequency offset.

In order to analytically evaluate the proposed scheme, both PD and PFA are derived for Rayleigh fading channels. The closed-form expression shows a clear relationship between the sensing performance and the cluster size, i.e., channel coherence bandwidth, which is an indicator of the diversity gain. Computer simulations are carried out in order to evaluate the effectiveness of the proposed approach and to compare the performance of the proposed scheme with state-of-the-art spectrum sensing schemes where up to 10 dB gain in performance can be observed.

The rest of this paper is organised as follows, in Section II, the system model and problem formulation for practical multi-carrier communication environment are provided. The proposed differential energy detection scheme is presented in Section III along with possible solutions for overcoming physical impairments and further discussions. The theoretical analysis of the proposed technique in terms of two classical metrics, PFA and PD, are provided in section IV followed by numerical results and computational complexity. Section V presents the simulation results, and finally conclusions are drawn in Section VI.

II. SYSTEM MODEL AND PROBLEM FORMULATION

A. Multi-Carrier Systems

1) Transmitted Signal: A general framework of multi-carrier systems has been presented in [26]. The transmitted signal can be expressed in the matrix form, $\mathbf{x}_k = \mathbf{\Psi}\mathbf{s}_k$, where $\mathbf{x}_k$ is an $J \times 1$ transmitted signal block, $\mathbf{s}_k$ is an $M \times 1$ information-bearing symbol block with the covariance $\sigma_s^2 \mathbf{I}_M$ (M stands for the number of subcarriers, and $\mathbf{I}_M$ for the identity matrix of size M), $\mathbf{\Psi}$ is an $J \times M$ ($J > M$) tall pre-coding matrix with full column rank, and subscript k is the block index. There are two conventional approaches for implementing the pre-coding matrix $\mathbf{\Psi}$, i.e.,

$$\text{CP}: \ \mathbf{\Psi} \triangleq \begin{bmatrix} \mathcal{F}_{\text{cp}}^H & \mathcal{F}^H \end{bmatrix}^T \mathbf{\Phi} \tag{1}$$

$$\text{ZP}: \ \mathbf{\Psi} \triangleq \begin{bmatrix} \mathcal{F}^H & \mathbf{0} \end{bmatrix}^T \mathbf{\Phi} \tag{2}$$

where $\mathcal{F}$ is the $M \times M$ normilized discrete Fourier transform (DFT) matrix [27], $\mathcal{F}_{\text{cp}}$ in (2) is formed by collecting the last $(J - M)$ columns of $\mathcal{F}$, $\mathbf{\Phi}$ is an $M \times M$ full rank matrix. This paper is focused on the cyclic prefix (CP) based system since it has been widely deployed in practical networks due to its advantages, e.g., eliminating inter-symbol interference and handling multi-path channels [28]. Nevertheless, it is shown in Section III-C how the proposed spectrum sensing scheme can be easily extended to the zero-padding (ZP) based system in (2).

2) Signal Analysis at the Sensing Device: Consider a wireless device sensing a particular frequency band, in the absence of the multi-carrier signal, the device can only receive noise, otherwise, it receives a signal distorted by the frequency-selective channel (denoted by h), timing offset (denoted by ϵ), frequency offset normalized by the subcarrier spacing (denoted by ε), and additive white Gaussian noise (denoted by v). Indeed, there are many other distortions such as phase noise

and non-linear distortions due to imperfect electro components [28]. But in this paper, we will focus on those major physical distortions, (i.e., frequency selective channel, noise, timing and frequency offsets) in order not to diverge the presentation of the key concept.

Given that the spectrum sensing device knows some key parameters of the operating air-interface such as the block length J, the number of subcarriers M, and the block duration T_b, the received continuous-time signal is sampled at the sampling period of $T_\text{s} = (T_\text{b})/(J)$. Hence, the timing offset can be expressed into two parts: the integer timing offset $n_\epsilon = \lfloor (\epsilon/T_\text{s}) \rfloor$ and the fractional timing offset $(\epsilon - n_\epsilon)$, where $\lfloor \cdot \rfloor$ denotes the floor operator. It is understood that the fractional timing offset can be incorporated into the channel impact. Hence, the discrete-time equivalent form of the received signal is [29]

$$\text{y}_n = \Omega^n \sum_{\ell=0}^{\mathcal{L}} h_\ell \text{x}_{kJ+n-\ell-n_\epsilon} + v_n, \tag{3}$$

where $\mathcal{L}$ denotes the upper bound of channel order ($\mathcal{L} \leq J - M$), $\Omega^n \triangleq \exp\left(\frac{j2\pi n\varepsilon}{M}\right)$, and the block index $k = \lfloor (n)/(J) \rfloor$. Consider an $J \times 1$ vector $\mathbf{y}_k = [y_{kJ+1}, y_{kJ+2}, \cdots, y_{kJ+J}]^T$, where $(\cdot)^T$ stands for the matrix transpose, then (3) can be expressed as the matrix form

$$\mathbf{y}_k = \mathbf{\Omega}_k(\mathbf{\Delta}(n_\varepsilon)\mathbf{x}_k + \underbrace{\mathbf{\nabla}(n_\varepsilon)\mathbf{x}_{k-1}}_{\text{IBI}}) + \boldsymbol{v}_k, \tag{4}$$

where $\mathbf{\Omega}_k \triangleq \text{diag}\{\Omega^{(kJ+1)}, \Omega^{(kJ+2)}, \cdots, \Omega^{(kJ+J)}\}$, $\boldsymbol{v}_k$ is the corresponding noise vector, $\mathbf{\Delta}(n_\varepsilon)$ is a lower triangle channel matrix, and $\mathbf{\nabla}(n_\varepsilon)$ is a upper triangle channel matrix. The detailed layout of both channel matrices depends on the timing offset n_ε, and the term $\mathbf{\nabla}(n_\varepsilon)\mathbf{x}_{k-1}$ is the inter-block interference (IBI).

B. Effect of Second Order Moment

The second-order moments of $\mathbf{y}_k$ in (4) can be computed as below

$$\mathbb{E}(\mathbf{y}_k \mathbf{y}_k^H) = \sigma_s^2 \mathbf{\Delta}(n_\varepsilon)\mathbf{\Psi}\mathbf{\Psi}^H \mathbf{\Delta}^H(n_\varepsilon) \\ + \sigma_s^2 \mathbf{\nabla}(n_\varepsilon)\mathbf{\Psi}\mathbf{\Psi}^H \mathbf{\nabla}^H(n_\varepsilon) + \mathcal{N}_0 \mathbf{I}_J, \tag{5}$$

where $\mathcal{N}_0$ is the noise power. It is observed that the above result is constant with respect to the block index k, and the carrier frequency offset (CFO) impact has been completely removed. This means that the second-order moments of y_n has a period of J. Furthermore, highly likely the diagonal entries of $\mathbb{E}(\mathbf{y}_k \mathbf{y}_k^H)$, for $n = 1, 2, \ldots, J$, are not constant with respect to the index n due to the frequency selectivity nature of the communication channels in high data rates.

Remark: In practice, the processing (5), i.e., ensemble average, is replaced by the time average

$$\mathbb{E}(\mathbf{y}_k \mathbf{y}_k^H) \approx \mathcal{C}_{\mathbf{y}}^{n_\epsilon} \triangleq \frac{1}{K} \sum_{k=0}^{K-1} (\mathbf{y}_k \mathbf{y}_k^H), \tag{6}$$

where K is the number of observation windows and $\mathcal{C}_{\mathbf{y}}^{n_\epsilon}$ is an $J \times J$ matrix. The above substitution is due to limited processing time available. This would result in fluctuation of

the ESD of AWGN (generation of impulse or heavily tailed noise). Hence, affecting the performance of any spectrum sensing algorithms regardless of what scheme is being employed.

C. Statement of The Spectrum Sensing Problem

The general problem of spectrum sensing is modelled as the binary hypothesis testing with hypothesises: $\mathcal{H}_0$, when the signal is absent; and $\mathcal{H}_1$, when the signal is present.

The specific problem of interest in this paper is: given the noise $\mathbf{v}_k$ to be white Gaussian, and independent of the multi-carrier signal $\mathbf{x}_k$ which is second-order cyclostationary with the period of N, what is the efficient way to determine the presence of the signal formulated in (3) specifically, in low SNR range?

It is understood that the random sequence $\mathbf{x}_n$ has the property $\mathbb{E}(\mathbf{x}_k\mathbf{x}_k^H) = \mathbb{E}(\mathbf{x}_{k+N}\mathbf{x}_{k+N}^H)$ due to it's cyclostationary property [27], and the random sequence v_k satisfies $\mathbb{E}(\mathbf{v}_k\mathbf{v}_k^H) = \mathcal{N}_0$. In other words, in the absence of signal, i.e., $\mathcal{H}_0$, the random sequence of observation $\{\mathbf{y}_k\}$ is a white process, otherwise, i.e., $\mathcal{H}_1$, a second-order cyclostationary process. Furthermore, as mentioned in Section II-B, making use of second-order moment of $\{\mathbf{y}_k\}$ results in overcoming the CFO phenomenon. Hence employing second order moment yields the hypothesises

$$\mathbb{E}(\mathbf{y}_k\mathbf{y}_k^H) \approx \mathcal{C}_{\mathbf{y}}^{n_\epsilon} \sim \begin{cases} \mathcal{X}_{2K}^2, & \mathcal{H}_0 \\ \mathcal{X}_{2K}^2\left(2\frac{\mathbf{\Lambda}}{\mathcal{N}_0\mathbf{I}}\right), & \mathcal{H}_1 \end{cases} \quad (7)$$

and

$$\mathbf{\Lambda} = \sigma_s^2\mathbf{\Delta}(n_\varepsilon)\mathbf{\Psi}\mathbf{\Psi}^H\mathbf{\Delta}^H(n_\varepsilon) + \sigma_s^2\mathbf{\nabla}(n_\varepsilon)\mathbf{\Psi}\mathbf{\Psi}^H\mathbf{\nabla}^H(n_\varepsilon) \quad (8)$$

where $\mathcal{X}_{2K}^2$ and $\mathcal{X}_{2K}^2(a)$ denote a central Chi squared distribution with $2K$ degrees of freedom, and non-central Chi squared with non-centrality factor a respectively. Thus, if the SNR of the received signal, i.e., $\frac{\mathbf{\Lambda}}{\mathcal{N}_0\mathbf{I}}$, was fairly high the hypothesis test in (7) will be trivial. The problem of interest in this paper is to consider spectrum sensing in very low SNR which, given (7), is a rather challenging problem.

III. CLUSTER-BASED DIFFERENTIAL ENERGY DETECTION

A. Sensing of CP-Based Multi-Carrier Signals

Form an $M \times M$ matrix $\bar{\mathcal{C}}_{\mathbf{y}}^{n_\epsilon}$ by collecting the last M columns and rows of $\mathcal{C}_{\mathbf{y}}^{n_\epsilon}$ defined in (6). Consider the special case where the timing offset $n_\epsilon = 0$, i.e., $\mathcal{C}_{\mathbf{y}}^0$. Due to the effect of CP, the second term at the right hand of (5) vanishes, i.e., IBI is removed, and the residual term can be written as

$$\bar{\mathcal{C}}_{\mathbf{y}}^0 = \sigma_s^2\mathbf{C}_M\mathcal{F}^H\mathbf{\Phi}\mathbf{\Phi}^H\mathcal{F}\mathbf{C}_M^H + \mathcal{N}_0\mathbf{I}_M, \quad (9)$$

where $\mathbf{C}_M$ is an $M \times M$ circulant channel matrix defined in [26]. Then, an M-point DFT operation is performed on $\bar{\mathcal{C}}_{\mathbf{y}}^0$ leading to

$$\tilde{\bar{\mathcal{C}}}_{\mathbf{y}}^0 \triangleq \mathcal{F}\bar{\mathcal{C}}_{\mathbf{y}}^0 = \sigma_s^2\mathcal{D}_M\mathbf{\Phi}\mathbf{\Phi}^H\mathcal{D}_M^H + \mathcal{N}_0\mathbf{I}_M, \quad (10)$$

where $\mathcal{D}_M \triangleq \mathcal{F}\mathbf{C}_M\mathcal{F}^H$ is an $M \times M$ diagonal matrix, whose diagonal entries are in fact the channel frequency response

(denoted by $\tilde{h}_m$). Let ϕ_m^T to be the mth row vector of $\mathbf{\Phi}$. Hence, the mth diagonal entry of $\tilde{\bar{\mathcal{C}}}_{\mathbf{y}}^0$ reads as

$$\left[\tilde{\bar{\mathcal{C}}}_{\mathbf{y}}^0\right]_m = \sigma_s^2|\tilde{h}_m|^2\|\phi_m\|^2 + \mathcal{N}_0, \ m = 1, 2, \cdots, M \quad (11)$$

where $\|\cdot\|^2$ denotes the Frobenius norm. Note that the above expression is equivalent to the ESD computation. In many multi-carrier systems such as OFDM, multi-carrier code division multiple access (MC-CDMA), and single-carrier frequency division multiple access (SC-FDMA), the term $\|\phi_m\|^2$ is normalized [28]. Therefore, the $\|\phi_m\|^2$ term in (11) can be ignored leading to

$$\left[\tilde{\bar{\mathcal{C}}}_{\mathbf{y}}^0\right]_m = \sigma_s^2|\tilde{h}_m|^2 + \mathcal{N}_0, \ m = 1, 2, \cdots, M. \quad (12)$$

Based on (12), we propose a cluster-based differential energy detection technique with the following steps. An overview of the proposed technique is illustrated in Fig. 1.

S1) Group $[\tilde{\bar{\mathcal{C}}}_{\mathbf{y}}^0]_m$, for $m = 1, 2, ..., M$, into $B = M/L$ clusters with each cluster having L elements. The mathematical form of each cluster can be expressed by

$$\mathbf{q}_i \triangleq [[\tilde{\bar{\mathcal{C}}}_{\mathbf{y}}^0]_i, [\tilde{\bar{\mathcal{C}}}_{\mathbf{y}}^0]_{i+B}, \cdots, [\tilde{\bar{\mathcal{C}}}_{\mathbf{y}}^0]_{i+(L-1)B}]^T, \ i = 1, 2, ..., B. \quad (13)$$

The grouping criteria are: c1) elements within each cluster are statistically uncorrelated or weakly correlated; c2) all clusters are almost identical or strongly correlated in the noiseless case, i.e., $\mathbf{q}_1 = \mathbf{q}_2 = ... = \mathbf{q}_B$. The criterion c1) is to assure that the channel gain within each cluster is sufficiently selective since the proposed differential energy detection technique aims to take advantage of the spectrum fluctuation induced by channel frequency selectivity. The criterion c2) is mainly for the purpose of de-noising through linear combination of all clusters on the step S3). Here, the noise is mainly referred to the residual noise after the second-order statistics (6).

In order to fulfil the criteria c1) and c2), we first divide the whole frequency band into L sub-bands with each accommodating B subcarriers. The mathematical form of the lth sub-band is expressible as: $\mathbf{p}_l \triangleq [[\tilde{\bar{\mathcal{C}}}_{\mathbf{y}}^0]_{(l-1)B+1}, [\tilde{\bar{\mathcal{C}}}_{\mathbf{y}}^0]_{(l-1)B+2}, \cdots, [\tilde{\bar{\mathcal{C}}}_{\mathbf{y}}^0]_{lB}]^T, \ l = 1, 2, ..., L$. When the bandwidth of each sub-band is smaller than the channel coherence bandwidth, all elements in $\mathbf{p}_l$ are highly correlated or approximately identical. Moreover, we can configure the parameter B such that the bandwidth of the group $[\mathbf{p}_l^T, [\tilde{\bar{\mathcal{C}}}_{\mathbf{y}}^0]_{lB+1}]^T$ is larger than the coherence bandwidth such that any two adjacent sub-bands are weakly correlated or even statistically independent. With the above configuration to be satisfied, the cluster $\mathbf{q}_i$ can be generated through block wise interleaving of $\mathbf{p}_l, \ l = 1, 2, ..., L$.

The above statement implicitly indicates that the clustering process requires the knowledge of the coherence bandwidth which can be computed assuming the availability of accurate channel models. In case the accurate channel models are not available at the sensing device, we can use the upper bound of channel order $\mathcal{L}$ to approximately estimate the coherence bandwidth (for instance we can let $B = \lfloor M/\mathcal{L} \rfloor$ since the coherence bandwidth is generally
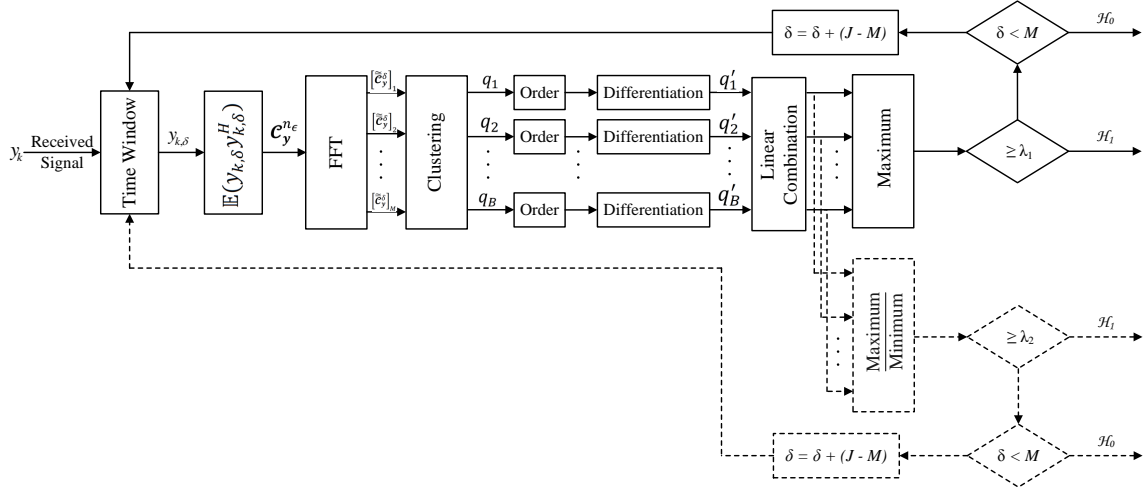
Fig. 1. Block diagram of the cluster-based differential energy detection algorithm.

inversely proportional to the channel order). Although, there is no optimal approach proposed to configure the parameter B, our simulation results in Section V demonstrate excellent performance when using the configuration $B = \lfloor M/\mathcal{L} \rfloor$.

It might also be worth mentioning that the idea of subcarrier clustering has recently received a lot of interests particularly for improving the communication quality and spectral efficiency in cognitive communications [30]-[33]. However, in our work, the subcarrier clustering is for improving the performance of spectrum sensing.

S2) Sort $\mathbf{q}_i$ in an ascending manner, and apply differentiation on each cluster respectively. This can be viewed as a rank conditioned rank selection process [34], where the order can change in an adaptive manner from zero to L. Advantages of such filtering process would be the insensitivity towards heavy tailed noise and impulsive noise while preserving the edge information [34]-[36]. The sorting operation allows smoothing of the input without affecting the statistics of the overall input. Furthermore, the differential operation allows us to observe the available second order moment diversity.

As it can be observed from (12), the sorting function will not have an effect in $\mathcal{H}_0$ scenario given that $\mathbf{q}_i|\mathcal{H}_0 = \mathcal{N}_0$. When considering a more practical scenario, i.e., limited number of samples, (ensemble average $\mathbb{E}(.)$ replaced by the time average (6)) we will experience noise power fluctuations. Thus, $\mathbf{q}_i|\mathcal{H}_0$ will no longer be constant and will follow the distribution described in (7). Given that the input signal at this stage, $\mathbf{q}_i$, is independent (due to the clustering operation performed in the previous stage) and identically distributed (i.i.d), with cumulative density function $F_q(\mathrm{q})$, the probability density function of the output of the sorting operation is given by [37]

$$f_{q_{r:L}}(\mathrm{q}) = r \binom{L}{r} F_q^{r-1}(\mathrm{q}) \left(1 - F_q(\mathrm{q})\right)^{L-r} f_q(\mathrm{q}), \quad (14)$$

where r $(1 \leq r \leq L)$ is the rth value returned after

the sorting operation, and $f_q(\mathrm{q})$ is the input probability density function. It can be observed from (14) that $f_{q_{r:L}}(\mathrm{q})$ is the product of the density function of the input, i.e., $f_q(\mathrm{q})$, and the function

$$w_{r:L}(\mathrm{q}) = r \binom{L}{r} F_q^{r-1}(\mathrm{q}) \left(1 - F_q(\mathrm{q})\right)^{L-r}. \quad (15)$$

It can be concluded that (15) is equivalent to beta probability density function [27]. Hence, the sorting operation is equivalent to multiplication of the input distribution function with a beta function, with shape parameters equal to r and $L - r + 1$. Replacing $u = F_q(\mathrm{q})$, the expression of the expected value of the rth value of the output can be calculated using

$$\mathbb{E}(q_{r:L}) = r \binom{L}{r} \int_{-\infty}^{\infty} \mathrm{q} F_q^{r-1}(\mathrm{q}) \left(1 - F_q(\mathrm{q})\right)^{L-r} f_q(\mathrm{q})$$
$$= \int_0^1 F_q^{-1}(u) \underbrace{r \binom{L}{r} u^{r-1} (1-u)^{L-r}}_{W_{r:L}(u)} \mathrm{d}u, \quad (16)$$

where $F_q^{-1}(u) = \mathrm{q}$ (since F_q is increasing in addition to being continuous) and $W_{r:L}(u)$ is the sorting function corresponding to rth highest value from set containing L elements. The above equation reveals that the expected value after sorting operation is the integral of the product between the sort function, $W_{r:L}(u)$, and the inverse distribution function. Fig. 2 shows the sorting function and the input distribution superimposed and further demonstrates how sorting operations allows focusing on a particular region. Thus, the sorting operation will reduce the effect of noise power fluctuation through smoothing the sudden changes by focusing on a specific region of the input density function out one time, this can be particularly useful when dealing with impulse/spike noise hence, having a direct effect on the error probability.

The sorting problem has attracted a great deal of research and since early 1950s many sorting algorithm
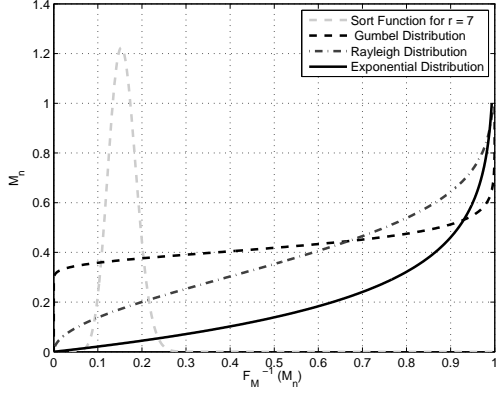
Fig. 2. Effect of the sort function on the output, for $N = 50$ and $r = 7$ on various distributions. This figure illustrates how the sort function focuses on a particular point of a distribution based on the value of r. Furthermore the shape difference for various distribution all having a mean value of 0.42 is also shown in this figure.

have been introduced in the literature, e.g., bucket sort, counting sort, spread sort. A comprehensive description of various search algorithms can be found in [38]. Hence, sorting operation in this step can be implemented using one of many developed sorting algorithm based on the memory/efficiency trade-off the spectrum sensing device requires. Therefore, the device does not need to perform the operations explained in (14)-(16) to sort the data.

The main objective of the differential operation, which is further performed in this stage, is to remove the constant noise floor, i.e., $\mathcal{N}_0$, contained in all $\mathbf{q}_i$ elements. The output of differentiation is denoted as $\mathbf{q}_i'$ with its lth element given by

$$[\mathbf{q}_i]_l' \triangleq [\mathbf{q}_i]_l - [\mathbf{q}_i]_{l-1} = \begin{cases} 0, & \mathcal{H}_0 \\ \sigma_s^2 \left(|\tilde{h}_{i,l}|^2 - |\tilde{h}_{i,l-1}|^2 \right), & \mathcal{H}_1 \end{cases}$$
(17)

It is clear that $[\mathbf{q}_i]_l'$ is zero for all l in the absence of the signal, and undergoes a fluctuation in the presence of the signal due to the channel frequency selectivity. This distinctive feature motivates the test statistics presented in S3) and allows us to overcome the noise uncertainty problem inherent in the conventional energy detection.

Furthermore, this stage is intended to exploit the second order moment diversity of the input signal distribution. Fig. 2 illustrates the shape/feature difference [39] (in terms of inverse CDF) which exists between various distributions. All three distributions in this figure have equal mean value, yet regions exist where the distributions are very distinct from one an other. In the case of no shape/feature difference, the performance of the proposed technique will degrade. Since today's high data rate communications always leads to frequency selective channel, we will experience shape difference and consequently second order moment diversity.

S3) Perform linear combination of $\mathbf{q}_i'$ for $i = 1, 2, \ldots, B$ for

the purpose of de-noising, and then the following test

$$\text{Test I}: \quad \max\left[\frac{1}{B} \sum_{i=1}^{B} \mathbf{q}_i' \right]_l \gtrless_{\mathcal{H}_0}^{\mathcal{H}_1} \lambda_1 \qquad (18)$$

$$\text{Test II}: \quad \frac{\max\left[\frac{1}{B} \sum_{i=1}^{B} \mathbf{q}_i' \right]_l}{\min\left[\frac{1}{B} \sum_{i=1}^{B} \mathbf{q}_i' \right]_l} \gtrless_{\mathcal{H}_0}^{\mathcal{H}_1} \lambda_2, \qquad (19)$$

where the threshold λ_1, λ_2 should be carefully configured to manage the PD and PFA, which will be discussed in the performance analysis (see Section IV). The test metrics presented in (18) and (19) represent the maximum and the maximum to minimum ratio of the clustered ESD respectively, which have been widely used for sub-optimum decision making with low computational cost [40]. It is shown in Section IV-C that the proposed differential energy detection technique can offer comparable performance to the optimal detector in Neyman-Pearson sense [41], however, the latter requires the knowledge of channel gain, noise power and signal power, which are often not available in practice for the spectrum sensing application.

B. Overcoming Timing Offset

As mentioned in Section II-B the effect of CFO has been already solved through employment of second-order statistics. Now, our main concern is to overcome the timing offset. In fact, the special case of $n_\epsilon = 0$ can be hardly captured due to the lack of timing synchronization mechanism before the spectrum sensing component. In order to handle the problem of unknown timing offset effectively, we propose an "***one ballot veto***" policy to reject the hypothesis $\mathcal{H}_0$. The policy is stated as follows:

S1) Form $J \times 1$ vectors, $\mathbf{y}_{k,\delta} = [y_{kJ+1+\delta}, y_{kJ+2+\delta}, \cdots, y_{kJ+J+\delta}]^T$, $k = 0, 1, \ldots, K$, where δ denotes the offset in time,

S2) Compute $\mathcal{C}_{\mathbf{y}}^{(n_\epsilon - \delta)} \triangleq \mathbb{E}(\mathbf{y}_{k,\delta} \mathbf{y}_{k,\delta}^H)$ according to (5), for $\delta = 0$, $(J - M)$, $2(J - M)$, ..., M;

S3) Apply the cluster-based differential energy detection explained in Section III-A on $\mathcal{C}_{\mathbf{y}}^{(n_\epsilon - \delta)}$, $\forall \delta$. If for any value of δ the test statistic satisfies $\mathcal{H}_1$ criterion it is understood that the signal is present and the cluster-based differential energy detection algorithm would not be applied on the input after detecting the first value of $\mathcal{C}_{\mathbf{y}}^{(n_\epsilon - \delta)}$ meeting the $\mathcal{H}_0$ condition.

The underlying idea is, in the presence of a signal, there exists such a δ fulfilling the condition $|n_\epsilon - \delta| < J - M$, and under this condition, the proposed spectrum sensing scheme can successfully reject the IBI. In the absence of signal, $\mathcal{C}_{\mathbf{y}}^{(n_\epsilon - \delta)}$ is approximately constant with respect to δ, due to constant energy of AWGN throughout the spectrum. Most certainly, this stage will add to the overall complexity of the algorithm which would be shown in Section IV-D. However, in order to increase the reliability of the sensing device, implementation of this stage is necessary.

C. Extension to the ZP-Based System

Let us start from the special case of $n_\epsilon = 0$. Using the result in [26], we can easily justify that the second term at the right hand of (5) vanishes due to the implementation of ZP, i.e., (2). Therefore, (5) can be expressed by

$$\mathbb{E}(\mathbf{y}_k \mathbf{y}_k^H) = \sigma_s^2 \boldsymbol{\Delta}(n_\varepsilon) \boldsymbol{\Psi} \boldsymbol{\Psi}^H \boldsymbol{\Delta}^H(n_\varepsilon) + \mathcal{N}_0 \mathbf{I}_J \quad (20)$$

$$= \sigma_s^2 \mathbf{C}_J \boldsymbol{\Psi} \boldsymbol{\Psi}^H \mathbf{C}_J^H + \mathcal{N}_0 \mathbf{I}_J. \quad (21)$$

Performing J-point DFT on (21) yields

$$\mathcal{F}_J \mathbb{E}(\mathbf{y}_k \mathbf{y}_k^H) \mathcal{F}_J^H = \sigma_s^2 \boldsymbol{\mathcal{D}}_J \bar{\boldsymbol{\Psi}} \bar{\boldsymbol{\Psi}}^H \boldsymbol{\mathcal{D}}_J^H + \mathcal{N}_0 \mathbf{I}_J, \quad (22)$$

where $\bar{\boldsymbol{\Psi}} \triangleq \mathcal{F}_J \boldsymbol{\Psi}$, $\mathcal{F}_J$ is an $J \times J$ DFT matrix normalized by the factor $(1)/(\sqrt{J})$, $\mathbf{C}_J$ is an $J \times J$ circulant channel matrix with $\boldsymbol{\mathcal{D}}_J$ formed by the corresponding channel frequency response. It is easy to observe that (22) has the same form as (10). Therefore, the three step spectrum sensing algorithm proposed in Section III-A for the CP-based system can, be straightforwardly, applied on (21).

Furthermore, the "one ballot veto" policy can be applied on the ZP-based system to handle the problem of unknown timing offset.

D. Knowledge of Key Parameters

The proposed spectrum sensing technique requires the knowledge of several key parameters about the operating air-interface as well as channel models (i.e.,, the block length J, the number of subcarriers M, the sampling rate T_s, as well as the upper bound of channel order $\mathcal{L}$). Those knowledge of parameters are very commonly assumed in almost all estimation and detection techniques including spectrum sensing, e.g., in [9] [14] [42] [43]. Lack of these parameters knowledge would result in performance degradation for all spectrum sensing techniques. Practically, it is possible to obtain the mentioned parameters through accessing a geo-location database. For example, the new Ofcom regulations [7] allow for sensing devices to access location-aided databases for obtaining key parameters about local air-interfaces and channel power delay profiles (PDPs). Design and maintenance of location-aided databases is an ongoing research activity in both Europe and US [4], [44]. Surely, the impact of imperfect knowledge of air-interface parameters on the spectrum sensing performance is of interest to telecommunication engineers, however in this paper we would focus onto the main technical problem due to the limited page budget.

IV. PERFORMANCE ANALYSIS

Conventionally, the metrics of interests for performance evaluation of spectrum sensing are mainly the PFA, PD, and computational complexity. The PFA is often formulated for the AWGN case since it would not be affected by the channel fading. However, the PD is related to the channel fading behaviour, and in this paper we are interested in the Rayleigh fading scenario. In addition to the PFA and PD analysis, we will present numerical results as well as the computational complexity of the proposed approach.

A. Probability of False Alarm

Mathematically, the probability of false alarm is defined by [9]

$$\text{PFA} \triangleq \Pr\left(\mathscr{T} \geq \lambda \mid \mathcal{H}_0\right) \quad (23)$$

where $\mathscr{T}$ denotes the test statistics. It reflects the probability of an event where the spectrum sensing device reports an alarm when the signal is actually not being transmitted.

Let's consider the special case of $n_\epsilon = 0$. It is understood that elements of $\mathbf{q}_i$ (see (13)) under the hypothesis $\mathcal{H}_0$ follow independent and identical central Chi squared distributions with $2K$ degrees of freedom [24], i.e.,

$$f_{\mathbf{q}_i | \mathcal{H}_0}(\alpha, K) = \frac{1}{2^K \Gamma(K)} \alpha^{K-1} \exp(-\alpha/2) \quad (24)$$

where $\Gamma(.)$ represents the Gamma function [27]. Hence, after the differentiation (ignoring the effect of the sorting operation), the l th element of $\mathbf{q}_i'$ based on Appendix A follows the p.d.f.

$$f_{[\mathbf{q}_i']_l | \mathcal{H}_0}(\alpha) \sim \frac{\left(\frac{1}{2}\right)^K \alpha^{K-1} \exp\left(-\frac{\alpha}{2}\right)}{(K-1)!}. \quad (25)$$

The linear combination $\overline{\mathbf{q}}_l' \triangleq \sum_{i=1}^B [\mathbf{q}_i']_l$ employed in (18)-(19) will result in the following MGF

$$\mathcal{M}([\overline{\mathbf{q}}']_l | \mathcal{H}_0) = \prod_{i=1}^B \mathcal{M}([\mathbf{q}_i']_l | \mathcal{H}_0) = \left((1 - 2t)^{-K}\right)^{-B} \quad (26)$$

It can be observed that the random variable $[\overline{\mathbf{q}}']_l | \mathcal{H}_0$ has an Erlang distribution [27] with the shape and rate parameter equal to $\varpi = KB$ and $\eta = 0.5$. Hence, its p.d.f is given by

$$f_{[\overline{\mathbf{q}}']_l | \mathcal{H}_0}(\alpha) = \frac{(\eta)^\varpi}{(\varpi - 1)!} (\alpha)^{\varpi - 1} \exp(-\eta \alpha) \quad (27)$$

Accordingly for Test II (see Appendix B), we can derive the p.d.f. of the ratio $[\overline{\mathbf{q}}']_l / [\overline{\mathbf{q}}']_j | \mathcal{H}_0$, $\forall\ 1 \leq l, j \leq L$ and $j \neq l$, bearing in mind that the values of $\overline{\mathbf{q}}_i'$ are non-negative, as [27]

$$f_{[\overline{\mathbf{q}}']_l / [\overline{\mathbf{q}}']_j | \mathcal{H}_0}(z) = \int_0^\infty \int_0^{\overline{q}_l' z} f_{[\overline{\mathbf{q}}']_l, [\overline{\mathbf{q}}']_j | \mathcal{H}_0}\left(\overline{q}_l', \overline{q}_j'\right) d\overline{q}_l'\ d\overline{q}_j'$$

$$= \int_0^\infty \overline{q}_l' f_{[\overline{\mathbf{q}}']_l, [\overline{\mathbf{q}}']_j | \mathcal{H}_0}\left(\overline{q}_l' z, \overline{q}_l'\right) d\overline{q}_l'$$

$$= \frac{z^{\varpi - 1} \Gamma(2\varpi)}{\Gamma(\varpi)^2 (1 + z)^{2\varpi}}. \quad (28)$$

Finally, applying (27) and (28) into (52) and (56) respectively, we can obtain the PFA as

$$\text{Test I}: \text{PFA} = 1 - \left(\frac{\mathcal{G}(\varpi, \eta \lambda_1)}{\Gamma(\varpi)}\right)^L \quad (29)$$

$$\text{Test II}: \text{PFA} =$$
$$1 - \left(1 - \frac{\lambda_2^\varpi \Gamma(2\varpi) {}_2F_1\left([\varpi, 2\varpi], B+1, -\lambda_2\right)}{\varpi\ \Gamma(\varpi)^2}\right)^\psi \quad (30)$$

where $\psi = \binom{L}{2}$, $\mathcal{G}(.,.)$ is the lower Gamma incomplete function, and ${}_2F_1([a, b], c, d)$ is the Gauss hypergeometric function [45].

The PFA formulas above indicate the probability where the second order moment diversity observed from the noise only input is higher than the test statistic. It can be observed from

(30) that Test II can only be applied and is meaningful if the channel order is, $L > 3$. Hence, given the maximum channel order one can choose which test to employ. Furthermore, it can be concluded from (29) and (30) that the PFA of proposed schemes is a function of the cluster size L, the number of clusters, B, and sample complexity K, as well as the thresholds λ_1, λ_2. Specifically, it is exponentially related to the inverse of the channel delay, i.e., L, implying that the performance is exponentially effected by the frequency selectivity of the environment. This was expected as the key idea behind the proposed spectrum sensing approach is to make a decision based on the observed second order moment diversity resulting from the frequency selective channel. Furthermore, PFA will be reduced dramatically as $K \to \infty$. Given that for practical applications, the PFA is often given a fixed value, such as 10% as per the FCCs requirement [4], (29) and (30) can be employed to determine the appropriate thresholds λ_1, λ_2 for a given air interface, channel order and the required observation length, i.e., $F(\lambda) = 1 - \text{PFA}$. The exact effect of threshold value on the performance of the proposed approach is shown in Section IV-C.

Remark: In the derivation of (25), we ignored the effect of the sorting operation. This is mainly because the exact probability density function of the rth order statistic from any continues population is rather difficult to deal with (see (14)) and in most cases requires numerical evaluation of a nontrivial integral [37]. Since the earliest known bounds for the expected value of highest order statistic with was derived by Gumbel and Hartley and David, many work has been done on statistical properties of order statistics. The summery of which can be found in [37]. Despite all the work carried out on the area of the order statistics still the only effective way for determining the distribution of $f_{r:L}(q)$ would be evaluating them numerically. However, using the probability-integral transformation we are able to approximate the variance of the rth order statistic, $\sigma^2_{r:L}(x)$, of any continuous distribution as

$$\sigma^2_{r:L}(q) \sim \frac{(L-r+1)}{(L+1)^2(L+2)^2}\left(f_x\left(E[q_{r:L}]\right)\right)^{-2}, \quad (31)$$

where $E[q_{r:L}]$, or in other words the expected value of rth order statistics, can be approximated by:

$$E[q_{r:L}] \sim F_q^{-1}\left(\frac{r}{L+1}\right), \quad (32)$$

where $F_q^{-1}(q)$ denotes the inverse cumulative distribution of the input signal. Please note that the above approximations will converge as $L \to \infty$ (see [37, Chapter 3] for proof). The above approximations indicate that the sorting operation will have a direct effect on the performance of the proposed algorithm since it will reduce the variance of the data significantly. Thus, it can be concluded that the sorting operation will reduce the effect of noise power fluctuation resulting from the limited observation length. Hence, having a direct effect on the error probability as the test statistic is subject to less variation. Since it is not mathematically feasible to derive the performance incorporating the sorting operation we have shown the effect of the sorting operation in Section V through simulations.

B. Probability of Detection

The probability of detection is defined by

$$\text{PD} \triangleq Pr\left(\mathcal{T} \geq \lambda \mid \mathcal{H}_1\right) \quad (33)$$

It reflects the probability of an event: the spectrum sensing device reports an alarm when the signal is indeed there. Similar to the analysis for PFA, we start the analysis of PD from the general model (7). It has been proved that the random variable $\mathbf{q}_i|\mathcal{H}_1$ follows non-central Chi squared distribution with the p.d.f. [9]

$$f_{\mathbf{q}_i|\mathcal{H}_1}(\alpha) = 0.5\left(\frac{\alpha}{2\gamma_i}\right)^{\frac{K-1}{2}}\exp\left(-\frac{2\gamma_i+\alpha}{2}\right)\mathcal{I}_{K-1}(\sqrt{2\gamma_i\alpha}) \quad (34)$$

where $\mathcal{I}(.)$ denotes the modified Bessel's function of the first kind, and γ_i the SNR affecting the $\mathbf{q}_i$ value.

Furthermore, we consider an interesting case when the SNR, γ, follows an independent and identical exponential distribution

$$f_\gamma(\alpha) = \frac{1}{\overline{\gamma}}\exp\left(-\frac{\alpha}{\overline{\gamma}}\right), \quad (35)$$

where $\overline{\gamma}$ denotes the SNR mean.

Remark: In fact, modelling the SNR as an i.i.d. exponential distribution implies that the communication channel is a Rayleigh fading channel. Rayleigh fading is considered as one of the most practical models for tropospheric and ionospheric signal propagation as well as for the effect of heavily built-up urban environments on radio signals. Rayleigh fading is mostly applicable when there is no dominant propagation along a line of sight between the transmitter and receiver [8]. Since, based on FCC regulations [4] there is no guarantee that there would exist a line of sight between the sensing device and the transmitter, it would be a reasonable assumption to model the fading channel as Rayleigh fading.

The distribution of $\Delta\gamma_l = \gamma_l - \gamma_{l-1}$, whose MGF is given by (refer to (48)-(50))

$$\mathcal{M}(\Delta\gamma_l) = \frac{1}{1+\frac{t^2}{1/\overline{\gamma}^2}}. \quad (36)$$

Hence, it can be concluded that $\Delta\gamma_l$ follows a Laplace distribution [27]. Considering that $\mathbf{q}_i'|\mathcal{H}_1$ follows a non-central Chi square distribution with $2K$ degrees of freedom and the non-centrality factor of $2\Delta\gamma_l$, and also the fact that $\Delta\gamma$ is non-negative, the term PD_l^{T1} in (53) is computed using the following

$$\text{PD}_l^{\text{T1}} = \frac{2}{\overline{\gamma}}\int_{\Delta\gamma_l=0}^{\infty}\mathcal{Q}_K\left(\varphi, \sqrt{\lambda_1}\right)\exp\left(-\frac{\Delta\gamma_l}{\overline{\gamma}}\right)\mathrm{d}\Delta\gamma_l \quad (37)$$

with $\varphi = 2\Delta\gamma_l$, where $\mathcal{Q}_K(a,b)$ denotes the generalized Marcum $\mathcal{Q}$-function defined by

$$\mathcal{Q}_K(a,b) = \frac{1}{a^{K-1}}\int_{x=b}^{\infty}x^K\exp\left(-\frac{x^2+a^2}{2}\right)\mathcal{I}_{K-1}(ax)\mathrm{d}x. \quad (38)$$

The PD for Test II can be evaluated using (57), where the

p.d.f. of $\Delta\gamma_l/\Delta\gamma_{d\neq l}|\mathcal{H}_1$ given by

$$f_{\Delta\gamma_l/\Delta\gamma_{d\neq l}|\mathcal{H}_1}(\alpha) = \int_{\beta=0}^{\infty} \beta f_{\Delta\gamma_l,\Delta\gamma_d|\mathcal{H}_0}(\beta,\beta\alpha)\,\mathrm{d}\beta$$
$$= \frac{1}{(1+\alpha)^2}. \qquad (39)$$

Hence, the term $\mathrm{PD}_l^{\mathrm{T2}}$ in (57) given Rayleigh fading is given by

$$\mathrm{PD}_l^{\mathrm{T2}} = \frac{2}{\overline{\gamma}} \int_{\Delta\gamma_l=0}^{\infty} \mathcal{Q}_K\left(\varphi,\sqrt{\lambda_2}\right) \frac{1}{(1+\Delta\gamma_l)^2} \mathrm{d}\Delta\gamma_l. \quad (40)$$

Once more considering the special case of $n_\epsilon = 0$, after the differentiation under the hypothesis $\mathcal{H}_1$, the lth differential SNR $\Delta\gamma_l$ corresponding to $\mathbf{q}_i$ (see (13)) follows the Laplace distribution with the p.d.f. based on the derivation in (26). Furthermore, the average of differential SNR $\overline{\Delta\gamma_l}$ can be computed by

$$f(\overline{\Delta\gamma_l}|\mathcal{H}_1) = \left(\frac{1}{\overline{\gamma}}\right)^{\varpi} \frac{\overline{\Delta\gamma_l}^{\varpi-1} e^{-\frac{\overline{\Delta\gamma_l}}{\overline{\gamma}}}}{(\varpi-1)!}. \qquad (41)$$

Then, the term $\mathrm{PD}_l^{\mathrm{T1}}$ can be evaluated by

$$\mathrm{PD}_l^{\mathrm{T1}} = \frac{1}{\overline{\gamma}(\varpi-1)!} \times \int_{\overline{\Delta\gamma_l}=0}^{\infty} \mathcal{Q}_K\left(\varphi,\sqrt{\lambda_1}\right)\left(\frac{\overline{\Delta\gamma_l}}{\overline{\gamma}}\right)^{\varpi-1}$$
$$\times \exp\left(-\frac{\overline{\Delta\gamma_l}}{\overline{\gamma}}\right) \mathrm{d}\overline{\Delta\gamma_l}. \qquad (42)$$

Based on the analysis in Appendix C, we can further write (42) into

$$\mathrm{PD}_l^{\mathrm{T1}} = \frac{\overline{\gamma}}{2+\overline{\gamma}} \exp\left(\frac{-\lambda_1}{2+\overline{\gamma}}\right) \times$$
$$\left[\sum_{k=0}^{\varpi-2}\left(\frac{2}{2+\overline{\gamma}}\right)^k L_k\left(\frac{-\lambda_1\overline{\gamma}}{2(2+\overline{\gamma})}\right)\left(1+\frac{2}{\overline{\gamma}}\right)\right.$$
$$\left.\left(\frac{2}{2+\overline{\gamma}}\right)^{\varpi-1} L_{\varpi-1}\left(-\frac{\lambda_1\overline{\gamma}}{2(2+\overline{\gamma})}\right)\right]$$
$$+ \left(\frac{2}{2+\overline{\gamma}}\right)^B \exp\left(-\frac{\lambda_1}{2}\right) \times$$
$$\left[\sum_{k=1}^{K-1}\frac{1}{k!}(0.5\lambda_1)^k \,_1F_1\left(KB;k+1;\frac{\lambda_1\overline{\gamma}}{2(2+\overline{\gamma})}\right)\right].$$

where $_1F_1(.;.;.)$ denotes the hypergeometric function [45], and $L_n(.)$ the Laguerre polynomial function defined by

$$L_n(x) = \sum_{r=1}^{n}(-1)^r\binom{n}{n-r}\frac{x^r}{r!}. \qquad (43)$$

We can obtain the PD for Test I by applying (43) into (53).

Evaluating the PD of Test II requires the p.d.f of the ratio of $\overline{\Delta\gamma_i}/\overline{\Delta\gamma_j}|\mathcal{H}_1$. Based on the derivation in (28), we have

$$f_{\overline{\Delta\gamma_i}/\overline{\Delta\gamma_j}|\mathcal{H}_1}(\alpha) = \frac{\alpha^{\varpi-1}\Gamma(2\varpi)}{\Gamma(B)^2(1+\alpha)^{2\varpi}}. \qquad (44)$$

Then, the term $\mathrm{PD}_l^{\mathrm{T2}}$ can be computed by

$$\mathrm{PD}_l^{\mathrm{T2}} = \frac{\Gamma(2\varpi)}{\Gamma(\varpi)^2} \int_{\overline{\Delta\gamma_n}=0}^{\infty} \mathcal{Q}_\varpi\left(\varphi,\sqrt{\lambda_2}\right) \frac{\overline{\Delta\gamma_l}^{\varpi-1}}{(1+\overline{\Delta\gamma_l})^{2\varpi}} \mathrm{d}\overline{\Delta\gamma_l}. \qquad (45)$$

Considering considerably Low SNR such that $1 \gg \overline{\Delta\gamma}$, the integration in (45) can be computed by using Appendix C and the analysis in [46, Appendix A]. Hence, $\mathrm{PD}_l^{\mathrm{T2}}$ can be expressed by:

$$\mathrm{PD}_l^{\mathrm{T2}} = \Phi\exp\left(\frac{-\lambda_2}{2\overline{\gamma}}\right)\left[\sum_{k=0}^{\varpi-2}L_k\left(\frac{-\lambda_2}{2}\right)+L_{\varpi-1}\left(-\frac{\lambda_2}{2}\right)\right]$$
$$+\Phi\exp\left(-\frac{\lambda_2}{2}\right)\left[\sum_{k=1}^{K-1}\frac{1}{k!}(0.5\lambda_2)^k \,_1F_1\left(\varpi;k+1;\frac{\lambda_2}{2}\right)\right] \qquad (46)$$

where $\Phi = \frac{\Gamma(2\varpi)}{\Gamma(\varpi)}$ (the full proof can be obtained by using[47, Eqn. (25)]). Finally, we can obtain the PD for Test II by applying (46) into (57).

It can be observed from (43) and (46) that the performance of the proposed spectrum sensing technique, in terms of PD, is affected by the average SNR value $\overline{\gamma}$, sample complexity K and the threshold value λ_1 and λ_2 and further exponentially effected by the channel order L. Moreover, it can be observed from (53) and (46) that the performance of Test II improves much faster with the increase in channel order, L. The effect of various parameters on the PD of the proposed approach will be discussed in detail and illustrated pictorially in Section IV-C.

C. Numerical Results and Discussions

Numerical results based on (29) and (30) for PFA and (53) and (57) for PD, are provided to visually demonstrate the effect of various factors. Fig. 3 illustrates how PD is affected by the observation length (latency) in Test I. The results are generated for the configuration where the number of sub-carriers $M = 64$, and the number of clusters $B = 6$. The threshold λ_1 was fixed for achieving PFA = 10% with the noise uncertainty factor set to 2 dB (the noise uncertainty factor in practical scenarios is typically between 1 to 2 dB [10]). The main factor causing noise uncertainty is the tem-
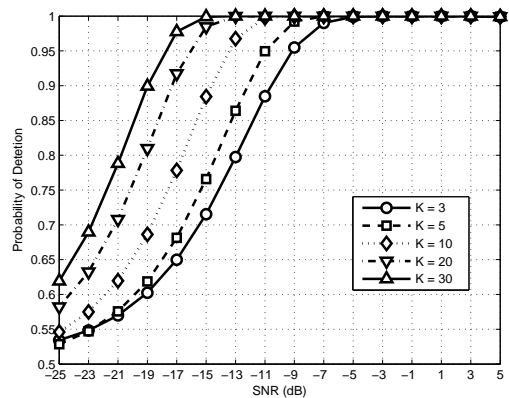

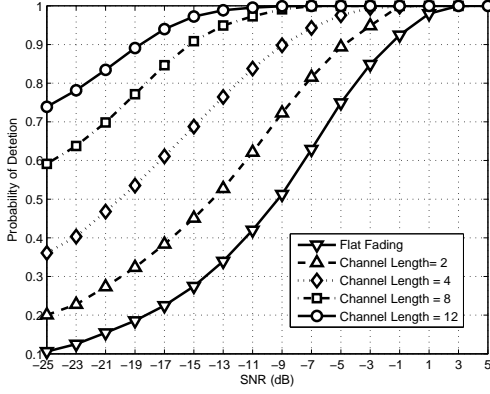Fig. 3. The relationship between the PD and the observation length for $M = 64$ and $L = 6$.

Fig. 4. The relationship between the PD and the coherence bandwidth, $\mathcal{L}$, and the observation length $K = 10$.
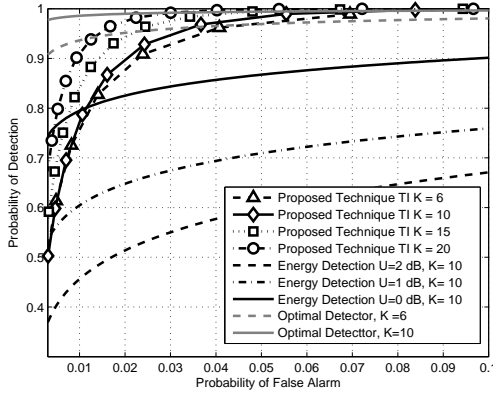

Fig. 5. Complementary ROC curves of the Test I and it's comparison with energy detection for various uncertainty factors (U), the the optimal detector based on Neyman-Pearson criteria. $\bar{\gamma} = -10dB$, $L = 8$ and $M = 64$.

Fig. 6. Complementary ROC curves of the Test II and it's comparison with energy detection for various uncertainty factors (U). $\bar{\gamma} = -10dB$, $L = 5$ and $M = 64$.

perature variations at the receiver which leads to inaccurate noise power measurements. The uncertainty in this paper is created by fixing assumed/estimated noise power based on the SNR value mentioned, while the real noise power varies with each realization by a certain degree according to the uncertainty factor. It is observed that the proposed approach features fast convergence rate. For example observing the point of PD = 90%, the PD improves by 5 dB in the SNR when the number of multi-carrier symbols K varies from 3 to 5, while this improvement is as small as approximately 1 dB when K varies from 20 to 30.

Fig. 4 shows how the channel length L would influence the PD when the observation length is set to $K = 10$. Take the point PD = 90% as an example, 8 dB gain in the SNR can be observed when L varies from 0 to 4. Furthermore, 10 dB improvement when it varies from 4 to 12. It is an interesting result which clearly indicates the channel frequency-diversity gain inherent in the proposed spectrum sensing scheme.

The complementary receiver operating characteristic (ROC) curve for both Test I and Test II (in Rayleigh fading channel) are shown in Fig. 5 and Fig. 6 respectively. These Figures
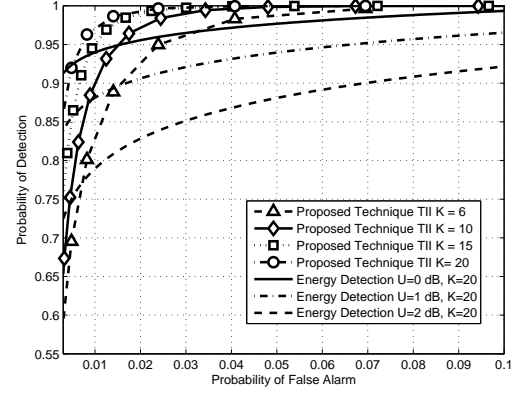
reflect a fundamental tradeoff between PFA and PD. Furthermore, the effect of the threshold value on both PFA and PD can be also observed, since different threshold values were employed to produce the PFA-PD tradeoff. In order to have a benchmark and also for performance comparison, the ROC curve for conventional energy detection with various uncertainty factors (U) are also illustrated. It is observed that the performance of the energy detection severely degrades as the uncertainty factor is introduced (this phenomenon has been fully investigated in [10]). While, due to differential stage of the proposed technique, it is considerably robust to uncertainty factor. For the sake of comprehensive performance comparison, Fig. 5 also illustrates the ROC of the optimal detector in Neyman-Pearson sense [41]. It should be noted that the optimal detector requires channel gain, noise power and the transmitted signal power (which is not a feasible solution in practical scenarios). Hence, as expected it delivers better performance.

D. Computational Complexity

The main complexity of the proposed scheme is due to the following stages:

1) The second-order time average: for the case of $n_\varepsilon = 0$, this stage requires $K^2 \times J$ complex multiplications and additions.
2) Discrete Fourier Transform: $M-$point DFT is implemented which introduces the complexity by $\mathcal{O}(M\log(M))$.
3) Sorting: there are B clusters consisting of L elements, hence, the complexity of this stage is $BL\mathcal{O}(L)$.
4) Differentiation: this stage consists of subtracting every element of $\mathbf{q}_i$ from its previous one for each cluster, hence the computational complexity is given by $B\mathcal{O}(L)$.
5) Linear combination: This would add a further complexity of $\mathcal{O}(B)$.
6) Decision making: Finally the extreme value(s) is selected and compared to the predetermined threshold value. Consequently adding a complexity factor of $\mathcal{O}(L)$.
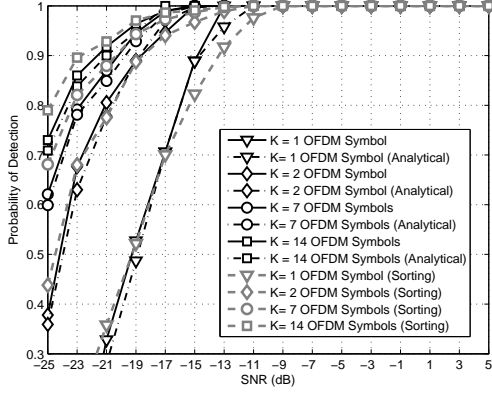
Fig. 7. Comparison of the simulation results and its equivalent analytical results for Test I. Furthermore the effect of the sorting operation is shown.
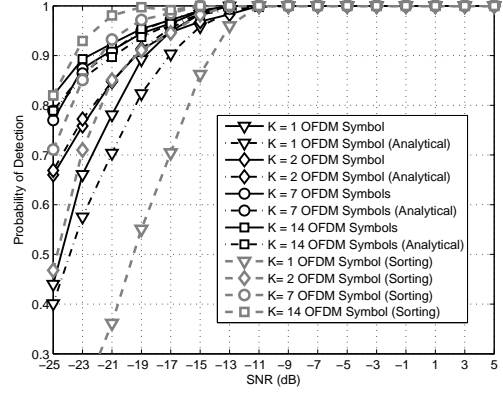

Fig. 8. Comparison of the simulation results of the proposed technique with and without the sorting operation and its equivalent analytical results for Test II.

Resulting in the overall computational complexity:

$$\mathcal{O}(K^2 J) + \mathcal{O}(M \log(M)) + B(L+1)\mathcal{O}(L) + \mathcal{O}(B) + \mathcal{O}(L) \tag{47}$$

Note that the above complexity is for the case $n_\epsilon = 0$. When employing the "one ballot veto" scheme for arbitrary n_ε (see Section III-A), the computational complexity is increased by a factor of $((M)/(J - M))$ (in the worst case scenario). As it can be observed the last three terms are negligible in comparison with other terms, the overall complexity of the proposed scheme can be approximately written by $M\left[\mathcal{O}(K^2 J) + \mathcal{O}(M \log(M))\right]/(J - M)$. This reflects that the proposed scheme requires a relatively low computational complexity, making it suitable for practical scenarios, where computational efficiency is a key issue.

V. SIMULATION RESULTS AND DISCUSSIONS

Computer simulations were performed to evaluate the proposed spectrum sensing scheme. The system investigated in this section has $M = 2,048$ sub-carriers with the sub-carrier spacing of 15 kHz (3GPP LTE-advanced system [48]), each frame consists of 7 OFDM blocks with the CP length of $J - M = 160$, the sampling frequency is the same as the signal bandwidth of 30.72 MHz. The carrier frequency is also set at 5 GHz. The communication channel is generated according to the WINNER channel model under B2 outdoor scenario [49], and the sensing device is moving at the speed of 3 km/h. The SNR is defined by the average received symbol energy to noise ratio at the sensing device. The threshold for hypothesis test is carefully chosen so that the PFA is fixed to 10%. All simulation results were obtained by averaging over 2,000 Monte Carlo realizations.

Experiment 1: The objective of this experiment is to examine the analytical analysis obtained in the previous sections by comparing them against the simulation results based on the configuration explained above and further to show the effect of the sorting operation on the performance of the proposed scheme. Fig. 7 and Fig. 8 demonstrate the probability of detection for different observation lengths given various
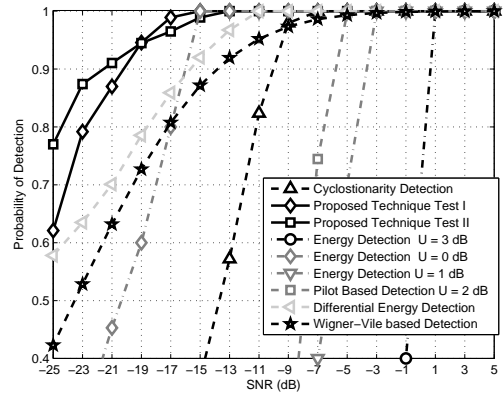

Fig. 9. The performance comparison of the proposed technique, frequency-domain energy detection, second order cyclostationarity, pilot based detection and differential energy detection for $K = 7$.

average SNRs for Test I and Test II, respectively. We can observe a very small difference between analytical results and simulation results when the observation length is larger than two symbols duration. The difference becomes large when the observation length is less than two symbols duration. This is mainly caused by the insufficient statistics used in signal processing. Comparing Fig. 7 and Fig. 8 verifies that Test II outperforms Test I, particularly, when the observation length is short. This difference is mitigated with the increase of observation length. We have also shown the effect of the sorting operation through simulations in Fig. 7 and Fig. 8. As it can be observed, the sorting operation can improve the performance as the observation length is increased. This was expected as previously explained in Section III-A and Section IV-A.

Experiment 2: The objective of this experiment is to examine the proposed scheme with respect to the state-of-the-art spectrum sensing approaches. Since, the proposed approach is based on exploiting the second order moment frequency diversity, it is essential to check how much gain is introduced
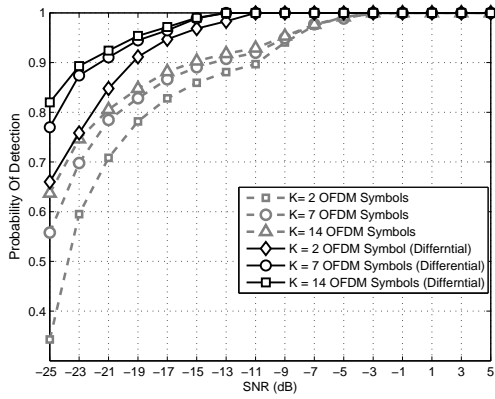
Fig. 10. The effect of the differential and clustering stages on the performance of the proposed spectrum sensing technique.

due to this exploitation by having the frequency-domain energy detection as a benchmark for performance comparison. The threshold setting for energy detection can be found in [9]. The simulation performed for energy detection are based on noise uncertainty factor, $U = 0, 1, 3$ dB and the threshold is based on the assumed/estimated noise power while the real noise power varies with each Monte Carlo realization by a certain degree depending on the uncertainty factor. Fig. 9 shows the performance comparison when observation length, $K = 7$ symbols. It can be observed that the performance of energy detection is considerably dependent on the noise uncertainty factor. It is further proved in [10] that increasing the observation length does not affect the performance of the energy detection scheme when the the exact noise power is not known, i.e., $U \neq 0$. Fig. 9 also illustrates the performance of the second order cyclostationarity based detection. The proposed approach is able to outperform the second order cyclostationarity by at least 8 dB when $K = 7$. Cyclostationarity based detection relies on the cyclic frequency of the received signal to determine existence of a source. Hence, deep fading at cyclic frequency can have a detrimental effect on its performance while the proposed technique takes advantage of this fading to exploit the moment diversity. The performance of the proposed approach is also compared to Wigner-Ville based spectrum sensing [17]. For this purpose, in order to have fair comparison, we have modified the original work in [17] to accommodate a SISO environment. As shown in Fig. 9 the mentioned approach can deliver acceptable performance up to SNR of -16 dB. However, as the SNR further decreases, the performance of Wigner-Ville transform based approach also decreases. This was expected since noise power fluctuation increases such that it makes the edge detection used in this approach reluctant. The performance comparison is also carried out for pilot based detection [50] and differential based energy detection [51]. In order to carry out simulations for the pilot based detection, it is assumed that the pilot symbols are embedded in each OFDM block, which are equally spaced for every 16 or 32 sub-carriers. Since, the mentioned pilot based detection is based on the energy of the pilot symbols

it is affected by the noise uncertainty factor. Therefore, the performance was evaluated for $U = 0$ dB and $U = 2$ dB. Fig. 9 shows that the proposed technique is able to outperform the mentioned technique by at least 10 dB. It should be also noted that pilot based detection requires synchronization and pilot information while this problem can be overcome in the proposed technique with the implementation of the "one ballot veto" policy.

Since the detection technique introduced in [51], also exploits the frequency diversity of the channel our main concern is to observe how much improvement can be delivered by the clustered-based energy detection. Fig. 9 further shows that the proposed approach can outperform the differential energy detection by at least 5 dB in low SNR environments. This improvement is mainly due to the clustering, the linear combination and the "one ballot veto" policy which are implemented. It is noteworthy to mention that the clustering operation not only improves the performance of the proposed techniques but also reduces the complexity by a factor of $O(2M)$.

Experiment 3: The objective of this experiment is to observe the performance improvement due to the clustering and differential stages. As it was explained, the main purpose of the differential stage is to remove the AWGN which is available in all the frequency bands and to further exploit the frequency diversity, while the clustering operation is to remove any possible correlation in the ESD due to the fading channel. In order to observe how much improvement can be achieved when incorporating these two stages (i.e., clustering and differential stages), we have set an experiment were the proposed technique in Section III-A is compared to its equivalent without the mentioned two stages. Without differential and clustering operations the proposed technique can be thought of as a simplified eigenvalue detection [19] where instead of making decision based on the ratio of the eigenvalues of the covariance matrix of the received signal, the decision is based on the ratio of the maximum and the minimum of the ESD of the received signal. This comparison is possible since, in multi-carrier systems, parallel transmission is performed, hence, the DFT decomposition can be considered as a special case of eigenvalue decomposition. The result of this performance comparison for different observation lengths is shown in Fig. 10. As it can be observed we are able to achieve up to 10 dB gain in performance. Furthermore, as expected, this gain is more apparent as the observation is increased.

VI. CONCLUSION

In this paper a novel differential energy detection scheme for multi-carrier systems, which can form fast and reliable decision of spectrum availability even in very low signal-to-noise ratio (SNR) environment has been proposed. The underlying initiative of the proposed scheme is applying order statistics onto the clustered differential energy spectral density in order to exploit the channel frequency diversity inherent in high data-rate communications. Specifically, the clustering operation is utilized to group uncorrelated subcarriers into clusters according to the coherence bandwidth in order to enjoy

good frequency diversity, and the differential operation applied onto each cluster effectively reduces the impact of noise uncertainty, consequently improves the sensing reliability. In order to analytically evaluate the proposed scheme, both the average PD and PFA were derived for Rayleigh fading channels. The closed-form expressions showed a clear relationship between the sensing performance and the cluster size (the indicator of diversity gain). It has been shown through simulations that the proposed approach provides up to 10 dB gain in comparison to the state-of-the-art spectrum sensing schemes.

APPENDIX A

The moment generating function (MGF) of the output of the differential process in step S2 in Section III-A, $q'_{i,l}|\mathcal{H}_0$, (which is expressed as the product of the MGF of two central Chi squared random variables shown in (24)) is given by [27]

$$\mathcal{M}(q'_i|\mathcal{H}_0) \triangleq \mathcal{M}(q_{i+1}|\mathcal{H}_0) \times \mathcal{M}(q_i|\mathcal{H}_0)$$
$$= (1 - 2t)^{-K} \times (1 + 2t)^{-K} = (1 - 4t^2)^{-K} \quad (48)$$

where $\mathcal{M}(a)$ denotes the MGF of random variable a, and t denotes the time-domain index. Hence, (48) indicates that $q'_i|\mathcal{H}_0$ follows a summation of K i.i.d. Laplace distributions (with location parameter $\mu = 0$ and scaling parameter $\xi = 2$), whose their MGF functions are given by [27]

$$\mathcal{M}(\mathscr{L}) = \frac{e^{\mu t}}{1 - \xi^2 t^2} \quad (49)$$

where $\mathscr{L}$ denotes a random variable following a Laplace distribution. Moreover, we know for a fact that $q'_{i,l}|\mathcal{H}_0$ is non-negative due to the sorting process. Therefore, it can be concluded that distribution of $q'_{i,l}|\mathcal{H}_0$ can be further simplified to summation of K exponential distributions with scale factor 0.5. Hence, the probability density function (p.d.f.) of

$$f_{q'_{i,l}|\mathcal{H}_0}(\alpha) \sim \frac{0.5^K \alpha^{K-1} \exp(-0.5\alpha)}{(K-1)!} \quad (50)$$

APPENDIX B

Since $\mathbf{q}'$ is i.i.d., for $1 \leq m \leq L$, we would have

$$F_{\max(\overline{\mathbf{q}'}|\mathcal{H}_0)}(\lambda_1) \triangleq \Pr(\max(\overline{\mathbf{q}'}|\mathcal{H}_0) \leq \lambda_1)$$
$$= \Pr(\overline{q}'_L \geq \lambda_1, \overline{q}'_L > \overline{q}'_m \ \forall m \neq L|\mathcal{H}_0)$$
$$\cup \Pr(\overline{q}'_{L-1} \geq \lambda_1, \overline{q}'_L > \overline{q}'_m \ \forall m \neq L-1|\mathcal{H}_0) \cup \ldots$$
$$\cup \Pr(\overline{q}'_1 \geq \lambda_1, q'_1 > \overline{q}'_m \ \forall m \neq 1|\mathcal{H}_0)$$
$$= \Pr(\overline{q}'_L \geq \lambda_1, q'_L > \overline{q}'_m \forall m \neq L|\mathcal{H}_0)$$
$$+ \Pr(\overline{q}'_{L-1} \geq \lambda_1, q'_L > \overline{q}'_m \forall m \neq L-1|\mathcal{H}_0) + \ldots$$
$$+ \Pr(\overline{q}'_1 \geq \lambda_1, q'_1 > \overline{q}'_m \forall m \neq 1|\mathcal{H}_0)$$
$$= [F_{\overline{q}'_i|\mathcal{H}_0}(\lambda_1)]^L. \quad (51)$$

PFA for Test I can then be evaluated using the cumulative density function (CDF) of $\max(\mathbf{q}'|\mathcal{H}_0)$, given $q_l|\mathcal{H}_0$ variables are i.i.d.

$$\text{Test I}: \text{PFA} = \Pr(\max(\overline{\mathbf{q}}'|\mathcal{H}_0) > \lambda_1)$$
$$= 1 - \underbrace{[F_{\max(\overline{\mathbf{q}}'|\mathcal{H}_0)}(\lambda_1)]}_{\triangleq \Pr(\max(\overline{\mathbf{q}}'|\mathcal{H}_0) \leq \lambda_1)}$$
$$= 1 - [F_{\overline{q}'|\mathcal{H}_0}(\lambda_1)]^L. \quad (52)$$

Furthermore, the PD for Test I can be expressed as

$$\text{Test I}: \text{PD} = \Pr(\max(\Delta\gamma|\mathcal{H}_1) > \lambda_1)$$
$$= 1 - [F_{\max(\overline{\mathbf{q}}')|\mathcal{H}_1}(\lambda_1)]$$
$$= 1 - \prod_{l=1}^{L} \underbrace{F_{q'_l|\mathcal{H}_1}(\lambda_1)}_{1 - \text{PD}_l^{\text{T1}}} \quad (53)$$

APPENDIX C

The Hypothesis for Test II is based on the ratio of maximum to minimum of $\mathbf{q}'_i|\mathcal{H}_0$, whose CDF can be computed using

$$F_{\max/\min(\overline{\mathbf{q}}'|\mathcal{H}_0)}(\lambda_2) \triangleq \Pr\left(\frac{\max(\overline{\mathbf{q}}'|\mathcal{H}_0)}{\min(\overline{\mathbf{q}}'|\mathcal{H}_0)} \leq \lambda_2\right)$$
$$= \Pr(\frac{\overline{q}'_L}{\overline{q}'_1} \geq \lambda_2, \overline{q}'_L > \overline{q}'_m \forall m \neq L \& \overline{q}'_1 < \overline{q}'_g \forall g \neq 1|\mathcal{H}_0) \cup$$
$$\ldots \cup \Pr(\frac{\overline{q}'_1}{\overline{q}'_L} \geq \lambda_2, \overline{q}'_1 > \overline{q}'_m \forall m \neq L \& \overline{q}'_L < \overline{q}'_g|\mathcal{H}_0)$$
$$= \Pr(\frac{\overline{q}'_L}{\overline{q}'_1} \geq \lambda_2, \overline{q}'_L > \overline{q}'_m \forall m \neq L \& \overline{q}'_1 < \overline{q}'_g|\mathcal{H}_0) +$$
$$\ldots + \Pr(\frac{\overline{q}'_1}{\overline{q}'_L} \geq \lambda_2, \overline{q}'_1 > \overline{q}'_m \forall m \neq L \& \overline{q}'_L < \overline{q}'_g|\mathcal{H}_0). \quad (54)$$

Due to $\max/\min(\mathbf{q}'|\mathcal{H}_0) \geq 0$, we can further express (54) as

$$F_{\max/\min(\overline{\mathbf{q}}'|\mathcal{H}_0)}(\lambda_2) = \int_0^\infty \int_{\alpha=0}^{\lambda_2\beta} f_{\overline{q}'_L \overline{q}'_1|\mathcal{H}_0}(\alpha, \beta) \, d\alpha \, d\beta$$
$$+ \int_0^\infty \int_{\alpha=0}^{\lambda_2\beta} f_{\overline{q}'_{L-1}\overline{q}'_1|\mathcal{H}_0}(\alpha, \beta) \, d\alpha \, d\beta + \ldots$$
$$+ \int_0^\infty \int_{\beta=0}^{\lambda_2\alpha} f_{\overline{q}'_1\overline{q}'_L|\mathcal{H}_0}(\alpha, \beta) \, d\beta \, d\alpha \quad (55)$$

Once $F_{\max/\min(\mathbf{q}'|\mathcal{H}_0)}(\lambda_2)$ is obtained, the PFA for Test II can be easily computed by

$$\text{Test II}: \text{PFA} = 1 - F_{\max/\min(\overline{\mathbf{q}}'|\mathcal{H}_0)}(\lambda_2). \quad (56)$$

Furthermore, for Text II we have

$$\text{Test II}: \text{PD} = \Pr(\max/\min(\overline{\mathbf{q}}'|\mathcal{H}_1) > \lambda_2)$$
$$= 1 - [F_{\max(\overline{q}'/\overline{q}'|\mathcal{H}_1)}(\lambda_2)]$$
$$= 1 - \prod_{l=1}^{\binom{L}{2}} \underbrace{F_{(\Delta\gamma_l/\Delta\gamma_{d\neq l})|\mathcal{H}_1}(\lambda_2)}_{1 - \text{PD}_l^{\text{T2}}} \quad (57)$$

APPENDIX D

Using the recursion for (38), the following result is obtained

$$Q_K(ax, b)$$
$$= \left(\frac{b}{ax}\right)^{K-1} e^{-\frac{b^2+(ax)^2}{2}} \mathcal{I}_{N/2-1}(axy) + Q_{K-1}(ax, b)$$
$$= \sum_{n=0}^{K-1} \left(\frac{b}{ax}\right)^n e^{-\frac{b^2+(ax)^2}{2}} I_n(axy) Q_1(ax, b). \quad (58)$$

Applying (58) into (42), the integration part can be computed by

$$\int_{x=0}^{\infty} Q_K(\alpha x, \beta) x^{2B-1} e^{-p^2 x^2/2}$$
$$= \sum_{n=0}^{K-1} \left(\frac{\beta}{a}\right)^n \int_{x=0}^{\infty} x^{2B-1-n} e^{-\frac{\beta^2+(p^2+\alpha^2)x^2}{2}} I_n(\alpha\beta x)\mathrm{d}x +$$
$$\int_{x=0}^{\infty} x^{2(B-1)} \exp^{-p^2 x^2/2} Q_1(\alpha x, \beta)\mathrm{d}x \quad (59)$$

where $x^2 = \frac{\overline{\Delta\gamma}}{\overline{\gamma}}$, $p = 2$, and $\beta = \sqrt{\lambda}$. Using combination of Bessel functions and exponentials in [45], the above terms can be evaluated.

REFERENCES

[1] Q. Zhao and B. M. Sadler, "A survey of dynamic spectrum access," *IEEE Signal Process. Mag.*, vol. 24, no. 3, pp. 79 –89, May. 2007.

[2] J. Mitola III and G. Q. Maguire, "Cognitive radio: making software radios more personal," *IEEE Pers. Commun.*, vol. 6, no. 4, pp. 13 –18, Aug. 1999.

[3] A. Ghasemi and E. S. Sousa, "Spectrum sensing in cognitive radio networks: requirements, challenges and design trade-offs," *IEEE Commun. Mag.*, vol. 46, no. 4, pp. 32 –39, Apr. 2008.

[4] Federal Communications Commission, "Notice of proposed rule making and order: Facilitating opportunities for flexible, efficient, and reliable spectrum use employing cognitive radio technologies," pp. 03 –108, Feb. 2005.

[5] "IEEE draft standard for information technology telecommunications and information exchange between systems local and metropolitan area networks specific requirements part 22.1: Standard to enhance harmful interference protection for low power licensed devices operating in tv broadcast bands," *IEEE P802.22.1/D8*, pp. 1 –151, Aug. 2010.

[6] V. Goncalves and S. Pollin, "The value of sensing for TV white spaces," *in IEEE DySPAN, Aachen, Germany*, pp. 231 –241, May 2011.

[7] Ofcom, "Implementing geolocation - summary of consultation responses and next steps [online] http://stakeholders.ofcom.org.uk/consultations/geolocation/statement," Sep. 2011.

[8] R. S. Rappaport, *Wireless Communications: Principles and Practice*, Number 4th. Prentice Hall, 2001.

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

[10] R. Tandra and A. Sahai, "SNR walls for signal detection," *IEEE J. of Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 4 –17, Feb. 2008.

[11] J. Wu, T. Luo, and G. Yue, "An energy detection algorithm based on double-threshold in cognitive radio systems," *in IEEE ICISE, Nanjing, China*, pp. 493 –496, Dec. 2009.

[12] Y. M. Kim, G. Zheng, S. Hwan Sohn, and J. M. Kim, "An alternative energy detection using sliding window for cognitive radio system," *in IEEE ICACT, Phoenix Park, Korea*, pp. 481 –485, Feb. 2008.

[13] K. Kim, Y. Xin, and S. Rangarajan, "Energy detection based spectrum sensing for cognitive radio: An experimental study," *in IEEE GLOBE-COM, Florida, USA*, pp. 1 –5, Dec. 2010.

[14] A. V. Dandawate and G. B. Giannakis, "Statistical tests for presence of cyclostationarity," *IEEE Trans. Signal Process.*, vol. 42, no. 9, pp. 2355 –2369, Sep. 1994.

[15] S. Chaudhari, V. Koivunen, and H. V. Poor, "Autocorrelation-based decentralized sequential detection of OFDM signals in cognitive radios," *IEEE Trans. Signal Process.*, vol. 57, no. 7, pp. 2690 –2700, Jul. 2009.

[16] S. Mallat and W. L. Hwang, "Singularity detection and processing with wavelets," *IEEE Trans. Inf. Theory*, vol. 38, no. 2, pp. 617 –643, Mar. 1992.

[17] M. Biagi, V. Polli, and J. A. A. Freitas, "An image processing approach to distributed access for multiantenna cognitive radios," *in IEEE ISWCS, York, UK*, pp. 621 –625, Sep. 2010.

[18] H. S. Chen, W. Gao, and D. G. Daut, "Signature based spectrum sensing algorithms for IEEE 802.22 WRAN," *in IEEE ICC, Glasgow, UK*, pp. 6487 –6492, Jun. 2007.

[19] Y. Zeng and Y. C. Liang, "Eigenvalue-based spectrum sensing algorithms for cognitive radio," *IEEE Trans. Commun.*, vol. 57, no. 6, pp. 1784 –1793, Jun. 2009.

[20] D. Denkovski, V. Atanasovski, and L. Gavrilovska, "Efficient mid-end spectrum sensing implementation for cognitive radio applications based on USRP2 devices," *in IEEE COCORA, Budapest, Hungary*, Apr. 2011.

[21] D. Denkovski, V. Atanasovski, and L. Gavrilovska, "HOS based goodness-of-fit testing signal detection," *IEEE Commun. Lett.*, pp. 1 –4, To appear 2012.

[22] Y. L. Zhang, Q. Y. Zhang, and T. Melodia, "A frequency-domain entropy-based detector for robust spectrum sensing in cognitive radio networks," *IEEE Commun. Lett.*, vol. 14, no. 6, pp. 533 –535, Jun. 2010.

[23] T. Yucek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *IEEE Commun. Surveys Tuts.*, vol. 11, no. 1, pp. 116 –130, 2009.

[24] J. Ma, G. Y. Li, and B. H. Juang, "Signal processing in cognitive radio," *Proc. IEEE*, vol. 97, no. 5, pp. 805 –823, May 2009.

[25] I. F. Akyildiz, W. Y. Lee, M. C. Vuran, and S. Mohanty., "A survey on spectrum management in cognitive radio networks," *IEEE Commun. Mag.*, vol. 46, no. 4, pp. 40 –48, Apr. 2008.

[26] Z. Wang and G. B. Giannakis, "Wireless multicarrier communications," *IEEE Signal Process. Mag.*, vol. 17, no. 3, pp. 29 –48, May 2000.

[27] A. Papoulis and S. U. Pillai, *Probability, Random Variables and Stochastic Processes*, Number 4th. McGraw-Hill New York, 2002.

[28] T. Keller and L. Hanzo, *OFDM and MC-CDMA: A Primer*, Wiley - IEEE, 2007.

[29] M. Movahhedian, Y. Ma, and R. Tafazolli, "Blind CFO estimation for linearly precoded OFDMA uplink," *IEEE Trans. Signal Process.*, vol. 58, no. 9, pp. 4698 – 4710, Sep. 2010.

[30] M. L. Gao, "The self-adapt spectrum management formula base on cluster model," *in IEEE ICMT, Zurich, Switzerland*, pp. 901 –903, Jul. 2011.

[31] A. B. McDonald and T. F. Znati, "A mobility-based framework for adaptive clustering in wireless ad hoc networks," *IEEE J. Sel. Areas in Commun.*, vol. 17, no. 8, pp. 1466 –1487, Aug. 1999.

[32] E. Dall'Anese and G. B. Giannakis, "Distributed cognitive spectrum sensing via group sparse total least-squares," *in IEEE CAMSAP, San Juan, Puerto Rico*, pp. 341 –344, Dec. 2011.

[33] C. Ragusa, A. Liotta, and G. Pavlou, "An adaptive clustering approach for the management of dynamic systems," *IEEE J. Sel. Areas in Commun.*, vol. 23, no. 12, pp. 2223 – 2235, Dec. 2005.

[34] R. C. Hardie and K. E. Barner, "Rank conditioned rank selection filters for signal restoration," *IEEE Trans. Image Process.*, vol. 3, no. 2, pp. 192 –206, Mar. 1994.

[35] R. C. Hardie and C. Boncelet, "LUM filters: a class of rank-order-based filters for smoothing and sharpening," *IEEE Trans. Signal Process.*, vol. 41, no. 3, pp. 1061 –1076, Mar. 1993.

[36] J. H. Lin, T. M. Sellke, and E. J. Coyle, "Adaptive stack filtering under the mean absolute error criterion," *IEEE Trans. Acoust., Speech and Signal Process.*, vol. 38, no. 6, pp. 938 –954, Jun. 1990.

[37] H. A. David and H. N. Nagaraja, *Order statistics*, John Wiley and sons Hoboken. New Jersey, 2003.

[38] R. Sedgewick and K. Wayne, *Algorithms*, Number 4th. Addison-Wesley Professional, 2011.

[39] J. Serra, *Image Analysis and Mathematical Morphology: Theoretical Advances*, vol. 2, New York: Academic, 1988.

[40] S. R. Sternberg, "Biological image processing," *Computer*, vol. 16, no. 1, pp. 22 –34, Jan. 1983.

[41] A. Taherpour, M. Nasiri-Kenari, and S. Gazor, "Multiple antenna spectrum sensing in cognitive radios," *IEEE Trans. Wireless Commun.*, vol. 9, no. 2, pp. 814 –823, Feb. 2010.

[42] Z. Quan, S. Cui, A. H. Sayed, and H. V. Poor, "Optimal multiband joint detection for spectrum sensing in cognitive radio networks," *IEEE Trans. Signal Process.*, vol. 57, no. 3, pp. 1128 –1140, Mar. 2009.

[43] B. Farhang-Boroujeny, "Filter bank spectrum sensing for cognitive radios," *IEEE Trans. Signal Process.*, vol. 56, no. 5, pp. 1801 –1811, May. 2008.

[44] H. R. Karimi, "Geolocation databases for white space devices in the UHF TV bands: Specification of maximum permitted emission levels," *in IEEE DySPAN, Aachen, Germany*, pp. 443 –454, May 2011.

[45] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, Academic Press, 2000.

[46] 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.

[47] A. H. Nuttall, "Some integrals involving the Q_M-function," *Naval Underwater System Center $(NUSC)$ Technical Report*, May 1974.

[48] 3GPP TS 25.814, "Physical layer aspects for evolved universal terrestrial radio access," vol. Rel. 7, 2007.

[49] Final Report on Link Level and System Level Channel Models, "IST-2003-507581 WINNER D 5.4," vol. V 1.4, no. Available: http://www.ist-winner.org, Nov. 2005.

[50] Z. Lu, Y. Ma, and R. Tafazolli, "A first-order cyclostationarity based energy detection approach for non-cooperative spectrum sensing," *in IEEE PIMRC, Istanbul, Turkey*, pp. 554 –559, Sep. 2010.

[51] P. Cheraghi, Y. Ma, and R. Tafazolli, "Frequency-domain differential energy detection based on extreme statistics for OFDM source sensing," *in IEEE VTC, Budapest, Hungary*, pp. 1 –5, May 2011.

Rahim Tafazolli (SM'09) is the Director of Centre for Communication Systems Research (CCSR), the University of Surrey, U.K. He has been active in research for over 20 years and has authored and co-authored more than 360 papers in refereed international journals and conferences. Professor Tafazolli is consultant to many mobile companies, has lectured, Chaired and being invited as keynote speaker to a number of IEE and IEEE workshops and conferences. He has been technical advisor to many mobile companies, all in the field of mobile communications. He is the Founder and past Chairman of IEE International Conference on 3rd Generation Mobile Communications. He is Chairman of EU Expert Group on Mobile Platform (e-mobility SRA) and Chairman of Post-IP working group in e-mobility, past Chairman of WG3 of WWRF, and editor of the Book Technologies for the Wireless Future Vol.1 2004 and Vol. 2 2006. He is nationally and internationally known in the field of mobile communications.

Parisa Cheraghi (S'05) received the B.Eng. (Hons.) degree in electronics and electrical engineering from the University of Westminster, London, U.K., in 2009.
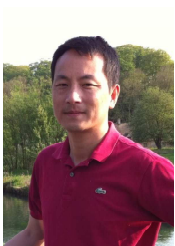
She is currently pursuing the Ph.D. degree in electrical engineering within the Centre for Communication Systems Research (CCSR), University of Surrey, U.K. Her main research interests are in the areas of signal processing advances in wireless communications, estimation and detection theory, mathematical modelling, and speech processing. She is particularly interested in fundamental research problems in dynamic spectrum sharing.

Zhengwei Lu (S'10) received his B.Eng. degree in electronic information engineering from North University of China, in 2006, and his M.Sc. degree in data communications from University of Sheffield, U.K., in 2007.

He is currently working towards the Ph.D. degree at Centre for Communication Systems Research (CCSR), University of Surrey, U.K. His research interests are mainly spectrum sensing for cognitive radios and statistical signal processing for wireless communications.

Yi Ma (M'04-SM'10) received the Ph.D. degree in electrical engineering and electronics from University of Liverpool, U.K., in 2004.

He joined the Centre for Communication Systems Research (CCSR), University of Surrey, Guildford, U.K., in 2004, where he is currently a Lecturer. He is leading a research team within the CCSR to conduct the basic-level communication systems research covering signal processing and applied information theory for wireless communications. He is the Chair of Air-Interface Club within the CCSR. He has authored and co-authored more than 70 peer reviewed international journal and conference papers in the areas of cooperative communications, cognitive radios, cooperative localization, radio resource allocation, MIMO, estimation, synchronization, modulation and detection techniques.