Research Article

Analysis of the Consecutive Mean Excision Algorithms

Johanna Vartiainen, Janne Lehtomäki, Harri Saarnisaari, and Markku Juntti

Centre for Wireless Communications (CWC), University of Oulu, P.O. Box 4500, Oulu 90014, Finland

Correspondence should be addressed to Johanna Vartiainen, jvartiai@ee.oulu.fi

The backward and forward consecutive mean excision (CME/FCME) algorithms are diagnostic methods for outlier (signal) detection. Since they are computationally simple, they have applications for both narrowband signal detection in cognitive radios and interference suppression. In this paper, a theoretical performance analysis framework of the CME algorithms is presented. The analysis provides simple tests of the detectability of the signals based on their shape in the considered domain (e.g., spectrum). As a consequence, results can be used to quickly check whether the CME/FCME algorithms are usable for a given problem or not without the need to resort to time consuming computer simulations. The computer simulations for random and orthogonal frequency division multiplexing (OFDM) signals show that the presented analysis is able to predict the detectability of signals well.

1. Introduction

Real-world data may contain samples that differ from the majority of data. These observations are called outliers [1–5]. In wireless communications, these coexisting samples are typically caused by signals that are narrow in the considered domain. Thus, narrowband (NB) interference suppression and NB signal detection methods can be classified to be outlier detection methods. Therefore, outlier detection is applicable also in cognitive radios to detect spectrum holes, that is, unused frequency bands [6, 7]. Additionally, outlier detection can be used to find time domain pulses/impulses which have relatively short duration compared with the inspection interval.

Outlier detection is usually based on some metric which is used to decide if a sample is an outlier or not. In the diagnostic outlier detection [8], the basic idea is to investigate normalized samples: the outliers are far from the mean as illustrated in Figure 1. A classical metric is the Mahalanobis squared distance (MSD) [9]

$$\chi_i = (\mathbf{r} - \bar{\mathbf{r}})^H \mathbf{C}^{-1} (\mathbf{r} - \bar{\mathbf{r}}) > T, \qquad (1)$$

where $\mathbf{r}$ is a sample vector, $\mathbf{C}$ is the shape parameter (usually the covariance matrix of $\mathbf{r}$), and $\bar{\mathbf{r}}$ is the location parameter (usually the mean of $\mathbf{r}$). The sample vector is classified as

an outlier if the MSD is larger than some predetermined threshold value T called the cutoff point. A threshold that separates the outliers from other samples can be solved using the statistics of (1). The main problem is that the mean and covariance are usually unknown and have to be estimated. Furthermore, the outliers can affect these estimates and that leads to unreliable results. This calls for iterative outlier detection and/or robust parameter estimation. Luckily, in wireless communications the signals are usually zero mean and the noise is white, that is, the covariance matrix is a scaled identity matrix which simplifies the procedure [10]. These assumptions are applied in this paper.

There exist several outlier detection methods, from which the iteratively operated consecutive mean excision (CME) algorithms [11–14] are among the most interesting methods. The CME algorithms are able to operate in any transform domain. The only requirement is that the outliers are from a signal that is "narrow" (i.e., *concentrated*) in the considered domain, for example, in the time or frequency or some other domain. Besides being computationally attractive, the CME methods operate blindly (i.e., unsupervised) without *a priori* knowledge about the noise level, the number of the concentrated signals or their characteristics. The CME algorithms can be seen to be unsupervised classification algorithms. In unsupervised classification, the observed
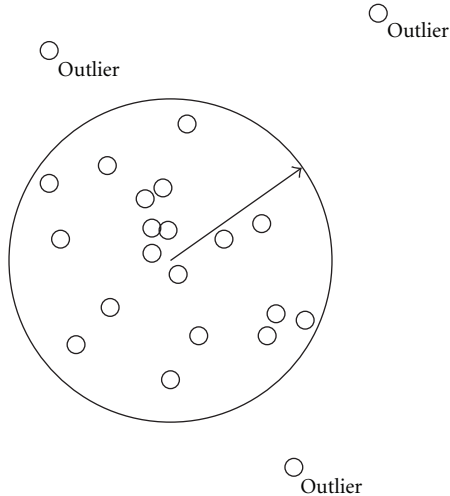
Figure 1: Outlier detection.

samples are divided into different subsets based on the properties of the samples, such that samples in the same cluster are, in some sense, similar to each other. Typically, the CME and FCME algorithms are used with magnitude-squared samples or with radiometer/energy detector outputs. Thus, the CME algorithms can be seen to correspond to simple energy detectors measuring the energy of the received samples. Thus, the CME algorithms are effective regardless of the NB signal type, used modulation or frequency-shifting. In addition, the CME algorithms are able to operate in any frequency range (i.e., from kHz to GHz) [13]. Both the CME and FCME methods and their applications have been investigated, for example, for concentrated interference suppression both in the time and frequency domains, and for narrowband signal detection in the frequency domain both in military and civilian applications [14–17]. The performance of the CME and FCME algorithms has already been compared to each other and to other methods, for example, in [12–15, 18, 19]. Therein, based on the statistical properties of the methods, it has been observed that the FCME algorithm outperforms the CME algorithm. It has also been found out that the performance of the simpler CME algorithm is adequate when the signal is very concentrated [8, 11]. Moreover, an example for real-world frequency-shifted signals is presented in [18]. Theoretical impulse detection performance of the CME and FCME algorithms was analyzed in [11]. There, the statistical analysis led up to the results of only sample-based probability of detection. No detection limits were derived. Asymptotic threshold setting for the FCME algorithm with the Welch spectrum estimator was presented in [20]. A simplified analysis of the CME algorithms have been considered in [19] as an example case. Therein, the signal consisted of only one lobe and no general detection limits were defined. In [21], some simple rules when a signal is detectable were considered briefly. The simulations were performed for random signals only. Enhancement of the CME algorithms called the localization algorithm based on double-thresholding (LAD) method was

proposed in [16]. The LAD method uses two thresholds and is able to localize the narrowband signal samples in the frequency. The performance analysis of the LAD method was considered in [22]. Therein, the clean sample rejection and detection rates were analyzed. The optimal upper and lower threshold values for the LAD method were analytically confirmed. Note that the analysis results presented in [22] are valid only for the LAD method, not for the CME algorithms. The LAD method has been investigated, for example, for signal detection in the frequency domain including spectrum sensing in cognitive radios [16, 17]. The LAD method has been implemented on the wireless open-access research platform (WARP) in [23]. Therein, it was noticed that the LAD method is able to sense the spectrum. In this paper, the performance of the CME algorithms is analyzed more widely and detailed. The analysis is based on signals *shape* in the considered domain. For example, in the frequency domain, the shape corresponds to spectrum. The main goal is to analytically characterize the conditions at which the CME algorithms find the outliers, that is, concentrated signals, and provide an easy-to-use tool for checking the detectability of signals by the CME algorithms. This leads to equations from which the detection limits can be derived. Herein, the term *detection limit* denotes at which detection parameters (height, width, and threshold parameter) the signal can be detected. The aim is not to compare the methods because the comparisons have been addressed in several papers. Further, the signal-to-noise ratio (SNR) values that limit the signal detection are derived. According to the authors best knowledge, the analytic detection limits for the CME and the FCME algorithms have not been presented earlier. Extensive simulations with random and orthogonal frequency division multiplexing (OFDM) signals confirm the validity of the analysis. It will be seen that the presented analysis leads to the simple limits of detectability. Therefore, the results of the analysis can be used, for example, in cognitive radios for fast checking if a signal is detectable or not without time-consuming simulations. A typical example is detecting future mobile digital video broadcasting Handheld (DVB-H) systems.

This paper is organized as follows. In Section 2, the CME algorithms are described. In Section 3, the CME and FCME algorithms are analyzed and general detection limits as well as detailed information about the detection alternatives are derived. Numerical results are presented in Section 4, and conclusions are drawn in Section 5.

2. Consecutive Mean Excision Algorithms

The considered received signal consists of outliers (i.e., signals to be detected), the noise, and a possible noise-like signal which is below the noise level (such as wideband signal in the frequency domain). Therein, the noise and noise-like signals form the base signal. The considered ith signal sample is assumed to be $r_i = c_i + s_i + \eta_i$, where c_i is a concentrated signal sample caused by outliers, s_i is a possible noise-like signal sample, and η_i is the noise process. Therein, the concentrated signal consists of $n \geq 0$ lobes. Note that the

notations are general and not specified to some particular domain, that is, samples r_i can be in time, frequency or in some other domain.

The observed ith scalar sample r_i, corresponding now to the base data without outliers, is assumed to be zero mean, independent, identically distributed random variable, that is, the covariance is a scaled identity matrix [10]. This means that (1) reduces to

$$\zeta_i = |r_i|^2 > T\overline{\zeta} = T_h, \qquad (2)$$

where

$$\overline{\zeta} = \frac{1}{Q}\sum_{i=1}^{Q}\zeta_i \qquad (3)$$

is an estimator corresponding to the shape parameter, $T > 1$ is a threshold, multiplier, T_h denotes the threshold and Q is the size of the so-called clean set (in the first iteration called an initial set).

The value of the threshold multiplier T can be calculated based on the desired false alarm probability $P_{FA,DES}$ in the outlier-free case [11]. Since the threshold depends on samples, the CME algorithms can be considered constant false alarm rate (CFAR) type detectors that often use the so-called reference samples to find the threshold. Gaussianity is a generally used noise model in communication [10]. For example, if it is assumed that noise sample r_i is a complex Gaussian variable, ζ_i follows a chi-squared distribution with two degrees of freedom. Thus, assuming that the sample mean converges to the actual value (ergodicy), the probability that a normalized sample $\xi_i = \zeta_i/\overline{\zeta}$ exceeds T is $P(\zeta_i > T\overline{\zeta}) = e^{-T}$. From that, $T = -\ln(P_{FA,DES})$. When the degrees of freedom is >2 (e.g., in a multiantenna case), the proper value of T has been determined in [14]. Respective values for T can be also defined for other distributions, for example, when the Welch spectrum estimator is used [20].

The CME algorithm operates backward. In the first iteration of the backward CME algorithm, estimate $\overline{\zeta}$ is based on all the samples, that is, $Q = N$. The CME algorithm operates, in this case, using the current mean energy of the samples ζ_i and by multiplying that mean value with the threshold multiplier T. In every iteration, sample r_i is rejected from the clean set Q if its energy $\zeta_i > T\overline{\zeta}$, after which the mean (energy) is estimated again from the remaining set. This continues until new outliers cannot be found.

Unlike the CME algorithm, the FCME algorithm operates forward. First, the samples are rearranged in an ascending order according to their energies. The sorting can be done, for example, using Heapsort or Quicksort, whose average computational complexities are $N\log_2 N$ and $2\log_2 N$, respectively, [24]. After that, the FCME algorithm calculates the mean $\overline{\zeta}$ of a small initial set consisting of samples that are the smallest in energy and are assumed to be free of the outliers. The size of the initial set is usually selected so that it is about 10% of the total data set [11]. Too small an initial set may lead to the situation that the algorithm does not converge. On the other hand, if the initial set is too large, it may be possible that it contains outlier samples so the

algorithm does not operate properly. In the first iteration, $Q = N_c$, where N_c is the number of samples in the initial set. The FCME algorithm iteratively calculates a new value for the mean $\overline{\zeta}$ and a new threshold until there are no samples below the threshold. That is, in every iteration, sample r_i is added to the clean set Q if $\zeta_i < T\overline{\zeta}$.

The number of considered samples has an effect on the performance of the CME algorithms. Usually, 1024 or 512 samples have been used [11, 19]. It has been noticed that the CME algorithms are able to operate properly when the number of samples is 256 or more. The more samples there are, the more samples are used when calculating the mean of the whole data set (CME)/initial set (FCME), so single strong outliers do not have very strong effect on the estimate of the mean [3]. In addition, if the number of samples is small, the initial set in the FCME algorithm may come too small, and the algorithm may not start to operate properly at all.

3. Performance Analysis

The aim of the performance analysis is to find simple rules for the outlier detection capability of the CME algorithms. Therefore, several simplifications have to be made, but their validity is confirmed by the computer simulations in Section 4. For simplicity and without the loss of generality, the signal samples are sorted in a descending order according to their heights (magnitudes) so that the sample with the largest magnitude is the first sample, and so on, until the sample with the lowest magnitude is the last sample. The purpose of the reordering is to clarify the analysis. Reordering does not have any impact on the analysis results nor the operation of the CME algorithms, because the CME algorithm does not care what the order of the samples is, and the FCME algorithm reorders the samples anyway. For those reasons, the analysis is valid also when the samples are not reordered. Subsequently, the signal samples are divided into different parts or *lobes* according to their amplitudes so that every lobe consists of signal samples with equal amplitude. The width of the lobe states how many samples with equal amplitude the lobe includes, while the height of the lobe reflects the sample magnitude. Thus, one lobe consists of samples (i.e., bins) which have exactly equal height (i.e., amplitude or magnitude). The presented shape-based analysis corresponds to the power spectral density (PSD) in the frequency domain, where one lobe includes the frequency domain signal samples with equal energy. A simplified example of reordering and developing the signal lobes is shown in Figure 2. In real life, the amplitudes of the lobe samples are not equal, because the probability that two realizations of continuous valued random variables are equal is zero. Therefore, average magnitude, that is, the mean of the lobe samples, may present the amplitude of the lobe.

Assume that the base signal has width W and height S. The outlier signal consists of $n \geq 1$ lobes with widths W_i, $\sum_{i=1}^{n} W_i \leq W$, and heights S_i, $i = 1, \ldots, n$, $S_1 > S_2 > \cdots > S_n$. The first signal lobe $S_1 W_1$ is called the main lobe and the other $n - 1$ lobes are the side lobes. The widths and heights of the lobes can be presented to be relative to the width and
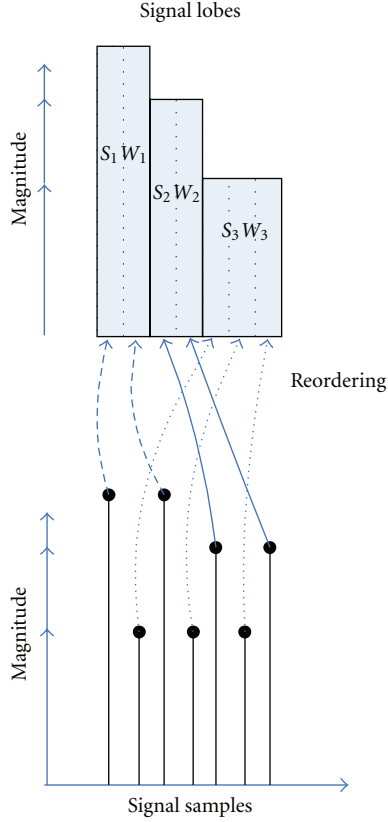
FIGURE 2: A simplified example of reordering and developing of signal lobes. There are 7 signal samples whose magnitudes are in three different levels. This leads to three signal lobes.

the height of the base signal, respectively. Thus, $S_i = \alpha_i S$ and $W_i = \beta_i W$, or $\alpha_i = S_i/S$ and $\beta_i = W_i/W$, $\alpha_i, \beta_i > 0$. For example, when the CME algorithm is used and $n = 3$ (Figure 3), we get from (2) that only the main lobe S_1 is detected in the first iteration if

$$S_1 + S > T\bar{\zeta} = T\left(\frac{SW + \sum_{i=1}^{3} S_i W_i}{W}\right) \qquad (4)$$

and if

$$S_2 + S < T\left(\frac{SW + \sum_{i=1}^{3} S_i W_i}{W}\right). \qquad (5)$$

In the practical applications, the relative signal powers are usually of interest. The power of the base signal is SW, that is, S presents the "density". The ith lobe power is $(S + S_i)W_i$. When S is the noise, SNR can be defined to be

$$\gamma = \frac{\sum_{i=1}^{n}(S + S_i)W_i}{SW} = \sum_{i=1}^{n}(1 + \alpha_i)\beta_i. \qquad (6)$$

Here, SNR is defined per total bandwidth. SNR could also be defined per NB signals bandwidth.

There are three different alternatives for the detection. The extreme alternatives are that (a1) all the lobes are detected at the same time or (a2) all the lobes are detected one-by-one. The third alternative includes all other possible
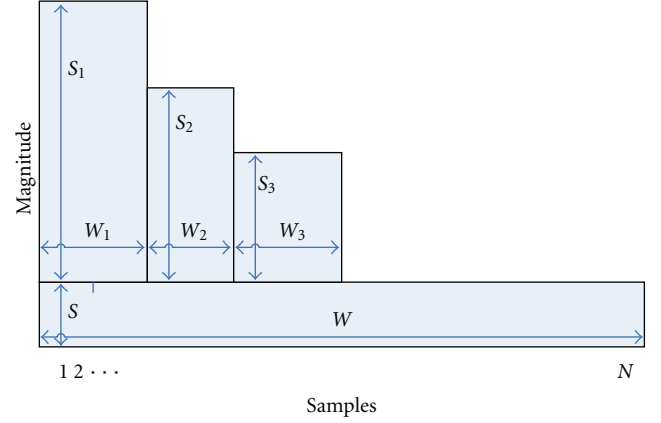

FIGURE 3: An example of base signal and three signal lobes. The total number of samples is N.

lobe detection combinations (a3). Naturally, it is also possible that only some of these alternatives are possible, or that the lobes are totally undetectable. However, it is not always necessary to detect all the lobes in practice. Instead, depending on the used application, the detection of only the highest(s) lobe(s) may be sufficient. In practice, the samples that are decided to belong to the signal can be defined, for example, using 3 dB bandwidth.

It appears that detectability can be solved with respect to the heights of the lobes and the used threshold multiplier. They depend on each other and their relationship is of interest. Next, the conditions at which the outlier signal can be detected are analysed in terms of the heights of the lobes and threshold multiplier. The main focus is to get conditions at which the CME and FCME algorithms find the signals. The different detection alternatives give additional information in what way the detection is performed.

Let $k = 0, \ldots, n - 1$ denote the number of lobes rejected (for the CME algorithm) or added (for the FCME algorithm) from/to the clean set Q in the previous iterations. Let $j = 1, \ldots, n$ denote the step size, that is, how many lobes are rejected/added at one iteration, and $k + j \leq n$. In our examples we have selected to use threshold multiplier values $T = 2.303, 4.6052$, and 6.9078 which have been noticed to be proper choices, for example, in [14, 18], and which have been used in the earlier papers. However, from the analysis results, the detection limits can be calculated using any other desired threshold multiplier values.

3.1. Analysis for the CME Algorithm. First, the effect of the heights of the detected lobes is considered. Using a geometrical approach for defining $\bar{\zeta}$, we get from (2) that the conditions for the heights of the detected lobes can be expressed using function

$$f(k, j) = \frac{T - 1 + T\sum_{i=k+1, i \neq k+j}^{n}\left(\alpha_i \beta_i / \left(1 - \sum_{i=1}^{k} \beta_i\right)\right)}{1 - T\left(\beta_{k+j} / \left(1 - \sum_{i=1}^{k} \beta_i\right)\right)}. \qquad (7)$$

The $(k + j)$th lobe $(S_{k+j} W_{k+j})$ is detected if $\alpha_{k+j} > f(k, j)$ and if $\alpha_{k+j+1} < f(k, j + 1)$. For example, in the first iteration,

only the highest lobe, that is, the main lobe (S_1W_1) is detected if $\alpha_1 > f(0,1)$ and if $\alpha_2 < f(0,2)$, as can be seen by (4) and (5), respectively. The corresponding alternatives for the detection as a function of the heights of the lobes are (a1) all the lobes are detected in the same (first) iteration when $\alpha_n > f(0,n)$, and (a2) all the lobes are detected one-by-one when $\alpha_{k+1} > f(k,1)$ and $\alpha_{k+2} < f(k,2)$ for all k. Importantly, the detection is not possible at all if

$$\alpha_1 < f(0,1). \tag{8}$$

In the special case when there are only the noise and the main lobe (S_1W_1), that is, $n = 1$, the main lobe is detected if $\alpha_1 > f(0,1) = (T-1)/(1-T\beta_1)$. This coincides with the result shown in [19].

The conditions for the used threshold multiplier T are considered next. From (2) we can define another function

$$g(k,j) = \frac{1 + \alpha_{k+j}}{1 + \left(\sum_{i=k+1}^{n} \alpha_i\beta_i / \left(1 - \sum_{i=1}^{k}\beta_i\right)\right)}. \tag{9}$$

The $(k+j)$th lobe ($S_{k+j}W_{k+j}$) is detected next if $g(k,j+1) < T < g(k,j)$. Because in every iteration the samples with the strongest energy are removed, the mean decreases as k increases. For example, in the first iteration, only the highest lobe, that is, the main lobe (S_1W_1) is detected if $g(0,2) < T < g(0,1)$.

From (9), the alternatives for the detection as a function of T are (a1) $1 < T < g(0,n)$ and (a2) $g(k,2) < T < g(k,1)$ for all k. Because $\alpha_i > \alpha_{i+1}$, we get from the above-mentioned equations that $g(k,2) > g(0,n)$. Therefore, (a3) is achieved when $g(0,n) < T < g(k,2)$. In other words, the smaller the threshold is, the more lobes are detected at the same time. However, small T leads to more false detections. The detection is not possible at all if

$$T > g(0,1). \tag{10}$$

Let us next consider the case when the detection is not possible at all. In general, from the denominator of (8), the limiting value is $\beta_1 > 1/T$. This means that the detection is impossible if (8) holds and if $\beta_1 \geq 0.4343$ (for $T = 2.303$), $\beta_1 \geq 0.217$ (for $T = 4.6052$) and $\beta_1 \geq 0.145$ (for $T = 6.9078$). This result holds in general regardless of the value of n. Since β_1 means the width of the main lobe, that main lobe can cover 43, 23, and 15% of the studied "spectrum", respectively.

Assume next that the signal has only the main lobe, that is, $n = 1$. It follows from (8) that the detection is impossible when $\alpha_1 < (T-1)/(1-T\beta_1)$, $\beta_1 > (1-T+\alpha_1)/T\alpha_1$ and $T > (1+\alpha_1)/(1+\alpha_1\beta_1)$, respectively. The values of α_1 when the signal detection is impossible via the CME algorithm are presented in Table 1 as a function of T and β_1. For example, if $T = 2.3026$ and $\beta_1 = 0.1$ (10% of the total width), the signal detection is impossible if $\alpha_1 < 1.692$. That is when the height of the signal lobe S_1 is smaller than 1.692 times the height of the base signal level S. In terms of SNR per total bandwidth, the signal detection is impossible if SNR < -6 dB. When considering SNR per NB signals bandwidth, 13 dB ($\beta_1 = 0.05$), 10 dB ($\beta_1 = 0.1$), 7 dB ($\beta_1 = 0.2$), 5 dB ($\beta_1 = 0.3$),

or 4 dB ($\beta_1 = 0.4$) should be added to the given SNR values. As seen, the narrower the signal is, the lower its height can be, that is, the lower SNR is required for the detection.

The relative height of the signal versus the relative width of the signal when detection is impossible via the CME algorithm is presented in Figure 4 such that above the curve is the area where the detection is impossible. As can be seen, the curves converge to the limiting values presented in Table 1. For example, if $T = 2.3026$ and the relative height of the signal is $\alpha_1 = 5$, the signal detection is impossible if relative width of the signal is $\beta_1 = 0.32 = 32\%$ or more.

Let us next consider the detection alternatives in more detail. Assume that the outlier signal has two lobes, that is, $n = 2$. Based on (7), there are two different possibilities to make the detection when the CME algorithm is used: (a1) Both the lobes are detected at the same time, if $\alpha_2 > f(0,2) = (T-1+T\alpha_1\beta_1)/(1-T\beta_2)$ holds, and (a2) both the lobes are detected separately if $\alpha_1 > f(0,1) = (T-1+T\alpha_2\beta_2)/(1-T\beta_1)$ and $\alpha_2 < f(0,2) = (T-1+T\alpha_1\beta_1)/(1-T\beta_2)$ holds. Next, the second lobe is detected if $\alpha_2 > f(1,1) = (T-1)/(1-T\beta_2)$. Note that in all the cases, $\alpha_1 > \alpha_2$ must hold.

Next, the case (a1) is considered. Note that $\beta_1 < 1/T$ must hold and $\alpha_1 > \alpha_2$ forms the upper limit. Let the width of the lobes be equal, that is, $\beta_1 = \beta_2$. The relative height of the main lobe α_1 versus the relative height of the second lobe α_2 when detection is possible via the CME algorithm is presented in Figures 5 and 6. Therein, the detection is possible above the curves but below the upper limit. In Figure 5, $\beta_1 = \beta_2 = 0.05$ (5%) and T varies [11]. For example, when $T = 4.6052$ and $\alpha_1 = 10$, $8 \leq \alpha_2 < 10$ in order to detect both the lobes at the same time. When $\beta_1 = \beta_2 \geq 0.1$ and $T = 4.6052$ or 6.9078, it follows that $\alpha_2 > \alpha_1$, so that only the results for $T = 2.3026$ can be presented. In Figure 6, $T = 2.3026$ and $\beta_1 = \beta_2$ varies. The case (a2) is straightforward but algebraically cumbersome because of several variables.

3.2. Analysis for the FCME Algorithm.

The analysis of the FCME algorithm is somewhat simpler than that of the CME algorithm. Let κW, $0 < \kappa < 1$ denote the size of the initial set and $W_{\text{clean}} = W - \sum_{i=1}^{n} W_i$. Assume first that the initial set is clean of outliers, that is, $\kappa W \leq W_{\text{clean}}$. Using a geometrical approach for defining $\bar{\zeta}$, we get from (2) that the mean of the clean set Q when it contains k lobes can be defined using function

$$f_F(k) = \frac{\beta_{\text{clean}} + \sum_{i=1}^{k} \beta_{n-i+1}(1 + \alpha_{n-i+1})}{\beta_{\text{clean}} + \sum_{i=1}^{k} \beta_{n-i+1}}. \tag{11}$$

In the considered iteration, the lobes $n-k-j+1, \ldots, n-k$ are added to the clean set Q if $\alpha_{n-k-j+1} < Tf_F(k)-1$ and $\alpha_{n-k-j} > Tf_F(k)-1$.

The alternatives for the detection as a function of heights of the lobes are (a1) $\alpha_n > Tf_F(0) - 1 = T - 1$ [19]. In that case, the clean set Q includes only outlier-free samples. The other alternative is that (a3) only part of the lobes, lobes $S_1W_1, \ldots, S_iW_i$, $i \in \{1, \ldots, n\}$, are detected when $\alpha_i > Tf_F(k) - 1$ and $\alpha_{i+1} < Tf_F(k) - 1$. In that case, the clean set Q includes both the outlier-free samples and some outlier

TABLE 1: The values of α_1 and SNR [dB] when signal detection via the CME algorithm is impossible with different values of T and β_1. The signal has only one lobe. Note that SNR is defined per total bandwidth.

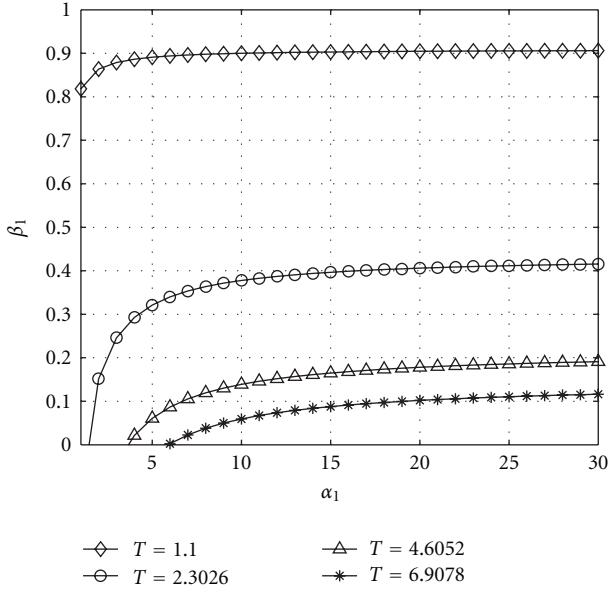| T	$\beta_1 = 0.05$ α_1 (SNR [dB])	$\beta_1 = 0.1$ α_1 (SNR [dB])	$\beta_1 = 0.2$ α_1 (SNR [dB])	$\beta_1 = 0.3$ α_1 (SNR [dB])	$\beta_1 = 0.4$ α_1 (SNR [dB])
2.3026	<1.472 (−9 dB)	<1.692 (−6 dB)	<2.415 (−2 dB)	<4.213 (2 dB)	<16.497 (8 dB)
4.6052	<4.684 (−5 dB)	<6.683 (−1 dB)	<45.659 (10 dB)	—*	—*
6.9078	<9.025 (−3 dB)	<19.106 (3 dB)	—*	—*	—*

*Detection is not possible.

FIGURE 4: The relative height of the main lobe versus the relative bandwidth of the main lobe. The signal has one lobe. The detection via the CME algorithm is impossible above the curves and possible below the curves.
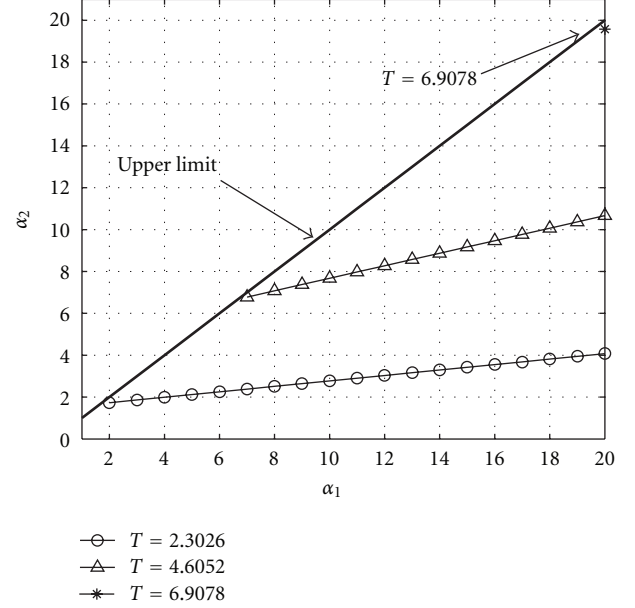

FIGURE 5: The relative height of the main lobe versus the relative height of the second lobe for joint detection of two lobes when $\beta_1 = \beta_2 = 0.05$ and $T = 2.3026, 4.6052$ or 6.9078. The limits of detection via the CME algorithm.

samples (lobes). Because the FCME algorithm operates in the forward direction increasing the clean set Q, the signal cannot be detected lobe by lobe. Furthermore, the signal cannot be detected at all when the main lobe is not detected, that is,

$$\alpha_1 < T f_F(k) - 1. \tag{12}$$

When considering the threshold multiplier T, the alternatives are (a1) $T < \alpha_n + 1$ and (a3) only part of the lobes, lobes $S_1 W_1, \ldots, S_i W_i, i \in \{1, \ldots, n\}$, are detected when $(\alpha_{i+1} + 1)/f_F(k) < T < (\alpha_i + 1)/f_F(k)$. The signal cannot be detected at all when when the threshold is larger than the main lobe, that is,

$$T > \frac{(\alpha_1 + 1)}{f_F(k)}. \tag{13}$$

The FCME algorithm is able to detect signal lobes regardless of their bandwidths when the initial set is clean, that is, when $\sum_{i=1}^{n} W_i < W - \kappa W$, assuming sufficient threshold multiplier T. Usually, the initial set size has been about 10% of the considered data set (samples), so the FCME algorithm is able to detect signals with widths less than 90% of the considered samples. Because the FCME algorithm is a concentrated signal detection method, it can be assumed that the initial set is usually clean.

If the initial set is not clean, it includes some samples from the signal lobes. Thus, the mean of the initial set is higher than the ideal value leading to too high a threshold. In that case, the signal lobes may be below the threshold, and, thus, the detection may fail. Assume that the number of iterations is one and the initial set includes part of the samples from hth lobe $(S_h W_h)$, $h \in \{1, \ldots, n\}$, and all the samples of smaller lobes $S_{h+1} W_{h+1}, \ldots, S_n W_n$. The initial set size is $\kappa W = W - \sum_{i=1}^{h} W_i + \nu W_h$, where $0 < \nu \le 1$. For the detection, there are two possibilities: the lobe $S_h W_h$ or lobe $S_{h+1} W_{h+1}$ is detected. Let

$$g_F(h) = \frac{\beta_{\text{clean}} + \sum_{i=h+1}^{n} \beta_i(1 + \alpha_i) + \nu \alpha_h \beta_h}{\beta_{\text{clean}} + \sum_{i=h+1}^{n} \beta_i + \nu \beta_h} \tag{14}$$

be the mean of the initial set. The hth lobe $(S_h W_h)$ is detected if $\alpha_h > T g_F(h) - 1$ and $\alpha_{h+1} < T g_F(h) - 1$. The $(h + 1)$st lobe $(S_{h+1} W_{h+1})$ is detected if $\alpha_{h+1} > T g_F(h) - 1$. It can be observed
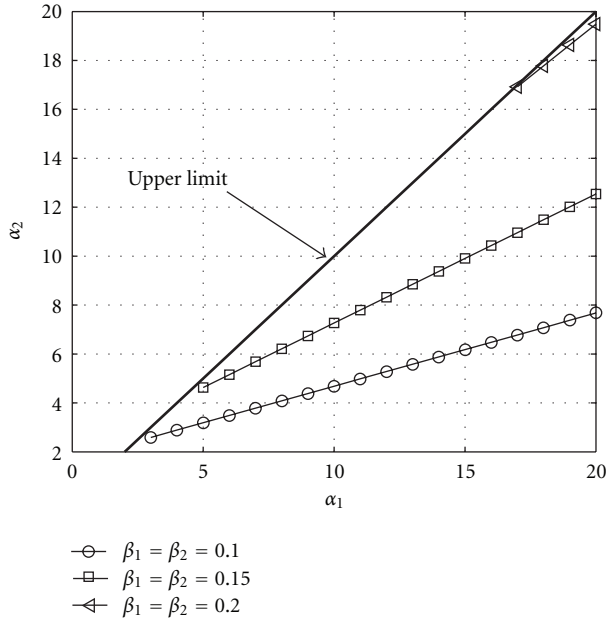
FIGURE 6: The relative height of the main lobe versus the relative height of the second lobe for joint detection of two lobes when $T = 2.3026$ and $\beta_1 = \beta_2 = 0.1$, 0.15 or 0.2. The limits of detection via the CME algorithm.

that the threshold for the hth lobe detection is higher than the threshold for the $(h + 1)$st lobe detection, as expected. In addition to, $g_F(h) > f_F(k)$.

Assume that there is one very high main lobe and $S \to 0$. Let us also assume that the initial set includes x samples of the main lobe. Consequently, from (14) we get that the main lobe can be detected if $S_1 > T(xS_1/W_{\text{init}})$, where W_{init} is the width of the initial set. It means that the main lobe can be detected if the initial set includes $x < W_{\text{init}}/T$ samples from the main lobe.

Next, the analysis results for the FCME algorithm are considered in the case when the detection is impossible. Assume that the initial set is clean of outliers. When the signal has only one lobe, it follows from (12) that the detection is impossible when $\alpha_1 < T - 1$ and $T > \alpha_1 + 1$. That means that the width of the signal lobe does not matter if the initial set is clean. For example, when $T = 2.3036$, 4.6052, and 6.9078, the main lobe cannot be detected if $\alpha_1 < 1.3026$, 3.6052 and 5.9078, respectively. The values of α_1 when signal detection is impossible via the FCME algorithm are presented in Table 2 as a function of T and β_1. Therein, the SNR values are per total bandwidth. When SNR is defined per the bandwidth of the NB signal, the detection is impossible when $\alpha_1 < 4$ dB ($T = 2.3036$), 6 dB ($T = 4.6052$), and 8 dB ($T = 6.9078$) with all values of β_i. Note that unlike in the case of the CME algorithm, the width of the signal does not matter when detection is performed via the FCME algorithm.

Next, the detection alternatives are considered more closely in the case of two lobes. Based on (11) there are two possible alternatives to make the detection: (a1) $\alpha_2 > T - 1$, that is, the detection depends only for the threshold

multiplier, and (a3) only the main lobe is detected, that is, if $\alpha_1 > T((\beta_{\text{clean}} + \beta_2(1 + \alpha_2))/(\beta_{\text{clean}} + \beta_2)) - 1 > \alpha_2$, that is, only the main lobe is detected if the clean set Q contains all other samples except the samples from the main lobe. It should be noted that the equation in the case (a1) is almost similar than the equation in the case where the signal has only one lobe ($\alpha_1 < T - 1$, results shown in Table 2).

The summary of the detection alternatives for both the CME and FCME algorithms is presented as a function of α and T at Table 3. As (a1)–(a3) give detailed information about the detection, alternative "impossible detection" gives the general limits when the signal is detectable and when it is not. The results in Tables 1 and 2 are derived from that.

4. Numerical Results

The theoretical detection limits derived in Section 3 were confirmed via computer simulations. The simulations were performed in the frequency domain for random and OFDM signals. In the first case, the channelized radiometer [25] was used, whereas in the latter case detection was performed using the Welch spectrum estimator which uses windowing and overlapping [26]. In practice, it may be the case that all the signal samples have different amplitudes. Thus, there are two possibilities: either one signal sample corresponds to one lobe, that is, the number of the signal samples equals the number of the lobes, or one lobe corresponds to all the signal samples, and the mean of the samples in that lobe presents the height (i.e., the energy in the frequency domain) of the lobe. Here, we have selected the latter approach.

4.1. Random Signal. The channelized radiometer uses several parallel total power radiometer receivers, that is, it integrates energy into several frequency bands simultaneously. It can be used, for example, in spectrum sensing. In total V radiometer channels were assumed. Each has integration time T_R and bandwidth W_R. One channel in the radiometer corresponds to one frequency domain sample. The noise was Gaussian. When there is only noise present, the channelized radiometer output follows the chi-square distribution with $\Theta = 2T_RW_R$ degrees of freedom. When there is both the signal and noise present, the output follows the noncentral chi-square distribution with Θ degrees of freedom with noncentrality parameter $\Gamma = 2E_i/N_0$, where E_i is the energy of the signal in the ith radiometer, $i = 1, 2, \ldots, V$ and N_0 is the noise power density. Thus, the SNR [dB] is $\gamma = 10\log_{10}(\Gamma/2)$. The average output value is $S = 2T_RW_R$ for the noise-only case and $2T_RW_R + \Gamma$ for the signal + noise case. Thus, the height of the ith lobe S_i is Γ and $\alpha_i = S_i/S = \Gamma/2T_RW_R$. Assume that there is one lobe with the noncentrality parameter Γ and the width of the lobe equals one radiometer channel (i.e., one sample). It follows from (10) that the signal is detected if

$$T < \frac{1 + \Gamma/2T_RW_R}{1 + \Gamma/2T_RW_RN}. \tag{15}$$

Table 2: The values of α_1 and SNR [dB] when signal detection via the FCME algorithm is impossible with different values of T and β_1. The signal has only one lobe. Note that SNR is defined per total bandwidth.

T	$\beta_1 = 0.05$ α_1 (SNR [dB])	$\beta_1 = 0.1$ α_1 (SNR [dB])	$\beta_1 = 0.2$ α_1 (SNR [dB])	$\beta_1 = 0.3$ α_1 (SNR [dB])	$\beta_1 = 0.4$ α_1 (SNR [dB])
2.3026	<1.303 (−9 dB)	<1.303 (−6 dB)	<1.303 (−3 dB)	<1.303 (−2 dB)	<1.303 (0 dB)
4.6052	<3.605 (−6 dB)	<3.605 (−3 dB)	<3.605 (0 dB)	<3.605 (2 dB)	<3.605 (3 dB)
6.9078	<5.908 (−5 dB)	<5.908 (−2 dB)	<5.908 (1 dB)	<5.908 (3 dB)	<5.908 (4 dB)

Table 3: Detection alternatives for the CME and FCME algorithms as a function of the relative height of the ith lobe α_i, $i = 1, \ldots, n$, and threshold multiplier T. The used functions are $f(k, j)$ (7), $g(k, j)$ (9), and $f_F(f)$ (11). Alternatives for the detection: all the lobes are detected (a1) at the same time, (a2) one-by-one, and (a3) other possible alternations. Alternative "impossible detection" gives general limits when the signal is detectable and when it is not.

	CME algorithm	FCME algorithm
(a1)	$\alpha_n > f(0, n)$	$\alpha_n > T - 1$
	$1 < T < g(0, n)$	$T < \alpha_n + 1$
(a2)	$\alpha_{k+1} > f(k, 1)$ and $\alpha_{k+1} < f(k, 2)$ for all k	—*
	$g(k, 2) < T < g(k, 1)$ for all k	—*
(a3)		$\alpha_i > T f_F(k) - 1$ and $\alpha_{i+1} < T f_F(k) - 1$**
	$g(0, n) < T < g(k, 2)$	$(\alpha_{i+1} + 1)/f_F(k) < T < (\alpha_i + 1)/f_F(k)$**
Impossible detection	$\alpha_i < f(0, 1)$	$\alpha_1 < T f_F(k) - 1$
	$T > g(0, 1)$	$T > (\alpha_1 + 1)/f_F(k)$

* Cannot be detected one-by-one.
** Lobes $S_1 W_1, \ldots, S_i W_i$ are detected.

The probability of finding the random signal for different values of T is presented in Figure 7. The theoretical detection limits are based on (10) and (13) for the CME and FCME algorithms, respectively. The total bandwidth is $W = 64$ samples, so the bandwidth of the signal is 1.6% of the system's bandwidth. Two cases are studied: the degrees of freedom is $\Theta = 10$ or 1000, and SNR is 7 dB or 34 dB, respectively. When SNR is 7 dB and $\Theta = 10$, only the results for the CME algorithm are presented because the results for the FCME algorithm are equivalent. When $\Theta = 10$ and SNR is 7 dB, the simulation results do not correspond to the theory as well as when $\Theta = 1000$ and SNR is 34 dB, where the simulation results and theory match very well, that is, the simple theoretical rules predict detectability (100%) quite reliably. It can be noticed that the higher SNR and Θ are, the better the theory and simulation results will match. In the case of smaller SNR = 7 dB and higher $\Theta = 1000$, the theoretical limit for detection is $T = 1.01$, and the simulated detection probability achieves 50% when $T = 1.04$. Instead, with higher SNR = 34 dB and smaller $\Theta = 10$, the theoretical limit for detection is $T = 56.8$ (CME algorithm) and the simulated detection probability drops to 50% when $T = 57$. It can also be seen that in all of the studied cases, the theoretical detection limits in terms of T correspond to around 50% probability of detection in the simulations. When SNR is 7 dB and $\Theta = 10$, the theoretical limit of detection using the CME algorithm is about $T = 1.97$ (10). This limit corresponds to the false alarm probability about 5% as can be seen from Figure 8, that shows the value of T versus the probability of false alarm for both the CME and FCME algorithms. The theoretical curve was calculated in the noise-only case and the simulations were

performed when both the signal and noise were present, but only noise-only radiometer channels were investigated. The larger W is, the better the results coincide with the theory. When the number of samples is small (e.g., $W = 10$), the obtained false alarm rate will differ from the nominal one. In that case, the FCME algorithm performs worse than the CME algorithm. This is because the size of the initial set of the FCME algorithm is too small (only one sample), and, consequently, the FCME algorithm does not converge. Obviously, the initial set size should be larger.

Next, a more realistic case when the width of the main lobe is more than one sample is studied. Herein, the mean of the lobe samples presents the amplitude of the lobe α_1. The probability of finding the signal for different values of T using the CME and FCME algorithms is presented in Figure 9. Here, $\Theta = 1000$, $\gamma = 34$ dB, $W = 64$, and the width of the main lobe is 4, 10, or 20 samples, that is, 6.25%, 15.5%, or 31.25% of the system's bandwidth, respectively. The theoretical detection limits were calculated based on (10) and (13). It can be seen that the theory and simulations match very well.

4.2. OFDM Signal. OFDM systems are used, for example, in high data rate applications, as in wireless local area networks (WLAN), and their detection is of interest in cognitive systems. SNR is defined to be per OFDM symbol in the whole bandwidth. The OFDM signal is, on average, well concentrated having only one lobe with relative height of $\alpha_1 = S_1/S = \gamma W/W_1$ and relative width of $\beta_1 = W_1/W$. Hence, the detection is possible if (9) $T < (1 + \alpha_1)/(1 + \alpha_1 \beta_1) = (1 + \gamma(W/W_1)/(1 + \gamma)$. When SNR γ is large, this reduces to $T < W/W_1$.
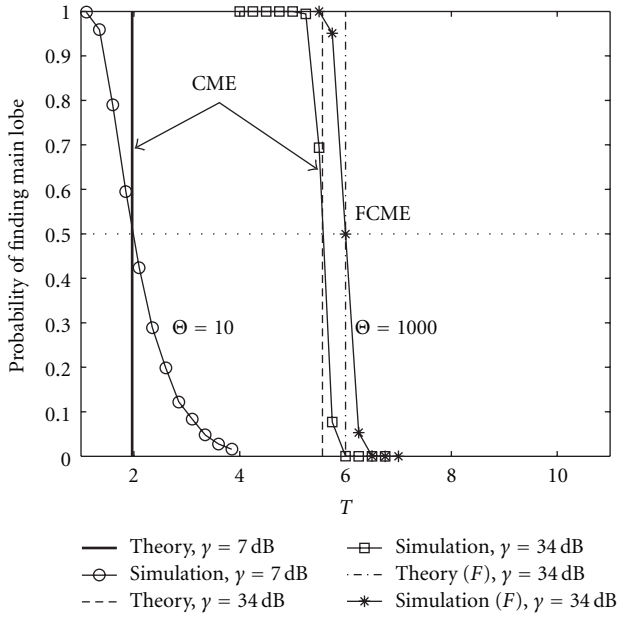
FIGURE 7: Threshold multiplier T versus the probability of finding the signal using the CME and FCME algorithms. Random signal and channelized radiometer. $\gamma = 7$ or 34 dB, $\Theta = 10$ or 1000, and $W = 64$. Signal bandwidth is 1.6% of the system's bandwidth.

FIGURE 8: Threshold multiplier T versus the probability of false alarm using the CME and FCME algorithms. Random signal and channelized radiometer. $\Theta = 10$, $\Gamma = 10$, $W = 10$ or 64.

In the simulations the system bandwidth is $W = 512$ samples, there is one OFDM signal with 68 active subcarriers and the prefix length P is 12.5% of the length of the total OFDM symbol length as, for example, in DVB-T (Terrestrial) systems, that is, oversampled OFDM signals. For these parameters, the theoretical limit of detection in the case of the CME algorithm is $T = 5.29$ (10). The Welch spectrum estimator with 50% overlapping was used. The length of the FFT was 512 samples. The signal has only the main lobe and W_1 is more than one sample. Average amplitude presents the amplitude of the lobe α_1.

The probability of finding all the samples of the main lobe as a function of the threshold multiplier T is presented in Figure 10. The theoretical detection limits that will express the upper bound of T (vertical lines) are based on (10). Here, the CME algorithm is used with two SNR values, $\gamma = 2$ and 10 dB. Signal bandwidth is 13% of the systems's bandwidth. Furthermore, $L = 1000$, that is, long averaging is used. It can be seen that the simulated detection limits are somewhat higher than the theoretical ones. This is mainly because the used signal is real-life, that is, random, or nonblock lobe like. However, the difference is not large and the theoretical limit explains rather well 100% detection point. The higher L is, the better the simulation results and theory will match each other. That is because the analysis is for flat signals, and the more averaging is used, the more flat-like the spectrum is.

The probability of finding all the samples of the main lobe as a function of the bandwidth of the signal using the CME and FCME algorithms is presented in Figure 11. In the case of the CME algorithm, the theoretical limit of detection that will express the upper bound of the signals bandwidth is calculated based on (8). In the case of the FCME algorithm,
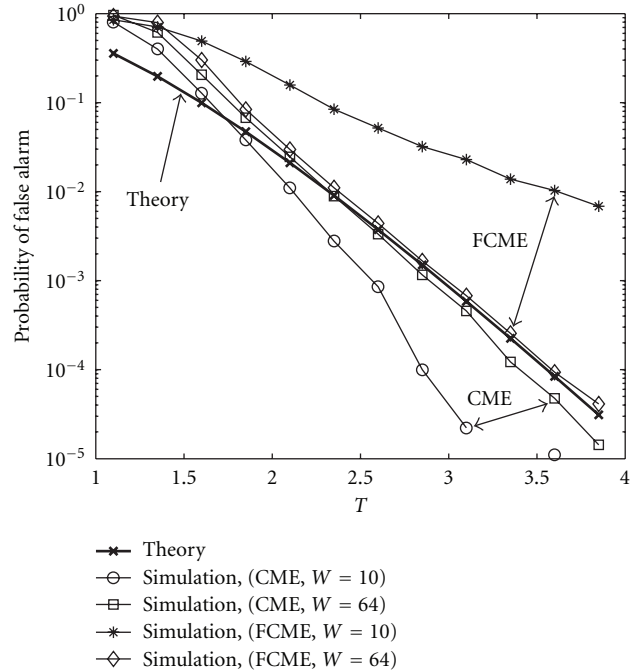
the size of the clean initial set defines the theoretical limit of detection. Here, $T = 2.3026$, $L = 100$ and the initial set of the FCME algorithm includes 64 samples, that is, 12.5% of the total number of the samples. Three different SNRs are used: $\gamma = -6$, 2, and 10 dB. It can be seen that the larger the SNR is, the better the simulation results match the theory (CME algorithm). In the case of the FCME algorithm, the cleanliness of the initial set is the limiting factor, that is, the theoretical limit is 87.5%. The theory and simulation matches almost perfectly.

It is interesting to observe that the SNR per OFDM symbol used in the simulations were lower than it is required in the real-life applications. In practical systems, SNR per subcarrier should be at least 0 dB for reliable communication with coding. This means that the corresponding required SNR per OFDM symbol in the whole bandwidth is about 18 dB with given parameters, whereas in simulations it was ≤10 dB. Therefore, it is expected that analysis is well valid in practice. Furthermore, it is noted that according to extensive simulations, the number of the active subcarriers, the length of the prefix as well as the length of the FFT in the detector have only a small impact (max about 10%) to the results. Therefore, the results are also valid for other OFDM-based systems.

5. Conclusions and Remarks

This paper addressed the analysis of the backward and forward CME algorithms. Computationally attractive CME algorithms are iteratively operating concentrated signal detection methods that, for example, are robust for
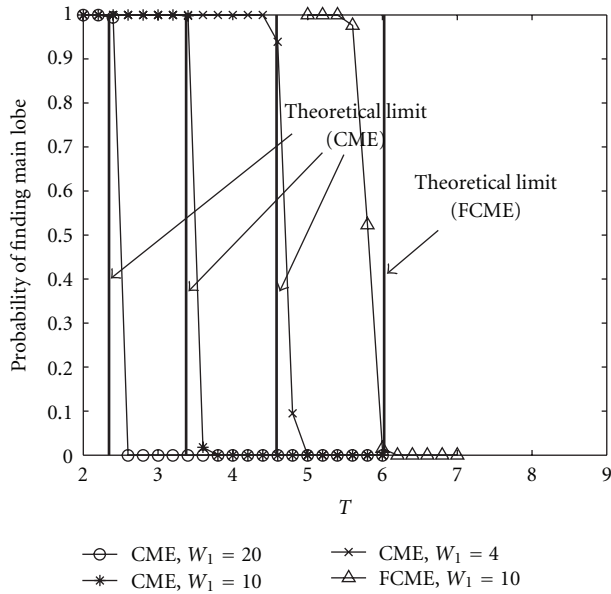
FIGURE 9: Threshold multiplier T versus the probability of finding the main lobe using the CME and FCME algorithms. Random signal and channelized radiometer. $\Theta = 1000$, $\gamma = 34$ dB, $W = 64$, $W_1 = 20, 10$, or 4 samples.
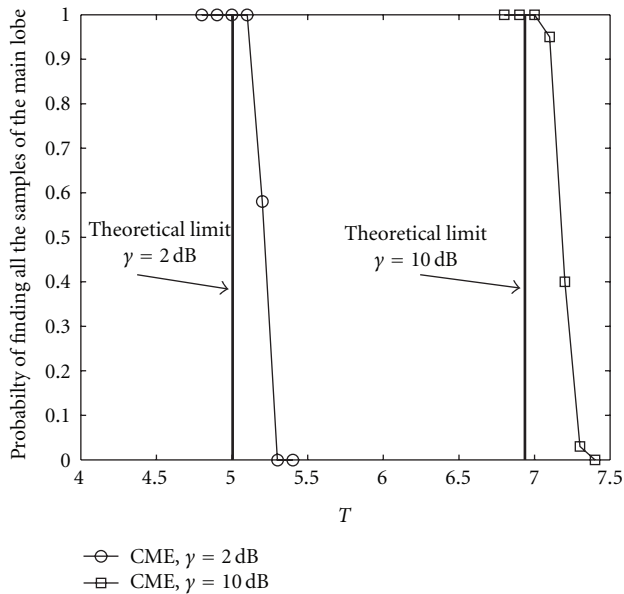

FIGURE 10: Probability of finding all the samples of the main lobe as a function of the threshold multiplier T using the CME algorithm. OFDM signal, $P = 64$, $L = 1000$, $\gamma = 2$ or 10 dB. Signal bandwidth is 13% of the systems' bandwidth.

frequency-shifting. The operation of the algorithms under different conditions was of interest. Simple limits for fast checking if a signal is detectable or not were presented and SNR values that limit the signal detection were derived. The validity of the analysis was confirmed with extensive computer simulations using both random and OFDM signals. It was noticed that the wider the bandwidth of the
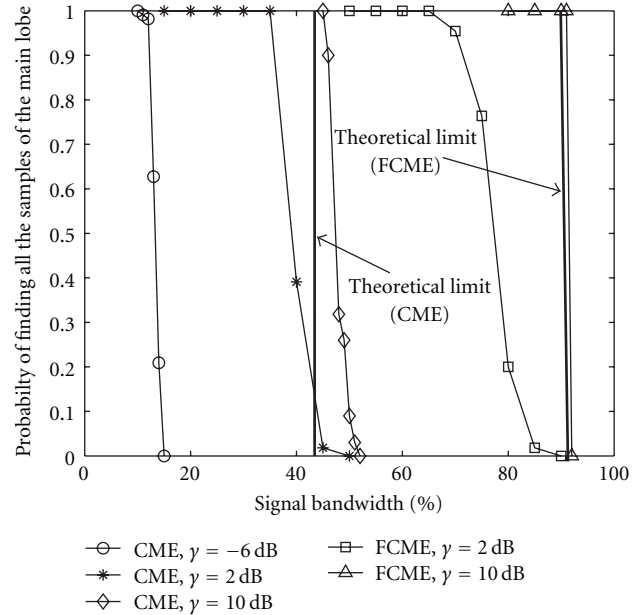

FIGURE 11: Probability of finding all the samples of the main lobe as a function of the width of the signal using the CME and FCME algorithms. OFDM signal, $P = 64$, $L = 100$ and $T = 2.3026$.

system and the larger the SNR, the better the simulations coincide with the theory. The SNR values used in the simulations are comparable with practical OFDM systems. It can be concluded that the presented detection limits can be used, for example, in real-life cognitive radio systems. In practical detection applications, false detections as well as signal separation may cause problems. However, these can be avoided using the extension of the CME algorithms, namely, the localization algorithm based on double-thresholding (LAD), which makes clustering after the CME/FCME detection. Thus, the analysis presented here is valid also for the detection part of the LAD method.

Appendix

The CME Algorithm (Derivation of (7)*).* In the case of the CME algorithm, the mean of the initial set is calculated based on all the samples, so in (3), $Q = N$. Geometrically, the mean of the initial set can be defined to be

$$\overline{\zeta} = \frac{SW + \sum_{i=1}^{n} S_i W_i}{W}. \tag{A.1}$$

Let $S_p \in \{S_1, \ldots, S_n\}$ denote one arbitrary lobe. When keeping in mind that $S_p = \alpha_p S$ and $W_p = \beta_p W$, we get from (2) that

$$\zeta_p > T\left(S + S\sum_{i=1}^{n} \alpha_i \beta_i\right). \tag{A.2}$$

Noticing that $\zeta_p = |r_p|^2 = S + S_p = S + \alpha_p S$, (A.2) reduces to the form

$$S + \alpha_p S > T\left(S + S\sum_{i=1}^{n} \alpha_i \beta_i\right) \tag{A.3}$$

which reduces into

$$1 + \alpha_p > T\left(1 + \sum_{i=1}^{n} \alpha_i \beta_i\right) \tag{A.4}$$

or

$$\alpha_p > T - 1 + T\sum_{i=1}^{n} \alpha_i \beta_i. \tag{A.5}$$

Furthermore,

$$\alpha_p - T\alpha_p\beta_p > T - 1 + T \sum_{i=1, i\neq p}^{n} \alpha_i \beta_i,$$

$$\alpha_p > \frac{T - 1 + T\sum_{i=1, i\neq p}^{n} \alpha_i \beta_i}{1 - T\beta_p}. \tag{A.6}$$

When taking into account that $j = 1,\dots,n$ is the step size and $k = 0,\dots,n$ is the number of lobes that have already been rejected in the previous iterations, $i = 1 \rightarrow i = k + 1$, $p \rightarrow k + j$, and the widths of the already rejected k lobes have to be taken into account. That is, W in (A.1) is replaced by $W - \sum_{i=1}^{k} W_i$ in the numerator and in the denominator. So, the conditions for the heights of the detected lobes can be expressed using function

$$f(k, j) = \frac{T - 1 + T\sum_{i=k+1, i\neq k+j}^{n}\left(\alpha_i\beta_i/\left(1 - \sum_{i=1}^{k}\beta_i\right)\right)}{1 - T\left(\beta_{k+j}/\left(1 - \sum_{i=1}^{k}\beta_i\right)\right)}. \tag{A.7}$$

Note that when $k = 0$, $1 - \sum_{i=1}^{k}\beta_i = 1$.

The FCME Algorithm (Derivation of (11)*).* The FCME algorithm operates forward, so at the first iteration, the clean (initial) set contains only part of the samples. As a forward-type method, the size of the clean set increases in every iteration. That is, when lobe S_p is added to the clean set, it means that also the smaller lobes $S_{p+1},\dots,S_n$ are added to the clean set. Let us assume that the initial set is clean of outliers. As a point of geometrical view, the mean of the clean set Q when it contains k lobes includes the outlier-free samples SW_{clean} which, for one, consists of initial set samples and the rest of the outlier-free samples; samples from k lobes, that is, $\sum_{i=1}^{k} S_{n-i+1} W_{n-i+1}$; and noise samples below these k lobes, that is, $\sum_{i=1}^{k} SW_{n-i+1}$. Thus, the width of the current clean set includes the samples without any lobes, that is, outlier-free samples, and the k weakest lobes $(n, n-1, \dots, n-k+1)$ already (falsely) added to the clean set. Given that, $\overline{\zeta}$ can be defined using function

$$\overline{\zeta} = \frac{SW_{\text{clean}} + \sum_{i=1}^{k} S_{n-i+1} W_{n-i+1} + \sum_{i=1}^{k} SW_{n-i+1}}{W_{\text{clean}} + \sum_{i=1}^{k} W_{n-i+1}}. \tag{A.8}$$

We get from (2) that lobe S_{n-k-j} is detected, that is, lobes $S_{n-k-j+1},\dots,S_{n-k}$ are added to the clean set, if

$$S_{n-k-j} + S > T$$

$$\times \left(\frac{SW_{\text{clean}} + \sum_{i=1}^{k} S_{n-i+1} W_{n-i+1} + \sum_{i=1}^{k} SW_{n-i+1}}{W_{\text{clean}} + \sum_{i=1}^{k} W_{n-i+1}}\right). \tag{A.9}$$

When keeping in mind that $S_p = \alpha_p S$ and $W_p = \beta_p W$, we get that

$$S\alpha_{n-k-j} + S > T$$

$$\times \left(\frac{SW\beta_{\text{clean}} + SW\sum_{i=1}^{k}\alpha_{n-i+1}\beta_{n-i+1} + SW\sum_{i=1}^{k}\beta_{n-i+1}}{W\beta_{\text{clean}} + W\sum_{i=1}^{k}\beta_{n-i+1}}\right) \tag{A.10}$$

or equivalently

$$\alpha_{n-k-j} + 1 > T\left(\frac{\beta_{\text{clean}} + \sum_{i=1}^{k}\alpha_{n-i+1}\beta_{n-i+1} + \sum_{i=1}^{k}\beta_{n-i+1}}{\beta_{\text{clean}} + \sum_{i=1}^{k}\beta_{n-i+1}}\right) \tag{A.11}$$

or

$$\alpha_{n-k-j} > T\left(\frac{\beta_{\text{clean}} + \sum_{i=1}^{k}\alpha_{n-i+1}\beta_{n-i+1} + \sum_{i=1}^{k}\beta_{n-i+1}}{\beta_{\text{clean}} + \sum_{i=1}^{k}\beta_{n-i+1}}\right) - 1 \tag{A.12}$$

which equals

$$\alpha_{n-k-j} > Tf_F(k) - 1. \tag{A.13}$$

Note that the FCME algorithm is able to detect all the lobes or only part of the lobes. Detecting lobes one-by-one is not possible.

List of Symbols

$\alpha_i = S_i/S$: Relative height of the ith lobe,
$\beta_{\text{clean}} = W_{\text{clean}}/W$: Relative width of the clean set,
$\beta_i = W_i/W$: Relative width of the ith lobe,
$\Gamma = 2E_i/N_0$: Noncentrality parameter,
γ: Signal-to-noise ratio (SNR),
$\zeta_i = |r_i|^2$: Squared sample,
$\overline{\zeta} = (1/Q)\sum_{i=1}^{Q}\zeta_i$: The mean,

η_i:	Noise process,
$\Theta = 2T_R W_R$:	Degrees of freedom,
$\xi_i = \zeta_i/\bar{\zeta}$:	Normalized sample,
$\mathbf{C}$:	Shape parameter (usually the covariance matrix of $\mathbf{r}$) (MSD),
c_i:	Concentrated signal sample,
E_i:	Energy of the signal in the ith radiometer,
$f_F(k)$:	Function for FCME,
$f(k, j)$:	Function for CME,
$g_F(k)$:	Another function for FCME,
$g(k, j)$:	Another function for CME,
h:	Initial set includes part of the samples from the hth lobe,
$j = 1, \ldots, n$:	Step size, $k + j \leqslant n$,
$k = 0, \ldots, n$:	Number of rejected (CME) or added (FCME) from/to the set Q in the previous iterations,
L:	Overlapping blocks in the Welch spectrum estimate,
n:	Number of the lobes,
N:	Total number of samples,
N_c:	Number of samples in the initial set,
N_0:	Noise power density,
P:	Length of the prefix,
Q:	Clean set (in the first iteration called as an initial set),
$\mathbf{r}$:	Sample vector (MSD),
$\bar{\mathbf{r}}$:	Location parameter (usually the mean of $\mathbf{r}$) (MSD),
r_i:	ith scalar sample,
s_i:	Noise-like signal sample,
S:	Height of the base lobe,
S_i:	Height of the ith lobe,
T:	Threshold multiplier,
T_h:	Threshold,
T_R:	Integration time of the channelized radiometer,
V:	Number of channels in the channelized radiometer,
W:	Width of the base signal (= total number of samples = BW),
$W_{\text{clean}} = W - \sum_{i=1}^{n} W_i$:	Number of clean samples,
W_i:	Width of the is lobe,
$W_{\text{init}} = \kappa W, 0 < \kappa < 1$:	Size of the initial set,
W_{OFDM}:	Bandwidth of the OFDM signal,
W_R:	Bandwidth of the one channel in the channelized radiometer,
x:	Initial set includes x samples of the main lobe.

Acknowledgments

References

[1] E. Pearson and C. C. Sekar, "The efficiency of statistical tools and a criterion for the rejection of outlying observations," *Biometrika*, vol. 28, pp. 308–320, 1936.

[2] D. M. Hawkins, *Identification of Outliers*, Chapman & Hall, Boca Raton, Fla, USA, 1980.

[3] A. S. Hadi, "Identifying multiple outliers in multivariate data," *Journal of the Royal Statistical Society*, vol. 54, no. 3, pp. 761–771, 1992.

[4] A. C. Atkinson, "Fast very robust methods for the detection of multiple outliers," *Journal of the American Statistical Association*, vol. 89, no. 428, pp. 1329–1339, 1994.

[5] A. C. Atkinson and M. Riani, "Bivariate boxplots, multiple outliers, multivariate transformations and discriminant analysis: the 1997 hunter lecture," *Environmetrics*, vol. 8, no. 6, pp. 583–602, 1997.

[6] J. Mitola and G. Q. Maguire Jr., "Cognitive radio: making software radios more personal," *IEEE Personal Communications*, vol. 6, no. 4, pp. 13–18, 1999.

[7] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 2, pp. 201–220, 2005.

[8] J. W. Wisnowski, D. C. Montgomery, and J. R. Simpson, "A comparative analysis of multiple outlier detection procedures in the linear regression model," *Computational Statistics and Data Analysis*, vol. 36, no. 3, pp. 351–382, 2001.

[9] J. Hardin and D. M. Rocke, "The distribution of robust distances," *Journal of Computational and Graphical Statistics*, vol. 14, no. 4, pp. 928–946, 2005.

[10] J. G. Proakis, *Digital Communications*, McGraw-Hill, New York, NY, USA, 3rd edition, 1995.

[11] H. Saarnisaari, P. Henttu, and M. Juntti, "Iterative multidimensional impulse detectors for communications based on the classical diagnostic methods," *IEEE Transactions on Communications*, vol. 53, no. 3, pp. 395–398, 2005.

[12] P. Henttu and S. Aromaa, "Consecutive mean excision algorithm," in *Proceedings of the IEEE International Symposium on Spread Spectrum Techniques and Applications (ISSSTA '02)*, pp. 450–454, Praha, Czech Republic, September 2002.

[13] J. Vartiainen, J. J. Lehtomäki, H. Saarnisaari, and P. Henttu, "Estimation of signal detection threshold by CME algorithms," in *Proceedings of the 59th IEEE Vehicular Technology Conference (VTC '04)*, pp. 1654–1658, Milan, Italy, May 2004.

[14] H. Saarnisaari and P. Henttu, "Impulse detection and rejection methods for radio systems," in *Proceedings of the IEEE Military Communications Conference (MILCOM '03)*, vol. 2, pp. 1126–1131, Boston, Mass, USA, October 2003.

[15] J. Vartiainen, S. Aromaa, H. Saarnisaari, and M. Juntti, "Performance evaluation of transform selective interference suppression," in *Proceedings of the IEEE Military Communications Conference (MILCOM '04)*, pp. 1422–1428, Monterey, Calif, USA, October-November 2004.

[16] J. Vartiainen, J. J. Lehtomäki, and H. Saarnisaari, "Double-threshold based narrowband signal extraction," in *Proceedings of the 61st IEEE Vehicular Technology Conference (VTC '05)*, vol. 2, pp. 1288–1292, Stockholm, Sweden, May-June 2005.

[17] J. Vartiainen, H. Sarvanko, J. Lehtomäki, M. Juntti, and M. Latva-Aho, "Spectrum sensing with LAD-based methods," in *Proceedings of the IEEE International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC '07)*, Athens, Greece, September 2007.

[18] J. Vartiainen, M. Alatossava, J. J. Lehtomäki, and H. Saarnisaari, "Interference suppression for measured radio channel

data at 2.45 GHz," in *Proceedings of the IEEE International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC '06)*, Helsinki, Finland, September 2006.

[19] H. Saarnisaari, "Consecutive mean excision algorithms in narrowband or short time interference mitigation," in *Proceedings of the Position Location and Navigation Symposium (PLANS '04)*, pp. 447–454, Monterey, Calif, USA, April 2004.

[20] J. Lehtomäki, S. Salmenkaita, J. Vartiainen, J.-P. Mäkelä, R. Vuohtoniemi, and M. Juntti, "Mea-surement studies of a spectrum sensing algorithm based on double thresholding," in *Proceedings of the Wireless Vitae*, Aalborg, Denmark, May 2009.

[21] J. Vartiainen, J. Lehtomäki, H. Saarnisaari, and M. Juntti, "Limits of detection for the consecutive mean excision algorithms," in *Proceedings of the International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM '10)*, Cannes, France, June 2010.

[22] J. J. Lehtomäki, J. Vartiainen, M. Juntti, and H. Saarnisaari, "Analysis of the LAD methods," *IEEE Signal Processing Letters*, vol. 15, pp. 237–240, 2008.

[23] T. Hänninen, J. Vartiainen, M. Juntti, and M. Raustia, "Implementation of spectrum sensing on wireless open-access research platform," in *Proceedings of the International Workshop on Cognitive Radio and Advanced Spectrum Management (CogART '10)*, Rome, Italy, November 2010.

[24] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C*, Cambridge University Press, Cambridge, UK, 2nd edition, 1992.

[25] J. Lehtomäki, *Analysis of energy based signal detection*, Ph.D. thesis, Acta Universitatis Ouluensis Technica C 229. Faculty of Technology, University of Oulu, Oulu, Finland, December 2005, http://herkules.oulu.fi/isbn9514279255/.

[26] H. Sarvanko, M. Mustonen, A. Hekkala, A. Mämmelä, M. Matinmikko, and M. Katz, "Cooperative and noncooperative spectrum sensing techniques using Welch's periodogram in cognitive radios," in *Proceedings of the International Workshop on Cognitive Radio and Advanced Spectrum Management (CogART '08)*, Aalborg, Denmark, February 2008.