Limits of Detection for the Consecutive Mean Excision Algorithms

Johanna Vartiainen, Janne Lehtomäki, Harri Saarnisaari and Markku Juntti

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

Abstract—Iterative backward and forward consecutive mean excision (CME/FCME) algorithms are efficient in suppressing and detecting narrowband signals. Being computationally relatively simple, they can be applied for spectrum sensing in cognitive radios. In this paper, theoretical limits of detection for the CME algorithms are derived. These limits predict whether a signal is detectable or not. More specifically, resulting signal-to-noise ratio (SNR) values at which the signals can be detected are presented. In addition, rules for choosing the optimal CME algorithm are conveyed. The validity of the presented limits of detection is confirmed with computer simulations.

I. INTRODUCTION

Outliers can be defined to be observations that differ from the distribution of the main (base) data [1], [2]. The outlier detection has numerous applications in various fields of science. Under the unifying umbrella of wireless communications, narrowband (NB) signal suppression and detection are probably among the most widely studied topics. More specifically, future cognitive radios [3–7] are a current example where signal detection methods can be applied. Therein, signal detection can be used to sense frequency bands in order to find un-occupied ones for transmission [8], [9]. Thus, signal samples can be considered as outliers (corresponding to occupied bands) and the noise as well as possible noise-like wideband information signal correspond to the base data (unoccupied bands).

The iteratively operating backward and forward consecutive mean excision (CME/FCME) algorithms [10], [11] are computationally simple but effective energy based diagnostic [12] outlier detection methods, so they can be used both in NB interference suppression and NB signal detection. In contrast to energy detector, the CME methods are blind in the sense that *a priori* knowledge about the noise level is not required. In addition, knowledge about the number of the NB signals or their characteristics is not required. Instead, the only requirement is that the outliers are from a signal that is 'narrow' (concentrated) in the considered domain. Thus, the CME algorithms can be used in time or frequency or in some other domain. One application of the CME algorithms called the localization algorithm based on double-thresholding (LAD) method can be used as a spectrum sensing technique

for spectrum measurements. Unlike the energy detector, the LAD method is able to estimate the number of NB signals and their parameters (bandwidths (BW), signal-to-noise ratio (SNR) values etc.) in any frequency range, thus being useful in future cognitive radios [13]. The CME algorithms and their applications have been widely studied and compared to other methods, see, for example, [14], [15] and references therein. A simplified analysis of the CME algorithms has been presented in [14] without detection limits. In [16], theoretical limits of detection were derived. Therein, the resulting SNR limits were not analyzed and the rules for choosing the optimal CME algorithm were not considered. It is known that the FCME algorithm outperforms the CME algorithm but it is also more complex [11]. However, in the case of very NB signals, the performance of the simpler CME algorithm is adequate.

In this paper, the theoretical detection capability of the CME algorithms is studied. The resulting theoretical limits predict whether a signal is detectable or not. These limits lead to theoretical SNR values at which the signals with different widths can be detected. This knowledge can be used, for example, in cognitive radios for fast checking if a signal is detectable or not without time consuming computer simulations. According to the authors' best knowledge, these kinds of detection limits have not been presented earlier for the CME algorithms. Here, we are focusing in analyzing the resulting SNR limits and presenting rules for choosing the optimal CME algorithm for the detection. The validity of the presented approximation is confirmed with extensive computer simulations using random signals and channelized radiometer.

II. CONSECUTIVE MEAN EXCISION ALGORITHMS

The ith observed sample r_i corresponding to the base data without outliers is assumed to be a zero mean, independent, i.i.d. Gaussian distributed complex random variable. Thus, $|r_i|^2 = \zeta_i$ follows a chi-squared distribution. Sample ζ_i is decided to be an outlier if [11]

$$\zeta_i > T\bar{\zeta}, \tag{1}$$

where $T > 1$ is a threshold parameter, $\bar{\varsigma} = 1/Q(\sum_{i=1}^{Q}\varsigma_i)$ is an estimator corresponding to the shape parameter (i.e., the average signal sample energy), and Q is the size of the so called clean set (in the first iteration called as an initial set) [11]. The value of the threshold parameter T is calculated based on the desired false alarm probability $P_{FA,DES}$ in the

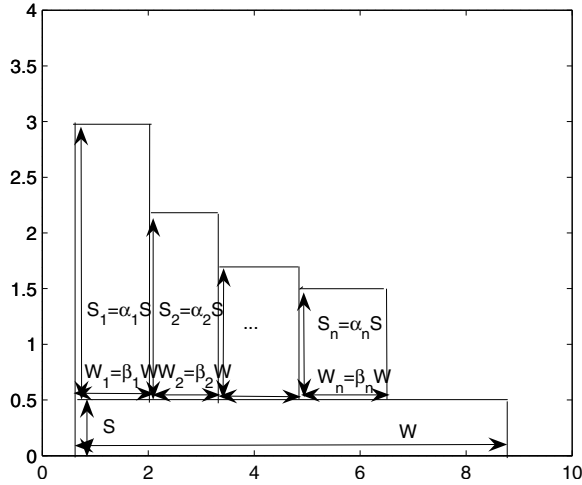|	CME	FCME
Sorting required	No	Yes
Initial set	All samples	ca. 10% of samples
Direction	Backward	Forward
Threshold T_h	$T\bar{\zeta}$	$T\bar{\zeta}$
Comparison	Reject $\zeta_i > T_h$	Add $\zeta_i < T_h$
Outliers	above T_h	above T_h

Fig. 1. An example of base signal (SW) and NB signal lobes (S_iW_i).

outlier-free case [11]. It means that the CME algorithms use constant false alarm rate (CFAR) principle used in the energy detection of unknown signals [17]. That is, if the required false alarm probability is high, it leads to a small value of threshold parameter, and, thus, to a low threshold value. Respectively, small false alarm probability leads to a high threshold value. The lower the threshold is, the better the signal samples can be detected but the more there will be false alarms, i.e., noise samples exceeding the threshold. This causes falsely detected signals. Instead, the higher the threshold is, the less there are false alarms. However, weak signal samples cannot be detected. Thus, a trade-off needs to be made.

Because the threshold parameter is a constant and depends only on the desired false alarm probability, it is calculated only once. As the CME algorithm operates backward rejecting NB signal samples samples from the clean set if their energy $\zeta_i > T\bar{\zeta}$, the FCME algorithm operates forward by adding samples to the clean set if $\zeta_i < T\bar{\zeta}$. At the end of this iterative processing, there are two sets of samples: noise and noise-like samples below the threshold value and outliers above the threshold value. The CME algorithms are summarized in Table I. For more details, see, [11].

III. THEORETICAL LIMITS OF DETECTION

The received signal consists of outliers (NB signal samples) and base signal (noise samples). In the analysis purposes,

the outliers can be assumed to be sorted in a descending order according to their heights (amplitudes). In the frequency domain, the height of the sample corresponds to the sample energy. Thus, the reordered samples can be divided into different parts (lobes) according to their heights, so every lobe consists of outliers with equal amplitude. The width of the lobe states how many samples with equal amplitude the lobe includes. If the width of the lobe is one, the lobe includes only one outlier sample. An example of lobes is shown in Fig. 1.

Assume that the base signal has width W and height S. Thus, the power of the base signal is SW. Assume also that the outlier signal consists of $n \geq 1$ lobes with widths $W_i = \beta_i W$, $\sum_{i=1}^{n} W_i \leq W$, and heights $S_i = \alpha_i S$, $i = 1, \cdots, n$, $S_1 > \cdots > S_n$. Therein, β_i and α_i denote the relative height and width of the ith lobe, respectively. Thus, we get that the SNR per total BW is

$$\gamma = \frac{\sum_{i=1}^{n}(S+S_i)W_i}{SW} = \sum_{i=1}^{n}(1+\alpha_i)\beta_i. \qquad (2)$$

The first lobe S_1W_1 is called as the main lobe. Because of the assumed sorting, the height (i.e., the energy) of the main lobe is the highest. Furthermore, because the CME algorithms operate as energy detectors, it can be assumed that the main lobe is the easiest lobe to be detected [17], [18].

Next, the limits of detection as a function of SNR are defined for the CME and FCME algorithms. Let $k = 0, \cdots, n-1$ denote the number of lobes already rejected (for the CME algorithm) or added (for the FCME algorithm) from/to the clean set Q in the previous iterations. Let $j = 1, \cdots, n$ denote the step size, i.e., how many lobes are rejected/added at current iteration, and $k+j \leq n$.

First, the case of the CME algorithm is considered. Using a geometrical approach for defining $\bar{\zeta}$, we get from (1) that the limits of detection as a function of α can be obtained from [16]

$$f(k,j) = \frac{T - 1 + T\sum_{i=k+1, i\neq k+j}^{n} \frac{\alpha_i\beta_i}{1-\sum_{i=1}^{k}\beta_i}}{1 - T\frac{\beta_{k+j}}{1-\sum_{i=1}^{k}\beta_i}}. \qquad (3)$$

It means that the $(k+j)$th lobe $(S_{k+j}W_{k+j})$ is detected if $\alpha_{k+j} > f(k,j)$ and if $\alpha_{k+j+1} < f(k, j+1)$, i.e., the $(k+j+1)$th lobe is not detected. In general, the detection is possible only if the main lobe is above the threshold, i.e., [16]

$$\alpha_1 > f(0,1). \qquad (4)$$

Because the FCME algorithm is able to detect signal lobes regardless of their BWs when the initial set is clean, the analysis of the FCME algorithm is simpler. Here, it is assumed that the initial set that usually covers only about 10% of the total BW is clean of outliers. The case when the initial set is not clean is considered in [16]. Let κW, $0 < \kappa < 1$ denote the size of the initial set and $W_{clean} = W - \sum_{i=1}^{n} W_i$ denote the number of outlier-free samples. It can be assumed that the initial set is clean of outliers, i.e., $\kappa W \leq W_{clean}$. The mean

TABLE II
THE VALUES OF SNR [DB] WHEN THE SIGNAL CAN BE DETECTED VIA THE CME ALGORITHM WITH DIFFERENT VALUES OF T AND β_1. THE SIGNAL HAS ONLY THE MAIN LOBE. SNR IS DEFINED PER TOTAL BANDWIDTH.

T	$\beta_1 = 0.05$	$\beta_1 = 0.1$	$\beta_1 = 0.2$	$\beta_1 = 0.3$	$\beta_1 = 0.4$	$\beta_1 = 0.5$
2.3	> -9 dB	> -6 dB	> -2 dB	> 2 dB	> 8 dB	$-$
4.6	> -5 dB	> -1 dB	> 10 dB	$-$	$-$	$-$
6.9	> -3 dB	> 3 dB	$-$	$-$	$-$	$-$

$-$ denotes that the detection is impossible

TABLE III
THE VALUES OF SNR [DB] WHEN THE SIGNAL CAN BE DETECTED VIA THE FCME ALGORITHM WITH DIFFERENT VALUES OF T AND β_1. THE SIGNAL HAS ONLY ONE LOBE. SNR IS DEFINED PER TOTAL BANDWIDTH.

T	$\beta_1 = 0.05$	$\beta_1 = 0.1$	$\beta_1 = 0.2$	$\beta_1 = 0.3$	$\beta_1 = 0.4$	$\beta_1 = 0.5$	$\beta_1 = 0.6$	$\beta_1 = 0.7$	$\beta_1 = 0.8$	$\beta_1 = 0.9$
2.3	> -9 dB	> -6 dB	> -3 dB	> -2 dB	> 0 dB	> 1 dB	> 1 dB	> 2 dB	> 3 dB	> 3 dB
4.6	> -6 dB	> -3 dB	> 0 dB	> 2 dB	> 3 dB	> 4 dB	> 4 dB	> 5 dB	> 6 dB	> 6 dB
6.9	> -5 dB	> -2 dB	> 1 dB	> 3 dB	> 4 dB	> 5 dB	> 6 dB	> 7 dB	> 7 dB	> 8 dB

of the clean set Q when it contains k lobes can be defined using function [16]

$$f_F(k) = \frac{\beta_{clean} + \sum_{i=1}^{k} \beta_{n-i+1}(1 + \alpha_{n-i+1})}{\beta_{clean} + \sum_{i=1}^{k} \beta_{n-i+1}}, \quad (5)$$

$\beta_{clean} = W_{clean}/W$, which has been derived from (1) using a geometrical approach for defining $\bar{\zeta}$. In the considered iteration, the lobes $n - k - j + 1, \cdots, n - k$ 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 clean set Q includes only outlier-free samples when $\alpha_n > Tf_F(0) - 1 = T - 1$. The signal cannot be detected at all when the main lobe is not detected or when the threshold is larger than the main lobe, i.e., $\alpha_1 < Tf_F(k) - 1$. It means that the signal can be detected only if [16]

$$\alpha_1 > Tf_F(k) - 1. \quad (6)$$

IV. NUMERICAL RESULTS

Here, the resulting limits of detection are considered in more detail. In practice, the amplitudes of the simulated signal samples are not equal. Hence, an average amplitude (mean of the lobe samples) presents the amplitude of the lobe. Assume first that all the signal samples belong to the same lobe. Thus, the signal consists of one lobe and an average amplitude of the signal samples is used to present the height of the lobe. The resulting SNR values at which the signals can be detected can be obtained from (4) and (6) using (2). In Tables II and III, the theoretical SNR values at which the NB signals with different widths can be detected for three commonly used values of threshold parameter T (2.3, 4.6 and 6.9 [15]) are presented. For example, in the case of the CME algorithm, the NB signal with width of 0.1 (10%) can be detected if the SNR is at least -6 dB when using $T = 2.3$, -1 dB when using $T = 4.6$ and 3 dB when using $T = 6.9$. In the case of the FCME algorithm, the corresponding SNR values are -6 dB, -3 dB and -2 dB. Naturally, from the presented equations, the detection limits can be derived for any desired value of T.

It can be seen from Tables II and III that the FCME algorithm is able to operate regardless of the width of the lobe. Instead, the CME algorithm has some restrictions. From (4)
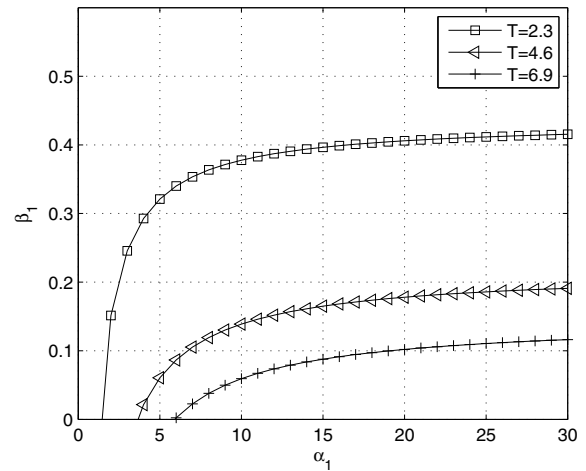

Fig. 2. The relative height of the main lobe vs. the relative bandwidth of the main lobe. The signal has one lobe. The detection via the CME algorithm is possible below the curves.

and denominator of (3), we get that the limiting values as a function of the width of the main lobe are $\beta_1 < 0.4343$ ($T = 2.3$), $\beta_1 < 0.217$ ($T = 4.6$) and $\beta_1 < 0.145$ ($T = 6.9$). In other words, the main lobe can cover 43, 23 and 15% of the spectrum, respectively. This result is confirmed in Fig. 2, where the relative height of the signal vs. the relative width of the signal when the signal can be detected via the CME algorithm is presented. Therein, it is assumed that the signal consists of one lobe. It can be seen that the curves converge to the limiting values. Below the curve is the area where the signal can be detected and above the curve is the area where the signal can not be detected. For example, if $T = 2.3$ and the relative height of the signal is $\alpha_1 = 5$, the signal can be detected if $\beta_1 < 0.32$.

As can be noticed, the narrower the signal is, the lower SNR is required for the detection. This holds for both the CME algorithms. When the signal is very NB, say about 5% of the total BW, the required SNR values for both the

CME algorithms are almost equal. Instead, when the BW of the signal gets wider, the CME algorithm requires larger SNR values for the detection than the FCME algorithm. However, it should keep in mind that the FCME algorithm is computationally more complex than the CME algorithm. Given that, when the signal is narrow and SNR is high enough, the CME algorithm could be a reasonable choice. When the BW of the NB signal is rather wide, say 20% of the total BW or more, the CME algorithm is not able to detect the NB signal when the threshold parameter T is relatively high. In addition, when the BW of the signal is about 50% of the total BW or more, the CME algorithm is not able to detect the signal at all. Instead, the FCME algorithm is still able to detect the signal. This is because the FCME algorithm is able to operate as long as the initial set is free of outliers. However, the larger the BW of the signal is, the larger the required SNR is. For example, when the threshold parameter $T = 2.3$ and the SNR is 2 dB, the CME algorithm is able to detect signals whose BW is at most 30% of the total BW. Instead, at the same parameters, the FCME algorithm is able to detect signals whose BW can be at most 70% of the total BW.

V. Simulation Results

The simulations were performed in the frequency domain using random signals and channelized radiometer in order to confirm the validity of the presented limits of detection. Note, that the purpose is not to compare the CME algorithms. In total, W radiometer channels were assumed. Each has integration time T_R, bandwidth W_R and $\Theta = 2T_RW_R$ denotes the degrees of freedom [16]. One channel in the radiometer corresponds one frequency domain sample. The noise was Gaussian.

First, the signal has only one lobe and the width of the lobe is one sample. The probability of finding the signal for different values of T is presented in Fig. 3. The theoretical limits are based on (4) and (6). The total BW is $W = 64$ samples, the signal has one sample, $\Theta = 10$ or 1000, and SNR is 7 dB or 34 dB. When SNR is 7 dB and $\Theta = 10$, the results for the CME and FCME algorithms 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, i.e., the simple theoretical limits 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. It can also be seen that in all of the cases, the theoretical limits of detection in terms of T corresponds around 50% probability of detection in the simulations. This phenomenon is confirmed in Fig. 4, where the probability of finding the signal at a specific threshold limit as a function of SNR [dB] is presented. Therein, $\Theta = 10$, 100 or 1000 and T is calculated based on (4). It can be clearly seen that the probability of detection converges to 50%.

Next, the width of the lobe α_1 is more than one sample, so an average amplitude (mean of the sample energies) presents the amplitude of the lobe α_1. The probability of finding the
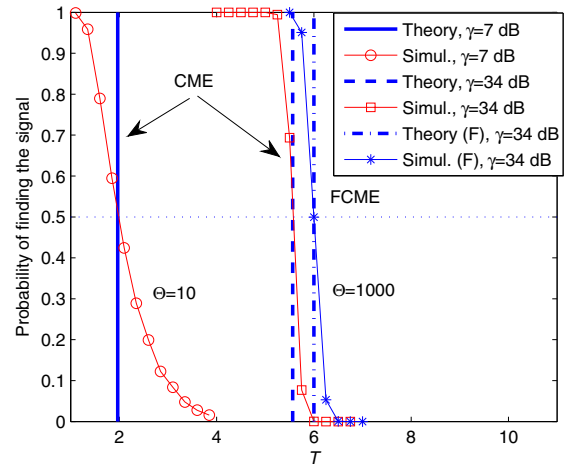

Fig. 3. Threshold parameter T vs. the probability of finding the signal using the CME algorithms. $\gamma = 7$ or 34 dB, $\Theta = 10$ or 1000, and $W = 64$. Signal bandwidth is 1.6% (one sample).
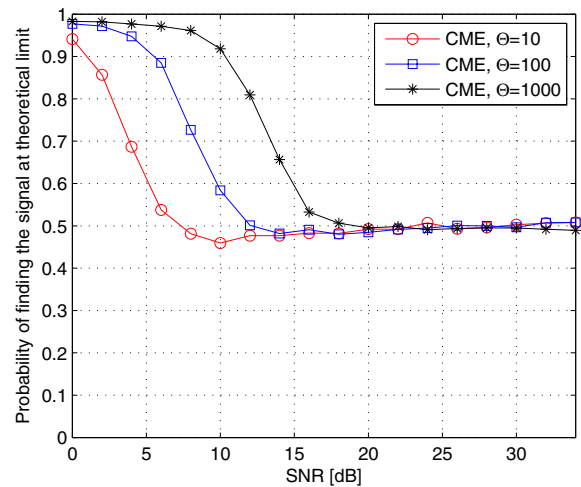

Fig. 4. SNR [dB] vs. the probability of finding main lobe at threshold limit using the CME algorithm. $\Theta = 10$, 100 or 1000 and $W = 64$. Signal bandwidth is 1.6% (one sample).

main lobe vs. threshold parameter T using the CME algorithm is presented in Fig. 5. Now, $\Theta = 1000$ and SNR (per sample) is 34 dB. The total width $W = 64$ samples and the width of the main lobe is 4 (6.25%), 8 (12.5%) or 16 (25%) samples. The theoretical limits were calculated based on (4). As in the previous case, the theory and simulation results match well.

Here, we have assumed that the signal consists of one lobe, so all the signal samples belong to the same lobe. However, it can also be assumed that the signal consists of several lobes. Next, it is assumed that the signal consists of two lobes and $\beta_1 = \beta_2$. Based on (3), we get also specific conditions for the detection. For example, the CME algorithm detects both the lobes at the same time if $\alpha_2 > f(0,2) = \frac{T-1+T\alpha_1\beta_1}{1-T\beta_2}$, $\alpha_1 >$
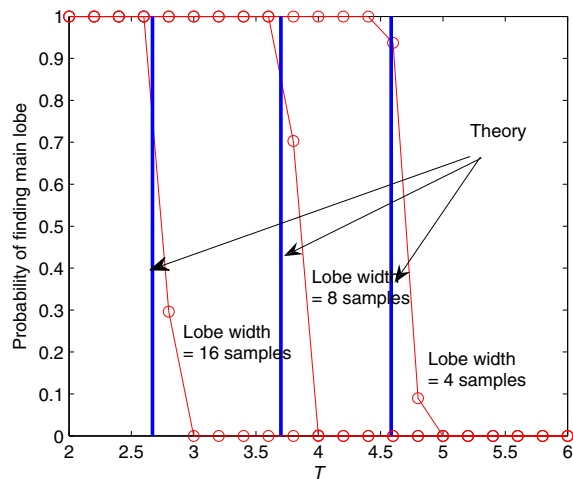
Fig. 5. SNR [dB] vs. the probability of finding main lobe at threshold limit using the CME algorithm. $\Theta = 1000$, $W = 64$ samples and SNR= 34 dB.
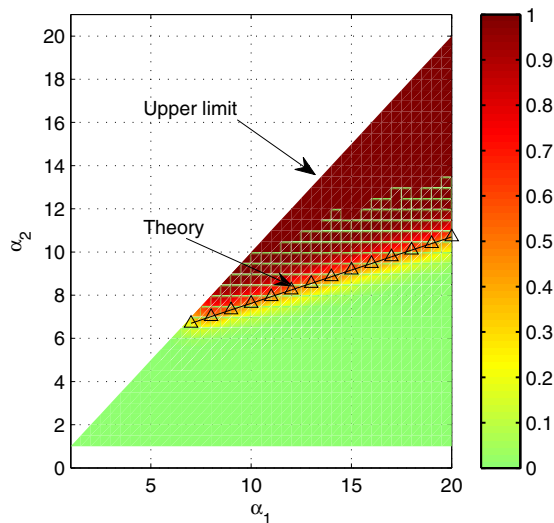

Fig. 6. The relative height of the main lobe vs. the relative height of the second lobe for joint detection of two lobes when $T = 4.6057$. $\beta_1 = \beta_2 = 0.05$, $\Theta = 100$ and $W = 20$ samples. The limits of detection via the CME algorithm.

α_2 and $\beta_1 < 1/T$ holds. This result is confirmed in Fig. 6, where the relative height of the main lobe α_1 vs. the relative height of the second lobe α_2 when detection is possible via the CME algorithm is presented. Therein, $\beta_1 = \beta_2 = 0.05$ (5%), $T = 4.6057$, $\Theta = 100$ and $W = 20$ samples. The detection is possible above the theoretical curve but below the upper limit $\alpha_1 > \alpha_2$. The simulated probability of detection is between 0 (green) and 1 (red). It can be seen that the simulation results coincide the theory very well, i.e., the probability of detection is one approximately between the theoretical curve and the upper limit.

VI. CONCLUSIONS

The theoretical limits of detection for the backward and forward CME algorithms were presented. The validity of the presented detection limits were confirmed via computer simulations. It was noticed that the wider the bandwidth of the system and the larger the SNR, the better the simulations coincide the theory. It can be concluded that the presented limits of detection can be used, e.g., in cognitive radio systems for fast checking if a signal is detectable or not.

REFERENCES

[1] A. S. Hadi, "Identifying multiple outliers in multivariate data," *J. Royal Stat. Soc.*, vol. 54, no. 3, pp. 761–771, 1992.
[2] A. C. Atkinson and M. Riani, "Bivariate boxplots, multiple outliers, multivariate trans- formations and discriminant analysis: the 1977 hunter lecture," *Environmetrics*, vol. 8, pp. 583–602, 1997.
[3] J. Mitola III, "Cognitive radio for flexible mobile multimedia communications," in *IEEE International Workshop on Mobile Multimedia Communications (MoMuC 99)*, San Diego, CA, USA, 1999, pp. 3–10.
[4] J. Mitola III and G. Q. Maguire Jr., "Cognitive radio: making software radios more personal," *IEEE Pers. Commun.*, vol. 6, no. 4, pp. 13–18, Aug. 1999.
[5] V.D. Chakravarthy, A.K. Shaw, M.A. Temple, and J.P Stephens, "Cognitive radio - an adaptive waveform with spectral sharing capability," in *Proc. IEEE Wireless Commun. and Networking Conf.*, New Orleans, LA, USA, Mar.13–17 2005, pp. 724–729.
[6] S. Haykin, "Cognitive radio: Brain-empowered wireless communications," *IEEE Journal in Selected Areas in Comm.*, vol. 23, no. 2, pp. 201–220, Feb. 2005.
[7] S. N. Shankar, C. Cordeiro, and K. Challapali, "Spectrum agile radios: Utilization and sensing architectures," in *IEEE International Symposium on Dynamic Spectrum Access Networks*, Baltimore, USA, Nov. 2005, vol. 1, pp. 160–169.
[8] FCC, "Notice for proposed rulemaking: Facilitating opportunities for flexible, efficient, and reliable spectrum use employing cognitive radio technologies," ET Docket No. 03-108, Dec. 2003.
[9] FCC, "Notice of proposed rule making: Unlicensed operation in the TV broadcast bands," ET Docket No. 04-186, May 2004.
[10] P. Henttu and S. Aromaa, "Consecutive mean excision algorithm," in *Proc. IEEE 7th International Symposium on Spread Spectrum Techniques and Application*, Praha, Czech Republic, 2002, pp. 450–454.
[11] H. Saarnisaari, P. Henttu, and M. Juntti, "Iterative multidimensional impulse detectors for communications based on the classical diagnostic methods," *IEEE Trans. Commun.*, vol. 53, no. 3, pp. 395–398, 2005.
[12] 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 & Data Analysis*, vol. 36, no. 3, pp. 351–382, June 2001.
[13] J. Vartiainen, H. Sarvanko, J. Lehtomäki, M. Juntti, and M. Latva-Aho, "Energy detection based spectrum sensing methods in cognitive radios," in *WWRF 2007*, Espoo, Finland, June 13–15 2007.
[14] H. Saarnisaari, "Consecutive mean excision algorithms in narrowband or short time interference mitigation," in *Position Location and Navigation Symposium PLANS 2004*, Monterey, CA, USA, April 2004, pp. 447–454.
[15] J. Vartiainen, M. Alatossava, J. J. Lehtomäki, and H. Saarnisaari, "Interference suppression for measured radio channel data at 2.45 GHz," in *Proc. IEEE Int. Symp. Pers., Indoor, Mobile Radio Commun.*, Helsinki, Finland, Sept. 11–14 2006, vol. 1.
[16] J. Vartiainen, J. J. Lehtomäki, H. Saarnisaari, and M. Juntti, "Analysis of the consecutive mean excision algorithms," *EURASIP J. Signal Processing*, 2010, submitted.
[17] S. M. Kay, *Fundamentals of statistical signal processing: Detection theory.*, Prentice Hall, Upple Saddle River, NJ, USA, 1998.
[18] H. Urkowitz, "Energy detection of unknown deterministic signals," vol. 55, no. 4, pp. 523–531, Apr. 1967.