DESCRIPTOR BASED ON SPECTRAL PEAKS CORRELOGRAMS

Brenda Lilia Valdez Reyna, Miroslav Zivanovic.
Ingeniería en Sistemas Computacionales, Ingeniería Elétrica y Eletrónico
Universidad Autónoma de Tamaulipas Campus Mante, Universidad Pública de Navarra
Blvd.E.Cárdenas Glz. 1201, 89840 Tamaulipas, México.
blvaldez@uat.edu.mx, miro@unavarra.

Abstract. We have presented in this report a new peak descriptor SD based on the lag plot and discussed its relationship to the NBD. The SD, defined as the slope of a linear regression model constructed through the data samples belonging to a time sequence, shows to be a good candidate for describing sinusoidal and noise peak classes.

A proper choice of data set is crucial for discerning correctly between the peak classes. The SD makes use only of the spectral peak shape and consequently this information can be explained by just two parameters, namely the root mean square BWrms and absolute bandwidth L. This is similar to the way the shape factor explains the filter spectral shape in the circuit theory. We have shown that BWrms and L, initially defined in the spectrum domain, hold relationship to the time duration of the data set and its sampling rate respectively.

Keywords: DESCRIPTOR, LAG PLOT, SPECTRAL

1 Introduction

This work presents the development of the application of descriptors, which allow them to know the characteristics of the signal, will be explored here the following analysis of descriptors:

1. LAG PLOT – BASICS, 2. LAG PLOT AND SPECTRAL PEAKS: THE SLOPE DESCRIPTOR (SD), 3. CHOOSING THE TIME-DOMAIN REPRESENTATION OF A SPECTRAL PEAK, 4. THE SLOPE DESCRIPTOR AND THE PARAMETERS BWrms AND L.

1. LAG PLOT – BASICS

The lag plot is a method in statistics that aims to determine whether a discrete data set is random or deterministic. For a data set x(n) the lag plot is conceived as a 2D representation of x(n) versus x(n-P), where P is a lag expressed in samples. Although P can be an arbitrary value, the lag plots are most commonly generated with P = 1. namely, the Gaussian noise and a pure tone of 100 samples each. In the case of the noise, the knowledge about the sample x(n-1) provides almost no knowledge about the following sample x(n). As for the sinusoid, the elliptical structure aligned with the diagonal (unitary slope) allows to precisely predict the next point in the data set and thus characterize the data set as deterministic.

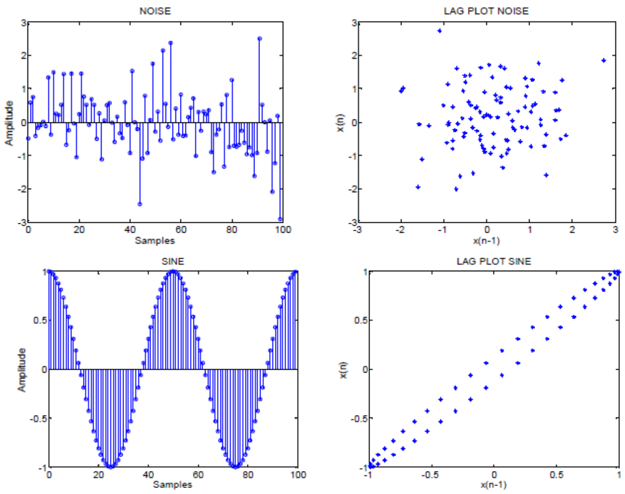

Figure 1. Lag plot representation for Gaussian and sinusoidal noise-free pattern.

2. LAG PLOT AND SPECTRAL PEAKS: THE SLOPE DESCRIPTOR (SD)

By applying a lag plot to the time domain representation of a spectral peak, we may eventually distinguish sinusoids from noise. The lag plot for the sinusoidal signal shown on Figure 1 suggests that a linear regression (least-squares) model might be appropriate for characterizing the data set. The linear regression models are simple to implement and for our case, we would need only one coefficient (parameter) that would represent the slope of the straight line fitted to the data set.

For a given data set x(n) we define the Slope Descriptor (SD) in the following way:

$$x(n) = ax(n-1) + b + e(n)\ ,\ a = SD\ . \quad (1)$$

The parameters a and b are the coefficients of the model while e(n) is the modeling error. Wemay expect a priori that the sinusoidal patterns will have the SD close to one while the noise patterns will present a wide range of values (including SD = 1). This means that this simple linear model cannot achieve a complete sine/noise separation, so we have to adapt it somehow in order to reduce/eliminate the overlap between the peak classes. Hence, the key issue is to correctly choose the time-domain representation of the spectral peaks i.e. to adapt the data set in such a way to be correctly handled by the SD.

3. CHOOSING THE TIME-DOMAIN REPRESENTATION OF A SPECTRAL PEAK

Approach 1: Complex narrow-band spectral peak and its time signal

In order to preserve the peak's central frequency we first set to zero all the bins in the spectrum except those belonging to the peak. Then we calculate the IDFT and use its real part as a data set for the lag plot. The problem with this approach is that the time domain pattern of any spectral peak would exhibit a strong deterministic behaviour. Let us recall our previous research where

we discovered that any spectral peak or, in general, a part of the spectrum with the bandwidth much smaller than the Nyquist frequency, can be considered a narrow-band process represented by a strong carrier with slowly varying amplitude and phase:

$$N(n) = A(n)\cos[\Omega_o n + \theta(n)]. \quad (2)$$

Therefore, even though the modulation parameters A(n) and $\Omega_o(n)$ are random variables, the presence of the carrier assures a deterministic behaviour of N(n). Hence, the lag plot for both sinusoidal and noise peaks will produce a tight clustering of the points along the diagonal and consequently there will be no way of distinguishing between the peak classes. As an example, we show on Figure 2 various narrow-band noise spectral peaks centred around 1.5 kHz and their corresponding real IDFTs. We observe that even though the time-domain patterns are quite different, the lag plots order the points as if they belonged to a purely deterministic signal.

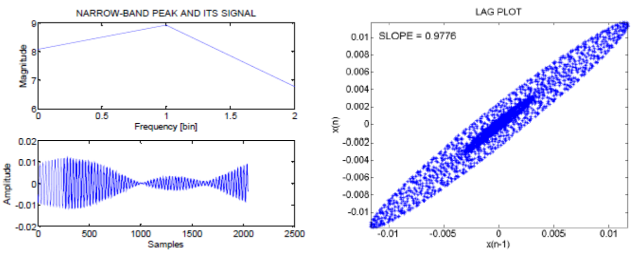

Figure 2. Various complex narrow-band noise peaks centered around 1.5 kHz, their real IDFTs and corresponding lag plots. The window is Hanning, the size of the widow is 1000 and the size of the FFT/IFFT is 2048. The IDFTs are computed by setting to zero all the bins in the FFT not belonging to the peak

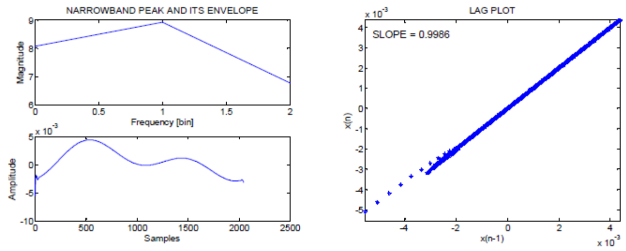

Figure 3. Various complex narrow-band noise peaks down-shifted to the origin, their real IDFTs and corresponding lag plots. The window is Hanning, the size of the widow is 1000 and the size of the FFT/IFFT is 2048. The IDFTs are computed by setting to zero all the bins in the FFT not belonging to the peak.

Approach 2: Complex narrow-band spectral peak and its time envelope

Another approach consists in first down shifting the peak to the origin (base-band) by typically multiplying the original signal with a corresponding complex exponential term. Then, like in Approach 1, the rest of the bins in the spectrum that are not included in the peak are set to zero

and the IDFT is calculated. The resulting pattern (the real part of the IDFT) is the amplitude envelope while the carrier is suppressed. For pure tones the real part of the IDFT will be exactly the analysis window. For modulated peaks it will be a product between the analysis window and the modulator of the sinusoidal signal. For out worst case sinusoidal signal the resulting pattern will still resemble the shape of the analysis window (e.g. see the windowed longest and shortest pattern for the NDD thresholds determination in our previous work). The noise peaks will produce a wide variety of patterns, which may seem beneficial for distinguishing noise from sinusoids. Unfortunately, this approach is neither good because the envelope from a noise peak is a low-pass function, i.e. the transition between the contiguous samples is quite smooth, which will again produce very “deterministic” lag plots. The same noise peaks from Figure 2 are represented on Figure 3 together with respective amplitude envelopes and lag plots. We can see at glance that the lag plots show very strong deterministic behaviour.

Approach 3: Complex broad-band spectral peak and its time signal

Therefore, we have to somehow "randomize" the time-domain representation of the noise spectral peaks. A way to achieve this is to describe spectral peaks by general broad-band, rather than narrow-band processes, i.e. the peak bandwidth is close to the Nyquist frequency. With this approach, the time-domain representations of spectral peaks would include many frequencies.

This is particularly beneficial for the noise peaks because this would allow to correctly describing randomness. For the sinusoidal peaks this approach is also good, because the deterministic signals are always correctly represented, either the model is narrow-band or broadband process.

A practical implementation of the above approach can be carried out by simply considering a spectral peak as a whole DFT, i.e. neglecting the presence of the rest of the bins in the spectrum. Then, we calculate the IDFT only over the bins in the peak and take its real part as

a data set for the lag plot.

If we take a look at Figure 4 we observe that the transitions between the contiguous samples in the real IDFTs of the noise peaks are abrupt, that is, the patterns are more "random". As a consequence, the corresponding lag plots show no apparent point clustering and the SD

will thus have a wide range of values. Unfortunately, the same effect is detected in the sinusoidal peaks. As an illustration, compare Figure 4 with Figure 5, where various spectral peaks belonging to our worst-case sinusoidal signal for the carrier frequency equal to 1.5 kHz are dealt with.

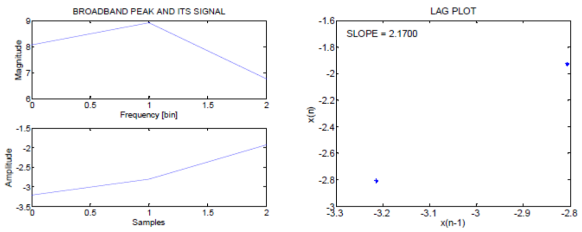

Figure 4. Various complex broad-band noise peaks, their real IDFTs and corresponding lag plots. The window is Hanning, the size of the widow is 1000 and the size of the FFT is 2048. The size of the IDFT is equal to the number of bins in the peak.

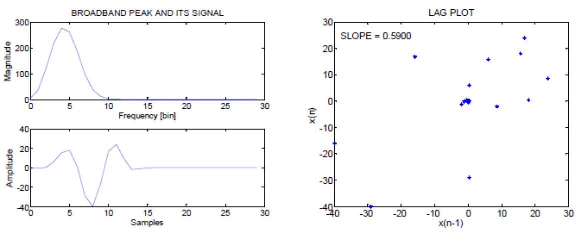

Figure 5. Various complex broad-band sine peaks from the worst case signal, their real IDFTs and corresponding lag plots. The window is Hanning, the size of the widow is 1000 and the size of the FFT is 2048. The size of the IDFT is equal to the number of bins in the peak.

Approach 4: Complex broad-band spectral peak and its time envelope

Hence, we now have to assure smoother transitions between the contiguous samples in thesinusoidal time-domain patterns, while preserving the randomness in the noise patterns. One way to do it is to deal with the envelope of the analytical signal, rather then its real part.

Namely, an isolated spectral peak can be considered unilateral spectrum of an analytical sequence (we recall that any real sequence will give a conjugated even absolute spectrum, while analytical sequences produce unilateral absolute spectra). Then, the IDFT of a complex broadband spectral peak is an analytical sequence which may be written as:

$$y(n) = x(n) + jx\hat{}(n)\ , \quad (3)$$

where x^(n) is the Hilbert transform of real sequence x(n). The expression (3) can also be written in the polar form:

$$y(n) = A(n)e^{j\phi(n)}\ , \qquad A(n) = \sqrt{x^2(n) + \hat{x}^2(n)}\ , \qquad \phi(n) = a\tan[\hat{x}(n)/x(n)]\ . \qquad (4)$$

The term A(n) is the low-pass component or envelope of the analytical signal i.e. it varies more slowly than the original signal x(n). In fact, this approach is similar to Approach 2 in the sense that the carrier is suppressed and the resulting pattern is a low-pass sequence. The difference is that herein the computational resolution of the IDFT is much smaller (the size of the IDFT), because we consider only the bins in the spectral peak. Therefore, we can say that the envelope has been sampled at a very low rate. How is this fact related to the lag plot?

A lag plot describes deterministic signals properly only if the sample rate is high enough to assure smooth transitions between adjacent samples. If this condition is satisfied, all the points in the lag plot will be clustered along the diagonal. If the sampling rate is reduced, the lag plot will exhibit a number of outliers and thus the slope of the linear model is reduced. According to the lag plot the signal is less deterministic, but this is not true! It is simply that we have fewer points per sinusoidal cycle and the transitions between the samples in the data set are more abrupt.

An extreme example would be a cosine sampled at exactly Nyquist frequency, that is, 2 samples per cycle. In this case all the points in the lag plot are concentrated in only 2 outliers: (-1, 1) and (1, -1). Figure 6 illustrates this trend for a sinusoidal signal sampled at different rates.

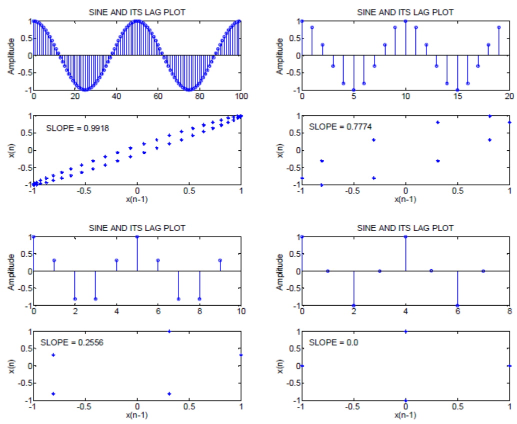

Figure 6. A sinusoid sampled at different rates: a) 50 samples/cycle, b) 10 samples/cycle, c) 5 samples/cycle, d) 4 samples/cycle, and respective lag plots.

Returning now to the issue of the time envelope, we may say that a low sampling rate will definitely randomize it up to some point. This sampling rate reduction affects principally the noise peaks, because they are in general narrower than the sinusoidal peaks and consequently the size of the corresponding IDFT is smaller. Therefore, the IDFT from a noise peak will put

fewer points on the corresponding time envelope. The sinusoidal peaks are in general broader, so the IDFT from a sinusoidal peak will put more samples on the corresponding envelope, thus rendering it smoother and more deterministic in the sense of the lag plot. Figures 7 and 8 illustrate the benefits of this approach for the same noise and sinusoidal peaks as those in Figure 4 and 5.

In spite of the advantages of this approach, there is a shortcoming that must be dealt with. While the values of the SD for sinusoidal peaks are close to one, the values for the noise peaks may be greater or lesser than one. As a consequence, we may thus expect some overlap between the sinusoidal and noise classes in the SD domain. Therefore, we still need to improve the data set in order to best separate the peak classes.

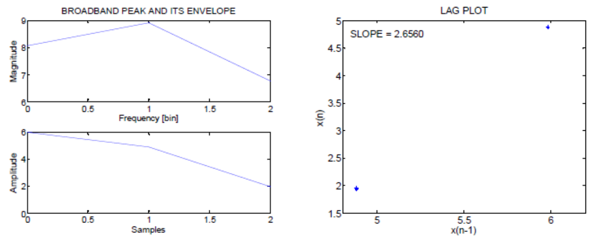

Figure 7. Various complex broad-band noise peaks, their absolute IDFTs and corresponding lag plots. The window is Hanning, the size of the widow is 1000 and the size of the FFT is 2048. The size of the IDFT is equal to the number of bins in the peak.

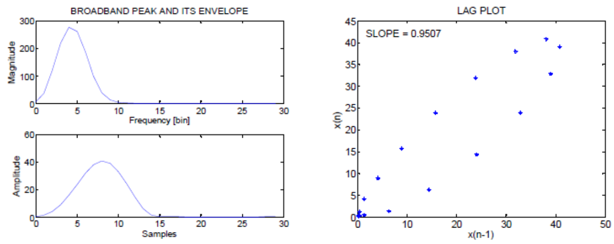

Figure 8. Various complex broad-band sine peaks from the worst case signal, their absolute IDFTs and corresponding lag plots. The window is Hanning, the size of the widow is 1000 and the size of the FFT is 2048. The size of the IDFT is equal to the number of bins in the peak.

Approach 5: Real broad-band spectral peak and its time envelope

Regarding Approach 4, we need to tighten the range of values of the SD for the noise peaks in order to reduce as much as possible the overlap between the peak classes. In fact, if we can reduce the variety of A(n) shapes for the noise peaks, it would automatically reduce the range of possible corresponding SD values. That is to say, we would let A(n) vary from peak to peak but always within some general pattern.

One way to do it is to process only the modulus of the spectral peak disregarding its spectral phase. As the spectral modulus is a real sequence, its IDFT will yield a complex sequence whose A(n) is conjugate even. The sequence A(n) is now periodic with period equal to the size of the IDFT. This periodicity will emphasize the deterministic aspect of the timedomain representation for the noise peaks and consequently the range of possible values for the SD reduces. If we compare the SD values from Figures 7 and 9 for the same noise peaks, we see how efficiently this approach delimits the SD values to a small range close to zero. Initially, we might have thought that this approach would put the noise peaks closer to the sinusoidal peaks in the SD domain. Recall, however, Approach 4 where we saw that due to larger IDFT the sinusoidal peaks yield smoother A(n) and thus more deterministic lag plots. Hence, the sinusoidal peaks will still have larger SD values (as an illustration, compare Figures 8 and 10) although the range of values is somewhat enlarged.

In view of the aforementioned discussion, we choose this approach as a reference for generating the data sets for the spectral peaks' lag plots. In the following sections we will prove analytically the hypothesis posed herein and show the benefits of this approach for the peak classification.

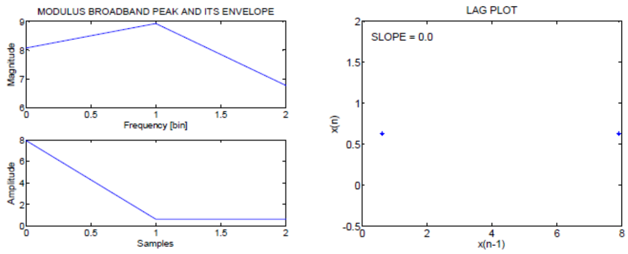

Figure 9. Various real broad-band noise peaks, their absolute IDFTs and corresponding lag plots. The window is Hanning, the size of the widow is 1000 and the size of the FFT is 2048. The size of the IDFT is equal to the number of bins in the peak.

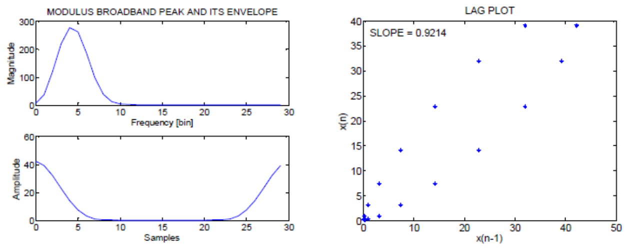

Figure 10. Various real broad-band sine peaks from the worst case signal, their absolute IDFTs and corresponding lag plots. The window is Hanning, the size of the widow is 1000 and the size of the FFT is 2048. The size of the IDFT is equal to the number of bins in the peak.

4. THE SLOPE DESCRIPTOR AND THE PARAMETERS BWrms AND L

Let us first recall the nomenclature we used in the previous research, because we are going to apply it herein too. The parameter BWrms was referred to as the root mean square bandwidth of a spectral peak expressed in bins. We called the parameter L the absolute bandwidth of a peak, being equal to the number of bins in the peak. We can now proceed to the discussion about the relationship between the SD and these spectral peak parameters.

The SD is the slope of a linear regression model which we have already defined in Section 2. For the sake of clarity we repeat herein the equation (1):

$$x(n) = ax(n-1) + b + e(n)\ , \ a = SD\ . \quad (5)$$

The expression (5) represents a first order linear predictor i.e. a first order AR model. Recalling the origin of x(n), we may say that a real broad-band spectral peak is the spectral estimate of the following LTI constant coefficient model H(z):

$$H(z) = \frac{1}{1 - az^{-1}} \quad , \quad z = e^{j\Omega} \ . \tag{6}$$

In the frequency domain the expression (6) becomes:

$$H(e^{j\Omega}) = \frac{1}{1 - ae^{-j\Omega}} = \frac{1}{1 - a\cos\Omega + ja\sin\Omega} \ . \tag{7}$$

The last expression says that a real broad-band spectral peak can be represented as a LTI lowpass system with one real pole which can be inside (a < 1), outside (a > 1) or onto the unitary circle (a = 1). The parameter a determines the sharpness of the peak and thus the bandwidth of the system. Therefore, we may expect the smallest bandwidth when the pole is on the unitary circle.

In order to estimate the bandwidth BWrms of the system, we can use the well-know 3dB criterion:

$$H(\Omega_c) = \frac{1}{\sqrt{2}} H(0) = \frac{1}{|1-a|\sqrt{2}} \times \frac{1}{\sqrt{1 - 2a\cos\Omega_c + a^2}} \ , \tag{8}$$

where Ω_c is the cut-off frequency. From (8) we can express the cut-off frequency as:

$$\Omega_c = \cos^{-1}\left(-\frac{a^2 - 4a + 1}{2a} \right) . \tag{9}$$

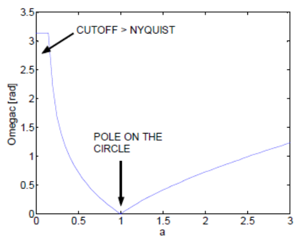

Figure 11. The cut-off frequency Ω_c versus the slope a. For a < 0.15 approximately, the bandwidth is greater than the Nyquist interval: we have aliasing.

If we put $BW_{rms} = 2\Omega_c$ and a = SD in (9) then we get straightforwardly an approximate relationship between the bandwidth and slope descriptor of a spectral peak. The expression (9) is plotted on Figure 11 showing the dependence of a on Wc. We observe that, according to our expectations, for $a = 1$ the bandwidth is theoretically zero while otherwise the bandwidth increases as the pole steps off the unitary circle. For small a the bandwidth spreads out of the Nyquist interval, giving rise to the aliasing effect. So, how is all this related to the peak classification?

Let us recall that in order to yield a deterministic characterization from the lag plot, the transitions between the contiguous samples in a time sequence must be smooth enough. We have already seen in the previous section that this condition can be met by having the sampling rate (the size of the IDFT) large enough. According to Approach 5, it means that larger peaks (L large) will have large SD independently of the shape of its time envelope.

However, a large SD can also be achieved with smaller L but larger time envelope, if we recall from the Fourier theory that a small BWrms correspond to a large time response and vice versa. This is conceptually similar to the process of decimating. For example, if a time sequence is decimated with a factor 2, then its bandwidth enlarges two times. But at the same time this enlargement is compensated by cutting down the number of bins in the DFT by two (assuming that we don't use zero padding). As a result, the number of DFT bins within the bandwidth remains the same.

We may now take a look at Figure 12 where various real broadband spectral peaks from the worst case signal and the corresponding time envelopes are shown. The envelopes are represented versus the normalized length, so that the effect of L is suppressed and only the effect of BWrms remains. We observe that both small BWrms and L produce almost the same shape of the envelope as larger BWrms and L, thus yielding similar SD. This issue is very important for the description of the sinusoids, because this approach always tends to compress the sinusoidal class by reducing the SD of the larger peaks (more modulation) and enlarging the SD of the narrower peaks (less modulation).

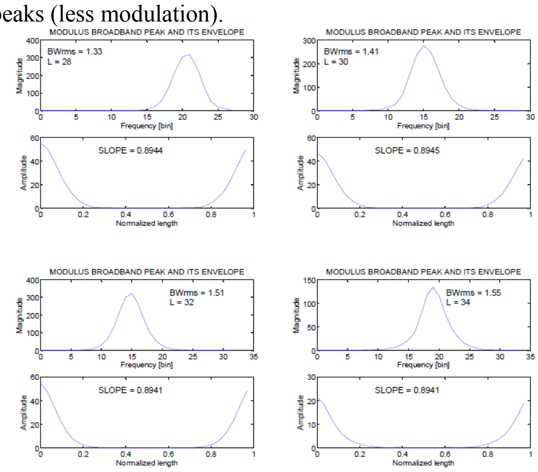

Figure 12. Various real broadband peaks from the worst-case signal and the corresponding envelopes. Normalized length is given as a sample index divided by the length of the envelope sequence.

However, this trend is also present in the noise peaks, because for any peak in general a large L corresponds to a large BWrms and the other way around. Fortunately, the relationship between those parameters is quite different for sinusoids and noise. Figure 13 shows the joint BWrms-L histograms for both peak classes (the axis are the same in both plots in order to facilitate the visual comparison).

We see that the noise histogram exhibit very strong correlation between the parameters, because the dependence region follows a steep linear trend and thus the parameters change proportionally. This is in fact very beneficial for the peak classification, because it will not allow the noise peaks go beyond a certain value of the SD. Regarding the sinusoidal peaks, the correlation is much weaker and doesn't seem to obey a linear trend. As a consequence, the beneficial trend (increments in L) is much more emphasized than the non-beneficial trend (increments in BWrms). Having in mind that the sinusoidal peaks are in general larger than the noise peaks, this approach will tend to push the sinusoidal peaks towards larger values of SD, that is, away from the noise peaks. In the next section we will show the histograms for the peak classes in the SD domain and establish a comparison between the SD and NBD.

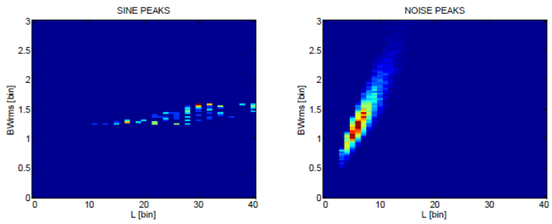

Figure 13. Joint BW_{rms}-L histograms for sinusoidal and noise peaks. For the sinusoidal peaks we used our noise-free worst case signal and for the noise peaks the Gaussian noise was used. The analysis window was the Hanning.

6. CONCLUSIONS

We have presented in this report a new peak descriptor SD based on the lag plot and discussed its relationship to the NBD. The SD, defined as the slope of a linear regression model constructed through the data samples belonging to a time sequence, shows to be a good candidate for describing sinusoidal and noise peak classes.

A proper choice of data set is crucial for discerning correctly between the peak classes. The SD makes use only of the spectral peak shape and consequently this information can be explained by just two parameters, namely the root mean square BWrms and absolute bandwidth L. This is similar to the way the shape factor explains the filter spectral shape in the circuit theory. We have shown that BWrms and L, initially defined in the spectrum domain, hold relationship to the time duration of the data set and its sampling rate respectively.

We have next shown that the mutual action of those parameters determines the smoothness of the data set and consequently its lag plot interpretation. In order to yield a deterministic interpretation, BWrms should be as small as possible and L as large as possible. This however does not correspond to a realistic scenario because those parameters are mutually coupled. However, this relationship is less strict for the sinusoidal peaks and thanks to this fact the sinusoidal class can be well distinguished from the noise.

Finally, we have shown that, regarding BWrms and L, the NBD behaves in a similar way the SD does but in the opposite direction. By inverting the roles of the parameters in the NBD we obtain a new description of the peak classes that resembles the SD. This shows that two apparently very different spectral peak descriptors, depending on the same parameters in a similar way, provide for a similar description of the peak classes. Their dependence on BWrms and L is of course not the same but it follows the same trends.

Referencias

[1] F. Plante, G. Meyer and W.A.Ainsworth, “Improvement of Speech Spectrogram Accuracy by the Method of Reassignment”, IEEE Transactions on Speech and Audio Processing. Vol.6 No.3, Mayo 1998.

[2] L. Cohen, “Time-frequency analysis”, Prentice Hall Signal Processing Series, 1995.

[3] S. Orphanidis, “Introduction to signal processing”. Prentice Hall International Editions,1996.

[4] C.Hory, N. Martin and A. Chehikian, “Spectrogram Segmentation by Means of Statistical Features for Non-Stationary Signal Interpretation”, IEEE Transactions on Signal Processing, Vol. 50, No.12, Diciembre 2002.

[5] C.Hory, N. Martin, “Mixture densities formulation of a spectrogram segmentation task”, Proceedings of EUSIPCO 2002, pp. 427-430, 3-6 September 2002.

[6] Miroslav Zivanovic, “Detection of Non-Stationary Sinosoids by Using Joint Frequency Reassignment and Null to Null Bandwith”, Proceedings of EUSIPCO 2008, Agosto 2008.