A MODIFIED TIME-FREQUENCY METHOD FOR TESTING WIDE-SENSE STATIONARITY

Douglas Baptista de Souza†, *Jocelyn Chanussot*†, *Anne-Catherine Favre*‡ *and Pierre Borgnat*§

†GIPSA-Lab, ‡LTHE, Grenoble INP, Domaine Universitaire, Saint Martin d'Hères - FRANCE
§CNRS, Laboratoire de Physique, École Normale Supérieure de Lyon, Lyon - FRANCE

ABSTRACT

Recently, a time-frequency approach for testing stationarity was proposed. However, this method inefficiently detects nonstationarities of the first-order. Here, we present two contributions that improve the test performance and allow the detection of first-order evolutions. The first one is to use an adequate distance measure. The second is a modification of the method in order to consider the spectral content from the signal itself when computing the distances.

Index Terms— First-order stationarity, Time-frequency analysis, Distances, Marginals.

1. INTRODUCTION

In statistical signal processing, many approaches rely on the assumption of stationarity. In the broad sense, it is defined as the invariance of the statistical properties relative to an absolute time [1]. In practice, this definition is relaxed and it is common to consider just the wide-sense stationarity (WSS), which is the invariance of second-order statistics.

General techniques for analyzing nonstationary signals do not exist. Here, we focus on approaches of frequency nature, by specifically analyzing the Power Spectrum Density (PSD). For WSS signals, the PSD is obtained by taking the Fourier transform of the autocorrelation sequence, since it depends on a single index. Nonstationary signals result in a time-varying PSD, which reflects the time evolution of the second properties of the process.

There are several ways to represent the time-varying spectrum of a nonstationary process. As a time-frequency (TF) representation, the estimates of the *Wigner-Ville Spectrum* (WVS) occupy a highlighted position. Other TF representations do not have as many good features as the WVS does. For only one realization one may find several estimators for the WVS. Due to its good properties, many authors have been proposing to use the WVS to represent the evolutionary PSD [1].

Among the recent articles, a very interesting stationarity test based on WVS estimates has been published [2]. The method tests stationarity given only one realization and relative to some observation scale. It is performed by gathering a collection of surrogates as a stationarized version of the signal. Then, by using an adequate dissimilarity measure, the distances between local and global spectra are computed for both the stationary version and the signal itself. The dispersion of these distances is used as a statistic for the stationarity test.

After applying the stationarity test to different signals, we observed an incapacity of detecting first-order evolutions. The purpose of this paper is to propose two modifications to the approach of [2], in order to increase the classification accuracies and allow the detection of an evolutionary mean.

The structure of the paper is as follows. Section 2 presents the stationarity test. Section 3 discusses the distance measures. Here, our first contribution is presented, which is to use an alternative distance. In Section 4 the test is applied to some synthetic signals. Our second contribution is shown in Section 5, which is taking the spectral content into account in order to detect first-order evolutions. Results obtained after this modification are also shown in this section. Finally, our conclusions are drawn in Section 6.

2. TESTING STATIONARITY WITH SURROGATES

According to [2], stationarity should not be seen as an absolute property, but related to a given duration or observation scale. Over this interval, a stationary signal should not exhibit evolution of its time-varying spectrum. It means that the local spectra $S(t_n, f)$ at all different time instants are statistically equal to the global average spectrum

$$\langle S(t_n, f)\rangle_n := \frac{1}{N}\sum_{n=1}^{N} S(t_n, f), \quad (1)$$

where N is the number of time positions $\{t_n, n = 1, ..., N\}$ in which the local spectra are computed. The spectra are obtained by estimates of the Wigner-Ville Spectrum (WVS), which reduces to the PSD when the signal turns out to be stationary. The WVS is estimated using multitaper spectrograms

$$S_K(t, f) = \frac{1}{K}\sum_{k=0}^{K-1} S^{h_k}(t, f). \quad (2)$$

In (2), for a given time point in $\{t_n, n = 1, ..., N\}$, the multitaper spectrogram $S_K(t, f)$ is obtained by averaging K dif-

978-1-4673-0046-9/12/$26.00 ©2012 IEEE

ferent spectrograms. This technique aims at reducing the variance of the estimation by projecting the observation on a family of orthonormal basis functions $\{h_k(t), k \in \mathbb{N}\}$. Here, the *Hermite functions* were chosen as such basis. The time positions t_n, the spacing $\Delta_t = t_{n+1} - t_n$, and the frequency range of the multitaper spectrograms are computed considering the length of the signal and an adjustable fraction of the chosen window.

The originality of the method is to give significance to the fluctuations of the local spectra by building stationarized references of the signal (surrogates) using just the available data. Each surrogate $s(t)$ has the same global PSD as the original signal while being stationary. By comparing the stationary version and the original signal, a hypothesis test is performed to check if fluctuations of local time-varying spectrum around the global average spectrum are statistically expected in the stationarized references. If they are not, we may infer that the variation between local and global spectra is due to the nonstationarity of the process.

Nonstationary signals have their spectral content spread in time differently from stationary ones. Hence, for an identical marginal spectrum over the same observation scale, we may expect some time-organized structures that are exclusive of nonstationary signals. The surrogates are obtained by destroying the organized phase structure controlling the supposed nonstationarity. The signal is Fourier transformed and its phase is replaced by a uniform distributed sequence, while keeping the magnitude unchanged. By applying the inverse Fourier transform we obtain as many stationary surrogates as phase randomizations are performed. After obtaining the J surrogates $s_j(t), j = 1, ..., J$, a distance between local and global spectra is computed for each of them.

$$\{c_n^{(s_j)} := D(S_{s_j,K}(t_n, f), \langle S_{s_j,K}(t_n, f)\rangle_n), n = 1, ..., N, \quad j = 1, ..., J\}. \tag{3}$$

In (3), the distance $D(.,.)$ denotes some dissimilarity measure in frequency. The null hypothesis of stationarity is characterized by the distribution of the variances of (3). For the signal itself, the distances between spectra are given in the vector

$$\{c_n^{(x)} := D(S_K(t_n, f), \langle S_K(t_n, f)\rangle_n), n = 1, ..., N\}, \tag{4}$$

corresponding to the same time instants where the multitaper spectrograms are being computed. We take the variance of the distance vector in (4) as a measure of the spectral fluctuations related to the signal itself:

$$\Theta_1 = \text{var}(c_n^{(x)})_{n=1,...,N},$$

where Θ_1 is the test statistic for the one-sided test:

$$d(x) = \begin{cases} 1 & \text{if } \Theta_1 > \gamma, \text{"nonstationary"}, \\ 0 & \text{if } \Theta_1 < \gamma, \text{"stationary"}. \end{cases} \tag{5}$$

The threshold γ above which the null hypothesis of stationarity is rejected is obtained from (3), the collection of distances regarding the surrogate set. To compute γ, first we take the variance of each distance vector $\{c_n^{(s_j)}, j = 1, ..., J\}$, which leads to a vector of J variances shown in (6).

$$\{\Theta_0(j) = \text{var}(c_n^{(s_j)})_{n=1,...,N}, j = 1, ..., J\}. \tag{6}$$

After computing (6), it is necessary to obtain its statistical distribution. In [2] it is proposed to model the distribution of $\Theta_0(j)$ by a gamma pdf, having its two parameters estimated in a maximum likelihood sense. It is reasonable, since *"the test statistic sums up squared, possibly dependent quantities which themselves result from a strong mixing likely to act as a Gaussianizer"*. Assuming the gamma model to hold, we derived a threshold $\gamma = 0.95$ above which the null hypothesis of stationarity is rejected.

3. DISTANCES

The dissimilarity measure between local and global spectra is a key point of the method. Distances can be separated in two broader classes: distances of probability or frequency nature.

Probability-based distances quantify the dissimilarity between two statistical objects, e.g., probability distributions. Alternatively, frequency-based distances are computed directly from the spectra, by comparing them in both shape and level [1].

Originally, it was proposed in [2] to mix both natures, forming a combined distance in order to take advantage from the different classes. Such dissimilarity measure is shown in (7), where $D_{LS}(G, H)$ and $D_{KL}(\widetilde{G}, \widetilde{H})$ denote the well-known log-spectral deviation and Kullback-Leibler divergence, respectively. The terms H and G are two positive spectra while $\widetilde{H}$ and $\widetilde{G}$ denote their normalized (to unity) versions.

$$D_{CB}(G, H) = D_{KL}(\widetilde{G}, \widetilde{H})(1 + D_{LS}(G, H)) \tag{7}$$

Unfortunately, the distance of (7) did not present a good performance with respect to first-order nonstationarities, even after the weighting procedure that will be introduced in Section 5. Hence, we propose to use a different dissimilarity measure called *Itakura-Saito distortion*, given by (8). It is a frequency-based distance which was chosen after an investigative study. Theoretically, we should expect better results using this measure, since it computes the matching error between an original spectrum $G(f)$ (the local spectrogram, for example), and its approximation $H(f)$ (the global spectrum, as being considered here) [3].

$$D_{IS}(G, H) = \int_{\Omega} \left[\frac{G(f)}{H(f)} - \log\frac{G(f)}{H(f)} - 1\right] df \tag{8}$$

4. APPLYING THE STATIONARITY TEST

After introducing the stationarity test, we apply it to some synthetic signals. These are Gaussian sequences with a varying mean, variance, and both, occurring in different positions over time. The testing signals, which have their mean component subtracted, are shown in Fig. 1, where the three adopted nonstationarity models are presented. In Fig. 1(a),(d),(g) and Fig. 1(b),(e),(h) we have, respectively, nonstationarities starting at the beginning and at the middle. In Fig. 1(c),(f),(i) we have nonstationarities starting and ending at certain positions, forming a sort of a *piecewise of stationarity*. Finally, we chose a short length ($N = 139$), in order to be in accordance with several areas of application like biomedicine, climatology and hydrology. Also, testing short signals allows to evaluate the test under unfavorable situations.

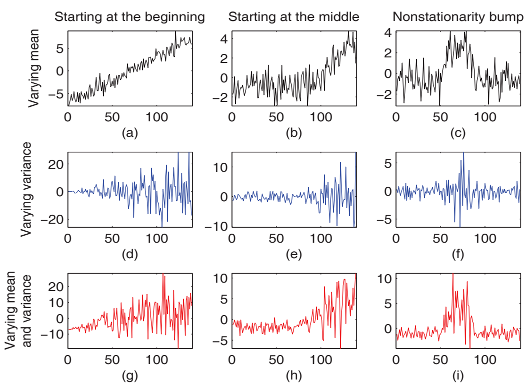

Fig. 1. Testing signals with a varying mean ((a),(b),(c)), variance ((d),(e),(f)) and both ((g),(h),(i)). From the first to the third column the adopted nonstationarity models are shown.

The stationarity test was applied sequentially (2500 times) in the signals presented in Fig. 1 and also in a stationary sequence of the same length ($N = 139$), drawn from a Gaussian distribution $\mathcal{N}(0, 1)$. The percentage of rights results (%) are shown in Table 1, where (μ), (σ) and (μ, σ) stand for nonstationarity in the mean, variance, and both, respectively. As it can be seen, the best performance is achieved by using the Itakura-Saito distance. The results were not constantly right for the combined distance. The changing outcomes are due to the randomness of the surrogate set, which allows fluctuations of the stationarity threshold if the test is being applied sequentially. Most important, none of the cases accused the signals having just first-order evolutions. In the following section we introduce a method to improve this situation.

5. WEIGHTING DISTANCES

When computing Θ_0, the vector of variances of the distances between local and global spectra, none consideration

Table 1. Percentage of right results (%) after applying the test 2500 times to the signals of Fig. 1 and in a stationary sequence. The Itakura-Saito and combined distances were used

Starting at the beginning			Starting at the middle			Nonstationarity bump		
μ	σ	μ, σ	μ	σ	μ, σ	μ	σ	μ, σ
Itakura-Saito distance (% right results)								
0	100	100	0	100	100	0	100	100
Combined distance (% right results)								
0	100	100	0	0.16	100	0	99.4	94.4
Stationary sequence $\mathcal{N}(0, 1)$ (% right results)								
100 (Itakura-Saito) and 100 (Combined)								

was taken about the spectral content itself. For instance, we would expect a signal with a varying mean having a spectrum more concentrated at low frequencies than one with a varying variance[1]. Also, the spectral content should be spread differently in time for both types of nonstationarities. As an example, let us consider the two signals having an exclusively varying mean and variance forming a *piecewise of stationarity* (Fig. 1(c) and (f)). They are shown again in Fig. 2(a) and (b), while their TF representations are shown in Fig. 2(c).

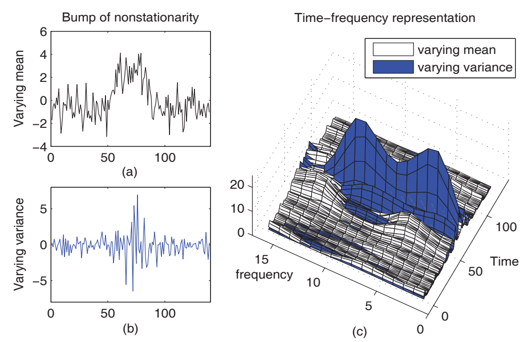

Fig. 2. Signals showing piecewise stationarity: (a) in the mean, (b) in the variance. (c) Their time-frequency representation.

Notice that the sequence with varying mean does not present a spectral content as significant at high frequencies as the one with varying variance. Therefore, we propose to use some more information from the signal and its spectrum to modify the distances (and consequently Θ_0), by building a weight vector either in time of frequency, so as to give more importance to those TF regions where the majority of signal is located.

[1] As it is known that the statistics of time-varying spectrum are specific at frequency near zero, this makes the method of [2] less applicable.

5.1. Marginals

To collect the required information from the TF plane, we might look at the influence of each frequency as the time runs *or* the significance of each time instant in the spectral domain. Actually, we are in fact recalling the marginal properties enjoyed by the Wigner-Ville representation. Since there are two sources from which we can extract an additional information, we need to choose one of them to build the weight vector.

The distances are computed from local time-varying spectrum (evaluated by multitaper spectrograms) at some time instants. So, while distances take as input functions of frequency, their output will be a function of time. Hence, the weight vector can be build in frequency or in time. If we sum over time (leading to a vector of weights in frequency), and multiply the spectra by the weight at the input of a given distance, we will be giving more importance to the most significant frequencies of the representation (globally). However, the spectral content from the surrogate is randomly spread over the TF plane, while properly equaling the global PSD of the original signal. By weighting the surrogate spectra, this random characteristic is affected. Moreover, the global PSD stops resembling the one from the original signal.

Conversely, summing over frequency leads to a vector representing the global spectral contribution in time. This vector can be normalized to unity, and thus be used multiplicatively to weight the distance output. *Doing so, we preserve the random spectral localization and the equality with the global PSD, which are the peculiarities from the surrogate set*. The scheme to choose the correct marginal and the obtained weight vector are shown in Fig. 3.

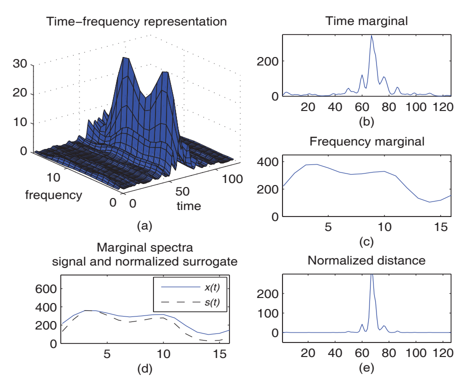

Fig. 3. Choosing the right marginal. (a) TF representation. (b) Marginal in time. (c) Marginal in frequency. (d) The global PSD from the surrogate is modified by spectral normalization. (e) Distance normalized by the correct marginal.

By collecting a vector of normalized distances for each surrogate and computing the variances, we obtain a new vector Θ_0', carrying information from the chosen marginal, where the gamma fit will be applied.

Considering these modifications, we applied the stationarity test as previously. The results are shown in Table 2.

Table 2. Percentage of right results ($\%$) after applying the test 2500 times in the signals of Fig. 1 and in a stationary sequence. The normalized distances were used

Starting at the beginning			Starting at the middle			Nonstationarity bump		
μ	σ	μ, σ	μ	σ	μ, σ	μ	σ	μ, σ
Itakura-Saito distance ($\%$ right results)								
23.6	100	100	100	100	100	98.7	100	100
Combined distance ($\%$ right results)								
0	100	100	40.3	99.7	100	95.6	100	100
Stationary sequence $\mathcal{N}(0, 1)$ ($\%$ right results)								
99 (Itakura-Saito) and 5.6 (Combined)								

Note that now we were able to detect a nonstationary mean. On the other hand, we observe a trade-off in weighting distances: although it allows the detection of first-order evolutions that were not accused by the original method, the test confidence related to the proper classification of a stationary sequence was significantly affected when using combined distance. In contrast, for the Itakura-Saito distance the test performance was improved significantly. The normalization allowed us to detect nonstationarities of both orders with just 1% of misclassification rate for the stationary sequence.

6. CONCLUSIONS

We observed that a recent stationarity test was not able to properly detect first-order nonstationarities. Here, we proposed two contributions that *improve the test performance* and *allow the detection of first-order nonstationarities*. First, we presented an adequate dissimilarity measure which increases the classification accuracies. Second, we performed a modification in the method in order to consider the spectral content from the signal itself. The latter was done by weighting the distances with the time marginal from the time-frequency distribution.

7. REFERENCES

[1] P. Flandrin, *Time-Frequency/Time-Scale Anaylsis*, Academic Press, San Diego, CA, 1999.

[2] P. Borgnat, P. Flandrin, P. Honeine, C. Richard, and J. Xiao, "Testing stationarity with surrogates: A time-frequency approach," *IEEE Trans. Signal Process.*, vol. 58, no. 7, pp. 3459–3470, Jul. 2010.

[3] M. Basseville, "Distances measures for signal processing and pattern recognition," *Signal Process.*, vol. 18, no. 4, pp. 349–369, 1989.