Open access • Journal Article • DOI:10.1109/TSP.2007.916125

On the Description of Spectrogram Probabilities With a Chi-Squared Law

Julien Huillery, Fabien Millioz, Nadine Martin

Institutions: Centre national de la recherche scientifique

Published on: 01 Jun 2008 - IEEE Transactions on Signal Processing (IEEE)

Topics: Spectrogram, Probability distribution, Likelihood-ratio test, Kullback–Leibler divergence and Law of total probability

On the Description of Spectrogram Probabilities with a Chi-Squared Law

Julien Huillery, Fabien Millioz, Nadine Martin

On the Description of Spectrogram Probabilities with a Chi-Squared Law

Julien Huillery, Fabien Millioz and Nadine Martin, *Member, IEEE*

***Abstract*—Given a correlated Gaussian signal, may a chi-squared law of probability always be used to describe a spectrogram coefficient distribution? If not, would a "chi-squared description" lead to an acceptable amount of error when detection problems are to be faced in the time-frequency domain? The two questions prompted the study reported in this paper. After deriving the probability distribution of spectrogram coefficients in the context of a non-centred Gaussian correlated signal, the Kullback-Leibler divergence is first used to evaluate to what extent the non-whiteness of the signal and the Fourier analysis window impact the probability distribution of the spectrogram. To complete the analysis, a detection task formulated as a binary hypothesis test is considered. We evaluate the error committed on the probability of false alarm when the likelihood ratio test is expressed with chi-squared laws. From these results, a chi-squared description of the spectrogram distribution appears accurate when the analysis window used to construct the spectrogram decreases to zero at its boundaries, regardless of the level of correlation contained in the signal. When other analysis windows are used, the length of the window and the correlation contained in the analysed signal impact the validity of the chi-squared description.**

***Index Terms*—Spectrogram probability distribution, Chi-squared law, Kullback-Leibler divergence, Time-Frequency statistical detection.**

I. Introduction

THIS paper focuses on the probability density function of spectrogram coefficients obtained by the squared modulus of a Discrete Fourier Transform (DFT). In the context of statistical signal processing, knowledge of these probability functions is necessary so as to develop detection and estimation methods dedicated to time-frequency analysis [1].

Under the assumptions of a white, centered and Gaussian signal, the spectrogram constructed with an infinite rectangular window is distributed as a chi-squared variable (noted χ^2) with 2 degrees of freedom [2]. Under the same assumptions, the impact of temporal windowing, zero-padding or spectral windowing was studied by Durrani [3][4]. His analysis reports a departure of the spectral coefficients distribution from the χ^2 law as soon as zero-padding or non-rectangular temporal windowing is used. The departure is evaluated in terms of a decrease in the equivalent degree of freedom. More recently, relaxing the whiteness assumption, Johnson and Long [5] derived a general form for the pdf of spectral estimates obtained by the Welch technique according to the number of periodograms averaged and their amount of overlap. Departure of the obtained pdf from a Gaussian distribution is evaluated with the Kullback-Leibler divergence.

Julien Huillery, Fabien Millioz and Nadine Martin are with the GIPSA-Lab, Signal and Image Department (DIS), INPG-CNRS, Grenoble, France.

As shown in these studies [3][4][5], the pdfs obtained are not trivial and of limited practical use. As a result, in applications, the spectrogram coefficients are classically described as χ^2 variables. For example in [6]($p262$), in the context of sinusoidal detection, the pdfs assumed under both hypotheses are χ^2 laws with 2 degrees of freedom. In [7] and [8], Martin also uses a χ^2 pdf with 2 degrees of freedom to describe the squared modulus of Fourier coefficients. He indicates the Fourier analysis window should be large enough and the signal correlation short enough so that this χ^2 description strictly holds.

This study aims at characterizing the differences between the spectrogram distribution and the χ^2 pdf for correlated Gaussian signals, as initiated in [9], and at giving a quantitative idea of the *short enough* or *long enough* conditions mentioned in [7].

We assume the signal observed, noted $x[m]$ in the discrete time domain, is composed of a deterministic part $d[m]$ embedded in a random additive perturbation $p[m]$:

$$x[m] = d[m] + p[m]. \tag{1}$$

The random perturbation is assumed stationary, Gaussian, centred, with an autocorrelation function $\Gamma_p[\tau]$. The symmetric covariance matrix associated with this autocorrelation function is noted $\mathbf{R}$. Each sample $x[m]$ of the signal is distributed as a Gaussian variable with mean $d[m]$ and autocorrelation function $\Gamma_p[\tau]$ and noted

$$x[m] \sim \mathcal{N}\left(d[m], \Gamma_p[\tau]\right). \tag{2}$$

From this temporal random model, the probability density function of a spectrogram coefficient can be evaluated. This probability function will depend on two kinds of parameters: the signal model parameters, namely $d[m]$ and $\mathbf{R}$, and the time-frequency transform parameters which, for the spectrogram, are the length and shape of the analysis window and the zero-padding.

The difference between χ^2 pdf and spectrogram pdf will first be evaluated in terms of the Kullback-Leibler (KL) divergence. Now, the amount of KL divergence between two pdfs is of great interest as a minimisation criterion, but does not easily provide a practical measure of accuracy as regards the approximation of one pdf with an other.

In order to give more insight to this question, a specific time-frequency detection task is investigated. For signal model (1), the detection task consists in deciding whether a considered time-frequency location contains energy originating from $d[m]$ (signal hypothesis H_1) or not (null hypothesis H_0). The Neyman-Pearson detector is used to evaluate the error

generated on the probability of false alarm when the χ^2 law is used instead of the exact spectrogram distribution.

This paper is organized as follows: Section II describes the time-frequency formulation of the signal random model. The χ^2 and spectrogram probability distributions are recalled in section III. Both central and non central cases are considered. In section IV, we describe how the parameters of the χ^2 law may be set to fit with the spectrogram pdf. In section V, the KL divergence is used as a first measure of accuracy of a χ^2 description of spectrogram. In section VI, the time-frequency detection problem is examined. Finally, a discussion on the respective sensibility of the two measures and the impact of the analysis window's shape is provided in section VII. Section VIII draws conclusions on the description of spectrogram probabilities with a χ^2 law for correlated signals.

II. Spectrogram and Random Time-Frequency Model

In this section, the random temporal model of signal (2) is formulated in the time-frequency domain according to the transform chosen in this work: the spectrogram.

Given a discrete analysis window $w[m]$ of length M, the Short-Time Fourier Transform $X[n,k]$ of a discrete signal $x[m]$ is formed by the successive Discrete Fourier Transforms of the windowed signal. Throughout this paper, the indices n and k will refer to the discrete time and frequency locations respectively. The spectrogram $S_x[n,k]$ corresponds to the squared modulus of the Short-Time Fourier Transform (STFT) or equivalently to the sum of the squares of the STFT real and imaginary parts, $X^r[n,k]$ and $X^i[n,k]$ respectively. We start with the following definitions:

$$S_x[n,k] = X^r[n,k]^2 + X^i[n,k]^2, \tag{3}$$

$$X^r[n,k] = \sum_{m=0}^{M-1} x[nD+m]w[m]\cos(-2\pi k\frac{m}{K}), \tag{4}$$

$$X^i[n,k] = \sum_{m=0}^{M-1} x[nD+m]w[m]\sin(-2\pi k\frac{m}{K}), \tag{5}$$

where K is the length of the computed DFT, D is the window sliding step and K/M corresponds to the factor of zero-padding.

Let us now define the STFT vector $\mathbf{X}[n,k]$ as

$$\mathbf{X}[n,k] = \begin{pmatrix} X^r[n,k] \\ X^i[n,k] \end{pmatrix}, \tag{6}$$

so that the spectrogram may be written

$$S_x[n,k] = \mathbf{X}^T[n,k]\mathbf{X}[n,k]. \tag{7}$$

Equations (4) and (5) express the real and imaginary parts of the STFT as linear combinations of signal samples. As $x[n]$ are Gaussian variables, $X^r[n,k]$ and $X^i[n,k]$ are also distributed with Gaussian laws and $\mathbf{X}[n,k]$ is a two-dimensional Gaussian vector characterized by five parameters: m_1, m_2, Σ_{11}, Σ_{22} and Σ_{12} defined as

$$\left\{\begin{array}{rcl} m_1 &=& \mathbb{E}\left\{X^r[n,k]\right\}, \\ m_2 &=& \mathbb{E}\left\{X^i[n,k]\right\}, \\ \Sigma_{11} &=& \mathbb{E}\left\{(X^r[n,k]-m_1)^2\right\}, \\ \Sigma_{22} &=& \mathbb{E}\left\{(X^i[n,k]-m_2)^2\right\}, \\ \Sigma_{12} &=& \mathbb{E}\left\{(X^r[n,k]-m_1)(X^i[n,k]-m_2)\right\}, \end{array}\right. \tag{8}$$

where $\mathbb{E}\{.\}$ denotes the expectation. For clarity, the time-frequency location $[n,k]$ in the probability parameter notations is dropped. The random time-frequency model can finally be written as

$$\mathbf{X}[n,k] \sim \mathcal{N}\left\{\begin{pmatrix} m_1 \\ m_2 \end{pmatrix}, \mathbf{\Sigma} = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12} & \Sigma_{22} \end{pmatrix}\right\}. \tag{9}$$

Let us now describe how the five parameters of this model can be written in terms of the signal and spectrogram parameters. Because the random perturbation $p[m]$ is assumed centred, the first order statistics m_1 and m_2 only depend on the deterministic STFT real and imaginary parts. From equations (1), (4) and (5), we have

$$\left\{\begin{array}{l} m_1 = Re(STFT\{d[m]\}), \\ m_2 = Im(STFT\{d[m]\}). \end{array}\right. \tag{10}$$

The centered second order statistics of $\mathbf{X}[n,k]$ are functions of the signal covariance matrix $\mathbf{R}$. Combined with the spectrogram parameters, we have in algebraic form

$$\left\{\begin{array}{rcl} \Sigma_{11} &=& \mathbf{W}^T\mathbf{C}_k\mathbf{R}\mathbf{C}_k\mathbf{W}, \\ \Sigma_{22} &=& \mathbf{W}^T\mathbf{S}_k\mathbf{R}\mathbf{S}_k\mathbf{W}, \\ \Sigma_{12} &=& \mathbf{W}^T\mathbf{C}_k\mathbf{R}\mathbf{S}_k\mathbf{W}, \end{array}\right. \tag{11}$$

where $\mathbf{W}^T = [w[0]\cdots w[M-1]]$ is the analysis window vector and $\mathbf{C}_k$ (resp. $\mathbf{S}_k$) is the cosine (resp. sine) diagonal matrix,

$$\mathbf{C}_k = diag\left[\cos(-2\pi k\frac{m}{K})\right]_{m=0,M-1}, \tag{12}$$

$$\mathbf{S}_k = diag\left[\sin(-2\pi k\frac{m}{K})\right]_{m=0,M-1}. \tag{13}$$

Note: as we are interested in a single time-frequency location $[n,k]$, the amount of overlap between the successive analysed frames of the signal (represented by the window sliding step D in (4) and (5)) is not involved in this time-frequency model and does not influence the results presented further.

III. χ^2 and spectrogram probability density functions

In this section the definition and pdf of a χ^2 variable are first recalled. Then the pdf of a spectrogram coefficient is expressed. For the central case ($m_1 = m_2 = 0$), an analytical formulation of the pdf is available. For the noncentral case however, the spectrogram pdf can not be written in a closed form. We present a numerical method based on geometrical considerations that can be used to compute the exact spectrogram pdf.

A. χ^2 law

Given N independent and homoscedastic (of same variance) Gaussian variables $G_i \sim \mathcal{N}\left(m_i, \sigma^2\right), i = 1, \ldots, N$, the sum of the squares of the variables G_i is distributed as a χ^2 variable. It is defined by three parameters and we note $\chi^2(\delta, \alpha, \theta)$ where

- δ is the degree of freedom. It corresponds to the number of independent Gaussian variables summed. $\delta = N$ in the example mentioned above.
- α denotes the coefficient of proportionality. It accounts for the common variance of the Gaussian variables. Here we have $\alpha = \sigma^2$.
- θ stands for the noncentrality parameter. It is defined in this work as $\theta = \sum_i m_i^2$.

For $\theta \neq 0$ and $x \geq 0$, the pdf of a χ^2 variable is

$$p_{\chi^2(\delta,\alpha,\theta)}(x) = \frac{1}{2\alpha}\left(\frac{x}{\theta}\right)^{\frac{\delta-2}{4}} \exp\left(-\frac{x+\theta}{2\alpha}\right) .I_{\frac{\delta-2}{2}}\left(\frac{\sqrt{x\theta}}{\alpha}\right), \quad (14)$$

where $I_n(.)$ stands for the n-order modified Bessel function of the first kind. In the central case ($\theta = 0$), the pdf is

$$p_{\chi^2(\delta,\alpha,\theta=0)}(x) = \frac{1}{(2\alpha)^{\frac{\delta}{2}}\Gamma(\frac{\delta}{2})} x^{\frac{\delta-2}{2}} \exp\left(-\frac{x}{2\alpha}\right), \quad (15)$$

where $\Gamma(n) = \int_0^{+\infty} e^{-x} x^{n-1} dx$ is the gamma function.

Note: the denomination "chi-squared variable" and its notation "χ^2" are usually dedicated to the central case only. In some reference literature [10], the central and noncentral cases are covered separately. The χ^2 variable is usually defined as the sum of the squares of Gaussian variables $\mathcal{N}(0,1)$ and requires only one parameter: the degree of freedom δ. The definition proposed here is based on non-centred and non-unit-variance Gaussian variables. This has led the authors to introduce three parameters (δ, α, θ) within the definition of the χ^2 variable. The reason for this choice is the informative nature of the means and variances of the Gaussian variables in the study presented here. They are part of the model and we found comfortable to handle them in a single formulation.

B. Spectrogram probability distribution

From the definition of the χ^2 law given above and the time-frequency model (9) described in section II, it follows that a spectrogram coefficient is distributed as a χ^2 variable as soon as the covariance matrix $\mathbf{\Sigma}$ is proportional to the identity matrix. As already reported in [9], this assumption is not always valid. The following pdfs are valid for any covariance matrix $\mathbf{\Sigma}$.

1) Central case, $m_1 = m_2 = 0$: Under the central case, the spectrogram pdf, noted $p_{S_x/0}(s)$, corresponds to the distribution of a quadratic form in two centered Gaussian variables. The moment-generating function of the random variable $\mathbf{X}[\mathbf{n},\mathbf{k}]^T\mathbf{X}[\mathbf{n},\mathbf{k}]$ is

$$M(\mu) = \mathbb{E}\left\{e^{-\mu \mathbf{X}^T\mathbf{X}}\right\} = \frac{1}{|\mathbf{I_2} + 2\mu\mathbf{\Sigma}|^{1/2}}. \quad (16)$$

The corresponding pdf is obtained by the inverse Laplace transform of $M(\mu)$ and is given by ([3], equation 106)

$$p_{S_x/0}(s) = \frac{1}{a}\exp(-bs) I_0(cs), s \geq 0. \quad (17)$$

Following [3] (equations 107 or 109), the three parameters a, b and c are obtained with

$$\begin{cases} a = \sqrt{4\det\mathbf{\Sigma}}, \\ b = \frac{\mathrm{tr}\,\mathbf{\Sigma}}{4\det\mathbf{\Sigma}}, \\ c = \frac{\left[(\mathrm{tr}\,\mathbf{\Sigma})^2 - 4\det\mathbf{\Sigma}\right]^{\frac{1}{2}}}{4\det\mathbf{\Sigma}}, \end{cases} \quad (18)$$

where '$\det\mathbf{\Sigma}$' and '$\mathrm{tr}\,\mathbf{\Sigma}$' stand for the determinant and trace of the covariance matrix $\mathbf{\Sigma}$ respectively. For $\mathbf{\Sigma}$ proportional to identity, we have $a = 1/b = 2\Sigma_{11}$ and $c = 0$, leading to the central χ^2 law (15) with $\delta = 2$ and $\alpha = \Sigma_{11} = \Sigma_{22}$.

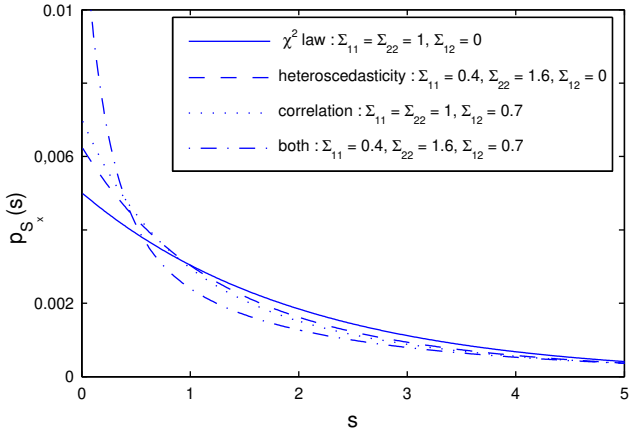

Fig. 1. Examples of central distributions corresponding to different covariance matrices $\mathbf{\Sigma}$.

Some examples of central spectrogram pdfs are displayed in figure 1. The continuous line plot corresponds to the χ^2 law. The three other dashed plots report the evolution of spectrogram probability law when heteroscedasticity (difference of variances, $\Sigma_{11} \neq \Sigma_{22}$) and/or correlation ($\Sigma_{12} \neq 0$) happens between the components of vector $\mathbf{X}[n,k]$.

2) Noncentral case, $m_1, m_2 \neq 0$: No closed form was derived for the pdf of a quadratic form in non-centered Gaussian variables. The literature provides many series expansions using Laguerre polynomials [11][12], χ^2 pdfs [13][14][15] or the hypergeometric function [5]. While one of the references cited above could have been used, in this paper we propose a simple geometrical approach to compute the pdf $p_{S_x}(s)$ of a spectrogram coefficient. The method is applicable to any squared modulus of a two-dimensional random vector, centred or not, Gaussian or not.

We look for the density of probability for which a spectrogram coefficient $S_x[n,k]$ equals a given positive value s.

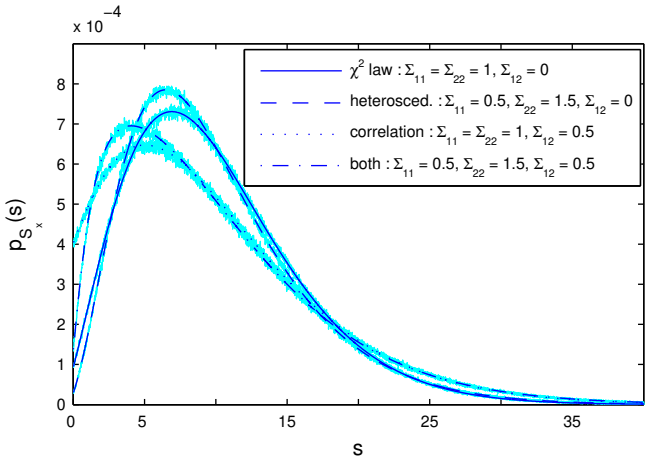

Fig. 2. Examples of noncentral distributions ($m_1^2 + m_2^2 = 8$) corresponding to different covariance matrices $\mathbf{\Sigma}$. The histograms of 10^6 equivalent random variables are also plotted (background colour).

From (3), this event happens when the realisations x_r and x_i of the two random variables $X^r[n,k]$ and $X^i[n,k]$ satisfy the equation $x_r^2 + x_i^2 = s$. Within the (X^r, X^i) plane, this describes the circle centred at (0,0) and with radius $\sqrt{s}$. The spectrogram pdf $p_{S_x}(s)$ can thus be obtained by integrating the joint pdf $p_{(X^r,X^i)}(x_r, x_i)$ of the two random variables X^r and X^i over this domain.

To adapt the notation to the geometry of the problem, the pdf $p_{(X^r,X^i)}(x_r, x_i)$ is expressed in polar coordinates (r, ϕ) using

$$\left\{ \begin{array}{rcl} x_r &=& r\cos\phi, \\ x_i &=& r\sin\phi. \end{array} \right. \tag{19}$$

The spectrogram pdf $p_{S_x}(s)$ is now obtained with integration over the angular coordinate ϕ over $[0;\ 2\pi]$ and is written

$$p_{S_x}(s) \;=\; \int_0^{2\pi} p_{(X^r,X^i)}(\sqrt{s}, \phi) d\phi. \tag{20}$$

In this work, the STFT vector $\mathbf{X}[n,k]$ is Gaussian. $p_{(X^r,X^i)}$ is a bi-dimensional Gaussian distribution which is formulated in polar coordinates as

$$p_{(X^r,X^i)}(r,\phi) \;=\; \frac{\exp\left(-\frac{A(r,\phi)}{2(1-\rho^2)}\right)}{2\pi\sqrt{\Sigma_{11}\Sigma_{22}(1-\rho^2)}}, \tag{21}$$

where

$$\rho \;=\; \frac{\Sigma_{12}}{\sqrt{\Sigma_{11}\Sigma_{22}}} \tag{22}$$

is the correlation coefficient between STFT real and imaginary parts and

$$\begin{aligned} A(r,\phi) \;=\; & \frac{(r\cos\phi - m_1)^2}{\Sigma_{11}} + \frac{(r\sin\phi - m_2)^2}{\Sigma_{22}} \\ & - \frac{2\rho(r\cos\phi - m_1)(r\sin\phi - m_2)}{\sqrt{\Sigma_{11}\Sigma_{22}}}. \end{aligned} \tag{23}$$

A discrete version of (20) is used to compute the spectrogram pdf. Some examples of noncentral distributions are displayed in figure 2 for different covariance matrices $\mathbf{\Sigma}$. In each configuration, the histogram of 10^6 runs of the corresponding random variable is also plotted so as to validate the accuracy of the method. The continuous line plot corresponds to a noncentral χ^2 distribution ($\mathbf{\Sigma} = I_2$). The dashed plots show the evolution of the spectrogram pdf when the components of the STFT vector $\mathbf{X}[n,k]$ have different variances and/or are correlated.

IV. Description of the spectrogram with a χ^2 law

In this section, we present two different settings of the χ^2 law parameters (δ, α and θ) that can be used to approximate the spectrogram distribution. The noncentrality parameter θ can be treated separately as it does not generate any difference between spectrogram and χ^2 pdfs. For all cases, it will be given by $\theta = m_1^2 + m_2^2$. The differences between spectrogram and χ^2 pdfs originate from the covariance matrix $\mathbf{\Sigma}$. Two different ways to link the χ^2 parameters α and δ to this matrix are proposed.

A. Setting 1: fixed δ

In a first approach, the degree of freedom is fixed at $\delta = 2$, which means considering STFT real and imaginary parts as independent and homoscedastic. This corresponds to the commonly used χ^2 law and leads to the most simple pdf. The proportionality coefficient α is set according to its maximum likelihood estimator for a central ($\theta = 0$) χ^2 law with $\delta = 2$ degrees of freedom. This estimator is the arithmetic mean of both real and imaginary STFT part variances. The three parameters of the first χ^2 law used to describe the spectrogram distribution are written as

$$\left\{ \begin{array}{rcl} \delta &=& 2, \\ \alpha &=& \frac{\Sigma_{11}+\Sigma_{22}}{2}, \\ \theta &=& m_1^2 + m_2^2. \end{array} \right. \tag{24}$$

The corresponding pdf is

$$\begin{aligned} p_{\chi^2}(s) \;=\; & \frac{1}{\Sigma_{11}+\Sigma_{22}} \exp\left(-\frac{s + m_1^2 + m_2^2}{\Sigma_{11}+\Sigma_{22}}\right) \\ & .I_0\left(\frac{\sqrt{2s(m_1^2+m_2^2)}}{\Sigma_{11}+\Sigma_{22}}\right). \end{aligned} \tag{25}$$

B. Setting 2: adapted δ

The degree of freedom δ of a χ^2 law reflects the number of independent variables that are summed. If the two Gaussian variables are correlated (*i.e* $\Sigma_{12} \neq 0$), the equivalent degree of freedom becomes smaller than 2. Also, if the two variances Σ_{11} and Σ_{22} are not equal, one of the Gaussian variable has more impact on the sum of the squares. Consequently the equivalent number of independant variables is not anymore 2 but lies between 1 ($\Sigma_{11} \gg \Sigma_{22}$) and 2 ($\Sigma_{11} = \Sigma_{22}$). As a result, adapting the degree of freedom δ according to the covariance matrix $\mathbf{\Sigma}$ of the time-frequency model should lead to a better description of the spectrogram.

The "method of moments" [1] is used for the joint estimation of the parameters δ and α. The first two statistical moments

[1]The method of moments consists in equating the statistical moments between the empirical observations and the model.

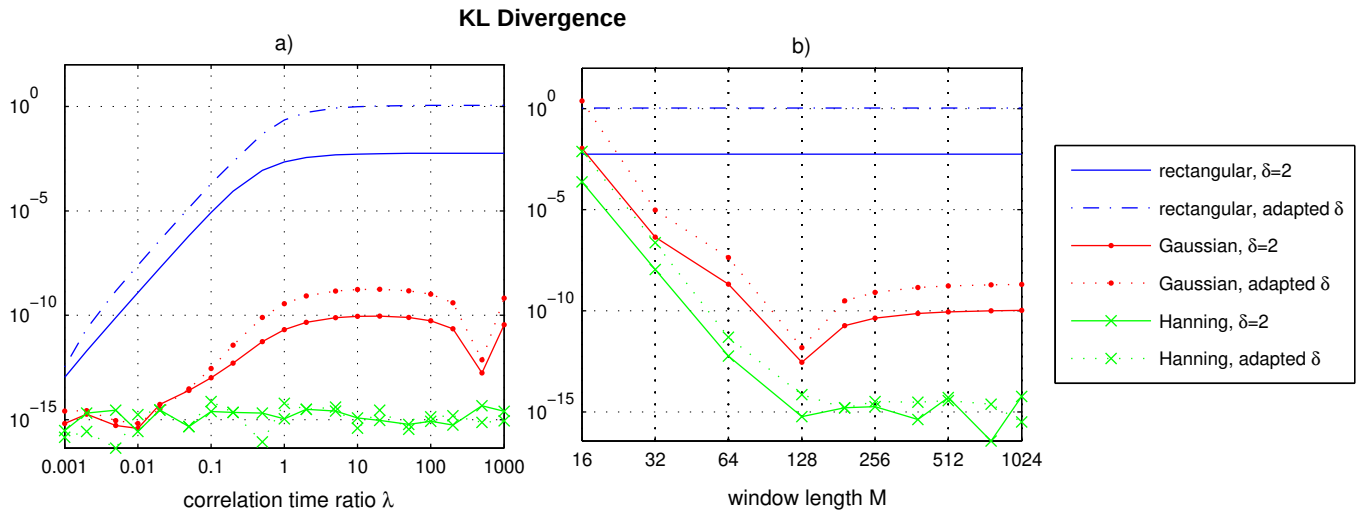

Fig. 3. Kullback-Leibler divergence between spectrogram and χ^2 distribution as a function of a) the correlation time ratio λ for 512-point long windows and b) the length M of the analysis window for a correlation time ratio $\lambda = 30$.

of a central χ^2 variable are

$$\left\{ \begin{array}{rcl} \mathbb{E}\left\{\chi^2\right\} & = & \alpha\delta, \\ \mathbb{V}ar\left\{\chi^2\right\} & = & 2\alpha^2\delta, \end{array} \right. \tag{26}$$

where $\mathbb{V}ar\{.\}$ stands for the variance operator. Now the first two statistical moments of a 'noise-only' spectrogram coefficient are

$$\left\{ \begin{array}{rcl} \mathbb{E}\left\{S_x\right\} & = & \Sigma_{11} + \Sigma_{22}, \\ \mathbb{V}ar\left\{S_x\right\} & = & 2[\Sigma_{11}^2 + \Sigma_{22}^2 + 2\Sigma_{12}^2]. \end{array} \right. \tag{27}$$

Equating the respective statistical moments leads to the new description

$$\left\{ \begin{array}{rcl} \delta & = & \frac{(\Sigma_{11}+\Sigma_{22})^2}{\Sigma_{11}^2+\Sigma_{22}^2+2\Sigma_{12}^2}, \\ \alpha & = & \frac{\Sigma_{11}+\Sigma_{22}}{\delta}, \\ \theta & = & m_1^2 + m_2^2. \end{array} \right. \tag{28}$$

The final corresponding χ^2 pdf is obtained through (14).

To obtain more insight into this new setting, let us define a coefficient of heteroscedasticity ξ between the STFT real and imaginary parts as

$$\xi = \frac{\frac{1}{2}(\Sigma_{11} + \Sigma_{22})}{\sqrt{\Sigma_{11}\Sigma_{22}}}. \tag{29}$$

The degree of freedom δ can be rewritten as

$$\delta = \frac{2}{2 - \frac{1-\rho^2}{\xi^2}}, \tag{30}$$

where ρ is the correlation coefficient (22). This formulation highlights the impact of both correlation and heteroscedasticity that tend to diminish the equivalent degree of freedom δ of the spectrogram coeffcient $S_x[n,k]$, as they increase.

The two proposed χ^2 parameter settings lead to two possible approximations of the spectrogram distribution. To evaluate these approximations, the χ^2 pdfs have to be compared with the exact spectrogram pdf obtained with (20). The Kullback-Leibler divergence is now used to this purpose.

V. KL DIVERGENCE MEASUREMENTS

In this section, the Kullback-Leibler divergence $\mathcal{J}(p_{\chi^2}||p_{S_x})$ is used as a measure of distance between the spectrogram and χ^2 distributions. It is calculated as [16]

$$\mathcal{J}(p_{\chi^2}||p_{S_x}) = \int_0^{+\infty} p_{\chi^2}(s) \log_e \left[\frac{p_{\chi^2}(s)}{p_{S_x}(s)}\right] ds. \tag{31}$$

As the natural logarithm is used in this expression, the KL divergence unit is the 'nat'.

In the sequel we consider a stationary, centred, Gaussian and exponentially correlated process $p[m]$. The autocorrelation function of this process is

$$\Gamma_p[\tau] = \Gamma_0 \exp\left(-\frac{|\tau|}{\tau_c}\right), \tag{32}$$

where τ_c is the correlation time and Γ_0 is set to 1.

1) Impact of correlation: For a given correlation time τ_c, the impact on the spectrogram pdf depends on the analysis window's length M. To analyse the impact of the correlation separately from the impact of the window's length, the correlation time ratio λ is defined as

$$\lambda = \frac{\tau_c}{M}. \tag{33}$$

This parameter represents the span of correlation at the scale of the analysis.

As shown in Figure 3-a), the KL divergence increases as the amount of correlation becomes higher. The spectrogram probability distribution moves away from a χ^2 law. When the correlation time τ_c is of the order of the window's length M (*i.e* $\lambda = 1$), the KL divergence is stabilized.

2) Impact of windowing: Figure 3-a) also highlights the different sensibilities of the spectrogram pdf to signal correlation according to the shape of the analysis window. The KL divergence increases and is maximal with the rectangular

window while it remains constant with the Hanning window. This point will be discussed in section VII.

Figure 3-b) reports the KL divergence between spectrogram and χ^2 pdfs according to the analysis window's length. The correlation time ratio is fixed at $\lambda = 30$. As in a white environment [3], the distance between spectrogram and χ^2 pdfs diminishes as the analysis window lengthens. Also, the KL divergence is stabilized for windows longer than a threshold that grows with the correlation time ratio λ (we found 64 samples for $\lambda = 1$, 128 for $\lambda = 30$ and 512 for $\lambda = 100$).

Dissimilarities with the white case also are to be noted. In a white environment, a long rectangular window assures the best fit between spectrogram and χ^2 pdfs as opposed to other windows [3]. The opposite phenomenon is observed in a correlated environment : the best fit is obtained for long and non-uniform windows. With the rectangular window, The KL divergence remains constant over the whole range of window lengths.

3) Impact of zero-padding: The impact of zero-padding was also investigated. The amount of KL divergence between spectrogram and χ^2 pdfs remained constant as the factor of zero-padding K/M went from 1 to 3. Our experiments do not show that this parameter influences the probability distribution of spectrogram coefficients.

4) Impact of the degree of freedom: An unexpected difference of KL divergence is observed between the two proposed χ^2 pdfs. The χ^2 law with $\delta = 2$ degrees of freedom appears closer to the spectrogram pdf than its counterpart with adapted δ. This point will be discussed in section VII.

The interpretation of a numerical amount of KL divergence between two probability distributions is difficult. This is due to a lack of a practical normalisation method and the non-symmetric behavior of this measure. Hence, only qualitative conclusions have been made from this point. In the following section, a time-frequency detection problem is investigated. This will provide more quantitative references so as to answer the question: what is the level of inaccuracy and on which parameters attention should specifically be paid if a χ^2 law is used to describe the spectrogram statistics?

VI. Approximation in a detection context

In this section, the impact of describing the spectrogram distribution with a χ^2 law is evaluated in a time-frequency detection context.

A. Problem formulation

We consider the model of observation (1) where the signal under interest is embedded in an additive random perturbation. The time-frequency detection task consists in determining whether the energy $S_x[n,k]$ observed at a particular time-frequency location $[n,k]$ originates from the perturbation $p[m]$ only (null hypothesis H_0) or is also due to the deterministic part $d[m]$ of the signal (signal hypothesis H_1). This time-frequency binary hypotheses test is formulated as

$$\begin{cases} H_0 : S_x[n,k] = S_p[n,k], \\ H_1 : S_x[n,k] = S_{d+p}[n,k]. \end{cases}$$

In the Neyman-Pearson detection approach, the detection threshold S^{th} is determined by means of a given probability of false alarm P_f. S^{th} is calculated so as to satisfy the relation

$$P_f = \int_{S^{th}}^{+\infty} p_{H_0}(s)ds, \tag{34}$$

where $p_{H_0}(s)$ is the spectrogram pdf under the null hypothesis H_0. The corresponding decision rule is

$$\begin{cases} \text{decide } H_0 \text{ if } S_x[n,k] < S^{th}, \\ \text{decide } H_1 \text{ if } S_x[n,k] > S^{th}. \end{cases}$$

B. Impact measurements

The overall probability of error of a binary hypotheses test is the sum of two terms: one is the probability of false alarm (PFA) while the other is the probability of miss-detection. The Neyman-Pearson detection strategy minimizes the probability of miss-detection while restraining the probability of false alarm to a given value. From this point of view, the respect of a chosen PFA is the main focus of this detection strategy. Now, if an approximation p_{ap} is used instead of the exact pdf p_{H_0}, the obtained probability of false alarm $P_{f,obtained}$ of the test will shift from the one initially desired $P_{f,wanted}$. More precisely, the detection threshold S^{th} will satisfy

$$P_{f,wanted} = \int_{S^{th}}^{+\infty} p_{ap}(s)ds, \tag{35}$$

whereas the effective size of the test will be

$$P_{f,obtained} = \int_{S^{th}}^{+\infty} p_{H_0}(s)ds. \tag{36}$$

We propose to consider the error Δ_{PFA} defined as

$$\Delta_{PFA} = |P_{f,obtained} - P_{f,wanted}| \tag{37}$$

$$= \int_{S^{th}}^{+\infty} |p_{H_0}(s) - p_{ap}(s)|ds \tag{38}$$

so as to evaluate the practical significance of the divergence observed between spectrogram and χ^2 pdfs. In a system, the error engendered by a modelisation mismatch has to be lower than the desired precision. So if the variation of PFA Δ_{PFA} appears higher than the desired PFA $P_{f,wanted}$, the χ^2 approximation has to be rejected. If the error Δ_{PFA} is much lower than $P_{f,wanted}$, the modelisation mismatch can be judged insignificant as it does not noticeably impact the detection test performances.

1) Impact of correlation: The error Δ_{PFA} on the PFA as a function of the correlation time ratio λ is reported in figure 4-a). The desired PFA is fixed at $P_{f,wanted} = 10^{-5}$. As the analysed signal becomes more correlated, the shift of PFA Δ_{PFA} reaches the significance level of 10^{-5} with the rectangular window only. For the Gaussian window, variations of PFA are stabilized at the level of 10^{-8}. Correspondingly,

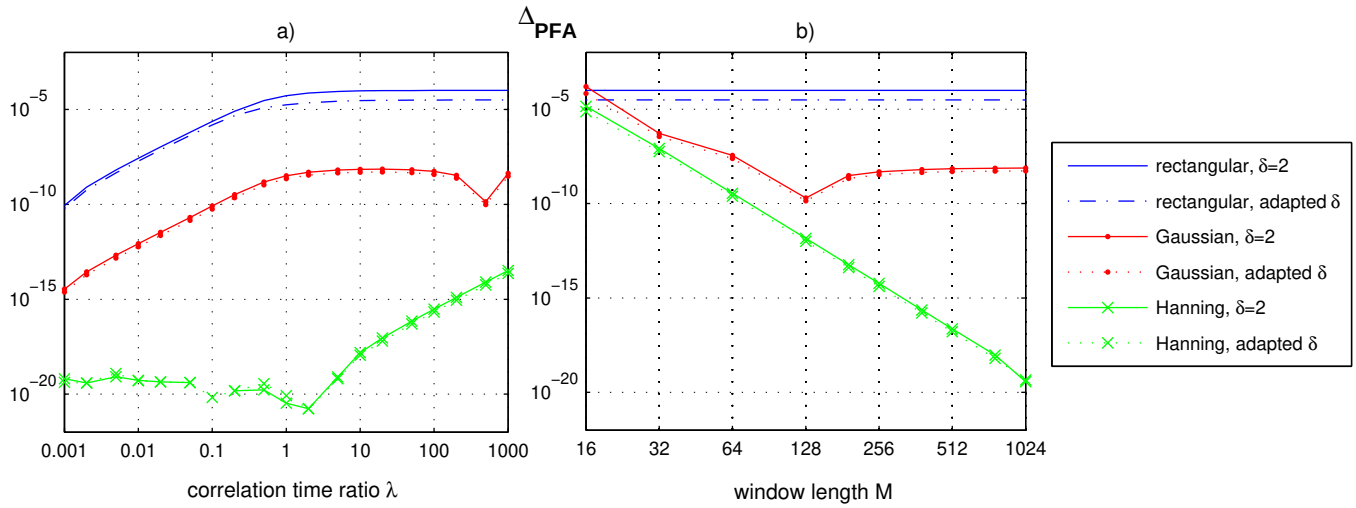

Fig. 4. Error on the PFA as a function of a) the correlation time ratio λ for 512-point long windows and b) the length M of the analysis window for correlation time ratio $\lambda = 30$. For both cases, the initial desired PFA is 10^{-5}.

the modelisation mismatch impact is of order 10^{-3} compared with the test precision and can be judge insignificant. The Hanning window generates even lower and insignificant impact.

2) Impact of the window's length: Errors on the PFA as a function of the window's length M are reported in figure 4-b). The correlation time ratio is fixed at $\lambda = 30$ and the desired PFA is $P_{f,wanted} = 10^{-5}$. For very short analysis windows (16 or 32 samples), the impact of a χ^2 approximation appears significant. However, as soon as the Gaussian or Hanning windows are longer than 64 samples, the effects observed appear non significant. The situation is different with the rectangular window: increasing the window's length has no effect and the variation of the PFA is always above the significance level.

3) Impact of the degree of freedom: For both figures 4-a) and -b), the continuous lines correspond to χ^2 laws with fixed $\delta = 2$ degree of freedom, while dotted lines stand for χ^2 laws with an adapted degree of freedom. Contrary to the KL divergence measurements, the errors on PFA are lower when the degree of freedom is adjusted to the covariance matrix $\mathbf{\Sigma}$. This is especially true as the analysed signal becomes more and more correlated and for high errors Δ_{PFA}. However, the improvement is not noticeable enough to generate non-significant errors.

VII. Discussion

A. Which χ^2 law?

We tried to approximate the spectrogram probability distribution with two different χ^2 laws according to a fixed or adapted degree of freedom. When the approximation was evaluated with the KL divergence, the χ^2 law with 2 degrees of freedom engendered the best fit with the spectrogram

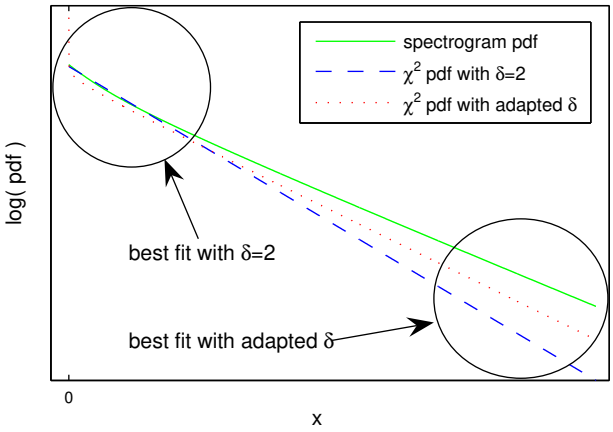

Fig. 5. Differences between the χ^2 pdfs with fixed or adapted degree of freedom δ and comparison with the exact spectrogram pdf.

pdf. Conversely, in the detection test, adapting the degree of freedom of the χ^2 law appeared favorable.

The different sensibilitiy of the two measures can explain this observation: our point is that KL divergence is particularly sensitive to differences located on the main body of the two probability distributions. Also, differences between the tails of the distributions have a small impact on the KL divergence. As an explanation, the KL divergence is an expectation and hence a summation weighted by a pdf. Consequently, the differences located where the pdf is high give more value to the KL divergence than the differences located where the pdf is small. On the other hand, it is intuitive that a quantile corresponding to a given PFA is particularly sensitive to the tail of the distribution. Henceforth, the two criteria used in this study mostly react to different modelisation mismatches.

The exact spectrogram pdf as well as the two proposed

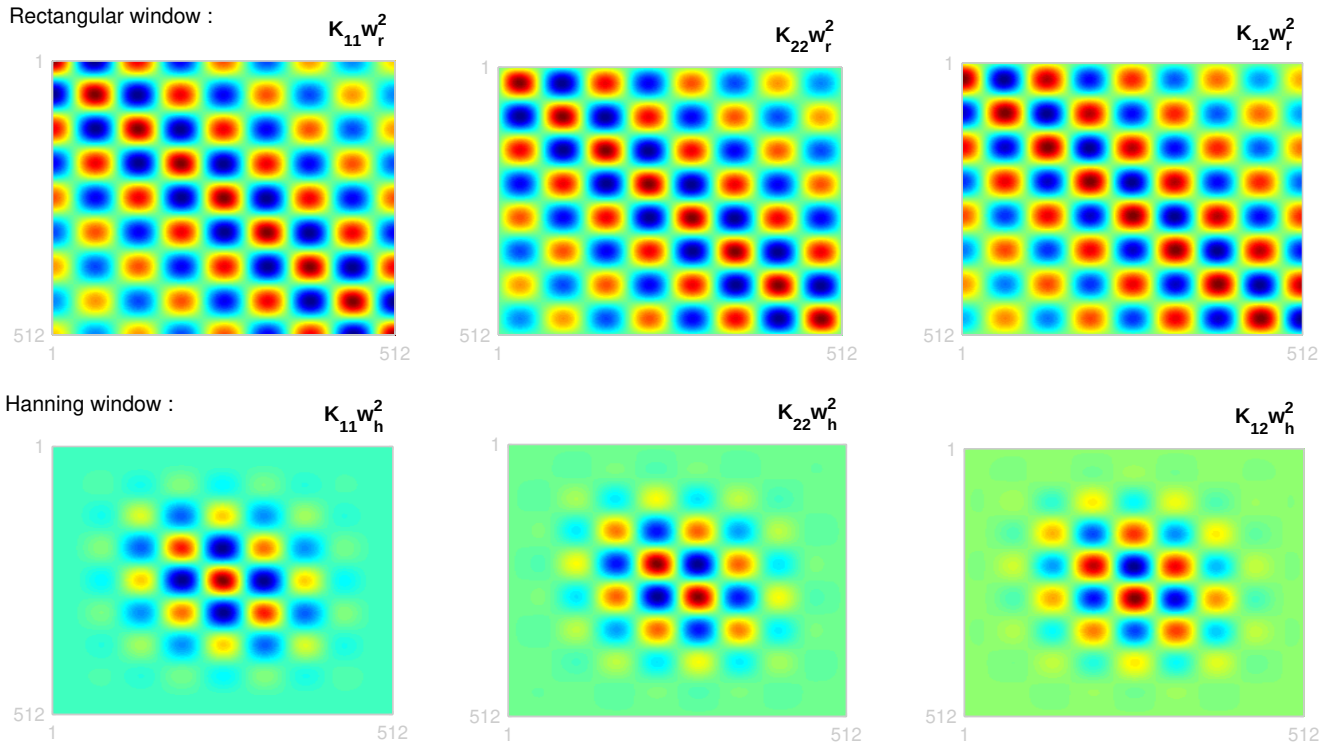

Fig. 6. Covariance kernel-window products before summation (42) as used to evaluate the covariances Σ_{11}, Σ_{22} and Σ_{12} for 512-point rectangular (top) and Hanning (bottom) windows. Observation frequency is $k = 4$, correlation time ratio is $\lambda = 1$. $\Sigma_{11}/\Sigma_{22} = 0.4415$ and $\rho = 0.0206$ for the rectangular window, $\Sigma_{11}/\Sigma_{22} = 0.9965$ and $\rho = 8, 3.10^{-5}$ for the Hanning window.

χ^2 approximations are plotted in log scale in figure 5. The spectrogram is constructed with a 512-point rectangular window and the correlation time ratio is $\lambda = 10$. The fixed degree of freedom produces the best fit regarding the main body of the distributions. Consequently, the KL divergence was smaller. On the other hand, the tail of the spectrogram pdf is better approximated when the degree of freedom of the χ^2 law is adapted, thus producing the smallest variations of PFA in the detection test.

Now, the use of a χ^2 law with fixed or adapted degree of freedom depends on the main focus of the approximation. If the approximation has to be global, the χ^2 law with $\delta = 2$ degrees of freedom is recommended. If the approximation has to respect the tail of the distribution, adapting the degree of freedom is favorable.

B. Impact of the window shape

Throughout this study, noticeable differences among the analysis windows were observed. The probability distribution of a spectrogram constructed with a Hanning window appeared insensitive to the correlation of the analysed signal. Conversely, the rectangular window reacted significantly to a correlated environment, leading to important modelisation mismatches. The sensibility of the Gaussian window stands in the middle.

As observed in [9], different windows lead to different covariance matrices $\mathbf{\Sigma}$. With a Hanning window, $\mathbf{\Sigma}$ is always nearly proportional to identity, *i.e* $\Sigma_{11}/\Sigma_{22} = 1$ and $\Sigma_{12} = 0$, regardless of the level of correlation contained in the signal. With the rectangular window, the correlation generates large differences between Σ_{11} and Σ_{22}, and non-null Σ_{12}. Consequently, the covariance matrix $\mathbf{\Sigma}$ is far from identity and the spectrogram distribution far from a χ^2 law.

Σ_{11}, Σ_{22} and Σ_{12}, defined in (11), are quadratic forms in the window vector $\mathbf{W}$. Let us call 'covariance kernels' the three matrices K_{11}, K_{22} and K_{12} associated with these quadratic forms, corresponding to

$$\mathbf{K_{11}} = \mathbf{C}_k\mathbf{R}\mathbf{C}_k, \quad (39)$$

$$\mathbf{K_{22}} = \mathbf{S}_k\mathbf{R}\mathbf{S}_k, \quad (40)$$

$$\mathbf{K_{12}} = \mathbf{C}_k\mathbf{R}\mathbf{S}_k. \quad (41)$$

From simple matrix manipulation, the definition of Σ_{11} can be reformulated as

$$\Sigma_{11} = \sum_{i,j} \mathbf{K_{11}}(i,j)\mathbf{W^2}(i,j), \quad (42)$$

where $\mathbf{W^2} = \mathbf{W}\mathbf{W}^T$. Σ_{11} appears as the summation of the covariance kernel elements weighted by the analysis window matrix $\mathbf{W^2}$. A similar formulation of Σ_{22} and Σ_{12} is also valid. Figure 6 displays the covariance kernel-window products before summation (42) in the case of a rectangular window (top three panels) and a Hanning window (bottom panels). The correlation time ratio is $\lambda = 1$.

The cosine and sine functions are dephased by $\pi/2$. Consequently, the covariance kernel $\mathbf{K_{22}}$ (formed with sines)

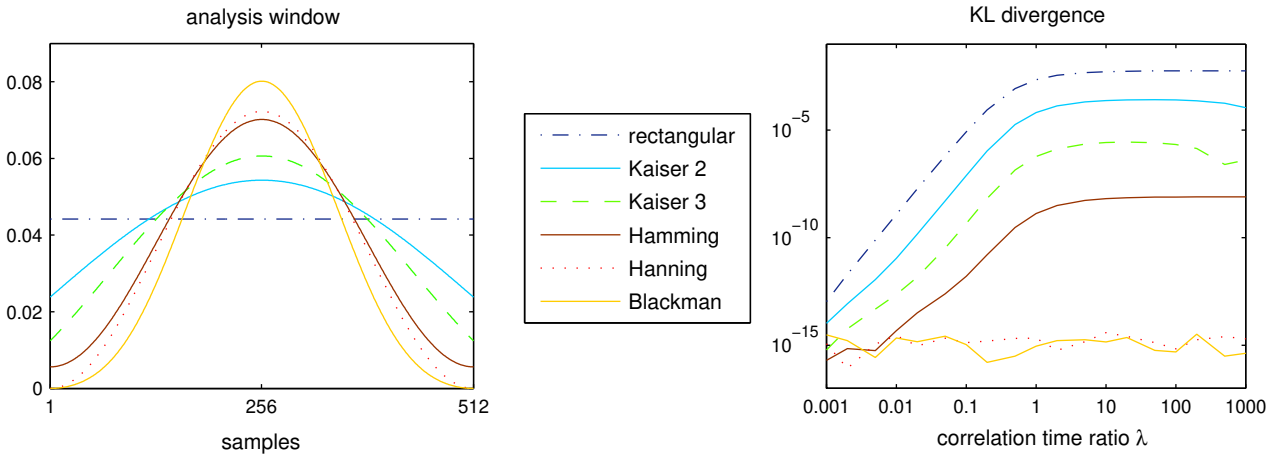

Fig. 7. Energy-normalized analysis windows and corresponding Kullback-Liebler divergences between spectrogram and χ^2 pdfs as a function of the correlation time ratio λ.

can be obtained from $\mathbf{K_{11}}$ (formed with cosines) with two translations, one horizontal and one vertical, of $T_k/4$ samples where T_k is the period of the sine and cosine at the observed frequency k. As seen in figure 6 (top panels), noticeable differences between $\mathbf{K_{11}}$ and $\mathbf{K_{22}}$ are present on the boundaries of the kernels.

If no correlation is present in the analysed signal, the kernels reduce to diagonal matrices, no differences exist and the spectrogram is distributed as a χ^2 variable. Now, if correlation exists in the analysed signal, the covariance kernels extend apart from the main diagonal and differences appear on the boundaries.

According to the shape of the analysis window, these differences will be preserved or not. The rectangular window represents a uniform weighting and preserves such differences. As a result, Σ_{11} differs from Σ_{22}. If the analysis window decreases to zero at its boundaries, as the Hanning window, the impact of the aforementioned differences is drastically reduced. The summation (42) leading to Σ_{11} and Σ_{22} only concerns the middle of the covariance kernels where no such differences exist (see figure 6, bottom panels). As a consequence, the correlation of the analysed signal does not lead to different Σ_{11} and Σ_{22}, and the spectrogram remains distributed as a χ^2 variable.

From these observations, the behavior of the analysis window at its boundaries appears to be responsible for the more or less important impact of correlation on the spectrogram probability distribution. Several currently used analysis windows are depicted in figure 7. For each of these windows, the KL divergence between the corresponding spectrogram pdf and a χ^2 law with $\delta = 2$ degrees of freedom is evaluated as a function of the correlation time ratio λ. This figure clearly shows the link between the behavior of the windows at their boundaries and the corresponding difference between spectrogram and χ^2 pdfs.

VIII. Conclusions

May the probability distribution of a spectrogram coefficient be accurately described with a χ^2 law when the analysed signal is embedded in a correlated centred Gaussian perturbation? What is the influence of the nature and length of the analysis window, the zero-padding or the amount of correlation in the signal?

In this paper, we present two experiments to answer these questions. The first consists in measuring the KL divergence between the exact spectrogram distribution and the χ^2 law whose parameters are tuned to match the first statistical moments of the spectrogram. The second focuses on a more practical situation: a detection task in the time-frequency domain. We evaluate the deviation of PFA engendered by the approximation of the spectrogram pdf with a χ^2 law. These two experiments lead to the following conclusions:

- No restrictions can be formulated as to the use of zero-padding. Its impact on the spectrogram probability distribution was found null.
- The spectrogram probability distribution differs from a χ^2 law as the amount of correlation in the signal increases. However, the difference is linked with the behavior of the analysis window at its boundaries. For windows with null boundaries (as the Hanning or Blackman windows), the spectrogram pdf remains insensitive to correlation.
- Longer non-uniform analysis windows increase the resemblance between spectrogram and χ^2 pdfs for correlated signals.
- The χ^2 law with 2 degrees of freedom provides the best fit with the overall spectrogram pdf. However, adapting the degree of freedom provides a better approximation of the tail of the distribution.
- The differences observed between spectrogram and χ^2 pdfs are significant only when the spectrogram is constructed with a rectangular window whose length is smaller than the correlation time of the signal. For windows with null boundaries, approximating the spectrogram pdf with a χ^2 law generates insignificant modelisation mismatches, regardless of the level of correlation.

ACKNOWLEDGMENT

The authors would like to thank the three anonymous reviewers for their comments and their help in the presentation of this manuscript.

REFERENCES

[1] R. A. Altes, "Detection, estimation and classification with spectrograms," *J. Acoust. Soc. Am.*, vol. 67, no. 4, pp. 1232–1245, April 1980.
[2] L. K. Koopman, *The spectral analysis of time series*. Academic Press, 1974.
[3] T. Durrani and J. Nightingale, "Probability distributions for discrete fourier spectra," *Proc. Inst. Elec. Eng*, vol. 120, no. 2, pp. 299–311, 1973.
[4] T. Durrani, "Joint density functions for digital spectra," *IEEE Trans. on Acoustics, Speech and Signal Processing*, vol. 22, no. 5, pp. 314–320, October 1974.
[5] P. Johnson and D. Long, "The probability density of spectral estimates based on modified periodogram averages," *IEEE Trans. on Signal Processing*, vol. 47, no. 5, pp. 1255–1261, May 1999.
[6] S. M. Kay, *Fundamentals of Statistical Signal Processing, vol.II - Detection Theory.* Prentice Hall, 1998.
[7] R. Martin, "Noise power spectral density estimation based on optimal smoothing and minimum statistics," *IEEE Trans. on Speech and Audio Processing*, vol. 9, no. 5, pp. 504–512, July 2001.
[8] ——, "Bias compensation methods for minimum statistics noise power spectral density estimation," *Signal Processing*, vol. 86, pp. 1215–1229, 2006.
[9] J. Huillery, F. Millioz, and N. Martin, "On the probability distributions of spectrogram coefficients for correlated gaussian process," *In proceedings of the ICASSP'06, Toulouse, France*, May 2006.
[10] N. Johnson, S. Kotz, and N. Balakrishnan, *Continuous univariate distributions*, 2nd ed. Wiley and sons, 1995, vol. 2.
[11] J. Gurland, "Distribution of definite and of indefinite quadratic forms," *Annals of Mathematical Statistics*, vol. 26, pp. 122–127, 1955.
[12] W. Freiberger and R. Jones, "Computation of the frequency function of a quadratic form in random normal variables," *Journal of the ACM*, vol. 7, no. 3, pp. 245–250, July 1960.
[13] H. Robbins and E. J. G. Pitman, "Application of the method of mixtures to quadratic forms in normal variates," *The annals of mathematical statistics*, vol. 20, no. 4, pp. 552–560, December 1949.
[14] J. Imhof, "Computing the distribution of quadratic forms in normal variables," *Biometrika*, vol. 48, no. 3, pp. 419–426, 1961.
[15] G. G. Tziritas, "On the distribution of positive-definite gaussian quadratic forms," *IEEE Trans. Information Theory*, vol. 33, no. 6, pp. 895–906, November 1987.
[16] T. Cover and J. Thomas, *Elements of information theory*. Wiley and Sons, New York, 1991.

Julien Huillery was born in France in 1979. He received the M.S. degree in electrical engineering from the École Supérieure d'Électricité (France) and the University of Wollongong (Australia) in 2002, and the master of research degree in signal processing from the university of Rennes (France) in 2004.
Since 2005, he is working toward the Ph.D. degree in signal processing with the Gipsa-lab, Grenoble, France.
His research interests concern time-frequency analysis, statistical signal processing and information theory.

Fabien Millioz graduated from the École Normale Supérieure de Cachan, France, and received the MS degree in Signal and Image Processing from Institut National Polytechnique de Grenoble, France, in 2005.
He is currently working toward the Ph.D. degree at the Gipsa-lab, Grenoble, France.
His research interests are in signal processing, more precisely time-frequency segmentation.

Nadine Martin received the Engineering degree in Physics and Electronics in 1980 and the Ph.D. degree in Signal Processing and Control in 1984. Currently, she is a senior researcher at the CNRS - National Centre of Scientific Research, and in charge of the research team SA-IGA - Signal and automatic for diagnostic and surveillance, a team within the Signal and Image Department, at GIPSA lab - Grenoble Image Speech Signal and Automatic, Grenoble, France. In the signal-processing domain, her research interests are the analysis and the interpretation of nonstationary signals. She is now working on signal detection, models of polynomial phase signals, and on time-frequency decision based on time-frequency random models. She is leading the TetrAS project on an automatic spectral analyzer. In addition of acoustic signals, vibratory signals are more particularly studied in relation with physical models. In 2004 and 2005, she was a scientific advisor for an automotive industry.
Dr. N. Martin is the author of more than 100 papers, of a French patent and of a book on "time-frequency decision". She was member of the National Council of Universities (CNU) in 2002. She was co-organizer of the Fourth European Signal Processing Conference (EUSIPCO'88), of a pre-doctoral course on recent advances in signal processing (Les Houches 93), of the Sixth French Symposium on Signal and Image Processing (GRETSI'97). She organized special sessions on signal processing for diagnosis (IEEE- SDEMPED'97, CM'07 and ACD'07).