Log-Rayleigh Distribution: A Simple and Efficient Statistical Representation of Log-Spectral Coefficients

Bertrand Rivet, Laurent Girin, *Member, IEEE*, and Christian Jutten, *Member, IEEE*

Abstract—In this paper, we study the distribution of the log-modulus of a Gaussian complex random variable. In the circular case, it is a Log-Rayleigh (LR) variable, whose probability distribution function (pdf) depends on only one parameter. In the noncircular case, the pdf is more complicated, although we show that it can be adequately modeled by an LR pdf, for which the optimal fitting parameter is derived. These results can be used in any application using the log-modulus of discrete Fourier transform coefficients, e.g., for speech/audio signals, and suggest that a mixture of LR pdf kernels is preferable to more classical models such as mixtures of Gaussian kernels, which are more costly and less efficient.

Index Terms—Discrete Fourier transform (DFT), Gaussian complex random variable, Rayleigh distribution, speech processing, statistical modeling.

I. INTRODUCTION

MANY SIGNAL processing applications (e.g., compression [1] and speech enhancement [2]) require statistical knowledge of the signal descriptors. Since the distribution of the considered coefficients is often intricate, the use of general models is a widely used solution. In particular, mixtures of (many) Gaussian kernels are very common models in many speech processing applications (e.g., voice conversion [3], speech enhancement [4], and speaker identification [5]). Such a general model provides a good fit for the data, but at the expense of a large number of parameters and a high computational cost.

In speech processing, as in many applications, the spectral domain is often considered, rather than the temporal domain. In particular, discrete Fourier transform (DFT) coefficients are widely used, since they provide a sparse and efficient representation of the signal. The real and imaginary parts of DFT coefficients of many frames of long-term continuous speech (which is *nonstationary* because it includes silences and all kind of speech sounds) are characterized by a "peaky" distribution. Thus, a *global* model of the signal distribution, such as a Laplacian model [6], can be used. However, this model does not take into account the fact that the long-term speech signal is a sequence of various sounds. Alternately, one may prefer to exploit the structure of speech to build a mixture of kernels in which each speech sound is expected to be modeled by one kernel. This would provide a more efficient physically based model if one is interested in retrieving "local" characterizations of spectral information from the complete mixture [3], [7].

In many sound processing applications, it is specifically the *logarithm of the modulus* of DFT coefficients that is often used, since the log-spectrum is related to perceptual properties of the human ear [8], and the log scale is computationally appropriate for the scaling of the absolute values of the coefficients. In this paper, we thus aim to find an efficient multisound model that exploits the structure of speech (as mentioned above) and whose kernels are appropriate to the log-modulus DFT coefficients. To do this, we first consider the basic case of quasi-stationary signals (i.e., one given speech sound) and we study the statistical properties of log-modulus DFT coefficients for this case. We then propose an efficient multisound model, derived from the single-sound case just mentioned.

More precisely, for quasi-stationary sections of speech, the DFT coefficients can be assumed to have a zero-mean Gaussian complex circular distribution [9]. Their *log-modulus* is therefore no longer Gaussian. In Section II, we thus address the theoretical issue of characterizing the logarithm of the modulus of a zero-mean circular Gaussian random variable by what will be referred to here as the Log-Rayleigh law. In Section III, we study the theoretical consequences for the pdf if circularity is not assumed, as is sometimes the case with nonstationary sections of real speech signals. Then, in Section IV, we exploit the results of Sections II and III in modeling the long-term speech log-modulus DFT coefficients using a very efficient model: a mixture of Log-Rayleigh kernels. This Log-Rayleigh mixture model (LRMM) is a weighted sum of Log-Rayleigh laws, in the same way that a Gaussian mixture model (GMM) is a sum of weighted Gaussian laws. Experiments prove that this approach provides very accurate modeling with a very low computational cost compared to GMMs, since 1) the Log-Rayleigh pdf depends on only one parameter, and 2) a single Log-Rayleigh pdf is able to model the DFT log-modulus of one kind of sound whereas several Gaussian kernels would be needed for the same task (since the log-modulus of DFT coefficients for each sound is no longer Gaussian).

II. COMPLEX RANDOM VARIABLES

In this section, we address basic theoretical issues with the characterization of complex variables that will be used in the following to model log-modulus DFT coefficients. For the sake
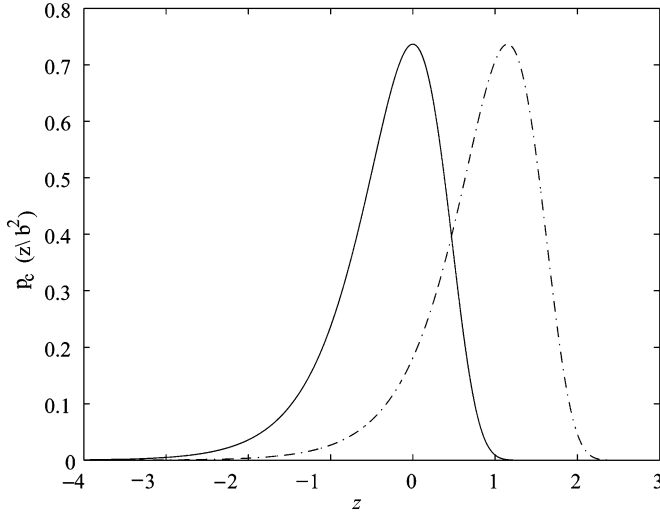
Fig. 1. Probability density functions of a circular Log-Rayleigh distribution of localization parameter $\beta^2 = 1$ (solid line) and $\beta^2 = 10$ (dashed-dotted line).

of simplicity, we will always consider the scalar (i.e., monodimensional) case, as opposed to the vector (i.e., multidimensional) case often developed in the literature [9]–[11]. In the following, we thus will statistically characterize the different bins of the log-modulus DFT coefficients separately, since it is a classic result ([12, ch. 12-4]) that quasi-stationary signals provide uncorrelated DFT coefficients. In addition, since elementary speech sounds are considered to be zero-mean signals, the corresponding DFT coefficients are also zero-mean, and we will therefore study zero-mean complex variables.

Let X denote a zero-mean circular (some authors prefer the term *proper* [10]) complex Gaussian random variable [9] with variance σ^2: $X \sim \mathcal{N}_{\mathbb{C}}(0, \sigma^2)$. The pdf of X is given by [9], [10]

$$p_X(x) = \left[\pi\sigma^2\right]^{-1} \exp\left\{-\frac{|x|^2}{\sigma^2}\right\} \quad (1)$$

where $|\cdot|$ denotes the modulus. This equation assumes that the real and imaginary parts of X are uncorrelated and of equal variance.

It is well known [12] that the modulus $Y = |X|$ of X is Rayleigh distributed with parameter $\sigma^2/2$: $Y \sim \mathrm{Ray}(\sigma^2/2)$. The pdf of a Rayleigh distribution of parameter β^2 is

$$p_Y(y) = \begin{cases} \frac{y}{\beta^2} \exp\left(-\frac{y^2}{2\beta^2}\right), & \text{if } y \geq 0 \\ 0, & \text{if } y < 0. \end{cases} \quad (2)$$

The natural logarithm $Z = \ln Y$ of Y follows what we call (see Def. 1) a Log-Rayleigh (LR) distribution of localization parameter $\sigma^2/2$: $Z \sim \mathrm{LogRay}(\sigma^2/2)$.

Definition 1: Let Z denote a circular Log-Rayleigh random variable of localization parameter β^2: $Z \sim \mathrm{LogRay}(\beta^2)$. The pdf of this variable is given by

$$\forall z \in \mathbb{R}, \quad p_Z(z) = p_c(z|\beta^2) = \frac{(e^z)^2}{\beta^2} \exp\left(-\frac{(e^z)^2}{2\beta^2}\right) \quad (3)$$

where $p_c(\cdot|\cdot)$ refers to the circular LR pdf. This distribution is plotted in Fig. 1.

The circular Log-Rayleigh distribution is characterized by a typical property: all the centered moments higher than one are independent of localization parameter β^2. Indeed, if we let Z_1 and Z_2 be two circular Log-Rayleigh random variables of localization parameters β_1^2 and β_2^2 respectively, then $p_{Z_1}(\cdot)$ and $p_{Z_2}(\cdot)$ verify

$$p_{Z_2}(z + a) = p_{Z_1}(z), \text{ with } a = \log\frac{\beta_2}{\beta_1}. \quad (4)$$

This means that each distribution is deduced from the other with a translation (see Fig. 1).

Moreover,
- the mean m_Z (the only moment which depends on the localization parameter) is equal to

$$m_Z = \ln\beta + \frac{\ln 2}{2} - \frac{\gamma}{2} \quad (5)$$

where γ is the Euler constant defined by

$$\gamma = -\int_0^{+\infty} \ln x \, \exp(-x)dx \quad (6)$$

- using the Euler–Mascheroni formula

$$\gamma^2 + \frac{\pi^2}{6} = \int_0^{+\infty} (\ln x)^2 \exp(-x)dx \quad (7)$$

we derive the variance v_Z

$$v_Z = \frac{\pi^2}{24}. \quad (8)$$

Note that, since X is a zero-mean circular complex Gaussian random variable, the square of its module, Y^2, is related to a chi-square distribution with two degrees of freedom: $\chi^2(2)$ [12], which is a special case of the gamma distribution: $\Gamma(1, 1/2)$ [12]. Thus, $\ln(Y^2)$, the natural logarithm of Y^2, is related to a log-gamma law [13]. Finally, the law of Z, the log-modulus of X, can also be deduced from the log-gamma distribution by a change of variable.

III. CONSEQUENCES OF NONCIRCULARITY

Until now, we have considered the case of a circular complex Gaussian random variable. However, in some cases, DFT coefficients can be noncircular, as illustrated in Fig. 3, where we can see that the real and imaginary parts of some of the plotted DFT coefficients are correlated. This may be due to several factors, notably the fact that the signal is locally quasi-stationary and not strictly stationary. Therefore, in this section, we study the consequences of the noncircularity of a zero-mean Gaussian complex random variable on the pdf of its log-modulus. We then estimate the closest circular LR law to the noncircular LR law according to the Kullback–Leibler divergence. To the best of our knowledge, this specific point has never been reported in the literature. This point will be useful in validating the general model proposed in Section II, for the case in which the DFT coefficients are not strictly circular, but close to circular.

The second order moments of a zero-mean complex random variable X are the covariance $v_X = E\{xx^*\}$ (* denotes the complex conjugate) and the pseudocovariance $c_X = E\{x^2\}$ (see [9], [10] where those moments are defined for the more
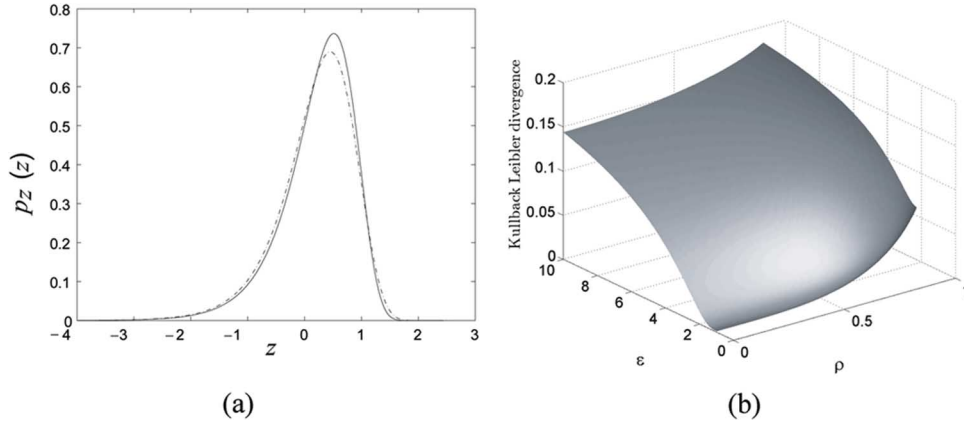
Fig. 2. Consequences of noncircularity. (a) $\rho = -0.4$ and $\epsilon = 0.8$, the dashed line represents the noncircular LR pdf (12) and the solid line the optimal circular Log-Rayleigh pdf of the localization parameter given by (17). (b) Kullback–Leibler divergence between the noncircular LR pdf and the optimal circular LR one as a function of ρ and ϵ.

general multi-dimensional case). Thus, in the scalar case, we have

$$v_X = \sigma^2_{\Re(X)} + \sigma^2_{\Im(X)} \qquad (9)$$

and

$$c_X = \sigma^2_{\Re(X)} - \sigma^2_{\Im(X)} + 2\jmath\rho\sigma_{\Re(X)}\sigma_{\Im(X)} \qquad (10)$$

where $\jmath = \sqrt{-1}$, ρ is the correlation coefficient between the real and imaginary part of X, and $\sigma^2_{\Re(X)}$ and $\sigma^2_{\Im(X)}$ are the variances of the real and imaginary parts of X, respectively.

In the circular case, the pseudocovariance is null ($c_X = 0$), which means that both $\sigma_{\Re(X)} = \sigma_{\Im(X)}$ and $\rho = 0$. Thus, the noncircularity can be due to different covariances for real and imaginary parts and/or to some correlation between real and imaginary parts. For simplification, we denote $\delta^2 = v_X$ (9), and we introduce ϵ such that

$$\sigma_{\Re(X)} = \epsilon\,\sigma_{\Im(X)}. \qquad (11)$$

In this case, we can show (see Appendix I) that the pdf of the noncircular LR random variable Z is equal to

$$p_Z(z) = p_c\left(z \left| \frac{\delta^2}{2}\right.\right) I(z, \delta^2, \rho, \epsilon) \qquad (12)$$

where $p_c(z|\beta^2)$ is the pdf of a circular Log-Rayleigh random variable of localization parameter β^2 given by (3), and

$$I(z, \delta^2, \rho, \epsilon) = \frac{\epsilon + \frac{1}{\epsilon}}{2\sqrt{1-\rho^2}} \exp\left\{-\frac{4\rho^2 + (\epsilon - \frac{1}{\epsilon})^2}{4(1-\rho^2)\delta^2}(e^z)^2\right\}$$
$$\times I_0\left(\left(\epsilon + \frac{1}{\epsilon}\right)\frac{\sqrt{(\epsilon - \frac{1}{\epsilon})^2 + 4\rho^2}}{4(1-\rho^2)\delta^2}(e^z)^2\right) \qquad (13)$$

where $I_0(x)$ is the modified Bessel function of the first kind: $I_0(x) = 1/(2\pi)\int_{-\pi}^{\pi}\exp\{x\sin\theta\}d\theta$. Note that in this expression, the pdf of a noncircular LR random variable is expressed as the product of a circular LR distribution whose localization parameter depends only on the covariance δ^2 and a correction term $I(\cdot, \cdot, \cdot, \cdot)$, which depends on the covariance δ^2 and the pseudo-covariance c_X.

Although much more complicated, this noncircular distribution is very close to a circular Log-Rayleigh distribution if ρ and $|\epsilon - 1|$ are small compared to 1 (see Fig. 2). Therefore, for this case, we again propose to model this noncircular LR pdf by a circular LR distribution $p_c(z|\alpha)$ that depends on a localization parameter α: $\mathrm{LogRay}(\alpha)$. This new localization parameter is estimated so that the two pdfs are as close as possible according to the following criterion:

$$\hat{\alpha} = \arg\min_{\alpha} g(\alpha) \qquad (14)$$

where $g(\alpha)$ is the Kullback–Leibler divergence[1]

$$g(\alpha) = \int_{-\infty}^{+\infty} p_Z(z) \ln\left(\frac{p_Z(z)}{p_c(z|\alpha)}\right) dz. \qquad (15)$$

In the following, the circular Log-Rayleigh distribution of optimal localization parameter $\hat{\alpha}$ is referred to as the optimal circular Log-Rayleigh distribution.

It is easy to show that

$$\hat{\alpha} = \frac{1}{2}\int_{-\infty}^{+\infty}(e^z)^2\,p_Z(z)\,dz \qquad (16)$$

which is equal to (see Appendix II)

$$\hat{\alpha} = \frac{\delta^2}{2}. \qquad (17)$$

Note that since δ^2 is the sum of the variances of the real and imaginary parts of X, the optimal localization parameter $\hat{\alpha}$ depends only on ϵ [see (9) and (11)]. Thus, this localization parameter $\hat{\alpha}$ defines a new circular LR random variable derived from the noncircular LR random variable by dropping the correction term. This implies that the new circular LR random variable can be considered as the log-modulus of a circular Gaussian random variable derived from a noncircular Gaussian random variable by cancelling the pseudocovariance and keeping the covariance. Schreier and Scharf obtain a similar result from derivations conducted directly on the Gaussian complex random variable in the vectorial case [11]. However, we derived our result on the log-modulus of the Gaussian random variable and we did so indirectly.

---

[1]Other criteria (e.g., MSE) could be used, but they lead to a more complicated derivation of optimal the localization parameter (without an analytic formula).

IV. APPLICATION TO SPEECH PROCESSING

In this section, we apply the theoretical derivations to a real speech corpus. We first present some results on the modeling of log-modulus DFT coefficients of a single isolated speech sound by a single LR kernel, and we then present some results obtained by the modeling of the log-modulus DFT coefficients of the entire corpus by a mixture of LR kernels.

A. Data

The corpus consists of nonsense utterances of the structure V1-C-V2-C-V1, where V1 and V2 are vowels from the set $\{[a],[i],[y],[u]\}$ and C is a plosive consonant from the set $\{[p],[t],[k],[b],[d],[g]\}$ or no plosive [#]. The 112 sequences were pronounced twice by a male native French speaker: the first series was used for training the model coefficients and the second series for the test. The signals were sampled at 16 kHz and the DFT coefficients were calculated with a Hamming analysis window of 320 lags (20 ms) with 0.75 overlap. About 48 s of speech were used, representing a total of about 9600 spectral vectors $\ln |\mathbf{S}|$, where $\mathbf{S} = [S(f_1), \ldots, S(f_N)]^T$ is the vector of the short-term Fourier transform coefficients over N bins. This corpus was used because it allowed us to efficiently test both the modeling of a single isolated sound and that of a controled number of sounds.

B. Modeling of a Single Speech Sound

First, all frames identified as a section of the vowel [a] (Fig. 3) were extracted from the database. About 4 s were used (representing a total of about 800 spectral vectors) to calculate the optimal localization parameter (17) [Fig. 3(a)] as well as ϵ and ρ, for each frequency bin [Fig. 3(b)]. The localization parameter can be interpreted as the spectral envelope of the sound. We can see in Fig. 3(c) that the estimated Log-Rayleigh distribution (solid line) closely follows the empirical distribution (estimated by a histogram) of the vowel coefficients. In contrast, at least two Gaussian kernels would be needed to correctly model the same data because of the pdf asymmetry [Fig. 3(c)]. Additionally, accurately modeling the LR pdf itself with Gaussian mixtures would require more Gaussian kernels (typically four kernels in our experiments).

C. Modeling of Continuous Speech

In this subsection, we characterize the overall corpus by a model that is expected to efficiently exploit the structure of speech: a mixture of optimal Log-Rayleigh kernels (LRMM). In this case, each kernel is expected to model a single given speech sound among all possible speech sounds. Moreover, each kernel k is multidimensional with $N \times N$ localization matrix Γ_k: each kernel gathers all N frequency bins. Since the length of the analysis window is 20 ms, the speech signal is assumed to be quasi-stationary. The DFT coefficients are thus assumed to be uncorrelated (and thus independent) at different frequency bins. This leads us to choose a diagonal localization matrix Γ_k for each kernel, with each diagonal entry
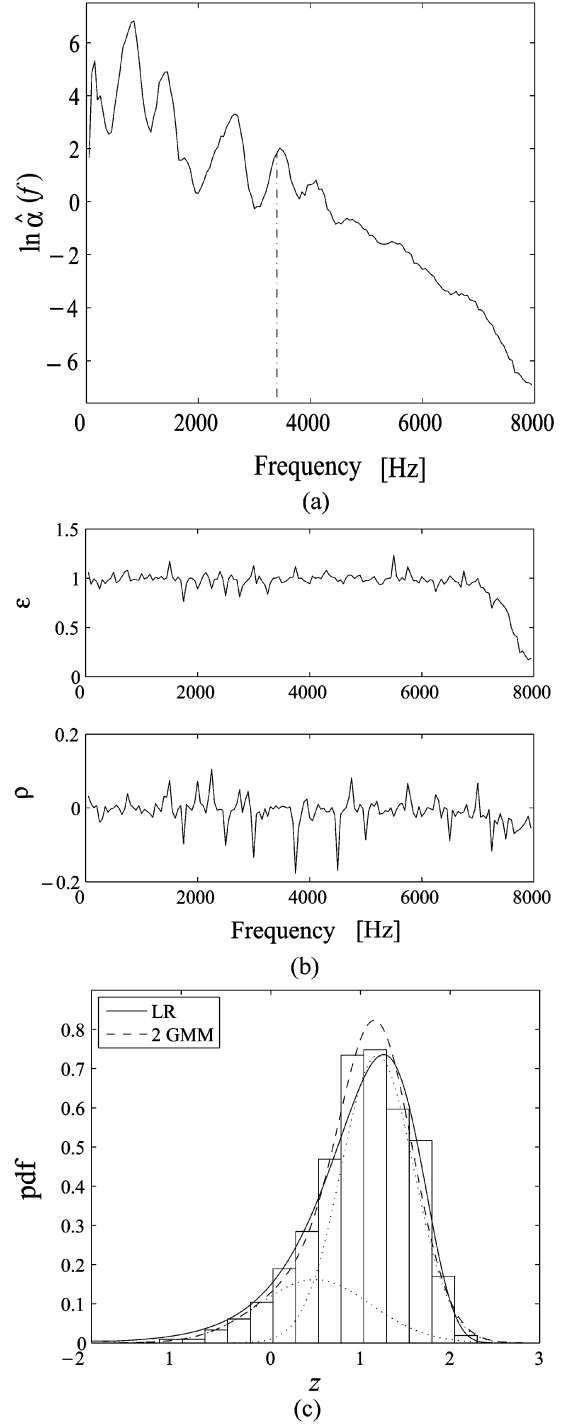

Fig. 3. Vowel [a]. (a) Logarithm of the optimal localization parameters versus the frequency, the dashed-dotted line corresponds to frequency bin 68. (b) Parameters ϵ and ρ versus frequency. (c) Histogram at frequency bin 68, the optimal Log-Rayleigh distribution (solid line), the approximation by two Gaussian kernels (dashed line) and the two corresponding Gaussian kernels (dotted line).

corresponding to a given frequency bin. Thus, the multikernel model is given by

$$\ln |\mathbf{S}| \sim \sum_{k=1}^{K} \omega_k \, \mathrm{LogRay}(\Gamma_k) \qquad (18)$$

where ω_k and Γ_k are the weight and the pdf localization parameter of the kth Log-Rayleigh kernel, respectively.[2] The set of parameters $\{\omega_k, \Gamma_k\}_k$ is estimated using the EM algorithm [14], where the initialization is done by vector quantization [15].

We compare the adequacy between the data and the model using the χ^2 test [12]: at each frequency bin f, we compute

$$\zeta(f) = N \sum_{i=1}^{m} \frac{(\hat{p}_i(f) - p_i(f))^2}{p_i(f)} \qquad (19)$$

where N is the number of data points, m the number of classes (i.e., bins) of the histogram used to calculate the empirical distribution $\hat{p}_i(f)$, and $p_i(f)$ the empirical and theoretical probabilities of the ith class, respectively. We then count the number of frequency bins for which (19) is smaller than a 5% confidence threshold defined by $\chi^2_{0.95}(m-1)$. To demonstrate the advantage of the LRMM over the GMM, we compare the results obtained with the two models (Fig. 4). We see [Fig. 4(a) and (b)] that when the number of kernels increases, the adequacy between the models and the data increases faster for the LRMM case than for the GMM case. Indeed, (19) decreases faster toward zero for the LRMM than for the GMM, as the number of kernel increases.

Fig. 4(c) shows that, for the same number of kernels K, the LRMM provides better adequacy than the GMM, except when K is less than 16. Indeed, since the corpus contains about ten phonemes, a minimum of ten kernels[3] is necessary for the LRMM. Alternately, for the same modeling accuracy, the LRMM requires a significantly smaller number of kernels than the GMM. For instance, the results obtained with 64 Gaussian kernels are the same as the results obtained with 16 Log-Rayleigh kernels. As mentioned above, this is because several Gaussian kernels are needed to correctly model the data set modeled by a single Log-Rayleigh kernel: on average, a ratio of the order of four in the multikernel case is coherent with the observations in the single kernel case (see the above-mentioned example of vowel [a]). Finally, since a Gaussian kernel requires three parameters (weight, mean vector, covariance matrix), and an LR kernel requires only two parameters (weight, localization matrix), the LRMM provides a much sparser model than the GMM for a given fitting performance.

V. CONCLUSION

In this paper, we have characterized the distribution of the logarithm of the modulus of a zero-mean Gaussian complex random variable in both circular and noncircular cases. Although the pdf is complicated in the noncircular case, it is close to—and then can be adequately modeled by—the simpler pdf obtained in the circular case, namely the Log-Rayleigh pdf.

[2]For the sake of simplicity, we note LogRay(Γ) the Log-Rayleigh distribution, with diagonal localization matrix $\Gamma = \mathrm{diag}(\gamma_1, \ldots, \gamma_N)$, which is equal to the product of the monodimensional Log-Rayleigh distributions: $\mathrm{LogRay}(\Gamma) = \prod_j \mathrm{LogRay}(\gamma_j)$.

[3]Actually, more than ten kernels are necessary since the corpus also contains transitions between phonemes. These transitions may provide additional speech sounds to be modeled, which explains that about 30 kernels are necessary to achieve a score of 90% in χ^2-test.
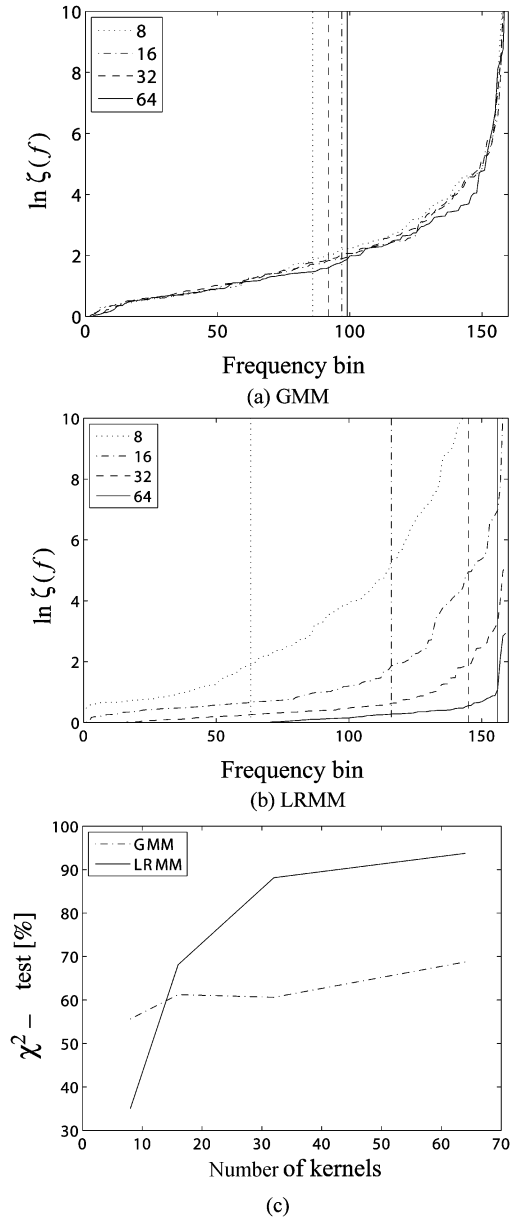

Fig. 4. Application to real speech signals. (a) [resp. (b)] shows the logarithm of (19) (sorted by ascending order) at each frequency bin for GMM (resp. LRMM). The vertical lines show the number of components which verify the χ^2 test. Several number of kernels were tested: 8 (dotted line), 16 (dashed-dotted line), 32 (dashed line), and 64 (solid line). (c) presents the percentage of frequency bins which satisfy the χ^2 test versus the number of kernels for both the GMM (dashed line) and the LRMM (solid line).

This latter pdf implies a very efficient representation of the DFT coefficients log-modulus since it provides an optimal or a quasi-optimal model with only one parameter, avoiding the use of the more complicated mixture of Gaussian kernels to model these coefficients. Experiments on speech signals provide a good illustration of these theoretical results. To conclude, such a characterization of pdf can be useful for many problems requiring a statistical model of log-modulus spectral coefficients, e.g., speech/audio spectral perceptual coding, enhancement, classification, and frequency domain source separation.

APPENDIX I
NONCIRCULAR LOG-RAYLEIGH DISTRIBUTION

In this Appendix, we derive the distribution of the log-modulus of a noncircular zero-mean complex random variable.

Let A and B denote the real and imaginary parts of X, respectively. We have

$$p_{A,B}(a,b) = \frac{\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{a^2}{\sigma_a^2} - \frac{2\rho ab}{\sigma_a \sigma_b} + \frac{b^2}{\sigma_b^2}\right]\right\}}{\left(2\pi \sigma_a \sigma_b \sqrt{1-\rho^2}\right)}. \quad (20)$$

Let R and Θ denote the modulus and the phase of X, respectively. We have $a = r\cos\theta$ and $b = r\sin\theta$, thus

$$p_{R,\Theta}(r,\theta) = r p_{A,B}(r\cos\theta, r\sin\theta)$$

leading to

$$p_R(r) = \int_{\theta=-\pi}^{\theta=\pi} p_{R,\Theta}(r,\theta) d\theta \quad (21)$$

so finally, using (20) and (21)

$$p_Z(z) = p_c\left(z \left| \frac{\delta^2}{2}\right.\right) I(z,\delta^2,\rho,\epsilon)$$

with $I(z,\delta^2,\rho,\epsilon)$ defined by (13).

APPENDIX II
OPTIMAL LOCALIZATION PARAMETER

In this Appendix, we detail the derivation of the optimal parameter $\hat{\alpha}$ defined by (14).

From the minimization of the Kullback–Leibler divergence $g(\alpha)$, where

$$g(\alpha) = \int_{-\infty}^{+\infty} p_Z(z) \ln\left(\frac{p_Z(z)}{p_c(z|\alpha)}\right) dz$$

it is easy to show that the optimal localization parameter $\hat{\alpha}$ is given by

$$\hat{\alpha} = \frac{1}{2}\int_{-\infty}^{+\infty} (e^z)^2\, p_Z(z)\, dz.$$

Substituting $p_Z(z)$ by (12), one can show that

$$\hat{\alpha} = \frac{\delta^2}{2} J(\rho,\epsilon)$$

with

$$J(\rho,\epsilon) = \left(\epsilon + \frac{1}{\epsilon}\right)(1-\rho^2)^{3/2} \frac{4}{\pi} \int_{-\pi}^{\pi} \frac{d\theta}{(a-b\sin\theta)^2} \quad (22)$$

where

$$a = 4 + \left(\epsilon - \frac{1}{\epsilon}\right)^2 = \left(\epsilon + \frac{1}{\epsilon}\right)^2 \quad (23)$$

and

$$b = \left(\epsilon + \frac{1}{\epsilon}\right)\sqrt{\left(\epsilon - \frac{1}{\epsilon}\right)^2 + 4\rho^2}. \quad (24)$$

Using [16, formula (14.361)], one can show that

$$\int_{-\pi}^{\pi} \frac{d\theta}{(a-b\sin\theta)^2} = \frac{a}{(a^2-b^2)} \int_{-\pi}^{\pi} \frac{d\theta}{a-b\,\sin\theta} \quad (25)$$

and using [16, formula (14.360)]

$$\int_{-\pi}^{\pi} \frac{d\theta}{a-b\,\sin\theta} = \frac{2\pi}{\sqrt{a^2-b^2}}. \quad (26)$$

Thus, substituting (25) and (26) in (22), we obtain

$$J(\epsilon,\rho) = \left(\epsilon + \frac{1}{\epsilon}\right)(1-\rho^2)^{3/2} \frac{8a}{(a^2-b^2)^{3/2}}. \quad (27)$$

By substituting (23) and (24) in (27), one can see that $J(\epsilon,\rho)$ is equal to 1.

REFERENCES

[1] A. Gersho and R. Gray, *Vector Quantization and Signal Compression*. Amsterdam, The Netherlands: Kluwer, 1992.

[2] Y. Ephraim and D. Malah, "Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 32, no. 6, pp. 1109–1121, Dec. 1984.

[3] Y. Stylianou, O. Cappé, and E. Moulines, "Continuous probabilistic transform for voice conversion," *IEEE Trans. Speech Audio Process.*, vol. 6, no. 2, pp. 131–142, Mar. 1998.

[4] D. Burshtein and S. Gannot, "Speech enhancement using a mixture-maximum model," *IEEE Trans. Speech Audio Process.*, vol. 10, no. 6, pp. 341–351, Sep. 2002.

[5] D. A. Reynolds and R. C. Rose, "Robust text-independent speaker identification using Gaussian mixture speaker models," *IEEE Trans. Speech Audio Process.*, vol. 3, no. 1, pp. 72–83, Jan. 1995.

[6] R. Martin and C. Breithaupt, "Speech enhancement in the DFT domain using Laplacian speech priors," in *Proc. IWAENC*, Kyoto, Japan, Sep. 2003, pp. 87–90.

[7] L. Benaroya and F. Bimbot, "Wiener based source separation with HMM/GMM using a single sensor," in *Proc. Int. Conf. Ind. Compon. Anal. Blind Source Separation (ICA)*, Nara, Japan, 2003, pp. 957–961.

[8] E. Zwicker and U. Zwicker, *Psychoacoustics, Facts and Models*. Berlin, Germany: Springer-Verlag, 1990.

[9] B. Picinbono, "Second-order complex random vectors and normal distributions," *IEEE Trans. Signal Process.*, vol. 44, no. 10, pp. 2637–2640, Oct. 1996.

[10] F. D. Neeser and J. L. Massey, "Proper complex random processes with applications to information theory," *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1293–1302, Jul. 1993.

[11] P. Schreier and L. Scharf, "Second-order analysis of improper complex random vectors and processes," *IEEE Trans. Signal Process.*, vol. 51, no. 3, pp. 714–725, Mar. 2003.

[12] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 3rd ed. New York: McGraw-Hill, 1991.

[13] J. F. Lawless, "Inference in the generalized gamma and log-gamma distribution," *Technometrics*, vol. 22, pp. 67–82, 1980.

[14] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum-likelihood from incomplete data via the EM algorithm," *J. R. Statist. Soc. Ser. B*, vol. 39, pp. 1–38, 1977.

[15] Y. Linde, A. Buzo, and R. M. Gray, "An algorithm for vector quantizer design," *IEEE Trans. Commun.*, vol. COM-28, no. 1, pp. 84–95, Jan. 1980.

[16] M. R. Spiegel, *Schaum's Mathematical Handbook of Formulas and Tables*. New York: McGraw-Hill, 2001.

Bertrand Rivet graduated from the École Normale Supérieure de Cachan, Cachan, France, and received the Agrégation de Physique Appliquée and the Master's degree from the University of Paris-XI, Orsay, France, in 2002 and 2003, respectively. He is currently pursuing the Ph.D. degree in signal processing at the Institut de la Communication Parlée (Speech Communication Laboratory) and at the Laboratoire des Images et des Signaux (Images and Signals Laboratory), Grenoble, France.

His research concerns audiovisual speech and blind source separation.

Laurent Girin received the M.Sc. and Ph.D. degrees in signal processing from the Institut National Polytechnique de Grenoble, Grenoble, France, in 1994 and 1997, respectively.

In 1997, he joined the Ecole Nationale d'Electronique et de Radioélectricité de Grenoble, where he is currently an Associate Professor in electrical engineering and signal processing. His research activity takes place in the Institut de la Communication Parlée (Speech Communication Laboratory) in Grenoble. His current research interests concern audiovisual speech processing with application to speech coding, speech enhancement, and audio/speech source separation, and also acoustic speech analysis, modeling and synthesis.

Christian Jutten received the Ph.D. and the Docteur ès Sciences degrees from the Institut National Polytechnique of Grenoble, Grenoble, France, in 1981 and 1987, respectively.

He was an Associate Professor with the Ecole Nationale Supérieure d'Electronique et de Radioélectricité, Grenoble, from 1982 to 1989. He was a Visiting Professor with the Swiss Federal Polytechnic Institute, Lausanne, Switzerland, in 1989, before becoming a Full Professor with the Université Joseph Fourier, Grenoble, more precisely in Polytech'Grenoble Institute. He is currently an Associate Director of the Images and Signals Laboratory (100 people). For 25 years, his research interests have been in blind source separation, independent component analysis, and learning in neural networks, including theoretical aspects (separability, source separation in nonlinear mixtures), applications in signal processing (biomedical, seismic, speech), and data analysis. He is author or coauthor of more than 40 papers in international journals, 16 invited paper, and 100 communications in international conferences.

Prof. Jutten was an Associate Editor of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS (1994–1995) and co-organizer the First International Conference on Blind Signal Separation and Independent Component Analysis (Aussois, France, January 1999). He is a reviewer of leading international journals (including IEEE TRANSACTIONS ON SIGNAL PROCESSING, IEEE SIGNAL PROCESSING LETTERS, IEEE TRANSACTIONS ON NEURAL NETWORKS, *Signal Processing*, *Neural Computation*, *Neurocomputing*) and conferences in signal processing and neural networks (including ICASSP, ISCASS, EUSIPCO, IJCNN, ICA, ESANN, IWANN).