ON SPECTROGRAM LOCAL MAXIMA

Patrick Flandrin

Univ Lyon, Ens de Lyon, Univ Claude Bernard, CNRS, Laboratoire de Physique, F-69342 Lyon, France
patrick.flandrin@ens-lyon.fr

ABSTRACT

In close connection with time-frequency uncertainty relations, spectrograms are known to have some built-in redundancy which constrains the landscape of their surface, thus calling for simplified descriptions based on a reduced number of salient features. This is investigated here in some detail for spectrogram local maxima in the generic case of white Gaussian noise. A simple model, based on a randomized hexagonal lattice structure, is proposed for the distribution of their time-frequency locations considered as realizations of a 2D point process in the plane. While the rationale of the model is discussed and its relevance tested, further consideration is also given to the distribution of maximal values as well as to that of zeros that can be inferred from the proposed model.

Keywords — time-frequency analysis; reassignment; spectrogram; white noise; Voronoi diagram

1. MOTIVATION

The time-frequency (TF) analysis of signals can be performed in many different ways, but spectrograms—i.e., squared magnitudes of Short-Time Fourier Transforms (STFTs)—are certainly among the simplest and most popular tools for this purpose. Whereas much is known about such transforms (see, e.g., [4, 10, 16]), specific properties of some of their geometrical attributes—in particular the distribution of their extrema, be they local maxima or zeros—are still to be better understood for a number of different reasons that can be listed as follows:

1. It is first known that, in close connection with TF uncertainty relations, spectrograms have some built-in redundancy [10]. This is expected to constrain the landscape of the surface, thus calling for simplified descriptions based on a reduced number of salient features that might offer a data-driven, adaptive way of getting a sparse representation of signals from a TF perspective.

2. More specifically, a suitable choice for the short-time analysis window—namely that of a properly scaled Gaussian function [2, 20]—happens to turn spectrograms into entire functions whose zeros can be used for TF filtering, as it has been recently proposed [11].

3. Besides zeros, local maxima also play an important role, especially with respect to improved versions of spectrograms based on reassignment (see, e.g., [13] for a survey), or on the closely related synchrosqueezing technique [6]. In fact, it has been shown in [3] that, under the above Gaussian assumption for the short-time window, the reassignment field of a spectrogram simply expresses as the gradient of a potential function which identifies to the logarithm of the spectrogram surface, with the consequence that reassignment tends to concentrate the signal energy around local maxima.

Extrema (zeros and local maxima) constitute therefore a very simplified description of a spectrogram that is yet expected to capture some significant information. The purpose of this paper is to have a closer look at the geometry which specifically governs the distribution of local maxima in the TF plane.

2. STFT AND SPECTROGRAM

2.1. Definitions

Given a signal $x(t)$ and a window $h(t)$, both supposed to belong to $L^2(\mathbb{R})$, the *Short-Time Fourier Transform* (STFT) $F_x^{(h)}(t,\omega)$ will be here defined as the inner product between $x(t)$ and shifted versions (in time and frequency) of $h(t)$:

$$F_x^{(h)}(t,\omega) = \langle x, \mathbf{T}_{t\omega} h \rangle, \quad (1)$$

where $\mathbf{T}_{t\omega}$ stands for the joint TF shift operator such that $(\mathbf{T}_{t\omega}h)(s) = h(s-t)\exp\{i\omega(s-t/2)\}$ [3]. The corresponding *spectrogram* simply follows as:

$$S_x^{(h)}(t,\omega) = \left| F_x^{(h)}(t,\omega) \right|^2. \quad (2)$$

2.2. Reproducing kernel, redundancy and uncertainty

While the STFT maps a 1D signal to a 2D function, any 2D function cannot be guaranteed to be the admissible transform of some signal. Indeed, STFTs—as well as the associated spectrograms—inherit some specific structure from their definition (1). More precisely, the STFT satisfies the identity:

$$F_x^{(h)}(t',\omega') = \iint_{-\infty}^{+\infty} K(t',\omega';t,\omega)\, F_x^{(h)}(t,\omega)\, dt\, \frac{d\omega}{2\pi}, \quad (3)$$

in which the *reproducing kernel* $K(t',\omega';t,\omega)$ is (up to a complex-valued multiplicative term) nothing but the STFT of the analyzing window:

$$K(t',\omega';t,\omega) \propto F_h^{(h)}(t'-t,\omega'-\omega). \tag{4}$$

It follows that any STFT (or spectrogram) has necessarily some local redundancy since the quantity (4) cannot be arbitrarily peaked in both time and frequency, a result which follows from general uncertainty relations (see, e.g., [22, 16]). In the particular case of the (unit energy) Gaussian window[1]

$$g(t) = \pi^{-1/4}\exp\{-t^2/2\}, \tag{5}$$

the so-called "Gabor's logon" [15] which is known to minimize classical measures of uncertainty [15] and is referred to as *circular* since

$$K(t',\omega';t,\omega) \propto \exp\left\{-\frac{1}{4}((t'-t)^2+(\omega'-\omega)^2)\right\}, \tag{6}$$

the reproducing kernel is maximally concentrated and defines an influence domain which is itself circular.

3. SPECTROGRAM OF WHITE GAUSSIAN NOISE

3.1. Covariance structure

Following first investigations reported in [14], we will concentrate here on the specific case of white Gaussian noise (wGn). Making use of noise as test signal is a convenient way of accessing configurations with no prescribed structure. It also allows to investigate the self-organizing properties of "generic" spectrogram surfaces whose geometry reflects uncertainty constraints.

Applying formally the definition (2) to zero-mean wGn $x(t)$, we readily get that:

$$\mathbb{E}\left\{S_x^{(h)}(t,\omega)\right\} = \gamma_0, \tag{7}$$

whereas the covariance between spectrogram values at two different locations in the TF plane only depends on the corresponding lags in both time and frequency, according to the relation:

$$\mathrm{cov}\left\{S_x^{(h)}(t,\omega), S_x^{(h)}(t',\omega')\right\} = \gamma_0^2\, S_h^{(h)}(t'-t,\omega-\omega'). \tag{8}$$

In the specific case of the circular Gaussian window (5), this covariance takes on the simple form

$$\mathrm{cov}\left\{S_x^{(g)}(t,\omega), S_x^{(g)}(t',\omega')\right\} = \gamma_0^2\,\exp\left\{-\frac{1}{2}d^2(t,\omega;t',\omega')\right\} \tag{9}$$

where $d(t,\omega;t',\omega') = \sqrt{(t-t')^2+(\omega-\omega')^2}$ measures the Euclidean distance in the plane between the two considered points, thus defining a second-order homogeneous (or *stationary*) field.

[1] In the Physics literature (see, e.g., [20]), the corresponding spectrogram is referred to as the "Husimi distribution function" [18].

4. LOCAL MAXIMA AS A 2D POINT PROCESS

4.1. Heuristics

It follows from the elements above that, in the generic case of wGn, the spectrogram of a realization should take on the form of a random, yet homogeneous distribution of energy "patches" whose extension, while fluctuating, is constrained by the reproducing kernel of the analysis. The locations of local maxima define therefore a collection of points in the TF plane that can be considered as a realization of some stochastic 2D point process. Given the covariance structure (8), such a process is expected to exhibit a *repulsive* structure and its characterization could therefore be envisioned within the framework of the so-called "determinantal point processes" [21, 23, 26]. We will however follow here a different approach, whose main purpose is to propose a *constructive modeling* that will eventually end up with the reported behavior attached to the covariance structure.

As discussed in Sect. 2.2, Gabor's logons (5) are elementary, minimum uncertainty, waveforms. They can therefore be viewed as *building blocks* whose suitably shifted versions, in both time and frequency, may add up to synthesize a more complicated waveform according to the tentative model

$$x(t) = \sum_m\sum_n x_{mn}\, g(t-t_m)\, e^{i(\omega_n t+\varphi_{mn})}, \tag{10}$$

where x_{mn} is some random weight, $\mathbf{x}_{mn} = \{(t_m,\omega_n); m,n \in \mathbb{Z}\}$ stands for the TF center of a logon $g(t)$ in the plane, and φ_{mn} for a possible phase term. It follows from this expansion that the associated spectrogram reads

$$S_x^{(g)}(t,\omega) = \left|\sum_m\sum_n x_{mn}\, F_g^{(g)}(\omega-\omega_n,t-t_m)\, e^{i\varphi_{mn}}\right|^2. \tag{11}$$

In the specific case of wGn, the TF spectrum (7) is expected to be flat. Together with the homogeneity evidenced by (9), this suggests some form of TF *equidistribution* for the logons that appear in (11), i.e., an identical mean value for all weights ($\mathbb{E}\{x_{mn}\} = $ cst for any m and $n \in \mathbb{Z}$) and a regular lattice structure for the centers $\mathbf{x}_{mn}$. To this end, it is proposed to choose a *regular triangular lattice* as a model, with a rationale that can be qualitatively justified for at least two complementary reasons:

1. As sketched in Figure 1-(a), given a logon centered at some TF location, the circular structure of the covariance (8) suggests that (on average) a neighbouring, independent logon is likely to show up anywhere on the limiting circle from which the covariance can be considered as negligible. Picking up at random such a point, the same argument applies to the next neighbours. Intersecting the two circles (see Figure 1-(b)) leads to logon TF centers located on the vertices of

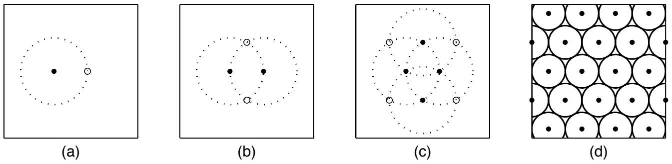

Fig. 1. *Mean model for the TF locations of logon centers in the case of wGn—Graphical heuristics (see text).*

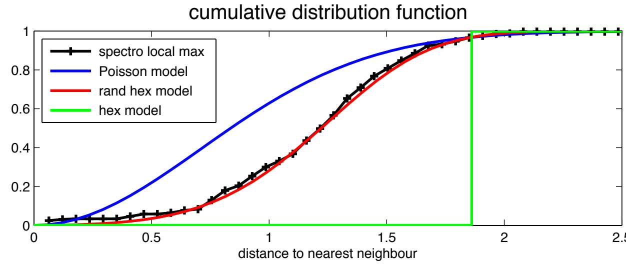

Fig. 2. *Cumulative distribution of the TF distance to nearest neighbour between spectrogram local maxima in the case of wGn—Comparison between actual data and reference models (see text).*

equilateral triangles and, proceeding further the same way, the average distribution of logon centers is expected to form a regular triangular lattice (see Figure 1-(c)). As a companion argument, one can also remark that, since individual logons concentrate most of their energy in TF domains taking the form of circular disks, organizing them on a regular triangular lattice corresponds to maximum packing [5] (see Figure 1-(d)).

2. A by-product of such an organization over a regular triangular lattice is that the associated *Voronoi* tessellation [25] should take the form of a *honeycomb* structure made of hexagons. This is amply supported by numerical experiments, as already mentioned in [14]: a typical result, obtained with $1,000$ independent realizations of wGn realizations of $2,048$ data points each, evidences that the average number of edges of the Voronoi cells centered on local maxima is 5.98, i.e., almost 6.

4.2. A randomized lattice model

Whereas the proposed model for the locations of the logon centers takes on the form of a regular triangular lattice that is deterministic, both the random nature of the phase references and the partial overlap between neighbouring logon STFTs introduce fluctuations whose outcome is that local extrema do not coincide with logon centers. The simplest model we can propose for taking into account such fluctuations is that of a perturbation of the regular lattice, with random time and frequency shifts that will be assumed to be i.i.d. Gaussian.

Building upon an approach outlined in [27], the relevance of this model can be tested by evaluating the *distribution of nearest neighbour distances*, as is commonly done for testing a Poisson distribution [7]. More precisely, given a fixed point arbitrarily chosen at the origin of the TF plane, the first step requires the evaluation of the probability of finding one of its 6 nearest neighbours within a given distance. To do so, we can first evaluate the probability that *one* of its neighbours lies at a distance at least d from the origin. This simply reads

$$P_1(d) = \mathrm{Prob}\{D > d\} = 1 - \iint_{\Omega} p(t,\omega)\, dt\, d\omega, \qquad (12)$$

where Ω stands for the disk of radius d centered at the origin of the plane and

$$p(t,\omega) = \frac{1}{2\pi\sigma^2} \exp\left\{-\frac{1}{2\sigma^2}\left[(t-m)^2+\omega^2\right]\right\} \qquad (13)$$

if the Gaussian fluctuations—around the chosen node of coordinates $(m, 0)$—are assumed to be of variance σ^2 in each direction. A change of variables to polar coordinates leads to:

$$P_1(d) = 1 - \int_0^d F(r;m,\sigma^2)dr, \qquad (14)$$

with

$$F(r;m,\sigma^2) = \frac{r}{\sigma^2}\exp\left\{-\frac{1}{2\sigma^2}\left(r^2+m^2\right)\right\} I_0\left(\frac{rm}{\sigma^2}\right) \qquad (15)$$

and

$$I_0(x) = \frac{1}{2\pi}\int_0^{2\pi} \exp\{x\cos\theta\}d\theta \qquad (16)$$

the modified Bessel function of first kind [1].

If we further assume that the variance is small enough to ensure that the nearest neighbour only comes from fluctuations around the vertices of the first hexagon around the point of reference, the total probability $P_6(d)$ that no neighbour can be found at a distance of a least d from any given point of the perturbed lattice can be approximated by $P_6(d) = (P_1(d))^6$, thus leading to the final result:

$$\mathrm{Prob}(D \leq d) = 1 - \left(1 - \int_0^d F(r;m,\sigma^2)dr\right)^6. \qquad (17)$$

A simulation example is given in Figure 2, which compares an actual distribution with its best fitted model according to (17), the lattice fluctuation turning out to be such that $\sigma \approx m/3.43$. For a sake of comparison, the Figure also reports the step distribution that would occur for a deterministic lattice (i.e., when $\sigma = 0$), as well as the one which would have resulted from a Poisson distribution with the same density (as expected from the independence assumption that cannot be valid for spectrogram maxima, the Poisson model is

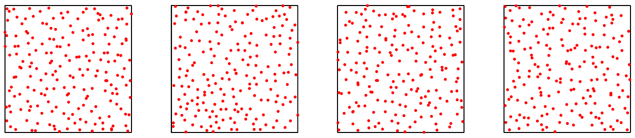

Fig. 3. *Locations of local spectrogram extrema in the case of wGn—One of the 4 diagrams corresponds to actual data whereas the 3 other ones result from the proposed model: who is who is left to the reader's appreciation (solution in [29]).*

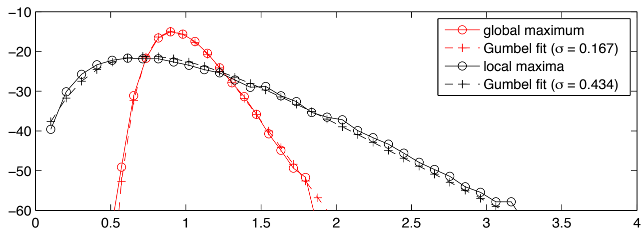

Fig. 4. *Probability distribution of (mean normalized) global and local spectrogram extrema in the case of wGn—Comparison between actual data and Gumbel fits (logarithmic scale, dB units).*

clearly ruled out). A visual illustration of the relevance of the approach is given in Figure 3, in which some actual data is proposed for an eyeballing comparison with 3 realizations of the proposed model, as synthetized in a similar configuration.

4.3. More on spectrogram maxima

Up to now, we have only been concerned by the location of spectrogram local maxima in the TF plane, and not by the values that are taken on at those locations. In order to characterize such values that are of importance too for a complete description, one can make use of known general results about the distribution of the periodogram ordinates of wGn [8, 17, 19] to claim (and easily check) that spectrogram values have an exponential distribution. The fact that spectrogram values are identically distributed according to an exponential law ends up with a Gumbel distribution of the form [9]:

$$p(x; \mu, \sigma) = \frac{1}{\sigma} \exp\{z - e^z\}; z = -(x-\mu)/\sigma, \qquad (18)$$

for the fluctuations of the global maximum as well as for local maxima, though with a larger (relative) variance in the latter case. This is illustrated in Figure 4 on synthetic data resulting from 5,000 realizations of wGn (each with 2,048 time samples), with all sequences of local maxima normalized to unity in the mean (i.e., such that $\mu = 1$). Plugging this result in the logon-based model (10)-(11) leads to synthetic spectrogram models that compare very favorably with actual data, as evidenced in Figure 5.

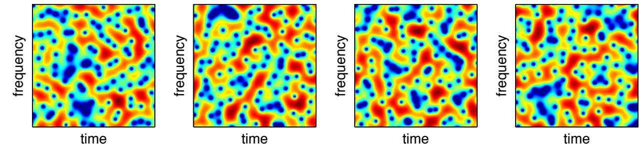

Fig. 5. *Spectrogram vs. synthetic models in the case of wGn—One of the 4 diagrams corresponds to actual data whereas the 3 other ones result from the proposed model: who is who is left to the reader's appreciation (solution in [29]).*

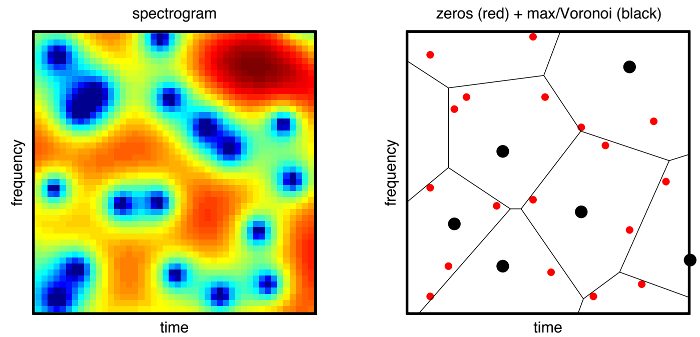

Fig. 6. *Left: spectrogram of wGn. Right: local maxima (black dots) and zeros (red dots), the latter ones turning out to be mostly located on the edges of the Voronoi cells associated to the former ones (black lines).*

5. CONCLUDING REMARKS

In this study, a number of results have been reported concerning the simplified description of spectrograms of wGn in geometrical terms, in particular with respect to the distribution of local extrema for which a simple, yet effective model has been proposed. Emphasis has been put here on maxima which are of special importance in reassignment methods (as attractors and fixed points of the reassigment vector field [24]), but zeros are known to be of importance too [11]. While zeros clearly deserve further investigations that cannot be reported here, it is worth noticing— as a brief remark—that geometrical features of the distribution of zeros can be seen as a by-product of the randomized lattice model proposed here. A first illustration is given in Figure 6, in which zeros are shown to be mostly located on the edges of the Voronoi cells attached to local maxima. This, along with other features related to spectrogram zeros that complete the present study, will be detailed further elsewhere [12].

Acknowledgements — This work is part of the ASTRES Project, supported by Agence Nationale de la Recherche under grant ANR-13-BS03-0002-01. Participants of this Project, as well as Pierre Chainais (CRIStAL, Lille), are gratefully acknowledged for discussions and comments.

6. REFERENCES

[1] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions*, Dover, 1965.

[2] V. Bargmann, "On a Hilbert space of analytic functions and an associated integral transform," *Commun. Pure Appl. Math.*, Vol. 14, pp. 187–214, 1961.

[3] E. Chassande-Mottin, I. Daubechies, F. Auger, and P. Flandrin, "Differential reassignment," *IEEE Signal Proc. Lett.*, Vol. 4, No. 10, pp. 293–294, 1997.

[4] L. Cohen, *Time-Frequency Analysis*, Prentice Hall,1995.

[5] J. Conway and N.J.A. Sloane, *Sphere Packings, Lattices and Groups* (3rd ed.), Springer, 1999.

[6] I. Daubechies, J. Lu, and H.-T. Wu, "Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool," *Appl. and Comp. Harm. Anal.*, Vol. 30, No. 1, pp. 243–261, 2011.

[7] P.J. Diggle, *Statistical Analysis of Spatial Point Patterns* (2nd ed.), Academic Press, 2003.

[8] T.S. Durrani and J.M. Nightingale, "Probability distributions for discrete Fourier spectra," *Proc. IRE*, Vol. 120, No. 2, pp. 299–311, 1973.

[9] P. Embrechts, C. Klüppelberg, and T. Mikosch, *Modeling Extremal Events for Insurance and Finance*, Springer, 1997.

[10] P. Flandrin, *Time-Frequency/Time-Scale Analysis*, Academic Press,1999.

[11] P. Flandrin, "Time-frequency filtering from spectrogram zeros," *IEEE Signal Proc. Lett.*, Vol. 22, No. 11, pp. 2137–2141, 2015.

[12] P. Flandrin, "The sound of silence: recovering signals from time-frequency zeros," in *Proc. 50th Asilomar Conf. on Signals, Systems, and Computers*, Pacific Grove (CA), 2016.

[13] P. Flandrin, F. Auger, and E. Chassande-Mottin, "Time-frequency reassignment — From principles to algorithms," in *Applications in Time-Frequency Signal Processing* (A. Papandreou-Suppappola, Ed.), Chapter 5, pp. 179–203, CRC Press, Boca Raton (FL), 2003.

[14] P. Flandrin, E. Chassande-Mottin, and F. Auger, "Uncertainty and spectrogram geometry," in *Proc. 20th European Signal Processing Conf. EUSIPCO-12*, pp. 794–798, Bucharest (RO), 2012.

[15] D. Gabor, "Theory of communication," *J. IEE*, Vol. 93, pp. 429–457, 1946.

[16] K. Gröchenig, *Foundations of Time-Frequency Analysis*, Birkhäuser, Boston (MA), 2011.

[17] J. Huillery, F. Millioz, and N. Martin, "On the description of spectrogram probabilities with a chi-squared law," *IEEE Trans. on Sig. Proc.*, Vol. 56, No. 6, pp. 2249–2258, 2008.

[18] K. Husimi, "Some formal properties of the density matrix," *Proc. Phys. Math. Soc. Jpn*, Vol. 22, pp 264–314, 1940.

[19] P.E. Johnson and D.G. Long, "The probability density of spectral estimates based on modified periodogram averages," *IEEE Trans. on Sig. Proc.*, Vol. 47, No. 5, pp. 1255–1261, 1999.

[20] H.J. Korsch, C. Müller, and H. Wiescher, "On the zeros of the Husimi distribution," *J. Phys. A: Math. Gen.*, Vol. 30, pp. L677–L684, 1997.

[21] F. Lavancier, J. Møller, and E. Rubak, "Determinantal point process models and statistical inference," *J. Roy. Stat. Soc. B*, Vol. 77, No. 4, pp. 853–877, 2015.

[22] E.H. Lieb, "Integral bounds for radar ambiguity functions and Wigner distributions," *J. Math. Phys.*, Vol. 31, pp. 594–599, 1990.

[23] O. Macchi, "The coincidence approach to stochastic point processes," *Adv. in Appl. Prob.*, Vol 7, No. 2, pp. 83–122, 1975.

[24] S. Meignen, T. Oberlin, Ph. Depalle, P. Flandrin, and S. McLaughlin, "Adaptive multimode signal reconstruction from time-frequency representations," *Proc. Phil. Trans. Roy. Soc. A*, Vol. 374: 20150205, 2016.

[25] A. Okabe, B. Boots, K. Sugihara, and S.N. Chiu, *Spatial Tessellations — Concepts and Applications of Voronoi Diagrams* (2nd ed.), John Wiley, 2000.

[26] A. Soshnikov, "Determinantal random point fields," *Russian Math. Surveys*, Vol. 55, No. 5, pp. 923–975, 2000.

[27] L. Stirling Churchman, H. Flyvberg, and J.A. Spudich, "A non-Gaussian distribution quantifies distances measured with fluorescence localization techniques," *Biophysical Journal*, Vol. 90, pp. 668–671, 2006.

[28] D. Stoyan, W.S. Kendall, and J. Mecke, *Stochastic Geometry and its Applications*, John Wiley, 1995.

[29] *Actual data corresponds to the second sub-diagram (from the left) in Figure 3, and to the fourth one (from the left too) in Figure 5.*

— ◊ —