Estimation of a white Gaussian noise in the Short Time Fourier Transform based on the spectral kurtosis of the minimal statistics: Application to underwater noise
— Source link ↗

Fabien Millioz, Nadine Martin

Related papers:

- Application to locally optimum detection of a new noise model

- Further results on the minimum variance time-frequency distribution kernels

- Robust estimation of non-stationary noise power spectrum for speech enhancement

- HOS-based noise models for signal detection optimization in non-Gaussian environments

- Estimation of Gaussian, Poissonian–Gaussian, and Processed Visual Noise and Its Level Function

Estimation of a white Gaussian noise in the Short Time Fourier Transform based on the spectral kurtosis of the minimal statistics: application to underwater noise

Fabien Millioz, Nadine Martin

ESTIMATION OF A WHITE GAUSSIAN NOISE IN THE SHORT TIME FOURIER TRANSFORM BASED ON THE SPECTRAL KURTOSIS OF THE MINIMAL STATISTICS: APPLICATION TO UNDERWATER NOISE

Fabien Millioz and Nadine Martin

Gipsa-lab DIS

961 rue de la Houille Blanche - Domaine universitaire - BP 46

38402 Saint Martin d'Hres cedex - France

email: 1stname.2ndname@gipsa-lab.grenoble-inp.fr

ABSTRACT

In this paper we present a noise level estimator using minimal values of the Short Time Fourier Transform of a signal embedded in a white Gaussian noise. The spectral kurtosis of the smallest values is used to estimate the variance of the noise without any *a priori* knowledge on the signal. This estimation is illustrated on both a synthetic and speech signal. A dolphin whistle detection in underwater noise is given as an application.

Index Terms— Short Time Fourier Transform, Noise Level Estimation, Spectral Kurtosis, Minimal Statistics

1. INTRODUCTION

The signal detection in the time-frequency plane can be defined by a hypothesis test: the test discriminates a point containing signal from a point containing noise only. Usually, this test leads to a threshold on the signal's energy. In a blind detection, the properties and the location in time and frequency of the signal are unknown. Consequently, the detection is based only on the noise properties. The problem is thus to estimate these properties to determinate the detector.

Assuming that the analyzed signal is a signal to detect embedded in a white Gaussian noise, the variance of the noise is to be estimated, without *a priori* knowledge on locations containing noise only. Two main techniques exist to estimate this variance.

First is based on an iterative principle [1, 2, 3]. The noise level is estimated with all time-frequency coefficients, including those containing signal. The noise variance is overestimated, but permits to detect a first set of points containing signal. The next iterations estimate the noise variance on points where no signal has been detected, leading to less overestimated noise level. Consequently more points containing signal may be detected. The iterations are stopped when one criterion on the non-detected points is satisfied: the convergence of their statistical characteristics [1]; a threshold on their skewness [2]; or a threshold on their kurtosis [3]. The drawback of theses methods is to be quite slow, depending on the quality of the stop criterion.

Second is based on the minimal statistics: considering that the points of lower energy contains noise only, the noise level is estimated with these points only. Rainer Martin [4, 5] uses the smallest value after a recursive smoothing. Ewans and Mason [6] uses the qth quantile in the context of speech recognition. Huillery [7] in his thesis (in French) uses of the Z smallest values. All theses methods need either *ad hoc* choices or the estimation of the number of points containing noise only.

This paper proposes a new general method for the noise estimation problem using the minimal statistics. The estimation is based on the linearity of the Short Time Fourier Transform (STFT), whose squared modulus is the spectrogram. The STFT of a white Gaussian noise is a complex Gaussian noise. When using a detector based on energy, a threshold on energy is equivalent to a threshold on the absolute value of the STFT. By rejecting points greater than the threshold, the STFT becomes a truncated complex Gaussian variable. In a first section, considering the circularity of the STFT, we study the truncated circular complex Gaussian distribution, whose variance and spectral kurtosis are given. This distribution is used in the case of the STFT to determinate an estimation of the noise level in a second section. The estimation is applied to a synthetic signal and to a speech signal embedded in a white Gaussian noise. The method is then illustrated with the detection a dolphin whistle in underwater noise.

A noise estimation based on the kurtosis of the truncated real and imaginary part of the STFT has already been published [8] (in French).

2. TRUNCATED CIRCULAR COMPLEX GAUSSIAN DISTRIBUTION

When a complex variable $Z = A + jB$ has a centered Gaussian circular distribution of variance σ^2, its real and imaginary
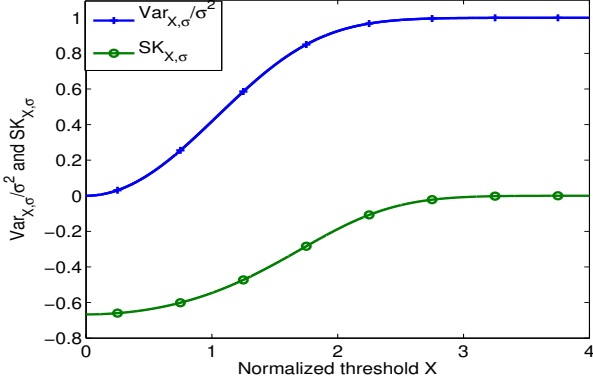
Fig. 1. Evolution of the normalized variance and the spectral kurtosis in function of the normalized threshold $X_n = \frac{X}{\sigma}$.

parts have both a Gaussian distribution of variance $\frac{\sigma^2}{2}$ and are independent. Its probability density function is:

$$f_Z(a,b) \quad = \quad \frac{1}{\pi\sigma^2} e^{-\frac{1}{\sigma^2}(a^2+b^2)} \tag{1}$$

Substituting polar variables (ρ, θ) for Cartesian variables (a, b), equation (1) becomes:

$$f_Z(\rho, \theta) \quad = \quad \frac{\rho}{\pi\sigma^2} e^{-\frac{\rho^2}{\sigma^2}} \tag{2}$$

Due to the circularity of z, $f_Z(\rho, \theta)$ is independent from θ:

$$f_Z(\rho) \quad = \quad \int_{-\pi}^{\pi} f_Z(\rho, \theta) d\theta = \frac{2\rho}{\sigma^2} e^{-\frac{\rho^2}{\sigma^2}} \tag{3}$$

A truncated circular Gaussian variable T_X is the random variable stemming from a circular Gaussian variable Z where realisations greater than a threshold X are rejected. Its probability density function is:

$$f_{T_X}(\rho) \quad = \quad \begin{cases} \frac{2}{1-e^{-X^2/\sigma^2}} \frac{\rho}{\sigma^2} e^{-\frac{\rho^2}{\sigma^2}} & \text{for } 0 \leq \rho \leq X, \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

Its variance and spectral kurtosis derive from the probability density function. Variance and kurtosis of real truncated Gaussian variables are detailed in [9]. In this paper we are concerned with the complex case only. The variance of T_x, written $Var_{X,\sigma}$, is:

$$\begin{aligned} Var_{X,\sigma} \quad &= \quad \mathrm{E}(TT^*) = \mathrm{E}(\rho^2) = \int_{-\infty}^{+\infty} \rho^2 f_{T_X}(\rho) d\rho \\ &= \quad \sigma^2 \left(1 - \frac{1}{e^{(X^2/\sigma^2)} - 1} \frac{X^2}{\sigma^2} \right) \end{aligned} \tag{5}$$

Its spectral kurtosis $\mathcal{SK}_{X,\sigma}$ is [10]:

$$\begin{aligned} \mathcal{SK}_{X,\sigma} \quad &= \quad \frac{E(T^2 T^{*2})}{\mathrm{E}(TT^*)^2} - 2 = \frac{E(\rho^4)}{E(\rho^2)^2} - 2 \\ &= \quad \frac{\left(e^{X^2/\sigma^2} - 1\right)\left(2\frac{X^2}{\sigma^2} - \frac{X^4}{\sigma^4}\right) - 2\frac{X^4}{\sigma^4}}{\left(e^{X^2/\sigma^2} - 1 - \frac{X^2}{\sigma^2}\right)^2} \end{aligned} \tag{6}$$

From equations (5,6), we notice that the normalized variance $Var_{X,\sigma}/\sigma^2$ and the spectral kurtosis $\mathcal{SK}_{X,\sigma}$ are not directly dependent on σ^2 but only on the normalized threshold X_n, defined by:

$$X_n \quad = \quad \frac{X}{\sigma} \tag{7}$$

Figure 1 shows the evolution of the normalized variance and the spectral kurtosis in function of the normalized threshold X_n. These two functions are both monotonic increasing functions. When the threshold increases, the truncated Gaussian variables tends to be a full Gaussian variable, its variance becomes σ^2 while its spectral kurtosis becomes 0.

3. NOISE LEVEL ESTIMATION FROM THE STFT OF A SIGNAL EMBEDDED IN AN ADDITIVE NOISE

3.1. Short Time Fourier Transform

The STFT of a discrete signal $x[n]$ is:

$$S_\phi[n,k] = \sum_{m=n-(M_\phi-1)/2}^{n+(M_\phi-1)/2} x[m]\phi[m-n]e^{-2j\pi k\frac{m}{K}} \tag{8}$$

with $\phi[n]$ is a normalized window of M_ϕ points and K the number of frequency bins. Continuous Fourier Transforms (FT) are circular [11], but the discrete ones are not at low and high frequencies [3]. In this paper, we will not consider frequency bin $k = 0$, where imaginary part of the STFT is null. All other time-frequency coefficients are approximated as circular variables.

Consequently, time-frequency points containing noise only have a complex circular Gaussian distribution.

3.2. Spectral kurtosis of the smallest values of the STFT of a signal embedded in an additive noise

In this section, we consider the signal $x[n] = s[n] + g[n]$, sum of an unknown signal $s[n]$ and a white Gaussian noise $g[n]$ of variance σ^2. The noise level is estimated using the minimal values of the STFT under the hypothesis that the smallest values of the STFT of $x[n]$ have the same distribution than the smallest values of the STFT of $g[n]$ only.

To illustrate this hypothesis, two signals $s[n]$ are used: the first one is synthetic, made of three chirps; the second is a speech signal. These signals are embedded in a white Gaussian noise $g[n]$ of known variance. Figure 2 shows the spectral kurtosis of the STFT of the signal and noise in relation to the threshold X and the spectral kurtosis of the truncated noise only. The STFT are computed with a Hanning window of 127 points, an overlap of 63 points and 256 frequency bins. This figure shows that the spectral kurtosis of small values is not altered by the presence of signal $s[n]$.
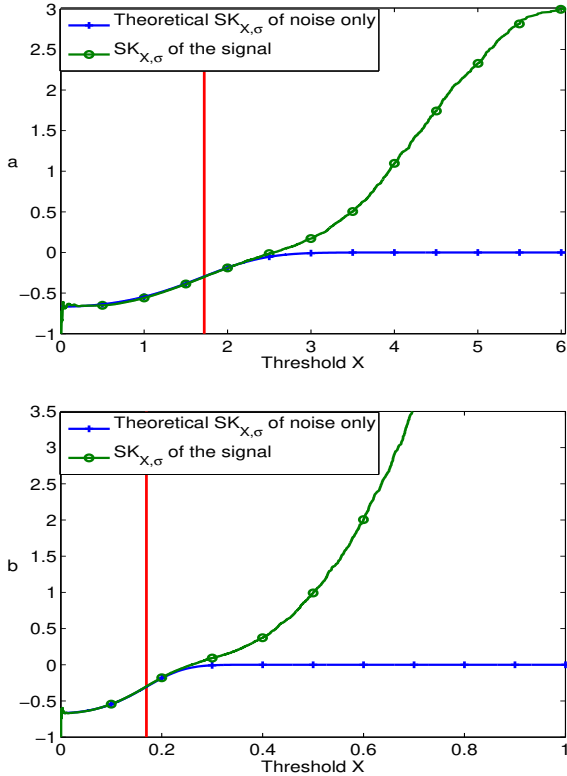
Fig. 2. In blue, spectral kurtosis of a truncated circular Gaussian variable of variance (a) 1 and (b) 10^{-2}. In green, spectral kurtoses of truncated sums of a white Gaussian noise of same variances and (a) a synthetic signal made of three linear chirps; (b) a speech signal. In any cases, spectral kurtoses of the truncated signals have the behavior for small X_n than spectral kurtoses of truncated Gaussian noise only. Red lines shows X such as $\mathcal{SK}_{X,\sigma} = -0.3$.

3.3. Noise level estimation

Spectral kurtosis is a normalized cumulant, in other words it does not depend on the variance of the random variable. In our context, spectral kurtosis of a truncated STFT depends on the normalized threshold $X_n = \frac{X}{\sigma}$ only, and is a monotonic increasing function. Consequently, there is only a threshold $X_n(\kappa)$ such as:

$$X_n(\kappa) \quad / \quad \mathcal{SK}_{X_n(\kappa),1} = \kappa \tag{9}$$

$X_n(\kappa)$ is computed by inversion of (6). Due to the complexity of the equation, theses values are numerically computed. Table 1 gives a few $X_n(\kappa)$.

Considering a STFT, it is possible to determine a non-normalized threshold $X(\kappa)$ such as the spectral kurtosis of the points smaller than the threshold is equal to a value κ:

$$X(\kappa) \quad / \quad \mathcal{SK}_{X(\kappa),\sigma} = \kappa \tag{10}$$

$\mathcal{SK}_{X,\sigma}$	-0.6	-0.5	-0.4	-0.3	-0.2	-0.1
X_n	0.755	1.164	1.455	1.709	1.965	2.279

Table 1. Rounded values of the normalized thresholds $X_n = \frac{X}{\sigma}$ corresponding to different values of the spectral kurtosis of a truncated circular Gaussian random variable.

Red lines in figure 2 shows the values $X(-0.3)$ in two cases.

From (7), the noise level estimator $\hat{\sigma^2}(\kappa)$ is:

$$\hat{\sigma^2}(\kappa) = \left(\frac{X(\kappa)}{X_n(\kappa)} \right)^2 \tag{11}$$

Using the two examples of figure 2, measured thresholds $X(-0.3)$ are equal from top to bottom to 1.7205 and 0.1705. Table 1 gives $X_n(-0.3) = 1.709$. From equation (11), estimated noise level $\hat{\sigma^2}(-0.3)$ are respectively 1.013 and $0.099.10^{-2}$, while the additive noises had variances of 1 and 10^{-2} respectively.

This estimator depends on two parameters: the choice of the estimator of the spectral kurtosis, which is not discussed in this paper, and the choice of κ, discussed in the following section.

3.4. Choice of κ

To determine the influence of the parameter κ, we define a normalized bias $b(\kappa)$ such as:

$$b(\kappa) = \frac{E\left(\hat{\sigma}^2(\kappa)\right) - \sigma^2}{\sigma^2} \tag{12}$$

For both the synthetic and speech signals, the estimation bias is estimated for 10000 realizations of the additive white Gaussian noise, at different noise levels and for the κ values displayed on table 1. Results are given on figure 3. Except for $\kappa = -0.6$ and $\kappa = -0.1$, this parameter has a little influence on the bias.

4. APPLICATION TO UNDERWATER NOISE

The relevance of such a noise estimation is illustrated on a signal detection in an underwater noise. The first image of figure 4 shows a spectrogram of a dolphin whistle, limited to the normalized bandwidth $[0.2, 0.34]$ such as noise is white. Given the estimated noise level from (11) and a probability of false alarm p_{fa}, a Neyman-Pearson criterion gives a detection threshold [3].

The second image of figure 4 shows the detection with a probability of false alarm of 10^{-3}. The modulation of the dolphin whistle are so detected. The only parameters needed to detect the signal are the user-chosen probability of false alarm, and the κ parameter of the noise level estimation, which has little influence on the estimation.
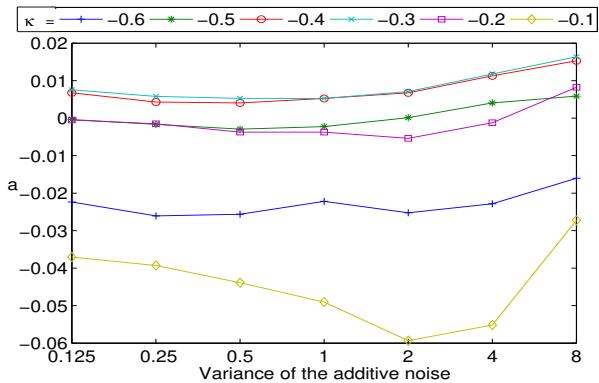
Fig. 3. Normalized biases (12) of the noise level estimator $\hat{\sigma^2}(\kappa)$ according to the noise level σ^2 and κ, for the synthetic signal. Except for high and low values, κ does not affect strongly the bias.

5. CONCLUSION

A new noise level estimator has been presented, based on the minimal statistics of the STFT. It is based on the spectral kurtosis of a truncated circular Gaussian random variable, consequently its properties depends on the spectral kurtosis estimation. This estimator has only one parameter to be chosen *a priori*, κ, which has little influence on the bias for medium values.

Future work will take in consideration the non-circularity of discrete STFT and the influence of the spectral kurtosis estimator on the noise level estimation. Variance of the noise level estimator should be studied. Finally, an extension to non-white or non-stationary Gaussian noise will follow, by estimating the noise level locally over the time-frequency plane.

6. REFERENCES

[1] C. Hory and N. Martin, "Maximum likelihood noise estimation for spectrogram segmentation control," in *Proceedings of IEEE Conference on Acoustics, Speech and Signal Processing*, Orlando, USA, May 2002, pp. 1581–1584.

[2] Chunghsin Yeh and Axel Röbel, "Adaptive noise level estimation," in *Proc. of the Int. Conf. on Digital Audio Effects (DAFx-06)*, Montreal, Quebec, Canada, Sept. 18–20, 2006, pp. 145–148.

[3] F. Millioz, J. Huillery, and N. Martin, "Short Time Fourier Transform Probability Distribution for Time-Frequency Segmentation," in *Proceedings of IEEE Conference on Acoustics, Speech and Signal Processing*, Toulouse, France, May 2006.

[4] Rainer Martin, "Noise power spectral density estimation based on optimal smoothingand minimum statistics," *Speech and Audio Processing, IEEE Trans. on*, vol. 9, no. 5, pp. 504–512, 2001.
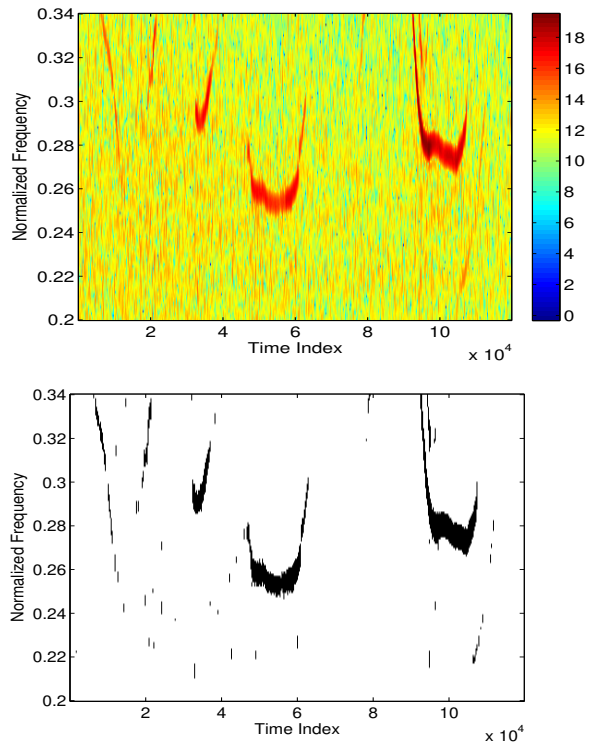
Fig. 4. Whistle of a dolphin. The spectrogram (top) is limited to the normalized frequency $[0.2, 0.34]$ to have a white Gaussian noise on the coefficients. Bottom, detection of signal with a probability of false alarm of 10^{-3}.

[5] Rainer Martin, "Bias compensation methods for minimum statistics noise power spectral density estimation," *Signal Process.*, vol. 86, no. 6, pp. 1215–1229, 2006.

[6] Nicholas W.D. Evans and John S. Mason, "Time-Frequency Quantile-Based Noise Estimation," in *Proceedings of EU-SIPCO*, Toulouse, France, 2002.

[7] Julien Huillery, *Support temps-fréquence d'un signal inconnu en présence de bruit additif gaussien*, Thèse de doctorat, INP Grenoble, July 2008.

[8] F. Millioz and N. Martin, "Estimation de la densité spectrale de puissance d'un bruit gaussien basée sur le kurtosis des statistiques minimales," in *22e colloque GRETSI sur le traitement du signal et des images*, Dijon, France, Sept. 2009.

[9] N.L. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions*, vol. 2, Wiley and sons, 2nd edition, 1994.

[10] V. Vrabie, P. Granjon, and C. Servière, "Spectral kurtosis: from definition to application," in *IEEE-EURASIP International Workshop on Nonlinear Signal and Image Processing*, Grado, Italie, jun 2003.

[11] B. Picinbono, "On Circularity," *Proceedings of the IEEE*, vol. 42, no. 12, pp. 3473–3482, Dec. 1994.