



HAL
open science

Estimation de la densité spectrale de puissance d'un bruit gaussien basée sur le kurtosis des statistiques minimales

Fabien Millioz, Nadine Martin

► **To cite this version:**

Fabien Millioz, Nadine Martin. Estimation de la densité spectrale de puissance d'un bruit gaussien basée sur le kurtosis des statistiques minimales. 22ème colloque GRETSI sur le Traitement du Signal et des Images, GRETSI 2009, Sep 2009, Dijon, France. hal-00385659v1

HAL Id: hal-00385659

<https://hal.science/hal-00385659v1>

Submitted on 19 May 2009 (v1), last revised 3 May 2011 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Estimation de la densité spectrale de puissance d'un bruit gaussien basée sur le kurtosis des statistiques minimales

Fabien MILLIOZ¹, Nadine MARTIN¹

¹GIPSA-lab DIS

BP 46 - 961 rue de la Houille Blanche, 38402 Saint Martin d'Hères Cedex

tél. 04 76 82 64 24 - fax. 04 76 82 63 84 - prénom.nom@gipsa-lab.inpg.fr

Thème – Analyse spectrale, temps-fréquence

Problème traité – L'estimation de la variance d'un bruit dans le plan temps-fréquence d'un mélange signal + bruit à l'aide des statistiques minimales.

Originalité – La loi des statistiques minimales d'une gaussienne sont décrites, pour être utilisée dans la transformée de Fourier à court terme.

Résultats – La méthode proposée estime le niveau de bruit en réduisant les choix *ad hoc* des techniques existantes.

Résumé

La détection aveugle d'un signal dans le plan temps-fréquence peut être défini à l'aide d'un test d'hypothèse, qui va différencier l'énergie d'un point temps-fréquence provenant d'un bruit uniquement ou contenant une contribution du signal à détecter. Ce test d'hypothèse prend classiquement la forme d'un seuil sur l'énergie.

Dans une détection aveugle, la localisation et les propriétés du signal sont inconnues, la détection se fait alors uniquement à partir des propriétés du bruit, qui sont à estimer. Le problème essentiel réside dans l'estimation de ces propriétés afin de construire le détecteur.

Sous l'hypothèse d'un signal additionné d'un bruit blanc gaussien stationnaire, il s'agit d'estimer la variance du bruit, la localisation des points ne contenant que du bruit étant *a priori* inconnue. Deux techniques principales existent pour pallier ce problème. La première [1, 2, 3] repose sur un algorithme itératif qui estime initialement la variance du bruit à partir de l'ensemble du plan temps-fréquence. Cette valeur initiale surestimée permet de détecter un premier sous-ensemble de points temps-fréquence contenant du signal. Les itérations suivantes réestiment la variance du bruit sur les points non détectés. Ces itérations s'arrêtent lorsqu'un critère concernant les points non détectés est rempli, que ce soit la convergence de la vraisemblance de leurs caractéristiques statiques [1], le passage d'un seuil sur l'asymétrie (*skewness*) [2] ou d'un seuil sur le kurtosis [3]. Afin de s'adapter aux bruits non blancs ou non stationnaire, les itérations de détections/estimations peuvent être effectuées sur des zones limitées du plan temps-fréquence.

Ces méthodes présentent l'inconvénient d'être relativement lentes, dépendant grandement de la qualité du critère d'arrêt.

La deuxième technique consiste à utiliser les statistiques minimales, c'est-à-dire à estimer la variance à partir des points de plus petite énergie, ceux-ci étant alors supposés ne provenir que du bruit. Rainer Martin [4, 5] utilise la plus petite valeur du plan temps-fréquence après un lissage récursif, tandis que Ewans et Mason [6] considèrent le $q^{\text{ième}}$ quantile comme estimation du bruit, dans le cadre de la reconnaissance vocale. Huillery [7] a dans sa thèse détaillé le cas de l'utilisation des Z plus petites valeurs, nécessitant alors la connaissance du nombre de points ne contenant que du bruit.

Toutes ces techniques nécessitent des choix *ad hoc*, ou l'estimation du nombre de points ne contenant que du bruit.

Cet article propose une nouvelle méthode générale d'estimation du bruit par les statistiques minimales, et s'appuie sur le caractère linéaire de la Transformée de Fourier à Court Terme (TFCT), dont le spectrogramme est le module carré. Par sa linéarité, la TFCT d'un bruit gaussien est un bruit gaussien complexe. Un seuil sur le spectrogramme revient à seuiller la valeur absolue de la TFCT. Les points de bruit sont alors issus d'une loi gaussienne tronquée, qui sera définie dans une première partie, et dont la variance et le kurtosis seront donnés. Cette loi sera étendue au cas de la TFCT d'un bruit blanc dans une seconde partie.

Une dernière partie illustrera l'utilité de cette loi dans le cadre de l'estimation par les statistiques minimales sur un signal synthétique et un signal de parole.

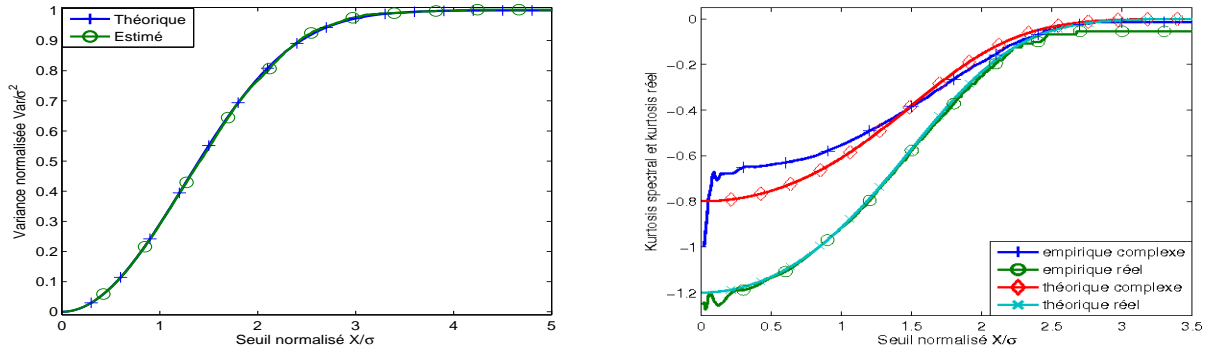


FIG. 1 : Variance et kurtosis de la partie réelle de la TFCT d'un bruit blanc gaussien de variance $\sigma^2 = 1$, pour une fenêtre de Hanning de 127 points, 63 points de recouvrement entre deux fenêtres successives, et 512 fréquences calculées. En bleu les valeurs théoriques, en vert les valeurs estimées.

1 Loi normale tronquée

Lorsque que l'on ne garde des réalisations d'une variable aléatoire gaussienne centrée de variance σ^2 que celles dont la valeur absolue est inférieure à un seuil X , on définit une variable aléatoire G dont la densité de probabilité $g_{X,\sigma^2}(x)$ s'écrit [8] :

$$g_{X,\sigma^2}(x) = \begin{cases} \frac{1}{\text{erf}(X/(\sigma\sqrt{2}))} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} & \text{si } -X \leq x \leq X, \text{ avec } \text{erf}(u) = \frac{2}{\sqrt{\pi}} \int_0^u e^{-t^2} dt \\ 0 & \text{sinon.} \end{cases} \quad (1)$$

Le troncature de la variable gaussienne centrée étant symétrique par rapport à zéro, la nouvelle variable G est également centrée. De l'expression de sa densité de probabilité découle celle de sa variances Var_{X,σ^2} :

$$Var_{X,\sigma^2} = E(x^2) = \int_{-\infty}^{+\infty} x^2 g_{X,\sigma^2}(x) dx = \sigma^2 \left(1 - \frac{2}{\sqrt{2\pi} \text{erf}(X/(\sigma\sqrt{2}))} e^{-\frac{1}{2} \frac{X^2}{\sigma^2}} \frac{X}{\sigma} \right) \quad (2)$$

Ainsi que de son cumulant d'ordre 4 $Cum4_{X,\sigma^2}$ [9] :

$$Cum4_{X,\sigma^2} = E(x^4) - 3E(x^2)^2 = \sigma^4 \left(\frac{2}{\sqrt{2\pi} \text{erf}(X/(\sigma\sqrt{2}))} e^{-\frac{1}{2} \frac{X^2}{\sigma^2}} \left(3 \frac{X}{\sigma} - \frac{X^3}{\sigma^3} \right) - \frac{6}{\pi \text{erf}(X/(\sigma\sqrt{2}))^2} e^{-\frac{X^2}{\sigma^2}} \frac{X^2}{\sigma^2} \right) \quad (3)$$

Ces deux expressions permettent de définir le kurtosis K , qui est le cumulant d'ordre 4 normalisé :

$$\mathcal{K}_{X,\sigma^2} = Cum4_{X,\sigma^2} / (Var_{X,\sigma^2})^2 \quad (4)$$

D'après les équations (2) et (3), la variance normalisée $\frac{Var_{X,\sigma^2}}{\sigma^2}$ et le kurtosis ne dépendent que du rapport $\frac{X}{\sigma}$.

2 Loi normale tronquée dans la TFCT

La Transformée de Fourier à Court Terme $S_\phi[n, k]$ d'un signal discret $s[n]$ s'écrit :

$$S_\phi[n, k] = \sum_{m=n-(M_\phi-1)/2}^{n+(M_\phi-1)/2} s[m] \phi[m-n] e^{-2j\pi k \frac{m}{K}}, 1 \leq k \leq K \quad (5)$$

Où $\phi[n]$ est une fenêtre d'apodisation de M_ϕ points, et K le nombre de fréquences calculées. La transformée de Fourier d'une variable aléatoire est une variable aléatoire complexe circulaire [10]. Dans le cas de la TFCT d'un bruit blanc gaussien, cela signifie que ses parties réelle et imaginaire sont indépendantes et de même variance.

Dans cet article, nous ferons l'approximation dans ce papier que la circularité se retrouve parfaitement dans le cas discret, bien que ce soit faux pour les valeurs de k proches de 0 et de K [3]. Dans ce cas, les parties réelles et imaginaire de la TFCT d'un bruit blanc gaussien centré de variance σ^2 suivent une loi gaussienne centrée de variance $\frac{\sigma^2}{2}$.

Afin de vérifier les équations (2,3,4) dans le cadre de la TFCT, on estime variance et kurtosis des coefficients tronqués de la TFCT d'un bruit blanc gaussien de 5000 points de variance $\sigma^2 = 1$. La TFCT est calculée avec une fenêtre de Hanning de 127 points, 63 points de recouvrement entre deux fenêtres successives, et 512 fréquences calculées. Le même seuil X est appliqué séparément sur la partie réelle et la partie imaginaire des coefficients. La figure 1 montre l'évolution de la variance et du kurtosis estimé de la partie réelle de la variable tronquée, comparés avec leurs valeurs théoriques. Les valeurs obtenues en pratique correspondent bien aux valeurs attendues en théorie, excepté pour le kurtosis avec un seuil X très faible. Le peu de points retenus dans ce cas conduit à une grande variance de la valeur estimée. La section suivante s'intéresse au problème du mélange de signal et de bruit, pour déduire une estimation de σ^2 .

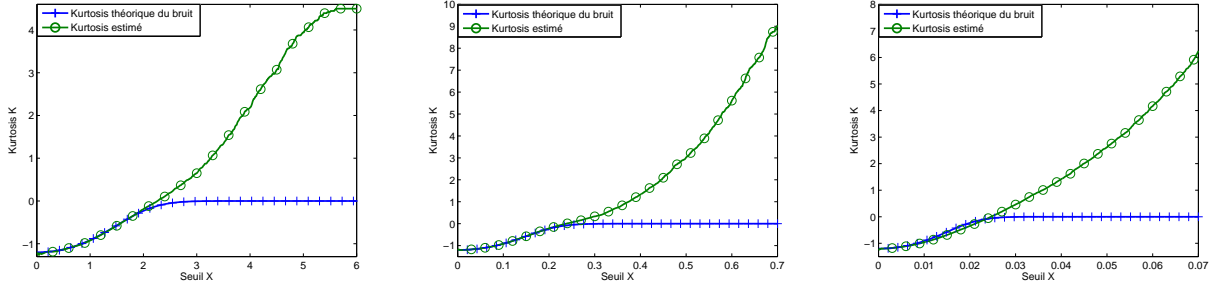


FIG. 2 : À gauche, kurtosis de la partie réelle d'un signal composé de trois modulations de fréquence additionné d'un bruit blanc gaussien de variance 1. Au centre et à droite, kurtosis de la partie réelle de la TFCT d'un signal parole auquel est ajouté respectivement un bruit blanc de variance 10^{-2} et 10^{-4} . Pour les faibles valeurs du seuil X , le kurtosis d'un mélange signal+bruit a la même allure que le kurtosis d'un bruit seul.

3 Estimation du niveau de bruit à partir d'un mélange signal et bruit

Le signal synthétique, composé de trois modulations de fréquence, et un signal de parole sont considérés ici. Les paramètres de la TFCT sont les mêmes que ceux utilisés pour la figure 1. Un bruit blanc de variance 1 est ajouté au signal synthétique, tandis que des bruits de variance 10^{-2} et de variance 10^{-4} sont ajoutés séparément au signal de parole. Le premier bruit noie les fricatives, contrairement au second. La figure 2 illustre l'évolution du kurtosis de la partie réelle de la TFCT des signaux bruités selon le seuil X . Pour les faibles valeurs de seuil, le kurtosis estimé correspond encore au kurtosis du bruit seul.

Il reste à déterminer quel point de la courbe théorique du kurtosis utiliser pour estimer la variance du bruit. L'idée retenue est de prendre la valeur correspondant au point d'inflexion, c'est-à-dire $K_{X,\sigma^2} = -0.6$. Le kurtosis ne dépendant que de $\frac{X}{\sigma}$, il est possible de déterminer quelle valeur de $\frac{X}{\sigma}$ correspond à un kurtosis de -0.6. De par la complexité de l'expression (4), cette valeur est calculée numériquement, et vaut 2.0655.

L'estimateur de σ^2 s'écrit en fonction de la valeur du seuil X_{06} tel que le kurtosis des points de plus faible valeur absolue soit égale à 0.6 :

$$\hat{\sigma} \quad / \quad \mathcal{K}_{X,\sigma^2}(X_{06}, \sigma) = \mathcal{K}_{X,\sigma^2}\left(\frac{X_{06}}{\sigma}\right) = -0.6 \quad \text{soit } \hat{\sigma} = \frac{X_{06}}{2.0655} \quad (6)$$

Pour les trois signaux de la figure 2 bruités respectivement avec des bruits blancs de variance 1, 10^{-2} et 10^{-4} , les seuils X_{06} sont 1.147, 0.146, 0.016. Les écarts-types des bruits dans la partie réelle de la TFCT sont alors de 0.714, 0.0707 et 0.0077 respectivement, ce qui estime les variance de bruit σ^2 à 1.02, 10^{-2} et $1.2 \cdot 10^{-4}$.

Cet estimateur, dont les propriétés reposent sur celles de l'estimation du kurtosis d'un ensemble de points issus d'une gaussienne tronquée présente l'avantage de n'avoir qu'une seule valeur fixée *a priori*, le choix du points $\mathcal{K} = 0.6$ tout en ne faisant aucune hypothèse sur le signal, au contraire des techniques existantes.

Références

- [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] Chungshin 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 smoothing and minimum statistics," *Speech and Audio Processing, IEEE Trans. on*, vol. 9, no. 5, pp. 504–512, 2001.
- [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. Ewans and John S. Mason, "Time-Frequency Quantile-Based Noise," in *Proceedings of EUSIPCO*, 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] N.L. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions*, vol. 2, Wiley and sons, 2nd edition, 1994.
- [9] J.-L. Lacoume, P.-O. Amblard, and P. Comon, *Statistiques d'ordre supérieur pour le traitement du signal*, Masson, 1997.
- [10] B. Picinbono, "On Circularity," *Proceedings of the IEEE*, vol. 42, no. 12, pp. 3473–3482, Dec. 1994.