

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

Fabien MILLIOZ<sup>1</sup>, Nadine MARTIN<sup>1</sup>

<sup>1</sup>GIPSA-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.grenoble-inp.fr

**Résumé** – Nous présentons dans ce papier un estimateur de la variance d'un bruit blanc gaussien à partir des valeurs minimales de la transformée de Fourier à court terme d'un signal composé d'un signal quelconque noyé dans un bruit à estimer. Le kurtosis des statistiques minimales dans le plan temps-fréquence est utilisé pour estimer la variance du bruit à partir de points ne contenant que du bruit.

**Abstract** – In this paper, we propose a variance estimator of a white Gaussian noise, using the minimal values of the Short Time Fourier Transform of a signal made of any signal embedded in a noise to estimate. The kurtosis of minimum statistics in the time-frequency plane is used to estimate the noise variance from points containing noise only.

## 1 Introduction

La détection d'un signal dans le plan temps-fréquence peut être définie à l'aide d'un test d'hypothèse, qui va différencier l'énergie d'un point temps-fréquence provenant d'un bruit uniquement de celle contenant en plus 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. 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 est nécessaire 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. La variance du bruit est d'abord estimée 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 satisfait, 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]. 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 du bruit 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 soit des choix *ad hoc*, soit 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 module carré est le spectrogramme. Par sa linéarité, la TFCT d'un bruit gaussien est un bruit gaussien complexe. Un seuil sur le spectrogramme, classiquement utilisé pour détecter les points contenant du signal, revient à seuiller la valeur absolue de la TFCT. En rejetant les points de valeur absolue supérieure au seuil, les points de bruit dans la TFCT voient leurs parties réelle et imaginaire provenir d'une loi gaussienne tronquée, qui sera définie dans une première partie, et dont la variance et le kurtosis seront donnés. Cette loi sera utilisée dans le cas de la TFCT d'un bruit blanc dans une seconde partie afin d'en déduire un estimateur de la variance du bruit.

Une dernière partie appliquera cet estimateur aux cas de la TFCT d'un signal synthétique et d'un signal de parole auxquels un bruit blanc gaussien est ajouté.

## 2 Loi normale tronquée

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

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

avec  $\operatorname{erf}(u) = \frac{2}{\sqrt{\pi}} \int_0^u e^{-t^2} dt$ .

La 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 variance  $Var_{X,\sigma^2}$  [8] :

$$\begin{aligned} Var_{X,\sigma^2}(|G| < X) &= E(x^2) = \int_{-\infty}^{+\infty} x^2 g_{X,\sigma^2}(x) dx \\ &= \sigma^2 \left( 1 - \frac{2}{\sqrt{2\pi} \operatorname{erf}(X/(\sigma\sqrt{2}))} e^{-\frac{1}{2} \frac{X^2}{\sigma^2}} \frac{X}{\sigma} \right) \end{aligned} \quad (2)$$

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

$$\begin{aligned} Cum4_{X,\sigma^2}(|G| < X) &= E(x^4) - 3E(x^2)^2 \\ &= \sigma^4 \left( \frac{2}{\sqrt{2\pi} \operatorname{erf}(X/(\sigma\sqrt{2}))} e^{-\frac{1}{2} \frac{X^2}{\sigma^2}} \left( 3 \frac{X}{\sigma} - \frac{X^3}{\sigma^3} \right) \dots \right. \\ &\quad \left. \dots - \frac{6}{\pi \operatorname{erf}(X/(\sigma\sqrt{2}))^2} e^{-\frac{X^2}{\sigma^2}} \frac{X^2}{\sigma^2} \right) \end{aligned} \quad (3)$$

Ces deux expressions permettent de définir le kurtosis  $\mathcal{K}$ , qui est le cumulant d'ordre 4 normalisé :

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

D'après les équations (2,3), la variance normalisée  $\frac{Var_{X,\sigma^2}}{\sigma^2}$  et le kurtosis  $\mathcal{K}$  ne dépendent plus directement de la variance du bruit  $\sigma^2$  mais uniquement du seuil normalisé  $X_n$ , défini par :

$$X_n = \frac{X}{\sigma} \quad (5)$$

### 3 Loi normale tronquée dans la TFCT et estimation du niveau de bruit

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

$$S_\phi[n, k] = \sum_{m=n-(M_\phi-1)/2}^{n+(M_\phi-1)/2} s[m] \phi[m-n] e^{-2j\pi k \frac{m}{K}} \quad (6)$$

Où  $\phi[n]$  est une fenêtre d'apodisation de  $M_\phi$  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 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]. Afin de vérifier les équations (2,3,4) avec l'approximation faite, on estime variance et kurtosis de la partie réelle de la TFCT d'un bruit blanc gaussien, notée  $\Re(S_\phi[n, k])$ , en fonction du seuil normalisé  $X_n = \frac{X}{\sigma}$ . La TFCT est calculée sur un bruit blanc de 5000 points et de variance  $\sigma^2 = 1$ , avec une fenêtre de Hanning de 127 points, 63 points de recouvrement entre deux fenêtres successives, et 512 fréquences calculées. La figure 1 montre l'évolution des estimations de la variance normalisée et du kurtosis de la partie réelle de la TFCT tronquée par le seuil  $X$ , comparés avec leurs valeurs théoriques. Les valeurs obtenues

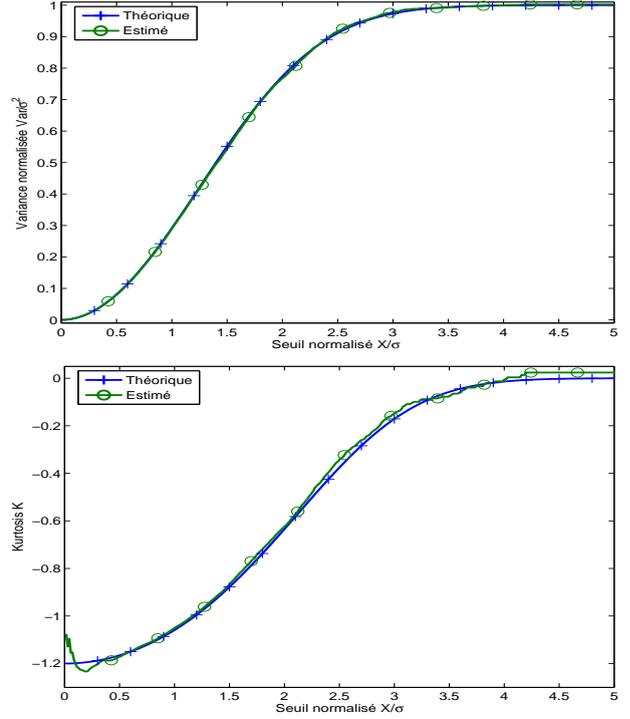


FIG. 1 : Variance et kurtosis de la partie réelle de la TFCT d'un bruit blanc gaussien de variance  $\sigma^2 = 1$ , pour une fenêtre de Hanning de 127 points, 63 points de recouvrement entre deux fenêtres successives, et 512 fréquences calculées. En bleu les valeurs théoriques, en vert les valeurs estimées.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

$\mathcal{K}_{X,\sigma^2}$	-1	-0.8	-0.6	-0.4	-0.2
$X_n$	1.1869	1.6712	2.0655	2.4506	2.9142

TAB. 1 : Valeurs approchées des seuils normalisés  $X_n = \frac{X}{\sigma}$  correspondants à différentes valeurs du kurtosis d'une gaussienne tronquée.

en pratique correspondent bien aux valeurs attendues en théorie, sauf pour le kurtosis avec un seuil  $X$  très faible. Dans ce cas, le peu de réalisations inférieures à ce seuil conduit à une grande variance de la valeur estimée.

Le kurtosis d'une variable aléatoire présente l'avantage de ne pas dépendre de la variance de cette variable aléatoire. De plus, dans le cadre de la loi gaussienne tronquée, le kurtosis varie de manière monotone selon le seuil normalisé  $X_n$ . Ainsi, il n'existe qu'un seul seuil  $X_n(\kappa)$  conduisant à la valeur du kurtosis  $\mathcal{K}_{X,\sigma^2} = \kappa$ , quelle que soit la variance  $\sigma^2$  de la gaussienne tronquée. Les valeurs  $X_n(\kappa)$  sont déterminées en inversant la fonction de l'expression (4). De par la complexité de cette équation, ces valeurs sont calculées numériquement. La tableau 1 donne quelques-unes de ces valeurs.

À partir d'une TFCT d'un bruit blanc gaussien de variance inconnue, il est possible de déterminer le seuil  $X_r(\kappa)$  non normalisé tel que le kurtosis des points de

la partie réelle de la TFCT,  $\Re(S_\phi[n, k])$ , de valeur absolue inférieure à ce seuil soit égal à  $\kappa$ .  $X_r(\kappa)$  s'écrit :

$$X_r(\kappa) / \mathcal{K}_{X_r(\kappa)} (|\Re(S_\phi)| < X_r(\kappa)) = \kappa \quad (7)$$

À partir du seuil normalisé  $X_n(\kappa)$  obtenu par inversion de l'équation  $\mathcal{K}_{X_n,1} = \kappa$ , de la valeur du seuil déterminé en pratique  $X_r(\kappa)$  et du lien entre les deux (5), on en déduit un estimateur de la variance de la partie réelle de la TFCT  $\sigma_r^2$  :

$$\hat{\sigma}_r^2(\kappa) = \left( \frac{X_r(\kappa)}{X_n(\kappa)} \right)^2 \quad (8)$$

La variance de la partie imaginaire de la TFCT,  $\Im(S_\phi[n, k])$ , est estimée de la même manière, et est notée  $\hat{\sigma}_i^2$ . La variance du bruit  $\sigma^2$  est égal à la somme des variances de ses parties réelle et imaginaire, son estimateur  $\hat{\sigma}^2$  s'écrit alors :

$$\hat{\sigma}^2(\kappa) = \hat{\sigma}_r^2(\kappa) + \hat{\sigma}_i^2(\kappa) = \left( \frac{X_r(\kappa)}{X_n(\kappa)} \right)^2 + \left( \frac{X_i(\kappa)}{X_n(\kappa)} \right)^2 \quad (9)$$

Cet estimateur permet de retrouver la valeur de  $\sigma^2$  à partir des échantillons des parties réelle et imaginaire de la TFCT de plus faibles valeurs absolues. Il ne présente aucun intérêt dans le cas de la TFCT d'un bruit seul, tous les échantillons provenant d'une même loi gaussienne dont la variance est à estimer. Toutefois, dans le cas d'une TFCT issue d'un signal quelconque auquel est ajouté un bruit blanc gaussien de variance inconnue à estimer, l'estimateur ne considère que les points de plus petite valeur absolue, autrement dit de plus petite énergie, permettant ainsi de rejeter les points contenant une contribution énergétique du signal.

La section suivante s'intéresse à ce problème du mélange d'un signal quelconque et d'un bruit blanc gaussien, en déterminant quelle valeur de  $\kappa$  utiliser.

## 4 Estimation du niveau de bruit d'un mélange signal et bruit

L'estimateur proposé est testé sur un signal synthétique de 8000 points composé de trois modulations de fréquence et un signal de parole de 36472 points échantillonné à 44.1kHz. Les paramètres de la TFCT sont les mêmes que ceux utilisés pour la figure 1, c'est à dire une fenêtre de Hanning de 127 points, un recouvrement de 63 points et 512 fréquences calculées. 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 distinctement 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  $X_r$ . Le kurtosis calculé à partir de la partie imaginaire de ces TFCT suit des variations identiques. Pour les faibles valeurs de seuil, le kurtosis estimé correspond au kurtosis du bruit uniquement. Ainsi, nous vérifions que les points de plus petite valeur absolue correspondent bien à des points de bruit seul, par conséquent

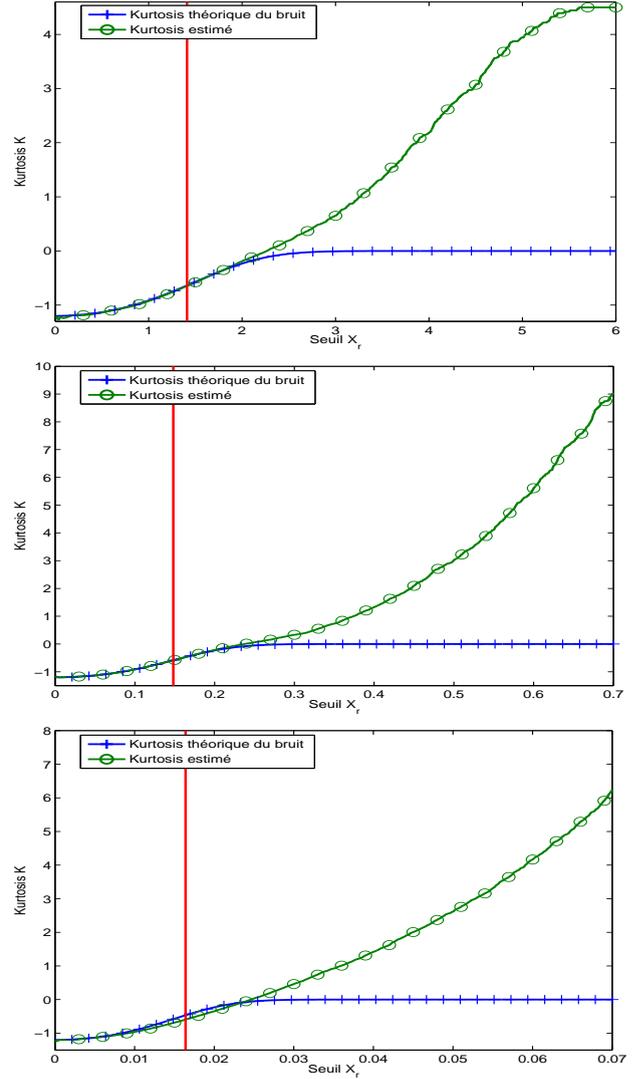


FIG. 2 : En haut, kurtosis de la partie réelle d'un signal composé de trois modulations de fréquence additionné d'un bruit blanc gaussien de variance 1. Au centre et en bas, kurtosis de la partie réelle de la TFCT d'un signal de parole auquel est ajouté respectivement un bruit blanc de variance  $10^{-2}$  et  $10^{-4}$ . Pour les faibles valeurs du seuil  $X_r$ , le kurtosis d'un mélange signal+bruit a la même allure que le kurtosis d'un bruit seul. Les droites verticales correspondent aux seuils  $X_r(-0.6)$  tels que  $\mathcal{K}_{X_r, \sigma^2} (|\Re(S_\phi[n, k])| < X_r) = -0.6$ .

que les statistiques minimales permettent d'estimer la variance de ce bruit, sans besoin de connaissance *a priori* sur la position dans le plan temps-fréquence des points contenant de l'énergie du signal.

Il reste à déterminer quel point de la courbe théorique du kurtosis utiliser pour estimer la variance du bruit, autrement dit quelle valeur de  $\kappa$  de l'équation (9) choisir. L'idée naturelle est de prendre une valeur proche du point d'inflexion. On retient alors  $\kappa = \mathcal{K}_{X_r, \sigma^2} = -0.6$ , correspondant à un seuil normalisé  $X_n \approx 2.0655$ . L'estimation de la variance de bruit se fait alors selon l'équation 9.

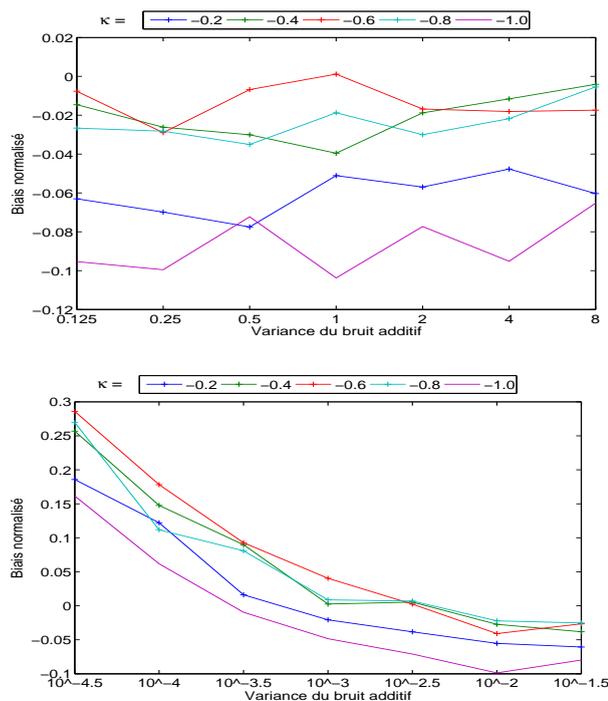


FIG. 3 : *Biais normalisés (10) des estimations des variances du bruit additif (9) à partir des TFCTs d'un signal synthétique (en haut) et d'un signal de parole (en bas), en fonction de la valeur de  $\kappa$  utilisée (en couleurs). L'estimateur du kurtosis utilisé possédant un biais positif, les estimations obtenues présentent un biais négatif, sauf dans le cas de faibles bruits ajoutés au signal de parole. Le bruit d'enregistrement ainsi que les fricatives rajoutent une énergie non négligeable aux points temps-fréquence ne contenant que du bruit additif, alors assimilable à un bruit additif de variance plus grande : une surestimation de la variance du bruit réellement ajouté apparaît.*

Pour les trois signaux de la figure 2, les seuils  $X_r(-0.6)$  sont respectivement 1.4099, 0.14827 et 0.016398, tandis que les seuils  $X_i(-0.6)$  sont respectivement 1.4489, 0.14755 et 0.015654. Les seuils calculés en pratique sur les parties réelle et imaginaire sont sensiblement égaux. D'après l'équation 9, les variances estimées  $\hat{\sigma}^2$  sont respectivement 0.958,  $1.026 \cdot 10^{-2}$  et  $1.205 \cdot 10^{-4}$ .

Pour vérifier le bon choix de  $\kappa = -0.6$ , nous nous intéressons au biais normalisé  $b(\kappa)$  de cet estimateur, défini par :

$$b(\kappa) = \frac{E(\hat{\sigma}^2(\kappa)) - \sigma^2}{\sigma^2} \quad (10)$$

La figure 3 montre l'évolution du biais normalisé selon la variance du bruit additif et pour différentes valeurs de  $\kappa$ , pour les signaux synthétique et le signal de parole. Le biais est estimé à partir de 100 réalisations de bruit. Pour le signal synthétique, le biais est plus faible à mesure que  $\kappa$  se rapproche de  $-0.6$ . Pour le signal de parole, un autre phénomène intervient, à savoir qu'il n'y a aucun point temps-fréquence d'énergie strictement nulle. Le bruit d'enregistrement et les fricatives augmentent la valeur des points de faible amplitude

de manière non négligeable lorsque le bruit additif est faible.

## 5 Conclusion et perspectives

Ce papier a présenté un estimateur de la variance d'un bruit blanc gaussien, à partir des statistiques minimales de la TFCT. Il est basé sur le kurtosis d'une loi gaussienne tronquée, par conséquent ses propriétés reposent sur celles de l'estimation du kurtosis. Cet estimateur présente l'avantage de n'avoir qu'une seule valeur fixée *a priori*, le choix du point de la courbe du kurtosis considéré, qui peut être choisi de manière à obtenir un biais le plus proche de zéro possible.

Une amélioration possible est de prendre en considération la non circularité de la TFCT [3], afin d'améliorer l'adéquation entre la loi théorique de la gaussienne tronquée et les lois des parties réelle et imaginaire tronquées de la TFCT. Une extension de ces travaux pour utiliser le kurtosis spectral des valeurs complexes de la TFCT plutôt que considérer séparément ses parties réelle et imaginaire est en cours.

## 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.