

# Approximation de la loi de probabilité du spectrogramme, KL Divergence et détection dans le plan temps-fréquence

Julien HUILLERY, Nadine MARTIN

Gipsa-lab, Département Image et Signal, INPG/CNRS  
961 rue de la Houille Blanche, 38402 Saint Martin d'Hères, France  
julien.huillery@gipsa-lab.inpg.fr, nadine.martin@gipsa-lab.inpg.fr

**Résumé** – La mise en place d'un test de détection dans le plan temps-fréquence sous la forme d'un test binaire d'hypothèses demande la connaissance des lois de probabilité de l'observation. Dans le cas du spectrogramme et pour un signal gaussien non-blanc, on cherche à approcher la loi des coefficients temps-fréquence avec une loi du  $\chi^2$ . À l'aide de la divergence de Kullback-Leibler, la validité de l'approximation est évaluée en fonction de la fenêtre utilisée pour l'analyse de Fourier et du degré de corrélation contenu dans le signal. En se replaçant dans le contexte de détection, la divergence de Kullback-Leibler correspond à la perte moyenne de puissance du test de détection. Les conclusions obtenues sur la validité de l'approximation de la distribution des coefficients du spectrogramme par une loi du  $\chi^2$  trouvent ainsi une interprétation pratique.

**Abstract** – A time-frequency detection task requires knowledge of the observation probability density functions under both null and signal hypotheses. Classically with spectrogram, the central and non-central  $\chi^2$  laws are employed. In the context of correlated Gaussian signal, those  $\chi^2$  laws appear to be an approximation of spectrogram coefficients probability laws. This approximation is first evaluated with the Kullback-Leibler Divergence as a function of the Fourier analysis window and the degree of correlation present in the signal. Back in the detection context, it turns out that this measure of divergence also indicates the average loss of power of the detection test. This allows to draw practical conclusions on the possible use of  $\chi^2$  laws to describe spectrogram probabilities in a detection task.

## 1 Introduction

La détection dans le plan temps-fréquence vise à déterminer les zones du plan temps-fréquence sur lesquelles le signal étudié présente de l'énergie. Interprété comme une signature temps-fréquence des signaux, le résultat de cette détection devient une caractéristique utilisable dans les problèmes de classification ou de reconnaissance. Des solutions aux problèmes de détection ont été trouvées dans les approches statistiques liées aux tests d'hypothèses [1]. La mise en place du détecteur demande alors la connaissance des lois de probabilité de l'observation sous les différentes hypothèses envisagées. Cependant, ces lois ne sont pas toujours faciles à formuler et une solution consiste à utiliser des lois plus "simples" ou mieux connues.

C'est le cas pour le spectrogramme. Sous l'hypothèse d'un signal noyé dans un bruit gaussien, centré et non-blanc, la loi de probabilité des coefficients du spectrogramme reste difficile à formuler et à utiliser [2]. On note que ce même contexte est étudié dans [3] pour le périodogramme estimé selon la méthode de Welch. Le cas du bruit blanc est reporté dans [4]. Aussi, la loi du spectrogramme est généralement approchée par une loi du  $\chi^2$ . Mieux connue, de formulation plus simple, cette loi correspond en outre à celle du spectrogramme lorsque le bruit additif est blanc et qu'une fenêtre rectangulaire est utilisée pour l'analyse. Pour d'autres fenêtres d'analyse et lorsque l'hypothèse de blancheur du bruit n'est plus valide, l'utilisation d'une loi du  $\chi^2$  pour décrire le spectrogramme devient une approximation. L'objectif de cet article est alors d'évaluer l'erreur engendrée par cette approximation

et de répondre à la question : sous quelles conditions peut-on utiliser une loi du  $\chi^2$  ?

Une première solution consiste à mesurer la différence entre la loi réelle et son approximation. Instructive sur le plan qualitatif, cette première approche ne permet cependant pas de répondre pratiquement à notre question. Parallèlement, remplacer la loi de probabilité relative à une hypothèse par une autre loi va avoir des répercussions sur les résultats du test de détection mis en place. De ce point de vue, une façon de mesurer l'impact de l'approximation réalisée sur les lois de probabilité consiste à évaluer les variations qu'elle engendre sur les performances du détecteur. Récemment, Eguchi et Copas [5] ont montré que la divergence de Kullback-Leibler, qui mesure la différence entre deux lois de probabilité, permet aussi de mesurer les variations des performances d'un test de détection lorsque l'une des lois vient remplacer la seconde. Cette mesure sera donc utilisée pour évaluer l'erreur d'approximation.

Cet article est organisé de la façon suivante : le contexte d'extraction du contenu temps-fréquence d'un signal à travers la mise en place d'un test de détection est d'abord précisé. Dans un second temps, la loi des coefficients d'un spectrogramme sous l'hypothèse d'un bruit non-blanc et son approximation par une loi du  $\chi^2$  sont présentées. Dans une troisième partie, l'erreur d'approximation et son impact sur les performances du test de détection sont évalués avec la divergence de Kullback-Leibler et finalement, les conclusions sur la validité de l'utilisation de la loi du  $\chi^2$  pour décrire les coefficients du spectrogramme sont données.

## 2 Détection de signaux dans le plan temps-fréquence

On suppose l'observation discrète  $x[m]$  composée d'une information  $i[m]$  perturbée par un signal  $p[m]$ , soit

$$x[m] = i[m] + p[m]. \quad (1)$$

L'information est supposée déterministe tandis que la perturbation revêt un caractère aléatoire. Un processus gaussien, centré, stationnaire et de fonction d'autocorrélation  $\Gamma_p[\tau]$  est adopté comme modèle pour cette perturbation.

Le problème de détection dans le plan temps-fréquence est posé de la manière suivante : étant donné le spectrogramme  $S_x[n, k]$  de l'observation  $x[m]$  ( $n$  est l'indice de la trame temporelle et  $k$  l'indice des fréquences), on veut décider si le coefficient observé à la position  $[n, k]$  résulte de la perturbation  $p[m]$  uniquement (hypothèse  $H_0$ ) ou reflète aussi la contribution de la composante informative  $i[m]$  (hypothèse  $H_1$ ). On formalise ce problème de détection par un test binaire d'hypothèses dans le plan temps-fréquence qui s'écrit

$$\begin{cases} H_0 : S_x[n, k] = S_p[n, k], \\ H_1 : S_x[n, k] = S_{i+p}[n, k]. \end{cases} \quad (2)$$

Le test du rapport des vraisemblances propose une solution pour la mise en place du détecteur. Il s'écrit

$$\mathcal{L}_{S_x[n, k]} = \frac{p_{H_1}(S_x[n, k])}{p_{H_0}(S_x[n, k])} \underset{H_0}{\overset{H_1}{\geq}} \delta. \quad (3)$$

Des considérations de coût de décision et/ou de probabilités *a priori* sur les hypothèses donnent la valeur du seuil  $\delta$ . Dans une approche Neyman-Pearson, ce seuil est fixé par le choix de la probabilité de fausse alarme [1]. Le support temps-fréquence du signal correspond finalement à l'ensemble des positions  $[n, k]$  pour lesquelles l'hypothèse  $H_1$  aura été retenue.

Ce test de détection requiert l'utilisation, et auparavant la détermination, des densités de probabilité  $p_{H_0}(S_x[n, k])$  et  $p_{H_1}(S_x[n, k])$  des coefficients du spectrogramme sous les hypothèses  $H_0$  et  $H_1$ .

## 3 Densité de probabilité du spectrogramme et approximation

Soit  $\mathbf{X}[n, k] = [X^r[n, k] \ X^i[n, k]]^T$  le vecteur issu des composantes réelle et imaginaire de la Transformée de Fourier à Court Terme (TFCT) avec

$$\begin{cases} X^r[n, k] = \sum_{m=0}^{M-1} x[n-m]w[m] \cos(-2\pi k \frac{m}{K}), \\ X^i[n, k] = \sum_{m=0}^{M-1} x[n-m]w[m] \sin(-2\pi k \frac{m}{K}), \end{cases} \quad (4)$$

où  $w[m]$  est la fenêtre d'analyse de longueur  $M$ ,  $K$  est la longueur de la FFT et  $K/M$  correspond au facteur de *zero-padding*.

Les composantes  $X^r[n, k]$  et  $X^i[n, k]$  sont des combinaisons linéaires des échantillons du signal (équation (4)). Sous l'hypothèse d'une perturbation  $p[m]$  gaussienne, le

signal  $x[m]$  est alors gaussien et par conséquent  $\mathbf{X}[n, k]$  est un vecteur gaussien à deux dimensions que l'on note

$$\mathbf{X}[n, k] \sim \mathcal{N} \left\{ \begin{pmatrix} m_r \\ m_i \end{pmatrix}, \boldsymbol{\Sigma} = \begin{pmatrix} \Sigma_{rr} & \Sigma_{ri} \\ \Sigma_{ri} & \Sigma_{ii} \end{pmatrix} \right\}. \quad (5)$$

D'après le modèle de signal adopté (équation (1)), les cinq paramètres de ce vecteur gaussien s'expriment de la façon suivante :

$$\begin{cases} m_r = \text{Re}(STFT\{i[m]\}), \\ m_i = \text{Im}(STFT\{i[m]\}), \\ \Sigma_{rr} = \mathbf{W}^T \mathbf{C}_k \mathbf{R} \mathbf{C}_k \mathbf{W}, \\ \Sigma_{ii} = \mathbf{W}^T \mathbf{S}_k \mathbf{R} \mathbf{S}_k \mathbf{W}, \\ \Sigma_{ri} = \mathbf{W}^T \mathbf{C}_k \mathbf{R} \mathbf{S}_k \mathbf{W}, \end{cases} \quad (6)$$

où le vecteur  $\mathbf{W}^T = [w[0] \cdots w[M-1]]$  représente la fenêtre d'analyse,  $\mathbf{C}_k$  et  $\mathbf{S}_k$  sont les matrices diagonales des cosinus et sinus,

$$\begin{cases} \mathbf{C}_k = \text{diag} [\cos(-2\pi k \frac{m}{K})]_{m=0, M-1}, \\ \mathbf{S}_k = \text{diag} [\sin(-2\pi k \frac{m}{K})]_{m=0, M-1}, \end{cases} \quad (7)$$

et  $\mathbf{R}$  est la matrice de covariance associée à la fonction d'autocorrélation  $\Gamma_p[\tau]$ , définie selon

$$\mathbf{R} = \begin{bmatrix} \Gamma_p[0] & \Gamma_p[1] & \cdots & \Gamma_p[M-1] \\ \Gamma_p[1] & \Gamma_p[0] & \cdots & \Gamma_p[M-2] \\ \Gamma_p[2] & \Gamma_p[1] & \cdots & \Gamma_p[M-3] \\ \vdots & \vdots & \ddots & \vdots \\ \Gamma_p[M-1] & \Gamma_p[M-2] & \cdots & \Gamma_p[0] \end{bmatrix}. \quad (8)$$

Le spectrogramme se définit comme le module carré du vecteur  $\mathbf{X}[n, k]$ , soit

$$S_x[n, k] = |\mathbf{X}[n, k]|^2 = X^r[n, k]^2 + X^i[n, k]^2. \quad (9)$$

Sa distribution de probabilité est donc celle du module carré d'un vecteur gaussien à deux dimensions, dénomé 'module carré gaussien' dans la suite, caractérisée par les cinq paramètres  $m_r$ ,  $m_i$ ,  $\Sigma_{rr}$ ,  $\Sigma_{ii}$  et  $\Sigma_{ri}$ . Lorsque  $\boldsymbol{\Sigma}$  est proportionnelle à la matrice identité ( $\Sigma_{rr} = \Sigma_{ii}$  et  $\Sigma_{ri} = 0$ ), le module carré gaussien devient proportionnel à une variable du  $\chi^2$ , notée  $\chi^2$ , à deux degrés de liberté et de paramètre de décentrage  $m_r^2 + m_i^2$ . Parmi les avantages de cette loi, sa densité possède une écriture analytique et des méthodes pour l'estimation de ses paramètres existent (voir par exemple [6] et [7]).

La figure 1 présente le rapport des variances  $\Sigma_{rr}/\Sigma_{ii}$  ainsi que le coefficient de corrélation  $\rho = \Sigma_{ri}/(\Sigma_{rr}\Sigma_{ii})^{1/2}$  entre les parties réelle et imaginaire de la STFT en fonction de la fréquence.

Lorsque que le signal analysé est blanc (figure 1-a)), la matrice  $\boldsymbol{\Sigma}$  est toujours très proche d'une matrice identité : quelle que soit la fenêtre d'analyse utilisée (ici rectangulaire, Kaiser ou Hanning), le rapport de variances est égale à 1 avec une marge d'erreur de  $10^{-3}$  et le coefficient de corrélation est inférieur à  $10^{-5}$ .

Dans un environnement non blanc en revanche, la matrice  $\boldsymbol{\Sigma}$  ne peut plus être considérée comme proportionnelle à la matrice identité. La figure 1-b) présente l'exemple

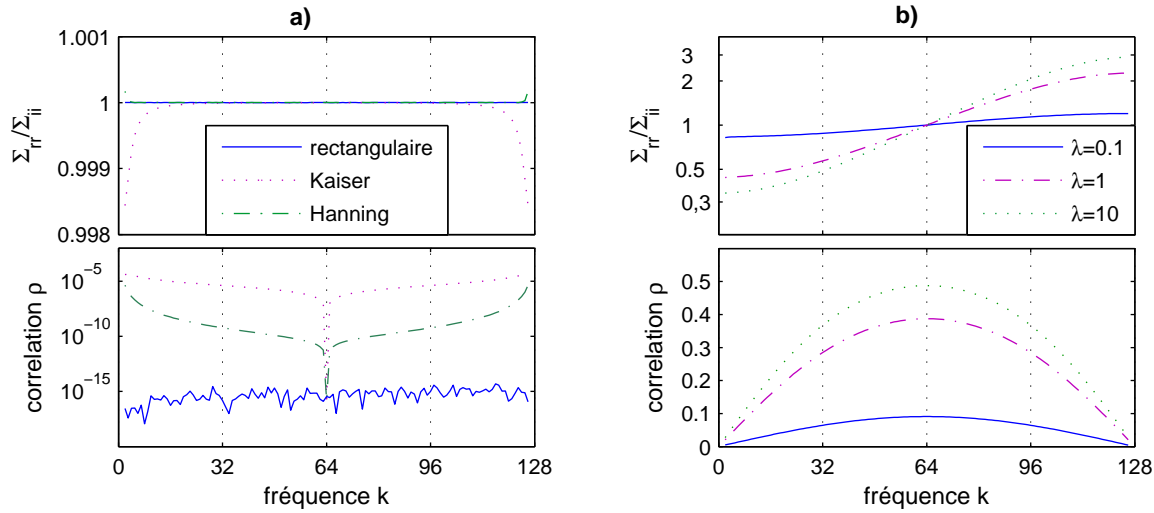


FIG. 1 – Rapport des variances  $\Sigma_{rr}/\Sigma_{ii}$  et coefficient de corrélation  $\rho = \Sigma_{ri}/(\Sigma_{rr}\Sigma_{ii})^{\frac{1}{2}}$  entre les parties réelle et imaginaire de la STFT pour un bruit blanc et différentes fenêtres d'analyse (a) et un bruit exponentiellement corrélé de paramètre  $\lambda$  et une fenêtre rectangulaire (b).

d'une perturbation exponentiellement corrélée. Plus le degré de corrélation du signal augmente (paramètre  $\lambda$ ), plus les parties réelle et imaginaire de la STFT deviennent corrélées et de variances différentes. La distribution de probabilité du spectrogramme n'est alors plus une loi du  $\chi^2$  à 2 degrés de liberté mais plus généralement une loi de module carré gaussien.

Cependant, une telle loi de probabilité n'est pas de toute simplicité [2]. Par exemple, sa densité ne possède pas toujours d'expression analytique finie ce qui rend assez difficile les tentatives de développements analytiques, comme pour ce qui nous concerne, la détermination d'un seuil de détection ou des performances du test de détection. Pour ces raisons, approcher la distribution du spectrogramme par une loi du  $\chi^2$  à 2 degrés de libertés semble intéressant. On se propose dans la suite d'évaluer les possibilités de cette approximation.

## 4 Évaluation de l'approximation

### 4.1 Mesure de divergence

La différence entre deux lois de probabilité  $p$  et  $q$  peut être mesurée avec la divergence de Kullback-Leibler, notée  $D_{KL}(p, q)$ . C'est une mesure positive qui égale zéro lorsque les lois  $p$  et  $q$  sont identiques. La divergence de Kullback-Leibler (KL) entre la distribution du spectrogramme, notée  $p_{S_x}$ , et la loi du  $\chi^2$ , notée  $p_{\chi^2}$ , se définit comme

$$D_{KL}(p_{S_x}, p_{\chi^2}) = \int p_{S_x}(x) \log \left[ \frac{p_{S_x}(x)}{p_{\chi^2}(x)} \right] dx. \quad (10)$$

La figure 2 donne l'évolution de la divergence KL entre la distribution du spectrogramme (calculée numériquement selon [2]) d'une part et celle d'un  $\chi^2$  à 2 degrés de liberté d'autre part, en fonction de la corrélation du signal. La fonction d'autocorrélation  $\Gamma_p[\tau]$  de la perturbation  $p[m]$  est ici exponentiellement décroissante, de paramètre  $\tau_c$ . Le paramètre  $\lambda$ , en abscisse des figures, est défini

comme le rapport entre la longueur  $\tau_c$  de cette corrélation et la longueur  $M$  de la fenêtre d'analyse, soit  $\lambda = \tau_c/M$ . Il mesure ainsi le degré de corrélation du signal à l'échelle de l'analyse.

La figure 2-a) présente le cas des lois centrées ( $m_r = m_i = 0$ ). Ce cas correspond à un point du plan temps-fréquence où seule la perturbation  $p[m]$  est présente. L'utilisation des fenêtres de Hanning ou Blackman engendre une erreur d'approximation faible et constante quel que soit le niveau de corrélation du signal. Pour les autres fenêtres, l'erreur augmente puis se stabilise lorsque la corrélation du signal devient aussi longue que la fenêtre ( $\lambda = 1$ ). L'utilisation d'une fenêtre rectangulaire donne une différence maximale entre la distribution du spectrogramme et celle du  $\chi^2$ . Le cas des lois décentrées ( $m_r, m_i \neq 0$ , l'information  $i[m]$  est présente) est présenté sur la figure 2-b). La divergence entre la loi du spectrogramme et celle d'un  $\chi^2$  apparaît élevée et indépendante de la fenêtre utilisée dès lors que la corrélation du signal est 30 fois plus longue que la fenêtre d'analyse. Les fenêtres de Hanning et Blackman restent intéressantes pour  $\lambda$  compris entre 0.03 et 3. Ici encore la divergence apparaît maximale avec la fenêtre rectangulaire.

Les différences observées entre les différentes fenêtres sont liées à leur comportement aux extrémités. Plus les valeurs aux extrémités se rapprochent de zéro et moins l'impact de la corrélation du signal sera important. On retrouve ainsi dans l'ordre d'efficacité croissante la fenêtre rectangulaire (pas d'atténuations aux bords), les fenêtres de Hamming et gaussiennes qui présentent une petite discontinuité puis les fenêtres de Blackman et Hanning dont les valeurs en bordure sont nulles.

### 4.2 Contexte de détection

Dans un test de détection, lorsqu'une loi de probabilité  $q$  est utilisée à la place de la loi de probabilité  $p$  pour décrire l'une des hypothèses, la différence entre les deux lois

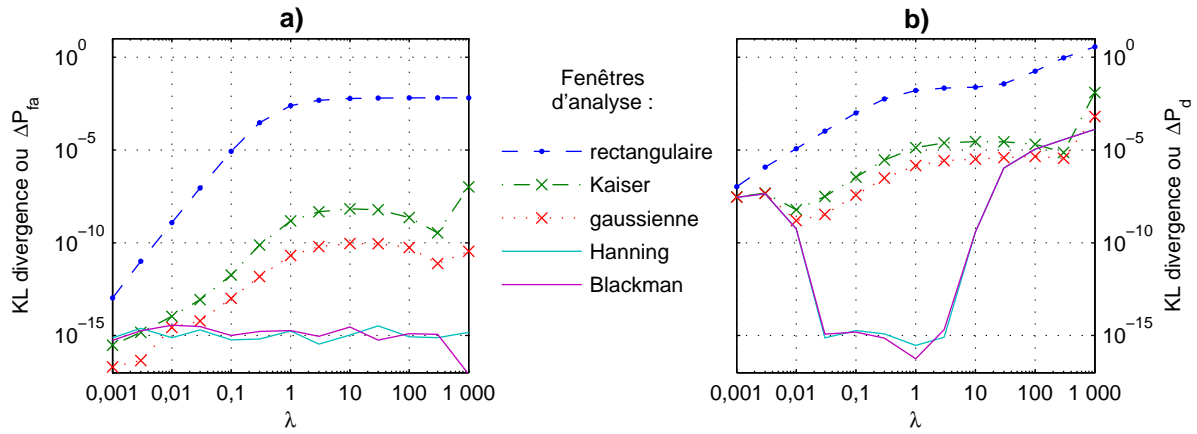


FIG. 2 – Divergence KL entre la loi du spectrogramme et un  $\chi^2$  à 2 degrés de liberté, en fonction de la corrélation du signal pour les cas a) centré (l’approximation porte sur l’hypothèse  $H_0$ ) et b) décentré (l’approximation porte sur l’hypothèse  $H_1$ ). Les distributions sont observées à la fréquence  $f_c/8$ .

entraîne un changement dans les performances du test. L’ampleur de ces changements détermine la validité pratique de l’approximation. Le résultat donné par Eguchi et Copas [5] est ici particulièrement intéressant car il permet d’interpréter les mesures de divergence présentées en section 4.1 en terme de variations moyennes des performances du détecteur.

Appliqué au cas de la détection dans le plan temps-fréquence, il se formule ainsi : la divergence KL entre la distribution du spectrogramme et celle d’un  $\chi^2$  mesure la perte moyenne de probabilité de fausse alarme  $\Delta P_{fa}$  lorsque la distribution du spectrogramme sous l’hypothèse  $H_0$  est approximée par une loi du  $\chi^2$ . Les mesures de divergence présentées sur la figure 2-a) correspondent donc à cette perte de probabilité de fausse alarme. De façon duale, lorsqu’une loi du  $\chi^2$  décrit le spectrogramme sous l’hypothèse  $H_1$ , la divergence KL mesure la perte moyenne  $\Delta P_d$  de probabilité de détection. La figure 2-b) se rapporte donc aux variations de la probabilité de détection du détecteur temps-fréquence.

Ceci nous permet de tirer quelques conclusions pratiques sur l’utilisabilité d’une loi du  $\chi^2$  à deux degrés de liberté pour décrire le spectrogramme dans un test de détection : les valeurs données par les figures 2-a) et 2-b) correspondent aux performances limites atteignables par le détecteur dans lequel des lois du  $\chi^2$  sont utilisées. Par exemple, lorsqu’une fenêtre rectangulaire est utilisée pour l’analyse d’un signal coloré, l’approximation des lois engendre une variation moyenne de la  $P_{fa}$  supérieure à  $10^{-5}$ . Une  $P_{fa}$  inférieure à ce seuil ne pourra dès lors pas être respectée. Aussi, il est à noter que l’impact sur les performances du test sont très faibles lorsque la fenêtre d’analyse présente une forte atténuation des bords. Pour les fenêtres de Hanning et Blackman par exemple, les variations de performance du test sont de l’ordre de  $10^{-15}$ .

## 5 Conclusion

Dans un environnement gaussien et corrélé, la loi de probabilité du spectrogramme est celle du module carré

d’un vecteur gaussien bidimensionnel. Face à une utilisation difficile de cette loi, nous avons cherché à approcher la distribution du spectrogramme par une loi du  $\chi^2$  à 2 degrés de liberté. À travers la divergence de Kullback-Leibler, l’évaluation de l’erreur d’approximation et son impact sur les performances d’un test de détection dans le plan temps-fréquence mène à la conclusion suivante : l’utilisation d’une fenêtre d’analyse dont les valeurs aux extrémités sont nulles (par exemple Hanning ou Blackman) permet de décrire ‘serenement’ (divergence KL ou variation de PFA égale à  $10^{-15}$ ) le spectrogramme avec une loi du  $\chi^2$  à deux degrés de liberté, quel que soit le degré de corrélation du signal analysé.

## Références

- [1] S. M. Kay, *Fundamentals of Statistical Signal Processing, vol.II - Detection Theory*. Prentice Hall, 1998.
- [2] J. Huillery, F. Millioz, and N. Martin, “On the probability distributions of spectrogram coefficients for correlated gaussian process,” *In proceedings of the ICASSP’06, Toulouse, France*, May 2006.
- [3] P. Johnson and D. Long, “The probability density of spectral estimates based on modified periodogram averages,” *IEEE Trans. on Signal Processing*, vol. 47, no. 5, pp. 1255–1261, May 1999.
- [4] T. Durrani and J. Nightingale, “Probability distributions for discrete fourier spectra,” *Proc. Inst. Elec. Eng.*, vol. 120, no. 2, pp. 299–311, 1973.
- [5] S. Egushi and J. Copas, “Interpreting kullback-leibler divergence with the neyman-pearson lemma,” *Journal of Multivariate Analysis*, vol. 97, pp. 2034–2040, 2006.
- [6] N. Johnson, S. Kotz, and Balakrishnan, *Continuous univariate distributions*, 2nd ed. Wiley and sons, 1995, vol. 2.
- [7] C. Hory and N. Martin, “Maximum likelihood noise estimation for spectrogram segmentation control,” *In Proceedings of the ICASSP’02, Orlando, USA*, pp. 1581–1584, 2002.