



HAL
open science

Support temps-fréquence d'un signal inconnu en présence de bruit additif gaussien

Julien Huillery

► **To cite this version:**

Julien Huillery. Support temps-fréquence d'un signal inconnu en présence de bruit additif gaussien. Traitement du signal et de l'image [eess.SP]. Institut National Polytechnique de Grenoble - INPG, 2008. Français. NNT: . tel-00301324

HAL Id: tel-00301324

<https://theses.hal.science/tel-00301324>

Submitted on 21 Jul 2008

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.

INSTITUT POLYTECHNIQUE DE GRENOBLE

N° attribué par la bibliothèque

THESE

pour obtenir le grade de

DOCTEUR DE L'INSTITUT POLYTECHNIQUE DE GRENOBLE

Spécialité : «Signal, Image, Parole, Télécoms»

préparée au Laboratoire GIPSA de Grenoble

dans le cadre de l'École Doctorale «Électronique, Électrotechnique, Automatique,
Télécommunications, Signal»

présentée et soutenue publiquement

par

Julien HUILLERY

le 09 Juillet 2008

Titre :

**Support temps-fréquence d'un signal inconnu
en présence de bruit additif gaussien**

Directrice de thèse : Nadine MARTIN

JURY

Monsieur	J.-L. Lacoume,	Président
Monsieur	L. Stankovic,	Rapporteur
Monsieur	B. Torresani,	Rapporteur
Monsieur	P. Flandrin,	Examineur
Monsieur	C. Richard,	Examineur
Madame	N. Martin,	Directrice de thèse

Table des matières

Introduction générale	1
1 Modèle, objectifs, méthode et outils	5
1.1 Introduction	7
1.2 Le modèle de signal en temps	7
1.2.1 Composition du modèle et informations <i>a priori</i>	7
1.2.2 Discussion	8
1.3 Travail proposé	8
1.3.1 Formulation	8
1.3.2 Motivations	11
1.3.3 Méthode proposée	12
1.4 Le modèle de signal en temps-fréquence	15
1.4.1 Choix d'une représentation temps-fréquence	15
1.4.2 Formulation temps-fréquence du modèle de signal	18
1.5 Résumé	19
2 Densité de probabilité du spectrogramme	21
2.1 Introduction	23
2.2 Densité de probabilité de la TFCT	24
2.3 Variable aléatoire du χ^2	25
2.3.1 Définition	25
2.3.2 Conditions pour le spectrogramme	26
2.4 (Non-)circularité de la TFCT	27
2.4.1 Observations	27
2.4.2 Circularité et forme de la fenêtre d'analyse	30
2.4.3 Conclusion	33
2.5 Densité de probabilité d'un module carré gaussien	33
2.5.1 Densité de probabilité dans le cas centré	33
2.5.2 Densité de probabilité dans le cas décentré	34
2.6 Le spectrogramme comme variable du χ^2	36
2.6.1 Quelle loi du χ^2 pour approcher le spectrogramme ?	37
2.6.2 Mesure de divergence	39
2.6.3 Impact de l'approximation dans un contexte de détection	43
2.7 Conclusion	47
3 Estimation de la densité spectrale de puissance du bruit	49
3.1 Introduction	51
3.2 Positionnement du problème	51
3.2.1 Les méthodes itératives	52
3.2.2 Les méthodes de lissage	52

3.2.3	Les statistiques minimales	53
3.3	Contribution : estimation par les Z valeurs minimales	54
3.3.1	Formulation	54
3.3.2	Estimateurs proposés	55
3.3.3	Évaluation des estimateurs	58
3.3.4	Détermination du voisinage et des paramètres Z et N	62
3.4	Exemples d'estimation dans le plan temps-fréquence	66
3.4.1	Perturbation blanche et stationnaire	67
3.4.2	Perturbation blanche ou stationnaire	69
3.4.3	Perturbation non blanche et non stationnaire	69
3.5	Conclusion	78
4	Détection du signal en un point du plan temps-fréquence	81
4.1	Introduction	83
4.2	Tests binaires d'hypothèses	83
4.2.1	Approche bayésienne	84
4.2.2	Approche de Neyman-Pearson	86
4.3	Critères de détection retenus et statistique de décision	87
4.3.1	Critère de Neyman-Pearson	88
4.3.2	Critère du maximum de vraisemblance	89
4.3.3	Rapport signal sur bruit de focus	91
4.3.4	Statistique de décision	92
4.4	Évaluation du détecteur	94
4.4.1	Courbes de performances dans le cas clairvoyant	94
4.4.2	Non optimalité et performances réelles	96
4.5	Conclusion	98
5	Localisation du signal dans le plan temps-fréquence	101
5.1	Introduction	103
5.2	Réduction des fausses alarmes	103
5.2.1	Prise en compte de la redondance du spectrogramme	104
5.2.2	Prise en compte du rsb de focus	105
5.2.3	Compromis fausses alarmes / signal	107
5.3	Algorithme de localisation	107
5.4	Évaluation de l'algorithme	111
5.4.1	Dépendance au rsb de focus	111
5.4.2	Dépendance à l'estimation du bruit	113
5.4.3	Limitations	118
5.4.4	Cas d'un signal d'intérêt gaussien	120
5.5	Conclusion	125
	Conclusions et perspectives	127
	Bibliographie	131
	Remerciements	135

Introduction générale

La thématique du travail rapporté dans ce manuscrit est la recherche d'information à partir d'un signal. La recherche d'information vise à acquérir de la connaissance vis-à-vis du ou des phénomènes qui engendrent le signal observé. Elle consiste à mettre à jour un certain nombre de caractéristiques qui seront utiles pour décrire, comprendre, expliquer ou encore discerner ces phénomènes sous-jacents. La recherche d'information est soumise à deux éléments : le mode de représentation du signal d'une part et l'extraction de l'information proprement dite d'autre part.

Le mode de représentation définit l'aspect sous lequel le signal est disponible. Dans une représentation temporelle, l'amplitude ou la puissance du signal est disponible en fonction du temps. Une représentation spatiale décrit l'amplitude du signal en différents points de l'espace. De la même manière, une représentation fréquentielle présente l'énergie du signal suivant les fréquences. Chaque espace de représentation permet ainsi de décrire le contenu du signal en fonction d'une variable d'évolution : le temps, l'espace ou la fréquence dans les exemples cités précédemment. À travers cette variable d'évolution, il apparaît que chaque espace de représentation met en valeur un type d'information : évolution dans le temps, répartition dans l'espace, distribution selon les fréquences... La recherche d'information est donc intimement liée aux possibilités offertes par l'espace dans lequel le signal est représenté. L'information recherchée doit en premier lieu être accessible dans cet espace.

Dans le cas contraire, il devient nécessaire de réorganiser le signal afin de mettre en évidence l'information souhaitée. Nous opérons ainsi un changement d'espace de représentation. Par exemple, nous pouvons passer d'une représentation temporelle du signal vers une représentation fréquentielle, une représentation temps-fréquence ou encore un espace des phases, etc. Précisons que changer d'espace de représentation ne modifie pas le contenu du signal et notamment ne crée pas d'information. Seul l'angle d'observation est modifié. L'information recherchée doit être initialement présente dans le signal et le choix d'un espace de représentation doit simplement permettre de la mettre en évidence. Précisons également que l'information contenue dans un signal est limitée. Plus précisément, l'espace d'acquisition initial dans lequel un phénomène est observé définit la nature des informations auxquelles il sera possible d'accéder. Par exemple, une représentation spatiale ne peut être obtenue à partir d'un enregistrement du signal sous forme temporelle.

Le second élément auquel est soumise la recherche d'information est son extraction proprement dite. Si l'information recherchée est accessible dans un espace de représentation correctement choisi, elle doit cependant être récupérée afin d'être utilisée par la suite. La mise en œuvre de cette extraction est alors contrainte par le degré d'accessibilité de l'information. Par exemple, est-il possible d'extraire l'information sans commettre d'erreur ? Toute l'information est-elle accessible ? Sur ce point, notons que le rapport signal sur bruit ou encore le degré de parcimonie sont des mesures permettant de quantifier ce degré d'accessibilité. Il découle de cette remarque que la méthode d'extraction à mettre en œuvre afin de récupérer l'information souhaitée est intimement liée à l'espace de représentation choisi, et notamment à sa faculté à rendre cette information accessible. Comme nous le verrons par la suite, c'est sur ce second aspect de la recherche d'information que se focalise le travail présenté dans ce manuscrit.

Objectif et contexte de travail :

Nous considérons les signaux pour lesquels une représentation temporelle est disponible. L'analyse de Fourier d'un signal temporel permet de mettre à jour la répartition de son énergie suivant ses fréquences d'oscillation. Le caractère physique et intuitif de la notion de fréquence a engendré le succès de cette analyse dans les domaines d'application les plus variés. Nous interprétons ce succès par le caractère informatif de la représentation fréquentielle d'un signal. Cependant, en révélant le contenu fréquentiel du signal, l'analyse de Fourier dissimule son caractère temporel et ne permet pas la description de l'évolution du signal dans le temps. Or, si les fréquences d'oscillations qui apparaissent dans un signal constituent une information pertinente, l'évolution dans le temps de ces fréquences, lorsqu'elle existe, apporte une information plus complète sur le phénomène étudié. Afin de rendre possible une description fréquentielle du signal tout en gardant l'information liée à son évolution dans le temps, des représentations temps-fréquence du signal ont été proposées [Fla93][Coh95]. Ces représentations visent à décrire la répartition de l'énergie du signal conjointement suivant le temps et la fréquence.

Les différentes possibilités offertes par plus d'un demi-siècle de recherche dans le but de représenter un signal dans le plan temps-fréquence ne sont pas le point central de ce travail. Dans la suite, nous ne considérerons qu'une seule représentation du signal dans l'espace temps-fréquence : le spectrogramme. Ce choix sera discuté dans le premier chapitre. Partant de cette représentation, nous nous sommes plutôt attachés à y décrire le comportement du signal. Le descripteur que nous avons retenu est le support du signal dans cet espace. Le support temps-fréquence d'un signal correspond à l'ensemble des régions du plan temps-fréquence sur lesquelles le signal existe. La détermination de ce support temps-fréquence correspond à l'objectif de notre travail.

Si chaque signal admet nécessairement un support dans le plan temps-fréquence, notons qu'une infinité de signaux peuvent admettre le même support. L'information de support n'est donc pas suffisante pour décrire complètement un signal. La manière dont l'énergie du signal s'y répartit en est effectivement absente. Sur ce point, la pertinence de cette information sera donnée par l'ensemble des problèmes pour lesquels elle apportera une solution suffisante. Aussi, notons que la détermination du support temps-fréquence du signal ne nécessite pas la perte de l'information énergétique qui reste dans tous les cas disponible.

Sans rentrer plus avant dans une méthode de détermination du support temps-fréquence, il apparaît dès à présent qu'un processus de décision doit être mis en place. Un support temps-fréquence est une information de nature binaire qui peut être représentée par une fonction indicatrice. À chaque point du plan temps-fréquence doit donc être associée la valeur 0 ou 1 en fonction de son appartenance au support du signal. Cette association est nécessairement le résultat d'une décision.

Le support temps-fréquence d'un signal correspond donc à l'information que nous nous proposons de rechercher dans le contexte spécifique que nous présentons maintenant.

D'une part, le signal d'intérêt sur lequel nous souhaitons gagner de l'information est entièrement inconnu. Il ne sera donc pas possible de formuler quelques hypothèses *a priori* sur son comportement dans le plan temps-fréquence, et notamment sur la forme ou la localisation de son support.

D'autre part, nous considérons que le signal temporel observé, c'est-à-dire le signal auquel nous avons accès, est une version perturbée du signal d'intérêt. Nous considérons que cette perturbation prend la forme d'un bruit gaussien additif dont le comportement temporel ainsi que fréquentiel est inconnu. Nous verrons que la présence de ce bruit doit être explicitement prise en compte dans la détermination du support temps-fréquence du signal d'intérêt. Notamment, elle

nécessite de prendre en compte les participations respectives du signal d'intérêt et du bruit dans les différentes régions du plan temps-fréquence. Aussi, la présence d'un bruit aléatoire impose un caractère incertain au processus de décision qui sera mis en place. Il sera alors nécessaire d'évaluer le degré d'incertitude qui accompagne la détermination du support temps-fréquence.

Organisation du manuscrit :

Dans le premier chapitre, nous posons les bases du travail qui a été mené. En point de départ, nous donnons la nature et la composition des signaux qui sont étudiés. Nous définissons un modèle temporel de signal pour lequel nous précisons les informations disponibles *a priori*. Dans un second temps, nous précisons la notion de support temps-fréquence dans le contexte d'un signal bruité. Aussi, la méthode d'extraction de cette information est introduite. Dans un troisième temps, quelques éléments sur les représentations temps-fréquence sont présentés afin de justifier le choix réalisé dans ce travail d'utiliser le spectrogramme. Enfin, les éléments du modèle de signal précédemment établi dans le domaine temporel sont traduits dans ce nouvel espace de représentation.

De par la présence du bruit, les échantillons temporels du signal observé ainsi que les coefficients du spectrogramme de ce même signal sont de nature aléatoire. Le second chapitre présente une étude menée sur le comportement statistique des coefficients du spectrogramme des signaux considérés. Ce comportement statistique caractérise d'une part l'incertitude qui doit être prise en compte dans la détermination du support temps-fréquence du signal d'intérêt. D'autre part, l'entière dépendance au bruit de ce comportement représente une porte d'entrée pour la caractérisation de ce bruit.

L'étude rapportée dans ce chapitre est la suivante. Dans le cas particulier d'un signal gaussien, blanc et stationnaire, les coefficients du spectrogramme sont distribués selon une loi du χ^2 à 2 degrés de liberté. Afin de nous replacer dans le contexte d'un bruit gaussien inconnu, nous évaluons si cette proposition reste vraie lorsque les hypothèses de blancheur et/ou de stationnarité sont relâchées. Ayant mis à jour les conditions pour lesquelles cette proposition n'est plus vraie, la densité de probabilité du spectrogramme est ensuite formulée dans le contexte d'un bruit gaussien non blanc et non stationnaire. Devant la complexité de la loi obtenue, nous mesurons sa divergence avec une loi du χ^2 afin d'évaluer les possibilités d'utiliser une loi du χ^2 pour décrire le spectrogramme. Notamment, nous mesurons l'impact de l'utilisation d'une loi du χ^2 dans le cadre d'un test binaire d'hypothèses.

La connaissance du comportement dans le plan temps-fréquence du bruit gaussien additif subordonne la possibilité de déterminer le support temps-fréquence du signal d'intérêt. Le troisième chapitre est consacré à l'estimation de ce comportement. À partir d'une réalisation du signal, l'objectif est de connaître l'espérance du spectrogramme du bruit. Dans ce but, un estimateur basé sur les plus petits coefficients observés dans une région prédéfinie du spectrogramme est développé. Cet estimateur est ensuite évalué sur des signaux réels et synthétiques.

Dans le quatrième chapitre, le processus de décision permettant de sélectionner les points du plan temps-fréquence qui appartiennent au support du signal d'intérêt est étudié au niveau d'un unique point temps-fréquence. Celui-ci prend la forme d'un test binaire d'hypothèses permettant de décider si l'énergie observée en un point du plan temps-fréquence est uniquement due au bruit additif ou bien révèle l'existence du signal d'intérêt. Notamment, un critère de rapport signal sur bruit est introduit dans cette prise de décision. Les performances du détecteur mis en place sont ensuite évaluées. De façon rétroactive, les performances obtenues caractérisent la représentation temps-fréquence choisie, c'est-à-dire le spectrogramme, pour la détection d'un signal inconnu en

présence de bruit gaussien additif.

Dans le cinquième et dernier chapitre, le processus de décision est ramené au niveau de l'ensemble de la représentation temps-fréquence afin de déterminer le support du signal d'intérêt. L'algorithme complet permettant l'extraction de cette information est présenté. L'impact et le choix des paramètres qui seront laissés libre à un utilisateur sont discutés, au même titre que les défauts et les limitations de la méthode.

Les conclusions de ce travail ainsi que quelques perspectives viennent finalement clore ce manuscrit.

Chapitre 1

Modèle, objectifs, méthode et outils

Sommaire

1.1	Introduction	7
1.2	Le modèle de signal en temps	7
1.2.1	Composition du modèle et informations <i>a priori</i>	7
1.2.2	Discussion	8
1.3	Travail proposé	8
1.3.1	Formulation	8
1.3.2	Motivations	11
1.3.2.1	Filtrage temps-fréquence	11
1.3.2.2	Caractérisation du signal	11
1.3.3	Méthode proposée	12
1.4	Le modèle de signal en temps-fréquence	15
1.4.1	Choix d'une représentation temps-fréquence	15
1.4.2	Formulation temps-fréquence du modèle de signal	18
1.5	Résumé	19

1.1 Introduction

Le rôle de ce premier chapitre est de présenter le modèle de signal, les objectifs et la méthode choisie, ainsi que les outils utilisés dans ce travail. Il en constitue donc une introduction technique et méthodologique.

Dans le premier paragraphe, nous présentons le modèle de signal sur lequel l'ensemble du travail rapporté dans ce manuscrit est fondé. Les choix réalisés sur le type et la nature des signaux qui seront pris en compte sont ainsi exposés. Également, nous précisons l'ensemble des informations que nous supposons connues *a priori*. Dans un second temps, nous précisons l'objectif et la motivation générale de ce travail. Nous donnons alors la démarche méthodologique que nous avons adoptée : la détermination d'un support temps-fréquence est formulée en tant que test statistique d'hypothèses dans le plan temps-fréquence. Le modèle de signal étant défini dans le domaine temporel, la troisième partie de ce chapitre concerne alors sa formulation dans le domaine temps-fréquence. Notamment, le choix du spectrogramme comme outil de représentation temps-fréquence est argumenté. Finalement, nous résumons les objectifs et définissons le travail qui sera présenté dans les chapitres suivants.

1.2 Le modèle de signal en temps

1.2.1 Composition du modèle et informations *a priori*

Dans tout ce travail, nous considérons des signaux réels formés par l'addition d'un signal déterministe et d'un signal aléatoire, indépendants l'un de l'autre. Le signal observé, noté $x[m]$ avec m désignant la variable temporelle discrète, s'écrit sous la forme

$$x[m] = d[m] + p[m], \quad (1.1)$$

avec $d[m]$ le signal déterministe et $p[m]$ la contribution aléatoire.

Le signal déterministe $d[m]$ est supposé non stationnaire au sens où la répartition de son énergie spectrale évolue avec le temps. Concernant le signal aléatoire $p[m]$, sa fonction d'autocorrélation, notée $\Gamma_p[m, \tau]$ à l'instant m et pour le décalage temporel τ , est définie selon

$$\Gamma_p[m, \tau] = \mathbb{E} \{p[m - \tau]p[m + \tau]\}. \quad (1.2)$$

C'est une fonction paire selon la variable de décalage temporel τ . Notons que la dépendance de la fonction d'autocorrélation à l'instant m traduit la non-stationnarité, tandis que la dépendance au décalage temporel τ caractérise la non-blanchéur du signal¹. La matrice de corrélation associée à cette fonction d'autocorrélation sera notée \mathbf{R}_p .

Nous choisissons de connaître la nature du processus aléatoire $p[m]$. Dans ce travail, la perturbation $p[m]$ est un processus gaussien, centré et inconnu au second ordre, soit

$$p[m] \sim \mathcal{N}(0, \Gamma_p[m, \tau]). \quad (1.3)$$

Le modèle aléatoire des signaux observés retenu pour ce travail correspond finalement à un processus gaussien, non stationnaire aux ordres 1 et 2, centré sur le signal déterministe $d[m]$, et de fonction d'autocorrélation $\Gamma_p[m, \tau]$. Ce modèle se résume sous l'écriture

$$x[m] \sim \mathcal{N}(d[m], \Gamma_p[m, \tau]). \quad (1.4)$$

¹Par définition, la fonction d'autocorrélation d'un signal blanc est nulle pour tous les décalages temporels excepté en $\tau = 0$

1.2.2 Discussion

a) Le traitement du signal a une grande tradition de modèles dits "signal + bruit". Le signal correspond alors à la composante informative qu'il faut généralement extraire du bruit, celui-ci représentant une perturbation physique ou plus symboliquement une dégradation de l'information. Cette perturbation peut souvent être modélisée par une contribution additive de nature aléatoire. Le modèle adopté dans ce travail est un cas particulier dans cette catégorie très générale de modèle "signal + bruit".

b) Le cas particulier évoqué ci-dessus provient du caractère déterministe attribué à la composante "signal". Ce choix peut s'apparenter à une considération "petite échelle" de l'observation $x[m]$, pour laquelle aucune hypothèse n'est réalisée sur la composante "signal". À l'échelle de l'échantillon, l'observation est formée d'une contribution $d[m]$ inconnue, aléatoirement perturbée par une variable gaussienne $p[m]$. Sans autre information, l'échantillon $d[m]$ ne peut être considéré que déterministe. Dans une vision à plus grande échelle, quelques caractéristiques statistiques et notamment un caractère centré peuvent rapidement venir décrire le signal. Maintenant, si le signal d'intérêt $d[m]$ présente un comportement aléatoire différent de celui porté par la composante $p[m]$, celui-ci ne sera tout simplement pas pris en compte. Notamment, ce comportement serait entièrement préservé si une reconstruction exacte du signal $d[m]$ était possible.

c) Le choix de donner au signal d'intérêt $d[m]$ une nature déterministe, ou un caractère inconnu, donne une grande importance aux hypothèses placées dans le comportement aléatoire du bruit : le signal sera reconnu comme ce qui ne correspond pas à ces hypothèses. La seule action possible correspond à identifier dans un signal donné ce qui ne correspond 'probablement' pas à la tendance aléatoire spécifiée par les hypothèses. Par exemple, faire une hypothèse de blancheur ou de stationnarité sur la composante aléatoire $p[m]$ caractérise le signal $d[m]$ comme ce qui n'est pas blanc ou stationnaire.

d) Le modèle gaussien choisi pour la composante aléatoire limite bien entendu la nature des bruits qui pourront être pris en compte. Les bruits de nature impulsive par exemple, présents dans les enregistrements de signaux atmosphériques ou sous-marins et modélisés par des lois α -stables [NS95] n'entrent pas dans le cadre de ce travail.

e) Pour rester familier aux habitudes de langage propres au traitement du signal, les dénominations "signal" ou "signal d'intérêt" ainsi que "bruit" ou "perturbation" seront communément employées en référence aux composantes déterministe $d[m]$ et aléatoire $p[m]$ respectivement. Précisons cependant que le signal $d[m]$ n'a pas nécessairement l'exclusivité du caractère informatif. Par exemple, l'estimation du comportement au second ordre de la perturbation $p[m]$, sous des hypothèses de non-blancher et/ou de non-stationnarité, permet une description du milieu environnant qui peut s'avérer utile.

1.3 Travail proposé

1.3.1 Formulation

Étant donné un enregistrement bruité $x[m]$ répondant au modèle (1.4) décrit ci-dessus, nous sommes à la recherche d'informations concernant le signal déterministe $d[m]$. La non-stationnarité de ce signal et de son environnement (le bruit) nous montre que l'information disponible évolue au cours du temps et que sa recherche doit prendre en compte cette évolution. Comme second élément de description, nous choisissons les fréquences au sens de Fourier, c'est à dire données par les oscillations des fonctions sinus et cosinus. Ce choix correspond à observer les signaux

dans le plan temps-fréquence, et plus précisément à observer la répartition de l'énergie du signal conjointement suivant le temps et la fréquence. De nombreuses fonctions mathématiques ont été proposées afin de permettre cette observation [Coh95]. Le choix d'une représentation temps-fréquence adaptée au travail présenté dans ce manuscrit est discuté dans le paragraphe 1.4.1.

Le point de départ de notre étude est ainsi une représentation temps-fréquence de l'observation bruitée $x[m]$. Cette représentation est notée $S_x[n, k]$. Dans tout le manuscrit, n désigne l'indice de temps et k l'indice des fréquences associés au plan temps-fréquence. À partir de cette représentation, l'information que nous nous proposons de retrouver est le "support temps-fréquence" du signal déterministe $d[m]$. Cette recherche s'apparente à localiser l'énergie du signal $d[m]$ dans le plan temps-fréquence ou encore à déterminer l'ensemble des points temps-fréquence sur lesquels ce signal est présent². Le support temps-fréquence d'un signal correspond donc à une région ou à un ensemble de régions du plan temps-fréquence. Il peut être représenté par la fonction indicatrice de cette(ces) région(s), c'est-à-dire la carte binaire sur laquelle la valeur 1 est attribuée aux points temps-fréquence appartenant à cette(ces) région(s), la valeur 0 étant attribuée dans le cas contraire. À titre d'exemple, la figure 1.1 présente les spectrogrammes et les supports temps-fréquence relatifs à un signal de parole (en haut), un signal synthétique composé de deux modulations de fréquence (au milieu) ainsi qu'un signal acoustique après propagation dans un milieu marin dispersif (en bas).

Nous avons précisé dans le premier paragraphe de ce chapitre que le signal $d[m]$ est entièrement inconnu. Il en découle qu'aucune hypothèse restrictive sur la forme de son support temps-fréquence ne peut être formulée. Tout point temps-fréquence $[n, k]$ a ainsi une probabilité *a priori* d'appartenir au support TF du signal égale à 1/2 et cette probabilité n'est pas modifiée à mesure que ce support devient connu dans une quelconque autre région du plan temps-fréquence. L'hypothèse d'un signal $d[m]$ inconnu impose de ce fait une approche locale : pour chaque point $[n, k]$ du plan temps-fréquence, son appartenance au support temps-fréquence du signal $d[m]$ doit être remise en question.

Remarque : la non-connaissance du phénomène observé doit d'abord être perçue comme une contrainte imposée à la méthode d'analyse, et non comme le reflet général d'une réalité pratique : il est souvent possible d'avoir quelques idées *a priori* sur les signaux analysés. Cependant, si la méthode a pour objectif une certaine adaptabilité à un grand nombre de signaux, elle doit, au départ, réduire les hypothèses sur la forme (spectrale par exemple) et la nature (déterministe/aléatoire) des signaux auxquels elle s'adresse. Une question se pose néanmoins : où est l'intérêt d'une méthode générale lorsque l'apport de connaissances sur le signal observé engendre des traitements certes spécifiques mais certainement beaucoup plus performants ? Dans un premier temps, nous répondons que la connaissance *a priori* n'est pas toujours suffisamment précise pour être directement utilisée par un algorithme. De plus, elle n'existe pas toujours exactement sous la forme nécessaire à l'algorithme. De ce point de vue, nous tentons de mettre en place un traitement qui permette d'obtenir automatiquement, rapidement et de façon suffisamment fiable une première information sur un signal, en l'occurrence son support dans le plan temps-fréquence et comme nous le verrons, une estimation du bruit. Cette information pourra ensuite être utilisée pour des traitements plus spécifiques. Dans une perspective d'évolution, un intérêt d'une méthode générale repose par contre sur sa capacité à supporter et même encourager l'intégration de connaissances sur le phénomène analysé. Lorsque certains détails sur le signal d'intérêt (par exemple une forme harmonique ou une nature aléatoire) ou sur la perturbation sont connus, la méthode doit permettre leur intégration dans le modèle et accroître ses performances sur les signaux considérés.

²Cette dernière définition sera modulée dans la suite en raison de la présence de bruit.

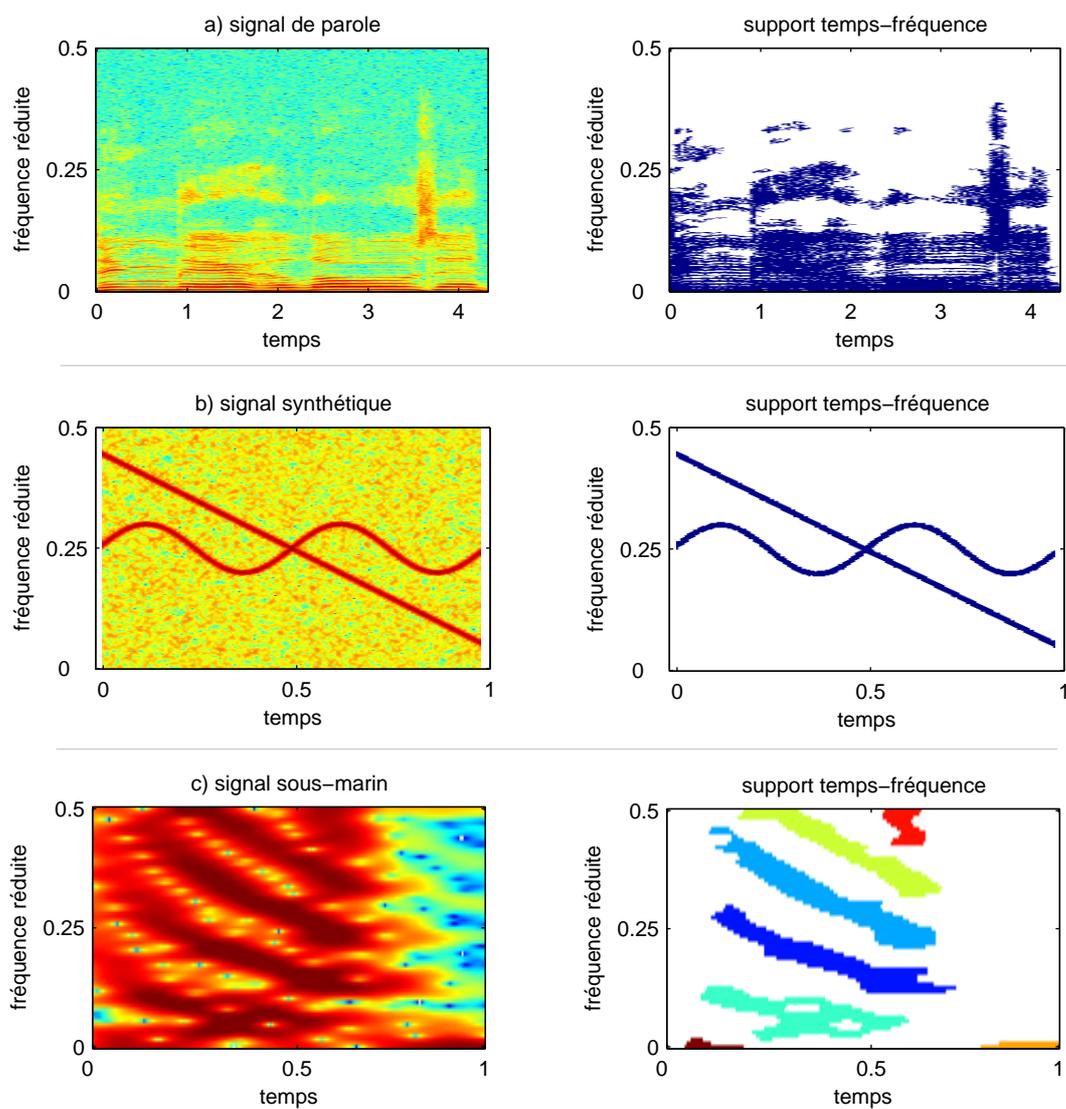


FIG. 1.1 – Exemples de représentations temps-fréquence et de supports temps-fréquence associés pour a) un signal de parole, b) un signal synthétique et c) un signal d'acoustique sous-marine.

1.3.2 Motivations

Nous avons relevé diverses problématiques de traitement du signal pour lesquelles la détermination du support temps-fréquence d'un signal présente un intérêt. Nous en citons deux grandes catégories qui sont le filtrage temps-fréquence et la détermination de paramètres caractéristiques d'un signal. Ces deux catégories correspondent respectivement à la transformation des signaux d'une part et à la recherche d'informations d'autre part.

1.3.2.1 Filtrage temps-fréquence

Plusieurs études ont montré que les représentations temps-fréquence sont des outils performants et très intuitifs pour la réalisation de filtre linéaire et variant dans le temps. La prise en compte de la dimension temporelle dans l'opération de filtrage est une caractéristique nécessaire lorsque les signaux sont de nature non stationnaire.

Une opération de filtrage temps-fréquence, basée sur le support temps-fréquence d'un signal, permet par exemple de réaliser un *débruitage* du signal. L'énergie de la perturbation additive qui n'appartient pas au support TF du signal peut effectivement être éliminée. Notons cependant que le support temps-fréquence du signal, seul, n'est pas une information suffisamment riche pour éliminer ou même atténuer le bruit contenu à l'intérieur même de ce support. Ce dernier point nécessite l'estimation de l'énergie propre au signal d'intérêt au sein du support. Une seconde application du filtrage temps-fréquence est possible lorsque le support du signal se compose de plusieurs régions distinctes dans le plan temps-fréquence. Il s'agit alors de construire le signal temporel correspondant à une partie seulement du support. Cette opération s'assimile à une *séparation de composantes disjointes dans le plan temps-fréquence*. La figure 1.1-c) donne un exemple de signal pour lequel cette opération est possible.

Nous mentionnons deux approches permettant de réaliser un filtre temps-fréquence. La première prend la forme générique d'une analyse-pondération-synthèse. Une représentation temps-fréquence inversible du signal est construite puis pondérée terme à terme par une fonction de gain (temps-fréquence). Une opération de synthèse temps-fréquence permet alors de reconstruire dans le domaine temporel le signal ainsi filtré. Cette technique est particulièrement utilisée pour les signaux audio sur la base de la Transformée de Fourier à Court Terme (des exemples de fonction de gain sont donnés dans [MM80, EM84], la question de la synthèse est traitée dans [All77, AR77]). Dans cette approche, la connaissance du support temps-fréquence du signal d'intérêt peut être une aide utile à la détermination de la fonction de gain temps-fréquence. La seconde approche permettant de définir un filtre linéaire et variant dans le temps, particulièrement développée par F. Hlawatsch [HK94, HMKK00], repose sur la notion de sous-espace signal dans le plan temps-fréquence. Elle consiste à projeter le signal sur une base de fonctions décrivant une région prédéfinie du plan temps-fréquence. L'opération a ici l'avantage d'être réalisée dans le domaine temporel tout en gardant une représentation intuitive liée au plan temps-fréquence. Pour cette méthode, la détermination du support temps-fréquence du signal d'intérêt donne explicitement la ou les régions du plan temps-fréquence sur lesquelles le signal doit être projeté.

1.3.2.2 Caractérisation du signal

Le traitement du signal porte une tradition d'extraction de l'information. Une partie de son activité est ainsi dédiée à la définition d'espaces de représentation dans lesquels l'information est accessible. À chaque espace de représentation correspond alors un type d'information. Parmi d'autres, les représentations temps-fréquence proposent de décrire un phénomène conjointement en temps et en fréquence. Le temps décrit la non-stationnarité, la fréquence décrit une nature oscillante. Si la lecture d'une représentation temps-fréquence est un élément pertinent quant à

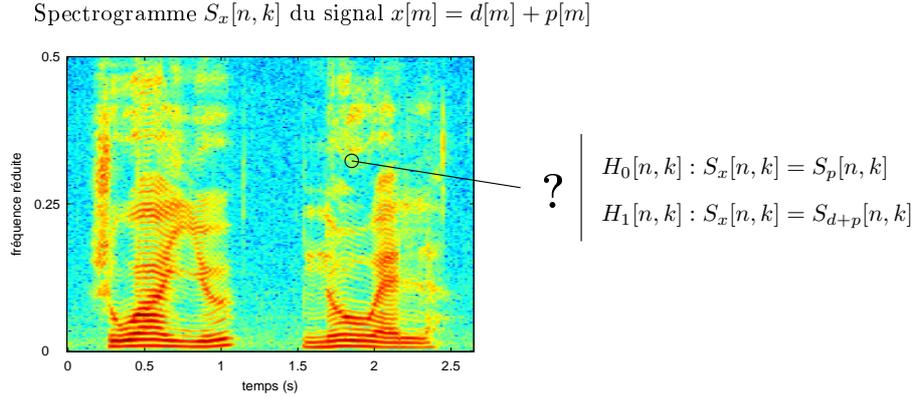


FIG. 1.2 – Principe du test binaire d'hypothèses mis en place en chaque point $[n, k]$ du plan temps-fréquence.

la connaissance d'un phénomène, l'emplacement dans le plan temps-fréquence de ce qu'on y lit contient une portion de cette information. Le support temps-fréquence d'un signal revêt donc dans certains cas un caractère informatif qui suggère son utilisation comme une caractéristique du signal. L'exemple typique d'une information particulièrement recherchée et qui s'apparente à un support temps-fréquence est la loi de fréquence d'un signal modulé en fréquence (voir par exemple le signal synthétique et son support TF présentés sur la figure 1.1-b)). L'utilisation de cette information dans une perspective de diagnostique est présentée dans [HM06b]. Le support temps-fréquence d'un signal peut également définir une caractéristique utilisable dans un problème de classification qui s'apparente alors à une reconnaissance de forme.

Toutefois, si ces perspectives paraissent attrayantes, elles sont conditionnées à la manière dont ce support temps-fréquence pourra être déterminé, et notamment à la possibilité de prendre en compte la perturbation additive $p[m]$, comme nous le verrons par la suite.

1.3.3 Méthode proposée

Afin de déterminer le support temps-fréquence du signal $d[m]$, nous choisissons de formuler, en chaque point $[n, k]$ du plan temps-fréquence, un test d'hypothèses binaires permettant de lui attribuer la valeur 0 ou 1. Plus précisément, étant donné une représentation temps-fréquence de l'observation $x[m]$, il s'agit de déterminer si le coefficient $S_x[n, k]$ observé à la position $[n, k]$ provient uniquement de la perturbation $p[m]$ (hypothèse H_0 dite "bruit seul") ou s'il reflète aussi la contribution du signal d'intérêt $d[m]$ (hypothèse H_1 dite "signal + bruit"). Illustré sur la figure 1.2, le test d'hypothèses au point $[n, k]$ s'écrit, en utilisant les notations introduites en (1.1),

$$\begin{cases} H_0 : S_x[n, k] = S_p[n, k], \\ H_1 : S_x[n, k] = S_{d+p}[n, k]. \end{cases} \quad (1.5)$$

Le support temps-fréquence du signal d'intérêt $d[m]$, noté $\mathcal{S}^{tf}(d)$, correspond à l'ensemble des positions $[n, k]$ pour lesquelles l'hypothèse H_1 est retenue.

Une manière intuitive de sélectionner les points temps-fréquence qui constituent le support d'un signal consiste à comparer le niveau d'énergie observé avec une référence prédéfinie. En un point du plan temps-fréquence, retenir l'une des hypothèses H_0 ou H_1 s'assimile alors à comparer le coefficient temps-fréquence $S_x[n, k]$ avec un seuil, noté $S^{\text{seuil}}[n, k]$, et le support temps-fréquence du signal $d[m]$ est obtenu selon

$$\mathcal{S}^{tf}(d) = \{[n, k] / S_x[n, k] > S^{\text{seuil}}[n, k]\}. \quad (1.6)$$

Suivant cette démarche et pour une représentation temps-fréquence $S_x[n, k]$ donnée, il apparaît que le support temps-fréquence $\mathcal{S}^{tf}(d)$ du signal $d[m]$ est entièrement déterminé par le choix du seuil $S^{\text{seuil}}[n, k]$. Ce "seuil de détection" est donc le seul inconnu du problème et l'objectif principal de ce travail est de proposer et évaluer une façon de le définir.

Dans la manière de formuler les deux hypothèses H_0 et H_1 , nous avons supposé que le bruit $p[m]$ et plus précisément sa représentation temps-fréquence $S_p[n, k]$ est non nulle sur tout le plan temps-fréquence. Notamment, l'hypothèse "signal seul" n'est pas formulée. Ce choix peut se justifier en considérant que l'enregistrement d'un signal dans un environnement réel contient à tous les instants un bruit, au minimum le bruit thermique lié au capteur, approximativement blanc et admettant ainsi une énergie non nulle sur l'ensemble du domaine fréquentiel.

Mais cette formulation possède surtout l'intérêt de donner explicitement au bruit le rôle de "référence" à partir de laquelle le signal sera détecté. Il paraît intuitif de considérer qu'un signal est accessible en un point du plan temps-fréquence lorsque l'énergie qu'il y développe est au moins supérieure à celle que le bruit présente en ce même point. Le signal accessible est ainsi celui qui *émerge* du bruit. De la même manière, si nous avons plus haut présenté le support temps-fréquence d'un signal comme l'ensemble des points temps-fréquence sur lesquels le signal *existe*, le support temps-fréquence *accessible* du signal d'intérêt correspond quant à lui à l'ensemble des points temps-fréquence pour lesquels le signal est lui-même accessible, c'est-à-dire admet une énergie au moins supérieure à celle du bruit. Nous voyons donc que le bruit tient une place tout à fait privilégiée, au sens où il détermine l'énergie minimale que doit posséder le signal pour être accessible. Il en découle aussi que la sélection des points temps-fréquence appartenant au support du signal se doit d'être plus liée à l'évaluation d'un rapport signal sur bruit qu'à la seule énergie du signal lui-même.

La figure 1.3 illustre quelques éléments de cette discussion. Pour chacun des trois cas a), b) et c), la figure de gauche représente le spectrogramme d'un signal synthétique auquel un bruit gaussien a été ajouté. À droite, deux possibilités de supports temps-fréquence sont présentées. Le premier, à gauche, ne prend pas en compte les éventuelles variations du bruit. Le seuil de décision est obtenu suivant un critère énergétique global³ appliqué sur la représentation temps-fréquence $S_x[n, k]$ du signal bruité $x[m]$. Il correspond donc à une "référence signal". Par opposition, le second support temps-fréquence, à droite, est obtenu en prenant explicitement le bruit comme référence. Dans ce cas, les points temps-fréquence sélectionnés admettent une énergie au moins supérieure à celle du bruit, celle-ci étant connue dans cet exemple.

Dans le cas d'une perturbation blanche et stationnaire, c'est-à-dire constante (en espérance) sur tout le plan temps-fréquence, les supports obtenus suivant l'une ou l'autre des références sont identiques, figure 1.3-a). La nécessité de prendre le bruit comme référence apparaît lorsque celui-ci est non stationnaire (figure 1.3-b)) ou non blanc (figure 1.3-c)), c'est-à-dire non constant sur tout le plan temps-fréquence. Dans ces situations, le support temps-fréquence du signal ne peut être extrait justement si la forme du bruit n'est pas prise en compte. Aussi, la figure 1.3-c) illustre la notion de support temps-fréquence *accessible*. Pour ce troisième cas, la perturbation non blanche recouvre certaines parties du signal. Celles-ci sont alors perdues et nous ne chercherons pas à les retrouver. Le support temps-fréquence de droite, obtenu avec une référence bruit, correspond alors au support accessible du signal.

Il ressort donc que le test de détection doit d'une part être formulé localement (pour chaque position $[n, k]$) et d'autre part que le seuil de détection doit être défini en relation avec le niveau de bruit au point $[n, k]$ considéré. Ces deux conditions sont naturellement prises en compte dans

³par exemple retenir les valeurs supérieures à $1/10^{\text{ième}}$ du maximum, ne garder que 20% de l'énergie totale, ou encore évaluer le maximum de l'histogramme de la représentation temps-fréquence...

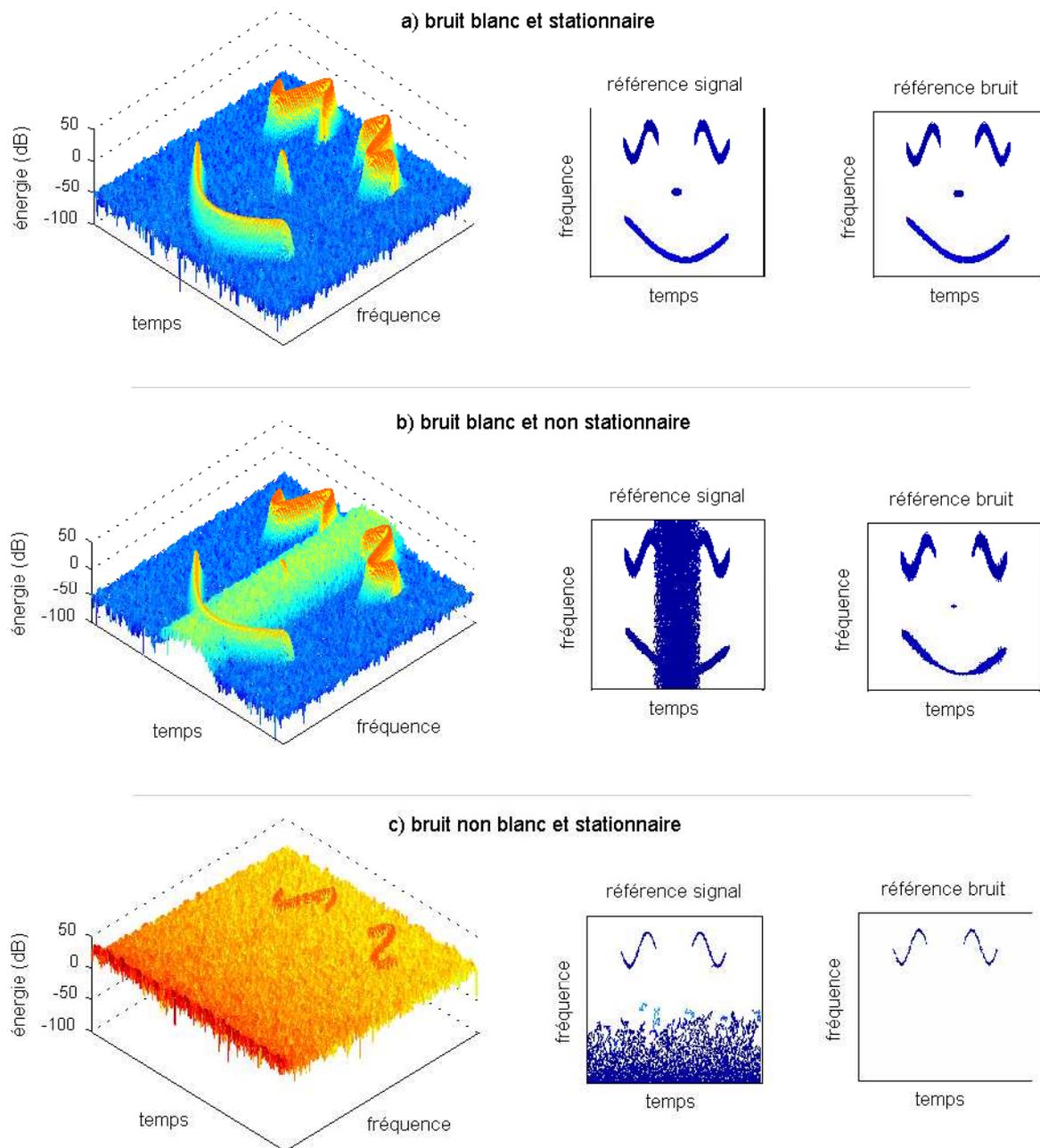


FIG. 1.3 – Illustration de l'intérêt de donner au bruit le rôle de *référence* et de la notion de support temps-fréquence *accessible*. À gauche, spectrogrammes du signal pour une perturbation a) blanche et stationnaire, b) blanche et non stationnaire et c) non blanche et stationnaire. Au centre, support temps-fréquence obtenu avec une référence "signal". À droite, support temps-fréquence obtenu avec une référence "bruit".

le test d'hypothèses formulé en (1.5). Maintenant, de par la nature aléatoire du signal observé $x[m]$, explicitée en (1.4), chaque coefficient $S_x[n, k]$ de la représentation temps-fréquence est également une réalisation d'une variable aléatoire. Le test d'hypothèses formulé en (1.5) est donc de nature statistique et rentre dans la classe de problèmes dit des "tests statistiques d'hypothèses" [VT68]. Nous pouvons ainsi utiliser à profit l'ensemble du formalisme qui a été développé dans le cadre de ces problèmes. Si ce formalisme propose une démarche générale menant exactement à la définition d'un seuil tel que celui utilisé dans l'expression (1.6), il propose surtout un ensemble de critères statistiques à partir desquels ce seuil pourra être défini. Nous verrons que certains de ces critères apportent une réponse tout à fait naturelle à notre problème. Outre la détermination du seuil, un second élément d'intérêt lié aux tests statistiques d'hypothèses concerne l'évaluation de la méthode. Afin de juger de la qualité de la prise de décision, un ensemble de critères a également été développé. L'observation de ces critères permettra alors de préciser les limites de validité et de juger des performances de l'ensemble de la démarche proposée pour déterminer le support temps-fréquence du signal $d[m]$.

Si la détermination du support temps-fréquence d'un signal aléatoire trouve une formulation en terme de test statistique d'hypothèses, il convient maintenant de préciser les besoins de ces méthodes, c'est à dire les connaissances qui sont nécessaires à leur mise en pratique. Dans ce but, nous mentionnons dès à présent le cadre général de solutions apporté par le test du rapport de vraisemblance [VT68]. Dans ce formalisme, le choix de l'hypothèse H_0 ou H_1 s'apparente à comparer le rapport des vraisemblances de l'observation $S_x[n, k]$ sous chaque hypothèse avec un seuil μ . En notant $\mathcal{L}_{[n,k]}$ ce rapport de vraisemblance au point $[n, k]$, la solution au test d'hypothèse est donnée selon

$$\mathcal{L}_{[n,k]} = \frac{p_{H_1}(S_x[n, k])}{p_{H_0}(S_x[n, k])} \underset{H_0}{\overset{H_1}{\geq}} \mu, \quad (1.7)$$

dans lequel $p_{H_0}(S_x[n, k])$ et $p_{H_1}(S_x[n, k])$ représentent les densités de probabilité des coefficients temps-fréquence $S_x[n, k]$ sous les hypothèses H_0 et H_1 ou 'conditionnellement' à ces hypothèses. Ces deux densités de probabilité constituent dès lors le point de départ de la méthode visant à résoudre le test d'hypothèses formulé en (1.5). Par ailleurs, elles représentent le strict équivalent, dans le domaine temps-fréquence, du modèle aléatoire de signal défini en (1.3) et (1.4) dans le domaine temporel. Bien entendu, elles sont entièrement dépendantes de la transformation temps-fréquence choisie.

1.4 Le modèle de signal en temps-fréquence

1.4.1 Choix d'une représentation temps-fréquence

Depuis plus d'un demi-siècle, de nombreuses transformations ont été proposées afin de représenter un signal conjointement en temps et en fréquence. Ces transformations peuvent être de natures différentes et posséder des caractéristiques très variables [Fla93][TMA00]. Comme nous nous intéressons à la localisation de l'énergie d'un signal, les transformations quadratiques et particulièrement celles appartenant à la classe de Cohen [Coh89] ont naturellement retenu notre attention. Une caractéristique majeure de cette classe de représentations est un compromis nécessaire entre localisation et présence de termes d'interférence. Une même représentation ne peut allier à la fois une localisation parfaite de l'énergie propre au signal dans le plan temps-fréquence et l'absence de termes d'interférence liés à la nature quadratique de la transformation d'un signal multi-composantes. Aux deux extrêmes de ce compromis se trouvent la distribution de Wigner-Ville et le spectrogramme. La distribution de Wigner-Ville présente une localisation

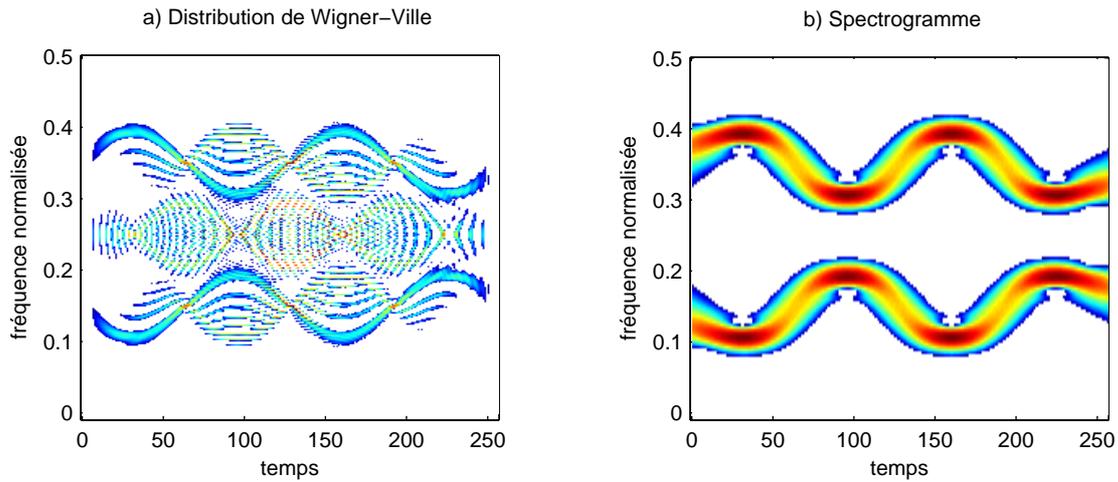


FIG. 1.4 – Comparaison de la distribution de Wigner-Ville (à gauche) et du spectrogramme (à droite) d'un signal composé de deux modulations sinusoidales de fréquences.

parfaite pour les signaux modulés linéairement en fréquence mais se trouve entachée de tous les termes d'interférence. À l'inverse, le spectrogramme d'un signal multi-composantes ne fait apparaître aucun terme d'interférence mais présente une mauvaise localisation de l'énergie des composantes. Pour illustration, la distribution de Wigner-Ville et le spectrogramme d'un signal contenant deux modulations sinusoidales de fréquences sont représentés sur la figure 1.4. Pour les autres transformations temps-fréquence de la classe de Cohen, le choix d'un noyau éventuellement adapté au signal [JB95] ou à la finalité de l'analyse [Dav00] module le compromis entre interférences et localisation.

Dans ce travail, nous avons choisi de représenter les signaux dans le plan temps-fréquence en utilisant le spectrogramme. Trois éléments ont motivé ce choix : l'absence de terme d'interférence d'une part, son inversibilité d'autre part et finalement les possibilités de calculs analytiques et de manipulation de sa densité de probabilités. Notons que ces trois caractéristiques proviennent de la nature linéaire de la transformée de Fourier à partir de laquelle le spectrogramme est construit.

Cherchant à localiser l'énergie d'un signal dans le plan temps-fréquence, l'utilisation d'une représentation présentant de bonnes propriétés de localisation semble être le choix le plus logique. Cependant, la méthode de détection du signal que nous nous proposons de mettre en œuvre est équivalente à un seuillage appliqué directement sur les coefficients de la représentation temps-fréquence. Un point du plan temps-fréquence sera considéré comme appartenant au support temps-fréquence du signal d'intérêt si l'énergie observée en ce point dépasse un certain seuil. Avec une telle méthode, les termes d'interférence seront inévitablement retenus dans le support temps-fréquence du signal. Bien que la nature quadratique du spectrogramme engendre nécessairement quelques interférences entre les différentes composantes d'un signal [JW92], ces interférences n'existent qu'au niveau de l'intersection du support temps-fréquence des composantes. En conséquence, un relief observé sur un spectrogramme traduit toujours l'existence, dans le signal analysé, d'une composante dont l'énergie individuelle appartient à cette région du plan temps-fréquence.

Nous pensons que la localisation d'un signal dans le plan temps-fréquence peut trouver un intérêt dans un processus d'interprétation automatique du contenu d'un signal. Ce processus d'interprétation consiste en une description informative du contenu temps-fréquence, permettant à terme

de mener à une prise de décision [MD04]. L'utilisation d'un masque temps-fréquence comme critère de classification est un exemple de ces processus. De ce point de vue, la présence de termes d'interférence complexifie le processus d'interprétation car elle élargit la notion de composante temps-fréquence : la caractérisation d'une structure temps-fréquence n'est plus possible individuellement, mais doit désormais prendre en compte les autres structures ou composantes présentes dans le signal et avec lesquelles elle est en interaction. Or, si cette caractérisation reste possible pour des signaux simples ou bien connus, elle semble peu réalisable dans le cas général d'un signal inconnu.

Outre le caractère informatif, un autre point d'intérêt d'une localisation du signal dans le plan temps-fréquence est lié à la définition d'une opération de filtrage temps-fréquence. Le support temps-fréquence du signal étant localisé, il semble intéressant de pouvoir reconstruire un signal temporel correspondant à tout ou partie de ce support. Or, cette opération nécessite que la transformation temps-fréquence utilisée soit inversible. Parmi les transformations quadratiques, le spectrogramme possède cette propriété sous une simple condition de normalisation de la fenêtre d'analyse [Coh95][All77].

Finalement, le traitement des signaux aléatoires nécessite la connaissance du comportement aléatoire de ces signaux sur le domaine dans lequel ils sont représentés. Notamment, le test d'hypothèse formulé en (1.7) nécessite la connaissance de la densité de probabilité des coefficients de la représentation temps-fréquence. Si pour l'ensemble des représentations de la classe de Cohen, les moments d'ordre 1 et 2 ont pu être exprimés [MF85][Sta02], une description complète du comportement statistique en terme de densité de probabilité d'un signal aléatoire, même gaussien, reste difficile. La situation du spectrogramme est un peu différente. Bien qu'appartenant à la classe de Cohen, son origine linéaire permet d'exprimer la densité de probabilité de ses coefficients sous l'hypothèse d'un signal gaussien, cette densité étant alors suffisamment simple pour être utilisée dans des développements analytiques.

Nous avons ainsi choisi d'utiliser une représentation temps-fréquence quadratique, libre de tous termes d'interférence (au sens de Wigner-Ville), inversible et pour laquelle le comportement aléatoire des signaux peut être complètement déterminé. Pour ces raisons nous retenons le spectrogramme. Les faibles propriétés de localisation de cette représentation constituent une limite intrinsèque à notre méthode.

Le spectrogramme est défini comme le module carré de la Transformée de Fourier à Court Terme (TFCT) [Fla93]. La Transformée de Fourier à Court Terme $F_x[n, k]$ du signal $x[m]$ est formée d'une succession de Transformées de Fourier Discrètes (TFD), calculées sur des portions de signal pondérées par une fenêtre d'analyse $w[m]$ de longueur M . $F_x[n, k]$ est un nombre complexe dont les parties réelle et imaginaire sont définies par

$$F_x^r[n, k] = \sum_{m=0}^{M-1} x[nD + m]w[m] \cos(-2\pi k \frac{m}{K}), \quad (1.8)$$

$$F_x^i[n, k] = \sum_{m=0}^{M-1} x[nD + m]w[m] \sin(-2\pi k \frac{m}{K}), \quad (1.9)$$

où K est la taille de la TFD, D est le pas de glissement de la fenêtre entre deux analyses successives et K/M correspond au facteur de *zero-padding*. Une condition nécessaire de reconstruction du signal temporel à partir de sa TFCT est l'utilisation d'une fenêtre d'analyse d'énergie unité [All77], soit

$$\sum_{m=0}^{M-1} w[m]^2 = 1. \quad (1.10)$$

Cette normalisation sera appliquée dans toute la suite de ce travail.

Le spectrogramme $S_x[n, k]$ du signal $x[m]$, module carré de la TFCT, s'écrit

$$\begin{aligned} S_x[n, k] &= |F_x[n, k]|^2 \\ &= F_x^r[n, k]^2 + F_x^i[n, k]^2 \\ &= \mathbf{F}_x^T[n, k]\mathbf{F}_x[n, k], \end{aligned} \quad (1.11)$$

dans lequel le vecteur $\mathbf{F}_x[n, k]$, défini selon

$$\mathbf{F}_x[n, k] = \begin{pmatrix} F_x^r[n, k] \\ F_x^i[n, k] \end{pmatrix}, \quad (1.12)$$

est introduit pour les besoins des analyses futures. C'est un vecteur réel de dimension 2 isomorphe au nombre complexe $F_x[n, k]$.

1.4.2 Formulation temps-fréquence du modèle de signal

L'objet de ce paragraphe est de donner la formulation temps-fréquence, basée sur le spectrogramme, du modèle de signal temporel introduit au paragraphe 1.2. Nous précisons comment les différentes composantes de ce modèle apparaissent dans cette représentation et comment les différentes hypothèses s'y traduisent.

En introduisant la décomposition temporelle du signal donnée en (1.1) dans la définition du spectrogramme, nous obtenons

$$\begin{aligned} S_x[n, k] &= S_{d+p}[n, k] \\ &= |F_{d+p}[n, k]|^2 \\ &= F_d^r[n, k]^2 + F_d^i[n, k]^2 + F_p^r[n, k]^2 + F_p^i[n, k]^2 + 2\left(F_d^r[n, k]F_p^r[n, k] + F_d^i[n, k]F_p^i[n, k]\right) \\ &= S_d[n, k] + S_p[n, k] + 2S_{dp}[n, k]. \end{aligned} \quad (1.13)$$

Cette dernière ligne donne la formulation temps-fréquence du modèle de signal (1.1) utilisé dans ce travail.

Le terme $S_d[n, k]$ correspond au spectrogramme du signal d'intérêt déterministe $d[m]$. C'est un terme déterministe. Le terme $S_p[n, k]$ représente le spectrogramme de la perturbation aléatoire $p[m]$. C'est un terme aléatoire. Le terme $S_{dp}[n, k]$ quant à lui, peut être interprété comme le spectrogramme croisé des signaux $d[m]$ et $p[m]$. Il représente le terme d'interférence lié à la nature quadratique du spectrogramme. Toutefois, à la différence des termes d'interférences observés sur une distribution de Wigner-Ville, ce terme n'existe que sur l'intersection des supports TF des deux signaux $d[m]$ et $p[m]$. C'est également un terme aléatoire.

De la même façon que la fonction d'autocorrélation $\Gamma_p[m, \tau]$ de la perturbation $p[m]$ a été introduite en (1.2) dans le domaine temporel, la densité "temps-fréquence" de puissance de cette perturbation, notée $\gamma_p[n, k]$, est ici introduite.

Dans le cas stationnaire, en suivant la relation de Wiener-Khintchine, la densité spectrale de puissance $\gamma_p[k]$ se définit comme la transformée de Fourier suivant la variable de décalage temporel τ de la fonction d'autocorrélation $\Gamma_p[\tau]$ (stationnaire). Dans ce cas, le périodogramme du signal, donné comme le module au carré de sa TFD, en est un estimateur asymptotiquement non biaisé⁴.

⁴Le biais de l'estimateur tend vers 0 lorsque la taille du signal devient infinie.

Dans le cas non stationnaire, une manière de définir une densité spectrale de puissance est de considérer à chaque instant m la transformée de Fourier suivant la variable de décalage temporel τ de la fonction d'autocorrélation $\Gamma_p[m, \tau]$. Cette densité spectrale appartient à la classe des représentations de type s-Wigner [Fla93], pour laquelle la valeur $-1/2$ est choisie pour le paramètre s . Cependant, la propriété d'estimateur non biaisé n'est dans ce cas pas vérifiée par le spectrogramme pour lequel la taille M des trames temporelles de signal analysé est nécessairement finie. La densité spectrale de puissance non stationnaire de la perturbation $\gamma_p[n, k]$, ne peut donc pas être correctement estimée à l'aide du spectrogramme. Devant cette impossibilité, nous choisissons de définir la quantité $\gamma_p[n, k]$ à partir du spectrogramme selon

$$\gamma_p[n, k] = \mathbb{E} \{S_p[n, k]\}. \quad (1.14)$$

S'il s'agit d'une estimation biaisée de la densité spectrale non stationnaire de la perturbation, l'introduction de cette quantité par Mark [Mar70] sous le nom de "spectre physique" repose sur son caractère très intuitif quant à la description des signaux aléatoires non stationnaires. En ce qui nous concerne, nous verrons que cette quantité constitue un élément tout à fait central dans ce travail. Dans la suite, la terminologie "densité spectrale de puissance du bruit" (ou dsp) sera utilisée, par abus donc, pour la désigner.

L'hypothèse de gaussianité de la perturbation $p[m]$ a permis d'établir en (1.4) la distribution de probabilité de l'échantillon $x[m]$. La formulation temps-fréquence de ce modèle aléatoire, pour le spectrogramme et dans le cas d'une perturbation non blanche et non stationnaire est l'objet du second chapitre de ce manuscrit.

1.5 Résumé

Nous proposons de déterminer le support temps-fréquence accessible d'un signal d'intérêt, non stationnaire, déterministe et inconnu, noyé dans un bruit gaussien additif, centré et de fonction d'autocorrélation inconnue. Le spectrogramme est choisi pour représenter les signaux dans le plan temps-fréquence. Le support temps-fréquence accessible d'un signal est défini comme l'ensemble des points temps-fréquence pour lesquels le signal d'intérêt admet une énergie au moins supérieure à celle du bruit. Nous choisissons alors de résoudre le problème de localisation au moyen d'un test de détection binaire, formulé en chaque point du plan temps-fréquence. Pour chaque coefficient du spectrogramme, un seuil de détection doit être défini afin de retenir l'une des hypothèses H_0 = "bruit seul" ou H_1 = "signal + bruit".

Afin de déterminer ce seuil de détection en lien avec la puissance du bruit et selon un critère de détection approprié, les lois de probabilité des observations, c'est-à-dire des coefficients du spectrogramme, doivent être connues. La détermination de ces densités de probabilité est l'objet du chapitre 2. La densité spectrale non stationnaire du bruit apparaît naturellement comme l'un des paramètres de la densité de probabilité des coefficients du spectrogramme. Aussi, le rôle central qu'il joue dans la définition du support temps-fréquence accessible, illustrée par sa position de référence dans le test d'hypothèse impose sa connaissance. Dans le chapitre 3, une méthode d'estimation de ce bruit est proposée. La mise en place et l'évaluation du test de détection à travers la détermination d'un seuil de détection est l'objet du chapitre 4. Finalement, l'algorithme complet de localisation temps-fréquence est présenté dans le chapitre 5.

Chapitre 2

Densité de probabilité du spectrogramme

Sommaire

2.1	Introduction	23
2.2	Densité de probabilité de la TFCT	24
2.3	Variable aléatoire du χ^2	25
2.3.1	Définition	25
2.3.2	Conditions pour le spectrogramme	26
2.4	(Non-)circularité de la TFCT	27
2.4.1	Observations	27
2.4.1.1	signal blanc et stationnaire	27
2.4.1.2	signal blanc et non stationnaire	28
2.4.1.3	signal corrélé et stationnaire	30
2.4.2	Circularité et forme de la fenêtre d'analyse	30
2.4.3	Conclusion	33
2.5	Densité de probabilité d'un module carré gaussien	33
2.5.1	Densité de probabilité dans le cas centré	33
2.5.2	Densité de probabilité dans le cas décentré	34
2.6	Le spectrogramme comme variable du χ^2	36
2.6.1	Quelle loi du χ^2 pour approcher le spectrogramme?	37
2.6.1.1	Paramètre de décentrage θ	37
2.6.1.2	Degré de liberté δ et coefficient de proportionnalité α	37
2.6.2	Mesure de divergence	39
2.6.3	Impact de l'approximation dans un contexte de détection	43
2.7	Conclusion	47

2.1 Introduction

Le but de ce chapitre est de préciser la loi de probabilité selon laquelle se distribuent les coefficients $S_x[n, k]$ du spectrogramme d'un signal $x[m]$ répondant au modèle (1.4) décrit dans le chapitre précédent.

Sous l'hypothèse d'un signal gaussien, la loi du χ^2 est classiquement utilisée pour décrire le module carré des coefficients d'une transformée de Fourier, et donc le spectrogramme [Koo74][Alt80][Kay98]. Cependant, l'exactitude de cette loi repose sur l'hypothèse de circularité de la distribution des coefficients de la transformée de Fourier.

La circularité d'une variable aléatoire correspond à une invariance par rotation de sa densité de probabilité. Dans le cas d'une variable aléatoire gaussienne à 2 dimensions $G = (G_1, G_2)$, la circularité se rapporte aux deux propriétés suivantes :

$$\begin{cases} \text{Cov} \{G_1, G_2\} = 0 \\ \text{Var} \{G_1\} = \text{Var} \{G_2\}, \end{cases} \quad (2.1)$$

dans lequel 'Var' et 'Cov' désignent respectivement la variance et la covariance.

Pour chaque coefficient de la transformée de Fourier d'un signal gaussien, l'hypothèse de circularité renvoie donc à la décorrélation des parties réelle et imaginaire d'une part et à l'égalité de leur variance d'autre part. Si en pratique la circularité de la TFCT est toujours supposée exacte, certains auteurs notent que cette hypothèse est strictement valide lorsque le signal analysé est blanc (échantillons temporels indépendants) et pour une fenêtre d'analyse rectangulaire de taille infinie [Koo74]. Dans le cadre d'un signal non-blanc, Martin précise que la longueur de corrélation du signal analysé doit être faible devant la taille de la fenêtre pour que cette hypothèse soit acceptable [Mar01].

Face à ces remarques, nous nous sommes intéressés à cette hypothèse de circularité de la TFCT [HMM06, HMM08]. Nous avons cherché à définir les conditions pour lesquelles la circularité de la TFCT peut être acceptée et par conséquent, pour lesquelles le spectrogramme peut être décrit par une loi du χ^2 . Les paramètres pris en compte dans cette étude sont :

- la forme de la fenêtre d'analyse,
- la taille de la fenêtre d'analyse,
- le degré de corrélation du signal analysé.
- la non-stationnarité du signal analysé.

L'influence du *zero-padding* a également été étudiée. N'ayant jamais pu mettre à défaut la circularité de la TFCT, l'étude de ce paramètre ne sera pas présentée dans la suite.

Dans les cas où la circularité de la TFCT n'est pas vérifiée, les coefficients du spectrogramme ne sont plus distribués selon une loi du χ^2 mais plus généralement selon la loi du module carré d'une variable gaussienne bidimensionnelle. Nous avons alors cherché à comparer ces deux lois et notamment à évaluer l'erreur commise lorsque la loi du χ^2 est malgré tout utilisée pour décrire le spectrogramme.

Dans ce but, nous avons d'une part évalué la différence entre la distribution du spectrogramme et la loi du χ^2 en utilisant la divergence de Kullback-Leibler [HM07]. Cette étude permet de mettre en évidence l'impact de différents paramètres comme la corrélation du signal analysé ainsi que la nature et la taille de la fenêtre d'analyse utilisée.

D'autre part, dans ce travail, la densité de probabilité du spectrogramme est utilisée dans le cadre d'un test de détection. Dans ce contexte, remplacer la loi du spectrogramme par une loi du χ^2 va entraîner des variations dans les performances du test : les probabilités de détection et de fausses alarmes vont par exemple être modifiées. Afin d'évaluer les conséquences de l'utilisation

d'une loi du χ^2 pour décrire le spectrogramme, nous pouvons donc évaluer les variations des performances du test de détection qui en résultent. Cette seconde étude permet de donner des conditions pratiques sur l'utilisation de la loi du χ^2 dans ce contexte.

L'organisation de ce chapitre est la suivante : dans un premier temps, la densité de probabilité de la TFCT est exprimée en fonction des paramètres du signal et des paramètres de l'analyse temps-fréquence. Ensuite, nous rappelons la définition d'une variable du χ^2 et formulons les conditions de circularité de la TFCT, conditions à respecter pour que le spectrogramme puisse être assimilé à une telle variable. Dans une troisième partie, ces conditions sont évaluées en fonction des caractéristiques du signal et des paramètres choisis pour l'analyse temps-fréquence. Ayant mis en évidence la possible non-circularité de la TFCT, la densité de probabilité d'un coefficient du spectrogramme est formulée dans le cadre général du module carré de variables gaussiennes. Finalement, dans la cinquième et dernière partie, la loi du spectrogramme est comparée à celle du χ^2 et les conséquences de l'approximation dans un test de détection sont présentées.

2.2 Densité de probabilité de la TFCT

L'objectif de ce paragraphe est de formuler la densité de probabilité d'un coefficient $F_x[n, k]$ de la TFCT du signal $x[m]$. Autrement dit, étant donné un processus aléatoire dont la loi de probabilité des échantillons temporels est donnée par (1.4), nous souhaitons connaître la densité de probabilité des coefficients de la TFCT de ce processus. Précisons que nous nous intéressons à la densité de probabilité d'un unique coefficient et que l'étude de la densité de probabilité conjointe à deux coefficients temps-fréquence dépasse le cadre de ce travail. Dans la suite de ce chapitre, nous avons supprimé l'indice temps-fréquence $[n, k]$ dans les notations relatives aux densités de probabilité ou à leurs paramètres. Nous insistons donc sur le fait que pour tout ce qui suit, un unique point $[n, k]$ du plan temps-fréquence est considéré et que tous les paramètres décrits sont *a priori* fonction de ces deux variables n et k .

Les parties réelle $F_x^r[n, k]$ et imaginaire $F_x^i[n, k]$ de la TFCT sont définies comme des combinaisons linéaires des échantillons du signal $x[m]$ (1.8 et 1.9). Chaque échantillon $x[m]$ étant par hypothèse une variable gaussienne, $F_x^r[n, k]$ et $F_x^i[n, k]$ sont elles-mêmes des variables gaussiennes. Le vecteur TFCT $\mathbf{F}_x[n, k]$ (1.12) est donc un vecteur gaussien que l'on note

$$\mathbf{F}_x[n, k] = \begin{pmatrix} F_x^r[n, k] \\ F_x^i[n, k] \end{pmatrix} \sim \mathcal{N} \left\{ \begin{pmatrix} m_1 \\ m_2 \end{pmatrix}, \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12} & \Sigma_{22} \end{pmatrix} \right\}. \quad (2.2)$$

Les cinq paramètres m_1 , m_2 , Σ_{11} , Σ_{22} et Σ_{12} sont définis par

$$\begin{cases} m_1 &= \mathbb{E} \{ F_x^r \}, \\ m_2 &= \mathbb{E} \{ F_x^i \}, \\ \Sigma_{11} &= \mathbb{E} \{ (F_x^r - m_1)^2 \}, \\ \Sigma_{22} &= \mathbb{E} \{ (F_x^i - m_2)^2 \}, \\ \Sigma_{12} &= \mathbb{E} \{ (F_x^r - m_1)(F_x^i - m_2) \}, \end{cases} \quad (2.3)$$

où $\mathbb{E} \{ . \}$ représente l'espérance mathématique.

Nous décrivons maintenant comment les paramètres de ce modèle temps-fréquence s'expriment en fonction des paramètres du modèle temporel de signal d'une part et de la TFCT d'autre part. Dans notre cas, les paramètres du signal sont les échantillons déterministes $d[m]$ et la fonction d'autocorrélation $\Gamma_p[m, \tau]$ de la perturbation. Les paramètres de la TFCT sont la

nature de la fenêtre d'analyse $w[m]$, sa taille M et le nombre K de coefficients spectraux calculés. Notons que le taux de recouvrement des fenêtres n'intervient pas dans la distribution d'un unique coefficient temps-fréquence¹.

D'après le modèle de signal (1.4) et la définition des parties réelle et imaginaire de la TFCT (1.8 et 1.9), les moments d'ordre 1, m_1 et m_2 , du vecteur \mathbf{F}_x sont respectivement les parties réelle et imaginaire de la TFCT du signal d'intérêt $d[m]$, soit

$$\begin{cases} m_1 = F_d^r[n, k], \\ m_2 = F_d^i[n, k]. \end{cases} \quad (2.4)$$

Les moments centrés d'ordre 2, soient Σ_{11} , Σ_{22} et Σ_{12} , sont quant à eux définis par la fonction d'autocorrélation $\Gamma_p[m, \tau]$ de la perturbation. En utilisant une notation matricielle, on obtient

$$\begin{cases} \Sigma_{11} = \mathbf{W}^T \mathbf{C}_k \mathbf{R}_p \mathbf{C}_k \mathbf{W}, \\ \Sigma_{22} = \mathbf{W}^T \mathbf{S}_k \mathbf{R}_p \mathbf{S}_k \mathbf{W}, \\ \Sigma_{12} = \mathbf{W}^T \mathbf{C}_k \mathbf{R}_p \mathbf{S}_k \mathbf{W}, \end{cases} \quad (2.5)$$

où le vecteur \mathbf{W} contient les échantillons de la fenêtre d'analyse $w[m]$, soit

$$\mathbf{W}^T = [w[0] \cdots w[M-1]], \quad (2.6)$$

\mathbf{C}_k et \mathbf{S}_k sont des matrices diagonales regroupant les coefficients du cosinus et du sinus pour la fréquence k , soit

$$\mathbf{C}_k = \text{diag} \left[\cos(-2\pi k \frac{m}{K}) \right]_{m=0, M-1}, \quad (2.7)$$

$$\mathbf{S}_k = \text{diag} \left[\sin(-2\pi k \frac{m}{K}) \right]_{m=0, M-1}, \quad (2.8)$$

et finalement \mathbf{R}_p , la matrice de corrélation associée à $\Gamma_p[m, \tau]$, est définie pour une analyse débutant à l'instant m et d'horizon temporel M selon

$$\mathbf{R}_p = \begin{bmatrix} \Gamma_p[m, 0] & \Gamma_p[m, 1] & \cdots & \Gamma_p[m, M-1] \\ \Gamma_p[m+1, 1] & \Gamma_p[m+1, 0] & \cdots & \Gamma_p[m+1, M-2] \\ \vdots & \vdots & \ddots & \vdots \\ \Gamma_p[m+M-1, M-1] & \Gamma_p[m+M-1, M-2] & \cdots & \Gamma_p[m+M-1, 0] \end{bmatrix}. \quad (2.9)$$

Nous notons que cette matrice \mathbf{R}_p , définie ici dans le cadre général d'une perturbation $p[m]$ non blanche et non stationnaire, devient symétrique pour une perturbation stationnaire à l'ordre 2 et diagonale pour une perturbation blanche.

2.3 Variable aléatoire du χ^2

2.3.1 Définition

Étant donné N variables gaussiennes $G_i \sim \mathcal{N}(m_i, \sigma^2)$ indépendantes et de variance identique σ^2 , la somme des carrés de ces variables $\sum_{i=1}^N G_i^2$ est distribuée selon une loi du χ^2 [JKB95]. Une variable du χ^2 , notée $\chi^2(\delta, \alpha, \theta)$, est caractérisée par trois paramètres :

¹Ce paramètre agit par contre sur la corrélation entre les coefficients spectraux des trames temporelles successives.

- le degré de liberté δ correspond au nombre de variables gaussiennes indépendantes. Dans le cas ci-dessus, $\delta = N$,
- le coefficient de proportionnalité α représente la variance des variables gaussiennes, soit ici $\alpha = \sigma^2$,
- le paramètre de décentrage θ regroupe les moyennes des variables gaussiennes sous la forme $\theta = \sum_{i=1}^N m_i^2$.

Dans le cas décentré ($\theta \neq 0$), la densité de probabilité $p_{\chi^2(\delta, \alpha, \theta)}(x)$ d'une variable du chi² s'écrit

$$p_{\chi^2(\delta, \alpha, \theta)}(x) = \frac{1}{2\alpha} \left(\frac{x}{\theta}\right)^{\frac{\delta-2}{4}} \exp\left(-\frac{x+\theta}{2\alpha}\right) I_{\frac{\delta-2}{2}}\left(\frac{\sqrt{x\theta}}{\alpha}\right), x \geq 0, \quad (2.10)$$

dans lequel $I_n(\cdot)$ est la fonction de Bessel modifiée du premier type et d'ordre n . Dans le cas centré ($\theta = 0$), la densité de probabilité devient

$$p_{\chi^2(\delta, \alpha, \theta=0)}(x) = \frac{1}{(2\alpha)^{\frac{\delta}{2}} \Gamma(\frac{\delta}{2})} x^{\frac{\delta-2}{2}} \exp\left(-\frac{x}{2\alpha}\right), x \geq 0, \quad (2.11)$$

ou $\Gamma(n) = \int_0^{+\infty} e^{-x} x^{n-1} dx$ est la fonction gamma [JKB95].

Les moments centrés d'ordre 1 et 2 d'une variable du chi² s'écrivent

$$\mathbb{E}\{\chi^2(\delta, \alpha, \theta)\} = \alpha\delta + \theta, \quad (2.12)$$

$$\text{Var}\{\chi^2(\delta, \alpha, \theta)\} = 2\alpha^2\delta + 4\alpha\theta, \quad (2.13)$$

où $\text{Var}\{\cdot\}$ dénote la variance.

Remarque sur cette définition : une variable aléatoire du chi² est classiquement définie à partir de variables gaussiennes de variance unité. Elle est alors caractérisée par deux paramètres : le degré de liberté δ et le paramètre de décentrage θ . Le coefficient de proportionnalité α étant alors égal à 1. Lorsque les variables gaussiennes ne sont pas de variance unité, la somme de leurs carrés est alors proportionnelle à une variable du chi², le coefficient de proportionnalité étant la variance des variables gaussiennes. Cette variance étant un paramètre important de notre modèle, nous avons choisi de l'intégrer explicitement dans l'écriture des variables du chi². Aussi, nous retenons l'appellation "variable du chi²" bien qu'il s'agisse alors d'un abus de langage. Les densités de probabilité formulées en (2.10) et (2.11) sont obtenues selon

$$p_{\chi^2(\delta, \alpha, \theta)}(x) = f_{\alpha\chi^2(\delta, \frac{\theta}{\alpha})}(x) = \frac{1}{\alpha} f_{\chi^2(\delta, \frac{\theta}{\alpha})}\left(\frac{x}{\alpha}\right), \quad (2.14)$$

dans lequel $f_{\chi^2(\delta, \theta)}(x)$ correspond à la densité de probabilité d'une variable du chi² à δ degrés de liberté et de paramètre de décentrage θ telle qu'exprimée dans la littérature [JKB95].

2.3.2 Conditions pour le spectrogramme

De par sa définition (1.11), le spectrogramme est la somme des carrés des parties réelle et imaginaire de la TFCT. Chacune de celles-ci étant dans notre cas une variable gaussienne, le spectrogramme semble naturellement suivre une loi du chi² à deux degrés de libertés. Toutefois, comme précisé dans la définition ci-dessus, une variable du chi² est issue de variables gaussiennes indépendantes et de même variance. Pour être assimilé à une variable du chi², le spectrogramme doit donc répondre à ces deux mêmes conditions : l'indépendance et l'homoscédasticité (l'égalité

des variances) des parties réelle et imaginaire de la TFCT, ou de manière équivalente, la circularité de la TFCT.

Reprenant la loi de probabilité de la TFCT établie en (2.2), ces deux conditions se rapportent à la matrice de variance-covariance Σ : le spectrogramme est distribué selon une loi du χ^2 si la matrice Σ est proportionnelle à la matrice identité, soit encore

$$\text{condition 1 : } \Sigma_{11} = \Sigma_{22}, \quad (2.15)$$

$$\text{condition 2 : } \Sigma_{12} = 0. \quad (2.16)$$

L'objet de la section suivante est de vérifier si ces deux conditions sont toujours respectées.

2.4 (Non-)circularité de la TFCT

Nous avons vu que la condition nécessaire au spectrogramme pour suivre une loi du χ^2 est la circularité de la distribution de la TFCT, condition également formulée par les deux égalités (2.15) et (2.16). Dans cette section, nous cherchons à savoir s'il existe des contextes pour lesquels ces deux égalités ne sont plus vérifiées. Par contexte, nous entendons un choix de paramètres concernant l'analyse temps-fréquence (forme et taille de fenêtre) ainsi qu'un niveau de corrélation ou de non-stationnarité pour le signal analysé.

Face à différents contextes donc, les deux critères suivants sont évalués :

$$h = \frac{\Sigma_{11}}{\Sigma_{22}}, \quad (2.17)$$

$$\rho = \frac{\Sigma_{12}}{\sqrt{\Sigma_{11}\Sigma_{22}}}. \quad (2.18)$$

Ces deux critères ne sont qu'une reformulation des deux conditions (2.15) et (2.16). h correspond à un critère d'hétéroscédasticité (différence de variance), il vaut 1 lorsque les parties réelle et imaginaire de la TFCT ont la même variance. ρ est le coefficient de corrélation entre les parties réelle et imaginaire de la TFCT. Les éléments Σ_{11} , Σ_{22} et Σ_{12} de la matrice de variance-covariance Σ sont évalués à partir des expressions (2.5) dans lesquelles le contexte intervient de manière explicite.

2.4.1 Observations

2.4.1.1 signal blanc et stationnaire

Pour une perturbation $p[m]$ blanche, stationnaire, et gaussienne de variance σ^2 , la matrice de corrélation \mathbf{R}_p du signal observé $x[m]$ est proportionnelle à la matrice identité, soit $\mathbf{R}_p = \sigma^2 \mathbf{I}$. les paramètres Σ_{11} , Σ_{22} et Σ_{12} sont alors définis par la forme et la taille de la fenêtre d'analyse $w[m]$.

La figure 2.1 présente les évolutions du critère d'hétéroscédasticité h et du coefficient de corrélation ρ le long de l'axe des fréquences, pour les fenêtres rectangulaire, Kaiser de paramètre 4 et Hanning, de longueur 64, 128, 256 et 512 échantillons. Pour la fenêtre rectangulaire, aucune différence entre les variances et aucune corrélation ne sont observées entre les parties réelle et imaginaire de la TFCT. Les mêmes tests sur des fenêtres de 16 et 32 points mènent aux mêmes résultats. Les coefficients de la TFCT sont donc distribués de façon circulaire. Pour les fenêtres d'analyse non rectangulaires, une différence de variance ainsi que de la corrélation apparaissent entre les parties réelle et imaginaire de la TFCT. Toutefois, le rapport des variances ne dépassant pas 10^{-3} et le coefficient de corrélation restant inférieur à 10^{-4} , la non-circularité de la distribution de la TFCT reste très faible.

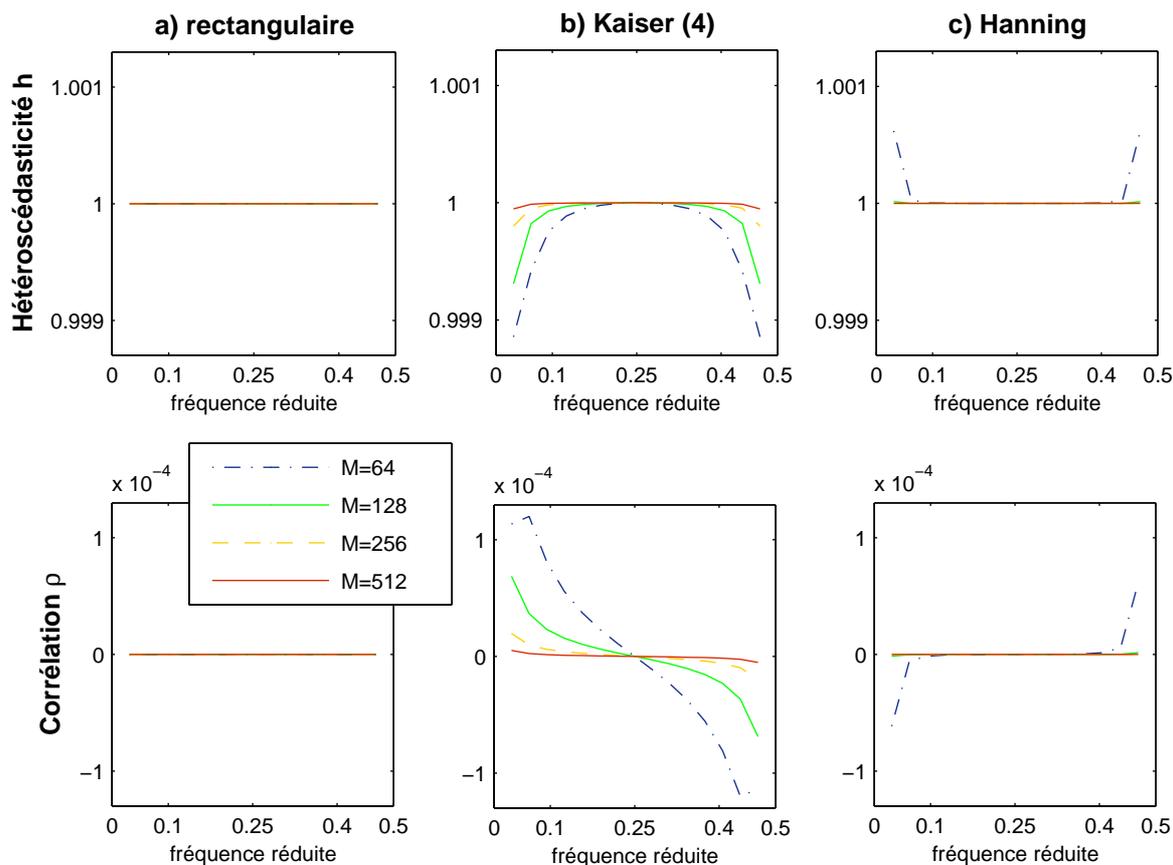


FIG. 2.1 – *Cas blanc et stationnaire*. Critère d’hétéroscedasticité h (rangée du haut) et coefficient de corrélation ρ (rangée du bas) entre les parties réelle et imaginaire de la TFCT d’un signal blanc et stationnaire pour les fenêtres d’analyse a) rectangulaire, b) Kaiser-4 et c) Hanning, de différentes longueurs M .

2.4.1.2 signal blanc et non stationnaire

L’analyse temps-fréquence est dédiée à l’analyse des signaux non stationnaires. Aussi, si l’interprétation physique du spectrogramme requiert la stationnarité du signal à l’échelle de la fenêtre d’analyse, celle-ci n’a pour ainsi dire aucune raison d’être vérifiée. Nous observons donc ici l’effet d’une non-stationnarité sur la circularité de la TFCT.

Soit $x[m]$ un signal blanc, gaussien, centré, dont la variance non stationnaire $\sigma^2[m]$ admet un saut plus ou moins abrupt dans l’intervalle de temps défini par la fenêtre d’analyse. La figure 2.2 donne le profil utilisé pour la variance $\sigma^2[m]$, pour une fenêtre de $M = 256$ points. Avec la taille M de la fenêtre, trois autres paramètres définissent ce profil : l’amplitude de la non-stationnarité (*saut*), son caractère abrupt (L) et sa position dans l’intervalle de la fenêtre (C).

La figure 2.3 présente les résultats obtenus pour différentes amplitudes de non-stationnarité. Les deux critères h et ρ observés sont comparables pour les trois fenêtres. Ils traduisent une plus grande non-circularité de la TFCT que dans le cas blanc et stationnaire précédent. On remarque que cette non-circularité apparaît pour une faible amplitude de non-stationnarité (*saut* = 1.1) et lui est ensuite proportionnelle. De la même manière, nous avons observé une évolution des

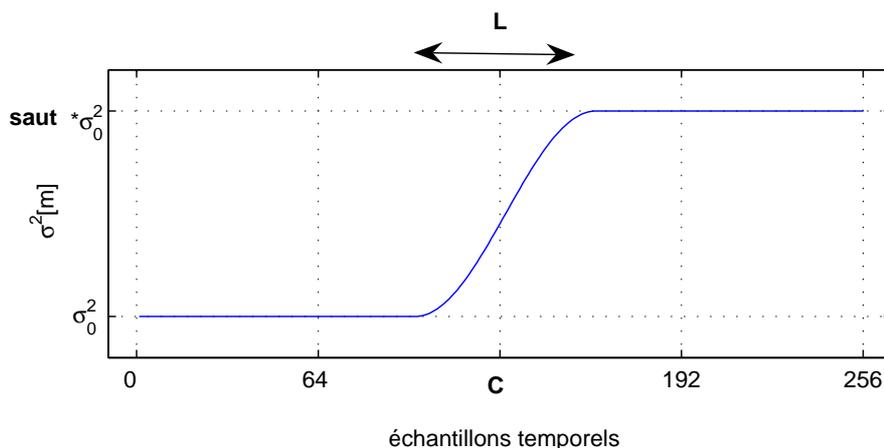


FIG. 2.2 – Profil de variance non stationnaire défini par les trois paramètres *saut*, *L* et *C*, pour une fenêtre de $M = 256$ points. Ici, $L = M/4 = 64$ et $C = M/2$.

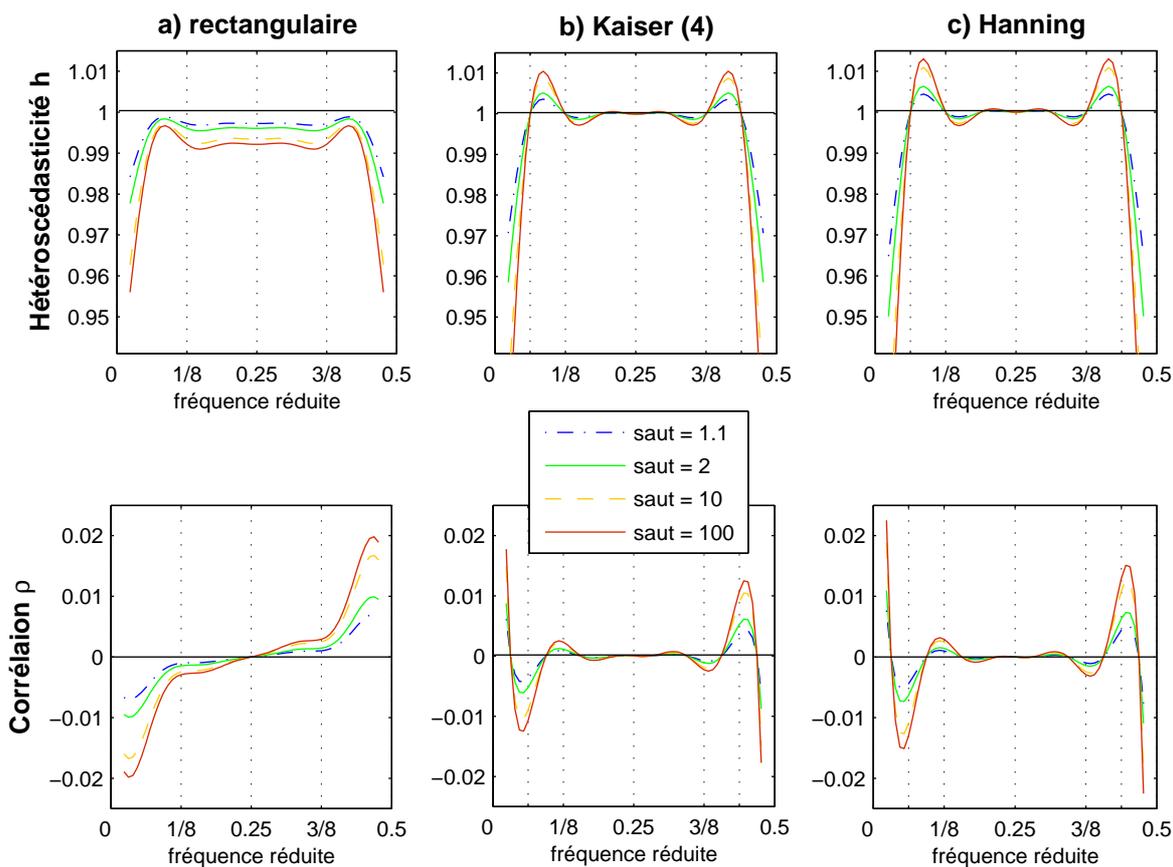


FIG. 2.3 – *Cas blanc et non stationnaire*. Critère d’hétérosécédasticité h (rangée du haut) et coefficient de corrélation ρ (rangée du bas) entre les parties réelle et imaginaire de la TFCT d’un signal blanc et non stationnaire pour les fenêtres d’analyse a) rectangulaire, b) Kaiser-4 et c) Hanning de $M = 256$ points et différentes valeurs de non-stationnarité.

critères h et ρ proportionnelle au caractère abrupt de la non-stationnarité. À ce titre, les résultats présentés sur la figure 2.3 pour $L = M/32 = 8$ échantillons correspondent à une non-stationnarité déjà très abrupte. Finalement, la position de la non-stationnarité (paramètre C , ici fixé à $M/2+5$) n'influence pas le degré de non-circularité mais introduit plus ou moins d'oscillations sur les profils fréquentiels tels que ceux présentés sur la figure 2.3.

Cependant, les oscillations observées pour les fenêtres non rectangulaires, soit ici les fenêtres de Hanning et de Kaiser, revêtent un caractère particulier : à l'image d'un cosinus et d'un sinus, elles sont déphasées de $\pi/2$ entre les critères h et ρ . Ainsi, lorsque l'un ou l'autre des critères est maximal, son homologue s'annule. Ce comportement tend alors à limiter la non-circularité puisque hétéroscédasticité et corrélation ne sont jamais fortement présents en même temps.

2.4.1.3 signal corrélé et stationnaire

Prenons maintenant l'exemple d'un signal stationnaire gaussien centré et exponentiellement corrélé. Sa fonction d'autocorrélation $\Gamma[\tau]$ s'écrit

$$\Gamma[\tau] = a \exp \left\{ -\frac{|\tau|}{\tau_0} \right\}, \quad (2.19)$$

dans lequel la variance $a = \Gamma[0]$ est fixée à 1 et τ_0 correspond au temps de corrélation. Afin de mettre en évidence le rôle de la corrélation du signal rapportée à la taille M de la fenêtre d'analyse, nous définissons le rapport fenêtre-corrélation, noté λ , tel que

$$\lambda = \frac{\tau_0}{M}. \quad (2.20)$$

L'évolution du critère d'hétéroscédasticité h et du coefficient de corrélation ρ le long de l'axe des fréquences, pour différentes fenêtres d'analyse et différents rapports fenêtre-corrélation λ est rapportée sur la figure 2.4. Contrairement au cas d'un signal blanc, la circularité de la distribution de la TFCT semble plus nettement mise à défaut lorsque le signal est analysé avec une fenêtre rectangulaire : pour une longueur de corrélation 10 fois plus grande que la taille de la fenêtre ($\lambda = 10$), le coefficient de corrélation atteint 0.5, tandis que le rapport des variances est de 3. Pour les fenêtres non rectangulaires, l'impact de la corrélation est moindre (fenêtre de type Kaiser) voire très faible pour la fenêtre de Hanning. L'explication de l'impact de la forme de la fenêtre sur la circularité de la TFCT en présence de corrélation est l'objet du paragraphe suivant.

Nous avons également testé le cas d'un signal corrélé dans lequel intervient une non-stationnarité identique à celle définie en 2.4.1.2. Les non-circularités observées sont alors la somme des cas blanc/non-stationnaire et corrélé/stationnaire.

2.4.2 Circularité et forme de la fenêtre d'analyse

La figure 2.4 montre que l'utilisation d'une fenêtre non rectangulaire permet de considérablement réduire l'impact de la corrélation sur la circularité de la TFCT. Nous tentons ici d'apporter un élément d'explication à cette observation.

Revenons à la définition (2.5) des éléments de la matrice de variance-covariance Σ . Chacun des éléments Σ_{ij} peut être vu comme une forme quadratique, fonction de la fenêtre d'analyse \mathbf{W} et d'un noyau \mathbf{K}_{ij} , et l'on peut écrire

$$\begin{cases} \Sigma_{11} &= \mathbf{W}^T \mathbf{K}_{11} \mathbf{W}, \\ \Sigma_{22} &= \mathbf{W}^T \mathbf{K}_{22} \mathbf{W}, \\ \Sigma_{12} &= \mathbf{W}^T \mathbf{K}_{12} \mathbf{W}, \end{cases} \quad (2.21)$$

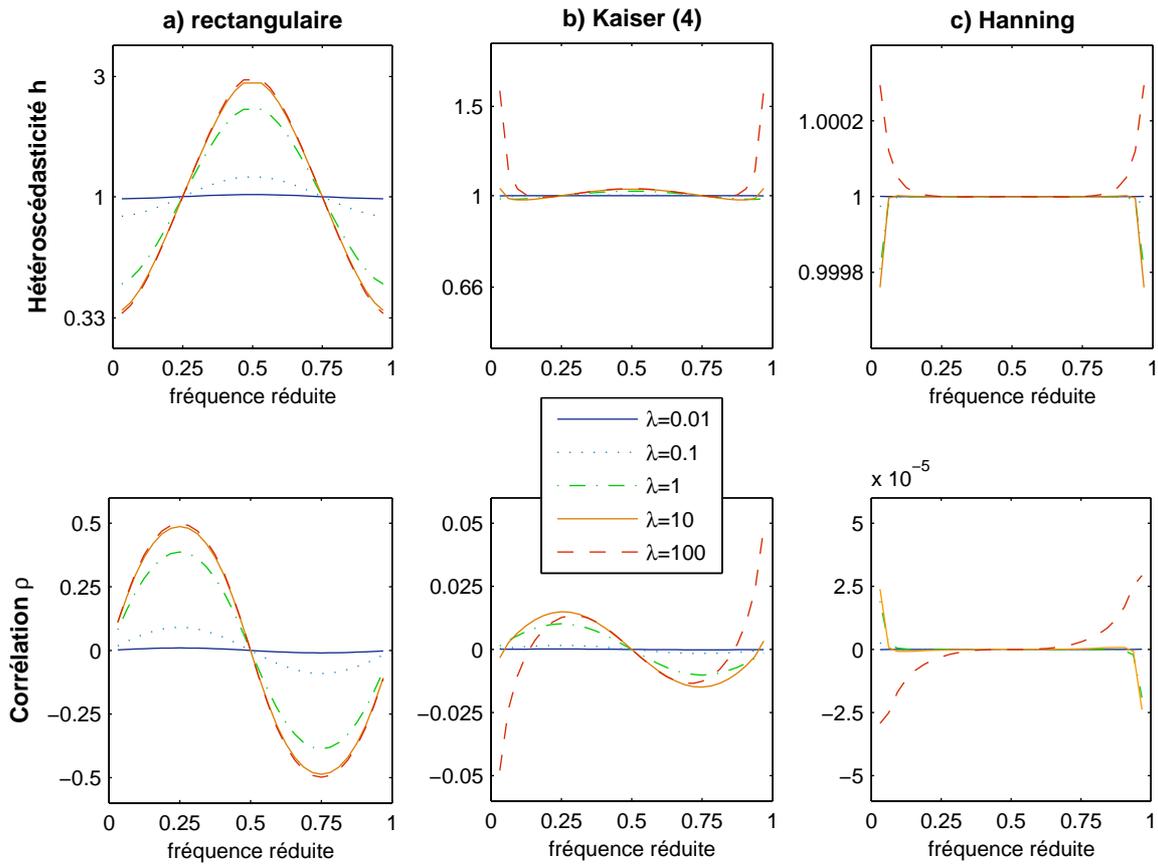


FIG. 2.4 – *Cas corrélé et stationnaire*. Critère d'hétéroscédasticité h (rangée du haut) et coefficient de corrélation ρ (rangée du bas) entre les parties réelle et imaginaire de la TFCT d'un signal exponentiellement corrélé et stationnaire pour les fenêtres d'analyse a) rectangulaire, b) Kaiser-4 et c) Hanning de $M = 256$ points et différents rapports fenêtre-corrélation λ .

dans lesquels les noyaux \mathbf{K}_{ij} , matrices carrées de taille $M * M$ avec M la taille de la fenêtre \mathbf{W} , sont fonctions de la fréquence k observée et de la matrice de corrélation \mathbf{R}_p de la perturbation gaussienne additive $p[m]$:

$$\begin{cases} \mathbf{K}_{11} &= \mathbf{C}_k \mathbf{R}_p \mathbf{C}_k, \\ \mathbf{K}_{22} &= \mathbf{S}_k \mathbf{R}_p \mathbf{S}_k, \\ \mathbf{K}_{12} &= \mathbf{C}_k \mathbf{R}_p \mathbf{S}_k. \end{cases} \quad (2.22)$$

En déployant explicitement les formes quadratiques (2.21), nous obtenons finalement

$$\begin{cases} \Sigma_{11} &= \sum_{i,j=1}^M \mathbf{K}_{11}(i, j) * w[i]w[j], \\ \Sigma_{22} &= \sum_{i,j=1}^M \mathbf{K}_{22}(i, j) * w[i]w[j], \\ \Sigma_{12} &= \sum_{i,j=1}^M \mathbf{K}_{12}(i, j) * w[i]w[j], \end{cases} \quad (2.23)$$

Il ressort de cette écriture que chaque élément Σ_{ij} de la matrice de variance-covariance Σ est la somme des éléments du noyau \mathbf{K}_{ij} pondérés par la fenêtre d'analyse $w[m]$ choisie.

Intéressons-nous au phénomène d'hétéroscédasticité qui s'exprime par l'inégalité $\Sigma_{11} \neq \Sigma_{22}$. Sur la figure 2.5 sont représentés les produits noyaux*fenêtre avant la sommation (2.23). La

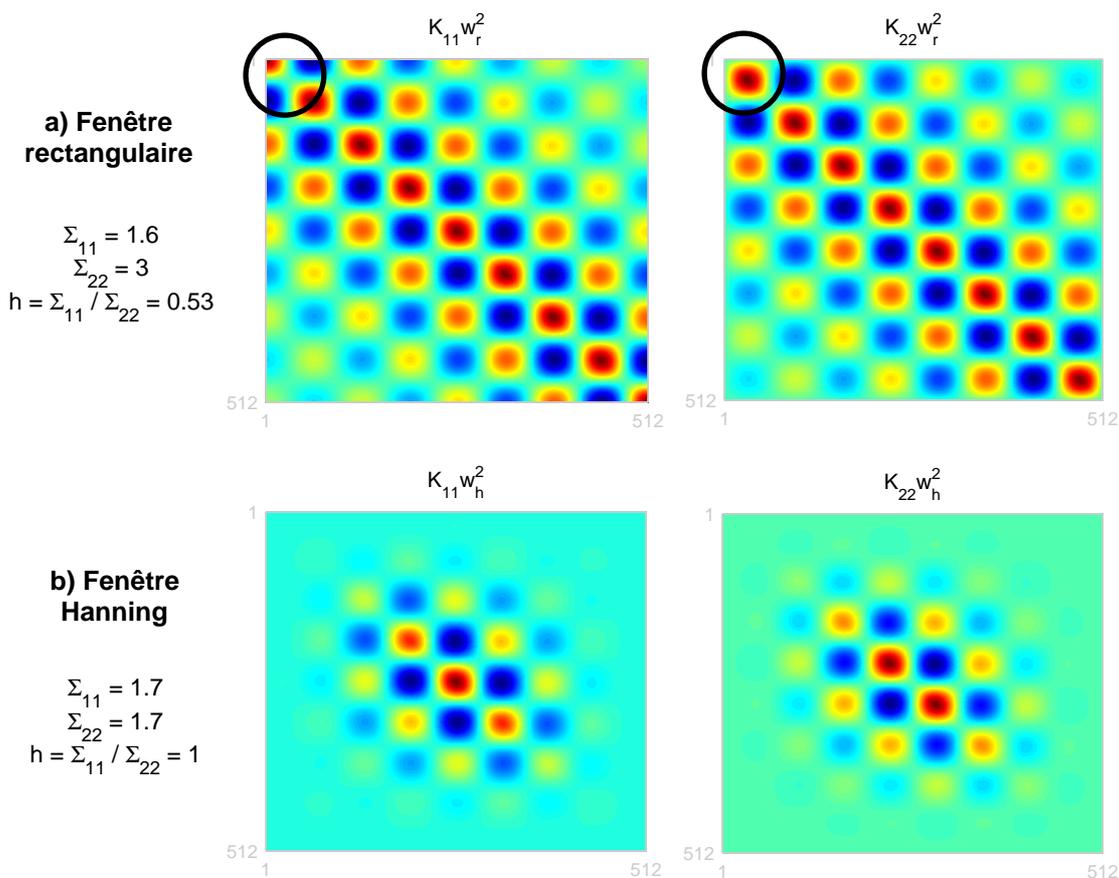


FIG. 2.5 – Noyaux de covariance \mathbf{K}_{11} et \mathbf{K}_{22} multipliés par a) une fenêtre rectangulaire et b) une fenêtre de Hanning de 512 points. Les cercles noirs indiquent les pixels qui engendrent une différence entre Σ_{11} et Σ_{22} pour la fenêtre rectangulaire. La fréquence observée est $k = 4$.

fenêtre rectangulaire (figures du haut) génère une pondération uniforme tandis que la fenêtre de Hanning (figures du bas) ne retient que la zone centrale du noyau.

Lorsque le signal est blanc, la matrice de corrélation \mathbf{R}_p et par conséquent les noyaux \mathbf{K}_{ij} sont diagonaux. La présence de corrélation dans le signal va étendre les noyaux \mathbf{K}_{ij} de chaque côté de la diagonale principale. Les noyaux présentés sur la figure 2.5 correspondent à une corrélation exponentiellement décroissante avec un rapport fenêtre-corrélation $\lambda = 0,5$.

On observe alors que la présence de corrélation introduit une différence entre les noyaux \mathbf{K}_{11} et \mathbf{K}_{22} aux abords de la diagonale principale (voir le cercle noir sur la figure). Les fonctions sinus et cosinus étant déphasées de $\pi/2$, le noyau \mathbf{K}_{22} est obtenu à partir de \mathbf{K}_{11} par deux translations de $T_k/4$ échantillons dans les sens horizontal et vertical, où T_k représente la période associée à la fréquence k observée. En conséquence, si 8 ronds (rouges) sont présents sur la diagonale principale du noyau \mathbf{K}_{22} (en haut à droite), ils sont au nombre de 7,5 pour le noyau \mathbf{K}_{11} (en haut à gauche). La pondération uniforme de la fenêtre rectangulaire conserve cette différence. Les parties réelle et imaginaire de la TFCT n'ont alors plus la même variance. Par contre, cette différence est éliminée par une fenêtre dont la valeur est nulle à ses extrémités (par exemple Hanning), et dont la pondération ne retient que la partie centrale des noyaux. Dans ce cas les parties réelles et imaginaire de la TFCT admettent une variance identique, quel que soit le degré de corrélation du signal analysé.

2.4.3 Conclusion

Nous avons montré que l'hypothèse de circularité de la distribution de la TFCT pouvait être mise en défaut par la présence de corrélation et de non-stationnarité dans le signal analysé. Nous avons également vu que l'utilisation d'une fenêtre à bordures nulles permet de considérablement diminuer l'impact de la corrélation et maintenir la circularité de la TFCT. D'une manière plus générale, nous mentionnons que cette "bonne" propriété semble liée à la régularité de l'ensemble de la fenêtre d'analyse.

Notre objectif est maintenant d'évaluer l'impact de l'utilisation d'une loi du χ^2 pour décrire le spectrogramme dans les cas où l'hypothèse de circularité de la distribution de la TFCT n'est pas vérifiée. Auparavant, la densité de probabilité d'un coefficient du spectrogramme doit être formulée pour ces mêmes cas. C'est l'objet de la prochaine partie.

2.5 Densité de probabilité d'un module carré gaussien

Dans la partie précédente, nous avons vu que la circularité de la distribution de la TFCT n'est pas toujours vérifiée, notamment lorsqu'une fenêtre rectangulaire est utilisée pour l'analyse de signaux non blancs ou non stationnaires. Dans de tels cas, le spectrogramme S_x (1.11), module carré du vecteur \mathbf{F}_x , n'est pas distribué selon une loi du χ^2 , mais plus généralement comme le module carré d'une variable gaussienne bidimensionnelle, que nous nommerons "module carré gaussien". L'objet de cette partie est de formuler la densité de probabilité de cette variable aléatoire.

Nous notons que le module carré d'un vecteur \mathbf{F}_x , soit $\mathbf{F}_x^T \mathbf{F}_x$, est un cas particulier de forme quadratique $\mathbf{F}_x^T A \mathbf{F}_x$, pour lequel la matrice A est la matrice identité. Nous pouvons donc avoir recours aux densités de probabilité des formes quadratiques de variables gaussiennes afin de formuler celle du module carré gaussien.

2.5.1 Densité de probabilité dans le cas centré

Dans le cas centré, c'est-à-dire pour $m_1 = m_2 = 0$, le spectrogramme $S_x[n, k]$ est une forme quadratique d'un vecteur gaussien centré. La densité de probabilité d'une telle variable est connue et peut être calculée de la manière suivante. La fonction génératrice des moments s'écrit ([Sch91] p.65) :

$$M(\mu) = \mathbb{E} \left\{ e^{-\mu \mathbf{F}_x^T \mathbf{F}_x} \right\} = \frac{1}{[\det(\mathbf{I}_2 + 2\mu \boldsymbol{\Sigma})]^{1/2}}, \quad (2.24)$$

dans lequel ' $\det(\cdot)$ ' correspond au déterminant. La transformée de Laplace inverse de cette fonction donne la densité de probabilité, qui en fonction de la matrice de variance-covariance $\boldsymbol{\Sigma}$ (2.2) s'écrit [DN73] :

$$p_{S_x}(s) = \frac{1}{\sqrt{4 \det \boldsymbol{\Sigma}}} \exp \left(-\frac{\text{tr} \boldsymbol{\Sigma}}{4 \det \boldsymbol{\Sigma}} s \right) I_0 \left(\frac{[(\text{tr} \boldsymbol{\Sigma})^2 - 4 \det \boldsymbol{\Sigma}]^{\frac{1}{2}}}{4 \det \boldsymbol{\Sigma}} s \right), \quad s \geq 0. \quad (2.25)$$

où $I_0(\cdot)$ est la fonction de Bessel modifiée du premier type et d'ordre 0, tandis que ' $\det \boldsymbol{\Sigma}$ ' et ' $\text{tr} \boldsymbol{\Sigma}$ ' sont les déterminant et trace de la matrice $\boldsymbol{\Sigma}$.

Des exemples de densités de probabilité dans le cas centré et correspondant à différentes matrices $\boldsymbol{\Sigma}$ sont présentés sur la figure 2.6. Par rapport à une loi du χ^2 (trait plein), la présence de corrélation et/ou d'hétéroscédasticité tend à concentrer la densité de probabilité vers l'origine.

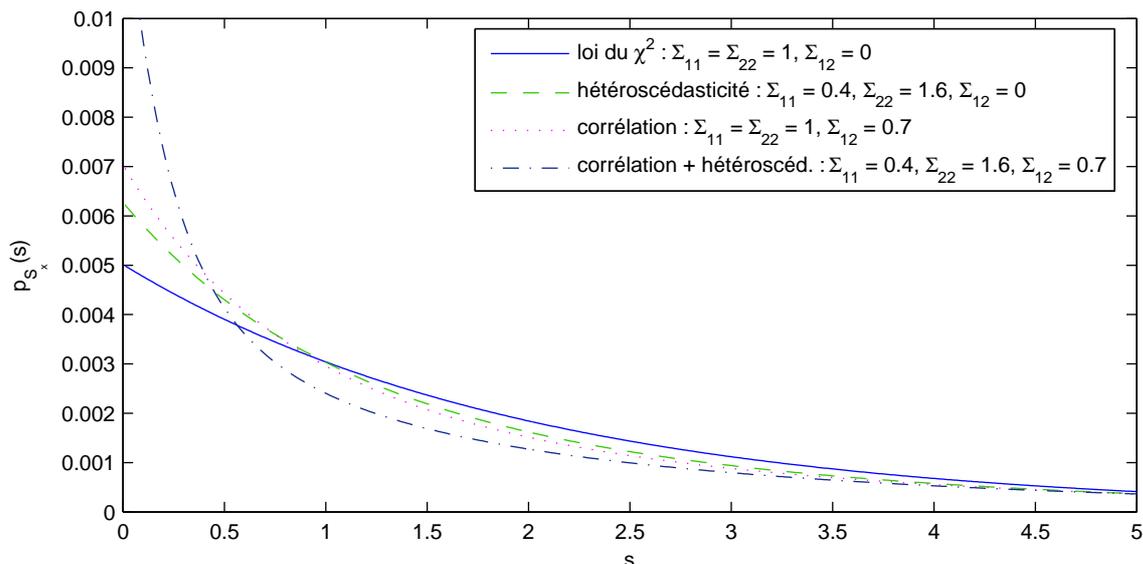


FIG. 2.6 – Exemples de densité de probabilité d'un module carré gaussien centré pour différentes matrices de variance-covariance Σ .

2.5.2 Densité de probabilité dans le cas décentré

Dans le cas décentré, la distribution d'une forme quadratique de variables gaussiennes n'a pas d'expression analytique finie. La littérature propose de nombreuses expressions et approximations de cette loi. On trouve notamment des développements en séries utilisant les polynômes de Laguerre [Gur55][FJ60] ou des lois du χ^2 [RP49][Imh61][Tzi87] ou encore la fonction hypergéométrique [JL99]. Bien que l'une de ces expressions eut pu être utilisée, nous nous sommes intéressés à une approche géométrique permettant de calculer la densité de probabilité du module carré d'une variable aléatoire bidimensionnelle, centré ou non, gaussienne ou non.

Reprécisons ici que les parties réelle et imaginaire de la TFCT du signal $x[m]$, notées respectivement F_x^r et F_x^i et définies en (1.8) et (1.9), sont des variables aléatoires, dont les réalisations sont notées respectivement f_x^r et f_x^i . De même, le spectrogramme S_x défini en (1.11) est une variable aléatoire dont les réalisations sont notées s .

Le coefficient temps-fréquence S_x est égal à une certaine valeur s si les réalisations f_x^r et f_x^i des parties réelle et imaginaire de la TFCT satisfont l'équation $f_x^{r^2} + f_x^{i^2} = s$. Dans le plan réel-imaginaire (F_x^r, F_x^i) , cette équation décrit le cercle de centre $(0,0)$ et de rayon \sqrt{s} . La densité de probabilité du spectrogramme correspond donc à l'intégrale sur ce domaine de la densité de probabilité conjointe des variables F_x^r et F_x^i , soit $p_{(F_x^r, F_x^i)}(f_x^r, f_x^i)$. La figure 2.7 illustre ce principe : à côté de la densité de probabilité conjointe des variables F_x^r et F_x^i , le cercle d'équation $s = f_x^{r^2} + f_x^{i^2}$ est tracé. Il correspond à l'ensemble des réalisations (f_x^r, f_x^i) pour lesquelles le spectrogramme prend la même valeur.

Pour adapter les notations à la géométrie du problème, la densité conjointe $p_{(F_x^r, F_x^i)}(f_x^r, f_x^i)$ est exprimée en coordonnées polaires (r, ϕ) par le changement de variables

$$\begin{cases} f_x^r &= r \cos \phi, \\ f_x^i &= r \sin \phi. \end{cases} \quad (2.26)$$

La densité de probabilité du spectrogramme, $p_{S_x}(s)$, est alors obtenue par intégration sur la

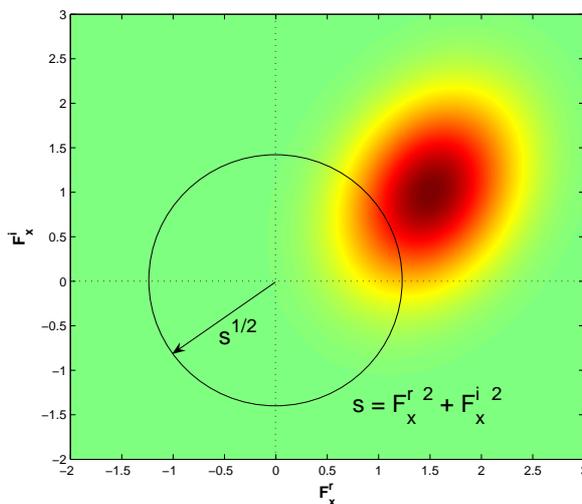


FIG. 2.7 – Illustration de l’approche géométrique utilisée pour le calcul de la densité de probabilité d’un module carré du vecteur \mathbf{F}_x à partir de sa densité de probabilité bidimensionnelle.

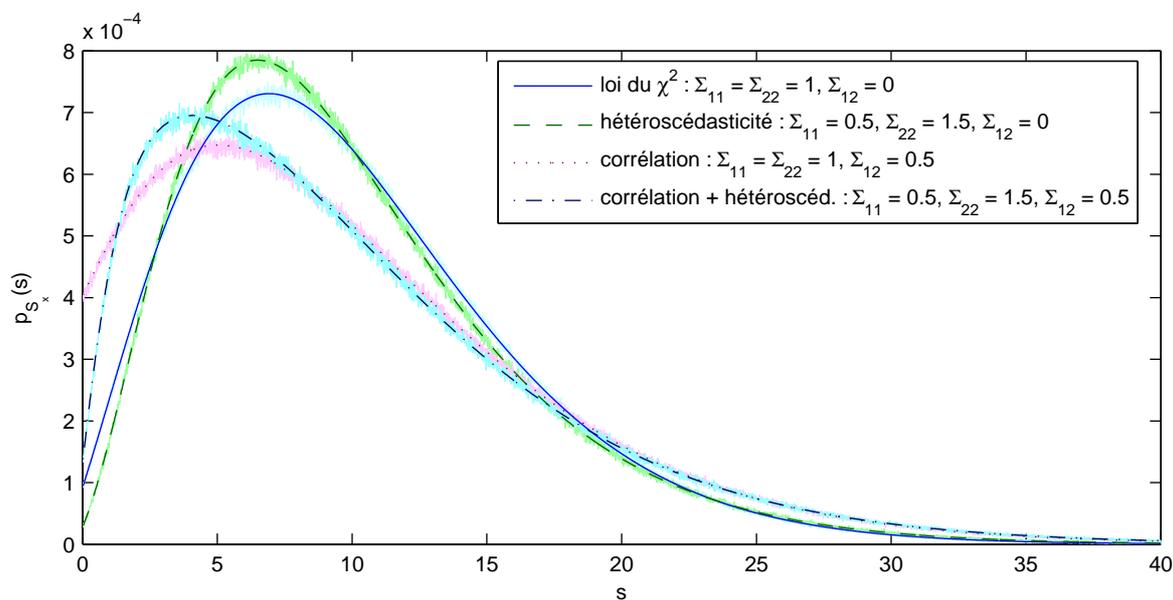


FIG. 2.8 – Exemples de distributions d’un module carré gaussien dans le cas décentré ($m_1^2 + m_2^2 = 8$) pour différentes matrices de variance-covariance Σ . Dans chaque cas, l’histogramme de 10^6 variables aléatoires équivalentes est également tracé en fond.

variable angulaire entre 0 et 2π et s'écrit

$$p_{S_x}(s) = \int_0^{2\pi} p_{(F_x^r, F_x^i)}(\sqrt{s}, \phi) d\phi, s \geq 0. \quad (2.27)$$

Dans notre cas, $p_{(F_x^r, F_x^i)}$ est une distribution gaussienne bidimensionnelle qui s'écrit en coordonnées polaires

$$p_{(F_x^r, F_x^i)}(r, \phi) = \frac{1}{2\pi \sqrt{\Sigma_{11}\Sigma_{22}(1-\rho^2)}} \exp\left(-\frac{A(r, \phi)}{2(1-\rho^2)}\right), \quad (2.28)$$

avec ρ le coefficient de corrélation (2.18) entre les parties réelle et imaginaire de la TFCT et

$$A(r, \phi) = \frac{(r \cos \phi - m_1)^2}{\Sigma_{11}} + \frac{(r \sin \phi - m_2)^2}{\Sigma_{22}} - \frac{2\rho(r \cos \phi - m_1)(r \sin \phi - m_2)}{\sqrt{\Sigma_{11}\Sigma_{22}}}. \quad (2.29)$$

Dans ce document, les densités de probabilité $p_{S_x}(s)$ sont calculées par discrétisation de l'équation (2.27). Typiquement, l'intervalle $[0, 2\pi]$ sur lequel l'intégration est effectuée est découpé en 2000 points équirépartis. Quelques exemples de distributions obtenues pour différentes matrices de variance-covariance Σ sont affichés sur la figure 2.8. Pour chaque cas, l'histogramme de 10^6 variables aléatoires équivalentes est affiché en arrière plan afin de justifier la validité de la méthode de calcul. De façon similaire au cas centré (figure 2.6), la présence de corrélation et/ou d'hétéroscédasticité tend à ramener les probabilités vers l'origine par rapport à la loi du χ^2 .

2.6 Le spectrogramme comme variable du χ^2

Nous avons montré dans les sections précédentes que le spectrogramme d'un signal gaussien n'était pas toujours exactement une variable du χ^2 , notamment lorsque le signal présente de la corrélation ou des non-stationnarités abruptes. Toutefois, plusieurs raisons incitent à utiliser une loi du χ^2 pour décrire la distribution de probabilité du spectrogramme :

- Les non-circularités observées sur la TFCT ne semblent pas très importantes dès lors qu'une fenêtre non rectangulaire est utilisée.
- L'expression générale de la densité de probabilité du spectrogramme, formulée en (2.27), limite les possibilités de développement analytique. En revanche, la loi du χ^2 possède une expression analytique plus simple et bénéficie de méthodes d'estimation de ses paramètres [JKB95][HM02].

Par ailleurs, la littérature faisant état d'une utilisation systématique de la loi du χ^2 pour décrire le spectrogramme, il nous a semblé intéressant d'évaluer les différences entre la loi du χ^2 d'une part et la loi du spectrogramme formulée dans le cas général d'autre part.

Afin d'effectuer cette comparaison, nous devons d'abord préciser comment les trois paramètres de la loi du χ^2 peuvent être définis. Ces paramètres pourront notamment être définis de manière à minimiser une distance entre les deux lois de probabilité. Ensuite, face aux différents contextes de corrélation, non-stationnarités et fenêtres d'analyse utilisés pour évaluer la circularité de la TFCT (cf. paragraphe 2.4), la différence entre la distribution du spectrogramme et la loi du χ^2 sera évaluée à l'aide de la divergence de Kullback-Leibler. Finalement, nous replacerons ces densités de probabilité dans le cadre du problème de détection dans le plan temps-fréquence qui nous concerne, afin d'évaluer les variations des performances du test de détection engendrée par l'utilisation de la loi du χ^2 à la place de la loi exacte du spectrogramme.

2.6.1 Quelle loi du χ^2 pour approcher le spectrogramme ?

Nous définissons ici comment les trois paramètres (δ, α, θ) de la loi du χ^2 (cf. paragraphe 2.3.1) peuvent être estimés afin d'approcher au mieux la densité de probabilité du spectrogramme. Il s'agit donc de mettre en relation les trois paramètres de la loi du χ^2 avec les cinq paramètres $(m_1, m_2, \Sigma_{11}, \Sigma_{22}, \Sigma_{12})$ qui caractérisent la loi du spectrogramme.

2.6.1.1 Paramètre de décentrage θ

Les paramètres influant sur la localisation des densités de probabilité sont le paramètre θ pour la loi du χ^2 et le couple (m_1, m_2) pour la loi du spectrogramme. En accord avec la définition de la loi du χ^2 (cf. paragraphe 2.3.1), nous gardons la relation

$$\theta = m_1^2 + m_2^2. \quad (2.30)$$

Remarque : le décentrage de la loi du spectrogramme est défini par les deux paramètres m_1 et m_2 et résumer ce décentrage par un unique paramètre θ induit nécessairement une perte d'information. L'information perdue est exactement la phase de la TFCT, perte inhérente à l'utilisation du spectrogramme en tant que module carré de la TFCT.

2.6.1.2 Degré de liberté δ et coefficient de proportionnalité α

Plusieurs solutions peuvent être envisagées pour relier les deux paramètres δ et α de la loi du χ^2 à la matrice de variance-covariance Σ .

Degré de liberté δ fixe

Une première solution consiste à maintenir le degré de liberté de la loi du χ^2 à la valeur fixe $\delta = 2$. C'est la solution retenue dans toute la littérature dans la mesure où elle découle naturellement de l'hypothèse de circularité de la TFCT. De plus, cette solution est très attractive car elle mène à la plus simple expression de la densité de probabilité de la loi du χ^2 . La valeur du coefficient de proportionnalité α peut alors être donnée par son estimateur au sens du maximum de vraisemblance ² et nous obtenons :

$$\begin{cases} \delta = 2, \\ \alpha = \frac{\Sigma_{11} + \Sigma_{22}}{2}. \end{cases} \quad (2.31)$$

La densité de probabilité de la loi du χ^2 associée à cette solution s'écrit :

$$p_{\chi^2}(s) = \frac{1}{\Sigma_{11} + \Sigma_{22}} \exp\left(-\frac{s + m_1^2 + m_2^2}{\Sigma_{11} + \Sigma_{22}}\right) I_0\left(\frac{2\sqrt{s(m_1^2 + m_2^2)}}{\Sigma_{11} + \Sigma_{22}}\right). \quad (2.32)$$

Degré de liberté δ adapté

Une seconde solution consiste à tenter de mieux prendre en compte la non-circularité de la TFCT en adaptant à la fois le coefficient de proportionnalité α et le degré de liberté δ à la matrice de variance-covariance Σ . La loi du χ^2 ne disposant que des deux paramètres δ et α pour traduire les trois paramètres Σ_{11} , Σ_{22} et Σ_{12} , une bijection exacte entre une loi du spectrogramme et une loi du χ^2 n'existe pas. Cependant, nous pouvons espérer une meilleure approximation de la loi du spectrogramme en adaptant les deux paramètres de la loi du χ^2 et en faisant notamment intervenir la covariance Σ_{12} des parties réelle et imaginaire de la TFCT.

²Cet estimateur est également celui qui minimise la divergence de Kullback-Leibler entre une loi du χ^2 centrée à 2 degrés de liberté et la loi du spectrogramme dans le cas centré (bruit seul).

Nous avons en premier lieu souhaité exprimer les solutions maximisant la vraisemblance ou minimisant la divergence de Kullback-Leibler entre les deux lois de probabilité. Malheureusement, les expressions des densités de probabilité du spectrogramme et d'un χ^2 de degré de liberté δ quelconque rendent ce calcul difficile à mener. Nous exposons ici la solution obtenue à partir de la méthode des moments qui consiste à égaliser les premiers moments statistiques des deux variables (χ^2 et spectrogramme) afin d'obtenir des équations liant les paramètres de la première variable à ceux de la seconde. Comme nous recherchons deux paramètres de la loi du χ^2 , cette méthode est appliquée sur les deux premiers moments.

Pour une variable du χ^2 centrée ($\theta = 0$), ces deux premiers moments sont

$$\begin{cases} \mathbb{E} \{ \chi^2 \} &= \alpha \delta, \\ \text{Var} \{ \chi^2 \} &= 2\alpha^2 \delta, \end{cases} \quad (2.33)$$

où $\text{Var} \{ \cdot \}$ désigne la variance. Les deux premiers moments du spectrogramme d'un signal centré s'écrivent quant à eux :

$$\begin{cases} \mathbb{E} \{ S_x \} &= \Sigma_{11} + \Sigma_{22}, \\ \text{Var} \{ S_x \} &= 2[\Sigma_{11}^2 + \Sigma_{22}^2 + 2\Sigma_{12}^2]. \end{cases} \quad (2.34)$$

L'égalisation des moments respectifs mène alors à formuler les paramètres δ et α de la loi du χ^2 de la façon suivante :

$$\begin{cases} \delta &= \frac{(\Sigma_{11} + \Sigma_{22})^2}{\Sigma_{11}^2 + \Sigma_{22}^2 + 2\Sigma_{12}^2}, \\ \alpha &= \frac{\Sigma_{11} + \Sigma_{22}}{\delta}. \end{cases} \quad (2.35)$$

et la densité de probabilité de cette variable du χ^2 est finalement obtenue selon (2.10).

Afin d'examiner cette nouvelle définition des paramètres de la loi du χ^2 , introduisons le paramètre ξ tel que :

$$\xi = \frac{\frac{1}{2}(\Sigma_{11} + \Sigma_{22})}{\sqrt{\Sigma_{11}\Sigma_{22}}}. \quad (2.36)$$

Ce paramètre reflète le niveau d'hétéroscédasticité entre les parties réelle et imaginaire de la TFCT. Il vaut 1 lorsque les variances Σ_{11} et Σ_{22} sont égales et tend lentement vers l'infini lorsque ces variances diffèrent. En utilisant ce paramètre, le degré de liberté δ exprimé selon (2.35) peut se récrire :

$$\delta = \frac{2}{2 - \frac{1-\rho^2}{\xi^2}}, \quad (2.37)$$

dans lequel nous retrouvons le coefficient de corrélation ρ entre les parties réelle et imaginaire de la TFCT. Cette formulation, illustré sur la figure 2.9, fait ressortir l'influence de la corrélation et de l'hétéroscédasticité qui agissent en diminuant le degré de liberté δ .

Interprétation :

Le degré de liberté reflète le nombre de variables gaussiennes, indépendantes et de même variance, qui forment une variable du χ^2 . Lorsque les parties réelle et imaginaire de la TFCT sont corrélées, le nombre équivalent de variables indépendantes est intuitivement inférieur à deux.

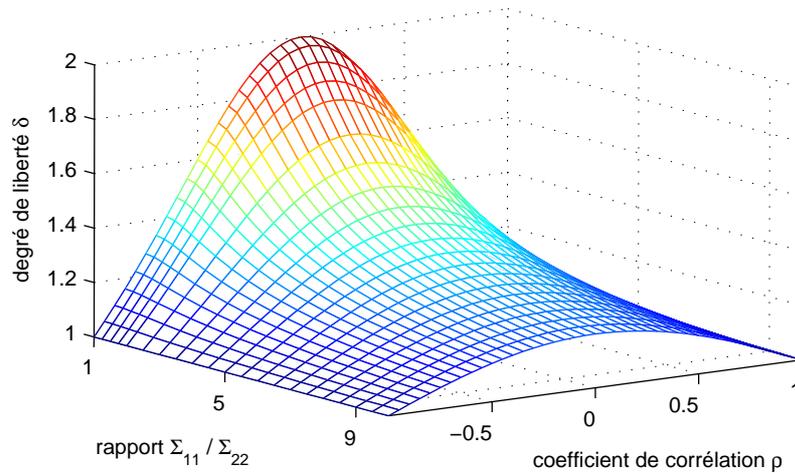


FIG. 2.9 – Degré de liberté δ (2.37) en fonction du coefficient de corrélation ρ et du rapport des variances $h = \Sigma_{11}/\Sigma_{22}$ entre les parties réelle et imaginaire de la TFCT.

Ce comportement est vérifié par le degré de liberté (2.37) qui évolue de 2 à 1 lorsque le coefficient de corrélation ρ évolue de 0 à ± 1 (cf. figure 2.9).

De même, lorsque l'une des variables gaussiennes possède une variance supérieure à son homologue, la somme des carrés de ces deux variables n'est plus influencée de façon équivalente par chacune des variables. Le nombre de variables équivalentes participant à cette somme devient alors inférieur à deux. Le degré de liberté (2.37) satisfait là encore l'intuition en évoluant de 2 vers 1 lorsque l'hétéroscédasticité ξ , ou de manière équivalente le rapport $h = \Sigma_{11}/\Sigma_{22}$, évoluent de 1 vers l'infini. La définition du degré de liberté donnée en (2.35) bénéficie ainsi d'une interprétation en accord avec l'intuition.

2.6.2 Mesure de divergence

Deux lois du χ^2 susceptibles de décrire la statistique du spectrogramme ont été définies dans le paragraphe précédent. L'objet de ce paragraphe est de mesurer l'adéquation entre ces deux lois du χ^2 et la distribution du spectrogramme. Nous nous intéressons aux cas pour lesquels la circularité de la TFCT n'est plus tout à fait exacte, c'est-à-dire lorsque le signal analysé contient de la corrélation, des non-stationnarités, et pour différentes fenêtres d'analyse temps-fréquence. A ce titre, le modèle d'une corrélation exponentiellement décroissante (de paramètre λ) ainsi que le profil de non-stationnarité employés au paragraphe 2.4 sont repris pour cette étude.

L'adéquation ou la dissimilarité entre deux lois de probabilité p et q peuvent être mesurées avec la divergence de Kullback-Leibler (KL), notée $D_{KL}(p, q)$. Cette mesure est positive et égale à zéro lorsque les lois p et q sont identiques. La divergence KL entre la distribution du spectrogramme, notée p_{S_x} , et la loi du χ^2 , notée p_{χ^2} , se définit comme [CT91] :

$$D_{KL}(p_{S_x}, p_{\chi^2}) = \int p_{S_x}(x) \log \left[\frac{p_{S_x}(x)}{p_{\chi^2}(x)} \right] dx. \quad (2.38)$$

Le logarithme népérien est utilisé pour cette expression et les divergences KL sont exprimées en 'nat'.

Pour un signal stationnaire et de corrélation exponentiellement décroissante, la figure 2.10 présente l'évolution de la divergence KL en fonction de la taille M de la fenêtre d'analyse pour

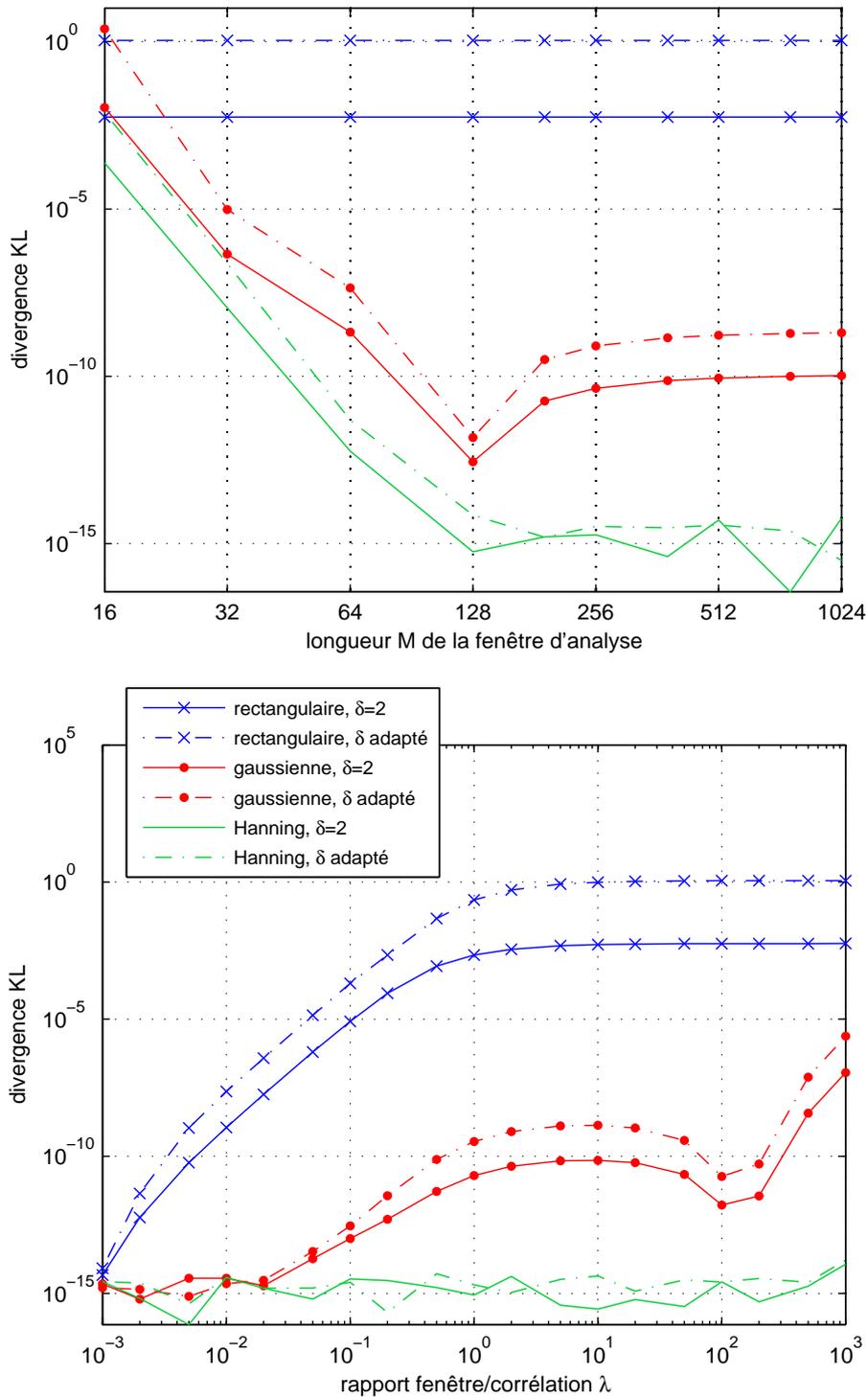


FIG. 2.10 – *Cas non blanc et stationnaire*. En haut : divergence KL entre les lois du χ^2 et la distribution du spectrogramme en fonction de la longueur M de la fenêtre d'analyse. Le rapport fenêtre-corrélation du signal est maintenu à $\lambda = 30$ et la fréquence observée est $k = M/8$. En bas : divergence KL entre les lois du χ^2 et la distribution du spectrogramme en fonction du degré de corrélation du signal analysé. Les fenêtres sont de 256 points et la fréquence observée est $k = M/8 = 32$.

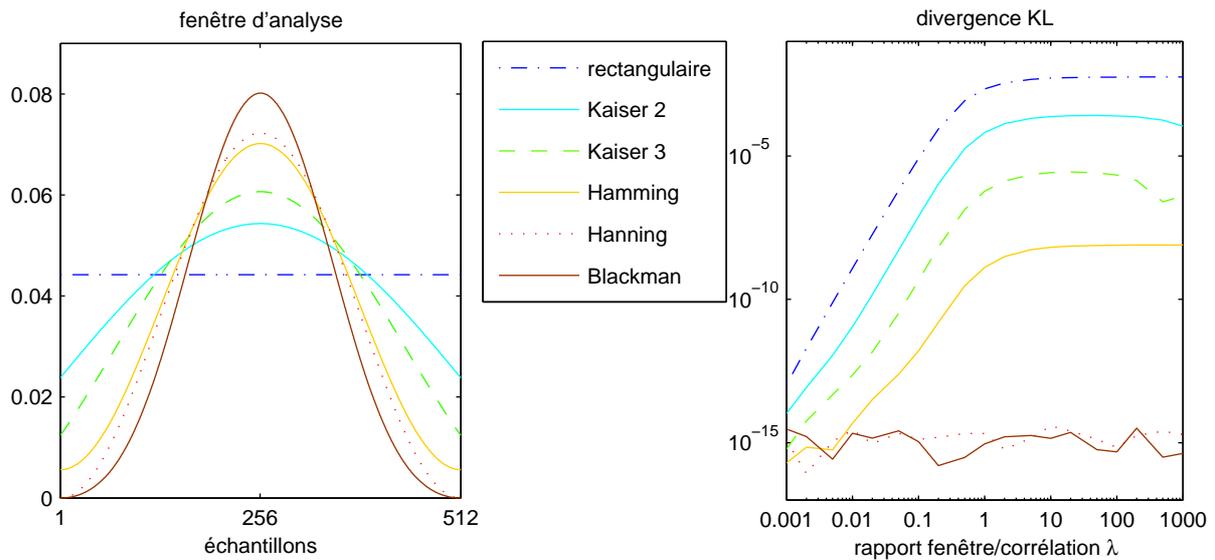


FIG. 2.11 – Profils de fenêtre d'analyse et divergence KL entre la loi du χ^2 à 2 degré de liberté et la distribution du spectrogramme en fonction du rapport fenêtre-corrélation λ . La fréquence observée est $k = M/8 = 64$.

la figure du haut, et en fonction du rapport fenêtre-corrélation λ défini en (2.20) pour la figure du bas. A l'exception de la fenêtre rectangulaire pour laquelle la taille de la fenêtre n'a pas d'impact, l'utilisation d'une longue fenêtre d'analyse améliore l'adéquation entre la loi du χ^2 et la distribution du spectrogramme (figure 2.10, en haut). Aussi, la loi du χ^2 décrit de moins en moins bien la distribution du spectrogramme lorsque le signal est de plus en plus corrélé (figure 2.10, en bas). Cependant, si la sensibilité à la corrélation est forte pour la fenêtre rectangulaire, elle l'est déjà moins pour la fenêtre gaussienne, et inexistante pour la fenêtre de Hanning. Nous retrouvons ici le phénomène observé et discuté aux paragraphes 2.4.1.3 et 2.4.2 : face à un signal corrélé, la circularité de la TFCT est liée à la valeur de la fenêtre d'analyse à ses extrémités. À ce titre, la figure 2.11 présente les profils de divergence pour six fenêtres d'analyse admettant différentes valeurs à leurs extrémités : plus la valeur aux extrémités est faible, moins l'adéquation χ^2 -spectrogramme est sensible à la corrélation du signal.

La figure 2.12 présente l'évolution de la divergence KL entre la loi du χ^2 et la distribution du spectrogramme en fonction de l'amplitude (figure du haut) et du caractère abrupt (figure du bas) d'une non-stationnarité présente dans l'intervalle de la fenêtre d'analyse (le profil de la non-stationnarité et ses paramètres sont donnés sur la figure 2.2). Pour les trois fenêtres, la divergence KL croît avec l'amplitude de la non-stationnarité (figure du haut). Lorsque la non-stationnarité devient de moins en moins abrupte (figure du bas), l'adéquation spectrogramme- χ^2 s'améliore rapidement pour les fenêtres non-rectangulaires, contrairement au cas de la fenêtre rectangulaire pour lequel la divergence KL reste constante. Cependant, notons qu'en présence d'une non-stationnarité fortement abrupte ($L < 6$ échantillons sur la figure), la fenêtre rectangulaire engendre la meilleure adéquation entre loi du χ^2 et distribution du spectrogramme.

Le troisième point d'intérêt de ces mesures de divergence concerne la paramétrisation de la loi du χ^2 . Deux possibilités ont été proposées : dans la première, le degré de liberté de la loi du χ^2 est fixé à $\delta = 2$. Dans la seconde, ce même degré de liberté est adapté à la matrice de

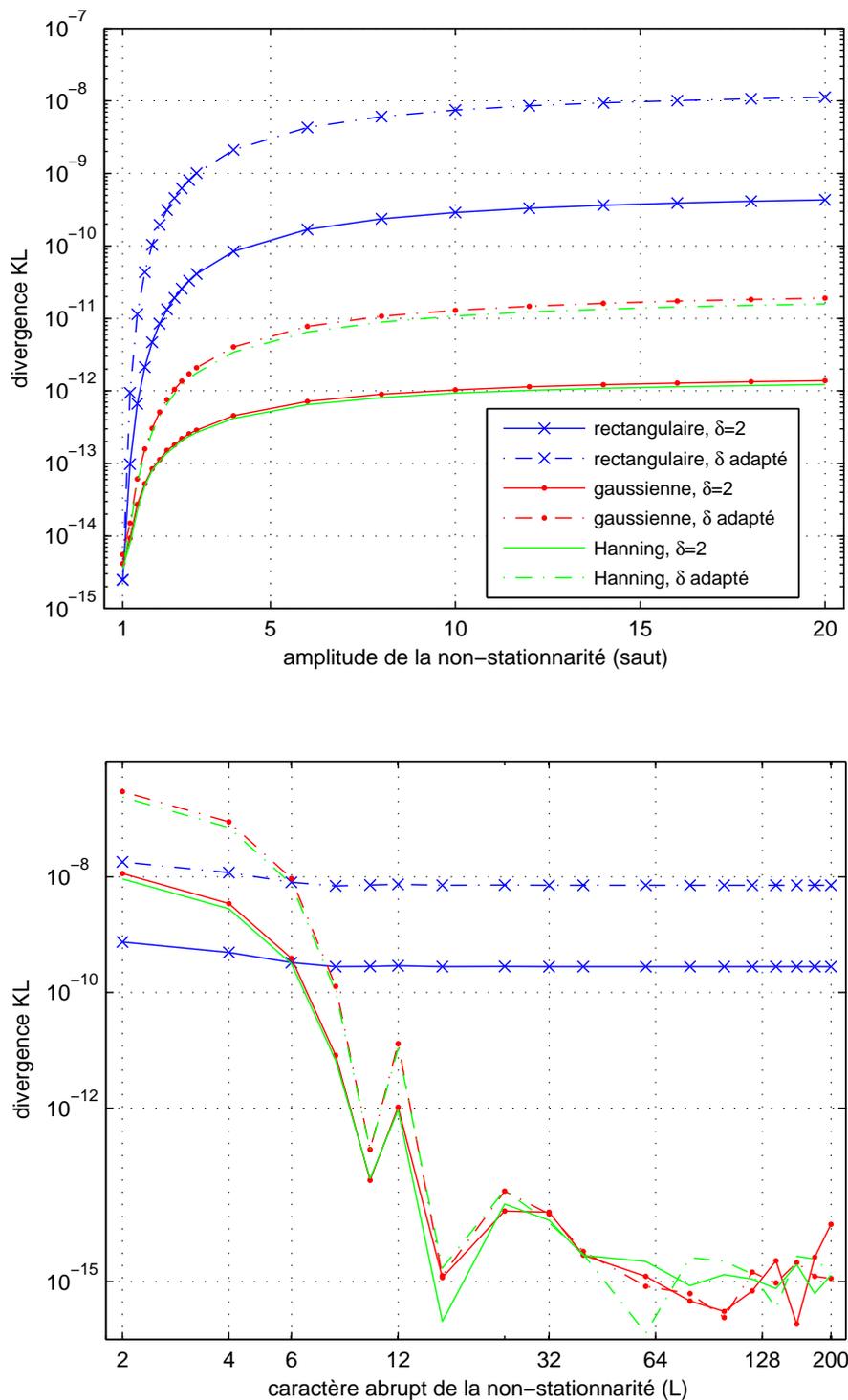


FIG. 2.12 – *Cas blanc et non stationnaire*. Divergence KL entre les lois du χ^2 et la distribution du spectrogramme en fonction de l'amplitude *saut* de la non-stationnarité pour $L = 12$ (en haut) et de sa largeur L pour *saut* = 10 (en bas). Les fenêtres sont de taille $M = 256$ points et la fréquence observée est $k = 20$.

variance-covariance Σ . En opposition avec l'effet escompté, adapter le degré de liberté de la loi du χ^2 ne permet pas une meilleure adéquation avec la distribution du spectrogramme. Dans les quatre cas de figure présentés (figures 2.10 et 2.12), la plus petite divergence KL est obtenue avec un degré de liberté fixé à $\delta = 2$. Nous reviendrons sur ce point à la fin du paragraphe 2.6.3.

Conclusion sur la mesure de divergence : cette étude basée sur la divergence de Kullback-Leibler permet de confirmer l'existence de contextes pour lesquels les coefficients du spectrogramme ne sont pas distribués selon une loi du χ^2 . De manière générale, nous pouvons conclure que la divergence augmente avec la corrélation et la non-stationnarité du signal. Cependant, cette conclusion doit être modérée par le choix de la fenêtre d'analyse : si une fenêtre rectangulaire impose à la distribution du spectrogramme une grande sensibilité aux signaux corrélés ou non stationnaires, l'utilisation de fenêtres non-uniformes réduit considérablement cette sensibilité. En particulier, en présence de corrélation, les fenêtres régulières admettant une valeur nulle à leurs extrémités (par exemple les fenêtres de Hanning ou Blackman) assurent une divergence faible et constante entre la loi du χ^2 et la distribution du spectrogramme.

Si les mesures de divergence permettent d'observer le comportement de la loi du spectrogramme vis à vis de différents signaux, et de mesurer sa dissimilarité avec une loi du χ^2 , une interprétation quantifiée des résultats reste difficile. Que signifie une divergence de 10^{-10} nats ? Peut-on malgré tout utiliser une loi du χ^2 pour décrire le spectrogramme ? Dans la section suivante, nous apportons quelques éléments de réponses dans le cadre d'un problème de détection.

2.6.3 Impact de l'approximation dans un contexte de détection

Nous reprenons ici le problème de détection dans le plan temps-fréquence tel qu'il a été formulé dans le chapitre précédent. L'objet de cette partie est d'évaluer l'impact de l'utilisation d'une loi du χ^2 pour décrire le spectrogramme sur les performances de ce test.

Dans l'approche de Neyman-Pearson, le choix d'une probabilité de fausse alarme P_{fa} permet de calculer le seuil de détection S^{seuil} en fonction de la densité de probabilité p_{H_0} des observations sous l'hypothèse H_0 . Cependant, l'utilisation d'une approximation p_{ap} de la densité de probabilité sous H_0 entraîne une variation des performances du test : la PFA réelle du test, notée $P_{fa,réelle}$, est différente de la PFA initialement désirée, notée $P_{fa,désirée}$. Plus précisément, le seuil de détection S^{seuil} est calculé de façon à satisfaire l'équation

$$P_{fa,désirée} = \int_{S^{seuil}}^{+\infty} p_{ap}(s) ds \quad (2.39)$$

dans laquelle l'approximation p_{ap} est utilisée. En revanche, les observations étant quant à elles distribuées selon la loi p_{H_0} , la PFA réelle du test devient

$$P_{fa,réelle} = \int_{S^{seuil}}^{+\infty} p_{H_0}(s) ds. \quad (2.40)$$

La variation de PFA, notée Δ_{PFA} , engendrée par l'approximation de p_{H_0} par la loi p_{ap} est alors évaluée selon

$$\Delta_{PFA} = |P_{fa,réelle} - P_{fa,désirée}|, \quad (2.41)$$

$$= \int_{S^{seuil}}^{+\infty} |p_{H_0}(s) - p_{ap}(s)| ds. \quad (2.42)$$

Cette variation de PFA peut servir à évaluer si l'approximation p_{ap} est acceptable ou non. Par exemple, si cette variation est supérieure à la PFA initialement souhaitée, nous pouvons

conclure que l'approximation de la loi du spectrogramme engendre une erreur supérieure à la précision souhaitée pour le test. Dans ce cas, l'approximation doit être rejetée. Dans la suite, nous calculons le taux d'erreur sur la PFA ϵ_{PFA} , définie selon

$$\epsilon_{PFA} = \frac{\Delta_{PFA}}{P_{fa,désirée}}, \quad (2.43)$$

pour évaluer l'approximation de la distribution du spectrogramme par une loi du χ^2 .

La figure 2.13 donne l'évolution du taux d'erreur ϵ_{PFA} en fonction de la taille de la fenêtre d'analyse M (en haut) et du rapport fenêtre-corrélation λ (en bas) pour un signal stationnaire et exponentiellement corrélé. Notons que nous avons observé une invariance de ces courbes face au choix de la PFA souhaitée $P_{fa,désirée}$. Pour la fenêtre de Hanning, une très courte fenêtre peut engendrer une erreur de PFA non négligeable (ici pour $M < 32$, figure du haut). Pour des fenêtres plus longues, le taux d'erreur ϵ_{PFA} reste très inférieur à 1, et ce quel que soit le degré de corrélation du signal (figure du bas). Pour la fenêtre gaussienne, le taux d'erreur ϵ_{PFA} se stabilise autour de 10^{-3} lorsque la longueur de la fenêtre augmente ou lorsque la longueur de corrélation du signal devient de l'ordre de grandeur de la taille de la fenêtre ($\lambda = 1$). L'erreur engendrée par l'approximation de la distribution du spectrogramme par une loi du χ^2 apparaît donc faible. Pour une fenêtre rectangulaire en revanche, le taux d'erreur devient supérieur à 1 lorsque le rapport fenêtre-corrélation est supérieur à $\lambda = 0.1$ (figure du bas). De plus, l'utilisation d'une fenêtre plus longue ne permet pas de réduire ce taux d'erreur (figure du haut). Ici l'utilisation d'une loi du χ^2 n'est pas acceptable.

Pour les signaux blancs et non-stationnaires, la figure 2.14 montre que quelle que soit la non-stationnarité, l'utilisation d'une loi du χ^2 engendre des taux d'erreur inférieurs ou de l'ordre de 10^{-3} . Ces erreurs peuvent être considérées comme négligeable et la loi du χ^2 peut être utilisée quelle que soit la fenêtre d'analyse choisie.

Curieusement, si l'adaptation du degré de liberté de la loi du χ^2 n'est pas bénéfique face à la mesure de divergence KL, elle permet ici de réduire le taux d'erreur de PFA. Dans toutes les situations (figures 2.13 et 2.14), le taux d'erreur ϵ_{PFA} introduit par la loi du χ^2 est légèrement inférieur lorsque le degré de liberté est adapté selon (2.35) plutôt que fixé à 2. L'explication, illustrée sur la figure 2.15, en est qu'adapter le degré de liberté de la loi du χ^2 permet une meilleure approximation de la queue de la distribution du spectrogramme. Or la mesure de PFA est spécifiquement liée à cette zone de la distribution. En revanche, une loi du χ^2 à 2 degré de liberté permet de mieux approcher le corps principal de la distribution du spectrogramme, zone à laquelle la divergence de Kullback-Leibler est particulièrement sensible³.

Conclusion sur l'erreur de PFA : cette étude du taux d'erreur de PFA introduit par l'approximation de la distribution du spectrogramme par une loi du χ^2 nous permet de tirer des conclusions pratiques sur l'utilisabilité de la loi du χ^2 dans un problème de détection temps-fréquence. En se basant sur un taux d'erreur ϵ_{PFA} inférieur à 10^{-3} , nous pouvons conclure qu'une loi du χ^2 peut toujours être utilisée lorsque le spectrogramme est construit avec une fenêtre non-rectangulaire de plus de 32 points. Les variations de la distribution du spectrogramme relatives

³La divergence KL correspond à l'espérance du rapport entre deux lois de probabilité (et plus précisément son logarithme). Le rapport donne une mesure de différence, l'espérance introduit une pondération. Dans la mesure de divergence, les différences entre les deux lois sont ainsi pondérées par la densité de probabilité de l'une des lois, disons la loi de référence. La pondération est donc importante là où la loi de référence est importante, soit sur son corps principal. La mesure de divergence est ainsi sensible aux différences qui interviennent dans cette zone. À l'inverse, les différences qui apparaissent sur des zones où la loi de référence est faible, et notamment sa queue, reçoivent un faible poids et ont peu d'impact sur la divergence KL. Mais elles sont plus rares...

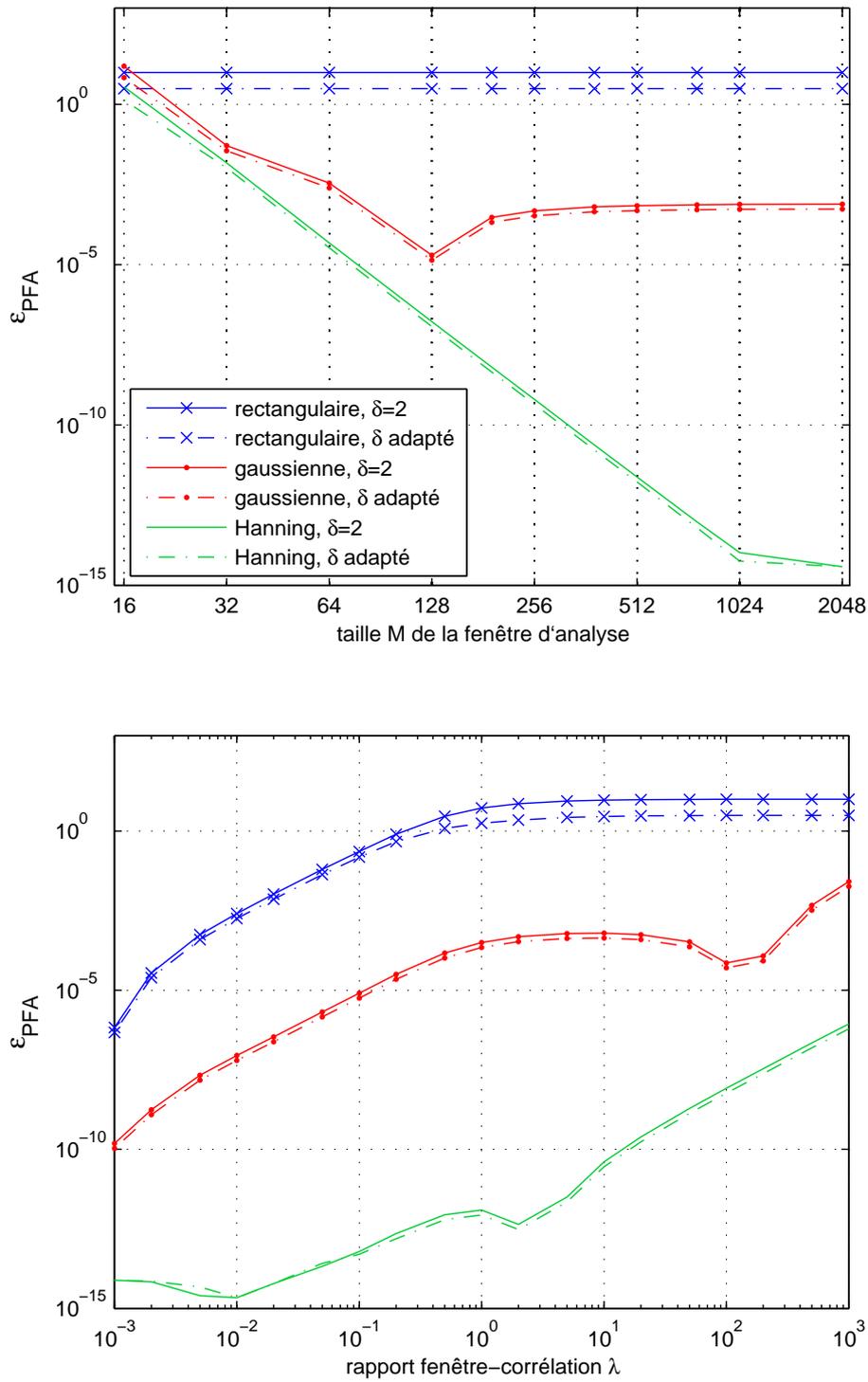


FIG. 2.13 – *Cas non blanc et stationnaire*. Taux d'erreur sur la PFA ϵ_{PFA} engendrée par l'utilisation d'une loi du χ^2 pour approcher la distribution du spectrogramme d'un signal exponentiellement corrélé en fonction de la longueur M de la fenêtre d'analyse pour $\lambda = 30$ (en haut) et du degré de corrélation λ du signal pour une fenêtre de $M = 256$. La fréquence observée est $k = M/8$.

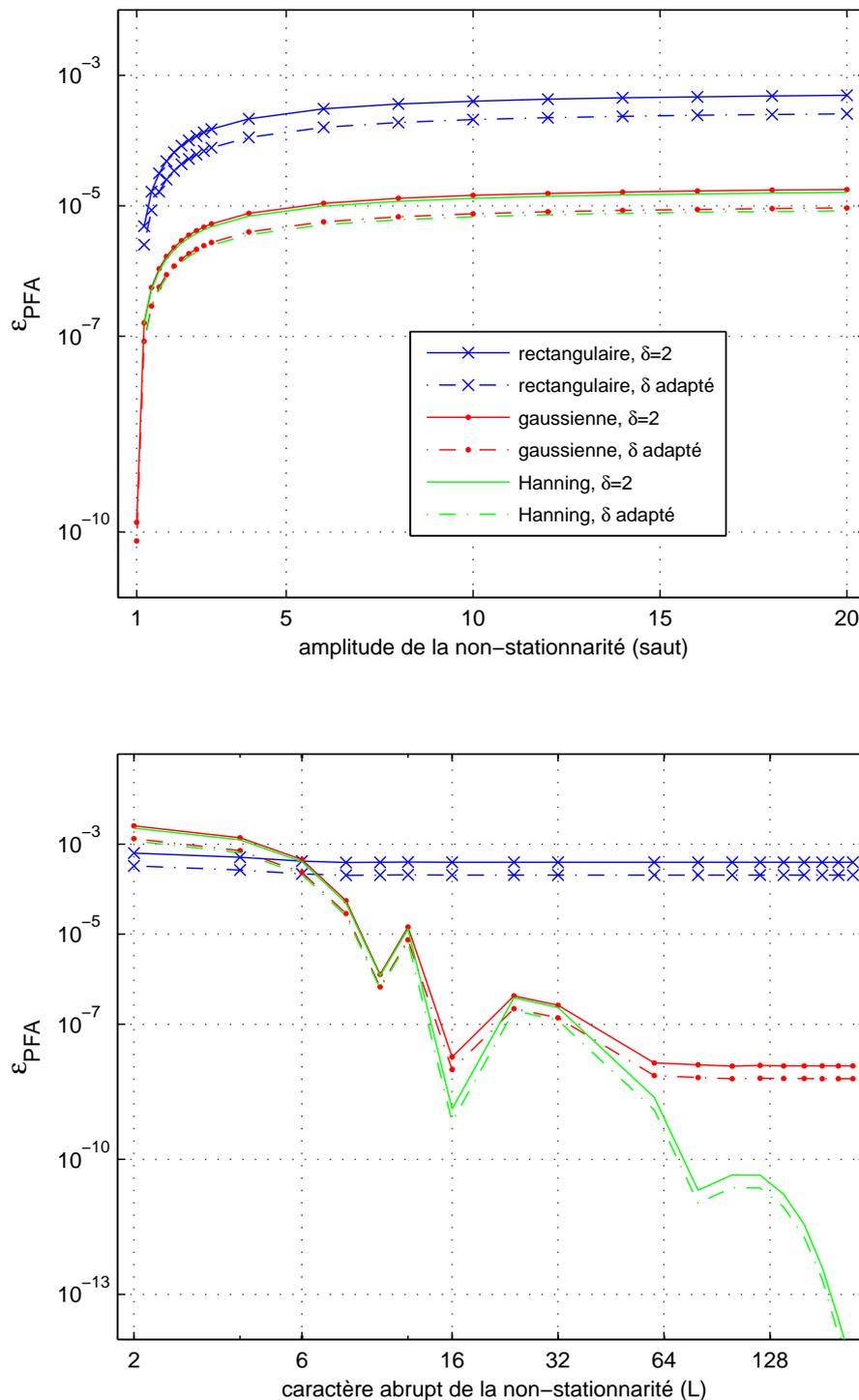


FIG. 2.14 – *Cas blanc et non stationnaire*. Taux d'erreur sur la PFA ϵ_{PFA} engendrée par l'utilisation d'une loi du χ^2 pour approcher la distribution du spectrogramme d'un signal blanc et non-stationnaire en fonction de l'amplitude *saut* pour $L = 12$ (en haut) et de la largeur L pour *saut* = 10 (en bas) de la non-stationnarité. Les fenêtres sont de $M = 256$ échantillons et la fréquence observée est $k = 20$.

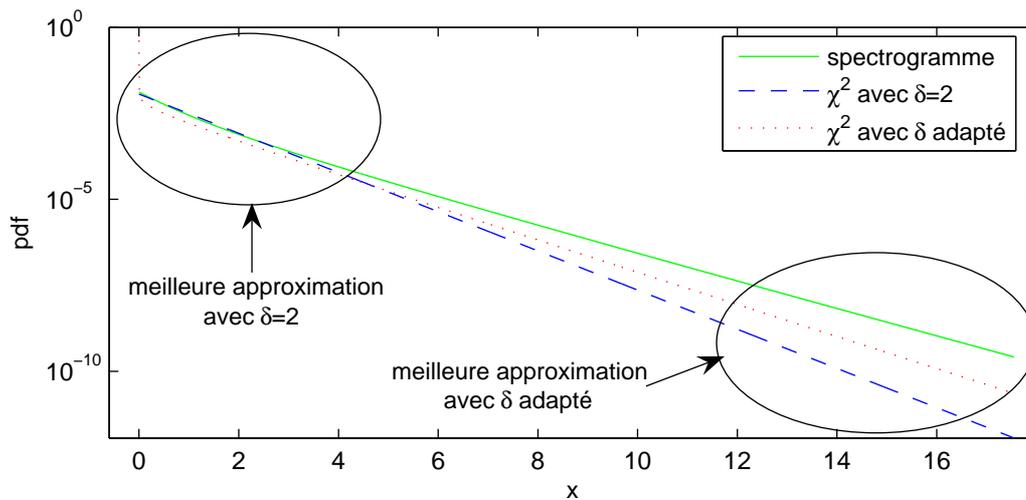


FIG. 2.15 – Densités de probabilité du spectrogramme d’un signal gaussien, centré et exponentiellement corrélé ($\lambda = 10$) et lois du χ^2 avec degré de liberté fixe ($\delta = 2$) ou adapté. Le degré de liberté fixé à 2 engendre la meilleure approximation du corps principal de la distribution, le degré de liberté adapté permet une meilleure approximation de la queue de la distribution.

à des signaux non-stationnaires ou fortement corrélés n’ont qu’un impact négligeable sur un cas pratique de détection. Lorsque la fenêtre rectangulaire est utilisée (pour l’analyse de signaux très courts par exemple), l’utilisation de la loi du χ^2 pour décrire les statistiques du spectrogramme dans un test de détection doit être soumise aux conditions de non-stationnarité et de corrélation des signaux à analyser.

2.7 Conclusion

La mise en place d’un test de détection requiert la connaissance de la densité de probabilité de la statistique de décision sous chacune des hypothèses envisagées. Dans le cas étudié, la statistique de décision correspond à un coefficient du spectrogramme et l’étude de sa densité de probabilité a fait l’objet de ce chapitre. Plus précisément, notre question était de savoir si la loi du χ^2 à 2 degrés de liberté, unanimement utilisée pour le traitement statistique du module carré de la transformée de Fourier d’un signal gaussien, restait toujours valable en présence de signaux fortement corrélés ou non-stationnaires. Cela n’étant pas le cas, nous avons voulu évaluer l’erreur introduite par l’utilisation de la loi du χ^2 pour décrire le spectrogramme, notamment dans le contexte de détection dans le plan temps-fréquence. L’étude présentée dans ce chapitre mène aux conclusions suivantes :

La distribution d’un coefficient du spectrogramme n’est pas une loi du χ^2 dès lors que le signal analysé est corrélé ou non-stationnaire. La distribution du spectrogramme est alors plus généralement celle du module carré d’une variable gaussienne bidimensionnelle.

La sensibilité de la distribution du spectrogramme à la corrélation ou aux non-stationnarités est fortement liée à la forme de la fenêtre d’analyse utilisée pour construire le spectrogramme. Les fenêtres non-uniformes assurent une meilleure robustesse à ces conditions, et plus précisément, le degré de sensibilité semble lié à la régularité de la fenêtre, à la valeur qu’elle présente à ses extrémités ainsi qu’à la valeur de sa dérivée à ces extrémités. Une grande robustesse est ainsi observée pour les fenêtres de Hanning et Blackman.

Si en présence de corrélation ou de non-stationnarités le spectrogramme n’est pas exactement

distribué selon une loi du χ^2 , l'utilisation de cette loi est cependant possible si une fenêtre non-rectangulaire est utilisée. Dans un contexte de détection et utilisant l'approche de Neyman-Pearson, l'erreur introduite par l'utilisation d'une loi du χ^2 à 2 degrés de liberté est apparue négligeable (toujours inférieure à 10^{-3}) pour les fenêtres non-uniformes. Si la fenêtre rectangulaire est utilisée, l'utilisation de la loi du χ^2 doit en revanche être soumise aux degrés de corrélation et de non-stationnarité des signaux analysés.

Dans la suite de ce travail, les coefficients $S_x[n, k]$ du spectrogramme du signal observé $x[m]$ (1.1) seront donc assimilés à des variables du χ^2 à $\delta = 2$ degrés de liberté, de coefficient de proportionnalité $\alpha = (\Sigma_{11} + \Sigma_{22})/2$ et de paramètre de décentrage $\theta = m_1^2 + m_2^2$. Les deux remarques suivantes permettent d'alléger les écritures et de redonner la signification des paramètres :

Comme nous l'avons vu en (2.4), les moyennes m_1 et m_2 sont respectivement les parties réelle et imaginaire de la TFCT du signal d'intérêt $d[m]$. Nous avons donc la relation

$$\begin{aligned}\theta &= m_1^2 + m_2^2 \\ &= F_d^r[n, k]^2 + F_d^i[n, k]^2 \\ &= S_d[n, k],\end{aligned}\tag{2.44}$$

où $S_d[n, k]$ est le coefficient au point $[n, k]$ du spectrogramme du signal d'intérêt $d[m]$.

En revenant aux définitions 2.3, nous avons

$$\begin{aligned}\Sigma_{11} + \Sigma_{22} &= \mathbb{E} \{ F_p^r[n, k]^2 \} + \mathbb{E} \{ F_p^i[n, k]^2 \} \\ &= \mathbb{E} \{ S_p[n, k] \} \\ &= \gamma_p[n, k],\end{aligned}\tag{2.45}$$

où $\gamma_p[n, k]$ est la densité spectrale de puissance de la perturbation $p[m]$ au point $[n, k]$. Le coefficient de proportionnalité α se réécrit alors

$$\begin{aligned}\alpha &= \frac{\Sigma_{11} + \Sigma_{22}}{2} \\ &= \frac{\gamma_p[n, k]}{2}.\end{aligned}\tag{2.46}$$

Ainsi, le coefficient $S_x[n, k]$ est une variable du χ^2 à 2 degrés de liberté, de coefficient de proportionnalité $\gamma_p[n, k]/2$ et de paramètre de décentrage $S_d[n, k]$. Dans le cadre du problème de détection formulé dans le chapitre précédent (1.5), les densités de probabilité p_{H_0} et p_{H_1} des observations sous les hypothèses H_0 et H_1 s'écrivent finalement en chaque point temps-fréquence $[n, k]$:

$$p_{H_0, [n, k]}(s) = \frac{1}{\gamma_p[n, k]} \exp\left(-\frac{s}{\gamma_p[n, k]}\right),\tag{2.47}$$

$$p_{H_1, [n, k]}(s) = \frac{1}{\gamma_p[n, k]} \exp\left(-\frac{s + S_d[n, k]}{\gamma_p[n, k]}\right) I_0\left(\frac{2}{\gamma_p[n, k]} \sqrt{S_d[n, k] s}\right).\tag{2.48}$$

Si la forme des lois de probabilité des observations est maintenant déterminée, il reste encore une inconnue avant de pouvoir résoudre le problème de détection temps-fréquence. La densité spectrale de puissance $\gamma_p[n, k]$ de la perturbation gaussienne n'étant pas supposée connue, elle est l'objet du problème d'estimation auquel se consacre le chapitre suivant.

Chapitre 3

Estimation de la densité spectrale de puissance du bruit

Sommaire

3.1	Introduction	51
3.2	Positionnement du problème	51
3.2.1	Les méthodes itératives	52
3.2.2	Les méthodes de lissage	52
3.2.3	Les statistiques minimales	53
3.3	Contribution : estimation par les Z valeurs minimales	54
3.3.1	Formulation	54
3.3.2	Estimateurs proposés	55
3.3.2.1	Utilisation de la fonction de répartition	55
3.3.2.2	Utilisation de la loi des minima	57
3.3.3	Évaluation des estimateurs	58
3.3.3.1	Biais et variance	58
3.3.3.2	Sensibilité aux paramètres	58
3.3.3.3	Récapitulatif	61
3.3.4	Détermination du voisinage et des paramètres Z et N	62
3.3.4.1	Détermination du voisinage	62
3.3.4.2	Détermination du paramètre Z	63
3.3.4.3	Estimation du paramètre N	63
3.4	Exemples d'estimation dans le plan temps-fréquence	66
3.4.1	Perturbation blanche et stationnaire	67
3.4.2	Perturbation blanche ou stationnaire	69
3.4.3	Perturbation non blanche et non stationnaire	69
3.5	Conclusion	78

3.1 Introduction

Dans le chapitre précédent, nous avons vu que les statistiques d'un coefficient du spectrogramme pouvaient être décrites par une loi du χ^2 . Replacées dans le contexte de détection formulé dans le chapitre 1, les lois sous les hypothèses "bruit seul" et "signal+bruit", H_0 et H_1 , sont formulées en (2.47) et (2.48). En chaque point $[n, k]$ du plan temps-fréquence, ces lois sont caractérisées par deux paramètres originaires du signal $x[m]$ observé : le spectrogramme $S_d[n, k]$ du signal d'intérêt $d[m]$ et la densité spectrale non stationnaire $\gamma_p[n, k]$ de la perturbation $p[m]$ (plus précisément son spectre physique). Le test de détection proposé vise à déterminer si le premier paramètre $S_d[n, k]$ est nul (hypothèse H_0) ou non (hypothèse H_1). Afin de mettre en place ce test, il est nécessaire de connaître la densité spectrale de puissance $\gamma_p[n, k]$ de la perturbation. Dans le cas où ce paramètre n'est pas connu *a priori*, il faut l'estimer, et cette question fait l'objet de ce troisième chapitre.

Dans un premier temps, nous décrivons les enjeux et les contraintes relatifs à ce problème d'estimation et présentons trois approches rencontrées dans la littérature afin de le résoudre. Dans un second temps, nous présentons notre contribution à ce problème. Nous formulons un estimateur de la densité spectrale de puissance du bruit basé sur la statistique des plus petits coefficients rencontrés dans une régions du spectrogramme. Cet estimateur est ensuite évalué en terme de biais et de variance et la détermination de ses paramètres constitutifs est discutée. Finalement, nous présentons dans un troisième temps quelques résultats d'estimation sur des signaux synthétiques et naturels.

3.2 Positionnement du problème

Soit $S_x[n, k]$ le spectrogramme du signal $x[m]$, définis respectivement en (1.11) et (1.1). En vertu du caractère centré du bruit $p[m]$ et de son indépendance vis à vis du signal déterministe $d[m]$, nous avons

$$\mathbb{E}\{S_{dp}[n, k]\} = 0, \quad (3.1)$$

et l'espérance du spectrogramme s'écrit

$$\mathbb{E}\{S_x[n, k]\} = S_d[n, k] + \gamma_p[n, k], \quad (3.2)$$

dans lequel $S_d[n, k]$ désigne le spectrogramme du signal $d[m]$ et $\gamma_p[n, k]$ le spectre physique de la perturbation $p[m]$ introduit en (1.14).

À partir du spectrogramme $S_x[n, k]$ observé, l'objectif présent est l'estimation de la densité spectrale de puissance $\gamma_p[n, k]$ de la perturbation $p[m]$. Afin de rester dans le cadre le plus général possible, aucune hypothèse de blancheur ou de stationnarité n'est dans un premier temps formulée concernant cette densité spectrale de puissance (dsp). Nous verrons en fin de chapitre comment l'une et l'autre de ces hypothèses augmentent les performances des estimateurs. De cette formulation, il apparaît que l'estimation de la dsp du bruit se heurte à un double problème :

- d'une part nous accédons à γ_p via une espérance, or nous ne considérons disponible qu'une seule réalisation de $x[m]$ et donc de $S_x[n, k]$.
- d'autre part la contribution $S_d[n, k]$ est inconnue et son estimation nécessite la connaissance de $\gamma_p[n, k]$...

Nous présentons dans la suite trois types d'approches proposées dans la littérature pour résoudre ce problème d'estimation.

3.2.1 Les méthodes itératives

Pour répondre aux problèmes de l'espérance et de l'indétermination sur le signal, une première stratégie consiste à étendre explicitement l'investigation sur un voisinage du point $[n, k]$ considéré. Après une recherche de l'ensemble des points "bruit seul" contenus dans ce voisinage, une statistique peut être appliquée afin d'estimer la dsp du bruit. La forme et la taille de ce voisinage sont alors régies par deux contraintes :

- la densité spectrale de puissance $\gamma_p[n, k]$ doit être constante sur l'ensemble du voisinage,
- le voisinage doit contenir des coefficients "bruit seul" pour lesquels $S_d[n, k] = 0$.

Sous ces deux conditions et sous couvert d'indépendance des coefficients temps-fréquence, la dsp $\gamma_p[n, k]$ peut alors être estimée par la moyenne arithmétique des coefficients "bruit seul" observés dans le voisinage. Notons que la moyenne arithmétique d'un ensemble de réalisations d'une variable du chi² centrée, à 2 degrés de liberté et de coefficient de proportionnalité $\gamma_p/2$ est l'estimateur de γ_p optimal au sens du maximum de vraisemblance.

L'existence de ce chapitre est lié à la nécessité de connaître la dsp γ_p du bruit afin de pouvoir séparer, dans un plan temps-fréquence, les zones "signal" des zones "bruit seul". Or, dans l'approche décrite ici, l'estimation de cette dsp demande la reconnaissance des coefficients "bruit seul". Face à ce problème qui reboucle sur lui-même, des algorithmes qui rebouclent également sur eux-même, c'est à dire itératif, ont été proposés [HMC02][YR06][MHM06] :

Le principe est d'éliminer petit à petit les points temps-fréquence contenant du signal et d'estimer le bruit sur les points restant. Les zones "signal" sont identifiées par le biais d'un seuil, lui-même défini d'après la dsp du bruit. La dsp du bruit étant estimée comme la moyenne des points temps-fréquence non éliminés, un processus itératif d'estimation-élimination est mis en place. Un critère d'arrêt permet d'identifier si les points temps-fréquence non éliminés correspondent uniquement à du bruit.

L'avantage de cette approche est de s'attaquer simultanément au problème de l'estimation du bruit et de la localisation du signal dans le plan temps-fréquence. Elle permet donc de tout résoudre (pour ce qui nous concerne...) dans un unique élan. Un point délicat concerne le critère d'arrêt de la procédure : une vraisemblance avec une loi du chi² centrée est utilisée dans [HMC02], le moment d'ordre trois est utilisé dans [YR06] et le moment d'ordre quatre est utilisé dans [MHM06]. Le principal inconvénient de cette approche en est le temps de calcul et la variance des critères d'arrêt lorsque la taille du voisinage est restreinte.

3.2.2 Les méthodes de lissage

Dans le domaine du traitement de la parole et des signaux audio, la mise au point d'algorithmes robustes au bruit nécessite l'estimation de sa densité spectrale de puissance. La communauté du traitement des signaux audio a donc proposé un certain nombre d'estimateurs basés sur le modèle "signal+bruit" considéré ici. Afin de prendre en compte l'éventuelle non blancheur du bruit, l'analyse est classiquement réalisée par bande de fréquence. L'enjeu est alors de pouvoir suivre les non stationnarités du bruit. Une caractéristique originale qui se retrouve dans la plupart de ces algorithmes est l'utilisation d'un lissage récursif au premier ordre suivant l'axe des temps. Un avantage de ce lissage est qu'il réalise une opération proche du moyennage, résolvant ainsi le problème de l'espérance décrit précédemment. De plus, ce lissage est contrôlé par un paramètre qui est généralement adapté afin de discriminer les coefficients "signal" d'un côté et "bruit seul" de l'autre.

Dans cette catégorie d'algorithmes, la dsp du bruit est finalement recréée à partir du spectrogramme observé en adaptant le coefficient de lissage à la présence de signal. Déterminer la

présence du signal devient donc l'aspect central de ces algorithmes et plusieurs techniques ont été proposées [CB02][Coh03][RL06], généralement basées sur l'estimation d'un rapport signal sur bruit ou d'une probabilité de présence lorsqu'un modèle aléatoire de signal est adopté. Un avantage important de cette approche est la possibilité de traitement en temps réel : l'estimation du bruit est réalisée à chaque nouvelle trame de signal. Cependant, le nombre et la sensibilité des paramètres de lissage représentent un frein quant à son utilisation et le discernement nécessaire entre signal et bruit reste une question ouverte.

3.2.3 Les statistiques minimales

Une troisième approche, également issue des techniques de traitement de la parole, a été initiée dans les années 1990 par Rainer Martin [Mar01][Mar06][MM06]. Le principe de l'estimateur, appliqué parallèlement à toutes les fréquences, est de récupérer le plus petit coefficient temps-fréquence contenu dans une fenêtre temporelle de taille prédéfinie. Ayant de forte chance de correspondre à un coefficient "bruit seul", un gain multiplicatif lui est appliqué afin de retrouver la moyenne. Ce gain est défini par la statistique du minimum de la distribution supposée pour le bruit. Dans l'approche initialement proposée, un lissage récursif du spectrogramme est appliqué. Ceci permet de diminuer la variance des coefficients temps-fréquence, et notamment la variance du minimum. Cependant, la méthode comporte trois points sensibles :

- le paramètre de lissage doit tenir compte du rapport signal sur bruit,
- le minimum possède tout de même une grande variance,
- la statistique du minimum doit tenir compte du lissage effectué, ce qui rend très difficile l'obtention du gain multiplicatif.

Afin de diminuer la variance de l'estimateur, Stahl et al. [SFB00] proposent d'utiliser non plus le minimum mais le $q^{ième}$ quantile observé dans la même fenêtre temporelle. Dans le cadre d'une tâche de reconnaissance vocale, ils montrent que le choix de $q = 0.5$ (c'est à dire la médiane) donne les meilleurs résultats. Cependant, notons que le $q^{ième}$ quantile est directement considéré comme l'estimateur de la dsp du bruit, sans gain multiplicatif.

Là encore, la taille de la fenêtre temporelle d'observation (équivalente au voisinage des méthodes itératives) est le résultat d'un compromis entre la présence obligatoire de coefficients "bruit seul" et le suivi des non stationnarités du bruit. Afin de gagner sur le nombre de points disponibles sans limiter les possibilités de suivi des non stationnarités, l'extension de cette fenêtre suivant l'axe fréquentiel est proposé dans [EM02]. Notons que dans ces deux dernières méthodes, le lissage récursif du spectrogramme a été supprimé.

L'approche des statistiques minimales semble donner une alternative intéressante face au problème de l'indétermination du signal. Contrairement à la recherche difficile de l'ensemble des points "bruit seul" contenus dans une région temps-fréquence, elle se base uniquement sur quelques coefficients pour lesquelles l'hypothèse "bruit seul" est très probable. À partir de ces coefficients minimaux, une statistique est développée afin d'estimer la dsp du bruit. Les travaux initiaux [Mar01][Mar06] montrent que l'opération de lissage, utile pour diminuer la variance du spectrogramme, rend difficile l'expression de cette statistique. D'un autre côté, ne plus baser cette statistique uniquement sur le minimum, mais sur un quantile plus élevé permet de diminuer la variance de cette statistique.

La suite de ce chapitre est dédiée à l'évaluation d'un estimateur de la densité spectrale de puissance $\gamma_p[n, k]$ retenant l'idée des statistiques minimales. À partir d'un spectrogramme non lissé, il se base sur les Z plus petits coefficients observés dans un voisinage temps-fréquence du point $[n, k]$.

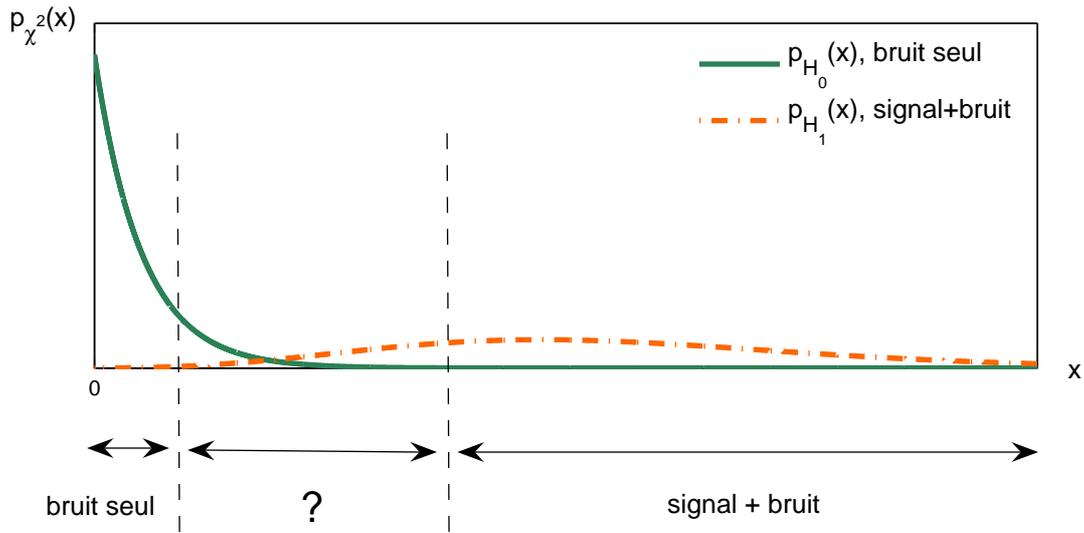


FIG. 3.1 – Densités de probabilité des coefficients du spectrogramme sous les deux hypothèses "bruit seul" et "signal+bruit". Identification de la zone dans laquelle l'hypothèse "bruit seul" est très fortement plus probable.

3.3 Contribution : estimation par les Z valeurs minimales

3.3.1 Formulation

Nous posons le problème de la manière suivante : soit l'ensemble $\mathcal{X} = \{X_1, \dots, X_T\}$ des T coefficients observés dans un voisinage du point temps-fréquence considéré $[n, k]$. Le choix du voisinage sera discuté par la suite. Parmi ces T coefficients, considérons que N d'entre eux correspondent à des réalisations "bruit seul", tandis que les $T - N$ restant sont "entachés" de signal. Si aux N coefficients "bruit seul" sont attribués les N premiers indices, l'ensemble \mathcal{X} se divise schématiquement selon

$$\mathcal{X} = \underbrace{\{X_1, \dots, X_N\}}_{\text{bruit seul}}, \underbrace{\{X_{N+1}, \dots, X_T\}}_{\text{signal+bruit}}.$$

Au regard des densités de probabilités des coefficients du spectrogramme sous les deux hypothèses "bruit seul" et "signal+bruit", nous pouvons supposer, comme illustré sur la figure 3.1, que les plus petits coefficients observés sont très probablement liés à la présence unique de bruit, tandis que les coefficients de très forte énergie indiquent très probablement la présence de signal. Entre ces deux extrêmes se trouve une zone de mélange dans laquelle des coefficients issus des deux hypothèses coexistent avec des probabilités comparables. En supposant que le plus petit coefficient contenant du signal est plus grand que le $Z^{\text{ième}}$ coefficient relatif au bruit seul, les trois zones de valeurs "bruit seul", "mixte" et "signal+bruit" se retrouvent, là encore schématiquement, dans l'ensemble \mathcal{X} selon

$$\mathcal{X} = \left\{ \begin{array}{c} X_1, \dots, X_Z, X_{Z+1}, \dots, X_N \\ X_{N+1}, \dots, X_S, X_{S+1}, \dots, X_T \\ \underbrace{\hspace{10em}}_{\text{bruit seul}} \quad \underbrace{\hspace{10em}}_{?} \quad \underbrace{\hspace{10em}}_{\text{signal+bruit}} \end{array} \right\}$$

Soit alors $\mathcal{X}_{ord} = \{X_{(1)}, \dots, X_{(T)}\}$ ce même ensemble \mathcal{X} dans lequel les valeurs sont entièrement ordonnées de façon croissante, *i.e* $X_{(1)} < \dots < X_{(T)}$. Selon le raisonnement précédent, cet ensemble se décompose selon

$$\mathcal{X}_{ord} = \underbrace{\{X_{(1)}, \dots, X_{(Z)}\}}_{\text{bruit seul}}, \underbrace{\{X_{(Z+1)}, \dots, X_{(S)}\}}_{?}, \underbrace{\{X_{(S+1)}, \dots, X_{(T)}\}}_{\text{signal+bruit}}.$$

Rappelons que la dsp recherchée du bruit correspond à la moyenne des N coefficients "bruit seul" contenus dans cette ensemble. L'opérateur de moyenne étant cependant une statistique globale, il peut être utilisé si les observations décrivent complètement et uniquement la loi sous-jacente qui les caractérise. Dans le cas présent, l'utilisation de cette statistique nécessite alors l'extraction des tous les coefficients "bruit seul" de la zone centrale de mélange. Sans hypothèses supplémentaires, cette tâche est impossible à réaliser de façon directe et fonde la motivation des algorithmes récursifs présentés auparavant. Maintenant, ne retenir que les Z plus petit coefficients permet d'isoler un échantillon de coefficients "bruit seul". L'estimation de la densité spectrale du bruit par les valeurs minimales se ramène alors au problème suivant : comment estimer la moyenne de N coefficients que nous supposons indépendants et identiquement distribués (iid) sachant que l'on observe uniquement les Z plus petits ?

3.3.2 Estimateurs proposés

La formulation mathématique du problème est la suivante : soient N réalisations ordonnées $\{x_{(1)}, \dots, x_{(N)}\}$ d'une variable du χ^2 centrée à deux degrés de liberté dont la densité de probabilité s'écrit (équation (2.47)) :

$$p(x) = \frac{1}{\gamma_p} \exp\left(-\frac{x}{\gamma_p}\right). \quad (3.3)$$

Comment estimer le paramètre γ_p à partir des Z plus petites réalisations $\{x_{(1)}, \dots, x_{(Z)}\}$ de cette variable X ? Dans la suite, deux estimateurs sont proposés. Le premier utilise la fonction de répartition d'une variable du χ^2 , le second est basé sur la densité de probabilité des Z plus petites réalisations.

3.3.2.1 Utilisation de la fonction de répartition

L'utilisation de la fonction de répartition $F_X(x) = P[X < x]$ permet de résoudre simplement ce problème. D'une part, son expression analytique est connue et il est facile d'en extraire le paramètre γ_p recherché : pour tout x positif, nous pouvons écrire

$$F_X(x) = 1 - \exp\left(-\frac{x}{\gamma_p}\right). \quad (3.4)$$

et donc

$$\gamma_p = \frac{x}{-\ln(1 - F_X(x))}. \quad (3.5)$$

D'autre part, une fonction de répartition possède l'estimateur empirique suivant : pour tout indice (i) compris entre 1 et N ,

$$\hat{F}_X(x_{(i)}) = \frac{i}{N}. \quad (3.6)$$

À partir de ces observations, il s'en suit qu'un estimateur de γ_p est disponible pour chacune des Z réalisations observées $\{x_{(1)}, \dots, x_{(Z)}\}$ de la variable du χ^2 . Il s'écrit, pour tout indice (i) compris entre 1 et Z ,

$$\hat{\gamma}_{p,i} = \frac{x_{(i)}}{-\ln\left(1 - \frac{i}{N}\right)}. \quad (3.7)$$

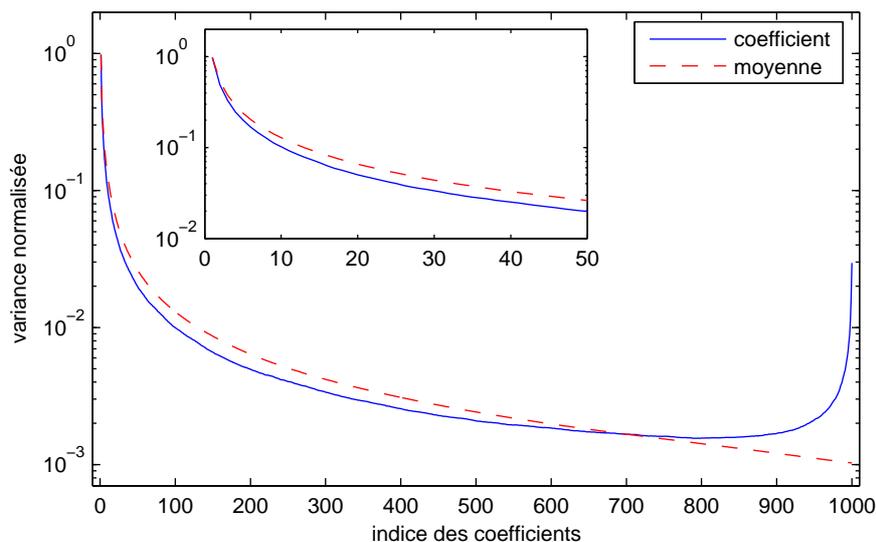


FIG. 3.2 – Variance des quantiles et de leur moyenne pour 10000 réalisations d'une variable du χ^2 centrée. Les variances sont normalisées par le carré de l'espérance.

Avec cette formulation, un estimateur de γ_p est disponible pour chacune des Z observations $x_{(i)}$. Il s'en suit qu'un grand nombre d'estimateurs peut être développé à partir d'une fonction de tout ou partie de ces Z observations $x_{(i)}$. Il est par exemple possible de formuler un estimateur basé sur la moyenne des $x_{(i)}$, ou encore sur l'intégrale de la fonction de répartition limitée à la donnée $x_{(Z)}$. La variance d'un tel estimateur est néanmoins directement reliée à la variance des quantiles qu'il utilise. Sur ce point, nous avons observé que l'estimateur n'utilisant que le $Z^{\text{ième}}$ quantile $x_{(Z)}$ admet la plus faible variance : la figure 3.2 présente la variance (sur 10000 évaluations) de chacun des 1000 quantiles d'une loi du χ^2 centrée (trait plein), ainsi que la variance de la moyenne des premiers quantiles (trait pointillé). Les variances sont normalisées par le carré de l'espérance de chaque quantile ou moyenne correspondant. Notons que la forme de cette courbe est dépendante de la densité de probabilité sous-jacente, ici la loi du χ^2 centrée à 2 degrés de liberté. Dans la zone des petits quantiles qui nous intéresse, nous observons une forte diminution de la variance lorsque l'ordre du quantile augmente. Par exemple, la variance du quantile $q = 0.1$ (soit d'indice $1/10^{\text{ième}}$ du nombre total de réalisations) est 100 fois plus petite que la variance du minimum. Aussi, la variance de la moyenne des i premiers quantiles est supérieure à la variance du $i^{\text{ième}}$ quantile (la moyenne utilise tous les termes précédents, chacun de variance plus élevée). L'estimateur de plus faible variance est ainsi celui qui utilise uniquement le plus grand des Z coefficients retenus pour l'estimation.

Finalement, le premier estimateur $\hat{\gamma}_p$ de la dsp du bruit γ_p s'écrit

$$\hat{\gamma}_p = \frac{x_{(Z)}}{-\ln\left(1 - \frac{Z}{N}\right)}, \quad (3.8)$$

dans lequel $x_{(Z)}$ est le $Z^{\text{ième}}$ plus petit coefficient parmi les N coefficients "bruit seul" contenus dans le voisinage.

3.3.2.2 Utilisation de la loi des minima

Nous proposons un second estimateur, cette fois basé sur le calcul de la densité de probabilité des Z plus petits coefficients. Son principe est de définir le comportement global de ces Z coefficients afin de pouvoir estimer le paramètre γ_p à l'aide d'une statistique d'ensemble comme une moyenne.

Nous cherchons donc la densité de probabilité d'une variable X du χ^2 centrée, à 2 degrés de liberté et de coefficient de proportionnalité $\gamma_p/2$, sous la condition que la variable X est inférieure à un seuil u qui est ici le $Z^{\text{ième}}$ coefficient $x_{(Z)}$. La notation u est cependant gardée dans la suite afin d'alléger les écritures. La fonction de répartition $F_Z(x)$ de cette variable aléatoire s'écrit, pour tout $x < u$:

$$\begin{aligned} F_Z(x) &= P\{X < x/X < u\} \\ &= \frac{P\{(X < x) \cap (X < u)\}}{P\{X < u\}} \\ &= \frac{P\{X < x\}}{P\{X < u\}}, \text{ car } x < u \\ &= \frac{F_X(x)}{F_X(u)}, \end{aligned} \quad (3.9)$$

dans lequel $F_X(x)$ est la fonction de répartition d'une variable du χ^2 donnée en (3.4). Par dérivation, la densité de probabilité s'écrit

$$\begin{aligned} p_Z(x) &= \frac{1}{F_X(u)} p_X(x) \\ &= \frac{1}{1 - \exp\{-\frac{u}{\gamma_p}\}} \frac{1}{\gamma_p} \exp\{-\frac{x}{\gamma_p}\}. \end{aligned} \quad (3.10)$$

En notant m_Z la moyenne empirique des Z coefficients $x_{(1)}, \dots, x_{(Z)}$, soit

$$m_Z = \frac{1}{Z} \sum_{i=1}^Z x_{(i)}, \quad (3.11)$$

l'estimateur $\hat{\gamma}_{p,mv}$ de γ_p au sens du maximum de vraisemblance est solution de l'équation :

$$m_Z - \hat{\gamma}_{p,mv} + \frac{u}{\exp\{\frac{u}{\hat{\gamma}_{p,mv}}\} - 1} = 0. \quad (3.12)$$

La solution $\hat{\gamma}_{p,mv}$ de cette équation ne possède pas d'expression analytique finie. Une solution approchée peut néanmoins être obtenue en récrivant l'équation (3.12) sous la forme

$$m_Z - \gamma_p + u \frac{1 - F_X(u)}{F_X(u)} = 0, \quad (3.13)$$

et en utilisant l'estimateur empirique donné en (3.6) pour la fonction de répartition $F_X(u)$. L'estimateur se formule finalement selon

$$\hat{\gamma}_p = m_Z + u \left(\frac{N}{Z} - 1 \right), \quad (3.14)$$

en rappelant que $u = x_{(Z)}$ est le $Z^{\text{ième}}$ plus petit coefficient observé, et N est le nombre de points "bruit seul" contenus dans le voisinage.

3.3.3 Évaluation des estimateurs

Dans ce paragraphe, les deux estimateurs (3.8) et (3.14) sont évalués en terme de biais et de variance et comparés à l'estimateur optimal qui consiste à réaliser la moyenne empirique des N coefficients "bruit seul". Dans un second temps, la sensibilité de l'estimation à une erreur sur le paramètre N et à la présence de signal parmi les Z plus petits coefficients est évaluée.

3.3.3.1 Biais et variance

Le biais et la variance normalisés des estimateurs, présentés sur la figure 3.3, sont évalués dans trois configurations :

- en fonction du nombre Z de coefficients retenus, pour un nombre constant $N = 100$ de coefficients "bruit seul" et pour $\gamma_p = 1$. Figure 3.3-a).
- en fonction du nombre N de coefficients "bruit seul", pour un nombre constant $Z = N/10$ de coefficients retenus et pour $\gamma_p = 1$. Figure 3.3-b).
- en fonction de la puissance spectrale γ_p à estimer, pour un nombre constant de coefficients retenus $Z = 10$ parmi $N = 100$ coefficients "bruit seul". Figure 3.3-c).

Le biais et la variance normalisés des estimateurs $\hat{\gamma}_p$ sont calculés de la manière suivante :

$$\text{biais normalisé } (\hat{\gamma}_p) = \frac{\mathbb{E}\{\hat{\gamma}_p\} - \gamma_p}{\gamma_p}, \quad (3.15)$$

$$\text{variance normalisée } (\hat{\gamma}_p) = \frac{\text{Var}\{\hat{\gamma}_p\}}{\gamma_p^2}, \quad (3.16)$$

dans lequel l'espérance et la variance sont évalués de façon empirique sur un ensemble de 5000 réalisations.

Les simulations présentées sur la figure 3.3 montrent que l'estimateur (3.8) basé sur la fonction de répartition (FdR) admet un léger biais et tend à sous-estimer la densité spectrale du bruit γ_p . Le second estimateur donné par (3.14), en revanche, ne présente pas cet inconvénient. Concernant la variance, les deux estimateurs montrent un comportement tout à fait similaire. La variance des estimateurs décroît logiquement lorsque le nombre Z de coefficients disponibles pour l'estimation augmente (figure 3.3-a) et -b)). Aussi, les performances des estimateurs sont invariants face à la quantité γ_p à estimer (figure 3.3-c). Logiquement, la variance des estimateurs proposés n'utilisant que Z coefficients est supérieure à la variance de l'estimateur idéal utilisant l'ensemble des N coefficients (figures 3.3-a), -b) et -c), courbes de droite). Pour tous les cas, nous observons que le rapport entre ces deux variances est égal au rapport entre N et Z .

À partir de cette petite analyse, nous choisissons d'utiliser l'estimateur (3.14) basé sur la loi des Z plus petits coefficients du spectrogramme afin d'estimer la densité spectrale de puissance $\gamma_p[n, k]$ de la perturbation.

3.3.3.2 Sensibilité aux paramètres

Le nombre N de coefficients "bruit seul" présents dans la région d'analyse est un paramètre inconnu. Afin de spécifier l'estimateur (3.14), ce paramètre doit être soit déterminé *a priori*, soit estimé. Avant d'envisager une manière de déterminer ce paramètre, nous évaluons ici la sensibilité de l'estimateur à une erreur sur ce paramètre. La figure 3.4 présente la moyenne sur 1000 réalisations du rapport $\hat{\gamma}_p/\gamma_p$ en fonction de l'erreur commise sur le paramètre N . Nous observons que l'erreur commise sur l'estimation est égale à l'erreur commise sur le paramètre N . La détermination de ce paramètre apparaît donc comme un élément important du processus

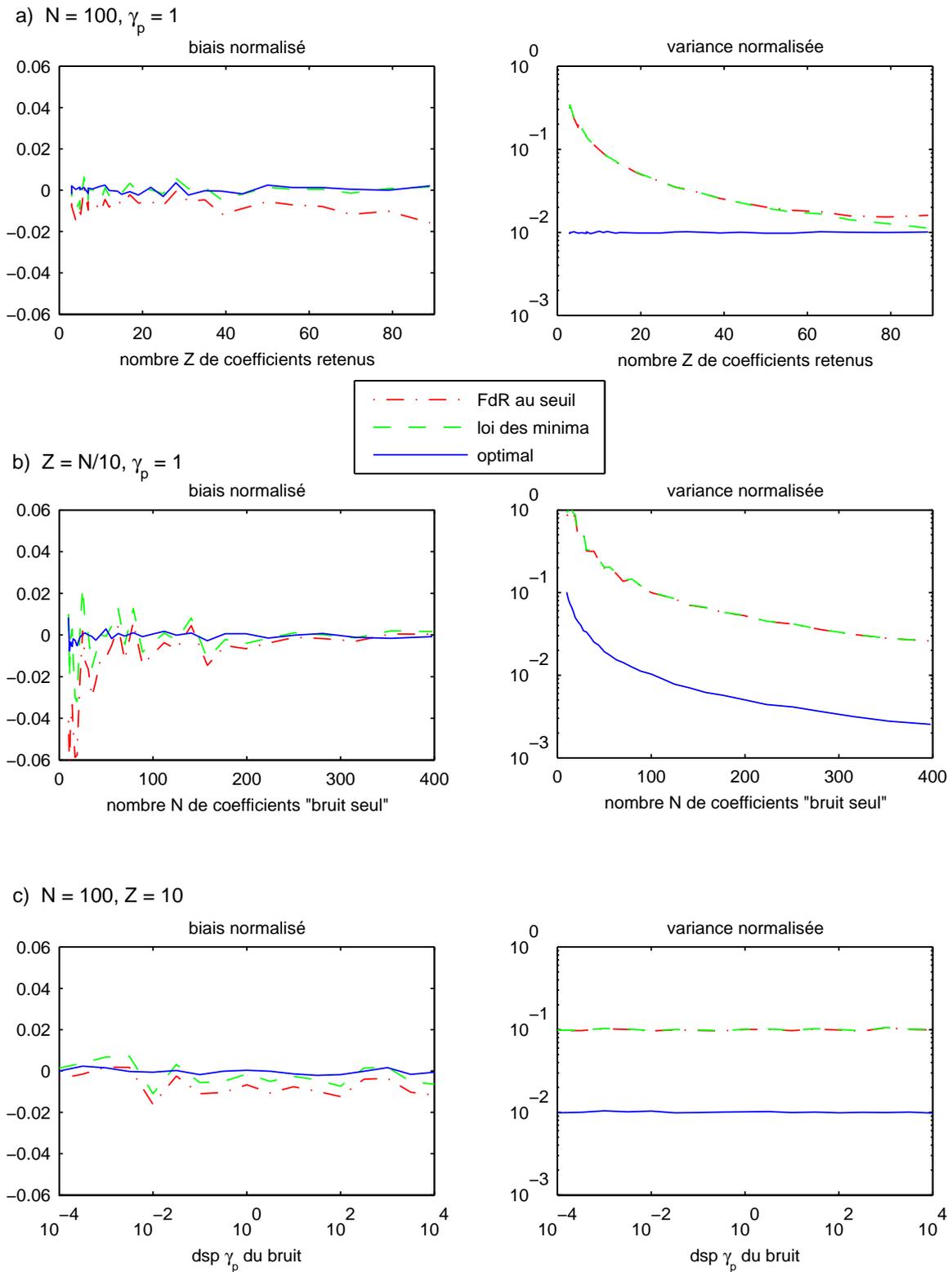


FIG. 3.3 – Biais et variance normalisés des estimateurs 'FdR au seuil' (3.8) et 'loi des minima' (3.14) (traits discontinus) et comparaison avec l'estimateur idéal utilisant l'ensemble des N coefficients (trait continu).

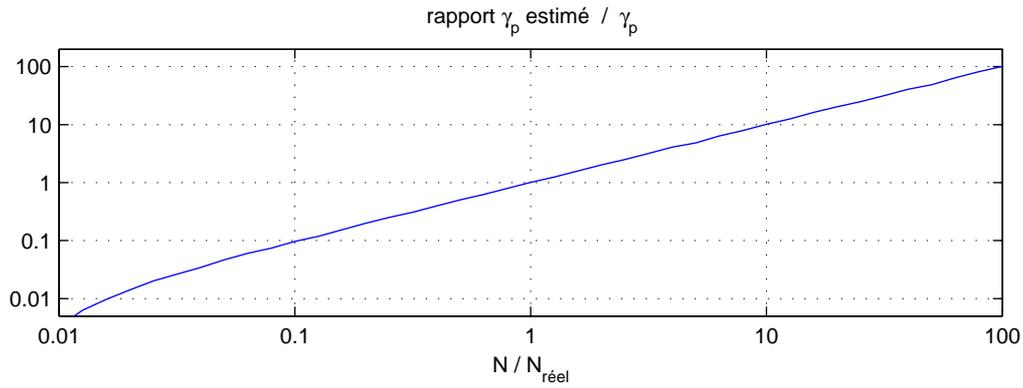


FIG. 3.4 – Sensibilité de l'estimateur (3.14) à une erreur sur le paramètre N . nombre de coefficients "bruit seul" $N_{\text{réel}} = 1000$, nombre de coefficients retenus $Z = 10$.

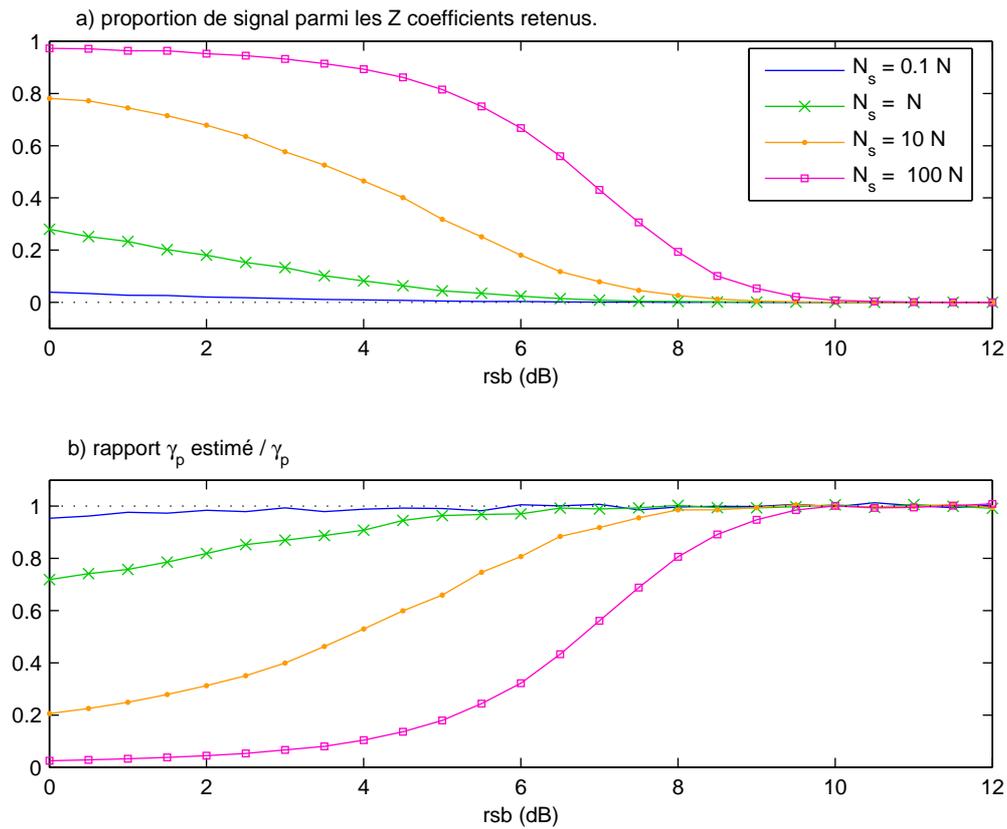


FIG. 3.5 – Sensibilité de l'estimateur (3.14) à la présence de signal dans les Z coefficients retenus pour l'estimation, en fonction du rapport signal sur bruit $\text{rsb}_{dB} = 10\log_{10}(S_d/\gamma_p)$. Nombre de coefficients "bruit seul" $N = 100$, nombre de coefficients retenus $Z = 10$.

d'estimation.

Nous évaluons maintenant la sensibilité de l'estimateur face à un second élément perturbateur : l'éventuelle **présence de coefficients "signal" parmi les Z plus petits coefficients** retenus pour l'estimation. Nous prenons l'exemple d'une région temps-fréquence au sein de laquelle $N = 100$ coefficients "bruit seul" sont présents. En fonction du nombre N_s de points signal également présents dans cette région et du rapport signal à bruit, les Z plus petits coefficients retenus pour l'estimation peuvent contenir certains de ces coefficients "signal", biaisant ainsi l'estimateur. Nous représentons sur la figure 3.5-a) la proportion de coefficients "signal" présents parmi les $Z = 10$ plus petits coefficients retenus. Quatre courbes sont données pour quatre proportions de points "signal" : $N_s = 0,1N$, $N_s = N$, $N_s = 10N$ et $N_s = 100N$. Rappelons que l'ensemble du voisinage est composé de $T = N + N_s$ coefficients. Le rapport signal sur bruit exprimé en dB est donné en abscisse. Logiquement, plus la proportion de points signal augmente dans la région temps-fréquence analysée, plus le risque d'en retrouver parmi les Z plus petits coefficients augmente. Ce risque décroît naturellement avec le rapport signal à bruit. Ainsi, lorsque l'énergie S_d du signal est environ 10 fois plus grande que l'énergie du bruit γ_p (rsb = 10 dB), le risque de prendre en compte des coefficients 'entachés' de signal est négligeable.

L'impact de ces éléments "signal" sur l'estimateur (3.14) est présenté sur la figure 3.5-b) en terme du rapport $\mathbb{E}\{\hat{\gamma}_p\}/\gamma_p$ estimé sur 2000 réalisations. La présence de signal parmi les Z coefficients retenus tend à sur-estimer la dsp γ_p du bruit. Aussi, nous observons que le biais engendré est exactement égal à la proportion de coefficients "signal" présents parmi les Z plus petits coefficients (cf figure 3.5-a)). Ainsi, si la moitié des Z coefficients retenus pour l'estimation contient du signal, le résultat de l'estimation sera $\hat{\gamma}_p = 2\gamma_p$.

3.3.3.3 Récapitulatif

L'évaluation des deux estimateurs proposés a mis en évidence les points suivants :

- Les deux estimateurs admettent une variance et une sensibilité aux paramètres similaires. Toutefois, un léger biais entache l'estimateur (3.8) et à ce titre le second estimateur (3.14) lui sera préféré.
- En comparaison avec la variance de l'estimateur idéal qui utiliserait la totalité des N coefficients "bruit seul", la variance des estimateurs basés sur les Z plus petits coefficients augmente d'un facteur égal à N/Z (figure 3.3). Soit

$$\text{Var}\{\hat{\gamma}_p\} = \frac{N}{Z} \cdot \text{Var}\{\hat{\gamma}_{p,\text{idéal}}\}. \quad (3.17)$$

L'estimateur idéal $\hat{\gamma}_{p,\text{idéal}}$ étant ici une moyenne empirique sur N termes, sa variance est égal à γ_p/N . Il en découle, dans le cas où N est déterministe (connu), que la variance de l'estimateur (3.14) est égale à

$$\text{Var}\{\hat{\gamma}_p\} = \frac{\gamma_p}{Z}. \quad (3.18)$$

- Le nombre N de coefficients "bruit seul" étant inconnu, il devra être déterminé ou estimé. En cas d'erreur sur ce paramètre, le biais introduit est égal au rapport entre le nombre N utilisé dans l'estimation et le nombre $N_{\text{réel}}$ de coefficients "bruit seul" réellement présents dans la région temps-fréquence (figure 3.4). Soit

$$\mathbb{E}\{\hat{\gamma}_p\} = \frac{N}{N_{\text{réel}}} \cdot \gamma_p. \quad (3.19)$$

- En cas de faible rapport signal à bruit, les Z coefficients retenus pour l'estimation peuvent contenir des points "signal". Dans ce cas, le biais introduit est égal à la proportion p de coefficients "signal" contenus dans les Z plus petits coefficients (figure 3.5) :

$$\mathbb{E}\{\hat{\gamma}_p\} = (1 - p) \cdot \gamma_p. \quad (3.20)$$

3.3.4 Détermination du voisinage et des paramètres Z et N

Dans ce paragraphe, nous discutons les trois aspects critiques relatifs à l'utilisation de l'estimateur (3.14), à savoir le choix du voisinage sur lequel l'estimation est réalisée ainsi que la détermination des paramètres Z et N .

3.3.4.1 Détermination du voisinage

Pour un point temps-fréquence $[n, k]$ donné, la taille et la forme du voisinage à partir duquel la densité spectrale de puissance du bruit $\gamma_p[n, k]$ est estimée est soumise à plusieurs contraintes.

Premièrement, la dsp $\gamma_p[n, k]$ de la perturbation doit y être la plus constante possible, permettant ainsi de considérer que les échantillons "bruit seul" sont identiquement distribués selon une loi du χ^2 centré. Globalement, cette contrainte tend à limiter la taille du voisinage. Néanmoins, un peu de connaissance *a priori* sur la perturbation permet d'adapter la forme du voisinage : pour une perturbation blanche, la région d'analyse peut s'étaler sur l'ensemble de l'axe fréquentiel. Pour une perturbation stationnaire, le bruit est identiquement distribué sur tout l'axe des temps sur lequel le voisinage peut donc s'étendre. En revanche, pour une perturbation colorée et non stationnaire, le recours à un voisinage centré en temps et en fréquence autour du point temps-fréquence considéré semble nécessaire. Dans ce cas, la dimension temporelle du voisinage représente une limite pour l'estimateur dans le suivi des non stationnarités. De façon équivalente, la dimension fréquentielle du voisinage se traduit par une limitation de l'estimateur à décrire fidèlement des variations fréquentielles importantes du spectre de la perturbation. Notons aussi que l'estimateur étant basé sur les plus petits coefficients temps-fréquence, l'estimation sera toujours biaisée vers le bas dans le cas où la dsp du bruit n'est pas constante dans le voisinage considéré.

Une seconde contrainte est liée à la présence obligatoire de coefficients "bruit seul" dans le voisinage. Sans hypothèse *a priori* sur le signal déterministe $d[m]$, l'existence de grandes régions temps-fréquence sur lesquelles ce signal est présent partout ne peut être exclue. Dans le cas idéal, le voisinage doit être plus grand que la plus grande de ces régions signal.

Une troisième contrainte est liée à la variance de l'estimation qui décroît naturellement avec le nombre de réalisations disponibles. À titre d'exemple, la décroissance de la variance de l'estimateur en fonction du nombre N de points "bruit seul" est observable sur la figure 3.3-b). Pour une dsp du bruit constante, cette contrainte tend à augmenter la taille du voisinage.

Pour finir, une quatrième contrainte associée aux "effets de bords" temporels peut être mentionnée. La représentation temps-fréquence d'un signal réel de durée temporelle finie étant également limitée dans le temps, la dimension temporelle du voisinage ne permet pas d'accéder aux bordures temporelles de l'image temps-fréquence. L'information contenue au tout début et à la toute fin du signal sera donc perdue. Par contre, concernant les "effets de bords" fréquentiels, l'utilisation de la périodicité en fréquence du spectre d'un signal discret permet d'atténuer cet aléa.

Il ressort de cette discussion que la détermination du voisinage est le résultat d'un compromis. Aussi, des connaissances *a priori* sur le signal et/ou sur les variations de la dsp du bruit sont

très avantageuses pour ce choix. En l'absence de toute information, la seule contrainte qui puisse être prise en compte concerne la variance de l'estimateur. Afin de ne pas voir celle-ci prendre des valeurs trop élevées, nous suggérons une taille minimale de 100 points pour le voisinage. Dans la suite, une cellule temps-fréquence de 21 points fréquentiels sur 21 points temporels (soit $T = 441$) et centrée sur le point $[n, k]$ considéré est utilisée. Notons que ce choix est purement arbitraire et doit dépendre en pratique des paramètres de construction du spectrogramme (notamment le nombre K de fréquences calculées et le décalage D entre deux analyses successives). Une formulation de la taille du voisinage en terme d'étendue temporelle (en seconde) et fréquentielle (en hertz) est alors mieux adaptée.

3.3.4.2 Détermination du paramètre Z

Le paramètre Z représente le nombre de coefficients minimaux retenus pour l'estimation de γ_p . Dans le meilleur des cas, il doit correspondre au plus grand coefficient "bruit seul" tel que tous les coefficients contenant du signal lui soient supérieurs. En pratique, il est bien entendu inconnu. Néanmoins, si le nombre de coefficients "bruit seul" observables ne peut être déterminé exactement, le nombre Z de coefficients retenus peut être fixé *a priori*. Le compromis repose entre une variance élevée de l'estimateur liée au choix d'un Z très petit et le risque de retenir des coefficients "signal" en cas de Z trop grand, ce qui induit un biais dans l'estimateur (figure 3.5). Dans la suite, le nombre Z est arbitrairement fixé à $1/10^{\text{ième}}$ du nombre total T de points contenus dans le voisinage.

3.3.4.3 Estimation du paramètre N

Le paramètre N correspond au nombre total de coefficients "bruit seul" contenus dans le voisinage du point $[n, k]$ considéré. Là encore, c'est un paramètre inconnu qu'il faut donc déterminer. L'évaluation de l'estimateur a montré qu'une erreur sur ce paramètre se reporte à l'identique dans le résultat de l'estimation (figure 3.4). La détermination de ce paramètre apparaît alors comme un élément tout à fait critique pour le bon comportement de l'estimateur et son utilisabilité.

La motivation première à l'utilisation des statistiques minimales est de contourner la détermination de l'ensemble des coefficients "bruit seul" contenus dans un voisinage. Force est de constater ici que la méthode proposée ne permet pas de s'affranchir complètement de cette difficulté. Toutefois, au lieu de rechercher quels sont les coefficients "bruit seul", nous devons trouver quel est leur nombre. Ce nouveau problème est un petit peu plus simple, mais comme nous allons le voir, ne peut être résolu sans la connaissance de la dsp γ_p du bruit. En conséquence, une méthode itérative permettant d'estimer conjointement γ_p et N est également nécessaire.

La détermination du paramètre N repose sur le raisonnement suivant : un voisinage donné contient T coefficients, parmi lesquels N sont du bruit et $T - N$ sont du signal. Si l'on peut ajuster un seuil correspondant à un rapport signal à bruit de 0dB¹, le nombre de coefficients N_{cand} supérieurs à ce seuil sera composé des $T - N$ coefficients "signal" auxquels s'ajoutent des coefficients "bruit seul" dont le nombre est donné par la probabilité de fausse alarme $P_{\text{fa},0dB}$ équivalente à ce seuil. Soit

$$N_{\text{cand}} = (T - N) + N.P_{\text{fa},0dB}. \quad (3.21)$$

¹Le rapport signal sur bruit considéré ici correspond au ratio $S_d[n, k]/\gamma_p[n, k]$ pour un unique point temps-fréquence. Un rapport signal sur bruit de 0dB semble alors intuitivement représenter la limite de détectabilité d'un signal. Les $T - N$ coefficients mentionnés ci-dessus correspondent alors aux coefficients "accessibles" du signal.

En manipulant cette expression, nous obtenons

$$N = \frac{T - N_{\text{cand}}}{1 - P_{\text{fa},0\text{dB}}} \quad (3.22)$$

Comme nous le verrons plus en détail dans le chapitre suivant, le seuil à 0dB est égal dans le cas d'une loi du chi² à $1,26\gamma_p$ et la PFA équivalente n'est autre que $P_{\text{fa},0\text{dB}} = \exp(-1,26)$. En utilisant l'estimateur (3.14) de γ_p , l'itération suivante est mise en place :

$$\left[\begin{array}{l} N_0 = T; \\ \text{for } i = 0 : 10 \\ \quad \hat{\gamma}_{p,i} = m_Z + u \left(\frac{N_i}{Z} - 1 \right); \\ \quad \text{seuil}_i = 1,26 \cdot \hat{\gamma}_{p,i}; \\ \quad N_{\text{cand},i} = \text{Card}\{X > \text{seuil}_i\}; \\ \quad \hat{N}_i = \frac{T - N_{\text{cand},i}}{1 - P_{\text{fa},0\text{dB}}}; \\ \text{end} \end{array} \right. \quad (3.23)$$

Le choix du nombre d'itérations (ici 10) est lié à l'observation que \hat{N} se stabilise généralement peu après 5 itérations. Notons aussi que le choix d'initialiser N à la valeur T impose une surestimation initiale de γ_p . Les itérations tendent alors à faire descendre conjointement $\hat{\gamma}_p$ et \hat{N} .

Afin d'évaluer la pertinence et les limites de cette méthode d'estimation, l'expérience suivante est réalisée : un signal sinusoïdal (fréquence pure) d'amplitude variable est ajouté à un bruit gaussien, blanc et stationnaire, de variance $\gamma_p = 1$. La sinusoïde est présente sur les 3/4 du signal temporel. À partir du spectrogramme du signal, la méthode d'estimation (3.23) est appliquée uniquement sur les T coefficients temps-fréquence observés à la fréquence de la sinusoïde. À cette fréquence, le nombre de coefficients "signal" est égal à $3T/4$ et le nombre N de coefficients "bruit seul" est égal à $T/4$. Aussi, les $Z = T/10$ plus petits coefficients sont retenus pour l'estimation. Les figures 3.6-e) et -f) donnent deux exemples de profils obtenus pour des rapports signal à bruit de 5 et 10 dB respectivement.

Les résultats obtenus pour 1000 estimations des paramètres N et γ_p , en fonction du rapport signal à bruit $\rho = 10\log_{10}(S_d/\gamma_p)$ sont présentés sur la figure 3.6. Les figures 3.6-a) et -b) donnent respectivement la moyenne et l'écart type de l'estimateur de N , tandis que les figures 3.6-c) et -d) concernent l'estimation de γ_p .

Ces figures indiquent clairement une limite de validité pour les estimateurs autour d'un rsb de 10 dB. Pour les rsb supérieurs à 10dB, les deux estimateurs \hat{N} et $\hat{\gamma}_p$ sont non biaisés. Dans cette zone de validité, l'estimateur \hat{N} admet un écart type constant égal à 15 (pour une espérance de 97). L'écart type de $\hat{\gamma}_p$ apparaît également constant et vaut 0.35, pour une espérance de 1 (figures 3.6-b) et -d)). Pour les rsb inférieurs à 10dB, les estimateurs sont fortement biaisés vers le haut (surestimation). La cause en est que la surestimation initiale de γ_p (imposée par l'initialisation $N_0 = T$) ne peut pas être rectifiée : le signal étant faible, le nombre de points N_{cand} supérieurs au seuil reste faible et \hat{N} reste bloqué aux alentours de T (figure 3.6-a)). La figure 3.6-c) indique une décroissance du biais de l'estimateur $\hat{\gamma}_p$ pour les très faibles rsb. Comme indiqué dans le paragraphe 3.3.3.2 sur la sensibilité aux paramètres et l'équation (3.20), ce comportement est lié à la présence de points "signal" parmi les Z coefficients retenus qui tend à diminuer la valeur estimée.

Il convient enfin de noter que ce rsb limite est fonction du rapport entre le nombre N de points "bruit seul" et la taille T du voisinage. Le rapport N/T fixe la surestimation initiale de la récursion (3.23). Plus la proportion de points "bruit seul" est importante et plus l'estimateur pourra prendre en compte des signaux de faible rsb. Ce comportement est illustré, pour la même

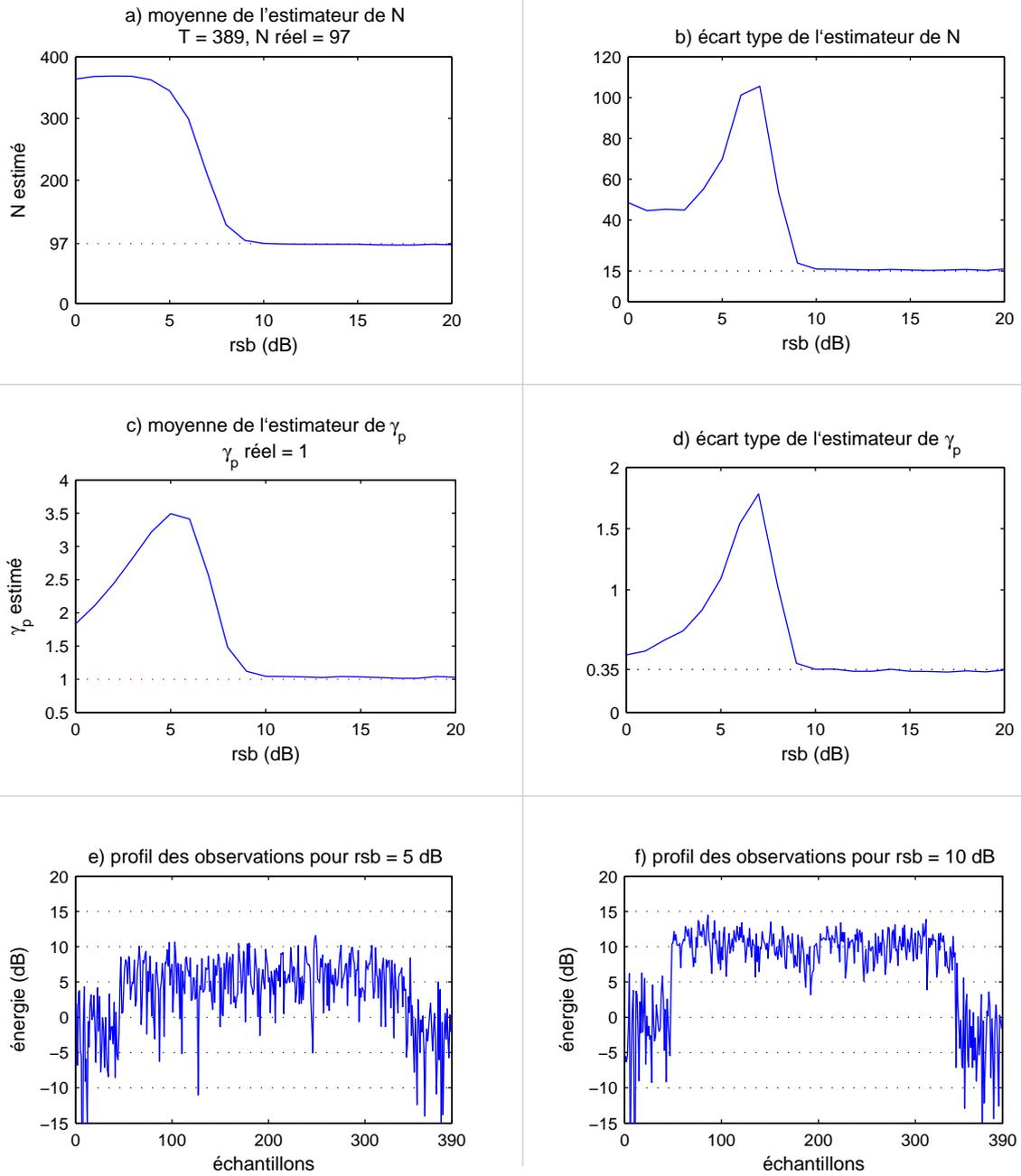


FIG. 3.6 – Évaluation de l'algorithme (3.23) proposé pour l'estimation du paramètre N et de γ_p en fonction du rapport signal sur bruit $10\log_{10}(S_d/\gamma_p)$. a) moyenne sur 1000 évaluations de l'estimateur \hat{N} , b) écart type de \hat{N} , c) moyenne de l'estimateur $\hat{\gamma}_p$, d) écart type de $\hat{\gamma}_p$, e) et f) coefficients temps-fréquence observés à la fréquence du signal pour un rsb de 5 et 10 dB respectivement.

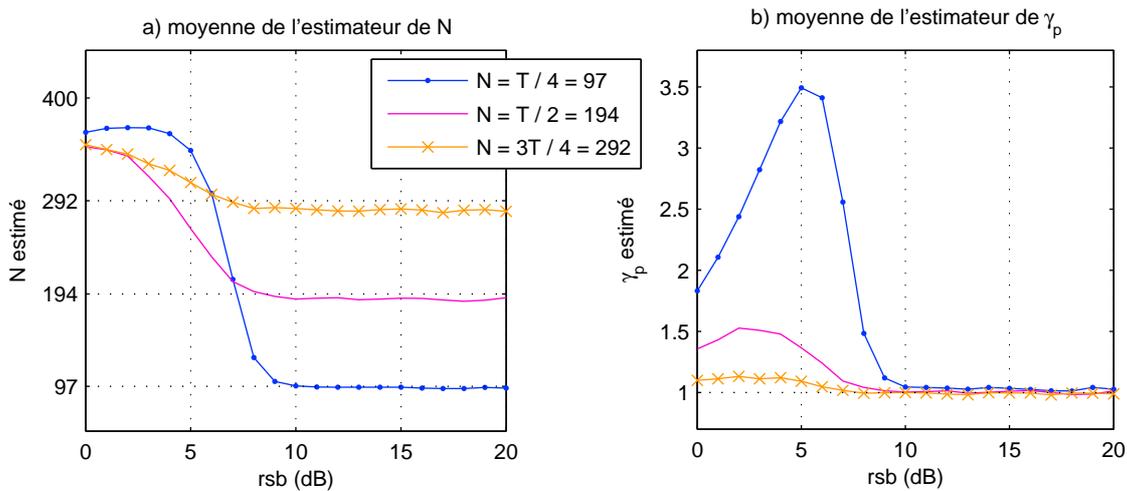


FIG. 3.7 – Évaluation de l'algorithme (3.23) proposé pour l'estimation du paramètre N et de γ_p en fonction du rapport signal sur bruit $10\log_{10}(S_d/\gamma_p)$, pour différents rapports N/T .

configuration que précédemment, sur la figure 3.7 pour les trois rapports $N/T = 0.25, 0.5$ et 0.75 . Dans le cas le plus favorable ($N/T = 3/4$), la limite de validité de l'estimateur correspond à un rsb de 7 dB.

La nécessité d'une estimation conjointe et itérative de N et γ_p remet bien sûr en balance l'avantage de la méthode basée sur les statistiques minimales par rapport à l'utilisation classique de la moyenne sur les zones "bruit seul". En reprenant la décomposition de l'ensemble ordonné \mathcal{X}_{ord} des coefficients observés dans un voisinage,

$$\mathcal{X}_{ord} = \underbrace{\{X_{(1)}, \dots, X_{(Z)}\}}_{\text{bruit seul}}, \underbrace{\{X_{(Z+1)}, \dots, X_{(S)}\}}_{?}, \underbrace{\{X_{(S+1)}, \dots, X_{(T)}\}}_{\text{signal+bruit}}$$

nous devons réaliser qu'une prise de décision sur les coefficients situés dans la zone centrale de mélange est inévitable. L'utilisation de la moyenne nécessite la détermination précise des coefficients "bruit seul" de cette zone. Dans les algorithmes itératifs dédiés à cette tâche, un critère de connexité temps-fréquence des coefficients "signal" est utilisé et chaque itération contient une recherche "spatiale" de ces coefficients dans le plan temps-fréquence. L'approche des statistiques minimales, quant à elle, ne nécessite plus cette recherche puisque seul le nombre de coefficients intervient. À ce titre, elle bénéficie d'un temps de calcul extrêmement avantageux.

3.4 Exemples d'estimation dans le plan temps-fréquence

Dans cette partie sont présentés des résultats d'estimation de la densité spectrale de puissance d'un bruit gaussien additif réalisée dans le plan temps-fréquence. Quatre signaux tests représentés par leur spectrogramme sur la figure 3.8 ont été choisis afin d'illustrer ces résultats : le premier est un signal synthétique composé de modulations de fréquence tandis que les trois autres correspondent respectivement à un signal acoustique de piano, un signal de parole "chantée" et un signal de parole "parlée". Le choix de ces signaux est principalement motivé par la

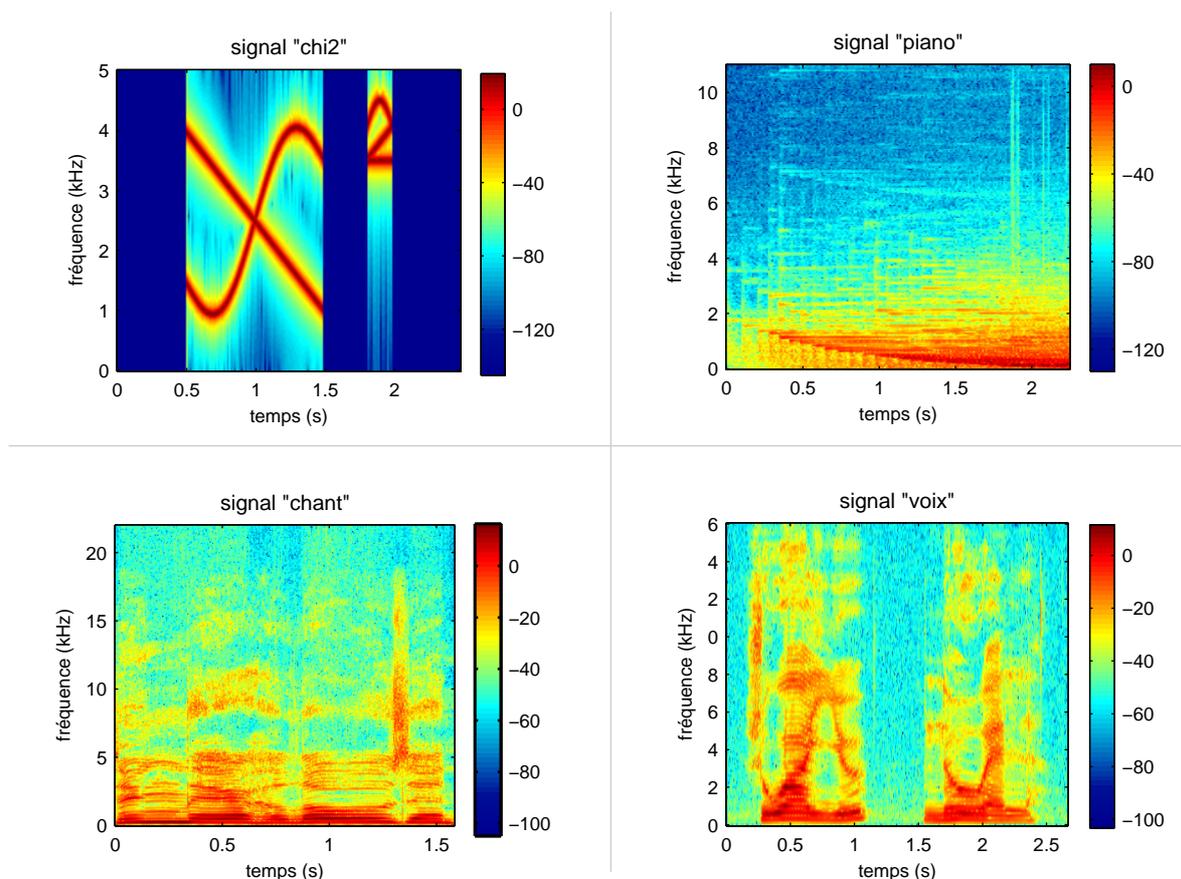


FIG. 3.8 – Spectrogrammes non bruités des quatre signaux tests utilisés pour l'évaluation de l'estimateur.

différence d'occupation du plan temps-fréquence et la diversité des motifs qu'ils présentent. Afin d'évaluer les performances, un bruit gaussien synthétique, dont les paramètres de blancheur et de stationnarité seront modifiés, est ajouté à ces signaux. Les cas d'une perturbation blanche et stationnaire, non blanche ou non stationnaire, puis non blanche et non stationnaire seront tour à tour présentés.

3.4.1 Perturbation blanche et stationnaire

Dans un premier temps, le cas d'une perturbation gaussienne blanche et stationnaire est étudié. C'est le cas le plus favorable puisque la perturbation étant constante sur l'ensemble du plan temps-fréquence, le voisinage d'estimation peut s'étendre sur la totalité du plan et contient fort probablement un nombre raisonnable de coefficients "bruit seul".

La figure 3.9 présente les résultats d'estimation obtenus sur chacun des quatre signaux tests. Pour chaque signal, l'histogramme de 10000 estimations est présenté. La valeur réelle de γ_p est indiquée par le trait pointillé. Dans chaque cas, le biais et la variance normalisés de l'estimateur sont indiqués.

Ce premier test met en évidence la dépendance de l'estimateur à la forme du signal et notamment à son degré de présence sur l'ensemble du spectrogramme (rapport N/T). Plus le signal est présent, plus l'estimateur est biaisé vers les grandes valeurs (voir par exemple les deux signaux de parole "chant" et "voix", figure 3.9-c) et -d)). À l'inverse, lorsque le spectrogramme contient

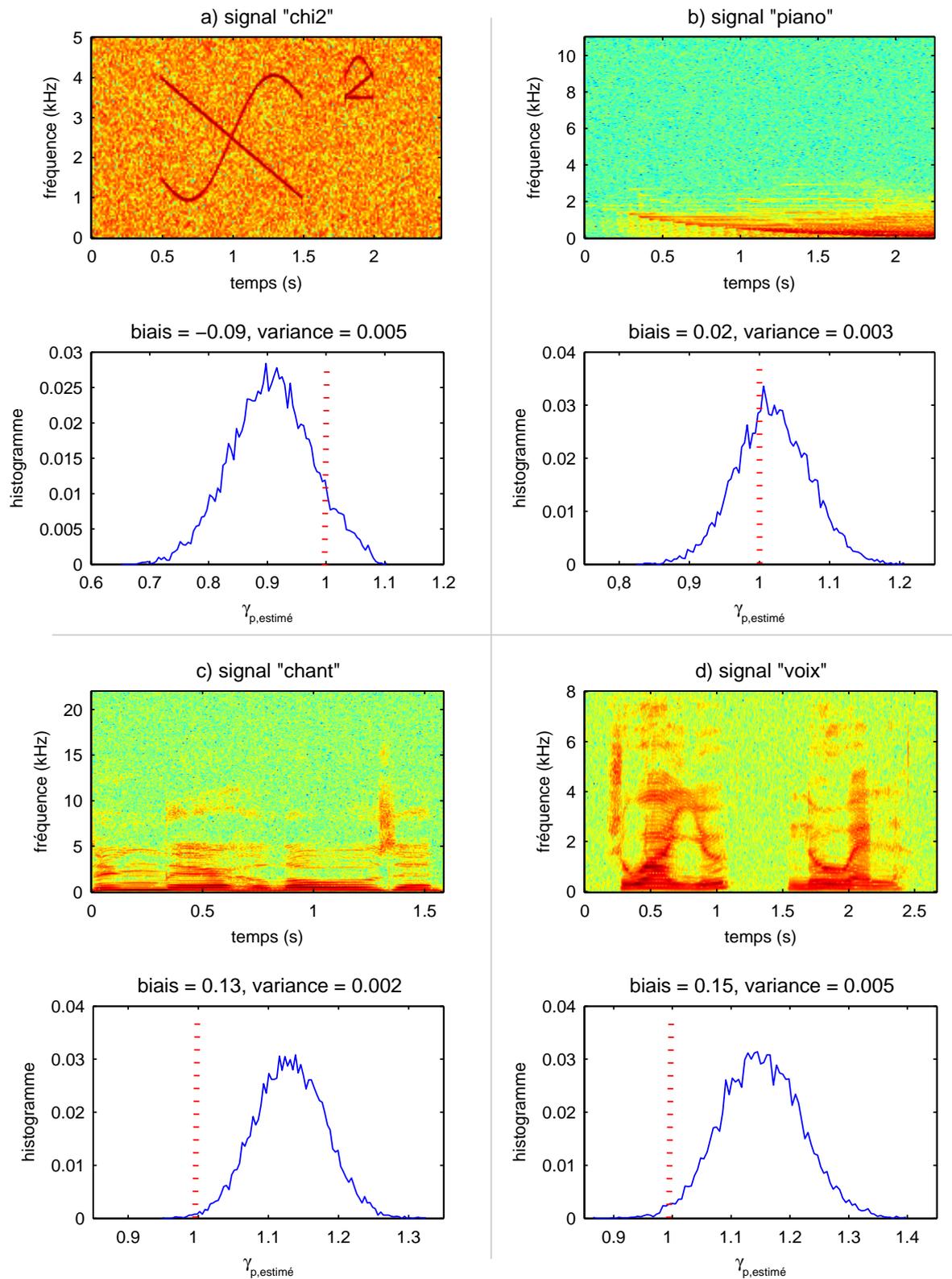


FIG. 3.9 – *Cas blanc et stationnaire.* Évaluation de l'estimateur dans le cas d'un bruit blanc et stationnaire. Pour chacun des quatre signaux test, le spectrogramme ainsi que l'histogramme issu de 10000 estimations sont représentés.

peu de coefficients "signal", l'estimateur tend à présenter un biais négatif (exemple du signal "chi2", figure 3.9-a)).

3.4.2 Perturbation blanche ou stationnaire

La connaissance *a priori* du caractère blanc ou stationnaire peut être prise en compte avec beaucoup d'avantage par l'ajustement de la forme du voisinage. Une perturbation blanche admettant une dsp constante sur tout l'axe fréquentiel, le voisinage retenu correspond à l'ensemble du spectre observé à la date n . De façon conjuguée, l'ensemble des temps à une fréquence donnée k sera retenu comme voisinage lorsque le caractère stationnaire de la perturbation est connu.

Les résultats d'estimation d'un profil fortement non stationnaire sont présentés sur les figures 3.10 et 3.11. Pour chaque signal, le résultat d'une estimation et de 100 estimations sont présentés afin d'évaluer respectivement la variance et le biais éventuel de l'estimateur. Pour les trois premiers signaux "chi²", "piano" et "chant", le suivi des non stationnarités est bien réalisé. Par contre, la forte variance de l'estimateur est nettement visible. Concernant le signal de parole "voix", trois cas limites mettant l'estimateur à défaut sont observés (figure 3.11-b)) :

- dans la zone autour de $n = 0.25$ sec, le rapport de présence entre signal et bruit est raisonnable ($N \sim T/4$). Par contre le rapport signal à bruit y est inférieur au rsb limite acceptable pour l'estimateur.
- pour la zone autour de $n = 0.6$ sec, le signal est très fortement présent sur tout le spectre. Le rapport T/N est trop important pour permettre une bonne estimation.
- pour la zone autour de $n = 1.8$ sec, même si le signal est moins énergétique que dans le cas précédent, le bruit y est tellement faible que le rapport N/T y est encore trop faible pour que le bruit soit estimé correctement.

La figure 3.12 présente les résultats obtenus pour des perturbations non blanches et stationnaires. Ici encore, si le biais de l'estimateur paraît acceptable, une grande variance est observée. La figure 3.13 présente le cas de trois signaux réels issus d'un environnement naturellement corrélié et stationnaire. Sur chaque signal, l'effet d'un filtrage anti-repliement accentue le caractère non blanc du bruit de fond. L'estimation du bruit et la prise en compte du filtrage anti-repliement semble ici acceptable.

3.4.3 Perturbation non blanche et non stationnaire

Le cas d'une perturbation non blanche et non stationnaire est bien entendu le cas le plus difficile. Néanmoins, il représente le critère final pour un estimateur n'utilisant aucune connaissance *a priori* sur le signal et sur le bruit.

La figure 3.14 présente le cas du signal synthétique "chi2" sur lequel un bruit violet (ascendant avec les fréquences) et non stationnaire a été ajouté. Pour chaque point temps-fréquence, la dsp du bruit $\gamma_p[n, k]$ est estimée à partir d'un voisinage de 21 points fréquentiels sur 21 points temporels et centré autour de $[n, k]$.

Les figures 3.14-a) et -b) présentent respectivement le spectrogramme du signal bruité ainsi que la densité spectrale de puissance de la perturbation. Le résultat global de l'estimation est donné sur la figure 3.14-d) où les effets de bord temporels sont laissés apparent. Pour ces trois figures, les échelles de couleur sont identiques. La figure 3.14-f) représente la carte des biais. Afin de discerner les biais négatifs et positifs tout en normalisant par la valeur réelle à estimer, le biais est calculé selon

$$\text{biais} = \log_{10} \left[\frac{\hat{\gamma}_p}{\gamma_p} \right].$$

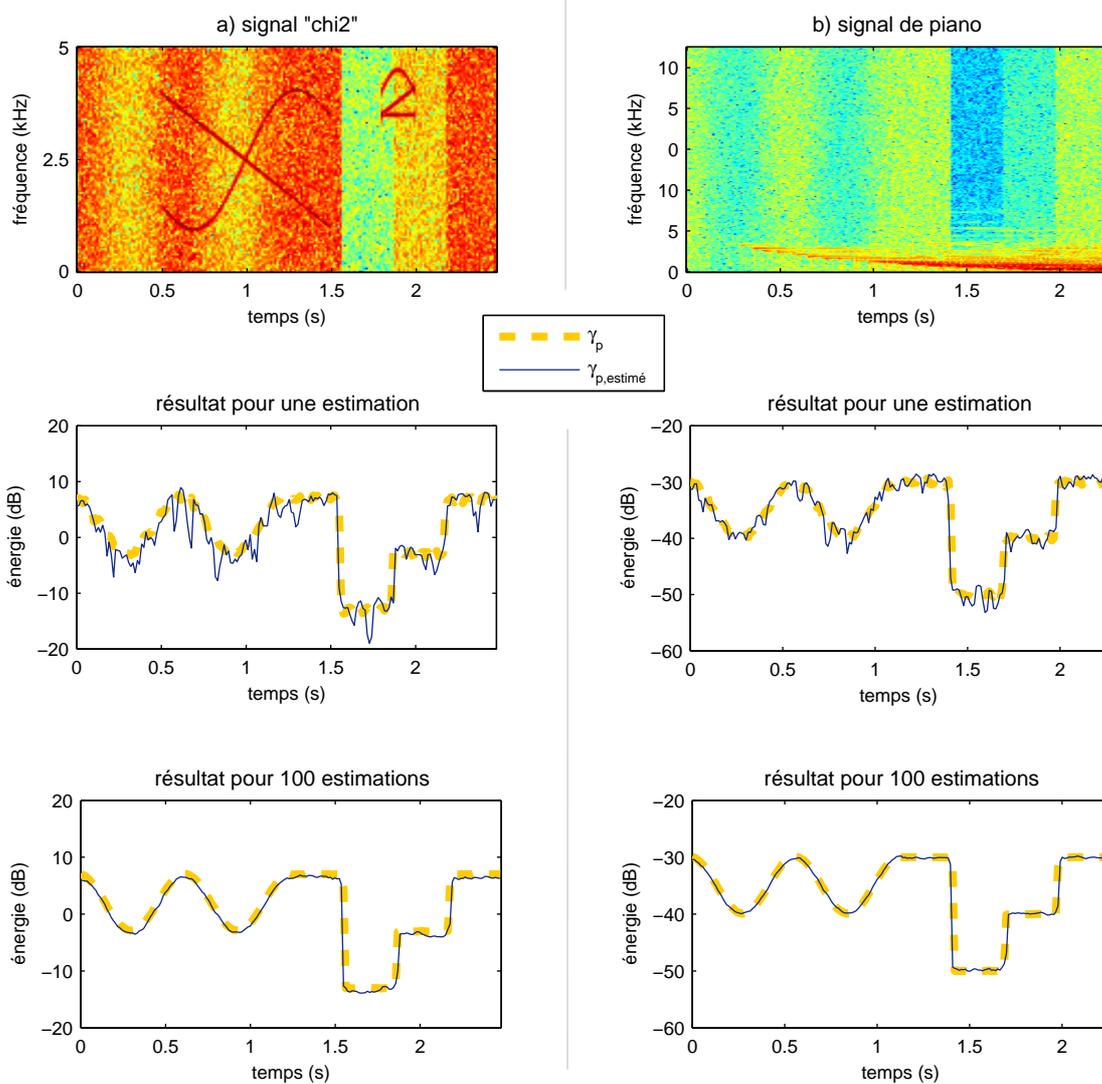


FIG. 3.10 – *Cas blanc et non stationnaire*. Évaluation de l'estimateur dans le cas d'un bruit blanc et non stationnaire. Pour chacun des deux signaux test, le résultat d'une estimation ainsi que la moyenne de 100 estimations sont données.

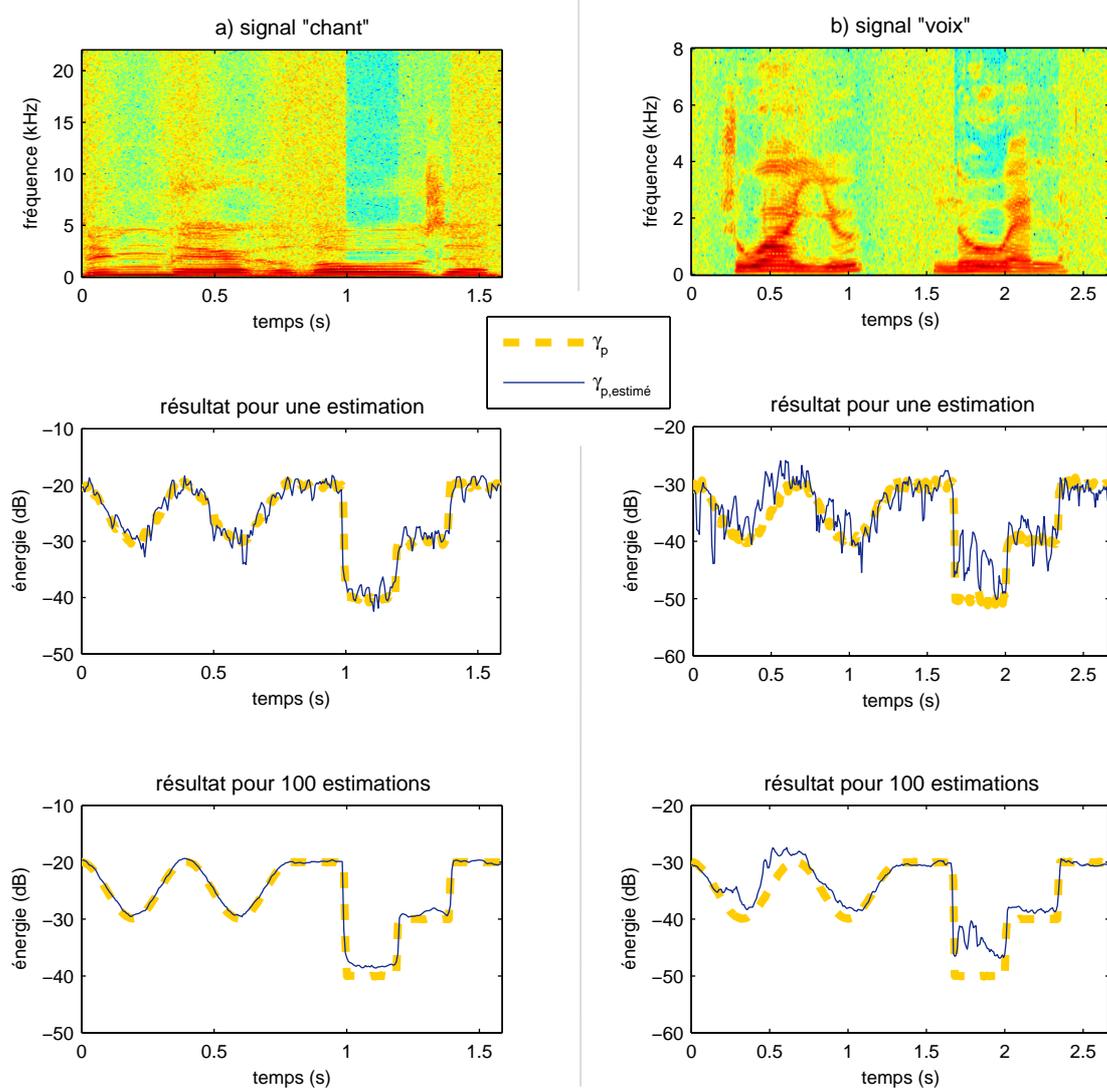


FIG. 3.11 – *Cas blanc et non stationnaire*. Évaluation de l'estimateur dans le cas d'un bruit blanc et non stationnaire. Pour chacun des deux signaux test, le résultat d'une estimation ainsi que la moyenne de 100 estimations sont données.

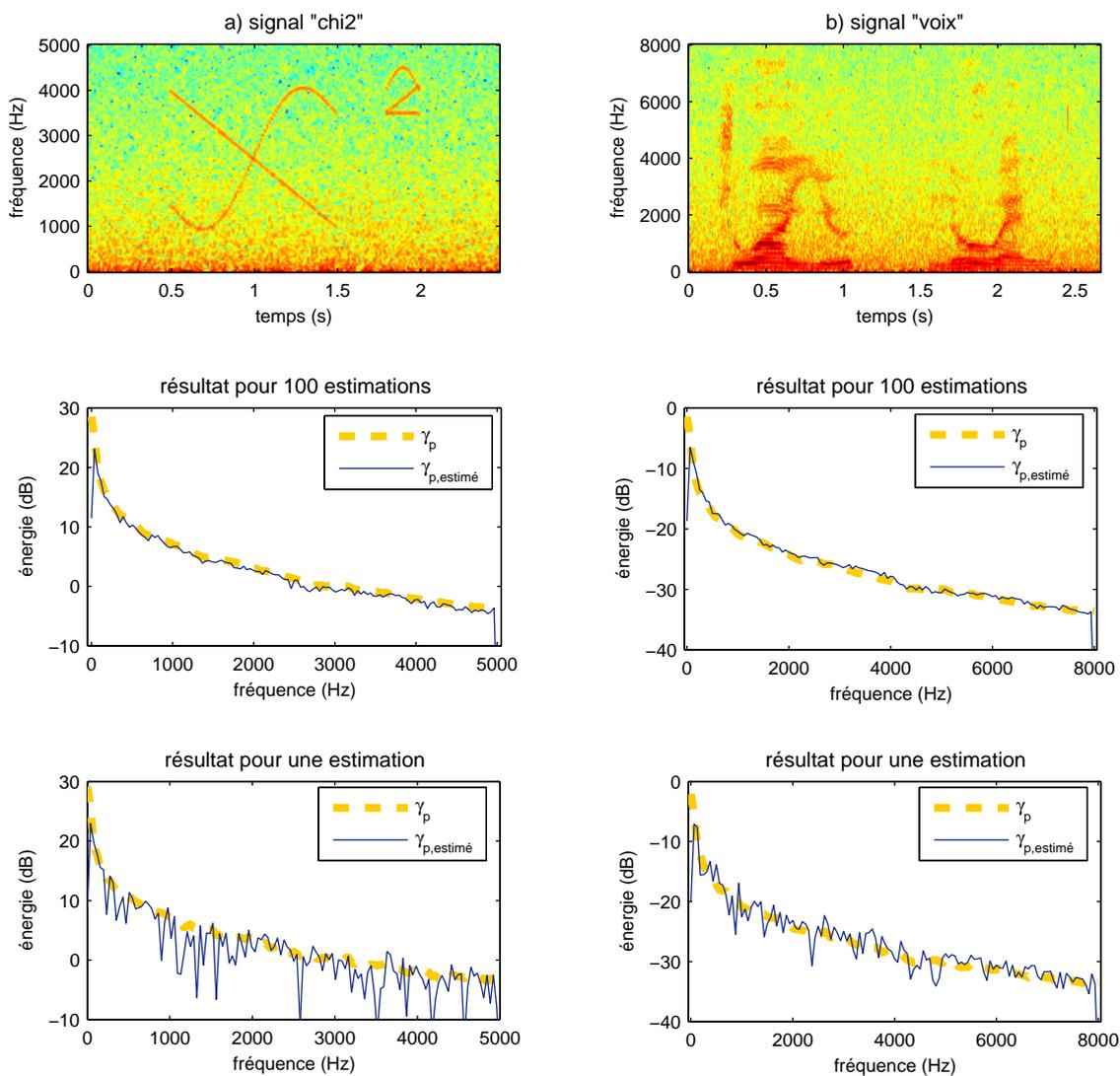


FIG. 3.12 – *Cas non blanc et stationnaire*. Évaluation de l'estimateur dans le cas d'un bruit non blanc et stationnaire. Pour chacun des deux signaux test, le résultat d'une estimation ainsi que la moyenne de 100 estimations sont données.

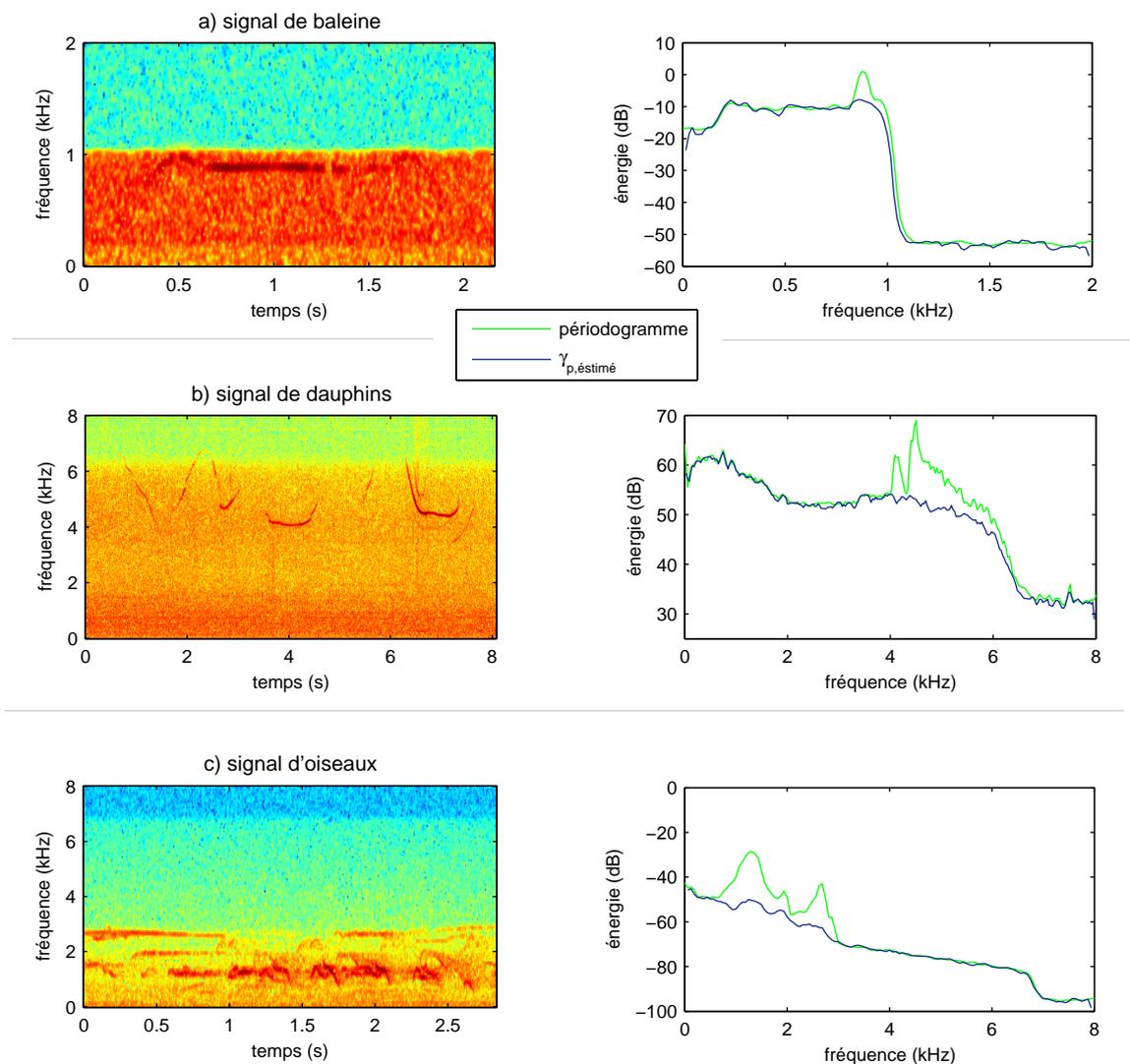


FIG. 3.13 – *Cas non blanc et stationnaire.* Évaluation de l'estimateur sur trois signaux réels présentant une perturbation naturelle non blanche et stationnaire. Sur les figures de droite, la courbe verte (claire), nommée 'périodogramme', correspond à la densité spectrale de puissance des signaux, estimée par la méthode de Welch, c'est à dire par moyennage du spectrogramme selon l'axe temporel. La courbe bleu marine (foncée) donne l'estimation obtenue de la ligne de fond.

Ainsi, un biais de -1 correspond à une estimation 10 fois plus petite que la valeur réelle. Un biais de $+0.3$ correspond à une estimation 2 fois plus grande que la valeur réelle.

Dans l'ensemble, une tendance à la sous-estimation peut être notée. Cette sous-estimation est notamment visible au niveau des fortes non-stationnarités ($t = 1.5, 1.9$ et 2.2 sec) pour lesquelles la dimension temporelle du voisinage ne permet pas une adaptation synchrone. Ce phénomène est plus précisément visible sur la figure 3.14-c), qui montre le résultat de l'estimation à une fréquence donnée (indiquée sur le spectrogramme). L'estimateur anticipe la diminution de la dsp du bruit ($t = 1.5$ sec) et présente ensuite un temps de latence lorsque la dsp du bruit augmente à nouveau ($t = 1.9$ et 2.2 sec). La figure 3.14-e) présente le résultat de l'estimation du spectre fréquentiel du bruit à un temps donné. La tendance à la sous-estimation est visible au niveau des basses fréquences pour lesquelles les variations d'amplitudes de la dsp du bruit sont plus importantes.

Cet exemple synthétique permet de mettre en avant les défauts inhérents à l'estimateur quant aux fortes variations de la dsp de la perturbation. Ce défaut est bien entendu à relier avec la taille du voisinage d'estimation. Il ne reflète cependant pas les défauts de l'estimateur liés au signal analysé. Le signal "chi2" est effectivement très simple au sens où il n'admet pas de variations d'amplitude et possède un support temps-fréquence bien défini et très étroit. La taille du voisinage utilisé est ici toujours suffisante pour englober les zones signal tout en gardant un nombre important de coefficients bruit seul.

La qualité de l'estimation est mise à défaut par le signal de parole "voix", pour lequel les résultats sont présentés sur la figure 3.15. Le même profil non-stationnaire est utilisé sur ce signal, par contre la corrélation est cette fois de type bruit rose (puissance décroissante avec les fréquences). La figure 3.15-e), autour de 6000 Hz, présente le cas typique d'une grande région "signal" sur laquelle l'estimateur proposé ne pourra pas fonctionner : la taille du voisinage est trop petite pour englober cette région et en plus admettre les Z points "bruit seul" nécessaires. L'estimation se base donc principalement sur des points "signal" et s'en trouve naturellement fortement biaisée. Ce type d'erreur d'estimation se trouve encore accentué pour le signal de piano, dont les résultats sont présentés sur la figure 3.16. La zone de signal (à partir de $t = 1$ sec et pour $f < 2$ kHz) étant très grande par rapport à la taille du voisinage, la valeur estimée n'a plus aucun rapport avec la densité spectrale du bruit.

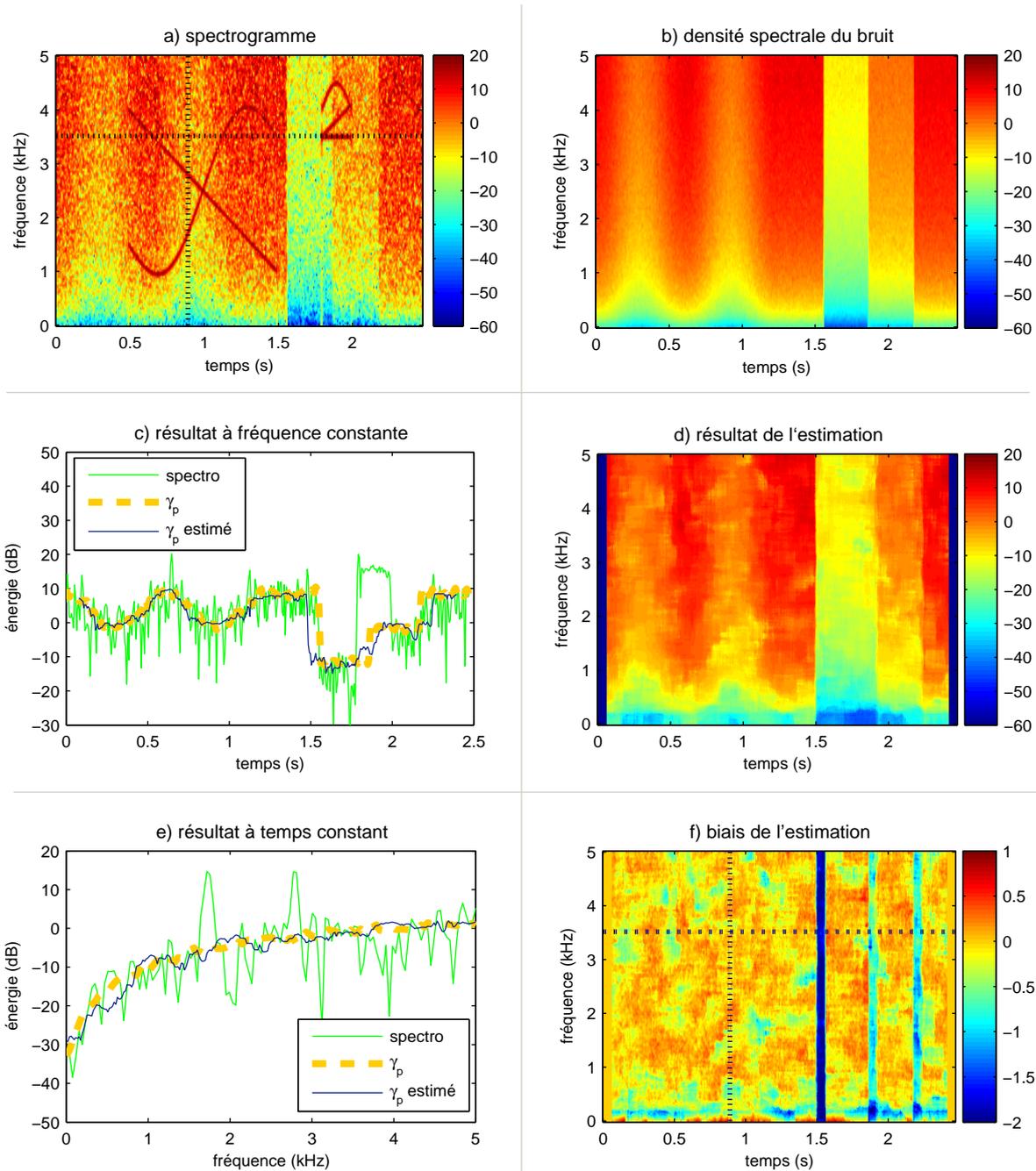


FIG. 3.14 – *Cas non blanc et non stationnaire*. a) spectrogramme du signal bruité, b) densité spectrale de puissance du bruit à estimer, c) résultat de l'estimation à la fréquence indiquée, d) résultat global de l'estimation, e) résultat de l'estimation au temps indiqué et f) carte des biais logarithmiques.

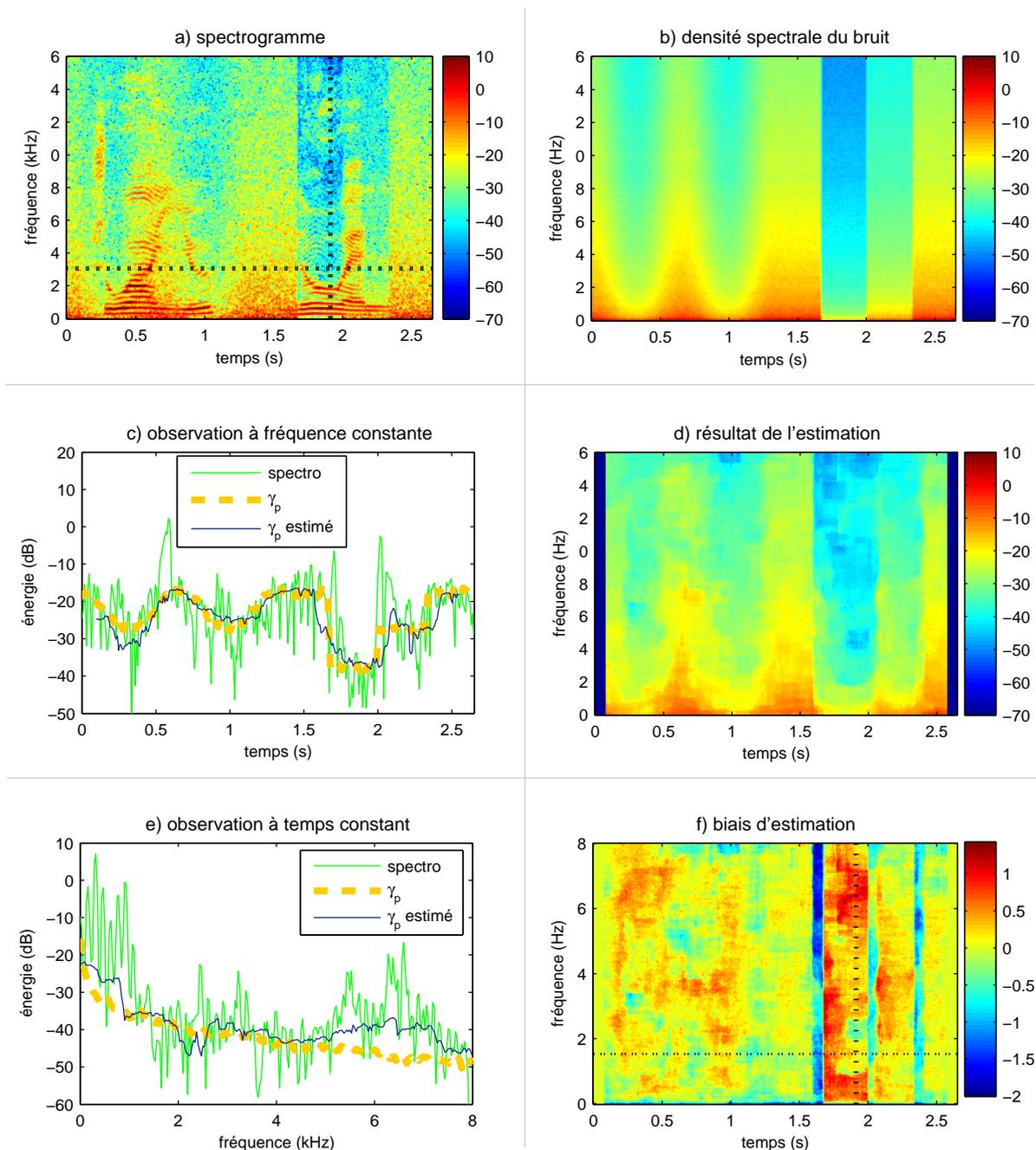


FIG. 3.15 – *Cas non blanc et non stationnaire.* a) spectrogramme du signal bruité, b) densité spectrale de puissance du bruit à estimer, c) résultat de l'estimation à la fréquence indiquée, d) résultat global de l'estimation, e) résultat de l'estimation au temps indiqué et f) carte des biais logarithmiques.

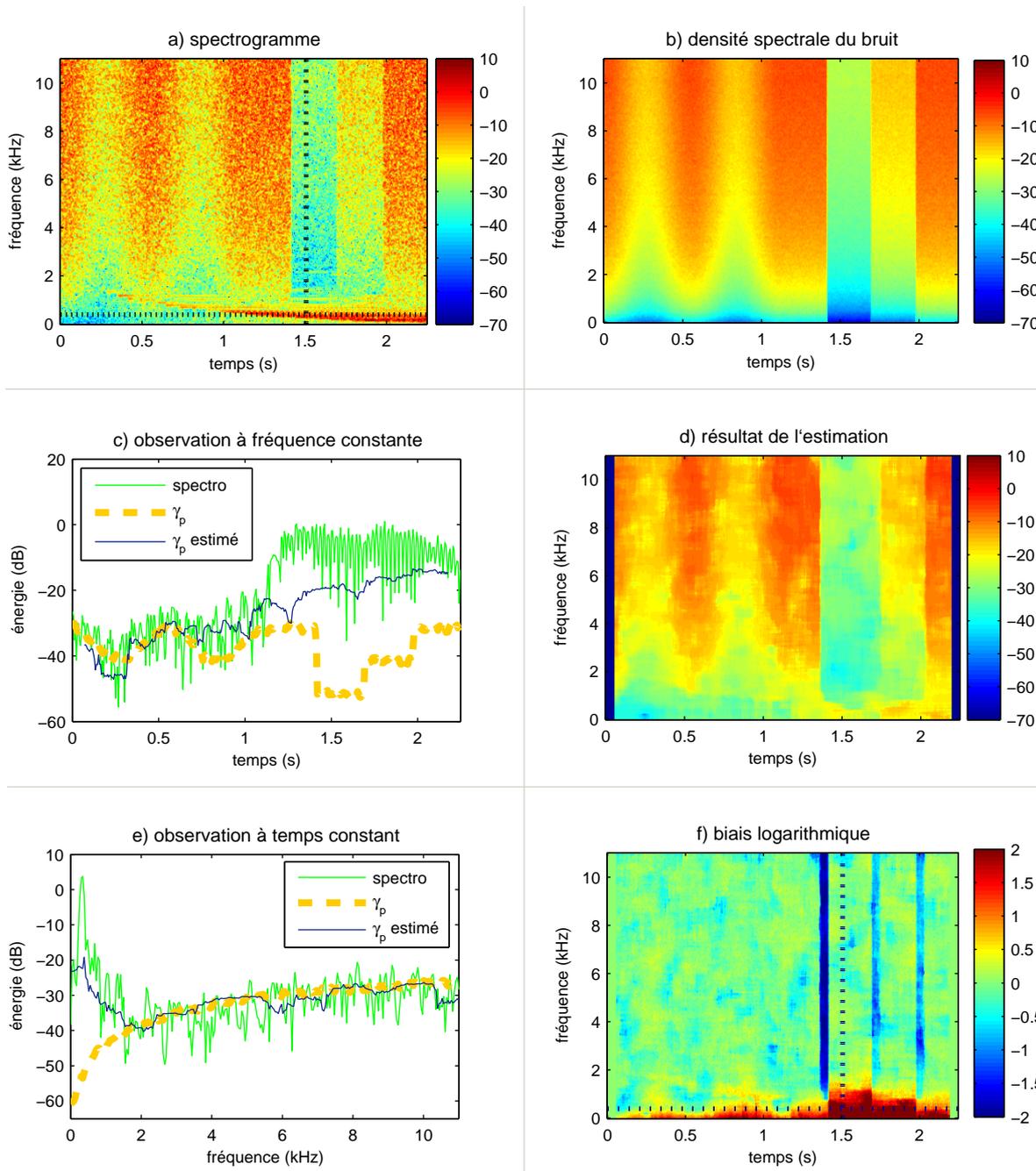


FIG. 3.16 – *Cas non blanc et non stationnaire*. a) spectrogramme du signal bruité, b) densité spectrale de puissance du bruit à estimer, c) résultat de l'estimation à la fréquence indiquée, d) résultat global de l'estimation, e) résultat de l'estimation au temps indiqué et f) carte des biais logarithmiques.

3.5 Conclusion

Dans le cadre d'un modèle d'observation "signal inconnu + bruit gaussien inconnu", l'estimation de la densité spectrale de puissance du bruit à partir d'un spectrogramme se heurte à deux problèmes majeurs : d'une part la dsp recherchée est accessible via une espérance, et d'autre part l'indétermination sur le signal impose à l'estimateur d'être uniquement basé sur des coefficients ne contenant que du bruit.

Sur la question de l'espérance, deux approches sont proposées dans la littérature. La première approche, notamment utilisée dans le traitement des signaux audio [Coh03][RL06], prend la forme d'un lissage temporel récursif, généralement d'ordre 1, des coefficients temps-fréquence. La seconde approche, adoptant un point de vue plus proche du traitement des images, consiste à définir un voisinage temps-fréquence dans lequel des statistiques peuvent être développées [HMC02][YR06][MHM06].

Dans l'approche "lissage", la prise en compte unique des coefficients "bruit seul" prend la forme d'une adaptation du paramètre de lissage à la présence ou non de signal dans l'échantillon rencontré. Ce dernier point représente alors le cœur du problème. Lorsqu'un voisinage est défini, l'indétermination sur le signal nécessite la recherche des coefficients "bruit seul" qu'il contient. Si tous les coefficients "bruit seul" peuvent être déterminés, l'estimateur optimal de la dsp du bruit consiste à réaliser la moyenne de ces coefficients.

Devant la difficulté de cette recherche exhaustive des coefficients bruit seul, une alternative basée sur la statistique du minimum a été proposée dans [Mar01] : le plus petit coefficient observé dans un voisinage étant le plus probable des coefficients "bruit seul", multiplier ce coefficient par un gain défini suivant la statistique du minimum permet de retrouver la moyenne de l'ensemble des coefficients "bruit seul", et donc la dsp du bruit. C'est cette approche, non plus restreinte au strict minimum mais aux Z plus petits coefficients, que nous avons étudié dans ce chapitre. Des conclusions sur trois points peuvent être formulées :

L'estimateur :

Un estimateur basé sur le comportement statistique des Z plus petites réalisations parmi N réalisations d'une variable aléatoire du χ^2 a été formulé. Dans le cas où le paramètre N est connu, cet estimateur est non biaisé et sa variance est égale à γ_p/Z . Cette variance est dans un rapport N/Z en comparaison avec l'estimateur optimal utilisant l'ensemble des N réalisations. Cette augmentation de variance est donc la contrepartie à l'utilisation d'une moindre quantité d'information.

La discrimination bruit/signal :

La recherche de l'ensemble des coefficients "bruit seul" au sein d'un voisinage représente une discrimination exhaustive entre bruit et signal. Si l'approche des statistiques minimales ne nécessite pas cette recherche, la prise de décision pour discriminer le bruit du signal n'en est cependant pas absente. Cette aspect apparaît au niveau de chacun des deux paramètres Z et N .

Les Z coefficients retenus pour l'estimation doivent être exclusivement issus du bruit. Le choix d'une valeur pour Z représente donc une première discrimination entre bruit et signal. Un avantage de l'approche étudiée est de restreindre ce premier choix à une zone de petites valeurs pour laquelle la discrimination est 'probablement' plus facile. La conséquence d'une mauvaise décision à ce niveau est un estimateur biaisé si des éléments signal ont été gardés, et une variance non minimale si quelques points "bruit seul" de plus auraient pu être gardés.

Le second lieu de discrimination bruit/signal est la détermination du paramètre N . Au sens où elle concerne l'ensemble des coefficients observés, cette décision revêt le même caractère exhaustif que la recherche de tous les coefficients "bruit seul". Néanmoins, seul leur nombre est dorénavant

nécessaire, ce qui simplifie un peu le problème. Les conséquences d'une mauvaise décision à ce niveau sont par contre importantes : l'erreur sur l'estimation de N se répercute à l'identique sur la valeur estimée de γ_p .

Pour finir sur ce point, notons que la proportion de coefficients "bruit seul" contenus dans le voisinage ainsi que le rapport signal sur bruit sont les deux facteurs qui déterminent la difficulté de cette tâche de discrimination bruit/signal. Aussi, l'évaluation de l'estimateur a montré qu'un rapport signal sur bruit supérieur à 10dB assure une bonne estimation de la densité spectrale de puissance du bruit (figures 3.5 et 3.7).

Le voisinage :

Le choix du voisinage sur lequel la dsp de la perturbation est estimée est un élément déterminant du processus d'estimation. Deux conditions majeures peuvent être formulées à son sujet :

- la présence de coefficients "bruit seul" dans le voisinage est une condition impérative au fonctionnement de l'estimateur,
- la dsp du bruit doit être constante sur tout le voisinage.

La première condition peut être mise en défaut par le signal. La taille du voisinage doit être suffisante pour englober les plus grandes composantes temps-fréquence du signal et en plus contenir quelques points "bruit seul". La seconde condition peut être mise en défaut par le bruit. La taille du voisinage ne doit pas être trop importante lorsque le bruit admet des variations de grande amplitude et localisées dans le plan temps-fréquence. Pour de tels cas, l'estimateur sera naturellement biaisé vers les basses valeurs.

Aussi, nous avons vu comment la forme du voisinage peut être adaptée lorsque le caractère blanc et/ou stationnaire de la perturbation sont connus. Les performances de l'estimateur en sont alors très favorablement améliorées. D'un point de vue plus général, ce dernier aspect renvoie à une caractérisation de ce qu'est le bruit par rapport à ce qu'est le signal. Il peut paraître utile que le bruit ne soit pas défini d'une façon figée et reste libre de contenir ce que nous ne souhaitons pas prendre en compte, c'est à dire ce qui n'est pas du type de l'information que l'on cherche. La spécification d'un caractère blanc ou stationnaire pour la perturbation traduit donc par opposition le caractère non blanc ou non stationnaire de l'information.

Chapitre 4

Détection du signal en un point du plan temps-fréquence

Sommaire

4.1	Introduction	83
4.2	Tests binaires d'hypothèses	83
4.2.1	Approche bayésienne	84
4.2.2	Approche de Neyman-Pearson	86
4.3	Critères de détection retenus et statistique de décision	87
4.3.1	Critère de Neyman-Pearson	88
4.3.2	Critère du maximum de vraisemblance	89
4.3.3	Rapport signal sur bruit de focus	91
4.3.4	Statistique de décision	92
4.4	Évaluation du détecteur	94
4.4.1	Courbes de performances dans le cas clairvoyant	94
4.4.2	Non optimalité et performances réelles	96
4.5	Conclusion	98

4.1 Introduction

Nous avons décrit dans le chapitre 1 comment la localisation du support temps-fréquence d'un signal déterministe noyé dans un bruit aléatoire et additif peut prendre la forme d'un test binaire d'hypothèses, formulé en chaque point du plan temps-fréquence. Pour mémoire, nous rappelons ici la formulation de ce test, déjà établie en (1.5),

$$\begin{cases} H_0[n, k] : S_x[n, k] = S_p[n, k], \\ H_1[n, k] : S_x[n, k] = S_{d+p}[n, k]. \end{cases} \quad (4.1)$$

Le support temps-fréquence du signal d'intérêt, ici $d[m]$, correspond alors à l'ensemble des positions $[n, k]$ pour lesquelles l'hypothèse H_1 est retenue.

Dans l'optique de résoudre ce test d'hypothèses, la densité de probabilité des coefficients $S_x[n, k]$ a été explicitée dans le chapitre 2 dans le cas d'une perturbation gaussienne. Il en est ressorti que cette densité de probabilité dépend du paramètre $\gamma_p[n, k]$ défini en (1.14) comme l'espérance du coefficient $S_x[n, k]$ sous l'hypothèse 'bruit seul' H_0 . Dans le chapitre 3, une méthode d'estimation de ce paramètre a été proposée.

Ayant en main ces deux pré requis, le présent chapitre est dédié à la détermination et à l'évaluation d'une fonction de décision permettant de décider entre les deux hypothèses H_0 et H_1 du test (4.1). Dans le cas présent, cette fonction de décision prend la forme d'un seuil de détection $S^{\text{seuil}}[n, k]$ auquel la valeur $S_x[n, k]$ doit être comparée. La détermination d'un seuil de détection résulte du choix d'un critère de décision. Ce critère correspond à une caractéristique que doit satisfaire la fonction de décision. Nous verrons que différents critères mènent à différentes expressions du seuil de détection faisant intervenir différents paramètres. Le caractère intuitif, pratique ou connu de l'un ou l'autre de ces paramètres permet alors de retenir l'expression correspondante pour la détermination du seuil de détection.

Ce 4^{ème} chapitre est construit comme suit. Dans un premier temps, les principes généraux de résolution d'un test binaire d'hypothèses sont exposés. Notamment, les approches de Bayes et Neyman-Pearson pour la définition d'un critère de décision sont présentées. Dans un second temps, ces principes sont dérivés dans le cas du test d'hypothèses (4.1). Les seuils de détection relatifs aux critères de Neyman-Pearson puis du maximum de vraisemblance sont explicités. Le choix de l'un ou l'autre de ces critères pour établir la fonction de décision entre les hypothèses H_0 et H_1 est ensuite discuté. Sur la base de ces résultats, les performances du détecteur ainsi construit en chaque point du plan temps-fréquence sont finalement évaluées.

4.2 Tests binaires d'hypothèses

Étant donnée deux hypothèses H_0 et H_1 relatives à une observation x de dimension finie, un test d'hypothèses consiste à établir une fonction de décision $\phi(x)$ permettant d'associer l'observation avec l'une ou l'autre des hypothèses. En notant \mathcal{R} l'espace dans lequel l'observation x est représentée, il s'agit de construire deux régions \mathcal{R}_0 et \mathcal{R}_1 formant une partition de \mathcal{R} , telles que la fonction de décision soit définie selon [Sch91]

$$\phi(x) = \begin{cases} 0 & \text{si } x \in \mathcal{R}_0 \text{ (} H_0 \text{ est décidée),} \\ 1 & \text{si } x \in \mathcal{R}_1 \text{ (} H_1 \text{ est décidée).} \end{cases} \quad (4.2)$$

Lorsque les observations sont aléatoires, nous verrons que la construction des deux régions \mathcal{R}_0 et \mathcal{R}_1 , et de manière équivalente le choix entre H_0 et H_1 , s'assimile à un test du rapport

de vraisemblance. Plus précisément, la décision consiste à comparer le rapport de vraisemblance $\mathcal{L}(x)$ de l'observation avec un seuil μ , soit

$$\mathcal{L}(x) = \frac{p_{H_1}(x)}{p_{H_0}(x)} \underset{H_0}{\overset{H_1}{\geq}} \mu, \quad (4.3)$$

dans lequel $p_{H_0}(x)$ et $p_{H_1}(x)$ représentent les densités de probabilité des observations x sous les hypothèses H_0 et H_1 ou 'conditionnellement' à ces hypothèses.

La prise de décision peut donner lieu à quatre situations, pour lesquelles les dénominations propres au traitement du signal seront utilisées :

- a) une 'non-détection' juste : l'hypothèse H_0 est décidée avec raison,
- b) un 'manque' : l'hypothèse H_0 est décidée tandis que l'hypothèse H_1 est vraie,
- c) une 'détection' juste : l'hypothèse H_1 est décidée avec raison,
- d) une 'fausse alarme' : l'hypothèse H_1 est décidée tandis que l'hypothèse H_0 est vraie.

À chacune des quatre situations correspond une probabilité d'occurrence. En fonction des régions de décision \mathcal{R}_0 et \mathcal{R}_1 , et de la probabilité des observations, nous avons pour la probabilité de non-détection juste,

$$P_{nd} = \int_{\mathcal{R}_0} p_{H_0}(x) dx, \quad (4.4)$$

la probabilité de manque,

$$P_m = \int_{\mathcal{R}_0} p_{H_1}(x) dx, \quad (4.5)$$

la probabilité de détection,

$$P_d = \int_{\mathcal{R}_1} p_{H_1}(x) dx, \quad (4.6)$$

et la probabilité de fausse alarme,

$$P_{fa} = \int_{\mathcal{R}_1} p_{H_0}(x) dx. \quad (4.7)$$

Van Trees [VT68] décrit alors l'enjeu d'un test binaire d'hypothèse comme la définition de l'importance relative qu'il faut attribuer à chacune des quatre issues, et la définition de la fonction de décision $\phi(x)$ qui en découle. Nous présentons maintenant deux approches qui vont dans ce sens, l'approche bayésienne et l'approche de Neyman-Pearson.

4.2.1 Approche bayésienne

Afin de déterminer les régions de décision \mathcal{R}_0 et \mathcal{R}_1 , le critère de Bayes consiste à minimiser le risque du test, le risque étant défini comme l'espérance du coût de la décision. Ce critère se fonde sur deux hypothèses :

- d'abord, une probabilité *a priori* peut être attribuée à chacune des deux hypothèses H_0 et H_1 . Elles sont respectivement notées p_0 et p_1 .
- ensuite, à chacune des quatre situations peut être attribué un coût de décision, notés C_{00} , C_{01} , C_{11} et C_{10} en respectant l'ordre de l'énumération donnée plus haut.

Le risque bayésien \mathfrak{R}_b associé à un test binaire d'hypothèses est alors défini par

$$\begin{aligned}\mathfrak{R}_b &= p_0 C_{00} \text{P}\{\text{décide } H_0 \text{ sachant que } H_0 \text{ est vraie}\} \\ &+ p_0 C_{10} \text{P}\{\text{décide } H_1 \text{ sachant que } H_0 \text{ est vraie}\} \\ &+ p_1 C_{11} \text{P}\{\text{décide } H_1 \text{ sachant que } H_1 \text{ est vraie}\} \\ &+ p_1 C_{01} \text{P}\{\text{décide } H_0 \text{ sachant que } H_1 \text{ est vraie}\},\end{aligned}\quad (4.8)$$

soit en utilisant les probabilités (4.4)-(4.7),

$$\mathfrak{R}_b = p_0[C_{00}P_{nd} + C_{10}P_{fa}] + p_1[C_{11}P_d + C_{01}P_m]. \quad (4.9)$$

En observant que $P_m = 1 - P_d$ et $P_{nd} = 1 - P_{fa}$ d'une part, et en utilisant la propriété de la partition de \mathcal{R} , soit $\mathcal{R}_1 = \mathcal{R} - \mathcal{R}_0$ d'autre part, le risque bayésien à minimiser peut s'écrire sous la forme

$$\mathfrak{R}_b = p_0 C_{10} + p_1 C_{11} + \int_{\mathcal{R}_0} [p_1(C_{01} - C_{11})p_{H_1}(x) - p_0(C_{10} - C_{00})p_{H_0}(x)] dx. \quad (4.10)$$

Si les deux premiers termes ne dépendent pas du choix des régions de décision et représentent un risque permanent, minimiser \mathfrak{R}_b revient à ne garder dans la région \mathcal{R}_0 que les termes participant négativement à l'intégrale. En supposant logiquement que $C_{01} > C_{11}$ et $C_{10} > C_{00}$ ¹, la région \mathcal{R}_0 est alors définie par l'ensemble des observations x de \mathcal{R} telles que

$$p_1(C_{01} - C_{11})p_{H_1}(x) < p_0(C_{10} - C_{00})p_{H_0}(x). \quad (4.11)$$

Ceci permet finalement d'établir la fonction de décision sous la forme d'un test du rapport de vraisemblance, soit

$$\frac{p_{H_1}(x)}{p_{H_0}(x)} \underset{H_0}{\overset{H_1}{\geq}} \frac{p_0(C_{10} - C_{00})}{p_1(C_{01} - C_{11})}. \quad (4.12)$$

où par analogie avec l'expression générale (4.3), le seuil de détection μ est défini ici par les connaissances supposées connues *a priori*, à savoir les coûts relatifs à chacune des issues possibles du test et les probabilités *a priori* des deux hypothèses.

Si le test donné en (4.12) apporte une solution générale quant au choix entre deux hypothèses, il permet d'englober plusieurs cas particuliers qui apparaissent lorsque les connaissances pré requises (probabilités *a priori* et coûts de décisions) ne sont pas toutes connues. Nous présentons deux de ces cas particuliers, correspondants à la minimisation de la probabilité d'erreur du test d'une part et au choix de la plus grande vraisemblance d'autre part. Le troisième cas particulier, apparaissant lorsque les probabilités *a priori* p_0 et p_1 ne sont pas connus et trouvant une solution de type minimax, n'est pas présenté (cette solution est par exemple décrite dans [VT68]).

Minimiser la probabilité d'erreur : Lorsque seules les probabilités *a priori* p_0 et p_1 des deux hypothèses sont connues, un test intéressant est obtenu en donnant un coût nulle aux décisions correctes, soit $C_{00} = C_{11} = 0$, et un coût égal aux décisions incorrectes, soit $C_{10} = C_{01}$. Dans ce cas, la minimisation du risque bayésien \mathfrak{R}_b conduit précisément à minimiser la probabilité d'erreur P_e du test, celle-ci étant composée de la probabilité de fausse alarme P_{fa} et de la probabilité de manque P_m , soit

$$P_e = \frac{P_m + P_{fa}}{2}. \quad (4.13)$$

¹C'est à dire que le coût d'une erreur est plus important que le coût lié à une bonne décision.

Le critère (4.12) devient alors

$$\frac{p_{H_1}(x)}{p_{H_0}(x)} \geq \frac{p_0}{p_1} \frac{H_1}{H_0}. \quad (4.14)$$

En utilisant la règle de Bayes pour les probabilités conditionnelles, on observe que cette règle de décision est équivalente à retenir l'hypothèse dont la probabilité, conditionnellement à l'observation x , est maximale, soit

$$p(H_1/x) \geq p(H_0/x), \quad (4.15)$$

dans lequel $p(H_0/x)$ et $p(H_1/x)$ dénotent les densités de probabilités des hypothèses H_0 et H_1 conditionnellement à l'observation x , appelées aussi densités de probabilité *a posteriori*. Pour cette raison, le critère (4.14) est couramment appelé critère du Maximum À Posteriori (MAP).

Maximum de vraisemblance : Dans certains cas, en plus des coûts de décision C_{ij} , les probabilités *a priori* p_0 et p_1 des deux hypothèses H_0 et H_1 ne sont pas connues. Il semble alors raisonnable de considérer les deux hypothèses comme équiprobables, et de poser ainsi $p_0 = p_1 = 1/2$. En affectant de même un coût nulle aux décisions correctes, soit $C_{00} = C_{11} = 0$, et un coût égal aux décisions incorrectes, soit $C_{10} = C_{01}$, le test (4.12) devient finalement

$$\frac{p_{H_1}(x)}{p_{H_0}(x)} \geq 1. \quad (4.16)$$

Dans ce cas, le critère de décision consiste à comparer entre elles les vraisemblances de l'observation x conditionnellement à chacune des deux hypothèses. Pour cette raison, le critère (4.16) est appelé critère du Maximum de Vraisemblance (MV) [Kay98].

4.2.2 Approche de Neyman-Pearson

L'approche bayésienne pour la détermination des régions de décision \mathcal{R}_0 et \mathcal{R}_1 se rapporte à la minimisation d'une fonction de risque, définie à partir des coûts assignés aux quatre issues du test et des probabilités *a priori* des deux hypothèses. Pour une situation où ces données ne s'appliquent pas au problème ou ne sont pas connues, l'approche de Neyman-Pearson permet de définir une stratégie de décision basée sur la minimisation sous contrainte de la probabilité d'erreur P_e du test, définie en (4.13). On se rend vite compte que la minimisation conjointe de la probabilité de fausse alarme et de la probabilité de manque (ou la maximisation de la probabilité de détection $P_d = 1 - P_m$) est impossible. l'approche de Neyman-Pearson consiste alors à construire les régions \mathcal{R}_0 et \mathcal{R}_1 de manière à maximiser la probabilité de détection P_d sous la contrainte d'une probabilité de fausse alarme P_{fa} inférieure à une certaine valeur α .

En utilisant la technique des multiplicateurs de Lagrange, ce problème se formule selon

$$\text{décide } \mathcal{R}_0 \text{ et } \mathcal{R}_1 \text{ tq } f = P_m + \mu(P_{fa} - \alpha'), \text{ avec } \alpha' < \alpha, \text{ soit minimale.} \quad (4.17)$$

En utilisant les expressions 4.5 et 4.7, la fonction f à minimiser s'écrit

$$f = \int_{\mathcal{R}_0} [p_{H_1}(x) - \mu p_{H_0}(x)] dx + \mu(1 - \alpha'). \quad (4.18)$$

Afin de minimiser f , la région \mathcal{R}_0 doit regrouper l'ensemble des observations x donnant une contribution négative à l'intégrale, c'est à dire pour lesquelles

$$p_{H_1}(x) - \mu p_{H_0}(x) < 0. \quad (4.19)$$

En assignant à \mathcal{R}_1 les observations x pour lesquelles cette contrainte n'est pas vérifiée, le critère de Neyman-Pearson mène alors au test du rapport de vraisemblance

$$\frac{p_{H_1}(x)}{p_{H_0}(x)} \underset{H_0}{\overset{H_1}{\geq}} \mu. \quad (4.20)$$

Sous réserve de continuité de la densité de probabilité $p_{H_0}(x)$, le seuil μ correspond au cas d'égalité $P_{fa} = \alpha$ [VT68]. Ceci permet de déterminer la région critique \mathcal{R}_1 , pour laquelle l'hypothèse H_1 est décidée, en vérifiant l'équation

$$P_{fa} = \int_{\mathcal{R}_1} p_{H_0}(x) dx = \alpha. \quad (4.21)$$

L'approche de Neyman-Pearson permet ainsi de définir une fonction de décision du type (4.2) en mettant l'accent sur les deux situations d'erreur pouvant intervenir dans un test binaire d'hypothèse. Plus précisément, elle maximise la probabilité de détection tout en limitant la probabilité de fausse alarme. Comme nous l'avons également vu, elle ne prend pas appuie sur un ensemble de connaissances *a priori*, relatives aux deux hypothèses et à l'impact de la décision, et diffère en cela de l'approche bayésienne exposée auparavant. Aussi, si d'un point de vue mathématique les deux critères de Bayes et de Neyman-Pearson mènent à la comparaison du rapport de vraisemblance de l'observation avec un seuil, la définition pratique des régions de décision reflète une autre différence importante entre les deux approches : seule la densité de probabilité des observations sous l'hypothèse H_0 est nécessaire dans le cadre de Neyman-Pearson, tandis que l'approche bayésienne nécessite la connaissance des probabilités conditionnellement aux deux hypothèses H_0 et H_1 . En contrepartie, les tests bayésiens assurent une décision cohérente envers chacune des hypothèses, tandis que le critère de Neyman-Pearson, exclusivement fondé sur l'hypothèse H_0 , ne permet pas ce compromis global.

4.3 Critères de détection retenus et statistique de décision

Dans cette seconde partie de chapitre, les principes généraux de résolution d'un test binaire d'hypothèses précédemment rappelés sont appliqués au cas des coefficients du spectrogramme d'un signal. Dans ce contexte temps-fréquence, le test du rapport de vraisemblance s'exprime à partir des densités de probabilités des coefficients du spectrogramme données dans le chapitre 2 par (2.47) et (2.48). Pour mémoire, nous rappelons ici ces deux densités de probabilités dont un exemple est donné sur la figure 4.1 :

$$p_{H_0}(S_x) = \frac{1}{\gamma_p} \exp\left(-\frac{S_x}{\gamma_p}\right), \quad (4.22)$$

et

$$p_{H_1}(S_x) = \frac{1}{\gamma_p} \exp\left(-\frac{S_x + S_d}{\gamma_p}\right) I_0\left(\frac{2}{\gamma_p} \sqrt{S_x S_d}\right). \quad (4.23)$$

La densité de probabilité $p_{H_0}(S_x)$ du spectrogramme sous l'hypothèse H_0 est strictement décroissante sur son intervalle de définition $[0, +\infty[$. De son côté, la densité de probabilité $p_{H_1}(S_x)$

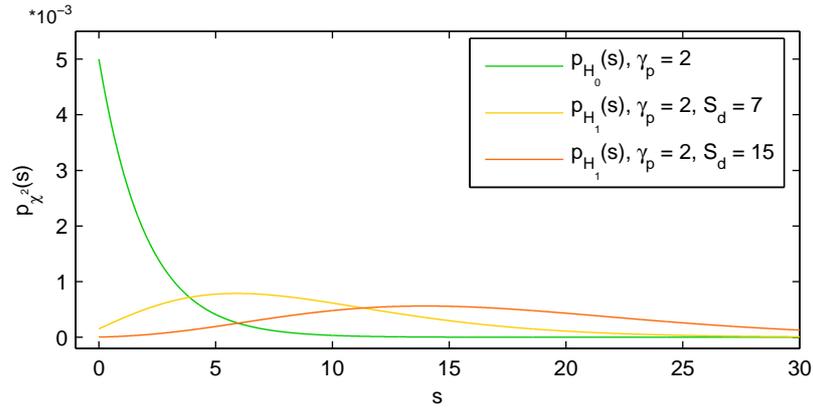


FIG. 4.1 – Exemples de lois de probabilité du spectrogramme sous les deux hypothèses H_0 et H_1 .

sous l'hypothèse H_1 , définie elle aussi sur l'intervalle $[0, +\infty[$, admet un mode à une abscisse proche du paramètre S_d (figure 4.1). Dans ces conditions, la région critique \mathcal{R}_1 que nous cherchons à déterminer prend la forme d'un intervalle ouvert de la forme $[s, +\infty[$.

Dans un premier temps, nous explicitons deux solutions pour la détermination des régions de décisions \mathcal{R}_0 et \mathcal{R}_1 . Celles-ci sont basées respectivement sur le critère de Neyman-Pearson (4.21) et le critère du maximum de vraisemblance (4.16). Dans un second temps, nous introduisons le rapport signal sur bruit de focus. Ce paramètre définit le degré d'émergence du signal par rapport au bruit à partir duquel la détection est souhaitée. Finalement, nous discutons de la statistique de décision qui intervient dans le test de détection (4.1).

4.3.1 Critère de Neyman-Pearson

Dans l'approche de Neyman-Pearson, l'enjeu est de maximiser la probabilité de détection après avoir fixé la probabilité de fausse alarme à un certain niveau. Ainsi, pour une probabilité de fausse alarme P_{fa} donnée, la région critique \mathcal{R}_1 pour laquelle l'hypothèse H_1 est décidée doit satisfaire la relation (4.21). Dans le cas du spectrogramme, la région critique \mathcal{R}_1 prend la forme d'un intervalle ouvert $[S^{seuil}, +\infty[$ et la relation (4.21) se réécrit

$$P_{fa} = \int_{S^{seuil}}^{+\infty} p_{H_0}(s) ds. \quad (4.24)$$

Le positionnement du seuil par le biais de cette relation est illustré sur la figure 4.2-a).

En introduisant dans l'équation (4.24) la densité de probabilité du spectrogramme sous l'hypothèse H_0 rappelée en (4.22), la probabilité de fausse alarme s'exprime selon

$$P_{fa} = \exp\left(-\frac{S^{seuil}}{\gamma_p}\right), \quad (4.25)$$

et le seuil de détection S_{np}^{seuil} obtenu d'après le critère de Neyman-Pearson, pour une probabilité de fausse alarme P_{fa} choisie *a priori*, est donné par

$$S_{np}^{seuil}[n, k] = \gamma_p[n, k] \ln\left(\frac{1}{P_{fa}}\right). \quad (4.26)$$

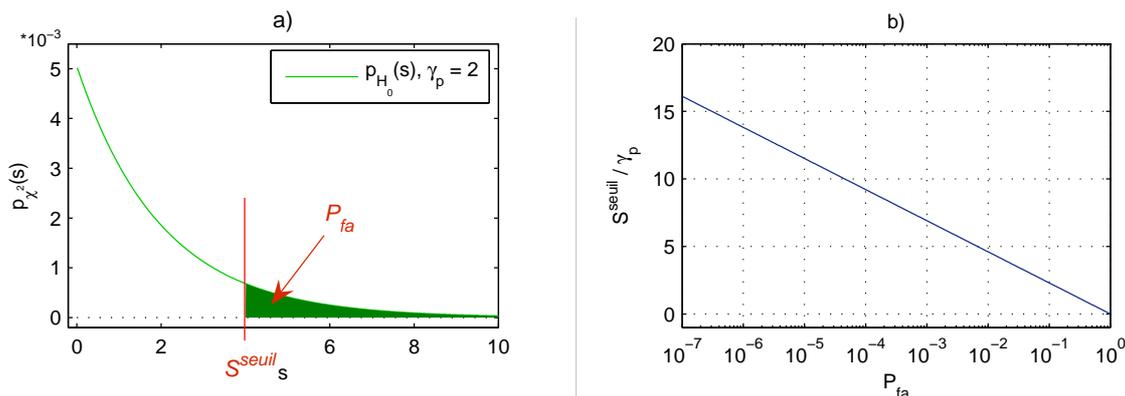


FIG. 4.2 – a) Illustration de la méthode de Neyman-Pearson pour la détermination du seuil de détection. b) Seuil de détection rapporté à la puissance spectrale du bruit, soit $S_{np}^{seuil} / \gamma_p$, en fonction de la probabilité de fausse alarme P_{fa} .

Pour illustration, la figure 4.2-b) présente la valeur du seuil de détection rapporté à la puissance spectrale du bruit, soit $S_{np}^{seuil} / \gamma_p$, en fonction de la probabilité de fausse alarme P_{fa} choisie.

Nous retenons que ce premier seuil de détection est proportionnel au niveau de bruit γ_p , le coefficient de proportionnalité étant déterminé par la PFA. Nous notons également que la PFA peut être choisie librement entre 0 et 1, ce qui autorise le seuil de détection à prendre ses valeurs sur l'intervalle $[0, +\infty[$.

4.3.2 Critère du maximum de vraisemblance

Dans ce paragraphe, un second seuil de détection est déterminé en utilisant cette fois-ci l'approche bayésienne [HM06a]. Pour un test binaire d'hypothèses, nous avons vu que la minimisation du risque bayésien (4.9) permet d'écrire le problème sous la forme du test du rapport de vraisemblance donné en (4.12) dans lequel le seuil de comparaison μ regroupe un ensemble de connaissances *a priori*.

Dans ce travail, nous ne souhaitons donner aucun *a priori* sur l'existence ou non du signal d'intérêt en un point du plan temps-fréquence. Les probabilités *a priori* sur les hypothèses sont en conséquence fixées à $p_0 = p_1 = 1/2$. De même, en l'absence de contexte applicatif, nous choisissons dans un premier temps de ne pas attribuer de coûts spécifiques aux différentes issues du test. Nous donnons ainsi un coût nulle aux décisions correctes, soit $C_{00} = C_{11} = 0$, et un coût égal aux décisions incorrectes, soit $C_{10} = C_{01}$. Par ces choix, le critère bayésien se résume au critère du maximum de vraisemblance pour lequel le test a été formulé en (4.16). La figure 4.3-a) illustre la détermination du seuil de décision correspondant à ce critère.

À partir des densités de probabilité du spectrogramme données en (4.22) et (4.23), le test du rapport de vraisemblance (4.3) utilisant le critère du maximum de vraisemblance ($\mu = 1$) s'écrit

$$\exp\left(-\frac{S_d}{\gamma_p}\right) I_0\left(\frac{2}{\gamma_p}\sqrt{S_x S_d}\right) \underset{H_0}{\overset{H_1}{\geq}} 1. \quad (4.27)$$

Afin de résoudre ce test, nous introduisons dans un premier temps le rapport signal sur bruit

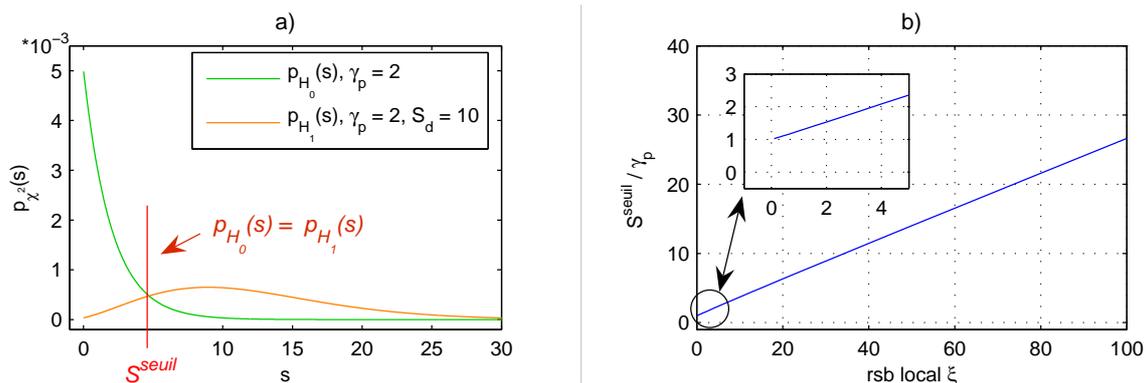


FIG. 4.3 – a) Illustration du critère du maximum de vraisemblance pour la détermination du seuil de détection. b) Solution en S^{seuil}/γ_p de l'équation (4.29) permettant d'obtenir le seuil de détection, en fonction du rapport signal sur bruit local ξ .

(rsb) local $\xi[n, k]$, défini en tout point du plan temps-fréquence selon

$$\xi[n, k] = \frac{S_d[n, k]}{\gamma_p[n, k]}. \quad (4.28)$$

En utilisant ce paramètre², le seuil de détection S^{seuil} défini par le test (4.27) doit satisfaire l'équation

$$\exp(-\xi) I_0 \left(2\sqrt{\xi \frac{S^{seuil}}{\gamma_p}} \right) = 1. \quad (4.29)$$

Nous n'avons pas pu formuler de solution analytique pour cette équation. La figure 4.3-b) en présente la solution en S^{seuil}/γ_p , obtenue numériquement en fonction du rsb local ξ . Une bonne approximation linéaire peut en être donnée sous la forme

$$\frac{S^{seuil}}{\gamma_p} = 0,26 \xi + 1, \quad (4.30)$$

ce qui conduit à définir le seuil de détection répondant au critère du maximum de vraisemblance selon

$$S_{mv}^{seuil}[n, k] = \gamma_p[n, k] (0,26 \xi[n, k] + 1). \quad (4.31)$$

De façon identique au critère de Neyman-Pearson, le seuil obtenu est à nouveau proportionnel à la densité spectrale du bruit γ_p . Par contre, le coefficient de proportionnalité est déterminé en fonction du rapport signal sur bruit local ξ . La dépendance à ce paramètre traduit son rôle discriminant dans l'analyse qui nous concerne. Il traduit l'information requise afin d'ajuster au mieux³ le seuil de détection.

L'équation (4.26) reliant le seuil de détection à la probabilité de fausse alarme reste bien entendue valable dans le cas présent. Le rsb local détermine le seuil de détection auquel correspond

²La dénomination de 'rapport signal sur bruit *a priori*' a été introduite pour ce paramètre par McAulay et Malpass [MM80] puis reprise par Ephraïm et Malah [EM84], en contraste avec le 'rapport signal sur bruit *a posteriori*' réservé au rapport $S_x[n, k]/\gamma_p[n, k]$.

³Au sens du maximum de vraisemblance...

une probabilité de fausse alarme. De même, le choix d'une PFA détermine un seuil de détection qui est adapté à un certains rsb local au sens du maximum de vraisemblance. Pour un seuil de détection donné, PFA et rsb local sont reliés selon :

$$\ln\left(\frac{1}{P_{fa}}\right) = 0,26 \xi[n, k] + 1. \quad (4.32)$$

Néanmoins, en comparaison avec l'approche de Neyman-Pearson présentée auparavant, l'apport d'information qui accompagne les approches bayésiennes⁴ se traduit dans le cas présent de la manière suivante. Les coefficients d'un spectrogramme étant définis non-négatifs, il en est de même pour le rsb local, soit $\xi[n, k] \geq 0$. Il découle alors de la relation (4.32) qu'une PFA supérieure à $1/e \simeq 0.368$ ne correspond à aucun rapport signal sur bruit local. Ceci nous informe que le gain de détection ne pourra être amélioré en élevant la PFA au-dessus de cette valeur. Ce comportement s'exprime également par le fait que le seuil de détection (4.31) est toujours plus grand que la puissance γ_p du bruit, ce qui apparaît sur la figure 4.3-b) lorsque ξ tend vers 0.

4.3.3 Rapport signal sur bruit de focus

En supposant la puissance γ_p du bruit connue ou estimée, le spectrogramme $S_d[n, k]$ du signal déterministe reste pour l'instant inconnu et il en est donc de même pour le rsb local $\xi[n, k]$. Par conséquent, la relation (4.31) n'est pas directement utilisable pour déterminer le seuil de détection. Une première solution consiste à estimer le rsb local. Une méthode d'estimation de ce paramètre est par exemple proposée dans [EM84] dans un cadre gaussien. Cependant, l'estimation du rsb local s'appuie nécessairement sur la connaissance de la puissance γ_p du bruit. Lorsque celle-ci provient elle-même d'un premier processus d'estimation, l'addition des variances des estimateurs peut porter atteinte à la validité du résultat.

Dans la volonté de déterminer le support d'un signal dans le plan temps-fréquence, il nous a semblé que le contrôle du degré d'émergence du signal par rapport au bruit constituait un degré de liberté pertinent. De plus, la formulation de ce degré d'émergence en terme de rapport signal sur bruit nous est apparue intuitive. Nous avons alors choisi d'adapter le seuil de détection à un degré d'émergence dont le choix est laissé libre. Ce degré est nommé *rapport signal sur bruit de focus* et noté ξ_0 . Il correspond au rapport signal sur bruit local ξ pour lequel le détecteur est réglé de façon optimale, au sens du maximum de vraisemblance. En fonction de ce paramètre, le seuil de détection appliqué aux coefficients du spectrogramme s'écrit

$$S^{\text{seuil}}[n, k] = \gamma_p[n, k] (0,26 \xi_0 + 1). \quad (4.33)$$

Notons que pour la majorité des points temps-fréquence, le rsb local $\xi[n, k]$ est différent du rsb de focus ξ_0 choisi et la décision ne satisfait plus le critère d'optimalité.

Dans cette nouvelle situation, un support temps-fréquence est caractérisé par le rsb de focus ξ_0 avec lequel il est obtenu. D'une façon analogue à la relation (4.32), le choix d'un rsb de focus détermine la PFA qui sera associée à ce même support. Ces deux paramètres, P_{fa} et ξ_0 , jouent dès lors un rôle tout à fait équivalent au sens où ils établissent le seuil de détection. Ils représentent un critère sur la base duquel la décision est prise. Les liens entre rsb de focus et PFA, illustrés sur la figure 4.4, se formulent suivant

$$\xi_0 = \frac{1}{0,26} \left[\ln\left(\frac{1}{P_{fa}}\right) - 1 \right], \quad (4.34)$$

$$P_{fa} = \exp\left(-\left(0,26 \xi_0 + 1\right)\right). \quad (4.35)$$

⁴Même réduites comme ici au cas le moins informé.

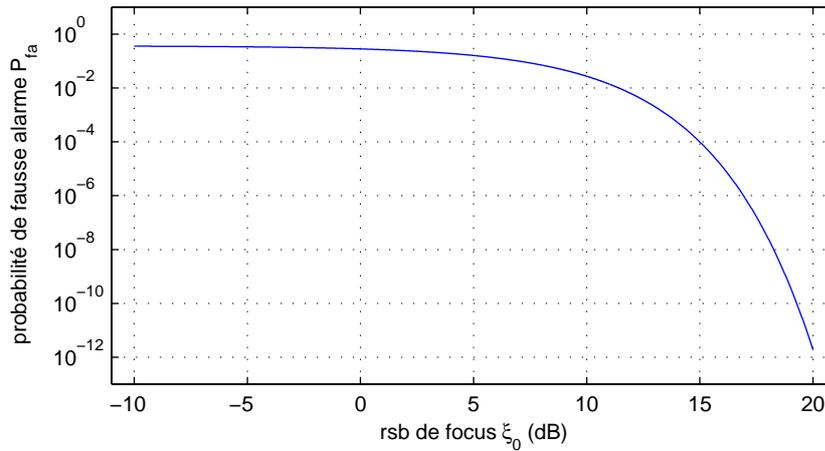


FIG. 4.4 – Lien entre la probabilité de fausse alarme P_{fa} et le rapport signal sur bruit de focus ξ_0 (en dB) relatif pour le test de détection mis en place en un point du spectrogramme.

P_{fa}	ξ_0 (dB)	ξ_0 (dB)	P_{fa}
10^{-1}	7	0	0,28
10^{-2}	11,4	5	0,16
10^{-3}	13,5	10	0,027
10^{-4}	15	20	$1,8 \cdot 10^{-12}$

TAB. 4.1 – Valeurs de référence pour la correspondance entre PFA et rsb de focus (en dB).

Le tableau 4.1 donne quelques valeurs de référence pour la correspondance entre la PFA et le rsb de focus (exprimé en dB soit $10 \log(\xi_0)$).

Concernant l'utilisation de l'un ou l'autre de ces critères, la probabilité de fausse alarme et le rapport signal sur bruit de focus renvoient chacun à des éléments différents et répondent ainsi à des objectifs différents. Le choix d'une probabilité de fausse alarme reflète l'importance de contrôler l'impact du bruit. De son côté, le choix d'un rsb de focus précise la puissance d'émergence minimale des éléments du signal sur lesquels l'intérêt est porté en priorité. Il permet la détermination du support temps-fréquence sur lequel le signal admet une énergie supérieure d'au moins ξ_0 dB à celle du bruit. Finalement, il permet d'indexer plusieurs supports temps-fréquence déterminés sur un même signal.

4.3.4 Statistique de décision

Dans ce paragraphe, nous discutons de la statistique de décision qui correspond au test d'hypothèse (4.1). Pour un test binaire d'hypothèses, une statistique de décision est une fonction de l'observation s dont la comparaison avec un seuil permet de décider entre les deux hypothèses en présence H_0 et H_1 . Une statistique de décision est dite *suffisante* lorsqu'elle contient toute l'information, nécessaire à la prise de décision et disponible dans les observations. La mise à jour d'une telle statistique est un élément pertinent car elle assure un traitement des observations adapté à la décision à venir. Dans un test du rapport de vraisemblance, le rapport de vraisemblance $\mathcal{L}(s)$

de l'observation s constitue une première statistique de décision dont on peut montrer qu'elle est suffisante [Sch91]. C'est sur ce rapport que se fonde la décision par comparaison avec le seuil μ (voir l'équation (4.3)).

La densité de probabilité d'une variable aléatoire fait intervenir deux éléments : le premier correspond à la (aux) réalisation(s) s de la variable aléatoire, et le second correspond à l'ensemble des paramètres de la loi. De même, le rapport de vraisemblance est fonction de ces deux types de paramètres. Lorsque le développement analytique du rapport de vraisemblance permet de séparer les contributions relatives à ces deux éléments, il en ressort une fonction $T(s)$ à appliquer sur l'observation afin de prendre la décision. Cette nouvelle fonction de l'observation constitue alors une nouvelle statistique de décision. Au même titre que le rapport de vraisemblance dont elle est issue, cette seconde statistique est également suffisante : elle contient toute l'information disponible dans les observations et nécessaire à la prise de décision.

Dans notre contexte, après développement du rapport de vraisemblance (présenté dans le paragraphe 4.3.2), le test de détection se résume à comparer le coefficient S_x du spectrogramme avec le seuil S^{seuil} donné en (4.31). En explicitant les dépendances aux différents paramètres, il y a alors équivalence entre les deux tests suivants :

$$\mathcal{L}(s, \gamma_p, S_d) = \frac{p_{H_1}(s, \gamma_p, S_d)}{p_{H_0}(s, \gamma_p)} \underset{H_0}{\overset{H_1}{\geq}} 1 \Leftrightarrow S_x \underset{H_0}{\overset{H_1}{\geq}} S_{mv}^{\text{seuil}}(\gamma_p, S_d). \quad (4.36)$$

Pour le test d'hypothèse formulé en (4.1), le coefficient S_x du spectrogramme correspond ainsi à une statistique de décision dont nous savons qu'elle est suffisante. Cependant, nous n'avons pas encore véritablement précisé le type des signaux auxquels cette décision fait référence. Autrement dit, pour quel type de signaux le spectrogramme correspond-t-il à une statistique suffisante pour leur détection dans le plan temps-fréquence ? Cette question est en réalité bien vaste et n'a pas été approfondie dans ce travail. Nous donnons simplement dans la suite un exemple de signal pour lequel le spectrogramme est une statistique de détection suffisante.

Considérons le problème de la détection d'une sinusoïde $d[m] = A \cos(2\pi k_0 m f_e / M + \varphi)$ de fréquence k_0 connue, de phase φ et d'amplitude A inconnues, dans un signal $x[m]$ de durée finie M , additivement perturbé par un bruit $p[m]$ gaussien, soit

$$\begin{cases} H_0 : x[m] = p[m] \\ H_1 : x[m] = A \cos\left(2\pi k_0 m \frac{f_e}{M} + \varphi\right) + p[m] \end{cases} \quad \text{avec } m = n_0, \dots, n_0 + M - 1. \quad (4.37)$$

Pour ce problème on montre que le module carré $|X(k_0)|^2$ de la transformée de Fourier discrète du signal $x[m]$, évaluée à la fréquence $k = k_0$, soit

$$|X(k_0)|^2 = \left| \sum_{m=1}^M x[m] e^{i2\pi k_0 m \frac{f_e}{M}} \right|^2, \quad (4.38)$$

est une statistique suffisante quant à la détection de $d[m]$ [Sch91][Kay98]. Nous reconnaissons dans cette statistique le coefficient $S_x[n_0, k_0]$ du spectrogramme du signal, construit avec une fenêtre rectangulaire. Si le spectrogramme est construit avec une fenêtre non rectangulaire, le caractère suffisant est conservé si la forme de la fenêtre s'apparente à l'enveloppe $A(m)$, dans ce cas non stationnaire, de la sinusoïde à détecter.

Le coefficient $S_x[n, k]$ du spectrogramme construit avec une fenêtre $w[m]$ de taille M est donc une statistique suffisante pour la détection d'une composante sinusoïdale de fréquence k^5 , d'amplitude proportionnelle à $w[m]$ et débutant à l'instant n^6 .

4.4 Évaluation du détecteur

En chaque point $[n, k]$ du plan temps-fréquence, les expressions (4.26) et (4.31) déterminent un seuil de détection qui permet de décider entre les hypothèses H_0 et H_1 du test (4.1). Dans cette troisième partie, nous évaluons les performances du détecteur ainsi mis en place. Notons qu'il s'agit ici d'évaluer la qualité de la décision prise en un unique point $[n, k]$ du plan temps-fréquence. La détermination du support temps-fréquence dans son ensemble sera évaluée dans le paragraphe 5.4 du chapitre suivant. Comme il l'a été décrit dans la première partie de ce chapitre, une prise de décision peut aboutir à quatre situations parmi lesquelles deux correspondent à des erreurs (les fausses alarmes et les manques) et deux correspondent à une bonne décision (les détections et non détections justes). Ayant en main l'ensemble des paramètres qui interviennent dans le test, les probabilités d'aboutir à chacune de ces quatre situations sont un ensemble de caractéristiques propres au détecteur qui permettent de juger de sa qualité. C'est donc sur ces caractéristiques que nous évaluons ici la décision qui découle de la comparaison du coefficient d'un spectrogramme avec le seuil défini selon (4.26) ou (4.31).

Le test de détection (4.1) fait intervenir deux paramètres : le spectrogramme S_d du signal déterministe et la puissance spectrale γ_p du bruit. Cependant, le seuil de détection étant déterminé en fonction γ_p , c'est uniquement de la valeur de S_d et plus précisément du rapport signal à bruit $\xi = S_d/\gamma_p$ défini en (4.28) que vont dépendre les performances du détecteur. Dans un premier temps, les caractéristiques du test seront ainsi évaluées en fonction de ce paramètre. Dans notre cas cependant, le rsb local ξ n'est pas connu. Le seuil de détection est effectivement déterminé d'après un rsb de focus ξ_0 , lui-même déterminé *a priori*. Dans un second temps, nous évaluons donc l'impact de ce choix sur les performances du test.

4.4.1 Courbes de performances dans le cas clairvoyant

Nous nous plaçons ici dans le cas idéal pour lequel le rapport signal sur bruit de focus ξ_0 correspond au rsb local $\xi[n, k]$. Ce cas correspond au détecteur "clairvoyant" au sens où l'ensemble des paramètres qui interviennent dans le test sont connus. Dans le cas présent, il est optimal au sens du maximum de vraisemblance et d'une probabilité d'erreur minimale. Notons que les performances obtenues caractérisent spécifiquement le spectrogramme en tant que détecteur temps-fréquence.

La figure 4.5 présente quelques courbes opérationnelles de réception (COR), c'est-à-dire P_d en fonction de P_{fa} , pour différents rapports signal sur bruit local ξ . Ces courbes ne tiennent pas compte du critère de détection retenu pour déterminer un seuil de détection. Elles donnent plutôt une vue d'ensemble sur les performances du détecteur. Dans le cas du spectrogramme, ces performances apparaissent très satisfaisantes pour un rsb local ξ supérieur à 10 dB, la probabilité de détection étant alors proche de 1 pour une très large gamme de probabilité de fausse alarme. Sur cette même figure, la courbe décroissante représente les points de fonctionnement du détecteur étudié ici. Ces points correspondent au choix spécifique du critère de maximum de vraisemblance pour la détermination du seuil de détection. C'est donc sur cette courbe de performance que le

⁵Plus précisément de fréquence $k \frac{f_e}{K}$ avec f_e la fréquence d'échantillonnage et K le nombre de fréquences calculées.

⁶Plus précisément à l'instant $m = nD$ avec D le pas de l'analyse glissante.

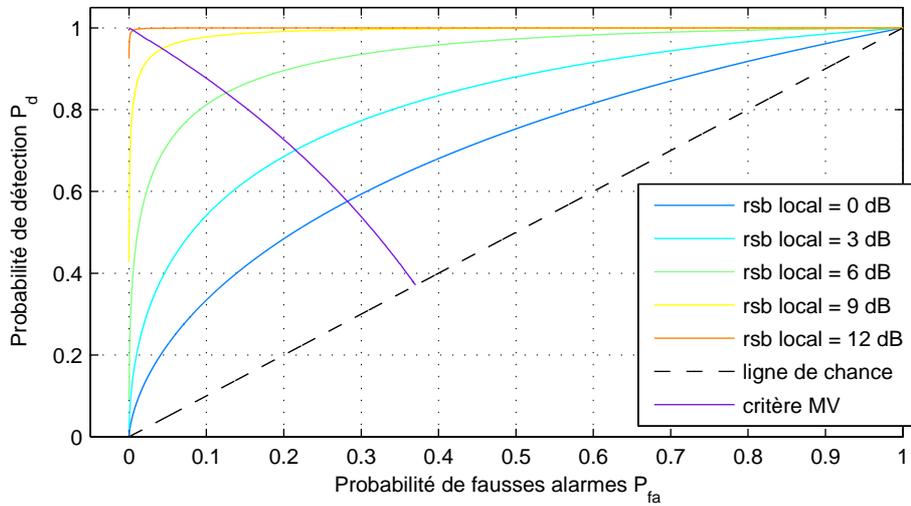


FIG. 4.5 – Courbe Opérationnelle de Réception du test de détection en un point du spectrogramme, pour différents rapports signal sur bruit local ξ . La courbe décroissante donne les points de fonctionnement correspondant au critère du maximum de vraisemblance.

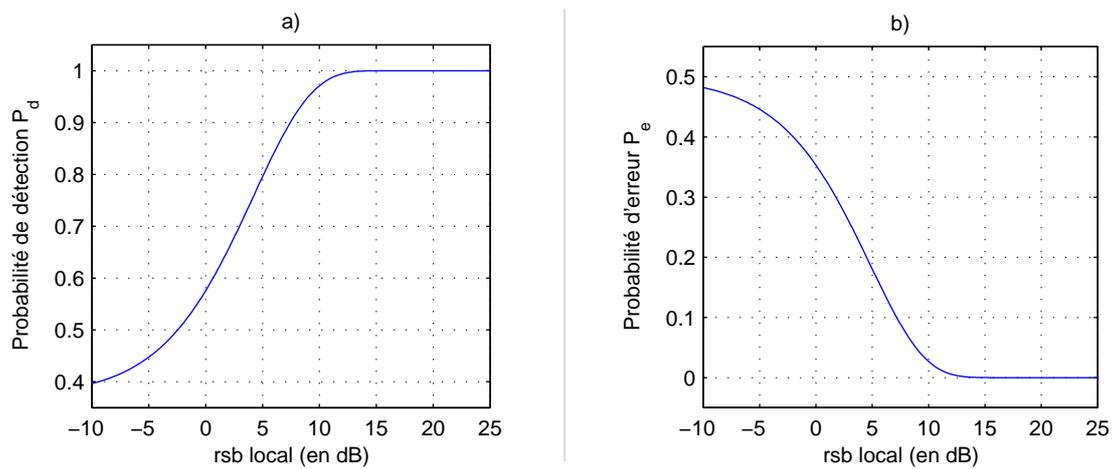


FIG. 4.6 – a) Probabilité de détection P_d et b) Probabilité d'erreur P_e en fonction du rapport signal sur bruit local ξ (exprimé en dB) pour le test de détection mis en place en un point du spectrogramme.

détecteur utilisé pour la détermination d'un support temps-fréquence se trouve⁷. Notons que la PFA se retrouve alors limitée, comme nous l'avons vu auparavant.

La figure 4.6 présente la probabilité de détection ainsi que la probabilité d'erreur (4.13) en fonction du rsb local ξ . Ces probabilités sont calculées pour le critère du maximum de vraisemblance. Nous pouvons donner quelques chiffres représentatifs : la probabilité de détection atteint 90% pour $\xi \simeq 7,4$ dB et 99% pour $\xi \simeq 11,4$ dB. De son côté, la probabilité d'erreur est inférieure à 1% à partir de $\xi \simeq 11,5$ dB, et inférieure à 0,1% à partir de $\xi \simeq 13,5$ dB. Pour ce même détecteur clairvoyant, l'évolution de la probabilité de fausse alarme en fonction du rsb local est donnée par la figure 4.4 présentée dans le paragraphe 4.3.3, en remplaçant le rsb de focus ξ_0 par le rsb local ξ .

4.4.2 Non optimalité et performances réelles

Nous insistons sur le fait que les courbes de performances présentées dans le paragraphe précédent ne sont effectives que si le rapport signal sur bruit local $\xi[n, k] = S_d[n, k]/\gamma_p[n, k]$ est connu. Ces courbes représentent donc les performances maximales que l'on peut espérer atteindre à partir du détecteur temps-fréquence utilisant le spectrogramme et le critère du maximum de vraisemblance. Dans la plupart des cas cependant, le rapport signal sur bruit local $\xi[n, k]$ est différent du rsb de focus ξ_0 choisi. Le détecteur fonctionne donc généralement de manière sous-optimal, c'est à dire que le risque bayésien (4.9) et plus précisément dans notre cas la probabilité d'erreur (4.13) ne sont plus minimisés.

Lorsque le rsb local $\xi[n, k]$ est inférieur au rsb de focus ξ_0 choisi (figure 4.7-a)), la probabilité de détection *réelle*⁸ est inférieure à la probabilité de détection *optimale*⁹. Par contre, la probabilité de fausse alarme *réelle* est inférieure à la PFA *optimale*. De façon opposée, lorsque le rsb local $\xi[n, k]$ est supérieur au rsb de focus ξ_0 choisi (figure 4.7-b)), la probabilité de détection *réelle* est supérieure à la probabilité de détection *optimale* et la probabilité de fausse alarme *réelle* est supérieure à celle obtenue par détecteur clairvoyant *optimal*. Notons que même si certaines caractéristiques semblent contribuer positivement aux performances réelles du détecteur, la probabilité d'erreur réelle reste toujours supérieure à celle obtenue lorsque le rsb local est connu. Pour les deux cas, le surplus de probabilité d'erreur correspond à l'aire des surfaces colorées de la figure 4.7.

Pour illustration, la figure 4.8 présente les probabilités de fausse alarme et de détection en fonction du rsb local ξ , obtenues pour différents choix du rsb de focus ξ_0 . Dans les mêmes conditions, la figure 4.9-a) présente la probabilité d'erreur en fonction rsb local ξ . Dans chaque cas, la caractéristique du détecteur clairvoyant est rappelée en trait discontinu.

Pour un rsb de focus ξ_0 donné, la probabilité d'erreur réelle $P_{e,r}$ du détecteur est supérieure ou égale à la probabilité d'erreur $P_{e,opt}$ du détecteur clairvoyant. Nous définissons le surplus de probabilité d'erreur, noté ΔP_e , comme la différence entre les deux probabilités d'erreur, soit

$$\Delta P_e(\xi_0, \xi) = P_{e,r}(\xi_0, \xi) - P_{e,opt}(\xi) \geq 0. \quad (4.39)$$

Nous posons alors la question de savoir s'il existe un rsb de focus ξ_0 qui minimise cette quantité. Afin d'exprimer une solution, il est cependant nécessaire de limiter l'étude à un intervalle

⁷ au mieux..., rappelons que nous supposons ici connu le rapport signal sur bruit local ξ .

⁸ mesurée à partir de la densité de probabilité paramétrée par le rsb local ξ tandis que le seuil de détection est fixé en accord avec le rsb de focus ξ_0 .

⁹ obtenue par le détecteur clairvoyant sur le critère du maximum de vraisemblance.

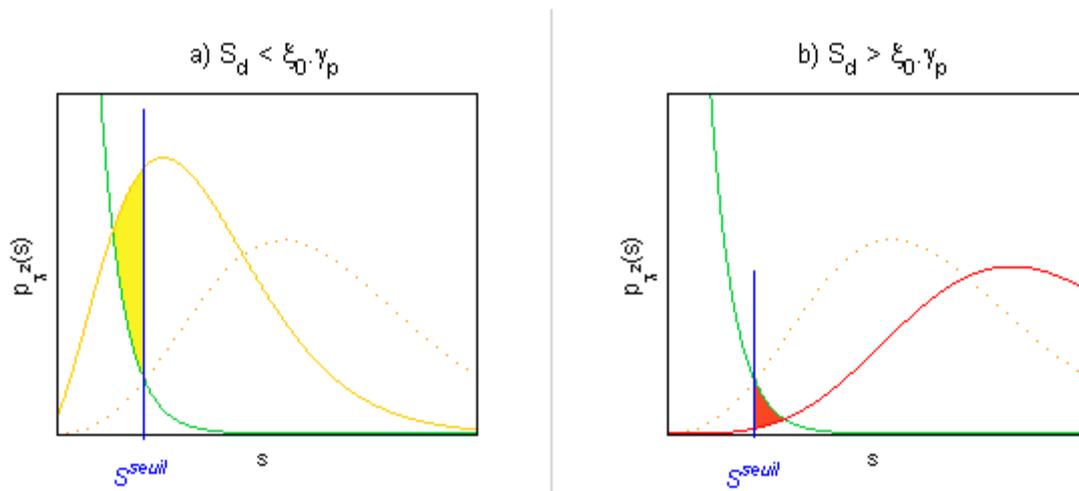


FIG. 4.7 – Variations des performances du détecteur dans le cas réel d'un rsb local inconnu. a) rsb de focus $\xi_0 >$ rsb local ξ et b) rsb de focus $\xi_0 <$ rsb local ξ . La courbe pointillée donne la densité de probabilité sous l'hypothèse H_1 qui correspond au rsb de focus ξ_0 choisi. Les zones colorées représentent le surplus de probabilité d'erreur.

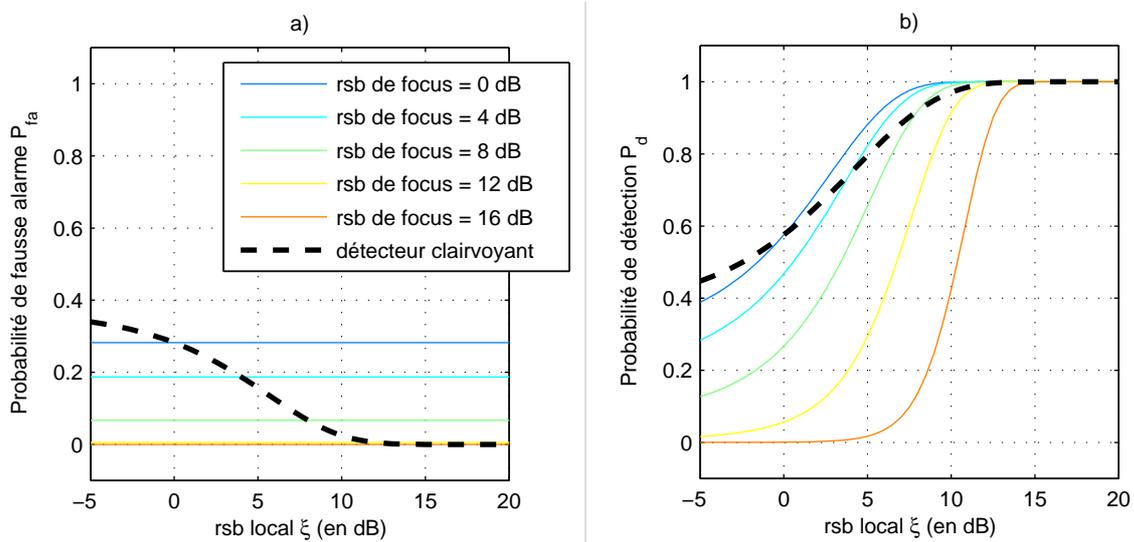


FIG. 4.8 – a) Probabilité de fausse alarme P_{fa} et b) probabilité de détection P_d en fonction du rsb local ξ , pour différents choix du rsb de focus ξ_0 .

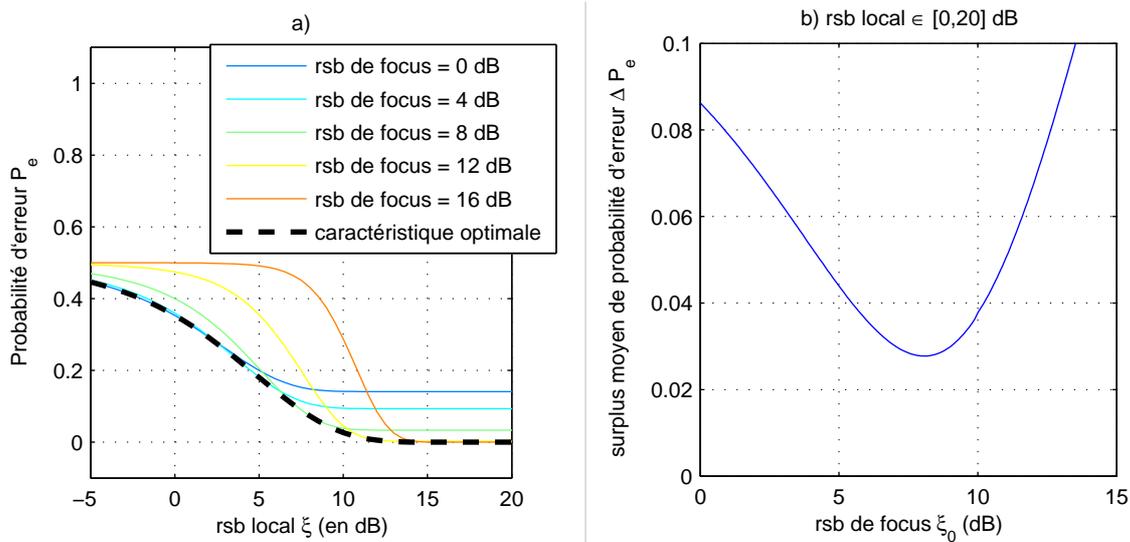


FIG. 4.9 – a) Probabilité d’erreur P_e en fonction du rsb local ξ , pour différents choix du rsb de focus ξ_0 . b) En fonction du rsb de focus ξ_0 , moyenne des surplus de probabilité d’erreur sur l’intervalle de rsb local $\xi \in [0, 20]$ dB.

$[\xi_{min}, \xi_{max}]$ de valeurs possibles pour le rsb local ξ . Nous proposons donc de minimiser la quantité

$$\frac{1}{\xi_{max} - \xi_{min}} \int_{\xi_{min}}^{\xi_{max}} \Delta P_e(\xi_0, \xi) d\xi \quad (4.40)$$

qui correspond au surplus moyen de probabilité d’erreur sur l’intervalle $[\xi_{min}, \xi_{max}]$ des rsb local ξ . L’intervalle $[\xi_{min}, \xi_{max}]$ définit ici une zone de focalisation sur laquelle le détecteur sera optimisé. La figure 4.9-b) présente l’évolution du surplus moyen de probabilité d’erreur, en fonction du rsb de focus ξ_0 , correspondant à l’intervalle de rsb local $\xi \in [0; 20]$ dB. Cet intervalle représente à nos yeux une zone critique pour laquelle une procédure d’optimisation peut bénéficier à la détermination du rsb de focus ξ_0 . Pour cet intervalle, un surplus moyen de probabilité d’erreur minimum égal à 0,028 est atteint pour un rsb de focus ξ_0 égal à 8,07 dB. Pour l’intervalle $\xi \in [-10; 60]$ dB, le surplus moyen de probabilité d’erreur est minimisé à 0,023 pour $\xi_0 = 10,1$ dB. Finalement pour l’intervalle $\xi \in [5; 10]$ dB, fixer le rsb de focus ξ_0 à 7,43 dB entraîne une augmentation moyenne de la probabilité d’erreur de 0,006.

4.5 Conclusion

Dans ce chapitre, nous avons explicité la fonction de décision permettant de décider si le point $[n, k]$ appartient au support temps-fréquence d’un signal déterministe noyé dans un bruit gaussien. Plus précisément, la décision se base sur la comparaison du coefficient $S_x[n, k]$ du spectrogramme observé avec un seuil de détection $S^{seuil}[n, k]$. À partir des critères de Neyman-Pearson et du maximum de vraisemblance, deux seuils de détection ont été formulés. Pour les deux seuils, la puissance du bruit $\gamma_p[n, k]$ joue le rôle de référence, ce qui définit le support temps-fréquence du signal comme l’ensemble des éléments qui émergent du bruit. De façon naturelle, cette émergence se caractérise par le rapport signal sur bruit au point temps-fréquence $[n, k]$ considéré qui apparaît dans la solution du test du rapport de vraisemblance.

C'est donc de ce rapport signal sur bruit local que dépendent les performances du détecteur mis en place au point $[n, k]$. Comme point de repère, nous retiendrons qu'au-delà d'un rapport signal sur bruit local de 10 dB, le spectrogramme admet les performances de détection suivantes : $P_{fa} \leq 2,4 \cdot 10^{-2}$, $P_d \geq 0.97$ et $P_e \leq 2,7 \cdot 10^{-2}$.

Dans notre cas, le rapport signal sur bruit local n'étant pas connu, nous avons choisi de l'intégrer en tant que paramètre du détecteur. Au même titre que la probabilité de fausse alarme dans l'approche de Neyman-Pearson, le choix du rsb de focus ainsi défini détermine le seuil de détection. Bien qu'imposant un caractère sous optimal au détecteur, ce paramètre lui donne cependant une flexibilité relativement intuitive. Il permet de poser explicitement le problème de la détermination d'un support temps-fréquence en terme d'un rapport signal sur bruit local à partir duquel les composantes du signal seront retenues.

Chapitre 5

Localisation du signal dans le plan temps-fréquence

Sommaire

5.1	Introduction	103
5.2	Réduction des fausses alarmes	103
5.2.1	Prise en compte de la redondance du spectrogramme	104
5.2.2	Prise en compte du rsb de focus	105
5.2.3	Compromis fausses alarmes / signal	107
5.3	Algorithme de localisation	107
5.4	Évaluation de l'algorithme	111
5.4.1	Dépendance au rsb de focus	111
5.4.2	Dépendance à l'estimation du bruit	113
5.4.2.1	Mauvaise estimation du bruit	113
5.4.2.2	Flexibilité de la discrimination bruit / signal	116
5.4.3	Limitations	118
5.4.4	Cas d'un signal d'intérêt gaussien	120
5.4.4.1	Cas général	121
5.4.4.2	Cas centré	122
5.5	Conclusion	125

5.1 Introduction

En se basant sur la caractérisation du comportement statistique des coefficients du spectrogramme établie dans le second chapitre de ce manuscrit, une fonction de décision permettant d'attribuer à un coefficient temps-fréquence une étiquette 'signal' ou 'bruit' a été décrite et évaluée dans le chapitre précédent. Afin de localiser le signal sur l'ensemble du plan temps-fréquence, cette fonction de décision doit être appliquée à tous les points de la représentation temps-fréquence. Ce processus permet d'obtenir une carte temps-fréquence regroupant l'ensemble des étiquettes 'bruit' ou 'signal' attribuées à chaque coefficient du spectrogramme. La mise en commun de ces étiquettes fait alors ressortir un ensemble de composantes temps-fréquence à partir duquel le support du signal peut être établi. C'est à cet objectif final que ce 5^{ème} et dernier chapitre est dédié. Notamment, nous abordons le problème des fausses alarmes inévitablement générées par le processus de décision. Sur ce point, nous verrons qu'un critère de taille de composante permet de réduire significativement leur nombre. Le traitement des fausses alarmes correspond à l'étape finale de l'algorithme de localisation temps-fréquence que nous proposons dans ce travail. La seconde partie du chapitre est dédiée à la présentation complète et à l'évaluation de cet algorithme. Plus précisément, nous tâcherons de décrire son comportement à travers le choix des paramètres qui sont laissés libres à l'utilisateur.

Ce chapitre est organisé de la manière suivante. Dans un premier temps, une méthode de réduction des fausses alarmes est présentée. Dans un second temps, nous donnons une vue d'ensemble sur l'algorithme de localisation temps-fréquence auquel l'ensemble de ce travail a mené. Dans un troisième temps, les résultats de cet algorithme en fonction du choix des paramètres sont présentés et discutés. Nous en donnons également quelques limitations et nous étudions finalement son comportement face à un signal d'intérêt de nature gaussienne.

5.2 Réduction des fausses alarmes

Jusqu'à présent, nous avons défini le support temps-fréquence $S^{tf}(d)$ du signal $d[m]$ comme l'ensemble des positions $[n, k]$ pour lesquelles l'observation $S_x[n, k]$ est supérieure au seuil de détection $S^{\text{seuil}}[n, k]$, soit

$$S^{tf}(d) = \{ [n, k] / S_x[n, k] > S^{\text{seuil}}[n, k] \}. \quad (5.1)$$

Cependant, nous avons vu qu'à un seuil de détection donné correspond une probabilité de fausse alarme P_{fa} , probabilité qu'un coefficient du spectrogramme uniquement représentatif du bruit $S_p[n, k]$ dépasse le seuil de détection, soit

$$P_{fa} = P \{ S_p[n, k] > S^{\text{seuil}}[n, k] \}. \quad (5.2)$$

Nous décrivons dans ce paragraphe un critère permettant de réduire sensiblement le nombre de ces points 'bruit seul' malencontreusement attribués au support temps-fréquence du signal $d[m]$ lors de sa détermination selon (5.1).

Dans la suite, l'ensemble des points du plan temps-fréquence retenus selon le critère (5.1) sera dénommé 'ensemble des points candidats'. Parmi ces points temps-fréquence candidats, les points connexes sont regroupés en composantes. La connexité est ici définie sur les 8 plus proches voisins d'un point temps-fréquence. À partir de l'ensemble des points candidats, nous définissons ainsi un ensemble de composantes temps-fréquence, candidates à devenir membre du support temps-fréquence du signal. Parmi ces composantes candidates, nous souhaitons éliminer celles qui ne proviennent que du bruit.

Le principe est d'exploiter le caractère non structuré du bruit dans la représentation temps-fréquence et parmi l'ensemble des points candidats. Ce phénomène est de plus accentué par la mise en référence du bruit dans la détermination du seuil de détection. Il y a effectivement de fortes chances que les fausses alarmes se répartissent aléatoirement et selon une distribution uniforme sur l'ensemble de plan temps-fréquence. Il en découle que les fausses alarmes consistent naturellement en un ensemble de points plus ou moins isolés les uns des autres et engendrent ainsi des composantes de taille réduites. D'un autre côté, le signal physique qui émerge de ce bruit admet par hypothèse une certaine structure dans le plan temps-fréquence. Celui-ci va donc générer une ou plusieurs composantes dont la taille, certes variable, ne peut être infiniment petite¹.

Il semble donc possible de réduire le nombre de fausses alarmes qui apparaissent avec le principe de détection (5.1) en définissant une taille minimale en dessous de laquelle une composante candidate ne sera pas retenue dans le support temps-fréquence du signal. Deux éléments doivent être pris en compte dans la définition de cette taille minimale. La redondance interne du spectrogramme d'une part et le rapport signal sur bruit de focus ξ_0 choisi pour la détection d'autre part. Ces deux points sont abordés dans les deux paragraphes qui suivent. Nous mentionnons finalement qu'une approche tout à fait similaire est proposée dans [MK07] afin de réduire le nombre de fausses alarmes issues du seuillage d'un scalogramme.

5.2.1 Prise en compte de la redondance du spectrogramme

Le spectrogramme est une représentation redondante du signal. Le théorème de Balian-Low [Bal81] stipule qu'il est impossible de créer une base de fonctions oscillantes orthogonales assurant un découpage non redondant du plan temps-fréquence. Il est ainsi impossible de créer un pavage du plan temps-fréquence suffisamment dense pour décrire toute l'information et tel que tous les coefficients soient indépendants deux à deux. Par conséquent, chaque coefficient partage une certaine quantité d'énergie avec l'ensemble des coefficients appartenant à son voisinage d'influence.

Afin de définir une taille minimale pour une composante signal, la taille de ce voisinage d'influence est un premier élément à prendre en compte car il définit un nombre de points en dessous duquel une composante détectée n'a plus de sens physique. La forme et la taille de ce voisinage d'influence sont données par le noyau reproduisant de la représentation, c'est à dire le spectrogramme de la fenêtre d'analyse $w[m]$. Ce noyau reproduisant est déterminé par tous les paramètres de l'analyse temps-fréquence qui sont, pour le spectrogramme, la forme et la taille M de la fenêtre $w[m]$, le nombre K de fréquences calculées et le pas de glissement D entre les analyses successives. Nous choisissons de définir la taille du noyau reproduisant, notée T_{noyau} , comme le nombre de coefficients de ce noyau admettant une énergie supérieure à $1/4$ de son maximum. À titre d'exemples, le tableau 5.1 donne la taille T_{noyau} du noyau reproduisant pour différentes fenêtres d'analyse de longueur $M = 256$, différents pas de glissement D et un nombre différent de fréquences calculées K ².

Nous observons logiquement que la taille du noyau est multiplié par 2 lorsque le facteur de *zero-padding* est doublé, et qu'elle diminue face à l'augmentation du pas de glissement. Dans notre situation, nous utilisons la taille du noyau reproduisant ainsi défini comme la taille minimale que doit avoir une composante détectée pour être retenu comme une composante 'signal', une composante de taille inférieure ne pouvant provenir d'un phénomène oscillant sinusoïdal.

¹La taille minimale d'une composante 'signal' est notamment limitée par le principe d'incertitude temps-fréquence : le signal ne peut être arbitrairement résolu à la fois en temps et en fréquence.

²Rappelons que le rapport K/M correspond au facteur de *zero-padding*.

D	$3M/4$	$M/2$	$M/4$	$M/8$	$M/2$	$M/4$	$M/8$	$M/4$
K	M	M	M	M	$2M$	$2M$	$2M$	$4M$
rectangulaire	2, 6	3, 4	6, 7	13, 1	5, 7	10, 9	20, 9	20, 6
hanning	3, 3	3, 8	6	11, 7	7, 2	11, 6	22, 9	22, 8
gaussienne	3, 1	3, 6	6	11, 7	7	11, 8	22, 9	23, 2
blackman	3, 7	4	6	11, 8	7, 7	11, 6	22, 6	23, 1

TAB. 5.1 – Taille T_{noyau} du noyau reproduisant pour différentes fenêtres de longueur $M = 256$ échantillons, différents pas de glissement D et nombre de fréquences calculées K .

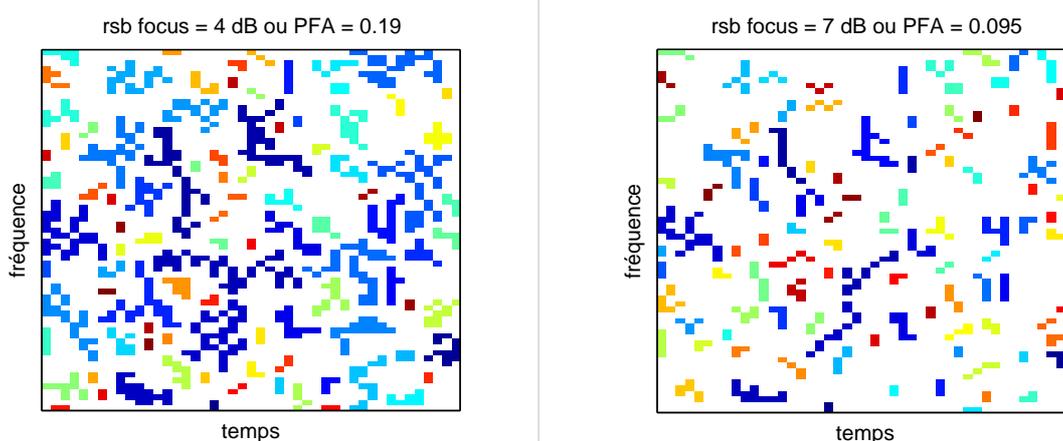


FIG. 5.1 – Illustration de la répartition et de la forme des fausses alarmes pour deux rsb de focus $\xi_0 = 4$ et 7 dB.

5.2.2 Prise en compte du rsb de focus

La figure 5.1 présente deux supports temps-fréquence obtenus sur un bruit blanc et stationnaire selon le principe de détection (5.1). Les points détectés correspondent ainsi uniquement à des fausses alarmes. Pour le support de gauche, le rapport signal sur bruit de focus ξ_0 choisi pour déterminer le seuil de détection est de 4 dB. Un rsb de focus $\xi_0 = 7$ dB a été choisi pour le support de droite. Les PFA équivalentes à ces rsb de focus sont respectivement $0,19$ et $0,095$. Cette figure laisse apparaître que les fausses alarmes ont effectivement tendance à se répartir de façon plutôt aléatoire et uniforme sur l'ensemble du spectrogramme. Cependant, elle indique également qu'en fonction du rsb de focus pour lequel le seuil de détection est ajusté, les fausses alarmes peuvent former des regroupements de taille plus ou moins importante. Plus le rsb de focus est faible, plus ces regroupements peuvent atteindre des tailles importantes. Afin d'éliminer ces composantes 'bruit seul', le critère de taille minimale doit donc tenir compte du rsb de focus.

N'ayant pu obtenir analytiquement une estimation de la taille maximale que peut atteindre une composante constituée uniquement de fausses alarmes, nous avons réalisé une série de simulations du type de celle présentée sur la figure 5.1 sur des bruits blancs et stationnaires. Pour différentes longueurs du signal et différents paramètres du spectrogramme, nous avons ainsi pu évaluer la taille minimale critique, notée T_{min} , permettant de rejeter un taux prédéfini de fausses alarmes, dénommé 'taux de rejet de fausse alarme' et noté $\tau_{\text{r,fa}}$. Un taux de rejet de 100% correspond à l'élimination de toutes les composantes 'fausses alarmes'.

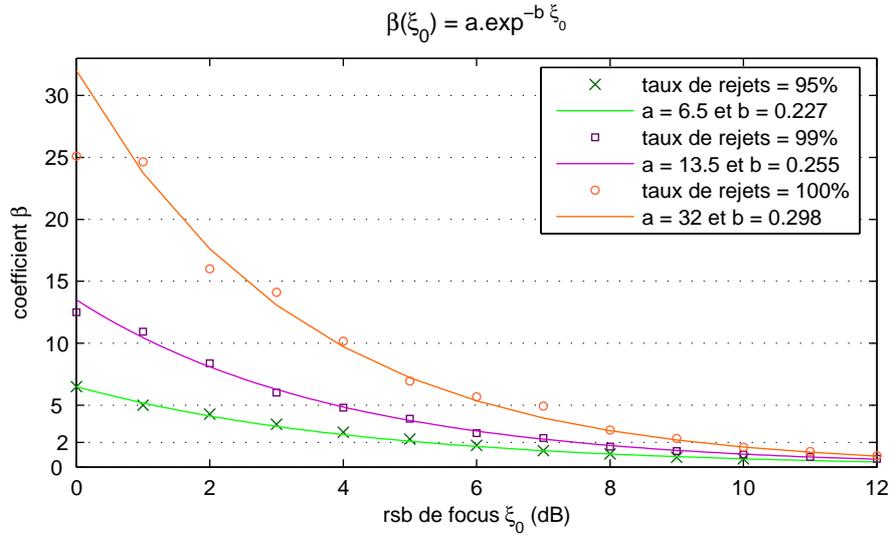


FIG. 5.2 – Coefficient multiplicateur β en fonction du rsb de focus ξ_0 , pour des taux de rejet de fausse alarme $\tau_{r,fa}$ de 95%, 99% et 100%. Les points marquent les moyennes obtenues sur 9 configurations différentes (taille de la fenêtre, recouvrement et *zero-padding*), les courbes donnent l'estimation continue obtenue avec le modèle exponentiel $\beta(\xi_0) = a \exp^{-b \xi_0}$.

$\tau_{r,fa}$	a	b
95%	6,5	0,227
99%	13,5	0,255
100%	32	0,298

TAB. 5.2 – Coefficients a et b du modèle exponentiel $\beta(\xi_0) = a \exp^{-b \xi_0}$, obtenus pour les taux de rejet de fausse alarme $\tau_{r,fa}$ de 95%, 99% et 100%.

Ces simulations ont d'une part permis de mettre en évidence que la taille critique T_{\min} est proportionnelle à la taille du noyau reproduisant T_{noyau} définie au paragraphe précédent, soit

$$T_{\min} = \beta * T_{\text{noyau}}. \quad (5.3)$$

D'autre part, le coefficient de proportionnalité β est une fonction du rsb de focus ξ_0 et du taux de rejet $\tau_{r,fa}$ souhaité. La figure 5.2 présente les valeurs moyennes obtenues pour le coefficient β en fonction du rsb de focus et pour des taux de rejet de 95%, 99% et 100%. Pour cette figure, chaque point correspond à la moyenne obtenue sur 9 configurations (taille de la fenêtre, recouvrement et *zero-padding*) de spectrogramme pour chacune desquelles 100 réalisations de bruit ont été générées. Sur la même figure, l'approximation obtenue en utilisant le modèle exponentiel

$$\beta(\xi_0, \tau_{r,fa}) = a \exp^{-b \xi_0} \quad (5.4)$$

est également présentée. le tableau 5.2 regroupe les valeurs des paramètres a et b utilisée pour les trois taux de rejet de fausse alarme.

Finalement, en fonction du rsb de focus choisi pour la détection et du taux de rejet de fausse alarme désiré, la taille T_{\min} est obtenue selon

$$T_{\min}(\xi_0, \tau_{r,fa}) = \beta(\xi_0, \tau_{r,fa}) * T_{\text{noyau}}, \quad (5.5)$$

en rappelant que la taille T_{noyau} du noyau reproduisant est fonction des paramètres de construction du spectrogramme. Seules les composantes candidates dont la taille est supérieure à T_{min} seront retenues dans le support temps-fréquence du signal.

5.2.3 Compromis fausses alarmes / signal

Nous avons défini une taille minimale pour une composante détectée afin d'être retenue dans le support temps-fréquence du signal. Cependant, l'application de ce critère entraîne inévitablement des 'faux rejets', c'est à dire l'élimination de composantes pourtant issues du signal mais dont la taille est inférieure à la taille minimale précédemment définie. La figure 5.3 donne un exemple de ce type d'erreur : la composante centrale du signal (le nez) n'étant pas assez étendue dans le plan temps-fréquence, elle ne peut être distinguée des fausses alarmes qui forment des composantes de grande taille.

Hormis la taille des composantes 'signal', ce type d'erreur est à mettre en relation avec le rapport signal sur bruit local des coefficients qui forment ces composantes. Pour des rsb local élevés, le seuil de détection peut être placé assez haut, diminuant ainsi la taille minimale critique T_{min} . La taille d'une composante signal étant dans une bien moindre mesure dépendante du niveau de seuillage, les composantes signal de petite taille mais de fort rsb pourront être retenues dans le support temps-fréquence en choisissant des rsb de focus élevés.

Le risque de 'faux rejets' des composantes signal est également lié au taux de rejet de fausse alarme que l'utilisateur souhaite obtenir. Abaisser ce taux permet de conserver des composantes signal de plus petite taille et peut donc paraître avantageux. Dans le choix de ce taux de rejet, il faut alors tenir compte de l'impact d'une fausse alarme dans le contexte où l'algorithme est utilisé. Dans le cadre d'un filtrage temps-fréquence notamment, la synthèse de quelques unes de ces composantes 'fausse alarme' a un impact très fort sur la qualité du signal reconstruit. Dans ce cadre, nous opterons pour un taux de rejet de fausse alarme de 100%, au détriment des composantes signal de petite taille qui ne seront pas conservées dans le support temps-fréquence du signal.

5.3 Algorithme de localisation

La prise en compte des fausses alarmes, abordée dans le chapitre précédent, constitue la dernière étape dans notre démarche de détermination du support temps-fréquence d'un signal. La figure 5.4 donne une vue d'ensemble sur l'algorithme complet qui résulte du travail présenté dans ce manuscrit. En parallèle, les différents paramètres mis en jeu à chaque étape de l'algorithme sont rappelés. Ces paramètres représentent les différents choix que l'utilisateur de l'algorithme doit réaliser. Ce sont les degrés de liberté de l'algorithme. Comme nous le verrons, ils représentent un éventail de possibilités qui peut être mis à profit pour les besoins de l'analyse.

Calcul du spectrogramme $S_x[n, k]$

La première étape consiste à calculer le spectrogramme du signal $x[m]$. Les paramètres en jeu sont la forme et la taille M de la fenêtre d'analyse $w[m]$, le nombre K de fréquences calculées pour chaque trame de signal, et le décalage D entre les trames successives. La fenêtre d'analyse détermine la résolution temps-fréquence du spectrogramme : un échantillon du spectrogramme résume l'information fréquentielle présente sur M échantillons du signal, et la résolution fréquentielle croît avec la taille de la fenêtre. Les deux autres paramètres déterminent la taille du spectrogramme, c'est à dire sa discrétisation temporelle (un échantillon du spectrogramme tous

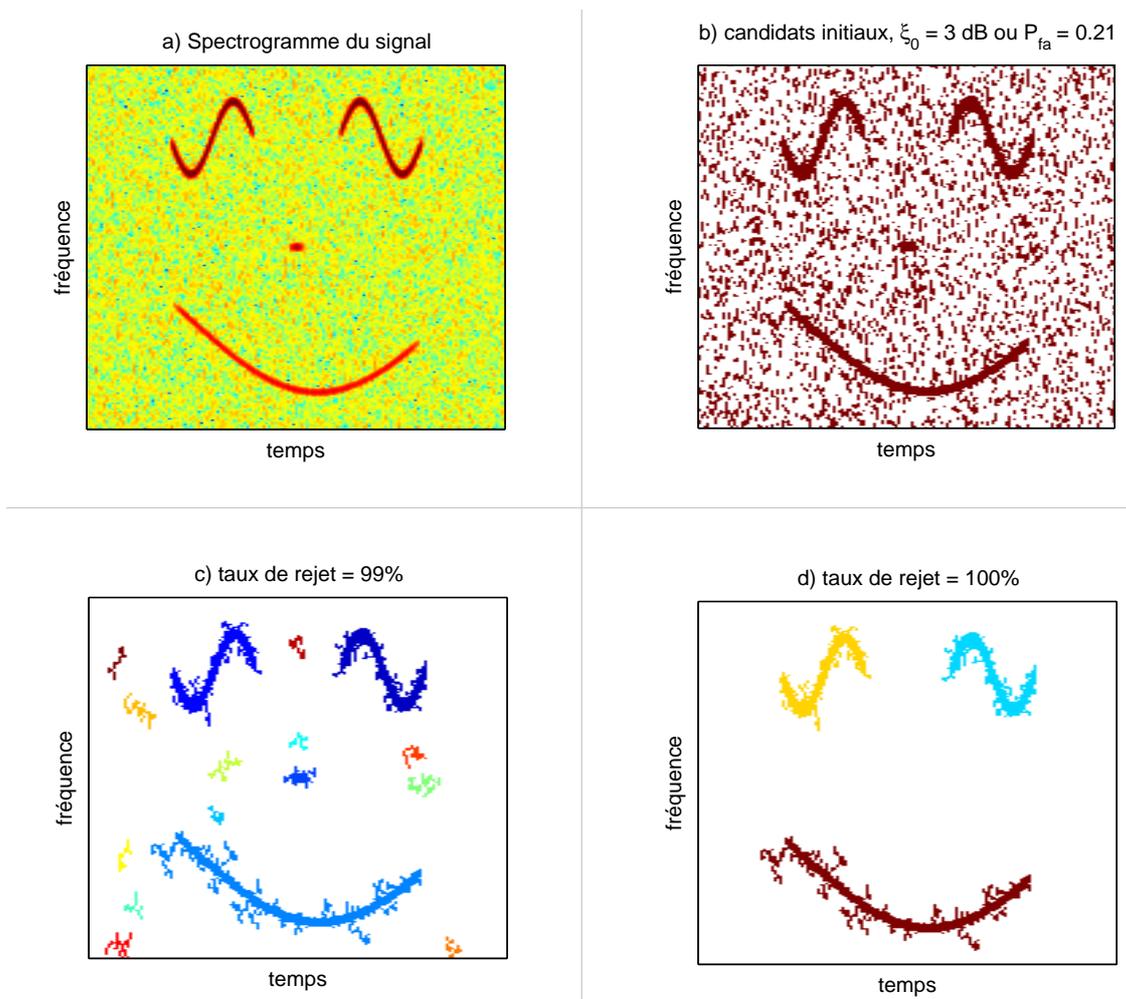


FIG. 5.3 – Illustration du compromis entre le taux de rejet des fausses alarmes et la taille des composantes 'signal'. a) spectrogramme du signal, b) ensemble des points candidats pour un rsb de focus $\xi_0 = 3$ dB, soit une P_{fa} de 0,21, c) réduction des fausses alarmes avec un taux de rejet de 99% et d) réduction des fausses alarmes avec un taux de rejet de 100%. Pour ce dernier cas, la composante signal centrale a été éliminée.

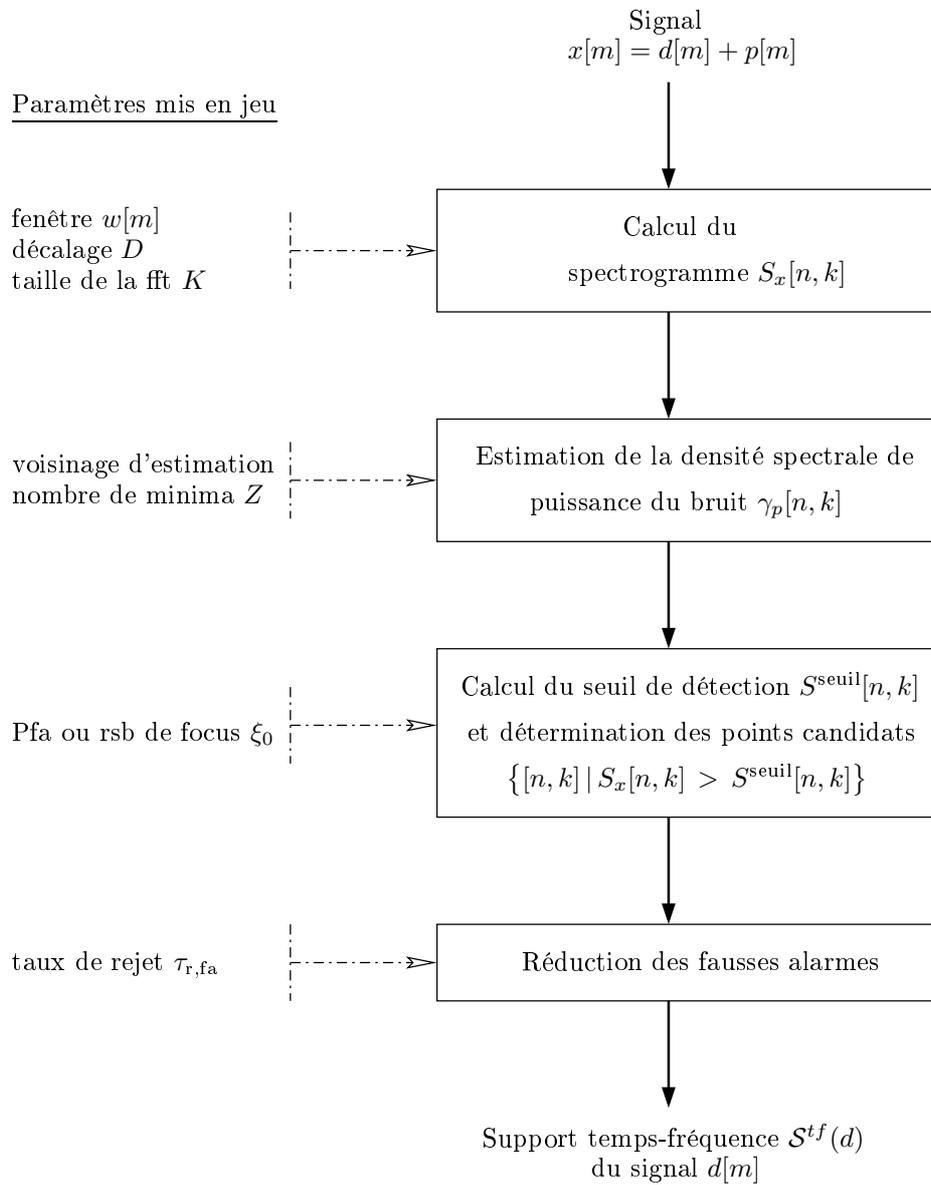


FIG. 5.4 – Schéma-bloc de l'algorithme de détermination du support temps-fréquence d'un signal et paramètres 'externes' mis en jeu à chaque étape.

les D échantillons du signal) et fréquentielle (la gamme fréquentielle $[0; f_e/2]^3$ est discrétisée en $K/2$ échantillons).

Remarque : si le calcul d'un spectrogramme ne représente aucun enjeu technique et l'impact de ses différents paramètres sont des éléments bien connus, cette première étape de l'algorithme n'en est pas pour autant dénuée d'enjeu. Une fois construit, le spectrogramme représente la seule source d'information pour toute la suite de l'algorithme. Ce qui n'apparaît pas sur ce spectrogramme ne sera pas 'retrouvé' par la suite. À ce titre, le choix des paramètres de construction sont déterminants.

Estimation de la densité spectrale du bruit $\gamma_p[n, k]$

La seconde étape consiste à estimer la densité spectrale du bruit en chaque point du plan temps-fréquence. Comme nous l'avons vu, la quantité estimée ici est plus précisément l'espérance du spectrogramme $S_p[n, k]$ du bruit $p[m]$, dit encore le 'spectre physique' (non stationnaire) du bruit. Pour chaque point $[n_0, k_0]$ du plan temps-fréquence, les Z plus petites valeurs du spectrogramme $S_x[n, k]$ observées dans un voisinage de ce point sont utilisées pour l'estimation. Cette étape nécessite deux paramètres, à savoir la forme et la taille du voisinage d'estimation d'une part et le nombre Z de coefficients retenus d'autre part. N'ayant pas mis au point de technique particulière pour le choix du paramètre Z , il est dans un premier temps fixé à 1/10^{ème} de la taille du voisinage.

Rappelons que pour ce problème d'estimation, l'apport de connaissances *a priori* sur la nature du bruit bénéficie grandement à l'estimateur en suggérant le voisinage d'estimation à retenir. Dans le cas d'un bruit blanc et stationnaire, la densité spectrale étant constante sur tout le plan temps-fréquence, le spectrogramme entier peut être utilisé comme voisinage. Dans le cas d'un bruit blanc non stationnaire, la densité spectrale de puissance du bruit varie uniquement dans le temps et reste constante suivant les fréquences. À chaque instant, l'ensemble de l'axe fréquentiel peut servir de voisinage. De la même manière, l'ensemble des points temporels à une fréquence donnée définit le voisinage d'estimation dans le cas d'un bruit non blanc et stationnaire.

Détermination des points candidats

Une fois la densité spectrale de puissance du bruit connue, le seuil de détection $S^{\text{seuil}}[n, k]$ est calculé pour chaque point $[n, k]$. Ce seuil est déterminé par le choix d'une probabilité de fausse alarme ou d'un rapport signal sur bruit de focus défini au paragraphe 4.3.3). L'ensemble des positions $[n, k]$ pour lesquelles le coefficient $S_x[n, k]$ du spectrogramme est supérieur au seuil est alors retenu comme l'ensemble des points candidats.

Réduction des fausses alarmes

La dernière étape de l'algorithme consiste à réduire le nombre de fausses alarmes Parmi les points candidats. Dans un premier temps, les points candidats connexes dans le plan temps-fréquence sont regroupés en composantes. Dans ce travail, la connexité est défini sur les 8 plus proches voisins. Dans un second temps, la taille de chaque composante est comparée avec une taille minimale T_{min} afin de rejeter les petites composantes. Le choix d'un taux de rejet de fausse alarme permet de déterminer cette taille minimale.

Support temps-fréquence du signal

L'ensemble de ces quatre étapes permet finalement d'obtenir le support temps-fréquence $\mathcal{S}^{tf}(d)$ du signal $d[m]$. Celui-ci est formé d'une ou plusieurs composantes disjointes dans le plan temps-fréquence. Dans le cas d'un signal multi-composantes, chacune des composantes est labélisée par un numéro, permettant de n'en retenir qu'une partie lors de traitements ultérieurs.

Nous notons finalement que cette algorithme de localisation d'un signal dans le plan temps-fréquence suggère une définition *a posteriori* de la notion de 'composante temps-fréquence' dans

³avec f_e la fréquence d'échantillonnage du signal $x[m]$.

un spectrogramme. Elle peut se traduire de la manière suivante :

Étant donnée le spectrogramme d'un signal bruité, une composante temps-fréquence de ce signal est un ensemble de points

- connexes dans le plan temps-fréquence,
- pour chacun desquels l'énergie observée est au moins supérieure à l'énergie moyenne du bruit,
- dont le nombre est suffisamment élevé pour qu'il ne puisse s'agir uniquement de bruit.

Cette définition comporte encore deux imprécisions sur le niveau d'énergie requis ainsi que sur la taille minimale de cet ensemble. Celles-ci sont déterminées par le choix du rapport signal sur bruit de focus.

5.4 Évaluation de l'algorithme

Dans cette partie, l'algorithme complet de localisation temps-fréquence est évalué. Dans un premier temps, nous tentons d'illustrer les variations de son comportement à travers le choix des différents paramètres. Nous présentons les supports temps-fréquence obtenus pour différents choix du rapport signal sur bruit de focus d'une part et du voisinage d'estimation du bruit d'autre part. Dans un second temps, nous présentons quelques limitations de l'algorithme, c'est-à-dire les objectifs que la localisation du support temps-fréquence d'un signal, telle que nous l'avons présentée, ne peut satisfaire. Finalement, dans un troisième temps, le cas d'un signal d'intérêt gaussien est étudié. L'importance de ce cas repose sur le fait que le signal d'intérêt et la perturbation présentent alors le même comportement statistique.

5.4.1 Dépendance au rsb de focus

Le support temps-fréquence extrait par l'algorithme est lié au choix d'un rapport signal sur bruit de focus. Celui-ci détermine le niveau du seuil de détection qui sera calculé pour chaque point TF en fonction de l'estimation du bruit. Dans ce paragraphe, nous présentons deux caractéristiques des composantes temps-fréquence extraites par l'algorithme directement liées au choix de ce paramètre.

Taille et nombre de composantes

Comme nous l'avons dit, le rapport signal sur bruit de focus détermine le niveau de seuillage du spectrogramme. Plus ce paramètre est choisi élevé, plus le support temps-fréquence obtenu se limite aux composantes les plus énergétiques. De même, en fonction du relief propre à chaque composante⁴, plus le rsb de focus est grand, plus les composantes extraites sont susceptibles de se réduire à un faible nombre de coefficients.

La figure 5.5 illustre ce comportement sur deux exemples de signaux. Les figures du haut présentent les spectrogrammes tandis que sur les figures du bas sont superposés les supports temps-fréquence obtenus pour différents rsb de focus. Cette dernière représentation permet de faire apparaître explicitement les variations du support en fonction du rsb de focus, et donne une représentation temps-fréquence graduée en rsb plutôt qu'en énergie 'brute'. Le premier signal (à gauche) se compose d'une seule composante temps-fréquence. En fonction du rsb de focus choisi, son support temps-fréquence se trouve plus ou moins étalé dans le plan. Outre la variation de taille, le second signal met en évidence la disparition de certaines composantes avec l'augmentation du rsb de focus. Ici, la composante 'oeil droit' n'appartient plus au support temps-fréquence pour les rsb de focus supérieur à 40 dB.

⁴c'est-à-dire l'amplitude des coefficients temps-fréquence qui forment cette composante

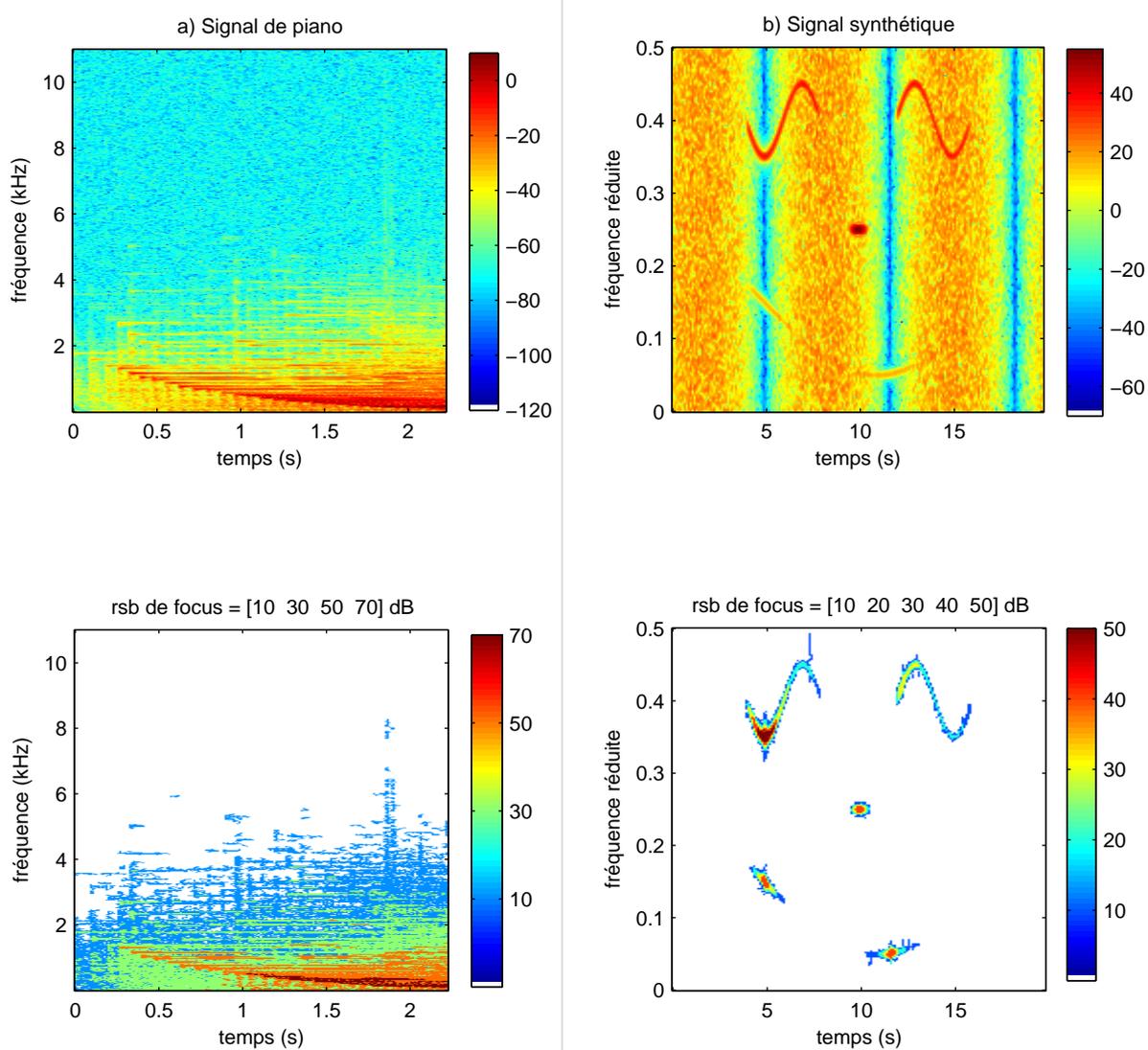


FIG. 5.5 – Spectrogramme (en haut) et supports temps-fréquence (en bas) extraits pour différents rsb de focus ξ_0 . a) signal de piano dans un bruit blanc et stationnaire, b) signal synthétique dans un bruit blanc et non stationnaire.

Présence de fausses alarmes connexes

Le critère sur la taille d'une composante présenté dans la première partie 5.2 de ce chapitre permet d'éliminer les composantes candidates constituées de l'agglomération aléatoire de plusieurs fausses alarmes. Cependant, ces fausses alarmes se répartissant aléatoirement sur l'ensemble du plan, il est inévitable que certaines d'entre elles apparaissent à des positions connexes aux composantes temps-fréquence liées au signal d'intérêt. Dans une telle situation, l'algorithme rattache ces fausses alarmes aux composantes signal et les laisse apparaître dans le support temps-fréquence du signal. Cette situation se rencontre principalement pour les faibles rsb de focus, pour lesquels la probabilité de fausse alarme est élevée et leur apparition à des points temps-fréquence connexes est fréquente.

Pour un signal synthétique perturbé par un bruit blanc et stationnaire, la figure 5.6 illustre la présence de fausses alarmes connexes dans le support temps-fréquence extrait par l'algorithme pour des rsb de focus ξ_0 de 0, 2, 5 et 10 dB. Pour $\xi_0 = 0$ et 2 dB, le support temps-fréquence se trouve fortement entaché par ces fausses alarmes et ne paraît pas acceptable. Nous notons qu'au-delà d'un rsb de focus de 10 dB, le phénomène disparaît et le support temps-fréquence extrait par l'algorithme devient uniquement représentatif du signal d'intérêt. Notons enfin que ce phénomène est également visible sur la figure 5.3 présentée auparavant, pour laquelle le rsb de focus est de 3 dB.

5.4.2 Dépendance à l'estimation du bruit

Après le rsb de focus, nous présentons dans ce paragraphe les variations du support temps-fréquence qui résultent de l'étape d'estimation de la densité spectrale de puissance du bruit. La détermination de la puissance du bruit est un élément tout à fait central de l'algorithme car il précise, en chaque point temps-fréquence, le niveau de référence à partir duquel les composantes 'signal' devront être recherchées. Le support temps-fréquence d'un signal tel qu'il est donné par l'algorithme est alors entièrement conditionné par ce niveau de référence. Dans la suite, nous présentons deux aspects de ce conditionnement. Le premier relève le risque et les erreurs liés à cette dépendance. Le second présente la flexibilité 'maîtrisée' qu'il est possible d'introduire dans l'extraction du support temps-fréquence d'un signal.

5.4.2.1 Mauvaise estimation du bruit

Lorsque la densité spectrale du bruit n'est pas connue *a priori*, il est nécessaire de l'estimer. Le support temps-fréquence obtenu est alors dépendant de la qualité de l'estimateur. Dans le chapitre 3, nous avons présenté un estimateur du bruit, qui comme nous l'avons vu présente une variance importante. Ce défaut induit une première source d'erreurs. Aussi, ce même estimateur nécessite que le bruit soit constant dans le voisinage temps-fréquence considéré. Cette condition introduit une seconde source d'erreur.

Variance de l'estimateur

L'estimateur proposé dans le chapitre 3 admet une forte variance. Étant donné un signal et son spectrogramme, il existe de nombreux points temps-fréquence pour lesquels la valeur estimée du bruit s'écarte fortement de la valeur réelle. Immanquablement, ces erreurs vont se retrouver au niveau de la détermination du support temps-fréquence.

La figure 5.7 présente le cas d'un signal synthétique noyé dans un bruit gaussien, blanc et non stationnaire. Le spectrogramme du signal est donné sur la figure 5.7-a). La figure 5.7-b) présente le résultat de l'estimation de la puissance spectrale du bruit en fonction du temps (courbe continue). La courbe pointillée donne la valeur exacte de cette dsp. Si l'estimation est globalement satisfaisante, il existe néanmoins quelques points, entourés sur la figure, pour lesquels le bruit est fortement sous estimé. Ces erreurs d'estimation se retrouvent alors directement sur le

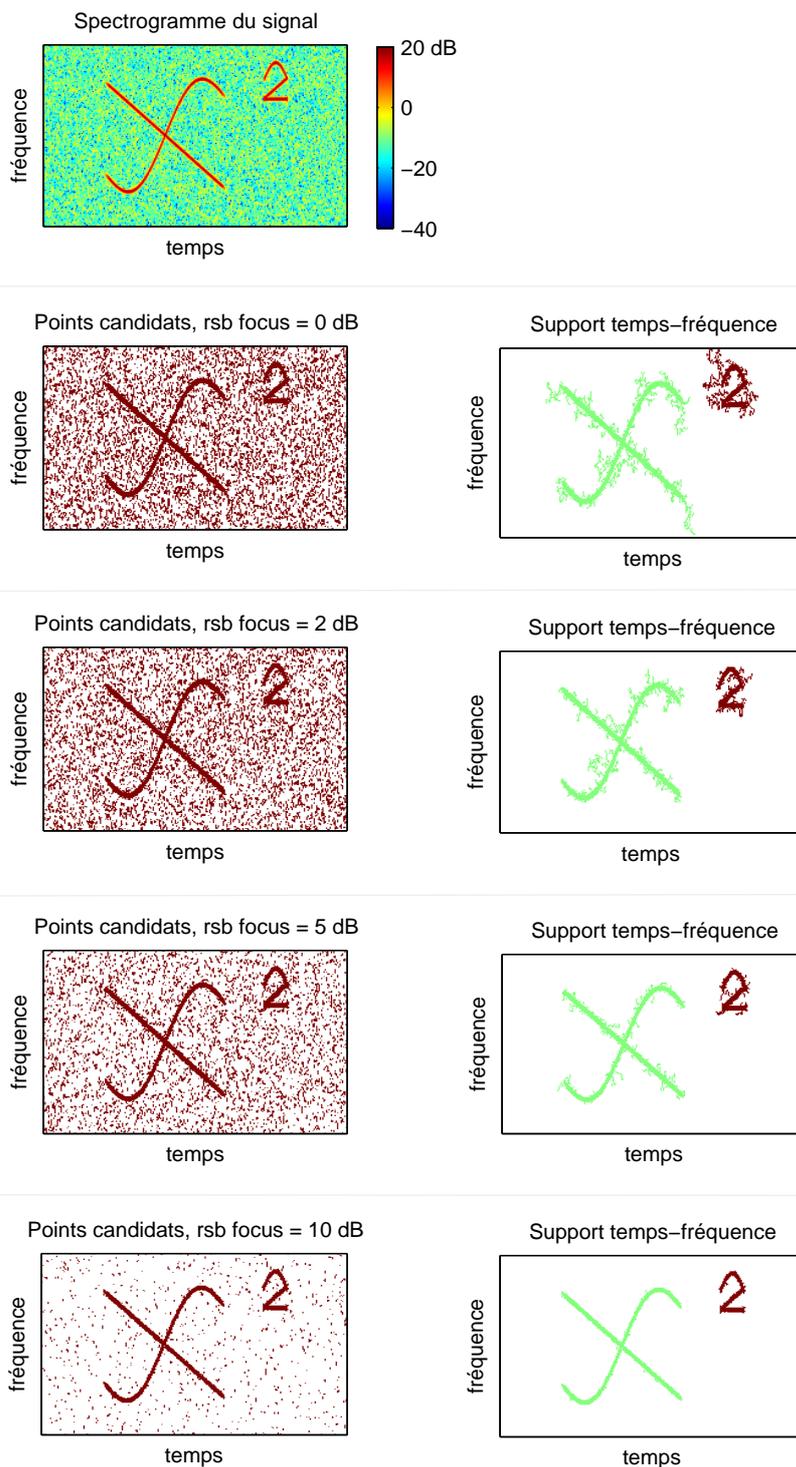


FIG. 5.6 – Illustration de la présence de fausses alarmes connexes au support temps-fréquence du signal. En haut, le spectrogramme d'un signal synthétique dans un bruit blanc et stationnaire. Colonne de gauche, les points candidats sélectionnés pour différents rsb de focus $\xi_0 = 0, 2, 5$ et 10 dB. À droite, le support temps-fréquence déterminé par l'algorithme après réduction des fausses alarmes.

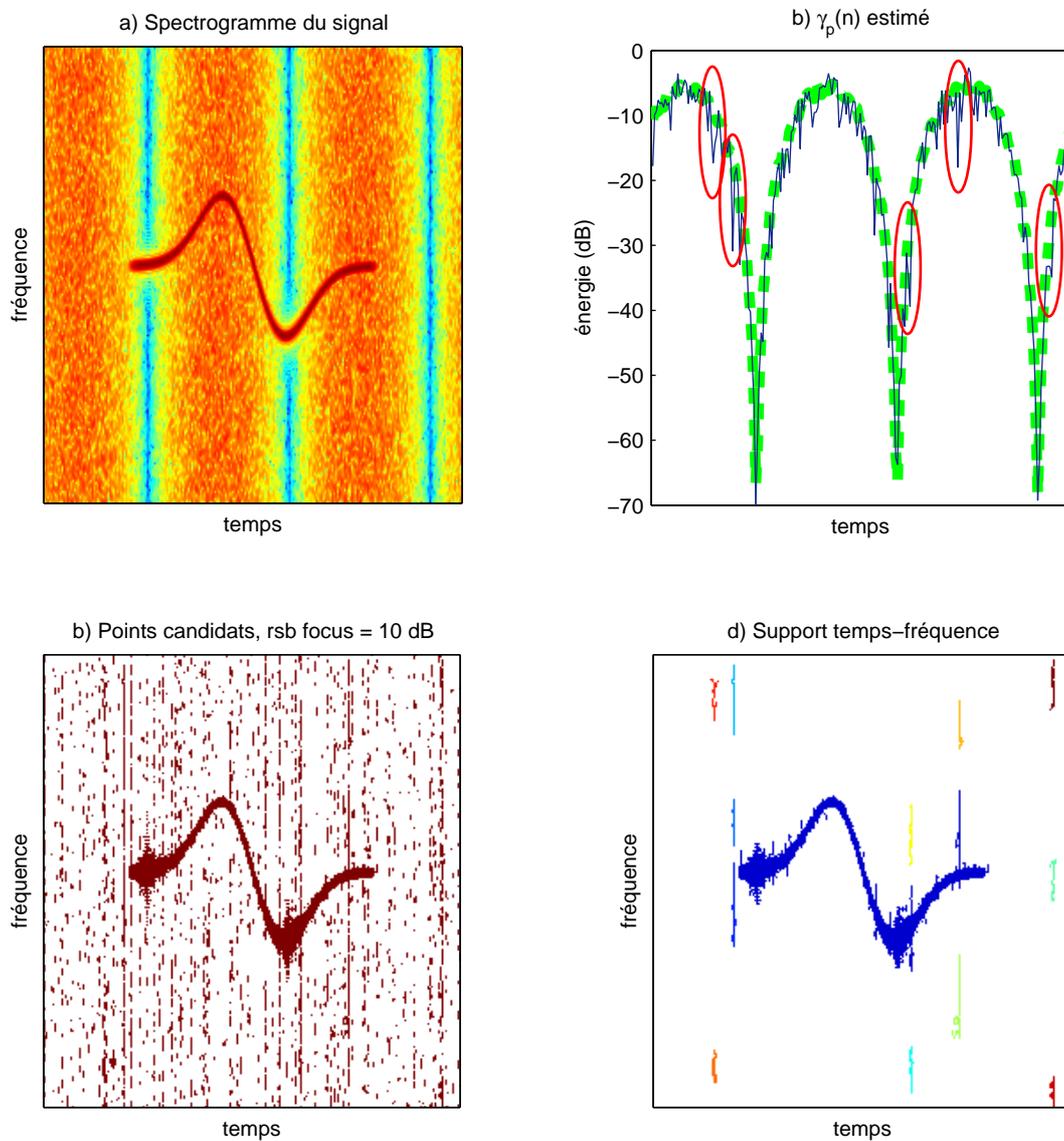


FIG. 5.7 – Impact de la variance de l'estimateur de la puissance spectrale du bruit sur le support temps-fréquence extrait par l'algorithme. Les ellipses donnent les instants temporels pour lesquels le bruit est sous estimé, entraînant l'apparition de motifs temps-fréquence indésirables dans le support obtenu.

support temps-fréquence extrait par l'algorithme, figure 5.7-d), sur lequel des composantes non désirables apparaissent, malgré un taux de rejet de fausses alarmes fixé à 100%.

Non-constance du bruit sur le voisinage d'estimation

L'estimateur du bruit présenté dans le chapitre 3 fait appel à un voisinage réparti autour du point temps-fréquence pour lequel le bruit est estimé. Afin que l'estimation soit juste, il est nécessaire que le bruit soit constant sur l'ensemble de ce voisinage. Lorsque cette condition n'est pas vérifiée, l'estimateur est biaisé. Rappelons que l'utilisation des valeurs minimales pour l'estimation du bruit conduit toujours, dans ce cas, à une sous estimation. Concernant la détection, cette tendance est très désavantageuse car elle conduit à un seuil de détection inférieur à ce qu'il aurait dû être et à une augmentation du nombre de fausses alarmes.

Pour illustration, nous reprenons le signal de la figure 5.7, en imposant en plus un caractère non blanc à la perturbation. Celle-ci devient donc non blanche et non stationnaire et son estimation passe par le choix d'un voisinage local et de taille réduite. Le spectrogramme du signal est donné sur la figure 5.8-a). Les résultats de l'estimation du bruit et de la localisation temps-fréquence sont respectivement présentés sur les figures 5.8-b) et -c). Nous observons que le support temps-fréquence obtenu contient de nombreuses composantes indésirables (précisées par les ellipses). Ces composantes correspondent aux zones temps-fréquence pour lesquelles le bruit admet de fortes variations et pour lesquelles son estimation est en conséquence fortement biaisée. Dans un tel cas de figure, la solution consiste à réduire la taille du voisinage d'estimation, au risque que le signal d'intérêt se retrouve confondu avec les variations du bruit.

5.4.2.2 Flexibilité de la discrimination bruit / signal

En chaque point du plan temps-fréquence, la puissance du bruit constitue le niveau de référence à partir duquel le signal est recherché. Toutefois, ce principe prend en compte deux entités, le 'signal' et le 'bruit', qu'il nous est laissé libre de définir. Dans bien des cas, ce que nous pourrions appeler le 'signal' se compose de plusieurs éléments, issus pourquoi pas de sources différentes. Visualisé dans le plan temps-fréquence, un tel signal peut présenter de multiples composantes, parmi lesquelles certaines seulement sont de nature informative. Dans une telle situation, il paraît intéressant de ne pas faire apparaître toutes les composantes du signal dans le support temps-fréquence. Ce support étant défini en référence au niveau de bruit, la modification possible de ce qui est jugé comme 'bruit' introduit de la flexibilité quant au support temps-fréquence extrait par l'algorithme.

En pratique, le bruit est estimé avec la méthode présentée dans le chapitre 3. Comme rappelé sur le schéma 5.4, les degrés de liberté concernant cette estimation sont le voisinage autour du point considéré et le nombre Z de coefficients appartenant à ce voisinage et retenus pour l'estimation. Parmi ces paramètres, le choix du voisinage d'estimation définit la discrimination entre bruit et signal à travers deux aspects :

- le voisinage précise la région temps-fréquence sur laquelle le bruit est jugé constant,
- le voisinage précise la taille maximale pour une composante 'signal'.

À travers ces deux aspects, une certaine flexibilité peut être introduite dans la discrimination entre bruit et signal. Par exemple, une composante du signal stationnaire pourra être éliminée en donnant au bruit ce même caractère stationnaire. Aussi, une composante signal très étalée dans le plan temps-fréquence pourra être éliminée en choisissant un voisinage d'estimation plus petit que cette composante.

Nous voyons donc que la discrimination entre le bruit et le signal se base sur un critère de taille de composante d'une part et d'énergie constante d'autre part. Nous formons ainsi deux classes, avec d'un côté les éléments de petite taille et d'énergie fluctuante, et de l'autre les éléments

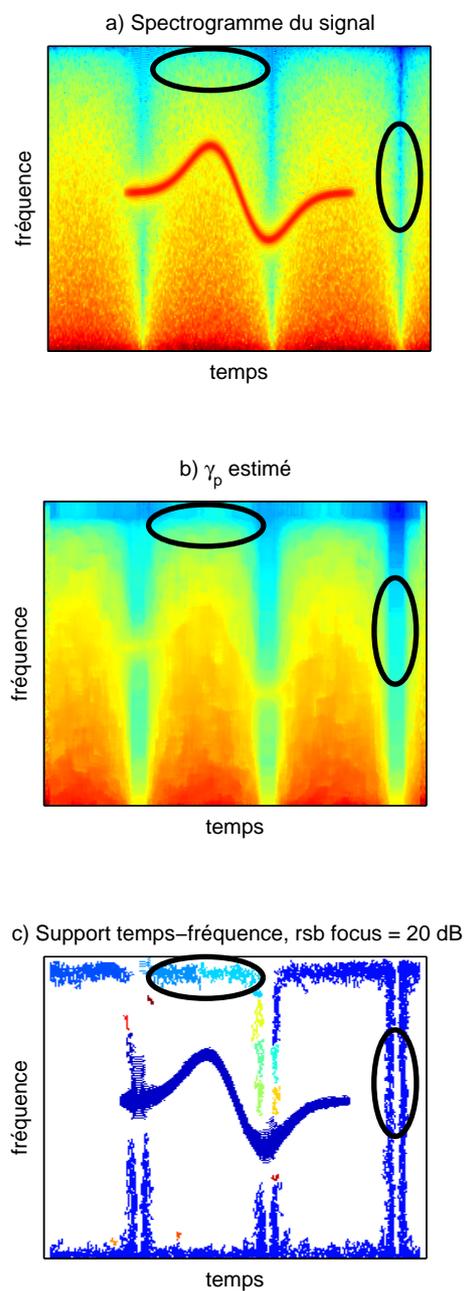


FIG. 5.8 – Impact d'un bruit non constant sur le voisinage d'estimation. Les ellipses donnent quelques exemples de zones temps-fréquence sur lesquelles le bruit est sous estimé. Ceci entraîne l'apparition de motifs temps-fréquence indésirables.

de taille plus importante et d'énergie moins fluctuante. Maintenant, si dans une dénomination générale la première classe est nommée 'signal' et la seconde 'bruit', il est possible que le caractère informatif soit porté par les grandes composantes rangées dans la classe bruit. Dans ce cas, plutôt que dans le support temps-fréquence extrait par l'algorithme, c'est dans le résultat de l'estimation du bruit que sera présente l'information recherchée.

Nous illustrons ces propos à travers l'exemple présenté sur la figure 5.9. Le signal observé correspond au niveau d'aimantation global d'un matériau, en réponse à un champ magnétique d'excitation de forme sinusoïdale. Ce signal, dit "bruit Barkhausen", permet de donner plusieurs caractéristiques sur l'état du matériau [PMH08]. La figure 5.9-a) montre que le spectrogramme du signal mesuré par le capteur présente de multiples composantes de diverses formes. La partie informative permettant de caractériser les matériaux se trouve quant-à-elle dans les composantes étalées selon les fréquences. La troisième de ces composantes est mise en évidence sur la figure 5.9-a).

La figure 5.9-b) présente le support temps-fréquence obtenu sous l'hypothèse d'un bruit blanc et stationnaire. Dans ce cas, un seuil unique (correspondant ici à un rsb de focus de 30 dB) est appliqué à tout le spectrogramme. Ce premier support met en évidence la présence d'un ensemble de composantes 'parasites'.

Ces composantes parasites étant plutôt localisées en fréquence et de nature stationnaire, les figures 5.9-c) et -d) présentent les résultats obtenus sous l'hypothèse d'un bruit non blanc et stationnaire. La figure 5.9-c) donne la puissance spectrale $\gamma_p[k]$ estimée pour ce bruit. Ce spectre de référence englobe alors tous les éléments stationnaires contenus dans le signal. La figure 5.9-d) présente le support temps-fréquence qui émerge de cette référence pour un rsb de focus de 20 dB. Sous ces hypothèses, une grande partie des composantes parasites ont pu être assimilés au bruit et éliminés du support temps-fréquence du signal. Celui-ci contient alors la majorité des composantes informatives, les éléments parasites ayant été éliminés. Toutefois, nous observons que l'algorithme ne peut rassembler dans une même composante un motif entier.

Afin de préserver une unité sur l'information extraite, l'hypothèse d'un bruit non blanc et non stationnaire a été réalisée. Les motifs d'intérêt étant plutôt étalés dans le plan temps-fréquence et de nature peu fluctuante, il semble possible de les isoler en les assimilant au bruit de référence. Le choix d'un voisinage d'estimation du bruit plus petit que ces motifs d'une part, et plus grand que les motifs parasites d'autre part, revient alors à estimer le spectre moyen des motifs d'intérêt en rejetant les motifs parasites qui en émergent. Les résultats sont présentés sur les figures 5.9-e) et -f). Pour ce troisième cas, le bruit correspond aux motifs d'intérêt étalés et peu fluctuants, tandis que les composantes parasites, plus petites et plus fluctuantes, constituent le support temps-fréquence qui en émerge.

5.4.3 Limitations

Après avoir présenté quelques caractéristiques relatives au choix des paramètres de l'algorithme, nous donnons ici deux limitations, ou plutôt deux exemples d'objectifs que l'algorithme proposé ne peut atteindre dans son état actuel. La première limitation concerne la distinction entre deux composantes qui se croisent ou se recouvrent dans le plan temps-fréquence. La seconde limitation concerne les fortes variations d'amplitude au sein d'une même composante temps-fréquence.

Composantes non disjointes dans le plan temps-fréquence

Le but de l'algorithme présenté dans ce travail est de localiser, dans le plan temps-fréquence, l'énergie d'un signal. Lorsque ce signal admet plusieurs composantes disjointes dans le plan temps-fréquence, celles-ci sont isolées par l'algorithme et peuvent être reconstruites séparément dans le domaine temporel. En revanche, lorsque deux composantes ne sont pas disjointes, c'est à

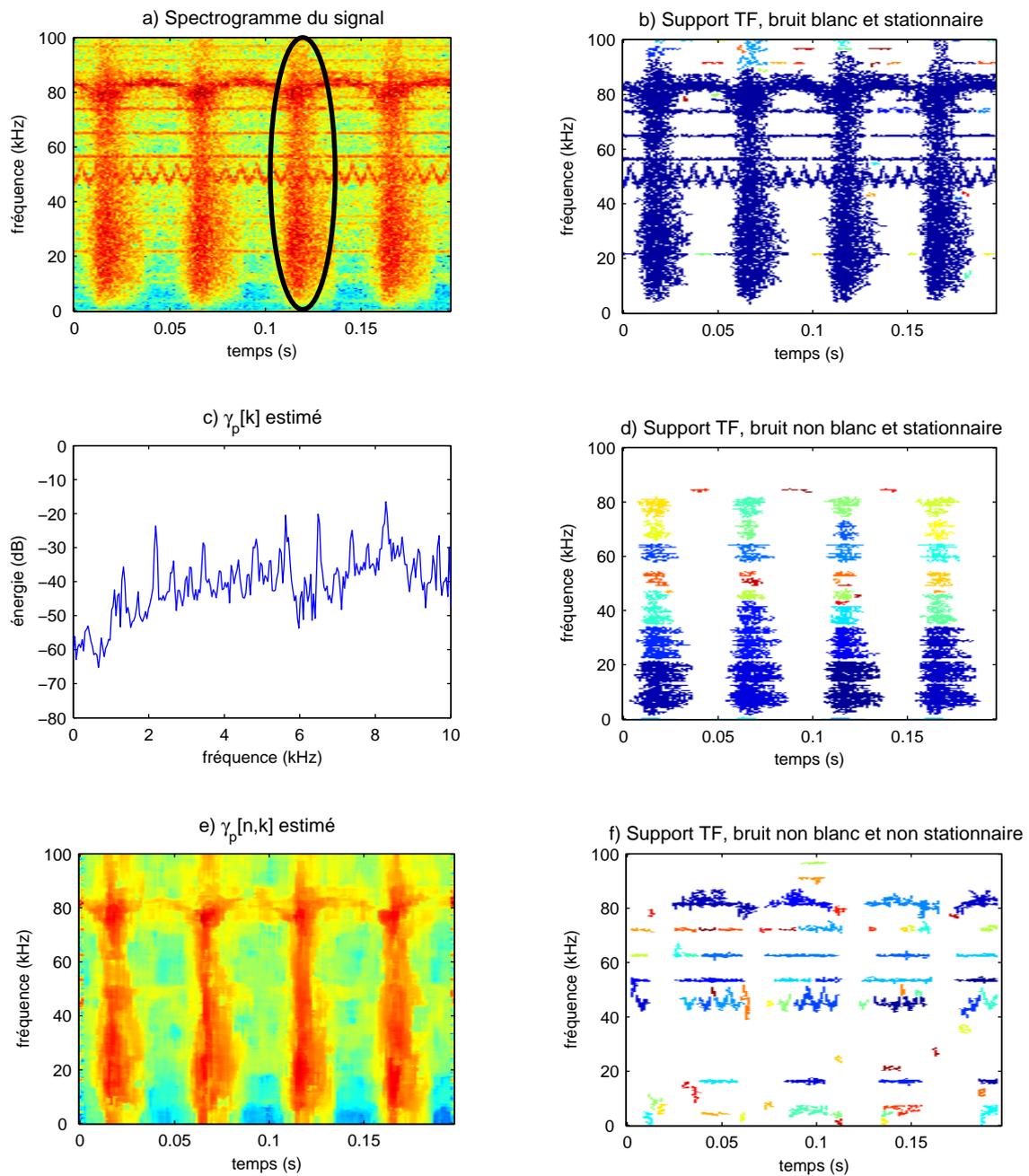


FIG. 5.9 – Extraction de l'information utile sur un signal "Barkhausen" en jouant sur le voisinage choisi pour estimer la puissance spectrale du 'bruit'. a) spectrogramme du signal, b) support temps-fréquence obtenu sous l'hypothèse d'un bruit blanc et stationnaire, c) et d) estimation de la densité spectrale de puissance et support temps-fréquence obtenus sous l'hypothèse d'un bruit non blanc et stationnaire. e) et f) estimation de la densité spectrale de puissance et support temps-fréquence obtenus sous l'hypothèse d'un bruit non blanc et non stationnaire.

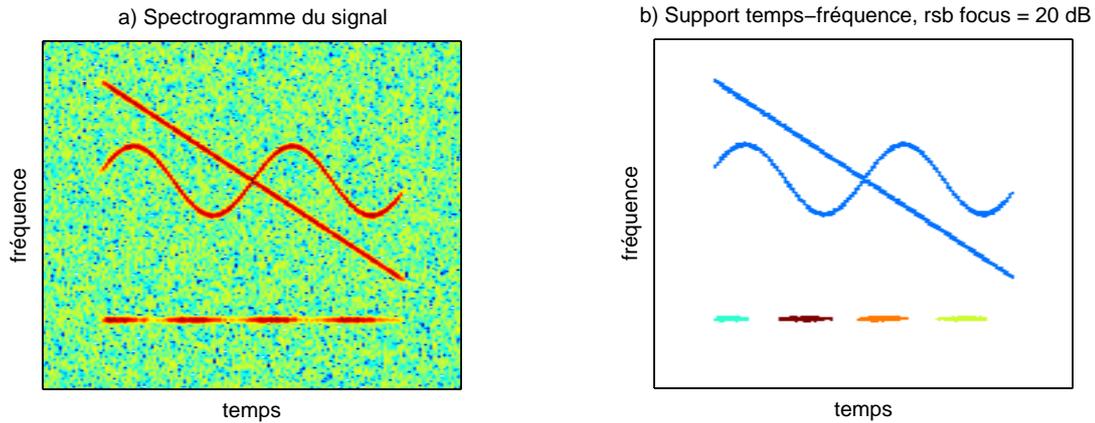


FIG. 5.10 – Limitations de l'algorithme : deux composantes non disjointes ne peuvent être séparées, une composante modulée en amplitude est scindée en plusieurs petites composantes. a) Spectrogramme du signal, b) support temps-fréquence obtenu pour un rsb de focus de 20 dB.

dire lorsque leurs supports temps-fréquence respectifs se croisent ou se recouvrent partiellement, l'algorithme proposé ne peut en aucun cas les distinguer. Dans ce cas, une seule composante temps-fréquence est reconnue et son support temps-fréquence correspond à l'union des deux supports individuels. La figure 5.10 présente un signal synthétique pour lequel il paraît intuitif de vouloir séparer les deux composantes qui se croisent, ce que l'algorithme ne pourra pas faire.

Composantes modulées en amplitude

Une seconde limitation de l'algorithme apparaît lorsque les coefficients temps-fréquence d'une même composante présentent des niveaux très différents. Pour un rsb de focus donné, il est probable qu'une partie seulement des coefficients soit détectée, scindant ainsi la composante d'origine en plusieurs petites composantes disjointes. Le signal présenté sur la figure 5.10 comporte une fréquence pure modulée en amplitude qui se retrouve découpée en plusieurs composantes par l'algorithme.

Ces deux limitations sont liées au caractère local du fonctionnement de l'algorithme. La décision de retenir ou non un coefficient temps-fréquence dans l'ensemble des points candidats est prise sur la base unique de ce coefficient. L'algorithme est donc tout à fait aveugle à ce qui se passe autour du point temps-fréquence considéré et plus généralement à l'ensemble du plan temps-fréquence. Or, séparer deux composantes qui se croisent dans le plan temps-fréquence, ou regrouper dans une même unité plusieurs composantes disjointes nécessitent une interprétation au moins 'régionale' de la représentation temps-fréquence. Dans son état actuel, l'algorithme permet simplement de détecter des zones énergétiques et d'éliminer celles qui correspondent à du bruit. Ces questions d'interprétation apparaissent plutôt comme des perspectives de ce travail.

5.4.4 Cas d'un signal d'intérêt gaussien

Nous abordons dans ce paragraphe le cas particulier d'un signal d'intérêt $d[m]$ gaussien. Dans ce cas le comportement statistique du signal d'intérêt est similaire à celui de la perturbation et il paraît intéressant de savoir si l'algorithme présenté dans ce travail peut s'adapter à cette nouvelle

configuration.

Nous considérons toujours l'observation $x[m]$ comme la somme de deux signaux indépendants, soit

$$x[m] = d[m] + p[m], \quad (5.6)$$

avec $p[m]$ une perturbation non stationnaire gaussienne, centrée et de fonction de corrélation $\Gamma_p[m, \tau]$, soit

$$p[m] \sim \mathcal{N}(0, \Gamma_p[m, \tau]). \quad (5.7)$$

Jusqu'alors jugé déterministe, nous supposons maintenant que le signal d'intérêt $d[m]$ est un processus gaussien. À partir de ce point nous distinguons le cas général pour lequel le processus est non centré, puis le cas particulier d'un processus centré.

5.4.4.1 Cas général

Dans le cas général, le signal d'intérêt $d[m]$ est un processus gaussien non stationnaire, de moyenne $\nu_d[m]$ et de fonction de corrélation $\Gamma_d[m, \tau]$, soit

$$d[m] \sim \mathcal{N}(\nu_d[m], \Gamma_d[m, \tau]). \quad (5.8)$$

Le signal observé $x[m] = d[m] + p[m]$ correspond alors à un processus gaussien, non stationnaire aux ordres 1 et 2, centré sur l'espérance $\nu_d[m]$ du signal $d[m]$, et de fonction d'autocorrélation $\Gamma_x[m, \tau] = \Gamma_d[m, \tau] + \Gamma_p[m, \tau]$ ⁵, soit

$$x[m] \sim \mathcal{N}(\nu_d[m], \Gamma_d[m, \tau] + \Gamma_p[m, \tau]). \quad (5.9)$$

Par comparaison avec le modèle (1.4) initialement choisi dans ce travail, l'introduction de l'hypothèse gaussienne pour le signal d'intérêt $d[m]$ modifie le moment d'ordre 1 du modèle puisqu'il ne s'agit plus du signal $d[m]$ lui-même mais de son espérance, et ajoute le terme $\Gamma_d[m, \tau]$ dans le moment d'ordre 2. Néanmoins, nous sommes toujours en présence d'une observation gaussienne et décentrée. En conséquence, l'algorithme développé s'adapte à cette nouvelle configuration et s'interprète de la manière suivante.

Soit $p_{eq}[m]$ le bruit gaussien centré équivalent à la somme de la perturbation $p[m]$ et des variations centrées $d[m] - \nu_d[m]$ du signal $d[m]$:

$$p_{eq}[m] = p[m] + d[m] - \nu_d[m], \quad (5.10)$$

tel que

$$x[m] = \nu_d[m] + p_{eq}[m]. \quad (5.11)$$

L'algorithme permet alors de déterminer le support temps-fréquence de l'espérance $\nu_d[m]$ du signal $d[m]$, accessible en référence au spectre physique $\gamma_{p_{eq}}[n, k] = \mathbb{E}\{S_{p_{eq}}[n, k]\}$ de la perturbation gaussienne $p_{eq}[m]$. Aussi, le rapport signal sur bruit local qui détermine les performances du détecteur est dans cas défini par

$$\xi[n, k] = \frac{S_{\nu_d}[n, k]}{\gamma_{p_{eq}}[n, k]}. \quad (5.12)$$

Il apparaît ainsi que l'utilisation de l'algorithme dans le cas général d'un signal $d[m]$ gaussien et non centré est directe. La figure 5.11 présente le résultat de l'algorithme pour un signal

⁵Les signaux $d[m]$ et $p[m]$ sont supposés indépendants.

$x[m]$ répondant au modèle (5.9). Pour ce signal, la perturbation $p[m]$ est un bruit gaussien, centré, blanc et stationnaire. Le signal $d[m]$ correspond à la somme d'une fréquence pure et d'une modulation de fréquence, à laquelle a été ajouté un bruit stationnaire gaussien, centré et coloré, de densité spectrale de puissance décroissante. Les deux modulations de fréquence correspondent donc à l'espérance du signal $d[m]$, dont le support temps-fréquence extrait par l'algorithme est donné sur la figure 5.11-d).

5.4.4.2 Cas centré

Dans de nombreux cas, l'hypothèse d'un signal gaussien est complétée par l'hypothèse que ce signal est centré, c'est à dire $\nu_d[m] = 0$. Nous avons alors

$$d[m] \sim \mathcal{N}(0, \Gamma_d[m, \tau]), \quad (5.13)$$

et toute l'information sur le signal $d[m]$ est portée par son moment d'ordre deux. Sous cette hypothèse, le signal observé est toujours un processus gaussien de fonction d'autocorrélation $\Gamma_x[m, \tau] = \Gamma_d[m, \tau] + \Gamma_p[m, \tau]$, mais cette fois-ci centré, soit

$$x[m] \sim \mathcal{N}(0, \Gamma_d[m, \tau] + \Gamma_p[m, \tau]). \quad (5.14)$$

L'algorithme présenté dans ce travail est dédié à extraire une information (temps-fréquence) concernant la moyenne d'un processus gaussien. Dans cette perspective, la variance du processus joue le rôle d'un obstacle perturbant la recherche d'information. Cette recherche se traduit alors par un test de décentrage sur une loi du χ^2 . Dans le cas du signal gaussien et centré (5.14), l'information n'est plus portée par la moyenne du processus, mais par les variations du moment d'ordre deux. Sous les deux hypothèses 'bruit seul' et 'signal+bruit', les coefficients du spectrogramme sont alors distribués selon une loi du χ^2 à 2 degrés de liberté et de paramètre de décentrage nul. Le test devient celui de deux lois du χ^2 centrées de coefficients de proportionnalité α différents. De ce point de vue, l'algorithme développé dans ce travail ne semble pas utilisable dans ce nouveau contexte.

Néanmoins, il est à noter que la densité de probabilité des coefficients du spectrogramme sous l'hypothèse 'bruit seul' n'est pas affectée par ce nouveau contexte d'un signal $d[m]$ gaussien et centré. Les coefficients temps-fréquence issus uniquement du bruit demeurent distribués selon une loi du χ^2 centrée à 2 degrés de liberté et de coefficient de proportionnalité $\alpha = \gamma_p/2$. Ceci implique que si le spectre physique $\gamma_p[n, k]$ du bruit est connu ou estimé, la détermination d'un seuil de détection à partir du choix d'une probabilité de fausse alarme P_{fa} reste valide en utilisant l'équation (4.26) donnée dans le chapitre 4. Par contre, la loi des coefficients sous l'hypothèse 'signal' étant modifiée, la détermination du seuil de détection suivant le critère du maximum de vraisemblance doit être adaptée et les performances de ce nouveau détecteur doivent être réévaluées. Dans ce nouveau contexte, nous concluons donc que la localisation du support temps-fréquence du signal d'intérêt $d[m]$ est possible avec l'algorithme développé dans ce travail, sous les conditions que le spectre physique $\gamma_p[n, k]$ du bruit puisse être estimé d'une part et que le seuil de détection soit déterminé à partir du critère de PFA d'autre part.

Nous avons appliqué l'algorithme de localisation sur un signal gaussien et centré composé de la somme de deux signaux gaussiens et centrés. Le premier signal, constituant le signal d'intérêt, est non blanc et non stationnaire. Il est obtenu par filtrage d'un bruit blanc et non stationnaire. Le second signal, représentant la perturbation additive, consiste en un bruit blanc et sinusoïdalement non stationnaire. Le spectrogramme du signal bruité est présenté sur la figure 5.12-a). En prenant appui sur l'hypothèse que le bruit additif est blanc, alors que le signal d'intérêt ne l'est pas, il a été possible d'estimer la densité spectrale de puissance du bruit au cours du temps (figure 5.12-b)). Pour une probabilité de fausse alarme fixée à 10^{-2} , les points candidats ainsi

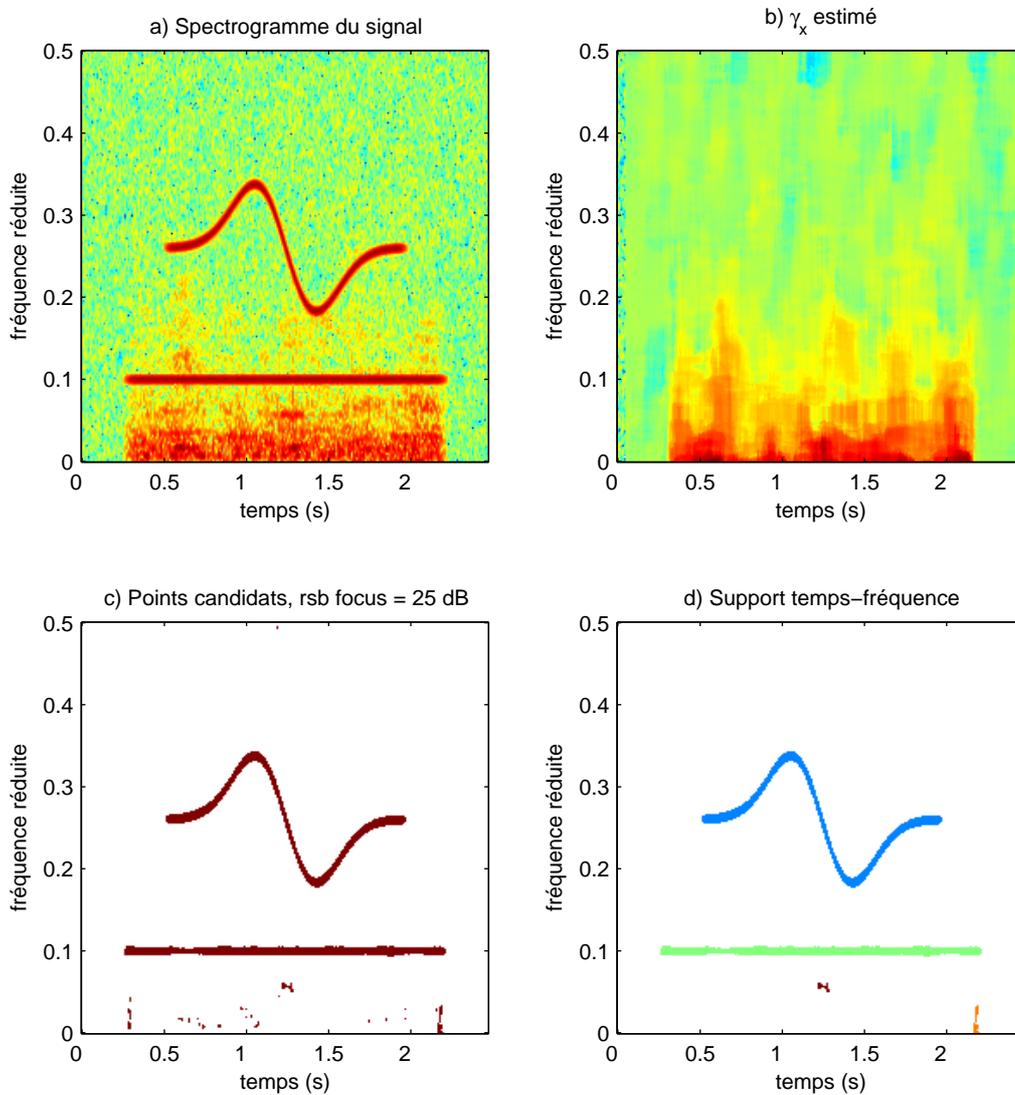


FIG. 5.11 – Support temps-fréquence de la moyenne d'un signal d'intérêt gaussien, non blanc et non stationnaire aux ordre 1 et 2, additivement bruité par une perturbation gaussienne, blanche et stationnaire. a) spectrogramme du signal, b) résultat de l'estimation de la densité spectrale du bruit global équivalent $\gamma_{peq}[n, k]$, c) Points candidats obtenu avec un rsb de focus de 25 dB et d) Support temps-fréquence retenu après réduction des fausses alarmes.

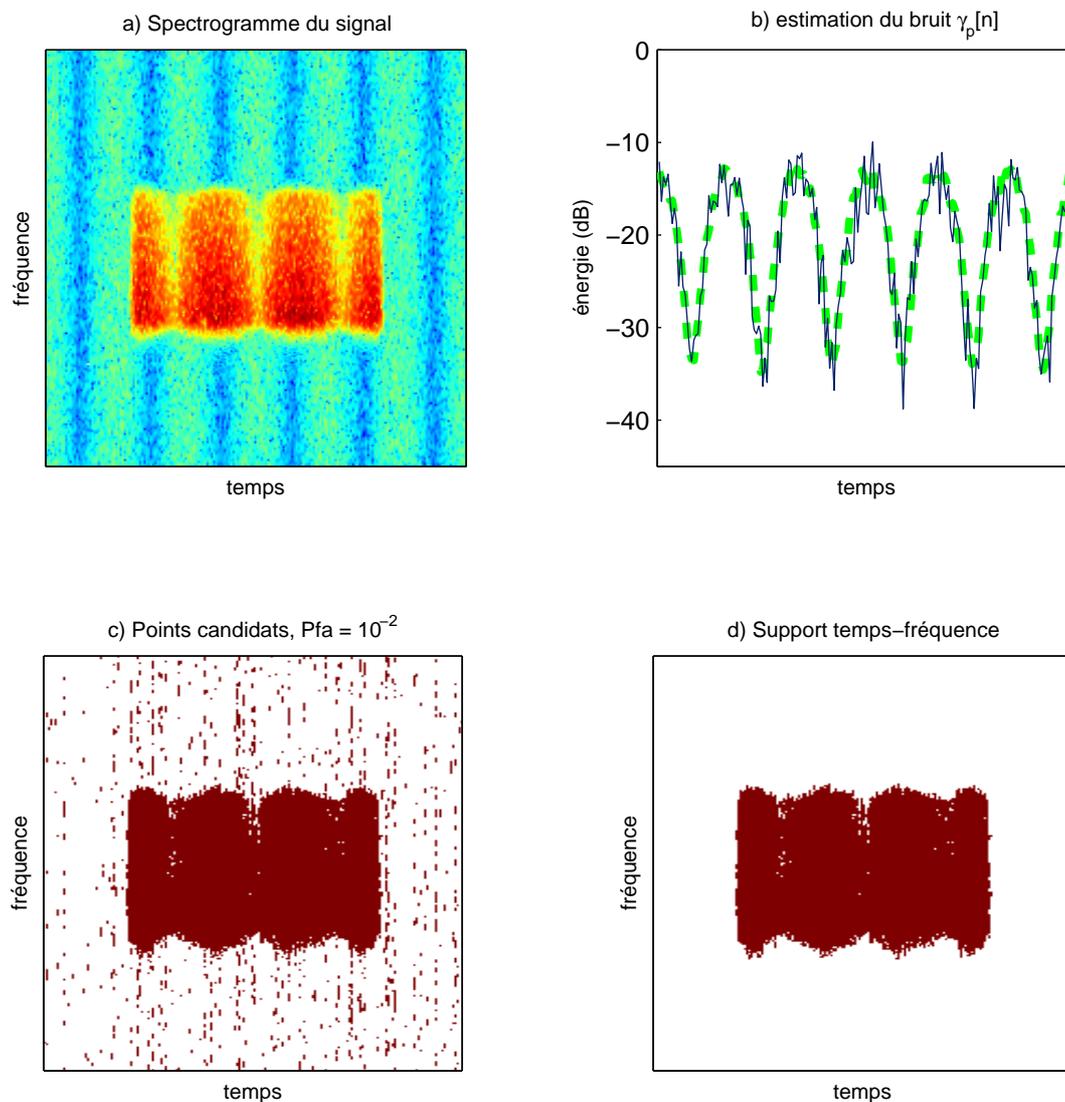


FIG. 5.12 – Support temps-fréquence d'un signal d'intérêt gaussien et centré, non blanc et non stationnaire, bruité par une perturbation gaussienne, blanche et non stationnaire. a) Spectrogramme du signal, b) résultat de l'estimation de la densité spectrale du bruit $\gamma_p[n]$, c) Points candidats pour un PFA de 10^{-2} d) Support temps-fréquence du signal.

que le support temps-fréquence du signal sont donnés sur les figures 5.12-c) et -d).

Pour finir, nous mentionnons que dans le traitement statistique des signaux de parole, les détecteurs d'activité vocale [SKS99], ou encore les algorithmes de réduction de bruit ou d'amélioration du signal [EM84][CB01] ont dans un premier temps été pour la grande majorité fondés sur ce modèle de signal gaussien et centré, noyé dans un bruit également gaussien et centré. Actuellement, ces mêmes algorithmes tendent à prendre appui sur des hypothèses différentes pour le signal d'intérêt, à savoir un modèle laplacien ou de loi gamma [CKM06] ou bien déterministe [HHJ07], ce qui renvoie au cadre initial de ce travail.

5.5 Conclusion

Un algorithme complet de localisation d'un signal dans le plan temps-fréquence a été présenté dans ce chapitre. Celui-ci s'appuie sur l'application à l'ensemble des points temps-fréquence du test d'hypothèses présenté dans le chapitre précédent. Ce processus permet de déterminer un ensemble de points candidats à devenir membre du support temps-fréquence du signal. En regroupant les points candidats connexes sous une même étiquette, cet ensemble peut être interprété comme un ensemble de composantes candidates. Parmi ces composantes, certaines correspondent à des fausses alarmes et nous nous sommes intéressés à leur possible suppression.

Dans ce but, un critère de taille minimale d'une composante a été défini, prenant en compte les paramètres de construction du spectrogramme ainsi que le niveau de seuillage déterminé par le rapport signal sur bruit de focus choisi. Si ce critère s'avère efficace, son application pour les rsb de focus inférieur à 10 dB peut entraîner l'élimination des petites composantes signal. Un compromis doit alors être réalisé entre rejet des fausses alarmes et suppression du signal. Aussi, ce critère de suppression est inefficace vis à vis des fausses alarmes qui apparaissent à des positions connexes aux composantes du signal. Là encore, ce second phénomène peut être négligé lorsque le rsb de focus est supérieur à 10 dB.

Dans la seconde partie de ce chapitre, nous avons présenté et caractérisé l'algorithme de localisation temps-fréquence qui finalise ce travail de thèse. Cet algorithme permet de déterminer le support temps-fréquence accessible d'un signal, ce support étant constitué des composantes du signal qui émergent du bruit environnant. Outre les paramètres de construction du spectrogramme, le support temps-fréquence est déterminé par deux paramètres : un rapport signal sur bruit de focus ou une probabilité de fausse alarme d'une part et le voisinage choisi pour l'estimation du bruit d'autre part.

Le rapport signal sur bruit de focus détermine le niveau de seuillage du spectrogramme. En référence au bruit environnant, il précise le degré d'émergence des composantes du signal qui sont retenues dans le support temps-fréquence. Il correspond à une discrimination énergétique entre bruit et signal retenu.

Le second paramètre, le 'voisinage d'estimation' du bruit joue le rôle le plus important en terme d'impact sur le résultat de l'algorithme. C'est effectivement sur ce paramètre que repose la détermination du bruit. Or, sans information *a priori* sur le signal d'intérêt, il apparaît naturel que son support temps-fréquence dit 'accessible' en référence au bruit environnant soit entièrement conditionnée par la façon dont ce bruit est défini. Le bruit constitue la référence à partir de laquelle le signal devient accessible. En déterminant ce qu'est le bruit, le 'voisinage d'estimation' détermine ce qu'est le signal. En complément du rsb de focus, le voisinage d'estimation introduit une discrimination de forme entre bruit et signal. Ce qui apparaît constant dans ce voisinage sera défini comme bruit tandis que les fluctuations internes au voisinage seront attribuées au signal.

Conclusions et perspectives

La détermination du support temps-fréquence d'un signal d'intérêt inconnu et additivement perturbé par un bruit gaussien a fait l'objet de cette thèse. Cet exercice nécessite la mise en place d'un processus de décision menant à décider si tel ou tel élément observé sur une représentation temps-fréquence traduit ou non la présence du signal. En l'absence d'information *a priori* sur le signal recherché, nous choisissons dans un premier temps de restreindre la prise de décision à un unique point du plan temps-fréquence et de lui donner la forme d'un test binaire d'hypothèses. Ce test est alors répété pour chaque coefficient de la représentation temps-fréquence. Dans un second temps, les résultats de ces tests sont regroupés et le processus de décision est complété au niveau du plan temps-fréquence dans son ensemble.

Le caractère inconnu du signal d'intérêt ainsi que du bruit environnant concentre la difficulté de la tâche et mène à reformuler le problème sous une forme interrogative : que peut-on connaître du support temps-fréquence d'un signal inconnu noyé dans un bruit additif gaussien ? Et dans ce but, quelles connaissances sur le signal et sur le bruit est-il nécessaire d'acquérir ? Dans le chapitre 1, nous avons mis en avant la notion de support temps-fréquence *accessible* d'un signal bruité. Celui-ci comprend les composantes temps-fréquence du signal d'intérêt qui émergent du bruit environnant. Cette définition met en avant la nécessité de connaître le bruit. Dans le cas d'un bruit inconnu, la qualité de son estimation conditionne l'utilisation possible des résultats de ce travail. De cette même définition découlent également deux éléments méthodologiques qui ont été adoptés dans ce travail. Le bruit joue d'une part le rôle de référence dans le processus de décision et le test de détection adopté consiste à décider entre une hypothèse "bruit seul" et une hypothèse "signal+bruit". D'autre part, le rapport signal sur bruit local⁶ qui représente intuitivement le degré d'émergence du signal par rapport au bruit devient le paramètre discriminant de la prise de décision.

L'estimation du bruit et la mise en place du test de détection font tous les deux appel au comportement statistique des coefficients de la représentation temps-fréquence. En raison de l'absence de termes d'interférences, nous avons retenu le spectrogramme pour représenter les signaux dans le plan temps-fréquence et le comportement statistique de ses coefficients dans le cas d'un signal gaussien a été étudié dans le chapitre 2.

Plus précisément, l'étude s'est portée sur la prise en compte des caractères non blanc et non stationnaire du bruit gaussien additif. Dans ces conditions, il a été mis en évidence que la forme de la fenêtre d'analyse utilisée dans la construction du spectrogramme influence la distribution de probabilité des coefficients temps-fréquence. Lorsque le bruit est non blanc, nous retenons que le comportement statistique du spectrogramme est fidèlement décrit par une loi du χ^2 à deux degrés de liberté lorsqu'une fenêtre d'analyse dont la valeur s'annule à ses extrémités est choisie. Le paramètre de décentrage de la loi correspond alors au spectrogramme du signal d'intérêt, tandis que le coefficient de proportionnalité correspond à l'espérance du spectrogramme du bruit. Cette dernière quantité correspond alors précisément à l'information concernant le bruit qu'il est

⁶La dénomination "local" renvoie en un point du plan temps-fréquence.

nécessaire d'acquérir.

Le chapitre 3 est ainsi dédié à l'estimation de l'espérance du spectrogramme du bruit, historiquement nommé "spectre physique" du bruit. Lorsque le bruit n'est ni blanc ni stationnaire, son spectre physique varie selon le temps et la fréquence. À partir d'une unique réalisation du spectrogramme du signal, son estimation possède alors un caractère insoluble. La solution adoptée dans ce travail consiste à définir un voisinage autour de chaque point temps-fréquence et à supposer que le spectre physique du bruit y est constant. Nous formons ainsi un ensemble de 'réalisations' identiquement distribuées à partir desquelles le bruit peut être estimé.

Cependant, parmi les coefficients temps-fréquence inclus dans ce voisinage d'estimation, seuls les coefficients représentatifs du bruit seul,⁷ doivent être retenus pour son estimation. La distribution de probabilité de ces coefficients "bruit seul" étant principalement concentrée dans les petites valeurs, nous proposons un estimateur basé sur les Z plus petites valeurs observées dans le voisinage d'estimation. L'évaluation de cet estimateur a cependant révélé sa grande variance et sa forte dépendance aux caractéristiques du voisinage d'estimation choisi. Notamment,

- le nombre de coefficients "signal+bruit" présents au sein du voisinage doit être connu. La variance de son estimation lorsque le rapport signal sur bruit local est inférieur à 10 dB représente une source d'erreur importante sur l'estimation du bruit.
- le caractère non biaisé de l'estimateur nécessite l'existence de coefficients "bruit seul" dans le voisinage d'une part et les caractères indépendants et identiques de leurs distributions d'autre part.

Le choix *a priori* d'un voisinage respectant ces deux dernières contraintes semble difficilement conciliable avec l'approche "sans information" privilégiée dans ce travail. Sur ce point, la connaissance du caractère blanc et/ou stationnaire du bruit permet de choisir *a priori* un voisinage d'estimation sur lequel le spectre physique du bruit est constant, ce qui améliore fortement la qualité de l'estimateur. Dans le cas d'une perturbation non blanche et non stationnaire, le choix du voisinage d'estimation représente une contrainte de régularité sur le bruit : les fluctuations d'énergie internes à ce voisinage seront attribuées au signal. Le voisinage choisi afin d'estimer le bruit représente alors un élément modulable pour la discrimination entre bruit et signal.

Dans le chapitre 4, le test de détection mis en place en chaque point du plan temps-fréquence a été présenté et évalué. Le seuil de détection correspondant au critère du maximum de vraisemblance a été formulé. Ce seuil est fonction du spectre physique du bruit d'une part et du rapport signal sur bruit local d'autre part. Comme point de repère, nous retenons qu'au-delà d'un rapport signal sur bruit local de 10 dB, les performances du détecteur sont les suivantes : $P_{fa} \leq 2,4 \cdot 10^{-2}$, $P_d \geq 0.97$ et $P_e \leq 2,7 \cdot 10^{-2}$.

Le signal étant inconnu, il en est cependant de même pour le rapport signal sur bruit local et le seuil de détection ne peut être ajusté de façon optimale en chaque point du plan temps-fréquence. le rapport signal sur bruit local traduit le degré d'émergence du signal par rapport au bruit en un point donné du plan temps-fréquence. Il représente une caractéristique très intuitive quant à la détermination d'un support temps-fréquence et nous avons choisi de le garder comme un paramètre propre à l'algorithme de localisation temps-fréquence. L'utilisateur est ainsi libre de choisir le rsb local dit "de focus" pour lequel le seuil de détection est ajusté au sens du maximum de vraisemblance. Lorsque l'intérêt est porté sur un intervalle de rapports signal sur bruit locaux, une procédure d'optimisation du choix du rsb de focus est également possible afin de minimiser la probabilité d'erreur relative à l'ensemble de l'intervalle. Le rsb de focus joue le même rôle qu'une probabilité de fausse alarme et une relation bijective existe entre ces deux paramètres. Cependant, si la PFA est dédiée spécifiquement au bruit, le rsb de focus précise une relation entre bruit

⁷Par opposition aux coefficients temps-fréquence pour lesquels la contribution du signal d'intérêt est non négligeable, soit les coefficients "signal+bruit".

et signal. Le rapport signal sur bruit de focus spécifie le degré d'émergence minimum à partir duquel le support temps-fréquence du signal est déterminé. Il précise le degré d'accessibilité de la notion de "support temps-fréquence accessible".

La mise en commun des résultats des tests d'hypothèses réalisés sur chaque coefficient du spectrogramme permet de faire évoluer le processus de décision au niveau de l'ensemble du plan temps-fréquence. Décrite dans le chapitre 5, cette évolution se traduit par l'introduction d'un critère de connexité temps-fréquence qui correspond à associer à un point du plan temps-fréquence ses 8 plus proches voisins. Cette connexité permet notamment de définir une composante temps-fréquence comme un ensemble de points connexes pour lesquels l'hypothèse "signal+bruit" a été retenue. Dans ce cadre, un critère de taille minimale de composante a été formulé afin de réduire la présence des fausses alarmes. Finalement, un algorithme complet permettant de déterminer le support temps-fréquence accessible d'un signal bruité a été décrit. Nous retenons qu'un degré d'accessibilité de 10 dB assure la détermination d'un support temps-fréquence cohérent avec le signal. Aussi, lorsque le signal observé est formé d'un ensemble de composantes dont le comportement diffère dans le plan temps-fréquence, le choix des paramètres de l'algorithme permet de discriminer certaines de ces composantes. Plus qu'une distinction entre signal et bruit, l'algorithme permet de séparer les composantes peu étalées en temps et/ou en fréquence vis-à-vis des composantes dont l'énergie peu fluctuante se répartie plus largement dans le plan temps-fréquence.

Perspectives :

La première perspective concerne le filtrage temps-fréquence. Cette question faisait partie des motivations initiales de ce travail que la durée impartie limitée renvoie aujourd'hui au rang des perspectives.

Plus précisément, le support temps-fréquence d'un signal délimite un ensemble de régions d'intérêt dans le plan temps-fréquence. Il est alors naturel de souhaiter reconstruire le signal qui correspond uniquement à ce support. Aussi, lorsque le support du signal fait apparaître plusieurs composantes disjointes dans le plan temps-fréquence, il est intéressant de reconstruire le signal qui ne correspond qu'à une seule de ces composantes.

Des algorithmes de synthèse d'un signal à partir d'une transformée de Fourier à court terme (TFCT) existent. Les méthodes dites *Filter-Bank Summation* et *OverLap and Add* sont par exemple décrites dans [AR77]. Afin de mener à bien ce filtrage, il est cependant nécessaire de prendre soigneusement en considération les paramètres de construction de la TFCT, ainsi que ceux de la transformée inverse. Aussi, le support temps-fréquence du signal est un masque binaire, équivalent à une fenêtre temps-fréquence rectangulaire. Dans une opération de synthèse, la nature de cette fenêtre engendre des modifications sur le signal reconstruit comme il est décrit dans [All77, ABR06]. Le second enjeu de ce filtrage concerne alors l'impact sur le signal synthétisé de la nature de la pondération appliquée à la TFCT avant la reconstruction. Une seconde stratégie de synthèse, dont un exemple est donné dans [HK95], consiste à utiliser le support temps-fréquence comme une fonction objectif. Dans ce cadre, une attention particulière doit être donnée à la discrétisation du plan temps-fréquence.

Une seconde perspective de ce travail concerne son extension à des représentations autres que le spectrogramme.

Une première adaptation semble immédiate au cas du scalogramme. Celui-ci est défini comme le module carré de la transformée en ondelette. Cette dernière étant une transformation linéaire du signal, le comportement statistique des coefficients du scalogramme est proche de celui étudié dans ce travail. Si la loi de probabilité reste donc la même, l'interprétation des paramètres de la

loi en fonction des paramètres de l'analyse temps-échelle doit cependant être revisitée. Une extension plus générale concerne la classe de Cohen. Pour cela, deux difficultés doivent être surmontées. La première difficulté est d'ordre mathématique et concerne la formulation du comportement statistique des coefficients. Les représentations de la classe de Cohen correspondent à des transformations quadratiques du signal. Pour un signal gaussien, la densité de probabilité recherchée correspond alors à la distribution de la somme de produits de gaussiennes⁸, ce qui complique considérablement sa formulation d'une part et son utilisation d'autre part. Cependant, la connaissance complète du comportement statistique d'un coefficient à travers sa densité de probabilité n'est pas une condition obligatoire à la mise en place d'un processus de décision. Par exemple, des critères de contrastes basés sur les deux premiers moments statistiques peuvent être utilisés. La seconde difficulté d'une extension aux éléments de la classe de Cohen concerne la prise en compte des termes d'interférence. Doivent-ils par exemple apparaître dans le support temps-fréquence ? Ils y seraient à juste titre en tant que représentant, dans l'espace mathématique défini par la représentation temps-fréquence choisie, de la présence du signal. Cependant la lecture et l'interprétation du support s'en trouveraient fortement complexifiées. Pour une interprétation plus physique, l'élimination de ces termes semble nécessaire. Du point de vue de la détection, les performances du spectrogramme en tant que détecteur d'un signal dans le plan temps-fréquence représente un élément de comparaison pour toutes autres représentations candidates.

Finalement, une troisième perspective est la construction d'une représentation temps-fréquence utilisant l'information donnée par le support temps-fréquence obtenu à partir du spectrogramme. Une première idée consiste à utiliser le support temps-fréquence obtenu sur le spectrogramme comme un masque binaire sur une autre représentation temps-fréquence. Notamment, le support temps-fréquence obtenu sur un spectrogramme ne comporte aucun termes d'interférence mais souffre des mauvaises propriétés de localisation du spectrogramme. L'utilisation de ce support comme un masque sur une représentation temps-fréquence admettant de bonnes propriétés de localisation permet alors de concilier localisation et absence de termes d'interférence. Pour être mis en oeuvre, cette méthode doit cependant faire face à la différence éventuelle entre le pavage temps-fréquence du spectrogramme et celui de la représentation visée.

Une seconde approche destinée aux représentations de la classe de Cohen concerne la définition de leur noyau. Elle correspond à la question suivante : est-il possible d'adapter le noyau d'une représentation temps-fréquence à l'information qui est donnée par le support temps-fréquence du signal obtenu avec son spectrogramme ? Dans le cas du spectrogramme, le noyau de la représentation est défini par la fenêtre d'analyse temporelle. Aussi, le support du spectrogramme étant lui-même obtenu à partir d'une fenêtre initiale, ce dernier cas renvoie naturellement aux stratégies d'analyse temps-fréquence multi-fenêtres [Tho98, BB96].

⁸Pour le spectrogramme, il s'agit de la somme de carrés de gaussiennes

Bibliographie

- [ABR06] F. Abda, D. Brie, and R. Ranta. Non-linear weighting function for non-stationary signal denoising. *In proc. of the Int. Conf. on Acoustic, Speech and Signal Processing, ICASSP'06. Toulouse, France*, May 2006.
- [All77] Jont B. Allen. Short term spectral analysis, synthesis and modification by discrete fourier transform. *IEEE Trans. on Acoustics, Speech and Signal Processing*, 25(3) :235–238, 1977.
- [Alt80] Richard A. Altes. Detection, estimation and classification with spectrograms. *J. Acoust. Soc. Am.*, 67(4) :1232–1245, April 1980.
- [AR77] Jont B. Allen and Lawrence R. Rabiner. A unified approach to short time-fourier analysis and synthesis. *Proceedings of the IEEE*, 65(11), November 1977.
- [Bal81] R. Balian. Un principe d'incertitude fort en théorie du signal ou en mécanique quantique. *Comptes Rendus de l'Académie des Sciences, Paris II*, 292(20) :1357–1362, Juin 1981.
- [BB96] Metin Bayram and Richard G. Baraniuk. Multiple window time-frequency analysis. *Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Paris, France*, pages 173–176, 1996.
- [CB01] I. Cohen and B. Berdugo. Speech enhancement for non-stationary noise environments. *Signal Processing*, 81 :2403–2418, 2001.
- [CB02] Israel Cohen and Baruch Berdugo. Noise estimation by minima controlled recursive averaging for robust speech enhancement. *IEEE Signal Processing Letters*, 9(1) :12–15, January 2002.
- [CKM06] Joon-Hyuk Chang, Nam Soo Kim, and Sanjit K. Mitra. Voice activity detection based on multiple statistical models. *IEEE Transaction on Signal Processing*, 54(6) :1965–1976, June 2006.
- [Coh89] Leon Cohen. Time-frequency distribution - a review. *Proceedings of the IEEE*, 77(7) :941–981, July 1989.
- [Coh95] L. Cohen. *Time-frequency analysis*. Prentice-Hall, 1995.
- [Coh03] Israel Cohen. Noise estimation in adverse environments : improved minima controlled recursive averaging. *IEEE Transactions on Speech and Audio Processing*, 11(5) :466–475, September 2003.
- [CT91] T. Cover and J. Thomas. *Elements of information theory*. Wiley and Sons, New York, 1991.
- [Dav00] Manuel Davy. *Noyaux optimisée pour la classification dans le plan temps-fréquence*. PhD thesis, Université de Nantes, Septembre 2000.
- [DN73] T.S. Durrani and J.M. Nightingale. Probability distributions for discrete fourier spectra. *Proc. Inst. Elec. Eng*, 120(2) :299–311, 1973.

- [EM84] Y. Ephraim and D. Malah. Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator. *IEEE Trans. on Acoustics, Speech and Signal Processing*, 32(6) :1109–1121, December 1984.
- [EM02] Nicholas Evans and John Mason. Time-frequency quantile-based noise estimation. In *European Signal and Image Processing Conference*, Toulouse, France, September 2002.
- [FJ60] W.F. Freiberger and R.H. Jones. Computation of the frequency function of a quadratic form in random normal variables. *Journal of the ACM*, 7(3) :245–250, July 1960.
- [Fla93] Patrick Flandrin. *Temps-Fréquence*. Traité des Nouvelles Technologies, série Traitement du Signal. Hermès, 1993.
- [Gur55] J. Gurland. Distribution of definite and of indefinite quadratic forms. *Annals of Mathematical Statistics*, 26 :122–127, 1955.
- [HHJ07] Richard C. Hendricks, Richard Heusdens, and Jensen Jesper. An mmse estimator for speech enhancement under a combined stochastic-deterministic speech model. *IEEE Transaction on Audio, Speech and Language processing*, 15(2) :406–415, February 2007.
- [HK94] Franz Hlawatsch and Werner Kozek. Time-frequency projection filters and time-frequency signal expansions. *IEEE Trans. on Signal Processing*, 42(12) :3321–3333, December 1994.
- [HK95] Franz Hlawatsch and Werner Kozek. Second-order time-frequency synthesis of nonstationary random processes. *IEEE Trans. on Information Theory*, 41(1) :255–267, January 1995.
- [HM02] C. Hory and N. Martin. Maximum likelihood noise estimation for spectrogram segmentation control. In *Proceedings of the ICASSP'02, Orlando, USA*, pages 1581–1584, 2002.
- [HM06a] Julien Huillery and Nadine Martin. Detection of time-frequency components of signals : Pfa and focus snr. In *proc. of the 7th Int. Conf. on mathematic in signal processing (IMA 2006), Cirencester, England.*, December 2006.
- [HM06b] Julien Huillery and Nadine Martin. Time-frequency modeling and detection of random non-stationary signals for monitoring purposes. In *Proc. of the 71th Int. Conf. of the Am. Inst. of Aeronautics and Astronautics (AIAA 2006), Newport, USA.*, May 2006.
- [HM07] Julien Huillery and Nadine Martin. Approximation de la loi de probabilité du spectrogramme, kl divergence et détection dans le plan temps-fréquence. *21e colloque sur le traitement du signal et des images (GRETSI 2007), Troyes, France.*, September 2007.
- [HMC02] Cyril Hory, Nadine Martin, and Alain Chehikian. Spectrogram segmentation by means of statistical features for non-stationary signal interpretation. *IEEE Trans. on Signal Processing*, 50(12) :2915–2925, December 2002.
- [HMKK00] Franz Hlawatsch, Gerald Matz, Heinrich Kirchauer, and Werner Kozek. Time-frequency formulation, design, and implementation of time-varying optimal filters for signal estimation. *IEEE Trans. on Signal Processing*, 48(5) :1417–1432, May 2000.
- [HMM06] Julien Huillery, Fabien Millioz, and Nadine Martin. On the probability distributions of spectrogram coefficients for correlated gaussian process. In *proceedings of the ICASSP'06, Toulouse, France*, May 2006.

- [HMM08] Julien Huillery, Fabien Millioz, and Nadine Martin. On the description of spectrogram probabilities with a chi-squared law. *IEEE Trans. Signal Processing*, 56(6) :2249–2258, June 2008.
- [Imh61] J.P. Imhof. Computing the distribution of quadratic forms in normal variables. *Biometrika*, 48(3) :419–426, 1961.
- [JB95] D.L. Jones and R.G. Baraniuk. An adaptative optimal-kernel time-frequency representation. *IEEE Trans. on Signal Processing*, 43(10) :2361–2371, October 1995.
- [JKB95] N.L. Johnson, S. Kotz, and N. Balakrishnan. *Continuous univariate distributions*, volume 2. Wiley and sons, second edition, 1995.
- [JL99] P.E. Johnson and D.G. Long. The probability density of spectral estimates based on modified periodogram averages. *IEEE Trans. on Signal Processing*, 47(5) :1255–1261, May 1999.
- [JW92] J. Jeong and J.W. Williams. Mechanism of the cross-terms in spectrograms. *IEEE Trans. on Signal Processing*, 40(10) :2608–2613, October 1992.
- [Kay98] Steven M. Kay. *Fundamentals of Statistical Signal Processing, vol.II - Detection Theory*. Prentice Hall, 1998.
- [Koo74] L. K. Koopman. *The spectral analysis of time series*. Academic Press, 1974.
- [Mar70] W. D. Mark. Spectral analysis of the convolution and filtering of non-stationary stochastic processes. *Journal of sound and vibration*, 11(1) :19–63, 1970.
- [Mar01] Rainer Martin. Noise power spectral density estimation based on optimal smoothing and minimum statistics. *IEEE Trans. on Speech and Audio Processing*, 9(5) :504–512, July 2001.
- [Mar06] Rainer Martin. Bias compensation methods for minimum statistics noise power spectral density estimation. *Signal Processing*, 86 :1215–1229, 2006.
- [MD04] N. Martin and C. Doncarli. *Décision dans le plan temps-fréquence*. Traité IC2, série Traitement du signal et des images. Hermes, 2004.
- [MF85] W. Martin and P. Flandrin. Wigner-ville spectral analysis of nonstationary processes. *IEEE Trans. on ASSP*, 33(6) :1461–1470, December 1985.
- [MHM06] Fabien Millioz, Julien Huillery, and Nadine Martin. Short-time fourier transform probability distribution for time-frequency segmentation. *In proceedings of the ICASSP'06, Toulouse, France*, May 2006.
- [MK07] D. Maraun and J. Kurths. Nonstationary gaussian processes in wavelet domain : synthesis, estimation and significance testing. *Physical review E*, 75(1) :016707, 2007.
- [MM80] R. J. McAulay and M. L. Malpass. Speech enhancement using a soft-decision noise suppression filter. *IEEE Trans. on Acoustics, Speech and Signal Processing*, 28(2) :137–145, April 1980.
- [MM06] D. Mauler and R. Martin. Noise power spectral density estimation on highly correlated data. *In Proc. of the Int. Workshop on Acoustic Echo and Noise Control, Paris*, September 2006.
- [NS95] C. Nikias and M. Shao. *Signal processing with α -stable distributions and applications*. Wiley, New-York, 1995.
- [PMH08] Linilson R. Padovese, Nadine Martin, and Julien Huillery. Analysing mbn signals of different materials by time-frequency methods. *In Proc. of the 5th Int. Conf. on condition monitoring, Edinburgh, Scotland.*, July 2008.
- [RL06] Sundarrajan Rangachari and Philipos Loizou. A noise-estimation algorithm for highly non-stationary environments. *Speech Communication*, 48 :220–231, 2006.

- [RP49] H. Robbins and E. J. G. Pitman. Application of the method of mixtures to quadratic forms in normal variates. *The annals of mathematical statistics*, 20(4) :552–560, December 1949.
- [Sch91] Louis L. Scharf. *Statistical Signal Processing : detection, estimation and time series analysis*. Addison-Wesley, 1991.
- [SFB00] V. Stahl, A. Fisher, and R. Bippus. Quantile based noise estimation for spectral subtraction and wiener filtering. *In proc. of the ICASSP*, 3 :1875–1878, 2000.
- [SKS99] Jongseo Sohn, Nam Soo Kim, and Wonyong Sung. A statistical model-based voic activity detection. *IEEE signal processing letters*, 6(1), January 1999.
- [Sta02] Ljubisa Stankovic. Analysis of noise in time-frequency distributions. *IEEE signal processing letters*, 9(9) :286–289, September 2002.
- [Tho98] David J. Thomson. Multiple-window spectrum estimates for non-stationary data. *Proc. Workshop on Statistical Signal and Array*, pages 344–347, 1998.
- [TMA00] Nadège Thirion-Moreau and Pierre-Yves Arquès. *Analyses temps-fréquence linéaires et quadratiques*, volume RAB of *Traité Télécoms*, pages 1–26. Techniques de l’ingénieur, 2000.
- [Tzi87] Georgios G. Tziritas. On the distribution of positive-definite gaussian quadratic forms. *IEEE Trans. Information Theory*, 33(6) :895–906, November 1987.
- [VT68] Harry L. Van Trees. *Detection, Estimation, and Modulation Theory*, volume Part 1. Wiley and Sons, New York, 1968.
- [YR06] Chunghsin Yeh and Axel Röbel. Adaptive noise level estimation. In *Workshop on Computer Music and Audio Technology (WOCMAT’06)*, Taipei, Taiwan, March 2006.

Remerciements

Voici donc trois ans et demi de travail qui se referment avec la dernière page de ce manuscrit. Toute une époque bien sûr, des joies, des doutes, des abandons, des reprises, des voyages, des discussions, des découvertes, des rencontres, du travail ... en un mot, une thèse ! Et puis à travers cette aventure il y a les personnes avec lesquels on discute, aux pauses café, dans un troquet, dans un salon ou au cours d'une soutenance, ceux qui nous ont fait confiance, ceux qui nous ont aidé, ceux qui nous ont motivé, repêché, appris. Alors pour ces multiples raisons mais aussi quelques autres, j'ai le plaisir de remercier, dans l'ordre et le désordre,

Pierre-Olivier Amblard, Cléo Barras, Patrick Bas, Daniel Baudois, Maxime Bérard, François Cayre, Jocelyn Chanussot, Jean-Marc Chassery, Cédric Duchêne, Matthieu Fauvel, Patrick Flandrin, Matthieu Gauthier, Pierre Gentil, Cédric Gérot, Pierre Granjon, Nathalie Guyader, Sébastien Harasse, Jérémy Huart, Cornel Ioana, Geneviève Jourdain, Jean-Louis Lacoume, Anthony Larue, Nicolas Le Bihan, Grégoire Le Touzé, Jérôme Mars, Nadine Martin, Fabien Millioz, Barbara Nicolas, Caroline Paulus, Cédric Richard, Laurent Ros, Benoit Roué, Christine Servièrre, Ljubisa Stankovic, Bruno Torrèsani, Jean-Michel Vanpé, Michelle Vieira et Steeve Zozor.

Résumé

Le travail de recherche présenté dans ce mémoire est dédié à la localisation d'un signal dans le plan temps-fréquence. Plus précisément, nous proposons de déterminer le support temps-fréquence accessible d'un signal d'intérêt non stationnaire, déterministe et inconnu, noyé dans un bruit additif gaussien, centré et de fonction d'autocorrélation inconnue. Le support temps-fréquence accessible est défini comme l'ensemble des points temps-fréquence pour lesquels le signal d'intérêt admet une énergie "au moins supérieure" à celle du bruit. Dans tout ce travail, le spectrogramme est choisi pour représenter les signaux dans le plan temps-fréquence.

Nous choisissons de résoudre ce problème de localisation au moyen d'un test binaire d'hypothèses, formulé en chaque point du plan temps-fréquence. Le seuil de détection correspondant à ce test est déterminé d'après les lois de probabilité des coefficients du spectrogramme, en lien avec la puissance du bruit et enfin selon un critère de détection approprié.

La première étude rapportée dans ce manuscrit concerne le comportement statistique des coefficients du spectrogramme dans le contexte d'un bruit non blanc et non stationnaire.

La densité spectrale de puissance du bruit apparaît naturellement comme l'un des paramètres des densités de probabilité des coefficients du spectrogramme. Nous proposons une méthode d'estimation de ce bruit, basée sur le comportement statistique des plus petits coefficients du spectrogramme.

Cet ensemble de connaissances nous permet finalement de résoudre le test d'hypothèses dont la solution naturelle au sens du maximum de vraisemblance fait apparaître le rapport d'énergie entre le signal et le bruit en chaque point du plan temps-fréquence. Ce rapport signal sur bruit local permet dès lors de préciser la condition "au moins supérieure" relative au support temps-fréquence accessible du signal.

Mots-clés : recherche d'information, représentations temps-fréquence, spectrogramme, loi du χ^2 , signaux aléatoires gaussiens, détection, estimation, support temps-fréquence.

Abstract

The concern of the studies reported in this manuscript is the localisation of a signal in the time-frequency plane. More precisely, our aim is to specify the obtainable time-frequency support for a non-stationary unknown deterministic signal of interest embedded in a centred Gaussian noise with an unknown autocorrelation function. The obtainable time-frequency support is defined as a set of time-frequency locations for which the energy provided by the signal of interest is "at least greater" than the energy provided by the noise. All over the manuscript, the spectrogram is used to represent the signals in the time-frequency plane.

In this work, this problem is formulated as a binary hypothesis test conducted on every time-frequency coefficients. According to an appropriate detection criterium, the detection threshold has to be determined according to the probability density function of the spectrogram coefficients on a first hand and according to the power spectral density of the noise on a second hand.

The first study is thus dedicated to the statistical behaviour of the spectrogram coefficients. Specifically, the case of a non white and non stationary noise is focused on.

The second study relates the estimation of the noise power spectral density. Based on the statistical behaviour of the minimal time-frequency coefficients, an unbiased estimator is proposed and evaluated.

As a third study, the solution for the detection problem that obeys the maximum likelihood criterion is formulated. From this study, the local (in a time-frequency meaning) signal to noise ratio appears as a natural specification for the "at least greater" condition carried by the definition of the obtainable time-frequency support.

keywords : time-frequency information, spectrogram, χ^2 law, Gaussian random signals, detection, estimation, time-frequency support.
