

TABLE DES MATIÈRES

REMERCIEMENTS	v
NOTATIONS	vii
AVANT-PROPOS	1
1 LES SIGNAUX NON STATIONNAIRES	5
1.1 Contexte	5
1.2 La stationnarité	7
1.2.1 Les processus stochastiques	7
1.2.2 Définitions de la stationnarité	9
1.2.3 L'analyse spectrale des processus stationnaires	10
1.3 Analyse spectrale en contexte non stationnaire	16
1.3.1 Le spectre instantané	16
1.3.2 Les outils d'analyse : temps-fréquence et temps-échelle	19
1.4 Des modèles de non-stationnarités	24
1.4.1 Des signaux localement stationnaires	24
1.4.2 Des signaux localement déformés	26
1.4.3 Des signaux localement harmoniques	31
1.5 Guide de lecture	33
2 ANALYSE DE SIGNAUX LOCALEMENT DÉFORMÉS	35
2.1 Introduction	35
2.2 Modèle et objectif	36
2.3 Approximation de la transformée en ondelettes	37
2.3.1 Première approche : une équation de transport	37
2.3.2 Deuxième approche : un théorème d'approximation	38
2.4 Estimation et algorithme	41
2.4.1 Procédure d'estimation	41
2.4.2 Algorithme	45
2.4.3 Performances des estimateurs et de l'algorithme	48
2.5 Résultats numériques	51
2.5.1 L'ondelette <i>sharp</i>	51
2.5.2 Estimateurs de référence	53
2.5.3 Signal synthétique	55

2.5.4	Application à l'analyse spectrale du son produit par un dauphin	57
2.5.5	Application à l'estimation de l'effet Doppler	58
2.5.6	Application à l'analyse d'un son de vent large bande	58
2.5.7	Application à la synthèse croisée de sons	62
2.6	Adaptation de l'algorithme à d'autres modèles de signaux	62
2.7	Bilan	63
3	SÉPARATION DE SOURCES DOUBLEMENT NON STATIONNAIRE	65
3.1	Introduction	65
3.2	La séparation de sources	67
3.2.1	Position du problème	67
3.2.2	En contexte stationnaire	68
3.2.3	En contexte non stationnaire	70
3.2.4	Indices de performances	72
3.3	Modèle et objectif	75
3.3.1	Une classe de signaux non stationnaires	75
3.3.2	Mélange instantané non stationnaire	75
3.4	Approximation des transformées en ondelettes	76
3.5	Procédure d'estimation et algorithme	77
3.5.1	Réduction de dimension	77
3.5.2	Procédure d'estimation	79
3.5.3	Algorithme	82
3.6	Résultats numériques	84
3.6.1	Application à un mélange synthétique	84
3.6.2	Application à un mélange de signaux audios	87
3.7	Bilan	89
4	SYNTHÈSE DE SIGNAUX LOCALEMENT DÉFORMÉS	91
4.1	Introduction : de l'analyse à la synthèse	91
4.2	L'approche à la synthèse	92
4.2.1	Limitations de l'approche à l'analyse	92
4.2.2	La synthèse probabiliste	95
4.3	Modèle probabiliste de synthèse pour les signaux localement dilatés	97
4.3.1	Modèle de synthèse	97
4.3.2	Modèle <i>a priori</i>	99
4.4	Estimation et algorithme	99
4.4.1	Espérance <i>a posteriori</i>	100
4.4.2	JEFAS à la synthèse	103
4.4.3	Algorithme	104
4.4.4	Reconstruction du signal original	106
4.5	Résultats numériques	108
4.5.1	Application à un signal synthétique	108
4.5.2	Application à un signal semi-réel à fréquence rapide	110
4.6	Bilan	114

5	SYNTHÈSE DE SIGNAUX LOCALEMENT HARMONIQUES	117
5.1	Introduction	117
5.2	Contexte	118
5.2.1	Modèle et objectif	118
5.2.2	L'analyse des signaux localement harmoniques	120
5.3	La représentation temps-fréquence idéale approchée	124
5.4	Formulation variationnelle et optimisation	127
5.4.1	Construction de la fonctionnelle	127
5.4.2	Algorithme	129
5.5	Résultats numériques	134
5.5.1	Application à un signal à une composante	134
5.5.2	Évaluations quantitatives sur un signal à deux composantes	137
5.6	Bilan	145
6	CONCLUSION ET PERSPECTIVES	147
6.1	Conclusion	147
6.2	Perspectives	148
A	DÉMONSTRATIONS	153
A.1	Démonstrations du chapitre 2	153
A.1.1	Démonstration du théorème d'approximation 2.1	153
A.1.2	Démonstration de la proposition 2.1	156
A.1.3	Démonstration de la proposition 2.2	157
A.2	Démonstration du chapitre 3	159
A.2.1	Démonstration du théorème d'approximation 3.1	159
A.3	Démonstrations du chapitre 4	161
A.3.1	Démonstration de la proposition 4.1	161
A.3.2	Démonstration de la proposition 4.2	164
A.4	Démonstrations du chapitre 5	165
A.4.1	Démonstration du théorème 5.2	165
B	RESSOURCES NUMÉRIQUES	169
B.1	Ressources relatives aux chapitres 2, 3 et 4	169
B.2	Ressources relatives au chapitre 5	170
	BIBLIOGRAPHIE	171

REMERCIEMENTS

Ce manuscrit constitue l'aboutissement de mes trois années de thèse. Il s'agit de la partie émergée de l'iceberg, et il ne reflète pas toute la diversité des aspects du doctorat. Je souhaite ici remercier toutes celles et tous ceux qui ont compté pour moi au cours de ces belles années à Marseille.

Au niveau des institutions, je remercie l'université d'Aix-Marseille, en particulier l'I2M qui m'a accueilli pendant ces trois ans. Je remercie également l'ENS Paris-Saclay pour le financement de mon contrat doctoral. Le GdR ISIS et l'école doctorale 184 ont également participé au financement des séjours effectués lors de ma thèse.

Sur le plan personnel, je tiens d'abord à remercier profondément mes directeurs de thèse qui m'ont accompagné de manière exceptionnelle, scientifiquement et amicalement. Matthieu KOWALSKI est celui avec qui j'ai démarré dans la recherche. C'est parce que mon TER puis mon stage de master 1 avec Matthieu se sont parfaitement déroulés que j'ai choisi de continuer dans le domaine, et que j'ai pu travailler avec Bruno TORRÉSANI. Les grandes qualités scientifiques de Bruno m'ont marqué et ont assurément déteint sur la rédaction de ce manuscrit. Humainement, j'ai adoré le temps passé avec lui à Marseille, au Canada et même en Estonie.

Je remercie également l'ensemble des membres du jury de ma thèse. Patrice ABRY et Monika DÖRFLER ont rapporté ma thèse avec soin et rigueur. J'ai apprécié la qualité et la justesse de leurs commentaires. Jean-François CARDOSO et Richard KRONLAND-MARTINET ont également participé au jury. Merci pour les questions soulevées et les commentaires enthousiastes sur mes travaux faits lors de ma soutenance.

Par ailleurs, je tiens à remercier ceux qui ont participé à l'élaboration des travaux présentés dans cette thèse. Bien que nous nous soyons peu croisés, les travaux d'Harold OMER ont été essentiels à la structuration de ma thèse puisque JEFAS constitue une extension directe de sa thèse. Avec Matthieu, Hau-Tieng WU a été à l'origine de l'approche variationnelle, qui a permis la construction de l'algorithme TYCOON. Je le remercie également de m'avoir accueilli pendant un mois à Toronto au début de ma thèse, ainsi que pour mon postdoc à venir. Vincent LOSTANLEN, par ces commentaires détaillés de l'article TASLP et l'intérêt qu'il a porté à notre sujet de recherche, nous a permis de rendre nos travaux plus abordables et de les mettre en valeur. D'ailleurs, le nom de l'algorithme JEFAS lui est dû.

Les enseignements à AMU représentent un autre aspect important de ma thèse. Je suis très reconnaissant envers les enseignants qui ont fait en sorte que mes cours se

soient bien passés. En particulier, je remercie Clotilde MÉLOT qui m'a guidé lors de mes premiers enseignements, j'ai pu apprécier sa rigueur et sa passion pour l'enseignement. Je remercie aussi Thomas WILLER avec qui j'ai travaillé en dernière année.

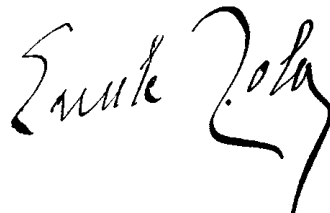
Je remercie l'ensemble des collègues et amis avec qui j'ai passé ces trois années. En particulier, je pense aux doctorants dont j'ai partagé le bureau. Par ordre d'arrivée en thèse, je remercie Hung qui fait voler les mouches, mais pas avant 11 heures du matin ! Je remercie bien sûr Marina, pour sa constante bonne humeur et sa générosité. Il s'agit certainement de la personne avec qui j'ai le plus échangé, sur nos thèses et sur le reste. J'ai aussi beaucoup aimé partager les enseignements avec elle. Les discussions avec Luc sur toutes sortes de sujets ont été très intéressantes et enrichissantes. Sa place a ensuite été reprise par Farah, qui a apporté encore plus de bonne humeur dans notre bureau. Je remercie également les autres doctorants et postdoctorants croisés durant ma thèse ou lors de PhBeer : Afef, Aurélien, Baptiste, Clément, Khuyen, Quentin, Riikka et Sokol. Je pense également aux chercheurs et enseignants-chercheurs que j'ai côtoyés dans l'équipe Signal et Image, qui méritent d'être cités notamment pour la bonne ambiance qui y règne : Caroline, Sandrine, Frédéric, Kai et Paul. En dehors de l'équipe, je voudrais mentionner Valentin pour sa participation aux Teufs et toute l'équipe QARMA pour leurs gentils mots de félicitations. Je remercie aussi les personnels administratifs de l'I2M pour leur professionnalisme.

Enfin, je remercie mes amis Clermontois, ex-Cachanais et l'ensemble de ma famille qui, bien qu'extérieurs à ma thèse, sont venus en nombre pour me soutenir le jour de ma soutenance. Leur présence durant ces trois années m'a permis de prendre régulièrement la distance nécessaire vis-à-vis de ma thèse pour reprendre mon souffle.

Pour finir, je trouve que cette citation de Zola résume bien ce que la pratique du vélo a pu m'apporter au cours de ces trois années.

Adrien MEYNARD

*J'aime la bicyclette pour
l'oubli qu'elle donne. J'ai beau
marcher, je pense. À bicyclette,
je vais dans le vent, je ne
pense plus, et rien n'est d'un
aussi délicieux repos.*



Émile Zola, *Le Cycle et la revue des sports*, numéro 239, 22 mars 1896.

NOTATIONS

Nombres réels et complexes

$\lfloor x \rfloor$	Plus grand entier inférieur ou égal à $x \in \mathbb{R}$.
$(x)^+$	Seuillage de $x \in \mathbb{R}$: $(x)^+ = \max(x, 0)$.
$\operatorname{Re}(z)$	Partie réelle $z \in \mathbb{C}$.
$\operatorname{Im}(z)$	Partie imaginaire de $z \in \mathbb{C}$.
\bar{z}	Complexe conjugué de $z \in \mathbb{C}$.

Matrices et vecteurs

$\langle \mathbf{x}, \mathbf{y} \rangle$	Produit scalaire entre deux vecteurs \mathbf{x} et \mathbf{y} .
$\ \mathbf{x}\ $	Norme de \mathbf{x} .
\mathbf{I}	Matrice identité.
$[\mathbf{M}]_{mn}$	Élément d'indices (m, n) de la matrice \mathbf{M} .
$[\mathbf{M}]_m$	m^{e} ligne de la matrice \mathbf{M} .
$[\mathbf{M}]_{\cdot n}$	n^{e} colonne de la matrice \mathbf{M} .
\mathbf{M}^T	Transposée de la matrice (ou du vecteur) \mathbf{M} .
\mathbf{M}^H	Transposée conjuguée de la matrice (ou du vecteur) \mathbf{M} .
\mathbf{M}^+	Pseudo-inverse de la matrice \mathbf{M} .
$\mathbf{M}^{\circ 2}$	Carré élément par élément de la matrice \mathbf{M} .
$\mathbf{M}_1 \circ \mathbf{M}_2$	Produit terme à terme des matrices de même taille \mathbf{M}_1 et \mathbf{M}_2 .

Signaux

δ	Distribution de Dirac.
$\mathbb{1}_{[a,b]}$	Fonction indicatrice, valant 1 dans $[a, b]$ et 0 en dehors.
$T_\tau x$	Signal x translaté en temps par τ (voir Définition 1.13).
$M_\nu x$	Signal x modulé en fréquence par ν (voir Définition 1.13).
$D_s x$	Signal x dilaté par s (voir Définition 1.13).

Opérateurs

f'	Dérivée de f .
------	------------------

$f^{(p)}$	Dérivée p -ième de f .
$\partial_t f$	Dérivée partielle de f par rapport à la variable t .
$x * y$	Produit de convolution entre les signaux x et y .

Espaces

$C^p(\mathbb{R})$	Espace des fonctions dont les p premières dérivées sont continues.
$L^p(\mathbb{R})$	Espace des fonctions dont la norme p existe (voir définition 1.8).
$H^2(\mathbb{R})$	Espace des fonctions de $L^2(\mathbb{R})$ dont le support de la transformée de Fourier est inclus dans \mathbb{R}_+ .
$\mathcal{S}(\mathbb{R})$	Espace de Schwartz, des fonctions infiniment dérivables à décroissance rapide, ainsi que leurs dérivées de tous ordres.

Représentations des signaux

\hat{x}	Transformée de Fourier du signal x (voir Définition 1.9).
V_x	Distribution de Wigner-Ville du signal x (voir Définition 1.12).
\mathcal{G}_x	Transformée de Fourier à court terme du signal x (voir Définition 1.14).
\mathcal{W}_x	Transformée en ondelettes du signal x (voir Définition 1.15).

Probabilités

$\mathbb{P}(\omega)$	Probabilité de l'événement ω .
p_X	Densité de probabilité de la variable aléatoire X .
$\mathbb{E}\{X\}$	Espérance de la variable aléatoire X .
$\mathcal{N}(\mu, \mathbf{C})$	Vecteur aléatoire gaussien réel de moyenne μ et de matrice de covariance \mathbf{C} .
$\mathcal{CN}_c(\mathbf{0}, \mathbf{C})$	Vecteur aléatoire gaussien complexe circulaire de moyenne nulle et de matrice de covariance \mathbf{C} .
$\mathcal{CN}(\mu, \mathbf{C}, \mathbf{R})$	Vecteur aléatoire gaussien complexe de moyenne μ , de matrice de covariance \mathbf{C} , et de matrice de relation \mathbf{R} .
k_X	Fonction d'autocovariance du processus stationnaire X .
\mathcal{S}_X	Densité spectrale de puissance du processus stationnaire X .

AVANT-PROPOS

Cette thèse traite de divers aspects de la non-stationnarité. Nous nous intéressons plus particulièrement à l'étude des signaux non stationnaires modélisés à partir d'un signal stationnaire sous-jacent dont la stationnarité a été brisée par une transformation dépendant d'une ou plusieurs variables fonctionnelles. La question à laquelle nous nous proposons de répondre est alors la suivante : comment étendre les outils d'analyse spectrale aux signaux non stationnaires de ce type ? Pour cela, nous présenterons diverses méthodes d'analyse et de synthèse spectrale appropriées à l'étude de plusieurs sous-classes de signaux non stationnaires pouvant être modélisés comme indiqué précédemment.

Nous verrons que l'estimation spectrale d'un signal non stationnaire appartenant à la classe considérée revient à déterminer simultanément la densité spectrale du signal stationnaire sous-jacent et l'opérateur de déformation ayant brisé la stationnarité. Nous illustrons schématiquement en figure 1 les grandes lignes des méthodes d'estimation spectrale mises en œuvre. Partant du signal non stationnaire dont le chronogramme est tracé en haut gauche sur la figure, nous lui appliquons une transformation bien choisie telle que, dans le domaine transformé, l'opération de déformation est caractérisée par une opération simple. Cette représentation est tracée en bas à gauche sur la figure 1. Dans cet exemple, la non-stationnarité est caractérisée, à chaque instant, par une translation verticale des coefficients d'ondelettes. Il s'agit alors d'estimer cette opération afin d'en déduire l'opérateur de déformation brisant la stationnarité. Ensuite, nous appliquons l'opération inverse afin d'obtenir une estimation de la représentation du signal stationnaire sous-jacent dans le domaine transformé. Il s'agit de l'étape de « stationnarisation ». Enfin, nous pouvons reconstruire ce signal stationnaire (chronogramme en haut à droite de la figure) en appliquant la transformée inverse puis estimer sa densité spectrale par une méthode d'analyse spectrale classique. Dans cet exemple, la densité spectrale estimée est tracée verticalement sur la droite de la figure 1.

Pour répondre plus précisément à la question que nous nous sommes posée, et présenter plus en détail les méthodes que nous proposons, le manuscrit a été organisé en six chapitres que nous décrivons brièvement dans ce qui suit.

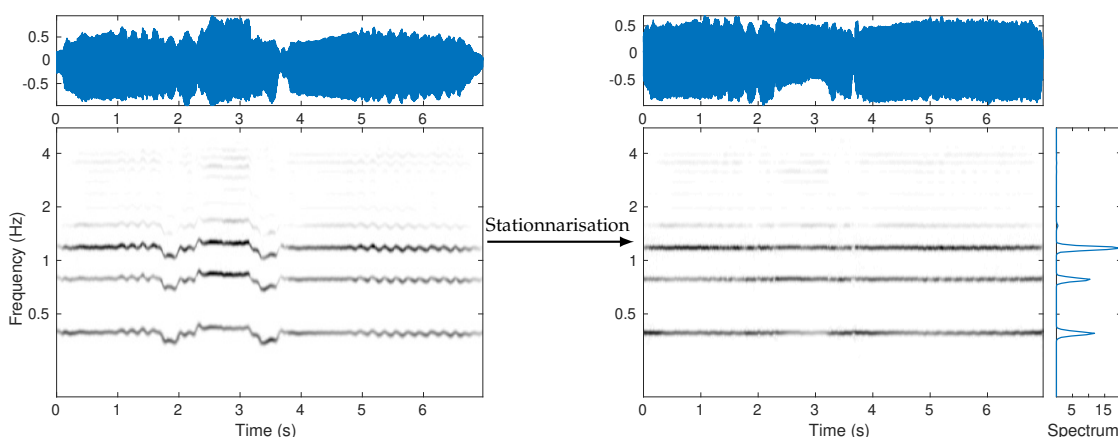


FIGURE 1 – Illustration des méthodes d’estimation développées dans la thèse. À gauche : chronogramme (en haut) et transformée en ondelette (en bas) du signal non stationnaire original. À droite : chronogramme (en haut), transformée en ondelette (en dessous) et densité spectrale de puissance estimée (à droite) du signal stationnaire sous-jacent estimé.

Chapitre 1. Dans ce chapitre introductif, nous présentons la stationnarité dans un cadre probabiliste. Ces signaux étant caractérisés par leur densité spectrale, nous passons en revue les techniques d’analyse spectrale existantes. L’extension de l’analyse spectrale aux signaux non stationnaires mène aux outils d’analyse temps-fréquence et temps-échelle que nous décrivons. Dans la fin de ce chapitre, nous nous intéressons à des modèles de non-stationnarités qui ont été proposés antérieurement. Nous décrivons plus particulièrement deux types de modèles physiquement réalistes : les signaux localement déformés et les signaux localement harmoniques.

Chapitre 2. Nous nous intéressons dans ce chapitre à un cas particulier de signaux localement déformés : les signaux localement dilatés et modulés en amplitude. Nous présentons alors un résultat d’approximation permettant de mettre en évidence l’opération brisant la stationnarité dans le plan temps-échelle. Les déformations sont caractérisées par des déplacements ou des multiplications de coefficients d’ondelettes. En supposant que les signaux stationnaires sous-jacents sont gaussiens, nous montrons que l’estimation des déformations peut être obtenue par maximum de vraisemblance. Nous présentons alors un algorithme d’estimation alternée permettant d’estimer à la fois l’opérateur de déformation et le spectre du signal stationnaire sous-jacent. Après avoir décrit les performances théoriques attendues, nous mettons cet algorithme en œuvre sur un signal synthétique puis sur des signaux audio réels variés tels que le son produit par un dauphin, ou le son produit par une voiture se déplaçant à vitesse constante et perçu par un observateur fixe.

Chapitre 3. Ce chapitre considère l’extension du modèle du chapitre précédent aux signaux multi-capteurs. Nous supposons maintenant que nous disposons d’un certain nombre de signaux localement dilatés qui sont amplifiés et mélangés à l’aide d’une fonction de modulation d’amplitude instantanée matricielle. Ce problème ainsi posé consiste en une séparation de source doublement non stationnaire. Il s’agit de retrouver chacune des sources non stationnaire à partir de l’observation de leurs mélanges. Après avoir présenté différentes méthodes de séparation des sources existantes, nous proposons une nouvelle méthode s’appuyant sur une approximation dans le plan temps-échelle similaire à celle du chapitre 2. L’algorithme que nous proposons est alors évalué et comparé aux autres algorithmes sur des mélanges synthétiques de signaux synthétiques et réels.

Chapitre 4. Comme le chapitre 2, ce chapitre porte sur l’étude des signaux localement dilatés. Cependant, l’approche proposée est assez différente car il s’agit d’une approche à la synthèse. Nous exposons d’abord les limites imposées par l’utilisation d’une approche à l’analyse. Nous proposons alors un modèle probabiliste de synthèse des signaux localement dilatés à partir d’une représentation temps-échelle. En choisissant bien *l’a priori* sur celle-ci, nous sommes capables de construire une représentation temps-échelle beaucoup plus concentrée que la transformée en ondelettes classique. En s’inspirant de l’algorithme du chapitre 2, nous construisons un algorithme d’estimation joint de cette représentation temps-échelle, de la dilatation instantanée et du spectre du signal stationnaire sous-jacent. Finalement, nous montrons sur un exemple synthétique que, bien que le temps de calcul soit fortement rallongé, cette approche à la synthèse permet d’étudier et caractériser une plus grande variété de signaux localement dilatés que l’approche à l’analyse du chapitre 2.

Chapitre 5. Comme le chapitre précédent, ce chapitre s’appuie sur une approche à la synthèse. En revanche, nous nous portons ici sur l’étude des signaux localement harmoniques. L’objectif est la recherche d’une représentation temps-fréquence idéale de tels signaux. Après avoir décrit les méthodes existantes permettant d’approcher cette représentation, nous répertorions les propriétés désirées de celle-ci (reconstruction du signal, régularité, concentration temps-fréquence). Nous construisons une fonctionnelle convexe dont cette représentation est le minimiseur. Nous proposons alors un algorithme d’optimisation convexe permet d’approcher ce minimiseur. Cet algorithme est ensuite mis en œuvre sur une exemple semi-réel. La qualité de la représentation temps-fréquence obtenue est comparée à celles données par d’autres méthodes.

Chapitre 6. Ce dernier chapitre est consacré aux conclusions et perspectives de cette thèse.

Remarque. Par souci de faciliter la lecture, les démonstrations des différents résultats théoriques établis dans cette thèse sont regroupées dans l’annexe [A](#).

Publications

Les travaux réalisés durant cette thèse se sont concrétisés par les publications mentionnées ci-dessous.

Chapitre 2. Les travaux présentés dans ce chapitre ont donné lieu à deux publications. Le premier article [\[45\]](#) a été présenté à la conférence SampTA 2017. Dans ce papier, nous avons détaillé le problème d’analyse des signaux localement dilatés et modulés en fréquence et en amplitude brièvement repris dans la section 2.6. Les limites de l’étude de ce modèle (présentées dans la section 6.2 ci-dessus) nous ont conduit à étudier le modèle plus simple des signaux localement dilatés et modulés en amplitude. C’est ce que nous avons présenté dans [\[46\]](#), qui a été publié dans la revue *IEEE/ACM Transactions on Audio, Speech and Language Processing*. Le chapitre 2 est en grande partie inspiré de cet article.

Chapitre 3. Les travaux concernant la séparation de sources ont été présentés lors du workshop iTWIST 2018 [\[44\]](#) et de la conférence francophone GRETSI 2019 [\[43\]](#). Une publication plus détaillée dans une revue est prévue.

Chapitre 4. Ce chapitre n’a pas encore donné lieu à une publication. Il s’agit des travaux les plus récents de cette thèse et les travaux sur ce sujet sont encore en cours.

Chapitre 5. Ce chapitre reprend principalement des résultats qui ont été publiés dans la revue *Applied and Computational Harmonic Analysis* [\[36\]](#).

Remarque. Par ailleurs, nous fournissons dans l’annexe [B](#) le lien vers les ressources numériques permettant de reproduire la plupart des figures illustrant ce manuscrit. Notamment, l’ensemble des résultats numériques présentés sont reproductibles à partir des codes MATLAB fournis en ligne.

Chapitre 1

LES SIGNAUX NON STATIONNAIRES

1.1 Contexte

Le traitement des signaux est l'un des maillons essentiels de la science expérimentale contemporaine. Par exemple, en physique la récente détection d'ondes gravitationnelles par les interféromètres LIGO et Virgo n'a pu se faire que grâce à des méthodes très évoluées de traitement de signal permettant d'extraire le signal d'intérêt à partir d'une mesure extrêmement bruitée [1]. Dans le domaine biomédical, les signaux captés par des électrocardiogrammes ou des électroencéphalogrammes demandent également de nombreuses étapes de traitement afin d'en extraire les caractéristiques utiles au médecin [65]. Les technologies modernes emploient également de nombreuses techniques issues de cette discipline : les systèmes audios pour obtenir des sons de très haute qualité ou les systèmes optiques pour le traitement d'images photographiques.

Le point commun à toutes les applications mentionnées ci-dessus concerne l'allure des signaux traités : ils ont la propriété d'être non stationnaires. Le concept de non-stationnarité est défini par opposition à celui de stationnarité. Celui-ci peut être présenté de deux points de vue différents. Premièrement, d'un point de vue mathématique, la stationnarité est bien définie comme une propriété des processus stochastiques. Les processus stationnaires sont ceux dont les propriétés statistiques sont invariantes par translation en temps. En toute rigueur, on ne peut donc pas parler de stationnarité lorsqu'on traite des signaux qui sont des grandeurs mesurées et donc déterministes. Deuxièmement, plaçons-nous d'un point de vue physique. On peut alors définir les signaux stationnaires comme étant ceux produits par des systèmes physiques stables ou ayant atteint un régime permanent. On peut songer, par exemple, au son émis par une trompette lorsqu'une note est tenue. On peut également penser au signal électrique produit par une machine électrique tournant à vitesse constante. Le lien qui est fait entre ces deux points de vue s'établit au niveau de la modélisation des signaux. En fait, on modélise systématiquement les signaux *physiquement* stationnaires comme étant des réalisations de processus *mathématiquement* stationnaires. Cette modélisation tient debout dans la mesure où, dans les deux cas, on conserve une propriété d'invariance par

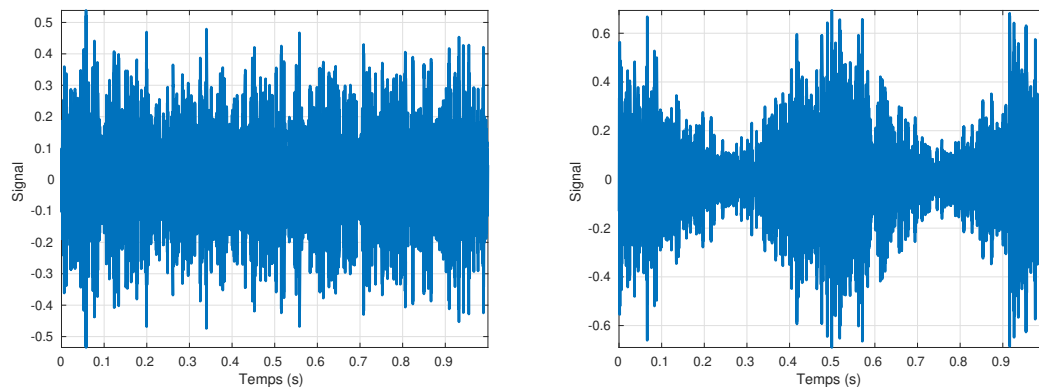


FIGURE 1.1 – Chronogrammes de signaux. À gauche : signal stationnaire. À droite : signal non stationnaire obtenu par modulation d’amplitude du signal stationnaire.

translation en temps.

Revenons aux exemples du premier paragraphe. Ce sont des signaux non stationnaires car ils sont produits par des systèmes physiques transitoires : l’onde gravitationnelle ne perdure qu’une fraction de seconde tandis que les informations neuronales ou cardiaques évoluent en permanence selon l’état du patient. De manière non rigoureuse, on pourrait définir la non-stationnarité comme le fait de pouvoir « percevoir » des différences sur le signal si l’on décale l’origine des temps d’une valeur arbitraire. La figure 1.1 permet de distinguer visuellement (directement sur les chronogrammes des signaux) la différence entre stationnarité et non-stationnarité. Pour cela, la non-stationnarité est introduite en modulant en amplitude un signal stationnaire. On « voit » alors que si nous décalons temporellement le signal de 0,3 secondes vers la gauche, le signal ainsi formé est facilement distinguable du signal non décalé. En effet, les positions des fortes amplitudes dans le signal décalé correspondent aux faibles amplitudes du signal non décalé (et vice versa). En revanche, une fois décalé d’un temps quelconque, le signal stationnaire (à gauche sur la figure) reste très similaire au signal non décalé. Cet exemple illustre un cas particulier de non-stationnarité : la modulation d’amplitude. Il s’agit d’un modèle très simple de non-stationnarité qui ne peut s’appliquer qu’à un faible nombre de signaux. On peut alors définir de nombreux autres types de déformations comme la modulation de fréquence ou la déformation temporelle (que nous définirons dans ce chapitre) qui une fois appliquées à un signal stationnaire le rendent non stationnaire. Comme il n’existe pas de modèle universel de non-stationnarité (si ce n’est par opposition à stationnarité), d’autres types de non-stationnarité peuvent être envisagés, ne s’appuyant pas nécessairement sur un processus stationnaire déformé. Nous nous intéresserons par la suite à des modèles de processus non stationnaires physiquement réalistes.

Dans ce chapitre, nous donnons des définitions plus rigoureuses des quantités que nous venons d’aborder. Dans un premier temps, nous définissons la stationnarité et nous abordons la question de l’analyse spectrale des signaux stationnaires. Ensuite,

nous décrivons une extension de l'analyse spectrale aux signaux non stationnaires, et nous présentons les représentations temps-fréquence et temps-échelle qui lui sont associées. Enfin, nous répertorions différentes classes de processus non stationnaires permettant la modélisation de nombreux signaux physiques.

1.2 La stationnarité

Comme nous l'avons vu dans la section précédente, la stationnarité au sens mathématique est une caractéristique de certains processus stochastiques. Présentons donc d'abord ceux-ci.

1.2.1 Les processus stochastiques

Dans ce manuscrit, nous manipulerons différentes grandeurs aléatoires. Les variables aléatoires seront notées en lettres capitales tandis que leurs réalisations seront notées en lettres minuscules. On notera p_X la densité de probabilité de la variable aléatoire X . Le caractère gaussien des variables et des vecteurs aléatoires sera souvent exploité dans les exemples présentés dans cette thèse. On distingue deux cas : celui où les variables aléatoires gaussiennes sont réelles (définition 1.1), et celui où elles sont complexes (définition 1.2).

Définition 1.1 (Variable aléatoire gaussienne réelle). Soit X une variable aléatoire réelle. Soit \mathbf{X} un vecteur aléatoire réel.

- La variable aléatoire X est dite gaussienne réelle si sa densité de probabilité s'écrit

$$p_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma^2}(x - \mu)^2\right),$$

où μ désigne sa moyenne et σ^2 sa variance.

- Le vecteur aléatoire \mathbf{X} est dit gaussien réel si toute combinaison linéaire de ses composantes est une variable aléatoire gaussienne réelle.

Soit m la taille du vecteur aléatoire \mathbf{X} . Notons $\boldsymbol{\mu} \in \mathbb{R}^m$ sa moyenne et $\mathbf{C} \in \mathbb{R}^{m \times m}$ sa matrice de covariance. Autrement dit :

$$\begin{aligned}\boldsymbol{\mu} &= \mathbb{E}\{\mathbf{X}\}, \\ \mathbf{C} &= \mathbb{E}\{(\mathbf{X} - \boldsymbol{\mu})(\mathbf{X} - \boldsymbol{\mu})^H\}.\end{aligned}$$

Alors, si la matrice de covariance du vecteur aléatoire gaussien réel \mathbf{X} est inversible, sa densité de probabilité s'écrit :

$$p_{\mathbf{X}}(\mathbf{x}) \triangleq \mathcal{N}(\boldsymbol{\mu}, \mathbf{C}) = \frac{1}{(2\pi)^{\frac{m}{2}} |\det(\mathbf{C})|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \mathbf{C}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right). \quad (1.1)$$

Lorsque qu'une variable aléatoire est à valeurs complexes, la circularité désigne la propriété d'invariance par rotation dans le plan complexe des caractéristiques statistiques de celle-ci [59]. En d'autres termes, cela signifie que les variables X et $Xe^{i\varphi}$ (où $\varphi \in \mathbb{R}$) sont régies par les mêmes lois de probabilité indépendamment de φ . Il découle de cette propriété que les variables aléatoires complexes circulaires sont nécessairement de moyenne nulle. De manière équivalente, on peut définir les variables aléatoires complexes circulaires comme celles vérifiant

$$\begin{cases} \mathbb{E}\{X\} = 0 ; \\ \mathbb{E}\{X^2\} = 0 . \end{cases}$$

Définition 1.2 (Variable aléatoire gaussienne complexe circulaire). Soit X une variable aléatoire complexe. Soit \mathbf{X} un vecteur aléatoire complexe.

- La variable aléatoire X est dite gaussienne complexe circulaire si sa densité de probabilité s'écrit

$$p_X(x) = \frac{1}{\sigma^2\pi} \exp\left(-\frac{1}{\sigma^2}|x|^2\right) ,$$

où σ^2 désigne sa variance.

- Le vecteur aléatoire \mathbf{X} est dit gaussien complexe circulaire si toute combinaison linéaire de ses composantes est une variable aléatoire gaussienne complexe circulaire.

Soit m la taille du vecteur aléatoire \mathbf{X} . Alors, si la matrice de covariance $\mathbf{C} \in \mathbb{C}^{m \times m}$ du vecteur aléatoire gaussien complexe circulaire \mathbf{X} est inversible, sa densité de probabilité s'écrit :

$$p_{\mathbf{X}}(\mathbf{x}) \triangleq \mathcal{CN}_c(\mathbf{0}, \mathbf{C}) = \frac{1}{\pi^m |\det(\mathbf{C})|} \exp\left(-\mathbf{x}^H \mathbf{C}^{-1} \mathbf{x}\right) . \quad (1.2)$$

Afin d'étendre la notion d'aléatoire aux fonctions à temps continu, on définit les processus stochastiques. Ces derniers décrivent l'évolution de variables aléatoires au cours du temps [67].

Définition 1.3 (Processus stochastique). Un processus stochastique X est une application $t \mapsto X(t)$ telle qu'à chaque instant $t \in \mathbb{R}$, $X(t)$ est une variable aléatoire à valeurs dans \mathbb{C} .

Par la suite, les signaux étudiés seront modélisés comme des réalisations de processus stochastiques. Autrement dit, si $x : t \mapsto x(t)$ est un signal mesuré, chaque valeur observée $x(t)$ sera vue comme une réalisation de la variable aléatoire $X(t)$. Les processus stochastiques gaussiens (réels ou complexes) seront utilisés dans les prochains chapitres. On les définit comme suit.

Définition 1.4 (Processus gaussien réel). Soit X un processus stochastique réel. Il est dit gaussien si pour toute partie finie $I = \{t_1, \dots, t_m\}$ d'instants sur \mathbb{R} , le vecteur aléatoire $\mathbf{X}_I = (X(t_1), \dots, X(t_m))$ est gaussien réel.

Définition 1.5 (Processus gaussien complexe circulaire). Soit X un processus stochastique complexe. Il est dit gaussien circulaire si pour toute partie finie $I = \{t_1, \dots, t_m\}$ d'instants sur \mathbb{R} , le vecteur aléatoire $\mathbf{X}_I = (X(t_1), \dots, X(t_m))$ est gaussien complexe circulaire.

1.2.2 Définitions de la stationnarité

Intéressons-nous maintenant au cadre théorique permettant de définir la stationnarité. Pour cela, nous nous plaçons dans le cadre des processus stochastiques définis précédemment.

La stationnarité, définie ci-dessous, permet de décrire une classe de processus stochastiques dont les propriétés sont présentées par la suite.

Définition 1.6 (Stationnarité au sens strict). Soit X un processus stochastique. X est dit stationnaire au sens strict si pour toute partie finie $\{t_1, \dots, t_m\}$ d'instants sur \mathbb{R} , et pour tout $u \in \mathbb{R}$, les vecteurs aléatoires $(X(t_1), \dots, X(t_m))$ et $(X(t_1 + u), \dots, X(t_m + u))$ ont la même loi de probabilité.

Ainsi, toutes les caractéristiques statistiques de X sont invariantes par translation en temps, ou de manière équivalente, par changement d'origine des temps. Par exemple, l'espérance, la variance et tous les moments d'ordres supérieurs de X sont indépendants du temps.

Dans la suite, nous nous placerons dans un cadre moins restrictif : celui de la stationnarité au sens large où l'on se contente de vérifier les propriétés de stationnarité des statistiques d'ordre un et deux du processus : sa moyenne et sa covariance.

Définition 1.7 (Stationnarité au sens large). Soit X un processus stochastique. X est dit stationnaire au sens large, ou en moyenne quadratique, si ses moments d'ordre un et deux existent et sont invariants par tout changement d'origine des temps. Autrement dit, les deux conditions suivantes doivent être vérifiées par X :

$$\begin{cases} \mathbb{E}\{X(t)\} = \mathbb{E}\{X(0)\} \in \mathbb{C}, & \forall t \in \mathbb{R} \\ \mathbb{E}\{X(t)\overline{X(\tau)}\} = \mathbb{E}\{X(t-\tau)\overline{X(0)}\} \in \mathbb{C}, & \forall t, \tau \in \mathbb{R}. \end{cases}$$

Dans ce cas, on note $k_X : \mathbb{R} \rightarrow \mathbb{C}$ la fonction d'autocovariance du processus X définie par

$$k_X(\tau) = \mathbb{E} \left\{ (X(\tau) - \mathbb{E}\{X(0)\}) (\overline{X(0) - \mathbb{E}\{X(0)\}}) \right\}.$$

Dans la suite, on appellera signal stationnaire, toute réalisation d'un processus stationnaire.

La stationnarité au sens large est en partie motivée par le fait que, pour les processus gaussiens (réels ou complexes circulaires), elle implique la stationnarité au sens strict. En effet, ces derniers sont entièrement déterminés par deux quantités : leurs moyenne et leurs fonctions d'autocovariance (voir les formules (1.1) et (1.2)).

Les définitions que nous présentons dans ce chapitre peuvent être étendues au cas plus général des processus stochastiques généralisés. Un processus stochastique généralisé associe une variable aléatoire à toute fonction de l'espace de Schwartz $\mathcal{S}(\mathbb{R})$ (c'est l'ensemble des fonctions infiniment dérivables à décroissance rapide, ainsi que leurs dérivées de tous ordres). Il s'agit de l'extension de la théorie des distributions dans un cadre probabiliste. Nous renvoyons le lecteur à [51] pour une présentation détaillée de la théorie des processus généralisés. Dans ce cas, la fonction d'autocovariance devient elle-même une distribution. Par conséquent, les matrices de covariance de vecteurs aléatoires issus de tels processus ne sont plus nécessairement définies. Dans ce qui suit, la notation $\langle X, \psi \rangle$ désigne la variable aléatoire issue du processus stochastique généralisé X appliqué à la fonction $\psi \in \mathcal{S}(\mathbb{R})$. Le bruit blanc, défini ci-dessous, constitue un exemple de processus stochastique généralisé stationnaire, dont la fonction d'autocovariance prend une forme particulière.

Exemple 1.1 (Bruit blanc réel). Soit X un processus généralisé réel stationnaire au sens large de fonction d'autocovariance k_X . On dit que X est un bruit blanc lorsqu'il vérifie les deux conditions suivantes :

$$\begin{cases} \mathbb{E} \{ \langle X, \psi \rangle \} = 0, & \forall \psi \in \mathcal{S}(\mathbb{R}), \\ k_X = \sigma_X^2 \delta, \end{cases}$$

où δ désigne la distribution de Dirac.

Cette forme particulière pour la fonction d'autocovariance d'un bruit blanc s'interprète comme le fait que toutes les valeurs de X sont temporellement décorréélées les unes des autres (deux à deux). De plus, comme c'est un processus stationnaire, sa variance est constante au cours du temps (et vaut σ_X^2).

1.2.3 L'analyse spectrale des processus stationnaires

1.2.3.1 La densité spectrale de puissance

Dans le cas des processus stationnaires, la densité spectrale de puissance, qui sera précisée dans la définition 1.10, est un outil suffisant pour la détermination du contenu

fréquentiel de ces processus. Il s'agit du pendant stochastique de la notion de transformée de Fourier définie pour les signaux déterministes. Intéressons-nous donc d'abord à la transformée de Fourier. Son utilisation se limite, dans un premier temps, aux signaux appartenant aux espaces $L^p(\mathbb{R})$, avec $p = 1$ ou 2 .

Définition 1.8 (Espace $L^p(\mathbb{R})$). On distingue deux cas.

— Soit $1 \leq p < \infty$. On appelle espace $L^p(\mathbb{R})$ l'espace vectoriel

$$L^p(\mathbb{R}) = \left\{ x : \mathbb{R} \rightarrow \mathbb{C} \text{ tel que } \int_{\mathbb{R}} |x(t)|^p dt < \infty \right\}.$$

On lui associe une norme notée $\|\cdot\|_p$ et définie par

$$\|x\|_p = \left(\int_{\mathbb{R}} |x(t)|^p dt \right)^{\frac{1}{p}}, \quad \forall x \in L^p(\mathbb{R}).$$

— Soit $p = \infty$. On appelle espace $L^\infty(\mathbb{R})$, l'espace vectoriel des fonction essentiellement bornées :

$$L^\infty(\mathbb{R}) = \{ x : \mathbb{R} \rightarrow \mathbb{C} \text{ tel que } \sup \text{ess } |x(t)| < \infty \}.$$

On lui associe une norme notée $\|\cdot\|_\infty$ et définie par $\|x\|_\infty = \sup \text{ess } |x|$.

Ces espaces sont naturellement associés aux signaux physiques. En particulier, l'espace $L^2(\mathbb{R})$ contient les signaux d'énergie finie (c'est-à-dire les signaux de carré intégrable). La transformée de Fourier est originellement définie pour les signaux intégrables, c'est-à-dire appartenant à l'espace $L^1(\mathbb{R})$. Sa définition s'étend ensuite à l'espace $L^2(\mathbb{R})$ par densité de $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ dans $L^2(\mathbb{R})$.

Définition 1.9 (Transformée de Fourier). Soit $x \in L^1(\mathbb{R})$. La transformée de Fourier \hat{x} du signal x est donnée par

$$\hat{x}(\xi) = \int_{\mathbb{R}} x(t) e^{-2i\pi \xi t} dt, \quad (1.3)$$

où $\xi \in \mathbb{R}$ désigne la fréquence.

La transformée de Fourier de $x \in L^1(\mathbb{R})$ est alors une application bornée et continue sur \mathbb{R} qui tend vers zéro en l'infini (c'est le théorème de Riemann-Lebesgue). On peut alors montrer que si elle est aussi intégrable, on a la théorème d'inversion suivant (voir [38] pour la démonstration).

Théorème 1.1 (Transformée de Fourier inverse). Soit $x \in L^1(\mathbb{R})$. Si $\hat{x} \in L^1(\mathbb{R})$ alors, le signal x s'exprime en fonction de sa transformée de Fourier. On a

$$x(t) = \int_{\mathbb{R}} \hat{x}(\xi) e^{2i\pi\xi t} d\xi, \quad (1.4)$$

en tout point de continuité de x .

Une fois la transformée de Fourier étendue à $L^2(\mathbb{R})$, on peut montrer que cette opération conserve le produit scalaire et la norme dans $L^2(\mathbb{R})$. Ce sont des propriétés qu'on utilisera par la suite.

Théorème 1.2 (Formules de Plancherel et de Parseval). Soit x et y dans $L^2(\mathbb{R})$. Alors, la formule de Plancherel est donnée par :

$$\int_{\mathbb{R}} x(t) \overline{y(t)} dt = \int_{\mathbb{R}} \hat{x}(\xi) \overline{\hat{y}(\xi)} d\xi. \quad (1.5)$$

On en déduit la formule de Parseval (en prenant $x = y$) :

$$\int_{\mathbb{R}} |x(t)|^2 dt = \int_{\mathbb{R}} |\hat{x}(\xi)|^2 d\xi. \quad (1.6)$$

Dans la suite, nous aurons régulièrement affaire à des signaux n'appartenant pas à $L^1(\mathbb{R})$ ou $L^2(\mathbb{R})$. La transformée de Fourier de ces signaux sera alors prise comme étant son extension au sens des distributions. En effet, comme la transformée de Fourier est définie pour les fonctions de l'espace de Schwartz (car il est inclus dans $L^1(\mathbb{R})$), la transformée de Fourier \hat{F} d'une distribution F est donnée par $\langle \hat{F}, g \rangle = \langle F, \hat{g} \rangle$, pour toute fonction g de l'espace de Schwartz.

Un processus stochastique stationnaire X possède une variance constante, soit : $\mathbb{E}\{X(t)\} = \sigma_X^2, \forall t \in \mathbb{R}$. De manière générale, les réalisations de ces processus ne sont donc pas des signaux d'énergie finie [26]. Néanmoins, la représentation de Cramér [17] permet de généraliser la transformée de Fourier aux processus stochastiques stationnaires. Cependant, avec la définition donnée ci-dessus, la transformée de Fourier d'un processus stochastique est elle-même un processus stochastique. Pour obtenir une grandeur déterministe qui quantifie le contenu fréquentiel d'un processus, on s'intéresse à la moyenne quadratique de la transformée de Fourier. On obtient alors la densité spectrale de puissance d'un processus stationnaire qui est une grandeur déterministe et qui permet également de caractériser le contenu fréquentiel de ce type de signaux. Pour définir cette dernière, énonçons le théorème de Wiener-Khintchine.

Théorème 1.3 (Wiener-Khintchine). *Soit X un processus stationnaire au sens large. Soit k_X son autocovariance telle que $k_X(\tau) < \infty, \forall \tau \in \mathbb{R}$. Si k_X est continue en 0, alors, il existe une unique mesure positive ou nulle η telle que*

$$k_X(\tau) = \int_{\mathbb{R}} e^{2i\pi\xi\tau} d\eta(\xi) . \quad (1.7)$$

On peut alors associer au processus stationnaire X sa densité spectrale de puissance définie ci-dessous.

Définition 1.10 (Densité spectrale de puissance). Lorsque la mesure η est continue par rapport à la mesure de Lebesgue, il existe \mathcal{S}_X mesurable tel que

$$d\eta(\xi) = \mathcal{S}_X(\xi) d\xi$$

La fonction \mathcal{S}_X est appelée densité spectrale de puissance du processus X . On appelle puissance de X la quantité :

$$\sigma_X^2 = \int_{\mathbb{R}} \mathcal{S}_X(\xi) d\xi = k_X(0) .$$

Dans la suite, on utilisera indistinctement les termes de spectre ou de densité spectrale pour désigner la densité spectrale de puissance des processus stationnaires. Le théorème de Wiener-Khintchine stipule donc que l'on obtient (sous certaines hypothèses) le spectre d'un processus stationnaire à partir de la transformée de Fourier de la fonction d'autocovariance. Dans un cadre plus général, la notion de densité spectrale peut être étendue aux processus stochastiques généralisés. C'est dans ce cadre que nous nous plaçons par la suite. On a alors $\mathcal{S}_X = \hat{k}_X$ au sens des distributions. Par exemple dans le cas du bruit blanc, étant donné que $k_X = \sigma_X^2 \delta$, on a $\mathcal{S}_X(\xi) = \sigma_X^2, \forall \xi \in \mathbb{R}$. En effet, la transformée de Fourier au sens des distributions de l'impulsion de Dirac est la « fonction » constante valant 1.

1.2.3.2 Estimation de la densité spectrale de puissance

Comme énoncé précédemment, la connaissance du spectre d'un processus stationnaire nous informe sur le contenu fréquentiel de ce dernier. Considérons, par exemple, un processus stochastique sinusoïdal donné par $X(t) = Ae^{i(2\pi\nu t + \varphi)}$, où φ est une variable aléatoire uniforme sur l'intervalle $[0, 2\pi]$. Les grandeurs que l'on souhaite déterminer sont la fréquence des oscillations ν et l'amplitude A du signal. Son spectre nous donne ces informations. En effet, il s'exprime : $\mathcal{S}_X(\xi) = A^2 \delta(\xi - \nu)$. Ainsi, pour

connaître ν , il suffit de déterminer la fréquence de la raie contenue dans la densité spectrale; pour connaître A , il suffit de déterminer son amplitude. Dans le but d'estimer la densité spectrale de puissance, de nombreuses méthodes d'estimation ont été développées. Nous allons présenter des méthodes s'appuyant sur le périodogramme ou le périodogramme modifié que l'on définit ci-dessous [67].

Définition 1.11 (Périodogramme). Soit X un processus stationnaire au sens large. Soit $L > 0$. Soit $h_L \in L^2(\mathbb{R})$ une fonction paire, à support compact dans $[-\frac{L}{2}, \frac{L}{2}]$ vérifiant les deux conditions suivantes :

$$\begin{cases} \|h_L\|_2 = 1; \\ \int_{\mathbb{R}} |\hat{h}_L(\xi)|^2 g(\xi) d\xi \xrightarrow{L \rightarrow \infty} g(0), \quad \forall g \in L^2(\mathbb{R}). \end{cases} \quad (1.8)$$

Soit \hat{X}_{h_L} la transformée de Fourier du processus fenêtré $X.h_L$:

$$\hat{X}_{h_L}(\xi) = \int_{-L/2}^{L/2} X(t) h_L(t) e^{-2i\pi \xi t} dt.$$

- Si $h_L(t) = \frac{1}{\sqrt{L}} \mathbf{1}_{[-\frac{L}{2}, \frac{L}{2}]}(t)$, le processus $|\hat{X}_{h_L}|^2$ est appelé périodogramme.
- Dans les autres cas, le processus $|\hat{X}_{h_L}|^2$ est appelé périodogramme modifié.

Une fenêtre couramment utilisée pour calculer le périodogramme modifié est la fenêtre de Hann. Elle est définie par

$$h_L(t) = \begin{cases} \frac{1}{L} + \frac{1}{L} \cos(2\pi \frac{t}{L}) & \text{si } t \in [-\frac{L}{2}, \frac{L}{2}] \\ 0 & \text{sinon.} \end{cases} \quad (1.9)$$

On peut montrer que le périodogramme est un estimateur biaisé de la densité spectrale de puissance. En effet, son espérance s'exprime ainsi :

$$\mathbb{E} \{ |\hat{X}_{h_L}(\xi)|^2 \} = (\mathcal{S}_X * |\hat{h}_L|^2)(\xi).$$

En revanche, cet estimateur est asymptotiquement non biaisé. En effet, la condition (1.8) nous permet d'avoir

$$\mathcal{S}_X(\xi) = \lim_{L \rightarrow \infty} \mathbb{E} \{ |\hat{X}_{h_L}(\xi)|^2 \}.$$

Cette propriété nous montre que plus la durée d'observation est longue, plus la précision de l'estimateur augmente. En effet, l'effet du fenêtrage est de lisser le spectre \mathcal{S}_X par $|\hat{h}_L|^2$ et donc, de diminuer la résolution fréquentielle de l'estimateur d'autant plus fortement que $|\hat{h}_L|^2$ est étalé autour de la fréquence nulle.

Différentes méthodes se sont appuyées sur cette propriété du périodogramme pour bâtir des estimateurs de la densité spectrale de puissance. Nous présentons ici la méthode de Welch qui est la méthode statistique d'analyse spectrale la plus performante.

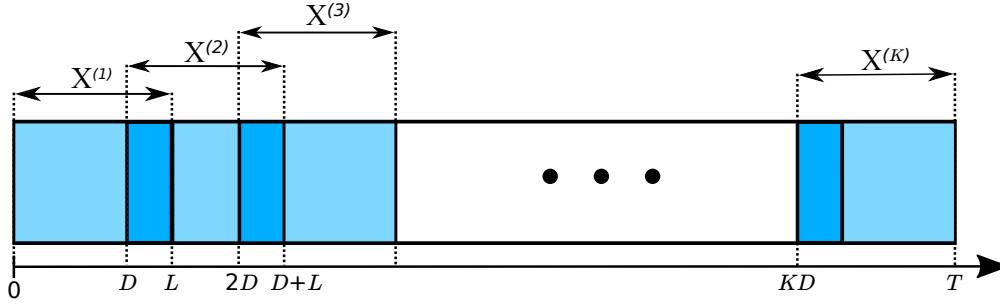


FIGURE 1.2 – Illustration des fenêtres chevauchantes de la méthode de Welch.

L'idée est de moyenner les périodogrammes de portions du signal complet afin de limiter les fluctuations de l'estimateur, c'est-à-dire sa variance. Pour définir cet estimateur, on introduit quelques grandeurs.

- Soit T la durée des observations. L'horizon temporel des mesures dont nous disposons est donc l'intervalle $[0, T]$.
- Le support de la fenêtre h_L utilisée pour calculer le périodogramme est de longueur L .
- Soit $D \leq L$ le pas avec lequel nous effectuons les estimations successives des périodogrammes. Nous allons calculer les périodogrammes successifs des processus $X^{(k)}$ définis par

$$X^{(k)}(t) = X\left(t + \frac{L}{2} + kD\right) \quad \text{avec } k \in \{0, \dots, K-1\}.$$

Ces périodogrammes sont notés $|\hat{X}_{h_L}^{(k)}|^2$. Le recouvrement temporel entre deux périodogrammes successifs est de $L - D$. Par ailleurs, le nombre K de périodogrammes calculés dépend des différents intervalles temporels comme suit : $K = \lfloor \frac{T-L}{D} \rfloor + 1$. L'illustration de la figure 1.2 nous permet de visualiser le sens à donner à ces différentes grandeurs.

L'estimateur de Welch [79] de la densité spectrale de puissance est alors donné par :

$$\tilde{\mathcal{S}}_X(\xi) = \frac{1}{K} \sum_{k=1}^K |\hat{X}_{h_L}^{(k)}(\xi)|^2. \quad (1.10)$$

Le biais de cet estimateur est le même que celui du périodogramme. En revanche, sa variance diminue d'autant plus fortement que le nombre de périodogrammes calculés K est grand. La longueur du signal étant fixée, une manière naturelle d'augmenter le nombre de périodogrammes K est d'augmenter le chevauchement de ceux-ci en diminuant le pas D . Cependant, un choix de D trop petit induit une quasi superposition entre deux périodogrammes successifs, ces derniers contiennent donc quasiment la même information. Ceci entraîne alors une très faible diminution de la variance tout en augmentant le temps de calcul de l'estimateur de Welch. Pour conserver un apport d'information non négligeable entre deux périodogrammes successifs, le choix

classiquement fait en pratique consiste à prendre $D = 0,5L$. Dans ce cas, on a alors $K = \lfloor \frac{2T}{L} \rfloor - 1$. Un compromis doit donc être fait sur la longueur des fenêtres L : suffisamment grande pour augmenter la résolution fréquentielle (c'est-à-dire diminuer le biais) et suffisamment petite pour diminuer les fluctuations de l'estimateur (K doit rester grand).

Remarque. Les estimateurs du périodogramme ou de Welch s'appuient sur des calculs de transformées de Fourier. En pratique, nous n'avons accès qu'à des signaux discrets. Les évaluations numériques de ces estimateurs sont donc obtenues en remplaçant les opérateurs intégraux par leurs versions discrétisées (comme, par exemple, la transformée de Fourier discrète).

1.3 Analyse spectrale en contexte non stationnaire

Comme nous l'avons justifié précédemment, l'hypothèse de stationnarité n'est pas raisonnable pour l'étude de la majorité des signaux physiques mesurés. C'est pourquoi nous allons nous intéresser à l'étude des processus stochastiques non stationnaires. Dans la suite, on considère uniquement des processus non stationnaires de moyenne nulle. On suppose donc que toute tendance a été éliminée des observations.

1.3.1 Le spectre instantané

En règle générale, associer un spectre à un processus non stationnaire est une question qui n'est pas aisée dans la mesure où il n'existe pas d'opérateur possédant toutes les propriétés désirées d'un spectre (positivité, covariance par décalage en temps et en fréquence, marginales) quel que soit le processus analysé. Priestley [61] a alors proposé une première définition de spectre évolutif pour une certaine classe de processus localement stationnaires (voir la section 1.4.1). Cette définition n'étant utilisable que pour une classe restreinte de processus, Matz *et al.* [41] ont étendu la notion de spectre évolutif à une plus grande gamme de processus non stationnaires. Dans la suite, la définition du spectre variant dans le temps que nous utiliserons sera celle du spectre de Wigner-Ville. Cette définition possède l'avantage d'être générale mais elle ne vérifie pas la propriété de positivité (entre autres).

Dans le cas des processus non stationnaires, l'autocovariance à un paramètre n'est plus définie. En revanche, on peut définir une nouvelle fonction d'autocovariance instantanée R_X dépendant de deux paramètres :

$$R_X(\tau, t) = \mathbb{E}\{X(t + \tau/2)\overline{X(t - \tau/2)}\}, \quad \forall t, \tau \in \mathbb{R}. \quad (1.11)$$

On pourra trouver, dans la littérature, d'autres conventions pour définir l'autocovariance instantanée comme $\mathbb{E}\{X(t)\overline{X(t - \tau)}\}$. Dans tout ce qui suit, nous conserverons la convention donnée par l'équation (1.11). Le paramètre τ est appelé variable stationnaire car l'autocovariance instantanée R_X d'un processus stationnaire ne dépend que du décalage temporel τ , et on a : $R_X(\tau, t) = k_X(\tau)$. Par suite, le paramètre t est appelé

variable non stationnaire. On parle aussi parfois de variable dynamique pour désigner le paramètre t .

Le théorème de Wiener-Khintchine n'étant plus valable pour les processus non stationnaires, nous ne pouvons plus associer de densité spectrale à de tels processus. Cependant, on peut leur associer un spectre variant dans le temps, appelé spectre instantané et noté $S_X(\xi, t)$. Par analogie avec la formule de Wiener-Khintchine, celui-ci sera la transformée de Fourier de R_X par rapport à sa variable stationnaire. On a alors

$$\begin{aligned} S_X(\xi, t) &= \int_{\mathbb{R}} R_X(\tau, t) e^{-2i\pi\xi\tau} d\tau \\ &= \mathbb{E} \left\{ \int_{\mathbb{R}} X(t + \tau/2) \overline{X(t - \tau/2)} e^{-2i\pi\xi\tau} d\tau \right\}. \end{aligned} \quad (1.12)$$

Le spectre instantané est alors obtenu à partir de l'espérance de la distribution de Wigner-Ville (cf. définition 1.12) du processus non stationnaire. À l'inverse de la densité spectrale des processus stationnaires, le spectre instantané peut prendre des valeurs négatives : on ne peut donc pas parler de densité spectrale instantanée.

Définition 1.12 (Distribution de Wigner-Ville). Soit $x \in L^2(\mathbb{R})$. Alors, la distribution de Wigner-Ville de x est donnée par :

$$V_x(\xi, t) = \int_{\mathbb{R}} x(t + \tau/2) \overline{x(t - \tau/2)} e^{-2i\pi\xi\tau} d\tau. \quad (1.13)$$

Cette expression de la distribution de Wigner-Ville peut être étendue aux cas où x est une distribution (voir [29]). Une étude complète des propriétés de la distribution de Wigner-Ville peut être trouvée dans [25]. Nous n'en présentons que quelques unes ici.

Covariance par translation et modulation. La distribution de Wigner-Ville est covariante par translation temporelle et par modulation fréquentielle. Soit x_{uv} le signal défini par :

$$x_{uv}(t) = e^{2i\pi vt} x(t - u).$$

C'est la translatée en temps (par u) et modulée en fréquence (par v) du signal x . Alors, la distribution de Wigner-Ville de x_{uv} et la translatée de celle de x dans le plan temps-fréquence, de sorte que

$$V_{x_{uv}}(\xi, t) = V_x(\xi - v, t - u).$$

Chirps linéaires. Soit x un chirp linéaire, c'est-à-dire un signal sinusoïdal dont la fréquence varie linéairement. Il s'écrit alors :

$$x(t) = e^{2i\pi\left(\frac{a}{2}t^2 + bt + c\right)}.$$

Sa distribution de Wigner-Ville prend alors la forme suivante :

$$V_x(\xi, t) = \delta(\xi - (at + b)) . \quad (1.14)$$

Il s'agit de la représentation idéale de la fréquence instantanée d'un chirp linéaire dans le plan temps-fréquence.

Marginales. À la manière d'une densité de probabilité, lorsqu'on « marginalise » la distribution de Wigner-Ville par rapport à la fréquence, on obtient le module carré du signal. De même, lorsqu'on la marginalise par rapport au temps, on obtient le module carré de la transformée de Fourier du signal. Autrement dit, on a :

$$\begin{aligned} |x(t)|^2 &= \int_{\mathbb{R}} V_x(\xi, t) d\xi \\ |\hat{x}(\xi)|^2 &= \int_{\mathbb{R}} V_x(\xi, t) dt . \end{aligned}$$

Lissage de la distribution de Wigner-Ville. La distribution de Wigner-Ville étant un opérateur quadratique, elle laisse apparaître des interférences entre signaux. Prenons en exemple deux sinusoides : $x_k(t) = e^{2i\pi\nu_k t}$ pour $k \in \{1, 2\}$. Alors :

$$V_{x_k}(\xi, t) = \delta(\xi - \nu_k), \quad \forall t \in \mathbb{R} .$$

Cependant, la distribution de Wigner-Ville de $x_1 + x_2$ s'écrit :

$$V_{x_1+x_2}(\xi, t) = V_{x_1}(\xi, t) + V_{x_2}(\xi, t) + V_{x_1x_2}(\xi, t) ,$$

où

$$V_{x_1x_2}(\xi, t) = \cos(2\pi(\nu_1 - \nu_2)t) \delta\left(\xi - \frac{\nu_1 + \nu_2}{2}\right) . \quad (1.15)$$

Ce dernier terme est un terme qui caractérise le phénomène d'interférence qui apparaît à la fréquence moyenne entre ν_1 et ν_2 . Par ailleurs, comme le montre l'expression précédente, la distribution de Wigner-Ville peut prendre des valeurs négatives dans les zones d'interférences du plan temps-fréquence. Elle ne peut donc pas être interprétée comme une densité spectrale. Une solution pour tenter de résoudre ces problèmes consiste à effectuer un lissage de la distribution de Wigner-Ville, l'objectif étant d'éliminer les oscillations introduites par les interférences en moyennant localement la distribution de Wigner-Ville. L'opération de lissage s'écrit :

$$\iint_{\mathbb{R}^2} H(t, t', \xi, \xi') V_x(t', \xi') dt' d\xi' . \quad (1.16)$$

où H constitue le filtre de lissage. Nous verrons dans la suite que la transformée de Fourier à court terme et la transformée en ondelettes conduisent naturellement à des expressions particulières du filtre H .

En pratique, nous n'observons qu'une réalisation x du processus X . On peut donc calculer sa distribution de Wigner-Ville V_x mais, nous ne pouvons pas approcher S_X à partir de celle-ci car, pour approcher $\mathbb{E}\{V_X\}$ nous aurions besoin de plusieurs réalisations de X . Par ailleurs, un moyennage temporel global semblable à celui effectué par la méthode de Welch serait inefficace ici du fait de la dépendance temporelle due à la non-stationnarité de V_x . En fait, les outils d'analyse temps-fréquence et temps-échelle, présentés dans la section suivante, peuvent être vus comme des estimateurs du spectre instantané (grâce à la propriété de lissage énoncée précédemment).

1.3.2 Les outils d'analyse : temps-fréquence et temps-échelle

Nous avons vu, dans la section précédente, que la transformée de Fourier constitue un outil d'analyse particulièrement adapté à la question de l'analyse spectrale lorsque l'on est confronté à un signal stationnaire. En effet, la méthode de Welch permet d'estimer la densité spectrale à partir de la transformée de Fourier de périodogrammes successifs. Dans ce qui suit, nous présentons la transformée de Fourier à court terme et la transformée en ondelettes qui sont des outils inspirés de la transformée de Fourier. Leur construction est analogue à celles des périodogrammes dans la mesure où on calcule des transformées de Fourier localement autour d'un instant donné, ce qui leur permet d'être mieux adaptées à l'analyse spectrale de signaux non stationnaires. Nous verrons pourquoi.

Afin de simplifier les expressions que nous aurons dans ce chapitre, introduisons d'abord quelques opérations élémentaires sur les signaux.

1.3.2.1 Opérateurs élémentaires sur les signaux

Les opérations suivantes sont des opérations élémentaires préservant la stationnarité.

Définition 1.13 (Opérations élémentaires sur les signaux). Soit x un signal donné. Soit $q > 0$. On définit les opérations élémentaires suivantes :

- l'opération de translation T_τ (avec $\tau \in \mathbb{R}$) telle que

$$T_\tau x(t) = x(t - \tau), \quad \forall t \in \mathbb{R}; \quad (1.17)$$

- l'opération de modulation d'amplitude A_α (avec $\alpha > 0$) telle que

$$A_\alpha x(t) = \alpha x(t), \quad \forall t \in \mathbb{R}; \quad (1.18)$$

- l'opération de modulation fréquentielle M_ν (avec $\nu \in \mathbb{R}$) telle que

$$M_\nu x(t) = x(t) e^{2i\pi \nu t}, \quad \forall t \in \mathbb{R}; \quad (1.19)$$

- l'opération de dilatation D_s (avec $s > 0$) telle que

$$D_s x(t) = q^{\frac{s}{2}} x(q^{-s} t), \quad \forall t \in \mathbb{R}. \quad (1.20)$$

À l'exception de la modulation d'amplitude, tous ces opérateurs sont unitaires sur $L^2(\mathbb{R})$. Par ailleurs, les opérations définies ci-dessus commutent de la manière suivante.

Proposition 1.1 (Commutation des opérations élémentaires). *Soit $q > 0$. On a les relations de commutation suivantes :*

$$T_\tau D_s = D_s T_{q^{-s}\tau}, \quad T_\tau M_\xi = e^{-2i\pi\xi\tau} M_\xi T_\tau, \quad M_\xi D_s = D_s M_{q^s\xi},$$

et l'opérateur de modulation d'amplitude commute avec tous les autres opérateurs.

1.3.2.2 Analyse temps-fréquence

La transformée de Fourier à court terme (abrégée TFCT dans ce manuscrit) est un outil d'analyse adapté aux signaux non stationnaires car elle permet une analyse locale des signaux.

Définition 1.14 (Transformée de Fourier à court terme). Soit $x \in L^2(\mathbb{R})$ le signal à analyser. Soit $g \in L^2(\mathbb{R})$ la fenêtre d'analyse. Alors, la transformée de Fourier à court terme \mathcal{G}_x de x est donnée par :

$$\mathcal{G}_x(\xi, \tau) = \langle x, T_\tau M_\xi g \rangle = \int_{\mathbb{R}} x(t) \overline{g}(t - \tau) e^{-2i\pi\xi(t-\tau)} dt, \quad \forall \xi, \tau \in \mathbb{R}.$$

Le paramètre ξ est appelé fréquence, tandis que τ est le paramètre de temps. On appelle spectrogramme le module carré de la transformée Fourier à court terme.

La convention de phase $-2\pi\xi(t - \tau)$ dans l'exponentielle complexe donne une transformée de Fourier à court terme covariante par translation temporelle. La convention de phase $-2\pi\xi t$ est une autre choix possible qui donne une transformée covariante par modulation de fréquence. Dans la suite, nous nous limitons à la convention retenue dans la définition 1.14.

La fenêtre d'analyse g est une fonction régulière qui décroît rapidement lorsqu'on s'éloigne de l'origine. La gaussienne ($g(t) = e^{-t^2}$) ou la fenêtre de Hann (voir l'équation (1.9)) constituent des choix standards pour la fenêtre d'analyse. Notons $g_{\tau\xi} = T_\tau M_\xi g$. La fonction $g_{\tau\xi}$ est alors concentrée temporellement autour de l'instant τ et fréquentiellement autour de la fréquence ξ . Faire varier les paramètres ξ et τ permet donc d'analyser localement le signal autour de chaque point temps-fréquence souhaité.

On peut également établir une formule d'inversion permettant de reconstruire le signal x à partir de sa transformée de Fourier à court terme \mathcal{G}_x . La formule d'inversion est donnée par le théorème suivant [29].

Théorème 1.4 (Transformée de Fourier à court terme inverse). Soit $x \in L^2(\mathbb{R})$. Soit $h \in L^2(\mathbb{R})$ la fenêtre de synthèse telle que $\langle h, g \rangle \neq 0$. Alors, le signal x s'exprime en fonction de sa transformée de Fourier à court terme \mathcal{G}_x par

$$x(t) = \frac{1}{\langle h, g \rangle} \iint_{\mathbb{R}^2} \mathcal{G}_x(\xi, \tau) h(t - \tau) e^{2i\pi\xi(t-\tau)} d\tau d\xi.$$

En pratique, on se place souvent dans le cas particulier où $h = g$. La formule d'inversion est alors donnée par

$$x(t) = \frac{1}{\|g\|_2^2} \iint_{\mathbb{R}^2} \mathcal{G}_x(\xi, \tau) g(t - \tau) e^{2i\pi\xi(t-\tau)} d\tau d\xi.$$

La définition de la transformée de Fourier à court terme exige que le signal à analyser soit dans l'espace $L^2(\mathbb{R})$. Toutefois, les signaux que nous allons analyser n'appartiennent généralement pas à cet espace. Cela ne constitue pas un problème car, à l'image de la transformée de Fourier, la définition de la transformée de Fourier à court terme est étendue aux distributions. Dans ce cas, les contraintes sur la décroissance de la fenêtre d'analyse sont plus fortes. On prend g dans l'espace de Schwartz. Alors, toute réalisation x d'un processus généralisé X étant une distribution, la transformée de Fourier à court terme de x est donnée par : $\mathcal{G}_x(\xi, \tau) = x(T_\tau M_\xi g)$. Nous nous plaçons dans ce cadre plus général dans la suite. Par ailleurs, la transformée de Fourier à court terme d'un processus stochastique sera elle-même un processus stochastique. Les transformées des signaux seront donc vues comme des réalisations de ces processus.

Dans le cas de signaux stationnaires, le spectrogramme permet d'obtenir une estimation du spectre. En effet, si la fenêtre d'analyse est paire et à support compact, le spectrogramme $|\mathcal{G}_X(\xi, \tau)|^2$ est identique (à un facteur de normalisation près) au périodogramme modifié $|\hat{X}_g|^2$ introduit dans la définition 1.11. Les propriétés statistiques du spectrogramme sont donc indépendantes de τ . On obtient alors une estimation du spectre de X grâce à l'estimateur de Welch qui est ici donné par :

$$\mathcal{S}_X(\xi) = \frac{1}{K\|g\|_2^2} \sum_{k=1}^K |\mathcal{G}_X(\xi, \tau_k)|^2,$$

où le choix des instants τ_k est issu d'un compromis discuté lors de la présentation de la méthode de Welch (voir le paragraphe 1.2.3.2).

Dans le cas de signaux non stationnaires, le spectrogramme constitue un estimateur du spectre instantané car il s'agit d'une version localement lissée de la distribution de Wigner-Ville du signal. Il s'agit d'un cas particulier de l'équation (1.16) où le filtre de lissage est donné. En effet, on a [3] :

$$|\mathcal{G}_x(\xi, \tau)|^2 = \iint_{\mathbb{R}^2} V_g(\nu - \xi, t - \tau) V_x(\nu, t) dt d\nu. \quad (1.21)$$

où V_g désigne la distribution de Wigner-Ville de la fenêtre d'analyse g . Étant donné que la fenêtre g est localisée en temps autour de l'origine et que c'est un filtre passe-bas, la distribution V_g sera concentrée autour de l'origine dans le plan temps-fréquence. C'est pourquoi V_g va agir sur V_x comme une opération de lissage local, estimant ainsi le spectre instantané par approximation de l'espérance par une moyenne locale. De plus, contrairement au spectre instantané, le spectrogramme ne prend pas de valeurs négatives. Son interprétation en terme de densité spectrale est donc plus aisée.

1.3.2.3 Analyse temps-échelle

La transformée en ondelettes est un outil d'analyse adapté aux signaux non stationnaires dont la construction est assez similaire à celle de la transformée de Fourier à court terme.

Définition 1.15 (Transformée en ondelettes). Soit $x \in L^2(\mathbb{R})$ le signal à analyser. Soit $\psi \in L^2(\mathbb{R})$ l'ondelette d'analyse. Soit $q > 1$. Alors, la transformée en ondelettes \mathcal{W}_x de x est donnée par :

$$\mathcal{W}_x(s, \tau) = \langle x, T_\tau D_s \psi \rangle = \int_{\mathbb{R}} x(t) q^{-\frac{s}{2}} \overline{\psi} \left(\frac{t - \tau}{q^s} \right) dt, \quad \forall s > 0, \tau \in \mathbb{R}.$$

Le paramètre s est appelé échelle, tandis que τ est le paramètre de temps. On appelle scalogramme le module carré de la transformée en ondelettes.

De la même manière que pour la fenêtre d'analyse, l'ondelette d'analyse ψ est une fonction régulière qui décroît rapidement lorsqu'on s'éloigne de l'origine. Par ailleurs, l'ondelette est une fonction oscillante (de moyenne nulle). Notons ν_0 sa fréquence centrale et $\psi_{s\tau} = T_\tau D_s \psi$. La fonction $\psi_{s\tau}$ est donc concentrée temporellement autour de l'instant τ et fréquentiellement autour de la fréquence $q^{-s}\nu_0$. Faire varier les paramètres s et τ permet donc d'analyser localement le signal autour de chaque point temps-échelle souhaitée. Il est à noter que l'analyse des hautes fréquences du signal est obtenue en faisant tendre le paramètre d'échelle s vers $-\infty$, et inversement.

L'ondelette de Morlet (gaussienne modulée par une exponentielle complexe) ou celle de Morse (voir [37]) constituent des choix standards pour l'ondelette d'analyse. Dans ce document, l'ondelette d'analyse sera choisie comme étant analytique c'est-à-dire appartenant à l'espace $H^2(\mathbb{R})$ tel que

$$H^2(\mathbb{R}) = \{ \psi \in L^2(\mathbb{R}) \mid \text{tel que } \text{supp}(\hat{\psi}) \subset \mathbb{R}_+ \},$$

où $\text{supp}(\hat{\psi})$ désigne le support de $\hat{\psi}$. Il peut être montré que, pour des choix appropriés de ψ , la transformée en ondelettes est inversible (voir [9, 38]). La formule d'inversion est alors donnée par le théorème suivant.

Théorème 1.5 (Transformée en ondelettes inverse). Soit $x \in L^2(\mathbb{R})$ un signal réel. Soit $\psi \in H^2(\mathbb{R})$ une ondelette analytique admissible, c'est-à-dire vérifiant :

$$C_\psi = \int_0^\infty \frac{|\hat{\psi}(\xi)|^2}{\xi} d\xi < \infty.$$

Alors, le signal x s'exprime en fonction de sa transformée en ondelettes \mathcal{W}_x par

$$x(t) = \frac{2 \ln(q)}{C_\psi} \operatorname{Re} \left(\iint_{\mathbb{R}^2} \mathcal{W}_x(s, \tau) q^{-\frac{s}{2}} \psi \left(\frac{t - \tau}{q^s} \right) d\tau \frac{ds}{q^s} \right).$$

La définition de la transformée en ondelettes exige que le signal à analyser soit dans l'espace $L^2(\mathbb{R})$. Toutefois, comme pour la transformée de Fourier à court terme, la définition de la transformée en ondelettes est étendue aux distributions. Dans ce cas, les contraintes sur la décroissance de l'ondelette d'analyse sont plus fortes. On prend ψ dans l'espace de Schwartz. Alors, toute réalisation x d'un processus généralisé X étant une distribution, la transformée en ondelettes de x est donnée par : $\mathcal{W}_x(s, \tau) = x(T_\tau D_s \psi)$. Nous nous plaçons dans ce cadre plus général dans la suite. Par ailleurs, la transformée en ondelettes d'un processus stochastique sera elle-même un processus stochastique. Les transformées des signaux seront donc vues comme des réalisations de ces processus.

Dans le cas de signaux stationnaires, le scalogramme permet d'obtenir une estimation du spectre. En effet, dans ce cadre, les propriétés statistiques du scalogramme $|\mathcal{W}_X(s, \tau)|^2$ sont indépendantes de τ . À la manière de l'estimateur de Welch, on peut donc effectuer une moyenne temporelle du scalogramme. On obtient alors l'estimateur du spectre-ondelettes $\tilde{\mathcal{S}}_X$ défini par :

$$\tilde{\mathcal{S}}_X(q^{-s} \nu_0) = \frac{1}{K \|\psi\|_2^2} \sum_{k=1}^K |\mathcal{W}_X(s, \tau_k)|^2. \quad (1.22)$$

Il s'agit également d'un estimateur du spectre biaisé du fait du filtrage par l'ondelette. On peut montrer que :

$$\mathbb{E} \{ \tilde{\mathcal{S}}_X(q^{-s} \nu_0) \} \stackrel{\Delta}{=} \mathcal{S}_{X, \psi}(q^{-s} \nu_0) = \frac{1}{\|\psi\|_2^2} \int_0^\infty \mathcal{S}_X(\nu) q^s |\hat{\psi}(q^s \nu)|^2 d\nu, \quad (1.23)$$

où ν_0 est la fréquence centrale de l'ondelette. $\mathcal{S}_{X, \psi}$ est une version filtrée passe-bande de \mathcal{S}_X centrée autour de la fréquence $\xi = q^{-s} \nu_0$. De plus, la bande passante du filtre est proportionnelle à la fréquence ν . Ainsi, à la différence de l'estimateur de Welch, le biais de l'estimateur du spectre-ondelettes est d'autant plus grand que l'on se trouve aux hautes fréquences.

Dans le cas de signaux non stationnaires, le scalogramme constitue un estimateur du spectre instantané car, à l'image du spectrogramme, il s'agit d'une version localement lissée de la distribution de Wigner-Ville du signal. Il s'agit d'un cas particulier de

l'équation (1.16) où le filtre de lissage est donné. En effet, on a [38] :

$$|\mathcal{W}_x(s, \tau)|^2 = \iint_{\mathbb{R}^2} V_\psi \left(q^s \nu, \frac{t - \tau}{q^s} \right) V_x(\nu, t) dt d\nu . \quad (1.24)$$

où V_ψ désigne la distribution de Wigner-Ville de l'ondelette d'analyse ψ (cf. la définition 1.12). Contrairement au spectre instantané, le scalogramme ne prend pas de valeurs négatives, son interprétation en terme de densité spectrale est donc plus aisée.

Remarque. Dans la suite de ce manuscrit, l'ensemble des évaluations numériques de la transformée en ondelettes seront effectuées en fixant la constante d'échelle q telle que $q = 2$. Celle-ci agit comme un sélecteur d'unité pour l'échelle s . Par exemple, en terminologie musicale, $q = 2$ signifie que s est mesurée en octaves, alors que lorsque $q = 2^{1/12}$, s est mesurée en demi-tons.

1.4 Des modèles de non-stationnarités

Il faut bien voir qu'il n'existe pas de modèle générique de non-stationnarité, et que chaque modèle que l'on peut proposer n'est adapté qu'à certains types de signaux. Dans la suite, différents modèles de non-stationnarités sont présentés. Deux approches peuvent être distinguées. Dans la première, on modélise des signaux approximativement stationnaires sur des intervalles de temps successifs, de longueurs bien choisies. Dans la seconde approche, on définit des modèles de non-stationnarités plus globaux ; les signaux modélisés sont obtenus en appliquant des déformations brisant la stationnarité à des processus stationnaires.

1.4.1 Des signaux localement stationnaires

Lorsque X est un processus stationnaire les valeurs de la fonction d'autocovariance $R_X(t, \tau)$ ne dépendent que du décalage temporel τ et l'on a $R_X(t, \tau) = k_X(\tau)$. Les premiers modèles de processus non stationnaires ont cherché à approcher localement les fonctions d'autocovariance de ces processus par des fonctions d'autocovariance de processus stationnaires. Ce sont des modèles non paramétriques.

Dans le but d'approcher la formule de Wiener-Khintchine (1.7) permettant d'exprimer l'autocovariance en fonction de la densité spectrale, Silverman [66] a proposé une première définition des processus *localement stationnaires*. Il s'agit des processus dont l'autocovariance instantanée peut se décomposer sous la forme

$$R_X(\tau, t) = R_1(t)R_2(\tau) .$$

où R_1 est une fonction positive et R_2 est une fonction d'autocovariance. Le spectre instantané de tels processus, défini au sens la formule (1.12), s'écrit donc

$$S_X(\xi, t) = R_1(t)\hat{R}_2(\xi) .$$

Cette expression de l'autocovariance instantanée va permettre d'approcher correctement les processus non stationnaires dont la variance instantanée varie lentement par rapport aux variations temporelles du processus-lui même. La fonction R_1 modélise, en quelque sorte, la variance instantanée de X tandis que R_2 modélise l'autocovariance moyenne à horizon court (une fois la variance instantanée corrigée). Ce modèle est analogue à celui des processus localement modulés en amplitude définis plus tard (voir section 1.4.2.3).

Dans le même esprit, Priestley [61, 62] s'intéresse aux processus pour lesquels il existe une mesure positive η et des fonctions $\{a_t\}_{t \in \mathbb{R}}$ telles que l'autocovariance se décompose sous la forme

$$R_X(\tau, t) = \int_{\mathbb{R}} a_{t+\tau/2}(\xi) a_{t-\tau/2}(\xi) e^{2i\pi\xi\tau} d\eta(\xi), \quad (1.25)$$

où les fonctions a_t sont lentement changeantes en fonctions du temps. Cela revient à dire que la transformée de Fourier $\int_{\mathbb{R}} a_t(\xi) e^{-2i\pi v t} dt$ est localisée autour de la fréquence nulle avec un maximum de son module en $v = 0$. Par analogie avec la formule de Wiener-Khintchine, le spectre instantané $S_X(\xi, t)$ associé à ce processus non stationnaire est donné par

$$S_X(\xi, t) d\xi = |a_t(\xi)|^2 d\eta(\xi). \quad (1.26)$$

L'inconvénient majeur de cette modélisation des processus non stationnaires est la non-unicité de la représentation donnée par l'équation (1.25). En effet, la fonction d'autocovariance d'un processus non stationnaire peut être décomposée grâce à différentes familles de fonction $\{a_t\}_t$. Priestley propose de sélectionner la famille dont la variation temporelle est la plus lente. Il quantifie les variations lentes en terme de concentration de la transformée de Fourier autour de l'origine : c'est cette quantité qui est maximisée.

Avec la même idée de représenter localement la fonction d'autocovariance de processus non stationnaires par celle de processus stationnaires, Mallat, Papanicolaou et Zhang [39] ont construit une classe de signaux non stationnaires pour lesquels l'autocovariance varie lentement par rapport au paramètre non stationnaire autour d'un point donné. On a alors l'approximation suivante autour du point u : pour tout $t \in \left[u - \frac{l(u)}{2}, u + \frac{l(u)}{2}\right]$, on a

$$R_X(\tau, t) \approx \begin{cases} C_X(\tau, u), & \text{si } |\tau| < \frac{l(u)}{2}, \\ 0 & \text{si } |\tau| \geq d(u), \end{cases}$$

avec $d(u) < \frac{l(u)}{2}$. Ici, $l(u)$ désigne la longueur de l'intervalle autour de u sur lequel le processus est localement stationnaire, et $d(u)$ désigne la longueur de décorrélation à partir de laquelle l'autocovariance s'annule approximativement. À partir de ce modèle, les auteurs montrent qu'on peut approcher le spectre instantané par un opérateur constant sur des rectangles temps-fréquence. Ainsi,

$$\forall (\xi, t) \in \left[v - \frac{1}{2d(u)}, v + \frac{1}{2d(u)}\right] \times \left[u - \frac{l(u)}{2}, u + \frac{l(u)}{2}\right],$$

on a :

$$S_X(\xi, t) = S_X(v, u).$$

Ces modèles, dont la définition s'appuie sur des formes particulières pour la fonction d'autocovariance, reposent sur une approximation locale des processus non stationnaires par des processus stationnaires. Ils permettent alors de définir un spectre variant dans le temps. Dans la suite, nous allons définir d'autres modèles de non-stationnarité qui reposent sur la déformation globale de processus stationnaires.

1.4.2 Des signaux localement déformés

Soit X un processus stationnaire de spectre \mathcal{S}_X donné. Dans cette partie, on suppose que le signal non stationnaire observé y provient d'un processus Y issu de la déformation de X par un opérateur de déformation \mathcal{T} brisant la stationnarité. Le modèle s'écrit alors, de façon générale, sous la forme :

$$Y = \mathcal{T}X. \quad (1.27)$$

Dans la suite, nous allons considérer un certain nombre de types d'opérateurs de déformation \mathcal{T} . Chacun des opérateurs va produire un modèle de processus non stationnaire adapté à la modélisation de certains signaux réels. Les modèles que nous présentons sont semi-paramétriques étant donné que les opérateurs de déformation sont paramétrés par une *fonction de déformation*.

1.4.2.1 Les processus localement dilatés

Les processus localement dilatés ont, à l'origine, été introduits et étudiés par Clerc et Mallat [15]. Ils sont définis comme suit.

Définition 1.16 (Processus localement dilaté). Soit X un processus stationnaire au sens large de densité spectrale \mathcal{S}_X . Soit $\gamma \in C^2(\mathbb{R})$ une fonction strictement croissante telle que :

$$\exists c_\gamma, C_\gamma \in \mathbb{R}_+^* : \quad \forall t \in \mathbb{R}, \quad c_\gamma \leq \gamma'(t) \leq C_\gamma. \quad (1.28)$$

Alors, on appelle processus localement dilaté, le processus Y tel que :

$$Y(t) = \mathcal{D}_\gamma X(t) = \sqrt{\gamma'(t)} X(\gamma(t)), \quad \forall t \in \mathbb{R}. \quad (1.29)$$

L'opération qui associe $\mathcal{D}_\gamma X$ à X est appelée déformation temporelle. Par suite, la fonction γ est alors appelée fonction de déformation temporelle (ou de changement d'horloge). De plus, sa dérivée γ' est appelée dilation instantanée.

La multiplication par $\sqrt{\gamma'(t)}$ permet de rendre l'opération de déformation temporelle unitaire sur $L^2(\mathbb{R})$, c'est-à-dire que les énergies (les normes $L^2(\mathbb{R})$) de x et de $\mathcal{D}_\gamma x$

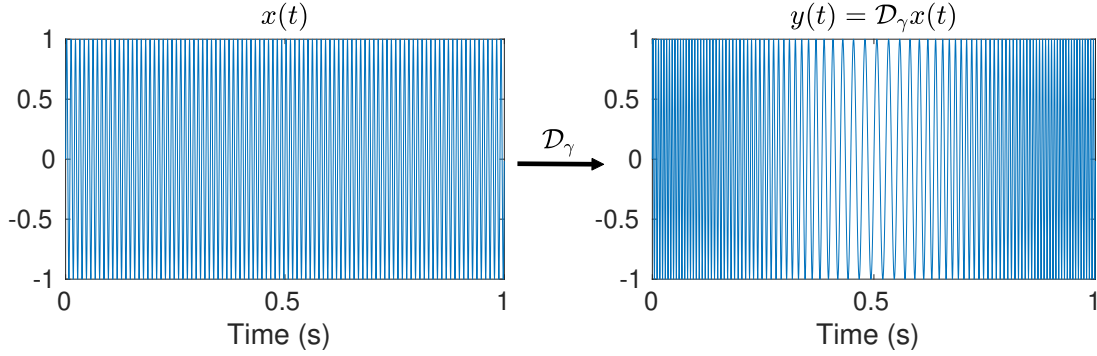


FIGURE 1.3 – Illustration de l'opération de déformation temporelle sur une sinusoïde.

sont égales pour tout $x \in L^2(\mathbb{R})$. La déformation temporelle a pour effet d'accélérer ou de ralentir les variations du signal. On illustre ce résultat sur la figure 1.3. Pour cela on applique l'opération de déformation temporelle (non normalisée par $\sqrt{\gamma'(t)}$) à une sinusoïde. Ainsi, lorsque $\gamma'(t) > 1$, les oscillations de la sinusoïde sont rapprochées : sa fréquence augmente, alors que, lorsque $\gamma'(t) < 1$, elles sont éloignées : sa fréquence diminue.

Les processus localement dilatés sont particulièrement pertinents pour la modélisation de nombreux signaux physiques. Par exemple, la déformation d'un signal stationnaire émis par un objet en mouvement (comme un signal audio ou électromagnétique) et perçu par un observateur fixe correspond à une dilatation locale. Il s'agit de l'effet Doppler : la fonction γ' associée dépend de la position et de la vitesse du mobile par rapport à l'observateur.

L'autocovariance instantanée de tels processus s'écrit

$$\begin{aligned} R_Y(\tau, t) &= \sqrt{\gamma'(t + \tau/2)\gamma'(t - \tau/2)} \mathbb{E} \left\{ X(\gamma(t + \tau/2)) \overline{X(\gamma(t - \tau/2))} \right\} \\ &= \sqrt{\gamma'(t + \tau/2)\gamma'(t - \tau/2)} k_X((\gamma(t + \tau/2)) - \gamma(t - \tau/2)) . \end{aligned}$$

Connaître le spectre instantané d'un tel processus non stationnaire revient donc à avoir la connaissance à la fois du spectre \mathcal{S}_X du processus stationnaire sous-jacent et de la fonction de déformation temporelle γ . Dans le chapitre 2, en approchant le spectre instantané au moyen de la transformée en ondelettes (et du scalogramme), nous verrons que l'on peut approcher cette covariance instantanée grâce à l'hypothèse de variations lentes de la fonction de dilatation instantanée. La covariance est alors donnée par :

$$R_Y(\tau, t) \approx \gamma'(t) k_X(\gamma'(t)\tau) .$$

Le spectre instantané associé à cette approximation est alors donné par

$$S_Y(\xi, t) \approx \mathcal{S}_X\left(\frac{\xi}{\gamma'(t)}\right) .$$

Remarque. Notons que lorsque la fonction de dilatation instantanée γ' est constante, le processus Y reste stationnaire. En effet, cette dernière condition est équivalente avec le fait d'avoir une fonction de déformation temporelle affine : $\gamma(t) = mt + p$. On a alors

$$R_Y(\tau, t) = m k_X(m\tau), \quad \forall \tau, t \in \mathbb{R}.$$

Comme la fonction d'autocovariance $R_Y(\tau, t)$ ne dépend que du paramètre τ , le processus Y est stationnaire.

1.4.2.2 Les processus localement modulés en fréquence

La modulation de fréquence est une opération classique dans les traitement des signaux en télécommunication.

Définition 1.17 (Processus localement modulé en fréquence). Soit X un processus stationnaire au sens large de densité spectrale \mathcal{S}_X . Soit $\phi \in C^2(\mathbb{R})$ telle que :

$$\exists c_\phi, C_\phi \in \mathbb{R}_+^* : \quad \forall t \in \mathbb{R}, \quad c_\phi \leq \phi'(t) \leq C_\phi. \quad (1.30)$$

Alors, on appelle processus localement modulé en fréquence, le processus Y tel que :

$$Y(t) = \mathcal{M}_\phi X(t) = e^{2i\pi\phi(t)} X(t), \quad \forall t \in \mathbb{R}. \quad (1.31)$$

L'opération qui associe $\mathcal{M}_\phi X$ à X est appelée modulation de fréquence instantanée. La fonction ϕ est alors appelée phase instantanée. Sa dérivée ϕ' est appelée fréquence instantanée de la modulante.

La modulation de fréquence est unitaire sur $L^2(\mathbb{R})$. Dans le cas où le processus stationnaire X est réel, on conserve un processus localement déformé réel en prenant la partie réelle du résultat obtenu via l'équation (1.31). On a donc :

$$Y(t) = \operatorname{Re} \left(e^{2i\pi\phi(t)} X(t) \right) = \cos(2\pi\phi(t)) X(t). \quad (1.32)$$

L'opération ainsi obtenue n'est plus unitaire sur $L^2(\mathbb{R})$.

L'effet visuel de la modulation de fréquence sur les chronogrammes des signaux est sensiblement similaire à celui de la déformation temporelle. En particulier, une sinusoïde (comme sur la figure 1.3) sera déformée de la même façon (à la normalisation près). La mise en évidence visuelle de la différence entre ces deux types de déformations sera présentée dans le chapitre 2 avec les outils d'analyse associés. Ce modèle s'applique davantage aux signaux artificiellement modulés en fréquence pour la transmission d'information (par exemple : le son modulé en fréquence dans les ondes radio).

L'autocovariance instantanée de tels processus s'écrit

$$R_Y(\tau, t) = e^{2i\pi(\phi(t+\tau/2) - \phi(t-\tau/2))} k_X(\tau).$$

Connaître le spectre instantané d'un tel processus non stationnaire revient donc à avoir la connaissance à la fois du spectre \mathcal{S}_X du processus stationnaire sous-jacent et de la fonction de phase instantanée ϕ (ou sa dérivée). En approchant le spectre instantané grâce à la transformée de Fourier à court terme (et au spectrogramme), on peut montrer [51] que, sous l'hypothèse de variations lentes de la fonction de dilatation instantanée par rapport à la fenêtre d'analyse, on peut approcher cette covariance instantanée par :

$$R_Y(\tau, t) \approx e^{2i\pi\phi'(t)\tau} k_X(\tau) .$$

Le spectre instantané associé à cette approximation est alors donné par

$$S_Y(\xi, t) \approx \mathcal{S}_X(\xi - \phi'(t)) .$$

Remarque. On peut noter que si la fréquence instantanée ϕ' est constante, le processus Y reste stationnaire. En effet, cette dernière condition est équivalente avec une fonction de déformation temporelle affine : $\phi(t) = mt + p$. L'autocovariance instantanée s'exprime alors

$$R_Y(\tau, t) = e^{2i\pi m\tau} k_X(\tau) , \quad \forall t, \tau \in \mathbb{R} .$$

Comme la fonction d'autocovariance $R_Y(\tau, t)$ ne dépend que du paramètre τ , le processus Y est stationnaire.

1.4.2.3 Les processus localement modulés en amplitude

Comme on l'a vu précédemment, la modulation en amplitude est une opération très classique qui brise la stationnarité. Dans [61], ces processus sont appelés « processus uniformément modulés ».

Définition 1.18 (Processus localement modulé en amplitude). Soit X un processus stationnaire au sens large de densité spectrale \mathcal{S}_X . Soit $\alpha \in C^1(\mathbb{R})$ telle que :

$$\exists c_\alpha, C_\alpha \in \mathbb{R}_+^* : \quad \forall t \in \mathbb{R}, \quad c_\alpha \leq \alpha(t) \leq C_\alpha . \quad (1.33)$$

Alors, on appelle processus localement modulé en amplitude, le processus Y tel que :

$$Y(t) = \mathcal{A}_\alpha X(t) = \alpha(t) X(t) , \quad \forall t \in \mathbb{R}. \quad (1.34)$$

La fonction α est alors appelée amplitude instantanée de la modulante.

Contrairement aux modèles d'opérateurs de déformation précédents, la modulation d'amplitude n'est pas unitaire sur $L^2(\mathbb{R})$. La figure 1.1 illustre cette déformation appliquée à un bruit blanc gaussien. Ce modèle peut décrire correctement de nombreuses situations physiques. Par exemple, en audio, le cas d'un son stationnaire dont l'intensité varie lentement.

L'autocovariance instantanée de tels processus s'écrit

$$R_Y(\tau, t) = \alpha(t + \tau/2) \alpha(t - \tau/2) k_X(\tau) .$$

Connaître le spectre instantané d'un tel processus non stationnaire revient donc à avoir la connaissance à la fois du spectre \mathcal{S}_X du processus stationnaire sous-jacent et de la fonction d'amplitude instantanée α . En approchant le spectre instantané grâce aux méthodes d'analyse locales (scalogramme ou spectrogramme), on peut montrer (voir chapitre 2) que, sous l'hypothèse de variations lentes de la fonction de modulation d'amplitude instantanée par rapport à la fenêtre (ou l'ondelette) d'analyse, on approche cette covariance instantanée par :

$$R_Y(\tau, t) \approx \alpha(t)^2 k_X(\tau) .$$

Le spectre instantané associé à cette approximation s'écrit

$$S_Y(\xi, t) \approx \alpha(t)^2 \mathcal{S}_X(\xi) .$$

En comparant cette dernière expression à l'équation (1.26), on remarque que les processus localement modulés en amplitude représentent un cas particulier du modèle (1.25) proposé par Priestley où $a_t(\xi) = \alpha(t)$ pour tout ξ .

Remarque. On peut noter que si l'amplitude instantanée est constante $\alpha(t) = A$, le processus Y reste stationnaire.

1.4.2.4 D'autres modèles de processus localement déformés

Wisdom *et al.* [81, 80] considèrent un modèle intermédiaire entre les processus localement modulés en amplitude et ceux localement modulés en fréquence. En effet, ils considèrent le cas où des processus non stationnaires Y s'écrivent

$$Y(t) = m(t)X(t) \quad \text{avec} \quad m(t) = \sum_{k=1}^K \alpha_k e^{2i\pi k \phi_0(t)} , \quad \forall t \in \mathbb{R},$$

où X est un processus complexe stationnaire au sens large, $\alpha_k \in \mathbb{C}$ et $\phi_0'(t)$ est une fonction variant lentement. On peut donc voir le signal non stationnaire ainsi obtenu comme un signal localement modulé en amplitude dont la fonction modulante est périodique (une somme de sinusoides de fréquences lentement variables). On peut également le voir comme un signal localement modulé en fréquence par une fonction harmonique de fréquence lentement variable, c'est-à-dire une somme de K sinusoides de fréquences $k\phi_0'$, $k \in \{1, \dots, K\}$.

Cet exemple nous montre que les différentes définitions de non-stationnarités proposées ci-dessus ne sont pas exclusives. Un signal non stationnaire pourra donc être modélisé à la fois comme réalisation de l'un ou l'autre des processus non stationnaires présentés ci-dessus. L'approximation du spectre instantané associée à ce modèle de non-stationnarité sera donc dépendante du modèle retenu. Dans le cas précédent, les deux approximations suivantes pourraient être valables :

- si on privilégie la modulation d'amplitude

$$S_Y(\xi, t) \approx |m(t)|^2 \mathcal{S}_X(\xi) ;$$

- si on privilégie la modulation de fréquence

$$S_Y(\xi, t) \approx \sum_{k=1}^K \alpha_k \mathcal{S}_X(\xi - k\phi_0(t)) .$$

Lors du choix du modèle mathématique, la pertinence physique (acoustique, mécanique, etc.) du modèle retenu doit donc être questionnée.

1.4.3 Des signaux localement harmoniques

Ce modèle a été proposé pour la première fois par McAulay et Quatieri [42] pour décrire des signaux de paroles. Il diffère des précédents modèles car, contrairement à ceux-ci, la construction des signaux non stationnaires localement harmoniques s'appuie sur des processus stationnaires possédant des spectres particuliers : les spectres de raies.

Le signal stationnaire sous-jacent à ce modèle est un signal périodique, somme de sinusoides. Il s'écrit donc

$$x(t) = \sum_{k=1}^K a_k \cos(2\pi\nu_k t) .$$

avec $A_k > 0$, $\nu_k > 0$, $\forall k \in \{1, \dots, K\}$. Les modèles décrits dans cette branche sont déterministes, la stationnarité n'a donc pas de définition précise dans ce cadre. Cependant, on peut voir le signal x donné par l'équation précédente comme étant une réalisation d'un processus stochastique stationnaire X tel que

$$X(t) = \sum_{k=1}^K a_k \cos(2\pi\nu_k t + \varphi_k) , \quad (1.35)$$

où les φ_k sont K variables aléatoires indépendantes et uniformes sur $[0, 2\pi]$. La densité spectrale associée à ce processus est alors donnée, au sens des distributions, par :

$$\mathcal{S}_X(\xi) = \frac{1}{4} \sum_{k=1}^K a_k^2 (\delta(\xi - \nu_k) + \delta(\xi + \nu_k)) .$$

Le modèle de non-stationnarité associé consiste à moduler chacune des sinusoides en amplitude et en fréquence (voir [3]). Le processus non stationnaire résultant Y s'écrit donc

$$Y(t) = \sum_{k=1}^K A_k(t) \cos(2\pi\phi_k(t) + \varphi_k) , \quad (1.36)$$

avec $\phi'_k(t) > 0$ et $A_k(t) > 0$, $\forall t \in \mathbb{R}$. On voit donc les signaux localement harmoniques comme une somme de sinusoides de fréquences et d'amplitudes variant dans le temps.

Cependant, cette définition n'est pas assez restrictive. En effet, à un signal donné, on pourra associer différentes combinaisons de fonction A_k et ϕ_k . Afin de garantir l'unicité du modèle, on fixe certaines contraintes sur les variations de l'amplitude et de la fréquence. Dans un cadre déterministe, différents types de contraintes peuvent être choisis. On s'appuie ici sur la définition donnée par Daubechies *et al.* [18] où ces signaux sont dénommés « *Superposition of Well-Separated Intrinsic Mode Components* ». Ici, les signaux localement harmoniques sont définis comme suit.

Définition 1.19 (Signaux localement harmoniques). Soit $K \geq 1$ un entier. Un signal localement harmonique est un signal s'écrivant sous la forme :

$$y(t) = \sum_{k=1}^K A_k(t) \cos(2\pi\phi_k(t)) , \quad (1.37)$$

où les amplitudes instantanées $A_k \in C^1(\mathbb{R})$ et les fréquences instantanées $\phi'_k \in C^1(\mathbb{R})$ satisfont les hypothèses suivantes :

1. Il existe des constantes c_1 et c_2 satisfaisant $0 < c_1 < c_2 < \infty$ telles que pour tout $t \in \mathbb{R}$:

$$c_1 \leq A_k(t) \leq c_2 , \quad (1.38)$$

$$c_1 \leq \phi'_k(t) \leq c_2 . \quad (1.39)$$

2. Il existe une constante $\epsilon \ll 1$ satisfaisant $0 \leq \epsilon < c_2$ telle que pour tout $t \in \mathbb{R}$:

$$|A'_k(t)| \leq \epsilon \phi'_k(t) , \quad (1.40)$$

$$|\phi''_k(t)| \leq \epsilon \phi'_k(t) . \quad (1.41)$$

3. Il existe une constante $d > 0$ telle que pour tout $t \in \mathbb{R}$:

$$\forall k \in \{1, \dots, K-1\}, \quad \phi'_{k+1}(t) - \phi'_k(t) > d . \quad (1.42)$$

La condition de séparation (1.42) empêche les croisements entre les différentes fréquences instantanées ϕ'_k . Combinée avec les conditions de régularité, elle permet d'approcher l'unicité de ce modèle. Autrement dit, si un signal localement harmonique y peut être décrit par deux combinaisons de K composantes sinusoïdales soit :

$$y(t) = \sum_{k=1}^K A_k \cos(2\pi\phi_k(t)) = \sum_{k=1}^K A_k \cos(2\pi\varphi_k(t)) ,$$

alors les différences $|A_k(t) - a_k(t)|$ et $|\phi'_k(t) - \varphi'_k(t)|$ tendent vers zéro lorsque ϵ tend vers zéro (cette condition est vraie pour d suffisamment grand, voir [18]).

On peut faire le lien entre ce modèle et les modèles de signaux localement dilatés et modulés en amplitude. En fait, chaque composante du processus stationnaire X défini en (1.35) est localement dilatée par la fonction γ_k et modulée en amplitude par la fonction α_k telles que

$$\gamma_k(t) = \frac{\phi_k(t)}{v_k},$$

$$\alpha_k(t) = \frac{A_k(t)}{a_k}.$$

Le spectre instantané correspondant serait alors approché de la manière suivante :

$$S_X(\xi, t) \approx \frac{1}{4} \sum_{k=1}^K A_k(t)^2 (\delta(\xi - \phi'_k(t)) + \delta(\xi + \phi'_k(t))) . \quad (1.43)$$

Connaître entièrement le spectre de tels processus nécessite donc la connaissance de l'ensemble des fonctions de modulation en amplitude A_k et en fréquence ϕ'_k . Nous développerons, dans le chapitre 5, une méthode permettant d'estimer le spectre instantané de tels processus.

1.5 Guide de lecture

Dans ce chapitre, nous avons passé en revue un certain nombre de notions utiles pour l'étude des signaux non stationnaires que nous proposons dans les chapitres suivants. Nous avons vu que la densité spectrale des signaux stationnaires est un outil bien défini, et son estimation est bien maîtrisée grâce à des procédures standards comme la méthode de Welch. En revanche, l'analyse spectrale n'est pas directe dans le cas des signaux non stationnaires car, en règle générale, il n'existe pas de quantité qualifiable de *densité de spectrale de puissance instantanée*, c'est-à-dire possédant toutes les propriétés désirées d'une densité spectrale : positivité, covariance par décalage en temps et en fréquence, localisation temps-fréquence, entre autres. C'est pourquoi plusieurs auteurs ont proposé des modèles de signaux localement stationnaires. En effet, ces modèles permettent de construire des classes de signaux non stationnaires pour lesquels le spectre instantané possède approximativement les qualités souhaitées.

Dans ce manuscrit, nous nous focaliserons sur d'autres classes de signaux non stationnaires dont l'interprétation physique est plus aisée que la classe précédente, tout en conservant la possibilité de le associer de manière naturelle une densité spectrale instantanée. Il s'agit d'abord des signaux localement déformés dont nous avons vu plusieurs sous-classes (signaux localement dilatés, modulés en amplitude ou en fréquence, etc.). Ces signaux seront étudiés dans les chapitres 2, 3 et 4. Nous développerons une méthode d'analyse spectrale appropriée aux signaux localement dilatés et modulés en amplitude (chapitre 2), que nous étendrons au cas multi-capteurs lorsque la modulation d'amplitude est matricielle (chapitre 3). Dans le chapitre 4, nous présenterons une méthode de « synthèse » spectrale adaptée aux signaux localement dilatés. L'autre classe

de signaux non stationnaires que l'on étudiera est celle des signaux localement harmoniques. Dans le chapitre 5, nous développerons un outil de synthèse spectrale construit pour être ajusté à cette classe de signaux.

Chapitre 2

ANALYSE DE SIGNAUX LOCALEMENT DÉFORMÉS

2.1 Introduction

Dans ce chapitre, nous nous concentrons l'analyse des signaux localement déformés. Nous présentons des méthodes d'estimations de spectres instantanés à partir de réalisations de processus non stationnaires. Prenons, par exemple, le son produit par le vent lorsqu'il souffle à travers une ouverture (comme une porte entrouverte). Le son que nous analyserons est un signal audio de 12 secondes (l'enregistrement audio complet est disponible sur *Freesound*¹). Nous avons représenté sur la figure 2.1 la transformée en ondelettes d'un tel enregistrement. Nous verrons que le modèle de processus localement dilatés (et modulés en amplitude) est particulièrement adapté à la description de ce signal non stationnaire. Nous montrerons, dans ce chapitre, qu'on peut approcher le comportement de cette transformée en ondelettes afin d'en extraire les grandeurs caractérisant le spectre instantané associé. Ce modèle se trouve être adapté à la description de nombreux autres types de signaux audios (voir la section 2.5).

Dans ce qui suit, après avoir décrit le modèle de processus localement déformés que nous souhaitons analyser, nous donnons un théorème d'approximation permettant de caractériser le comportement des coefficients d'ondelettes de tels processus. Ensuite, nous présentons la procédure d'estimation et l'algorithme mis en œuvre permettant l'estimation des quantités qui déterminent le spectre instantané de ces processus : les fonctions de déformations et les spectres des processus stationnaires sous-jacents. Des résultats permettant de quantifier les performances de l'algorithme sont également présentés. L'algorithme est ensuite appliqué à un signal synthétique afin de comparer les résultats aux performances attendues. Enfin, nous analysons des signaux audios réels au moyen de notre algorithme. L'idée est alors d'approcher les sources non stationnaires par des signaux stationnaires par morceaux. Notons que dans ce chapitre, nous reprenons et développons les résultats que nous avons présentés dans [45] et [46].

1. <https://freesound.org/people/Bosk1/sounds/144083/>

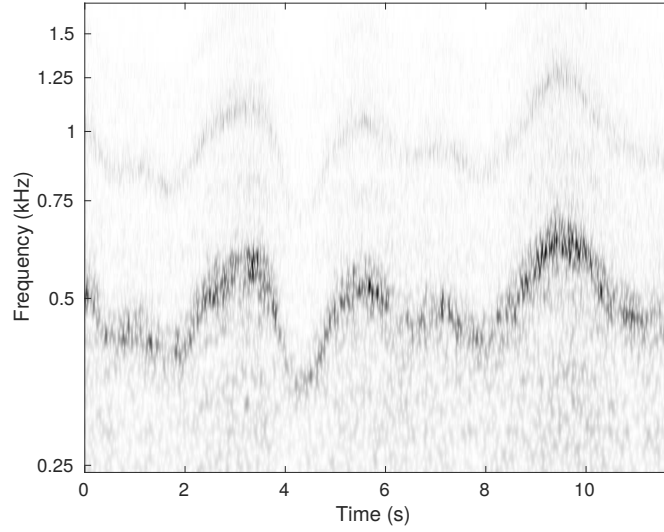


FIGURE 2.1 – Scalogramme du son produit par du vent.

2.2 Modèle et objectif

Le modèle de non-stationnarité pris en compte ici sera la combinaison d'une modulation d'amplitude et d'une déformation temporelle. C'est le type de non-stationnarité qui a été étudié dans [46]. Par ailleurs, dans [45], on a considéré le cas de la combinaison de la modulation de fréquence avec la déformation temporelle. Auparavant, Clerc et Mallat [14, 15] avaient proposé une méthode d'estimation spectrale pour les processus localement dilatés. Omer et Torrèsani [52, 51, 53] ont également étudié indépendamment l'analyse des processus localement dilatés, et celle des processus localement modulés en fréquence.

Soit y le signal non stationnaire que nous cherchons à analyser. On le modélise comme une réalisation du processus Y tel que

$$Y(t) = \mathcal{A}_\alpha \mathcal{D}_\gamma X(t), \quad \forall t \in \mathbb{R}, \quad (2.1)$$

où X est un processus stochastique réel, stationnaire au sens large de spectre \mathcal{S}_X et de moyenne nulle. Les opérations de déformation temporelle \mathcal{D}_γ et de modulation d'amplitude \mathcal{A}_α étant définies au chapitre précédent (cf. les définitions 1.16 et 1.18), les fonctions de dilatation instantanée γ' et d'amplitude instantanée α vérifient les conditions de régularité associées (respectivement données par les équations (1.28) et (1.33)). Nous verrons dans la partie 2.5 que ce modèle est adapté pour représenter divers types de signaux audios comme certaines vocalisations animales (par exemple des dauphins) ou les sons déformés par effet Doppler.

L'objectif de l'algorithme d'estimation que nous allons présenter sera donc d'estimer simultanément les fonctions de déformations γ et α , ainsi que le spectre \mathcal{S}_X du processus stationnaire sous-jacent. Pour ce faire, nous disposons d'une réalisation y du processus non stationnaire Y .

2.3 Approximation de la transformée en ondelettes

La transformée en ondelettes est un outil particulièrement adapté à l'analyse des processus localement dilatés car l'effet de la déformation temporelle est directement perceptible dans le plan temps-échelle comme un déplacement des coefficients d'ondelettes.

2.3.1 Première approche : une équation de transport

Clerc et Mallat [13, 15] ont mis en évidence le déplacement des coefficients d'ondelettes en établissant une équation de transport (approximativement) vérifiée par les coefficients du scalogramme des processus localement dilatés. Ils s'intéressent aux processus localement déformés de la forme :

$$Y(t) = X(\gamma(t)) ,$$

avec $\gamma \in C^3(\mathbb{R})$, $\gamma' > 0$ et X un processus stationnaire. Il s'agit des processus localement dilatés, mais, à la différence du modèle donné par la définition 1.16, l'opération de déformation temporelle n'est pas rendue unitaire sur $L^2(\mathbb{R})$.

Alors, on établit une équation aux dérivées partielles vérifiée par le scalogramme moyen A_Y défini par :

$$A_Y(s, \tau) = \mathbb{E} \{ |\mathcal{W}_Y(s, \tau)|^2 \} ,$$

En faisant une hypothèse sur la régularité de la fonction d'autocovariance de X , et sur la concentration de l'ondelette d'analyse (voir le théorème 3.1 dans [15] pour plus de détails sur les hypothèses), il est établi que pour tout $\tau \in \mathbb{R}$ tel que $\gamma''(\tau) \neq 0$, on a, lorsque $s \rightarrow -\infty$:

$$\partial_\tau A_Y(s, \tau) - (\ln \gamma')'(\tau) \partial_s A_Y(s, \tau) = \mathcal{O}(q^s) \partial_\tau A_Y(s, \tau) . \quad (2.2)$$

En négligeant le second membre aux petites échelles, l'équation vérifiée par le scalogramme du signal dilaté est une équation de transport. Cela signifie que la déformation temporelle d'un processus stationnaire induit un déplacement des coefficients du scalogramme dans le plan temps-échelle.

Lorsque l'équation de transport est valide, on peut obtenir un estimateur de la fonction de déformation temporelle. En effet, si l'on dispose de N réalisations $y_1(t)$, $y_2(t)$, ..., $y_N(t)$ du signal aléatoire, on peut estimer le scalogramme moyen A_Y grâce à la moyenne empirique :

$$\tilde{A}_Y(s, \tau) = \frac{1}{N} \sum_{n=1}^N |\mathcal{W}_{y_n}(s, \tau)|^2 .$$

Alors, lorsque l'équation de transport (2.2) est valide, un estimateur naturel du rapport $\frac{\gamma''(\tau)}{\gamma'(\tau)} = (\ln \gamma')'(\tau)$ est donné, pour une échelle fixé s_0 , par :

$$\widetilde{\frac{\gamma''(\tau)}{\gamma'(\tau)}} = \frac{\partial_\tau \tilde{A}_Y(s_0, \tau)}{\partial_s \tilde{A}_Y(s_0, \tau)} .$$

En pratique, pour obtenir un bon estimateur de la dilatation, on doit considérer N grand : de l'ordre de $N = 1000$. Or, nous nous intéressons ici au problème où l'on ne dispose que d'une réalisation du processus Y . On ne peut donc pas calculer l'estimateur empirique du scalogramme moyen. Les estimées des différentes dérivées du scalogramme sont alors obtenues par lissage par un noyau de convolution d'une réalisation du processus aléatoire $|\mathcal{W}_Y|^2$. Cette méthode a pour effet d'augmenter le biais de l'estimateur. Cette méthode d'estimation de la dilatation instantanée est donc limitée par certaines contraintes énumérées ci-dessous.

- L'estimation est valide lorsque $s \rightarrow -\infty$. Or, le théorème d'échantillonnage limite l'analyse à des échelles suffisamment grandes. De plus, lorsque $s \rightarrow -\infty$ on analyse les hautes fréquences contenues dans le signal. En pratique, ces fréquences sont souvent fortement corrompues par le bruit. On ne travaille donc pas aux petites échelles.
- De plus, le lissage d'une réalisation de $|\mathcal{W}_Y|^2$ a pour effet de biaiser l'estimateur. En effet, le lissage par un noyau de convolution agit sur le scalogramme comme un moyennage local par rapport au temps et à l'échelle. Or, le processus $|\mathcal{W}_Y|^2$ n'est pas stationnaire. On ne peut donc pas identifier ces moyennes statistiques avec ces moyennes temporelles (et d'échelles). Un noyau de convolution trop large aura alors pour effet d'augmenter le biais de l'estimateur tout en diminuant sa variance. Cette méthode oblige donc à faire un compromis entre biais et variance de l'estimateur. Il est cependant montré (théorème 4.1 dans [15]) que l'estimateur est consistant. Ainsi, si T est la longueur du signal y alors :

$$\lim_{T \rightarrow \infty} \mathbb{P} \left\{ \left| \frac{\widetilde{\gamma''(\tau)}}{\gamma'(\tau)} - \frac{\gamma''(\tau)}{\gamma'(\tau)} \right| \leq 2(\ln T)T^{-1/5} \right\} = 1.$$

Pour les raisons énoncées ci-dessus, cette approche n'est pas adaptée au problème que l'on se pose. Par ailleurs, cette méthode d'estimation exploite peu la stationnarité du processus X sous-jacent. En effet, elle ne cherche pas à déterminer les caractéristiques du processus stationnaire comme son spectre. Nous allons alors présenter une deuxième approche permettant de décrire le déplacement des coefficients d'ondelettes en s'appuyant sur la stationnarité de X .

2.3.2 Deuxième approche : un théorème d'approximation

Une deuxième approche, proposée par Omer et Torrèsani, consiste à approcher directement le comportement de la transformée en ondelettes en approchant localement les opérations de déformations. Dans [51], des théorèmes d'approximations dans les cas des processus localement dilatés ou localement modulés en fréquence. Nous donnons ici un théorème d'approximation des processus suivants le modèle (2.1). Celui-ci montre que l'effet de la déformation temporelle des signaux se caractérise par une translation des coefficients d'ondelettes dans le plan temps-échelle. Les ingrédients clés justifiant l'intérêt du théorème d'approximation sont la régularité des fonctions α et γ ,

et leurs variations lentes. Cela nous permet d'effectuer une analyse locale à l'aide de fonctions d'analyse régulières et localisées, grâce auxquelles \mathcal{A}_α et \mathcal{D}_γ peuvent être approchés par leurs opérateurs tangents $\widetilde{\mathcal{A}}_\alpha^\tau$ et $\widetilde{\mathcal{D}}_\gamma^\tau$ (voir [15, 52, 53, 51]). Étant donnée une fonction g localisée autour de $t = \tau$ (c'est-à-dire décroissant rapidement en fonction de $|t - \tau|$), Un développement de Taylor en $t = \tau$ conduit à

$$\mathcal{A}_\alpha g(t) \approx \widetilde{\mathcal{A}}_\alpha^\tau g(t), \quad \text{avec} \quad \widetilde{\mathcal{A}}_\alpha^\tau \triangleq \mathcal{A}_{\alpha(\tau)}, \quad (2.3)$$

$$\mathcal{D}_\gamma g(t) \approx \widetilde{\mathcal{D}}_\gamma^\tau g(t), \quad \text{avec} \quad \widetilde{\mathcal{D}}_\gamma^\tau \triangleq T_\tau D_{-\log_q(\gamma'(\tau))} T_{-\gamma(\tau)}. \quad (2.4)$$

Par conséquent, on approche la transformée en ondelettes de Y par

$$\mathcal{W}_Y(s, \tau) \approx \widetilde{\mathcal{W}}_Y(s, \tau) \triangleq \left\langle \widetilde{\mathcal{A}}_\alpha^\tau \widetilde{\mathcal{D}}_\gamma^\tau X, T_\tau D_s \psi \right\rangle.$$

La transformée en ondelettes approchée $\widetilde{\mathcal{W}}_Y$ s'exprime donc en fonction de celle du processus stationnaire X par

$$\widetilde{\mathcal{W}}_Y(s, \tau) = \alpha(\tau) \mathcal{W}_X \left(s + \log_q(\gamma'(\tau)), \gamma(\tau) \right). \quad (2.5)$$

Nous avons utilisé ici les règles de commutations des opérateurs élémentaires données dans la proposition 1.1. Le théorème 2.1 donne les conditions de validité de l'équation d'approximation (2.5).

Théorème 2.1 (Approximation). *Soit X un processus stochastique réel stationnaire au sens large de spectre \mathcal{S}_X et de moyenne nulle. Soit Y le processus non stationnaire défini par l'équation (2.1). Soit ψ une ondelette analytique, localisée autour de l'origine de sorte que $|\psi(t)| \leq 1/(1 + |t|^\beta)$ avec $\beta > 2$. Soit \mathcal{W}_Y la transformée en ondelettes de Y , $\widetilde{\mathcal{W}}_Y$ son approximation donnée par l'équation (2.5). Soit $\varepsilon = \mathcal{W}_Y - \widetilde{\mathcal{W}}_Y$ l'erreur d'approximation associée. On suppose que ψ et \mathcal{S}_X sont tels que*

$$I_X^{(\rho)} \triangleq \sqrt{\int_0^\infty \xi^{2\rho} \mathcal{S}_X(\xi) d\xi} < \infty, \quad \text{avec} \quad \rho = \frac{\beta - 1}{\beta + 2}.$$

Alors, l'erreur d'approximation ε est un champ aléatoire complexe, à deux dimensions, de moyenne nulle et dont la variance vérifie

$$\mathbb{E} \{ |\varepsilon(s, \tau)|^2 \} \leq C_\alpha^2 q^{3s} (K_1 \|\gamma''\|_\infty + K_2 q^{\mu s} \|\gamma''\|_\infty^\rho + K_3 \|\alpha'\|_\infty)^2$$

où

$$\begin{aligned} K_1 &= \frac{\beta \sigma_X}{2(\beta - 2) \sqrt{c_\gamma}}, & K_2 &= I_X^{(\rho)} \left(\frac{\pi}{2} \right)^\rho \frac{4 \sqrt{2C_\gamma}}{3\rho}, \\ K_3 &= \frac{\sqrt{C_\gamma} \beta \sigma_X}{(\beta - 2) c_a}, & \mu &= \frac{\beta - 4}{\beta + 2}, \end{aligned}$$

σ_X^2 désignant la variance de X .

La démonstration du théorème 2.1 est donnée en annexe A.1.1. Elle s'appuie principalement sur des développements de Taylor des fonctions de déformations autour de l'instant τ .

Intéressons-nous aux conditions de validité de la formule d'approximation. Pour que celle-ci soit valable, il faut que la variance de l'erreur d'approximation soit faible. Plusieurs conditions sont alors nécessaires.

- Dans le plan temps-échelle, la borne supérieure de la variance est indépendante du temps τ mais elle dépend de l'échelle s . Le terme q^{3s} nous indique que l'erreur reste faible pour les petites échelles mais qu'elle n'est plus contrôlée lorsque l'échelle est grande, c'est-à-dire aux basses fréquences. Par la suite, nous limiterons donc nos analyses en ondelettes à des valeurs d'échelles suffisamment petites pour que (2.5) reste valable.
- L'ondelette d'analyse doit être suffisamment concentrée autour de l'origine pour que les constantes K_1 et K_3 restent faibles. Pour borner K_1 et K_3 , la décroissance de l'ondelette à l'infini doit être suffisamment forte. En effet, quand $\beta \rightarrow 2$, $K_1 \rightarrow \infty$ et $K_3 \rightarrow \infty$.
- La constante μ doit rester positive pour éviter que le terme $q^{\mu s}$ soit trop grand aux petites échelles. Par construction, cette constante vérifie $-1/2 < \mu < 1$. La condition $\mu > 0$ revient alors à prendre $\beta > 4$.
- Pour que la constante K_2 reste faible, on doit s'assurer que l'intégrale $I_X^{(\rho)}$ reste bornée. On doit donc avoir une décroissance suffisamment rapide du spectre \mathcal{S}_X aux hautes fréquences. Or, comme $\beta > 2$, on a $1/4 < \rho < 1$. Dans le cas le plus défavorable, pour que l'intégrale converge, on doit au moins avoir $\mathcal{S}_X(\xi) = o(\xi^{-3})$ lorsque $\xi \rightarrow \infty$.
- Les termes en $\|\gamma'\|_\infty$ et $\|\gamma'\|_\infty^\rho$ nous imposent d'avoir une fonction de dilatation instantanée γ' dont les variations sont contrôlées. Dans le cas contraire, $\|\gamma'\|_\infty$ peut être grand et la variance de l'erreur n'est plus bornée raisonnablement. De la même manière, le troisième terme en $\|\alpha'\|_\infty$ impose que la modulation d'amplitude instantanée α varie suffisamment lentement.

Illustrons le théorème en simulant un signal jouet dont nous allons évaluer la transformée en ondelettes. Pour cela, on synthétise un signal x issu d'un processus stationnaire X de spectre \mathcal{S}_X constitué de trois lobes (des fenêtres de Hann) non chevauchants. Le signal non stationnaire y est obtenu en appliquant une fonction de dilatation instantanée γ' sinusoïdale à x . Le spectre \mathcal{S}_X , ainsi que les scalogrammes de x et y sont représentés figure 2.2. Alors que la transformée en ondelettes \mathcal{W}_x du signal stationnaire est composée de trois bandes horizontales correspondant aux trois lobes du spectre de X , la transformée en ondelettes \mathcal{W}_y du signal non stationnaire y est composée de trois bandes dont l'allure décrit la courbe de $\log_q(\gamma')$. Pour obtenir \mathcal{W}_y , les coefficients d'ondelettes de \mathcal{W}_x semblent donc avoir été translatés verticalement (c'est-à-dire le long de l'axe des échelles). En réalité, le déplacement des coefficients d'ondelettes est plus complexe qu'une simple translation verticale. Effectivement, la formule (2.5) nous montre que l'effet de la déformation temporelle est également visible le long de l'axe temporel du fait de la composition par γ . Ce déplacement horizontal n'est pas visible directement

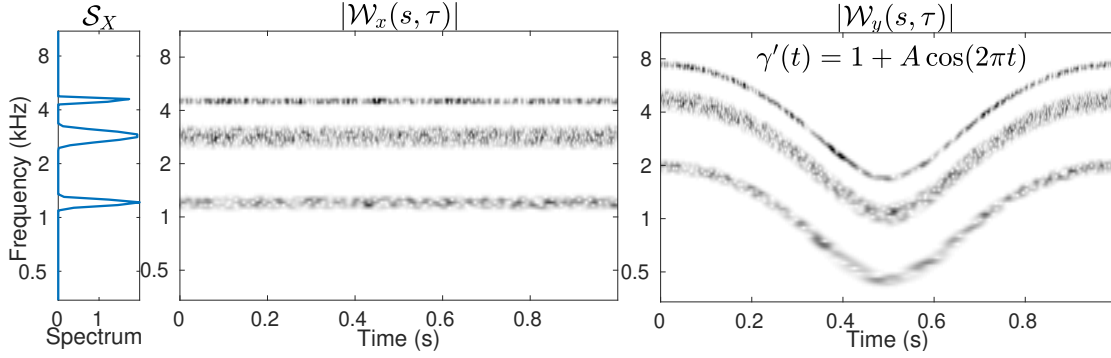


FIGURE 2.2 – Illustration du théorème d'approximation. À gauche : spectre \mathcal{S}_X . Au milieu : scalogramme du signal stationnaire x . À droite : scalogramme du signal non stationnaire y .

sur les scalogrammes en raison de la stationnarité de X qui entraîne une invariance par translation horizontale des propriétés statistiques de sa transformée en ondelettes.

La procédure d'estimation jointe repose sur cette formule d'approximation. Dans la suite, nous supposons donc que cette approximation est suffisamment précise pour pouvoir être exploitée.

2.4 Estimation et algorithme

2.4.1 Procédure d'estimation

Nous disposons dorénavant de la transformée en ondelettes \mathcal{W}_Y du signal observé y et on suppose qu'elle vérifie $\mathcal{W}_Y = \widetilde{\mathcal{W}}_Y$. L'objectif est alors d'estimer conjointement les fonctions de déformations γ et α , et le spectre \mathcal{S}_X du processus stationnaire sous-jacent. En vue de mettre en œuvre des méthodes d'estimation statistique, nous avons besoin de connaître la loi de probabilité de la transformée en ondelettes des observations.

Proposition 2.1 (Loi de la transformée en ondelettes). *Soit X un processus stochastique réel, stationnaire au sens large, de moyenne nulle et de spectre \mathcal{S}_X . Soit ψ une ondelette analytique. Soit $\widetilde{\mathcal{W}}_Y$ la transformée en ondelettes approchée définie par l'équation (2.5). Alors, cette dernière est un champ aléatoire à deux dimensions, de moyenne nulle et de covariance :*

$$\begin{aligned} \mathbb{E} \left\{ \widetilde{\mathcal{W}}_Y(s, \tau) \overline{\widetilde{\mathcal{W}}_Y(s', \tau')} \right\} &= C_Y(s, s', \tau, \tau') \\ &= \alpha(\tau) \alpha(\tau') q^{\frac{s+s'}{2}} \sqrt{\gamma'(\tau) \gamma'(\tau')} \\ &\quad \times \int_0^\infty \mathcal{S}_X(\xi) \widehat{\psi}(q^s \gamma(\tau) \xi) \overline{\widehat{\psi}(q^{s'} \gamma(\tau') \xi)} e^{2i\pi(\gamma(\tau) - \gamma(\tau'))} d\xi. \end{aligned} \quad (2.6)$$

D'autre part, si X est un processus stochastique gaussien réel, alors $\widetilde{\mathcal{W}}_Y$ est un champ aléatoire gaussien complexe circulaire.

La démonstration de cette proposition est donnée en annexe A.1.2.

Dorénavant, nous faisons l'hypothèse supplémentaire que X est un processus stochastique gaussien réel. Par conséquent, la proposition ci-dessus nous montre que la densité de probabilité de $\widetilde{\mathcal{W}}_Y$ est caractérisée uniquement par sa fonction de covariance donnée par (2.6). Toutefois, cette expression montre qu'en plus des fonctions de déformation, la covariance dépend également du spectre \mathcal{S}_X du signal stationnaire sous-jacent X , qui est lui aussi inconnu. C'est pourquoi une évaluation par une méthode d'estimation statistique (comme le maximum de vraisemblance) des fonctions α et γ nécessite d'avoir une estimation préalable de \mathcal{S}_X . Cette contrainte nous conduit naturellement à opter pour un algorithme alterné comme stratégie d'estimation. Dans [45], une estimée de \mathcal{S}_X est obtenue à chaque itération en appliquant la méthode de Welch au signal « stationnarisé » $\mathcal{A}_{\tilde{\alpha}}^{-1} \mathcal{D}_{\tilde{\gamma}}^{-1} Y$, $\tilde{\alpha}$ et $\tilde{\gamma}$ étant les estimées courantes des fonctions α et γ . Nous mettrons en œuvre ici une estimée plus simple, dans la mesure où elle s'obtient directement à partir de la transformée en ondelettes des observations. Les deux étapes de l'estimation alternée sont détaillées dans ce qui suit.

La stratégie de maximisation alternée de la vraisemblance se rapproche de la structure de l'algorithme *Espérance-Maximisation* (EM, introduit dans [24]) où le spectre de puissance serait le paramètre caché. Cependant, s'il serait souhaitable de se placer directement notre problème dans le cadre de l'algorithme EM (dont la convergence est prouvée), la dimensionnalité de ce dernier, et la taille correspondante des matrices de covariance, nous obligent à faire des simplifications supplémentaires qui s'écartent du schéma EM. Nous nous tournons donc vers une approche plus simple avec plusieurs étapes de réduction des dimensions. Dans le chapitre 4, le point de vue « à la synthèse » nous permettra une adaptation de l'algorithme que nous allons présenter, qui se placera dans le cadre de l'algorithme EM. La procédure d'estimation consiste donc en une estimation alternée des déformation et du spectre.

(a) Estimation des déformations. Supposons que la densité spectrale \mathcal{S}_X soit connue (en réalité, seulement une estimée $\hat{\mathcal{S}}_X$ de celle-ci est connue). Nous pouvons alors écrire la vraisemblance des coefficients d'ondelettes observés. L'estimateur du maximum de vraisemblance est alors mis en œuvre pour déterminer les fonctions inconnues γ et α .

En pratique, la transformée en ondelettes est calculée sur une grille temps-échelle régulière $\Lambda = \mathbf{s} \times \tau$. On note δ_s le pas d'échantillonnage des échelles et F_s la fréquence d'échantillonnage en temps. Les tailles des vecteurs \mathbf{s} et τ sont notées respectivement M_s et N_τ .

Si l'on considère l'expression de la covariance (2.6), nous devons estimer simultanément le vecteur de paramètres

$$\Theta = (\theta_1, \theta_2, \theta_3) \triangleq \left(\alpha(\tau)^2, \log_q(\gamma'(\tau)), \gamma(\tau) \right), \quad (2.7)$$

ce qui donne un vecteur de paramètres de taille $3N_\tau$. Notons par $\mathbf{W}_y = \widetilde{\mathcal{W}}_y(\boldsymbol{\Lambda})$ la transformée en ondelettes discrétisée, et notons par $\mathbf{C}_\mathbf{W}(\boldsymbol{\Theta})$ la matrice de covariance correspondante. La log-vraisemblance associée s'écrit

$$\begin{aligned} \mathcal{L}(\boldsymbol{\Theta}) &\triangleq \ln(p_{\mathbf{W}_Y}(\mathbf{W}_y; \boldsymbol{\Theta})) + M_s N_\tau \ln(\pi) \\ &= -\ln|\det(\mathbf{C}_\mathbf{W}(\boldsymbol{\Theta}))| - \left\langle \mathbf{C}_\mathbf{W}(\boldsymbol{\Theta})^{-1} \mathbf{W}_y, \mathbf{W}_y \right\rangle. \end{aligned} \quad (2.8)$$

La matrice $\mathbf{C}_\mathbf{W}(\boldsymbol{\Theta})$ est une matrice de taille $M_s N_\tau \times M_s N_\tau$, qui est généralement grande. Par exemple, pour un signal de 1 seconde, échantillonné à la fréquence $F_s = 44,1$ kHz, et si la transformée en ondelettes est calculée sur un vecteur d'échelles de longueur 100, la matrice $\mathbf{C}_\mathbf{W}(\boldsymbol{\Theta})$ possède environ 2.10^{13} éléments ce qui rend le problème impossible à traiter numériquement. Par ailleurs, en raison de la redondance de la transformée en ondelettes, $\mathbf{C}_\mathbf{W}(\boldsymbol{\Theta})$ se révèle être singulière et le calcul de la log-vraisemblance est alors impossible.

Pour surmonter ces problèmes de dimension, nous utilisons une régularisation bloc-diagonale de la matrice de covariance, obtenue en forçant à zéro les éléments correspondant à des instants distincts. En d'autres termes, nous ne tenons pas compte des corrélations temporelles dans le domaine des ondelettes. La covariance régularisée devient alors :

$$\mathbb{E} \left\{ \widetilde{\mathcal{W}}_Y(s, \tau) \overline{\widetilde{\mathcal{W}}_Y(s', \tau')} \right\} = C(s, s', \tau, \tau') \mathbb{1}_\tau(\tau').$$

Cela revient à considérer les vecteurs $\mathbf{w}_{y, \tau_n} = \widetilde{\mathcal{W}}_y(\mathbf{s}, \tau_n)$ où $\tau_n = \tau(n)$ comme gaussiens complexes circulaires, indépendants, de moyenne nulle et de matrice de covariance de taille $M_s \times M_s$ donnée par

$$\mathbf{C}(\boldsymbol{\Theta}_n) = \theta_{n,1} \mathbf{C}_0(\theta_{n,2}), \quad (2.9)$$

où

$$[\mathbf{C}_0(\theta_{n,2})]_{ij} = q^{(s_i + s_j + 2\theta_{n,2})/2} \int_0^\infty \mathcal{S}_X(\xi) \bar{\psi}(q^{s_i + \theta_{n,2}} \xi) \hat{\psi}(q^{s_j + \theta_{n,2}} \xi) d\xi \quad (2.10)$$

$$= q^{(s_i + s_j)/2} \int_0^\infty \mathcal{S}_X(q^{-\theta_{n,2}} \xi) \bar{\psi}(q^{s_i} \xi) \hat{\psi}(q^{s_j} \xi) d\xi, \quad 1 \leq i, j \leq M_s, \quad (2.11)$$

avec $\boldsymbol{\Theta}_n = (\theta_{n,1}, \theta_{n,2}) \triangleq (\theta_1(n), \theta_2(n))$ correspondant aux paramètres de modulation d'amplitude et de dilatation instantanée à l'instant donné τ_n . Dans cette situation, la log-vraisemblance régularisée \mathcal{L}^r se scinde en une somme de termes indépendants

$$\mathcal{L}^r(\boldsymbol{\Theta}) = \sum_{n=1}^{N_\tau} \mathcal{L}(\boldsymbol{\Theta}_n).$$

Notons que, dans un tel formalisme, $\theta_{n,3} = \gamma(\tau_n)$ n'apparaît plus dans l'expression de la covariance. Nous sommes alors amenés à estimer indépendamment, à chaque instant

τ_n , un vecteur à deux paramètres maximisant la log-vraisemblance de la n^e colonne de la transformée en ondelettes discrétisée \mathbf{W}_y . Cette fonction est donnée par :

$$\mathcal{L}(\Theta_n) = -\ln |\det(\mathbf{C}(\Theta_n))| - \mathbf{w}_{y,\tau_n}^H \mathbf{C}(\Theta_n)^{-1} \mathbf{w}_{y,\tau_n} . \quad (2.12)$$

Par souci de simplicité, la procédure d'estimation est effectuée au moyen d'un algorithme itératif (détailé dans la partie 2.4.2), qui repose sur deux étapes principales.

1. D'abord, la log-vraisemblance est maximisée par rapport au paramètre de dilatation instantanée $\theta_{n,2}$ en fixant $\theta_{n,1}$. Comme on ne peut pas déterminer d'expression explicite de ce maximiseur, nous allons l'approcher grâce à une méthode de remontée de gradient. L'algorithme 2.1 résume les étapes permettant d'obtenir l'estimation $\tilde{\theta}_{n,2}$ de $\theta_{n,2}$. Dans cet algorithme, d désigne le pas de gradient, il est fixé. Le calcul effectif de la direction de descente requiert la connaissance de la dérivée partielle $\partial_{\theta_{n,2}} \mathcal{L}$ de \mathcal{L} par rapport au paramètre $\theta_{n,2}$. Son expression est donnée par :

$$\begin{aligned} \partial_{\theta_{n,2}} \mathcal{L}(\Theta_n) = & \text{Trace} \left(\mathbf{C}(\Theta_n)^{-1} \frac{\partial \mathbf{C}}{\partial \theta_{n,2}}(\Theta_n) \right) \\ & + \mathbf{w}_{y,\tau_n}^H \mathbf{C}(\Theta_n)^{-1} \frac{\partial \mathbf{C}}{\partial \theta_{n,2}}(\Theta_n) \mathbf{C}(\Theta_n)^{-1} \mathbf{w}_{y,\tau_n} , \end{aligned} \quad (2.13)$$

où l'expression de $\frac{\partial \mathbf{C}}{\partial \theta_{n,2}}$ s'obtient directement à partir de la dérivation de l'expression (2.10) par rapport à la variable $\theta_{n,2}$. Le résultat fait intervenir la dérivée de $\hat{\psi}$ que nous supposons connue (voir la section 2.5.1 pour l'expression de $\hat{\psi}$ utilisée en pratique).

Algorithme 2.1 Remontée de gradient. $\tilde{\theta}_{n,2} = \text{grad}(\theta_0, \theta_{n,1}, d)$

```

 $\theta^{(0)} = \theta_0$ 
 $q \leftarrow 0$ 
tant que  $\frac{\theta^{(q+1)} - \theta^{(q)}}{\theta^{(q)}} > \Lambda_{\text{grad}}$  faire
     $\theta^{(q+1)} = \theta^{(q)} + d \partial_{\theta_{n,2}} \mathcal{L}(\theta_{n,1}, \theta^{(q)})$ 
     $q \leftarrow q + 1$ 
fin tant que
Sortie :  $\tilde{\theta}_{n,1} = \theta^{(q)}$ 

```

2. Ensuite, nous fixons $\theta_{n,2}$, on obtient alors directement l'estimée de $\theta_{n,1}$ en annulant la dérivée partielle de \mathcal{L} par rapport à $\theta_{n,1}$. On obtient alors :

$$\tilde{\theta}_{n,1} = \frac{1}{M_s} \mathbf{w}_{y,\tau_n}^H \mathbf{C}_0^{-1}(\theta_{n,2}) \mathbf{w}_{y,\tau_n} . \quad (2.14)$$

(b) Estimation de la densité spectrale. Supposons que les paramètres de modulation d'amplitude θ_1 et de dilatation instantanée θ_2 sont connus (en réalité, seuls leurs estimées $\tilde{\theta}_1$ et $\tilde{\theta}_2$ sont connues). Pour tout n , nous pouvons calculer la transformée en ondelettes « rectifiée »

$$\frac{1}{\theta_{n,1}^{1/2}} \widetilde{\mathcal{W}}_y(s - \theta_{n,2}, \tau_n) = \mathcal{W}_x(s, \gamma(\tau_n)) . \quad (2.15)$$

La rectification permet d'avoir un champs aléatoire dont la dépendance par rapport aux échelles est invariante par translation temporelle. Ainsi, pour une échelle fixée s_m , $\mathbf{w}_{x,s_m} \triangleq \mathcal{W}_x(s_m, \gamma(\tau)) \in \mathbb{C}^{N_\tau}$ est un vecteur aléatoire gaussien complexe circulaire de moyenne nulle dont la variance est indépendante du temps. Par conséquent, bien que l'on n'ait pas accès à la transformée en ondelettes d'un processus stationnaire, l'estimateur du spectre-ondelettes $\tilde{\mathcal{S}}_X$, défini dans l'équation (1.22), peut être mis en œuvre. Avec les notations de ce chapitre, cet estimateur s'écrit :

$$\tilde{\mathcal{S}}_X(q^{-s_m} \omega_0) = \frac{1}{N_\tau \|\psi\|_2^2} \|\mathbf{w}_{x,s_m}\|^2 . \quad (2.16)$$

Enfin, l'estimée $\tilde{\mathcal{S}}_X$ est étendue à toutes les fréquences $\zeta \in [0, F_s/2]$ par interpolation linéaire.

2.4.2 Algorithme

La procédure d'estimation est implémentée dans un algorithme d'optimisation alternée itératif. Cet algorithme dont le pseudo-code est fourni dans l'algorithme 2.2 est appelé JEFAS (*Joint Estimation of Frequency, Amplitude, and Spectrum*). L'initialisation requiert une première estimation pour la densité spectrale de puissance \mathcal{S}_X de X . Pour cela, nous utilisons l'estimateur de spectre (2.16) appliqué à la transformée en ondelettes des observations Y . Le passage de la procédure d'estimation que nous venons de détailler à la mise en œuvre numérique de l'algorithme d'estimation requiert quelques adaptations. Nous les donnons dans ce qui suit.

Expression de la log-vraisemblance. Après k itérations de l'algorithme, les estimées $\tilde{\Theta}_n^{(k)}$ et $\tilde{\mathcal{S}}_X^{(k)}$ de Θ_n et \mathcal{S}_X sont disponibles. Par conséquent, nous ne pouvons évaluer que l'estimation $\tilde{\mathbf{C}}_0^{(k)}$ de la matrice de covariance \mathbf{C}_0 , obtenue en remplaçant le spectre par son estimée dans l'expression de la matrice de covariance (2.11). Cela conduit à une expression approchée $\mathcal{L}^{(k)}$ de la log-vraisemblance, qui est utilisée à la place de \mathcal{L} dans (2.12) pour l'estimation du maximum de vraisemblance. L'influence de telles approximations sur les performances de l'algorithme est discutée en section 2.4.3.

Échantillonnage temporel. Afin d'accélérer la vitesse de l'algorithme, l'estimation des paramètres de déformation peut n'être faite que sur une grille de temps sous-échantillonnée. C'est pourquoi on effectue une estimation tous les D échantillons. Le

Algorithme 2.2 JEFAS (*Joint Estimation of Frequency, Amplitude and Spectrum*). $(\tilde{\alpha}, \tilde{\gamma}) = \text{JEFAS}(y, \mathbf{s}, D, p, \Lambda, k_{\max})$

- Calculer les transformées en ondelettes discrétisées de y : \mathbf{W}_y sur les échelles \mathbf{s} et $\mathbf{W}_{y,p}$ sur les échelles \mathbf{s}_p .

Initialisations :

- Grâce au spectre-ondelettes, évaluer une estimée $\tilde{\mathcal{S}}_Y$ de la densité spectrale de Y comme première estimée $\tilde{\mathcal{S}}_X^{(0)}$ de \mathcal{S}_X .
- Initialiser l'estimée de la modulation d'amplitude carrée avec $\tilde{\theta}_{n,1}^{(0)} = 1, \forall n$.
- $k \leftarrow 1$

tant que le critère d'arrêt (2.18) est faux **et** $k \leq k_{\max}$ **faire**

pour tout $n \in \{1, 1 + D, 1 + 2D, \dots, N_\tau\}$ **faire**

- À partir de $\mathbf{W}_{y,p}$: estimer le paramètre de dilatation instantanée $\tilde{\theta}_{n,2}^{(k+1)} = \text{grad}(\tilde{\theta}_{n,2}^{(k)}, \tilde{\theta}_{n,1}^{(k)}, d)$ (voir algorithme 2.1).
- À partir de \mathbf{W}_y : estimer le paramètre de modulation d'amplitude $\tilde{\theta}_{n,1}^{(k+1)}$ donné par la formule (2.14) en remplaçant $\theta_{n,2}$ par son estimée courante $\tilde{\theta}_{n,2}^{(k+1)}$.

fin pour

- Construire la transformée en ondelettes rectifiée qui est donnée par l'équation (2.15) à partir de $\tilde{\theta}^{(k)}$ et en interpolant \mathbf{W}_y sur les échelles \mathbf{s} .
- Estimer le spectre correspondant $\tilde{\mathcal{S}}_X^{(k+1)}$ au moyen du spectre-ondelettes donné par (2.16).
- $k \leftarrow k + 1$

fin tant que

- Calculer $\tilde{\alpha}$ et $\tilde{\gamma}$ par interpolation à partir de $\tilde{\Theta}^{(k)}$.
-

principal effet négatif de ce choix sur l'algorithme concerne l'estimation finale de α et γ qui est plus sensible à l'opération d'interpolation.

Échantillonnage de l'axe des échelles. Pour permettre le contrôle de la variance des estimateurs et du coût de calcul, deux discrétisations différentes de l'axe des échelles sont utilisées pour estimer $\tilde{\theta}_1$ et $\tilde{\theta}_2$. En effet, le calcul de la log-vraisemblance implique d'évaluer l'inverse de la matrice de covariance. Dans [51], une condition suffisante pour l'inversibilité est donnée (en présence de bruit). La conséquence majeure induite par cette condition est que lorsque δ_s se rapproche de zéro (c'est-à-dire que le pas d'échantillonnage des échelles est petit), la matrice de covariance peut ne pas être inversible numériquement. La discrétisation des échelles doit donc être suffisamment grossière pour assurer un bon conditionnement de cette matrice. Bien que cette condition puisse être raisonnablement remplie pour estimer le paramètre de dilatation instantanée $\theta_{n,2}$ sans nuire aux performances de l'estimateur, cette discrétisation ne peut pas être appliquée à l'estimation du paramètre de modulation d'amplitude $\theta_{n,1}$ à cause de l'influence de M_s sur la variance de son estimateur (voir l'expression de la borne de Cramér-Rao dans la section 2.4.3). Le choix que nous avons fait est de maximiser $\mathcal{L}(\Theta_n)$ par rapport à $\theta_{n,2}$ avec \mathbf{w}_{y,τ_n} correspondant à une échantillonnage grossier \mathbf{s}_p (qui est une version sous-échantillonnée du vecteur original \mathbf{s}). Le pas d'échantillonnage des échelles et la taille de \mathbf{s}_p sont notés respectivement $p\delta_s$ et $\lfloor M_s/p \rfloor$ pour $p \in \mathbb{N}^*$ donné. En revanche, $\mathcal{L}(\Theta_n)$ est maximisé par rapport à $\theta_{n,1}$ sur l'échantillonnage fin original \mathbf{s} . Une régularisation de la matrice de covariance doit alors être effectuée pour en assurer l'inversibilité. La matrice régularisée est construite en remplaçant la matrice de covariance $\mathbf{C}_0(\theta_{n,2})$ dans (2.11) par sa version régularisée $\mathbf{C}_{0,r}(\theta_{n,2})$, donnée par

$$\mathbf{C}_{0,r}(\theta_{n,2}) = (1 - r)\mathbf{C}_0(\theta_{n,2}) + r\mathbf{I}, \quad (2.17)$$

avec $0 \leq r \leq 1$.

Critère d'arrêt. Pour évaluer la convergence de l'algorithme, un seuil minimal sur la mise à jour relative des paramètres est fixé comme critère d'arrêt :

$$\frac{\|\tilde{\theta}_j^{(k)} - \tilde{\theta}_j^{(k-1)}\|_2^2}{\|\tilde{\theta}_j^{(k-1)}\|_2^2} < \Lambda, \text{ pour } j \in \{1, 2\}, \quad (2.18)$$

où $0 < \Lambda < 1$ désigne le seuil choisi.

Interpolation. Enfin, une fois que l'algorithme a convergé vers la valeur estimée $\tilde{\Theta}^{(k)}$, les fonctions $\log_q(\gamma')$ et α^2 sont estimées à chaque instant par interpolation par des splines cubiques. De plus, γ est obtenu par intégration numérique en fixant sa valeur initiale à $\gamma(0) = 0$.

Estimation du spectre Après convergence de l'algorithme, les fonctions estimées $\tilde{\alpha}$ et $\tilde{\gamma}$ nous permettent de construire un signal « stationnarisé »

$$\tilde{x} = \mathcal{D}_{\tilde{\gamma}^{-1}} \mathcal{A}_{\tilde{\alpha}^{-1}} y .$$

Le signal \tilde{x} constitue une estimation du signal stationnaire original x . Par conséquent, l'estimation spectrale par la méthode de Welch peut être calculée à partir de \tilde{x} pour obtenir une estimée $\mathcal{S}_{\tilde{x}}$ dont le biais ne dépend pas de la fréquence (contrairement à l'estimée du spectre-ondelettes utilisée dans l'algorithme JEFAS).

Dans la section suivante, nous évaluons des grandeurs (biais, variance) permettant de quantifier les performances attendues des estimateurs et leur influence sur la convergence de l'algorithme. Le lecteur qui n'est pas directement intéressé par les justifications statistiques de JEFAS peut sauter cette section et se rendre directement aux résultats numériques présentés dans la partie 2.5.

2.4.3 Performances des estimateurs et de l'algorithme

Pour quantifier les performances des estimateurs, on va s'intéresser à trois quantités : le biais, la variance et la robustesse au bruit.

(a) Biais. Concernant le paramètre de modulation d'amplitude $\theta_{n,1}$, l'estimateur est non biaisé lorsque les vraies valeurs de $\theta_{n,2}$ et \mathcal{S}_X sont connues. Dans notre cas, on ne possède qu'une estimation de ces grandeurs, le biais $b_{n,1}^{(k)}(\theta_{n,1}) = \mathbb{E} \left\{ \tilde{\theta}_{n,1}^{(k)} \right\} - \theta_{n,1}$ s'écrit alors

$$b_{n,1}^{(k)}(\theta_{n,1}) = \frac{\theta_{n,1}}{M_s} \text{Trace} \left\{ \tilde{\mathbf{C}}_0^{(k)} \left(\tilde{\theta}_{n,2}^{(k)} \right)^{-1} \mathbf{C}_0(\theta_{n,2}) - \mathbf{I} \right\} . \quad (2.19)$$

Comme l'on pouvait s'y attendre, le biais $|b_{n,1}^{(k)}|$ est d'autant plus faible que le produit matriciel $\tilde{\mathbf{C}}_0^{(k)} \left(\tilde{\theta}_{n,2}^{(k)} \right)^{-1} \mathbf{C}_0(\theta_{n,2})$ se rapproche de la matrice identité. Le biais dépend donc de la qualité de l'estimation de la matrice de covariance de la transformée en ondelettes à chaque instant.

Concernant le paramètre de dilatation instantanée $\theta_{n,2}$, dans la mesure où nous ne possédons d'expression explicite pour son estimateur nous ne sommes pas capables de donner une expression du biais associé. L'évaluation de la qualité de l'estimation de $\theta_{n,2}$ sera donc principalement vérifiée numériquement (voir la section 2.5). Néanmoins, si l'on suppose que les valeurs exactes des deux autres variables sont connues, comme l'estimateur $\tilde{\theta}_{n,2}$ est un estimateur du maximum de vraisemblance, nous nous assurons que l'estimateur est asymptotiquement non biaisé (c'est-à-dire $\tilde{\theta}_{n,2} \rightarrow \theta_{n,2}$ quand $M_s \rightarrow \infty$).

Concernant le spectre \mathcal{S}_X , l'estimateur du spectre-ondelettes introduit un biais dépendant de la fréquence. En effet, on a :

$$\begin{aligned}\mathcal{S}_{X,\psi}(q^{-s_m}v_0) &= \mathbb{E} \left\{ \tilde{\mathcal{S}}_X(q^{-s_m}v_0) \right\} \\ &= \frac{1}{\|\psi\|_2^2} \int_0^\infty \mathcal{S}_X(\xi) q^{s_m} |\hat{\psi}(q^{s_m}\xi)|^2 d\xi.\end{aligned}\quad (2.20)$$

Ceci montre que l'estimateur du spectre en donne une version lissée, donc biaisée. En réalité, un terme de biais supplémentaire est introduit par l'erreur d'estimation des paramètres de déformation. En effet, à chaque itération de JEFAS, l'estimation du spectre $\tilde{\mathcal{S}}_X^{(k)}$ s'obtient au moyen de la transformée en ondelettes rectifiée à partir des paramètres de déformation estimés. On a donc :

$$\tilde{\mathcal{S}}_X^{(k)}(q^{-s_m}\omega_0) = \frac{1}{N_\tau \|\psi\|_2^2} \sum_{n=1}^{N_\tau} \frac{1}{\tilde{\theta}_{n,1}^{(k)}} \left| \tilde{\mathcal{W}}_Y(s_m - \tilde{\theta}_{n,2}^{(k)}, \tau_n) \right|^2. \quad (2.21)$$

La proposition 2.2 ci-dessous montre que l'espérance de la densité spectrale de puissance estimée $\tilde{\mathcal{S}}_X^{(k)}$ converge vers le spectre-ondelettes $\mathcal{S}_{X,\psi}$ lorsque les paramètres de déformation convergent vers les vraies valeurs.

Proposition 2.2 (Biais de l'estimateur du spectre). *Soit $\psi \in H^2(\mathbb{R})$ une ondelette analytique telle que $\hat{\psi}$ est borné et dérivable de dérivée bornée, avec $|\hat{\psi}(u)| = o(u^{-2})$ quand $u \rightarrow \infty$. Soit φ_1 et φ_2 les fonctions bornées définies sur \mathbb{R}_+ par $\varphi_1(u) = u |\hat{\psi}(u)|^2$ et $\varphi_2(u) = u^2 |\hat{\psi}(u)|$. Supposons que le spectre \mathcal{S}_X est tel que*

$$J_X = \int_0^\infty \xi^{-1} \mathcal{S}_X(\xi) d\xi < \infty.$$

Soit $\mathcal{S}_X^{(k)}$ l'estimée du spectre après k itérations de l'algorithme JEFAS. Soit $b_{\mathcal{S}_X}^{(k)}$ le biais défini pour tout $m \in \llbracket 1, M_s \rrbracket$ par

$$b_{\mathcal{S}_X}^{(k)}(m) = \mathbb{E} \left\{ \tilde{\mathcal{S}}_X^{(k)}(q^{-s_m}v_0) \right\} - \mathcal{S}_{X,\psi}(q^{-s_m}v_0).$$

Supposons qu'il existe une constante $c_{\theta_1} > 0$ telle que $\tilde{\theta}_{n,1}^{(k)} > c_{\theta_1}, \forall n, k$. Alors

$$\left\| b_{\mathcal{S}_X}^{(k)} \right\|_\infty \leq \frac{J_X}{\|\psi\|_2^2} \left(K'_1 \left\| \theta_1 - \tilde{\theta}_1^{(k)} \right\|_\infty + K'_2 \left\| \tilde{\theta}_2^{(k)} - \theta_2 \right\|_\infty \right), \quad (2.22)$$

où

$$\begin{aligned}K'_1 &= \frac{\|\varphi_1\|_\infty}{c_{\theta_1}} < \infty, \\ K'_2 &= \ln(q) (\|\varphi_1\|_\infty + 2\|\hat{\psi}'\|_\infty \|\varphi_2\|_\infty) < \infty.\end{aligned}$$

La preuve de la proposition est donnée en annexe A.1.3. On note que si $\theta_1^{(k)} \rightarrow \theta_1$ et $\theta_2^{(k)} \rightarrow \theta_2$ quand $k \rightarrow \infty$, on a bien $\mathbb{E}\{\mathcal{S}_X^{(k)}(\nu_m)\} \rightarrow \mathcal{S}_{X,\psi}(\nu_m)$, comme espéré.

La formule (2.22) permet le contrôle du biais aux fréquences $\nu_m = q^{-s_m}\omega_0$ uniquement. Notons également que la condition $J_X < \infty$ force \mathcal{S}_X à s'annuler à la fréquence nulle. En effet, l'estimation du spectre est obtenue grâce à la transformée en ondelettes qui est un filtre passe-bande. Ainsi, la fréquence nulle n'est pas analysée et lorsque $\mathcal{S}_X(0) = 0$ aucune information n'est perdue lors du filtrage : l'estimateur du spectre-ondelettes peut alors estimer correctement $\mathcal{S}_X(\xi)$ pour tout $\xi > 0$. D'autre part, le contrôle du biais de l'estimateur du spectre-ondelettes nécessite des valeurs de K'_1 et K'_2 raisonnables. Outre le contrôle des fonctions φ_1 , φ_2 et $\hat{\psi}'$ (obtenu grâce à la décroissance rapide de l'ondelette), cela passe par le contrôle de l'estimation de la modulation d'amplitude. En effet, une valeur trop faible de $\tilde{\theta}_{n,1}^{(k)}$ perturbera fortement l'estimateur du spectre-ondelettes du fait de la division par ce terme.

(b) Variance. La borne de Cramér-Rao (BCR) donne la variance minimale qui peut être atteinte par un estimateur non biaisé. La formule de Slepian-Bangs (voir son expression dans [68], page 363) nous donne directement l'expression de celle-ci lorsqu'on a affaire à un estimateur du maximum de vraisemblance de données gaussiennes complexes circulaires. On obtient alors directement la borne de Cramér-Rao de l'estimateur de $\theta_{n,i}$ suivante :

$$\text{BCR}(\theta_{n,i}) = \left(\text{Trace} \left\{ \left(\mathbf{C}(\Theta_n)^{-1} \frac{\partial \mathbf{C}(\Theta_n)}{\partial \theta_{n,i}} \right)^2 \right\} \right)^{-1}.$$

Cette borne nous donne une information sur la variance de l'estimateur de $\theta_{n,i}$ lorsque l'algorithme JEFAS a convergé, c'est-à-dire quand à la fois \mathcal{S}_X et les paramètres de déformation sont bien estimés.

L'application de cette formule au paramètre de modulation d'amplitude $\theta_{n,1}$ nous donne

$$\mathbb{E} \left\{ (\tilde{\theta}_{n,1} - \mathbb{E} \{ \tilde{\theta}_{n,1} \})^2 \right\} \geq \text{BCR}(\theta_{n,1}) = \frac{2\theta_{n,1}^2}{M_s}.$$

Cela implique que le nombre d'échelles M_s de la transformée en ondelettes doit être suffisamment grand pour que la variance de l'estimateur soit faible.

Concernant le paramètre de dilatation instantanée $\theta_{n,2}$, nous n'avons d'expression explicite de la BCR. C'est pourquoi l'évaluation de cette borne et sa comparaison avec la variance de l'estimateur $\tilde{\theta}_{n,2}$ ne peut se faire qu'à partir de résultats numériques. Nous reportons donc cette analyse à la section 2.5.

(c) Robustesse au bruit. Supposons maintenant que les observations soient corrompues par un bruit blanc gaussien W de variance σ_W^2 (supposée connue). Le modèle devient

$$Y = \mathcal{A}_\alpha \mathcal{D}_\gamma X + W. \quad (2.23)$$

L'estimateur de la modulation d'amplitude $\tilde{\theta}_{n,1}$ n'est pas robuste au bruit. En effet, en présence de bruit blanc, un nouveau terme vient s'ajouter à l'expression du biais (2.19), qui devient

$$b_{n,1}^{(k)}(\theta_{n,1}) = b_0^{(k)}(\theta_{n,1}) + b_W^{(k)},$$

où $b_0^{(k)}(\theta_{n,1})$ est l'expression (2.19) du biais intrinsèque à l'estimateur, et $b_W^{(k)}$ est le biais additif introduit par le bruit. Il vérifie :

$$b_W^{(k)} = \frac{\sigma_W^2}{M_s} \text{Trace} \left\{ \tilde{\mathbf{C}}_0^{(k)} \left(\tilde{\theta}_{n,2}^{(k)} \right)^{-1} \mathbf{C}_\psi \right\},$$

où $[\mathbf{C}_\psi]_{ij} = q^{(s_i+s_j)/2} \int_0^\infty \bar{\psi}(q^{s_i}\xi) \hat{\psi}(q^{s_j}\xi) d\xi$. Naturellement, ce terme de biais supplémentaire est d'autant plus grand que la variance du bruit augmente.

En revanche, l'estimateur de la dilatation instantanée $\tilde{\theta}_{n,2}$ est très robuste au bruit. En effet, l'unique changement apporté par l'ajout du bruit dans l'expression de la log-vraisemblance concerne le remplacement du spectre \mathcal{S}_X par $\mathcal{S}_Z = \mathcal{S}_X + \frac{\sigma_W^2}{\theta_{n,2}}$. Ce terme constant additif n'affecte pas l'estimateur tant qu'il est petit par rapport aux valeurs maximales de \mathcal{S}_X .

En outre, l'estimateur $\tilde{\mathcal{S}}_X$ est modifié car lorsqu'on effectue l'opération de rectification des coefficients d'ondelettes $\frac{1}{\theta_{n,1}^{1/2}} \tilde{\mathcal{W}}_y(s_m - \theta_{n,2}, \tau_n)$ sur l'échelle s_m , on calcule en réalité :

$$\mathbf{w}_{Z,s_m} = \mathbf{w}_{X,s_m} + \mathbf{w}_{W,s_m},$$

où $\mathbf{w}_{W,s_m} = \frac{1}{\theta_1^{1/2}} \tilde{\mathcal{W}}_w(s_m - \theta_2, \tau)$ désigne la transformée en ondelettes d'un bruit blanc modulé en amplitude par a^{-1} . Ainsi, un nouveau terme $\tilde{\sigma}_W$ indépendant de la fréquence est ajouté au nouvel estimateur du spectre $\tilde{\mathcal{S}}_Z$, de sorte

$$\mathbb{E} \{ \tilde{\mathcal{S}}_Z \} = \mathcal{S}_{X,\psi} + \tilde{\sigma}_W^2 \quad \text{où} \quad \tilde{\sigma}_W^2 = \sigma_W^2 \frac{1}{N_\tau} \sum_{n=1}^{N_\tau} \frac{1}{\theta_{n,1}}.$$

Un nouveau terme de biais s'ajoute donc à celui introduit par l'estimateur du spectre-ondelettes.

2.5 Résultats numériques

Nous passons maintenant aux simulations numériques et aux applications. Ici on utilisera toujours l'ondelette *sharp* ψ_\sharp (voir la définition 2.1) pour évaluer les transformées en ondelettes.

2.5.1 L'ondelette *sharp*

L'outil d'analyse relatif à ce chapitre étant la transformée en ondelettes (voir la définition 1.15), le choix d'ondelette d'analyse qui est fait pour les évaluations numériques

est celui de l'ondelette *sharp*, notée ψ_{\sharp} , que nous avons introduite dans [45]. Elle est définie par sa transformée de Fourier sur les fréquences positives comme suit.

Définition 2.1 (Ondelette *sharp*). Soit $0 < \epsilon \ll 1$. Soit $\nu_0 > 0$ la fréquence propre de l'ondelette. Soit $\nu_1 > \nu_0$ une fréquence de coupure. Soit Δ la semimétrie définie par :

$$\Delta(a, b) = \frac{1}{2} \left(\frac{a}{b} + \frac{b}{a} \right) - 1 \quad \forall a, b > 0.$$

Alors, la transformée de Fourier de l'ondelette *sharp* ψ_{\sharp} est donnée sur les fréquences positives par

$$\hat{\psi}_{\sharp}(\nu) = e^{\frac{\Delta(\nu, \nu_0)}{\Delta(\nu_1, \nu_0)}} , \quad \nu > 0, \quad (2.24)$$

et elle s'annule aux fréquences négatives ($\nu \leq 0$).

Cette ondelette a la propriété de posséder une infinité de moments nuls, c'est-à-dire que

$$\int_{\mathbb{R}} t^k \psi_{\sharp}(t) dt = 0, \quad \forall k \in \mathbb{N}.$$

De manière équivalente, cette dernière propriété s'exprime sur la transformée de Fourier de l'ondelette de la façon suivante :

$$\hat{\psi}_{\sharp}^{(k)}(0) = 0, \quad \forall k \in \mathbb{N}.$$

On peut vérifier que $\hat{\psi}_{\sharp}(\nu) \rightarrow 0$ quand $\nu \rightarrow 0^+$. La transformée de Fourier $\hat{\psi}_{\sharp}$ est donc continue en zéro. De même, les dérivées successives de $\hat{\psi}_{\sharp}$ s'annulent en zéro en les prolongeant par continuité.

Ici, ν_0 désigne la fréquence propre de l'ondelette ψ_{\sharp} . Elle vérifie $\hat{\psi}_{\sharp}(\nu_0) = 1$ qui est la valeur maximale prise par $\hat{\psi}_{\sharp}$. La fréquence ν_1 correspond quant à elle, à la fréquence de coupure vérifiant $\hat{\psi}_{\sharp}(\nu_1) = \epsilon$. Dans les applications, nous choisirons $\nu_1 = F_s/2$ où F_s correspond à la fréquence d'échantillonnage. Nous choisissons une valeur de ϵ petite de manière à éviter le repliement de spectre dans la transformée en ondelettes à l'échelle $s = 0$.

En outre, il peut être plus intéressant d'exprimer $\hat{\psi}_{\sharp}$ en fonction du facteur de qualité de l'ondelette. Le facteur de qualité Q correspond au rapport entre la fréquence centrale ν_0 et sa bande passante $\Delta\nu$ donnée par $\Delta\nu = \nu_+ - \nu_-$ où $\hat{\psi}_{\sharp}(\nu_+) = \hat{\psi}_{\sharp}(\nu_-) = \frac{1}{\sqrt{2}}$ et $\nu_- < \nu_0 < \nu_+$. Le facteur de qualité est donc d'autant plus grand que la bande passante de l'ondelette est étroite. On peut montrer qu'on a le résultat suivant :

$$Q = \frac{\nu_0}{\Delta\nu} = \frac{1}{\sqrt{C(C+4)}} \quad \text{avec} \quad C = -\frac{\ln(2)}{\ln(\epsilon)} \Delta(\nu_1, \nu_0).$$

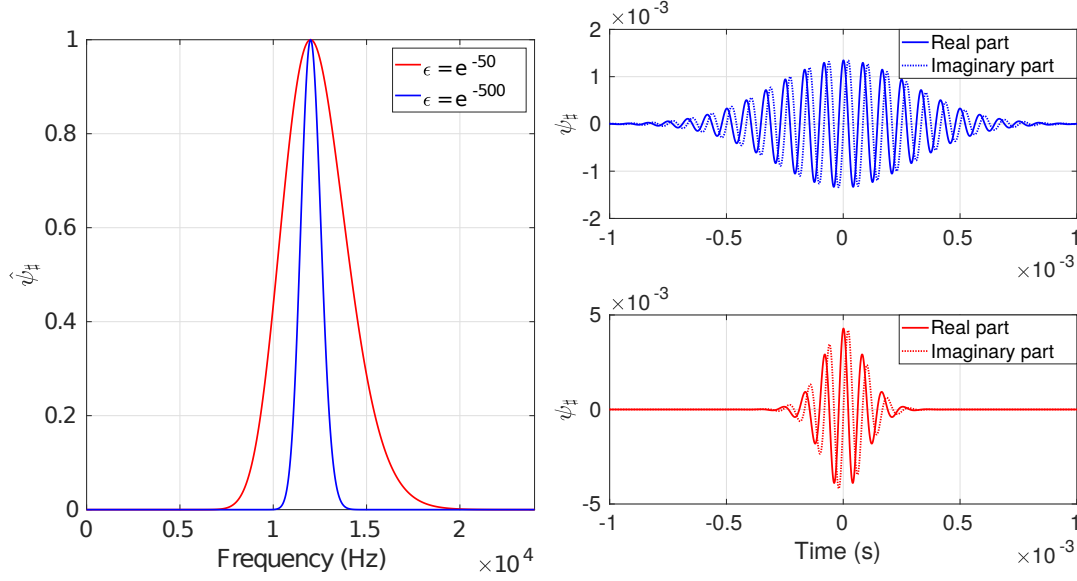


FIGURE 2.3 – À gauche : transformée de Fourier de ψ_{\sharp} deux valeurs différentes de ϵ . À droite : ondelettes correspondantes dans le domaine temporel : $\epsilon = e^{-500}$ soit $Q = 13,43$ (en haut) et $\epsilon = e^{-50}$ soit $Q = 4,24$ (en bas). Ici, $\nu_0 = F_s/4 = 11025$ Hz et $\nu_1 = F_s/2 = 22050$ Hz.

On peut alors exprimer $\hat{\psi}_{\sharp}$ en fonction du facteur de qualité et de la fréquence centrale :

$$\hat{\psi}_{\sharp}(\nu) = 2^{-Q(2Q + \sqrt{1+4Q^2})\Delta(\nu, \nu_0)}, \quad \forall \nu > 0.$$

Pour montrer l'influence du paramètre ϵ (ou de Q , de manière équivalente) sur la transformée en ondelettes, on affiche à gauche de la figure 2.3, la transformée de Fourier de l'ondelette *sharp* obtenue pour deux valeurs différentes de ϵ . Plus ϵ est petit, plus la transformée de Fourier de ψ_{\sharp} est étroite. De plus, le côté droit de la figure 2.3 montre les représentations temporelles de ces deux ondelettes. Il illustre le fait que lorsque ϵ diminue, l'étalement temporel de ψ_{\sharp} devient plus grand. Ensuite, nous calculons des transformées en ondelettes correspondant à ces deux valeurs de ϵ . Les scalogrammes sont affichés sur la figure 2.4. Le signal utilisé pour évaluer l'effet de ϵ est un enregistrement de hurlement de loup. En conséquence du résultat précédent, une valeur plus petite de ϵ améliore la localisation de la transformée en ondelettes le long de l'axe des échelles, tandis que la localisation le long de l'axe temporel est dégradée.

2.5.2 Estimateurs de référence

Nous comparons systématiquement notre approche à des estimateurs simples de la modulation d'amplitude et de la dilatation instantanée, couramment utilisés dans les applications, et définies ci-dessous.

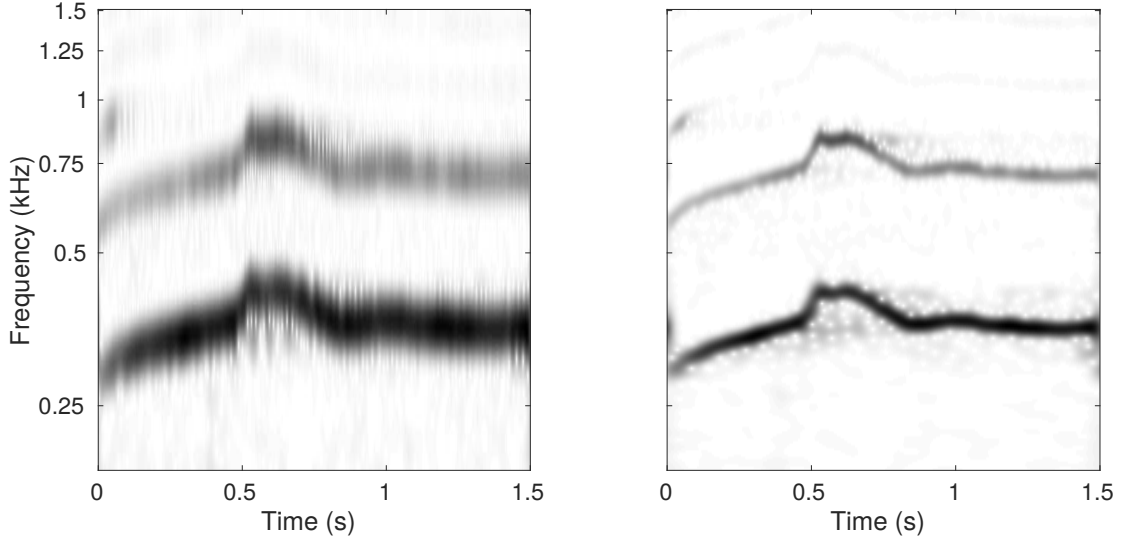


FIGURE 2.4 – Scalogrammes un enregistrement de hurlement de loup en utilisant l’ondelette d’analyse ψ_{\sharp} avec pour paramètre $\epsilon = e^{-50}$ (à gauche) et $\epsilon = e^{-500}$ (à droite).

L’approche de Clerc et Mallat (voir section 2.3.1) conduit à un estimateur de la fonction de dilatation instantanée des processus localement dilatés. Cependant, on a vu que sa mise en œuvre n’est pas adaptée à notre problème. En particulier, cette approche n’est pas valide aux échelles d’analyse standards. C’est pourquoi les estimateurs de référence des fonctions de déformation qui seront mis en œuvre sont les suivants.

Modulation d’amplitude : l’estimateur de référence $\tilde{\theta}_{n,1}^{(B)}$ du paramètre $\theta_{n,1} = \alpha(\tau_n)^2$ est tiré de l’énergie instantanée de la transformée en ondelettes moyennée entre les échelles, définie de la façon suivante :

$$\tilde{\theta}_{n,1}^{(B)} = \frac{1}{M_s} \|\mathbf{w}_{y,\tau_n}\|^2. \quad (2.25)$$

Cela revient à remplacer la matrice de covariance estimée dans (2.14) par la matrice identité. Notons que $\tilde{\theta}_{n,1}^{(B)}$ ne dépend ni de l’opérateur de dilatation instantanée γ' ni du spectre \mathcal{S}_X , et peut donc être calculé directement à partir des observations.

Dilatation instantanée : l’estimateur de référence $\tilde{\theta}_{n,2}^{(B)}$ est le centre de gravité par rapport aux échelles du scalogramme à temps fixé. Il est donné, à chaque instant, par :

$$\tilde{\theta}_{n,2}^{(B)} = C_0 + \frac{1}{\|\mathbf{w}_{y,\tau_n}\|^2} \sum_{m=1}^{M_s} s[m] |\mathbf{w}_{y,\tau_n}[m]|^2. \quad (2.26)$$

où la constante C_0 est choisie telle que le vecteur $\tilde{\theta}_2^{(B)}$ soit de moyenne nulle.

L’évaluation numérique de l’algorithme JEFAS est effectuée à la fois sur des signaux synthétiques et sur des signaux audio réels.

TABLEAU 2.1 – Erreur quadratique moyenne d'estimation pour les deux déformations.

Méthode d'estimation	Modulation d'amplitude	Dilatation instantanée
Référence	0.2015	0.0232
JEFAS	0.0701	0.0005

2.5.3 Signal synthétique

Nous évaluons d'abord les performances de l'algorithme sur un signal synthétique. La connaissance des vraies valeurs des fonctions de déformation et du spectre nous permet de comparer la variance et le biais empiriques avec leurs valeurs théoriques.

Le signal simulé est constitué de $N_\tau = 2^{16}$ échantillons. La fréquence d'échantillonnage étant de $F_s = 8$ kHz, cela signifie que la longueur du signal est de $t_F = (N_\tau - 1)/F_s \approx 8.2$ s. Le spectre \mathcal{S}_X se décompose sous la forme

$$\mathcal{S}_X = S_1 + S_2 \quad \text{où} \quad S_l(\nu) = \begin{cases} 1 + \cos\left(2\pi \frac{\nu - \nu_0^{(l)}}{\Delta \nu^{(l)}}\right) & \text{si } |\nu - \nu_0^{(l)}| < \frac{\Delta \nu^{(l)}}{2} \\ 0 & \text{sinon,} \end{cases} \quad (2.27)$$

avec $l \in \{1, 2\}$. La fonction de modulation d'amplitude instantanée α est une sinusoïde

$$\alpha(t) = \alpha_0 (1 + \alpha_1 \cos(2\pi t/T_1)) ,$$

où α_0 est choisi de sorte que $t_F^{-1} \int_0^{t_F} \alpha^2(t) dt = 1$. La fonction de déformation temporelle γ est telle que

$$\log_q(\gamma'(t)) = \Gamma + \cos(2\pi t/T_2) e^{-t/T_3} , \quad (2.28)$$

où Γ est choisi de sorte que $t_F^{-1} \int_0^{t_F} \gamma'(t) dt = 1$.

JEFAS est implémenté dans l'environnement scientifique MATLAB. Les dimensions ont été définies comme suit : $M_s = 106$ et $p = 7$. La transformée en ondelettes est calculée en utilisant l'ondelette *sharp* avec $\ln(\epsilon) = -25$ correspondant à un facteur de qualité $Q = 6$. Dans ce problème, il a fallu 67 secondes à l'algorithme pour converger sur un ordinateur de bureau standard (CPU Intel Core à 3,20 GHz \times 4 et 7,7 Go de RAM). Les résultats obtenus sont tracés en figure 2.5, ils sont superposés aux tracés des estimateurs de référence présentés précédemment (voir les équations (2.25) et (2.26)). Par souci de visibilité, l'estimateur de référence de la modulation d'amplitude (qui est très oscillant) n'est pas affiché, mais des évaluations numériques sont fournies dans le tableau 2.1, qui donne les valeurs d'erreurs quadratiques moyennes pour les différentes estimations. JEFAS est nettement plus précis que les estimateurs de référence. De plus, sa précision est bien expliquée par sa borne de Cramér-Rao : sur la figure 2.5, l'estimation est essentiellement comprise dans l'intervalle de confiance à 95 % fourni par la BCR (sous les hypothèses de gaussianité et de biais nul des estimateurs).

Le côté gauche de la figure 2.6 montre le spectre-ondelettes estimé à partir de la formule (2.16). L'accord avec le spectre réel est très bon, avec un léger effet d'élargissement des lobes dû au filtrage par $|\hat{\psi}|^2$. Le côté droit de la figure 2.6 montre l'évolution

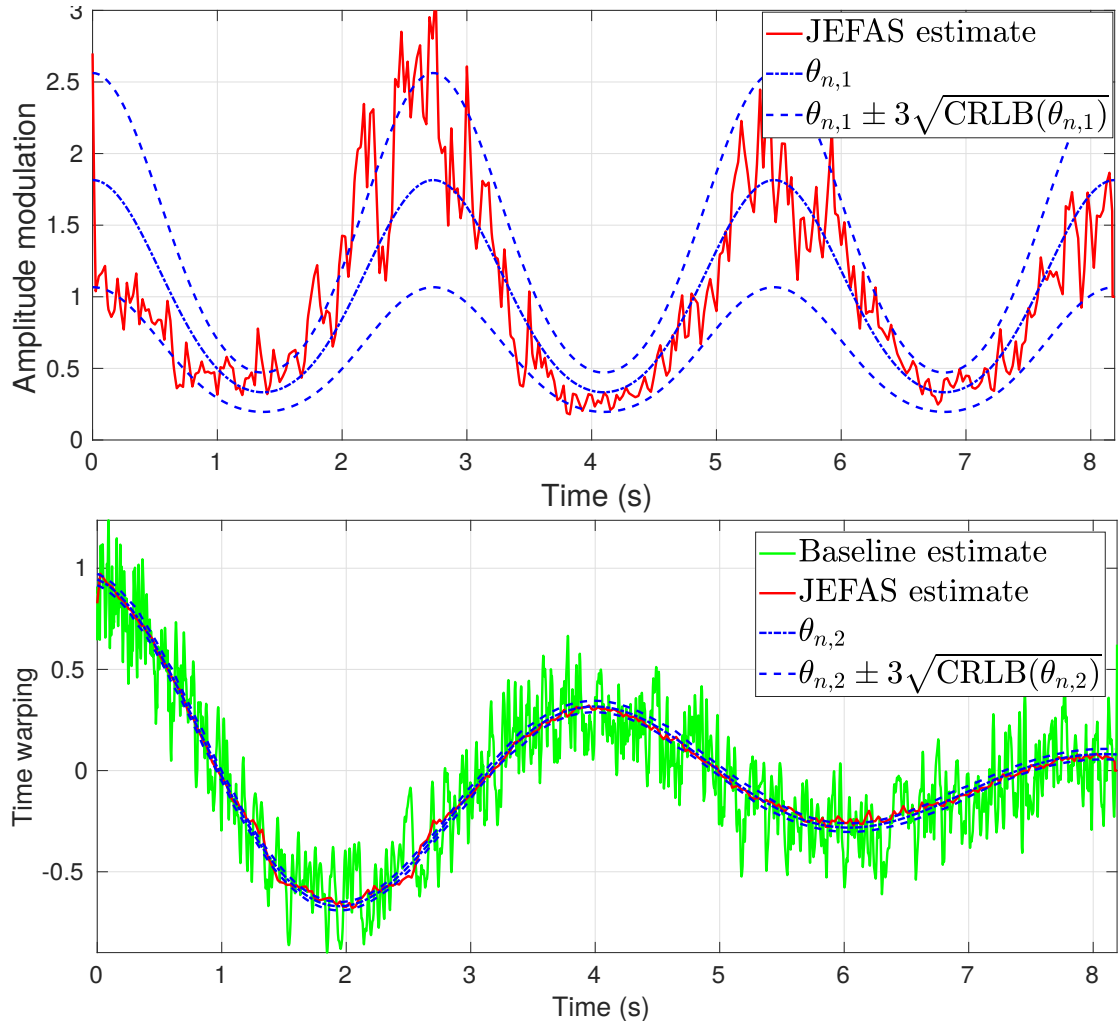


FIGURE 2.5 – Résultats de JEFAS sur un signal synthétique. En haut : estimation de la modulation d'amplitude ($\alpha_1 = 0.4$ et $T_1 = t_F/3$). En bas : estimation de la dilation instantanée ($T_2 = t_F/2$ et $T_3 = t_F/2$).

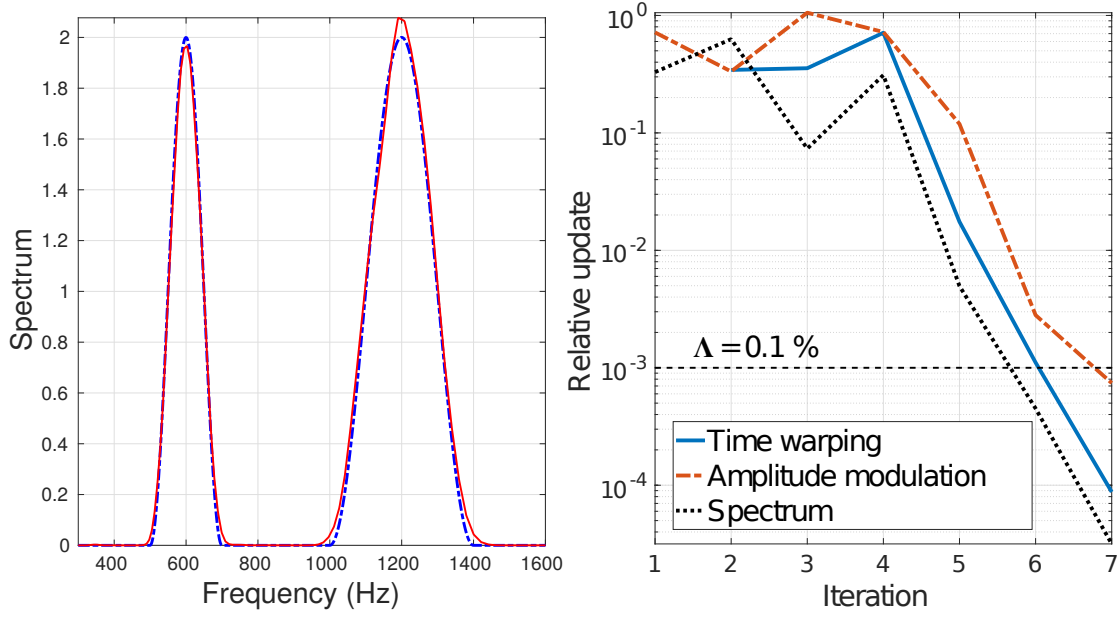


FIGURE 2.6 – À gauche : estimation du spectre ($\nu_0^{(1)} = 600$ Hz, $\Delta_\nu^{(1)} = 200$ Hz, $\nu_0^{(2)} = 1.2$ kHz, $\Delta_\nu^{(2)} = 400$ Hz) : spectres réel (ligne bleue pointillée) et estimé (ligne rouge continue). À droite : évolution de la mise à jour relative.

de la quantité déterminant le critère d'arrêt (2.18) aux cours des itérations de JEFAS. Les résultats numériques montrent que l'estimation de la dilatation instantanée converge plus rapidement que l'estimation de la modulation d'amplitude. Néanmoins, lorsqu'on fixe le seuil du critère d'arrêt à 0,1 %, seulement 7 itérations sont nécessaires pour obtenir la convergence de JEFAS.

2.5.4 Application à l'analyse spectrale du son produit par un dauphin

Après avoir étudié l'influence des divers paramètres, nous passons maintenant à des exemples audio réels. Tout d'abord, nous analysons un enregistrement d'un son de deux secondes de vocalisation du dauphin qui est décrit dans [69]. La transformée en ondelettes du signal, tracée sur la figure 2.7, montre que le modèle de déformation temporelle (2.1) correspond bien à ce type de signal, à l'exception des clics transitoires non pris en compte. JEFAS permet alors l'estimation du spectre du processus stationnaire sous-jacent.

En haut à droite de la figure 2.7, nous présentons la transformée en ondelettes du signal obtenu par application des déformations inverses estimées par JEFAS. Notez que la présence de clics perturbe légèrement le processus de « stationnarisation ». Néanmoins, il est naturel d'estimer une densité spectrale à partir de ce signal, car la dépendance temporelle de la transformée en ondelettes est négligeable par rapport à la

dépendance en échelle. Les spectres estimés à partir du signal d'origine (qui n'a pas grand sens mais qui est donné pour comparaison) et du signal stationnaire sous-jacent estimé sont affichés au milieu et en bas de la figure 2.7. Grâce à JEFAS, la structure harmonique apparaît clairement (tracé du bas).

2.5.5 Application à l'estimation de l'effet Doppler

Nous analysons maintenant un son qui est un enregistrement (depuis un emplacement fixe) d'une voiture de course, se déplaçant à une vitesse constante. Le son du moteur de la voiture est alors déformé par l'effet Doppler, ce qui entraîne une déformation temporelle, comme expliqué ci-dessous. En outre, comme la voiture est en mouvement, plus la voiture est proche du microphone, plus l'amplitude du son enregistré est grande. Ainsi, notre modèle correspond bien à ce signal.

Les transformées en ondelettes du signal d'origine et des deux estimations du signal stationnaire sous-jacent sont affichées à la figure 2.8. Alors que l'estimation de la déformation temporelle corrige uniquement le déplacement des coefficients d'ondelettes dans le domaine temporel, l'estimation conjointe de la déformation temporelle et de la modulation d'amplitude corrige également de manière approximative les variations non stationnaires des amplitudes.

La pertinence physique du modèle peut être vérifiée. En particulier, on peut donner une expression théorique de la fonction de dilatation instantanée. En effet, soit V la vitesse (constante) de la voiture et soit c la vitesse du son. Alors, en fixant l'origine du temps à l'instant auquel la voiture passe devant l'observateur à une distance d , il est possible de montrer que la fonction de dilatation instantanée due à l'effet Doppler est

$$\gamma'(t) = \frac{c^2}{c^2 - V^2} \left(1 - \frac{V^2 t}{\sqrt{d^2(c^2 - V^2) + (cVt)^2}} \right). \quad (2.29)$$

On trace en figure 2.8 (en bas à droite), l'estimée de $\tilde{\gamma}'$ superposée à son allure théorique lorsque $d = 5$ m et $V = 54$ m/s. Il est clair que l'estimation est proche de la courbe théorique correspondante obtenue avec ces données, qui sont des valeurs réalistes.

Néanmoins, un examen plus approfondi des scalogrammes de la figure 2.8 montre que la correction de la modulation d'amplitude n'est toujours pas parfaite, en raison de la présence de bruit, et du fait que le modèle reste trop simple : la modulation d'amplitude dépend en réalité de la fréquence, ce qui n'est pas pris en compte par le modèle.

2.5.6 Application à l'analyse d'un son de vent large bande

Nous illustrons ici la capacité de JEFAS à estimer le spectre instantané d'un signal non stationnaire dont le spectre du signal stationnaire sous-jacent est large bande. Le son étudié ici est le son produit par du vent soufflant à travers une porte, et présenté dans l'introduction de ce chapitre. Contrairement aux sons analysés précédemment, ce son n'affiche pas une structure harmonique claire. En effet, comme on peut le voir sur la

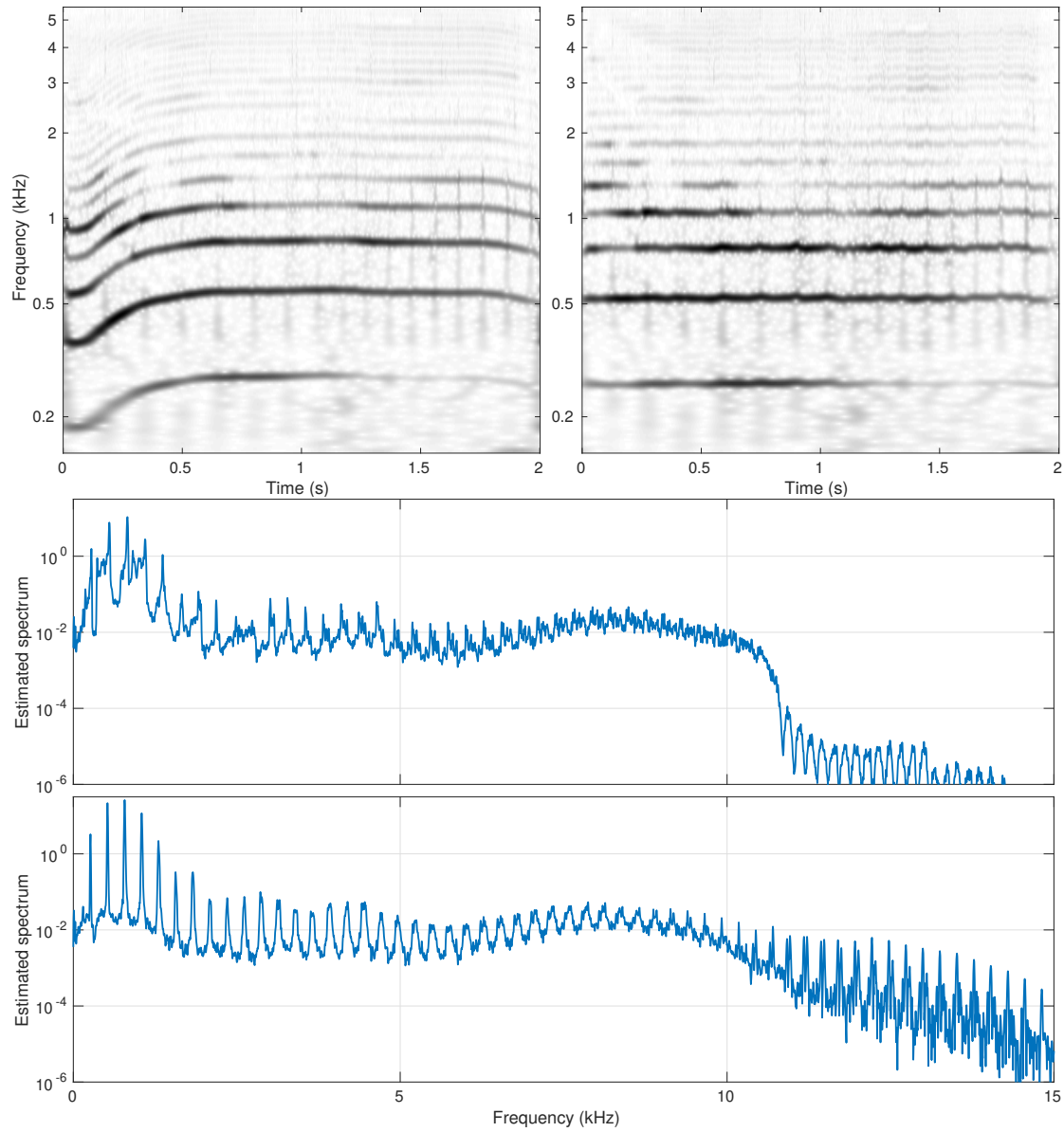


FIGURE 2.7 – Analyse spectrale du son produit par un dauphin. En haut à gauche : log-scalogramme du signal original. En haut à droite : log-scalogramme du signal démodulé en amplitude et « dé-dilaté ». Au milieu : Spectre estimé à partir du signal original. En bas : spectre estimé à partir du signal stationnaire sous-jacent obtenu par JEFAS.

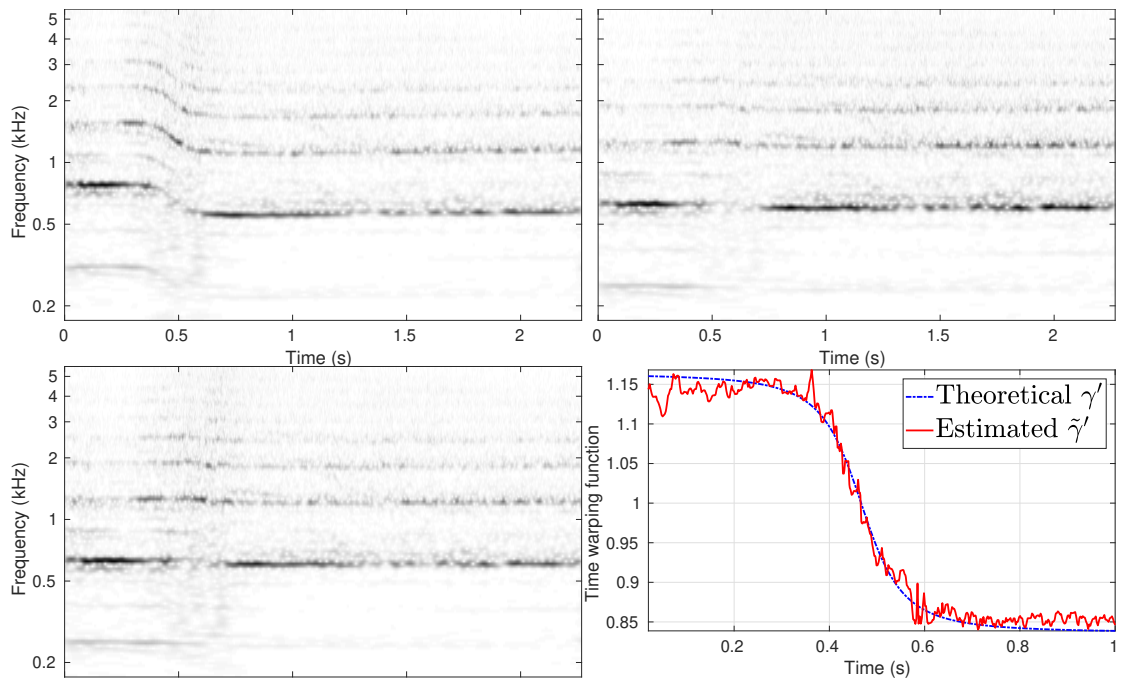


FIGURE 2.8 – Estimation de l'effet Doppler. En haut : log-scalogrammes du signal original (à gauche), et démodulé et « dé-dilaté » (à droite). En bas à gauche : log-scalogramme du signal « dé-dilaté ». En bas à droite : Dilatation instantanée estimée superposée à son allure théorique donnée en (2.29).

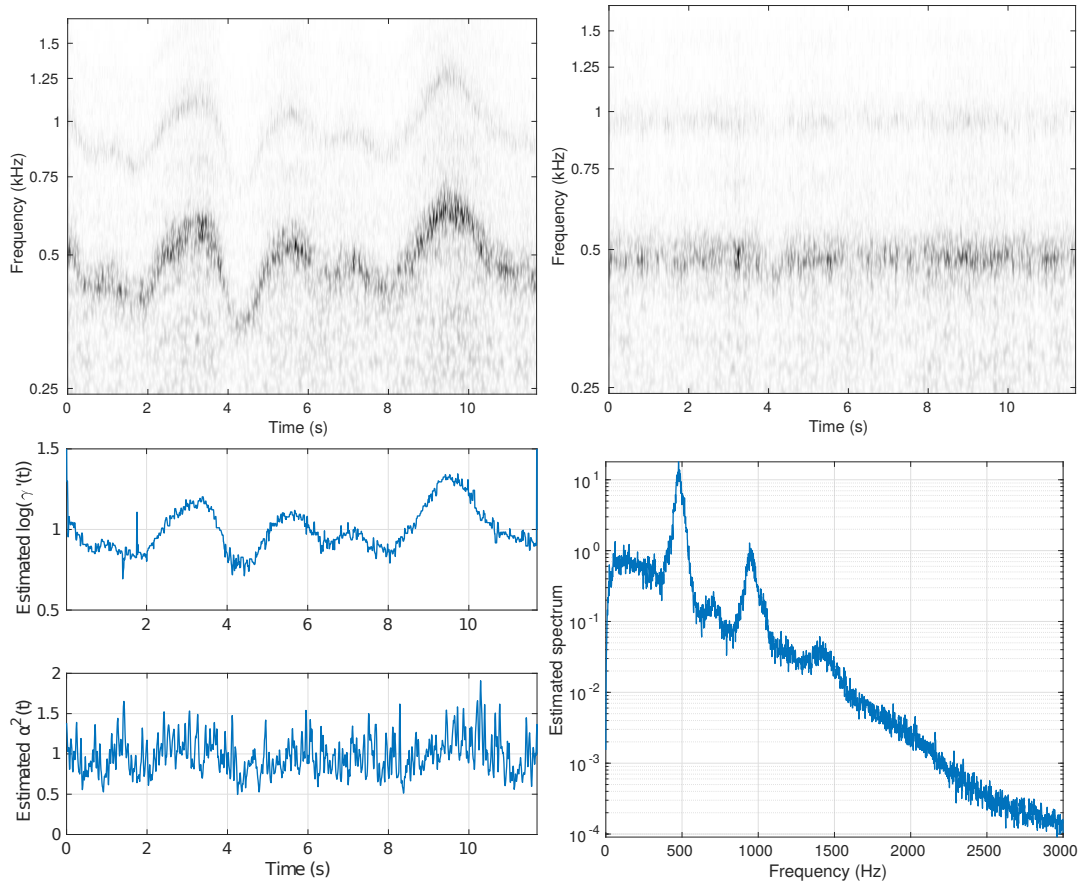


FIGURE 2.9 – Analyse d'un son de vent. En haut à gauche : transformée en ondelettes (module) du signal original. En haut à gauche : transformée en ondelettes (module) du signal démodulé et dé-dilaté. En bas à gauche : estimations des fonctions de dilatation instantanée et de modulation d'amplitude. En bas à droite : Estimation du spectre du processus stationnaire sous-jacent.

transformée en ondelettes du son original affiché en haut à gauche de la figure 2.9, deux lobes sont présents mais ils sont assez larges et le contenu fréquentiel est vraiment très large bande, avec une forte composante basse fréquence. L'estimation par JEFAS du spectre du processus stationnaire sous-jacent estimé est affichée en bas à droite de la figure 2.9. Cela confirme la nature large bande du signal et le fait que les deux lobes sont effectivement mal localisés en fréquence et ne pourraient pas être correctement modélisés par des sinusoides. De plus, la fonction de dilatation instantanée estimée est affichée en bas à gauche. Elle donne une image de la vitesse instantanée du vent.

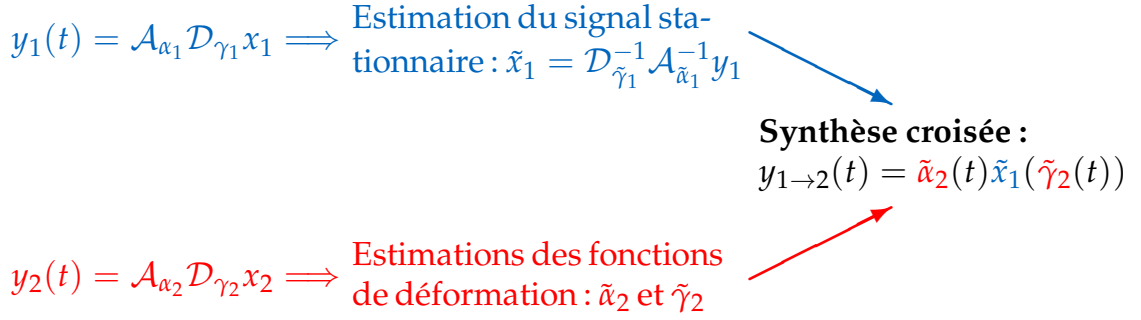


FIGURE 2.10 – Schéma illustratif des étapes de la synthèse croisée.

2.5.7 Application à la synthèse croisée de sons

Une application amusante de JEFAS est la synthèse croisée. Pour la réaliser, on applique JEFAS à deux signaux distincts. Pour le premier signal y_1 , on extrait l'estimation \tilde{x}_1 du signal stationnaire sous-jacent, tandis que pour le second signal y_2 , on extrait l'estimation des fonctions déformation. On synthétise alors un nouveau signal localement déformé en appliquant les déformations du second signal au premier signal stationnaire. On obtient alors le signal $y_{1 \rightarrow 2}$ dont le spectre du processus stationnaire sous-jacent est celui du premier signal alors que les déformations associées sont celles du second signal. La figure 2.10 illustre le procédé d'obtention du signal synthétisé.

Quelques exemples de synthèses croisées réalisées en audio sont disponibles en ligne² (voir annexe B.1). Par exemple, en partant du signal stationnaire de vent estimé précédemment, et en lui appliquant les déformations estimées par JEFAS à partir d'un enregistrement des vocalises d'une chanteuse, on obtient un son synthétisé qui paraît provenir d'un « vent chantant ».

2.6 Adaptation de l'algorithme à d'autres modèles de signaux

Le modèle retenu pour la mise en œuvre de JEFAS peut parfois être trop simpliste et non adapté à certains signaux. Pour décrire le comportement non stationnaire d'autres types de signaux, de nouvelles combinaisons des opérateurs de non-stationnarité locale peuvent être analysées. Par exemple, la combinaison de la déformation temporelle et de la modulation de fréquence peut être envisagée, comme nous l'avons fait dans [45]. Nous allons détailler rapidement ce cas par souci d'exhaustivité. En reprenant les notations présentés dans la section 1.4.2, le modèle de déformation dans [45] s'écrit

$$Y = \mathcal{A}_{\alpha} \mathcal{M}_{\phi} \mathcal{D}_{\gamma} X. \quad (2.30)$$

Pour réaliser l'estimation jointe des modulations d'amplitude et de fréquence et de la déformation temporelle à chaque instant, une transformée temps-fréquence-échelle appropriée \mathcal{V} est introduite. Elle est définie par $\mathcal{V}_X(\nu, s, \tau) = \langle X, \psi_{\nu s \tau} \rangle$, avec $\psi_{\nu s \tau} =$

2. <https://meynard.perso.math.cnrs.fr/paperJEFAS/NonStationaryAudio.html>

$T_\tau M_\nu D_s \psi$. Dans ce cas, un résultat d'approximation similaire au théorème 2.1 peut être obtenu. On approche la transformée temps-fréquence-échelle grâce aux opérateurs tangents. Les opérateurs tangents de la modulation d'amplitude et de la déformation temporelle sont donnés dans les équations (2.3) et (2.4). Celui de la modulation de fréquence est donné, pour toute fonction g localisée autour de $t = \tau$, par :

$$\mathcal{M}_\phi g(t) \approx \widetilde{\mathcal{M}}_\phi^\tau g(t), \quad \text{avec} \quad \widetilde{\mathcal{M}}_\phi^\tau \triangleq e^{2i\pi\phi(\tau)} T_\tau M_{\phi'(\tau)} T_{-\tau}. \quad (2.31)$$

On obtient alors la transformée approchée suivante :

$$\mathcal{V}_Y(\nu, s, \tau) \approx \widetilde{\mathcal{V}}_Y(\nu, s, \tau) \triangleq \left\langle \widetilde{\mathcal{A}}_\alpha \widetilde{\mathcal{M}}_\phi^\tau \widetilde{\mathcal{D}}_\gamma^\tau X, T_\tau M_\nu D_s \psi \right\rangle.$$

On peut réécrire cette dernière en fonction de la transformée temps-fréquence-échelle de X . On a :

$$\widetilde{\mathcal{V}}_Y(\nu, s, \tau) = \alpha(\tau) e^{2i\pi\phi(\tau)} \mathcal{V}_X \left(\frac{\nu - \phi'(\tau)}{\gamma'(\tau)}, s + \log_q(\gamma'(\tau)), \gamma(\tau) \right).$$

À partir de là, la log-vraisemblance associée peut être calculée. À τ fixé, la stratégie d'estimation est alors la même que précédemment, mais l'espace des paramètres est de plus grande dimension, et le paramètre supplémentaire $\theta_{n,3} = \phi'(\tau)$ complique la maximisation de log-vraisemblance. En particulier, le choix de la discrétisation des axes des échelles s et des fréquences ν influence les performances de l'estimateur, notamment la borne de Cramér-Rao.

2.7 Bilan

Dans ce chapitre, nous avons développé une méthode d'analyse spectrale des signaux localement dilatés et modulés en amplitude. Pour cela, nous nous sommes appuyés sur une idée originellement proposée par Clerc et Mallat. Il s'agit de voir la dilatation instantanée comme un opérateur déplaçant les coefficients d'ondelettes dans le plan temps-échelle. Cependant, la méthode qu'ils ont développée pour estimer ces déplacements n'est pas adaptée à l'analyse des signaux qui nous intéressent. C'est pourquoi nous avons ici repris et amélioré la procédure d'estimation développée par Omer et Torrèsani. La méthode qu'ils ont proposée consiste à décrire le déplacement approché des coefficients d'ondelettes dans le plan temps-échelle dans un cadre probabiliste en résolvant un problème de maximum de vraisemblance. Par rapport à cette méthode, les principaux points que nous avons apportés sont :

- l'ajout de la modulation d'amplitude au modèle de non-stationnarité. Ce nouveau modèle de signaux offre un degré de liberté supplémentaire et permet de prendre en compte une plus grande classe de signaux. Nous avons modifié l'algorithme en conséquence, afin d'estimer la fonction de modulation d'amplitude instantanée.

- l'estimation du paramètre de dilatation instantanée est réalisée sur un espace de paramètres continu. Auparavant, Omer considérait une estimation sur un espace de paramètres discret ce qui génère un effet de quantification, la fonction de dilatation instantanée estimée étant constante par morceaux. Cet effet, qui est perceptible à l'oreille, n'est plus présent dans notre approche.
- l'étape d'estimation spectrale est effectuée à partir de la transformée en ondelettes rectifiée du signal non stationnaire, c'est-à-dire après correction par les non-stationnarités inverses estimées. Ceci permet de conserver la cohérence de JEFAS qui n'opère qu'à partir de la transformée en ondelettes du signal. En effet, cette étape était auparavant réalisée à partir de la méthode de Welch appliquée au signal stationnarisé. Ceci conduisait à un algorithme s'appuyant à la fois sur la transformée en ondelettes du signal pour l'étape d'estimation des déformations et sur le signal lui-même pour l'étape d'estimation spectrale.
- l'étude théorique des performances de l'algorithme JEFAS. En conséquence de la résolution du problème de maximum de vraisemblance sur un espace de paramètres continu, nous avons pu établir quelques expressions permettant de quantifier les performances de l'algorithme. Entre autres, une expression de la borne de Cramér-Rao permettant de majorer la variance des estimations des fonctions de déformation instantanées a été donnée.

L'algorithme ainsi construit a pu être appliqué à des signaux audios synthétiques et réels. Les codes MATLAB de JEFAS et les scripts utilisés pour obtenir l'ensemble des figures de ce chapitre sont disponibles en ligne (voir l'annexe [B.1](#)). Par ailleurs, par rapport au pseudo-code de l'algorithme [2.2](#), nous avons ajouté une étape initiale de prétraitement : elle permet de détecter la présence du signal. Ainsi, l'estimation du spectre et des déformation n'est effectuée que sur les portions du signal où sa présence a été détectée. Ceci peut s'avérer utile lorsqu'on a affaire à des enregistrements longs de signaux brefs comme les cris d'oiseaux ou les hurlements de loups. L'application de JEFAS à ce type de signaux est également proposée en ligne.

On peut se demander comment étendre JEFAS à la situation où plusieurs signaux localement déformés et modulés en amplitude sont perçus simultanément. Il s'agit alors d'isoler chacune des sources. Ce problème de séparation de sources est étudié dans le chapitre suivant.

SÉPARATION DE SOURCES DOUBLEMENT NON STATIONNAIRE

3.1 Introduction

Dans le chapitre 2, l'algorithme JEFAS nous a permis d'estimer les fonctions de dilatation et de modulation d'amplitude instantanées caractérisant un processus localement dilaté et modulé en amplitude, à partir d'une réalisation de ce processus. Ce modèle suppose que le signal mesuré ne contient que le signal d'intérêt (plus un éventuel bruit blanc). On peut imaginer des situations où le signal perçu provient d'un certain nombre de signaux localement dilatés et modulés en amplitude qui sont mélangés. Considérons ici un exemple de mélange synthétique de signaux audio. Nous disposons de deux signaux mesurés (comme en stéréo), chacun d'entre eux étant obtenu à partir de deux sources sonores mélangées. Les deux sources que nous avons mélangées sont : le son de vent analysé dans le chapitre précédent (voir la section 2.5.6) et les vocalises d'une chanteuse. On représente en figure 3.1 les transformées en ondelettes (en module) des signaux mesurés. On peut alors distinguer deux cas de figure. Dans le premier cas, le mélange des signaux est linéaire invariant dans le temps. De nombreuses méthodes de séparation de sources permettent d'estimer les deux signaux sources (voir la section 3.2). On peut ensuite leur appliquer JEFAS afin d'estimer les fonctions recherchées. Dans le deuxième cas plus général, le mélange n'est plus constant au cours du temps. Cela revient alors à estimer des fonctions de modulation d'amplitude croisées en plus des termes donnés par JEFAS. En effet, en reprenant l'exemple de la figure 3.1, et en notant x_1 et x_2 les deux signaux stationnaires sources, chacun des deux signaux observés, notés respectivement z_1 et z_2 , se décomposent sous la forme :

$$\begin{aligned} z_1(t) &= \mathcal{A}_{\alpha_{11}} \mathcal{D}_{\gamma_1} x_1 + \mathcal{A}_{\alpha_{12}} \mathcal{D}_{\gamma_2} x_2 \\ z_2(t) &= \mathcal{A}_{\alpha_{21}} \mathcal{D}_{\gamma_1} x_1 + \mathcal{A}_{\alpha_{22}} \mathcal{D}_{\gamma_2} x_2 , \end{aligned}$$

où γ_1 et γ_2 désignent les fonctions de déformations temporelles. Les quatre fonctions α_{ij} combinent les modulations d'amplitudes provenant des sources avec celles dues

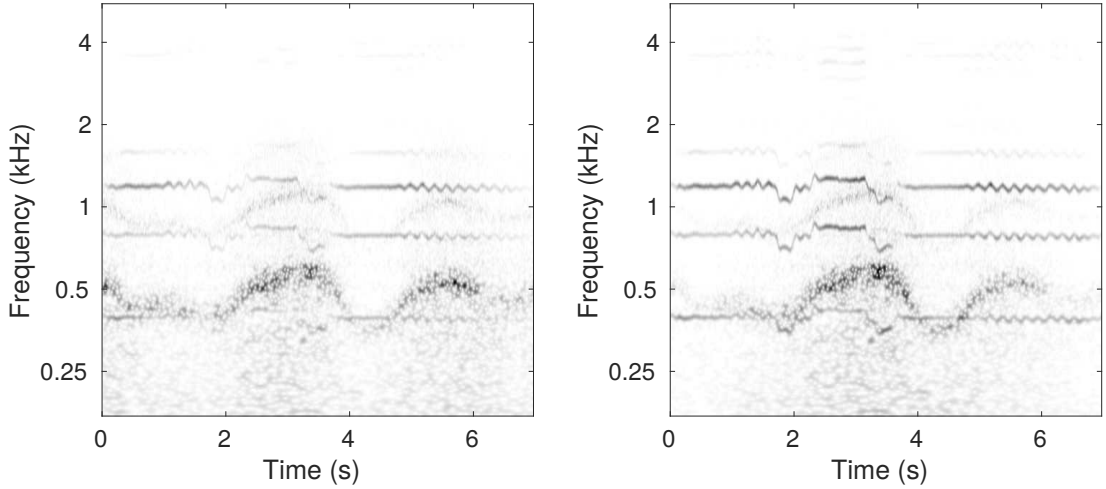


FIGURE 3.1 – Modules des transformées en ondelettes de deux mélanges de deux signaux audio.

au mélange non constant. On peut donc voir ce problème d'analyse des signaux non stationnaires comme une extension du modèle de JEFAS aux signaux multi-capteurs dans le cas où la modulation d'amplitude instantanée est un opérateur matriciel. Dans ce chapitre, nous présenterons un algorithme permettant d'analyser et d'estimer les grandeurs caractérisant ce type de modèle.

Dans ce qui suit, nous présentons d'abord les différents modèles de mélanges linéaires existants. On détaille également quelques méthodes de séparation de sources qui ont été développées pour être adaptées à ces modèles (par exemple : cas de sources stationnaires ou non). Ensuite, nous nous intéressons au problème de séparation de sources doublement non stationnaire que nous venons d'ébaucher. Nous parlons de non-stationnarité *double* car ce problème traite des signaux non stationnaires mélangés par une matrice de mélange non stationnaire. Nous montrerons que si les matrices de mélange et les fonctions de déformation temporelle sont suffisamment lisses et lentement variables, la vraisemblance approchée des coefficients d'ondelettes des observations peut s'exprimer sous une hypothèse de sources gaussiennes. Par conséquent, l'estimation conjointe de la matrice de mélange et des fonctions de déformation temporelle constitue un problème de maximum de vraisemblance. L'algorithme de séparation de sources proposé consiste en une estimation alternée de la matrice de mélange, des fonctions de déformation et des spectres des sources stationnaires sous-jacentes. Les estimations de ces deux dernières grandeurs sont effectuées par l'algorithme JEFAS (*Joint Estimation of Frequency, Amplitude and Spectrum*) introduit dans le chapitre 2. Enfin, nous évaluons les performances de l'algorithme proposé (appelé JEFAS-BSS) sur un mélange synthétique en le comparant à d'autres algorithmes de séparation de sources.

3.2 La séparation de sources

3.2.1 Position du problème

La séparation de sources est un problème classique en traitement du signal [16, 12]. Il s'agit de restituer des signaux sources à partir de signaux mesurés qui sont composés de mélanges des signaux sources. Les techniques de séparation aveugle de sources (*blind source separation*, BSS) visent à estimer conjointement des signaux sources et l'opérateur de mélange (par exemple une matrice de mélange dans le cas des mélanges linéaires instantanés) à partir des seules mesures. Entre autres, ce problème peut apparaître :

- en audio [76] lorsqu'on enregistre des sons provenant de plusieurs origines (par exemple les différents instruments d'un orchestre) ;
- en ingénierie biomédicale lorsqu'on souhaite séparer les signaux émis par différentes zones du cerveau à partir de mesures d'électroencéphalographie [21] ;
- en télécoms lorsqu'on souhaite récupérer le signal émis par une antenne au milieu d'autres signaux ;
- en astronomie, en chimie, etc.

Le problème de la séparation de sources s'est d'abord posé dans le contexte d'un mélange linéaire instantané (voir [16] pour une étude approfondie du sujet). Soit N_y le nombre de signaux sources que l'on souhaite extraire à partir des mesures de N_z mélanges. Notons par $\mathbf{y}(t) \in \mathbb{R}^{N_y}$ le vecteur contenant l'ensemble des sources à l'instant t et par $\mathbf{z} \in \mathbb{R}^{N_z}$ le vecteur contenant l'ensemble des mélanges observés à l'instant t . En supposant que les observations sont issues d'un mélange linéaire instantané des sources, elles s'expriment en fonction des sources sous la forme

$$\mathbf{z}(t) = \mathbf{A}\mathbf{y}(t), \quad (3.1)$$

où $\mathbf{A} \in \mathbb{R}^{N_z \times N_y}$ désigne la matrice de mélange. Le problème de la séparation aveugle des sources consiste alors à estimer les sources \mathbf{y} à partir des observations \mathbf{z} , la matrice de mélange \mathbf{A} étant inconnue.

Dans ce qui suit, nous nous focaliserons sur les méthodes de séparation de sources dans le cas où le nombre d'observations est égal au nombre de sources : $N_z = N_y$. Dans le cas surdéterminé (c'est-à-dire $N_z > N_y$), on effectue souvent une réduction de dimension pour revenir au cas $N_z = N_y$. Dans le cas sous-déterminé, le problème admet en règle générale une infinité de solutions. Afin d'assurer l'unicité de la solution, il est nécessaire de réaliser des hypothèses supplémentaires sur les sources. En particulier, l'hypothèse qu'il existe une représentation parcimonieuse des sources s'avère être particulièrement adaptée à ce problème [28]. Les méthodes de séparation de sources qui en découlent dépassent le cadre du présent travail.

Bien que les nombres de mélanges et de sources soient égaux, le problème de séparation aveugle des sources est mal posé. En effet, soit \mathbf{A}_0 et \mathbf{y}_0 un couple de solutions du problème de séparation de sources, alors, pour toute matrice \mathbf{M} inversible, le couple $\mathbf{A}_1 = \mathbf{A}_0\mathbf{M}^{-1}$ et $\mathbf{y}_1 = \mathbf{M}\mathbf{y}_0$ est également solution du problème. Effectivement, pour

tout t , on a

$$\mathbf{A}_1 \mathbf{y}_1(t) = \mathbf{A}_0 \mathbf{M}^{-1} \mathbf{M} \mathbf{y}_0(t) = \mathbf{A}_0 \mathbf{y}_0(t) = \mathbf{z}(t) .$$

Nous devons donc faire quelques hypothèses sur les sources pour restreindre les degrés de libertés laissés par le problème. Pour cela, nous allons considérer un modèle probabiliste où chacune des sources est supposée être une réalisation d'un processus stochastique. Alors, dans ce qui suit, il sera systématiquement fait l'hypothèse que les sources sont décorrélées les unes des autres. Autrement dit, la matrice de covariance instantanée des sources donnée par :

$$\mathbb{E} \left\{ (\mathbf{Y}(t) - \mathbb{E}\{\mathbf{Y}(t)\}) (\mathbf{Y}(t) - \mathbb{E}\{\mathbf{Y}(t)\})^H \right\}$$

est diagonale pour tout $t \in \mathbb{R}$. Malgré cette hypothèse, des indéterminations restent présentes. D'abord, toute permutation dans l'ordre des sources peut être compensée en permutant les colonnes correspondantes dans la matrice de mélange. D'autre part, toute multiplication du vecteur source \mathbf{y} par une matrice diagonale inversible peut être compensée par la multiplication à droite de la matrice de mélange par l'inverse cette matrice. Pour lever cette indétermination sur l'amplitude des sources, nous supposons que toutes les sources sont de variance unitaire. Finalement, on considérera donc comme solution du problème de séparation de sources tous les signaux $\tilde{\mathbf{y}}$ qui s'écrivent sous la forme $\tilde{\mathbf{y}} = \mathbf{P} \mathbf{D} \mathbf{y}$ où \mathbf{P} est une matrice de permutation et \mathbf{D} une matrice diagonale dont les éléments valent ± 1 (indétermination restante sur le signe des sources).

3.2.2 En contexte stationnaire

Les algorithmes de séparation de sources les plus utilisés s'appuient sur une hypothèse supplémentaire sur les sources à estimer : la stationnarité. Comme les observations dépendent linéairement des sources (avec des coefficients constants), elles sont également stationnaires. Cette nouvelle hypothèse permet donc d'avoir accès aux outils d'analyse des signaux stationnaires. Par exemple, on peut s'intéresser aux statistiques de second ordre. La matrice de covariance $\mathbf{K}_\mathbf{X}$ est définie pour tout vecteur \mathbf{X} de processus stochastiques stationnaires par :

$$\mathbf{K}_\mathbf{X}(\tau) = \mathbb{E} \left\{ (\mathbf{X}(\tau) - \mathbb{E}\{\mathbf{X}(0)\}) (\mathbf{X}(0) - \mathbb{E}\{\mathbf{X}(0)\})^H \right\} , \quad \forall \tau \in \mathbb{R}.$$

Notons que les éléments diagonaux de cette matrice de covariance contiennent les fonctions d'autocovariance de chacun des processus stochastiques formant le vecteur \mathbf{X} . Les éléments non diagonaux contiennent les covariances croisées entre les différents processus formant le vecteur \mathbf{X} .

Par exemple, l'algorithme SOBI [5] constitue une méthode de référence pour la résolution de ce problème. La première étape de SOBI consiste à blanchir les observations, c'est-à-dire rechercher une matrice $\mathbf{L} \in \mathbb{C}^{N \times N}$ telle que les éléments de $\mathbf{z}^\perp = \mathbf{L} \mathbf{z}$ soit décorrélés et de variance unitaire. Autrement dit, on veut avoir l'égalité $\mathbf{K}_{\mathbf{z}^\perp}(0) = \mathbf{I}$. Or,

la matrice $\mathbf{K}_Z(0)$ est hermitienne et semi-définie positive. Elle est donc diagonalisable avec :

$$\mathbf{K}_Z(0) = \mathbf{U}_0 \mathbf{D}_0 \mathbf{U}_0^H,$$

où \mathbf{U}_0 est une matrice unitaire, et \mathbf{D}_0 est une matrice diagonale dont les éléments diagonaux sont positifs ou nuls. Le choix $\mathbf{L} = \mathbf{D}_0^{-1/2} \mathbf{U}_0^H$ convient alors pour blanchir les observations. En effet :

$$\begin{aligned} \mathbf{K}_{Z^\perp}(0) &= \mathbf{L} \mathbf{K}_Z(0) \mathbf{L}^H \\ &= \mathbf{D}_0^{-1/2} \mathbf{U}_0^H \mathbf{U}_0 \mathbf{D}_0 \mathbf{U}_0^H \mathbf{U}_0 \mathbf{D}_0^{-1/2} = \mathbf{I}. \end{aligned} \quad (3.2)$$

En pratique, on obtient la matrice de blanchiment en diagonalisant l'estimateur empirique de la matrice de covariance $\mathbf{K}_Z(0)$. Notons que le choix de \mathbf{L} est unique (aux permutations près) à condition que les valeurs propres de $\mathbf{K}_Z(0)$ soient distinctes. Par ailleurs, étant donné que les sources sont décorrélées, on a $\mathbf{K}_Y(0) = \mathbf{I}$, et donc : $\mathbf{K}_Z(0) = \mathbf{A} \mathbf{A}^H$. Alors, en combinant ce résultat avec l'expression de $\mathbf{K}_{Z^\perp}(0)$ dans (3.2), on a $\mathbf{L} \mathbf{A} \mathbf{A}^H \mathbf{L}^H = \mathbf{I}$. La matrice $\mathbf{U} = \mathbf{L} \mathbf{A}$ est donc unitaire. Cette étape de blanchiment ramène donc la séparation de sources à la recherche d'une matrice unitaire car le modèle de mélange s'écrit désormais :

$$\mathbf{z}^\perp(t) = \mathbf{U} \mathbf{y}(t). \quad (3.3)$$

Les matrices de covariance des sources et des observations blanchies sont donc liées par :

$$\mathbf{K}_{Z^\perp}(\tau) = \mathbf{U} \mathbf{K}_Y(\tau) \mathbf{U}^H. \quad (3.4)$$

Les sources étant décorrélées, la matrice $\mathbf{K}_Y(\tau)$ est diagonale pour tout τ . L'équation matricielle (3.4) est donc une équation de diagonalisation de $\mathbf{K}_{Z^\perp}(\tau)$. L'estimation de la matrice de mélange correspondante repose donc sur les diagonalisations jointes des matrices de covariance des observations à différents instants τ .

Pour cela, on approche la matrice de covariance $\mathbf{K}_{Z^\perp}(\tau)$ grâce à l'estimateur empirique $\tilde{\mathbf{K}}_{Z^\perp}(\tau)$ de celle-ci :

$$\tilde{\mathbf{K}}_{Z^\perp}(\tau) = \frac{1}{P} \sum_{t=t_1}^{t_p} \mathbf{z}^\perp(t) \mathbf{z}^\perp(t - \tau)^H. \quad (3.5)$$

L'algorithme AMUSE [72], antérieur à SOBI, consiste à choisir une valeur τ_0 pour τ et à effectuer la diagonalisation de $\tilde{\mathbf{K}}_{Z^\perp}(\tau_0)$. Cette méthode possède l'inconvénient d'être sensible au bruit et à une mauvaise estimation de $\mathbf{K}_{Z^\perp}(\tau_0)$. Afin d'obtenir une estimation plus robuste, l'algorithme SOBI s'appuie sur l'estimée de $\mathbf{K}_{Z^\perp}(\tau)$ pour différentes valeurs de τ . En supposant que l'on a évalué les estimateurs empiriques de la matrice de covariance de \mathbf{z}^\perp pour un ensemble d'instant $\{\tau_1, \dots, \tau_J\}$, la diagonalisation jointe des matrices $\tilde{\mathbf{K}}_{Z^\perp}(\tau_j)$ est obtenue en minimisant le critère

$$g(\mathbf{U}) = \sum_{\tau=\tau_1}^{\tau_J} \left\| \text{off} \left(\mathbf{U}^H \tilde{\mathbf{K}}_{Z^\perp}(\tau_j) \mathbf{U} \right) \right\|_2^2, \quad (3.6)$$

où l'opération « off » consiste à mettre à zéro les éléments diagonaux d'une matrice. On cherche donc à minimiser simultanément la norme des éléments non diagonaux de l'ensemble des matrices de covariances des sources estimées. Soit $\tilde{\mathbf{U}}$ le minimiseur du critère (3.6), on obtient finalement l'estimation $\tilde{\mathbf{A}}$ de la matrice de mélange par : $\tilde{\mathbf{A}} = \tilde{\mathbf{L}}^{\dagger} \tilde{\mathbf{U}}$. Les auteurs démontrent que la séparation de sources ainsi obtenue est efficace à condition que les densités spectrales des différentes sources stationnaires soient distincts.

D'autres modèles de mélanges stationnaires ont été proposés. En effet, afin de modéliser des situations physiques plus complexes (par exemple : réverbération de signaux audio), des modèles de mélange non instantanés sont plus adéquats. De manière à préserver la stationnarité des observations, le mélange linéaire retenu pour modéliser de tels signaux est le mélange convolutif. Il s'écrit :

$$\mathbf{z}(t) = \int_{\mathbb{R}} \mathbf{H}(u) \mathbf{y}(t - u) du ,$$

où \mathbf{H} désigne un système de mélange contenant des filtres linéaires. Ce modèle est particulièrement adapté à la séparation de signaux de parole. De nombreuses méthodes de séparation de sources dans ce cadre ont été développées. On peut se référer à [56] pour une revue des algorithmes existants pour ce modèle.

Cependant, les hypothèses de stationnarité sur les sources et les mélanges limitent l'application de la séparation aveugle des sources à un petit nombre de signaux et de situations. Par la suite, des algorithmes de séparation de sources dans un contexte non stationnaire ont donc été proposés.

3.2.3 En contexte non stationnaire

On distingue deux cas de mélanges non stationnaires. Dans le premier cas, ce sont les sources qui sont non stationnaires tandis que le mélange préserve la stationnarité (mélange linéaire instantané ou convolutif). Dans le second cas, c'est le mélange qui rend les observations non stationnaires. Dans tous les cas, on doit maintenant s'intéresser à la matrice de covariance instantanée définie pour un vecteur \mathbf{X} de processus stochastiques par :

$$\mathbf{R}_{\mathbf{X}}(\tau, t) = \mathbb{E} \left\{ \mathbf{X}(t + \tau/2) \mathbf{X}(t - \tau/2)^H \right\} , \quad \forall \tau \in \mathbb{R}.$$

Mélange linéaire instantané de signaux non stationnaires. Plusieurs classes d'algorithmes adaptés à ce modèle ont été proposés. Une première classe repose sur les représentations temps-fréquence quadratique croisées entre les observations. Ces algorithmes présentent plusieurs variantes (en particulier en ce qui concerne l'étape 3 décrite ci-dessous). Ces différentes variantes sont passées en revue par Thirion-Moreau et Amin dans [70]. Néanmoins, ces algorithmes s'appuient en général sur quatre étapes détaillées dans ce qui suit.

1. *Blanchiment.* À la manière de l'algorithme SOBI, on se ramène à un problème de séparation de sources du même type que (3.3) où la matrice de mélange à estimer \mathbf{U} est unitaire et les observations \mathbf{z}^\perp sont décorrélées et de variance unitaire. L'opération de blanchiment s'effectue alors en diagonalisant la matrice de covariance instantanée moyennée par rapport à la variable non stationnaire et définie par :

$$\langle \mathbf{R}_X(\tau) \rangle = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} \mathbf{R}_X(\tau, t) dt .$$

2. *Analyse temps-fréquence quadratique.* L'outil d'analyse temps-fréquence employé ici (voir par exemple [70, 6]) est l'extension de la distribution de pseudo Wigner-Ville lissée (voir la définition 1.12) à l'analyse des signaux vectoriels (ou multicapteurs). Il s'agit d'analyser des « spectres instantanés croisés » entre les observations. C'est une matrice carrée, elle est donnée, pour tout vecteur de signaux \mathbf{x} , par :

$$\mathbf{V}_x(\xi, t) = \int_{\mathbb{R}} \mathbf{x}(t + \tau/2) \mathbf{x}(t - \tau/2)^H h(\tau) e^{-2i\pi\xi\tau} d\tau , \quad (3.7)$$

où h est un filtre de lissage. Il est choisi réel de sorte que $\mathbf{V}_x(\xi, t)$ soit une matrice hermitienne. Les éléments diagonaux de $\mathbf{V}_x(\xi, t)$ informent sur le spectre local de chacune des composantes de \mathbf{x} , tandis que les éléments non diagonaux contiennent les interférences en temps-fréquence entre les spectres instantanés des composantes de \mathbf{x} . De plus, la relation de mélange implique que les matrices de « corrélation temps-fréquence » $\mathbf{V}_{\mathbf{z}^\perp}(\xi, t)$ et $\mathbf{V}_y(\xi, t)$ sont liées comme suit :

$$\mathbf{V}_{\mathbf{z}^\perp}(\xi, t) = \mathbf{U} \mathbf{V}_y(\xi, t) \mathbf{U}^H . \quad (3.8)$$

3. *Sélection de points temps-fréquence.* On recherche des points dans le plan temps-fréquence pour lesquels un seul des spectres instantanés parmi ceux des sources prend une valeur significative. On note ces points (ξ_j, t_j) . En ces points, aucun terme d'interférences entre les sources n'est donc présent. Les matrices $\mathbf{V}_y(\xi_j, t_j)$ sont alors diagonales avec un seul élément diagonal non nul (celui de la source « active »). L'équation précédente (3.8) est donc une équation de diagonalisation. On peut se référer à [70] pour plus de détails sur les méthodes mises en œuvre pour détecter et sélectionner les points temps-fréquence (ξ_j, t_j) .
4. *Diagonalisation jointe.* Une fois les points (ξ_j, t_j) connus, il reste à effectuer une diagonalisation jointe des matrices $\mathbf{V}_{\mathbf{z}^\perp}(\xi_j, t_j)$. On peut la réaliser en minimisant un critère identique à celui obtenu dans l'équation (3.6) pour l'algorithme SOBI.

Dans ce chapitre, pour comparaison avec les méthodes de séparation de sources que nous proposons, nous mettons en œuvre une de ces techniques, appelée QTF-BSS (abréviation de *Quadratic Time-Frequency BSS*) dans le document. Ces techniques sont efficaces à condition que les signatures temps-fréquences des sources soient suffisamment disjointes pour que des points temps-fréquence où une seule source est présente existent.

Pour le même modèle de mélange, Parra et Spence [55] proposent une méthode sur la minimisation d'un critère basé sur la distance entre la covariance instantanée empirique et son expression théorique. Cette méthode est étendue au cas du mélange convolutif. Un algorithme de séparation de sources a également été proposé par Pham et Cardoso [58]. Il s'appuie sur un modèle particulier pour les sources non stationnaires. Il s'agit de bruits blancs localement modulés en amplitude (cf. définition 1.18). Les auteurs précisent que l'algorithme peut être étendu aux sources non stationnaires formées à partir de bruits colorés localement modulés en amplitude. La matrice de séparation est obtenue en résolvant un problème de séparation de sources pour lesquels les sources non stationnaires sont approchées par des signaux stationnaires par morceaux. Autrement dit, la modulation d'amplitude instantanée est constante par morceaux. Sous cette hypothèse, l'algorithme de séparation de sources repose, à l'instar de SOBI, sur une diagonalisation jointe de matrices de covariances.

Mélange non stationnaire (de sources stationnaires ou non). Plusieurs modèles ont été proposés pour ce type de problème. Dans [33], les auteurs considèrent le cas très général d'un mélange convolutif variant dans le temps. Le mélange est alors donné par :

$$\mathbf{z}(t) = \int_{\mathbb{R}} \mathbf{H}(t, u) \mathbf{y}(u) du .$$

Le système de mélange \mathbf{H} contient ici des filtres variant dans le temps. Ce modèle inclut, entre autres, le mélange instantané non constant. Pour cela, il suffit de prendre $\mathbf{H}(t, u) = \mathbf{A}(t)\delta(u)$ dans l'équation précédente. Néanmoins, la méthode de séparation de sources associée repose sur l'hypothèse qu'il existe une représentation des sources dans laquelle elles sont parcimonieuses. Il s'agit alors, comme pour QTF-BSS, de sélectionner les points où une seule source est active. Concernant la BSS d'un mélange linéaire variant dans le temps, l'approche de différents algorithmes de BSS (voir par exemple [63]) consiste à effectuer une analyse en composantes indépendantes sur de courtes portions du mélange sur lesquelles la stationnarité du mélange peut être supposée. L'idée est alors d'approcher les sources non stationnaires par des signaux stationnaires par morceaux.

Dans le même esprit, on peut adapter SOBI à ce modèle en effectuant une estimation SOBI par morceaux sur des sous-intervalles de longueur T_f ne se chevauchant pas. Cette méthode sera notée p-SOBI dans le manuscrit. Le pseudo-code correspondant à cette méthode est donné dans l'algorithme 3.1.

3.2.4 Indices de performances

Quels que soient le modèle de mélange et la méthode de séparation de sources appliquée, il est important d'évaluer quantitativement les performances de celle-ci. Pour cela, différents indices de performances ont été introduits. Nous présentons ici les indices que nous utiliserons dans la suite.

Algorithme 3.1 p-SOBI : $\tilde{\mathbf{B}} = \text{p-SOBI}(\mathbf{z}, T_f)$

```

 $p \leftarrow 0$ 
tant que  $pT_f < T$  faire
    Conserver  $\mathbf{z}_p$  : la restriction des observations  $\mathbf{z}$  à l'intervalle  $[pT_f, (p+1)T_f[$ .
    Estimation de la matrice de séparation :  $\mathbf{B}_p = \text{SOBI}(\mathbf{z}_p)$ .
fin tant que
Sortie :  $\tilde{\mathbf{B}}(\tau) = \mathbf{B}_p$  si  $\tau \in [pT_f, (p+1)T_f[$ 
    
```

3.2.4.1 Des indices de performance globaux

Indépendamment du modèle de mélange et de la méthode de séparation de sources, Vincent *et al.* [77] ont introduit différents indices permettant de mesurer la qualité de la séparation de sources effectuée. Nous présentons ici deux de ces indices, que nous exploiterons par la suite. Avec les notations (3.1), soit \tilde{y}_i l'estimée de la i^{e} source. Les auteurs la décomposent sous la forme :

$$\tilde{y}_i(t) = y_{\text{target}}(t) + e_{\text{interf}}(t) + e_{\text{noise}}(t) + e_{\text{artif}}(t),$$

où y_{target} désigne la projection de la source estimée \tilde{y}_i sur le vrai signal source y_i (y_{target} et y_i sont généralement égaux à un facteur multiplicatif près). e_{interf} désigne le terme d'interférences provenant des sources véritables y_j , $j \neq i$ dans la i^{e} source estimée. Le terme e_{noise} désigne la contribution du bruit. Le signal e_{artif} contient ce qui a été créé artificiellement par l'algorithme de séparation de sources. Le rapport source à interférences (SIR) quantifie le rapport entre l'énergie du terme que l'on cherche à estimer y_{target} et celle du terme e_{interf} provenant des autres sources. Ce terme s'exprime en décibels, on a donc :

$$\text{SIR}(\tilde{y}_i, \mathbf{y}) = 10 \log_{10} \left(\frac{\|y_{\text{target}}\|_2^2}{\|e_{\text{interf}}\|_2^2} \right). \quad (3.9)$$

Le rapport source à distorsion (SDR) quantifie le rapport entre l'énergie du terme y_{target} et celle du terme contenant toutes les signaux non désirés. En décibels, ce terme s'exprime :

$$\text{SDR}(\tilde{y}_i, \mathbf{y}) = 10 \log_{10} \left(\frac{\|y_{\text{target}}\|_2^2}{\|e_{\text{interf}} + e_{\text{noise}} + e_{\text{artif}}\|_2^2} \right). \quad (3.10)$$

Par définition le SDR est donc inférieur au SIR.

Les auteurs de [77] définissent un troisième indice de performance : le rapport source à artefacts (SAR). Cet indice est donné par :

$$\text{SAR}(\tilde{y}_i, \mathbf{y}) = 10 \log_{10} \left(\frac{\|y_{\text{target}} + e_{\text{interf}} + e_{\text{noise}}\|_2^2}{\|e_{\text{artif}}\|_2^2} \right). \quad (3.11)$$

Le SAR quantifie la présence d'artefacts générés par l'algorithme de séparation de sources dans les sources estimées. Cet indice étant plus difficilement interprétable que

les deux indices précédents, on se limitera en pratique à l'évaluation numérique du SIR et du SDR.

3.2.4.2 Un indice de performance local

Dans le cadre moins général du mélange linéaire instantané où la matrice de mélange varie dans le temps, les indices de performances globaux ne sont plus suffisants. En effet, on voudrait pouvoir quantifier la qualité de la séparation de sources à chaque instant. Des adaptations des indices de performances tels que le SIR, le SDR et le SAR à cette situation ont été proposés. Pour cela, il s'agit d'évaluer les variations de ces indices au cours du temps en les calculant sur des portions successives des sources estimées. Cependant, ces indices sont sensibles à la largeur de la fenêtre utilisée pour obtenir les différentes portions des sources estimées.

Pour contourner ce problème, on introduit un indice de performance qui évalue la qualité de l'estimation de la matrice de mélange à chaque instant. Dans le cas où la matrice de mélange varie dans le temps, cette façon d'évaluer la BSS nous permet de suivre au cours du temps la qualité de la séparation de sources. Soit $\mathbf{B}(t) = \mathbf{A}^{-1}(t)$ la matrice de séparation et $\tilde{\mathbf{B}}(t)$ son estimation. Alors, pour que la séparation de sources soit bien réalisée il faut que la matrice $\mathbf{G}(t) = \tilde{\mathbf{B}}(t)\mathbf{A}(t)$ soit la plus proche possible d'une matrice diagonale dont les éléments valent ± 1 . L'indice d'Amari normalisé (initialement introduit par Moreau et Macchi [47] pour la version de l'indice non normalisé) permet de quantifier cela. Il est noté $\rho(t)$ et est défini par

$$\rho(t) = \frac{1}{2N_y(N_y - 1)} \left(\sum_{i=1}^{N_y} \left(\sum_{j=1}^{N_y} \frac{|[\mathbf{G}(t)]_{ij}|}{\max_k |[\mathbf{G}(t)]_{ik}|} - 1 \right) + \sum_{j=1}^{N_y} \left(\sum_{i=1}^{N_y} \frac{|[\mathbf{G}(t)]_{ij}|}{\max_k |[\mathbf{G}(t)]_{kj}|} - 1 \right) \right). \quad (3.12)$$

Lorsque la matrice $\mathbf{G}(t)$ satisfait la condition décrite ci-dessus, c'est à dire $|\mathbf{G}(t)| = \mathbf{I}$, l'indice d'Amari s'annule : $\rho(t) = 0$. Cet indice est donc d'autant plus faible que la séparation de sources est de bonne qualité. De plus, on a toujours $\rho(t) \in [0, 1]$ avec $\rho(t) = 1$ dans le cas le plus défavorable à la séparation de sources, c'est-à-dire lorsque tous les coefficients de $|\mathbf{G}(t)|$ sont identiques.

Il est à noter que le calcul des différentes mesures de performance présentées dans cette section requiert la connaissance des vraies sources ou de la vraie matrice de mélange. Pour nous permettre de comparer les algorithmes de séparation de sources, nous appliquerons donc ces algorithmes à des mélanges synthétiques pour lesquels nous fixons ces grandeurs.

Dans la suite, nous présentons un modèle de mélange où la matrice de mélange varie dans le temps, et les sources suivent un modèle de non-stationnarité particulier : la déformation temporelle.

3.3 Modèle et objectif

3.3.1 Une classe de signaux non stationnaires

Dans ce qui suit, nous considérons les signaux non stationnaires comme des réalisations de processus localement déformés. En effet, comme nous l'avons vu dans le chapitre 1, cette hypothèse est appropriée pour décrire de nombreux signaux de la vie réelle, et nous avons présenté différentes classes d'opérateurs brisant la stationnarité qui sont pertinentes pour modéliser des phénomènes physiques. Nous nous focalisons ici sur l'opérateur de déformation temporelle. Un tel opérateur peut modéliser des phénomènes physiques non stationnaires aussi divers que l'effet Doppler, les variations de vitesse d'un moteur, la vocalisation d'animaux ou la parole (voir le chapitre 2).

Introduisons maintenant quelques notations. Soit x_i un signal stationnaire, modélisé comme la réalisation d'un processus aléatoire stationnaire X_i dont la densité spectrale de puissance est notée \mathcal{S}_{X_i} . En agissant sur x_i avec l'opérateur de déformation temporelle noté \mathcal{D}_{γ_i} (voir la définition 1.16), on obtient un signal non stationnaire noté y_i . Dans ce chapitre les signaux non stationnaires sources y_i seront donc modélisés comme étant des réalisations de processus localement dilatés Y_i donnés par :

$$Y_i(t) = \mathcal{D}_{\gamma_i} X_i(t), \quad \forall t \in \mathbb{R}. \quad (3.13)$$

3.3.2 Mélange instantané non stationnaire

Le problème que nous considérons est la séparation de N_y sources non stationnaires modélisées par l'équation (3.13). Chacune d'entre elles étant formée à partir d'un processus non stationnaire et d'une fonction de déformation temporelle distincts. Les sources sont en outre supposées non corrélées. Soit $\mathbf{y}(t) \in \mathbb{R}^{N_y}$ le vecteur colonne contenant toutes les sources à l'instant t . Le mélange considéré est lui-même non stationnaire dans la mesure où la matrice de mélange varie au cours du temps. Dans la suite, nous supposons que le nombre d'observations est au moins égal au nombre de sources. Soit $\mathbf{z}(t) \in \mathbb{R}^{N_z}$ (avec $N_z \geq N_y$) le vecteur colonne contenant les observations à l'instant t . Le modèle de mélange s'écrit alors

$$\mathbf{z}(t) = \mathbf{A}(t)\mathbf{y}(t), \quad (3.14)$$

où $\mathbf{A}(t) \in \mathbb{R}^{N_z \times N_y}$ désigne la matrice de mélange variant dans le temps. Ce modèle généralise le modèle de modulation d'amplitude instantanée correspondant au cas où $\mathbf{A}(t)$ est diagonale. En effet, dans cette situation, chaque observation z_i peut se mettre sous la forme :

$$z_i(t) = a_i(t)y_i(t) = \mathcal{A}_{a_i} \mathcal{D}_{\gamma_i} x_i(t).$$

Cette dernière situation correspondant au modèle présenté dans le chapitre précédent (voir équation (2.1)). L'algorithme JEFAS nous a alors permis d'estimer les fonctions de déformation. L'estimation des sources stationnaires x_i et non stationnaires y_i est alors effectuée en appliquant les opérations de déformations inverses.

Dans le cas général, notre objectif sera de déterminer conjointement et à chaque instant : la matrice de mélange instantané $\mathbf{A}(t)$, les fonctions de déformation temporelle $\gamma_i(t)$, et le spectre des sources stationnaires \mathcal{S}_{X_i} pour $i \in \{1, \dots, N_y\}$. Toutes ces grandeurs doivent être obtenues uniquement à partir des observations $\mathbf{z}(t)$.

3.4 Approximation des transformées en ondelettes

Comme nous l'avons vu au chapitre 2, le comportement de la transformée en ondelettes d'un signal localement dilaté s'interprète en terme de déplacement des coefficients d'ondelettes dans le plans temps-échelle. C'est pourquoi nous allons chercher à exprimer l'équation de mélange en fonction des transformées en ondelettes des sources et des observations.

Pour cela, plaçons-nous à un instant fixe τ . Puis, pour chaque observation z_i , on note $\mathbf{w}_{z_i, \tau} = \mathcal{W}_{z_i}(\mathbf{s}, \tau)$ le vecteur ligne contenant les valeurs de la transformée en ondelettes à l'instant τ pour un vecteur d'échelles \mathbf{s} (dont la taille est notée M_s). Ensuite, tous ces vecteurs sont rassemblés dans une matrice de taille $N_z \times M_s$ notée $\mathbf{w}_{\mathbf{z}, \tau}$ telle que

$$\mathbf{w}_{\mathbf{z}, \tau} = \begin{pmatrix} \mathbf{w}_{z_1, \tau} \\ \vdots \\ \mathbf{w}_{z_{N_z}, \tau} \end{pmatrix}.$$

La même notation est utilisée pour la transformée en ondelettes des sources $\mathbf{w}_{\mathbf{y}, \tau}$. On suppose que la matrice $\mathbf{A}(t)$ varie lentement vis-à-vis des oscillations des signaux. Nous allons montrer qu'on peut alors approcher le mélange en ondelettes comme étant un mélange linéaire instantané. Sous cette hypothèse, la relation linéaire (3.14) devient une relation entre les transformées en ondelettes de \mathbf{y} et \mathbf{z} de la forme

$$\mathbf{w}_{\mathbf{z}, \tau} \approx \mathbf{A}(\tau) \mathbf{w}_{\mathbf{y}, \tau}. \quad (3.15)$$

Le théorème d'approximation suivant donne une borne quantitative sur l'erreur d'approximation dans l'équation (3.15). Remarquons que la notation $\mathbf{M}^{\circ 2}$ désigne ici le carré élément par élément d'une matrice \mathbf{M} .

Théorème 3.1 (Approximation). Soient X_i ($i \in \{1, \dots, N_y\}$) des processus stochastiques réels stationnaires au sens large, de moyenne nulle, de puissance finie σ_X^2 . Soient Y_i les processus localement dilatés définis par (3.13) et Z_i les mélanges non stationnaires donnés par (3.14). Soit ψ l'ondelette d'analyse régulière, localisée autour de l'origine de sorte que $|\psi(t)| \leq 1/(1 + |t|^\beta)$ avec $\beta > 1$. Avec les notations ci-dessus, soit $\epsilon_\tau \in \mathbb{C}^{N_z \times M_s}$ l'erreur d'approximation dans le domaine des ondelettes, définie ainsi :

$$\epsilon_\tau = \mathbf{w}_{\mathbf{z}, \tau} - \mathbf{A}(\tau) \mathbf{w}_{\mathbf{y}, \tau}. \quad (3.16)$$

Alors l'erreur d'approximation ϵ_τ est une matrice aléatoire du second ordre, complexe circulaire de moyenne nulle. De plus, la variance des coefficients de la matrice d'erreur est bornée comme suit :

$$\mathbb{E} \{ |\epsilon_\tau|^2 \} \leq \frac{\sigma_x^2 \beta^2}{(\beta - 2)^2} \mathbf{A}'_{\infty} \gamma'_{\infty} (q^{3s})^T, \quad (3.17)$$

où

$$\begin{aligned} \mathbf{A}'_{\infty} &\in \mathbb{R}_+^{N_z \times N_y} : & (\mathbf{A}'_{\infty})_{ij} &= \sup_t |\mathbf{A}'_{ij}(t)|, \\ \gamma'_{\infty} &\in \mathbb{R}_+^{N_y \times 1} : & (\gamma'_{\infty})_i &= \sup_t |\gamma'_i(t)|. \end{aligned}$$

La démonstration repose sur des arguments similaires à celle du théorème d'approximation 3.1. Un développement de Taylor de \mathbf{A} en τ permet la construction de la borne donnée par l'équation (3.17). On détaille cette démonstration en annexe A.2.1. Notons que, mis à part les termes contrôlant la borne sur l'erreur dans (2.5), c'est-à-dire la résolution temporelle de l'ondelette et les variations des fonctions de déformation temporelle, la borne sur l'erreur dans (3.15) est également contrôlée par les variations des coefficients de la matrice de mélange.

Cette nouvelle relation de mélange va nous permettre de réaliser la séparation de sources dans le domaine ondelettes ce qui rend possible l'estimation de la matrice de mélange instantané à chaque instant τ « indépendamment » des autres instants.

3.5 Procédure d'estimation et algorithme

La procédure d'estimation repose sur les relations approchées (2.5) et (3.15), que nous supposons valides, c'est-à-dire que \mathbf{A} et γ' varient suffisamment lentement. Cette hypothèse nous permet d'écrire une vraisemblance approchée sur les transformées en ondelettes des observations dans le cas gaussien. L'estimation s'appuie sur des transformées en ondelettes discrètes, les paramètres évoluant dans le temps sont donc estimés sur une grille de temps discrète τ . Dans ce qui suit, nous décrivons la procédure d'estimation pour un $\tau \in \tau$ donné.

3.5.1 Réduction de dimension

Dans le cas où le nombre de sources à estimer est strictement inférieur au nombre de signaux observés, on doit procéder à une réduction de dimension préalable permettant de ramener le nombre de signaux mesurés au nombre de sources à estimer. En effet, on verra dans ce qui suit que la méthode de séparation de sources retenue nécessite que la matrice de mélange considérée soit carrée (pour calculer, par exemple, son déterminant).

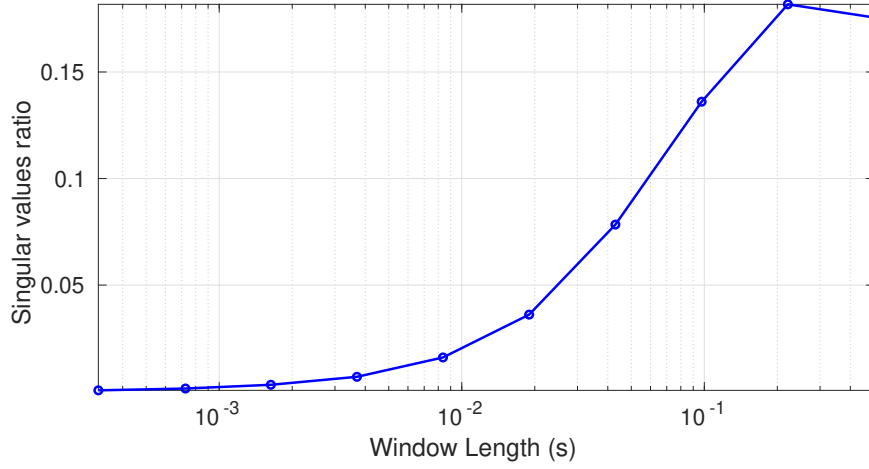


FIGURE 3.2 – Estimation du nombre de sources. Évolution du rapport entre les deux dernières valeurs singulières en fonction de la longueur des portions analysées.

Lors d'un mélange linéaire instantané de sources stationnaires, une décomposition en valeurs singulières est appliquée à la matrice des observations : $\mathbf{Z} = (\mathbf{z}(\tau))_{\tau \in \mathcal{T}} \in \mathbb{R}^{N_z \times N_\tau}$. On obtient $\mathbf{Z} = \mathbf{U}\mathbf{\Delta}\mathbf{V}^T$ où $\mathbf{U} \in \mathbb{R}^{N_z \times N_z}$ et $\mathbf{V} \in \mathbb{R}^{N_\tau \times N_\tau}$ sont des matrices orthogonales, $\mathbf{\Delta} \in \mathbb{R}^{N_z \times N_\tau}$ est une matrice dont seuls les éléments diagonaux peuvent être non nuls : ils sont positifs ou nuls. Ces coefficients sont appelés valeurs singulières de \mathbf{Z} . Le nombre de valeurs singulières non nulles de \mathbf{Z} est égal à son rang. Or, par construction, le rang de \mathbf{Z} vaut N_y . En pratique, toutes les valeurs singulières de \mathbf{Z} sont non nulles (présence de bruit, précision numérique, ...). L'estimation de N_y est alors obtenue en comptant le nombre de valeurs singulières de \mathbf{Z} supérieures à un seuil choisi.

Dans le cas qui nous intéresse, le mélange étant dépendant du temps, on se ramène au cas du mélange constant en effectuant des décompositions en valeurs singulières sur des courtes portions successives des signaux observés où l'on suppose la matrice de mélange constante. Ceci étant justifié par l'équation d'approximation (3.15). Alors, chaque portion analysée nous donne une estimation de N_y . On choisit finalement d'estimer le nombre de sources en conservant la valeur la plus souvent obtenue lors des décomposition en valeurs singulières successives. Pour illustrer ce choix, on prend pour exemple une situation où $N_y = 2$ et $N_z = 3$. On trace en figure 3.2 le rapport entre la plus petite valeur singulière et la deuxième valeur singulière en fonction de la longueur des portions analysées. On observe que pour une longueur de portion suffisamment petite, la dernière valeur singulière est bien plus faible que les deux premières. En effet, si les portions sont trop longues, l'hypothèse de stationnarité approchée sur cette portion n'est plus valide et l'information apportée par la décomposition en valeurs singulières n'est pas exploitable pour déterminer N_y . C'est pourquoi en choisissant un seuil adapté (par exemple 3% de la seconde valeur singulière) et une longueur de portion assez faible (moins de 10 ms dans cet exemple), on estime bien que le nombre de sources est égal à

2.

Enfin, la procédure d'estimation décrite en section 3.5.2 est appliquée à N_y observations choisies parmi les N_z disponibles. Dans la section 3.5.2, nous présentons la procédure d'estimation des sources une fois la réduction de dimension effectuée. Nous verrons ensuite, dans la section 3.5.3 comment l'algorithme global permet de profiter de l'information supplémentaire apportée par les $N_z - N_y$ sources non exploitées lors de la procédure d'estimation.

3.5.2 Procédure d'estimation

Dorénavant, on considère que le nombre d'observations N_z est égal au nombre de sources N_y . La matrice de mélange est alors supposée inversible. Par souci de simplicité, nous introduisons les notations suivantes. Soit $\mathbf{B}_\tau = \mathbf{A}(\tau)^{-1}$ la matrice de séparation. Soit $\theta_{i,\tau} = \log_q(\gamma'_i(\tau))$ le paramètre de déformation à l'instant τ de la i^e source et $\boldsymbol{\theta}_\tau = (\theta_{1,\tau} \cdots \theta_{N_y,\tau})^T$ le vecteur contenant l'ensemble de ces paramètres pour chacune des sources.

Désormais, nous supposons que les conditions de validité du théorème 2.1 sont vérifiées. Il découle de cette hypothèse que la transformée en ondelettes de la i^e source Y_i peut être approchée par la formule d'approximation (2.5). Afin de pouvoir écrire la vraisemblance des coefficients d'ondelettes des observations en fonction des paramètres inconnus, les processus stationnaires X_i , sous-jacents à chacune des sources, sont dorénavant supposés gaussiens. Alors, à τ fixé, la colonne de la transformée en ondelettes de Y_i est un vecteur aléatoire gaussien, complexe circulaire (cf. la proposition 2.1). On a donc $\mathbf{w}_{Y_i,\tau} \sim \mathcal{CN}_c(\mathbf{0}, \mathbf{C}_i(\theta_{i,\tau}))$, avec

$$[\mathbf{C}_i(\theta_{i,\tau})]_{kk'} = q^{\frac{s_k + s_{k'}}{2}} \int_{\mathbb{R}} \mathcal{S}_{X_i}(q^{-\theta_{i,\tau}} \xi) \bar{\psi}(q^{s_k} \xi) \hat{\psi}(q^{s_{k'}} \xi) d\xi.$$

Par ailleurs, comme l'équation de mélange (3.15) approchée dans le domaine des ondelettes est supposée vraie, la densité de probabilité de $\mathbf{w}_{\mathbf{z},\tau}$ s'exprime en fonction de $\mathbf{w}_{\mathbf{y},\tau}$ par

$$p_{\mathbf{w}_{\mathbf{z},\tau}}(\mathbf{w}_{\mathbf{z},\tau}) = |\det(\mathbf{B}_\tau)|^{2M_s} p_{\mathbf{w}_{\mathbf{y},\tau}}(\mathbf{B}_\tau \mathbf{w}_{\mathbf{z},\tau}).$$

Alors, comme l'hypothèse de décorrélation des sources est équivalente à une hypothèse d'indépendance dans le cas gaussien, cela conduit à la log-vraisemblance des observations suivante :

$$\begin{aligned} \mathcal{L}_\tau(\mathbf{B}_\tau, \boldsymbol{\theta}_\tau) &\triangleq \ln(p_{\mathbf{w}_{\mathbf{z},\tau}|\mathbf{B}_\tau, \boldsymbol{\theta}_\tau}(\mathbf{w}_{\mathbf{z},\tau}; \mathbf{B}_\tau, \boldsymbol{\theta}_\tau)) + N_y M_s \ln(\pi) \\ &= 2M_s \ln |\det(\mathbf{B}_\tau)| + \sum_{i=1}^{N_y} \ln(p_{\mathbf{w}_{Y_i,\tau}|\theta_{i,\tau}}(\mathbf{w}_{Y_i,\tau}; \theta_{i,\tau})) + N_y M_s \ln(\pi) \\ &= 2M_s \ln |\det(\mathbf{B}_\tau)| - \sum_{i=1}^{N_y} \ln |\det(\mathbf{C}_i(\theta_{i,\tau}))| \\ &\quad - \sum_{i=1}^{N_y} [\mathbf{B}_\tau \mathbf{w}_{\mathbf{z},\tau}]_i \cdot \mathbf{C}_i(\theta_{i,\tau})^{-1} [\mathbf{B}_\tau \mathbf{w}_{\mathbf{z},\tau}]_i^H, \end{aligned} \quad (3.18)$$

où $[\mathbf{M}]_i$ indique la i^e ligne de la matrice \mathbf{M} . Les estimations du maximum de vraisemblance sont les maximiseurs de $\mathcal{L}_\tau(\mathbf{B}_\tau, \boldsymbol{\theta}_\tau)$, ils sont notés respectivement $\tilde{\mathbf{B}}_\tau$ et $\tilde{\boldsymbol{\theta}}_\tau$. Par ailleurs, les spectres \mathcal{S}_{X_i} des différentes sources stationnaires sous-jacentes interviennent de manière cachée dans l'expression de la vraisemblance : leur connaissance est effectivement nécessaire pour le calcul des matrices de covariance $\mathbf{C}_i(\theta_{i,\tau})$. À l'image de l'algorithme JEFAS présenté dans le chapitre précédent, la stratégie d'estimation sera ici d'alterner l'estimation de la matrice de séparation instantanée et celles des déformations et des spectres associés.

(a) Estimation de la matrice de séparation. Supposons que le paramètre de déformation $\boldsymbol{\theta}_\tau$ et les spectres \mathcal{S}_{X_i} sont connus (en réalité seules les estimées $\tilde{\boldsymbol{\theta}}_\tau$ et $\tilde{\mathcal{S}}_{X_i}$ sont connues). Les différentes matrices de covariance $\mathbf{C}_i(\theta_{i,\tau})$ sont donc connues. Afin de simplifier les calculs, réécrivons la log-vraisemblance

$$\begin{aligned} \mathcal{L}_\tau(\mathbf{B}_\tau, \boldsymbol{\theta}_\tau) = & 2M_s \ln |\det(\mathbf{B}_\tau)| - \sum_{i=1}^{N_y} \ln (\det(\mathbf{C}_i(\theta_{i,\tau}))) \\ & - \sum_{i=1}^{N_y} \mathbf{J}_i^T \mathbf{B}_\tau \mathbf{w}_{z,\tau} \mathbf{C}_i(\theta_{i,\tau})^{-1} \mathbf{w}_{z,\tau}^H \mathbf{B}_\tau \mathbf{J}_i \end{aligned}$$

où $\mathbf{J}_i \in \mathbb{R}^{N_y}$ désigne le vecteur colonne tel que $[\mathbf{J}_i]_k = \mathbb{1}_i(k)$ (ici, la notation $\mathbb{1}_i$ désigne la fonction indicatrice de la valeur i). On peut alors dériver la log-vraisemblance par rapport à \mathbf{B}_τ pour obtenir l'équation qui doit être vérifiée par cette matrice. On note cette dérivée $\nabla_{\mathbf{B}_\tau} \mathcal{L}_\tau$. Alors, en appliquant les formules de dérivées matricielles (57) et (118) dans *The Matrix Cookbook* [57], on obtient :

$$\nabla_{\mathbf{B}_\tau} \mathcal{L}_\tau(\mathbf{B}_\tau, \boldsymbol{\theta}_\tau) = 2M_s \left(\mathbf{B}_\tau^{-1} \right)^T - \sum_{i=1}^{N_y} \mathbf{J}_i \mathbf{J}_i^T \mathbf{B}_\tau 2 \operatorname{Re}(\mathbf{w}_{z,\tau} \mathbf{C}_i(\theta_{i,\tau})^{-1} \mathbf{w}_{z,\tau}^H) .$$

L'estimateur du maximum de vraisemblance vérifie alors $\nabla_{\mathbf{B}_\tau} \mathcal{L}_\tau(\tilde{\mathbf{B}}_\tau, \boldsymbol{\theta}_\tau) = \mathbf{0}$. On a donc :

$$\frac{1}{M_s} \sum_{i=1}^{N_y} \mathbf{J}_i \mathbf{J}_i^T \operatorname{Re} \left(\tilde{\mathbf{B}}_\tau \mathbf{w}_{z,\tau} \mathbf{C}_i(\theta_{i,\tau})^{-1} (\tilde{\mathbf{B}}_\tau \mathbf{w}_{z,\tau})^H \right) = \mathbf{I}$$

Comme les matrices $\mathbf{C}_i(\theta_{i,\tau})^{-1}$ sont semi-définies positives, cette condition se réécrit :

$$\frac{1}{M_s} \tilde{\mathbf{w}}_{y_i,\tau} \mathbf{C}_i(\theta_{i,\tau})^{-1} \tilde{\mathbf{w}}_{y_i,\tau}^H = \mathbb{1}_i(j), \quad \forall i, j \in \{1, \dots, N_y\} . \quad (3.19)$$

où $\tilde{\mathbf{w}}_{y_i,\tau} = [\tilde{\mathbf{B}}_\tau \mathbf{w}_{z,\tau}]_i$ est l'estimation de la transformée en ondelettes à l'instant τ de la i^e source. Cette équation montre que lorsque $\tilde{\mathbf{B}}_\tau$ est proche de \mathbf{B}_τ , les sources estimées sont décorréliées et la matrice de covariance de $\tilde{\mathbf{w}}_{y_i,\tau}$ doit être proche de la véritable matrice de covariance de la transformée en ondelettes $\mathbf{w}_{y_i,\tau}$, c'est-à-dire $\mathbf{C}_i(\theta_{i,\tau})$. Par

ailleurs, la relation (3.19) fournit N_y^2 équations, et nous devons estimer N_y^2 paramètres (les éléments de la matrice \mathbf{B}_τ). De plus, ces équations sont quadratiques par rapport aux éléments de \mathbf{B}_τ . C'est pourquoi la matrice $\tilde{\mathbf{B}}_\tau$ peut être obtenue de façon approchée par une méthode de Newton (voir le chapitre 4 dans [8]). Pour cela, on considère la variable vectorielle $\mathbf{b}_\tau = \text{Vec}(\mathbf{B}_\tau^T) \in \mathbb{R}^{N_y^2}$ où l'opération « Vec » consiste à concaténer les colonnes d'une matrice dans un vecteur colonne. Ici, en notant β_i la i -ème ligne de \mathbf{B}_τ , on a donc : $\mathbf{b}_\tau = (\beta_1 \cdots \beta_{N_y})^T$. Il s'agit alors d'annuler la fonction vectorielle $\mathbf{f} : \mathbb{R}^{N_y^2} \rightarrow \mathbb{R}^{N_y^2}$ dont la n -ème composante est définie par :

$$\mathbf{f}_n(\mathbf{b}_\tau) = \frac{1}{M_s} \beta_j \mathbf{w}_{z,\tau} \mathbf{C}_i(\theta_{i,\tau})^{-1} \tilde{\mathbf{w}}_{z,\tau}^H \beta_i^T - \mathbb{1}_i(j) \quad \text{où} \quad n = N_y(i-1) + j, \quad (3.20)$$

pour tout $i, j \in \{1, \dots, N_y\}$. On donne dans l'algorithme 3.2 le pseudo-code associé à la méthode de Newton permettant de donner une solution approchée de l'équation $\mathbf{f}(\mathbf{b}) = \mathbf{0}$. Cet algorithme nécessite le calcul de la matrice jacobienne \mathbf{J}_f de \mathbf{f} . C'est une matrice carrée de taille $N_y^2 \times N_y^2$ définie par $[\mathbf{J}_f]_{mn} = \partial_m \mathbf{f}_n$ où ∂_m désigne la dérivation partielle par rapport à la m -ème composante de \mathbf{b}_τ . La solution de l'équation $\mathbf{f} = \mathbf{0}$

Algorithme 3.2 Méthode de Newton : $\tilde{\mathbf{B}} = \text{Newton}(\mathbf{B}_0, \epsilon_N)$

```

l ← 0
b(0) = Vec(B0)
tant que ||b(l+1) - b(l)|| > εN faire
    b(l+1) = b(l) - Jf(b(l))f(b(l))
    l ← l + 1
fin tant que
Sortie :  $\tilde{\mathbf{B}} = \text{Vec}^{-1}(\mathbf{b}^{(l)})$ 

```

n'est pas unique. En particulier, les indéterminations du problème de séparation de sources (liées à l'ordre arbitraire des sources) sont toujours présentes. C'est pourquoi on détaillera dans la section 3.5.3 la stratégie mise en œuvre pour ne pas modifier l'ordre des sources estimées tout au long de l'algorithme d'estimation.

(b) Estimation des déformations et des spectres associés. Supposons que la matrice de séparation \mathbf{B}_τ est connue à chaque instant (en réalité, seule son estimée $\tilde{\mathbf{B}}_\tau$ est connue). Dans l'expression de la log-vraisemblance (3.18), seules les matrices $\mathbf{C}_i(\theta_{i,\tau})$ dépendent des déformations et des spectres associés. Or, grâce à la décorrélation des sources, cette expression est séparable vis-à-vis des différentes matrices de covariance $\mathbf{C}_i(\theta_{i,\tau})$. On peut donc estimer les déformations et les spectres de chaque source indépendamment les uns des autres. Pour la i -ème source, il faut donc maximiser la fonction :

$$\mathcal{L}_{i,\tau}(\theta_\tau) = -\log |\det(\mathbf{C}_i(\theta_{i,\tau}))| - [\mathbf{B}_\tau \mathbf{w}_{z,\tau}]_i \mathbf{C}_i(\theta_{i,\tau})^{-1} [\mathbf{B}_\tau \mathbf{w}_{z,\tau}]_i^H.$$

Il s'agit de maximiser exactement la même vraisemblance que celle obtenue au chapitre 2 (voir l'équation (2.12)) en remplaçant la transformée en ondelettes de la source y_i

(à laquelle nous n'avons pas accès) par sa transformée en ondelettes approchée à partir de l'équation d'approximation (3.15), c'est-à-dire $[\mathbf{B}_\tau \mathbf{w}_{z,\tau}]_{i\cdot}$. La procédure d'estimation de $\theta_{i,\tau}$ sera donc identique à celle présentée au chapitre 2.

3.5.3 Algorithme

Nous venons de décrire la procédure d'estimation dans le cas où $N_z = N_y$. Dans le cas où $N_z > N_y$, nous avons accès à $\binom{N_z}{N_y}$ combinaisons de N_y signaux sources différentes pour effectuer la séparation de sources. On choisit d'appliquer l'algorithme de séparation de sources à chacune de ces combinaisons. On se ramène ainsi à la résolution de $\binom{N_z}{N_y}$ problèmes de séparation de sources successifs pour lesquels les nombres d'observations et de sources sont égaux. Notons \mathbf{z}_l le vecteur contenant les observations de la l -ème combinaison. De même, on note $\mathbf{B}_l \in \mathbb{R}^{N_y \times N_y}$ la matrice de séparation contenant les N_y colonnes correspondant aux observations conservées.

L'algorithme 3.3, nommé JEFAS-BSS, synthétise toutes les étapes de l'estimation. En ce qui concerne l'initialisation de la séparation de sources, nous utilisons le résultat de p-SOBI (cf. section 3.2.3) qui fournit une méthode de base donnant une matrice de mélange initiale variant dans le temps.

Notons $\tilde{\mathbf{B}}_{l,\tau^{(k-1)}}$ l'estimée de la matrice de séparation après $(k-1)$ itérations de l'algorithme JEFAS-BSS. Ensuite, l'algorithme 2.2 (JEFAS) est mis en œuvre à partir de chacune des estimées des transformées en ondelettes des sources données, à chaque instant τ , par $[\tilde{\mathbf{B}}_{l,\tau}^{(k-1)} \mathbf{w}_{z_l,\tau}]_{i\cdot}$. Il permet l'estimation des déformations temporelles $\tilde{\theta}_i$ et des spectres \mathcal{S}_{X_i} . Remarquons également que l'estimation de $\mathbf{B}_{l,\tau}$ s'effectue sur une grille d'instantanés espacés d'un pas temporel Δ_τ . L'estimation des sources à chaque instant s'effectue alors en considérant la matrice de séparation $\mathbf{B}_{l,\tau}$ constante sur l'intervalle $I_\tau = [\tau - \Delta_\tau/2, \tau + \Delta_\tau/2[$. En outre, l'estimation de la matrice de séparation s'effectue de la même manière en remplaçant les vraies valeurs de θ_τ et \mathcal{S}_X (qui sont inconnues) par leurs estimations à l'itération k dans la relation (3.19) que nous approchons grâce à la méthode de Newton.

Enfin, la convergence est contrôlée au moyen du rapport source à interférence (SIR) défini par (3.9). Comme nous n'avons pas accès aux véritables sources, nous utilisons comme critère d'arrêt le SIR entre $\tilde{\mathbf{y}}^{(k-1)}$ et $\tilde{\mathbf{y}}^{(k)}$ (au lieu de \mathbf{y}) qui donne une évaluation de la mise à jour de la BSS et constitue donc une évaluation de convergence pertinente. Pratiquement, le critère de convergence est testé sur le SIR moyenné entre les sources. Il est défini comme suit :

$$\frac{1}{N_y} \sum_{i=1}^{N_y} \text{SIR} \left(\tilde{y}_i^{(k-1)}, \tilde{\mathbf{y}}^{(k)} \right) > \Lambda, \quad (3.21)$$

où Λ est choisi suffisamment grand pour que l'amélioration de la qualité de la BSS entre la $(k-1)$ et la k ème itération de JEFAS-BSS ne soit pas significative.

Une fois que chacun des $\binom{N_z}{N_y}$ algorithmes de séparation de sources a convergé, on estime la matrice de séparation globale $\tilde{\mathbf{B}}_\tau$ en moyennant chaque colonne par l'en-

Algorithme 3.3 JEFAS-BSS

- Estimer le nombre de sources N_y (voir section 3.5.1).
- pour** chacune des combinaisons d'observations parmi $\binom{N_y}{N_z}$ **faire**
 - Évaluer $\tilde{\mathbf{B}}_{l,\cdot}^{(0)} = \text{p-SOBI}(\mathbf{z}_l, T_f)$ (voir algorithme 3.1).
 - Calculer les sources estimées $\tilde{\mathbf{y}}^{(0)}(\tau) = \tilde{\mathbf{B}}_{l,\tau}^{(0)} \mathbf{z}_l(\tau)$.
 - $k \leftarrow 1$
 - tant que** le critère d'arrêt (3.21) est faux **et** $k \leq k_{\max}$ **faire**
 - pour** $i = 1, \dots, N_y$ **faire**
 - Estimer les paramètres $\tilde{\theta}_{i,\tau}^{(k)}, \forall \tau \in \tau$ et le spectre $\tilde{\mathcal{S}}_{X_i}^{(k)}$ en appliquant JEFAS à $\tilde{\mathbf{y}}_i^{(k-1)}$ (voir algorithme 2.2).
 - fin pour**
 - pour** $\tau = 0, \Delta_\tau, \dots, T$ **faire**
 - Calculer $\tilde{\mathbf{B}}_{l,\tau}^{(k)} = \text{Newton}(\tilde{\mathbf{B}}_{l,\tau}^{(k-1)}, \epsilon_N)$ (voir l'algorithme 3.2) en remplaçant θ_τ et $\{\mathcal{S}_{X_i}\}_{i=1,\dots,N_y}$ par leurs estimations actuelles $\tilde{\theta}_\tau^{(k)}$ et $\{\tilde{\mathcal{S}}_{X_i}^{(k)}\}_{i=1,\dots,N_y}$ dans l'expression de \mathbf{f} (3.20).
 - fin pour**
 - Estimer les sources $\tilde{\mathbf{y}}^{(k)}(\tau) = \tilde{\mathbf{B}}_{l,\tau}^{(k)} \mathbf{z}_l(\tau)$.
 - $k \leftarrow k + 1$
 - fin tant que**
 - $\tilde{\Theta}_l = \left(\tilde{\theta}_\tau^{(k)} \right)_{\tau \in \tau}$
 - $\forall i \in \{1, \dots, N_y\}, \tilde{\mathcal{S}}_{X_i,l} = \tilde{\mathcal{S}}_{X_i}^{(k)}$
 - fin pour**
 - Estimer la matrice de séparation globale au moyen de la formule (3.22).
 - Estimer les fonctions de déformation et les spectre en moyennant leurs estimations (par rapport à l).

semble des $\binom{N_y-1}{N_z-1}$ estimations que l'on a fait de celle-ci. Autrement dit, on a :

$$\tilde{\mathbf{B}}_\tau = \frac{1}{\binom{N_y-1}{N_z-1}} \sum_{l=1}^{\binom{N_y}{N_z}} \tilde{\mathbf{B}}_{l,\tau} \mathbf{E}_l, \quad (3.22)$$

où $\mathbf{E}_l \in \mathbb{R}^{N_y \times N_z}$. La n -ème colonne de la matrice \mathbf{E}_l est non nulle si et seulement si la n -ème observation de \mathbf{z} est présente dans la l -ème combinaison analysée. Dans ce cas, cette colonne contient uniquement un élément non nul égal à 1 permettant de placer la colonne $\tilde{\mathbf{B}}_{l,\tau}$ au bon endroit dans $\tilde{\mathbf{B}}_\tau$.

D'autre part, cette stratégie d'estimation nous donne accès à $\binom{N_y}{N_z}$ estimations de chacune des fonctions de déformation γ_i et de chacun des spectres \mathcal{S}_{X_i} . On peut alors obtenir de nouvelles estimées de variance plus faible en moyennant toutes ces estimations.

3.6 Résultats numériques

3.6.1 Application à un mélange synthétique

Nous construisons un exemple synthétique avec $N_y = N_z = 3$ pour évaluer les performances de l'algorithme JEFAS-BSS. La durée des signaux est de 1 seconde et ils sont échantillonnés à $F_s = 44,1$ kHz. Les deux sources gaussiennes sont non stationnaires selon le modèle (3.13), leurs spectres de puissance sous-jacents \mathcal{S}_{X_i} sont constitués de différentes fenêtres de Hann ne se chevauchant pas. Les coefficients de la matrice de mélange évoluent sinusoïdalement dans le temps (avec des fréquences différentes). Les transformées en ondelettes des sources et des observations sont affichées sur la figure 3.3. Leurs supports étant superposés, un algorithme de BSS s'appuyant sur la parcimonie d'une représentation des signaux ne serait pas efficace ici.

L'algorithme JEFAS-BSS est appliqué aux observations, il converge en 5 itérations. Les scalogrammes des sources estimées sont affichés sur la figure 3.4. Pour éviter la redondance avec les résultats du chapitre 2, les fonctions de déformation temporelle et les spectres estimés par JEFAS ne sont pas tracés ici. Pour une étude complète des performances de JEFAS, nous renvoyons le lecteur au chapitre 2. Dans ce qui suit, nous nous concentrons sur les performances de la séparation de sources.

Afin d'évaluer la BSS obtenue via JEFAS-BSS, nous la comparons à d'autres algorithmes de séparation de sources (SOBI, p-SOBI, QTF-BSS). Nous appliquons ces algorithmes à 20 réalisations de l'exemple synthétique ci-dessus. Pour évaluer la qualité des algorithmes BSS, nous calculons le SIR et le SDR entre les vraies sources et leurs estimations en moyenne sur 20 simulations (voir les formules (3.9) et (3.10) pour les définitions de ces indices de performance). Les résultats ainsi que les écarts types (ET) correspondants sont donnés dans le tableau 3.1. Néanmoins, cette quantité est un indicateur global de la qualité de la séparation de sources. Afin de suivre l'évolution temporelle de la qualité de la BSS, nous évaluons à chaque instant l'indice d'Amari normalisé $\rho(t)$ (défini par (3.12)). L'évolution de ρ (en décibels) pour les algorithmes de

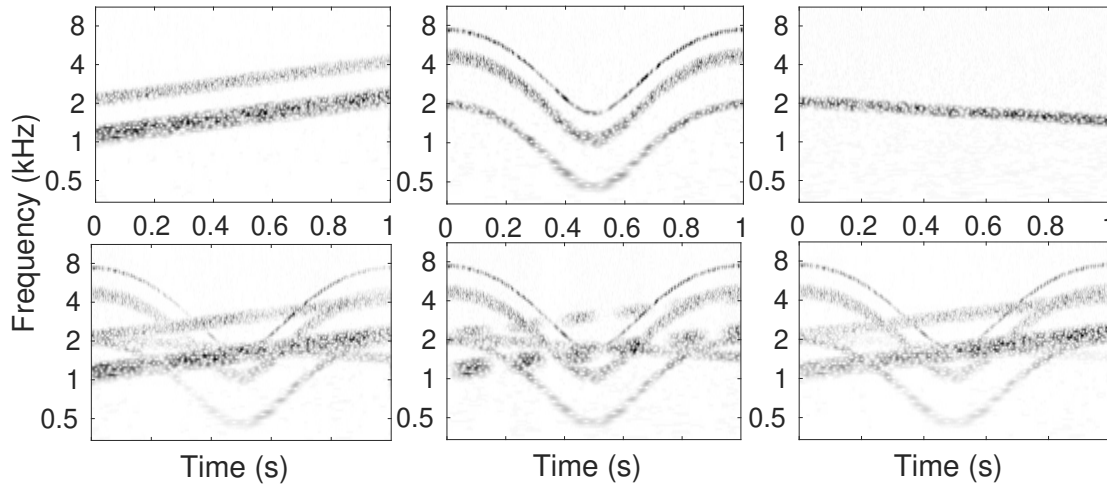


FIGURE 3.3 – Mélange synthétique. En haut : scalogrammes des trois sources non stationnaires. En bas : scalogrammes des trois observations.

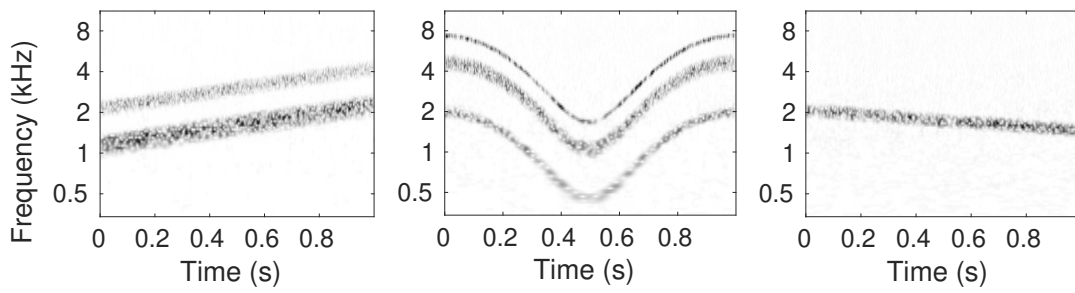


FIGURE 3.4 – Mélange synthétique. Scalogrammes des trois sources estimées.

TABLEAU 3.1 – Mélange synthétique (20 réalisations). Comparaison du SiR, du SDR et de l'indice d'Amari moyen pour quatre algorithmes de BSS.

Algorithme	SIR (dB)		SDR (dB)		ρ moyen (dB)	
	Moyenne	ET	Moyenne	ET	Moyenne	ET
SOBI	11,82	3,65	6,20	1,94	-6,54	0,86
p-SOBI	3,58	1,93	-7,99	1,91	-9,06	0,21
QTF-BSS	0,79	3,88	-2,12	3,25	-3,87	0,42
JEFAS-BSS	30,26	2,37	16,30	2,54	-15,36	0,70

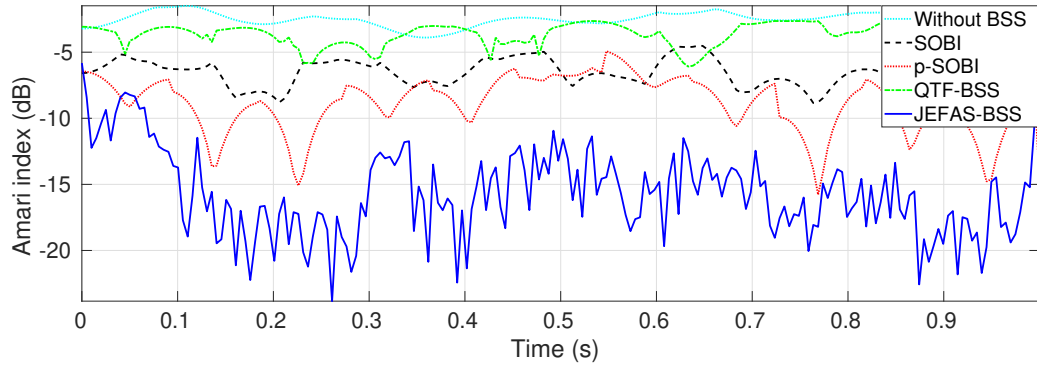


FIGURE 3.5 – Mélange synthétique. Évolution temporelle de l'indice d'Amari pour différentes méthodes de séparation de sources.

séparation de sources que nous évaluons est tracée sur la figure 3.5. D'autres part, les moyennes et écarts-types temporels de cet indice sont mesurés. Les valeurs moyennes obtenues sur 20 simulations sont données dans le tableau 3.1.

On peut alors comparer les performances des différents algorithmes de séparation de sources dans cette situation.

JEFAS-BSS. Comme on pouvait s'y attendre, l'indice d'Amari de JEFAS-BSS est nettement inférieur à ceux des autres méthodes de séparation de sources. En effet, la connaissance *a priori* du modèle de non-stationnarité des sources avantage cet algorithme. On peut noter que l'indice d'Amari reste faible (aux alentours de -15 dB) en dehors des plages de temps initiales et finales (de 0 à 0,1 s et de 0,9 à 1 s) où la qualité de la séparation de sources décroît : la valeur de l'indice d'Amari remonte autour de -10 dB. Ceci est probablement principalement dû aux effets de bords.

p-SOBI. C'est le deuxième algorithme le plus performant au niveau de l'indice d'Amari. On observe des ruptures de pente brutales de l'évolution de l'indice d'Amari. Cela est dû à la construction de l'algorithme p-SOBI où la matrice de séparation estimée est constante par morceaux. Le passage d'un segment à l'autre peut donc entraîner une modification importante de la matrice de sépa-

ration, et par conséquent, de l'indice d'Amari. L'effet de cette segmentation de la BSS est également visible sur le SDR de p-SOBI qui est négatif (-2 dB environ) car la mise bout à bout des segments de source estimées entraîne des discontinuités sur le signal reconstitué. Ces distorsions proviennent des indéterminations sur l'amplitude des sources, elles entraînent la baisse de la valeur du SDR.

SOBI. Cet algorithme étant adapté au modèle stationnaire, il est global. C'est pourquoi la valeur de l'indice d'Amari de SOBI est haute. De même, le SIR est faible car la séparation de sources est mal réalisée mais le SDR reste supérieur à p-SOBI (6 dB environ). Cela met en évidence le fait qu'un compromis est nécessaire sur le choix de la taille de la fenêtre de p-SOBI. Une fenêtre trop petite entraîne un SDR très faible mais l'indice d'Amari sera bon. À l'inverse, une fenêtre trop large entraîne un mauvais indice d'Amari mais le SDR reste raisonnable.

QTF-BSS. Plus surprenant, les performances QTF-BSS sont inférieures à celles de SOBI. Par exemple, la valeur de l'indice d'Amari est seulement légèrement inférieure au cas où l'on effectue aucune séparation de sources, c'est-à-dire $\hat{\mathbf{B}}_\tau = \mathbf{I}$ pour tout τ . Cela peut être dû au fait que, comme pour SOBI, l'estimation de QTF-BSS repose sur un modèle matriciel de mélange constant. De plus, cet exemple de mélange est particulièrement défavorable à la méthode QTF-BSS car les scalogrammes des différentes sources se superposent. En particulier, il existe très peu de points du plan temps-fréquence (ou temps-échelle) où la troisième source est présente seule. L'étape de sélection des points dans QTF-BSS devient donc délicate.

3.6.2 Application à un mélange de signaux audios

Nous nous intéressons maintenant au mélange synthétique de signaux audios qui a été présenté dans l'introduction de ce chapitre. Les transformées en ondelettes des signaux observés sont représentées figure 3.1. Ce mélange est intéressant car il est composé de signaux sources dont les transformées en ondelettes se superposent sur de grandes zones temps-échelle. Il permet donc de mettre en évidence la capacité de JEFAS-BSS à séparer des sources large bande. Les transformées en ondelettes des deux sources estimées sont représentées figure 3.6. Le scalogramme correspondant au son de vent estimé est à gauche, alors que celui correspondant au chant est à droite. L'algorithme JEFAS-BSS nous fournit également les déformations et les spectres estimés pour chacune de deux sources. Les deux spectres estimés sont superposés figure 3.7. Cette figure montre la superposition des contenus spectraux des deux sources, notamment entre 0 et 2 kHz (au delà, les valeurs du spectre du chant sont encore significatives alors que les valeurs du spectre du vent ne le sont plus). D'autre part, comme pour le mélange synthétique de signaux synthétiques, les indices de performances de la BSS montrent une nette amélioration de la qualité de la BSS par rapport à p-SOBI ou QTF-BSS. En effet, l'indice d'Amari de JEFAS-BSS vaut -21 dB contre $-17,3$ dB pour p-SOBI, $-7,4$ dB pour SOBI et -4 dB pour QTF-BSS. De plus, le SIR de JEFAS-BSS vaut 44,8 dB contre

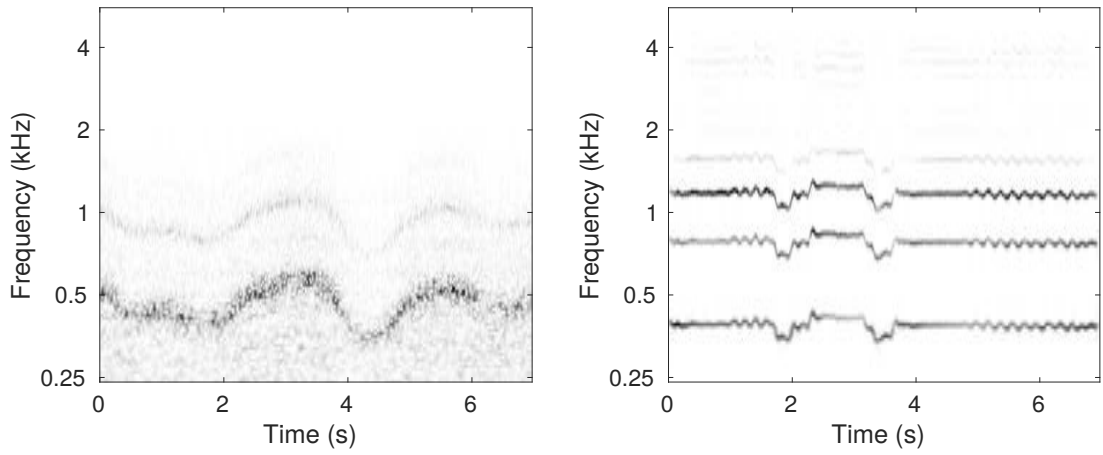


FIGURE 3.6 – Mélange de signaux audios. Modules des transformées en ondelettes de deux sources estimées par JEFAS-BSS.

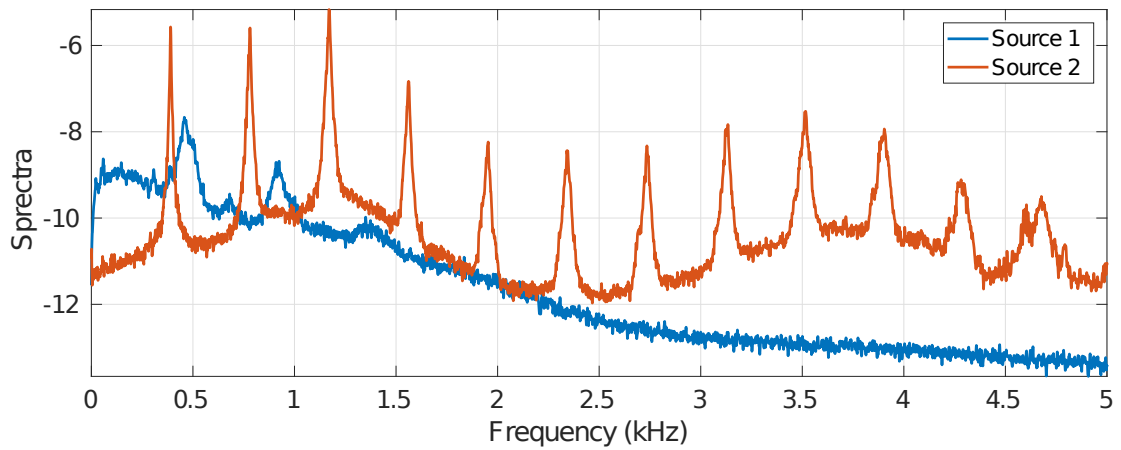


FIGURE 3.7 – Mélange de signaux audios. Estimations des spectre des sources stationnaires sous-jacents.

5,6 dB pour p-SOBI, 20,5 dB pour SOBI et 7,6 dB pour QTF-BSS.

3.7 Bilan

Dans ce chapitre, nous avons proposé une méthode de séparation de sources doublement non stationnaire. Le modèle que nous avons considéré est celui du mélange linéaire instantané de signaux localement dilatés par une matrice de mélange variant dans le temps. Ce modèle de mélange est fait l’extension du modèle du chapitre 2 aux signaux multi-capteurs et à une opération de modulation d’amplitude instantanée matricielle. L’objectif que nous nous sommes fixé est celui de l’analyse spectrale de ces signaux multi-capteurs localement déformés, c’est-à-dire l’estimation jointe de la matrice de mélange, des fonctions de dilatation instantanées, et des densités spectrales des signaux stationnaires sous-jacents. C’est pourquoi l’algorithme de séparation de sources que nous proposons opère dans le domaine des ondelettes, et est fortement inspiré de JEFAS. Par rapport à cet algorithme, les principales extensions apportées à JEFAS pour l’adapter à ce nouveau cadre sont :

- l’étape préalable de réduction de dimension. Cette étape nous permet, de considérer des situations où le nombre de signaux observés est supérieur au nombre de signaux sources.
- l’étape d’estimation de la matrice de séparation. Contrairement à l’étape d’estimation de la modulation d’amplitude dans JEFAS, nous n’avons pas d’expression explicite de la matrice de séparation maximisant la log-vraisemblance. Nous procédons alors à l’estimation de cette matrice à l’aide d’une méthode de Newton.

Par ailleurs, JEFAS-BSS diffère de JEFAS dans la mesure où il requiert une étape initiale de séparation de sources. Nous envisageons de modifier la fonction à maximiser de manière à se passer de cette étape d’initialisation (se reporter les perspectives de la thèse en section 6.2).

Finalement, nous avons vu que JEFAS-BSS permet d’améliorer nettement les performances des méthodes de séparation de sources existantes. Ce résultat est principalement dû à la connaissance supplémentaire apportée par la connaissance du modèle de non-stationnarité. Notons que l’application numérique de JEFAS-BSS à une situation où le nombre d’observations est supérieur au nombre de sources n’a pas été proposée dans ce chapitre. En effet, l’étude de cette situation n’apporte pas d’informations supplémentaire sur l’algorithme, seule l’étape initiale d’estimation du nombre de sources étant nouvelle (ensuite il s’agit d’appliquer successivement JEFAS-BSS à plusieurs combinaisons d’observations). Néanmoins, la mise en œuvre de JEFAS-BSS sur un exemple synthétique de ce type est proposée en ligne (cf. annexe B.1).

Dans les chapitres 2 et 3, nous avons décrit des méthodes d’analyse des signaux non stationnaires. Dans les chapitres 4 et 5 à venir, nous allons présenter des méthodes reposant sur des approches à la synthèse de ces signaux. Intéressons-nous d’abord aux signaux localement déformés.

Chapitre 4

SYNTHÈSE DE SIGNAUX LOCALEMENT DÉFORMÉS

4.1 Introduction : de l'analyse à la synthèse

Dans les chapitres 2 et 3, nous avons vu des méthodes d'analyse des signaux localement déformés. En effet, les différentes estimations de spectres instantanés et de matrices de séparation s'appuient sur un outil d'analyse : la transformée en ondelettes. Les différentes grandeurs que nous estimions étaient ensuite extraites de cette analyse temps-échelle. Dans les deux chapitres qui viennent, l'approche est assez différente. En effet, nous proposons ici deux approches à la synthèse. Cela consiste à modéliser le signal à partir d'une formule de synthèse (ou de reconstruction), l'outil à partir duquel le signal est synthétisé sera l'objet recherché. Le modèle de synthèse repose sur le principe des problèmes inverses où l'on recherche ce qui a produit un signal à partir de l'observation de ce dernier. Dans ce chapitre, nous nous focalisons sur une approche à la synthèse probabiliste, tandis que dans le chapitre 5, la synthèse sera étudiée d'un point de vue déterministe. En outre, ce chapitre se concentre sur l'étude des signaux localement dilatés alors que le chapitre suivant considère le cas des signaux localement harmoniques. Dans les deux situations, l'objectif sera de d'estimer une représentation permettant de déterminer le spectre instantané de ces signaux avec le plus de précision possible et suivant certaines contraintes que l'on se fixe (par exemple : une contrainte de régularité ou de parcimonie).

Dans ce chapitre, après avoir détaillé les limites des méthodes d'analyse, nous présenterons quelques méthodes de synthèse temps-fréquence et temps-échelle existantes. Ensuite, nous décrirons le modèle de synthèse temps-échelle probabiliste adapté aux signaux localement déformés. En s'appuyant sur la structure de l'algorithme EM, la méthode d'estimation bayésienne qui en découle nous amène à construire un algorithme d'estimation alterné entre le spectre instantané de tels processus et la représentation temps-fréquence correspondante. Enfin, les performances de cet algorithme, appelé JEFAS-S, sont évaluées et comparées à celle de JEFAS sur une exemple synthétique. La capacité de cette approche à étendre la quantité de signaux analysables est également détaillée sur une exemple semi-réel.

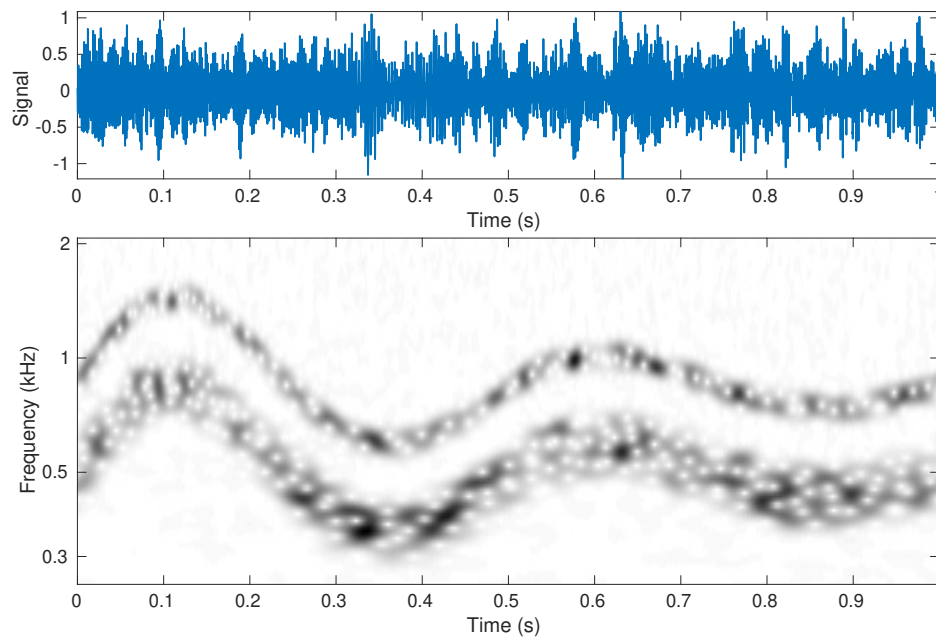


FIGURE 4.1 – Signal synthétique. En haut : chronogramme. En bas : scalogramme.

Au cours de chapitre, nous illustrerons l’allure de certaines grandeurs décrites à partir d’un exemple synthétique de signal localement déformé. Nous donnons en figure 4.1 le chronogramme et le scalogramme de ce dernier (se référer à la section 4.5.1 pour plus de détails sur la construction de ce signal).

4.2 L’approche à la synthèse

4.2.1 Limitations de l’approche à l’analyse

Jusqu’ici, nous avons proposé des méthodes d’analyse spectrale de signaux non stationnaires telles que la transformée en ondelettes ou la transformée de Fourier à court terme. Cependant, celles-ci sont limitées par des contraintes intrinsèques comme le principe d’incertitude ou les interférences.

Principe d’incertitude. Il s’agit d’un principe qui limite la résolution des opérateurs linéaires comme la transformée en ondelettes ou la transformée de Fourier à court terme. Celui-ci stipule que la résolution temps-fréquence de ces représentations ne peut pas être arbitrairement fine. Il peut s’écrire sous différentes formes, on peut se référer à [27] ou [64] pour plus de généralité sur ce sujet. Nous donnons ici un exemple de principe d’incertitude proposé par Gröchenig dans [29]. Celui-ci porte sur la concentration de la transformée de Fourier à court terme. Il est énoncé comme suit.

Théorème 4.1 (Principe d'incertitude pour la TFCT). Soit $0 \leq \epsilon \leq 1$. Soit $U \subseteq \mathbb{R}^2$ une partie du plan temps-fréquence mesurable. Soit $g \in L^2(\mathbb{R})$ une fenêtre d'analyse avec $\|g\|_2 = 1$. Soit $y \in L^2(\mathbb{R})$ un signal avec $\|y\|_2 = 1$. Si

$$\iint_U |\mathcal{G}_y(\xi, \tau)|^2 d\xi d\tau \geq 1 - \epsilon,$$

alors, la mesure $|U|$ de U vérifie l'inégalité suivante :

$$|U| \geq 2(1 - \epsilon)^2.$$

Comme le signal y et la fenêtre g sont unitaires : $\int_{\mathbb{R}^2} |\mathcal{G}_y(\xi, \tau)|^2 d\xi d\tau = \|y\|_2^2 \|g\|_2^2 = 1$. Le résultat du théorème s'interprète donc comme le fait que 95 % de l'énergie du signal ne peut pas être contenue dans une zone du plan temps-fréquence de surface inférieure à 1,8 (correspond au cas $\epsilon = 0,05$). Ce principe rend donc la parcimonie du spectre instantané estimé impossible. Par exemple, l'énergie de la TFCT ne peut pas être concentrée sur une courbe décrivant une fréquence instantanée d'un signal localement harmonique dans le plan temps-fréquence (voir chapitre 5). Un principe d'incertitude similaire limite la précision temps-échelle de la transformée en ondelettes. La partie gauche de la figure 4.2 illustre cette caractéristique de la transformée en ondelettes. Deux chirps linéaires sont présents dans le signal : à chaque instant, les valeurs significatives du scalogramme sont situées dans deux bandes de fréquence assez large situées autour des fréquences instantanées de chacun des chirps, limitant la détermination exacte de celles-ci.

Bulles ou Interférences. Une deuxième limitation de ces méthodes concerne l'apparition d'interférences. En plus de la dispersion des coefficients des représentations temps-fréquence imposée par le principe d'incertitude, celles-ci sont également limitées par la présence d'interférences. Par exemple, la distribution de Wigner-Ville (définition 1.12) d'un chirp linéaire n'est pas limitée par le principe d'incertitude. En effet, comme nous l'avons vu dans l'équation (1.14), on obtient une représentation temps-fréquence idéale où la fréquence instantanée est parfaitement localisée. En revanche, dès que le signal est composé d'au moins deux chirps linéaires, sa distribution de Wigner-Ville fait apparaître des interférences aux fréquences instantanées intermédiaires situées entre les différentes composantes. Des points temps-fréquence pour lesquels on souhaiterait voir le spectre instantané s'annuler sont alors activés (voir l'équation (1.15)). De même, comme illustré par la figure 4.2, lorsque deux composantes fréquentielles sont trop proches, la transformée en ondelettes fait apparaître des « bulles » le long de la fréquence intermédiaire et périodiquement disposées à la fréquence de battement [22]. Il en est de même pour la TFCT. Dans cette illustration, sur la partie droite de la figure 4.2, les fréquences instantanées des deux chirps linéaires sont espacées de 50 Hz. C'est pourquoi 50 « bulles » sont présentes sur un intervalle d'une seconde. Ce même phénomène peut

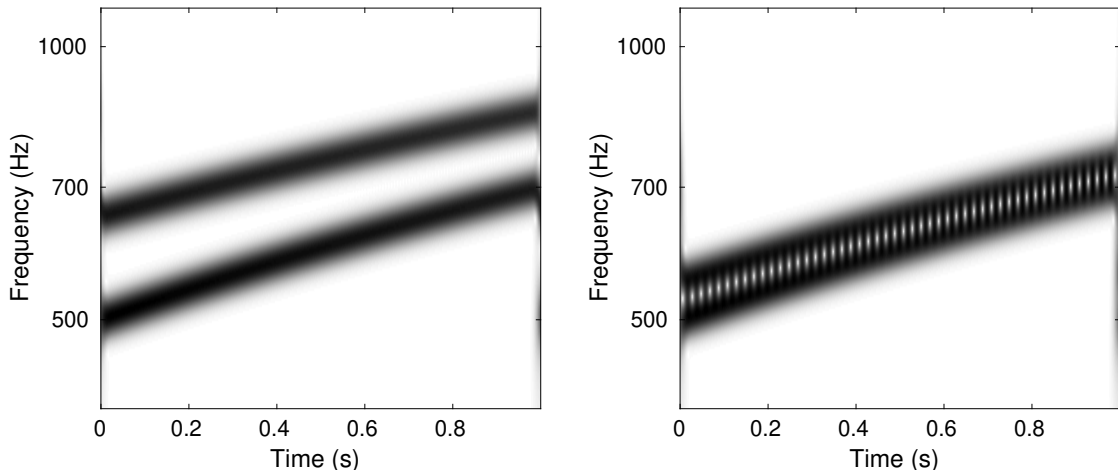


FIGURE 4.2 – Illustration du phénomène d’interférences. À gauche : scalogramme de deux chirps linéaires éloignés. À droite : scalogramme de deux chirps linéaires rapprochés.

aussi apparaître lorsque le spectre instantané d’un signal localement déformé varie rapidement, avec l’apparition d’interférences au niveau temporel.

Par conséquent, JEFAS est limité par ces différents phénomènes. Cela entraîne les conséquences suivantes.

- *Impossibilité d’estimer précisément des spectres parcimonieux.* C’est une conséquence du principe d’incertitude. La résolution fréquentielle de l’ondelette produit un lissage du spectre comme on l’a montré avec l’estimateur du spectre-ondelettes (cf. équation (1.23)).
- *Sensibilité aux interférences.* Si le spectre du processus stationnaire est composé de deux lobes assez proches, on peut voir apparaître des bulles sur la transformée en ondelettes du signal non stationnaire. La transformée en ondelettes du signal stationnaire sous-jacent n’est alors plus invariante par rapport au temps. Le théorème d’approximation 2.1 n’est donc plus exploitable et l’estimation de la dilatation instantanée par JEFAS devient impossible.
- *Sensibilité aux variations rapides.* C’est également une conséquence des conditions de validité du théorème 2.1, lorsque la fonction de dilatation instantanée varie trop rapidement chaque composante du spectre instantané peut interférer avec elle-même. Dans ce cas, l’estimation de JEFAS peut s’avérer délicate voire divergente.

Finalement, l’idée de l’approche à la synthèse est de construire une représentation temps-fréquence ou temps-échelle qui ne soit pas limitée par ces diverses contraintes.

4.2.2 La synthèse probabiliste

L'idée originale de voir la représentation temps-fréquence comme un outil de synthèse probabiliste a été proposée par Turner et Sahani [73]. Dans cet article, l'ensemble des quantités présentées sont données comme des grandeurs discrètes. Ainsi, les signaux sont définis à temps discret, les instants de mesures seront notés $\tau_1, \dots, \tau_{N_\tau}$. De même, l'ensemble des fréquences considérées sera noté $\{\xi_m\}_{m \in \{1, \dots, M_\xi\}}$.

Les auteurs proposent d'abord une formule de reconstruction inspirée de celle caractérisant la transformée de Fourier à court terme. Soit $\mathbf{y} \in \mathbb{R}^{N_\tau}$ le signal observé. Notons $\mathbf{G} \in \mathbb{C}^{M_\xi \times N_\tau}$ sa représentation temps-fréquence recherchée. Notons \mathbf{g}_m le vecteur ligne tel que $\mathbf{g}_m(\tau_n) = \mathbf{G}(\xi_m, \tau_n)$. Alors, la formule de reconstruction à partir de la représentation temps-fréquence discrétisée \mathbf{G} est donnée par :

$$\mathbf{y}(\tau_n) = \sum_{m=1}^M \operatorname{Re} \left(\mathbf{g}_m(\tau_n) e^{2i\pi \xi_m \tau_n} \right) + \epsilon(\tau_n), \quad (4.1)$$

où $\epsilon \in \mathbb{R}^{N_\tau}$ modélise un bruit de mesure indépendant de \mathbf{G} . Il est choisi de représenter ce bruit par un bruit blanc gaussien, donc : $\epsilon \sim \mathcal{N}(\mathbf{0}, \sigma_y^2 \mathbf{I})$ où σ_y^2 désigne la variance du bruit. Remarquons que cette formule de synthèse s'inspire de la formule de reconstruction à partir de la TFCT donnée dans le théorème 1.4 où la fenêtre de synthèse serait l'impulsion de Dirac. Cependant, la convention de phase retenue par les auteurs est différente de celle que nous avons donnée dans la définition 1.14 (voir la remarque qui suit cette définition).

Cas stationnaire. Dans un premier temps, les auteurs font l'hypothèse que le signal observé \mathbf{y} est stationnaire de spectre \mathcal{S}_Y . Ainsi, on peut voir chaque terme \mathbf{g}_m comme un signal basse fréquence stationnaire telle que la composante $\operatorname{Re}(\mathbf{g}_m e^{2i\pi \xi_m \cdot})$ constitue la partie du signal \mathbf{y} dont le spectre contient la sous-bande fréquentielle de \mathcal{S}_Y située autour de la fréquence ξ_m . Dans cette optique, l'*a priori* sur la composante \mathbf{g}_m est tel que $\mathbf{g}_m \sim \mathcal{CN}_c(\mathbf{0}, \mathbf{\Gamma}_m)$ où la matrice de covariance $\mathbf{\Gamma}_m$ d'un filtre linéaire noté κ_m , c'est-à-dire une matrice circulante. Elle est alors diagonalisée dans la base de Fourier, et ses valeurs propres sont données par la transformée de Fourier discrète $\hat{\kappa}_m$ de ce filtre linéaire. Ici, on choisit κ_m comme étant un filtre passe-bas de sorte que le spectre de la composante $\operatorname{Re}(\mathbf{g}_m e^{2i\pi \xi_m \cdot})$ soit situé autour de la fréquence ξ_m . Dans ce cadre, l'estimation du maximum *a posteriori* $\tilde{\mathbf{g}}_m$ de la composante \mathbf{g}_m est donnée par le filtrage de Wiener (voir [74], section 3.8.5) :

$$\tilde{\mathbf{g}}_m(\tau_n) = \sum_k \mathbf{y}(\tau_k) \mathbf{h}_m(\tau_k - \tau_n) e^{-2i\pi \xi_m \tau_k},$$

avec

$$\hat{\mathbf{h}}_m(v_n) = \frac{\hat{\kappa}_m(v_n)}{\sigma_y^2 + \frac{1}{2} \sum_{k=1}^M \hat{\kappa}_k(v_n - \xi_k)}.$$

Finalement, l'expression de $\tilde{\mathbf{g}}_m$ est similaire à celle de la transformée de Fourier à court terme discrétisée où la fenêtre \mathbf{h}_m dépend de la fréquence considérée grâce à l'*a priori*.

On peut alors parler de représentation temps-fréquence adaptative car la fenêtre est adaptée au spectre attendu du signal observé.

Cas non stationnaire. Turner et Sahani font le choix d'introduire la non-stationnarité du signal y en modifiant *a priori* sur chacune des composantes \mathbf{g}_m . Le type de non-stationnarité qu'ils prennent en compte est celui de la modulation d'amplitude. L'amplitude de chaque composante \mathbf{g}_m est supposée varier au cours du temps. Autrement dit, on voudrait pouvoir écrire $\mathbf{g}_m(\tau_n) = \alpha_m(\tau_n)e^{2i\pi\phi_m(\tau_n)}$ où α_m modélise l'amplitude variable de la composante \mathbf{g}_m et ϕ_m sa phase aléatoire. Pour cela, le modèle *a priori* choisi est de la forme :

$$\mathbf{g}_m(\tau_n) = b_m \mathbf{g}_m(\tau_{n-1}) + \epsilon_m(\tau_n) , \quad (4.2)$$

avec $\epsilon_m \sim \mathcal{CN}_c(\mathbf{0}, \sigma_m^2 \mathbf{I})$ et $0 < b_m < 1$. Il s'agit d'un modèle autorégressif d'ordre 1. Il s'agit d'un modèle conduisant à une décroissance exponentielle de l'amplitude, il est pertinent pour modéliser des signaux audios et des signaux issus de processus vibratoires. De plus, ce modèle présente l'avantage de conduire à un calcul rapide et aisé de la représentation temps-fréquence \mathbf{G} (type filtre de Kalman).

Autre modèle non stationnaire. Enfin, les auteurs proposent de modifier la formule de reconstruction du signal à partir la représentation temps-fréquence en introduisant une nouvelle structure. La factorisation en matrices non négatives (abrégé en NMF pour *Non-negative Matrix Factorization*) est une décomposition pertinente pour modéliser le spectrogramme de certains signaux audios comme les sons produits par des instruments de musique [78]. Il s'agit de rechercher des vecteurs colonnes $\mathbf{h}_l \in \mathbb{R}_+^{M_\xi}$ et $\mathbf{w}_l \in \mathbb{R}_+^{N_\tau}$ tels que le spectrogramme, noté $\mathbf{A}^2 \in \mathbb{R}_+^{M_\xi \times N_\tau}$, se décompose ainsi :

$$\mathbf{A}^2 = \sum_{l=1}^L \mathbf{h}_l \mathbf{w}_l^T . \quad (4.3)$$

Intuitivement, \mathbf{w}_l contient l'information d'activation temporelles de la composante fréquentielle \mathbf{h}_l . Dans le cas d'un signal musical, chaque vecteur \mathbf{h}_l décrit le spectre caractéristique d'une note tandis que \mathbf{w}_l caractérise les instants où cette note est jouée. L'idée est alors d'ajouter cette décomposition dans le modèle de synthèse. On considère alors que la représentation temps-fréquence s'écrit sous la forme :

$$\mathbf{G}(\xi_m, \tau_n) = \mathbf{A}(\xi_m, \tau_n) \mathbf{H}(\xi_m, \tau_n) ,$$

avec \mathbf{A} se décomposant sous la forme (4.3), et \mathbf{H} portant l'information de phase. De plus, les auteurs choisissent de considérer *a priori* autorégressif d'ordre 1 du même type que (4.2) sur $\mathbf{h}_m = \mathbf{H}(\xi_m, \cdot) \in \mathbb{C}^{N_\tau}$. Finalement, le nouveau modèle de synthèse s'écrit maintenant :

$$\mathbf{y}(\tau_n) = \sum_{m=1}^M \mathbf{a}_m(\tau_n) \operatorname{Re} \left(\mathbf{g}_m(\tau_n) e^{2i\pi\xi_m\tau_n} \right) + \epsilon(\tau_n) ,$$

où $\mathbf{a}_m = \mathbf{A}(\xi_m, \cdot) \in \mathbb{R}_+^{N_\tau}$. Les auteurs proposent alors une méthode permettant d'estimer entièrement \mathbf{G} , c'est-à-dire son amplitude \mathbf{A} qui suit le modèle de NMF (4.3) et sa phase \mathbf{H} .

Bien qu'il puisse modéliser la représentation temps-fréquence de signaux non stationnaires, le modèle proposé par Turner et Sahani reste peu satisfaisant en vue de représenter le spectre instantané de processus localement déformés. En effet, cette modélisation est limitée par le fait que l'on néglige la corrélation entre les composantes fréquentielles. C'est pourquoi nous allons adopter un point de vue différent où la corrélation fréquentielle sera bien prise en compte tandis que la corrélation temporelle ne découlera pas d'une hypothèse sur *l'a priori* mais d'une approche type « chapitre 2 » où on suppose que le déplacement des coefficients varie lentement.

4.3 Modèle probabiliste de synthèse pour les signaux localement dilatés

On reprend ici le problème de modélisation des signaux non stationnaires abordé dans le chapitre 2, en adoptant une approche inspirée de celle proposée par Tuner et Sahani.

4.3.1 Modèle de synthèse

Nous nous focalisons désormais sur la construction d'une représentation du spectre instantané d'un signal localement dilaté. Motivés par les résultats du chapitre 2, nous allons donc proposer une formule de synthèse similaire à la formule de reconstruction à partir de la transformée en ondelettes. Ainsi, nous chercherons à estimer une représentation temps-échelle ayant des propriétés similaires à la transformée en ondelettes.

Soit y le signal observé. Comme dans les chapitres précédents, y sera modélisé comme une réalisation d'un processus stochastique Y . Soit ψ une ondelette, on note $\psi_s(t) = q^{-s/2}\psi(q^{-s}t)$ l'ondelette dilatée. Soit \mathbf{s} le vecteur de taille M_s contenant l'ensemble des échelles que l'on souhaite analyser. Alors, par analogie avec la transformée en ondelettes, la formule de reconstruction s'écrit :

$$Y(t) = \text{Re} \left(\sum_{s \in \mathbf{s}} (\psi_s * W_s)(t) \right) + \epsilon(t), \quad (4.4)$$

où W_s désigne une représentation temps-échelle stochastique du processus aléatoire Y à l'échelle s et ϵ désigne un bruit de mesure. Dans ce qui suit, l'objectif est d'estimer la représentation temps-échelle W_s associée au signal observée y pour tout $s \in \mathbf{s}$. Pour cela, on se fixera certaines contraintes supplémentaires devant être respectées par cette représentation temps-échelle. En effet, l'idée est d'obtenir une représentation temps-échelle « idéale » possédant un certain nombre de propriétés désirées qu'une approche à l'analyse (comme la transformée en ondelettes classique) ne permet pas. Pour cela,

nous allons mettre en œuvre des estimations statistiques, en particulier bayésiennes. Les propriétés désirées de la représentation temps-échelle sont alors introduites comme des *a priori* sur le processus stochastique W_s .

Dans ce qui suit, l'ensemble des observations est réalisé à temps discret et sur un intervalle de longueur finie. Notons τ , le vecteur de taille N_τ contenant l'ensemble des instants de mesures (vecteur d'instants régulièrement espacés par un pas valant F_s^{-1} , où F_s désigne la fréquence d'échantillonnage). La discrétisation s'effectue en considérant des conditions aux bords périodiques. En particulier, pour l'ondelette dilatée ψ_s , on aura : $\psi_s(\tau_n) = \psi_s(\tau_{n+N_\tau}), \forall n$. De plus, les mesures étant discrètes nous chercherons à estimer W_s aux instants de mesures uniquement, et nous raisonnerons en terme de vecteurs et matrices aléatoires. Soit $n \in \{1, \dots, N_\tau\}$, on s'intéresse alors au vecteur $\mathbf{w}_n \in \mathbb{C}^{M_s}$ tel que $[\mathbf{w}_n]_m = W_{s_m}(\tau_n)$. La transformée en ondelettes discrétisée est donnée par la matrice $\underline{\mathbf{W}} = (\mathbf{w}_1 \cdots \mathbf{w}_{N_\tau}) \in \mathbb{C}^{M_s \times N_\tau}$. Soit $\mathbf{y} \in \mathbb{R}^{N_\tau}$ le signal discrétisé et $\epsilon \in \mathbb{R}^{N_\tau}$ le bruit associé, indépendant de $\underline{\mathbf{W}}$. Le modèle de synthèse se réécrit alors sous forme vectorielle :

$$\mathbf{y} = \text{Re} \left(\sum_{n=1}^{N_\tau} \Psi_n \mathbf{w}_n \right) + \epsilon, \quad (4.5)$$

où les Ψ_n sont les matrices de taille $N_\tau \times M_s$ telles que :

$$\Psi_n = \begin{pmatrix} \psi_{s_1}(\tau_{1-n}) & \cdots & \psi_{s_{M_s}}(\tau_{1-n}) \\ \vdots & & \vdots \\ \psi_{s_1}(\tau_{N_\tau-n}) & \cdots & \psi_{s_{M_s}}(\tau_{N_\tau-n}) \end{pmatrix}.$$

De manière encore plus compacte, soit \mathbf{W} la vectorisation de la matrice de transformée en ondelettes discrétisée. Il s'agit d'un vecteur de taille $N_\tau M_s$ formé par la concaténation des colonnes de la matrice $\underline{\mathbf{W}}$. Ainsi :

$$\mathbf{W} = \text{vec}(\underline{\mathbf{W}}) = \begin{pmatrix} \mathbf{w}_1 \\ \vdots \\ \mathbf{w}_{N_\tau} \end{pmatrix}.$$

Alors, le modèle de synthèse s'écrit

$$\mathbf{y} = \text{Re}(\mathbf{D}\mathbf{W}) + \epsilon. \quad (4.6)$$

où $\mathbf{D} = (\Psi_1 \cdots \Psi_{N_\tau})$ est une matrice de taille $N_\tau \times N_\tau M_s$. L'équation de synthèse (4.5) doit nous permettre d'estimer la représentation temps-échelle adaptée au signal. Pour cela, nous utiliserons les outils d'inférence statistique. On modélise alors le bruit ϵ comme étant un bruit blanc gaussien stationnaire de variance σ_y^2 , soit : $\epsilon \sim \mathcal{N}(\mathbf{0}, \sigma_y^2 \mathbf{I})$.

Dans l'esprit de l'algorithme JEFAS présenté au chapitre 2, nous voulons décrire l'effet de la déformation temporelle par un déplacement de coefficients dans le plan temps-échelle. Nous l'introduisons dans le modèle sous forme d'un *a priori* sur \mathbf{W} .

4.3.2 Modèle *a priori*

En plus de l'équation de synthèse, nous supposons que le signal observé provient d'un processus localement dilaté. Pour tenir compte de cette hypothèse dans le modèle de synthèse, nous devons ajouter un *a priori* sur la forme prise par la transformée temps-échelle adaptée recherchée. La propriété majeure que l'on souhaite voir vérifiée par la représentation temps-échelle de signaux localement déformés est celle du déplacement des coefficients dépendant de la fonction de dilatation instantanée tel que prévu par le théorème 2.1. Pour cela, nous construisons l'*a priori* de la manière suivante.

1. Par analogie avec les résultats du chapitre 2, les colonnes \mathbf{w}_n sont supposées décorréliées les unes des autres.
2. Pour chaque colonne, On suppose le vecteur \mathbf{w}_n gaussien complexe circulaire de moyenne nulle et dont la covariance est de la même forme que celle obtenue pour la transformée en ondelettes approchée dans la proposition 2.1. Ainsi $\mathbf{w}_n \sim \mathcal{CN}_c(\mathbf{0}, \mathbf{C}_n)$, où :

$$[\mathbf{C}_n]_{mm'} \triangleq [\mathbf{C}(\theta_n)]_{mm'} = f(s_m + \theta_n, s_{m'} + \theta_n) \quad \text{avec } f : (\mathbb{R}_+^*)^2 \rightarrow \mathbb{C}. \quad (4.7)$$

où θ_n désigne le paramètre de dilatation instantanée à l'instant τ_n .

3. Comme la matrice \mathbf{C}_n doit être une matrice de covariance, elle doit présenter une symétrie hermitienne, soit : $f(s, s') = \overline{f(s', s)}$ pour tous $s, s' > 0$. D'autre part, comme \mathbf{C}_n est une matrice semi-définie positive, la fonction f doit vérifier la condition suivante :

$$\sum_{k=1}^K \sum_{k'=1}^K \alpha_k \overline{\alpha_{k'}} f(s_k, s_{k'}) , \quad \forall (s_k)_{k=1, \dots, K} \in (\mathbb{R}_+^*)^K, \quad \forall (\alpha_k)_{k=1, \dots, K} \in \mathbb{C}^K .$$

Le choix de cette fonction dépend des propriétés supplémentaires que l'on veut donner à la représentation temps-échelle (comme, par exemple, la parcimonie). Le choix naturel pour la fonction f consiste à prendre cette fonction telle que

$$f(s, s') = q^{\frac{s+s'}{2}} \int_0^\infty \mathcal{S}_X(\xi) \overline{\hat{\psi}(q^s \xi)} \hat{\psi}(q^{s'} \xi) d\xi . \quad (4.8)$$

En effet, ceci conduit à la matrice de covariance d'une colonne de la transformée en ondelettes approchée du signal localement dilaté y (se reporter à l'équation (2.10)).

4.4 Estimation et algorithme

Nous décrivons ici une approche bayésienne pour l'estimation de la représentation temps-échelle. Cela suppose que les densités de probabilité du bruit et *a priori* sont parfaitement déterminées. Il faut donc connaître les paramètres suivants :

- la variance du bruit σ_y^2 ;

- le paramètre de dilatation instantanée θ_n pour tout $n \in \{1, \dots, N_\tau\}$;
- la fonction f qui peut elle-même dépendre d'autres paramètres, comme dans l'expression (4.8) où elle dépend du spectre \mathcal{S}_X du processus stationnaire sous-jacent.

Supposons dans un premier temps que ces paramètres sont connus. Nous détaillons dans la section 4.4.1 l'estimation *a posteriori* de \mathbf{W} qui en découle. Dans un second temps, nous proposons un algorithme d'estimation alternée de la représentation temps-échelle et de la fonction de dilatation. Il s'agit de l'analogue adaptatif de JEFAS pour une approche à la synthèse.

4.4.1 Espérance *a posteriori*

L'estimateur bayésien naturel mis en œuvre est l'espérance *a posteriori* qui coïncide ici avec celui du maximum *a posteriori* car toutes les variables aléatoires sont gaussiennes. Nous voulons estimer le vecteur \mathbf{W} . Les différentes colonnes \mathbf{w}_n de la représentation temps-échelle recherchée étant supposées indépendantes, la loi *a priori* du vecteur \mathbf{W} (où ces différentes colonnes sont concaténées) s'écrit alors (voir l'équation (1.2)) :

$$p(\mathbf{W}) = \mathcal{CN}_c(\mathbf{0}, \mathbf{\Gamma}_{pri}) = \frac{1}{\pi^{ST} \det(\mathbf{\Gamma}_{pri})} e^{-\mathbf{W}^H \mathbf{\Gamma}_{pri}^{-1} \mathbf{W}}, \quad (4.9)$$

avec

$$\mathbf{\Gamma}_{pri} = \begin{pmatrix} \mathbf{C}_1 & & & \\ & \mathbf{C}_2 & & \\ & & \ddots & \\ & & & \mathbf{C}_{N_\tau} \end{pmatrix} \in \mathbb{C}^{N_\tau M_s \times N_\tau M_s}.$$

D'autre part, comme le bruit blanc additif ϵ suit une loi gaussienne de variance σ_y^2 , la vraisemblance des observations s'écrit

$$p(\mathbf{y}|\mathbf{W}) = \mathcal{N}(\text{Re}(\mathbf{D}\mathbf{W}), \sigma_y^2 \mathbf{I}). \quad (4.10)$$

On déduit alors la loi *a posteriori* sur la représentation temps-échelle \mathbf{W} à partir de la règle de Bayes. On a :

$$p(\mathbf{W}|\mathbf{y}) = \frac{p(\mathbf{y}|\mathbf{W})p(\mathbf{W})}{p(\mathbf{y})}. \quad (4.11)$$

Dans le cas réel, un résultat classique est que toute composition de loi gaussiennes et une loi gaussienne. Dans notre cas, la loi *a posteriori* $p(\mathbf{W}|\mathbf{y})$ est formée à partir d'une composition de lois gaussiennes (réelle et complexe), on peut alors montrer que c'est une loi gaussienne complexe mais non circulaire. Dans ce cas, en plus de la connaissance de la moyenne et de la matrice de covariance, la loi est entièrement déterminée par la matrice de relation \mathbf{R} définie pour tout vecteur aléatoire complexe \mathbf{X} par :

$$\mathbf{R} = \mathbb{E} \left\{ (\mathbf{X} - \mathbb{E}\{\mathbf{X}\})(\mathbf{X} - \mathbb{E}\{\mathbf{X}\})^T \right\}.$$

La proposition 4.1 fournit l'expression de la loi *a posteriori*.

Proposition 4.1 (Loi a posteriori). Soit $\mathbf{W} \in \mathbb{C}^{N_\tau M_s}$ un vecteur aléatoire gaussien complexe circulaire dont la loi est donnée par (4.9). Soit $\mathbf{y} \in \mathbb{R}^{N_\tau}$ un vecteur aléatoire tel que la vraisemblance $p(\mathbf{y}|\mathbf{W})$ soit donnée par (4.10). Alors, la loi a posteriori $p(\mathbf{W}|\mathbf{y})$ est une loi gaussienne complexe, soit :

$$p(\mathbf{W}|\mathbf{y}) \sim \mathcal{CN}(\boldsymbol{\mu}_{pos}, \mathbf{C}_{pos}, \mathbf{R}_{pos}) ,$$

avec

$$\boldsymbol{\mu}_{pos} = \boldsymbol{\Gamma}_{pri} \mathbf{D}^H \mathbf{C}_y^{-1} \mathbf{y} \quad (4.12)$$

$$\boldsymbol{\Gamma}_{pos} = \boldsymbol{\Gamma}_{pri} - \frac{1}{4} \boldsymbol{\Gamma}_{pri} \mathbf{D}^H \mathbf{C}_y^{-1} \mathbf{D} \boldsymbol{\Gamma}_{pri} \quad (4.13)$$

$$\mathbf{R}_{pos} = - \left(\mathbf{D}^H \mathbf{D} + 4\sigma_y^2 \boldsymbol{\Gamma}_{pri}^{-1} \right)^{-1} \overline{\mathbf{D}^T \mathbf{D} \boldsymbol{\Gamma}_{pos}} , \quad (4.14)$$

où

$$\mathbf{C}_y = \sigma_y^2 \mathbf{I} + \frac{1}{2} \text{Re} \left(\mathbf{D} \boldsymbol{\Gamma}_{pri} \mathbf{D}^H \right) . \quad (4.15)$$

La démonstration de cette proposition est présentée en annexe A.3.1. L'expression de la matrice de relation \mathbf{R}_{pos} est donnée mais elle n'est pas utile pour la suite.

En notant $\tilde{\mathbf{W}} = \mathbb{E}_{\mathbf{W}|\mathbf{y}}\{\mathbf{W}\}$, l'estimateur de l'espérance a posteriori de \mathbf{W} , celui-ci correspond donc à la moyenne $\boldsymbol{\mu}_{pos}$ de la loi a posteriori $p(\mathbf{W}|\mathbf{y})$: $\tilde{\mathbf{W}} = \boldsymbol{\mu}_{pos}$. Chaque colonne de la représentation temps-échelle adaptée s'exprime donc sous la forme

$$\tilde{\mathbf{w}}_n = \frac{1}{2} \mathbf{C}_n \boldsymbol{\Psi}_n^H \mathbf{C}_y^{-1} \mathbf{y} \quad \text{avec} \quad \mathbf{C}_y = \sigma_y^2 \mathbf{I} + \frac{1}{2} \text{Re} \left(\sum_{n=1}^{N_\tau} \boldsymbol{\Psi}_n \mathbf{C}_n \boldsymbol{\Psi}_n^H \right) . \quad (4.16)$$

Bien que le point de vue soit assez différent, l'expression obtenue est assez similaire à celle obtenue par Turner et Sahani. Cependant, dans leur cas, la matrice \mathbf{C}_y définit la matrice de covariance d'un processus stationnaire, elle est donc diagonalisable dans la base de Fourier où son inversion s'effectue facilement. Notons d'ores et déjà que, dans notre situation, le calcul de l'estimation $\tilde{\mathbf{w}}_n$ sera numériquement compliqué par la nécessité de calculer la matrice \mathbf{C}_y . Celle-ci étant de grande taille ($N_\tau \times N_\tau$), son stockage peut donc être délicat, et son inversion s'avérer coûteuse en temps de calcul. Sur la figure 4.3, nous représentons la matrice de covariance \mathbf{C}_y associée au signal synthétique de la figure 4.1. Ce signal étant constitué de 8192 échantillons, la matrice \mathbf{C}_y contient donc 67 millions d'éléments environ. Toutefois, comme l'illustre la figure 4.1, seuls les éléments de cette matrice situés autour de la diagonale prennent des valeurs significatives. Cela vient du spectre instantané du signal : lorsque celui-ci se déplace vers les basses fréquences, la corrélation temporelle du signal augmente et la bande de coefficients significatifs situés autour de la diagonale s'élargit (autour de 0,4 secondes par exemple). À l'inverse, lorsque le spectre instantané se décale vers les hautes fréquences,

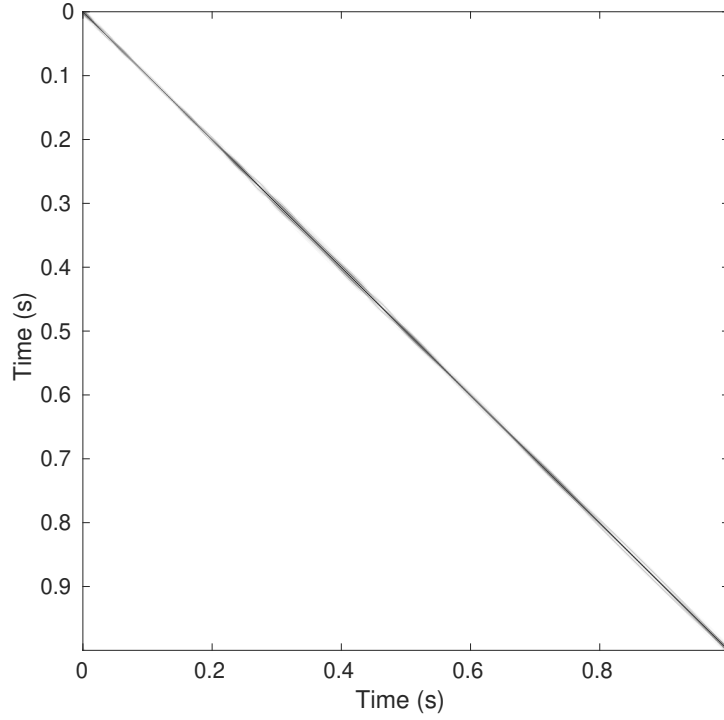


FIGURE 4.3 – Exemple synthétique. Module de la matrice de covariance \mathbf{C}_y associée à ce signal.

la corrélation temporelle diminue et seulement quelques éléments non diagonaux sont significatifs (autour de 0,1 seconde par exemple).

Remarque. La représentation temps-échelle estimée est corrélée temporellement. En effet, soit τ_n et $\tau_{n'}$ deux instants de mesure, alors :

$$\mathbb{E}\{\tilde{\mathbf{w}}_n \tilde{\mathbf{w}}_{n'}^H\} = \mathbb{1}_n(n') \mathbf{C}_n - \frac{1}{4} \mathbf{C}_n \mathbf{\Psi}_n^H \mathbf{C}_y^{-1} \mathbf{\Psi}_{n'} \mathbf{C}_{n'} . \quad (4.17)$$

Or, lorsque $n' \neq n$, ce terme ne s'annule généralement pas.

Quel choix faire pour la fonction f ? L'idée naturelle consiste à prendre celle donnée par l'expression (4.8), qui correspond à la fonction que donnerait la transformée en ondelettes. Pour déterminer entièrement cette fonction et la covariance associée, il nous reste néanmoins à connaître la fonction de dilatation instantanée et le spectre du processus stationnaire sous-jacent. Pour estimer ces deux grandeurs, on peut préalablement appliquer JEFAS au signal. On peut également directement les estimer à partir du modèle de synthèse. Dans ce qui suit, on montre que c'est une algorithm EM qui donne cette estimation.

4.4.2 JEFAS à la synthèse

Notons $\boldsymbol{\theta} = (\theta_1, \dots, \theta_{N_\tau})^T$ les paramètres de dilatation temporelle qui sont inconnus. Le calcul de $\tilde{\mathbf{w}}_n$ requiert donc la connaissance de $\boldsymbol{\theta}$. C'est pourquoi nous allons mettre en œuvre une estimation alternée des paramètres de dilatation temporelle. L'algorithme EM (abréviation de *Espérance-Maximisation*) est particulièrement adapté au modèle considéré. En effet, cet algorithme introduit par Dempster *et al.* [24] propose d'estimer un paramètre par maximum de vraisemblance lorsque celle-ci dépend de variables cachées inconnues. En utilisant le vocabulaire de l'algorithme EM, \mathbf{y} représente ici les observations, \mathbf{W} les variables cachées et $\boldsymbol{\theta}$ les paramètres à estimer. Dans ce cas, étant donnée une estimée courante $\tilde{\boldsymbol{\theta}}^{(k-1)}$ des paramètres, la mise à jour se fait de la manière suivante :

1. Espérance :

$$Q(\boldsymbol{\theta}, \tilde{\boldsymbol{\theta}}^{(k-1)}) = \mathbb{E}_{\mathbf{W}|\mathbf{y}, \tilde{\boldsymbol{\theta}}^{(k-1)}} \{\log p(\mathbf{y}, \mathbf{W}|\boldsymbol{\theta})\}$$

2. Maximisation :

$$\tilde{\boldsymbol{\theta}}^{(k)} = \arg \max_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}, \tilde{\boldsymbol{\theta}}^{(k-1)}) .$$

La proposition 4.2 donne une réécriture de l'algorithme EM dans le cas qui nous intéresse.

Proposition 4.2 (Étapes de l'algorithme EM). Soient $\tilde{\boldsymbol{\theta}}^{(k-1)}$ le vecteur des paramètres de dilatation estimé après $k - 1$ itérations de l'algorithme EM. Soit $\tilde{\mathbf{W}}^{(k-1)}$ la représentation temps-échelle estimée après $k - 1$ itérations de l'algorithme EM. Alors, la k -ème itération de l'algorithme EM met à jour les valeurs de ces deux grandeurs en effectuant les deux étapes suivantes.

1. Pour tout $n \in \{1, \dots, N_\tau\}$, l'estimateur (4.16) de la représentation adaptée à l'instant τ_n est donné par :

$$\tilde{\mathbf{w}}_n^{(k)} = \frac{1}{2} \mathbf{C}(\tilde{\theta}_n^{(k-1)}) \boldsymbol{\Psi}_n^H \mathbf{C}_y(\tilde{\boldsymbol{\theta}}^{(k-1)})^{-1} \mathbf{y} . \quad (4.18)$$

2. Le problème d'estimation du paramètre de dilatation $\tilde{\boldsymbol{\theta}}$ se réécrit :

$$\tilde{\boldsymbol{\theta}}^{(k)} = \arg \min_{\boldsymbol{\theta}} \sum_{n=1}^{N_\tau} \left[\log |\det(\mathbf{C}(\theta_n))| + \tilde{\mathbf{w}}_n^{(k)H} \mathbf{C}(\theta_n)^{-1} \tilde{\mathbf{w}}_n^{(k)} + \text{Trace} \left(\mathbf{C}(\theta_n)^{-1} \boldsymbol{\Gamma}_n(\tilde{\boldsymbol{\theta}}^{(k-1)}) \right) \right] , \quad (4.19)$$

où $\boldsymbol{\Gamma}_n(\tilde{\boldsymbol{\theta}}^{(k-1)})$ désigne le n -ème bloc de taille $M_s \times M_s$ le long de la diagonale de la matrice $\boldsymbol{\Gamma}_{\text{pos}}(\tilde{\boldsymbol{\theta}}^{(k-1)})$.

Les différents calculs conduisant (4.19) au problème sont détaillés en annexe A.3.2. Ainsi, l'expression (4.19) montre que l'estimation du paramètre de dilatation instantanée θ peut s'effectuer indépendamment à chaque instant. L'étape de maximisation s'écrit alors comme la résolution d'un problème très proche de celui de JEFAS :

$$\tilde{\theta}_n^{(k)} = \arg \max_{\theta_n} \ell(\theta_n), \quad (4.20)$$

où :

$$\ell(\theta_n) = -\log |\det(\mathbf{C}(\theta_n))| - \tilde{\mathbf{w}}_n^{(k-1)H} \mathbf{C}(\theta_n)^{-1} \tilde{\mathbf{w}}_n^{(k-1)} - \text{Trace} \left(\mathbf{C}(\theta_n)^{-1} \Gamma_n \left(\tilde{\theta}^{(k-1)} \right) \right).$$

Par ailleurs, l'équation (4.13) nous donne

$$\Gamma_n \left(\tilde{\theta}^{(k-1)} \right) = \mathbf{C} \left(\tilde{\theta}_n^{(k-1)} \right) - \frac{1}{4} \mathbf{C} \left(\tilde{\theta}_n^{(k-1)} \right) \Psi_n^H \mathbf{C}_y \left(\tilde{\theta}^{(k-1)} \right)^{-1} \Psi_n \mathbf{C} \left(\tilde{\theta}_n^{(k-1)} \right).$$

4.4.3 Algorithme

L'algorithme EM suggère donc une procédure d'estimation jointe de la déformation instantanée caractéristique du signal localement déformé et de la représentation temps-échelle adaptée à ce signal. De plus, la construction de la matrice de covariance de la représentation temps-échelle adaptée requiert la connaissance du spectre du signal stationnaire sous-jacent (voir l'équation (4.8)), il faudra donc également l'estimer. Il s'agit donc d'une adaptation de JEFAS par un approche à la synthèse. C'est pourquoi l'algorithme proposé est appelé JEFAS-S (pour abréger JEFAS-Synthèse). Le pseudo-code correspondant est donné dans l'algorithme 4.1. Précisons la mise en œuvre de chacune des étapes de l'algorithme.

Algorithme 4.1 JEFAS-S (*JEFAS-Synthèse*). $(\tilde{\mathbf{W}}, \tilde{\theta}, \tilde{\mathcal{S}}_X) = \text{JEFAS-S}(\mathbf{y}, \sigma_y^2)$

- **Initialisation** : À partir de \mathbf{y} , estimer $\tilde{\theta}^{(0)}$ et $\tilde{\mathcal{S}}_X^{(0)}$ par l'algorithme JEFAS.
 - $k \leftarrow 1$.
 - tant que** le critère d'arrêt (4.22) est faux **faire**
 - pour tout** $n \in \{1, \dots, N_\tau\}$ **faire**
 - Restreindre $\Psi_n, \mathbf{C}_y \left(\tilde{\theta}^{(k-1)} \right)$ et \mathbf{y} à l'intervalle de taille N'_τ autour de n .
 - Obtenir $\tilde{\mathbf{w}}_n^{(k)}$ par l'équation (4.18).
 - fin pour**
 - Estimer $\tilde{\theta}^{(k)}$ en approchant la solution du problème (4.20) par la méthode de quasi-Newton.
 - Estimer $\tilde{\mathcal{S}}_X^{(k)}$ par l'estimateur du spectre-ondelettes (1.22).
 - $k \leftarrow k + 1$.
 - fin tant que**
-

Initialisation. Afin d'estimer une première représentation temps-échelle adaptée au signal y , nous devons connaître une première estimation de la fonction de dilatation instantanée. Or, l'initialisation peut jouer un rôle crucial sur les performances de l'algorithme. En effet, l'algorithme EM n'assure la convergence que vers un maximum local de la vraisemblance. Une mauvaise initialisation de JEFAS-S peut donc conduire à une estimation des paramètres de dilatation instantanée très éloignée des vraies valeurs. Or, lorsqu'elle est disponible, l'estimation donnée par JEFAS est la meilleure estimation de cette fonction. C'est pourquoi on privilégiera, lorsque c'est possible, l'initialisation de JEFAS-S par le résultat de JEFAS aux autres méthodes. Dans le cas où l'estimation par JEFAS est impossible, d'autres initialisations sont possibles comme le barycentre local du scalogramme (voir l'expression (2.26)) ou la fonction nulle.

Calcul de la représentation temps-échelle. Numériquement, le calcul de la représentation temps-échelle adaptée grâce à l'expression (4.18) conduit à quelques contraintes numériques.

- Premièrement, le calcul de $\mathbf{C}(\tilde{\theta}_n^{(k-1)})$ nécessite la connaissance de la fonction f dans l'équation (4.7). Concernant l'algorithme JEFAS-S, nous choisirons systématiquement d'estimer la représentation temps-échelle adaptée dans le cas où cette fonction est la fonction caractéristique de la transformée en ondelettes approchée, c'est-à-dire celle donnée par l'expression (4.8). Cependant, la détermination de cette fonction requiert la connaissance du spectre \mathcal{S}_X du signal stationnaire sous-jacent. C'est pourquoi nous ajoutons une étape d'estimation spectrale fondée sur le spectre-ondelettes (voir l'expression (1.22)). Lorsque le résultat de JEFAS est accessible, on fournira le spectre estimé par JEFAS comme initialisation du spectre. Dans le cas où cette estimée n'est pas disponible, on peut initialiser \mathcal{S}_X par l'estimation de Welch appliquée au signal non stationnaire \mathbf{y} , ou sans *a priori* par une fonction constante (caractéristique du spectre d'un bruit blanc). Néanmoins, ces deux dernières initialisations peuvent conduire à une convergence très lente de l'algorithme.
- Deuxièmement, la matrice \mathbf{C}_y est une matrice de grande taille. Son inversion et son stockage peuvent donc s'avérer délicats. Cependant, on peut faire l'approximation que $\tilde{\mathbf{w}}_n^{(k)}$ ne dépendra que des valeurs de \mathbf{y} autour de τ_n . On peut s'en convaincre en vérifiant que la corrélation entre deux colonnes de cette représentation tend vers zéro dès que ces colonnes sont suffisamment éloignées temporellement (voir l'équation (4.17) et l'exemple proposé en figure 4.3). Nous profitons de cette propriété pour restreindre chacune des matrices dans l'expression (4.18) à la portion utile de taille $N'_\tau \ll N_\tau$. Ainsi, la matrice \mathbf{C}_y devient plus facilement stockable et inversible.

Optimisation. L'estimation du paramètre de dilatation instantanée $\tilde{\theta}_n^{(k)}$ nécessite la résolution du problème d'optimisation (4.8). Pour cela, nous procédons grâce à la méthode de quasi-Newton (voir le chapitre 4 dans [8]). À chaque itération de cette mé-

thode de descente, la direction de descente est donnée par le gradient de la fonction à minimiser, dont l'expression est donnée par

$$\begin{aligned} \ell'(\theta_n) = & \text{Trace} \left(\mathbf{C}(\theta_n)^{-1} \frac{\partial \mathbf{C}}{\partial \theta_n}(\theta_n) \right) + \mathbf{w}_{y,\tau_n}^H \mathbf{C}(\theta_n)^{-1} \frac{\partial \mathbf{C}}{\partial \theta_n}(\theta_n) \mathbf{C}(\theta_n)^{-1} \mathbf{w}_{y,\tau_n} \\ & - \text{Trace} \left(\mathbf{C}(\theta_n)^{-1} \frac{\partial \mathbf{C}}{\partial \theta_{n,2}}(\theta_n) \mathbf{C}(\theta_n)^{-1} \mathbf{\Gamma}_n \left(\tilde{\boldsymbol{\theta}}^{(k-1)} \right) \right) . \end{aligned} \quad (4.21)$$

Critère d'arrêt. La convergence de JEFAS-S peut être vérifiée de deux manières différentes. Premièrement, comme pour JEFAS, nous pouvons calculer l'innovation relative du paramètre de dilatation instantanée et arrêter l'algorithme dès que celle-ci est inférieure à un seuil donné. Deuxièmement, l'algorithme EM assure la décroissance de la vraisemblance des observations, notée $\mathcal{L}(\boldsymbol{\theta})$. Dans notre situation, cette dernière est donnée par :

$$\begin{aligned} \mathcal{L}(\boldsymbol{\theta}) & \triangleq p(\mathbf{y}|\boldsymbol{\theta}) = \int p(\mathbf{y}, \mathbf{W}|\boldsymbol{\theta}) d\mathbf{W} \\ & = \mathcal{N}(\mathbf{0}, \mathbf{C}_y(\boldsymbol{\theta})) . \end{aligned}$$

On peut alors établir un critère de convergence en vérifiant que l'innovation absolue de cette vraisemblance passe sous un seuil donné. C'est ce critère d'arrêt qui sera mis en œuvre en pratique. À la k -ème itération de JEFAS-S, on vérifie donc si l'inégalité suivante est vraie :

$$\mathcal{L}(\boldsymbol{\theta}^{(k)}) - \mathcal{L}(\boldsymbol{\theta}^{(k-1)}) < \Lambda_y , \quad (4.22)$$

où Λ_y désigne le seuil choisi. Le principal inconvénient de ce critère d'arrêt concerne le choix de Λ_y car la décroissance de la vraisemblance varie fortement d'un signal à l'autre. Il convient donc d'observer les valeurs successivement prises par la vraisemblance au cours des premières itérations avant de choisir Λ_y .

4.4.4 Reconstruction du signal original

Une fois que l'algorithme JEFAS-S a convergé, nous avons accès au spectre et à la dilatation instantanée estimés. Nous pouvons alors modifier la fonction f proposée en (4.8) afin de synthétiser de nouvelles représentations temps-échelles adaptées. Cependant, on doit se limiter à des choix raisonnables dans la sens où la formule de synthèse (4.4) reste approximativement vraie. Pour vérifier cela, on considère le biais et la variance de l'estimateur du signal original.

En s'appuyant sur l'équation (4.5) décrivant le modèle, le signal original non bruité \mathbf{y}_0 est donné par :

$$\mathbf{y}_0 = \text{Re} \left(\sum_{n=1}^{N_\tau} \mathbf{\Psi}_n \mathbf{w}_n \right) .$$

L'estimateur $\tilde{\mathbf{y}}_0$ de \mathbf{y}_0 est alors donné par :

$$\tilde{\mathbf{y}}_0 = \text{Re} \left(\sum_{n=1}^{N_\tau} \mathbf{\Psi}_n \tilde{\mathbf{w}}_n \right) = \frac{1}{2} \text{Re} \left(\sum_{n=1}^{N_\tau} \mathbf{\Psi}_n \mathbf{C}_n \mathbf{\Psi}_n^H \right) \mathbf{C}_y^{-1} \mathbf{y}. \quad (4.23)$$

Afin d'étudier les performances de cet estimateur, calculons son biais.

$$\begin{aligned} \mathbb{E} \{ \tilde{\mathbf{y}}_0 | \mathbf{y}_0 \} - \mathbf{y}_0 &= \frac{1}{2} \text{Re} \left(\sum_{n=1}^{N_\tau} \mathbf{\Psi}_n \mathbf{C}_n \mathbf{\Psi}_n^H \right) \mathbf{C}_y^{-1} \underbrace{\mathbb{E} \{ \mathbf{y} \}}_{=\mathbf{y}_0} - \mathbf{y}_0 \\ &= \left(\frac{1}{2} \text{Re} \left(\sum_{n=1}^{N_\tau} \mathbf{\Psi}_n \mathbf{C}_n \mathbf{\Psi}_n^H \right) \mathbf{C}_y^{-1} - \mathbf{I} \right) \mathbf{y}_0. \end{aligned}$$

Or, par définition de \mathbf{C}_y on a :

$$\frac{1}{2} \text{Re} \left(\sum_{n=1}^{N_\tau} \mathbf{\Psi}_n \mathbf{C}_n \mathbf{\Psi}_n^H \right) \mathbf{C}_y^{-1} - \mathbf{I} = -\sigma_y^2 \mathbf{C}_y^{-1},$$

d'où l'expression du biais :

$$\mathbb{E} \{ \tilde{\mathbf{y}}_0 | \mathbf{y}_0 \} - \mathbf{y}_0 = -\sigma_y^2 \mathbf{C}_y^{-1} \mathbf{y}_0. \quad (4.24)$$

On peut également calculer la matrice de covariance de l'estimateur du signal original. Elle est définie par $\mathbf{R}(\tilde{\mathbf{y}}_0 | \mathbf{y}_0) \triangleq \mathbb{E} \left\{ (\tilde{\mathbf{y}}_0 - \mathbb{E} \{ \tilde{\mathbf{y}}_0 | \mathbf{y}_0 \}) (\tilde{\mathbf{y}}_0 - \mathbb{E} \{ \tilde{\mathbf{y}}_0 | \mathbf{y}_0 \})^T \middle| \mathbf{y}_0 \right\}$ et s'exprime :

$$\begin{aligned} \mathbf{R}(\tilde{\mathbf{y}}_0 | \mathbf{y}_0) &= \underbrace{\frac{1}{2} \text{Re} \left(\sum_{n=1}^{N_\tau} \mathbf{\Psi}_n \tilde{\mathbf{C}}_n \mathbf{\Psi}_n^H \right) \mathbf{C}_y^{-1}}_{=\mathbf{I} - \sigma_y^2 \mathbf{C}_y^{-1}} \underbrace{\mathbb{E} \{ (\mathbf{y} - \mathbf{y}_0)(\mathbf{y} - \mathbf{y}_0)^T \}}_{=\sigma_y^2 \mathbf{I}} \underbrace{\mathbf{C}_y^{-1} \frac{1}{2} \text{Re} \left(\sum_{n=1}^{N_\tau} \mathbf{\Psi}_n \tilde{\mathbf{C}}_n \mathbf{\Psi}_n^H \right)}_{=\mathbf{I} - \sigma_y^2 \mathbf{C}_y^{-1}} \\ &= \sigma_y^2 \left(\mathbf{I} - \sigma_y^2 \mathbf{C}_y^{-1} \right)^2 \end{aligned}$$

Application au débruitage. L'objectif de déterminer cette représentation peut aussi être de débruiter le signal observé afin d'en extraire uniquement la composante d'intérêt, c'est-à-dire le signal localement dilaté. Pour quantifier la qualité du débruitage, déterminons l'amélioration du rapport signal sur bruit entre le signal mesuré \mathbf{y} et le signal estimé $\tilde{\mathbf{y}}_0$. Notons Δ_{SNR} cette quantité, elle s'exprime :

$$\Delta_{\text{SNR}} = \text{SNR}(\tilde{\mathbf{y}}_0) - \text{SNR}(\mathbf{y}) = 10 \log_{10} \left(\frac{\mathbb{E} \{ \|\tilde{\mathbf{y}}_0\|_2^2 \}}{\mathbb{E} \{ \|\tilde{\mathbf{y}}_0 - \mathbf{y}_0\|_2^2 \}} \right) - 10 \log_{10} \left(\frac{\mathbb{E} \{ \|\mathbf{y}\|_2^2 \}}{\mathbb{E} \{ \|\epsilon\|_2^2 \}} \right). \quad (4.25)$$

Concernant le rapport signal sur bruit original, on a :

$$\begin{aligned} \text{SNR}(\mathbf{y}) &= 10 \log_{10} \left(\frac{\mathbb{E} \{ \|\mathbf{y}_0 + \epsilon\|_2^2 \}}{N_\tau \sigma_y^2} \right) = 10 \log_{10} \left(\frac{\|\mathbf{y}_0\|_2^2 + \mathbb{E} \{ \|\epsilon\|_2^2 \}}{N_\tau \sigma_y^2} \right) \\ &= 10 \log_{10} \left(1 + \frac{\|\mathbf{y}_0\|_2^2}{N_\tau \sigma_y^2} \right). \end{aligned}$$

Concernant le rapport signal sur bruit après reconstruction du signal synthétisé, son expression se déduit de celles du biais et de la variance de l'estimateur $\tilde{\mathbf{y}}_0$. En effet :

$$\begin{aligned} \text{SNR}(\tilde{\mathbf{y}}_0) &= 10 \log_{10} \left(\frac{\mathbb{E} \{ \|\tilde{\mathbf{y}}_0 - \mathbb{E} \{ \tilde{\mathbf{y}}_0 \} \|_2^2 \} + \|\mathbb{E} \{ \tilde{\mathbf{y}}_0 \} \|_2^2}{\mathbb{E} \{ \|\tilde{\mathbf{y}}_0 - \mathbb{E} \{ \tilde{\mathbf{y}}_0 \} \|_2^2 \} + \|\mathbb{E} \{ \tilde{\mathbf{y}}_0 \} - \mathbf{y}_0\|_2^2} \right) \\ &= 10 \log_{10} \left(\frac{\text{Trace} \left(\mathbf{R}(\tilde{\mathbf{y}}_0) + \mathbb{E} \{ \tilde{\mathbf{y}}_0 \} \mathbb{E} \{ \tilde{\mathbf{y}}_0 | \mathbf{y}_0 \}^T \right)}{\text{Trace} \left(\mathbf{R}(\tilde{\mathbf{y}}_0) + (\mathbb{E} \{ \tilde{\mathbf{y}}_0 \} - \mathbf{y}_0) (\mathbb{E} \{ \tilde{\mathbf{y}}_0 | \mathbf{y}_0 \} - \mathbf{y}_0)^T \right)} \right) \\ &= 10 \log_{10} \left(\frac{\text{Trace} \left(\left(\mathbf{I} - \sigma_y^2 \mathbf{C}_y^{-1} \right)^2 (\mathbf{I} + \mathbf{y}_0 \mathbf{y}_0^T) \right)}{\sigma_y^2 \text{Trace} \left(\left(\mathbf{I} - \sigma_y^2 \mathbf{C}_y^{-1} \right)^2 + \sigma_y^2 \mathbf{C}_y^{-2} \mathbf{y}_0 \mathbf{y}_0^T \right)} \right). \end{aligned}$$

Compte tenu des expressions de $\text{SNR}(\mathbf{y})$ et $\text{SNR}(\tilde{\mathbf{y}}_0)$, l'expression de Δ_{SNR} est difficilement interprétable. En particulier, nous ne pouvons pas déterminer aisément dans quelles conditions cette quantité est positive. C'est pourquoi dans la section 4.5, nous évaluons numériquement ces différentes quantités sur des exemples synthétiques et nous comparons les résultats obtenus aux expressions théoriques.

4.5 Résultats numériques

4.5.1 Application à une signal synthétique

Dans un premier temps, nous allons étudier les performances de ce nouvel algorithme sur un signal synthétique. On construit un signal de $N_\tau = 8192$ échantillons, échantillonné à la fréquence $F_s = 8192$ Hz. Le signal localement dilaté y est construit de la même manière que l'exemple synthétique de la section 2.5.3 (hormis la modulation d'amplitude). On rappelle que le spectre du processus stationnaire sous-jacent est constitué de deux lobes formés par des fenêtres de Hann (voir l'expression (2.27) et la figure 2.6). La fonction de dilatation instantanée γ' est une fonction oscillante donnée par l'expression (2.28). Ce signal est ensuite bruité par l'ajout d'un bruit blanc gaussien de variance $\sigma_y^2 = 5 \times 10^{-2}$, de sorte que le rapport signal sur bruit est de 16,2 dB.

On applique alors JEFAS-S à ce signal. Après initialisation par le résultat de JEFAS, l'algorithme converge en 3 itérations. Le temps de calcul sur un processeur Intel Xeon CPU E5-2680 v4 à 2,40 GHz est de 347 secondes. On superpose en figure 4.4 la fonction

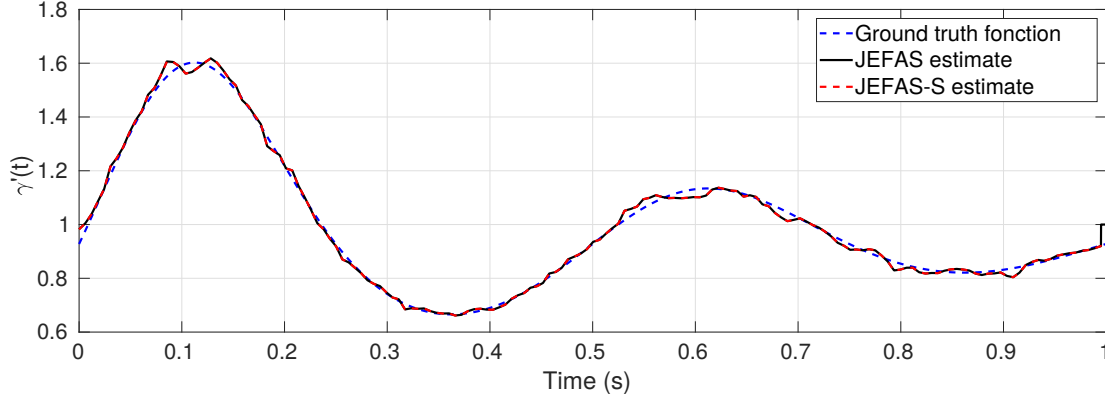


FIGURE 4.4 – Signal synthétique. fonction de dilatation instantanée estimée par JEFAS-S superposée avec celle estimée par JEFAS.

de dilatation estimée par JEFAS-S à celle obtenue par JEFAS, et à la vraie fonction. On remarque que les résultats des deux estimations sont quasiment superposés. JEFAS-S ne permet donc pas d'améliorer significativement le résultat d'estimation donné par JEFAS. On peut d'ailleurs noter que l'erreur quadratique moyenne d'estimation de la fonction de dilatation instantanée par JEFAS est de $2,250.10^{-4}$ tandis que celle résultant de JEFAS-S est de $2,238.10^{-4}$. En outre, le temps de calcul est fortement augmenté. On peut en conclure que l'utilisation de JEFAS-S n'est pas nécessaire lorsque le résultat de JEFAS est satisfaisant.

Néanmoins, JEFAS-S reste intéressant dans la mesure où cet algorithme permet d'obtenir une représentation temps-échelle adaptée au signal considéré. La représentation ainsi obtenue est affichée à côté du scalogramme sur la figure 4.5. D'abord, on note que les deux représentations temps-échelles sont très similaires. C'est ce qui était attendu car nous avons choisi un *a priori* tel que les covariances inter-échelles de \mathbf{W} et de la transformée en ondelettes soient égales. En revanche, le modèle retenu au cours de ce chapitre est la décorrélation temporelle *a priori* des colonnes de la représentation temps-échelle. C'est pourquoi on observe une oscillation beaucoup plus forte de cette représentation par rapport à la transformée en ondelettes dont les variations temporelles sont beaucoup plus lisses (et régulières) du fait de la corrélation temporelle existante.

Enfin, la représentation adaptée pouvant être vue comme un filtre de Wiener, remarquons que le bruit blanc présent à toutes les fréquences de la transformée en ondelettes est fortement atténué (voir invisible) sur la représentation temps-échelle adaptée. JEFAS-S permet donc d'effectuer un débruitage du signal mesuré grâce à ce filtrage non stationnaire. Après avoir reconstruit le signal original grâce à la formule de synthèse (4.23), on note d'ailleurs que le rapport signal sur bruit a été amélioré passant de 16,08 dB pour le signal mesuré à 23,14 dB pour le signal reconstruit, ce qui correspond à une amélioration du SNR de 7,06 dB. L'expression théorique (4.25) donne, pour cet exemple, une amélioration du rapport signal sur bruit de 13,40 dB. Cette valeur est lé-

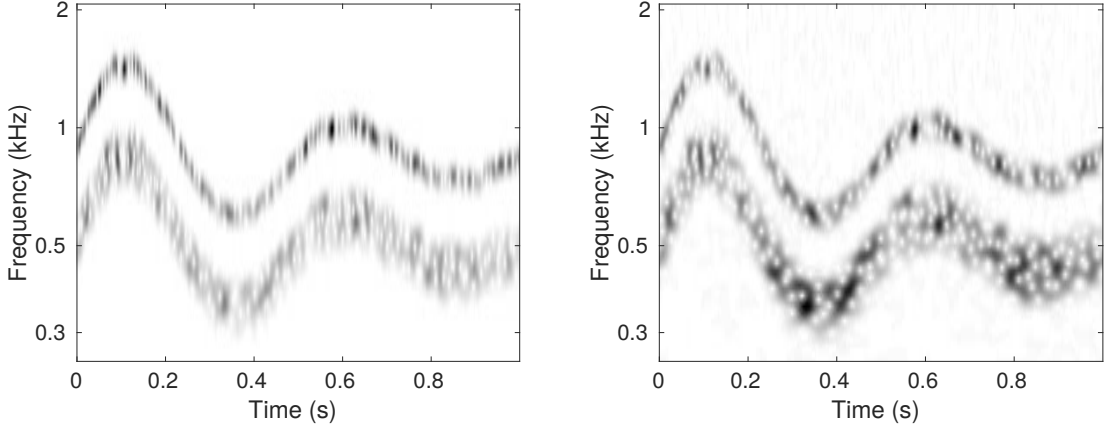


FIGURE 4.5 – Signal synthétique. À gauche : représentation temps-échelle adaptée donnée par JEFAS-S. À droite : scalogramme (transformée en ondelettes).

gèrement supérieure à la valeur mesurée car le calcul de l'expression (4.25) ne tient pas compte du fait que le paramètre de dilatation instantanée θ est également issu d'une estimation.

4.5.2 Application à un signal semi-réel à fréquence rapide

Considérons maintenant un signal semi-réel, c'est-à-dire un signal synthétisé à partir de mesures réelles. L'idée est d'étudier un signal localement déformé dont le spectre du processus stationnaire sous-jacent possède une structure parcimonieuse comme, par exemple, celui d'une sinusoïde. C'est pourquoi nous considérons un signal y localement harmonique à une seule composante. On peut donc l'écrire sous la forme

$$y(t) = A_1(t) \cos(2\pi\phi_1(t)) .$$

Ici, la fréquence instantanée ϕ_1' provient de données réelles mesurées. Il s'agit de la mesure de la fréquence cardiaque d'une personne atteinte de fibrillation atriale. L'amplitude instantanée est quant à elle simulée à partir d'une réalisation d'un mouvement brownien lissé puis normalisé et centré de sorte que $1 \leq A_1(t) \leq 5/3$ pour tout t . Le signal ainsi créé est échantillonné à la fréquence $F_s = 10$ Hz sur une durée de 83,1 secondes (soit $N_t = 832$). Le chronogramme du signal est tracé en figure 4.6. Les graphes de la fréquence et de l'amplitude instantanées ainsi que la transformée en ondelettes du signal y sont également tracés. Ce signal est un signal localement déformé car il s'agit d'un signal à une composante donc on peut écrire

$$y(t) = \mathcal{A}_\alpha \mathcal{D}_\gamma x(t) ,$$

où $x(t) = \cos(2\pi t)$, $\gamma(t) = \phi_1(t)$, et $\alpha(t) = \frac{A(t)}{\sqrt{\phi_1'(t)}}$. Pour appliquer JEFAS-S, nous allons négliger la modulation amplitude instantanée α et nous concentrer sur la dilatation instantanée γ' .

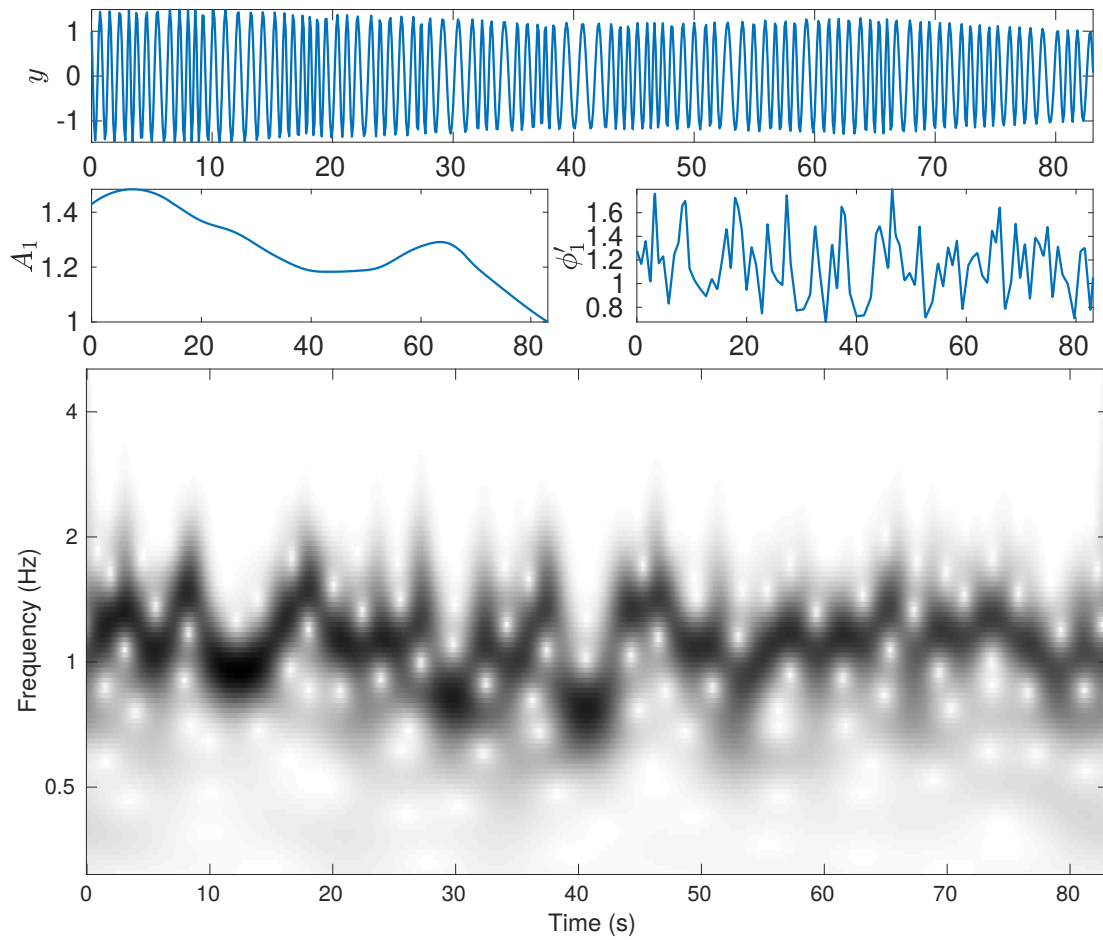


FIGURE 4.6 – Signal semi-réel. En haut : chronogramme Au milieu à gauche : amplitude instantanée simulée. Au milieu à droite : fréquence instantanée mesurée. En bas : scalogramme.

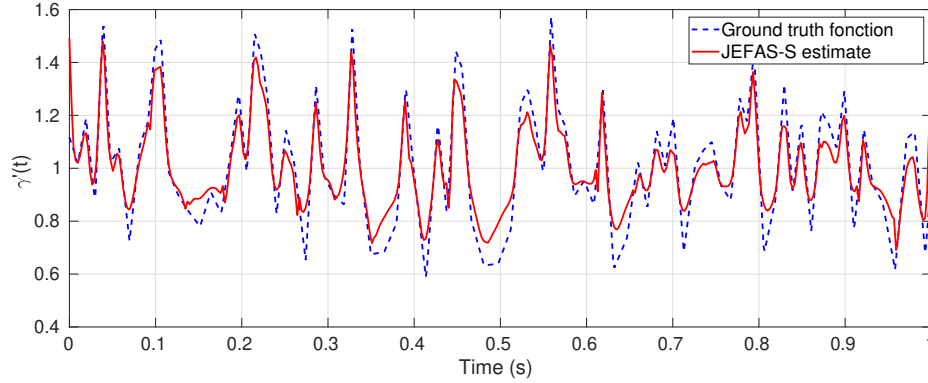


FIGURE 4.7 – Signal semi-réel. Fonction de dilatation instantanée estimée superposée à la fréquence instantanée normalisée.

Pour ce signal, l'algorithme JEFAS diverge. En effet, la transformée en ondelettes tracée en bas de la figure 4.6 présente des interférences car les variations temporelles de ϕ_1' sont trop rapides. Ainsi, le déplacement des coefficients d'ondelettes approché par le théorème 2.1 n'est pas valable pour décrire cette transformée en ondelettes. En fait, l'erreur d'approximation donnée par ce théorème est très grande du fait que la valeur de $\|\gamma''\|_\infty$ est grande. On initialise alors JEFAS-S par la fonction nulle pour première estimée de $\log_2(\gamma')$. Le spectre du processus stationnaire sous-jacent est initialisé par une fonction constante. Étant données ces initialisations très éloignées des valeurs attendues, la convergence de JEFAS-S est très lente. Le critère d'arrêt est atteint après 72 itérations. On représente en figure 4.7 la fonction de dilatation instantanée estimée superposée à la fréquence instantanée normalisée du signal original. JEFAS-S permet d'obtenir une estimation performante de la fréquence instantanée (à une constante multiplicative près). On bénéficie ici de l'*a priori* qui force la représentation temps-échelle obtenue à vérifier la propriété de déplacement des coefficients. Ainsi, nous ne sommes plus limités par les interférences visibles sur la transformée en ondelettes du fait des variations rapides de γ' . Cet exemple nous permet d'illustrer l'intérêt de JEFAS-S vis-à-vis de JEFAS. Au prix d'un temps de calcul fortement augmenté, JEFAS-S permet d'étudier une catégorie de signaux localement dilatés plus importante.

On peut profiter de ce résultat pour obtenir une représentation temps-échelle plus parcimonieuse que la transformée en ondelettes. Pour cela, introduisons l'*a priori* que le spectre \mathcal{S}_X présente une structure parcimonieuse, et notons ν_1 la fréquence de la sinusoïde correspondante. On peut estimer ν_1 en considérant, par exemple, la fréquence pour laquelle le spectre \mathcal{S}_X estimé par JEFAS-S est maximale. On construit alors une nouvelle représentation temps-échelle en prenant pour la fonction f définie en (4.7) l'expression suivante :

$$f(s, s') = e^{-\frac{(s-\zeta)^2}{\sigma_s^2}} \mathbf{1}_s(s') .$$

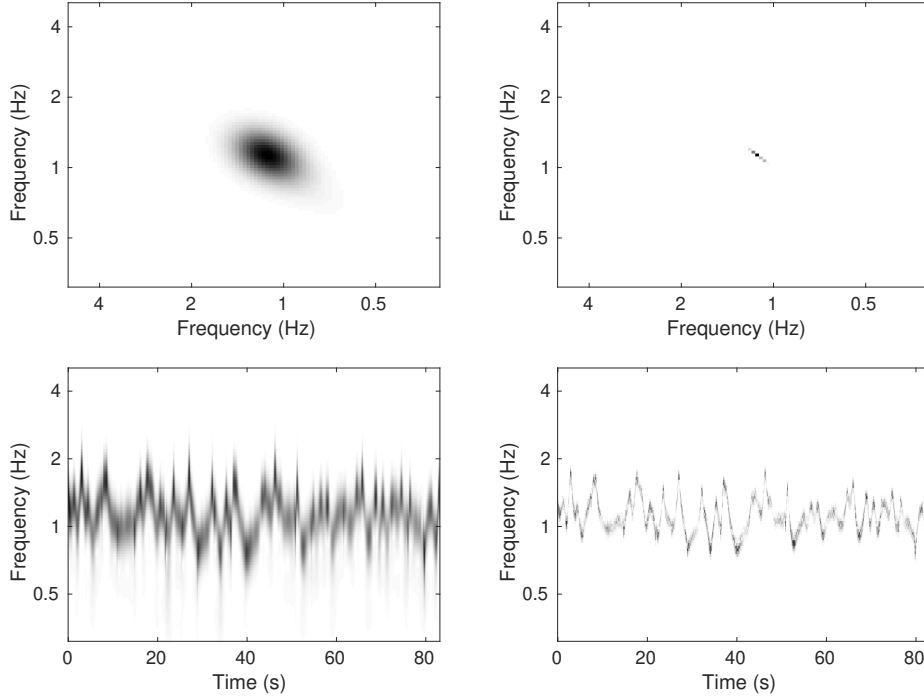


FIGURE 4.8 – Signal semi-réel. En haut : Deux formes de matrices de covariance *a priori*. En bas : Représentations temps-échelle associées.

où ς est choisie telle que $\varsigma = \log_q \left(\frac{\xi_0}{\nu_1} \right)$, et σ_s est un paramètre de réglage de la concentration fréquentielle. Cette forme doit nous permettre d'obtenir une représentation très parcimonieuse car, à chaque instant, l'énergie est *a priori* concentrée sur un petit nombre d'échelles centrées autour de ς . On représente en haut à droite de la figure 4.8 la matrice de covariance de covariance $\mathbf{C}(0)$ associée à cette expression de f en comparaison à celle associée à l'expression (4.8) (en haut à gauche). En dessous, les représentations temps-échelle correspondant à chacune de ces expressions sont tracées. La deuxième forme *a priori* permet bien de rendre plus parcimonieuse la représentation temps-échelle. Elle est clairement adaptée à la représentation temps-échelle de ce type de signaux appelés localement harmoniques à fréquence rapide (voir chapitre 5). Par ailleurs, cette modification de la matrice de covariance *a priori* ne détériore pas la qualité de la reconstruction comme le montre la superposition en figure 4.9 des signaux reconstruits à partir des deux représentations temps-échelles obtenues au signal original.

JEFAS-S permet donc de construire des représentations temps-échelle suivant une approche à la synthèse, avec des propriétés de concentration temps-échelle très fortes. Sur ce type de signaux, JEFAS-S est compétitif (et parfois supérieur) à des techniques du type synchrosqueezing, abordées au chapitre 5.

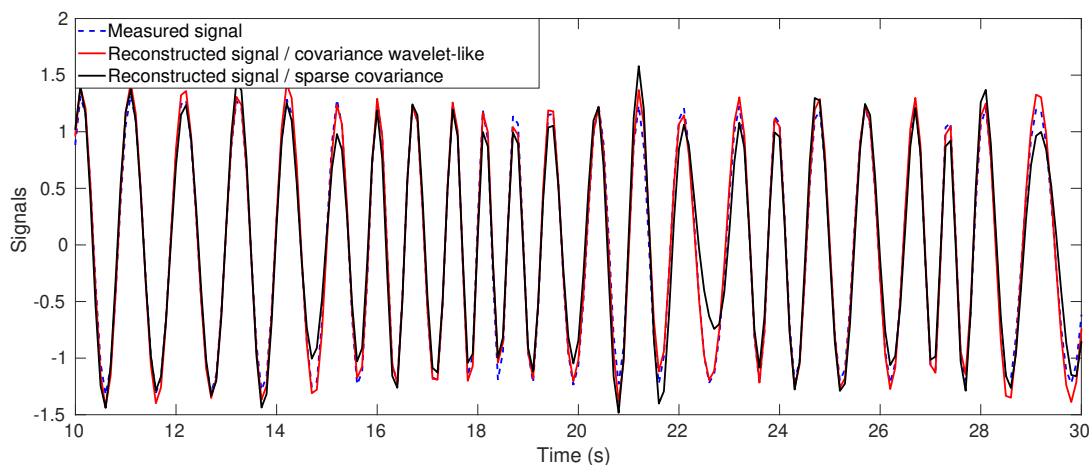


FIGURE 4.9 – Signal semi-réel. Zoom sur la superposition des signaux reconstruits à partir des représentations adaptées au signal original.

4.6 Bilan

Dans ce chapitre, nous avons présenté une approche à la synthèse pour l'étude spectrale des signaux localement dilatés. En s'inspirant des travaux de Turner et Sahani, nous avons proposé un modèle probabiliste de synthèse temps-échelle pour les signaux localement dilatés. Cependant, à l'inverse du modèle de Turner et Sahani qui considère la non-stationnarité dans le plan temps-échelle comme une modulation d'amplitude de chacune des lignes de la représentation temps-fréquence, nous la modélisons comme une translation verticale des coefficients temps-échelle. Ce modèle nous permet de s'inspirer des résultats du chapitre 2 pour construire un algorithme d'estimation spectrale adapté au signaux localement dilatés. Listons les différences et points commun entre JEFAS-S et JEFAS.

1. La différence majeure entre ces deux algorithmes concerne la construction simultanée d'une représentation temps-échelle adaptée au signal fournie par l'algorithme JEFAS-S grâce à l'approche à la synthèse.
2. L'approche à la synthèse permet d'effectuer un débruitage non stationnaire des signaux localement dilatés.
3. Lorsque les variations de la dilation instantanée sont rapides, la présence d'interférences empêche la convergence de JEFAS tandis que JEFAS-S permet l'estimation de la fonction de dilatation instantanée.
4. Lorsque les variations de la dilation instantanée sont contrôlées et que l'on n'observe pas d'interférences sur la transformée en ondelettes du signal, JEFAS-S n'est pas plus performant que JEFAS.
5. Lorsque JEFAS et JEFAS-S convergent, le temps de calcul nécessaire à JEFAS-S est bien supérieur à celui de JEFAS.

6. La convergence de JEFAS-S vers un maximum local de la vraisemblance est prouvée comme cas particulier d'un algorithme EM. À l'inverse, la convergence de JEFAS reste empirique.

En résumé, JEFAS-S est particulièrement utile pour étudier les signaux de courte durée pour lesquels JEFAS ne converge pas. Dans les autres cas, on peut se contenter de l'estimation obtenue par JEFAS car celle-ci est satisfaisante.

Par ailleurs, nous pouvons profiter des degrés de liberté laissés par le choix de l'*a priori* sur la représentation temps-échelle pour construire des représentations temps-échelle nettement plus concentrées que la transformée en ondelettes. Cette propriété est particulièrement intéressante pour représenter le spectre instantané idéal des signaux localement harmoniques (voir la définition 1.19). Cependant, un signal localement harmonique peut être vu comme un signal localement dilaté uniquement à condition qu'il soit constitué d'une unique composante : c'est le cas du signal étudié dans le paragraphe 4.5.2. C'est pourquoi dès que le nombre de composantes formant le signal localement harmonique étudié est supérieur ou égal à deux, la mise en œuvre de JEFAS-S n'est plus possible. La construction d'une représentation temps-échelle concentrée adaptée à de tels signaux n'est donc pas permise par le modèle de synthèse considéré dans ce chapitre. Ainsi, dans le chapitre 5, nous allons présenter une approche à la synthèse adaptée à l'étude des signaux localement harmoniques, et qui n'est, en principe, pas limitée par le nombre de composantes. Contrairement à ce que nous venons de présenter, nous verrons que cette approche repose sur une approche à la synthèse déterministe.

SYNTHÈSE DE SIGNAUX LOCALEMENT HARMONIQUES

5.1 Introduction

Dans ce chapitre, on s'intéresse aux signaux localement harmoniques (voir la définition 1.19). Il s'agit d'un modèle permettant de décrire des signaux décomposables sous la forme d'une somme d'un nombre fini de composantes sinusoïdales d'amplitudes et de fréquences instantanées variant au cours du temps. Par exemple, ce modèle est utilisé en audio notamment en traitement de la parole [42, 48]. Il s'agit aussi d'un modèle de référence très utilisé en traitement des signaux musicaux [20]. Il est également utilisé pour modéliser certains types de signaux médicaux comme les électrocardiogrammes pour suivre la fréquence cardiaque notamment [82].

L'objectif de l'analyse de tels signaux sera de suivre l'évolution au cours du temps de l'amplitude et de la fréquence instantanées de chacune des composantes. Représenter le spectre instantané de ces signaux constitue une idée naturelle pour résoudre ce problème. En effet, on peut s'attendre à voir apparaître sur le spectrogramme (ou le scalogramme) de ces signaux, des courbes dans le plan temps-fréquence permettant de suivre l'évolution des fréquence-instantanées et dont l'intensité dépend de l'amplitude de ces composantes. Toutefois, ces représentations temps-fréquence sont peu précises car limitées par le principe d'incertitude. Par exemple, on a tracé sur la figure 5.1 le spectrogramme d'un signal localement harmonique à deux composantes (voir la section 5.5.2 pour la description de ce signal). Cette figure montre qu'il est difficile de suivre simultanément l'évolution des deux fréquence instantanées de ce signal localement harmonique, en particulier lorsque sa fréquence varie rapidement.

C'est pourquoi, dans ce chapitre, nous décrivons d'abord l'ensemble des méthodes d'analyse existantes permettant d'obtenir des représentations temps-fréquence pour lesquelles l'énergie est concentrée sur un petit nombre de coefficients (situés le long des fréquences instantanées). Nous proposons ensuite de construire une représentation temps-fréquence idéale approchée. Pour cela, nous adoptons une approche à la synthèse. Ainsi, après avoir décrit les différentes propriétés que l'on souhaite voir vérifiées

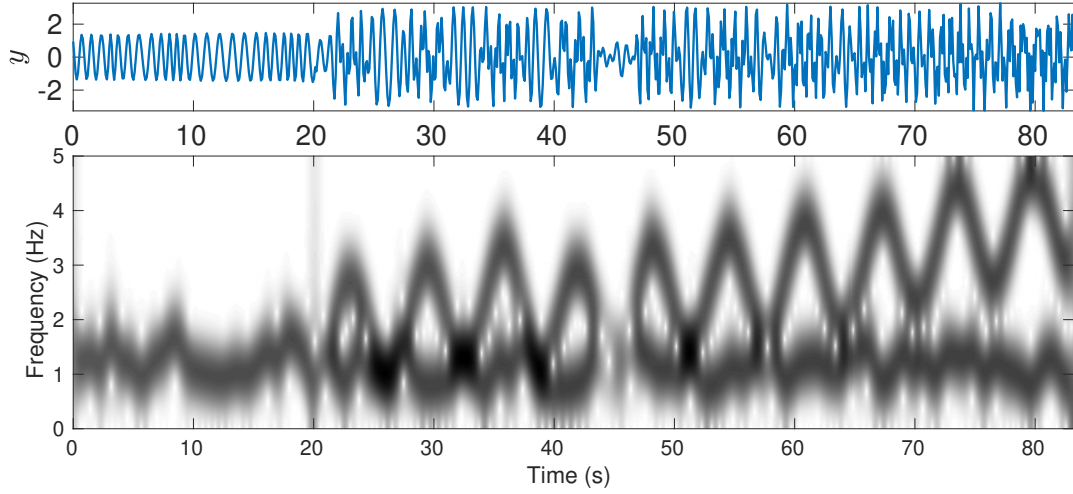


FIGURE 5.1 – Signal localement harmonique à deux composantes. Chronogramme (en haut) et spectrogramme (en bas).

par celle-ci, nous formulons un problème d’optimisation convexe dont elle est solution. Nous détaillons ensuite l’algorithme permettant de résoudre ce problème. Enfin, nous mettons cet algorithme en œuvre sur des signaux localement harmoniques et le comparons aux méthodes préexistantes.

5.2 Contexte

5.2.1 Modèle et objectif

Dans ce chapitre, on s’intéresse désormais aux signaux localement harmoniques décrits par la définition 1.19. On rappelle que ces signaux peuvent s’écrire sous la forme

$$y(t) = \sum_{k=1}^K A_k(t) \cos(2\pi\phi_k(t)) ,$$

où K désigne le nombre de composantes. Pour chacune des composantes, on appelle A_k son amplitude instantanée, ϕ_k sa phase instantanée, ϕ'_k sa fréquence instantanée, et ϕ''_k son *chirp factor*.

Pour se placer dans un cadre plus général que celui donné par la définition 1.19, on s’intéressera aussi aux signaux localement harmoniques à fréquences rapides. Les conditions définissant ce type de signaux permettent de prendre en compte un plus grand nombre de signaux présentant la même structure. Elles sont données par la définition 5.1, et sont moins contraignantes que celles données par la définition 1.19. En particulier, les variations des fréquences instantanées de chacune des composantes peuvent être grandes.

Définition 5.1 (Signaux localement harmoniques à fréquences rapides). Soit $K \geq 1$ un entier. Un signal localement harmonique à fréquences rapides est un signal s'écrivant sous la forme :

$$y(t) = \sum_{k=1}^K A_k(t) \cos(2\pi\phi_k(t)) , \quad (5.1)$$

où les amplitudes instantanées $A_k \in C^1(\mathbb{R})$ et les fréquences instantanées $\phi'_k \in C^2(\mathbb{R})$ satisfont les hypothèses suivantes :

1. Il existe des constantes c_1, c_2 et c_3 satisfaisant $0 < c_1 < c_2 < c_3 < \infty$ telles que pour tout $t \in \mathbb{R}$:

$$c_1 \leq A_k(t) \leq c_2 , \quad (5.2)$$

$$c_1 \leq \phi'_k(t) \leq c_2 , \quad (5.3)$$

$$|\phi''_k(t)| \leq c_3 . \quad (5.4)$$

2. Il existe une constante $\epsilon \ll 1$ satisfaisant $0 \leq \epsilon < c_2$ telle que pour tout $t \in \mathbb{R}$:

$$|A'_k(t)| \leq \epsilon \phi'_k(t) , \quad (5.5)$$

$$|\phi'''_k(t)| \leq \epsilon \phi'_k(t) . \quad (5.6)$$

3. Il existe une constante $d > 0$ telle que pour tout $t \in \mathbb{R}$:

$$\forall k \in \{1, \dots, K-1\}, \quad \phi'_{k+1}(t) - \phi'_k(t) > d . \quad (5.7)$$

Les conditions (5.2) et (5.3) assurent que chaque amplitude et fréquence instantanée reste strictement positive et bornée, tandis que les conditions (5.5) et (5.6) sont des conditions de régularité sur l'amplitude instantanée et le *chirp factor*. Sous ces différentes hypothèses, on peut montrer que l'unicité du modèle permettant de décrire un signal de ce type est approchée lorsque la constante de séparation d dans (5.7) est suffisamment grande (voir le théorème 2.2 dans [36] pour plus de détails).

L'objectif est d'obtenir une représentation temps-fréquence idéale de ce type de signaux. Dans ce cas, le spectre instantané idéal S_Y associé à ce type de signaux est donné par (1.43), ou de manière équivalente, en ne s'intéressant qu'aux fréquences positives, par

$$S_Y(\xi, t) = \sum_{k=1}^K A_k(t)^2 \delta(\xi - \phi'_k(t)) . \quad (5.8)$$

Dans ce chapitre, nous allons plutôt nous focaliser sur la construction d'une représentation temps-fréquence idéale de tels signaux. En effet, contrairement à l'expression (5.8) qui est quadratique vis-à-vis de l'amplitude instantanée, la représentation temps-fréquence est linéaire et inversible. Autrement dit, on peut reconstruire le signal

à partir de cette représentation. Par ailleurs, dans les chapitres précédents, la notion de spectre (instantané ou non) a été associée à des processus stochastiques. Or, dans ce chapitre, les signaux sont modélisés comme étant déterministes. C'est pourquoi dans ce cadre, il est plus pertinent de chercher à obtenir une représentation temps-fréquence idéale (abrégée RTFi). La RTFi des signaux localement harmoniques est alors définie, pour toute fréquence ξ positive, par

$$G_y(\xi, t) = \sum_{k=1}^K A_k(t) e^{2i\pi\phi_k(t)} \delta(\xi - \phi'_k(t)) . \quad (5.9)$$

Dans la section suivante, nous présentons différentes méthodes permettant d'estimer le plus précisément possible la RTFi d'un signal localement harmonique.

5.2.2 L'analyse des signaux localement harmoniques

L'objectif de l'analyse des signaux localement harmoniques est d'approcher le plus précisément possible la RTFi définie par (5.9). De manière équivalente, il s'agit de déterminer, à chaque instant, l'amplitude instantanée et la fréquence instantanée de chacune des composantes formant le signal à analyser.

Les outils d'analyse temps-fréquence et temps-échelle, comme la transformée de Fourier à court terme (cf. définition 1.14) et la transformée en ondelettes (cf. définition 1.15), peuvent naturellement fournir ce type d'informations. Néanmoins, la précision de ces représentations est limitée par le principe d'incertitude. Le théorème 4.1, énoncé au chapitre précédent, fournit une borne minimale qui ne peut pas être franchie par la concentration de la TFCT d'un signal quelconque de $L^2(\mathbb{R})$. Pour illustrer ce résultat, on peut approcher l'expression de la transformée de Fourier à court terme \mathcal{G}_y d'un signal localement harmonique y à une seule composante. Delprat *et al.* [23] ont établi le résultat du théorème 5.1 portant sur une approximation de \mathcal{G}_y (voir également [38] pour plus de détails).

Théorème 5.1. Soit $g \in L^2(\mathbb{R})$ une fenêtre d'analyse symétrique à support sur $[\frac{1}{2}, \frac{1}{2}]$, telle que sa transformée de Fourier \hat{g} soit symétrique réelle et maximale en zéro. Soit $y \in L^2(\mathbb{R})$ un signal. S'il existe des fonctions A et ϕ telles que $y(t) = A(t) \cos(2\pi\phi(t))$, alors, pour tout $\xi \geq 0$, on a :

$$\mathcal{G}_y(\xi, \tau) = \frac{1}{2} A(\tau) e^{2i\pi\phi(\tau)} (\hat{g}(\xi - \phi'(\tau)) + \epsilon(\xi, \tau)) . \quad (5.10)$$

Le terme correctif $\epsilon(\xi, \tau)$ vérifie

$$|\epsilon(\xi, \tau)| \leq \epsilon_{A,1} + \epsilon_{A,2} + \epsilon_{\phi,2} + \sup_{|v| > \phi'(\tau)} |\hat{g}(v)| ,$$

où

$$\begin{aligned}\epsilon_{A,1} &= \sup_{u \in [\tau - \frac{1}{2}, \tau + \frac{1}{2}]} \frac{A'(u)}{A(u)} , & \epsilon_{A,2} &= \sup_{u \in [\tau - \frac{1}{2}, \tau + \frac{1}{2}]} \frac{A''(u)}{A(u)} , \\ \epsilon_{\phi,2} &= \sup_{u \in [\tau - \frac{1}{2}, \tau + \frac{1}{2}]} \phi''(u) .\end{aligned}$$

Le terme correctif de l'expression (5.10) est d'autant plus faible que les fonctions A et ϕ' sont régulières et varient lentement vis-à-vis de la largeur de la fenêtre g . Si l'on néglige ce terme, on peut voir que lorsque deux composantes d'un signal localement harmonique présentent des fréquences instantanées proches, les bandes temps-fréquence associées à ces deux composantes se superposent. Ce phénomène apparaît dès que la différence entre les deux fréquences instantanées concernées est inférieure à la bande passante de la fenêtre d'analyse g . On voit alors apparaître le phénomène de battements aussi bien sur le chronogramme du signal avec la présence d'oscillations à la fréquence égale à la demi différence des deux fréquences instantanées, que sur le spectrogramme avec l'apparition de « bulles » [22]. Dans cette situation, on ne peut alors plus distinguer les deux fréquences instantanées.

Pour contourner ces différentes contraintes inhérentes à ces outils d'analyse, nous donnons ici une liste non exhaustive de méthodes qui ont été proposées.

Détection de crêtes. Cette méthode a été introduite dans [23]. Elle s'intéresse aux points de crêtes de la transformée de Fourier à court terme, c'est-à-dire les points où le spectrogramme présente un maximum local. À partir de l'expression approchée (5.10), on montre qu'à chaque instant τ , on trouve un point de crête à la fréquence $\xi(\tau) = \phi'(\tau)$ car \hat{g} est maximal en zéro. On estime alors l'amplitude instantanée à l'instant τ grâce à relation suivante :

$$A(t) = \frac{2 |\mathcal{G}_y(\xi(\tau), \tau)|}{\hat{g}(0)} .$$

D'autre part, si \hat{g} est réel, les points de crêtes sont également des points où la phase de \mathcal{G}_y vérifie l'équation suivante :

$$\partial_\tau \Phi_{\mathcal{G}_y}(\xi(\tau), \tau) = 2\pi \xi(\tau) .$$

où $\Phi_{\mathcal{G}_y}$ désigne la phase complexe de \mathcal{G}_y . Cette relation permet d'affiner l'estimation de ϕ' . On adapte ce résultat au cas des signaux localement harmoniques à K composantes en recherchant, à chaque instant, les maxima locaux du spectrogramme où la dérivée de la phase complexe s'annule. Notons que des résultats similaires peuvent être établis à partir de la transformée en ondelettes. En outre, d'autres méthodes de détections de crêtes ultérieures ont été proposées (voir [10], par exemple).

Réallocation. Cette méthode a été initialement proposée par Kodera *et al.* [34], puis reprise et généralisée par Auger et Flandrin [2]. La principale motivation de la méthode de réallocation consiste à réduire et à compenser la limitation de la résolution temps-fréquence introduite par le principe d'incertitude. En d'autres termes, on souhaite obtenir une représentation temps-fréquence parcimonieuse. Pour cela, il s'agit de déplacer les coefficients temps-fréquence du spectrogramme dont les valeurs sont non négligeables vers la fréquence instantanée estimée dont ils sont le plus proche. L'idée est alors de concentrer l'énergie du spectrogramme sur un petit nombre de coefficients temps-fréquence. Formellement, en notant $\check{\mathcal{R}}_y$ le spectrogramme du signal y « réalloué », il s'exprime :

$$\check{\mathcal{R}}_y(\xi, \tau) = \iint_{\mathbb{R}^2} |\mathcal{G}(\nu, t)|^2 \delta(\tau - \check{\tau}_y(\nu, t)) \delta(\xi - \check{\xi}_y(\nu, t)) dt d\nu ,$$

où $\check{\tau}_y$ et $\check{\xi}_y$ désignent les opérations de réallocation temporelle et fréquentielle. Ainsi, la réallocation préserve l'énergie du spectrogramme. En effet, on a

$$\int_{\mathbb{R}^2} \check{\mathcal{R}}_y(\xi, \tau) d\tau d\xi = \int_{\mathbb{R}^2} |\mathcal{G}_y(\xi, \tau)|^2 d\tau d\xi .$$

Par analogie avec l'équation (1.21) décrivant le spectrogramme comme le lissage de la distribution de Wigner-Ville V_y de y par celle de la fenêtre d'analyse g , on définit les opérations de réallocation comme étant les centres de gravité de cette distribution [3] :

$$\check{\tau}_y(\xi, \tau) = \frac{1}{|\mathcal{G}_y(\xi, \tau)|^2} \iint_{\mathbb{R}^2} \nu V_g(\nu - \xi, t - \tau) V_y(\nu, t) dt d\nu ; \quad (5.11)$$

$$\check{\xi}_y(\xi, \tau) = \frac{1}{|\mathcal{G}_y(\xi, \tau)|^2} \iint_{\mathbb{R}^2} t V_g(\nu - \xi, t - \tau) V_y(\nu, t) dt d\nu . \quad (5.12)$$

Ces opérations peuvent également s'exprimer à partir des transformées de Fourier à court terme du signal contre différentes fenêtres qui s'expriment en fonction de g . En pratique, cette reformulation permet le calcul de ces opérations de déplacement des coefficients. L'inconvénient majeur de cette méthode provient de la perte de l'information portée par la phase de la TFCT, celle-ci empêchant la reconstruction du signal y à partir de $\check{\mathcal{R}}_y$. Enfin, notons que la réallocation peut s'appliquer à d'autres représentations temps-fréquence quadratiques comme le scalogramme.

Synchsqueezing. Cette méthode a été proposée par Daubechies et Maes [19] puis revisitée plus récemment [18]. Le principe du synchsqueezing est le même que celui de la réallocation. Cette technique diffère de la précédente dans le sens où le déplacement des coefficients est effectué directement sur la transformée de Fourier à court terme (et non sur le spectrogramme) de sorte que l'information portée par la phase de la TFCT n'est pas perdue. Par ailleurs, le déplacement des coefficients s'effectue uniquement le long de l'axe fréquentiel. Notons $\check{\mathcal{T}}_y$ le synchsqueezing du signal y , il s'exprime donc par :

$$\check{\mathcal{T}}_y(\xi, \tau) = \frac{1}{g(0)} \int_{\mathbb{R}} \mathcal{G}_y(\nu, \tau) \delta(\xi - \check{\xi}_y(\nu, \tau)) d\nu , \quad (5.13)$$

où $\check{\xi}_y$ est donné par (5.12). On appellera cette représentation SST-TFCT. Un déplacement sur l'axe fréquentiel uniquement permet de conserver la causalité de la représentation, et on a alors une formule de reconstruction du signal original :

$$y(t) = \int_{\mathbb{R}} \check{\mathcal{T}}_y(\xi, t) d\xi. \quad (5.14)$$

Dans le cas d'un signal localement harmonique et avec une fenêtre bien choisie, cette formule permet de reconstruire chacune des composantes du signal en limitant l'intégrale (5.14) aux fréquences voisines de la composante à isoler. Récemment, plusieurs variantes du synchrosqueezing ont été proposées comme, par exemple, le synchrosqueezing au deuxième ordre [50] ou le synchrosqueezing horizontal. Remarquons que le synchrosqueezing peut également s'appliquer à la transformée en ondelettes, cette représentation sera abrégée SST-CWT.

Décomposition modale empirique (EMD). Cette méthode a été introduite en 1998 par Huang *et al.* [32]. Elle est assez différente des méthodes présentées précédemment dans la mesure où il s'agit d'une transformation non linéaire qui ne s'appuie pas sur une analyse temps-fréquence (linéaire ou quadratique) préalable. Il s'agit d'une méthode de synthèse. L'algorithme EMD permet de décomposer le signal y en une somme de K_{EMD} composantes oscillantes, notées \tilde{s}_k . Chacune des composante étant obtenue par synthèse soustractive. Le signal y est donc décomposé ainsi :

$$y(t) = \sum_{k=1}^{K_{EMD}} \tilde{s}_k(t) + r(t),$$

où r désigne le résidu produit par l'algorithme. Les détails sur le fonctionnement de l'algorithme EMD permettant d'extraire ces différentes composantes peuvent être trouvés dans [40] ou [49]. Signalons que la convergence de l'algorithme EMD reste mal comprise au niveau mathématique et diverses questions numériques restent ouvertes (comme par exemple la robustesse au bruit). L'étape suivante consiste à calculer la transformée de Hilbert $\mathcal{H}\tilde{s}_k$ de chacune des composantes \tilde{s}_k . Ceci permet d'obtenir le signal analytique (complexe) $\tilde{s}_k^a = \tilde{s}_k + i\mathcal{H}\tilde{s}_k$ dont la transformée de Fourier coïncide avec celle de s_k aux fréquences positives et s'annule aux fréquences négatives. Enfin, on peut décomposer \tilde{s}_k^a sous la forme $\tilde{s}_k^a(t) = \tilde{A}_k(t)e^{2i\pi\tilde{\phi}_k(t)}$. On obtient alors la représentation temps-fréquence issue de l'algorithme EMD puis de la transformée de Hilbert des composantes estimées, notée \mathcal{E} et abrégée EMD-HT telle que

$$\mathcal{E}(\xi, \tau) = \sum_{k=1}^{K_{EMD}} \tilde{A}_k(t) \delta(\xi - \tilde{\phi}'_k(t)).$$

Parmi les différentes méthodes d'analyse des signaux localement harmoniques que nous avons détaillées, le synchrosqueezing est celui qui répond le mieux aux différentes contraintes posées : représentation parcimonieuse, possibilité d'analyser des signaux

dont une composante n'est présente que sur un sous-intervalle de temps, robustesse au bruit. Cependant, cette représentation présente une limitation importante. Alors que des arguments mathématiques nous assurent que le synchrosqueezing permet d'analyser les signaux localement harmoniques dont les variations de fréquence sont contrôlées, l'analyse des signaux localement harmoniques à fréquences rapides est plus délicate. En effet, il peut arriver que le synchrosqueezing ne capture pas les fréquences instantanées variant rapidement et fasse apparaître des motifs indésirés dans le plan temps-fréquence.

Ces différents besoins que l'on a listé pour la représentation temps-fréquence idéale peuvent être formulés comme un certain nombre de contraintes à respecter par une représentation que l'on cherche. C'est pourquoi dans ce chapitre, nous allons suivre une approche à la synthèse et formuler un problème d'optimisation. Nous avons présenté l'ensemble des résultats développés dans ce qui suit dans l'article [36].

5.3 La représentation temps-fréquence idéale approchée

Dans le but de suivre une approche à la synthèse, nous n'allons pas construire une représentation temps-fréquence directement à partir du signal, mais nous allons définir une représentation temps-fréquence idéale approchée que l'on pourrait construire si l'on avait la connaissance exacte des amplitudes et des fréquences instantanées de chacune des composantes du signal localement harmonique. On déterminera ensuite des propriétés vérifiées par cette représentation comme par exemple : la formule de reconstruction du signal original ou une équation aux dérivées partielles. L'idée sera ensuite de chercher une fonction vérifiant le plus précisément possible ces propriétés afin qu'elle se rapproche le plus possible de la représentation temps-fréquence idéale.

La représentation temps-fréquence que l'on cherchera à obtenir doit permettre de distinguer l'évolution des K fréquences instantanées du signal y dans le plan temps-fréquence. Par ailleurs, elle doit permettre de reconstruire le signal. En pratique, on peut espérer approcher la RTFi G_y qui est une la distribution par une fonction à deux dimensions présentant des caractéristiques similaires. Nous allons donc chercher à obtenir la RTFi approchée à la résolution η dont on donne la définition.

Définition 5.2 (RTFi approchée à la résolution η). Soit y un signal local harmonique (cf. définition 1.19). Soit $h \in \mathcal{S}(\mathbb{R})$ une fonction à support compact dans $[-b, b]$ (où $b > 0$) vérifiant $\int h = 1$. La représentation temps-fréquence idéale approchée à la résolution η de y , notée \tilde{G}_y , est définie par

$$\tilde{G}_y(\xi, t) = \sum_{k=1}^K A_k(t) e^{2i\pi\phi_k(t)} h_\eta(\xi - \phi'_k(t)) , \quad (5.15)$$

$$\text{où } h_\eta(\xi) = \frac{1}{\eta} h\left(\frac{\xi}{\eta}\right).$$

Sous les hypothèses faites, h_η converge au sens faible vers la distribution de Dirac δ quand $\eta \rightarrow 0$, c'est à dire qu'on a $f * h_\eta \rightarrow f$ quand $\eta \rightarrow 0$ pour tout $f \in L^2(\mathbb{R})$. C'est pourquoi la représentation approchée \tilde{G}_y converge au sens faible vers G_y quand $\eta \rightarrow 0$. Dans ce qui suit, on se concentre donc sur la construction d'un problème d'optimisation dont la RTFi approchée serait solution.

Dans un premier temps, déterminons la formule de reconstruction associée à cette représentation. Soit d la distance minimale entre les composantes fréquentielles (voir la condition (5.7)). Si η est suffisamment petit pour que $b\eta < \frac{d}{2}$, on peut alors reconstruire le signal localement harmonique en sommant la partie réelle de \tilde{G}_y sur les fréquences. En effet dans ce cas, les composantes de la RTFi approchée sont disjointes, et on a :

$$\begin{aligned} \operatorname{Re} \left(\int_{\mathbb{R}} \tilde{G}_y(\xi, t) d\xi \right) &= \sum_{k=1}^K \int_{\mathbb{R}} A_k(t) \cos(2\pi\phi_k(t)) \frac{1}{\eta} h \left(\frac{\xi - \phi'_k(t)}{\eta} \right) d\xi \\ &= \sum_{k=1}^K A_k(t) \cos(2\pi\phi_k(t)) \int_{\phi'_k(t)-b\eta}^{\phi'_k(t)+b\eta} \frac{1}{\eta} h \left(\frac{\xi}{\eta} \right) d\xi \\ &= \sum_{k=1}^K A_k(t) \cos(2\pi\phi_k(t)) \\ &= y(t). \end{aligned} \tag{5.16}$$

D'autre part, le comportement fréquentiel de la RTFi approchée \tilde{G}_y n'étant pas caractérisé par la formule de reconstruction, on peut chercher à établir une relation de régularité vérifiée par \tilde{G}_y . Nous allons voir que cette régularité peut s'exprimer sous la forme d'une équation aux dérivées partielles. En effet, comme $A_k \in C^1(\mathbb{R})$, $\phi_k \in C^2(\mathbb{R})$ et $h \in \mathcal{S}(\mathbb{R})$, la RTFi approchée est dérivable par rapport au temps et à la fréquence. On montre alors que lorsque les fréquences instantanées vérifient la condition de régularité (1.41) (c'est-à-dire $|\phi''_k(t)| < \epsilon\phi'(k)$ où $0 < \epsilon \ll 1$), \tilde{G}_y vérifie l'équation aux dérivées partielles approchée (5.17) donnée par le théorème 5.2. Dans le cas moins contraignant des signaux localement harmoniques à fréquences rapides, \tilde{G}_y vérifie approximativement l'équation aux dérivées partielles (5.18), également fournie dans le théorème 5.2.

Théorème 5.2 (Équation aux dérivées partielles sur la RTFi approchée). *On distingue deux cas.*

1. Soit y un signal localement harmonique (cf. définition 1.19). Soit \tilde{G}_y sa représentation temps-fréquence idéale approchée à la résolution η définie par la définition (5.2). Alors, si $b\eta < \frac{d}{2}$, la RTFi approchée \tilde{G}_y vérifie approximativement l'équation aux dérivées partielles suivante :

$$\partial_t \tilde{G}_y(\xi, t) = 2i\pi\xi \tilde{G}_y(\xi, t) + E_1(\xi, t). \tag{5.17}$$

où E_1 désigne le terme correctif sur l'équation aux dérivées partielles. Pour tout intervalle $I \subset \mathbb{R}$ non vide, sa norme 2 est bornée comme suit :

$$\eta \|E_1\|_{L^2(\mathbb{R} \times I)}^2 \leq K|I|c_2^2 \left(\epsilon^2 J_{0,0,2} + 4\pi^2 \eta^2 J_{2,0,2} + \frac{c_2^2 \epsilon^2}{\eta^2} J_{0,1,2} \right) ,$$

où

$$J_{l,m,n} = \int_{\mathbb{R}} u^l |\partial_u^m h(u)|^n du .$$

2. Soit y un signal localement harmonique à fréquences rapides (cf. définition 5.1). Soit \tilde{G}_y sa RTFi approchée à la résolution η . Alors si $b\eta < \frac{d}{2}$, la RTFi approchée \tilde{G}_y vérifie l'équation aux dérivées partielles suivante :

$$\partial_t \tilde{G}_y(\xi, t) = 2i\pi \xi \tilde{G}_y(\xi, t) - \Omega_y(\xi, t) \partial_\xi \tilde{G}_y(\xi, t) + E_2(\xi, t) , \quad (5.18)$$

où

$$\Omega_y(\xi, t) = \begin{cases} \phi_k''(t) & \text{si } \xi \in [\phi_k'(t) - b\eta, \phi_k'(t) + b\eta] \\ 0 & \text{sinon} , \end{cases} \quad (5.19)$$

et E_2 désigne le terme correctif sur l'équation aux dérivées partielles. Pour tout intervalle $I \subset \mathbb{R}$ non vide, sa norme 2 est bornée comme suit :

$$\eta \|E_2\|_{L^2(\mathbb{R} \times I)}^2 \leq K|I|c_2^2 (\epsilon^2 J_{0,0,2} + 4\pi^2 \eta^2 J_{2,0,2}) .$$

On peut se référer à l'annexe A.4.1 pour les détails des calculs conduisant à ces équations aux dérivées partielles et pour la démonstration des expressions des bornes sur les termes correctifs. Dans les deux cas, ces équations aux dérivées partielles sont des équations de transport et modélise le déplacement des coefficients temps-fréquence le long de la fréquence instantanée. Notons d'ailleurs que l'équation (5.18) est assez proche de l'équation aux dérivées partielles établie par Clerc et Mallat sur les coefficients du scalogramme de signaux localement dilatés (voir l'équation (2.2)). En effet, le terme $2i\pi \xi \tilde{G}_y(\xi, t)$ est un paramètre permettant de décrire l'évolution de la phase de la RTFi, l'évolution de son amplitude est écrite par la relation entre les dérivées temporelle et fréquentielle. Celles-ci sont reliées par le paramètre de transport Ω_y contenant les *chirp factors*, tandis que dans l'équation (2.2) les dérivées temporelle et par rapport à l'échelle du scalogramme sont reliées par le paramètre de transport $(\ln(\gamma'))'$.

Pour les deux équations aux dérivées partielles du théorème 5.2, l'erreur d'approximation sur les équations aux dérivées partielles est d'autant plus faible que la constante ϵ est faible. En effet, cette constante dépend de la régularité des amplitudes instantanées et des fréquences instantanées (cas des signaux localement harmoniques) ou des *chirp factors* instantanés (cas des signaux localement harmoniques à fréquences rapides). Ainsi, si ces fonctions sont bien régulières, le modèle (5.15) est bien adapté pour décrire le comportement temps-fréquence du signal. Les erreurs d'approximation dépendent

également des intégrales $J_{l,m,n}$ qui caractérisent la concentration autour de l'origine de la fonction h et de ses dérivées. Ainsi, une erreur d'approximation faible demande que la fonction h et ses dérivées soient assez concentrées autour de l'origine pour rendre ces intégrales faibles. De la même manière, une valeur de η faible entraîne une fonction h concentrée autour de l'origine et une erreur d'approximation faible car la RTFi approchée tend vers la RTFi.

Le cas extrême où le terme correctif est nul est obtenu lorsque $\epsilon = 0$ et $\eta = 0$. Dans le cas d'un signal localement harmonique, la condition sur ϵ implique que ses amplitudes et fréquences instantanées sont constantes : $A_k(t) = A_k$ et $\phi'_k(t) = \phi'_k$ pour tout t . Il s'agit donc d'un signal stationnaire. Dans ce cas, sa RTFi s'écrit :

$$\begin{aligned}\tilde{G}_y(\xi, t) &= \sum_{k=1}^K A_k e^{2i\pi\phi_k(t)} \delta(\xi - \phi'_k) \\ &= \sum_{k=1}^K A_k e^{2i\pi\phi'_k t} \delta(\xi - \phi'_k) = \sum_{k=1}^K A_k e^{2i\pi\xi t} \delta(\xi - \phi'_k) \\ &= \left(\sum_{k=1}^K A_k \delta(\xi - \phi'_k) \right) e^{2i\pi\xi t}.\end{aligned}$$

L'équation aux dérivées partielles (5.17) devient donc exacte car on a bien :

$$\partial_t \tilde{G}_y(\xi, t) = 2i\pi\xi \tilde{G}_y(\xi, t).$$

Notons que la solution générale de cette équation se met sous la forme $C(\xi)e^{2i\pi\xi t}$ où C est une fonction déterminée par une condition initiale. Ici, il s'agit de la formule de reconstruction (5.16) qui permet de déterminer C . On obtient alors $C(\xi) = \sum_{k=1}^K A_k \delta(\xi - \phi'_k)$.

Dans le cas des signaux localement harmoniques à fréquences rapides, le terme correctif E_2 dans (5.18) s'annule lorsque ϵ et η sont nuls. Cela signifie que les amplitudes instantanées et les *chirp factors* sont constants. Il s'agit alors d'un signal composé de chirps linéaires :

$$y(t) = \sum_{k=1}^K A_k \cos(2\pi(a_k t^2/2 + b_k t + c_k)).$$

Le même type de raisonnement que précédemment nous permet de vérifier que la RTFi de y satisfait bien l'équation aux dérivées partielles (5.18) de manière exacte et que la résolution de celle-ci combinée avec la formule de reconstruction (5.16) permet de retrouver l'expression de la RTFi.

5.4 Formulation variationnelle et optimisation

5.4.1 Construction de la fonctionnelle

Dans une approche d'optimisation convexe, nous allons construire une fonctionnelle sur un espace de représentations temps-fréquence que nous chercherons à minimiser. L'idée est de construire une fonctionnelle convexe dont le minimum sera atteint

en \check{G} , où \check{G} sera aussi proche que possible de la RTFi approchée \check{G}_y du signal y . Notons $I \subset \mathbb{R}$ l'intervalle de temps sur lequel nous cherchons à estimer la RTFi approchée, la formule de reconstruction (5.16) suggère alors le terme d'attache aux données suivant :

$$\int_I \left| y(t) - \operatorname{Re} \left(\int_{\mathbb{R}} G(\xi, t) d\xi \right) \right|^2 dt. \quad (5.20)$$

Ce terme s'annule bien lorsque $G = \check{G}_y$. De plus, les équations aux dérivées partielles donnée par le théorème 5.2 (cf. équations (5.17) et (5.18)) suggèrent un terme de régularisation de la forme :

$$\eta \iint_{I \times \mathbb{R}} \left| \partial_t G(\xi, t) - 2i\pi\xi G(\xi, t) + \Omega(\xi, t) \partial_\xi G(\xi, t) \right|^2 dt d\xi, \quad (5.21)$$

où Ω est une fonction inconnue dans le cas des signaux localement harmoniques à fréquences rapides et $\Omega = 0$ dans les cas où les fréquences sont lentes. Il faudra donc également optimiser la fonctionnelle par rapport à Ω . La fonctionnelle doit également être construite de sorte que son minimiseur $\check{\Omega}$ soit le plus proche possible de Ω_y . Dans ce cas, le théorème 5.2 nous assure que ce terme de régularisation est faible lorsque $G = \check{G}_y$.

D'autre part, la RTFi approchée que l'on souhaite obtenir présente une structure parcimonieuse. En effet, la plupart des points temps-fréquence doivent être associés à une valeur nulle de la RTFi approchée. Pour obtenir une représentation parcimonieuse, on introduit un terme de régularisation supplémentaire : il s'agit de la norme 1 de G . Il a effectivement été montré qu'introduire un terme de pénalisation par la norme 1 dans un problème d'optimisation conduit à une solution parcimonieuse (voir [71]).

La fonctionnelle globale est alors donnée par :

$$\begin{aligned} \mathcal{H}(G, \Omega) = & \int_{\mathbb{R}} \left| y(t) - \operatorname{Re} \left(\int_{\mathbb{R}} G(\xi, t) d\xi \right) \right|^2 dt \\ & + \mu\eta \left(\lambda \iint_{\mathbb{R}^2} \left| \partial_t G(\xi, t) - 2i\pi\xi G(\xi, t) + \Omega(\xi, t) \partial_\xi G(\xi, t) \right|^2 dt d\xi \right. \\ & \left. + (1 - \lambda) \|G\|_{L^1(I \times \mathbb{R})} \right) + \gamma \|\Omega\|_{L^2(I \times \mathbb{R})}^2, \end{aligned} \quad (5.22)$$

où μ , λ et γ sont appelés hyperparamètres et sont choisis de façon à donner plus ou moins de poids à chacun des termes de la fonctionnelle.

Dans ce qui suit, nous considérerons deux cas particuliers :

- le cas des signaux localement harmoniques à fréquences rapides formés d'une unique composante. Dans ce cas, la fonction temps-fréquence Ω par rapport à laquelle on souhaite optimiser \mathcal{H} peut être remplacée par une fonction α dépendant uniquement du temps dans la mesure où dans ce cas, $\Omega_y \partial_\xi \check{G}_y = \phi'' \partial_\xi \check{G}_y$. On voudra donc obtenir un minimiseur $\check{\alpha}$ aussi proche que possible du *chirp factor* ϕ'' .

- le cas des signaux localement harmoniques (dont les fréquences instantanées sont contrôlées) à plusieurs composantes. Puisque ces signaux vérifient l'équation aux dérivées partielles (5.17), l'algorithme associé à cette situation est identique à ce qui est présenté dans la suite à la différence que l'étape d'optimisation par rapport à Ω est négligée. Autrement dit, on conserve $\Omega = 0$ tout au long de l'algorithme d'optimisation.

5.4.2 Algorithme

5.4.2.1 Discrétisation de la fonctionnelle

Numériquement, nous avons affaire à des signaux discrets. C'est pourquoi la RTFi que l'on va estimer sera discrétisée. Soit \mathbf{t} le vecteur de taille N_t contenant l'ensemble des instants de mesure et ξ le vecteur de taille M_ξ contenant l'ensemble des fréquences pour lesquelles on estime G . On se limite ici au cas d'un échantillonnage uniforme, la fréquence d'échantillonnage étant notée F_s . On notera $\Delta_t = \frac{1}{F_s}$ le pas d'échantillonnage temporel. Par ailleurs, on supposera que la fréquence d'échantillonnage est suffisamment grande pour que le signal échantillonné ne présente pas de repliement de spectre. Par ailleurs, la RTFi approchée \tilde{G}_y s'annulant aux fréquences nulles, on construit le vecteur de fréquences ξ comme un ensemble de M_ξ fréquences linéairement espacées allant de la fréquence nulle à la fréquence de Nyquist $F_s/2$. On notera $\Delta_\xi = \frac{F_s}{2(M_\xi-1)}$ le pas d'échantillonnage fréquentiel correspondant.

On va alors rechercher à minimiser une fonction coût discrétisée par rapport à la matrice temps-fréquence $\mathbf{G} \in \mathbb{R}^{M_\xi \times N_t}$ et au *chirp factor* discrétisé $\alpha \in \mathbb{R}^{N_t}$. Soit $\mathbf{y} = (y(t_1) \cdots y(t_{N_t}))^T$ le signal discrétisé. Les différentes intégrales constituant la fonctionnelle \mathcal{H} sont approchées grâce à la méthode des rectangles. Après avoir redéfini les hyperparamètres, la nouvelle fonction coût discrétisée \mathcal{H}_d s'exprime sous la forme :

$$\begin{aligned} \mathcal{H}_d(\mathbf{G}, \alpha) = & \|\mathbf{y}^T - \Delta_\xi \operatorname{Re}(\mathbf{1}^T \mathbf{G})\|_2^2 + \tilde{\mu} \lambda \|\partial_t \mathbf{G} - 2i\pi \operatorname{diag}(\xi) \mathbf{G} + \partial_\xi \mathbf{G} \operatorname{diag}(\alpha)\|_2^2 \\ & + \tilde{\mu}(1 - \lambda) \|\mathbf{G}\|_1 + \tilde{\gamma} \|\alpha\|_2^2, \end{aligned} \quad (5.23)$$

où $\mathbf{1}$ est le vecteur colonne de taille M_ξ dont tous les coefficients sont égaux à 1. On peut exprimer les hyperparamètres de la fonction coût (5.23) en fonction de ceux de la fonctionnelle (5.22). Alors, on a $\tilde{\mu} = \mu \eta \Delta_\xi$ et $\tilde{\gamma} = \gamma \Delta_\xi$.

Les opérateurs linéaires de dérivées partielles discrétisées $\partial_t G$ et $\partial_\xi G$ peuvent être mis en œuvre grâce à la méthode des différences finies. Cependant, ce choix risque de conduire à des instabilités numériques. Ici, nous nous placerons plutôt dans le domaine de Fourier discret où les calculs de dérivées partielles reviennent à effectuer des multiplications par des matrices diagonales. En effet, à temps continu, on a :

$$\mathcal{F} [\partial_t G(\xi, \cdot)] (\nu) = 2i\pi \nu \mathcal{F} [G(\xi, \cdot)] (\nu),$$

où \mathcal{F} désigne la transformée de Fourier. Exprimons ce dernier résultat sous forme discrétisée. Soit \mathbf{F}_t la matrice de Fourier de taille $N_t \times N_t$ définie par $[\mathbf{F}_t]_{mn} = e^{2i\pi nm/N_t}$.

Alors, la transformée de Fourier discrète de $\partial_t \mathbf{G}$ par rapport à la variable temporelle t s'obtient par

$$\mathbf{F}_t \partial_t \mathbf{G}^T = \text{diag}(2i\pi\nu) \mathbf{F}_t \mathbf{G}^T, \quad \text{où } \nu_k = \frac{(k-1)F_s}{N_t}, \quad k \in \{1, \dots, N_t\}.$$

Enfin, on obtient la transformée de Fourier inverse discrète en appliquant la matrice $\mathbf{F}_t^{-1} = \frac{1}{N_t} \mathbf{F}_t^H$. Des calculs similaires donnent la dérivée partielle par rapport à la fréquence $\partial_\xi \mathbf{G}$ en prenant la matrice de Fourier de taille $M_\xi \times M_\xi$ telle que $[\mathbf{F}_\xi]_{mn} = e^{2i\pi nm/M_\xi}$. Finalement, en notant τ le vecteur de taille M_ξ telle que $\tau_k = \frac{(k-1)}{\Delta_\xi M_\xi}$, les dérivées partielles discrétisées sont données par :

$$\begin{aligned} \partial_t \mathbf{G} &= \frac{1}{N_t} \mathbf{G} \mathbf{F}_t^T \text{diag}(2i\pi\nu) \overline{\mathbf{F}_t}; \\ \partial_\xi \mathbf{G} &= \frac{1}{M_\xi} \mathbf{F}_\xi^H \text{diag}(2i\pi\tau) \mathbf{F}_\xi \mathbf{G}. \end{aligned}$$

5.4.2.2 Minimisation de la fonctionnelle

La fonction \mathcal{H}_d dépend de deux variables : \mathbf{G} et α . Bien que l'algorithme PALM étudié dans [7] fournisse une procédure simple pour la minimisation de \mathcal{H}_d , cet algorithme converge très lentement en pratique. Néanmoins, comme la fonctionnelle \mathcal{H}_d est convexe par rapport à chacun des deux termes à estimer, il est naturel de mettre en œuvre une stratégie d'optimisation alternée. En d'autres termes, la k -ème itération de l'algorithme se décompose de la manière suivante :

$$\begin{cases} \mathbf{G}^{(k+1)} = \arg \min_{\mathbf{G}} \mathcal{H}(\mathbf{G}, \alpha^{(k)}) \\ \alpha^{(k+1)} = \arg \min_{\alpha} \mathcal{H}(\mathbf{G}^{(k+1)}, \alpha) \end{cases} \quad (5.24)$$

On trouvera un résultat sur la convergence de cette procédure d'optimisation dans [36].

Minimisation de $\mathcal{H}_d(\mathbf{G}^{(k+1)}, \cdot)$. Vis-à-vis de α , la fonction est quadratique, on obtient donc une forme explicite pour le minimiseur. En annulant le gradient de \mathcal{H}_d par rapport à α , on trouve :

$$\alpha_n^{(k+1)} = - \frac{\sum_{m=0}^{M_\xi} \text{Re} \left(\overline{[\partial_\xi \mathbf{G}^{(k)}]_{mn}} \left([\partial_t \mathbf{G}^{(k)}]_{mn} - 2i\pi \xi_m [\mathbf{G}^{(k)}]_{mn} \right) \right)}{\sum_{m=0}^{M_\xi} |[\partial_\xi \mathbf{G}^{(k)}]_{mn}|^2 + \frac{\tilde{\gamma}}{\tilde{\mu}\lambda}}.$$

Sous forme compacte, cela donne :

$$\alpha^{(k+1)} = - \frac{\mathbf{1}^T \text{Re} \left(\overline{\partial_\xi \mathbf{G}^{(k)}} \circ \left(\partial_t \mathbf{G}^{(k)} - 2i\pi \text{diag}(\xi) \mathbf{G}^{(k)} \right) \right)}{\mathbf{1}^T |\partial_\xi \mathbf{G}^{(k)}|^2 + \frac{\tilde{\gamma}}{\tilde{\mu}\lambda}}, \quad (5.25)$$

où le signe \circ désigne le produit terme à terme, aussi appelé produit de Hadamard.

Minimisation de $\mathcal{H}_d(\cdot, \alpha^{(k)})$. Lorsque α est fixé, il s'agit maintenant de minimiser la fonction coût

$$\mathcal{H}_{\alpha^{(k)}} = \mathcal{G}_{\alpha^{(k)}}(\mathbf{G}) + \tilde{\mu}(1 - \lambda)\Psi(\mathbf{G}),$$

où

$$\begin{aligned} \mathcal{G}_{\alpha^{(k)}}(\mathbf{G}) &= \|\mathbf{y}^T - \Delta_{\xi} \operatorname{Re}(\mathbf{1}^T \mathbf{G})\|_2^2 + \tilde{\mu}\lambda \left\| \partial_t \mathbf{G} - 2i\pi \operatorname{diag}(\xi) \mathbf{G} + \partial_{\xi} \mathbf{G} \operatorname{diag}(\alpha^{(k)}) \right\|_2^2, \\ \Psi(\mathbf{G}) &= \|\mathbf{G}\|_1. \end{aligned}$$

La fonction $\mathcal{H}_{\alpha^{(k)}}$ est convexe mais n'est pas différentiable. En effet, le terme $\mathcal{G}_{\alpha^{(k)}}$ est convexe et Lipschitz différentiable tandis que le terme Ψ est un terme convexe non différentiable. Différents algorithmes, appelés algorithmes proximaux, ont été proposés pour résoudre ce type de problèmes d'optimisation convexe. Ici, nous mettons en œuvre l'algorithme FISTA (*Fast Iterative Shrinkage Algorithm*, voir [4]). Cet algorithme possède l'avantage de converger rapidement. En effet, si l'on note par $\check{\mathbf{G}}$ un point de convergence de FISTA et par \mathbf{G}_l la valeur donnée par l'algorithme après l itérations, on a : $\mathcal{H}_{\alpha^{(k)}}(\mathbf{G}_l) - \mathcal{H}_{\alpha^{(k)}}(\check{\mathbf{G}}) = \mathcal{O}\left(\frac{1}{k^2}\right)$ [11]. Le schéma global de FISTA est résumé par l'algorithme 5.1. FISTA repose principalement sur trois étapes que l'on détaille ci-

Algorithme 5.1 Algorithme FISTA pour $\mathcal{H}_{\alpha^{(k)}} : \mathbf{G}^{(k+1)} = \text{FISTA}(\mathbf{G}_0, \alpha^{(k)}, T_1)$

Initialiser \mathbf{G} par \mathbf{G}_0 et $\mathbf{z}_0 = \mathbf{G}_0$.

Évaluer la constante de Lipschitz L de $\nabla \mathcal{G}_{\alpha^{(k)}}$ par la méthode des puissances itérées.

$l \leftarrow 0$

tant que $\frac{\|\mathbf{G}_{l+1} - \mathbf{G}_l\|}{\|\mathbf{G}_l\|} > T_1$ **faire**

 Étape de gradient : $\mathbf{G}_{l+1/2} \leftarrow \mathbf{z}_l - \frac{1}{L} \nabla \mathcal{G}_{\alpha^{(k)}}(\mathbf{z}_l)$ (voir l'équation (5.28)).

 Étape proximale : $\mathbf{G}_{l+1/2} \leftarrow \mathbf{G}_{l+1/2} \left(1 - \frac{\tilde{\mu}(1-\lambda)/L}{\|\mathbf{G}_{l+1/2}\|}\right)^+$.

 Étape monotone :

si $\mathcal{G}_{\alpha^{(k)}}(\mathbf{G}_{l+1/2}) < \mathcal{G}_{\alpha^{(k)}}(\mathbf{G}_l)$ **alors**

$\mathbf{G}_{l+1} \leftarrow \mathbf{G}_{l+1/2}$,

sinon

$\mathbf{G}_{l+1} \leftarrow \mathbf{G}_l$.

fin si

 Étape de relaxation : $\mathbf{z}_{l+1} \leftarrow \mathbf{G}_l + \frac{l}{l+2}(\mathbf{G}_{l+1} - \mathbf{G}_l) + \frac{l+1}{l+2}(\mathbf{G}_{l+1/2} - \mathbf{G}_l)$.

$l \leftarrow l + 1$.

fin tant que

$\mathbf{G}^{(k+1)} = \mathbf{G}_l$.

dessous.

1. *Descente de gradient.* Il s'agit d'effectuer d'effectuer une étape de descente de gradient sur le terme différentiable \mathcal{G}_{α} . Cette étape nécessite donc la connaissance du gradient de $\mathcal{G}_{\alpha^{(k)}}$. Pour cela décomposons la fonctionnelle $\mathcal{G}_{\alpha^{(k)}}$ en deux termes :

$$\mathcal{G}_{\alpha^{(k)}}(\mathbf{G}) = \mathcal{G}_1(\mathbf{G}) + \tilde{\mu}\lambda \mathcal{G}_2(\mathbf{G}),$$

avec

$$\begin{aligned}\mathcal{G}_1(\mathbf{G}) &= \|\mathbf{y}^T - \Delta_{\xi} \text{Re}(\mathbf{1}^T \mathbf{G})\|_2^2 \\ \mathcal{G}_2(\mathbf{G}) &= \|\mathcal{J}\mathbf{G}\|_2^2 ,\end{aligned}$$

où

$$\begin{aligned}\mathcal{J} : \mathbb{C}^{M_{\xi} \times N_t} &\rightarrow \mathbb{C}^{M_{\xi} \times N_t} \\ \mathbf{G} &\mapsto \partial_t \mathbf{G} - 2i\pi \text{diag}(\xi) \mathbf{G} + \partial_{\xi} \mathbf{G} \text{diag}(\alpha^{(k)}) .\end{aligned}$$

On note $\nabla \mathcal{G}_{\alpha^{(k)}} \in \mathbb{C}^{M_{\xi} \times N_t}$ le gradient de $\mathcal{G}_{\alpha^{(k)}}$, il s'écrit comme la somme des gradients de chacun des deux termes. Concernant le premier terme, un calcul de dérivée matricielle nous donne :

$$\nabla \mathcal{G}_1(\mathbf{G}) = 2\Delta_{\xi} \mathbf{1} \left(\Delta_{\xi} \text{Re}(\mathbf{1}^T \mathbf{G}) - \mathbf{y}^T \right) . \quad (5.26)$$

Concernant le second terme, on peut écrire :

$$\nabla \mathcal{G}_2(\mathbf{G}) = 2\mathcal{J}^* \mathcal{J} \mathbf{G} ,$$

où \mathcal{J}^* désigne l'adjoint de \mathcal{J} . Or, pour tout \mathbf{G} et $\mathbf{H} \in \mathbb{C}^{M_{\xi} \times N_t}$:

$$\begin{aligned}\langle \mathcal{J}\mathbf{G}, \mathbf{H} \rangle &= \left\langle \partial_t \mathbf{G} - 2i\pi \text{diag}(\xi) \mathbf{G} + \partial_{\xi} \mathbf{G} \text{diag}(\alpha^{(k)}) , \mathbf{H} \right\rangle \\ &= \left\langle \mathbf{G} , -\partial_t \mathbf{H} + 2i\pi \text{diag}(\xi) \mathbf{H} - \partial_{\xi} \mathbf{H} \text{diag}(\alpha^{(k)}) \right\rangle \\ &= \langle \mathbf{G} , \mathcal{J}^* \mathbf{H} \rangle .\end{aligned}$$

On en déduit que :

$$\begin{aligned}\nabla \mathcal{G}_2(\mathbf{G}) &= 2 \left(-\partial_t^2 \mathbf{G} + 4i\pi \text{diag}(\xi) \partial_t \mathbf{G} - \partial_t \left(\partial_{\xi} \mathbf{G} \text{diag}(\alpha^{(k)}) \right) + 4\pi^2 \text{diag}(\xi^2) \mathbf{G} \right. \\ &\quad + 2i\pi \text{diag}(\xi) \partial_{\xi} \mathbf{G} \text{diag}(\alpha^{(k)}) - \partial_{\xi} \partial_t \mathbf{G} \text{diag}(\alpha^{(k)}) \\ &\quad \left. + 2i\pi \partial_{\xi} (\text{diag}(\xi) \mathbf{G}) \text{diag}(\alpha^{(k)}) - \partial_{\xi}^2 \mathbf{G} \text{diag}(\alpha^{(k)})^2 \right) .\end{aligned} \quad (5.27)$$

Enfin, on obtient $\nabla \mathcal{G}_{\alpha^{(k)}}$ en combinant les résultats des équations (5.26) et (5.27) pour calculer :

$$\nabla \mathcal{G}_{\alpha^{(k)}}(\mathbf{G}) = \nabla \mathcal{G}_1(\mathbf{G}) + \tilde{\mu} \lambda \nabla \mathcal{G}_2(\mathbf{G}) . \quad (5.28)$$

Notons que l'étape de descente de gradient nécessite également la connaissance de la constante de Lipschitz L du gradient de la fonctionnelle $\mathcal{G}_{\alpha^{(k)}}$, celle-ci peut être évaluée grâce à la méthode des puissances itérées.

2. *Seuillage doux*. Cette étape, aussi appelée étape proximale, consiste à effectuer un seuillage des coefficients du résultat de l'étape précédente. Les coefficients dont la valeur absolue est inférieure à un seuil donné sont alors mis à zéro. Plus précisément, l'étape de seuillage doux de la matrice temps-fréquence \mathbf{G} s'écrit :

$$\text{signe}(\mathbf{G}) \left| \mathbf{G} - \frac{\tilde{\mu}(1-\lambda)}{L} \right|^+,$$

où $(x)^+ = \max(x, 0)$ pour tout $x \in \mathbb{R}$.

3. *Relaxation*. Cette étape consiste à amplifier « l'élan » des deux étapes précédentes si elles ont permis de faire décroître la fonctionnelle. Soit \mathbf{G}_l le résultat précédent les étapes de descente de gradient et de seuillage doux, et soit $\mathbf{G}_{l+1/2}$ le résultat obtenue à l'issue de ces deux étapes. Dans le cas où la fonctionnelle prend une valeur plus faible en $\mathbf{G}_{l+1/2}$ qu'en \mathbf{G}_l , on prolonge la direction de descente $\mathbf{G}_{l+1/2} - \mathbf{G}_l$, en se plaçant au point :

$$\mathbf{z}_{l+1} = \mathbf{G}_{l+1/2} + \frac{l-1}{l+2}(\mathbf{G}_{l+1/2} - \mathbf{G}_l).$$

Dans le cas où la fonctionnelle n'a pas décro, on se place en un point situé entre \mathbf{G}_l et $\mathbf{G}_{l+1/2}$ tel que :

$$\mathbf{z}_{l+1} = \frac{l+1}{l+2}\mathbf{G}_{l+1/2} + \frac{1}{l+2}\mathbf{G}_l.$$

Dans les deux cas, l'algorithme se poursuit par l'application de l'étape de gradient au point \mathbf{z}_{l+1} .

Algorithme général. L'algorithme global est nommé TYCOON qui est l'abréviation de *Time-frequency bY CONVex OptimizatiON*. Les étapes de TYCOON sont synthétisées dans l'algorithme 5.2. On discute des choix qui sont fait pour les différents paramètres et pour l'initialisation de l'algorithme.

Critère d'arrêt. Pour stopper les algorithmes FISTA et TYCOON, on a choisi un critère d'arrêt classique : il s'agit de comparer l'innovation relative à un seuil donné. En pratique, le choix $T_1 = T_2 = 5 \times 10^{-4}$ convient. Des valeurs plus faibles de ces seuils conduisent à un algorithme plus lent pour des résultats similaires.

Ensemble de valeurs $I_{\tilde{\mu}}$. Comme on l'a montré précédemment, on a $\tilde{\mu} = \mu\eta\Delta_{\xi}$. Ainsi, comme l'on souhaite obtenir une RTFi approchée la plus fine possible, η doit être petit. Par conséquent, $\tilde{\mu}$ doit également être très faible. Cependant, un choix de $\tilde{\mu}$ trop faible entraîne une convergence très lente de TYCOON. C'est pourquoi un choix efficace en pratique est de prendre un ensemble de valeurs successives décroissantes pour $\tilde{\mu}$, les premières valeurs, plutôt grandes, entraînent une convergence rapide vers une solution grossière, tandis que les dernière valeurs, plutôt

faibles, permette d'affiner la concentration de la RTFi approchée estimée. En pratique, $I_{\tilde{\mu}}$ est un ensemble de valeurs décroissantes, régulièrement espacées sur une échelle logarithmique.

Paramètre λ . Ce paramètre doit être choisi entre 0 et 1. λ est d'autant plus proche de 1 que l'on donne de l'importance aux terme de régularisation sur les dérivées partielles. Comme l'équation aux dérivées partielles doit être satisfaite avec autant de précision que possible, en pratique, on choisit $\lambda \approx 0,95$. Les résultats numériques présentés dans la section 5.5.1 permettront de mettre en évidence comment effectuer qualitativement le réglage de λ .

Paramètre $\tilde{\gamma}$. En pratique, l'influence de ce paramètre n'est pas décisive sur le résultat de TYCOON. On fixe $\gamma \approx 10^{-3}$ de façon à éviter toute division par zéro pendant l'étape d'estimation de α donné par (5.25).

Initialisation. Le choix $\alpha^{(0)} = \mathbf{0}$ semble naturel dans la mesure où nous n'avons pas accès au *chirp factor*. La première itération de TYCOON est donc équivalente à une estimation sans prendre en compte le *chirp factor*. Néanmoins, l'initialisation a une influence sur la vitesse de l'algorithme. Comme la solution \mathbf{G} est censée être parcimonieuse, le choix $\mathbf{G}^{(0)} = \mathbf{0}$ semble être raisonnable.

Algorithme 5.2 Algorithme TYCOON pour la minimisation de \mathcal{H}

Choisir une seuil T_1 pour FISTA.
 Choisir une seuil T_2 pour la minimisation alternée.
 Choisir un ensemble $I_{\tilde{\mu}}$ de valeurs décroissantes pour le paramètre $\tilde{\mu}$.
 Choisir les paramètres $\tilde{\lambda}$ et $\tilde{\gamma}$.
 Initialiser : $k \leftarrow 0$, $\mathbf{G}^{(0)} \leftarrow \mathbf{0}$ et $\alpha^{(0)} \leftarrow \mathbf{0}$
tant que $\frac{\|\mathbf{G}^{(k)} - \mathbf{G}^{(k-1)}\|}{\|\mathbf{G}^{(k)}\|} > T_2$ **faire**
 Étape FISTA : $\mathbf{G}^{(k+1)} = \text{FISTA}(\mathbf{G}^{(k)}, \alpha^{(k)}, T_1)$.
 Étape d'estimation de $\alpha^{(k+1)}$: voir l'équation (5.25).
 $k \leftarrow k + 1$.
fin tant que
 Sorties : $\mathbf{G}^{(k)}$ et $\alpha^{(k)}$.

5.5 Résultats numériques

Dans cette section, nous présentons des résultats numériques donné par l'algorithme TYCOON appliqué à différents signaux simulés.

5.5.1 Application à un signal à une composante

Le premier exemple sur lequel nous évaluons TYCOON est un exemple semi-réel. Pour cela, on considère le signal localement harmonique à une seule composante y

décrit dans la section 4.5.2. On rappelle qu'on peut l'écrire sous la forme

$$y(t) = A_1(t) \cos(2\pi\phi_1(t)) .$$

La fréquence instantanée ϕ_1' provient de la mesure de la fréquence cardiaque d'une personne atteinte de fibrillation atriale. Or, cette pathologie entraîne des variations brutales de la fréquence cardiaque. C'est pourquoi le modèle de signal localement harmonique à fréquence rapide paraît adapté à la description de ce signal. L'amplitude instantanée est quant à elle simulée à partir d'une réalisation d'un mouvement brownien lissé. Le chronogramme du signal, ainsi que les graphes de la fréquence et de l'amplitude instantanées sont tracés sur la figure 4.6.

On applique alors TYCOON à ce signal en utilisant différents jeux d'hyperparamètres $\tilde{\mu}$ et λ . Notons $\tilde{\mathbf{G}}$ la représentation temps-fréquence issue de l'algorithme après convergence. Pour amplifier l'effet de contraste et atténuer la dynamique de l'image temps-fréquence, on trace $\mathcal{R}(\tilde{\mathbf{G}})$ donné par :

$$\mathcal{R}(\tilde{\mathbf{G}}) = \log(1 + |\tilde{\mathbf{G}}|)$$

Ainsi, les coefficients de $|\tilde{\mathbf{G}}|$ prenant des petites valeurs restent visibles dans le tracé de $\mathcal{R}(\tilde{\mathbf{G}})$ alors qu'ils sont indistinguables dans celui de $|\tilde{\mathbf{G}}|$. Les représentations temps-fréquence issues de ces différentes estimations de TYCOON sont ainsi tracées sur la figure 5.2. Ces résultats montrent qu'un choix trop petit pour $\tilde{\mu}$ (voir la colonne de droite où $\tilde{\mu} = 10^{-6}$) entraîne la construction d'une représentation temps-fréquence très peu concentrée en fréquence. En effet, dans ce cas, l'importance donnée au terme d'attache aux données est trop importante. Ainsi, la formule de reconstruction est bien respectée par la représentation ainsi obtenue mais la régularité temporelle et la parcimonie recherchées ne sont pas obtenues. À l'inverse, un choix de $\tilde{\mu}$ trop grand (voir la colonne de gauche où $\tilde{\mu} = 1$) entraîne une trop grande parcimonie de la représentation temps-fréquence obtenue, qui à l'extrême peut devenir identiquement nulle (cas où $\lambda = 0,95, 0,98$ ou $0,995$). Un choix de $\tilde{\mu}$ intermédiaire (voir la colonne centrale où $\tilde{\mu} = 10^{-3}$) constitue un compromis entre le recherche de parcimonie et la fidélité au terme d'attache aux données. Le réglage du paramètre λ permet ensuite de décider du poids à accorder à la parcimonie par rapport à la régularité temporelle. Un choix de λ trop grand (comme $\lambda = 0,9999$) conduit à une trop grande régularité temporelle, ce qui fait perdre la parcimonie fréquentielle recherchée. À l'inverse, un choix de λ trop petit (comme $\lambda = 0,95$) conduit à une trop grande variabilité temporelle de la représentation obtenue. Finalement, dans cet exemple, le choix $\tilde{\mu} = 10^{-3}$ et $\tilde{\lambda} = 0,98$ est celui qui permet d'avoir une représentation temps-fréquence satisfaisant le plus correctement possible l'ensemble des contraintes fixées.

Intéressons-nous maintenant aux résultats de TYCOON obtenus pour $\tilde{\mu}$ décroissant exponentiellement au fur et à mesure des itérations comme proposé en section 5.4.2. Les valeurs successivement prises par $\tilde{\mu}$ sont $1, 10^{-1}, 10^{-2}, \dots, 10^{-8}$. Ici, λ est fixé tel que $\lambda = 0,98$. Les résultats obtenus sont présentés en figure 5.3. À droite de la figure 5.3, on superpose cette image temps-fréquence à la fréquence instantanée ϕ_1' afin de mettre en évidence le fait que cette représentation temps-fréquence permet un bon suivi de

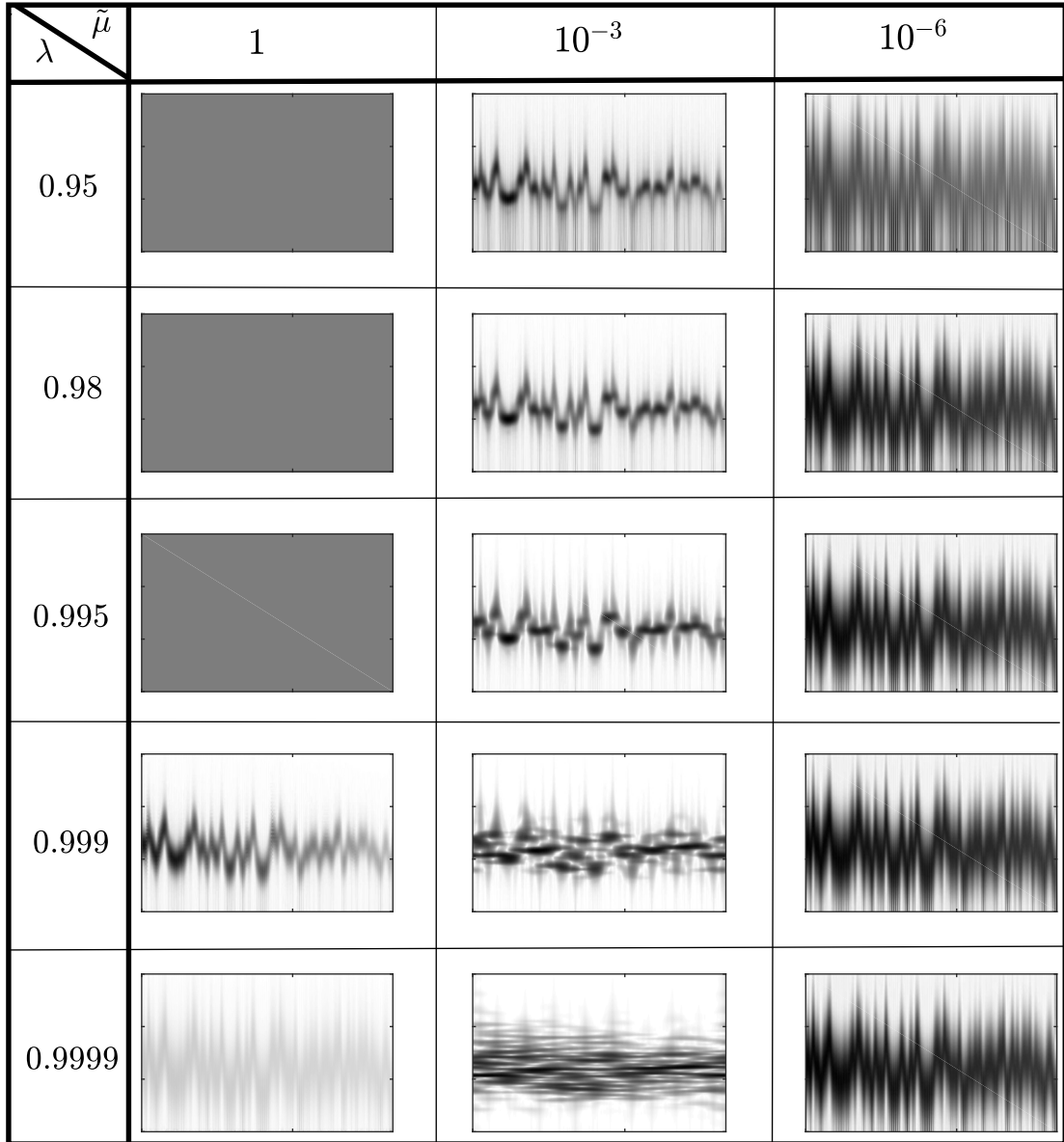


FIGURE 5.2 – Signal à une composante. Représentations temps-fréquences estimées par TYCOON pour différentes combinaisons d’hyperparamètres $\tilde{\mu}$ et λ .

l'évolution de la fréquence instantanée. Pour comparaison, nous avons tracé en bas de la même figure les résultats obtenus à partir de la TFCT avec une fenêtre gaussienne (de variance valant 1 seconde) et du synchrosqueezing qui en résulte. Dans les deux cas, le suivi de l'évolution de la fréquence instantanée reste beaucoup plus délicat que dans celui de TYCOON. En effet, bien que l'on puisse suivre une allure générale de ϕ'_1 , l'information sur la fréquence instantanée reste floue et des artefacts apparaissent. Enfin, au milieu la figure 5.3, nous avons tracé l'estimation du *chirp factor* ϕ''_1 superposé à la vraie fonction. TYCOON permet donc d'approcher ce *chirp factor* dans une certaine mesure.

On peut également comparer ce résultat à la représentation temps-échelle construite au chapitre 4. On remarque d'abord que la représentation obtenue grâce à un *a priori* de concentration (tracée en bas à droite de la figure 4.8) permet une meilleure localisation que celle obtenue par TYCOON. En effet, la représentation donnée par TYCOON reste assez étalée du fait de l'influence du paramètre η . Néanmoins, JEFAS-S reste limité aux signaux localement harmoniques à une seule composante. Étudions maintenant TYCOON sur un signal à deux composantes.

5.5.2 Évaluations quantitatives sur un signal à deux composantes

Dans cette section, nous allons considérer un signal localement harmonique à fréquence rapide y constitué de deux composantes. Il se décompose donc ainsi :

$$y(t) = A_1(t) \cos(2\pi\phi_1(t)) + A_2(t) \cos(2\pi\phi_2(t)) , \quad (5.29)$$

où l'amplitude instantanée A_1 et la fréquence instantanée ϕ'_1 sont identiques à l'exemple précédent. Concernant la deuxième composante, l'amplitude instantanée s'exprime

$$A_2(t) = a_2(t) \mathbb{1}_{[20, \infty[}(t) ,$$

où l'amplitude a_2 est construite de la même manière que A_1 (c'est-à-dire à partir d'une autre réalisation du même mouvement brownien lissé). La multiplication par l'indicatrice $\mathbb{1}_{[20, \infty[}$ conduit à deux régimes : lorsque $t < 20$ une seule composante est présente (la première), lorsque $t \geq 20$, les deux composantes sont présentes dans le signal y . Afin de conserver la validité du problème d'optimisation (5.23), on se limitera au cas où une seule composante possède une fréquence variant rapidement. Ceci permet de faire l'optimisation par rapport à la fonction à une variable α plutôt que par rapport à la fonction à deux variables Ω . On construit alors la fréquence instantanée ϕ'_2 telle que

$$\phi'_2(t) = \phi'_2 - \sin(t) ,$$

où ϕ'_2 est également construit à partir d'un mouvement brownien lissé puis normalisé et centré de sorte que $1 \leq \phi'_2(t) \leq 5$ pour tout t .

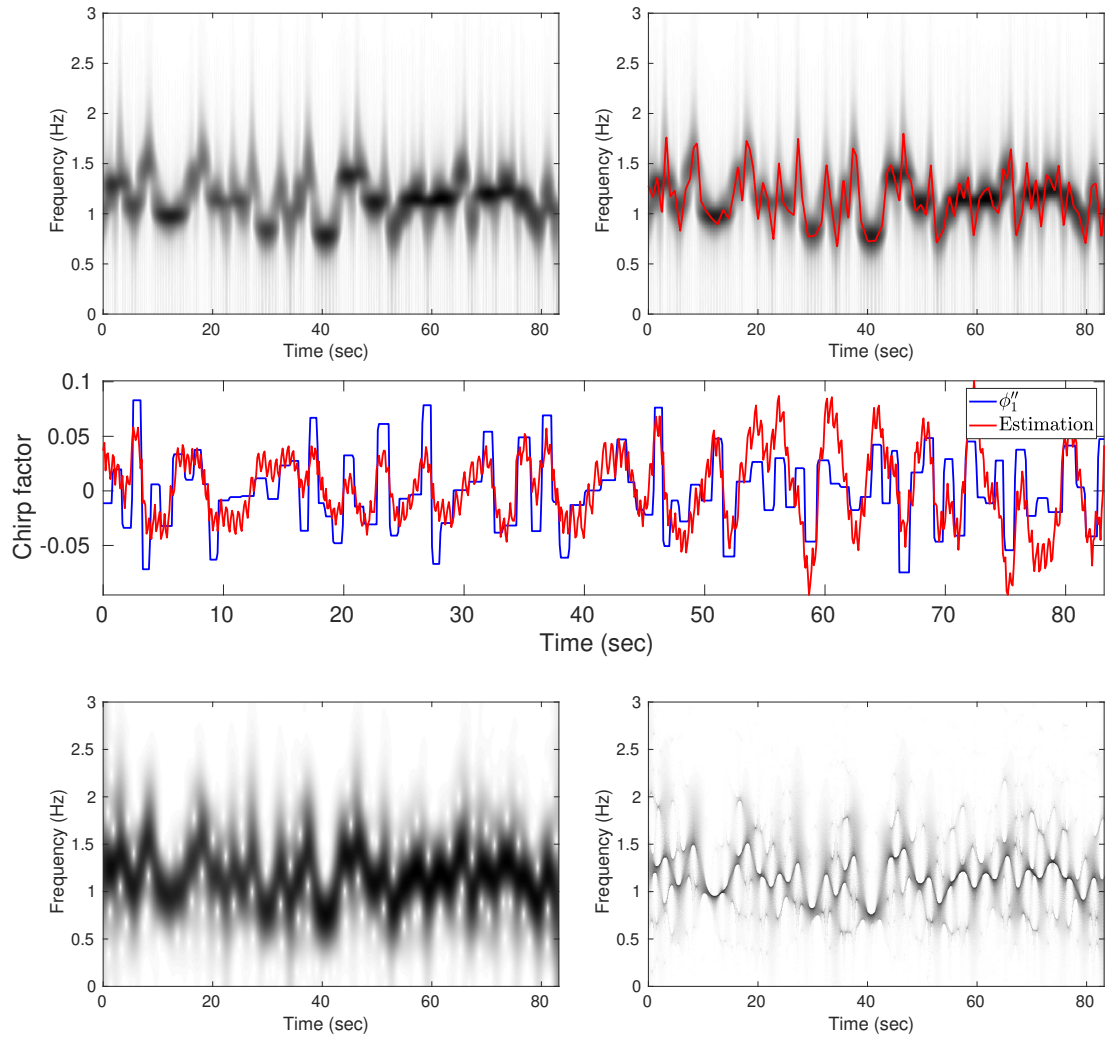


FIGURE 5.3 – Représentations temps fréquence du signal à une composante. En haut : résultat de TYCOON (à gauche), même résultat superposé à la fréquence instantanée (à droite). Au milieu : *chirp factor* estimé par TYCOON. En bas à gauche : TFCT. En bas à droite : synchrosqueezing de la TFCT.

5.5.2.1 Évaluation des performances

Afin de comparer les performances de TYCOON aux autres représentations temps-fréquence, il peut être utile de quantifier la qualité de l'estimation. C'est pourquoi nous allons établir un critère permettant de quantifier la distance du spectre instantané estimé par chacune de ces représentations temps-fréquence au spectre instantané idéal. Pour cela, nous allons raisonner par analogie avec la distance de transport optimal qui permet de quantifier la distance entre deux densités de probabilité. De manière générique, notons \mathcal{Q}_y la représentation temps-fréquence étudiée. Soit S_Y le spectre instantané idéal donnée par (5.8) et G_y la RTFi donnée par (5.9). Alors, pour t fixé, on considère les pseudo densités de probabilité suivantes :

$$p_y^t(\xi) = \frac{S_Y(\xi, t)}{\int_{\mathbb{R}} S_Y(\nu, t) d\nu}$$

$$\tilde{p}_y^t(\xi) = \frac{|\mathcal{Q}_y(\xi, t)|^2}{\int_{\mathbb{R}} |\mathcal{Q}_y(\nu, t)|^2 d\nu} ,$$

À chaque instant t , on peut alors déterminer la distance de transport optimal d_t entre ces deux pseudo densités. Elle est donnée par la norme 1 de la différence entre les fonctions de répartitions associées. Autrement dit, soit $P_y^t(\xi) = \int_{-\infty}^{\xi} p_y^t(\nu) d\nu$ et $\tilde{P}_y^t(\xi) = \int_{-\infty}^{\xi} \tilde{p}_y^t(\nu) d\nu$, on a

$$d_t(\mathcal{Q}_y, G_y) = \int_{\mathbb{R}} |\tilde{P}_y^t(\xi) - P_y^t(\xi)| d\xi .$$

Finalement, la distance entre les deux représentations temps-fréquence est obtenue en moyennant l'ensemble des distances de transport optimal par rapport au temps :

$$D(\mathcal{Q}_y, G_y) = 100 \times \frac{1}{|I|} \int_I d_t(\mathcal{Q}_y, G_y) dt . \quad (5.30)$$

Afin de justifier qualitativement le choix de cette distance, plaçons-nous à un instant t fixé. Supposons que le signal ne contient qu'une seule composante localement harmonique. Alors $p_Y^t(\xi) = \delta(\xi - \phi'(t))$. Sa primitive, P_Y^t est donc une fonction échelon : $P_Y^t = \mathbb{1}_{[\phi'(t), \infty[}(\xi)$. La figure 5.4 illustre la distance de transport optimal entre cette RTFi (en bleu sur la figure) et deux représentations temps-fréquences fictives. Pour la première (en rouge sur la figure), la pseudo densité \tilde{p}_1^t est également formée d'une impulsion de Dirac, mais elle est décalée d'une fréquence notée Δ_ξ par rapport à la vraie valeur $\phi'(t)$. Pour la seconde (en magenta sur la figure), la pseudo densité est composée d'une gaussienne centrée sur la bonne fréquence $\phi'(t)$. À droite de cette figure, on trace les fonctions $|\tilde{P}_j^t - P_Y^t|$ où $j = 1$ ou 2 . Dans le premier cas, cette fonction est non nulle uniquement sur la plage de fréquences située entre $\phi'(t)$ et $\phi'(t) + \Delta_\xi$. Pour le deuxième exemple, cette fonction est lentement décroissante autour de $\phi'(t)$. La distance d_t mesure la norme 1 de cette fonction, elle quantifie donc la parcimonie de celle-ci. C'est pourquoi cette distance est plus faible dans le premier exemple que dans

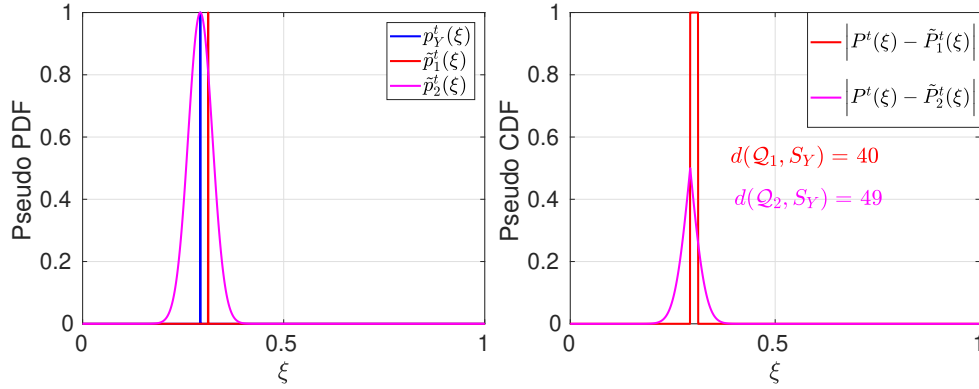


FIGURE 5.4 – Exemple illustration de la distance de transport optimal. À gauche : exemples de pseudo densités de probabilité normalisée. À droite : différences en valeur absolue entre les pseudos fonctions de répartition estimées et réelle.

le second. Ainsi, la distance de transport optimal permet de quantifier la proximité entre les fréquences instantanées estimées et réelles tout en favorisant la parcimonie de la représentation temps-fréquence estimée. Une étude plus détaillée de cette distance pourra être trouvée dans [75].

On calcule alors la distance D entre le spectre instantané idéal et les différentes représentations temps-fréquence calculées. Chacune des méthode est mise en œuvre pour plusieurs valeurs des paramètres dont elle dépend. Comme nous l'avons présenté au chapitre 4, l'algorithme JEFAS-S ne permet pas de construire la représentation temps-fréquence idéale de signaux localement harmoniques comportant plus d'une composante, il ne sera donc pas mis en œuvre ici. Néanmoins, la représentation donnée par TYCOON est comparée aux méthodes suivantes :

- la TFCT pour une fenêtre gaussienne g_σ de la forme :

$$g_\sigma(t) = e^{-\frac{1}{2\sigma^2}t^2}; \quad (5.31)$$

- le synchrosqueezing associé à la TFCT;
- le synchrosqueezing associé à la transformée en ondelettes pour une ondelette ψ_d dont la transformée de Fourier est donnée par :

$$\hat{\psi}_d(\xi) = \exp\left(\frac{1}{\left(\frac{\xi - \xi_0}{d}\right)^2 - 1}\right) \mathbb{1}_{[\xi_0 - d, \xi_0 + d]}(\xi), \quad (5.32)$$

où ξ_0 désigne la fréquence centrale de l'ondelette, il s'agit d'une ondelette dont on peut facilement contrôler la bande passante à l'aide du paramètre d car elle est d'autant plus concentrée en fréquence que d est petit;

- la décomposition modale empirique avec K_{EMD} composantes.

TABLEAU 5.1 – Signal à deux composantes. Distance des spectres instantanés estimés au spectre instantané idéal.

Méthode	Paramètre associé	Distance D à la RTFi	
		Moyenne	Écart-type
TYCOON	$\lambda = 0,95$	7,18	0,90
	$\lambda = 0,98$	4,80	0,30
	$\lambda = 0,999$	7,16	0,48
TFCT	$\sigma = 0,125$	5,60	0,38
	$\sigma = 0,25$	4,49	0,18
	$\sigma = 0,5$	6,16	0,20
	$\sigma = 1$	10,82	0,30
SST-TFCT	$\sigma = 0,125$	8,88	1,02
	$\sigma = 0,25$	8,90	1,50
	$\sigma = 0,5$	7,34	1,51
	$\sigma = 1$	8,12	0,73
SST-CWT	$d = \xi_0/10$	7,69	0,68
	$d = \xi_0/5$	6,42	0,68
	$d = 2\xi_0/5$	5,04	0,84
	$d = 3\xi_0/5$	5,47	1,27
EMD-HT	$K_{EMD} = 1$	15,40	2,55
	$K_{EMD} = 2$	5,65	0,96
	$K_{EMD} = 4$	5,56	0,85
	$K_{EMD} = 6$	5,69	0,90

TYCOON est mis en œuvre pour trois valeurs différentes de l’hyperparamètre λ . Les valeurs de l’hyperparamètre $\tilde{\mu}$ sont fixées comme dans l’exemple précédent, c’est-à-dire que les valeurs successives de $\tilde{\mu}$ sont celles contenues dans le vecteur $I_{\tilde{\mu}}$ tel que $I_{\tilde{\mu}} = (1, 10^{-1}, \dots, 10^{-8})$. La TFCT et le synchrosqueezing associé sont testés pour quatre valeurs différentes de la largeur σ de la fenêtre d’analyse g_σ donnée par (5.31). Le méthode SST-CWT est évaluée pour quatre largeurs différentes d de l’ondelette ψ_d donnée par (5.32). Enfin, la décomposition modale empirique est effectuée pour quatre valeurs différentes de K_{EMD} . On effectue alors le calcul de la distance de transport optimal de chacune des représentations temps-fréquence calculées. Cette mesure de distance est répétée pour 100 réalisations du signal y . À chaque réalisation, les amplitudes instantanées A_1 et A_2 et la fréquence instantanée ϕ'_1 sont modifiées car elles proviennent de nouvelles réalisations de mouvements browniens lissés. Les moyennes et écarts-types des distances mesurées sont rassemblés dans le tableau 5.1. Pour la valeur optimale $\lambda = 0,98$, les résultats montrent que représentation donnée par TYCOON est aussi proche de la RTFi au sens de la distance D que la plupart autres représentations quelles que soient les valeurs des paramètres choisies. Seule la représentation donnée par le synchrosqueezing de la TFCT ne permet pas d’obtenir une représentation de même

qualité. Il est à noter que chacune des méthodes possède une valeur optimale du paramètre dont elle dépend permettant de minimiser la distance à la RTFi.

TYCOON. Comme on l'a vu dans la section 5.5.1, le paramètre λ règle la régularité temporelle et le parcimonie de la représentation temps-fréquence obtenue par TYCOON. Lorsque λ est trop grand, la représentation temps-fréquence devient trop régulière tandis que lorsque λ est trop petit. On renvoie le lecteur à la figure 5.2 pour observer l'effet de λ sur le résultat de TYCOON.

TFCT. Une fenêtre trop large ne permet pas de suivre l'évolution rapide des fréquences instantanées car des « bulles » dues aux interférences d'une composante localement harmoniques avec elle-même apparaissent sur le spectrogramme. À l'inverse, une fenêtre d'analyse trop étroite conduit à élargir la bande fréquentielle suivant chacune des composantes. La concentration des coefficients significatifs de la TFCT le long des fréquences instantanées est donc fortement réduite.

SST-TFCT et SST-CWT. Les remarques sont les mêmes que pour la TFCT. Lorsque que la fenêtre ou l'ondelette d'analyse est trop large, les bulles présentes sur le spectrogramme ou le scalogramme conduisent à l'estimation par le synchro-queezing d'une fréquence instantanée fictive passant par ces bulles.

EMD-HT. Lorsque le nombre de composantes estimées est inférieur au nombre réels de composantes, l'algorithme EMD-HT n'estime qu'une partie des composantes constituant le signal ce qui rend la représentation temps-fréquence estimée trop parcimonieuse. Un choix de K_{EMD} supérieur au nombre réel de composantes ne semble pas détériorer significativement la qualité de la représentation temps-fréquence. Cette remarque reste vraie tant que le signal analysé n'est pas bruité, la présence de bruit pouvant amener à l'estimation de composantes « fantômes » à la place des composantes réelles (voir le paragraphe 5.5.2.2).

Parmi les combinaisons de paramètres testées, on conserve les valeurs permettant d'obtenir une distance au spectre instantané idéal minimale pour chacune des méthodes mise en œuvre. Les représentations temps-fréquence correspondantes sont alors tracées en figure 5.5. Cet exemple illustre clairement le fait que TYCOON permet de suivre les variations fréquentielles de la composante dont la fréquence varie rapidement, même lorsque qu'une seconde composante est présente. En revanche, les autres représentations temps-fréquence laissent assez bien apparaître la deuxième composante dont la fréquence varie lentement mais ne permettent pas d'isoler nettement la première composante dont la fréquence varie rapidement. Par contre, nous ne traçons pas ici l'estimation du *chirp factor* obtenu par TYCOON car la signification à apporter à ce terme dans cette situation reste floue. Dès que la deuxième composante apparaît (c'est-à-dire dès que $t \geq 20$ s), le *chirp factor* estimé ne suit plus la fonction ϕ_1'' bien que la deuxième composante vérifie $\phi_2'' < \epsilon\phi_2'$. Une optimisation par rapport à Ω serait plus facilement interprétable ici, bien que plus délicate à obtenir numériquement.

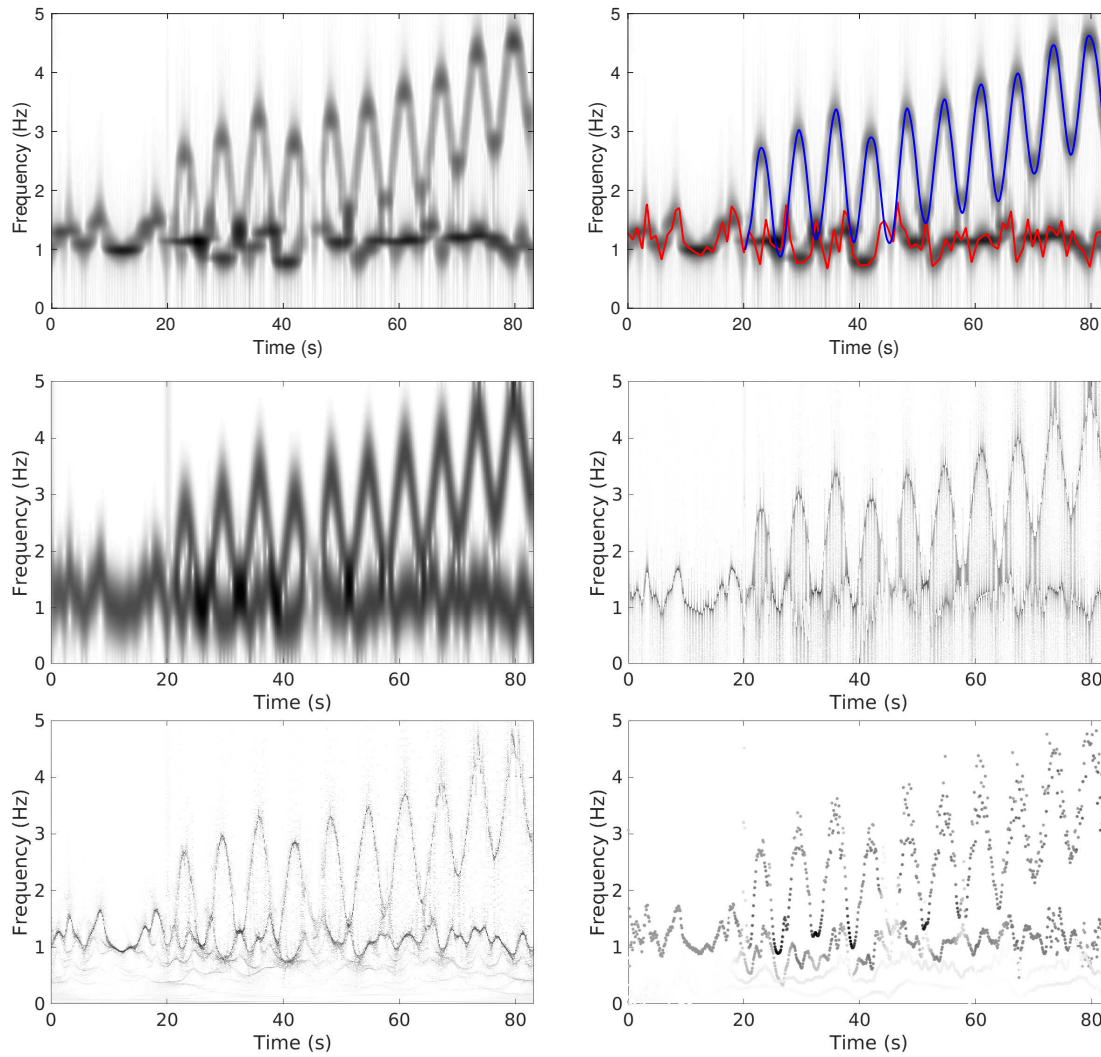


FIGURE 5.5 – Représentations temps fréquence du signal à deux composantes. En haut : résultat de TYCOON (à gauche), même résultats superposé aux fréquences instantanées (à droite). Au milieu à gauche : TFCT. Au milieu à droite : synchrosqueezing de la TFCT. En bas à gauche : synchrosqueezing de la transformée en ondelettes. En bas à droite : décomposition modale empirique.

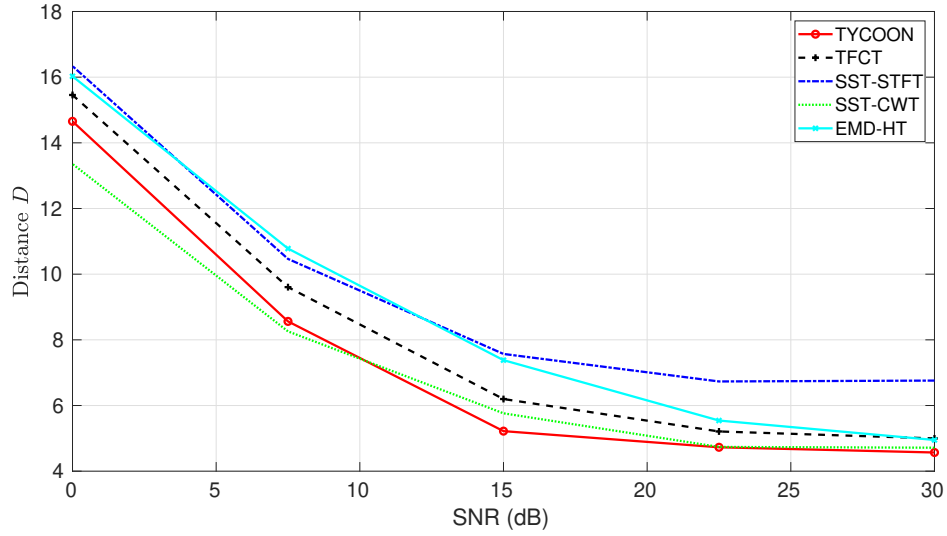


FIGURE 5.6 – Signal à deux composantes. Évolution de la distance du spectre instantané estimé par différentes méthodes en fonction du rapport signal sur bruit.

5.5.2.2 Robustesse au bruit

On ajoute maintenant du bruit au signal à deux composantes défini en (5.29). On considère donc le signal y_b donné par

$$y_b(t) = y(t) + e(t) ,$$

où y est le signal défini par l'équation (5.29) et e est un bruit blanc gaussien dont la variance est choisie en fonction du rapport signal sur bruit. Soit SNR le rapport signal sur bruit exprimé en décibels et défini par :

$$\text{SNR} = 10 \log_{10} \left(\frac{\text{Var}(y_b)}{\text{Var}(e)} \right) .$$

On applique alors les différentes méthodes au signal bruité pour différentes valeurs du rapport signal sur bruit. Les paramètres dont dépend chacune de ces méthodes sont choisis comme étant ceux dont les performances sont les meilleures parmi celles testées dans le paragraphe précédent (voir le tableau 5.1). On trace en figure 5.6 l'évolution de la distance D en fonction du SNR pour chacune de ces méthodes. Ce graphique montre que TYCOON reste performant tant que le niveau de bruit reste raisonnable. En effet, tant que le SNR reste supérieur à 5 dB environ, la distance à la RTFi de la représentation temps-fréquence issue de TYCOON est du même ordre que celle de la meilleure représentation donnée par une autre méthode (la SST-CWT dans ce cas). Notons également que la décomposition modale empirique est très sensible à la présence de bruit, la qualité de la représentation temps-fréquence obtenue s'altère rapidement dès que le rapport signal sur bruit diminue.

5.6 Bilan

Dans ce chapitre, nous avons proposé une approche à la synthèse afin d’approcher la représentation temps-fréquence instantanée idéale de signaux localement harmoniques dont les fréquences instantanées varient rapidement. Cette approche est motivée par le souhait de contourner les artefacts produits par les méthodes d’analyse spectrale des signaux de cette forme. L’algorithme TYCOON que nous avons construit présente différents avantages sur ces méthodes.

- La représentation temps-fréquence déduite de TYCOON permet de conduire à plus de régularité dans le suivi des fréquences instantanées qu’une méthode d’analyse comme le synchrosqueezing.
- La concentration des coefficients temps-fréquence significatifs le long des fréquences instantanées est supérieure à des méthodes d’analyse comme la TFCT ou la transformée en ondelettes.
- la méthode d’optimisation alternée que nous avons proposée permet d’estimer conjointement la RTFi et le *chirp factor*. TYCOON donne donc accès à une grandeur supplémentaire par rapport aux autres méthodes : la dérivée de la fréquence instantanée.

Néanmoins, TYCOON reste peu performant sur certains points :

- La concentration des coefficients significatifs dans le plan temps-fréquence reste faible comparativement à des méthodes comme le synchrosqueezing ou l’EMD. En particulier, lorsque les fréquences instantanées varient lentement, ces méthodes restent efficaces et peuvent surpasser TYCOON.
- L’influence du jeu d’hyperparamètres sur l’allure de la représentation temps-fréquence obtenue par TYCOON est bien comprise, mais l’ajustement conduisant au compromis entre la parcimonie temps-fréquence et la régularité temporelle peut être délicat.
- TYCOON repose sur un algorithme d’optimisation convexe. C’est pourquoi le temps de calcul associé est bien supérieur à celui de méthodes directes telles que le synchrosqueezing.

Nous renvoyons le lecteur à la section 6.2 où nous proposons des perspectives de recherches futures pour atténuer ces effets négatifs, notamment le manque de concentration temps-fréquence.

En outre, cette approche est complémentaire de celle proposée dans le chapitre 4 car elle permet d’étudier un autre type de signaux : les signaux localement harmoniques. Lorsqu’on a affaire à des signaux pouvant être modélisés par les deux classes définies, c’est-à-dire localement déformés et localement harmoniques, l’approche probabiliste permet un meilleur contrôle des propriétés recherchées pour la représentation temps-fréquence idéale. JEFAS-S donne donc une représentation plus concentrée que TYCOON lorsque les deux algorithmes sont applicables (signaux à une composante). Cependant, TYCOON est applicable à des signaux plus « compliqués » car ils peuvent posséder plusieurs composantes. La complexité de ces signaux vient du fait qu’ils dépendent de $2K$ fonctions : K amplitudes et K fréquences instantanées, contre seulement

deux fonctions pour les signaux localement dilatés : le spectre et la dilation instantanée.

En résumé, TYCOON est une méthode intéressante dans les situations où le signal comporte aux moins deux composantes localement sinusoïdales et qu'au moins une de ces composantes possède une fréquence instantanée variant rapidement.

CONCLUSION ET PERSPECTIVES

6.1 Conclusion

Dans cette thèse, nous avons étudié des signaux non stationnaires présentant une structure particulière : il s'agit d'un modèle de non-stationnarité considérant que ces signaux proviennent de signaux stationnaires dont la stationnarité a été brisée par un opérateur de déformation. Ce modèle offre alors deux degrés de libertés distincts. Le premier concerne le choix de la densité spectrale de puissance du signal stationnaire sous-jacent, le second concerne le choix de l'opérateur de déformation brisant la stationnarité. En jouant sur ces degrés de liberté, nous avons analysé en détail deux modèles de signaux. D'une part, le modèle de signal localement déformé considère un spectre du signal stationnaire sous-jacent quelconque mais se limite à des opérateurs de déformations paramétrés par une fonction (dilatation instantanée, modulation d'amplitude et de fréquence). D'autre part, le modèle de signal localement harmonique considère uniquement un spectre de raies pour le signal stationnaire sous-jacent mais offre une plus large gamme d'opérateurs brisant la stationnarité en permettant de moduler en fréquence chacune des raies composant ce spectre par une fréquence et une amplitude instantanées distinctes.

La contribution principale de cette thèse a été de définir, analyser et mettre en œuvre des méthodes d'analyse et de synthèse spectrale appropriées à l'étude de ces signaux. En particulier les chapitres 2, 3 et 4 se sont concentrés sur l'étude des signaux localement déformés tandis que le chapitre 5 s'est focalisé sur l'étude des signaux localement harmoniques.

Concernant les signaux localement déformés, nous avons vu que déterminer le spectre instantané de ce type de signal revenait à connaître simultanément le spectre du signal stationnaire sous-jacent et les fonctions associées aux opérateurs de déformations. Les méthodes d'estimation développées dans les chapitres 2, 3 et 4 ont consisté à déterminer ces quantités pour différents modèles de déformations. Dans le chapitre 2, nous avons développé une méthode d'analyse des signaux localement dilatés et modulés en amplitude. Dans le chapitre 3, nous avons étendu cette méthode d'analyse

au cas des signaux multi-capteurs et pour un opérateur de modulation d'amplitude matricielle. Nous avons vu que le problème ainsi posé revenait à développer une méthode de séparation de source doublement non stationnaire. Dans le chapitre 4, nous avons étudié les signaux mono-capteur localement dilatés. En revanche, contrairement à l'approche des deux chapitres précédents, nous avons proposé dans ce chapitre une méthode de synthèse spectrale.

Concernant les signaux localement harmoniques, la détermination du spectre instantané revient à connaître, à chaque instant, les valeurs prises par chacune des amplitudes et fréquences instantanées composant ces signaux. C'est pourquoi nous avons développé dans le chapitre 5 une méthode de synthèse spectrale adaptée à ces signaux.

Finalement, la prise en compte du modèle de « stationnarité brisée » a été caractérisée dans ces différents chapitres à partir d'un déplacement de coefficients dans le plan temps-échelle ou le plan temps-fréquence. En effet, dans les chapitres 2 à 4 la non-stationnarité des signaux localement dilatés a été caractérisée par une translation des coefficients d'ondelettes ou de la représentation temps-échelle synthétisée. De même, dans le chapitre 5 la régularité de la représentation temps-fréquence obtenue a été caractérisée par une équation de transport décrivant le déplacement des coefficients temps-fréquence en fonction des variations de la fréquence instantanée. C'est cette propriété commune qui nous a permis de développer et mettre en œuvre, dans chacune des configurations considérées, un algorithme d'estimation correspondant, que nous avons testé sur des signaux synthétiques et réels.

Ainsi, l'algorithme JEFAS nous a permis de déterminer le spectre instantané de divers signaux audios comme le son produit par une voiture en déplacement et déformé par effet Doppler. JEFAS-S nous permet d'étudier le spectre instantané d'une plus grande gamme de signaux localement dilatés : ceux pour lesquels la déformation temporelle varie rapidement, comme les signaux cardiaques dont la fréquence instantanée varie rapidement. JEFAS-BSS nous a permis de profiter de l'estimation produite par JEFAS pour améliorer la qualité de la séparation de sources produite par les méthodes de référence. Enfin, TYCOON fournit une estimation de la représentation temps-fréquence idéale des signaux localement harmoniques à fréquences rapides, alors que les méthodes d'analyse classiques, comme le synchrosqueezing, sont défaillantes pour étudier ce type de signaux.

6.2 Perspectives

Finalement, ce manuscrit nous a permis de proposer des méthodes assez diverses pour étudier les signaux localement déformés et localement harmoniques. Chacun des chapitres étudiés donne lieu à diverses questions, et différentes perspectives de recherches sur ces sujets peuvent être envisagées. Nous en proposons ici quelques unes.

Chapitre 2. Ce chapitre s'est focalisé sur un cas particulier de signal localement déformé : celui des signaux localement dilatés et modulés en amplitude. Comme nous l'avons mentionné dans la section 2.6, d'autres modèles de déformations peuvent être

intéressants pour décrire des non-stationnarités plus complexes. Nous avons considéré dans [45] le cas des signaux localement dilatés et modulés en fréquence. Nous avons alors proposé d’analyser ces signaux au moyen d’une analyse temps-fréquence-échelle. Toutefois, l’utilisation de cet outil d’analyse reste délicate numériquement. En particulier le choix de la discrétisation doit être étudié afin de contrôler le redondance des coefficients d’analyse. C’est pourquoi nous n’avons mis en œuvre l’étude de ce modèle que sur un cas particulier où la modulation de fréquence dépend de la dilatation instantanée. Par ailleurs, nous n’avons pas rencontré de signaux réels (audio ou non) pour lesquels ce modèle plus complexe serait adéquat.

En revanche, l’étude d’un autre modèle de déformation nous semble plus pertinente. En traitement de la parole, le modèle classique de production de sons voisés suppose que quelle que soit la hauteur du son produit, celui-ci passe par un filtre caractéristique du conduit vocal et de la voyelle produite. Ce modèle est appelé *source-filtre* (voir [30] pour plus de détails). Dans notre formalisme, cela reviendrait à considérer un filtrage postérieur à la dilatation instantanée et à la modulation d’amplitude. En reprenant les notations du chapitre 2, cela reviendrait à approcher la transformée en ondelettes du signal non stationnaire y comme suit :

$$\mathcal{W}_y(s, \tau) \approx \hat{h}(q^{-s}v_0)\alpha(\tau)\mathcal{W}_x\left(s + \log_q(\gamma'(\tau)), \gamma(\tau)\right), \quad (6.1)$$

où \hat{h} désigne la réponse fréquentielle du filtre caractérisant ce nouveau modèle de non-stationnarité. Il s’agirait alors d’estimer un nouveau coefficient multiplicatif dans le plan temps-échelle. L’étude de ce modèle serait alors pertinente pour une meilleure analyse de signaux vocaux. D’ailleurs, nous proposons en ligne l’analyse par JEFAS d’un son de vocalises produit une chanteuse (voir annexe B.1). Ceci met en évidence le fait que JEFAS estime la principale source de non-stationnarité venant de ce signal. Mais, nous pouvons toujours percevoir sur la transformée en ondelettes du signal stationnarisé quelques variations d’amplitude entre les harmoniques qui pourrait être estimées en considérant le modèle (6.1).

Chapitre 3. Dans ce chapitre, nous nous sommes bornés à l’étude de JEFAS-BSS sur des mélanges synthétiques. Il serait intéressant d’évaluer les performances de JEFAS-BSS sur des mélanges réels. Or, nous n’avons pas trouvé de signaux réels sur lesquels JEFAS-BSS serait applicable. La bio-acoustique pour l’étude des sons d’animaux en multi-capteurs pourrait être un domaine d’application de JEFAS-BSS. Par exemple, le hurlement de loup est un son non stationnaire dont l’analyse par JEFAS est possible car le modèle de dilatation instantané est pertinent pour ce type de signal (voir la figure 2.4). Or, lorsque dans une meute de loups se déplaçant, plusieurs loups hurlent simultanément le modèle de ce chapitre pourrait être adapté pour isoler le cri de chacun des loups. Le matrice de mélange variant dans le temps modéliserait alors les variations d’amplitude dues aux déplacements des loups. En éco-acoustique, ce type de problème a d’ailleurs été traité pour estimer la taille d’une population de loups [54].

En outre, on pourrait également envisager de complexifier le modèle en modifiant l’opérateur de déformation comme nous l’avons proposé pour le chapitre précédent

(ajout d'un filtrage par exemple). Mais, l'étude du modèle où ce n'est plus l'opération de modulation d'amplitude instantanée qui devient matricielle mais celle de déformation temporelle nous paraît la plus intéressante. Par exemple, négligeons la modulation d'amplitude, et supposons que les nombres de sources et d'observations sont égaux à 2, alors le modèle décrivant chacun des deux signaux observés z_k serait le suivant :

$$z_k(t) = \mathcal{D}_{\gamma_{k1}} x_1(t) + \mathcal{D}_{\gamma_{k2}} x_2(t) .$$

il s'agirait alors d'estimer quatre fonctions de dilatation instantanées et deux spectres des signaux stationnaires sous-jacents. Ce modèle pourrait, par exemple, être particulièrement approprié à la situation où nous possédons deux enregistrements pris depuis deux points différents de deux voitures se déplaçant à vitesse constante. Ainsi, la fonction de déformation due à l'effet Doppler dépend de la position des observateurs, ce que ce modèle décrit bien.

Chapitre 4. Dans ce chapitre, nous sommes limités à l'étude des signaux localement dilatés. Afin d'obtenir le pendant exact de JEFAS pour une approche à la synthèse, il nous faudra considérer l'ajout de la modulation d'amplitude au modèle de non-stationnarité, et ajuster l'algorithme JEFAS-S en conséquent. Comme pour JEFAS, on peut également envisager d'étendre JEFAS-S aux signaux localement déformés puis filtrés, l'*a priori* sur la représentation temps-échelle recherchée s'écrit alors sous la même forme que dans l'équation (6.1). Plus généralement, l'approche à la synthèse conduisant à exprimer la déformation dans le domaine temps-échelle, nous pouvons considérer de nouveaux modèles de déformation s'exprimant directement dans le domaine transformé. Par exemple, on peut considérer le modèle *a priori* tel que les coefficients temps-échelle W_y du signal non stationnaire y s'expriment en fonctions des coefficients W_x du signal stationnaire x selon :

$$W_y(s, \tau) = m(s, \tau) W_x(s, \tau) ,$$

où m désigne un multiplicateur temps-échelle inconnu.

On peut également envisager des modifications plus mineures de JEFAS-S. En particulier, le modèle de bruit additif que nous avons choisi est celui du bruit blanc. C'est un choix qui a été fait par souci de simplicité mais il peut très bien être modifié. Nous pourrions alors considérer le cas d'un bruit coloré, présent principalement aux hautes fréquences par exemple, afin de forcer JEFAS-S à estimer la dilatation instantanée à partir des échelles où le signal est le moins bruité. Cela permettrait donc d'exclure des bandes de fréquences pour lesquelles nous considérons que le modèle de dilatation n'est pas approprié pour décrire le signal, et de privilégier les bandes de fréquences où le modèle est convenable.

En outre, comme JEFAS-BSS apparaît comme une extension naturelle de JEFAS à la séparation de sources, il nous semble intéressant d'essayer d'étendre JEFAS-S à la séparation de sources. Bien que nous soyons pessimiste quant à l'utilisation en pratique d'une tel algorithme compte tenu du temps de calcul demandé par JEFAS-S, cette extension de JEFAS-S aurait pour intérêt de fournir une méthode d'estimation à la synthèse qui serait le pendant de JEFAS-BSS.

Chapitre 5. Dans ce chapitre, nous avons mis en œuvre l'algorithme TYCOON qui permet d'approcher la représentation temps-fréquence idéale des signaux localement harmoniques. Cependant, même dans la situation où nous avons affaire à un signal à une composante, il paraît clair que TYCOON ne permet pas d'obtenir une version véritablement concentrée de la représentation temps-fréquence. Il nous semble alors intéressant de comprendre pourquoi cette incertitude reste présente et si nous avons la possibilité de la réduire. Une possibilité provient du fait que le terme de régularisation par la norme L^1 favorise une parcimonie dans le plan temps-fréquence. Ce terme serait donc plutôt adapté aux cas où la représentation recherchée serait constituée de quelques points localisées dans le plan temps-fréquence, et non de lignes comme l'est la RTFi où chaque ligne caractérise une fréquence instantanée. Une possibilité envisageable pour améliorer la concentration de la représentation temps-fréquence serait de remplacer le terme de régularisation sur la représentation recherchée. On peut alors imaginer placer une norme mixte [35] qui contiendrait un terme de régularisation L^1 par rapport à la variable fréquentielle et un terme de régularisation L^2 par rapport à la variable temporelle. Cela aurait pour but de forcer la parcimonie uniquement le long de l'axe fréquentiel, et ainsi ne pas pénaliser l'apparition de lignes suivant l'axe temporel.

En outre, nous avons mis en œuvre TYCOON en remplaçant la fonction du plan temps-fréquence Ω devant contenir l'ensemble des *chirp factors* (c'est-à-dire les dérivées des fréquences instantanées) par la fonction α dépendant uniquement du temps. Cette dernière est censée estimer le *chirp factor* lorsqu'on étudie un signal localement harmonique à fréquence rapide ne comportant qu'une seule composante. L'adaptation de TYCOON au cas plus général, adapté aux signaux multi-composantes, est une suite logique de ce travail. Pour réaliser cette modification, nous devrions en tenir compte en changeant la régularisation sur cette fonction. En effet, à l'image de la représentation temps-fréquence idéale, les valeurs significatives de la fonction Ω doivent être concentrées le long des fréquences instantanées. Un terme de régularisation L^1 (qui force la parcimonie) serait donc plus approprié que le terme de régularisation L^2 actuellement placé sur la fonction α pour des questions numériques. Néanmoins, cela conduirait à un algorithme de régularisation plus complexe car nous devrions alors effectuer une descente proximale pour estimer Ω (en plus de la RTFi) alors que l'estimation de α à chaque étape de l'algorithme alterné est actuellement donnée par une expression explicite.

Annexe A

DÉMONSTRATIONS

Nous donnons ici les démonstrations des théorèmes et propositions présentés dans ce manuscrit.

A.1 Démonstrations du chapitre 2

Plusieurs des démonstrations présentées font appel au lemme suivant.

Lemme A.1. *Soit X un processus stochastique réel, stationnaire au sens large de spectre \mathcal{S}_X et de moyenne nulle. Soient f et g des fonctions telles que les variables aléatoires $\langle X, f \rangle$ et $\langle X, g \rangle$ soient bien définies et bornées. Alors*

$$\begin{cases} \mathbb{E} \{ \langle X, f \rangle \overline{\langle X, g \rangle} \} &= \langle \mathcal{S}_X, \hat{f} \bar{\hat{g}} \rangle \\ \mathbb{E} \{ \langle X, f \rangle \langle X, g \rangle \} &= \langle \mathcal{S}_X, \hat{f} \hat{g} \rangle, \end{cases}$$

où $\hat{g}(\xi) = \hat{g}(-\xi)$.

La démonstration de ces égalités s'obtient directement en développant l'expression à gauche de l'équation. En reconnaissant l'expression de la fonction d'autocovariance et en appliquant la formule de Plancherel, on obtient l'expression à droite de l'équation.

A.1.1 Démonstration du théorème d'approximation 2.1

Nous fournissons ici plus de détails sur l'approximation de la transformée en ondelettes des signaux non stationnaires localement dilatés et modulés en amplitude.

Sous l'hypothèse de régularité de α et γ , les transformées en ondelettes du processus non stationnaire Y et du processus stationnaire sous-jacent X sont approximativement reliées comme suit :

$$\mathcal{W}_Y(s, \tau) \approx \widetilde{\mathcal{W}}_Y(s, \tau) = \alpha(\tau) \mathcal{W}_X \left(s + \log_q(\gamma'(\tau)), \gamma(\tau) \right).$$

Le théorème 2.1, que nous allons démontrer fournit une évaluation quantitative de la qualité de l'approximation.

Démonstration du théorème 2.1. Simplifions les notations. Soit \mathcal{B}_γ l'opérateur de déformation temporelle non normalisé et $\widetilde{\mathcal{B}}_\gamma^\tau$ l'opérateur tangent associé, c'est-à-dire :

$$\begin{aligned}\mathcal{B}_\gamma &= \frac{1}{\sqrt{\gamma'}} \mathcal{D}_\gamma, \\ \widetilde{\mathcal{B}}_\gamma^\tau &= \frac{1}{\sqrt{\gamma'(\tau)}} \widetilde{\mathcal{D}}_\gamma^\tau.\end{aligned}$$

Nous décomposons l'erreur d'approximation en trois termes de sorte que :

$$\begin{aligned}\varepsilon(s, \tau) &= \langle \mathcal{A}_\alpha \mathcal{D}_\gamma X, \psi_{s\tau} \rangle - \langle \widetilde{\mathcal{A}}_\alpha^\tau \widetilde{\mathcal{D}}_\gamma^\tau X, \psi_{s\tau} \rangle \\ &= \widetilde{\mathcal{A}}_\alpha^\tau \left(\langle \widetilde{\mathcal{A}}_\alpha^{\tau^{-1}} \mathcal{A}_\alpha \mathcal{D}_\gamma X, \psi_{s\tau} \rangle - \langle \widetilde{\mathcal{D}}_\gamma^\tau X, \psi_{s\tau} \rangle \right) \\ &= \alpha(\tau) \left(\left\langle \sqrt{\gamma'} \mathcal{B}_\gamma X, \psi_{s\tau} \right\rangle - \left\langle \sqrt{\gamma'(\tau)} \widetilde{\mathcal{B}}_\gamma^\tau X, \psi_{s\tau} \right\rangle \right. \\ &\quad \left. + \left\langle \left(\widetilde{\mathcal{A}}_\alpha^{\tau^{-1}} \mathcal{A}_\alpha - 1 \right) \mathcal{D}_\gamma X, \psi_{s\tau} \right\rangle \right) \\ &= \alpha(\tau) \left(\left\langle \left(\sqrt{\gamma'} - \sqrt{\gamma'(\tau)} \right) \mathcal{B}_\gamma X, \psi_{s\tau} \right\rangle + \sqrt{\gamma'(\tau)} \left\langle \left(\mathcal{B}_\gamma - \widetilde{\mathcal{B}}_\gamma^\tau \right) X, \psi_{s\tau} \right\rangle \right. \\ &\quad \left. + \left\langle \left(\widetilde{\mathcal{A}}_\alpha^{\tau^{-1}} \mathcal{A}_\alpha - 1 \right) \mathcal{D}_\gamma X, \psi_{s\tau} \right\rangle \right) \\ &= \alpha(\tau) \left(\varepsilon^{(1)}(s, \tau) + \sqrt{\gamma'(\tau)} \varepsilon^{(2)}(s, \tau) + \varepsilon^{(3)}(s, \tau) \right),\end{aligned}$$

où

$$\begin{aligned}\varepsilon^{(1)}(s, \tau) &\triangleq \left\langle \left(\sqrt{\gamma'} - \sqrt{\gamma'(\tau)} \right) \mathcal{B}_\gamma X, \psi_{s\tau} \right\rangle = \left\langle X, \mathcal{B}_{\gamma^{-1}} \left(\sqrt{\gamma'} - \sqrt{\gamma'(\tau)} \right) \psi_{s\tau} \right\rangle, \\ \varepsilon^{(2)}(s, \tau) &\triangleq \left\langle \left(\mathcal{B}_\gamma - \widetilde{\mathcal{B}}_\gamma^\tau \right) X, \psi_{s\tau} \right\rangle = \left\langle X, \left(\mathcal{B}_{\gamma^{-1}} - \widetilde{\mathcal{B}}_\gamma^{\tau^{-1}} \right) \psi_{s\tau} \right\rangle, \\ \varepsilon^{(3)}(s, \tau) &\triangleq \left\langle \left(\mathcal{A}_\alpha \widetilde{\mathcal{A}}_\alpha^{\tau^{-1}} - 1 \right) \mathcal{D}_\gamma X, \psi_{s\tau} \right\rangle = \left\langle X, \mathcal{D}_{\gamma^{-1}} \left(\mathcal{A}_\alpha \widetilde{\mathcal{A}}_\alpha^{\tau^{-1}} - 1 \right) \psi_{s\tau} \right\rangle.\end{aligned}$$

Dans la suite, le terme de droite dans le produit scalaire définissant $\varepsilon^{(k)}(s, \tau)$ sera noté $f_{s\tau}^{(k)}$ de sorte que $\varepsilon^{(k)}(s, \tau) = \langle X, f_{s\tau}^{(k)} \rangle$ (avec $k \in \{1, 2, 3\}$).

En outre, l'inégalité triangulaire donne :

$$\begin{aligned}\mathbb{E} \left\{ |\varepsilon(s, \tau)|^2 \right\} &\leq C_\alpha^2 \left(\sqrt{\mathbb{E} \left\{ |\varepsilon^{(1)}(s, \tau)|^2 \right\}} + \sqrt{C_\gamma \mathbb{E} \left\{ |\varepsilon^{(2)}(s, \tau)|^2 \right\}} \right. \\ &\quad \left. + \sqrt{\mathbb{E} \left\{ |\varepsilon^{(3)}(s, \tau)|^2 \right\}} \right)^2.\end{aligned}\tag{A.1}$$

Déterminons maintenant une borne supérieure pour chacun des termes d'erreur. Pour cela, réécrivons ceux-ci au moyens du résultat du lemme A.1 :

$$\mathbb{E} \left\{ \left| \varepsilon^{(k)}(s, \tau) \right|^2 \right\} = \mathbb{E} \left\{ \left| \left\langle X, f_{s\tau}^{(k)} \right\rangle \right|^2 \right\} = \left\langle \mathcal{S}_X, \left| f_{s\tau}^{(k)} \right|^2 \right\rangle = \int_{\mathbb{R}} \mathcal{S}_X(\xi) \left| \hat{f}_{s\tau}^{(k)}(\xi) \right|^2 d\xi,$$

avec $k \in \{1, 2, 3\}$.

Concernant le premier terme d'erreur, un développement de Taylor de $\sqrt{\gamma'}$ en τ donne

$$\begin{aligned} \left| \hat{f}_{s\tau}^{(1)}(\xi) \right| &= \left| \int_{\mathbb{R}} \left(\sqrt{\gamma'(t)} - \sqrt{\gamma'(\tau)} \right) \psi_{s\tau}(t) e^{-2i\pi\gamma(t)\xi} dt \right| \\ &\leq \int_{\mathbb{R}} \left\| \frac{\gamma''}{2\sqrt{\gamma'}} \right\|_{\infty} |t - \tau| q^{-s/2} \left| \psi \left(\frac{t - \tau}{q^s} \right) \right| dt \\ &\leq q^{\frac{3s}{2}} \frac{\|\gamma''\|_{\infty}}{2\sqrt{c_{\gamma}}} I_{\psi}, \end{aligned}$$

où $I_{\psi} = \int_{\mathbb{R}} |t\psi(t)| dt$. De plus, l'hypothèse de localisation de l'ondelette ψ nous permet d'écrire

$$I_{\psi} \leq 2 \int_0^{\infty} \frac{t}{1+t^{\beta}} dt \leq 2 \left(\int_0^1 t dt + \int_1^{\infty} \frac{1}{t^{\beta-1}} dt \right) = \frac{\beta}{\beta-2}.$$

Finalement, on peut contrôler le premier terme d'erreur ainsi :

$$\mathbb{E} \left\{ \left| \varepsilon^{(1)}(s, \tau) \right|^2 \right\} \leq \left(q^{3s/2} \frac{\|\gamma''\|_{\infty}}{\sqrt{c_{\gamma}}} \frac{\beta \sigma_X}{2(\beta-2)} \right)^2.$$

Concernant le deuxième terme d'erreur, les applications successives de la formule d'Euler et de celle de Taylor-Lagrange donnent

$$\begin{aligned} \left| \hat{f}_{s\tau}^{(2)}(\xi) \right| &= \left| \int_{\mathbb{R}} \left(e^{-2i\pi\gamma(t)\xi} - e^{-2i\pi(\gamma(\tau) + (t-\tau)\gamma'(\tau))\xi} \right) \psi_{s\tau}(t) dt \right| \\ &\leq \int_{\mathbb{R}} \left| 1 - e^{-2i\pi(\gamma(\tau) + (t-\tau)\gamma'(\tau) - \gamma(t))\xi} \right| |\psi_{s\tau}(t)| dt \\ &\leq \int_{\mathbb{R}} 2 \left| \sin \left(\pi (\gamma(\tau) + (t-\tau)\gamma'(\tau) - \gamma(t)) \xi \right) \right| |\psi_{s\tau}(t)| dt \\ &\leq \int_{\mathbb{R}} 2 \left| \sin \left(\frac{\pi}{2} \xi (t-\tau)^2 \gamma''(t_*) \right) \right| |\psi_{s\tau}(t)| dt, \end{aligned}$$

pour un certain t_* entre t et τ . De plus, on a $|\sin(u)| \leq |u|$ et $|\sin(u)| \leq 1$, de sorte que :

$$\left| \hat{f}_{s\tau}^{(2)}(\xi) \right| \leq 2q^{s/2} \left(\int_J \frac{\pi}{2} |\xi| \|\gamma''\|_{\infty} q^{2s} t^2 |\psi(t)| dt + \int_{\mathbb{R} \setminus J} |\psi(t)| dt \right).$$

où $J = [-v, v]$. Alors, en bornant $|\psi(t)|$ par 1 sur J et par $|t|^{-\beta}$ sur $\mathbb{R} \setminus J$, on obtient :

$$\left| \hat{f}_{s\tau}^{(2)}(\xi) \right| \leq 2q^{s/2} \left(\frac{\pi}{3} |\xi| \|\gamma''\|_{\infty} q^{2s} v^3 + \frac{2}{(\beta-1)v^{\beta-1}} \right).$$

En recherchant à annuler la dérivée de cette expression par rapport à v , on peut montrer que la valeur de v minimisant le membre de droite de l'équation ci-dessus est $v = \left(\frac{\pi}{2}|\xi| \|\gamma''\|_\infty q^{2s}\right)^{-1/(\beta+2)}$. Ainsi, avec ce choix de v , on a :

$$\left| \hat{f}_{s\tau}^{(2)}(\xi) \right| \leq q^{\frac{5\beta-2}{2(\beta+2)}s} \frac{4(\beta+2)}{3(\beta-1)} \left(\frac{\pi}{2}|\xi| \|\gamma''\|_\infty \right)^{\frac{\beta-1}{\beta+2}}.$$

Finalement, le deuxième terme d'erreur est contrôlé ainsi :

$$\mathbb{E} \left\{ \left| \varepsilon^{(2)}(s, \tau) \right|^2 \right\} \leq \left(q^{\frac{5\beta-2}{2(\beta+2)}s} \frac{4(\beta+2)}{3(\beta-1)} \left(\frac{\pi}{2} \|\gamma''\|_\infty \right)^{\frac{\beta-1}{\beta+2}} \sqrt{2} I_X^{(\rho)} \right)^2.$$

Concernant le troisième terme d'erreur, l'application à α de l'inégalité de Taylor-Lagrange donne :

$$\begin{aligned} \left| \hat{f}_{s\tau}^{(3)}(\xi) \right| &= \left| \int_{\mathbb{R}} \sqrt{\gamma'(t)} \left(\frac{\alpha(t)}{\alpha(\tau)} - 1 \right) \psi_{s\tau}(t) e^{-2i\pi\gamma(t)\xi} dt \right| \\ &\leq \sqrt{C_\gamma} \int_{\mathbb{R}} \frac{\|\alpha'\|_\infty}{c_\alpha} |t - \tau| |\psi_{s\tau}(t)| dt = q^{3s/2} \sqrt{C_\gamma} \frac{\|\alpha'\|_\infty}{c_\alpha} I_\psi, \end{aligned}$$

Finalement, le troisième terme d'erreur est contrôlé ainsi :

$$\mathbb{E} \left\{ \left| \varepsilon^{(3)}(s, \tau) \right|^2 \right\} \leq \left(q^{3s/2} \frac{\sqrt{C_\gamma}}{c_\alpha} \frac{\beta\sigma_X}{\beta-2} \|\alpha'\|_\infty \right)^2.$$

Pour conclure la démonstration, les trois termes d'erreur dans l'équation (A.1) sont remplacés par leurs bornes supérieures. On obtient alors l'erreur d'approximation donnée dans le théorème. ■

A.1.2 Démonstration de la proposition 2.1

Cette proposition caractérise la loi de la transformée en ondelettes approchée $\widetilde{\mathcal{W}}_\gamma$ en fonction du spectre du signal stationnaire sous-jacent, de l'ondelette d'analyse ψ et des fonctions de déformation α et γ .

Démonstration de la proposition 2.1. Soit X une processus stochastique, stationnaire au sens large de moyenne nulle et de spectre \mathcal{S}_X . Notons les paramètres :

$$\begin{cases} \tilde{s} &= s + \log_q(\gamma'(\tau)), \\ \tilde{\tau} &= \gamma(\tau). \end{cases}$$

Alors,

$$\begin{aligned} \widetilde{\mathcal{W}}_\gamma(s, \tau) &= \alpha(\tau) \mathcal{W}_X \left(s + \log_q(\gamma'(\tau)), \gamma(\tau) \right) \\ &= \alpha(\tau) \langle X, \psi_{\tilde{s}\tilde{\tau}} \rangle. \end{aligned}$$

$\widetilde{\mathcal{W}}_Y$ est donc un champ aléatoire à deux dimensions et

$$\mathbb{E} \left\{ \widetilde{\mathcal{W}}_Y(s, \tau) \right\} = \alpha(\tau) \mathbb{E} \left\{ \langle X, \psi_{\bar{s}\bar{\tau}} \rangle \right\} = 0,$$

car X est de moyenne nulle. Par ailleurs, à l'aide du lemme A.1, on peut calculer sa covariance comme suit :

$$\begin{aligned} \mathbb{E} \left\{ \widetilde{\mathcal{W}}_Y(s, \tau) \overline{\widetilde{\mathcal{W}}_Y(s', \tau')} \right\} &= \alpha(\tau) \alpha(\tau') \mathbb{E} \left\{ \langle X, \psi_{\bar{s}\bar{\tau}} \rangle \overline{\langle X, \psi_{\bar{s}'\bar{\tau}'} \rangle} \right\} \\ &= \alpha(\tau) \alpha(\tau') \left\langle \mathcal{S}_X, \overline{\hat{\psi}_{\bar{s}\bar{\tau}}} \hat{\psi}_{\bar{s}'\bar{\tau}'} \right\rangle \\ &= \alpha(\tau) \alpha(\tau') \int_0^\infty \mathcal{S}_X(\xi) q^{\frac{\bar{s}+\bar{s}'}{2}} \overline{\hat{\psi}(q^{\bar{s}}\xi)} \hat{\psi}(q^{\bar{s}'}\xi) e^{2i\pi\xi(\bar{\tau}-\bar{\tau}')} d\xi. \end{aligned}$$

On retrouve la formule (2.6).

En outre, si l'on suppose maintenant que X est un processus gaussien, alors $\widetilde{\mathcal{W}}_Y$ est un champ aléatoire gaussien complexe. En effet, $\widetilde{\mathcal{W}}_Y$ est une transformation linéaire du processus X . Or, le caractère gaussien est stable par combinaison linéaire. On a également :

$$\begin{aligned} \mathbb{E} \left\{ \widetilde{\mathcal{W}}_Y(s, \tau) \widetilde{\mathcal{W}}_Y(s', \tau') \right\} &= \alpha(\tau) \alpha(\tau') \mathbb{E} \left\{ \langle X, \psi_{\bar{s}\bar{\tau}} \rangle \langle X, \psi_{\bar{s}'\bar{\tau}'} \rangle \right\} \\ &= \alpha(\tau) \alpha(\tau') \left\langle \mathcal{S}_X, \overline{\hat{\psi}_{\bar{s}\bar{\tau}}} \hat{\psi}_{\bar{s}'\bar{\tau}'} \right\rangle. \end{aligned}$$

Or, l'ondelette ψ étant analytique, les supports de $\hat{\psi}_{\bar{s}\bar{\tau}}$ et $\hat{\psi}_{\bar{s}'\bar{\tau}'}$ sont disjoints. Leur produit est donc identiquement nul. Ainsi :

$$\mathbb{E} \left\{ \widetilde{\mathcal{W}}_Y(s, \tau) \widetilde{\mathcal{W}}_Y(s', \tau') \right\} = 0, \quad \forall s, s', \tau, \tau'.$$

Le champ $\widetilde{\mathcal{W}}_Y$ est donc circulaire, ce qui conclue la démonstration. ■

A.1.3 Démonstration de la proposition 2.2

Cette proposition fournit une borne maximale pouvant être atteinte par le biais de l'estimateur du spectre-ondelettes de \mathcal{S}_X au cours des itérations de JEFAS. Celle-ci s'exprime en fonction de l'erreur maximale des estimées courantes des fonctions de déformation.

Démonstration de la proposition 2.2. Supposons que les estimées des paramètres de déformation $\tilde{\theta}_1^{(k)}$ et $\tilde{\theta}_2^{(k)}$ sont disponibles. Soit $\mathbf{w}_{x,s_m} \in \mathbb{C}^{N_\tau}$ le vecteur contenant les coefficients d'ondelettes du signal stationnaire original x à l'échelle s_m , et soit $\tilde{\mathbf{w}}_{x,s_m}^{(k)} \in \mathbb{C}^{N_\tau}$ l'estimée de \mathbf{w}_{x,s_m} après k itérations de l'algorithme. En considérant l'équation (2.15), nous avons

$$\tilde{\mathbf{w}}_{x,s_m}^{(k)} = \frac{1}{\sqrt{\tilde{\theta}_1^{(k)}}} \widetilde{\mathcal{W}}_y \left(s_m - \tilde{\theta}_2^{(k)}, \tau \right).$$

De plus, à partir de la définition de l'estimateur du spectre donné dans (2.16), on obtient

$$\mathbb{E} \left\{ \mathcal{J}_X^{(k)}(q^{-s_m} \nu_0) \right\} = \frac{1}{N_\tau \|\psi\|_2^2} \mathbb{E} \left\{ \left\| \tilde{\mathbf{w}}_{x, s_m}^{(k)} \right\|^2 \right\}.$$

Afin de simplifier les notations, définissons quelques grandeurs. Soient $s_m^{(\xi)} = s_m + \log_q(\xi)$ et $h(x) = \varphi_1(q^x) = q^x |\hat{\psi}(q^x)|^2$ pour $x \in \mathbb{R}$.

À l'aide de l'expression de la covariance dans l'équation (2.6), on peut écrire

$$\begin{aligned} \mathbb{E} \left\{ \left\| \tilde{\mathbf{w}}_{x, s_m}^{(k)} \right\|^2 \right\} &= \sum_{n=1}^{N_\tau} \frac{1}{\tilde{\theta}_{n,1}^{(k)}} \mathbb{E} \left\{ \widetilde{\mathcal{W}}_y \left(s_m - \tilde{\theta}_{n,2}^{(k)}, \tau_n \right) \overline{\widetilde{\mathcal{W}}_y \left(s_m - \tilde{\theta}_{n,2}^{(k)}, \tau_n \right)} \right\} \\ &= \sum_{n=1}^{N_\tau} \frac{\theta_{n,1}}{\tilde{\theta}_{n,1}^{(k)}} \int_0^\infty \frac{\mathcal{J}_X(\xi)}{\xi} \xi q^{s_m - \tilde{\theta}_{n,2}^{(k)} + \theta_{n,2}} \left| \hat{\psi} \left(\xi q^{s_m - \tilde{\theta}_{n,2}^{(k)} + \theta_{n,2}} \right) \right|^2 d\xi \\ &= \sum_{n=1}^{N_\tau} \frac{\theta_{n,1}}{\tilde{\theta}_{n,1}^{(k)}} \int_0^\infty \frac{\mathcal{J}_X(\xi)}{\xi} h \left(s_m^{(\xi)} + \theta_{n,2} - \tilde{\theta}_{n,2}^{(k)} \right) d\xi \\ &= \int_0^\infty \frac{\mathcal{J}_X(\xi)}{\xi} \left(\sum_{n=1}^{N_\tau} \frac{\theta_{n,1}}{\tilde{\theta}_{n,1}^{(k)}} h \left(s_m^{(\xi)} + \theta_{n,2} - \tilde{\theta}_{n,2}^{(k)} \right) \right) d\xi. \end{aligned}$$

Scindons maintenant le biais en deux termes tels que $b_{\mathcal{J}_X}^{(k)}(m) = g_1(m) + g_2(m)$, où g_1 et g_2 sont donnés par

$$\begin{aligned} g_1(m) &= \frac{N_\tau^{-1}}{\|\psi\|_2^2} \int_0^\infty \frac{\mathcal{J}_X(\xi)}{\xi} \left(\sum_{n=1}^{N_\tau} \left(\frac{\theta_{n,1}}{\tilde{\theta}_{n,1}^{(k)}} - 1 \right) h \left(s_m^{(\xi)} + \theta_{n,2} - \tilde{\theta}_{n,2}^{(k)} \right) \right) d\xi, \\ g_2(m) &= \frac{N_\tau^{-1}}{\|\psi\|_2^2} \int_0^\infty \frac{\mathcal{J}_X(\xi)}{\xi} \left(\sum_{n=1}^{N_\tau} \left(h \left(s_m^{(\xi)} + \theta_{n,2} - \tilde{\theta}_{n,2}^{(k)} \right) - h \left(s_m^{(\xi)} \right) \right) \right) d\xi. \end{aligned}$$

En ce qui concerne le premier terme, nous avons directement

$$\begin{aligned} |g_1(m)| &\leq \frac{\|h\|_\infty}{\|\psi\|_2^2} \frac{1}{N_\tau} \sum_{n=1}^{N_\tau} \frac{|\theta_{n,1} - \tilde{\theta}_{n,1}^{(k)}|}{\tilde{\theta}_{n,1}^{(k)}} \int_0^\infty \frac{\mathcal{J}_X(\xi)}{\xi} d\xi \\ &\leq \frac{\|h\|_\infty}{\|\psi\|_2^2 c_{\theta_1}} \left\| \theta_1 - \tilde{\theta}_1^{(k)} \right\|_\infty J_X. \end{aligned}$$

De plus, on a $\|h\|_\infty = \|\varphi_1\|_\infty$, et la régularité et l'hypothèse de décroissance sur $\hat{\psi}$ nous permettent d'écrire $\varphi_1(u) = o_{u \rightarrow \infty}(u^{-3}) \xrightarrow{u \rightarrow \infty} 0$. Ainsi, φ_1 est bornée et $K'_1 < \infty$. Cela mène à :

$$|g_1(m)| \leq \frac{J_X K'_1}{\|\psi\|_2^2} \left\| \theta_1 - \tilde{\theta}_1^{(k)} \right\|_\infty.$$

En ce qui concerne le second terme, un développement de Taylor de h autour de $s_m^{(\xi)}$ donne

$$\begin{aligned} |g_2(m)| &\leq \frac{N_\tau^{-1}}{\|\psi\|_2^2} \int_0^\infty \frac{\mathcal{S}_X(\xi)}{\xi} \sum_{n=1}^{N_\tau} \left| h\left(s_m^{(\xi)} + \theta_{n,2} - \tilde{\theta}_{n,2}^{(k)}\right) - h\left(s_m^{(\xi)}\right) \right| d\xi \\ &\leq \frac{\|h'\|_\infty}{\|\psi\|_2^2} \frac{1}{N_\tau} \sum_{n=1}^{N_\tau} \left| \theta_{n,2} - \tilde{\theta}_{n,2}^{(k)} \right| \int_0^\infty \frac{\mathcal{S}_X(\xi)}{\xi} d\xi \\ &\leq \frac{\|h'\|_\infty}{\|\psi\|_2^2} \left\| \theta_2 - \tilde{\theta}_2^{(k)} \right\|_\infty J_X. \end{aligned}$$

Par ailleurs, $\forall x \in \mathbb{R}$

$$\begin{aligned} |h'(x)| &= |\ln(q) q^x \varphi_1'(q^x)| \leq \ln(q) \left(q^x |\hat{\psi}(q^x)|^2 + 2q^{2x} |\hat{\psi}(q^x) \hat{\psi}'(q^x)| \right) \\ &\leq \ln(q) (\|\varphi_1\|_\infty + 2\|\hat{\psi}'\|_\infty \|\varphi_2\|_\infty) = K_2'. \end{aligned}$$

De plus, l'hypothèse de décroissance sur $\hat{\psi}$ donne $|\varphi_2(u)| = o_{u \rightarrow \infty}(1) \xrightarrow{u \rightarrow \infty} 0$. Ainsi φ_2 est borné et $K_2' < \infty$. Cela mène à :

$$|g_2(m)| \leq \frac{J_X K_2'}{\|\psi\|_2^2} \left\| \theta_2 - \tilde{\theta}_2^{(k)} \right\|_\infty.$$

On conclue la démonstration en sommant les bornes supérieures de $|g_1|$ et $|g_2|$ afin d'obtenir la borne supérieure de $b_{\mathcal{S}_X}^{(k)}$. Il est à noter que cette borne ne dépend pas de m . ■

A.2 Démonstration du chapitre 3

A.2.1 Démonstration du théorème d'approximation 3.1

Démonstration du théorème 3.1. L'erreur d'approximation s'écrit :

$$\begin{aligned} \epsilon_\tau &= \mathbf{w}_{\mathbf{z},\tau} - \mathbf{A}(\tau) \mathbf{w}_{\mathbf{y},\tau} \\ &= \langle \mathbf{A} \mathbf{y}, \psi_{\mathbf{s}\tau} \rangle - \langle \mathbf{A}(\tau) \mathbf{y}, \psi_{\mathbf{s}\tau} \rangle \\ &= \langle (\mathbf{A} - \mathbf{A}(\tau)) \mathbf{y}, \psi_{\mathbf{s}\tau} \rangle. \end{aligned}$$

Or, la formule de Taylor-Lagrange nous indique que pour toutes les valeurs de t , il existe une valeur $u_\tau(t)$ comprise en t et τ telle que :

$$\mathbf{A}(t) - \mathbf{A}(\tau) = \mathbf{A}'(u_\tau(t))(t - \tau).$$

Notons $\mathbf{M}_\tau(t)$ la matrice réelle : $\mathbf{M}_\tau(t) = \mathbf{A}'(u_\tau(t))(t - \tau) \in \mathbb{R}^{N_z \times N_y}$. On a donc :

$$\begin{aligned} [\epsilon_\tau]_{im} &= \sum_{j=1}^{N_y} \langle [\mathbf{M}_\tau]_{ij} \gamma_j, \psi_{s_m \tau} \rangle \\ &= \sum_{j=1}^{N_y} \langle \mathcal{D}_{\gamma_j} X_j, [\mathbf{M}_\tau]_{ij} \psi_{s_m \tau} \rangle \\ &= \sum_{j=1}^{N_y} \langle X_j, \mathcal{D}_{\gamma_j^{-1}} [\mathbf{M}_\tau]_{ij} \psi_{s_m \tau} \rangle . \end{aligned}$$

Notons $g_{ij,s\tau} = \mathcal{D}_{\gamma_j^{-1}} [\mathbf{M}_\tau]_{ij} \psi_{s\tau}$, alors

$$\mathbb{E} \left\{ |[\epsilon_\tau]_{im}|^2 \right\} = \sum_{j=1}^{N_y} \sum_{k=1}^{N_y} \mathbb{E} \left\{ \langle X_j, g_{ik,s_m \tau} \rangle \overline{\langle X_k, g_{ij,s_m \tau} \rangle} \right\} .$$

Comme les sources stationnaires X_j sont indépendantes les unes des autres, seuls les termes $k = j$ sont non nuls dans la somme précédente. Par ailleurs, grâce au résultat du lemme A.1, on a :

$$\begin{aligned} \mathbb{E} \left\{ |[\epsilon_\tau]_{im}|^2 \right\} &= \sum_{j=1}^{N_y} \mathbb{E} \left\{ \langle X_j, g_{ij,s_m \tau} \rangle \overline{\langle X_j, g_{ij,s_m \tau} \rangle} \right\} \\ &= \sum_{j=1}^{N_y} \left\langle \mathcal{S}_{X_j}, |\hat{g}_{ij,s_m \tau}|^2 \right\rangle . \end{aligned}$$

Bornons $|\hat{g}_{ij,s\tau}|$:

$$\begin{aligned} |\hat{g}_{ij,s_m \tau}(\xi)| &= \left| \int_{\mathbb{R}} [\mathbf{M}_\tau]_{ij}(t) \psi_{s_m \tau}(t) \sqrt{\gamma_j'(t)} e^{-2i\pi \gamma_j(t) \xi} dt \right| \\ &\leq \int_{\mathbb{R}} [\mathbf{A}'_\infty]_{ij} |t - \tau| |\psi_{s_m \tau}(t)| \|\gamma_j'\|_\infty^{\frac{1}{2}} dt \\ &\leq [\mathbf{A}'_\infty]_{ij} [\gamma'_\infty]_j^{\frac{1}{2}} q^{\frac{3s_m}{2}} \int_{\mathbb{R}} |t| |\psi(t)| dt . \end{aligned}$$

On a donc

$$\begin{aligned} \mathbb{E} \left\{ |[\epsilon_\tau]_{im}|^2 \right\} &\leq \sum_{j=1}^{N_y} q^{3s_m} \sigma_X^2 [\mathbf{A}'_\infty \circ 2]_{ij} [\gamma'_\infty]_j I_\psi^2 \\ &\leq q^{3s_m} \sigma_X^2 [\mathbf{A}'_\infty \circ 2]_{i \cdot} \gamma'_\infty I_\psi^2 , \end{aligned}$$

où

$$\begin{aligned} I_\psi &= \int_{\mathbb{R}} |t \psi(t)| dt \\ &\leq 2 \left(\int_0^1 t dt + \int_1^\infty \frac{1}{t^{\beta-1}} dt \right) \\ &\leq \frac{\beta}{\beta-2}. \end{aligned}$$

On a finalement

$$\mathbb{E} \{ |\epsilon_\tau|^2 \} \leq \frac{\sigma_X^2 \beta^2}{(\beta-2)^2} \mathbf{A}'_{\infty} \gamma'_{\infty} (q^{3s})^T. \blacksquare$$

A.3 Démonstrations du chapitre 4

A.3.1 Démonstration de la proposition 4.1

On démontre ici l'expression de la loi *a posteriori*. En particulier, on démontre les équations (4.12) et (4.13).

Démonstration. Comme composition de lois gaussiennes, la densité de probabilité *a posteriori* $p(\mathbf{W}|\mathbf{y})$ est une loi gaussienne complexe (mais pas circulaire). En toute généralité, la densité de probabilité d'un vecteur aléatoire gaussien complexe s'écrit sous la forme suivante [60] :

$$p(\mathbf{W}|\mathbf{y}) = K \exp \left(-\frac{1}{2} \begin{pmatrix} \overline{\mathbf{W}} - \overline{\boldsymbol{\mu}_{pos}} \\ \mathbf{W} - \boldsymbol{\mu}_{pos} \end{pmatrix}^T \begin{pmatrix} \boldsymbol{\Gamma}_{pos} & \mathbf{R}_{pos} \\ \mathbf{R}_{pos}^T & \overline{\boldsymbol{\Gamma}_{pos}} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{W} - \boldsymbol{\mu}_{pos} \\ \overline{\mathbf{W}} - \overline{\boldsymbol{\mu}_{pos}} \end{pmatrix} \right)$$

avec $\boldsymbol{\mu}_{pos}$ la moyenne, $\boldsymbol{\Gamma}_{pos}$ la matrice de covariance et \mathbf{R}_{pos} la matrice de relation de la loi *a posteriori*. K est une constante de normalisation. On obtient l'inverse de la matrice définie par blocs grâce au complément de Schur. En utilisant également les propriétés des matrices de covariance et de relation ($\boldsymbol{\Gamma}_{pos}^T = \overline{\boldsymbol{\Gamma}_{pos}}$ et $\mathbf{R}_{pos}^T = \mathbf{R}_{pos}$), on peut montrer que :

$$\begin{pmatrix} \boldsymbol{\Gamma}_{pos} & \mathbf{R}_{pos} \\ \mathbf{R}_{pos}^T & \overline{\boldsymbol{\Gamma}_{pos}} \end{pmatrix}^{-1} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^T & \mathbf{A} \end{pmatrix}$$

avec :

$$\mathbf{A} = (\boldsymbol{\Gamma}_{pos} - \mathbf{R}_{pos} \overline{\boldsymbol{\Gamma}_{pos}}^{-1} \mathbf{R}_{pos}^T)^{-1}, \quad (\text{A.2})$$

$$\mathbf{B} = -(\boldsymbol{\Gamma}_{pos} - \mathbf{R}_{pos} \overline{\boldsymbol{\Gamma}_{pos}}^{-1} \mathbf{R}_{pos}^T)^{-1} \mathbf{R}_{pos} \overline{\boldsymbol{\Gamma}_{pos}}^{-1}, \quad (\text{A.3})$$

et on a les propriétés $\mathbf{A}^T = \overline{\mathbf{A}}$ et $\mathbf{B}^T = \mathbf{B}$. Après développement, on a :

$$\ln(p(\mathbf{W}|\mathbf{y})) = K' - \text{Re}(\mathbf{W}^T \overline{\mathbf{B}} \mathbf{W}) - \mathbf{W}^H \mathbf{A} \mathbf{W} + 2 \text{Re}(\mathbf{W}^H (\mathbf{A} \boldsymbol{\mu}_{pos} + \mathbf{B} \overline{\boldsymbol{\mu}_{pos}})).$$

On va maintenant chercher à identifier le vecteur $\boldsymbol{\mu}_{pos}$ et les matrices $\boldsymbol{\Gamma}_{pos}$ et \mathbf{R}_{pos} . L'équation (4.11), formée à partir de la règle de Bayes, nous permet d'obtenir :

$$\ln(p(\mathbf{W}|\mathbf{y})) = K'' - \frac{1}{2\sigma_y^2} \left((\mathbf{y} - \text{Re}(\mathbf{D}\mathbf{W}))^T (\mathbf{y} - \text{Re}(\mathbf{D}\mathbf{W})) \right) - \mathbf{W}^H \boldsymbol{\Gamma}_{pri}^{-1} \mathbf{W}.$$

Avec l'identité $\text{Re}(z) = \frac{z + \bar{z}}{2}$, on obtient :

$$\begin{aligned} \ln(p(\mathbf{W}|\mathbf{y})) = & K''' - \text{Re} \left(\mathbf{W}^T \left(\frac{\mathbf{D}^T \mathbf{D}}{4\sigma_y^2} \right) \mathbf{W} \right) - \mathbf{W}^H \left(\frac{\mathbf{D}^H \mathbf{D}}{4\sigma_y^2} - \boldsymbol{\Gamma}_{pri}^{-1} \right) \mathbf{W} \\ & + 2 \text{Re} \left(\mathbf{W}^H \left(\frac{\mathbf{D}^H \mathbf{y}}{2\sigma_y^2} \right) \right). \end{aligned}$$

Par identification, on a :

$$\begin{cases} \mathbf{A} = \frac{\mathbf{D}^H \mathbf{D}}{4\sigma_y^2} + \boldsymbol{\Gamma}_{pri}^{-1} \\ \mathbf{B} = \frac{\mathbf{D}^T \mathbf{D}}{4\sigma_y^2} \\ \mathbf{A} \boldsymbol{\mu}_{pos} + \mathbf{B} \overline{\boldsymbol{\mu}_{pos}} = \frac{\mathbf{D}^H \mathbf{y}}{2\sigma_y^2}. \end{cases} \quad (\text{A.4})$$

Intéressons-nous à la moyenne *a posteriori* :

$$\left(\mathbf{D}^H \mathbf{D} + 4\sigma_y^2 \boldsymbol{\Gamma}_{pri}^{-1} \right) \boldsymbol{\mu}_{pos} + \overline{\mathbf{D}^T \mathbf{D} \overline{\boldsymbol{\mu}_{pos}}} = 2\mathbf{D}^H \mathbf{y}$$

ou de manière équivalente :

$$\mathbf{D}^H \text{Re}(\mathbf{D} \boldsymbol{\mu}_{pos}) + 2\sigma_y^2 \boldsymbol{\Gamma}_{pri}^{-1} \boldsymbol{\mu}_{pos} = \mathbf{D}^H \mathbf{y},$$

d'où :

$$\boldsymbol{\mu}_{pos} = \frac{1}{2\sigma_y^2} \boldsymbol{\Gamma}_{pri} \mathbf{D}^H (\mathbf{y} - \text{Re}(\mathbf{D} \boldsymbol{\mu}_{pos})).$$

Notons $\mathbf{z} = \text{Re}(\mathbf{D} \boldsymbol{\mu}_{pos})$ alors :

$$\mathbf{z} = \frac{1}{2\sigma_y^2} \text{Re}(\mathbf{D} \boldsymbol{\Gamma}_{pri} \mathbf{D}^H) (\mathbf{y} - \mathbf{z})$$

d'où :

$$\mathbf{z} = \left(2\sigma_y^2 \mathbf{I} + \text{Re}(\mathbf{D} \boldsymbol{\Gamma}_{pri} \mathbf{D}^H) \right)^{-1} \text{Re}(\mathbf{D} \boldsymbol{\Gamma}_{pri} \mathbf{D}^H) \mathbf{y},$$

et donc :

$$\boldsymbol{\mu}_{pos} = \frac{1}{2\sigma_y^2} \boldsymbol{\Gamma}_{pri} \mathbf{D}^H \left(\mathbf{I} - \left(2\sigma_y^2 \mathbf{I} + \text{Re}(\mathbf{D} \boldsymbol{\Gamma}_{pri} \mathbf{D}^H) \right)^{-1} \text{Re}(\mathbf{D} \boldsymbol{\Gamma}_{pri} \mathbf{D}^H) \right) \mathbf{y}.$$

On peut donc écrire μ_{pos} sous la forme :

$$\mu_{pos} = \frac{1}{2} \Gamma_{pri} \mathbf{D}^H \mathbf{C}_y^{-1} \mathbf{y}.$$

avec :

$$\mathbf{C}_y = \sigma_y^2 \left(\mathbf{I} - \left(2\sigma_y^2 \mathbf{I} + \text{Re} \left(\mathbf{D} \Gamma_{pri} \mathbf{D}^H \right) \right)^{-1} \text{Re} \left(\mathbf{D} \Gamma_{pri} \mathbf{D}^H \right) \right)^{-1},$$

que l'on réécrit grâce à la formule de Woodbury [31] :

$$\mathbf{C}_y = \sigma_y^2 \mathbf{I} + \frac{1}{2} \text{Re} \left(\mathbf{D} \Gamma_{pri} \mathbf{D}^H \right).$$

Ceci démontre la formule (4.12).

Intéressons-nous maintenant à l'expression de la covariance *a posteriori*. Les deux premières équations du système (A.4) combinées avec les expressions (A.2) et (A.3) de **A** et **B** donnent :

$$\begin{cases} \Gamma_{pos} - \mathbf{R}_{pos} \overline{\Gamma_{pos}}^{-1} \overline{\mathbf{R}_{pos}} &= 4\sigma_y^2 \left(\mathbf{D}^H \mathbf{D} + 4\sigma_y^2 \Gamma_{pri}^{-1} \right)^{-1} \\ \mathbf{R}_{pos} &= - \left(\mathbf{D}^H \mathbf{D} + 4\sigma_y^2 \Gamma_{pri}^{-1} \right)^{-1} \overline{\mathbf{D}^T \mathbf{D} \Gamma_{pos}} \end{cases}$$

soit :

$$\begin{aligned} \Gamma_{pos} &= 4\sigma_y^2 \left(\mathbf{D}^H \mathbf{D} + 4\sigma_y^2 \Gamma_{pri}^{-1} - \overline{\mathbf{D}^T \mathbf{D}} \left(\mathbf{D}^T \overline{\mathbf{D}} + 4\sigma_y^2 \overline{\Gamma_{pri}^{-1}} \right)^{-1} \mathbf{D}^T \mathbf{D} \right)^{-1} \\ &= \left(\Gamma_{pri}^{-1} + \frac{1}{4\sigma_y^2} \mathbf{D}^H \underbrace{\left(\mathbf{I} - \overline{\mathbf{D}} \left(\mathbf{D}^H \overline{\mathbf{D}} + 4\sigma_y^2 \overline{\Gamma_{pri}^{-1}} \right)^{-1} \mathbf{D}^T \right) \mathbf{D}}_{\triangleq \mathbf{M}^{-1}} \right)^{-1} \\ &= \Gamma_{pri} - \Gamma_{pri} \mathbf{D}^H \left(4\sigma_y^2 \mathbf{M} + \mathbf{D} \Gamma_{pri} \mathbf{D}^H \right)^{-1} \mathbf{D} \Gamma_{pri}. \end{aligned}$$

Or :

$$\begin{aligned} \overline{\mathbf{M}} &= \left(\mathbf{I} - \mathbf{D} \left(\mathbf{D}^H \mathbf{D} + 4\sigma_y^2 \Gamma_{pri}^{-1} \right)^{-1} \mathbf{D}^H \right)^{-1} \\ &= \mathbf{I} + \mathbf{D} \left(\mathbf{D}^H \mathbf{D} + 4\sigma_y^2 \Gamma_{pri}^{-1} - \mathbf{D}^H \mathbf{D} \right)^{-1} \mathbf{D}^H \\ &= \mathbf{I} + \frac{1}{4\sigma_y^2} \mathbf{D} \Gamma_{pri} \mathbf{D}^H, \end{aligned}$$

d'où

$$\Gamma_{pos} = \Gamma_{pri} - \Gamma_{pri} \mathbf{D}^H \left(4\sigma_y^2 \mathbf{I} + 2 \text{Re} \left(\mathbf{D} \Gamma_{pri} \mathbf{D}^H \right) \right)^{-1} \mathbf{D} \Gamma_{pri}.$$

Ceci démontre l'expression (4.13). ■

A.3.2 Démonstration de la proposition 4.2

On démontre les formulations 4.19 et (4.19) du problème EM.

Démonstration. La loi jointe s'écrit :

$$p(\mathbf{y}, \mathbf{W} | \boldsymbol{\theta}) = p(\mathbf{y} | \mathbf{W}) p(\mathbf{W} | \boldsymbol{\theta}) = p(\mathbf{W} | \mathbf{y}, \boldsymbol{\theta}) p(\mathbf{y})$$

avec :

$$p(\mathbf{y} | \mathbf{W}) = \mathcal{N}(\text{Re}(\mathbf{D}\mathbf{W}), \sigma_y^2 \mathbf{I}) \quad (\text{A.5})$$

$$p(\mathbf{W} | \boldsymbol{\theta}) = \mathcal{CN}(\mathbf{0}, \boldsymbol{\Gamma}_{pri}(\boldsymbol{\theta})) \quad (\text{A.6})$$

$$p(\mathbf{W} | \mathbf{y}, \boldsymbol{\theta}) = \mathcal{CN}(\boldsymbol{\mu}_{pos}(\boldsymbol{\theta}), \boldsymbol{\Gamma}_{pos}(\boldsymbol{\theta}), \mathbf{R}_{pos}(\boldsymbol{\theta})) \quad (\text{A.7})$$

Calculons $Q(\boldsymbol{\theta}, \tilde{\boldsymbol{\theta}}^{(k-1)})$:

$$\begin{aligned} Q(\boldsymbol{\theta}, \tilde{\boldsymbol{\theta}}^{(k-1)}) &= \mathbb{E}_{\mathbf{W} | \mathbf{y}, \tilde{\boldsymbol{\theta}}^{(k-1)}} \{ \log p(\mathbf{y}, \mathbf{W} | \boldsymbol{\theta}) \} \\ &= \mathbb{E}_{\mathbf{W} | \mathbf{y}, \tilde{\boldsymbol{\theta}}^{(k-1)}} \{ \log p(\mathbf{y} | \mathbf{W}, \boldsymbol{\theta}) \} + \mathbb{E}_{\mathbf{W} | \mathbf{y}, \tilde{\boldsymbol{\theta}}^{(k-1)}} \{ \log p(\mathbf{W} | \boldsymbol{\theta}) \} . \end{aligned}$$

Le premier terme étant indépendant de $\boldsymbol{\theta}$ (cf. équation (A.5)), il n'intervient pas dans l'étape de maximisation. Regardons le second terme.

$$\begin{aligned} \mathbb{E}_{\mathbf{W} | \mathbf{y}, \tilde{\boldsymbol{\theta}}^{(k-1)}} \{ \log p(\mathbf{W} | \boldsymbol{\theta}) \} &= \mathbb{E}_{\mathbf{W} | \mathbf{y}, \tilde{\boldsymbol{\theta}}^{(k-1)}} \left\{ \log \left(\frac{1}{\pi^{ST} \det(\boldsymbol{\Gamma}_{pri}(\boldsymbol{\theta}))} e^{-\mathbf{W}^H \boldsymbol{\Gamma}_{pri}(\boldsymbol{\theta})^{-1} \mathbf{W}} \right) \right\} \\ &= -\log |\det(\boldsymbol{\Gamma}_{pri}(\boldsymbol{\theta}))| - \mathbb{E}_{\mathbf{W} | \mathbf{y}, \tilde{\boldsymbol{\theta}}^{(k-1)}} \left\{ \mathbf{W}^H \boldsymbol{\Gamma}_{pri}(\boldsymbol{\theta})^{-1} \mathbf{W} \right\} - ST \log(\pi) \end{aligned} \quad (\text{A.8})$$

Afin de calculer le terme $\mathbb{E}_{\mathbf{W} | \mathbf{y}, \tilde{\boldsymbol{\theta}}^{(k-1)}} \left\{ \mathbf{W}^H \boldsymbol{\Gamma}_{pri}(\boldsymbol{\theta})^{-1} \mathbf{W} \right\}$, notons $\mathbf{X} = \text{Re}(\mathbf{W})$ et $\mathbf{Y} = \text{Im}(\mathbf{W})$ et $\mathbf{Z} = (\mathbf{X}^T \ \mathbf{Y}^T)^T$. On a

$$\begin{aligned} \mathbb{E}_{\mathbf{W} | \mathbf{y}, \tilde{\boldsymbol{\theta}}^{(k-1)}} \left\{ \mathbf{W}^H \boldsymbol{\Gamma}_{pri}(\boldsymbol{\theta})^{-1} \mathbf{W} \right\} &= \mathbb{E}_{\mathbf{W} | \mathbf{y}, \tilde{\boldsymbol{\theta}}^{(k-1)}} \left\{ \mathbf{X}^T \boldsymbol{\Gamma}_{pri}(\boldsymbol{\theta})^{-1} \mathbf{X} \right\} \\ &\quad + \mathbb{E}_{\mathbf{W} | \mathbf{y}, \tilde{\boldsymbol{\theta}}^{(k-1)}} \left\{ \mathbf{Y}^T \boldsymbol{\Gamma}_{pri}(\boldsymbol{\theta})^{-1} \mathbf{Y} \right\} \\ &= \mathbb{E}_{\mathbf{W} | \mathbf{y}, \tilde{\boldsymbol{\theta}}^{(k-1)}} \left\{ \mathbf{Z}^T \begin{pmatrix} \boldsymbol{\Gamma}_{pri}(\boldsymbol{\theta})^{-1} & 0 \\ 0 & \boldsymbol{\Gamma}_{pri}(\boldsymbol{\theta})^{-1} \end{pmatrix} \mathbf{Z} \right\} . \end{aligned}$$

Or

$$\mathbf{Z} | \mathbf{y}, \tilde{\boldsymbol{\theta}}^{(k-1)} \sim \mathcal{N} \left(\begin{pmatrix} \text{Re} \left(\boldsymbol{\mu}_{pos} \left(\tilde{\boldsymbol{\theta}}^{(k-1)} \right) \right) \\ \text{Im} \left(\boldsymbol{\mu}_{pos} \left(\tilde{\boldsymbol{\theta}}^{(k-1)} \right) \right) \end{pmatrix}, \mathbf{G} \right)$$

avec

$$\mathbf{G} = \frac{1}{2} \begin{pmatrix} \text{Re} \left(\Gamma_{pos} \left(\tilde{\boldsymbol{\theta}}^{(k-1)} \right) + \mathbf{R}_{pos} \left(\tilde{\boldsymbol{\theta}}^{(k-1)} \right) \right) & \text{Im} \left(-\Gamma_{pos} \left(\tilde{\boldsymbol{\theta}}^{(k-1)} \right) + \mathbf{R}_{pos} \left(\tilde{\boldsymbol{\theta}}^{(k-1)} \right) \right) \\ \text{Im} \left(\Gamma_{pos} \left(\tilde{\boldsymbol{\theta}}^{(k-1)} \right) + \mathbf{R}_{pos} \left(\tilde{\boldsymbol{\theta}}^{(k-1)} \right) \right) & \text{Re} \left(\Gamma_{pos} \left(\tilde{\boldsymbol{\theta}}^{(k-1)} \right) - \mathbf{R}_{pos} \left(\tilde{\boldsymbol{\theta}}^{(k-1)} \right) \right) \end{pmatrix},$$

d'où

$$\begin{aligned} \mathbb{E}_{\mathbf{W}|\mathbf{y}, \tilde{\boldsymbol{\theta}}^{(k-1)}} \left\{ \mathbf{W}^H \Gamma_{pri}(\boldsymbol{\theta})^{-1} \mathbf{W} \right\} &= \text{Trace} \left(\begin{pmatrix} \Gamma_{pri}(\boldsymbol{\theta})^{-1} & 0 \\ 0 & \Gamma_{pri}(\boldsymbol{\theta})^{-1} \end{pmatrix} \mathbf{G} \right) \\ &+ \begin{pmatrix} \text{Re} \left(\boldsymbol{\mu}_{pos} \left(\tilde{\boldsymbol{\theta}}^{(k-1)} \right) \right) \\ \text{Im} \left(\boldsymbol{\mu}_{pos} \left(\tilde{\boldsymbol{\theta}}^{(k-1)} \right) \right) \end{pmatrix}^T \Gamma_{pri}(\boldsymbol{\theta})^{-1} \begin{pmatrix} \text{Re} \left(\boldsymbol{\mu}_{pos} \left(\tilde{\boldsymbol{\theta}}^{(k-1)} \right) \right) \\ \text{Im} \left(\boldsymbol{\mu}_{pos} \left(\tilde{\boldsymbol{\theta}}^{(k-1)} \right) \right) \end{pmatrix} \\ &= \text{Trace} \left(\Gamma_{pri}(\boldsymbol{\theta})^{-1} \text{Re} \left(\Gamma_{pos} \left(\tilde{\boldsymbol{\theta}}^{(k-1)} \right) \right) \right) \\ &+ \boldsymbol{\mu}_{pos} \left(\tilde{\boldsymbol{\theta}}^{(k-1)} \right)^H \Gamma_{pri}(\boldsymbol{\theta})^{-1} \boldsymbol{\mu}_{pos} \left(\tilde{\boldsymbol{\theta}}^{(k-1)} \right). \end{aligned}$$

On a alors

$$\begin{aligned} \tilde{\boldsymbol{\theta}}^{(k)} &= \arg \max_{\boldsymbol{\theta}} \left[-\log |\det(\Gamma_{pri}(\boldsymbol{\theta}))| + \boldsymbol{\mu}_{pos} \left(\tilde{\boldsymbol{\theta}}^{(k-1)} \right)^H \Gamma_{pri}(\boldsymbol{\theta})^{-1} \boldsymbol{\mu}_{pos} \left(\tilde{\boldsymbol{\theta}}^{(k-1)} \right) \right. \\ &\quad \left. + \text{Trace} \left(\Gamma_{pri}(\boldsymbol{\theta})^{-1} \text{Re} \left(\Gamma_{pos} \left(\tilde{\boldsymbol{\theta}}^{(k-1)} \right) \right) \right) \right] \\ &= \arg \min_{\boldsymbol{\theta}} \left[\sum_{n=1}^{N_\tau} \log |\det(\mathbf{C}(\theta_n))| + \tilde{\mathbf{w}}_n^{(k-1)H} \Gamma_{pri}(\boldsymbol{\theta})^{-1} \tilde{\mathbf{w}}_n^{(k-1)} \right. \\ &\quad \left. + \text{Trace} \left(\mathbf{C}(\theta_n)^{-1} \text{Re} \left(\Gamma_n \left(\tilde{\boldsymbol{\theta}}^{(k-1)} \right) \right) \right) \right]. \end{aligned}$$

Ceci conclut la démonstration. ■

A.4 Démonstrations du chapitre 5

A.4.1 Démonstration du théorème 5.2

On démontre ici les équations aux dérivées partielles (5.17) et (5.18), qui sont approximativement vérifiées par \tilde{G}_y , la représentation temps-fréquence idéale approchée à la résolution η d'un signal localement harmonique (à fréquences rapides ou non).

Démonstration du point 1. On rappelle l'expression de \tilde{G}_y :

$$\tilde{G}_y(\xi, t) = \sum_{k=0}^K A_k(t) e^{2i\pi\phi_k(t)} \frac{1}{\eta} h \left(\frac{\xi - \phi'_k(t)}{\eta} \right).$$

On en déduit que :

$$\begin{aligned}\partial_t \tilde{G}_y(\xi, t) &= \sum_{k=1}^K A'_k(t) e^{2i\pi\phi_k(t)} \frac{1}{\eta} h \left(\frac{\xi - \phi'_k(t)}{\eta} \right) \\ &\quad + 2i\pi \sum_{k=1}^K A_k(t) \phi'_k(t) e^{2i\pi\phi_k(t)} \frac{1}{\eta} h \left(\frac{\xi - \phi'_k(t)}{\eta} \right) \\ &\quad - \sum_{k=1}^K A_k(t) e^{2i\pi\phi_k(t)} \frac{1}{\eta^2} \phi''_k(t) h' \left(\frac{\xi - \phi'_k(t)}{\eta} \right) .\end{aligned}$$

Nous nous plaçons, dans un premier temps, dans le cas où y est un signal localement harmonique. Notons $E_1(\xi, t)$ l'erreur d'approximation sur l'équation aux dérivées partielles correspondante (5.17). Alors :

$$E_1(\xi, t) = \partial_t \tilde{G}_y(\xi, t) - 2i\pi \xi \tilde{G}_y(\xi, t) .$$

L'erreur d'approximation se décompose sous la forme :

$$E_1(\xi, t) = e_1(\xi, t) + e_2(\xi, t) + e_3(\xi, t) ,$$

où

$$e_1(\xi, t) = \sum_{k=1}^K A'_k(t) e^{2i\pi\phi_k(t)} \frac{1}{\eta} h \left(\frac{\xi - \phi'_k(t)}{\eta} \right) ; \quad (\text{A.9})$$

$$e_2(\xi, t) = 2i\pi \sum_{k=1}^K A_k(t) (\xi - \phi'_k(t)) e^{2i\pi\phi_k(t)} \frac{1}{\eta} h \left(\frac{\xi - \phi'_k(t)}{\eta} \right) ; \quad (\text{A.10})$$

$$e_3(\xi, t) = - \sum_{k=1}^K A_k(t) e^{2i\pi\phi_k(t)} \phi''_k(t) \frac{1}{\eta} \partial_\xi h \left(\frac{\xi - \phi'_k(t)}{\eta} \right) . \quad (\text{A.11})$$

Le module carré de cette erreur s'exprime alors comme suit :

$$\begin{aligned}|E_1(\xi, t)|^2 &= |e_1(\xi, t)|^2 + |e_2(\xi, t)|^2 + |e_3(\xi, t)|^2 + 2 \operatorname{Re}(e_1(\xi, t) \overline{e_2(\xi, t)}) \\ &\quad + 2 \operatorname{Re}(e_1(\xi, t) \overline{e_3(\xi, t)}) + 2 \operatorname{Re}(e_2(\xi, t) \overline{e_3(\xi, t)}) .\end{aligned} \quad (\text{A.12})$$

Exprimons chacun des termes de cette somme. Pour cela, on rappelle que le support de h est noté $[-b, b]$. Alors, du fait de l'hypothèse de séparation des composantes des signaux localement harmoniques (1.42), les supports des différents termes formant chacune des sommes sont disjoints dès lors que $b\eta < d/2$, d'où :

$$\begin{aligned}|e_1(\xi, t)|^2 &= \left| \sum_{k=1}^K A'_k(t) e^{2i\pi\phi_k(t)} \frac{1}{\eta} h \left(\frac{\xi - \phi'_k(t)}{\eta} \right) \right|^2 \\ &= \sum_{k=1}^K |A'_k(t)|^2 \frac{1}{\eta^2} \left| h \left(\frac{\xi - \phi'_k(t)}{\eta} \right) \right|^2 \\ &\leq \epsilon^2 c_2^2 \sum_{k=1}^K \frac{1}{\eta^2} \left| h \left(\frac{\xi - \phi'_k(t)}{\eta} \right) \right|^2 ,\end{aligned}$$

car, par définition des signaux localement harmoniques, on a $|A'_k(t)| \leq \epsilon \phi'_k(t)$ et $\phi'_k(t) \leq c_2$ pour tout $t \in \mathbb{R}$ (voir les équations (1.40) et (1.39)). On a alors, pour tout intervalle non vide $I \subset \mathbb{R}$:

$$\iint_{I \times \mathbb{R}} |e_1(\xi, t)|^2 dt d\xi \leq K |I| \epsilon^2 c_2^2 \frac{1}{\eta} J_{0,0,2}. \quad (\text{A.13})$$

On a de même :

$$\begin{aligned} |e_2(\xi, t)|^2 &= 4\pi^2 \sum_{k=1}^K |A_k(t)|^2 |\xi - \phi'_k(t)|^2 \frac{1}{\eta^2} \left| h \left(\frac{\xi - \phi'_k(t)}{\eta} \right) \right|^2 \\ &\leq 4\pi^2 c_2^2 \sum_{k=1}^K |\xi - \phi'_k(t)|^2 \frac{1}{\eta^2} \left| h \left(\frac{\xi - \phi'_k(t)}{\eta} \right) \right|^2, \end{aligned}$$

car, par définition des signaux localement harmoniques, $|A_k(t)| \leq c_2$ pour tout $t \in \mathbb{R}$ (se référer à l'équation (1.38)). On a alors :

$$\iint_{I \times \mathbb{R}} |e_2(\xi, t)|^2 dt d\xi \leq 4\pi^2 K |I| c_2^2 \eta J_{2,0,2}. \quad (\text{A.14})$$

On a également :

$$\begin{aligned} |e_3(\xi, t)|^2 &= \sum_{k=1}^K |A_k(t)|^2 |\phi''_k(t)|^2 \frac{1}{\eta^2} \left| \partial_\xi h \left(\frac{\xi - \phi'_k(t)}{\eta} \right) \right|^2 \\ &\leq \epsilon^2 c_2^4 \sum_{k=1}^K \frac{1}{\eta^2} \left| \partial_\xi h \left(\frac{\xi - \phi'_k(t)}{\eta} \right) \right|^2, \end{aligned}$$

car, par définition des signaux localement harmoniques, $\phi''(t) < \epsilon \phi'(t)$ pour tout $t \in \mathbb{R}$ (se référer à l'équation (1.41)). On a alors :

$$\iint_{I \times \mathbb{R}} |e_3(\xi, t)|^2 dt d\xi \leq K |I| \epsilon^2 c_2^4 \frac{1}{\eta^3} J_{0,1,2}. \quad (\text{A.15})$$

On a également :

$$\begin{aligned} \operatorname{Re} \left(e_1(\xi, t) \overline{e_2(\xi, t)} \right) &= \operatorname{Re} \left(2i\pi \sum_{k=1}^K A'_k(t) A_k(t) \frac{1}{\eta^2} \left| h \left(\frac{\xi - \phi'_k(t)}{\eta} \right) \right|^2 \right) \\ &= 0. \end{aligned}$$

De la même façon, on a :

$$\operatorname{Re} \left(e_2(\xi, t) \overline{e_3(\xi, t)} \right) = 0.$$

D'autre part, on a :

$$\begin{aligned} e_1(\xi, t) \overline{e_3(\xi, t)} &= - \sum_{k=1}^K A_k(t) A'_k(t) \phi''_k(t) \frac{1}{\eta^2} h \left(\frac{\xi - \phi'_k(t)}{\eta} \right) \partial_\xi h \left(\frac{\xi - \phi'_k(t)}{\eta} \right) \\ &= - \sum_{k=1}^K A_k(t) A'_k(t) \phi''_k(t) \frac{1}{2\eta^2} \partial_\xi h^2 \left(\frac{\xi - \phi'_k(t)}{\eta} \right), \end{aligned}$$

On a alors :

$$\begin{aligned} \iint_{I \times \mathbb{R}} 2 \operatorname{Re} \left(e_1(\xi, t) \overline{e_3(\xi, t)} \right) dt d\xi &= - \int_I \sum_{k=1}^K A_k(t) A'_k(t) \phi''_k(t) \\ &\quad \times \frac{1}{\eta^2} \underbrace{\left(\int_{\mathbb{R}} \partial_{\xi} h^2 \left(\frac{\xi - \phi'_k(t)}{\eta} \right) d\xi \right)}_{=[h^2]_{\mathbb{R}}=0 \text{ car } h \in \mathcal{S}(\mathbb{R})} dt \\ &= 0 . \end{aligned}$$

Finalement, pour tout intervalle non vide $I \subset \mathbb{R}$, l'intégrale sur $I \times \mathbb{R}$ des trois termes croisés dans (A.12) est nulle. Alors, en combinant ce résultat avec les expressions des bornes (A.13), (A.14) et (A.15), on a :

$$\begin{aligned} \eta \iint_{I \times \mathbb{R}} |E(\xi, t)|^2 dt d\xi &= \eta \left(\iint_{I \times \mathbb{R}} |e_1(\xi, t)|^2 dt d\xi + \iint_{I \times \mathbb{R}} |e_2(\xi, t)|^2 dt d\xi \right. \\ &\quad \left. + \iint_{I \times \mathbb{R}} |e_3(\xi, t)|^2 dt d\xi \right) \\ &\leq K |I| c_2^2 \left(\epsilon^2 J_{0,0,2} + 4\pi^2 \eta^2 J_{2,0,2} + \frac{c_2^2 \epsilon^2}{\eta^2} J_{0,1,2} \right) . \blacksquare \end{aligned}$$

Démonstration du point 2. Dans le cas où on a affaire à un signal localement harmonique à fréquences rapides, on s'intéresse à l'erreur suivante :

$$E_2(\xi, t) = \partial_t \tilde{G}_y(\xi, t) - 2i\pi \xi \tilde{G}_y(\xi, t) + \Omega(\xi, t) \partial_{\xi} \tilde{G}_y(\xi, t) .$$

Avec les mêmes calculs que précédemment, on obtient :

$$E_2(\xi, t) = e_1(\xi, t) + e_2(\xi, t) ,$$

où e_1 et e_2 sont définis en (A.9) et (A.10). On a donc :

$$|E_2(\xi, t)|^2 = |e_1(\xi, t)|^2 + |e_2(\xi, t)|^2 + 2 \underbrace{\operatorname{Re}(e_1(\xi, t) \overline{e_2(\xi, t)})}_{=0}$$

Donc, en reprenant les calculs précédents, on a :

$$\eta \iint_{I \times \mathbb{R}} |E_2(\xi, t)|^2 dt d\xi \leq K |I| c_2^2 (\epsilon^2 J_{0,0,2} + 4\pi^2 \eta^2 J_{2,0,2}) . \blacksquare$$

RESSOURCES NUMÉRIQUES

L'ensemble des fichiers MATLAB ayant servi à créer les figures et à obtenir les résultats numériques présentés dans ce manuscrit sont disponibles en ligne sur GitHub.

B.1 Ressources relatives aux chapitres 2, 3 et 4

Les fichiers MATLAB permettant de mettre en œuvre les algorithmes 2.2 (JEFAS) et 3.3 (JEFAS-BSS) sont disponibles à l'adresse suivante :

<https://github.com/AdMeynard/JEFAS> .

Ce répertoire contient différents dossiers décrits ci-dessous.

- `cwt` : contient les fonctions permettant de calculer la transformée en ondelettes (et son inverse) à l'aide de l'ondelette *sharp*.
- `JEFASalgo` : contient les fonctions nécessaires à la mise en œuvre de JEFAS et à l'analyse de ses performances.
- `signals` : contient les signaux audios et synthétiques utilisés dans le manuscrit.
- `scriptsIEEE_TASLP` : contient les scripts permettant de réaliser les figures du chapitre 2.
- `JEFAS-BSS` : contient les fonctions permettant la mise en œuvre de JEFAS-BSS.
- `scriptsBSS` : contient les scripts permettant de réaliser les figures du chapitre 3.
- `JEFAS-S` : contient les fonctions permettant la mise en œuvre de JEFAS-S.
- `scriptsJEFAS-S` : contient les scripts permettant de réaliser les figures du chapitre 4.

Les versions publiées des scripts MATLAB, les sons analysés par JEFAS, ainsi que ceux obtenus par synthèse croisée sont disponibles à l'adresse suivante :

<http://meynard.perso.math.cnrs.fr/paperJEFAS/NonStationaryAudio.html> .

B.2 Ressources relatives au chapitre 5

Les fichiers MATLAB permettant de mettre en œuvre l’algorithme TYCOON sont disponibles à l’adresse suivante :

<https://github.com/AdMeynard/TYCOON> .

Ce répertoire contient différents dossiers décrits ci-dessous.

- `AlgorithmTYCOON` : contient l’ensemble de fonctions permettant la mise en œuvre de TYCOON.
- `OtherMethods` : contient les fichiers MATLAB permettant la mise en œuvre du synchrosqueezing (à partir de la TFCT et de la transformée en ondelettes) et de la décomposition modale empirique. Les codes relatifs à l’EMD ont été produits par Zhaohua Wu [83], ceux relatifs au synchrosqueezing par Hau-Tieng Wu [82].
- `SynthSig` : contient les fonctions permettant de synthétiser les signaux utilisés dans le chapitre 5.
- `PerfEvaluation` : contient les fonctions permettant la quantification des performances des représentations temps-fréquence.
- `scriptsTYCOON` : contient les scripts permettant de réaliser les figures et d’obtenir les résultats du chapitre 5.

BIBLIOGRAPHIE

- [1] Benjamin P. ABBOTT et al. (LIGO SCIENTIFIC COLLABORATION AND VIRGO COLLABORATION) : Observation of Gravitational Waves from a Binary Black Hole Merger. *Phys. Rev. Lett.*, 116:061102, février 2016.
- [2] François AUGER et Patrick FLANDRIN : Improving the readability of time-frequency and time-scale representations by the reassignment method. *IEEE Transactions on Signal Processing*, 43(5):1068–1089, mai 1995.
- [3] François AUGER, Patrick FLANDRIN, Yu-Ting LIN, Steve MCLAUGHLIN, Sylvain MEIGNEN, Thomas OBERLIN et Hau-Tieng WU : Time-Frequency Reassignment and Synchrosqueezing : An Overview. *IEEE Signal Processing Magazine*, 30(6):32–41, 2013.
- [4] Amir BECK et Marc TEBoulLE : A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems. *SIAM Journal on Imaging Sciences*, 2(1):183–202, 2009.
- [5] Adel BELOUCHRANI, Karim ABED-MERAIM, Jean-François CARDOSO et Eric MOULINES : A blind source separation technique using second-order statistics. *IEEE Transactions on Signal Processing*, 45(2):434–444, février 1997.
- [6] Adel BELOUCHRANI, Moeness G. AMIN, Nadège THIRION-MOREAU et Yimin D. ZHANG : Source Separation and Localization Using Time-Frequency Distributions : An Overview. *IEEE Signal Processing Magazine*, 30(6):97–107, novembre 2013.
- [7] Jérôme BOLTE, Shoham SABACH et Marc TEBoulLE : Proximal alternating linearized minimization for nonconvex and nonsmooth problems. *Mathematical Programming*, 146(1):459–494, août 2014.
- [8] Joseph-Frédéric BONNANS, Jean Charles GILBERT, Claude LEMARÉCHAL et Claudia A. SAGASTIZÁBAL : *Numerical Optimization : Theoretical and Practical Aspects*. Springer, 2^e édition, 2006.
- [9] René CARMONA, Wen-Liang HWANG et Bruno TORRÉSANI : *Practical time-frequency analysis : Gabor and Wavelet Transforms With an Implementation in S*. Academic Press, 1998.

- [10] René A. CARMONA, Wen-Liang HWANG et Bruno TORRÉSANI : Multiridge detection and time-frequency reconstruction. *IEEE Transactions on Signal Processing*, 47(2):480–492, Feb 1999.
- [11] Antoine CHAMBOLLE et Charles DOSSAL : On the Convergence of the iterates of the “Fast Iterative Shrinkage/Thresholding Algorithm”. *Journal of Optimization Theory and Applications*, 166(3):968–982, septembre 2015.
- [12] Andrzej CICHOCKI et Shun-ichi AMARI : *Adaptive Blind Signal and Image Processing : Learning Algorithms and Applications*. Wiley, juin 2002.
- [13] Maureen CLERC : *Analyse par ondelettes de processus localement dilatés, et application au gradient de texture*. Thèse de doctorat, École Polytechnique, Palaiseau, 1999.
- [14] Maureen CLERC et Stéphane MALLAT : The texture gradient equation for recovering shape from texture. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(4):536–549, 2002.
- [15] Maureen CLERC et Stéphane MALLAT : Estimating deformations of stationary processes. *Ann. Statist.*, 31(6):1772–1821, décembre 2003.
- [16] Pierre COMON et Christian JUTTEN : *Handbook of Blind Source Separation, Independent Component Analysis and Applications*. Academic Press (Elsevier), février 2010.
- [17] Harald CRAMÉR : A Contribution to the Theory of Stochastic Processes. In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, pages 329–339, Berkeley, Calif., 1951. University of California Press.
- [18] Ingrid DAUBECHIES, Jianfeng LU et Hau-Tieng WU : Synchrosqueezed wavelet transforms : An empirical mode decomposition-like tool. *Applied and Computational Harmonic Analysis*, 30(2):243 – 261, 2011.
- [19] Ingrid DAUBECHIES et Stéphane MAES : A nonlinear squeezing of the continuous wavelet transform based on auditory nerve models. In Akram ALDROUBI et Michael UNSER, éditeurs : *Wavelets in Medicine and Biology*, pages 527–546. CRC Press, 1996.
- [20] Manuel DAVY et Simon J. GODSILL : Bayesian harmonic models for musical signal analysis. In José M. BERNARDO, M. J. BAYARRI, James O. BERGER, A. Philip DAWID, David HECKERMANN, Adrian F. M. SMITH et Mike WEST, éditeurs : *Bayesian Statistics 7 : Proceedings of the Seventh Valencia International*, 2003.
- [21] Arnaud DELORME et Scott MAKEIG : EEGLAB : an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1):9 – 21, 2004.
- [22] Nathalie DELPRAT : *Analyse temps-fréquence de sons musicaux : exploration d’une nouvelle méthode d’extraction de données pertinentes pour un modèle de synthèse*. Thèse de doctorat, Université d’Aix-Marseille II, avril 1992.
- [23] Nathalie DELPRAT, Bernard ESCUDIÉ, Philippe GUILLEMAIN, Richard KRONLAND-MARTINET, Philippe TCHAMITCHIAN et Bruno TORRÉSANI :

- Asymptotic wavelet and Gabor analysis : extraction of instantaneous frequencies. *IEEE Transactions on Information Theory*, 38(2):644–664, mars 1992.
- [24] Arthur P. DEMPSTER, Nan M. LAIRD et Donald B. RUBIN : Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, 39(1):1–38, 1977.
- [25] Patrick FLANDRIN : *Time-Frequency/Time-Scale Analysis*. Academic Press, 1999.
- [26] Patrick FLANDRIN : *Explorations in Time-Frequency Analysis*. Cambridge University Press, 2018.
- [27] Gerald B. FOLLAND et Alladi SITARAM : The uncertainty principle : A mathematical survey. *Journal of Fourier Analysis and Applications*, 3(3):207–238, mai 1997.
- [28] Rémi GRIBONVAL et Michael ZIBULEVSKY : Sparse Component Analysis. In COMMON, PIERRE, JUTTEN et CHRISTIAN, éditeurs : *Handbook of Blind Source Separation, Independent Component Analysis and Applications*, pages 367–420. Academic Press, 2010.
- [29] Karlheinz GRÖCHENIG : *Foundations of time-frequency analysis*. Applied and Numerical Harmonic Analysis. Birkhäuser Inc., Boston, MA, 2001.
- [30] Jonathan HARRINGTON et Steve CASSIDY : *Techniques in Speech Acoustics*, volume 8 de *Text, Speech and Language Technology*. Springer Netherlands, 1999.
- [31] Harold V. HENDERSON et Shayle R. SEARLE : On Deriving the Inverse of a Sum of Matrices. *SIAM Review*, 23(1):53–60, 1981.
- [32] Norden E. HUANG, Zheng SHEN, Steven R. LONG, Manli C. WU, Hsing H. SHIH, Quanan ZHENG, Nai-Chyuan YEN, Chi Chao TUNG et Henry H. LIU : The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London. Series A : Mathematical, Physical and Engineering Sciences*, 454(1971):903–995, 1998.
- [33] Ran KAFTORY et Yehoshua Y. ZEEVI : Blind Separation of Time/Position Varying Mixtures. *IEEE Transactions on Image Processing*, 22(1):104–118, janvier 2013.
- [34] Kunihiko KODERA, Claude De VILLEDARY et Roger GENDRIN : A new method for the numerical analysis of non-stationary signals. *Physics of the Earth and Planetary Interiors*, 12(2):142 – 150, 1976.
- [35] Matthieu KOWALSKI : Sparse Regression Using Mixed Norms. *Applied and Computational Harmonic Analysis*, 27(3):303–324, septembre 2009.
- [36] Matthieu KOWALSKI, Adrien MEYNARD et Hau-tieng WU : Convex Optimization approach to signals with fast varying instantaneous frequency. *Applied and Computational Harmonic Analysis*, 44(1):89 – 122, 2018.
- [37] Jonathan M. LILLY et Sofia C. OLHEDE : Higher-Order Properties of Analytic Wavelets. *IEEE Transactions on Signal Processing*, 57(1):146–160, janvier 2009.
- [38] Stéphane MALLAT : *A Wavelet Tour of Signal Processing (Second Edition)*. Academic Press, San Diego, 1999.

- [39] Stéphane MALLAT, George PAPANICOLAOU et Zhifeng ZHANG : Adaptive covariance estimation of locally stationary processes. *Ann. Statist.*, 26(1):1–47, février 1998.
- [40] Danilo P. MANDIC, Naveed ur REHMAN, Zhaohua WU et Norden E. HUANG : Empirical Mode Decomposition-Based Time-Frequency Analysis of Multivariate Signals : The Power of Adaptive Data Analysis. *IEEE Signal Processing Magazine*, 30(6):74–86, novembre 2013.
- [41] Gerald MATZ, Franz HLAWATSCH et Werner KOZEK : Generalized evolutionary spectral analysis and the Weyl spectrum of nonstationary random processes. *IEEE Transactions on Signal Processing*, 45(6):1520–1534, juin 1997.
- [42] Robert J. MCAULAY et Thomas F. QUATIERI : Speech analysis/synthesis based on a sinusoidal representation. *IEEE Transactions on Acoustics, Speech, Signal Processing*, 34:744–754, août 1986.
- [43] Adrien MEYNARD : Joint nonstationary blind source separation and spectral analysis. In *iTWIST : international Traveling Workshop on Interactions between low-complexity data models and Sensing Techniques*, Proceedings of iTWIST’18, Marseille, France, novembre 2018.
- [44] Adrien MEYNARD : Séparation de sources doublement non stationnaire. In *GRETSI 2019 - XXVIIème Colloque francophone de traitement du signal et des images*, Lille, France, août 2019.
- [45] Adrien MEYNARD et Bruno TORRÉSANI : Spectral estimation for non-stationary signal classes. In *Sampling Theory and Applications*, Proceedings of SampTA17, Tallinn, Estonia, juillet 2017.
- [46] Adrien MEYNARD et Bruno TORRÉSANI : Spectral Analysis for Nonstationary Audio. *IEEE/ACM Transactions on Audio, Speech and Language Processing*, 26(12):2371 – 2380, décembre 2018.
- [47] Eric MOREAU et Odile MACCHI : A one stage self-adaptive algorithm for source separation. In *1994 IEEE International Conference on Acoustics, Speech and Signal Processing. Proceedings of ICASSP ’94.*, volume iii, pages III/49–III/52, avril 1994.
- [48] Eric MOULINES et Jean LAROCHE : Non-parametric techniques for pitch-scale and time-scale modification of speech. *Speech Communication*, 16(2):175 – 205, 1995. Voice Conversion : State of the Art and Perspectives.
- [49] Thomas OBERLIN : *Analyse de Signaux Multicomposantes : Contributions à la Décomposition Modale Empirique, aux Représentations temps-fréquence et au Synchronqueezing*. Thèse de doctorat, Université de Grenoble, novembre 2013.
- [50] Thomas OBERLIN, Sylvain MEIGNEN et Valérie PERRIER : Second-order synchronqueezing transform or invertible reassignment ? Towards ideal time-frequency representations. *IEEE Transactions on Signal Processing*, 63(5):1335–1344, mars 2015.
- [51] Harold OMER : *Modèles de déformation de processus stochastiques généralisés. Application à l’estimation des non stationnarités dans les signaux audio*. Thèse de doctorat, Aix-Marseille Université, 2015.

- [52] Harold OMER et Bruno TORRÉSANI : Estimation of frequency modulations on wideband signals ; applications to audio signal analysis. In Goetz PFANDER, éditeur : *Proceedings of the 10th International Conference on Sampling Theory and Applications (SampTA)*, pages 29–32. Eurasip Open Library, 2013.
- [53] Harold OMER et Bruno TORRÉSANI : Time-frequency and time-scale analysis of deformed stationary processes, with application to non-stationary sound modeling. *Applied and Computational Harmonic Analysis*, 43(1):1 – 22, 2017.
- [54] Morgane PAPIN, Julian PICHENOT, François GUÉROLD et Estelle GERMAIN : Acoustic localization at large scales : a promising method for grey wolf monitoring. *Frontiers in Zoology*, 15(1):11, avril 2018.
- [55] Lucas PARRA et Clay SPENCE : Convolutional blind separation of non-stationary sources. *IEEE Transactions on Speech and Audio Processing*, 8(3):320–327, mai 2000.
- [56] Michael Syskind PEDERSEN, Jan LARSEN, Ulrik KJEMS et Lucas C. PARRA : *Convolutional Blind Source Separation Methods*, chapitre 52, pages 1065–1094. Springer, Berlin, Heidelberg, 2008.
- [57] Kaare Brandt PETERSEN et Michael Syskind PEDERSEN : *The Matrix Cookbook*, novembre 2012. Version 2012-11-15.
- [58] Dinh-Tuan PHAM et Jean-François CARDOSO : Blind separation of instantaneous mixtures of nonstationary sources. *IEEE Transactions on Signal Processing*, 49(9): 1837–1848, septembre 2001.
- [59] Bernard PICINBONO : On circularity. *IEEE Transactions on Signal Processing*, 42(12): 3473–3482, 1994.
- [60] Bernard PICINBONO : Second-order complex random vectors and normal distributions. *IEEE Transactions on Signal Processing*, 44(10):2637–2640, 1996.
- [61] Maurice B. PRIESTLEY : *Spectral analysis and time series*. Probability and mathematical statistics. Academic Press, 1982.
- [62] Maurice B. PRIESTLEY : *Non-linear and non-stationary time series analysis*. Academic Press, 1988.
- [63] Ramon E. PRIETO et Pamornpol JINACHITRA : Blind source separation for time-variant mixing systems using piecewise linear approximations. In *2005 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings. (ICASSP '05).*, volume 5, pages v/301–v/304, mars 2005.
- [64] Benjamin RICAUD et Bruno TORRÉSANI : A survey of uncertainty principles and some signal processing applications. *Advances in Computational Mathematics*, 40(3): 629–650, 2014.
- [65] John L. SEMMLOW et Benjamin GRIFFEL : *Biosignal and Medical Image Processing*. CRC Press, 3^e édition, février 2014.
- [66] Richard A. SILVERMAN : Locally stationary random processes. *IRE Transactions on Information Theory*, 3(3):182–187, septembre 1957.

- [67] Christian SOIZE : *Méthodes mathématiques en analyse du signal*. Masson, Paris, 1993.
- [68] Petre STOICA et Randolph MOSES : *Spectral analysis of signals*. Pearson/Prentice Hall, Upper Saddle River, N.J., 2005.
- [69] Dan STOWELL : *Computational Analysis of Sound Scenes and Events*, chapitre Computational Bioacoustic Scene Analysis, pages 303–333. Springer, 2018.
- [70] Nadège THIRION-MOREAU et Moeness G. AMIN : Chapter 11 - quadratic time-frequency domain methods. In Pierre COMON et Christian JUTTEN, éditeurs : *Handbook of Blind Source Separation*, pages 421 – 466. Academic Press, Oxford, 2010.
- [71] Robert TIBSHIRANI : Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 58(1):267–288, 1996.
- [72] Lang TONG, Victor C. SOON, Yih-Fang HUANG et Ruey-Wen LIU : AMUSE : a new blind identification algorithm. In *IEEE International Symposium on Circuits and Systems*, pages 1784–1787 vol.3, mai 1990.
- [73] Richard E. TURNER et Maneesh SAHANI : Time-Frequency Analysis as Probabilistic Inference. *IEEE Transactions on Signal Processing*, 62(23):6171–6183, décembre 2014.
- [74] Martin VETTERLI, Jelena KOVAČEVIĆ et Vivek K GOYAL : *Foundations of Signal Processing*. Cambridge University Press, 2014.
- [75] Cédric VILLANI : *Topics in Optimal Transportation*, volume 58 de *Graduate Studies in Mathematics*. American Mathematical Society, Providence, Rhodes Island, 2003.
- [76] Emmanuel VINCENT : *Contributions to audio source separation and content description*. Habilitation à diriger des recherches, Université Rennes 1, novembre 2012.
- [77] Emmanuel VINCENT, Rémi GRIBONVAL et Cédric FÉVOTTE : Performance measurement in blind audio source separation. *IEEE Transactions on Audio, Speech, and Language Processing*, 14(4):1462–1469, juillet 2006.
- [78] Tuomas VIRTANEN : Monaural sound source separation by nonnegative matrix factorization with temporal continuity and sparseness criteria. *IEEE Transactions on Audio, Speech, and Language Processing*, 15(3):1066–1074, mars 2007.
- [79] Peter D. WELCH : The use of fast Fourier transform for the estimation of power spectra : A method based on time averaging over short, modified periodograms. *IEEE Transactions on Audio and Electroacoustics*, 15(2):70–73, juin 1967.
- [80] Scott T. WISDOM : *Improving and Unfolding Statistical Models of Nonstationary Signals*. Thèse de doctorat, Department of Electrical Engineering, University of Washington, Seattle, WA, août 2017.
- [81] Scott T. WISDOM, Les ATLAS et James PITTON : Extending coherence time for analysis of modulated random processes. In *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 340–344, mai 2014.
- [82] Hau-Tieng WU, Yi-Hsin CHAN, Yu-Ting LIN et Yung-Hsin YEH : Using synchrosqueezing transform to discover breathing dynamics from ECG signals. *Applied and Computational Harmonic Analysis*, 36(2):354 – 359, 2014.

- [83] Zhaohua WU et Norden E. HUANG : Ensemble Empirical Mode Decomposition : A Noise-Assisted Data Analysis Method. *Advances in Adaptive Data Analysis*, 01(01): 1–41, 2009.

Résumé

La non-stationnarité est caractéristique des phénomènes physiques transitoires. Par exemple, elle peut être engendrée par la variation de vitesse d'un moteur lors d'une accélération. De même, du fait de l'effet Doppler, un son stationnaire émis par une source en mouvement sera perçu comme étant non stationnaire par un observateur fixe. Ces exemples nous conduisent à considérer une classe de non-stationnarité formée des signaux stationnaires dont la stationnarité a été brisée par un opérateur de déformation physiquement pertinent. Après avoir décrit les modèles de déformation considérés (chapitre 1), nous présentons différentes méthodes permettant d'étendre l'analyse et la synthèse spectrale à de tels signaux. L'estimation spectrale des signaux revient à déterminer le spectre du processus stationnaire sous-jacent et la déformation ayant brisé sa stationnarité.

Ainsi, dans le chapitre 2, nous nous intéressons à l'analyse de signaux localement déformés pour lesquels la déformation subie s'exprime simplement comme un déplacement des coefficients d'ondelettes dans le plan temps-échelle. Nous tirons profit de cette propriété pour proposer l'algorithme d'estimation du spectre instantané JEFAS. Dans le chapitre 3, nous étendons cette analyse spectrale aux signaux multi-capteurs pour lesquels l'opérateur de déformation prend une forme matricielle. Il s'agit d'un problème de séparation de sources doublement non stationnaire. Dans le chapitre 4, nous proposons une approche à la synthèse pour étudier des signaux localement déformés. Enfin, dans le chapitre 5, nous construisons une représentation temps-fréquence adaptée à l'étude des signaux localement harmoniques.

Mots-clés : non-stationnarité, spectre instantané, déformation, temps-échelle, temps-fréquence, signaux audio.

Abstract

Nonstationarity characterizes transient physical phenomena. For example, it may be caused by a speed variation of an accelerating engine. Similarly, because of the Doppler effect, a stationary sound emitted by a moving source is perceived as being nonstationary by a motionless observer. These examples lead us to consider a class of nonstationary signals formed from stationary signals whose stationarity has been broken by a physically relevant deformation operator. After describing the considered deformation models (Chapter 1), we present different methods that extend the spectral analysis and synthesis to such signals. The spectral estimation amounts to determining simultaneously the spectrum of the underlying stationary process and the deformation breaking its stationarity. To this end, we consider representations of the signal in which this deformation is characterized by a simple operation.

Thus, in Chapter 2, we are interested in the analysis of locally deformed signals. The deformation describing these signals is simply expressed as a displacement of the wavelet coefficients in the time-scale domain. We take advantage of this property to develop a method for the estimation of these displacements. Then, we propose an instantaneous spectrum estimation algorithm, named JEFAS. In Chapter 3, we extend this spectral analysis to multi-sensor signals where the deformation operator takes a matrix form. This is a doubly nonstationary blind source separation problem. In Chapter 4, we propose a synthesis approach to study locally deformed signals. Finally, in Chapter 5, we construct a time-frequency representation adapted to the description of locally harmonic signals.

Keywords: nonstationarity, instantaneous spectrum, deformation, time-scale, time-frequency, audio signals.