



HAL
open science

Processus alpha-stables pour le traitement du signal

Mathieu Fontaine

► **To cite this version:**

Mathieu Fontaine. Processus alpha-stables pour le traitement du signal. Traitement du signal et de l'image [eess.SP]. Université de Lorraine, 2019. Français. NNT : 2019LORR0037 . tel-02188304

HAL Id: tel-02188304

<https://theses.hal.science/tel-02188304>

Submitted on 18 Jul 2019

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

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

Processus α -stables pour le traitement du signal

THÈSE

présentée et soutenue publiquement le 12 juin 2019

pour l'obtention du

Doctorat de l'Université de Lorraine
(mention Informatique)

par

Mathieu Fontaine

Composition du jury

Rapporteurs : **Rémi GRIBONVAL**, Directeur de Recherche, Inria Rennes
Hermine BIERMÉ, Professeure, Université de Poitiers

Examineurs : **Antoine LEJAY**, Directeur de Recherche, Inria Nancy Grand-Est
Angélique DRÉMEAU, Enseignante Chercheuse, ENSTA Bretagne

Encadrants : **Roland BADEAU**, Professeur, Télécom ParisTech
Antoine LIUTKUS, Chargé de Recherche, Inria Sophia Antipolis Méditerranée

Mis en page avec la classe thesul.

Remerciements

Je tiens tout d'abord à remercier très chaleureusement mes encadrants Antoine et Roland, pour m'avoir permis de découvrir le sujet du traitement du signal audio, de l'enseignement dans le supérieur et avec qui j'ai appris énormément durant ces trois années. Merci aux rapporteurs et examinateurs pour leurs conseils et avis sur mon sujet, dans le but de sublimer ce mémoire de thèse. Je ne peux évidemment pas oublier toutes les équipes avec lesquelles j'ai pu travailler. Notamment, je remercie pour MULTISPEECH à Nancy Baldwin, Ken, Diego, Antoine, Aman, Arie, Sunit, Nicolas, Matthieu, Romain, Emmanuel, Hélène mais également Denis pour ses précieux conseils. Également à Télécom ParisTech, je remercie Umut, Simon, Victor, Paul, Magdalena, Laurent, Bertrand, Pascal et Gaël pour toutes les collaborations sur Paris et fabuleux moments passés ensemble. Merci à l'équipe ZENITH de Montpellier en particulier Fabian, Titouan, Gaëtan, Benjamin, Valentin, Antoine, Jean-Christophe, Alexis, Renan, Nathalie et Patrick pour m'avoir accueilli. Je tiens par ailleurs à remercier Kazuyoshi de l'Université de Kyoto et toute son équipe du traitement du signal pour me donner la merveilleuse opportunité d'effectuer un post-doctorat au Japon. Un grand merci par ailleurs à Charles et Laurent Girin avec qui j'ai pu collaborer sur plusieurs articles de conférences et dont l'échange de connaissance fût superbe.

Plus personnellement, je ne peux pas oublier mes acolytes musiciens de Presumption avec qui j'ai brûlé les planches de nombreux scènes en Europe : Marvin, mon très cher frère Anaël et Adrien. Également mes collègues musiciens que j'ai pu rencontré jusqu'à maintenant et qui représente une très grande partie de ma vie : Aymerick, Ronan, Stéphane I,II & III, Jérémie, Thomas, Rodolphe, Jérôme, Emeric, Guillaume I & II, Benjamin, Pierre, Rémi, Marco, Cédric, Serge, Jean-Bernard, Romain, Bastien, Nicolas, Emile, Adrien, Yoann, Victor, David et tous ceux que j'ai oublié. Également, à mes très chers amis de l'ouest : Antoine, Morane, Guillaume, Richard, Gaëlle, Brigitte, Patrice, Nikita, Olivier, Paco, Lydia, Brigitte, Stéphane, Nathalie, Dominique, Audric, Emilie, Jean-Pierre, Bruno, Fabien, Mathieu, Clément, Thomas, Mélissa, Bastien, Arnold, Pauline, Simon, Marie, Mélanie, Carolanne, Tony, Sylvain, Krisbet et Vivien. Un très grand merci à mes charmants colocataires Lilian, Céline, Cosy, Davai et leurs amis avec qui j'ai passé de merveilleux moments. En vrac, je remercie également Camille et Patricia pour m'avoir accueilli pendant mes missions sur Paris, Cheryl, Claudia et les petits pour les bons moments. Une pensée à mes amis Jean-Luc, Cyprien, Marcel, Jeannot, Thomas, Arnaud et Didier mais également à mon grand-père Gustave. Enfin, je terminerai par remercier ma soeur Laëtitia, mon frère Anaël, mes neveux Hugo et Cyprien pour me supporter depuis tout ce temps et à mes parents Ghislaine et Daniel qui m'ont toujours encouragé dans mes études faites et sur les différents chemins que j'ai pu emprunter.

*Je dédie cette thèse à mes parents, mon frère et ma sœur.
Également à certains amis partis trop tôt.*

Résumé

Le sujet scientifique de la *séparation de sources sonores* (SSS) vise à décomposer les signaux audio en leurs éléments constitutants, par exemple en séparant la voix du chanteur principal de son accompagnement musical ou du bruit de fond. Dans le cas d'enregistrements historiques très anciens et très dégradés, la SSS étend de manière significative les méthodes classiques de débruitage en permettant de prendre en compte des motifs complexes de signal et de bruit et de réaliser efficacement la séparation, là où les approches traditionnelles sont tenues en échec.

En traitement du signal audio, le signal observé est souvent supposé être égal à la somme des signaux que nous souhaitons obtenir. Dans le cadre d'une modélisation probabiliste, il est alors primordial que les processus stochastiques préservent leur loi par sommation. Le processus le plus employé et vérifiant cette stabilité est le processus gaussien. Comparé aux autres processus α -stables vérifiant la même stabilité, les processus gaussiens ont la particularité d'admettre des outils statistiques facilement interprétables comme la moyenne et la covariance. L'existence de ces moments permet d'esquisser des méthodes statistiques en SSS et plus généralement, en traitement du signal. La faiblesse de ces processus réside néanmoins dans l'incapacité à s'écarter trop loin de leurs moyennes. Cela limite la dynamique des signaux modélisables et peut provoquer des instabilités dans les méthodes d'inférence considérées.

Les processus α -stables non-gaussiens soulèvent des défis mathématiques et ont déjà démontré leur efficacité dans des applications de filtrage et en terme algorithmique. En dépit de non-existence d'une forme analytique des densités de probabilités, les processus α -stables jouissent de résultats non valables dans le cas gaussien. Par exemple, un vecteur α -stable non-gaussien admet une représentation spatiale unique. En résumé, le comportement d'une distribution multivariée α -stable est contrôlé par deux opérateurs. Une mesure dite «spectrale» informant sur l'énergie globale venant de chaque direction de l'espace et un vecteur localisant le centroïde de sa densité de probabilité. Cette représentation spatiale a notamment montré son efficacité dans le cas de la SSS pour la célèbre *analyse en composantes indépendantes* (ACI). Les modèles pour cette ACI α -stable ne sont cependant proposés que dans le cas de mélanges linéaires instantanés. Par conséquent, l'information dans le domaine fréquentiel est omise pour ce type de mélange.

Ce mémoire de thèse introduit différents modèles α -stables d'un point de vue théorique et les développe dans plusieurs directions. Nous proposons notamment une extension de la théorie de filtrage α -stable monocanal au cas multicanal. En particulier, une nouvelle représentation spatiale pour les vecteurs α -stables est adoptée. Nous développons en outre un modèle de débruitage où le bruit et la parole découlent de distributions α -stables mais ayant un exposant caractéristique α différent. La valeur d' α permet de contrôler la stationnarité de chaque source. Grâce à ce modèle hybride, nous avons également déduit une explication rigoureuse sur des filtres de Wiener heuristiques esquissés dans les années 80. Une autre partie de ce manuscrit décrit en outre comment la théorie α -stable permet de fournir une méthode pour la localisation de sources sonores. Pour ce faire, nous employons la représentation spatiale d'un vecteur α -stable non-gaussien afin d'exploiter sa mesure spectrale. En pratique, elle nous permet d'en déduire si une source est active à un endroit précis de l'espace.

Au final, nos travaux ont consisté à étudier la théorie α -stable pour le traitement du signal. Nous avons abouti à de nombreux modèles pour un large panel d'applications. En dehors des quelques applications que nous avons déjà considérées, nous pouvons étendre les modèles multicanaux α -stables à la déréverbération ou à la SSS. Quant au modèle sur la localisation des sources sonores, il pourrait être employé afin de déterminer la géométrie d'une salle.

Mots-clés: α -stable, probabilité, traitement du signal, localisation, débruitage

Abstract

The scientific topic of *sound source separation* (SSS) aims at decomposing audio signals into their constitutive components, e.g. separate singing voice from background music or from background noise. In the case of very old and degraded historical recordings, SSS strongly extends classical denoising methods by being able to account for complex signal or noise patterns and achieve efficient separation where traditional approaches fail.

It is classic in signal processing to model the observed signal as the sum of desired signals. If we adopt a probabilistic model, it is preferable that law of the additive processes is stable by summation. The Gaussian process notoriously satisfies this condition. It admits useful statistical operators as the covariance and the mean. The existence of those moments allows to provide a statistical model for SSS. However, Gaussian process has difficulty to deviate from its mean. This drawback limits signal dynamics and may cause unstable inference methods. On the contrary, non-Gaussian α -stable processes are stable under addition, and permit the modeling of signals with considerable dynamics.

For the last few decades, α -stable theory have raised mathematical challenges and have already been shown to be effective in filtering applications. This class of processes enjoys outstanding properties, not available in the Gaussian case. A major asset for signal processing is the unique spatial representation of a multivariate α -stable vector, controlled by a so-called spectral measure and a deterministic vector. The spectral measure provides information on the global energy coming from all space directions while the vector localizes the centroid of the probability density function. It shows its usefulness for the so-called *independent component analysis* (ICA) topic and the SSS. However, those models are only linear instantaneous mixture and do not consider the frequency domain.

This thesis introduces several α -stable models, with the aim of extending them in several directions. First, we propose an extension of single-channel α -stable filtering theory to a multichannel one. In particular, a novel spatial representation for α -stable vectors is proposed. Secondly, we develop α -stable models for denoising where each component could admit a different α . This hybrid model provides a rigorous explanation of some heuristic Wiener filters outlined in the 1980s. We also describe how the α -stable theory yields a new method for audio source localization. We use the spectral measure resulting from the spatial representation of α -stable vectors. In practice, it leads to determine whether a source is active at a specific location.

Our work consisted in investigating the α -stable theory for signal processing and developing several models for a wide range of applications. The models introduced in this thesis could also be extend to more signal processing tasks. We could use our multivariate α -stable models to dereverberation or SSS. Moreover, the localization algorithm is implementable for room geometry estimation.

Keywords: α -stable, probability, signal processing, localization, denoising

Table des matières

Remerciements	i
Résumé	v
Abstract	vi
Table des figures	xiii
Liste des tableaux	xv
Acronymes	xvii
Notations	xix
Introduction générale	xxi
Partie I Contextes de modèles probabilistes pour la séparation et la localisation de sources sonores	1
1 Premières définitions	3
1.1 Théorie probabiliste	3
1.1.1 Fonction de répartition et densité de probabilité	3
1.1.2 Espérance, moments et fonction caractéristique	4
1.1.3 Indépendance	5
1.1.4 Conditionnalité et inférence Bayésienne	5
1.2 Processus stochastiques	6
1.2.1 Définitions et propriétés	6
1.2.2 Mesures aléatoires et représentation stochastique	7
2 Processus α-stables	9
2.1 Distributions α -stables	9

2.1.1	Contexte historique	9
2.1.2	Distributions α -stables symétriques réelles	9
2.1.3	Exposant de Lévy: une caractérisation de la parcimonie	12
2.1.4	Distributions α -stables symétriques complexes isotropes	14
2.2	Processus gaussiens	16
2.2.1	Processus gaussiens & stationnaires au sens large	16
2.2.2	Analyse spectrale dans le cas SSL	18
2.3	Processus α -stables	19
2.3.1	Processus symétrique α -stable	20
2.3.2	Processus α -harmonisables	21
3	Mélanges et séparation de processus α-stables	23
3.1	Séparation de processus gaussiens	23
3.1.1	Cas général	24
3.1.2	Cas stationnaire au sens large	24
3.1.3	Séparation de variables α -stables complexes isotropes	25
3.1.4	Le cas α -harmonisable et DSP fractionnaire	26
4	Séparation de sources sonores : estimation des paramètres	29
4.1	Paradigme pour la séparation de sources sonores	29
4.2	Cas gaussien : divergence d'Itakura-Saito	30
4.2.1	Maximum de vraisemblance et algorithme d'espérance-maximisation	30
4.2.2	Modèle de factorisation matricielle non-négative pour la DSP	33
4.3	Le cas Cauchy	35
4.3.1	Estimation des paramètres par les règles de mises à jour multiplicatives	36
4.3.2	Technique de majoration-égalisation	37
4.4	Cas Général	38
4.4.1	Méthode des Moments d'ordres faibles fractionnaires (MOFF)	38
4.4.2	Gaussianité conditionnelle et algorithme de Métropolis-Hastings	39
5	Méthodes et modèles probabilistes pour la localisation de sources acoustiques	41
5.1	TDOAs exploitant les moments d'ordre 2	41
5.2	TDOAs dans le cas α -stable stéréophonique	43
Partie II	Mélanges discrets et application à la localisation de sources audio	45
6	Mélanges multivariés linéaires de sources α-stables scalaires	47
6.1	Modèle et notations	47

6.2	Fonction caractéristique et mesure spatiale discrète (MSD)	48
6.3	Utilisation de la MSD en traitement du signal audio: historique	50
7	Estimation de la MSD	53
7.1	Introduction	53
7.2	Estimation de l'exposant de Lévy	55
7.2.1	Estimation « naïve »	55
7.2.2	Estimation dans le cas symétrique α -stable	55
7.2.3	Problème d'optimisation	56
7.2.4	Mise à jour des paramètres	57
8	Application à la localisation de sources acoustiques	59
8.1	Modèle pour la localisation	59
8.1.1	Positionnement du problème et hypothèses	59
8.1.2	Modèle probabiliste	60
8.1.3	Modèle acoustique	62
8.1.3.1	Réponses impulsionnelles de salle	62
8.1.3.2	Modèle de champ proche	63
8.2	Expériences	64
8.2.1	Méthodes de références	64
8.2.2	Résultats avec un modèle de champ proche	64
8.2.3	Résultats avec un modèle de réponse acoustique de salle	66
	Conclusion de la Partie II	69
	Partie III Filtrage de processus α-stables vectoriels	71
9	Représentation spatiale de vecteurs α-stables	73
9.1	Distribution multivariée α -stable	73
9.1.1	Premier théorème de représentation	73
9.1.2	Propriétés et motivations	74
9.2	Théorème de représentation spatiale	75
9.2.1	Spectre spatial	75
9.2.2	Énoncé et démonstration	75
9.2.3	Interprétation du théorème de représentation	77
10	Filtrage de mélange α-stables vectoriels: cas général	81
10.1	Modèle de mélange	81
10.2	Estimation du spectre spatial	82

10.2.1	Critère d'estimation	82
10.2.2	Continuité de g_χ	84
10.3	Simplification de g_χ	85
10.3.1	Motivations et description	85
10.3.2	Modélisation de séries numériques	87
10.4	Algorithme de Séparation	88
10.4.1	Reconstruction des sources images	88
10.4.2	Résumé de la méthode de filtrage	90
11	Filtrage linéaire α-stable: minimisation de la norme de covariation	91
11.1	Introduction: un nouvel estimateur linéaire	91
11.2	Covariation	92
11.2.1	Définitions, exemples et propriétés	92
11.2.2	Norme de covariation	94
11.3	Problème d'optimisation	95
11.3.1	Filtre linéaire et existence d'une solution	95
11.3.2	Résolution du problème	96
11.3.3	Cas limite quand $\alpha \rightarrow 2$	97
12	Évaluation des performances	99
12.1	Méthode de filtrage, hypothèses et score d'erreur	99
12.2	Écart angulaire et pouvoir séparateur	100
12.3	Qualité de séparation en fonction du nombre de composantes	102
12.4	Score de séparation en fonction de la complexité du modèle spatial	103
	Conclusion de la partie III	105
Partie IV	Modèles hybrides et applications au rehaussement de la parole	107
13	Modèles multi-alphas	109
13.1	Contexte	109
13.2	Modèles de mélange monocanal	111
13.2.1	Exemples de filtre de Wiener paramétrique	111
13.2.2	Modèle de filtrage α -stable monophonique hybride	112
13.3	Modèle de mélange multicanal	113
13.3.1	Distribution multivariée α -stable à contour elliptique	113
13.3.2	Modèle de filtrage multicanal: gaussien + α -stable	115

14 Multi-Alpha Denoising (MAD): application au débruitage monocanal	117
14.1 Estimation des variables impulsionnelles	117
14.1.1 Présentation de l'estimation	117
14.1.2 Évaluation empirique des masques de Wiener paramétriques	118
14.2 Estimation des paramètres d'échelles	119
14.2.1 Exploitation des moments logarithmiques	119
14.2.2 Lissage des paramètres	120
14.3 Résumé de l'algorithme MAD	120
14.4 Évaluation	121
14.4.1 Description de l'expérience	121
14.4.2 Résultats et commentaires	122
14.4.3 Restauration audio du patrimoine ethnomusicologique	122
15 Gaussien + Elliptique: application au débruitage multicanal	125
15.1 Estimation des paramètres	125
15.1.1 Espérance-Maximisation	125
15.1.1.1 Étape-E	126
15.1.1.2 Étape-M	127
15.1.2 Algorithme de Metropolis-Hastings	127
15.1.3 Estimation des paramètres σ	128
15.1.4 Reconstruction monocanal : Maximisation de l'énergie	128
15.1.5 Algorithme Alpha Residual Component (ARC)	129
15.2 Évaluation	130
15.2.1 Configuration et état de l'art	130
15.2.2 Mesure de performance	130
15.2.3 Première expérience: résultats généraux	131
15.2.4 Deuxième expérience: score de rehaussement de la parole en milieu réverbérant	132
Conclusion de la partie IV	135
Conclusion générale	137
Index	143
Bibliographie	145

Table des figures

1	Comportement asymptotique de la queue à droite de trois densités de probabilités. La gaussienne ($\alpha = 2$) est qualifiée de queue légère car elle est située en dessous de la loi exponentielle. Se localisant au-dessus de l'exponentielle, la distribution de Cauchy ($\alpha = 1$) est quand à elle dite à queue lourde.	xxii
2	Structure du manuscrit. Les flèches désignent la dépendance entre chaque chapitre ou partie. Les chiffres entre crochets font référence aux articles publiés durant cette thèse (voir ci-dessous).	xxiv
2.1	À gauche, densités de probabilités de variables $S\alpha S(\sigma^\alpha)$ avec $\sigma^\alpha = 1$. À droite, densités de probabilités de variables $S\alpha(\sigma^\alpha, \beta, \mu)$ avec $\alpha = 0.9$, $\sigma^\alpha = 1$ et $\mu = 2$	12
2.2	Réalisation de distributions scalaires $S\alpha S$: plus α est petit, plus l'échantillonnage devient « parcimonieux ».	13
2.3	Densités de probabilité de lois $\mathcal{P}\alpha S$	15
2.4	Réalisations d'un processus gaussien avec la covariance décrite en (2.21) pour différents λ : plus λ augmente, plus le signal semble corrélé.	17
2.5	log-spectrogramme d'un signal de parole.	19
4.1	Paradigme classique pour la séparation des sources sonores dans un cadre Bayésien.	30
4.2	Algorithme d'espérance-maximisation en SSS.	33
4.3	Log-spectrogramme de la note de DO ₃ d'un piano droit.	33
4.4	Factorisation matricielle non-négative d'un signal de parole bruité.	34
5.1	Configuration de deux microphones m_1, m_2 et d'une source s_l localisée en l	42
6.1	Estimation de la mesure spectrale pour un mélange linéaire instantanée.	50
7.1	Stratégie d'estimation de la MSD	53
7.2	courbe des β -divergences $v \mapsto d_\beta(1 v)$	57
8.1	Modèle de localisation et identification avec la MSD	60
8.2	Réponse impulsionnelle de salle pour un temps de réverbération (avec diminution du niveau sonore de 60 dB) de 64 ms.	63
8.3	Représentation de la MSD Υ , de la PRD et des algorithmes CLEAN/RELAX. La vraie position des sources sur le plan est annotée par des triangles blancs.	65
8.4	Scores de corrélations pour chaque algorithme. Les moustaches noires affichent la variance en fonction du nombre de microphones.	66
8.5	Scores de corrélations pour chaque algorithme. Les moustaches noires affichent la variance en fonction du nombre de microphones.	67

8.6	MSD (haut) vs PRD (bas) avec $K = 30$ microphones.	67
1	Localisation pour $\alpha = 1.9$ (à gauche) et $\alpha = 1.2$ (à droite): quand α devient petit, les sources de faible amplitude sont moins convenablement localisées.	69
9.1	Représentation spatiale d'un tirage de vecteurs $S\alpha S_c^2$ dans le cas réel: dans chaque direction, <i>le spectre spatial</i> est une collection de variables $S\alpha S_c$ indépendantes, avec une énergie contrôlée par la <i>mesure spatiale</i>	77
9.2	Mélange de deux densités de Von-Mises Fisher, de paramètres $\{\mu = \frac{\pi}{3}, \kappa = 15\}$ et $\{\mu = \frac{2\pi}{3}, \kappa = 30\}$	78
9.3	Représentation spatiale d'un tirage de vecteurs $S\alpha S_c^2$ dans le cas réel. Sur la figure de gauche, les deux tirages de $\mathbf{y}_1 \sim S\alpha S^2(\Gamma_1)$ (en vert) et $\mathbf{y}_2 \sim S\alpha S^2(\Gamma_2)$ (en rouge) admettent des densités spatiales Γ_1 et Γ_2 égales à la somme de deux distributions de Von-Mises Fisher, dont les maxima sont localisés en $\{\frac{5\pi}{8}, \frac{5\pi}{8} + \frac{\pi}{3}\}$ et $\{\frac{5\pi}{8} + \frac{\pi}{6}, \frac{5\pi}{8} + \frac{\pi}{2}\}$. Sur la figure de droite, nous voyons que la représentation des lignes de niveaux du mélange $\mathbf{x} = \mathbf{y}_1 + \mathbf{y}_2$ montre la propriété d'additivité des densités spatiales.	78
10.1	Séparation multidimensionnelle α -stable via l'estimation de spectres spatiaux.	82
10.2	Évaluation d'estimation des paramètres du modèle $\hat{\eta}$, pour η dans (10.18)	88
12.1	Deux densités de Von-Mises Fisher Γ_1, Γ_2 sur le demi-cercle, avec les directions respectives $\mu_1 = \frac{\pi}{3}, \mu_2 = \frac{\pi}{2}$ et le paramètre de concentration $\kappa = 15$. L'aire en rouge indique le chevauchement entre les deux <i>mesures spatiales</i>	100
12.2	Diagramme en boîte des scores d'EAM (plus le score est bas, meilleure est la séparation) des trois algorithmes, en fonction de l'écart angulaire entre les deux <i>mesures spatiales</i> avec $\alpha = 1.6$	101
12.3	Scores EAM en fonction d' α , pour un écart angulaire de 25 degrés entre les sources.	101
12.4	EAM moyennée sur $\alpha \in [1.2, 2]$, en fonction du nombre de composantes. Les lignes représentent la médiane tandis que les aires transparentes décrivent l'écart inter-quartiles de chaque méthode.	102
12.5	Diagramme en boîte de l'EAM pour $\alpha = 1.4$ en fonction du nombre de directions d'arrivées C , pour chaque composante.	104
13.1	Schéma général des algorithmes de débruitage.	110
13.2	Amplitudes de densités de probabilités de distributions α -MSCE (symétrique), pour $\alpha = 1.1$ à gauche et $\alpha = 2$ à droite.	114
14.1	Masque de Wiener théorique (figures du haut) et paramétrique (figures du bas).	119
14.2	Lissage des paramètres d'échelle.	120
14.3	Capture d'écran : Version de MAD en ligne pour la restauration d'archives du CREM.	123
15.1	Configuration de la pièce simulée pour l'évaluation des algorithmes de débruitage.	131
15.2	Moyenne de RSB-IP (plus le score est élevé, meilleure est l'amélioration) et DSP-IP (plus le score est faible, meilleure est la qualité de débruitage) en fonction du RSB en entrée. Les barres indiquent l'écart type de chaque scénario.	132

Liste des tableaux

1.1	Exemples de fonctions caractéristiques et de densités de probabilité associées à des variables aléatoires réelles.	5
12.1	Temps de calcul (en seconde) pour chaque méthode de filtrage, en fonction du nombre de composantes J à séparer.	103
14.1	Scores PESQ pour chaque algorithme. Les écritures en caractères gras indiquent les meilleurs scores.	122
15.1	Amélioration du RSB-IP quand le score de DSP-IP est identique pour chaque méthode. Les meilleures performances sont en police de caractère grasse.	132

Acronymes

- SSL : stationnaire au sens large
- α -MSCE : multivarié α -stable à contour elliptique
- EQM : erreur quadratique moyenne
- EMA : erreur absolue moyenne
- MDV : maximum de vraisemblance
- KKT : Karush-Khun Tucker
- NMF : *nonnegative matrix factorization*
- MJM : mise à jour multiplicative
- MOFF : méthodes des moments d'ordres faibles fractionnaires
- EM : espérance-maximisation
- MCMC : *Markov chain Monte-Carlo*
- MH : *Metropolis-Hastings*
- ACI : analyse en composantes indépendantes
- TF : temps-fréquence
- DSP : densité spectrale de puissance
- TFCT : transformée de Fourier à court terme
- SSS : séparation de sources sonores
- DAV : détecteur d'activité vocale
- RSB : rapport signal à bruit
- PESQ : *perceptual evaluation of speech quality*
- DSP-IP : distorsion spectrale pondérée par intelligibilité de la parole
- RSB-IP : rapport signal à bruit avec intelligibilité de la parole pondérée
- PDS-FWM : poids de distorsion spectrale du filtrage de Wiener multicanal
- MSD : mesure spatiale discrète

- TDOA : *time differential of arrival*
- PRD : puissance de réponse directionnelle
- PRD-HAT : puissance de réponse directionnelle avec phase transformée
- KAM : *kernel additive modelling*
- SSA : algorithme de soustraction spectrale d'amplitude
- SSG : algorithme de soustraction spectrale généralisée
- MEQM : algorithme minimum de l'erreur moyenne quadratique via l'estimation de l'amplitude spectrale à court terme
- MAD : *multialpha denoising*
- GDVP : filtrage de Wiener multicanal généralisé via une décomposition des valeurs propres
- ARC : *alpha residual component*
- FWM : filtrage de Wiener multicanal
- FWP : filtrage de Wiener paramétrique
- α -FWM : α -filtrage de Wiener multicanal
- α -FSS : filtrage α -stable multivarié via une estimation du spectre spatial
- α -FMC : filtrage α -stable via une minimisation de la covariation

Notations

Dans ce mémoire, tout scalaire est noté en police de caractères normale, e.g. $\alpha \in]0, 2]$, $F \in \mathbb{N}$. Un vecteur est quant à lui en minuscule et en police de caractères grasse, e.g. $\mathbf{u} \in \mathbb{C}^F$, $\boldsymbol{\theta} \in \mathbb{R}^K$. En outre, un tenseur ou une matrice est typographié en majuscule et en police de caractères grasse, e.g. $\mathbf{R} \in \mathbb{C}^{K \times K}$, $\mathbf{X} \in \mathbb{C}^{F \times T \times K}$. Les notations suivantes seront également utilisées :

- $\mathbb{N}, \mathbb{R}, \mathbb{C}$: respectivement l'ensemble des entiers naturels, nombres réels et nombres complexes
- \mathbb{K} : ensemble des nombres réels ou ensemble des nombres complexes
- \mathbb{A}^B : ensemble des vecteurs de dimension B dont chaque composante est définie sur l'ensemble \mathbb{A}
- \Re : partie réelle d'un nombre complexe
- \Im : partie imaginaire d'un nombre complexe
- \cdot^\top : transposé d'un vecteur
- \cdot^* : transposé hermitien (respectivement conjugué) d'un vecteur complexe (respectivement nombre complexe)
- $\langle \cdot, \cdot \rangle$: produit scalaire $\langle \mathbf{a}, \mathbf{b} \rangle = \mathbf{a}^* \mathbf{b}$
- $\|\cdot\|_H$: norme hermitienne $\|\mathbf{a}\|_H = \sqrt{\mathbf{a}^* \mathbf{a}}$
- ∇ : gradient
- $\#$: cardinal (nombre d'éléments) d'un ensemble
- $\llbracket a, b \rrbracket$: ensemble des entiers compris entre a et b
- $\mathcal{S}_{\mathbb{R}}^K = \{(\theta_1, \dots, \theta_K) \in \mathbb{R}^K, \sum_{k=1}^K \theta_k^2 = 1\}$: hypersphère réelle de dimension K
- $\mathcal{S}_{\mathbb{C}}^K = \{(\theta_1, \dots, \theta_K) \in \mathbb{C}^K, \sum_{k=1}^K |\theta_k|^2 = 1\}$: hypersphère complexe de dimension K
- \mathcal{S}^K : hypersphère réelle ou complexe de dimension K
- \mathbb{E} : espérance mathématique
- \mathbb{P} : probabilité d'un événement ou mesure de probabilité
- \mathbb{M} : médiane
- \mathcal{N} : loi gaussienne réelle
- \mathcal{N}_c : loi gaussienne complexe isotrope

- \mathcal{N}_c^K : loi gaussienne complexe multivariée isotrope de dimension K
- \mathcal{PN} : processus gaussien
- $S\alpha$: loi α -stable
- $S\alpha S$: loi α -stable symétrique (habituellement réelle)
- $S\alpha S_c$: loi α -stable symétrique complexe isotrope
- $S\alpha S_c^K$: loi α -stable symétrique complexe isotrope de dimension K
- $\mathcal{P}\alpha S$: loi positive α -stable
- $\mathcal{E}\alpha S^K$: loi multivariée α -stable à contour elliptique de dimension K
- g : densité de probabilité
- \triangleq : « égalité » par définition
- $\stackrel{d}{=}$: « égalité » entre deux distributions
- \propto : égalité à une constante multiplicative près
- \simeq : égalité numérique approximative
- \equiv : égalité fonctionnelle. (sur le même espace de définition, les fonctions sont égales)
- \odot : opération « terme à terme »
- \leftarrow : mise à jour des paramètres

Introduction générale

L'aléatoire, antonyme au déterminisme, n'a cessé d'être étudié depuis des siècles. L'une des applications majeures est la modélisation de données afin d'extraire une information souhaitée. Un exemple connu est la prédiction d'un événement futur afin de construire d'autres données à partir de celle qui sont observées [4, 26]. Il est par ailleurs courant d'employer des modèles capables de détecter aisément des événements rares, entendus comme se produisant peu souvent dans le temps. Ces informations inhabituelles peuvent être en général plus complexes à obtenir qu'un type de données se répétant régulièrement.

Dans le cas d'un signal acquis au cours du temps, il existe de nombreuses applications aboutissant à l'observation d'événements rares. Il peut alors se caractériser par l'apparition d'une valeur aberrante, au sens d'inhabituelle, dans les données reçues [96, 160]. Par exemple, si nous souhaitons obtenir la position géométrique d'une source sonore dans une pièce, nous pouvons nous ramener à étudier l'énergie en chaque point de l'espace [101]. Une variation soudaine de l'énergie en un point de l'espace peut alors être vue comme un événement rare : nous pouvons alors en conclure que la position en question est une source potentiellement active. Dans un tout autre registre, l'étude d'électroencéphalogrammes peut également contenir des valeurs aberrantes [190]. Ici, il s'agit d'une activité inhabituelle du cerveau qui cependant fournit des données dites parasites ou bruitées. Globalement, la restauration de données reste à ce jour un problème difficile à résoudre selon le niveau de dégradation et le type de données.

La restauration d'enregistrements est l'une des directions principales de cette thèse. Le centre de recherche ethnomusicologique (CREM)¹ assure la gestion des archives d'enregistrements anciens très dégradés. Plus particulièrement, il s'agit de fichiers audiofréquences monophoniques ou plus rarement stéréophoniques enregistrés dans différents pays. Ces archives du CNRS — Musée de l'Homme datent de périodes ultérieures à la fin du XIX^{ème} siècle jusqu'à aujourd'hui et furent acquises grâce à un ou plusieurs microphones. Malheureusement, les microphones de l'époque génèrent un bruit électrostatique qui dégrade les acquisitions. Ils pouvaient en outre saturer facilement si le volume sonore acquis était trop élevé [73]. C'est pourquoi, l'agence nationale de recherche (ANR) a financé le projet KAMoulox dans le but de mettre en relation des chercheurs du CREM et d'Inria, pour la restauration de ce patrimoine audiofréquence. Comme sus-mentionné, les données dégradées nécessitent des modèles robustes (qui prennent en compte des valeurs aberrantes) pour obtenir des résultats satisfaisants. En probabilité, les modèles robustes peuvent faire appel à des distributions dites à queues lourdes [153, 2]. En d'autres termes, la probabilité qu'une valeur aberrante se produise n'est pas négligeable. Un modèle classique du signal consiste alors à supposer que les données observées sont simplement la somme du signal recherché et d'un bruit additif [175]. L'addition de ces termes est appelée un mélange [20, 41, 158]. Nous souhaiterions alors que la loi considérée du mélange soit la somme des lois des deux sources. En prenant en compte un tel modèle probabiliste, deux objectifs doivent être atteints. Premièrement, arriver à estimer les paramètres des signaux additif à partir de l'observation des mélanges. Deuxièmement, élaborer une méthode de filtrage basée sur cette estimation paramétrique. Le cas des distributions gaussiennes est un candidat possible. La gaussienne a la particularité de n'être contrôlée que par deux paramètres : la covariance et la moyenne. Néanmoins, ce type de distribution admet une densité de probabilité dont

1. Centre de recherche ethnomusicologique : <http://lesc-cnrs.fr/crem>

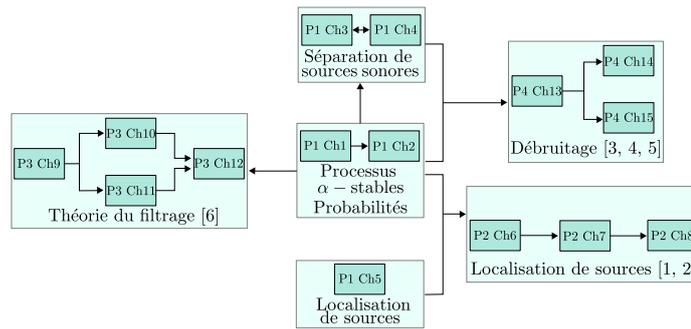


FIGURE 1 – Comportement asymptotique de la queue à droite de trois densités de probabilités. La gaussienne ($\alpha = 2$) est qualifiée de queue légère car elle est située en dessous de la loi exponentielle. Se localisant au-dessus de l'exponentielle, la distribution de Cauchy ($\alpha = 1$) est quand à elle dite à queue lourde.

le comportement asymptotique fait défaut pour l'acceptation de valeurs éloignées de sa moyenne. Plus particulièrement, les queues de cette distributions ne sont pas exponentiellement bornées (voir Fig. 1). Dans ce cas précis, la densité de probabilité d'une gaussienne est dite à queue légère.

Théorie α -stable : une modélisation robuste

Nous cherchons à proposer des distributions dites à queues lourdes et stables par combinaisons linéaires. Les lois vérifiant ces conditions furent proposées en premier par Paul Lévy durant les années 30. Cette théorie a été développée davantage pendant les années 60 par Mandelbrot [34, 125], pour la modélisation de marchés financiers. Le but était de pouvoir prédire les Krachs boursiers comme celui ayant eu lieu en 1929. D'un certain point de vue, il était alors nécessaire de modéliser un processus, qui contenait de nombreuses valeurs négligeables face à un nombre peu important de valeurs extrêmes. En somme, les premiers travaux de Mandelbrot introduisent une classe de processus plus générale appelée processus parcimonieux [172]. Ses recherches ont plus précisément esquissé des processus où le taux d'acceptation de valeurs aberrantes peut être contrôlé par un nombre $\alpha \in]0, 2]$ appelé exposant caractéristique [153]. C'est ainsi que les processus α -stables sont nés.

Dans le cadre de la théorie α -stable, la lourdeur des queues est contrôlée par le paramètre α (voir Fig. 1). Si α est maximum ($\alpha = 2$), il s'agit des distributions Gaussiennes. Si $\alpha = 1$, il s'agit des distributions de Cauchy. Pour résumer, la décroissance d' α augmente la lourdeur des queues. D'un certain point de vue, l'étude statistique des distributions α -stables rentre dans un cadre plus général appelé statistique robuste [78]. Ce domaine regroupe également l'étude d'autres distributions à queues lourdes comme les lois t de Student, de Pareto ou encore de Weibull. Il est intéressant de noter que certaines distributions α -stables multivariées et t de Student sont incluses dans une même classe de distributions : les distributions à contour elliptique [137]. La différence majeure réside dans le fait que la somme de distributions de t de Student n'est pas nécessairement une distribution t de Student. Dans les années 1970-1980, de nouveaux résultats montrant toute la richesse des distributions α -stables ont vu le jour. Citons par exemple le théorème de représentation spatiale, uniquement valable dans le cas non-Gaussien [100]. Ce dernier a permis de donner des interprétations uniques sur les distributions α -stables dans plusieurs domaines (financier [138], traitement du signal [95]...). En outre, une nouvelle représentation d'un processus α -stable grâce à l'apparition de la théorie des mesures aléatoires [87, 153] a également été esquissée. Il est en effet possible d'illustrer un processus α -stable sous la forme d'une intégrale paramétrisée par des distributions finies

α -stables.

Processus α -stables pour le traitement du signal

Comme mentionné précédemment, il devient naturel de modéliser des signaux impulsifs grâce à la théorie α -stable. Dans le cas du débruitage, les premiers travaux eurent lieu dans les années 70 [168]. En effet, lors de transmissions téléphoniques, des recherches ont employé les distributions α -stables pour modéliser des bruits dits impulsifs et ainsi réduire les interférences. Parallèlement, l'évolution de la théorie α -stable et du traitement du signal ont été complémentaires pendant ces dernières décennies. Citons par exemple l'introduction d'une nouvelle notion théorique appelée *covariation* dans [133] et son utilisation dans [65] pour le débruitage de signaux α -stables. Explicitement, la covariation vise à palier le talon d'Achille suivant : tout moment d'ordre plus grand ou égal à α n'existe pas pour une distribution α -stable [153]. La seule exception est celle des distributions gaussiennes ($\alpha = 2$) admettant des moments de tous ordres. Mais dans le cas de la covariation, nous n'employons pas des méthodes du traitement du signal basées sur des corrélations ou des erreurs quadratiques, qui sont liées à des moments d'ordre 2.

Depuis le milieu des années 2000 jusqu'à aujourd'hui, de nombreuses recherches en modélisation α -stable pour la *séparation de sources sonores* (SSS) a eu lieu. Le domaine de la SSS vise, à partir d'un enregistrement le plus souvent stéréophonique, à extraire plusieurs instruments de musique. À l'instar du débruitage, toutes les sources du mélange sont estimées et récupérées. Ainsi, l'usage habituel est de considérer à nouveau un modèle de mélange additif. Dans ce contexte, les processus Gaussiens ont bénéficié de certains développements [18, 33, 119, 117], et la théorie a été généralisée au cas de signaux plus impulsifs récemment [118, 162, 106, 161].

Certains travaux dans le cas multicanal α -stable ont aussi été traités ces 20 dernières années. Citons par exemple les travaux de Kidmose dans [94, 95] exploitant les distributions multivariées réelles α -stables pour l'*analyse en composantes indépendantes* (ACI) [33, 86] appliquée à la SSS. Ce principe repose sur la supposition que le mélange observé \mathbf{x} est une combinaison linéaire des sources \mathbf{s} supposées i.i.d. . En résumé, une relation du type $\mathbf{x} = \mathbf{A}\mathbf{s}$ où \mathbf{A} est appelée matrice de mélange. Afin de retrouver les sources \mathbf{s} , il faut que la matrice \mathbf{A} soit inversible. En pratique, les premières approches supposaient que le nombre de sources est inférieur ou égal au nombre d'observations : nous disons que le mélange est respectivement sur-déterminé ou déterminé. Sous hypothèse d'un modèle α -stable, Kidmose montre comment estimer la matrice de mélange avec précision, en mettant en œuvre la mesure spectrale. Introduit initialement par Kuelbs dans les années 70 [100], elle caractérise la répartition de l'énergie d'une distribution α -stable multivariée dans l'espace. Dans un tel mélange, la mesure spectrale pointe exactement dans la direction dont le vecteur directeur est proportionnel aux vecteurs colonnes de la matrice \mathbf{A} . Cependant, ce type de modèle reste limité dans le sens où seuls des mélanges linéaires instantanés peuvent être considérés. Pour résumer, la représentation riche du signal dans le domaine temps-fréquence n'est pas traitée comme dans le cas monocanal de [118].

Résumé des apports de la thèse

L'un des souhaits majeurs de cette thèse a été de continuer à « défricher » la théorie α -stable pour le traitement du signal. Notamment, la généralisation de [118] dans le cas multicanal était un des buts majeurs : fournir une méthode de filtrage pour un mélange de processus α -stables stationnaires non-scalaires. Nous avons tout d'abord montré qu'un vecteur α -stable admet une représentation sous forme d'intégrale stochastique. Plus précisément, la distribution d'un vecteur α -stable peut être vue comme la contribution d'une infinité de variables α -stables et pointant dans toutes les directions. La répartition spatiale est résumée par une mesure aléatoire que nous avons appelée *spectre spatial* et l'énergie de chaque variable est contenue dans la *mesure spatiale*. Cette nouvelle représentation nous a alors permis de fournir

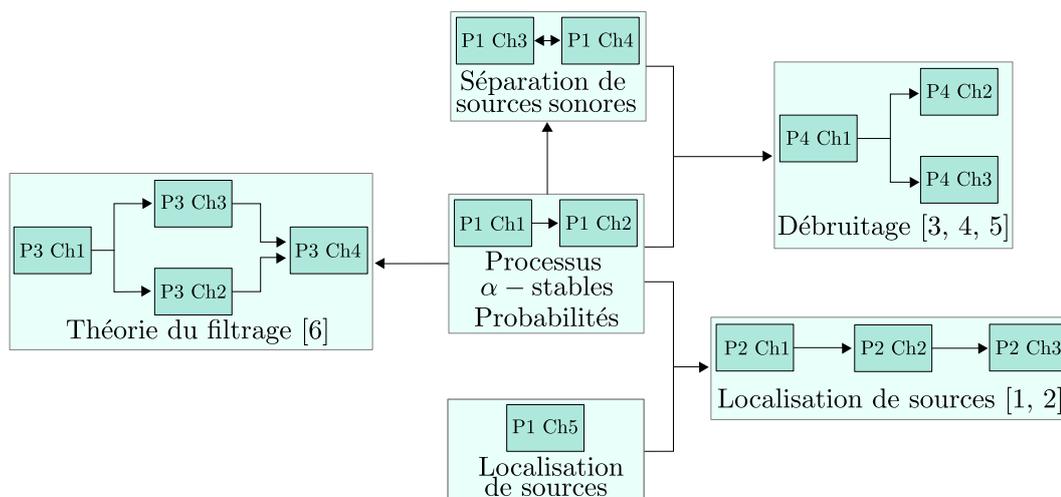


FIGURE 2 – Structure du manuscrit. Les flèches désignent la dépendance entre chaque chapitre ou partie. Les chiffres entre crochets font référence aux articles publiés durant cette thèse (voir ci-dessous).

une méthode pour obtenir des réalisations de vecteurs α -stables. En outre, deux théories du filtrage ont été établies. La première est basée sur une technique de filtrage non-linéaire. Elle découle d'une nouvelle représentation spatiale d'un vecteur α -stable multivarié. La seconde est quant à elle une généralisation au cadre α -stable du filtrage linéaire multicanal minimisant l'*erreur quadratique moyenne* (EQM). Dans la littérature, ce critère de minimisation de l'EQM via un opérateur linéaire est appelé *filtrage de Wiener*.

Nous avons souhaité par ailleurs approfondir les travaux dans [162] sur le débruitage de signal monophonique en α -stable. Pour ce faire, nous avons développé un modèle de mélange hybride où la parole et le bruit n'admettent pas le même exposant caractéristique α [58]. Il en découle une plus grande souplesse sur la mise au point des filtres pour un meilleur rehaussement de la parole. En parallèle, une nouvelle technique de débruitage multicanal employant un modèle hybride gaussien (pour la parole) plus α -stable (pour le bruit) a été élaborée [59]. Les résultats apparaissent en général meilleurs en présence d'un scénario de débruitage où le bruit à atténuer est plus complexe.

Un autre aspect exploré est celui de la localisation de sources sonores. Rappelons que la mesure spectrale utilisée par Kidmose dans [94] est un opérateur « pointant » spatialement vers chaque source supposée ponctuelle et évoluant dans un espace multidimensionnel. Nous avons alors montré que la mesure spectrale contient exactement la position spatiale de sources sonores dans une salle. Plusieurs résultats dans [153, p 69] ont pu renforcer notre conviction sur la pertinence de la mesure spectrale pour la localisation de sources sonores en milieu réverbérant. La nouveauté dans [60, 61] par rapport à [95] est l'étude des signaux dans le domaine temps-fréquence et leur application pour la localisation de sources sonores.

Structure du manuscrit et annonce du plan

Ce mémoire de thèse est constitué de quatre parties, chacune décomposée en chapitres, sections et sous-sections. Certains de ces chapitres sont totalement indépendants et peuvent être ignorés par le lecteur ne souhaitant s'intéresser qu'à l'une des parties du manuscrit. Les chapitres 3 et 5 de la partie I ne sont pas indispensables pour une bonne compréhension de la partie IV. En outre, le chapitre 5 est nécessaire pour la description des méthodes de l'état de l'art employées dans la partie II. Le chapitre 2 de la partie

Il est important pour se familiariser avec la théorie α -stable et les notations employées dans ce mémoire de thèse. Autrement, les parties II, III et IV sont indépendantes. La structure générale est détaillée dans la Fig. 2.

Partie 1 : Contextes de modèles probabilistes pour la séparation et la localisation de sources sonores

Nous commençons dans un premier temps par rappeler les outils classiques en probabilités. La raison de ce premier chapitre est de familiariser le lecteur avec les outils mathématiques et notations employés. Nous rappelons par exemple la fonction caractéristique, opérateur essentiel pour l'étude des distributions α -stables. De plus, il nous a paru important de définir les notions de processus stochastique et de mesure aléatoire, nécessaires à la représentation introduite dans la partie III. Un exemple illustrant l'utilité de la mesure aléatoire est par ailleurs ajouté à la fin du premier chapitre.

Le second chapitre décrit différents résultats sur les processus α -stables, en orientant la discussion vers un cadre adapté au traitement du signal. Une première section définit les bases de la théorie α -stable tandis que la seconde aborde le sujet de la représentation d'un signal par un processus gaussien stationnaire ou non-stationnaire. Nous finissons par décrire la façon dont un signal dans le domaine temps-fréquence peut être représenté via un processus dit harmonisable.

Nous considérons ensuite le mélange de processus α -stables pour la SSS. Nous rappelons entre autres les résultats essentiels de filtrage de processus gaussiens dans le cas stationnaire. En particulier, ceux qui se ramènent au filtrage de Wiener classique et d'autres extensions proposées dans les années 2000. De plus, nous explicitons les travaux de Liutkus & Badeau dans [118] et l'apport de ces derniers quant à la proposition d'une nouvelle technique de filtrage et de résultats théoriques prometteurs.

Une fois la technique de séparation établie, il reste alors à définir une méthode pour estimer les paramètres régissant les modèles probabilistes pris en compte pour chaque signal. Une piste classique dans ce sens est d'adopter un modèle à rang faible pour les *densités spectrales de puissance* (DSP) des signaux. Nous montrons comment cette piste conduit à l'utilisation d'algorithmes de type *nonnegative matrix factorization* (NMF) appliqués aux spectrogrammes des signaux, en nous concentrant sur le cas de processus Gaussiens et Cauchy.

Le dernier chapitre, relativement indépendant des autres, traite des techniques de localisation dans le cas α -stable. En premier lieu, nous évoquons la localisation de sources acoustiques, où l'emploi de moment d'ordre 2 est fait pour introduire la *puissance de réponse directionnelle* (PRD). Nous décrivons également des recherches menées dans [65] pour la localisation en présence de deux microphones avec un modèle α -stable de sources.

Partie 2 : Mélanges discrets et application à la localisation de sources audio

Cette seconde partie amène tout d'abord divers résultats théoriques que nous avons démontrés sur les distributions α -stables. Nous verrons qu'un mélange linéaire vectoriel de variables α -stables peut être spatialement caractérisé grâce à un opérateur nommé *mesure spatiale discrète* (MSD). Plus précisément, la MSD contient l'énergie de chaque source d'un mélange.

Nous expliquons alors comment la MSD peut être estimée. Ce problème est résolu via une relation linéaire entre la MSD, un opérateur non-négatif appelé *exposant de Lévy* et une matrice à coefficients non-négatifs que nous explicitons. L'intérêt de considérer une telle équation est de permettre d'employer une technique d'optimisation consistant en des mises à jour multiplicatives des paramètres, garantissant la non-négativité théorique de la MSD. Nous commentons également sa complexité calculatoire relativement faible. En effet, la dimension de l'*exposant de Lévy* est en général plus petite qu'un fichier audiofréquence multicanal : la MSD peut alors être qualifiée d'opérateur de sketching [92].

Enfin, nous employons la MSD pour la localisation de source audio. En particulier, nous remarquons que l'estimation découle également du modèle acoustique considéré. La MSD fournit une carte de chaleur des énergies de chaque point d'une pièce : plus la valeur de la MSD en un point est importante, plus la possibilité qu'une source soit présente est grande. Nous comparons finalement notre algorithme avec trois autres méthodes classiques : deux approches gloutonnes utilisées en astronomie et nommées CLEAN et RELAX ainsi que la PRD rattachée aux moments d'ordre 2.

Partie 3 : Filtrage de processus α -stable vectoriels

Après avoir considéré des mélanges linéaires vectoriels de variables α -stables, nous décrivons comment il est possible de filtrer directement un mélange de vecteurs α -stables. Une nouvelle représentation spatiale d'un vecteur aléatoire α -stable est introduite dans un premier temps. Nous montrons qu'un tel vecteur peut être vu comme engendré par la superposition d'une infinité de variables α -stables scalaires indépendantes, en provenance de toutes les directions. L'objet contenant l'énergie globale de chacune des directions est appelé *mesure spatiale* et la répartition des contributions de chaque direction est résumée via un opérateur intitulé *spectre spatial*. Notons qu'une application directe de ce théorème de représentation est de fournir une méthode pour générer des distributions multivariées α -stables.

Pour continuer, nous proposons une première technique de filtrage pour le problème suivant : étant donné un vecteur α -stable observé égal à une somme de vecteurs α -stables, comment retrouver chaque composante ? Nous décrivons alors une méthode de filtrage non-linéaire basée sur une estimation directe du *spectre spatial*. La forme de base de l'estimateur proposé n'est en pratique pas implémentable. C'est pourquoi nous transformons l'estimateur et en donnons une version équivalente calculable en pratique. Nous aboutissons finalement à un algorithme de séparation.

Une deuxième approche est aussi proposée pour la séparation de vecteurs α -stables. Dans celle-ci, la technique de filtrage est linéaire et inspirée des célèbres travaux de Norbert Wiener [179], visant à minimiser l'erreur moyenne quadratique entre une source et son estimée. Dans le cas α -stable, la covariance n'existe pas sauf si $\alpha = 2$. Nous faisons cependant appel à un outil similaire appelé *covariation* [27]. Celle-ci a la particularité de jouir de propriétés proches de la covariance. La méthode sera alors équivalente à des mises à jour de paramètres, rappelant une méthode dite du point fixe en mathématiques. Nous prouvons également que dans le cas où α tend vers 2, le filtrage multicanal obtenu est celui de Wiener.

Afin de comparer les différentes approches, nous comparons chaque méthode de filtrage avec celui de Wiener multicanal quand α tend vers 2. Nous supposons cependant connues les mesures spatiales pour chaque algorithme. Trois types d'évaluations donnent alors des scores informant sur le pouvoir séparateur en fonction respectivement de la distance spatiale, du nombre et de la complexité spatiale des composantes.

Partie 4 : Modèles hybrides et applications au rehaussement de la parole

Cette dernière partie décrit deux algorithmes de débruitage hybrides. Le premier est monocanal et considère que le bruit et la parole sont des distributions α -stables n'admettant pas le même exposant caractéristique α . Nous remarquons que considérer un tel mélange revient à prendre en compte un filtrage de Wiener dit paramétrique, faisant référence à des méthodes heuristiques ayant émergé dans les années 80. Dans un autre registre, un filtrage multicanal est proposé où la composante de bruit est régie par un modèle α -stable particulier appelé *multivarié α -stable à contour elliptique*. La parole est cette fois-ci modélisée via des distributions gaussiennes admettant un grand nombre de paramètres. Nous supposons en effet que les covariances sont de rang plein. À contrario, le bruit additif α -stable admet un modèle paramétrique très simple. Dans ce contexte, nous supposons que le modèle de voix ne s'écarte pas beaucoup du comportement attendu alors que celui du bruit peut s'éloigner significativement des paramètres estimés.

Nous décrivons alors la méthode d'estimation des paramètres en monocanal. Habituellement, l'estimation de certains d'entre eux est effectuée en utilisant un algorithme de type *Metroplis-Hastings*. Nous avons plutôt considéré un estimateur plus simple et robuste qui a d'excellentes performances pratiques. Parallèlement, les paramètres d'échelle contenant l'énergie des sources sont calculés en utilisant des moyennes glissantes le long de l'axe temporel. Nous comparons finalement notre algorithme à plusieurs techniques classiques de filtrage de Wiener paramétriques. En outre, cet algorithme a été testé pour la restauration des archives du CREM et a prouvé son efficacité pour la restauration des archives dégradées.

Le dernier chapitre est dédié à l'estimation des paramètres du modèle multicanal. Ici, les paramètres du modèle gaussien sont estimés à l'aide d'un modèle classique où la covariance est décomposée comme un produit entre un scalaire positif et une matrice ne dépendant que de la fréquence. Ce scalaire est alors factorisé via une NMF. Les matrices, appelées matrices de *covariance spatiale*, sont quant à elles évaluées via des mises à jour découlant de techniques d'optimisation classiques. Enfin, une partie des paramètres du résidu α -stable est estimée en début d'algorithme et le reste en parallèle avec les paramètres du modèle gaussien, à l'aide d'un algorithme de type *espérance-maximisation* couplé à un algorithme de *Metropolis-Hastings*.

Publications

Articles de Conférences

- [1] M. Fontaine, C. Vanwysberghe, A. Liutkus, and R. Badeau. Sketching for nearfield acoustic imaging of heavy-tailed sources. In *International Conference on Latent Variable Analysis and Signal Separation*, pages 80-88. Springer, 2017.
- [2] M. Fontaine, C. Vanwysberghe, A. Liutkus, and R. Badeau. Scalable source localization with multichannel alpha-stable distributions. In *Proc. European Signal Processing Conf. (EUSIPCO)*, pages 11-15. IEEE, 2017.
- [3] M. Fontaine, A. Liutkus, L. Girin, and R. Badeau. Explaining the parameterized Wiener filter with alpha-stable processes. In *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, 2017.
- [4] M. Fontaine, F-R. Stöter, A. Liutkus, U. Şimşekli, R. Serizel, and R. Badeau. Multichannel audio modeling with elliptically stable tensor decomposition. In *International Conference on Latent Variable Analysis and Signal Separation*, pages 13-23. Springer, 2018.
- [5] M. Fontaine, A.A. Nugraha, R. Badeau, K. Yoshii and A. Liutkus. Cauchy Multichannel Speech Enhancement with a Deep Speech Prior. In *Proc. European Signal Processing Conf. (EUSIPCO)*, 2019.

Article de Revue

- [6] M. Fontaine, R. Badeau and A. Liutkus. Multichannel alpha-stable filtering. In *Signal Processing*. (Soumis).

Première partie

**Contextes de modèles probabilistes
pour la séparation et la localisation de
sources sonores**

Chapitre 1

Premières définitions

Résumé

Dans ce chapitre, nous introduisons plusieurs définitions préliminaires et notions en probabilités. Il est en effet important de souligner l'importance d'outils classiques, utilisables par la suite pour le développement de modèles probabilistes. En particulier, la fonction caractéristique est un outil indispensable pour l'étude des distributions α -stables. Nous définissons en outre le terme de mesure aléatoire. Il s'agit d'un opérateur permettant d'esquisser une représentation des processus α -stables qui s'avère pratique pour l'estimation de ses paramètres.

1.1 Théorie probabiliste

1.1.1 Fonction de répartition et densité de probabilité

Par souci de compréhension et de clarté, la théorie présentée dans cette section sera considérée dans \mathbb{R}^K . Nous définissons la tribu Borélienne de \mathbb{R}^K , noté \mathcal{B}^K , comme étant celle engendrée par les pavés de \mathbb{R}^K . Nous disons que \mathbb{P} est une *mesure de probabilité* définie sur la tribu borélienne \mathcal{B}^K à valeurs dans $[0, 1]$ si :

- (*σ -additivité*) : $\forall B_1, B_2, \dots \in \mathcal{B}^K$ dénombrables et deux à deux disjoints, $\mathbb{P}(\cup_i B_i) = \sum_i \mathbb{P}(B_i)$;
- (*mesure unitaire*) : $\mathbb{P}(\mathbb{R}^K) = 1$.

L'espace probabilisé associé à la tribu Borélienne est donc le triplet $(\mathbb{R}^K, \mathcal{B}^K, \mathbb{P})$. Par ailleurs, les éléments de \mathcal{B}^K sont appelés des événements. Rappelons en outre qu'un *vecteur aléatoire réel* $\mathbf{x} \triangleq [x_1, \dots, x_K]^T : \mathcal{B}^K \rightarrow \mathbb{R}$ est une fonction qui à tout événement $B \in \mathcal{B}^K$ associe un nombre réel $\mathbf{x}(B) \in \mathbb{R}^K$. La loi de probabilité de \mathbf{x} , notée $\mathbb{P}_{\mathbf{x}}$, est une mesure de probabilité définie comme la mesure image de \mathbb{P} par \mathbf{x} . Une manière de décrire sa loi de probabilité est d'employer sa *fonction de répartition* comme suit :

$$\forall \mathbf{u} \triangleq [u_1, \dots, u_K]^T \in \mathbb{R}^K, G_{\mathbf{x}}(\mathbf{u}) = \mathbb{P}(x_1 \leq u_1, \dots, x_K \leq u_K), \quad (1.1)$$

où \cdot^T représente l'opérateur de transposition et \triangleq l'« égalité » par définition. De ce fait, la fonction de répartition décrit la façon dont la probabilité croît quand \mathbf{u} augmente. Autrement dit, elle décrit l'évolution de la loi de probabilité de \mathbf{x} . La *densité de probabilité* de \mathbf{x} , quand elle existe, est la fonction $g_{\mathbf{x}}$ définie par

$$\forall \mathbf{u} \triangleq [u_1, \dots, u_K]^T \in \mathbb{R}^K, G_{\mathbf{x}}(\mathbf{u}) = \int_{v_1=-\infty}^{u_1} \dots \int_{v_K=-\infty}^{u_K} g_{\mathbf{x}}(v_1, \dots, v_K) dv_1 \dots dv_K, \quad (1.2)$$

où $d_{v_1} \dots d_{v_K}$ est la mesure de Lebesgue sur \mathbb{R}^K . Soulignons le fait que G_x n'est pas en général la primitive de g_x . L'égalité *presque partout* au sens de la mesure de Lebesgue est vraie, mais rien ne garantit l'égalité entre la différentielle de G_x et la fonction g_x sur tout \mathbb{R}^K .

1.1.2 Espérance, moments et fonction caractéristique

Bien que la fonction de répartition décrit le comportement d'une distribution, elle n'admet pas toujours de forme analytique. Pour pallier ce problème, un outil alternatif appelé fonction caractéristique est définie. Le but de cette section est d'en fournir une description et quelques exemples classiques.

Soit \mathbf{x} un vecteur aléatoire réel et l'espace probabilisé $E_K \triangleq (\mathbb{R}^K, \mathcal{B}^K, \mathbb{P})$. Liée à la loi de probabilité \mathbb{P}_x et sous réserve que \mathbf{x} soit intégrable selon la mesure de probabilité \mathbb{P} , nous définissons l'espérance de \mathbf{x} comme suit :

$$\mathbb{E}(\mathbf{x}) = \int_{\mathbb{R}^K} \mathbf{v} \mathbb{P}_x(d\mathbf{v}). \quad (1.3)$$

Si de plus \mathbf{x} admet une densité de probabilité g_x , alors $\mathbb{E}(\mathbf{x}) = \int_{\mathbb{R}^K} \mathbf{v} g_x(\mathbf{v}) d\mathbf{v}$ est également l'espérance mathématique de \mathbf{x} , sous réserve d'absolue convergence de l'intégrale. De la définition de l'espérance en (1.3), il découle que :

- \mathbb{E} est un opérateur linéaire ;
- (*théorème du transfert*) : Soit \mathbf{x} un vecteur aléatoire à valeurs dans $(\mathbb{R}_{|\mathbf{x}}^K, \mathcal{B}_{|\mathbf{x}}^K)$; c'est à dire restreint à l'espace image de \mathbf{x} et f une fonction mesurable sur $(\mathbb{R}_{|\mathbf{x}}^K, \mathcal{B}_{|\mathbf{x}}^K)$. Sous réserve d'existence, l'espérance de $f(\mathbf{x})$ est définie par :

$$\mathbb{E}(f(\mathbf{x})) = \int_{\mathbb{R}_{|\mathbf{x}}^K} f(\mathbf{v}) \mathbb{P}_x(d\mathbf{v}). \quad (1.4)$$

Grâce au théorème du transfert, on peut spécifier quand il existe le *moment d'ordre* $p \in \mathbb{R}$ de $\mathbf{x} \triangleq [x_1, \dots, x_K]^T$:

$$\mathbb{E}(\mathbf{x}^p) \triangleq \int_{\mathbb{R}_{|\mathbf{x}}^K} x_1^p x_2^p \dots x_K^p \mathbb{P}_x(d_{x_1} \dots d_{x_p}). \quad (1.5)$$

Comme mentionné précédemment, le moment d'un certain ordre peut ne pas exister. Cependant, la transformée de Fourier inverse de la densité de probabilité de \mathbf{x} est toujours définie. Elle est appelée *fonction caractéristique* et se décrit, grâce au théorème de transfert, comme suit pour \mathbf{x} :

$$\forall \mathbf{u} \in \mathbb{R}^K, \varphi_x(\mathbf{u}) \triangleq \mathbb{E}(\exp(i\mathbf{u}^T \mathbf{x})) = \int_{\mathbb{R}_{|\mathbf{x}}^K} \exp(i\mathbf{u}^T \mathbf{v}) \mathbb{P}_x(d\mathbf{v}). \quad (1.6)$$

Remarque Dans le cas d'un vecteur aléatoire complexe \mathbf{x} , sa fonction caractéristique est égale à

$$\forall \mathbf{u} \in \mathbb{C}^K, \varphi_x(\mathbf{u}) \triangleq \mathbb{E}(\exp(i\Re[\mathbf{u}^* \mathbf{x}])) = \int_{\mathbb{C}_{|\mathbf{x}}^K} \exp(i\mathbf{u}^* \mathbf{v}) \mathbb{P}_x(d\mathbf{v}),$$

où \cdot^* est la transposée Hermitienne d'un vecteur complexe.

Le tableau 1.1 répertorie la densité de probabilité et la fonction caractéristique de différentes lois centrées réduites classiques. Citons par ailleurs quelques propriétés importantes de la fonction caractéristique pour un vecteur aléatoire réel \mathbf{x} de dimension K :

- $\forall a, b \in \mathbb{R}, \forall \mathbf{u} \in \mathbb{R}^K, \varphi_{a\mathbf{x}+b}(\mathbf{u}) = \varphi_{\mathbf{x}}(a\mathbf{u})\exp(ib\mathbf{u})$;
- $\forall \mathbf{u} \in \mathbb{R}^K, |\varphi_{\mathbf{x}}(\mathbf{u})| \leq 1$;
- Sous réserve d'existence, la relation suivante permet de relier la fonction caractéristique et tout moment d'ordre p :

$$\forall p \in \mathbb{R}, \nabla_{\mathbf{x}}^p \varphi_{\mathbf{x}}(0) = (-i)^{-p} \mathbb{E}(\mathbf{x}^p). \quad (1.7)$$

Loi (<i>centrée réduite</i>)	Densité de probabilité $g_{\mathbf{x}}$	Fonction caractéristique $\varphi_{\mathbf{x}}$
Gaussienne $\mathcal{N}(0, 1)$	$v \mapsto \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{v^2}{2}\right)$	$u \mapsto \exp\left(-\frac{u^2}{2}\right)$
Cauchy $\mathcal{C}(0, 1)$	$v \mapsto \frac{1}{\pi(1+v^2)}$	$u \mapsto \exp(- u)$
Laplace	$v \mapsto \frac{1}{2} \exp(- v)$	$u \mapsto \frac{1}{1+u^2}$

TABLE 1.1 – Exemples de fonctions caractéristiques et de densités de probabilité associées à des variables aléatoires réelles.

1.1.3 Indépendance

On dit que N vecteurs aléatoires $\mathbf{x}_1, \dots, \mathbf{x}_N$ définis sur E_K sont indépendants si :

$$\forall A_1, \dots, A_N \in \mathcal{B}^K, \mathbb{P}(\mathbf{x}_1 \in A_1, \dots, \mathbf{x}_N \in A_N) = \prod_{n=1}^N \mathbb{P}(\mathbf{x}_n \in A_n). \quad (1.8)$$

Dans le cas où x_1, \dots, x_N sont simplement des variables aléatoires sur $(\mathbb{R}, \mathcal{B})$ et si, pour toute fonction Borélienne f_1, \dots, f_N (i.e. de \mathbb{R} dans \mathbb{R} et \mathcal{B} -mesurable), l'égalité

$$\mathbb{E}\left(\prod_{n=1}^N f_n(x_n)\right) = \prod_{n=1}^N \mathbb{E}(f_n(x_n)) \quad (1.9)$$

est vérifiée, alors x_1, \dots, x_N sont indépendantes. Rappelons également une propriété qui nous sera utile par la suite. Si N vecteurs aléatoires $\mathbf{x}_1, \dots, \mathbf{x}_N$ sont indépendants, alors on a :

$$\varphi_{\mathbf{x}_1+\dots+\mathbf{x}_N} \equiv \prod_{n=1}^N \varphi_{\mathbf{x}_n}. \quad (1.10)$$

où \equiv représente l'égalité entre deux applications.

1.1.4 Conditionnalité et inférence Bayésienne

En mathématiques appliquées, la probabilité conditionnelle est une notion primordiale pour l'*inférence Bayésienne* [85]. Considérons $H \in \mathcal{B}^K$ ayant le sens d'hypothèse d'une expérience. En général, les hypothèses d'une expérience sont connues et par conséquent l'*a priori* $\mathbb{P}(H)$ l'est également. La probabilité *a posteriori* est quant à elle une probabilité des hypothèses mise à jour suite à une observation $O \in \mathcal{B}^K$. Elle est notée $\mathbb{P}(H | O)$ et vérifie la relation :

$$\mathbb{P}(H | O) = \frac{\mathbb{P}(O, H)}{\mathbb{P}(O)}. \quad (1.11)$$

Le but de l'inférence Bayésienne repose sur le calcul de (1.11). En somme, vérifier si les hypothèses restent valables lors de nouvelles observations. Le théorème de Bayes

$$\mathbb{P}(H | O) = \frac{\mathbb{P}(O | H) \times \mathbb{P}(H)}{\mathbb{P}(O)} \propto \mathbb{P}(O | H) \times \mathbb{P}(H) \quad (1.12)$$

nous permet de relier l'*a posteriori* avec l'*a priori* et $\mathbb{P}(O | H)$ appelé *vraisemblance*. En particulier, la quantité $\mathbb{P}(O | H)$ permet de proposer un modèle à l'expérience considérée. Nous détaillons dans le chapitre 3 comment le principe d'inférence Bayésienne peut être utilisé pour des modèles de séparation de sources sonores. Intéressons-nous maintenant à la notion de processus stochastique scalaire, autrement appelé fonction aléatoire.

1.2 Processus stochastiques

1.2.1 Définitions et propriétés

Soit $E = (\mathbb{R}^K, \mathcal{B}^K, \mathbb{P})$ un espace probabilisé. Il existe de nombreuses définitions d'un processus stochastique dans la littérature. L'une des plus classiques est la suivante [57] :

Définition 1.2.1. Soit \mathfrak{N} un ensemble dénombrable ou non. Un **processus stochastique** peut être vu comme une collection de variables aléatoires $\mathbf{x} \triangleq (x(n))_{n \in \mathfrak{N}}$, où chacune d'entre elle est définie sur E et à valeurs dans son espace d'état $(\mathbb{R}^K, \mathcal{B}^K)$.

Nous notons abusivement la valeur d'une variable aléatoire $x(n)$ pour un certain $\omega \in \mathcal{B}^K$ par $x(n, \omega)$ ou $(x(n))(\omega)$. Un processus stochastique peut être vu comme la généralisation d'un vecteur aléatoire où la dimension devient infinie. Si nous fixons une réalisation $\omega \in \mathcal{B}^K$, alors la fonction

$$\begin{aligned} x(., \omega) : \mathfrak{N} &\rightarrow (\mathbb{R}^K, \mathcal{B}^K) \\ n &\mapsto (x(n))(\omega) \end{aligned} \quad (1.13)$$

est appelée une *trajectoire* du processus \mathbf{x} [57]. Nous allons maintenant énoncer la généralisation d'une loi pour un processus stochastique. Pour ce faire, soit $\mathbb{R}^{\mathfrak{N}}$ l'ensemble des applications définies sur \mathfrak{N} en tout point et à valeurs dans \mathbb{R} .

Définition 1.2.2. Soit $\mathbf{x} \triangleq (x(n))_{n \in \mathfrak{N}}$ un processus stochastique comme défini en 1.2.1. La **loi de probabilité** de \mathbf{x} est alors [57] :

$$\mathbb{P}_{\mathbf{x}} = \mathbb{P} \circ \mathbf{x}^{-1} \quad (1.14)$$

où :

$$\forall B \subset \mathbb{R}^{\mathfrak{N}}, \mathbf{x}^{-1}(B) = \{\omega \in \mathcal{B}^K; \mathbf{x}(\omega) \in B\} \quad (1.15)$$

représente la pré-image de \mathbf{x} . Cette définition est valable sous condition que $\mathbf{x}^{-1}(B)$ existe (au sens d'être mesurable).

La loi de probabilité, comme dans le cas vectoriel ou scalaire, permet elle aussi de caractériser le comportement d'un processus stochastique. En pratique, nous sommes amenés à ne considérer les processus que sur un nombre fini de points de mesures. Autrement dit, à étudier la loi jointe d'un nombre de réalisations fini d'une trajectoire². Nous allons expliquer rigoureusement pourquoi une telle restriction fait sens afin de se rapprocher de la loi de probabilité $\mathbb{P}_{\mathbf{x}}$:

2. https://perso.univ-rennes1.fr/jean-christophe.breton/Fichiers/processus_M2.pdf

Définition 1.2.3. On appelle *lois fini-dimensionnelles* d'un processus l'ensemble des lois

$$\mathcal{L}_x = \left\{ \mathbb{P}_{(x(n_1), \dots, x(n_N)); n_1, \dots, n_N \in \mathfrak{N}, N \in \mathbb{N}^*} \right\}. \quad (1.16)$$

En d'autres termes, \mathcal{L}_x résume l'ensemble des lois d'un nombre fini de réalisations de x . Nous avons alors la propriété suivante :

Proposition 1.2.1. La loi \mathbb{P}_x d'un processus x est entièrement caractérisée par ses lois fini-dimensionnelles \mathcal{L}_x .

Il en découle directement que deux processus x et y ont la même loi si et seulement si $\mathcal{L}_x = \mathcal{L}_y$. Il est donc justifié d'étudier la loi jointe d'un nombre fini de réalisations d'un processus. La notion d'indépendance entre deux processus stochastiques s'étend de la manière suivante [108] :

Définition 1.2.4. On dit que x et y sont *indépendants*, si tout tirage en nombre fini $(x(n))_n$ et $(y(n))_n$ respectivement des processus x et y est mutuellement indépendant.

1.2.2 Mesures aléatoires et représentation stochastique

Nous illustrons dans cette section la mesure aléatoire en employant un exemple intuitif. Pour une présentation rigoureuse, le lecteur est par exemple invité à consulter [88]. Considérons $F = (\mathbb{R}, \mathcal{B}, m)$ un espace mesurable et $E = (\mathbb{R}^K, \mathcal{B}^K, \mathbb{P})$ un espace probabilisé. Comparé à E , la mesure m est ici une mesure de probabilité généralement non-unitaire (voir définition de l'unitarité dans la sous-section 1.1.1). Soit $M \triangleq \{M_B\}_{B \in \mathcal{B}}$ une collection de variables aléatoires définie sur E . On dit que M est une *mesure aléatoire de Poisson* associée à la *mesure de contrôle* m si :

- $\forall B \in \mathcal{B}, M_B$ suit une loi de Poisson de paramètre $m(B)$;
- Si les ensembles $B_1, \dots, B_N \in \mathcal{B}$ sont disjoints, alors M_{B_1}, \dots, M_{B_N} sont indépendants ;
- $\forall \omega \in \mathbb{R}^K, M(\cdot, \omega)$ est une mesurévaluonse sur (F, \mathcal{B}, m) .

Plaçons nous maintenant dans le cas d'une famille de processus définie sur \mathbb{R}_+ . Les processus de Lévy peuvent être décrits à partir de la mesure aléatoire de Poisson via la célèbre décomposition de Lévy-Itô [6]. En d'autres termes, ils sont l'extension continue d'une marche aléatoire. Un processus de Lévy peut alors faire apparaître une intégrale stochastique contre une mesure de Poisson aléatoire, décrivant la partie ponctuelle du processus. Il peut donc s'interpréter comme la somme de variables aléatoires de Poisson indépendantes. Rappelons que les lois de Poisson comptent le nombre d'événements pouvant se produire sur un espace fixé. Ici, elles jouent le rôle de comptage des discontinuités du processus. En traitement du signal, la mesure aléatoire de Poisson a été employée dans [12] afin d'unifier plusieurs modèles de réverbération classique. En outre, leur utilisation pour la construction de processus α -stables est également explicitée dans [153]. En conclusion, une mesure aléatoire peut être vue comme une collection de mesures indexées par Ω . Plus particulièrement dans le cas qui nous intéresse, une mesure aléatoire peut être une fonction générant des lois de probabilités.

Dans le cas où la mesure aléatoire est complexe, la mesure de contrôle est alors définie sous la forme suivante [153] :

Définition 1.2.5. Soit \mathcal{B}^2 la tribu borélienne engendrée par les pavés de \mathbb{R}^2 et S^1 le cercle unité. On appelle *mesure circulaire* k , une mesure définie sur l'espace $(F \times S^1, \mathcal{A} \otimes \mathcal{B}^2)$ telle que :

$$\forall A \in \mathcal{A}, k\left(A \times S^1\right) < +\infty. \quad (1.17)$$

Nous emploierons cette définition pour décrire une classe de processus α -stables complexes.

Chapitre 2

Processus α -stables

Résumé

La motivation de ce chapitre est d'une part d'appuyer le choix de différents modèles probabilistes dans le cadre du traitement du signal et d'autre part d'apporter plusieurs éléments de comparaison et de justification quant à la motivation d'employer la théorie α -stable pour la modélisation de signaux numériques. En effet, nous verrons que les processus α -stables englobent bon nombre de « processus classiques » comme les processus gaussiens ou encore les processus de Cauchy. L'exclusion du cas Cauchy ($\alpha = 1$) et gaussien ($\alpha = 2$) voit naître l'utilisation d'autres méthodes moins courantes mais qui prennent tout leur sens, notamment pour le traitement du signal audio, vis-à-vis d'expériences préliminaires effectuées il y a quelques années. L'apport majeur de la théorie α -stable est de proposer des distributions à queues lourdes pour la séparation de sources sonores. Plus précisément, la théorie introduit une séparation plus robuste en présence de bruit impulsif ou encore de sources audiofréquences ayant des caractéristiques sonores éloignées.

2.1 Distributions α -stables

2.1.1 Contexte historique

Les premières applications de la théorie α -stable furent effectuées par Mandelbrot [34, 125] dans les années 60, afin de modéliser des séries temporelles financières. En effet, l'évolution des marchés boursiers était auparavant modélisée par les célèbres mouvements Browniens, qui découlaient des représentations de processus gaussiens. Cependant, afin d'avoir un modèle capable de prendre en compte la possibilité d'un « krach boursier », Mandelbrot proposa un nouveau type de processus appelé *processus autosimilaire α -stable* [153] pouvant refléter l'aspect discontinu et brutal des mouvements boursiers. Par la suite, les premiers travaux utilisant la modélisation probabiliste α -stable ont fait surface dans différentes applications en traitement du signal.

Nous pouvons citer par exemple des travaux développés dans les années 70 afin de modéliser le bruit parfois jugé « impulsif » dans les appels téléphoniques [168]. Toutes ces études de statistiques robustes ont été poursuivies [65, 89], afin d'améliorer l'utilisation de la théorie α -stable dans de multiples problèmes classiques en traitement du signal [161, 171].

2.1.2 Distributions α -stables symétriques réelles

La théorie des distributions stables univariées fut introduite et étudiée en premier par Paul Lévy dans les années 30 [109]. Donnons en premier lieu la définition historique d'une variable stable et symétrique :

Définition 2.1.1. Une variable aléatoire x est dite **stable** si pour tous nombres $a, b > 0$, il existe $c > 0$ et un nombre réel d tel que

$$ax^{(1)} + bx^{(2)} \stackrel{d}{=} cx + d, \quad (2.1)$$

où $x^{(1)}$ et $x^{(2)}$ sont indépendants et ont la même distribution que x et $\stackrel{d}{=}$ définit l'égalité entre deux distributions. De plus, si les variables $x \stackrel{d}{=} -x$, on dira que x est **symétrique stable**.

Il en découle un théorème permettant de caractériser les distributions stables [45] :

Théorème 2.1.1. Soit x une variable aléatoire stable. Il existe un nombre $\alpha \in]0, 2]$, reliant les constantes a, b et c de l'équation (2.1) :

$$c^\alpha = a^\alpha + b^\alpha. \quad (2.2)$$

Le nombre α est appelé **exposant caractéristique**.

De ce fait, une variable aléatoire stable associée à un exposant caractéristique α sera dite **α -stable**. En outre, une variable aléatoire α -stable et symétrique sera dite **symétrique α -stable**. Les deux exemples suivants illustrent l'appartenance des distributions gaussiennes et Cauchy à celles des α -stables (cf. Table 1.1).

1. **Distribution gaussienne** ($\alpha = 2$) : Nous rappelons que \mathcal{N} représente la loi gaussienne. Si $x \sim \mathcal{N}(\mu, \sigma^2)$ avec μ sa moyenne et σ^2 sa variance, alors l'équation (2.1) est vérifiée pour $c = \sqrt{a^2 + b^2}$ et $d = (a + b - c)\mu$. En effet :

$$\begin{aligned} cx + d &\sim \mathcal{N}\left(\left[\sqrt{a^2 + b^2} + (a + b - c)\right]\mu, (a^2 + b^2)\sigma^2\right) \\ \Leftrightarrow cx + d &\sim \mathcal{N}\left((a + b)\mu, (a^2 + b^2)\sigma^2\right) \\ \Leftrightarrow cx + d &\stackrel{d}{=} ax^{(1)} + bx^{(2)}. \end{aligned}$$

On peut donc affirmer qu'une loi gaussienne est une *loi 2-stable*.

2. **Distribution de Cauchy** ($\alpha = 1$) : Pour rappel, C désigne la loi de Cauchy. Si $x \sim C(\mu, \sigma)$ avec μ son paramètre de position et σ son paramètre d'échelle alors, l'équation 2.1 est vérifiée pour $c = a + b$ et $d = 0$. En effet :

$$\begin{aligned} cx + d &\sim C((a + b)\mu, (a + b)\sigma) \\ \Leftrightarrow cx + d &\stackrel{d}{=} ax^{(1)} + bx^{(2)}. \end{aligned}$$

Ainsi, une loi de Cauchy est une *loi 1-stable*.

Les deux distributions présentées précédemment font partie des quelques rares cas de distributions α -stables à admettre une expression analytique de leur densité de probabilité. Dans le cas général, la densité de probabilité ne peut être explicitement exprimée comme une fonction analytique [188].

Remarque : Des travaux récents ont montré que la densité de probabilité d'une distribution α -stable peut s'écrire grâce à une fonction plus générale : la *fonction H* de Fox [129, 148]. Cette dernière est elle-même une généralisation de la *fonction G*, proposée par Meijer dans les années 40 [131]. La fonction H fut l'objet de bon nombre d'applications en sciences appliquées, car elle généralise parfaitement de nombreuses fonctions classiques (fonction de Bessel [184], fonctions hypergéométriques [183] ou encore les fonctions de Krätzel [99]) utiles en ingénierie.

Malgré la non-existence d'une forme analytique pour la densité de probabilité d'une loi α -stable, la fonction caractéristique admet pour sa part une expression analytique :

Une loi stable avec $d = 0$ est appelée loi strictement stable.

Définition 2.1.2. Une variable aléatoire x est dite stable, au sens de la définition 2.1.1, s'il existe quatre paramètres $0 < \alpha \leq 2$, $\sigma \geq 0$, $-1 \leq \beta \leq 1$ et $\mu \in \mathbb{R}$ tels que sa fonction caractéristique φ soit de la forme suivante :

$$\forall \theta \in \mathbb{R}, \varphi_x(\theta) = \begin{cases} \exp(-\sigma^\alpha |\theta|^\alpha (1 - i\beta \operatorname{sig}(\theta) \tan(\frac{\pi\alpha}{2})) + i\mu\theta) & \text{si } \alpha \neq 1 \\ \exp(-\sigma^\alpha |\theta|^\alpha (1 + i\beta \frac{2}{\pi} \operatorname{sig}(\theta) \ln(|\theta|)) + i\mu\theta) & \text{si } \alpha = 1 \end{cases}, \quad (2.3)$$

avec $\theta \mapsto \operatorname{sig}(\theta) \in \{-1, 0, 1\}$ la fonction signe. Les paramètres sont respectivement l'**exposant caractéristique** α , le **paramètre d'échelle** σ , le **paramètre d'asymétrie** β et le **paramètre de localisation** μ de la variable x . On notera alors que $x \sim S\alpha(\sigma^\alpha, \beta, \mu)$. Si x est symétrique ($\beta = 0$) et centrée ($\mu = 0$), alors φ s'écrit comme suit :

$$\forall \theta \in \mathbb{R}, \varphi_x(\theta) = \exp(-\sigma^\alpha |\theta|^\alpha) \quad (2.4)$$

et x ne dépend plus que de α et de σ . La notation adoptée pour une telle distribution sera $x \sim S\alpha S(\sigma^\alpha)$.

Les quatre paramètres peuvent être interprétés de la façon suivante [153] : (voir Fig. 2.1) :

- α : Contrôle la lourdeur des queues des distributions α -stables. Plus α est petit, plus les distributions seront dites à queues lourdes. L'expression asymptotique des queues d'une variable $x \sim S\alpha(\sigma^\alpha, \beta, \mu)$ est donnée par :

$$\begin{cases} \lim_{\lambda \rightarrow +\infty} \lambda^\alpha \mathbb{P}(X > \lambda) = C_\alpha \frac{1+\beta}{2} \sigma^\alpha \\ \lim_{\lambda \rightarrow -\infty} \lambda^\alpha \mathbb{P}(X < -\lambda) = C_\alpha \frac{1-\beta}{2} \sigma^\alpha \end{cases} \quad (2.5)$$

où

$$C_\alpha = \begin{cases} \frac{1-\alpha}{\Gamma(2-\alpha)\cos(\frac{\pi\alpha}{2})} & \text{si } \alpha \neq 1 \\ \frac{2}{\pi} & \text{si } \alpha = 1 \end{cases} \quad (2.6)$$

- μ : Au même titre que la moyenne pour une distribution gaussienne, μ localise « globalement » la position du mode (maximum de la densité de probabilité d'une variable aléatoire) de la distribution. Plus α est petit et $|\beta|$ grand, moins le paramètre μ joue le rôle de localisation du mode.
- σ : Ce paramètre positif, appelé paramètre d'échelle, joue un rôle similaire à l'écart-type dans le cas gaussien. Il caractérise l'étalement de la distribution par rapport à sa localisation μ .
- β : Définit l'asymétrie de la distribution. Si $\beta = 0$, la distribution est symétrique autour de μ . Si $\beta > 0$, l'asymétrie apparaît vers la droite. Si $\beta < 0$, l'asymétrie est orientée vers la gauche.

Dans la suite, nous ne considérerons que des distributions symétriques α -stables (dites $S\alpha S$). Énonçons diverses propriétés intéressantes de ces distributions :

Proposition 2.1.1. Soit $x \sim S\alpha S(\sigma_x^\alpha)$, $y \sim S\alpha S(\sigma_y^\alpha)$ deux variables symétriques α -stables indépendantes et $a \in \mathbb{R}^*$. Alors :

1. $x + y \sim S\alpha S(\sigma_x^\alpha + \sigma_y^\alpha)$;
2. $ax \sim S\alpha S(|a|^\alpha \sigma_x^\alpha)$.

À l'instar d'une distribution gaussienne ou de Cauchy, la proposition précédente démontre qu'une combinaison linéaire de variables $S\alpha S$ reste une variable $S\alpha S$. D'une certaine manière, les distributions α -stables définissent presque sûrement un espace vectoriel. Par ailleurs, le théorème central limite se généralise également dans le cas α -stable :

Dans le cas gaussien, il s'agit du théorème central limite classique.

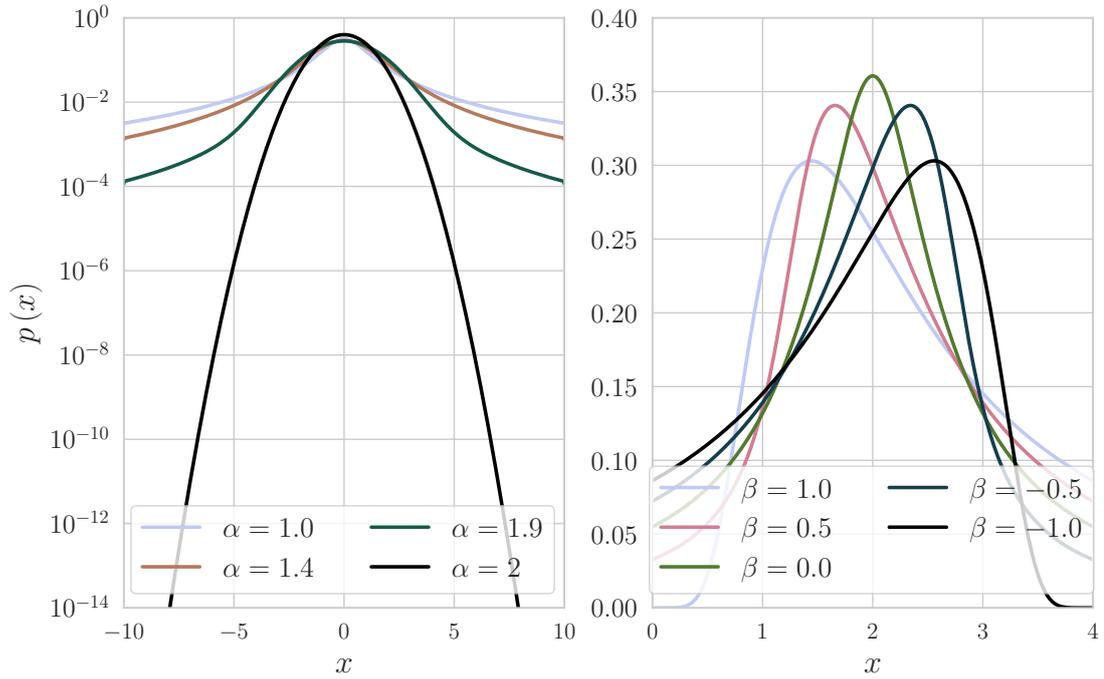


FIGURE 2.1 – À gauche, densités de probabilités de variables $S\alpha S(\sigma^\alpha)$ avec $\sigma^\alpha = 1$. À droite, densités de probabilités de variables $S\alpha(\sigma^\alpha, \beta, \mu)$ avec $\alpha = 0.9$, $\sigma^\alpha = 1$ et $\mu = 2$.

Théorème 2.1.2 (Théorème central limite généralisé). x est stable si et seulement s'il existe une suite de variables aléatoires i.i.d. y_1, y_2, \dots , une suite de nombres positifs $(u_n)_n$ et une suite de nombres réels $(v_n)_n$ tels que :

$$\frac{y_1 + \dots + y_n}{u_n} + v_n \xrightarrow{d} x, \quad (2.7)$$

où \xrightarrow{d} signifie « convergence en distribution ».

Ce théorème, prouvé par Gnedenko & Kolmogorov [97] nous affirme que la stabilité par somme est vérifiée mais avec une convergence en distribution dans l'espace vectoriel des distributions α -stables. Le théorème central limite classique suppose néanmoins que les y_m dans (2.7) ont une variance finie.

2.1.3 Exposant de Lévy : une caractérisation de la parcimonie

Dans la littérature, différentes représentations de processus stochastiques ont été proposées afin de les classifier [169, 172]. L'une d'entre elles consiste à prendre en compte l'aspect dit « parcimonieux ». Des données sont dites parcimonieuses si elles contiennent un nombre important de valeurs proches (ou égales) à zéro [40]. Dans le cas où toute réalisation d'un processus stochastique admet une représentation parcimonieuse, on dira alors qu'il s'agit d'un *processus parcimonieux*. Reprenons maintenant l'attente de la représentation d'un processus α -stable décrit dans la section 2.1.1. Nous souhaitons qu'un tel processus admette des réalisations où un grand nombre de valeurs acquises sont négligeables face aux valeurs aberrantes. Les processus α -stables non-gaussiens sont donc intuitivement des processus parcimonieux (cf Fig. 2.2). Le lecteur pourra trouver la démonstration de l'affirmation précédente dans [172].

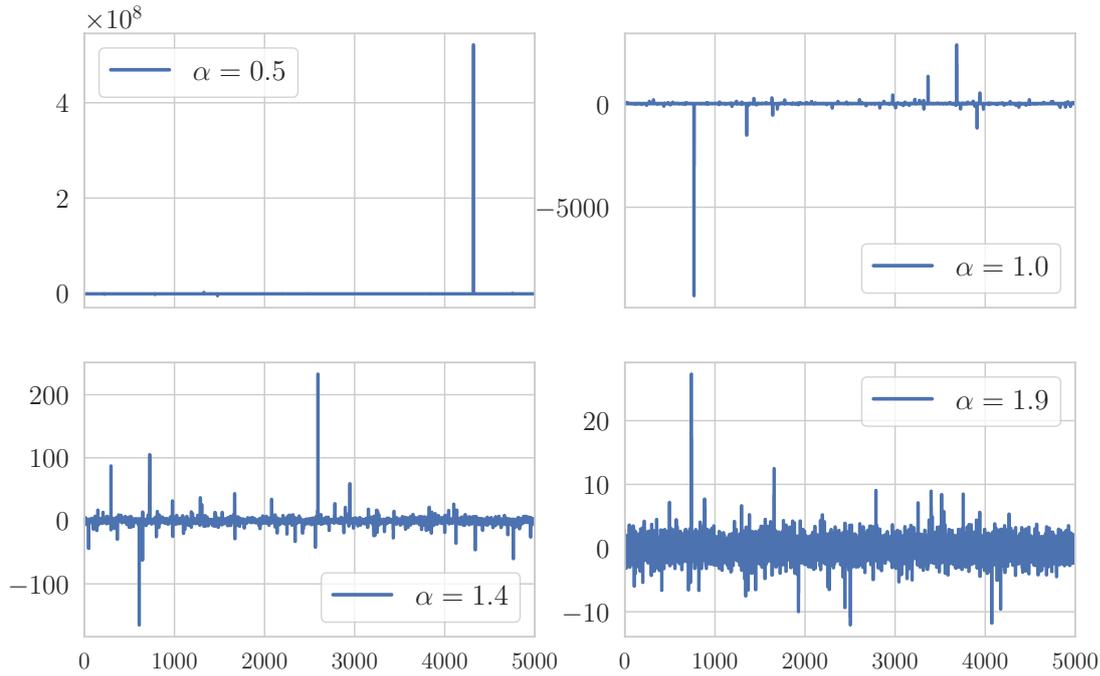


FIGURE 2.2 – Réalisation de distributions scalaires $S\alpha S$: plus α est petit, plus l'échantillonnage devient « parcimonieux ».

Afin de caractériser au mieux le critère de parcimonie, la notion d'exposant de Lévy a été définie de la manière suivante [172] :

Définition 2.1.3 (Exposant de Lévy). *On dit qu'une fonction continue $f : \mathbb{R} \rightarrow \mathbb{C}$ avec $f(0) = 0$ est un exposant de Lévy si pour tout choix possible de $\theta_1, \dots, \theta_N \in \mathbb{R}$ et $v_1, \dots, v_N \in \mathbb{C}$ avec $\sum_{n=1}^N v_n = 0$:*

$$\sum_{n'=1}^N \sum_{n=1}^N f(\theta_{n'} - \theta_n) v_{n'} v_n^* \geq 0, \quad (2.8)$$

où \cdot^* est le conjugué d'un nombre complexe.

Intuitivement, un exposant de Lévy représente une fonction définie positive d'ordre 1, c'est-à-dire fournissant la « corrélation » entre deux événements sur la trajectoire d'un processus ainsi que les valeurs (ou énergies) prises. En soi, il décrit comment la parcimonie opère sur le processus considéré. Dans le cas des distributions α -stables, il a été démontré le théorème suivant dans [172] :

Théorème 2.1.3 (Lévy-Schoenberg). *Soit φ_x , la fonction caractéristique d'une variable stable x . Alors,*

$$\forall \theta \in \mathbb{R}, I(\theta) = -\log \varphi_x(\theta) \quad (2.9)$$

est un exposant de Lévy, au sens de la définition 2.1.3. La fonction \log représente le logarithme d'un nombre complexe.

En réalité, ce résultat s'étend à une classe plus grande de distributions appelée lois *infiniment divisibles*³. Le potentiel de cet objet est important dans le cadre du traitement du signal. En effet, l'exposant

3. x est une loi infiniment divisible si pour tout $N \geq 0$, il existe des variables aléatoires i.i.d. x_1, \dots, x_N telles que $x \stackrel{d}{=} \sum_{n=1}^N x_n$.

de Lévy décrit alors la répartition des valeurs aberrantes d'un signal considéré.

2.1.4 Distributions α -stables symétriques complexes isotropes

Commençons par généraliser la définition 2.1.1 au cas multivarié [153] :

Définition 2.1.4. Un vecteur \mathbf{x} est dit **multivarié α -stable**, si pour $a, b > 0$ il existe un nombre $c \in \mathbb{R}$ et un vecteur \mathbf{d} tels que :

$$a\mathbf{x}^{(1)} + b\mathbf{x}^{(2)} \stackrel{d}{=} c\mathbf{x} + \mathbf{d}, \quad (2.10)$$

où $\mathbf{x}^{(1)}$ et $\mathbf{x}^{(2)}$ ont la même distribution que \mathbf{x} . Si de plus $\mathbf{x} \stackrel{d}{=} -\mathbf{x}$, on dira que \mathbf{x} est multivarié symétrique α -stable.

L'analyse spectrale en traitement du signal audio prend en compte l'information dite d'amplitude et de phase d'une source. Il vient naturellement que ces deux informations, l'une étant positive et l'autre correspondant à une exponentielle complexe, soient numériquement représentées par un nombre complexe. Tout comme les distributions gaussiennes complexes isotropes supposant l'invariance par rotation, la notion de distribution α -stable complexe isotrope est également définie comme suit [153] :

Définition 2.1.5. On dit qu'une variable complexe $x = \Re(x) + i\Im(x)$ est α -stable complexe isotrope si :

1. le vecteur $[\Re(x), \Im(x)]^\top$ est symétrique α -stable comme décrit dans la définition (2.1.4) ;
2. x est invariant par rotation : $\exp(i\omega)x \stackrel{d}{=} x$ pour tout $\omega \in [0, 2\pi[$.

Remarque Dans le cas où, pour une variable complexe x , la condition 2 est remplacée par $-x \stackrel{d}{=} x$, on dira que x est symétrique complexe α -stable anisotrope.

Pareillement à (2.4), une variable α -stable complexe isotrope admet une formule explicite pour sa fonction caractéristique :

Proposition 2.1.2. Soit x une variable α -stable complexe isotrope et φ_x sa fonction caractéristique. Alors :

$$\forall \theta \in \mathbb{C}, \varphi_x(\theta) \triangleq \mathbb{E}(\exp(i\theta^*x)) = \exp(-\sigma^\alpha |\theta|^\alpha). \quad (2.11)$$

La distribution de x est alors caractérisée par son paramètre d'échelle σ et son exposant caractéristique α . On notera $x \sim S\alpha\mathcal{S}_c(\sigma^\alpha)$.

Remarque : Les propriétés de stabilité par combinaison linéaire énoncées dans la proposition 2.1.1 restent valables dans le cas de variables $S\alpha\mathcal{S}_c$.

Pour une variable aléatoire gaussienne complexe isotrope x , la partie réelle et imaginaire de x sont indépendantes [64]. Si $\alpha < 2$, il a été démontré le résultat suivant [153] :

Proposition 2.1.3. Si une variable aléatoire complexe α -stable symétrique est isotrope et non-dégénérée alors ses parties réelle et imaginaire sont dépendantes.

La proposition 2.1.3 implique donc que l'information de phase est davantage conservée car, les parties réelles et imaginaires sont dépendantes. La phase peut donc a fortiori être utilisée dans différents modèles probabilistes pour le traitement du signal. Par ailleurs, il existe un résultat permettant d'affirmer qu'une variable $S\alpha\mathcal{S}_c$ est conditionnellement gaussienne. Définissons au préalable une nouvelle classe de variables α -stables :

Définition 2.1.6. On dit que ϕ suit une loi **positive** α -stable [153, 162] si :

$$\phi \sim S\frac{\alpha}{2} \left(\left[\cos\left(\frac{\pi\alpha}{4}\right) \right]^{2/\alpha}, 1, 0 \right). \quad (2.12)$$

Le support d'une telle distribution est inclus dans \mathbb{R}_+ (c.f. [153, p16-18]). La notation adoptée dans ce manuscrit est alors $\phi \sim \mathcal{P}\alpha\mathcal{S} \left(\left[\cos\left(\frac{\pi\alpha}{4}\right) \right]^{2/\alpha} \right)$.

La Fig. 2.3 affiche plusieurs densités de probabilités de lois positives α -stables. La caractérisation d'une loi $S\alpha\mathcal{S}_c$ comme conditionnellement gaussienne se fait alors comme suit :

Proposition 2.1.4. Une variable $x \sim S\alpha\mathcal{S}_c(\sigma^\alpha)$ est isotrope si et seulement s'il existe une variable gaussienne complexe G de moyenne nulle, une variable $\phi \sim \mathcal{P}\alpha\mathcal{S} \left(2 \left[\cos\left(\frac{\pi\alpha}{4}\right) \right]^{2/\alpha} \right)$ indépendante de G telles que [153] :

$$x = \sqrt{2\phi}G. \quad (2.13)$$

Nous notons alors

$$x | \phi \sim \mathcal{N}_c \left(0, \phi\sigma^2 \right). \quad (2.14)$$

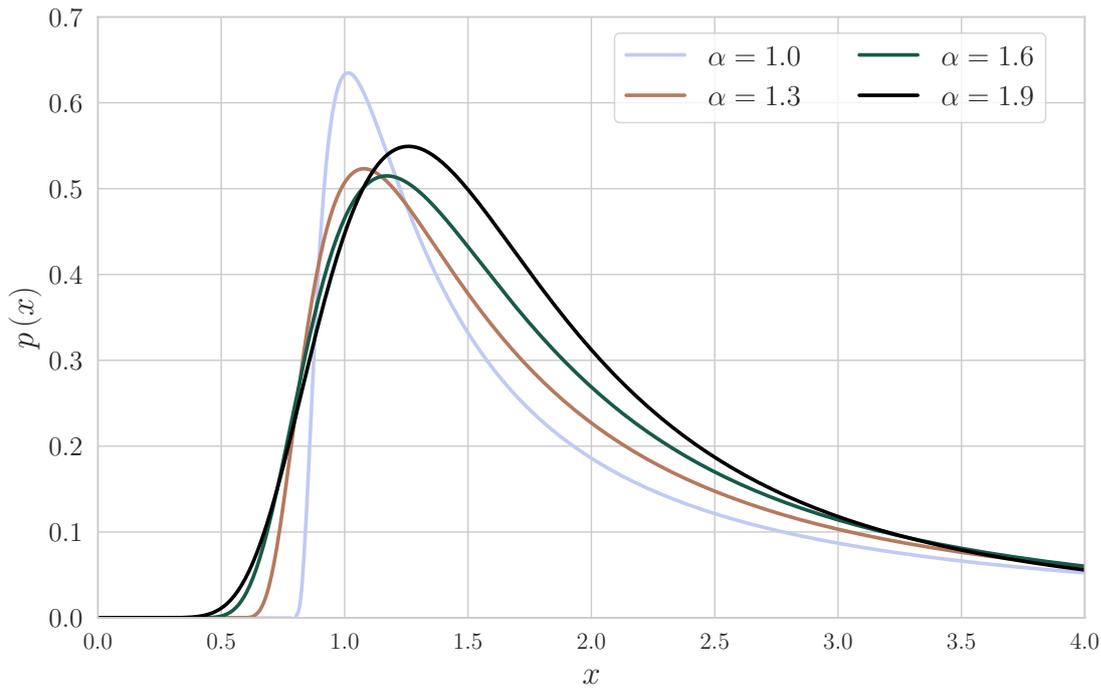


FIGURE 2.3 – Densités de probabilité de lois $\mathcal{P}\alpha\mathcal{S}$.

Remarque L'équation (2.13) permet d'établir une méthode pour échantillonner des variables isotropes α -stables simplement. En effet, il suffit de tirer des gaussiennes centrées et une variable positive α -stable afin d'obtenir des réalisations $S\alpha\mathcal{S}_c$.

Le vecteur $[\Re(x), \Im(x)]^\top$ associé à une variable $x \sim S\alpha S_c(\sigma^\alpha)$ est parfois qualifié dans la littérature de vecteur *sous-gaussien* [153]. Néanmoins, cette définition peut apparaître ambiguë. En effet, un vecteur \mathbf{x} sous-gaussien [77] de dimension K vérifie habituellement l'inégalité suivante sous réserve d'existence de $\sigma \geq 0$:

$$\forall \boldsymbol{\theta} \in \mathbb{R}^K, \mathbb{E}(\exp(\boldsymbol{\theta}^\top \mathbf{x})) \leq \exp\left(-\frac{\sigma^2 \|\boldsymbol{\theta}\|_2^2}{2}\right). \quad (2.15)$$

Les moments exponentiels n'existant pas pour $\alpha < 2$, nous préférons adopter le terme de distribution *multivariée stable à contour elliptique* (MSCE), correspondant davantage à la représentation spatiale (sous forme elliptique, cf. Fig. 13.2) d'un tel vecteur. Cette catégorie de distribution sera en outre appliquée au débruitage dans la partie IV page 107.

Les sections suivantes introduisent les processus α -stables et leurs utilisations pour la théorie du filtrage. Commençons par un cas célèbre dans le traitement du signal : les processus gaussiens.

2.2 Processus gaussiens

Tout d'abord, rappelons quelques résultats sur les distributions gaussiennes. Pour les preuves, le lecteur pourra se référer aux articles et livres suivants : [3, 33, 180].

2.2.1 Processus gaussiens & stationnaires au sens large

Rappelons la définition générale d'un processus gaussien :

Définition 2.2.1. *Un processus stochastique $\mathbf{x} = (x(t))_{t \in \mathcal{T}}$ est dit **gaussien** si chacune de ses lois finidimensionnelles $\mathcal{L}_{\mathbf{x}}$ est gaussienne. Autrement dit, si toute combinaison linéaire de ses marginales $a_1 x(t_1) + \dots + a_T x(t_T)$ est une loi gaussienne pour tous $T \in \mathbb{N}^*$, $t_1, \dots, t_T \in \mathcal{T}$ et $a_1, \dots, a_T \in \mathbb{C}$.*

Nous pouvons ajouter à la définition 2.2.1 que les processus gaussiens sont entièrement décrits par les fonctions ci-dessous :

$$\forall t \in \mathcal{T}, \mu(t) = \mathbb{E}[x(t)], \quad (2.16)$$

$$\forall t, t' \in \mathcal{T}, r(t, t') = \mathbb{E}[(x(t) - \mu(t))(x(t') - \mu(t'))^*] \quad (2.17)$$

appelées *moyenne* et *covariance*. Nous noterons par conséquent un processus gaussien \mathbf{x} de moyenne μ et de covariance r comme suit :

$$\mathbf{x} \sim \mathcal{PN}(\mu, r). \quad (2.18)$$

Considérons l'ensemble $\tau = \{t_1, \dots, t_T\} \in \mathcal{T}$. Ainsi, tout échantillon issu de la trajectoire d'un processus gaussien $\mathbf{x}_\tau \triangleq [x(t_1), \dots, x(t_T)]^\top$ est distribué comme suit :

$$\begin{bmatrix} x(t_1) \\ \vdots \\ x(t_T) \end{bmatrix} \sim \mathcal{N}^N \left(\underbrace{\begin{bmatrix} \mu(t_1) \\ \vdots \\ \mu(t_T) \end{bmatrix}}_{\triangleq \boldsymbol{\mu}}, \underbrace{\begin{bmatrix} r(t_1, t_1) & \dots & r(t_1, t_N) \\ \vdots & \ddots & \vdots \\ r(t_N, t_1) & \dots & r(t_N, t_N) \end{bmatrix}}_{\triangleq \mathbf{R}_{T,x}} \right) \quad (2.19)$$

autrement écrit

$$\mathbf{x}_\tau \sim \mathcal{N}^T(\boldsymbol{\mu}, \mathbf{R}_{T,x}), \quad (2.20)$$

où μ et $R_{T,x}$ sont respectivement le vecteur *moyenne* et la matrice de *covariance* du vecteur gaussien $x(v)$. Un choix classique de modélisation de la fonction de covariance exponentielle carrée présentée ci-dessous [180] :

$$\forall t, t' r(t, t') = \sigma^2 \exp\left(-\frac{(t-t')^2}{\lambda^2}\right). \quad (2.21)$$

En effet, tandis que la moyenne d'un processus gaussien joue le rôle d'indicateur de la valeur la plus plausible que peut prendre le processus, la covariance peut avoir différents rôles. La modélisation de la covariance présentée en (2.21) indique que la corrélation entre deux points proches sera grande. En d'autres termes, que la dépendance entre deux valeurs proches est importante, comparée à deux valeurs éloignées (cf Fig. 2.4).

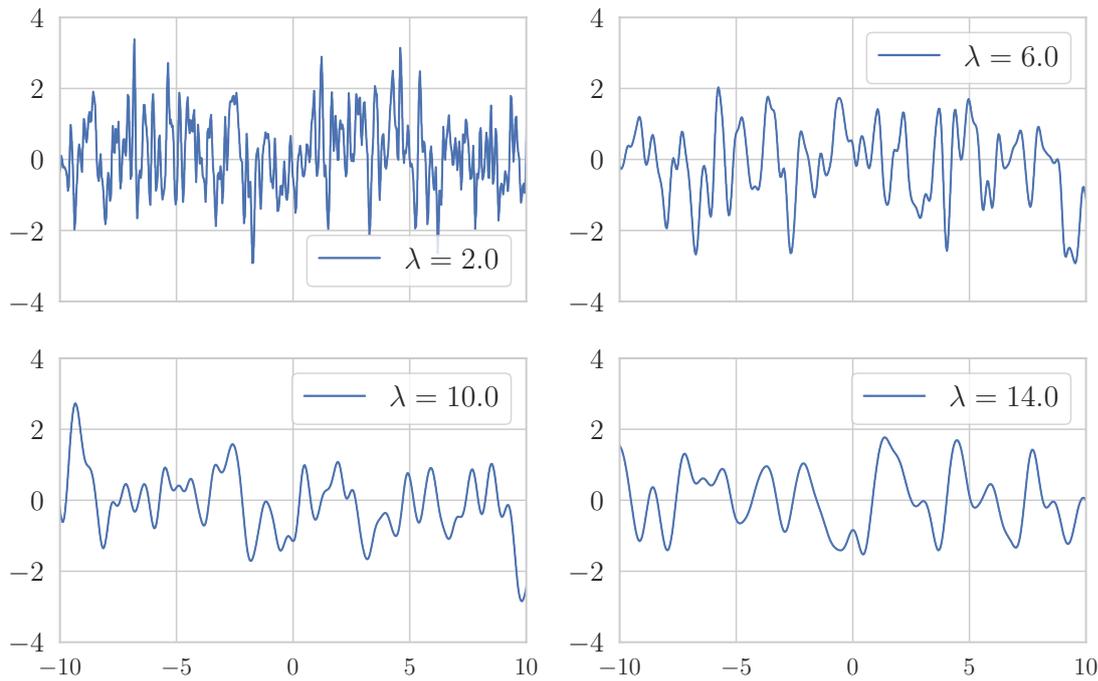


FIGURE 2.4 – Réalisations d'un processus gaussien avec la covariance décrite en (2.21) pour différents λ : plus λ augmente, plus le signal semble corrélé.

Il existe bien entendu d'autres façons de modéliser la covariance [117, 180]. Celle en (2.21) reste cependant la plus naturelle dans le cas de modélisation en audio [119]. Présentons maintenant une classe plus spécifique de processus :

Définition 2.2.2. Soit $x = (x(t))_{t \in \mathcal{T}}$ un processus stochastique. On dit que x est *stationnaire au sens strict* si pour tout $T \in \mathbb{N}^*$, $t_1, \dots, t_T \in \mathcal{T}$ et $\tau \in \mathbb{R}$:

$$(x(t_1), \dots, x(t_T)) \stackrel{d}{=} (x(t_1 + \tau), \dots, x(t_T + \tau)). \quad (2.22)$$

Si de plus x admet une moyenne $t \rightarrow \mu(t)$ et une covariance $(t, t') \rightarrow r(t, t')$, on dira que x est *stationnaire au sens large (SSL)* si :

1. sa moyenne $\mathbb{E}(x(t)) = \mu$ ne dépend pas de t ;

2. $(t, t') \rightarrow r(t, t')$ est une fonction dépendant uniquement de la différence $t - t'$.

Un processus gaussien et SSL sera appelé processus **gaussien stationnaire au sens large**.

Nous montrons par la suite que de tels processus permettent d'obtenir des résultats intéressants pour l'analyse d'un signal dans le domaine fréquentiel [38, 155].

2.2.2 Analyse spectrale dans le cas SSL

Nous justifions dans cette section le lien pouvant être établi entre les processus gaussien SSL et l'analyse spectrale. Supposons en premier lieu que pour un signal observé \mathbf{x} , les observations sont en nombre fini et sont des échantillons issus d'une réalisation d'un processus gaussien SSL $\mathbf{x} \sim \mathcal{PN}(\boldsymbol{\mu}, r)$. Soit $\nu = \{t_1, \dots, t_T\} \in \mathcal{T}$ et T observations $\mathbf{x}_\nu = [x(t_1), \dots, x(t_T)]$ à valeurs dans \mathbb{R} , échantillonnées de manière régulière. Alors

$$\mathbf{x}_\nu \sim \mathcal{N}^T(\boldsymbol{\mu}, \mathbf{R}_{T,x}) \quad (2.23)$$

où, grâce à l'hypothèse de stationnarité, la matrice de covariance $\mathbf{R}_{T,x}$ est de la forme

$$\mathbf{R}_{T,x} = \begin{bmatrix} r(0) & r(-1) & \dots & r(-(T-1)) \\ r(1) & r(0) & \dots & r(-(T-2)) \\ \vdots & \dots & \ddots & \vdots \\ r(T-1) & r(T-2) & \dots & r(0) \end{bmatrix}. \quad (2.24)$$

Ce type de matrice, plus communément appelée *matrice de Toeplitz*, admet une propriété remarquable. Considérons une base de Fourier :

$$\forall l, l' \in \llbracket 1, L \rrbracket, [F_T]_{l,l'} = \frac{1}{\sqrt{L}} \exp\left(-2i\pi \frac{(l-1)(l'-1)}{T}\right) \quad (2.25)$$

et soit $\mathbf{r}_T \triangleq [r(0), \dots, r(-(T-1))]^\top$ la première colonne de $\mathbf{R}_{T,x}$. La propriété suivante est alors vérifiée :

Proposition 2.2.1. *Si le nombre de composantes N est suffisamment grand, la matrice $\mathbf{R}_{T,x}$ est diagonalisable de la manière suivante dans la base orthonormée \mathbf{F}_T , composée des coefficients de l'équation (2.25) :*

$$\mathbf{R}_{T,x} \simeq \mathbf{F}_T^* \text{diag}(\mathbf{F}_T \mathbf{r}_T) \mathbf{F}_T, \quad (2.26)$$

où \simeq en (2.26) est l'opérateur d'équivalence entre deux matrices.

Ce résultat découle du fait qu'une *matrice de Toeplitz* très grande esquissée en (2.24) est approximativement équivalente [70] à une *matrice circulante*. De plus, une matrice circulante est toujours diagonalisable dans la base de Fourier [68].

De ce théorème découlent deux résultats très importants :

- Les coefficients de la matrice $\text{diag}(\mathbf{F}_T \mathbf{r}_T)$ sont ceux de la transformée de Fourier de \mathbf{r}_T . Le théorème de Bochner [152] nous assure qu'ils sont positifs. Nous les noterons $\sigma^2(f)$. Appelés en traitement du signal *densité spectrale de puissance* (DSP) [166], ils traduisent la puissance du signal pour une fréquence donnée.
- Supposons que la transformée de Fourier à temps discret contienne un nombre fini de termes. Alors, la transformée de Fourier de \mathbf{x}_ν , que nous noterons abusivement $\mathbf{x}(F) \triangleq [x(f)]_{1 \leq f \leq F}$, est un vecteur gaussien isotrope centré où chaque marginale est indépendante des autres :

$$\forall f, x(f) \sim \mathcal{N}_c\left(0, \sigma^2(f)\right). \quad (2.27)$$

On dit alors que la transformée de Fourier « blanchit » le vecteur \mathbf{x}_ν .

Ainsi, l'analyse spectrale d'un processus gaussien SSL décrit une conservation de ses lois finies dimensionnelles. Dans le cadre de modélisation d'un signal, il apparaît plutôt naturel de l'analyser sur plusieurs trames temporelles. Pour ce faire, nous utilisons la *transformée de Fourier à court terme* (TFCT) sur les observations. Soit le vecteur $\mathbf{x}_T = [x(t)]_{1 \leq t \leq T}$. Nous appliquons la TFCT sur ce vecteur, en considérant une fenêtre $\tau \rightarrow w(\tau)$ de longueur finie. En pratique, cela implique que nous obtenions :

$$\forall (f, t) \in \llbracket 1, F \rrbracket \times \llbracket 1, T_{\text{TFCT}} \rrbracket, x(f, t) = \sum_{n=1}^N x(n) w(n-t) \exp(-2i\pi fn). \quad (2.28)$$

Le couple (f, t) est appelé *point temps-fréquence* (TF). En omettant le recouvrement entre les différentes trames temporelles causées par (2.28), il en découle que :

$$\forall (f, t), x(f, t) \sim \mathcal{N}_c(0, \sigma^2(f, t)) \quad (2.29)$$

et que les variables $x(f, t)$ peuvent être supposées indépendants entre eux. Grâce à (2.29), les coefficients $|x(f, t)|^2$ peuvent être interprétés comme une estimation de la DSP du processus Gaussien. L'ensemble des coefficients $|x(f, t)|^2$ forment la *spectrogramme de puissance* du signal (cf. Fig. 2.5) .

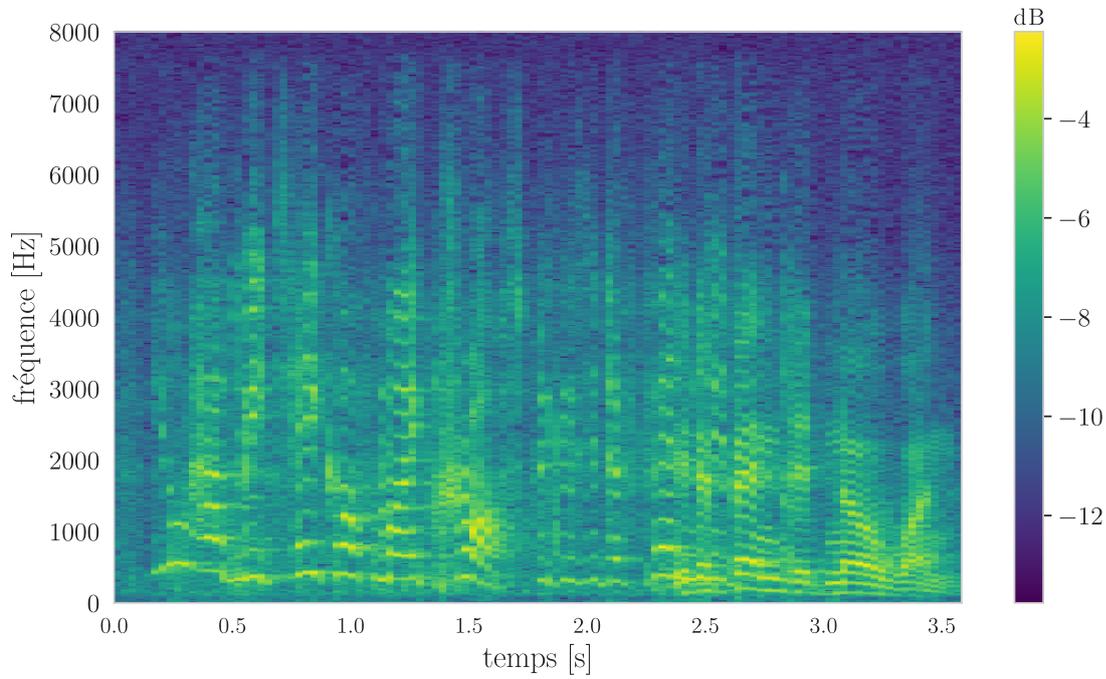


FIGURE 2.5 – log-spectrogramme d'un signal de parole.

Nous pouvons alors nous demander si une telle représentation d'un processus gaussien SSL dans le domaine TF est valable dans le cas d'un processus α -stable quelconque.

2.3 Processus α -stables

Considérons les ensembles $E \triangleq (\mathbb{R}, \mathcal{B}, \mathbb{P})$ et $E_C \triangleq (\mathbb{R}^2, \mathcal{B}^2, \mathbb{P})$, où \mathcal{B}^K est la tribu Borélienne engendrée par les pavés de \mathbb{R}^K .

2.3.1 Processus symétrique α -stable

Il existe plusieurs définitions équivalentes d'un processus α -stable. La première est la plus classique et intuitive :

Définition 2.3.1. *Un processus stochastique $x = (x(t))_{t \in \mathcal{T}}$ est dit **symétrique α -stable** (noté $S\alpha S$) si chacune de ses lois fini-dimensionnelles \mathcal{L}_x définit une loi symétrique α -stable. Autrement dit, si toute combinaison linéaire de ses marginales $a_1 x(t_1) + \dots + a_T x(t_T)$ est une loi symétrique α -stable, pour tous $T \in \mathbb{N}^*$, $t_1, \dots, t_T \in \mathcal{T}$ et $a_1, \dots, a_T \in \mathbb{C}$.*

La définition précédente peut paraître convenable de part sa simplicité. Cependant, cette représentation d'un processus $S\alpha S$ n'est pas suffisante pour en déduire une représentation spectrale équivalente à celle d'un processus gaussien SSL [153]. Pour ce faire, nous allons utiliser une définition alternative, mettant en jeu la notion de mesure aléatoire pour caractériser les processus α -stables. Afin d'alléger la présentation, nous nous restreignons au cas où la mesure de contrôle circulaire k de la définition 1.2.5 s'écrit :

$$k = m \times \gamma \quad (2.30)$$

où γ est la mesure de probabilité uniforme sur le cercle unité, \times est le produit entre deux mesures et m une mesure quelconque. Par souci de clarté, nous introduisons une mesure aléatoire complexe $S\alpha S$ dans ce manuscrit comme suit :

Définition 2.3.2. *Considérons les notations de la définition 1.2.5 page 7. Soit k une mesure de contrôle circulaire définie en (2.30) sur $(F \times S^1, \mathcal{A} \otimes \mathcal{B}^2)$ avec \mathcal{B}^2 la tribu borélienne engendrée par les pavés de \mathbb{R}^2 , S^1 le cercle unité et $F = (\mathcal{A}, \Sigma, m)$ un espace mesurable. Notons par ailleurs $\mathcal{R}_{\mathbb{C}}$ l'ensemble des variables aléatoires réelles définies sur \mathbb{C} . Soit M une mesure aléatoire complexe. Alors $M : \mathcal{A} \rightarrow \mathcal{R}_{\mathbb{C}}$ est qualifiée d'être **α -stable complexe** (ou mesure aléatoire complexe $S\alpha S$) de mesure de contrôle m si d'une part :*

$$\forall A \in \mathcal{A}, M(A) \sim S\alpha S(m(A)) \quad (2.31)$$

et d'autre part, M est à incréments indépendants et σ -additive. Cela signifie que

1. pour $A_1, \dots, A_n \in \mathcal{A}$ disjoints, les variables $M(A_1), \dots, M(A_n)$ sont indépendantes ;
2. pour toute union dénombrable disjointe, $\bigcup_{n=1}^{+\infty} A_n \in \mathcal{A}$, $M(\bigcup_{n=1}^{+\infty} A_n) = \sum_{n=1}^{+\infty} M(A_n)$ p.s.

De cette définition, il peut être alors démontré la représentation suivante d'un processus symétrique α -stable :

Théorème 2.3.1. *Soit M une mesure aléatoire $S\alpha S$, de mesure de contrôle circulaire $k = m \times \gamma$. Alors, pour toute fonction $\phi : F \rightarrow \mathbb{C}$ telle que $\int_F |\phi(x)|^\alpha m(dx) < +\infty$, l'intégrale*

$$x(\phi) \triangleq \int_{f \in F} \phi(f) M(df) \quad (2.32)$$

est un processus $S\alpha S$. Comme dans la définition 2.3.1, nous utilisons également l'appellation « x est un processus $S\alpha S$ ».

Rappelons que nous cherchons à obtenir une transformée de Fourier de processus α -stables conservant les lois finies dimensionnelles comme dans le cas gaussien SSL. Introduisons au préalable une nouvelle classe de processus stochastiques : les processus α -harmonisables.

2.3.2 Processus α -harmonisables

Nous définissons une mesure aléatoire isotrope M , comme vérifiant :

$$\exp(i\phi) M \stackrel{d}{=} M \forall \phi \in \mathbb{R}. \quad (2.33)$$

Il apparaît alors l'équivalence ci-dessous :

Théorème 2.3.2. *M est une mesure aléatoire isotrope avec une mesure de contrôle circulaire k si et seulement si :*

$$k = m \times \gamma, \quad (2.34)$$

où γ est la mesure de probabilité uniforme sur le cercle unité, \times est le produit d'une mesure produit et m est une mesure sur (F, \mathcal{A}) , appelée naturellement **mesure de contrôle** de M .

La définition (2.3.2) est donc équivalente à la représentation en (2.33) et les notions de mesure aléatoire $S\alpha S$ complexe et mesure aléatoire isotrope peuvent être confondues. Nous pouvons en outre définir, sans perte de généralité sur la mesure de contrôle circulaire, le processus stochastique suivant :

Définition 2.3.3 (processus α -harmonisable). *On dira que $\mathbf{x} = (x(t))_{t \in \mathcal{T}}$ est un **processus α -harmonisable** s'il existe une mesure de contrôle m associée à une mesure aléatoire isotrope M tels que :*

$$\forall t \in \mathcal{T}, x(t) = \int_{-\infty}^{+\infty} \exp(itf) M(df). \quad (2.35)$$

Autrement dit, \mathbf{x} est la transformée de Fourier inverse d'une mesure aléatoire isotrope. Il en découle par ailleurs le résultat remarquable suivant :

Théorème 2.3.3. *Soit M une mesure aléatoire $S\alpha S$ quelconque et soit \mathbf{x} , un processus α -harmonisable. Si M est une mesure aléatoire isotrope $S\alpha S_c$ alors \mathbf{x} est stationnaire. Autrement dit, un processus α -harmonisable est stationnaire si chacune des variables aléatoires pour chaque tirage du processus α -harmonisable suit une loi $S\alpha S_c$.*

Ainsi pour tout tirage $\mathbf{x}_\tau = [x(t_1), \dots, x(t_T)]^\top$ avec $\tau = \{t_1, \dots, t_T\}$ d'un processus α -harmonisable \mathbf{x} , nous supposons que :

$$\forall t \in \tau, x(t) \sim S\alpha S_c(\sigma_x^\alpha(t)), \quad (2.36)$$

où les $x(t)$ sont indépendants entre eux. La notion de processus harmonisable apparaît dans le cas α -stable comme la généralisation pertinente de la stationnarité au sens large du cas Gaussien, qui permet un traitement efficace dans le domaine fréquentiel, puisque les composantes de son spectre sont indépendantes.

Chapitre 3

Mélanges et séparation de processus α -stables

Résumé

Les processus α -stables ont la propriété de conserver leur nature par combinaison linéaire. Par exemple, une somme de processus α -stables reste un processus α -stable. Dans un contexte de séparation des processus, il est alors légitime de se demander comment extraire chaque processus α -stable d'une somme. En particulier, dans le cadre d'un traitement dans le plan temps-fréquence : est-il possible de récupérer chaque composante α -harmonisable de la somme de processus α -harmonisables ? Pour un mélange de processus gaussiens ($\alpha = 2$) *stationnaires au sens large* (SSL), la réponse est positive et exploite notamment les travaux précurseurs de Norbert Wiener [179]. On peut en effet montrer que minimiser l'erreur quadratique moyenne revient à employer le très célèbre filtre de Wiener pour la séparation. Un cadre plus général de processus gaussiens dits localement stationnaires a été également proposé dans [119], qui permet en somme de généraliser le cas SSL. En outre, une stratégie de séparation de processus α -harmonisables a été proposée dans [118]. Il en découle une méthode de filtrage similaire au cas gaussien et appelé « filtrage de Wiener généralisé ». Dans le cas où $\alpha \neq 2$, il est à noter que les paramètres d'échelles obtenus ne correspondent pas à *une densité spectrale de puissance*, mais à l'énergie « fractionnaire » : au lieu de parler de *densité spectrale de puissance*, nous nommerons les paramètres d'échelles associés à l'analyse spectrale *densité spectrale de puissance fractionnaire*.

3.1 Séparation de processus gaussiens

Afin d'alléger les notations dans cette section, nous nous limitons au cas où le processus gaussien $\mathbf{x} = (x(t))_{t \in \mathcal{T}}$ est la somme de seulement deux processus gaussiens $\mathbf{s} = (s(t))_{t \in \mathcal{T}}$ et $\mathbf{n} = (n(t))_{t \in \mathcal{T}}$:

$$\mathbf{x} = \mathbf{s} + \mathbf{n}. \quad (3.1)$$

Dans un contexte de débruitage, ce modèle de mélange est très courant (observation = signal cible + bruit additif). Dorénavant dans ce chapitre, l'indice d'échantillonnage ne sera plus noté n mais t , en rapport avec la mesure physique d'un instant temporel. La lettre f quant à elle désignera une fréquence donnée du signal étudié et (f, t) un point *temps-fréquence* (TF).

3.1.1 Cas général

Dans un premier temps, nous nous plaçons dans le cas général où les processus gaussiens ne sont pas forcément stationnaires. Soit $s \sim \mathcal{PN}(\mu_s, r_s)$ et $\mathbf{n} \sim \mathcal{PN}(\mu_n, r_n)$ deux processus gaussiens indépendants respectivement de moyennes μ_s, μ_n et d'opérateurs de covariance r_s, r_n . Alors, $\mathbf{x} = s + \mathbf{n}$ est un processus gaussien de moyenne $\mu_s + \mu_n$ et de covariance $r_s + r_n$. Prenons en compte une trajectoire de chaque processus, à laquelle nous considérons un nombre fini de T instants $\tau = \{t_1, \dots, t_T\} \in \mathcal{T}$. Ces échantillons, supposés indépendants, sont respectivement regroupés dans les vecteurs $\mathbf{x}_\tau \triangleq [x(t_1), \dots, x(t_T)]^\top$, $s_\tau \triangleq [s(t_1), \dots, s(t_T)]^\top$ et $\mathbf{n}_\tau \triangleq [n(t_1), \dots, n(t_T)]^\top$. D'après (2.20), nous savons alors que

$$\mathbf{x}_\tau \sim \mathcal{N}^T(\mu_x(\tau), \mathbf{R}_{T,x}(\tau)); s_\tau \sim \mathcal{N}^T(\mu_s(\tau), \mathbf{R}_{T,s}(\tau)); \mathbf{n}_\tau \sim \mathcal{N}^T(\mu_n(\tau), \mathbf{R}_{T,n}(\tau)). \quad (3.2)$$

Dans un cadre d'inférence Bayésienne, nous souhaitons utiliser les observations pour en déduire la distribution *a posteriori* des sources. Plus précisément, nous supposons connu le vecteur \mathbf{x}_τ : quelle loi *a posteriori* suit alors le vecteur $s_\tau \mid \mathbf{x}_\tau$?

Proposition 3.1.1. *Soit $\mathbf{x}_\tau \sim \mathcal{N}^T(\mu_x(\tau), \mathbf{R}_{T,x}(\tau))$ et $s_\tau \sim \mathcal{N}^T(\mu_s(\tau), \mathbf{R}_{T,s}(\tau))$, deux vecteurs aléatoires multivariés gaussiens. Alors, la loi jointe de s_τ et \mathbf{x}_τ vaut*

$$\begin{bmatrix} s_\tau \\ \mathbf{x}_\tau \end{bmatrix} \sim \mathcal{N}^{2T} \left(\begin{bmatrix} \mu_s(\tau) \\ \mu_x(\tau) \end{bmatrix}, \begin{bmatrix} \mathbf{R}_{T,s}(\tau) & \mathbf{R}_{T,sx}(\tau) \\ \mathbf{R}_{T,xs}(\tau) & \mathbf{R}_{T,x}(\tau) \end{bmatrix} \right) \quad (3.3)$$

où $\mathbf{R}_{T,sx}, \mathbf{R}_{T,xs}$ sont respectivement les matrices de covariances de s_τ et \mathbf{x}_τ et de \mathbf{x}_τ et s_τ . De plus, la distribution *a posteriori* $s_\tau \mid \mathbf{x}_\tau$ est gaussienne :

$$s_\tau \mid \mathbf{x}_\tau \sim \mathcal{N}^T(\mu_{s|x}(\tau), \mathbf{R}_{s|x}(\tau)) \quad (3.4)$$

où

$$\mu_{s|x}(\tau) = \mu_s(\tau) + \mathbf{R}_{T,sx}(\tau) \mathbf{R}_{T,x}^{-1}(\tau) (\mathbf{x}_\tau - \mu_x(\tau)) \quad (3.5)$$

s'appelle la moyenne *a posteriori* et

$$\mathbf{R}_{s|x}(\tau) = \mathbf{R}_{T,s}(\tau) - \mathbf{R}_{T,sx}(\tau) \mathbf{R}_{T,x}^{-1}(\tau) \mathbf{R}_{T,xs}(\tau) \quad (3.6)$$

définit la covariance *a posteriori* [39, 180].

Grâce à la distribution *a posteriori* gaussienne, nous avons d'une part une indication sur la valeur moyenne du tirage s en chaque point temporel, mais également une information d'incertitude de l'estimation de la valeur donnée, par la matrice de covariance *a posteriori* $\mathbf{R}_{s|x}(\tau)$. En l'état, ce modèle n'est pas implémentable car T est en pratique relativement grand (en moyenne, $\approx 10^6$ échantillons pour un morceau de musique). En effet, cela nécessiterait l'inversion des matrices $\mathbf{R}_{T,x}^{-1}(\tau)$ qui est un problème de complexité algorithmique d'ordre $O(T^3)$.

3.1.2 Cas stationnaire au sens large

Plusieurs travaux dans [10, 117] ont décrit différentes méthodes s'affranchissant de la *stationnarité au sens large* (SSL) des processus, notamment en proposant une stationnarité seulement locale. Cette astuce permet d'appliquer la méthode de séparation d'un processus gaussien SSL mais de ne considérer les dépendances temporelles que sur des durées de l'ordre de quelques dizaines de millisecondes [119]. Nous décrivons comment le problème se simplifie dans le cas de processus gaussiens SSL. Supposons maintenant que :

$$s \sim \mathcal{PN}(0, r_s); \mathbf{n} \sim \mathcal{PN}(0, r_n), \quad (3.7)$$

où s et n sont stationnaires (cf. définition 2.2.2). Par stabilité sommatoire, on a

$$\mathbf{x} = \mathbf{s} + \mathbf{n} \sim \mathcal{PN}(0, r_s + r_n). \quad (3.8)$$

D'après la sous-section 2.2.2 page 18, nous pouvons considérer les signaux indépendants dans le domaine temps-fréquence comme suit :

$$\forall (f, t), s(f, t) \sim \mathcal{N}_c \left(\sigma_s^2(f, t) \right); \quad (3.9)$$

$$\forall (f, t), n(f, t) \sim \mathcal{N}_c \left(\sigma_n^2(f, t) \right), \quad (3.10)$$

où $\sigma_s^2(f, t)$ et $\sigma_n^2(f, t)$ sont les *densités spectrales de puissance* (DSP) respectives de n et s pour le point TF (f, t) . Par stabilité de sommation de variables gaussiennes, cela implique que

$$\forall (f, t), x(f, t) \sim \mathcal{N}_c \left(\sigma_x^2(f, t) \right) \quad (3.11)$$

où $\sigma_x^2(f, t) = \sigma_s^2(f, t) + \sigma_n^2(f, t)$. En appliquant la propriété 3.1.1 pour chaque $x(f, t)$ et $s(f, t)$, nous en déduisons

$$\begin{bmatrix} s(f, t) \\ x(f, t) \end{bmatrix} \sim \mathcal{N}_c^2 \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_s^2(f, t) & \sigma_s^2(f, t) \\ \sigma_s^2(f, t) & \sigma_s^2(f, t) + \sigma_n^2(f, t) \end{bmatrix} \right) \quad (3.12)$$

$$s(f, t) | x(f, t) \sim \mathcal{N}_c \left(\mu_{s|x}(f, t), \sigma_{s|x}^2(f, t) \right) \quad (3.13)$$

avec

$$\mu_{s|x}(f, t) = \frac{\sigma_s^2(f, t)}{\sigma_s^2(f, t) + \sigma_n^2(f, t)} x(f, t) \quad (3.14)$$

et

$$\sigma_{s|x}^2(f, t) = \sigma_s^2(f, t) - \frac{\sigma_s^4(f, t)}{\sigma_s^2(f, t) + \sigma_n^2(f, t)} = \frac{\sigma_s^2(f, t)\sigma_n^2(f, t)}{\sigma_s^2(f, t) + \sigma_n^2(f, t)}. \quad (3.15)$$

Par conséquent, la moyenne *a posteriori* $\mu_{s|x}(f, t)$ est simplement calculée indépendamment en chaque point TF. Comparé aux équations (3.5) et (3.6), le problème est beaucoup plus simple à traiter car ici il ne s'agit plus d'inversion de matrice mais de calculs scalaires à effectuer pour chaque point TF. Cette technique de filtrage peut être vue comme étant la célèbre méthode appelée *filtrage de Wiener*. Il est à noter que le filtrage de Wiener reste valable dès lors que les moments d'ordre 2 existent. Dans le cas α -stable non-gaussien, il est donc nécessaire d'introduire un filtrage exploitant des moments d'ordre strictement plus petit que l'exposant caractéristique α .

3.1.3 Séparation de variables α -stables complexes isotropes

Étudions comment le filtrage de Wiener peut se généraliser dans le cadre des processus α -stables harmonisables. Tous les travaux présentés dans cette sous-section sont tirés des deux références [118] et [13]. Considérons de prime abord $J + 1$ variables aléatoires x et s_1, \dots, s_J telles que :

$$x = \sum_{j=1}^J s_j; \quad (3.16)$$

$$s_j \sim S\alpha S_c \left(\sigma_j^\alpha \right). \quad (3.17)$$

avec les s_j indépendantes entre eux. Il en découle, par stabilité sommatoire, que

$$x \sim S\alpha S_c \left(\sum_j \sigma_j^\alpha \right). \quad (3.18)$$

Très souvent, nous ne pouvons conclure simplement sur la loi a posteriori de $s_j | x$. Nous avons vu que cette loi *a posteriori* reste gaussienne dans le cas d'un modèle gaussien, mais ce résultat ne se généralise pas dans le cas α -stable, sauf pour certains cas dont les hypothèses ne sont pas adaptées aux modèles proposés dans le cadre du traitement du signal (voir par exemple [153, chap.4]). Alternativement, en utilisant minutieusement l'équation (1.7) reliant l'espérance d'une variable aléatoire et sa fonction caractéristique, nous pouvons en déduire le théorème suivant démontré dans [13] :

Théorème 3.1.1. *Soit des variables aléatoires x et s_1, \dots, s_J comme mentionnées ci-avant dans (3.16) et (3.17). Alors, l'espérance a posteriori $\mathbb{E}[s_j | x]$ est donnée par la relation pour tout $\alpha > 1$:*

$$\mathbb{E}[s_j | x] = \frac{\sigma_j^\alpha}{\sum_j \sigma_j^\alpha} x. \quad (3.19)$$

Ainsi, l'équation (3.19) permet de généraliser le filtrage de Wiener classique du cas gaussien à un α quelconque. En effet, en posant $\alpha = 2$, on retombe exactement sur la moyenne *a posteriori* de (3.5). Grâce à l'équation (3.19), nous sommes en mesure de proposer une méthode de filtrage pour les processus α -harmonisables généralisant celle appliquée sur des processus gaussiens SSL.

3.1.4 Le cas α -harmonisable et DSP fractionnaire

Considérons $\mathbf{s}_1 = (s_1(f, t))_{(f, t) \in \mathcal{F} \times \mathcal{T}}, \dots, \mathbf{s}_J = (s_J(f, t))_{(f, t) \in \mathcal{F} \times \mathcal{T}}$, des représentations spectrales de signaux audiofréquences monophoniques. Modélisons-les comme étant des processus α -harmonisables décrits à la définition 2.3.3 page 21. D'après le théorème 2.3.3 page 21, nous pouvons considérer pour tout j un échantillonnage vectoriel $s_{j, \tau\omega}$ où $\tau \subset \mathcal{T}$ et $\omega \subset \mathcal{F}$. De plus, les marginales de chaque vecteur $s_{j, \tau\omega}$ sont des variables indépendantes et suivant une loi $S\alpha S_c$. Autrement dit :

$$\forall (f, t), s_j(f, t) \sim S\alpha S_c \left(\sigma_j^\alpha(f, t) \right), \quad (3.20)$$

avec les paramètres $\sigma_j^\alpha(f, t)$ appelés *DSP fractionnaire* [118]. La stabilité par somme de variables aléatoires isotropes entraîne par ailleurs que le mélange $x(f, t) = \sum_j s_j(f, t)$ est également isotrope :

$$\forall (f, t), x(f, t) \sim S\alpha S_c \left(\sum_j \sigma_j^\alpha(f, t) \right). \quad (3.21)$$

Comme les sources s_j sont indépendantes, nous pouvons employer le théorème 3.1.1 afin de fournir une méthode de filtrage :

$$\forall (f, t), \mathbb{E}[s_j(f, t) | x(f, t)] = \frac{\sigma_j^\alpha(f, t)}{\sum_j \sigma_j^\alpha(f, t)} x(f, t). \quad (3.22)$$

Cette espérance *a posteriori*, proche de celle gaussienne en (3.14), est appelée filtrage α -Wiener [118]. Elle étend ainsi le filtrage de Wiener classique au cas de la théorie α -stable.

Interprétons l'apport du filtrage α -Wiener en (3.22) par rapport au filtrage 2-Wiener en (3.14). Rappelons que dans le cas gaussien, nous supposons que la DSP est approchée par le spectrogramme de puissance (voir la sous-section 2.2.2 page 18) pour tout (f, t) :

$$\forall (f, t), |x(f, t)|^2 \simeq \sigma^2(f, t) \quad (3.23)$$

avec

$$\sigma^2(f, t) = \sum_{j=1}^J \sigma_j^2(f, t) \quad (3.24)$$

L'approximation naturelle proposée est alors de supposer que le spectrogramme de puissance du mélange est la somme des spectrogrammes de puissance de chaque source [119, 52, 127] :

$$\forall(f, t), |x(f, t)|^2 \simeq \sum_{j=1}^J |s_j(f, t)|^2. \quad (3.25)$$

A contrario, d'autres recherches visant à l'amélioration de la qualité de séparation adoptent heuristiquement l'hypothèse suivante [163, 22]

$$\forall(f, t), |x(f, t)|^\alpha \simeq \sum_{j=1}^J |s_j(f, t)|^\alpha, \quad (3.26)$$

où $\{|x(f, t)|^\alpha\}_{f, t}$ a été dénommé plus tard α -spectrogramme du mélange dans [118]. Nous voyons que de manière similaire au modèle gaussien, l'égalité (3.24) se généralisant à

$$\sigma^\alpha(f, t) = \sum_{j=1}^J \sigma_j^\alpha(f, t) \quad (3.27)$$

sous-entend de considérer le modèle $S\alpha S_c$ en (3.20) et (3.21). Les paramètres d'échelles $\sigma^\alpha(f, t)$ sont alors appelés DSP *fractionnaire* [118]. Il devient légitime de se demander quelle approximation entre (3.25) et (3.26) est la plus justifiée ? Les expériences menées dans [118] ont alors considéré le critère

$$\forall(f, t), L_\alpha(f, t) = \left| |x(f, t)|^\alpha - \sum_{j=1}^J |s_j(f, t)|^\alpha \right|^{1/\alpha} \quad (3.28)$$

pour $\alpha \in]0, 2]$, appelé α -dispersion. Également, les divergences de Kullback-Leibler D_{KL} et d'Itakura-Saito D_{IS} ci-dessous

$$\forall u, v, D_{KL}(u|v) = u \ln\left(\frac{u}{v}\right) - u + v \quad (3.29)$$

$$\forall u, v, D_{IS}(u|v) = \frac{u}{v} - \ln\left(\frac{u}{v}\right) - 1 \quad (3.30)$$

ont été calculées. Souvent utilisées dans le cadre de la séparation audio [51, 54], ces divergences sont des outils pratiques afin d'évaluer l'erreur entre deux distributions. En particulier, il est connu que la divergence d'Itakura-Saito apparaît naturellement lors de la résolution du maximum de vraisemblance d'un modèle probabiliste gaussien. La partie sur l'estimation des paramètres sera davantage développée dans le chapitre suivant.

L'expérience faite a été alors de considérer huit œuvres musicales différentes⁴ tirées de phonogrammes aux styles musicaux variés. Chaque morceau est décomposé en plusieurs pistes (percussions, guitares, basse et chant) monophoniques. Les α -spectrogrammes sont alors calculés pour chaque piste, indiquant les valeurs $\{|x(f, t)|^\alpha\}_{f, t}$ et $\{|s_j(f, t)|^\alpha\}_{f, t}$. Les trois divergences des équations (3.28), (3.29) et (3.30) sont alors évaluées.

Il a été observé que les divergences sont minimales pour α proche de 1.2. En conséquence, cela induit qu'en pratique l'approximation (3.26) semble plutôt justifiable dans un cadre non-gaussien et donc a fortiori souligne l'importance de considérer les processus α -harmonisables pour différents problèmes de séparation.

4. Plus précisément, dans la base QUASI <http://www.tsi.telecom-paristech.fr/aao/en/2012/03/12/quasi/>

Chapitre 4

Séparation de sources sonores : estimation des paramètres

Résumé

Les chapitres précédents visaient à présenter, dans le cas des modèles α -stables, différentes méthodes de séparation. En général, les paramètres du modèle sont inconnus. Il est alors naturel que l'étape suivante soit celle de l'estimation des paramètres. Autrement dit, connaissant les observations, comment estimer les paramètres du modèle probabiliste considéré ? Pour ce faire, nous présentons dans un premier temps le contexte probabiliste dans lequel nous nous plaçons pour la séparation des sources sonores. Ensuite, nous exposerons dans le cas de modélisation α -stables la façon dont les paramètres de ces derniers sont estimés. Notamment, nous introduirons un modèle paramétrique très courant en apprentissage automatique appelée *nonnegative matrix factorization* (NMF). Dans le cas de l'audio, cet algorithme permet de décomposer en rang faible une matrice renseignant l'activation au cours du temps de « motifs spectraux » comme suit : l'une dépendant de la fréquence et l'autre dépendant du temps.

4.1 Paradigme pour la séparation de sources sonores

Considérons en premier lieu un signal observé. Couramment dans le cadre du traitement du signal audio, l'acquisition se fait à l'aide d'un ou plusieurs microphones. Par exemple, plusieurs personnes pratiquant simultanément un instrument de musique sont enregistrées et les données numérisées sont sauvegardées. Les acquisitions numériques fournissent alors un nombre fini de données. Contrairement à des acquisitions analogiques, les données numériques ne sont pas « continues ». Dans les modèles que nous considérons dans ce manuscrit, nous négligeons cet aspect discret.

De cette supposition, on en déduit que les signaux observés sont des échantillons d'une réalisation particulière d'un processus stochastique. Comme une variable aléatoire, un vecteur aléatoire ou un processus stochastique admet dans un certain nombre de cas une loi caractérisée entièrement par un ou plusieurs paramètres (*e.g.* la variance et la moyenne pour une variable aléatoire gaussienne). Il suffit alors d'estimer ces paramètres pour en déduire la valeur d'un signal recherché. C'est également l'un des aspects particuliers des distributions α -stables : l'estimation des paramètres impliquerait une méthode de séparation. Le but de cette estimation est d'arriver à inférer les paramètres Θ relatifs aux sources individuelles, à partir de l'observation du mélange x seulement. On dit alors que le modèle probabiliste pris en compte est paramétrique. Le paradigme pour la séparation de sources sonores (SSS) est le suivant :

1. Choisir un modèle probabiliste sur les sources désirées, paramétré par Θ .

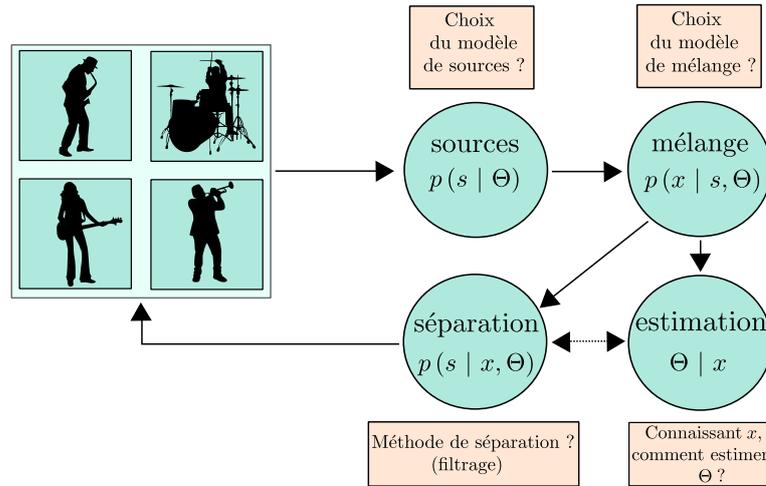


FIGURE 4.1 – Paradigme classique pour la séparation des sources sonores dans un cadre Bayésien.

2. En supposant les sources s et les paramètres Θ connus, en déduire un modèle applicable au mélange x .
3. En supposant les données observées x , estimer les paramètres Θ .
4. En supposant les données observées x et les paramètres Θ connus, extraire les sources s .

Ce schéma est résumé sur la Fig 4.1. Les procédés présentés dans [31, 175] sont l'essence même de tout algorithme utilisé en *séparation de sources sonores* (SSS) et faisant usage d'un modèle probabiliste. Dans les modèles vus au chapitre 3 page 23, nous considérons simplement que les observations x étaient la somme de chaque source souhaitée. Grâce à la stabilité par somme des modèles choisis, les lois probabilistes étaient simplement conservées et induisaient alors une méthode de séparation (ou filtrage) relativement simple. Les points 1, 2 et 4 pour un processus α -stable dans le cadre de la SSS ont donc été présentés dans le chapitre précédent. Le troisième point fera l'objet de ce chapitre, traitant de l'estimation des paramètres et des modèles pouvant être mis en place sur ces derniers.

4.2 Cas gaussien : divergence d'Itakura-Saito

Dans cette section, nous reprenons le modèle présenté dans la sous-section 3.1.2 page 24 du chapitre 3. Définissons à nouveau trois processus x , s et n respectivement l'observation, la source et le bruit. Nous supposons par ailleurs que l'observation est la somme de la source cible et du bruit :

$$x = s + n. \quad (4.1)$$

4.2.1 Maximum de vraisemblance et algorithme d'espérance-maximisation

Considérons le modèle proposé par les équations (3.9) et (3.10). Nous supposons que s et n sont des processus gaussiens localement SSL et indépendants. Nous nous restreignons alors à l'observation d'un nombre fini de points pour chaque composante régulièrement échantillonnée. Les signaux sont alors observés dans le domaine *temps-fréquence* (TF), donnant lieu à F bandes fréquentielles et T trames temporelles observables, dont chaque point TF est une gaussienne centrée de variance sa *densité spectrale*

de puissance (DSP). Il en résulte que les point TF sont indépendants. Pour résumer, on a :

$$\forall (f, t), s(f, t) \sim \mathcal{N}_c \left(\sigma_s^2(f, t) \right) \quad (4.2)$$

$$\forall (f, t), n(f, t) \sim \mathcal{N}_c \left(\sigma_n^2(f, t) \right) \quad (4.3)$$

impliquant par stabilité sommatoire que :

$$\forall f, t, x(f, t) \sim \mathcal{N}_c \left(\sigma_s^2(f, t) + \sigma_n^2(f, t) \right). \quad (4.4)$$

Notons $\Theta = \{\sigma_n^2(f, t), \sigma_s^2(f, t)\}_{f, t}$ l'ensemble des paramètres de ce modèle. On rappelle alors que le filtrage de Wiener nous donnant la source estimée aux moindres carrés et notée $\hat{s}(f, t) \triangleq \mathbb{E}[s(f, t) | x(f, t), \Theta]$ est :

$$\forall (f, t), \hat{s}(f, t) = \frac{\sigma_s^2(f, t)}{\sigma_s^2(f, t) + \sigma_n^2(f, t)} x(f, t). \quad (4.5)$$

Comme expliqué dans la section précédente, nous supposons connus pour tout (f, t) uniquement les observations $x(f, t)$ et le modèle sous-jacent appliqué aux sources $s(f, t)$ et au bruit $n(f, t)$. Afin d'estimer au mieux Θ , nous pouvons chercher à maximiser la densité de probabilité $g(x(f, t) | \Theta)$ de la variable aléatoire $x(f, t)$. Cette idée, appelée recherche du *maximum de vraisemblance* (MDV) [28, 29, 49, 81] revient alors à résoudre le problème ci-dessous :

$$\Theta^* \triangleq \operatorname{argmax}_{\Theta} \prod_{f, t} g(x(f, t) | \Theta). \quad (4.6)$$

En composant $g(x(f, t) | \Theta)$ avec l'opposé du logarithme népérien pour obtenir la log-vraisemblance, le problème dans le cas gaussien revient à minimiser la divergence d'Itakura-Saito comme définie à l'équation (3.30) :

$$\Theta^* = \operatorname{argmin}_{\Theta} \sum_{f, t} d_{IS} [|x(f, t)|^2 | \sigma_s^2(f, t) + \sigma_n^2(f, t)]. \quad (4.7)$$

En pratique, nous avons moins d'observations ($|x(f, t)|$) que d'inconnues ($\sigma_s^2(f, t)$ et $\sigma_n^2(f, t)$). Il est donc nécessaire de rajouter des hypothèses à notre modèle paramétrique. Dans le cadre du rehaussement de la parole, différentes initialisations peuvent être proposées :

- On peut supposer que le bruit \mathbf{n} est un bruit blanc gaussien [63, 113].
- À l'aide d'un *détecteur d'activité vocale* [14, 164, 62] (DAV), nous identifions les zones ne comportant que du bruit. On peut alors affirmer à ces positions que $\mathbf{x} = \mathbf{n}$ et a fortiori estimer les $\sigma_n^2(f, t)$ en supposant une évolution faible dans le temps. [22, 126, 90] (cette méthode sera d'avantage détaillée dans la partie IV).

Connaissant une estimation $\sigma_n^2(f, t)$, le problème devient alors convexe. Nous arrivons alors à la technique dite de *soustraction spectrale* :

$$\sigma_s^2(f, t) = \max \left(0, |x(f, t)|^2 - \sigma_n^2(f, t) \right). \quad (4.8)$$

Il suffit d'appliquer le filtre de Wiener (4.5) pour obtenir un signal débruité. Dans le cas où l'on souhaite estimer les paramètres Θ , on peut employer par exemple un algorithme itératif présenté dans [35] et appelé *espérance-maximisation* (EM). Il intègre plusieurs étapes intermédiaires pour estimer Θ . Dans notre exemple, on rajoute aux données observées $\{x(f, t)\}_{f, t}$ les sources non observées $\{s(f, t), n(f, t)\}_{f, t}$. Il peut alors être démontré que l'algorithme s'écrivant à l'itération i (avec $\Theta_{(0)}$ initialisé au préalable)

$$\text{Etape-E : } \mathcal{Q}_i(\Theta, \Theta_{(i)}) \triangleq \mathbb{E} \left[\sum_{f,t} \ln [g(x(f,t) | \Theta)] + \ln [g(s(f,t) | x(f,t), \Theta)] \right. \\ \left. + \ln [g(n(f,t) | x(f,t), \Theta)] | \Theta_{(i)} \right] \quad (4.9)$$

$$\text{Etape-M : } \Theta_{(i+1)} \leftarrow \underset{\Theta}{\operatorname{argmax}} (\mathcal{Q}_i(\Theta, \Theta_{(i)})) \quad (4.10)$$

permet de construire une suite $(\Theta_{(i)})_i$ faisant tendre $\sum_{f,t} \ln [g(x(f,t) | \Theta_{(i)})]$ vers un maximum local.

Souvent, l'initialisation des paramètres $\Theta^{(0)}$ est déterminé de manière empirique.

Il existe en outre plusieurs variantes de l'algorithme EM [30, 135], visant à améliorer la probabilité d'atteindre un maximum global ou encore d'accélérer la convergence du dit algorithme. Posons $\hat{\sigma}_s(f,t)$, $\hat{\sigma}_n(f,t)$ et $\hat{s}(f,t)$ respectivement les estimées de $\sigma_s(f,t)$, $\sigma_n(f,t)$ et $s(f,t)$. Grâce à la stabilité par sommation du modèle Gaussien, l'étape E peut se résumer par les mises à jour successives suivantes [39], impliquant entre autre le filtrage de Wiener classique

$$\text{Etape-E : } \begin{cases} \hat{\mu}_{s|x}(f,t) \triangleq \frac{\hat{\sigma}_s^2(f,t)}{\hat{\sigma}_s^2(f,t) + \hat{\sigma}_n^2(f,t)}, & \text{(masque de Wiener)} \\ \hat{s}(f,t) = \hat{\mu}_{s|x}(f,t) x(f,t), & \text{(espérance a posteriori)} \\ \hat{\sigma}_{s|x}^2(f,t) = \hat{s}(f,t) \hat{s}^*(f,t) + (1 - \hat{\mu}_{s|x}(f,t)) \hat{\sigma}_s^2(f,t) & \text{(variance a posteriori)} \end{cases} \quad (4.11)$$

et l'étape M résume la mise à jour des paramètres ci-dessous

$$\text{Etape-M : } \begin{cases} \hat{\sigma}_s^2(f,t) = \sigma_{s|x}^2(f,t), & \text{(variance de la source)} \\ \hat{\sigma}_n^2(f,t) = x(f,t)(x^*(f,t) - \hat{s}^*(f,t)) - \hat{s}(f,t)(\hat{s}^*(f,t) - x^*(f,t)). & \text{(variance du bruit)} \end{cases} \quad (4.12)$$

La stratégie de « filtrage itératif » basé sur l'algorithme EM peut alors se résumer comme suit, dans le cas d'un modèle gaussien :

Algorithme 1 Exemple d'algorithme de débruitage EM-gaussien

1. Entrées

- Données observées $\{x(f,t)\}_{f,t}$.
- Nombre d'itérations I de l'algorithme EM.

2. Initialisation

- Initialiser $\{\sigma_{n,(0)}^2(f,t)\}_{f,t}$ soit à l'aide d'un détecteur d'activité vocale (DAV) soit comme étant un bruit gaussien.
- Initialiser $\{\sigma_{s,(0)}^2(f,t)\}_{f,t}$ (dans [39], $\sigma_{s,(0)}^2(f,t) = x(f,t)x^*(f,t)$)

3. Algorithme EM, pour $i = 1, \dots, I$

- Etape-E en (4.11) afin d'obtenir $\hat{s}_{(i)}(f,t)$ et $\hat{n}_{(i)}(f,t) \triangleq x(f,t) - \hat{s}_{(i)}(f,t)$
 - Etape-M en (4.12) afin de calculer $\hat{\sigma}_{s,(i)}^2(f,t)$ et $\hat{\sigma}_{n,(i)}^2(f,t)$.
-

Cet algorithme itératif est également résumé dans la Fig. 4.2.

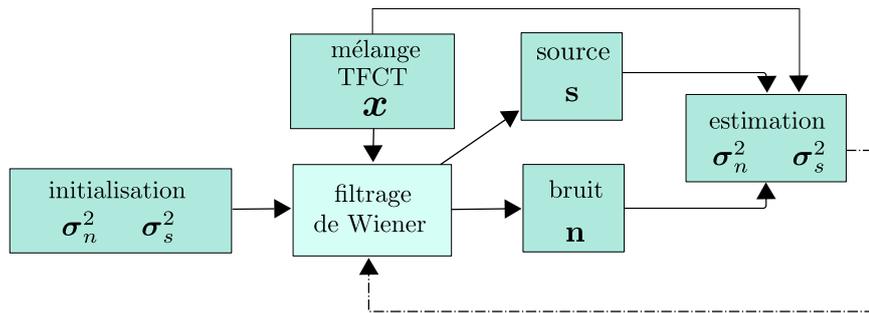
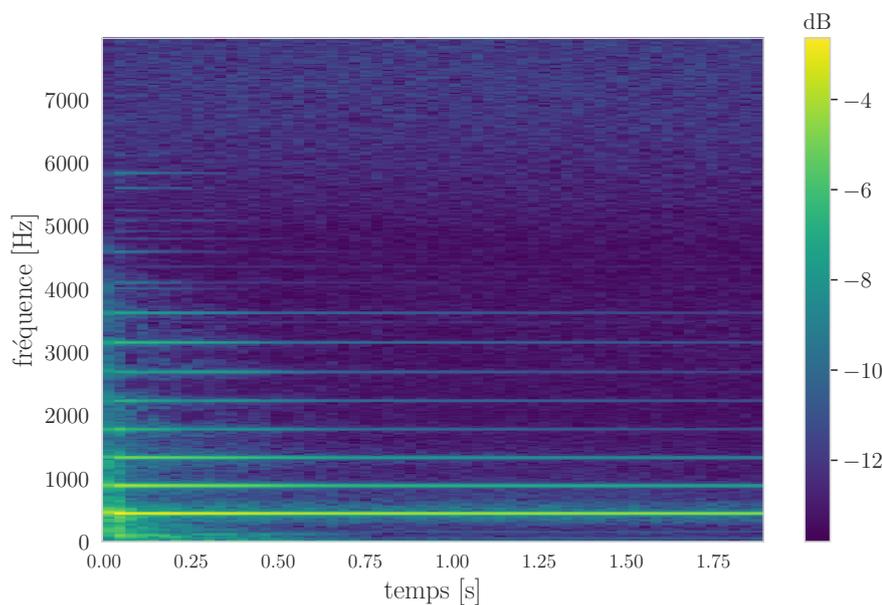


FIGURE 4.2 – Algorithme d’espérance-maximisation en SSS.

Malgré un modèle proposant une indépendance forte entre chaque point TF d’une source, il n’en demeure pas moins que les paramètres sont souvent en réalité dépendants entre eux. Par exemple, pour une note jouée comme représentée sur la Fig. 4.3, la fréquence dite fondamentale et ses harmoniques sont étroitement liées. En effet, si l’amplitude de la note fondamentale augmente, il en sera de même pour les harmoniques. Pour un signal plus complexe tel que la voix, un instrument à vent, un instrument à corde, d’autres types de caractérisations propres à chaque famille peuvent être considérées. On dit que les caractérisations forment un *dictionnaire*. Nous allons présenter un modèle particulier de DSP développé ces dix dernières années pour la séparation des sources sonores.

4.2.2 Modèle de factorisation matricielle non-négative pour la DSP

FIGURE 4.3 – Log-spectrogramme de la note de DO_3 d’un piano droit.

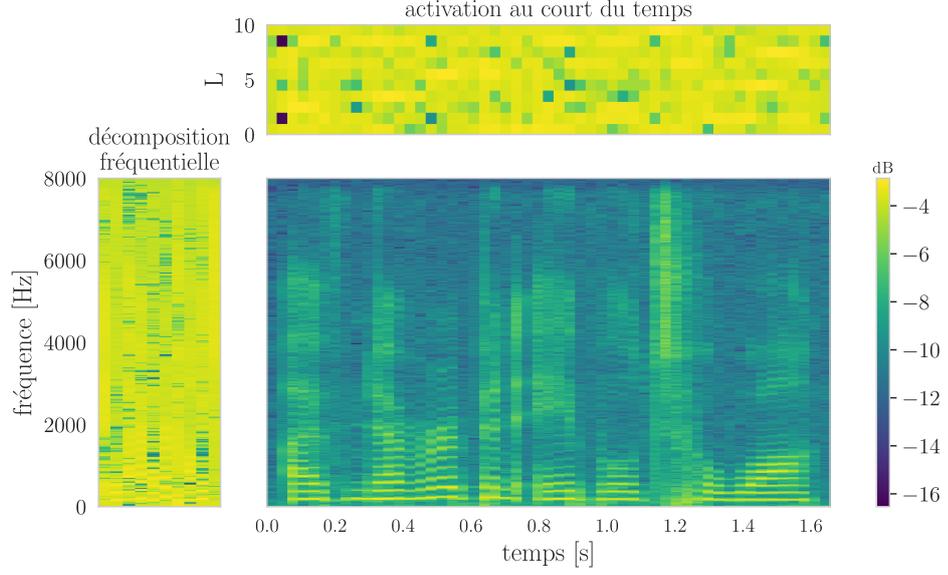


FIGURE 4.4 – Factorisation matricielle non-négative d'un signal de parole bruité.

Sur l'exemple de la Fig. 4.3, nous observons un spectre modulé dans le temps par une « activation ». Le but de la célèbre *nonnegative matrix factorization* (NMF) est de modéliser chaque source $s \triangleq \{s(f, t)\}_{f, t}$ comme la superposition de L spectres modulés dans le temps. Ces informations sont réunies dans une matrice non-négative \mathbf{W} de dimension $F \times L$ où F est le nombre de bandes fréquentielles contenues dans s . Les activations au cours du temps sont quant à elles résumées dans une matrice non-négative $\mathbf{H} \in \mathbb{R}_+^{T \times L}$ où T est le nombre de trames temporelles contenues dans s . Finalement, posons $\Sigma \triangleq \{\sigma_s^2(f, t)\}_{f, t}$ l'ensemble des *densités spectrales de puissances* (DSP) de la source s . Nous supposons alors le modèle de DSP suivant :

$$\Sigma \simeq \mathbf{W}\mathbf{H}. \quad (4.13)$$

Cette NMF illustrée dans Fig. 4.4 est devenue très populaire dans les années 2000 en SSS [52, 162, 11, 39, 105] et continue d'être développée aujourd'hui [80, 150].

Plaçons-nous maintenant dans un cadre plus général de SSS. Soit J sources monophoniques s_j , chacune étant issue de la trajectoire de J processus gaussiens indépendants. Chaque source est étudiée dans le domaine TF, où F représente le nombre de bande fréquentielle et T le nombre de trames temporelles. Nous appliquons de plus un modèle NMF, décomposé sur une base de taille L pour chaque DSP des sources. Pour résumer, nous avons pour tout $(j, f, t) \in \llbracket 1, J \rrbracket \times \llbracket 1, F \rrbracket \times \llbracket 1, T \rrbracket$:

$$s_j(f, t) \sim \mathcal{N}_c^1 \left(\sigma_j^2(f, t) \triangleq \sum_{l=1}^L w_j(f, l) h_j(l, t) \right). \quad (4.14)$$

Il en découle naturellement que l'observation $x(f, t) = \sum s_j(f, t)$ pour chaque point TF (f, t) est

$$\forall (f, t), x(f, t) \sim \mathcal{N}_c^1 \left(p(f, t) \triangleq \sum_{j=1}^J \sum_{l=1}^L w_j(f, l) h_j(l, t) \right). \quad (4.15)$$

Les paramètres $\Theta = \{\mathbf{W}_j, \mathbf{H}_j\}_j$ où $\mathbf{W}_j \in \mathbb{R}_+^{F \times L}$, $\mathbf{H}_j \in \mathbb{R}_+^{L \times T}$ sont respectivement la concaténation des coefficients $w_j(f, l) \geq 0$ et $h_j(l, t) \geq 0$. Nous pouvons rechercher le MDV comme en (4.6) et (4.7), et

retomber à nouveau sur une minimisation de la divergence d'Itakura-Saito en considérant la fonction de coût C suivante :

$$C(\Theta) = \sum_{f,t} d_{IS} \left[|x(f,t)|^2 \mid \sum_{j=1}^J \mathbf{w}_j^\top(f) \mathbf{h}_j(t) \triangleq [\mathbf{W}_j \mathbf{H}_j]_{ft} \right]. \quad (4.16)$$

Le calcul de $\frac{\partial C}{\partial \mathbf{W}_j}(\Theta)$ et de $\frac{\partial C}{\partial \mathbf{H}_j}(\Theta)$ nous donne respectivement pour tout j [51] :

$$\frac{\partial C}{\partial \mathbf{W}_j}(\mathbf{W}_j, \mathbf{H}_j) = [\mathbf{W}_j \mathbf{H}_j]^{\circ-1} \mathbf{H}_j^\top - \left([\mathbf{W}_j \mathbf{H}_j]^{\circ-2} \odot |\mathbf{X}|^{\odot 2} \right) \mathbf{H}_j^\top; \quad (4.17)$$

$$\frac{\partial C}{\partial \mathbf{H}_j}(\mathbf{W}_j, \mathbf{H}_j) = \mathbf{W}_j^\top [\mathbf{W}_j \mathbf{H}_j]^{\circ-1} - \mathbf{W}_j^\top \left([\mathbf{W}_j \mathbf{H}_j]^{\circ-2} |\mathbf{X}|^{\odot 2} \right) \quad (4.18)$$

où l'opérateur, " \odot " indique un produit terme à terme, " \circ^a " indique une puissance a appliquée terme à terme sur chaque coefficient d'une matrice, " $-$ " factorise l'expression terme à terme et $\forall f, t, [|\mathbf{X}|]_{f,t} \triangleq |x(f,t)|$. La solution des équations (4.17) et (4.18) peut être calculée en faisant une descente de gradient classique [9]. Malheureusement, cela n'assure pas la non-négativité des matrices \mathbf{W}_j et \mathbf{H}_j .

Afin de respecter cette contrainte, une astuce proposée dans [102] permet d'atteindre un minimum local de la fonction de coût C . En effet, il est possible d'assurer une descente de gradient garantissant la non-négativité des matrices en multipliant chaque matrice par le rapport de la partie négative sur la partie positive du gradient :

$$\mathbf{W}_j \leftarrow \mathbf{W}_j \odot \frac{\left([\mathbf{W}_j \mathbf{H}_j]^{\circ-2} \odot |\mathbf{X}|^{\odot 2} \right) \mathbf{H}_j^\top}{[\mathbf{W}_j \mathbf{H}_j]^{\circ-1} \mathbf{H}_j^\top}; \quad (4.19)$$

$$\mathbf{H}_j \leftarrow \mathbf{H}_j \odot \frac{\mathbf{W}_j^\top \left([\mathbf{W}_j \mathbf{H}_j]^{\circ-2} |\mathbf{X}|^{\odot 2} \right)}{[\mathbf{W}_j \mathbf{H}_j]^{\circ-1}}. \quad (4.20)$$

On s'assure ainsi d'une descente vers un minimum local, tout en préservant la non-négativité des matrices [102, 51]. Cette règle, dite règle de *mise à jour multiplicative* (MJM) peut être appliquée à d'autres divergences comme nous le verrons plus tard. L'un des problèmes majeurs de cette représentation réside dans l'initialisation des paramètres. En pratique, proposer une initialisation de matrice non-négative aléatoire avec un tel modèle gaussien n'assure pas forcément une convergence rapide et convenable vers un minimum global.

4.3 Le cas Cauchy

Intéressons-nous à l'estimation des paramètres dans le cas d'un modèle α -stable. En premier lieu, nous verrons comment exploiter la densité de probabilité d'un modèle de Cauchy ($\alpha = 1$) pour l'estimation des paramètres, au même titre que le cas gaussien. Nous en déduirons notamment qu'une factorisation de matrice non-négative fait également sens.

Nous considérons J sources comme étant des processus α -harmonisables (voir définition 2.3.3 page 21) avec $\alpha = 1$. D'après le théorème 2.3.3 page 21, nous pouvons considérer J signaux inconnus $s_1, \dots, s_j \in \mathbb{C}^{F \times T}$ dans le domaine TF comme suit pour tout $(j, f, t) \in \llbracket 1, J \rrbracket \times \llbracket 1, F \rrbracket \times \llbracket 1, T \rrbracket$:

$$s_j(f, t) \sim S1S_c(\sigma_j(f, t)), \quad (4.21)$$

où F, T représentent respectivement le nombre de bandes fréquentielles et le nombre de trames temporelles. Les $\{\sigma_j(f, t)\}_{f,t}$ sont les DSP fractionnaires de la $j^{\text{ème}}$ source. Par conséquent, le mélange $x(f, t) =$

$\sum_j s_j(f, t)$ suit une loi de Cauchy isotrope pour tout (f, t) :

$$x(f, t) \sim S1S_c \left(\sum_{j=1}^J \sigma_j(f, t) \right) \quad (4.22)$$

Rappelons que le théorème 3.1.1 page 26 nous donne le filtrage α -Wiener suivant :

$$\forall (f, t), \mathbb{E} \left[s_j(f, t) \mid x(f, t), \{\sigma_j(f, t)\}_{j,f,t} \right] = \frac{\sigma_j(f, t)}{\sum_{j'} \sigma_{j'}(f, t)} x(f, t) \quad (4.23)$$

Comme dans le cas Gaussien, nous pouvons imposer un modèle NMF au DSP des sources [120] :

$$\sigma_j(f, t) = \sum_{l=1}^L w_j(f, l) h_j(l, t); \quad (4.24)$$

$$\sigma_x(f, t) = \sum_{j=1}^J \sum_{l=1}^L w_j(f, l) h_j(l, t). \quad (4.25)$$

Notons $\Theta = \{\mathbf{W}_j, \mathbf{H}_j\}_{j,f,t}$ les matrices composées respectivement des coefficients $w_j(f, l) \geq 0$ et $h_j(l, t) \geq 0$. Pour l'estimation de Θ , la forme analytique connue de la densité de probabilité g_x d'une loi de Cauchy isotrope $x \sim S1S_c(\sigma_x)$ [153, p 81]

$$\forall u, g_x(u \mid \sigma_x) = \frac{\sigma_x}{2\pi (u^2 + \sigma_x^2)^{3/2}} \quad (4.26)$$

est associée à la recherche d'un MDV, afin d'estimer les paramètres Θ :

$$\Theta^* \leftarrow \underset{\Theta}{\operatorname{argmin}}(C(\Theta)) \quad (4.27)$$

où C est la fonction de coût global définie comme

$$C(\Theta) = \sum_{f,t} -\ln [g_{f,t}(p(f, t) \mid \Theta)] \quad (4.28)$$

avec $g_{f,t}$ la densité de probabilité de $x(f, t)$ et $p(f, t) = |x(f, t)|$ l'amplitude du spectrogramme au point TF (f, t) . L'équation (4.28) se réécrit alors à une constante additive près

$$C(\Theta) \stackrel{\text{c}}{=} \sum_{f,t} \left(\frac{3}{2} \ln [p(f, t)^2 + \sigma(f, t)^2] - \ln [\sigma(f, t)] \right). \quad (4.29)$$

4.3.1 Estimation des paramètres par les règles de mises à jour multiplicatives

Afin de résoudre ce problème de minimisation, une première « approche naïve » de mise à jour des paramètres, énoncée dans la sous-section 4.2.2, peut être proposée. En calculant $\frac{\partial C}{\partial \mathbf{W}}(\Theta)$ et $\frac{\partial C}{\partial \mathbf{H}}(\Theta)$ et en adoptant la technique de [102], nous obtenons les mises à jour suivantes :

$$\mathbf{W}_j \leftarrow \mathbf{W}_j \odot \frac{[\mathbf{W}_j \mathbf{H}_j]^{\ominus 1} \mathbf{H}_j^\top}{\mathbf{Z}_j \mathbf{H}_j^\top}, \quad (4.30)$$

$$\mathbf{H}_j \leftarrow \mathbf{H}_j \odot \frac{\mathbf{W}_j^\top [\mathbf{W}_j \mathbf{H}_j]^{\ominus 1}}{\mathbf{W}_j^\top \mathbf{Z}_j}, \quad (4.31)$$

où

$$\mathbf{Z}_j = \frac{3\mathbf{W}_j\mathbf{H}_j}{|\mathbf{X}|^{\odot 2} + (\mathbf{W}_j\mathbf{H}_j)^{\odot 2}}. \quad (4.32)$$

Malgré un algorithme relativement simple à appliquer et régi par un modèle robuste aux valeurs aberrantes (bruit impulsif, variation anormale de l'amplitude du signal observé *etc.*), il n'en demeure pas moins une justification théorique maigre de la convergence de l'algorithme vers un minimum local. Nous décrivons de manière plus théorique une autre exploitation de $C(\theta)$ dans la sous-section suivante.

4.3.2 Technique de majoration-égalisation

Cette technique proposée également dans [51] repose sur la majoration de la fonction de coût par une fonction f tel que :

$$\forall (\Theta, \Theta'), C(\Theta) \leq f(\Theta', \Theta); \quad (4.33)$$

$$\forall \Theta, C(\Theta) = f(\Theta, \Theta). \quad (4.34)$$

En explorant d'autres paramètres Θ' sur une telle fonction f , la non-croissance de C est assurée. Elle est en effet continuellement majorée par f et vérifie à terme $f(\Theta', \Theta) = C(\Theta)$. De plus, il a été démontré dans [51] que si la fonction majorée est une β -divergence [15] (*e.g.* divergence d'Itakura-Saito, divergence de Kullback-Leiber), alors une descente de gradient de f possible garantissant la non-négativité des paramètres Θ peut se résumer à nouveau par une mise à jour des paramètres multiplicatifs comme en [51]. La fonction f esquissée dans [120] est :

$$f(\Theta', \Theta) = C(\Theta) + \sum_{f,t} \left[\frac{3}{2} \frac{\sigma'(f,t)^2 - \sigma(f,t)^2}{\sigma(f,t)^2 + p(f,t)^2} + \frac{\sigma(f,t)}{\sigma'(f,t)} - 1 \right]. \quad (4.35)$$

Nous en déduisons les mises à jour suivantes :

$$\mathbf{W}_j \leftarrow \mathbf{W}_j \odot \frac{[\mathbf{W}_j\mathbf{H}_j]^{\odot -1} \mathbf{H}_j^\top}{\check{\mathbf{Z}}_j \mathbf{H}_j^\top + \mathbf{F}_j}, \quad (4.36)$$

$$\mathbf{H}_j \leftarrow \mathbf{H}_j \odot \frac{\mathbf{W}_j^\top [\mathbf{W}_j\mathbf{H}_j]^{\odot -1}}{\mathbf{W}_j^\top \check{\mathbf{Z}}_j + \mathbf{G}_j}. \quad (4.37)$$

où les matrices, \mathbf{F}_j et \mathbf{G}_j dépendant de \mathbf{Z}_j (4.32) sont de la forme suivante

$$\mathbf{F}_j = \sqrt{\left[\frac{\mathbf{Z}_j\mathbf{H}_j}{4} \right]^{\odot 2} + 2 \left([\mathbf{W}_j\mathbf{H}_j]^{\odot -1} \mathbf{H}_j^\top \right) \odot \left[\frac{\mathbf{Z}_j\mathbf{H}_j}{4} \right]^{\odot 2}}$$

$$\mathbf{G}_j = \sqrt{\left[\frac{\mathbf{W}_j^\top \mathbf{Z}_j}{4} \right]^{\odot 2} + 2 \left(\mathbf{W}_j^\top [\mathbf{W}_j\mathbf{H}_j]^{\odot -1} \right) \odot \left[\frac{\mathbf{W}_j^\top \mathbf{Z}_j}{4} \right]^{\odot 2}}$$

avec " $\sqrt{\cdot}$ " étant l'opérateur racine carrée terme à terme.

Les deux méthodes particulières présentées dans cette section sont difficilement applicables pour $\alpha \neq 1$ et $\alpha \neq 2$. En effet, il n'existe pas de forme analytique d'une densité de probabilité de variables α -stables. Néanmoins, nous pouvons nous affranchir d'estimer les paramètres d'échelle σ en utilisant d'autres techniques alternatives : les moments d'ordre faibles fractionnaires ou encore la propriété 2.1.4 page 15 permettant d'affirmer qu'une variable α -stable est conditionnellement gaussienne.

4.4 Cas Général

On se replace à nouveau dans le cas d'un mélange de J réalisations de processus α -harmonisables indépendants. Les sources dans le domaine TF vérifient donc :

$$\forall (j, f, t) \in \llbracket 1, J \rrbracket \times \llbracket 1, F \rrbracket \times \llbracket 1, T \rrbracket, s_j(f, t) \sim S\alpha S_c \left(\sigma_j^\alpha(f, t) \right) \quad (4.38)$$

De nouveau, $x(f, t) = \sum_j s_j(f, t)$ suit une loi α -stable isotrope pour tout (f, t) qui s'écrit

$$x(f, t) \sim S\alpha S_c \left(\sum_{j=1}^J \sigma_j^\alpha(f, t) \right) \quad (4.39)$$

et le théorème 3.1.1 nous donne le filtrage α -Wiener suivant :

$$\forall (f, t), \mathbb{E} \left[s_j(f, t) \mid x(f, t), \left\{ \sigma_{j'}^\alpha(f, t) \right\}_{j', f, t} \right] = \frac{\sigma_j^\alpha(f, t)}{\sum_{j'} \sigma_{j'}^\alpha(f, t)} x(f, t). \quad (4.40)$$

Nous allons voir comment estimer $\sigma_j^\alpha(f, t)$ pour tout α .

4.4.1 Méthode des Moments d'ordres faibles fractionnaires (MOFF)

Une première méthode permettant une estimation des paramètres $\sigma_j^\alpha(f, t)$ consiste à exploiter les moments existants d'une variable α -stable. En effet, rappelons que si $x \sim S\alpha S_c(\sigma_x^\alpha)$, alors [153]

$$\begin{aligned} \forall p < \alpha, \mathbb{E}[|x|^p] < \infty; \\ \forall p \geq \alpha, \mathbb{E}[|x|^p] = \infty. \end{aligned}$$

Par exemple, une loi de Cauchy n'admet pas de moyenne.

Il existe de plus une relation entre le moment d'ordre $p < \alpha$ et σ :

Proposition 4.4.1. Soit $x \sim S\alpha S_c(\sigma_x^\alpha)$. Alors, pour $0 < p < \alpha$, on a :

$$\mathbb{E}[|x|^p]^{\alpha/p} = c_{\alpha, p} \sigma_x^\alpha, \quad (4.41)$$

avec

$$c_{\alpha, p} = \left(\frac{p2^p}{\alpha} \Gamma\left(\frac{p}{2}\right) \frac{\Gamma(-p/\alpha)}{\Gamma(-p/2)} \right)^{1/p}. \quad (4.42)$$

On peut alors estimer σ_x en calculant empiriquement plusieurs moments d'ordre p . Cette méthode des *moments d'ordres faibles fractionnaires* (MOFF) [136, 170, 116] consiste à choisir R moments $p_1, \dots, p_R \in]0, \alpha[$ et utiliser N échantillons x_1, \dots, x_N de façon à obtenir une estimation de σ_x^α :

$$\hat{\sigma}_{MOFF, x}^\alpha = \frac{1}{R} \sum_{r=1}^R \left(\frac{1}{c_{\alpha, p_r} N} \sum_{n=1}^N |x_n|^{p_r} \right)^{\alpha/p_r}. \quad (4.43)$$

En pratique, cette méthode n'est utilisable que dans le cas de séparation de deux sources très stationnaires [122, 43].

Remarque Il existe par ailleurs une méthode de séparation multicanal où, après projection dans différentes directions, les variables obtenues sont α -stables isotropes. Ces dernières admettent dans ce cas un paramètre d'échelle estimable grâce aux MOFF [55, 56].

Algorithme 2 Algorithme de Métropolis-Hastings

1. Générer $\phi'_{n+1} \sim q(\cdot, \phi_n)$.
2. Calculer la probabilité d'acceptation

$$a(\phi_n \rightarrow \phi'_{n+1}) = \min \left(\frac{\pi_\phi(\phi'_{n+1}) q(\phi_n, \phi'_{n+1})}{\pi_\phi(\phi_n) q(\phi'_{n+1}, \phi_n)}, 1 \right) \quad (4.44)$$

3. Tirer un nombre u uniformément entre 0 et 1, puis
 - (a) Si $u > a(\phi_n \rightarrow \phi'_{n+1})$, alors $\phi_{n+1} = \phi'_{n+1}$; (acceptation)
 - (b) Sinon, $\phi_{n+1} = \phi_n$. (rejet)
4. Recommencer l'algorithme, en fonction du nombre de ϕ_n voulu.

4.4.2 Gaussianité conditionnelle et algorithme de Métropolis-Hastings

La méthode ci-dessous utilise la propriété 2.1.4 page 15 représentant une variable α -stable isotrope comme conditionnellement gaussienne. Considérons à nouveau les modèles décrits en (4.38), (4.39) et (4.40). L'utilisation de la propriété 2.1.4 nous permet de transformer les modèles α -stables en modèles conditionnellement gaussiens :

$$\forall (f, t), s_j(f, t) \mid \phi_j(f, t) \sim \mathcal{N}_c(\phi_j(f, t) \sigma_j^2(f, t)), \quad (4.45)$$

$$\forall (f, t), x(f, t) \mid \{\phi_j(f, t)\}_j \sim \mathcal{N}_c \left(\sum_{j=1}^J \phi_j(f, t) \sigma_j^2(f, t) \right), \quad (4.46)$$

où $\phi_j(f, t) \sim \mathcal{P}\alpha S \left(2 \left[\cos \left(\frac{\pi\alpha}{4} \right) \right]^{2/\alpha} \right)$ (cf. définition 2.1.6). Étant données les variables impulsionnelles $\phi_j(f, t)$, nous pouvons reconstruire chaque source en utilisant un filtrage de Wiener comme suit :

$$\begin{aligned} \forall (f, t), \hat{s}_j(f, t) &\triangleq \mathbb{E} [s_j(f, t) \mid x(f, t), \Phi, \Theta], \\ &= \frac{\sigma_j^2(f, t) \phi_j(f, t)}{\sum_{j'} \sigma_{j'}^2(f, t) \phi_{j'}(f, t)} x(f, t), \end{aligned} \quad (4.47)$$

où $\Theta = \left\{ \sigma_j^2(f, t) \right\}_{j, f, t}$ et $\Phi = \left\{ \phi_j(f, t) \right\}_{j, f, t}$ sont les paramètres du modèle. Concernant l'estimation des paramètres Θ , une méthode utilise la supposition que chaque $\sigma_j^2(f, t)$ suit une loi gamma généralisée [161, 165]. Si de plus nous supposons une décomposition NMF des paramètres Θ , la loi inverse-gamma [182] peut être employée afin de modéliser la distribution a priori des coefficients non-négatifs [47]. Nous invitons le lecteur à se référer aux articles [47, 48, 161] pour la méthode d'estimation des paramètres Θ .

La partie qui nous intéresse est celle de l'estimation des variables impulsionnelles Φ . Afin de simplifier les notations, nous fixons j, f et t . Comme mentionné précédemment, il n'existe pas de forme analytique de la densité de probabilité d'une variable α -stable. Néanmoins, des méthodes comme celles dites de *Markov chain Monte-Carlo* (MCMC) ont été proposées dans [162, 161]. Les algorithmes MCMC impliquent d'approcher une loi de probabilité π_ϕ , en générant ϕ_{n+1} uniquement par rapport à ϕ_n avec le souhait d'obtenir au bout d'un grand nombre d'itérations un échantillonnage suivant la loi π cible.

Un algorithme MCMC très employé est celui de *Metropolis-Hastings* [132](MH) où la règle dite de *rejet* est considérée pour accepter ou non un nouvel échantillon tiré en fonction du précédent. Soit

$q(\phi_n, \phi_{n+1}) \triangleq q(\phi_n | \phi_{n+1})$ la densité de probabilité de la variable $\phi_n | \phi_{n+1}$. Grâce au modèle conditionnellement gaussien présenté dans [162], l'équation (4.44) peut simplement s'écrire comme le rapport de deux densités de probabilités gaussiennes. L'algorithme 2 décrit l'approche MH employée.

Chapitre 5

Méthodes et modèles probabilistes pour la localisation de sources acoustiques

Résumé

Potentiellement aidée de certaines informations de contexte comme la taille de la pièce, le temps de réverbération, ou la position des microphones, l'objectif de la *localisation de sources sonores* est d'estimer la position dans l'espace des sources acoustiques d'une source potentielle dans un milieu physique. Une méthode classique vise à exploiter les *time difference of arrival* (TDOA) entre plusieurs microphones d'une antenne. Par exemple, dans le cas d'un modèle de source gaussien, l'utilisation de la covariance entre les microphones s'avère naturelle pour estimer les TDOA. On montre alors que la corrélation entre les signaux captés par les différents microphones fournit l'information de localisation recherchée. Cependant, des études dans [65] ont souligné qu'en présence d'un bruit relativement important ou d'un temps de réverbération non négligeable, la corrélation ne suffit pas à fournir une localisation satisfaisante. Ceci a motivé l'introduction de méthodes de localisation basée également sur les TDOA mais dans le cas d'une modélisation α -stable.

5.1 TDOAs exploitant les moments d'ordre 2

Les travaux présentés ici sont inspirés entre autres de [185, 189]. Soit \tilde{s}_l une source représentée dans le domaine temporel par sa forme d'onde et localisée en un point l dans l'espace \mathbb{R}^3 . Par ailleurs, nous supposons que l'acquisition du signal acoustique émis par cette source l se fait à l'aide de K microphones enregistrant pour tout temps t un signal noté $\tilde{x}_k(t)$ supposé indépendant et modélisé comme suit :

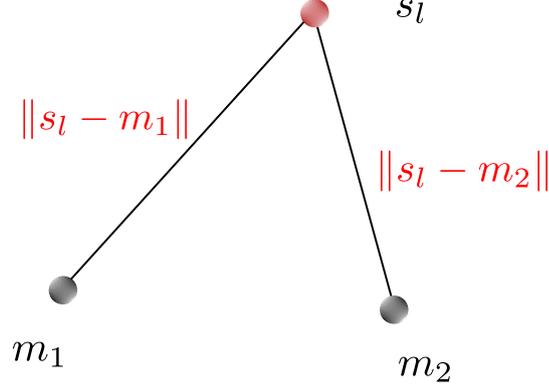
$$\forall k > 0, \tilde{x}_k(t) = (a_{k,l} * \tilde{s}_l)(t) \quad (5.1)$$

où $a_{k,l}$ désigne le filtre modélisant la propagation acoustique de la source \tilde{s}_l vers le micro k et $*$ l'opérateur de convolution. L'approche classique est alors de décomposer le filtre a_k comme la somme de son premier pic (associée à une amplitude $c_{k,l}$ au temps $\tau_{k,l} = c_{k,l}/v$ où v est la vitesse du son dans l'air) et du reste du filtre $a_{k,l}$ (noté $\hat{a}_{k,l}$) :

$$\forall k > 0, \tilde{x}_{k,l}(t) = c_{k,l}\tilde{s}_l(t - \tau_{k,l}) + \tilde{n}_{k,l}(t)$$

où $\tilde{n}_{l,k}(t) = (\hat{a}_{k,l} * \tilde{s}_l)(t)$ est le reste du filtre a_k de la source l . Une transformée de Fourier de (5.1) nous donne

$$\forall k > 0, \forall f, x_k(f) = c_{k,l}s_l(f) \exp(-i2\pi\tau_{k,l}) + n_{k,l}(f) = a_{k,l}(f)s_l(f) + n_{k,l}(f). \quad (5.2)$$

FIGURE 5.1 – Configuration de deux microphones m_1, m_2 et d'une source s_l localisée en l .

On note $\tau_{kk',l} = \tau_{k,l} - \tau_{k',l}$ le *time differential of arrival* (TDOA) entre les microphones k et k' via l'émission de la source l . Si on repère par exemple deux microphones spatialement par les points m_k et $m_{k'}$ et que l'on considère uniquement le chemin direct de la source l , on a la relation [143] (voir Fig. 5.1 pour l'illustration) :

$$\tau_{kk',l} = \frac{\|s_l - m_k\|}{v} - \frac{\|s_l - m_{k'}\|}{v}$$

où l'ensemble des solutions correspond à deux hyperboles, communément appelées *direction of arrival* (DOA) dans la littérature [73]. En utilisant plus de deux microphones, nous contraignons davantage le problème et localisons avec plus de précision la source l . Pour déterminer les TDOA, il est expliqué dans [189] qu'ils maximisent les fonctions de corrélations entre deux microphones :

$$\tau_{kk',l} = \operatorname{argmax}_{\tau} \left(\int x_k(f) x_{k'}^*(f) \exp(2\pi i f \tau) df \right).$$

Il est donc important que les moments d'ordre 2 existent pour que cette méthode soit valable. Pour prendre en compte l'information globale entre chaque microphone, nous pouvons sommer les corrélations suivantes et déterminer le maximum de la fonction suivante :

$$\begin{aligned} P(l) &= \sum_{k=1}^K \sum_{k'=1}^K \int x_k(f) x_{k'}^*(f) \exp(2\pi i f \tau_{kk',l}) df \\ &= \int \left| \sum_{k=1}^K x_k(f) \exp(2\pi i f \tau_{k,l}) \right|^2 df. \end{aligned}$$

Si nous ajoutons l'hypothèse forte que le mélange est anéchoïque ($\hat{a}_{k,l}(t) = 0$) et sans bruit additif ($n_k(t) = 0$), on obtient :

$$P(l) = \int \left| \sum_{k=1}^K \frac{x_k(f) a_{k,l}^*(f)}{c_{k,l}} \right|^2 df.$$

En général, $c_{k,l}$ est connu suivant le modèle acoustique sous-jacent (nous reviendrons sur ces détails dans la partie II). Il en découle alors

$$P(l) = \int |\langle \mathbf{x}(f), \mathbf{a}_l(f) \rangle|^2 df, \quad (5.3)$$

où $\mathbf{x}(f) = [x_1(f), \dots, x_k(f)]^\top$ et $\mathbf{a}_l(f) = \left[\frac{a_{1,l}(f)}{c_{1,l}(f)}, \dots, \frac{a_{K,l}(f)}{c_{K,l}(f)} \right]^\top$ pour tout f . L'équation (5.3) est appelé *puissance de réponse directionnelle* (PRD). En calculant cette valeur pour plusieurs positions l et de ce fait pour différentes valeurs de $\mathbf{a}(f)$ et $\mathbf{x}(f)$, nous pouvons alors choisir la ou les valeurs maximales du PRD en différents points l comme étant une source potentiellement active.

Remarque Il existe également une méthode améliorée de (5.3) appelée *puissance de réponse directionnelle avec phase transformée* [24] (PRD-PHAT) où l'équation (5.3) se réécrit :

$$P_{\text{HAT}}(l) = \int \left| \frac{\langle \mathbf{x}(f), \mathbf{a}_l(f) \rangle}{\|\mathbf{x}(f)\|_2} \right|^2 df.$$

Cette approche donne en pratique de meilleurs résultats que l'estimation de la PRD.

Il est également possible d'utiliser d'autres opérateurs que la corrélation pour calculer la PRD. Nous allons voir notamment dans la section suivante que la covariation est un outil alternatif pour la localisation de sources acoustiques.

5.2 TDOAs dans le cas α -stable stéréophonique

Les travaux présentés dans cette section sont inspirés de [65]. Dans la section 5.1, la localisation de sources acoustiques repose essentiellement sur l'exploitation des moments statistiques d'ordre 2. Étant donné que dans le cas non-gaussien α -stable seul les moments d'ordre $p < \alpha$ existent, cette stratégie tombe en échec. Il existe cependant une alternative faisant appel à la *covariation* [65, 136, 153]. Nous informons le lecteur qu'une discussion plus approfondie sur le concept de covariation ainsi que différents exemples sont traités au Chapitre 11 de ce manuscrit. Pour simplifier les explications, nous supposons simplement observer le chemin direct d'une source s . De nouveau, nous nous intéressons aux TDOAs entre deux microphones en considérant le modèle suivant pour une source s_l et pour $\alpha \in]1, 2[$:

$$\forall f, \begin{cases} x_1(f) = s_l(f) \\ x_2(f) = \exp(-i2\pi f \tau_l) s_l(f) \end{cases} \quad (5.4)$$

où chaque $s_l(f) \sim S\alpha S(\sigma_l^\alpha(f))$ est une variable α -stable symétrique complexe (pas forcément isotrope).

À la place de la corrélation entre les deux microphones, nous utilisons alors l'estimateur de *covariation* définie comme suit pour deux variables jointement α -stables $x_1 \sim S\alpha S(\sigma_{x_1}^\alpha)$ et $x_2 \sim S\alpha S(\sigma_{x_2}^\alpha)$ [136] :

$$\forall 0 < p < \alpha, [x_1, x_2]_\alpha = \frac{\mathbb{E}(x_1 x_2^{(p-1)})}{\mathbb{E}(|x_2|^p)} \sigma_{x_2}^\alpha \quad (5.5)$$

qui admet des propriétés similaires à la covariance (linéarité à gauche, égale à zéro si x_1 et x_2 sont indépendantes etc.) avec $z^{(\alpha)} \triangleq z|z|^{\alpha-1}$ la *fonction puissance signée* [153]. En s'inspirant alors de l'exploitation des TDOAs via la PRD, on peut calculer la covariation suivante :

$$\begin{aligned} \forall f, \gamma_{x_1, x_2}(l, f) &= [x_1(f), x_2(f)]_\alpha \\ &= [s_l(f), s_l(f)]_\alpha |\exp(-if\tau_l)|^{\alpha-2} \exp(if\tau_l) \\ &= \sigma_l^\alpha(f) \exp(if\tau_l). \end{aligned} \quad (5.6)$$

À partir du calcul précédent, nous pouvons alors proposer une mesure de covariation contenant l'information du TDOA en normalisant (5.6) :

$$\forall f, d_{x_1, x_2}(l, f) \triangleq \frac{\gamma_{x_1, x_2}(l, f)}{|\gamma_{x_1, x_2}(l, f)|} = \exp(if\tau_l) \quad (5.7)$$

Comme dans le cas du PRD-PHAT, le délai τ correspond à la transformée de Fourier inverse de (5.7). La méthode revient alors à calculer les covariations $d_{x_1, x_2}(l, f)$, de les normaliser et d'effectuer une transformée de Fourier inverse de (5.7) pour obtenir τ_l . Il a été démontré dans [65] que les TDOAs étaient mieux estimés avec cette méthode qu'avec la PRD, surtout en présence d'une forte réverbération. Malgré des hypothèses fortes de ne considérer comme observation que le chemin direct de chaque source, la localisation est assurée via cette nouvelle méthode.

Deuxième partie

Mélanges discrets et application à la localisation de sources audio

Chapitre 6

Mélanges multivariés linéaires de sources α -stables scalaires

Résumé

Dans la première partie, l'état de l'art que nous avons développé se concentrait sur les distributions α -stables scalaires. Dans ce chapitre, nous commencerons par introduire plusieurs résultats et propriétés sur les mélanges multivariés linéaires de variables α -stables scalaires. Toujours dans un contexte de modélisation de signaux, ces variables auront le sens de source et les vecteurs déterministes par lesquels ces variables seront multipliées contiendront l'information spatiale. Nous verrons qu'un tel mélange vectoriel peut être résumé, dans le cas α -stable non-gaussien, par un vecteur non-négatif appelé *mesure spatiale discrète* (MSD) dont l'estimation sera présentée au chapitre 7. Nous présentons comment cette théorie a été instanciée par Kidmose dans [94] à un cas particulier de l'*analyse en composantes indépendantes* (ACI), dont le but est la séparation de mélanges linéaires instantanés de sources indépendantes et identiquement distribuées (i.i.d.) α -stables.

6.1 Modèle et notations

Introduisons le modèle de mélange considéré ainsi que les notations utilisées tout au long de cette seconde partie. Soit s_1, \dots, s_L des variables aléatoires complexes indépendantes et $0 < \alpha < 2$ tels que :

$$\forall l \in \llbracket 1, L \rrbracket, s_l \sim S\alpha S_c(\Upsilon_l), \quad (6.1)$$

où $\Upsilon_l \geq 0$. Considérons en outre L vecteurs complexes déterministes $\mathbf{a}_1, \dots, \mathbf{a}_L \in \mathbb{C}^K$ de dimension K . Le mélange linéaire \mathbf{x} considéré est alors le suivant :

$$\mathbf{x} = \sum_{l=1}^L \mathbf{a}_l s_l. \quad (6.2)$$

Nous allons montrer que le vecteur \mathbf{x} de l'équation (6.2) est en général anisotrope mais symétrique α -stable ($S\alpha S$). Si on revient à la définition 2.1.5 page 14, on devrait montrer que $\forall \phi \in [0, 2\pi[$, $e^{i\phi} \mathbf{x} \stackrel{d}{=} \mathbf{x}$. Nous allons utiliser le résultat [153, p79-80 prop 2.5.5] affirmant que la fonction caractéristique $\varphi_{\mathbf{x}}$ d'un vecteur isotrope \mathbf{x} doit avoir la forme ci-dessous :

$$\forall \mathbf{u} \in \mathbb{C}^K, \varphi_{\mathbf{x}}(\mathbf{u}) = \exp(-c^\alpha \|\mathbf{u}\|_H^\alpha) \quad (6.3)$$

où $\|\mathbf{u}\|_H = \sqrt{\mathbf{u}^* \mathbf{u}}$ est la norme hermitienne. Pour commencer, plaçons-nous dans le cas particulier où $\mathbf{x} = [s_1, \dots, s_K]^T$ avec les Υ_k distincts entre eux. Grâce à l'indépendance des s_k et à leurs caractères isotropes, on a pour tout $\phi \in \mathbb{R}$ et pour tout $\mathbf{u} \triangleq [u_1, \dots, u_K]^T \in \mathbb{C}^K$:

$$\begin{aligned}
\varphi_{e^{i\phi} \mathbf{x}}(\mathbf{u}) &= \mathbb{E} \left[\exp \left(i \Re \left(e^{i\phi} \mathbf{u}^* \mathbf{x} \right) \right) \right] \\
&= \prod_{k=1}^K \mathbb{E} \left[\exp \left(i \Re \left(e^{i\phi} u_k^* s_k \right) \right) \right] \\
&= \prod_{k=1}^K \varphi_{e^{i\phi} s_k}(u_k) \\
&= \prod_{k=1}^K \varphi_{s_k}(u_k) \\
&= \prod_{k=1}^K \mathbb{E} \left[\exp \left(i \Re \left(u_k^* s_k \right) \right) \right] \\
&= \prod_{k=1}^K \exp \left(-\Upsilon_k |u_k|^\alpha \right) \\
&= \exp \left(-\sum_{k=1}^K \Upsilon_k |u_k|^\alpha \right). \tag{6.4}
\end{aligned}$$

Ainsi, si l'on compare les fonctions caractéristiques en (6.4) et (6.3), nous avons prouvé que le vecteur \mathbf{x} est dans ce cas anisotrope. Cependant, il est relativement simple de remarquer que \mathbf{x} est $S\alpha S$. En effet, si on reprend un vecteur $\mathbf{x} = \sum_{l=1}^L \mathbf{a}_l s_l$ et pour tout l $s_l \sim S\alpha S_c(\Upsilon_l)$, nous voyons bien que :

- \mathbf{x} et $-\mathbf{x}$ ont même distribution car les s_l sont $S\alpha S_c$.
- Montrons que \mathbf{x} reste une distribution α -stable. Si nous posons $\mathbf{a}_l = [a_{l1}, \dots, a_{lK}]^T \in \mathbb{C}^K$, alors d'après 2.1.1 page 11 $a_{lk} s_l \sim S\alpha S(|a_{lk}|^\alpha \Upsilon_l)$. Pour tout k , la somme $\sum_{l=1}^L a_{lk} s_l$ des distributions $S\alpha S$ permet d'en déduire que chaque composante de $\mathbf{x} = [x_1, \dots, x_K]^T$ est une variable $S\alpha S$. Les coefficients \mathbf{c} et \mathbf{d} dans la définition 2.1.4 dépendent alors de ceux associés aux variables $S\alpha S$ x_1, \dots, x_K .

6.2 Fonction caractéristique et mesure spatiale discrète (MSD)

Reprenons le vecteur aléatoire $\mathbf{x} \triangleq [x_1, \dots, x_K]^T$ et les variables aléatoires s_1, \dots, s_L définis respectivement en (6.2) et (6.1). Les travaux démontrés et présentés ici sont inspirés de résultats dans le cas de mélanges linéaires $S\alpha S$ réels [153, p69-70]. Le but de cette section est d'expliquer comment les paramètres Υ_l peuvent correspondre à une information sur la localisation spatiale des sources. Pour ce faire, calculons de prime abord la fonction caractéristique de \mathbf{x} pour tout $\mathbf{u} \triangleq [u_1, \dots, u_K]^T \in \mathbb{C}^K$:

$$\varphi_{\mathbf{x}}(\mathbf{u}) = \mathbb{E} \left[\exp \left(i \Re \left(\mathbf{u}^* \mathbf{x} \right) \right) \right].$$

Nous obtenons les égalités successives suivantes :

$$\begin{aligned} \forall \mathbf{u} \triangleq [u_1, \dots, u_K]^\top \in \mathbb{C}^K, \varphi_{\mathbf{x}}(\mathbf{u}) &= \prod_{l=1}^L \mathbb{E} \left[\exp \left(i \Re \left(\sum_{k=1}^K u_k^* a_{kl} s_l \right) \right) \right] \\ &= \prod_{l=1}^L \varphi_{s_l} \left(\sum_{k=1}^K u_k^* a_{kl} \right) \\ &= \prod_{l=1}^L \exp \left(-\Upsilon_l \left| \sum_{k=1}^K u_k^* a_{kl} \right|^\alpha \right). \end{aligned}$$

Si on pose $\langle \mathbf{u}, \mathbf{a}_l \rangle = \sum_{k=1}^K u_k^* a_{kl}$ comme étant le produit scalaire sur \mathbb{C}^K , on obtient alors que :

$$\forall \mathbf{u} \in \mathbb{C}^K, \varphi_{\mathbf{x}}(\mathbf{u}) = \exp \left(- \sum_{l=1}^L \Upsilon_l |\langle \mathbf{u}, \mathbf{a}_l \rangle|^\alpha \right). \quad (6.5)$$

L'équation (6.5) nous permet alors de démontrer la représentation spatiale suivante pour \mathbf{x} :

Proposition 6.2.1 (Représentation spatiale d'un mélange linéaire α -stable). *En reprenant le modèle en (6.2), la fonction caractéristique $\varphi_{\mathbf{x}}$ induit la représentation spatiale*

$$\forall \mathbf{u} \in \mathbb{C}^K, \varphi_{\mathbf{x}}(\mathbf{u}) = \exp \left(- \int_{\mathcal{S}_{\mathbb{C}}^K} |\langle \mathbf{u}, \boldsymbol{\theta} \rangle|^\alpha \Gamma(d\boldsymbol{\theta}) \right) \quad (6.6)$$

où

$$\Gamma \stackrel{\approx}{=} \frac{1}{2} \sum_{l=1}^L \Upsilon_l \|\mathbf{a}_l\|_H^\alpha (\delta_{\tilde{\mathbf{a}}_l} + \delta_{-\tilde{\mathbf{a}}_l}) \quad (6.7)$$

avec $\delta_{\tilde{\mathbf{a}}_l}$ la mesure de Dirac sur $\mathcal{S}_{\mathbb{C}}^K$ centrée en $\tilde{\mathbf{a}}_l \triangleq \mathbf{a}_l / \|\mathbf{a}_l\|_H$. La relation d'équivalence $\stackrel{\approx}{=}$, issue du caractère anisotrope de \mathbf{x} , est définie comme suit : pour toute fonction continue f définie sur $\mathcal{S}_{\mathbb{C}}^K$, $\int_{\mathbf{a} \in \mathcal{S}_{\mathbb{C}}^K} f(\mathbf{a}) \delta_{\tilde{\mathbf{a}}_l}(\mathbf{a}) = \frac{1}{2\pi} \int_{z \in \mathcal{S}_{\mathbb{C}}^1} f(z\tilde{\mathbf{a}}_l) dz$. Autrement dit, la relation $\stackrel{\approx}{=}$ affirme que la mesure Γ est définie à changement de phase près. Le vecteur $\boldsymbol{\Upsilon} = [\Upsilon_1, \dots, \Upsilon_L]^\top \in \mathbb{R}_+^L$ est appelé mesure spatiale discrète (MSD).

Démonstration. En injectant l'équation (6.7) dans (6.6), on obtient pour tout $\mathbf{u} \in \mathbb{C}^K$:

$$\begin{aligned} \varphi_{\mathbf{x}}(\mathbf{u}) &= \exp \left(- \int_{\mathcal{S}_{\mathbb{C}}^K} |\langle \mathbf{u}, \boldsymbol{\theta} \rangle|^\alpha \Gamma(d\boldsymbol{\theta}) \right) \\ &= \exp \left(- \frac{1}{2} \sum_{l=1}^L \Upsilon_l \|\mathbf{a}_l\|_H^\alpha \left[\left| \left\langle \mathbf{u}, \frac{\mathbf{a}_l}{\|\mathbf{a}_l\|_H} \right\rangle \right|^\alpha + \left| \left\langle \mathbf{u}, \frac{-\mathbf{a}_l}{\|\mathbf{a}_l\|_H} \right\rangle \right|^\alpha \right] \right) \\ &= \exp \left(- \sum_{l=1}^L \Upsilon_l |\langle \mathbf{u}, \mathbf{a}_l \rangle|^\alpha \right). \end{aligned}$$

La dernière égalité est bien celle de la fonction caractéristique d'un mélange linéaire de variables $S\alpha S_{\mathbb{C}}$ (6.5). □

Ainsi, un tel mélange multivarié peut être vu comme la décomposition selon plusieurs directions d'arrivées \mathbf{a}_l , chacune pondérée par l'énergie propre Υ_l de la source s_l correspondante. Nous expliquons dans le chapitre 7 comment l'estimation de $\boldsymbol{\Upsilon}$ peut être effectuée grâce à la proposition 6.2.1.

La somme des diracs de direction opposées est instaurée dans (6.7) pour que \mathbf{x} reste symétrique.

6.3 Utilisation de la MSD en traitement du signal audio : historique

Différents travaux ont visé à exploiter la MSD pour la séparation de signaux de paroles [94, 95]. Dans [95], le modèle considéré est un modèle similaire à (6.1) et (6.2) mais dans le cas de variables aléatoires $S\alpha S$ réelles. On dit qu'un tel modèle de mélange est un *mélange linéaire instantané* [44, 7] :

$$\forall n, \begin{cases} \mathbf{x}(n) = \mathbf{A} \mathbf{s}(n) \\ \mathbf{s}(n) = [s_1(n), \dots, s_L(n)]^\top \\ s_l(n) \sim S\alpha S(\Upsilon_l) \end{cases}, \quad (6.8)$$

où les sources $s_l(n)$ sont supposés indépendantes et identiquement distribuées (i.i.d.). Durant les années 80, certains chercheurs se sont intéressés à résoudre le problème dit de « soirée cocktail ». Dans une soirée composée de plusieurs personnes discutant en petit groupe, on cherche à séparer le signal de parole de chaque personne (le reste étant considéré comme un bruit parasite). Sous le formalisme de l'*analyse en composantes indépendantes* (ACI), nous supposons que les sources s_l sont indépendantes (c'est à dire, chaque discours est indépendant). De plus, il est généralement admis que le nombre de capteurs $K \geq L$ pour pouvoir séparer chaque s_l en utilisant la pseudo-inverse \mathbf{A}^\dagger . Le problème se ramène alors à estimer la matrice \mathbf{A} . Le modèle en (6.8) est dit linéaire instantané non bruité. Plusieurs algorithmes basés sur des modèles gaussiens et des calculs de cumulants ont vu le jour [32, 86, 167]. Malheureusement dans le cas α -stable, ces méthodes ne sont pas applicables étant donné que les moments d'ordre $p \geq \alpha$ n'existent pas. Cependant, comme nous venons de voir dans le cas complexe, nous pouvons tenter d'estimer Υ afin de trouver les directions d'arrivées de chaque source. Grâce à cette information, la *matrice de mélange* \mathbf{A} peut être estimée et les sources s_l par la même occasion. Nous laissons le soin au lecteur de découvrir la méthode employée dans [95] pour estimer Υ . À noter que nous nous inspirons de quelques éléments de cette méthode pour l'estimation dans le cas $S\alpha S_c$ développée au chapitre suivant.

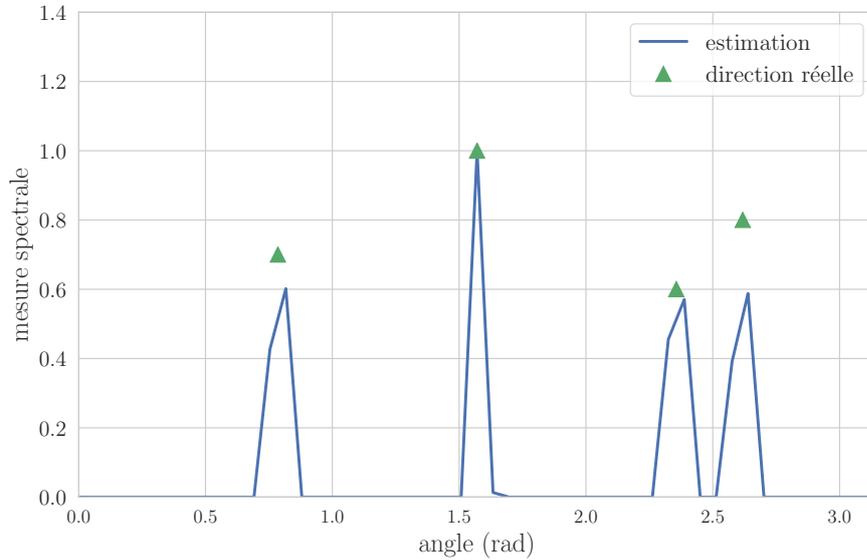


FIGURE 6.1 – Estimation de la mesure spectrale pour un mélange linéaire instantané.

Afin d'illustrer les travaux effectués par Kidmose dans [94], nous avons considéré 4 sources scalaires comme étant des variables $S\alpha S$. Elles sont placées à différents angles d'arrivée $\frac{\pi}{4}$, $\frac{\pi}{2}$, $\frac{3\pi}{4}$ et $\frac{5\pi}{6}$ dans un

cadre stéréophonique ($K = 2$). Enfin, la *matrice de mélange* [5, 17] $\mathbf{A} \in \mathbb{R}^{2 \times L}$ est

$$\mathbf{A} = \begin{bmatrix} \cos\left(\frac{\pi}{4}\right) & \cos\left(\frac{\pi}{2}\right) & \cos\left(\frac{3\pi}{4}\right) & \cos\left(\frac{5\pi}{6}\right) \\ \sin\left(\frac{\pi}{4}\right) & \sin\left(\frac{\pi}{2}\right) & \sin\left(\frac{3\pi}{4}\right) & \sin\left(\frac{5\pi}{6}\right) \end{bmatrix}. \quad (6.9)$$

En audio stéréophonique, le placement de sources dans l'espace sonore est appelé panoramique.

Pour cette expérience, un nombre de $N = 500$ échantillons α -stables a été considéré et la MSD est égale à $\Upsilon = [0.7, 1, 0.6, 0.8]^\top$.

Nous employons la méthode de [94] visant à résoudre le problème d'optimisation suivant

$$\min_{\Upsilon \geq 0} \|\hat{\mathbf{I}} - 2\mathfrak{K}(\Psi)\Upsilon\|_2 \quad (6.10)$$

où $\hat{\mathbf{I}}$ est un estimateur de l'exposant de Lévy $-\log \varphi_x$, Ψ est une matrice composée des coefficients $\langle \mathbf{u}, \boldsymbol{\theta} \rangle^\alpha$ où $\boldsymbol{\theta} \in \mathbb{S}_{\mathbb{R}}^K$ et $\mathbf{u} \in \mathbb{R}^K$. Nous constatons alors sur la Fig. 6.1 que la MSD pointe bien vers les angles annoncés de la matrice \mathbf{A} . Cependant, ce modèle reste limité dans le sens où seules les sources réelles et i.i.d. ainsi que les mélanges linéaires instantanés sont pris en compte, ce qui sont des hypothèses assez limitantes en audio. Nos travaux dans [60, 61] étendent la méthode d'estimation dans le cadre temps-fréquence. Par ailleurs, nous avons développé notre propre approche pour estimer Υ . Cette dernière décrit une complexité algorithmique plus faible que celles proposées dans l'état de l'art [94, 138].

Chapitre 7

Estimation de la MSD

Résumé

Comme établi précédemment, l'information spatiale pour un mélange linéaire α -stable multivarié peut être caractérisée grâce au vecteur composé des énergies de chaque source s_l , dans les directions induites par les vecteurs \mathbf{a}_l . Ce vecteur, appelé *mesure spatiale discrète* (MSD), peut être estimé grâce à une stratégie peu coûteuse en calcul. Tout d'abord, les observations \mathbf{x} sont résumées via un opérateur appelé *exposant de Lévy* [172] défini en (2.9). Ensuite, cet *exposant de Lévy* est utilisé pour estimer la MSD. Un tel opérateur résumant les observations et permettant l'obtention des paramètres souhaités est dit de *sketching* [92]. Les contributions présentées ici incluent un nouvel estimateur de l'*exposant de Lévy* spécifique au cas α -stable ainsi qu'une nouvelle technique pour l'estimation de la MSD, plus stable et moins coûteuse en temps de calcul que celles présentées en [138, 95].

7.1 Introduction

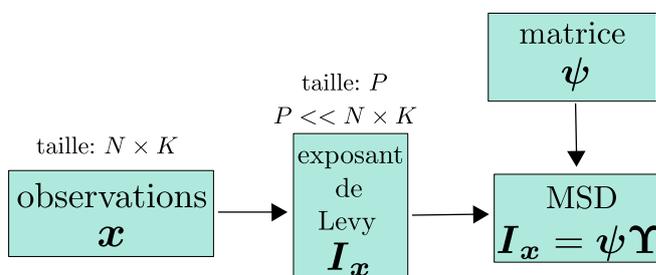


FIGURE 7.1 – Stratégie d'estimation de la MSD

Supposons que l'on observe N vecteurs comme suit :

$$\forall n \in \llbracket 1, N \rrbracket, \mathbf{x}(n) = \mathbf{A} \mathbf{s}(n) \quad (7.1)$$

où $\mathbf{A} \in \mathbb{C}^{K \times L}$ est une matrice fixée et les $\mathbf{s}(n) \triangleq [s_1(n), \dots, s_L(n)]^\top$ sont des vecteurs aléatoires indépendants et identiquement distribués (i.i.d.). En adéquation avec (6.8), les composantes vectorielles des $\mathbf{s}(n)$ sont $S\alpha S_C$:

$$\forall n \in \llbracket 1, N \rrbracket, \forall l \in \llbracket 1, L \rrbracket, s_l(n) \sim S\alpha S_C(\Upsilon_l). \quad (7.2)$$

Dorénavant, nous supposons être dans le cas non-gaussien ($\alpha < 2$). Notre but est d'estimer la *mesure spatiale discrète* (MSD) $\Upsilon = [\Upsilon_1, \dots, \Upsilon_L]^T$. La méthode que nous décrivons par la suite est résumée via la Fig. 7.1. En général, nous pouvons proposer une technique de maximum de vraisemblance. Autrement dit, maximiser $\mathbb{P}(\mathbf{X} \triangleq [\mathbf{x}(1), \dots, \mathbf{x}(N)] \mid \Upsilon)$. Malheureusement, cette technique tombe en défaut dans la plupart des cas car il n'existe pas d'expression analytique pour la vraisemblance. De plus, une méthode employant des moments d'ordre ≥ 2 n'est pas envisageable. Autrement, nous pouvons considérer les *moments d'ordres faibles fractionnaires* (MOFF) présentés dans la section 4.4.1. Nous avons cependant expliqué dans la section 4.4.1 que l'exploitation des MOFF n'est valable qu'en présence d'un faible nombre de sources s_l . La fonction caractéristique peut s'avérer une solution élégante car, d'après (1.7), elle contient sous réserve d'existence les moments à tout ordre d'un vecteur aléatoire. Dans le cas d'un mélange multivarié linéaire de sources α -stables scalaires non-gaussien \mathbf{x} , nous avons montré en (6.5) que la fonction caractéristique $\varphi_{\mathbf{x}}$ admet une forme analytique simple. D'après l'équation (6.5), la fonction caractéristique $\varphi_{\mathbf{x}}$ évaluée en $\mathbf{u} \in \mathbb{C}^K$ est l'exponentielle d'une combinaison linéaire des vecteurs colonnes $\mathbf{a}_1, \dots, \mathbf{a}_L \in \mathbb{C}^K$ de la matrice \mathbf{A} et de la MSD Υ . Les vecteurs \mathbf{u} sont alors appelés *fréquences caractéristiques*. L'équation (6.5) peut se réécrire :

$$\forall \mathbf{u} \in \mathbb{C}^K, I_{\mathbf{x}}(\mathbf{u}) = \sum_{l=1}^L \Upsilon_l |\langle \mathbf{u}, \mathbf{a}_l \rangle|^\alpha, \quad (7.3)$$

où $I_{\mathbf{x}}(\mathbf{u}) \triangleq -\ln(\varphi_{\mathbf{x}}(\mathbf{u}))$ est l'*exposant de Lévy* du vecteur \mathbf{x} évalué en \mathbf{u} . La technique d'estimation proposée est alors de choisir un certain nombre de fréquences caractéristiques \mathbf{u} , et mesurer les $I_{\mathbf{x}}(\mathbf{u})$ correspondants aux données observées \mathbf{X} . Considérons P vecteurs $\mathbf{u}_1, \dots, \mathbf{u}_P \in \mathbb{C}^K$ et $\mathbf{I}_{\mathbf{x}} \triangleq [I_{\mathbf{x}}(\mathbf{u}_1), \dots, I_{\mathbf{x}}(\mathbf{u}_P)]^T \in \mathbb{R}_+^P$. Définissons alors la matrice :

$$\Psi = \begin{bmatrix} |\langle \mathbf{u}_1, \mathbf{a}_1 \rangle|^\alpha & \dots & |\langle \mathbf{u}_1, \mathbf{a}_P \rangle|^\alpha \\ \vdots & \vdots & \vdots \\ |\langle \mathbf{u}_P, \mathbf{a}_1 \rangle|^\alpha & \dots & |\langle \mathbf{u}_P, \mathbf{a}_P \rangle|^\alpha \end{bmatrix} \in \mathbb{R}_+^{P \times L}. \quad (7.4)$$

D'après (7.3), il en découle l'égalité

$$\mathbf{I}_{\mathbf{x}} = \Psi \Upsilon. \quad (7.5)$$

L'équation (7.5) nous révèle une procédure naturelle pour l'estimation de Υ . Après avoir fixé la matrice Ψ et estimé $\mathbf{I}_{\mathbf{x}}$, nous pouvons proposer un estimateur du vecteur Υ . La méthode d'estimation sera décrite plus en détail dans la section 7.2.3. .

Notre souhait est de pouvoir contrôler le nombre P de *fréquences caractéristiques* nécessaires à l'estimation du vecteur $\mathbf{I}_{\mathbf{x}}$. Il est important de remarquer que P est totalement indépendant du nombre d'observations N et de la dimension K des vecteurs. En somme, le vecteur $\mathbf{I}_{\mathbf{x}}$ résume les données \mathbf{X} . Dans la littérature, l'*exposant de Lévy* peut alors être qualifié d'opérateur de *sketching* [92]. Brièvement, le sketching est une méthode proposant un opérateur résumant un grand nombre de données issues d'un modèle de mélange probabiliste. Par exemple, le but peut être de réduire la complexité d'un algorithme exploitant les données observées. En effet, si la taille de l'opérateur de *sketching* est plus petite par rapport à celle des données observées, cela réduira considérablement le temps de calcul. Un exemple concret est la fonction caractéristique d'un mélange gaussien qui est présentée comme étant un opérateur dit de *sketching* dans [92]. Un autre aspect important des méthodes dites de *sketching* est que les données observées ne doivent être utilisées qu'une seule fois pour l'estimation des paramètres. C'est le cas ici car les *exposants de Lévy* ne sont calculés qu'une seule fois pour l'estimation d' Υ . Enfin, il est important de choisir un échantillonnage de *fréquences caractéristiques* $\mathbf{u}_1, \dots, \mathbf{u}_P$ pour résumer convenablement les données \mathbf{X} . Par exemple dans [98], Koutrouvelis souligne l'importance de sélectionner un grand nombre de *fréquences caractéristiques* autour de zéro pour estimer une fonction caractéristique. Pour des variables $S\alpha S$ réelles, Nikias décrit dans [136] un échantillonnage astucieux afin d'obtenir la meilleure estimation

« sketching »
pourrait ici être
traduit comme
« résumé ».

de la fonction caractéristique. Si nous supposons la matrice \mathbf{A} connue, un choix convenable serait de prendre les colonnes de \mathbf{A} comme fréquences caractéristiques. Cela engendrerait une matrice Ψ carrée et symétrique et a fortiori une plus grande stabilité pour l'estimation de Υ .

7.2 Estimation de l'exposant de Lévy

Intéressons-nous à l'estimation des exposants de Lévy pour un certain nombre de *fréquences caractéristiques*. Nous proposons deux estimateurs. Le premier est basé sur une estimation « naïve » ne prenant pas en compte le modèle probabiliste. Proche de l'estimateur général appelé *fonction caractéristique empirique* [46], nous verrons que ce dernier souffre de quelques défauts. Le second estimateur exploite la forme particulière des fonctions caractéristiques de distributions $S\alpha S$. L'avantage de ce deuxième est alors de conserver la non-négativité de l'*exposant de Lévy* théorique.

7.2.1 Estimation « naïve »

Soit $\mathbf{x}(1), \dots, \mathbf{x}(N)$ des réalisations i.i.d. d'un vecteur aléatoire \mathbf{x} . L'estimateur ci-dessous permet une première estimation de $I_{\mathbf{x}}$:

$$\forall \mathbf{u} \in \mathbb{C}^K, \hat{I}_{\mathbf{x}}^{(1)}(\mathbf{u}) \triangleq -\ln \left[\frac{1}{N} \sum_{n=1}^N \exp(i\Re(\langle \mathbf{u}, \mathbf{x}(n) \rangle)) \right]. \quad (7.6)$$

Remarque Dans le cas d'un vecteur aléatoire \mathbf{x} symétrique α -stable complexe comme décrit en (6.8), l'équation (7.5) affirme que l'exposant de Lévy théorique est réel. Pour prendre en compte cet aspect, une estimation alternative utilisée dans [95, 136, 138] peut être proposée :

$$\forall \mathbf{u} \in \mathbb{C}^K, \hat{I}_{\mathbf{x}}^{(2)}(\mathbf{u}) \triangleq -\ln \left[\frac{1}{N} \sum_{n=1}^N \cos(\Re(\langle \mathbf{u}, \mathbf{x}(n) \rangle)) \right]. \quad (7.7)$$

Même si les deux *exposants de Lévy* empiriques en (7.6) et (7.7) semblent satisfaisants, il n'en demeure pas moins qu'ils souffrent en pratique de différents problèmes qu'il est important de souligner. En théorie, d'après (7.5), l'*exposant de Lévy* se doit d'être non-négatif. Or, rien n'assure cette positivité dans les deux formules précédentes. Considérer le cosinus dans (7.7) pour s'assurer une estimation à valeurs réelles reste plus judicieux. Afin de palier à ce défaut, nous proposons un nouvel estimateur valable uniquement dans la théorie α -stable.

7.2.2 Estimation dans le cas symétrique α -stable

L'estimateur que nous proposons satisfait les conditions de non-négativité de l'*exposant de Lévy* :

Proposition 7.2.1 (Estimateur de l'exposant de Lévy $S\alpha S$). *Soit $\mathbf{x}(1), \dots, \mathbf{x}(N)$ des réalisations i.i.d. d'un vecteur $S\alpha S$ complexe \mathbf{x} . Alors l'estimateur :*

$$\forall \mathbf{u} \in \mathbb{C}^K, \hat{I}_{\mathbf{x}}(\mathbf{u}) = -\ln \left[\left| \frac{1}{N} \sum_{n=1}^N \exp\left(\frac{i\Re(\langle \mathbf{u}, \mathbf{x}(n) \rangle)}{2^{1/\alpha}}\right) \right|^2 \right] \quad (7.8)$$

est un estimateur de l'exposant de Lévy. Il est positif, car nous calculons l'opposé logarithmique d'un nombre réel compris entre 0 et 1.

Démonstration. Pour tout $n \in \llbracket 1, N \rrbracket$ et $n_1 \neq n_2 \in \llbracket 1, N \rrbracket$, il suffit de remarquer que les $\mathbf{x}(n)$ et $\frac{\mathbf{x}(n_1) - \mathbf{x}(n_2)}{2^{1/\alpha}}$ suivent la même loi. L'estimateur (7.6) des vecteurs $\frac{\mathbf{x}(n_1) - \mathbf{x}(n_2)}{2^{1/\alpha}}$ est exactement de la forme souhaitée en (7.8). Il en est de même pour $\mathbf{x}(n)$, car deux vecteurs ont même lois si et seulement s'ils ont même exposant de Lévy. \square

Ainsi, nous avons un estimateur de I_x non-négatif uniquement valable dans le cas α -stable.

Remarque Il serait intéressant d'étudier le biais et la variance de l'estimateur en (7.8) par rapport à celui de (7.6) ou (7.7). Cependant, les calculs effectués ne nous ont pas permis de comparer les deux valeurs statistiques sus-mentionnées pour chaque estimateur.

7.2.3 Problème d'optimisation

Via la relation (7.5), le problème d'optimisation que nous souhaitons résoudre est le suivant

$$\hat{\Upsilon} \leftarrow \underset{\Upsilon \geq 0}{\operatorname{argmin}} d(\hat{I}_x | \Psi \Upsilon), \quad (7.9)$$

où d est une divergence quelconque. Grâce à l'estimateur présenté en (7.8), chaque terme de l'équation (7.5) est assuré d'être non-négatif. C'est pourquoi nous employons une technique d'optimisation garantissant cette contrainte. Considérons à nouveau le modèle décrit en (7.1) et (7.2). Nous sélectionnons $\mathbf{u}_1, \dots, \mathbf{u}_P$ des fréquences caractéristiques pour lesquelles l'exposant de Lévy \hat{I}_x sera calculé. Nous obtenons alors :

$$\hat{I}_x \simeq \Psi \Upsilon \quad (7.10)$$

où

$$\hat{I}_x = [I_x(\mathbf{u}_1), \dots, I_x(\mathbf{u}_P)]^T \in \mathbb{R}_+^{P \times 1}; \quad (7.11)$$

$$\Psi = \begin{bmatrix} |\langle \mathbf{u}_1, \mathbf{a}_1 \rangle|^\alpha & \dots & |\langle \mathbf{u}_1, \mathbf{a}_L \rangle|^\alpha \\ \vdots & \vdots & \vdots \\ |\langle \mathbf{u}_P, \mathbf{a}_1 \rangle|^\alpha & \dots & |\langle \mathbf{u}_P, \mathbf{a}_L \rangle|^\alpha \end{bmatrix} \in \mathbb{R}_+^{P \times L}; \quad (7.12)$$

$$\Upsilon = [\Upsilon_1, \dots, \Upsilon_L]^T \in \mathbb{R}_+^{L \times 1}. \quad (7.13)$$

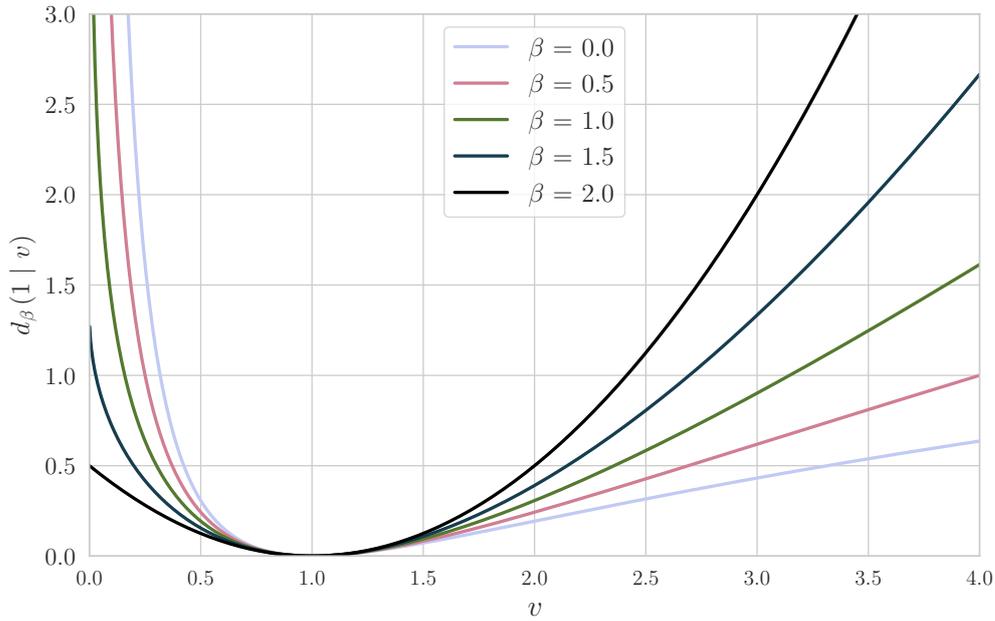
En supposant connus Ψ et \hat{I}_x , le but est d'estimer Υ . Une première idée pour résoudre ce problème serait d'appliquer une pseudo-inverse de Moore-Penrose [144] à la matrice Ψ . Cependant, Ψ est mal conditionnée (au sens du rapport entre la valeur singulière maximale et la valeur singulière minimale [147]). Une telle inversion directe pourrait entraîner des erreurs d'estimation importantes. Afin de résoudre ce problème, Nolan propose dans [138] une matrice Ψ mieux conditionnée en utilisant une grille $([\mathbf{u}_p, \mathbf{a}_l])_{p,l}$ symétrique et en exploitant des propriétés de l'estimateur de Lévy $\hat{I}_x^{(1)}$ (7.6). Cependant, cette méthode n'est valable que dans le cas de vecteurs $S\alpha S$ réels et ne s'étend pas naturellement au cas complexe.

Notons de plus que comparé à l'approche de Nolan, l'estimateur \hat{I}_x proposé dans ce manuscrit est non-négatif. Par ailleurs, les matrices Ψ et la MSD Υ sont également non-négatifs. C'est pourquoi un schéma similaire aux méthodes de *nonnegative matrix factorization* (NMF) [51] présentées dans (4.28) est proposé. Définissons dans un premier temps une classe de divergences appelée β -divergence [51] :

Définition 7.2.1 (β -divergence). Une β -divergence est définie comme suit :

$$d_\beta(u | v) \triangleq \begin{cases} \frac{1}{\beta(\beta-1)} (u^\beta + (\beta-1)v^\beta - \beta uv^{\beta-1}) & \beta \in \mathbb{R} \setminus \{0, 1\} \\ u \ln\left(\frac{u}{v}\right) - u + v & \beta = 1 \text{ (Kullback-Leibler)} \\ \frac{u}{v} - \ln\left(\frac{u}{v}\right) - 1 & \beta = 0 \text{ (Itakura-Saito)} \end{cases} \quad (7.14)$$

Kidmose propose la même année une méthode similaire dans [95].

FIGURE 7.2 – courbe des β -divergences $v \mapsto d_\beta(1|v)$.

Ces divergences (cf. Fig. 7.2) sont un cas particulier des *divergences de Bregman* [25], possédant diverses propriétés de convexité et de linéarité (pour $\beta \in [1, 2]$, d_β est une fonction convexe). Dans le cas d'un problème de NMF, il a été prouvé dans [51] que pour $\beta \in [0, 2]$, les solutions obtenues par les méthodes de mises à jour multiplicatives convergent vers le même minimum local.

Nous choisissons d'utiliser les β -divergences dans (7.9) afin de minimiser la « distance » (au sens des β -divergences) entre l'estimateur $\hat{\mathbf{I}}_x$ et l'expression théorique $\Psi\mathbf{Y}$ dans (7.10) :

$$\hat{\mathbf{Y}} \leftarrow \underset{\mathbf{Y} \geq 0}{\operatorname{argmin}} d_\beta(\hat{\mathbf{I}}_x | \Psi\mathbf{Y}). \quad (7.15)$$

Notre choix se justifie par le fait que notre problème est convexe car la matrice Ψ est supposée connue et que seul \mathbf{Y} reste inconnu [51]. Nous sommes par conséquent garantis d'une convergence vers un minimum global.

7.2.4 Mise à jour des paramètres

Abordons maintenant la résolution du problème d'optimisation en (7.15). Nous posons

$$C(\mathbf{Y}) = d_\beta(\hat{\mathbf{I}}_x | \Psi\mathbf{Y}). \quad (7.16)$$

Le calcul du gradient $\nabla_{\mathbf{Y}}[C(\mathbf{Y})]$ donne

$$\forall \mathbf{Y} \in \mathbb{R}_+^L, \nabla_{\mathbf{Y}}[C(\mathbf{Y})] = \Psi^\top (\Psi\mathbf{Y})^{\odot(\beta-1)} - \Psi^\top \left((\Psi\mathbf{Y})^{\odot(\beta-2)} \odot \hat{\mathbf{I}}_x \right), \quad (7.17)$$

où $a \odot b$ et $a^{\odot b}$ désignent respectivement la multiplication et l'exponentiation terme à terme. Comme dans le cas d'une divergence d'Itakura-Saito de la sous-section 4.2.2 page 33, nous appliquons une règle de

mise à jour multiplicative. Cela garanti la non-négativité de $\hat{\Upsilon}$ d'une part mais également la convergence de $\hat{\Upsilon}$ vers une unique solution pour tout $\beta \in [0, 2]$. La méthode pour calculer $\hat{\Upsilon}$ est donnée par la mise à jour

$$\hat{\Upsilon} \leftarrow \hat{\Upsilon} \odot \frac{\Psi^\top \left((\Psi \Upsilon)^{\odot(\beta-2)} \odot \hat{\mathbf{I}}_x \right)}{\Psi^\top (\Psi \Upsilon)^{\odot(\beta-1)}}. \quad (7.18)$$

Algorithme 3 Estimation de la MSD Υ

1. Entrées

- Données observées $\mathbf{x}(1), \dots, \mathbf{x}(N)$.
- Fréquences caractéristiques $\mathbf{u}_1, \dots, \mathbf{u}_P$.
- Matrice \mathbf{A} .
- Nombre d'itérations pour les mise à jour multiplicatives.
- Paramètre d'échelle α et divergence ($\beta \in [0, 2]$) à utiliser.

2. Sketching

Calculer $\hat{\mathbf{I}}_x \triangleq [\hat{I}_x(u_1), \dots, \hat{I}_x(u_P)]^\top$ en utilisant (7.8).

3. Estimation des paramètres

- Calculer Ψ relatif à (7.12).
 - Calculer $\hat{\Upsilon}$ en itérant l'équation (7.18).
-

Ainsi, à partir d'un mélange linéaire vectoriel de variables $S\alpha S_c$, nous pouvons estimer l'énergie de chaque source s_l dans ses directions d'arrivées \mathbf{a}_l correspondantes. Soulignons à nouveau que toute la méthode d'estimation développée dans cette partie n'est valable que dans le cas où $\alpha < 2$. En réalité, dans le cas où $\alpha = 2$, la MSD existe mais n'est pas unique (cf. [153]). Le chapitre suivant montre comment le modèle α -stable associé à sa MSD introduit dans les équations (6.7), (6.2) et (6.1) peut être employé dans une tâche très courante en traitement du signal audio : la localisation de sources sonores acoustiques.

Chapitre 8

Application à la localisation de sources acoustiques

Résumé

La *mesure spatiale discrète* (MSD) précédemment définie va maintenant être appliquée à la localisation de sources acoustiques. Cette application consiste à déterminer les coordonnées spatiales d'un point ou d'un ensemble de points pouvant être considérés comme propageant des ondes acoustiques dans un milieu physique. Associé à un modèle probabiliste α -stable adéquat, nous exploitons le domaine temps-fréquence de l'observation captée par plusieurs microphones. On montre qu'estimer la MSD revient exactement à fournir l'énergie globale émise, au sens des paramètres d'une distribution α -stable, dans la pièce en chaque point d'une grille. Cette information permet de déterminer précisément où les sources « potentiellement actives » se situent. Notre approche de localisation sera comparée à trois algorithmes classiques de l'état de l'art. Les deux premiers sont des algorithmes dits de *matching pursuit* : CLEAN et RELAX. Le dernier est présenté dans le chapitre 5 page 41 et exploite les moments d'ordre 2.

Dans tout ce chapitre, on prendra l'exposant caractéristique $0 < \alpha < 2$ pour assurer l'unicité garantie de la *mesure spatiale discrète* (MSD) [153].

8.1 Modèle pour la localisation

8.1.1 Positionnement du problème et hypothèses

La localisation de sources acoustiques vise à déduire, à partir d'un modèle acoustique prédéfini et d'enregistrements audio, la position d'une ou plusieurs sources. En trois dimensions, la localisation des sources est souvent formulée comme la recherche des coordonnées sphériques des sources via l'obtention de deux angles (appelés *Azimuth* et *Elevation*) et de leur distance par rapport à une origine [146]. Une autre option est de découper l'espace considéré en un nombre fini de segments et d'inférer une énergie acoustique émise par chacun des éléments de la grille obtenue [159, 176]. Cette deuxième option aboutit à une *carte thermique* permettant la localisation : typiquement, les positions dépassant un certain seuil énergétique sont supposées contenir des « sources actives ». La Fig. 8.1 illustre cette représentation. Dans ce travail, nous choisissons cette dernière approche qui a comme avantage principal de ne pas nécessiter la connaissance *a priori* du nombre de sources. Nous faisons comme hypothèse qu'à chaque point de la grille, une source émise produit un signal α -stable harmonisable indépendant des autres. De plus, les microphones repérés également sur la grille seront supposés sans géométrie particulière mais leurs

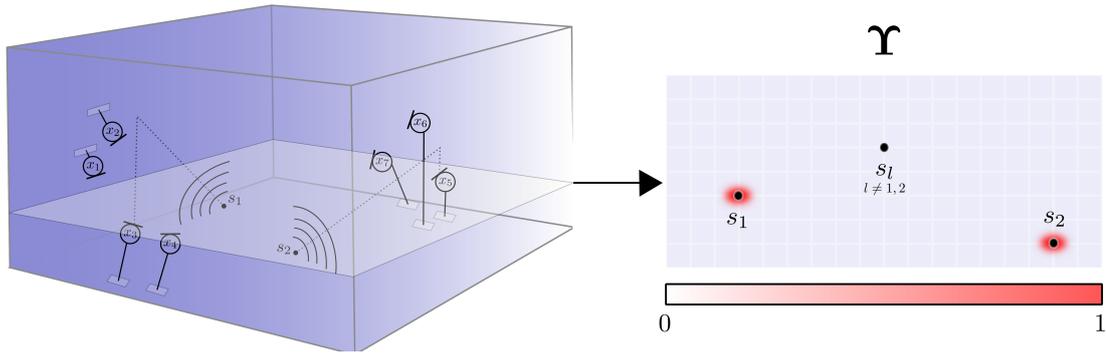


FIGURE 8.1 – Modèle de localisation et identification avec la MSD

positions sont connues. En résumé, l'estimation de la MSD nous donne l'amplitude de chaque source sur la grille sélectionnée. La réunion de toutes ces amplitudes formera alors la *carte thermique* où nous supposons que les valeurs les plus importantes seront celles révélant l'existence d'une « source active ».

8.1.2 Modèle probabiliste

Considérons une grille contenant L points. À chaque cellule $l \in \llbracket 1, L \rrbracket$ se trouve une source s_l émettant un processus α -harmonisable. Pris dans le domaine *temps-fréquence* (TF), nous avons donc pour tout $(l, f, t) \in \llbracket 1, L \rrbracket \times \llbracket 1, F \rrbracket \times \llbracket 1, T \rrbracket$:

$$s_l(f, t) \sim S\alpha S_c(\Upsilon_l(f, t)). \quad (8.1)$$

Si nous supposons que les sources sont immobiles, nous pouvons nous affranchir de la dépendance en t des paramètres d'échelle $\Upsilon_l(f, t)$. En outre, nous admettons que les paramètres d'échelles ne dépendent pas de f . Cela est possible grâce au fait que les distributions α -stables sont à queues lourdes et donc permettent une dynamique très forte. Ainsi le même paramètre d'échelle pourra suffisamment bien rendre compte de l'ensemble du spectre par le biais d'une échelle commune Υ_l . Dans le cas gaussien, au contraire, la faible capacité exploratoire des réalisations par rapport à la variance rend nécessaire un modèle plus fin et donc plus paramétré [119]. Le modèle de sources que nous considérons sera alors :

$$\forall (l, f, t), s_l(f, t) \sim S\alpha S_c(\Upsilon_l). \quad (8.2)$$

Vecteur de pointage se traduit en anglais par *steering vector* [66, 187]. Associé à ce modèle où L sources sont situées en chaque cellule de la grille, un modèle acoustique précise la propagation des signaux correspondants vers K microphones, conduisant aux observations $\mathbf{x} \triangleq [x_1, \dots, x_K]$. Pour tout point TF (f, t) , le signal observé $\mathbf{x}(f, t)$ est une combinaison linéaire des L sources contre L vecteurs $\mathbf{a}_1(f), \dots, \mathbf{a}_L(f)$. En acoustique, les vecteurs $\mathbf{a}_l(f)$ sont appelés *vecteurs de pointages* [21, 143], et *filtres de mélange* en traitement du signal audio [103]. Chaque $\mathbf{a}_l(f)$ précise le chemin acoustique emprunté par la source l pour aboutir au K microphones. Nous décrivons plusieurs modèles permettant de calculer ces vecteurs $\mathbf{a}_l(f)$ en section suivante. Pour résumer :

$$\forall (f, t), \mathbf{x}(f, t) = \sum_{l=1}^L \mathbf{a}_l(f) s_l(f, t). \quad (8.3)$$

Remarque Il est important de souligner qu'un tel modèle acoustique en (8.3) est issu de la simplification d'un modèle convolutif entre les vecteurs de pointage $\tilde{\mathbf{a}}_l$ et les sources dans le domaine temporel \tilde{s}_l . Pour que le modèle de l'équation (8.3) reste valable, la longueur de chaque filtre $\tilde{\mathbf{a}}_l$ se doit d'être plus petite que la taille de fenêtrage considérée dans la *transformée de Fourier à court*

terme (TFCT) [103]. En effet, cela permet d'approximer un modèle de mélange convolutif dans le domaine temporel en mélange instantané dans la domaine fréquentiel.

D'après les résultats présentés dans la proposition 6.2.1 page 49, étant donnés les vecteurs de pointage $\mathbf{a}_l(f)$, la MSD $\Upsilon \triangleq [\Upsilon_1, \dots, \Upsilon_L]^\top$ décrit l'énergie émise par chacune des cellules de la grille. Le but de la localisation est donc d'estimer Υ . La méthode fournie dans le chapitre 7 est adaptée à ce problème. Ici pour toute fréquence f , nous estimons Υ en exploitant la relation (7.10) pour différentes fréquences caractéristiques $\mathbf{u}_1(f), \dots, \mathbf{u}_P(f)$. Afin de simplifier le problème et les notations, nous supposons que le nombre de fréquences caractéristiques est le même que le nombre de sources (i.e. $P = L$ avec les notations de (7.11) et (7.12)). Nous définissons alors

$$\mathbf{I}_f = [I_f(\mathbf{u}_1(f)), \dots, I_f(\mathbf{u}_L(f))]^\top \quad (8.4)$$

comme étant le vecteur non-négatif de taille $L \times 1$ contenant les *exposants de Lévy* du mélange à la fréquence f pour chaque $\mathbf{u}_l(f)$. Autrement dit

$$\forall \mathbf{u} \in \mathbb{C}^K, I_f(\mathbf{u}) = \sum_{l=1}^L \Upsilon_l |\langle \mathbf{u}, \mathbf{a}_l(f) \rangle|^\alpha. \quad (8.5)$$

Les matrices carrées Ψ_f de taille $L \times L$ sont définies par

$$[\Psi_f]_{l,l'} = |\langle \mathbf{u}_l(f), \mathbf{a}_{l'}(f) \rangle|^\alpha. \quad (8.6)$$

Discutons brièvement du choix des fréquences caractéristiques $\mathbf{u}_l(f)$. En pratique, les vecteurs de pointages $\mathbf{a}_l(f)$ peuvent être connus si l'on applique un modèle acoustique à notre scénario. C'est pourquoi, nous choisissons volontairement de poser pour tout f, l , $\mathbf{u}_l(f) = \mathbf{a}_l(f)$. Comme mentionné dans la section 7.1, les matrices Ψ_f sont par conséquent symétriques et carrées. Une autre raison est que les vecteurs de pointages sont en pratiques échantillonnés dans l'espace réel de dimension 3. Les fréquences caractéristiques quelconques sont quant à elles échantillonnées sur \mathbb{C}^K . Grâce à cette astuce d'exploiter un modèle acoustique particulier, nous soulevons le problème dit du *fléau de la dimension* [16, 139]. Dans tous les cas, la combinaison des équations (8.4) et (8.6), reliées à l'égalité (8.5) permet d'affirmer que $\mathbf{I}_f = \Psi_f \Upsilon$. En concaténant tous les vecteurs \mathbf{I}_f en un vecteur \mathbf{I}_x de taille FL et tous les Ψ_f en une matrice Ψ de taille $FL \times L$, nous obtenons

$$\mathbf{I}_x = \Psi \Upsilon. \quad (8.7)$$

La stratégie d'estimation de l'exposant de Lévy est semblable à celle présentée dans la sous-section 7.2.2 page 55 du chapitre précédent. Pour toute fréquence caractéristique \mathbf{u} , nous obtenons alors une estimée de l'exposant de Lévy $I_f(\mathbf{u})$ par :

$$\forall \mathbf{u} \in \mathbb{C}^K, \hat{I}_f(\mathbf{u}) = -\ln \left| \frac{1}{T} \sum_{t=1}^T \exp \left(i \frac{\Re \langle \mathbf{u}, \mathbf{x}(f, t) \rangle}{2^{1/\alpha}} \right) \right|^2 \quad (8.8)$$

où nous le calculons pour les fréquences caractéristiques $\mathbf{u}_1(f), \dots, \mathbf{u}_L(f)$. L'équation (8.7) se réécrit

$$\hat{\mathbf{I}}_x \simeq \Psi \Upsilon \quad (8.9)$$

où $\hat{\mathbf{I}}_x$ est la concaténation de tous les $\hat{\mathbf{I}}_f$ de taille FL .

Remarque On voit apparaître ici clairement l'opérateur de *sketching* mentionné dans le chapitre précédent. En effet, l'observation de taille $F \times T \times K$ est résumée dans l'exposant de Lévy de taille $FL \times 1$. La procédure d'estimation de ce *sketch* est parallélisable à la fois le long de l'axe temporel et fréquentiel.

Considérons le problème d'optimisation proposé en (7.9). Une contrainte que nous pouvons ajouter sur l'estimation de Υ est celle d'être parcimonieux (c'est-à-dire, nul presque partout). En effet, le nombre de points sur la grille est bien supérieur au nombre de sources à localiser. Un exemple simple d'une telle régularisation est la norme ℓ_1 [23, 134] :

$$\hat{\Upsilon} \leftarrow \underset{\Upsilon \geq 0}{\operatorname{argmin}} d_\beta(\hat{\mathbf{I}}_x | \Psi \Upsilon) + \lambda \|\Upsilon\|_1, \quad (8.10)$$

où $\|\Upsilon\|_1 \triangleq \sum_l |\Upsilon_l|$. En nous inspirant des calculs développés en (7.17), une stratégie de mise à jour multiplicative déduite du calcul du gradient $\nabla_\Upsilon \left[d_\beta(\hat{\mathbf{I}}_x | \Psi \Upsilon) + \lambda \|\Upsilon\|_1 \right]$ donne

$$\hat{\Upsilon} \leftarrow \hat{\Upsilon} \odot \frac{\Psi^\top \left((\Psi \Upsilon)^{\odot(\beta-2)} \odot \hat{\mathbf{I}}_x \right)}{\Psi^\top (\Psi \Upsilon)^{\odot(\beta-1)} + \lambda}. \quad (8.11)$$

Nous allons maintenant introduire plusieurs modèles acoustiques pouvant être utilisés pour notre méthode de localisation, notamment pour l'échantillonnage des *fréquences caractéristiques*.

8.1.3 Modèle acoustique

8.1.3.1 Réponses impulsionnelles de salle

Nous ne présentons ici que deux modèles acoustiques utilisés dans les articles [60, 61]. Le premier modèle choisi pour les *vecteurs de pointages* est l'hypothèse forte de connaître la *réponse impulsionnelle* [156, 145] de la propagation de chaque endroit d'une salle vers chacun des K microphones. La Fig. 8.2 représente une telle réponse impulsionnelle. Le premier pic est appelé *son direct* et décrit le chemin direct entre la source et le microphone. Les pics suivants appelés *réflexions précoces* représentent le signal obtenu après réflexion de la source primaire sur les parois de la salle. La dernière partie est quant à elle composée de *réflexions tardives* et englobe les dernières réflexions jusqu'à absorption complète du signal sonore par les parois de la salle ou le milieu physique.

En somme, une *réponse impulsionnelle* renferme exactement l'information de propagation d'une onde sonore dans un milieu acoustique. C'est pourquoi un premier choix proposé de *vecteurs de pointages* $\mathbf{a}_l(f)$ est le suivant :

- Pour chaque point de la grille $l \in \llbracket 1, L \rrbracket$, nous supposons connue la *réponse impulsionnelle* de la position l au micro $k : \tau \mapsto \mathbf{a}_{lk}(\tau)$ pour chaque microphones.
- En effectuant une transformée de Fourier discrète à temps fini pour chaque *réponse impulsionnelle*, nous obtenons une *réponse en fréquence* $\mathbf{a}_{lk}(f)$.
- On pose pour tout l, f , $\mathbf{a}_l(f) = [\mathbf{a}_{l1}(f), \dots, \mathbf{a}_{lK}(f)]^\top$.

Nous sélectionnons alors chaque *fréquence caractéristique* $\mathbf{u}_l(f)$ comme étant égale à $\mathbf{a}_l(f)$. Pour rappel, cela engendre des matrices Ψ_f carrées et symétriques. De plus, le choix des $\mathbf{a}_l(f)$ nous permet d'échantillonner \mathbb{R}^3 à la place de \mathbb{C}^K . L'algorithme 4 résume alors la méthode complète de localisation associée à ce modèle acoustique et au modèle probabiliste en (8.2) et (8.3).

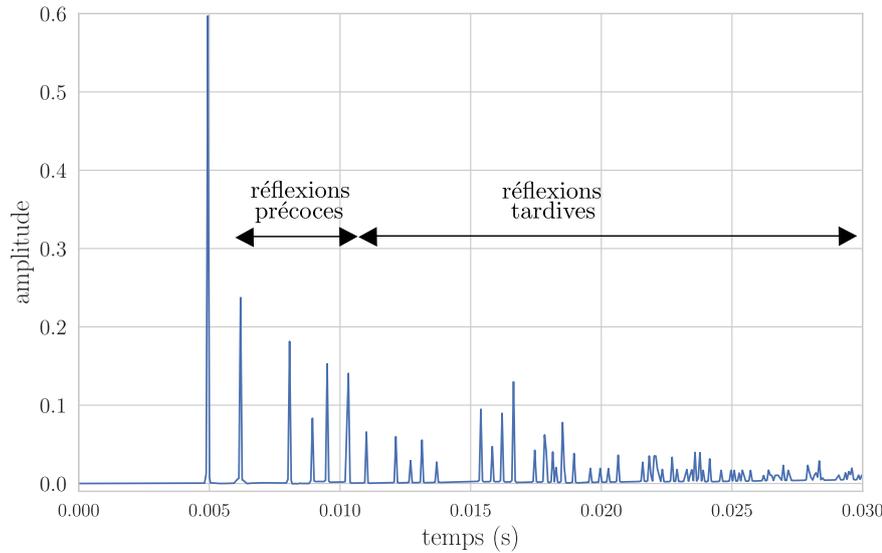


FIGURE 8.2 – Réponse impulsionnelle de salle pour un temps de réverbération (avec diminution du niveau sonore de 60 dB) de 64 ms.

Algorithme 4 Estimation de la MSD $\hat{\mathbf{Y}}$ pour la localisation de sources sonores.

1. Entrées

- Nombre de positions possibles L .
- Vecteurs de pointages $\mathbf{a}_l(f)$ construits à partir du modèle acoustique choisi.
- TFCT \mathbf{x} du mélange, de taille $F \times T \times K$.
- Nombre d'itérations pour les mises à jour multiplicatives.
- Paramètre d'échelle α et divergence ($\beta \in [0, 2]$) à utiliser.

2. Sketching

Pour tout f , calculer $\hat{\mathbf{I}}_f(\mathbf{a}_l(f))$ en utilisant (8.8).

3. Estimation des paramètres

- Calculer Ψ_f relatif à (8.6) avec $\mathbf{u}_l(f) = \mathbf{a}_l(f)$.
 - Calculer $\hat{\mathbf{Y}}$ en itérant l'équation (8.11).
-

8.1.3.2 Modèle de champ proche

En acoustique des salles, de nombreuses modélisations de la propagation d'onde acoustique sont possibles. Dans ce manuscrit, on supposera que les sources, relativement proches des microphones, ne sont pas à l'infini. Nous adoptons alors un modèle de *champ proche* [73]. Il décrit la propagation des ondes sonores dans l'air grâce aux *vecteurs de pointages* suivants :

$$\forall l, f \mathbf{a}_l(f) = \left[\frac{1}{r_{l1}} \exp\left(-2i\pi \frac{fr_{l1}}{\nu c_0} F_s\right), \dots, \frac{1}{r_{lK}} \exp\left(-2i\pi \frac{fr_{lK}}{\nu c_0} F_s\right) \right]^\top, \quad (8.12)$$

où r_{kl} représente la distance en mètres entre la $l^{\text{ème}}$ source potentielle et le $k^{\text{ème}}$ microphone, $c_0 \approx 340$ m/s est la vitesse du son dans l'air à 15°C, ν la fréquence en Hertz et F_s la fréquence d'échantillonnage

du signal. En posant de nouveau pour tout (l, f) , $\mathbf{u}_l(f) = \mathbf{a}_l(f)$, nous obtenons des *fréquences caractéristiques* échantillonnables sur \mathbb{R}^3 à la place de \mathbb{C}^K pour l'estimation de l'*exposant de Lévy*. À noter que généralement, la non-connaissance des *réponses impulsionnelles* entraînera des localisations moins bonnes. En revanche, les hypothèses pour employer un tel modèle acoustique sont moins fortes : seule la position des microphones est requise. La méthode de localisation est alors exactement la même que celle de l'algorithme 4. La seule différence réside dans la sélection des vecteurs de pointages $\mathbf{a}_l(f)$.

D'autres modèles acoustiques auraient pu être également considérés dans cette tâche de localisation. Citons par exemple le modèle de *champ lointain*, relativement similaire au modèle de champ proche décrit ici. La seule différence étant que les sources sont supposées à l'infini. Dans ce cas, les ondes acoustiques se propageant en leur provenance deviennent planes et mettent en relation des équations légèrement différentes de celle proposée en (8.12) (cf. [73] pour plus de détails).

Afin de comparer notre méthode de localisation à d'autres de l'état de l'art, nous présentons dans la section suivante différentes évaluations tirées d'expériences faites en [60, 61].

8.2 Expériences

8.2.1 Méthodes de références

Les trois méthodes de localisation que nous comparons sont les suivantes :

PRD La *puissance de réponse directionnelle* (PRD), appelée également retard-somme [181], est celle décrite dans le chapitre 5 page 41. Elle peut être ici vue comme une restriction de notre approche mais au cas Gaussien ($\alpha = 2$). Nous résumons la PRD par l'équation $\forall l, PRD(l) = \frac{1}{FT} \sum_{f,t} \frac{|\mathbf{a}_l^*(f)\mathbf{x}(f,t)|}{\|\mathbf{a}_l(f)\|_H}$.

CLEAN Cet algorithme glouton [177] se décrit de la manière suivante [124] : à la première itération, il identifie successivement la source ayant l'amplitude la plus forte grâce au calcul de la PRD, et supprime la contribution de cette amplitude. L'algorithme est répété jusqu'à ce que toutes les sources soient identifiées. Il nécessite cependant la connaissance a priori du nombre de sources à estimer.

RELAX est une version améliorée de CLEAN [110] présentée dans [177]. Comparé à CLEAN, il estime à chaque étape la localisation de toutes les sources. Par exemple, si RELAX a localisé une source et en cherche une deuxième, il emploie tout d'abord les coordonnées de localisation et le résidu obtenus lors de la première étape pour localiser la deuxième comme dans CLEAN. Ensuite, il corrige les résidus et localisations obtenus lors de la première étape. Ce dynamisme permet en pratique une meilleure localisation.

8.2.2 Résultats avec un modèle de champ proche

Les premières évaluations considèrent un modèle de *champ proche* pour la localisation de sources sonores. Nous prenons en compte $J = 5$ signaux de paroles monophoniques d'une durée de 10s, provenant de la base de données CMU⁵. La fréquence d'échantillonnage est de 16 kHz et les sources sont positionnées aléatoirement dans une salle simulée de dimensions $5 \times 4 \times 3$ mètres, captées par $K = 4, 5, 10, 20$ et 50 microphones positionnés également de manière aléatoire. Afin d'obtenir un scénario davantage réaliste, nous créons des *réponses impulsionnelles* de salle contenant 4096 échantillons temporels grâce au simulateur présenté dans [72], avec un temps de réverbération associé à une diminution de niveau de 60 dB égal à 0.4 s. En raison du fait que la complexité algorithmique pour l'estimation de la MSD peut être de $\mathcal{O}(FL^2N)$ avec N le nombre d'itérations de (8.11), les sources sont contenues sur un plan à 1.5

5. Carnegie Mellon University dataset : <http://www.festvox.org/cmufaf/>

CLEAN et RELAX sont ici des versions à larges bandes (elles prennent plusieurs fréquences en compte pour la localisation). Les premières versions dans [75, 124] étaient à bandes étroites.

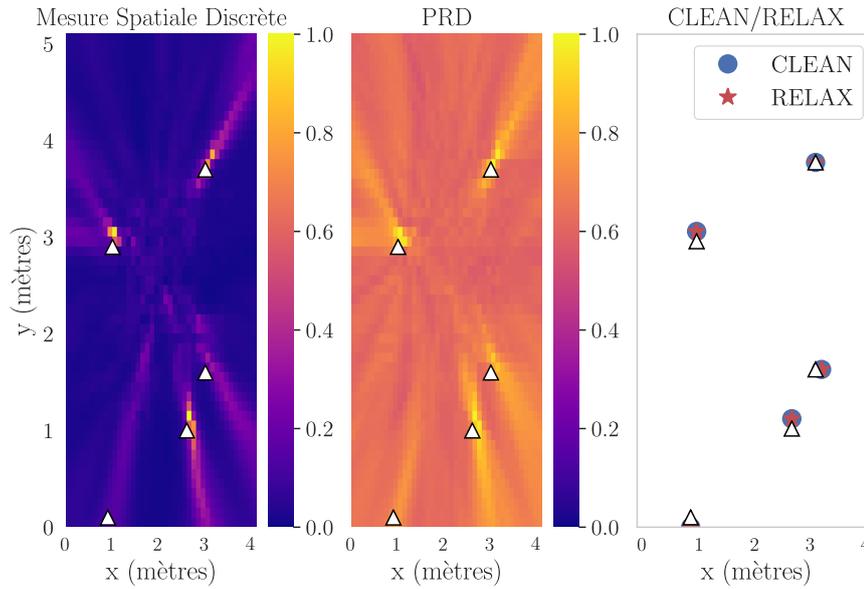


FIGURE 8.3 – Représentation de la MSD \hat{Y} , de la PRD et des algorithmes CLEAN/RELAX. La vraie position des sources sur le plan est annotée par des triangles blancs.

mètres de distance par rapport au sol. Ainsi, $L = 2091$ localisations possibles sont considérées, consistant en un maillage de 10 cm entre chaque position. La grille d'analyse sera décalée de 5 cm par rapport à celle où sont placés les microphones et les sources. Pour diminuer le temps de calcul, nous ne prenons pas toutes les fréquences du signal observé. À la place, nous considérons une plage de fréquences pour l'analyse entre 1000 et 3000 Hz.

Dans cette expérience, les paramètres pris en compte dans l'algorithme 4 sont $\alpha = 1$ pour l'exposant caractéristique, $\lambda = 1$ pour la régularisation parcimonieuse via la norme ℓ_1 et $\beta = 0$ correspondant à la divergence d'Itakura-Saito. Un total de 80 itérations est effectué pour le calcul de la MSD \hat{Y} en (8.11).

La Fig. 8.3 montre un exemple d'estimation de la MSD, de la PRD et des algorithmes CLEAN/RELAX avec 50 microphones. On peut apercevoir que la *carte thermique* de notre méthode est moins bruitée que la méthode PRD. Cela s'explique notamment par la présence de lobes secondaires pour cette dernière.

Pour chaque nombre de microphones, nous exécutons 50 simulations de positionnement aléatoire de sources sur le plan de dimensions 5×4 mètres. Un point important à souligner est que la matrice Ψ ne dépend que de la configuration de la salle. Ainsi, le calcul de Ψ de complexité algorithmique $\mathcal{O}(FL^2K)$ sera effectué une seule fois en fonction du nombre de microphones, et utilisé pour les 50 simulations. Le score employé pour l'évaluation est la corrélation entre la vérité terrain de la *carte thermique* des sources, convoluée avec un noyau gaussien 2D de taille 10 centimètres, et les *cartes thermiques* obtenues pour chaque méthode. À noter que CLEAN et RELAX donnent des estimées ponctuelles et non des *cartes thermiques*. Nous les convoluons cependant afin d'obtenir une évaluation équitable.

Le noyau gaussien permet ici de « lisser » les cartes obtenues.

Nous observons sur la Fig. 8.4 que, pour $K \geq 10$ microphones, CLEAN et RELAX ont les meilleurs scores de corrélation, notamment grâce à la connaissance *a priori* du nombre de sources J . Toutefois, on peut également remarquer une décroissance rapide des résultats quand le nombre de microphones diminue. La MSD quant à elle semble rester plus robuste face à ce retrait d'information et montre également une variance plus faible que le duo CLEAN/RELAX. Pour chaque cas de figure, nous observons que la MSD présente un meilleur score de corrélation contre la PRD.

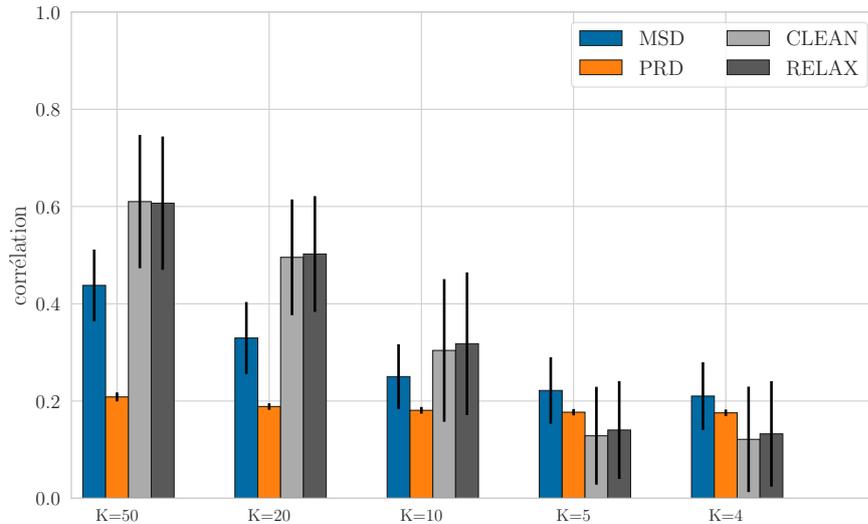


FIGURE 8.4 – Scores de corrélations pour chaque algorithme. Les moustaches noires affichent la variance en fonction du nombre de microphones.

8.2.3 Résultats avec un modèle de réponse acoustique de salle

Nous reprenons exactement les mêmes configurations que l'expérience précédente mais en considérant cette fois-ci $J = 4$ sources. La comparaison avec notre méthode sera uniquement faite avec l'estimation de la PRD.

Les paramètres pour l'estimation de la MSD sont $\alpha = 1.2$, $\beta = 0$ et $\lambda = 0$ (pas de parcimonie via la norme ℓ_1). Cette fois-ci, 70 itérations ont été effectuées pour le calcul de $\hat{\mathbf{Y}}$. À l'aide des $\mathbf{a}_l(f)$ obtenues grâce aux *réponses impulsionnelles* de salle, nous calculons à nouveau la matrice Ψ une seule fois. La Fig. 8.6 montre un exemple de comparaison entre notre méthode et celle de la PRD. Nous pouvons voir davantage l'obtention d'une carte bruitée pour la localisation avec la PRD.

Afin d'évaluer le pouvoir de localisation entre les deux méthodes, nous avons calculé la corrélation entre la vérité terrain convoluée de nouveau avec un noyau gaussien 2D de taille 10 centimètres. Nous avons alors considéré 20 itérations pour un nombre de $K = 3, 4, 10, 20, 30$ microphones. Les résultats sont présentés dans la Fig. 8.5.

À nouveau, le score de corrélation de la MSD est plus important que celui de la PRD. Si l'on compare aux scores établis dans la Fig. 8.4, nous nous apercevons que l'utilisation des véritables filtres de mélange permet de meilleurs résultats pour la MSD. Notons cependant que le nombre de microphones reste important afin d'obtenir des scores de corrélation convenables ($\approx 65\%$ pour $K = 30$ et $\approx 21\%$ pour $K = 3$). Les scores relativement faibles de la méthode PRD sont à nouveau dus à l'existence de lobes secondaires comme le montre la Fig. 8.6.

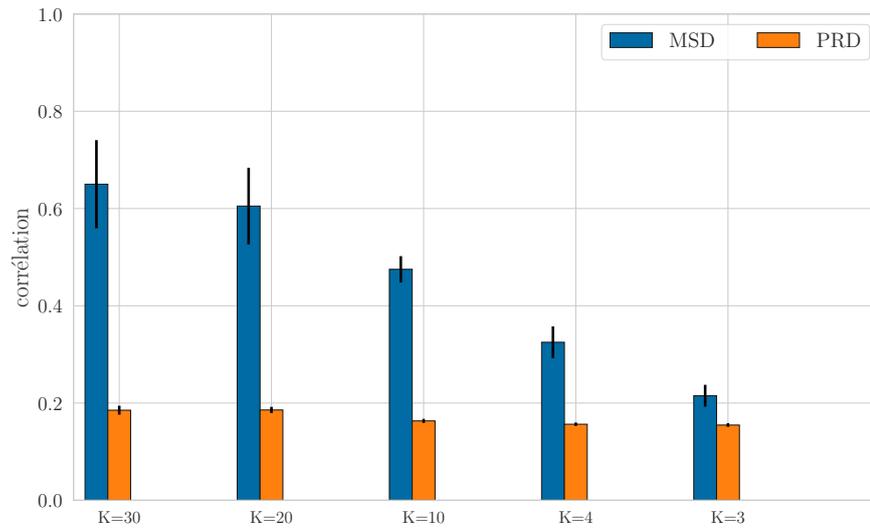


FIGURE 8.5 – Scores de corrélations pour chaque algorithme. Les moustaches noires affichent la variance en fonction du nombre de microphones.

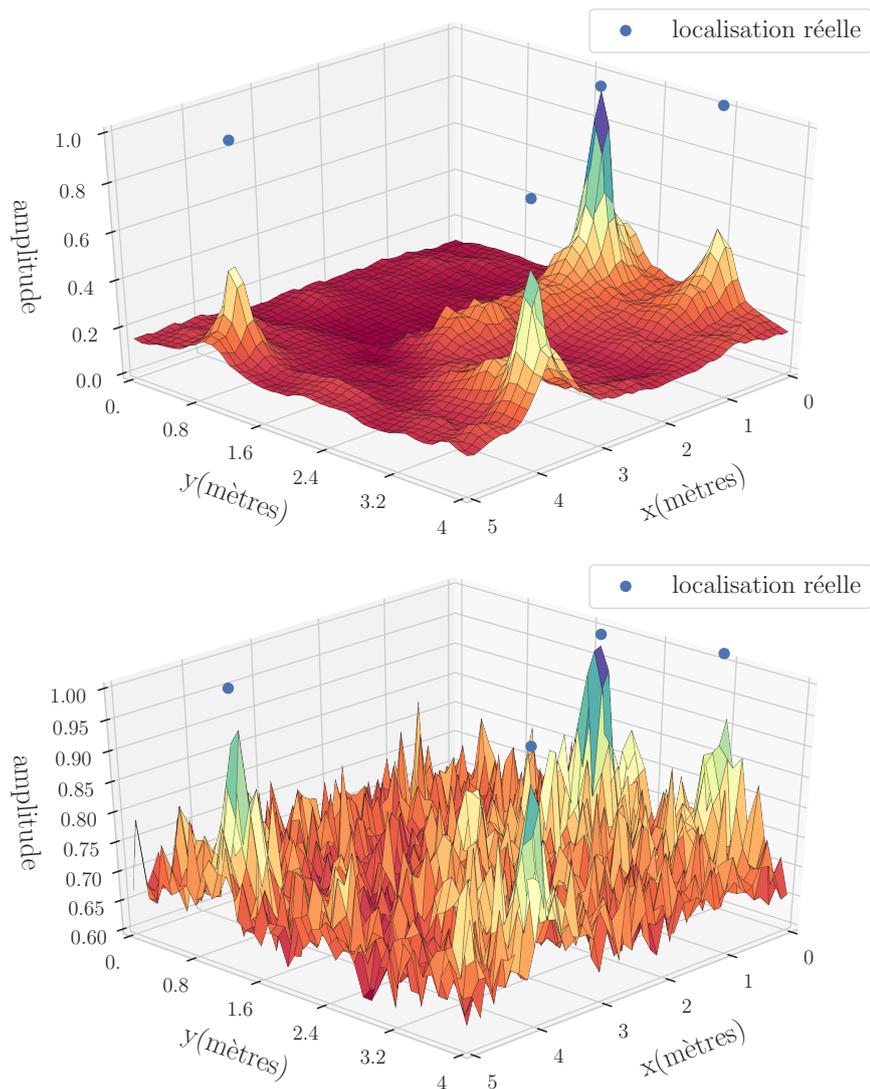


FIGURE 8.6 – MSD (haut) vs PRD (bas) avec $K = 30$ microphones.

Conclusion de la Partie II

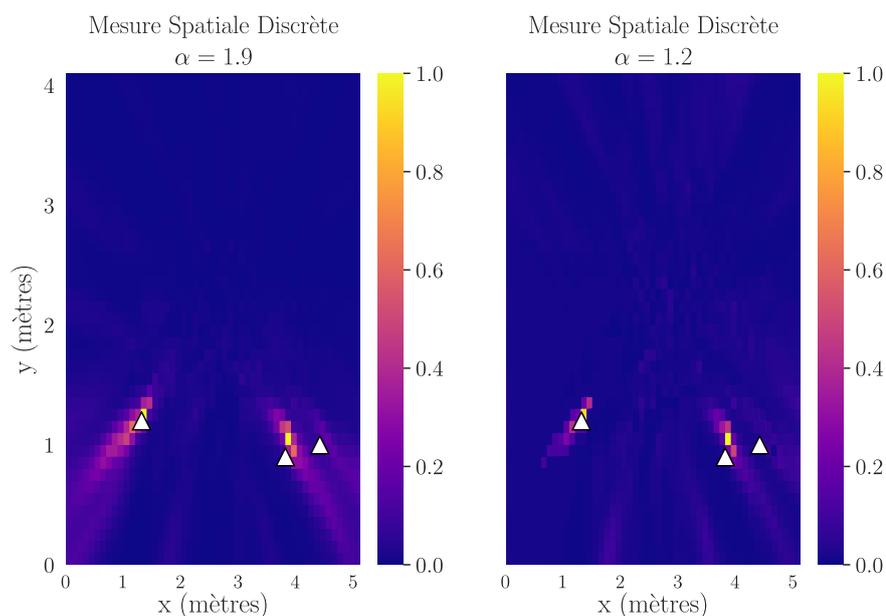


FIGURE 1 – Localisation pour $\alpha = 1.9$ (à gauche) et $\alpha = 1.2$ (à droite) : quand α devient petit, les sources de faible amplitude sont moins convenablement localisées.

Les travaux réalisés dans cette seconde partie ont concerné des mélanges linéaires multivariés d'un nombre fini de sources α -stables complexes isotropes scalaires. Nous avons vu que dans le cas non-gaussien, il existe un vecteur unique que nous avons nommé *mesure spatiale discrète* (MSD), regroupant l'énergie de chaque source du mélange et pour chaque direction d'arrivée possible. Nous avons montré comment exploiter la MSD pour la localisation de sources acoustiques en milieu réverbérant. Il se trouve que la considération d'un opérateur appelé *exposant de Lévy*, résumant les observations pour un nombre déterminé de *fréquences caractéristiques* et indépendant du temps, est suffisante pour estimer la MSD. Nous avons en outre fourni un estimateur de l'*exposant de Lévy* non-négatif garantissant la non-négativité de l'*exposant de Lévy* théorique. Il s'ensuit une méthode proche des règles de mise à jour multiplicatives que l'on retrouve dans le cas de *nonnegative matrix factorization* (NMF), à laquelle nous pouvons ajouter une contrainte de parcimonie. En effet, dans le cas de modélisation acoustique, la MSD Υ résume l'amplitude de chaque position dans une salle, supposant a fortiori que chaque position est une source potentielle active ou non : cela entraîne que le vecteur Υ contient beaucoup de valeurs nulles.

Malgré une méthode affichant de bons résultats comparativement à plusieurs méthodes de l'état de

l'art, il n'en demeure pas moins que l'algorithme peut rester gourmand en calculs. Cela se justifie par la considération d'un grand nombre de sources potentielles. Par ailleurs, le choix d' α est totalement arbitraire et reste un paramètre défini par l'utilisateur. Il existe cependant différentes méthodes pour estimer l'exposant caractéristique α . Par exemple, si l'on s'en réfère à [136, p 61], on peut exploiter l'*exposant de Lévy* et effectuer une régression afin d'estimer α . Le choix d' α est en effet très important comme le montre la Fig. 1. Plus α est petit et plus les pics importants seront pris en compte. À contrario, cela fera omettre les sources potentielles de faible amplitude. Dans un tout autre registre, il existe des cas particuliers de distributions multicanales α -stables que nous introduirons dans la partie IV, appelées *distributions multivariées stables à contour elliptique*. Dans ce cas, l'aspect spatial du mélange sera directement caractérisé par des matrices appelées *matrice de dispersion* et une méthode pour estimer ces dernières sera envisagée.

Troisième partie

**Filtrage de processus α -stables
vectoriels**

Chapitre 9

Représentation spatiale de vecteurs α -stables

Résumé

Ce chapitre sera consacré à l'étude de deux représentations spatiales d'un vecteur α -stable. Nous rappelons dans un premier temps des résultats classiques faisant usage de la *mesure spatiale* qui caractérise entièrement la distribution d'un vecteur α -stable. Notamment le fait que cette mesure, non-négative, est définie sur la sphère, ce qui peut être interprété comme l'ensemble des directions d'arrivées. Même si cette représentation spatiale décrit le comportement d'une distribution multivariée α -stable, elle ne permet pas d'aboutir à une méthode de filtrage d'un mélange de vecteurs α -stables. Nous démontrons que la répartition spatiale peut être décrite plus précisément en décomposant un vecteur α -stable à valeurs dans \mathbb{K} comme une intégrale stochastique le long de l'ensemble des directions d'arrivées. Cette décomposition fait intervenir une mesure aléatoire à valeurs dans \mathbb{K} que nous appelons *spectre spatial* : il précise la contribution de chaque direction dans le vecteur α -stable considéré. Nous aboutissons ainsi à une nouvelle représentation d'un vecteur α -stable : il est la contribution d'un nombre infini de variables scalaires α -stables, chacune en provenance d'une direction donnée et dont le paramètre d'échelle est contrôlé par la *mesure spatiale*.

Dans tout ce chapitre, le cas gaussien ($\alpha = 2$) est exclu afin de manipuler une *mesure spatiale* unique.

9.1 Distribution multivariée α -stable

9.1.1 Premier théorème de représentation

Soit \mathbf{x} un vecteur symétrique $S\alpha S$ complexe ou réel, au sens de la définition 2.1.4 page 14. Si le vecteur aléatoire \mathbf{x} est complexe isotrope, alors

$$\forall \phi \in [0, 2\pi[, e^{i\phi} \mathbf{x} \stackrel{d}{=} \mathbf{x}. \quad (9.1)$$

Excepté pour le cas Cauchy ($\alpha = 1$) et gaussien ($\alpha = 2$), il n'existe pas de forme analytique de la densité de probabilité d'un vecteur $S\alpha S$ isotrope. Cependant, on peut également considérer la fonction caractéristique $\varphi_{\mathbf{x}}$ de \mathbf{x} . Dans la partie II, la fonction caractéristique d'un vecteur $S\alpha S$ de dimension K issu d'un mélange linéaire de variables scalaires $S\alpha S_c$ se décomposait comme l'exponentielle d'une intégrale le long de l'hypersphère S^K contre une mesure réelle (cf. équation (6.6)). Il a été démontré dans [100] qu'une telle représentation spatiale reste valable pour un vecteur symétrique α -stable isotrope :

Définition 9.1.1. Soit \mathbf{x} un vecteur aléatoire dans \mathbb{K}^K ($\mathbb{K} = \mathbb{R}$ ou \mathbb{C}) et $\varphi_{\mathbf{x}} : \mathbf{u} \in \mathbb{K}^K \mapsto \mathbb{E}[\exp(i\Re(\langle \mathbf{u}, \mathbf{x} \rangle))]$ sa fonction caractéristique. Une distribution α -stable symétrique isotrope est entièrement décrite par la représentation unique suivante :

$$\forall \mathbf{u} \in \mathbb{K}^K, \varphi_{\mathbf{x}}(\mathbf{u}) = \exp\left(-\int_{\boldsymbol{\theta} \in \mathcal{S}^K} |\langle \mathbf{u}, \boldsymbol{\theta} \rangle|^\alpha \Gamma_{\mathbf{x}}(d\boldsymbol{\theta})\right) \quad (9.2)$$

où $\Gamma_{\mathbf{x}}$ est une mesure symétrique ⁶ sur la sphère \mathcal{S}^K appelée mesure spatiale. De ce fait, nous notons $\mathbf{x} \sim S\alpha S_c^K(\Gamma_{\mathbf{x}})$ lorsque \mathbf{x} aura une fonction caractéristique égale à (9.2).

Remarque Dans la littérature, $\Gamma_{\mathbf{x}}$ est appelée *mesure spectrale* [153, 136, 138, 94]. Notre choix d'utiliser une autre terminologie se justifie d'une part pour appuyer son importance quant à la description des dépendances entre les différentes composantes de \mathbf{x} , souvent reliées dans un contexte de filtrage aux propriétés spatiales (e.g. les matrices de covariances spatiales dans [39]). D'autre part, il est classique en audio de plutôt réserver le terme de spectre aux transformées de Fourier des formes d'onde, prêtant ainsi facilement à confusion.

Le cas de la *mesure spatiale* présentée dans (6.7) correspondait à un cas particulier où elle est une somme de Dirac. Ici, la mesure $\Gamma_{\mathbf{x}}$ n'a comme condition particulière que d'être symétrique sur \mathcal{S}^K . Cela implique qu'une modélisation plus large de la configuration spatiale d'un vecteur α -stable est envisageable. Présentons quelques propriétés utiles sur les vecteurs $S\alpha S_c^K$.

9.1.2 Propriétés et motivations

Une première propriété classique des vecteurs α -stables isotropes est celle de la stabilité par linéarité.

Proposition 9.1.1. Soit $\mathbf{x} \sim S\alpha S_c^K(\Gamma_{\mathbf{x}})$ et $\mathbf{y} \sim S\alpha S_c^K(\Gamma_{\mathbf{y}})$ deux vecteurs aléatoires indépendants et $a \in \mathbb{R}$ une constante. Alors :

1. $\mathbf{x} + \mathbf{y} \sim S\alpha S_c^K(\Gamma_{\mathbf{x}} + \Gamma_{\mathbf{y}})$;
2. $a\mathbf{x} \sim S\alpha S_c^K(|a|^\alpha \Gamma_{\mathbf{x}})$.

Démonstration. Nous voyons directement en calculant la fonction caractéristique de $\varphi_{\mathbf{x}+\mathbf{y}}$ que :

$$\begin{aligned} \forall \mathbf{u} \in \mathbb{K}^K, \varphi_{\mathbf{x}+\mathbf{y}}(\mathbf{u}) &= \varphi_{\mathbf{x}}(\mathbf{u}) \times \varphi_{\mathbf{y}}(\mathbf{u}) \\ &= \exp\left(-\int_{\boldsymbol{\theta}' \in \mathcal{S}^K} |\langle \mathbf{u}, \boldsymbol{\theta}' \rangle|^\alpha \Gamma_{\mathbf{x}}(d\boldsymbol{\theta}')\right) \times \exp\left(-\int_{\boldsymbol{\theta}' \in \mathcal{S}^K} |\langle \mathbf{u}, \boldsymbol{\theta}' \rangle|^\alpha \Gamma_{\mathbf{y}}(d\boldsymbol{\theta}')\right) \\ &= \exp\left(-\int_{\boldsymbol{\theta}' \in \mathcal{S}^K} |\langle \mathbf{u}, \boldsymbol{\theta}' \rangle|^\alpha (\Gamma_{\mathbf{x}}(d\boldsymbol{\theta}') + \Gamma_{\mathbf{y}}(d\boldsymbol{\theta}'))\right). \end{aligned}$$

De même, nous avons

$$\begin{aligned} \forall \mathbf{u} \in \mathbb{K}^K, \varphi_{a\mathbf{x}}(\mathbf{u}) &= \varphi_{\mathbf{x}}(a\mathbf{u}) \\ &= \exp\left(-\int_{\boldsymbol{\theta}' \in \mathcal{S}^K} |\langle a\mathbf{u}, \boldsymbol{\theta}' \rangle|^\alpha \Gamma_{\mathbf{x}}(d\boldsymbol{\theta}')\right) \\ &= \exp\left(-\int_{\boldsymbol{\theta}' \in \mathcal{S}^K} |\langle \mathbf{u}, \boldsymbol{\theta}' \rangle|^\alpha |a|^\alpha \Gamma_{\mathbf{x}}(d\boldsymbol{\theta}')\right). \end{aligned}$$

□

6. L'aspect symétrique de la mesure $\Gamma_{\mathbf{x}}$ sera décrite de la manière suivante : pour toute fonction f continue et pour $z \in \mathbb{K}$ tel que $|z| = 1$, on a : $\int_{\boldsymbol{\theta} \in \mathcal{S}^K} f(\boldsymbol{\theta}) \Gamma_{\mathbf{x}}(z d\boldsymbol{\theta}) = \int_{\boldsymbol{\theta} \in \mathcal{S}^K} f(\boldsymbol{\theta}) \Gamma_{\mathbf{x}}(d\boldsymbol{\theta})$.

Cette propriété est utile en théorie du filtrage car elle nous permet d'établir une relation entre les *mesures spatiales* de sources et celles d'un mélange dans un contexte de séparation. Cependant, affirmer par analogie avec le cas scalaire présenté dans [118] qu'un filtrage α -Wiener multicanal serait faisable par un rapport de *mesures spatiales* n'est pas possible mathématiquement. Il existe donc deux solutions possibles : la première idée serait de contraindre davantage les *mesures spatiales*, afin de les décrire comme des fonctions mesurables (cf. Théorème de Radon-Nikodym-Lebesgue et densité de mesure [151]). La seconde serait d'esquisser une nouvelle représentation spatiale d'un vecteur $S\alpha S_c^K$.

Notre choix s'est orienté vers cette deuxième approche. Rappelons divers résultats déjà vus dans la section 2.3 page 19 sur la façon de représenter un processus α -stable via une *mesure aléatoire* M . La représentation que nous obtenons permet d'écrire un vecteur $\mathbf{x} \sim S\alpha S_c^K(\Gamma_{\mathbf{x}})$ de la manière suivante :

$$\mathbf{x} \stackrel{d}{=} \int_{\theta \in S^K} \theta M(d\theta) \quad (9.3)$$

où M est une *mesure aléatoire* avec comme *mesure de contrôle* $\Gamma_{\mathbf{x}}$. Cette représentation (9.3) signifie que le vecteur \mathbf{x} peut être vu comme la somme d'une infinité de contributions $M(d\theta) \in \mathbb{K} \sim S\alpha S_c(\Gamma_{\mathbf{x}}(d\theta))$, pondérant la direction θ . Nous verrons plus tard que cette représentation est appropriée pour mettre au point des méthodes de filtrage.

9.2 Théorème de représentation spatiale

9.2.1 Spectre spatial

Nous avons explicité dans la section 3.1.3 page 25 différents travaux de Liutkus & Badeau [118] portant sur la théorie du filtrage associée à l'utilisation de processus α -harmonisables. Nous rappelons également qu'un processus α -harmonisable a une représentation sous forme d'intégrale stochastique, contrôlée par une *mesure aléatoire isotrope* M (cf. définition 2.3.3 page 21). Malheureusement, notre souhait d'étendre la représentation de processus α -stables harmonisable scalaire au cas d'un champ vectoriel α -harmonisable en considérant (9.3) et le couple $(M, \Gamma_{\mathbf{x}})$ n'est pas réalisable. Même si les processus obtenus ne sont pas α -harmonisables, nous pouvons néanmoins construire étape par étape la représentation souhaitée. Commençons par définir la mesure aléatoire pour obtenir (9.3) :

Définition 9.2.1 (Spectre spatial). *Nous appelons **spectre spatial** de la distribution $S\alpha S_c^K(\Gamma_{\mathbf{x}})$, associé à la mesure spatiale $\Gamma_{\mathbf{x}}$, la mesure aléatoire $S\alpha S_c$ de mesure de contrôle $\Gamma_{\mathbf{x}}$, notée χ , satisfaisant les conditions particulières suivantes :*

1. *Pour tout ensemble Borélien $A \subset \mathcal{B}(S^K)$, la variable aléatoire scalaire $\chi(A) \in \mathbb{K}$ est distribuée comme $\chi(A) \sim S\alpha S_c(\Gamma_{\mathbf{x}}(A))$;*
2. *$\chi(A)$ et $\chi(B)$ sont indépendantes quand $A \cap B = \emptyset$ pour tout ensemble Borélien A et $B \subset \mathcal{B}(S^K)$.*

Avec une telle mesure aléatoire, nous allons démontrer qu'il est possible d'obtenir la représentation en (9.3).

9.2.2 Énoncé et démonstration

Le théorème ci-après énonce une nouvelle représentation spatiale d'un vecteur $S\alpha S_c^K$:

Théorème 9.2.1 (Représentation spatiale). *Soit $\mathbf{x} \sim S\alpha S_c^K(\Gamma_{\mathbf{x}})$, associé à la mesure spatiale $\Gamma_{\mathbf{x}}$. Alors \mathbf{x} admet la représentation spatiale suivante :*

$$\mathbf{x} \stackrel{d}{=} \int_{\theta \in S^K} \theta \chi(d\theta), \quad (9.4)$$

où χ est le spectre spatial, associé à la mesure spatiale $\Gamma_{\mathbf{x}}$ et « $\chi(d\theta) \sim S\alpha S_c(\Gamma_{\mathbf{x}}(d\theta))$ ».

Nous confondons les termes « mesure de contrôle » et « mesure spatiale » pour cette représentation particulière.

Démonstration. Nous traitons les cas réel et complexe séparément. Pour commencer, dans le cas où $\mathbb{K} = \mathbb{C}$, nous pouvons obtenir directement le résultat en appliquant le théorème 6.3.4 de [153, p. 284] avec les paramètres suivants

$$E = \mathcal{S}^K, x = \theta, M(dx) = \mathcal{X}(d\theta), m(dx) = \Gamma_x(d\theta), d = K, \\ j = k, z_j = u_k, T = [1 \dots K], t_j = k, f_{t_j}(x) = x_{t_j} = \theta_k$$

et ainsi, obtenir l'égalité du théorème 6.3.4 :

$$\forall \mathbf{u} \in \mathbb{C}^K, \mathbb{E} \left[\exp \left(i \Re \left\langle \mathbf{u}, \int_{\theta \in \mathcal{S}^K} \theta \mathcal{X}(d\theta) \right\rangle \right) \right] = \exp \left(- \int_{\theta \in \mathcal{S}^K} |\langle \mathbf{u}, \theta \rangle|^\alpha \Gamma_x(d\theta) \right). \quad (9.5)$$

Le terme de droite représente la fonction caractéristique connue d'un vecteur $\mathbf{x} \sim S\alpha S_c^K(\Gamma_x)$, issue de la représentation spatiale (9.2). Celui de gauche correspond à la fonction caractéristique du vecteur $\int_{\theta \in \mathcal{S}^K} \theta \mathcal{X}(d\theta)$. Ainsi, on en conclut que \mathbf{x} et $\int_{\theta \in \mathcal{S}^K} \theta \mathcal{X}(d\theta)$ ont la même distribution. Le cas où $\mathbb{K} = \mathbb{R}$ peut être démontré en appliquant le théorème de représentation dans \mathbb{R}^K [153, p. 131, thm 3.5.6]. \square

Remarque Si $\mathbb{K} = \mathbb{C}$, nous définissons implicitement, comme mentionné dans la sous-section 2.3.2 page 21, une *mesure de contrôle circulaire* $k = \Gamma_x \times \gamma$. Ici, γ est la mesure uniforme sur le cercle unité S_c^1 , telle que $\gamma(S_c^1) = \frac{2\pi}{\int_0^\pi |\cos(\phi)|^\alpha d\phi}$. Alors, pour toute fonction ψ telle que $\int_{\theta \in \mathcal{S}^K} |\psi(\theta)|^\alpha \Gamma_x(d\theta) < +\infty$, nous pouvons démontrer que l'intégrale stochastique $\int_{\theta \in \mathcal{S}^K} \psi(\theta) \mathcal{X}(d\theta)$ définit une variable aléatoire $S\alpha S_c \left(\int_{\theta \in \mathcal{S}^K} |\psi(\theta)|^\alpha \Gamma_x(d\theta) \right)$ (cf. [153, thm 6.3.1 p. 281]). Dans le cas où $\mathbb{K} = \mathbb{R}$, nous avons également (cf. [153, prop 3.2.2 p. 117])

$$\int_{\theta \in \mathcal{S}^K} \psi(\theta) \mathcal{X}(d\theta) \sim S\alpha S \left(\int_{\theta \in \mathcal{S}^K} |\psi(\theta)|^\alpha \Gamma_x(d\theta) \right). \quad (9.6)$$

En plus de fournir une représentation spatiale adéquate pour développer une théorie du filtrage, le théorème 9.2.1 nous donne une méthode élégante pour échantillonner un vecteur $S\alpha S_c^K$ à partir de la connaissance de sa *mesure spatiale* Γ_x . En effet, la méthode peut se résumer en deux étapes : générer le spectre spatial \mathcal{X} comme étant une variable aléatoire $\mathcal{X}_p \sim S\alpha S_c(\Gamma_x(\Theta_p))$ pour tout borélien Θ_p de l'hypersphère \mathcal{S}_K , et utiliser (9.4) pour construire \mathbf{x} . Nous avons résumé cette approche dans l'algorithme 5

Algorithme 5 Échantillonnage de vecteurs aléatoires $S\alpha S_c^K(\Gamma_x)$ via leur représentation spatiale.

1. **Entrées**

- Nombre N de réalisations désirées
- Partition $\Theta = \{\Theta_1, \dots, \Theta_P\}$ régulière de \mathcal{S}^K , chacune de mesure de Lebesgue Δ_Θ
- Exposant caractéristique $\alpha \in (0, 2)$
- Mesure spatiale Γ_x

2. **Création du spectre spatial discret**

$$\forall n, p, \mathcal{X}_{np} \sim S\alpha S_c(\Gamma_x(\Theta_p))$$

3. **Synthèse** $\forall n, \mathbf{x}_n \leftarrow \sum_p \theta_p \mathcal{X}_{np}$

En pratique, il est à noter que l'intégrale le long de l'hypersphère doit être remplacée par une somme finie. Par exemple, ce calcul peut être effectué en construisant une partition régulière $\Theta \triangleq \{\Theta_p\}_{p \in \llbracket 1, P \rrbracket}$ de \mathcal{S}^K , au sens de la mesure de Lebesgue. Ainsi, l'intégrale dans (9.4) serait substituée par une somme sur

les Θ_p . Considérons une fonction f mesurable sur \mathcal{S}^K et une mesure M définie sur ce même ensemble et dominée par la mesure de Lebesgue. Comme M est dominée par la mesure de Lebesgue, il existe une fonction mesurable m sur \mathcal{S}^K telle que $M(d\theta) = m(\theta)d\theta$. Pour un nombre entier P suffisamment grand, nous appliquons l'approximation suivante :

$$\int_{\theta \in \mathcal{S}^K} \theta f(\theta) M(d\theta) \approx \sum_{p=1}^P \theta_p f(\theta_p) M(\Theta_p) \tag{9.7}$$

où $M(\Theta_p)$ peut être approximée sous la forme

$$M(\Theta_p) \approx m(\theta_p) \Delta_{\Theta} \tag{9.8}$$

avec Δ_{Θ} la mesure de Lebesgue d'un ensemble quelconque Θ_p et $\theta_p \in \Theta_p$.

9.2.3 Interprétation du théorème de représentation

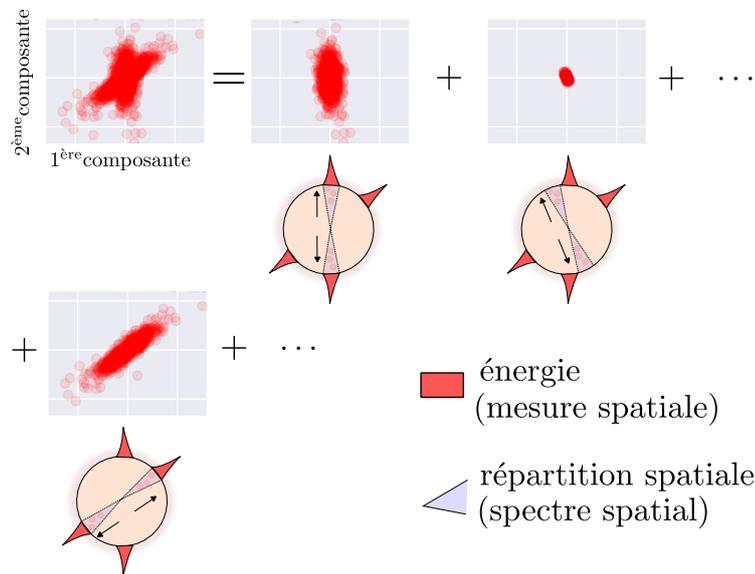


FIGURE 9.1 – Représentation spatiale d'un tirage de vecteurs $S\alpha S_c^2$ dans le cas réel : dans chaque direction, le *spectre spatial* est une collection de variables $S\alpha S_c$ indépendantes, avec une énergie contrôlée par la *mesure spatiale*.

Le théorème (9.2.1) stipule qu'un vecteur $\mathbf{x} \sim S\alpha S_c^K(\Gamma_x)$ peut se décomposer comme une somme infinie de variables scalaires $S\alpha S_c$ venant de toutes les directions $\theta \in \mathcal{S}^K$. Toutes ces variables prises ensemble portent le nom de *spectre spatial*. La *mesure spatiale* $\Gamma_x(d\theta)$ est quant à elle interprétable comme l'énergie des contributions en provenance d'une direction θ .

Le point fort de cette représentation est le niveau de sophistication spatiale que peut avoir un vecteur $S\alpha S_c^K$. Sous la seule contrainte que Γ_x soit symétrique, cela peut conduire à des courbes de niveaux de \mathbf{x} non-elliptiques. En d'autres termes, un vecteur $\mathbf{x} \sim S\alpha S_c^K$ peut admettre une représentation spatiale plus complexe que celle d'une distribution multivariée gaussienne isotrope (qui admet une courbe de niveaux elliptique). En somme, Γ_x nous permet de contrôler précisément la répartition de \mathbf{x} dans l'espace.

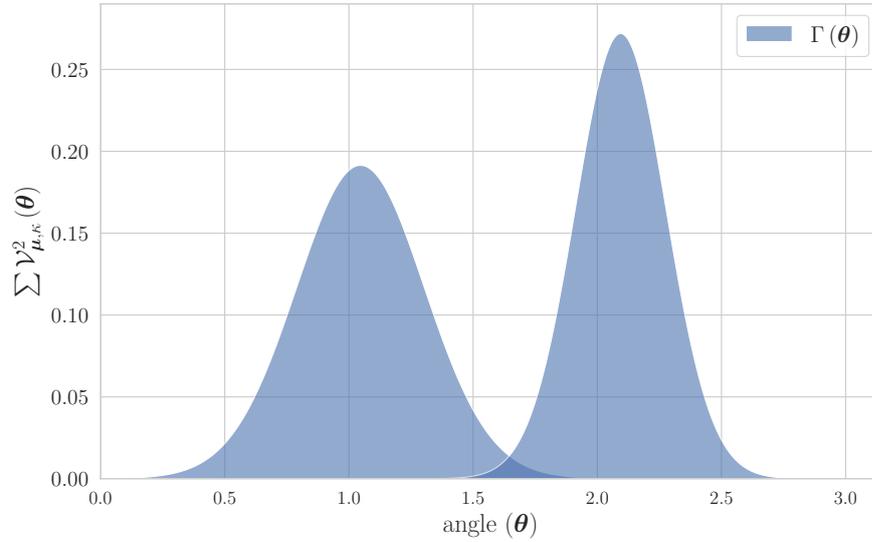


FIGURE 9.2 – Mélange de deux densités de Von-Mises Fisher, de paramètres $\{\mu = \frac{\pi}{3}, \kappa = 15\}$ et $\{\mu = \frac{2\pi}{3}, \kappa = 30\}$.

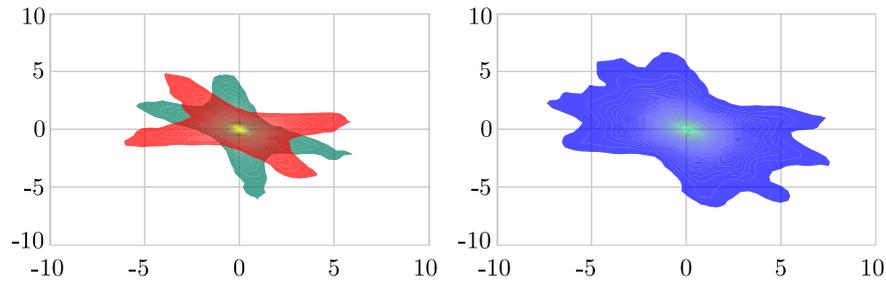


FIGURE 9.3 – Représentation spatiale d'un tirage de vecteurs $S\alpha S_c^2$ dans le cas réel. Sur la figure de gauche, les deux tirages de $y_1 \sim S\alpha S^2(\Gamma_1)$ (en vert) et $y_2 \sim S\alpha S^2(\Gamma_2)$ (en rouge) admettent des densités spatiales Γ_1 et Γ_2 égales à la somme de deux distributions de Von-Mises Fisher, dont les maxima sont localisés en $\{\frac{5\pi}{8}, \frac{5\pi}{8} + \frac{\pi}{3}\}$ et $\{\frac{5\pi}{8} + \frac{\pi}{6}, \frac{5\pi}{8} + \frac{\pi}{2}\}$. Sur la figure de droite, nous voyons que la représentation des lignes de niveaux du mélange $x = y_1 + y_2$ montre la propriété d'additivité des densités spatiales.

Dans la Fig. 9.3, nous avons par exemple considéré une mesure spatiale Γ_x définie comme le mélange de 4 distributions de *Von-Mises Fisher* [53]. Une telle fonction, noté $\mathcal{V}_{\mu, \kappa}^K$, est définie comme suit sur l'hypersphère réelle $\mathcal{S}_{\mathbb{R}}^K$:

$$\forall \theta \in \mathcal{S}_{\mathbb{R}}^K, \mathcal{V}_{\mu, \kappa}^K(\theta) = \frac{\kappa^{K/2-1}}{(2\pi)^{K/2} \mathfrak{B}_{K/2-1}(\kappa)} \exp(\kappa \mu^\top \theta) d\theta \quad (9.9)$$

où $\mu \in \mathbb{R}^K$ est un vecteur de norme 1 appelé *paramètre de localisation*, $\kappa > 0$ est le *paramètre de concentration* (voir la Fig. 9.2 pour un mélange de distributions de Von-Mises Fisher) et $\mathfrak{B}_{K/2-1}$ est la fonction de Bessel modifiée du premier type à l'ordre $K/2 - 1$ [1]. Nous avons également représenté une

interprétation de la décomposition du théorème de représentation spatiale sur la Fig. 9.1 : nous pouvons se représenter le tirage d'un vecteur $S\alpha S_c^K$ comme un « canon », alias *spectre spatial*, projetant un nombre infini de réalisations $S\alpha S_c$ dans toutes les directions avec une énergie, alias *mesure spatiale*, contrôlée par Γ_x .

Dans le chapitre suivant, nous abordons une première stratégie exploitant le théorème 9.2.1 pour la séparation d'un mélange de vecteurs $S\alpha S_c^K$.

Chapitre 10

Filtrage de mélange α -stables vectoriels : cas général

Résumé

En exploitant la nouvelle représentation spatiale présentée au théorème 9.2.1 page 75, nous proposons une nouvelle méthode de filtrage permettant d'estimer chaque terme d'une somme de vecteurs $S\alpha S_c^K$. En supposant connues les *mesures spatiales* de chaque composante, la première étape consistera à estimer le *spectre spatial* du mélange. Dans un second temps, la connaissance des *mesures spatiales* et du *spectre spatial* nous permettra d'établir une méthode d'estimation de chaque composante α -stable. Cependant, une partie des calculs nécessaires à l'estimation du *spectre spatial* demeure trop complexe pour être implémentée. Afin d'obtenir une estimation plus simple, nous proposons une modification de l'estimateur initial.

Dans tout ce chapitre, on supposera que $\alpha \in]1, 2[$.

10.1 Modèle de mélange

Soit J vecteurs aléatoires $y_j \sim S\alpha S_c^K(\Gamma_j)$ et le modèle de mélange suivant :

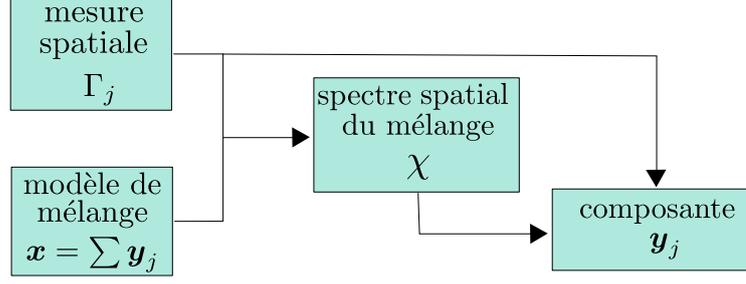
$$\mathbf{x} = \sum_{j=1}^J y_j. \quad (10.1)$$

Le but de notre méthode de filtrage est, à partir de la connaissance du vecteur \mathbf{x} et des mesures spatiales Γ_j , d'estimer les vecteurs y_j . La technique de séparation employée est résumée via la Fig. 10.1. D'après la propriété 9.1.1 page 74, le vecteur aléatoire \mathbf{x} est un vecteur $S\alpha S_c^K$ paramétré par

$$\begin{cases} \mathbf{x} \sim S\alpha S_c^K(\Gamma_{\mathbf{x}}) \\ \Gamma_{\mathbf{x}} \triangleq \sum_{j=1}^J \Gamma_j \end{cases}. \quad (10.2)$$

Si nous appliquons le théorème 9.2.1 page 75 de représentation spatiale pour chaque composante y_j , nous obtenons

$$y_j \stackrel{d}{=} \int_{\theta \in S^K} \theta \mathcal{Y}_j(d\theta) \quad (10.3)$$

FIGURE 10.1 – Séparation multidimensionnelle α -stable via l'estimation de spectres spatiaux.

où \mathcal{Y}_j est le *spectre spatial* associé à la *mesure spatiale* Γ_j du vecteur \mathbf{y}_j . D'après la définition 9.2.1 page 75, il est clair que $\chi \triangleq \sum_j \mathcal{Y}_j$ est également le spectre spatial de \mathbf{x} associée à la mesure spatiale $\Gamma_{\mathbf{x}}$. En effet,

$$\begin{aligned}
 \mathbf{x} &= \sum_{j=1}^J \mathbf{y}_j \\
 &\stackrel{d}{=} \sum_{j=1}^J \int_{\boldsymbol{\theta} \in \mathcal{S}^K} \boldsymbol{\theta} \mathcal{Y}_j(d\boldsymbol{\theta}) \\
 &= \int_{\boldsymbol{\theta} \in \mathcal{S}^K} \boldsymbol{\theta} \sum_{j=1}^J \mathcal{Y}_j(d\boldsymbol{\theta}) = \int_{\boldsymbol{\theta} \in \mathcal{S}^K} \boldsymbol{\theta} \chi(d\boldsymbol{\theta}).
 \end{aligned}$$

Informellement, cela signifie que \mathbf{x} est aussi une somme infinie de contributions $\chi(d\boldsymbol{\theta})$, venant de toutes les directions $\boldsymbol{\theta} \in \mathcal{S}^K$ de la sphère. Pour un $\boldsymbol{\theta} \in \mathcal{S}^K$ donné, chacune de ces contributions est à son tour une somme des contributions $\mathcal{Y}_j(d\boldsymbol{\theta})$ de chaque composante venant de cette direction particulière. La stratégie que nous adoptons pour la séparation sera alors la suivante :

- à partir du modèle de mélange $\mathbf{x} = \sum \mathbf{y}_j$ et de la connaissance des mesures spatiales Γ_j , nous estimons le spectre spatial χ ;
- grâce à χ et aux Γ_j , nous reconstruisons chaque \mathbf{y}_j .

Les prochaines sections sont consacrées au développement de la méthode de filtrage proposée. Nous verrons notamment comment considérer le spectre spatial χ différemment, dans le but d'obtenir une forme implémentable pour son estimation.

10.2 Estimation du spectre spatial

10.2.1 Critère d'estimation

Supposons que l'observation $\mathbf{x} \sim S\alpha\mathcal{S}_C^K(\Gamma_{\mathbf{x}})$ et sa mesure spatiale $\Gamma_{\mathbf{x}}$ sont connus. La première étape est d'estimer le spectre spatial χ . Pour ce faire, nous choisissons comme estimateur l'espérance *a posteriori* $\hat{\chi}(d\boldsymbol{\theta})$ de $\chi(d\boldsymbol{\theta})$ sachant \mathbf{x} . C'est à dire, nous cherchons $\hat{\chi}$ tel que pour toute fonction continue ψ sur \mathcal{S}^K satisfaisant

$$\int_{\boldsymbol{\theta} \in \mathcal{S}^K} |\psi(\boldsymbol{\theta})|^\alpha \Gamma_{\mathbf{x}}(d\boldsymbol{\theta}) < +\infty, \quad (10.4)$$

nous ayons

$$\mathbb{E} \left[\int_{\theta \in \mathcal{S}^K} \psi(\theta) \mathcal{X}(d\theta) \mid \mathbf{x} \right] = \int_{\theta \in \mathcal{S}^K} \psi(\theta) \hat{\chi}(d\theta). \quad (10.5)$$

En résumé, l'estimateur $\hat{\chi}(d\theta)$ conduit à la reconstruction parfaite $\int_{\theta \in \mathcal{S}^K} \psi(\theta) \hat{\chi}(d\theta)$ pour toute fonction ψ satisfaisant (10.4). En particulier, pour $\psi(\theta) = \theta$, nous avons que (10.5) est identiquement distribué comme le vecteur \mathbf{x} . Notons que la condition (10.4) est nécessaire pour que chaque intégrale contre la *mesure spatiale* soit bien définie (cf. la construction de distributions $S\alpha S$ via une représentation en intégrale stochastique [153, p. 275-281, section 6.2]). Dans l'état, un tel estimateur n'est pas implémentable. Néanmoins, comme $\mathbf{x} \sim S\alpha S_c^K(\Gamma_x)$, $\hat{\chi}(d\theta)$ peut s'exprimer comme le produit d'une fonction contre la *mesure spatiale* Γ_x :

Proposition 10.2.1. *Supposons que l'équation (1.7) se généralise pour la fonction caractéristique de $\int_{\theta \in \mathcal{S}^K} \psi(\theta) \hat{\chi}(d\theta) \mid \mathbf{x}$ (au sens des dérivées de Wirtinger [79]) :*

$$\mathbb{E} \left[\int_{\theta \in \mathcal{S}^K} \psi(\theta) \mathcal{X}(d\theta) \mid \mathbf{x} \right] = -i \frac{dc}{dv^\star}(v=0), \quad (10.6)$$

où

$$\forall v \in \mathbb{K}, c(v) = \varphi_{\int_{\theta \in \mathcal{S}^K} \psi(\theta) \hat{\chi}(d\theta) \mid \mathbf{x}}(v). \quad (10.7)$$

Alors, ajouté aux hypothèses données sur le critère d'estimation, nous pouvons réécrire $\hat{\chi}(d\theta)$ comme suit pour tout $\alpha > 1$:

$$\hat{\chi}(d\theta) = g_\chi(\mathbf{x}, \theta) \Gamma_x(d\theta) \text{ p.s.} \quad (10.8)$$

avec :

$$g_\chi(\mathbf{x}, \theta) \triangleq \frac{N(\mathbf{x}, \theta)}{D(\mathbf{x})} \triangleq \frac{\alpha i \int_{\mathbf{u} \in \mathbb{K}^K} \langle \theta, \mathbf{u} \rangle^{\langle \alpha-1 \rangle} \varphi_x(\mathbf{u}) e^{-i\Re(\langle \mathbf{u}, \mathbf{x} \rangle)} d\mathbf{u}}{\int_{\mathbf{u} \in \mathbb{K}^K} \varphi_x(\mathbf{u}) e^{-i\Re(\langle \mathbf{u}, \mathbf{x} \rangle)} d\mathbf{u}} \quad (10.9)$$

où pour tout $z \in \mathbb{K}$,

$$z^{\langle \alpha \rangle} \triangleq z |z|^{\alpha-1} \quad (10.10)$$

est la fonction puissance signée [153] et $\mathbf{u} \mapsto \varphi_x(\mathbf{u})$ la fonction caractéristique du vecteur \mathbf{x} .

Démonstration. Démontrons que $-i \frac{dc}{dv^\star}(v=0)$ dans (10.6) est égal, presque sûrement, au terme à droite de l'égalité (10.8). Calculons en premier lieu la fonction caractéristique jointe de $\int_{\theta \in \mathcal{S}^K} \psi(\theta) \mathcal{X}(d\theta)$ et \mathbf{x} . Nous avons pour tout $v \in \mathbb{K}$ et $\mathbf{u} \in \mathbb{K}^K$

$$\begin{aligned} \varphi_{\left(\int_{\theta \in \mathcal{S}^K} \psi(\theta) \mathcal{X}(d\theta), \mathbf{x}\right)}(v, \mathbf{u}) &\triangleq \mathbb{E} \left[e^{i\Re\left(v^* \int_{\theta \in \mathcal{S}^K} \psi(\theta) \mathcal{X}(d\theta) + \langle \mathbf{u}, \mathbf{x} \rangle\right)} \right] \\ &= \mathbb{E} \left[e^{i\Re\left(\int_{\theta \in \mathcal{S}^K} (v^* \psi(\theta) + \langle \mathbf{u}, \theta \rangle) \mathcal{X}(d\theta)\right)} \right] \\ &= \varphi_{\int_{\theta \in \mathcal{S}^K} (v^* \psi(\theta) + \langle \mathbf{u}, \theta \rangle) \mathcal{X}(d\theta)}(1) \\ &= e^{-\int_{\theta \in \mathcal{S}^K} |v^* \psi(\theta) + \langle \mathbf{u}, \theta \rangle|^\alpha \Gamma_x(d\theta)}, \end{aligned}$$

où la dernière égalité se justifie en utilisant l'équation (9.6) :

$$\int_{\theta \in \mathcal{S}^K} (v^* \psi(\theta) + \langle \mathbf{u}, \theta \rangle) \mathcal{X}(d\theta) \sim S\alpha S \left(\int_{\theta \in \mathcal{S}^K} |v^* \psi(\theta) + \langle \mathbf{u}, \theta \rangle|^\alpha \Gamma_x(d\theta) \right). \quad (10.11)$$

Ainsi, la fonction c admet la forme suivante pour tout $v \in \mathbb{K}$:

$$\begin{aligned} c(v) &= \frac{\int_{\mathbf{u} \in \mathbb{K}^K} \varphi \int_{\boldsymbol{\theta} \in \mathcal{S}^K} \psi(\boldsymbol{\theta}) \mathcal{X}(d\boldsymbol{\theta})_{\mathbf{x}}(v, \mathbf{u}) e^{-i\Re(\langle \mathbf{u}, \mathbf{x} \rangle)} d\mathbf{u}}{\int_{\mathbf{u} \in \mathbb{K}^K} \varphi_{\mathbf{x}}(\mathbf{u}) e^{-i\Re(\langle \mathbf{u}, \mathbf{x} \rangle)} d\mathbf{u}} \\ &= \frac{\int_{\mathbf{u} \in \mathbb{K}^K} e^{-\int_{\boldsymbol{\theta} \in \mathcal{S}^K} |v^* \psi(\boldsymbol{\theta}) + \langle \mathbf{u}, \boldsymbol{\theta} \rangle|^\alpha \Gamma_{\mathbf{x}}(d\boldsymbol{\theta})} e^{-i\Re(\langle \mathbf{u}, \mathbf{x} \rangle)} d\mathbf{u}}{\int_{\mathbf{u} \in \mathbb{K}^K} \varphi_{\mathbf{x}}(\mathbf{u}) e^{-i\Re(\langle \mathbf{u}, \mathbf{x} \rangle)} d\mathbf{u}}. \end{aligned} \quad (10.12)$$

L'existence de $\frac{dc}{dv^*}(v=0)$ se justifie par le fait que $b : v \mapsto \int_{\mathbf{u} \in \mathbb{K}^K} \varphi \int_{\boldsymbol{\theta} \in \mathcal{S}^K} \psi(\boldsymbol{\theta}) \mathcal{X}(d\boldsymbol{\theta})_{\mathbf{x}}(v, \mathbf{u}) e^{-i\Re(\langle \mathbf{u}, \mathbf{x} \rangle)} d\mathbf{u}$ est la transformée de Fourier d'une fonction caractéristique, admettant une dérivée à l'ordre 1 car $\alpha > 1$. En effet, si un vecteur aléatoire admet un moment d'ordre 1, alors sa fonction caractéristique est différentiable et sa dérivée est continue en 0. Par conséquent, nous pouvons employer le théorème de dérivation sous le signe somme (la domination étant induite par le fait qu'une fonction caractéristique est bornée par 1) pour prouver l'existence de la dérivée. Enfin, en combinant les équations (10.6) et (10.12), nous obtenons :

$$\begin{aligned} \mathbb{E} \left[\int_{\boldsymbol{\theta} \in \mathcal{S}^K} \psi(\boldsymbol{\theta}) \mathcal{X}(d\boldsymbol{\theta}) \mid \mathbf{x} \right] &= \frac{\alpha i \int_{\mathbf{u} \in \mathbb{K}^K} \left(\int_{\boldsymbol{\theta} \in \mathcal{S}^K} \psi(\boldsymbol{\theta}) \langle \boldsymbol{\theta}, \mathbf{u} \rangle^{(\alpha-1)} \Gamma_{\mathbf{x}}(d\boldsymbol{\theta}) \right) \varphi_{\mathbf{x}}(\mathbf{u}) e^{-i\Re(\langle \mathbf{u}, \mathbf{x} \rangle)} d\mathbf{u}}{\int_{\mathbf{u} \in \mathbb{K}^K} \varphi_{\mathbf{x}}(\mathbf{u}) e^{-i\Re(\langle \mathbf{u}, \mathbf{x} \rangle)} d\mathbf{u}} \\ &= \int_{\boldsymbol{\theta} \in \mathcal{S}^K} \psi(\boldsymbol{\theta}) \left[\frac{\alpha i \int_{\mathbf{u} \in \mathbb{K}^K} \langle \boldsymbol{\theta}, \mathbf{u} \rangle^{(\alpha-1)} \varphi_{\mathbf{x}}(\mathbf{u}) e^{-i\Re(\langle \mathbf{u}, \mathbf{x} \rangle)} d\mathbf{u}}{\int_{\mathbf{u} \in \mathbb{K}^K} \varphi_{\mathbf{x}}(\mathbf{u}) e^{-i\Re(\langle \mathbf{u}, \mathbf{x} \rangle)} d\mathbf{u}} \right] \Gamma_{\mathbf{x}}(d\boldsymbol{\theta}) \\ &= \int_{\boldsymbol{\theta} \in \mathcal{S}^K} \psi(\boldsymbol{\theta}) g_{\chi}(\mathbf{x}, \boldsymbol{\theta}) \Gamma_{\mathbf{x}}(d\boldsymbol{\theta}), \end{aligned} \quad (10.13)$$

où nous avons utilisé la définition de g_{χ} en (10.9) dans la dernière égalité. L'équation (10.8) est obtenue en identifiant (10.13) avec (10.5), pour n'importe quelle fonction continue ψ . \square

Ainsi, ce théorème nous décrit une relation entre l'estimateur choisi $\hat{\chi}$ du *spectre spatial* et la *mesure spatiale* $\Gamma_{\mathbf{x}}$. Intéressons-nous à une propriété particulière de g_{χ} .

10.2.2 Continuité de g_{χ}

Nous allons démontrer le résultat suivant :

Proposition 10.2.2. *La fonction g_{χ} définie dans (10.9) est une fonction continue de $\boldsymbol{\theta}$.*

Démonstration. $\forall \mathbf{u} \in \mathbb{K}^K, \boldsymbol{\theta} \mapsto h(\boldsymbol{\theta}, \mathbf{u}) = \langle \boldsymbol{\theta}, \mathbf{u} \rangle^{(\alpha-1)} \varphi_{\mathbf{x}}(\mathbf{u}) e^{-i\Re(\langle \mathbf{u}, \mathbf{x} \rangle)}$ est continue. Par ailleurs, la densité de probabilité (non-dégénérée) d'une distribution $S\alpha S^K$ est infiniment continue. Cela implique que pour tout $\mathbf{u} \in \mathbb{K}^K$, la fonction caractéristique $\varphi_{\mathbf{x}}(\mathbf{u})$ décroît plus rapidement que n'importe quelle puissance de $\|\mathbf{u}\|$ (avec $\|\cdot\|, \cdot$, une norme quelconque sur l'espace de dimension fini \mathbb{K}^K) [82]. En particulier, $\mathbf{u} \mapsto f(\mathbf{u}) = \|\mathbf{u}\|^{\alpha-1} |\varphi_{\mathbf{x}}(\mathbf{u})|$ est intégrable. De plus, pour tout $\boldsymbol{\theta} \in \mathcal{S}^K$ et $\mathbf{u} \in \mathbb{K}^K$

$$\left| \langle \boldsymbol{\theta}, \mathbf{u} \rangle^{(\alpha-1)} \right| \leq \|\mathbf{u}\|^{\alpha-1} \quad (10.14)$$

conduisant finalement à l'inégalité

$$h(\boldsymbol{\theta}, \mathbf{u}) \leq f(\mathbf{u}). \quad (10.15)$$

En appliquant le théorème de continuité sous le signe somme de l'intégrale, g_{χ} est bien une fonction continue en $\boldsymbol{\theta}$. \square

Ce théorème apporte alors l'information importante que l'estimateur choisi $\hat{\chi}$ est dominé par $\Gamma_{\mathbf{x}}$.

10.3 Simplification de g_χ

10.3.1 Motivations et description

Comme le montre l'équation (10.9), le calcul de g_χ requiert deux intégrations sur \mathbb{K}^K . Il est en pratique impossible d'effectuer de tels calculs. Pour résoudre ce problème, nous devons nous ramener à des calculs d'intégrales sur des ensembles finis. Nous allons donc démontrer qu'il est possible de se ramener à des intégrales sur \mathcal{S}^K :

Proposition 10.3.1. *Soit*

$$\epsilon = \begin{cases} 1 & \text{si } \mathbb{K} = \mathbb{R} \\ 2 & \text{si } \mathbb{K} = \mathbb{C} \end{cases} \quad (10.16)$$

Le numérateur de $g_\chi(\mathbf{x}, \boldsymbol{\theta})$ dans (10.9) se réécrit

$$\forall \boldsymbol{\theta} \in \mathbb{K}^K, N(\mathbf{x}, \boldsymbol{\theta}) = \int_{\boldsymbol{\theta}' \in \mathcal{S}^K} \frac{\langle \boldsymbol{\theta}, \boldsymbol{\theta}' \rangle^{\langle \alpha-1 \rangle} \langle \boldsymbol{\theta}', \mathbf{x} \rangle}{I_{\mathbf{x}}(\boldsymbol{\theta}')^{\frac{\epsilon K + \alpha}{\alpha}}} \eta \left(\frac{|\langle \boldsymbol{\theta}', \mathbf{x} \rangle|^2}{I_{\mathbf{x}}(\boldsymbol{\theta}')^{\frac{2}{\alpha}}} \right) d\boldsymbol{\theta}', \quad (10.17)$$

où :

$$\forall \rho \in \mathbb{K}, \eta(\rho) = \sum_{n=0}^{+\infty} \frac{(-1)^n f_\Gamma \left(\frac{2n + \epsilon K + \alpha}{\alpha} \right)}{2^{2n+1} n! (n+1)!} \rho^n, \quad (10.18)$$

f_Γ désignant la fonction gamma usuelle [8] et $I_{\mathbf{x}} \equiv -\ln(\varphi_{\mathbf{x}})$ est l'exposant de Lévy du vecteur \mathbf{x} . De même, une forme équivalente du dénominateur de g_χ est la suivante :

$$D(\mathbf{x}) = \frac{\left\| \int_{\mathcal{S}^K} \boldsymbol{\theta} N(\mathbf{x}, \boldsymbol{\theta}) \Gamma_{\mathbf{x}}(d\boldsymbol{\theta}) \right\|_1}{\|\mathbf{x}\|_1}, \quad (10.19)$$

où pour $\mathbf{u} = (u_1, \dots, u_K)^\top \in \mathbb{K}^K$, $\|\mathbf{u}\|_1 \triangleq \sum_{k=1}^K |u_k|$ est la norme ℓ_1 usuelle.

Démonstration. Considérons d'abord $N(\mathbf{x}, \boldsymbol{\theta})$ dans l'expression (10.17) de $g_\chi(\mathbf{x}, \boldsymbol{\theta})$, sur lequel est appliqué le changement de variable $\mathbf{u} \triangleq r\boldsymbol{\theta}' \in \mathbb{K}^K$ où $r \in \mathbb{R}^+$ et $\boldsymbol{\theta}' \in \mathcal{S}^K$, entraînant la forme différentielle $d\mathbf{u} = r^{\epsilon K-1} dr d\boldsymbol{\theta}'$. Alors, nous obtenons pour tout $\boldsymbol{\theta} \in \mathbb{K}^K$:

$$N(\mathbf{x}, \boldsymbol{\theta}) = i \int_{\boldsymbol{\theta}' \in \mathcal{S}^K} \langle \boldsymbol{\theta}, \boldsymbol{\theta}' \rangle^{\langle \alpha-1 \rangle} \left(\int_{r \in \mathbb{R}_+} \left(\alpha r^{\alpha-1} e^{-r^\alpha I_{\mathbf{x}}(\boldsymbol{\theta}')} \right) \left(r^{\epsilon K-1} e^{-ir \Re(\langle \boldsymbol{\theta}', \mathbf{x} \rangle)} \right) dr \right) d\boldsymbol{\theta}'.$$

Une intégration par parties sur $\int_{r \in \mathbb{R}_+} \left(\alpha r^{\alpha-1} e^{-r^\alpha I_{\mathbf{x}}(\boldsymbol{\theta}')} \right) \left(r^{\epsilon K-1} e^{-ir \Re(\langle \boldsymbol{\theta}', \mathbf{x} \rangle)} \right) dr$ entraîne :

$$N(\mathbf{x}, \boldsymbol{\theta}) = A(\boldsymbol{\theta}, \mathbf{x}) + (\epsilon K - 1) i B(\boldsymbol{\theta}, \mathbf{x}) \quad (10.20)$$

où

$$A(\boldsymbol{\theta}, \mathbf{x}) = \int_{\boldsymbol{\theta}' \in \mathcal{S}^K} \frac{\langle \boldsymbol{\theta}, \boldsymbol{\theta}' \rangle^{\langle \alpha-1 \rangle} \Re(\langle \boldsymbol{\theta}', \mathbf{x} \rangle)}{I_{\mathbf{x}}(\boldsymbol{\theta}')} \left(\int_{r \in \mathbb{R}_+} r^{\epsilon K-1} e^{-r^\alpha I_{\mathbf{x}}(\boldsymbol{\theta}')} e^{-ir \Re(\langle \boldsymbol{\theta}', \mathbf{x} \rangle)} dr \right) d\boldsymbol{\theta}'$$

$$B(\boldsymbol{\theta}, \mathbf{x}) = \int_{\boldsymbol{\theta}' \in \mathcal{S}^K} \frac{\langle \boldsymbol{\theta}, \boldsymbol{\theta}' \rangle^{\langle \alpha-1 \rangle}}{I_{\mathbf{x}}(\boldsymbol{\theta}')} \left(\int_{r \in \mathbb{R}_+} r^{\epsilon K-2} e^{-r^\alpha I_{\mathbf{x}}(\boldsymbol{\theta}')} e^{-ir \Re(\langle \boldsymbol{\theta}', \mathbf{x} \rangle)} dr \right) d\boldsymbol{\theta}'.$$

En développant en série entière l'exponentielle complexe puis en appliquant l'identité $\int_0^{+\infty} x^n e^{-ax^b} dx = \frac{1}{b} a^{-\frac{n+1}{b}} f_\Gamma \left(\frac{n+1}{b} \right)$ aux intégrales $A(\boldsymbol{\theta}, \mathbf{x})$ et $B(\boldsymbol{\theta}, \mathbf{x})$, il en découle que :

$$A(\boldsymbol{\theta}, \mathbf{x}) = \frac{1}{\alpha} \int_{\boldsymbol{\theta}' \in \mathcal{S}^K} \sum_{n=0}^{+\infty} \frac{(-i)^n \text{fr}\left(\frac{n+\epsilon K}{\alpha}\right) \langle \boldsymbol{\theta}, \boldsymbol{\theta}' \rangle^{<\alpha-1>}}{n!} \frac{\frac{1}{2^{n+1}} \mathcal{B}_{n+1}(\langle \boldsymbol{\theta}', \mathbf{x} \rangle, \langle \mathbf{x}, \boldsymbol{\theta}' \rangle)}{I_{\mathbf{x}}(\boldsymbol{\theta}')^{\frac{n+\epsilon K+\alpha}{\alpha}}} d\boldsymbol{\theta}' \quad (10.21)$$

$$B(\boldsymbol{\theta}, \mathbf{x}) = \frac{1}{\alpha} \int_{\boldsymbol{\theta}' \in \mathcal{S}^K} \sum_{n=0}^{+\infty} \frac{(-i)^n \text{fr}\left(\frac{n+\epsilon K-1}{\alpha}\right) \langle \boldsymbol{\theta}, \boldsymbol{\theta}' \rangle^{<\alpha-1>}}{n!} \frac{\frac{1}{2^n} \mathcal{B}_n(\langle \boldsymbol{\theta}', \mathbf{x} \rangle, \langle \mathbf{x}, \boldsymbol{\theta}' \rangle)}{I_{\mathbf{x}}(\boldsymbol{\theta}')^{\frac{n+\epsilon K-1+\alpha}{\alpha}}} d\boldsymbol{\theta}', \quad (10.22)$$

où $\mathcal{B}_n(u, v) = (u+v)^n = \sum_{k=0}^n \binom{n}{k} u^k v^{n-k}$ désigne le binôme de Newton. Avant de poursuivre le calcul de l'intégrale, nous remarquons que

$$N(z\mathbf{x}, \boldsymbol{\theta}) = zN(\mathbf{x}, \boldsymbol{\theta}), \forall z \in \mathcal{S}_{\mathbb{K}}^1. \quad (10.23)$$

En effet, en reprenant la formule (10.9), nous avons le résultat suivant :

$$\begin{aligned} N(z\mathbf{x}, \boldsymbol{\theta}) &= \alpha i \int_{\mathbf{u} \in \mathbb{K}^K} \langle \boldsymbol{\theta}, \mathbf{u} \rangle^{<\alpha-1>} \varphi_{z\mathbf{x}}(\mathbf{u}) e^{-i\Re(\langle \mathbf{u}, z\mathbf{x} \rangle)} d\mathbf{u} \\ &= \alpha i \int_{\mathbf{u} \in \mathbb{K}^K} \langle \boldsymbol{\theta}, \mathbf{u} \rangle^{<\alpha-1>} \varphi_{z\mathbf{x}}(\mathbf{u}) e^{-i\Re(\langle z^{-1}\mathbf{u}, \mathbf{x} \rangle)} d\mathbf{u} \\ &= \alpha i \int_{\mathbf{v} \in \mathbb{K}^K} \langle \boldsymbol{\theta}, z\mathbf{v} \rangle^{<\alpha-1>} \varphi_{z\mathbf{x}}(z\mathbf{v}) e^{-i\Re(\langle \mathbf{v}, \mathbf{x} \rangle)} d\mathbf{v} \\ &= z\alpha i \int_{\mathbf{v} \in \mathbb{K}^K} \langle \boldsymbol{\theta}, \mathbf{v} \rangle^{<\alpha-1>} \varphi_{\mathbf{x}}(\mathbf{v}) e^{-i\Re(\langle \mathbf{v}, \mathbf{x} \rangle)} d\mathbf{v} \\ &= zN(\mathbf{x}, \boldsymbol{\theta}) \end{aligned}$$

où le passage de la première égalité à la seconde utilise le fait que $z \in \mathcal{S}_{\mathbb{K}}^1$. Autrement dit, $z^* = z^{-1}$ pour $\mathbb{K} = \mathbb{C}$ et $z = z^{-1}$ pour $\mathbb{K} = \mathbb{R}$ (au sens de l'inverse dans le groupe additif $(\mathbb{R}, +)$). Le passage de la deuxième à la troisième égalité s'effectue par le changement de variable $\mathbf{u} = z\mathbf{v}$. Enfin, de la troisième à la quatrième égalité, nous utilisons le fait que \mathbf{x} est un vecteur isotrope. En effet, on a $z\mathbf{x} \stackrel{d}{=} \mathbf{x}$ et $\varphi_{z\mathbf{x}}(z\mathbf{v}) = \varphi_{\mathbf{x}}(\mathbf{v})$ pour tout \mathbf{v} . Au final, la fonction de puissance signée est la seule à faire sortir un coefficient z . En combinant les équations (10.21) et (10.22) dans l'équation (10.20), nous obtenons :

$$\begin{aligned} A(\boldsymbol{\theta}, \mathbf{x}) + (1 - \epsilon K)(-i)B(\boldsymbol{\theta}, \mathbf{x}) &= \frac{1}{\alpha} \int_{\boldsymbol{\theta}' \in \mathcal{S}^K} \frac{\langle \boldsymbol{\theta}, \boldsymbol{\theta}' \rangle^{<\alpha-1>}}{I_{\mathbf{x}}(\boldsymbol{\theta}')^{\frac{\epsilon K + \alpha}{\alpha}}} \left(\sum_{n=0}^{+\infty} \frac{(-i)^n \text{fr}\left(\frac{n+\epsilon K}{\alpha}\right)}{n!} \frac{\frac{1}{2^{n+1}} \mathcal{B}_{n+1}(\langle \boldsymbol{\theta}', \mathbf{x} \rangle, \langle \mathbf{x}, \boldsymbol{\theta}' \rangle)}{I_{\mathbf{x}}(\boldsymbol{\theta}')^{\frac{n}{\alpha}}} \right. \\ &\quad \left. + (1 - \epsilon K) \sum_{n=0}^{+\infty} \frac{(-i)^{n+1} \text{fr}\left(\frac{n+\epsilon K-1}{\alpha}\right)}{n!} \frac{\frac{1}{2^n} \mathcal{B}_n(\langle \boldsymbol{\theta}', \mathbf{x} \rangle, \langle \mathbf{x}, \boldsymbol{\theta}' \rangle)}{I_{\mathbf{x}}(\boldsymbol{\theta}')^{\frac{n-1}{\alpha}}} \right) d\boldsymbol{\theta}'. \end{aligned}$$

Le résultat de l'équation (10.23) justifie alors que seuls les termes de chaque binôme de Newton seront modifiés. En effet, $I_{\mathbf{x}}$ est l'exposant de Lévy d'un vecteur $S\alpha S_c^K$: il restera invariant. Pour la première série dans $A(\boldsymbol{\theta}, \mathbf{x})$, on peut affirmer sans perte de généralité que

$$\begin{aligned} z\mathcal{B}_{n+1}(\langle \boldsymbol{\theta}', \mathbf{x} \rangle, \langle \mathbf{x}, \boldsymbol{\theta}' \rangle) &= \mathcal{B}_{n+1}(\langle \boldsymbol{\theta}', z\mathbf{x} \rangle, \langle z\mathbf{x}, \boldsymbol{\theta}' \rangle) \\ \Leftrightarrow z \sum_{k=0}^{n+1} \binom{n+1}{k} \langle \boldsymbol{\theta}', \mathbf{x} \rangle^k \langle \mathbf{x}, \boldsymbol{\theta}' \rangle^{n+1-k} &= \sum_{k=0}^{n+1} \binom{n+1}{k} z^k (z^*)^{n+1-k} \langle \boldsymbol{\theta}', \mathbf{x} \rangle^k \langle \mathbf{x}, \boldsymbol{\theta}' \rangle^{n+1-k}. \end{aligned}$$

C'est-à-dire

$$z^{k-1} \left(z^{n+1-k} \right)^* = 1 \Leftrightarrow \left(k = \frac{n}{2} + 1 \right) \text{ et } (n \text{ est pair})$$

ou encore

$$\mathcal{B}_{n+1}(\langle \theta', \mathbf{x} \rangle, \langle \mathbf{x}, \theta' \rangle) = \frac{(n+1)!}{\left(\frac{n}{2}\right)! \left(\frac{n+1}{2}\right)!} |\langle \theta', \mathbf{x} \rangle|^n \langle \theta', \mathbf{x} \rangle.$$

Avec des arguments similaires, nous pouvons démontrer que la condition pour le deuxième binôme de Newton est équivalente à supposer que

$$\left(k = \frac{n+1}{2}\right) \text{ et } (n \text{ est impair}),$$

d'où

$$\mathcal{B}_n(\langle \theta', \mathbf{x} \rangle, \langle \mathbf{x}, \theta' \rangle) = \frac{n!}{\left(\frac{n-1}{2}\right)! \left(\frac{n+1}{2}\right)!} |\langle \theta', \mathbf{x} \rangle|^n \langle \theta', \mathbf{x} \rangle.$$

Ainsi, en posant $n = 2n$ et $n = 2n + 1$ respectivement dans la première série et dans la deuxième série, nous en déduisons que

$$\begin{aligned} A(\theta, \mathbf{x}) + (\epsilon K - 1) i B(\theta, \mathbf{x}) &= \frac{1}{\alpha} \int_{\theta' \in \mathcal{S}^K} \frac{\langle \theta, \theta' \rangle^{<\alpha-1>}}{I_X(\theta')^{\frac{\epsilon K + \alpha}{\alpha}}} \left(\sum_{n=0}^{+\infty} \frac{(-1)^n (2n+1) f_{\Gamma}\left(\frac{2n+\epsilon K}{\alpha}\right) |\langle \theta', \mathbf{x} \rangle|^{2n} \langle \theta', \mathbf{x} \rangle}{n!(n+1)! 2^{2n+1} I_X(\theta')^{\frac{2n}{\alpha}}} + \right. \\ &\quad \left. (\epsilon K - 1) \sum_{n=0}^{+\infty} \frac{(-1)^n f_{\Gamma}\left(\frac{2n+\epsilon K}{\alpha}\right) |\langle \theta', \mathbf{x} \rangle|^{2n} \langle \theta', \mathbf{x} \rangle}{n!(n+1)! 2^{2n+1} I_X(\theta')^{\frac{2n}{\alpha}}} \right) d\theta' \\ &= \int_{\theta' \in \mathcal{S}^K} \frac{\langle \theta, \theta' \rangle^{<\alpha-1>} \langle \theta', \mathbf{x} \rangle}{I_X(\theta')^{\frac{\epsilon K + \alpha}{\alpha}}} \left(\sum_{n=0}^{+\infty} \frac{(-1)^n \left[n + \frac{\epsilon K}{2}\right] f_{\Gamma}\left(\frac{2n+\epsilon K}{\alpha}\right) |\langle \theta', \mathbf{x} \rangle|^{2n}}{n!(n+1)! 2^{2n} I_X(\theta')^{\frac{2n}{\alpha}}} \right) d\theta' \end{aligned}$$

d'où le résultat final annoncé en utilisant $z f_{\Gamma}(z) = f_{\Gamma}(z+1)$, $\forall z$; $\Re(z) > 0$ et en réajustant les indices. Ajoutons à cela que l'équation (10.18) définit une série entière avec un rayon de convergence infini quand $\alpha > 1$. Cela peut par exemple se démontrer en employant la règle de D'Alembert, où un équivalent à l'infini des fonctions f_{Γ} doit être fourni, indépendant du modèle de mélange.

L'estimation de $\mathbf{x} \mapsto D(\mathbf{x})$ est obtenue en remarquant que :

$$\mathbf{x} = \int_{\theta \in \mathcal{S}^K} \theta g_X(\mathbf{x}, \theta) \Gamma_X(d\theta) = \int_{\theta \in \mathcal{S}^K} \theta \frac{N(\mathbf{x}, \theta)}{D(\mathbf{x})} \Gamma_X(d\theta). \quad (10.24)$$

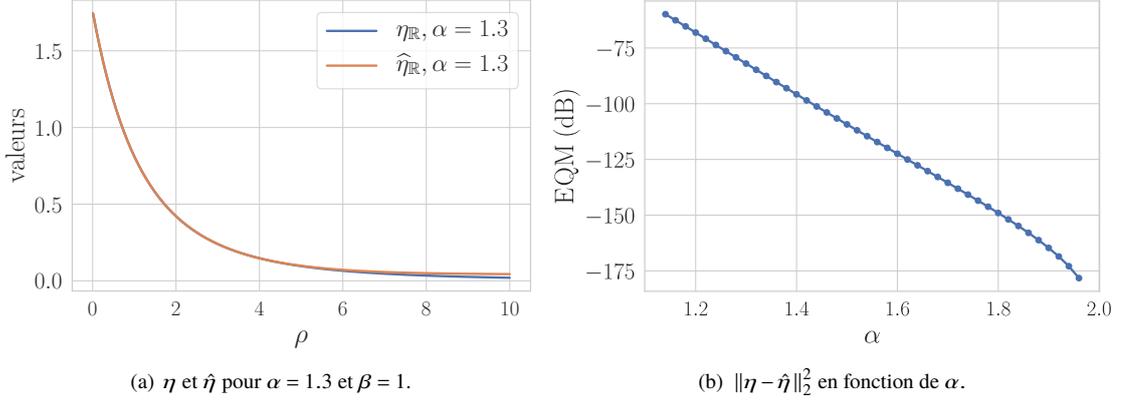
Le choix de la norme ℓ_1 a été établie de manière empirique en testant plusieurs normes. À noter que n'importe quelle norme peut ici être valable. □

Grâce à cette proposition, la complexité est réduite aux calculs d'intégrales le long de la sphère. Cependant, un autre problème surgissant est celui du calcul de la série en (10.18). Explicitons comment une telle série sera modélisée et approximée via une fonction analytique.

10.3.2 Modélisation de séries numériques

La série de l'équation (10.18) demande de considérer un nombre de termes très important pour obtenir une valeur numériquement exploitable. En effet, cette série alternée peut en pratique osciller vers des valeurs importantes quand ρ devient grand. Cela entraîne naturellement une mauvaise estimation. Nous allons nous inspirer de la forme analytique prise par la série quand $\alpha = 2$. Donnons par exemple η quand $\alpha = 2, K = 2$ et $\epsilon = 2$:

$$\forall \rho \in \mathbb{C}, \eta(\rho) = \frac{1}{16} \left(\rho^2 - 20\rho + 64 \right) e^{-\rho/4}. \quad (10.25)$$

FIGURE 10.2 – Évaluation d'estimation des paramètres du modèle $\hat{\eta}$, pour η dans (10.18)

Proche de la forme analytique esquissée en (10.25), nous faisons le choix d'approximer η par la fonction suivante pour tout α, K, β :

$$\forall \rho \in \mathbb{R}_+, \eta(\rho) \simeq \hat{\eta}(\rho) = \left(a\rho^2 + b\rho + c \right) e^{-d\rho} \quad (10.26)$$

où $a, b, c, d \in \mathbb{R}$ sont des paramètres du modèle. Le choix de $\rho \in \mathbb{R}_+$ se justifie par le fait que nous évaluons cette série aux points $\frac{I(\theta', \mathbf{x})^2}{I_x(\theta')^{\frac{2}{\alpha}}} \geq 0$. Grâce à (10.26), le temps total d'estimation de χ sera considérablement réduit. Nous avons choisi pour l'estimation des paramètres le critère de minimisation de l'erreur quadratique moyenne (EQM) $\|\eta - \hat{\eta}\|_2^2$. Les graphiques de η et $\hat{\eta}$ pour $\beta = 1$ sont affichés dans la Fig. 10.2(a), et l'erreur d'ajustement de η pour $\beta = 2$ en fonction de α est décrite dans la Fig. 10.2(b). Les résultats sont similaires pour le cas où $\beta = 1$ et $\beta = 2$.

La série η a été calculée pour 50 exposants caractéristiques $\alpha \in]1, 2[$ régulièrement espacés et pour $\rho \in]0, 10[$. La borne de ρ à 10 se justifie par les valeurs atteintes par la série : en effet pour $\rho \simeq 10$, les valeurs proches de zéro atteignent l'équivalent de 5 fois la précision machine (en flottant 32 bits, soit environ 3×10^{-7}). Ajoutons à cela que plus ρ devient grand et plus le nombre de termes à calculer pour obtenir une valeur numérique convenable grandit aussi. Pour l'obtention de courbes appropriées à l'estimation des paramètres, 10^5 termes ont été calculés.

Nous observons que l'erreur induite par notre modèle paramétrique est minimale (≤ -75 dB pour $\alpha > 1.3$) pour une accélération notable en terme de temps de calcul. Par conséquent, nous considérons que l'approximation $\hat{\eta}$ en (10.26) est suffisante.

10.4 Algorithme de Séparation

10.4.1 Reconstruction des sources images

Maintenant que nous sommes en mesure d'estimer $\hat{\chi}$ via un calcul de g_{χ} , il nous reste à fournir un estimateur des vecteurs aléatoires \mathbf{y}_j . Nous choisissons de considérer l'espérance *a posteriori*

$$\forall j, \hat{\mathbf{y}}_j \triangleq \mathbb{E}[\mathbf{y}_j | \mathbf{x}] \quad (10.27)$$

comme estimateur. Nous allons démontrer qu'elle s'exprime comme suit :

Théorème 10.4.1. Soit \mathbf{x} la somme de vecteurs aléatoires $\mathbf{y}_j \sim S\alpha S_c^K(\Gamma_j)$, chacun associé à sa mesure spatiale Γ_j . Alors, l'espérance a posteriori de chaque composante \mathbf{y}_j sachant \mathbf{x} est :

$$\forall j, \widehat{\mathbf{y}}_j \triangleq \mathbb{E}[\mathbf{y}_j | \mathbf{x}] = \int_{\boldsymbol{\theta} \in \mathcal{S}^K} \boldsymbol{\theta} g_{\chi}(\mathbf{x}, \boldsymbol{\theta}) \Gamma_j(d\boldsymbol{\theta}), \quad (10.28)$$

où g_{χ} est définie en (10.9).

Démonstration. Comme dans la démonstration de la propriété 10.2.1, nous déterminons de prime abord la distribution jointe de \mathbf{y}_j et \mathbf{x} . Ensuite, nous utilisons l'égalité (1.7) reliant la fonction caractéristique et l'espérance du vecteur aléatoire $\mathbf{y}_j | \mathbf{x}$. Nous avons pour tous $\mathbf{v}, \mathbf{u} \in \mathbb{K}^K$:

$$\begin{aligned} \varphi_{(\mathbf{y}_j, \mathbf{x})}(\mathbf{v}, \mathbf{u}) &\triangleq \mathbb{E} \left[e^{i\Re(\langle \mathbf{v}, \mathbf{y}_j \rangle + \langle \mathbf{u}, \mathbf{x} \rangle)} \right] \\ &= \mathbb{E} \left[e^{i\Re(\langle \mathbf{v} + \mathbf{u}, \mathbf{y}_j \rangle + \sum_{j' \neq j} \langle \mathbf{u}, \mathbf{y}_{j'} \rangle)} \right] \\ &= \varphi_{\mathbf{y}_j}(\mathbf{v} + \mathbf{u}) \prod_{j' \neq j} \varphi_{\mathbf{y}_{j'}}(\mathbf{u}) \\ &= e^{-\int_{\boldsymbol{\theta} \in \mathcal{S}^K} |\langle \mathbf{v} + \mathbf{u}, \boldsymbol{\theta} \rangle|^{\alpha} \Gamma_j(d\boldsymbol{\theta}) - \sum_{j' \neq j} \int_{\boldsymbol{\theta} \in \mathcal{S}^K} |\langle \mathbf{u}, \boldsymbol{\theta} \rangle|^{\alpha} \Gamma_{j'}(d\boldsymbol{\theta})}. \end{aligned}$$

Cela nous permet de déduire la fonction caractéristique de $\mathbf{y}_j | \mathbf{x}$ pour tout $\mathbf{u} \in \mathbb{K}^K$:

$$\varphi_{\mathbf{y}_j | \mathbf{x}}(\mathbf{u}) = \frac{\int_{\mathbf{v} \in \mathbb{K}^K} e^{-\int_{\boldsymbol{\theta} \in \mathcal{S}^K} |\langle \mathbf{u} + \mathbf{v}, \boldsymbol{\theta} \rangle|^{\alpha} \Gamma_j(d\boldsymbol{\theta}) - \sum_{j' \neq j} \int_{\boldsymbol{\theta} \in \mathcal{S}^K} |\langle \mathbf{v}, \boldsymbol{\theta} \rangle|^{\alpha} \Gamma_{j'}(d\boldsymbol{\theta})} e^{-i\Re(\langle \mathbf{v}, \mathbf{x} \rangle)} d\mathbf{v}}{\int_{\mathbf{v} \in \mathbb{K}^K} \varphi_{\mathbf{x}}(\mathbf{v}) e^{-i\Re(\langle \mathbf{v}, \mathbf{x} \rangle)} d\mathbf{v}}.$$

La preuve de l'existence de la dérivée de Wirtinger $\nabla_{\mathbf{v}^*} \varphi_{\mathbf{y}_j | \mathbf{x}}$ utilise exactement les mêmes arguments que ceux de la démonstration de la proposition 10.2.1. Ainsi, nous obtenons :

$$\begin{aligned} \mathbb{E}[\mathbf{y}_j | \mathbf{x}] &= -i \nabla_{\mathbf{v}^*} \varphi_{\mathbf{y}_j | \mathbf{x}}(\mathbf{v} = \mathbf{0}) \\ &= \int_{\boldsymbol{\theta} \in \mathcal{S}^K} \boldsymbol{\theta} \left[\frac{\alpha i \int_{\mathbf{u} \in \mathbb{K}^K} \langle \mathbf{u}, \boldsymbol{\theta} \rangle^{\alpha-1} \varphi_{\mathbf{x}}(\mathbf{u}) e^{-i\Re(\langle \mathbf{u}, \mathbf{x} \rangle)} d\mathbf{u}}{\int_{\mathbf{u} \in \mathbb{K}^K} \varphi_{\mathbf{x}}(\mathbf{u}) e^{-i\Re(\langle \mathbf{u}, \mathbf{x} \rangle)} d\mathbf{u}} \right] \Gamma_j(d\boldsymbol{\theta}) \\ &= \int_{\boldsymbol{\theta} \in \mathcal{S}^K} \boldsymbol{\theta} g_{\chi}(\mathbf{x}, \boldsymbol{\theta}) \Gamma_j(d\boldsymbol{\theta}), \end{aligned}$$

ce qui achève la démonstration. □

Ce théorème de séparation nous affirme ainsi que chaque \mathbf{y}_j peut être estimé grâce à la connaissance de g_{χ} , de la mesure spatiale Γ_j et de l'observation \mathbf{x} . Il est à noter que cette méthode de filtrage n'est pas linéaire en \mathbf{x} .

10.4.2 Résumé de la méthode de filtrage

Algorithme 6 α -FSS : Filtrage α -stable multivarié via une estimation du spectre spatial.

1. Entrées

- Observation \mathbf{x} de dimension K
- Partition régulière $\Theta = \{\Theta_1, \dots, \Theta_p\}$ de \mathcal{S}^K
- Exposant caractéristique $\alpha \in (1, 2)$
- Mesures spatiales Γ_j

2. Estimation du spectre spatial

- Utilisation de $\hat{\eta}$ en (10.26), calcul $\forall p, N(\mathbf{x}, \theta_p)$ dans (10.17)
- Calcul de $D(\mathbf{x})$ dans (10.19)
- Calcul de $g_\chi(\mathbf{x}, \theta_p) = \frac{N(\mathbf{x}, \theta_p)}{D(\mathbf{x})}$

3. Reconstruction : $\hat{\mathbf{y}}_j = \sum_p \theta_p g_\chi(\mathbf{x}, \theta_p) \Gamma_j(\Theta_p)$

L'algorithme 6 résume la méthode de séparation d'un mélange de vecteurs $S\alpha S_c^K$. Soulignons le fait que le calcul de g_χ , décomposé via son numérateur en (10.17) et son dénominateur en (10.19), fait intervenir des calculs d'intégrales le long de la sphère \mathcal{S}^K . Nous utilisons la méthode d'intégration décrite à l'équation (9.7) en employant une approximation discrète. En pratique, l'approximation des séries numériques présentée dans la sous-section 10.3.2 est pré-calculée et sauvegardée dans un fichier, afin d'accélérer l'exécution de l'algorithme.

Chapitre 11

Filtrage linéaire α -stable : minimisation de la norme de covariation

Résumé

Dans le cas où les moments d'ordre 2 sont définis, l'estimateur linéaire d'une composante au sein d'un mélange, minimisant l'erreur quadratique moyenne, porte le nom de filtre de Wiener. Au final, le masque de Wiener dépend de la covariance du mélange et de celle entre le mélange et la source souhaitée. Si nous prenons des vecteurs α -stables non-gaussiens, la covariance n'est pas définie. Pour $\alpha > 1$, nous la remplaçons par la covariation qui joue un rôle similaire à de nombreux égards. La méthode de filtrage proposée est alors un problème d'optimisation où nous minimisons la norme de covariation entre le signal désiré et son estimé. Ce problème admet une solution unique et se résume à une méthode dite de point fixe. Nous montrons également que le cas limite quand $\alpha \rightarrow 2$ coïncide avec le filtre de Wiener multicanal, réécrit en exploitant le théorème 9.2.1 page 75 de représentation spatiale.

Dans tout ce chapitre, nous supposons que $\alpha \in]1, 2[$.

11.1 Introduction : un nouvel estimateur linéaire

Rappelons tout d'abord le filtrage de Wiener multicanal. Soit le modèle de mélange de dimension K suivant :

$$\left\{ \mathbf{x} = \sum_{j=1}^J \mathbf{y}_j \right. \quad (11.1)$$

où les sources $\mathbf{y}_j \in \mathbb{K}^K$ sont décorrélées. Pour la suite, nous supposons par ailleurs que chaque source admet ses moments d'ordre 2. Un filtrage linéaire peut alors être vu comme le produit matriciel entre des matrices $\mathbf{W}_j \in \mathbb{K}^{K \times K}$ dépendant de la $j^{\text{ème}}$ source et le mélange \mathbf{x} . Autrement dit, l'estimateur $\hat{\mathbf{y}}_j \in \mathbb{K}^K$ souhaité serait de la forme :

$$\forall j, \hat{\mathbf{y}}_j = \mathbf{W}_j \mathbf{x}. \quad (11.2)$$

Cherchons à minimiser l'erreur quadratique moyenne (EQM) entre l'estimation $\hat{\mathbf{y}}_j$ et la vérité terrain \mathbf{y}_j . Ou encore

$$\begin{aligned} &\text{minimiser} && \sum_j \mathbb{E} \left[\|\mathbf{y}_j - \hat{\mathbf{y}}_j\|_2^2 \right] \text{ avec } \mathbf{W}_j \\ &\text{sous contrainte que} && \hat{\mathbf{y}}_j = \mathbf{W}_j \mathbf{x}. \end{aligned} \quad (11.3)$$

Il est alors prouvé que ce problème admet une unique solution . En résolvant $\nabla_{\mathbf{W}_j} \sum_j \mathbb{E} \left[\|\mathbf{y} - \mathbf{W}_j \mathbf{x}\|_2^2 \right] = 0$, nous pouvons montrer que ce problème est alors équivalent à considérer les *masques de Wiener* suivants [179] :

$$\mathbf{W}_j = \mathbb{E} [\mathbf{y}_j \mathbf{x}^*] \mathbb{E} [\mathbf{x} \mathbf{x}^*]^{-1}. \quad (11.4)$$

Une question naturelle serait de se demander si cette approche est valable dans le cas où les sources considérées sont supposées suivre une loi α -stable. Pour le cas Gaussien, nous savons que les moments à tout ordre existent, donc le filtrage de Wiener est applicable. Pour $\alpha < 2$, les moments d'ordre 2 ne sont pas définis. Dans l'état, ce problème d'optimisation d'un filtrage α -stable multicanal linéaire ne peut être proposé. Pour généraliser une telle approche, Miller proposa en 1978 [133] un tout nouvel opérateur dans le cas où $\alpha > 1$, satisfaisant des propriétés relativement proches de la covariance.

11.2 Covariation

11.2.1 Définitions, exemples et propriétés

Dans le cadre de la séparation, il existe de nombreuses techniques utilisant les moments statistiques d'ordre 2 [29, 39]. Malheureusement, dans le cas des distributions α -stables non-gaussiennes, les moments statistiques d'ordre $p \geq \alpha$ n'existent pas. Cependant, en exploitant la mesure spatiale $\Gamma_{\mathbf{x}}$ d'un vecteur $\mathbf{x} \sim S\alpha S^2(\Gamma_{\mathbf{x}})$, il est possible de proposer une définition proche de la covariance entre deux variables aléatoires [153, 136] :

Définition 11.2.1 (Covariation). *Soit $\mathbf{x} \triangleq [x_1, x_2]^T \sim S\alpha S^2(\Gamma_{\mathbf{x}})$ un vecteur aléatoire. La covariation entre les variables x_1 et x_2 est alors définie comme suit :*

$$[x_1, x_2]_{\alpha} \triangleq \int_{\boldsymbol{\theta}=(\theta_1, \theta_2) \in \mathcal{S}^2} \theta_1^* \theta_2^{(\alpha-1)} \Gamma_{\mathbf{x}}(d\boldsymbol{\theta}). \quad (11.5)$$

où pour tout $z \in \mathbb{K}, z^{(\alpha)} \triangleq z|z|^{\alpha-1}$ est la fonction puissance signée.

Nous verrons que la covariation a des propriétés communes à l'opérateur de covariance. Afin d'amener une interprétation plus simple au lecteur, donnons quelques exemples de calculs de covariations.

Exemple 1 (variable gaussienne) Étant donné que la densité spatiale $\Gamma_{\mathbf{x}}$ n'est pas unique dans le cas gaussien, nous supposons que $\alpha \rightarrow 2$. Nous notons abusivement $[x_1, x_2]_2$ à la place de $[x_1, x_2]_{\alpha \rightarrow 2}$. Soit $\Gamma_{\mathbf{x}}$, une mesure spatiale d'un vecteur gaussien centré réel $\mathbf{x} \triangleq (x_1, x_2)$. Alors, en utilisant la définition 9.1.1 page 74, la fonction caractéristique $\varphi_{\mathbf{x}}$ de \mathbf{x} est :

$$\begin{aligned} \forall \mathbf{u} \in \mathbb{R}^2, \varphi_{\mathbf{x}}(\mathbf{u}) &= \exp \left(- \int_{\boldsymbol{\theta}=(\theta_1, \theta_2) \in \mathcal{S}_{\mathbb{R}}^2} (u_1 \theta_1 + u_2 \theta_2)^2 \Gamma_{\mathbf{x}}(d\boldsymbol{\theta}) \right) \\ &= \exp \left(\left[-u_1^2 \int_{\boldsymbol{\theta}=(\theta_1, \theta_2) \in \mathcal{S}_{\mathbb{R}}^2} \theta_1^2 \Gamma_{\mathbf{x}}(d\boldsymbol{\theta}) + \right. \right. \\ &\quad \left. \left. 2u_1 u_2 [x_1, x_2]_2 + u_2^2 \int_{\boldsymbol{\theta}=(\theta_1, \theta_2) \in \mathcal{S}_{\mathbb{R}}^2} \theta_2^2 \Gamma_{\mathbf{x}}(d\boldsymbol{\theta}) \right] \right). \end{aligned} \quad (11.6)$$

D'un autre côté, la fonction caractéristique d'un vecteur gaussien est connue et vaut

$$\forall \mathbf{u} \in \mathbb{R}^2, \varphi_{\mathbf{x}}(\mathbf{u}) = \exp \left(-\frac{1}{2} \left[u_1^2 \mathbb{E}(x_1^2) + 2u_1 u_2 \mathbb{E}(x_1 x_2) + u_2^2 \mathbb{E}(x_2^2) \right] \right). \quad (11.7)$$

En évaluant en $u_1 = 0$ et en identifiant (11.6) et (11.7), nous obtenons

$$\int_{\theta=(\theta_1, \theta_2) \in \mathcal{S}_{\mathbb{R}}^2} \theta_2^2 \Gamma_x(d\theta) = \frac{1}{2} \mathbb{E}(x_2^2). \quad (11.8)$$

En faisant de même pour $u_2 = 0$

$$\int_{\theta=(\theta_1, \theta_2) \in \mathcal{S}_{\mathbb{R}}^2} \theta_1^2 \Gamma_x(d\theta) = \frac{1}{2} \mathbb{E}(x_1^2). \quad (11.9)$$

En substituant (11.9), (11.8) dans (11.6), nous avons

$$\mathbb{E}(x_1 x_2) = 2[x_1, x_2]_2. \quad (11.10)$$

Autrement dit, la covariation est égale à deux fois la covariance quand $\alpha \rightarrow 2$.

Exemple 2 (covariation d'une variable complexe $S\alpha S$) Soit $x \sim S\alpha S(\sigma_x^\alpha)$. Alors :

$$[x, x]_\alpha \triangleq \int_{\theta \in \mathcal{S}_{\mathbb{C}}^1} |\theta|^\alpha \Gamma_x(d\theta). \quad (11.11)$$

Si nous appliquons la définition 9.1.1 page 74 à x , sa fonction caractéristique est

$$\forall u \in \mathbb{C}, \varphi_x(u) = \exp\left(-\int_{\theta \in \mathcal{S}_{\mathbb{C}}^1} |u^* \theta|^\alpha \Gamma_x(d\theta)\right). \quad (11.12)$$

En outre, l'expression explicite de la fonction caractéristique d'une variable symétrique α -stable x est également connue et égal à :

$$\forall u \in \mathbb{C}, \varphi_x(u) = \exp(-\sigma_x^\alpha |u|^\alpha). \quad (11.13)$$

En posant $u = 1$, nous avons l'égalité

$$\exp(-\sigma_x^\alpha) = \exp\left(-\int_{\theta \in \mathcal{S}_{\mathbb{C}}^1} |\theta|^\alpha \Gamma_x(d\theta)\right) = \exp(-[x, x]_\alpha) \quad (11.14)$$

et donc

$$[x, x]_\alpha = \sigma_x^\alpha. \quad (11.15)$$

En interprétant (11.15), nous pouvons affirmer que la covariation d'une variable α stable calcule sa « variance d'ordre α ».

Rappelons quelques propriétés sur la covariation. Le lecteur pourra se référer aux preuves dans [133, 153].

Proposition 11.2.1. Soient x_1, x_2, x_3 trois variables aléatoires sur \mathbb{K} telles que $\mathbf{x} \triangleq (x_1, x_2, x_3) \sim S\alpha S^3(\Gamma_{\mathbf{x}})$. Alors :

1. L'opérateur de covariation est antilinéaire à gauche. C'est-à-dire, pour tout $a, b \in \mathbb{K}$:

$$[ax_1 + bx_2, x_3]_\alpha = a^* [x_1, x_3]_\alpha + b^* [x_2, x_3]_\alpha. \quad (11.16)$$

2. Pour tout $a, b \in \mathbb{K}$:

$$[ax_1, bx_2]_\alpha = ab^{(\alpha-1)} [x_1, x_2]_\alpha. \quad (11.17)$$

3. Si x_2 et x_3 sont indépendants, alors :

$$[x_2, x_3]_\alpha = 0. \quad (11.18)$$

Nous allons proposer une généralisation de l'approche présentée à la section 11.1 visant à minimiser l'erreur quadratique moyenne (EQM) et conduisant à un filtrage de Wiener. Pour ce faire, nous définissons une norme appelée *norme de covariation*.

11.2.2 Norme de covariation

Nous savons que la combinaison linéaire de variables α -stables symétriques reste une variable α -stable symétrique. Cela implique que l'ensemble des variables α -stables symétriques induit un espace vectoriel. Notons

$$\forall \alpha \in]0, 2[, \mathfrak{S}_\alpha \triangleq \left\{ \sum_n a_n x_n \mid \forall n, x_n \sim S\alpha S_c(\sigma_{x_n}^\alpha); a_n \in \mathbb{K} \right\} \quad (11.19)$$

l'espace vectoriel de toutes les combinaisons linéaires de variables α -stables isotropes. Dans le cas où $\alpha > 1$, l'espace induit une norme [153, p 95] :

Proposition 11.2.2 (Norme de covariation). *Pour tout $x \in \mathfrak{S}_\alpha$, nous définissons la fonction :*

$$\|x\|_\alpha = ([x, x]_\alpha)^{1/\alpha}. \quad (11.20)$$

Cette fonction est une norme appelée norme de covariation.

Si on reprend le résultat en (11.15), la norme de covariation d'une variable α -stable symétrique est égale à son paramètre d'échelle :

$$\begin{cases} x \sim S\alpha S(\sigma_x^\alpha) \\ \|x\|_\alpha^\alpha = \sigma_x^\alpha \end{cases}. \quad (11.21)$$

En outre, d'autres propriétés que nous employons pour notre méthode de filtrage sont valables :

Proposition 11.2.3. 1. *Soit x et y , deux variables $S\alpha S$ indépendantes. Alors :*

$$\|x + y\|_\alpha^\alpha = \|x\|_\alpha^\alpha + \|y\|_\alpha^\alpha. \quad (11.22)$$

2. *Soit un vecteur $\mathbf{x} \sim S\alpha S^K(\Gamma_x)$ et $\mathbf{x} \stackrel{d}{=} \int_{\boldsymbol{\theta} \in S^K} \boldsymbol{\theta} \chi(d\boldsymbol{\theta})$ sa représentation spatiale. Alors, pour toute fonction continue ψ définie sur la sphère S^K :*

$$\left\| \int_{\boldsymbol{\theta} \in S^K} \psi(\boldsymbol{\theta}) \chi(d\boldsymbol{\theta}) \right\|_\alpha^\alpha = \int_{\boldsymbol{\theta} \in S^K} |\psi(\boldsymbol{\theta})|^\alpha \Gamma_x(d\boldsymbol{\theta}). \quad (11.23)$$

Démonstration. Pour le premier résultat, un développement de $\|x + y\|_\alpha^\alpha$ entraîne

$$\begin{aligned} \|x + y\|_\alpha^\alpha &\triangleq [x + y, x + y]_\alpha^\alpha \\ &= [x, x]_\alpha + [x, y]_\alpha + [y, x]_\alpha + [y, y]_\alpha \\ &= [x, x]_\alpha + [y, y]_\alpha \\ &= \|x\|_\alpha^\alpha + \|y\|_\alpha^\alpha, \end{aligned}$$

où le passage de la deuxième égalité à la troisième est obtenu grâce à l'indépendance entre x et y de l'équation (11.18).

Le second résultat nécessite l'identification en (9.6) avec $\left\| \int_{\boldsymbol{\theta} \in S^K} |\psi(\boldsymbol{\theta})|^\alpha \chi(d\boldsymbol{\theta}) \right\|_\alpha^\alpha$ et $\int_{\boldsymbol{\theta} \in S^K} |\psi(\boldsymbol{\theta})|^\alpha \Gamma_x(d\boldsymbol{\theta})$. Si on utilise (11.21), il s'ensuit :

$$\left\| \int_{\boldsymbol{\theta} \in S^K} |\psi(\boldsymbol{\theta})|^\alpha \chi(d\boldsymbol{\theta}) \right\|_\alpha^\alpha = \int_{\boldsymbol{\theta} \in S^K} |\psi(\boldsymbol{\theta})|^\alpha \Gamma_x(d\boldsymbol{\theta}),$$

d'où l'égalité annoncée. □

11.3 Problème d'optimisation

11.3.1 Filtre linéaire et existence d'une solution

Considérons un modèle de mélange équivalent à (11.1) mais dans le cas de sources α -stables indépendantes :

$$\begin{cases} \mathbf{x} \triangleq [x_1, \dots, x_K]^\top = \sum_{j=1}^J \mathbf{y}_j \\ \mathbf{y}_j \triangleq [y_{j1}, \dots, y_{jK}]^\top \sim S\alpha S_c^K(\Gamma_j) \end{cases} \quad (11.24)$$

Rappelons que le but est d'extraire \mathbf{y}_j à partir du vecteur observé \mathbf{x} . Soit $\Gamma_{\mathbf{x}} \triangleq \sum_j \Gamma_j$, la mesure spatiale associée à \mathbf{x} , vecteur aléatoire $S\alpha S_c^K$. Appliquons les théorèmes de représentation à $\mathbf{x} \stackrel{d}{=} \int_{\theta \in \mathcal{S}^K} \theta \mathcal{X}(d\theta)$ et $\mathbf{y}_j \stackrel{d}{=} \int_{\theta \in \mathcal{S}^K} \theta \mathcal{Y}_j(d\theta)$, pour tout $j \in \llbracket 1, J \rrbracket$. Nous proposons l'estimateur linéaire suivant :

$$\forall j, k, \hat{y}_{jk} = \langle \mathbf{w}_{jk}, \mathbf{x} \rangle \quad (11.25)$$

avec

$$\sum_j \mathbf{w}_{jk} = \mathbf{e}_k, \quad (11.26)$$

où \mathbf{e}_k représente le $k^{\text{ème}}$ vecteur de la base canonique. Rappelons que la condition (11.26) induit une parfaite reconstruction $\sum_j \hat{y}_{jk} = x_k$. Le problème d'optimisation à résoudre, se faisant selon les \mathbf{w}_{jk} , est donc le suivant :

$$\begin{array}{ll} \text{minimiser} & \sum_j \|y_{jk} - \langle \mathbf{w}_{jk}, \mathbf{x} \rangle\|_\alpha^\alpha \text{ associé à } \mathbf{w}_{jk} \\ \text{sous contrainte} & \sum_j \mathbf{w}_{jk} = \mathbf{e}_k. \end{array} \quad (11.27)$$

De tels travaux dans [128] ont déjà fait part de la minimisation de la norme de covariation $\|y_{jk} - \hat{y}_{jk}\|_\alpha^\alpha$. Il est en effet fait usage de la relation (5.5) afin de résoudre le problème de minimisation. Notre approche exploitera d'avantage le théorème 9.2.1 page 75 de représentation spatiale, faisant intervenir les *spectres spatiaux* et *mesures spatiales* des sources \mathbf{y}_j . Nous pouvons affirmer que la méthode employée ci-dessous est contrainte davantage par sa configuration spatiale que celle présentée dans [128]. Montrons que ce problème sous contrainte admet une unique solution. Pour cela, nous utilisons les conditions de Karush-Khun Tucker (KKT) [23]. Nous remarquons que :

1. Les contraintes $\sum_j \mathbf{w}_{jk} = \mathbf{e}_k$ sont linéaires, donc elles définissent un ensemble convexe.
2. La fonction $\mathbf{w}_{jk} \mapsto \sum_j \|y_{jk} - \langle \mathbf{w}_{jk}, \mathbf{x} \rangle\|_\alpha^\alpha$ est de classe C^1 (différentiable et dont la différentielle est continue) comme somme de fonctions de classe C^1 . Par ailleurs, chaque terme de la somme sur j est une norme. Donc, la fonction est convexe comme somme de fonctions convexes.
3. La fonction $\mathbf{w}_{jk} \mapsto \sum_j \|y_{jk} - \langle \mathbf{w}_{jk}, \mathbf{x} \rangle\|_\alpha^\alpha$ étant la somme de normes est coercive. C'est à dire que :

$$\lim_{\|\mathbf{w}_{jk}\| \rightarrow \infty} \left[\sum_j \|y_{jk} - \langle \mathbf{w}_{jk}, \mathbf{x} \rangle\|_\alpha^\alpha \right] = +\infty. \quad (11.28)$$

Ainsi, les trois conditions précédentes nous permettent d'appliquer le théorème de KKT. Nous en concluons que le problème d'optimisation sous contraintes présentée en (11.27) admet une unique solution, caractérisée par les conditions de KKT.

11.3.2 Résolution du problème

Dans toute cette sous-section, les intégrales de long de la sphère \mathcal{S}^K seront simplement notées " \int ". Exprimons tout d'abord le Lagrangien [23] du problème (11.27) :

$$\mathcal{L}(\{\mathbf{w}_{jk}\}_j, \lambda_k) = \sum_j \|y_{jk} - \langle \mathbf{w}_{jk}, \mathbf{x} \rangle\|_\alpha^\alpha + \alpha \mathfrak{R} \left(\lambda_k^* \left(\sum_j \mathbf{w}_{jk} - \mathbf{e}_k \right) \right), \quad (11.29)$$

où pour tout k , $\lambda_k \in \mathbb{C}^K$ est un multiplicateur de Lagrange. Sans perte de généralité, le facteur α est simplement introduit afin de simplifier les calculs qui suivront. Un développement de la quantité $\|y_{jk} - \hat{y}_{jk}\|_\alpha^\alpha$ entraîne les égalités suivantes :

$$\begin{aligned} & \|y_{jk} - \hat{y}_{jk}\|_\alpha^\alpha \\ &= \left\| y_{jk} - \sum_{j'} \langle \mathbf{w}_{jk}, \mathbf{y}_{j'} \rangle \right\|_\alpha^\alpha \\ &= \left\| \int (\theta_k - \langle \mathbf{w}_{jk}, \boldsymbol{\theta} \rangle) \mathcal{Y}_j(d\boldsymbol{\theta}) \right\|_\alpha^\alpha + \sum_{j' \neq j} \left\| \int \langle \mathbf{w}_{jk}, \boldsymbol{\theta} \rangle \mathcal{Y}_{j'}(d\boldsymbol{\theta}) \right\|_\alpha^\alpha \\ &= \int |\theta_k - \langle \mathbf{w}_{jk}, \boldsymbol{\theta} \rangle|^\alpha \Gamma_j(d\boldsymbol{\theta}) + \sum_{j' \neq j} \int |\langle \mathbf{w}_{jk}, \boldsymbol{\theta} \rangle|^\alpha \Gamma_{j'}(d\boldsymbol{\theta}) \\ &= \int (|\theta_k^* - \langle \boldsymbol{\theta}, \mathbf{w}_{jk} \rangle|^\alpha - |\langle \boldsymbol{\theta}, \mathbf{w}_{jk} \rangle|^\alpha) \Gamma_j(d\boldsymbol{\theta}) + \int |\langle \boldsymbol{\theta}, \mathbf{w}_{jk} \rangle|^\alpha \Gamma_x(d\boldsymbol{\theta}), \end{aligned} \quad (11.30)$$

où le passage de la première égalité à la seconde utilise l'équation (11.22), et celui de la seconde égalité à la troisième utilise l'égalité de l'équation (11.23). La substitution de l'équation (11.30) dans l'équation (11.29) et le calcul du gradient de \mathcal{L} entraînent :

$$\begin{aligned} \nabla_{\mathbf{w}_{jk}} \mathcal{L}(\{\mathbf{w}_{jk}\}_j, \lambda_k) &= -\alpha \int \boldsymbol{\theta} \left[\frac{(\theta_k^* - \langle \boldsymbol{\theta}, \mathbf{w}_{jk} \rangle) |\theta_k^* - \langle \boldsymbol{\theta}, \mathbf{w}_{jk} \rangle|^\alpha}{|\theta_k^* - \langle \boldsymbol{\theta}, \mathbf{w}_{jk} \rangle|} + \frac{(\langle \boldsymbol{\theta}, \mathbf{w}_{jk} \rangle) |\langle \boldsymbol{\theta}, \mathbf{w}_{jk} \rangle|^\alpha}{|\langle \boldsymbol{\theta}, \mathbf{w}_{jk} \rangle|} \right] \Gamma_j(d\boldsymbol{\theta}) \\ &\quad + \alpha \int \boldsymbol{\theta} \frac{(\langle \boldsymbol{\theta}, \mathbf{w}_{jk} \rangle) |\langle \boldsymbol{\theta}, \mathbf{w}_{jk} \rangle|^\alpha}{|\langle \boldsymbol{\theta}, \mathbf{w}_{jk} \rangle|} \Gamma_x(d\boldsymbol{\theta}) + \alpha \lambda_k \\ &= -\alpha \left[\int \boldsymbol{\theta} \left((\theta_k^* - \langle \boldsymbol{\theta}, \mathbf{w}_{jk} \rangle)^{\langle \alpha-1 \rangle} + \langle \boldsymbol{\theta}, \mathbf{w}_{jk} \rangle^{\langle \alpha-1 \rangle} \right) \Gamma_j(d\boldsymbol{\theta}) \right. \\ &\quad \left. - \int \boldsymbol{\theta} \langle \boldsymbol{\theta}, \mathbf{w}_{jk} \rangle^{\langle \alpha-1 \rangle} \Gamma_x(d\boldsymbol{\theta}) + \lambda_k \right], \end{aligned} \quad (11.31)$$

où à nouveau $\forall z \in \mathbb{C}$, $z^{\langle \alpha \rangle} \triangleq z|z|^{\alpha-1}$. En calculant $\nabla_{\mathbf{w}_{jk}} \mathcal{L}(\{\mathbf{w}_{jk}\}_j, \lambda_k) = \mathbf{0}$, nous obtenons pour tout j :

$$\lambda_k = \int \boldsymbol{\theta} \left((\theta_k^* - \langle \boldsymbol{\theta}, \mathbf{w}_{jk} \rangle)^{\langle \alpha-1 \rangle} + \langle \boldsymbol{\theta}, \mathbf{w}_{jk} \rangle^{\langle \alpha-1 \rangle} \right) \Gamma_j(d\boldsymbol{\theta}) - \int \boldsymbol{\theta} \langle \boldsymbol{\theta}, \mathbf{w}_{jk} \rangle^{\langle \alpha-1 \rangle} \Gamma_x(d\boldsymbol{\theta}). \quad (11.32)$$

En introduisant la matrice \mathbf{P}_{jk} , de taille $K \times K$ comme :

$$\mathbf{P}_{jk} = \int \left(\frac{\boldsymbol{\theta} \boldsymbol{\theta}^*}{|\theta_k - \langle \mathbf{w}_{jk}, \boldsymbol{\theta} \rangle|^{2-\alpha}} - \frac{\boldsymbol{\theta} \boldsymbol{\theta}^*}{|\langle \mathbf{w}_{jk}, \boldsymbol{\theta} \rangle|^{2-\alpha}} \right) \Gamma_j(d\boldsymbol{\theta}) + \int \frac{\boldsymbol{\theta} \boldsymbol{\theta}^*}{|\langle \mathbf{w}_{jk}, \boldsymbol{\theta} \rangle|^{2-\alpha}} \Gamma_x(d\boldsymbol{\theta}) \quad (11.33)$$

et le vecteur \mathbf{r}_{jk} de taille K par :

$$\mathbf{r}_{jk} = \int \frac{\theta \theta_k^*}{|\theta_k - \langle \mathbf{w}_{jk}, \theta \rangle|^{2-\alpha}} \Gamma_j(d\theta), \quad (11.34)$$

nous en déduisons la relation linéaire suivante entre le multiplicateur de Lagrange λ_k et \mathbf{w}_{jk} pour tout k

$$\lambda_k = -\mathbf{P}_{jk} \mathbf{w}_{jk} + \mathbf{r}_{jk}. \quad (11.35)$$

Autrement dit, \mathbf{w}_{jk} est un point fixe de l'équation suivante pour tout j, k :

$$\mathbf{w}_{jk} = \mathbf{P}_{jk}^{-1} (\mathbf{r}_{jk} - \lambda_k). \quad (11.36)$$

De même, en vertu de la propriété de reconstruction parfaite, une somme sur l'index j dans (11.36) entraîne que :

$$\lambda_k = \left(\sum_j \mathbf{P}_{jk}^{-1} \right)^{-1} \left(\left(\sum_j \mathbf{P}_{jk}^{-1} \mathbf{r}_{jk} \right) - \mathbf{e}_k \right). \quad (11.37)$$

Il est alors possible d'estimer les \mathbf{w}_{jk} en mettant à jour successivement les équations de (11.33) à (11.37). Cette méthode de *Filtrage via une Minimisation de la Covariation* (α -FMC) est résumée via l'algorithme 7. À nouveau, les calculs d'intégrales appartenant respectivement aux équations (11.33), (11.36) et (11.37) seront calculés, via la discrétisation de l'équation (9.7) :

Algorithme 7 α -FMC : Filtrage via une Minimisation de la Covariation.

1. Entrées

- Observation \mathbf{x} de dimension K
- Partition régulière $\Theta = \{\Theta_1, \dots, \Theta_P\}$ de \mathcal{S}^K
- Exposant caractéristique $\alpha \in]1, 2[$
- Mesures spatiales Γ_j
- Nombre d'itérations pour la mise à jour

2. Initialisation

- $\forall j, k, \mathbf{w}_{jk} = \frac{1}{j} \mathbf{e}_k$
- $\forall j, k$, les entrées $\mathbf{P}_{j,k}$ sont indépendamment tirés selon une distribution gaussienne centrée réduite
- $\forall k, \lambda_k = 0$

3. Mises à jour de \mathbf{P}_{jk} , $\mathbf{w}_{j,k}$ et λ_k

- $\forall j, k$, mettre à jour \mathbf{P}_{jk} comme dans (11.33)
- $\forall k$, mettre à jour λ_k comme dans (11.37)
- $\forall j, k$, mettre à jour \mathbf{w}_{jk} comme dans (11.36)

4. Reconstruction : $\forall j, k, \hat{y}_{jk} = \langle \mathbf{w}_{jk}, \mathbf{x} \rangle$

Intéressons nous au cas « presque » gaussien quand $\alpha \rightarrow 2$.

11.3.3 Cas limite quand $\alpha \rightarrow 2$

En raison de la non-unicité des *mesures spatiales* dans le cas gaussien, la méthode proposée dans la sous-section 11.3.2 n'est valable que pour $\alpha \in]1, 2[$. Nous pouvons néanmoins nous intéresser au

comportement de la méthode de filtrage proposée lorsque $\alpha \rightarrow 2$ dans les mises à jour (11.33) et (11.37). En premier lieu, l'équation (11.33) se transforme comme

$$\forall j, k, \mathbf{P}_{jk} = \mathbf{P} = \int \boldsymbol{\theta} \boldsymbol{\theta}^* \Gamma_x(d\boldsymbol{\theta}), \quad (11.38)$$

où \mathbf{P} devient la matrice de covariance du mélange \mathbf{x} . En effet, si on utilise la fonction caractéristique φ_x représentée via l'équation (9.2), nous avons

$$\begin{aligned} \forall \mathbf{u} \in \mathbb{K}^K, \varphi_x(\mathbf{u}) &= \exp\left(-\int |\langle \mathbf{u}, \boldsymbol{\theta} \rangle|^2 \Gamma_x(d\boldsymbol{\theta})\right) \\ &= \exp\left(-\int \mathbf{u}^* \boldsymbol{\theta} (\mathbf{u}^* \boldsymbol{\theta})^* \Gamma_x(d\boldsymbol{\theta})\right) \\ &= \exp\left(-\mathbf{u}^* \left[\int \boldsymbol{\theta} \boldsymbol{\theta}^* \Gamma_x(d\boldsymbol{\theta})\right] \mathbf{u}\right) \\ &= \exp(-\mathbf{u}^* \mathbf{P} \mathbf{u}), \end{aligned} \quad (11.39)$$

qui est la fonction caractéristique d'un vecteur gaussien de matrice de covariance \mathbf{P} . Par ailleurs, il est clair que $\mathbf{P} = \sum_j \mathbf{P}_j$, où \mathbf{P}_j est la matrice de covariance de la $j^{\text{ème}}$ composante. Contrairement à (11.38), les matrices \mathbf{P}_j sont calculées contre la *mesure spatiale* Γ_j :

$$\mathbf{P}_j = \int \boldsymbol{\theta} \boldsymbol{\theta}^* \Gamma_j(d\boldsymbol{\theta}). \quad (11.40)$$

Nous observons alors que le vecteur \mathbf{r}_{jk} dans (11.34) devient la $k^{\text{ème}}$ colonne de la matrice \mathbf{P}_j . Donc, chaque $\lambda_k = 0$ et

$$\forall j, k, \mathbf{w}_{jk} = \mathbf{P}^{-1} \mathbf{r}_{jk}. \quad (11.41)$$

Finalement, en réinjectant l'équation (11.41) dans l'estimateur $\hat{\mathbf{y}}_{jk} = \langle \mathbf{w}_{jk}, \mathbf{x} \rangle$, nous avons l'estimateur linéaire

$$\forall j, \hat{\mathbf{y}}_j = \mathbf{P}_j \left(\sum_{j'} \mathbf{P}_{j'} \right)^{-1} \mathbf{x} \quad (11.42)$$

correspondant exactement au filtrage de Wiener multicanal. La différence réside dans le fait que (11.42) exploite les *mesures spatiales* Γ_j pour le filtrage. Ceci montre que le filtre proposé dans ce chapitre généralise le filtre de Wiener multicanal dans le cas α -stable. À noter que comme $\alpha < 2$, l'expression $\mathbf{P}_j = \mathbb{E}[\mathbf{y}_j \mathbf{y}_j^*]$ n'existe pas, car il s'agit d'un moment d'ordre 2. Néanmoins, il est très intéressant que (11.38) et (11.40) restent valables et calculables quel que soit $\alpha \in]1, 2[$, mettant en scène un filtrage *ad-hoc*. Par abus de notation, nous appelons cette méthode *α -Filtrage de Wiener Multicanal* (α -FWM).

L'algorithme 8 résume α -FWM. Les intégrales de l'équation (11.40) seront calculées en utilisant la discrétisation esquissée en (9.7) :

Algorithme 8 α -FWM : α -Filtrage de Wiener multicanal.

1. **Entrées**

- Observation \mathbf{x} de dimension K
- Partition régulière $\Theta = \{\Theta_1, \dots, \Theta_P\}$ de \mathcal{S}^K
- Mesures spatiales Γ_j

2. **Estimation des « matrices de covariances »**

- Calcul des \mathbf{P}_j dans (11.40)

3. **Reconstruction** : $\hat{\mathbf{y}}_j = \mathbf{P}_j \left(\sum_{j'} \mathbf{P}_{j'} \right)^{-1} \mathbf{x}$

Chapitre 12

Évaluation des performances

Résumé

Nous comparons les performances des algorithmes α -*Filtrage de Wiener Multicanal* (α -FWM), *Filtrage via l'estimation du Spectre Spatial* (α -FSS) et *Filtrage via la Minimisation de la Covariation* (α -FMC). Toutes les évaluations supposent connues les *mesures spatiales*. En effet, nous souhaitons évaluer les méthodes dans un contexte « oracle » et ainsi dans un premier temps mesurer le « pouvoir séparateur » de chaque algorithme, ses paramètres étant connus. Nous considérons plusieurs scénarios de séparation. Tout d'abord, nous envisageons la séparation de 2 vecteurs α -stables dont les mesures spatiales sont des Von-Mises Fisher avec des modes choisies plus ou moins proches. Plus le recouvrement entre les densités est grand, plus dur sera la séparation. Ensuite, nous évaluons comment se comporte l'erreur en fonction du nombre de composantes. Enfin, 4 vecteurs α -stables avec des *mesures spatiales* davantage complexes sont considérées, dans le but de se rapprocher d'un scénario classique de séparation de 4 sources sonores.

12.1 Méthode de filtrage, hypothèses et score d'erreur

Pour évaluer les travaux de la partie III, nous comparons les trois algorithmes suivants :

- α -FMC L'algorithme 7 résume cette technique de filtrage. Inspirée du filtrage de Wiener minimisant l'*erreur quadratique moyenne* (EQM) entre sources et estimées, cette technique de filtrage linéaire minimise son pendant en α -stable : la norme de covariation. La méthode se résume alors par la mise à jour des paramètres. Nous considérons 50 itérations pour les mises à jour.
- α -FWM Explicité dans la sous-section 11.3.3 page 97 et plus précisément dans l'algorithme 8, cette méthode est un cas particulier de α -FMC, quand $\alpha \rightarrow 2$. Elle est équivalente au filtrage de Wiener multicanal, mais exploite la connaissance des mesures spatiales pour le calcul des covariances.
- α -FSS est décrit dans l'algorithme 6 et dans le chapitre 10 page 81. Cette méthode estime directement le spectre spatial et fournit un algorithme de séparation non linéaire.

Dans toutes les évaluations qui suivent, nous supposons connues les densités spatiales Γ_j . Notons qu'un tel cas d'évaluation *oracle* est parfois considéré dans le cas Gaussien [39, 119]. L'évaluation sera ici concentrée sur des données synthétiques. Nous employons l'algorithme 5 page 76 pour construire des échantillons issus d'une loi $S\alpha S_c^K$. La procédure d'évaluation sera la suivante :

1. Générer des réalisations de chaque vecteur $\mathbf{y}_j \sim S\alpha S_c^K(\Gamma_j)$ grâce à l'algorithme 5.

2. Sommer les réalisations pour obtenir des échantillons d'un vecteur aléatoire $\mathbf{x} \sim S\alpha S_c^K \left(\sum_j \Gamma_j \right)$.
3. Le score comparant les estimées avec les vraies sources générées à la première étape sera calculé.

Tout calcul d'intégral le long de la sphère S_K utilise la technique de discrétisation présentée en (9.7). Cependant, étant donné que les sources sont supposées symétriques (ainsi que la *mesure spatiale* et le *spectre spatial*), nous échantillonons uniquement l'hyper demi-sphère. Le score pour évaluer la performance de chaque méthode sera le calcul empirique de l'*erreur absolue moyenne* (EAM) entre une composante y et son estimée \hat{y} :

$$\text{EAM}(y, \hat{y}) = \frac{\sum_j \mathbb{E}(|y_j - \hat{y}_j|)}{\sum_j \mathbb{E}(|y_j|)}, \quad (12.1)$$

où les espérances seront approximées par des moyennes empiriques. Les mesures spatiales seront prises comme étant des mélanges de distributions de *Von-Mises Fisher* (VMF) comme décrit dans la section 9.2.3 page 77.

12.2 Écart angulaire et pouvoir séparateur

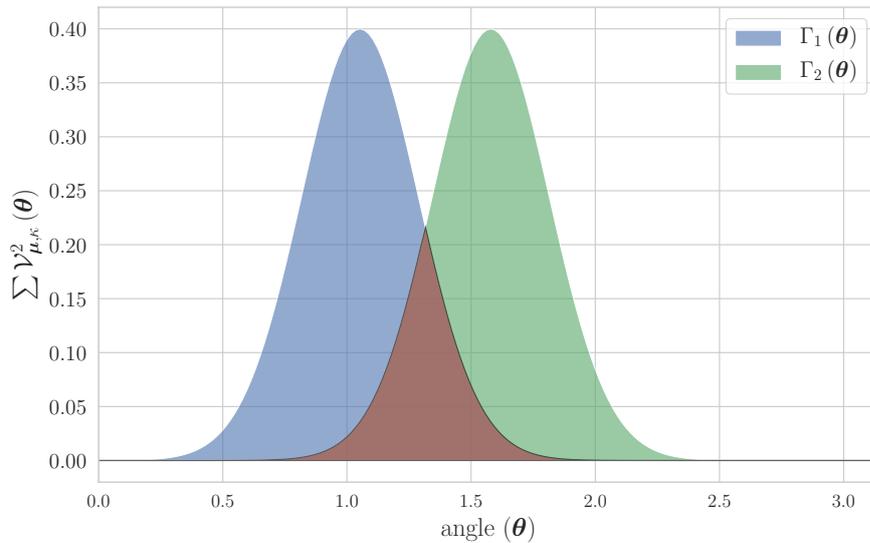


FIGURE 12.1 – Deux densités de Von-Mises Fisher Γ_1, Γ_2 sur le demi-cercle, avec les directions respectives $\mu_1 = \frac{\pi}{3}, \mu_2 = \frac{\pi}{2}$ et le paramètre de concentration $\kappa = 15$. L'aire en rouge indique le chevauchement entre les deux *mesures spatiales*.

Dans cette première évaluation, nous étudions la résolution spatiale de chaque algorithme. Pour ce faire, nous considérons un mélange de $J = 2$ composantes réelles de dimension $K = 2$. Ici, une direction $\theta \in S^K$ peut être vue comme un point du cercle unité.

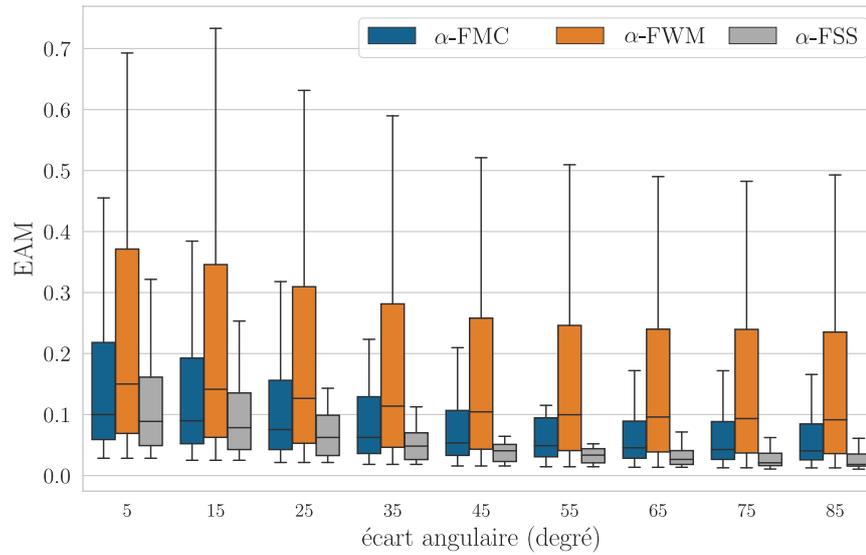


FIGURE 12.2 – Diagramme en boîte des scores d’EAM (plus le score est bas, meilleure est la séparation) des trois algorithmes, en fonction de l’écart angulaire entre les deux *mesures spatiales* avec $\alpha = 1.6$.

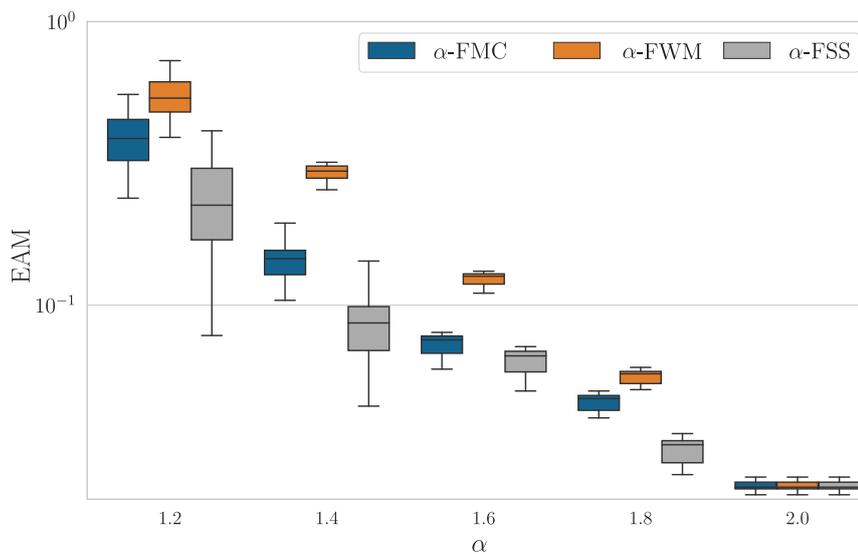


FIGURE 12.3 – Scores EAM en fonction d’ α , pour un écart angulaire de 25 degrés entre les sources.

Chaque composante y_j a comme *mesure spatiale* une distribution de VMF de direction μ_j et de concentration $\kappa = 15$, de telle manière que $\Gamma_j = \mathcal{V}_{\mu_j, \kappa}^2$. Selon le choix des directions μ_j , un chevauchement entre les mesures spatiales Γ_j peut être engendré (voir Fig. 12.1). Nous étudions les performances pour $\alpha \in \{1.2, 1.4, 1.6, 1.8\}$. Les directions μ_j sont quant à elles séparées de $\{5, 15, \dots, 85\}$ degrés. La première direction μ_1 est positionnée aléatoirement sur le demi-cercle.

Les algorithmes α -FMC et α -FSS sont alors exécutés avec une partition Θ composée de $P = 180$ arcs de cercles. Pour chaque configuration de μ_j , l'EAM est évaluée en prenant en compte $N = 2000$ réalisations indépendantes. Nous répétons l'expérience 100 fois pour chaque configuration.

Les EAM pour $\alpha = 1.6$ sont reportées dans la Fig. 12.2. Nous pouvons observer que la méthode non linéaire α -FSS admet une meilleure performance que les autres méthodes pour chaque écart angulaire entre les sources.

Dans la Fig. 12.3, nous avons relevé les différents scores d'erreur en fonction de α . Nous avons fixé l'écart angulaire entre les sources à 25 degrés. Ici encore, la méthode α -FSS semble plus robuste quand α diminue. Cependant, toutes les méthodes restent relativement compétitives. Notons que dans le cas $\alpha = 2$, les trois méthodes donnent des scores équivalents. En effet, les trois méthodes coïncident en théorie quand $\alpha = 2$.

12.3 Qualité de séparation en fonction du nombre de composantes

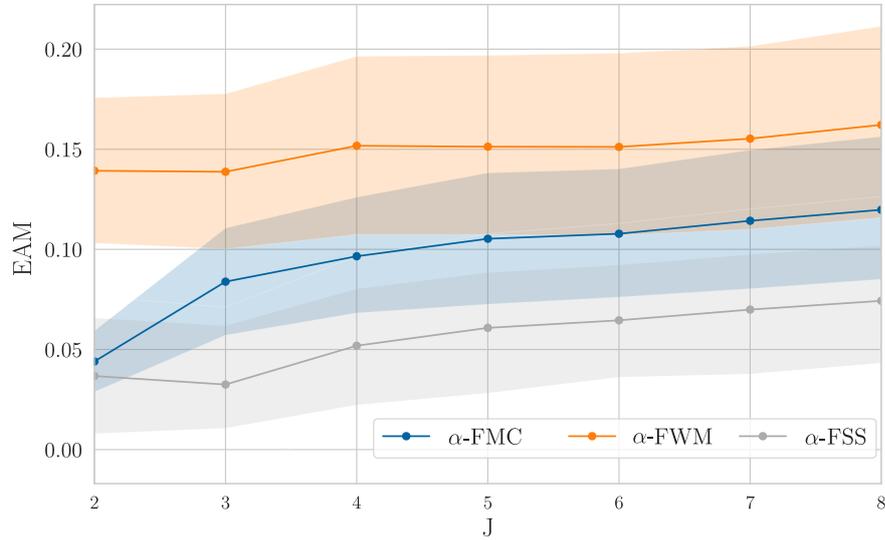


FIGURE 12.4 – EAM moyennée sur $\alpha \in [1.2, 2]$, en fonction du nombre de composantes. Les lignes représentent la médiane tandis que les aires transparentes décrivent l'écart inter-quartiles de chaque méthode.

Dans cette seconde expérience, nous évaluons nos méthodes de filtrage dans le cas complexe où $K = 2$. Afin de partitionner convenablement l'hyper demi-sphère complexe, nous rappelons les *coordonnées de Hopf* [76] pour $\mathcal{S}_{\mathbb{C}}^2$:

$$\begin{aligned} \left[0, \frac{\pi}{2}\right] \times [-\pi, \pi[\times [-\pi, \pi[&\rightarrow \mathcal{S}_{\mathbb{C}}^2 \\ (\theta, \phi_1, \phi_2) &\mapsto (\sin \theta e^{i\phi_1}, \cos \theta e^{i\phi_2}). \end{aligned} \quad (12.2)$$

Les *coordonnées de Hopf* pour l'hyper demi-sphère consistent à poser $\phi_2 = 0$ [107, p. 156] (ou $\phi_2 = \pi$) :

$$\begin{aligned} \left[0, \frac{\pi}{2}\right] \times [-\pi, \pi[&\rightarrow \mathcal{S}_{\mathbb{C}}^2 \\ (\theta, \phi) &\mapsto (\sin \theta e^{i\phi}, \cos \theta). \end{aligned} \quad (12.3)$$

-	α -FWM	α -FMC	α -FSS
$J = 2$	0.02	0.18	1.02
$J = 3$	0.02	0.20	1.11
$J = 5$	0.02	0.45	1.12
$J = 8$	0.02	0.65	1.16

TABLE 12.1 – Temps de calcul (en seconde) pour chaque méthode de filtrage, en fonction du nombre de composantes J à séparer.

Nous considérons $P = 400$ valeurs de la partition Θ , en utilisant le paramétrage particulier présenté dans l'équation (12.3). Nous employons $J \in \{2, \dots, 8\}$ composantes, représentées chacune par une direction μ_j et un paramètre de concentration κ_j distincts ou non. L'exposant caractéristique α sera à nouveau pris dans $\{1.2, \dots, 2\}$. Nous tirons $N = 2000$ échantillons indépendants. À nouveau, 100 expériences indépendantes pour chaque paramétrage de α et J sont effectuées, soit 3500 expériences au total. Les résultats de la Fig. 12.4 montrent que la méthode α -FSS donne de nouveaux de meilleurs résultats. De plus, son EAM évolue plus lentement que les méthodes α -FMC et α -FWM.

Soulignons qu'en dépit d'un meilleur pouvoir séparateur, la méthode α -FSS demeure la plus lente. En effet, la complexité algorithmique pour α -FSS est en $O\left(P^4 (P')^4 K^5 N^3 + JP^3 K^2 N\right)$, où P' est le nombre d'échantillons nécessaires pour la discrétisation des intégrales dans (10.17). À titre de comparaison, la complexité algorithmique d' α -FMC est de $O(IJK^4 P^4 N)$ où I est le nombre d'itérations réalisées dans l'algorithme 7. Nous avons reporté dans le Tableau 12.1 les temps de calculs moyens pour $N = 2000$ échantillons. Pour information, notre implémentation a été effectuée en Python sur un ordinateur portable avec un CPU i7-4810MQ et 32 GB de RAM. Finalement, α -FWM est l'algorithme le plus rapide, suivi de α -FMC et α -FSS. En analysant de plus près la complexité algorithmique, nous pouvons observer que le calcul de g_χ dans α -FSS est la cause de ce temps de calcul. C'est pourquoi augmenter le nombre de composantes J n'augmente que très légèrement la complexité algorithmique de α -FSS.

12.4 Score de séparation en fonction de la complexité du modèle spatial

Nous considérons maintenant que chaque composante n'est plus décrite par une VMF comme provenant principalement d'une direction d'arrivée. Au contraire, nous supposons à présent chaque composante comme ayant une configuration spatiale complexe. En pratique, nous prenons chaque Γ_j comme un mélange de C VMF comme suit :

$$\forall j, \Gamma_j(d\theta) = \sum_c w_{jc} \mathcal{V}_{\mu_{jc}, \kappa_{jc}}(d\theta), \quad (12.4)$$

où $w_{jc} \in [0, 1]$ pondère l'importance de $\mathcal{V}_{\mu_{jc}, \kappa_{jc}}$ dans Γ_j tel que $\sum_c w_{jc} = 1$. Nous sélectionnons à nouveau $\alpha \in \{1.2, 1.4, \dots, 2\}$ et $N = 2000$ échantillons indépendants de $J = 4$ sources. Les *mesures spatiales* sont le mélange de $C = 2, 3, 4$ distributions de VMF, avec des paramètres w_{jc}, μ_{jc} et κ_{jc} tirés aléatoirement (d'une distribution gaussienne centrée réduite) pour 100 expériences. Nous nous plaçons dans le cas réel avec $K = 2$ et $P = 360$ arcs de cercles.

Comme le montre la Fig. 12.5, la méthode α -FSS donne à nouveau de meilleurs résultats que les autres algorithmes. Cette légère différence peut s'expliquer par le fait qu' α -FSS est un filtrage non-linéaire, et donc intègre des configurations spatiales plus complexes dans son modèle.

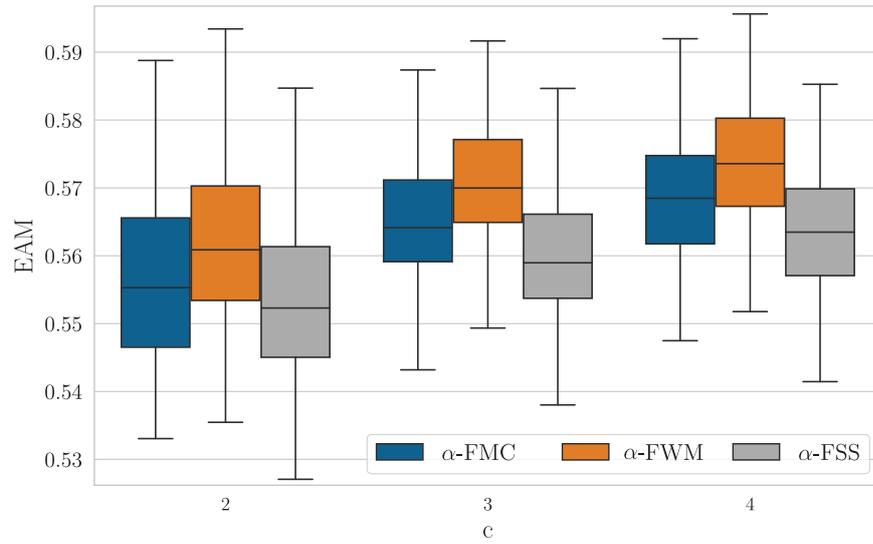


FIGURE 12.5 – Diagramme en boîte de l'EAM pour $\alpha = 1.4$ en fonction du nombre de directions d'arrivées C , pour chaque composante.

Conclusion de la partie III

Nous avons présenté une nouvelle représentation spatiale d'un vecteur α -stable dans le but d'établir une nouvelle méthode de filtrage. Ce théorème de représentation propose une représentation stochastique faisant intervenir un opérateur aléatoire qui permet de générer un vecteur $S_\alpha S_c^K$ dans l'espace de dimension K associé. De cette représentation découlent deux méthodes de filtrage multivarié. La première, présentée au chapitre 10 page 81 et intitulée *Filtrage via l'estimation du Spectre Spatial* (α -FSS) est une estimation directe de cet opérateur aléatoire appelé *spectre spatial*. La seconde, introduite au chapitre 11 page 91 et appelée *Filtrage via une Minimisation de la Covariation* (α -FMC) utilise cette même représentation mais en considérant un problème d'optimisation visant à minimiser une norme définie sur mesure pour les distributions α -stables et appelée norme de covariation. Nous avons d'ailleurs montré que α -FMC tend vers un filtrage de Wiener quand $\alpha \rightarrow 2$ et appelé *α -Filtrage de Wiener Multicanal* (α -FWM). Comparé au α -FWM, le pouvoir séparateur de α -FMC est en moyenne plus important en présence d'un mélange plus complexe. Cependant, les meilleurs scores en termes de séparation sont obtenus avec la méthode α -FSS.

En outre, la méthode α -FSS souffre d'un temps de calcul élevé mais évoluant très peu en fonction du nombre de composantes à séparer. C'est pourquoi une perspective serait de proposer de nouvelles méthodes d'estimations du spectre spatial, pour obtenir un algorithme performant et rapide. À terme, nous souhaitons que de telles méthodes comme α -FSS soient appliquées dans différents cas en traitement du signal (séparation en imagerie hyperspectrale, débruitage audio ou encore séparation de sources sonores).

Quatrième partie

Modèles hybrides et applications au rehaussement de la parole

Chapitre 13

Modèles multi-alphas

Résumé

Les parties précédentes supposaient que l'exposant caractéristique α était le même pour chaque composante d'un mélange. Nous présentons ici comment un modèle dont les sources admettent des exposants caractéristiques différents peut être employé. En particulier, nous mettons en œuvre ce choix dans le cas du débruitage. Deux modèles hybrides issus de distributions α -stables dites *multivariées stables à contour elliptique* (MSCE) sont décrits. Plus précisément, nous utilisons le fait qu'une distribution MSCE est conditionnellement gaussienne. Comme nous le montrons, ce modèle hybride engendre un filtrage de Wiener dit *paramétrique*. Ce faisant, nous revisitons des travaux anciens en débruitage, tels que ceux menés par Ephraim & Malah [41, 42].

13.1 Contexte

Le domaine du débruitage consiste à extraire d'un signal observé bruité un signal souhaité (e.g. signal de parole, instrument de musique etc.). Il est courant de supposer que pour tout instant t , l'observation $\mathbf{x}(t) \in \mathbb{R}^K$ multicanal de dimension K est la somme du signal de parole $\mathbf{s}(t) \in \mathbb{R}^K$ et d'un bruit additif $\mathbf{n}(t) \in \mathbb{R}^K$:

$$\forall t, \mathbf{x}(t) = \mathbf{s}(t) + \mathbf{n}(t). \quad (13.1)$$

L'estimation du bruit est en général faite grâce à un *détecteur d'activité vocale* (DAV), qui permet d'isoler des parties où seul le bruit est présent et où ses caractéristiques peuvent donc être estimées. Ensuite, étant donné un modèle de bruit, le second aspect consiste à filtrer le signal bruité \mathbf{x} . Une approche historique est la *soustraction spectrale* [20, 22, 115, 114, 130], dont le grand intérêt est de ne pas nécessiter de modèle pour la partie voix, mais juste pour le bruit. La Fig. 13.1 résume le procédé général de débruitage que nous emploierons. Un tel schéma a déjà été présenté dans des travaux comme [162], pouvant se référer à une méthode dite d'*espérance-maximisation* (EM) comme décrite dans la sous-section 4.2.1 page 30 pour les distributions α -stables. À nouveau, chaque composante est étudiée dans le domaine d'une *transformée de Fourier à court-terme* (TFCT). Pour commencer, le modèle de mélange se résume via des tenseurs respectivement pour l'observation, la parole et le bruit $\mathbf{X}, \mathbf{S}, \mathbf{N} \in \mathbb{C}^{F \times T \times K}$, où F est le nombre de bandes fréquentielles, T le nombre de trames temporelles et K le nombre de microphones. En résumé, nous avons :

$$\mathbf{X} = \mathbf{S} + \mathbf{N} \quad (13.2)$$

ou encore, pour un point temps-fréquence (TF) (f, t) , la relation suivante entre l'observation $\mathbf{x}(f, t)$, la parole $\mathbf{s}(f, t)$ et le bruit $\mathbf{n}(f, t)$

$$\mathbf{x}(f, t) = \mathbf{s}(f, t) + \mathbf{n}(f, t). \quad (13.3)$$

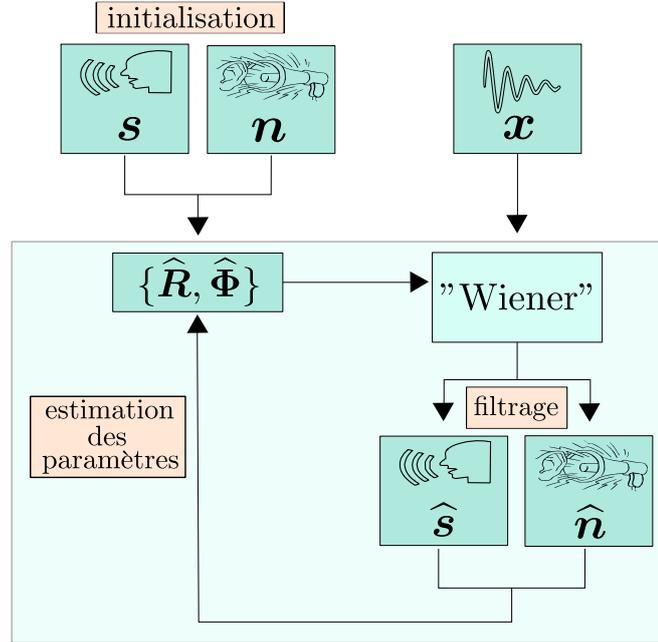


FIGURE 13.1 – Schéma général des algorithmes de débruitage.

Comme dans le cas des précédents travaux, un modèle probabiliste pour chaque point TF peut être imposé. Inspiré de [39, 162], nous supposons que pour tout point TF (f, t) , chaque vecteur sera une distribution conditionnellement gaussienne complexe et indépendante des autres points TF :

$$s(f, t) | \phi_s(f, t) \sim \mathcal{N}_c(\phi_s(f, t) \mathbf{R}_s(f)) \quad (13.4)$$

$$n(f, t) | \phi_n(f, t) \sim \mathcal{N}_c(\phi_n(f, t) \mathbf{R}_n(f)) \quad (13.5)$$

où un aspect majeur de notre travail est de choisir pour les variables $\phi_s(f, t), \phi_n(f, t)$ un modèle α -stable particulier que nous décrivons par la suite. $\mathbf{R}_s(f), \mathbf{R}_n(f)$ sont des matrices dites de *dispersion spatiale*. L'intérêt d'un tel modèle gaussien est de mener à un filtrage de Wiener classique, afin de séparer chaque composante vectorielle.

Pour résumer ce modèle de séparation, nous estimons les paramètres $\Phi \triangleq \{\phi_s(f, t), \phi_n(f, t)\}_{f, t}$ et $\mathbf{R} \triangleq \{\mathbf{R}_s(f), \mathbf{R}_n(f)\}_f$ à partir des observations x . Grâce aux paramètres estimés, nous proposons une méthode de séparation via un filtrage de Wiener paramétrique pour extraire les tenseurs estimés de la parole et du bruit notés respectivement \hat{S} et \hat{N} . Nous répétons alors le procédé d'estimation des paramètres mais en utilisant cette fois-ci \hat{S} et \hat{N} .

Dans la littérature, les variables ϕ_s, ϕ_n sont appelés des distributions *a priori* du modèle [50, 112]. Par exemple dans [104], les distributions ϕ_s, ϕ_n suivent une loi de gamma inversé et, associée à un modèle de Student t, sont utilisées pour la séparation de sources sonores. Également dans [186], il est décrit un modèle de Laplace et des distributions *a priori* suivant une loi exponentielle sont considérées pour le rehaussement de la parole. Dans ce chapitre, nous allons introduire un modèle α -stable faisant intervenir des distributions *a priori* particulières : les loi positives α -stables.

13.2 Modèles de mélange monocanal

13.2.1 Exemples de filtre de Wiener paramétrique

Avant d'explicitier le modèle de mélange multi-alphas monocanal, nous allons motiver un tel choix. Nous commençons par définir un filtre de Wiener paramétrique. Reprenons le modèle de mélange pour tout (f, t) , $x(f, t) = s(f, t) + n(f, t)$ introduit en 13.3. Les moments d'ordre α seront supposés exister et notés pour tout (f, t) , $\sigma_x^\alpha(f, t)$, $\sigma_s^\alpha(f, t)$ et $\sigma_n^\alpha(f, t)$ respectivement pour l'observation $x(f, t)$, la parole $s(f, t)$ et le bruit $n(f, t)$. Un filtrage de Wiener paramétrique sera alors une estimation de $s(f, t)$ noté $\hat{s}(f, t)$ tel que :

$$\hat{s}(f, t) = \frac{\sigma_s^\alpha(f, t)}{\sigma_s^\alpha(f, t) + k\sigma_n^\alpha(f, t)} x(f, t). \quad (13.6)$$

avec k une constante pouvant être fixée heuristiquement [22] ou évaluée via un modèle *a priori* sur le bruit [41, 42]. Durant les années 70-80, plusieurs méthodes conduisant à des modèles de débruitage ont fait surface [41, 42, 114, 130, 113]. Les premiers travaux employaient une technique de soustraction d'énergie du bruit, en appliquant au préalable une TFCT. Considérons le modèle TF présenté dans la section 13.1 à l'équation (13.3). En supposant que les moments d'ordre 2 existent, nous posons alors pour un signal v (le bruit ou le signal cible)

$$\forall (f, t), \sigma_v^2(f, t) = \mathbb{E}(|v(f, t)|^2) \quad (13.7)$$

comme étant la *densité spectrale de puissance* (DSP) du processus, supposé localement stationnaire au sens large. Nous supposons en outre que :

$$\sigma_x^2(f, t) = \sigma_s^2(f, t) + \sigma_n^2(f, t). \quad (13.8)$$

Une estimation $\hat{\sigma}_n^2(f, t)$ de $\sigma_n(f, t)$ est alors calculée sur les segments temporels de $x(t)$ où le signal de parole est supposé absent. De plus, nous admettons que le bruit seul dure assez longtemps pour qu'un *détecteur d'activité vocale* (DAV) repère ses segments. L'amplitude spectrale à court-terme de la parole, qui vaut $|s(f, t)|$ pour un certain (f, t) , est alors estimée grâce aux $\hat{\sigma}_n(f, t)$. Cette méthode, appelée *soustraction spectrale* [114], se résume alors par la relation suivante entre l'estimation $|\hat{s}(f, t)|^2$ de $|s(f, t)|^2$ et $|x(f, t)|$, $\sigma_n(f, t)$:

$$\forall (f, t), |\hat{s}(f, t)|^2 = \sqrt{||x(f, t)|^2 - \sigma_n^2(f, t)|}. \quad (13.9)$$

Elle dérive directement de 13.8 et de l'approximation $|x(f, t)|^2 \simeq |s(f, t)|^2 + |n(f, t)|^2$. Parallèlement, plusieurs méthodes « heuristiques » comme dans [22], ont montré que considérer une soustraction spectrale de la forme

$$\forall (f, t), |\hat{s}(f, t)| = (|x(f, t)|^\alpha - \hat{\sigma}_n^\alpha(f, t))^{1/\alpha} \quad (13.10)$$

avec $\alpha > 0$ améliorerait considérablement la qualité de débruitage, comparée à une *soustraction spectrale de puissance* en (13.9) quand $\alpha = 2$. Par ailleurs, d'autres méthodes dans [41, 42] proposent d'étendre (13.10) mais dans un cadre plus général :

$$\forall (f, t), |\hat{s}(f, t)|^\alpha = |x(f, t)|^\alpha - k\hat{\sigma}_n^\alpha(f, t), \quad (13.11)$$

où $0 < k < 1$ est déterminé à l'aide de connaissances *a priori* sur $\hat{\sigma}_n(f, t)$. L'estimation de la phase de la parole s est estimée en employant uniquement la phase de l'observation x . En somme, ces soustractions spectrales en (13.10) et (13.11) peuvent être vues comme des *filtrages de Wiener paramétrique* (FWP). En effet, rappelons que le FWP en (13.6) avec $k = 1$ et $\alpha = 2$ est donné comme suit :

$$\forall (f, t), \hat{s}(f, t) = \frac{\sigma_s^2(f, t)}{\sigma_s^2(f, t) + \sigma_n^2(f, t)} x(f, t). \quad (13.12)$$

De son côté, la *soustraction spectrale de puissance* ($\alpha = 2$ dans (13.10)) est justifiable en théorie par :

1. la décorrélation entre le bruit et la parole. Autrement dit, la somme des DSPs des sources est égale à la DSP du mélange ;
2. le fait que $|\hat{s}(f, t)|^2$ est l'estimateur du maximum de vraisemblance de $\sigma_s^2(f, t)$, sous hypothèse que les $s(f, t) \sim \mathcal{N}(\sigma_s^2(f, t))$ soient indépendants pour tout f, t .

Comme présentée dans la sous-section 3.1.2 page 24, cette théorie conduit naturellement à considérer (13.12) et à prouver qu'il existe un lien entre les techniques de soustractions spectrales et le filtrage de Wiener, méthode basée sur une théorie probabiliste solide. Dans le cas où $\alpha \in]0, 2]$ et $k = 1$, les travaux en [118] prouvent qu'en théorie l'équation (13.11) peut être expliquée en considérant des processus α -harmonisables. Cela implique un modèle de variables $S\alpha S_c$ associé à leur densités spectrales de puissance fractionnaires $\sigma_s^\alpha(f, t)$ et $\sigma_n^\alpha(f, t)$. De plus, nous avons vu qu'il en résulte un filtrage α -Wiener adapté :

$$\forall(f, t), \hat{s}(f, t) = \frac{\sigma_s^\alpha(f, t)}{\sigma_s^\alpha(f, t) + \sigma_n^\alpha(f, t)} x(f, t). \quad (13.13)$$

13.2.2 Modèle de filtrage α -stable monophonique hybride

Nous reprenons dans cette sous-section l'explication théorique que nous avons fourni dans [58], comparable à celle de (13.12) et (13.13) décrivant théoriquement le modèle de soustraction spectrale en (13.11). Nous justifions par ailleurs le choix de la constante k . En d'autres termes, nous expliquons de quel modèle probabiliste le FWP suivant

$$\forall(f, t), \hat{s}(f, t) = \frac{\sigma_s^2(f, t)}{\sigma_s^2(f, t) + k\sigma_n^2(f, t)} x(f, t) \quad (13.14)$$

découle. Nous montrons que le FWP de l'équation (13.14) peut être expliqué en considérant le modèle conditionnellement gaussien en (13.4) et (13.5). En effet, supposons que pour chaque point TF nous avons le modèle symétrique isotrope α -stable suivant :

$$\forall(f, t), s(f, t) \sim S\alpha_s S_c(\sigma_s(f, t)) \quad (13.15)$$

$$\forall(f, t), n(f, t) \sim S\alpha_n S_c(\sigma_n(f, t)) \quad (13.16)$$

et que le mélange soit la somme des deux sources, comme explicité dans (13.3). Si $\alpha_s \neq \alpha_n$, nous sommes alors en présence d'un *modèle hybride*, dans le sens où les distributions ne sont pas identiques. Par ailleurs pour tout (f, t) , le mélange $x(f, t)$ n'est pas nécessairement une distribution α -stable. Comparé au cas où $\alpha_s = \alpha_n$ dans [118], on ne peut pas obtenir directement un filtrage α -Wiener. Cependant, en utilisant la propriété 2.1.4 page 15, les équations (13.15) et (13.16) induisent le modèle conditionnellement gaussien suivant pour tout (f, t) :

$$s(f, t) | \phi_s(f, t) \sim \mathcal{N}_c\left(\phi_s(f, t) \sigma_s^2(f, t)\right), \quad (13.17)$$

$$n(f, t) | \phi_n(f, t) \sim \mathcal{N}_c\left(\phi_n(f, t) \sigma_n^2(f, t)\right) \quad (13.18)$$

où

$$\phi_s(f, t) \sim \mathcal{P}_{\alpha_s} S\left(2 \left[\cos\left(\frac{\pi\alpha_s}{4}\right)\right]^{2/\alpha_s}\right) \quad (13.19)$$

$$\phi_n(f, t) \sim \mathcal{P}_{\alpha_n} S\left(2 \left[\cos\left(\frac{\pi\alpha_n}{4}\right)\right]^{2/\alpha_n}\right). \quad (13.20)$$

Un tel modèle conditionnel a déjà été exploité dans [162]. Cependant, la source et le bruit avaient le même exposant caractéristique α . Les variables définies par l'équation (13.19) sont appelées *variables*

impulsionnelles. En traitement du signal audio, elles peuvent être interprétées comme mesurant l'activation du spectrogramme en chaque point TF : plus α_s sera petit, et plus $\phi_s(f, t)$ est susceptible de prendre des valeurs importantes. Le modèle de mélange $x(f, t) = s(f, t) + n(f, t)$ est finalement le suivant :

$$\forall (f, t), x(f, t) | \phi_n(f, t), \phi_s(f, t) \sim \mathcal{N}_c \left(\phi_s(f, t) \sigma_s^2(f, t) + \phi_n(f, t) \sigma_n^2(f, t) \right). \quad (13.21)$$

Supposons que les paramètres $\Phi = \{\phi_s(f, t), \phi_n(f, t)\}_{f, t}$ et $\Sigma = \{\sigma_s(f, t), \sigma_n(f, t)\}_{f, t}$ du modèle de mélange en (13.21) soient connus. En appliquant la propriété 3.1.1 page 24 pour tout (f, t) , nous avons la distribution *a posteriori* de $s(f, t)$

$$s(f, t) | x(f, t), \Phi, \Sigma \sim \mathcal{N}_c \left(\mu_{s|x, \Phi, \Sigma}(f, t), c_{s|x, \Phi, \Sigma}(f, t) \right) \quad (13.22)$$

où

$$\hat{s}(f, t) \triangleq \mu_{s|x, \Phi, \Sigma}(f, t) = \frac{\phi_s(f, t) \sigma_s^2(f, t)}{\phi_s(f, t) \sigma_s^2(f, t) + \phi_n(f, t) \sigma_n^2(f, t)} x(f, t) \quad (13.23)$$

$$c_{s|x, \Phi, \Sigma}(f, t) = \frac{\phi_s(f, t) \phi_n(f, t) \sigma_s^2(f, t) \sigma_n^2(f, t)}{\phi_s(f, t) \sigma_s^2(f, t) + \phi_n(f, t) \sigma_n^2(f, t)}. \quad (13.24)$$

Comparons le filtrage (13.14) et l'équation (13.23). Il apparaît que la constante k en (13.14) dépend ici de (f, t) et vérifie

$$k(f, t) = \frac{\phi_n(f, t)}{\phi_s(f, t)}. \quad (13.25)$$

Nous verrons comment (13.25) est estimée dans le chapitre 14. Ajoutons à cela que (13.23) décrit des paramètres d'échelle Σ ne dépendant pas des exposants caractéristiques. Il pourrait être employé les équations (13.17) et (13.18) afin d'obtenir (13.14). Nous insistons sur l'interprétation des paramètres comme suit : la structure prédictible est contenue dans Σ , tandis que les *variables impulsionnelles* $\phi_v(f, t)$ contiennent uniquement l'information d'impulsivité du signal, au sens des écarts par rapport au modèle.

13.3 Modèle de mélange multicanal

13.3.1 Distribution multivariée α -stable à contour elliptique

Dans la proposition 2.1.4 page 15, une variable $x \sim S\alpha S_c(\sigma_x^\alpha)$ peut être vue comme conditionnellement gaussienne isotrope de variance $\phi_x \sigma_x^\alpha$ où ϕ_x est une variable positive α -stable ($\mathcal{P}\alpha S$). Cette logique peut être généralisée au cas vectoriel avec les distributions suivantes :

Définition 13.3.1 (Distribution multivariée α -stable à contour elliptique). *Nous disons qu'un vecteur aléatoire \mathbf{x} de dimension K suit une loi (symétrique) multivariée α -stable à contour elliptique (α -MSCE) [106, 137], si sa fonction caractéristique s'écrit comme suit :*

$$\forall \mathbf{u} \in \mathbb{C}^K, \varphi_{\mathbf{x}}(\mathbf{u}) = \exp \left(-2^{-\alpha/2} (\mathbf{u}^* \mathbf{R}_{\mathbf{x}} \mathbf{u})^{\alpha/2} \right), \quad (13.26)$$

avec $\mathbf{R}_{\mathbf{x}}$ une matrice hermitienne définie positive appelé matrice de dispersion. Nous noterons alors $\mathbf{x} \sim \mathcal{E}\alpha S^K(\mathbf{R}_{\mathbf{x}})$.

La Fig. 13.2 montre un exemple de distribution α -MSCE. Nous attirons l'attention du lecteur sur les propriétés suivantes, dont la seconde découle naturellement du fait qu'une α -MSCE est une distribution α -stable particulière :

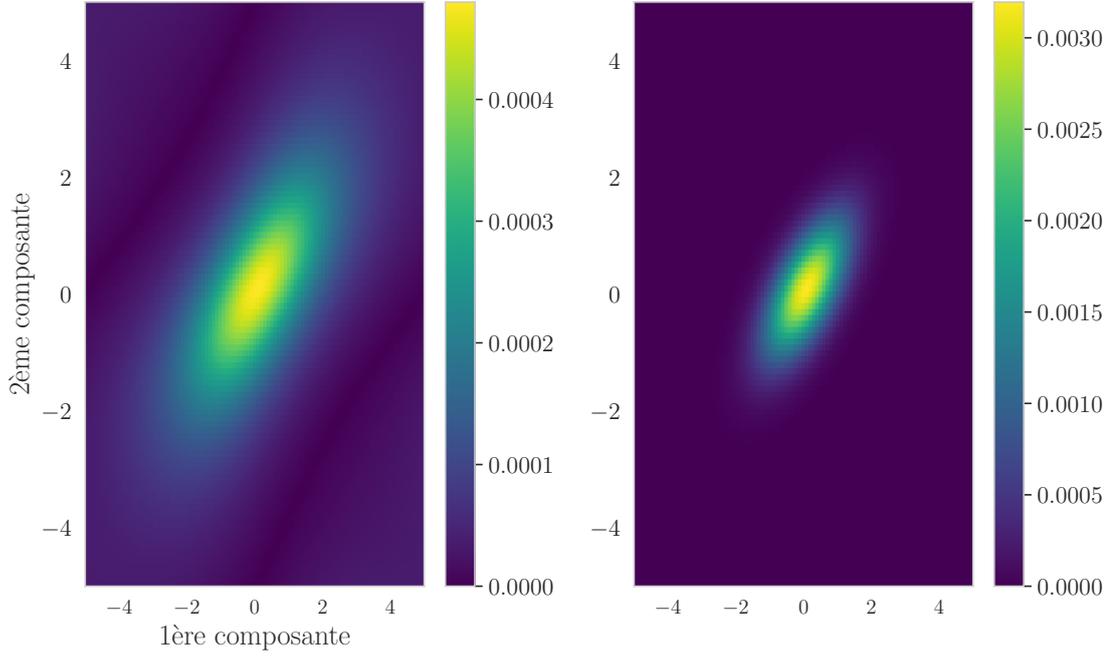


FIGURE 13.2 – Amplitudes de densités de probabilités de distributions α -MSCE (symétrique), pour $\alpha = 1.1$ à gauche et $\alpha = 2$ à droite.

1. Si $\mathbf{x} \sim \mathcal{E}\alpha S^K(\mathbf{R}_x)$ et $\mathbf{y} \sim \mathcal{E}\alpha S^K(\mathbf{R}_y)$ sont indépendants, alors $\mathbf{x} + \mathbf{y} \not\sim \mathcal{E}\alpha S^K(\mathbf{R}_x + \mathbf{R}_y)$ où $\not\sim$ signifie « ne suit pas la même loi » ;
2. Si $\mathbf{x} \sim \mathcal{E}\alpha S^K(\mathbf{R}_x)$ alors pour tout $a \in \mathbb{R}^*$, $a\mathbf{x} \sim \mathcal{E}\alpha S^K(|a|^\alpha \mathbf{R}_x)$.

Plus généralement, la somme de deux vecteurs α -MSCE n'est pas α -MSCE, sauf dans le cas Gaussien. La méthode la plus générale pour sommer les paramètres d'une distribution α -stable reste l'utilisation des *mesures spatiales* introduites dans la partie III. Néanmoins, il est intéressant de remarquer que dans le cas univarié, une variable $x \sim \mathcal{E}\alpha S^1(r_x)$ est en réalité une distribution $S\alpha S_c$. En effet, si nous calculons sa fonction caractéristique $\varphi_{e^{i\phi}x}$ pour $\phi \in [0, 2\pi[$, nous obtenons :

$$\forall u \in \mathbb{C}, \varphi_{e^{i\phi}x}(u) = \exp\left(-2^{-\alpha/2} \left(e^{-i\phi} u^* r_x e^{i\phi} u\right)^{\alpha/2}\right) = \varphi_x(u). \quad (13.27)$$

Donc $x \sim S\alpha S_c(\sigma_x^\alpha)$ où $\sigma_x = \sqrt{\frac{r_x}{2}}$. Comme dans le cas des distributions $S\alpha S_c$, il a été démontré qu'un vecteur α -MSCE peut être considéré comme conditionnellement gaussien [153] :

Proposition 13.3.1. *Un vecteur $\mathbf{x} \sim \mathcal{E}\alpha S^K(\mathbf{R}_x)$ est α -MSCE si et seulement s'il existe un vecteur \mathbf{G} gaussien de moyenne nulle et une variable $\phi \sim \mathcal{P}\alpha S\left(2\left[\cos\left(\frac{\pi\alpha}{4}\right)\right]^{2/\alpha}\right)$ (voir définition 2.1.6 page 15) indépendante de \mathbf{G} tels que :*

$$\mathbf{x} = \sqrt{2\phi}\mathbf{G}. \quad (13.28)$$

Nous notons alors $\mathbf{x} | \phi \sim \mathcal{N}_c^K(\phi\mathbf{R}_x)$.

La distribution α -MSCE peut donc être interprétée comme gaussienne dont la covariance peut être amplifiée de manière aléatoire. Cette perturbation est régie par une variable positive ϕ que nous appelons à nouveau *variable impulsionnelle* [58, 162]. En pratique, la covariance peut être infinie.

13.3.2 Modèle de filtrage multicanal : gaussien + α -stable

Le modèle hybride se rapportant aux travaux de [59] sera interprété autrement que celui présenté dans le cas monocanal. Comme décrit dans [161], nous pouvons supposer que la composante de bruit additif dans un modèle de mélange présente un caractère plus impulsif et non-stationnaire que le signal de parole. En somme, il est nécessaire de contrôler le bruit via un modèle à queue lourde permettant de supposer que cette composante peut être extrêmement permissive. La parole quant à elle sera supposée ne pas trop s'éloigner du modèle qui lui est imposé. Pour ce faire, un modèle Gaussien comprenant de nombreux paramètres (*e.g.* matrices de covariance pas forcément diagonale) sera choisi. Nous avons ainsi un modèle de parole réaliste et ne s'éloignant pas du modèle. Nous choisissons donc d'imposer un modèle gaussien à la source. Ce modèle avec un *bruit résiduel* est donc différent de la modélisation proposée dans la sous-section 13.2.1. Pour résumer, le modèle probabiliste que nous considérons suppose

- chaque point TF (f, t) du *bruit résiduel* $\mathbf{n}(f, t) \in \mathbb{C}^K$ est un vecteur α -MSCE et chaque point TF (f, t) du signal de parole $\mathbf{y}(f, t) \in \mathbb{C}^K$ suit une loi gaussienne isotrope ;
- comme dans [141], nous supposons que la matrice de dispersion de $\mathbf{n}(f, t)$ ne dépend que de la fréquence. De plus, on suppose qu'elle est diagonale ;
- une décomposition des matrices de covariances de chaque $\mathbf{y}(f, t)$ est proposée comme dans [39]. Chacune est égale à une matrice dite de *covariance spatiale* $\mathbf{R}(f)$, multipliée par un scalaire positif $v(f, t)$. En outre, les $v(f, t)$ seront décomposés via une *nonnegative matrix factorization* (NMF)[39, 47, 141, 161].

Autrement dit, nous employons le modèle probabiliste suivant pour tout (f, t) :

$$\forall (f, t), \mathbf{y}(f, t) \sim \mathcal{N}_c^K(\mathbf{R}(f)v(f, t)) \quad (13.29)$$

$$\forall (f, t), \mathbf{n}(f, t) \sim \mathcal{E}\alpha S^K(\sigma(f)\mathbf{I}_k) \quad (13.30)$$

où la matrice $\mathbf{V} \triangleq [v(f, t)]_{f,t} \in \mathbb{R}_+^{F \times T}$ composé de scalaires non-négatifs vérifie le modèle NMF suivant

$$\mathbf{V} \simeq \mathbf{W}\mathbf{H}$$

avec $\mathbf{W} \triangleq [w_l(f)] \in \mathbb{R}_+^{F \times L}$ et $\mathbf{H} \triangleq [h_l(t)] \in \mathbb{R}_+^{L \times T}$ les matrices décomposant les $v(f, t)$ dans respectivement une base fréquentielle et temporelle. Les scalaires non-négatifs $\sigma(f)$ dépendent uniquement de la fréquence. En appliquant la proposition 13.3.1 sur chaque composante $\mathbf{n}(f, t)$, nous obtenons

$$\forall (f, t), \mathbf{n}(f, t) | \phi(f, t) \sim \mathcal{N}_c^K(\phi(f, t)\sigma(f)\mathbf{I}_k). \quad (13.31)$$

Au final, le mélange $\mathbf{x}(f, t) = \mathbf{y}(f, t) + \mathbf{n}(f, t)$ admet pour tout f, t , le modèle probabiliste

$$\mathbf{x}(f, t) | \phi(f, t) \sim \mathcal{N}_c^K(\mathbf{C}_{\mathbf{x}|\phi}(f, t)) \quad (13.32)$$

$$\mathbf{C}_{\mathbf{x}|\phi}(f, t) = \mathbf{R}(f) \sum_{l=1}^L w_l(f) h_l(t) + \phi(f, t)\sigma(f)\mathbf{I}_k. \quad (13.33)$$

Nous supposons à nouveau que les paramètres $\Phi = \{\phi_s(f, t), \phi_n(f, t)\}_{f,t}$, $\sigma = \{\sigma(f)\}_f$ et $\Theta = \{\mathbf{W}, \mathbf{H}, \mathbf{R}\}$ où $\mathbf{R} = \{\mathbf{R}(f)\}_f$ sont connus. Posons $\mathbf{C}_y(f, t) \triangleq \mathbf{R}(f) \sum_{l=1}^L w_l(f) h_l(t)$, la matrice de covariance de $\mathbf{y}(f, t)$. Nous appliquons la propriété 3.1.1 pour obtenir

$$\forall (f, t), \mathbf{y}(f, t) | \mathbf{x}(f, t), \Phi, \sigma, \Theta \sim \mathcal{N}_c(\mu_{\mathbf{y}|\mathbf{x}, \Phi, \Theta}(f, t), \mathbf{C}_{\mathbf{y}|\mathbf{x}, \Phi, \Theta}(f, t)) \quad (13.34)$$

où

$$\begin{aligned}\boldsymbol{\mu}_{y|x, \boldsymbol{\Phi}, \sigma, \boldsymbol{\Theta}}(f, t) &= \mathbf{C}_y(f, t) (\mathbf{C}_{x|\phi}(f, t))^{-1} \mathbf{x}(f, t) \\ &\triangleq \tilde{\mathbf{G}}(f, t) \mathbf{x}(f, t)\end{aligned}\quad (13.35)$$

$$\begin{aligned}\mathbf{C}_{y|x, \boldsymbol{\Phi}, \sigma, \boldsymbol{\Theta}}(f, t) &= \mathbf{C}_y(f, t) - \mathbf{C}_y(f, t) (\mathbf{C}_{x|\phi}(f, t))^{-1} \mathbf{C}_y(f, t) \\ &= \mathbf{C}_y(f, t) - \tilde{\mathbf{G}}(f, t) \mathbf{C}_y(f, t).\end{aligned}\quad (13.36)$$

En résumé, nous proposons l'estimateur $\hat{\mathbf{y}}(f, t)$ de $\mathbf{y}(f, t)$ suivant

$$\forall(f, t), \hat{\mathbf{y}}(f, t) \triangleq \mathbb{E}_{\phi|x} [\mathbb{E}[\mathbf{y}(f, t) | \mathbf{x}(f, t), \boldsymbol{\Phi}, \boldsymbol{\Theta}]] = \mathbf{G}(f, t) \mathbf{x}(f, t) \quad (13.37)$$

avec

$$\begin{aligned}\forall(f, t), \mathbf{G}(f, t) &\triangleq \mathbb{E}_{\phi|x} \left[\mathbf{C}_y(f, t) (\mathbf{C}_{x|\phi}(f, t))^{-1} \right] \\ &= \mathbf{C}_y(f, t) \mathbb{E}_{\phi|x} \left[(\mathbf{C}_{x|\phi}(f, t))^{-1} \right] \\ &\triangleq \mathbf{C}_y(f, t) \boldsymbol{\Xi}(f, t).\end{aligned}\quad (13.38)$$

La matrice $\boldsymbol{\Xi}(f, t) \triangleq \mathbb{E}_{\phi|x} \left[(\mathbf{C}_{x|\phi}(f, t))^{-1} \right]$ est la marginale selon ϕ de l'inverse de la matrice de covariance du mélange. Notons que le passage de la première à la seconde égalité dans (13.38) découle du fait que la matrice $\mathbf{C}_y(f, t)$ ne dépend pas de $\phi | \mathbf{x}$.

Dans les deux prochains chapitres, nous traiterons des méthodes d'estimations des paramètres.

Chapitre 14

Multi-Alpha Denoising (MAD) : application au débruitage monocanal

Résumé

Dans le chapitre précédent, nous avons établi un lien entre un modèle hybride α -stable et différents travaux sur le rehaussement de la parole datant des années 70-80. Dans un premier temps, nous présentons les techniques employées pour l'estimation des paramètres associés au modèle. Deux familles de paramètres sont à estimer. La première est celle des *variables impulsives*, caractérisant la dynamique à court terme des signaux. La seconde rassemble les paramètres d'échelles, décrivant l'énergie du spectrogramme sur le moyen terme. Pour l'estimation des *variables impulsives*, un algorithme de type *Markov Chain Monte-Carlo* (MCMC) a été proposé dans [162]. Nous établissons dans ce chapitre d'autres hypothèses afin de s'affranchir d'un algorithme MCMC. L'estimation des paramètres d'échelles est obtenue grâce à une relation entre ces derniers et les moments logarithmiques [136]. En pratique, ces moments sont moyennés sur des horizons temporels plus ou moins longs, en fonction des spécificités de chaque source à séparer. Cette approche est réminiscente des travaux de Liutkus dans [121].

14.1 Estimation des variables impulsives

14.1.1 Présentation de l'estimation

Commençons par développer un estimateur pour les *variables impulsives*. Pour ce faire, nous conditionnons les sources s selon les *variables impulsives*, sur lesquelles nous marginalisons [134]. Autrement dit, nous avons la relation suivante :

$$\begin{aligned} \forall (f, t), \mathbb{E}[s(f, t) | x(f, t), \Sigma] &= \mathbb{E}_{\Phi} [\mathbb{E}[s(f, t) | x(f, t), \Sigma, \Phi]] \\ &= \int_{\phi \in \Phi} \mathbb{E}[s(f, t) | x(f, t), \Sigma] g(\phi | x(f, t), \Sigma) d\phi \end{aligned} \quad (14.1)$$

où $g(\phi | x(f, t), \Sigma)$ est la densité de probabilité *a posteriori* de $\phi | x(f, t), \Sigma$. Malheureusement, g n'admet pas de forme analytique. Comme dans [162], nous pourrions estimer cette densité en utilisant un algorithme de type *Markov Chain Monte-Carlo* (MCMC). Nous proposons ici de remplacer la fonction $g(\phi | x(f, t), \Sigma)$ par une distribution dite dégénérée. Cette caractérisation suppose que la loi de probabilité

est une masse unitaire présente sur la valeur la plus probable de ϕ . Il en résulte l'approximation ci-après :

$$\forall (f, t), \int_{\phi \in \Phi} \mathbb{E}[s(f, t) | x(f, t), \Sigma] g(\phi | x(f, t), \Sigma) d\phi \simeq \mathbb{E}[s(f, t) | x(f, t), \Sigma, \hat{\phi}] \quad (14.2)$$

où $\hat{\phi} \triangleq \{\hat{\phi}_s, \hat{\phi}_n\}$ est une estimée ponctuelle. Pour résumer, nous supposons que les *variables impulsionnelles* $\phi_s(f, t), \phi_n(f, t)$ sont égales à $\hat{\phi}_s, \hat{\phi}_n$, étant donnés x, Σ et $\hat{\phi}$, sans aucune incertitude. Ce choix est en effet justifiable car :

- la distribution *a priori* des variables impulsionnelles

$$\phi_s(f, t) \sim \mathcal{P}_{\alpha_s} S \left(2 \left[\cos \left(\frac{\pi \alpha_s}{4} \right) \right]^{2/\alpha_s} \right); \quad (14.3)$$

$$\phi_n(f, t) \sim \mathcal{P}_{\alpha_n} S \left(2 \left[\cos \left(\frac{\pi \alpha_n}{4} \right) \right]^{2/\alpha_n} \right) \quad (14.4)$$

ne dépend pas du point *temps-fréquence* (TF) considéré ;

- les équations (14.3) et (14.4) décrivent des variables i.i.d. et indépendantes des *exposants caractéristiques* α_s, α_n .

Notons qu'une telle approximation permet de réduire considérablement le temps de calcul comparé à une méthode MCMC. En effet, les points $\hat{\phi}_s, \hat{\phi}_n$ ne dépendent pas de la fréquence et du temps. Nous choisissons d'évaluer $\hat{\phi}_s, \hat{\phi}_n$ dans (14.2) par leurs médianes respectives $\mathbb{M}(\phi_n)$ et $\mathbb{M}(\phi_s)$. Il en résulte alors le *filtrage de Wiener paramétrique* (FWP) suivant :

$$\forall f, t, \hat{s}(f, t) \simeq \frac{\mathbb{M}(\phi_s) \sigma_s^2(f, t)}{\mathbb{M}(\phi_s) \sigma_s^2(f, t) + \mathbb{M}(\phi_n) \sigma_n^2(f, t)} x(f, t). \quad (14.5)$$

Nous retrouvons dans (14.5) la constante k esquissée en (13.14) :

$$k = \frac{\mathbb{M}(\phi_n)}{\mathbb{M}(\phi_s)}. \quad (14.6)$$

Justifions empiriquement le choix d'une telle approximation.

14.1.2 Évaluation empirique des masques de Wiener paramétriques

En pratique, les médianes seront calculées grâce à des réalisations i.i.d. de variables aléatoires suivant l'équation (13.19). Les réalisations seront tirées au sort grâce aux méthodes présentées dans [67, 178]. Nous choisissons $\alpha_s = 1.2$ et $\alpha_n = 1.89$ pour notre évaluation empirique. Ce choix d'exposants (α_s, α_n) modélise convenablement le caractère impulsif de la parole et d'un bruit non-stationnaire d'après [162].

Nous avons considéré plusieurs couples $(\sigma_s, \sigma_n) \in [0, 1]^2$ et avons calculé le FWP théorique présenté en (13.23). Les ϕ_s et ϕ_n associés à (13.23) sont issus d'un unique tirage dépendant de (α_s, α_n) . En second lieu pour les mêmes couples (σ_s, σ_n) , nous avons calculé le FWP approximatif donné dans l'équation (14.5). Ici les ϕ_s et ϕ_n sont remplacés par les médianes $\mathbb{M}(\phi_s)$ et $\mathbb{M}(\phi_n)$. Ces deux médianes sont calculées en tirant 10000 réalisations de variables positives α -stables $\mathcal{P}_{\alpha_s} S$ et $\mathcal{P}_{\alpha_n} S$. La Fig. 14.1 montre alors les masques de Wiener obtenus pour les différents couples $(\sigma_s, \sigma_n) \in [0, 1]^2$. Nous employons l'erreur relative moyenne pour évaluer empiriquement notre approximation. L'erreur obtenue entre le FWP théorique et le FWP approché est inférieure à 1.8%. Par conséquent, nous considérons que le choix de la médiane est suffisant pour l'approximation du FWP.

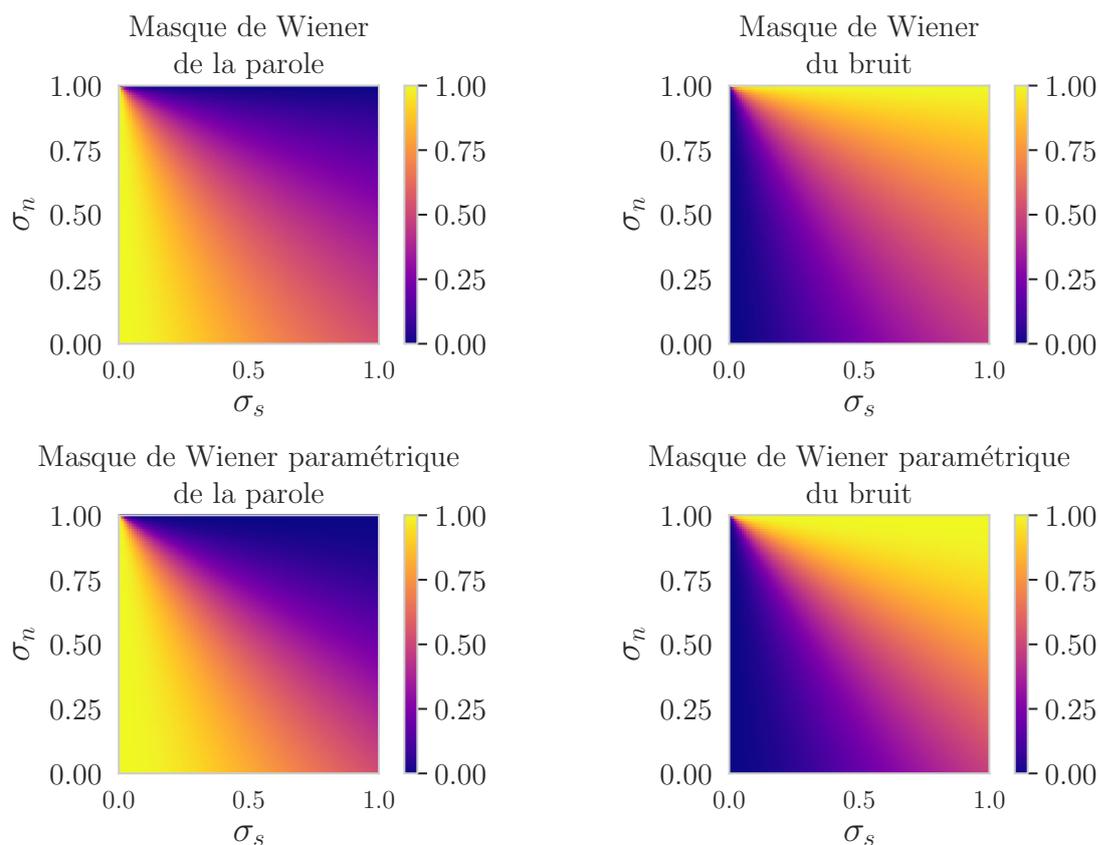


FIGURE 14.1 – Masque de Wiener théorique (figures du haut) et paramétrique (figures du bas).

14.2 Estimation des paramètres d'échelles

14.2.1 Exploitation des moments logarithmiques

Supposons que le signal de parole s et le bruit additif n ont été séparés grâce à la méthode de filtrage en (14.5). Afin de ré-estimer les paramètres d'échelles, nous utilisons le résultat suivant [136, p. 69] :

Proposition 14.2.1. *Soit $x \sim S\alpha S_c(\sigma)$ une variable symétrique isotrope α -stable. Alors, le moment logarithmique de x existe et satisfait la relation suivante :*

$$\mathbb{E}[\ln(|x|)] = \gamma \left(\frac{1}{\alpha} - 1 \right) + \alpha \ln(\sigma) + \ln(2) \quad (14.7)$$

où $\gamma \approx 0.5772156$ est la constante d'Euler.

Remarque La relation (14.7) se démontre en considérant la proposition 4.4.1 page 38 sur les *moments d'ordres faibles fractionnaires* (MOFF). Si nous posons $y = \ln(|x|)$, nous obtenons alors l'égalité entre la fonction génératrice des moments $\mathbb{E}(\exp(py))$ et (4.41). Nous pouvons alors effectuer un développement en série entière et évaluer la dérivée en $p = 0$ de la relation (4.41). Il en découle finalement le résultat souhaité en (14.7).

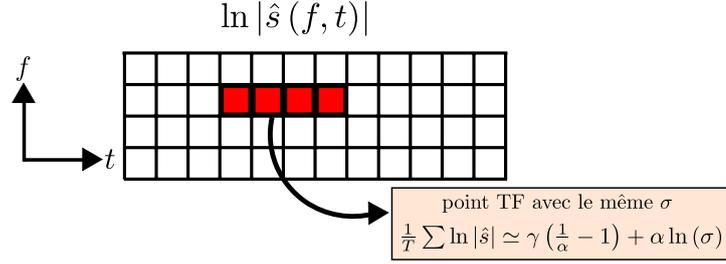


FIGURE 14.2 – Lissage des paramètres d'échelle.

Si nous appliquons la proposition 14.2.1 à chaque source pour chaque point TF, nous obtenons

$$\forall (f, t), \mathbb{E}[\ln(|s(f, t)|)] = \gamma \left(\frac{1}{\alpha_s} - 1 \right) + \alpha_s \ln(\sigma_s(f, t)) + \ln(2), \quad (14.8)$$

$$\forall (f, t), \mathbb{E}[\ln(|n(f, t)|)] = \gamma \left(\frac{1}{\alpha_n} - 1 \right) + \alpha_n \ln(\sigma_n(f, t)) + \ln(2). \quad (14.9)$$

Nous pouvons alors estimer ponctuellement chaque paramètre d'échelle. Cependant, un tel modèle pour les paramètres peut dégrader sérieusement les sources sonores séparées. Cela se justifie notamment par le caractère impulsif des modèles α -stables choisis et par le fait que nous n'avons qu'une seule observation pour chaque point TF. En statistique robuste, il est important de trouver un nombre d'échantillons permettant de préserver l'aspect « à queue lourde » d'une distribution sans pour autant fournir une mauvaise estimation de ses paramètres [78, chap. 1]. Pour ce faire, nous allons introduire une méthode dite de lissage des paramètres.

14.2.2 Lissage des paramètres

Le lissage de paramètres que nous allons présenter maintenant est illustré dans la Fig. 14.2. L'idée globale est de supposer que plusieurs points TF partagent le même paramètre d'échelle. Le lissage est effectué uniquement le long de l'axe temporel. En effet, nous avons observé qu'un lissage le long de l'axe fréquentiel donnait en pratique plus d'artefacts sur le signal de parole obtenu. Soit τ_s, τ_n deux ensembles d'indices de trames temporelles. Nous estimons les paramètres communs $\{\sigma_s(f, t)\}_{t \in \tau_s}$ et $\{\sigma_n(f, t)\}_{t \in \tau_n}$ en effectuant une moyenne des $\{\ln(|s(f, t)|)\}_{t \in \tau_s}$ et $\{\ln(|n(f, t)|)\}_{t \in \tau_n}$. Afin d'obtenir un lissage des paramètres, nous effectuons une moyenne glissante le long de l'axe temporel. La longueur des filtres pour ces moyennes sera résumée par les paramètres $\Delta_s \triangleq \#\tau_s$ et $\Delta_n \triangleq \#\tau_n$ où $\#$ est le cardinal d'un ensemble. Une stratégie d'estimation relativement similaire appelée *Kernel Additive Modeling* (KAM) a déjà été proposée dans [121]. Cependant, KAM propose une méthode utilisant un filtre médian. Notre méthode est quant à elle un filtrage linéaire par rapport au log-spectrogramme.

14.3 Résumé de l'algorithme MAD

L'algorithme itératif que nous proposons est nommé *Multi-Alpha Denoising* (MAD). Il nécessite en premier lieu deux exposants caractéristiques α_s, α_n ainsi que deux estimations de la parole \hat{s} et du bruit \hat{n} . Une fois ces informations données, une estimation des *variables impulsives* est effectuée via un calcul de médianes d'un tirage de variables $\mathcal{P}\alpha\mathcal{S}$, comme décrit dans la sous-section 14.1.2 page 118.

Les paramètres d'échelles sont estimés et lissés grâce à la méthode énoncée dans la section 14.2. Nous procédons alors au FWP de l'équation (14.5) fournissant deux nouvelles estimation de la parole \hat{s} et du bruit \hat{n} . Nous réitérons la procédure d'estimation des paramètres et de filtrage. Le choix du nombre d'itérations est fixé arbitrairement par l'utilisateur. L'algorithme 9 décrit en détail cette approche.

Algorithme 9 MAD : Algorithme MultiAlpha Denoising .

1. Entrées

- TFCT x du signal monocanal
- Nombre d'itérations
- Exposants caractéristiques α_s et α_n
- Horizons de taille fixe Δ_s et Δ_n pour le lissage du log-spectrogramme

2. Initialisation

- $\forall (f, t), \hat{s}(f, t) = \hat{n}(f, t) = x(f, t) / 2$
- Calcul de $\hat{\phi}_s$ et $\hat{\phi}_n$ obtenues à partir de la distribution (14.5)

3. Estimation des paramètres d'échelles

- Calculer les log-spectrogrammes $\ln(|\hat{s}(f, t)|)$ et $\ln(|\hat{n}(f, t)|)$. Les lisser en considérant les paramètres Δ_s et Δ_n : il en résulte $\hat{\sigma}_s(f, t)$ et $\hat{\sigma}_n(f, t)$ en utilisant (14.8) et (14.9)
-

14.4 Évaluation

14.4.1 Description de l'expérience

Dans le but d'évaluer MAD, nous avons pris en compte 30 signaux de paroles extraits de la base de données NOIZEUS⁷. Ces signaux de paroles durent approximativement 3 secondes et sont échantillonnés à 8 kHz. Chacun des 30 extraits est alors corrompu par un bruit additif avec différents *rapports de signal à bruit* (RSB) : 0, 5, 10 et 15 dB. Trois bruits de NOIZEUS sont utilisés : le moteur d'une voiture (*Voiture*), un bruit de foule (*Foule*) et le hall d'un aéroport (*Aéroport*). La TFCT est faite avec des fenêtres de 1024 échantillons en employant une fenêtre de Hamming et 85% de chevauchement entre les trames. La métrique employée pour l'évaluation est appelée *perceptual evaluation of speech quality* (PESQ) [149]. Elle fournit un score entre -0.5 et 4.5, correspondant respectivement au pire et à la meilleure performance de débruitage. Les algorithmes suivants seront comparés :

MAD Notre algorithme proposé dans la section 14.3. Dans cette évaluation, nous considérons $\alpha_s = 1.3$ et $\alpha_n = 1.8$ pour les exposants caractéristiques. Les longueurs des filtres de lissage sont de $\Delta_s = 0.09$ s et $\Delta_n = 0.16$ s pour l'estimation des paramètres d'échelles. Ces longueurs signifient que le bruit est supposé plus stationnaire que la parole. Nous itérons le procédé d'estimation de paramètre et de filtrage 4 fois. Comme mentionné dans l'algorithme 9, nous initialisons la parole et le bruit en posant $\hat{s} = \hat{n} = x/2$.

SSA L'algorithme de *soustraction spectrale d'amplitude* (SSA) est le cas particulier où $\alpha = 1$ dans (13.10). Nous utilisons un *détecteur d'activité vocale* (DAV) dit « parfait », dans le sens où nous l'appliquons sur la parole non bruitée.

SSG La *soustraction spectrale généralisée* (SSG) est celle décrite dans l'équation (13.11). À nouveau, un DAV « parfait » est employé et les paramètres choisis dans (13.11). sont $\alpha = 1.2$ et $k = 0.8$. Cette sélection des paramètres a été effectuée de manière empirique, pour fournir le meilleur score PESQ possible.

7. NOIZEUS <http://ecs.utdallas.edu/loizou/speech/noizeus/>

MEQM Ce dernier algorithme de l'état de l'art présenté dans [41] est appelé *Minimum de l'Erreur Quadratique Moyenne* (MEQM) via l'estimation de l'Amplitude Spectrale à Court Terme. Il s'agit également d'un FWP mais paramétré à la fois par un *a priori* et un *a posteriori* du RSB qui sont estimés au préalable. Cet algorithme est muni également d'un DAV supposé « parfait ». Nous considérons les paramètres conseillés dans [41], c'est à dire $\alpha = 0.92$ et $\beta = 2$.

14.4.2 Résultats et commentaires

Les scores sont affichés dans le tableau 14.1. L'algorithme MAD produit un meilleur score, quelles que soient les configurations et les RSB considérés. Nous pouvons par ailleurs observer de manière générale que la méthode MEQM obtient également des scores compétitifs avec MAD. Après écoute, il apparaît que cette différence de scores peut être perçue. En effet, la méthode MEQM présente un bruit résiduel significatif comparé à la méthode MAD. Ce bruit résiduel était d'ailleurs déjà mentionné dans [41].

Bruit	RSB	MAD	SSA	SSG	MEQM
Voiture	0 dB	1.312	1.183	1.201	1.178
	5 dB	1.558	1.300	1.343	1.282
	10 dB	1.913	1.405	1.560	1.515
	15 dB	2.208	1.504	1.755	1.889
Aéroport	0 dB	1.308	1.174	1.192	1.208
	5 dB	1.561	1.301	1.390	1.368
	10 dB	1.844	1.389	1.556	1.642
	15 dB	2.102	1.507	1.755	2.000
Foule	0 dB	1.272	1.170	1.200	1.199
	5 dB	1.519	1.272	1.351	1.357
	10 dB	1.816	1.390	1.576	1.603
	15 dB	2.121	1.494	1.771	2.042

TABLE 14.1 – Scores PESQ pour chaque algorithme. Les écritures en caractères gras indiquent les meilleurs scores.

14.4.3 Restauration audio du patrimoine ethnomusicologique

Le centre de recherche en ethnomusicologie (CREM) détient plusieurs dizaines de milliers d'enregistrements audio monocanal très dégradés datant de la fin du *XIX^{ème}*, début du *XX^{ème}* siècle. Les différentes dégradations peuvent être atmosphériques (vent, pluie etc.) ou dues au matériel employé générant différents bruits parasites. Par exemple, plusieurs enregistrements anciens utilisent des micros à ruban générant un bruit dit électrostatique. Ce problème s'explique par la dégradation de charge de polarisation dans le temps. D'autres bruits sont également ajoutés à cause de l'enregistrement vinyle. En effet, ce type d'enregistrement induit une composante périodique de bruit, calculable selon le nombre de tours par minutes que le disque effectue. Enfin les supports d'enregistrements se dégradent également avec le temps.

En général, les enregistrements antérieurs à 1947 n'admettaient pas un nombre de tours par minute constant.

L'algorithme MAD a alors été utilisé afin de débruiter des données du CREM. À l'heure où ce manuscrit est rédigé, une version allégée de MAD où l'utilisateur peut modifier manuellement les paramètres de longueur du filtre de lissage Δ_n et Δ_s est en cours de création (cf. Fig. 14.3).

The screenshot displays the MAD online interface for audio restoration. At the top, the title "Conte : 'Le rusé Barma'" is centered. Below it is a timeline from 0:00 to 0:22. The interface is divided into two main sections: "Reference" and "Denoised". The "Reference" section has a "SOLO" button, and the "Denoised" section has an "SOLO" button. A green play button is located to the right of the waveforms. Below the waveforms, there is a text area titled "MAD Denoise" with a description of the method and a citation: "Fontaine, Mathieu, et al. 'Explaining the parameterized wiener filter with alpha-stable processes.' Applications of Signal Processing to Audio and Acoustics (WASPAA), 2017 IEEE Workshop on. IEEE, 2017." Underneath the text is a "Parameters" section with two sliders: "Speech Stability (ms)" set to 85 and "Noise Stability (ms)" set to 680. At the bottom, there is a large orange "PROCESS" button.

FIGURE 14.3 – Capture d'écran : Version de MAD en ligne pour la restauration d'archives du CREM.

Chapitre 15

Gaussien + Elliptique : application au débruitage multicanal

Résumé

Ce chapitre traite de l'application pour le débruitage du modèle de mélange hybride multicanal introduit dans la section 13.3 page 113. Inspiré de [161], nous présentons un algorithme de type *Espérance-Maximisation* (EM). Cette méthode a pour but l'estimation des paramètres du signal de parole multicanal supposé gaussien. Nous employons cette fois-ci un algorithme de type *Markov Chain Monte-Carlo* (MCMC) pour l'estimation des *variables impulsionnelles* associées au bruit résiduel. Plus particulièrement, l'algorithme de *Metropolis-Hastings* offre dans ce modèle hybride une estimation relativement rapide des *variables impulsionnelles*. Dans le contexte du rehaussement de la parole, il est souvent plus approprié d'obtenir un signal de voix monocanal (traduction, sous-titrage etc.). Nous proposons alors une technique de filtrage spatial pour récupérer le signal monocanal désiré.

15.1 Estimation des paramètres

15.1.1 Espérance-Maximisation

Reprenons le modèle présenté dans les équations (13.29) et (13.31). Supposons que les observations \mathbf{x} et les variables impulsionnelles Φ soient connues. Le but est alors d'estimer les paramètres Θ . Afin d'obtenir les paramètres de *nonnegative matrix factorization* (NMF) \mathbf{W}, \mathbf{H} les plus probables, nous estimons les paramètres au maximum de vraisemblance :

$$(\mathbf{W}^*, \mathbf{H}^*, \mathbf{R}^*) = \underset{\mathbf{W}, \mathbf{H}, \mathbf{R}}{\operatorname{argmax}} [\ln(g(\mathbf{x}, \Phi | \Theta, \sigma))], \quad (15.1)$$

où $\Phi \triangleq \{\phi(f, t)\}_{f,t}$ regroupe ici les *variables impulsionnelles* du bruit résiduel et $\ln(g(\mathbf{x}, \Phi | \Theta, \sigma))$ est la log-probabilité jointe des données. Comme dans [161], nous proposons un algorithme d'*espérance-maximisation* (EM) consistant à minimiser une borne supérieure de $\mathcal{L}_n(\mathbf{W}, \mathbf{H}, \mathbf{R}) = -\ln(g(\mathbf{x}, \Phi | \Theta, \sigma))$. Cette approche peut être décrite par les deux étapes suivantes :

$$\text{Étape-E : } Q_n(\mathbf{W}, \mathbf{H}, \mathbf{R}) = -\mathbb{E}_{\Phi | \mathbf{x}, \mathbf{W}^{(n-1)}, \mathbf{H}^{(n-1)}} [\mathcal{L}_n(\mathbf{W}, \mathbf{H}, \mathbf{R})], \quad (15.2)$$

$$\text{Étape-M : } (\mathbf{W}^{(n)}, \mathbf{H}^{(n)}, \mathbf{R}^{(n)}) = \underset{\mathbf{W}, \mathbf{H}, \mathbf{R}}{\operatorname{argmax}} Q_n(\mathbf{W}, \mathbf{H}, \mathbf{R}). \quad (15.3)$$

Analysons en détail la résolution de (15.2) et (15.3).

15.1.1.1 Étape-E

Nous définissons en premier lieu une fonction positive qui domine la log-probabilité jointe négative $\mathcal{L}_n(\mathbf{W}, \mathbf{H}, \mathbf{R})$ à l'itération n . Cette fonction proposée dans [154] est la suivante :

$$\mathcal{L}_n(\mathbf{W}, \mathbf{H}, \mathbf{R}) = \sum_{f,t} \left[\text{tr} \left(\tilde{\mathbf{X}}(f,t) (\mathbf{C}_{\mathbf{x}|\phi}(f,t))^{-1} \right) + \ln(\det[\mathbf{C}_{\mathbf{x}|\phi}(f,t)]) \right] \quad (15.4)$$

où $\tilde{\mathbf{X}}(f,t) \triangleq \mathbf{x}(f,t) \mathbf{x}^*(f,t)$ et $\mathbf{C}^{\mathbf{x}|\phi}(f,t)$ est la matrice de covariance du mélange sachant ϕ décrite dans l'équation (13.33). La fonction dominant \mathcal{L}_n présentée également dans [154] est alors la suivante :

$$\begin{aligned} \mathcal{L}_n^+(\mathbf{W}, \mathbf{H}, \mathbf{R}, \mathbf{U}, \mathbf{V}) = \sum_{f,t} \left[\sum_l \frac{\text{tr} \left(\tilde{\mathbf{X}}(f,t) \mathbf{U}_l(f,t) (\mathbf{C}_{\mathbf{x}|\phi}(f,t))^{-1} \mathbf{U}_l(f,t) \right)}{w_l(f) h_l(t)} + \frac{\text{tr} \left(\tilde{\mathbf{X}}(f,t) \mathbf{U}_r^2(f,t) \right)}{\sigma(f) \phi(f,t)} + \right. \\ \left. + \ln(\det[\mathbf{V}(f,t)]) + \frac{\det[\mathbf{C}_{\mathbf{x}|\phi}(f,t)] - \det[\mathbf{V}(f,t)]}{\det[\mathbf{V}(f,t)]} \right] \quad (15.5) \end{aligned}$$

où $\mathbf{U} \triangleq \{\mathbf{U}_l(f,t), \mathbf{U}_r(f,t)\}_{l,f,t}$, $\mathbf{V} \triangleq \{\mathbf{V}(f,t)\}_{f,t}$ sont des matrices hermitiennes définies positives dans $\mathbb{C}^{K \times K}$ telles que

$$\forall(f,t), \mathbf{U}_l(f,t) + \mathbf{U}_r(f,t) = \mathbf{I}_K. \quad (15.6)$$

Il est prouvé en outre que pour tout $\mathbf{W}, \mathbf{H}, \mathbf{R}, \mathbf{U}, \mathbf{V}$ comme défini précédemment, nous avons

$$\mathcal{L}_n^+(\mathbf{W}, \mathbf{H}, \mathbf{R}, \mathbf{U}, \mathbf{V}) \geq \mathcal{L}_n(\mathbf{W}, \mathbf{H}, \mathbf{R}). \quad (15.7)$$

L'égalité est vérifiée lorsque

$$\mathbf{U}_l(f,t) = w_l(f) h_l(f) \mathbf{R}(f) (\mathbf{C}_{\mathbf{x}|\phi}(f,t))^{-1}; \quad \mathbf{U}_r(f,t) = \sigma(f) \phi(f,t) (\mathbf{C}_{\mathbf{x}|\phi}(f,t))^{-1}; \quad (15.8)$$

$$\mathbf{V}(f,t) = \mathbf{C}^{\mathbf{x}|\phi}(f,t). \quad (15.9)$$

En utilisant (15.7) et la définition de \mathcal{Q}_n dans (15.2), l'inégalité

$$\mathbb{E}_{\Phi|\mathbf{x}} \mathcal{L}_n(\cdot) \leq \mathbb{E}_{\Phi|\mathbf{x}} \mathcal{L}_n^+(\cdot) \triangleq \mathcal{Q}_n^+(\cdot) \quad (15.10)$$

est vraie pour

$$\begin{aligned} \mathcal{Q}_n^+(\mathbf{W}, \mathbf{H}, \mathbf{R}, \mathbf{U}, \mathbf{V}) = \sum_{f,t} \left[\sum_l \frac{\mathbb{E}_{\Phi|\mathbf{x}} \left(\text{tr} \left[\tilde{\mathbf{X}}(f,t) \mathbf{U}_l(f,t) (\mathbf{C}_{\mathbf{x}|\phi}(f,t))^{-1} \mathbf{U}_l(f,t) \right] \right)}{w_l(f) h_l(t)} \right. \\ \left. + \mathbb{E}_{\Phi|\mathbf{x}} \left(\text{tr} \left[\frac{\tilde{\mathbf{X}}(f,t) \mathbf{U}_r^2(f,t)}{\sigma(f) \phi(f,t)} \right] \right) + \mathbb{E}_{\Phi|\mathbf{x}} \left(\ln(\det[\mathbf{V}(f,t)]) + \frac{\det[\mathbf{C}_{\mathbf{x}|\phi}(f,t)] - \det[\mathbf{V}(f,t)]}{\det[\mathbf{V}(f,t)]} \right) \right]. \quad (15.11) \end{aligned}$$

L'équation (15.11) admet alors des dérivées partielles qui sont calculées pour la résolution de l'étape-M de l'algorithme en (15.3).

15.1.1.2 Étape-M

Les résultats dans [154] nous fournissent les dérivées partielles suivantes de Q_n^+ contre $w_l(f)$ et $h_l(t)$:

$$\begin{aligned} \frac{\partial Q_n^+}{\partial w_l(f)} = \sum_t \left[\mathbb{E}_{\Phi|x} \left(\frac{\det[\mathbf{C}_{x|\phi}(f,t)]}{\det[\mathbf{V}(f,t)]} \operatorname{tr} \left((\mathbf{C}_{x|\phi}(f,t))^{-1} \mathbf{R}(f) \right) h_l(t) \right) \right. \\ \left. - \mathbb{E}_{\Phi|x} \left(\frac{\operatorname{tr}(\mathbf{U}_l(f,t) \mathbf{R}^{-1}(f) \mathbf{U}_l(f,t) \tilde{\mathbf{X}}(f,t))}{w_l^2(f) h_l(t)} \right) \right]; \end{aligned} \quad (15.12)$$

$$\begin{aligned} \frac{\partial Q_n^+}{\partial h_l(t)} = \sum_f \left[\mathbb{E}_{\Phi|x} \left(\frac{\det[\mathbf{C}_{x|\phi}(f,t)]}{\det[\mathbf{V}(f,t)]} \operatorname{tr} \left((\mathbf{C}_{x|\phi}(f,t))^{-1} \mathbf{R}(f) \right) w_l(f) \right) \right. \\ \left. - \mathbb{E}_{\Phi|x} \left(\frac{\operatorname{tr}(\mathbf{U}_l(f,t) \mathbf{R}^{-1}(f) \mathbf{U}_l(f,t) \tilde{\mathbf{X}}(f,t))}{w_l(f) h_l^2(t)} \right) \right]. \end{aligned} \quad (15.13)$$

Nous effectuons alors une descente de gradient en considérant la règle de mises à jour multiplicatives des paramètres dans [51]. Cela a pour effet de minimiser Q_n^+ et donc de considérer les équations (15.8) et (15.9) car l'égalité dans (15.10) est atteinte

$$\forall f, w_l(f) \leftarrow w_l(f) \sqrt{\frac{\sum_t h_l(t) \operatorname{tr}(\mathbf{R}(f) \mathbf{P}(f,t))}{\sum_t h_l(t) \operatorname{tr}(\mathbf{R}(f) \Xi(f,t))}}; \forall t, h_l(t) \leftarrow h_l(t) \sqrt{\frac{\sum_f w_l(f) \operatorname{tr}(\mathbf{R}(f) \mathbf{P}(f,t))}{\sum_f w_l(f) \operatorname{tr}(\mathbf{R}(f) \Xi(f,t))}} \quad (15.14)$$

où la quantité

$$\forall(f,t), \Xi(f,t) \triangleq \mathbb{E}_{\Phi|x} \left[(\mathbf{C}_{x|\phi}(f,t))^{-1} \right] \quad (15.15)$$

est définie pour la méthode de filtrage dans (13.38) et

$$\forall(f,t), \mathbf{P}(f,t) \triangleq \mathbb{E}_{\Phi|x} \left[(\mathbf{C}_{x|\phi}(f,t))^{-1} \tilde{\mathbf{X}}(f,t) (\mathbf{C}_{x|\phi}(f,t))^{-1} \right]. \quad (15.16)$$

Pour les résultats concernant l'inversion de l'opérateur de trace et de l'espérance, nous renvoyons le lecteur aux résultats dans [157]. La méthode d'estimation des matrices $\Xi(f,t)$ et $\mathbf{P}(f,t)$ est esquissée dans la sous-section 15.1.2. Il nous reste à décrire une méthode d'estimation des matrices $\mathbf{R}(f)$. Nous utilisons la mise à jour des matrices de covariance spatiales proposée dans [140], plus robustes en pratique que celles proposées dans [39] :

$$\forall f, \mathbf{R}(f) \leftarrow \left(\sum_t v(f,t) \right)^{-1} \times \sum_t \left(\mathbf{C}^{yy^*|x}(f,t) \right), \quad (15.17)$$

où $v(f,t) \triangleq \sum_{l=1}^L w_l(f) h_l(t)$ est définie dans (13.29) et

$$\forall(f,t), \mathbf{C}^{yy^*|x}(f,t) \triangleq \mathbf{G}(f,t) \tilde{\mathbf{X}}(f,t) \mathbf{G}^*(f,t) + \mathbf{C}^y(f,t) - \mathbf{G}(f,t) \mathbf{C}^y(f,t) \quad (15.18)$$

est la variance *a posteriori* totale du signal de parole avec $\mathbf{C}^y(f,t) \triangleq \mathbf{R}(f) v(f,t)$, la matrice de covariance de la source \mathbf{y} .

15.1.2 Algorithme de Metropolis-Hastings

Nous établissons maintenant notre stratégie d'estimation des matrices $\Xi(f,t)$ et $\mathbf{P}(f,t)$ définie en (15.15) et (15.16). En pratique, ces matrices dépendent des paramètres Φ . Cependant, les *variables impulsionnelles* n'admettent en général pas de forme analytique de leur densité de probabilité. Afin de

palier à ce problème, nous proposons un algorithme de type MCMC. Nous employons en particulier un algorithme de *Metropolis-Hastings* via une estimation empirique $\bar{\Xi}(f, t)$ et $\bar{\mathbf{P}}(f, t)$ respectivement des matrices $\Xi(f, t)$ et $\mathbf{P}(f, t)$ pour tout (f, t) :

$$\bar{\Xi}(f, t) \simeq \frac{1}{I} \sum_{i=1}^I \left(\mathbf{C}^{\mathbf{x}|\varphi_i}(f, t) \right)^{-1} ; \bar{\mathbf{P}}(f, t) \simeq \frac{1}{I} \sum_{i=1}^I \left(\left(\mathbf{C}^{\mathbf{x}|\varphi_i}(f, t) \right)^{-1} \tilde{\mathbf{X}}(f, t) \left(\mathbf{C}^{\mathbf{x}|\varphi_i}(f, t) \right)^{-1} \right) \quad (15.19)$$

avec $\left(\mathbf{C}^{\mathbf{x}|\varphi_i}(f, t) \right)^{-1} = [\mathbf{R}(f) \sum_l (w_l(f) h_l(t)) + \varphi_i(f, t) \sigma(f) \mathbf{I}_K]^{-1}$ et $\varphi_i(f, t)$ calculé de la manière suivante pour tout (f, t) :

Première étape (Échantillonnage) :

Générer un échantillonnage grâce à la distribution de l'a priori $\varphi'(f, t) \sim \mathcal{P}\alpha S\left(2 \cos\left(\frac{\pi\alpha}{4}\right)^{2/\alpha}\right)$.

Seconde étape (Acceptation) :

- Tirer $u \sim \mathcal{U}([0, 1])$ avec \mathcal{U} décrivant une distribution uniforme.
- Calculer la probabilité d'acceptation ci-dessous :

$$\text{acc}(\varphi(f, t) \rightarrow \varphi'(f, t)) = \min\left(1, \frac{\mathcal{N}_c(\mathbf{x}(f, t); 0, \varphi'(f, t) \sigma(f) \mathbf{I}_K + \mathbf{C}^{\mathbf{y}}(f, t))}{\mathcal{N}_c(\mathbf{x}(f, t); 0, \varphi(f, t) \sigma(f) \mathbf{I}_K + \mathbf{C}^{\mathbf{y}}(f, t))}\right). \quad (15.20)$$

- Tester alors l'acceptation :
 - si $u < \text{acc}(\varphi_{i-1}(f, t) \rightarrow \varphi'(f, t))$, alors $\varphi_i(f, t) = \varphi'(f, t)$ (acceptation)
 - sinon, $\varphi_i(f, t) = \varphi_{i-1}(f, t)$ (rejet)

Il nous reste à décrire une méthode d'estimation des paramètres $\sigma(f)$ pour le bruit résiduel.

15.1.3 Estimation des paramètres σ

Nous choisissons pour l'estimation des coefficients positifs $\sigma = \{\sigma(f)\}_f$ d'employer un résultat présenté dans [27]. Il affirme que si $z \sim \mathcal{E}\alpha S^K(\sigma)$, alors $\mathbb{E}[\|z\|^p]^{1/p} \propto \sigma$ pour tout $0 < p < \alpha$, où $\|\cdot\|$ représente la norme Hermitienne et \propto « proportionnel à ». La stratégie que nous adoptons est alors d'appliquer cette estimation au mélange et uniquement au début de l'algorithme. Afin d'obtenir une estimation robuste, nous prenons la médiane \mathbb{M} au lieu de la moyenne. Le choix de prendre la médiane sur \mathbf{x} peut être également justifié par le fait que les $\sigma(f)$ se rapportent au modèle du bruit résiduel supposé plus impulsif que la parole. C'est pourquoi, il est important d'utiliser un estimateur robuste comme la médiane pour décrire $\sigma(f)$ qui est associée à une distribution α -stable. Nous décidons de choisir $p = \alpha/2$. Pour résumer, nous faisons l'approximation

$$\sigma(f) \simeq \mathbb{M}_t \left(\|\mathbf{x}(f, t)\|^{\alpha/2} \right)^2. \quad (15.21)$$

15.1.4 Reconstruction monocanal : Maximisation de l'énergie

Soit $\hat{\mathbf{y}}$ le signal multicanal obtenu après le *filtrage de Wiener paramétrique* décrit en (13.37). Dans un contexte de rehaussement de la parole, on souhaite souvent récupérer plutôt un signal monocanal $\hat{\mathbf{s}} \in \mathbb{C}^{F \times T}$. Dans cette étude, nous prenons $\hat{\mathbf{s}}$ comme étant obtenu à partir de $\hat{\mathbf{y}}$ par une *formation de voie* linéaire, invariant dans le temps $\mathbf{B}(f) \in \mathbb{C}^K$ [173] :

$$\forall(f, t), \hat{\mathbf{s}}(f, t) \triangleq \mathbf{B}^*(f) \hat{\mathbf{y}}(f, t). \quad (15.22)$$

Plusieurs méthodes existent afin de construire un *tel filtre spatial* $\mathbf{B}(f)$. Ici, nous choisissons de maximiser l'énergie moyenne du vecteur $\mathbf{B}^*(f)\mathbf{y}(f,t) | \mathbf{x}$. Calculons en premier lieu l'énergie moyenne dans le temps de $\mathbf{B}^*(f)\mathbf{y}(f,t) | \mathbf{x}$. Autrement dit, la quantité $\frac{1}{T} \sum_t \mathbb{E} \left(|\mathbf{B}^*(f)\mathbf{y}(f,t)|^2 | \mathbf{x}(f,t) \right)$:

$$\begin{aligned} \frac{1}{T} \sum_t \mathbb{E} \left(|\mathbf{B}^*(f)\mathbf{y}(f,t)|^2 | \mathbf{x}(f,t) \right) &= \frac{1}{T} \sum_t \mathbb{E} \left(\mathbf{B}^*(f)\mathbf{y}(f,t)\mathbf{y}^*(f,t)\mathbf{B}(f) | \mathbf{x}(f,t) \right) \\ &= \frac{1}{T} \sum_t \mathbf{B}^*(f) \mathbb{E} \left(\mathbf{y}(f,t)\mathbf{y}^*(f,t) | \mathbf{x}(f,t) \right) \mathbf{B}(f) \\ &= \mathbf{B}^*(f) \frac{1}{T} \sum_t \left(\mathbf{C}^{yy^*|x}(f,t) \right) \mathbf{B}(f) \end{aligned} \quad (15.23)$$

où $\mathbf{C}^{yy^*|x}(f,t)$ est définie en (15.18). Ce problème connu [39, 173] peut alors être résolu en prenant $\mathbf{B}(f)$ comme le vecteur propre associé à la plus grande valeur propre de la matrice Hermitienne $\frac{1}{T} \sum_t \left(\mathbf{C}^{yy^*|x}(f,t) \right)$, sous contrainte que $\|\mathbf{B}(f)\| = 1$ (avec $\|\cdot\|$ la norme Hermitienne).

15.1.5 Algorithme Alpha Residual Component (ARC)

Algorithme 10 ARC : *Alpha Residual Component* (ARC).

1. Entrées

- Observations \mathbf{x} obtenues après une TFCT sur un signal multicanal
- Nombre de variables latentes L pour la NMF
- Nombre N d'itérations pour l'algorithme EM
- Nombre I d'itérations pour l'algorithme de MH

2. Initialisation

- Calculer σ comme dans (15.21)
- Initialisation aléatoire des matrices \mathbf{W} et \mathbf{H}
- $\mathbf{R}(f) \leftarrow \mathbf{I}_K$ (où \mathbf{I}_K est la matrice identité de taille $K \times K$)
- $\phi(f,t) \sim \mathcal{P}\alpha S \left(2 \cos \left(\frac{\pi\alpha}{4} \right)^{2/\alpha} \right)$

3. Algorithme EM, pour $n = 1 \dots N$

- Algorithme de MH :
 - (a) Mettre à jour $\varphi_i(f,t)$ via l'algorithme MH (sous-section 15.1.2)
 - (b) Calculer $\bar{\mathbf{E}}$ et $\bar{\mathbf{P}}$ (15.19)
- Mettre à jour \mathbf{W} , \mathbf{H} (15.14) et \mathbf{R} (15.17)

4. Reconstruction de la source image

Calculer $\hat{\mathbf{y}}$ comme décrit en (13.37)

5. Filtrage spatial

- Considérer \mathbf{B}_f comme le vecteur propre principal de norme 1 de la matrice $\frac{1}{T} \sum_t \mathbf{C}^{yy^*|x}(f,t)$
 - Calculer $\hat{s}(f,t) = \mathbf{B}^*(f)\hat{\mathbf{y}}(f,t)$
-

Les étapes sont résumées dans l'algorithme 10 appelé *Alpha Residual Component* (ARC). Nous pouvons observer que la technique de débruitage est assez similaire à celle proposée en monocanal dans le chapitre 14. Les différences majeures sont l'utilisation d'un algorithme de type MH pour l'estimation des *variables impulsionnelles* ϕ et la reconstruction monophonique proposée dans la sous-section 15.1.4.

15.2 Évaluation

15.2.1 Configuration et état de l'art

Nous évaluons dans les expériences suivantes la qualité de rehaussement de la parole de l'algorithme ARC. Pour ce faire, nous le comparons aux deux méthodes ci-après :

- ARC** Le filtrage *alpha residual component* proposé dans ce chapitre. Nous posons $\alpha = 1.8$ pour la composante de bruit résiduel. En outre, 5 itérations de l'algorithme sont exécutés. L'usage de l'algorithme de Metropolis-Hastings à l'étape 3 de l'algorithme 10 est effectué 6 fois.
- FWM** Le filtrage de Wiener multicanal classique (FWM) [19, 36] supposant que le signal de parole et de bruit suivent des distributions gaussiennes, pour tout point *temps-fréquence* (TF).
- GDVP** Introduit dans [158], le filtrage de Wiener multicanal *généralisé via une décomposition des valeurs propres* (GDVP) est basé sur l'hypothèse que la matrice d'autocorrélation de la parole est une matrice de rang faible. Cette modélisation permet de fournir ainsi une réduction du bruit plus robuste.

Le corpus pour l'évaluation est construit à partir de la base de données Librispeech [142]. Il contient plusieurs signaux de paroles monocanal avec une fréquence d'échantillonnage de 16 kHz . Nous concaténons plusieurs de ces données en insérant des périodes de silences afin d'obtenir des signaux durant environ 3 minutes chacun. À cela, nous ajoutons différents bruits environnementaux extrait d'Aurora [74] : bruit de foule, ambiance de restaurant et bruit de train. Nous appliquons alors une TFCT utilisant un fenêtrage de Hann de longueur 1024 et un chevauchement de 50%. Un *détecteur d'activité vocale* (DAV) « parfait », dans le sens où le DAV est estimé à partir de la voix non bruitée, est utilisé dans les trois algorithmes.

Les extraits de 3 minutes bruités sont alors convolués avec différentes *réponses impulsives* de salles fournies via Roomsimove [174]. Les dimensions de la salle simulée sont de $5 \times 4 \times 3$ mètres et les temps de réverbération, basés sur une diminution du niveau sonore à 60 dB , sont de 0,250 et 500 ms. Nous considérons deux microphones ($K = 2$) séparés de 15 cm et placés au milieu de la pièce à 1.5 m de hauteur. Les sources sont quant à elles situées à 1 m de distance des microphones. L'angle solide entre les sources vu par les microphones est pris de 30° ou 90° . Plusieurs scénarios de *rappports signal à bruit* (RSB) sont considérées : -5 dB , 0 dB , 5 dB et 10 dB . Pour résumer, un total de 48 sources bruitées ont été rehaussées par les trois méthodes proposées.

Comme les sources sont immobiles, les deux algorithmes de l'état de l'art emploient des matrices d'autocorrélations ne dépendant pas du temps. Elles sont ainsi moyennées le long de l'axe temporel. La Fig. 15.1 résume la pièce simulée et les configurations de cette évaluation.

15.2.2 Mesure de performance

Nous considérons deux scores courants en rehaussement de la parole. Le premier se nomme la *mesure de distorsion spectrale pondérée par l'intelligibilité de la parole* (DSP-IP) et la seconde le *rappport signal à bruit avec intelligibilité de la parole pondérée* (RSB-IP) [69]. La DSP-IP est définie comme suit :

$$\text{DSP-IP} = \sum_i I_i \text{DSP}_i \quad (15.24)$$

où I_i est la *fonction d'importance de bandes fréquentielles* [83] et DSP_i la distorsion spectrale moyenne (en dB) de la $i^{\text{ème}}$ bande d'un tiers d'octave :

$$\text{DSP}_i = \frac{1}{(2^{1/6} - 2^{-1/6}) f_i^c} \int_{2^{-1/6} f_i^c}^{2^{1/6} f_i^c} |10 \log_{10} G^y(f)| df \quad (15.25)$$

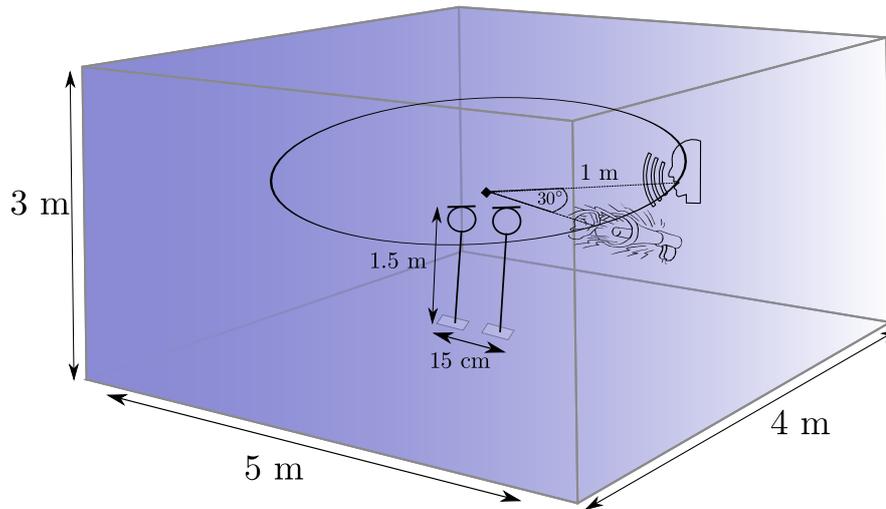


FIGURE 15.1 – Configuration de la pièce simulée pour l'évaluation des algorithmes de débruitage.

avec les fréquences centrées f_i^c et $G^y(f)$ obtenues comme suit :

$$\forall f, G^y(f) = \frac{P_y(f)}{P_{\hat{y}}(f)} \quad (15.26)$$

dans lequel $P_y(f)$ et $P_{\hat{y}}(f)$ sont respectivement la puissance, pour la fréquence f , de la parole non bruitée y et de la parole débruitée \hat{y} . Le RSB-IP est quant à lui calculé afin de fournir le score d'amélioration du débruitage que chaque méthode a pu fournir :

$$\Delta RSB - IP = \sum_i (I_i RSB_{i,sortie} - RSB_{i,entrée}) \quad (15.27)$$

où $RSB_{i,sortie}$ et $RSB_{i,entrée}$ représentent respectivement le RSB du filtre de réduction du bruit obtenu via les différentes méthodes et le RSB du signal obtenu dans le premier microphone, à la $i^{\text{ème}}$ bande fréquentielle.

15.2.3 Première expérience : résultats généraux

Nous étudions en premier lieu l'impact de la réverbération sur la performance de chaque algorithme. La Fig. 15.2 représente les deux scores décrits dans la sous-section 15.2.2, en fonction du temps de réverbération et du RSB entre la parole et le bruit résiduel. Nous observons que dans le cas anéchoïque ($RT_{60} = 0ms$), les algorithmes de l'état de l'art apparaissent plus performant que celui proposé. Cependant, en présence d'une réverbération ($RT_{60} = 250ms$ et $RT_{60} = 500ms$), l'algorithme ARC présente un RSB-IP plus élevé et un DSP-IP plus faible. En somme, un meilleur rehaussement de la parole. Concentrons-nous davantage sur l'évolution de la RSB-IP en fonction de la configuration spatiale des sources.

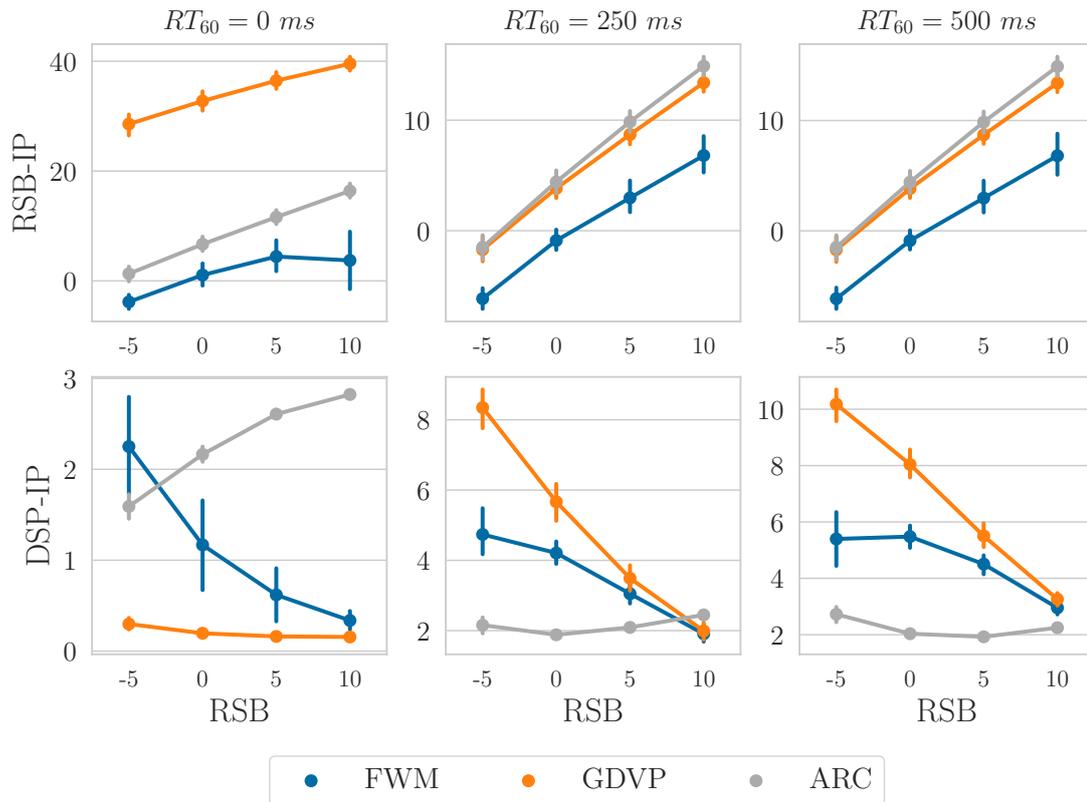


FIGURE 15.2 – Moyenne de RSB-IP (plus le score est élevé, meilleure est l'amélioration) et DSP-IP (plus le score est faible, meilleure est la qualité de débruitage) en fonction du RSB en entrée. Les barres indiquent l'écart type de chaque scénario.

15.2.4 Deuxième expérience : score de rehaussement de la parole en milieu réverbérant

Position du bruit	RSB	FWM	GDVP	ARC
30°	-5 dB	0	2.2	3.7
	0 dB	1.0	3.1	4.2
	5 dB	1.2	5.2	5.4
90°	-5 dB	0.2	2.7	2.9
	0 dB	1.3	5.0	4.8
	5 dB	1.9	5.4	5.7

TABLE 15.1 – Amélioration du RSB-IP quand le score de DSP-IP est identique pour chaque méthode. Les meilleures performances sont en police de caractère grasse.

Dans les algorithmes de l'état de l'art, il subsiste un compromis entre la quantité de bruit retirée et la distorsion spectrale introduite dans le signal de parole estimé. Il est possible de modifier explicitement cet équilibre, appelé dans la littérature *poids de distorsion spectrale du filtre de Wiener multicanal* (PDS-FWM) [158]. Ici, le PDS-FWM sera ajusté pour obtenir un score de DSP-IP identique dans chaque scénario (anéchoïque, $RT_{60} = 250ms$ et $RT_{60} = 500ms$).

Les résultats en terme de RSB-IP sont affichés dans le tableau 15.1. Nous observons que globalement ARC est plus performant que les autres méthodes, notamment quand les sources à séparer sont proches (30° entre le bruit et la source).

Conclusion de la partie IV

Deux méthodes de débruitage de la parole ont été introduites dans cette partie. La première fournit un algorithme monocanal hybride où l'impulsivité de chaque source est contrôlée par des exposants caractéristiques différents pour le bruit et la parole. Nous avons alors démontré qu'un tel modèle induisait une méthode relativement proche d'un filtrage de Wiener usuel mais pour laquelle la contribution de chaque source est pondérée différemment, en fonction de son exposant caractéristique. Cette technique fournit une base théorique où plusieurs améliorations empiriques ont été proposées dans les années 80 par Ephraim & Mallah [41]. Dans ce papier, il est fait usage d'un *a priori* uniquement sur la parole en supposant un mélange gaussien. Nous avons alors fait le lien entre notre modèle α -stable hybride et un cas particulier de *filtrage de Wiener paramétrique* (FWP) décrit dans [41]. Dans notre cas, les distributions *a priori* de la parole et du bruit sont cependant des distributions positives α -stables nommées *variables impulsionsnelles*. Elles décrivent exactement les variations du spectrogramme. Dans notre modélisation, nous avons supposé les distributions des *variables impulsionsnelles* indépendantes du temps et de la fréquence. Nous employons alors simplement la médiane pour les estimer. En outre, l'énergie du spectrogramme résumée dans les paramètres d'échelles a été lissée en considérant le log-spectrogramme de chaque source. Afin d'obtenir un algorithme plus flexible face à différents scénarios de signaux corrompus, les longueurs des filtres de lissages pour les paramètres d'échelles peuvent être différentes pour chaque source. Au final, l'algorithme baptisé *Multi-Alpha Denoising* (MAD) a permis de prouver qu'un tel choix de paramètre pour le FWP établissait un meilleur score perceptuel que les méthodes historiques de soustraction spectrale [22] ou estimant un *a priori* gaussien sur la parole [41]. Nous avons également implémenté MAD dans le cadre du projet KAMoulox visant à la restauration d'archives du centre de recherche ethnomusicologique (CREM). Nos écoutes ont pu montrer qu'un tel algorithme serait adapté pour le débruitage des données du CREM.

Dans un second temps, nous avons décrit un modèle hybride permettant le rehaussement de la parole d'un mélange multicanal. Tandis que le signal de parole est supposé gaussien, nous avons modélisé la partie résiduelle de l'observation comme étant α -stable. Dans cette approche, il est alors pris en compte que le modèle de parole, gaussien, doit rendre compte fidèlement du signal à extraire. Les covariances du signal de parole sont égales au produit d'un scalaire non-négatif représentant l'énergie du spectrogramme et d'une matrice de covariance spatiale dépendant uniquement de la fréquence [39]. Le bruit résiduel est quant à lui supposé être paramétrisé par une matrice diagonale ne dépendant que de la fréquence. Cela nous permet de modéliser un bruit pouvant être très permissif. La gaussianité conditionnelle des variables α -stables est alors exploitée de nouveau. Contrairement au modèle monocanal, les *variables impulsionsnelles* de la partie résiduelle sont estimées grâce à l'utilisation d'un algorithme de *Metropolis-Hastings*. Dans ce type de modélisation, la convergence est rapide sous certaines hypothèses liées aux échantillonnages de Gibbs [162]. L'estimation des paramètres gaussiens est obtenue en appliquant un algorithme de type *espérance-maximisation* (EM). Cette méthode appelée *Alpha Residual Component* (ARC) prouve son efficacité dans des scénarios de débruitage complexes.

Dans le cadre d'amélioration de la parole, il serait intéressant de proposer une méthode où les *exposants caractéristiques* seraient estimés au cours des itérations de l'algorithme. Durant nos expériences, nous avons remarqué que la distance entre les *exposants caractéristiques* apparaît plus importante que leur

valeur précise. Il serait également intéressant de considérer ces mêmes *exposants caractéristiques* comme dépendant de la fréquence. Des travaux prometteurs dans [93] ont déjà pris en compte un tel *exposant caractéristique* variant en fonction du temps et de la fréquence .

Conclusion générale

Bilan des contributions et discussion

Les quatre parties de ce mémoire, distinctes par leurs applications et la théorie qu'elles mettent en œuvre, ont permis d'esquisser le potentiel de la théorie α -stable en traitement du signal. Dans la première partie, nous avons rappelé des résultats théoriques généraux sur l'application des processus α -stables pour la *séparation de sources sonores* (SSS). En particulier, l'extension des processus gaussiens stationnaires au sens large aux processus α -harmonisables a fait la preuve de son utilité pour la SSS. Il a été notamment montré que l'hypothèse d'additivité des densités spectrales (fractionnaires) est davantage valable pour un exposant de $\alpha = 1.2$ que pour le cas gaussien ($\alpha = 2$). Nous avons également introduit deux techniques de localisation de sources acoustiques. L'une basée sur les moments d'ordre 2 pour les sources et l'autre sur la théorie α -stable et l'estimation de la covariation.

L'apport de cette première partie est également à visée pédagogique : réunir toutes les théories en séparation et localisation sur les processus α -stables. Pour autant que nous sachions, un tel catalogue n'était pas disponible jusqu'à présent. De plus, les mesures aléatoires forment un outil très peu connu mais pourtant pratique, pour peu que l'on s'intéresse à la modélisation de processus stochastiques pour le traitement du signal. Nous espérons qu'une telle analyse sera une inspiration pour le lecteur, afin de l'aider à développer son propre modèle pour le traitement du signal. Nous pouvons par exemple penser à l'étude des processus parcimonieux, classe dans laquelle sont inclus les processus α -stables. Le schéma serait donc en premier lieu une description des processus parcimonieux via une représentation utilisant une mesure aléatoire. Ensuite, une comparaison à l'état de l'art incluant les processus α -stables.

Dans la seconde partie, nous avons étudié l'application des processus α -stables à la localisation de sources sonores acoustiques. Il a été démontré qu'il existait une manière élégante et nouvelle d'utiliser les modèles α -stables non-gaussiens pour la localisation dans un milieu réverbérant. Nous avons introduit en premier lieu la *mesure spatiale discrète* (MSD) et ses premières utilisations dans l'état de l'art. Notamment, l'emploi de la MSD pour l'analyse en composantes indépendantes et la SSS. Nous avons expliqué qu'elle peut s'interpréter comme une carte de chaleur fournissant l'activité d'une source dans l'espace. L'intérêt alors du dernier chapitre est de montrer la flexibilité du modèle face à différents scénarios de localisation en milieu réverbérant. Nous pouvons appliquer n'importe quel modèle acoustique pour l'estimation de la MSD. Soulignons cependant que l'estimation de la MSD souffre de certains défauts. En particulier si α devient petit, les sources de faibles énergies ne sont pas convenablement localisées mais qu'en revanche, les sources de fortes énergies le sont plus précisément. Ce problème pourrait être résolu en employant une technique dite de *matching pursuit* : nous prenons α petit et retirons la source prédominante de la carte de chaleur. Nous augmentons α et continuons de soustraire ainsi chaque source. Comme les méthodes de localisation CLEAN ou RELAX, ce type d'algorithme peut avoir cependant le désavantage de nécessiter la connaissance *a priori* du nombre de sources à localiser.

Ensuite, la troisième partie a permis de décrire différents résultats sur de nouvelles techniques de filtrage pour les distributions multivariées α -stables. Il est à noter que le filtrage de processus α -stables n'avait jusqu'alors été considéré qu'en monocal (cas scalaire). Nous avons donc défini une nouvelle

manière de décrire des distributions symétriques multivariée α -stables et de générer des tirages d'une telle distribution. Cette décomposition contrôlée par deux objets appelés *mesure spatiale* et *spectre spatial* nous a permis d'esquisser deux algorithmes de séparation. Le premier estime tout d'abord le spectre spatial du mélange, qui est séparé pour aboutir aux spectres spatiaux des différentes sources pour la reconstruction. Le second algorithme est quant à lui une généralisation de la minimisation de l'erreur quadratique moyenne, aussi appelée filtrage de Wiener multicanal. Dans le cas α -stable, nous avons contourné le problème de non-existence de moment d'ordre 2 en utilisant la covariation, jouissant de propriétés similaires. Malgré ces deux nouvelles approches produisant de meilleurs résultats que l'état de l'art, le temps d'exécution reste relativement important. Notamment, la méthode d'estimation directe du spectre spatial reste lente.

Enfin, dans la dernière partie, nous avons explicité deux algorithmes de débruitage. Celui dans le cas monocanal apporte une interprétation théorique rigoureuse d'une amélioration empirique de filtres de Wiener décrite dans les années 1980. L'algorithme appelé *Multialpha Denoising Algorithm* (MAD) que nous avons proposé s'avère très rapide. Il a été mis en œuvre pour le débruitage des archives du CNRS — centre de recherche ethnomusicologique (CREM) avec MAD. En considérant un jeu de paramètres adapté, les résultats de débruitage sont très satisfaisants. L'inconvénient demeure la présence d'un bruit périodique provoqué par le sillon entrant en contact avec le vinyle. Une réflexion autour de la considération d'une troisième source de bruit rendant compte de cette dégradation est envisageable. Dans le cas multicanal, nous avons également fourni un algorithme pour le rehaussement de la parole. Ici la composante de parole est supposée gaussienne et le résidu suit une distribution *multivariée α -stable à contour elliptique* (α -MSCE). Les distributions α -MSCE sont paramétrées par une matrice de dispersion définie positive, analogues aux matrices de covariances dans le modèle gaussien. Dans le cadre d'un environnement bruité, notre algorithme *Alpha Residual Component* (ARC) est en moyenne plus performant que son homologue gaussien.

Perspectives

Nous avons résumé dans cette section plusieurs perspectives de recherche, découlant des travaux présentés dans ce manuscrit.

Modèle α -stable multicanal pour la séparation de sources sonores

Dans la partie III, le cadre de recherche était théorique et les évaluations ne portaient que sur la séparation de mélanges α -stables synthétiques. La motivation derrière un tel modèle est de pouvoir l'appliquer à des problèmes concrets en traitement du signal comme la *séparation de sources sonores* (SSS). Une direction intéressante serait la généralisation du modèle gaussien suivant proposé dans [39] :

$$\forall (f, t), \begin{cases} \mathbf{x}(f, t) = \sum_j \mathbf{y}_j(f, t) \\ \mathbf{y}_j(f, t) \sim \mathcal{N}_c^K \left(\underbrace{v_j(f, t)}_{\geq 0} \mathbf{R}_j(f) \right) \end{cases}, \quad (1)$$

où l'énergie est contenue dans le scalaire non-négatif $v_j(f, t)$ dépendant de la fréquence et du temps et l'information spatiale dans la matrice hermitienne $\mathbf{R}_j(f)$ définie positive appelée matrice de covariance spatiale, dépendant uniquement de la fréquence. Une généralisation de ce modèle dans le cas α -stable

serait alors la suivante :

$$\forall (f, t), \begin{cases} \mathbf{x}(f, t) = \sum_j \mathbf{y}_j(f, t) \\ \mathbf{y}_j(f, t) \sim S\alpha S_c^K \left(\underbrace{v_j(f, t)}_{\geq 0} \Gamma_j(f) \right) \end{cases} \quad (2)$$

avec $\Gamma_j(f)$, la *mesure spatiale* d'une distribution isotrope symétrique α -stable de dimension K . Nous pouvons alors imaginer plusieurs choix possibles pour l'estimation des sources $\mathbf{y}_j(f, t)$, connaissant $\mathbf{x}(f, t)$.

Estimation non-paramétrique

Nous souhaitons procéder à l'estimation des paramètres $\Gamma_j(f)$ et $v_j(f, t)$ à tour de rôle. Supposons connus chaque $\Gamma_j(f)$ et $\mathbf{x}(f, t)$. Nous pouvons par exemple discrétiser comme dans la partie II page 45 la forme particulière de l'*exposant de Lévy* des $\mathbf{y}_j(f, t)$ (noté ci-après $I_{j,ft}$) pour tout j, f, t

$$\forall \mathbf{u} \in \mathbb{C}^K \quad I_{j,ft}(\mathbf{u}) = |v_j(f, t)|^\alpha \int_{\boldsymbol{\theta} \in S^K} |\langle \mathbf{u}, \boldsymbol{\theta} \rangle|^\alpha \Gamma_j(f)(d\boldsymbol{\theta}) \quad (3)$$

afin d'estimer les $v_j(f, t)$. Malheureusement, nous n'avons qu'une seule observation par point temps-fréquence et par source. Comme dans la section 14.2.2, nous pouvons imaginer un lissage des paramètres particulier en fonction du type de source à séparer.

A contrario, supposons maintenant que les $v_j(f, t)$ et $\mathbf{y}_j(f, t)$ soient connus. Nous pouvons alors employer une technique proche de celle décrite dans la partie II page 45, en discrétisant à nouveau les intégrales. Cela est possible car

$$\forall j, f, t, \quad \frac{\mathbf{y}_j(f, t)}{v_j^{1/\alpha}(f, t)} \sim S\alpha S_c^K(\Gamma_j(f)). \quad (4)$$

Cependant, de tels modèles dit non-paramétrique pour la densité spatiale Γ_j ou pour les scalaires v_j peut s'avérer trop général pour la SSS. Cela peut en effet complexifier la tâche d'obtenir un modèle efficace pour la séparation. Une autre approche serait donc de prendre en compte un modèle paramétrique.

Modèle paramétrique de $v_j(f, t)$

Les stratégies suivantes d'estimation des paramètres $v_j(f, t)$ sont possibles :

- une factorisation matricielle non-négative (NMF) $\mathbf{V}_j \simeq \mathbf{W}_j \mathbf{H}_j$ où $\mathbf{V}_j \triangleq \{v_j(f, t)\}_{f,t} \in \mathbb{R}_+^{F \times T}$, $\mathbf{W}_j \triangleq \{w_j(f, l)\}_{f,l} \in \mathbb{R}_+^{F \times L}$ et $\mathbf{H}_j \triangleq \{h_j(l, t)\}_{l,t} \in \mathbb{R}_+^{L \times T}$;
- une estimation via un réseau de neurones profond. En entrée, les observations $\mathbf{x}(f, t)$ et *mesures spatiales* $\Gamma_j(f)$. En sortie, les paramètres $v_j(f, t)$.

Modèle paramétrique de $\Gamma_j(f)$

Nous esquissons ici deux grandes familles paramétriques envisageables pour la densité spatiale. La première consisterait à décomposer les *mesures spatiales* $\Gamma_j(f)$ dans une base d'harmoniques sphériques. Elles généralisent la transformée de Fourier mais dans un cadre multidimensionnel. Nous savons en particulier qu'elles sont parfois utilisées en localisation de sources audio [84] et améliorent les performances des algorithmes. Par conséquent, nous pourrions paramétrer les *mesures spatiales* comme des fonctions polynomiales décomposées dans la base d'harmoniques sphériques, contre une mesure de Lebesgue.

Les harmoniques sphériques généralisent la transformée de Fourier, dans un espace de dimension supérieure à 2.

Une autre approche paramétrique serait de supposer, comme dans les expériences du chapitre 12, que les $\Gamma_j(f)$ peuvent être représentés par des densités de probabilités. Par exemple, les distributions de Von-Mises Fisher, ou encore les distributions de Kent [91] généralisant les distributions de Von-Mises Fisher.

Modèle α -stable multicanal pour la déréverbération

Un prolongement original du travail présenté ici serait d'exploiter la théorie vectorielle α -stable pour aborder le problème de la déréverbération [71, 111, 123], dans le cas multicanal. L'idée serait d'envisager l'observation comme une somme de L échos indépendants, et de procéder à leur séparation pour ne garder que le chemin direct. En pratique, notons toujours $X \in \mathbb{C}^{F \times T \times K}$ la TFCT de l'observation, où F, T, K correspondent respectivement au nombre de bande fréquentielle, de trame temporelle et de microphones. Pour un point temps-fréquence (f, t) , la $k^{\text{ème}}$ composante $x_k(f, t)$ du vecteur $x(f, t)$ est donnée comme suit :

$$x_k(f, t) = \sum_l s_l(f, t) H_{kl}(f) \quad (5)$$

où $H_{kl}(f)$ est la réponse fréquentielle du chemin acoustique de la source virtuelle s_l vers x_k . L'objectif deviendrait d'associer les différents membres constitutifs de l'observation en (5) avec son spectre spatial χ présenté dans la section 9.2. Comme nous savons estimer χ en fonction des observations α -stables, nous pourrions a fortiori estimer les $H_k(f)$, qui seraient supposés les mêmes sur l'ensemble des observations.

Débruitage hybride en monocanal : estimation des exposants caractéristiques et dépendance fréquentielle et/ou temporelle

Au cours des recherches effectuées dans la partie IV, nous avons pu observer que l'écart entre les deux *exposants caractéristiques* du bruit et de la parole, notés respectivement α_s et α_n , était très important. Nous avons cependant supposé que ces constantes étaient indépendantes de la fréquence et du temps. De plus, le choix des *exposants caractéristiques* était effectué de manière arbitraire. Précisons alors que l'extension de modèles hybrides où $\alpha_s(f, t)$ et $\alpha_n(f, t)$ dépendraient de la fréquence et du temps pourrait se justifier par les points suivants :

- la loi d'une *variable impulsionnelle* est théoriquement explicitée en fonction de son exposant caractéristique. Supposons par exemple que $\alpha_s(f, t)$ soit l'*exposant caractéristique* au point temps-fréquence (f, t) . Dans ce cas, les variables impulsionnelles auraient une accroche aux données plus importante. Le modèle est alors plus flexible et le filtrage de Wiener paramétrique en (13.23) davantage justifié.
- un voisinage où l'on suppose le même $\alpha_s(f, t)$ pourrait également permettre de proposer une manière élégante de calculer les *variables impulsionnelles*. Nous pouvons avoir pour un certain voisinage de taille $\Delta_f \times \Delta_t$ des exposants $\alpha_s(f, t)$ identiques. Nous pouvons alors prendre un opérateur statistique (médiane, moyenne, variance etc.) d'un tirage de variable

$$\varphi_s(f, t) \sim \mathcal{P}_{\alpha_s(f, t)} S \left(2 \left[\cos \left(\frac{\pi \alpha_s(f, t)}{4} \right) \right]^{2/\alpha_s(f, t)} \right). \quad (6)$$

Pour l'estimation des $\alpha_s(f, t)$, $\alpha_n(f, t)$, il est possible d'employer une méthode de régression linéaire via la relation ci-dessous [136] :

$$\forall \theta \in \mathbb{C}, \ln [I_{s(f, t)}(\theta)] = \alpha_s(f, t) [\ln(\sigma_s(f, t)) + \ln(|\theta|)] \quad (7)$$

où $I_{s(f, t)}(\theta)$ est l'exposant de Lévy évalué à la *fréquence caractéristique* $\theta \in \mathbb{C}$. Des travaux dans [93] ont prouvé avec succès que faire dépendre α de la fréquence permet d'obtenir des résultats en SSS honorables.

Mesure spatiale discrète pour l'identification de la forme d'une salle

L'utilisation d'échos acoustiques et de source virtuelles pour la reconstruction géométrique d'une salle est un sujet déjà traité en 2013 [37]. Étant donné que la *mesure spatiale discrète* nous informe de l'énergie de chaque source dans une salle, il est possible qu'en étendant la localisation hors des parois d'une salle, nous puissions identifier les sources virtuelles. Supposons par exemple avoir une seule source et plusieurs microphones. Alors, en identifiant les différentes symétries entre chaque source virtuelle, il serait possible de reconstituer la géométrie d'une salle. Le problème reste ouvert quant au choix du modèle acoustique pour l'estimation de la mesure spatiale discrète « étendue » (modèle champ proche, modèle champ lointain etc.).

Index

- algorithme
 - espérance-maximisation, 31
 - Markov chain Monte-Carlo*, 39, 117, 128
 - Metropolis-Hastings*, 39, 128
- α -spectrogramme, 27
- analyse en composantes indépendantes, 50
- carte thermique, 59
- champ proche, 63
- covariance, 16, 24, 98, 114, 115
- covariation, 43, 92
- covariation, norme de, 94
- détecteur d'activité vocale, 31, 111
- densité spectrale de puissance, 18, 25, 111
- distribution
 - α -stable, 10
 - de Cauchy, 5, 10
 - de Laplace, 5
 - de Von-Mises Fisher, 78, 100
 - gaussienne, 5, 10
 - infiniment divisible, 13
 - multivarié α -stable à contours elliptique, 113, 115
 - positive α -stable, 15, 118
 - symétrique α -stable, 10, 11, 43
 - symétrique isotrope α -stable, 14, 47, 73, 81, 112, 119
- divergence
 - β -divergence, 56
 - Itakura-Saito, 27, 31, 37, 65
 - Kullback-Leibler, 27
- espérance-maximisation, 125
- exposant
 - caractéristique, 10, 11, 14, 65, 101, 118
 - de Lévy, 13, 54, 56, 61, 85
- filtrage
 - α -Wiener, 26, 38, 75, 112
 - de Wiener, 91, 98
 - de Wiener paramétrique, 111, 118
- filtre de mélange, 60
- fléau de la dimension, 61, 85
- fonction caractéristique, 4, 11, 14, 47, 83, 98, 113
- fréquence caractéristique, 54, 61
- lois fini-dimensionnelles, 7
- matrice
 - circulante, 18
 - de dispersion, 113, 115
 - de Toeplitz, 18
- maximum de vraisemblance, 31, 126
- mesure
 - de contrôle, 75
 - de contrôle circulaire, 7, 20, 76
 - spatiale, 74–77, 82, 84, 92, 95, 98, 100, 114
 - spatiale discrète, 49, 54, 60, 69
- mesure aléatoire, 75
 - isotrope, 21, 75
 - symétrique α -stable, 20
- moments d'ordres faibles fractionnaires, 38, 119
- moyenne, 16, 24
- nonnegative matrix factorization*, 34, 56, 115, 125
- paramètre d'échelle, 11, 14, 119
- paramètre d'asymétrie, 11
- paramètre de localisation, 11
- processus
 - α -harmonisable, 21
 - α -harmonisable, 25, 75, 112
 - gaussien, 16, 23
 - parcimonieux, 12
 - stationnaire, 17
 - stationnaire au sens large, 17, 18, 24
 - symétrique α -stable, 20
- puissance de réponse directionnelle, 43, 64
- réponse
 - en fréquence, 62

- impulsionnelle, 62, 130
- sketching, 54, 61
- soustraction spectrale, 31, 111
- spectre spatial, 75, 77, 82, 84, 90, 95, 100
- spectrogramme, 19
- transformée de Fourier
 - à court terme, 19, 25, 109
 - à temps discret, 18
- variable impulsionnelle, 113, 114, 117, 118, 129
- vecteur de pointage, 60, 62, 63

Bibliographie

- [1] M. Abramowitz and I.A. Stegun. *Handbook of mathematical functions : with formulas, graphs, and mathematical tables*, volume 55. Courier Corporation, 1965.
- [2] R. Adler, R. Feldman, and M. Taqqu. *A practical guide to heavy tails : statistical techniques and applications*. Springer Science & Business Media, 1998.
- [3] R. J. Adler and J. E. Taylor. *Random fields and geometry*. Springer Science & Business Media, 2009.
- [4] M. Akhtar, J. Epps, and E. Ambikairajah. Signal processing in sequence analysis : advances in eukaryotic gene prediction. *IEEE journal of selected topics in signal processing*, 2(3) :310–321, 2008.
- [5] S. Amari, A. Cichocki, and H.H. Yang. A new learning algorithm for blind signal separation. In *Advances in neural information processing systems*, pages 757–763, 1996.
- [6] David Applebaum. *Lévy processes and stochastic calculus*. Cambridge university press, 2009.
- [7] S. Arberet, R. Gribonval, and F. Bimbot. A robust method to count and locate audio sources in a stereophonic linear instantaneous mixture. In *Proc. Int. Workshop on Independent Component Analysis and Signal Separation (ICA2006)*, pages 536–543. Springer, 2006.
- [8] E. Artin. *The gamma function*. Courier Dover Publications, 2015.
- [9] Mordecai Avriel. *Nonlinear programming : analysis and methods*. Courier Corporation, 2003.
- [10] R. Badeau. *Méthodes à haute résolution pour l'estimation et le suivi de sinusoïdes modulées. Application aux signaux de musique*. PhD thesis, Télécom ParisTech, 2005.
- [11] R. Badeau. Gaussian modeling of mixtures of non-stationary signals in the time-frequency domain (hr-nmf). In *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, pages 253–256, 2011.
- [12] R. Badeau. Unified stochastic reverberation modeling. In *Proc. European Signal Processing Conf. (EUSIPCO)*, pages 2175–2179. IEEE, 2018.
- [13] R. Badeau and A. Liutkus. Proof of Wiener-like linear regression of isotropic complex symmetric alpha-stable random variables. 2014.
- [14] P.A. Barrett. Voice activity detector, May 9 2000. US Patent 6,061,647.
- [15] H.H. Bauschke and J.M. Borwein. Joint and separate convexity of the Bregman distance. In *Studies in Computational Mathematics*, volume 8, pages 23–36. Elsevier, 2001.
- [16] R. Bellman. *Dynamic programming*. Courier Corporation, 2013.
- [17] A. Belouchrani, K. Abed-Meraim, J-F. Cardoso, and E. Moulines. A blind source separation technique using second-order statistics. *IEEE Trans. Signal Process.*, 45(2) :434–444, 1997.
- [18] L. Benaroya, F. Bimbot, and R. Gribonval. Audio source separation with a single sensor. *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, 14(1) :191–199, 2006.

- [19] J. Benesty, J. Chen, and Y. Huang. Noncausal (frequency-domain) optimal filters. pages 115–137, 2008.
- [20] M. Berouti, R. Schwartz, and J. Makhoul. Enhancement of speech corrupted by acoustic noise. In *Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, volume 4, pages 208–211. IEEE, 1979.
- [21] S. Bidon, O. Besson, and J-Y. Touneret. Algorithme STAP pour fouillis non stationnaire en distance. In *21° Colloque GRETSI, Troyes, FRA, 11-14 septembre 2007*. GRETSI, Groupe d’Etudes du Traitement du Signal et des Images, 2007.
- [22] S. Boll. Suppression of acoustic noise in speech using spectral subtraction. *IEEE Trans. Acoust., Speech, Signal Process.*, 27(2) :113–120, 1979.
- [23] S. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge university press, 2004.
- [24] M.S. Brandstein and H.F. Silverman. A robust method for speech signal time-delay estimation in reverberant rooms. In *Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, volume 1, pages 375–378. IEEE, 1997.
- [25] L.M. Bregman. The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming. *USSR computational mathematics and mathematical physics*, 7(3) :200–217, 1967.
- [26] R.G. Brown, P. YC. Hwang, et al. *Introduction to random signals and applied Kalman filtering*, volume 3. Wiley New York, 1992.
- [27] S. Cambanis, R. Keener, and G. Simons. On α -symmetric multivariate distributions. *Journal of Multivariate Analysis*, 13(2) :213–233, 1983.
- [28] J-F. Cardoso. Infomax and maximum likelihood for blind source separation. *IEEE Signal Process. Lett.*, 4(4) :112–114, 1997.
- [29] J-F. Cardoso. Blind signal separation : statistical principles. *Proc. IEEE*, 86(10) :2009–2025, 1998.
- [30] G. Celeux and G. Govaert. A classification EM algorithm for clustering and two stochastic versions. *Computational statistics & Data analysis*, 14(3) :315–332, 1992.
- [31] A.T. Cemgil, S.J. Godsill, P.H. Peeling, and N. Whiteley. The Oxford handbook of applied Bayesian analysis, no. isbn13 : 978-0-19-954890-3, ch. Bayesian statistical methods for audio and music processing, 2010.
- [32] P. Comon. Independent component analysis, a new concept? *Signal processing*, 36(3) :287–314, 1994.
- [33] P. Comon and C. Jutten. *Handbook of Blind Source Separation : Independent component analysis and applications*. Academic press, 2010.
- [34] R. Cont. Benoit Mandelbrot et la modélisation mathématique des risques financiers.
- [35] A.P. Dempster, N.M. Laird, and D.B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the royal statistical society. Series B (methodological)*, pages 1–38, 1977.
- [36] S. Doclo and M. Moonen. GSVD-based optimal filtering for single and multimicrophone speech enhancement. *IEEE Trans. Signal Process.*, 50(9) :2230–2244, 2002.
- [37] I. Dokmanić, R. Parhizkar, A. Walther, Y.M. Lu, and M. Vetterli. Acoustic echoes reveal room shape. *Proc. of the National Academy of Sciences*, 110(30) :12186–12191, 2013.
- [38] E.R. Dougherty. *Random processes for image and signal processing*. SPIE Optical Engineering Press, 1999.
- [39] N. Q. K. Duong, E. Vincent, and R. Gribonval. Under-determined reverberant audio source separation using a full-rank spatial covariance model. *IEEE Trans. Audio, Speech, Lang. Process.*, 18(7) :1830–1840, Sept. 2010.

- [40] Y.C. Eldar and G. Kutyniok. *Compressed sensing : theory and applications*. Cambridge University Press, 2012.
- [41] Y. Ephraim and D. Malah. Speech enhancement using a minimum-mean square error short-time spectral amplitude estimator. *IEEE Trans. Acoust., Speech, Signal Process.*, 32(6) :1109–1121, 1984.
- [42] Y. Ephraim and D. Malah. Speech enhancement using a minimum mean-square error log-spectral amplitude estimator. *IEEE Trans. Acoust., Speech, Signal Process.*, 33(2) :443–445, 1985.
- [43] T. Fabricius, P. Kidmose, and L.K. Hansen. Dynamic components of linear stable mixtures from fractional low order moments. In *Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, volume 6, pages 3957–3960. IEEE, 2001.
- [44] D. Farina, C. Févotte, C. Doncarli, and R. Merletti. Blind separation of linear instantaneous mixtures of nonstationary surface myoelectric signals. *IEEE Trans. Bio-Med. Eng.*, 51(9) :1555–1567, 2004.
- [45] W. Feller. *An introduction to probability theory and its applications*, volume 1. Wiley, New York, 1968.
- [46] A. Feuerverger and R.A. Mureika. The empirical characteristic function and its applications. *The annals of statistics*, pages 88–97, 1977.
- [47] C. Févotte, N. Bertin, and J.-L. Durrieu. Nonnegative matrix factorization with the Itakura-Saito divergence : With application to music analysis. *Neural computation*, 21(3) :793–830, 2009.
- [48] C. Févotte, O. Cappe, and A.T. Cemgil. Efficient Markov Chain Monte Carlo inference in composite models with space alternating data augmentation. In *Proc. IEEE/SP Workshop on Statistical Signal Processing*, pages 221–224. IEEE, 2011.
- [49] C. Févotte and J-F. Cardoso. Maximum likelihood approach for blind audio source separation using time-frequency Gaussian source models. In *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, pages 78–81. IEEE, 2005.
- [50] C. Fevotte and S.J. Godsill. A bayesian approach for blind separation of sparse sources. *IEEE Trans. Audio, Speech, Lang. Process.*, 14(6) :2174–2188, 2006.
- [51] C. Févotte and J. Idier. Algorithms for nonnegative matrix factorization with the β -divergence. *Neural computation*, 23(9) :2421–2456, 2011.
- [52] C. Févotte, E. Vincent, and A. Ozerov. Single-channel audio source separation with NMF : divergences, constraints and algorithms. In *Audio Source Separation*, pages 1–24. Springer, 2018.
- [53] N.I. Fisher, T. Lewis, and B. JJ. Embleton. *Statistical analysis of spherical data*. Cambridge university press, 1993.
- [54] D. FitzGerald, M. Cranitch, and E. Coyle. On the use of the beta divergence for musical source separation. 2009.
- [55] D. Fitzgerald, A. Liutkus, and R. Badeau. Projection-based demixing of spatial audio. *IEEE Trans. Audio, Speech, Lang. Process.*, 24(9) :1556–1568, 2016.
- [56] D. Fitzgerald, A. Liutkus, and R. Badeau. Projet—spatial audio separation using projections. In *Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, pages 36–40. IEEE, 2016.
- [57] Ionut Florescu. *Probability and stochastic processes*. John Wiley & Sons, 2014.
- [58] M. Fontaine, A. Liutkus, L. Girin, and R. Badeau. Explaining the parameterized Wiener filter with alpha-stable processes. In *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, 2017.

- [59] M. Fontaine, F-R. Stöter, A. Liutkus, U. Şimşekli, R. Serizel, and R. Badeau. Multichannel audio modeling with elliptically stable tensor decomposition. In *International Conference on Latent Variable Analysis and Signal Separation*, pages 13–23. Springer, 2018.
- [60] M. Fontaine, C. Vanwynsberghe, A. Liutkus, and R. Badeau. Scalable source localization with multichannel α -stable distributions. In *Proc. European Signal Processing Conf. (EUSIPCO)*, pages 11–15. IEEE, 2017.
- [61] M. Fontaine, C. Vanwynsberghe, A. Liutkus, and R. Badeau. Sketching for nearfield acoustic imaging of heavy-tailed sources. In *International Conference on Latent Variable Analysis and Signal Separation*, pages 80–88. Springer, 2017.
- [62] D.K. Freeman and I. Boyd. Voice activity detection, January 4 1994. US Patent 5,276,765.
- [63] P. Frost and T. Kailath. An innovations approach to least-squares estimation–part iii : Nonlinear estimation in white Gaussian noise. *IEEE Trans. Autom. Control*, 16(3) :217–226, 1971.
- [64] R. G. Gallager. Circularly-symmetric Gaussian random vectors. *preprint*, pages 1–9, 2008.
- [65] P. G. Georgiou, P. Tsakalides, and C. Kyriakakis. Alpha-stable modeling of noise and robust time-delay estimation in the presence of impulsive noise. *IEEE Trans. Multimedia*, 1(3) :291–301, 1999.
- [66] L.C. Godara and A. Cantoni. Uniqueness and linear independence of steering vectors in array space. *The Journal of the Acoustical Society of America*, 70(2) :467–475, 1981.
- [67] S. Godsill and E.E. Kuruoglu. Bayesian inference for time series with heavy-tailed symmetric α -stable noise processes. *Proc. Applications of heavy tailed distributions in economics, engineering and statistics*, 1999.
- [68] R. M. Gray et al. Toeplitz and circulant matrices : A review. *Foundations and Trends in Communications and Information Theory*, 2(3) :155–239, 2006.
- [69] JE Greenberg, PM Peterson, and PM Zurek. Intelligibility-weighted measures of speech-to-interference ratio and speech system performance. *The Journal of the Acoustical Society of America*, 94(5) :3009–3010, 1993.
- [70] J. Grifone. *Algèbre linéaire*. Cépaduès, 1990.
- [71] E.A.P. Habets. Single- and multi-microphone speech dereverberation using spectral enhancement, 2007.
- [72] EAP. Habets. Room impulse response (RIR) generator, 2008.
- [73] D. Havelock, S. Kuwano, and M. Vorländer. *Handbook of signal processing in acoustics*. Springer Science & Business Media, 2008.
- [74] H-G. Hirsch and D. Pearce. The aurora experimental framework for the performance evaluation of speech recognition systems under noisy conditions. In *ASR2000-Automatic Speech Recognition : Challenges for the new Millenium ISCA Tutorial and Research Workshop (ITRW)*, 2000.
- [75] J.A. Högbom. Aperture synthesis with a non-regular distribution of interferometer baselines. *Astronomy and Astrophysics Supplement Series*, 15 :417, 1974.
- [76] H. Hopf. Über die abbildungen der dreidimensionalen sphäre auf die kugelfläche. *Mathematische Annalen*, 104(1) :637–665, 1931.
- [77] D. Hsu, S. Kakade, T. Zhang, et al. A tail inequality for quadratic forms of subgaussian random vectors. *Electronic Communications in Probability*, 17, 2012.
- [78] P.J. Huber. Robust statistics. In *International Encyclopedia of Statistical Science*, pages 1248–1251. Springer, 2011.

- [79] R. Hunger. *An introduction to complex differentials and complex differentiability*. Munich University of Technology, Inst. for Circuit Theory and Signal Processing, 2007.
- [80] K. Itakura, Y. Bando, E. Nakamura, K. Itoyama, K. Yoshii, and T. Kawahara. Bayesian multichannel audio source separation based on integrated source and spatial models. *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, 26(4) :831–846, 2018.
- [81] G-J. Jang and T-W. Lee. A maximum likelihood approach to single-channel source separation. *Journal of Machine Learning Research*, 4(Dec) :1365–1392, 2003.
- [82] S. Janson. Stable distributions. *arXiv preprint arXiv :1112.0220*, 2011.
- [83] JH. Janssen. A method for the calculation of the speech intelligibility under conditions of reverberation and noise. *Acta Acustica united with Acustica*, 7(5) :305–310, 1957.
- [84] D.P. Jarrett, E. AP. Habets, and P.A. Naylor. 3d source localization in the spherical harmonic domain using a pseudointensity vector. pages 442–446, 2010.
- [85] E.T. Jaynes. *Probability theory : The logic of science*. Cambridge university press, 2003.
- [86] C. Jutten and A. Taleb. Source separation : from dusk till dawn. In *Proc. 2nd Int. Workshop on Independent Component Analysis and Blind Source Separation (ICA2000)*, pages 15–26, 2000.
- [87] O. Kallenberg. *Random measures*. Academic Press, 1983.
- [88] Olav Kallenberg. *Random measures, theory and applications*. Springer, 2017.
- [89] S. Kalluri and G. R. Arce. Adaptive weighted myriad filter algorithms for robust signal processing in/spl alpha/-stable noise environments. *IEEE Trans. Signal Process.*, 46(2) :322–334, 1998.
- [90] S. Kamath and P. Loizou. A multi-band spectral subtraction method for enhancing speech corrupted by colored noise. In *Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, volume 4, pages 44164–44164. Citeseer, 2002.
- [91] J.T. Kent. The Fisher-Bingham distribution on the sphere. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 71–80, 1982.
- [92] N. Keriven, A. Bourrier, R. Gribonval, and P. Pérez. Sketching for large-scale learning of mixture models. *Information and Inference : A Journal of the IMA*, 7(3) :447–508, 2017.
- [93] N. Keriven, A. Deleforge, and A. Liutkus. Blind source separation using mixtures of alpha-stable distributions. In *Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, pages 771–775. IEEE, 2018.
- [94] P. Kidmose. *Blind separation of heavy tail signals*. Institute of Mathematical Modelling, Technical University of Denmark, 2001.
- [95] P. Kidmose. Independent component analysis using the spectral measure for alpha-stable distributions. In *Proc. Intl. Workshop IEEE on Nonlinear Signal and Image Processing*, volume 400, 2001.
- [96] S. Kitić, N. Bertin, and R. Gribonval. Audio declipping by cosparseness hard thresholding. In *iTwist-2nd international-Traveling Workshop on Interactions between Sparse models and Technology*, 2014.
- [97] A. N. Kolmogorov and B. V. Gnedenko. *Limit distributions for sums of independent random variables*. Addison-Wesley, 1968.
- [98] I.A. Koutrouvelis. A goodness-of-fit test of simple hypotheses based on the empirical characteristic function. *Biometrika*, 67(1) :238–240, 1980.
- [99] E. Krätzel. Integral transformations of Bessel-type. *Generalized functions and operational calculus*, pages 148–155, 1979.
- [100] J. Kuelbs. A representation theorem for symmetric stable processes and stable measures on H . *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 26(4) :259–271, 1973.

- [101] B. Kunka, B. Kostek, M. Kulesza, P. Szczuko, and A. Czyzewski. Gaze-tracking-based audio-visual correlation analysis employing quality of experience methodology. *Intelligent Decision Technologies*, 4(3) :217–227, 2010.
- [102] D.D. Lee and H.S. Seung. Algorithms for non-negative matrix factorization. In *Advances in neural information processing systems*, pages 556–562, 2001.
- [103] S. Leglaive, R. Badeau, and G. Richard. Multichannel audio source separation with probabilistic reverberation priors. *IEEE Trans. Audio, Speech, Lang. Process.*, 24(12) :2453–2465, 2016.
- [104] S. Leglaive, R. Badeau, and G. Richard. Semi-blind Student’s t source separation for multichannel audio convolutive mixtures. In *Proc. European Signal Processing Conf. (EUSIPCO)*, pages 2259–2263. IEEE, 2017.
- [105] S. Leglaive, R. Badeau, and G. Richard. Separating time-frequency sources from time-domain convolutive mixtures using non-negative matrix factorization. In *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, pages 264–268. IEEE, 2017.
- [106] S. Leglaive, U. Şimşekli, A. Liutkus, R. Badeau, and G. Richard. Alpha-stable multichannel audio source separation. In *Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, pages 576–580. IEEE, 2017.
- [107] U. Leonhardt and T. Philbin. *Geometry and light : the science of invisibility*. Courier Corporation, 2010.
- [108] Sabin Lessard. *Processus stochastiques : cours et exercices corrigés*. Ellipses, 2014.
- [109] P. Lévy. Théorie de l’addition des variables aléatoires. *Gauthier-Villars*, 1937.
- [110] J. Li and P. Stoica. Efficient mixed-spectrum estimation with applications to target feature extraction. *IEEE Trans. Signal Process.*, 44(2) :281–295, 1996.
- [111] X. Li, S. Gannot, L. Girin, and R. Horaud. Multichannel identification and nonnegative equalization for dereverberation and noise reduction based on convolutive transfer function. *IEEE Trans. Audio, Speech, Lang. Process.*, 2018.
- [112] Y. Liang, G. Chen, SMR Naqvi, and J.A. Chambers. Independent vector analysis with multivariate Student’s t-distribution source prior for speech separation. *IEE Electronics Lett.*, 49(16) :1035–1036, 2013.
- [113] J. Lim. Evaluation of a correlation subtraction method for enhancing speech degraded by additive white noise. *IEEE Trans. Acoust., Speech, Signal Process.*, 26(5) :471–472, 1978.
- [114] J.S. Lim. *Speech enhancement*. Prentice-Hall, 1983.
- [115] J.S. Lim and A.V. Oppenheim. Enhancement and bandwidth compression of noisy speech. *Proceedings of the IEEE*, 67(12) :1586–1604, 1979.
- [116] T-H. Liu and Jerry M. Mendel. A subspace-based direction finding algorithm using fractional lower order statistics. *IEEE Trans. Signal Process.*, 49(8) :1605–1613, 2001.
- [117] A. Liutkus. *Processus gaussiens pour la séparation de sources et le codage informé*. PhD thesis, Télécom ParisTech, 2012.
- [118] A. Liutkus and R. Badeau. Generalized Wiener filtering with fractional power spectrograms. In *Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, pages 266–270. IEEE, 2015.
- [119] A. Liutkus, R. Badeau, and G. Richard. Gaussian processes for underdetermined source separation. *IEEE Trans. Signal Process.*, 59(7) :3155–3167, 2011.
- [120] A. Liutkus, D. Fitzgerald, and R. Badeau. Cauchy nonnegative matrix factorization. In *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, pages 1–5. IEEE, 2015.

- [121] A. Liutkus, D. Fitzgerald, Z. Rafii, B. Pardo, and L. Daudet. Kernel additive models for source separation. *IEEE Trans. Signal Process.*, 62(16) :4298–4310, 2014.
- [122] X. Ma and C.L. Nikias. Joint estimation of time delay and frequency delay in impulsive noise using fractional lower order statistics. *IEEE Trans. Signal Process.*, 44(11) :2669–2687, 1996.
- [123] W. Mack, S. Chakrabarty, F-R. Stöter, S. Braun, B. Edler, and E. Habets. Single-channel dereverberation using direct MMSE optimization and bidirectional LSTM networks. *Proc. Interspeech Conf.*, pages 1314–1318, 2018.
- [124] S. Mallat and Z. Zhang. Matching pursuit with time-frequency dictionaries. Technical report, Courant Institute of Mathematical Sciences New York United States, 1993.
- [125] B. Mandelbrot. The Pareto-Lévy law and the distribution of income. *International Economic Review*, 1(2) :79–106, May 1960.
- [126] R. Martin. Spectral subtraction based on minimum statistics. *Proc. European Signal Processing Conf. (EUSIPCO)*, 6 :8, 1994.
- [127] R. Martin. Noise power spectral density estimation based on optimal smoothing and minimum statistics. *IEEE Trans. Speech Audio Process.*, 9(5) :504–512, 2001.
- [128] E. Masry. Alpha-stable signals and adaptive filtering. *IEEE Trans. Signal Process.*, 48(11) :3011–3016, 2000.
- [129] A. M. Mathai, R. K. Saxena, and H. J. Haubold. *The H-function : theory and applications*. Springer Science & Business Media, 2009.
- [130] R. McAulay and M. Malpass. Speech enhancement using a soft-decision noise suppression filter. *IEEE Trans. Acoust., Speech, Signal Process.*, 28(2) :137–145, 1980.
- [131] C.S. Meijer. *Eine neue Erweiterung der Laplace-Transformation*. NV Noord-Hollandsche Uitgevers Maatschappij, 1941.
- [132] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *The journal of chemical physics*, 21(6) :1087–1092, 1953.
- [133] G. Miller. Properties of certain symmetric stable distributions. *Journal of Multivariate Analysis*, 8(3) :346–360, 1978.
- [134] N.M. Nasrabadi. Pattern recognition and machine learning. *Journal of electronic imaging*, 16(4) :049901, 2007.
- [135] F. Nielsen. k-mle : A fast algorithm for learning statistical mixture models. In *Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, pages 869–872. IEEE, 2012.
- [136] C.L. Nikias and M. Shao. *Signal processing with alpha-stable distributions and applications*. Wiley-Interscience, 1995.
- [137] J.P. Nolan. Multivariate elliptically contoured stable distributions : theory and estimation. *Computational Statistics*, 28(5) :2067–2089, 2013.
- [138] J.P. Nolan, A.K. Panorska, and J.H. McCulloch. Estimation of stable spectral measures. *Mathematical and Computer Modelling*, 34(9-11) :1113–1122, 2001.
- [139] E. Novak and K. Ritter. The curse of dimension and a universal method for numerical integration. In *Multivariate approximation and splines*, pages 177–187. Springer, 1997.
- [140] A.A. Nugraha, A. Liutkus, and E. Vincent. Multichannel music separation with deep neural networks. In *Proc. European Signal Processing Conf. (EUSIPCO)*, pages 1748–1752. IEEE, 2016.
- [141] A. Ozerov and C. Févotte. Multichannel nonnegative matrix factorization in convolutive mixtures for audio source separation. *IEEE Trans. Audio, Speech, Lang. Process.*, 18(3) :550–563, 2010.

- [142] V. Panayotov, G. Chen, D. Povey, and S. Khudanpur. Librispeech : an ASR corpus based on public domain audio books. In *Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, pages 5206–5210. IEEE, 2015.
- [143] J.C. Pascal. Formation de voies : détermination du vecteur de pointage. *Cours de Master*.
- [144] R. Penrose. A generalized inverse for matrices. In *Mathematical proceedings of the Cambridge philosophical society*, volume 51, pages 406–413. Cambridge University Press, 1955.
- [145] J-D. Polack. Modifying chambers to play billiards : the foundations of reverberation theory. *Acta Acustica united with Acustica*, 76(6) :256–272, 1992.
- [146] A.N. Popper and R.R. Fay. *Sound source localization*. Springer, 2005.
- [147] J. Rappaz and M. Picasso. *Introduction à l'analyse numérique*. Lausanne (Suisse) : Presses polytechniques et universitaires romandes, 1998.
- [148] P.N. Rathie, L.C. de SM Ozelim, and C.E.G. Otiniano. Exact distribution of the product and the quotient of two stable lévy random variables. *Communications in Nonlinear Science and Numerical Simulation*, 36 :204–218, 2016.
- [149] A.W. Rix, J.G. Beerends, M.P. Hollier, and A.P. Hekstra. Perceptual evaluation of speech quality (pesq)-a new method for speech quality assessment of telephone networks and codecs. In *Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, volume 2, pages 749–752. IEEE, 2001.
- [150] A. Rolet, V. Seguy, M. Blondel, and H. Sawada. Blind source separation with optimal transport non-negative matrix factorization. *arXiv preprint arXiv :1802.05429*, 2018.
- [151] W. Rudin. *Analyse réelle et complexe*. 1975.
- [152] W. Rudin. *Fourier analysis on groups*, ser. a wiley-interscience publication, 1990.
- [153] G. Samoradnitsky and M. Taqu. *Stable non-Gaussian random processes : stochastic models with infinite variance*, volume 1. CRC Press, 1994.
- [154] H. Sawada, H. Kameoka, S. Araki, and N. Ueda. Efficient algorithms for multichannel extensions of Itakura-Saito nonnegative matrix factorization. In *Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, pages 261–264. IEEE, 2012.
- [155] P.J. Schreier and L.L. Scharf. *Statistical signal processing of complex-valued data : the theory of improper and noncircular signals*. Cambridge University Press, 2010.
- [156] M.R. Schroeder. Frequency-correlation functions of frequency responses in rooms. *The Journal of the Acoustical Society of America*, 34(12) :1819–1823, 1962.
- [157] G. AF. Seber and A.J. Lee. *Linear regression analysis*, volume 329. John Wiley & Sons, 2012.
- [158] R. Serizel, M. Moonen, B. Van Dijk, and J. Wouters. Low-rank approximation based multichannel Wiener filter algorithms for noise reduction with application in cochlear implants. *IEEE Trans. Audio, Speech, Lang. Process.*, 22(4) :785–799, 2014.
- [159] X. Sheng and Y-H. Hu. Maximum likelihood multiple-source localization using acoustic energy measurements with wireless sensor networks. *IEEE Trans. Signal Process.*, 53(1) :44–53, 2005.
- [160] K. Siedenburg and M. Dörfler. Structured sparsity for audio signals. In *Proc. Conf. on Digital Audio Effects*, 2011.
- [161] U. Simsekli, H. Erdogan, S. Leglaive, A. Liutkus, R. Badeau, and G. Richard. Alpha-stable low-rank plus residual decomposition for speech enhancement. In *Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, 2018.
- [162] U. Şimşekli, A. Liutkus, and A.-T. Cemgil. Alpha-stable matrix factorization. *IEEE Signal Process. Lett.*, 22(12) :2289–2293, 2015.

- [163] P. Smaragdis, C. Fevotte, G. J. Mysore, N. Mohammadiha, and M. Hoffman. Static and dynamic source separation using nonnegative factorizations : A unified view. *IEEE Signal Process. Mag.*, 31(3) :66–75, 2014.
- [164] J. Sohn, N.S. Kim, and W. Sung. A statistical model-based voice activity detection. *IEEE Signal Process. Lett.*, 6(1) :1–3, 1999.
- [165] E.W. Stacy et al. A generalization of the gamma distribution. *The Annals of mathematical statistics*, 33(3) :1187–1192, 1962.
- [166] P. Stoica, R.L. Moses, et al. Spectral analysis of signals. 2005.
- [167] J.V. Stone. *Independent component analysis : a tutorial introduction*. MIT press, 2004.
- [168] B. W. Stuck and B. Kleiner. A statistical analysis of telephone noise. *Bell Labs Technical Journal*, 53(7) :1263–1320, 1974.
- [169] P. Tankov. *Financial modelling with jump processes*, volume 2. CRC press, 2003.
- [170] G. A. Tsihrintzis, P. Tsakalides, and C. L. Nikias. Spectral methods for stationary harmonizable alpha-stable processes. In *Proc. European Signal Processing Conf. (EUSIPCO)*, pages 1–4. IEEE, 1998.
- [171] G.A. Tsihrintzis and C.L. Nikias. Fast estimation of the parameters of alpha-stable impulsive interference. *IEEE Trans. Signal Process.*, 44(6) :1492–1503, 1996.
- [172] M. Unser and P.D. Tafti. *An introduction to sparse stochastic processes*. Cambridge University Press, 2014.
- [173] B.D. Van Veen and K.M. Buckley. Beamforming : A versatile approach to spatial filtering. *IEEE ASSP Mag.*, 5(2) :4–24, 1988.
- [174] E. Vincent and D.R. Campbell. Roomsmove. www.irisa.fr/metiss/members, 2008.
- [175] E. Vincent, M.G. Jafari, S.A. Abdallah, M.D. Plumbley, and M.E. Davies. Probabilistic modeling paradigms for audio source separation. In *Machine Audition : Principles, Algorithms and Systems*, pages 162–185. IGI Global, 2011.
- [176] E. Vincent, A. Sini, and F. Charpillat. Audio source localization by optimal control of a mobile robot. In *Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, pages 5630–5634. IEEE, 2015.
- [177] Y. Wang, J. Li, P. Stoica, M. Sheplak, and T. Nishida. Wideband relax and wideband clean for aeroacoustic imaging. *The Journal of the Acoustical Society of America*, 115(2) :757–767, 2004.
- [178] R. Weron et al. On the chambers-mallows-stuck method for simulating skewed stable random variables. *Statistics and Probability Letters*, 28(2) :165–171, 1996.
- [179] N. Wiener. *Extrapolation, interpolation, and smoothing of stationary time series*, volume 7. MIT press Cambridge, MA, 1949.
- [180] C. K. Williams and C. E. Rasmussen. Gaussian processes for machine learning. *the MIT Press*, 2(3) :4, 2006.
- [181] E.G. Williams. *Fourier acoustics : sound radiation and nearfield acoustical holography*. Elsevier, 1999.
- [182] V. Witkovský. Computing the distribution of a linear combination of inverted gamma variables. pages 79–80, 2001.
- [183] E. M. Wright. The asymptotic expansion of the generalized hypergeometric function. *Journal of the London Mathematical Society*, 1(4) :286–293, 1935.
- [184] E.M. Wright. The asymptotic expansion of the generalized Bessel function. 2(1) :257–270, 1935.

- [185] C. Zhang, D. Florêncio, D.E. Ba, and Z. Zhang. Maximum likelihood sound source localization and beamforming for directional microphone arrays in distributed meetings. *IEEE Trans. Multimedia*, 10(3) :538–548, 2008.
- [186] L. Zhang, X. Zhang, X. Zou, and G. Min. An improved bayesian NMF-based speech enhancement method using multivariate laplace distribution. In *Proc. Intl. Conf. Wireless Communications and Signal Processing*, pages 1–6. IEEE, 2014.
- [187] I. Ziskind and M. Wax. Maximum likelihood localization of multiple sources by alternating projection. *IEEE Trans. Acoust., Speech, Signal Process.*, 36(10) :1553–1560, 1988.
- [188] V. M. Zolotarev. *One-dimensional stable distributions*, volume 65. American Mathematical Soc., 1986.
- [189] D.N. Zotkin and R. Duraiswami. Accelerated speech source localization via a hierarchical search of steered response power. *IEEE Trans. Speech Audio Process.*, 12(5) :499–508, 2004.
- [190] J.M. Zumer, H.T. Attias, K. Sekihara, and S.S. Nagarajan. A probabilistic algorithm integrating source localization and noise suppression of MEG and EEG data. In *Proc. Intl. Conf in Neural Information Processing Systems*, pages 1625–1632, 2007.