



**HAL**  
open science

# Restauration de sources radioastronomiques en milieu radioélectrique hostile : Implantation de détecteurs temps réel sur des spectres dynamiques.

Cedric Dumez-Viou

► **To cite this version:**

Cedric Dumez-Viou. Restauration de sources radioastronomiques en milieu radioélectrique hostile : Implantation de détecteurs temps réel sur des spectres dynamiques.. Sciences de l'ingénieur [physics]. Université d'Orléans, 2007. Français. NNT : . tel-00319939

**HAL Id: tel-00319939**

**<https://theses.hal.science/tel-00319939>**

Submitted on 9 Sep 2008

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

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



Thèse présentée à l'Université d'Orléans  
pour obtenir le grade de  
**Docteur de l'Université d'Orléans**

Discipline : Systèmes électroniques de traitement de l'information

Restauration de sources radioastronomiques  
en milieu radioélectrique hostile :  
Implantation de détecteurs temps réel  
sur des spectres dynamiques.

Cedric DUMEZ-VIOU

<Cedric.Dumez-Viou@obs-nancay.fr>

Cette thèse a été soutenue le 28 septembre 2007. Le jury était constitué de :

Présidente du jury	Élisabeth Lallier-Vergès	Directrice de recherche CNRS à l'ISTO Directrice de la Fédération STUC
Rapporteur	Christian Faye	Maître de conférences à l'ENSEA, Cergy-Pontoise
Rapporteur	Jean-Yves Tournet	Professeur à l'Université de Toulouse III
Directeur de thèse	Rachid Harba	Professeur à l'Université d'Orléans
Coencadrant	Pierre Colom	Astronome-Adjoint à l'Observatoire de Paris-Meudon
Membre du jury	Philippe Ravier	Maître de conférences à l'Université d'Orléans
Membre du jury	Laurent Denis	Ingénieur de recherche CNRS à l'Observatoire de Paris, Station de radioastronomie de Nançay
Membre du jury	Philippe Zarka	Directeur de recherche CNRS à l'Observatoire de Paris-Meudon



---

# REMERCIEMENTS

---

Je remercie d'abord mes directeurs de thèse, Rachid Harba, professeur au LESI (Université d'Orléans) et Pierre Colom, astronome au LESIA (Observatoire de Paris). Je tiens à remercier Rodolphe Weber, enseignant-chercheur au LESI pour m'avoir donné l'opportunité de travailler sur ce sujet.

J'exprime aussi toute ma sympathie à Christian Fayes et Jean-Yves Tourneret qui ont acceptés d'être les rapporteurs de mon travail. Leurs remarques ont permis d'améliorer grandement la qualité de ce mémoire. J'associe également l'ensemble du jury à mes remerciements.

J'ai vraiment à cœur de remercier Andrée Coffre (Ingénieur de recherche au CNRS), Laurent Denis (Ingénieur de recherche au CNRS) et Philippe Ravier (Enseignant-chercheur au LESI) pour m'avoir permis de mener à bien ces travaux. Rien n'aurait été possible sans leur soutien, leurs remarques et leurs conseils.

Je remercie encore Pierre Colom, Alain Lecacheux, Philippe Ravier et Rodolphe Weber pour l'aide qu'ils m'ont apporté lors de la relecture de mon mémoire.

Mon travail a été financé par l'Observatoire de Paris et les Fonds de Solidarité Européens. J'ai été accueilli à l'Unité Scientifique de la Station de Nançay (Laboratoire de l'Observatoire Paris), Unité de Service et de Recherche Associée au CNRS no. 704. Je remercie les personnes qui ont mis en œuvre les moyens financiers, logistiques et techniques pour assurer l'avancement de mes travaux.

Je tiens à citer la société ALSE et en particulier Ludovic Bacquart qui, sur un cahier des charges minimaliste, a su produire en VHDL un espace de développement souple et agréable.

Je souhaite exprimer ma reconnaissance envers les personnes qui m'ont conseillé et guidé lors de mes observations. Sur le Grand Radiotélescope de Nançay, j'ai bénéficié de l'aide de Laurence Alsac, Dominique Aubry, Pierre Colom, Éric Gérard, Jean-Michel Martin et bien sûr, des Observateurs. Sur le Réseau Décamétrique, j'ai eu recours à l'aide de Laurent Denis, que je remercie encore, et à l'expertise de Philippe Zarka lors de l'observation des S-bursts joviens.

Je remercie toute l'équipe de l'Électronique, en particulier Yvette et Michel Chapuis pour avoir sérieusement étoffé mes connaissances en RF et Patricia pour son aide lors du maquetage. Suit évidemment ma reconnaissance à la Mécanique qui m'a permis de réaliser ce qui n'aurait pu rester que concepts.

Je présente ma sympathie à Emmanuel Thetas pour ses talents et sa ténacité lors de la résolution de problèmes informatiques. J'en profite pour saluer tous les informaticiens de l'ombre qui codent du Logiciel Libre.

Je salue les copains de l'ESPEO et les thésards : Marie-Line, Véro et Mickaël, Ivan et les autres. Je remercie encore Stéphanie Bretteil et Gabriel Aufort pour avoir initié la

création du canevas L<sup>A</sup>T<sub>E</sub>X du document que vous tenez entre vos mains.

Merci à l'équipe des cuisines de la station pour ses copieux et délicieux repas (en particulier les pommes de terre rôties). Merci aux jeunes de la station (et au(sic) moins jeune(sic)) pour la bonne ambiance et les barbecues au bord de l'étang. Merci aux copains-copines qui m'ont autorisé à m'aérer l'esprit lorsque nous partagions plongées, apnées ou fichiers informatiques distrayants (1.5 Go de vidéos et 475 Mo de photos et diaporamas comptabilisés sur 14 mois, soit 4.8 Mo par jour. Est-ce une des raisons de l'essor de l'ADSL?).

Je remercie mes grands parents, mes parents qui m'ont aidé à me construire, mon frère pour tout les moments passés ensemble, mes parents-hôtes pour l'ouverture sur le monde qu'ils m'ont apporté. Finalement, je remercie mon épouse pour sa confiance et son attention, Margaux pour sa vivace joie de vivre et Antonin pour sa sérénité de nouveau né.

---

# TABLE DES MATIÈRES

---

<b>Table des matières</b>	<b>6</b>
<b>Table des figures</b>	<b>9</b>
<b>Liste des tableaux</b>	<b>11</b>
<b>Glossaire</b>	<b>13</b>
<b>Introduction</b>	<b>17</b>
<b>1 Introduction à la Radioastronomie</b>	<b>19</b>
1.1 Sources étudiées . . . . .	19
1.1.1 Fenêtres de transmission . . . . .	19
1.1.2 Procédés d'émission . . . . .	20
1.1.3 Bandes protégées . . . . .	23
1.2 Zoologie des interférences radioélectriques . . . . .	25
1.3 Antennes et radiotélescopes . . . . .	28
1.3.1 Types d'antennes . . . . .	28
1.3.2 Télescopes à antenne unique . . . . .	28
1.3.3 Interféromètre . . . . .	33
1.4 Instrumentation spécifique au radiotélescope à antenne unique . . . . .	36
1.4.1 Mesures et calibrations . . . . .	36
1.4.2 Récepteur numérique de test . . . . .	42
<b>2 Détection en puissance sur des spectres dynamiques</b>	<b>47</b>
2.1 Position du problème . . . . .	48
2.1.1 Contraintes opérationnelles . . . . .	48
2.1.2 Problème de détection . . . . .	50
2.1.3 Critère de décision . . . . .	51
2.1.4 Calcul de seuil sur une loi du Chi2 . . . . .	53
2.2 Estimateurs de moyenne . . . . .	56
2.2.1 Critères de comparaison . . . . .	57
2.2.2 Moyenne arithmétique . . . . .	60
2.2.3 Médiane . . . . .	63
2.2.4 Médiane corrigée . . . . .	67
2.2.5 Moyenne tronquée . . . . .	74
2.2.6 Moyenne tronquée symétrique . . . . .	80

2.2.7	Moyenne du Petit Poucet	85
2.2.8	Moyenne de classe	89
2.2.9	Filtre passe-bas IIR	94
2.2.10	Filtre passe-bas IIR seuillé	98
2.3	Comparatif des estimateurs	103
<b>3</b>	<b>Applications : blanking commandé en puissance</b>	<b>107</b>
3.1	Parasites large bande intermittents : Radar au NRT	107
3.1.1	Contraintes	109
3.1.2	Algorithme	109
3.1.3	Résultats	115
3.1.4	Conclusion	116
3.2	Parasites à bande étroite intermittents : TFDMA au NRT	123
3.2.1	Contraintes	124
3.2.2	Algorithme	125
3.2.3	Résultats	130
3.2.4	Conclusion	130
3.3	Parasites bandes étroites continus : AM au NDA	133
3.3.1	Contraintes	133
3.3.2	Algorithmes	135
3.3.3	Conclusion	140
<b>4</b>	<b>Détection de signaux brefs et intenses en milieu parasité</b>	<b>149</b>
4.1	Caractéristiques des signaux	150
4.2	Algorithme	151
4.2.1	Périodogramme moyenné	151
4.2.2	Calcul de log2 en virgule fixe et FFT-2D	153
4.2.3	Lissage du plan u-v et atténuation des RFI	153
4.2.4	Transformée de Radon	153
4.2.5	Calcul d'index	157
4.2.6	Implantation matérielle	157
4.3	Résultats	158
4.4	Perspectives	160
	<b>Conclusion</b>	<b>163</b>
	<b>A Rapport d'implantation en FPGA du "blanqueur" radar</b>	<b>167</b>
	<b>B Calcul de log2 en virgule fixe</b>	<b>169</b>
	<b>C Table d'allocation des fréquences US</b>	<b>181</b>
	<b>D Outils utilisés</b>	<b>183</b>
	<b>Bibliographie</b>	<b>184</b>
	<b>Résumé</b>	<b>190</b>

---

# TABLE DES FIGURES

---

1.1	Atténuation de l’atmosphère . . . . .	20
1.2	Émission radio à 408 MHz . . . . .	22
1.3	Interférence à bande étroite intermittente . . . . .	26
1.4	Interférence à large bande . . . . .	26
1.5	Interférence à large bande intermittente . . . . .	26
1.6	Interférence continue à bande étroite . . . . .	27
1.7	Interférence continue à bande large . . . . .	27
1.8	Grands instruments de Nançay . . . . .	29
1.9	Fonctionnement d’un réseau phasé . . . . .	30
1.10	Chaîne de réception simplifiée d’un télescope à antenne unique . . . . .	31
1.11	Radiotélescope Décimétrique de Nançay (NRT) . . . . .	32
1.12	Réseau Décamétrique de Nançay (NDA) . . . . .	33
1.13	Chaîne de réception simplifiée d’un télescope interféromètre . . . . .	34
1.14	Antennes NS du NRH . . . . .	34
1.15	Configuration du Radiohéliographe de Nançay (NRH) . . . . .	35
1.16	Température d’antenne en fonction de la fréquence . . . . .	39
1.17	Sources noyées dans du bruit . . . . .	40
1.18	Limite de détection d’une source . . . . .	41
1.19	Structure du récepteur RDH . . . . .	42
1.20	Test comparatifs des ADCs . . . . .	43
1.21	Variance de Allan du récepteur RDH . . . . .	44
2.1	Temps différé <i>versus</i> temps réel . . . . .	49
2.2	Illustration de la $P_{fa}$ et $P_{nd}$ . . . . .	51
2.3	DSP de la loi du $\chi^2$ . . . . .	54
2.4	Détermination du seuil de “blanking” $S$ . . . . .	56
2.5	Effet d’un parasite additif sur la densité de probabilité . . . . .	58
2.6	Densité de probabilité d’une loi en $\chi^2$ corrompue par une loi de Dirac . . . . .	59
2.7	Estimateur de moyenne par moyenne arithmétique . . . . .	61
2.8	Moyenne et écart type de l’estimateur $\hat{\mu}_{MN}$ en présence d’un parasite . . . . .	62
2.9	Estimateur de moyenne par médiane . . . . .	65
2.10	Moyenne et écart type de l’estimateur $\hat{\mu}_{MD}$ en présence d’un parasite . . . . .	66
2.11	Fonction correctrice $K(\nu)$ pour $\nu = 2^{1 \dots 10}$ . . . . .	68
2.12	Fonction correctrice $K(\nu, N)$ . . . . .	69
2.13	Erreur relative de $K(\nu)$ par rapport à $K(\nu, N)$ . . . . .	70
2.14	Estimateur de moyenne par médiane corrigée ( $K(\nu, N)$ ) . . . . .	71



2.15	Estimateur de moyenne par médiane corrigée ( $K(\nu)$ ) . . . . .	72
2.16	Moyenne et écart type de l'estimateur $\hat{\mu}_{MD_{\nu,N}}$ en présence d'un parasite . . . . .	73
2.17	Densité de probabilité pour le calcul de la moyenne tronquée . . . . .	74
2.18	Densité de probabilité $f_{\delta,N}(x)$ . . . . .	75
2.19	Estimateur de moyenne par moyenne tronquée . . . . .	77
2.20	$L(\nu, \delta)$ : rapport de sous-estimation de $\mu_{MT_{\delta}}$ par rapport à $\mu_{MN}$ . . . . .	78
2.21	Moyenne et écart type de l'estimateur $\hat{\mu}_{MT_{\delta}}$ en présence d'un parasite . . . . .	79
2.22	Densité de probabilité pour le calcul de la moyenne tronquée symétrique . . . . .	80
2.23	Moyenne tronquée symétrique : Calcul de $a_{\delta'}$ et $b_{\delta''}$ . . . . .	81
2.24	Estimateur de moyenne par moyenne tronquée symétrique . . . . .	82
2.25	Moyenne et écart type de l'estimateur $\hat{\mu}_{MTS_{\delta}}$ en présence d'un parasite . . . . .	84
2.26	Fonction correctrice $M(\nu, N)$ . . . . .	85
2.27	Estimateur de moyenne par moyenne du Petit Poucet . . . . .	86
2.28	Moyenne et écart type de l'estimateur $\hat{\mu}_{PP}$ en présence d'un parasite . . . . .	87
2.29	Calcul de $\mu_{02C}$ . . . . .	89
2.30	$C(\nu, N)$ : rapport pour calculer le seuil $S$ . . . . .	90
2.31	Estimateur de moyenne par moyenne de classe . . . . .	92
2.32	Moyenne et écart type de l'estimateur $\hat{\mu}_{2C}$ en présence d'un parasite . . . . .	93
2.33	Comparaison des estimateurs $\hat{\mu}_{2C}$ et $\hat{\mu}_{PP}$ face aux interférences . . . . .	93
2.34	$\hat{\mu}_{0LP}$ en présence d'interférences . . . . .	96
2.35	$\langle \Theta \rangle$ et $\eta$ en fonction de $\alpha$ et $\epsilon$ . . . . .	96
2.36	Moyenne et écart type de l'estimateur $\hat{\mu}_{LP}$ en présence d'un parasite . . . . .	97
2.37	Densité de probabilité $f_{[0]}(x)$ . . . . .	98
2.38	$\mu_{0lps}$ en fonction de $C$ . . . . .	99
2.39	Fonction correctrice $L(C)$ . . . . .	100
2.40	$\Omega(C)$ : variance normalisée de l'estimateur $\hat{\mu}_{0LPS}$ . . . . .	101
2.41	Moyenne et écart type de l'estimateur $\hat{\mu}_{LPS}$ en présence d'un parasite . . . . .	102
2.42	Choix d'un estimateur de moyenne . . . . .	105
3.1	Puissance instantanée d'une interférence radar . . . . .	108
3.2	Radars observés dans la bande 1.300 GHz-1.375 GHz . . . . .	109
3.3	Résidus radars après un ON-OFF . . . . .	110
3.4	Synoptique du "blankeur" radar . . . . .	111
3.5	Schéma du détecteur d'impulsions radar . . . . .	112
3.6	$P_{VP}$ en fonction de $INR_{dB}$ pour les détecteurs d'impulsions . . . . .	114
3.7	Courbes COR du détecteur d'impulsions faibles . . . . .	114
3.8	Simulation de "blanking" sur données synthétiques . . . . .	115
3.9	Simulation de "blanking" sur données réelles . . . . .	116
3.10	Observation de la raie HI de ESO 474-G26 . . . . .	118
3.11	Observation de la raie HI de Malin I . . . . .	119
3.12	Observation de la raie HI de PGC 12382 . . . . .	120
3.13	Observation de la raie HI de PGC 19534 . . . . .	121
3.14	Observation de la raie HI de PGC 51094 . . . . .	122
3.15	Émissions dans la bande de téléphonie mobile par satellite . . . . .	124
3.16	MégaMaser IIIZw35 avant la mise en service d'Iridium . . . . .	124
3.17	MégaMaser IIIZw35 depuis la mise en service d'Iridium . . . . .	125
3.18	Blancheur imparfaite de la bande d'analyse . . . . .	126

3.19	Réobservation de la raie à 1667 MHz en temps réel . . . . .	131
3.20	Réobservation des raies à 1667 MHz et 1667 MHz de III Zw 35 . . . . .	132
3.21	Plan d'occupation temps-fréquence de la routine classique . . . . .	133
3.22	Plan d'occupation temps-fréquence de la nouvelle routine RDH . . . . .	134
3.23	Effet de la résolution fréquentielle sur des parasites à bande étroite . . . . .	135
3.24	Fenêtre de Blackman-Harris . . . . .	137
3.25	Diverses estimations spectrales d'une raie . . . . .	137
3.26	Simulation du nombre de canaux occupés par une raie . . . . .	138
3.27	Comparaison des observations "routine" et "Rdh" . . . . .	142
3.28	Quicklook d'une émission Jovienne . . . . .	143
3.29	Données haute résolution d'une émission Jovienne . . . . .	144
3.30	Données "routine" d'une émission Jovienne . . . . .	145
3.31	Données "routine" estimées par médiane . . . . .	146
3.32	Données "routine" estimées par moyenne du Petit Poucet . . . . .	147
3.33	Données "routine" estimées par moyenne de classe . . . . .	148
4.1	Exemples de structures fines à enregistrer . . . . .	150
4.2	Principe du détecteur de structures fines . . . . .	152
4.3	Algorithme du détecteur de structures fines FISSA. . . . .	152
4.4	Réduction de l'espace de recherche dans le plan de Radon . . . . .	154
4.5	Indices de tableau de la transformée de Radon . . . . .	155
4.6	Accélération du calcul de la transformée de Radon . . . . .	156
4.7	Effet de $\Delta\theta$ sur l'estimation de $\mathcal{R}_{dis} \{I\}_{r=0}(\theta)$ . . . . .	156
4.8	Densité de probabilité de la variable aléatoire $SF_{Index}$ . . . . .	157
4.9	FISSA en présence et en absence d'interférences . . . . .	158
4.10	Ordonnancement des tâches et gestion mémoire de l'algorithme FISSA. . . . .	159
4.11	Exemple d'observations effectuées avec FISSA . . . . .	161
B.1	Approximation de $\log_2(x)$ . . . . .	171
B.2	Erreur quadratique moyenne du calcul de $\log_2$ . . . . .	174
B.3	Approximation de $\log_2(x)$ avec arrondis convergent . . . . .	176
B.4	Erreur quadratique moyenne du calcul de $\log_2$ avec arrondis convergent. . . . .	177
B.5	Hardware d'un calculateur de $\log_2$ . . . . .	178



---

## LISTE DES TABLEAUX

---

1.1	Bandes protégées pour la radioastronomie . . . . .	24
1.2	Résolutions d'analyse disponibles avec RDH. . . . .	44
2.1	Tabulation de $K(\nu)$ pour $\nu = 2^{1\cdots 10}$ . . . . .	68
2.2	Tabulation de $L(C)$ . . . . .	100
2.3	Comparatif des estimateurs de moyenne . . . . .	103
3.1	Comparaison des $P_{fa}$ des détecteurs d'impulsions . . . . .	115
3.2	Sources observées avec le "blanqueur" radar . . . . .	116
3.3	Comparaison des estimateurs face à un opérateur TDMA . . . . .	127
3.4	Comparaison de la complexité des algorithmes de classement . . . . .	127
3.5	Code C du classement par insertion directe . . . . .	128
3.6	Code C du classement progressif d'un N-échantillon . . . . .	129
3.7	Source observée avec le "blanqueur" Iridium . . . . .	130
3.8	Comparaison des SNRs Routine classique et Routine RDH . . . . .	134
3.9	Taux d'occupation spectrale des interférences entre 20 et 40 MHz . . . . .	140
4.1	Observation de structures fines : comparatif des volumes acquis . . . . .	159
B.1	Code C du calcul du $\log_2$ . . . . .	179



---

# GLOSSAIRE

---

**ADC** (Analog to Digital Converter) Voir CAN.

**CAN** (Convertisseur Analogique Numérique) Composant électronique qui échantillonne un signal analogique pour le coder sous un format numérique.

**CIADT** Comité interministériel pour l'aménagement et le développement du territoire

**COR** (Caractéristique Opérationnelle de Réception) Graphique représentant la probabilité de détection en fonction de la probabilité de fausse alarme pour différentes valeurs du seuil de discrimination.

**DDC** (Digital Down Conversion) Conversion en bande de base d'un signal sur porteuse effectuée en numérique.

**DMA** (Direct Memory Access) Procédé informatique où des données sont transférées d'un périphérique à l'autre *via* un contrôleur externe. Le processeur n'intervient plus que pour initier ou conclure le transfert et peut se consacrer à d'autres tâches.

**DSP** (Digital Signal Processor) Microprocesseur intégrant des fonctionnalités optimisées pour le traitement du signal.

**Effet Faraday** Interaction entre une onde électromagnétique et un champ magnétique dans un matériau : la polarisation de l'onde effectue une rotation proportionnelle à la composante du champ magnétique sur la direction de propagation de l'onde.

**FDMA** (Frequency Division Multiple Access) Multiplexage fréquentiel de signaux pour répartir une bande de fréquence entre plusieurs utilisateurs.

**FPGA** (Field Programmable Gate Array) Composant électronique reconfigurable au niveau porte logique permettant l'implantation d'architectures dédiées pouvant présenter un très haut niveau de parallélisme algorithmique.

**GLONASS** (GLObal NAVigation Satellite System) Système de positionnement russe basé sur de la triangulation par satellite pour assister au positionnement et à la navigation.

**GPS** (Global Positioning System) Système de positionnement américain basé sur de la triangulation par satellite pour assister au positionnement et à la navigation.

**IIR** (Infinite Impulse Response) Qualifie un filtre numérique à réponse impulsionnelle infinie. L'"infinité" de la réponse est obtenue par un algorithme calculatoire récursif.

**INR** (Interference to Noise Ratio) Rapport de la puissance de l'interférence sur celle du bruit, dB.

**ITU** (International Telecommunication Union) Organisation internationale appartenant aux Nations Unis constituée de représentants de gouvernements et de sociétés privées. Elle coordonne les réseaux de télécommunication et de services.

- Jitter** (Gigue d'horloge) Variation de phase des fronts d'une horloge. Génère une imprécision sur l'instant d'échantillonnage d'un signal et y introduit du bruit.
- LEO** (Low Earth Orbit) L'orbite LEO est une orbite circulaire, située entre 500 à 2000 km d'altitude. Cette proximité offre deux avantages : un temps de latence (temps que met à parcourir un signal) très court et une puissance réduite pour entrer en contact avec le satellite. [[www.futura-sciences.com](http://www.futura-sciences.com)]
- LLS( $k$ )** (Logical Left Shift) Décalage de  $k$  bits d'un mot binaire vers la gauche. Équivaut à la multiplication du mot en décimal par  $2^k$ .
- LOFAR** (LOW Frequency ARray) Interféromètre basse fréquence hollandais. (fonctionnement prévu fin 2008)
- LRS( $k$ )** (Logical Right Shift) Décalage de  $k$  bits d'un mot binaire vers la droite. Équivaut à la division du mot en décimal par  $2^k$  suivi de la troncature de la partie décimale.
- LUT** (Look Up Table) Table de correspondances numériques stockée dans une mémoire.
- LVDS** (Low Voltage Differential Signaling) Bus de transfert de données numériques utilisant des paires différentielles. Il permet l'utilisation de câble torsadé bon marché tout en assurant des débits élevés sur de longues distances tout en gardant une faible consommation.
- Mégamaser** Maser astrophysique extra-galactique. La luminosité de la cinquantaine de mégamasers connus est de 6 ordres au-dessus celle des masers galactiques, ce qui explique leur nom. Il s'agit d'un nuage de gaz pompé thermiquement qui amplifie par émission stimulée un rayonnement incident externe [51, §6]. L'observateur, situé dans l'alignement maser-source externe, perçoit un flux inhabituellement intense.
- NDA** (Nançay Decameter Array) Télescope opérant dans la bande décamétrique principalement pour l'étude du Soleil et de Jupiter.
- NRH** (Nançay Radio Heliograph) Interféromètre, imageur 2D, voué à l'observation du Soleil dans la bande 150-450 MHz.
- NRT** (Nançay Radio Telescope) Télescope opérant dans la bande décimétrique principalement pour l'étude des galaxies, de la cosmologie, des enveloppes d'étoiles, de l'effet Zeeman, des pulsars et des comètes.
- PCI** (Peripheral Component Interconnect) Norme de bus interne permettant de connecter des cartes filles à la carte mère d'un ordinateur. Le DMA est supporté par ce matériel.
- RDH** (Reconquête du Domaine Hertzien) Récepteur numérique à grande dynamique pour la radioastronomie.
- Red-Shift** (Décalage vers le rouge) Diminution, due à l'effet Doppler, de la fréquence d'observation d'une raie émise par un corps en mouvement.
- RFI** (Radio Frequency Interference) Interférence radioélectrique.
- SAO** (Spectromètre Acousto-Optique) Analyseur spectral basé sur une cellule de Bragg dont la réfraction variable est contrôlée par le signal à étudier. Le faisceau lumineux traversant la cellule est réfracté, la lecture de sa position permet de connaître le contenu spectral du signal.
- SNR** (Signal to Noise Ratio) Rapport de la puissance du signal sur celle du bruit, dB.

**TDMA** (Time Division Multiple Access) Multiplexage temporel de signaux pour répartir le temps d'utilisation d'un canal fréquentiel entre plusieurs utilisateurs.

**TFDMA** (Time-Frequency Division Multiple Access) Multiplexage temporel et fréquentiel de signaux pour répartir une bande de fréquence entre plusieurs utilisateurs.

**UMTS** (Universal Mobile Telecommunications System) Système de téléphonie mobile de troisième génération basée sur la technologie W-CDMA (Accès multiple par répartition en code en bande large).

**VLBI** (Very Long Base Interferometry) Interférométrie à très grande ligne de base, donc observation à fort pouvoir de résolution.





---

# INTRODUCTION

---

Le domaine radio des ondes électromagnétiques est une des ressources naturelles que l'Homme a vandalisé avant même de s'apercevoir qu'il pouvait l'étudier. Quand en 1928, les laboratoires Bell demandèrent à Karl Jansky d'étudier l'utilisation des ondes courtes pour la téléphonie transatlantique et qu'à cette occasion fut découvert Sagittarius A\*, la source radio la plus intense dans le ciel située au centre de notre Galaxie, Marconi avait déposé depuis 32 ans son premier brevet de radiotéléphonie et depuis 8 ans, Writtle (Angleterre) émettait les premiers programmes radiophoniques réguliers. Il fallut attendre 1937 pour voir la construction par Grote Reber sur ses propres finances [42] de la première parabole dédiée à l'observation des ondes radios d'origine naturelle. Il ne reçut guère de support pour ses recherches et c'est surtout après la seconde guerre mondiale et l'expérience acquise grâce aux radars que la radioastronomie se développe.

Depuis lors, les observations radioastronomiques ont toujours dû faire avec les interférences. Les astronomes disposent évidemment de bandes protégées qui leur sont exclusivement réservées mais la proximité de bandes allouées aux télécommunications rend difficile certaines observations. De plus, certains phénomènes physiques ne se réduisent pas à docilement émettre dans les bandes protégées. Il faut donc scruter tout ou partie de la bande radio au risque de faire face à de fortes perturbations.

L'utilisation de filtres analogiques permet, certes, de limiter au mieux la bande d'analyse et de s'affranchir ainsi d'un certain nombre d'interférences mais les observations ont été tout de même effectuées de préférence dans des lieux reculés, loin des activités humaines. Malheureusement, si l'avènement du satellite a permis à l'Humanité de se sentir chez Elle partout sur Terre, il a aussi placé une myriade d'émetteurs intenses devant le Zodiaque des Anciens.

L'essor actuel des télécommunications obscurcit encore cet horizon peu réjouissant. Néanmoins, il offre aussi de nouveaux outils qui peuvent aider la radioastronomie à continuer d'être. Les nouvelles technologies numériques développées pour le grand public peuvent être réutilisées en instrumentation pour restaurer quelques pans du spectre radio.

Nous verrons dans une première partie pourquoi l'étude des astres *via* la radioastronomie est si importante, comment elle est actuellement effectuée et quelles sont les caractéristiques des interférences qui nous gênent tant.

Dans une deuxième partie, nous exposerons les bases du blanking et le problème de classification généré. La statistique de détection choisie, la puissance, sera justifiée et nous montrerons que le problème se résume à une estimation robuste de moyenne. Divers estimateurs de moyenne, classiques ou de mon cru, seront comparés en terme d'efficacité, de robustesse aux interférences et de coût d'implantation matérielle.

La troisième partie présentera des applications opérationnelles de traitements d'interférences basées sur les travaux précédemment décrits.

Enfin, dans une dernière partie, nous exposerons une méthode originale de détection et d'enregistrement de signaux naturels brefs en présence d'interférences intenses. L'enregistrement de ces types de signaux serait extrêmement trivial –un simple détecteur de puissance– si la bande étudiée n'était pas utilisée par les transmissions par ondes courtes. Leur présence masque de plusieurs dizaines de décibels les signaux naturels et implique un critère de détection plus complexe.

# CHAPITRE 1

---

## INTRODUCTION À LA RADIOASTRONOMIE

---

Afin de fournir des méthodes appropriées à la restauration de signaux de sources radio-astronomiques altérés par des interférences, un minimum de connaissance est nécessaire de façon à introduire les bonnes connaissances *a priori* qui permettront de spécialiser au mieux les implantations d’algorithmes.

Nous présenterons très brièvement quels signaux sont accessibles à la radioastronomie et quelles en sont les origines. Nous exposerons les types d’interférences auxquels se confrontent les observations. Nous décrirons ensuite les méthodes, les instruments et les unités communément utilisés dans ce domaine. Étant donné que le travail a porté sur les radiotélescopes dit “antenne unique”, nous développerons plus spécifiquement les méthodes liées à ce type d’instrument.

### 1.1 Sources étudiées

La quasi-totalité des connaissances accumulées par l’humanité sur l’Univers – les galaxies, les étoiles et leur planètes, les comètes, l’espace qui les sépare – a été obtenue par l’intermédiaire de l’observation des ondes électromagnétiques qui émanent de ces corps. Le peu de matière sur laquelle nous avons un peu étendu notre savoir provient des météorites qui ont pénétré l’atmosphère terrestre, des particules cosmiques et des échantillons collectés par des sondes automatisées ou des missions habitées.

Le spectre électromagnétique couvre le domaine des ondes radios, des micro-ondes, de l’infrarouge, du visible, de l’ultraviolet, des rayons X et des rayons  $\gamma$ . Les divers phénomènes physiques qui se produisent dans l’Univers émettent une contrepartie électromagnétique dans diverses bandes de fréquence (partie basse de la figure 1.1). L’étude d’un corps nécessite la collecte d’observations dans le plus grand nombre de domaines possible afin d’obtenir le panorama le plus complet des phénomènes mis en jeu.

Malheureusement, cette collecte n’est pas forcément évidente, car la totalité du spectre électromagnétique ne parvient pas intacte à notre portée.

#### 1.1.1 Fenêtres de transmission

L’étude de l’Univers a été entamée par les Anciens lors d’observations oculaires directes du Soleil, des étoiles et des planètes. Sans le savoir, Ils utilisaient l’une des deux fenêtres de transmission laissée à la Terre par son atmosphère.

En effet, les composés de l'atmosphère terrestre (vapeur d'eau, dioxygène, ozone, dioxyde de carbone, azote) interagissent avec les ondes électromagnétiques et les atténuent ou les réfléchissent dans certaines bandes. Les deux seules zones laissées presque intactes sont celle de la fenêtre du visible et celle de la fenêtre radio (figure 1.1). L'observation des ondes électromagnétiques situées en dehors de ces fenêtres n'est possible qu'à haute altitude.

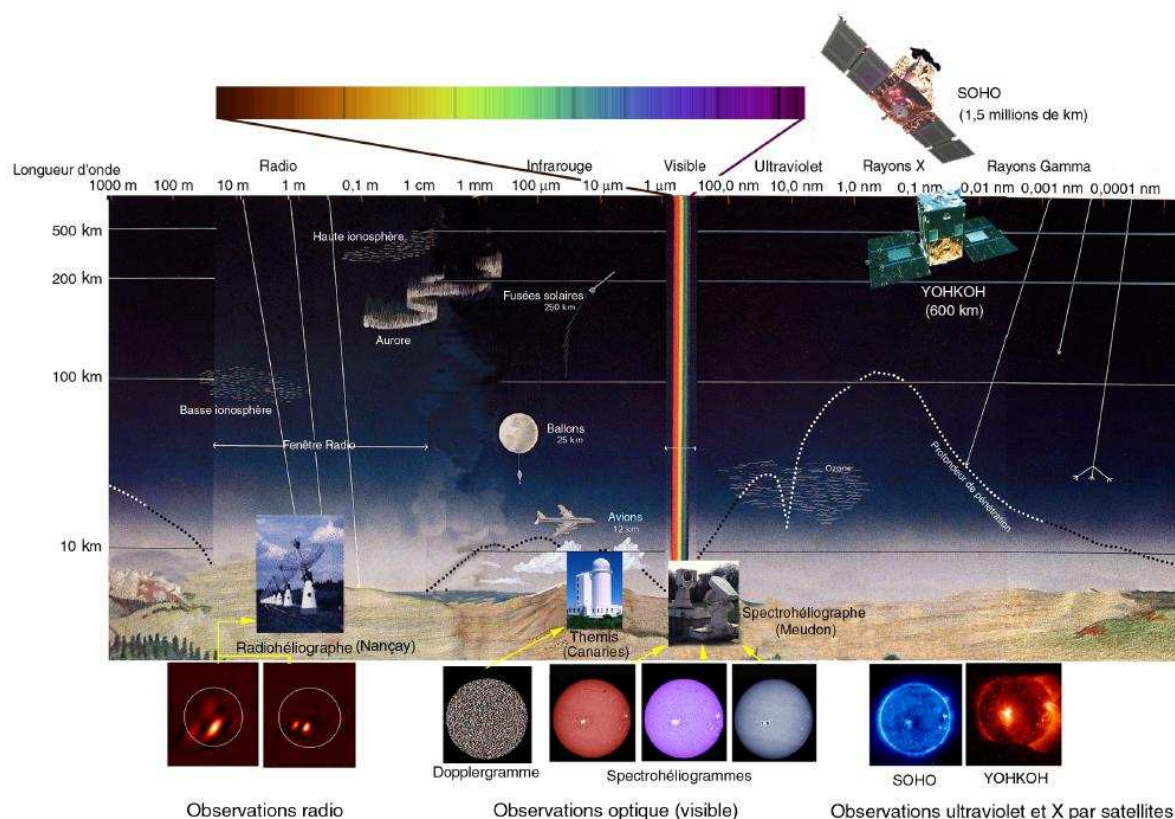


FIG. 1.1 – Atténuation des ondes électromagnétiques par l'atmosphère  
[\[www.lesia.obspm.fr\]](http://www.lesia.obspm.fr)

La fenêtre radio s'étend de quelques mégahertz jusqu'à 300 GHz. La limite haute est fixée par l'absorption des ondes par la vapeur d'eau et le dioxygène. La limite basse est fonction de la concentration électronique  $N_e$  du plasma ionosphérique qui réfléchit les ondes basse fréquence. La fréquence de coupure est définie par  $F_p = 9\sqrt{N_e}$ .  $N_e$  varie en fonction de l'interaction Soleil-ionosphère. Les conditions à la verticale du lieu d'observation varient donc avec le jour et la nuit et le niveau d'activité du Soleil. La nuit, la fréquence de coupure basse est typiquement de 4.5 MHz et atteint 11 MHz de jour.

### 1.1.2 Procédés d'émission

Nous avons vu que les corps célestes rayonnaient dans un très grand domaine de fréquences mais que seules certaines bandes étaient accessibles au sol. La fenêtre radio donne accès à divers corps célestes dont voici une énumération non-exhaustive :

- Étoiles (dont le Soleil)

- Planètes
- Comètes
- Galaxies
- Mégamasers
- Quasars
- Pulsars

Les phénomènes physiques internes à ces corps produisent des émissions radio-électriques classiquement réparties de la façon suivante :

- émission thermique
- émission non-thermique
  - raies spectrales
  - rayonnement synchrotron
  - rayonnement “Bremßstrahlung”

Ces procédés d’émission peuvent présenter un degré de polarisation linéaire et/ou circulaire.

Nous réorganisons ces phénomènes selon qu’ils émettent à bande large ou à bande étroite. Bien que plus abrupte, cette classification permet de mieux faire le parallèle avec les classes d’interférences qui seront décrites plus loin dans ce chapitre.

### 1.1.2.1 ÉMISSION CONTINUUM

Les multiples interactions ions-électrons et électrons-champs magnétiques produisent des impulsions radioélectriques d’amplitude et largeur diverses. La superposition d’un grand nombre de tels événements dans un corps produit un bruit large bande appelé émission continuum.

Le continuum radio est produit par 3 mécanismes principaux : l’**émission thermique** dont l’intensité est proportionnelle à la température du corps observé (Loi de Rayleigh-Jeans valable pour  $h\nu \ll kT$ , en particulier en radioastronomie), l’**émission non-thermique** qui résulte de l’interaction de particules très énergétiques avec des champs magnétiques. Dans le cas d’électrons quasi-relativistes, on parle d’émission synchrotron. Le dernier type d’émission est le **rayonnement continu de freinage** (Bremßstrahlung), créé lors de l’accélération de particules chargées dans un champ électrique. L’étude de ces émissions radio permet d’estimer les densités électroniques et les champs magnétiques en présence.

La figure 1.2 est le résultat de la constitution d’une carte de l’émission continuum par les observations à 408 MHz des radiotélescopes Jodrell Bank, Effelsberg et Parkes [37].

Les émissions à bande large ont été détectées et étudiées les premières. En 1944, Hendrick van de Hulst suggéra la présence d’une raie hydrogène observable à 1420 MHz. Il propose que sa recherche en soit effectuée.

### 1.1.2.2 ÉMISSION DE RAIES SPECTRALES

L’espace entre les étoiles n’est pas vide mais parsemé de gaz et de poussières. Ces nuages peuvent être denses, froids et le siège de création d’étoiles et de planètes. Ailleurs, des étoiles chauffent et nourrissent d’énergie ces nuages. Dictés par la physique quantique, les changements de niveaux d’énergie au niveau des atomes du nuage provoquent donc l’émission d’ondes radioélectriques à des fréquences discrètes [62, §10] [45, §8.24-8.25].

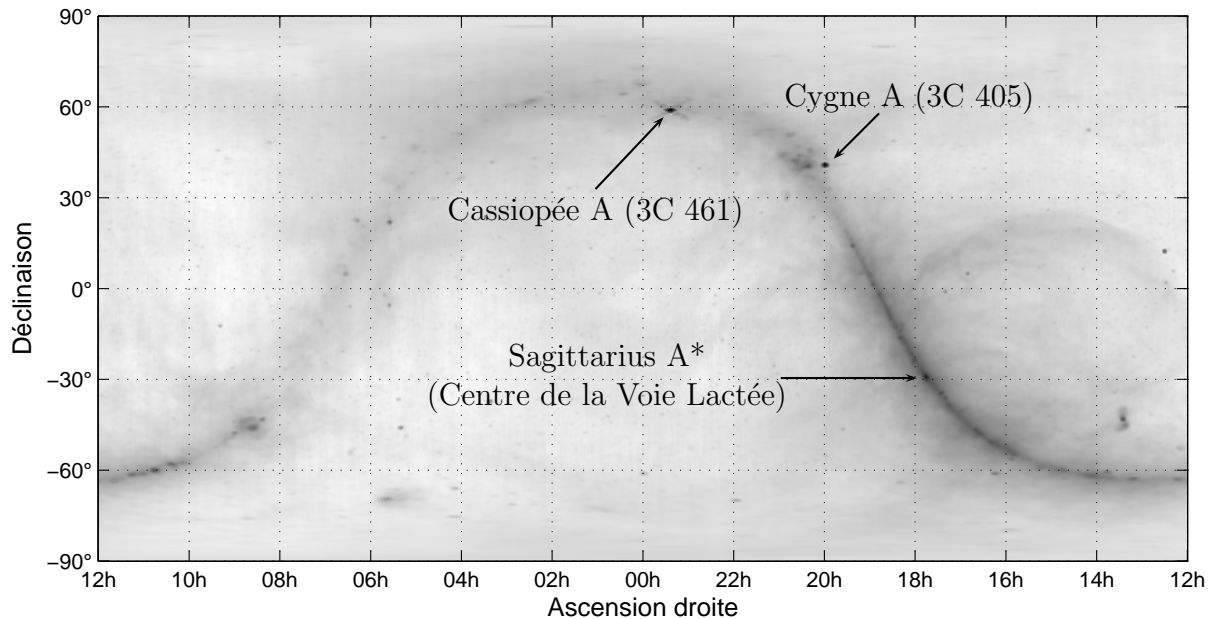


FIG. 1.2 – Carte de l'émission continuum mesurée à 408 MHz. (coordonnées équatoriales, équinoxe de référence : J2000) Trois sources particulièrement intenses se dégagent de la carte. Cassiopée A, la plus forte source radio du ciel après le Soleil, est le rémanent d'une supernova qui a explosé dans la Voie Lactée dans la seconde moitié du XVIIe siècle. Cygne A est une radio galaxie subissant une explosion massive ou une collision avec une autre galaxie. Sagittarius A\* occupe le centre de notre galaxie (coordonnées J2000 : ascension droite 17h 45m 40,045s, déclinaison  $-29,00775^\circ$ ). Sa masse estimée à  $3,7 \pm 1,5$  million de masses solaires, confinée dans un rayon de moins de 120 unités astronomiques en fait le candidat le plus probable au titre de trou noir super massif. <http://skyview.gsfc.nasa.gov>

Si le nuage est en mouvement, l'ensemble est soumis à un effet Doppler qui décale les fréquences observées. Il est donc possible de remonter à la vitesse relative du nuage ainsi qu'à sa dynamique interne grâce à une étude spectroscopique des émissions radios.

La raie à 1420.4057 MHz de l'hydrogène neutre (HI) au repos est la signature de la matière la plus présente dans l'Univers (75% en masse et 90% en nombre d'atomes). C'est donc aussi la plus observée en radioastronomie. Elle fut détectée le 25 Mars 1951 par Ewen et Purcell (Université d'Harvard). Depuis, plusieurs milliers d'autres raies atomiques et moléculaires ont été détectées dans le milieu interstellaire et sont utilisées, par exemple, en exobiologie.

Si les atomes peuvent émettre des photons aux niveaux d'énergie quantifiés, ils peuvent aussi en absorber en suivant les mêmes règles. On peut donc observer des raies d'absorption.

### 1.1.2.3 ABSORPTION DE RAIES SPECTRALES

Ce phénomène n'est pas, à proprement parler, le siège d'une émission radio mais il trahit la présence et la composition chimique d'un nuage de gaz froid. En effet, si un nuage de température  $T_c$  est présent sur la ligne de visée d'une source continuum, *i.e.*

large bande, de température  $T_{sc}$  telle que  $T_{sc} > T_c$ , les atomes et molécules présents vont absorber une partie du rayonnement à certaines fréquences discrètes et laisser passer le reste du rayonnement. L'estimation spectrale du rayonnement va donc faire apparaître dans la bande d'analyse une série de "creux" correspondant à la signature des éléments chimiques observés [45, §8.26].

Afin d'observer ces signaux en toute sérénité, la communauté radioastronomique a demandé l'établissement de protections juridiques sur des zones du spectre radioélectrique particulièrement intéressantes.

### 1.1.3 Bandes protégées

Après la détection en 1932 par Karl G. Jansky de l'émission du centre de notre galaxie, après la cartographie du ciel à 160 MHz en 1939 par Grote Reber, après la découverte du rayonnement électromagnétique de Jupiter en 1955 puis de la détection de la raie de l'hydrogène neutre en 1951, la communauté scientifique réclama un statut officiel auprès de l'ITU afin que la radioastronomie soit considérée comme un "service" à part entière de la bande radio-fréquence. Les scientifiques demandèrent la protection d'une série de bandes espacées afin d'observer les émissions continuum et de bandes spécifiquement centrées autour des raies d'émission de l'hydrogène neutre (1420.406 MHz), du radical OH (1612.231 MHz, 1665.402 MHz, 1667.359 MHz, 1720.530 MHz) et de la raie deutérium (327.384 MHz, non observée à l'époque, seulement prédite).

Les recherches se poursuivant, de nouvelles bandes furent négociées pour arriver au panorama présenté dans la table 1.1.

Mais la radioastronomie n'est pas toujours utilisateur primaire dans ces bandes et certaines bandes autrefois considérées comme peu intéressantes sont maintenant très utiles. Les observations doivent parfois se faire en présence d'émetteurs radio artificiels.

De plus, les observations de sources extra-galactiques sont des observations de sources de vitesse apparente non négligeable, due majoritairement à l'expansion de l'Univers. Ces sources sont fortement affectées par l'effet Doppler, c'est à dire que la fréquence d'observation des fréquences d'émission des raies fondamentales se décale fortement vers les basses fréquences (red-shift) et sort des bandes protégées. La Science n'a que faire des réglementations humaines et les réponses à certains problèmes astrophysiques ne peuvent être recherchées que dans des bandes occupées : il faut donc observer en présence d'émetteurs artificiels dont une classification est donnée à la partie suivante.



TAB. 1.1 – Bandes dédiées à la radioastronomie (extraits des allocations ITU).

Allocation ITU RR	Types d'observation
13.36 - 13.41 MHz 25.55 - 25.67 MHz	Soleil, Jupiter
37.5 - 38.25 MHz	Jupiter
73.0 - 74.6 MHz	Soleil
150.05 - 153.0 MHz	Continuum, pulsar, Soleil
322 - 328.6 MHz	Continuum, deutérium
406.1 - 410 MHz	Continuum
608 - 614 MHz	VLBI
1330 - 1400 MHz	Raie HI red-shiftée
1400 - 1427 MHz	Raie HI
1610.6 - 1613.8 MHz 1660 - 1670 MHz 1718.8 - 1722.2 MHz	Raies OH
2655.0 - 2700.0 MHz	Continuum, HII
3100.0 - 3400.0 MHz	Raies CH
4800.0 - 5000.0 MHz	VLBI, HII, raies H <sub>2</sub> CO et HCOH
6650.0 - 6675.2 MHz	CH <sub>3</sub> OH, VLBI
10.60 - 10.70 GHz 14.47 - 14.50 GHz 15.35 - 15.40 GHz	Quasar, raies H <sub>2</sub> CO, Continuum
22.01 - 22.21 GHz	Raie H <sub>2</sub> O red-shiftée
22.21 - 22.5 GHz	Raies H <sub>2</sub> O
22.81 - 22.86 GHz	Raies NH <sub>3</sub> , HCOOCH <sub>3</sub>
23.07 - 23.12 GHz	Raies NH <sub>3</sub>
23.6 - 24.0 GHz	Raie NH <sub>3</sub> , Continuum
31.3 - 31.8 GHz	Continuum
36.43 - 36.5 GHz	Raies HC <sub>3</sub> N, OH
42.5 - 43.5 GHz	Raie SiO
47.2 - 50.2 GHz	Raies CS, H <sub>2</sub> CO, CH <sub>3</sub> OH, OCS
51.4 - 59.0 GHz	
76.0 - 116.0 GHz	Continuum, raies moléculaires
123 - 158.5 GHz	Raies H <sub>2</sub> CO, DCN, H <sub>2</sub> CO, CS
164.0 - 167.0 GHz	Continuum
168.0 - 185.0 GHz	H <sub>2</sub> O, O <sub>3</sub> , multiples raies
191.8 - 231.5 GHz	Raie CO à 230.5 GHz
241 - 275 GHz	Raies C <sub>2</sub> H, HCN, HCO <sup>+</sup>
275 - 1000 GHz	Continuum, Raies moléculaires

## 1.2 Zoologie des interférences radioélectriques

La zoologie des interférences radioélectriques a été détaillée dans [30]. Nous ne retiendrons que deux critères généraux pour ces interférences. Premièrement, l’occupation spectrale que l’on peut qualifier soit à bande étroite ou à bande large (figure 1.6 et 1.7). Deuxièmement, l’occupation temporelle que l’on peut qualifier de continue ou d’intermittente.

Ces deux paires de critères binaires génèrent quatre classes d’interférences :

- Interférences intermittentes à bande étroite.
- Interférences intermittentes à bande large.
- Interférences continues à bande étroite.
- Interférences continues à bande large.

Les deux premières classes d’interférences (figures 1.3, 1.4 et 1.5) permettent de continuer l’observation de sources astronomiques dans toute la bande à condition de disposer d’un système permettant de détecter et rejeter les tranches temporelles corrompues. Cette procédure simple est appelée “blanking”. La largeur des canaux du système d’analyse doit être adaptée aux interférences pour retirer du plan temps-fréquence le minimum de données intègres contigues aux données corrompues. Le point crucial de cette procédure est la classification appropriée intègre/corrompue des zones du plan temps-fréquence. Des méthodes développées pour ces deux classes d’interférences sont présentées en sections 3.1 et 3.2.

Les deux dernières classes d’interférences ne permettent pas l’utilisation du “blanking” puisqu’à chaque instant le parasite est présent dans le canal analysé : la source est donc constamment masquée et ne peut être restaurée.

Pour des émetteurs à large bande de type GPS, GLONASS, UMTS, ... (figure 1.7) qui utilisent plusieurs mégahertz de bande chacun, il faut donc utiliser des méthodes plus élaborées comme l’annulation d’interférences, l’estimation d’interférence, l’annulation spatiale (en interférométrie), ... pour restaurer les signaux masqués [3, 7, 52, 57, 56, 8, 6].

Pour des interférences à bande étroite  $\delta f$  parasitant des sources naturelles à bande large  $\Delta f$  (figure 1.6), il est possible d’effectuer une analyse fréquentielle de la bande à une résolution très inférieure à  $\Delta f$  afin de résoudre au mieux les interférences. On gagne ainsi un maximum de canaux libres qui permettent de reconstruire une pseudo-bande intègre à la résolution nécessaire à l’étude des signaux naturels. Cette classe d’interférences est à rapprocher des classes à critère d’occupation temporelle intermittente dans la mesure où l’on considère que le processus est ergodique, ce qui est le cas dans l’hypothèse d’un bruit blanc gaussien. On se permet alors, dans un bloc de données, de substituer une série fréquentielle à une série temporelle. Une implémentation de ce type de traitement est étudiée en section 3.3.

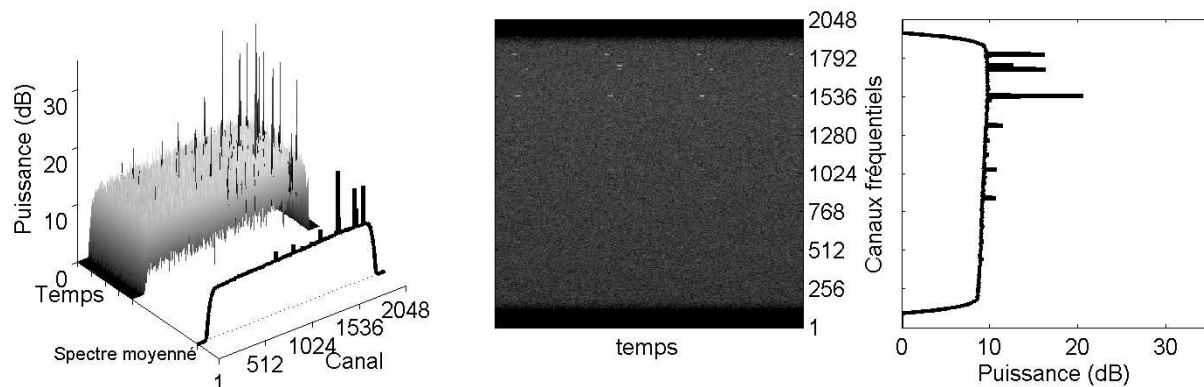


FIG. 1.3 – Interférence générée par une modulation TFDMA sur une observation au NRT

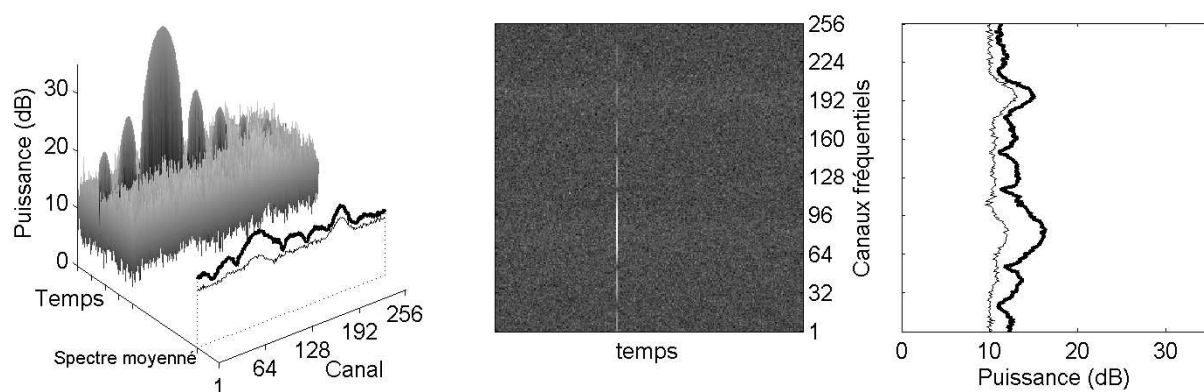


FIG. 1.4 – Interférence à large bande. Une impulsion courte occupe une large bande spectrale. Les sources (trait fin) sont fortement distordues (trait fort). L'intégration est décalée d'1 dB par rapport à la référence pour affichage.

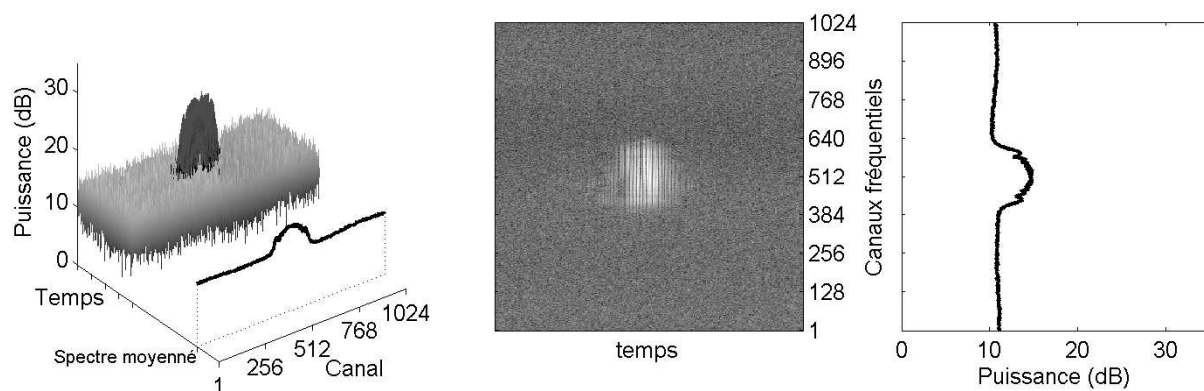


FIG. 1.5 – Interférence générée par un radar sur une observation au NRT

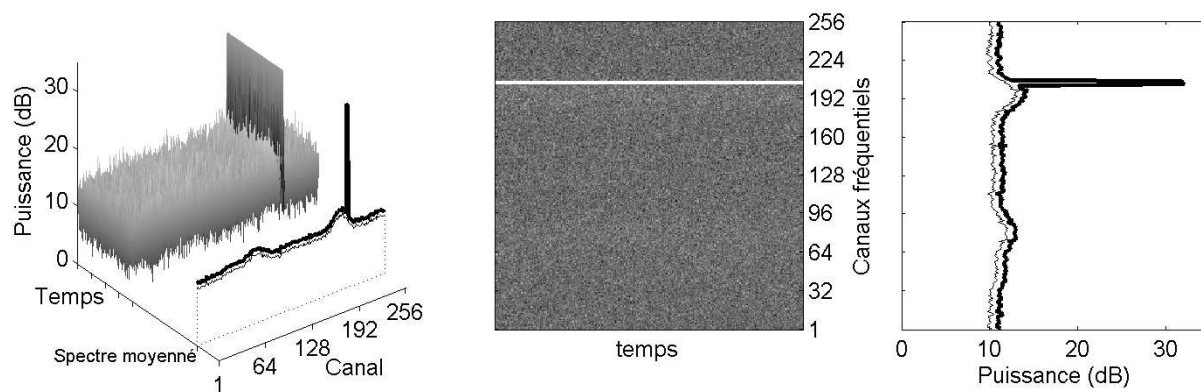


FIG. 1.6 – Interférence continue à bande étroite. Les sources sont majoritairement préservées mais les canaux 199 à 201 sont définitivement perdus. L'intégration est décalée d'1 dB par rapport à la référence pour affichage.

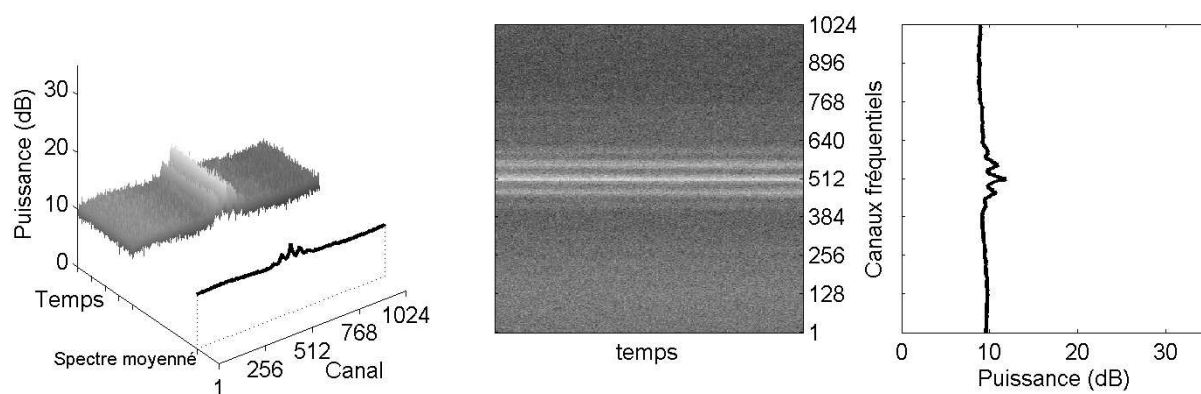


FIG. 1.7 – Interférence continue à bande large : porteuse L1 du GPS (1575.42 MHz). Une large bande fréquentielle n'est pas disponible pour les observations.

## 1.3 Antennes et radiotélescopes

Selon Grote Reber [59], “*la combinaison antenne-récepteur se comporte comme un bolomètre, un instrument de mesure de chaleur, dans laquelle la résistance de rayonnement de l’antenne mesure la température équivalente d’objets distants projetés par la réponse du diagramme d’antenne*”.

La nature des objets observés implique divers types d’antennes associées à divers types de récepteurs pour former des radiotélescopes aux diverses utilisations [5].

### 1.3.1 Types d’antennes

Une antenne est la transition entre une zone d’ondes libres et un guide d’ondes. Son rôle en radioastronomie est de concentrer les ondes électromagnétiques générées par des radio-sources et de les transformer en signal électrique. Ses principales caractéristiques sont sa bande de fréquence d’utilisation, son gain, son diagramme de rayonnement, sa surface collectrice, sa taille et son type de réalisation mécanique.

Il faut noter que surface collectrice et taille d’antenne ne vont pas forcément de paire. Une antenne dite “multiple” est composée de plusieurs antennes éloignées ce qui confère à l’ensemble une grande taille et donc un pouvoir de résolution important, mais la somme  $A_e$  des surfaces collectrices de ces antennes peut rester faible face à l’encombrement total de l’instrument et rendre inaccessible l’observation de sources peu intenses (voir section 1.4.1.4).

Les antennes de radiotélescopes se répartissent en trois classes principales :

- Les antennes fixes qui ne comportent aucune pièce mobile. Ce sont des réseaux dont le lobe principal du diagramme d’antenne est synthétisé par la mise en phase des signaux d’antennes.
- Les antennes partiellement manoeuvrables, souvent pour des raisons mécaniques.
- Les antennes totalement manoeuvrables.

Les télescopes se fondent sur les “jeux” d’antennes décrites précédemment et se divisent eux-mêmes en deux classes:

- Les télescopes à antennes uniques.
- Les interféromètres.

Le site de Nançay (figure 1.8) regroupe un large panel de télescopes différents en cours d’exploitation, détaillés ci-après.

### 1.3.2 Télescopes à antenne unique

Les télescopes dits “antenne unique” ne sont pas forcément construits sur une unique antenne physique mais les données qu’ils fournissent sont celles qui seraient issues d’une telle antenne virtuelle. L’antenne physique peut donc être un unique réflecteur plein comme c’est le cas pour le NRT (France), Arecibo (Porto Rico), Effelsberg (Allemagne), Lovell (Angleterre) ou Green Bank (USA), ou bien un ensemble d’antennes fixes sommées en phase (nommé réseau phasé, voir figure 1.9) comme pour le NDA ou les stations de base de l’interféromètre LOFAR.

Ce type de télescope permet d’observer un point particulier du ciel. Le lobe principal de l’antenne est pointé sur une zone de la source à étudier. La chaîne de réception (figure 1.10) filtre le signal pour ne conserver que la bande voulue qui est amplifiée puis

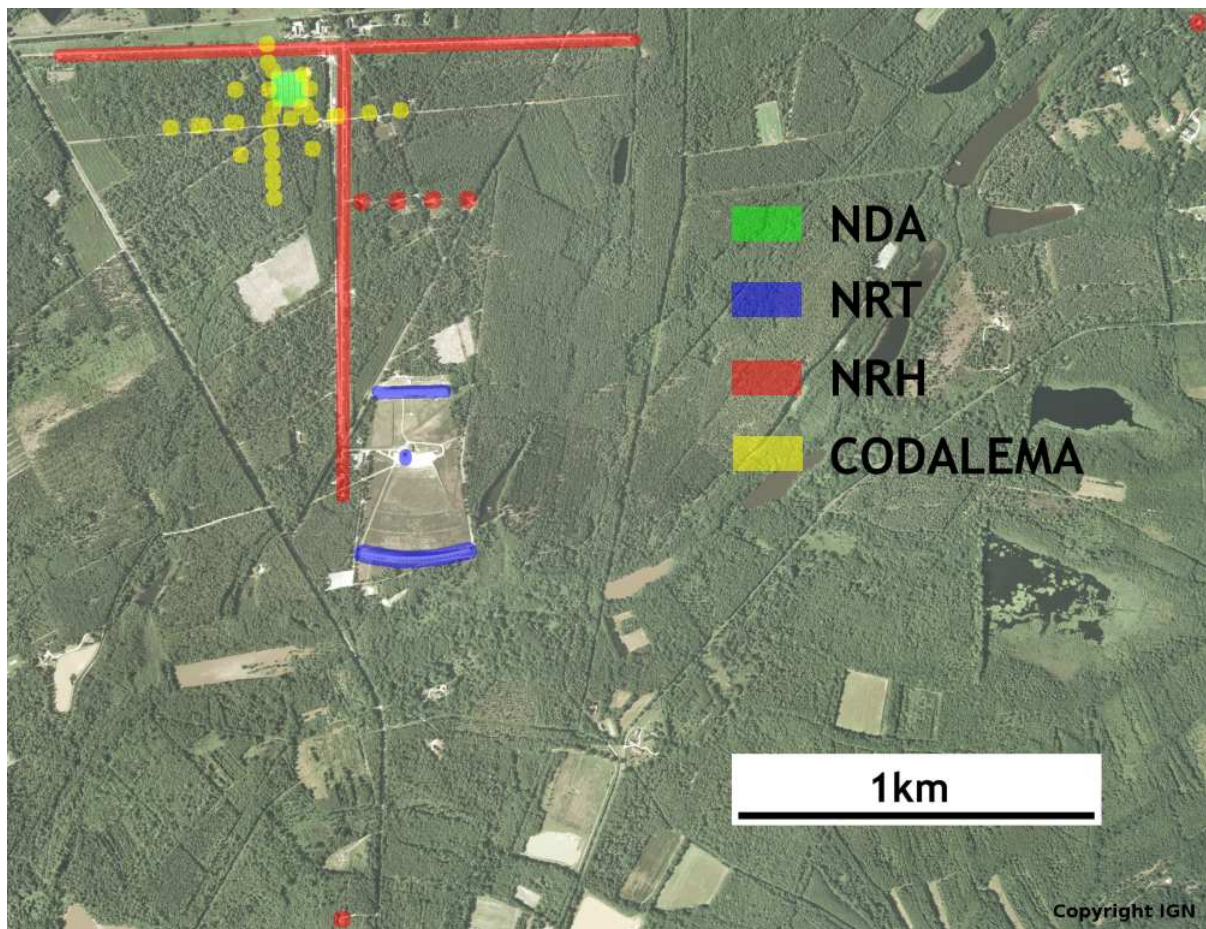


FIG. 1.8 – Grands instruments de Nançay. NDA : Réseau Décimétrique, NRT : Radiotélescope Décimétrique, NRH : Radio-Héliographe, CODALEMA : astronomie impulsionnelle

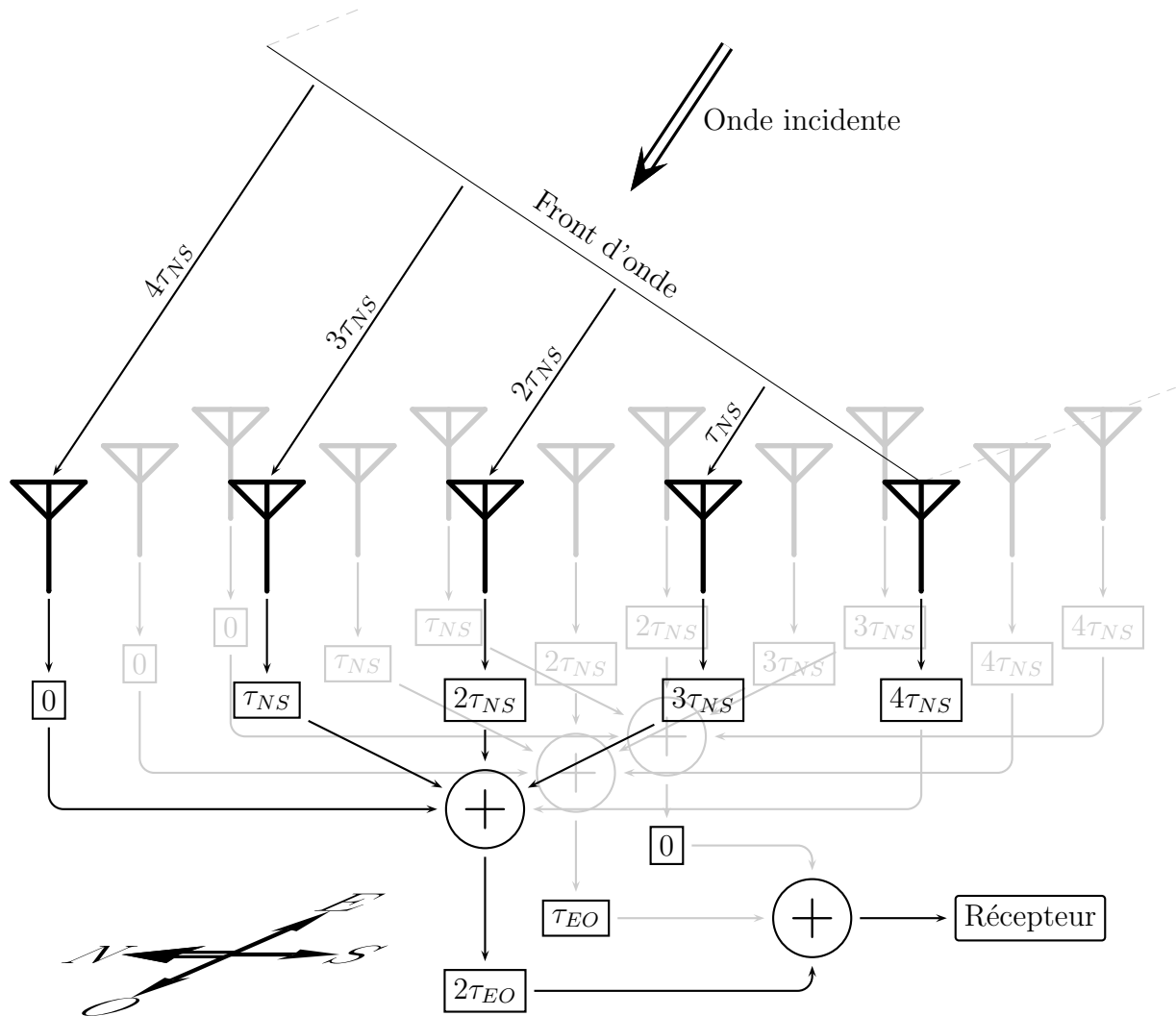


FIG. 1.9 – Fonctionnement d'un réseau phasé : L'onde incidente balaye successivement les antennes d'une ligne du réseau (en noir). Les retards d'arrivée sur la ligne Nord-Sud sont compensés par les déphasages  $\tau_{NS}$  avant d'effectuer une première sommation des signaux. Le choix des retards permet de pointer le télescope dans la direction voulue située dans le plan d'une ligne antennes. L'ensemble du ciel est accessible via l'addition phasée ( $\tau_{EO}$ ) de ces lignes d'antennes (en noir et gris). Le signal résultant est envoyé vers le récepteur. Le réseau multi-antenne est donc vu par le récepteur comme une unique antenne.

transposée en bande de base. Le signal peut maintenant subir une analyse spectrale (SAO, autocorrélateur, FFTeur)<sup>1</sup>. La chaîne de réception peut être dédoublée pour mutualiser le réflecteur. Il est facile de mesurer le champ électrique sur les deux vecteurs d'une base orthogonale. Dans le cas d'une paire d'antennes à polarisation linéaire, on mesure la composante électrique horizontale et la composante électrique verticale (cas du NRT). Pour des antennes de polarisation circulaire, on mesure la composante électrique circulaire droite et la composante électrique circulaire gauche (cas du NDA). Dans les deux cas, la

<sup>1</sup>D'autres types de traitements sont envisageables : dédispersion pour la recherche de radar, détection d'événements brefs, transit de sources, ...

paire de mesure permet de remonter aux paramètres de polarisation de l'onde électrique incidente. Une simple matrice de passage permet de passer d'une représentation à l'autre.

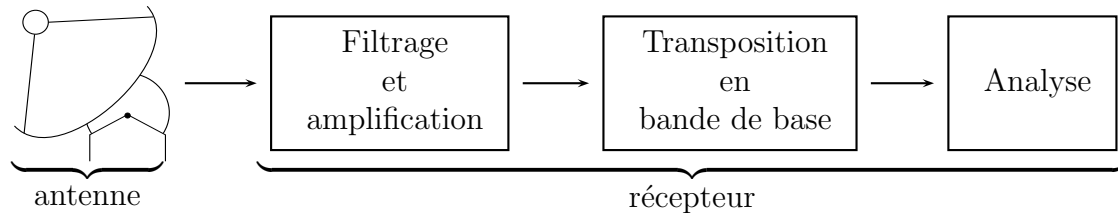


FIG. 1.10 – Chaîne de réception simplifiée d'un télescope à antenne unique

Le système doit être calibré pour assurer la blancheur de la chaîne de réception et pour connaître le niveau de puissance absolu mesuré par l'instrument afin d'assurer une exploitation cohérente des données avec d'autres instruments (§1.4.1, p.36).

Le traitement du signal sur ce type d'antenne requière peu de duplication de matériel car toute l'information est contenue sur une, voire 4 voies (cas d'un signal polarisé). C'est sur le cas des antennes uniques qu'ont été développés et implantés les algorithmes d'atténuation d'interférence de cette thèse.

Le site de Nançay présente deux exemples de ce type de télescope : le NRT, un télescope à antenne unique fondée sur des réflecteurs pleins et le NDA, un réseau phasé.

### 1.3.2.1 PRÉSENTATION DU NRT

Il est utilisé pour l'étude des Galaxies et de la cosmologie, les enveloppes d'étoiles, les objets proto-stellaires, les Pulsars et les Comètes dans la bande décimétrique.

Il faut pour observer ces ondes radioélectriques souvent très faibles une grande surface collectrice. C'est ainsi que fut décidée la construction d'un grand radiotélescope décimétrique. Une première tranche représentant le cinquième de l'instrument définitif a d'abord été construite en 1961, par la Compagnie Française d'Entreprises. Les essais ayant montré que son fonctionnement était correct, la construction était achevée en 1964.

Il est symétrique par rapport à un plan Nord-Sud. On ne pourra observer un astre que lorsqu'il passe au voisinage de ce plan, appelé plan méridien (c'est cette restriction qui a permis de construire une très grande surface pour un prix modéré (15 M€ équivalent en 2005)). Les ondes reçues tombent sur un premier miroir (figure 1.11) qui est un plan de 200 m x 40 m, et dont on peut régler l'inclinaison : on choisit ainsi le point du ciel à étudier. Ce réflecteur, comme le suivant, est recouvert d'un grillage<sup>2</sup> dont la maille est de 12.5 mm, et sa forme est précise à mieux que 5 mm près. Les ondes arrivent ensuite sur un miroir fixe, situé à 460 m du premier, et mesurant 300 m x 35 m. Celui-ci est sphérique, de 560 m de rayon. Le Radiotélescope Décimétrique (ou NRT) est l'instrument le plus imposant de l'Observatoire de Nançay. Avec sa surface collectrice effective de 7000 m<sup>2</sup>, il est l'un des plus grands du monde. Il concentre les ondes en un point, son foyer, situé entre les deux miroirs. On les y recueille dans des cornets. À cause de la rotation de la Terre, la zone de ciel visée se déplace lentement, d'Est en Ouest. Le foyer où se concentrent les ondes se déplace donc sur une "surface focale", et les cornets doivent suivre ce mouvement.

<sup>2</sup>vu comme un réflecteur "plein" pour les ondes décimétriques



Pour cela ils sont portés par un ascenseur, lui même situé sur un chariot guidé par des rails.

Le récepteur est conçu pour couvrir de façon continue la bande de fréquences de 1,060 GHz à 3,500 GHz, permettant l'étude d'émissions continuum mais aussi des raies spectrales HI (hydrogène neutre), OH (radical hydroxyle) et du radical CH subissant un fort décalage Doppler. La température système (voir 1.4.1.2) vaut environ 35 K dans la partie basse de cette bande et monte vers le haut de cette bande, en partie à cause de la "transparence" des miroirs qui augmente lorsque la longueur d'onde diminue.

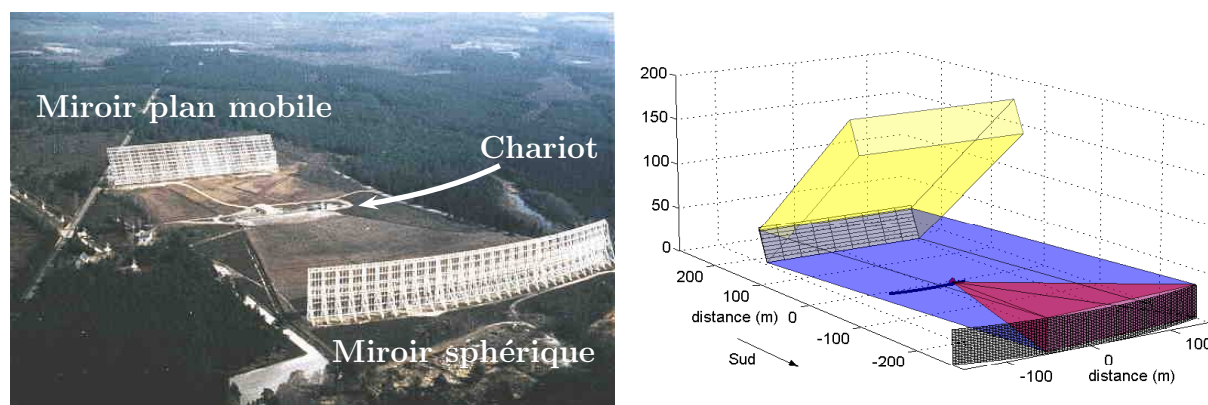


FIG. 1.11 – Radiotélescope Décimétrique de Nançay (NRT) : Le signal radioastronomique est réfléchi par le miroir plan vers le miroir sphérique qui concentre le faisceau sur la voie focale où se trouve le chariot qui porte l'étage de réception. Le positionnement du miroir plan et du chariot permet de décrire le ciel.

### 1.3.2.2 PRÉSENTATION DU NDA

Le Réseau Décimétrique de Nançay (NDA) est dédié à l'étude des environnements magnétisés et ionisés des planètes (principalement Jupiter) et du Soleil.

Cet instrument, conçu pendant les années 1975-78 sous la direction d'André BOISCHOT, fonctionne entre 3 et 30 mètres de longueur d'onde, comporte 144 antennes log-spirales de 9 mètres de hauteur et de 5 mètres de diamètre à la base, réparties sur une surface physique de 10.000 m<sup>2</sup> (figure 1.12). Il s'agit d'un réseau phasé qui observe en polarisation circulaire droite et gauche. La surface physique de l'antenne est répartie en deux surfaces effectives de 4000 m<sup>2</sup> chacune.

Les observations radio de ces objets révèlent des phénomènes variables (jusqu'à des échelles de temps très inférieures à la seconde : sursauts "millisecondes" de Jupiter, éruptions solaires et stellaires) et permettent donc d'étudier la dynamique de ces environnements.

Pour les longueurs d'ondes plus grandes, inobservables depuis le sol, les observations doivent être réalisées à l'aide des sondes spatiales telles que Voyager, Ulysses, Galileo ou STEREO auxquelles l'instrument de Nançay est scientifiquement associé.



FIG. 1.12 – Les antennes log-spirales fixes du réseau décamétrique de Nançay. La direction de visée est fixée par le déphasage des signaux provenant des antennes. Les pylônes sont pointés vers l'élévation moyenne de passage des sources étudiées (Soleil, Jupiter).

### 1.3.3 Interféromètre

Les interféromètres sont des imageurs radioélectriques 2D (voire 3D si on inclut le temps) construits autour d'un ensemble d'antennes corrélées entre elles. En une observation, ils fournissent des cartes d'intensité d'émissions pour une ou plusieurs bandes de fréquence. Le lobe des antennes est pointé vers les sources à étudier. Les signaux reçus sont pré-traités, intercorrélés (figure 1.13) puis intégrés pour participer à l'échantillonnage de la fonction complexe de visibilité dans le plan  $uv$  (figure 1.15). En effet, à chaque paire d'antennes est associé un point dans le plan  $uv$  et ce point se voit attribuer le résultat complexe de l'intercorrélés  $C_{ij}$  des deux antennes. La transformée de Fourier 2D inverse de la fonction de visibilité permet de restituer l'image des sources observées [68].

Le traitement des interférences nécessite l'analyse simultanée des signaux issus de toutes les antennes. Il s'agit du traitement d'un signal à  $N$  dimensions <sup>3</sup> traité en particulier par [6].

#### Présentation du NRH

Le Radiohéliographe de Nançay (figures 1.14 et 1.15) observe le Soleil de façon quotidienne et simultanée en 5 longueurs d'onde (0.7 m à 2 m, soit 450 à 150 MHz). Ses images sondent la couronne solaire à différentes altitudes entre 50 000 km et 500 000 km au dessus de la photosphère car le rayonnement provient de couches de la couronne de plus en plus extérieures à mesure que la fréquence d'observation décroît.

Le récepteur à corrélation calcule 2344 intercorrélations et permet une imagerie à deux dimensions. Il peut acquérir 200 images par seconde à plusieurs fréquences simultanément. De ce fait, il est particulièrement adapté à l'observation des émissions solaires les plus rapides.

En terme d'interférence, Cet instrument est perturbé par la télévision analogique terrestre, en particulier par la première chaîne cryptée française. Une partie des canaux d'observation est corrompue par des émetteurs à bande étroite utilisés pour les

<sup>3</sup> $N$  désigne le nombre d'intercorrélations possible.

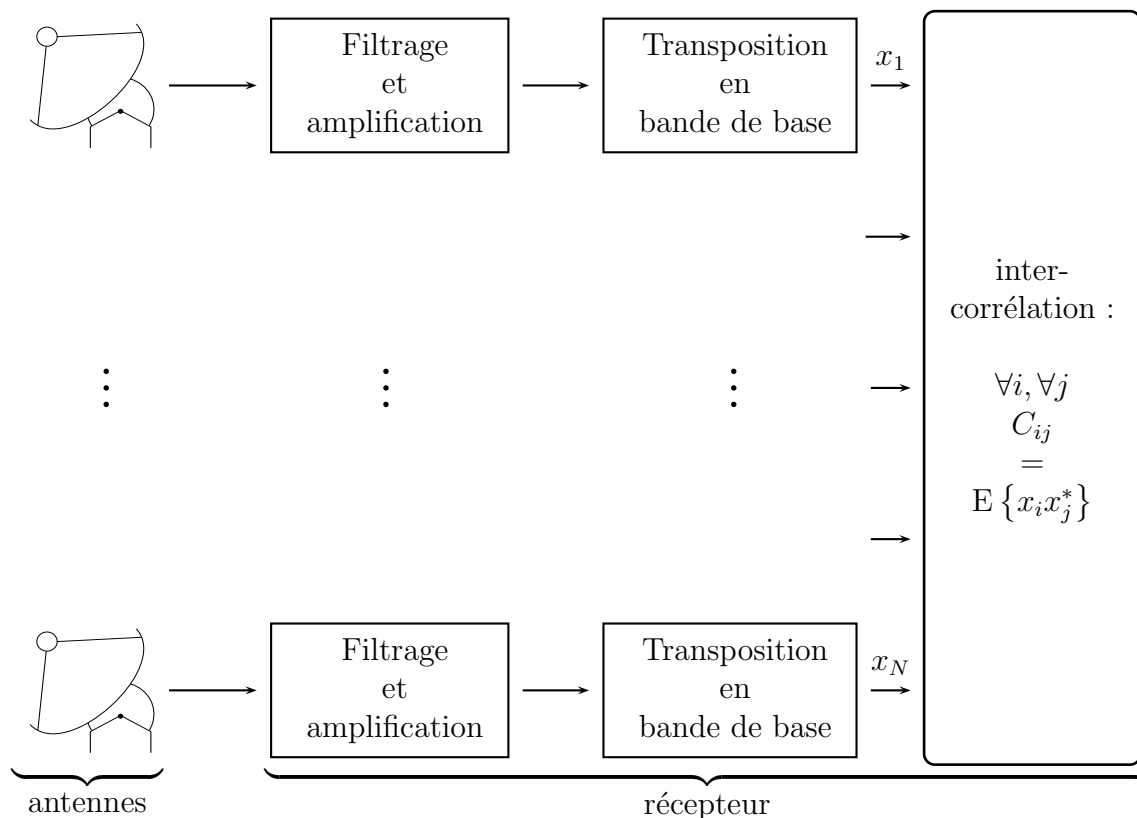


FIG. 1.13 – Chaîne de réception simplifiée d'un télescope interféromètre



FIG. 1.14 – Antennes du bras Nord-Sud du NRH.

télécommunications militaires et maritimes, pour la radiolocalisation et par les radio-amateurs.

Des filtres analogiques sont insérés sur les voies d'antennes pour limiter les phénomènes d'intermodulation et seules des bandes étroites choisies dans des canaux propres sont analysées. Cette lutte analogique contre les interférences qui opère en suivant une tactique d'évitement permet d'assurer la plupart des observations.

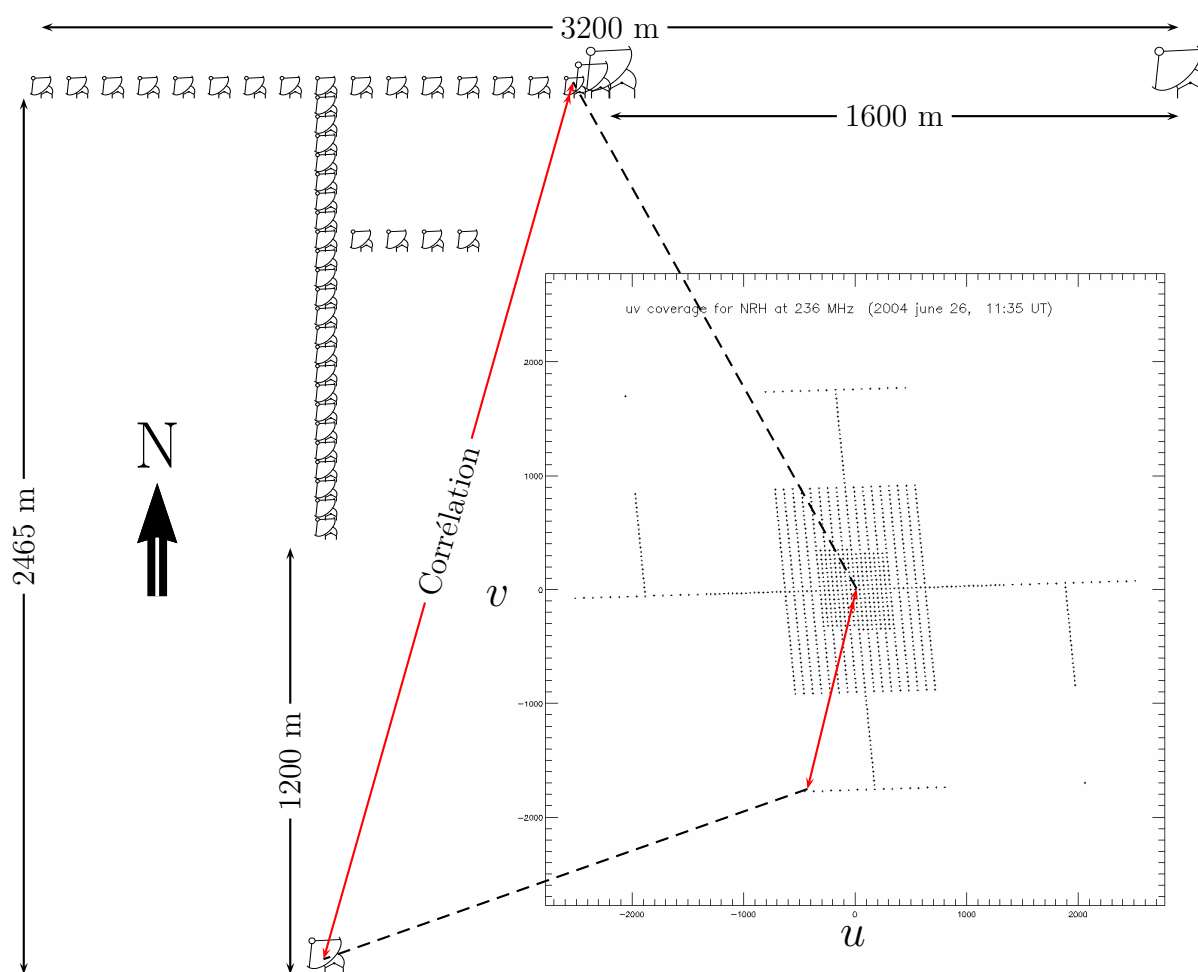


FIG. 1.15 – Configuration physique du Radiohéliographe de Nançay (NRH) et couverture instantanée du plan  $uv$  associé.

## 1.4 Instrumentation spécifique au radiotélescope à antenne unique

Les études présentées dans ce document sont appliquées à des observations effectuées *via* un télescope à antenne unique. Nous présentons ici les concepts et les unités couramment utilisés dans ce domaine. Nous exposerons aussi les caractéristiques du récepteur numérique qui a servi de banc d'essai aux traitements décrits dans ce mémoire.

### 1.4.1 Mesures et calibrations

Afin de mieux comprendre les mesures effectuées par les astronomes, il est nécessaire de se familiariser avec les unités et procédures utilisées dans la spectrométrie de signaux radioastronomiques.

#### 1.4.1.1 FRÉQUENCE ET VITESSE DE SOURCE

L'expansion de l'univers se manifeste par l'observation d'une vitesse de récession des objets astrophysiques lointains par rapport à la Galaxie. Si l'on n'observe pas directement un déplacement de ceux-ci, on observe un décalage de leurs raies d'émission et de leurs raies d'absorption que l'on peut interpréter en terme d'effet Doppler. Ce décalage étant presque systématiquement vers le rouge, l'on en conclut que les objets s'éloignent de nous. De plus, ce mouvement d'éloignement relatif est homogène et isotrope dans l'Univers : une galaxie située à une distance donnée s'éloigne de la nôtre à la même vitesse, quelle que soit la direction où elle se trouve, et il en est de même pour tout observateur qui serait situé dans une autre galaxie. Il existe donc une relation entre la vitesse de récession des galaxies et la distance qui nous sépare d'elles, connue sous le nom de Loi de Hubble :

$$v = H_0 d \quad (1.1)$$

avec  $v$  : Vitesse de la source ( $\text{km.s}^{-1}$ )

$H_0$  : Constante de Hubble ( $= 72 \pm 10\% \text{ km.s}^{-1}.\text{Mpc}^{-1}$ )

$d$  : Distance de la source (Mpc)

La relation liant vitesse et décalage spectral dans le cadre de la relativité restreinte est donnée par [40, Chap. 11] :

$$f = f_0 \frac{\sqrt{1 - \left(\frac{V}{c}\right)^2}}{1 + \frac{V}{c} \cos(\alpha)} \quad (1.2)$$

avec  $f$  : Fréquence d'observation de la raie (Hz)

$f_0$  : Fréquence au repos de la raie (Hz)

$V$  : Vitesse absolue de la source ( $\text{m.s}^{-1}$ )

$c$  : Vitesse de la lumière dans le vide ( $= 299792458\text{m.s}^{-1}$ )

$\alpha$  : Angle entre le vecteur de vitesse absolue et la ligne de visée (radian)

Néanmoins, il est rarement possible de connaître  $\alpha$ . La vitesse absolue  $V$  n'est donc pas accessible. L'une des 4 conventions majeures de vitesse est donc utilisée [22] : la vitesse relativiste ( $V_{rel}$ ), la vitesse radio ( $V_{rad}$ ), la vitesse optique ( $V_{opt}$ ) et le décalage spectral vers le rouge ( $z$ ).

$$V_{rad} = \frac{f_0 - f}{f_0}c = \frac{\lambda - \lambda_0}{\lambda}c \quad (1.3)$$

$$V_{opt} = \frac{\lambda - \lambda_0}{\lambda_0}c = \frac{f_0 - f}{f}c \quad (1.4)$$

$$V_{rel} = \frac{f_0^2 - f^2}{f_0^2 + f^2}c \quad (1.5)$$

$$z = \frac{f_0 - f}{f} \quad (1.6)$$

La vitesse relativiste, rarement utilisée, est un cas particulier de 1.2 pour  $\alpha = 0$ . Pour éviter la confusion entre la vitesse optique issue de l'astronomie "optique" où une onde est caractérisée par sa longueur d'onde  $\lambda$  et la vitesse radio issue de l'astronomie où l'onde est cette fois caractérisée par sa fréquence  $f$ , on préférera utiliser le décalage spectral  $z$  qui est communément utilisé dans les deux communautés.

#### 1.4.1.2 TEMPÉRATURE DE SOURCE

Un radiotélescope est un instrument qui détermine, à distance, la température équivalente d'un objet. La puissance rayonnée par la zone pointée, projetée par l'intermédiaire du lobe d'antenne et mesurée aux bornes de la résistance de rayonnement de l'antenne peut s'écrire comme suit [45, 62] :

$$P_{sys} = kT_{sys}\Delta\nu \quad (1.7)$$

avec  $P_{sys}$  : Puissance mesurée (W)

$k$  : Constante de Boltzmann ( $= 1.38 \times 10^{-23} \text{J.K}^{-1}$ )

$T_{sys}$  : Température système totale (K)

$\Delta\nu$  : Bande passante (Hz)

La puissance des bruits externes à la source peut se modéliser comme le bruit thermique équivalent présent dans une résistance parasite prise en série avec la résistance de rayonnement de l'antenne. La température système  $T_{sys}$  est le résultat de la somme de deux sources de bruit  $T_a$  et  $T_{rec}$ , avec  $T_a$  la température équivalente de ce qui est collecté par l'antenne et  $T_{rec}$  la température équivalant au bruit généré par l'électronique du récepteur.  $T_a$  peut se décomposer comme suit :

$$T_a = T_{CMB} + T_{GBN} + T_{atm} + T_G + \Delta T + T_{RFI} \quad (1.8)$$

avec  $T_a$  : Température d'antenne

$T_{CMB}$  : Bruit de fond cosmique (Cosmic Microwave background)

$T_{GBN}$  : Bruit de fond galactique (Galactic Background Noise) (La Voie Lactée)

$T_{atm}$  : Température de l'atmosphère à la fréquence observée

$T_G$  : Température de sol (Ground)

$\Delta T$  : Température de l'éventuelle source étudiée

$T_{RFI}$  : Température équivalente des interférences artificielles

Mis à part la température  $\Delta T$  qui est l'essence même de la radioastronomie, les autres températures ne sont que des sources de bruit indésirables.

La  $T_{rec}$  correspond au bruit apporté à la mesure par les composants électroniques de la chaîne de réception. On veille évidemment à soigner la qualité des amplificateurs situés en tête de chaîne de réception. Ce matériel peut nécessiter une cryogénéisation pour atteindre les performances requises. Au NRT,  $T_{rec}$  vaut 15K.

La  $T_G$  provient de la projection sur l'antenne, via les chemins secondaires, d'un angle solide plus ou moins important de sol à  $\sim 300K$  (température ambiante). On peut aussi y inclure la température de transparence des réflecteurs qui participent à l'augmentation du bruit système par leur opacité incomplète. Un flux non-négligeable provenant de la matière située au-delà du miroir pénètre ainsi dans le système.

Les trois dernières nuisances correspondent à tout ce qui est présent sur la ligne de visée mis à part la source elle-même. Du plus éloigné au plus proche, on trouve (figure 1.16) :

$T_{CMB}$ : c'est la température du reliquat de l'époque chaude et dense qu'a connue l'Univers il y a 13 milliards d'années. C'est le moment où l'Univers a suffisamment refroidi pour que les photons n'interagissent plus totalement avec la matière et commencent à pouvoir circuler dans la matière et donc à être visible. Ce bruit de fond vaut 2.7K avec un maximum d'émission à 160 GHz (domaine microonde).

$T_{GBN}$ : Le radiotélescope est localisé sur Terre ou à grande proximité. *A fortiori*, il est donc aussi situé dans "notre" galaxie, à savoir, la Voie Lactée. Quelque soit la direction visée, il y aura donc de la matière appartenant à la Voie Lactée sur l'alignement observateur-source. Le cas le plus favorable est celui où le moins de matière est visée, c'est à dire vers les pôles galactiques. Le cas le plus défavorable est celui d'un pointage vers le centre de la galaxie. Cette température augmente en basse fréquence.

$T_{atm}$ : c'est le bruit thermique de notre atmosphère. Plus la couche traversée est fine, moins le niveau de bruit est élevé. Le minimum est donc obtenu lors d'une observation au zénith et le maximum lors de l'observation d'une source à l'horizon.

Lorsque l'antenne pointe sur le fond du ciel, hors source, la température  $T_{mes1} = T_{rec} + T_{CMB} + T_{GBN} + T_{atm} + T_G$  est mesurée par le récepteur. Lorsque l'antenne pointe sur une source, la température mesurée est maintenant  $T_{mes2} = T_{mes1} + \Delta T$ .

Quand  $\Delta T \gg T_{mes1}$  (émissions non-thermiques par exemple), les températures parasites peuvent être ignorées et, une fois calibré sur une source connue, le récepteur mesure directement la température  $\Delta T$  de la zone pointée.

Quand  $\Delta T \ll T_{mes1}$ , c'est à dire que le signal utile est noyé dans le bruit (figure 1.17), il faut procéder à une observation de type *ON-OFF*. Une première observation est menée sur la source (*ON*-source) puis une seconde observation est immédiatement effectuée en dehors de la source (*OFF*-source) pour assurer des conditions d'observation quasiment

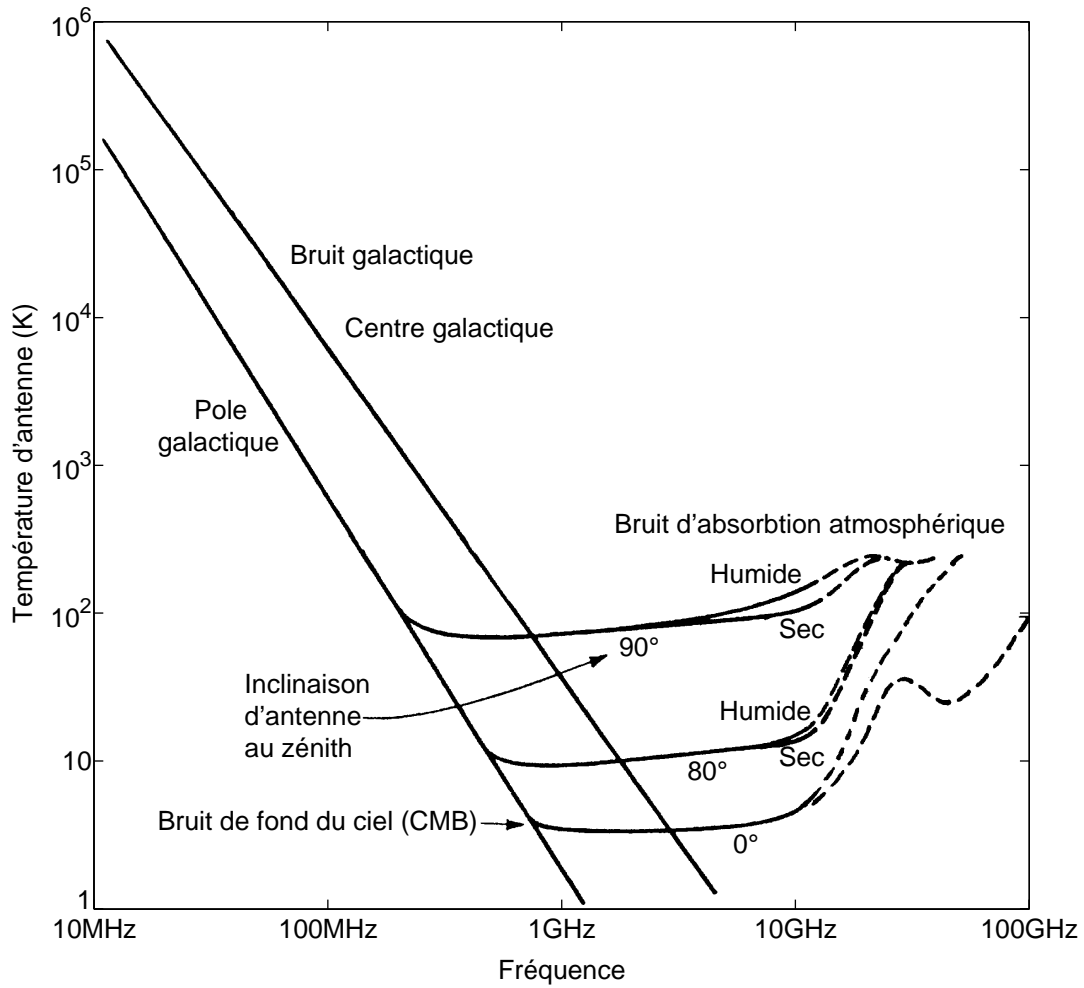


FIG. 1.16 – Température d'antenne en fonction de la fréquence [45, Fig. 7-1]

identiques. On calcule ensuite le rapport  $(ON - OFF)/OFF$  pour obtenir :

$$\frac{ON - OFF}{OFF} = \frac{T_{mes2} - T_{mes1}}{T_{mes1}} = \frac{\Delta T}{T_{sys}} \quad (1.9)$$

Dans ce cas, la détermination de  $\Delta T$  passe par la bonne connaissance de la température système  $T_{sys}$ . La calibration de l'instrument doit être régulièrement effectuée et les variations de gain doivent être prises en compte. On peut notamment souligner les variations diurnes/nocturnes de  $T_{atm}$ .

### 1.4.1.3 DENSITÉ DE FLUX D'UNE SOURCE

La contribution de la source à la densité de flux en provenance de la zone pointée est notée  $S$  telle que :

$$S = \frac{2k\Delta T}{A_e} \quad (1.10)$$



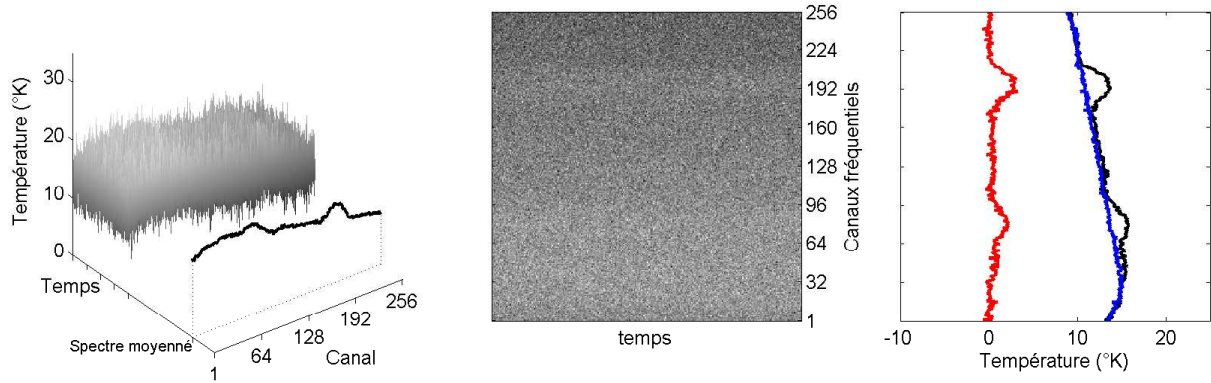


FIG. 1.17 – Sources noyées dans du bruit. Après intégration des spectres de l’observation *ON*-source (figure centrale) au cours du temps, il est parfois possible de discerner les sources (sur le spectre noir, centré sur les canaux 80 et 192). L’intégration des spectres de l’observation *OFF*-source permet d’obtenir la contribution de la température des sources de bruits parasites (spectre bleu). La différence permet la restitution de  $\Delta T$ , la température de la source observée (spectre rouge).

avec  $S$  : Densité de flux en provenance de la zone observée ( $\text{W}\cdot\text{m}^{-2}\cdot\text{Hz}^{-1}$ )

$k$  : Constante de Boltzmann ( $= 1.38 \times 10^{-23} \text{J}\cdot\text{K}^{-1}$ )

$\Delta T$  : Température de source (K)

$A_e$  : Aire effective de l’antenne ( $\text{m}^2$ )

Puisque les flux mesurés en radioastronomie sont souvent très faibles, une unité spécifique au domaine a été introduite sous le nom de Jansky (abrégée Jy), telle que  $1 \text{ Jy} = 10^{-26} \text{W}\cdot\text{m}^{-2}\cdot\text{Hz}^{-1}$ .

#### 1.4.1.4 LIMITE DE DÉTECTION D’UNE SOURCE

On considère qu’une source est convenablement détectée quand elle induit, sur la ligne de base de l’estimation spectrale  $\Delta T$  de la source, une variation  $\Delta T_{min}$  de l’ordre de l’amplitude rms  $\Delta T_{rms}$  du bruit du spectre  $\Delta T[k]$  [45] :

$$\Delta T_{min} = \Delta T_{rms} = \frac{K_s T_{sys}}{\sqrt{\Delta\nu N \tau}} \quad (1.11)$$

avec  $K_s$  : Coefficient de sensibilité de l’instrument (proche de 1, sans unité)

$T_{sys}$  : Température système (K)

$\Delta\nu$  : Bande de fréquence d’analyse (Hz)

$N$  : Nombre de séquences d’intégration

$\tau$  : Durée d’une intégration (s)

En substituant 1.11 dans 1.10, la densité de flux minimum détectable devient :

$$\Delta S_{min} = \frac{2k}{A_e} \frac{K_s T_{sys}}{\sqrt{\Delta\nu N \tau}} \quad (1.12)$$

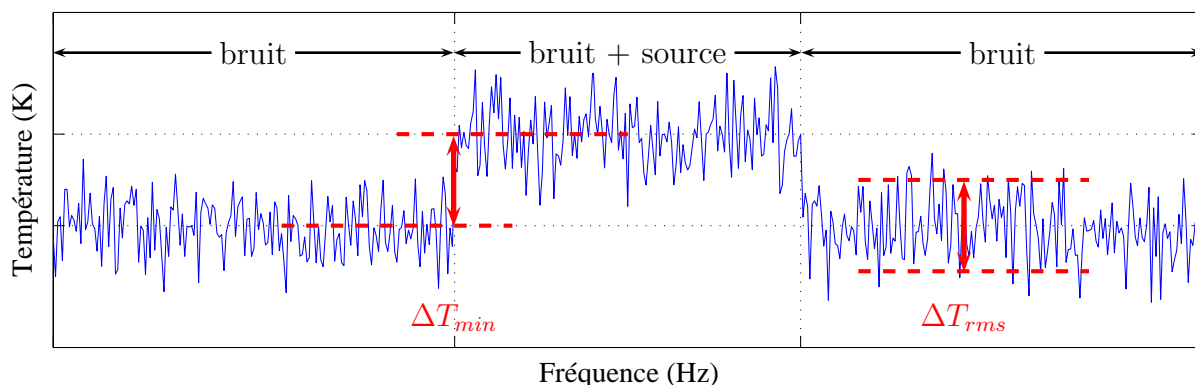


FIG. 1.18 – Limite de détection d'une source

Pour détecter des sources de très faible densité de flux, on peut donc diminuer  $T_{sys}$  en réduisant le bruit du récepteur par exemple, ou augmenter  $A_e$ ,  $\Delta\nu$ ,  $N$  ou  $\tau$ .

Néanmoins, la durée d'intégration  $\tau$  d'une séquence ne peut pas être allongée à volonté. Elle est techniquement limitée par la stabilité des horloges et mixeurs du récepteur à un temps d'intégration optimal  $T_A$  connu sous le nom de temps de stabilité d'Allan ([44] et [2], p.60). Cela se manifeste par une décroissance normale du bruit rms en rapport proportionnel avec la racine carré du temps d'intégration, puis lorsque  $\tau$  dépasse le temps  $T_A$ , la tendance s'inverse et le niveau de bruit rms remonte. Il est donc judicieux de travailler avec  $\tau \approx T_A$  et d'augmenter  $N$  si besoin, c'est à dire alterner les séquences *ON*-source et *OFF*-source.

Le même type de limitation s'applique à  $\Delta\nu$ . Une grande bande passante est intéressante du point de vue sensibilité pour l'observation d'émissions continuum mais elle est difficile à manipuler lors des changements de fréquences, filtrages et numérisations requis pour son analyse. De même, une large bande peut augmenter la présence d'interférences qui vont rendre les observations difficiles.

### 1.4.2 Récepteur numérique de test

Le récepteur RDH (Reconquête du Domaine Hertzien) a été conçu grâce au financement des mesures CIADT de 1998. C'est un récepteur numérique grande dynamique multi-instruments, multi-utilisateurs. Il permet l'observation simultanée dans 8 bandes de 14 MHz que l'on peut juxtaposer pour augmenter la bande totale d'analyse. Il est utilisable pour trois instruments d'observation de la Station de Nançay, le Radiotélescope Décimétrique, le Réseau Décamétrique et l'antenne de surveillance parasite. Le récepteur numérique est construit autour de cartes numériques d'usage général. Le système consiste en une collection de cartes filles portées par une carte principale de type PCI qui gère les transferts de données et la communication avec l'ordinateur hôte. Les cartes filles peuvent être choisies parmi une large gamme de fonctionnalités ce qui permet au client de construire la solution qui lui convient le mieux. Nançay a construit RDH, un système basé sur 8 voies comportant chacune un module de conversion analogique, des modules FPGA et des modules DSP (figure 1.19).

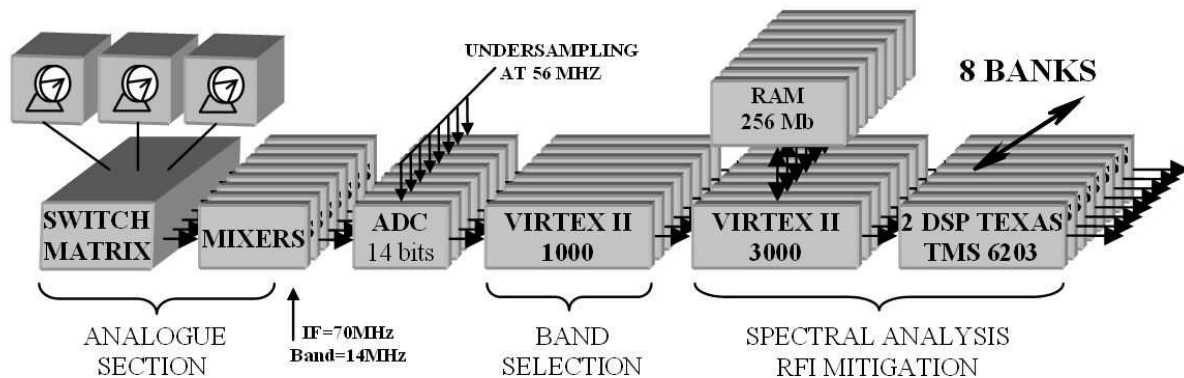


FIG. 1.19 – Structure du récepteur RDH

Les signaux en provenance des instruments sont numérisés par des convertisseurs analogique-numérique 14 bits. Chaque voie passe par un filtre anti-repliement passe-bas de 14 MHz de bande passante centrée sur 70 MHz puis sous-échantillonnée à 56 MHz. Cet étage doit être soigneusement réalisé afin de ne pas ajouter aux interférences observées par les antennes, les interférences produites par les PCs hôtes (figure 1.20). De plus, le sous-échantillonnage nécessite une horloge à faible gigue (jitter) [12, 66, 67] et il est indispensable de disposer d'une alimentation filtrée et correctement distribuée.

La bande numérique obtenue est convertie en représentation complexe IQ [49] puis traitée par un DDC (Digital Down Converter) afin de réduire le nombre d'échantillons par seconde. Nous disposons maintenant d'une forme d'onde complexe de 14 MHz de bande dont les composantes sont codées en virgule fixe sur 16 bits.

Ce signal est transmis à un FPGA qui se charge de l'estimation de la densité spectrale de puissance par méthode de Welch, conduisant à une structure de récepteur de type FX<sup>4</sup> [73, 10, 9, 74]. La résolution fréquentielle est fixée par le nombre de canaux de la FFT et la résolution temporelle par le nombre d'accumulations de spectres élémentaires.

<sup>4</sup>Estimation spectrale par accumulation de transformées de Fourier. En opposition aux récepteur de type XF où l'estimation s'effectue par la transformée de Fourier de l'autocorrélation du signal d'étude.

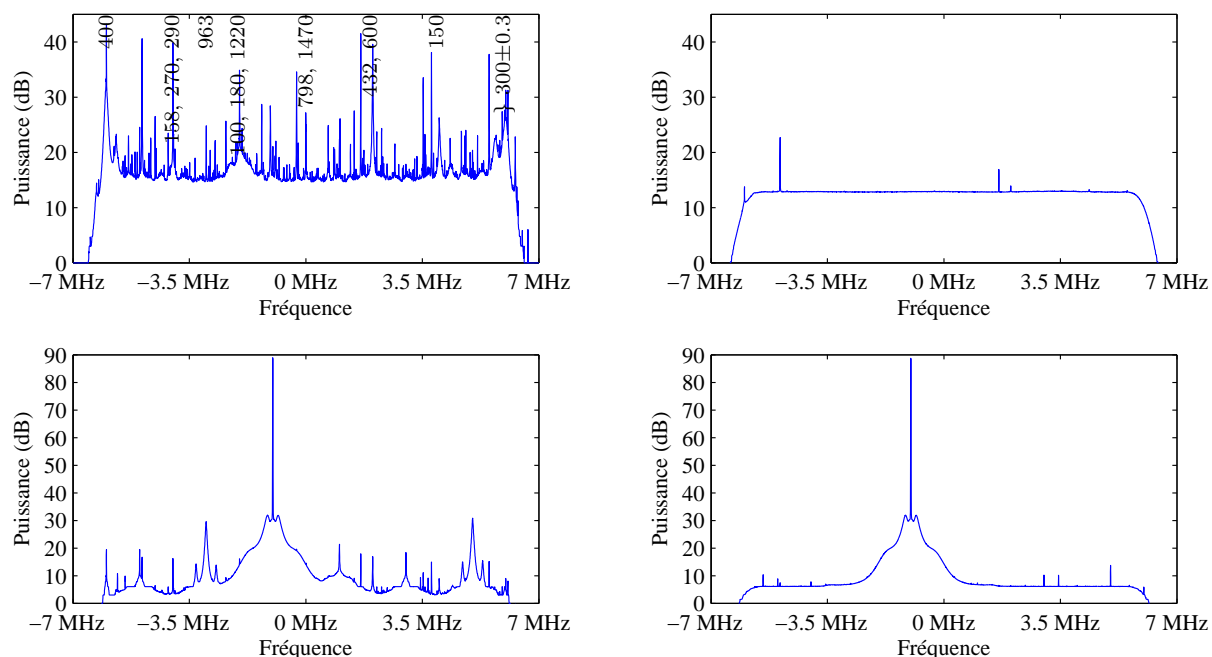


FIG. 1.20 – Test comparatifs des ADCs : (gauche) ADC imparfaitement mis en œuvre. (droite) ADC en environnement blindé. (haut) Chargé sur 50 ohms. (bas) Raie à 69 MHz d’amplitude proche de la pleine échelle, repliée à -1 MHz après échantillonnage et DDC. Le large pied de raie est dû au bruit de phase du générateur et au jitter d’horloge. Les nombres sur la figure en haut à gauche représentent la fréquence des raies (en MHz) avant repliement.

Les cartes de conversion analogique-numérique internes aux PCs ont été remplacés par des modules externes. Ils sont constitués de cartes d’évaluation d’ADCs AD9446 (Analog Devices) scellées dans des caissons blindés avec alimentations séparées. Un bus LVDS transporte l’horloge, les données 16 bits et le bit de dépassement vers chaque PC hôte. Les nombreuses raies systèmes identifiées comme propre au PC (bus PCI 66 MHz, horloge PC à 800 MHz, et leurs nombreuses harmoniques) sont maintenant supprimées. Le système voit sa dynamique d’entrée utilisable augmentée de 20 dB.

Il faut noter que le nombre d’accumulations est limité par les imperfections du système (figure 1.21).

Le flux de données devient une succession de spectres qui forment un plan temps fréquence codé en virgule fixe sur 31 bits<sup>5</sup> auxquels sont ajoutés des trames intégrant contrôles d’erreur et informations sur le flux. Il est possible d’effectuer des traitements d’interférences juste en amont de cet endroit en utilisant les ressources restantes du FPGA (§3.1).

<sup>5</sup>Chaque accumulation nécessite une augmentation de la largeur du bus de donnée. Il est commun en télécommunication de supprimer les bits de poids faibles mais puisque la radioastronomie est principalement intéressée par des signaux situés juste au dessus du bruit, voire dans le bruit, il convient plutôt de supprimer les bits de poids fort en effectuant une saturation du signal.

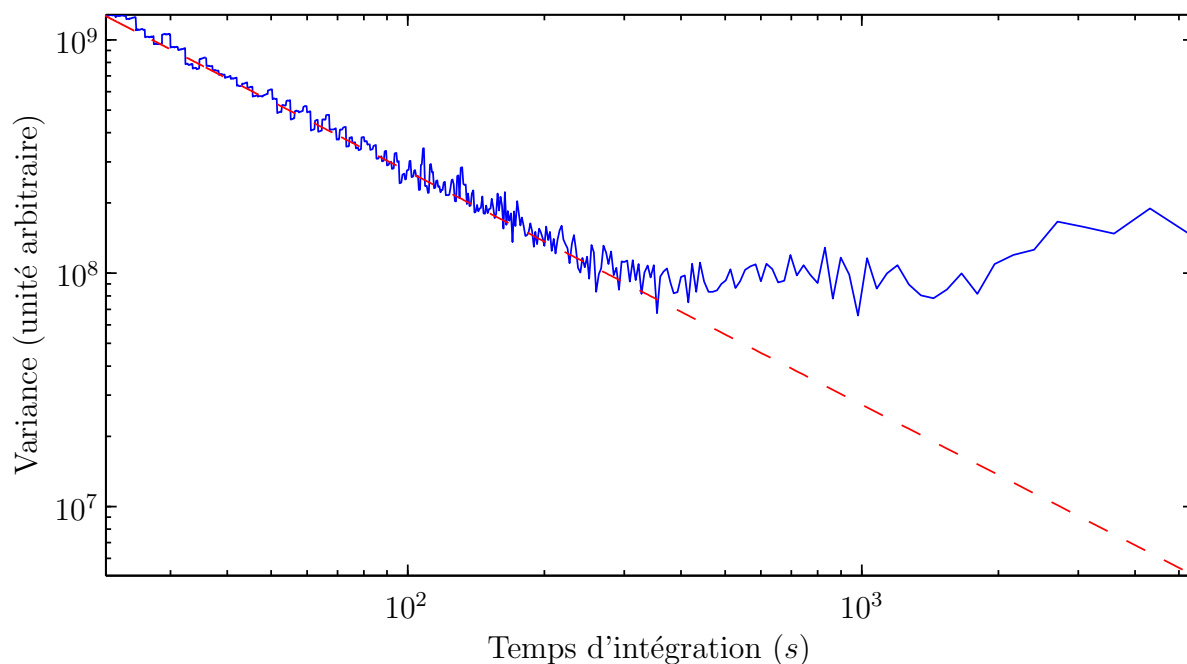


FIG. 1.21 – Variance de Allan du récepteur RDH. Lors d’une intégration de durée  $\tau$ , l’écart type du bruit de mesure diminue en  $1/\sqrt{\tau}$  (§1.4.1.4) mais les imperfections du système limitent cette décroissance dans le temps jusqu’à inverser la tendance et faire croître de nouveau l’écart type. La variance de Allan est communément utilisée pour estimer la stabilité d’un enregistreur [2]. Elle est tracée ici pour le récepteur RDH. Le temps garanti de stabilité est de 200 s, durée suffisante pour les utilisations effectuées sur le NRT ou le NDA. La durée de stabilité concorde avec la valeur attendue pour l’horloge rubidium servant de référence.

Les plans temps-fréquence sont transférés au PC *via* les deux DSPs du système qui peuvent héberger d’autres algorithmes de traitement.

Une des autres raisons qui a poussé à l’utilisation de la méthode FX plutôt que XF est la possibilité, en modifiant légèrement la structure, de coupler deux voies pour traiter une paire de polarisations orthogonales et obtenir ainsi l’information de polarisation du signal. Le flux de données devient un quadruplet de plans temps-fréquence : deux densités spectrales de puissance et le produit croisé complexe. Ces données sont suffisantes pour calculer les types et taux de polarisation ou les paramètres de Stokes [34, 64, 33, 32] qui ne sont que des représentations diverses de la même information.

Une présentation détaillée du récepteur est donnée en [72]. Les résolutions d’analyse, principalement limitées par les PC hôtes, sont données dans le tableau 1.2 :

TAB. 1.2 – Résolutions d’analyse disponibles avec RDH.

Paramètres	Config 1	Config 2	Config 3
Bande	14 MHz	14 MHz	875 kHz
Résolution temporelle	585 $\mu$ s	73 $\mu$ s	9 ms
Résolution fréquentielle	6.8 kHz	54.7 kHz	106.8 Hz

Ce récepteur entièrement construit autour de cartes reconfigurables permet l'implantation des algorithmes de traitement des interférences décrits ci-après.



## CHAPITRE 2

---

# DÉTECTION EN PUISSANCE SUR DES SPECTRES DYNAMIQUES

---

Le traitement du signal est une discipline rendue très populaire dans le milieu universitaire par les nouvelles applications qu'il a permis de développer en télécommunication. Dans ce domaine, les procédés mis en œuvre visent à générer un signal contenant de l'information, l'encoder physiquement sur une onde électromagnétique (par exemple), étudier son passage dans le canal de transmission, effectuer des traitements visant à annuler les effets du canal, atténuer le bruit ajouté par la chaîne de transmission et décoder l'information. Les interférences qui peuvent apparaître sur le signal reçu sont issues de phénomènes naturels ou de la présence d'autres opérateurs de télécommunication.

Le traitement du signal en radioastronomie a la particularité de ne pouvoir agir sur le signal généré. De plus, l'information n'est pas codée dans une onde électromagnétique mais c'est l'onde elle-même qui est l'information. Les caractéristiques temporelles et fréquentielles de ces ondes sont variables mais on peut généralement considérer que nous sommes en présence d'un bruit. Le traiteur de signal en radioastronomie sera donc particulièrement précautionneux vis-à-vis du bruit et ne cherchera en aucun cas à l'éliminer comme un résidu sans valeur.

Un exemple classique de cette opposition de procédés est trouvé lors du développement des premiers systèmes radar qui se virent particulièrement gênés par des émissions solaires. A la fin de la guerre, c'est ce même type de matériel qui fut utilisé pour effectuer les premières observations scientifiques de notre astre.

Ce chapitre présente l'étude d'estimateurs de moyenne et leurs utilisations dans la restauration de sources naturelles qui auraient été corrompues par des interférences intermittentes, en temps et/ou en fréquence. On traitera le cas général de données représentées en plan temps-fréquence, la forme d'onde ne sera qu'un cas particulier où le nombre de canaux vaut un.

Une statistique de décision est calculée sur la puissance du signal afin de décider de la présence ou de l'absence d'une interférence dans le bloc de données étudié. Si le classifieur penche en faveur de la présence d'une interférence, nous considérerons le bloc comme irrémédiablement perdu et il sera écarté des données scientifiques par sa mise à zéro. On parle de "blanking".



## 2.1 Position du problème

Dans la suite de ce chapitre, nous verrons que la discrimination de la présence ou de l'absence d'interférences n'est pas aisée du fait de la multiplicité des types d'émissions et donc des modèles qu'il faudrait élaborer. On préférera plutôt vérifier que nous sommes en présence d'un bruit seul ou bien d'un bruit auquel s'est ajoutée une contribution que l'on ne cherchera pas à exactement modéliser.

Nous attacherons donc une attention particulière à l'étude du bruit que nous supposons blanc gaussien. La statistique de décision  $C_s$  sera choisie comme la puissance du signal. La distribution étudiée sera donc une loi du  $\chi^2$  et nous verrons qu'il est possible de calculer un seuil  $S$  auquel sera confronté la statistique de décision  $C_s$  en effectuant seulement une estimation de la moyenne de la distribution. Le point délicat sera de pouvoir assurer cette estimation de façon exacte que l'on soit en présence d'interférences ou non.

### 2.1.1 Contraintes opérationnelles

Les travaux de cette thèse ont une coloration instrumentale forte. De très nombreux travaux ont étudié, et étudient encore, le problème de l'atténuation d'interférences pour la restauration des sources radioastronomiques [3, 24, 23, 18, 43, 14, 6, 8, 56]. Néanmoins, ces travaux peuvent être très complexes à implanter et peuvent imposer une modification profonde du récepteur pour fonctionner, ou bien nécessitent des moyens que la technologie ne nous fournira que dans une dizaine d'années. Le problème levé par les interférences radio-électriques affecte les observations depuis plusieurs années. Pour permettre aux instruments de jouer pleinement leur rôle, nous verrons que dans certains cas, il est possible aujourd'hui d'appliquer des traitements suffisamment avancés pour garantir des observations de qualité dans des bandes réputées difficilement utilisables.

#### 2.1.1.1 CHOIX DES TRAITEMENTS

Les observations avec traitements d'interférences doivent rester le plus transparent possible vis-à-vis des utilisateurs finaux que sont les radio-astronomes. Il est toutefois dangereux de vouloir enfouir complètement de tels outils dans les systèmes au risque d'oublier leur présence : les traitements, utilisés dans de mauvaises conditions, peuvent rendre le récepteur aveugle à des signaux naturels. C'est pourquoi il est laissé à l'utilisateur le choix de tel ou tel outil pour contrer un type de parasite donné. A lui d'être sûr que les sources naturelles observées ne seront pas altérées par les traitements.

#### 2.1.1.2 INTÉGRATION AUX INSTRUMENTS

Néanmoins, la mise en place d'une observation ne doit pas devenir rédhibitoire de par une configuration du système trop lourde ou des manœuvres annexes trop coûteuses en temps. C'est pourquoi les tâches sont entièrement automatisées et configurées dans un court fichier texte facilement modifiable et très souvent réutilisable tel quel. Le récepteur numérique n'est vu par les télescopes que comme un nouvel enregistreur analogique classique.

Le récepteur numérique de son côté fonctionne en condition réelle, sur des signaux qui sont loin d'être idéaux. Par exemple, nous verrons que l'hypothèse d'un bruit blanc

gaussien n'est plus vérifiée lors du traitement des interférences intermittentes à bande étroite (section 3.2). Cela nécessitera des traitements appropriés pour garantir l'intégrité des sources naturelles.

### 2.1.1.3 TRAITEMENTS TEMPS RÉEL

Une dernière contrainte est celle de la disponibilité du système. Il serait techniquement possible d'enregistrer le signal d'antenne à très haute résolution pour ensuite appliquer des méthodes très évoluées de débruitage (figure 2.1, haut). Les volumes de données acquis et le coût du matériel informatique nécessaire à la gestion de ces données deviendraient prohibitifs. De plus, l'utilisation intensive d'un tel système serait dégradée : en plus de l'acquisition des données se grefferaient des temps de transferts par réseau et des temps de post-traitements rédhibitoires qui rendraient le système inutilisable pour une exploitation courante. À titre d'exemple, la résolution requise pour résoudre une impulsion radar est inférieure à la microseconde et quelques dizaines de nanosecondes sont nécessaires pour une détection efficace. Sur RDH, cela correspond à un flot de 56 Mo/s. Une observation typique sur le RT dure 1 heure avec un temps effectif d'observation de 35 à 45 minutes, soit  $\approx 120$  Go de données qu'il faudrait éventuellement re-transférer et débruiter.

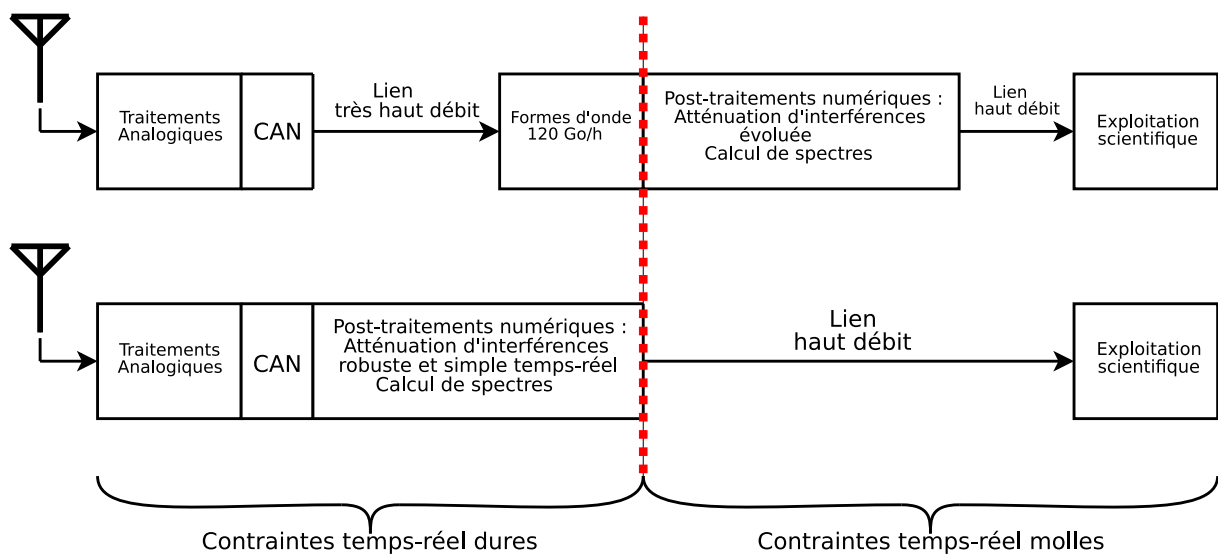


FIG. 2.1 – Illustration des traitements en temps différé ou en temps réel : les traitements en temps différé nécessitent la gestion de très haut débits et le stockage de volumes de données importants. Le temps réel peut réduire ces contraintes mais impose des contraintes temporelles serrées sur l'exécution des calculs.

Dans le cas d'observations effectuées par une antenne unique, l'information principalement recherchée par les radioastronomes (si on exclut la recherche et le chronométrage de pulsars) est le profil spectral des sources. Comme l'estimation du profil est généralement obtenue par  $Acc$  intégrations de spectres élémentaires issus de la forme d'onde, on réduit le débit de données d'un facteur  $\approx Acc$ . On peut contourner la contrainte de flux et de stockage de forme d'onde si on effectue le traitement des interférences et l'estimation spectrale à flux tendu, pendant l'observation, sans devoir effectuer un stockage massif intermédiaire (figure 2.1, bas). Cela introduit une nouvelle contrainte : des traitements relativement

simples qui doivent permettre un fonctionnement en temps réel de l'ensemble du système. La notion de temps réel implique que le temps de calcul affecté à une procédure doit rester inférieur à une limite haute afin de permettre un fonctionnement continu de la chaîne globale sans aucune perte de données. Les calculs les plus lourds sont confiés à des composants numériques programmables de type DSP et/ou FPGA, seuls à pouvoir maintenir le rythme soutenu des traitements envisagés. Les composants programmables présentent néanmoins deux principaux inconvénients : des ressources mémoires limitées et des outils de développement plus complexes à mettre en œuvre. Les tâches présentant des contraintes temporelles plus souples telles que la programmation du système ou des traitements de faibles flots de données sont laissées à des logiciels informatiques exécutés sur des architectures de type PC.

### 2.1.2 Problème de détection

La décision de “blanking” pose, en fait, un problème classique de détection à deux hypothèses. Il s'agit, à partir d'une observation, de déceler la présence ou l'absence d'un brouilleur radioélectrique dans un échantillon en vue de décider ou non, par sa mise à zéro, de la prendre en compte dans la suite des traitements. Le problème se résume ainsi :

$$\begin{cases} \text{Hypothèse } H_0 : s = u & \rightarrow \text{signal utile seul} \\ \text{Hypothèse } H_1 : s = u + RFI & \rightarrow \text{signal utile parasité} \end{cases} \quad (2.1)$$

avec  $s$ , le signal reçu

$u$ , le signal utile

et  $RFI$ , le signal interférant (Radio Frequency Interference)

Pour tester quelle hypothèse est vraie, une statistique de détection  $C_S$  est calculée sur l'observation  $s$ . Le critère de décision utilisé est une comparaison à un seuil  $S$  ce qui génère une variable de décision binaire  $d_s$  telle que :

$$\begin{cases} C_S < S & \rightarrow d_s = 0 & \rightarrow \text{Décision en faveur de } H_0 \\ C_S \geq S & \rightarrow d_s = 1 & \rightarrow \text{Décision en faveur de } H_1 \end{cases} \quad (2.2)$$

Une fois que la décision est prise quant au contenu de  $s$ , quatre cas sont envisageables:

- L'hypothèse vraie est  $H_0$  et  $d_s = 0$
- L'hypothèse vraie est  $H_1$  et  $d_s = 1$
- L'hypothèse vraie est  $H_0$  et  $d_s = 1$   $\rightarrow$  Fausse alarme ( $P_{fa}$ )
- L'hypothèse vraie est  $H_1$  et  $d_s = 0$   $\rightarrow$  Non-détection ( $P_{nd}$ )

Il existe donc deux cas qui correspondent à des erreurs de classification : les fausses alarmes et les non-détections (figure 2.2).

Les non-détections (surface bleu) correspondent aux parasites qui réussissent à franchir le classifieur sans être détectés. L'incidence sur les observations est une augmentation de l'INR (Interference to Noise Ratio), soit une distorsion de la bande d'analyse.

Les fausses alarmes (surface rouge) correspondent aux échantillons qui sont classifiés comme parasités alors que l'hypothèse vraie est  $H_0$ . L'incidence sur le résultat final est une diminution du temps d'observation du signal utile.

En général, pour la radioastronomie, on privilégie une  $P_{nd}$  la plus faible possible pour éviter de perdre toute l'observation, au détriment de la  $P_{fa}$  qui réduit le temps d'observation de façon mineure.

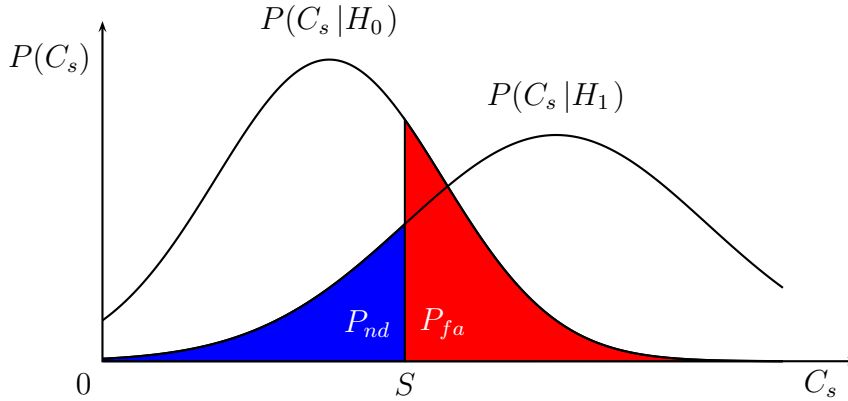


FIG. 2.2 – Les deux classes  $H_0$  et  $H_1$  se superposent. Les échantillons de  $H_0$  dont la statistique de détection est supérieure à  $S$  sont donc mal classifiés. Ce sont des fausses alarmes. À l’opposé, les échantillons mal classifiés appartenant à  $H_1$  sont des non-détections.

La connaissance des densités de probabilité de la statistique  $C_s$  sous  $H_0$  et  $H_1$  et l’importance (la pondération) accordée aux fausses alarmes et aux non-détections peuvent permettre le choix optimal du seuil  $S$ . Le problème se réduit à minimiser la somme des deux surfaces pondérées [71].

Au vu de la diversité des densités de probabilité de parasites, il n’est pas envisageable de construire un modèle commun sous peine de dégrader les performances du détecteur.

On peut envisager des lois  $P(C_s | H_1)$  spécifiques à chaque type d’interférence. Par exemple, en choisissant une résolution temporelle et fréquentielle suffisamment fine, on peut considérer que l’interférence est constante dans le bloc de données. Son comportement est donc modélisable par une distribution certaine sous la forme d’un Dirac avec comme seul paramètre la puissance  $P_{RFI}$  de l’interférence. Mais même ce modèle simpliste rend impossible le calcul d’un seuil optimal car on ne peut que difficilement connaître l’évolution de  $P_{RFI}$  au cours du temps (son estimation est possible si on utilise des méthodes itératives / recuit simulé). Bref, les lois  $P(C_s | H_1)$  ne peuvent être fiables.

Pour résumer, il est théoriquement possible de parfaitement modéliser  $P(C_s | H_0)$  et  $P(C_s | H_1)$  mais, en pratique, il est difficile de connaître  $P(C_s | H_0)$  et impossible de maîtriser  $P(C_s | H_1)$ .

On décide donc de se baser sur la connaissance seule de  $P(C_s | H_0)$ . Cette connaissance sera entaché par un volume de parasites plus ou moins présents dans l’observation brute. Il faudra développer différentes versions d’estimateurs de paramètres pour la loi  $H_0$  perturbée par un mélange partiel avec  $H_1$ .

### 2.1.3 Critère de décision

Pour les besoins scientifiques courants, l’analyse spectrale des signaux radioastronomiques est une étape souvent nécessaire. Les récepteurs sont donc menés à effectuer en routine l’estimation de densités spectrales de puissance et d’intercorrélations. Il se trouve

que l'observation effectuée constitue une assez bonne statistique de détection. De plus, nous n'avons pas vraiment le choix à cause des contraintes temps réel qui interdisent de prendre des statistiques plus sophistiquées. Dans le cadre de cette étude, ce sont les puissances instantanées ou les densités spectrales de puissance qui vont être utilisées comme statistiques de détection. Selon le cas, le problème de détection 2.1 devient :

$$\begin{cases} H_0 : P_s(t) < S \\ H_1 : P_s(t) > S \end{cases} \quad (2.3)$$

ou

$$\begin{cases} H_0 : P_{Acc}(f) < S \\ H_1 : P_{Acc}(f) > S \end{cases} \quad (2.4)$$

avec  $P_s(t)$ , la puissance du signal  $s$  à l'instant  $t$ ,  
et  $P_{Acc}(f)$ , le périodogramme moyenné du signal  $s$ .

Dans le problème 2.3, on estime la puissance instantanée du signal  $s$  dans une bande de fréquence centrée sur  $f_0$ . La bande peut être la bande complète de  $s$  (application 3.1) ou seulement une de ses sous-bandes qui aura été obtenue par filtrage ou par FFT (application 3.2). La résolution de ce problème permet de traiter le cas des interférences intermittentes. Le découpage en sous-bande du signal  $s$  comportant des interférences intermittentes à bande étroite permet d'obtenir un ensemble de signaux comportant des interférences intermittentes à bande large. Il faudra résoudre le problème de détection pour chacune des sous-bandes pour traiter  $s$  dans sa globalité.

Le problème 2.4 traite de l'estimation robuste du spectre de  $s$  à l'instant  $t = t_0$ . Il sera appliqué en section 3.3.

Le récepteur numérique est configuré pour fournir des plans temps-fréquence selon la méthode indiquée en section 1.4.2, page 42 en notant que les études seront effectuées pour un recouvrement de 0%.

Les plans temps-fréquence  $P_{Acc}[t, f]$  sont constitués, entre autres, de juxtaposition temporelle de densités spectrales de puissance  $P_{Acc}[f]$  constituées de l'accumulation des modules carrés de  $Acc$  spectres élémentaires  $X_i[f]$  suivant la technique du périodogramme moyenné sans recouvrement tel que :

$$P_{Acc}[f] = \sum_{i=1}^{Acc} X_i[f] X_i^*[f] \quad (2.5)$$

$$\text{ou } P_{Acc}[f] = \sum_{i=1}^{Acc} (\Re(X_i[f])^2 + \Im(X_i[f])^2) \quad (2.6)$$

Dans cette écriture, les spectres élémentaires  $X_i[f]$  sont calculés par FFT sur des tranches temporelles successives. À la fin des  $Acc$  accumulations, une densité spectrale de puissance est générée à la date  $t$  de sorte qu'un plan temps-fréquence  $P_{Acc}[t, f]$  puisse être rempli, pour des dates  $t$  successives. Le facteur de normalisation de 2.6 par  $Acc^{-1}$  n'est pas inclus ici car la normalisation est effectuée sur des critères techniques visant à éviter

la saturation numérique des bus de données ou l'arrondi *sauvage* de l'information. Les niveaux absolus sont connus par injection de signaux calibrés dans la chaîne analogique.

Selon les dimensions des vecteurs  $t$  et  $f$ ,  $C_s$  pourra prendre les formes suivantes :

$$C_s = P_{Acc}[f] \text{ pour le périodogramme moyenné} \quad (2.7)$$

$$C_s = P_{Acc}[t] = |x_{IQ}[t]|^2 \text{ pour la puissance instantanée} \quad (2.8)$$

de la forme d'onde complexe  $x_{IQ}[t]$

Pour étudier la distribution des valeurs de  $P_{Acc}$ , indépendamment des dimensions de  $t$  et  $f$ , nous traiterons le cas du  $N$ -échantillon  $P_{Acc}$  constitué de  $N$  valeurs  $P_{Acc}[k], k \in [1, \dots, N]$ . Selon que l'on traitera le cas de la puissance d'une forme d'onde ou d'un périodogramme moyenné,  $P_{Acc}[k]$  aura pour expression 2.8 ou 2.5.

Nous avons vu en section 1.4.1.2 que le bruit mesuré par le système pouvait être dominé par le bruit thermique de l'électronique ou bien par le bruit du ciel. Le bruit thermique de l'électronique, ou bruit de Johnson-Nyquist, est modélisé comme un bruit blanc gaussien. De même, la nature de la génération d'une majorité des signaux naturels en fait aussi un bruit blanc gaussien [62, 5]. Le signal perçu est la somme de multiples interactions indiscernables dont les densités de probabilité individuelles se convoluent pour tendre vers une distribution gaussienne.

Néanmoins, il existe certains phénomènes qui ne peuvent pas être décrits par ce modèle. On peut citer l'exemple des raies d'émission qui sont bien évidemment à bande étroite, mais leur intensité est si ténue qu'elles sont encore noyées dans le bruit système alors que s'effectueront les procédures d'atténuation d'interférences. L'hypothèse du bruit blanc gaussien tient donc toujours. Il nous faut néanmoins présenter ce dernier exemple qui montrera les précautions qu'il faut prendre avant d'utiliser des algorithmes de traitement d'interférences : les signaux provenant du couple Jupiter-Io sont en partie polarisés linéairement [16]. Leur interaction avec notre ionosphère connus sous le nom d'effet Faraday produit une modulation du spectre en une série de franges. Le signal n'est donc plus blanc. Il le redevient si la bande d'analyse est réduite.

Nous pouvons donc considérer que l'hypothèse d'un bruit blanc gaussien est valide dans une grande majorité des cas mais parfois il faudra prêter attention aux corrections à effectuer pour pouvoir continuer d'utiliser cette propriété.

L'expression 2.6 montre que  $P_{Acc}$  est obtenu par la sommation de  $2Acc$  variables ( $\Re(X_i)$  ou  $\Im(X_i)$ ) élevées au carré. Les variables aléatoires  $\Re(X_i)$  et  $\Im(X_i)$  suivent une loi gaussienne centrée, sont indépendantes (cas du recouvrement nul lors de l'estimation du périodogramme) et de variance  $\sigma^2$ . La variable aléatoire  $P_{Acc}$  suit donc une loi du  $\chi^2$  à  $2Acc$  degrés de liberté.

### 2.1.4 Calcul de seuil sur une loi du $\chi^2$

La loi du  $\chi^2$  est abondamment décrite dans la littérature [41]. Nous redécrivons néanmoins la loi pour un signal de distribution gaussienne centrée réduite puis pour un signal de distribution gaussienne centrée non-réduite afin de montrer qu'il est possible de calculer un seuil  $S$  basé uniquement sur l'estimation de la moyenne  $E\{\chi_{\nu,\sigma}^2\}$  lorsque  $\nu$  est fixé et que la fausse alarme  $P_{fa}$  est choisie.

Soient  $X_1, X_2, \dots, X_\nu$ ,  $\nu$  variables aléatoires indépendantes de distribution normale centrée réduite ( $E\{X\} = 0$  et  $Var(X) = E\{X^2\} = 1$ ). Soit  $\chi_\nu^2$ , la somme de leur carré :

$$\chi_\nu^2 = X_1^2 + X_2^2 + \dots + X_\nu^2 \quad (2.9)$$

La loi du  $\chi^2$  à  $\nu$  degrés de liberté est la loi de probabilité de la variable  $\chi_\nu^2$  ainsi construite. Sa densité de probabilité est donnée par 2.10 et est visible en figure 2.3 pour diverses valeurs de  $\nu$ . Dans le cadre de cette étude, seules les valeurs paires de  $\nu$  sont d'intérêt.

$$f_{\chi_\nu^2}(x) = \begin{cases} \frac{1}{2\Gamma(\nu/2)} \left(\frac{x}{2}\right)^{\nu/2-1} e^{-x/2} & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (2.10)$$

avec  $\Gamma(\nu)$ , la fonction Gamma [1].

Quand  $\nu$  augmente, la densité de probabilité  $f_{\chi_\nu^2}(x)$  tend rapidement vers une distribution gaussienne.

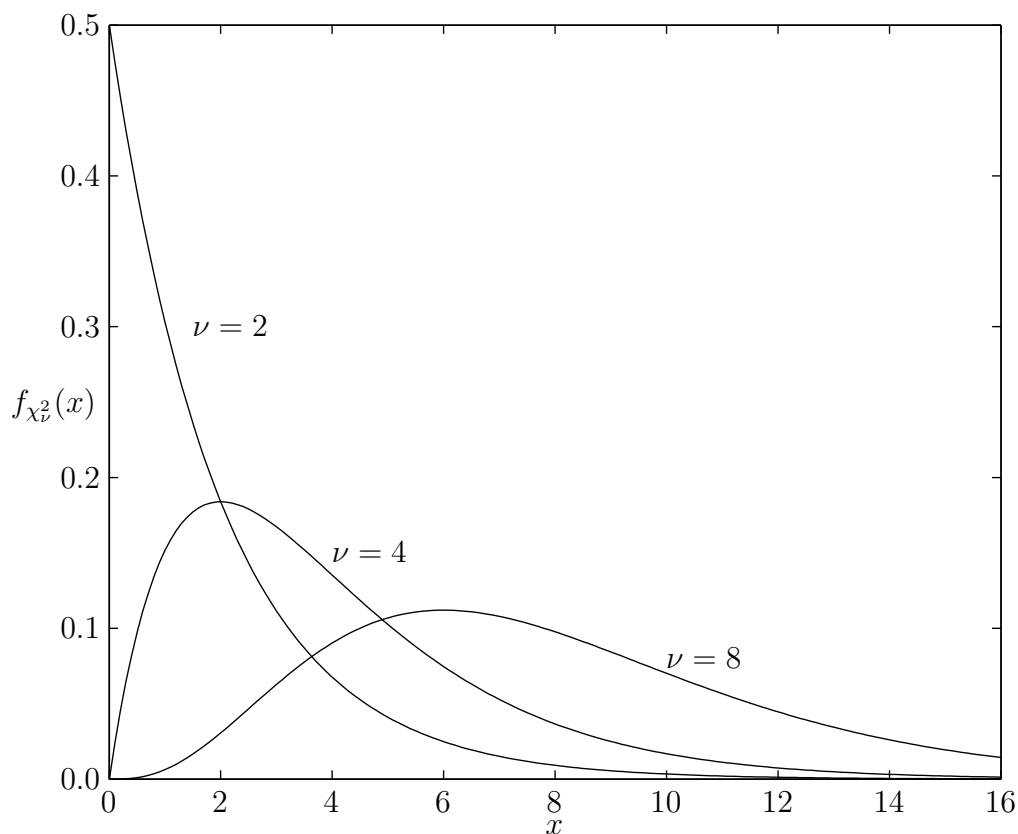


FIG. 2.3 – Densité de probabilité de la loi du  $\chi^2$  pour  $\nu=[2,4,8]$  degrés de liberté

La fonction de répartition de la loi du  $\chi^2$  est donnée par :

$$F_{\chi_\nu^2}(x) = \begin{cases} P_{\Gamma(\frac{\nu}{2}, \frac{x}{2})} & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (2.11)$$

où la fonction gamma incomplète normalisée  $P_\Gamma$  s'exprime comme :

$$P_\Gamma(a, b) = \frac{\Gamma(a, b)}{\Gamma(a)} = \frac{1}{\Gamma(a)} \int_0^b t^{a-1} e^{-t} dt \quad (2.12)$$

Le moment non-centré d'ordre 1,  $\mu'_1$ , c'est à dire la moyenne  $\mu$ , de la loi du  $\chi^2$  à  $\nu$  degrés de liberté vaut  $\nu$  :

$$\mu'_1 = E\{\chi_\nu^2\} = E\{X_1^2\} + E\{X_2^2\} + \dots + E\{X_\nu^2\} = \nu \quad (2.13)$$

car  $E\{X^2\} = 1$ .

Les moments centrés  $\mu_i$  de cette loi sont donnés par la formule de récurrence suivante [55] :

$$\begin{aligned} \mu_0 &= 1 \\ \mu_1 &= 0 \\ \mu_j &= 2(j-1)(\mu_{j-1} + \nu\mu_{j-2}) \quad , j \geq 2 \end{aligned} \quad (2.14)$$

En particulier,  $Var(\chi_\nu^2) = \mu_2 = 2\nu$ .

Dans le cas de variables aléatoires indépendantes  $Y_1, Y_2, \dots, Y_\nu$ , de distribution normale mais seulement centrée et présentant une variance  $\sigma^2$  commune,  $\chi_{\nu, \sigma}^2$ , la somme de leur carré vaut :

$$\chi_{\nu, \sigma}^2 = Y_1^2 + Y_1^2 + \dots + Y_\nu^2 \quad (2.15)$$

$$= \sigma^2 (X_1^2 + X_1^2 + \dots + X_\nu^2) \quad (2.16)$$

$$= \sigma^2 \cdot \chi_\nu^2 \quad (2.17)$$

La différence entre les distributions de  $\chi_{\nu, \sigma}^2$  et de  $\chi_\nu^2$  est donc un simple facteur d'échelle. L'espérance et la variance de la variable aléatoire  $\chi_{\nu, \sigma}^2$  valent :

$$E\{\chi_{\nu, \sigma}^2\} = \nu\sigma^2 = \mu_{\text{théo}} \quad (2.18)$$

$$Var(\chi_{\nu, \sigma}^2) = 2\nu\sigma^4 = \sigma_{\text{théo}}^2 \quad (2.19)$$

Pour un taux de fausse alarme  $P_{fa} = P(P_{Acc} > S | H_0) = \delta$  donné où  $P_{Acc} | H_0$  suit une loi du  $\chi_{\nu, \sigma}^2$  et  $\nu$  connu, on calcule un seuil  $S$  tel que (figure 2.4):

$$F_{\chi_{\nu, \sigma}^2}(S) = 1 - \delta \quad (2.20)$$

$$S = F_{\chi_{\nu, \sigma}^2}^{-1}(1 - \delta) \quad (2.21)$$

$$S = 2\sigma^2 P_\Gamma^{-1}(\nu/2, 1 - \delta) \quad (2.22)$$

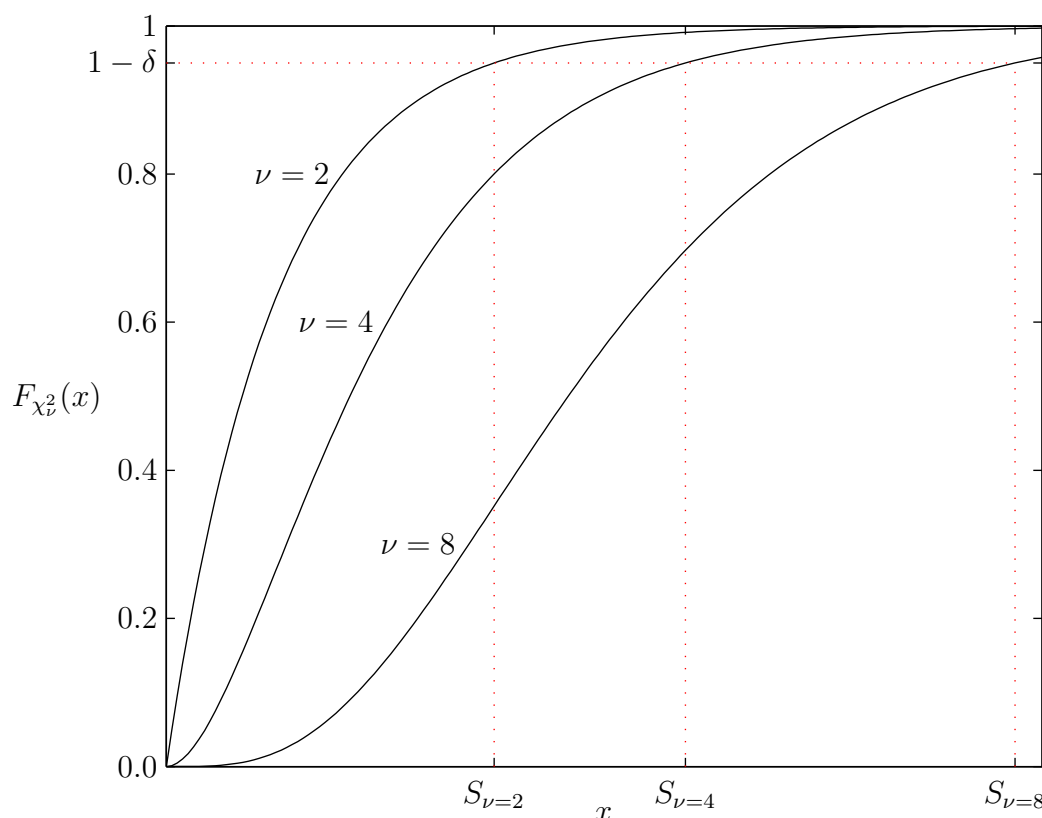
où  $P_\Gamma^{-1}$  est la fonction inverse de  $P_\Gamma(a, b)$  à  $a$  fixé.

Néanmoins,  $\sigma$  n'est pas toujours directement accessible. En introduisant (2.18) dans (2.22), on obtient :

$$\boxed{S = 2 \frac{\mu_{\text{théo}}}{\nu} P_\Gamma^{-1}(\nu/2, 1 - \delta)} \quad (2.23)$$

Le calcul d'un seuil est maintenant uniquement fonction de l'estimation  $\hat{\mu}$  correcte de  $\mu_{\text{théo}}$  car les autres paramètres sont connus. Contrairement à un signal de distribution



FIG. 2.4 – Détermination du seuil de “blanking”  $S$ 

gaussien, le calcul d’un seuil sur une distribution en  $\chi^2$  ne nécessite plus le calcul de la variance, simplifiant ainsi les algorithmes d’estimation.

Le point délicat reste d’effectuer l’estimation de  $\mu_{\text{théo}}$  en présence de parasites. On notera cette estimation  $\hat{\mu}$ . Nous verrons qu’il existe des estimateurs de moyenne qui sont plus ou moins robustes à une certaine quantité de points aberrants. Il faut les étudier en détail pour choisir celui qui conviendra le mieux à un problème.

## 2.2 Estimateurs de moyenne

L’étude de divers estimateurs de la moyenne d’un échantillon  $P_{Acc}[k]$  dont la distribution suit une loi du  $\chi^2$  à  $\nu$  degrés de liberté est présentée dans les sections suivantes. L’influence des interférences radioélectriques sur les estimateurs sera traitée.

De plus, les contraintes d’implémentation comme le nombre d’échantillons disponibles, le pourcentage possible de points parasités ... font qu’un estimateur sera plus ou moins bien adapté à un cas précis.

L’étude de divers estimateurs de moyenne fiables sur diverses distributions a été conduite en [55, section 12.5 : Les statistiques fiables]. Nous étendons cette étude à des estimateurs de notre cru sur une distribution suivant une loi du  $\chi^2$  et sur cette même distribution intégrant des interférences.

### 2.2.1 Critères de comparaison

Pour faciliter la comparaison, les expressions de l'espérance et de la variance de chaque estimateur de moyenne sont dérivées en fonction du degré de liberté  $\nu$  et de la variance  $\sigma^2$  du bruit gaussien centré sur lequel est basée la loi du  $\chi^2$ .

$$E\{\hat{\mu}\} = E\{\hat{\mu}\}_{\nu,\sigma} \quad (2.24)$$

$$Var(\hat{\mu}) = Var(\hat{\mu})_{\nu,\sigma} \quad (2.25)$$

$$Var(\hat{\mu}) = E\{\hat{\mu}^2\}_{\nu,\sigma} - E\{\hat{\mu}\}_{\nu,\sigma}^2 \quad (2.26)$$

La robustesse des divers estimateurs face à des interférences de divers niveaux et de divers taux d'occupation est étudiée.

On pose  $f_{RFI}(x)$  (respectivement  $f_{RFI}(x)$ ,  $P_{RFI}$  et  $\sigma_{RFI}^2$ ), la densité de probabilité (respectivement la fonction de répartition, la moyenne et la variance) de la loi de la puissance du signal parasitant. Lorsque l'interférence est présente, sa puissance est sommée au niveau de l'antenne à la puissance du bruit blanc gaussien du système (bruit électronique + bruit du ciel). Nous effectuons cette hypothèse en considérant que l'interférence n'est pas corrélée au bruit mesuré. Donc, sous l'hypothèse  $H_1$ , la densité de probabilité  $f_1(x)$  de la variable aléatoire résultante est la convolution des deux densités de probabilité [55, section 4.3.1]. Par symétrie, on pose  $f_0(x) = f_{\chi^2_{\nu,\sigma}}(x)$  la densité de probabilité de la variable aléatoire sous l'hypothèse  $H_0$  et on définit de la même manière  $P_0(x)$ ,  $P_1(x)$ ,  $F_0(x)$  et  $F_1(x)$ . De même, on notera  $P_{Acc_0}[x]$  le  $N$ -échantillon composé des  $N$  valeurs réalisées sous  $H_0$ .

$$f_1(x) = f_0(x) * f_{RFI}(x) \quad (2.27)$$

Nous avons vu que les traitements envisagés n'étaient possibles qu'en cas de parasites à occupation temporelle intermittente (et, par extension, aux parasites à bande étroite). La variable aléatoire  $X$  est donc en partie issue d'une loi de densité de probabilité  $f_0(x)$  et d'une loi de densité de probabilité  $f_1(x)$ . On pose  $\alpha$ , la probabilité d'occurrence d'une interférence dans le jeu de données tel que :

$$P(H_1) = \alpha \quad (2.28)$$

$$\text{et } P(H_0) = 1 - \alpha \quad (2.29)$$

La probabilité  $P(X < x)$  pour que la variable aléatoire  $X$  issue à  $\alpha\%$  d'une loi de densité de probabilité  $f_1(x)$  et à  $(1-\alpha)\%$  d'une loi de densité de probabilité  $f_0(x)$  vaut :

$$P(X < x) = P\left[\left((X < x) \cap (x \in H_0)\right) \cup \left((X < x) \cap (x \in H_1)\right)\right] \quad (2.30)$$

$$P(X < x) = P(X | H_0 < x) + P(X | H_1 < x) \quad (2.31)$$

$$P(X < x) = P(H_0)P_0(X < x) + P(H_1)P_1(X < x) \quad (2.32)$$

On peut donc en déduire que la densité de probabilité  $f(x)$  de la loi d'une variable aléatoire issue d'une loi du  $\chi^2$  corrompue à  $\alpha\%$  par une interférence de densité de probabilité  $f_{RFI}(x)$  vaut :

$$f(x) = (1 - \alpha)f_0(x) + \alpha f_1(x) \quad (2.33)$$

Sur le  $N$ -échantillon  $P_{Acc}[k]$ , seuls  $N_{RFI}$  points subissent un parasitage additif. On peut donc écrire :

$$f(x) = \frac{N - N_{RFI}}{N} f_0(x) + \frac{N_{RFI}}{N} f_1(x) \quad (2.34)$$

La densité de probabilité  $f(x)$  est la dérivée de la fonction de répartition  $F(x)$ . La fonction de répartition vaut donc:

$$f(x) = \int_{-\infty}^x f(t) dt \quad (2.35)$$

$$f(x) = (1 - \alpha) \int_{-\infty}^x f_0(t) dt + \alpha \int_{-\infty}^x f_{RFI}(t) * f_0(t) dt \quad (2.36)$$

$$f(x) = (1 - \alpha) f_0(x) + \alpha f_{RFI}(x) * f_0(x) \quad (2.37)$$

$$F(x) = (1 - \alpha) F_0(x) + \alpha F_1(x) \quad (2.38)$$

La figure 2.5 illustre les effets du mélange additif d'une loi du  $\chi^2$  avec d'autres lois pour  $\alpha = 10\%$ .

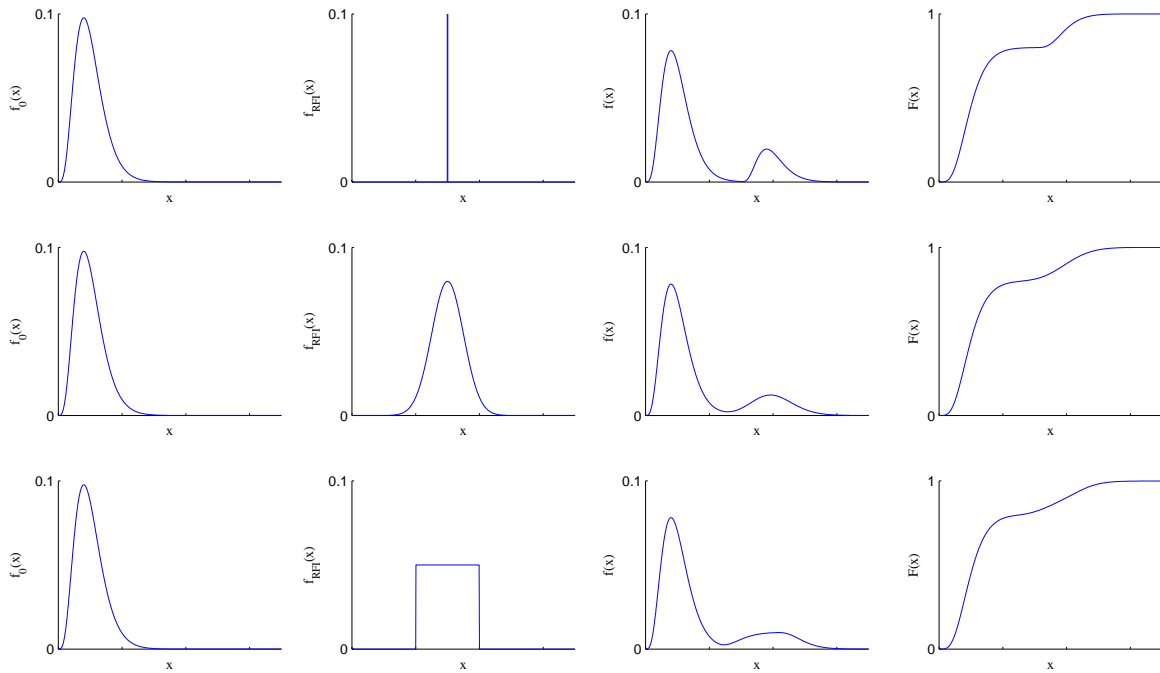


FIG. 2.5 – Densité de probabilité  $f(x)$  et fonction de répartition  $F(x)$  obtenues par mélange du signal naturel de densité de probabilité  $f_0(x)$  à un parasite de densité de probabilité  $f_{RFI}(x)$  dans le rapport 90%/10%. En haut : loi de Dirac, au milieu : loi gaussienne et en bas : loi uniforme.

L'estimation de la moyenne  $\hat{\mu}$  sur les données brutes polluées partiellement par un niveau  $\alpha\%$  de parasites vient modifier les propriétés statistiques de l'estimateur  $\hat{\mu}_0$  exempt de parasites. Pour quantifier et comparer la dérive des estimateurs en présence de parasites, on dérivera les expressions de l'espérance  $E\{\hat{\mu}\}$  et de la variance  $Var(\hat{\mu})$  pour pouvoir

évaluer les rapports suivants en fonction de  $\alpha = N_{RFI}/N$  et  $P_{RFI}$  :

$$\Delta_\mu = \frac{E\{\widehat{\mu}\}}{E\{\widehat{\mu}_0\}} \quad (2.39)$$

$$\Delta_\sigma = \frac{\sqrt{Var(\widehat{\mu})}}{\sqrt{Var(\widehat{\mu}_0)}} \quad (2.40)$$

L'écart par rapport à l'unité de ces indicateurs quantifie la non-robustesse de l'estimateur vis-à-vis des parasites. Une fonction de Dirac est choisie pour la densité de probabilité<sup>1</sup> du parasite afin de simplifier les calculs et les simulations. Avec  $f_{RFI}(x) = \delta(x - P_{RFI})$ , la densité de probabilité et la fonction de répartition à étudier deviennent :

$$f(x) = (1 - \alpha)f_0(x) + \alpha f_0(x - P_{RFI}) \quad (2.41)$$

$$F(x) = (1 - \alpha)f_0(x) + \alpha f_0(x - P_{RFI}) \quad (2.42)$$

Un exemple d'une telle densité de probabilité est disponible en figure 2.6 pour  $\sigma = 1$ ,  $\mu_{\chi^2} = \nu = 16$  et  $P_{RFI} = 6\mu_{\chi^2}$ .

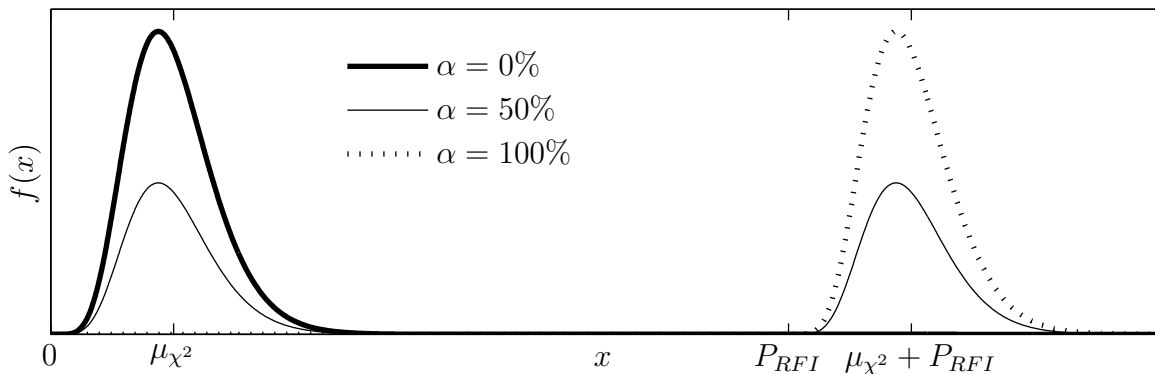


FIG. 2.6 – Densité de probabilité d'une loi en  $\chi^2$  corrompue par une loi de Dirac

Les estimateurs de moyenne décrits ci-après poursuivent le même but mais dans des conditions expérimentales différentes. Il s'agit d'estimer  $\mu_{théo}$  de l'équation 2.23 de façon la plus exacte possible (biais nul), précise (variance faible), en garantissant la meilleure insensibilité et robustesse aux RFI tout en assurant la préservation des sources.

<sup>1</sup>Densité de probabilité correspondant un processus certain puisque  $P(x = P_{RFI}) = 1$

## 2.2.2 Moyenne arithmétique

La moyenne de référence de l'équation 2.23 est théoriquement exprimée comme l'espérance de la distribution de  $P_{Acc_0}$  :

$$\mu_{0MN} = E \{P_{Acc_0}[k]\} = \mu_{théo} \quad (2.43)$$

$$(2.44)$$

### 2.2.2.1 ESPÉRANCE ET VARIANCE DE L'ESTIMATEUR

$\mu_{0MN}$  est classiquement estimé comme 2.45. Son espérance vaut 2.46 et la variance de cet estimateur vaut 2.47.

$$\hat{\mu}_{0MN} = \frac{1}{N} \sum_{k=1}^N P_{Acc_0}[k] \quad (2.45)$$

$$E \{\hat{\mu}_{0MN}\} = \frac{1}{N} \sum_{k=1}^N E \{P_{Acc_0}[k]\} = \mu_{théo} \quad (2.46)$$

$$Var(\hat{\mu}_{0MN}) = \frac{1}{N^2} \sum_{k=1}^N Var(P_{Acc_0}[k]) = \frac{1}{N^2} (N \cdot 2\nu\sigma^2) = \frac{\sigma_{théo}^2}{N} \quad (2.47)$$

La figure 2.7 illustre l'absence de biais de l'estimateur. Sa variance diminue quand  $N$  augmente. Il est donc consistant.

### 2.2.2.2 DÉRIVE DE L'ESTIMATEUR

L'estimateur de moyenne travaille maintenant sur la distribution parasitée de  $P_{Acc}[k]$ . On peut facilement mener les calculs en conservant  $f(x)$  quelconque. La moyenne théorique  $\mu_{MN}$  de cette distribution vaut :

$$\mu_{MN} = E \{P_{Acc}[k]\} \quad (2.48)$$

$$\mu_{MN} = \int_{\mathbb{R}} x [(1 - \alpha) f_0(x) + \alpha f_{RFI}(x) * f_0(x)] dx \quad (2.49)$$

$$\mu_{MN} = (1 - \alpha) \mu_{théo} + \alpha (P_{RFI} + \mu_{théo}) \quad (2.50)$$

$$\mu_{MN} = \mu_{théo} + \alpha P_{RFI} \quad (2.51)$$

La méthode d'estimation de  $\tilde{\mu}_{MN}$  ne change pas. Par contre, la distribution de  $P_{Acc}[k]$  diffère de celle de  $P_{Acc_0}[k]$ . Il faut donc évaluer les expressions d'espérance et de variance de l'estimateur  $\hat{\mu}_{MN}$  :

$$\hat{\mu}_{MN} = \frac{1}{N} \sum_{k=1}^N P_{Acc}[k] \quad (2.52)$$

$$E \{\hat{\mu}_{MN}\} = \frac{1}{N} \sum_{k=1}^N E \{P_{Acc}[k]\} \quad (2.53)$$

$$E \{\hat{\mu}_{MN}\} = \mu_{MN} = \mu_{théo} + \alpha P_{RFI} \quad (2.54)$$

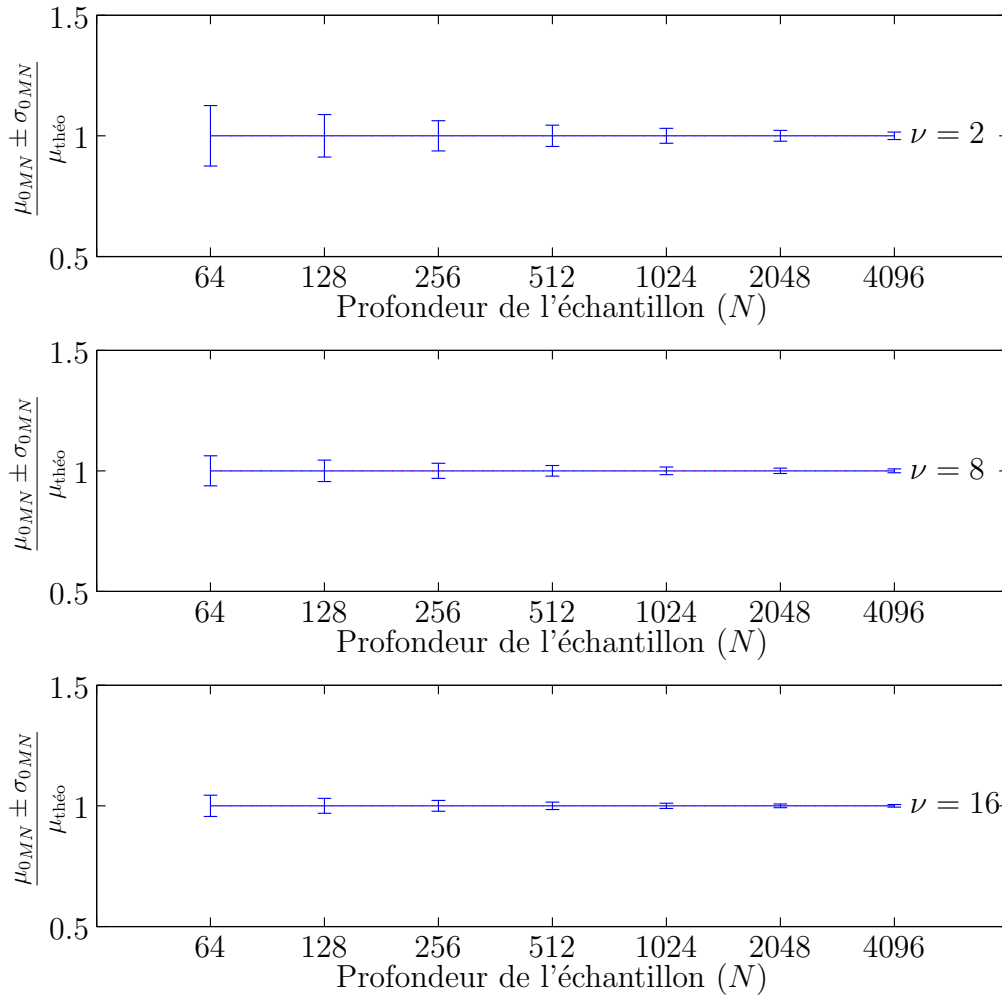


FIG. 2.7 – Moyenne  $\pm$  variance de l'estimateur de moyenne par moyenne arithmétique normalisée par la moyenne théorique de la distribution en  $\chi^2$  à  $\nu$  degrés de liberté.

$$\text{Var}(\hat{\mu}_{MN}) = \frac{1}{N^2} \sum_{k=1}^N \text{Var}(P_{Acc}[k]) \quad (2.55)$$

$$\text{Var}(\hat{\mu}_{MN}) = \sigma_{\text{théo}}^2 + \frac{1}{N} [\alpha \sigma_{RFI}^2 + \alpha(1 - \alpha) P_{RFI}^2] \quad (2.56)$$

L'équation 2.54 montre que l'estimateur de moyenne diverge de la valeur de  $\mu_{\text{théo}}$  dans la proportion de la présence de l'interférence.

La variance de l'estimateur donnée par 2.56 augmente avec la variance  $\sigma_{RFI}^2$  de l'interférence mais aussi avec sa valeur moyenne  $P_{RFI}$ . On peut noter le coefficient  $\alpha(1 - \alpha)$  qui maximise l'effet de la moyenne de l'interférence sur la variance de l'estimateur pour des valeurs intermédiaires de  $N_{RFI}$  (figure 2.8). Cela s'explique facilement par la transformation de  $f(x)$  suivant une loi mono-modale pour  $\alpha \approx 0\%$  ou  $\alpha \approx 100\%$ , vers une loi bi-modale pour  $\alpha \approx 50\%$ .

On cherche à évaluer la qualité de l'estimateur en fonction de l'amplitude moyenne

$P_{RFI}$  des interférences et de la proportion  $\alpha$  de points parasités sur l'échantillon observé. On présente pour chaque estimateur les plans  $\Delta_{\mu}^{log}(P_{RFI}, \alpha)$  et  $\Delta_{\sigma}^{log}(P_{RFI}, \alpha)$  résultant du calcul ou de la simulation des expressions 2.39 et 2.40. Comme les valeurs de ces plans peuvent évoluer dans une grande dynamique et qu'elles ont la dimension d'un rapport de puissance, on optera pour une représentation en dB.

La moyenne et la variance de l'estimateur sont calculés analytiquement pour obtenir la figure 2.8 avec les paramètres suivants :  $f_{RFI}(x) = \delta(x - P_{RFI})$ ,  $N = 4096$ ,  $\alpha = N_{RFI}/N = 0 \dots 100\%$  et  $\text{INR}_{dB} = 10 \log_{10}(P_{RFI}/\mu_{\text{théo}}) \in [-20, \dots, 70]$ .

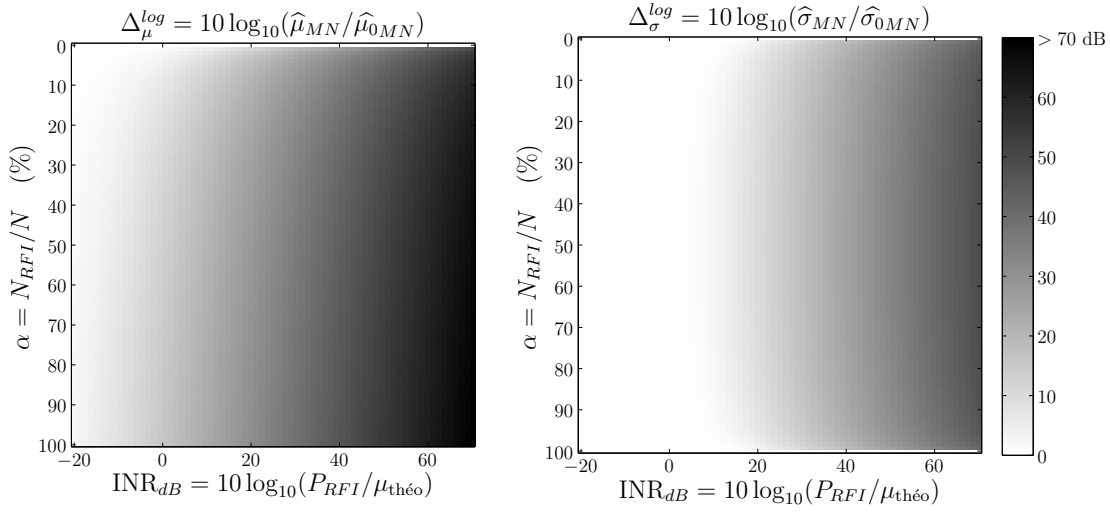


FIG. 2.8 – Moyenne et écart type de l'estimateur  $\hat{\mu}_{MN}$  en présence d'un parasite de taux de présence  $\alpha = N_{RFI}/N$  et de densité de probabilité  $\delta(x - P_{RFI})$ .

Par elle-seule, cette représentation pour l'estimateur de moyenne par moyenne arithmétique n'a pas grand intérêt. Elle servira surtout de référence lors de la comparaison de la robustesse des autres estimateurs. Dans l'idéal, un estimateur robuste ne devrait jamais dériver en présence d'interférences. La figure correspondante serait alors les courbes  $\Delta_{\mu}^{log} = 0 \text{ dB}$  et  $\Delta_{\sigma}^{log} = 0 \text{ dB}$ .

### 2.2.2.3 CONTRAINTES DE CALCUL

L'algorithme ne pose pas de problème particulier. La somme est effectuée dans un registre. La division finale est facilitée si le nombre de points moyennés est une puissance entière de 2. Il s'agit d'un simple décalage vers la droite.

### 2.2.3 Médiane

Afin de limiter l'influence des interférences qui dénaturent la moyenne à cause de la prise en compte de certains points aberrants de  $P_{Acc}[k]$ , on envisage un calcul de médiane pour estimer  $\mu_{théo}$  de la façon suivante sur un échantillon de taille  $N$  :

$$\hat{\mu}_{MD} = median_{k=1\dots N} \{P_{Acc}[k]\}. \quad (2.57)$$

L'algorithme consiste à classer l'échantillon par valeur puis à sélectionner le point central, ou la moyenne des deux points centraux dans le cas d'un échantillon de profondeur paire [55, p.194]. Le classement de l'échantillon implique donc d'étudier la statistique d'ordre [13].

$F_{X(r:N)}(x)$  est la fonction de répartition de la  $r^{\text{ième}}$  variable ordonnée  $X_{(r:N)}$  issue du  $N$ -échantillon ordonné ( $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(N)}$ ) lui-même issu du classement du  $N$ -échantillon ( $X_1, X_2, \dots, X_N$ ) de fonction de répartition  $F_X(x)$ .

Selon [20, 55], la fonction de répartition vaut :

$$F_{X(r:N)}(x) = \sum_{k=r}^N C_N^k F_X(x)^k [1 - F_X(x)]^{N-k} \quad (2.58)$$

$$F_{X(r:N)}(x) = I_{F_X(x)}(r, N - r + 1) \quad (2.59)$$

$$F_{X(r:N)}(x) = \frac{1}{B(r, N - r + 1)} \int_0^{F_X(x)} t^{r-1} (1-t)^{N-r} dt \quad (2.60)$$

avec  $I_x(a, b)$ : Fonction bêta incomplète

$B(a, b)$ : Fonction bêta

La densité de probabilité de la statistique d'ordre de rang  $r$  est obtenue par la dérivée de 2.60 :

$$f_{X(r:N)}(x) = \frac{1}{B(r, N - r + 1)} \frac{d}{dx} \left( \int_0^{F_X(x)} t^{r-1} (1-t)^{N-r} dt \right) \quad (2.61)$$

$$f_{X(r:N)}(x) = \frac{1}{B(r, N - r + 1)} F_X(x)^{r-1} (1 - F_X(x))^{N-r} f_X(x) \quad (2.62)$$

avec  $f_X(x)$ : Densité de probabilité de la loi parente de  $X$

$F_X(x)$ : Fonction de répartition de la loi parente de  $X$

#### 2.2.3.1 ESPÉRANCE ET VARIANCE DE L'ESTIMATEUR

La médiane est un cas particulier des statistiques d'ordre. Par convenance, nous choisissons  $r = N/2$  même en cas d'échantillons de profondeur impaire afin de simplifier simulations et implantations matérielles.

Le calcul du biais et de la variance de l'estimateur par médiane de la moyenne d'une



distribution de type  $\chi^2_{\nu,\sigma}$  nécessite d'établir les expressions de ses moments non-centrés :

$$E \left\{ \widehat{\mu}_{MD}^k \right\} = \int_{-\infty}^{+\infty} x^k f_{0(\frac{N}{2}:N)}(x) dx \quad (2.63)$$

$$E \left\{ \widehat{\mu}_{MD}^k \right\} = \frac{1}{B(\frac{N}{2}, \frac{N}{2} + 1)} \int_{-\infty}^{+\infty} x^k F_0(x)^{\frac{N}{2}-1} (1 - F_0(x))^{\frac{N}{2}} f_0(x) dx \quad (2.64)$$

avec  $f_0(x)$ : Densité de probabilité de la loi du  $\chi^2$  à  $\nu$  degrés de liberté.

$F_0(x)$ : Fonction de répartition de la loi du  $\chi^2$  à  $\nu$  degrés de liberté.

Le biais et la variance sont calculés en discrétisant l'intégrale 2.64. Les résultats, disponibles en figure 2.9, montrent que la variance de l'estimateur diminue quand la profondeur  $N$  de l'échantillon augmente. Par contre, un fort biais dépendant largement de  $\nu$  (et de façon moindre, de  $N$ , voir section 2.2.4, p.67) subsiste du fait de la dissymétrie de la loi et conduit le pourcentage de points rejetés à dépasser significativement la valeur de probabilité de fausse alarme  $\delta$  souhaitée.

La diminution du biais corrélée avec l'augmentation du degré de liberté est légitime car la distribution tend vers une distribution gaussienne qui est symétrique et égalise ainsi moyenne et médiane.

On étudiera en section 2.2.4 une version non-biaisée de cet estimateur.

### 2.2.3.2 DÉRIVE DE L'ESTIMATEUR

Nous avons vu que l'estimateur de moyenne par médiane présentait un biais que nous nous proposons de corriger plus tard. Son comportement imparfait en absence d'interférences pourrait laisser croire qu'il serait oisieux de poursuivre par l'étude de sa dérive face à des interférences. Or, pour des valeurs de  $\nu$  élevées, le biais devient suffisamment négligeable pour que la correction n'apporte qu'un sur-coût de calcul. Nous poursuivons donc plus en avant l'étude du comportement de l'estimateur qui est donc maintenant appliqué à un échantillon  $P_{Acc}[k]$  qui contient une portion de points parasités. L'expression des moments centrés de  $\widehat{\mu}_{MD}^k$  est basée sur celle de  $\widehat{\mu}_{MD}^k$  (équation 2.64) mais la densité de probabilité  $f_0(x)$  est remplacé par  $f(x)$  et la fonction de répartition  $F_0(x)$  par  $F(x)$  :

$$E \left\{ \widehat{\mu}_{MD}^k \right\} = \frac{1}{B(\frac{N}{2}, \frac{N}{2} + 1)} \int_{-\infty}^{+\infty} x^k F(x)^{\frac{N}{2}-1} (1 - F(x))^{\frac{N}{2}} f(x) dx \quad (2.65)$$

La moyenne et la variance de l'estimateur sont calculés numériquement pour obtenir la figure 2.10 avec les paramètres suivants :  $f_{RFI}(x) = \delta(x - P_{RFI})$ ,  $\nu = 16$ ,  $N = 4096$ ,  $\alpha = N_{RFI}/N = 0 \dots 100\%$  et  $\text{INR}_{dB} = 10 \log_{10}(P_{RFI}/\mu_{\text{théo}}) \in [-20, \dots, 70]$ .

Pour  $\alpha < 45\%$ , l'espérance de l'estimateur ne subit aucune des variations que pourrait créer le signal interférant. Par contre, pour  $\alpha > 50\%$ , elle dévie totalement pour ne prendre en compte que les valeurs parasitées. C'est un comportement normal de l'algorithme de médiane qui est couramment utilisé en traitement d'image pour retirer les points aberrants. Quand  $N_{RFI}$  dépasse  $N/2$ , les valeurs appartenant au signal intègre deviennent minoritaires, donc aberrants, et sont alors ignorées par l'estimateur.

La variance de l'estimateur présente un maximum pour des valeurs de  $N_{RFI}$  proches de  $N/2$ . En effet, la densité de probabilité présente dans ces cas-là une équi-répartition

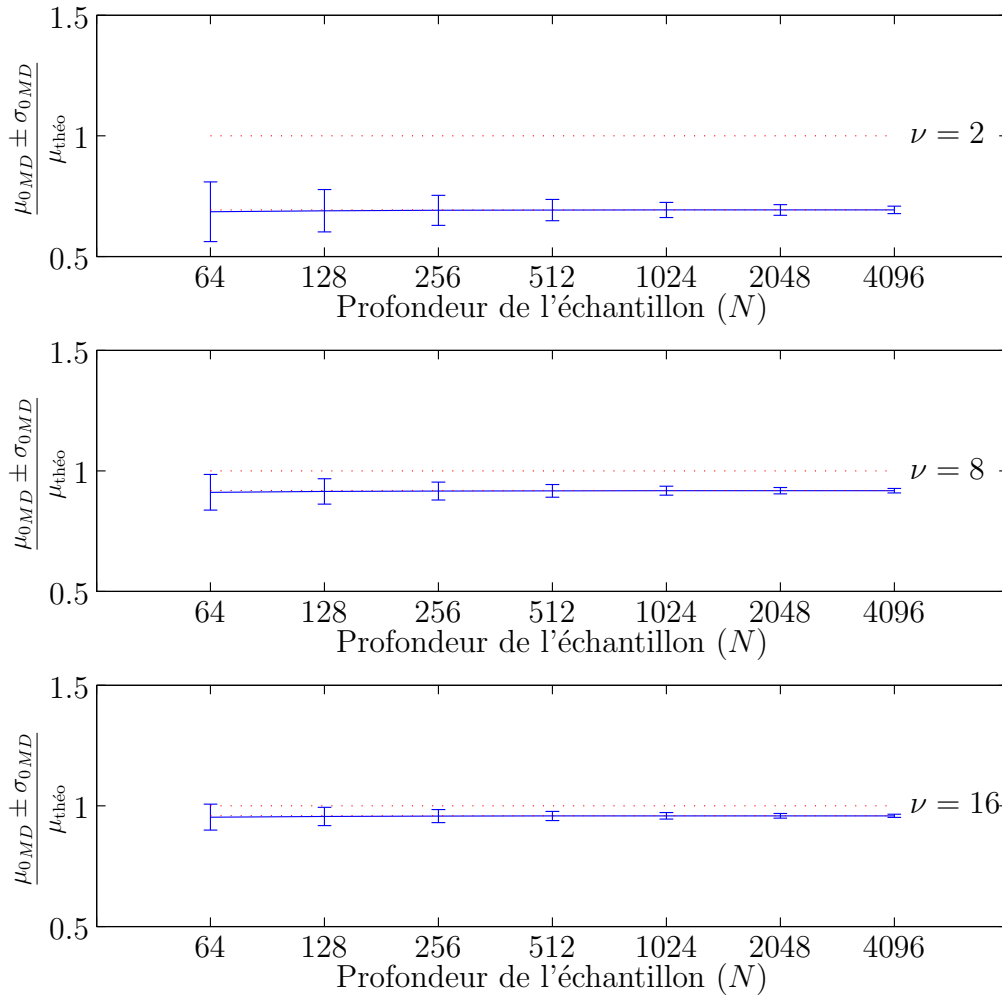


FIG. 2.9 – Moyenne  $\pm$  variance de l'estimateur de moyenne par médiane normalisée par la moyenne théorique de la distribution en  $\chi^2$  à  $\nu$  degrés de liberté.

entre la densité de probabilité du signal intègre  $f_0(x)$  et la densité de probabilité du signal intègre convoluée par la densité de probabilité du signal corrompu  $f_{RFI}(x) * f_0(x)$  (cas  $\alpha = 50\%$  de la figure 2.6, p.59). La bimodalité de la loi  $f(x)$  fait exploser la variance. En revanche, pour les valeurs faibles (respectivement fortes) de  $N_{RFI}$ ,  $f(x)$  tend vers  $f_0(x)$  (respectivement  $\delta(x - P_{RFI}) * f_0(x)$ ). La variance de l'estimateur appliqué sur la densité de probabilité  $f(x)$  approche alors la variance de l'estimateur appliqué sur la densité de probabilité  $f_0(x)$  car la variance est insensible à la moyenne de la loi  $f(x)$  qui devient monomodale dans ces deux cas extrêmes ( $\alpha = 0\%$  et  $\alpha = 100\%$ ).

### 2.2.3.3 CONTRAINTES DE CALCUL

L'estimateur de moyenne par médiane impose un classement du  $N$ -échantillon. Ce type de traitement est lourd à mettre en place dans un FPGA et, comme sur un DSP ou un processeur classique, présente une durée d'exécution dépendante des données à traiter, donc imprédictible, tout au mieux maximisable.

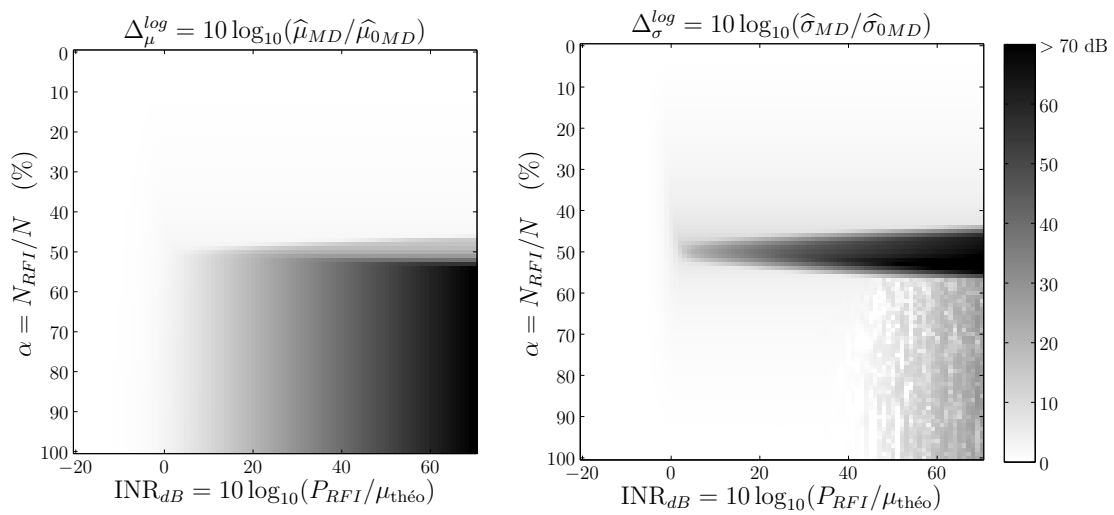


FIG. 2.10 – Moyenne et écart type de l'estimateur  $\hat{\mu}_{MD}$  ( $\nu = 16$  et  $N = 4096$ ) en présence d'un parasite de taux de présence  $\alpha = N_{RFI}/N$  suivant une loi de type Dirac de moyenne  $P_{RFI}$ . (La partie grisée pour  $INR > 50$  dB et  $\alpha > 50\%$  est le résultat de calculs erronés suite à une limitation de la précision machine.)

### 2.2.4 Médiane corrigée

Nous avons vu que l'estimateur de moyenne par médiane était biaisé, c'est-à-dire que  $b(\widehat{\mu}_{0MD}, \mu_{\text{théo}}) = E\{\widehat{\mu}_{0MD}\} - \mu_{\text{théo}} \neq 0$ . Nous allons calculer une *correction additive* de biais  $CA(\nu, N)$  pour  $b(\widehat{\mu}_{0MD}, \mu_{\text{théo}})$  afin de créer un estimateur  $\widehat{\mu}_{0MD\nu, N} = \widehat{\mu}_{0MD} - CA(\nu, N)$  qui serait non-biaisé.

$$CA(\nu, N) = E\{\widehat{\mu}_{0MD}\} - \mu_{\text{théo}} \quad (2.66)$$

$$CA(\nu, N) = \int_{-\infty}^{+\infty} x f_{0(N/2:N)}(x) dx - \nu\sigma^2 \quad (2.67)$$

$$CA(\nu, N) = \int_{-\infty}^{+\infty} (t\sigma) f_{0(N/2:N)}(t\sigma) \sigma dt - \nu\sigma^2 \quad (2.68)$$

$$\text{avec le changement de variable : } x = t\sigma \quad (2.69)$$

$$CA(\nu, N) = \sigma^2 \left[ \int_{-\infty}^{+\infty} t f_{\chi_{\nu}^2(N/2:N)}(t) dt - \nu \right] \quad (2.70)$$

$$(2.71)$$

Mais  $CA(\nu, N)$  est alors fonction de  $\sigma^2$  qui est exactement le paramètre que l'on cherche à évaluer par l'intermédiaire de  $\mu_{\text{théo}}$ . Il n'est pas possible de procéder directement de cette façon.

En revanche, on peut construire l'estimateur  $\widehat{\mu}_{0MD\nu, N} = K(\nu, N) \cdot \widehat{\mu}_{0MD}$  avec le *correcteur multiplicatif* de biais  $K(\nu, N)$  tel que :

$$K(\nu, N) = \frac{\mu_{\text{théo}}}{E\{\widehat{\mu}_{0MD}\}} \quad (2.72)$$

$$K(\nu, N) = \frac{\nu\sigma^2}{\sigma^2 \int_{-\infty}^{+\infty} t f_{\chi_{\nu}^2(N/2:N)}(t) dt} \quad (2.73)$$

$$K(\nu, N) = \frac{\nu}{\int_{-\infty}^{+\infty} t f_{\chi_{\nu}^2(N/2:N)}(t) dt} \quad \text{pour } \nu > 1 \quad (2.74)$$

Pour  $N$  grand ( $N > 128$ ), la correction à apporter est principalement dépendant de  $\nu$ . Dans ce cas, l'estimateur  $\widehat{\mu}_{0MD}$  tend vers la médiane théorique  $\mu_{0MD}$  appliquée au cas continu :

$$\lim_{N \rightarrow \infty} (K(\nu, N)) = K(\nu) \quad \text{pour } \nu \text{ constant.} \quad (2.75)$$

On peut ainsi calculer explicitement  $K(\nu)$  tel que :

$$K(\nu) = \frac{\mu_{\text{théo}}}{\mu_{0MD}} \quad (2.76)$$

$$K(\nu) = \frac{\nu\sigma^2}{2\sigma^2 P_{\Gamma}^{-1}\left(\frac{\nu}{2}, \frac{1}{2}\right)} \quad (2.77)$$

$$K(\nu) = \frac{\nu}{2P_{\Gamma}^{-1}\left(\frac{\nu}{2}, \frac{1}{2}\right)} \quad \text{pour } \nu > 1 \quad (2.78)$$

$$(2.79)$$

$K(\nu)$  est bien un sous-cas de  $K(\nu, N)$  mais, si son utilisation est possible ( $N > 128$ ), l'implémentation matérielle de la version de cet estimateur permet réduire l'empreinte mémoire du stockage des coefficients  $K(\nu, N)$  d'un facteur  $N$ .  $K(\nu)$  est tabulé en table 2.1 et visible en figure 2.11. Pour  $\nu$  grand, la distribution  $\chi_{\nu, \sigma}^2$  tend vers une distribution symétrique de type gaussienne, réduisant ainsi l'utilité d'une correction sur l'estimation de la moyenne. L'utilisation d'un coefficient de correction est donc de moins en moins pertinent : il tend vers 1.

TAB. 2.1 – Tabulation de  $K(\nu)$  pour  $\nu = 2^{1 \dots 10}$

$\nu$	$K(\nu)$
2	1.44371155382361
4	1.19239849334379
8	1.08984648434222
16	1.04351689349685
32	1.02147620388388
64	1.01070966553792
128	1.00537674290324
256	1.00271430399948
512	1.00130347833915
1024	1.000651390294

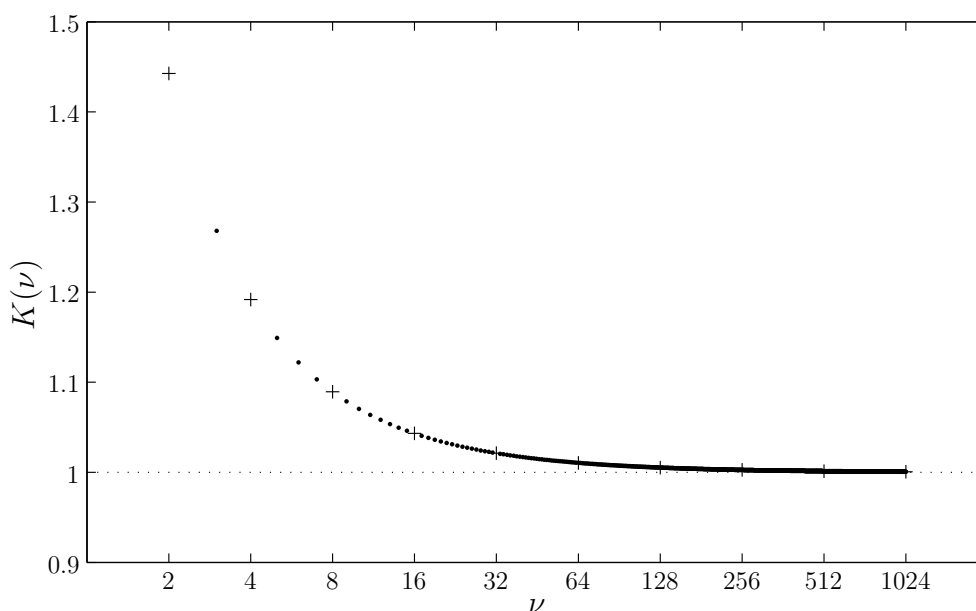


FIG. 2.11 – Fonction correctrice  $K(\nu)$  affichée pour  $\nu = 2^{1 \dots 10}$ .  $K(\nu)$  tend rapidement vers 1 quand  $\nu$  augmente.

$K(\nu, N)$  est affiché en figure 2.12. Comme pour  $K(\nu)$ , le coefficient correcteur tend vers 1 pour de grandes valeurs de  $\nu$ . Les coupes  $K(\nu, N)$  pour  $N \geq 128$  sont très proches de la fonction  $K(\nu)$  de la figure 2.11.

La figure 2.13 présente l'erreur relative de  $K(\nu)$  par rapport à  $K(\nu, N)$ . Pour  $\nu$  ou  $N$  grand, l'erreur devient négligeable.

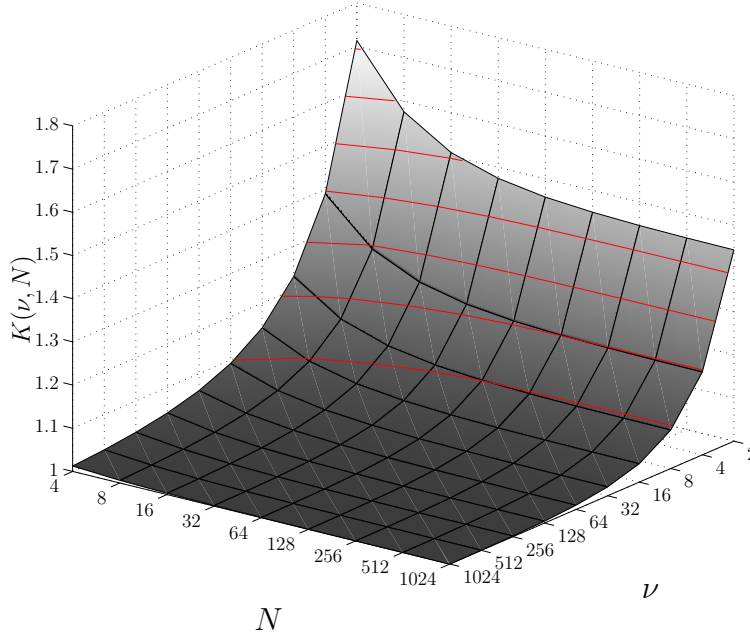


FIG. 2.12 – Fonction correctrice  $K(\nu, N)$  affichée pour  $\nu = 2^{1 \dots 10}$  et  $N = 2^{2 \dots 10}$ .

#### 2.2.4.1 ESPÉRANCE ET VARIANCE DE LA MÉDIANE CORRIGÉE PAR $K(\nu, N)$

L'estimation de la moyenne à l'aide de la médiane corrigée par  $K(\nu, N)$  peut s'effectuer comme :

$$\hat{\mu}_{0MD_{\nu,N}} = K(\nu, N) \cdot \text{median}_{k=1 \dots N} \{P_{Acc0}[k]\}. \quad (2.80)$$

L'expression des moments non-centrés de  $\hat{\mu}_{0MD_{\nu,N}}$  est donnée par :

$$E \left\{ \hat{\mu}_{0MD_{\nu,N}}^k \right\} = \int_{-\infty}^{+\infty} (xK(\nu, N))^k f_{0(N/2:N)}(x) \, dx \quad (2.81)$$

$$E \left\{ \hat{\mu}_{0MD_{\nu,N}}^k \right\} = \frac{1}{B\left(\frac{N}{2}, \frac{N}{2} + 1\right)} \int_{-\infty}^{+\infty} (xK(\nu, N))^k F_0(x)^{\frac{N}{2}-1} (1 - F_0(x))^{\frac{N}{2}} F_0(x) \, dx \quad (2.82)$$

La figure 2.14 illustre les modifications apportées par l'intégration du facteur de correction  $K(\nu, N)$ . L'estimateur  $\hat{\mu}_{0MD_{\nu,N}}$  n'est pas biaisé, quelles que soient les valeurs de  $\nu$  et  $N$ . On note néanmoins, comme pour  $\hat{\mu}_{0MD_{\nu}}$  et  $\hat{\mu}_{0MD}$ , que l'estimateur  $\hat{\mu}_{0MD_{\nu,N}}$  présente une variance plus conséquente que la variance de  $\hat{\mu}_{0MN}$  pour des  $\nu$  faibles.

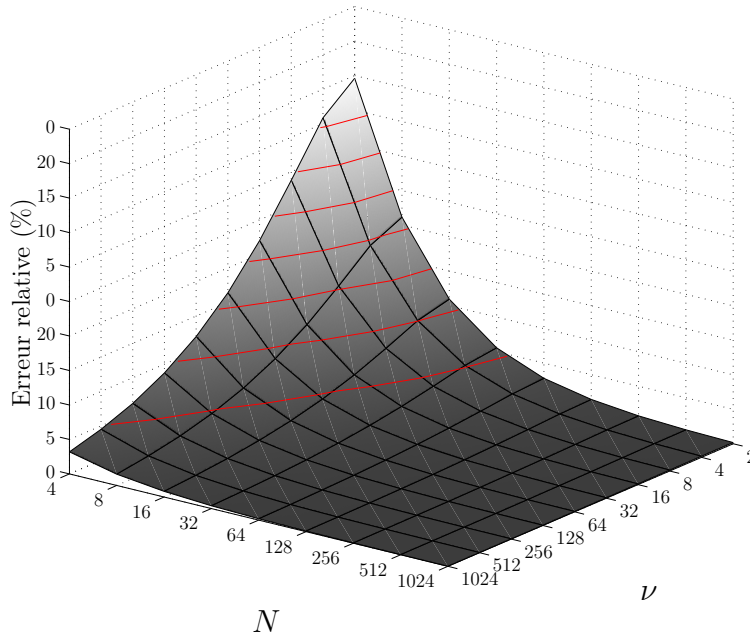


FIG. 2.13 – Erreur relative de  $K(\nu)$  par rapport à  $K(\nu, N)$  affichée pour  $\nu = 2^{1 \dots 10}$  et  $N = 2^{2 \dots 10}$ .

### 2.2.4.2 ESPÉRANCE ET VARIANCE DE LA MÉDIANE CORRIGÉE PAR $K(\nu)$

L'estimation de la moyenne par médiane corrigée par  $K(\nu)$  peut s'effectuer comme :

$$\hat{\mu}_{0MD\nu} = K(\nu) \cdot \text{median}_{k=1 \dots N} \{P_{Acc0}[k]\}. \quad (2.83)$$

L'expression des moments non-centrés de  $\hat{\mu}_{0MD\nu}$  est donnée par :

$$E \left\{ \hat{\mu}_{0MD\nu}^k \right\} = \int_{-\infty}^{+\infty} (xK(\nu))^k f_{0(N/2:N)}(x) dx \quad (2.84)$$

$$E \left\{ \hat{\mu}_{0MD\nu}^k \right\} = \frac{1}{B\left(\frac{N}{2}, \frac{N}{2} + 1\right)} \int_{-\infty}^{+\infty} (xK(\nu))^k F_0(x)^{\frac{N}{2}-1} (1 - F_0(x))^{\frac{N}{2}} f_0(x) dx \quad (2.85)$$

La figure 2.15 illustre les modifications apportées par l'intégration du facteur de correction  $K(\nu)$ . Le biais diminue fortement et ne semble plus fonction de  $\nu$ . Toutefois, pour les faibles valeurs de  $N$ , l'estimateur est encore légèrement biaisé. On peut noter une augmentation de la variance pour les faibles valeurs de  $\nu$ .

### 2.2.4.3 DÉRIVE DE L'ESTIMATEUR

La loi  $f_0(x)$  (respectivement  $F_0(x)$ ) est remplacée par la loi de mélange associée  $f(x)$  (respectivement  $F(x)$ ) dans les équations 2.84 et 2.85.

La moyenne et la variance de l'estimateur  $\hat{\mu}_{MD\nu,N}$  sont calculées numériquement pour obtenir la figure 2.16 avec les paramètres suivants :  $f_{RFI}(x) = \delta(x - P_{RFI})$ ,  $\nu = 16$ ,  $N = 4096$ ,  $\alpha = \frac{N_{RFI}}{N} = 0 \dots 100\%$  et  $INR_{dB} = 10 \log_{10}(P_{RFI}/\mu_{théo}) \in [-20, \dots, 70]$ .

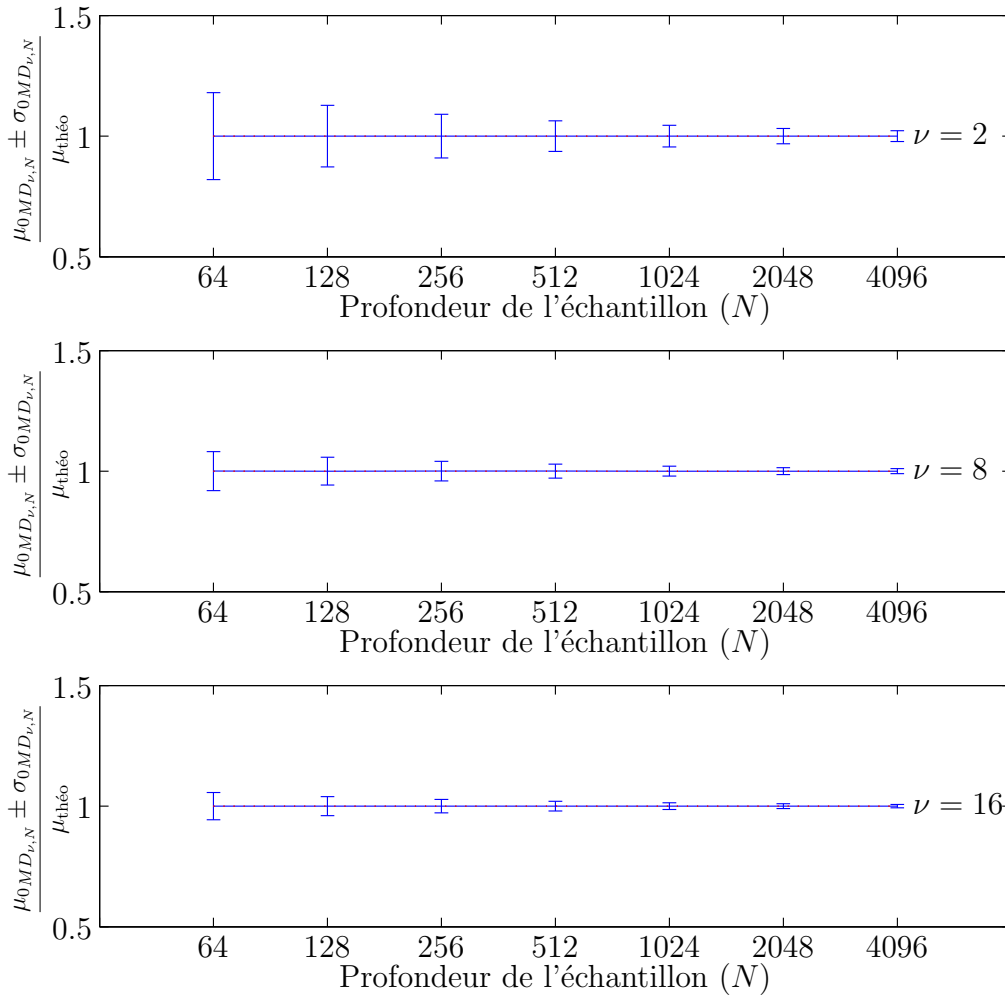


FIG. 2.14 – Moyenne  $\pm$  variance de l'estimateur de moyenne par médiane avec correction de biais  $K(\nu, N)$ , normalisée par la moyenne théorique de la distribution en  $\chi^2$  à  $\nu$  degrés de liberté.

Comme pour  $\hat{\mu}_{MD\nu}$  (non représenté ici), la dérive de l'estimateur est très proche de celle de  $\hat{\mu}_{MD}$ . Pour  $\alpha < 45\%$ , la dérive est négligeable mais elle devient totale en faveur des interférences pour  $\alpha > 50\%$ .

#### 2.2.4.4 CONTRAINTES DE CALCUL

La contrainte du classement est identique à celle de l'estimateur par médiane. La prise en compte du coefficient correcteur  $K(\nu, N)$  ou  $K(\nu)$  est possible lors du calcul du seuil avec 2.23 :

$$S = 2 \frac{K(\nu, N) \cdot \text{median}_{k=1 \dots N} \{P_{Acco}[k]\}}{\nu} P_{\Gamma}^{-1}(\nu/2, 1 - \alpha) \quad (2.86)$$

Dans le cas d'une application configurable, les coefficients correcteurs doivent donc



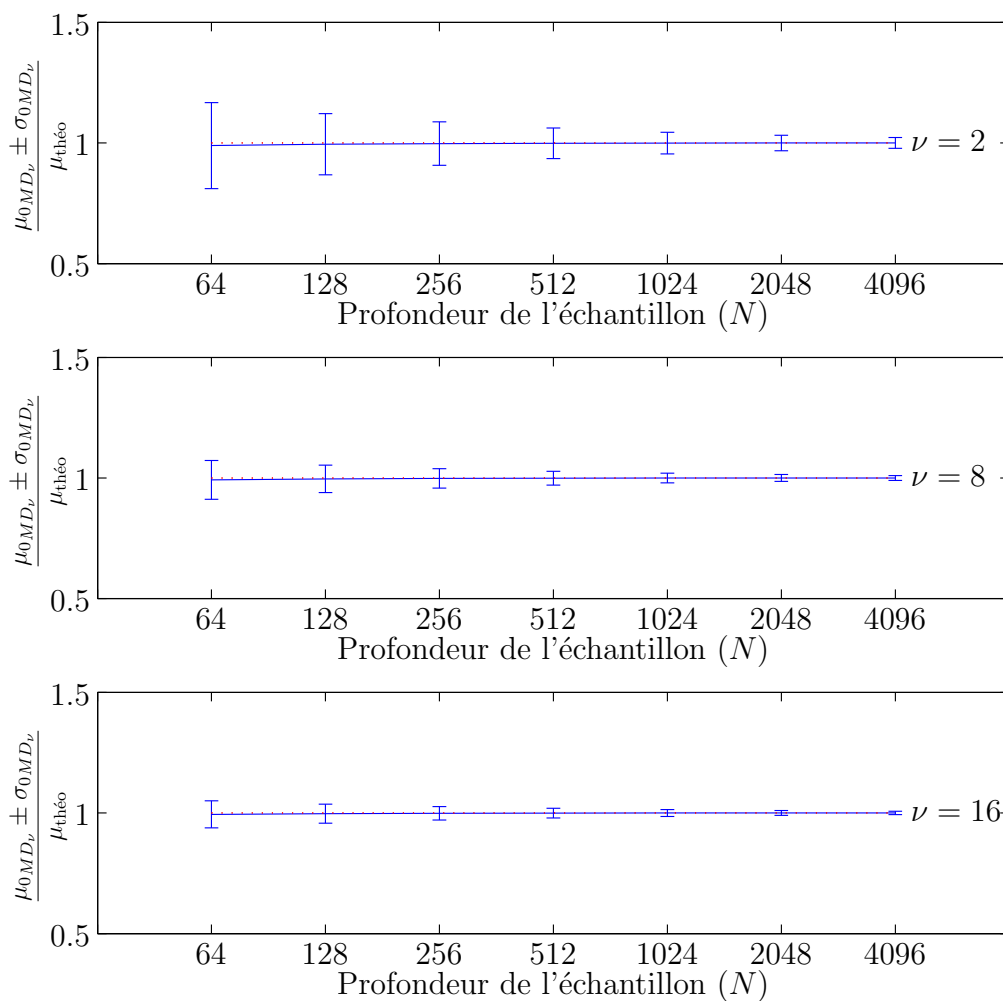


FIG. 2.15 – Moyenne  $\pm$  variance de l'estimateur de moyenne par médiane avec correction de biais  $K(\nu)$ , normalisée par la moyenne théorique de la distribution en  $\chi^2$  à  $\nu$  degrés de liberté.

être stockés dans une LUT. Il est possible de réduire l'empreinte mémoire de cette table en ne conservant que  $K(\nu)$  car le biais est principalement sensible à  $\nu$ , directement lié au nombre d'accumulations  $Acc$  effectuées lors de l'estimation de la densité spectrale de puissance.

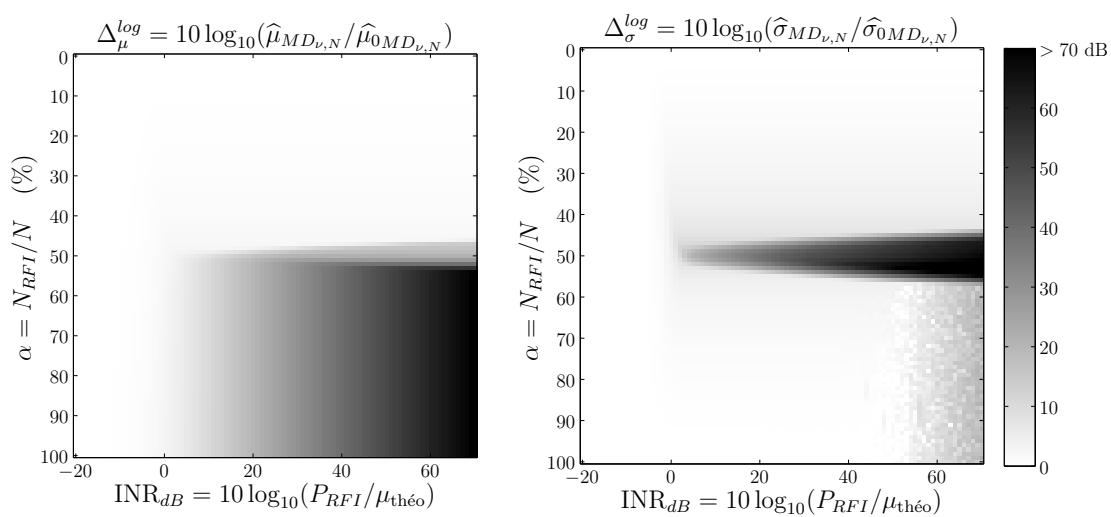


FIG. 2.16 – Moyenne et écart type de l'estimateur  $\hat{\mu}_{MD\nu,N}$  ( $\nu = 16$  et  $N = 4096$ ) en présence d'un parasite de taux de présence  $\alpha = N_{RFI}/N$  suivant une loi de type Dirac de moyenne  $P_{RFI}$ . (La partie grisée pour  $INR > 50$  dB et  $\alpha > 50\%$  est le résultat de calculs erronés suite à une limitation de la précision machine.)

### 2.2.5 Moyenne tronquée

Nous avons vu dans les sections précédentes des estimateurs présentant chacun divers avantages. L'estimateur de moyenne par moyenne tronquée à  $\delta\%$ , noté  $\mu_{MT_\delta}$ , cherche à associer à la robustesse de la médiane face aux interférences une faible variance d'estimation.

Considérons  $f_\delta(x)$ , la densité de probabilité du  $[(1 - 2\delta)N]$ -échantillon  $P_{Acco,\delta}[k]$  obtenue par la troncature des  $\lfloor N\delta \rfloor^2$  plus fortes et des  $\lfloor N\delta \rfloor$  plus faibles valeurs de  $P_{Acco}[k]$  (figure 2.17), telle que :

$$f_\delta(x) = \begin{cases} K \cdot f_0(x) & a_\delta \leq x \leq b_\delta \\ 0 & \text{ailleurs} \end{cases} \quad (2.87)$$

avec  $K = \frac{1}{F_0(b_\delta) - F_0(a_\delta)}$

Les paramètres  $a$  et  $b$  correspondent au seuil bas et au seuil haut de troncature de la densité de probabilité de la distribution de loi du  $\chi_\nu^2$  :

$$a_\delta = 2\sigma^2 P_\Gamma^{-1}\left(\frac{\nu}{2}, \delta\right) \quad (2.88)$$

$$b_\delta = 2\sigma^2 P_\Gamma^{-1}\left(\frac{\nu}{2}, 1 - \delta\right) \quad (2.89)$$

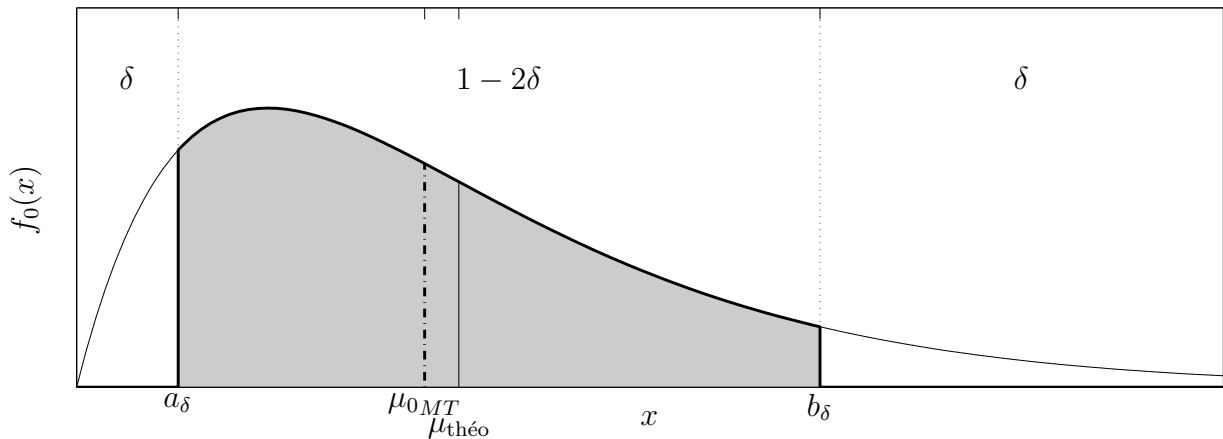


FIG. 2.17 – Densité de probabilité pour le calcul de la moyenne tronquée de  $\mu_{0MT_\delta}$ . On observe un biais dû à la dissymétrie de la distribution.

L'estimateur théorique de moyenne par moyenne tronquée s'écrit :

<sup>2</sup>L'arrondi de  $x$  à l'entier le plus proche est noté  $\lfloor x \rfloor$

$$\mu_{0MT_\delta} = E \left\{ P_{Acc0,\delta}[k] \right\} \quad (2.90)$$

$$= K \sigma^2 \int_{a_\delta}^{b_\delta} x f_{\chi_{\nu,\sigma}^2}(x) dx \quad (2.91)$$

$$= K \nu \sigma^2 \int_{a_\delta}^{b_\delta} f_{\chi_{\nu+2,\sigma}^2}(x) dx \quad (2.92)$$

$$\mu_{0MT_\delta} = \frac{F_{\chi_{\nu+2,\sigma}^2}(b_\delta) - F_{\chi_{\nu+2,\sigma}^2}(a_\delta)}{F_{\chi_{\nu,\sigma}^2}(b_\delta) - F_{\chi_{\nu,\sigma}^2}(a_\delta)} \mu_{\text{théo}} \quad (2.93)$$

On peut noter que  $\mu_{0MT_0} = \mu_{0MN}$ . De même,  $\mu_{0MT_\delta} \xrightarrow{\delta \rightarrow 0.5} \mu_{0MD}$ . L'estimateur  $\mu_{0MT_\delta}$  est bien un croisement des estimateurs  $\mu_{0MN}$  et  $\mu_{0MD}$ .

La procédure d'estimation pratique consiste à classer le  $N$ -échantillon  $P_{Acc0}[k]$ , retirer les  $[\delta N]$  échantillons de plus faible valeur et les  $[\delta N]$  échantillons de plus forte valeur du  $N$ -échantillon ordonné précédemment obtenu puis effectuer la moyenne des échantillons restants[50, p.31] :

$$\hat{\mu}_{0MT_\delta} = \frac{1}{N - 2q} \sum_{r=q+1}^{N-q} P_{Acc0(r:N)}, \quad \text{avec } q = [\delta N] \quad (2.94)$$

Pour être rigoureux, il faudrait donc étudier les statistiques d'une variable aléatoire composée de la somme de  $(1 - 2\delta)N$  variables aléatoires dont le comportement suit celui d'une variable aléatoire ordonnée  $X_{(r:N)}$ , avec  $r \in [1 + [\delta N], \dots, N - [\delta N]]$  et dont une description est effectuée en début de section 2.2.3. La figure 2.18 montre l'estimation par histogrammes moyennés des densités de probabilité  $f_{\delta,N}(x)$  pour diverses valeurs de  $N$ . Néanmoins, pour faciliter les calculs et les simulations, il peut être plus commode de considérer comme densité de probabilité d'usage celle du cas continu  $f_{\delta,\infty}(x) = f_\delta(x)$ .

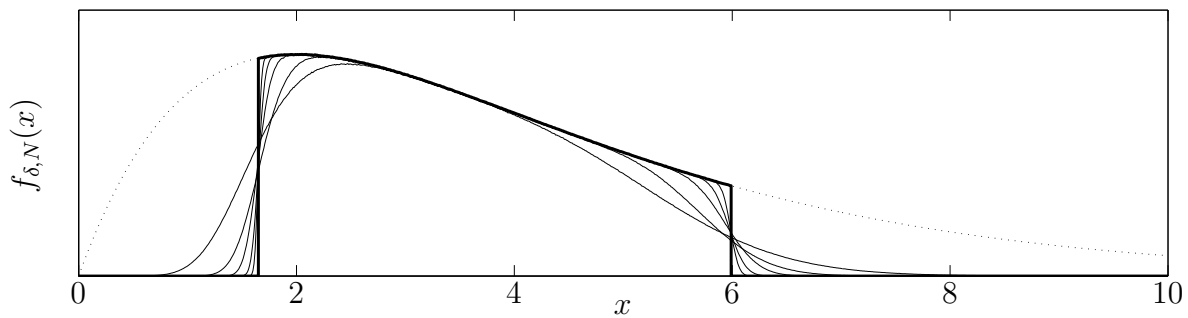


FIG. 2.18 – Densités de probabilité  $f_{\delta,N}(x)$  pour  $N=32, 128, 512, 2048, 8192$  tendant vers  $f_{\delta,\infty}(x)$  (en gras) basées sur la troncature à 20% de  $f_0(x)$  avec  $\nu = 4$  et  $\sigma = 1$  (en pointillé). Les aires des différentes densités de probabilité ne sont pas normalisées pour faciliter la comparaison.

### 2.2.5.1 ESPÉRANCE ET VARIANCE DE L'ESTIMATEUR

L'espérance et la variance de l'estimateur de moyenne par moyenne tronquée sont données par les équations 2.96 et 2.98.

$$E \{ \widehat{\mu}_{0MT_\delta} \} = \frac{1}{N-2q} \sum_{r=q+1}^{N-q} E \{ P_{Acco(r:N)} \} \quad (2.95)$$

$$= \frac{1}{N-2q} \sum_{r=q+1}^{N-q} \int_{\mathbb{R}} x f_{X(r:N)}(x) \, dx \quad (2.96)$$

$$Var \{ \widehat{\mu}_{0MT_\delta} \} = \left( \frac{1}{N-2q} \right)^2 \sum_{r=q+1}^{N-q} Var \{ P_{Acco(r:N)} \} \quad (2.97)$$

$$= \left( \frac{1}{N-2q} \right)^2 \sum_{r=q+1}^{N-q} \left[ \int_{\mathbb{R}} x^2 f_{X(r:N)}(x) \, dx - \left[ \int_{\mathbb{R}} x f_{X(r:N)}(x) \, dx \right]^2 \right] \quad (2.98)$$

$$\text{avec } f_{X(r:N)}(x) = \frac{1}{B(r, N-r+1)} F_0(x)^{r-1} (1-F_0(x))^{N-r} f_0(x)$$

Les deux paramètres sont estimés par intégration numérique des relations 2.96 et 2.98 pour  $\delta = [10\%, 20\%, 45\%]$  en figure 2.19. La variance est considérablement plus faible que pour les estimateurs précédents mais il faut noter un fort biais  $E \{ \widehat{\mu}_{0MT_\delta} \} - \mu_{\text{théo}}$  dépendant de  $\nu$  et de  $\delta$ .

On peut approcher l'espérance et la variance en remplaçant les densités de probabilité de la statistique d'ordre  $f_{X(r:N)}(x)$  par  $f_\delta(x)$  :

$$E \{ \widehat{\mu}_{0MT_\delta} \} \cong \frac{1}{N-2q} \sum_{r=q+1}^{N-q} E \{ P_{Acco,\delta}[k] \} \quad (2.99)$$

$$\cong \int_{\mathbb{R}} x f_\delta(x) \, dx \quad (2.100)$$

$$\cong \frac{F_{\chi_{\nu+2,\sigma}^2}(b_\delta) - F_{\chi_{\nu+2,\sigma}^2}(a_\delta)}{F_{\chi_{\nu,\sigma}^2}(b_\delta) - F_{\chi_{\nu,\sigma}^2}(a_\delta)} \mu_{\text{théo}} \quad (2.101)$$

$$Var \{ \widehat{\mu}_{0MT_\delta} \} \cong \left( \frac{1}{N-2q} \right)^2 \sum_{r=q+1}^{N-q} Var \{ P_{Acco,\delta}[k] \} \quad (2.102)$$

$$\cong \frac{1}{N-2q} \left[ \int_{\mathbb{R}} x^2 f_\delta(x) \, dx - \left( \int_{\mathbb{R}} x f_\delta(x) \, dx \right)^2 \right] \quad (2.103)$$

$$\cong \frac{1}{N-2q} \left( \nu(\nu+2)\sigma^4 \frac{F_{\chi_{\nu+4,\sigma}^2}(b_\delta) - F_{\chi_{\nu+4,\sigma}^2}(a_\delta)}{F_{\chi_{\nu,\sigma}^2}(b_\delta) - F_{\chi_{\nu,\sigma}^2}(a_\delta)} - \mu_{\text{théo}}^2 \right) \quad (2.104)$$

On conclut que le biais  $E \{ \widehat{\mu}_{0MT_\delta} \} - \mu_{0MT_\delta}$  est approximativement nul lors de l'estimation de  $\widehat{\mu}_{0MT_\delta}$  avec la procédure donnée en 2.94. Néanmoins, un biais  $E \{ \widehat{\mu}_{0MT_\delta} \} - \mu_{\text{théo}}$ ,

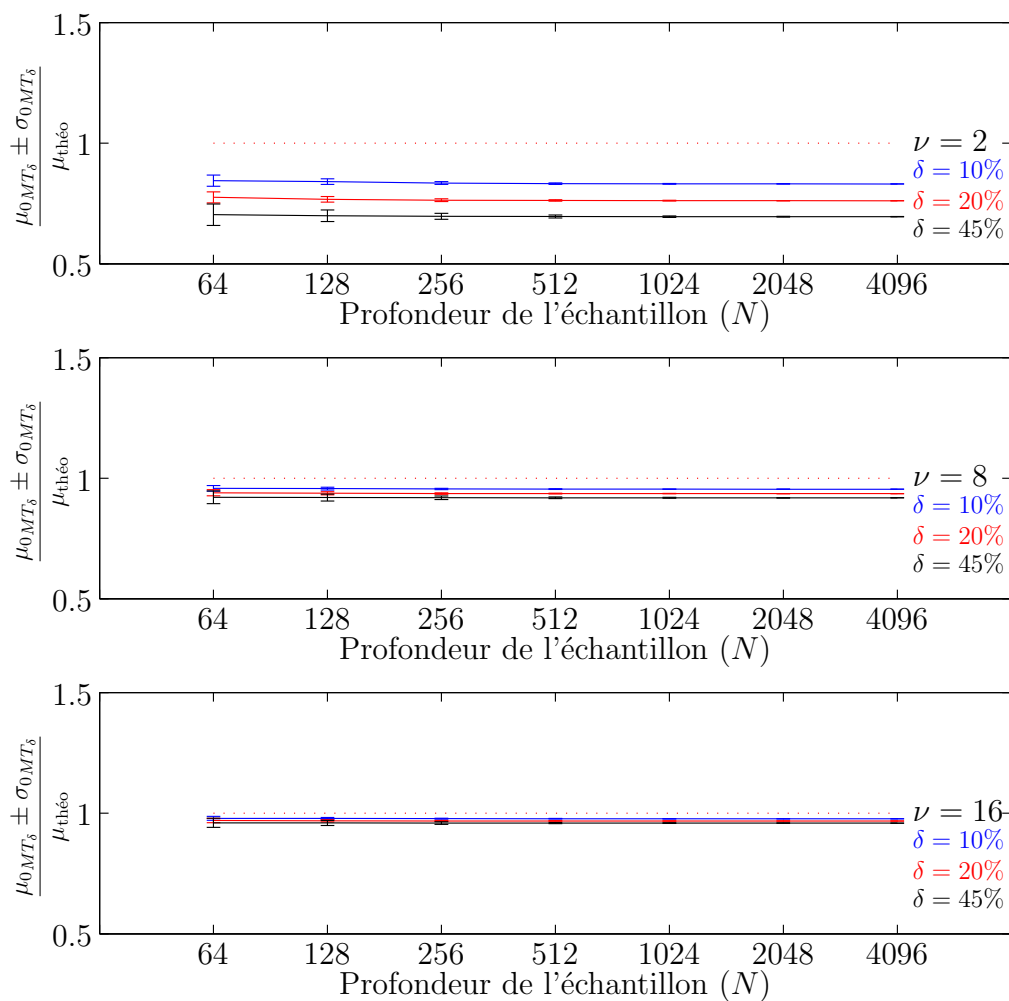


FIG. 2.19 – Moyenne  $\pm$  variance de l'estimateur de moyenne par moyenne tronquée normalisée par la moyenne théorique de la distribution en  $\chi^2$  à  $\nu$  degrés de liberté.

issu de la construction de  $\mu_{0MT_\delta}$ , implique une divergence de l'estimateur de la valeur  $\mu_{théo}$  par un facteur  $L(\nu, \delta)$  :

$$L(\nu, \delta) = \frac{\mu_{0MT_\delta}}{\mu_{théo}} \quad (2.105)$$

$$= \frac{F_{\chi_{\nu+2,\sigma}^2}(b_\delta) - F_{\chi_{\nu+2,\sigma}^2}(a_\delta)}{F_{\chi_{\nu,\sigma}^2}(b_\delta) - F_{\chi_{\nu,\sigma}^2}(a_\delta)} \quad (2.106)$$

Il est affiché en figure 2.20 pour  $\nu = [2, \dots, 32]$  et  $\delta = [0\%, \dots, 50\%[$  et montre que  $\mu_{MT_\delta}$  sous-estime  $\mu_{théo}$  d'autant plus que  $\nu$  est faible et  $\delta$  fort, confirmant ainsi les courbes de la figure 2.19.

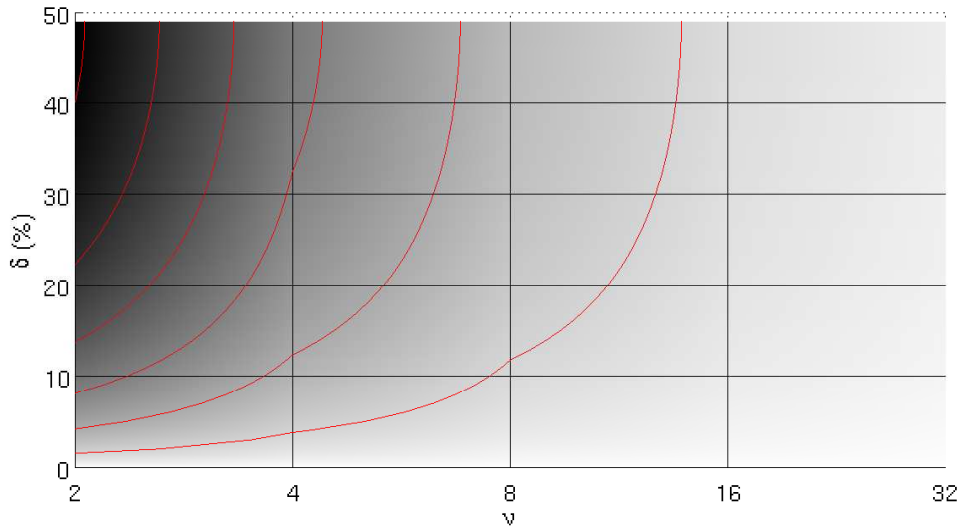


FIG. 2.20 –  $L(\nu, \delta)$  : rapport de sous-estimation de  $\mu_{MT_\delta}$  par rapport à  $\mu_{MN}$

### 2.2.5.2 DÉRIVE DE L'ESTIMATEUR

L'étude de l'effet des interférences sur l'estimateur n'est pas conduite de façon analytique. Cela nécessiterait l'évaluation de  $a$  et  $b$  sur la densité de probabilité donnée en 2.41 puis à l'évaluation des moments non-centrés par intégration numérique.

La figure 2.21 est donc issue d'une simulation numérique où des jeux de  $N$ -échantillons suivant une loi du  $\chi^2$  sont générés puis parasités à divers taux  $\alpha$  par des interférences d'intensité  $P_{RFI}$ . L'estimation pour  $\delta = [10\%, 20\%, 49\%]$  est effectuée avec les paramètres suivants :  $f_{RFI}(x) = \delta(x - P_{RFI})$ ,  $\nu = 16$ ,  $N = 4096$ ,  $\alpha = \frac{N_{RFI}}{N} = 0 \dots 100\%$  et  $INR_{dB} = 10 \log_{10}(P_{RFI}/\mu_{théo}) \in [-20, \dots, 70]$ . La moyenne et l'écart type sont estimés puis comparés à aux valeurs issues de données intègres.

Tant que le taux d'occupation parasite  $\alpha$  ne dépasse pas le pourcentage de troncature  $\delta$ , l'espérance et l'écart type de l'estimateur restent invariants. Passé ce stade, l'estimateur bascule vers la valeur des interférences. Pour  $\alpha > 1 - \delta$ , on note pour l'écart type un retour à une valeur normale qui suit la même explication que pour la médiane (section 2.2.3.2, p.64).

### 2.2.5.3 CONTRAINTES DE CALCUL

La contrainte du classement est proche de celle de l'estimateur par médiane. En toute rigueur, il suffit de retirer les  $\lfloor 2\delta N \rfloor$  plus fortes valeurs et les  $\lfloor 2\delta N \rfloor$  plus faibles valeurs du  $N$ -échantillon sans que le classement du reste des données ne soit nécessaire. Quelque soit le choix effectué, il faut ajouter au coût de calcul la somme des  $\lfloor (1 - 2\delta)N \rfloor$  points centraux de l'échantillon ordonné.

La division de l'équation 2.94 peut ne pas être calculable par un simple décalage vers la droite car  $\lfloor (1 - 2\delta)N \rfloor$  n'est pas systématiquement une puissance entière de 2. On peut néanmoins inclure cette division dans le calcul du seuil.

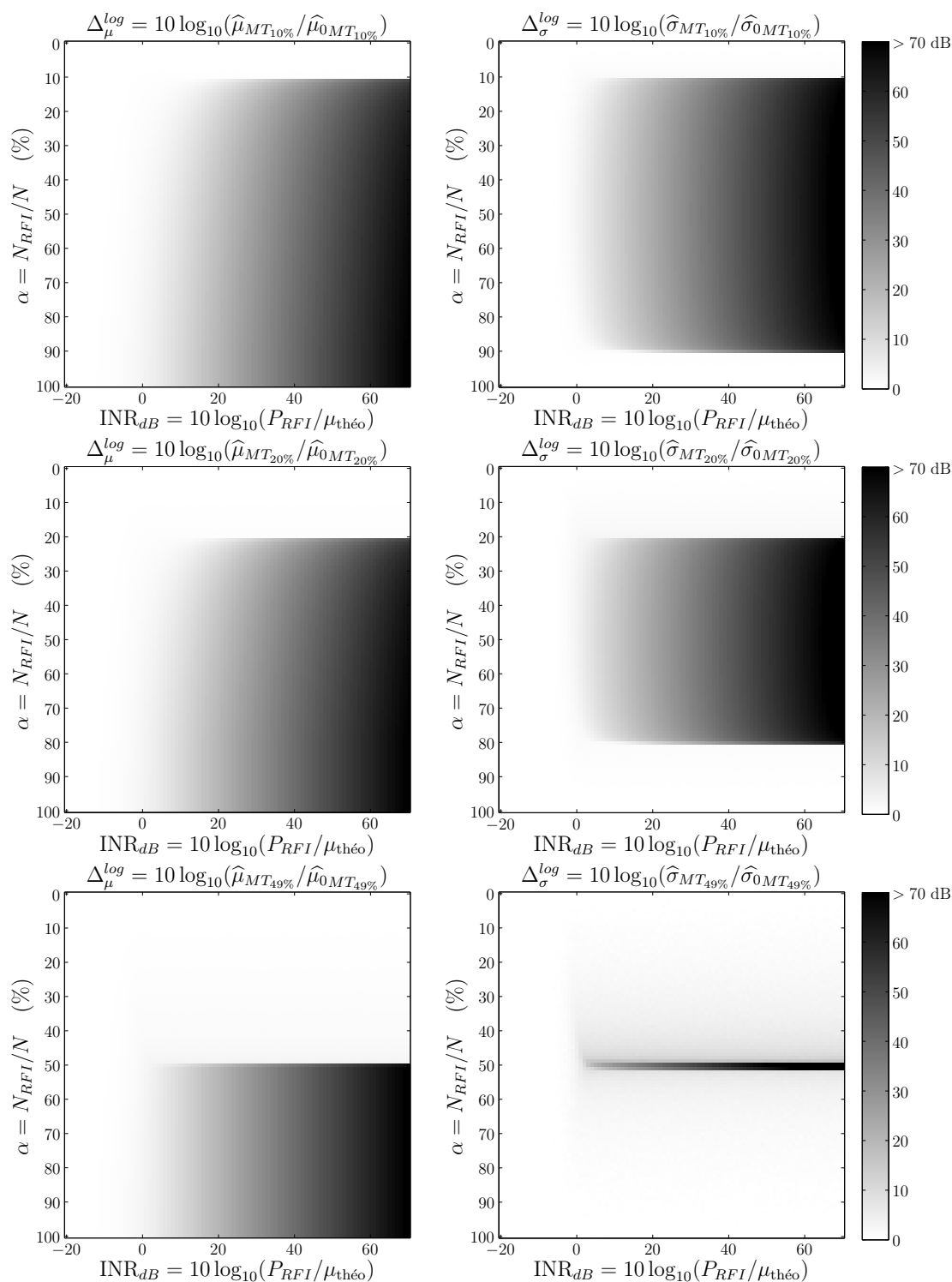


FIG. 2.21 – Moyenne et écart type de l'estimateur  $\hat{\mu}_{MT_{\delta}}$  ( $\nu = 16$  et  $N = 4096$ ) en présence d'un parasite de taux de présence  $\alpha = N_{RFI}/N$  suivant une loi de type Dirac de moyenne  $P_{RFI}$ . (haut :  $\delta = 10\%$ , milieu :  $\delta = 20\%$ , bas :  $\delta = 49\%$ )



### 2.2.6 Moyenne tronquée symétrique

La moyenne tronquée consiste à supprimer les  $[N\delta]$  plus petites valeurs et les  $[N\delta]$  plus grandes valeurs du  $N$ -échantillon. Les points restants sont moyennés (figure 2.17). L'inconvénient de cet estimateur est que la dissymétrie de la densité de probabilité induit un biais dans le calcul de la moyenne. On conçoit donc un nouvel estimateur nommé "moyenne tronquée symétrique" qui est de biais nul par construction. Pour ce faire, les  $[N\delta']$  plus petites valeurs et les  $[N\delta'']$  plus grandes valeurs du  $N$ -échantillon sont supprimés (avec  $\delta' + \delta'' = 2\delta$ , figure 2.22). Le choix correct de  $\delta'$  et  $\delta''$  permet d'obtenir un estimateur équivalent à la moyenne tronquée à  $\delta\%$  mais sans le défaut de biais.

La densité de probabilité du nouvel estimateur est celle d'un  $\chi^2$  à  $\nu$  degrés de liberté tronquée de part et d'autre des aires  $\delta'$  et  $\delta''$  :

$$f_{MTS}(x) = \begin{cases} K \cdot f_0(x), & x \in [a_{\delta'}, b_{\delta''}] \\ 0, & x \notin [a_{\delta'}, b_{\delta''}] \end{cases} \quad (2.107)$$

Avec:

$$a_{\delta'} = 2\sigma^2 P_{\Gamma}^{-1}\left(\frac{\nu}{2}, \delta'\right) \quad (2.108)$$

$$b_{\delta''} = 2\sigma^2 P_{\Gamma}^{-1}\left(\frac{\nu}{2}, 1 - \delta''\right) \quad (2.109)$$

$$K = \frac{1}{F_0(b_{\delta''}) - F_0(a_{\delta'})} \quad (2.110)$$

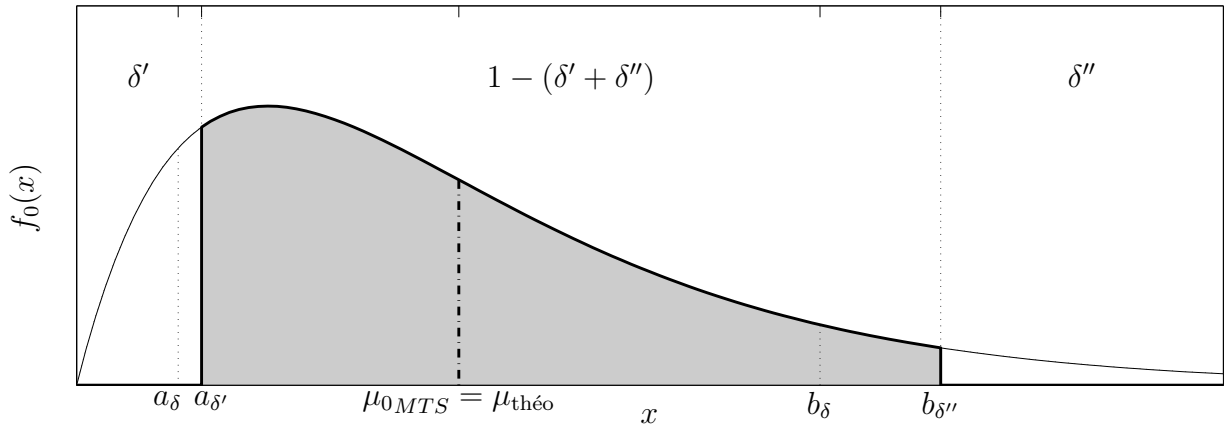


FIG. 2.22 – Densité de probabilité pour le calcul de la moyenne tronquée symétrique de  $\mu_{0MTS}$ .

La détermination de  $\delta'$  et  $\delta''$  passe par la connaissance de  $a_{\delta'}$  et  $b_{\delta''}$  qui sont les solutions du système d'équation suivant :

$$\begin{cases} \text{L'aire sous } f_0(x) \text{ entre } a_{\delta} \text{ et } b_{\delta''} \text{ égale l'aire sous } f_0(x) - 2\delta \\ \text{La moyenne de } f_{MTS}(x) \text{ doit égaler celle de } f_0(x) \end{cases} \quad (2.111)$$

$$\begin{cases} \int_{a_{\delta'}}^{b_{\delta''}} f_0(x) dx = 1 - 2\delta \\ \int_{a_{\delta'}}^{b_{\delta''}} x K f_0(x) dx = \int_0^{+\infty} x f_0(x) dx \end{cases} \quad (2.112)$$

$$\begin{cases} F_{\chi_{\nu,\sigma}^2}(b_{\delta''}) - F_{\chi_{\nu,\sigma}^2}(a_{\delta'}) - (1 - 2\delta) = 0 \\ F_{\chi_{\nu+2,\sigma}^2}(b_{\delta''}) - F_{\chi_{\nu+2,\sigma}^2}(a_{\delta'}) - \left( F_{\chi_{\nu,\sigma}^2}(b_{\delta''}) - F_{\chi_{\nu,\sigma}^2}(a_{\delta'}) \right) = 0 \end{cases} \quad (2.113)$$

La résolution est obtenue numériquement par la localisation de l'intersection des trois surfaces de la figure 2.23 en fixant  $\delta$  à 1 et en imposant  $\delta$  et  $\nu$ .

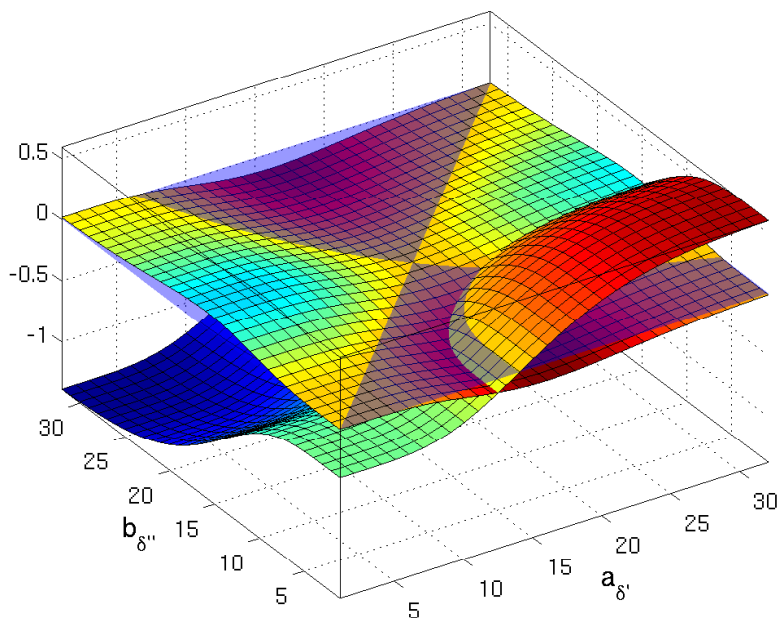


FIG. 2.23 – Moyenne tronquée symétrique : Le point  $[a_{\delta'}, b_{\delta''}]$  est situé à l'intersection des trois surfaces

### 2.2.6.1 ESPÉRANCE ET VARIANCE DE L'ESTIMATEUR

L'espérance et la variance de l'estimateur de moyenne par moyenne tronquée symétrique sont très proches de celles de l'estimateur par moyenne tronquée. Elles sont données par les équations 2.115 et 2.117 où  $q' = \lfloor N\delta' \rfloor$  et  $q'' = \lfloor N\delta'' \rfloor$  :

$$E \{ \hat{\mu}_{0MTS_{\delta}} \} = \frac{1}{N - (q' + q'')} \sum_{r=q'+1}^{N-q''} E \{ P_{Acco(r:N)} \} \quad (2.114)$$

$$= \frac{1}{N - (q' + q'')} \sum_{r=q'+1}^{N-q''} \int_{\mathbb{R}} x f_{X(r:N)}(x) \, dx \quad (2.115)$$

$$Var \{ \hat{\mu}_{0MTS_{\delta}} \} = \left( \frac{1}{N - (q' + q'')} \right)^2 \sum_{r=q'+1}^{N-q''} Var \{ P_{Acco(r:N)} \} \quad (2.116)$$

$$= \left( \frac{1}{N - (q' + q'')} \right)^2 \sum_{r=q'+1}^{N-q''} \left[ \int_{\mathbb{R}} x^2 f_{X(r:N)}(x) \, dx - \left[ \int_{\mathbb{R}} x f_{X(r:N)}(x) \, dx \right]^2 \right] \quad (2.117)$$

$$\text{avec } f_{X(r:N)}(x) = \frac{1}{B(r, N-r+1)} F_0(x)^{r-1} (1-F_0(x))^{N-r} f_0(x)$$

Les paramètres sont estimés numériquement en figure 2.24. Les légers écarts assimilables à un biais pour des valeurs de  $N$  faibles sont dus aux arrondis de  $q'$  et  $q''$ .

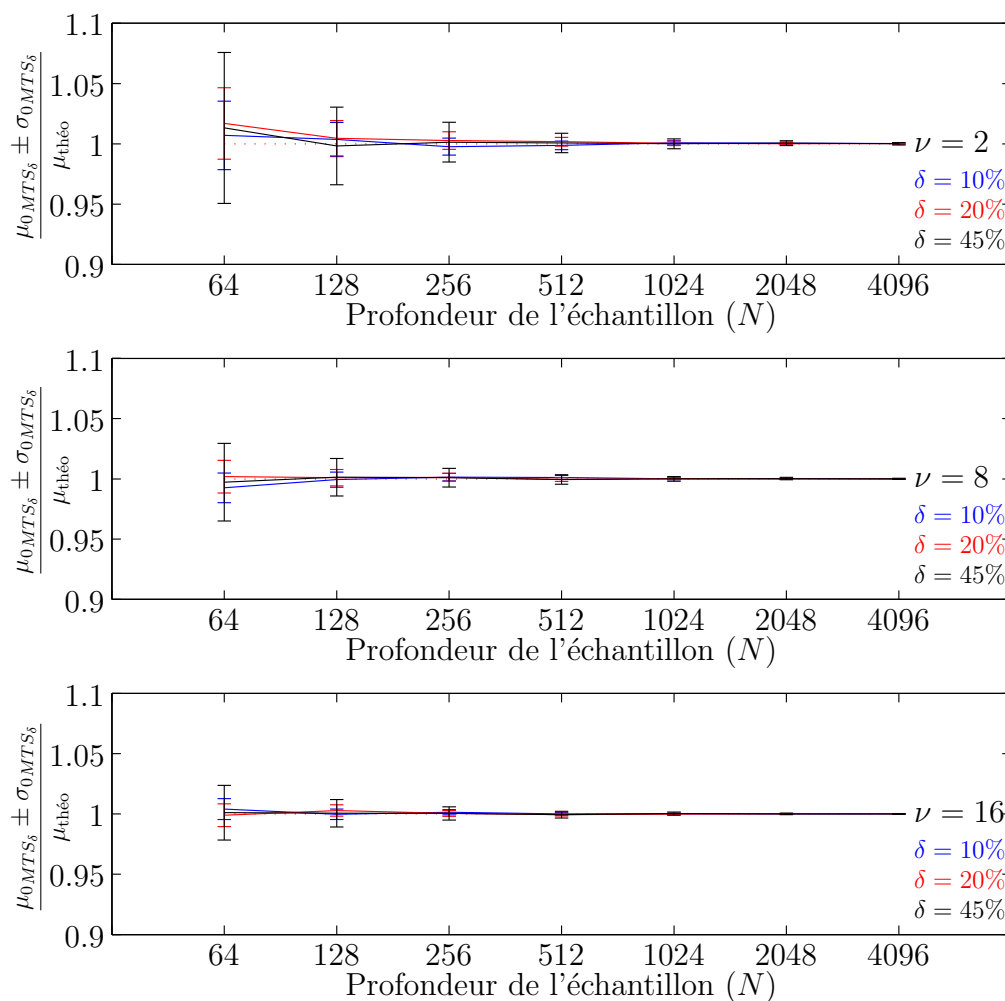


FIG. 2.24 – Moyenne  $\pm$  variance de l'estimateur de moyenne par moyenne tronquée symétrique normalisée par la moyenne théorique de la distribution en  $\chi^2$  à  $\nu$  degrés de liberté.

Comme pour l'estimateur par moyenne tronquée, espérance et variance peuvent être approchées en remplaçant les densités de probabilité de la statistique d'ordre  $f_{X(r:N)}(x)$  par  $f_\delta(x)$  :

$$E \{ \widehat{\mu}_{0MTS_\delta} \} \cong \frac{1}{N - (q' + q'')} \sum_{r=q'+1}^{N-q''} E \{ P_{Acc0(r:N)}[k] \} \quad (2.118)$$

$$\cong \int_{\mathcal{R}} x f_\delta(x) \, dx \quad (2.119)$$

$$\cong \frac{F_{\chi_{\nu+2,\sigma}^2}(b_{\delta''}) - F_{\chi_{\nu+2,\sigma}^2}(a_{\delta'})}{F_{\chi_{\nu,\sigma}^2}(b_{\delta''}) - F_{\chi_{\nu,\sigma}^2}(a_{\delta'})} \mu_{\text{théo}} \quad (2.120)$$

$$Var \{ \widehat{\mu}_{0MTS_\delta} \} \cong \left( \frac{1}{N - (q' + q'')} \right)^2 \sum_{r=q'+1}^{N-q''} Var \{ P_{Acc0,\delta}[k] \} \quad (2.121)$$

$$\cong \frac{1}{N - (q' + q'')} \left[ \int_{\mathcal{R}} x^2 f_\delta(x) \, dx - \left( \int_{\mathcal{R}} x f_\delta(x) \, dx \right)^2 \right] \quad (2.122)$$

$$\cong \frac{1}{N - (q' + q'')} \left( \nu(\nu + 2) \sigma^4 \frac{F_{\chi_{\nu+4,\sigma}^2}(b_{\delta''}) - F_{\chi_{\nu+4,\sigma}^2}(a_{\delta'})}{F_{\chi_{\nu,\sigma}^2}(b_{\delta''}) - F_{\chi_{\nu,\sigma}^2}(a_{\delta'})} - \mu_{\text{théo}}^2 \right) \quad (2.123)$$

Maintenant,  $\mu_{0MTS_\delta}$  présente un biais nul lors de l'estimation de  $\mu_{\text{théo}}$ . Le facteur  $L'(\nu, \delta)$  vaut 1 partout :

$$L'(\nu, \delta) = \frac{\mu_{0MTS_\delta}}{\mu_{\text{théo}}} \quad (2.124)$$

$$= \frac{F_{\chi_{\nu+2,\sigma}^2}(b_{\delta''}) - F_{\chi_{\nu+2,\sigma}^2}(a_{\delta'})}{F_{\chi_{\nu,\sigma}^2}(b_{\delta''}) - F_{\chi_{\nu,\sigma}^2}(a_{\delta'})} \quad (2.125)$$

$$= 1 \quad (2^{\text{e}} \text{ condition du système d'équation 2.113}) \quad (2.126)$$

### 2.2.6.2 DÉRIVE DE L'ESTIMATEUR

Comme pour l'estimateur de moyenne par moyenne tronquée, l'étude de l'effet des interférences est effectuée par simulation pour  $\delta = [10\%, 20\%, 49\%]$  avec les paramètres suivants :  $f_{RFI}(x) = \delta(x - P_{RFI})$ ,  $\nu = 16$ ,  $N = 4096$ ,  $\alpha = \frac{N_{RFI}}{N} = 0 \dots 100\%$  et  $INR_{dB} = 10 \log_{10} \left( \frac{P_{RFI}}{\mu_{\text{théo}}} \right) \in [-20, \dots, 70]$ . Les résultats, disponibles en figure 2.25, affichent une robustesse proche de celle de l'estimateur asymétrique décrit précédemment. On note néanmoins un basculement de la moyenne  $\widehat{\mu}_{0MTS_\delta}$  vers la valeur de l'interférence à un pourcentage  $\alpha$  de présence parasites plus faible de quelques points que pour  $\widehat{\mu}_{0MT}$ . C'est le résultat direct d'un seuillage haut  $b_{\delta''}$  plus élevé que  $b$ , celui de l'estimateur de moyenne par moyenne tronquée.

### 2.2.6.3 CONTRAINTES DE CALCUL

Les contraintes sont les mêmes que pour celles de l'estimation par moyenne tronquée en section 2.2.5.3.

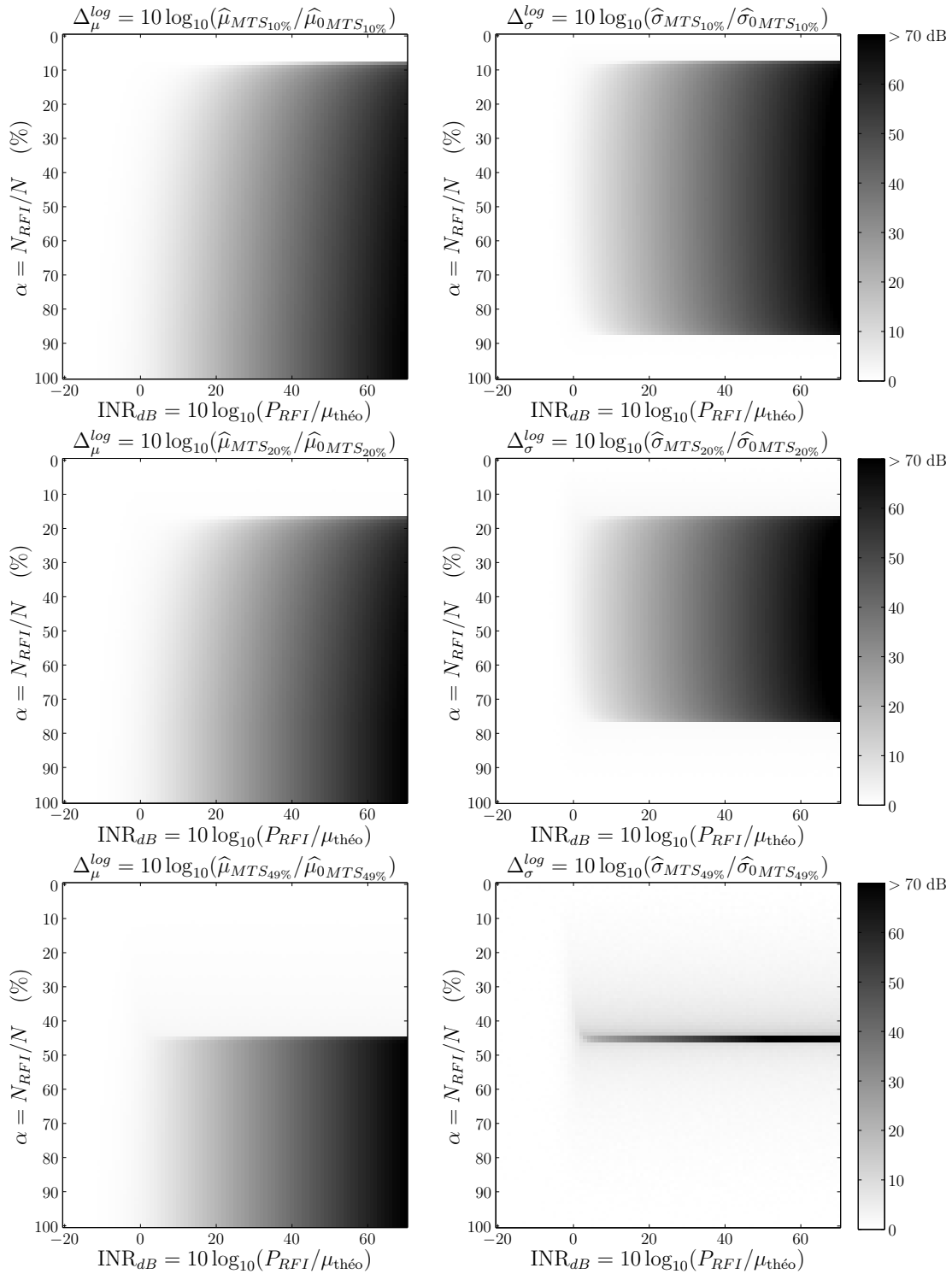


FIG. 2.25 – Moyenne et écart type de l'estimateur  $\hat{\mu}_{MTS_{\delta}}$  ( $\nu = 16$  et  $N = 4096$ ) en présence d'un parasite de taux de présence  $\alpha = N_{RFI}/N$  suivant une loi de type Dirac de moyenne  $P_{RFI}$ . (haut :  $\delta = 10\%$ , milieu :  $\delta = 20\%$ , bas :  $\delta = 49\%$ )

### 2.2.7 Moyenne du Petit Poucet

Le défaut de tous les estimateurs étudiés dans les sections précédentes est de fortement dévier quand le taux de parasites  $\alpha$  dépasse 50% dans les meilleurs cas. Pour certaines applications, il serait souhaitable de disposer d'un estimateur utilisable même pour  $\alpha > 50\%$ .

Les interférences que nous étudions produisent une hausse du niveau de puissance sur les échantillons du bloc de données étudié. Si tous les points du  $N$ -échantillon sauf un subissent l'interférence, le point de plus petite valeur a une forte probabilité d'être le point intègre. Cela nous conduit donc à utiliser encore les statistiques d'ordre.

Si l'hypothèse vraie est  $H_0$  (pas de parasite présent), l'estimation de la moyenne par la valeur du plus petit point du  $N$ -échantillon conduit nécessairement à une sous-évaluation de la moyenne. Il faut donc introduire un coefficient correcteur  $M(\nu, N)$  tel que (figure 2.26):

$$M(\nu, N) = \frac{\mu_{\text{théo}}}{E\{X_{(1:N)}\}} \quad (2.127)$$

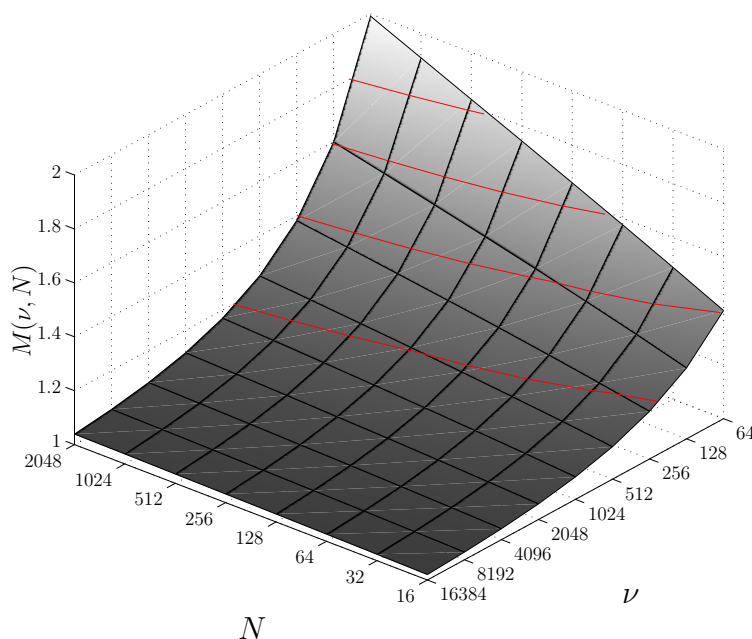


FIG. 2.26 – Fonction correctrice  $M(\nu, N)$  pour  $\nu = 2^{6 \cdots 16}$  et  $N = 2^{4 \cdots 11}$ .

Nous allons donc nous intéresser à l'estimateur de moyenne par moyenne du Petit Poucet  $\mu_{0PP}$  dont l'expression théorique est donnée en 2.128 et son estimation effectuée comme en 2.129.

$$\mu_{0PP} = N \int_{-\infty}^{+\infty} x \cdot M(\nu, N) [1 - F_0(x)]^{N-1} f_0(x) dx \quad (2.128)$$

$$\hat{\mu}_{0PP} = M(\nu, N) X_{(1:N)} \quad (2.129)$$

### 2.2.7.1 ESPÉRANCE ET VARIANCE DE L'ESTIMATEUR

L'expression des moments non-centrés de  $\widehat{\mu}_{0PP}^k$  est donnée par :

$$E \left\{ \widehat{\mu}_{0PP}^k \right\} = \int_{-\infty}^{+\infty} [M(\nu, N)x]^k f_{0(1:N)}(x) \, dx \quad (2.130)$$

$$E \left\{ \widehat{\mu}_{0PP}^k \right\} = N \int_{-\infty}^{+\infty} [M(\nu, N)x]^k [1 - F_0(x)]^{N-1} f_0(x) \, dx \quad (2.131)$$

L'espérance et l'écart type de l'estimateur sont évalués numériquement en figure 2.27.

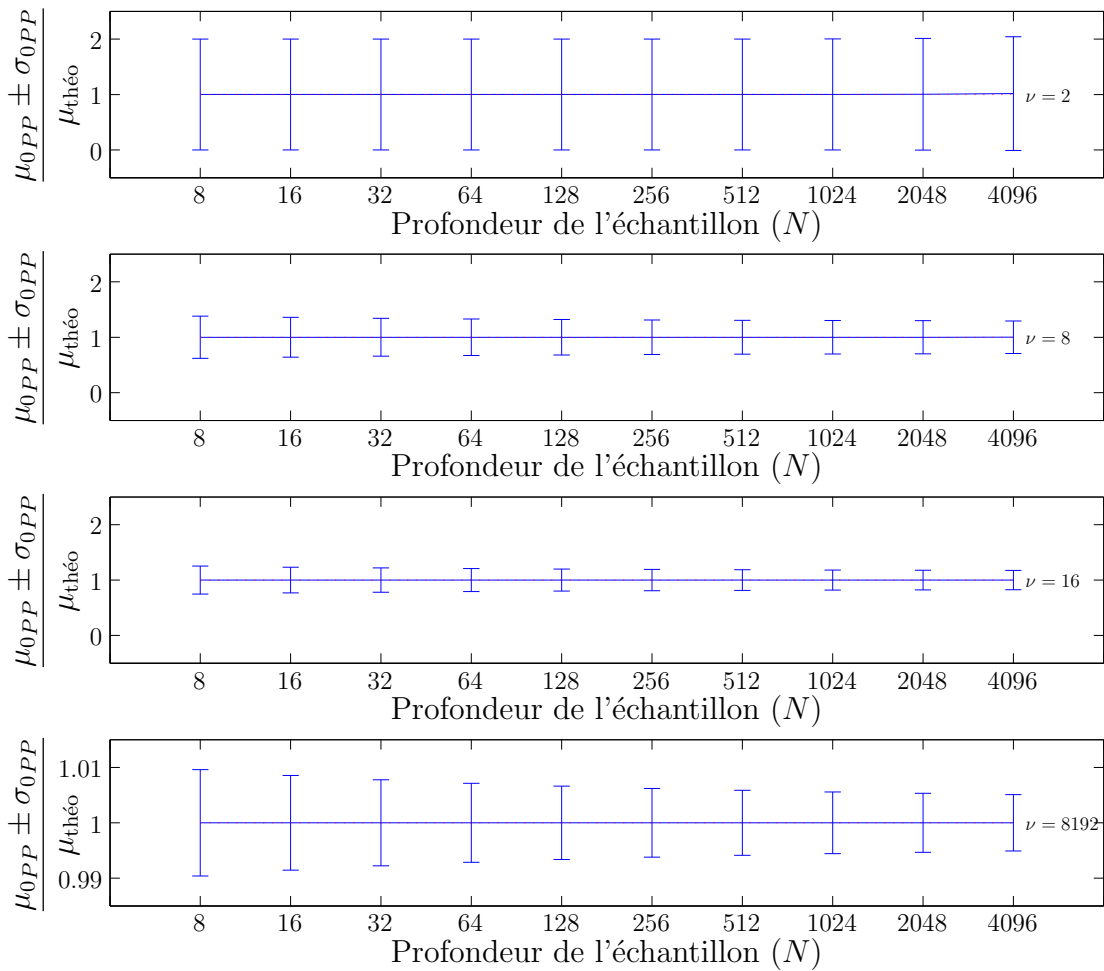


FIG. 2.27 – Moyenne  $\pm$  variance de l'estimateur de moyenne par moyenne du Petit Poucet normalisée par la moyenne théorique de la distribution en  $\chi^2$  à  $\nu$  degrés de liberté.

Le biais est nul par construction grâce à la correction apportée par  $M(\nu, N)$ . Par contre, l'écart type est désastreux pour de faibles valeurs de  $\nu$ . Il sera donc réservé pour un usage sur des spectres ayant nécessité de longues accumulations. Le point positif de cet estimateur est que ses performances d'estimation ne sont quasiment pas dépendantes de  $N$ . On pourra l'utiliser en particulier dans des applications où  $N$  est de valeur faible.

### 2.2.7.2 DÉRIVE DE L'ESTIMATEUR

Si l'hypothèse vraie est  $H_1$  (parasite présent), l'utilisation du coefficient correcteur  $M(\nu, N)$  surestime la moyenne car la variable ordonnée  $X_{(1:N)}$  devient  $X_{(1:N-N_{RFI})}$ . L'influence des interférences sur l'espérance et l'écart type de l'estimateur est obtenu par simulation avec les paramètres suivants :  $f_{RFI}(x) = \delta(x - P_{RFI})$ ,  $\nu = 8192$ ,  $N = 19$ ,  $\alpha = \frac{N_{RFI}}{N} = 0 \cdots 100\%$  et  $\text{INR}_{dB} = 10 \log_{10}(P_{RFI}/\mu_{\text{théo}}) \in [-20, \cdots, 70]$ . Nous noterons que la valeur de  $\nu$  utilisée pour décrire la dérive des estimateurs précédents ( $\nu = 16$ ) n'est pas applicable à l'étude de cet estimateur. Dans ces conditions, il se comporte déjà d'une façon abominable sans interférences. Il est donc inutile de le tester plus en avant dans des conditions plus sévères. Nous utiliserons plutôt  $N = 8192$  qui correspond à une application décrite en section 3.3.

On peut observer la grande robustesse de l'algorithme face aux interférences. L'estimateur ne dévie que pour  $\alpha = 100\%$ . L'écart type ne dépend pas du niveau des interférences mais seulement du taux de présence des interférences, comme prévu dans le paragraphe précédent.

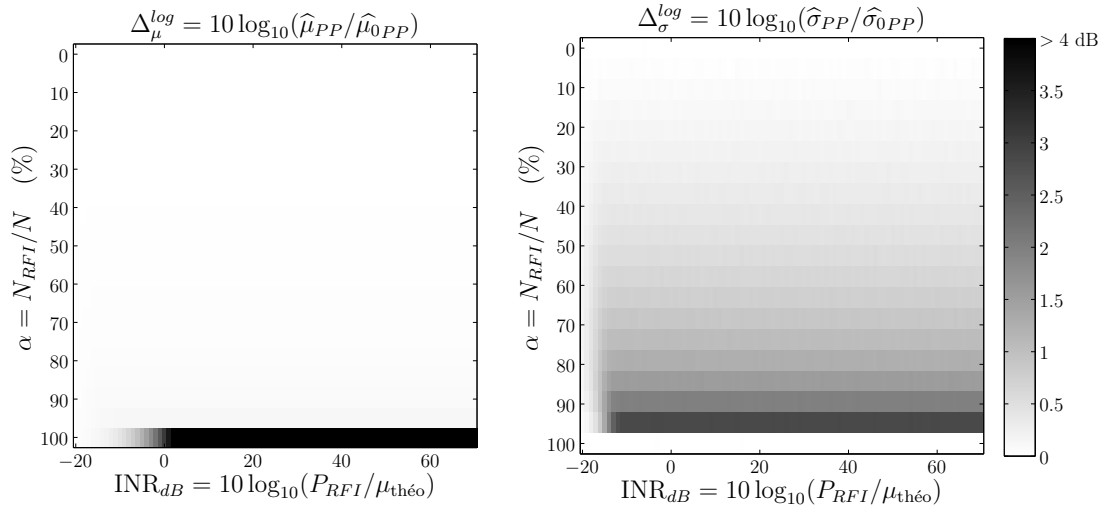
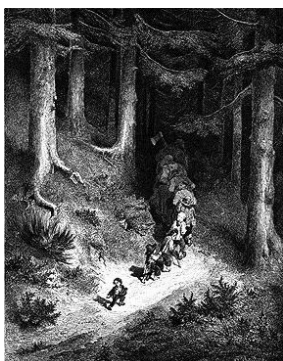


FIG. 2.28 – Moyenne et écart type de l'estimateur  $\hat{\mu}_{PP}$  ( $\nu = 8192$  et  $N = 19$ ) en présence d'un parasite de taux de présence  $\alpha = N_{RFI}/N$  suivant une loi de type Dirac de moyenne  $P_{RFI}$ .

### 2.2.7.3 CONTRAINTES DE CALCUL

Cet algorithme ne nécessite pas de classement. Il suffit de parcourir le  $N$ -échantillon à la recherche du point de plus faible valeur. Le produit de cette valeur par  $M(\nu, N)$  donne la moyenne recherchée. Encore une fois, il est possible d'inclure ce coefficient dans le calcul de seuil. C'est un algorithme très simple à implanter mais qui n'est robuste que pour  $\nu$  grand.





### MORALITE

*On ne s'afflige point d'avoir beaucoup d'enfants,  
Quand ils sont tous beaux, bien faits et bien grands,  
Et d'un extérieur qui brille;  
Mais si l'un d'eux est faible, ou ne dit mot,  
On le méprise, on le raille, on le pille :  
Quelquefois, cependant, c'est ce petit marmot  
Qui fera le bonheur de toute la famille.*

### 2.2.8 Moyenne de classe

L'estimateur de moyenne par moyenne du Petit Poucet est intéressant pour sa robustesse face aux interférences mais il présente néanmoins une forte variance. Tout comme les moyennes tronquées ont été un moyen de réduire la variance de l'estimateur par médiane, l'estimateur de moyenne par moyenne de classe  $\mu_{2C}$  cherche à allier la robustesse de la moyenne du Petit Poucet à une variance réduite.

L'estimateur du Petit Poucet travaille toujours dans le cas le plus défavorable assurant néanmoins un résultat utilisable, c'est à dire un taux  $\alpha$  de présence des interférences de  $\frac{N-1}{N}$ . Bien heureusement, les bandes d'observation ne sont pas systématiquement d'aussi piètre qualité et dans ces cas, il serait possible de réduire la variance de l'estimateur en moyennant l'ensemble des canaux intègres.

Le tout est de distinguer la classe des canaux intègres de celle des canaux parasités. De nouveau, il s'agit d'un problème de détection où un seuil va séparer les deux classes (voir section 2.1.2, page 50). La première classe est constituée des points non-corrompus du signal utile dont la distribution suit une loi du  $\chi^2$  (figure 2.29). La seconde classe est constituée des points corrompus (distribution non représentée sur la figure). Il reste à séparer ces 2 classes par un seuil que l'on calcule sur les données les moins sujettes aux altérations. On choisit ici le point du jeu de données qui présente la plus faible valeur ( $X_{(1:N)}$ ). Si ce point-là est corrompu alors cela signifie que tout le jeu de données est probablement altéré et qu'il ne faut pas espérer restaurer le signal utile par cette méthode. Le seuil est choisi comme l'espérance du point du jeu de données de plus forte valeur ( $X_{(N:N)}$ ) sous l'hypothèse  $H_0$  (sans interférences).

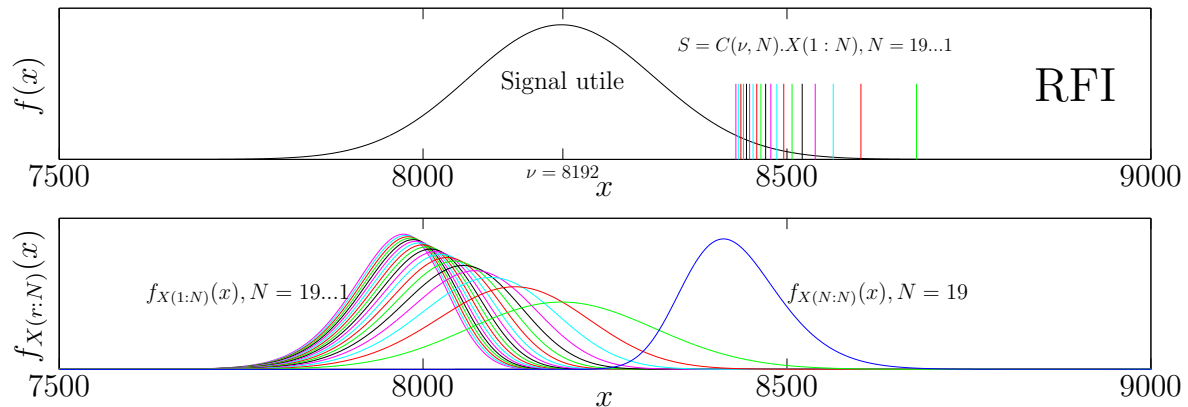


FIG. 2.29 – Principe de calcul de  $\mu_{02C}$  : la moyenne est obtenue par la somme normalisée des points inférieurs au seuil  $S$ . Ce seuil est calculé à l'aide de  $X_{(1:N)}$  et des lois  $f_{X_{(1:N)}}(x)$  et  $f_{X_{(N:N)}}(x)$  sous  $H_0$  où  $N$  est la taille du jeu de données. Dans le cas où une partie des points est corrompue,  $N$  diminue et la loi  $f_{X_{(1:N)}}(x)$  glisse vers les plus fortes valeurs. Cela implique une augmentation de la valeur du seuil mais, d'une part, cette variation reste dans un domaine de travail utilisable pour l'application décrite plus loin et, d'autre part, le taux de présence des interférences n'est pas connu, rendant impossible sa prise en compte dans une éventuelle correction.

Considérons le  $N$ -échantillon ordonné  $X_{(r:N)}$  composé de points entiers. Il est possible

de calculer en fonction de  $\nu$  et  $N$ , le rapport  $C(\nu, N)$  qui existe entre  $X_{(N:N)}$  et  $X_{(1:N)}$  :

$$C(\nu, N) = \frac{E \{X_{(N:N)}\}}{E \{X_{(1:N)}\}} \quad (2.132)$$

$$= \frac{\int_{-\infty}^{+\infty} x f_{0(N:N)}(x) dx}{\int_{-\infty}^{+\infty} x f_{0(1:N)}(x) dx} \quad (2.133)$$

où  $f_{0(r:N)}(x)$  est défini par 2.62, page 63.

Le rapport  $C(\nu, N)$  peut s'interpréter comme le rapport des valeurs moyennes des extrêmes du  $N$ -échantillon ordonné. Ses valeurs sont données en figure 2.30 pour  $\nu = 2^{6 \dots 14}$  et  $N = 2^{4 \dots 11}$ .

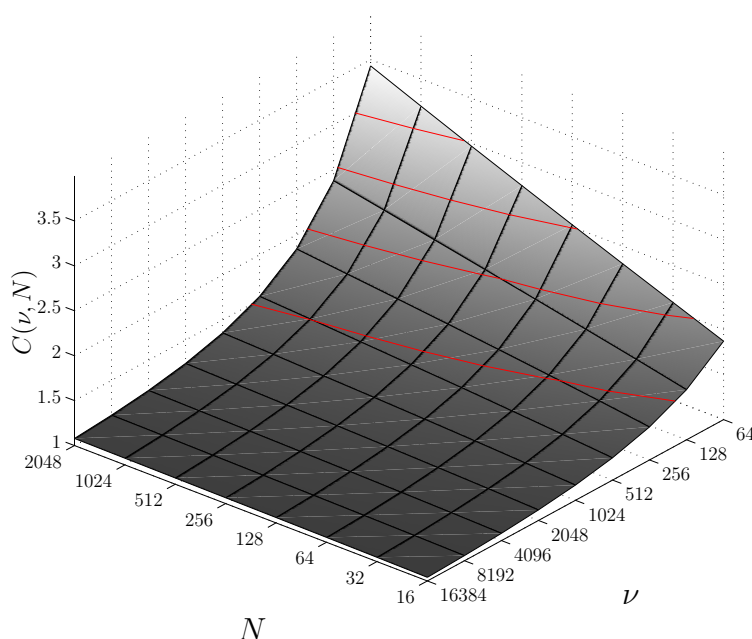


FIG. 2.30 – Le seuil  $S$  est calculé avec le rapport  $C(\nu, N)$  affiché pour  $\nu = 2^{6 \dots 14}$  et  $N = 2^{4 \dots 11}$ .

En présence exclusive d'échantillons intègres, on peut utiliser  $C(\nu, N)$  pour calculer un seuil  $S$  basé sur la valeur du plus petit échantillon  $X_{(1:N)}$  de  $X$  tel que :

$$S = C(\nu, N) \cdot X_{(1:N)} \quad (2.134)$$

La quasi-totalité des valeurs du  $N$ -échantillon va se situer sous la valeur de  $S$ . Il suffit d'effectuer la moyenne de ces points pour obtenir l'estimation de la moyenne :

$$\hat{\mu}_{2C} = \int_0^S x f(x) dx \quad (2.135)$$

$$\approx \int_{\mathbb{R}} x f_0(x) dx \quad (2.136)$$

Quand des émissions parasites altèrent une partie du  $N$ -échantillon, la valeur de ces échantillons dépasse le seuil rejetant ainsi leur usage pour le calcul de la moyenne. La

distribution étudiée n'est plus  $f_0(x)$  mais  $f(x)$  (voir section 2.2.1). On peut alors noter un défaut dans cet estimateur : le terme  $E \{X_{(1:N)}\}$  du calcul de  $C(\nu, N)$  devient  $E \{X_{(1:N-N_{RFI})}\}$  (figure 2.29). La valeur de  $C(\nu, N)$  devrait donc être adaptée pour tenir compte de  $N_{RFI}$  mais cette valeur n'est pas connue *a priori*. Nous devons donc travailler avec cette erreur et en étudier les effets (voir section 2.2.8.2).

### 2.2.8.1 ESPÉRANCE ET VARIANCE DE L'ESTIMATEUR

L'espérance et la variance de l'estimateur de moyenne par moyenne de classe ne sont pas dérivées analytiquement. Les résultats de simulation sont visible en figure 2.31.

Un biais est présent sur l'estimateur. Il diminue quand  $N$  ou  $\nu$  augmente. La variance de l'estimateur diminue quand  $N$  ou  $\nu$  augmente mais ses valeurs sont systématiquement plus faibles que celles de la variance du Petit Poucet (voir figure 2.27, p. 86). D'une manière générale, la barre d'erreur de  $\hat{\mu}_{02C}$  est incluse dans celle de  $\hat{\mu}_{0PP}$ .

### 2.2.8.2 DÉRIVE DE L'ESTIMATEUR

L'étude de l'effet des interférences est effectué par simulation avec les paramètres suivants :  $f_{RFI}(x) = \delta(P_{RFI})$ ,  $\nu = 8192$ ,  $N = 19$ ,  $\alpha = \frac{N_{RFI}}{N} = 0 \dots 100\%$  et  $INR_{dB} = 10 \log_{10}(\frac{P_{RFI}}{\mu_{théo}}) \in [-20, \dots, 70]$ . Les résultats, disponibles en figure 2.32, affichent une robustesse de la moyenne proche de celle de l'estimateur moyenne du Petit Poucet. On peut noter pour  $INR \approx -15dB$  que la moyenne augmente sensiblement pour toutes les valeurs de  $\alpha$ . La variance reste insensible au niveau des interférences mais on observe aussi une augmentation du niveau pour  $INR \approx -15dB$ . Ces effets sont dus aux non-détections, lors du seuillage par  $S$ , des échantillons faiblement parasités. Il faut néanmoins relativiser ces écarts qui restent inférieurs au décibel alors que l'estimateur est insensible à des interférences de plusieurs dizaines de décibels.

La figure 2.33 compare les dérivées des estimateurs  $\hat{\mu}_{02C}$  et  $\hat{\mu}_{0PP}$  en présence d'interférences. L'espérance de  $\hat{\mu}_{02C}$  vaut celle de  $\hat{\mu}_{0PP}$  à quelques centièmes de décibel près. La variance de  $\hat{\mu}_{02C}$  reste majoritairement plus faible que celle de  $\hat{\mu}_{0PP}$ . Elles sont strictement égales pour  $\alpha = \frac{N-1}{N}$ . La zone où  $INR \approx -15dB$  et  $\alpha$  grand met en valeur la faible robustesse de  $\hat{\mu}_{02C}$  par rapport à  $\hat{\mu}_{0PP}$  pour des interférences de faibles niveaux.

### 2.2.8.3 CONTRAINTES DE CALCUL

Cette extension de l'estimateur par moyenne du Petit Poucet est intéressante lorsque le meilleur SNR possible est recherché. Il a comme désavantage de produire une hausse de la variance lorsqu'une interférence de faible niveau se présente.

L'algorithme consiste en la recherche du point de plus faible valeur dans le  $N$ -échantillon. Le produit de ce point par  $C(\nu, N)$  fournit un seuil qui permet de décider de l'accumulation des autres points dans un registre. La division nécessaire à la moyenne dépend du jeu de donnée traité. Il est donc impossible d'inclure ce facteur dans le calcul de seuil final.

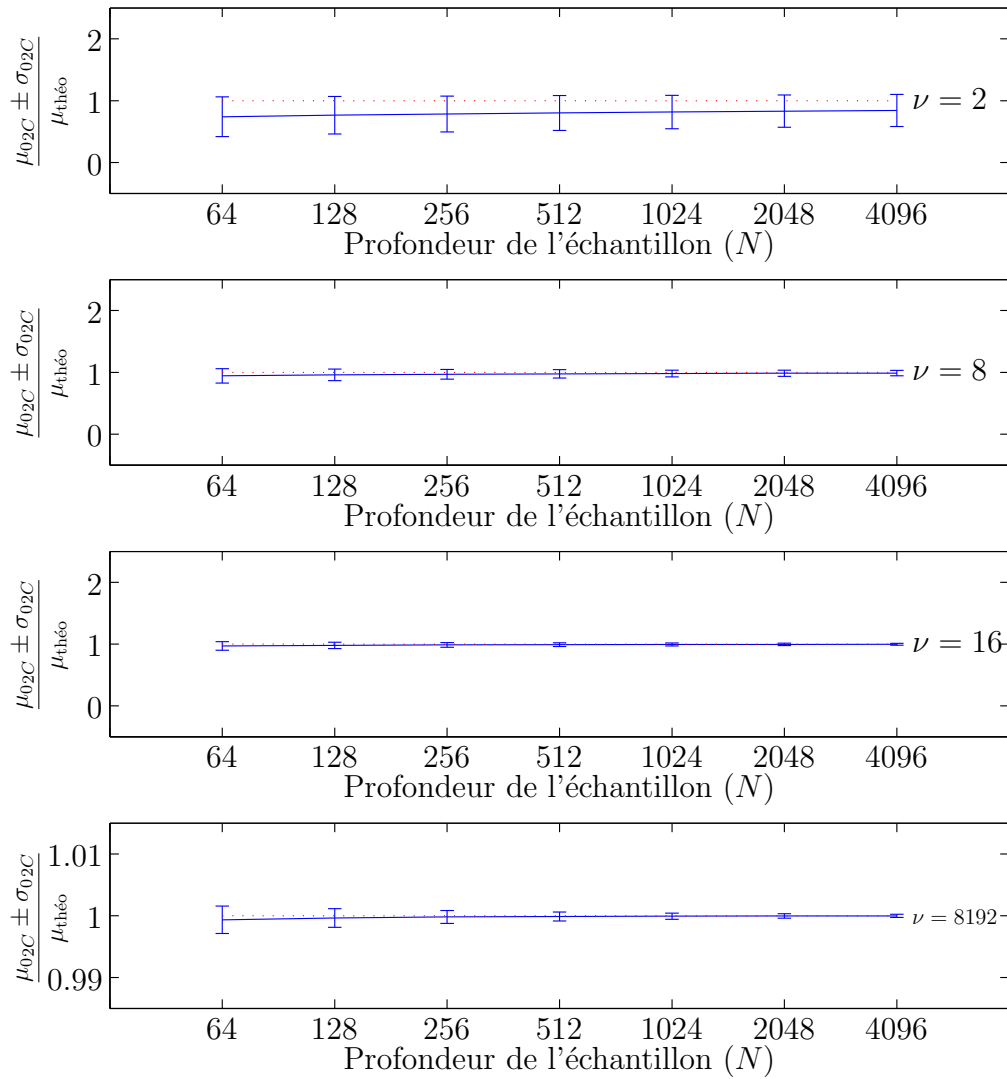


FIG. 2.31 – Moyenne  $\pm$  variance de l'estimateur de moyenne par moyenne de classe normalisée par la moyenne théorique de la distribution en  $\chi^2$  à  $\nu$  degrés de liberté.

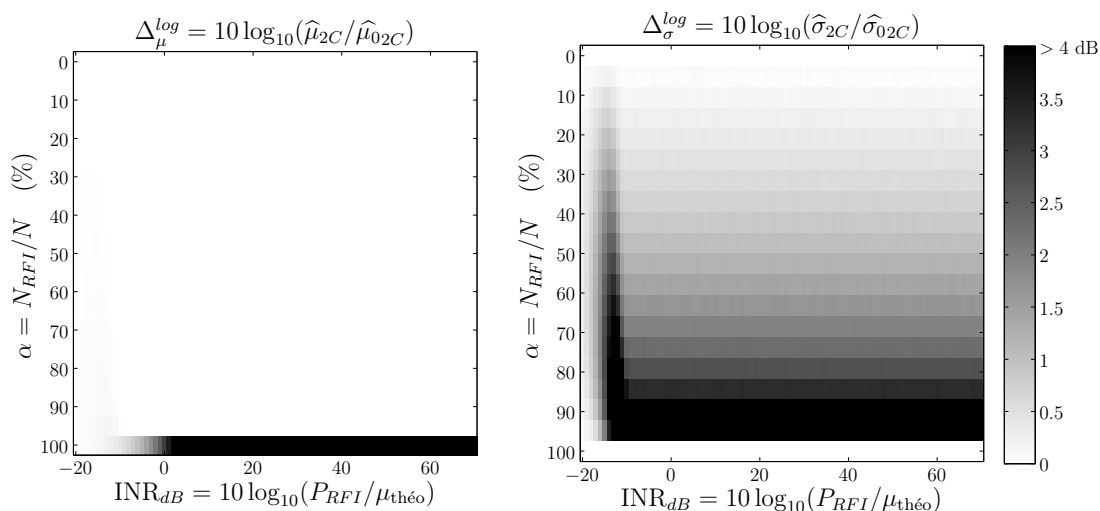


FIG. 2.32 – Moyenne et écart type de l'estimateur  $\hat{\mu}_{2C}$  ( $\nu = 8192$  et  $N = 19$ ) en présence d'un parasite de taux de présence  $\alpha = N_{RFI}/N$  suivant une loi de type Dirac de moyenne  $P_{RFI}$ .

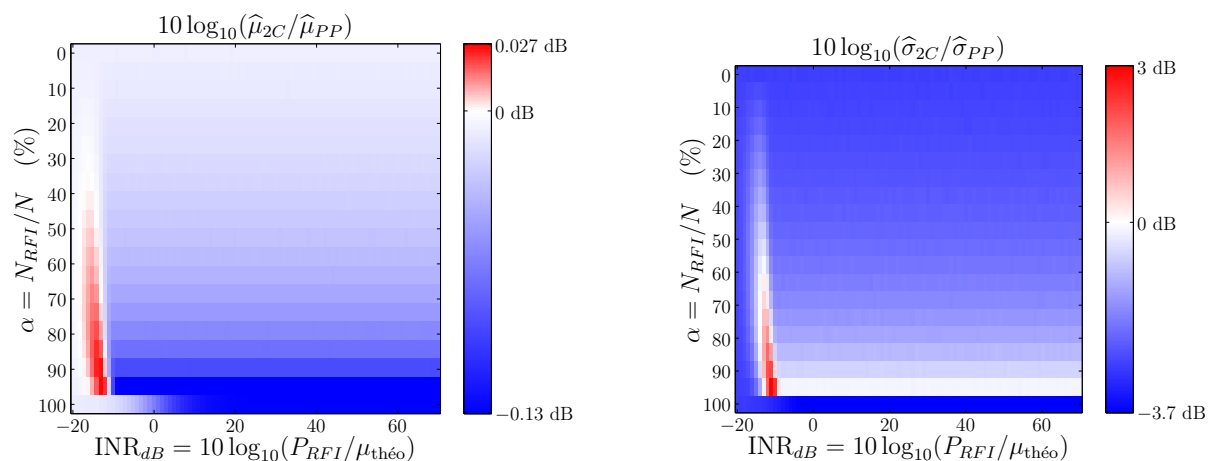


FIG. 2.33 – Comparaison des estimateurs  $\hat{\mu}_{2C}$  et  $\hat{\mu}_{PP}$  ( $\nu = 8192$  et  $N = 19$ ) en présence d'un parasite de taux de présence  $\alpha = N_{RFI}/N$  suivant une loi de type Dirac de moyenne  $P_{RFI}$ .

### 2.2.9 Filtre passe-bas IIR

Cet estimateur, noté  $\widehat{\mu}_{LP}$ , est basé sur l'estimation récursive de la moyenne de la puissance instantanée d'une forme d'onde complexe ( $\nu = 2$ ) :

$$\widehat{\mu}_{LP}[t] = \widehat{\mu}_{LP}[t-1] + \epsilon \cdot (x[t] - \widehat{\mu}_{LP}[t-1]) \quad \text{avec } \epsilon < 1 \quad (2.137)$$

En utilisant la transformée en  $z$  pour passer dans le domaine fréquentiel, 2.137 devient :

$$\widehat{M}_{LP}(z) = \widehat{M}_{LP}(z) \cdot z^{-1}(1 - \epsilon) + \epsilon X(z) \quad (2.138)$$

On peut ainsi facilement définir  $H(z)$  et dériver sa transformée en  $z$  inverse  $h[t]$  telles que :

$$H(z) = \frac{\widehat{M}_{LP}(z)}{X(z)} = \frac{\epsilon}{1 - (1 - \epsilon)z^{-1}} \quad (2.139)$$

$$h[t] = \epsilon(1 - \epsilon)^t \mathcal{U}[t] \quad (2.140)$$

On peut alors récrire 2.137 comme le filtrage non-récursif suivant :

$$\widehat{\mu}_{LP}[t] = \sum_{n=0}^{\infty} h[n] \cdot x[t-n] \quad (2.141)$$

#### 2.2.9.1 ESPÉRANCE ET VARIANCE DE L'ESTIMATEUR

En présence d'un signal intègre,  $x[t]$  est la puissance instantanée d'un bruit blanc gaussien de moyenne nulle. L'espérance et la variance de l'estimateur valent :

$$E\{\widehat{\mu}_{0LP}\} = \sum_{n=-\infty}^{\infty} E\{h[n] \cdot x_0[t-n]\} \quad (2.142)$$

$$= E\{\chi_{\nu=2,\sigma}^2\} \sum_{n=-\infty}^{\infty} h[n] \quad (2.143)$$

$$= E\{\chi_{\nu=2,\sigma}^2\} H(z)|_{z=e^{i0}} \quad (2.144)$$

$$= E\{\chi_{\nu=2,\sigma}^2\} \quad (2.145)$$

$$\text{Var}(\widehat{\mu}_{0LP}) = \sum_{n=-\infty}^{\infty} \text{Var}(h[n] \cdot x_0[t-n]) \quad (2.146)$$

$$= \text{Var}(\chi_{\nu=2,\sigma}^2) \sum_{n=-\infty}^{\infty} h^2[n] \quad (2.147)$$

$$= \frac{1}{2\epsilon^{-1} - 1} \text{Var}(\chi_{\nu=2,\sigma}^2) \quad (2.148)$$

Pour pouvoir comparer la variance de cet estimateur avec celle obtenue par la moyenne arithmétique classique, il faut établir un lien entre  $N$  et  $\epsilon$ . On peut écrire la moyenne arithmétique comme un filtrage par la réponse impulsionnelle  $g[t] = \frac{1}{N}\Pi_N[t]$  où  $\Pi_N[t]$  est la fonction porte de largeur  $N$ . Les deux filtres-estimateurs de moyenne auront les mêmes

performances si on égalise les erreurs quadratiques moyennes des signaux issus du filtrage par  $g[t]$  et  $h[t]$  d'un bruit suivant une loi du  $\chi_{\nu,\sigma}^2$ . On obtient :

$$EQM(g[t] * x_0[t]) = EQM(h[t] * x_0[t]) \quad (2.149)$$

$$\frac{1}{N} \sigma_{\text{théo}}^2 + \mu_{\text{théo}} = \frac{1}{2\epsilon^{-1} - 1} \sigma_{\text{théo}}^2 + \mu_{\text{théo}} \quad (2.150)$$

$$N = \frac{2}{\epsilon} - 1 \quad (2.151)$$

### 2.2.9.2 DÉRIVE DE L'ESTIMATEUR

Les estimateurs antérieurs supposaient un taux  $\alpha$  de points parasités dans l'observation sans que leur position dans l'échantillon ne soit critique. Ici, l'estimation est calculée au fil de l'eau : l'ordre d'apparition des échantillons parasités est donc crucial quant à la dérive de l'estimateur. Nous nous placerons dans un cas réaliste d'une interférence intermittente périodique observée en forme d'onde qui s'ajoute au bruit de mesure (figure 2.34). Nous considérerons que celle-ci présente des caractéristiques stables pendant un temps très supérieur aux constantes de temps du système.

$$x[t] = \mathcal{N}[t] + P_{RFI} \cdot \Pi_{N_{RFI}}[t \% T_0] \quad (2.152)$$

avec  $\mathcal{N}[t]$ : puissance instantané d'un bruit blanc gaussien de moyenne nulle intègre

$P_{RFI}$ : puissance instantanée de l'interférence

$N_{RFI}$ : durée de l'interférence

$T_0$ : période de l'interférence

$a \% b$ : opérateur modulo, le reste de la division euclidienne de  $a$  par  $b$

La formule précédente et l'équation 2.141 conduisent à l'espérance et la variance de l'estimateur en présence de parasite :

$$E\{\hat{\mu}_{LP}[t]\} = E\{\hat{\mu}_{0LP}\} + \Theta[t] \quad (2.153)$$

$$\text{avec } \Theta[t] = P_{RFI} \cdot \sum_{n=0}^{\infty} h[n] \cdot \Pi_{N_{RFI}}[(t-n) \% T_0] \quad (2.154)$$

$$\text{et } Var(\hat{\mu}_{LP}[t]) = Var(\hat{\mu}_{0LP}) \quad (2.155)$$

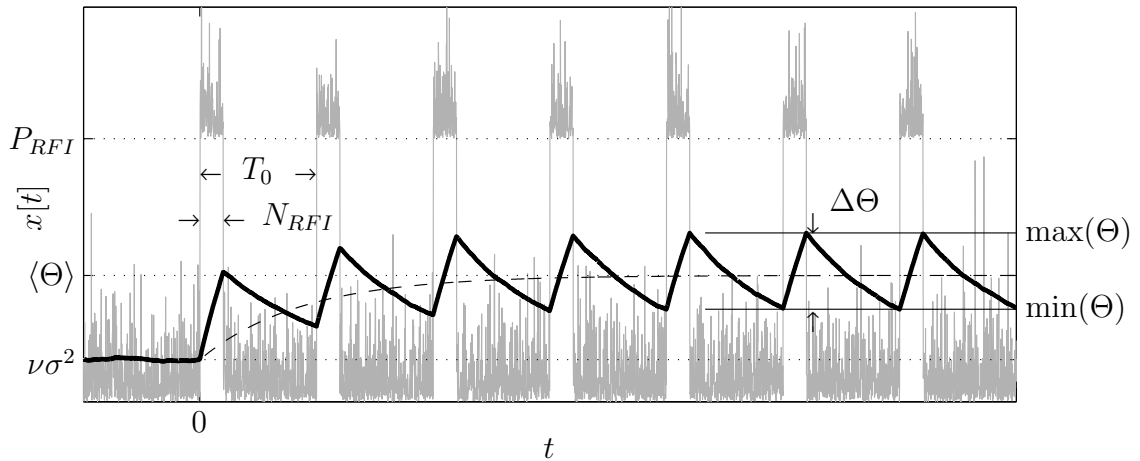
La puissance moyenne est donc systématiquement surévaluée. Lorsque le système "Interfère-Estimateur" est stabilisé (figure 2.34 pour  $t$  grand), la surévaluation  $\Theta[t]$  peut être encadrée en résolvant le système d'équations suivant :

$$\begin{cases} \max(\Theta) = \min(\Theta) + (P_{RFI} - \min(\Theta)) \sum_{n=0}^{N_{RFI}-1} h[n] \\ \min(\Theta) = \max(\Theta) \left(1 - \sum_{n=0}^{T_0-N_{RFI}-1} h[n]\right) \end{cases} \quad (2.156)$$

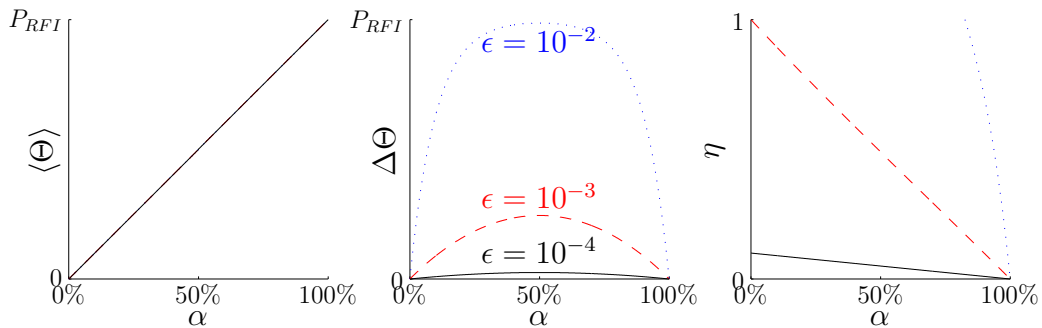
On peut alors calculer la surévaluation moyenne  $\langle \Theta \rangle$  et le taux d'ondulation  $\eta$  par :

$$\langle \Theta \rangle = \frac{1}{T_0} \sum_{t=0}^{T_0-1} \Theta[t] = P_{RFI} \frac{N_{RFI}}{T_0} \quad (2.157)$$



FIG. 2.34 –  $\hat{\mu}_{0LP}$  en présence d'interférences

$$\eta = \frac{\Delta\Theta}{\langle\Theta\rangle} = \frac{T_0}{N_{RFI}} \cdot \frac{\sum_{n=0}^{T_0-N_{RFI}-1} h[n] \sum_{n=0}^{N_{RFI}-1} h[n]}{\sum_{n=0}^{N_{RFI}-1} h[n] + \left(1 - \sum_{n=0}^{N_{RFI}-1} h[n]\right) \sum_{n=0}^{T_0-N_{RFI}-1} h[n]} \quad (2.158)$$

FIG. 2.35 – Evolution de  $\langle\Theta\rangle$  et  $\eta$  en fonction de  $\alpha = \frac{N_{RFI}}{T_0}$  et  $\epsilon$ 

D'une part, la figure 2.35 montre que  $\langle\Theta\rangle$  présente ses plus faibles valeurs pour  $\alpha$  faible. Le taux d'ondulation  $\eta$  est minimal pour  $\alpha = 100\%$  mais c'est une valeur peu utilisable car cela correspond à une corruption totale, et irréversible, des données. Cet estimateur n'est donc utilisable qu'avec des parasites à faibles taux de présence  $\alpha$ .

D'autre part, un  $\epsilon$  faible est nécessaire pour assurer un taux d'ondulation réduit. Néanmoins, excéder en ce sens conduit à une stabilisation trop lente du système lors de l'initialisation. On peut choisir comme temps d'initialisation la période  $T_0$  de l'interférence. Dans les deux cas, on assistera à un changement brusque de la moyenne du signal d'entrée et à la stabilisation de l'estimateur de moyenne sur sa nouvelle valeur. Il faut donc se placer dans le cadre d'étude où les caractéristiques de l'interfèreux changent. Nous considérerons alors le cas de l'apparition à  $t = 0$  (figure 2.34) d'une interférence sur le signal étudié.

$\Theta[t]$  devient :

$$\Theta'[t] = P_{RFI} \cdot \sum_{n=-\infty}^{\infty} h[n] [\Pi_{N_{RFI}}[(t-n)\%T_0] \cdot \mathcal{U}[t-n]] \quad (2.159)$$

Pour s'affranchir des variations de période  $T_0$  de  $\Theta'[t]$ , on étudie  $\langle \Theta' \rangle [t]$ , la moyenne glissante sur  $T_0$  échantillons de  $\Theta'[t]$  :

$$\langle \Theta' \rangle [t] = \frac{1}{T_0} \sum_{m=0}^{T_0-1} \Theta'[t-m] \quad (2.160)$$

$$= P_{RFI} \frac{N_{RFI}}{T_0} [1 - (1 - \epsilon)^{t+1}] \quad (2.161)$$

$\langle \Theta' \rangle [t]$  est visible en pointillé sur la figure 2.34. La stabilisation de l'estimateur au bout d'un temps  $T_0$  à  $100\beta\%$  près est obtenue lorsque la relation suivant est vérifiée :

$$\langle \Theta' \rangle [T_0] = (1 - \beta) \langle \Theta \rangle \quad (2.162)$$

$$\epsilon = 1 - 10^{\left(\frac{\log_{10}(\beta)}{T_0 + 1}\right)} \quad (2.163)$$

Pour faciliter la comparaison avec les autres estimateurs, la figure 2.36 a été fournie. On notera que l'espérance de  $\hat{\mu}_{LP}$  qui est comparée à celle de  $\mu_{\text{théo}}$  est en fait l'espérance moyenne de  $\hat{\mu}_{LP}$ . On notera aussi que la variance de l'estimateur n'est pas affectée par les interférences.

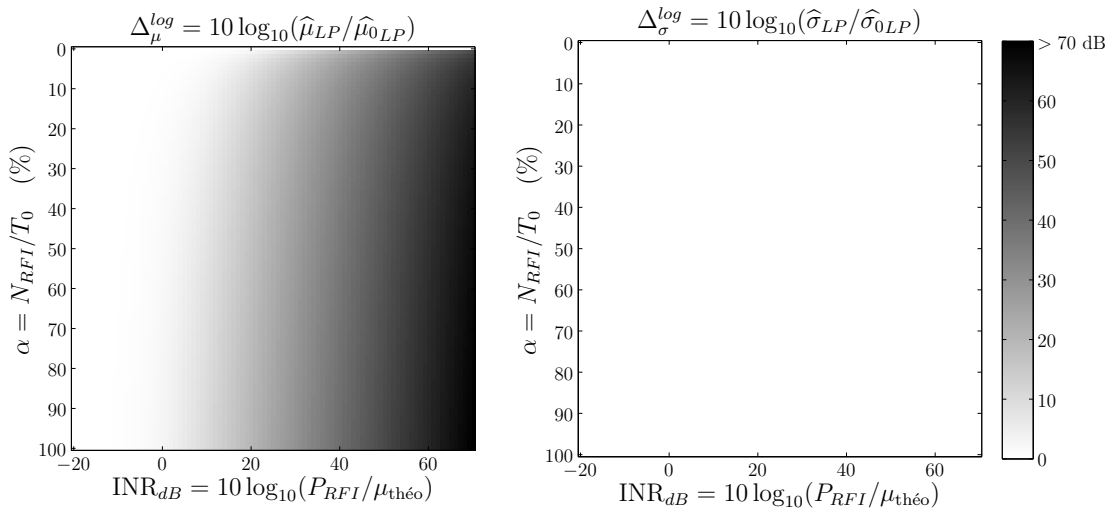


FIG. 2.36 – Moyenne et écart type de l'estimateur  $\hat{\mu}_{LP}$  en présence d'un parasite de moyenne  $P_{RFI}$  et de taux de présence  $\alpha = N_{RFI}/T_0$ .

### 2.2.9.3 CONTRAINTES DE CALCUL

L'implantation en FPGA d'un tel algorithme est très classique et ne nécessite que peu de matériel. Il faudra être attentif aux erreurs générées par la quantification lors de la multiplication par  $\epsilon$ .

### 2.2.10 Filtre passe-bas IIR seuillé

L'estimateur précédent présente une bonne robustesse mais son espérance est majorée par l'ajout de puissance des interférences. Il serait souhaitable de réduire cet effet. L'estimateur retenu,  $\hat{\mu}_{LPS}$ , ne prend pas en compte les points aberrants lors de la mise à jour du calcul de la moyenne. On calcule un seuil  $C \cdot \hat{\mu}_{lps}$  (figure 2.37) à partir de l'estimée de la moyenne  $\hat{\mu}_{lps}$  qui permet de rejeter les parasites les plus intenses. L'estimateur de moyenne  $\hat{\mu}_{lps}$  est donné par :

$$\hat{\mu}_{lps}[t] = \begin{cases} \hat{\mu}_{lps}[t-1] & \text{si } x[t] \geq C \cdot \hat{\mu}_{lps}[t-1] \\ \hat{\mu}_{lps}[t-1] + \epsilon \cdot (x[t] - \hat{\mu}_{lps}[t-1]) & \text{si } x[t] < C \cdot \hat{\mu}_{lps}[t-1] \\ \text{avec } 0 < \epsilon < 1 \end{cases} \quad (2.164)$$

La valeur de la moyenne est bien évidemment sous-estimée par  $\hat{\mu}_{lps}$  puisque les points intègres présents dans la "queue" de la distribution ( $\geq C \cdot \hat{\mu}_{lps}$ ) ne sont pas utilisés dans le calcul de la moyenne. Il faut donc corriger ce biais en estimant la dérive que produit  $C$  sur l'estimateur et fournir l'estimateur corrigé  $\hat{\mu}_{LPS}$  tel que :

$$\hat{\mu}_{LPS}[t] = L(C) \cdot \hat{\mu}_{lps}[t] \quad (2.165)$$

$$\text{avec } L(C) = \frac{\mu_{\text{théo}}}{\mu_{0lps}} = \frac{\nu\sigma^2}{\mu_{0lps}} \quad (2.166)$$

Pour évaluer  $L(C)$ , il faut calculer  $\mu_{0lps}$  pour  $\sigma = 1$ . La figure 2.37 présente la densité de probabilité  $f_{[0]}(x)$  issue du seuillage de  $f_0(x)$ . Les points supérieurs à  $C \cdot \mu_{0lps}$  ne sont pas simplement supprimés mais remplacés par des points issus du calcul de moyenne. Ces points suivent une distribution qui n'est pas encore connue mais, pour  $\epsilon$  faible et en s'appuyant sur l'équation 2.148, on peut considérer que sa variance est très faible. On considérera que cette distribution suit une loi de Dirac  $(1 - F_0(C \cdot \mu_{0lps})) \cdot \delta(x - \mu_{0lps})$ .

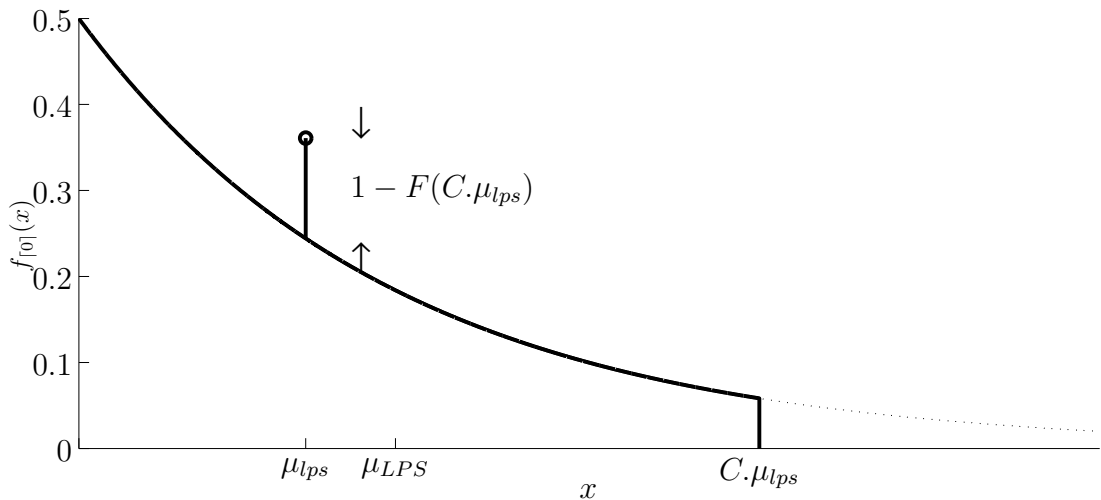


FIG. 2.37 – Densité de probabilité  $f_{[0]}(x)$  issue du seuillage de  $f_0(x)$ . Les points de la queue de la distribution sont reportés autour de  $\mu_{0lps}$ .

L'application visée est le traitement d'une forme d'onde au format IQ [49]. Nous serons donc dans le cas particulier où  $\nu = 2$ . La moyenne  $\mu_{0lps}$  est obtenu en intégrant  $xf_{[0]}(x)$  sur  $\mathbb{R}$ . Mais la densité de probabilité  $f_{[0]}(x)$  est elle-même dépendante de  $\mu_{0lps}$ . Il faut donc résoudre 2.168 qui, avec  $\nu = 2$  pour une forme d'onde complexe, se simplifie en 2.169.

$$\mu_{0lps} = \int_{\mathbb{R}} x f_{[\chi_{\nu=2, \sigma=1}^2]}(x) dx \quad (2.167)$$

$$\mu_{0lps} = \int_0^{C \cdot \mu_{0lps}} x \left[ f_{\chi_{\nu=2, \sigma=1}^2}(x) + \left(1 - F_{\chi_{\nu=2, \sigma=1}^2}(C \cdot \mu_{0lps})\right) \cdot \delta(x - \mu_{0lps}) \right] dx \quad (2.168)$$

$$\left[ C \cdot \mu_{0lps} + 2 - \mu_{0lps} \right] e^{-\frac{C \cdot \mu_{0lps}}{2}} - 2 + \mu_{0lps} = 0 \quad (2.169)$$

Le membre de gauche de 2.169 s'annule aux lieux  $[C, \mu_{0lps}]$  visibles en pointillé sur la figure 2.38. Les solutions pour  $\mu_{0lps} = 0$  ou  $C = 0$  ne sont pas utilisables pour résoudre notre problème. Un critère  $C$  strictement supérieur à 2 est alors nécessaire pour qu'une solution pour  $\mu_{0lps}$  existe.

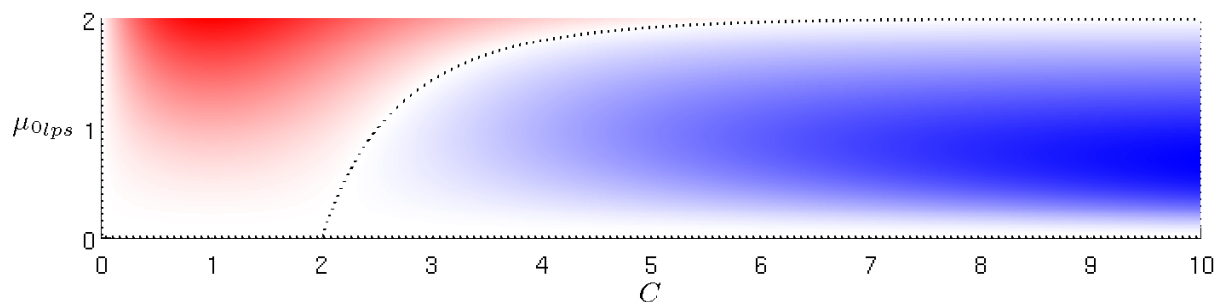
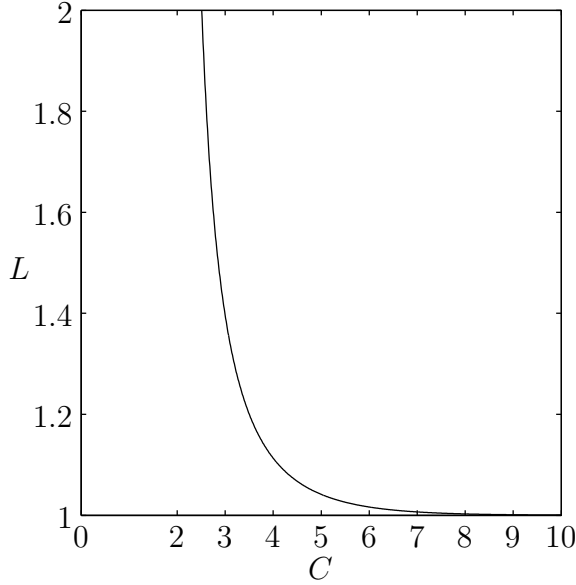


FIG. 2.38 –  $\mu_{0lps}$  en fonction de  $C$

On peut alors tabuler et tracer  $L(C)$  pour  $C$ . Pour des valeurs de  $C$  supérieures à 5,  $L(C)$  tend vers 1, soit une correction quasi-inexistante justifiée par une troncature minimale de la distribution  $f_{\chi_{\nu=2, \sigma}^2}(x)$ .

FIG. 2.39 – Fonction correctrice  $L(C)$ TAB. 2.2 – Tabulation de  $L(C)$ 

$C$	$L(C)$
2.5	2.032631
3.0	1.395916
3.5	1.200481
4.0	1.113117
4.5	1.067419
5.0	1.041448
5.5	1.025955
6.0	1.016432
6.5	1.010470
7.0	1.006694
7.5	1.004285
8.0	1.002744
8.5	1.001756
9.0	1.001122
9.5	1.000716
10.0	1.000456

### 2.2.10.1 ESPÉRANCE ET VARIANCE DE L'ESTIMATEUR

Pour calculer l'espérance et la variance de l'estimateur  $\hat{\mu}_{0lps}$ , on se place dans le même formalisme que la section 2.2.9.1 à la différence près que  $x[t]$  suit maintenant la distribution  $f_{[\chi_{\nu=2,\sigma}^2]}(x)$  de la figure 2.37.

$$E\{\hat{\mu}_{0LPS}\} = L(C) \sum_{n=-\infty}^{\infty} E\{h[n] \cdot x[t-n]\} \quad (2.170)$$

$$= L(C) \mu_{0lps} \sum_{n=-\infty}^{\infty} h[n] \quad (2.171)$$

$$= E\{\chi_{\nu=2,\sigma}^2\} \quad (2.172)$$

$$Var(\hat{\mu}_{0LPS}) = L(C)^2 \sum_{n=-\infty}^{\infty} Var(h[n] \cdot x[t-n]) \quad (2.173)$$

$$= L(C)^2 Var([\chi_{\nu=2,\sigma}^2]) \sum_{n=-\infty}^{\infty} h^2[n] \quad (2.174)$$

$$= \frac{1}{2\epsilon^{-1} - 1} L(C)^2 Var([\chi_{\nu=2,\sigma}^2]) \quad (2.175)$$

$Var([\chi_{\nu=2,\sigma}^2])$ , la variance de la distribution suivant la densité de probabilité  $f_{[\chi_{\nu=2,\sigma}^2]}(x)$  est estimée de façon classique :

$$Var([\chi_{\nu=2,\sigma}^2]) = E\{\hat{\mu}_{0lps}^2\} - E\{\hat{\mu}_{0lps}\}^2 \quad (2.176)$$

D'où:

$$\text{Var}(\hat{\mu}_{0LPS}) = \Omega(C) \frac{1}{2^{\epsilon-1} - 1} \text{Var}(\chi_{\nu=2,\sigma}^2) \quad (2.177)$$

$$\text{et } \text{Var}(\hat{\mu}_{0LPS}) = \Omega(C) \text{Var}(\hat{\mu}_{0LP}) \quad (2.178)$$

$$\text{avec : } \Omega(C) = 2L(C)^2 - 1 - [(C + L(C))^2 + L(C)^2 - 1] e^{-\frac{C}{L(C)}} \quad (2.179)$$

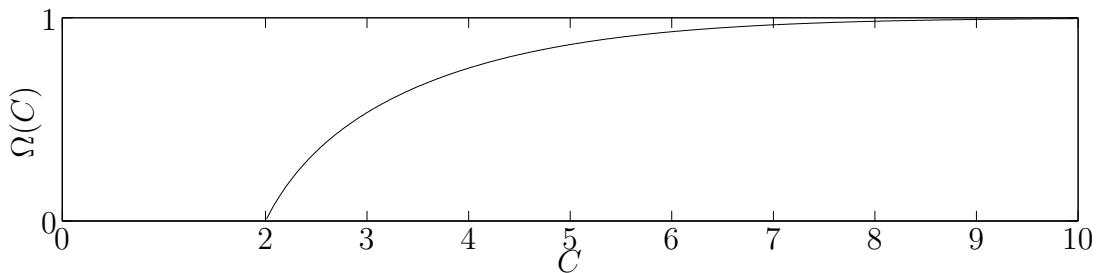


FIG. 2.40 –  $\Omega(C)$  : variance de l'estimateur  $\hat{\mu}_{0LPS}$  normalisée par la variance de  $\hat{\mu}_{0LP}$  en fonction du critère de seuillage  $C$

On peut déduire des expressions et figures précédentes que l'estimateur  $\hat{\mu}_{0LPS}$  est non-biaisé et que sa variance est toujours plus faible que celle de  $\hat{\mu}_{0LP}$ .

### 2.2.10.2 DÉRIVE DE L'ESTIMATEUR

La simulation de la dérive de l'estimateur en présence d'interférences qui permet d'obtenir la figure 2.41 a été menée avec les paramètres suivants:  $f_{RFI}(x) = \delta(x - P_{RFI})$ ,  $\nu = 2$ ,  $\alpha = \frac{N_{RFI}}{N} = 0 \dots 100\%$  et  $INR_{dB} = 10 \log_{10}(\frac{P_{RFI}}{\mu_{théo}}) \in [-20, \dots, 70]$  pour 3 valeurs de critère différentes :  $C = \{3, 4, 5\}$ .

Afin d'obtenir des résultats comparables avec les autres estimateurs, c'est l'espérance moyenne de  $\hat{\mu}_{LPS}$  qui est comparée à  $\mu_{théo}$ . Afin de ne pas améliorer artificiellement les résultats, les échantillons de l'estimateur de moyenne  $\hat{\mu}_{LPS}[i]$  qui ont été figés pour cause d'interférences ne sont pas inclus dans le calcul de la variance.

La figure montre une insensibilité de l'estimateur aux interférences de fort niveau même pour un taux d'occupation élevé (mais inférieur à 100%).

Pour des interférences de faible niveau (-10 à 5 dB), l'espérance et la variance de l'estimateur dérivent légèrement à cause des non-détections d'interférences.

### 2.2.10.3 CONTRAINTES DE CALCUL

L'implantation en FPGA d'un tel algorithme est très classique et ne nécessite que peu de matériel. Il est intéressant pour  $C < 5$  car  $\text{Var}(\hat{\mu}_{0LPS}) < \text{Var}(\hat{\mu}_{0LP})$  mais la correction  $L(C)$  est nécessaire. On peut néanmoins l'inclure dans l'étape du calcul de  $S$  de l'équation 2.23.

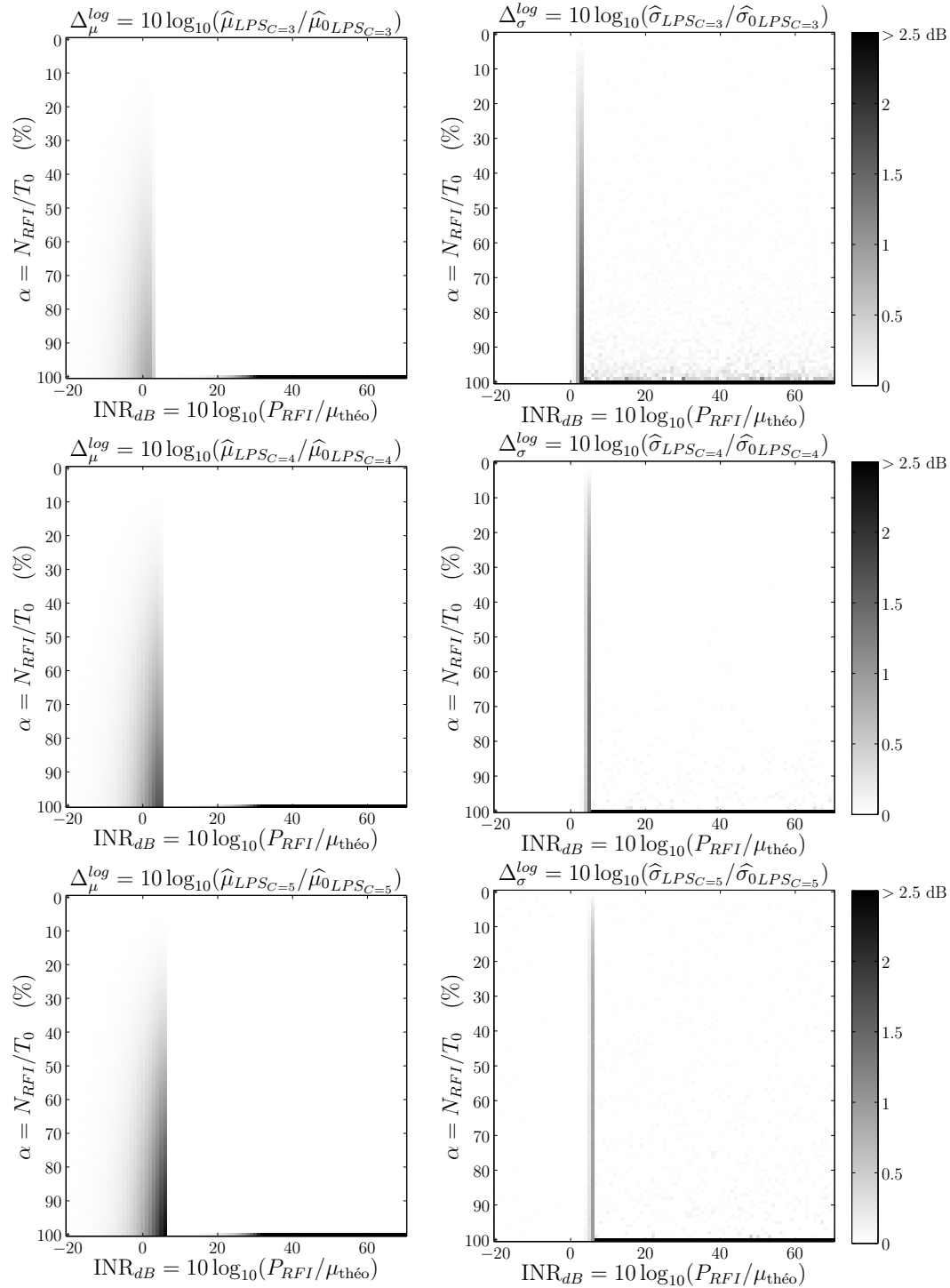


FIG. 2.41 – Moyenne et écart type de l'estimateur  $\hat{\mu}_{LPS}$  en présence d'un parasite de moyenne  $P_{RFI}$  et de taux de présence  $\alpha = N_{RFI}/N$  pour  $C = \{3, 4, 5\}$ .

## 2.3 Comparatif des estimateurs

Le tableau 2.3 résume les spécificités des estimateurs précédemment étudiés. Le premier critère rend compte de la qualité de l'estimation en absence d'interférence : biais et écart type de l'estimateur. Le deuxième paramètre renseigne sur la qualité de la robustesse de l'estimateur face aux RFI et le dernier paramètre résume les étapes principales de l'algorithme et le coût calculatoire engendré. Les étapes de calcul sont indiquées avec la notation  $\text{CALCUL}[x]$ , où  $\text{CALCUL}$  identifie le calcul effectué sur  $x$  points. Par exemple,  $\text{Min}[N]$  indique la recherche de la valeur minimum d'un tableau de  $N$  points.

TAB. 2.3 – Comparatif des estimateurs de moyenne

Estimateur	Qualité sans RFI		Qualité avec RFI	Implantation	
	Biais	écart type	Robustesse	Coût de calcul	Types de calculs
Moyenne arithmétique	0	faible	faible	faible	Accumulation[ $N$ ]
Médiane	$> 0$ pour $\nu$ faible	faible	forte	élevé	Classement[ $N$ ]
Médiane corrigée	0	faible	forte	élevé	Classement[ $N$ ] Produit[1]
Moyenne tronquée	$> 0$ pour $\nu$ faible	très faible	forte	élevé	Classement[ $N$ ] Accumulation[ $\delta N$ ]
Moyenne tronquée symétrique	0	très faible	forte	élevé	Classement[ $N$ ] Accumulation[ $\delta N$ ]
Moyenne du Petit Poucet	0	fort	très forte	faible	Min[ $N$ ]
Moyenne de classe	$> 0$ pour $\nu$ faible	très faible	très forte	assez faible	Min[ $N$ ] Seuillage Accumulation [ $< N$ ]
Filtre passe-bas IIR	0	$\propto \epsilon$	forte	faible	Différence[1] Produit[1] Somme[1]
Filtre passe-bas IIR seuillé	0	$\propto \epsilon$	très forte	faible	Différence[1] Produit[1] Somme[1] Seuillage[1]

Nous pouvons voir que quatre classes principales d'estimateurs se détachent. La première est constituée du simple estimateur de moyenne par moyenne arithmétique. La deuxième classe regroupe les algorithmes basés sur les statistiques d'ordre. La troisième classe base son estimation sur les *minima* du jeu de données, sans chercher à en effectuer un classement exhaustif. La dernière classe regroupe deux estimateurs itératifs : le classique filtre passe-bas et sa version seuillée.



Il est possible de réorganiser ce tableau en deux groupes principaux composés des estimateurs itératifs et des algorithmes non-itératifs. On voit alors d'une part que le premier groupe comporte le filtre passe-bas seuillé qui apporte des améliorations à la robustesse du filtre passe bas classique. Et d'autre part, le second groupe présente une succession d'estimateurs qui améliorent les qualités de leurs prédécesseurs : la moyenne arithmétique est surpassée en robustesse par la médiane. Le biais d'estimation peut être réduit si on utilise la médiane corrigée. L'écart type en est réduit si la moyenne tronquée est utilisée, au détriment d'un léger biais qui disparaît lors de l'utilisation de la moyenne tronquée symétrique. Dans certaines configurations ( $\nu$  élevé), il est encore possible d'améliorer la robustesse aux points aberrants en utilisant la moyenne du Petit Poucet. Son écart type relativement élevé peut être réduit en utilisant la moyenne de classe, au prix d'un léger biais supplémentaire.

Bien évidemment ces multiples améliorations ont souvent comme prix à payer une augmentation de la charge de calcul. De plus, certaines applications seront moins sensibles que d'autres à un biais ou une écart type élevé. C'est donc au cas par cas que sera choisi un estimateur.

Nous verrons dans la partie suivante que tous ces estimateurs n'ont pas été développés en vain lors de la recherche d'un estimateur universel introuvable (mais l'auteur envisage volontiers le fait d'avoir mal cherché). Au contraire, presque tous les estimateurs trouvent leur utilité dans l'une des applications présentées ci-après.

Le diagramme de la figure 2.42 permet de choisir parmi ceux précédemment présentés le meilleur pour résoudre un problème d'atténuation d'interférence, ou plus généralement, le meilleur estimateur de moyenne en présence de points aberrants. La sélection pourra s'opérer selon que l'on a besoin d'effectuer l'estimation de façon récursive ou non, en présence de rares ou de nombreux points aberrants et selon si l'on souhaite augmenter encore la charge de calcul pour obtenir un estimateur de plus faible variance et/ou faiblement biaisé. Les trois applications notées en grisé sont présentées au chapitre 3

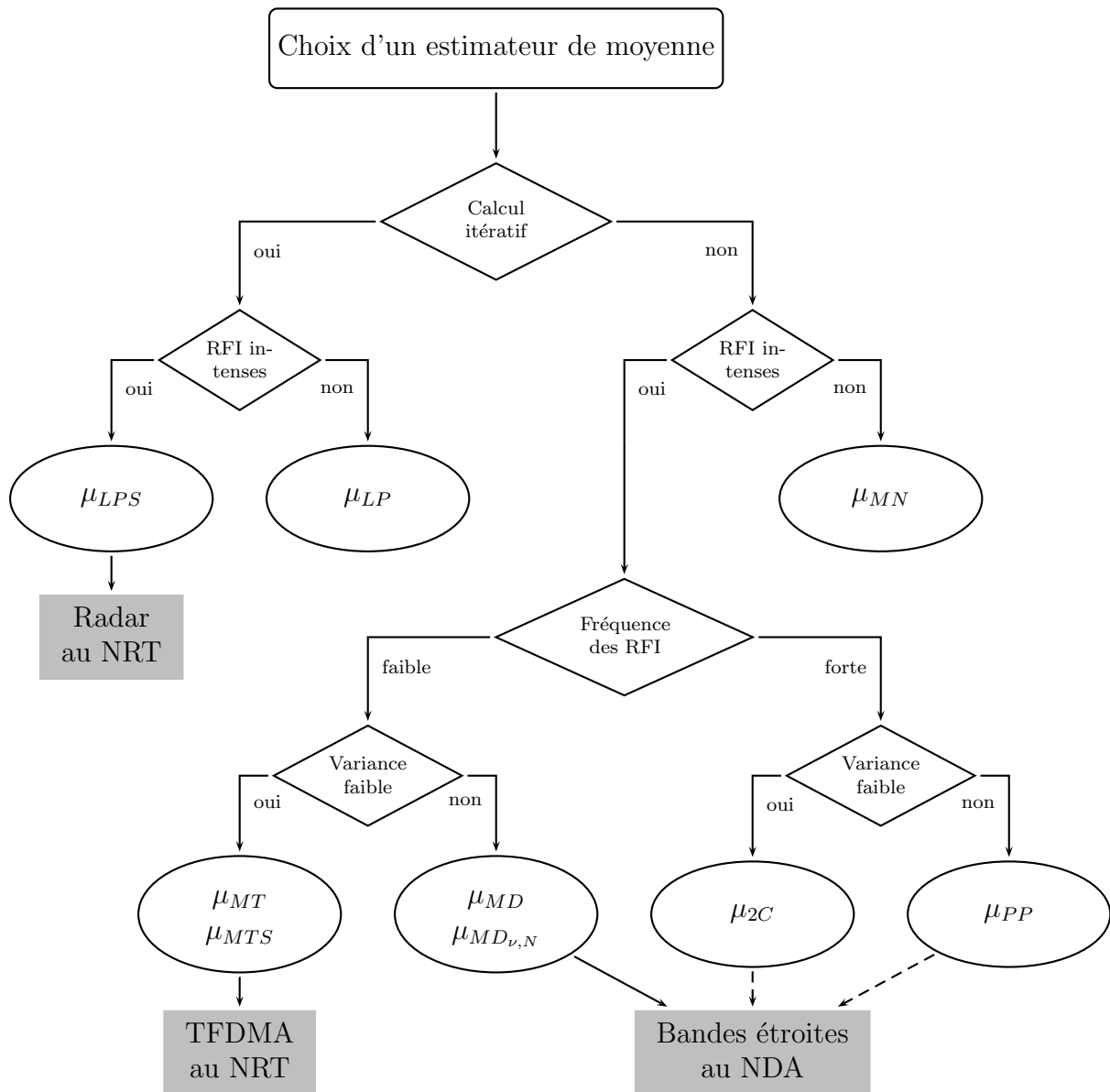


FIG. 2.42 – Choix d'un estimateur de moyenne



## CHAPITRE 3

---

# APPLICATIONS : BLANKING COMMANDÉ EN PUISSANCE

---

Tous les estimateurs de moyenne précédemment exposés n'ont pas été étudiés au hasard. Ils résultent d'une heuristique guidée par la recherche d'une solution à des problèmes instrumentaux. Certains estimateurs ont l'air très mauvais, mais ce sont parfois les meilleurs utilisables dans une situation donnée.

Trois applications sont présentées. La première fait usage de l'estimateur itératif  $\hat{\mu}_{LPS}$  pour supprimer dans le domaine temporel des interférences générées par des systèmes radar au sol qui parasitent des observations de raies HI. La deuxième application utilise l'estimateur de moyenne par moyenne tronquée symétrique  $\hat{\mu}_{MTS}$  pour détecter et “blanker”, dans chaque sous bande du plan temps-fréquence, les émissions d'un système de communication satellitaire à partage de canal TFDMA lors des observations de raies OH. La dernière application traite de l'atténuation des interférences à bande étroite dans le domaine fréquentiel rencontrées lors des observations du Soleil et de Jupiter dans la bande décimétrique.

### 3.1 Parasites large bande intermittents : Radar au NRT

Au NRT, une source majeure d'interférences provient des systèmes de radio-localisation aérien au sol (radars). Ces systèmes sont constitués d'un lobe tournant (pointage mécanique ou électronique) d'une période de quelques secondes. Des impulsions très brèves ( $\approx \mu s$  soient quelques mégahertz de bande) et modulées en fréquence sont émises avec une période de répétition de l'ordre de la milliseconde. La mesure du temps de vol de l'impulsion permet de déterminer la distance aéronef-antenne. Le signal s'introduit principalement dans la chaîne de réception du télescope *via* les lobes secondaires du diagramme d'antenne et de façon occasionnelle par le lobe principal, lorsque le signal radar subit une réflexion directe sur un aéronef.

Le signal interférant est donc caractérisé par trois pseudo-constantes de temps comme le montre la figure 3.1. Il est constitué d'une enveloppe de période de l'ordre de 4 à 5 secondes modulant un train d'impulsions de période  $\approx 1 ms$  et de durée  $\approx 1$  à  $60 \mu s$  (non visible sur la figure).

Lorsque plusieurs radars sont présents dans la même bande de fréquence (fig. 3.2), il arrive que les impulsions s'enchevêtrent (fig. 3.1, courbe du bas).

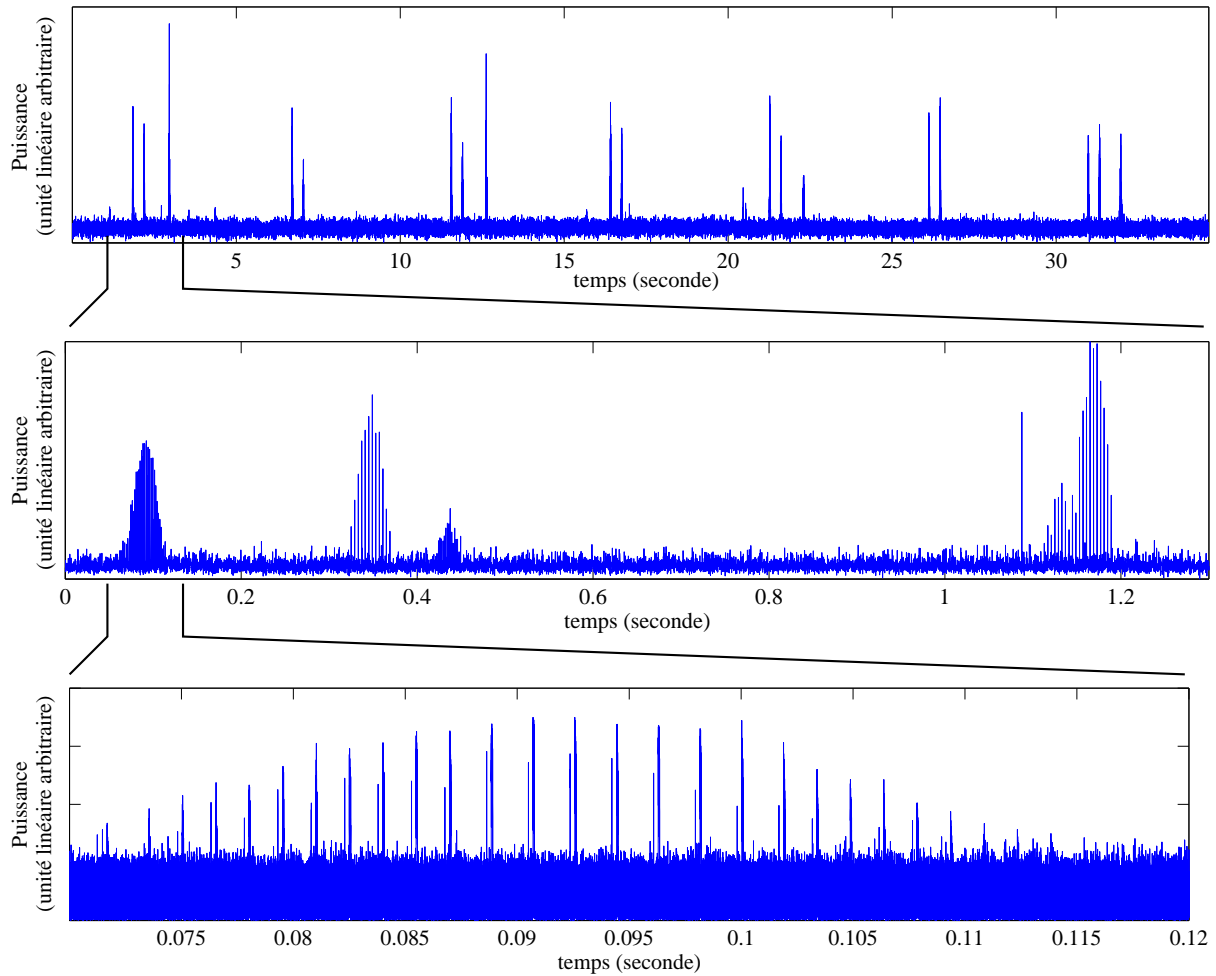


FIG. 3.1 – Puissance instantanée d’une interférence radar.

Le problème posé par ce type d’interférence est flagrant lors d’observations spectroscopiques de raies HI “red-shiftées” (voir sections 1.4.1.1 et 1.1.3). La raie au repos est localisée à 1420.406 MHz, dans une bande protégée. Mais son observation dans des objets extragalactiques implique des vitesses relatives de sources importantes et donc un effet Doppler important de sorte que la fréquence observée sort de la bande protégée. Les observations de test ont été effectuées sur des sources dont les vitesses variaient entre 12000 et 23000 km/s soient des raies observées entre 1310 MHz à 1365 MHz (Tableau 3.2). C’est dans cette bande, présentée en figure 3.2, que se répartissent les émetteurs radar.

La procédure d’observation de type ON-OFF tend à masquer très largement les distortions produites par ce type d’émetteur. L’émission parasitante étant similaire d’un cycle ON à un cycle OFF, la différence des deux atténue naturellement l’amplitude de l’interférence (figure 3.3). Néanmoins, les résidus dus à la différence de niveau des émissions (aussi faible soit-elle) sont suffisants pour masquer totalement le signal utile ou, pire, pour créer de parfaits artéfacts de radiosources (voir la courbe rouge de la figure 3.14 pour un exemple typique).

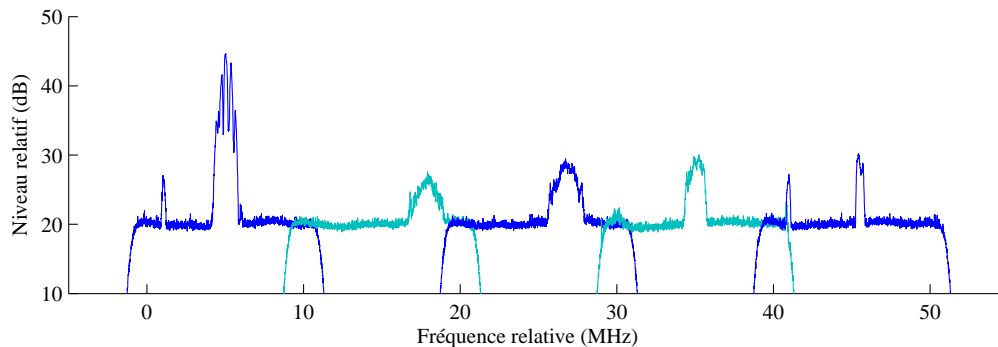


FIG. 3.2 – Radars observés dans la bande 1.300 GHz-1.375 GHz. Les fréquences exactes ne sont pas indiquées pour raison de confidentialité. La bande observée de 50 MHz est composée de la juxtaposition de 5 bandes de 10 MHz.

### 3.1.1 Contraintes

Le système doit s’adapter à la principale procédure d’observation existante qui consiste à effectuer du ON-OFF spatial. Le lobe principal du télescope suit la source à étudier dans le ciel pour intégrer son signal pendant quelques dizaines de secondes (séquence ON) puis, après un temps mort d’une quinzaine de secondes, il pointe à proximité de la source pendant la même durée (séquence OFF). La différence des deux séquences permet de retirer du signal les contributions instrumentales qui masquent le signal utile. La procédure est répétée jusqu’à l’obtention d’une durée d’observation suffisante (voir section 1.4.1.4, p.40). À ces séquences d’observation s’ajoutent systématiquement 3 séquences de calibration d’une seconde séparées du même temps consistant à injecter un bruit de niveau connu dans le cornet du récepteur qui continue d’observer le ciel et donc de recevoir des interférences.

Il faut donc que le détecteur atteigne un régime de fonctionnement stable en un temps très inférieur à la seconde et que les traitements soient terminés dans la demi-seconde suivant la fin d’une séquence pour permettre la reconfiguration partielle du système pour la séquence suivante.

### 3.1.2 Algorithme

L’algorithme, développé dans le but de détecter et supprimer des impulsions radar en temps réel sans informations *a priori*, propose une alternative à [35, 19, 75, 14, 15, 53]. Les différences majeures avec [35, 53] sont une estimation de seuil basée sur l’étude d’un  $\chi^2$  plutôt que par l’estimation de moyenne et variance sur une distribution gaussienne. Le matériel nécessaire est réduit de plus de moitié impliquant au final le traitement possible d’une plus large bande d’analyse. L’autre différence majeure est l’usage de “blanking” synchronisé sur les FFTs (voir section 3.1.2.1).

#### 3.1.2.1 SYNOPTIQUE

La détection des impulsions radar est effectuée sur la forme d’onde. Une détection dans le plan temps-fréquence n’est pas adaptée car la puissance de l’impulsion est alors

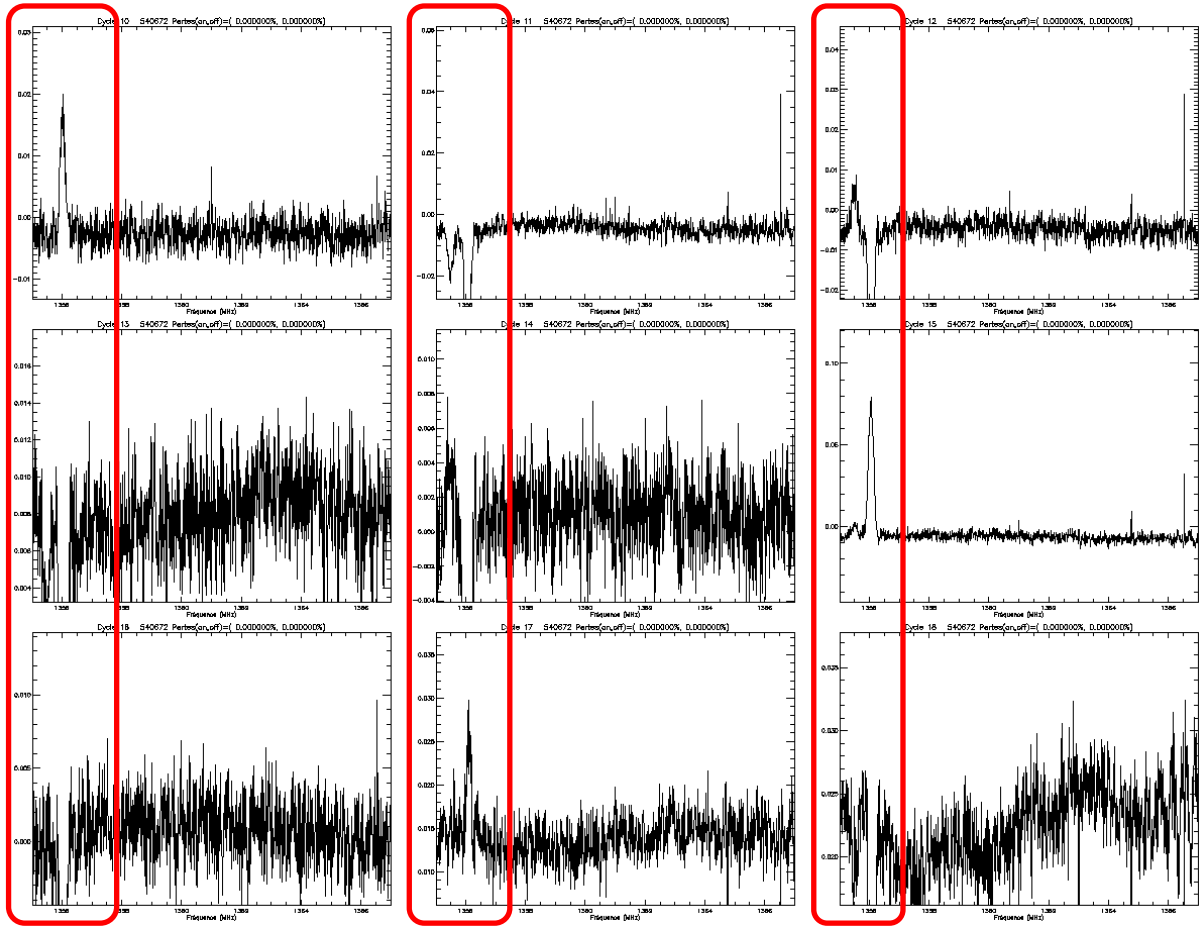


FIG. 3.3 – Résidus laissés par des impulsions radars après une séquence d’observation de type ON-OFF. L’échelle verticale est linéaire.

diluée sur plusieurs canaux fréquentiels. La base qui offre le plus de contraste est donc la représentation temporelle. Le “blankeur” radar est par conséquent inséré sur le flot de données de la forme d’onde complexe juste avant que l’estimation spectrale par FFT soit effectuée (figure 3.4). La forme d’onde est dupliquée par blocs de taille  $N$  pour effectuer le recouvrement précédant la FFT sur  $N$  canaux. Une copie de la forme d’onde est envoyée vers le détecteur d’impulsions qui prend la décision de “blanker” les blocs entiers situés autour de l’interférence. Les algorithmes de “blanking” visent à supprimer les points détectés comme aberrants mais il est commun d’éliminer aussi les données à proximité qui *semblent* intègres. En travaillant de cette manière, on augmente le nombre de points “blankés” mais on supprime des interférences faibles qui, isolées, n’auraient pas été correctement classifiées par notre détecteur non-idéal. Cette méthodologie est particulièrement applicable au traitement des impulsions radar qui subissent nécessairement une propagation multi-trajet [19].

Le “blanking” d’un tel volume de données pour de courtes impulsions radar ne se justifie pas pour l’unique raison d’élimination d’interférences. Il sert surtout à contrôler la position des discontinuités issues du “blanking” de parties de la forme d’onde. En effet,

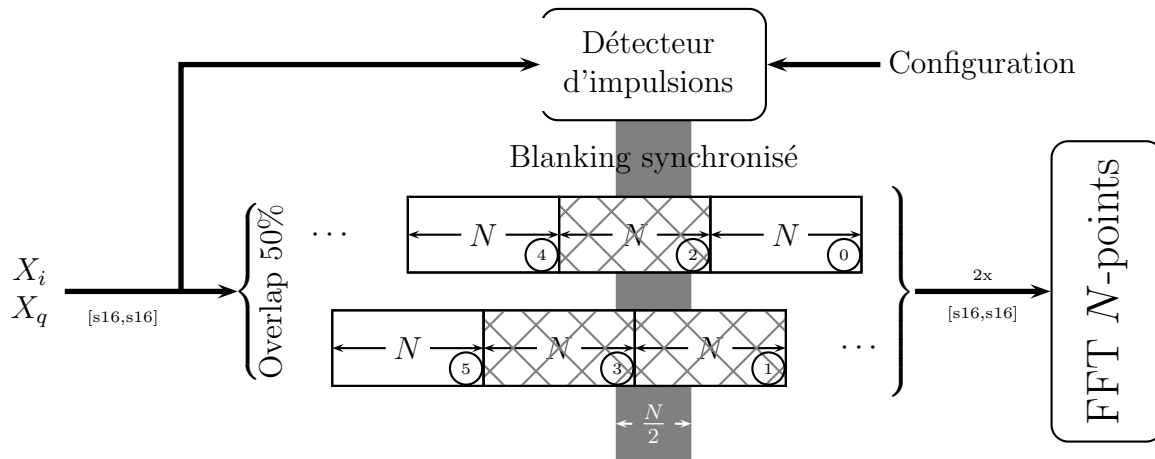


FIG. 3.4 – Synoptique du “blanqueur” radar

la FFT sur  $N$  points est précédée d'un fenêtrage par une fonction de Blackman-Harris qui limite l'effet de la discontinuité entre les points d'indice  $N - 1$  et 0 (pour la FFT, le vecteur représente une période d'un signal périodique) [36]. Si le “blanking” n'était pas effectué de façon synchronisée avec les FFTs, des discontinuités apparaîtraient également entre d'autres indices que  $N - 1$  et 0. Elles seraient non masquables par fenêtrage avant le calcul de la FFT et une distorsion de l'estimation spectrale apparaîtrait.

### 3.1.2.2 DÉTECTEUR D'IMPULSIONS

Le bloc “détecteur de d'impulsions” de la figure 3.4, décrit en figure 3.5, analyse la forme d'onde et commande le “blanking” de certaines zones de la forme d'onde *via* le signal “Blank”. L'analyse consiste d'abord en l'estimation robuste de la moyenne du signal (bloc ① : Estimateur de moyenne). La moyenne est estimée de façon récursive grâce à un filtre de type passe-bas IIR seuillé décrit en section 2.2.10.

On choisit  $\epsilon = 32/2^{16} = 4.88e^{-4}$  de façon à assurer une constante de temps d'établissement du système légèrement inférieure à la milliseconde (0.67 ms), soit la période typique des radars rencontrés.

Le signal  $\text{Freeze-}\mu_{lps}$  est issu du bloc ② qui effectue la détection des forts points aberrants. Le critère  $C$  du filtre passe-bas IIR seuillé est fixé à 4 en raison de la facilité à implanter cette multiplication en électronique (LLS(2)). La moyenne  $\mu_{lps}$  sous-estime  $\mu_{LPS}$  d'un facteur  $L(4)^{-1} = 0.898$  et la probabilité de fausse alarme  $P_{fa}$  vaut 2.75%. Un tel niveau de fausse alarme se manifeste grossièrement par la détection moyenne d'un point aberrant tous les 36 échantillons et chaque fausse alarme se traduit par une décision de “blanking” qui élimine un bloc de  $3N/2$  points. Pour  $N$  typiquement fixé à 2048, cela produit un “blanking” quasi-complet des données (En utilisant la loi binomiale, nous trouvons que la probabilité pour n'avoir aucun points aberrants parmi les  $3N/2$  échantillons, et ainsi ne pas “blanker” le bloc de données, vaut  $6.2644e^{-36}\%$ ). Pour pallier ce problème, trois résultats de seuillage successifs sont scrutés dans le registre à décalage vers la gauche  $\text{LRS}[0 \dots 2]$ . Si le seuil est dépassé consécutivement trois fois cela signifie qu'une impulsion d'une durée supérieure à  $3 \times \frac{1}{14 \text{ MHz}} = 210 \text{ ns}$  est apparue. La probabilité de fausse



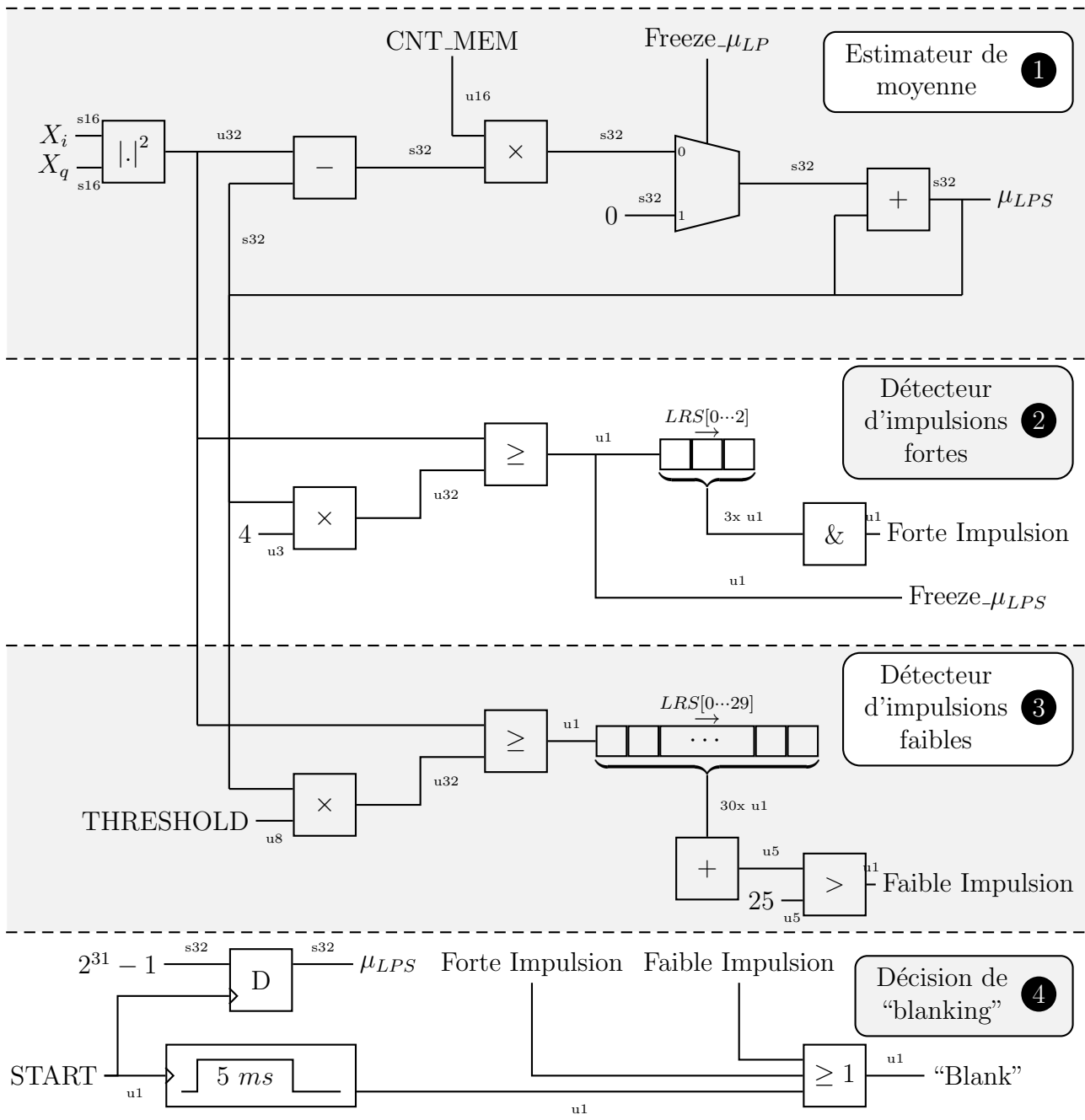


FIG. 3.5 – Schéma du détecteur d'impulsions radar

alarme pour ce type d'évènement vaut  $(2.75\%)^3 = 2.1e^{-3}\%$ . Les fausses alarmes ne sont maintenant responsables que de 6.4% de données "blankées" pour  $N = 2048$ .

Toutefois, l'utilisation du registre à décalage LRS n'améliore pas la détection des impulsions de faible niveau (figure 3.6). Le "blanking" des impulsions supérieures à  $4\mu_{LPS}$  améliore significativement la qualité des données observées mais de plus faibles impulsions sont encore présentes et perturbent encore les données.

Le bloc ③ "détecteur d'impulsions faibles" est voué à la détection de ce type d'interférence. La simple diminution du seuil de détection provoquerait une trop forte hausse

du pourcentage de fausses alarmes individuelles  $P_{fa}^i$ . Il faut néanmoins diminuer ce seuil de façon à aller chercher des impulsions faibles noyées dans le bruit. Nous exploitons la propriété d'ordre temporelle qu'ont les interférences : plutôt que de rechercher les points parasites un à un, nous allons étudier la détection de points parasites sur un plus long terme. Le principe du détecteur de impulsions fortes est étendu à la résolution du problème. Une séquence d'échantillons est seuillée et lorsque le nombre de détections dans la fenêtre d'étude devient trop important, on peut conclure que l'on se trouve en présence d'une faible impulsion. La fenêtre choisie couvre  $2 \mu s$  (la largeur typique d'une impulsion radar) soit  $L_{win} = 30$  échantillons. Il est décidé qu'une impulsion faible est présente dans la fenêtre quand le nombre de détection est supérieur à  $N_{win} = 25$ . Le seuil est fixé à  $\frac{26}{32}$  fois la moyenne.

Le calcul de la probabilité de fausse alarme globale  $P_{fa}^g$  se résume à l'étude d'une loi de probabilité binomiale où la probabilité pour que  $N_{win}$  soit égal à  $k$  est donnée par [55, p.139-140] :

$$Pr\{N_{win} = k\} = C_{L_{win}}^k P_{fa}^i{}^k (1 - P_{fa}^i)^{L_{win}-k} \quad k = 0, 1, 2, \dots, L_{win} \quad (3.1)$$

La fonction de répartition de cette loi, nécessaire pour obtenir la probabilité de fausse alarme globale  $P_{fa}^g = 1 - F(x)$ , est donnée par :

$$F(x) = \begin{cases} 0 & \text{si } x < 0; \\ 1 - I_{P_{fa}^i}(\lfloor x \rfloor + 1, L_{win} - \lfloor x \rfloor) & \text{si } 0 \leq x \leq L_{win}; \\ 1 & \text{si } x > L_{win}. \end{cases} \quad (3.2)$$

où  $I_p(a, b)$  est la fonction Béta incomplète normalisée définie pour  $0 \leq p \leq 1$  et  $a, b \in \mathbb{R}^+$  [55, p.319].

L'application numérique donne :

$$P_{fa}^g = 1 - F(25) \quad (3.3)$$

$$= I_{P_{fa}^i}(25 + 1, 30 - 25) \quad (3.4)$$

$$\text{avec } P_{fa}^i = 1 - F_{\chi_{\nu=2, \sigma}^2}\left(\frac{26}{32} \cdot \frac{\nu}{L(4)}\right) \quad (3.5)$$

$$P_{fa}^g = 1.3e^{-3\%} \quad (3.6)$$

$$(3.7)$$

Dans l'hypothèse où  $N = 2048$ , le pourcentage de données "blankées" à cause des fausses alarmes vaut 4.0%.

Pour finir, le bloc 4 concentre les décisions de "blanking" des différents blocs. Au démarrage de l'observation, la moyenne  $\mu_{LPS}$  est initialisée à la plus grande valeur stockable dans un registre u32. Son initialisation à zéro interdirait son établissement car `Freeze_μLP` serait systématiquement forcé à '1'. Le seuil est alors très fortement surestimé. Afin de préserver la qualité des observations au cours de l'établissement de  $\mu_{LPS}$ , une temporisation force le "blanking" des données pendant 5 ms.

### 3.1.2.3 TESTS ET IMPLÉMENTATION

Le taux de Vrais Positifs  $P_{VP}$  est donné en figure 3.6 pour divers niveaux d'interférences. Le détecteur d'impulsions fortes devient inefficace pour des interférences

d'INR inférieur à 5.8 dB. Il est alors épaulé par le détecteur d'impulsions faibles qui continue de fonctionner jusqu'à des INRs de 0 dB. En contrepartie, les impulsions brèves ( $< 2 \mu s$ ), même intenses, ne sont pas détectées. Les deux détecteurs se complètent donc.

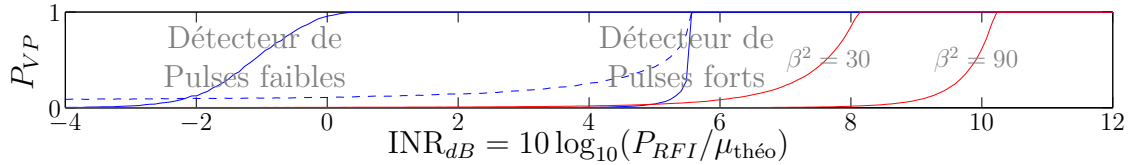


FIG. 3.6 –  $P_{VP}$  en fonction de  $INR_{dB}$  pour les détecteurs d'impulsions. La courbe pointillée bleue trace les performances du détecteur d'impulsions fortes sans utiliser de registre LSR. Les courbes rouges présentent les performances du détecteur décrit dans [35, 53] pour les valeurs de  $\beta$  indiquées.

Les détecteurs ont été comparés avec celui décrit dans [35, 53]. L'estimation robuste des valeurs de la moyenne et de la variance qui ont permis de construire le signal de test n'a pas été simulée. Ces valeurs, connues, ont été directement utilisée par la fonction de décision pour calculer un seuil  $S = \mu + \beta\sigma$ . Les impulsions détectables sont 2 à 4 dB plus intenses que ce que permet le détecteur d'impulsions fortes. L'utilisation de plus faibles valeurs pour  $\beta$  est limitée par un taux "blanking" trop élevé. La comparaison avec le détecteur d'impulsions faibles ne serait pas honnête car celui-ci utilise 30 échantillons pour prendre une décision alors que [35, 53] n'en considère qu'un seul à la fois. Néanmoins le document cité spécifie des durées d'impulsions radar de 200 à 40000 échantillons et cette simple information *a priori* n'est pas exploitée.

La figure 3.7 présente les courbes COR du détecteur d'impulsions faibles obtenues par simulation pour différents niveaux d'interférences. Les courbes pour des INR de 1.0 et 0.8 sont quasi-idéales. Par exemple, le détecteur écarte 100% des interférences d'un INR de 0 dB avec une probabilité de fausse alarme de  $10^{-3}\%$ . Son efficacité par rapport au détecteur d'impulsions fortes est illustrée par le "blanking" d'une série d'impulsions synthétiques en figure 3.8. Pour de plus faibles INR, le détecteur s'éloigne de l'idéal.

La simulation du "blanking" sur des impulsions synthétiques permet de confirmer les résultats précédents. La figure 3.8 présente un cas typique du fonctionnement du détecteur d'impulsions.

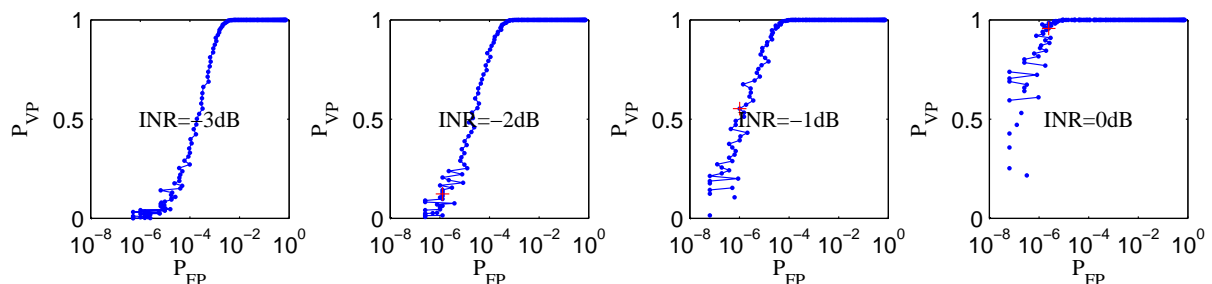


FIG. 3.7 – Courbes COR du détecteur d'impulsions faibles pour différents niveaux d'interférences. La croix rouge indique la position du critère choisi ( $C = \frac{26}{32} \frac{1}{L(4)}$ )

Le tableau 3.1 montre que les taux de fausses alarmes obtenus par simulation concordent avec les précédents calculs.

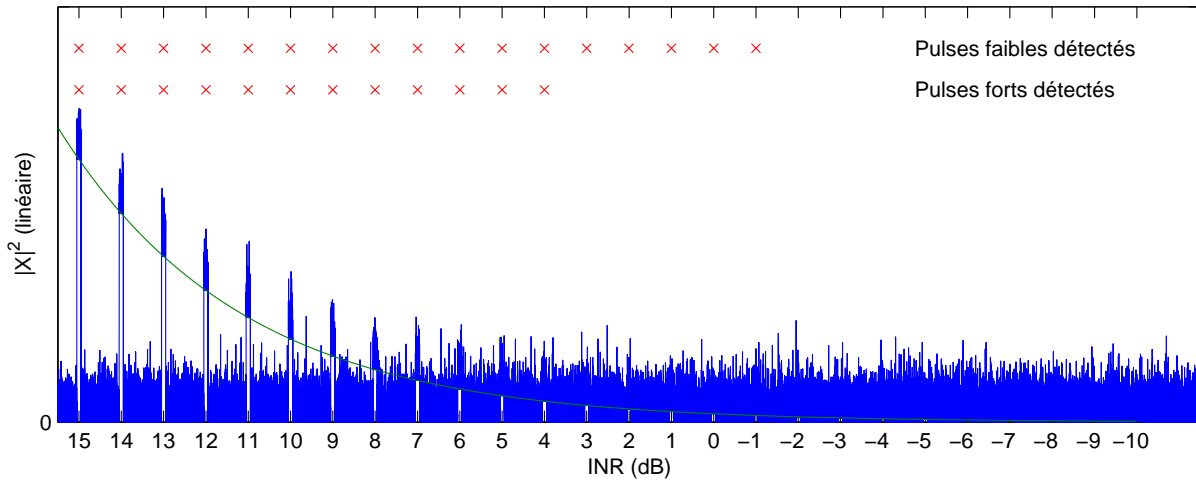


FIG. 3.8 – Simulation de la détection d’impulsions synthétiques par les deux détecteurs. Les fausses alarmes des détecteurs d’impulsions ont été retirées de la figure.

TAB. 3.1 – Comparaison des  $P_{fa}$  des détecteurs d’impulsions

Détecteur	Calcul analytique	Simulation	
	$P_{fa}$	$P_{fa}$	Taille d’échantillon
Impulsions fortes	$2.1 \times 10^{-3}\%$	$2.0963 \times 10^{-3}\%$	$13.5 \times 10^6$
Impulsions faibles	$1.3 \times 10^{-3}\%$	$1.3962 \times 10^{-3}\%$	$8 \times 10^6$

La figure 3.9 est le résultat d’une simulation de “blanking” sur des données réelles. Le détecteur d’impulsions fortes fonctionne dans la zone centrale de l’émission ( $|X_{SP}|^2$ ) mais le détecteur d’impulsions faibles retrouve d’autres impulsions de faible niveau ( $|X_{WP}|^2$ ). Néanmoins, ce détecteur n’élimine pas certaines impulsions fortes car elles sont trop brèves pour être détectées. La forme d’onde  $|X_{SP+WP}|^2$  issue du “blanking” contrôlé par les détections combinées est envoyée au module d’analyse spectrale.

Le rapport d’implantation hardware dans un FPGA 3 millions de portes (2v3000ff1152-5) est disponible en annexe A. Le détecteur a été utilisé sur des signaux de 14 MHz de bande mais il est techniquement possible de le faire fonctionner à 145.709 MEch/s. Le remplacement des multiplieurs implantés en cellules logiques par des multiplieurs câblés épargnerait l’usage de 776 bascules et 392 slices tout en offrant une vitesse de fonctionnement accrue. Ces multiplieurs ne sont pas disponibles dans l’application actuelle car ils sont utilisés pour l’estimation spectrale.

### 3.1.3 Résultats

Les résultats d’une partie des observations menées sur le NRT avec le “blanqueur” numérique sont présentés dans cette partie. Les sources visées sont listées dans le tableau 3.2.

La ligne de base de certains spectres obtenus a été estimée par un polynôme, puis soustraite. Les acquisitions ont été comparées avec d’autres observations effectuées à Arecibo, Green Bank ou Nançay dans des condition différentes. Mis à part l’observation de PGC 12382 présentée en figure 3.12, on ne note aucune distorsion notable. Les vitesses

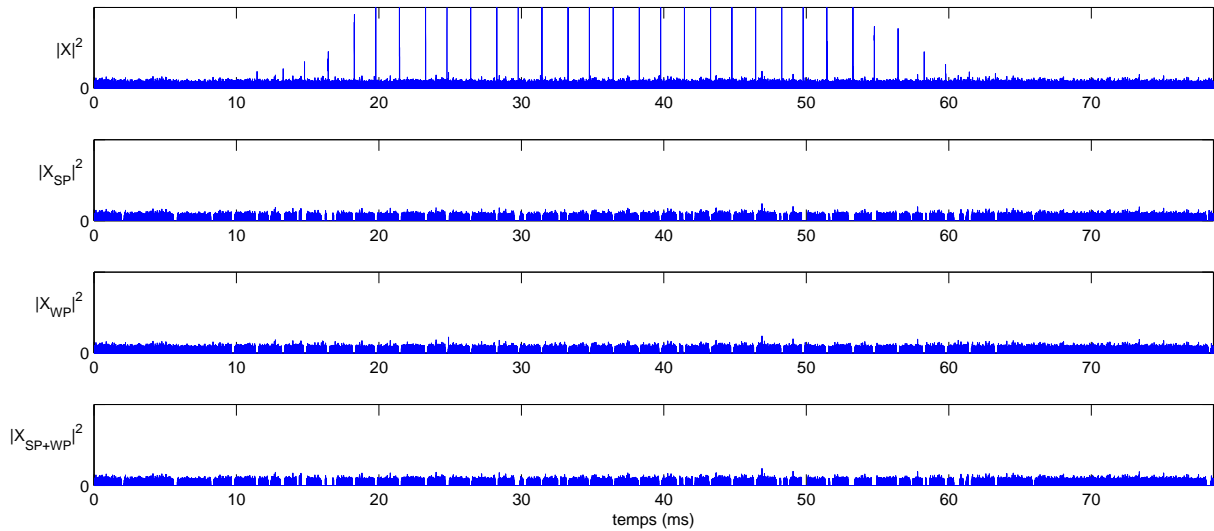


FIG. 3.9 – Simulation de la détection et du “blanking” d’impulsions radars issus de données réelles.  $|X_{SP}|^2$  correspond aux données traitées par le détecteur d’impulsions fortes seul.  $|X_{WP}|^2$  correspond aux données traitées par le détecteur d’impulsions faibles seul.  $|X_{SP+WP}|^2$  correspond aux données traitées par les efforts joints des deux détecteurs.

TAB. 3.2 – Sources observées avec le “blankeur radar”. Les positions sont données en coordonnées équatoriales référencées J2000.

Nom	Alpha	Déclinaison	z	$V_{opt}$ (km/s)	Fréquence (MHz)
ESO 474-G26	00h47m07,53s	-24d22m14,5s	0,053460	16027	1348,3
Malin I	12h36m59,3s	+14d19m49s	0,082557	24750	1312,1
PGC 12382	03h19m22,3s	+18d12m28s	0,044481	13335	1359,9
PGC 19534	06h44m58,08s	+35d18m07,8s	0,043854	13147	1360,7
PGC 51094	14h18m18,5s	+28d58m06s	0,042216	12656	1362,9

des sources sont indiquées sous convention optique.

L’observation de PGC 19534 (Figure 3.13) semble la plus représentative des capacités du “blankeur” numérique. La source restaurée est localisée exactement à la fréquence du radar le plus intense observé. Nous nous sommes assuré de ce résultat en effectuant une observation similaire hors source : la ligne de base ne laissait pas entrevoir de déviations soupçonnables.

### 3.1.4 Conclusion

Les limitations du “blankeur” sont atteintes pour des sources de niveau inférieur à 5 mJy, soit le niveau de Malin I (Figure 3.11). De plus longs temps d’intégration introduisent la présence de distorsions autour des fréquences radar. Ces distorsions ne peuvent être confondues avec des sources faibles mais la qualité des mesures pourrait être affectée.

Au NRT, les sources dont la densité de flux est 2-3 mJy peuvent être observée mais l’intérêt de la source ne vaut pas nécessairement le coût du temps d’occupation de l’ins-

trument. Pour une observation de surveillance, il sera rare de mener systématiquement de très longues intégrations. Par contre, si un astronome est demandeur d'informations sur une source en particulier, il est envisageable de bloquer l'instrument plus longtemps. Le "blankeur" permet de fournir des données valides pour ces utilisations.

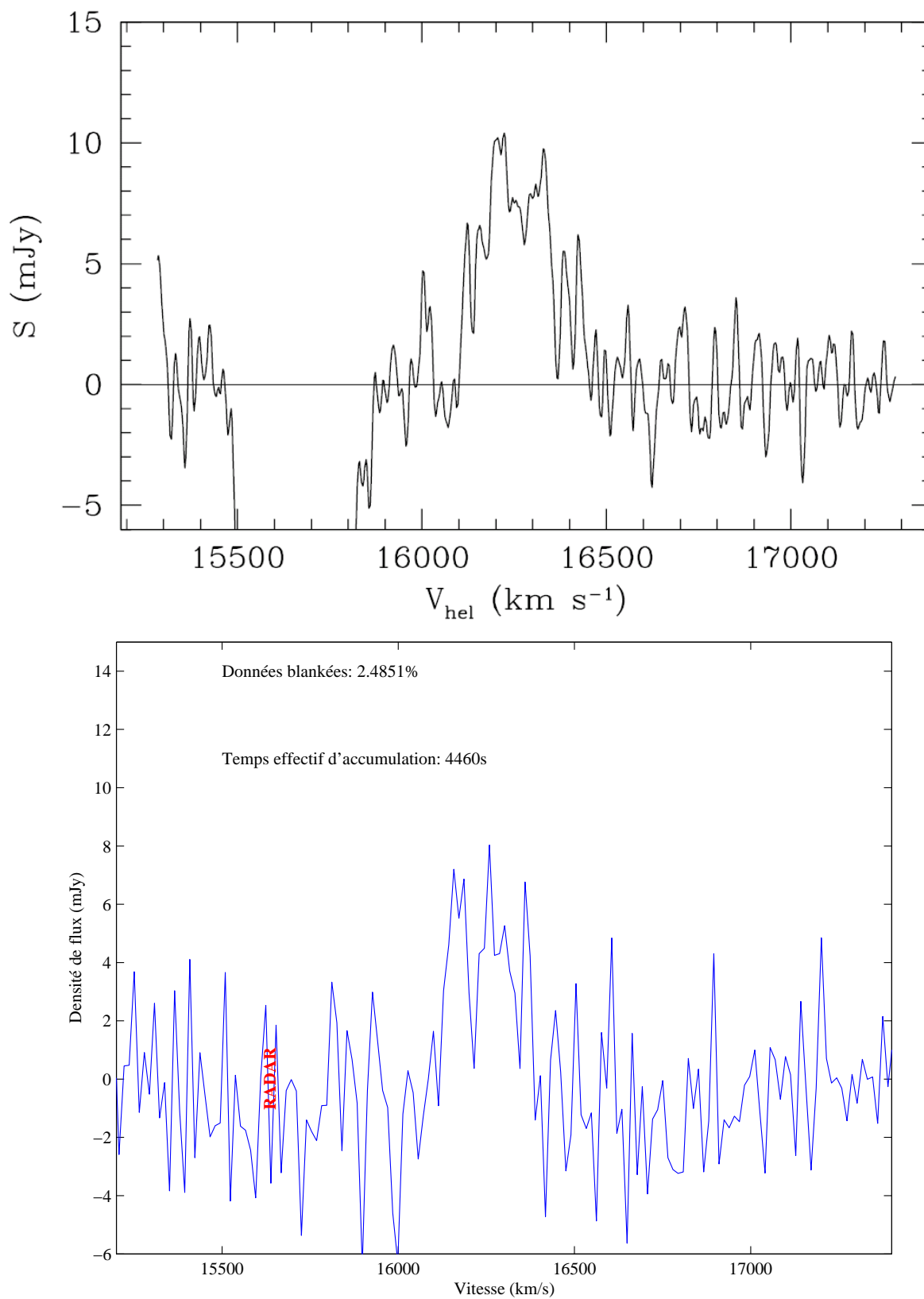


FIG. 3.10 – Raie HI de ESO 474-G26. Haut : observée au NRT sans élimination radar [60]. Bas : Observée au NRT avec élimination radar.

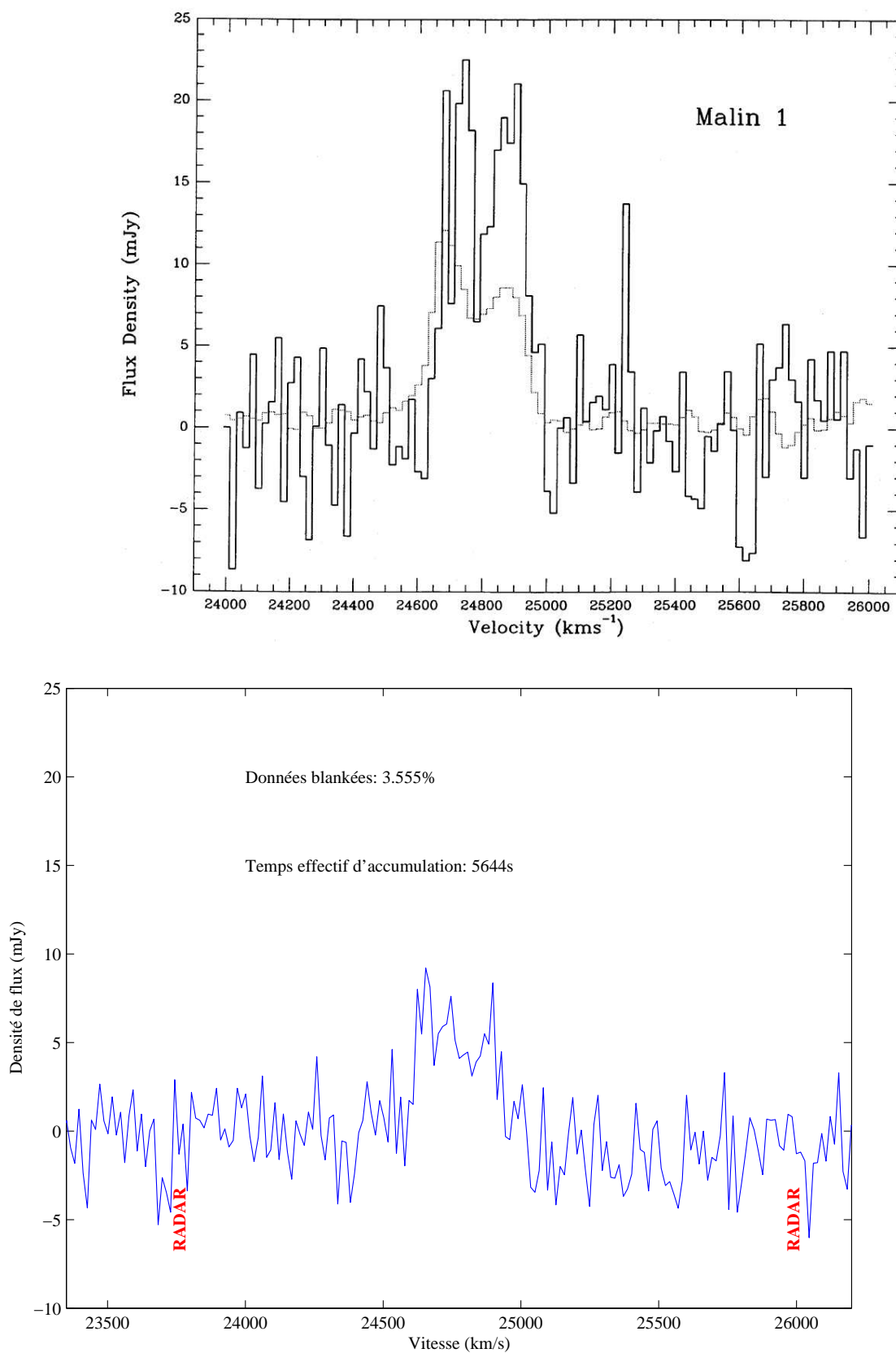


FIG. 3.11 – Raie HI de Malin I. Haut : observée à Green Bank (trait continu) et à Arecibo (trait pointillé)[39]. Bas : Observée au NRT avec élimination radar.



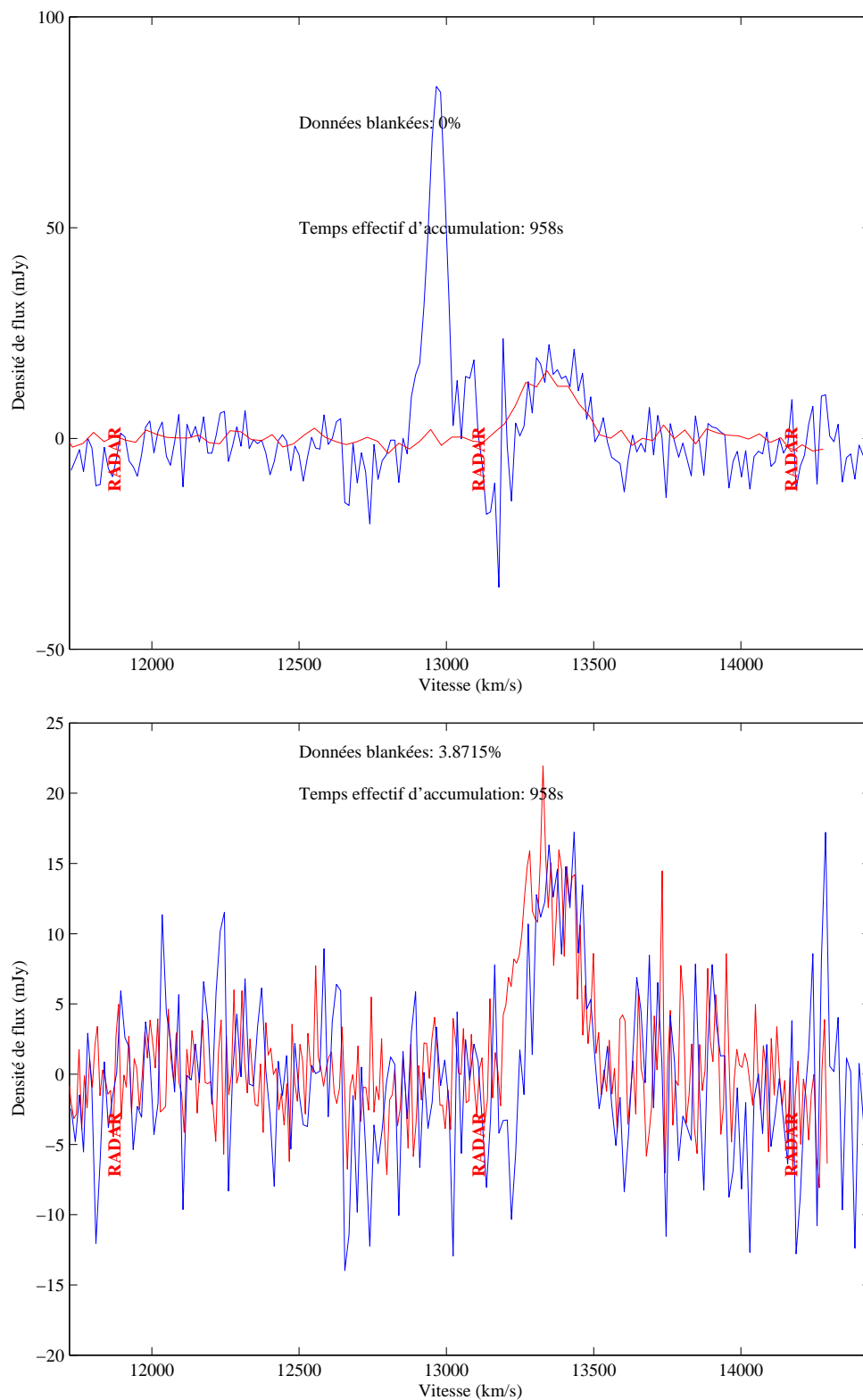


FIG. 3.12 – Raie HI de PGC 12382. Haut : observée au NRT sans élimination radar. Bas : Observée au NRT avec élimination radar. La courbe rouge observée à Arcibo est issue de [65, AGC/UGC:130166] [Cornell Digital HI Archive](#).

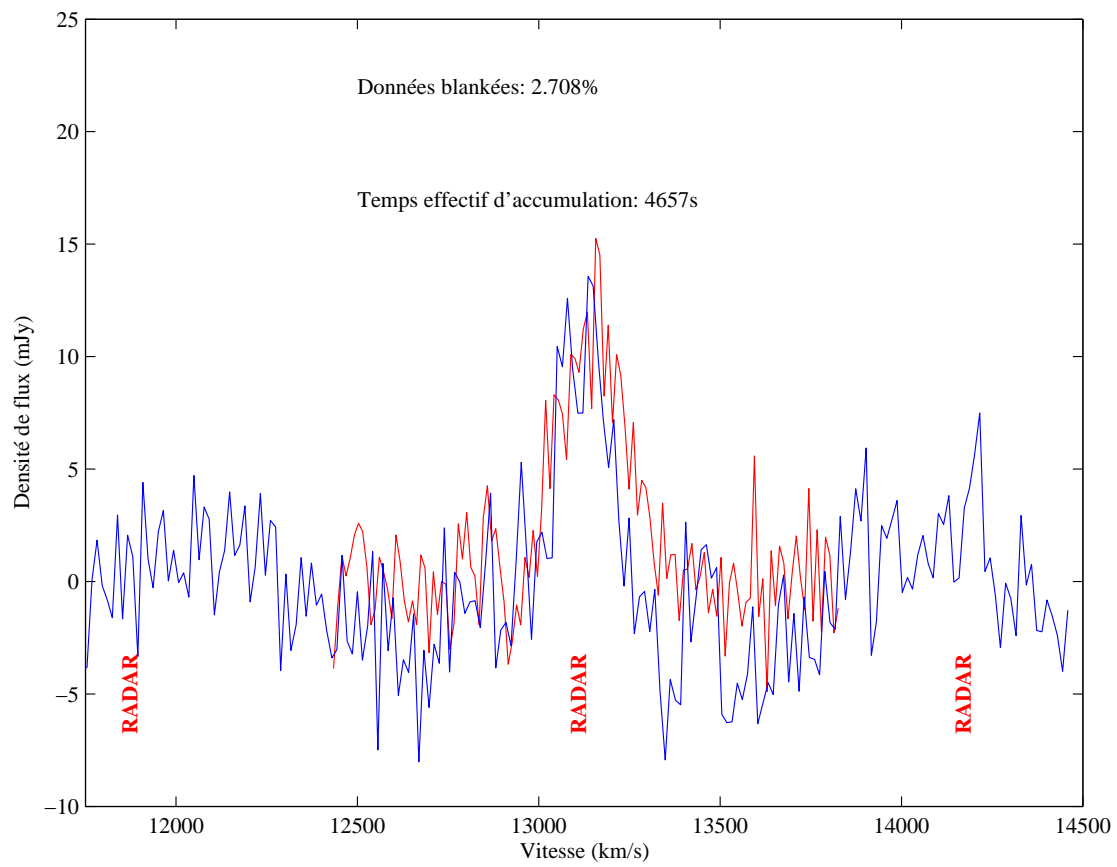


FIG. 3.13 – Raie HI de PGC 19534 observée au NRT avec élimination radar. La courbe rouge observée au NRT en l’absence de radar est issue de [HI Database](#) [54].

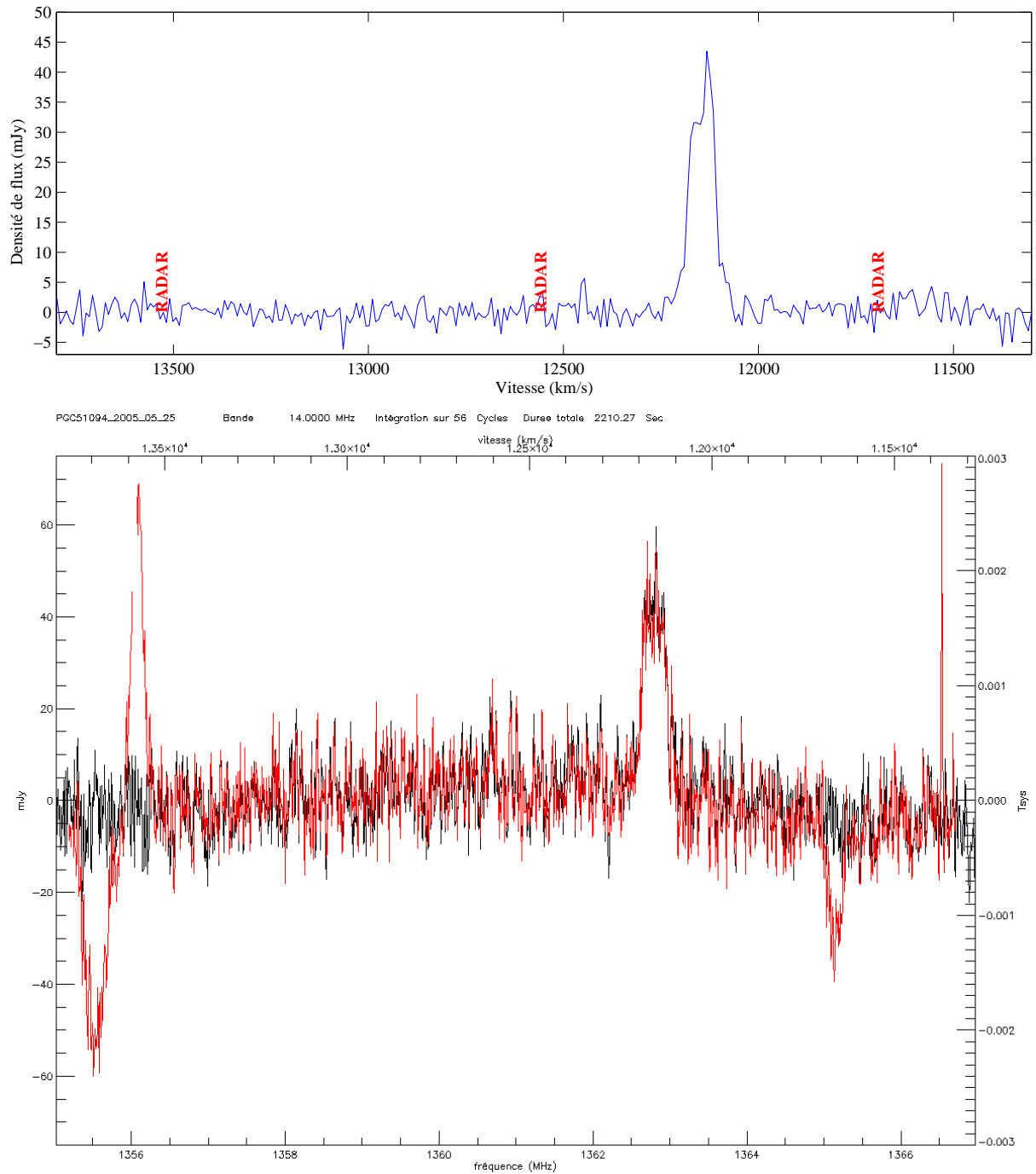


FIG. 3.14 – Raie HI de PGC 51094. Bas : Observée au NRT avec (noir) et sans (rouge) élimination radar. Haut : Observée à Arecibo [65, AGC/UGC:009155] Cornell Digital HI Archive.

## 3.2 Parasites à bande étroite intermittents : TFDMA au NRT

Une autre source d'interférences au NRT est rencontrée lors de l'observation des raies 1665 et 1667 MHz de sources OH redshiftées. Il s'agit de la bande 1613.8 - 1626.5 MHz allouée à la téléphonie satellitaire.

Le principal utilisateur est le consortium Iridium. Le système consiste en une constellation de 66 satellites LEO (+6 de réserve) circulant à une altitude de 780 km sur 6 plans orbitaux inclinés de  $86.4^\circ$  (passage approximatif aux pôles) offrant chacun une couverture de 4000 km. Chaque satellite n'est visible que 20 min par l'utilisateur à cause de la courte période orbitale spécifique aux satellites LEO ( $\approx 100$  min), mais un des 10 autres satellites de l'orbite prend en charge la suite du transfert de données. Quelque soit sa position sur le globe terrestre, un utilisateur est ainsi toujours en connexion avec au moins un satellite.

L'étude du système a été entamée en 1987 par Motorola. L'exploitation commerciale commence fin 1998 mais le groupe fait faillite en août 1999 à cause d'un trop faible nombre de clients. Le système, remis en service en 2001, est maintenant largement utilisé par le ministère de la défense américain ainsi que par des contrats commerciaux dans des domaines tels que le maritime, l'aviation, les gouvernements, l'industrie pétrolière, les sciences, ...

Les autorités de régulation des télécommunications ont attribué la bande 1621.35 - 1626.5 MHz à Iridium. Le partage de ce médium de communication est effectué par TFDMA (figure 3.15). Les 5.15 MHz de bande sont répartis en 120 canaux de 31.5 kHz de largeur séparés par 10.17 kHz de garde. Chaque canal est divisé en trames de 90 ms de long comprenant 9 blocs de 8.28 ms (un bloc d'entête, 4 blocs utilisateurs pour communications montantes et 4 blocs utilisateurs pour communications descendantes). Le temps restant est réparti entre les temps de garde. Les trames montantes ne sont captées par le radiotélescope qu'en présence et à proximité d'un téléphone satellitaire, situation rendue exceptionnelle grâce au silence radio requis à proximité des installations de radioastronomie et au faible nombre actuel d'utilisateurs. Les trames descendantes sont la principale source d'interférences car à tout moment des satellites sont vus par l'antenne.

En théorie, le système peut émettre des trames descendantes pendant 46% du temps. Les observations courantes montrent que cette valeur est rarement atteinte (Les statistiques à ce sujet ne sont pas fournies car elles sont propriétés de l'opérateur).

Le niveau des interférences peut dépasser le bruit système du télescope de 35 dB bien que le satellite soit au dehors du lobe principal de l'antenne. Cela correspond à un flux de 50000 Jy alors que les mégamasers observées dans cette bande ne dépassent jamais le Jansky bien qu'ils soient parmi les sources des raies les plus intenses observées au NRT.

L'élimination de ce type d'interférences TFDMA a été traitée par Vincent CLERC [11, p.189-195]. Les plans temps-fréquence parasités étaient enregistrés sur la machine d'acquisition, puis les échantillons dont l'écart type dépassait de  $2\sigma$  le bruit théorique étaient "blankés" en temps différé avant d'effectuer l'intégration qui permettait d'obtenir le profil de la source III Zw 35 (figure 3.16 et 3.17). Néanmoins, rien n'est dit sur l'adaptabilité de l'algorithme au niveau de bruit, ni sur les imperfections de blancheur de la bande. Quelques améliorations doivent être ajoutées pour permettre un fonctionnement en routine du système.

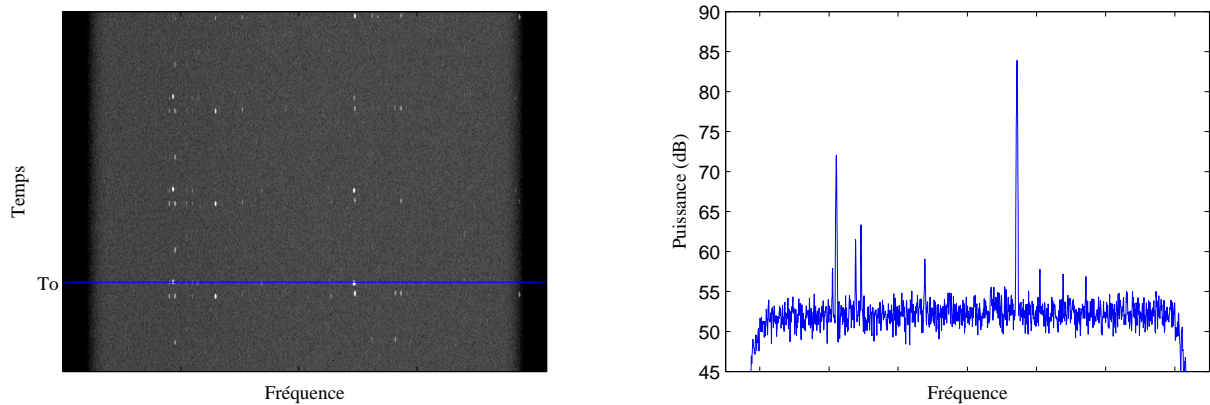


FIG. 3.15 – Émissions enregistrées au NRT dans la bande de téléphonie mobile par satellite. Gauche : plan temps-fréquence typique des interférences. Droite : spectre lors de l’émission de blocs de données (à  $t = T_0$ ). Les graduations sont masquées pour cause de confidentialité.

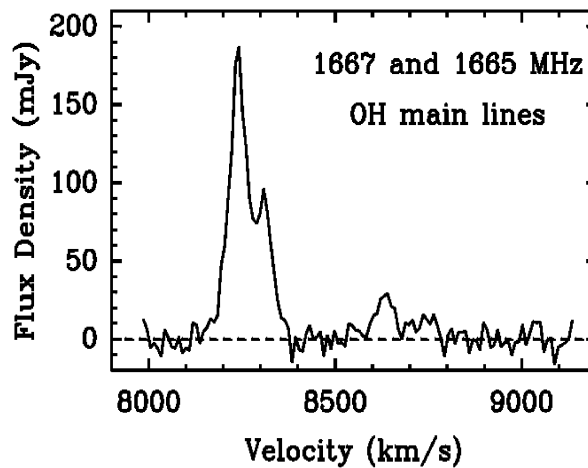


FIG. 3.16 – MégaMaser III Zw35 en 1987 au NRT avant la mise en service d’Iridium [51]

### 3.2.1 Contraintes

Tout comme le “blankeur” radar, le système doit s’adapter à la principale procédure d’observation existante qui consiste à effectuer du ON-OFF spatial (section 3.1.1).

Il faut donc que le détecteur atteigne un régime de fonctionnement stable en un temps inférieur à la seconde et que les traitements soient terminés dans la demi-seconde suivant la fin d’une séquence pour permettre la reconfiguration partielle du système pour la séquence suivante.

De plus, les seuils de “blanking” doivent être ajustés automatiquement pour s’affranchir des variations de niveau en temps et en fréquence. D’une part, la température système du récepteur est fonction de la position des miroirs. Au cours du pointage d’une source –et d’une source à l’autre– la puissance du bruit système varie donc légèrement. D’autre part, la chaîne analogique est construite sur des composants électroniques de bonne qualité mais dont les inévitables tolérances éloignent le système de l’idéal. Cela se caractérise en particulier par un imparfaite blancheur de bruit. La bande d’analyse est le siège d’ondulations

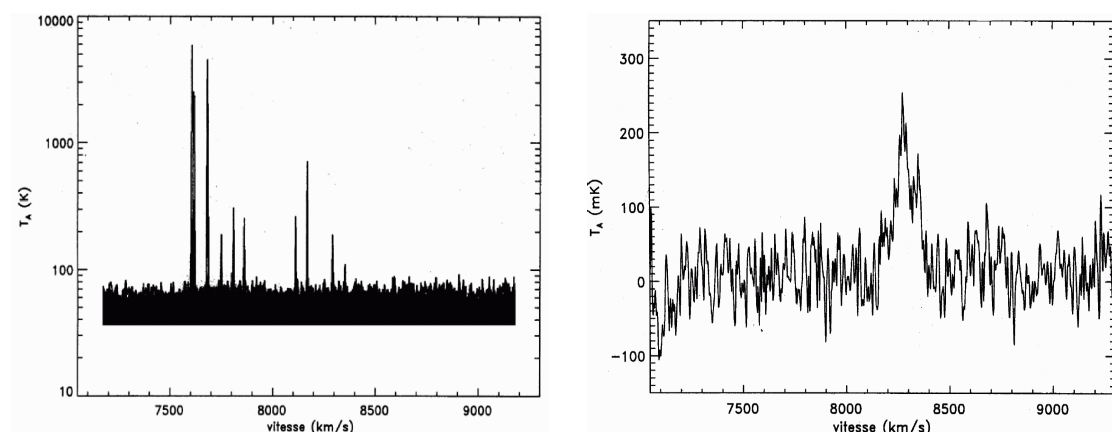


FIG. 3.17 – MégaMaser III Zw 35. Gauche : inobservable par méthode classique depuis la mise en service d’Iridium à cause d’interférences 10000 fois plus puissantes que le bruit système. Droite : ré-observé après l’élimination en temps différé des parasites (960 s ON - 960 s OFF, 80% de taux d’observation, 45 min de post-traitement avec un PII à 450 MHz) [63, 11]

d’amplitude de 0.5 dB. Ces variations sont produites par la combinaison des fonctions de transfert des filtres insérés sur la chaîne analogique. Une éventuelle calibration serait rendue ardue à cause de la présence des étages de changement de fréquence. C’est pour pallier ce type d’imperfection que la procédure de ON-OFF est prévue.

Néanmoins, le “blanking” s’effectue sur des séquences ON et des séquences OFF individuelles. La correction apportée par le ON-OFF n’est pas disponible au moment du “blanking”. Il faut donc que le détecteur fonctionne dans ces conditions dégradées.

### 3.2.2 Algorithme

L’estimation spectrale est configurée pour fournir des plans temps-fréquence suffisamment fins pour résoudre les blocs de données Iridium. Les paramètres choisis sont 7 MHz de bande analysée en 2048 canaux pour fournir une résolution fréquentielle de 3.4 kHz et un facteur d’accumulation de  $2^4$  –soit un degré de liberté de 32– pour produire des spectres à la résolution temporelle de 2.34 ms. Le débit de ces données est de 3.5 Mo/s par voie. Les traitements sont effectués en temps réel sur un DSP TMS320C6203 cadencé à 300 MHz.

Dans cette partie, 200 canaux à chaque extrémité de la bande seront écartés des traitements car ce sont des zones fortement atténuées par les filtres analogiques et numériques du système. Ces données seront systématiquement laissées telles quelles. A charge, lors du dépouillement des données, de limiter la bande à la zone d’intérêt des 1648 canaux centraux.

Une acquisition débutera par une phase de calibration qui visera à estimer du mieux possible un seuil de “blanking”. Ensuite, cette information sera utilisée pour “blanker” les données spectre par spectre jusqu’à ce que la séquence d’observation se termine.

### 3.2.2.1 CALCUL DE SEUIL

Dans un premier temps, il a été envisagé d'estimer un seuil scalaire pour toute la bande. Cette méthode s'est révélée décevante à cause des ondulations de bande décrites dans la partie précédente (figure 3.18). Pour une  $P_{fa}$  fixée, le taux de "blanking" effectif était hétérogène sur la bande. Les parties de niveau supérieur à la moyenne présentaient un taux de "blanking" élevé, soit une variance plus forte après accumulation. Les parties de niveau inférieur à la moyenne présentaient un taux de "blanking" très faible, ce qui n'est pas un problème en soi, mais étaient en contrepartie plus sujettes aux interférences de faible niveau.

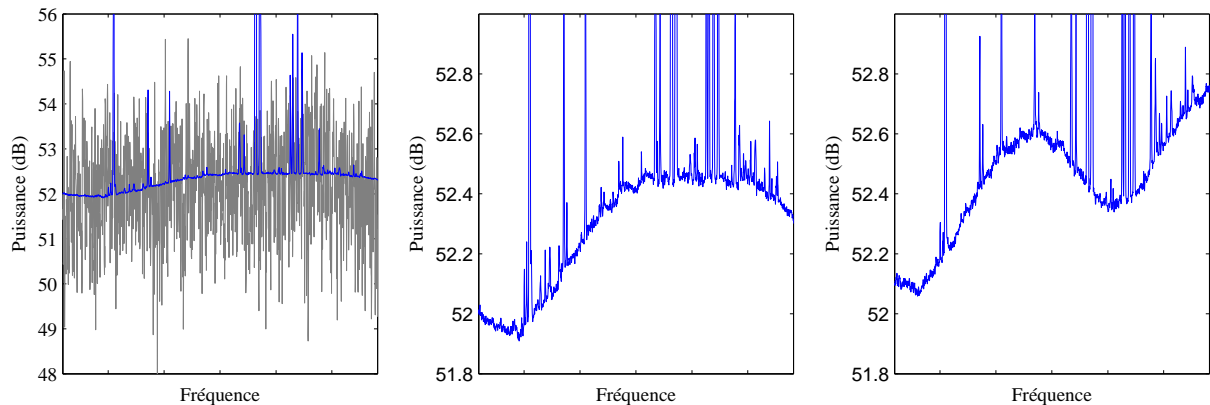


FIG. 3.18 – Blancher imparfaite de la bande d'analyse : la non blancher du bruit système est difficilement visible sur un spectre élémentaire (à gauche en gris) mais une fois l'écart type du bruit réduit par intégration de 12200 spectres durant 29 s, l'ondulation de la bande est flagrante (à gauche en bleu et au milieu). Deux intégrations effectuées au même instant mais sur deux polarisations linéaires différentes (milieu et droite) montrent l'effet sur la bande de deux chaînes analogiques de conception identique mais physiquement différentes. Les points aberrants de fortes valeurs sont imputables à l'apparition de signaux TFDMA pendant l'intégration.

Il a donc été décidé de substituer le seuil scalaire  $S$  par un seuil vectoriel  $S[k]$ . La solution prenant la forme d'une interpolation pour approcher la bande en n'estimant des seuils qu'en quelques points de la bande a été abandonnée à cause du degré trop élevé du polynôme à estimer en virgule fixe ainsi qu'à cause de la présence de forts points aberrants. Les seuils vont donc être estimés individuellement sur chaque canal.

Les 128 premiers spectres de chaque séquence sont utilisés pour estimer le seuil de "blanking". Tous ces points sont systématiquement "blankés" pour interdire l'entrée dans la chaîne d'inévitables points aberrants. Cela correspond à 150 ms de signal, soit une durée suffisamment courte pour conserver suffisamment de signal pendant les séquences de calibration de durée d'une seconde.

À la vue de la faible occupation d'un canal au cours du temps, une méthode de type médiane ou une de ses dérivées est utilisée (figure 2.42)). Les estimateurs sont comparés dans le tableau 3.3.

L'estimateur de moyenne par moyenne tronquée symétrique est choisie pour son faible biais et sa faible variance.

TAB. 3.3 – Comparaison des estimateurs de moyenne envisagés sur un 128-échantillon suivant une loi du  $\chi^2$  à 32 degrés de liberté, corrompu par un signal TDMA

	Moyenne arithmétique	Médiane	$\mu_{MD_{\nu,N}}$	$\mu_{MFS_{20\%}}$
Moyenne de l'estimateur	32	31.2668	32.0000	32.0055
Variance de l'estimateur	0.5	0.7615	0.7977	0.0113

Cet estimateur implique le classement de 1648 128-échantillons. Les algorithmes de tri les plus classiques ont été comparés (tableau 3.4). La méthode de tri bulle est présentée ici dans le seul but d'exposer son inefficacité. La méthode de classement par insertion directe [58, 8.1], éventuellement optimisée par la méthode de Shell, est la plus simple pour résoudre le problème, mais d'autres méthodes sont plus optimisées pour travailler sur des échantillons de taille supérieure à 50 comme c'est le cas de "Quicksort" et "Heapsort" [58, sec. 8.2 et 8.3]. Dans certains cas, "Heapsort" peut être plus rapide que "Quicksort" mais il présente l'inconvénient de nécessiter une zone mémoire additionnelle qui est une ressource précieuse dans un système embarqué.

TAB. 3.4 – Comparaison de la complexité des algorithmes de classement. L'utilisation de l'applet Java de [http://interstices.info/display.jsp?id=c\\_6973](http://interstices.info/display.jsp?id=c_6973) a permis de faire quelques simulations comparatives de 100 classements de 128 spectres à 1628 canaux.

	tri bulle	Insertion directe	Méthode de Shell	Quicksort	Heapsort
Complexité moyenne	$O(N^2)$	$O(N^2/4)$	$O(N^{1.25})$	$O(1.4N \log_2(N))$	$O(N \log_2(N))$
Stockage	$O(N)$	$O(N)$	$O(N)$	$O(N)$	$O(2 \log_2(N))$
Nombre de comparaison	$1.3 \times 10^8$	$6.8 \times 10^7$	-	$1.4 \times 10^7$	-
Nombre de copies	$7.0 \times 10^7$	$7.0 \times 10^7$	-	$7.8 \times 10^6$	-
Temps de calcul (indicatif)	4.90 s	1.94 s	-	0.58 s	-

Une spécificité du récepteur RDH propre aux systèmes temps réel est exploitée pour accélérer encore le classement. Plutôt que d'attendre la réception complète des données avant de commencer le classement, il est possible de faire le tri au cours de l'envoi. Cela équivaut un peu à effectuer la méthode de "Shell", une variante de l'insertion directe qui consiste à effectuer des pré-classements sur des sous-ensembles de l'échantillon avant d'effectuer le classement final. Les 128 spectres sont envoyés par 4 au DSP tout au long des 150 ms. À chaque fois qu'un bloc de 4 spectres est reçu, les 4 nouveaux points de chaque canal sont classés puis insérés dans les spectres déjà traités.

Une autre subtilité consiste à insérer les 4 points en même temps plutôt que l'un après l'autre. Le code de "Numerical Recipes" listé ci-après (table 3.5) est appelé 31 fois ( $128/4 - 1$ ) avec  $n = \{8, 12, \dots, 128\}$ . À chaque appel, 4 points sont à classer. Donc,



par quatre fois, une partie plus ou moins importante du vecteur est décalée d'une case mémoire vers les indices forts.

TAB. 3.5 – Code C du classement par insertion directe [58, 8.1]. Le code a été modifié par rapport à la version d'origine pour faire commencer les indices de tableau à 0 .

```

1 void piksrt(int n, float arr []) {
2     int i, j;
3     float a;
4     // Pick out each element in turn.
5     for (j=1; j<n; j++) {
6         a=arr [ j ];
7         i=j-1;
8         // Look for the place to insert it.
9         while ( i >= 0 && arr [ i ] > a ) {
10            arr [ i+1]=arr [ i ];
11            i--;
12        }
13        // Insert it.
14        arr [ i+1]=a;
15    }
16 }
```

Dans la nouvelle version de classement (listing de la table 3.6), les 4 points sont pré-classés (ligne 4) puis ils sont insérés en décalant une partie du vecteur de 4 cases mémoire lors de l'insertion du plus fort point pré-classé (ligne 15), puis la partie suivante de 3 cases pour le deuxième plus fort point, puis 2, puis 1 case. Au final, le vecteur complet n'aura été partiellement recopié qu'une fois lors de l'appel de la fonction au lieu de quatre pour la méthode précédente.

La suite de l'estimation de moyenne par moyenne tronquée symétrique consiste à sommer les échantillons centraux du vecteur pour chaque canal. Une troncature à 20% impose  $\delta' = 25.7$  et  $\delta'' = 39.3$  (section 2.2.6). Dans chaque canal, les points issus des 128 spectres sont classés dans un tableau indexé de 0 à 127. Les points d'indice 28 à 106 sont moyennés pour obtenir  $\hat{\mu}_{MTS_{20\%}} [k]$ .

### 3.2.2.2 “BLANKING” DES SPECTRES

Le plan temps fréquence est traité spectre par spectre. Chaque canal est comparé à un élément du vecteur de seuils. Quand un canal semble aberrant, il est mis à zéro (“blanké”) ainsi que ses deux proches voisins. La probabilité de fausse alarme globale est fixée à 5%. En l'absence d'interférences, deux points marqués par le seuillage ont peu de chance d'être contigus. On peut alors, sans faire trop d'erreur, considérer la valeur de la  $P_{fa}$  du seuillage comme égale à  $\frac{5\%}{3}$ . La probabilité d'avoir deux points aberrants contigus vaut donc  $(\frac{5\%}{3})^2 = 0.27\%$ , ce qui confirme l'hypothèse de départ. Le seuil est calculé

TAB. 3.6 – Code C du classement progressif d'un N-échantillon.

```

1 void my_sorting(int n, float arr [], float C[]) {
2     int i, j, k;
3     float a;
4     // Presorting of 4 new samples recorded in table C[]
5     for (j=1; j<4; j++){
6         a=C[j];
7         i=j-1;
8         while (i>=0 && C[i]>temp) {
9             C[i+1] = C[i];
10            i--;
11        }
12        C[i+1]=a;
13    }
14
15    // One by one insertion for each of the 4 samples
16    i=n-1;
17    for (j=n; j<n+4; j++){
18        k = 4-(j-n);
19        a=C[k-1];
20        while (i>=0 && arr[i]>temp) {
21            // Move a subpart of C[] by k = 4 memory cells
22            // then by 3, 2 and 1
23            datar[i+k] = datar[i];
24            i--;
25        }
26        datar[i+k]=a;
27    }
28 }

```

comme suit :

$$S[k] = \frac{\hat{\mu}_{MTS_{20\%}}[k]}{\nu} F_{\chi^2_{\nu, \sigma}}^{-1} \left( \frac{5\%}{3} \right) \Big|_{\nu=32} \quad (3.8)$$

$$S[k] = 1.6029 \hat{\mu}_{MTS_{20\%}}[k] \quad (3.9)$$

Les plans temps-fréquence complets sont écrits sur disque. Sans “blanking”, il est inconcevable, sauf grosse erreur de configuration du système, que des échantillons nuls soient présents dans la bande. A la fin de l’observation, les spectres dynamiques sont moyennés temporellement tout en rejetant les points nuls correspondant à des zones “blankées”. Chaque canal est donc issu de la moyenne d’un échantillon de taille variable. L’écart type de l’estimateur de moyenne n’est donc pas constant sur la bande. Cette information est connue de l’observateur mais pour des taux de “blanking” faibles, la dispersion de l’écart type entre les canaux est négligeable. Par exemple, pour un taux de “blanking” dans un

canal de 10% (5% de fausses alarmes et 5% d'interférences), l'écart type augmente de 5.4%.

L'accumulation pourrait être facilement effectuée en temps réel, pendant l'acquisition, mais le but principal de ce travail était de montrer la faisabilité de l'atténuation temps réel d'Iridium. Cette évolution permettrait néanmoins de réduire les temps de transfert de données en fin d'observation.

### 3.2.3 Résultats

La seule source que nous ayons observée est III Zw 35, un mégamaser situé à  $380 \times 10^6$  années-lumière de notre planète, mais si intense qu'il est possible de l'observer en quelques minutes de pose. C'est donc le candidat idéal pour nos tests (tableau 3.7). Le faible nombre de candidats aux observations de test s'explique par la difficulté de trouver des sources localisées dans cette étroite bande parasitée.

TAB. 3.7 – Source observée avec le “blankeur” Iridium. Les positions sont données en coordonnées équatoriales référencées J2000.

Nom	Alpha	Déclinaison	$z$	$V_{opt}$ (km/s)	Fréquence (MHz)
III Zw 35	01h44m30.56s	+17d06m05.0s	0.027436	8225	1622.8

Plusieurs observations de la source ont été effectuées. La première, le 8 janvier 2004 (figure 3.19), a permis d'observer la raie 1667 MHz décalée vers le rouge.

Une des dernières observations, effectuée le 24 février 2005 (figure 3.20), permet par une plus longue pose de distinguer maintenant les deux raies OH à 1665 MHz et 1667 MHz. La probabilité de fausse alarme a été fixée très basse par erreur. On observe certains pics produits par un “blanking” imparfait mais la source est bien discernable.

### 3.2.4 Conclusion

Cette méthode simple permet de réobserver des sources situées dans la bande de l'opérateur Iridium.

Les observations actuelles ont montré que le taux de couverture du spectre était plus faible que ce que l'opérateur pourrait réellement se permettre. Cela nous a permis d'utiliser un estimateur dont la robustesse restait faible face un fort taux d'interférences. Si l'utilisation de la bande devait augmenter, un autre estimateur devra être utilisé. L'estimateur de moyenne par moyenne de classe semble être un bon candidat.

L'atténuation des interférences dépend d'abord de la capacité du système à les résoudre. Le système actuel est limité par le DSP qui effectue l'algorithme de “blanking”. L'implantation en FPGA n'est pas aisée à cause du tri nécessaire au calcul du seuil mais elle permettrait d'analyser une bande plus large dans le même temps. Cette migration pourrait être motivée par l'utilisation de l'estimateur par moyenne de classe qui est plus facilement implantable sur une telle plateforme.

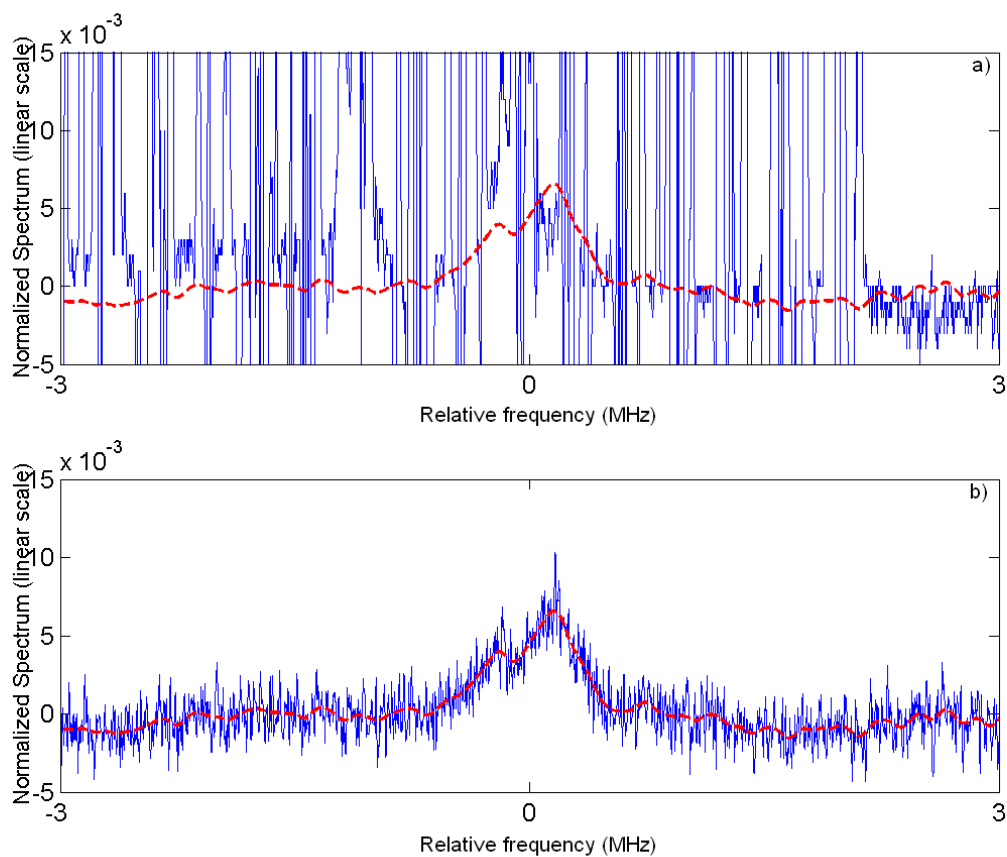


FIG. 3.19 – Réobservation le 8 janvier 2004 de III Zw 35 en temps réel après 14 minutes d'intégration au NRT : (a) Sans "blanking". L'axe vertical est ajusté de façon à rendre visible le profil attendu de la source. (b) Avec "blanking" (3.5% de données éliminées). La source est de nouveau clairement visible.

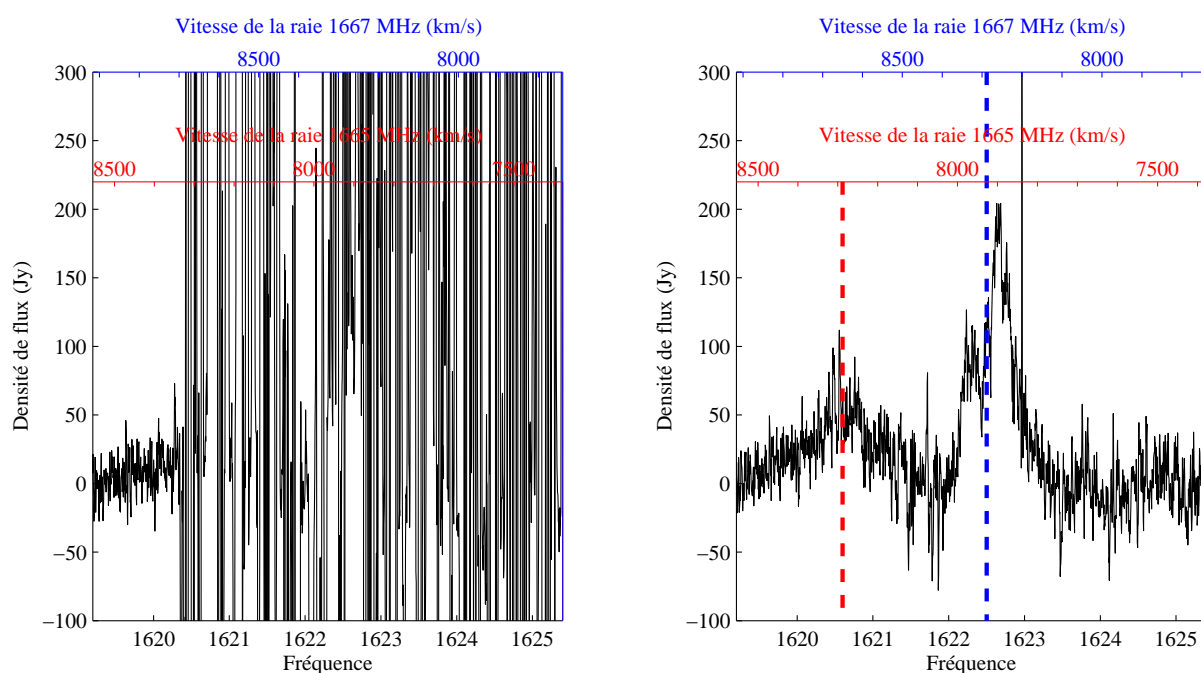


FIG. 3.20 – Réobservation le 24 février 2005 de III Zw 35 en temps réel après 29 minutes d'intégration au NRT : Gauche : Sans "blinking". Droite : Avec "blinking" (0.55% de données éliminées). Les deux raies OH sont clairement visibles. La très faible probabilité de fausse alarme choisie permet tout de même d'observer mais on visualise quelques distortions.

### 3.3 Parasites bandes étroites continus: AM au NDA

Dans le cadre des observations en routine du Soleil, le réseau décimétrique de Nançay fournit une partie des données mises à disposition sur [BASS2000](#). Les données consistent en des plans temps-fréquence couvrant la bande 20-80 MHz au rythme d'un spectre par seconde (polarisation droite et gauche alternées toutes les 494 ms).

Il est prévu que les observations Joviennes subissent le même sort. La bande observée sur cette source est limitée à 10-40 MHz.

L'observation de ces deux astres est effectuée 4 heures de part et d'autre du moment de leur passage au méridien. Néanmoins, le jeu de rotation de la Terre et de Jupiter autour du Soleil implique un décalage relatif des passages au méridien de près de 4 minutes. Les plages d'observation se chevauchent donc de temps en temps et ne permettent pas l'enregistrement complet des signaux issus des deux sources. Dans ce cas, la priorité d'étude est fixée en tenant compte des émissions probables calculées de Jupiter et du niveau d'activité du Soleil observé lors des jours précédents.

Les signaux issus du réseau polarisé circulaire gauche et du réseau polarisé circulaire droit sont alternativement enregistrés par un analyseur de spectre du commerce contrôlé par PC *via* une interface GPIB. Le principe de fonctionnement de l'analyseur est basé sur un filtrage hétérodyne : la bande totale est analysée point à point dans une sous-bande de 30 kHz. Les paramètres sont fixés à 400 points de mesure balayés en 350 ms dans la bande d'étude (50 MHz pour le Soleil et 30 MHz pour Jupiter, figure 3.21).

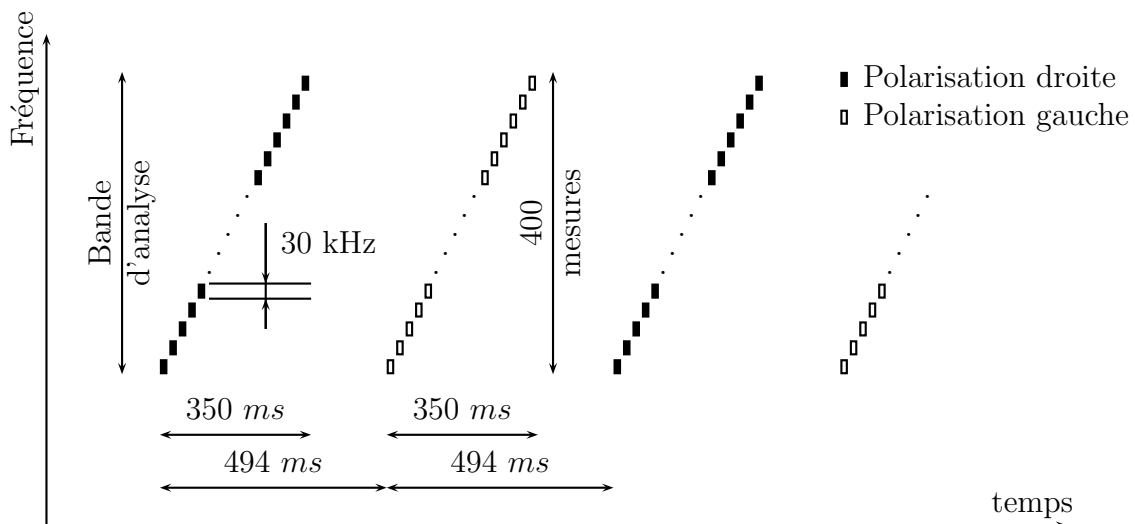


FIG. 3.21 – Plan d'occupation temps-fréquence de la routine classique

#### 3.3.1 Contraintes

Une nouvelle routine numérique décimétrique est demandée pour répondre aux nouveaux besoins d'observation.

L'implication de Nançay comme soutien sol de la mission STEREO nécessite une observation maximale du Soleil. Les observations de Jupiter et de Saturne deviendraient

secondaires. Une observation multi-sources est donc souhaitable. Le réseau décimétrique est pointé tour à tour sur les deux ou trois sources observables afin d'établir une couverture quasi continue des émissions de toutes les sources visibles.

Le récepteur numérique permet maintenant d'accéder à l'information de phase du signal grâce à l'enregistrement simultané de la cross-corrélation avec les densités spectrales de puissance des polarisations droite et gauche. Le volume de données est donc doublé.

La bande d'analyse de 14 MHz du récepteur numérique RDH impose un balayage en fréquence pour couvrir toute la bande d'étude des sources (50 MHz pour le Soleil, 30 MHz pour Jupiter). La brique de base des observations est fixée comme étant un spectre de largeur de bande 10 MHz, accumulé durant quelques centaines de millisecondes et centré à  $(10N + 5)$  MHz avec  $N \in [1, \dots, 7]$ . L'entrelacement des observations unitaires est modulable à souhait et permet par exemple le plan d'occupation temps-fréquence de la figure 3.22. Dans chaque bande, le signal est intégré plus longtemps que lors des observations avec la routine classique. Cela permet d'atteindre un SNR plus élevé comme le montrent les calculs et mesures regroupés dans le tableau 3.8.

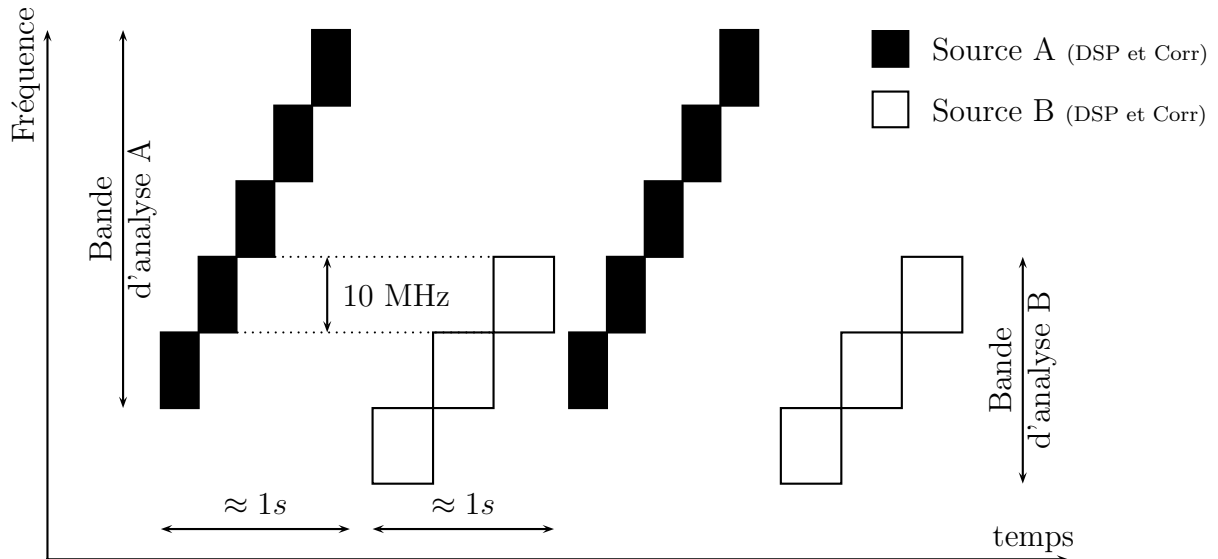


FIG. 3.22 – Plan d'occupation temps fréquence de la nouvelle routine RDH

TAB. 3.8 – Comparaison des SNRs Routine classique et Routine RDH

	Calculs			Mesures	
	Bande	$\tau$	Gain	SNR	Gain
Routine classique	30 kHz	0.875 ms	25.7 dB	-0.4 dB	19.2 dB
Routine RDH	65 kHz	150 ms		18.8 dB	

La résolution fréquentielle retenue reste du même ordre que celle de la routine classique afin de limiter la quantité de données à enregistrer sur disque. Néanmoins, les capacités du récepteur rendent possible l'élimination en temps réel des parasites à bande étroite. Les acquisitions sont effectuées à haute résolution fréquentielle afin de mieux pouvoir résoudre ces parasites (figure 3.23) puis le nombre de canaux est réduit à l'aide des algorithmes décrits dans la partie suivante qui cherchent à éliminer les émissions intenses à bande

étroite –signature la plus typique des parasites en bande décamétrique– tout en conservant intactes les statistiques du bruit qui contient l’information utile.

Les observations sont effectuées avec le paramétrage suivants:

- Bande d’acquisition : 14 MHz
- Nombre de canaux de la FFT : 4096
- Bande utile d’observation : 10 MHz
- Nombre de canaux utiles : 2926
- Résolution fréquentielle d’étude : 3.42 kHz (= 14 MHz/4096)
- Facteur de réduction de canaux : 19
- Nombre de canaux enregistrés : 154
- Résolution fréquentielle finale : 65 kHz (= 19 × 3.42 kHz)

Cette routine ne vise qu’à enregistrer les fluctuations des sources à une résolution temporelle de l’ordre de la seconde. A cette échelle, les signaux naturels présents sont de bande bien plus large que les parasites. Ils ne sauraient donc être confondus avec eux et éliminés comme des émissions sans intérêt.

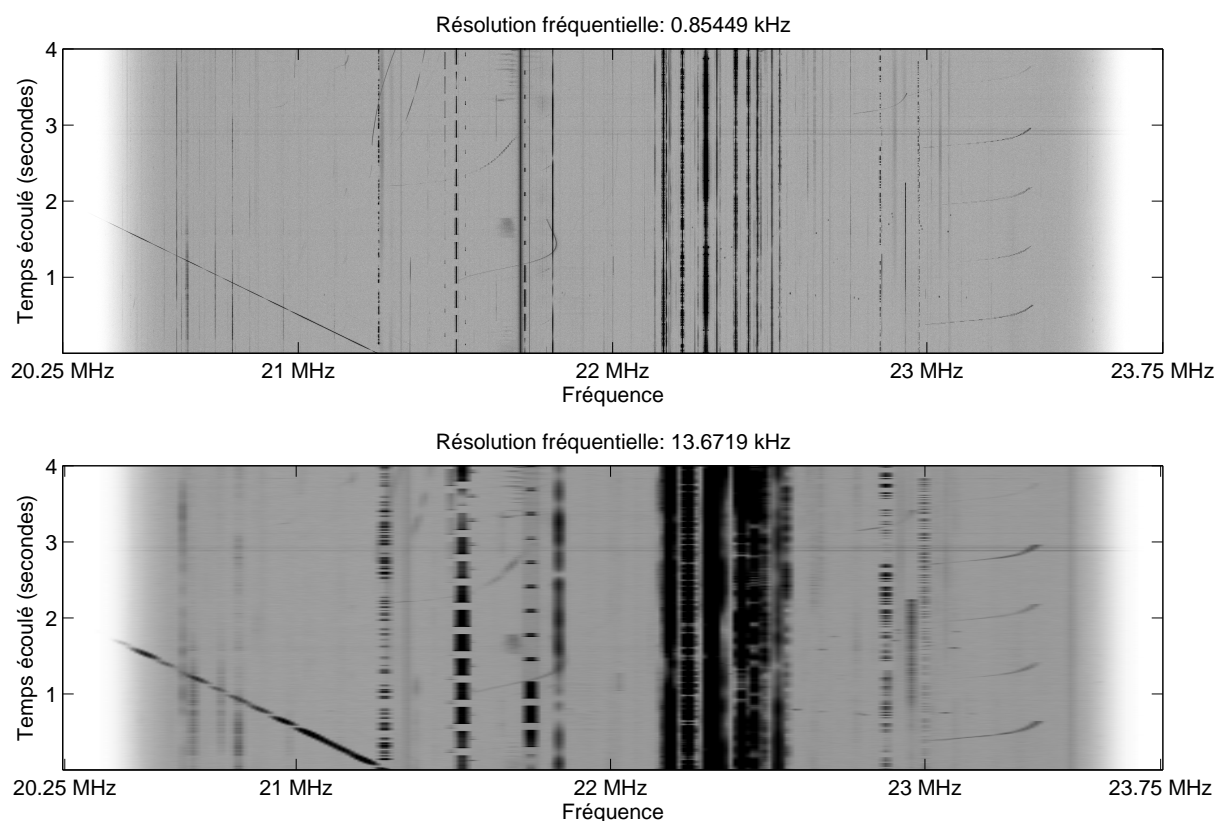


FIG. 3.23 – Effet de la résolution fréquentielle sur des parasites à bande étroite. L’analyse à haute résolution du signal permet de mieux résoudre les interférences : le nombre de canaux libres augmente.

### 3.3.2 Algorithmes

Les acquisitions multi-bandes et multi-sources rendent le flot de données trop hétérogène pour facilement traiter des plans temps-fréquence. Ces contraintes orientent un traitement



des données par spectre unitaire de 10 MHz.

Soit  $x_l[t]$  et  $x_r[t]$ , deux signaux temporels provenant d'une paire d'antenne à polarisations orthogonales. Les spectres complexes  $X_l[f_{étude}]$  et  $X_r[f_{étude}]$ , avec  $f_{étude} \in [0, \dots, 4095]$ , sont calculés par une FFT fenêtrée par une fonction de Blackman-Harris à 3 termes et les produits  $X_l[f_{étude}].X_l^*[f_{étude}]$ ,  $X_r[f_{étude}].X_r^*[f_{étude}]$  et  $X_l[f_{étude}].X_r[f_{étude}]$  sont accumulés pendant la durée d'observation de la bande dans  $\langle X_l.X_l^*[f_{étude}] \rangle$ ,  $\langle X_r.X_r^*[f_{étude}] \rangle$  et  $\langle X_l.X_r^*[f_{étude}] \rangle$ .

Ces trois vecteurs sont traités par l'un des algorithmes pour fournir  $\langle X_l.X_l^*[f_{routine}] \rangle$ ,  $\langle X_r.X_r^*[f_{routine}] \rangle$  et  $\langle X_l.X_r^*[f_{routine}] \rangle$ , avec  $f_{routine} \in [0, \dots, 153]$ .

La sélection des canaux non corrompus est effectuée par traitement des densités spectrales de puissance  $\langle X_l.X_l^*[f_{étude}] \rangle$  et  $\langle X_r.X_r^*[f_{étude}] \rangle$ , mais le résultat de la sélection est aussi appliqué à la corrélation  $\langle X_l.X_r^*[f_{étude}] \rangle$  pour garantir la cohérence du triplet  $\{\langle X_l.X_l^*[i] \rangle, \langle X_r.X_r^*[i] \rangle, \langle X_l.X_r^*[i] \rangle\}$  qui décrit le signal polarisé dans le canal  $i$ .

Les algorithmes décrits ci-après fonctionnent en aveugle et sans information *a priori*. Pour faciliter les développements, les algorithmes ont été implantés dans le PC d'acquisition. Le flot de données est suffisamment faible pour permettre un traitement temps réel sur une machine de faible puissance (PIII 800MHz).

### 3.3.2.1 MOYENNE HAUTE RÉOLUTION

Le fenêtrage retenu avant le calcul de la FFT est un Blackman-Harris à 3 termes car sa réjection des lobes secondaire de 67 dB le rend performant pour la séparation de raies voisines et il se comporte bien en arithmétique à virgule fixe dans les calculs effectués par les FPGAs [36]. La formule qui décrit la fenêtre d'apodisation représentée figure 3.24 est :

$$w[n] = 0.42 - 0.5 \cos\left(\frac{2\pi}{N-1}n\right) + 0.08 \cos\left(\frac{4\pi}{N-1}n\right) \quad (3.10)$$

avec  $n \in [0, \dots, N-1]$

Si l'on souhaite analyser directement la bande de 10 MHz à une résolution proche de la résolution de routine (154 canaux), on pourrait directement effectuer une FFT de 256 points sur une bande d'analyse de 14 MHz pour recueillir 183 points utiles couvrant la bande de 10 MHz. L'inconvénient de cette méthode est que toute raie parasite de puissance supérieur à 20 dB au dessus du bruit va occuper un minimum de 5 canaux fréquentiels, soit 273 kHz, alors que l'occupation spectrale d'une raie (infiniment étroite en théorie) portant une modulation de type "modulation d'amplitude" peut n'occuper que quelques kilohertz. En effet, le fenêtrage de Blackman-Harris a pour inconvénient de présenter un large lobe central. En présence de raies voisines, même séparées de plusieurs canaux, le plan temps-fréquence est entièrement parasité (figure 3.23, plan temps-fréquence du bas à 22.3 MHz), le bruit du fond du ciel est inaccessible : toute source éventuelle serait masquée par les parasites.

L'analyse de la bande à 4096 canaux permet de réduire l'occupation spectrale d'une raie à 21 kHz . 19 canaux contigus sont ensuite moyennés pour respecter la résolution routine (figure 3.25). Si la raie est entièrement incluse dans les 19 canaux d'analyse, l'occupation spectrale routine vaut un canal, soit 65 kHz. Si la raie chevauche deux sous-bandes de 19 canaux d'analyse, l'occupation spectrale est doublée pour atteindre 130 kHz.

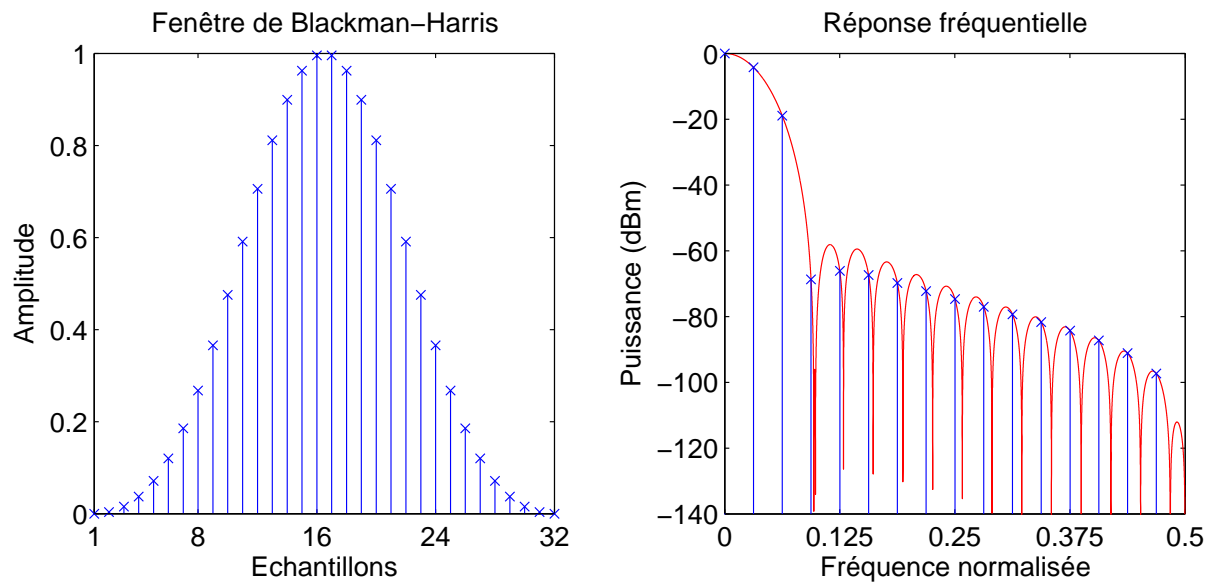


FIG. 3.24 – Fenêtre de Blackman-Harris

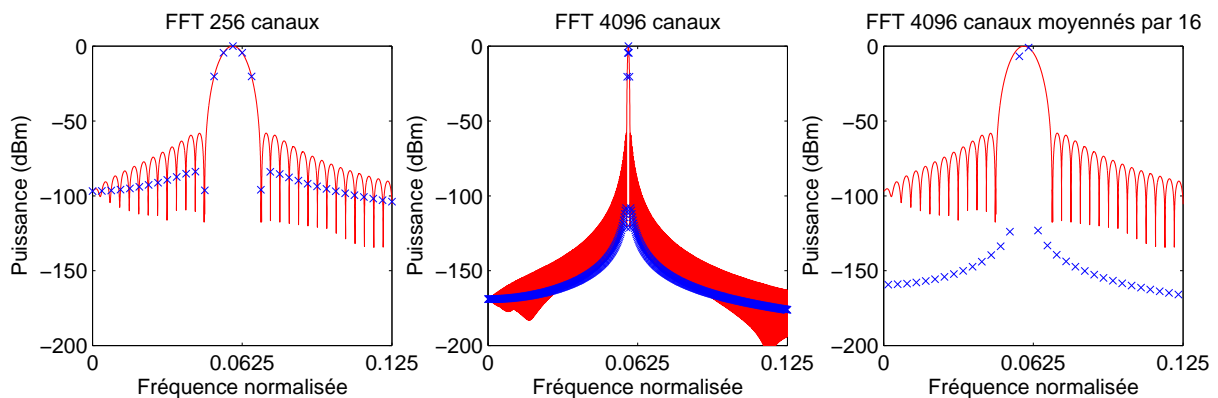


FIG. 3.25 – Diverses estimations spectrales d'une raie par fenêtrage par Blackman-Harris suivit d'une FFT (zoom en abscisse sur la raie). Gauche : FFT sur 256 points, (au centre) FFT sur 4096 points. Droite : FFT sur 4096 points moyennés par 16 points

Les résultats de la simulation des canaux occupés lors de l'analyse d'une raie de fréquence choisie aléatoirement dans la bande équivalente de 19 canaux à haute résolution sont disponibles en figure 3.26. La réduction moyenne de la largeur d'un lobe vaut 73%.

Cette procédure simple réduit d'un facteur 4 l'occupation spectrale d'une raie dans la bande. Elle présente néanmoins un gros défaut : la moindre raie parasite condamne irrémédiablement le canal routine qui la contient alors que l'analyse à haute résolution permettrait d'entrevoir le bruit de fond du ciel. Le cas extrême est celui d'une bande où seraient disposées des raies espacées en fréquence d'un intervalle proche de la résolution routine. En haute résolution, l'occupation spectrale moyenne resterait inférieur à  $6 \text{ canaux} / 19 \text{ canaux} = 32\%$ , soit 68% de ciel visible mais le passage à la résolution routine catapulterait l'occupation spectrale à 100%, rendant toute observation impossible.

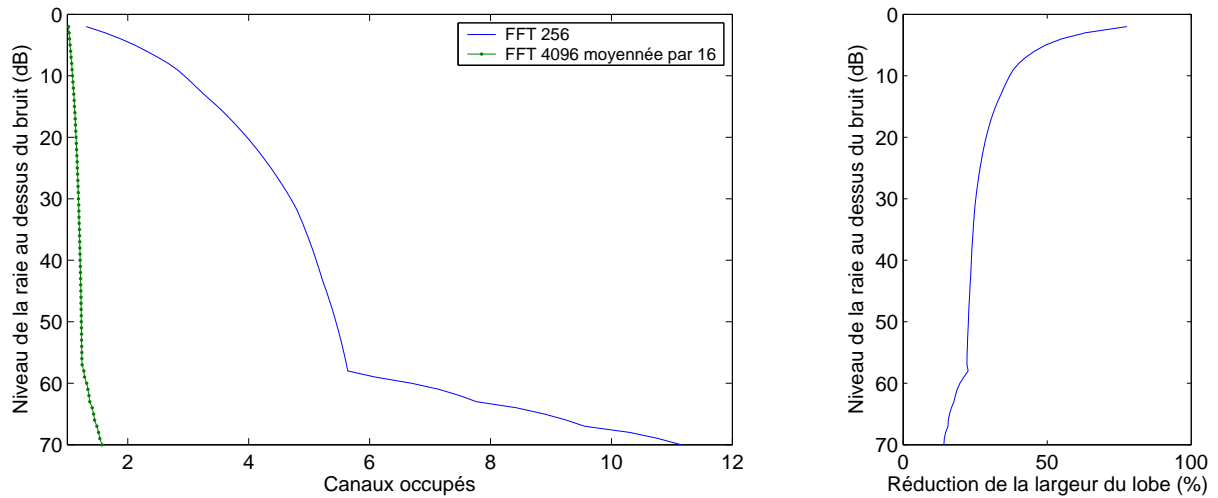


FIG. 3.26 – Simulation du nombre de canaux occupés par une raie lors de l'analyse spectrale par une FFT 256 points et par une FFT 4096 points suivie d'une moyenne par bloc de 19 points.

### 3.3.2.2 MÉDIANE

Il faut donc trouver un moyen pour estimer correctement le niveau de bruit d'origine d'un canal en présence de raies parasites. Les travaux effectués sur la médiane en section 2.2.3 peuvent nous aider dans ce sens.

La première étape du calcul est la création du vecteur  $\langle X_{lr}.X_{lr}^*[f_{étude}] \rangle$  tel que :

$$\langle X_{lr}.X_{lr}^*[k] \rangle = \langle X_l.X_l^*[k] \rangle + \langle X_r.X_r^*[k] \rangle \quad \text{pour } k \in f_{étude}. \quad (3.11)$$

Cela permet de rechercher un canal intègre simultanément en polarisations droite et gauche.

Ensuite, la médiane est calculée sur des jeux consécutifs de 19 canaux de  $\langle X_{lr}.X_{lr}^*[étude] \rangle$ . Ici, nul besoin des corrections apportées par la médiane corrigée car le temps d'intégration utilisé pour observer le Soleil (respectivement Jupiter) est de 150 ms (respectivement 300 ms), soit une accumulation de 1024 (respectivement 2048) spectres élémentaires. Le degré de liberté de la loi du  $\chi^2$  vaut 4096 (respectivement 8192) soit une sous-estimation de la puissance d'un facteur  $K(\nu, N) = 1.000150 \pm 12 \cdot 10^{-6}$  (respectivement  $1.000077 \pm 29 \cdot 10^{-6}$ ). L'erreur au final ne vaut que  $0.7 \cdot 10^{-3}$  dB (respectivement  $0.3 \cdot 10^{-3}$  dB). Le coefficient correcteur  $K(\nu, N)$  a été estimé par simulation car le calcul analytique sur machine de type PC produit des valeurs hors de l'espace de représentation flottant 64 bits IEEE 754-1985.

Un exemple d'observation sur le Soleil est donné en figure 3.27. Les raies parasites à 36 MHz sont entièrement retirées. Celles autour de 23 et 27 MHz occupent maintenant une bande bien plus faible. Les signaux à bande large (2 Types III et une calibration) ne sont pas altérés par le traitement. En contrepartie, des signaux à bande large et courte durée sont maintenant plus souvent présents dans la bande d'analyse.

### 3.3.2.3 MOYENNE DU PETIT POUCKET ET MOYENNE DE CLASSE

La médiane ne peut restaurer un canal que si le pourcentage de canaux haute résolution parasités est inférieur à 50%. Pourtant, même à de plus fort taux occupation, il est toujours possible d’entrevoir le bruit de fond du ciel. Les algorithmes d’estimation de moyenne par moyenne du Petit Poucet et moyenne de classe ont été testés sur le ciel. Leurs effets sur des signaux à bande large comme des Types III ne sont pas décelables mais une observation particulière remet en cause leur utilisation en routine.

Le signal étudié est une observation de Jupiter le 25 octobre 2006 de 12h00TU à 12h15TU entre 20 et 35 MHz (voir le "Quicklook routine" en figure 3.28). C’est une émission de faible niveau acquise peu de temps après le passage au méridien du Soleil (11h35TU). Les observations Joviennes y sont habituellement de faible qualité. De plus, l’émission est noyée au milieu de forts émetteurs artificiels qui font intermoduler une partie de la chaîne analogique.

La portion de signal étudiée est disponible en figure 3.29 et figure 3.30. Les données routine classiques ne permettent pas de distinguer l’émission noyée dans le bruit. L’acquisition avec RDH (figure 3.29) permet un meilleur SNR et donc un meilleur détail de la source. De plus, la corrélation permet de mieux localiser une source grâce à ses franges d’interférence. Nous ne nous intéresserons pas aux émissions de très faibles niveaux qui ne sont pas discernables en densité spectrale de puissance. Après moyennage des canaux pour fournir des plans temps-fréquence à la résolution "routine" (figure 3.30), le nombre de canaux occupés par les interférences reste élevé mais la source est toujours visible. Dans les trois cas, le contraste est très accentué pour faciliter la localisation des signaux naturels.

Afin de comparer le comportement des différents algorithmes, il faut disposer d’un jeu de donnée commun. Puisque la source n’est pas clairement visible dans les données "routine" classique et que les données RDH hautes résolutions ne peuvent être comparées directement avec les données à résolution routine du fait de la différence de résolution, les données obtenues par moyenne haute résolution seront la référence pour la suite de cette section. Un autre argument en faveur de l’utilisation de ces données pour référence est qu’il s’agit d’un algorithme aux propriétés strictement linéaires, ce qui n’est pas le cas des autres méthodes.

Le plan temps-fréquence généré par estimation de moyenne par médiane est donné en figure 3.31, joint à sa différence avec le plan TF de référence. La source ne semble pas avoir été altérée par le calcul. De nombreuses raies ont subi une atténuation atteignant jusqu’à 40 dB.

Le plan temps-fréquence généré par estimation de moyenne par moyenne du Petit Poucet est donné en figure 3.32, joint à sa différence avec le plan TF de référence. L’atténuation des raies est encore plus radicale mais la source est maintenant altérée. Le même comportement est présent lors de l’estimation de moyenne par moyenne de classe (figure 3.33). La différence entre ces deux dernières méthodes se limite à un rapport signal à bruit légèrement meilleur pour l’estimation par moyenne de classe. Elle n’est pas visible sur les figures.

L’altération de la source par les algorithmes du Petit Poucet et de la moyenne de classe provient de la haute résolution des franges Faraday présentes dans la source. Sur les données étudiées, l’interfrange à 18.5 MHz vaut 40 kHz et celle à 32 MHz vaut 300 kHz. Le domaine de variation des interfranges est donc très proche de la résolution "routine".

Les algorithmes pris en défaut sont tous deux basés sur la recherche d'un minimum dans le canal final. Ce minimum correspond aux minima locaux des franges. Le reste du signal contenu dans le canal est considéré comme des interférences. Les maxima locaux des franges sont donc quasi systématiquement écartés lors de l'estimation de la moyenne.

L'estimation de moyenne par médiane prend en compte l'ensemble des points du canal final. En cas de présence exclusive de franges dans le canal, la distribution du bruit est convoluée par la distribution de l'enveloppe des franges que l'on peut approcher par une portion de sinusöide. Les deux distributions étant symétriques, le résultat est lui aussi symétrique : la médiane coïncide avec la valeur de la moyenne.

### 3.3.3 Conclusion

L'occupation spectrale des raies parasites a été calculée pour différents estimateurs de moyenne. Les résultats sont présentés dans la table 3.9. La routine classique et l'estimation de moyenne par moyenne arithmétique donnent des résultats similaires pour des résolutions spectrales équivalentes (75 kHz / 65 kHz). Le passage en haute résolution diminue sensiblement l'occupation spectrale des fortes raies mais affecte à peine les raies de faible niveau. La médiane fait gagner un rapport 4 sur les raies de forts niveaux et un rapport 2 sur les raies de très faibles niveaux. Les estimateurs par moyenne du Petit Poucet et par moyenne de classe suppriment totalement les raies de niveaux supérieurs à 5 dB. Les raies de niveaux compris entre 1 et 5 dB empiètent maintenant 20 fois moins sur la bande d'analyse que lors d'une réduction de canaux par simple moyenne. La comparaison avec la routine classique n'a pas été possible car son bruit d'estimation est supérieur à 1 dB. Le SNR de la moyenne de classe est très légèrement inférieur à celui de la moyenne du Petit Poucet (1 à 2 dB). Néanmoins la charge de calcul supplémentaire ne vaut peut être pas l'effort dans la majorité des situations.

TAB. 3.9 – Taux d'occupation spectrale par des interférences à bande étroite entre 20 et 40 MHz avant et après traitements pendant 60 s d'observation. La valeur manquante pour la routine classique est imputée à un trop faible SNR du plan temps-fréquence : les parasites sont noyés sous le bruit (1.17 dB contre moins de 0.1 dB pour les autres données).

	Niveau des interférences			
	≥30 dB	≥15 dB	≥5 dB	≥1 dB
Routine classique	1.4%	2.6%	11.1%	-
Routine haute résolution	0.2%	0.1%	4.5%	14.7%
Moyenne arithmétique	0.6%	1.9%	7.9%	22.9%
Médiane	0.0%	0.3%	1.8%	9.6%
Petit Poucet	0.0%	0.0%	0.0%	1.2%
Moyenne de classe	0.0%	0.0%	0.0%	1.0%

Quelque soit la méthode de réduction de donnée utilisée par la nouvelle routine, le long temps d'intégration des spectres vis à vis du temps d'intégration de la routine analogique augmente la probabilité d'observer des parasites à bande large. Là où la routine analogique n'était parasitée que durant le temps de présence de ce type d'interférence, la routine

numérique rend inutilisable un bloc entier de 10 MHz. Il pourrait être intéressant de supprimer ces interférences en amont de l'analyse spectrale avec un détecteur du même type que le "blankeur" radar (voir section 3.1).

Pour faciliter les développements techniques, la résolution d'analyse commune de 3.4 kHz a été choisie pour toute la bande 10-70 MHz. Néanmoins, la bande 10-20 MHz utilisée par les ondes courtes nécessite une résolution encore plus faible pour correctement résoudre les émetteurs. La thèse de Vincent Clerc [11, p.136-144] traite ce sujet en détail en estimant les plans temps-fréquence par l'accumulation de FFTs fenêtrées (Hamming ou Hanning). L'utilisation d'un filtre polyphase serait ici indiquée pour augmenter le nombre de canaux libres avant traitement grâce son excellente réjection inter-canal.

L'augmentation de résolution aux basses fréquences se justifie aussi pour la restitution des sources subissant l'effet Faraday. C'est le cas pour les émissions joviennes qui sont fortement polarisées elliptiquement. Aux résolutions actuellement choisies, ces observations ne pourront se faire qu'avec l'algorithme d'estimation de moyenne par médiane. Les observations solaires ne souffrent pas de cette limitation car les émissions du Soleil calme ne sont pas polarisées et celles des sursauts solaires sont polarisées circulairement [31]. Ces émissions ne subissent donc pas l'effet Faraday.

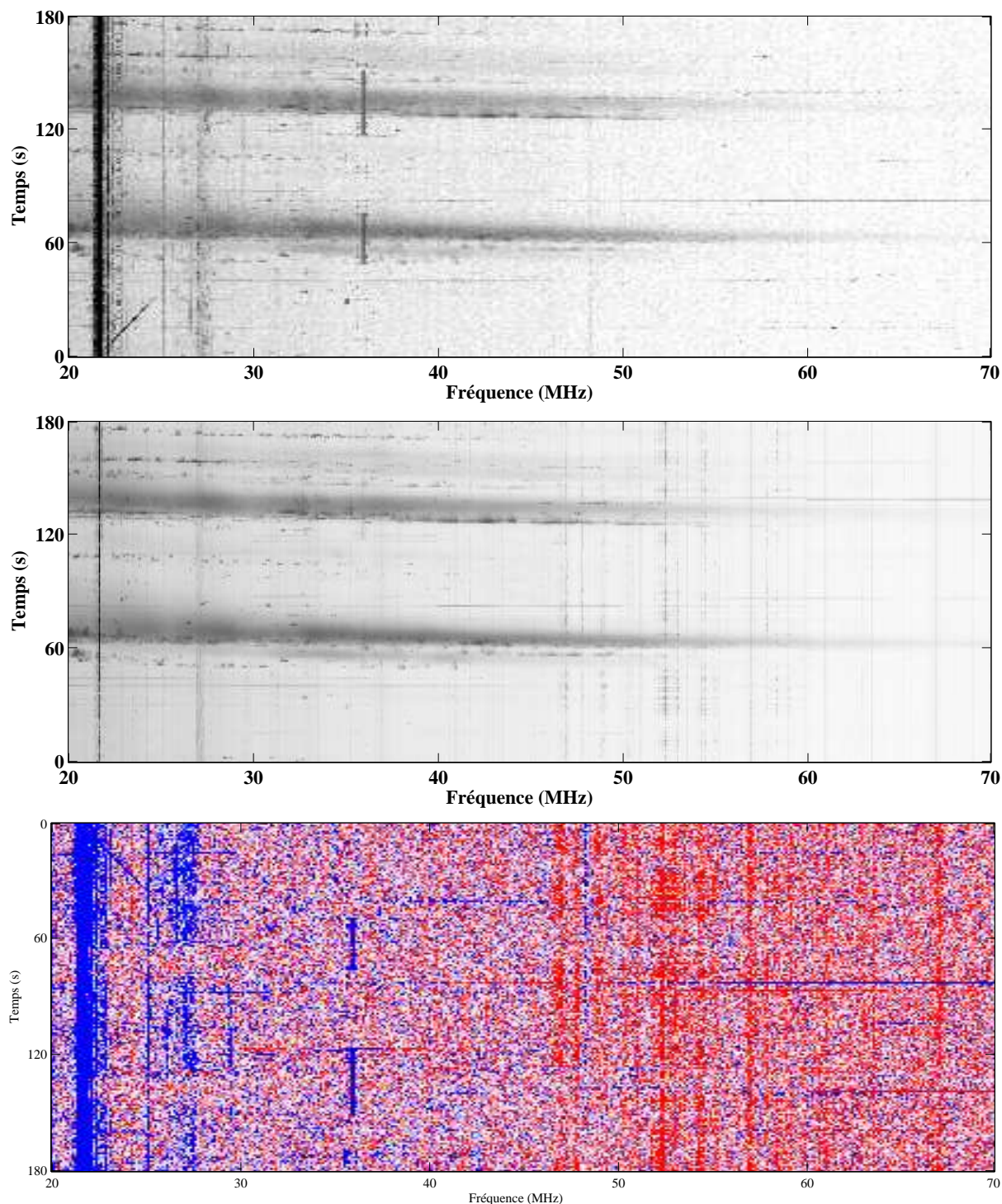


FIG. 3.27 – Comparaison des observations "routine" (en haut) et "RDH" (au milieu). Le plan temps-fréquence "RDH" a été ré-échantillonné au format "routine" pour calculer la différence saturée à  $\pm 2$  dB entre les deux jeux de données (en bas). Les zones bleues verticales correspondent aux interférences atténuées jusqu'à 40 dB par l'algorithme. Les zones horizontales sont des éclairs d'orages enregistrés uniquement par l'un ou l'autre des systèmes. Les zones rouges verticales sont des raies systèmes parasites. Les trois signaux à bande large (deux Types III et une calibration) ne sont pas altérés pendant le traitement.

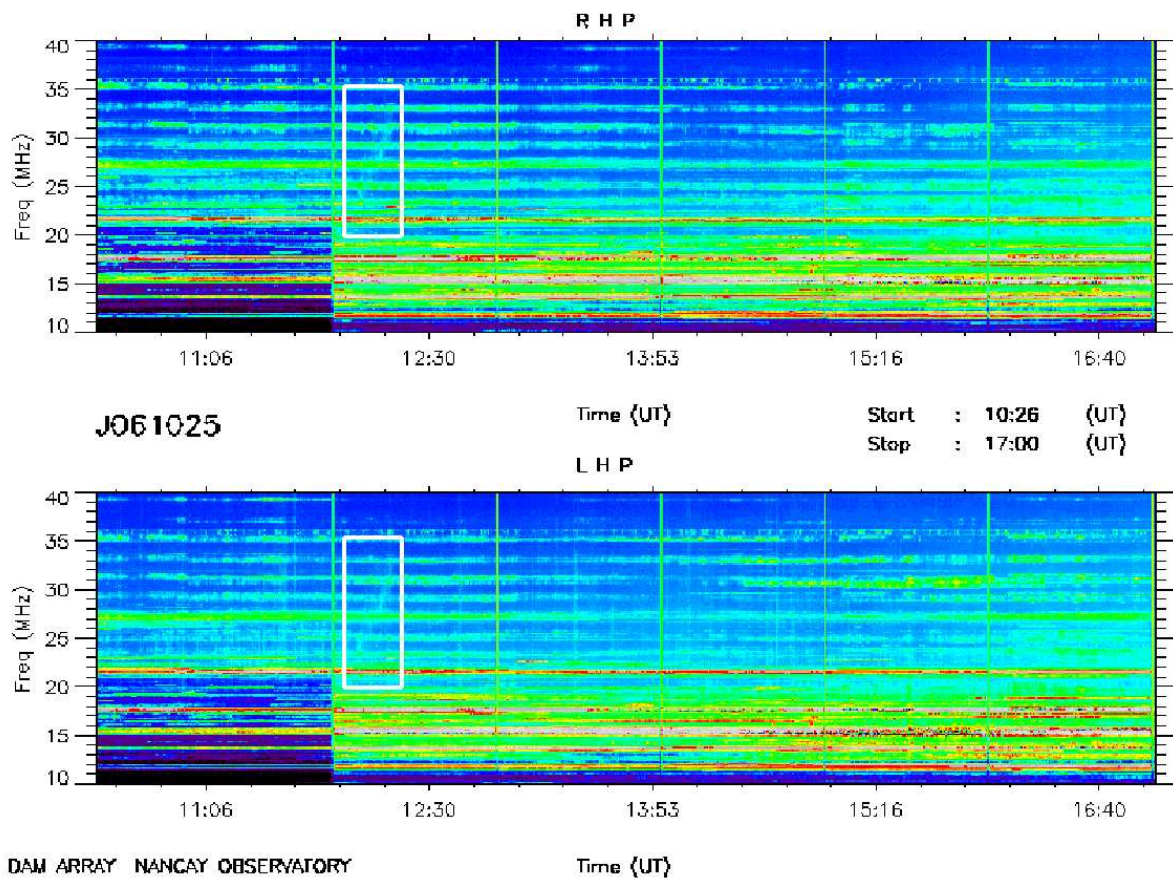


FIG. 3.28 – Quicklook d’une observation Jovienne obtenue par moyennage des données “routine” enregistrées au NDA.

<http://www.obs-nancay.fr/dam> ”Quicklooks Soleil-Jupiter”



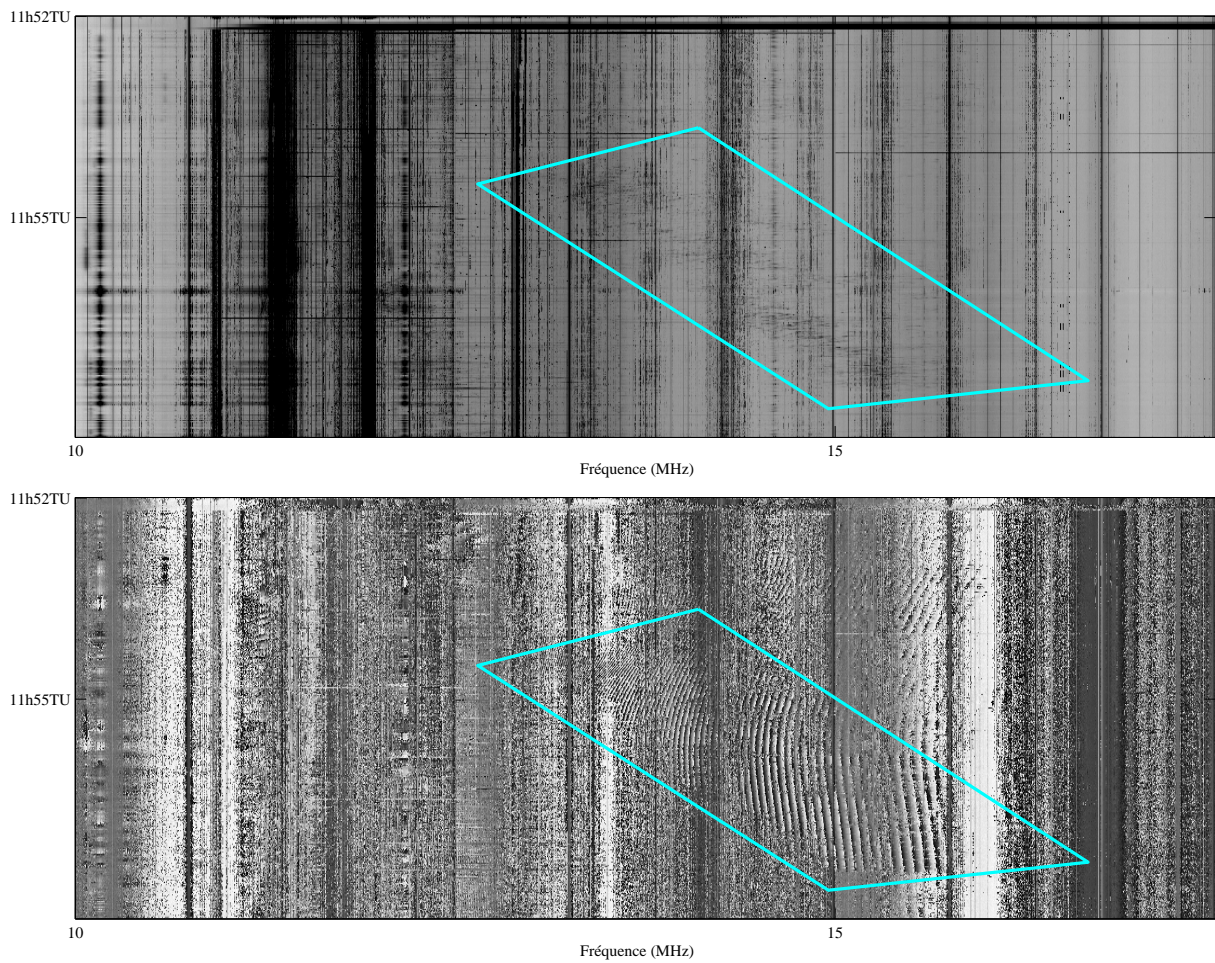


FIG. 3.29 – Données haute résolution d’une émission Jovienne. Haut : densité spectrale de puissance de la polarisation gauche. Bas : phase du signal issue des données de la corrélation de la polarisation droite avec la polarisation gauche.

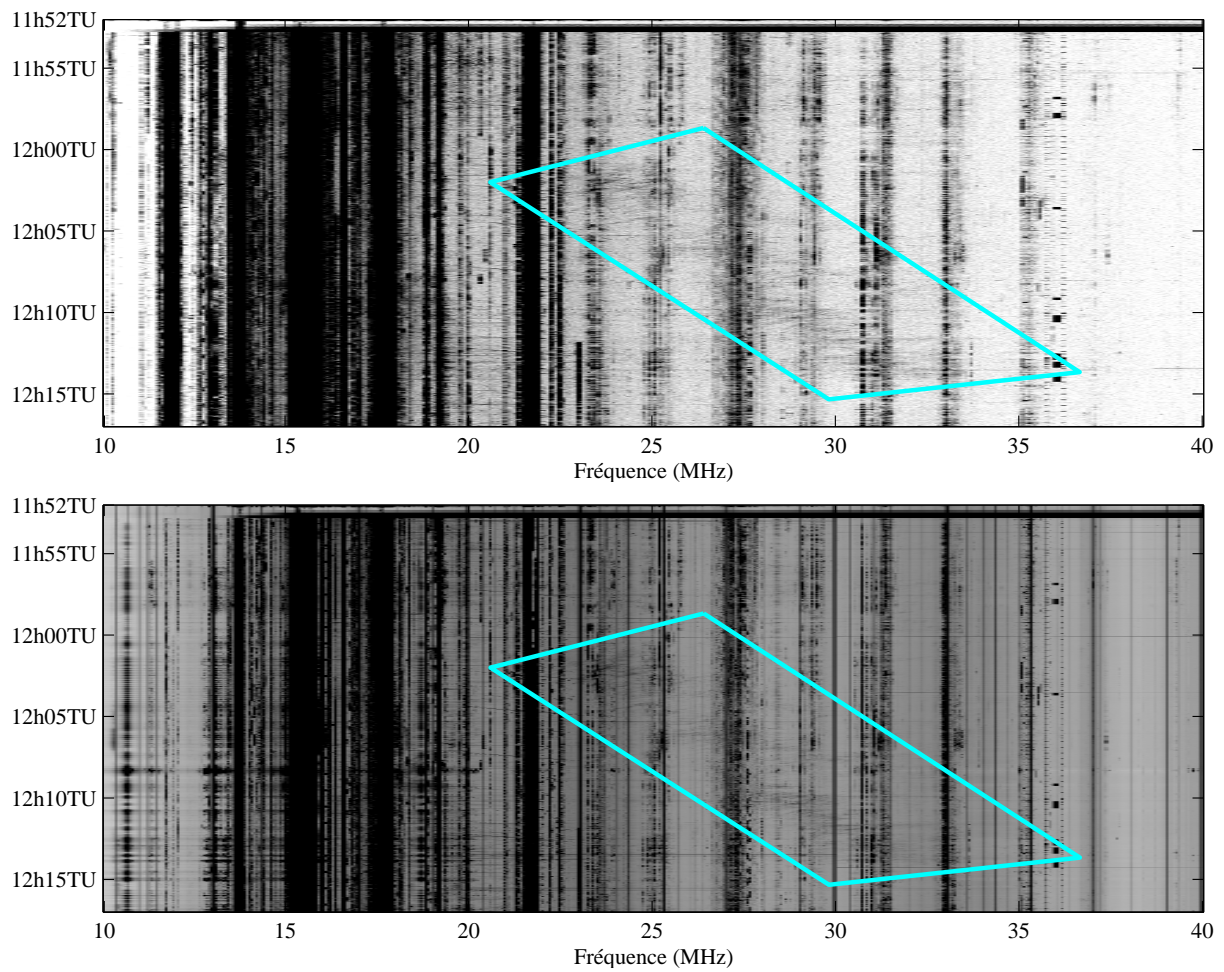


FIG. 3.30 – Émission Jovienne. Haut : Données “routine” classiques. Bas : données RDH après moyenne haute résolution.

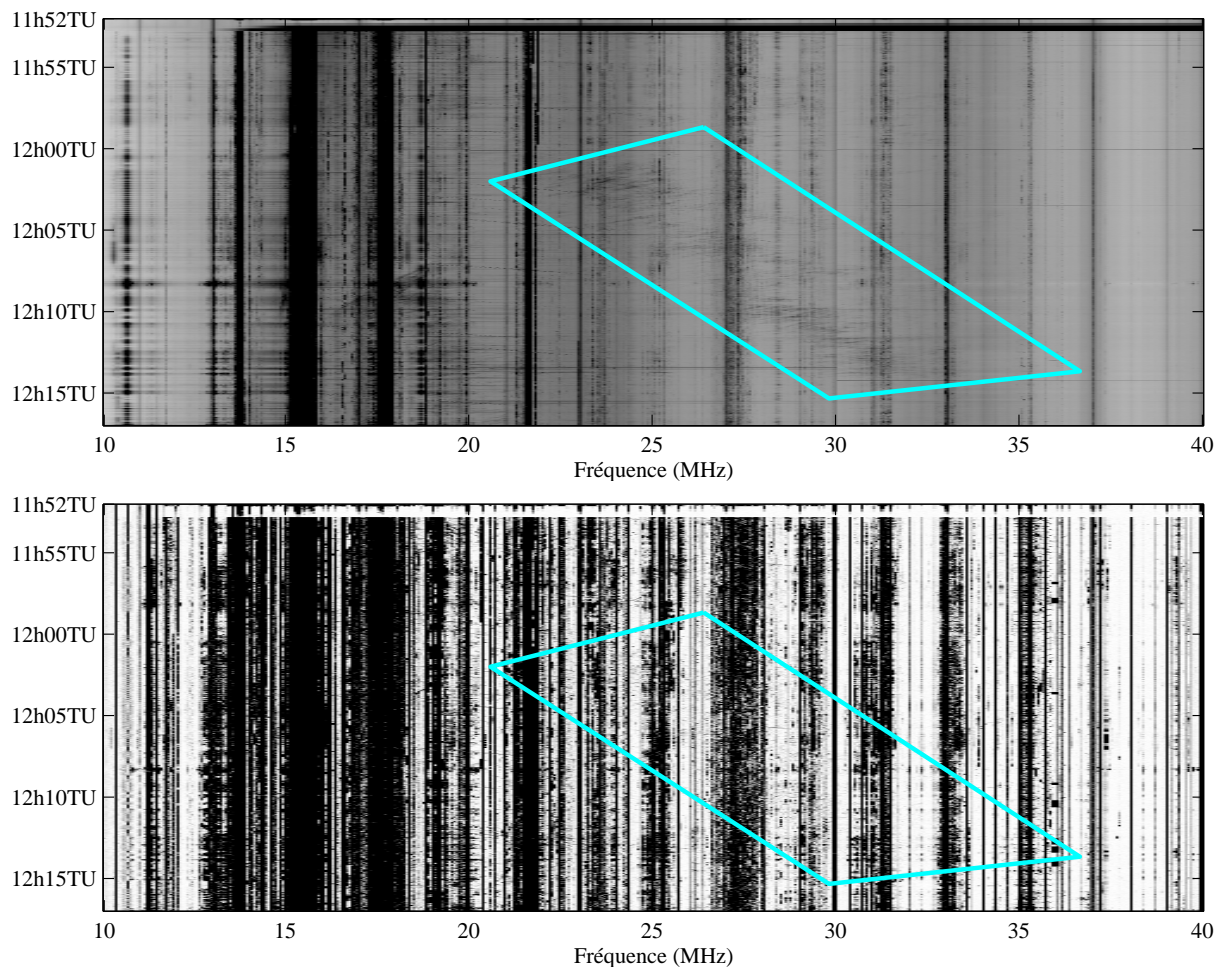


FIG. 3.31 – Données “routine” d’une émission Jovienne estimées par médiane (haut) et sa différence avec le plan TF de référence saturé à 1 dB (bas).

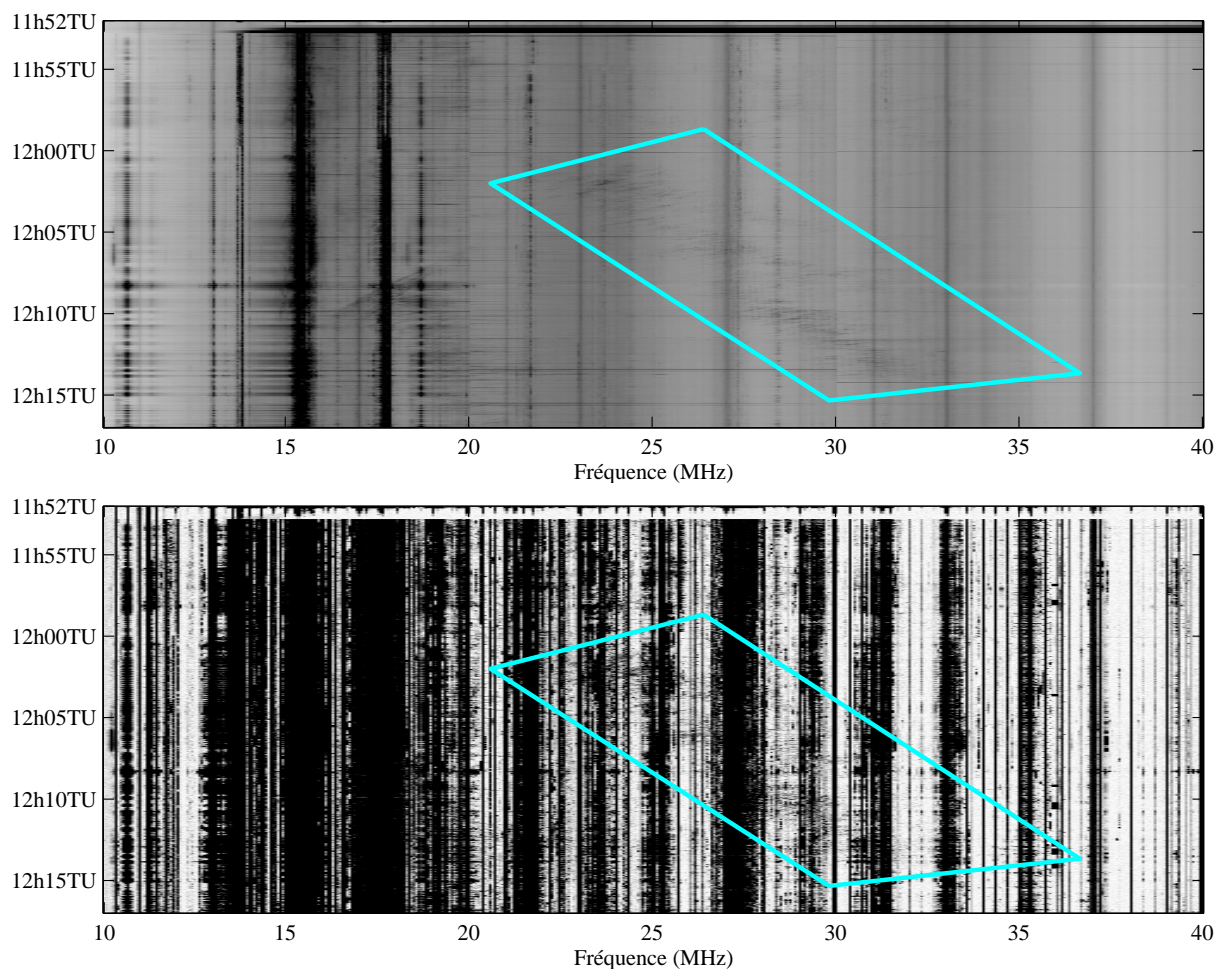


FIG. 3.32 – Données “routine” d’une émission Jovienne estimées par moyenne du Petit Poucet (haut) et sa différence avec le plan TF de référence saturé à 1 dB (bas).

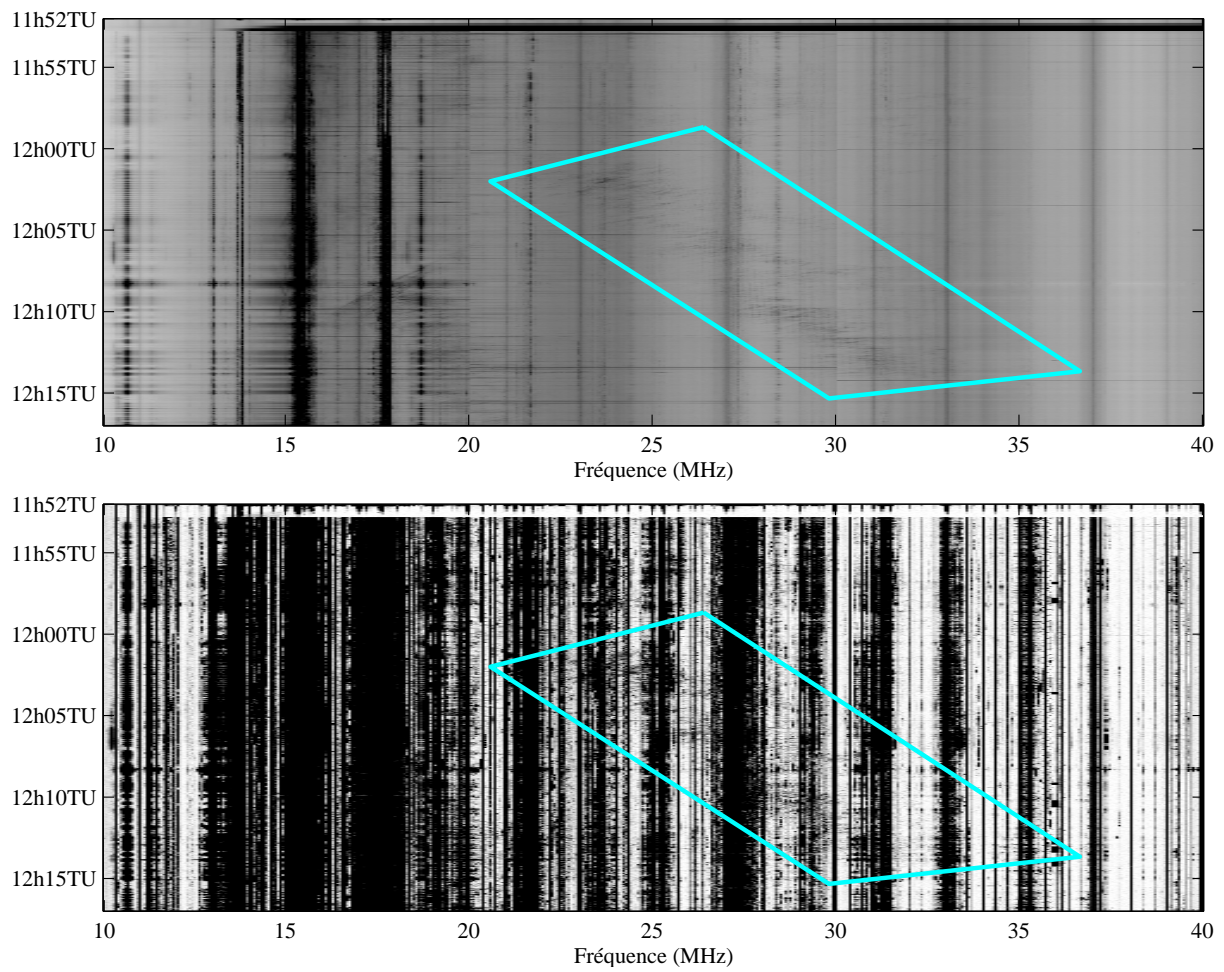


FIG. 3.33 – Données “routine” d’une émission Jovienne estimées par moyenne de classe (haut) et sa différence avec le plan TF de référence saturé à 1 dB (bas).

## CHAPITRE 4

---

# DÉTECTION DE SIGNAUX BREFS ET INTENSES EN MILIEU PARASITÉ

---

La bande décamétrique fournit plusieurs exemples d'émissions brèves dérivant en fréquence. Le meilleur exemple est l'émission radio de l'interaction Io-Jupiter, appelée structures fines ou S-bursts, qui dérive généralement des hautes vers les basses fréquences au taux de quelques mégahertz par seconde pendant une durée de quelques dizaines de millisecondes. La bande de fréquence instantanée du signal est de quelques kilohertz. Un autre exemple est l'émission de structures fines solaires qui est parfois associée à l'émission de Types III [4]. On peut prédire les émissions joviennes pour certaines configurations géométriques observateur-Jupiter-Io mais il est possible que des émissions soient générées pour d'autres géométries et qu'elles aient une dérive fréquentielle éventuellement positive. L'activité des structures fines solaires, quant à elle, est plutôt imprédictible. Les bases de données sur ces types d'émissions sont donc trop incomplètes pour effectuer des études statistiques.

Une routine permanente serait utile pour rassembler des informations sur de tels événements. Néanmoins, les résolutions temporelle et fréquentielle requises pour permettre l'explication des phénomènes physiques sous-jacents produiraient de faramineux volumes de données. De plus, les traitements en temps différé effectués pour détecter les événements dignes d'intérêt nécessiteraient l'utilisation d'un ordinateur et une connexion réseau à grande bande passante pour transférer les données. Un système qui aurait la faculté d'examiner de façon quasi-permanente une source donnée en vue de détecter des structures fines pour déclencher leur enregistrement à haute résolution serait un atout majeur. En effet, l'étude du phénomène serait approfondie car basée sur de nombreux enregistrements.

Une solution est d'effectuer la détection sur la plateforme temps réel RDH qui déclencherait les enregistrements. L'application FISSA (FIne Spectral Structures Algorithm) a été développée dans ce sens. Un indice qualifiant la présence de structures fines est calculé en temps réel et permet de prendre la décision d'une pose sur disque des données.

Les travaux précédemment exposés portent sur la détection et l'élimination d'interférences en vue de restaurer les sources radio naturelles présentes. Néanmoins, il est évident que la détection en puissance seule présente de fortes limitations. Par exemple, les travaux de [8] montrent que dans des conditions d'observations identiques, il est possible de détecter des interférences d'INR plus faibles qu'en utilisant de simples détecteurs de puissance. Cela implique, bien sûr, un coût de calcul et une complexité d'implantation

plus élevés. Si l'on souhaite appliquer des traitements plus élaborés, il peut être judicieux de ne traiter que les données où des signaux d'intérêt auraient été identifiés. On peut alors décider d'étendre le sujet du traitement des interférences à la détection de signaux naturels en présence quasi-continue de brouilleurs radioélectriques.

## 4.1 Caractéristiques des signaux

L'étude de la détection a été menée sur les signaux joviens car des zones où les probabilités sont fortes ont déjà été identifiées expérimentalement [27]. Il est donc aisé de planifier des observations de tests. Ces signaux sont observés depuis la fin des années 60 [29, 61, 48, 47, 28, 16] mais les phénomènes physiques ne sont toujours pas compris et des études sur des signaux acquis avec des méthodes plus modernes sont encore en cours [38]. L'un des besoins qui se fait sentir est la disponibilité de signaux enregistrés sur toute la durée de l'émission afin d'étudier leur évolution au cours du temps.

Des exemples de signaux étudiés sont donnés en figure 4.1. Ils sont produits par l'interaction d'ions générés par l'éjection de matière due à l'activité volcanique de Io avec le puissant champ magnétique de Jupiter. La totalité des phénomènes mis en jeu ne sont pas encore compris mais plus d'informations sont disponibles dans les références citées plus haut. La puissance des S-bursts dépasse rarement 15dB au dessus du bruit du ciel. Les interférences, elles, sont typiquement à 30dB et peuvent atteindre 70dB. On notera la non-stationnarité des fréquences des émetteurs à bande étroite et la présence ponctuelle de parasites à large bande.

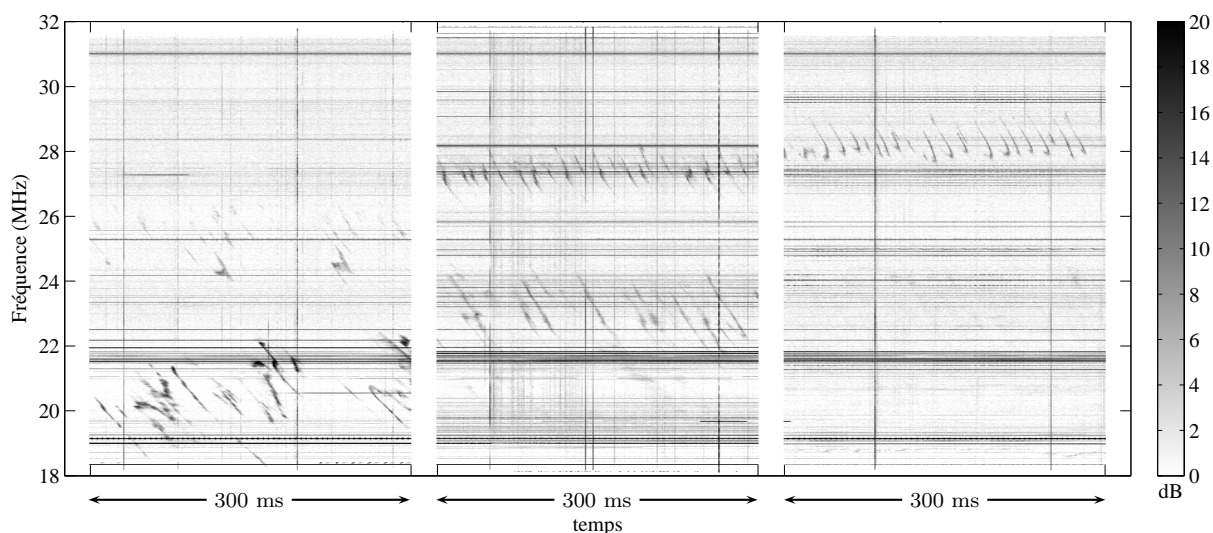


FIG. 4.1 – Exemples de structures fines à détecter enregistrés en polarisation circulaire gauche le 03/12/2006, centrés à 09:38:39,6TU, 09:58:27,2TU et 10:08:44,9TU.

La différence de niveau entre les signaux et les interférences rend très difficile la détection des structures fines à l'aide d'un simple critère de puissance. Une pré-étude de détection basée sur une procédure de seuillage adapté, d'érosions morphologiques utilisant des éléments structurants obliques puis d'une intégration de l'image résultante n'a permis de détecter que les émissions les plus intenses. De plus, l'algorithme était trop

lourd et de durée d'exécution trop imprédictible pour permettre facilement son implantation matérielle sur une plateforme temps réel. Finalement, seules les structures d'angle définies par les éléments structurants étaient détectables. Il était donc impossible avec un tel système d'explorer le plan temps-fréquence à la recherche de toutes les structures dérivantes.

La recherche des lignes inclinées par transformée de Radon [70] est très perturbée par la présence des interférences fortes, et ce, même après passage en échelle logarithmique.

Un système de détection similaire basé sur des réseaux de neurones a été développé par [17] pour détecter des signaux transitoires de type sifflement dans la bande radio très basse fréquence (2-9 kHz). Cette solution n'a pas été envisagée à cause de la faible quantité de données disponibles pour construire une base d'apprentissage.

Une approche fréquentielle 2D utilisant des filtres de Gabor a été envisagée [46, 25, 21] mais, comme pour les traitements morphologiques, la nécessité de construire un filtre pour chaque valeur de dérive recherchée nous a très vite fait reculer. Néanmoins, cette approche nous a permis d'envisager la solution que fut finalement retenue.

## 4.2 Algorithmes

Les structures fines peuvent être assimilées à une texture ayant une direction privilégiée oblique. Le passage dans le plan  $u-v$  (*via* une FFT-2D définie par l'équation 4.1) fait apparaître une ellipse centrée sur le DC<sup>1</sup> et tournée de  $\pi/2$  par rapport à la direction privilégiée de la texture. Les interférences sont soit à bande étroite, soit à bande large. Dans le plan  $u-v$ , elles se retrouvent donc localisées dans une croix centrée sur le DC. Alors que structures fines et interférences étaient spatialement mélangées dans le plan  $t-f$ , leur représentation dans le plan  $u-v$  permet leur séparation (figure 4.2).

$$\mathcal{TF}\{I[f, t]\} = \mathcal{I}[u, v] = \sum_{f=0}^{N-1} \sum_{t=0}^{N-1} I[f, t] e^{-i2\pi \frac{f \cdot u}{N}} e^{-i2\pi \frac{t \cdot v}{N}} \quad (4.1)$$

Il suffit donc de détecter si une ellipse se détache au centre du plan  $u-v$ . La mesure de son angle d'inclinaison permet de connaître le taux de dérive des structures fines et éventuellement de limiter l'enregistrement à certains types d'émissions.

L'algorithme complet est exposé en figure 4.3.

### 4.2.1 Périodogramme moyenné

L'analyse spectrale haute résolution est effectuée par les FPGAs du système sur les deux voies d'entrée. Des blocs de 2048 points de données échantillonnées à 14 mégaséchantillons complexes par seconde sont projetés dans l'espace de Fourier. Les périodogrammes calculés sont moyennés d'un facteur 8. La densité spectrale de puissance des deux polarisations circulaires d'antenne ainsi que l'inter-produit complexe (information de phase)

<sup>1</sup>On note DC le pixel de coordonnées  $[u, v] = [0, 0]$ . Il est commun de représenter le plan  $u-v$  avec des fréquences spatiales dans le domaine  $[-0.5, 0.5]$ . Dans cette représentation, le DC est donc situé au centre du plan  $u-v$ .



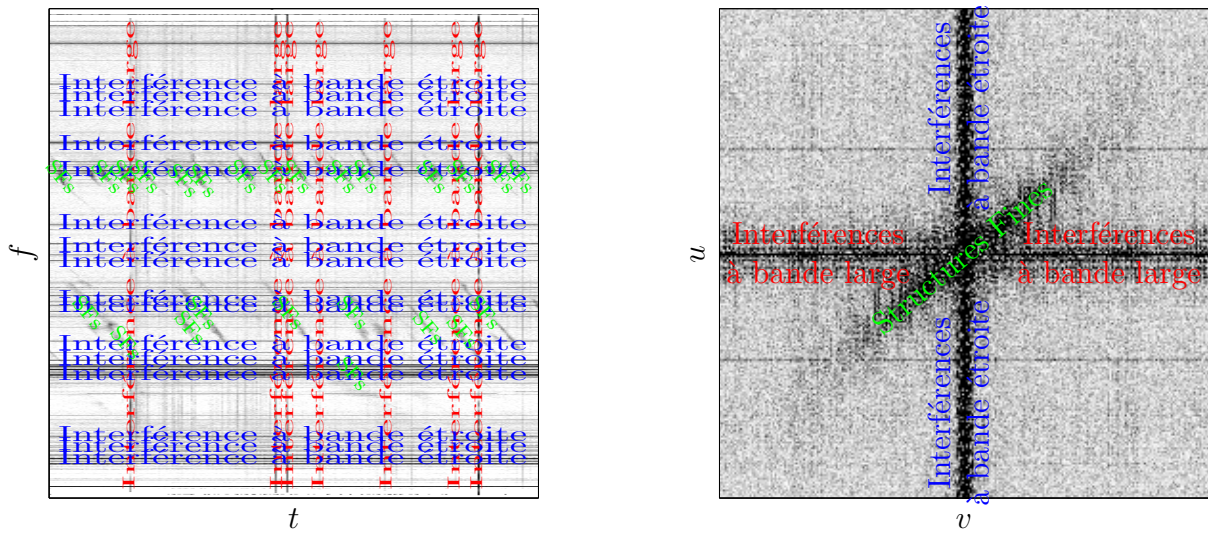


FIG. 4.2 – Principe du détecteur de structures fines : signaux et interférences sont mélangés dans le plan  $t$ - $f$  (à gauche), le passage dans le plan  $u$ - $v$  (à droite) permet leur séparation.

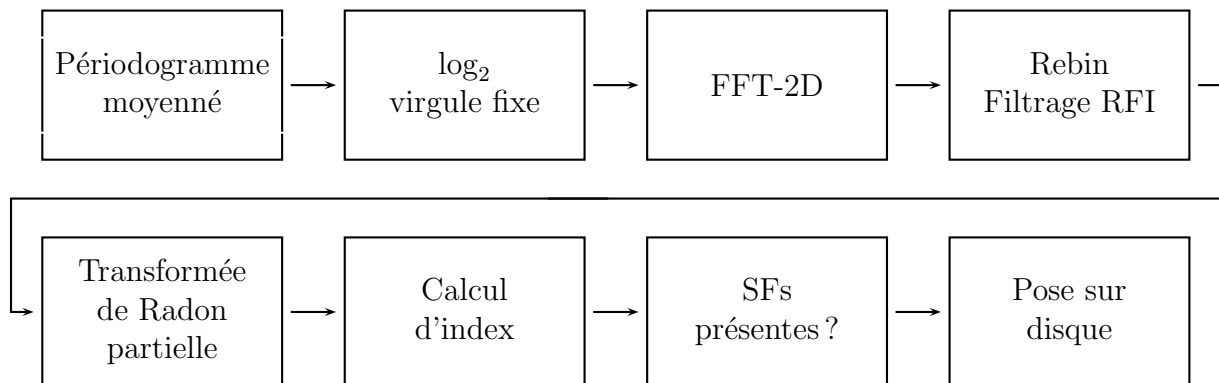


FIG. 4.3 – Algorithme du détecteur de structures fines FISSA.

sont calculés pour la future exploitation scientifique des données. Les résolutions temporelle  $\Delta t = 1.17$  ms et fréquentielle  $\Delta f = 6.8$  kHz induites ont été choisies de manière à exploiter scientifiquement les structures fines de façon satisfaisante.

Par contre, détecter les structures fines ne nécessite pas une si forte résolution fréquentielle. Ces valeurs sont même préjudiciables à la détection des signaux. Pour les signaux joviens, nous avons choisi  $\Delta t = 1.17$  ms et  $\Delta f = 55$  kHz qui affiche une dérive optimale d'un canal par spectre pour des signaux de dérive de  $47 \text{ MHz}\cdot\text{s}^{-1}$ . L'élargissement de la résolution spectrale est effectué par sommation successive de 8 canaux consécutifs au cours de leur arrivée dans le système. L'inter-produit complexe n'est pas traité pour la détection. Les plans temps-fréquence sont traités par blocs de 256 spectres, soient des imageries  $256 \times 256$  couvrant une durée de 300 ms. Les voies polarisées circulairement droite puis gauche sont alternativement traitées.

### 4.2.2 Calcul de $\log_2$ en virgule fixe et FFT-2D

La FFT-2D se calcule sur la base de la FFT-1D. La FFT-1D utilisée est issue de la “TMS320C62x DSP Library” [69] et prend comme opérandes des données 16 bits. Le traitement en échelle linéaire nécessiterait la troncature sur 16 bits de données codées sur 32 bits. Cette étape doit être effectuée de manière à conserver le bruit du ciel et quelques décibel de signaux supplémentaires contenant éventuellement des structures fines. Les signaux plus intenses, principalement des interférences, sont saturés. Pour éviter au système d’avoir à décider de la meilleure troncature à effectuer, on code les plans temps-fréquence en échelle logarithmique. Cela a aussi pour avantage de limiter l’influence des interférences saturées : assimilables à des fonctions “porte” d’amplitudes égales à la pleine échelle leurs transformées de Fourier occuperaient une large part du plan  $u-v$ . Le calcul du  $\log_2$  sans partie décimale, conduit par la simple recherche de la position du bit de poids fort, mène à une erreur RMS de 1.88 dB (annexe B) encadrée par les *extrema* [0 dB, +3 dB] : les structures fines de faible niveau pourraient être masquées.

Une méthode de calcul de  $\log_2$  en virgule fixe a été développée (annexe B: Calcul de  $\log_2$  avec partie décimale convergente, p.175). Il s’agit d’un calcul non-itératif utilisant une empreinte mémoire limitée. Nous avons choisi  $S = 5$  et  $Md = 5$ , menant à une erreur RMS de 0.0388 dB encadrée par  $[-0.0470 \text{ dB}, +0.0866 \text{ dB}]$ . Le calcul nécessite une table de  $2^S = 32$  mots de  $Md + 1 = 6$  bits.

Le code présenté en tableau B.1 effectue l’élargissement de la résolution fréquentielle d’un facteur 8, le codage en  $\log_2$  dans le format Q(11,5) et l’écriture du résultat sur 32 bits. On s’assure ainsi que les 16 bits de poids fort sont nuls : la FFT impose un tableau d’entrée composé de 256 paires mots de 16 bits codant partie réelle et partie imaginaire. Les deux composantes du nombre complexe sont initialisées en une unique écriture.

La FFT-2D est ensuite effectuée par deux séries de 256 FFT-1D intercalées avec une transposition. Aucune apodisation n’a été envisagée car les plans  $t-f$  sont naturellement atténués aux *extrema* fréquentiels d’une part et, d’autre part, le niveau de bruit du ciel, constant tout au long des 300 ms de données, limite la présence de discontinuités majeures.

### 4.2.3 Lissage du plan $u-v$ et atténuation des RFI

Le lissage de l’image  $u-v$  est effectué par le moyennage de blocs 4x4 pour arriver à une image 64x64. Il permet de réduire le bruit d’estimation et la taille du jeu de donnée à traiter lors de la prochaine étape. Il n’est pas possible d’effectuer un élargissement de résolution de cette ampleur avant la FFT-2D car les structures fines ne seraient pas résolues.

Afin de ne pas être perturbé par les interférences et puisque nous connaissons maintenant leur localisation majeure dans le plan  $u-v$ , la zone correspondant aux basses fréquences  $|u| < 5/64$  et  $|v| < 5/64$  est fixée à zéro (figure 4.4, à gauche). Elle n’est en fait même pas calculée, ce qui réduit d’autant le temps de calcul.

### 4.2.4 Transformée de Radon

La recherche de l’inclinaison de l’ellipse est effectuée par la localisation du maximum de l’image  $I$  contenue dans le plan  $r-\theta$  obtenu par la transformée de Radon du module carré lissé de la FFT-2D du  $\log_2$  du plan  $t-f$ .

La recherche de ce maximum est facilitée car nous connaissons la position de l'ellipse sur la dimension  $r$  (figure 4.4). Le calcul de la surface  $\mathcal{R}\{I\}(r, \theta)$  dans le plan  $r$ - $\theta$  complet n'est pas utile. On peut se limiter à son estimation pour  $r = 0$ , soit la courbe  $\mathcal{R}\{I\}_{r=0}(\theta)$ . La recherche se réduit à l'exploration de l'axe  $\theta$ .

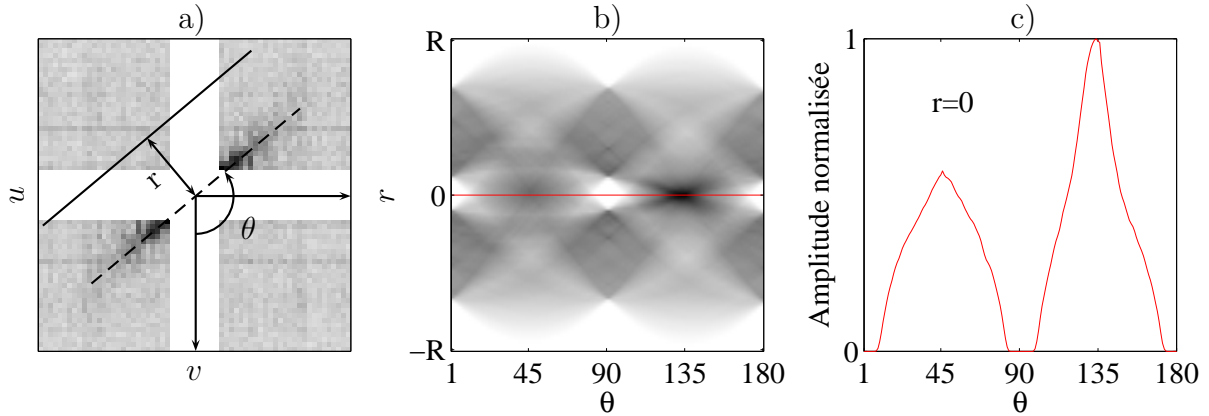


FIG. 4.4 – Réduction de l'espace de recherche dans le plan de Radon. a) La structure recherchée est localisée à  $r = 0$ , comme le montre la surface  $\mathcal{R}\{I\}(r, \theta)$  de b) issue de la transformée de Radon de l'image de a). On se limite en c) au calcul de la courbe  $\mathcal{R}\{I\}_{r=0}(\theta)$ .

Pour des soucis de rapidité d'exécution, aucun facteur de pondération n'est appliqué lors du calcul des intégrations aux divers angles choisis. Cela se manifeste en partie par les artéfacts situés aux angles  $45^\circ$  et  $135^\circ$ . Ces *maxima* locaux indésirables seront en partie gommés lors de la discrétisation de la transformée de Radon. Pour les mêmes soucis d'optimisation, les pixels à sommer sur l'angle  $\theta$  sont sélectionnés par la méthode du plus proche voisin. Leurs indices de tableau ne sont pas calculés par trigonométrie dans le DSP. Ils sont codés en dur dans un tableau. Ce tableau et toutes les informations nécessaires à la compilation du projet sont contenus dans un fichier ".h" généré par un script. La figure 4.5-a-b montre les positions des éléments à sommer pour calculer la transformée de Radon pour  $\theta = [0^\circ, 10^\circ, 20^\circ, \dots, 170^\circ]$ . Nous utiliserons à présent le terme de transformée de Radon pour désigner ce calcul par somme de la transformée de Radon "discrète"  $\mathcal{R}_{dis}\{I\}$  par opposition au terme de transformée de Radon "continue"  $\mathcal{R}_{cont}\{I\}$  qui fait usage d'intégrations :

$$\mathcal{R}_{cont}\{I\}(r, \theta) = \int_{\mathbb{R}} I(r \cos \theta - s \sin \theta, r \sin \theta + s \cos \theta) ds \quad (4.2)$$

$$\mathcal{R}_{cont}\{I\}_{r=0}(\theta) = \int_{\mathbb{R}} I(-s \sin \theta, s \cos \theta) ds \quad (4.3)$$

$$\mathcal{R}_{dis}\{I\}_{r=0}(\theta) = \sum_{s=1}^{63} I(\lfloor -s \sin \theta \rfloor, \lfloor s \cos \theta \rfloor) \quad (4.4)$$

Le nombre de points utilisés pour chaque sommation est constant ( $DC+2 \times 31 = 63$ ) pour éviter de normaliser les différentes sommes. Néanmoins, certains points sont toujours à zéros et pourraient être retirés du calcul. De même, il pourrait être plus juste de sommer

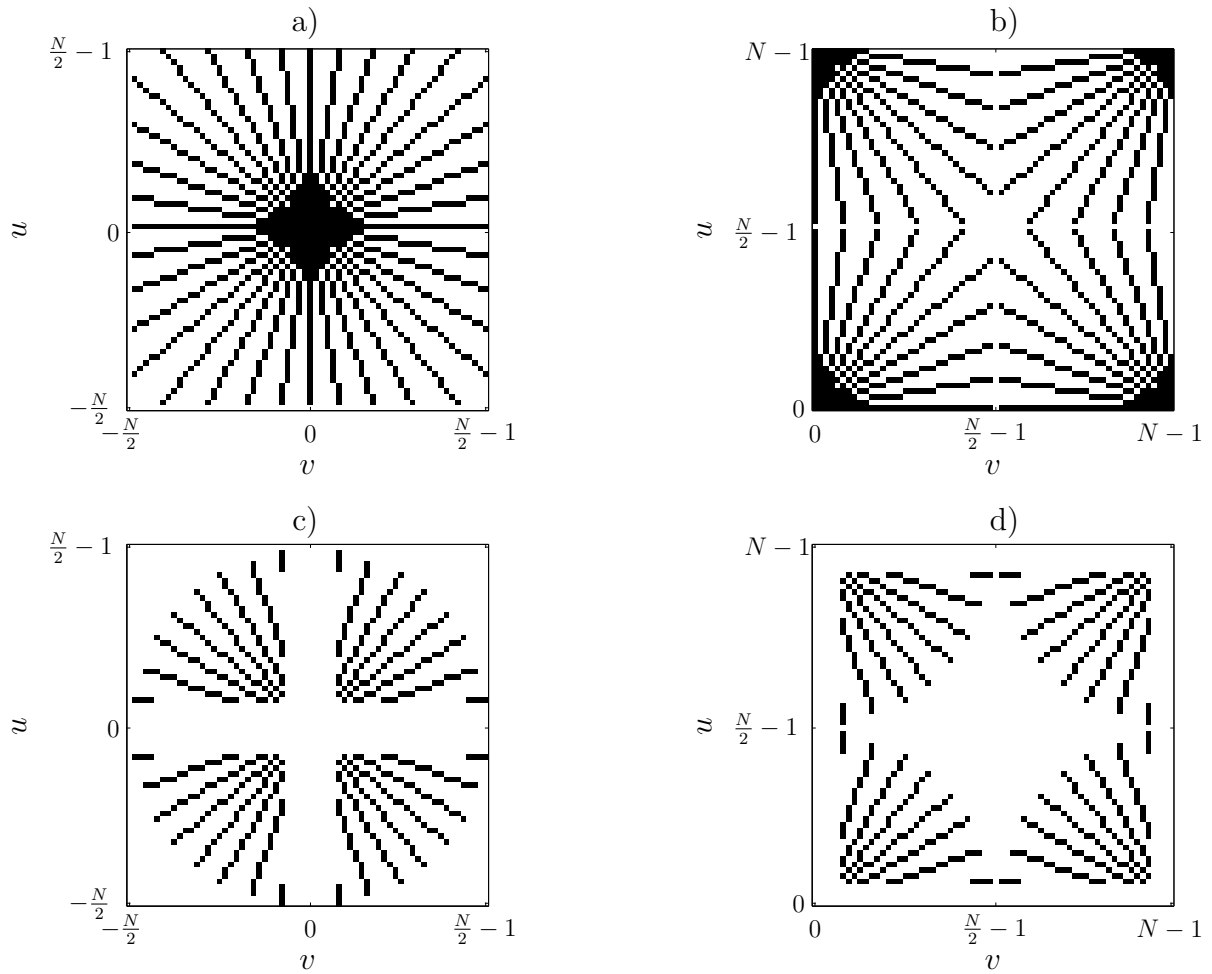


FIG. 4.5 – Calcul des indices de tableau de la transformée de Radon pour  $\theta = [0^\circ, 10^\circ, 20^\circ, \dots, 170^\circ]$ . a) les pixels noirs sont utilisés pour effectuer les sommations. b) Leur positions sont recalculées pour coïncider avec le format retourné par la FFT-2D. Mais certains points sont inutiles ou appartiennent aux interférences, on obtient alors c) et d).

à  $\sqrt{u^2 + v^2}$  constant plutôt qu'à  $\max(u, v)$ . En combinant ces deux optimisations, on obtient la figure 4.5-c-d. Le nombre de points à sommer passe de 1049 à 472.

Dans les cas où le pas  $\Delta\theta$  est petit, il a été envisagé de calculer la somme de manière itérative car une bonne partie des points sont communs d'un angle à l'autre. Pour le calcul de  $\mathcal{R}_{dis}^{itt} \{I\}_{r=0}(\theta_i)$  à l'angle  $\theta_i$ , on utilise la valeur précédemment calculée  $\mathcal{R}_{dis}^{itt} \{I\}_{r=0}(\theta_{i-1})$  à laquelle on retranche les points  $X_{\theta_{i-1} \setminus \theta_i}$ <sup>2</sup> appartenant à l'angle  $\theta_{i-1}$  mais n'appartenant plus à  $\theta_i$  et on ajoute les points  $X_{\theta_i \setminus \theta_{i-1}}$  de  $\theta_i$  qui n'appartenaient pas à  $\theta_{i-1}$  :

$$\begin{cases} \mathcal{R}_{dis}^{itt} \{I\}_{r=0}(\theta_0) &= \mathcal{R}_{dis} \{I\}_{r=0}(\theta_0) \\ \mathcal{R}_{dis}^{itt} \{I\}_{r=0}(\theta_i) &= \mathcal{R}_{dis}^{itt} \{I\}_{r=0}(\theta_{i-1}) + X_{\theta_i \setminus \theta_{i-1}} - X_{\theta_{i-1} \setminus \theta_i} \quad \text{pour } i > 0 \end{cases} \quad (4.5)$$

Nous avons montré que cette technique accélère le calcul quand l'angle  $\Delta\theta$  devient supérieur à  $\Delta\theta_{seuil}$  (figure 4.6) mais c'est malheureusement le moment où il devient inutile

<sup>2</sup>L'ensemble des atomes de  $A$  qui n'appartiennent pas à  $B$  est noté  $A \setminus B$

de diminuer encore la valeur de ce pas car la résolution du plan  $u-v$  limite la résolution angulaire à  $\Delta\theta_{\min} = \tan^{-1}\left(\frac{2}{N}\right)$ .

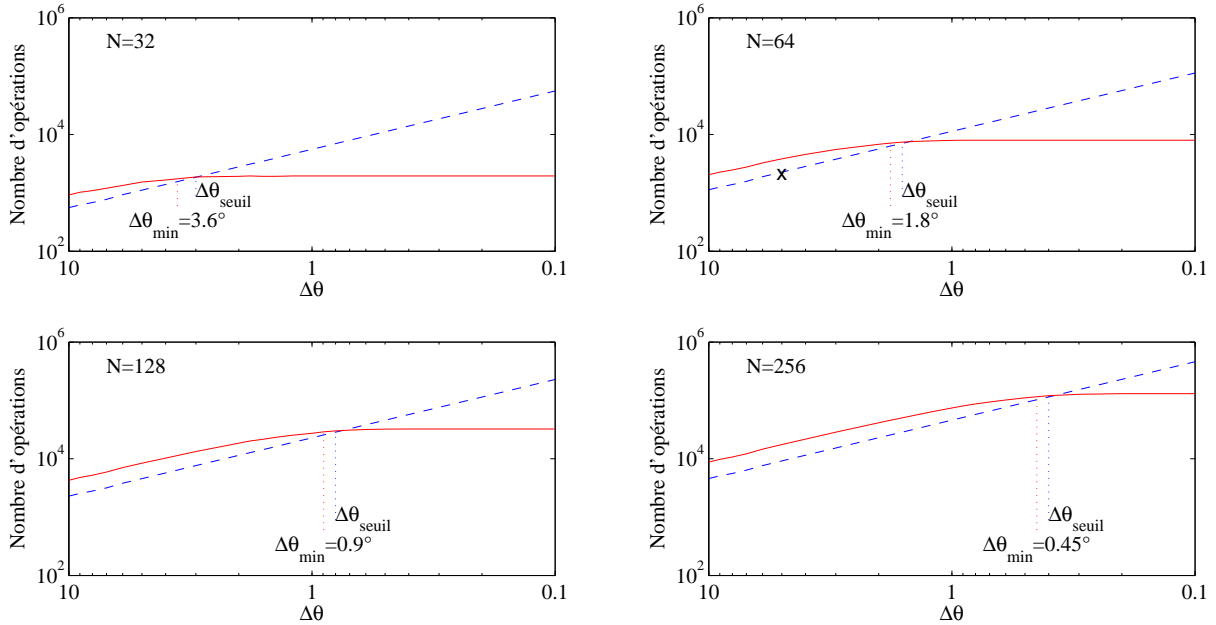


FIG. 4.6 – Accélération du calcul de la transformée de Radon (ligne bleue pointillée) par méthode itérative (ligne rouge pleine). La méthode itérative serait bénéfique pour  $\Delta\theta < \Delta\theta_{\text{seuil}}$  mais la limite de résolution angulaire  $\Delta\theta_{\min}$  représentable par une image  $N$  par  $N$  se trouve au-dessus de  $\Delta\theta_{\text{seuil}}$ .

L'application actuelle fonctionne avec une résolution angulaire de  $5^\circ$  (figure 4.6,  $\times$  en haut à droite). Cette valeur est un compromis satisfaisant entre coût de calcul et précision sur la position du maximum (figure 4.7).

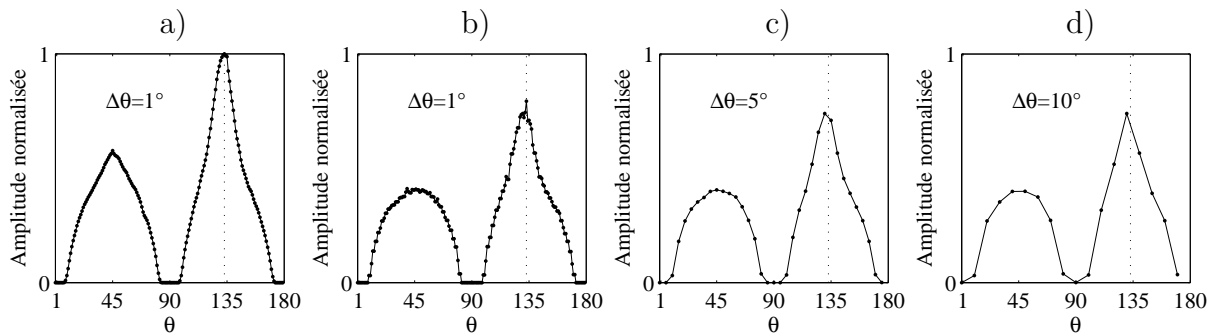


FIG. 4.7 – Effet de  $\Delta\theta$  sur l'estimation de la courbe  $\mathcal{R}_{dis}\{I\}_{r=0}(\theta)$  : a) Transformée de Radon “continue” de la figure 4.4-a pour  $\Delta\theta = 1^\circ$ . b,c,d) Transformée de Radon “discrète” pour  $\Delta\theta = [1^\circ, 5^\circ, 10^\circ]$

### 4.2.5 Calcul d'index

On calcule un index en se basant sur la valeur moyenne et la valeur du maximum de la courbe  $\mathcal{R}_{dis} \{I\}_{r=0}(\theta)$  :

$$SF_{Index} = \frac{\max(\mathcal{R}_{dis} \{I\}_{r=0}(\theta))}{\langle \mathcal{R}_{dis} \{I\}_{r=0}(\theta) \rangle} \quad (4.6)$$

Pour ne pas effectuer la division en virgule fixe sur le DSP, les deux valeurs sont envoyées au PC qui effectue le calcul. En absence de structures fines, la variable aléatoire  $SF_{Index}$  suit la distribution estimée empiriquement en figure 4.8 avec 14000 réalisations obtenues en 1.2 heures sur des données réelles (ciel + interférences de nuit, pas de S-Bursts). D'autres observations effectuées de jours ou de nuit (on note une variation diurne des caractéristiques des interférences) présentent des caractéristiques similaires. On fixe le seuil de détection à 2, soit une  $P_{fa}$  de 1.1%. Pour éviter les déclenchements dus à des événements isolés, on démarre l'enregistrement des données quand la proportion de détection dépasse 33% sur les 3.6 dernières secondes. L'enregistrement se termine quand la proportion de positifs reste inférieure à 33% pendant plus de 3 secondes. Quand un enregistrement prend fin, la pose des données sur disque est interdite pendant 3 secondes afin de laisser au système le temps de finaliser les écritures disques courantes. Cependant l'analyse de la bande continue d'être effectuée sans interruption.

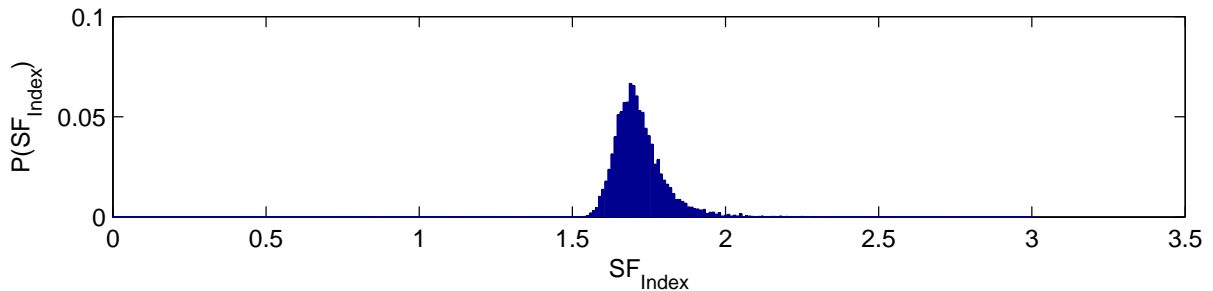


FIG. 4.8 – Estimation de la densité de probabilité de la variable aléatoire  $SF_{Index}$  en absence de structures fines.

La figure 4.9 résume le comportement de l'algorithme en présence ou en l'absence d'interférences.

### 4.2.6 Implantation matérielle

L'algorithme FISSA est exécuté par le noyau temps réel DSP-BIOS sur un DSP à virgule fixe TMS320C6203.

Trois tâches de traitement synchronisées par sémaphores sont utilisées ainsi que deux zones mémoires principales (figure 4.10). Une zone de mémoire donnée externe au DSP de 30 Mo sera utilisée pour stocker les données intermédiaires et une mémoire donnée interne au DSP de 17 ko permet d'effectuer les calculs. Les transferts entre mémoire interne et externe sont minimisés et effectués par bloc *via* des canaux DMA.

La première tâche effectuée la lecture des spectres haute résolution provenant du FPGA et place les données dans un buffer circulaire situé en mémoire externe.

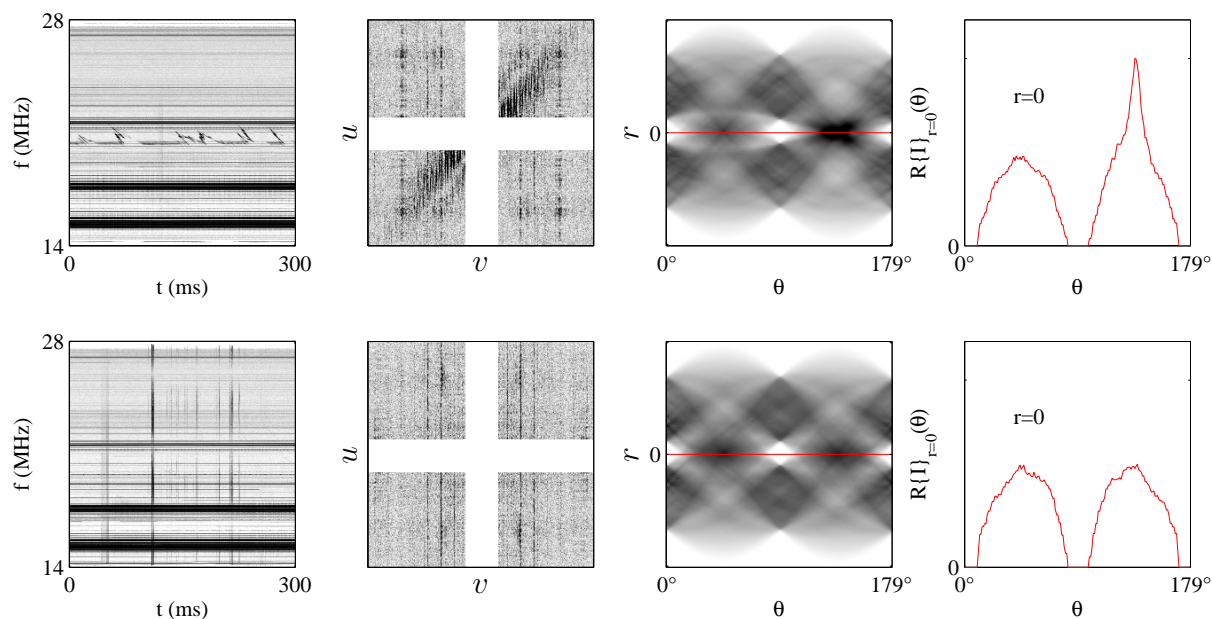


FIG. 4.9 – L’algorithme FISSA de détection de structures fines en présence (haut) et en absence de S-Burst (bas).

La deuxième tâche est responsable du fonctionnement de FISSA. Les données sont lues progressivement du buffer circulaire pour effectuer la première partie de la FFT-2D en mémoire interne. Les résultats sont replacés en mémoire externe jusqu’au moment où il est possible d’effectuer la seconde partie de la FFT-2D. Le calcul est de nouveau effectué progressivement et les résultats, de faibles volumes, peuvent maintenant être conservés en mémoire interne pour la transformée de Radon et le calcul de  $SF_{Index}$ .

Ces indicateurs sont insérés dans les plans temps-fréquence haute résolution et il est laissé aux soins de la troisième tâche d’effectuer le transfert des données vers le PC qui décidera de la pertinence de la pose sur disque des données.

### 4.3 Résultats

Une longue mise au point imputable à des erreurs logicielles a permis de produire une application stable. Elle a été utilisée pour la dernière fois au mois de décembre 2006 pour effectuer 4 observations (table 4.1 et figure 4.11) qui ont permis de récolter des signaux de structures fines.

Les enregistrements ont été effectués en parallèle avec des observations à basse résolution temporelle (1 spectre/s) afin d’annoter les types d’émissions rencontrées par le détecteur. Les émissions joviennes lentes, des types III solaires (*via* les lobes secondaires), les diodes à bruits de calibration système, les interférences à bande large et une grande majorité des interférences à bande étroite n’ont jamais fait déclencher l’enregistrement de façon intempestive.

Un type de signal a su tromper FISSA : le début des émissions de balises autoroutières se caractérisant par le balayage dans le plan temps-fréquence de la porteuse en cours de stabilisation. La faible fréquence d’occurrence de ces phénomènes dans la fenêtre d’analyse

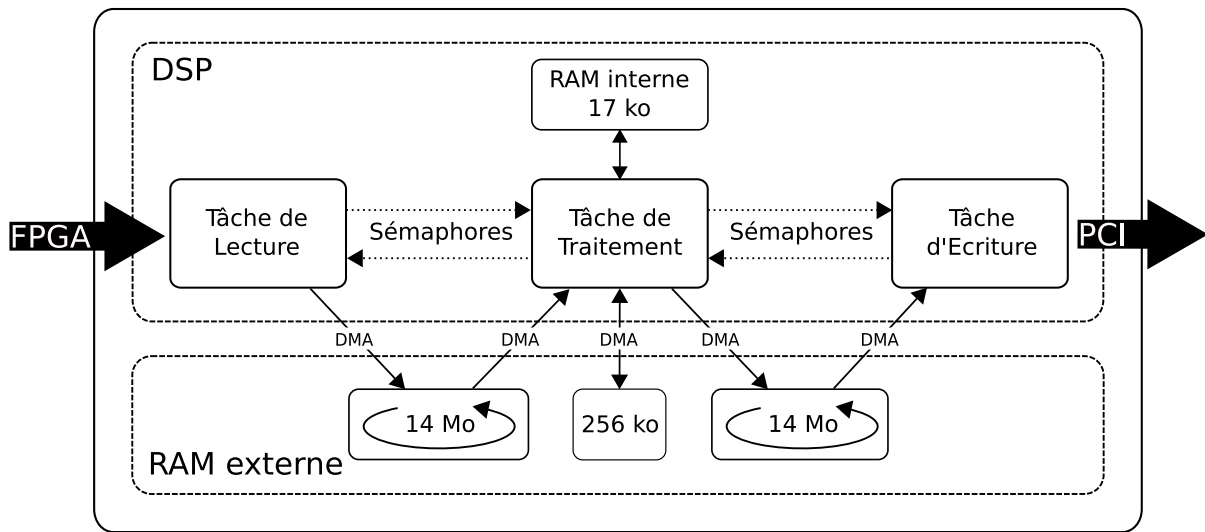


FIG. 4.10 – Ordonnancement des tâches et gestion mémoire de l'algorithme FISSA.

du calcul d'index (§ 4.2.5) a permis d'éliminer ce problème.

De même, en cas de très fortes intermodulations ou d'activités sources d'émissions à bande large quasi-continues (orages intenses, engins motorisés à proximité immédiate,...), la pseudo-périodicité des interférences produit de nombreuses raies dans le plan  $u-v$  qui trompent parfois FISSA suffisamment longtemps pour lancer une acquisition.

TAB. 4.1 – Exemples d'observation de structures fines. L'utilisation de FISSA permet de réduire l'enregistrement d'un flux constant de 50 Go/h à l'enregistrement réduit des seules données utiles.

N°	Date (TU)	Passage au méridien	Durée	Volume virtuel continu	Volume réel acquis	Volume réel utile
1	2006-12-02T07:06:00 2006-12-02T10:35:00	11:06	03:29:00	171 Go	4.6 Go	4.4 Go
2	2006-12-03T07:03:00 2006-12-03T11:30:00	11:03	04:27:00	219 Go	15.4 Go	13.8 Go
3	2006-12-09T06:45:00 2006-12-09T10:45:00	10:45	04:00:00	197 Go	3.1 Go	0.6 Go
4	2006-12-10T06:42:00 2006-12-10T11:42:00	10:45	05:00:00	246 Go	5.7 Go	0.5 Go

L'enregistrement en continu du signal produirait des volumes de données avoisinant 50 Go/h. L'utilisation du détecteur permet de réduire le volume de donnée au strict minimum (observation 1 et 2). Les observations 3 et 4 ont été menées dans des environnements très dégradés ce qui explique la forte quantité de données inutilement enregistrées. Ce sont néanmoins des volumes qui restent aisément maniables.



## 4.4 Perspectives

L'algorithme FISSA est suffisamment fonctionnel pour être utilisé en routine. Certains raffinements seraient néanmoins souhaitables.

Tout d'abord, il serait intéressant de pouvoir bénéficier d'une FIFO de grande profondeur (30-40 secondes) de façon à pouvoir enregistrer le signal avant que des structures fines ne soient clairement détectées. On pourrait ainsi avoir accès au début des émissions qui peuvent être faibles et présenter d'autres caractéristiques. De plus, cette FIFO permettrait de pouvoir continuer l'enregistrement même en cas d'atténuation forte du signal pendant des durées de plus de quelques secondes.

Pour des limitations purement techniques, les polarisations gauche et droite sont traitées alternativement par blocs de 300 ms. Le traitement simultané des deux voies pourrait permettre un fonctionnement plus robuste de la procédure de décision.

## PROBABILITE JUPITER A NANCAY

Decembre 2006

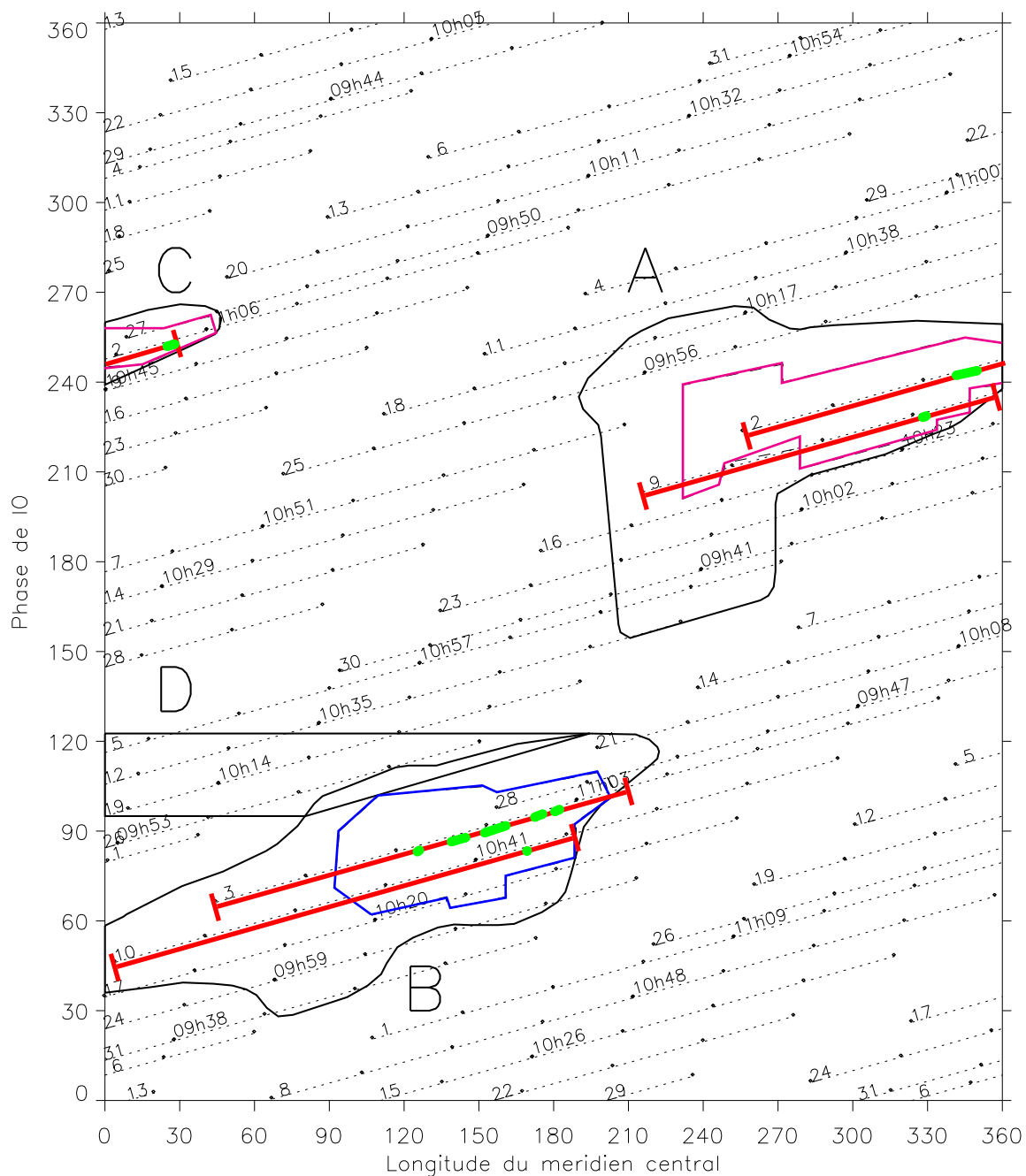


FIG. 4.11 – Exemple d'observations effectuées avec FISSA pendant le mois de décembre 2006. Le système de coordonnées de la figure est exposé dans [26, § 3.2.3]. Les observations ont été menées les 2, 3, 9 et 10 (lignes rouges). Des structures fines ont été détectées et enregistrées (lignes vertes) dans les géométries où leur probabilité d'occurrence est connue comme élevée (émissions contrôlées Io-C (cadre magenta) et Io-B (cadre bleu))



---

## CONCLUSION

---

Ce mémoire a présenté l'étude de méthodes d'implantation matérielle aisée qui permettent :

- de ré-observer dans des conditions où cela n'était plus possible,
- d'améliorer la qualité des observations dans le cas de conditions difficiles,
- de programmer de nouveaux types d'observations mis à disposition par les avancées récentes des systèmes numériques.

La première partie a exposé les bases nécessaires à la compréhension des observations effectuées en radioastronomie et des problèmes rencontrés.

La deuxième partie du document a présenté et traité du problème de classification inhérent aux méthodes de masquage (, ou blanking) basées sur un critère de puissance. La solution choisie pour ce problème est principalement conditionnée par l'estimation performante de la puissance de bruit du signal utile. Nous avons montré que la recherche de points aberrants dans des plans temps-fréquence (, ou tout autre jeu de données suivant une loi de distribution en  $\chi^2$ ) pouvait être effectuée par un seuillage calculé sur l'estimation seule de la moyenne. La contrainte essentielle est de pouvoir mener cette estimation en présence des points aberrants générés par les interférences radioélectriques. Les performances d'estimateurs classiques et d'algorithmes de notre cru ont été comparés sur des critères d'efficacité d'estimation (biais et variance), de robustesse aux interférences et de coût d'implantation matérielle. Une méthodologie d'analyse des performances a été clairement décrite pour facilement tester de nouveaux algorithmes d'estimation. On y évalue le niveau de qualité des critères choisis en fonction de la taille du jeu de données utilisé pour effectuer l'estimation, du taux de présence du brouilleur dans ce jeu de données et de son niveau de puissance relatif au bruit (INR). On a choisi de traiter le cas d'une interférence de niveau constant dans le jeu de données. La densité de probabilité de cette loi certaine est modélisée par un dirac. **Même si le modèle d'interférence choisit est minimaliste (caractérisé uniquement par son apport d'énergie), il ressort de cette étude que plusieurs estimateurs sont envisageables et que, pour optimiser les performances, l'estimateur choisi doit être adapté au taux de présence et au niveau d'interférence. Un arbre de sélection d'estimateurs a été dérivé selon des critères propres aux observations radioastronomiques effectuées.**

Les estimateurs étudiés dans la deuxième partie sont utilisés dans trois applications concrètes décrites au cours d'une troisième partie. Tous les travaux présentés ont fait l'objet d'une implantation matérielle ou logicielle et ils sont tous opérationnels. En vingt mois, les observations où ils sont impliqués totalisent 5000 h de fonctionnement.

Au Radiotélescope Décimétrique de Nançay, les outils développés permettent de ré-observer des objets astrophysiques dans la bande de l'opérateur de téléphonie satellitaire Iridium ( $\approx 1.6$  GHz) et dans les bandes allouées aux systèmes radar (1.2-1.4 GHz). Ces

nouveaux traitements vont faciliter un programme de 500 h d'observations de raies d'absorption fortement *red-shiftées* dans les bandes radar.

Le traitement des interférences radar est effectué par blanking de forme d'onde. Une stratégie de détection d'impulsions brèves mais intenses ou ténues mais longues a été mise en œuvre. L'implantation de cet algorithme dans un FPGA ne requiert que peu de ressources matérielles tout en assurant un fonctionnement à 100 MEch/s soit 400 Mo/s.

Le traitement des interférences Iridium est basé sur une analyse du plan temps-fréquence aux résolutions adaptées. Le masquage est décidé par seuillage. Un seuil doit être calculé pour chaque canal car la bande analysée n'est pas blanche. La principale contrainte est donc d'effectuer avec une barre d'erreur décente les 1600 estimations de moyenne suffisamment rapidement pour conserver un peu de temps d'observation. L'ensemble du traitement est implanté dans un DSP pour assurer le traitement des plans temps-fréquence au rythme de 14 Mo/s.

Ces travaux sont aussi utilisés systématiquement au Réseau Décamétrique de Nançay (10-80 MHz), en particulier pour effectuer des observations conjointes avec les missions satellitaires **STEREO** et **NEW HORIZONS**. La qualité des observations est améliorée par l'analyse à très haute résolution fréquentielle du signal de façon à résoudre les parasites à bande étroite qui constituent la majeure partie des problèmes rencontrés dans cette bande. Une résolution fréquentielle plus faible mais adaptée aux études astrophysiques et aux contraintes de stockage de données est obtenue par la sélection des meilleurs canaux issus de l'analyse qui a précédé. De nombreux émetteurs isolés disparaissent et le niveau au dessus du bruit du ciel des interférences les plus intenses passe de 60 dB à 30 dB.

Enfin, dans la quatrième partie, nous avons abordé la détection de signaux naturels dispersés en environnement parasite afin de déclencher l'enregistrement à haute résolution de l'ensemble. La statistique de décision est construite sur une FFT-2D et une transformée de Radon. Cette application innovante ouvre de nouvelles possibilités d'observations offertes par les récepteurs numériques reconfigurables. À terme, les observations systématiques à basses résolutions du Soleil et de Jupiter seraient avantageusement épaulées par une détection d'événements naturels brefs pour enregistrer, à haute résolution, des sursauts solaires rapides ou des S-Bursts joviens. On bénéficierait ainsi d'observations multi-échelles : en basse résolution (temporelle et fréquentielle) pour disposer du contexte radioélectrique (présence d'émissions naturelles exploitables en basse résolution, qualité de la bande radio, ...) et en haute résolution uniquement dans les zones du plan temps-fréquence présentant un intérêt pour l'astrophysique.

Nous nous sommes restreints dans notre étude au cas d'observations effectuées avec une antenne unique en présence d'interférences peu étendues en temps et/ou en fréquence. Néanmoins, les cas plus complexes d'observations avec antennes multiples pour de l'interférométrie et/ou en présence d'interférences continues et à bande large n'ont pas ici été abordés et sont loin d'être résolus pour des fonctionnements en routine. Dans le champ d'applications sur antenne unique, l'utilisation d'autres statistiques de décision, comme la polarisation ou la cyclostationnarité par exemple, pourraient permettre de séparer interférences et sources radioastronomiques avec plus d'efficacité qu'un critère de puissance.

En effet, le détecteur le plus robuste est aussi celui qui intègre à son fonctionnement le moins d'information sur le signal à détecter. Mais on peut améliorer les performances du détecteur en lui fournissant ces informations. Bien entendu, on restreint la réponse du détecteur aux signaux pour lesquels il a été construit.

Un critère de décision envisagé pour détecter des interférences fut l'information de polarisation. Cette information est obtenue en corrélant deux antennes de polarisations orthogonales. On peut l'utiliser pour calculer un taux de corrélation qui est un indicateur du degré de polarisation du signal. Les sources radioastronomiques présentent des taux de polarisation variables (et donc des taux de corrélation variables) mais les signaux d'origine artificielle affichent des taux relativement forts. Calculer le taux de corrélation dans le plan temps-fréquence et masquer les données présentant des taux trop élevés serait une alternative au blanking contrôlé par un critère de puissance. L'expérimentation de cette méthode n'a pas été envisagée au cours de cette thèse car le récepteur n'est pas facilement configurable pour calculer des taux de corrélation en temps-réel. Néanmoins, sur des flux de données faibles (comme pour l'application du traitement des interférences à bande étroite sur le NDA, section 3.3) une implantation dans le PC d'acquisition est envisageable.

Il est aussi possible de simplifier les calculs tout en conservant la même philosophie : plutôt que de traiter le taux de corrélation comme classiquement défini dans la littérature, on peut envisager de ne traiter que la phase de la corrélation. Bonne partie des sources radioastronomiques ne sont pas polarisées ce qui équivaut à dire que l'angle de polarisation observé est aléatoire pour ces radiosources. Si dans un jeu de données, la phase ne suit plus une loi de distribution uniforme, on peut conclure que nous sommes en présence d'une source polarisée dont l'angle est donné par le mode de la distribution observée. Il faut néanmoins être sûr que la zone observée ne contient pas de sources naturelles polarisées sous peine de les voir, elles aussi, disparaître.

Nous avons vu que puissance et polarisation étaient des critères communs aux radiosources naturelles et aux émetteurs artificiels. Il serait évidemment souhaitable de disposer d'un critère qui ne soit caractéristique que de l'une de ces deux classes de signaux afin d'effectuer la classification avec la plus grande confiance possible. Un tel critère existe si on met de côté le cas des pulsars, il s'agit de la cyclostationnarité. On peut classer les signaux dans deux catégories : ceux dont les caractéristiques sont stables au cours du temps et que l'on qualifie de stationnaires et ceux dont les caractéristiques évoluent au cours du temps et que l'on qualifie de non stationnaires. Si les caractéristiques du signal évoluent de manière périodique (l'inverse d'une fréquence cyclique), on peut qualifier ces signaux de cyclostationnaires. Un pulsar, qui trahit sa présence par une émission pulsée de période extrêmement stable entre donc dans cette catégorie. Les signaux artificiels possèdent intrinsèquement cette propriété. On retrouve des périodicités dans les impulsions radars, dans la fréquence d'émission des symboles des télécommunications numériques et même dans les modulations analogiques. Des méthodes de blanking contrôlées par cyclostationnarité ainsi que des méthodes plus audacieuses de soustraction ou d'annulation d'interférences ont été étudiées [8, 56, 3, 7, 57, 52], tant sur des données simulées que sur des données réelles mais leur coût de calcul important a considérablement limité les implantations matérielles. De plus, l'utilisation de telles méthodes implique de conditionner l'observation autour des interférences à éliminer : la radio-source (qui reste pour l'astrophysicien le cœur de ses recherches) passe au second plan !

C'est pour cette principale raison qu'un minimum d'automatisme doit être inclus dans les nouveaux récepteurs numériques intégrant des méthodes de traitement des interférences. Même si les radioastronomes sont familiers avec les problèmes induits par les interférences radio, ils n'ont pas à devenir experts en télécommunication pour ajuster les configurations des observations. Il serait utile de disposer d'un système qui choisirait les algorithmes à appliquer tout en informant l'utilisateur du type d'altération qui pourrait corrompre les signaux de son sujet d'étude. La sélection des algorithmes à utiliser pourrait se baser dans un premier temps sur le rapport de l'ERC (European Radiocommunications Committee). Le choix des méthodes pourrait ensuite s'affiner en fonction des informations remontées par les observations et les programmes de surveillance.

## ANNEXE A

---

# RAPPORT D'IMPLANTATION EN FPGA DU "BLANKEUR" RADAR

---

HDL Synthesis Report:

Macro Statistics:

-----

Registers	: 91
1-bit register	: 84
12-bit register	: 1
16-bit register	: 2
32-bit register	: 2
5-bit register	: 2
Counters	: 1
32-bit up counter	: 1
Adders/Subtractors	: 7
12-bit adder	: 1
16-bit adder	: 2
32-bit adder	: 1
32-bit subtractor	: 1
5-bit adder	: 2
multiplier	: 4
16x16 multiplier	: 2
32x16 multiplier	: 1
32x8 multiplier	: 1
Comparators	: 34
32-bit comparator less	: 2
5-bit comparator greater	: 32

Device utilization summary:

-----

Selected Device : 2v3000ff1152-5

Logics:

Number of Slices: 249 out of 14336 1.7%



Number of Slice Flip Flops:	275 out of 28672	1%
Multipliers:		
Number of Slices:	392 out of 14336	2.7%
Number of Slice Flip Flops:	776 out of 28,672	2.7%
Number of MULT18X18s:	2 out of 96	2%

**Timing Summary:**

-----

Speed Grade: -5  
Minimum period: 6.863ns (Maximum Frequency : 145.709MHz)  
Minimum input arrival time before clock : 6.189ns  
Maximum output required time after clock : 4.840ns  
Maximum combinational path delay : 2.008ns

## ANNEXE B

---

# CALCUL DE $\log_2$ EN VIRGULE FIXE

---

### Problème

Nous souhaitons calculer la densité spectrale de puissance (DSP) d'un signal, en dB, au cours du temps. Le calcul de la DSP en échelle linéaire  $X_t(k)$  est effectué par une FFT à  $N_{Chan}$  canaux dans un composant programmable de type FPGA.

$$X_t(k) = \left| \sum_{j=1}^{N_{Chan}} x(j) \omega_{N_{Chan}}^{(j-1)(k-1)} \right|^2 \quad \text{avec } \omega_{N_{Chan}} = \exp^{-2\pi i / N_{Chan}} \quad (\text{B.1})$$

Les calculs sont conduits en virgule fixe sur des mots non-signés de 32 bits et le fort débit de données à traiter en temps réel impose une routine optimisée.

Il s'agit donc d'évaluer  $P_t(k)$  pour  $k = [1, \dots, N_{Chan}]$  et aux instants  $t = 1, \dots, t_0$  :

$$P_t(k) = 10 \log_{10}(X_t(k)) \quad \text{pour } X_t(k) > 0 \quad (\text{B.2})$$

Comme les calculs sont effectués en binaire, le logarithme en base 2 ( $\log_2$ ) est plus approprié.

$$\begin{aligned} P_t(k) &= 10 \left( \frac{\log_2(X_t(k))}{\log_2(10)} \right) \\ P_t(k) &= \frac{10}{\log_2(10)} \log_2(X_t(k)) \\ P_t(k) &\approx 3.0103 \log_2(X_t(k)) \end{aligned}$$

La suite de cette annexe étudie les effets du calcul en virgule fixe de  $\log_2$ . Une première partie décrit l'erreur obtenue quand la partie décimale du résultat est inexploitée. La deuxième expose une méthode facilement implantable pour évaluer la partie décimale du  $\log_2$  ainsi que les calculs d'erreur associés. La dernière partie apporte une légère complexification à l'algorithme qui permet de réduire encore l'erreur de calcul.

## Calcul de log<sub>2</sub> sans partie décimale

Nous pouvons dans un premier temps approcher le résultat  $R_{\log_2}$  du calcul du log<sub>2</sub> en tronquant la partie décimale du résultat du calcul effectué en flottant pour ne garder que la partie entière  $E_{\log_2}$ .

$$\begin{aligned} \lfloor \log_2(x) \rfloor &= E_{\log_2} \\ \log_2(x) &= E_{\log_2} + \epsilon_{\log_2}(x) \quad \text{pour } x \geq 1 \\ 10 \log_{10}(x) &= 3.0103 E_{\log_2} + \epsilon_{\log_{10}}(x) \quad \text{pour } x \geq 1 \\ \epsilon_{\log_{10}}(x) &= 3.0103 \epsilon_{\log_2}(x) \end{aligned}$$

La quantification du résultat peut alors s'écrire  $Q(\log_2(M_{E_{max}}), 0)$ <sup>1</sup> avec  $M_{E_{max}} = \lfloor \log_2(x_{max}) \rfloor$ .  $M_{E_{max}}$  est le nombre de niveau minimum nécessaire pour représenter la partie entière du log<sub>2</sub> d'un nombre. Il faut  $\log_2(M_{E_{max}})$  bits pour représenter  $M_{E_{max}}$ . Pour stocker le résultat du log<sub>2</sub> d'un mot codé sur 32 bits,  $M_{E_{max}}$  vaut 32 donc le résultat du calcul peut être stocké dans un format  $Q(5, 0)$ . L'implantation matérielle équivaut à trouver dans le mot de 32 bits la position  $E_{\log_2}$  du premier bit de poids fort à 1.

Dérivons les bornes de  $\epsilon_{\log_{base}}(x)$  ainsi que l'erreur quadratique moyenne  $RMS(\epsilon_{\log_{base}}) = \sqrt{\langle \epsilon_{\log_{base}}^2(x) \rangle}$ . Nous pouvons restreindre l'étude de  $\epsilon_{\log_{base}}(x)$  dans l'intervalle  $[2^{E_{\log_2}}, 2^{E_{\log_2}+1}]$  :

$$\begin{aligned} \epsilon_{\log_2}(x) &= \log_2(x) - E_{\log_2} \\ \epsilon_{\log_2}(x) &= \log_2\left(\frac{x}{2^{E_{\log_2}}}\right) \\ 2^{E_{\log_2}} &\leq x < 2^{E_{\log_2}+1} \\ 0 &\leq \epsilon_{\log_2}(x) < 1 \\ 0 &\leq \epsilon_{\log_{10}}(x) < 3.0103 \text{ dB} \end{aligned} \tag{B.3}$$

$$\begin{aligned} RMS(\epsilon_{\log_2}) &= \sqrt{\frac{1}{2^{E_{\log_2}+1} - 2^{E_{\log_2}}} \int_{2^{E_{\log_2}}}^{2^{E_{\log_2}+1}} [\epsilon_{\log_2}(x)]^2 dx} \\ RMS(\epsilon_{\log_2}) &= \sqrt{\frac{1}{2^{E_{\log_2}} \log_{10}^2(2)} \int_{2^{E_{\log_2}}}^{2^{E_{\log_2}+1}} \left[ \log_{10}\left(\frac{x}{2^{E_{\log_2}}}\right) \right]^2 dx} \\ RMS(\epsilon_{\log_2}) &= \sqrt{\frac{1}{2^{E_{\log_2}} \log_{10}^2(2)} \left[ x \left( \log_{10}^2\left(\frac{x}{2^{E_{\log_2}}}\right) - 2 \log_{10}\left(\frac{x}{2^{E_{\log_2}}}\right) + 2 \right) \right]_{2^{E_{\log_2}}}^{2^{E_{\log_2}+1}}} \\ RMS(\epsilon_{\log_2}) &= \sqrt{2} \frac{|\log_{10}(2) - 1|}{\log_{10}(2)} \\ RMS(\epsilon_{\log_2}) &= \sqrt{2} (\log_2(e) - 1) \\ RMS(\epsilon_{\log_2}) &\approx 0.6261 \\ RMS(\epsilon_{\log_{10}}) &\approx 1.8846 \text{ dB} \end{aligned}$$

<sup>1</sup> $Q(a, b)$  est la représentation entière d'un nombre sur  $a$  bits dont  $b$  bits pour la partie décimale

## Calcul de $\log_2$ avec partie décimale

Pour certaines applications, une erreur quadratique moyenne de 1.9 dB est rédhitoire. Chaque intervalle  $[2^{E_{\log_2}}, 2^{E_{\log_2}+1}[$  est donc subdivisé en  $2^S$  niveaux égaux (voir figure B.1) pour augmenter la finesse de la quantification de  $x$  et pour ainsi augmenter le nombre de niveaux de représentation du  $\log_2$ .

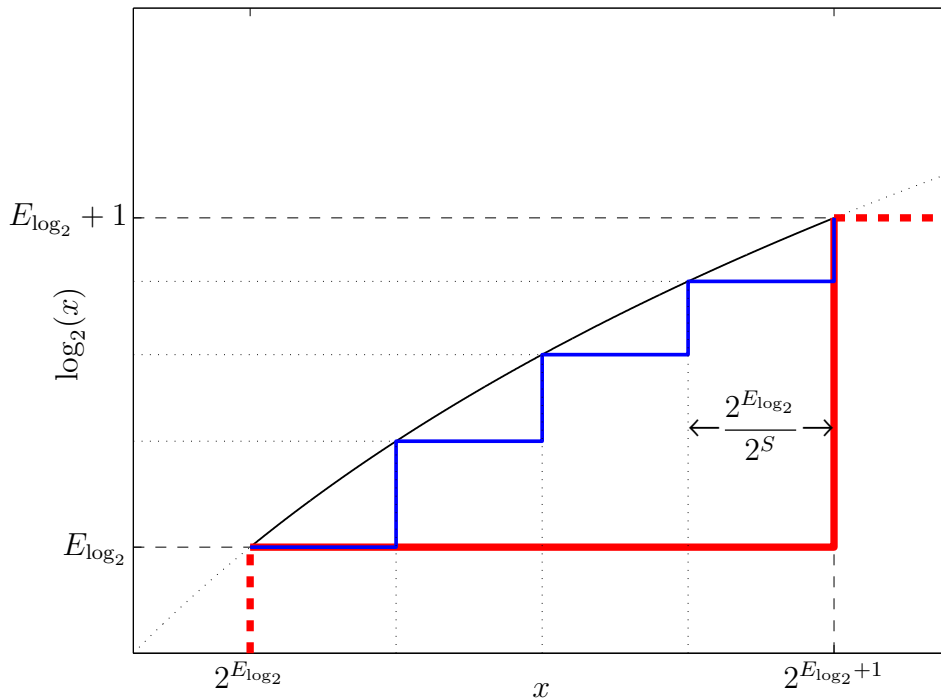


FIG. B.1 – Approximation de  $\log_2(x)$  (en noir fin) pour  $S = 0$  (en rouge très gras) et  $S = 2$  (en bleu gras)

Il faut, en contrepartie, augmenter la précision du résultat  $R_{\log_2}$  du calcul du  $\log_2$  en ajoutant à la partie entière  $E_{\log_2}$  une partie décimale  $D_{\log_2}$  codée sur  $M_D$  bits. La quantification du résultat  $R_{\log_2}$  s'écrit maintenant  $Q(\lfloor \log_2(M_{E_{max}}) \rfloor + 1 + M_D, M_D)$  soit  $Q(5 + M_D, M_D)$  pour stocker le résultat du  $\log_2$  d'un mot sur 32 bits. Il faut noter que le format de stockage de  $R_{\log_2}$  dépend, comme précédemment, de la dynamique de  $x$  mais aussi de la taille  $M_D$  en bit de  $D_{\log_2}$  requise pour stocker avec la précision voulue les  $2^S$  subdivisions de  $2^{E_{\log_2}}$ .

$$\begin{aligned}
 x &= 2^{E_{\log_2} + D_{\log_2}(S, M_D)} + \xi \\
 \log_2(x) &= E_{\log_2} + D_{\log_2}(S, M_D) + \epsilon_{\log_2}(x) \quad \text{pour } x > 0 \\
 \text{avec } E_{\log_2} &= \lfloor \log_2(x) \rfloor \\
 \text{et } D_{\log_2}(S, M_D) &= 2^{-M_D} \left[ 2^{M_D} \log_2 \left( 1 + \left\lfloor \frac{x - 2^{E_{\log_2}}}{2^{E_{\log_2}}} \cdot 2^S \right\rfloor 2^{-S} \right) \right]
 \end{aligned}$$

L'implantation matérielle du calcul de  $D_{\log_2}(S, M_D)$  se résume à deux étapes:

- Masquer les  $S$  bits qui suivent le bit  $E_{\log_2}$  dans  $k$ . Cette opération évalue  $k = \left\lfloor \frac{x - 2^{E_{\log_2}}}{2^{E_{\log_2}}} \cdot 2^S \right\rfloor$ .

- Utiliser  $k$  pour entrer dans une table  $Tab\_D_{\log_2}^{S,M_D}[k]$  contenant les valeurs tabulées, arrondies sur  $M_D$  bits, de  $D_{\log_2}(S, M_D) = \log_2(1 + k \cdot 2^{-S})$  pour  $k \in [0, 1, \dots, 2^S[$ .

Il suffit ensuite de concaténer  $E_{\log_2}$  et  $D_{\log_2}(S, M_D)$  dans  $R_{\log_2}$ .

Le choix de  $S$  et  $M_D$  doit se faire judicieusement pour atteindre la précision voulue tout en ne sur-spécifiant pas un paramètre par rapport à l'autre. L'erreur  $\epsilon_{\log_2}^{S,M_D}(x)$  s'écrit maintenant :

$$\epsilon_{\log_2}^{S,M_D}(x) = \log_2(x) - (E_{\log_2} + D_{\log_2}(S, M_D))$$

Dans un premier temps, il faut déterminer la valeur de  $S$  nécessaire à l'obtention d'un pas d'échantillonnage suffisamment fin pour  $x$ . Nous faisons tendre  $M_D \rightarrow \infty$ , ce qui correspond à une représentation en précision infinie. L'erreur peut se récrire comme suit :

$$\begin{aligned} \epsilon_{\log_2}^{S,\infty}(x) &= \log_2(x) - (E_{\log_2} + D_{\log_2}(S, \infty)) \\ \epsilon_{\log_2}^{S,\infty}(x) &= \log_2(x) - (E_{\log_2} + Tab\_D_{\log_2}^{S,\infty}[k]) \\ \epsilon_{\log_2}^{S,\infty}(x) &= \log_2(x) - (E_{\log_2} + \log_2(1 + k \cdot 2^{-S})) \\ \epsilon_{\log_2}^{S,\infty}(x) &= \log_2\left(\frac{x}{2^{E_{\log_2}}(1 + k \cdot 2^{-S})}\right) \end{aligned}$$

L'étude de  $\epsilon_{\log_{base}}^{S,\infty}(x)$  peut se restreindre dans l'intervalle  $I = [2^{E_{\log_2}}, 2^{E_{\log_2}+1}[$  qui peut lui-même se décomposer sous la forme  $\{(1 + k \cdot 2^{-S}) 2^{E_{\log_2}}, (1 + (k + 1) 2^{-S}) 2^{E_{\log_2}}\}$ ,  $k = [0, \dots, 2^S - 1]$ . Ceci implique les expressions suivantes pour les bornes de  $\epsilon_{\log_{base}}^{S,\infty}(x)$  et l'erreur quadratique moyenne  $RMS(\epsilon_{\log_{base}}^{S,\infty})$  :

$$\begin{aligned} (1 + k \cdot 2^{-S}) 2^{E_{\log_2}} &\leq x < (1 + (k + 1) 2^{-S}) 2^{E_{\log_2}} \\ 0 \leq \epsilon_{\log_2}^{S,\infty}(x) &< \log_2\left(\frac{2^S + (k + 1)}{2^S + k}\right) < \log_2(1 + 2^{-S}) \quad \text{pour } k = 0 \\ 0 \leq \epsilon_{\log_{10}}^{S,\infty}(x) &< 3.0103 \log_2(1 + 2^{-S}) \quad \text{pour } k = 0 \end{aligned} \tag{B.4}$$

$$\begin{aligned}
RMS(\epsilon_{\log_2}^{S,\infty}) &= \sqrt{\frac{1}{2^{E_{\log_2}+1} - 2^{E_{\log_2}}} \int_{2^{E_{\log_2}}}^{2^{E_{\log_2}+1}} [\epsilon_{\log_2}(x)]^2 dx} \\
RMS(\epsilon_{\log_2}^{S,\infty}) &= \sqrt{2^{-E_{\log_2}} \sum_{k=0}^{2^S-1} \int_{2^{E_{\log_2}+k} \cdot 2^{E_{\log_2}-S}}^{2^{E_{\log_2}+(k+1)} \cdot 2^{E_{\log_2}-S}} [\epsilon_{\log_2}(x)]^2 dx} \\
RMS(\epsilon_{\log_2}^{S,\infty}) &= \sqrt{2^{-E_{\log_2}} \sum_{k=0}^{2^S-1} \int_{(1+k \cdot 2^{-S}) 2^{E_{\log_2}}}^{(1+(k+1) 2^{-S}) 2^{E_{\log_2}}} \left[ \begin{array}{l} \log_2(x) \\ -E_{\log_2} \\ -\log_2 \left( 1 + \left[ \frac{x - 2^{E_{\log_2}}}{2^{E_{\log_2}}} \cdot 2^S \right] 2^{-S} \right) \end{array} \right]^2 dx} \\
RMS(\epsilon_{\log_2}^{S,\infty}) &= \sqrt{2^{-E_{\log_2}} \sum_{k=0}^{2^S-1} \left[ x \left( \begin{array}{l} (\log_2(x) - \log_2(e) - R_{\log_2}(k))^2 \\ + \log_2^2(e) \end{array} \right) \right]_{(1+k \cdot 2^{-S}) 2^{E_{\log_2}}}^{(1+(k+1) 2^{-S}) 2^{E_{\log_2}}} \\
&\quad \text{avec } R_{\log_2}(k) = E_{\log_2} + Tab\_D_{\log_2}^{S,\infty}[k] \\
RMS(\epsilon_{\log_2}^{S,\infty}) &= \sqrt{\sum_{k=0}^{2^S-1} \left[ x \left( \left( \log_2(x) - \log_2(e) - Tab\_D_{\log_2}^{S,\infty}[k] \right)^2 + \log_2^2(e) \right) \right]_{1+k \cdot 2^{-S}}^{1+(k+1) 2^{-S}}}
\end{aligned}$$

Pour  $S = 0$ , on retrouve  $RMS(\epsilon_{\log_2}^{0,\infty}) = RMS(\epsilon_{\log_2}) = \sqrt{2}(\log_2(e) - 1)$ . L'erreur  $RMS(\epsilon_{\log_2}^{S,\infty})$  peut s'approcher par l'expression suivante :

$$RMS(\epsilon_{\log_2}^{S,\infty}) \approx \left[ \sqrt{2}(\log_2(e) - 1) \right] 2^{-S}$$

Dans un second temps,  $M_D$  est ajusté afin que l'erreur de quantification du résultat reste négligeable face à l'erreur du pas d'échantillonnage précédemment déterminé par  $S$ . L'expression exacte de  $RMS(\epsilon_{\log_2}^{S,M_D})$  est donnée par :

$$RMS(\epsilon_{\log_2}^{S,M_D}) = \sqrt{\sum_{k=0}^{2^S-1} \left[ x \left( \left( \log_2(x) - \log_2(e) - Tab\_D_{\log_2}^{S,M_D}[k] \right)^2 + \log_2^2(e) \right) \right]_{1+k \cdot 2^{-S}}^{1+(k+1) 2^{-S}}}$$

$RMS(\epsilon_{\log_2}^{S,M_D})$  est tabulé en figure B.2 pour  $S \in [0, \dots, 10]$  et  $M_D \in [0, \dots, 16]$ . Les lignes isocontours de la taille de mémoire requise pour stocker la table  $Tab\_D_{\log_2}^{S,M_D}[k]$  sont superposées à la surface d'erreur.

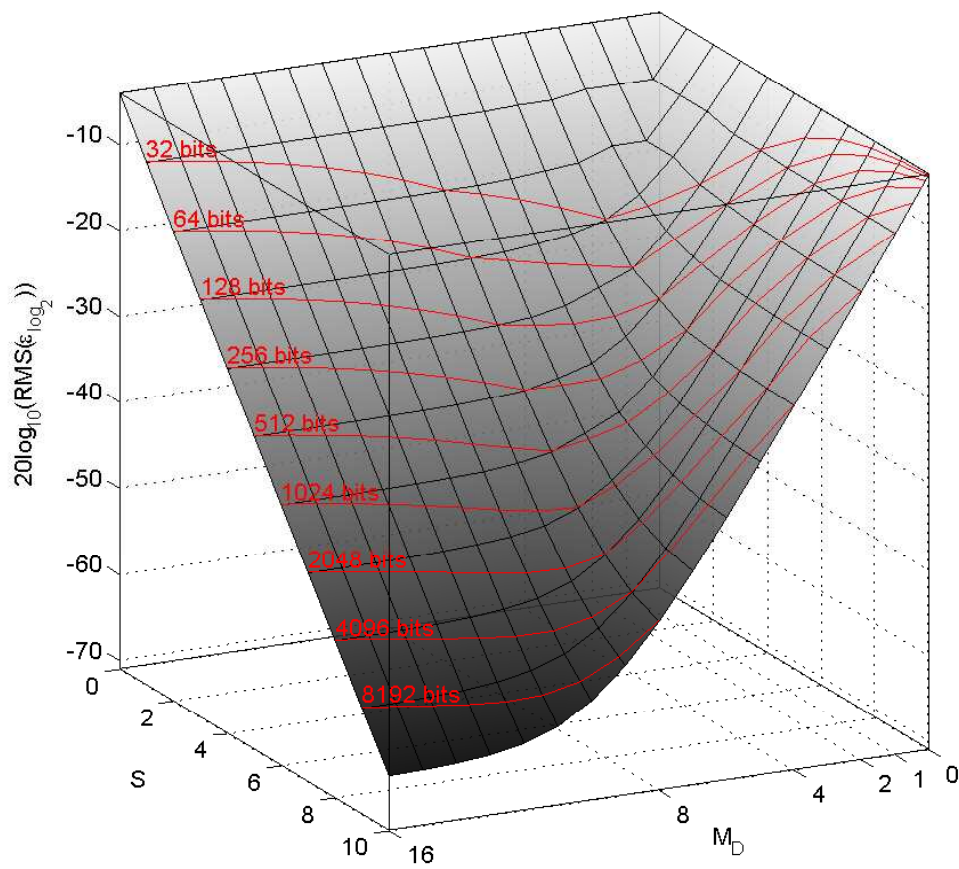


FIG. B.2 – Tabulation de l'erreur  $\text{RMS}(\epsilon_{\log_2}^{S, M_D})$  du calcul de  $\log_2$  en fonction de  $S$  et  $M_D$ . Les lignes rouges transversales sont les isocontours de la taille de la mémoire requise pour stocker la table.

## Calcul de $\log_2$ avec partie décimale convergente

Il est encore possible de réduire l'erreur sans augmenter l'empreinte mémoire en utilisant les arrondis convergents pour coder  $Tab\_Conv\_D_{\log_2}^{S,M_D}[k]$ . Cette table contient les valeurs  $2^{-(S+1)} + \log_2(1 + k \cdot 2^{-S})$  codées sur  $(M_D + 1)$  bits pour  $k \in [0, 1, \dots, 2^S]$ .

La méthode d'approximation de  $\log_2(x)$  diffère légèrement de la précédente comme montré en figure B.3. Maintenant, les résultats ne sont plus systématiquement sous-estimés si l'on compare B.3 et B.4 avec le nouvel encadrement d'erreur :

$$\begin{aligned}
 -2^{-(S+1)} &\leq \epsilon_{\log_2}^{S,\infty}(x) < \log_2 \left( \frac{2^S + (k+1)}{2^S + k} \right) - 2^{-(S+1)} \\
 -2^{-(S+1)} &\leq \epsilon_{\log_2}^{S,\infty}(x) < \log_2(1 + 2^{-S}) - 2^{-(S+1)} \quad \text{pour } k = 0 \\
 -2^{-(S+1)} &\leq \epsilon_{\log_{10}}^{S,\infty}(x) < 3.0103 [\log_2(1 + 2^{-S}) - 2^{-(S+1)}] \quad \text{pour } k = 0
 \end{aligned} \tag{B.5}$$

Ceci se fait au détriment de la simplicité de l'algorithme qui voit la simple concaténation de  $E_{\log_2}$  et  $D_{\log_2}(S, M_D)$  dans  $R_{\log_2}$  remplacée par leur addition (voir figure B.5). L'erreur s'écrit maintenant comme suit et est tabulée en figure B.4.

$$RMS(\epsilon_{\log_2}^{S,\infty}) = \sqrt{\sum_{k=0}^{2^S-1} \left[ x \left( \begin{array}{c} (\log_2(x) - \log_2(e) - Tab\_Conv\_D_{\log_2}^{S,M_D}[k])^2 \\ + \log_2^2(e) \end{array} \right) \right]_{1+k \cdot 2^{-S}}^{1+(k+1)2^{-S}}}$$

avec  $Tab\_Conv\_D_{\log_2}^{S,M_D}[k] = 2^{-M_D} [2^{M_D} [2^{-(S+1)} + \log_2(1 + k \cdot 2^{-S})]]$



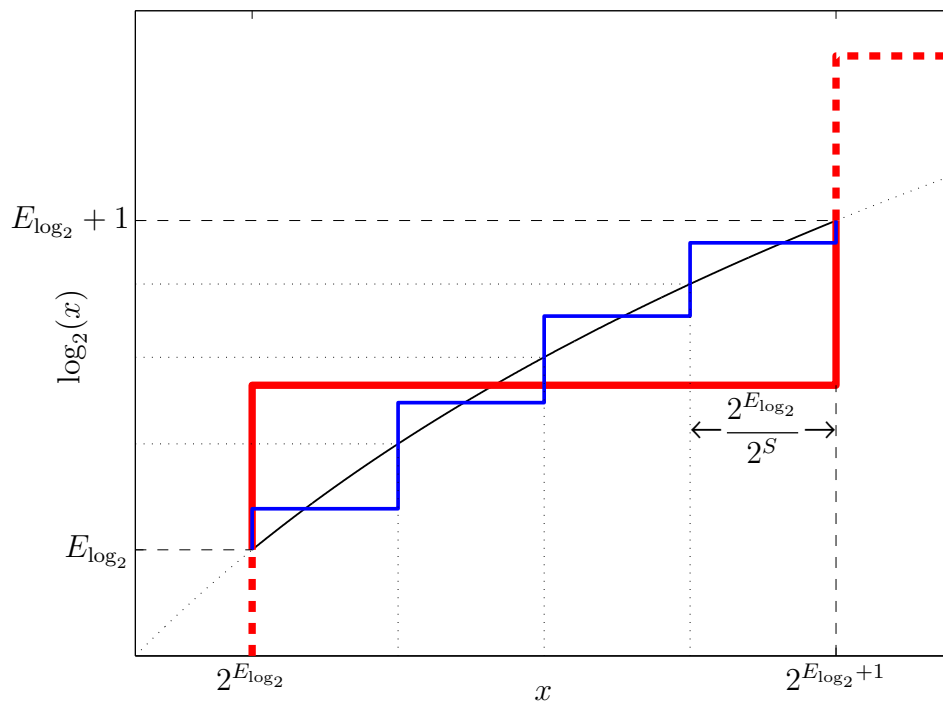


FIG. B.3 – Approximation avec arrondis convergeant de  $\log_2(x)$  (en noir fin) pour  $S = 0$  (en rouge très gras) et  $S = 2$  (en bleu gras)

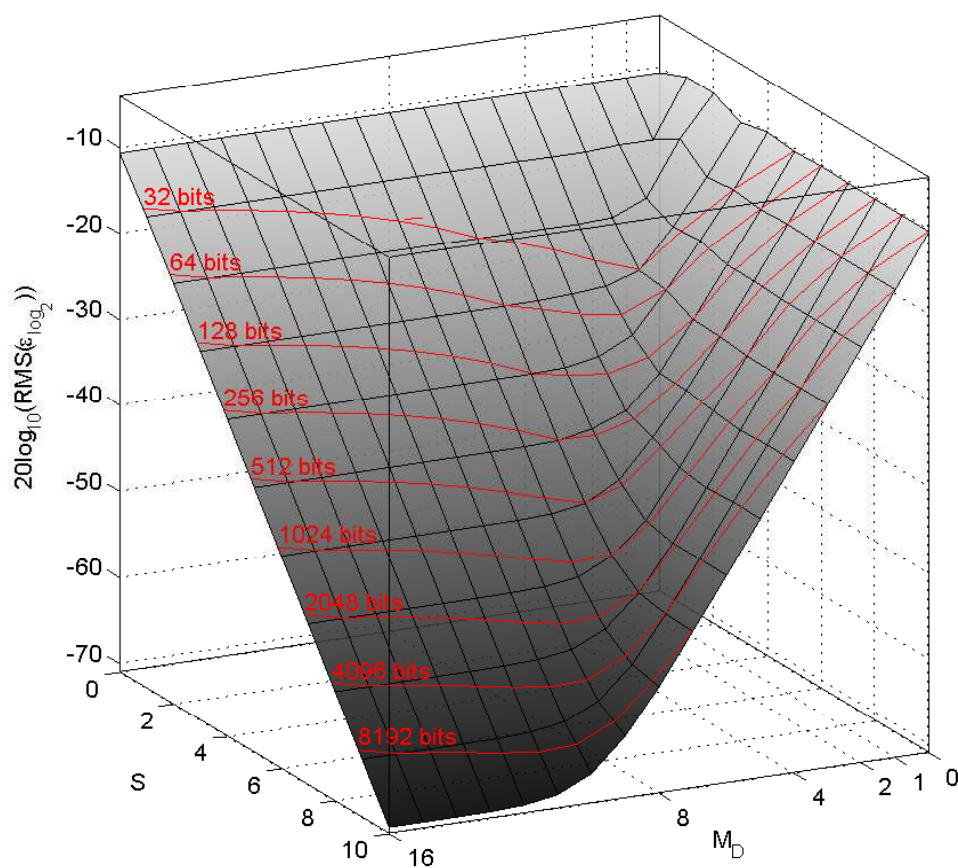


FIG. B.4 – Tabulation de l'erreur  $\text{RMS}(\epsilon_{\log_2}^{S, M_D})$  du calcul de  $\log_2$  avec arrondis convergent en fonction de  $S$  et  $M_D$ . Les lignes rouges transversales sont les isocontours de la taille de mémoire requise pour stocker la table.

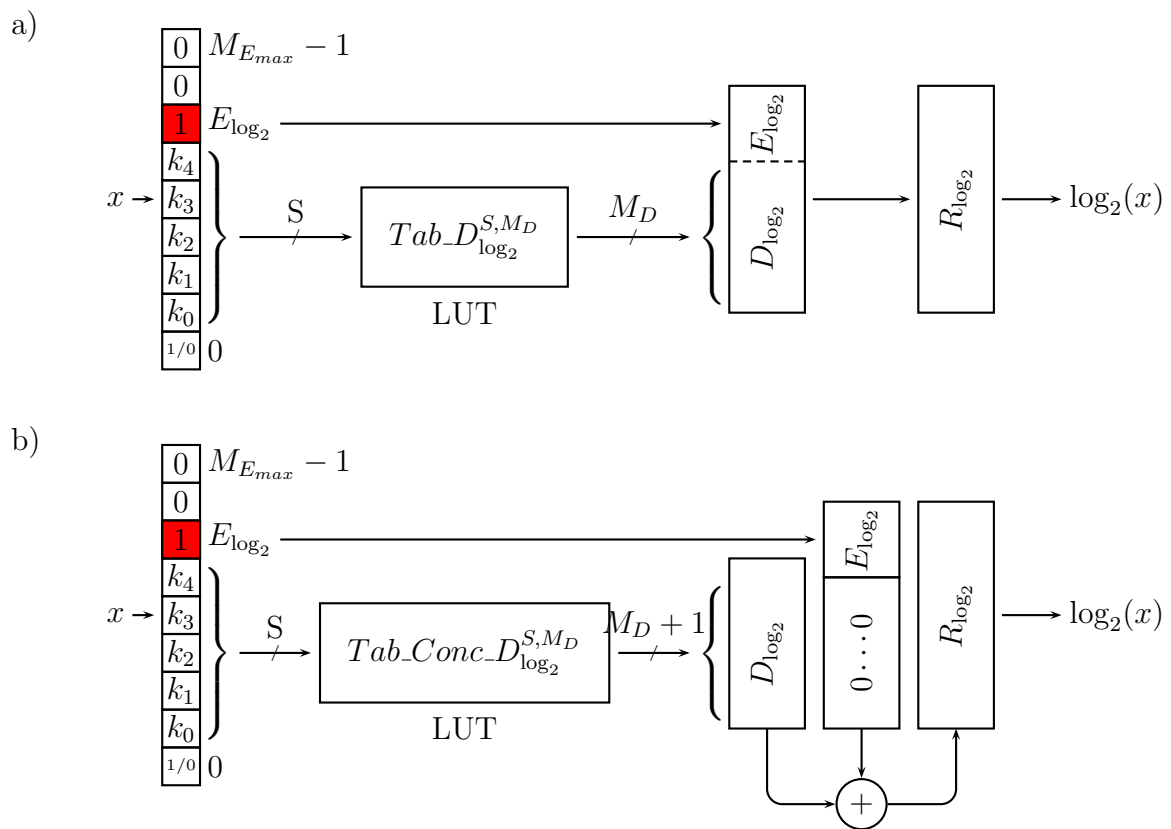


FIG. B.5 – Hardware d'un calculateur de  $\log_2$  : a) sans, b) avec codage convergent de partie décimale.

TAB. B.1 – Code C du calcul du  $\log_2$ . La table *Tab\_D* est issue des calculs présentés en page 175. Les lignes 19 à 23 incluse concernent uniquement l'application décrite dans le chapitre 4.

```

1 #define S 5
2 #define Md 5
3 #define SIZE_SPECTRES 256
4 #define RES_REDUCTION 8
5 unsigned int Tab_D[] = { 1, 2, 3, 5, 6, 7, 8,10,\
6                          11,12,13,14,15,16,17,18,\
7                          19,20,21,22,23,24,25,26,\
8                          26,27,28,29,30,30,31,32};
9
10 void approx_log2(unsigned int dest, unsigned int orig){
11     register unsigned int j, k, acc,tmp;
12     register unsigned int *in, *out;
13     register int N;
14
15     in = (unsigned int*)orig;
16     out = (unsigned int*)dest;
17
18     for(j=0;j<SIZE_SPECTRES;j++){
19         // Frequency resolution widening using summation
20         acc = 0;
21         for (k=0;k<RES_REDUCTION;k++){
22             acc += *(in++);
23         }
24
25         // Integer part calculation of log2
26         N = 0;
27         tmp = acc;
28         while (acc > 1){
29             acc >>= 1;
30             N ++;
31         }
32
33         // Decimal part calculation of log2 (Tab_D[acc])
34         acc = tmp - (1<<N);
35         acc = ((N-S)>0)? acc >> (N-S) : acc << (S-N);
36
37         // log2 calculation
38         *(out++) = (N<<Md) + Tab_D[acc];
39     }
40 }

```



## ANNEXE C

---

# TABLE D'ALLOCATION DES FRÉQUENCES US

---

Cette table représente graphiquement les bandes allouées aux États-Unis pour les divers utilisateurs du spectre radioélectrique.

On pourra noter quelques différences avec les allocations européennes disponibles dans <http://www.ero.dk/doc98/official/pdf/rep025.pdf>.

Les bandes allouées plus ou moins exclusivement à la radioastronomie sont les zones jaunes encadrées de rouge.



## ANNEXE D

---

# OUTILS UTILISÉS

---

Bien que les traitements proposés dans ce mémoire soient destinés à être implantés dans du matériel, quantité de suites de développement et de bureautique sont nécessaires à la réalisation de cette tâche. Nous établissons ici un descriptif des logiciels les plus utilisés.

Les systèmes d'exploitation des machines de développement et des machines cibles sont fortement contraints par les cartes électroniques embarquées ou les versions d'outils de traitement disponibles. Les OS utilisés couvrent la gamme des Windows NT, 2000, XP Pro et des Linux Ubuntu, Kubuntu, openSUSE et CentOS.

### Traitement de données

Les simulations, la génération des figures de ce mémoire, l'interprétation des fichiers binaires générés par le récepteur numérique a été effectuée majoritairement avec [MATLAB®](#) 6.5, R13 de MathWorks.

[GNU Octave](#) 2.1.72, grâce à son déploiement aisé sur de multiples PCs fonctionnant aussi bien sous Linux que sous Windows, a permis d'exécuter sur plusieurs machines en parallèle des simulations simples mais coûteuses en temps initialement écrites sous [MATLAB®](#).

[IDL](#) 6.3 de ITT Visual Information Solutions a été utilisé pour quelques traitements personnalisés et dépouillements d'observations mais surtout pour visualiser les données acquises grâce aux formidables outils de visualisation écrits par Andrée.

### Développement

Les algorithmes ont été implantés selon les contraintes techniques sur des plateformes programmables diverses décrites par des langages différents.

Le code VHDL pour les FPGAs a été écrit et synthétisé sous [ISE](#) 6.3 de Xilinx et simulé *via* des testbenchs exécutés sous [ModelSim](#).

Le code C des DSPs a été écrit, compilé et débogué sous [CCS](#) 2.2 de Texas Instruments.

Le code C des applications PC-Windows a été écrit, compilé et débogué sous [Borland C++ Builder](#).



## Bureautique

La rédaction de courts mémos ou de documentations utilisateurs, la gestion de feuilles de calculs et l'édition de présentations a été effectuée avec beaucoup de plaisir sous [OpenOffice.org](http://OpenOffice.org) 2.0 (OOo). On notera la fonction d'export au format PDF (copiée depuis par d'autres suites bureautiques) qui permet un rendu parfait des documents livrés, gage d'une sérénité idyllique lors des présentations publiques sur des machines de configuration inconnue. On notera également que le standard ODF, supporté en natif par OOo, est certifié ISO tout comme PDF et HTML. Lui seul permet l'interopérabilité entre toutes les bonnes suites bureautiques qui en offrent le support.

La rédaction de ce mémoire de thèse a été effectuée en  $\LaTeX$ . L'édition a été entamée sous Win2000 avec le couple [TExnicCenter/MiKTeX](http://TExnicCenter/MiKTeX), MiKTeX effectuant la gestion des paquets  $\LaTeX$  et TExnicCenter, l'interface utilisateur, facilitant la gestion du projet (profils de compilation, gestion de l'arborescence du document, aide à la syntaxe, correction orthographique, ...). Après migration du système d'exploitation vers Kubuntu/Linux 6.10 "Edgy Eft", la rédaction a continué sous [Kile](http://Kile) après la simple conversion du jeu de caractère du contenu des fichiers texte avec `recode ISO-8859-15..UTF-8 filename` et `convmv -r --notest -f cp850 -t UTF-8 repertoire` pour leur nom. Une partie des diagrammes a été codé directement avec des directives [PSTricks](http://PSTricks), la majorité ayant été édités avec [Inkscape](http://Inkscape) sous Windows ou Linux puis exportés en EPS. Les figures MATLAB® et les images ont aussi été converties en EPS avant leur inclusion dans le code  $\LaTeX$ .

---

## BIBLIOGRAPHIE

---

- [1] Milton ABRAMOWITZ et Irene A. STEGUN. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover Publications, New York, 1964.
- [2] David W. ALLAN, Neil ASHBY et Clifford C. HODGE. *The Science of Timekeeping, Application Note 1289* *ii*. Hewlett Packard, 1997.
- [3] C. BARNBAUM et R. F. BRADLEY. *A New Approach to Interference Excision in Radio Astronomy: Real-Time Adaptive Cancellation* *ii*. *Astronomical Journal*, 115:2598–2614, 1998.
- [4] C. H. BARROW, Ph. ZARKA et M. G. AUBIER. *Fine structures in solar radio emission at decametre wavelengths* *ii*. *Astronomy and Astrophysics*, 286:597–606, juin 1994.
- [5] François BIRAUD et Nguyen quang RIEU. *Techniques de l'Ingénieur, traité Électronique, Radioastronomie* *ii*. Rapport Technique E 6850, Techniques de l'ingénieur, 1994.
- [6] Albert-Jan BOONSTRA. *Radio Frequency Interference Mitigation in Radio Astronomy* *ii*. Thèse de doctorat, ASTRON, 2005.
- [7] G. C. BOWER. *Application of Wiener and Adaptive Filters to GPS and Glonass Data from the Rapid Prototyping Array* *ii*. Allen Telescope Array (ATA), août 2001.
- [8] Stéphanie BRETTEIL. *Estimation spectrale en temps réel de sources radioastronomiques en présence de brouilleurs cyclostationnaires : Application au radiotélescope décimétrique de Nançay* *ii*. Thèse de doctorat, Université d'Orléans, décembre 2005.
- [9] John BUNTON. *Memo 342 - An Improved FX Correlator* *ii*. Rapport Technique, ALMA, 2000.
- [10] Y. CHIKADA, M. ISHIGURO, H. HIRABAYASHI, M. MORIMOTO, K. I. MORITA, K. MIYAZAWA, K. NAGANE, K. MURATA, A. TOJO, S. INOUE, T. KANZAWA et H. IWASHITA. *A Digital FFT Spectro-Correlator for Radio Astronomy* *ii*. Dans *Indirect Imaging. Measurement and Processing for Indirect Imaging. Proceedings of an International Symposium held in Sydney, Australia, August 30-September 2, 1983*. Editor, J.A. Roberts; Publisher, Cambridge University Press, Cambridge, England, New York, NY, 1984. LC # QB51.3.E43 I53 1984. ISBN # 0-521-26282-8. P.387, 1983, pages 387–+, 1983.
- [11] Vincent CLERC. *Implémentation de processus hétérogènes de traitement du signal en temps réel à contrainte moyenne. Application à la radioastronomie.* *ii*. Thèse de doctorat, Université de Paris 6, 2003.

- [12] Nicola DA DALT, Moritz HARTENECK, Christoph SANDNER et Andreas WIESBAUER. ¶ On the jitter requirements of the sampling clock for analog-to-digital converters ¶. *IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS*, 49:1354–1360, septembre 2002.
- [13] H. A. DAVID et H. N. NAGARAJA. *Order Statistics*. John Wiley & Sons Canada, Ltd., 3 édition, jul 2003.
- [14] W. DONG, B. D. JEFFS et J. Richard FISHER. Kalman Tracking and Bayesian Detection for Radar RFI Blanking. Dans S. W. ELLINGSON, éditeur, *RFI2004, Workshop in Mitigation of Radio Frequency Interference in Radio Astronomy*. Ellingson, S. W., Penticton, British Columbia, Canada, juillet 2004.
- [15] Weizhen DONG. ¶ Time Blanking for GBT Data with Radar RFI ¶. Thèse de master recherche, Brigham Young University, 2004.
- [16] G. A. DULK, A. LECACHEUX et Y. LEBLANC. ¶ The complete polarization state of a storm of millisecond bursts from Jupiter ¶. *Astronomy and Astrophysics*, 253:292–306, janvier 1992.
- [17] Franck ELIE. ¶ *Conception et réalisation d’un système utilisant des réseaux de neurones pour l’identification et la caractérisation, à bord de satellites, de signaux transitoires de type sifflement* ¶. Thèse de doctorat, Université d’Orléans, décembre 1997.
- [18] S. W. ELLINGSON. ¶ A Study of Adaptive Canceling for Microwave Radiometry and Spectrometry ¶. Rapport Technique 743467-1, The Ohio State University ElectroScience Laboratory, septembre 2002.
- [19] S. W. ELLINGSON et G. A. HAMPSON. ¶ Mitigation of radar interference in L-band radio astronomy ¶. *Astrophysical Journal Supplement Series*, 147:167–176, 2003.
- [20] Jani EVEN. ¶ *Contribution à la séparation de sources à l’aide de statistiques d’ordre* ¶. Thèse de doctorat, Université Joseph Fournier, Grenoble, décembre 2003. p. 21-47.
- [21] G. FERRÉ. ¶ *Segmentation d’images de documents anciens par approche texture* ¶. Thèse de master, Université de la Rochelle, 2006.
- [22] J. Richard FISHER. ¶ GBT LO and Doppler Correction ¶. Rapport Technique, NRAO, 2000. <http://www.cv.nrao.edu/~rfisher/Ephemerides/doppler.ps>.
- [23] P. A. FRIDMAN. ¶ RFI excision using a higher order statistics analysis of the power spectrum ¶. *aap*, 368:369–376, septembre 2001.
- [24] P. A. FRIDMAN et W. A. BAAN. ¶ RFI mitigation methods in radio astronomy ¶. *Astronomy and Astrophysics*, 378:327–344, 2001.
- [25] Djamel GACEB. ¶ *Extraction de similarités de formes dans les images de traits* ¶. Thèse de master, Université de Lyon, 2005.
- [26] F. GENOVA. ¶ Les émissions radio “aurorales” des planètes ¶. *Annales de Physique*, 12:57–107, avril 1987.
- [27] F. GENOVA et M. G. AUBIER. ¶ Io-dependent sources of the Jovian decameter emission ¶. *Astronomy and Astrophysics*, 150:139–150, septembre 1985.

- [28] F. GENOVA, Ph. ZARKA et Alain LECACHEUX. ¶ Jupiter Decametric Radiation ¶. Dans M. J. S. BELTON, R. A. WEST, J. RAHE et M. PEREYDA, éditeurs, *NASA SPECIAL PUBLICATION SERIES, NASA-SP-494, 156-174. Document Section.*, pages 156–174, 1989.
- [29] M. A. GORDON et J. W. WARWICK. ¶ High Time-Resolution Studies of Jupiter’s Radio Bursts ¶. *The Astrophysical Journal*, 148:511–+, mai 1967.
- [30] Malcolm J. GORIS. ¶ Categories of Radio-Frequency Interference ¶. Rapport Technique, NFRA, 1998.
- [31] J.-M. GRIESSMEIER, U. MOTSCHMANN, G. MANN et H. O. RUCKER. ¶ The influence of stellar wind conditions on the detectability of planetary radio emissions ¶. *Astronomy and Astrophysics*, 437:717–726, juillet 2005.
- [32] J. P. HAMAKER. ¶ Understanding radio polarimetry. IV. The full-coherency analogue of scalar self-calibration: Self-alignment, dynamic range and polarimetric fidelity ¶. *Astronomy and Astrophysics Suppl. Ser.*, 143:515–534, mai 2000.
- [33] J. P. HAMAKER et J. D. BREGMAN. ¶ Understanding radio polarimetry. III. Interpreting the IAU/IEEE definitions of the Stokes parameters. ¶. *Astronomy and Astrophysics Suppl. Ser.*, 117:161–165, mai 1996.
- [34] J. P. HAMAKER, J. D. BREGMAN et R. J. SAULT. ¶ Understanding radio polarimetry. I. Mathematical foundations. ¶. *Astronomy and Astrophysics Suppl. Ser.*, 117:137–147, mai 1996.
- [35] Grant A. HAMPSON. ¶ *Implementation of the Asynchronous Pulse Blanker* ¶, juin 2002. <http://esl.eng.ohio-state.edu/~rsttheory/iip/docserv.html>.
- [36] Fredric J. HARRIS. ¶ On the Use of Windows for Harmonic Analysis with Discrete Fourier Transform ¶. *Proceedings of the IEEE*, 66(1):51–83, janvier 1978.
- [37] C. G. T. HASLAM, C. J. SALTER, H. STOFFEL et W. E. WILSON. ¶ A 408 MHz all-sky continuum survey ¶. *Astronomy and Astrophysics Suppl.*, 47(1):1–142, janvier 1982.
- [38] S. HESS, Ph. ZARKA et F. MOTTEZ. ¶ Io-Jupiter Interaction, Millisecond Bursts and Field Aligned Potentials. ¶. *AGU Fall Meeting Abstracts*, pages A1275+, décembre 2005.
- [39] C. IMPEY et G. BOTHUN. ¶ Malin 1 - A quiescent disk galaxy ¶. *Astrophysical Journal*, 341:89–104, juin 1989.
- [40] John David JACKSON. *Classical Electrodynamics*. John Wiley & Sons, Inc., third edition édition, 1998.
- [41] P. E. JOHNSON et D. G. LONG. ¶ The probability density of spectral estimates based on modified periodogram averages ¶. *IEEE Transactions on Signal Processing*, 47(5):1255–1261, may 1999.
- [42] K. I. KELLERMANN. ¶ Grote Reber (1911-2002) ¶. *The Publications of the Astronomical Society of the Pacific*, 116:703–711, août 2004.
- [43] M. J. KESTEVEN. ¶ Adaptive Filters for RFI Mitigation in Radioastronomy ¶. Dans *IVS Symposium, New Technologies in VLBI*, Gyeong-Ju, Korea, novembre 2002. IVS Symposium, New Technologies in VLBI.

- [44] J. W. KOOI, G. CHATTOPADHYAY, M. THIELMAN, T. G. PHILLIPS et R. SCHIEDER. ¶ Noise stability of SIS receivers ¶. *International Journal of Infrared and Millimeter Waves*, 21(5):689–716, mai 2000.
- [45] John D. KRAUS. *Radio Astronomy*. Cygnus-Quasar Books, 1986.
- [46] David LAST. ¶ *Caractérisation de l'architecture tridimensionnelle du tissu trabéculaire in vitro et in vivo par IRM* ¶. Thèse de doctorat, Université Paris XI, 2003.
- [47] Y. LEBLANC, M. G. AUBIER, C. ROSOLEN, F. GENOVA et J. DE LA NOE. ¶ The Jovian S-bursts. II - Frequency drift measurements at different frequencies throughout several storms ¶. *Astronomy and Astrophysics*, 86:349–354, juin 1980.
- [48] Y. LEBLANC, F. GENOVA et J. DE LA NOE. ¶ The Jovian S-bursts. I - Occurrence with L-bursts and frequency limit ¶. *Astronomy and Astrophysics*, 86:342–348, juin 1980.
- [49] Richard LYONS. ¶ *Quadrature Signals: Complex, But Not Complicated* ¶, novembre 2000. <http://www.dspguru.com/info/tutor/QuadSignals.pdf>.
- [50] Ricardo A. MARONNA, Douglas R. MARTIN et Victor J. YOHAI. *Robust Statistics: Theory and Methods (Wiley Series in Probability and Statistics)*. Wiley, juin 2006.
- [51] Jean-Michel MARTIN. ¶ *Physique du gaz interstellaire dans les galaxies fortement émettrices dans l'infrarouge lointain et étude particulière des mégamasers OH* ¶. Thèse de doctorat, Université de Paris VII, février 1989.
- [52] D. A. MITCHELL et G. C. BOWER. ¶ *Cancelling Satellite Interference at the Rapid Prototyping Array. A Comparison of Voltage and Power Domain Techniques* ¶. Allen Telescope Array (ATA), décembre 2001.
- [53] Noppasin NIAMSUWAN, Joel T. JOHNSON et Steven W. ELLINGSON. ¶ Examination of a simple pulse-blanking technique for radio frequency interference mitigation ¶. *RADIO SCIENCE*, 40(3), juin 2005.
- [54] G. PATUREL, G. THEUREAU, L. BOTTINELLI, L. GOUGUENHEIM, N. COUDREAU-DURAND, N. HALLET et C. PETIT. ¶ HYPERLEDA. II. The homogenized HI data ¶. *Astronomy and Astrophysics*, 412:57–67, décembre 2003.
- [55] Didier PELAT. ¶ *Bruits et Signaux* ¶, École doctorale d'Île-de-france édition, 2002.
- [56] A. J. POULSEN, B. D. JEFFS, K. F. WARNICK et J. R. FISHER. ¶ Programmable Real-Time Cancellation of GLONASS Interference With the Green Bank Telescope ¶. *The Astronomical Journal*, 130:2916–2927, 2005.
- [57] Andrew J. POULSEN. ¶ *Real-time Adaptive Cancellation of Satellite Interference in Radio Astronomy* ¶. Thèse de doctorat, Brigham Young University, août 2003.
- [58] William H. PRESS, Brian P. FLANNERY, Saul A. TEUKOLSKY et William T. VETTERLING. *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, 2 édition édition, septembre 1992.
- [59] G. REBER. ¶ Cosmic Static ¶. Dans *Astrophys. J.*, volume 30, pages 367–378. Proc. IRE (Institute of Radio Engineers), août 1942.
- [60] V. RESHETNIKOV, F. BOURNAUD, F. COMBES, M. FAÚNDEZ-ABANS, M. DE OLIVEIRA-ABANS, W. VAN DRIEL et S. E. SCHNEIDER. ¶ Galaxy transmutations: The double ringed galaxy ESO 474-G26 ¶. *Astronomy and Astrophysics*, 431:503–510, février 2005.

- [61] J. J. RIIHIMAA. ;; Spectra of Jovian Millisecond Pulsars *ii*. *Nature*, 218:1141–+, 1968.
- [62] K. ROHLFS et T. L. WILSON. *Tools of Radio Astronomy (Astronomy & Astrophysics Library)*. Springer-Verlag Berlin and Heidelberg GmbH & Co. K, 2003.
- [63] Carlo ROSOLEN, Laurent DENIS, Vincent CLERC et Rodolphe WEBER. ;; New generation of robust receiver at Nançay radio astronomy observatory *ii*. Rapport Technique, RFI mitigation workshop, Bonn, mars 2001.
- [64] R. J. SAULT, J. P. HAMAKER et J. D. BREGMAN. ;; Understanding radio polarimetry. II. Instrumental calibration of an interferometer array. *ii*. *Astronomy and Astrophysics Suppl. Ser.*, 117:149–159, mai 1996.
- [65] C. M. SPRINGOB, M. P. HAYNES, R. GIOVANELLI et B. R. KENT. ;; A Digital Archive of H I 21 Centimeter Line Spectra of Optically Targeted Galaxies *ii*. *The Astrophysical Journal Supplement Series*, 160:149–162, septembre 2005.
- [66] Engineering STAFF. ;; Application Note 728: Defining and Testing Dynamic Parameters in High-Speed ADCs, Part 1 *ii*. Rapport Technique, Maxim Integrated Products, 2001. [http://www.maxim-ic.com/appnotes.cfm/an\\_pk/728](http://www.maxim-ic.com/appnotes.cfm/an_pk/728).
- [67] Engineering STAFF. ;; Application Note 729: Dynamic Testing of High-Speed ADCs, Part 2 *ii*. Rapport Technique, Maxim Integrated Products, 2001. [http://www.maxim-ic.com/appnotes.cfm/an\\_pk/729](http://www.maxim-ic.com/appnotes.cfm/an_pk/729).
- [68] A. Richard THOMPSON, James M. MORAN et George W. SWENSON, JR.. *Interferometry and Synthesis in Radio Astronomy*. Krieger Pub Co, 1986.
- [69] TI ENGINEERING STAFF. ;; TMS320C62x DSP Library Programmer’s Reference *ii*. Rapport Technique, Texas Instruments, 2003. <http://focus.ti.com/lit/ug/spru402b/spru402b.pdf>.
- [70] M. VAN GINKEL, C. L. LUENGO HENDRIKS et L. J. VAN VLIET. ;; A short introduction to the Radon and Hough transforms and how they relate to each other *ii*. Rapport Technique, Delft University of Technology, Quantitative Imaging Group, 2004. [citeseer.ist.psu.edu/vanginkel04short.html](http://citeseer.ist.psu.edu/vanginkel04short.html).
- [71] H. VAN-TREES. *Detection, Estimation, and Modulation Theory. Part I*. John Wiley & Sons, 1968.
- [72] Rodolphe WEBER, Cédric VIOU, Andrée COFFRE, Laurent DENIS, Philippe ZARKA et Alain LECACHEUX. ;; DSP-Enabled Radio Astronomy: Towards IIIZW35 Reconquest *ii*. *EURASIP Journal on Applied Signal Processing*, 2005(16):2686–2693, 2005.
- [73] Peter D. WELCH. ;; The Use of Fast Fourier Transform for the Estimation of Power Spectra: A Method Based on Time Averaging Over Short, Modified Periodograms *ii*. *IEEE Trans. Audio Electroacoustics*, AU-15:70–73, juin 1967.
- [74] C. YOSHIHIRO. ;; Correlators for interferometry - Today and tomorrow *ii*. Dans T. J. CORNWELL et R. A. PERLEY, éditeurs, *ASP Conf. Ser. 19: IAU Colloq. 131: Radio Interferometry. Theory, Techniques, and Applications*, pages 26–32, 1991.
- [75] Qing ZHANG, Yibin ZHENG, Stephen G. WILSON, J. Richard FISHER et Richard BRADLEY. ;; Combating Pulsed Radar Interference in Radio Astronomy *ii*. *The Astronomical Journal*, 126:1588–1594, septembre 2003.

---

Restauration de sources radioastronomiques en milieu radioélectrique hostile :  
Implantation de détecteurs temps réel sur des spectres dynamiques.

---

L'étude de l'Univers est effectuée de façon quasi-exclusive grâce aux ondes électromagnétiques. La radioastronomie qui étudie la bande radio du spectre est couramment obligée d'observer hors de ses bandes protégées. Il faut donc pouvoir observer des sources dans des bandes où les émissions des systèmes de télécommunication sont souvent proches et donc très puissantes.

Les travaux de cette thèse ont eu pour but de développer et d'implanter des méthodes temps réel de traitements numériques visant à atténuer les interférences ambiantes afin de restaurer les radiosources naturelles lors d'observation avec une antenne unique sans informations *a priori*. Le cœur de ces traitements est basé sur l'estimation de la moyenne d'un échantillon suivant une loi du  $\chi^2$  en présence de points aberrants. Un récepteur numérique multi-instruments à grande dynamique élaboré à la station de radioastronomie de Nançay a servi de banc de test à ces divers algorithmes.

La bande corrompue par Iridium est maintenant observable et les sources HI situées dans la bande allouée aux radars sont de nouveau accessibles. Les observations du Soleil et de Jupiter effectuées en bande décimétrique sont maintenant de bien meilleure qualité.

Un algorithme a été développé pour configurer le récepteur en tant que détecteur et enregistreur rapide d'évènements très brefs ( $\approx 100$  ms) en milieu parasité. Il est aujourd'hui utilisé pour l'enregistrement de structures fines Joviennes.

**Mots-Clés : Estimation robuste, Détection, Implantation matérielle, Instrumentation, Radioastronomie**

---

Radioastronomical sources restoration from hostile radioelectric environment:  
Implementation of real-time detectors for dynamic spectra analysis.

---

The Universe is mostly studied thanks to electromagnetic waves. Radioastronomy studies the radio band of the spectrum and therefore must look beyond its protected bands. Thus, sources have to be observed in bands where radiocommunication emissions are often both near and powerful.

This PhD thesis aimed at both developing and implementing real time digital processing methods. Such methods intend to mitigate surrounding interferences in order to restore natural radiosources. Instrumental limitations include single antenna observations without any *a priori* knowledge about interferers. The keystone of all processings is an accurate estimation of the mean for a  $\chi^2$ -distribution based set of data partly corrupted by outlayer values. The various algorithms we used, ran on a test bench elaborated at Nançay. It consists of a high-dynamic multi-instruments digital receiver that is highly reconfigurable.

Both OH sources in Iridium frequency band and HI sources located in ground based radar frequency bands are now available to radioastronomy observations. Sun and Jupiter daily recordings are now of much better quality.

A particular algorithm has been developed in order to detect brief events ( $\approx 100$  ms) in a corrupted environment and then trigger high resolution recordings. This algorithm is now used on a regular basis at the Nançay Radio Station to gather Jovian S-bursts.

**Keywords: Robust estimation, Detection, Hardware Implementation, Instrumentation, Radio-Astronomie**