



**Développement d'une technique numérique  
d'autocorrélation d'intensité pour la recherche  
de minilentilles gravitationnelles dans le radio et  
de faibles variations temporelles rapides**

**Thèse**

**Éric Trottier**

**Doctorat en physique**  
Philosophiae doctor (Ph.D.)

Québec, Canada

© Éric Trottier, 2018

**Développement d'une technique numérique  
d'autocorrélation d'intensité pour la recherche  
de minilentilles gravitationnelles dans le radio et  
de faibles variations temporelles rapides**

**Thèse**

**Éric Trottier**

Sous la direction de:

Ermanno F. Borra, directeur de recherche

## Résumé

Pour la première fois, on propose d'appliquer une technique numérique d'autocorrélation d'intensité pour rechercher des minilentilles gravitationnelles dans des observations radio de noyaux actifs de galaxies (NAGs), et aussi pour trouver de faibles variations temporelles rapides. Essentiellement, la technique consiste à détecter un signal (délai temporel) cosmique (ou artificiel) à partir des fluctuations temporelles d'intensité. La programmation des algorithmes est codée en Matlab sur un microordinateur et en C++ pour le traitement en parallèle multi-cœur sur un superordinateur. On valide la technique et les logiciels développés en montrant la détection de faibles signaux artificiels périodiques et non périodiques. On pourrait aussi trouver des variations temporelles très rapides puisque les données radioastronomiques interférométriques sont échantillonnées sur de très courts laps de temps (ex., 15.625 ns). Enfin, des simulations numériques permettent d'évaluer l'efficacité de détection de notre technique innovatrice.

La transformée de Fourier (TF) peut aussi détecter des signaux périodiques. Cependant, l'autocorrélation d'intensité a l'avantage majeur de détecter des signaux non périodiques, alors que la TF ne le peut pas. De plus, notre technique permet de visualiser directement la forme des signaux, tandis qu'il est très difficile de voir la forme avec la transformée de Fourier. Une nouveauté de la démonstration de notre technique est qu'on utilise des données prises par des radiointéromètres (ex., NRAO) déjà existants dans un format standard (ex., Mark 5B), de sorte que l'on n'a pas besoin d'instrumentation spécialisée.

On a analysé 22 objets. Aucun signal cosmique n'a été détecté pendant les dix minutes d'observation VLBA du NRAO de chaque objet. À partir du résultat nul obtenu et d'une analyse statistique simplifiée par la binomiale, on a fait un estimé approximatif du paramètre actuel de densité cosmique de minilentilles compactes non-baryoniques, noté  $\Omega_{l,0}$ , uniformément distribuées dans l'intervalle de masse,  $10 - 10^7$  masses solaires, lequel correspond à des délais de  $\sim 1$  ms à  $\sim 600$  s. Pour le modèle

plat (actuellement favorisé) avec constante cosmologique  $\lambda_0 = 0.7$  ( $\Omega_0 = 0.3$ ), la limite estimée est  $\Omega_{l,0} \leq 0.2 - 0.3$ . Vu le faible échantillon-test,  $n = 22$ , les incertitudes (intervalles de confiance) obtenues sont trop élevées pour fixer une contrainte statistiquement significative: statistique de petit nombre.

## Abstract

For the first time, we propose to apply a digital autocorrelation of intensity technique to gravitational minilensing in active galactic nuclei (AGNs) radio observations, and also to search weak rapid time variations. Essentially, the technique consists of detecting a cosmic (or artificial) signal (time delay) using intensity fluctuations as a function of time. The programming algorithm is coded in Matlab on a microcomputer and in C++ for multi-core parallel processing on a supercomputer. We validate the technique and softwares by showing weak periodic and non-periodic artificial signals detections. We could also find very fast time variations because radioastronomical interferometric data are sampled over very short times (e.g., 15.625 ns). Finally, numerical simulations make it possible to evaluate the detection efficiency of our innovative technique.

The Fourier transform (FT) can also detect periodic signals. However, the main advantage of the autocorrelation of intensity is that it can detect non-periodic signals, while the FT cannot detect these non-periodic signals. Moreover, our technique allows a direct visualization of the shape of the signals, while it is very difficult to see the shape with a Fourier Transform. A novelty of the demonstration of our technique is that we use data taken with existing radio-interferometers (e.g., NRAO) in a standard format (e.g., Mark 5B) and therefore no specialized instrumentation is needed.

We analyzed 22 objects. No signal has been detected for the ten minutes of VLBA of the NRAO observation for each object. From the null result obtained and a simplified statistical analysis by the binomial, we made a rough estimate of the present cosmic density parameter of non-baryonic compact minilenses, named  $\Omega_{l,0}$ , uniformly distributed in the mass range,  $10 - 10^7$  solar masses, corresponding to delays of  $\sim 1$  ms to  $\sim 600$  s. For a flat model (currently favored) with cosmological constant  $\lambda_0 = 0.7$  and with  $\Omega_0 = 0.3$ , the estimated limit is  $\Omega_{l,0} \leq 0.2 - 0.3$ . Given the low test sample, n

= 22, the uncertainties (confidence intervals) obtained are too high to set a statistically significant constraint: small number statistics.

# Table des matières

<b>RÉSUMÉ</b> .....	<b>III</b>
<b>ABSTRACT</b> .....	<b>V</b>
<b>TABLE DES MATIÈRES</b> .....	<b>VII</b>
<b>LISTE DES TABLEAUX</b> .....	<b>XI</b>
<b>LISTE DES FIGURES</b> .....	<b>XIII</b>
<b>REMERCIEMENTS</b> .....	<b>XV</b>
<b>1 INTRODUCTION</b> .....	<b>1</b>
1.1 MISE EN CONTEXTE, PROBLÉMATIQUE ET OBJECTIFS .....	1
1.2 DEMANDE D'ALLOCATION DE RESSOURCES ADDITIONNELLES AU CAR .....	7
1.3 LA LONGUE HISTOIRE DE LA DÉFLEXION LUMINEUSE GRAVITATIONNELLE .....	7
1.4 ORGANISATION DE LA THÈSE .....	9
1.5 CONTRIBUTIONS SCIENTIFIQUES .....	12
<b>2 RAPPELS DES PRINCIPAUX OUTILS DE RADIOASTRONOMIE</b> .....	<b>13</b>
2.1 ÉLÉMENTS DE RADIOASTRONOMIE QU'IL FAUT MAÎTRISER POUR LE PROJET .....	13
2.1.1 Sources radio .....	13
2.1.1.1 Généralités sur l'émission thermique et non thermique des radiosources .....	13
2.1.1.2 Sources radio cosmologiques: les quasars .....	15
2.1.1.3 Émission radio des quasars: le rayonnement synchrotron .....	18
2.1.1.4 Spectres des quasars .....	20
2.1.1.5 Choix des candidats pour les observations .....	22
2.1.2 Radiotélescopes .....	23
2.1.3 Bruit de photons et bruit d'ondes .....	25
2.1.4 Principales sources dominantes de bruit et température de bruit .....	27
2.1.5 Sensibilité du système VLBI du NRAO .....	29
2.2 MILIEUX DE PROPAGATION DES ONDES RADIO .....	31
2.2.1 Dispersion .....	31
2.2.2 Contributions à la mesure de dispersion totale .....	32
2.2.3 Délai temporel dû à la DM .....	34
<b>3 NOTIONS SUR LES LENTILLES GRAVITATIONNELLES</b> .....	<b>35</b>
3.1 THÉORIE DE BASE.....	35
3.1.1 Régimes des lentilles gravitationnelles .....	35
3.1.2 Principes de base des lentilles ponctuelles .....	36
3.1.3 Estimation de l'ordre de grandeur des délais temporels .....	41
3.2 LENTILLES GRAVITATIONNELLES FORTES ET FAIBLES.....	43
3.2.1 Effets de lentilles fortes .....	43
3.2.2 Effets de lentilles faibles .....	44
3.3 EFFETS DE MICROLENTILLES COSMOLOGIQUES SUR LES QUASARS .....	45
3.4 COURBES DE LUMIÈRE .....	47
3.4.1 Courbes de lumière standards.....	47
3.4.2 Courbes de lumière non standards .....	49
<b>4 AUTOCORRÉLATION D'INTENSITÉ ET SIGNAUX NUMÉRISÉS</b> .....	<b>51</b>
4.1 MODÈLE DU CHAMP ÉLECTRIQUE D'UN RAYONNEMENT SYNCHROTRON .....	51

4.2 INTENSITÉ ET FLUCTUATIONS D'INTENSITÉ.....	54
4.3 MODÈLE AVEC L'INTERFÉROMÈTRE DE YOUNG ET MODULATIONS SPECTRALES.....	55
4.4 ANALYSE BASÉE SUR L'AUTOCORRÉLATION DES FLUCTUATIONS D'INTENSITÉ .....	58
4.4.1 Signal dans l'autocorrélation d'intensité dû à l'effet d'une lentille .....	59
4.4.2 Avantages de la technique d'autocorrélation d'intensité.....	61
4.4.3 Inconvénients de la technique d'autocorrélation d'intensité.....	63
4.5 SIGNAUX NUMÉRISÉS.....	64
4.5.1 Notions de base sur les signaux.....	64
4.5.2 Échantillonnage.....	66
4.5.3 Quantification .....	67
4.5.4 Bruit aléatoire gaussien.....	68
4.5.5 Bruit et efficacité de quantification des données du NRAO .....	69
<b>5 MÉTHODOLOGIE ET TRAITEMENT DES DONNÉES MATLAB/C++ .....</b>	<b>71</b>
5.1 IMPLANTATION DES ALGORITHMES EN MATLAB .....	71
5.1.1 Lecture des données des data frames de format Mark 5B.....	72
5.1.2 Calcul et lissage des intensités .....	72
5.1.3 Autocorrélation temporelle d'intensité .....	74
5.1.4 Validation du code Matlab par la détection d'un signal périodique .....	75
5.1.5 Distribution du bruit dans l'autocorrélation d'intensité.....	77
5.2 TRADUCTION DU CODE MATLAB EN CODE C++.....	79
5.2.1 Lecture des données et autocorrélation d'intensité en C++ .....	79
5.2.2 Validation du code C++.....	80
5.2.3 Soumission des calculs sur un superordinateur.....	80
5.2.4 Récupération avec Matlab des données traitées en C++ .....	81
5.2.5 Détection d'un signal périodique rapide de faible amplitude .....	83
5.2.6 Détection de signaux non périodiques faibles .....	85
5.2.7 Détection de signaux d'interférences radio (RFIs) .....	87
5.3 TECHNIQUES DE DÉTECTION DE CANDIDATS (SIGNAUX).....	88
5.3.1 Choix des lissages et récupération des données .....	89
5.3.2 Évaluation du bruit dans l'autocorrélation d'intensité .....	90
5.3.3 Étapes de la détection d'un signal .....	93
<b>6 SIMULATIONS NUMÉRIQUES, ANALYSE STATISTIQUE DES RÉSULTATS ET ESTIMATION DE LA DENSITÉ COSMOLOGIQUE DE MINILENTILLES .....</b>	<b>97</b>
6.1 MISE EN CONTEXTE ET PROBLÉMATIQUE.....	97
6.2 SIMULATIONS NUMÉRIQUES DE DÉLAIS TEMPORELS .....	98
6.3 INTERVALLE DE MASSE APPLICABLE .....	108
6.4 MINILENTILLES DISTRIBUÉES LE LONG DE LA LIGNE DE VISÉE .....	110
6.4.1 Équations de la profondeur optique et des distances de diamètre .....	110
6.4.2 Rappel de la loi binomiale .....	112
6.4.3 Hypothèses simplificatrices et calculs des profondeurs optiques pour le modèle plat avec $\Omega_0 = 0.3$ et $\lambda_0 = 0.7$ .....	113
6.4.4 Intervalles de confiance à 95 % et estimé de la densité cosmologique d'objets compacts pour le modèle plat avec $\Omega_0 = 0.3$ et $\lambda_0 = 0.7$ .....	114
<b>7 CONCLUSION .....</b>	<b>117</b>
7.1 FORMAT DES DONNÉES, ALGORITHMES DE BASE ET LOGICIELS MATLAB/C++ .....	117
7.2 DÉTECTIONS DE FAIBLES VARIATIONS PÉRIODIQUES ET NON PÉRIODIQUES.....	118

7.3 RAPPEL DES TROIS ÉTAPES DE DÉTECTIONS DE SIGNAUX.....	122
7.4 SIMULATIONS NUMÉRIQUES DE DÉLAIS TEMPORELS .....	122
7.5 ESTIMÉ DE LA DENSITÉ COSMOLOGIQUE DE MINILENTILLES .....	123
7.6 LIMITATIONS DU PROJET DE RECHERCHE.....	124
7.7 OUVERTURE SUR LA DÉTECTION DE FRBS ET SUR L'INTELLIGENCE EXTRATERRESTRE.....	125
<b>BIBLIOGRAPHIE.....</b>	<b>127</b>
<b>A FORMAT ET TABLEAUX DES DONNÉES.....</b>	<b>135</b>
<b>B MÉTHODES D'INTERVALLES DE CONFIANCE D'UNE BINOMIALE.....</b>	<b>141</b>
B.1 INTERVALLE DE CLOPPER-PEARSON .....	141
B.2 INTERVALLE DE WILSON .....	142
B.3 INTERVALLE DE JEFFREYS .....	143
<b>C CODE MATLAB.....</b>	<b>145</b>
C.1 PROGRAMME APPELANT .....	145
C.2 FONCTION DE LECTURE DREAD_MK5 .....	147
C.3 FONCTION DE CONVERSION REPLACE_2BITS_VALUES.....	150
C.4 PROGRAMME POUR EXTRAIRE LES PICS DES INTENSITÉS AUTO-CORRÉLÉES.....	151
<b>D CODE ORIENTÉ OBJET C++ .....</b>	<b>155</b>
D.1 FICHIER DREAD_MK5.CPP POUR LA LECTURE DES DONNÉES .....	155
D.2 FICHIER MARK5B_AUTOCORR.CPP POUR LES AUTOCORRÉLATIONS.....	161
<b>E CODE DU FICHIER DE SOUMISSION (SIMULATIONS).....</b>	<b>167</b>
<b>F LISTE DES ÉQUATIONS RÉFÉRENCÉES DANS CETTE THÈSE .....</b>	<b>169</b>



## Liste des tableaux

<b>Tableau 2.1.</b> Valeurs typiques de température du bruit dans les observations radio du NRAO .....	29
<b>Tableau 6.1.</b> Redshift, délai temporel et masse minimum dus au milieu intergalactique .....	109
<b>Tableau 6.2.</b> Redshift et profondeur optique ( $\Omega_{1,0} = 1, 0.3$ et $0.2$ ) des 22 radiosources pour le modèle dominé par une constante cosmologique ( $\Omega_0 = 0.3$ et $\lambda_0 = 0.7$ ) .....	114
<b>Tableau 6.3.</b> Résultats pour trois méthodes de calcul à 95 % de confiance pour 22 objets .....	115
<b>Tableau A.1.</b> Polarisation circulaire droite ou gauche et fréquence maximale et minimale .....	136
<b>Tableau A.2.</b> Nom, coordonnées équatoriales, redshift et flux des 22 sources radio observées .....	137
<b>Tableau A.3.</b> Époque et durée d'observation des 22 radiosources. ....	138
<b>Tableau A.4.</b> Lieu, code, coordonnées géographiques et élévation des sept stations VLBA .....	138
<b>Tableau A.5.</b> Code des stations VLBA, nom et taille des fichiers d'entrée des 22 objets .....	139



# Liste des figures

<b>Figure 1.1.</b> Plages de limites supérieures sur la densité cosmique de minilentilles compactes .....	6
<b>Figure 2.1.</b> Modèle unifié d'un noyau galactique actif.....	18
<b>Figure 2.2.</b> Spectre logarithmique synchrotron générique.....	21
<b>Figure 2.3.</b> Schéma simplifié d'un système de réception hétérodyne.....	25
<b>Figure 3.1.</b> Schéma d'un système composé d'une seule lentille gravitationnelle ponctuelle .....	36
<b>Figure 3.2.</b> Image d'une source avec une lentille gravitationnelle à symétrie axiale au centre .....	39
<b>Figure 3.3.</b> Variation logarithmique temporelle d'amplification due à l'effet de microlentille .....	48
<b>Figure 4.1.</b> Schéma simplifié d'un interféromètre à une lentille gravitationnelle.....	56
<b>Figure 4.2.</b> Numérisation d'un champ électrique sur une chaîne de bits .....	65
<b>Figure 4.3.</b> Quantification sur deux bits menant à quatre niveaux de codage.....	68
<b>Figure 5.1.</b> Profil temporel d'autocorrélation d'intensité, obtenu pour un sous-échantillon de 1000 <i>data frames</i> (2.5 millions de points ou 39 ms), résultant de l'observation de 3C 345.....	75
<b>Figure 5.2.</b> Profil temporel d'intensité, obtenu pour un sous-échantillon de 2.5 millions de points (39 ms), correspondant à l'observation de l'objet PKS 1622-253. ....	76
<b>Figure 5.3.</b> Profil temporel d'autocorrélation d'intensité, obtenu pour un sous-échantillon de 2.5 millions de données (39 ms), provenant de l'observation du quasar PKS 1622-253.....	77
<b>Figure 5.4.</b> Histogramme de la fréquence relative des amplitudes des intensités auto-corrélées pour un sous-échantillon de 2.5 millions de points de l'objet 3C 345.....	79
<b>Figure 5.5.</b> Illustration du profil temporel (temps positifs) des intensités auto-corrélées, provenant du fichier de sortie (600 Go) du quasar 3C 345, pour la station SC. ....	82
<b>Figure 5.6.</b> 70 points ( $\sim 1 \mu\text{s}$ ) du profil d'autocorrélation d'intensité résultant du faible signal périodique de $\sim 5 \text{ MHz}$ ( $\sim 200 \text{ ns}$ ) inséré pour observer PKS 1622-253. ....	84
<b>Figure 5.7.</b> Profil temporel d'autocorrélation d'intensité résultant de 650 secondes d'observation de la radiosource PMN 0134-0931. ....	86
<b>Figure 5.8.</b> Profil temporel (de 20 à 140 ms) d'autocorrélation d'intensité correspondant à l'observation de 3C 345 à la station SC (canaux impairs seulement).....	88
<b>Figure 5.9.</b> Demi profil temporel uniforme d'autocorrélation d'intensité de 3C 345 .....	93
<b>Figure 6.1.</b> Exemple d'histogramme de la fréquence relative des deux intensités possibles.....	102
<b>Figure 6.2.</b> Exemple de profil temporel d'autocorrélation d'intensité obtenu en simulant un délai temporel de 280 secondes et du bruit environnemental... ..	103
<b>Figure 6.3.</b> Agrandissement de la figure 6.2. ....	104
<b>Figure 6.4.</b> Exemple de profil temporel d'autocorrélation d'intensité obtenu en simulant un délai temporel de 280 secondes et du bruit environnemental fort. ....	105
<b>Figure 6.5.</b> Agrandissement de la figure 6.4. ....	106
<b>Figure 6.6.</b> Exemple des 10 premières millisecondes du profil temporel d'autocorrélation d'intensité obtenu en simulant un très bref délai temporel de 1 ms et du bruit local fort.....	107

**Figure A.1.** En-tête d'un data frame de format Mark 5B..... 135

## Remerciements

L'écriture de cette thèse de doctorat concrétise quatre années de travail pour réaliser ce projet expérimental de physique, mais aussi d'informatique. Avant tout, mes remerciements s'adressent à mon directeur de recherche, Ermanno F. Borra, pour sa confiance témoignée et son financement (<sup>1</sup>CNRSG) afin que je réalise ce projet. De par sa confiance en moi, j'ai pu entre autres développer une autonomie d'expérimentation. Merci aussi Ermanno pour ta grande disponibilité et ton dynamisme contagieux. Merci encore d'avoir fait de ma thèse une expérience tout aussi divertissante qu'enrichissante.

Puis, je tiens à remercier l'astrophysicien du <sup>2</sup>NRAO, le Dr. Jonathan Romney, pour avoir accepté de prendre pour nous les observations radio d'objets astronomiques. Sans sa collaboration, ce projet n'aurait jamais pu se concrétiser.

Par suite, je désire aussi remercier les experts en calcul haute performance de Calcul Québec, spécialement Florent Parent et Maxime Boissonneault, lesquels ont accepté de prendre de leur précieux temps pour m'aider à adapter mon code Matlab en code C++ applicable pour le traitement en parallèle (dont je ne suis pas familier) sur le superordinateur Colosse de l'Université Laval.

Après quoi, je m'en voudrais de ne pas mentionner le soutien apporté par le département de physique, de génie physique et d'optique. Plus particulièrement, j'aimerais remercier le technicien en informatique Martin Blouin pour sa grande disponibilité, surtout lors du téléchargement de la quantité gigantesque de données requises pour la réalisation du projet.

Finalement, merci à ma conjointe Tania, à ma famille et à mes amis, surtout Charles Saint-Hilaire, pour m'avoir soutenu et pour avoir cru en mes capacités.

---

<sup>1</sup> Cette recherche a été soutenue par le Conseil de recherches en sciences naturelles et en génie du Canada CRSNG.

<sup>2</sup> Le National Radio Astronomy Observatory est un établissement de la US National Science Foundation exploité aux termes de l'accord de coopération par l'Associated Universities, Inc.



# 1 Introduction

## 1.1 Mise en contexte, problématique et objectifs

Le sujet de cette thèse expérimentale de doctorat porte sur le développement informatique d'une technique numérique visant à détecter des effets de minilentilles gravitationnelles sur des objets radio, et aussi à détecter des variations temporelles rapides. Essentiellement, cette technique de détection indirecte est basée sur l'autocorrélation de fluctuations temporelles d'intensité pour détecter et interpréter la présence d'un signal (délai temporel) cosmique (ou artificiel). Le lecteur doit savoir que la théorie de base des fluctuations d'intensité a été diffusée et confirmée expérimentalement par des expériences de laboratoire publiées (ex., Borra 2011). Cela dit, les principaux objectifs de ce projet expérimental sont les suivants:

Objectif scientifique général:

Appliquer, pour la première fois, la technique numérique d'autocorrélation d'intensité à des données interférométriques radio standards issues d'observations de noyaux actifs de galaxies (NAGs) tels que des quasars.

Objectifs spécifiques:

- Concevoir les algorithmes et les logiciels de la technique numérique;
- Valider la technique et les logiciels développés par la détection de signaux artificiels périodiques faibles et rapides;
- Montrer l'efficacité de détection de signaux de la technique au moyen de simulations numériques;
- Appliquer la technique pour détecter des effets de minilentilles gravitationnelles et/ou de faibles variations temporelles non périodiques;
- Confirmer ou infirmer les effets de minilentilles gravitationnelles via l'autocorrélation d'intensité sur de courtes échelles de temps (0.001 – 600 s);
- Faire une estimation des limites supérieures de la densité de minilentilles cosmologiques à partir d'une analyse statistique simplifiée basée sur seulement 22 objets.

Un modèle pratique consiste à supposer qu'un défecteur gravitationnel génère deux faisceaux, ayant une différence de parcours optique (DPO), qui interfèrent. Dans cette hypothèse, le défecteur peut être modélisé par un interféromètre de Young. Dès lors, la technique d'autocorrélation d'intensité devient particulièrement intéressante dans ce domaine, car elle permet de mesurer des effets d'interférence pour de grandes DPOs (délais temporels), à savoir, des DPOs plus grandes que la longueur (temps) de cohérence de l'onde électromagnétique, comme on s'y attend pour des effets de lentilles gravitationnelles. À l'opposé, les techniques standards d'interférométrie ont des limites, par exemple, les délais temporels doivent être plus courts que le temps de cohérence, et ne sont pas vraiment utiles dans ce domaine. Ceci étant acquis, les délais temporels entre les images multiples produites par des effets de lentilles gravitationnelles peuvent être utilisés pour diverses études (Blanford & Narayan 1992). À présent, pour trouver ces délais, il faut mesurer des variations temporelles de luminosité d'objets astronomiques. Ces mesures prennent beaucoup de temps (ex., plusieurs mois) et sont limitées aux lentilles résolues dont les sources doivent être variables. En pratique, cela nécessite des lentilles de masse quasiment galactique, lesquelles génèrent des délais temporels sur de longues échelles de temps (jours, mois, années). En contrepartie, notre technique d'autocorrélation des fluctuations d'intensité, fluctuations causées par des effets de superpositions d'ondes, permet également de trouver de courts délais temporels, aussi petits que la milliseconde pour les ondes radio (la dispersion causée par le milieu intergalactique empêchant d'avoir de plus petits délais), caractérisant la signature de lentilles gravitationnelles non résolues telles que des minilentilles cosmologiques (Borra 2008).

La matière sombre compacte, qui pourrait représenter près du quart du contenu de l'univers, est baptisée matière sombre puisqu'elle ne semble pas émettre, ni même interagir, par l'intermédiaire de la force électromagnétique. Elle pourrait constituer la majorité de la masse des galaxies, y compris celle de la Voie lactée, et pourrait former des halos de matière sombre. En outre, parmi les hypothèses émises pour expliquer la matière sombre, l'une d'entre elles stipule que cette matière serait due à la présence de trous noirs dans les galaxies. Par ailleurs, de nombreuses études

sur la croissance des structures à grandes échelles de l'univers montrent que les halos de matière sombre s'étendent sur une large gamme de masses, allant des grandes masses associées aux galaxies et aux amas de galaxies jusqu'à de plus petites masses associées aux halos de galaxies naines et aux petites structures comme des sous-halos. Plus encore, ces études prédisent qu'une fraction importante de la matière cosmique serait constituée de sous-structures de halos de matière sombre se formant à tous les décalages vers le rouge (le terme anglais, *redshift*, sera dorénavant utilisé pour simplifier l'écriture). Également, les objets sombres compacts pourraient être aléatoirement distribués le long de la ligne de visée, de sorte que l'on pourrait spéculer qu'ils «flottent» librement et qu'ils remplissent l'espace intergalactique. La très faible luminosité de ces objets fait en sorte qu'ils sont pratiquement indétectables avec les techniques conventionnelles. Pour cette raison, une méthode de détection indirecte est requise.

La technique d'autocorrélation d'intensité développée dans cette thèse pourra servir à détecter des effets de minilentilles gravitationnelles causés par cette matière sombre compacte. En ce qui concerne la détection, la nature de ces objets sombres importe peu, à condition qu'ils soient suffisamment compacts pour produire des effets de minilentilles à des distances cosmologiques et qu'ils aient des masses comprises entre 10 et 10 millions de masses solaires (voir section 6.3). Notez que par objet compact, on entend ici tout objet dont la taille est inférieure à son rayon d'Einstein (Schneider 1993).

Les candidats pour les minilentilles que notre technique cherche à détecter sont en majorité des objets sombres non-baryoniques compte tenu que plus de 80 % de la matière sombre est non-baryonique et que la densité des baryons dans l'univers est déjà connue via la nucléosynthèse primordiale: elle correspond à environ 4 % de la densité critique. Par exemple, les minilentilles ciblées pourraient être des sous-halos contenant des concentrations compactes de masse de particules exotiques non-baryoniques, mais aussi de trous noirs primordiaux suffisamment massifs (ex.,  $10 M_{\odot}$  –  $10^3 M_{\odot}$ ), quoique plus rares, formés au tout début de l'univers juste après le Big

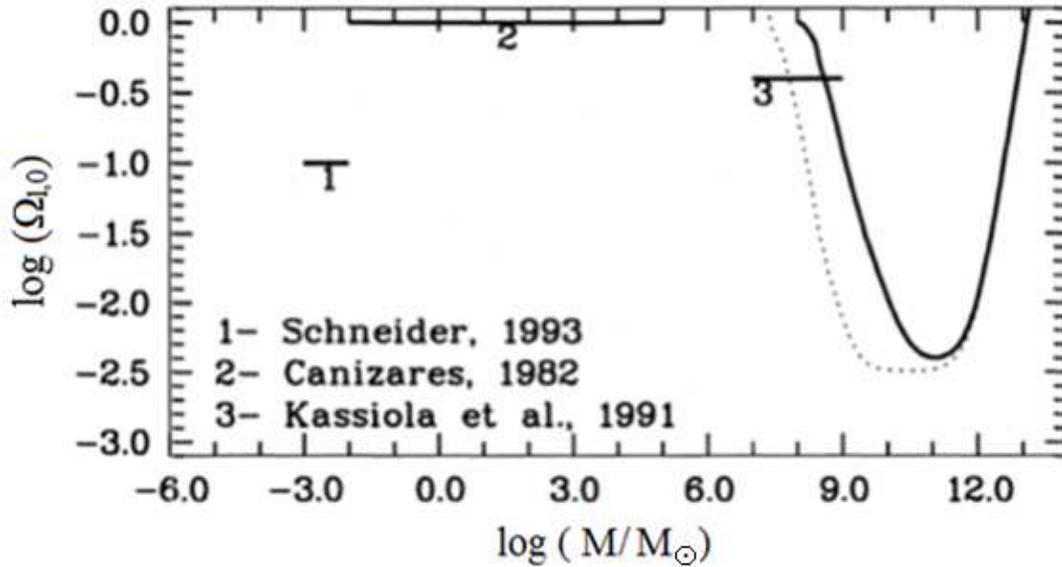
Bang. À ce propos, notre technique permettrait de confirmer ou d'infirmier la présence d'objets sombres non-baryoniques. Dans le cas où on ne découvre pas de tels objets, le résultat nul de notre recherche pourrait tout de même servir à mettre des contraintes supérieures intéressantes sur la densité cosmologique nette de ces objets condensés non-baryoniques dans l'intervalle de masse considéré.

Notre technique de détection indirecte permettrait aussi de détecter des objets exotiques, tels que des quasars ou des pulsars, qui émettraient des impulsions périodiques très rapides. Cette hypothèse est d'ailleurs validée par la découverte de modulations périodiques dans les spectres de galaxies et de certaines étoiles (Trottier 2012; Borra 2013; Borra & Trottier 2016): ces modulations pourraient être causées par l'émission d'impulsions séparées par moins d'un dixième de nanoseconde (Borra 2010). Avec les techniques courantes, il serait très ardu de trouver ces impulsions périodiques, car les périodes peuvent être très courtes et la plupart de ces objets exotiques pourraient être de faibles objets lointains (ex., quasars). Enfin, il serait possible que certains objets exotiques émettent des impulsions rapides non périodiques, comme les impulsions générées par des sursauts radio rapides (FRBs pour *Fast Radio Bursts*). Ces FRBs seraient alors très difficiles à détecter avec les techniques conventionnelles.

L'utilisation de la fonction d'autocorrélation pour trouver des variations périodiques est une technique connue. Par contre, la portée innovatrice de notre travail est que nous la démontrons en utilisant des données radioastronomiques standards. Un avantage majeur de l'utilisation de l'autocorrélation, que nous allons d'ailleurs démontrer, est qu'elle peut trouver des variations temporelles non périodiques, ce que, par exemple, les transformées de Fourier ne peuvent pas faire. À notre connaissance, l'autocorrélation n'a jamais été utilisée pour trouver des variations non périodiques, encore moins dans le radio.

D'autres techniques permettent également de trouver des signaux périodiques (Ransom et al. 2002) ou transitoires (Law & Bower 2012) dans des données numériques. Dans la conclusion, nous comparons notre technique aux autres techniques en illustrant les avantages de la nôtre.

Des recherches approfondies sur les quasars subissant des effets de lentilles gravitationnelles ont déjà imposé des limites sur les objets compacts de masse supérieure à  $10^7$  masses solaires (ex., Kassiola et al. 1991; Wilkinson et al. 2001). De plus, d'autres études théoriques telles que celles de Canizares (1982), Schneider (1993), Dalcanton et al. (1994) et Zackrisson & Bergvall (2003) ont permis de contraindre la densité cosmique d'objets compacts de masse solaire et sous-solaire (ex.,  $10^{-4} M_{\odot} - 10 M_{\odot}$ ). En revanche, il n'y a actuellement pas de limites vraiment strictes sur les objets compacts entre  $10^2$  et  $10^7$  masses solaires (voir figure 1.1). Voilà pourquoi, malgré le faible échantillon-test de seulement 22 objets, on a décidé de faire, via une analyse statistique simple, un estimé de l'abondance actuelle cosmologique de minilentilles, essentiellement non-baryoniques, intergalactiques uniformément distribuées dans l'intervalle de masse,  $10 M_{\odot} \leq M \leq 10^7 M_{\odot}$  (ce qui correspond grosso modo à la région 2 de la figure 1.1). Pour connaître les masses limites et les contraintes imposées sur les objets cosmologiques, pour une vaste plage de masses s'étendant de  $10^{-18} M_{\odot}$  à  $10^{16} M_{\odot}$ , le lecteur est invité à consulter les tables 4 et 5 de Dalcanton et al. (1994).



**Figure 1.1.** Plages de limites supérieures sur la densité cosmologique de minilentes compactes,  $\Omega_{l,0}$ , en fonction de la masse  $M/M_{\odot}$  (en unités de masse solaire). – Sources (Claeskens & Surdej 1996; Claeskens 2004).

Afin de pouvoir détecter des effets de minilentes gravitationnelles, il faut mettre au point les algorithmes et les logiciels de base. Pour ce faire, on développe d’abord un code Matlab pour traiter les données reçues du NRAO (*National Radio Astronomy Observatory*) sur un microordinateur. À cet égard, les données radio ont été prises par un astrophysicien du NRAO (Dr. J. Romney) qui a accepté de collaborer avec nous pour ce projet.

Du point de vue technique le problème et la solution sont succinctement décrits dans les paragraphes qui suivent:

- Le problème réside dans le fait que les microordinateurs, utilisés avec Matlab, ne sont pas assez puissants et n’ont pas suffisamment de mémoire pour traiter la grande quantité de données.

- La solution est d'utiliser un superordinateur et un autre logiciel pour traiter les volumineux fichiers de données. Après quoi, on traduit (adapte) le code développé en Matlab en un code de langage de programmation orienté objet (POO) C++ utilisable sur un superordinateur (voir section 5.2.3). Finalement, ce logiciel C++ est appliqué pour traiter (auto-corréler) la grande quantité de données provenant des gros fichiers du NRAO.

## **1.2 Demande d'allocation de ressources additionnelles au CAR**

Le traitement des données avec le langage POO C++ requiert plus de mille cœurs (1024) de calcul sur le superordinateur Colosse de l'Université Laval (voir section 5.2.3). Malheureusement, l'allocation par défaut des ressources de Calcul Canada est limitée à 240 cœurs sur Colosse. De ce fait, il a fallu soumettre notre projet au CAR (concours d'allocation des ressources) pour demander les 1024 cœurs de mémoire requis. Après avoir étudié notre requête, le jury du CAR, composé d'un physicien et d'un expert informatique en calcul haute performance (CHP), a jugé que le mérite scientifique et l'aspect technique de notre projet répondaient aux conditions d'attribution de ces 1024 cœurs de calcul. N'empêche qu'il a fallu annuellement renouveler notre demande au CAR afin de conserver les ressources supplémentaires.

## **1.3 La longue histoire de la déflexion lumineuse gravitationnelle**

Les effets de lentilles gravitationnelles sont considérés comme un domaine relativement nouveau en astrophysique. Pourtant, l'histoire reliée à la déviation de la lumière remonte à plus de trois siècles. Considérant que la lumière pouvait être composée d'un ensemble de corpuscules en mouvement (les photons), Newton a d'abord suggéré, dès 1704, que le champ gravitationnel d'un objet massif pourrait éventuellement dévier des rayons lumineux, au même titre qu'il fait dévier la trajectoire des particules matérielles. Un siècle plus tard, Soldner calcule expressément l'angle de déflexion de la lumière en suivant la mécanique

newtonienne. Il dérive alors un angle de déviation de 0.875 seconde d'arc pour un rayon lumineux rasant le limbe solaire.

En 1911, Einstein utilise son principe d'équivalence, tout en supposant une métrique euclidienne, pour étudier ce phénomène. Par exemple, il retrouve les résultats de Soldner (1804). En 1916, Einstein termine la plus grande partie de sa théorie de la relativité générale. De par cette théorie, l'angle de déflexion trouvé est de 1.75 seconde d'arc, correspondant au double de celui dérivé par Soldner.

Quelques années plus tard, Eddington et ses collaborateurs (Dyson et al. 1920) confirment, avec une incertitude de  $\lesssim 30\%$ , l'angle de déviation prédit par la relativité générale. En fait, aujourd'hui, l'incertitude avec la valeur théorique de la relativité générale est plutôt,  $\lesssim 1\%$ , permettant ainsi d'exclure quasiment toutes les autres théories de la gravitation.

En 1937, Zwicky propose de rechercher des effets de lentilles gravitationnelles entre des galaxies angulairement proches, mais situées à des redshifts différents. Dans ce cas, les séparations entre les images multiples produites peuvent être détectables à partir d'un observateur terrestre. Zwicky (1937a, 1937b) dévoile entre autres que les observations de sources subissant des effets de lentilles permettraient de mesurer directement la masse des galaxies, incluant celle de la matière sombre, et de détecter des galaxies distantes.

Après les publications de Zwicky en 1937, le sujet des lentilles gravitationnelles s'est éteint, avec pratiquement aucune nouvelle publication pendant plus de deux décennies. Heureusement, en 1964, Refsdal propose une méthode permettant de mesurer les distances des quasars (découverts en 1963) à des échelles cosmologiques. En effet, il remarque qu'il y a un délai temporel entre les composantes distinctes de certaines sources lointaines (quasars, supernovas), et que ce délai pourrait être mesuré. Compte tenu que ce délai temporel est inversement proportionnel à la constante de Hubble  $H_0$ , Refsdal suggère d'utiliser les effets de

défecteurs gravitationnels pour mesurer  $H_0$ . Plus encore, il s'avance même à déclarer que ces effets pourraient être appliqués pour vérifier la validité de différentes théories cosmologiques.

Après maintes tentatives infructueuses de détections de systèmes de lentilles gravitationnelles, Walsh, Carswell & Weymann découvrent finalement, en 1979, le premier système de lentilles gravitationnelles, soit le quasar à structure double Q0957+561. Depuis ce temps, la recherche en astronomie sur les lentilles gravitationnelles est devenue un sujet d'actualité (ex., Blanford & Naryan 1992). Justement, au cours des dernières décennies, plusieurs centaines de nouveaux systèmes de lentilles gravitationnelles ont été découverts donnant, par exemple, des arcs gravitationnels géants et des anneaux lumineux complets d'Einstein. À titre d'exemple, l'effet d'une microlentille, résultant d'un système formé d'une étoile source et d'une étoile lentille (système qu'Einstein avait rejeté), a été employé pour détecter des planètes extrasolaires de même que pour sonder des candidats à des objets sombres.

## **1.4 Organisation de la thèse**

Le chapitre 2 traite des outils de la radioastronomie qui sont les plus pertinents pour le projet. De prime abord, quelques aspects scientifiques fondamentaux sont exposés. On décrit entre autres les processus d'émission impliqués, les sources radio extragalactiques, le rayonnement synchrotron, les spectres et les deux principales propriétés (brillance et compacité) que doivent avoir les radiosources observées. Puis, les aspects instrumentaux et techniques sont discutés. On retrouve notamment des notions sur les télescopes radio (antenne, récepteur, détecteur, puissance collectée, température, etc.), sur les trois sources dominantes de bruit (bruit du récepteur, du sol et de l'atmosphère) et sur la sensibilité d'un système d'interférométrie à très longue base. Finalement, on discute de la dispersion des ondes radio résultant des différents milieux de propagation, particulièrement le plasma intergalactique, mais aussi le milieu galactique et le milieu de la galaxie hôte.

Le troisième chapitre donne seulement les aspects de la vaste théorie des lentilles gravitationnelles qui sont utiles au projet. On présente d'abord les différents régimes des lentilles, la théorie de base des lentilles ponctuelles et les délais temporels causés par ces dernières. Par suite, on discute des lentilles fortes et faibles. Après quoi, on explique les effets de microlentilles (ou minilentilles) extragalactiques sur les quasars à image simple et à images multiples. On distingue les variations de luminosité intrinsèques des variations extrinsèques induites par des microlentilles. Pour clore le chapitre, on explique les principes de base d'une méthode, très répandue, utilisant les courbes de lumière pour détecter des événements de microlentilles gravitationnelles dans les quasars.

Le chapitre 4 porte essentiellement sur l'autocorrélation d'intensité et sur le traitement numérique des signaux. La première partie traite de la modélisation du champ électrique d'un rayonnement synchrotron, des intensités et des fluctuations d'intensité, de l'interféromètre de Young comme modèle d'un défecteur gravitationnel, du signal auto-corrélé issu de sources compactes et étendues, ainsi que des avantages et des désavantages de la technique d'autocorrélation d'intensité. Pour ce qui est de la seconde partie, elle porte sur les signaux numérisés. Cette partie est loin d'être exhaustive, mais dévoile plutôt les notions de base sur les signaux qu'il faut savoir en radioastronomie, comme les notions de signaux aléatoires, d'échantillonnage, de quantification sur deux bits, de bruit gaussien et de bruit de quantification.

Le cinquième chapitre constitue en quelque sorte le cœur du travail de cette thèse expérimentale de physique, mais aussi d'informatique. Il est principalement consacré à la méthodologie et au traitement rigoureux des données, à savoir, le développement informatique de la technique numérique d'autocorrélation d'intensité, les validations des deux logiciels (Matlab et C++) développés, de même que les techniques utilisées pour extraire des candidats (signaux) potentiels aux événements de minilentilles gravitationnelles. Entre autres choses, on décrit comment se fait la lecture et l'autocorrélation des données dans les deux logiciels. Également, des

exemples de détections de signaux artificiels forts, faibles, périodiques et non périodiques valident la robustesse de la technique numérique et des codes Matlab/C++ développés. Le chapitre se termine en décrivant les trois principales étapes pour détecter un signal auto-corrélé (délai temporel) à travers le bruit, idéalement d'origine cosmique tel qu'un effet de minilente, mais aussi d'origine artificielle comme un effet instrumental ou un signal RFI ou encore des trous dans les données.

Le chapitre 6 est surtout consacré aux simulations et à l'analyse des résultats. Après un sommaire sur l'analyse (statistique) simplifiée utilisée pour estimer la densité cosmologique de minilentes, on décrit les quatre principales étapes des simulations numériques, lesquelles se résument à imposer artificiellement un délai temporel et du bruit local. Le but de ces simulations consiste à montrer l'efficacité de détection d'un signal (délai temporel) par l'autocorrélation d'intensité. Dans un deuxième temps, on fait une analyse statistique, via la loi binomiale, des probabilités d'avoir des événements de minilentes, pour un intervalle de masse allant de 10 à 10 millions de masses solaires, dans vingt-deux sources radio. Enfin, à partir de la profondeur optique, du résultat nul obtenu et de trois méthodes de calcul pour l'intervalle de confiance (barres d'incertitude) d'une binomiale, nommément les méthodes de Clopper-Pearson, de Wilson et de Jeffreys, on fait un estimé approximatif de la limite supérieure de la densité cosmologique de minilentes compactes non-baryoniques intergalactiques pour le modèle critique (actuellement favorisé) avec constante cosmologique  $\lambda_0 = 0.7$  et avec  $\Omega_0 = 0.3$ .

## 1.5 Contributions scientifiques

Ci-dessous, les deux publications qui ont été réalisées directement ou indirectement dans le cadre de cette thèse de doctorat:

1. Borra, E. F., Romney, J., & Trottier, E., 2018, “A technique to detect rapid flux time variations with radioastronomical data”, JOAA, accepté, 2018-04-23
2. Borra, E. F., & Trottier, E., 2016, “Discovery of Peculiar Spectral Modulations in a Small Fraction of Solar-type Stars”, PASP, 128, 114201

## 2 Rappels des principaux outils de radioastronomie

La radioastronomie permet, entre autres, de détecter l'émission radio de diverses sources comme la radiation issue de planètes, d'étoiles, de nébuleuses gazeuses, de nuages de gaz atomique ou moléculaire, de pulsars, de galaxies et de noyaux actifs de galaxies (NAGs) comme des quasars.

### 2.1 Éléments de radioastronomie qu'il faut maîtriser pour le projet

#### 2.1.1 Sources radio

##### 2.1.1.1 Généralités sur l'émission thermique et non thermique des radiosources

Le rayonnement des sources radio est souvent caractérisé par sa puissance transportée par unité de surface et par unité de fréquence. Cette puissance, nommée indûment, densité de flux ou tout simplement flux (le terme employé devrait plutôt être éclaircissement monochromatique), s'exprime généralement en  $\text{Wm}^{-2}\text{Hz}^{-1}$ . Pour les radiosources, les flux émis sont généralement très faibles, à tel point qu'en pratique l'unité de mesure utilisée est plutôt le Jansky, voire le mJy ( $1 \text{ Jy} = 10^{-26} \text{ Wm}^{-2}\text{Hz}^{-1}$ ). Pour un rayonnement de continuum, le flux radio est quasi constant pour une plage de fréquences proches de la fréquence où le flux est mesuré (Biraud & Rieu 1994).

Dans le cas d'un radiotélescope, l'antenne ou le capteur, collectant la puissance de la radiation est, électroniquement parlant, similaire à une résistance portée à la température  $T$ . Aux bornes de la résistance, il y a alors une différence de tension (un signal) dont la puissance  $P$ , fournie dans une bande de fréquences  $B$ , est donnée par la relation,

$$P = kTB, \tag{2.1}$$

où  $k$  est la constante de Boltzmann.

En radioastronomie, il est pratique de définir la densité de flux radio par une température équivalente de bruit, même si les mécanismes dominants d'émission radio sont souvent non thermiques. Pour l'émission continue thermique, on retrouve particulièrement l'émission d'un corps noir ou l'émission des électrons en équilibre thermique dans un gaz chaud ionisé. Sachant que la température de brillance d'une source radio thermique est physiquement limitée, comme une contrainte sur la température de brillance imposée par la température des électrons dans le gaz ionisé, il faut alors recourir à des types de rayonnement radio de nature non thermique, comme le rayonnement synchrotron (voir section 2.1.1.3), pour justifier les cas où les températures de brillance des radiosources excèdent les températures limites physiques (Biraud & Rieu 1994).

La radiation des sources de température physique  $T$  peut être spécifiée par la brillance d'un corps noir. Ceci implique que la brillance, notée  $B_\nu$ , du rayonnement thermique d'un objet source à une fréquence  $\nu$ , est déterminée par la distribution de Planck,

$$B_\nu = \frac{2h\nu^3}{c^2} \left[ \exp\left(\frac{h\nu}{kT}\right) - 1 \right]^{-1}, \quad (2.2)$$

avec  $c$  et  $h$  représentant respectivement la vitesse de la lumière et la constante de Planck.

Aux longueurs d'onde centimétriques, le terme  $h\nu$  est beaucoup plus petit que le terme  $kT$ , si bien qu'on peut approximer la loi de Planck par celle de Rayleigh-Jeans. Sachant cela, la relation précédente se réduit à,

$$B_\nu = \frac{2\nu^2 kT}{c^2} = \frac{2kT}{\lambda^2}. \quad (2.3)$$

Pour l'émission non thermique des objets, comme le rayonnement synchrotron des 22 sources radio observées (voir tableau A.2), on peut tout de même définir une température de brillance, notée  $T_b$ , (contrairement au rayonnement thermique où  $T_b$  est une température physique, ici  $T_b$  est une température radiative apparente), de façon à ce que le terme de droite de l'équation (2.3) devienne égal à la luminance de l'objet source à la fréquence d'observation (8.4 GHz pour le NRAO). Mathématiquement, cela se traduit par,

$$T_b = \frac{B_\nu c^2}{2\nu^2 k} = \frac{B_\nu \lambda^2}{2k}. \quad (2.4)$$

### 2.1.1.2 Sources radio cosmologiques: les quasars

Plusieurs des radiosources extragalactiques fournies par le NRAO sont des quasars. Ces derniers sont les sources radio les plus lumineuses et les plus distantes détectées dans les observatoires radio. Actuellement, on connaît plus de cent mille quasars et on en estime autour d'un milliard. Or, pour observer ces quasars très lointains, il faut recourir à de l'instrumentation à forte résolution spatiale, par exemple, l'interférométrie à très longue base (VLBI pour *Very Long Baseline Interferometry*) du NRAO (Kellermann & Frazer 1988; Thompson et al. 2004; Léna 2008).

La lumière reçue des quasars contient des informations sur la distribution de la matière sombre le long de la ligne de visée. Dans le contexte de la matière sombre, deux phénomènes importants impliquent les propriétés observées des quasars, à savoir, l'absorption de gaz et, prioritairement dans cet ouvrage, les effets de lentilles gravitationnelles.

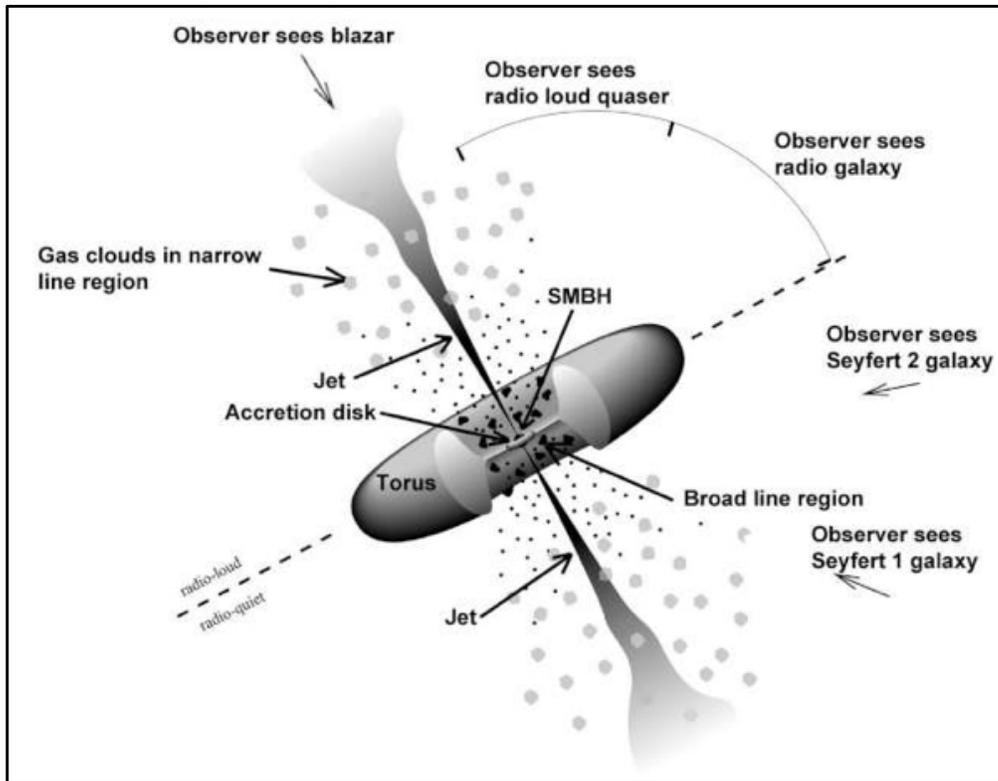
La structure globale typique d'un quasar est composée, entre autres, 1) du quasar lui-même, soit une structure centrale ultra compacte, de 0.01 à 1 pc, située dans le noyau de la galaxie hôte (noyau du quasar et la galaxie qui l'héberge), 2)

d'une structure étendue comprenant deux grandes régions, dénommées lobes radio, plus ou moins régulières et symétriques par rapport au quasar, et 3) de deux jets relativistes radio, plus ou moins visibles, dont chacun d'eux relie le quasar à l'un des deux lobes (Kellermann & Frazer 1988; Collin-Zahn 2009; Burke & Graham-Smith 2010).

Dans le modèle standard, les quasars représentent un type (voir figure 2.1) de galaxies à noyau actif (Elvis 2000; Beckmann & Shrader 2013). Leur activité ne résulte pas d'étoiles ordinaires, elle fait plutôt intervenir d'autres phénomènes de la physique. Leurs spectres ne montrent pas les caractéristiques des étoiles: ils émettent des raies qui ne peuvent pas se former dans les atmosphères stellaires. L'extrême luminosité intrinsèque totale des quasars varie d'un quasar à l'autre, de  $\sim 10^{45}$  à  $\sim 10^{48}$  ergs/s, mais leur émission est théoriquement limitée par la luminosité d'Eddington. Ceci étant acquis, il devient alors légitime de se demander quel mécanisme pourrait fournir une telle puissance. Un des mécanismes plausibles stipule que les quasars sont alimentés au moyen de l'accrétion de la matière environnante (étoiles, gaz et poussières) par un attracteur supermassif central, généralement un trou noir, ou, à tout le moins, une forte concentration de masse de  $10^6$  à  $10^9$  masses solaires. Pour tout dire, le trou noir est un corps très compact dont la gravité extrême empêche même la lumière de s'en échapper. À titre d'exemple, la masse du trou noir de notre Galaxie est estimée à 4.5 millions de masses solaires (Ghez et al. 2008), correspondant à une luminosité d'autour de  $5 \times 10^{44}$  ergs/s. Vu que le trou noir exerce une attraction gravitationnelle sur la matière qui l'entoure, celle-ci se rapproche graduellement du trou en spiralant avant de s'y engouffrer. De là, le moment cinétique non nul de la matière retombant sur elle-même fait en sorte qu'il se forme un disque tournoyant autour de ce corps invisible massif, appelé disque d'accrétion (Lynden-Bell 1969). En se rapprochant progressivement du trou noir, la dissipation (frottement et cisaillement) visqueuse au sein de la matière, qui est probablement due au transfert du moment cinétique vers l'extérieur du disque, réchauffe le disque ( $\geq 100000$  K) jusqu'à le porter à l'incandescence. L'énergie générée par frottement et par cisaillement quitte alors le disque en rotation en émettant du rayonnement intense sur une large bande de longueurs d'onde (Sheilds 1978). De chaque côté du trou noir

central, la matière s'échappe par l'intermédiaire d'une paire de jets cosmiques relativistes étroitement collimatés et pouvant parcourir de très grandes distances. Ces jets augmentent l'intensité du rayonnement du faisceau qui s'approche de l'observateur, mais diminuent celle du faisceau opposé (qui s'éloigne) au point où il peut devenir très difficile d'observer ce faisceau. Disons enfin que ce sont ces jets relativistes, lesquels se nourrissent de particules énergétiques et de l'énergie des champs magnétiques dans les lobes radio, qui forment la structure à doubles lobes de la radiosource. Dans l'ensemble, retenons que les quasars sont des galaxies (actives) dont le noyau renferme un trou noir supermassif qui produit une gigantesque quantité d'énergie essentiellement due à l'accrétion de la matière environnante.

Il existe quatre principaux types de noyaux galactiques actifs, de formes différentes, comme les galaxies de Seyfert (I et II), les blazars, les radiogalaxies, et évidemment, les quasars. On pourrait aussi ajouter les LINERS pour *Low-Ionization Nuclear Emission-line Region*. Les types (formes) de NAGs dépendent de l'importance relative des différentes composantes. Plusieurs de ces formes distinctes proviennent tout simplement des différences intrinsèques dans l'intensité relative de l'activité des diverses composantes: la fraction de l'intensité de l'émission totale provenant du coeur, du disque, de la paire de jets et des lobes. Pourtant, les plus importantes différences de formes des NAGs proviennent de l'angle avec lequel l'observateur voit ces NAGs. À cet égard, la figure 2.1 montre un modèle unifié d'un noyau actif de galaxie et les types possibles de NAGs selon l'angle de vue. Notez que pour les quasars, l'angle de vue est inférieur à 45 degrés (Zackrisson 2005; Beckmann & Sharader 2013).



**Figure 2.1.** Modèle unifié d'un noyau galactique actif. Le type de NAG dépend essentiellement de l'angle de vue de l'observateur. Selon sa position, l'observateur aperçoit différentes facettes du noyau actif. Par exemple, si l'angle de vue est perpendiculaire au disque, soit exactement dans l'axe des jets cosmiques, il s'agit alors d'un blazar. (SMBH pour *Super Massive Black Hole*). – Source (<http://fermi.gsfc.nasa.gov/science/eteu/agn>).

### 2.1.1.3 Émission radio des quasars: le rayonnement synchrotron

Les régions d'émission des sources radio comprennent généralement une région compacte, opaque à son propre rayonnement, et une région étendue transparente. La plupart des radiosources compactes, dominées par leur cœur, sont identifiées avec les noyaux galactiques actifs ou avec les quasars radio-silencieux (QSOs). Pour leur part, les radiosources étendues, dominées par leurs lobes radio, sont associées aux galaxies, mais aussi aux quasars qui n'ont aucune étendue optique apparente. La plupart de ces objets étendus contiennent tout de même une composante centrale compacte radio (Kellermann & Frazer 1988; Burke & Graham-Smith 2010).

Le mécanisme d'émission de continuum radio (sans fréquences privilégiées) résulte de processus où les échanges d'énergie ne sont pas quantifiés, si bien que les photons émis peuvent avoir une distribution d'énergie continue. Les processus qui entraînent des émissions de continuum sont essentiellement les processus thermiques (ex., rayonnement d'un corps noir) et ceux non thermiques (ex., rayonnement synchrotron). Dans le domaine des observations centimétriques du NRAO (ex., 3.6 cm), le mécanisme dominant d'émission de continuum radioélectrique des quasars n'est pas de nature thermique. Cela signifie que les échanges entre le rayonnement et les particules élémentaires (ex., électrons) ne se contrebalancent pas. En d'autres termes, le fait que les électrons soient beaucoup plus énergétiques que les photons empêche tout équilibre thermique possible (Collin-Zahn 2009).

L'émission radio des quasars possède entre autres deux signatures permettant de la reconnaître. D'une part, elle est polarisée. La majorité des quasars ont de faibles polarisations ( $\leq 3\%$ ). Par exemple, une forte proportion de l'émission synchrotron montre un faible degré de polarisation, ce qui indique la distribution des champs magnétiques au sein du quasar. Une minorité de quasars, plus polarisés, sont des sources radio compactes dont l'émission donne généralement des spectres plats. D'autre part, la distribution spectrale de ce rayonnement radio suit une loi de puissance (voir équation (2.5)). Au fond, ces deux signatures sont typiques d'une émission non thermique, à savoir, l'émission radio synchrotron (Kellermann & Frazer 1988; Collin-Zahn 2009).

Le rayonnement de continuum radioélectrique dominant des quasars est donc le rayonnement synchrotron. Cette forte émission radio provient d'une paire de jets énergétiques de particules élémentaires (électrons) jaillissant du centre de la galaxie hébergeant le quasar. Elle est émise par des électrons se déplaçant à des vitesses relativistes dans un champ magnétique de l'ordre de quelques micro-Gauss. L'énergie totale des électrons atteint jusqu'à  $10^{60}$  ergs. L'émission synchrotron est très intense et fortement concentrée, ce qui signifie qu'elle semble provenir d'un cône très étroit. Elle est émise sur une grande fourchette d'énergies, produisant ainsi un spectre

d'énergie large. Ce type d'émission peut être observée partout où il y a des électrons relativistes et des champs magnétiques. Dans notre cas, le rayonnement synchrotron provient de jets cosmiques relativistes, émanant de NAGs (ex., quasars), qui sont situés à proximité des trous noirs.

Notez enfin que le lecteur intéressé aux principes physiques et aux équations mathématiques de base de la radiation synchrotron est invité à consulter les manuels de Rybicki & Lightman (1979), de Wilson, Rohlfs & Huttemeister (2009) et de Longair (2011).

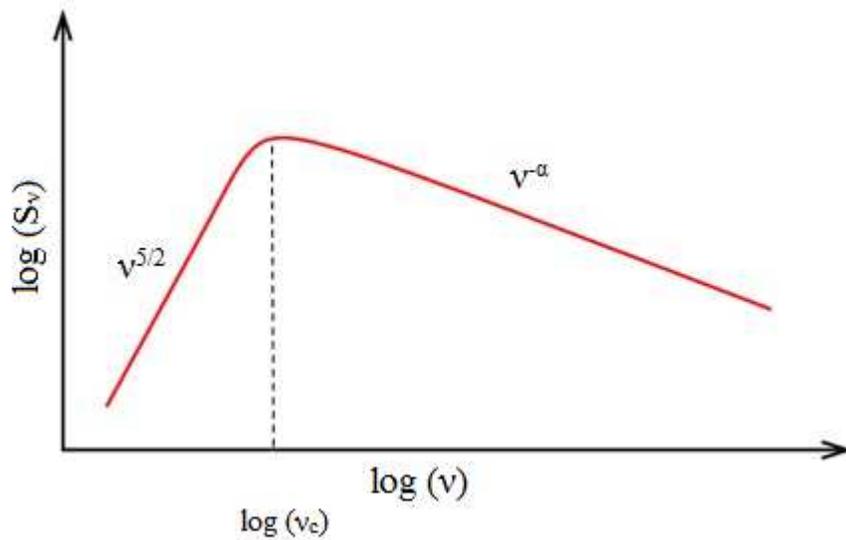
#### 2.1.1.4 Spectres des quasars

Le fait que les spectres de quasars ne présentent pas vraiment de traits caractéristiques (hormis la raie 21 cm de HI) limite les observations aux mesures de spectres continus. Cela dit, du rayonnement radio jusqu'au rayonnement gamma, les spectres de maints quasars s'étendent sur plus de dix échelles de grandeur du spectre électromagnétique. Dans cette plage, l'émission provient de différentes sources, certaines thermiques d'autres non thermiques. L'intensité relative de ces diverses sources d'émission (cœur, disque, jets et lobes) dépend de l'orientation (angle de vue) de la galaxie à noyau actif. En dépit du mélange des différentes sources d'émission, l'effet global de toutes ces radiations produit un spectre dont le flux à une fréquence  $\nu$ , noté  $S_\nu$ , obéit à une loi en puissance de la forme,

$$S_\nu \propto \nu^{-\alpha}, \quad (2.5)$$

où  $\alpha$  représente l'indice spectral. La figure 2.2 montre la forme logarithmique générale d'un spectre d'origine synchrotron. La première partie du spectre, dont le rayonnement est d'origine thermique, est associée au phénomène d'auto-absorption. En effet, aux basses fréquences, on observe une coupure dans le spectre d'origine synchrotron. Le coefficient d'absorption diminue lorsque la fréquence augmente. Ceci implique qu'il y a une fréquence de coupure ( $\nu_c$ ) sous laquelle le milieu apparaît

comme optiquement épais. En deçà de cette fréquence de coupure, il se produit une réabsorption de l'émission par les électrons synchrotron eux-mêmes. C'est ce qu'on appelle l'auto-absorption synchrotron. Dans ce cas, le flux augmente avec la fréquence au lieu de diminuer. Quant à la seconde partie, de pente négative (et habituellement voisine de l'unité), elle provient du rayonnement non thermique synchrotron proprement dit. Cela est causé par le fait que les électrons n'ont pas une seule énergie. Dans les faits, les électrons obéissent eux-mêmes à une loi de puissance: le nombre d'électrons à une énergie donnée est proportionnel à leur énergie élevée à la puissance,  $p = 2\alpha + 1$  (Wilson, Rohlfs & Huttemeister 2009; Longair 2011).



**Figure 2.2.** Spectre logarithmique synchrotron générique. Le terme  $\nu^{5/2}$  est associé à l'auto-absorption, alors que le terme  $\nu^{-\alpha}$  est associé à l'émission synchrotron. La fréquence de coupure  $\nu_c$  sépare les milieux optiquement épais et minces.

### 2.1.1.5 Choix des candidats pour les observations

Dans le but d'optimiser les chances de détecter des événements de minilentilles gravitationnelles, l'échantillon des objets choisis est majoritairement composé de sources radio brillantes et compactes. Leur flux moyen (voir tableau A.2) dépasse amplement l'intensité (1 mJy) du bruit du système (voir équation (2.16)).

Du point de vue de l'autocorrélation d'intensité, le choix de sources radio brillantes est très recommandé afin de détecter des signaux (ex., minilentilles). En effet, à l'encontre des objets trop faibles, les radiosources suffisamment brillantes permettent aux fluctuations d'intensité, à partir desquelles les détections d'effets de minilentilles (délais temporels) deviennent possibles, de sortir du bruit (voir section 4.4.3).

Les radiosources compactes subissant l'effet d'un défecteur gravitationnel donnent, comparativement aux radiosources étendues, un signal des intensités auto-corrélées de plus grande amplitude et ayant un seul délai temporel (voir section 4.4.1). Par ailleurs, les spectres plats de ces objets condensés, d'indices spectraux  $\alpha$  voisins de zéro, conviennent très bien à la recherche de minilentilles gravitationnelles puisqu'ils ont tendance à être dominés par un seul cœur compact, de taille angulaire ordinairement inférieure à la milliseconde d'arc (sous-mas), qui est situé près du trou noir central dans le noyau de la galaxie hôte (Wilkinson et al. 2001). Outre cela, les sources radio compactes sont fréquemment variables et sont identifiées avec des galaxies montrant des spectres à raies d'émission larges dans leur noyau. En fait, plusieurs radiosources pourraient présenter des variations de flux dans leur noyau, mais ces variations sont plus faciles à distinguer dans les radiosources compactes où l'émission du noyau variable n'est pas combinée avec celle d'intensité constante issue des jets cosmiques. Finalement, compte tenu que le contraste de luminosité entre le cœur et les autres régions d'émission des sources radio (disque, jets et lobes) augmente avec la fréquence, il est aussi recommandé d'observer aux plus hautes radiofréquences possibles (ex., 8.4 GHz pour les observations de nos objets).

## 2.1.2 Radiotélescopes

Afin de détecter les signaux radio des sources astronomiques, les radiotélescopes utilisent principalement trois composantes, notamment l'antenne parabolique, le récepteur hétérodyne et le détecteur (Thompson et al. 2004; Wilson, Rohfls & Huttemeister 2009; Burke & Graham-Smith 2010).

L'antenne parabolique est l'antenne la plus utilisée en radioastronomie. Elle permet de convertir la radiation électromagnétique incidente en tension (courant) mesurable par les diverses composantes électroniques des télescopes radio. Elle consiste surtout à évaluer la puissance radiative totale reçue en sommant toutes les puissances radiatives ponctuelles d'un même front d'onde (Léna 2008).

Pour une radiosource compacte, le flux émis  $S$  (en Jy), est entièrement collecté par l'antenne du radiotélescope. Sachant cela, la puissance reçue par unité de surface, notée  $P^*$ , pour une largeur de bande  $B$ , est donnée par,

$$P^* = \frac{1}{2}SB, \quad (2.6)$$

où la présence du facteur  $\frac{1}{2}$  est pour tenir compte que l'on traite séparément la puissance reçue dans chacune des deux polarisations. Par exemple, le NRAO traite les deux polarisations du rayonnement émis avec des paires de canaux, soit des canaux pairs et impairs correspondant respectivement aux polarisations circulaires gauches et droites (voir tableau A.1).

L'antenne parabolique de surface  $A$ , qui collecte le signal de puissance,  $P = P^* A$ , dans  $B$ , est définie par sa température d'antenne (capteur)  $T_a$ , laquelle est obtenue des relations (2.1) et (2.6), soit,

$$T_a = \frac{AS}{2k}. \quad (2.7)$$

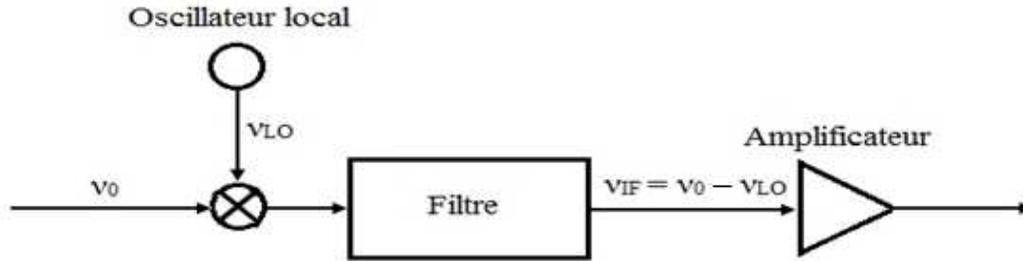
Pour une source radio étendue, les radioastronomes préfèrent travailler avec la brillance (monochromatique) plutôt qu'avec le flux. Dans ce cas, la puissance  $P$  captée par l'antenne est donnée par,

$$P = kT_a B = \frac{1}{2} B_\nu A \Omega B, \quad (2.8)$$

où, aux basses fréquences radio,  $B_\nu$  suit la loi de Rayleigh-Jeans (voir équation 2.3) et,  $A \Omega = \lambda^2$  ( $\Omega$  étant l'angle solide du lobe principale de l'antenne), donne l'étendue du rayonnement (Léna 2008). Il en ressort que les températures d'antenne et de brillance s'égalisent,  $T_a = T_b$ , de sorte que l'antenne parabolique et l'objet source s'équilibrent thermodynamiquement. Ceci étant dit, de l'équation précédente, il vient que,

$$T_a = \frac{1}{2k} B_\nu A \Omega = \frac{1}{2k} B_\nu \lambda^2. \quad (2.9)$$

L'émission des sources radio produit des signaux électriques peu intenses, mais de fréquences élevées. Sans entrer dans les détails techniques de l'électronique sous-jacente, disons simplement que la fonction du récepteur hétérodyne (voir figure 2.3) consiste à convertir les signaux de hautes fréquences reçus en des signaux de fréquences intermédiaires (IF). Ces signaux IF pourront alors être amplifiés et détectés par un détecteur quadratique ou par radiointerférométrie (Thompson et al. 2004; Burke & Graham-Smith 2010), comme le VLBI du NRAO, ou, en ce qui nous concerne, par la technique d'autocorrélation des fluctuations d'intensité (voir les chapitres 4 et 5).



**Figure 2.3.** Schéma simplifié d'un système de réception hétérodyne. Ce système fait appel à un oscillateur local (LO), de fréquence  $\nu_{LO}$  voisine à la fréquence d'observation  $\nu_0$ , puis à un filtre (passe-bas) pour transposer les signaux de fréquences élevées vers des signaux de fréquences intermédiaires (IF),  $\nu_{IF} = \nu_0 - \nu_{LO}$ , qui peuvent être amplifiés. Notez que la figure 6.2 de Thompson et al. (2004) montre un système de réception pour de la radiointerférométrie à deux antennes.

### 2.1.3 Bruit de photons et bruit d'ondes

La lumière et ses interactions avec la matière sont ici traitées à l'aide du modèle physique semi-classique (Mandel & Wolf 1965). Ce modèle suit un processus aléatoire de Poisson pour décrire la propagation du champ électrique par la théorie électromagnétique ondulatoire classique et les interactions avec la matière. Pour un tel modèle, la nature physique de la lumière mène à une source de fluctuations fondamentales lors de ses interactions avec la matière, ceci peu importe la nature du récepteur. Ceci étant acquis, d'après ce modèle, on tire, en insérant l'équation (2.2) dans l'équation (2.8) et en se rappelant que l'étendue du rayonnement est,  $A\Omega = \lambda^2 = (c/\nu)^2$ , une puissance, dans l'une des deux polarisations, par hertz valant (Léna 2008),

$$P_\nu = h\nu \left[ \exp\left(\frac{h\nu}{kT}\right) - 1 \right]^{-1}. \quad (2.10)$$

La variance moyenne temporelle des fluctuations thermodynamiques autour de la puissance, notée  $\langle \Delta P_v^2 \rangle$  (les chevrons  $\langle \rangle$  indiquant une valeur moyenne), est donnée par (Landau & Lifchitz 2004),

$$\langle \Delta P_v^2 \rangle = kT^2 \frac{dP_v}{dv} = P_v hv + P_v hv \left[ \exp\left(\frac{hv}{kT}\right) - 1 \right]^{-1}. \quad (2.11)$$

Le terme,  $hvP_v$ , découlant des fluctuations aléatoires de Poisson autour du nombre moyen de photons (ou de Gauss lorsque le nombre de photons est suffisamment grand), est appelé bruit de photons ou bruit quantique. Quant au terme,  $hvP_v[\exp(hv/kT)-1]^{-1}$ , désigné bruit d'ondes (*wave noise*), il résulte du phénomène de battements d'ondes, c'est-à-dire des battements entre les champs de fréquences et de phases temporelles différentes. Ce dernier type de bruit apparaît surtout dans le domaine des basses fréquences, domaine où le terme exponentiel ( $hv/kT \ll 1$ ) domine. Or, compte tenu que les observations du NRAO sont faites à basse fréquence (8.4 GHz), seul le bruit d'ondes sera considéré pour la suite de cet ouvrage. D'ailleurs, contrairement au domaine optique (ou ultraviolet) où le bruit de photons est dominant aux hautes fréquences, c'est le bruit d'ondes qui domine aux radiofréquences (Léna 2008; Wilson, Rohfls & Huttemeister 2009).

Le bruit d'ondes peut se subdiviser en deux types de bruit. Le premier, qui provient de l'objet source observé, est un bruit où les variations sont plus lentes à cause des effets des milieux de propagation (voir section 2.2), par-dessus tout le plasma intergalactique, qui ont tendance à élargir temporellement les fluctuations d'intensité du bruit (et aussi celles du signal). Pour sa part, le second type de bruit d'ondes, couramment dénommé bruit thermique, est une conséquence de l'environnement local dans lequel les observations sont effectuées. D'ailleurs, contrairement au bruit issu des radiosources cosmiques proprement dites, les fluctuations de ce bruit local sont plutôt rapides: elles ne sont pas temporellement étalées par le milieu intergalactique.

## 2.1.4 Principales sources dominantes de bruit et température de bruit

En pratique, lorsque l'antenne parabolique pointe vers une région du ciel ne contenant pas d'objets cosmiques émissifs, on obtient tout de même un signal résiduel, une tension de bruit mesurée, d'origine thermique de puissance non nulle. Ceci étant dit, les sources dominantes de ce signal de bruit, d'où l'appellation bruit thermique, proviennent de l'émission du sol, de l'atmosphère et du récepteur hétérodyne. Le bruit émanant du sol s'apparente à celui causé par le rayonnement d'un corps noir, à une température d'autour de 300 K, lequel est réfléchi au large de la structure de l'antenne (émission détectée dans les lobes secondaires) dans l'alimentation du récepteur. Quant au bruit atmosphérique, il provient des effets de l'absorption des ondes radio par les molécules d'oxygène et d'eau dans la troposphère (élargissement des raies spectrales), ainsi que des effets de perturbations dans l'ionosphère. Par contre, tous ces effets demeurent faibles à la fréquence d'observation du NRAO (8.4 GHz). Finalement, le bruit du récepteur survient spécialement dans l'alimentation (diodes-mélangeuses incluses dans le système hétérodyne), dans les guides d'onde et dans les amplificateurs (Thompson et al. 2004; Léna 2008; Wilson, Rohfls & Huttemeister 2009; Burke & Graham-Smith 2010).

Les champs électriques résultant de ces différents types de bruit ne sont pas cohérents. Par conséquent, ceux-ci peuvent s'additionner quadratiquement pour donner une puissance de bruit. Cependant, en radioastronomie, il est plus pratique de travailler avec des températures qu'avec des puissances. Plus précisément, les radioastronomes associent à la puissance radiative de l'environnement (sol, atmosphère, récepteur) une puissance de bruit issue d'une charge (résistance) à la température  $T_{\text{bruit}}$ . La puissance résistive par unité de largeur de bande  $B$ , notée  $P_R$ , se trouve simplement avec,  $P_R = kT_{\text{bruit}}$  (Carilli 2010; Hayward 2012).

En général, on appelle température de bruit du système, notée  $T_{\text{sys}}$ , la température équivalente à la puissance totale de bruit générée par toutes les sources de bruit reliées à l'entrée d'un récepteur et connectées à la sortie d'un radiotélescope.

Cette température inclut évidemment les trois principales températures de bruit, notées respectivement  $T_{\text{sol}}$ ,  $T_{\text{atm}}$  et  $T_{\text{réc}}$ , mais aussi, la température d'antenne  $T_a$  (issue de l'objet source observé), la température du fond diffus cosmologique, notée  $T_{\text{cmb}}$ , et toutes autres températures ( $T_{\text{diverses}}$ ) de bruit (voir tableau 2.1). Mathématiquement, tout cela se traduit par (Thompson et al. 2004; Carilli 2010; Hayward 2012),

$$T_{\text{sys}} = \left( T_{\text{sol}} + T_{\text{atm}} + T_{\text{réc}} \right) + T_a + T_{\text{cmb}} + T_{\text{diverses}} . \quad (2.12)$$

La température système ne représente pas une température physique proprement dite, mais plutôt une puissance, ramenée en unités de température, à partir de laquelle il faut isoler le signal. Par exemple, la température d'un récepteur hétérodyne du NRAO est à l'entour de 15 K (leurs récepteurs sont refroidis dans des enceintes cryogéniques). En pratique, la température du récepteur est toujours beaucoup plus grande que la température minimale théorique du système. D'ailleurs, la limite théorique minimale de  $T_{\text{sys}}$  est fixée soit par la limite quantique,  $T_{\text{min}} = hv/k$  (= 395 mK à la fréquence d'observation du NRAO), ou bien soit par le rayonnement fossile à  $T_{\text{cmb}} = 2.728$  K (Wilson, Rohfls & Huttemeister 2009; Carilli 2010).

Pour terminer, il faut souligner qu'il existe également du bruit résultant de l'émission de raies spectrales (inclus dans  $T_{\text{diverses}}$ ). À la fréquence d'observation du NRAO, l'émission de raies provient surtout des transmissions radio et radar de systèmes terrestres proches de la station d'observation, ou de systèmes basés dans l'espace, ou encore des installations (ex., calibrations) internes des observatoires radio. Par exemple, le système de calibration de la phase instrumentale (voir section 5.2.5) et les nouveaux systèmes satellites radio (voir section 5.2.7) peuvent causer des signaux interférentiels parasites (Thompson et al. 2004).

**Tableau 2.1.** Valeurs typiques de température de bruit dans les observations radio du NRAO.

$T_{\text{sol}}$ (K)	$T_{\text{atm}}$ (K)	$T_{\text{réc}}$ (K)	$T_a$ (K)	$T_{\text{cmb}}$ (K)	$T_{\text{diverses}}$ (K)
5	5	15	$\leq 1$	3	$\geq 1$

### 2.1.5 Sensibilité du système VLBI du NRAO

La notion de sensibilité caractérise le signal minimum détectable par le VLBI, à deux antennes, c'est-à-dire le bruit théorique minimal (rapport signal à bruit valant un). Elle est directement rattachée à la sensibilité propre du système de réception hétérodyne et à la sensibilité de la surface collectrice effective de l'antenne parabolique. Cette surface, notée  $A_e$ , est égale la surface géométrique  $A$  multipliée par le facteur d'efficacité d'ouverture de l'antenne  $\eta_A$  qui varie typiquement de 0.3 à 0.8 (Thompson et al. 2004).

Il est très pratique d'indiquer la sensibilité combinée (récepteur et antenne) par la notion de système équivalent de densité de flux (SEFD). Le SEFD se définit comme étant la densité de flux d'une source radio ponctuelle qui porterait la puissance du bruit du récepteur au double de celle du bruit du système en l'absence de radiosource. Partant de ce fait, en égalisant l'équation (2.1) à l'équation (2.6) avec,  $P = P'' A$ , puis en substituant respectivement  $T$  par  $T_{\text{sys}}$  et  $A$  par  $A_e$ , il vient que, pour une seule polarisation (Thompson et al. 2004; Hayward 2012; Witz et Momjian 2016),

$$\text{SEFD} = \frac{2kT_{\text{sys}}}{A_e}. \quad (2.13)$$

Pour nos observations, le NRAO a utilisé des antennes paraboliques de diamètre,  $D = 25$  m. Avec un tel diamètre, la relation précédente peut alors se récrire sous la forme plus pratique suivante (Hayward 2012; Witz et Momjian 2016),

$$\text{SEFD}(\text{Jy}) = \frac{5.62T_{\text{sys}}(\text{K})}{\eta_A}. \quad (2.14)$$

La nature aléatoire du signal et du bruit permet de considérer les mesures séparées par des laps de temps de  $1/(2B)$  comme étant indépendantes. Autrement dit, une mesure où le signal est moyenné pendant un temps d'observation  $\tau$  contient  $2B\tau$  échantillons indépendants. Ceci étant acquis, la sensibilité (bruit minimum) théorique d'un système VLBI à deux antennes en unités de flux, notée  $\Delta S_{\text{lim}}$ , est donnée par (Thompson et al. 2004; Carilli 2010; Witz et Momjian 2016),

$$\Delta S_{\text{lim}} = (10^{29}) \frac{kT_{\text{sys}}}{A_e \sqrt{B\tau}} \text{ (mJy)}, \quad (2.15)$$

où  $T_{\text{sys}}$  et  $B\tau$  sont respectivement associés au bruit du système et au nombre de mesures indépendantes (le facteur  $10^{29}$  dérivant de la conversion,  $\text{Wm}^{-2}\text{Hz}^{-1}$  en mJy).

Pour les 22 objets observés (voir tableau A.2) à 8.4 GHz,  $\eta_A = 0.56$  et  $\text{SEFD} \approx 300$  Jy (Hayward 2012; Witz et Momjian 2016). On obtient alors de l'équation (2.14), une température théorique du système (de bruit) d'environ 30 K. Pour cette valeur de  $T_{\text{sys}}$  et pour une surface collectrice effective,  $A_e = \eta_A(\pi D^2/4) = 275 \text{ m}^2$ , on obtient, pour dix minutes d'observation dans une largeur de bande  $B = 32$  MHz (voir tableau A.1), de l'équation (2.15) une sensibilité (flux) minimale de,

$$\Delta S_{\text{lim}} = (10^{29} \text{ mJy/Wm}^{-2}\text{Hz}^{-1}) \frac{(1.38 \times 10^{-23} \text{ WK}^{-1}\text{Hz}^{-1})(30\text{K})}{(275 \text{ m}^2)(\sqrt{(32 \times 10^6 \text{ s}^{-1})(600\text{s})})} = 1.09 \text{ mJy}. \quad (2.16)$$

Il est important de savoir que cette sensibilité minimale théorique d'environ un millijansky correspond au plus faible bruit possible. En réalité, il y a plusieurs facteurs d'efficacité qui contribueront à augmenter l'intensité du bruit. Parmi ces facteurs, on retrouve entre autres l'efficacité de quantification (voir section 4.5.5), du

système d'enregistrement, des bandes passantes (multicanaux) et des systèmes de calibration (ex., calibration des phases). Compte tenu que tous ces facteurs d'efficacité se multiplient entre eux, on comprend bien que le facteur d'efficacité (de pertes) global  $\eta$  augmentera l'intensité du bruit: la limite de sensibilité passera de  $\Delta S_{\text{lim}}$  à  $\Delta S_{\text{lim}}/\eta$  pour  $0 < \eta < 1$ . À ce propos, pour connaître la nature de la majorité des facteurs de pertes possibles pour le VLBI du NRAO, on réfère le lecteur aux sections 9.6 à 9.9 de Thompson et al. (2004) et à Witz & Momjian (2016).

## 2.2 Milieux de propagation des ondes radio

### 2.2.1 Dispersion

Un modèle simple pour décrire un milieu dissipant de l'énergie est celui du plasma tenu dans lequel les électrons libres et les ions sont uniformément distribués, si bien que la densité spatiale totale de charge est nulle. Dans ce modèle, la masse élevée des ions, comparativement à celle des électrons, permet de négliger leur contribution (faible vitesse induite) par rapport à celle des électrons libres, lesquels sont accélérés par un champ électrique (Wilson, Rohfls & Huttemeister 2009).

Lorsque les ondes radio se propagent à travers un milieu dissipatif, comme un plasma intergalactique, elles sont altérées par les effets de dispersion (et de diffusion) dus à ce milieu. Ces effets font en sorte que les différentes longueurs d'onde radio (fréquences) se propagent à des vitesses différentes. Le fait que chacune des radiofréquences est dispersée et prend un temps différent pour parvenir jusqu'à l'observateur génère un délai temporel et un élargissement du signal électromagnétique émis par la radiosource (Wilson, Rohfls & Huttemeister 2009; Clarke et al. 2014).

Généralement, la mesure de dispersion, notée DM, est définie comme le délai temporel,  $\Delta t_{\text{diss-vid}}$ , entre les temps d'arrivée d'une onde radio dans un milieu dissipatif et dans le vide, soit (Ioka 2003; Wayth et al. 2011; Deng & Zhang 2014),

$$\Delta t_{\text{diss-vidé}} ; \int \frac{dl}{c} \frac{v_p^2}{2v^2} = 4.15 \left( \frac{v}{\text{GHz}} \right)^2 \left( \frac{\text{DM}}{10^3 \text{pc/cm}^3} \right) \text{ s}, \quad (2.17)$$

où  $v_p (= e(n_e/\pi m_e)^{1/2} = 8.98 \times 10^3 n_e^{1/2} \text{ Hz})$  est la fréquence du plasma ténu. Ici, la mesure de dispersion est normalisée à la valeur de référence,  $10^3 \text{ pc/cm}^3$ , pour le milieu intergalactique (MIG) d'un objet radio situé à une distance cosmologique. Pour un observateur terrestre, la DM est donnée par (Deng & Zhang 2014),

$$\text{DM} = \int \frac{n_{e,z}}{1+z} dl, \quad (2.18)$$

où  $z$  et  $n_{e,z}$  sont respectivement le redshift du plasma (de la radiosource) et le nombre de densité d'électrons libres ( $\text{cm}^{-3}$ ) à ce redshift, tandis que le facteur  $1/(1+z)$  découle du passage du référentiel de l'objet source à celui de l'observateur (Deng & Zhang 2014).

### 2.2.2 Contributions à la mesure de dispersion totale

La mesure de dispersion totale  $\text{DM}_{\text{tot}}$  entre l'observateur et la source radio provient de deux contributions galactiques, soit celle de notre Galaxie  $\text{DM}_G$  et celle de la galaxie hôte  $\text{DM}_{\text{hôte}}$ , de même que d'une troisième contribution, la plus importante d'ailleurs, résultant de la présence du très long milieu intergalactique  $\text{DM}_{\text{MIG}}$ . Cela dit, en sommant ces trois contributions, il vient que (Ioka 2003; Deng & Zhang 2014),

$$\text{DM}_{\text{tot}} = \text{DM}_G + \text{DM}_{\text{hôte}} + \text{DM}_{\text{MIG}}. \quad (2.19)$$

L'étude des pulsars permet de bien connaître la valeur de la mesure de dispersion dans notre Galaxie.  $\text{DM}_G$  atteint la valeur maximale de  $\sim 1000 \text{ pc/cm}^3$  dans une direction parallèle au plan galactique (à une latitude galactique,  $|b| \sim 0^\circ$ ) et vaut

moins que  $100 \text{ pc/cm}^3$  à mesure que l'on s'éloigne du plan galactique, soit à  $|b| > 10^\circ$  (Taylor & Cordes 1993; Thornton et al. 2013; Deng & Zhang 2014). À l'inverse, la mesure de dispersion dans la galaxie hôte est mal connue. D'ailleurs, Ioka (2003) fait l'hypothèse que les galaxies hôtes ressemblent à la nôtre, de telle sorte que  $DM_{\text{hôte}} \cong DM_G$ .

Pour dériver l'équation donnant la contribution du milieu intergalactique  $DM_{\text{MIG}}$ , on utilise l'hypothèse simplificatrice d'un plasma constitué d'hydrogène pur complètement ionisé (Ioka 2003; Inoue 2004). Dans ce cas, le nombre de densité d'électrons libres, à un redshift  $z$ , est donné par,

$$n_{e,z} = n_{e,0} (1+z)^3, \quad (2.20)$$

avec  $n_{e,0} \cong 2.4 \times 10^{-7} \text{ cm}^{-3}$  ( $= 3H_0^2 \Omega_b / (8\pi G m_p)$ ) représentant le nombre de densité d'électrons libres au temps présent ( $z = 0$ ).

Pour le modèle d'Einstein- de Sitter,  $dl(z) = cH_0^{-1}(1+z)^{-5/2} dz$  (Schneider 1993). Dès lors, en insérant l'expression de  $dl$  et l'équation précédente dans l'équation (2.18), il vient que,

$$DM_{\text{MIG}} = \frac{3cH_0 \Omega_{b,0}}{8\pi G m_p} \int_0^z \frac{dz'}{\sqrt{1+z'}} \cong 10^3 \left( \text{pc/cm}^3 \right) \int_0^z \frac{dz'}{\sqrt{1+z'}} \quad (2.21a)$$

$$\cong 10^3 \left( \text{pc/cm}^3 \right) 2\sqrt{1+z} \Big|_{z=0}^{z \equiv z_S}, \quad (2.21b)$$

où  $H_0 = 71 \text{ km/s/Mpc}$ ,  $\Omega_{b,0} \approx 0.04$  est la fraction actuelle de masse baryonique dans l'univers et  $m_p$  est la masse d'un proton ( $c$  et  $G$  étant respectivement la vitesse de la lumière et la constante gravitationnelle).

### 2.2.3 Délai temporel dû à la DM

Pour le projet, ce n'est pas le délai temporel dû à la différence entre les temps de propagation d'une onde radio dans un plasma dissipatif et dans le vide (voir équation (2.17)) qui est utile, mais c'est plutôt le délai temporel causé par la  $DM_{\text{tot}}$  entre la fréquence minimale et la fréquence maximale d'un canal du NRAO (voir tableau A.1), et cela pour une longueur donnée du milieu dissipatif d'un objet donné.

Soit,  $v_{n,\text{min}}$  et  $v_{n,\text{max}}$ , respectivement la fréquence minimale et maximale d'un canal  $n$  donné. Dès lors, le délai temporel  $\Delta t$  entre ces deux fréquences est donné par (Wilson, Rohfls & Huttemeister 2009; Wayth et al. 2011; Deng & Zhang 2014),

$$\Delta t = \frac{e^2}{2\pi m_e c} DM_{\text{tot}} (v_{n,\text{min}}^{-2} - v_{n,\text{max}}^{-2}) \quad (2.22a)$$

$$= 4.148 \times 10^{12} \left( \frac{DM_{\text{tot}}}{10^3 \text{ pc/cm}^3} \right) \left[ \left( \frac{v_{n,\text{min}}}{\text{MHz}} \right)^{-2} - \left( \frac{v_{n,\text{max}}}{\text{MHz}} \right)^{-2} \right] \mu\text{s}, \quad (2.22b)$$

où le passage de l'équation (2.22a) à l'équation (2.22b) découle de l'évaluation numérique du terme constant  $e^2/(2\pi m_e c)$  et du réglage des unités afin d'obtenir un  $\Delta t$  en microsecondes.

Pour terminer sur les mesures de dispersion, il faut mentionner qu'étant donné que les objets observés sont des radiosources cosmologiques (voir tableau A.2), seule la contribution à la mesure de dispersion totale issue du MIG sera considérée dans les traitements numériques tels que les lissages pour les calculs des autocorrélations et pour les simulations (voir chapitres 5 et 6). Les contributions,  $DM_{\text{hôte}}$  et  $DM_G$ , sont négligées dans cette thèse considérant qu'en général  $DM_{\text{hôte}} \cong DM_G = 30 - 100 \text{ pc/cm}^3$  (Ioka 2003; Deng & Zhang 2014), tandis que  $DM_{\text{MIG}}$  varie de 4000 à 50000  $\text{pc/cm}^3$  pour les 22 sources radio traitées (voir tableau 6.1). Qui plus est, rappelons que la valeur de  $DM_{\text{hôte}}$  est encore mal connue (Ioka 2003).

## 3 Notions sur les lentilles gravitationnelles

Le sujet traitant des effets de lentilles gravitationnelles est beaucoup trop vaste pour être discuté en profondeur dans ce chapitre. Pour cette raison, le chapitre se limite à décrire uniquement les aspects de la théorie des lentilles gravitationnelles qui sont les plus pertinents pour le projet. N'empêche que le lecteur intéressé à approfondir le sujet doit savoir qu'il existe d'excellentes références, en particulier, celles de Schneider, Ehlers & Falco (1992), Blandford & Nayaran (1992), Refsdal & Surdej (1994), Petters, Levine & Wambsganss (2001), Courbin, Saha & Schechter (2002) et Schneider, Kochanek & Wambsganss (2006).

### 3.1 Théorie de base

#### 3.1.1 Régimes des lentilles gravitationnelles

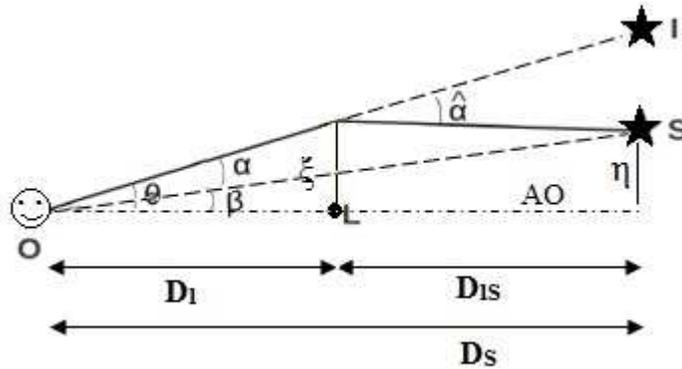
Selon le niveau de distorsions des images formées, les effets de lentilles gravitationnelles peuvent être divisés en deux principaux régimes, à savoir, les régimes de lentilles fortes (*strong lensing*) et de lentilles faibles (*weak lensing*). Les effets de lentilles fortes font référence aux cas où des images multiples sont produites, tandis que ceux de lentilles faibles font plutôt référence aux faibles amplifications et aux petites distorsions d'images individuelles (Narayan & Bartelmann 1997; Schneider, Kochanek & Wambsganss Part 1 2006; Zackrisson & Riehm 2010).

D'après les valeurs typiques des séparations angulaires d'images, on retrouve notamment les effets de macrolentilles, qui conduisent à des séparations de l'ordre des secondes d'arc (macroimages) et les effets de minilentilles, lesquels peuvent se subdiviser en des effets de millilentilles, conduisant en des séparations de l'ordre des millisecondes d'arc et en des effets de microlentilles dont les séparations d'images (microimages) ne sont que de l'ordre des microsecondes d'arc.

### 3.1.2 Principes de base des lentilles ponctuelles

La physique de base d'un système de lentilles gravitationnelles consiste essentiellement à la déflexion de la lumière (et des ondes électromagnétiques en général) dans un champ gravitationnel tel que prédit par la théorie de la relativité générale d'Einstein (Einstein 1911, 1916).

Pour les besoins de cette thèse, il est suffisant d'étudier le cas simple d'un système constitué d'un objet lentille d'aspect ponctuel qui est décrit par une géométrie euclidienne plate. La figure 3.1 illustre le principe d'une source ponctuelle (ex., quasar) subissant l'effet d'un déflecteur gravitationnel ponctuel.



**Figure 3.1.** Schéma d'un système composé d'une seule lentille gravitationnelle ponctuelle. Les lettres majuscules O, L, S et I représentent respectivement l'observateur, la lentille, la source ponctuelle (position réelle) et l'image de la source (position apparente). Voir le texte pour les autres détails. Par souci de clarté, tous les angles sont fortement exagérés.

Dans le régime des petites déviations et des faibles champs gravitationnels, un rayon lumineux issu d'une source située au point S, passant à une distance minimale  $\xi$  d'une lentille L de masse M, est dévié d'un angle  $\hat{\alpha}$ , dit angle de déflexion d'Einstein, qui est donné par,

$$\hat{\alpha}(\xi) = \frac{4GM}{c^2\xi} = \frac{2R_g}{\xi} = 1, \quad (3.1)$$

où le rayon de Schwarzschild,  $R_g \equiv 2GM/c^2$ , désigne le rayon gravitationnel de la masse ponctuelle  $M$  ( $c$  et  $G$  étant respectivement la vitesse de la lumière et la constante de gravitation universelle).

Le rayon lumineux perturbé par le faible champ gravitationnel de l'objet défecteur se propage jusqu'à l'observateur au point  $O$ , lequel voit la lumière issue d'une image  $I$  (position apparente de la source) située à un angle  $\theta$  de l'axe optique  $AO$  (arbitrairement choisi), au lieu de la voir provenir de la vraie position  $S$  de la source située à un angle  $\beta$  de l' $AO$ .

En désignant respectivement par  $D_l$ ,  $D_s$  et  $D_{ls}$  les distances de diamètre angulaire entre l'observateur et la lentille, entre l'observateur et la source ainsi qu'entre la lentille et la source, on obtient facilement la relation suivante (voir figure 3.1),

$$\theta D_s = \beta D_s + \hat{\alpha} D_{ls}. \quad (3.2)$$

En introduisant l'angle de déflexion réduit valant,

$$\alpha = \frac{D_{ls}}{D_s} \hat{\alpha}, \quad (3.3)$$

on peut alors reformuler l'équation (3.2) comme suit,

$$\beta = \theta - \alpha(\theta). \quad (3.4)$$

L'équation (3.4), reliant la position (angulaire) de la source à celle de l'image, porte le nom d'équation de la lentille (mince), et représente l'équation fondamentale de la théorie des lentilles gravitationnelles.

Pour un déflecteur ponctuel, en insérant les équations (3.1) et (3.3) dans l'équation de la lentille (3.4) et en utilisant la relation linéaire,  $\xi = D_l\theta$ , il vient que,

$$\beta(\theta) = \theta - \frac{D_{is}}{D_l D_s} \frac{2R_g}{\theta} = \theta - \frac{\theta_E^2}{\theta}, \quad (3.5)$$

où le terme,

$$\theta_E = \sqrt{2R_g \frac{D_{is}}{D_l D_s}}, \quad (3.6)$$

appelé rayon angulaire d'Einstein, donne l'échelle angulaire d'un système de lentilles gravitationnelles. Après quoi, en résolvant l'équation quadratique (3.5), on tire que les deux positions angulaires d'images, notées  $\theta_+$  et  $\theta_-$ , valent,

$$\theta_{\pm} = \frac{1}{2} \left( \beta \pm \sqrt{\beta^2 + 4\theta_E^2} \right), \quad (3.7)$$

et que la séparation angulaire  $\Delta\theta$  des deux images formées est donnée par,

$$\Delta\theta = \sqrt{\beta^2 + 4\theta_E^2}. \quad (3.8)$$

Les deux images du même objet source apparaissent de part et d'autre du déflecteur ponctuel, une image étant à l'intérieur du rayon d'Einstein et l'autre étant à l'extérieur (Blandford & Narayan 1992; Narayan & Bartelmann 1997).

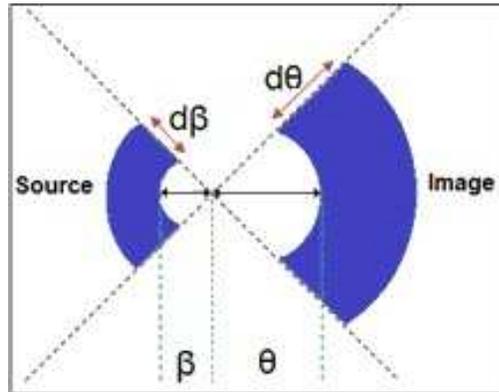
Dans le cas d'un alignement parfait entre la position réelle de la source et celle de la lentille ( $\beta = 0$ ), les deux images se fusionnent à une distance angulaire  $\theta_E$  de la lentille. L'observateur entrevoit alors un mirage gravitationnel prenant la forme d'un anneau lumineux, dit anneau d'Einstein, qui entoure l'image de la source. D'un autre côté, quand la source s'éloigne de l'objet déflecteur (augmentation de  $\beta$  dans la figure

3.1), une des images s'approche du déflecteur et devient très faible, tandis que l'autre image s'approche de plus en plus de la source, au point où son intensité devient pour ainsi dire égale à celle de la source (Narayan & Bartelmann 1997; Schneider, Kochanek & Wambsganss Part 1 2006).

Le champ de gravitation d'une lentille peut non seulement modifier la position angulaire apparente d'une source, mais aussi affecter l'intensité (flux) observée. En fait, bien que l'effet d'une lentille gravitationnelle ne change pas la brillance de surface d'une source, il en modifie l'angle solide apparent. L'intensité totale reçue de l'image d'une source est alors changée. Sachant cela, le rapport entre l'angle solide de l'image et celui de la source est donné par la relation suivante (Refsdal & Surdej 1994; Paczynski 1996; Narayan & Bartelmann 1997),

$$\mu = \frac{\text{surface de l'image}}{\text{surface de la source}} = \frac{\theta d\theta}{\beta d\beta}, \quad (3.9)$$

où le paramètre  $\mu$  représente l'amplification de l'image (facteur d'amplification) due à la présence de la lentille ponctuelle (voir figure 3.2).



**Figure 3.2.** Image d'une source avec lentille gravitationnelle à symétrie axiale au centre.

L'équation (3.5) permet de connaître la fonction  $\beta = \beta(\theta)$ , soit, en inversant l'ordre des paramètres,  $\theta = \theta(\beta)$ . De là, il suffit de dériver  $\theta$  par rapport à  $\beta$  pour trouver les deux amplifications, notées  $\mu_{\pm}$ , associées aux deux images, à savoir,

$$\mu_{\pm} = \left[ 1 - \left( \frac{\theta_E}{\theta_{\pm}} \right)^4 \right]^{-1} = \frac{u^2 + 2}{2u\sqrt{u^2 + 4}} \pm \frac{1}{2}, \quad (3.10)$$

avec  $u = \beta/\theta_E$  représentant la séparation angulaire (adimensionnelle) entre l'objet source et l'objet lentille.

Pour calculer l'amplification totale d'une source à images multiples, il suffit de sommer toutes les valeurs absolues des amplifications individuelles des images créées. Une amplification négative signifie tout simplement que la parité de l'image est inversée par rapport à celle de la source (Narayan & Bartelmann 1997).

Pour clore cette section, il est important de souligner qu'il va de soi que les systèmes réels de défecteurs gravitationnels sont plus élaborés que le système simple composé d'un seul défecteur ponctuel (décrit par une géométrie euclidienne plate) étudié ici. En vérité, les systèmes réels sont plutôt composés de défecteurs plus ou moins étendus et sont décrits par des métriques cosmologiques. Malgré cela, il est possible de simplifier ces systèmes réels en utilisant l'approximation de la lentille mince (voir équation (3.4)) et des distances de diamètre angulaire (voir figure 3.1). En effet, dans les systèmes réels, la taille de la lentille est beaucoup plus petite que les distances cosmologiques impliquées. Sachant cela, on peut à juste titre supposer que la déviation se produit strictement dans un mince plan local à la lentille, dit plan de la lentille, quasi perpendiculaire à la ligne de visée: c'est ce qu'on appelle communément l'approximation de la lentille mince. Avec cette approximation (et avec l'utilisation des distances de diamètre angulaire), il est encore possible d'utiliser l'équation simple de la lentille, pour la résolution de systèmes réels sophistiqués (ex., amplifications et délais temporels), en substituant les angles scalaires de l'équation

(3.4) par des angles vectoriels bidimensionnels, à savoir,  $\theta$  par  $\overset{\cdot}{\theta}$ ,  $\alpha$  par  $\overset{\cdot}{\alpha}$  et  $\beta$  par  $\overset{\cdot}{\beta}$  (Blandford & Nayaran 1992; Schneider, Ehlers & Falco 1992; Petters, Levine & Wambsganss 2001).

### 3.1.3 Estimation de l'ordre de grandeur des délais temporels

Refsdal (1964b, 1966), Press & Gunn (1973) et Spillar (1993) ont formulé des équations analytiques permettant d'estimer le délai temporel produit par un défecteur ponctuel situé à proximité de la ligne de visée reliant l'observateur à une source également ponctuelle. Tel qu'attendu, les relations trouvées par ces auteurs dépendent de la masse de la lentille, des distances impliquées ( $D_l$ ,  $D_{ls}$  et  $D_s$ ) et des paramètres cosmologiques (ex.,  $H_0$ ).

Afin d'estimer l'ordre de grandeur des délais temporel, on part du principe que la séparation angulaire est déjà connue ou, à vrai dire, estimée à partir de la séparation angulaire la plus probable. À ce propos, l'équation donnant cette séparation angulaire est donnée à l'annexe F (équation (5.1) de Press & Gunn (1973)). Cela dit, la démarche fait appel à une équation simple permettant d'approximer la séparation angulaire entre deux images d'intensité comparable produites par une source lointaine de redshift  $z_s \simeq 2$  subissant l'effet d'une lentille ponctuelle. Pour une telle source, la séparation angulaire la plus probable  $\Delta\theta$  est donnée par Press & Gunn (1973),

$$\Delta\theta ; 4(M_{PG}H_{0,PG})^{1/2} \quad (3.11a)$$

$$; 4\left(\frac{GMH_0}{c^3}\right)^{1/2} \quad (3.11b)$$

$$; 1.36 \times 10^{-11} \left(\frac{M}{M_e}\right)^{1/2} \text{ rad} = 2.8 \times 10^{-6} \left(\frac{M}{M_e}\right)^{1/2} \text{ arcsec} . \quad (3.11c)$$

Ici, le passage de l'équation (3.11a) à l'équation (3.11b) se fait en substituant respectivement  $M_{PG}$ , caractérisant la longueur associée à la masse  $M$  de la lentille en unités géométriques, par  $GM/c^2$  et  $H_{0,PG}$  par  $H_0/c$  (la notation de Press & Gunn (1973) incluant la vitesse de la lumière  $c$  dans l'expression de la constante de Hubble, soit  $cH_{0,PG}$ ). Notez que la légère différence (facteur de 2.8 au lieu de 2.6) entre l'équation (3.11c) et l'équation (F.1) résulte simplement de l'utilisation d'une valeur plus récente de la constante de Hubble, soit  $H_0 = 71$  km/s/Mpc au lieu de la valeur  $cH_{0,PG} = 60$  km/s/Mpc utilisée dans Press & Gunn (1973).

Pour une géométrie euclidienne plate, le délai temporel, noté  $\tau$ , entre deux images de luminosité comparable est donné par,

$$\tau ; \frac{1}{2c} \Delta\theta^2 D_s \frac{(D_s - D_1)}{D_1}. \quad (3.12)$$

Notez que l'équation précédente est tirée de l'équation (5.2) de Press & Gunn (1973) (voir équation (F.2)) où on a substitué les variables (paramètres affines)  $L$  et  $\lambda$  de la notation de Press & Gunn (1973) par  $D_s$  et  $D_1$  de notre notation.

Pour une lentille typique située à un tiers de la distance séparant la source de l'observateur, soit  $D_1 = D_s/3$ , le délai temporel se réduit à,

$$\tau ; \frac{1}{2c} \Delta\theta^2 D_s \frac{(D_s - D_s/3)}{(D_s/3)} ; \frac{1}{c} \Delta\theta^2 D_s. \quad (3.13)$$

Pour un objet source situé en moyenne à mi-horizon, soit  $D_s \sim R_0/2 = c/(2H_0) \sim 1/4$  Gpc, où  $R_0$  est le rayon de Hubble, on obtient des équations (3.11c) et (3.13) un délai de,

$$\tau \approx 40 \left( \frac{M}{M_c} \right) \mu\text{sec}. \quad (3.14)$$

En terminant, signalons que l'on verra à la section (6.2) qu'à partir de cette dernière relation, on pourra aisément calculer l'intervalle de masse applicable pour les minilentilles.

## **3.2 Lentilles gravitationnelles fortes et faibles**

### **3.2.1 Effets de lentilles fortes**

Les effets de lentilles gravitationnelles fortes peuvent survenir lorsque la densité de masse surfacique, le long de la ligne de visée (LDV), dépasse une certaine valeur critique (cas sur-critique). Ces effets produisent alors de grandes amplifications et des images multiples comme des arcs gravitationnels, des arcs géants ou même des anneaux complets dans le plan de la lentille. Concrètement, les effets de lentilles fortes sont des événements rares, particulièrement lorsque les sources sont des quasars plutôt que des galaxies (Marshall, Blanford & Sako 2005), où la déflexion produite par des objets massifs en avant-plan, comme des galaxies ou des amas de galaxies, est suffisamment grande pour produire des images multiples résolues (macroimages) d'une source lointaine. En pratique, le phénomène de lentilles fortes arrive quand l'objet source, l'objet lentille et l'observateur sont suffisamment alignés ( $\beta \cong 0$ ), ce qui correspond alors au cas sur-critique de densité de surface (Blandford & Nayaran 1992; Courbin, Saha & Schechter 2002; Narayan & Bartelmann 1997; Treu 2010; Zackrisson & Riehm 2010).

Les lentilles gravitationnelles fortes sont très utiles pour mesurer et comprendre l'univers. Avant tout, les propriétés observables, comme les séparations d'images, les amplifications et les délais temporels, dépendent du potentiel gravitationnel du défecteur (et de ses dérivées). Puis, ces observables dépendent également de la géométrie globale de l'univers par l'intermédiaire des distances de diamètre angulaire impliquées. Enfin, la source éloignée, vue par l'observateur, apparaît souvent amplifiée, parfois même de plus d'un ordre de grandeur. Ceci étant dit, les effets de lentilles fortes constituent des outils puissants pour entre autres 1)

comprendre la distribution spatiale de la masse aux échelles des kiloparsecs et des sous-kiloparsecs, échelles où la matière baryonique et la matière sombre interagissent afin de former les galaxies, 2) déterminer la géométrie globale, le contenu et la cinématique de l'univers, puis 3) étudier les galaxies lointaines, les trous noirs et les galaxies actives comme les quasars; tous ces objets sont trop petits ou bien trop faibles pour être résolus ou détectés avec les instruments actuels (Treu 2010).

### 3.2.2 Effets de lentilles faibles

À l'inverse des lentilles fortes, les effets de lentilles faibles se produisent dans les régions où la densité de masse surfacique est inférieure à une valeur critique (cas sous-critique), ce qui donne lieu à de petites amplifications et à de légères distorsions d'images. En général, les effets de lentilles faibles sont très courants dans l'univers, mais peu visibles, de sorte qu'ils peuvent uniquement être détectés statistiquement en étudiant un grand nombre de sources (Refregier 2003; Schneider, Kochanek & Wambsganss 2006 Part 3; Zackrisson & Riehm 2010).

Il y a plusieurs cosmologistes qui s'intéressent aux effets de lentilles faibles pour étudier la structure générale à grande échelle de l'univers. Dans ces circonstances, leur intention n'est pas vraiment de trouver la distribution de la matière sombre, mais plutôt d'obtenir des mesures statistiques (corrélations) afin de faire des comparaisons avec la théorie prédite. En réalité, ils visent à associer les observations de distorsions d'images des galaxies ou des quasars aux spectres de puissance de masse sous-jacents (Dodelson 2003).

### 3.3 Effets de microlentilles cosmologiques sur les quasars

D'un point de vue général, l'effet d'une microlentille survient lorsqu'un objet peu massif (ex., étoile), suite à son mouvement propre, passe devant une source de fond et amplifie temporairement son intensité.

Les effets de microlentilles (ou de minilentilles) cosmologiques sur les NAGs, tels que les quasars, traitent des effets d'objets compacts (objets dont la taille est inférieure à leur rayon d'Einstein) situés à proximité de la LDV. Notez que le terme anglais *cosmological microlensing* est attribuable au fait que le rayon angulaire d'Einstein est à l'échelle de la microseconde d'arc.

Les effets de microlentilles proviennent de leur granularité. Les microlentilles peuvent être des objets composés de matière baryonique comme des planètes, des étoiles ordinaires, des naines brunes, des trous noirs stellaires, ou encore des objets sombres non-baryoniques tels que des sous-halos ou micro-halos dont le potentiel gravitationnel est trop faible pour renfermer toutes composantes baryoniques, des trous noirs primordiaux dont la fraction des baryons n'est pas définie ou tout autre objet dont la nature n'est toujours pas claire. Ces effets peuvent survenir autant dans les quasars à images multiples qu'à image simple. Pour les quasars à macroimages multiples, les microlentilles sont possiblement situées dans le halo sombre de la galaxie déflectrice principale interposée entre le quasar et l'observateur. Quant aux quasars à une seule macroimage, les microlentilles peuvent être situées dans des amas de galaxies ou bien uniformément distribuées selon l'approximation de Press & Gunn (1973). Dans ce dernier cas, on peut imaginer que les microlentilles «flottent» librement et remplissent l'espace intergalactique (Wambsganss 2000).

La séparation moyenne (projetée) des objets de masse stellaire situés dans les galaxies lointaines est faible comparativement aux rayons d'Einstein stellaires typiques. Ceci étant acquis, tous les quasars peuvent subir des effets de microlentilles. Cependant, la résolution limitée des télescopes contemporains ne permet pas de

détecter les événements de microlentilles extragalactiques à partir de variations astrométriques des positions d'images de quasars (variations infimes de quelques microsecondes d'arc). N'empêche que pour les quasars compacts (de taille inférieure aux rayons d'Einstein stellaires), il est possible de les détecter en observant les variations de flux dans les courbes de lumière (voir section 3.4). En effet, les mouvements relatifs entre le quasar, les microlentilles et l'observateur modifient temporellement l'amplification totale, à cause des variations temporelles des amplifications individuelles des microimages, ce qui mène alors à des variations non corrélées de luminosité dans les différentes images d'un quasar (Paczynski 1986a, 1986b; Blandford & Nayaran 1992; Paczynski 1996).

Les variations temporelles de la luminosité des quasars peuvent tout simplement provenir de fluctuations intrinsèques de leur luminosité, mais elles peuvent aussi provenir de variabilités temporelles extrinsèques induites par des événements de microlentilles (Narayan & Bartelmann 1997; Zackrisson & Bergvall 2003; Schneider, Kochanek & Wambsganss 2006 Part 4). Pour les quasars à une seule image, ces événements ne sont pas aisément différenciables des variabilités intrinsèques, au point où il faut recourir aux analyses statistiques pour obtenir des informations à propos des événements de microlentilles sur ces quasars. Quant aux quasars à images multiples, il est possible de différencier les fluctuations extrinsèques, découlant de la présence de microlentilles dans la macrolentille, des fluctuations intrinsèques. Effectivement, pour ces dernières, les variations sont corrélées, c'est-à-dire qu'elles apparaissent dans toutes les images produites, avec un certain délai temporel dû à la différence de parcours des rayons lumineux, alors que les variations induites par les événements de microlentilles sont non corrélées entre les différentes images. D'ailleurs, l'intensité de ces fluctuations de luminosité non corrélées peut dépasser une magnitude avec des échelles de temps s'étendant sur des semaines, des mois ou même des années. Par exemple, pour le quasar à structure quadruple, Q2237+0305, communément connu sous le nom de Croix d'Einstein, une seule de ces images montrait des variations temporelles, tandis que les trois autres demeuraient pratiquement constantes (Irwin et al. 1989).

## 3.4 Courbes de lumière

### 3.4.1 Courbes de lumière standards

Environ 90 % des événements de microlentilles s'approximent bien par le modèle standard des courbes de lumière. Sachant cela, il s'avère pertinent d'expliquer les principes de base de ce modèle simple, même si les courbes de lumière de quasars subissant des effets de microlentilles sont généralement non standards (voir section 3.4.2). Essentiellement, dans ce modèle, on suppose qu'il y a une seule source, une seule microlentille (ces deux objets sont supposés ponctuels) et que tous les mouvements sont linéaires. Avec ces hypothèses simplificatrices, la courbe de lumière peut facilement être décrite par seulement quelques paramètres, pour une trajectoire donnée de la source, lesquels sont néanmoins contraints à la dégénérescence des effets de microlentilles (Shude 2008).

Dans le cas d'un objet source à petit redshift ( $z_S \sim 0$ ), le temps requis pour se déplacer d'un rayon angulaire d'Einstein  $\theta_E$ , dit temps caractéristique et noté  $t_E$ , est donné, en termes de  $D_l$  et de la vitesse transversale relative  $v_t$  entre la source et la microlentille, par,

$$t_E = \frac{D_l \theta_E}{v_t}. \quad (3.15)$$

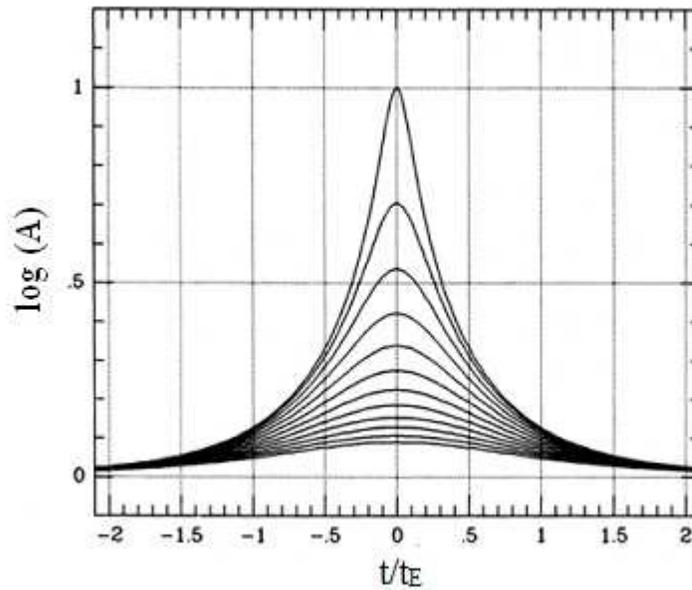
Soit  $p$ , le paramètre d'impact adimensionnel, normalisé en unités de  $\theta_E$ , lequel représente la plus petite distance angulaire entre la source à deux images et la microlentille, et soit  $t_{\max}$ , le temps requis pour produire une amplification maximale. Dès lors, pour déterminer quantitativement la courbe de lumière standard, il suffit d'additionner les valeurs absolues des deux amplifications (voir équation (3.10)) comme suit,

$$\mu(t) = |\mu_+(t)| + |\mu_-(t)| = \frac{u^2(t) + 2}{u(t)\sqrt{u^2(t) + 4}}, \quad (3.16)$$

puis de calculer la séparation angulaire  $u(t)$  avec (Paczynski 1986b, 1996),

$$u(t) = \sqrt{p^2 + \left(\frac{t - t_{\max}}{t_E}\right)^2}. \quad (3.17)$$

La figure 3.3 illustre les courbes de lumière produites par un événement de microlentille pour douze séparations différentes minimales entre l'objet source et l'objet déflecteur. Tel que prévu, l'amplification maximale est plus forte lorsque la distance minimale d'approche  $p$  est plus petite.



**Figure 3.3.** Variation logarithmique d'amplification  $A$ , due à un événement de microlentille gravitationnelle, en fonction du temps  $t$  normalisé à  $t_E$  ( $t_{\max} = 0$ ). Les douze courbes de lumière standards correspondent, de haut en bas, aux douze valeurs de  $p$ , notamment 0.1, 0.2, ..., 1.1 et 1.2. – Source (Paczynski 1986b).

De l'équation (3.17), on en déduit que pour modéliser une courbe de lumière observée, il y a trois paramètres essentiels, à savoir,  $p$ ,  $t_{\max}$  et  $t_E$ . Malheureusement, parmi ces trois paramètres, seul  $t_E$  est un paramètre physique. Il dépend de la masse

de la microlentille ( $\theta_E$  étant fonction de la masse), des distances impliquées (voir figure 3.1) et de la vitesse transversale relative  $v_t$  (voir équation (3.15)). Il en résulte que pour qu'une courbe de lumière observée soit correctement approximée par le modèle standard, il n'y a généralement pas une seule masse et une seule distance de la microlentille qui satisfont la courbe de lumière. En réalité, vu qu'il y a plutôt plusieurs combinaisons possibles de masse et de distances pouvant satisfaire une courbe de lumière donnée, on dira qu'il y a une dégénérescence des effets de microlentilles. Heureusement, si on connaît le spectre de masse des microlentilles, et certains modèles cinématiques de notre Galaxie, il devient alors possible de déduire la masse des microlentilles par une analyse statistique (Shude 2008).

### 3.4.2 Courbes de lumière non standards

Le 10 % restant de courbes de lumière sont dites non standards ou exotiques, car elles violent une ou plusieurs des hypothèses simplificatrices (source ponctuelle, microlentille ponctuelle simple, linéarité des mouvements, etc.). Nonobstant, les effets de microlentilles non standards permettent de dériver des contraintes supplémentaires et de lever, du moins partiellement, la dégénérescence des effets de microlentilles. Ci-dessous, on retrouve une énumération concise de trois cas pratiques, résultant en des courbes de lumière non standards, que l'on peut rencontrer dans les événements de microlentilles extragalactiques (pour plus de détails, voir Dominik 1999; Dominik et al. 2008).

- 1) La présence de deux ou de maintes microlentilles génère des formes de courbes de lumières très variées.
- 2) Lorsque la taille angulaire du quasar est comparable au paramètre d'impact  $p$ , le quasar ne peut plus être supposé comme étant un objet ponctuel. Dans ce cas, l'effet de la taille finie du quasar change considérablement les formes des courbes de lumière associées.

- 3) L'événement de microlentille peut «se répéter», entre autres, quand la microlentille ou le quasar constitue un système binaire suffisamment large. Pour ce cas, l'événement de microlentille peut se manifester par deux maximums (pics) séparés dans la courbe de lumière, ce qui donne l'impression qu'il se produit une autre fois.

## 4 Autocorrélation d'intensité et signaux numérisés

### 4.1 Modèle du champ électrique d'un rayonnement synchrotron

Le rayonnement d'origine synchrotron issu d'une source radio présente des variations erratiques temporelles et fréquentielles. Malgré tout, on suppose un modèle où l'émission synchrotron radio suit un processus aléatoire gaussien modulé en amplitude. Dans ce modèle, les électrons émettent leur rayonnement indépendamment les uns des autres. La radiation synchrotron n'a donc pas de cohérence temporelle, mis à part une certaine cohérence partielle introduite par le filtre du récepteur hétérodyne (Klein & Furtak 1986; Spillar 1993; Loudon 2000; Richards 2000).

Pour un tel modèle, le champ électrique total, noté  $E(t)$ , de ce rayonnement synchrotron, mesuré par un radiotélescope idéal dont le récepteur aurait une largeur de bande infinie, peut mathématiquement être approximé par (Rickett 1975; Lovelace & Richards 2013),

$$E(t) = \sum_i E_i(t) = \sum_i \dot{E}_{0i}^{\circ}(t) \delta(t-t_i). \quad (4.1)$$

Ici, pour un électron donné  $i$ ,  $E_i(t)$  est sa contribution à  $E(t)$ ,  $\dot{E}_{0i}^{\circ}(t)$  est l'amplitude complexe de la composante du champ électrique dont la polarisation correspond à celle de l'antenne (ex., RCP ou LCP),  $\delta(t-t_i)$  est la fonction delta et  $t_i$  est le temps où l'antenne reçoit l'impulsion.

Le rayonnement synchrotron est détecté uniquement lorsque la direction de l'observateur est à l'intérieur du faisceau d'émission. La radiation synchrotron instantanée est, en général, elliptiquement polarisée, mais vu que l'angle de position de polarisation de l'ellipse est en rotation avec l'électron, la polarisation moyennée dans le temps est linéaire. Ceci est aussi vrai pour le rayonnement émis par un ensemble d'électrons mono-énergétiques se déplaçant sur des orbites parallèles (Wilson, Rohfls & Huttemeister 2009). Les récepteurs hétérodynes de

radiofréquences ne détectent qu'une polarisation de champ électrique. Comme ils ne sont sensibles qu'à une polarisation particulière, ils combinent donc la détection et la mesure de la polarisation (Léna 2008, Wilson, Rohfls & Huttemeister 2009). Les polarisateurs VLBA du NRAO génèrent deux polarisations circulaires inverses, droites et gauches (RCP et LCP), à partir de cette polarisation (moyenne) linéaire. Dans les huit canaux impairs RCP (voir tableau A.1), la polarisation linéaire est retardée de  $90^\circ$  (positive), tandis qu'elle est avancée de  $90^\circ$  (négative) dans les huit canaux pairs LCP (Hayward 2012). Ceci étant acquis, rappelons comment une polarisation rectiligne peut être décomposée en une somme de deux ondes circulairement polarisées, en rotation opposées, de même amplitude.

Compte tenu que la polarisation circulaire (droite ou gauche) est une polarisation où la norme du vecteur du champ électrique ne change pas (seulement son orientation change selon un mouvement de rotation), on traite donc le vecteur  $\vec{E}$  d'amplitude  $E_0$  comme deux scalaires d'amplitude  $E_0/2$ . Le facteur  $1/2$  correspondant à la sélection d'une polarisation (droite ou gauche). (Notez que dans l'équation (4.1) et dans notre analyse, on a absorbé le facteur  $1/2$ , de sorte que l'on considère plutôt une onde de polarisation linéaire d'amplitude double,  $2E_0$ , décomposée en deux ondes de polarisations circulaires d'amplitude identique  $E_0$ .) Par ailleurs, notre technique est basée sur l'autocorrélation de l'intensité  $I(t)$  (voir équation (4.5) et non pas sur l'autocorrélation du champ électrique  $\vec{E}(t)$ . Cela dit, vu que  $I(t)$  est proportionnel à  $|\vec{E}(t)|^2$ , on s'intéresse seulement à la norme (au carré) du champ électrique, autrement dit à l'amplitude  $E_0^2$ .

Dans l'analyse des données (voir chapitre 5), on tient compte des deux polarisations circulaires R et L en traitant indépendamment les huit canaux pairs (L) et les huit canaux impairs (R). À vrai dire, les 16 canaux sont considérés comme indépendants, car on analyse (auto-corrèle) un seul canal à la fois. Par exemple, pour l'intervalle 8450 – 8512 MHz correspondant aux canaux contigus 1 et 2 (voir tableau

A.1), on effectue deux autocorrélations d'intensité, l'une pour la partie RCP de l'onde (canal 1) et l'autre pour sa partie inverse LCP (canal 2).

En pratique, le système de réception d'un radiotélescope réel possède une bande passante de largeur  $B$  finie centrée autour d'une fréquence d'observation, notée  $\nu_0$ . Ceci implique que le filtre passe-bande de largeur  $B$  du récepteur fait en sorte que seules les fréquences  $\nu_0 \pm B/2$  ( $B \ll \nu_0$ ) du champ électrique (signal) peuvent le traverser. Cela dit, après avoir été filtré, le signal est ensuite mélangé avec un signal exponentiel,  $\exp(2\pi i \nu_{LO} t)$ , de référence provenant d'un oscillateur local (voir figure 2.3) dont la fréquence  $\nu_{LO}$  est voisine de la fréquence d'observation  $\nu_0$ . Dès lors, le champ électrique moyenné total  $\bar{E}(t)$  reçu par le radiotélescope est maintenant donné par (Cordes 1976; Cordes et al. 2004; Lovelace & Richards 2013),

$$\bar{E}(t) = \sum_i \hat{E}_{0i}^{\circ}(t) \Delta(t-t_i), \quad (4.2)$$

où la substitution de  $\delta(t-t_i)$  par  $\Delta(t-t_i)$  provient du passage d'un récepteur idéal ( $B \rightarrow \infty$ ) à un récepteur réel de largeur  $B$  finie. La largeur temporelle,  $\Delta(t-t_i) = \Delta\tau$ , est de l'ordre de la réciproque de la largeur de bande  $B$  du récepteur utilisé pour trouver  $\bar{E}(t)$ . Dans l'équation (4.2), la barre supérieure sur  $E(t)$  signifie que le champ électrique est moyenné sur un intervalle de temps  $\Delta\tau$ , dit période ou pas d'échantillonnage, supérieur à la période des oscillations de  $E(t)$ , lesquelles sont beaucoup trop rapides pour être observées directement. Par exemple, pour les observations du NRAO,  $\Delta\tau = 1/2B = 1/(2 \times 32 \text{ MHz}) = 15.625$  nanosecondes et  $\nu_0 = 8.4 \text{ GHz}$  (voir tableau A.1). Avec une telle radiofréquence, la période des oscillations de  $E(t)$  vaut,  $1/\nu_0 = 0.119 \text{ ns}$ , laquelle est de deux ordres de grandeur plus courte que la période d'échantillonnage de  $15.625 \text{ ns}$ .

## 4.2 Intensité et fluctuations d'intensité

On a déjà dit à la section précédente qu'un rayonnement synchrotron, modélisé par un processus aléatoire gaussien, implique que les différents électrons individuels ne sont pas reliés. Il n'y a donc pas d'interférence entre les champs électriques des électrons individuels. Conséquemment, chaque électron indépendant contribue séparément à l'intensité radio. Pour cette raison, on calcule l'intensité temporelle totale, notée  $\bar{I}(t)$ , tout simplement en additionnant les carrés des champs électriques individuels, soit mathématiquement (en ignorant les constantes multiplicatives sans conséquence),

$$\bar{I}(t) = \sum_i \bar{I}_i(t) = \sum_i |\bar{E}_i(t)|^2 = |\bar{E}(t)|^2. \quad (4.3)$$

L'équation précédente montre bien qu'on peut approximer l'intensité totale de l'émission synchrotron par la superposition des intensités résultant d'un très grand nombre d'électrons indépendants. Par ailleurs, afin de déterminer les fluctuations temporelles aléatoires de  $\bar{I}(t)$ , on suppose que l'intervalle de temps  $\Delta\tau$  pris pour évaluer chaque fluctuation individuelle est beaucoup plus court que le temps de cohérence  $\tau_c$  de la radiation (Klein & Furtak 1986; Mandel & Wolf 1995; Loudon 2000). De cette façon, il devient possible de calculer des fluctuations instantanées d'intensité avec un détecteur dont le temps de réponse  $\Delta\tau$  serait beaucoup moindre que  $\tau_c$ . Mais, pour le projet, le temps  $\Delta\tau$  minimum pour calculer les fluctuations d'intensité est évidemment limité au pas d'échantillonnage de 15.625 ns du NRAO.

Les fluctuations instantanées d'intensité au temps  $t$ , notées  $\Delta I(t)$ , sont données par (Klein & Furtak 1986; Loudon 2000),

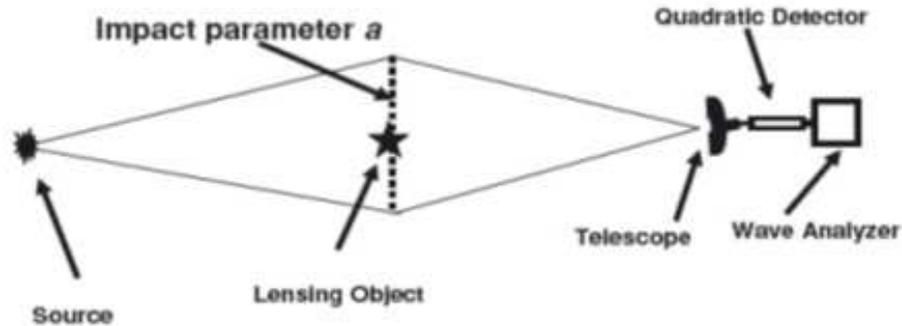
$$\Delta I(t) = I(t) - \langle I(t) \rangle, \quad (4.4)$$

avec les chevrons  $\langle \rangle$  indiquant que l'intensité est évaluée sur le temps total (ex., 600 s) d'intégration. Notez que pour simplifier la notation, on a laissé tomber la barre supérieure sur  $I(t)$ .

### **4.3 Modèle avec l'interféromètre de Young et modulations spectrales**

On a déjà mentionné au chapitre 3 que l'effet d'une lentille gravitationnelle sur une source lointaine pouvait produire des arcs, des anneaux ou des images multiples. Le cas le plus souvent observé en pratique est celui où deux images de la source sont créées. Ce cas particulier peut être modélisé par un interféromètre de Young, lequel est un modèle standard pour les lentilles gravitationnelles. Essentiellement, dans l'interféromètre de Young (ou de Michelson), deux faisceaux lumineux identiques, provenant de la source, sont recombinaés après avoir été divisés et avoir parcouru des longueurs de chemins optiques différentes (Borra 2008).

À l'aide du modèle géométrique classique simple pour les deux composantes (images) d'un défecteur gravitationnel, il est possible de modéliser le défecteur par un interféromètre de Young, qui introduit un délai temporel entre les deux faisceaux qui interfèrent. La figure 4.1 schématise un interféromètre pour la détection d'une lentille gravitationnelle. Le paramètre d'impact, noté  $a$ , analogue à la séparation entre les fentes d'un interféromètre de Young, est responsable du délai temporel et dépend de la masse de l'objet défecteur (Borra 2008).



**Figure 4.1.** Schéma simplifié d'un interféromètre à une lentille gravitationnelle. La lentille est modélisée par un interféromètre de Young comportant deux faisceaux de même intensité. Un télescope est utilisé pour observer l'image issue des deux faisceaux qui interfèrent. Un détecteur quadratique permet de mesurer l'intensité temporelle moyennée des faisceaux. Un analyseur d'ondes permet d'obtenir la distribution spectrale des fluctuations d'intensité. – Source (Borra 2008).

Les techniques interférométriques classiques permettent d'observer les franges d'intensité. Par contre, lorsque la différence de parcours optique (DPO) entre les deux faisceaux interférant est plus grande que la longueur de cohérence de l'onde électromagnétique, les franges ne sont plus visibles (Loudon 2000; Borra 2008). Nonobstant, les faisceaux d'interférence séparés par de très grandes DPOs produisent un faisceau recombinaison présentant une modulation spectrale (voir la figure 3 de Borra (2011)). En d'autres termes, la modulation spectrale de l'intensité se produit pour des DPOs beaucoup plus grandes que la longueur de cohérence des faisceaux qui interfèrent. En dépit du fait que cette assertion semble à priori violer des théories physiques fondées, il reste qu'elle est validée par des expériences publiées d'Alford & Gold (1958), de Cielo, Brochu & Delisle (1975), de Basano & Ottonello (2000) et de Borra (2011), et qu'elle repose sur des bases théoriques fermes (Givens 1961; Mandel 1962a, 1962b; Baczynski et Oster 1965; Delisle & Brochu 1972; Zardecki & Delisle 1974). Selon ces théories, la modulation spectrale est conservée pour des DPOs arbitrairement longues. À ce propos, les expériences de laboratoire, faites dans

le visible et avec des DPOs de plusieurs mètres, montrent la présence d'un signal (ex., figure 2 de Borra (2011)) pour des DPOs beaucoup plus grandes que la longueur de cohérence des faisceaux. Sachant cela, il n'y a aucune raison physique pour que les conditions de laboratoire ne soient pas applicables (extrapolables) aux systèmes astrophysiques dont les DPOs sont de plusieurs kilo-parsecs pour les effets de minilentilles cosmologiques. À ce propos, Borra (1997) et les références incluses montrent qu'on peut trouver des DPOs (délais temporels) de longueur arbitraire, d'où des lentilles gravitationnelles de masse arbitraire (voir équation (3.14)). Par contre, pour les ondes radio, la présence du milieu intergalactique aura pour effet d'étaler temporellement le signal et d'en diminuer l'amplitude.

L'article de Borra (1997) souligne que les fluctuations d'intensité causées par des superpositions d'ondes produisent une signature de lentilles qui peut être utilisée pour identifier des lentilles non résolues et mesurer des délais temporels. Bracewell (1986) décrit la base de la méthode qui permet de déterminer le délai temporel dans une lentille gravitationnelle. Givens (1961) et Mandel (1962a, 1962b) soutiennent que les expériences d'Alford & Gold (1958), faites pour des objets pulsés, s'appliquent également aux objets non pulsés. Dès lors, les expériences sont aussi réalisables avec de la lumière non modulée, pourvu que les deux faisceaux soient cohérents. L'effet AG est donc directement applicable aux sources astronomiques à lentilles gravitationnelles, lesquelles ne sont pas pulsées (continues). La discussion menant à l'équation (5) de Borra (1997), reproduite à l'annexe F, est empruntée de l'effet AG et de Givens (1961), permet de mesurer des délais temporels de lentilles gravitationnelles pour une vaste gamme de délais allant de seulement une dizaine de nanosecondes jusqu'à plus d'une année. Ajoutons enfin que la discussion menant à l'équation (5) de Borra (1997) est valide pour des délais temporels arbitrairement longs, cela indépendamment du temps de cohérence donné par la bande passante.

Pour les sources astronomiques, comme les quasars qui sont des sources d'émission de continuum, l'interféromètre divise les fluctuations d'intensité du faisceau d'entrée en deux fluctuations aléatoires identiques temporellement séparées

de  $\tau = \text{DPO}/c$ . Par le fait même, le faisceau recombinaé contient également des fluctuations aléatoires identiques: chaque fluctuation d'intensité au temps  $t$  est suivie de sa réplique au temps  $t+\tau$  (Basano & Ottonello 2000; Borra 2011, 2013, 2014).

#### 4.4 Analyse basée sur l'autocorrélation des fluctuations d'intensité

Borra (2008) présente une théorie innovatrice basée sur les effets de superpositions d'ondes. Pour faire son analyse théorique, il suit un traitement mathématique rigoureux, basé sur le principe de superposition d'ondes de la théorie électromagnétique classique de Maxwell, qui fait intervenir des notions d'optiques statistiques avancées (O'Neill 2003; Goodman 2015). En revanche, l'analyse beaucoup plus intuitive, basée sur l'autocorrélation de  $I(t)$ , ou à dire vrai de  $\Delta I(t)$ , permet de détecter plus simplement des effets de lentilles gravitationnelles, soit des signatures sans équivoques de délais temporels, que celle de Borra (2008), laquelle utilise, entre autres choses, la transformée de Fourier de la fonction de cohérence spatiale pour obtenir la distribution spectrale d'intensité (voir les équations (13) et (14) de Borra (2008), lesquelles sont reproduites à l'annexe F). Ceci étant acquis, la fonction d'autocorrélation des fluctuations d'intensité, notée  $R(\tau)$ , s'écrit mathématiquement comme suit,

$$R(\tau) = \int_{-\infty}^{+\infty} \Delta I(t+\tau)\Delta I(t) dt. \quad (4.5)$$

La raison pour laquelle on utilise  $\Delta I(t)$  au lieu de  $I(t)$  provient du fait que l'autocorrélation de  $I(t)$  donne une forme contenant une pente descendante très abrupte, tandis que l'autocorrélation de  $\Delta I(t)$  donne une forme horizontale de  $R(\tau)$  avec une valeur moyenne à peu près nulle, facilitant ainsi la détection des signaux (voir section 5.3).

Dans l'équation précédente, toute l'information pertinente, comme le délai temporel  $\tau$ , est comprise dans la dépendance en temps de l'intensité. Pour tout dire, la

fonction d'autocorrélation effectue une intégrale de  $\Delta I(t)$  multipliée par sa réplique,  $\Delta I(t+\tau)$ , décalée de  $\tau$ . On a déjà dit que c'est dans  $\Delta I(t)$  que se retrouvent les fluctuations aléatoires instantanées d'intensité. Ces fluctuations instantanées résultent des effets de superpositions d'ondes électromagnétiques variant sur des temps  $t$  beaucoup plus courts que le temps de cohérence  $\tau_c$ . Elles proviennent de la superposition d'ondes électromagnétiques ayant différentes fréquences et différentes phases (Borra 2013, 2014).

#### 4.4.1 Signal dans l'autocorrélation d'intensité dû à l'effet d'une lentille

Dans l'éventualité de l'effet d'un défecteur gravitationnel (modélisé par un interféromètre de Young), le signal, contenu dans les fluctuations d'intensité, serait composé de la somme de deux signaux identiques séparés d'un délai  $\tau$  causé par le défecteur. Par conséquent, le signal observé (un pic élargi), dans le profil d'autocorrélation d'intensité, serait décalé de l'origine du même intervalle de temps que le délai  $\tau$  séparant une fluctuation de sa réplique. Pour s'en convaincre, on recommande fortement au lecteur de consulter les expériences de Borra (2011), lesquelles valident formellement que la technique d'autocorrélation d'intensité peut détecter des effets de lentilles gravitationnelles.

Pour une source compacte, le signal d'autocorrélation d'intensité comprend un pic central situé à  $t = 0$  et deux pics symétriques situés à  $t = \pm DPO/c$ . À titre d'exemple, dans l'expérimentation de Borra (2011), la différence de parcours optique (délai temporel) causée par la présence d'un objet défecteur, modélisé par un interféromètre de Mach-Zehnder ( $DPO = 3.12$  m) dans la figure 1 de Borra (2011), génère deux pics symétriques dans la figure 2 de Borra (2011).

Une source d'aspect étendu peut être représentée par une superposition de sources ponctuelles individuelles, chacune d'elles étant associée à une sous-région infinitésimale  $dA_i$  de la source étendue. Pour cette raison, le signal résultant de l'autocorrélation d'intensité d'une source étendue découle de la somme des signaux

des autocorrélations d'intensité des sources ponctuelles individuelles (Klein & Furtak 1986; Zhdanov 1999; Loudon 2000; Borra 2008). Cette affirmation est d'autant plus vraie que les corrélations-croisées (*cross-correlation*) des intensités des signaux provenant de plusieurs sources ponctuelles indépendantes et incohérentes sont toujours nulles: aucune contribution de la part des termes croisés. Cela revient à dire que toutes les contributions à l'intensité du signal provenant des différentes sous-régions infinitésimales  $dA_i$  de la source étendue sont auto-corrélées avec elles-mêmes. Ceci entraîne que l'autocorrélation globale comporte un pic central à  $t = 0$  (à l'instar de la source ponctuelle) et des pics symétriques à  $t_i = \pm DPO_i/c$  pour une sous-région  $dA_i$  de la source étendue. Pour un exemple de source étendue, revenons à l'expérimentation de Borra (2011) dans laquelle la source étendue est tout simplement modélisée par deux sources ponctuelles (voir la figure 1 de Borra (2011)). L'objet défecteur, modélisé par les deux interféromètres, produit deux pics symétriques aux  $DPO_{s1} = \pm 3.12$  m (premier interféromètre) pour la sous-région  $dA_1$ , tandis qu'il produit deux autres pics symétriques aux  $DPO_{s2} = \pm 6.35$  m (deuxième interféromètre) pour l'autre sous-région  $dA_2$ . À partir de cet exemple, on peut aisément en déduire que  $N$  sources ponctuelles indépendantes généreraient  $N$  pics distincts symétriques de chaque côté de l'axe des ordonnées dans le profil de l'autocorrélation globale.

Pour terminer sur les sources étendues, soulevons que Spillar (1993) et Zhdanov (1999) ont étudié l'effet de l'étendue de la source sur l'autocorrélation du champ électrique, tandis que Borra (2008, 2011, 2014) a étudié cet effet sur l'autocorrélation de l'intensité. Ils ont conclu qu'au fur et à mesure que la taille de la source augmente, l'amplitude du signal de l'autocorrélation globale, autant celle du champ électrique que celle de l'intensité, diminue et que le signal est étalé sur une grande gamme de délais temporels: le signal dans l'autocorrélation globale est composé d'une somme de pics de plus faibles amplitudes situés à différents délais temporels.

#### 4.4.2 Avantages de la technique d'autocorrélation d'intensité

La technique d'autocorrélation d'intensité développée dans cette thèse permet d'aisément détecter des variations temporelles d'intensité très rapides et de très faibles amplitudes, qui sont totalement perdues dans le bruit dans  $I(t)$ . En effet, comme l'autocorrélation peut en moyenne se faire sur une longue période (ex., 600 s), cela permet de détecter de très faibles signaux, ainsi que des variations d'intensité se produisant sur des temps qui ne sont que modérément plus longs que le temps d'échantillonnage. De plus, notre technique peut utiliser des données déjà obtenues avec des radiointerféromètres existants dans leur format standard (ex., Mark 5B). Le fait d'utiliser des données interférométriques standard ne requière pas d'instrumentation spécialisée, et comme l'échantillonnage est fait sur de très courts temps (ex., 15.625 ns), on peut détecter des variations de flux très rapides.

À présent, les délais temporels impliquent des mesures de variations temporelles de flux parmi les images résolues découlant d'effets de lentilles fortes (amplifications gravitationnelles). Ceci entraîne que l'on peut seulement mesurer de longs délais temporels résultant de sources de luminosité variable. En revanche, notre technique d'autocorrélation d'intensité s'applique à des sources de luminosité constante (ex., quasar), car elle fait appel aux fluctuations d'intensité générées par des effets de superpositions d'ondes qui se retrouvent également dans les objets de luminosité constante (Borra 2008, 2014).

Les méthodes actuelles de détection, telles que celles basées sur les séparations d'images ou sur les courbes de lumière sont limitées à mesurer des délais temporels sur de longues échelles de temps. Par exemple, les effets de microlentilles extragalactiques sur les quasars engendrent des courbes de lumières (voir section 3.4) variant sur des échelles de temps allant des semaines aux mois, voire aux années (voir section 3.3). En conséquence, il faut observer les variations de luminosité des quasars pendant très longtemps, soit, en pratique, faire maintes observations sur plusieurs jours, voire quelques mois. En contrepartie, notre technique permet de mesurer de très

courts délais temporels. En théorie, on pourrait mesurer des délais aussi courts que le pas d'échantillonnage de 15.625 nanosecondes des observations du NRAO. Cependant, d'aussi brefs délais ne sont pas atteignables dans la région radio (voir section 4.4.3). En plus, notre technique nécessite, pour une station donnée, une seule observation (la numérisation des données étant faite par logiciel) de courte durée, par exemple, dix minutes pour les observations du NRAO (voir tableau A.3).

La détection de la signature du délai temporel se fait facilement par simple inspection visuelle du profil d'autocorrélation et identifie, contrairement à la méthode standard de la variabilité des sources (courbes de lumière), sans ambiguïté un événement de lentilles gravitationnelles. De fait, dans le cas des sources (ex., quasars) à une seule image, il est souvent difficile avec la technique basée sur les courbes de lumière, de distinguer les variations de flux induites par des effets de minilentilles des variations intrinsèques de l'objet source (Canizares 1982; Schneider 1993; Narayan & Bartelmann 1997; Zackrisson & Bergvall 2003; Schneider, Kochanek & Wambsganss Part 4 2006). Par contre, notre technique d'autocorrélation permet de distinguer, sans ambiguïté, les variations intrinsèques, périodiques ou non périodiques, de celles induites par un événement de minilentille. En effet, pour les sources radio comprenant une variation périodique (ex., pulsars), le signal auto-corrélé détecté contiendrait plusieurs pics répétitifs caractérisant la périodicité. Les figures 5.3 et 5.6 illustrent bien ces pics pour un signal périodique fort et faible. Puis, pour les sources radio (compactes) à minilentilles gravitationnelles, l'effet d'une minilentille (ponctuelle) générerait deux fluctuations d'intensité identiques temporellement décalées (voir section 4.4.1). La figure 2 de Borra (2011) donne un exemple du signal auto-corrélé attendu dans le cas d'un événement de minilentille. Enfin, pour les radiosources comportant des variations non périodiques rapides, par exemple les objets à sursauts radio rapides (FRBs) comme le FRB121102, il n'y aurait pas une réplique retardée de la première fluctuation (pas de jumeaux ou de dédoublement du signal). La figure 5.7 donne un exemple de profil d'autocorrélation pour des signaux non périodiques.

### 4.4.3 Inconvénients de la technique d'autocorrélation d'intensité

À l'instar de l'interférométrie d'intensité (Hanbury Brown 1968), la technique d'autocorrélation d'intensité utilise la théorie de superposition d'ondes, qui est basée sur les équations de Maxwell. De ce fait, la technique requiert des sources brillantes afin que les fluctuations d'intensité, découlant des effets de superpositions d'ondes, puissent sortir du bruit. Le fait de travailler dans la région radio permet justement d'avoir un bon rapport signal à bruit (S/B). Le paragraphe suivant explique pourquoi.

Une discussion explicite du S/B dépend, entre autres, des instruments et des techniques utilisés, lesquels dépendent eux-mêmes du domaine du spectre électromagnétique dans lequel on travaille, et cela dépasse le cadre de cet ouvrage. Ceci étant acquis, on peut tout de même estimer le S/B en se basant sur le bruit de photons. D'après les articles de Forrester et al. (1955), Purcell (1956), Mandel (1962b) et Borra (1997), on obtient un rapport  $R \leq \delta = \alpha I / \Delta\nu$  entre le bruit de superposition d'ondes et le bruit de photons en termes de densités spectrales des fluctuations. La valeur de R prédit que pour une intensité I donnée, le paramètre de dégénérescence  $\delta$  ( $\alpha$  étant l'efficacité quantique du système) est plus grand dans le radio que dans l'optique, car l'énergie d'un photon ( $h\nu$ ) est plus faible et que le temps de cohérence ( $\sim 1/\Delta\nu$ ) est plus grand à des fréquences inférieures. Or, comme  $\delta$  augmente lorsque la fréquence diminue (Mandel 1962b; Borra 1997), le signal recherché (les fluctuations) sera plus facile à détecter aux basses fréquences, c'est-à-dire que le S/B sera meilleur aux radiofréquences.

En plus, vu que les sources brillantes connues subissant des effets de lentilles fortes sont principalement des NAGs, comme les quasars lointains, il est alors recommandé d'observer dans le domaine radio (Borra 2014). Par contre, les observations radio introduisent du bruit système ou bruit thermique (sol, atmosphère et récepteur), plus présent dans le radio que dans l'optique. Heureusement, cet inconvénient est minimisé par le fait que les observations du NRAO se font dans des conditions où le bruit système est le plus faible possible (ex., récepteurs refroidis dans

des enceintes cryogéniques). Sur ce point, on a déjà vu, à la section 2.1.5, que l'intensité du bruit système d'un millijansky, calculée par l'équation (2.16), était beaucoup plus faible que l'intensité d'une source radio brillante (ex., 1 Jy) observée pendant une dizaine de minutes.

Les milieux de propagation des ondes radio, principalement le plasma intergalactique, dispersent le signal (l'impulsion) de l'ordre de la milliseconde. Dans les faits, une impulsion initiale, de largeur temporelle aussi courte que les 15.625 nanosecondes nécessaires pour échantillonner les données, sera étalée de quelques centaines de microsecondes, dépendamment du redshift de la radiosource (voir tableau 6.1).

Relevons, comme dernier inconvénient de la technique, qu'auto-corréler numériquement des intensités, au lieu de les auto-corréler avec de l'instrumentation spécialisée comme des corrélateurs utilisés en radiointerférométrie, requiert le stockage d'une grande quantité de données. À titre d'exemple, considérant l'échantillonnage très rapide de 15.625 ns du NRAO, il faut 150 gigaoctets de mémoire (voir tableau A.5) afin de stocker les données de seulement dix minutes d'observation, et cela même pour des données enregistrées sur uniquement deux bits (voir section 4.5.3).

## **4.5 Signaux numérisés**

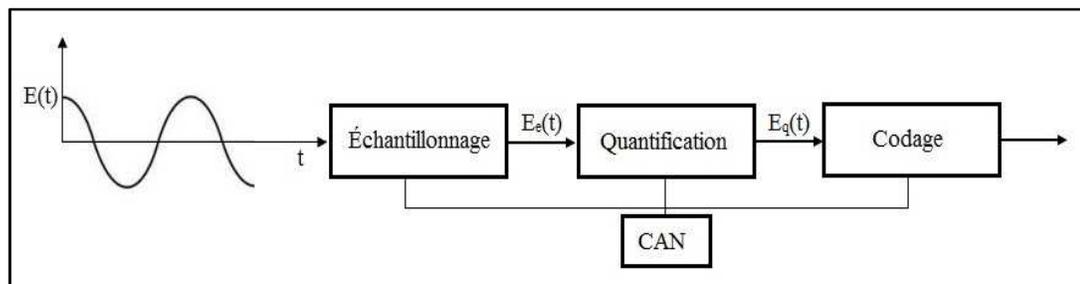
### **4.5.1 Notions de base sur les signaux**

On distingue deux principaux types de signaux électroniques, à savoir, les signaux analogiques et les signaux numériques. Les signaux analogiques sont des signaux dont les valeurs varient de façon continue dans le temps. Ce sont donc des signaux à temps continu où il est possible de mesurer les valeurs à tout instant, c'est-à-dire qu'il y a une infinité de valeurs possibles. D'un autre côté, si on peut seulement mesurer les valeurs des signaux reçus à des intervalles de temps discrets fixes, il

s'agit alors de signaux à temps discret, communément appelés signaux numériques (Leland 1996; Ambardar 2003; Blanchet & Charbit 2013; Oppenheim 2014). Par exemple, les signaux numériques du NRAO sont des champs électriques discrétisés à chaque 15.625 nanosecondes.

Le traitement numérique du signal, abrégé par TNS, est une discipline technique qui consiste à traiter des signaux à temps discrets. Dans ce type de traitement, les informations sont représentées par un nombre fini de valeurs discrètes bien déterminées, c'est-à-dire par une suite de valeurs numériques.

Le passage d'un signal analogique d'entrée vers un signal numérisé de sortie s'effectue généralement en deux opérations successives (ces deux opérations sont parfois intégrées dans la même composante électronique), soit l'échantillonnage et la quantification (voir figure 4.2). L'opération d'échantillonnage d'un signal analogique consiste à prélever périodiquement des valeurs du signal selon une période d'échantillonnage notée  $T_e$ . Elle correspond donc à une discrétisation en temps. Quant à l'opération de quantification et/ou de codage, elle permet d'associer une valeur numérique binaire à chacun des échantillons prélevés. Elle correspond donc à une discrétisation en amplitude où les valeurs sont codées sur un nombre fini de bits.



**Figure 4.2.** Numérisation d'un signal (champ électrique) sur une chaîne de bits. Généralement, (ex., en génie électrique) un convertisseur analogique-numérique (CAN) est utilisé pour les opérations d'échantillonnage et de quantification et/ou de codage. Les symboles  $E(t)$ ,  $E_e(t)$  et  $E_q(t)$  signifient respectivement le champ électrique analogique, échantillonné et quantifié.

## 4.5.2 Échantillonnage

Compte tenu que le principal but du TNS est l'extraction de l'information comprise dans le signal analogique originel, il va de soi qu'idéalement ces informations doivent être conservées après avoir échantillonné ce signal source.

Soit  $E(t)$  et  $E_e(t)$  représentant respectivement les champs électriques analogiques et échantillonnés (voir figure 4.2). Le signal échantillonné est représenté par le signal analogique composé de la superposition linéaire d'impulsions survenant à des temps,  $t = nT_e$ , soit mathématiquement par,

$$E_e(t) = \sum_{n=-\infty}^{\infty} E(nT_e) \delta(t-nT_e), \quad (4.6)$$

où le terme de sommation,  $\sum_{n=-\infty}^{\infty} \delta(t-nT_e)$ , est le peigne de Dirac (une somme d'impulsions), alors que  $E(nT_e)$  fixe l'amplitude de chaque impulsion.

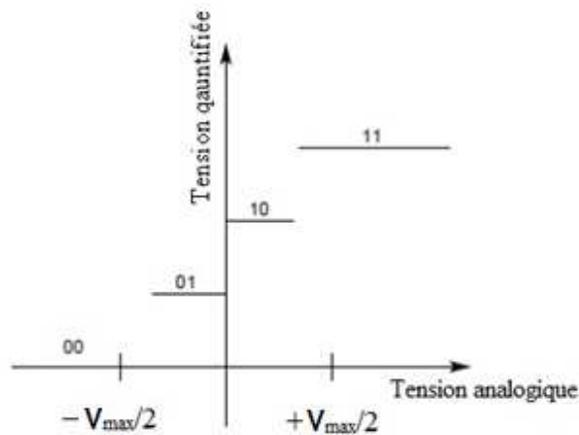
Afin de s'assurer que le signal échantillonné soit une réplique (quasi) parfaite du signal analogique, il faut échantillonner à un taux minimal de Nyquist. Avec un tel échantillonnage, le théorème de Shannon est satisfait. De fait, ce théorème stipule que, pour un spectre borné à une largeur de bande  $B$ , la condition  $F_e \geq 2B$  doit être respectée afin de reconstruire un signal de sortie de manière fidèle au signal d'entrée ( $F_e = 1/T_e$  étant la fréquence d'échantillonnage). Dans le cas où cette condition n'est pas complètement satisfaite, il se produit un repliement du spectre, c'est-à-dire que l'on retrouve la présence de fréquences parasites n'appartenant pas au signal initial (Blanchet & Charbit 2013; Oppenheim 2014).

### 4.5.3 Quantification

Après avoir échantillonné le signal analogique d'entrée, l'étape suivante consiste à quantifier et/ou coder les échantillons. L'opération de quantification consiste à approximer le signal échantillonné, discrétisé en temps, par un signal composé d'amplitudes discrètes codées sur une chaîne finie de bits.

Les données reçues du NRAO sont codées sur deux bits et mènent à quatre niveaux de quantification, lesquels sont représentés par les quatre valeurs binaires possibles des champs électriques, soit 00, 01, 10 et 11 ou encore 0, 1, 2 et 3 en base décimale (voir figure 4.3). Le fait de quantifier les données sur un si petit nombre de bits (ex., deux) a l'avantage de générer un débit de transfert moins élevé, donc facilement gérable. En revanche, la quantification sur si peu de bits cause une perte de sensibilité: la modification des niveaux du champ électrique aux valeurs quantifiées entraîne l'ajout d'une composante de bruit de quantification. En d'autres termes, la quantification sur peu de bits introduit une erreur, dite bruit de quantification (voir section 4.5.5), entre les champs électriques quantifiés et échantillonnés.

À la différence des quatre valeurs décimales entières souvent utilisées pour la quantification à quatre niveaux,  $\pm 1$  et  $\pm 3$  (il est plus facile pour un instrument, tel un corrélateur, de corrélérer des valeurs entières), les valeurs des niveaux inférieurs et supérieurs des champs électriques du NRAO sont respectivement  $\pm 1$  (inchangées) et  $\pm 3.3359$  (table A8.1 et annexe 8.3 de Thompson et al. 2004). Or, puisque, parmi les quatre valeurs décimales possibles de la tension (champ électrique) quantifiée, seule la valeur 1 peut être représentée sur deux bits (avec un niveau de codage en complément à deux pour le signe négatif, la valeur -1 peut également être codée sur deux bits), on doit implanter une fonction simple, mais primordiale (voir annexe B.3), permettant de convertir les tensions positives quantifiées, sur seulement deux bits (voir figure 4.3), en quatre valeurs de tensions décimales signées, soit -3.3359, -1, 1 et 3.3359.



**Figure 4.3.** Quantification sur deux bits menant à quatre niveaux de codage. Pour une tension analogique d'entrée comprise dans la fourchette de tensions continues,  $-V_{\max}/2 - V_{\max}/2$ , la tension de sortie quantifiée est une des quatre valeurs binaires, 00, 01, 10, 11.

#### 4.5.4 Bruit aléatoire gaussien

Par opposition aux signaux temporels déterministes (ex., sinusoïdes, échelons, impulsions), où il est possible de représenter leur évolution grâce à des fonctions mathématiques, les signaux astronomiques radio (ex., le rayonnement synchrotron) sont de nature aléatoire dont l'évolution est imprévisible: la connaissance de la valeur du signal à l'instant  $t$  ne permet pas de prédire sa valeur à l'instant  $t+\Delta t$ .

Le bruit est considéré ici comme un signal aléatoire (un signal aléatoire n'est pas forcément un bruit, puis un bruit peut aussi être déterministe), c'est-à-dire que le signal bruité est dû au hasard. Ce bruit peut alors être modélisé par des processus aléatoires et traité par les lois de probabilités classiques, entre autres pour un bruit aléatoire gaussien, la loi de Gauss. C'est ainsi qu'une variable aléatoire stationnaire normalement distribuée  $X$ , ou tout simplement un bruit gaussien, suit une loi normale de moyenne  $\mu$  et de variance  $\sigma^2$ , dont la densité de probabilité est donnée par,

$$p(X) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{X-\mu}{\sigma}\right)^2}. \quad (4.7)$$

#### 4.5.5 Bruit et efficacité de quantification des données du NRAO

Pour des signaux aléatoires ou pour du bruit aléatoire, la série d'erreurs de quantification est également aléatoire, à tel point que l'on parle alors de bruit de quantification. Cela dit, Thompson & Emerson (2005) montrent que le bruit de quantification, noté  $q$ , associé à quatre niveaux de codage des données (ex., champs électriques sur deux bits du NRAO) vaut,

$$q = V_{\text{anal}} - 0.5130V_{\text{quant}}, \quad (4.8)$$

avec  $V_{\text{anal}}$  et  $V_{\text{quant}}$  représentant respectivement les tensions analogiques et quantifiées. (Notez que dans l'équation précédente, on a remplacé leur constante 0.5651, calculée pour un seuil de  $\pm 0.996$ , par la constante 0.5130, calculée pour un seuil de  $\pm 0.982$ .)

Bien que l'on connaisse maintenant les notions de quantification et de bruit ou d'erreur de quantification, il n'en reste pas moins que la notion la plus utile à maîtriser pour le projet est celle de l'efficacité de quantification d'un signal. Pour tout dire, cette efficacité est définie en comparant le TNS quantifié à un traitement idéal non quantifié. Mathématiquement, cela signifie que le facteur d'efficacité s'exprime par le ratio entre les rapports signal à bruit avec et sans quantification. De cette façon, pour la quantification à quatre niveaux du NRAO, où l'amplitude du plateau supérieur vaut  $\pm 3.3359$ , on montre que l'efficacité de quantification est de 88.25 % (voir table A8.1 de Thompson et al. 2004).

Pour terminer, ce qui est essentiel de connaître pour le projet est la diminution du rapport signal sur bruit causé par la quantification. Ceci implique que lorsqu'on fixe un seuil de détection, basé sur le rapport signal à bruit (S/B) du signal quantifié, on doit toujours garder à l'esprit que le seuil réel sans quantification correspond à  $1/0.8825 = 1.133$  fois le seuil fixé avec quantification. Par exemple, une limite de détection,  $S/B = 4.0$  (voir section 5.3.3), signifie que l'amplitude du signal réel, échantillonné et non quantifié, doit être  $1.133 \times 4.0 = 4.53$  fois celle du bruit pour que le signal soit détecté, et non pas 4.0. À vrai dire, le seuil de détection réel est supérieur à 4.53, car il faut aussi inclure tous les autres facteurs multiplicatifs d'efficacité (pertes) qui ont tendance à faire diminuer le S/B, c'est-à-dire à faire augmenter la sensibilité (bruit) du système (voir section 2.1.5).

## 5 Méthodologie et traitement des données Matlab/C++

Ce chapitre est d'abord consacré à l'implantation des algorithmes et des logiciels. Sur ce point, le lecteur intéressé à en savoir davantage sur les programmes implantés est invité à consulter les annexes C et D qui donnent respectivement l'écriture des codes commentés Matlab et C++. Ensuite, la robustesse des codes implémentés est décrite en montrant quelques exemples de validations et de détections de signaux artificiels. Pour clore le chapitre, les techniques quantitatives développées, afin de détecter des candidats raisonnables aux effets de lentilles gravitationnelles (ou de faibles variations rapides de flux), sont explicitées.

### 5.1 Implantation des algorithmes en Matlab

Pour commencer le traitement informatique, on met au point les algorithmes de base avec le logiciel Matlab sur un microordinateur, mais en n'utilisant qu'une très petite partie (un sous-échantillon) des volumineux fichiers de données (voir tableau A.5). En fait, comme l'espace alloué pour l'allocation dynamique de tableaux sous Matlab, avec le système d'exploitation Windows de 64 bits, est limité à une dizaine de gigaoctets, on comprend bien qu'on ne peut pas stocker simultanément toutes les données des gros fichiers du NRAO dans un vecteur de Matlab. Or, comme ces fichiers sont segmentés en *data frames* (voir figure A.1), il devient alors possible de traiter sous Matlab un nombre limité de *data frames*.

Malgré que la mémoire vive de Matlab permet de traiter, pour un canal donné, jusqu'autour de 20000 *data frames*, soit 20000 *data frames* X 2500 données/*data frame* = 50 millions de données (voir section 5.1.1), on commence la mise au point du logiciel Matlab en se limitant à traiter un sous-échantillon de seulement 1000 *data frames* (2.5 millions de données). L'avantage de travailler avec un si petit sous-échantillon provient du fait que le temps de calcul n'est que de quelques secondes comparativement à quelques heures avec un échantillon de 20000 *data frames*.

### 5.1.1 Lecture des données des *data frames* de format Mark 5B

Afin de traiter les données d'un ou de plusieurs *data frames*, on doit implanter une fonction permettant de lire sous Matlab les données de format Mark 5B (voir figure A.1). Cette fonction de lecture, nommée *dread\_mk5()* (voir annexe C.2), lit d'abord l'en-tête (16 octets) de tous les *data frames* à traiter. Après quoi, elle lit et extrait les données brutes, c'est-à-dire les champs électriques, de chaque *data frame*. Pour cela, la fonction lit les deux segments (4992 et 5008 octets) par paquets de deux bits pour un total de 40000 échantillons répartis dans 16 canaux, soit  $40000/16 = 2500$  échantillons lus par canal et par *data frame* traité. En terminant, rappelons qu'on a déjà mentionné à la section 4.5.3 que l'on doit ajouter à la fonction de lecture une fonction simple (voir annexe C.3), mais fondamentale, permettant de convertir les quatre valeurs possibles positives sur deux bits du champ électrique (0, 1, 2 et 3) en quatre valeurs décimales quantifiées, soit  $\pm 1$  et  $\pm 3.3359$ .

### 5.1.2 Calcul et lissage des intensités

Le calcul du vecteur intensité, noté  $I(t)$ , est trivial puisqu'il s'agit tout simplement de calculer le carré du vecteur champ électrique, noté  $E(t)$ , du canal traité, soit mathématiquement (en ignorant les constantes multiplicatives sans conséquence),  $I(t) = E^2(t)$ .

Le lissage des intensités est effectué à l'aide d'un filtre passe-bas RIF (à réponse impulsionnelle finie) à moyenne mobile (*moving averaging FIR filter*). Le filtre RIF est non récursif et est simple à utiliser. Il n'est donc pas surprenant qu'il soit couramment utilisé en traitement numérique des signaux. Nonobstant sa simplicité d'utilisation, ce type de lissage est parmi les plus optimaux pour atténuer l'amplitude du bruit aléatoire tel que le bruit gaussien auquel on est confronté. Notez que deux autres types de filtrage ont été essayés, à savoir, le filtre RII (à réponse impulsionnelle infinie) et le filtre RIF de lissage par moindres carrés (filtre Savitzky-Golar). Par contre, ces deux filtres étaient plus difficiles à implanter et à appliquer et

ne donnaient pas de résultats plus optimaux que le simple filtre passe-bas RIF utilisé. (Pour une explication détaillée sur le design de filtre, on invite l'examineur à consulter l'excellent manuel d'Orfanidis (2010), qui très utilisé en ingénierie électrique et informatique.)

Pour un filtre RIF à moyenne mobile de longueur  $N$ , tous les coefficients (réponse impulsionnelle)  $h_n$  sont égaux et valent (Orfanidis 2010, Oppenheim 2014),

$$h_n = \frac{1}{N}, \quad n = 0, 1, 2, \dots, N-1. \quad (5.1)$$

Avec cette valeur de  $h_n$ , l'équation de convolution de ce type de filtrage se réduit à,

$$y(n) = \frac{1}{N} \sum_{m=-\left(\frac{N-1}{2}\right)}^{m=\left(\frac{N-1}{2}\right)} x(n-m). \quad (5.2)$$

Par exemple, dans le cas simple d'un lissage avec  $N = 5$  points, en développant l'équation précédente, il vient que,

$$y(n) = \frac{1}{5} [x(n+2) + x(n+1) + x(n) + x(n-1) + x(n-2)], \quad (5.3)$$

où, dans notre cas, l'échantillon  $x(n)$  est équivalent à l'intensité  $I$  à la position  $n$ , alors que  $y(n)$  équivaut à l'intensité de sortie, dite intensité lissée, obtenue en remplaçant, pour chaque  $n$ ,  $I(n)$  par une moyenne calculée avec  $I(n)$  et avec les deux intensités en avant et en arrière de  $I(n)$ .

Les intensités lissées, ou plus justement, les fluctuations des intensités lissées, sont celles où les fluctuations rapides instantanées sont moyennées sur une portée de

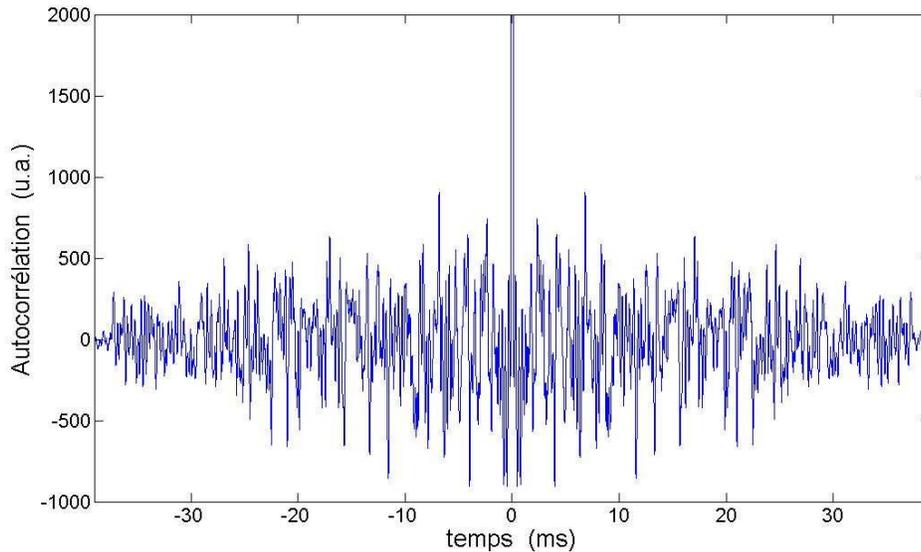
N points. Il en résulte que leurs variations temporelles sont plus lentes et que leurs amplitudes sont plus faibles.

En raison de l'utilisation courante du filtre RIF à moyenne mobile, il va de soi qu'il existe déjà une fonction prédéfinie Matlab pour ce type de filtrage, dénommée *smooth()* (voir annexe C.1), laquelle effectue un lissage du signal (des intensités) d'entrée et retourne en sortie un signal moyenné sur les N points, passés en paramètre, du filtre.

### 5.1.3 Autocorrélation temporelle d'intensité

Pour effectuer l'autocorrélation temporelle on utilise la fonction prédéfinie Matlab, *xcorr()* (voir annexe C.1). Cette fonction auto-corrèle directement le vecteur I passé comme argument. Par ailleurs, vu que *xcorr()* donne une forme d'autocorrélation comprenant une pente descendante très abrupte (ex., voir figure 5.3), on soustrait la moyenne des intensités afin de simplifier les techniques de détection des signaux: on auto-corrèle les fluctuations d'intensité  $\Delta I(t)$  (voir équation (4.5)). En Matlab, ceci s'écrit: *xcorr (I - mean(I))*.

La figure 5.1 montre un profil temporel d'autocorrélation d'intensité, obtenu pour un sous-échantillon de 2.5 millions de points, découlant de l'observation du quasar 3C 345. Notez que l'amplitude nulle aux extrémités provient de l'effet d'un échantillonnage fini. Dans les faits, on effectue une somme de produit sur un intervalle de temps fini, c'est-à-dire qu'on fait une autocorrélation tronquée, tandis qu'en théorie l'autocorrélation d'intensité devrait se faire sur un intervalle de temps infini (voir équation (4.5)).

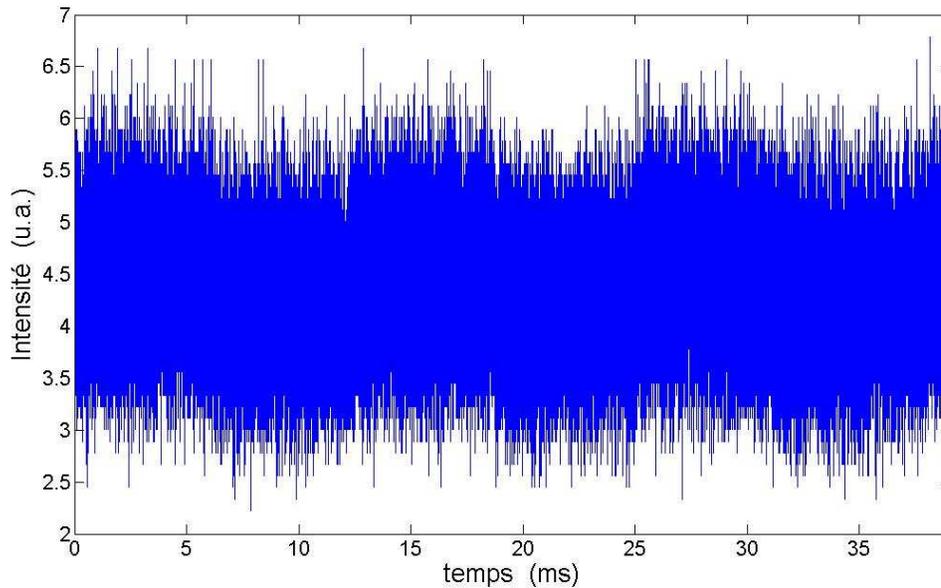


**Figure 5.1.** Profil temporel d'autocorrélation d'intensité (en unités arbitraires), obtenu pour un sous-échantillon de 1000 *data frames* ou 2.5 millions de points (39 ms), résultant de l'observation de l'objet source 3C 345. Le lissage de 7000 points correspond à environ la moitié de l'élargissement temporel dû au milieu intergalactique (voir tableau 6.1).

#### 5.1.4 Validation du code Matlab par la détection d'un signal périodique

Une façon simple et efficace de valider le bon fonctionnement du code Matlab développé consiste à vérifier si le code pour l'autocorrélation d'intensité est a priori capable de détecter un signal périodique de forte amplitude. Pour cela, on utilise encore un sous-échantillon de seulement 2.5 millions de données provenant, cette fois-ci, du quasar PKS 1622-253.

Pendant l'observation du tout premier objet, PKS 1622-253, un signal de commutation-bruit (*switch-noise*) périodique de 80 Hz a été introduit par le NRAO. La figure 5.2 donne l'intensité temporelle du sous-échantillon de PKS 1622-253. On remarque que bien qu'il s'agisse d'un signal de grande amplitude, la périodicité du signal est difficilement distinguable. En vérité, dans le profil d'intensité, les demi-périodes (6.25 ms) sont totalement perdues dans le bruit: l'unique raison pour laquelle ces demi-périodes ou discontinuités sont devenues visibles, à travers le bruit, découle de l'application d'un lissage qui diminue l'amplitude de ce bruit.

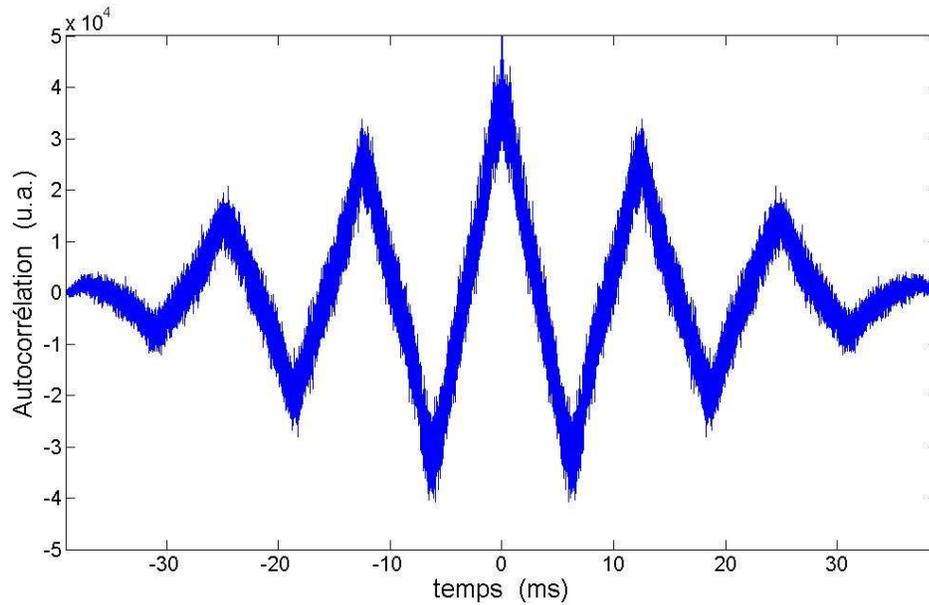


**Figure 5.2.** Profil temporel d'intensité (en unités arbitraires), obtenu pour un sous-échantillon de 2.5 millions de points (39 ms), correspondant à l'observation de l'objet PKS 1622-253. Malgré un lissage de 90 points réduisant l'amplitude du bruit, les demi-périodes (6.25 ms) restent difficilement identifiables.

Toujours pour le même sous-échantillon de PKS 1622-253, on retrouve, à la figure 5.3, le profil temporel d'autocorrélation d'intensité dans lequel on a enlevé la moyenne. À l'encontre du profil d'intensité, le profil de la figure 5.3 illustre clairement la signature d'un signal périodique: les pics (triangles) sortent nettement du bruit. Il faut dire que l'autocorrélation globale sur le domaine temporel du sous-échantillon (39 ms) résulte des six autocorrélations individuelles associées aux six discontinuités de 6.25 ms présentes dans le signal d'entrée.

Les grandes amplitudes des sommets des triangles valident judicieusement que la technique d'autocorrélation d'intensité permet de détecter la présence de signaux périodiques. Qui plus est, cela valide implicitement la fonction *dread\_mk5()* développée. En effet, une erreur dans la lecture des données, comme insérer (au lieu d'ignorer) une partie des champs d'identification du format Mark 5B (voir figure A.1), ne serait-ce que quelques bits, à la limite un seul, dans les données brutes

(champs électriques), donnerait, dans le profil d'autocorrélation, des pics d'intensité générés par la présence de faux signaux issus des en-têtes.



**Figure 5.3.** Profil temporel d'autocorrélation d'intensité (en unités arbitraires), obtenu pour un sous-échantillon de 2.5 millions de données (39 ms), provenant de l'observation du quasar PKS 1622-253. Afin de bien voir le signal auto-corrélé, la moyenne des intensités a été soustraite. La séparation de 12.5 ms entre les pics (triangles) de l'autocorrélation correspond à la périodicité du signal de commutation-bruit de 80 Hz introduit par le NRAO. Contrairement au profil d'intensité, l'amplitude des pics de l'autocorrélation indique distinctement la présence de la périodicité.

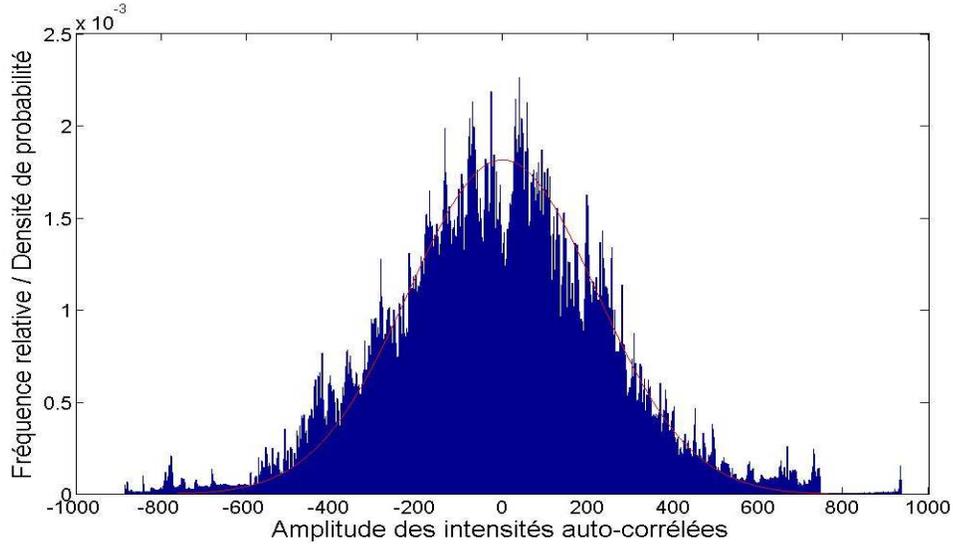
### 5.1.5 Distribution du bruit dans l'autocorrélation d'intensité

Vu que la méthode de détection appliquée dans cette thèse repose sur l'autocorrélation d'intensité, il est plus pertinent de vérifier la distribution du bruit dans le signal auto-corrélé que dans le signal d'entrée. Pour cela, on détermine d'abord l'histogramme de la distribution de l'amplitude des intensités auto-corrélées (où on a soustrait la moyenne). Ensuite, pour qualitativement illustrer jusqu'à quel point cette distribution peut être comparée à une gaussienne, on détermine à partir de l'histogramme la valeur de la variance (largeur à mi-hauteur). Enfin, on utilise cette variance pour obtenir la distribution théorique de Gauss.

La figure 5.4 illustre l'histogramme de la distribution de l'amplitude des intensités auto-corrélées et la densité de probabilité théorique gaussienne (courbe rouge). En examinant cette figure, on constate que malgré quelques différences au sommet, comme de fortes fluctuations d'amplitudes, et à la base de l'histogramme, comme la présence d'ailes, la distribution du bruit auto-corrélé suit assez bien celle d'un bruit gaussien.

Le fait que le bruit (auto-corrélé) ne soit pas parfaitement gaussien nuit en rien à l'analyse des données. C'est pourquoi, une comparaison qualitative (une estimation) avec un bruit gaussien est amplement suffisante: on n'a pas besoin de quantifier précisément et systématiquement le niveau du bruit (gaussien). Qui plus est, les multiples sources de bruit possibles (sol, atmosphère, récepteur, antenne, CMB, etc.) empêchent de connaître la distribution théorique exacte du bruit (voir sections 2.1.3 et 2.1.4).

La principale raison de connaître, ou à dire vrai d'estimer, la distribution du bruit réside dans le fait qu'elle permet de fixer une limite de détection (basée sur le S/B) pour l'extraction de signaux (pics), ainsi que de comparer le nombre réel de pics extraits au nombre prévu pour une gaussienne. Par exemple, pour l'objet source 3C 345, on a détecté, avec un seuil  $S/B = 4.0$ , 180 pics alors que la gaussienne en prédisait 174 (voir section 5.3.3).



**Figure 5.4.** Histogramme de la fréquence relative des amplitudes des intensités auto-corrélées (on a enlevé la moyenne) pour un sous-échantillon de 2.5 millions de points de l'objet source 3C 345. La courbe rouge superposée représente la densité de probabilité théorique de Gauss calculée à partir de l'écart-type tiré de l'histogramme (2.3548 fois la largeur à mi-hauteur). Afin d'augmenter la précision, l'histogramme comprend  $2 \times 2500 = 5000$  classes de largeur identique (1000 points) non v sur cette figure. Le lissage est de 7000 points.

## 5.2 Traduction du code Matlab en code C++

### 5.2.1 Lecture des données et autocorrélation d'intensité en C++

Afin de lire et d'extraire les données d'un ou de plusieurs *data frames*, mais cette fois-ci, avec le langage de POO C++, on doit d'abord traduire (adapter) la fonction Matlab *dread\_mk5()* en C++. À ce propos, le fichier *dread\_mk5.cpp* contient le code C++ commenté pour lire et extraire exclusivement les données de format Mark 5B (voir annexe D.1). Par suite, on doit écrire le code C++ permettant d'autocorréler les intensités (voir annexe D.2). Somme toute, à partir d'un fichier d'entrée (voir tableau A.5), cette fonction auto-corrèle les intensités, puis écrit les résultats de l'autocorrélation dans un fichier de sortie de format inhabituel, hdf5, et de taille quatre fois plus grande que celle du fichier d'entrée (voir section 5.2.4).

## 5.2.2 Validation du code C++

Compte tenu que le développement du code C++ (sur le superordinateur) est plutôt complexe (présence d’erreurs et d’artéfacts possibles), il est essentiel de le valider en comparant les résultats C++ aux résultats Matlab (sur un microordinateur), dont le code est beaucoup plus simple à implanter, donc pratiquement sans risques d’erreurs. Par exemple, l’autocorrélation en Matlab s’effectue directement via la fonction prédéfinie *xcorr* (), laquelle est évidemment exempte d’erreur (validée par les développeurs de Matlab). Pour ce faire, il faut, bien entendu, n’utiliser que de petits échantillons de données vu la mémoire vive limitée de Matlab. Cela dit, en faisant la comparaison des résultats pour diverses combinaisons de fichiers d’entrée, c’est-à-dire de divers objets observés, et du nombre de *data frames* (maximum de ~20000 avec Matlab), on a obtenu une concordance parfaite des résultats. À titre d’exemple, en recommençant en C++ l’autocorrélation du sous-échantillon de 1000 *data frames* du quasar PKS 1622-253, on a détecté la même périodicité de 12.5 millisecondes du signal de commutation-bruit inséré par le NRAO (voir figure 5.3). En un mot, le fait de retrouver dans les deux profils d’autocorrélation d’intensité, Matlab et C++, la signature du signal périodique de 12.5 ms valide le code C++ développé.

## 5.2.3 Soumission des calculs sur un superordinateur

Pour exécuter le code C++ sur le système d’exploitation Linux d’un superordinateur, il faut créer un fichier de soumission en lot (*batch file*). Ce fichier, similaire à un script *Shell* de Linux, permet de soumettre des tâches pour le traitement en parallèle (plusieurs cœurs de calcul) sur le superordinateur Colosse de l’Université Laval. Par exemple, pour traiter dix minutes d’observation d’un objet donné, on a besoin de 1024 cœurs de calcul. Malheureusement, même si le temps de calcul des tâches soumises ne prend que ~30 minutes, il faut souvent attendre une semaine avant d’obtenir les résultats puisque l’ordonnanceur automatique des tâches du superordinateur donne évidemment priorité, dans la file d’attente, aux tâches

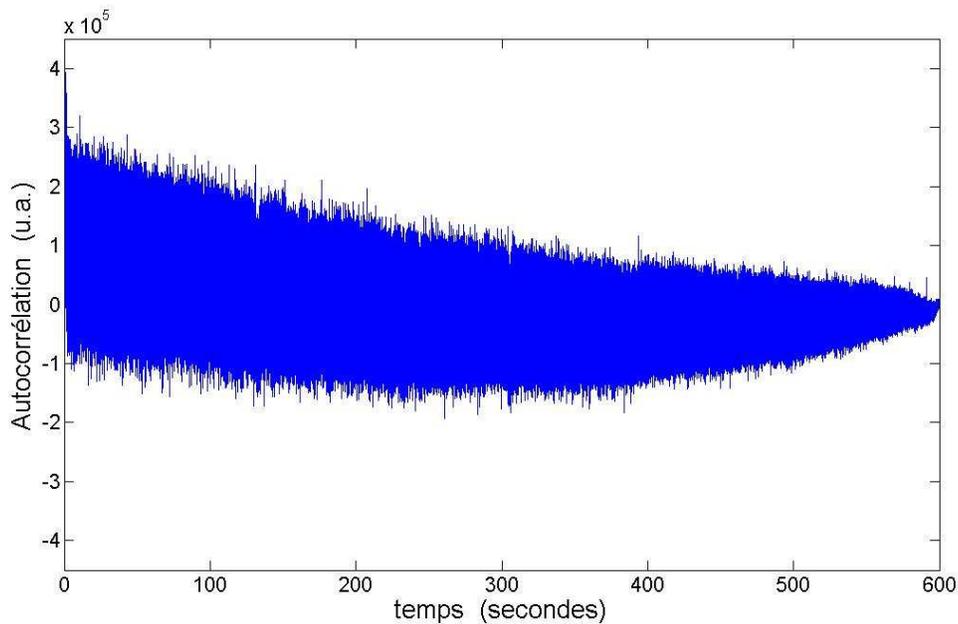
requérant peu de cœurs de calcul. Enfin, le nombre d'analyses effectuées (nombre de canaux auto-corrélés) est au minimum de 22 objets X 4 stations/objet X 16 canaux/station = 1408 canaux. À vrai dire, il y a eu plus que 1400 traitements (autocorrélations), disons 1500 pour fixer les idées, car certains canaux (de certains radiotélescopes) problématiques ont dû être repris, particulièrement à cause de la présence de signaux artificiels parasites.

La nécessité d'utiliser 1024 cœurs de calcul provient du fait que le superordinateur Colosse ne possède pas de disques locaux. Par ailleurs, la limitation de mémoire (ex., Matlab) ne peut pas être contournée par l'utilisation de disques SSDs (*Solide State Drive*) étant donné que les SSDs n'offrent pas les mêmes performances que la mémoire vive de Colosse. En effet, les meilleurs SSDs de 2017 offrent ~6 Gbits/s de bande passante, alors que la mémoire vive de Colosse (datant de 2009) en offre ~30 Gbits/s. Qui plus est, la librairie *FFTW*, laquelle est nécessaire pour faire l'autocorrélation en C++ (une TF et une TFI), fonctionne en mémoire vive et non pas sur disque. En somme, il n'est pas possible de faire les calculs avec de l'espace de stockage sur disque puisqu'il faut stocker le jeu de données en mémoire vive.

#### 5.2.4 Récupération avec Matlab des données traitées en C++

Les dizaines de milliards d'intensités auto-corrélées, résultant de l'exécution du code C++, sont sauvegardées dans un fichier de sortie de format atypique hdf5, qui permet la structuration et l'enregistrement de fichiers, et récupérer sous Matlab avec la fonction prédéfinie *h5read()* (voir annexe C.4). En raison de la mémoire vive limitée de Matlab (~10 Go), on comprend bien qu'il n'est pas possible de récupérer, dans une seule opération de lecture, les dizaines de milliards de données de sortie résultant de l'exécution du programme C++. D'où, afin de récupérer toutes ces données, il est nécessaire d'implémenter une boucle parcourant tous les blocs de données, lus par *h5read()*, du fichier des résultats. Par exemple, la récupération de ~38 milliards de données, via des blocs de 50 millions, requiert une boucle itérative de 760 opérations de lecture.

La figure 5.5 montre un exemple du profil d'autocorrélation d'intensité découlant de l'observation, à la station SC, de l'objet source 3C 345. Notez que la pente dans ce profil non uniforme découle de la présence de signaux artificiels non périodiques (voir section 5.2.6), et non pas du fait que la moyenne de l'intensité est soustraite avant de faire l'autocorrélation. Dans le cas où il n'y aurait pas de signaux (non périodiques) artificiels à la station SC, le profil de la figure 5.5 serait similaire à celui de la figure 5.9. Le profil d'autocorrélation, qui serait alors uniforme (sans pente), n'aurait pas besoin d'être rectifier avant de procéder à la technique d'extraction des pics (voir section 5.3.3).



**Figure 5.5.** Illustration du profil temporel (temps positifs) des intensités auto-corrélées (en unités arbitraires), provenant du fichier de sortie de 3C 345, pour la station SC. La pente du profil est générée par des signaux artificiels non périodiques. Le pas d'échantillonnage est de 1000 points (15.625  $\mu$ s). Cet échantillonnage permet de visualiser l'allure de l'autocorrélation complète en utilisant seulement un millième des données (vecteur de  $\sim$ 38 millions de points stockable en Matlab). Le lissage de 7000 points correspond à la moitié de l'étalement temporel causé par l'espace intergalactique (voir tableau 6.1).

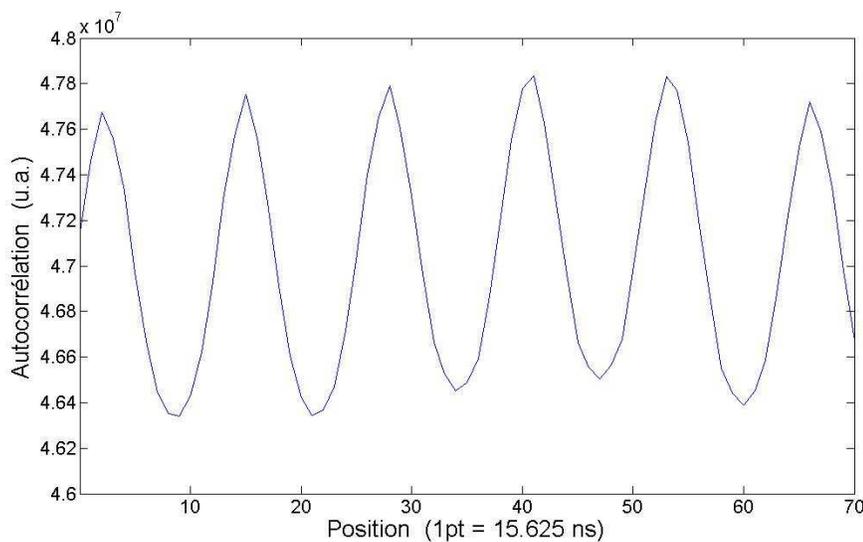
### 5.2.5 Détection d'un signal périodique rapide de faible amplitude

En reprenant l'autocorrélation de la radiosource PKS 1622-253 (voir section 5.1.4), toujours pour un lissage de 90 points, avec, cette fois-ci, toutes les données (~3 milliards), on détecte un deuxième signal périodique de très faible amplitude d'environ 5 MHz, dont seulement les 70 premiers points du profil de l'autocorrélation sont illustrés à la figure 5.6. Ce faible signal est composé d'impulsions étroites périodiques, injectées par le système de calibration d'impulsions du NRAO à toutes les ~200 nanosecondes, afin de générer des tonalités monochromes (de phases stables) nécessaires à la calibration de la phase instrumentale.

Le fait qu'on n'a pas détecté avec Matlab cette périodicité rapide de ~200 ns (5 MHz), laquelle est comparable (treize fois supérieure) à la période d'échantillonnage du NRAO (15.625 ns), s'explique par le nombre très restreint de données (50 millions ou 39 ms) utilisées dans l'autocorrélation du signal. En réalité, avec Matlab, la périodicité du faible signal auto-corrélé se trouve perdue dans le bruit. Or, considérant que la détection de la très faible amplitude de ce signal auto-corrélé nécessite de traiter un très grand nombre de données du fichier d'entrée associé à PKS 1622-253 (voir tableau A.5), voire la totalité du fichier (~3 milliards de données), seul le code de l'autocorrélation développé en C++ permet d'en détecter la périodicité. De fait, dans le profil des intensités auto-corrélées en C++, la faible amplitude du signal est bien visible (voir figure 5.6), tandis que les impulsions périodiques étroites sont totalement perdues dans le bruit dans le profil des intensités, même en C++. À ce point-ci, il est important de souligner (rappeler) que malgré la mémoire vive limitée de Matlab, qui dans ce cas-ci n'a pu permettre de détecter le faible signal de 5 MHz, on doit aussi développer un code en Matlab, car il permet entre autres de valider le code (complexe) C++ (voir section 5.2.2), de récupérer les données de format non usuel hdf5 sur un microordinateur, d'afficher les données traitées (ex., les figures d'autocorrélation) sur une interface graphique (le code C++ sur le superordinateur est en mode commande sur Linux et ne permet donc pas l'affichage) et d'extraire les pics des intensités auto-corrélées (voir annexe D.4).

Il est aussi important de noter que même si l'autocorrélation et la transformée de Fourier (TF) sont deux méthodes très efficaces pour détecter des signaux périodiques, l'autocorrélation a l'avantage d'être plus intuitive que la TF (Borra 2014). Un autre avantage de l'autocorrélation d'intensité provient du fait que son profil temporel permet de bien visualiser la forme du signal périodique, à travers le bruit, tout au long du temps (ex., 600 s). À l'inverse, dans le profil de la TF de l'intensité, la forme du signal a tendance à se perdre rapidement dans le bruit radio dès que l'on s'éloigne quelque peu de la fréquence fondamentale du signal (1<sup>ère</sup> harmonique, 2<sup>ème</sup> harmonique, etc.).

En somme, la détection de ce signal artificiel de 5 MHz valide que notre logiciel C++ d'autocorrélation d'intensité peut facilement détecter des signaux périodiques très rapides et de très faibles amplitudes. Par exemple, ceux-ci pourraient provenir d'objets exotiques (ex., trous noirs dans les quasars) émettant de faibles signaux à pulsations rapides.



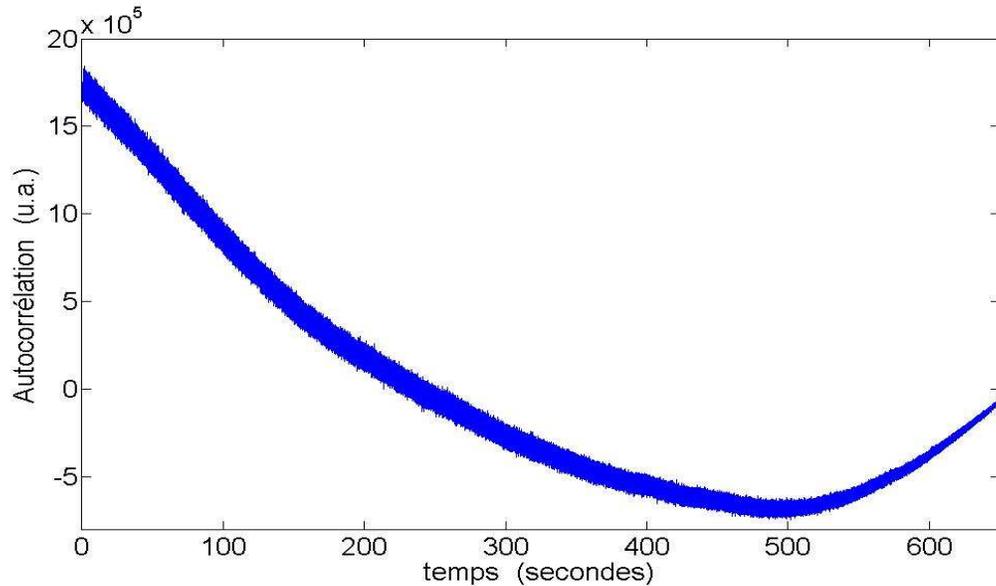
**Figure 5.6.** 70 points ( $\sim 1 \mu\text{s}$ ) du profil d'autocorrélation d'intensité (u. a.) résultant du faible signal périodique de  $\sim 5 \text{ MHz}$  ( $\sim 200 \text{ ns}$ ) inséré pour observer PKS 1622-253. L'abscisse est étiquetée en points afin de mieux distinguer les six impulsions apparaissant à tous les  $\sim 12$  points. Le lissage est de 90 points.

### 5.2.6 Détection de signaux non périodiques faibles

Pour la première fois, on montre que notre technique d'autocorrélation peut aussi être utilisée pour détecter des variations temporelles d'intensité non périodiques de faibles amplitudes variant sur de courtes échelles de temps. Par exemple, il pourrait s'agir d'objets exotiques émettant de brèves impulsions radio non périodiques ou de sursauts radio rapides comme le FRB121102 (voir section 7.7). Ces impulsions rapides seraient très difficiles à détecter au moyen des techniques conventionnelles. Ceci étant dit, le paragraphe suivant décrit comment l'autocorrélation d'intensité arrive à détecter de faibles signaux non périodiques.

Lorsque  $I(t)$  contient des fluctuations d'intensité non périodiques minimales,  $\Delta I(t)$  dans l'autocorrélation  $R(\tau)$  (voir équation (4.5)) contient évidemment ces fluctuations, mais contient aussi des fluctuations aléatoires découlant du bruit gaussien dans les données. L'intégrale dans l'équation (4.5) est effectuée sur le temps de mesure  $T$  (ex., 600 s) et via le produit  $\Delta I(t)\Delta I(t+\tau)$ . Pour un délai temporel  $\tau$  nul, l'intégrale se fait via  $\Delta I(t)\Delta I(t)$  et sur  $T$ . Les fluctuations positives et négatives donnent alors une contribution nette positive puisque l'intégrale s'effectue sur le carré de  $\Delta I(t)$ , soit  $\Delta I^2(t)$ . Ainsi, dans le cas d'un très grand nombre de fluctuations aléatoires, toutes ces fluctuations s'additionnent dans l'intégrale et donnent une grande valeur de  $R(\tau)$ , nettement au-dessus du bruit. Par conséquent,  $R(\tau)$  atteint sa valeur maximale à  $\tau = 0$ . Quand le délai  $\tau$  augmente, l'intégrale est plutôt faite sur  $\Delta I(t)\Delta I(t+\tau)$ , de telle sorte que les fluctuations ne sont plus aux mêmes positions. Dès lors,  $\Delta I(t)\Delta I(t+\tau)$  donnent des contributions positives et négatives dans l'intégrale, ce qui a pour effet de diminuer  $R(\tau)$ . Qui plus est, lorsque  $\tau$  augmente, les fluctuations comprises dans les plus petites et les plus grandes valeurs du temps  $t$  dans  $I(t)$  se chevauchent de moins en moins. Dans le cas limite où  $\tau = T$ , les fluctuations ne se chevauchent pas du tout, ce qui donne  $R(\tau) = 0$ .

La figure 5.7 montre un exemple du profil d'autocorrélation d'intensité obtenu lors de l'observation de la source radio PMN 0134-0931. Ce profil non uniforme (la pente) résulte de la présence de signaux non périodiques artificiels.



**Figure 5.7.** Profil temporel d'autocorrélation d'intensité (en unités arbitraires) résultant de 650 secondes d'observation de la radiosource PMN 0134-0931. Le pas d'échantillonnage est de 1000 points ( $15.625 \mu\text{s}$ ). L'allure non uniforme du profil résulte de la présence d'un grand nombre de signaux artificiels non périodiques. Le lissage de 22000 points correspond à la moitié de l'étalement temporel causé par le MIG (voir tableau 6.1).

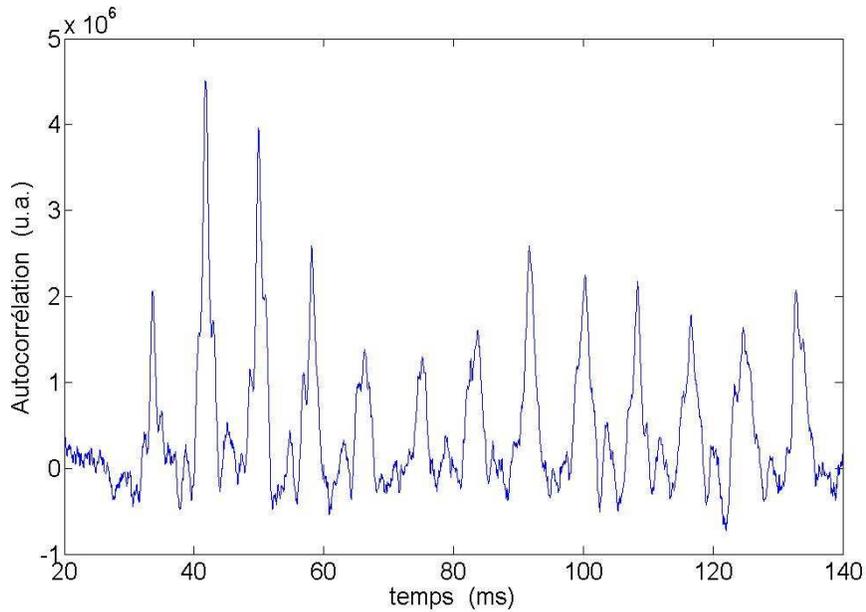
Il est important de souligner que la façon de distinguer les signaux non périodiques qui seraient d'origine cosmique (ex., l'objet à sursauts radio rapides FRB121102) de ceux de nature artificielle consiste à vérifier si les signaux non périodiques se retrouvent dans plusieurs radiotélescopes, au minimum deux, et dans plusieurs canaux, idéalement les seize canaux (voir la seconde étape de la section 5.3.3).

### 5.2.7 Détection de signaux d'interférences radio (RFIs)

Il faut savoir que les signaux parasites dus à des interférences locales de fréquences radio (RFI), comme des systèmes radars au sol ou dans les airs et des satellites, sont une réalité en radioastronomie tout autant que la pollution lumineuse et les raies spectrales atmosphériques sont une réalité dans l'optique. Compte tenu que ces signaux parasites indésirables peuvent s'introduire durant les observations du NRAO, il est primordial que notre logiciel C++ puisse les détecter de façon à les distinguer des signaux d'origine cosmique.

Les émissions cosmiques synchrotron circulairement polarisées issues des observations radio doivent généralement être détectables dans plusieurs télescopes radio (au minimum deux), dans les deux polarisations circulaires (RCP et LCP) et dans la majorité des seize canaux, idéalement les seize. À l'opposé des signaux cosmiques issus de radiosources, les signaux RFI se retrouvent habituellement à l'observatoire radio qui est le site de ces interférences locales et fréquemment dans une seule polarisation circulaire. À titre d'exemple, lors de l'observation du quasar 3C 345, le radiotélescope de la station SC (voir tableau A.4) a détecté un signal RFI, généré par la présence d'un radar à proximité de SC, quasi périodique d'environ 8 ms de polarisation circulaire droite dans les huit canaux impairs (voir tableau A.1), tandis que les trois autres stations VLBA du même objet, BR, HN et OV (et les canaux pairs de SC), n'ont pas détecté ce signal RFI.

La figure 5.8 montre l'allure de ce signal d'interférences locales dans une petite partie du profil temporel d'autocorrélation d'intensité. Au contraire du profil d'intensité (non illustré ici) où le signal RFI est complètement perdu dans le bruit (même après un lissage), les pics de plus grandes amplitudes indiquent clairement la signature du signal d'interférences radio.



**Figure 5.8.** Région du profil temporel (de 20 à 140 ms) d'autocorrélation d'intensité (en unités arbitraires) correspondant à l'observation de 3C 345 à la station SC (canaux impairs seulement). Les pics de plus hautes amplitudes correspondent à une périodicité de  $\sim 8$  ms du signal RFI. À l'inverse du profil d'intensité (non illustré ici), dans lequel ce signal local est totalement perdu dans le bruit, on distingue ici nettement la présence du signal RFI.

### 5.3 Techniques de détection de candidats (signaux)

La méthodologie utilisée pour détecter des candidats raisonnables, à savoir, les candidats donnant un signal cosmique similaire à celui causé par la présence d'une minilente, soit une bosse de largeur  $DM_{MIG}$  (voir section 5.3.1) dans l'autocorrélation, ou encore les candidats (objets exotiques) émettant de brèves pulsations rapides périodiques ou non, consiste à développer des techniques et des algorithmes (programmes) permettant de détecter et d'extraire des signaux (pics) répondant à des conditions préétablies (voir section 5.3.3).

### 5.3.1 Choix des lissages et récupération des données

Pour chacune des 22 radiosources observées, la longueur des quatre lissages en intensité essayés correspond respectivement au quart, à la moitié, aux trois quarts et à la totalité de l'étalement du signal dû au milieu intergalactique. Ces lissages permettent de convertir les deux valeurs possibles d'intensité, soit 1 et 11.1282, en un continuum de valeurs possibles. Ce continuum de valeurs a pour effet de diminuer le niveau de bruit, et donc d'améliorer la précision nécessaire aux traitements des données. Par exemple, un lissage correspondant à tout l'élargissement du signal causé par le MIG semble à priori plus précis (meilleur rapport signal à bruit), mais en contrepartie, un lissage aussi élevé risque de réduire ou d'aplatir, à la limite d'annihiler, le signal dans le cas où la dispersion intergalactique du signal serait plus faible qu'anticipée (voir tableau 6.1): l'amplitude du signal recherché pourrait être diminuée au point où le signal ne serait plus visible à travers le bruit. Pour comprendre cela, il faut savoir que le calcul de la mesure de dispersion DM, donc du délai temporel (voir équation (2.22b)), reste un calcul approximatif basé sur des hypothèses simplificatrices faites sur les milieux de propagation. En particulier, pour calculer la  $DM_{MIG}$ , on avait supposé un plasma intergalactique composé d'hydrogène pur complètement ionisé (voir section 2.2.2). Pour terminer, ajoutons que la dispersion causée par la présence du MIG n'altère pas le dédoublement du signal. En effet, vu que la DPO entre le signal et sa réplique est de beaucoup inférieure à la longueur du trajet total de chacun des deux signaux, c'est-à-dire  $DPO/\text{trajet total} \ll 1$ , la différence des effets dispersifs (due MIG) entre les deux signaux est totalement négligeable.

Pour la majorité des exemples déjà donnés dans ce chapitre, le lissage retenu correspondait à la moitié de l'étalement temporel du signal causé par le MIG. Cela permettait de se donner un facteur de sécurité de 2 sur le calcul approximatif de la  $DM_{MIG}$  associée à chaque objet: ce facteur évite de perdre le signal dans le bruit.

À la section 5.2.4, on a décrit la procédure utilisée pour récupérer les quelques milliards de données du vecteur des intensités auto-corrélées par blocs (ex., 760) de 50 millions de données. Pourtant, ici, on utilise une autre alternative qui consiste à récupérer uniquement les données à tous les  $N$  points (pas de  $N$  points dans `h5read()`). Cette nouvelle approche d'échantillonnage ne nécessite pas de boucles (ex. 760) pour récupérer les données: toutes les données récupérées sont placées dans un seul vecteur Matlab (en mémoire vive). De plus, elle se justifie par le fait que la largeur minimale du signal radio recherché (le nombre minimal de points) est fixée par l'effet de lissage venant de la dispersion causée par le plasma intergalactique ionisé (hydrogène) le long du trajet des rayons issus de la radiosource cosmique. Voilà pourquoi, le pas  $N$  choisi correspond à la longueur du lissage équivalant à la moitié du MIG, soit un facteur de sécurité de 2 par rapport à la largeur minimale du signal recherché. Ainsi donc, avec ce nouveau pas  $N$  d'échantillonnage (ex.,  $N = 7000$  points), on a beaucoup moins de données à récupérer et à traiter (mémoire vive limitée de Matlab) comparativement au cas où on avait un pas d'échantillonnage unitaire. À titre d'exemple, pour le quasar 3C 345, dont le redshift donne une dispersion de  $\sim 14000$  points (voir tableau 6.1), on récupère avec Matlab un nouveau vecteur comprenant seulement  $\sim 5.5$  millions ( $\sim 38 \times 10^9 / 7000$ ) d'intensités auto-corrélées au lieu de  $\sim 38$  milliards (non stockable dans un vecteur Matlab). C'est d'ailleurs à partir de ce nouveau vecteur qu'on pourra éventuellement évaluer le niveau de bruit, et par conséquent, extraire les candidats possibles.

### 5.3.2 Évaluation du bruit dans l'autocorrélation d'intensité

Tel que déjà mentionné à la section 5.2.6, la présence indésirable de signaux artificiels non périodiques génère des profils d'autocorrélation non uniformes. C'est pourquoi, afin de bien évaluer le bruit pratiquement gaussien (voir section 5.1.5), et aussi de pouvoir appliquer correctement nos techniques de détections de signaux (voir section 5.3.3), il est d'abord nécessaire de rendre le profil d'autocorrélation le plus uniforme possible, c'est-à-dire de le rectifier afin qu'il devienne semblable au profil tronqué d'autocorrélation d'un bruit gaussien (voir figure 5.9). Une façon simple et

pratique de rectifier le profil d'autocorrélation consiste à utiliser la régression linéaire (voir annexe C.4). À cette fin, pour tenir compte des changements de pente (diverses sections de droite) causés par des signaux artificiels dans l'autocorrélation de certaines radiosources, on subdivise d'abord le nouveau vecteur, soit celui obtenu par pas de  $N$  points (où  $N$  correspond à un lissage valant la moitié du MIG) dans  $h5read()$ , en  $n$  blocs (ex., 500) de  $M$  points (ex.,  $M = 5.5 \times 10^6 / 500 = 11000$  points). La taille  $n$  choisie du bloc doit être suffisamment petite pour que le bruit soit considéré comme blanc (pas de diminution monotone de l'amplitude), mais assez grande afin qu'elle soit statistiquement significative. Notez que l'utilisation de la méthode courante du *fit* polynomial sur la moyenne permet également de rectifier le profil d'autocorrélation. Cependant, compte tenu des nombreux changements de pente qui ont tendance à diminuer la justesse du *fit* polynomial, on a préféré utiliser une méthode rectifiant le profil d'autocorrélation en soustrayant des sections de droite obtenus dans  $n$  blocs de  $N$  points. Par suite, pour chaque bloc  $i$ , on détermine le vecteur droite de régression ( $mat\_yfit_i$ ) et on calcule le vecteur résidu ( $mat\_r_i$ ) en soustrayant le vecteur droite de régression au vecteur des intensités auto-corrélées ( $mat\_y_i$ ), soit  $mat\_r_i = mat\_y_i - mat\_yfit_i$  (voir annexe C.4). Enfin, puisque le vecteur  $mat\_r_i$  donne le profil d'autocorrélation uniforme (rectifié) voulu, on peut alors évaluer le bruit, dans chacun des blocs  $i$ , à partir de la relation bien connue de l'écart-type ( $\sigma_{mat\_r,i}$ ) d'une distribution gaussienne, qui est donnée par,

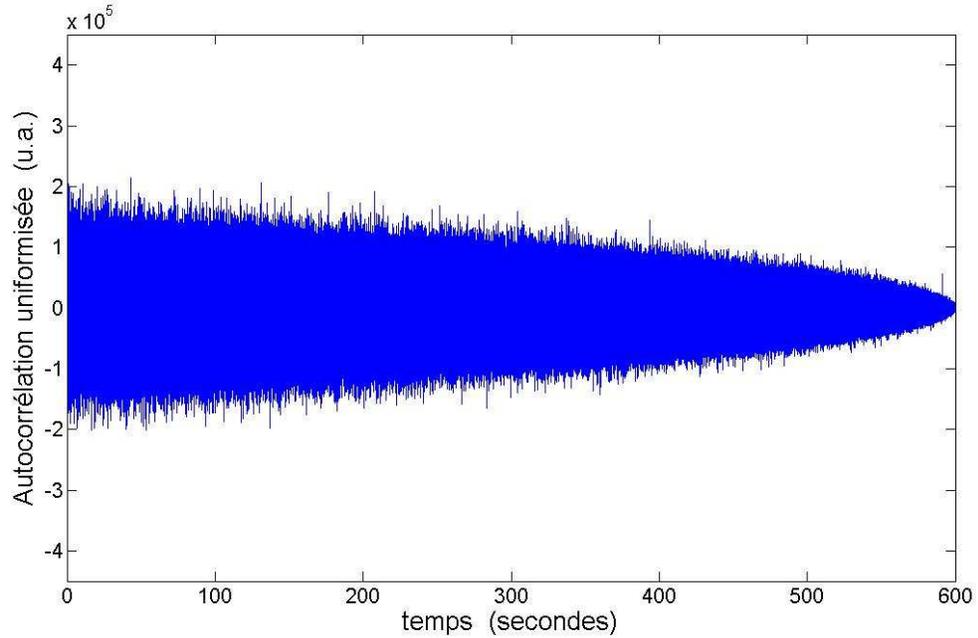
$$\sigma_{mat\_r,i} = \sqrt{\frac{1}{M-1} \sum_{k=1}^M (mat\_r_{i,k} - \overline{mat\_r_i})^2} . \quad (5.4)$$

À titre illustratif, la figure 5.9 indique l'allure du profil rectifié d'autocorrélation d'intensité du quasar 3C 345. Notez l'uniformité de ce nouveau profil comparativement au profil original (voir figure 5.5).

Il est important de mentionner que le fait de rectifier le profil d'autocorrélation permet d'appliquer facilement et précisément notre technique (programme) d'extraction de pics (voir section 5.3.3 et annexe C.4): pour sortir les

pics, il serait plus difficile (et plus imprécis) de parcourir un profil courbé (ex., voir figures 5.5 et 5.7) qu'un profil horizontal (ex., voir figure 5.9). Enfin, comme le signal radio recherché est beaucoup plus étroit que le nouveau vecteur, il n'y a pas de risque que le processus de rectification affecte, voire détruise, le signal. En effet, la largeur du signal recherché dans le profil d'autocorrélation rectifié équivaut à un seul point, correspondant au  $DM_{MIG}$  (ex., 7000 points), parmi les 5.5 millions de points (regroupement de 38 milliards / 7 mille points pour 3C 345). De là, bien que la technique d'évaluation du bruit (voir équation (5.4)) subdivise le profil rectifié en petits blocs (ex., 11000 points), ces blocs demeurent largement plus grand que le signal recherché: le rapport de la largeur du signal sur celle du bloc étant de l'ordre de  $10^{-4}$  (ex., 1 point / 11000 points).

En raison de la petite taille de chaque bloc (ex., 11000 points) comparativement à la taille du profil d'autocorrélation rectifié (ex.,  $5.5 \times 10^6$  points), on comprend bien que l'effet de l'échantillonnage fini, qui produit une diminution monotone de l'amplitude des intensités auto-corrélées de l'origine jusqu'à l'extrémité, devient négligeable, de sorte que le bruit calculé via l'équation (5.4) est quasi blanc dans cette petite région du profil: l'amplitude du bruit est constante dans une boîte de largeur égale à 1/500 du profil d'autocorrélation.



**Figure 5.9.** Profil temporel (temps positifs) uniforme d'autocorrélation d'intensité (en unités arbitraires) de l'objet 3C 345. Le pas utilisé pour l'affichage est égal au lissage de 7000 points ( $109 \mu\text{s}$ ). Ce lissage correspond à environ la moitié de la dispersion intergalactique. Notez la différence avec le profil non uniforme d'origine de la figure 5.5.

### 5.3.3 Étapes de la détection d'un signal

Cette sous-section décrit la façon de mesurer, plus précisément de détecter et d'extraire, un «vrai» signal cosmique, lequel pourrait provenir de candidats aux effets de minilentilles gravitationnelles ou de candidats ayant des variations temporelles périodiques ou non périodiques.

La signature d'un signal cosmique doit se retrouver dans plus d'une station VLBA, idéalement toutes (voir la deuxième étape), tandis que celle d'un signal artificiel se retrouve généralement dans une seule station. Sachant cela, les paragraphes qui suivent décrivent comment on peut détecter des signaux (pics) et comment on peut savoir si les signaux détectés sont de nature cosmique ou de nature artificielle. Cela dit, les difficultés pratiques associées à l'implantation des techniques et des algorithmes (programmes) de détection sont essentiellement de fixer un rapport signal à bruit (S/B) pour l'extraction d'un nombre raisonnable de pics, et après quoi de

vérifier si les pics (signaux potentiels) extraits, selon ce S/B, sont dans au moins deux télescopes radio et dans au moins deux canaux (RCP et LCP) de ces télescopes.

Parmi les deux méthodes appliquées pour détecter et extraire des candidats raisonnables, seules les principales étapes de la méthode retenue, où à vrai dire la plus utilisée, sont décrites dans les paragraphes qui suivent.

La première étape, cruciale pour la détection d'un signal cosmique, mais aussi, inopportunément, d'un signal artificiel (ex., RFI, effet instrumental), consiste à imposer un critère de sélection, basé sur le rapport signal à bruit S/B, permettant d'extraire un nombre limité de vrais signaux possibles (voir les deux autres étapes) parmi les quelques millions de données du profil d'autocorrélation d'intensité uniforme (rectifié). À ce propos, rappelons que ces quelques millions (ex., ~5.5 millions) de données ont été obtenues avec la méthode (la plus employée) utilisant un nouvel échantillonnage dont le pas  $N$  (ex., 7000 points) correspond à la moitié de la  $DM_{MIG}$  (voir section 5.3.1). Rappelons également que la seconde méthode, de pas d'échantillonnage unitaire, requière plusieurs boucles (ex., 760) pour récupérer avec Matlab toutes les dizaines de milliards (ex., ~38 milliards) de données (voir section 5.2.4). (Voilà l'une des raisons pour laquelle cette longue méthode n'est pas décrite dans cette thèse.)

D'une part, le choix d'un tel critère de sélection est basé sur la nécessité de faire un compromis entre deux restrictions diamétralement opposées, à savoir, maximiser le nombre de «vrai» signaux possibles et minimiser le nombre de faux signaux issus de fortes fluctuations statistiques. D'autre part, sachant que le bruit suit approximativement une distribution gaussienne (voir figure 5.4), il va de soi que la valeur du seuil de détection à fixer est choisie en fonction de la théorie de la loi de Gauss. Cela dit, la valeur du seuil (S/B) retenue, donnant un nombre raisonnable de pics à inspecter et à analyser individuellement, a été fixée à 4.0 (voir annexe C.4). En fait, avec cette valeur de S/B, on détecte quelques centaines de pics, généralement entre 100 et 300.. Par exemple, pour le quasar 3C 345, on a obtenu en moyenne (sur

les 16 canaux des deux radiotélescopes) 180 détections. Ce nombre concorde bien avec les  $5.5 \times 10^6$  points  $\times 3.167 \times 10^{-5}$  détection/point = 174 détections prévues par une gaussienne ( $S/B = 4.0$ ). Par ailleurs, la présence de signaux artificiels parasites génère un surplus de fausses détections, qui s'ajoutent aux fausses alarmes statistiques, allant de plusieurs centaines à un millier. Rappelons, en terminant, que si l'on tient compte de l'efficacité de quantification (88.25 %), la limite de détection fixée à 4.0 correspond en réalité à un seuil de 4.53 (voir section 4.5.5).

La seconde étape a pour but de vérifier si le signal cosmique est détecté simultanément dans au moins deux stations VLBA. En théorie, ce signal cosmique devrait être détectable simultanément à tous les radio-observatoires et dans tous les canaux (voir la 3<sup>ième</sup> étape). En pratique, cependant, le dysfonctionnement de certains radiotélescopes (ex., de l'électronique ou des guides d'onde), les problèmes dans la sauvegarde des données (ex., données mal formatées ou corrompues) et la présence de bruit élevé, peuvent limiter la détection d'un signal dans seulement deux télescopes radio (la détection dans un seul radiotélescope n'étant pas retenue). En outre, dans le cas où les signaux sont détectés uniquement dans un seul radiotélescope (parfois deux), il s'agit alors, la plupart du temps, de signaux artificiels indésirables.

Quand le signal détecté lors de l'observation du même objet source se retrouve dans deux stations VLBA ou plus, on impose alors une troisième étape de détection consistant, d'une part, à vérifier la présence du signal dans au moins deux canaux (RCP et LCP) de chaque radiotélescope. Le numéro des canaux (voir tableau A.1) où le signal est présent doit être le même (pour chaque polarisation) dans chacun des télescopes concernés. Qui plus est, la position (le temps  $t$ ) du signal auto-corrélé détecté doit également être la même dans chaque canal. En théorie, tout signal cosmique (et parfois artificiel) détecté devrait s'étendre sur une grande région de la bande passante (512 MHz) d'un radiotélescope, c'est-à-dire que le signal devrait être détecté dans la majorité, idéalement la totalité, des 16 canaux du télescope radio. En pratique, si pour un intervalle de fréquences donné, le rapport signal sur bruit tombe sous la limite de détection de 4.0 (voir la 2<sup>ième</sup> étape), le signal ne sera pas détecté

dans les canaux dont la bande passante est comprise dans cet intervalle de fréquences. Aux fins d'illustration seulement, prenons l'exemple simpliste où un signal serait détecté ( $S/B > 4.0$ ) dans la moitié supérieure de la largeur de bande, soit 8384 – 8512 MHz (canaux 1 à 8), mais ne le serait pas ( $S/B < 4.0$ ) dans la moitié inférieure, soit 8256 – 8384 MHz (canaux 9 à 16). D'autre part, un autre critère de sélection, baptisé critère de priorité, est imposé afin d'ordonner les investigations. En effet, comme le signal d'origine cosmique est généralement présent dans plusieurs canaux et dans les deux polarisations circulaires (RCP et LCP), on impose un critère qui ordonne la priorité des investigations, comme les inspections individuelles des profils d'autocorrélation, d'après le nombre de canaux où des candidats ( $S/B > 4.0$ ) ont été détectés: allant de la détection la plus prioritaire dans les seize canaux jusqu'à celle la moins prioritaire dans seulement deux canaux.

## 6 Simulations numériques, analyse statistique des résultats et estimation de la densité cosmologique de minilentilles

### 6.1 Mise en contexte et problématique

Rappelons que le but essentiel de cette thèse de doctorat est le développement d'une technique numérique et l'écriture de logiciels innovateurs (Matlab et C++) afin de rechercher des événements de minilentilles gravitationnelles dans des observations radio de noyaux actifs de galaxies (NAGs) et/ou de faibles variations temporelles rapides. Ceci étant acquis, le but de l'analyse statistique simplifiée de ce chapitre se limite à faire une estimation (approximative, voire grossière) de la densité cosmologique d'objets compacts intergalactiques, dans l'intervalle de masse,  $10 - 10^7$  masses solaires (voir section 6.2).

L'estimé de la densité cosmique de minilentilles obtenue, soit  $\Omega_{l,0} \leq 0.2 - 0.3$  (voir section 6.3.4), via l'analyse statistique simplifiée donne une trop grande incertitude (voir figure 6.7). Plus précisément, les barres d'incertitude (intervalles de confiance) obtenues sont trop larges pour nous permettre de fixer une contrainte statistique stricte sur  $\Omega_{l,0}$ . Il fallait s'y attendre considérant que le très faible échantillon-test de 22 objets n'était pas assez statistiquement significatif.

Les trop grandes barres d'incertitude expliquent, en grande partie, pourquoi on n'approfondit pas l'analyse en considérant des modèles réalistes plus complexes de minilentilles gravitationnelles. Une analyse statistique plus poussée des minilentilles gravitationnelles exigerait une modélisation qui tiendrait compte de divers paramètres physiques comme le paramètre d'impact et l'angle d'approche de la minilentille par rapport au faisceau lumineux. Il faudrait aussi considérer qu'un effet de minilentille détecté individuellement ne permet pas de quantifier une seule minilentille, ce que l'analyse simplifiée de ce chapitre suppose, mais plutôt de comparer le délai temporel trouvé (ou simulé) à une distribution possible de minilentilles, à partir de laquelle on

pourrait fixer des limites statistiques. Enfin, il faudrait tenir compte des diverses inhomogénéités de densité le long de la ligne de visée: on suppose que les minilentilles compactes sont uniformément distribuées, tandis que la matière dans l'univers est plutôt distribuée de façon inhomogène.

En somme, il ne sert à rien de faire une analyse statistique très rigoureuse utilisant une modélisation sophistiquée des systèmes astrophysiques de minilentilles puisque la principale source d'incertitude restera tout de même le nombre très limité de radiosources traitées: l'échantillon-test de 22 objets n'étant pas assez statistiquement significatif (statistique de petit nombre).

## 6.2 Simulations numériques de délais temporels

Le but des simulations consiste à montrer l'efficacité de détection ( $S/B \geq 4$ ) d'un signal (délai temporel) par notre technique d'autocorrélation d'intensité dans l'intervalle de masse applicable de  $\sim 10$  à  $\sim 10^7$  masses solaires (voir section 6.2). Par exemple, ce signal pourrait être causé par la présence d'une minilentille gravitationnelle extragalactique le long du parcours reliant une source radio à image simple à l'observateur (voir section 6.4). Pour cela, on a effectué, pour différentes masses, qui se traduisent par différents délais temporels  $\tau$  (voir équation (3.14)), une centaine de simulations numériques. Ces simulations sont basées sur un modèle simple qui approxime seulement la physique réelle complexe des systèmes de lentilles gravitationnelles. Le modèle utilisé suppose une géométrie simple où seulement deux faisceaux dominants (la contribution de tous les faisceaux additionnels étant négligée) d'intensité identique ou comparable interfèrent, tandis que les systèmes (modèles) réels de lentilles peuvent avoir de multiples faisceaux d'intensité inégale (Borra 2008). Notez que le fait d'avoir deux faisceaux d'intensité différente n'a aucune incidence sur la valeur du délai temporel (positionnement dans le profil d'autocorrélation des simulations). En revanche, l'amplitude du signal diminue lorsque l'intensité des deux faisceaux est différente.

Nos simulations sont comparables à celles de Splliar (1993) et aux expériences publiées dans Borra (2011). De fait, dans Borra (2011), le délai temporel (DPO) est imposé par la présence d'un interféromètre divisant le faisceau initial en deux faisceaux identiques (l'un étant une réplique retardée de l'autre), tandis que dans nos simulations le délai est artificiellement imposé en retardant (avançant) un champ électrique (faisceau) par rapport à l'autre.

Les paragraphes qui suivent décrivent les quatre principales étapes de simulations de délais temporels, qui pourraient être causés par des minilentilles cosmologiques, et dans l'hypothèse d'une géométrie limitée à deux faisceaux d'intensité identique ou comparable. À titre complémentaire, les lignes commentées du fichier de soumission de l'annexe E résume bien les étapes et les paramètres d'entrée des simulations (pour la simulation correspondant aux figures 6.4 et 6.5).

- 1) Imposer un délai temporel artificiel entre les champs électriques  $E(t)$  des deux faisceaux, et non pas entre les intensités  $I(t)$  puisque les données brutes reçues sont des champs électriques. Au niveau algorithmique, cette première étape peut s'abréger par l'instruction simple suivante,

$$I(t) = (f * E(t) + (1 - f) * E(t + \text{nbr\_pts\_délai}))^2.$$

Ici,  $E(t + \text{nbr\_pts\_délai})$  est la réplique de  $E(t)$  décalée de  $\text{nbr\_pts\_délai}$  points,  $f$  ( $1 - f$ ) est la fraction (entre 0 et 1) du champ électrique total du premier (second) faisceau et  $I(t)$  est l'intensité du faisceau recombinaison résultant. Par contre, l'implantation du code C++ requis pour les simulations, code déjà sophistiqué pour le traitement d'un seul signal, se complique davantage puisqu'ici il faut faire l'addition de deux signaux: on traite simultanément deux fichiers d'entrée identiques, simulant deux faisceaux comparables, à des temps différents. Qui plus est, comme l'autocorrélation d'intensité de deux fichiers nécessite deux fois plus de mémoire, on doit, afin de respecter la limite de 1024 cœurs allouée par le CAR de Calcul Canada, restreindre le

traitement à environ la moitié des 600 secondes d'observation du NRAO. Notez qu'il faudrait plus de deux mille cœurs de calcul pour exécuter chaque autocorrélation simulée sur 600 secondes.

- 2) Simuler l'effet de l'espace intergalactique (voir section 2.2) en effectuant un lissage de  $I(t)$  dont le nombre de points correspond à environ la dispersion intergalactique  $DM_{MIG}$  entre la radiosource et l'observateur (voir tableau 6.1). Le lissage des intensités transforme les deux valeurs possibles d'intensité (voir étape suivante) en un continuum de valeurs possibles, qui améliore la précision requise aux traitements. Sous forme de pseudo-code, cette étape se résume simplement par,

$I(t)_{\text{lissée}} = \text{lissage}(I(t), DM_{MIG} \text{ points})$ .

- 3) Générer du bruit artificiel similaire aux sources dominantes de bruit environnemental (sol, atmosphère et récepteur) présentes lors des observations du NRAO.

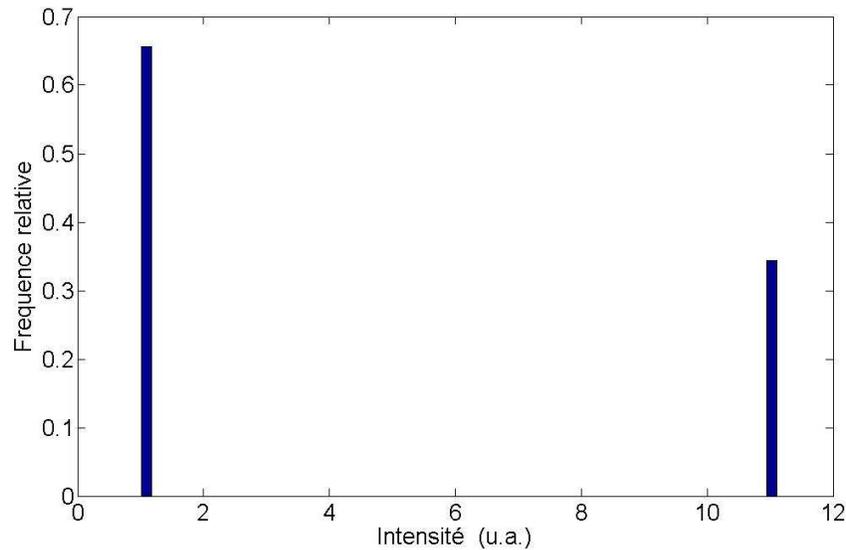
À la section 4.5.3, on a déjà dit qu'il n'y avait que quatre valeurs possibles de champs électriques quantifiés,  $\pm 1$  et  $\pm 3.3359$ , lesquelles se traduisent par seulement deux valeurs possibles des intensités, 1 et 11.1282 (avant le lissage). Sachant cela, pour trouver la fréquence relative de ces deux intensités, on détermine la distribution moyenne d'intensité pour chacune des 22 sources radio. En réalité, pour chaque objet, il y a 4 stations/canal X 8 canaux/polarisation X 2 polarisations/station = 64 distributions d'intensité possibles par objet (22 X 64 = 1408 au total). En pratique, il est préférable de choisir les objets les plus faibles (ex.,  $\sim 10$  –  $\sim 100$  mJy). Voilà pourquoi, les exemples de simulations (les figures 6.2 à 6.6) discutés dans cette section ont été reproduits pour le deuxième objet le plus faible observé, le B1030+074 dont le flux moyen  $S$  n'est que de 203 mJy. (On a choisi cet objet au lieu du plus faible objet, soit le quasar Q0957+561 avec  $S = 29$  mJy, car la durée

d'observation de Q0957+561 n'est que de 180 secondes au lieu de 600 s pour les autres objets.)

L'histogramme de la figure 6.1 indique un exemple de distribution (moyennée sur les seize canaux des quatre radiotélescopes, soit sur 64 canaux) des deux intensités possibles (avant le lissage) pour la radiosource B1034+074. Notez que la distribution est quasi identique pour les 21 autres sources radio. En effet, pour les deux objets les plus brillants, soit 3C 273 et 3C 279 dont le flux moyen est près de 20 Jy (voir tableau A.2), la fréquence (proportion) de  $I = 11.1282$  demeure inférieure à 35 %. Il faut dire qu'avec ces objets brillants, on s'attendait plutôt d'avoir une plus grande fréquence de  $I = 11.1282$ . Cela dit, on génère artificiellement l'intensité du bruit local avec la ligne d'instructions qui suit,

$I(t)_{\text{bruit\_local\_simulé}} = \text{générateur\_aléatoire} (p*1 + (1-p)*11.1282),$

où  $p$  et  $(1-p)$  représentent respectivement la proportion choisie de  $I = 1$  ( $E = \pm 1$ ) et de  $I = 11.1282$  ( $E = \pm 3.3359$ ), tandis que la fonction `générateur_aléatoire` génère aléatoirement les deux (quatre) valeurs discrètes 1 ( $\pm 1$ ) et 11.1282 ( $\pm 3.3359$ ) selon la fraction  $p$  choisie. En clair, le vecteur,  $I(t)_{\text{bruit\_local\_simulé}}$ , contient, à un facteur d'amplitude près donné par  $S/B$  (voir étape 4), le bruit local simulé à partir de seulement deux valeurs discrètes d'intensité (pas de continuum de valeurs).



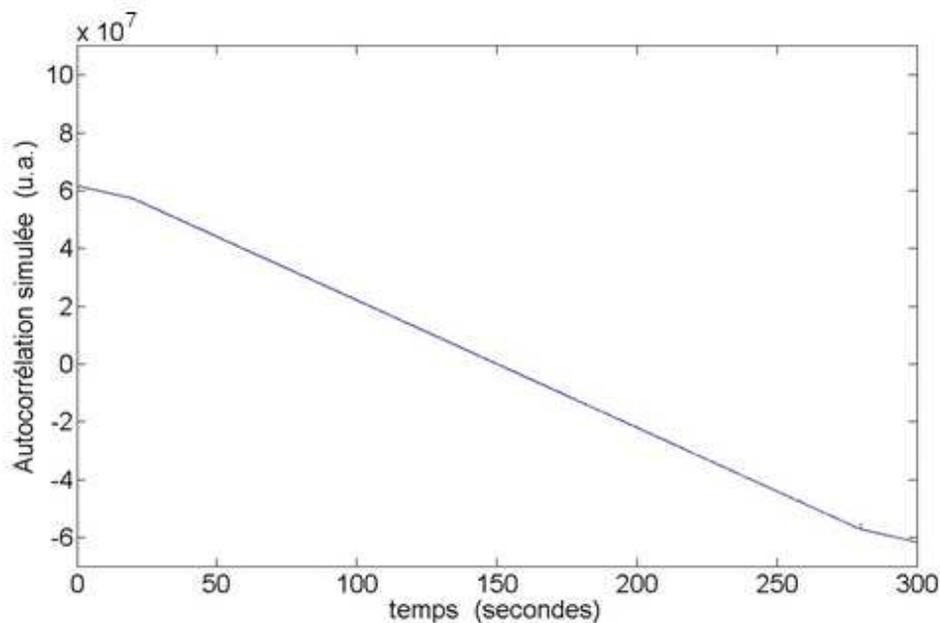
**Figure 6.1.** Exemple d’histogramme donnant la fréquence relative (moyennée sur les seize canaux des quatre stations, soit sur 64 canaux) de chacune des deux intensités possibles (en unités arbitraires) de l’objet B1030+074. 65.6 % des intensités valent 1 et 34.4% valent 11.1282.

- 4) Additionner l’intensité du bruit environnemental artificiel à celle de la radiosource simulée subissant un effet de minilente (délai temporel imposé).  
Sous forme algorithmique, ceci se traduit par,

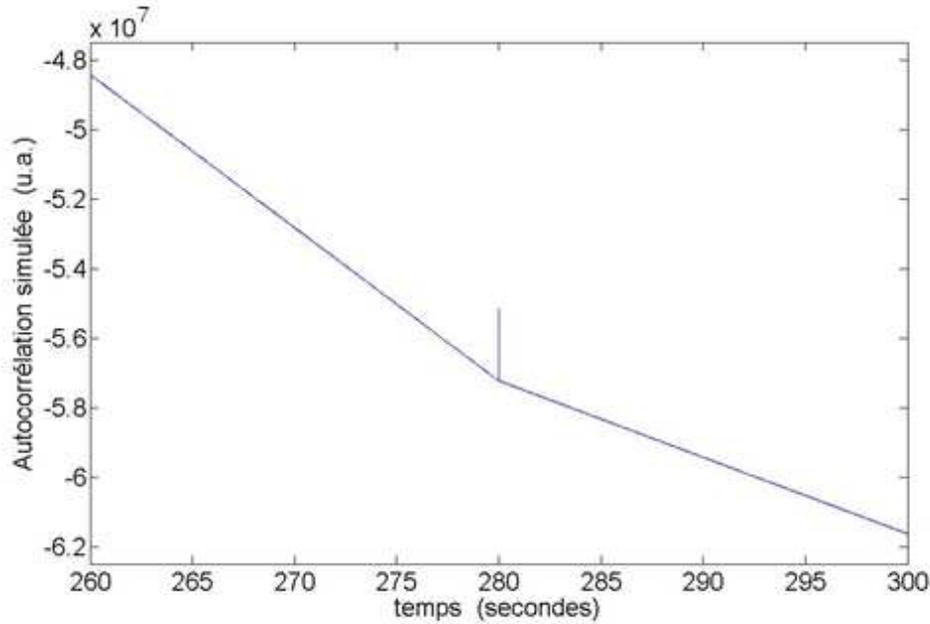
$$I(t)_{\text{lissée\_bruit}} = I(t)_{\text{lissée}} + (S/B) * I(t)_{\text{bruit\_local\_simulé}}.$$

Ici,  $I(t)_{\text{lissée}}$  est l’intensité du signal reçu de l’objet source (le lissage simulant l’étalement du signal causé par le MIG) et  $S/B$  est le rapport signal à bruit moyen. Ce rapport est calculé en divisant l’amplitude moyenne de l’intensité du signal de l’objet,  $\langle I(t)_{\text{lissée}} \rangle$ , par l’amplitude moyenne de l’intensité du bruit local simulé,  $\langle I(t)_{\text{bruit\_local\_simulé}} \rangle$  (les chevrons  $\langle \rangle$  indiquant une valeur moyenne). Par exemple, pour B1030+074 (dans les mêmes conditions que celles des figures 6.4 et 6.5), on obtient une intensité moyenne du bruit d’environ 11, soit  $(19.422408 - 8.314869 = 11.107539)$ , voir la fin de l’annexe E), ce qui donne un  $S/B$  moyen d’environ 1.75  $(19.422408 / 11.107539)$ .

Les figures 6.2 et 6.3 illustrent un exemple de simulation numérique de l'autocorrélation d'intensité pour deux faisceaux (champs électriques) comparables ( $f = 2/3$ ) artificiellement séparés de 280 secondes. Par exemple, ce délai temporel pourrait être causé par la présence d'un trou noir d'environ sept millions de masses solaires (voir équation (3.14)). On a fixé un délai temporel presque aussi grand que la durée (demi-durée) d'observation (300 s) afin que l'amplitude du signal auto-corrélé soit la plus faible possible. Dans cet exemple, pour simuler le bruit local, la proportion  $p$  choisie (0.656) pour le générateur aléatoire de bruit est identique à la distribution d'intensité moyenne de l'objet B1030+074 (voir figure 6.1). La figure 6.3 (agrandissement de la figure 6.2) montre que même pour un signal situé près de l'extrémité du profil d'autocorrélation, la signature du signal (délai temporel) est bien distinguable (S/B élevé).

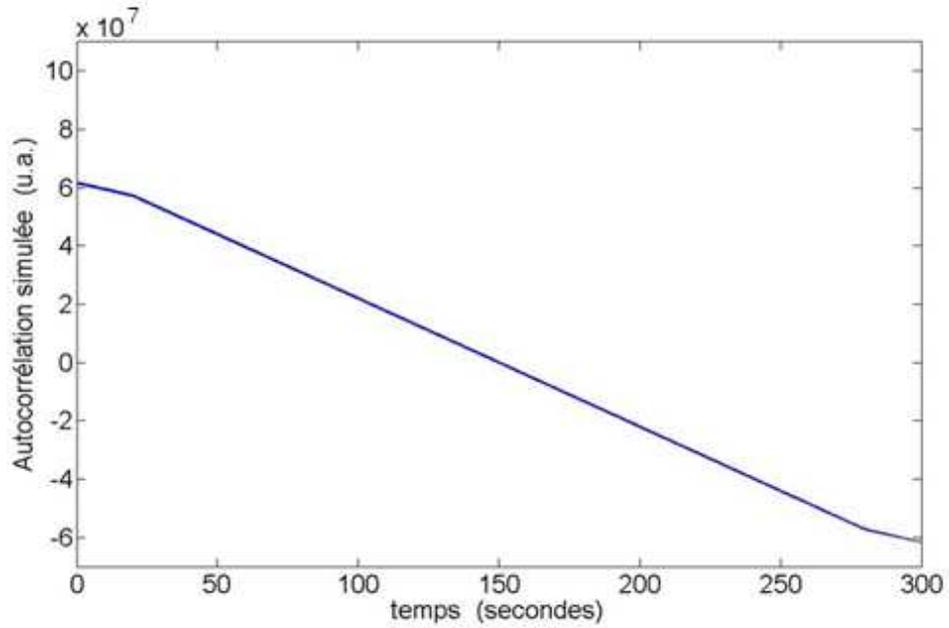


**Figure 6.2.** Exemple de profil temporel d'autocorrélation d'intensité (u. a.) obtenu en simulant un délai temporel de 280 secondes (ex., minilentille de  $\sim 10^6 - 10^7$  masses solaires) et du bruit environnemental (sol, atmosphère et récepteur). Le rapport des intensités (champs électriques) des deux faisceaux est de 4:1 (2:1). La proportion  $p$  choisie pour simuler le bruit local est la même que la fréquence relative des deux intensités possibles, soit  $p = 0.656$ . L'agrandissement de la figure 6.3 permet de bien distinguer le pic à  $t = 280$  s. L'objet choisi est le B1030+074 ( $S = 203$  mJy).

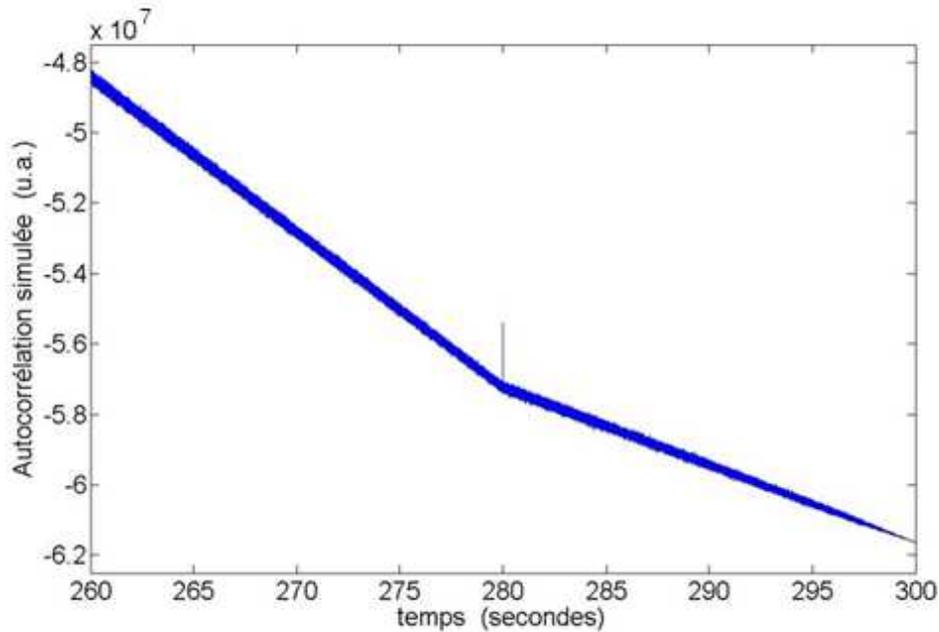


**Figure 6.3.** Agrandissement de la figure 6.2. Bien que délai de 280s soit près de l'extrémité du profil d'autocorrélation, on distingue clairement la signature du signal.

L'exemple de simulation menant aux figures 6.4 et 6.5 reprend les conditions de la simulation précédente (B1030+074,  $f = 0.667$  et  $\tau = 280$  s), mais en fixant le niveau de bruit local pratiquement au maximum, soit  $p = 0.05$  de  $I = 1$  et  $(1-p) = 0.95$  de  $I = 11.1282$ . La figure 6.5 (agrandissement de la figure 6.4) indique que malgré le bruit élevé, le signal (pic) demeure distinguable. Plus précisément, le rapport signal à bruit du pic est  $\sim 5$ , ce qui satisfait le premier critère de sélection d'extraction d'un signal cosmique potentiel, soit  $S/B \geq 4.0$  (voir section 5.3.3). Par contre, pour des rapports de champs (des intensités) supérieurs à  $\sim 2:1$  ( $\sim 4:1$ ), l'intensité du pic est diminuée au point où le  $S/B < 4.0$ . Dans ce cas, on ne peut plus distinguer sans ambiguïté le pic auto-corrélé des fortes fluctuations statistiques de bruit.

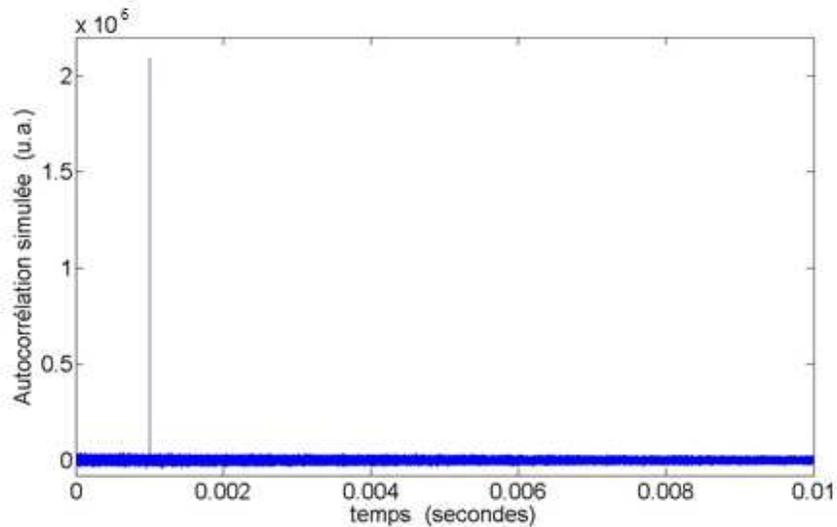


**Figure 6.4.** Exemple de profil temporel d'autocorrélation d'intensité (u. a.) obtenu en simulant un délai temporel de 280 secondes (ex., minilentille de  $\sim 10^6 - 10^7$  masses solaires) et du bruit environnemental fort (sol, atmosphère et récepteur). Le rapport des intensités (champs électriques) des deux faisceaux est de 4:1 (2:1). La proportion  $p$  choisie pour simuler un bruit local fort est  $p = 0.05$  ( $1-p = 0.95$ ). Sans l'agrandissement de la figure 6.5, le pic à  $t = 280$  s est à peine perceptible. L'objet choisi est le B1030+074 ( $S = 203$  mJy).



**Figure 6.5.** Agrandissement de la figure 6.4. Bien que délai de 280s soit près de l'extrémité du profil d'autocorrélation et que le niveau de bruit est quasi maximal ( $1-p = 0.95$ ), l'amplitude du signal est  $\sim 5$  fois celle du bruit.

L'exemple de simulation de la figure 6.6 reprend les conditions de la simulation précédente (B1030+034,  $f = 2/3$  et bruit fort), mais pour un très court délai artificiel d'une milliseconde. Ce délai temporel est du même ordre de grandeur que la dispersion temporelle intergalactique moyenne de 0.455 ms (voir tableau 6.1). Qui plus est, il correspond à une masse ( $25 M_{\odot}$  selon l'équation (3.14)) du même ordre de grandeur que la masse minimale applicable d'une dizaine de masses solaires (voir section 6.2). Pour ce délai minimal (masse minimale), la figure 6.6 montre que la signature du signal sort nettement du bruit.



**Figure 6.6.** Exemple des 10 premières millisecondes du profil temporel d'autocorrélation d'intensité (u. a.) obtenu en simulant un très bref délai temporel de 1 ms (ex., minilentille de  $\sim 10$  masses solaires) et du bruit environnemental fort (sol, atmosphère et récepteur). Le rapport des intensités (champs électriques) des deux faisceaux est de 4:1 (2:1). La proportion  $p$  choisie pour simuler un bruit local fort est  $p = 0.05$  ( $1-p = 0.95$ ). Notez que la signature du signal sort clairement du bruit. L'objet choisi est le B1030+074 ( $S = 203$  mJy).

En fin de compte, dans l'hypothèse d'une géométrie à deux faisceaux d'intensité identique ou comparable, les simulations numériques ont montré que la détection d'un signal par l'autocorrélation d'intensité est efficace ( $S/B \geq 4$ , voir section 5.3.3) pour l'intervalle de masse applicable,  $10 \leq M / M_e \leq 10^7$ . Pour les délais temporels, de l'ordre de  $\sim 10^2$  secondes, la technique de détection est efficace jusqu'à des rapports de champs (des intensités) de  $\sim 2:1$  ( $\sim 4:1$ ), même davantage pour de plus faibles niveaux de bruit. Pour de petits délais,  $< 0.01$  s – 10 s, la méthode demeure efficace jusqu'à des rapports de champs (des intensités) de  $\sim 10:1$  ( $\sim 100:1$ ), voire davantage pour les plus courts délais de quelques centaines de microsecondes (voir tableau 6.1) atteignables dans le radio: de tels délais donnent des signaux auto-corrélés de plus grandes amplitudes (signaux forts près de l'origine).

### 6.3 Intervalle de masse applicable

Pour évaluer la borne supérieure de l'intervalle de masse applicable pour les minilentilles, on considère un délai temporel équivalant à la durée d'observation d'un objet source, soit généralement 600 secondes (voir tableau A.3). De ce délai maximal, on tire de la relation (3.14) une masse limite de  $\sim 10^7$  masses solaires. Sachant que l'amplitude du signal détecté diminue au fur et à mesure que l'on s'approche de l'extrémité du profil de l'autocorrélation (ex., voir figure 5.3), le pic correspondant à ce délai de 600 secondes serait évidemment plus faible. Malgré cela, les simulations (section 6.2) montrent que ce pic éloigné serait visible à travers le bruit.

Afin d'estimer la masse minimale, on considère, cette fois-ci, un délai équivalant à l'étalement temporel dû au plasma intergalactique (PIG). Cela dit, le signal (ex., une impulsion) le plus étroit possible aurait théoriquement une largeur de 15.625 ns, qui correspond au temps d'échantillonnage du NRAO. Mais, après avoir traversé le PIG pour se rendre jusqu'au radio-observatoire, le signal est étalé d'une largeur qui correspond au délai de dispersion intergalactique de la radiosource (voir tableau 6.1). Dès lors, en utilisant le délai temporel moyen de dispersion (455  $\mu$ s) des 22 sources radio, on tire de l'expression (3.14) une limite inférieure moyenne de masse d'environ 10 masses solaires.

En résumé, notre technique d'autocorrélation d'intensité permet de détecter des minilentilles dans l'intervalle de masse,  $10 \leq M / M_{\odot} \leq 10^7$ . La limite inférieure de masse étant due à la présence du PIG.

**Tableau 6.1.** Redshift, délai temporel et masse minimale résultant du milieu intergalactique  $DM_{MIG}$  des 22 radiosources.

Objet	Redshift	Délai dû à la $DM_{MIG}$ (points)	Délai dû à la $DM_{MIG}$ ( $\mu s$ )	Masse MIN due à la $DM_{MIG}$ ( $M_e$ )
PKS 1622-253	0.786	18 641	291	7.3
4C +01.02	2.099	42 134	658	16.5
0957+561AB	1.413	30 664	479	12
OJ 287	0.306	7 913	124	3.1
3C 273	0.1583	4 225	66	1.7
3C 279	0.536	13 263	207	5.2
OT 081	0.322	8 300	130	3.3
3C 345	0.59347	14 536	227	5.7
NRAO 530	0.902	21 008	328	8.2
PKS 1830-211	2.507	48 356	756	18.9
B0218+357	0.96	22 164	346	8.7
B1030+074	1.535	32 811	513	12.8
PKS 0528+134	2.07	41 677	651	16.3
4C +73.18	0.302	7 816	122	3
PMN 0414+0534	2.639	50 304	786	19.7
PMN 0134-0931	2.225	44 096	689	17.2
B1422+231	3.62	63 680	995	24.9
4C +38.41	1.813	37 504	586	14.7
NRAO 150	1.52	32 576	509	12.7
PKS 2134+004	1.932	39 488	617	15.4
3C 454.3	0.859	20 096	314	7.9
3C 418	1.686	38 720	605	15.1

Délai moyen dû à la  $DM_{MIG}$ : 455 microsecondes

Masse minimale moyenne due à la  $DM_{MIG}$ : 11.4  $M_e$

## 6.4 Minilentilles distribuées le long de la ligne de visée

### 6.4.1 Équations de la profondeur optique et des distances de diamètre

Une quantité utile pour les calculs de probabilités d'événements de lentilles gravitationnelles est la profondeur optique, notée  $\tau$  (Schneider 1993, ci-après S93). La profondeur optique donne la probabilité d'un événement de lentille et/ou le nombre moyen de lentilles le long d'une ligne de visée (LDV). Dans la limite des faibles profondeurs optiques, on peut directement utiliser  $\tau$  pour estimer la probabilité que survienne un événement de minilentille. En fait, pour de faibles profondeurs optiques ( $\tau < 0.01 - 0.1$ ), on peut faire l'hypothèse qu'une seule minilentille à la fois contribue significativement le long de la LDV (S93; Dalcanton et al. 1994; Zackrisson & Riehm 2007; Holopainen et al. 2008). En effet, les événements de minilentilles multiples, où il y a des chevauchements, sont statistiquement rares, particulièrement pour de faibles valeurs de densité cosmique nette de minilentilles. Par exemple, dans l'hypothèse où la probabilité d'avoir plusieurs minidéfecteurs à la fois suit une distribution de Poisson, on obtiendrait, pour une profondeur optique de 0.1, une probabilité d'obtenir deux minilentilles à la fois dans le faisceau lumineux de  $0.1^2 e^{-0.1}/2 = 0.0045$  (Vietri & Ostriker 1983).

La probabilité différentielle  $d\tau$  d'avoir un effet de minilentille sur une partie  $dz_1$  de la ligne de visée vaut (Turner et al. 1984, ci-après TOG; Fukugita et al. 1992),

$$d\tau = n(z_1, M) \frac{dV}{dz_1} dz_1 = n(z_1, M) \sigma(z_1, z_S, M) \frac{cdt}{dz_1} dz_1, \quad (6.1)$$

où  $n(z_1, M)$  et  $\sigma(z_1, z_S, M)$  représentent respectivement le nombre de densité de minilentilles et la section efficace d'un minidéfecteur de masse  $M$ ,  $dV$  étant l'élément de volume d'un tube fictif de rayon  $R_E = D_l \theta_E$  entourant la LDV, tandis que  $z_1$  et  $z_S$  sont respectivement les redshifts de la minilentille et de la source. Ainsi donc,

la section efficace du tube associée à un événement de minilente est donnée par (Press & Gunn 1973; Fukugita et al. 1992; S93; Zackrisson & Riehm 2007),

$$\sigma(z_1, z_S, M) = \pi R_E^2 = \frac{4\pi G M D_{is} D_1}{c^2 D_s}. \quad (6.2)$$

Avec l'hypothèse, souvent utilisée dans la littérature, que les minilentes sont uniformément distribuées le long de la LDV, et avec une densité comobile constante dans l'univers homogène de Friedmann-Lemaître, il vient que (Press & Gunn 1973; TOG; Kassiola & Kovner 1991; Fukugita et al. 1992; S93; Zackrisson & Riehm 2007),

$$n(z_1, M) = n_{1,0} (1+z_1)^3 = \Omega_{1,0} \frac{3H_0^2}{8\pi G M} (1+z_1)^3, \quad (6.3)$$

avec  $n_{1,0}$  et  $\Omega_{1,0}$  représentant respectivement le nombre et le paramètre de densité cosmique de minilentes au temps présent.

L'expression différentielle (6.1) ne dépend pas de la masse des minilentes, car le changement de taille de  $\sigma$  avec la masse est exactement compensé par le changement de  $n$  (Dalcanton et al. 1994; Zackrisson & Riehm 2007). Autrement dit, la longueur du parcours le long de la LDV où se produit un événement de minilente, soit  $(n\sigma)^{-1}$ , est indépendante de la masse de la minilente (Wilkinson et al. 2001).

Pour l'univers, présentement le plus plausible, sans courbure ( $k = 0$ ) et avec une constante cosmologique  $\lambda_0$ , le terme  $c dt/dz_1$  de l'équation (6.1) est donné par (Zackrisson & Riehm 2007),

$$\frac{c dt}{dz_1} = \frac{R_0}{(1+z_1)} \left[ \Omega_0 (1+z_1)^3 + (1 - \Omega_0) \right]^{-1/2} \quad (R_0 = c/H_0 \text{ étant le rayon de Hubble}), \quad (6.4)$$

où  $\Omega_0 = 1 - \lambda_0$  est le paramètre de densité de matière cosmique. De là, en combinant les quatre équations précédentes, l'intégrale donnant la profondeur optique totale s'écrit comme suit (Zackrisson & Riehm 2007; Holopainen et al. 2008),

$$\tau = \frac{3\Omega_{1,0}}{2R_0} \int_0^{z_s} dz_1 (1+z_1)^2 \left[ \Omega_0 (1+z_1)^3 + (1-\Omega_0) \right]^{-1/2} \frac{D_1 D_{1s}}{D_s}, \quad (6.5)$$

où, en utilisant l'approximation du faisceau plein (TOG; Kayser et al. 1997), la distance de diamètre angulaire, notée  $D$ , entre deux redshifts,  $z_i$  et  $z_j$ , est donnée par (Fukugita et al. 1992; Zackrisson & Riehm 2007),

$$D(z_i, z_j) = \frac{R_0}{(1+z_j)} \int_{z_i}^{z_j} dz_1 \left[ \Omega_0 (1+z_1)^3 + (1-\Omega_0) \right]^{-1/2}. \quad (6.6)$$

#### 6.4.2 Rappel de la loi binomiale

La loi des événements rares de Poisson est souvent employée pour l'analyse probabiliste et statistique d'événements de lentilles (ex., Canizares 1982; Kassiola et al. 1991; S93; Dalcanton et al. 1994; Zackrisson & Bergvall 2003). Par contre, en raison du petit échantillon de radiosources traitées,  $n = 22$ , on devra plutôt appliquer la loi binomiale. Dans les faits, pour bien approximer une binomiale par une Poisson, il faudrait avoir un échantillon tel que  $n \geq 30 - 50$  objets et  $np < 5 - 10$  (Bélisle 2009).

Considérons d'abord une expérience aléatoire avec seulement deux résultats possibles, soit le résultat «succès» survenant avec une probabilité  $p$  et le résultat «échec» survenant avec une probabilité  $1 - p$ . Une telle expérience aléatoire est baptisée épreuve de Bernoulli avec une probabilité de réussite ou proportion, notée  $p$ . Maintenant, considérons  $n$  épreuves de Bernoulli identiques et indépendantes les unes des autres et ayant toutes la même probabilité de réalisation  $p$ . Désignons par  $X$ , la variable aléatoire égale au nombre de réussites parmi ces  $n$  épreuves. Les valeurs possibles de  $X$  sont les entiers  $0, 1, 2, \dots, n$ . En théorie des probabilités, on dit alors

que la variable aléatoire  $X$  suit une loi binomiale, notée  $B(n, p)$ , de paramètres  $n$  et  $p$ . Dans ce cas, la loi de probabilité de la variable  $X$ , qui à chaque issue associe  $k$  succès, s'écrit comme suit,

$$P(X = k) = \begin{cases} \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} & \text{si } k \in \{0, 1, 2, \dots, n\}, \\ 0 & \text{autrement.} \end{cases} \quad (6.7)$$

### 6.4.3 Hypothèses simplificatrices et calculs des profondeurs optiques pour le modèle plat avec $\Omega_0 = 0.3$ et $\lambda_0 = 0.7$

Pour être en mesure d'appliquer la loi binomiale à notre analyse statistique, on doit associer les 22 objets observés aux 22 épreuves indépendantes et identiques de Bernoulli, de même que la proportion  $p$  à la probabilité que se produise un événement de minilentille, soit plus simplement, la probabilité d'avoir une minilentille en se rappelant l'hypothèse d'une seule minilentille à la fois par événement. Or, l'application rigoureuse de cette loi requiert un échantillon idéal contenant 22 objets (hypothétiques) de redshifts identiques (épreuves identiques) et ayant tous le même paramètre  $p$ . Ceci étant acquis, pour se conformer le mieux possible aux exigences de la loi binomiale, on approxime notre échantillon-test réel, composé d'objets de redshifts  $z_S$  différents et de probabilités de succès différentes,  $p_i = \tau_i$ , à un échantillon-test idéal de 22 objets qui auraient le même redshift,  $\langle z_S \rangle = z_{S,\text{tot}}/n = 1.399$  et la même proportion  $p$ , celle-ci étant estimée par la moyenne des proportions de chaque objet, soit  $\langle p_i \rangle$ .

Afin de calculer la profondeur optique  $\tau$  pour l'univers plus plausible avec constante cosmologique ( $\Omega_0 = 0.3$  et  $\lambda_0 = 0.7$ ), on intègre numériquement l'équation (6.6) pour trouver les trois distances de diamètre angulaire, puis on intègre aussi numériquement l'équation (6.5) pour trouver  $\tau$ . À ce propos, le tableau 6.2 indique les profondeurs optiques calculées pour trois valeurs du paramètre de densité cosmique nette de minilentilles.

**Tableau 6.2.** Redshift  $z_S$  et profondeur optique  $\tau_i$  ( $\Omega_{l,0} = 1, 0.3$  et  $0.2$ ) des 22 radiosources pour le modèle dominé par une constante cosmologique ( $\Omega_0 = 0.3$  et  $\lambda_0 = 0.7$ ).

Objet	$z_S$	$\tau_i (\Omega_{l,0} = 1)$	$\tau_i (\Omega_{l,0} = 0.3)$	$\tau_i (\Omega_{l,0} = 0.2)$
PKS 1622-253	0.786	0.1626	0.0478	0.0325
4C +01.02	2.099	0.7987	0.2396	0.1597
Q0957+561AB	1.413	0.4456	0.1337	0.0891
OJ 287	0.306	0.02686	0.0081	0.0054
3C 273	0.1583	0.0097	0.0029	0.0019
3C 279	0.536	0.0788	0.0236	0.0158
OT 081	0.322	0.02882	0.0086	0.0058
3C 345	0.5935	0.0922	0.0277	0.0184
NRAO 530	0.902	0.2164	0.0649	0.0433
B0218+357	0.96	0.2497	0.0749	0.0499
B1030+074	1.535	0.5137	0.1541	0.1027
PKS 0528+134	2.07	0.7915	0.2375	0.1583
4C +73.18	0.302	0.02656	0.008	0.0053
PMN 0414+0534	2.639	1.033	0.3099	0.2066
PMN 0134-0931	2.225	0.9137	0.2741	0.1827
B1422+231	3.62	1.6242	0.4873	0.3248
PKS 1830-211	2.507	0.9812	0.2944	0.1962
4C +38.41	1.813	0.594	0.1782	0.1188
NRAO 150	1.52	0.507	0.1521	0.1014
PKS 2134+004	1.932	0.687	0.2061	0.1374
3C 454.3	0.859	0.197	0.0591	0.0394
3C 418	1.686	0.544	0.1632	0.1088
<b>Valeurs moyennes</b>	1.399	0.478	0.144	0.096

#### 6.4.4 Intervalles de confiance à 95 % et estimé de la densité cosmologique d'objets compacts pour le modèle plat avec $\Omega_0 = 0.3$ et $\lambda_0 = 0.7$

Le caractère discret de la loi binomiale fait en sorte que le calcul exact d'un intervalle de confiance (barre d'incertitude) est difficile. Dans les faits, la façon adéquate de procéder pour calculer un intervalle de confiance, spécialement pour de petits échantillons, est sujette à des controverses parmi les statisticiens (Brenner & Quan 1990). Sachant cela, on doit confectionner un intervalle de confiance

approximant le plus justement possible l'intervalle de confiance "exact", qui est inconnu.

À partir de la fréquence observée dans notre échantillon-test de seulement 22 objets, on applique trois méthodes statistiques reconnues pour calculer l'intervalle de confiance d'une binomiale dans le cas d'échantillons de petite taille ( $n < 20 - 30$ ). De toute évidence, il faut ici exclure les méthodes courantes nécessitant de grands échantillons, telles que celles basées sur des approximations de la loi normale ou de Poisson. Cela dit, les trois méthodes de calcul en question sont respectivement les méthodes de Clopper-Pearson, aussi appelée de la binomiale "exacte", de Wilson et de Jeffreys. La méthode de Clopper-Pearson est très conservatrice, au point où elle a tendance à donner des intervalles trop larges. Voilà pourquoi, Brown et al. (2001) recommandent plutôt les méthodes plus précises de Wilson et de Jeffreys. L'annexe B résume les trois méthodes de calcul appliquées pour trouver les intervalles de confiance, à un seuil de 95 %, pour notre échantillon-test,  $n = 22$ . Le tableau 6.3 et montre les résultats obtenus en appliquant ces trois méthodes.

**Tableau 6.3.** Résultats obtenus par trois méthodes de calcul différentes à 95 % de confiance pour  $n = 22$  objets.

	<b>Clopper-Pearson</b>	<b>Wilson</b>	<b>Jeffreys</b>
Nombre de minilentilles détectées: X	0	0	0
Nombre d'objets: n	22	22	22
Proportion observée: $p = X/n$	0	0	0
Proportion minimale de l'intervalle	0	0	0
Proportion maximale de l'intervalle	0.154	0.149	0.107

Les intervalles de confiance calculés par les méthodes de Clopper-Pearson et de Wilson se trouvent respectivement par les équations (B.4) et (B.5), et donnent,  $IC_{C-P} = [0, 0.154]$  et  $IC_W = [0, 0.149]$ . Vu que les limites supérieures de  $IC_{C-P}$  et de  $IC_W$  avoisinent  $\langle p_i \rangle = \langle \tau_i \rangle = 0.144$ , elles correspondent alors au cas où  $\Omega_{l,0} = 0.3$  (voir tableau 6.2). Pour sa part, l'intervalle de confiance de Jeffreys se trouve par les

équations (B.6a) et (B.6b), et donne,  $IC_J = [0, 0.107]$ . La borne supérieure de  $IC_J$  correspond à environ  $\langle p_i \rangle = \langle \tau_i \rangle = 0.096$ , soit le cas où  $\Omega_{1,0} = 0.2$  (voir tableau 6.2).

Pour la cosmologie dominée par une constante cosmologique ( $\Omega_0 = 0.3$  et  $\lambda_0 = 0.7$ ), les résultats obtenus de l'analyse statistique simplifiée, de seulement 22 radiosources, basée sur trois intervalles de confiance d'une binomiale, *estiment que les objets compacts sombres non-baryoniques intergalactiques uniformément distribués, dans l'intervalle de masse  $10 \leq M / M_e \leq 10^7$ , contribueraient au plus de 20 à 30 % du contenu de l'univers.* L'estimé le plus optimiste, correspondant à la méthode de Jeffreys donnant la plus faible incertitude, vaut  $\Omega_{1,0} \leq 20 \%$ . N'empêche que l'estimé de Jeffreys donne tout de même un intervalle de confiance (barre d'incertitude) trop grand ( $p = 0.107$ ). La contrainte statistique estimée,  $\Omega_{1,0} \leq 20 \%$ , n'est donc pas assez stricte: statistique de petit nombre.

## 7 Conclusion

Dans cette thèse expérimentale de doctorat, on a développé une technique numérique d'autocorrélation d'intensité afin de détecter des événements de minilentilles gravitationnelles, et également de détecter de brèves variations temporelles (périodiques ou non) d'intensité. Pour la première fois, la technique a été appliquée à des observations radio de NAGs (ex., quasars) prises dans un format standard (ex., Mark 5B) pour des observations interférométriques.

### 7.1 Format des données, algorithmes de base et logiciels Matlab/C++

Dans le but de mettre au point les algorithmes et les logiciels associés, on s'est avant tout assuré de bien maîtriser le format des données stockées, échantillonnées et quantifiées et/ou codées sur le système de lecture Mark 5B. Les données utilisées provenaient d'observations faites parmi les sept stations VLBA du NRAO (voir tableau A.4). Ces données, qui sont en fait des champs électriques, ont été échantillonnées, à une fréquence de Nyquist, à toutes les 15.625 nanosecondes et ont été enregistrées sur seulement deux bits. Les champs électriques sont compris dans huit canaux RCP et huit autres LCP de 32 MHz de largeur, et s'étendent sur la gamme de fréquences, 8256 – 8512 MHz (voir tableau A.1). Compte tenu de la limite d'enregistrement imposée par les deux bits, les champs électriques étaient alors restreints aux quatre valeurs binaires 00, 01, 10 et 11 ou décimales 0, 1, 2 et 3. Cependant, on a développé une fonction simple (voir annexe B.3) permettant la conversion de ces quatre champs positifs en quatre champs électriques signés,  $\pm 3.3359$  et  $\pm 1$ , sur une échelle conforme aux seuils d'échantillonnage du VLBA (voir sections 4.5.3 et 4.5.5). La compréhension des données étant acquise, on a commencé en programmant les algorithmes de base en Matlab sur un microordinateur avec un nombre très limité de données (ex., 2.5 millions) afin de d'implémenter notre technique de calcul. Après avoir validé le code Matlab développé (voir section 5.2.4), on a ensuite traduit (adapté) ce code Matlab en langage POO C++ pour traiter en

parallèle (ex., sur 1024 cœurs de calcul) la grande quantité de données (ex., 38 milliards) des fichiers d'entrée sur le superordinateur Colosse de l'Université Laval.

## **7.2 Détections de faibles variations périodiques et non périodiques**

On a validé la technique et les logiciels Matlab et C++ développés en montrant que l'on pouvait facilement détecter des signaux périodiques très rapides et de très faibles amplitudes, lesquels seraient complètement perdus dans le bruit dans le profil temporel d'intensité. À ce propos, bien que la thèse présente seulement un test validant la détection de signaux périodiques faibles (voir section 5.2.5), la technique d'autocorrélation a permis de détecter plus d'une dizaine d'autres signaux périodiques de faibles amplitudes (non décrits dans la thèse).

On a aussi appliqué, pour la première fois, l'autocorrélation d'intensité pour détecter des variations temporelles non périodiques. Comme exemple de signal périodique, on aurait pu trouver un signal généré par de brèves pulsations périodiques provenant de trous noirs dans les quasars. Un exemple de signal non périodique possible aurait été un signal produit par des pulsations non périodiques émises par certains objets exotiques. Cela dit, rappelons trois détections importantes de signaux artificiels, car celles-ci valident la technique et les logiciels développés.

- 1) Lors de l'analyse du tout premier objet, le quasar PKS 1622-253, on a détecté de faibles impulsions périodiques rapides de 5 MHz (voir section 5.2.5). Ces impulsions étaient totalement perdues dans le bruit dans le profil temporel d'intensité. Elles étaient également invisibles si l'autocorrélation s'effectuait sur des temps trop courts. Par exemple, une autocorrélation en Matlab sur 39 ms ne permettait pas de faire ressortir le signal impulsionnel. En revanche, ces impulsions devenaient très visibles (voir figure 5.6) en exécutant une autocorrélation en C++ sur des temps plus longs (ex., 50 secondes).

- 2) Lors de l'analyse du quasar 3C 345, on a détecté un signal RFI quasi périodique rapide de  $\sim 8$  ms résultant d'un radar situé à proximité d'une station VLBA (voir section 5.2.7). Ce signal était clairement visible dans le profil d'autocorrélation d'intensité (voir figure 5.8), mais totalement invisible dans celui de l'intensité, et cela même après un lissage pour diminuer l'amplitude du bruit.
- 3) Lors de l'analyse du quasar PMN 0134-0931, on a détecté des variations temporelles d'intensité non périodiques rapides et de faibles amplitudes (voir section 5.2.6). En fait, c'était la première fois que l'autocorrélation d'intensité était appliquée à la détection de variations non périodiques, particulièrement dans le radio. La figure 5.7 donne un exemple du profil d'autocorrélation d'intensité découlant de la présence de plusieurs signaux artificiels non périodiques.

La principale conclusion qu'il faut tirer des trois détections (validations) précédentes est que notre technique d'autocorrélation d'intensité peut aisément être appliquée pour détecter de faibles variations temporelles périodiques et non périodiques de l'ordre de la période d'échantillonnage. On a pu constater cela à la figure 5.6 dans laquelle la signature d'un signal artificiel, de période (200 ns)  $\sim 13$  fois plus grande que le temps d'échantillonnage (15.625 ns), était nettement visible dans le profil d'autocorrélation. Rappelons toutefois que la limite inférieure de la période, de l'ordre de la milliseconde (voir tableau 6.1), est imposée par l'effet de lissage issu de la dispersion due aux milieux de propagation (voir section 2.2), essentiellement le plasma intergalactique ionisé. Cet effet de lissage vient du fait que les ondes radio de fréquences différentes se propagent à des vitesses différentes dans le MIG, de sorte que les champs électriques ayant différentes fréquences dans la bande passante, 32 MHz par canal (voir tableau A.1), arrivent à des temps différents au radio-observatoire. Le MIG étale les impulsions, ce qui diminue leur amplitude au point où elles deviennent plus difficiles à détecter (voir section 5.3.1). Heureusement, l'effet de lissage est moins important dans l'autocorrélation en raison de son effet de

moyenne. En l'occurrence, il faudrait un étalement beaucoup plus grand que la largeur des impulsions pour que les variations temporelles d'intensité deviennent indétectables par l'autocorrélation. D'ailleurs, la figure 5.6 montrait formellement que le signal impulsionnel était détectable malgré un lissage temporel ( $\sim 1400$  ns) sept fois plus grand que la période du signal (200 ns). Sachant que le délai temporel dû à la dispersion intergalactique diminue lorsque la fréquence augmente (voir équation (2.17)), des observations faites à des radiofréquences supérieures à 8.4 GHz devraient permettre de détecter des variations périodiques plus brèves. Qui plus est, l'effet de dispersion serait moindre en utilisant des bandes passantes plus étroites que 32 MHz, car l'étalement temporel dépend de la propagation de différentes fréquences à différentes vitesses dans le plasma intergalactique. Entre autres choses, notre technique d'autocorrélation d'intensité pourrait aussi servir à détecter des objets variant rapidement dans notre Galaxie ou dans des galaxies voisines. Pour ces objets, la dispersion serait faible puisqu'elle proviendrait uniquement du milieu galactique ( $DM_G \leq 100$  pc/cm<sup>3</sup>), si bien que l'on pourrait trouver des variations temporelles encore plus rapides. La technique pourrait également être utilisée pour rechercher et étudier des pulsars millisecondes radio (ex., Hessels et al. 2006).

Un des avantages d'utiliser l'autocorrélation d'intensité vient de sa simplicité, car on n'a pas besoin d'instrumentation spécialisée telle que celle utilisée pour observer les pulsars. On utilise simplement les données obtenues avec des radiotélescopes déjà existants dans leur format numérique standard (voir section 4.4.2). L'utilisation de signaux artificiels pour valider notre technique peut être discutable, surtout si l'on considère les fausses détections venant de tels signaux. À ce sujet, disons que les signaux artificiels de calibration (voir figures 5.3 et 5.6) peuvent facilement être désactivés pendant les observations afin de ne pas entraver la détection de variations temporelles. D'ailleurs, ils ont effectivement été désactivés dans la majorité de nos observations. Les signaux artificiels résultant des autres sources artificielles, comme le signal RFI qui a généré le profil de la figure 5.8, peuvent aisément être exclus en utilisant, comme on l'a fait, des observations simultanées provenant de maintes stations VLBA situées à différents endroits (voir

tableau A.4). Un signal artificiel serait généralement présent dans un seul radio-observatoire. Vu que les radiotélescopes du NRAO sont séparés par de très grandes distances, même les signaux atmosphériques ne seraient pas présents dans les radiotélescopes séparés. Sur ce point, rappelons que le signal RFI a seulement été détecté à la station SC, et par surcroît, uniquement dans les canaux RCP (voir section 5.2.7). Précisons enfin que même si un très grand nombre d'observations, utilisées à des fins différentes, n'ont pas de signaux artificiels, on discute, dans cet ouvrage, des observations comportant des signaux artificiels vu qu'elles valident notre technique.

D'autres techniques appliquant la transformée de Fourier (TF) telles que celles utilisées dans le logiciel PRESTO (ex., Ransom et al. 2002) permettent aussi de trouver des signaux périodiques dans des données numériques. Pourtant, un aspect innovateur majeur des résultats obtenus en appliquant l'autocorrélation d'intensité, comparativement à ceux obtenus si on appliquait la TF, réside dans le fait que l'autocorrélation peut détecter des signaux non périodiques, alors que la TF est impuissante à le faire. Tel que déjà dit à la section 5.5.5, un avantage de l'autocorrélation vient du fait qu'elle permet de facilement visualiser la forme et les variations temporelles des signaux auto-corrélés (voir figures 5.3, 5.6 et 5.8). Malgré que la forme de l'autocorrélation n'est pas la même que celle des signaux de base qui dépendent du temps, elle donne les caractéristiques générales des signaux comme la périodicité et la durée moyenne des impulsions. Ajoutons que même si la transformée de Fourier identifie aisément la période des impulsions, il est très difficile de visualiser leur forme dans le spectre de Fourier.

La technique discutée dans Borra (2010) et utilisée par Trottier (2012), Borra (2013) et Borra & Trottier (2016), soit la découverte de modulations périodiques dans des spectres de galaxies et de certaines étoiles, pourrait également permettre de trouver des variations rapides d'intensité. Toutefois, cette technique ne pourrait pas servir à détecter des impulsions temporellement séparées par moins de 0.1 nanoseconde dans les spectres optiques. En revanche, notre technique serait en mesure de trouver des variations périodiques de temps beaucoup plus courtes que 0.1

ns, tout comme elle pouvait détecter des variations temporelles quasi périodiques (voir figure 5.8) et non périodiques (voir figure 5.7).

### **7.3 Rappel des trois étapes de détections de signaux**

- 1) On a imposé un critère de sélection basé sur le rapport signal à bruit (S/B). En se basant sur la loi d'une gaussienne, on a retenu,  $S/B = 4.0$ , comme condition d'extraction de signaux dans le profil d'autocorrélation d'intensité rectifié (voir section 5.3.2). Cette valeur de S/B a permis de limiter l'inspection et l'analyse individuelle à 150 – 250 pics et de restreindre les fausses identifications dues aux fluctuations statistiques.
- 2) On a vérifié si un signal cosmique a été détecté simultanément dans un minimum de deux observatoires radio. On n'a pas retenu les signaux détectés dans un seul radio-observatoire, car ces signaux étaient généralement des signaux artificiels locaux parasites (ex., effets instrumentaux, signaux RFI, etc.).
- 3) Si des signaux sont détectés dans aux moins deux stations VLBA (l'étape 2 serait ainsi satisfaite), on vérifie alors s'ils sont présents dans un minimum de deux canaux (le canal RCP et le canal LCP adjacent) de chaque télescope radio (mêmes canaux et la même position dans chaque canal). Si tel est le cas, on impose un critère de priorité pour ordonnancer les investigations par ordre décroissant du nombre de canaux où sont les signaux: de seize (priorité n° 1) à deux canaux (priorité n° 8).

### **7.4 Simulations numériques de délais temporels**

À la section 6.2, on a effectué une centaine de simulations numériques afin de montrer l'efficacité de détection de signaux (délais temporels) via la technique d'autocorrélation d'intensité. En bref, les quatre étapes des simulations consistaient à

1) introduire un délai temporel arbitraire  $\Delta t_{\text{arb}}$  entre les champs électriques de deux faisceaux:  $E(t) = E_1(t) + E_1(t+\Delta t_{\text{arb}})$ , 2) simuler l'effet de dispersion intergalactique avec un lissage de  $DM_{\text{MIG}}$  points (voir tableau 6.1), 3) générer du bruit local artificiel (sol, atmosphère et récepteur), 4) sommer l'intensité du bruit local simulé à l'intensité de l'objet et à auto-corréler cette somme.

Pour une géométrie à deux faisceaux d'intensité comparable, les résultats des simulations ont révélé que la détection d'un signal par l'autocorrélation d'intensité est efficace ( $S/B \geq 4$ ) pour l'intervalle de masse applicable,  $10 \leq M / M_e \leq 10^7$ . Par exemple, même pour un signal auto-corrélé éloigné de l'origine (ex., 280 secondes), la signature du signal était bien visible (voir figures 6.3 et 6.5).

## 7.5 Estimé de la densité cosmologique de minilentilles

On n'a pas détecté d'événements de minilentilles gravitationnelles dans l'échantillon-test de vingt-deux objets.

À partir d'une analyse statistique simplifiée de notre résultat nul basée sur trois méthodes de calcul d'intervalles de confiance (Clopper-Pearson, Wilson et Jeffreys) d'une binomiale, on a fait un estimé (approximatif) de la valeur de  $\Omega_{1,0}$  dans l'intervalle de masse,  $10 M_e \leq M \leq 10^7 M_e$ . Dans le cadre d'un modèle plat avec constante cosmologique ( $\lambda_0 = 0.7$  et  $\Omega_0 = 0.3$ ), la limite estimée vaut  $\Omega_{1,0} \leq 0.2 - 0.3$ . L'échantillon-test de seulement 22 objets fait en sorte que cette limite comporte de trop grandes incertitudes. En bref, la contrainte statistique estimée de  $\Omega_{1,0}$  n'est pas suffisamment sévère (voir section 6.4.4).

Le modèle simple d'une distribution uniforme d'objets compacts (ex., Press & Gunn 1973), comparativement à un modèle, plus réaliste, où les minilentilles seraient regroupées dans des halos sombres aléatoirement distribués, aurait un impact plutôt modeste sur la probabilité moyenne d'avoir des événements de minilentilles pour des

radiosources à redshift élevé et des minilentilles sombres compactes. En revanche, la distribution (inhomogène) de matière sombre dans des halos donnerait une dispersion notable de la profondeur optique, particulièrement pour les objets à petit redshift (Zackrisson & Riehm 2007; Holopainen et al. 2008).

## 7.6 Limitations du projet de recherche

La principale limitation de cette thèse expérimentale de physique, mais aussi d'informatique, vient du faible nombre de sources radio traitées, soit 22 seulement. Ce nombre limité d'objets s'explique en grande partie par la difficulté à obtenir des observations radio du NRAO. Les longs délais des demandes et d'attente nous ont limités à recevoir deux séries d'observations par année pour un total d'environ cinq objets par an. La grande quantité de données à traiter exigeait l'emploi d'un superordinateur. Or, comme les tâches soumises requéraient plusieurs cœurs de calcul (ex., 1024), l'attente des résultats était généralement d'une semaine. Au début de cet ambitieux projet, on n'avait pas prévu de telles complications, car c'était la première fois que ce type de travail s'effectuait, à savoir, l'application d'une technique numérique d'autocorrélation d'intensité à des observations utilisant des données radioastronomiques standards. À ce propos, le fait de travailler dans le domaine radio entraîne une autre limitation provenant de la dispersion intergalactique qui contraint la détection des délais temporels à l'ordre de la milliseconde (le délai temporel moyen dû au MIG étant de 0.455 ms), et du même coup restreint la borne inférieure de l'intervalle de masse applicable à une dizaine de masses solaires (voir section 6.2). Par exemple, dans l'optique, domaine où l'influence du MIG est négligeable, notre technique permettrait d'obtenir des délais de l'ordre de la période d'échantillonnage du NRAO (15.625 ns), si bien que d'après l'équation (3.14) elle permettrait de détecter des objets compacts sombres peu massifs de l'ordre de  $10^{-4}M_{\odot}$ , ce qu'on ne peut pas faire dans le radio.

La taille très limitée de l'échantillon-test (22 objets) implique que la signification statistique des résultats (contraintes) ne s'est pas avérée suffisamment précise. De fait, l'analyse statistique simplifiée par la binomiale a donné de trop

grandes incertitudes. Voilà pourquoi, même le meilleur estimé de la limite supérieure de la densité cosmique de minilentes,  $\Omega_{l,0} \leq 0.2$  (méthode de Jeffreys), pour un univers critique à constante cosmologique ( $\lambda_0 = 0.7$  et  $\Omega_0 = 0.3$ ), n'est pas assez stricte: statistique de petit nombre.

## 7.7 Ouverture sur la détection de FRBs et sur l'intelligence extraterrestre

Les sursauts radio rapides (FRBs) sont une récente découverte en astronomie. Compte tenu que le ciel transitoire clignotant radio n'est toujours pas bien compris, il s'ensuit que l'analyse des données via de nouvelles techniques peut permettre de découvrir des phénomènes physiques fortuits. Ceci étant dit, le premier FRB a été découvert en 2001. Depuis ce temps, seulement 18 objets émettant des FRBs ont été découverts. Ces FRBs sont puissants et durent quelques millisecondes. La plupart des détections de FRBs comprenait une seule impulsion, hormis le FRB121102 (Spitler et al. 2016) qui comportait dix sursauts répétés, avec six sursauts survenant dans un intervalle de dix minutes et trois sursauts à des semaines d'intervalle. Malgré que l'autocorrélation d'intensité n'ait pas vraiment d'avantages pour détecter un seul sursaut radio rapide, elle permettrait de détecter des FRBs se répétant à l'intérieur de la durée d'observation. Le FRB121102 a six sursauts brusques dans un intervalle de 600 secondes, lequel est compris dans les 650 secondes d'observation du profil de la figure 5.7 résultant de signaux non périodiques. Bien que beaucoup moins nombreux que les impulsions qui auraient généré le profil de la figure 5.7, les six sursauts de FRB121102 donneraient des signaux détectables étant donné qu'ils sont beaucoup plus puissants. Comme l'origine des FRBs est encore inconnue, il en découle que toutes les hypothèses sont possibles. Il est alors possible que certains objets astronomiques émettent un grand nombre de FRBs très faibles, lesquels ne sont pas détectables avec les techniques standards. Ces nombreux FRBs faibles généreraient précisément le même type de signal que celui illustré à la figure 5.7.

Law & Bower (2012) discutent d'une technique permettant de détecter des transitoires radio. Par contre, leur technique est très ardue comparativement à la nôtre qui est très simple. Contrairement à notre technique, celle de Law & Bower (2012) dépend des résultats interférométriques. Qui plus est, Law & Bower (2012) ne démontrent pas l'utilisation de leur technique pour détecter les nombreuses impulsions non périodiques faibles, tandis que nous on l'a fait.

Le lecteur doit savoir qu'il y a eu des spéculations soulevant que les sursauts radio rapides pourraient provenir d'une forme d'intelligence extraterrestre (New Scientist March 31 2015). Si une telle forme d'ETI (*extraterrestrial intelligence*) envoie des messages encodés dans maintes impulsions radio plus faibles que les FRBs, l'autocorrélation d'intensité pourrait les détecter sous formes d'impulsions non périodiques issues des étoiles. Cela est évidemment très spéculatif, mais également très intéressant advenant le cas où des signaux ETI sont découverts.

## Bibliographie

- Alford, W. P., Gold, A., 1958, Am. J. Phys., 26, 481
- Ambardar, A., 2003, Digital Signal Processing: A Modern Introduction, Michigan. Univ.
- Baczynski, A., Oster, G., 1965, J. Opt. Soc. Am., 55, 1325
- Basano, L., Ottonello, P., 2000, Am. J. Phys., 68, 325
- Beckmann, V., Shrader, R., 2013, arXiv:astro-ph.HE/1302.1397v1
- Bélisle, C., 2009, Probabilités pour ingénieurs. Université Laval. Québec
- Biraud, F., Rieu, N. Q., 1994, Radioastronomie, Généralités. Techniques de l'Ingénieur, E 6850v1
- Blanchet, G., Charbit, M, 2013, Traitement Numérique du Signal. Techniques de l'Ingénieur, E3087v2
- Blandford, R. D., & Nayaran, R., 1992, ARA&A, 30, 311
- Borra, E. F., 1997, MNRAS, 289, 660
- Borra, E. F., 2008, MNRAS, 389, 364
- Borra, E. F., 2010, A&A, 511, L6
- Borra, E. F., 2011, MNRAS, 411, 1695
- Borra, E. F., 2013, ApJ, 774, 142
- Borra, E. F., 2014, MNRAS, 440, 1993
- Borra, E. F., Trottier, E., 2016, PASP, 128, 114201
- Bracewell, R. N., 1986, The Fourier Transform and its Applications. McGraw-Hill, New York
- Brenner, D. J., Quan H., 1990, J. R. Stat. Soc., Series D, 39, 391
- Brown, L. D., Cat, T.T., DasGupta, A., 2001, Statist. Sci., 16, 101

Burke, B., Graham-Smith, F., 2010, *An Introduction to Radio Astronomy*, Third Edition. Cambridge Univ. Press, New York

Canizares, C. R., 1981, *Nature*, 291, 620

Canizares, C. R., 1982, *ApJ*, 263, 508

Carilli, C., 2010, *Noise and Interferometry*, Thirteenth Synthesis Imaging Workshop, NRAO, Socorro, NM

Chang, K., Refsdal, S., 1979, *Nature*, 282, 561

Cielo, P., Brochu, M., Delisle, C., 1975, *Can. J. Phys.*, 53, 1047

Claeskens, J. F., Surdej, J., 1996, *Determination of the Cosmological Density of Compact Objects using Gravitational Lensing and the HST*, STScI, 96

Claeskens, J. F., 2004, *Cosmological Density of Dark Compact Objects from Observations of Lensed QSOS*, *“THINKING, OBSERVING AND MINING THE UNIVERSE”*, pp. 73-78

Clarke, N., D’Addario, L., Navarro, R., Trinh J., 2014, *JAI*, 3, 1

Collin-Zahn, S., 2009, *Des quasars aux trous noirs*. EDP Sciences, France

Cordes, J. M., 1976, *ApJ*, 210, 780

Cordes, J.M., Bhat, N. D. R., Hankins, T. H., McLaughlin, M. A., Kern, J. 2004, *ApJ*, 612, 375

Courbin F., Saha P., Schechter P. L., 2002, arXiv:astro-ph/0208043v1

Dalcanton, J. J., Canizares, C. R., Granados, A., Steidel, C. C., Stocke, J. T., 1994, *ApJ*, 424, 550

Delisle, C., Brochu, M., 1972, *Can. J. Phys.*, 50, 219

Deng, W., Zhang, B., 2014, *ApJ*, 783, L35

Dodelson, S., 2003, *Modern Cosmology*, First Edition. Academic. Press, London

Dominik, M. 1999, *A&A*, 349, 108

Dominik, M., Jorgensen, G., Horne, K., et al. 2008, arXiv:astro-ph/0808.0004v1

Dusart, P., 2012, *Statistiques Inférentielles*. Univ. Limoges. France  
[http://www.unilim.fr/pages\\_perso/pierre.dusart/Probab/document\\_stat\\_irem.pdf](http://www.unilim.fr/pages_perso/pierre.dusart/Probab/document_stat_irem.pdf)

Dyson, F. W., Eddington, A. S., Davidson, C. R., 1920, Mem. R. Astron. Soc., 62, 291

Einstein, A., 1911, Ann. Phys., 35, 898

Einstein, A., 1916, Ann. Phys., 49, 769

Elvis, M., 2000, ApJ, 545, 63

Forrester, A. I., Gudmundsen, R. A., Johnson, P. O., 1955, Phys. Rev., 99, 1691

Fukugita, M., Futamase, T., Kasai, M., Turner, E. L., 1992, ApJ, 393, 3

Ghez, A. M., Salim, S., Weinberg, N. N., Lu, J. R., Do, T., Dunn, J. K., Matthews, K., Morris, M., Yelda, S., Becklin, E. E., Kremenek, T., Milosavljevic, M., Naiman, J., 2008, arXiv:astro-ph:0808.2870v1

Giuliano, S., Berardi, V., Shih, Y., 2006, Phys. Rev. Lett. 96, 063602

Givens, M. P., 1961, J. Opt. Soc. Am., 51, 1030

Goodman, J. W., 2015, Statistical Optics, Second Edition. John Wiley & Sons, New York

Gott J. R., 1981, ApJ, 243, 140

Hanbury Brown, R., 1968, ARA&A, 6, 13

Hayward, B., 2012, Antenna & Receivers in Radio Astronomy, Thirteenth Synthesis Imaging Workshop, NRAO, Socorro, NM

Hessels, J. W. T., Ransom, S. M., Stairs, I. H., Freire, P. C., Kaspi, V. M., Camilo, F. 2006, Science, 311, 1901

Holopainen, J., Zackrisson, E., Knebe, A., Nurmi, P., Heinamaki, P., Flynn, C., Gill, S., Riehm, T., 2008, MNRAS, 383, 720

Inoue, S., 2004, MNRAS, 348, 999

Ioka, K., 2003, arXiv:astro-ph/0309200v2

Irwin, M. J., Webster, R. L., Hewett, P. C., Corrigan, R. T., Jedrzejewski, R. I., 1989, AJ, 98, 1989

Kassiola, A., Kovner, I., Blanford, R. D., 1991, ApJ, 381, 6

Kayser, R., Helbig, P., Schramm T. 1997, A&A, 318, 680

Kellermann K. I., Frazer, N.O, 1988, Radio Galaxy and Quasars. Published in Galactic and Extragalactic Radio Astronomy, Second Edition. Berlin: Springer

Krishnamoorthy, K., Peng, J., 2007, Communications in Statistics – Simulation and Computation, 36, 1171

Landau, L., Lifshitz, E., 2004, The Classical Theory of Fields, Fourth Edition, Vol. 2, Theoretical physics. Field theory. Pergamon. Press Ltd., Elsevier, London

Law, C. J., Bower, G. C., 2012, ApJ, 749, 143

Leland, B. J., 1996, Digital Filters and Signal Processing, Third Edition. Academic Publishers, New York

Léna, P., Rouan, D. 2008, L'OBSERVATION EN ASTROPHYSIQUE, EDP Sciences, France

Lewis, G. F., Ibata, R. A., 2004, MNRAS, 348, 24

Longair, M. S., 2011, High Energy Astrophysics, Third Edition. Cambridge Univ. Press, New York

Loudon, R., 2000, The Quantum Theory of Light, Third Edition. Oxford Univ. Press, Oxford

Lovelace, R. V. E., Richards, D. W., 2013, MNRAS, arXiv:astro-ph.CO/1302.3457v2

Lyndel-Bell, D., 1969, Nature, 223, 690

Mandel, L., 1962a, Prog. Opt., 2, 183

Mandel, L., 1962b, J. Opt. Soc. Am., 52, 1335

Mandel, L., Wolf, E., 1965, Rev. Mod. Phys., 37, 231

Mandel, L. Wolf, E., 1995, Optical Coherence and Quantum Optics. Cambridge Univ. Press, New York

Marshall, P., Blandford, R., Sako, M., 2005, N. Astron. Rev. 49, 387

MIT Haystack Observatory, 2008, Mark 5C Data-Frame Specification. Mark 5 Memo 58, NRAO, Socorro, NM

Napier, P. J., et al., 1994, Proc. IEEE 82, 658

Narayan, R., Bartelmann, M., 1997, Lectures on gravitational lensing. arXiv:astro-ph/9606001v2

Nityananda, R., Ostriker, J. P., 1984, *J. Ap. Astr.*, 5, 235

Olivier, J., May, W. L., 2006, *Stat. Methods in Med. Res.*, 15(1), 37

O'Neill, E. L., 2003, *Introduction to Statistical Optics*. Addison Wesley, Boston,

Oppenheim, A. V., 2014, *Signals and Systems, Second Edition*. Pearson Education Limited, England

Orfanidis, J., 2010, *Introduction to Signal Processing*. Pearson Education, Inc. Rutgers Univ., New Jersey

Paczynski, B., 1986a, *ApJ*, 301, 503

Paczynski, B., 1986b, *ApJ*, 304, 1

Paczynski, B., 1996, *ARA&A*, 34, 419

Petters, A. O., Levine, H., Wambsganss, J., 2001, *Singularity Theory and Gravitational Lensing*. New York: Springer.

Pires, A. M., Conceição, A., 2008, *Revstat Stat J.*, 6(2), 165

Press, W. H., & Gunn, J. E., 1973, *ApJ*, 185, 397

Ransom, S. M., Eikenberry, S. S., Middleditch, J., 2002, *AJ*, 124, 1788

Refregier, A. 2003, *ARA&A*, 41, 645, arXiv:astro-ph/0307212v1

Refsdal, S., 1964b, *MNRAS*, 128, 307

Refsdal, S., 1966, *MNRAS*, 132, 101

Refsdal, S., Stabell, R., Pelt, J., Schild, R., 2000, *A&A*, 360, 10

Refsdal, S., Surdej, J., 1994, *Rep. Prog. Phys.* 56, 117

Richards, E. A., 2000, *ApJ*, 533, 611

Rickett, B. J., 1975, *ApJ*, 197, 185

Romney, J., Reid, M., 2005, *ASP Conference Proceedings* 340

Rybicki, G. B., Lightman, A. P., 1979: *Radiative Processes in Astrophysics*. Wiley, New York

Schneider, P., 1993, *A&A.*, 279, 1 (S93)

Schneider, P., Ehlers, J., & Falco, E. E., 1992, Gravitational Lenses. Berlin: Springer

Schneider, P., Kochanek, C. S., Wambsganss, J., 2006, Gravitational Lensing: Strong, Weak and Micro: Saas-Fee Advanced Course 33, Berlin: Springer

Schneider, P., Wambsganss, J., 1990, A&A, 237, 42

Shields, G. A., 1978, Nature, 272, 706

Shude, M., 2008, arXiv:astro-ph/08110441v1

Soldner, J., 1804, Berliner Astronomisches Jahrbuch, p. 161

Spillar, E. J., 1993, ApJ, 403, 20

Spitler, L. G., Scholz, P., Hessels, J. W. T., et al. 2016, Nature, 531, 202

Taylor, J. H., Cordes, J. M., 1993, ApJ, 411, 674

Thompson, A. R., Emerson, D. T., 2005, Quantization Noise, EVLA Memo 88, NRAO, Socorro, NM.

Thompson, A. R., Moran, J. M., Swenson Jr, G. W., 2004, Interferometry and Synthesis in Radio Astronomy, Second Edition. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

Thornton, D., Stappers, B., Bailes, M., et al. 2013, Sci, 341, 53

Thulin, M., 2013, arXiv:1303.1288v1

Treu, T., 2010, ARA&A, 48, 87

Trottier, E., 2012, M.Sc., Univ. Laval

Turner, E. L., 1990, ApJ, 365, L43

Turner, E. L., Ostriker, J. P., & Gott, J. R., 1984, ApJ, 284, 1 (TOG)

Vietri, M., Ostriker, J. P., 1983, ApJ, 267, 488

Wallis, S., 2013, JQL, 20, 178

Walsh, D., Carswell, R. F., Weymann, R. J., 1979, Nature, 279, 381

Wambsganss, J. 2000, arXiv:astro-ph/0008419v1

Wambsganss, J., Paczynski, B., 1992, ApJ, 397, 1

- Wambsganss, J., Paczynski, B., Katz, N., 1990a, *ApJ*, 352, 407
- Wambsganss, J., Schmidt, R.W., Colley, W., Kundic, T., Turner, E. L., 2000, *A&A*, 362, L37
- Wayth, R. B., Brisken, W. F., Deller, A. T., Majid, W. A., Thompson, D. R., Tingay, S. J., Wagstaff, K. L., 2011, *ApJ*, 735, 97
- Wilkinson, P. N., Henstock, D. R., Browne, I. W. A., Polatidis, A.G., Augusto, P., Readhead, A. C. S., Pearson, T. J., Xu, W., Taylor, G. B., Vermeulen, R. C., 2001, *arXiv:astro-ph/0101328v1*
- Wilson, T. L., Rohfls, K., Huttemeister, S., 2009, *Tools of Radio Astronomy*, Fifth Edition. Berlin: Springer
- Witt, H. J., Kayser, R., Refsdal, S., 1993, *A&A*, 268, 501
- Witz, S.W., Momjian, E, 2016,  
<https://science.nrao.edu/facilities/vla/docs/manuals/oss/performance/sensitivity>
- Wyithe, J. S. B., Turner, E. L., 2002b, *ApJ*, 575, 650
- Young, P., 1981, *ApJ*, 244, 756
- Zackrisson, E., 2005, *Quasars and Low Surface Brightness Galaxies as Probes of Dark Matter*. Uppsala Univ., Sweden
- Zackrisson, E., Bergvall, N., 2003, *A&A*, 399, 23
- Zackrisson, E., Bergvall, N., Helbig, P., 2002, Chapter 87, *Constraining Dark Matter with the Long-Term Variability of quasars*, in *Lighthouses of the Universe: The Most Luminous Celestial Objects and Their Use for Cosmology*, ed. Gilfanov, M., sunyaev, R., Churazov, E., Springer, 550
- Zackrisson, E., Bergvall, N., Marquart, T., Helbig, P., 2003, *A&A*, 408, 17
- Zackrisson, E., Riehm, T., 2007, *A&A*, 475, 453
- Zackrisson, E., Riehm, T., 2010, *Adv. Astron. Article 478910*
- Zardecki, A., Delisle, C., 1974, *Can. J. Phys.*, 52, 1694
- Zhdanov, V. I., 1999, *Astron. Lett.*, 25, 12
- Zhou X. H., Li, C. M., Yang, Z., 2008, *Phil. Trans. R. Soc. A*, 366, 2405

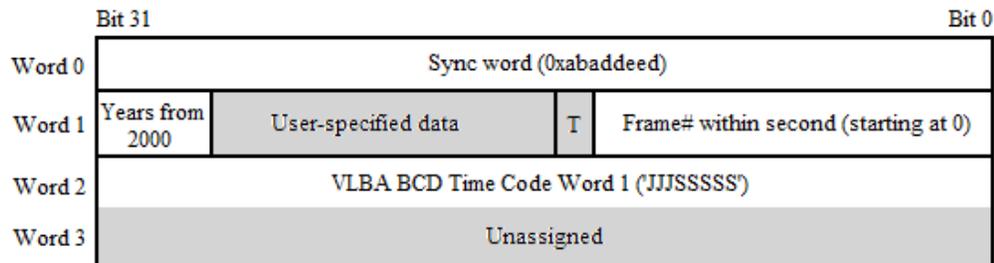
Zwicky, F., 1937a, Phys. Rev., 51, 290

Zwicky, F., 1937b, Phys. Rev., 51, 679

## A Format et tableaux des données

Les données utilisées proviennent d'observations de radiosources par un réseau à très longue base (VLBA pour *Very Long Baseline Array*) exploité par le NRAO (Napier et al. 1994; Romney & Reid 2005). Elles ont été prises dans un mode de fonctionnement standard du VLBA pour l'interférométrie à corrélation-croisée, en segments de temps d'essai (ex., 10 minutes), afin d'évaluer notre technique. Un total de 16 canaux à double polarisation circulaire de 32 MHz de largeur de bande, s'étendent sur des fréquences couvrant une plage contiguë de 8256 à 8512 MHz (voir tableau 5.1).

La structure des données d'entrée du NRAO est de format Mark 5B (voir figure A.1). Cette structure est agencée en trames de données (*data frames*) contenant deux longs paquets *UDPs* (*User Datagram Protocol*). Chaque *data frame* comprend un en-tête de quatre mots suivi de 10000 octets de données brutes. Ces 10000 octets sont décomposés en deux segments de taille presque identique de 5 ko.



**Figure A.1.** En-tête d'un *data frame* de format Mark 5B.-Source (MIT HIT Haystack Observatory Mark Memo #582008).

**Tableau A.1.** Polarisation circulaire droite ou gauche (R ou L), fréquence maximale et minimale d'observation des 16 canaux du système de données VLBA du NRAO.

Numéro du canal	Polarisation circulaire	Fréq. max (MHz)	Fréq. min (MHz)
1	R	8512	8480
2	L	8512	8480
3	R	8480	8448
4	L	8480	8448
5	R	8448	8416
6	L	8448	8416
7	R	8416	8384
8	L	8416	8384
9	R	8384	8352
10	L	8384	8352
11	R	8352	8320
12	L	8352	8320
13	R	8320	8288
14	L	8320	8288
15	R	8288	8256
16	L	8288	8256

Largeur de bande des seize canaux: 32 MHz

Bande passante totale: 2 X 256 MHz = 512 MHz

Pas échantillonnage: 15.625 ns

Fréquence d'observation moyenne: 8.4 GHz

**Tableau A.2.** Nom (commun) des 22 sources radio observées par le NRAO. Chaque objet est identifié par ses coordonnées équatoriales, son redshift, le redshift d'une galaxie lentille associée (si présente) et son flux total moyen.

<b>Objet</b>	<b>R.A. (J2000)</b> <b>(hh mm ss)</b>	<b>Dec (J2000)</b> <b>(dd mm ss)</b>	<b>Redshift</b>	<b>Redshift</b> <b>(macrolentille)</b>	<b>S</b> <b>(mJy)</b>
PKS 1622-253	16 25 46.892	-25 27 38.327	0.786		1918
4C +01.02	01 08 38.771	+01 35 00.317	2.099		2184
0957+561AB	10 01 20.691	+55 53 55.61	1.413	0.355	29
OJ 287	08 54 48.875	+20 06 30.641	0.306		3263
3C 273	12 29 06.700	+02 03 08.598	0.1583		19722
3C 279	12 56 11.167	-05 47 21.525	0.536		18893
OT 081	17 51 32.819	+09 39 00.728	0.322		3968
3C 345	16 42 58.810	+39 48 36.994	0.5935		8551
NRAO 530	17 33 02.706	-13 04 49.548	0.902		4492
B0218+357	02 21 05.473	+35 42 31.90	0.96	0.685	997
B1030+074	10 33 34.024	+07 11 26.15	1.535	0.599	203
PKS 0528+134	05 30 56.417	+13 31 55.150	2.07		3517
4C +73.18	19 27 48.495	+73 58 01.570	0.302		3603
PMN 0414+0534	04 14 37.550	+05 34 42.361	2.639	0.9584	559
PMN 0134-0931	01 34 35.6668	-09 31 02.873	2.225	0.7645	228
B1422+231	14 24 38.0940	+22 56 00.590	3.62	0.338	384
PKS 1830-211	18 33 39.92	-21 03 39.90	2.507	0.886	10000
4C +38.41	16 35 15.493	+38 08 04.501	1.813		2830
NRAO 150	03 59 29.747	+50 57 50.162	1.52		8633
PKS 2134+004	21 36 38.586	+00 41 54.213	1.932		6030
3C 454.3	22 53 57.748	+16 08 53.561	0.859		10060
3C 418	20 38 37.035	+51 19 12.663	1.686		2745

Notez que tous les objets à image simple (sans macrolentille) sont très brillants. Leur flux moyen,  $S \sim 1$  à  $10$  Jy, dépasse de trois à quatre ordres de grandeur l'intensité du bruit du système VLBI du NRAO valant 1 mJy (voir section 2.1.5). Même l'objet le plus faible, le quasar jumeau 0957+561AB, donne un rapport signal à bruit moyen de presque 30.

**Tableau A.3.** Époque et durée d'observation des 22 radiosources.

<b>Objet</b>	<b>Époque</b>	<b>Durée</b> <b>(secondes)</b>
PKS 1622-253	2013-10-24	50
4C +01.02	2013-11-27	180
0957+561AB	2014-05-25	180
OJ 287	2014-09-24	600
3C 273	2014-09-24	600
3C 279	2014-09-26	600
OT 081	2014-09-26	600
3C 345	2015-02-16	600
NRAO 530	2015-02-17	600
B0218+357	2015-06-04	600
B1030+074	2015-06-04	600
PKS 0528+134	2015-07-30	600
4C +73.18	2015-08-01	600
PMN 0414+0534	2016-05-22	650
PMN 0134-0931	2016-05-24	650
B1422+231	2016-08-05	600
PKS 1830-211	2016-08-05	600
4C +38.41	2016-11-23	600
NRAO 150	2016-11-23	600
PKS 2134+004	2016-11-25	600
3C 454.3	2016-11-28	600
3C 418	2016-11-28	600

**Tableau A.4.** Lieu, code, coordonnées géographiques et élévation des sept stations VLBA utilisées par le NRAO pour observer les 22 objets. – Source (site web du NRAO).

<b>Lieu</b>	<b>Code</b> <b>(XX)</b>	<b>Latitude Nord</b> <b>(° ' ")</b>	<b>Longitude Ouest</b> <b>(° ' ")</b>	<b>Élévation</b> <b>(m)</b>
Brewster, WA	BR	48:07:52.42	119:40:59.80	250
Hancock, NH	HN	42:56:00.99	71:59:11.69	296
Kitt Peak, AZ	KP	31:57:22.70	111:36:44.72	1902
Mauna Kea, HI	MK	19:48:04.97	155:27:19.81	3763
Owens Valley, CA	OV	37:13:53.95	118:16:37.37	1196
North Liberty, IA	NL	41:46:17.13	91:34:26.88	222
Saint Croix, VI	SC	17:45:23.68	64:35:01.07	-15

**Tableau A.5.** Code des stations VLBA, nom et taille des fichiers d'entrée des 22 objets.

<b>Objet</b>	<b>Code (XX)</b>	<b>Nom des fichiers</b>	<b>Taille (Go)</b>
PKS 1622-253	MK, SC	BL175BG_xx.dat	13
4C +01.02	BR, HN, OV, SC	xx_tb018_0002.m5	45
Q0957+561AB	BR, HN, OV, SC	xx_tb018_0005.m5	45
OJ 287	BR, HN, OV, SC	xx_tb020_0001.m5	150
3C 273	BR, HN, OV, SC	xx_tb020_0002.m5	150
3C 279	BR, HN, OV, SC	xx_tb020_0003.m5	150
OT 081	HN, SC	xx_tb020_0004.m5	150
3C 345	BR, HN, OV, SC	tb021_xx_0002.m5	150
NRAO 530	HN, SC	tb021_xx_0003.m5	150
B0218+357	BR, HN, KP, SC	tb022_xx_a.m5	150
B1030+074	BR, HN, KP, SC	tb022_xx_b.m5	150
PKS 0528+134	BR, HN, KP, SC	tb022_xx_e.m5	150
4C +73.18	BR, HN, OV, SC	tb022_xx_f.m5	150
PMN 0414+0534	BR, HN, KP, OV, SC	tb022j_xx_02.m5	160
PMN 0134-0931	BR, HN, KP, OV, SC	tb022j_xx_03.m5	160
B1422+231	BR, HN	tb022m_xx.m5	150
PKS 1830-211	BR, HN	tb022n_xx.m5	150
4C +38.41	BR, HN, NL, OV	tb022p_xx_no0001.m5	150
NRAO 150	BR, HN, NL, OV	tb022q_xx_no0001.m5	150
PKS 2134+004	BR, HN, OV	tb022r_xx_no0001.m5	150
3C 454.3	BR, HN, OV	tb022r_xx_no0002.m5	150
3C 418	BR, HN, OV	tb022r_xx_no0003.m5	150



## B Méthodes d'intervalles de confiance d'une binomiale

Rappelons que le principe général d'un intervalle de confiance consiste à construire, à partir des résultats obtenus d'un échantillon-test, un intervalle dans lequel la caractéristique d'intérêt, au sein de la population dont on a extrait l'échantillon-test, a beaucoup de chances d'appartenir. Dans notre cas, il s'agit de confectionner un intervalle de confiance, à un niveau de risque  $\alpha = 5\%$ , connaissant la proportion observée d'événements de minilentilles ( $p_{\text{obs}} = 0$ ) dans notre petit échantillon-test ( $n = 22$ ), dans lequel la proportion théorique inconnue  $p$  se situe avec une probabilité  $(1 - \alpha) = 0.95$ . À ce propos, les sections qui suivent résument trois méthodes courantes différentes pour calculer l'intervalle de confiance d'une binomiale d'un échantillon de petite taille ( $n < 20 - 30$ ).

### B.1 Intervalle de Clopper-Pearson

L'intervalle de Clopper-Pearson est une méthode régulièrement utilisée afin de calculer des intervalles de confiance pour une binomiale. On dit souvent de cette méthode qu'elle est "exacte", car elle se base précisément sur les probabilités cumulatives de la loi binomiale plutôt que sur des probabilités issues de distributions approximant la binomiale, comme la loi normale pour de grands échantillons. N'empêche que les intervalles de confiance ne sont pas vraiment exacts en ce sens que la nature discrète de la loi binomiale exclut tous les intervalles recouvrant exactement toutes les proportions  $p$ .

L'intervalle de confiance "exact" bilatéral symétrique de Clopper-Pearson d'une binomiale, noté  $IC_{C-P}$ , au seuil de risque de  $\alpha\%$ , est un intervalle s'étendant de  $p_{li}$  à  $p_{ls}$  et qui satisfait les deux restrictions suivantes (Krishnamoorthy & Peng 2007; Thulin et al. 2013; Wallis 2013),

$$\sum_{i=0}^k \frac{n!}{i!(n-i)!} p_{ls}^i (1-p_{ls})^{n-i} = \alpha/2, \quad (\text{B.1a})$$

$$\sum_{i=k}^n \frac{n!}{i!(n-i)!} p_{li}^i (1-p_{li})^{n-i} = \alpha/2, \quad (\text{B.1b})$$

où  $p_{li}$  et  $p_{ls}$  sont respectivement les limites inférieures et supérieures des proportions de l' $IC_{C-P}$ , tandis que  $k$  est le nombre de réussites dans l'échantillon  $n$ . En général, le calcul des valeurs de  $p_{li}$  et de  $p_{ls}$  est non trivial et requiert d'utiliser un ordinateur. Heureusement, puisque les équations (B.1a) et (B.1b) reposent sur la fonction de distribution cumulative de la binomiale, il devient alors possible de simplement utiliser les quantiles de la distribution continue bêta pour représenter l'intervalle de Clopper-Pearson de la binomiale. Dans ce cas, l'intervalle s'écrit sous la forme suivante,

$$IC_{C-P} = [p_{li}, p_{ls}], \quad (B.2)$$

où les valeurs  $p_{li}$  et  $p_{ls}$  de l' $IC_{C-P}$  sont calculées par (Thulin et al. 2013; Wallis 2013),

$$p_{li} = \text{Bêta}(\alpha/2, k, n - k + 1), \quad (B.3a)$$

$$p_{ls} = \text{Bêta}(1 - \alpha/2, k + 1, n - k), \quad (B.3b)$$

avec Bêta ( $q, a, b$ ) représentant le  $q^{\text{ième}}$  quantile d'une loi bêta de paramètres de formes  $a$  et  $b$ .

Avec notre échantillon-test, on n'a obtenu aucune détection ( $k = 0$  succès). De là, l'intervalle de confiance se réduit à (Olivier & May 2006; Pires & Conceição 2008),

$$IC_{C-P} = [p_{li}, p_{ls}] = [0, 1 - (\alpha/2)^{1/n}]. \quad (B.4)$$

## B.2 Intervalle de Wilson

L'intervalle de confiance de Wilson d'une binomiale, noté  $IC_w$ , au seuil de risque de  $\alpha$  %, s'écrit sous la forme suivante (Brown et al. 2001; Olivier & May 2006; Krishnamoorthy & Peng 2007; Zhou et al. 2008; Thulin et al. 2013; Wallis 2013),

$$IC_w = \frac{1}{1 + \frac{1}{n} z_{\alpha/2}^2} \left[ p_{obs} + \frac{1}{2n} z_{\alpha/2}^2 \pm z_{\alpha/2} \left( \frac{1}{n} p_{obs} (1 - p_{obs}) + \frac{z_{\alpha/2}^2}{4n^2} \right)^{1/2} \right], \quad (B.5)$$

où  $z_{\alpha/2}$  est le  $1 - \alpha/2$  quantile de la distribution normale centrée à 0 et réduite à 1. Avec un risque  $\alpha = 5$  %, on tire de la table de la loi normale standard,  $z_{\alpha/2} \equiv z_{0.025} = 1.96$ . En insérant cette valeur de  $z_{0.025}$ , ainsi que les valeurs  $n = 22$  et  $p_{obs} = 0$  dans l'expression (B.5), on trouve,  $IC_w = 0.0745 \pm 0.0745$ .

### B.3 Intervalle de Jeffreys

À l'instar de l'intervalle de Clopper-Pearson, on utilise les quantiles de la distribution continue bêta pour représenter l'intervalle de Jeffreys d'une binomiale, noté  $IC_J$ , à un niveau de risque de  $\alpha$  %. Par contre, avec l' $IC_J$ , les deux équations donnant les limites inférieures et supérieures ( $p_{li}$  et  $p_{ls}$ ) sont différentes des équations (B.3a) et (B.3b) de l' $IC_{C-P}$ . Cela dit, les relations donnant  $p_{li}$  et  $p_{ls}$  s'écrivent respectivement pour l' $IC_J$  (Brown et al. 2001; Zhou et al. 2008; Thulin et al. 2013),

$$p_{li} = \text{Bêta} (\alpha/2, k + 1/2, n - k + 1/2), \quad (\text{B.6a})$$

$$p_{ls} = \text{Bêta} (1 - \alpha/2, k + 1/2, n - k + 1/2). \quad (\text{B.6b})$$

Pour  $k = 0$  succès, la valeur de  $p_{li}$  est fixée à 0 afin d'éviter des résultats non désirés comme la probabilité de couverture tendant vers zéro (Brown et al. 2001). La valeur de  $p_{ls}$  requière d'être calculée numériquement ou d'utiliser des tables. Ceci dit, avec notre échantillon-test,  $n = 22$  et  $k = 0$ , on tire de la table 5 de Brown et al. (2001),  $p_{ls} = 0.107$  pour  $\alpha = 5$  %.



# C Code Matlab

## C.1 Programme appelant

```
%=====
%Programme permettant d'exécuter l'autocorrélation des intensités.
%
%=====DÉBUT DU PROGRAMME PRINCIPAL=====

clear all;
close all;
clc;

%Lecture du fichier d'entrée (ex., l'objet source 3C 345).
filename = 'tb021_sc_0002.m5';

%Nombre de data frames traités (1000).
number_df = 1000;

%Nombre de canaux total (16).
number_ch = 16;

%Appel de la fonction dread_mk5(...) qui retourne un tableau:
%nombre de data frames X nombre de canaux (16) et un vecteur
%contenant les numéros des data frames (0, 1, 2, ...).
[data_all_chan, df_number] =
dread_mk5(filename, number_df, number_ch);

%Appel de la fonction replace_2bits_values(...) qui remplace
%les quatre valeurs positives sur deux bits (3, 1, 2 et 0)
%par les quatre valeurs signées des champs électriques
%(3.3359, 1, -1 et -3.3359).
tab_E_signed_values = replace_2bits_values(data_all_chan, number_ch);

%Calul des intensités de tous les canaux (16).
for i = 1:number_ch
    tab_I(:,i) = tab_E_signed_values(:,i).^2;
end

%=====TRAITEMENT D'UN SEUL CANAL=====

%Extraire les intensités d'un seul canal(ex., canal 1 de 3C 345).
I = tab_I(:,1);

%Lissage (filtre RIF) des intensités dont le nombre de points
%correspond à environ la moitié du milieu intergalactique
%(7000 points pour le quasar 3C 345).
I = smooth(I,7000);
```

```

%Appel de la fonction prédéfinie xcorr(...) pour calculer
%l'autocorrélation de l'intensité en enlevant la moyenne.
xcorr_I_moy = xcorr(I-mean(I));

%Abscisse des temps (secondes) pour les 1000 dataframes
%(25 millions de points).
t = linspace(-15.625e-9*(length(I)-1),15.625e-9*length(I),
2*length(I)-1);

% Vecteurs colonnes.
I = I(:);
xcorr_I_moy = xcorr_I_moy(:);
t = t(:);

%Affichage de l'autocorrélation d'intensité en fonction du temps.
figure (1);
plot(t,xcorr_I_moy);
axis tight;
xlabel ('temps (secondes)');
ylabel ('Autocorrélation (u.a.)');

%=====PARTIE OPTIONNELLE: Spectre de puissance=====
%Appel de la fonction prédéfinie fft(...) (fast fourier transform)
%pour calculer la transformée de Fourier (TF) de l'autocorrélation
%d'intensité.
FFT_xcorr_I_moy = (fft(xcorr_I_moy));

%Spectre de puissance calculé avec le module (abs (...)) de la TF.
pow_spec = abs(FFT_xcorr_I_moy);

%Calcul de la fréquence d'échantillonnage (domaine de Fourier).
fs = 1/(max(t) - abs(min(t)));

%Abscisse des fréquences en Hz.
f = linspace(0,fs*length(pow_spec)/2,length(pow_spec)/2);

% Vecteurs colonnes.
FFT_xcorr_I_moy = FFT_xcorr_I_moy(:);
pow_spec = pow_spec(:);
f = f(:);

%Affichage du spectre de puissance (partie positive) en fonction de
%la fréquence (Hz).
figure (2);
pow_spec = pow_spec(:);
plot(f,pow_spec(1:floor(length(pow_spec)/2)));
axis tight;
xlabel ('Fréquence (Hz)');
ylabel ('Spectre de puissance (u.a.)');
exit;

=====FIN DU PROGRAMMEAPPELANT=====

```

## C.2 Fonction de lecture dread\_mk5

```
%=====
%La fonction dread_mk5(...) permet de lire les données des data
frames de %format Mark 5B.
%
%Entrée: filename:          nom du fichier
%         number_dataframes:  nombre de data frames à lire
%         number_channels = 16:  nombre de canaux
%
%Sortie: da_ch: nombre de données lues dans un tableau
%           de (nombre de data frames) X 16
%         df_nu: numéro du data frame
%=====
%Début

function [da_ch,df_nu] =
dread_mk5(filename,number_dataframes,number_channels)

%Ouvrir le fichier d'entrée pour la lecture en format big-endian.
fid = fopen(filename,'r','b');

%Déclaration et initialisation d'un tableau vide de 16 colonnes qui
%contiendra les données des 16 canaux d'un data frame.
%number_channels = 16.
channel_dataframe = zeros(1,number_channels);

%Boucle pour lire les données de chacun des data frames
%du fichier d'entrée.
for inc_df = 1:number_dataframes

%=====LECTURE DE L'ENTÊTE DE 16 OCTETS DE CHAQUE DATA FRAME=====
%Lire le premier mot (32 bits), Sync word, de chaque data frame en
%convertissant les entiers décimaux (non signés sur 32 bits) en
%chaîneshexadécimales. L'option 'ieee-le' permet de convertir le
%format big-endian en format little-endian.
SyncWord(inc_df,:) = dec2hex(fread(fid,1,'uint32','ieee-le'),8);

%Lire le second mot (4 octets) en format little-endian.
RawWord = fread(fid,1,'ubit32','ieee-le');

%Construire l'objet numérique à virgule fixe, RawWord_q, à partir du
%mot de 32 bits, RawWord.
RawWord_q = fi(RawWord,0,32,0);

%Extraire le champ d'identification 'Years from 2000' qui commence
%au bit 28 (position 29) et se termine au bit 31(position 32).
YearsFrom2000(inc_df) = bitsliceget(RawWord_q,32,29);

%Convertir le champ 'Years from 2000' (contenant 4 bits) en chaîne
%hexadécimale.
YearsFrom2000_h = YearsFrom2000.hex;
```

```

%Extraire le champ d'identification 'User-specified data' commençant
%au bit 16 (position 17) et se terminant au bit 27(position 28).
UserSpecifiedData = bitsliceget (RawWord_q,28,17);

%Convertir le champ 'User-specified data' en chaîne hexadécimale.
UserSpecifiedData_h = UserSpecifiedData.hex;

%Extraire le bit 15 représentant un bit de validation d'un data
%frame.
T = bitsliceget (RawWord_q,16,16);

%Extraire le numéro du data frame qui va du bit 0 au bit 15.
frame_number (inc_df) = bitsliceget (RawWord_q,16,1);

%Convertir le numéro du data frame en chaîne hexadécimale.
frame_number_h = frame_number.hex;

%Lire le troisième mot (4 octets) en format little-endian.
RawWord1 = fread(fid,1,'ubit32','ieee-le');

%Construire l'objet numérique à virgule fixe, VLBA TimeCode, à partir
%du mot de 32 bits, RawWord1.
VLBA TimeCode = fi (RawWord1,0,32,0);

%Extraire (sur 1 octet) les trois derniers chiffres JJJ du jour
%julien de l'observation de l'objet source.
VLBA TimeJJJ (inc_df) = bitsliceget (VLBA TimeCode,32,23);

%Extraire (sur 3 octets) les cinq chiffres SSSSS représentant
%les secondes de la journée de l'observation de l'objet source.
VLBA TimeSSS (inc_df) = bitsliceget (VLBA TimeCode,24,1);

%Lire le quatrième mot et construire l'objet numérique à virgule
%fixe, Unassigned.
Unassigned = fi (fread(fid,1,'uint32','ieee-le'),0,32,0);

%=====

%===LECTURE DU PREMIER SEGMENT DE 4992 OCTETS DE CHAQUE DATA FRAME===
%Lire 4992(octets) = 4992(octets)*8(bits)/2(bits) = 19968 valeurs
%sur 2 bits.
[channel_dataframe_temp, NumRead] =
fread(fid, 19968,'ubit2','ieee-le');

%Convertir le vecteur de 19968 valeurs en un tableau de 1248 X 16
%valeurs (number_channels = 16).
channel_dataframe_temp =
vec2mat (channel_dataframe_temp,number_channels);

%Inverser l'ordre des canaux: le canal 16 devient le canal 1,
%le canal 15 %devient canal 2, etc.
channel_dataframe_firsthalf =
channel_dataframe_temp (:,number_channels:-1:1);

```

```

%Mettre le premier segment de données (1248 X 16) dans le tableau
%des données des seize canaux.
channel_dataframe = [channel_dataframe;
channel_dataframe_firsthalf];

%==LECTURE DU SECOND SEGMENT DE 5008 OCTETS DE CHAQUE DATA FRAME===
%Lire 5008(octets) = 5008(octets)*8(bits)/2(bits) = 20032 valeurs
%sur 2 bits.
[channel_dataframe_temp, NumRead] =
fread(fid, 20032, 'ubit2', 'ieee-le');

%Convertir le vecteur de 20032 valeurs en un tableau de 1252 X 16
%valeurs (number_channels = 16).
channel_dataframe_temp =
vec2mat(channel_dataframe_temp, number_channels);

%Inverser l'ordre des canaux: le canal 16 devient le canal 1,
%le canal 15 %devient canal 2, etc.
channel_dataframe_secondhalf =
channel_dataframe_temp(:, number_channels:-1:1);

%Ajouter (concaténer) le deuxième segment au tableau de données des
%16 canaux. La première valeur de chaque canal étant nulle.
channel_dataframe =
cat(1, channel_dataframe, channel_dataframe_secondhalf );
end
%Fermer le fichier.
fclose(fid);

%Nombre de données pour tous les data frames, mais pour un seul
%canal.
da_ch_len = length(channel_dataframe);

%Déclaration et initialisation d'un tableau vide contenant les
%données de tous les data frames passés en paramètres (les rangées)
%pour les 16 canaux (les colonnes).
da_ch = zeros(da_ch_len, number_channels);

%Remettre les 16 canaux de tous les data frames dans l'ordre.
%La première valeur de chaque canal étant nulle.
for ii = 1:number_channels
    %Retourne un tableau de données: nombre de data frames X 16.
    da_ch(:, ii) = channel_dataframe(:, (number_channels+1-ii));
end

%Retourne un vecteur contenant les numéros des data frames (0, 1,
%2, ...).
df_nu = FrameNum.data;

%Fin

```

### C.3 Fonction de conversion `replace_2bits_values`

```
%=====
%La fonction replace_2bits_values(...) permet de remplacer les quatre
%valeurs positives sur deux bits (3, 1, 2 et 0) par les quatre
%valeurs %signées des champs électriques(3.3359, 1, -1 et -3.3359).
%
%Entrée: data_all_channels:    nombre de données totales
%                               (dans les 16 canaux) codées
%                               sur 2 bits (00, 01, 10, 11)
%
%                               number_channels = 16: nombre de canaux
%
%Sortie: Données converties en champs électriques signés
%        (3.3359, 1, -1, -3.3359)
%=====
%Début

function                               [tab_E_signed_values]           =
replace_2bits_values (data_all_channels, number_channels)

%Déclaration et initialisation d'un tableau de dimension
%nombre de données par canal X nombre de canal (16).
[size_tab_E_signed_values, ~] = size(data_all_channels);
tab_E_signed_values           =
zeros(size_tab_E_signed_values, number_channels);

%Parcourir tous les canaux (16).
for jj = 1:number_channels
    %Parcourir toutes les données (champs électriques) d'un canal.
    for ii = 1:size_tab_E_signed_values
        %Remplacer la valeur sur deux bits (11) 3 par 3.3359.
        if(data_all_channels(ii, jj) == 3)
            tab_E_signed_values(ii, jj) = 3.3359;
        %La valeur sur deux bits (01) 1 reste 1.
        elseif(data_all_channels(ii, jj) == 1)
            tab_E_signed_values(ii, jj) = 1;
        %Remplacer la valeur sur deux bits (10) 2 par -1.
        elseif(data_all_channels(ii, jj) == 2)
            tab_E_signed_values(ii, jj) = -1;
        %Remplacer la valeur sur deux bits (00) 0 par -3.3359.
        else
            tab_E_signed_values(ii, jj) = -3.3359;
        end
    end
end

%Fin
```

## C.4 Programme pour extraire les pics des intensités auto-corrélées

```
%=====
%Programme permettant de récupérer les données (intensités auto-
%corrélées) traitées en C++ et d'extraire les pics ayant un rapport
%signal à bruit (S/B) supérieur à 4.0. Pour évaluer le bruit, ce
%programme fait notamment appel à la RÉGRESSION LINÉAIRE pour
%rectifier le profil d'autocorrélation.
%=====
%DÉBUT

clear all;
close all;
clc;

%Pas d'échantillonnage correspondant à la moitié du nombre de
%points dû à l'espace intergalactique (ex., 7000).
STRIDE = 7000;
%Position de départ.
START = 2*STRIDE;
%Nombre d'éléments à lire (taille du bloc de données à lire).
COUNT = floor(3.84e10./STRIDE);

%Appel de la fonction prédéfinie h5read(...) pour récupérer
%les intensités auto-corrélées en C++ dans le vecteur ordonné y.
y = h5read('xc_tb021_sc_0002_m5.h5','/Autocorrelation',
START+1,COUNT,STRIDE);

%Déclarer un vecteur abscisse x contenant le même nombre d'éléments
%que le vecteur ordonné y (ex., 3.84e10./STRIDE = 5.485e6 éléments).
x = 1:COUNT;
x = transpose(x);

%Nombre de blocs de données (ex., 500).
n = 500;
%Nombre d'éléments dans un bloc de données (ex., 10970).
M = floor(COUNT./n);
%Redéclarer le vecteur abscisse x contenant le même nombre
%d'éléments qu'un bloc de données.
x = 1:M;
x = transpose (x);

%Parcourir les n (ex., 500) blocs de données.
for i = 1:n

    %Placer les éléments du vecteur ordonné y dans un tableau M X n:
    %nombre d'éléments par bloc X nombre de blocs.
    mat_y(:,i) = y(i*M-M+1:i*M);

    %Calculer l'ordonnée à l'origine b et la pente m.
    b_m = [ones(length(x),1) x] \ mat_y(:,i);

    %Extraire l'ordonnée à l'origine b du bloc i.
    b(i) = m(1,1);
end
```

```

%Extraire la pente m du bloc i.
m(i) = b_m(2,1);

%RÉGRESSION LINÉAIRE.
mat_yfit(:,i) = m(i)*x+b(i);

%Calculer la matrice de résidus: données - droite de régression
%linéaire.
mat_r(:,i) = mat_y(:,i)-mat_yfit(:,i);

%Calculer l'écart-type de la matrice des résidus.
sigma_r(i) = std(mat_r(:,i));
end
sigma_r = transpose (sigma_r);

%Convertir la matrice des droites de régression en un vecteur droite
%de régression linéaire yfit contenant COUNT (ex.,5.485e6) éléments.
yfit = reshape (mat_yfit,COUNT,1);

%Convertir la matrice de résidus en un vecteur résidu r contenant
%COUNT(ex, 5.485e6) éléments.
r = reshape (mat_r,COUNT,1);

%Déclaration d'un vecteur qui contiendra les pics extraits
%(position, S/B et amplitude).
posPicMAX = zeros(1,3);
iii = 0;

%Parcourir les n (ex., 500) blocs de données.
for i = 1:n;
    %Parcourir les M (ex., 10970) éléments d'un bloc.
    for ii = i*M-M+1:i*M

        %Extraire les pics excédant 4 fois le bruit dans le vecteur r.
        if(r(ii) > 4*sigma_r(i))

            %Compter les pics extraits.
            iii = iii+1;
            %Position du pic sorti.
            posPicMAX(iii,1) = (STRIDE*ii)-START+1;
            %S/B du pic sorti.
            posPicMAX(iii,2) = r(ii)./sigma_r(i);
            %Amplitude du pic sorti.
            posPicMAX(iii,3) = r(ii);

        end

    end

end

%Ouverture d'un fichier en écriture.
fid = fopen('FILE_DETEC_BOSES.txt', 'w');

```

```

%Vérifier s'il y a des pics extraits.
if ((posPicMAX(1,1) ~= 0) || (posPicMAX(1,2) ~= 0))
    tailleposPicMAX = size(posPicMAX);
    nLignes = tailleposPicMAX(1,1);
    if (nLignes >= 1)
        %Parcourir le vecteur de pics sortis.
        for j = 1:nLignes
            %Écrire la position, le S/B et l'amplitude.
            fprintf(fid, '%i %5.4f %14E %d\n', posPicMAX(j,1),
                posPicMAX(j,2), posPicMAX(j,3), j);
        end
    end

    %Fermeture du fichier ouvert en écriture.
    fclose('all');
end

%Affichage de l'autocorrélation d'intensité en fonction de la
%position en points.
figure (1);
plot (y);
axis tight;
xlabel ('Position (pts)');
ylabel ('Autocorrélation (a.u.)');

```



## D Code orienté objet C++

Notez que pour développer le code C++, utilisé avec le superordinateur Colosse de l'Université Laval, on a collaboré avec des experts en calcul haute performance (CHP) de Calcul Québec.

- 1) On a d'abord adapté le code Matlab en parties de code C++ provisoires en utilisant Microsoft Visual Studio C++ 2012 de Windows sur un microordinateur (à mémoire vive limitée). Par exemple, plusieurs parties du code du fichier `dread_mk5.cpp` ont au préalable été implantées avec Visual Studio C++ sur un microordinateur, puis corrigées et adaptées afin de s'exécuter sur le superordinateur.
- 2) Les experts de Calcul Québec ont ensuite corrigés ces parties de code C++ provisoires et développés d'autres parties de code C++ afin de pouvoir exécuter un programme C++ (`mark5b_autocorr.cpp`) sur le superordinateur avec le système d'exploitation Linux. En fait, toutes les parties du code impliquant le traitement en parallèle de données (sur plusieurs nœuds de calcul), la gestion optimale de la mémoire, de même que les applications MPI (Message Passing Interface) ont principalement été implantées par les experts de Calcul Québec. Notez enfin que tout le code C++ développé (ex., toutes les classes) est disponible sur demande.

### D.1 Fichier `dread_mk5.cpp` pour la lecture des données

```
//=====fichier dread_mk5.cpp=====
//Appel de 11 librairies C++.

#include <iostream>//Pour le traitement du flux du fichier de sortie
#include <fstream>//Pour le traitement des flux (fichiers) d'entrée/sortie
#include <vector>//Pour l'allocation et le traitement de tableaux dynamiques
#include <bitset>//Pour le traitement des tableaux de bits
#include <sstream>//Pour le traitement de flux des chaînes de caractères
#include <iomanip>//Pour manipuler les formats d'entrée/sortie
#include <string>//Pour le traitement de chaînes de caractères
#include <stdio.h>//Pour les opérations d'entrée/sortie
#include <stdlib.h>//Pour exécuter diverses opérations (conversion, allocation
//de mémoire, contrôle de processus, gérer l'environnement, etc.)
#include <cstdlib>//Pour utiliser le fichier d'en-tête stdlib.h dans l'espace
//des noms (namespace)
#include <hex_binary_utils.h>//Pour la conversion hexadécimal-binaire

//Se placer dans l'espace de noms std pour utiliser les symboles déclarés dans
//le fichier iostream (librairie <iostream>) et pour résoudre les conflits de
//nom.
using namespace std;
```

```

//=====
//Déclaration de la fonction inline(pour accélérer le programme) mask(...).
//ENTRÉE : const char * c. (1 octet). Pointeur (vecteur) sur une chaîne de
// caractères constante.
//SORTIE : unsigned char (1 octet). Retourne un caractère non signé qui est
// défini par le patron de la classe bitset<8>. Il y a d'abord une
// conversion en entier long non signé (4 octets), dont la valeur
// entière est définie par le patron de classe bitset<8>. Ensuite, il y
// a une conversion de la valeur entière de type entier long non signé
// au type caractères non signé (1 octet).
//=====
inline unsigned char mask(const char * c) { return (unsigned
char)bitset<8>(string(c)).to_ulong(); }

//Déclaration d'une structure nommée format contenant 5 champs de type entier.
struct format {
int df_header_size; //Taille de l'en-tête d'un data frame
int df_data_size; //Taille d'un data frame sans l'en-tête
int nblocks; //Nombre de blocs par data frame
int bit_per_chan; //Nombre de bits pour représenter les données
//((champs électriques) d'un canal
int num_chan; //Numéro d'un canal
};

//Déclaration d'une énumération nommée mode.
enum mode { SUM=0, SINGLE_CHAN=1 };

//Déclaration d'une constante statique format_mark_5b de type format. La
//constante est affectée des 5 valeurs (16 octets, 10000 octets, 2 blocs, 2
//bits et 16 canaux) pour les 5 champs de la structure format.
static const struct format format_mark5b = {16,10000,2,2,16};

//Déclaration de la constante statique (permanente) SUM de type entier.
//La valeur de SUM est -1. (-1 indique la somme des 16 canaux).
static const int SUM = -1;

//=====
//Déclaration de la fonction extract_data_from_Mark5B_dataframe(...).
//Extraire les données d'un data frame dans le format Mark 5B.
//Le format d'un data frame est spécifié par le format de struct passé en
//paramètre. La longueur de raw_data est fmt.df_header_size + fmt.df_data_size,
//qui contient 16 + 10000 = 10016 caractères pour le format Mark5B.
//ENTRÉE : const struct format & fmt = format_mark_5b. Déclaration d'une
// constante fmt de type format et initialiser à format_mark5b.
// const char * raw_data (1 octet). Déclaration d'un pointeur constant
// (vecteur) sur des caractères. La taille du vecteur raw_data est de
// 16 octets + 10000 octets = 10016 octets.
//SORTIE : unsigned char * processed_data (1 octet). Déclaration d'un pointeur
// (vecteur) des caractères non signés. Le vecteur contient les valeurs
// (-3.3359, -1, 1, 3.3359) de chaque canal pour chaque pas de temps
// d'un data frame.
// unsigned short * frame_number (2 octets). Déclaration d'un pointeur
// (vecteur) sur des entiers courts non signés. Contiendra la valeur
// extraite du numéro d'un data frame.
//=====
void extract_data_from_Mark5B_dataframe(const char * raw_data, unsigned short *
frame_number, unsigned char * processed_data, const struct format & fmt =
format_mark5b) {

```

```

//Déclaration et affectation d'un data frame avec un en-tête
//(10016 octets) de type entier.
//Taille d'un data frame complet : taille de l'en-tête du data frame +
//taille des données actuelles (champs électriques) du data frame.
int full_df_size = fmt.df_header_size + fmt.df_data_size;

//Déclaration d'un compteur d'octets de type entier (2 octets) pour
//un data frame donné du fichier d'entrée.
    int file_ptr = 0;

//Ignorer les 4 premiers octets de l'en-tête (Sync word).
    file_ptr += 4;

//Le numéro du data frame s'obtient en ajoutant les 2 prochains octets
//(15 prochains bits) dans le format little-endian.
//Déclaration du vecteur raw_data de type entier non signé (2 octets).
//Déclaration du pointeur*frame_number (vecteur) sur des entiers non
//signés(2 octets).
//*frame_number = 4 octets + 1 octet + 1 octet (1 bit concaténé à 7
//bits) = 6 octets (taille du vecteur raw_data).
*frame_number = (unsigned short)raw_data[file_ptr] + (unsigned
short)((raw_data[file_ptr+1] & mask("01111111")) * 255);

//Ignorer le champ the "Years from 2000".
//Incrémenter de 2 octets.
file_ptr += 2;
//Ignorer le champ "User-specified data".
//Incrémenter de 2 octets.
file_ptr += 2;
//Ignorer les 2 prochains mots.
//Incrémenter de 8 octets.
file_ptr += 8;
//Déclaration de chan_per_byte de type entier (2 octets).
//Maintenant: le traitement des données actuelles (champs électriques),
//4 canaux par octet (8 bits), soit 16 canaux sur 2 bits.
int chan_per_byte = 8/fmt.bit_per_chan;

//Déclaration et affectation à 0 de mem_ptr de type entier pour
//compter le nombre d'octets (sans l'en-tête) d'un data frame (10000).
int mem_ptr = 0;
//Déclaration de datum de type caractère non signé (1 octet).
unsigned char datum;

//Déclaration de mask_chan_val de type caractère non signé (1 octet).
//Masquer le bit de validité d'un data frame.
unsigned char mask_chan_val = (1 << fmt.bit_per_chan)

//Le premier bloc de 4992 octets contient un bit de moins
//file_ptr = 16 au départ (on a ignoré 16 octets en tout).
//Parcourir les 5008-16 = 4992 octets du premier bloc.
//file_ptr = 4992+16 = 5008 à la fin de la boucle.
//mem_ptr vaut 0 au début et vaut 4992 à la fin de la boucle.
for (; file_ptr<(fmt.df_data_size/fmt.nblocks
fmt.df_header_size/fmt.nblocks); file_ptr++, mem_ptr++)
{
//Extraire les données quantifiées du premier bloc.
datum = (unsigned char)(raw_data[file_ptr]);

```

```

//Parcourir les 16 canaux du premier bloc.
for (int chan_i=0; chan_i<chan_per_byte; chan_i++)
{
//Données converties pour chaque pas de temps de chaque canal
//du premier bloc d'un data frame.
processed_data[mem_ptr*chan_per_byte + chan_i] = (datum >>
(fmt.bit_per_chan*chan_i)) & mask_chan_val;
}
}
//Deuxième bloc de 5008 octets.
//file_ptr = 5008 au départ.
//Parcourir les 5000-16 = 5008 octets du deuxième bloc.
//file_ptr = 5008+5008 = à la fin de la boucle.
//mem_ptr vaut 4992 au début et 4992+5008 = 10000 à la fin de la boucle.

//Parcourir les 5000-16 = 5008 octets du deuxième bloc.
for (; file_ptr<full_df_size; file_ptr++, mem_ptr++)
{
//Extraire une à une les données quantifiées du second bloc.
datum = (unsigned char)(raw_data[file_ptr]);
//Données converties pour chaque pas de temps de chaque canal
//du second bloc d'un data frame.
for (int chan_i=0; chan_i<chan_per_byte; chan_i++)
processed_data[mem_ptr*chan_per_byte + chan_i] = (datum >>
(fmt.bit_per_chan*chan_i)) & mask_chan_val;
}
}
//=====
//Déclaration de la fonction read_mark5b_channel(...).
//Lit un seul canal de Mark 5B. Le canal (numéro = -1) peut aussi lire la somme
//des 16 canaux (valeur par défaut). Lit un nombre donné de data frames.
//Affiche les numéros de data frames et les valeurs des données converties
// (-3.3359, -1, 1, 3.3359).
//ENTRÉE : ifstream & if. Objet if de la classe prédéfinie ifstream pour faire
// des lectures (read(...)) et des accès directs (seekg(...)) dans le
// fichier d'entrée.
// unsigned long num_datadrames. Déclaration d'une variable de type
// entier long non signé (4 octets) contenant le nombre de data frames
// à lire.
// const int channel. Déclaration d'une constante de type entier
// (2 octets) contenant le numéro du canal à traiter. La valeur par
// défaut est le canal = SUM = -1 (somme des 16 canaux).
// unsigned long file_dataframe_offset. Déclaration d'une variable de
// de décalage (offset) (en unités de longueur de data frames) de type
// entier long non signé (4 octets) pour chercher dans le flux (dans le
// fichier ouvert).
// unsigned long memory_dataframe_offset. Déclaration d'une variable de
// de décalage (offset) (en unités de longueur de data frames) de type
// entier long non signé (4 octets) pour chercher dans la mémoire.
// const struct format & fmt = format_mark_5b. Déclaration d'une
// constante fmt de type format et initialiser à format_mark_5b.
//SORTIE: unsigned short & frame_numbers. Déclaration d'un vecteur de type
// entier court non signé (1 octet) contenant le nombre de data frames
// qui ont été extraits.
// float & processed_data. Déclaration d'un vecteur de type simple
// précision (4 octets) contenant les données converties
// (-3.3359, -1, 1, 3.3359) pour un seul canal ou la somme des 16 canaux.
//=====

```

```

void read_mark5b_channel(istream & is, unsigned long num_dataframes,
std::vector<unsigned short> & frame_numbers, std::vector<float> &
processed_data, const int channel=SUM, unsigned long file_dataframe_offset=0,
unsigned long memory_dataframe_offset=0, const struct format & fmt =
format_mark5b) {

//Déclaration de la variable full_df_size de type entier (2 octets) qui
//est affectée de la taille de l'entête (16 octets) et de la taille des
//données actuelles (10000 octets).
int full_df_size = fmt.df_header_size + fmt.df_data_size;
//Message disant de lire un bloc 15000000*10016 = environ 150 milliards
//d'octets, i.e. 37.5 milliards de données de 4 octets chacune.
std::cerr << "Lire un bloc de " << num_dataframes*full_df_size << " octets."
<< std::endl;
//Déplacement vers le décalage initial.
is.seekg(file_dataframe_offset*(full_df_size));

//CETTE PARTIE DU CODE PORTE SUR LA GESTION OPTIMALE DE LA MÉMOIRE=====
//Cette partie du code a principalement été implantée par des experts en CHP
//de Calcul Québec.
//Début
char * buffer = new char[num_dataframes*full_df_size];
is.read(buffer,num_dataframes*full_df_size);
unsigned long ntimesteps_per_dataframe = fmt.df_data_size/(fmt.num_chan
/ (8/fmt.bit_per_chan));
try {
if (processed_data.size() < (unsigned long)(num_dataframes *
ntimesteps_per_dataframe) )
processed_data.resize((unsigned long)(num_dataframes *
ntimesteps_per_dataframe));
} catch (std::bad_alloc) {
std::cerr << "Erreur, pas assez de mémoire." << std::endl;
}
if (frame_numbers.size() < num_dataframes)
frame_numbers.resize(num_dataframes);

//Fin=====

//CETTE PARTIE DU CODE PORTE SUR LE TRAITEMENT EN PARALLÈLE=====
//Partage de la tâche sur plusieurs nœuds.
//Cette partie du code a principalement été implantée par des experts en CHP de
//Calcul Québec.
//Début
unsigned long file_ptr = 0;
unsigned long data_ptr = 0;
unsigned long df_iter = 0;
#ifdef _OPENMP
#pragma omp parallel
shared(ntimesteps_per_dataframe,fmt,processed_data,memory_dataframe_offset,num_
dataframes,buffer,frame_numbers,full_df_size) private(df_iter)
{
#endif
std::vector<unsigned char> extracted_data(ntimesteps_per_dataframe *
fmt.num_chan);
#ifdef _OPENMP
#pragma omp for
#endif
//Fin=====

```

```

//Parcourir tous (15 millions) les data frames du fichier d'entrée.
for (df_iter = 0; df_iter < num_dataframes; df_iter++) {

//Extraire les données de chaque data frame.
//Appel de la fonction qui extrait les données d'un data frame
//dans le format Mark 5B. On passe comme argument à la fonction,
//la référence & (l'adresse) du tampon buffer dont la taille
//correspond est de df_iter*taille d'un data frame, la référence
//& du vecteur frame_number qui contiendra les valeurs extraites
//des numéros de data frames, ainsi que la référence & du premier
//premier élément du vecteur des données converties extracted_data
//qui contiendra la valeur de chacun des canaux pour chaque pas de
//de temps (échantillon) dans le data frame.

extract_data_from_Mark5B_dataframe(
    &(buffer[(df_iter)*full_df_size]),
    &(frame_numbers[(memory_dataframe_offset+df_iter)]),
    &(extracted_data[0])
);

//Déclaration et initialisation d'une variable réelle qui contiendra
//une des quatre valeurs (-3.3359,-1, 1, 3.3359) du champ électrique.
float val = 0;
//Variable contenant une donnée (champ électrique) quantifiée.
unsigned char datum;
//Pour la somme des 16 canaux
if (channel == SUM) {
//Parcourir les (2500) pas de temps par canal par data frame.
for (size_t i=0; i<ntimesteps_per_dataframe; i++)
{
    val = 0;
//Parcourir les 16 canaux d'un data frame.
for (size_t j=0; j<fmt.num_chan; j++)
{
//Extraire une à une les données quantifiées
//des 16 canaux d'un data frame
//(16*2500) = 40000 échantillons.
datum = extracted_data[i*fmt.num_chan + j];

//2 valeurs positives possibles des
//des champs électriques quantifiés.
if (datum == 3) val += 3.3359;
else if (datum == 1) val += 1;
//2 valeurs négatives possibles des
//des champs électriques quantifiés.
else if (datum == 2) val -= 1;
else val -= 3.3359;
}

//Affectation des champs électriques quantifiés (-3.3359, -1, 1, 3.3359)
//au vecteur des données converties processed_data.
processed_data[(memory_dataframe_offset+df_iter)*ntimesteps_per_dataframe + i]
= val;
}
}
//Pour les canaux individuels (pris un à un).
else if (channel <= fmt.num_chan) {

```

```

//Parcourir les (2500) pas de temps par canal par data frame.
for (size_t i=0; i<ntimesteps_per_dataframe; i++)
{
    val = 0;
//Extraire une à une les données quantifiées
//d'un seul canal d'un data frame
//(1*2500) = 2500 échantillons.
        datum = extracted_data[i*fmt.num_chan + channel];

//2 valeurs positives possibles des
//des champs électriques quantifiés.
        if (datum == 3) val = 3.3359;
    else if (datum == 1) val = 1;

//2 valeurs négatives possibles des
//des champs électriques quantifiés.
    else if (datum == 2) val = -1;
    else val = -3.3359;
//Affectation des champs électriques quantifiés (-3.3359, -1, 1, 3.3359)
//au vecteur des données converties processed_data.
    processed_data[(memory_dataframe_offset+df_iter)*ntimesteps_per_dataframe + i]
= val;
}
}
}
#ifdef _OPENMP
}
#endif

//Détruire le tampon ayant servi à l'allocation dynamique de la mémoire.
delete[] buffer;
}

```

## D.2 Fichier mark5b\_autocorr.cpp pour les autocorrélations

```

//=====fichier mark5b_autocorr.cpp=====
//Appel de huit bibliothèques/fichiers d'en-tête (classes).
#include "mpi.h"//Classe développée par des experts de Calcul Québec pour les
//manipulations d'application MPI
#include "xcorr.h"//Classe de transformées de Fourier directes et inverses de
//C++ permettant de faire l'autocorrélation
#include <iostream>//Pour le traitement du flux du fichier de sortie

#include "Mark5BDataProvider.h"//Classe Mark5BDataProvider.h développée par des
//experts de Calcul Québec.
#include "HDF5DataStore.h"//Classe HDF5DataStore.h développée par des
//experts de Calcul Québec.
#include <string>//Pour le traitement de chaînes de caractères
#include <unistd.h>//Pour utiliser l'interface de programmation du système
//d'exploitation Linux de Colosse POSIX/SUS.
#include "boost/program_options.hpp"//Pour gérer les options de commandes C++
//en ligne

```

```

//Se placer dans l'espace de noms std pour utiliser les symboles déclarés dans
//le fichier iostream.
using namespace std;

//=====PROGRAMME PRINCIPAL=====
//ENTRÉE : char ** argv. Tableau (2D) de chaînes de caractères contenant les
//          les paramètres passés au programme. argv[1] contient le nom du
//          fichier d'entrée. (argv[0] contient le nom de l'exécutable).
//          int argc. Contient le nombre de paramètres passés au programme.
//SORTIE : char ** argv. Tableau (2D) de chaînes de caractères contenant les
//          paramètres passés au programme. argv[2] contient le nom du fichier
//          de sortie qui contiendra les données converties (-3.3359, -1, 1,
//          3.3359).
//          (argv[0] contient le nom de l'exécutable).
//          Retourne 1 (vrai).
//=====

int main(int argc, char ** argv) {

//Appel de l'application MPI en lui passant le nombre de paramètres et
//le nom des fichiers traités (fichier d'entrée et fichier de sortie).
    MPI_Init(&argc, &argv);

//Déclaration de variables booléennes pour vérifier si
//les opérations suivantes sont exécutées ou non.
bool output_autocorrelation;//Autocorrélation
bool output_spectrum;//Spectre (nombres réels et complexes)
bool output_power_spectrum;//Spectre de puissance
bool output_initial_data;//Données initiales: champs électriques
bool compute_intensity;//Intensité
bool add_intensity_noise;//Intensité du bruit artificiel ajouté

int channel;//Numéro du canal à traiter
int num_df_background_avg;//Nombre de data frames pour enlever
//la moyenne
int smoothing_len;//Nombre de points pour le lissage
int num_pts_shift_in_intensity;//Nombre de points pour imposer un délai
long num_df;//Nombre de data frames traités.
double weight_low_intensity_noise;//Proportion de l'intensité du bruit
//ajouté (proportion de valeurs 1).

//Déclaration d'une variable de type chaînes de caractères contenant
//le nom du fichier d'entrée du NRAO. Ex.: tb021_sc_0002.m5 (150 Go).
string input_filename;

//Déclaration d'une variable de type chaînes de caractères contenant
//le nom du fichier de sortie (résultats) de format hdf5.
//Ex.: xc_tb021_sc_0002_m5.h5 (600 Go).
string output_filename;

//Se placer dans l'espace des noms pour manipuler les applications MPI.
namespace po = boost::program_options;

//Liste des options possibles.
po::options_description desc("Options");
desc.add_options()
//Valider la présence du fichier d'entrée.
("input",po::value<string>(&input_filename)->required(),"Input file name")

```

```

//Valider la présence du fichier de sortie.
("output",po::value<string>(&output_filename)->required(),"Output file name")
//Numéro du canal traité (-1 pour la somme de tous les canaux).
("channel",po::value<int>(&channel)->default_value(-1),"Channel to use. -1
means the sum of all channels.")
//Nombre de data frames traités (-1 pour tous les data frames,
//soit tout le fichier).
("numdf",po::value<long>(&num_df)->default_value(-1),"Number of dataframes. -1
means the whole file.")
//Nombre de points pour le lissage (filtre RIF à moyenne mobile).
//Par défaut, pas de lissage (1 point).
("smoothing_len",po::value<int>(&smoothing_len)->default_value(1),"Number of
points to use for the running average length.")
//Nombre de data frames pour calculer la moyenne (le continu).
//-1: on enlève la moyenne sur tous les data frames (le fichier).
//0: on ne soustrait pas la moyenne.
("numdf_background_avg",po::value<int>(&num_df_background_avg)-
>default_value(1),"Number of dataframe to use to average background. Value of 0
means do not remove background. Value of -1 means remove the background
globally")
//Activer ou désactiver l'option pour calculer les intensités.
("compute_intensity",po::value<bool>(&compute_intensity)-
>default_value(true),"Enables (if true/yes/1/on or empty) or disable (if
off/no/0/false) computation of intensity.")

//Partie seulement utilisée pour les SIMULATIONS seulement=====
//Début
//Nombre de points de décalage pour SIMULER un délai.
//I(t) = E(t)+ E(t + délai artificiel imposé).
("num_pts_shift_in_intensity",po::value<int>(&num_pts_shift_in_intensity)-
>default_value(0),"Enables a shift in the intensity computation. If intensity
computation is enabled, will compute I(t) = E(t) +
E(t+num_pts_shift_in_intensity). This number can be positive or negative. ")
//Fin

//Activer ou désactiver l'option pour calculer l'autocorrélation.
("output_autocorrelation",po::value<bool>(&output_autocorrelation)-
>default_value(true),"Enables or disables output of autocorrelation.")
//Activer ou désactiver l'option pour le spectre
//(TF avec des nombres réels et complexes).
("output_spectrum",po::value<bool>(&output_spectrum)-
>default_value(false),"Enables or disables output of amplitude spectrum.")
//Activer ou désactiver l'option pour le spectre de puissance.
("output_power_spectrum",po::value<bool>(&output_power_spectrum)-
>default_value(true),"Enables or disables output intensity spectrum.")
//Activer ou désactiver l'option pour obtenir les données
//initiales soit les champs électriques.
("output_initial_data",po::value<bool>(&output_initial_data)-
>default_value(false),"Enables (if true/yes/1/on or empty) or disable (if
off/no/0/false) outputing of initial data.")
("verbosity_level",po::value<int>(&VERBOSITY_LEVEL)->default_value(0),"Enables
more verbose output")

```

```

//Partie seulement utilisée pour les SIMULATIONS seulement=====
//Début
//Activer ou désactiver l'option pour ajouter du bruit artificiel
//lors des simulations.
("add_intensity_noise",po::value<bool>(&add_intensity_noise)-
>default_value(false),"Enables or disables adding intensity noise to the
signal.")
//Pondérer l'ajout des deux valeurs possibles de l'intensité du bruit
//simulé (1 ou 11.1282). L'utilisateur entre la proportion (entre 0 et 1) de
//du bruit de plus faible intensité (valeur = 1).
("weight_low_intensity_noise",po::value<double>(&weight_low_intensity_noise)-
>default_value(1),"Sets the weight (between 0 and 1) of the noise for the low
value of intensity.")
;
//Fin

//CETTE PARTIE DU CODE QUI PORTE SUR LA MANIPULATION DES APPLICATIONS MPI=====
//Cette partie du code a principalement été implantée par des experts en CHP
//de Calcul Québec.
//Début
    int comm_size, comm_rank;
    MPI_Comm_size(MPI_COMM_WORLD, &comm_size);
    MPI_Comm_rank(MPI_COMM_WORLD, &comm_rank);
//Fin=====

//Définir l'objet template T.
typedef double T;

//Définir l'objet l'objet dp de la classe fournissant les données.
Mark5BDataProvider<T> dp(input_filename,channel,num_df,channels_to_sum);

//Faire le lissage par filtre moyen mobile sur le nombre de points passé
//en paramètre.
dp.setRunningAverageLength(smoothing_len);

    if (compute_intensity) {
//Calculer l'intensité en activant la méthode pour le calcul des
//intensités.
dp.enableIntensityComputation();

//Partie seulement utilisée pour les SIMULATIONS seulement=====
//Début
//Calculer le décalage en points en activant la méthode pour le
//le calcul des décalages.
        dp.setShiftForIntensityComputation(num_pts_shift_in_intensity);
//Fin
    }
    else {

//Désactiver la méthode du calcul des intensités.
dp.disableIntensityComputation();
    }
    if (num_df_background_avg >= 1) {

//Évaluer la moyenne sur le nombre de data frames passés
//en paramètres (-1 pour tous les data frames, tout le fichier).
dp.setBackgroundPeriodicity(num_df_background_avg);

```

```

//Enlever la moyenne avant d'effectuer l'autocorrélation.
    dp.enableBackgroundRemoval();
}
else if (num_df_background_avg == -1) {

//Enlever la moyenne sur tous les data frames.
    dp.enableGlobalBackgroundRemoval();
}
else {
//Désactiver la méthode du calcul pour enlever la moyenne.
    dp.disableBackgroundRemoval();
}
//Partie seulement utilisée pour les SIMULATIONS seulement=====
//Début
    if (add_intensity_noise) {

//Ajouter du bruit artificiel en activant la méthode pour le calcul de
//l'intensités du bruit (deux valeurs possibles : 1 ou 11.1282)
        dp.enableAdditionOfIntensityNoise();
        dp.setWeightForLowIntensity(weight_low_intensity_noise);
    }
//Fin

//PARTIE DU CODE QUI PORTE SUR LA GESTION DE LA MÉMOIRE=====
//Partage de la tâche sur plusieurs nœuds.
//Cette partie du code a principalement été implantée par des experts en CHP
//de Calcul Québec.
//Début
    size_t mem_per_rank=size_t(2*1024)*size_t(1024*1024);
    if (comm_rank == 0) std::cout << "dp.size():" << dp.size() << "
numDataFrames():" << dp.getNumDataFrames() << std::endl;
    if (dp.size()*sizeof(std::complex<T>) > 1*comm_size*mem_per_rank)
    {
        if (comm_rank == 0) {
            std::cerr << "Error, not enough memory to make the computation." <<
std::endl;
            std::cerr << "Need at least " << dp.size()*sizeof(std::complex<T>) /
(1*8*mem_per_rank)+1 << " nodes to make the computation." << std::endl;
            std::cout << "Error, not enough memory to make the computation." <<
std::endl;
            std::cout << "Need at least " << dp.size()*sizeof(std::complex<T>) /
(1*8*mem_per_rank)+1 << " nodes to make the computation." << std::endl;
        }
        MPI_Finalize();
        exit(1);
    }
    for (const auto& it : vm) {
        auto & value = it.second.value();
        if (auto v = boost::any_cast<int>(&value))
            store.setProperty(it.first.c_str(),std::to_string(*v));
        else if (auto v = boost::any_cast<bool>(&value))
            store.setProperty(it.first.c_str(),std::to_string(*v));
        else if (auto v = boost::any_cast<string>(&value))
            store.setProperty(it.first.c_str(),*v);
        else if (auto v = boost::any_cast<ptrdiff_t>(&value))
            store.setProperty(it.first.c_str(),std::to_string(*v));
    }
//Fin=====

```

```

while (dp.size() % (comm_size-1) != 0) {
    dp.set_size(dp.size()-2);
}
//Déclaration de la variable status de type entier représentant une
//condition pour faire l'autocorrélation (1 = oui, 0 = non).
//Affecter le résultat (1 ou 0) de la fonction autocorr_c2c_inplace_mpi
//à status.
int status =
autocorr_c2c_inplace_mpi(&dp,store,output_autocorrelation,output_spectrum,output
t_power_spectrum,output_initial_data);

//Tant que la variable status n'est pas nulle (est vrai, vaut 1) faire
//l'autocorrélation des intensités.
while (status != 0) {
    if (comm_rank == 0) std::cerr << "Autocorrelation with " << dp.size() << "
data points could not be performed because it is prime. Performing
autocorrelation with " << dp.size()-2 << " data points." << std::endl;

    //Enlever 2 à la taille de l'objet dp (pour les nombres premiers).
    dp.set_size(dp.size()-2);

    if (comm_rank == 0) std::cout << "dp.size()=" << dp.size() << "," << comm_size
<< endl;

    //Affecter le résultat (1 ou 0) de la fonction autocorr_c2c_inplace_mpi (...)
    //à statut.
    status =
autocorr_c2c_inplace_mpi(&dp,store,output_autocorrelation,output_spectrum,output
t_power_spectrum,output_initial_data);
}
//Partie seulement utilisée pour les SIMULATIONS seulement=====
//Début
//Calculer l'intensité du signal avant l'ajout du bruit simulé.
store.setProperty("average_signal_before_noise",std::to_string(real(dp.getSigna
lAverageBeforeNoise())));

//Calculer l'intensité du signal après l'ajout du bruit simulé.
store.setProperty("average_signal_after_noise",std::to_string(real(dp.getSigna
lAverageAfterNoise())));
//Fin

//Destruction de l'objet store.
store.close();
//Fermer l'application MPI.
MPI_Finalize();

```

## E Code du fichier de soumission (simulations)

```
[etrottier@colosse2 mark5b_correlator]$ vim submit_two_signal.sh
```

```
1 #! /bin/bash
2 #PBS -S /bin/bash
3 #PBS -A dwx-253-aa
4 #PBS -l walltime=00:30:00
5 #PBS -m bea
6 #PBS -M eric.trottier.1@ulaval.ca
7 #PBS -l qos=SPJ1024
8 #PBS -l gattr=ckpt
9 ## 128 nœuds = 1024 cœurs
10 #PBS -l nodes=128:ppn=8
11
12 cd ${PBS_O_WORKDIR}
13 ## librairie fftw nécessaire à l'autocorrélation (TF suivie d'une TFI)
14 module load libs/fftw libs/phdf5 libs/boost apps/gnu-parallel
15
16 mpiexec \
17     ./mark5b_two_signals.out \
18     --input1 ~/scratch/fichiers_entree/ tb022_hn_b.m5 \    ## faisceau 1 de B1030+074
19     --input2 ~/scratch/fichiers_entree/tb022_hn_b.m5 \    ## faisceau 2 de B1030+074
20 ##   fichier de sortie contenant la simulation désirée (ex., figure 6.2)
21     --output ~/scratch/fichiers_sortie/simul_sh_17_92G_280sec_sl_32811_B1030_074.h5 \
22     --channel 0 \    ## canal 1
23     --numdf 7680000 \    ## correspond à 300 secondes d'observation (19.2 milliards de données)
24     --smoothing_len 32811\    ## lissage correspondant à la  $D_{MIC}$  de B1030+074 (voir tableau 6.2)
```

```
25 --numdf_background_avg -1 \ ## enlever la moyenne sur les 300 secondes
26 --output_initial_data_of_sum 1 \ ## activer la somme des deux intensités (champs)
27 --output_autocorrelation 1 \ ## activer l'autocorrélation d'intensité
28 --shift1 1792000000 \ ## délai (avant) imposé de 280 s (17.92 milliards points) au faisceau 1
29 --shift2 0 \
30 --intensity_of_signal1 0.667 \ ## fraction f de l'intensité I (i.e. E) du faisceau 1
31 --intensity_of_signal2 0.333 \ ## fraction 1-f de l'intensité I (i.e. E) du faisceau 2
32 --compute_intensity_of_sum 1 \ ## additionner les intensités des deux faisceaux
33 --add_intensity_noise 1 \ ## activer l'ajout de bruit local simulé
34 --weight_low_intensity_noise 0.05 ## p = 0.05 de bruit I = 1, (1-p) = 0.95 de bruit I = 11.1282
```

Le fichier de sortie (.h5) contient entre autres les deux attributs nommés, `average_signal_before_noise` et `average_signal_after_noise`, qui contiennent respectivement le signal moyen d'intensité avant et après l'ajout de bruit. On obtient évidemment le bruit moyen en faisant la soustraction. Par exemple, pour l'objet B1030+074, on obtient comme valeurs des deux attributs ( $f = 0.667$ ,  $p = 0.05$ ):

`average_signal_before_noise = 8.314869,`

`average_signal_after_noise = 19.422408.`

## F Liste des équations référencées dans cette thèse

Cette annexe permet au lecteur de retrouver (rapidement) la liste des cinq équations, tirées de publications externes, dont on a fait référence dans le corps du manuscrit, sans toutefois les écrire.

1. L'équation (5.1) de Press & Gunn (1973), qui donne la séparation angulaire la plus probable, s'écrit comme suit,

$$\theta_{\text{sep}} ; 2.6 \times 10^{-6} \left( \frac{M}{M_e} \right)^{1/2} \text{ arcsec}, \quad (\text{F.1})$$

avec  $\theta_{\text{sep}}$  représentant la séparation angulaire la plus probable entre deux images.

2. L'équation (5.2) de Press & Gunn (1973), qui donne le délai temporel entre deux images d'intensité comparable, s'écrit comme suit,

$$\Delta t ; \frac{1}{2} \theta_{\text{sep}}^2 L \frac{(L - \lambda)}{\lambda}; \theta^2, \quad (\text{F.2})$$

où  $\Delta t$  est le délai temporel (entre deux images), tandis que  $L$  et  $\lambda$  sont les paramètres affines de distance (en Mpc):  $L$  séparant l'observateur de la source et  $\lambda$  séparant l'observateur de la lentille. Notez la correction d'une erreur typographique dans Press & Gunn (1973): compte tenu que les unités de  $L$  et  $\lambda$  sont des unités de longueur, il faudrait ajouter le facteur  $1/c$  dans l'équation (F.2) pour obtenir des unités de temps.

3. L'équation (5) de Borra (1997), qui donne le spectre en fréquence, s'écrit comme suit,

$$I(\omega) = K \cos(\omega\tau/2) \exp(-1/2\omega^2/\sigma^2), \quad (\text{F.3})$$

où  $K$  est un regroupement de constantes d'intégration,  $\omega$  est la fréquence de d'observation,  $\tau$  est le délai temporel entre deux impulsions,  $\sigma$  est la dispersion gaussienne et  $I(\omega)$  est l'intensité (spectre de puissance) dans le domaine de Fourier.

4. L'équation (13) de Borra (2008), s'écrit comme suit,

$$\gamma(a, \omega') = \text{Re} \left[ \int_{-\infty}^{\infty} i(\theta) e^{-i\omega' a\theta/c} d\theta \right], \quad (\text{F.4})$$

avec  $i(\theta)$  représentant la distribution d'intensité angulaire normalisée. La fonction  $\gamma(a, \omega')$  représente la transformée de Fourier de  $i(\theta)$ , et peut être aisément identifiée comme étant la fonction de cohérence spatiale  $\gamma(0)$  évaluée à la fréquence (faible) de battements  $\omega'$ . Enfin, le paramètre  $a$  est similaire à la séparation entre les fentes d'un interféromètre de Young.

5. L'équation (14) de Borra (2008), qui quantifie le contraste spectral  $D(\omega')$  dans le spectre de battements, s'écrit comme suit,

$$D(\omega') = S_0 \psi(\omega') \left[ 1 + \gamma(a, \omega') \cos(\omega' \tau) \right], \quad (\text{F.5})$$

où  $S_0$  est l'intensité totale de la source, tandis que  $\psi(\omega')$  représente la densité spectrale du spectre de battements non modulé. (Voir éq. (F.3) et éq. (F.4) pour la signification  $\tau$ ,  $\omega'$  et  $\gamma(a, \omega')$ .)