



Du HI radio à la mesure de la croissance des bassins gravitationnels

Alexandra Dupuy

► To cite this version:

Alexandra Dupuy. Du HI radio à la mesure de la croissance des bassins gravitationnels. Astrophysique [astro-ph]. Université de Lyon, 2018. Français. NNT : 2018LYSE1201 . tel-02108079v2

HAL Id: tel-02108079

<https://theses.hal.science/tel-02108079v2>

Submitted on 25 Apr 2019

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

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



N° d'ordre NNT : 2018LYSE1201

THÈSE DE DOCTORAT DE L'UNIVERSITÉ DE LYON

opérée au sein de
l'Université Claude Bernard Lyon 1

École Doctorale ED52
École Doctorale de Physique et Astrophysique

Spécialité de doctorat : Physique
Discipline : Astrophysique

Soutenue publiquement le 23/10/2018, par :
Alexandra Dupuy

Du HI radio à la mesure de la croissance des bassins gravitationnels

Devant le jury composé de :

Mme KRAAN-KORTEWEG Renée
M. ANSARI Réza
M. DUPONT Florent
M. POMARÈDE Daniel
Mme COURTOIS Hélène

Rapporteure
Rapporteur
Examineur
Examineur
Directrice de thèse

Résumé

Les mouvements des galaxies dans l'Univers sont causés par le tir à la corde qui a lieu entre l'expansion de l'univers et la gravitation. Cette rivalité a un impact sur la formation et la croissance des grandes structures de l'univers. Par conséquent, l'univers est plus ou moins compact selon le gagnant de ce duel. Dans ce contexte, cette thèse est divisée en trois parties, allant des observations à 21 cm à la mesure du taux de croissance actuel des grandes structures de l'univers et de la compacité de l'univers local.

La collaboration Cosmicflows prépare des catalogues de distances de galaxies. Jusqu'à présent, trois catalogues ont été publiés, le dernier étant Cosmicflows-3. Cette thèse présente et analyse les données obtenues à partir d'observations à 21 cm. Ces nouvelles données seront utilisées pour former le prochain catalogue de distances Cosmicflows-4 à l'aide de la relation de Tully-Fisher, afin de corriger le manque de données dans l'hémisphère céleste Nord dans le catalogue actuel.

À partir de la distance d'une galaxie, il est possible d'en déduire la partie radiale de sa vitesse particulière, correspondant à la composante de sa vitesse totale causée par la gravitation. Les vitesses particulières des galaxies sont des sondes non-baisées de la matière et permettent d'extraire des informations sur les grandes structures de l'univers. Deux méthodologies utilisant les vitesses particulières ont été développées durant cette thèse pour caractériser les grandes structures de l'univers local. D'une part, les catalogues de vitesses particulières peuvent être utilisés pour reconstruire des champs de vitesse tridimensionnels. De tels champs de vitesse permettent de cartographier la structure de l'univers local et sont exploités dans cette thèse, à l'aide des lignes de flux, pour identifier des bassins et vallées gravitationnels dans l'univers local. D'autre part, une méthode basée sur l'analyse des corrélations à deux points des vitesses particulières des galaxies a été développée afin d'exhiber le taux de croissance actuel des grandes structures de l'univers local à partir de données observationnelles. Cette méthodologie est appliquée au catalogue observationnel de vitesses particulières Cosmicflows-3.

Abstract

Motions of galaxies in the universe are due to the rivalry between the expansion of the universe and gravitation. This tug-of-war impacts the formation and the growth of large scale structures of the universe. Thus, depending on the identity of the winner of this duel, the universe is more or less compact. Within this context, this PhD thesis is divided into three parts, spanning from HI observations to the estimate of the growth of rate of large scale structures of the universe and the compactness of the local universe.

The Cosmicflows collaboration assembles catalogues of galaxy distances. Up to now, three catalogues have been published, the last one being Cosmicflows-3. This thesis presents and analyses observational data obtained from HI observations. These new data will be used to construct the new compilation of distances Cosmicflows-4 by the use of the Tully-Fisher relation, in order to correct the lack of data in the Northern celestial hemisphere in the current catalog.

From the distance of a galaxy, one can derive the radial part of its peculiar velocity corresponding to the component of its total velocity caused by gravitation. Peculiar velocities allow to probe the matter content of the universe and to extract information on large scale structures of the universe. Two methodologies using peculiar velocities have been developed during this thesis to characterize large scale structures of the local universe. On the one hand, peculiar velocity catalogues can be used to reconstruct tri-dimensional velocity fields. These velocity field allow one to map the structure of the local universe and are used in this thesis to identify gravitational basins and valleys within the local universe by computing streamlines. On the other hand, a method based on the analysis of two-point galaxy peculiar velocity correlations has been developed in order to constrain the growth rate of large scale structures of the local universe from observational data. This method is applied to the Cosmicflows-3 catalogue of observed peculiar velocities.

Remerciements

Par ces quelques mots je souhaite remercier toutes les personnes qui m'ont aidée durant ces trois années de thèse de doctorat.

Je commence bien sûr par remercier ma directrice de thèse, Hélène Courtois, sans qui ce travail n'aurait pas été possible. Grâce à son aide et son soutien tout au long de ces années difficiles, j'ai pu mener jusqu'au bout mes travaux de recherche. Malgré ses nombreuses occupations et son emploi du temps débordant, elle a toujours été présente. Son exigence à mon égard a contribué certainement à ma réussite. Il est vrai que durant ces trois années, les semaines de travail se poursuivaient souvent le week-end, et tard le soir, y compris la veille de la soutenance de thèse à 23h... Cette rigueur dans le travail m'a cependant permis d'évoluer. Hélène, je te suis reconnaissante de m'avoir fait confiance il y a trois ans. Je te remercie aussi d'avoir rendu possibles tous les déplacements dans le monde entier auxquels j'ai pu participer.

Je tiens aussi à remercier toute l'équipe de cosmologie observationnelle de l'Institut de Physique Nucléaire de Lyon : Yannick Copin, Gérard Smadja, Bogna Kubik, Mickael Rigault, Clément Buton, et Daniel Guinet, pour leurs nombreux conseils et plus particulièrement pour avoir participé chacun à la relecture de mon manuscrit de thèse et de mes articles ; ainsi que les doctorants, avec qui j'ai partagé un bureau, des discussions et des moments agréables durant toutes ces années : Romain Graziani, Mehdi Outini et Martin Briday. Je souhaite également remercier les stagiaires qui ont rejoint l'équipe ces trois dernières années : Savitri Gallego, Benjamin Vigneron, et Jean-Baptiste Boutillier, qui m'ont également aidée à avancer tout au long de ma thèse. Un énorme merci encore à Daniel Guinet, pour m'avoir accompagnée à Washington et à Greenbank, et par la suite d'avoir participé avec Hélène et moi-même aux observations au GBT, malgré les difficultés. Daniel, ta collaboration a été précieuse.

Un très grand merci aussi aux collaborations Cosmicflows et CLUES, et plus particulièrement à Brent Tully, Daniel Pomarède, Yehuda Hoffman, Stefan Gottlöber, Noam Libeskind, Edoardo Carlesi, Jenny Sorce, Ehsan Kourkchi, pour leurs nombreux conseils et discussions qui m'ont permis d'avancer dans mes travaux de recherche. Brent, merci de m'avoir accueillie à Hawaï durant un mois. Je tiens aussi à remercier Florent Dupont et Florence Denis du LIRIS, sans qui le chapitre 4 de cette thèse n'aurait pas pu se réaliser.

Je remercie particulièrement les membres de mon jury de thèse : les rapporteurs Renée Kraan-Korteweg et Réza Ansari, et les examinateurs Florent Dupont et Daniel Pomarède. Merci également à toutes les personnes qui ont assisté à ma soutenance.

Et comment ne pas oublier dans ces remerciements mes parents, pour m'avoir soutenue

et supportée (dans les deux sens du terme) durant ces trois années de thèse. Vous avez toujours cru en moi, malgré mes moments de doute. Merci ! Et merci aussi à ma tante Martine, mon oncle Ghislain et ma cousine Lisa, pour m'avoir aidée à évacuer mon stress par leur présence, pendant les quelques jours passés ensemble chez les indiens cet été, la veille et le jour de ma soutenance. Merci à vous !

Merci encore à vous tous, et à ceux que j'ai peut-être oublié.

Table des matières

1	Cosmologie et grandes structures de l'univers	17
1.1	Modèle standard cosmologique Λ CDM	18
1.1.1	Cinématique de l'univers	18
1.1.2	Contenu de l'univers	22
1.2	Grandes structures de l'univers	24
1.2.1	Régime linéaire et taux de croissance des grandes structures	27
1.2.2	Croissance non-linéaire : à l'échelle des galaxies	31
2	Données observationnelles Cosmicflows	35
2.1	Distances et vitesses particulières observées	36
2.1.1	Distances	36
2.1.2	Vitesses particulières	38
2.2	Catalogues de distances de galaxies Cosmicflows	39
2.2.1	Les débuts : Cosmicflows-1 et Cosmicflows-2	39
2.2.2	<i>Longius</i> — Cosmicflows-3 : encore plus loin ... vers le Sud	40
2.2.3	<i>Altius</i> — Cosmicflows-4 : encore plus haut ... vers le Nord	47
3	Vers Cosmicflows-4 : observations HI	51
3.1	La relation de Tully-Fisher	52
3.1.1	Radioastronomie : observations à 21 cm	54
3.2	Données provenant du radiotélescope Nançay	61
3.2.1	Kinematic in the Local UNiverse	62

3.2.2	Nançay Interstellar Baryons Legacy Extragalactic Survey	67
3.3	Données provenant du radiotélescope Arecibo	72
3.3.1	Arecibo Legacy Fast ALFA : description	72
3.3.2	Intégration des données ALFALFA à Cosmicflows-4	73
3.4	Projet GBT18A-021 : Cosmic flows in the Local Universe	80
3.4.1	Sélection des sources HI	80
3.4.2	Observations à 21 cm au GBT	85
3.4.3	Observations au GBT : réduction des données	90
3.4.4	Premiers résultats	94
3.5	Conclusion : données HI pour Cosmicflows-4	97
4	Identification, caractérisation et visualisation des grandes structures	101
4.1	Champs de vitesse particulière tri-dimensionnels	102
4.1.1	Technique du filtre de Wiener et des réalisations contraintes	102
4.1.2	Reconstruction par échantillonnage de Gibbs	105
4.2	Exploitation des champs de vitesse particulière	106
4.2.1	Lignes de flux — <i>streamlines</i>	107
4.2.2	Segmentation des bassins gravitationnels	113
4.2.3	Vallées gravitationnelles	117
4.2.4	Implémentation en C++	118
4.3	Simulation	120
4.4	Application aux données <i>Cosmicflows</i>	129
4.4.1	Champs de surdensité et de vitesse <i>Cosmicflows</i>	129
4.4.2	Cosmicflows-2 : notre superamas Laniakea	132
4.4.3	Cosmicflows-3 : résultats préliminaires	136
5	Taux de croissance local	141
5.1	Bibliographie	142

5.1.1	Distorsions dans l'espace des redshifts	142
5.1.2	Oscillations acoustiques des baryons	145
5.1.3	Vitesses particulières	147
5.2	Statistiques à deux points des vitesses particulières	148
5.2.1	Vitesse relative de paires	149
5.2.2	Corrélations à deux points des vitesses particulières	151
5.2.3	Estimation des erreurs sur les statistiques à deux points	154
5.3	Taux de croissance sur les données Cosmicflows-3	155
5.3.1	Validation sur des catalogues fictifs	156
5.3.2	Variance cosmique	159
5.3.3	Constante de Hubble	161
5.3.4	Taux de croissance local avec les données Cosmicflows-3	163
6	Conclusion et perspectives	167
	Appendices	173
A	Observations au GBT : paramètres HI des galaxies adéquates pour Cosmicflows-4	175
B	Observations au GBT : spectres HI adéquats pour Cosmicflows-4	179
C	Observations au GBT : spectres HI des sources confuses ou non détectées	191

Introduction

L'Univers fascine et a toujours fasciné. L'être humain a constamment cherché à en percer les mystères. Cela a donné naissance à la cosmologie : l'étude de l'origine de l'univers et de son évolution. Depuis des siècles, les scientifiques cherchent à le comprendre et à l'expliquer par diverses théories, parfois farfelues, satisfaisant les observations. Petit à petit, avec les progrès de la science et de la technologie, la qualité des observations s'est améliorée et a permis d'affiner les modèles cosmologiques. Certains furent abandonnés, au profit d'autres théories, plus concordantes. Ainsi, le modèle standard de la cosmologie moderne utilisé de nos jours s'est mis en place : le modèle Λ CDM ; reposant sur la Relativité Générale d'Albert Einstein.

Mais que sait-on réellement sur l'Univers aujourd'hui ? Les observations jouent un rôle crucial dans l'étude de l'univers. Les études de décalages vers le rouge des galaxies montrent que non seulement l'univers est en expansion (Hubble, 1929), mais que cette expansion est accélérée Riess et al. (1998); Perlmutter et al. (1999). Cela implique l'existence d'une forme inconnue de matière nommée *énergie noire*, qui a une pression négative et donc un effet répulsif, contrairement à la matière ordinaire. Par ailleurs, l'analyse des courbes de rotation des galaxies spirales (Rubin and Ford, 1970; Rubin et al., 1980; Roberts and Rots, 1973) et l'étude des amas de galaxies Einasto et al. (1974); Hubble and Humason (1931); Zwicky (1933); Brownstein and Moffat (2007) ont mis en évidence la présence dans l'univers d'une autre forme de matière : la *matière noire*, qui n'agit que gravitationnellement et dont la nature est encore inconnue à ce jour. L'univers est composé de seulement 5% de matière ordinaire, nommée la matière baryonique. Les 95% restants sont la matière noire et l'énergie noire. Par conséquent, la majorité de la composition de l'univers est inconnue.

Les études du fond diffus cosmologique ont montré que l'univers était quasiment homogène et isotrope il y a 13 millions d'années, avec de très faibles fluctuations de densité. Ces fluctuations ont donné naissance aux structures à grande échelle que l'on observe aujourd'hui, à l'aide des catalogues de redshift : des galaxies, des amas et superamas de galaxies, et des vides cosmiques. Les mouvements des galaxies dans l'univers sont causés par le *tir à la corde* qui a lieu entre l'expansion de l'univers et la gravitation, et sont responsables de la formation des grandes structures. Ces mouvements permettent ainsi de cartographier la matière totale de l'univers. Les galaxies ont alors une vitesse qui peut être décomposée en deux parties : une vitesse due à l'expansion, et une vitesse due à la gravitation, faisant dévier la galaxie du mouvement d'expansion général. Cette vitesse, appelée vitesse particulière, peut être obtenue observationnellement à partir de la distance et de la vitesse radiale d'une galaxie, son redshift. Les vitesses particulières peuvent être considérées comme traceurs non biaisés de la matière dans l'univers, et donc comme une sonde permettant de tester le modèle standard de la cosmologie Λ CDM.

Dans cette thèse, les vitesses particulières des galaxies, tirées des catalogues observationnels *Cosmicflows*, permettent de caractériser les grandes structures de l'univers. D'une part, les grandes structures de l'univers seront identifiées en tant que bassins et vallées gravitationnelles. D'autre part, les vitesses particulières observées seront utilisées comme sondes cosmologiques afin de mesurer le taux de croissance des grandes structures de l'univers, quantité représentant la rivalité entre l'expansion et la gravitation.

Plan de la thèse

L'organisation de cette thèse est la suivante.

Le premier chapitre met en place le contexte dans lequel s'inscrit cette thèse : la cosmologie. Les notions de base concernant le modèle standard de la cosmologie, nécessaires à la lecture de ce manuscrit, seront décrites rapidement. Puis seront mises en avant les grandes structures de l'univers, objets-clé de cette thèse. Les théories linéaire et non-linéaire décrivant la formation de ces structures à deux échelles différentes seront détaillées. Le taux de croissance des grandes structures sera notamment introduit dans le cadre du régime linéaire.

Le deuxième chapitre décrit les données observationnelles qui sont utilisées tout au long de cette thèse. Les méthodes permettant de mesurer les distances de galaxies sont tout d'abord évoquées, puis les équations permettant d'estimer les vitesses particulières des galaxies à partir des distances observées sont données. Les catalogues observationnels *Cosmicflows* sont ensuite détaillés, notamment le catalogue *Cosmicflows-3*, sur lequel est basée cette thèse. La prochaine compilation de distance de galaxies, *Cosmicflows-4* est ensuite présentée.

Cette dernière fait l'objet du troisième chapitre, qui expose les mesures de largeurs de raie à 21 cm agglomérées durant cette thèse. Après avoir introduit la relation de Tully-Fisher, l'analyse de données des collaborations KLUN et NIBLES provenant du radiotélescope Nançay, puis celles de la collaboration ALFALFA provenant du télescope Arecibo, sont présentées. Ensuite, les observations qui ont été conduites au radiotélescope de Greenbank durant cette thèse sont détaillées, allant de la sélection des sources HI à observer aux mesures des largeurs de raies HI. Toutes ces mesures seront ainsi ajoutées à *Cosmicflows-3* pour former *Cosmicflows-4*. Un bilan sera dressé pour exhiber l'évolution entre ces deux catalogues.

Le quatrième chapitre présente une méthode permettant d'identifier et caractériser ces grandes structures. Les techniques permettant de reconstruire le champ de vitesse tridimensionnel à partir de catalogues de vitesses particulières observées sont tout d'abord décrites. Puis la méthode utilisée dans cette thèse, exploitant un champ de vitesse et basée sur les lignes de flux, est décrite et testée sur un champ de vitesse synthétique. Cette méthodologie permettant d'identifier des bassins et vallées gravitationnelles dans un champ de vitesse est ensuite appliquée à une simulation Λ CDM puis aux champs de vitesses reconstruits à partir des données *Cosmicflows*.

Le cinquième chapitre est consacré à la mesure du taux de croissance des grandes structures de l'univers local à partir des données *Cosmicflows-3*. Après avoir listé quelques méthodes permettant d'estimer le taux de croissance selon les sondes cosmologiques consi-

dérées, la méthode utilisée dans cette thèse permettant d'estimer le taux de croissance à partir des données observationnelles à l'aide de statistiques à deux points des vitesses particulières est décrite. Puis cette méthode est validée sur des catalogues fictifs. Les effets provoqués par la variance cosmique et le choix de la valeur de la constante de Hubble sont discutés. Enfin, le taux de croissance local sera mesuré sur les données du catalogue observationnel *Cosmicflows-3*, ce qui clôturera ce chapitre et le corps de cette thèse.

Pour terminer, le sixième chapitre résume le travail observationnel et d'analyse effectué durant cette thèse, et propose des ouvertures et perspectives qui pourraient être envisagées.

Chapitre 1

Cosmologie et grandes structures de l'univers

Les mouvements des galaxies dans l'univers sont causés par la compétition entre l'expansion de l'univers et la gravitation. Ce *tir à la corde* a un impact sur la formation, sur la croissance des grandes structures de l'univers. L'univers est ainsi localement plus ou moins compact, selon le *gagnant* de ce duel. Par ailleurs, les observations donnent le décalage spectral (redshift) des galaxies qui correspond à la somme des vitesses dues à l'expansion et à la gravitation. En utilisant la constante de Hubble dont la valeur sera discutée dans cette thèse, il est possible de déduire la partie gravitationnelle de la vitesse, notée vitesse particulière. Ces vitesses particulières des galaxies sont des traceurs non biaisés de la matière totale dans l'univers (matière noire et matière baryonique). Par conséquent, elles agissent comme une sonde cosmologique et permettent ainsi d'obtenir des estimations du taux de compacité actuel de notre univers local, ainsi que son architecture composée de bassins versants gravitationnels.

Ce chapitre présente tout d'abord le modèle standard de la cosmologie. Puis il met en avant les grandes structures de l'univers, objets clés de cette thèse, qui seront identifiées plus tard dans le chapitre 4 à l'aide des données de la collaboration *Cosmicflows*. Les bases ainsi posées seront utilisées pour définir le taux de croissance des grandes structures un paramètre symbolisant la compacité de l'univers, qui sera estimé avec les mêmes données observationnelles *Cosmicflows* dans le chapitre 5.

Sommaire

1.1	Modèle standard cosmologique ΛCDM	18
1.1.1	Cinématique de l'univers	18
1.1.2	Contenu de l'univers	22
1.2	Grandes structures de l'univers	24
1.2.1	Régime linéaire et taux de croissance des grandes structures	27
1.2.2	Croissance non-linéaire : à l'échelle des galaxies	31

1.1 Modèle standard cosmologique Λ CDM

Les études de l'univers observable montrent à la fois que nous ne sommes pas au centre de l'univers, et qu'il n'y a pas de centre. À grande échelle, l'univers paraît isotrope et homogène, c'est-à-dire qu'il est le même partout quelle que soit la direction dans laquelle on regarde. Cependant ceci n'est pas vrai à toutes les échelles. D'une part, si on regarde vers l'amas de Virgo situé à 10 Mpc de nous, on ne voit pas d'autres amas situés à la même distance dans d'autres directions : cela signifie qu'à l'échelle des amas de galaxies, l'univers n'est ni homogène ni isotrope. Mais à l'échelle de la carte du fond diffus (CMB) de Planck, l'univers paraît homogène. Il existe donc une échelle d'inhomogénéité située entre ces deux échelles et correspondant à l'échelle à partir de laquelle l'univers paraît homogène et isotrope.

Les observations montrent aussi que l'univers est en expansion accélérée, et que des structures se sont formées. L'univers actuel comporte un grand nombre de structures telles que des étoiles et systèmes solaires, des galaxies, des amas de galaxies, superamas de galaxies. Les observations montrent aussi que l'univers est géométriquement plat, c'est-à-dire qu'il n'a pas de courbure, un terme introduit par la relativité générale d'Einstein.

Un modèle cosmologique décrit la dynamique d'un tel univers. Il peut être basé sur la relativité générale, ou bien sur une autre théorie de gravitation modifiée. Le modèle décrit aussi le contenu de cet univers : de quels types de matière, fluide, etc. notre univers est-il composé, à quelles proportions et quelles sont leurs propriétés ?

Le modèle cosmologique décrivant actuellement de la manière la plus satisfaisante l'univers observable selon les observations astronomiques est le modèle cosmologique standard *Λ Cold Dark Matter* (Λ CDM). Ce modèle décrit ainsi un univers né du BigBang, homogène et isotrope à grande échelle, et dans lequel des structures, telles que nous les connaissons aujourd'hui, se sont formées. De plus, l'univers est composé de matière baryonique, ainsi que de matière noire froide non relativiste, et d'énergie noire responsable de l'expansion de l'univers et représentée par une constante cosmologique Λ .

Toutes les informations données par la suite, sauf indiqué, sont extraites principalement de Peebles (1980); Peacock (1999).

1.1.1 Cinématique de l'univers

Métrique de Friedman-Lemaitre-Robertson-Walker

La théorie de la relativité générale a été développée par Einstein en 1905. Ce concept révolutionne la physique moderne. En effet, la gravitation était auparavant modélisée comme une force par Newton. La relativité générale d'Einstein décrit à présent la gravitation comme un effet de géométrie, une déformation de l'espace temps.

La notion d'espace-temps décrit une représentation mathématique de l'espace et du temps, qui sont deux notions inseparables en relativité générale. Cette représentation comporte quatre dimensions : le temps (1D, t) et l'espace (3D, x , y et z) qui dépendent tous deux du référentiel considéré.

Habituellement, dans l'espace euclidien, une distance ds entre deux points séparés par un vecteur $d\vec{x} = (dx, dy, dz) = (dx^1, dx^2, dx^3)$ est définie par $ds^2 = (dx^1)^2 + (dx^2)^2 + (dx^3)^2$. En introduisant le tenseur métrique $g_{\mu\nu}$ qui relie les coordonnées de deux points à leur distance :

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu, \quad (1.1)$$

on a $g_{\mu\nu} = \text{diag}(1, 1, 1)$. Ici seules les variables *espace* (x, y, z) sont considérées.

Cependant, en relativité générale, les variables *temps* et *espace* étant liées, il faut passer dans l'espace de Minkowski, qui est un espace pseudo-euclidien à quatre dimensions. Ainsi les points de l'espace euclidien deviennent des *événements* dans l'espace de Minkowski. La métrique devient $g_{\mu\nu} = \text{diag}(-c^2, 1, 1, 1)$ où c est la vitesse de la lumière. Le facteur c^2 permet d'avoir les quatre dimensions homogènes à une distance. Pour simplifier, on fixera par la suite $c = 1$. La distance entre deux événements de l'espace de Minkowski est donc $ds^2 = g_{\mu\nu} dx^\mu dx^\nu = -dt^2 + (dx^1)^2 + (dx^2)^2 + (dx^3)^2$.

Les métriques décrivant un univers homogène et isotrope tel que le décrit le modèle cosmologique standard sont celles de Friedman-Lemaître-Robertson-Walker (FLRW), définies par :

$$ds^2 = -dt^2 + a(t)^2 \left[\frac{dr^2}{1 - kr^2} + r^2(d\theta^2 + \sin^2 \theta d\phi^2) \right], \quad (1.2)$$

où (r, θ, ϕ) sont les coordonnées sphériques et sont indépendantes du temps. Le paramètre $a(t)$ est le facteur d'échelle et décrit l'expansion de l'univers au temps t . Il est normalisé à sa valeur actuelle (d'aujourd'hui) $a_0 = a(t = 0) = 1$. Le paramètre k est le paramètre de courbure. Son signe donne des détails sur la courbure de l'univers :

- si $k < 0$, l'univers est négativement courbé. La somme des angles d'un triangle est inférieure à 180° . On dit dans ce cas que l'univers est ouvert. La selle de cheval est une forme géométrique représentant un tel univers ;
- si $k = 0$, la courbure de l'univers est nulle et la somme des angles d'un triangle est égale à 180° . L'univers est géométriquement plat ;
- si $k > 0$, l'univers est positivement courbé. Dans ce cas, la somme des angles d'un triangle est supérieure à 180° , et l'univers est fermé, tel qu'une sphère.

Équation d'Einstein

En présence de masse ou d'énergie, l'espace-temps se courbe. Dans ce cas, le chemin le plus court entre deux points ds^2 n'est plus une droite mais une courbe appelée *géodésique*. C'est pour cette raison que la trajectoire des photons est "déviée" lorsqu'ils sont proches d'un objet de masse importante qui va ainsi *déformer* localement l'espace-temps (effet de lentille gravitationnelle).

L'équation d'Einstein est l'équation fondamentale de la relativité générale. Elle est donnée par :

$$\left[R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R \right] - \Lambda g_{\mu\nu} \equiv G_{\mu\nu} - \Lambda g_{\mu\nu} = 8\pi G T_{\mu\nu}, \quad (1.3)$$

où g est la constante gravitationnelle. Le tenseur $G_{\mu\nu}$ est le tenseur d'Einstein. Il est défini par le tenseur de Ricci $R_{\mu\nu}$ et le scalaire de Ricci R , qui dépendent seulement de la métrique $g_{\mu\nu}$ (et de ses dérivées). Le tenseur $G_{\mu\nu}$ représente l'effet de la gravitation qui courbe l'espace-temps. Le tenseur $T_{\mu\nu}$ est le tenseur énergie-impulsion. Celui-ci décrit le contenu de l'univers en matière et énergie. Enfin, le terme $\Lambda g_{\mu\nu}$ représente l'énergie du vide, où la constante Λ est la constante cosmologique. Ce terme $\Lambda g_{\mu\nu}$ peut appartenir aussi bien au terme de droite qu'au terme de gauche. Cela signifie que ce terme apporte soit une modification à la théorie de la gravitation (terme du côté de $G_{\mu\nu}$), soit une nouvelle source d'énergie contenue dans l'univers (terme du côté de $T_{\mu\nu}$).

Equations de Friedman

Dans le cadre d'une métrique de FLRW, il est possible de déduire les équations de Friedman. Celles-ci décrivent l'évolution du facteur d'échelle $a(t)$. Dans la métrique de FLRW, le tenseur de Ricci $R_{\mu\nu}$ est diagonal et a pour composantes non nulles :

$$R_{tt} = -3\frac{\ddot{a}}{a}; \quad (1.4)$$

$$R_{rr} = \frac{a^2}{1 - kr^2} \left[\frac{\ddot{a}}{a} + 2 \left(\frac{\dot{a}}{a} \right)^2 + 2 \frac{k}{a^2} \right]; \quad (1.5)$$

$$R_{\theta\theta} = r^2 \left[\frac{\ddot{a}}{a} + 2 \left(\frac{\dot{a}}{a} \right)^2 + 2 \frac{k}{a^2} \right]; \quad (1.6)$$

$$R_{\phi\phi} = r^2 \sin^2 \theta \left[\frac{\ddot{a}}{a} + 2 \left(\frac{\dot{a}}{a} \right)^2 + 2 \frac{k}{a^2} \right]. \quad (1.7)$$

Pour un fluide de densité d'énergie ρ et de pression P , le tenseur énergie-impulsion $T_{\mu\nu}$ est défini par :

$$T_{\mu\nu} = (\rho + P)U_\mu U_\nu + P g_{\mu\nu}, \quad (1.8)$$

où U_ν et U_μ sont les quadri-vitesses de ce fluide. Dans le cadre d'un univers homogène et isotrope rempli de ce fluide, le tenseur $T_{\mu\nu}$ est diagonal : $T_{\mu\nu} = \text{diag}(\rho, P, P, P)$.

Enfin, en utilisant les cinq équations précédentes, on peut écrire les deux équations de Friedman :

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3P) + \frac{\Lambda}{3}; \quad (1.9)$$

$$H(a)^2 = \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G\rho}{3} + \frac{\Lambda}{3} - \frac{k}{a^2}, \quad (1.10)$$

où $H(a)$ est le paramètre de Hubble et décrit l'évolution de l'expansion de l'univers. Les équations de Friedman décrivent l'évolution dynamique de l'univers.

Décalage vers le rouge et loi de Hubble

Considérons un photon de longueur d'onde λ émis à un temps t_e et observé à un temps t_o . Comme l'univers est en expansion, il s'est étendu entre t_e et t_o . La longueur d'onde du photon a elle aussi subi l'expansion : elle est proportionnelle au facteur d'échelle. On a donc :

$$\frac{\lambda_o}{\lambda_e} = \frac{a(t_o)}{a(t_e)} = 1 + z \quad (1.11)$$

où $a(t_e)$ est le facteur d'échelle au moment où le photon a été émis, et $a(t_o)$ sa valeur au moment où le photon a été observé. Le terme $\frac{a(t_o)}{a(t_e)}$ est supérieur à 1 car le facteur $a(t)$ croît avec le temps à cause de l'expansion de l'univers. Le paramètre z est donc positif, et la longueur d'onde observée λ_o est donc plus grande que la longueur d'onde émise λ_e : on observe un décalage vers le rouge, *redshift* en anglais. En effet l'expansion de l'univers *allonge* la longueur d'onde du photon entre t_e et t_o . On observe donc une longueur d'onde plus grande que prévu.

En posant $a(t_o) = a_0 = 1$, on peut relier le facteur d'échelle a au redshift z :

$$a = (1 + z)^{-1}. \quad (1.12)$$

En 1929, l'astronome Edwin Hubble remarque que la lumière provenant des galaxies est soumise à un décalage vers le rouge proportionnel à leur distance. Cela veut dire que les galaxies s'éloignent de la Terre à une vitesse de recession $v_{\text{recession}}$ proportionnelle à leur distance d . Il énonce alors la loi de Hubble :

$$v_{\text{recession}} = cz = H_0 d, \quad (1.13)$$

où le facteur $H_0 = H(a = a_0)$ est la constante de Hubble, i.e la valeur actuelle du paramètre de Hubble.

1.1.2 Contenu de l'univers

Équation d'état de l'univers

À partir des équations de Friedman, et en postulant une equation d'état $P = w\rho$, l'équation de conservation de l'énergie de l'univers est donnée par :

$$\frac{\dot{\rho}}{\rho} = -3(1+w)\frac{\dot{a}}{a}. \quad (1.14)$$

La valeur du paramètre w varie en fonction du type de fluide considéré. La solution de cette equation est :

$$\rho \propto a^{-3(1+w)}. \quad (1.15)$$

En effet, en supposant que le fluide d'équation d'état $\rho = wP$ est contenu dans une boîte de volume $V = a^3$ qui est soumise à l'expansion de l'univers, on peut déduire les propriétés de ce fluide selon la valeur de w considérée :

- si $w = 0$, alors $\rho \propto a^{-3} \propto V^{-1}$. Cela signifie que la densité d'énergie est inversement proportionnelle au volume. C'est-à-dire que si la quantité d'énergie est conservée, la densité décroît comme a^{-3} tandis que le volume croît comme a^3 . Ceci concerne la matière ordinaire, non-relativiste, aussi appelée matière baryonique ;
- si $w = \frac{1}{3}$, alors $\rho \propto a^{-4}$. Dans ce cas, la densité d'énergie décroît plus rapidement que le volume de la boîte ne s'étend. Un tel fluide correspond à de la matière ultra-relativiste, ou radiation, telle que des photons ou des neutrinos ;
- si $w = -\frac{1}{3}$, alors $\rho \propto a^{-2}$. Ce cas correspond au terme de courbure ;
- si $w = -1$, alors $\rho = \text{const.}$ Ceci est un cas inhabituel, car ce fluide a une pression négative. De plus, la densité d'énergie reste constante et le fluide ne se dilue pas avec l'expansion. Ce cas correspond à la constante cosmologique Λ .

Densité critique et paramètres de densité d'énergie Ω_i

Pour un univers géométriquement plat ($k = 0$), la densité critique est définie par :

$$\rho_c(t) \equiv \frac{3H^2(t)}{8\pi G}. \quad (1.16)$$

Un univers ayant une densité supérieure à ρ_c correspond à un univers fermé, tandis qu'un univers de densité inférieure à ρ_c est un univers ouvert. En introduisant la densité critique dans la seconde équation de Friedman, on obtient, en négligeant les interactions matière-radiation après le découplage :

$$\frac{\rho_{\text{rad}}^0}{\rho_c^0} + \frac{\rho_m^0}{\rho_c^0} + \frac{\rho_\Lambda^0}{\rho_c^0} - \frac{\rho_k^0}{\rho_c^0} = 1 \quad (1.17)$$

où les exposants 0 dénotent la valeur actuelle d'un paramètre. La quantité $\rho = \rho_{\text{rad}} + \rho_m$ correspond à la somme de la densité de radiation (matière ultra-relativiste) et de la densité de matière baryonique. Les paramètres $\rho_\Lambda = \frac{\Lambda}{8\pi G}$ et $\rho_k = 3\frac{k}{8\pi G}$ sont les densités d'énergie dues à la constante cosmologique et à la courbure de l'univers respectivement. En introduisant un nouveau paramètre appelé paramètre de densité d'énergie pour un fluide i et défini par :

$$\Omega_i = \frac{\rho_i^0}{\rho_c^0}, \quad (1.18)$$

on obtient :

$$\Omega_{\text{rad}} + \Omega_m + \Omega_\Lambda - \Omega_k = 1. \quad (1.19)$$

On peut ainsi réécrire la seconde équation de Friedman en utilisant à la fois l'évolution des différentes densités d'énergie discutées un peu plus haut et les nouveaux paramètres de densité d'énergie Ω_i :

$$H^2(a) = H_0^2 [\Omega_{\text{rad}} a^{-4} + \Omega_m a^{-3} - \Omega_k a^{-2} + \Omega_\Lambda], \quad (1.20)$$

où $H_0 = 100 h \text{ km s}^{-1} \text{ Mpc}^{-1}$ est la constante de Hubble et est égale à la valeur actuelle du paramètre de Hubble.

Valeurs actuelles des paramètres du modèle Λ CDM

Les valeurs des paramètres cosmologiques Λ CDM utilisées tout au long de cette thèse (sauf pour le calcul des vitesses particulières, voir chapitre 2 sont celles de [Planck Collaboration et al. \(2016\)](#). La constante de Hubble est donnée par $H_0 = 67.74 \pm 0.46$. Les différents paramètres de densité ont pour valeur $\Omega_m = 0.3089 \pm 0.0062$ pour la densité de matière, $\Omega_\Lambda = 0.6911 \pm 0.0062$ pour la densité d'énergie noire, et $\Omega_k = 0.0008_{-0.0039}^{+0.0040}$ pour le paramètre de courbure. Enfin, [Planck Collaboration et al. \(2016\)](#) ont obtenu $w_\Lambda = -1.019_{-0.080}^{+0.075}$ pour le paramètre de l'équation d'état.

La matière noire

En 1933, alors qu'il étudiait les galaxies de l'amas Coma, Fritz Zwicky a obtenu une masse 400 fois plus grande que la valeur attendue à partir de la luminosité des galaxies. Il en a déduit l'existence d'une matière *invisible, noire* entre les galaxies. La notion de cette matière noire fut abandonnée puis réintroduite en 1970 par Vera Rubin et Kent Ford. Ils ont constaté que la vitesse de rotation des galaxies reste constante à grande distance du bulbe alors que les lois de Newton prédisent une baisse de la vitesse avec la distance par rapport au bulbe. Mais Vera Rubin, plus tard, ne croyait plus à la matière noire et penchait pour une modification de la gravitation

On doit donc introduire un nouveau paramètre de densité Ω_i correspondant à la matière noire. La quantité Ω_m devient le paramètre de densité de matière totale dans l'univers :

$$\Omega_m = \Omega_b + \Omega_c, \quad (1.21)$$

où Ω_b décrit la densité de matière baryonique et Ω_c la densité de matière noire.

Les derniers résultats de Planck donnent $\Omega_b = 0.04860 \pm 0.00073$ ainsi que $\Omega_c = 0.2589 \pm 0.0041$. Ces paramètres de densité donnent ainsi la composition de l'univers :

- 4.86% de matière baryonique,
- 25.89% de matière noire,
- et 69.11% d'énergie noire.

1.2 Grandes structures de l'univers

En 1964, les astronomes américains Penzias et Wilson mettent en évidence accidentellement le fond diffus cosmologique (*Cosmic Microwave Background*, CMB). Quelques années plus tard, la collaboration Planck publie la fameuse carte du CMB de Planck. On voit sur cette carte un univers homogène, avec de très petites variations de température autour de 2.7 K. Ceci est la preuve d'un univers jeune homogène.

Cependant les observations aux télescopes montrent l'existence de petites (galaxies) et grandes (super-amas de galaxies) structures. La figure 1.1, extraite de [Huchra et al. \(1985\)](#), présente une tranche du ciel de six degrés d'épaisseur et centrée en $\delta = 29.5^\circ$. L'observateur est situé dans le coin en bas de la figure. La coordonnée radiale correspond au redshift, et la coordonnée angulaire correspond à l'ascension droite. Les points disséminés représentent les galaxies du sondage de redshift CfA. Cette *tranche d'univers* montre que la distribution des galaxies dans l'univers n'est pas aléatoire. Les galaxies semblent former une structure filamentaire : les amas sont reliés par des filaments, qui sont eux-mêmes séparés par des vides.

Pour expliquer comment, dans un univers au départ homogène, de telles structures ont pu se former, le théoricien Alan Guth propose le modèle de l'inflation cosmique en 1981. Selon ce modèle, l'univers aurait subi une phase d'expansion très rapide, nommée *inflation*. Les fluctuations quantiques présentes dans le plasma primordial (avant le découplage matière-radiation) auraient été amplifiées, étendues par l'inflation, et ont ainsi pu donner naissance aux grandes structures que l'on connaît aujourd'hui. L'inflation explique ainsi pourquoi des régions de l'espace, qui n'ont pas de relation causale aujourd'hui, peuvent être à la même température (CMB) : elles l'ont été dans le passé, avant le découplage matière-radiation.

La figure 1.2 montre l'évolution en fonction du temps d'une tranche de surdensité d'une simulation Λ CDM de côté 500 Mpc/h. La tranche de gauche correspond à l'univers il y a 10 milliard d'années (i.e $z \approx 2$), tandis que la tranche de droite représente l'univers actuel (i.e $z = 0$). On peut constater qu'à partir d'un univers homogène, des structures se

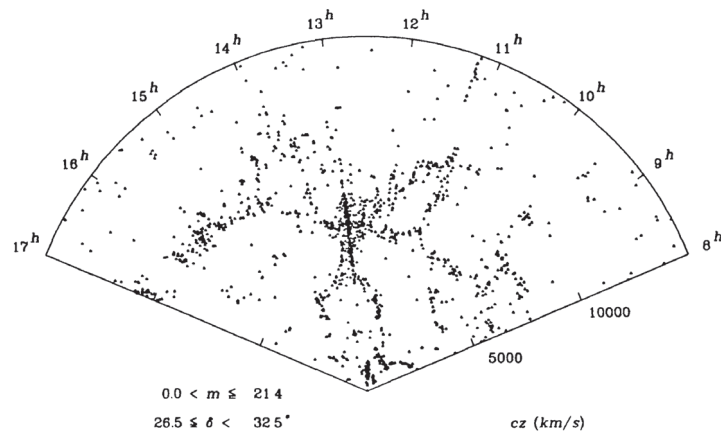


FIGURE 1.1 – Tranche du ciel centrée en $\delta = 29.5^\circ \pm 6^\circ$. L'observateur est situé dans le coin en bas de la figure. La coordonnée radiale correspond au redshift, et la coordonnée angulaire correspond à l'ascension droite. Les points disséminés représentent les galaxies du sondage de redshift CfA. Figure extraite de [Huchra et al. \(1985\)](#)

forment petit à petit, sous l'effet de la compétition entre expansion et la gravitation. En effet, on voit apparaître avec le temps des amas et une structure filamentaire, qui sont de plus en plus prononcés, tandis que les vides s'évacuent.

La *compacité* de l'univers donne des informations sur ces deux lois rivales de la nature qui sont la gravité et l'expansion. En effet, si la gravitation est plus forte et gagne le duel qui a lieu entre ces deux lois, l'univers sera plus compact, tandis que si l'expansion joue un rôle plus important, l'univers sera plus dilué.

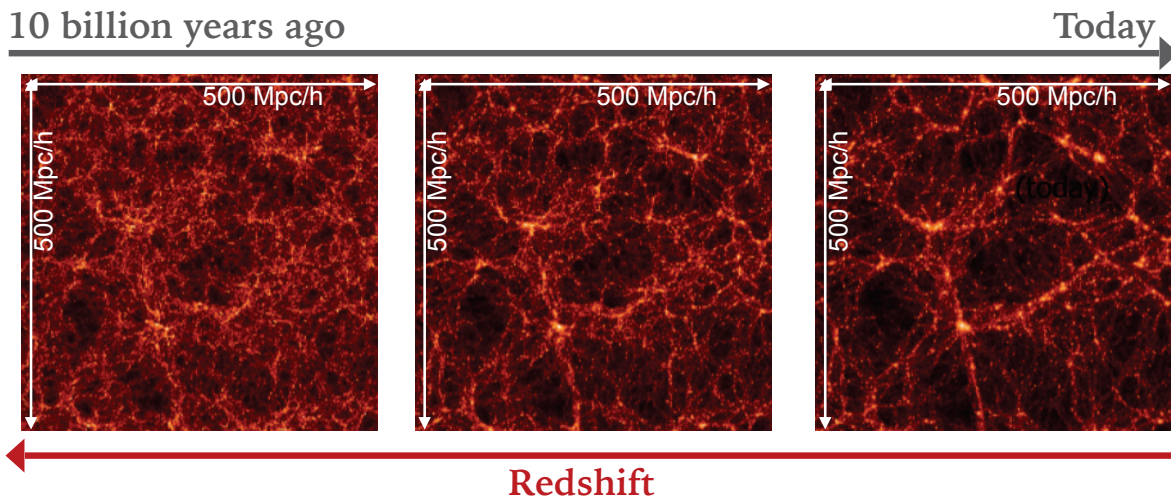


FIGURE 1.2 – Évolution en fonction du temps d'une tranche de surdensité d'une simulation Λ CDM de côté 500 Mpc/h. La tranche de gauche correspond à l'univers il y a 10 milliard d'années (i.e $z \approx 2$), tandis que la tranche de droite représente l'univers actuel (i.e $z = 0$). Sous l'effet de la compétition entre l'expansion et la gravitation, des amas, des vides et une structure filamentaire se forment dans un univers auparavant homogène.

Une méthode permettant de caractériser les grandes structures à l'aide des vitesses par-

ticulières des galaxies est présentée dans le chapitre 4. Cette méthode permet d'identifier des attracteurs et des repoussoirs, ainsi que des bassins et des vallées gravitationnelles.

Il existe deux régimes, valables à deux échelles distinctes, permettant de décrire la formation des structures : le régime linéaire valable à grande échelle (échelles des amas), et le régime non-linéaire, qui est lui valide à l'échelle des galaxies, là où la théorie linéaire n'est plus vérifiée.

1.2.1 Régime linéaire et taux de croissance des grandes structures

Dynamique des perturbations linéaires et coordonnées comobiles

Si l'on modélise la matière en tant que fluide idéal, non relativiste, de densité d'énergie ρ et de pression nulle, située dans un univers de FLRW, alors les trois équations gouvernant le comportement de ce fluide sont les suivantes :

$$\text{Équation de conservation} \quad \frac{\partial \rho}{\partial t}(\vec{r}, t) + \frac{\partial}{\partial \vec{r}}(\rho \vec{u})(\vec{r}, t) = 0; \quad (1.22)$$

$$\text{Équation d'Euler} \quad \frac{\partial \vec{u}}{\partial t}(\vec{r}, t) + \vec{u} \frac{\partial}{\partial \vec{r}} \vec{u}(\vec{r}, t) = - \frac{\partial}{\partial \vec{r}} \Phi(\vec{r}, t); \quad (1.23)$$

$$\text{Équation de Poisson} \quad \frac{\partial^2}{\partial \vec{r}^2} \Phi(\vec{r}, t) = 4\pi G \rho(\vec{r}, t) - \Lambda, \quad (1.24)$$

où \vec{r} représente la position d'une particule du fluide, \vec{u} la vitesse de cette particule et Φ le potentiel gravitationnel. Ces équations sont valables aux échelles inférieures au rayon de Hubble défini par $R_H = \frac{c}{H_0}$ et lorsque ce fluide est éloigné d'un trou noir.

Les équations précédentes sont écrites dans la description eulérienne : les particules étant soumises à l'expansion de l'univers, leurs coordonnées et les distances entre les particules changent avec le temps. Il existe une deuxième description, souvent utilisée en mécanique des fluides : la description lagrangienne. Dans ce cas, les coordonnées d'une particule du fluide ne varient pas en fonction du temps.

En cosmologie, on utilise un compromis entre ces deux descriptions : les coordonnées comobiles. Dans cette description, si les particules sont situées dans un univers non perturbé et sont ainsi seulement soumises à l'expansion de l'univers et à aucune autre force extérieure, alors leurs coordonnées ne changent pas en fonction du temps (cf. description lagrangienne). Cependant si l'univers est perturbé, i.e ces particules subissent une force extérieure qui modifie leurs trajectoires dues à l'expansion, leurs coordonnées varient en fonction du temps. Ces coordonnées permettent ainsi en quelque sorte d'omettre l'expansion de l'univers, tout en prenant en compte les deviations de l'expansion dues aux perturbations locales de l'univers.

Les coordonnées comobiles sont obtenues en divisant les coordonnées eulériennes par le facteur d'échelle $a(t)$:

$$\vec{x} = \vec{r}/a(t). \quad (1.25)$$

La vitesse des particules du fluide peut se décomposer en deux composantes, dont l'une correspond à l'expansion et l'autre à la gravitation :

$$\vec{u} = \dot{a}\vec{x} + \vec{v}(\vec{x}, t), \quad (1.26)$$

où $\dot{a}\vec{x}$ est le flot de Hubble, c'est à dire la composante de la vitesse \vec{v} due à l'expansion de l'univers i.e la vitesse de recession, tandis que $\vec{v}(\vec{x}, t) = \dot{\vec{x}}$ est la vitesse particulière

d'une particule de ce fluide, i.e la composante de \vec{v} déviant de l'expansion de l'univers et attirée par la gravité vers une perturbation.

Les trois equations fondamentales 1.22, 1.23 et 1.24 peuvent être réécrites en coordonnées comobiles :

$$\text{Conservation} \quad \frac{\partial \rho}{\partial t}(\vec{x}, t) + \frac{3\dot{a}}{a}\rho(\vec{x}, t) + \frac{1}{a}\frac{\partial}{\partial \vec{x}}(\rho\vec{v})(\vec{x}, t) = 0; \quad (1.27)$$

$$\text{Euler} \quad \frac{\partial \vec{v}}{\partial t}(\vec{x}, t) + \frac{1}{a}\vec{v}\frac{\partial}{\partial \vec{x}}\vec{v}(\vec{x}, t) + \frac{\dot{a}}{a}\vec{v}(\vec{x}, t) = -\frac{1}{a}\frac{\partial}{\partial \vec{x}}\Phi'(\vec{x}, t); \quad (1.28)$$

$$\text{Poisson} \quad \frac{1}{a^2}\frac{\partial^2}{\partial \vec{x}^2}\Phi'(\vec{x}, t) = 4\pi G\rho(\vec{x}, t) - \Lambda + \frac{3\ddot{a}}{a}, \quad (1.29)$$

où $\Phi'(\vec{x}, t)$ est le nouveau potentiel gravitationnel défini par $\Phi' = \Phi + \frac{1}{2}a\ddot{a}\vec{x}^2$.

Le contraste de densité δ en fonction de la densité moyenne de l'univers $\bar{\rho}$ est donné par :

$$\delta(\vec{x}, t) = \frac{\rho(\vec{x}, t) - \bar{\rho}(t)}{\bar{\rho}(t)}. \quad (1.30)$$

Ce paramètre permet d'indiquer où se situent les déviations de la densité moyenne dans l'univers.

Considérons de petites perturbations par rapport à l'univers homogène, c'est-à-dire pour de faibles contrastes de densité $\delta(\vec{x}, t) \ll 1$, et supposons que les vitesses sont suffisamment faibles ($v t_{\text{exp}}/d_{\text{coh}} \ll \delta$ où v est la vitesse caractéristique de la perturbation, $t_e \approx (G\bar{\rho})^{-1/2}$ le temps d'expansion et d_{coh} la longueur de cohérence pour les variations spatiales de δ). Il est possible dans ce cas de linéariser le groupe de trois équations (1.27, 1.28, 1.29). On obtient :

$$\text{Conservation} \quad \frac{\partial \delta}{\partial t}(\vec{x}, t) + \frac{1}{a}\frac{\partial \vec{v}}{\partial \vec{x}}(\vec{x}, t) = 0; \quad (1.31)$$

$$\text{Euler} \quad \frac{\partial^2 \delta}{\partial t^2}(\vec{x}, t) + 2\frac{\dot{a}}{a}\frac{\partial \delta}{\partial t}(\vec{x}, t) = 4\pi G\rho_b\delta(\vec{x}, t); \quad (1.32)$$

$$\text{Poisson} \quad \left(\frac{\partial}{\partial \vec{x}}\right)^2 \Phi'(\vec{x}, t) = 4\pi G\rho_b a^2 \delta(\vec{x}, t). \quad (1.33)$$

Les trois équations ci-dessus décrivent la croissance linéaire des fluctuations de densité, pour un fluide idéal sans pression contenu dans un univers en expansion. On peut montrer que, dans l'approximation linéaire, les deux solutions de ces équations pour δ ont la forme :

$$\delta \propto D_d(a) = H_0 (\Omega_m a^{-3} + \Omega_k a^{-2} + \Omega_\Lambda)^{1/2} \quad \text{Mode décroissant}; \quad (1.34)$$

$$\delta \propto D_c(a) = D_d(a) \int_0^a \frac{1}{a'^3 D_d^3(a')} da' \quad \text{Mode croissant}. \quad (1.35)$$

Taux de croissance des grandes structures

Le mode croissant, défini dans l'équation 1.35, permet de mettre en place un nouveau paramètre cosmologique appelé *taux de croissance des structures*, qui sera mesuré sur des données observationnelles dans le chapitre 5 de cette thèse. Ce paramètre symbolise le duel qui a lieu entre l'expansion et la gravitation, et est ainsi lié à la compacité de l'univers.

Le taux de croissance des grandes structures f paramètre est donné par :

$$f \equiv \frac{d \ln D_c(a)}{d \ln a} = \frac{a}{D_c(a)} \frac{dD_c(a)}{da}. \quad (1.36)$$

On peut exprimer f en fonction de a , Ω_m et Ω_Λ simplement en dérivant le mode croissant par a . Cela donne :

$$f(\Omega_m, \Omega_\Lambda, a) = X^{-2} \left(\Omega_\Lambda - \frac{\Omega_m}{2} a^{-3} \right) - 1 + \frac{X^{-3} a^{-2}}{\int_0^a X^{-3} a^{-3} da}, \quad (1.37)$$

avec $X(a) = a^{-1} \sqrt{1 + \Omega_m(a^{-1} - 1) + \Omega_\Lambda(a^2 - 1)}$.

Peebles (1980) et Lahav et al. (1991) ont proposé une approximation : $f = \Omega_m^{0.6}$. D'autre part, Linder (2005) propose d'introduire un nouveau paramètre cosmologique : l'indice de croissance γ , qui permet d'écrire le taux de croissance f comme une puissance de Ω_m :

$$f = \Omega_m^\gamma \quad (1.38)$$

La valeur de γ change selon la théorie de gravité considérée. Les paramètres f et γ sont ainsi utilisés pour sonder la théorie de gravité sous-jacente, et la cause de l'accélération de l'expansion de l'univers. Voici quelques valeurs de γ pour différents modèles cosmologiques et différentes théories de la gravitation (Linder, 2005) :

- $\gamma = \frac{6}{11} = 0.545$ pour les modèles de Relativité Générale, dont Λ CDM ;
- $\gamma = \frac{11}{10} = 0.687$ pour le modèle DGP, modèle de gravité modifiée.

La figure 1.3 montre le taux de croissance $f = \Omega_m^\gamma$ en fonction du redshift z pour différentes valeurs de γ , allant de $\gamma = 0.3$ à $\gamma = 0.9$. On voit que la valeur du taux de croissance augmente avec γ quel que soit le redshift. De plus, on peut remarquer que les valeurs obtenues avec différents γ sont plus espacées à petit redshift qu'à grand redshift. Il serait donc plus facile de contraindre les modèles cosmologiques à petit redshift. On verra cependant dans le chapitre 5 que les mesures locales de f sont limitées par la taille des sondages cosmologiques proches, à la fois en nombre de points considérés (erreurs statistiques) et en volume (variance cosmique), en plus des erreurs sur les mesures observationnelles.

Cependant le taux de croissance f est dégénéré avec le paramètre σ_8 , représentant les fluctuations de densité dans une sphère de rayon 8 Mpc/h. On mesure donc observationnellement la combinaison $f\sigma_8$.

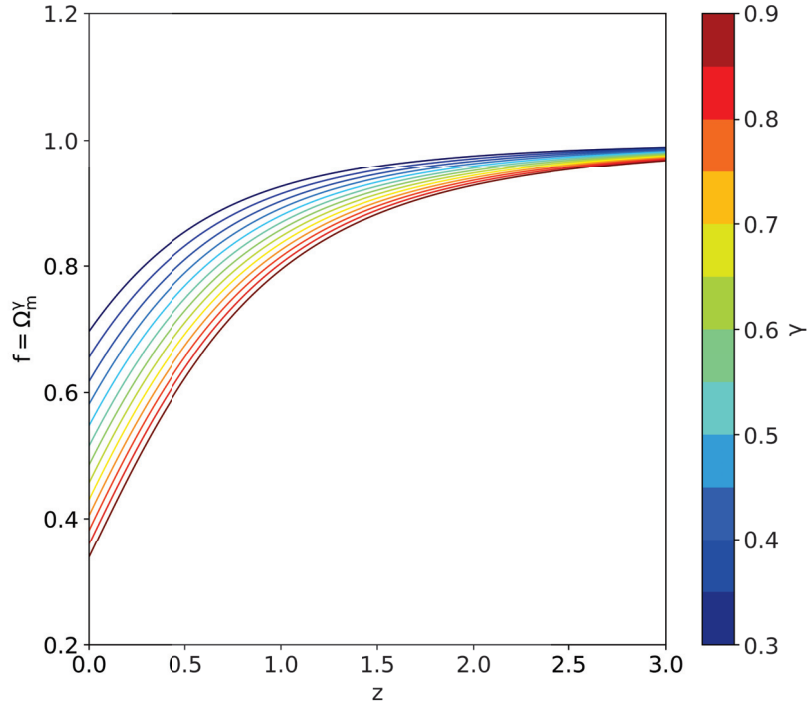


FIGURE 1.3 – Taux de croissance $f = \Omega_m^\gamma$ en fonction du redshift z et pour différentes valeurs de γ , allant de $\gamma = 0.3$ à $\gamma = 0.9$.

La figure 1.4 montre le taux de croissance normalisé $f\sigma_8$ en fonction du redshift z , donné pour différentes valeurs de γ , allant de 0.45 (tirets) à 0.70 (pointillés). La courbe en trait plein correspond au modèle Λ CDM avec $\gamma = 0.55$. Les points disséminés avec barres d'erreur montrent diverses contraintes du taux de croissance extraites de la littérature qui sont présentées dans la table 1.1. Les couleurs dépendent de la sonde cosmologique utilisée pour estimer le taux de croissance : les SN Ia en jaune (Turnbull et al., 2012), les vitesses particulières en bleu (Johnson et al., 2014), les oscillations acoustiques des baryons en vert (Blake et al., 2011; Reid et al., 2012) et les distortions dans l'espace des redshifts en rose (Hawkins et al., 2003; Samushia et al., 2012; Beutler et al., 2012; de la Torre et al., 2013). Les prévisions pour de futurs relevés cosmologiques sont représentés en rouge, avec une erreur de mesure prédite de 10 % (McConnachie et al., 2016; Amendola et al., 2016; da Cunha et al., 2017). Aucune de ces mesures de $f\sigma_8$ ne permet de contraindre fortement γ , les incertitudes de mesure étant beaucoup trop grandes.

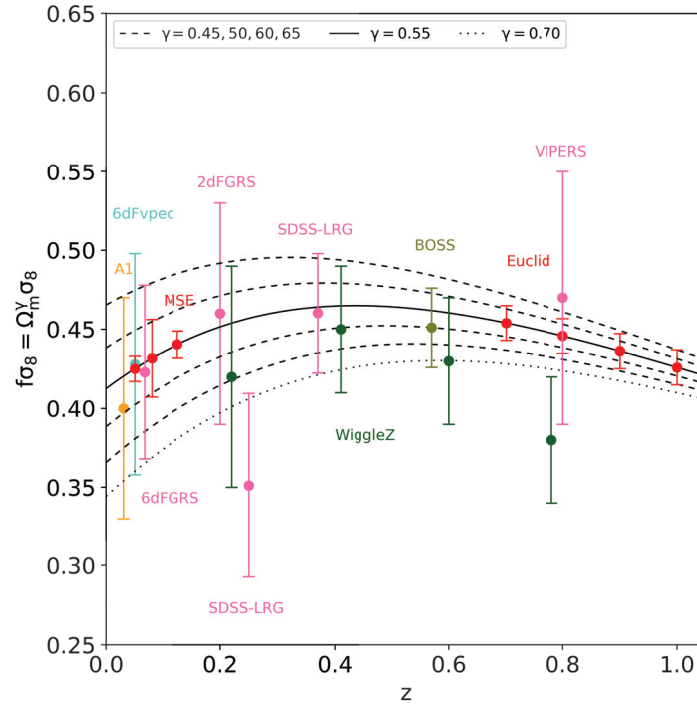


FIGURE 1.4 – Taux de croissance normalisé $f\sigma_8$ en fonction du redshift z , donné pour différentes valeurs de γ , allant de 0.45 (tirets) à 0.70 (pointillés). La courbe en trait plein correspond au modèle Λ CDM avec $\gamma = 0.55$. Les points avec barres d'erreur correspondent à diverses contraintes de $f\sigma_8$ issues de la littérature. Les couleurs dépendent de la sonde cosmologique utilisée pour estimer le taux de croissance : les SN Ia en jaune, les vitesses particulières en bleu, les oscillations acoustiques des baryons en vert et les distortions dans l'espace des redshifts en rose. Les prévisions pour de futurs relevés cosmologiques sont représentées en rouge, avec une erreur de mesure prédite de 10%.

1.2.2 Croissance non-linéaire : à l'échelle des galaxies

La théorie linéaire décrite dans la section ci-dessus a tout de même une limite : elle n'est plus valable au-delà d'une certaine échelle, celle des galaxies, i.e $\delta \gg 1$. Afin de décrire l'évolution de l'univers et des structures au-delà de cette échelle, on utilise la théorie *non-linéaire* de la croissance des structures dans l'univers.

Il y a cependant plusieurs solutions : soit on utilise des modèles analytiques décrivant l'évolution non-linéaire des grandes structures de l'univers, soit on utilise des simulations numériques.

Il existe en effet une variété de modèles permettant de décrire l'évolution non-linéaire d'une perturbation de manière complètement analytique. Ces deux théories vont être détaillées rapidement par la suite : le modèle d'effondrement sphérique et l'approximation de Zel'dovich.

TABLE 1.1 – Contraintes de $f\sigma_8$ issues de la littérature. Ces mesures sont obtenues à partir de diverses sondes cosmologiques : les vitesses particulières (Vpec), les supernovae de type Ia (SN Ia), les distortions dans l'espace des redshifts (RSD) et les oscillations acoustiques des baryons (BAO).

	Redshift z	Growth rate $f\sigma_8$	Publication	Cosmological probe
6dFGS	0.05	$0.428^{+0.079}_{-0.068}$	Johnson et al. (2014)	Vpec
A1	0.03	0.40 ± 0.07	Turnbull et al. (2012)	SN Ia
2dFGRS	0.20	0.46 ± 0.07	Hawkins et al. (2003)	RSD
6dFGRS	0.067	0.423 ± 0.055	Beutler et al. (2012)	RSD
WiggleZ	0.22	0.42 ± 0.07	Blake et al. (2011)	BAO
	0.41	0.45 ± 0.04		
	0.60	0.43 ± 0.04		
	0.78	0.38 ± 0.04		
SDSS-LRG	0.25	0.3512 ± 0.0583	Samushia et al. (2012)	RSD
	0.37	0.4602 ± 0.0378		
BOSS	0.57	0.451 ± 0.025	Reid et al. (2012)	BAO
VIPERS	0.80	0.47 ± 0.08	de la Torre et al. (2013)	RSD

Modèle d'effondrement sphérique "top-hat"

En considérant une perturbation surdense, de symétrie sphérique et évoluant dans un univers homogène, dans le régime linéaire, la surdensité de cette perturbation est donnée par $\delta_{\text{lin}} = \delta_0 a(t)$.

Le théorème de Birkoff dit qu'une telle perturbation se comporte comme un univers indépendant. Cela signifie que le champ de matière extérieur n'a aucune influence sur l'évolution de la perturbation. L'évolution d'une perturbation surdense peut être décrite par la forme paramétrique suivante :

$$r(\theta) = A(1 - \cos \theta); \quad (1.39)$$

$$t(\theta) = B(\theta - \sin \theta), \quad (1.40)$$

où θ correspond au rayon de la perturbation sphérique. L'équation de mouvement de la couche sphérique externe de la perturbation $\ddot{r} = -GM/r^2$ donne $A^3 = GMB^2$. Pour t petit, il est possible de développer l'équation précédente pour obtenir :

$$r \approx \frac{A}{2} \left(\frac{6t}{B} \right)^{2/3} \left[1 - \frac{1}{20} \left(\frac{6t}{B} \right)^{2/3} \right] \quad (1.41)$$

On peut identifier la surdensité de la perturbation sphérique :

$$\delta \approx \frac{3}{20} \left(\frac{6t}{B} \right)^{2/3}. \quad (1.42)$$

Par conséquent, dans un premier temps, la perturbation sphérique suit le flot de Hubble ($a \propto t^{2/3}$) et sa surdensité évolue proportionnellement à a . La perturbation évolue donc dans le régime linéaire jusqu'à ce que $\delta \approx 1$. À ce moment-là, la perturbation s'effondre et son évolution n'est plus décrite par le régime linéaire mais par le régime non-linéaire.

Approximation de Zel'dovich

En considérant que le déplacement initial des particules se fait dans un champ de matière homogène et sans collision et en supposant que ces particules continuent de se déplacer dans leur direction initiale, les coordonnées eulériennes d'une particule de ce champ à un temps t sont données par, dans l'approximation de Zel'dovich :

$$\vec{r}(\vec{q}, t) = a(t)(\vec{q} + b(t)\vec{s}(\vec{q})). \quad (1.43)$$

L'équation 1.43 correspond au flot de Hubble avec une perturbation qui devient négligeable quand $t \rightarrow 0$. La quantité \vec{q} correspond aux coordonnées lagrangiennes. La fonction $b(t) = a(t)D_c(t)$ mesure l'ampleur du déplacement $\vec{s}(\vec{q}) = \nabla\Phi(\vec{q})$ de la particule.

A $t > 0$, des inhomogénéités se forment dans le champ de densité. L'équation de conservation de la masse est donnée par :

$$\rho(\vec{r}, t)d\vec{r} = \rho_0 d\vec{q}, \quad (1.44)$$

et cela donne :

$$\rho(\vec{q}, t) = \rho_0 \left| \frac{d\vec{r}}{d\vec{q}} \right| = \frac{\bar{\rho}(t)}{\left| \delta_{ij} - b(t) \frac{\partial s_i}{\partial q_j} \right|}. \quad (1.45)$$

La quantité $\bar{\rho}(t)$ est la densité moyenne au temps t et δ_{ij} est le symbole de Kronecker où i et j sont les composantes cartésiennes. La quantité $\frac{\partial s_i}{\partial q_j}$ définit le tenseur des déformations et prend en compte l'évolution du champ de densité sous l'effet de la gravitation. Ce tenseur est une matrice réelle symétrique. L'équation 1.45 peut ainsi être réécrite en fonction des valeurs propres $(-\alpha(\vec{q}), -\beta(\vec{q}), -\gamma(\vec{q}))$ du tenseur des déformations, qui représentent les déformations principales selon les trois axes principaux donnés par les vecteurs propres de $\frac{\partial s_i}{\partial q_j}$:

$$\frac{\rho(\vec{q}, t)}{\bar{\rho}} = [(1 - b(t)\alpha)(1 - b(t)\beta)(1 - b(t)\gamma)]^{-1}. \quad (1.46)$$

En conséquence, on observe tout d'abord un effondrement dans la direction de l'axe principal défini par la plus grande valeur propre négative. Cela correspond à la formation de structures aplaties appelées *pancakes* (ou *sheets*) par contraction dans la direction d'un des trois axes principaux.

Chapitre 2

Données observationnelles Cosmicflows

Dans cette thèse, ce sont les vitesses particulières des galaxies qui vont permettre de caractériser les grandes structures de l'univers. Ces vitesses correspondent à la composante de la vitesse totale d'une galaxie résultant donc de la gravitation et non pas de l'expansion de l'univers. Cependant, les observations donnent accès uniquement à la partie radiale de la vitesse totale d'une galaxie. En connaissant la distance d'une galaxie et sa vitesse radiale, il est possible d'en déduire la composante radiale de sa vitesse particulière.

Ce chapitre explique tout d'abord comment les distances de galaxies par rapport à la Terre sont obtenues, et comment les vitesses particulières radiales des galaxies peuvent être déduites de ces distances. Puis il décrit les catalogues de distances *Cosmicflows*, dont la dernière version *Cosmicflows-3*, qui seront par la suite utilisés dans les chapitres 4 et 5. Pour terminer, seront présentées une future version de la série de catalogues *Cosmicflows* : *Cosmicflows-4*. La préparation de ce catalogue à travers les observations à 21 cm sera plus particulièrement détaillée dans le chapitre 3.

Sommaire

2.1	Distances et vitesses particulières observées	36
2.1.1	Distances	36
2.1.2	Vitesses particulières	38
2.2	Catalogues de distances de galaxies Cosmicflows	39
2.2.1	Les débuts : Cosmicflows-1 et Cosmicflows-2	39
2.2.2	<i>Longius</i> — Cosmicflows-3 : encore plus loin ... vers le Sud . . .	40
2.2.3	<i>Altius</i> — Cosmicflows-4 : encore plus haut ... vers le Nord . . .	47

2.1 Distances et vitesses particulières observées

2.1.1 Distances

Pour estimer la vitesse particulière d'une galaxie, il faut tout d'abord connaître sa distance. Cependant les distances des galaxies dans l'univers ne sont pas accessibles simplement par mesure directe. Il faut donc trouver des outils permettant d'accéder indirectement à ces distances : leur luminosité.

Distance de luminosité : la relation entre le flux F et la luminosité L d'un objet céleste est donnée par :

$$D_L = \sqrt{\frac{L}{4\pi F}} , \quad (2.1)$$

où la quantité D_L correspond à la distance de luminosité.

Échelle logarithmique des distances : module de distance et magnitudes

Le module de distance μ décrit les distances des objets célestes sur une échelle logarithmique. Il est défini par :

$$\mu = m - M = 5 \log_{10}(D_L[\text{pc}]) - 5 = 5 \log_{10}(D_L[\text{Mpc}]) + 25, \quad (2.2)$$

où $D_L[\text{pc}]$ et $D_L[\text{Mpc}]$ correspondent à la distance de luminosité exprimée respectivement en pc et en Mpc.

Par ailleurs, la magnitude apparente m représente la luminosité d'un objet céleste perçue par un observateur depuis la Terre. Elle est inversement reliée à la brillance apparente d'un objet : plus un objet apparaît brillant dans le ciel, plus sa magnitude apparente est petite. Les magnitudes apparentes sont obtenues par photométrie dans une bande spectrale donnée, comme nous le verrons avec les images obtenues par PanSTARRS dans les bandes g , r , i , z et y dans le chapitre 3. La magnitude apparente est définie par :

$$m = -2.5 \log_{10} \left(\frac{F}{F_0} \right), \quad (2.3)$$

où F est le flux mesuré et F_0 le flux de référence dont la valeur dépend de la bande spectrale considérée.

Enfin, le flux d'un objet céleste étant inversement proportionnel à sa distance au carré, il est possible d'obtenir la même magnitude apparente pour un objet situé à une distance donnée et un autre objet quatre fois plus lumineux et deux fois plus lointain. Afin de mesurer la luminosité intrinsèque d'un objet, on utilise la magnitude absolue M donnée

par :

$$M = -2.5 \log_{10} \left(\frac{L}{L_0} \right), \quad (2.4)$$

où L est la luminosité mesurée et $L_0 = 3.0128 \times 10^{28}$ W est la luminosité de référence. La magnitude absolue correspond à la magnitude apparente qu'un objet aurait si il était placé à une distance de 10 pc.

Pour la suite, la distance de luminosité D_L d'une galaxie représente sa distance par rapport à la Terre et sera notée d .

Méthodes de mesure des distances extragalactiques

Il existe une grande variété de méthodes permettant de déterminer les distances des galaxies. Une de ces méthodes sera présentée plus en détails dans le chapitre 3 : la méthode Tully-Fisher. Les autres méthodes utilisées par la collaboration *Cosmicflows* sont évoquées ci-dessous.

Relation période-luminosité des Céphéides : les céphéides sont des étoiles jeunes variables, i.e à pulsations périodiques. Cela signifie que l'on observe des variations de diamètre et de température, et donc de luminosité. Ces étoiles ont une période de pulsation, qui correspond au temps qu'il faut pour retrouver la magnitude de départ. Dans les années 1910-1920, Henrietta Leavitt remarque que les Céphéides ayant des périodes longues sont intrinsèquement plus lumineuses que celles ayant une période plus courte. Elle vient de découvrir la relation période-luminosité (PL) des Céphéides. Cette relation permet de relier la période de pulsation au module de distance et ainsi estimer la distance d'une galaxie par rapport à la Terre.

Sommet de la branche des géantes rouges, ou *Tip of the Red Giant Branch* (TRGB) : cette méthode utilise la luminosité des étoiles appartenant à la branche des géantes rouges en tant que chandelle standard pour estimer la distance de la galaxie hôte. La collaboration *Cosmicflows* utilise le télescope spatial Hubble (*Hubble Space Telescope*, HST) pour effectuer ce type de mesure.

Ces deux méthodes sont très précises mais ne s'appliquent qu'aux galaxies proches. Elles servent par ailleurs à calibrer des méthodes moins précises mais qui permettent d'estimer les distances de galaxies plus lointaines.

Fluctuations de la brillance de surface, ou *Surface Brightness Fluctuations* (SBF) : cette méthode s'applique aux galaxies elliptiques, car la morphologie des galaxies spirales peuvent perturber leur brillance de surface (poussière, amas globulaires). Les galaxies elliptiques ont une brillance de surface lisse et régulière, et les galaxies plus lointaines apparaissent plus lisses que les galaxies proches. Cependant chaque élément de volume contient un nombre fini d'étoiles, et ce nombre varie d'un élément de volume à un autre. Cela génère des fluctuations dans la brillance de surface de la galaxie, qui peuvent être décrites par une distribution de Poisson. Plus la galaxie est lointaine, plus il y a d'étoiles dans les différents éléments de volume, et donc plus les fluctuations de la brillance de

surface sont petites. Ces variations de la brillance de surface dépendent de l'éloignement d'une galaxie et permettent d'extraire sa distance (Tonry and Schneider, 1988; Tonry et al., 1997, 2001).

Supernovae de type Ia (SNIa) : les supernovae de type Ia correspondent à l'explosion d'une naine blanche en fin de vie. L'explosion se produit lorsque la masse de la naine blanche dépasse une certaine masse limite. Elle a donc toujours lieu à la même masse, et a donc toujours la même luminosité. Cela permet d'utiliser les SNIa comme chandelles standards : on connaît toujours leur luminosité dont on peut déduire la distance des galaxies hôtes. De plus, les explosions de supernovae étant très brillantes, elles permettent d'estimer les distances de galaxies très lointaines, allant jusqu'à environ 10 Gpc.

Plan fondamental des galaxies elliptiques, *Fundamental Plane* (FP) : cette méthode utilise la relation entre le rayon effectif R_{eff} (rayon auquel la moitié de la brillance de surface est confinée), la brillance de surface moyenne $\langle \mu_{\text{eff}} \rangle$ dans R_{eff} et la dispersion des vitesses des étoiles au centre d'une galaxie elliptique σ_0 . Ces trois caractéristiques sont liées, et permettent ainsi d'estimer la distance des galaxies elliptiques. Le plan fondamental représente la surface sur laquelle est placée une galaxie elliptique, dans un espace de paramètres à trois dimensions ($R_{\text{eff}}, \langle \mu_{\text{eff}} \rangle, \sigma_0$). On a alors :

$$\log_{10}(R_{\text{eff}}) = a \log_{10} \sigma_0 + b \langle \mu_{\text{eff}} \rangle + c, \quad (2.5)$$

où a , b et c sont des constantes. Ainsi, en mesurant la brillance de surface, la dispersion des vitesses et la taille angulaire de la galaxie, il est possible de déduire le rayon effectif, et ainsi sa distance par trigonométrie.

2.1.2 Vitesses particulières

Les distances d des galaxies permettent d'estimer la partie radiale v_p seulement de leur vitesse particulière. Il existe plusieurs estimateurs. Le plus utilisé mais le moins précis est le suivant :

$$v_p = cz - H_0 d, \quad (2.6)$$

où z est le redshift observé, c la vitesse de la lumière dans le vide et H_0 la constante de Hubble. L'incertitude sur v_p est $\delta v_p = H_0 \delta d$ où δd est l'erreur sur la mesure de d (en general $\delta d \sim 20\%$). Cependant cet estimateur n'est plus valable à grand redshift.

Un autre estimateur prenant en compte cette fois-ci l'accélération de l'expansion de l'univers et donc valable à grand redshift a été introduit par Davis and Scrimgeour (2014) :

$$v_p = \frac{cz_{\text{mod}} - H_0 d}{1 + z_{\text{mod}}}, \quad (2.7)$$

où $z_{\text{mod}} = z \left(1 + 0.5(1 - q_0)z - \frac{1}{6}(1 - q_0 - 3q_0^3 + 1)z^2 \right)$ dépend d'un modèle cosmologique à travers le paramètre de décélération $q_0 = 0.5(\Omega_m - 2\Omega_\Lambda)$. L'incertitude est alors définie

ainsi : $\delta v_p = \frac{H_0 \delta d}{1+z_{\text{mod}}}$.

Seulement, un problème se pose. La distribution des erreurs sur v_p obtenues par ces deux estimateurs n'est pas Gaussienne pour les raisons suivantes : les erreurs sur le module de distance μ sont, elles, Gaussiennes ; mais ce n'est pas le cas pour les erreurs sur la distance. En effet, passer par $10''$ pour obtenir la distance biaise la distribution des erreurs, et biaise ainsi la distribution des erreurs sur la vitesse particulière.

[Watkins and Feldman \(2015\)](#) proposent par la suite un nouvel estimateur permettant de corriger ce problème :

$$v_p = cz \log \left(\frac{cz}{H_0 d} \right). \quad (2.8)$$

Les erreurs sur les vitesses particulières calculées avec cet estimateur sont bien Gaussiennes à cause de la présence du logarithme dans l'équation 2.8. L'incertitude est alors de $\delta v_p = cz \delta(\log d)$ où $\delta(\log d)$ est l'erreur sur le logarithme de la distance. L'équation 2.9 ci-dessous, valable à grand redshift, est celle qui sera utilisée pour calculer les vitesses particulières dans cette thèse :

$$v_p = \frac{cz_{\text{mod}}}{1+z_{\text{mod}}} \log \left(\frac{cz_{\text{mod}}}{H_0 d} \right), \quad (2.9)$$

avec comme incertitude $\delta v_{\text{mod}} = \frac{cz_{\text{mod}} \delta(\log d)}{1+z_{\text{mod}}}$.

2.2 Catalogues de distances de galaxies *Cosmicflows*

Les compilations de distances de galaxies permettent ainsi d'avoir accès à leur vitesse particulière. La collaboration *Cosmicflows* prépare de telles compilations.

2.2.1 Les débuts : *Cosmicflows-1* et *Cosmicflows-2*

Cosmicflows-1

Cosmicflows-1 (CF1, [Tully et al., 2008](#)) est le tout premier catalogue du projet *Cosmicflows*. C'est une extension du *Nearby Galaxies Catalog* ([Tully, 1988](#)). Il s'agit d'une compilation de nouvelles distances et de distances extraites de la littérature. CF1 fournit des distances précises pour 1,797 galaxies dans 743 groupes et jusqu'à 3,000 km s⁻¹. Les distances sont obtenues par quatre méthodes différentes. La majorité des distances sont obtenues par la relation de Tully-Fisher et proviennent de [Tully and Pierce \(2000\)](#) (1,023 distances) et [Karachentsev et al. \(2002\)](#) (402 distances utilisant la photométrie de 2MASS dans la bande k). Cette compilation contient aussi 353 distances estimées par SBF ([Tonry et al., 2001](#); [Mei et al., 2007](#)), 209 distances obtenues par la méthode TRGB ([Karachentsev et al., 2004, 2006](#)) et 29 distances mesurées avec la relation PL des Céphéides

(*HST Cepheid Key Project* [Freedman et al., 2001](#)). Une incertitude de mesure de 20% en moyenne a été assignée à toutes les distances obtenues par la relation Tully-Fisher. Les autres méthodes étant plus précises, une erreur de 10 % leur a été attribuée.

Cosmicflows-2

Le catalogue *Cosmicflows-2* (CF2, [Tully et al., 2013](#)) est une extension de CF1. Ce catalogue contient à la fois des distances extraites de la littérature et des nouvelles distances obtenues par les auteurs. Il donne accès à une distance pour 8,188 galaxies dans 4,899 groupes, en allant jusqu'à 30,000 km s⁻¹. Cependant la majorité des galaxies vont seulement jusqu'à 8,000 km s⁻¹. Un total de 60 distances a été obtenu par la relation PL des Céphéides et 304 distances TRGB ont été acquises par des observations au télescope spatial Hubble [Jacobs et al. \(2009\)](#). De plus, 101 distances TRGB ont été extraites de la littérature, ainsi que 20 distances obtenues par méthodes diverses (*RR Lyrae*, *Horizontal Branch*, *Eclipsing Binary*), 387 distances mesurées par SBF et 280 distances tirées d'observations de SNIa (et provenant en majorité de Union2, [Amanullah et al., 2010](#)). Par ailleurs, 5,954 distances ont été obtenues par la relation Tully-Fisher. Les observations HI ont été effectuées aux radiotélescopes de Green Bank, Arecibo et Parkes, et l'analyse des données provient des archives Arecibo, Nancay, Parkes et Effelsberg. Par ailleurs, la photométrie vient de Spitzer ([Courtois et al., 2011a](#); [Springob et al., 2007](#); [Courtois et al., 2011b](#)) et SDSS [Hall et al. \(2012\)](#). Enfin, 1,458 distances ont été estimées par la méthode FP.

2.2.2 *Longius* — Cosmicflows-3 : encore plus loin ... vers le Sud

Dans cette thèse, on utilisera le tout dernier catalogue du projet *Cosmicflows* : le catalogue *Cosmicflows-3* (CF3, [Tully et al., 2016](#)). Il s'agit d'une extension de CF2, qui donne cette fois-ci accès aux distances d'un total de 17,648 galaxies dans 11,936 groupes, avec une couverture en redshift allant jusqu'à 16,000 km s⁻¹.

Les ajouts majeurs de ce nouveau catalogue sont les suivants :

- 2,257 distances obtenues par Tully-Fisher. La photométrie provenant du télescope spatial Spitzer à 3.6 μm a été utilisée et extraite en majorité des programmes S4G (*Spitzer Survey of Stellar Structure in Galaxies*, [Sheth et al., 2010](#)) et CFS (Cosmic Flows with Spitzer, [Sorce et al., 2014](#)) ;
- 8,885 distances obtenues avec la méthode FP par la collaboration *Six Degree Field Galaxy Survey* (6dFGS, [Springob et al., 2014](#)). Les mesures de dispersions des vitesses centrales de galaxies elliptiques de 6dFGS proviennent de [Jones et al. \(2009\)](#), et la photométrie est issue de 2MASS ([Jarrett et al., 2000b](#)). Cet échantillon contient des galaxies lointaines. Il a en effet une couverture en redshift allant jusqu'à 16,000 km/s ;
- 87 nouvelles distances TRGB obtenues en utilisant les données provenant des archives HST ;
- enfin, la dernière contribution est celle des SNIa. Deux échantillons ont été retenus, tous les deux dans un redshift $z < 0.1$. Le premier est issu de [Rest et al. \(2014\)](#), avec

58 nouvelles distances. Le second échantillon provient de [Walker et al. \(2015\)](#) avec 29 distances. Malgré leur nombre, les distances mesurées avec les SNIa constituent tout de même un apport important. En effet, quelle que soit la méthode utilisée, les incertitudes sur les mesures de vitesses particulières augmentent avec le redshift. Or, cette erreur est au moins deux fois moins grande en utilisant les SNIa qu’avec les relations TF ou FP.

La compilation *Cosmicflows-3* fournit donc à la fois un catalogue non groupé contenant des distances de galaxies, et un catalogue groupé donnant accès aux distances pour des ensembles de galaxies. Les groupes ont été construits par [Tully \(2015\)](#) à partir des 43,526 galaxies de 2MASS (catalogue le plus complet en magnitude $K_s < 11.75$). Cela permet de diminuer les erreurs sur la distance d’un groupe, i.e d’un ensemble de galaxies (l’erreur est divisée par \sqrt{N} où N est le nombre de galaxies dans un groupe donné).

La figure [2.1](#) présente la distribution des galaxies (haut) et des groupes de galaxies (bas) du catalogue *Cosmicflows-3* en coordonnées supergalactiques. La couleur correspond à l’éloignement des galaxies (ou des groupes) en redshift. Certains superamas sont identifiés parmi les galaxies, tels que Virgo, Fornax, Persée-Poisson et Shapley. On voit que l’on passe ainsi d’environ 18,000 galaxies individuelles avec une erreur moyenne sur la distance de 26 Mpc, à 12,000 groupes dont l’erreur moyenne est de 24 Mpc. Cependant, il devient plus difficile d’identifier les amas de galaxies dans la distribution des groupes dans le ciel. En effet, il n’y a plus qu’une seule mesure de distance pour l’amas Virgo dans le catalogue groupé, et cela ne correspond plus qu’à un point sur la carte.

Comparaison des catalogues *Cosmicflows*

La table [2.1](#) résume toutes les caractéristiques des trois catalogues CF1, CF2 et CF3 présentés ci-dessus.

TABLE 2.1 – Évolution des catalogues *Cosmicflows-1*, *Cosmicflows-2* et *Cosmicflows-3*. Des prévisions pour le prochain catalogue *Cosmicflows-4* sont aussi données.

	Référence	Galaxies	Groupes	Limite
CF1	Tully et al. (2008)	1,797	743	3,000
CF2	Tully et al. (2013)	8,188	4,899	8,000
CF3	Tully et al. (2016)	17,648	11,936	16,000
CF4	vers 2019	environ 25,000		16,000

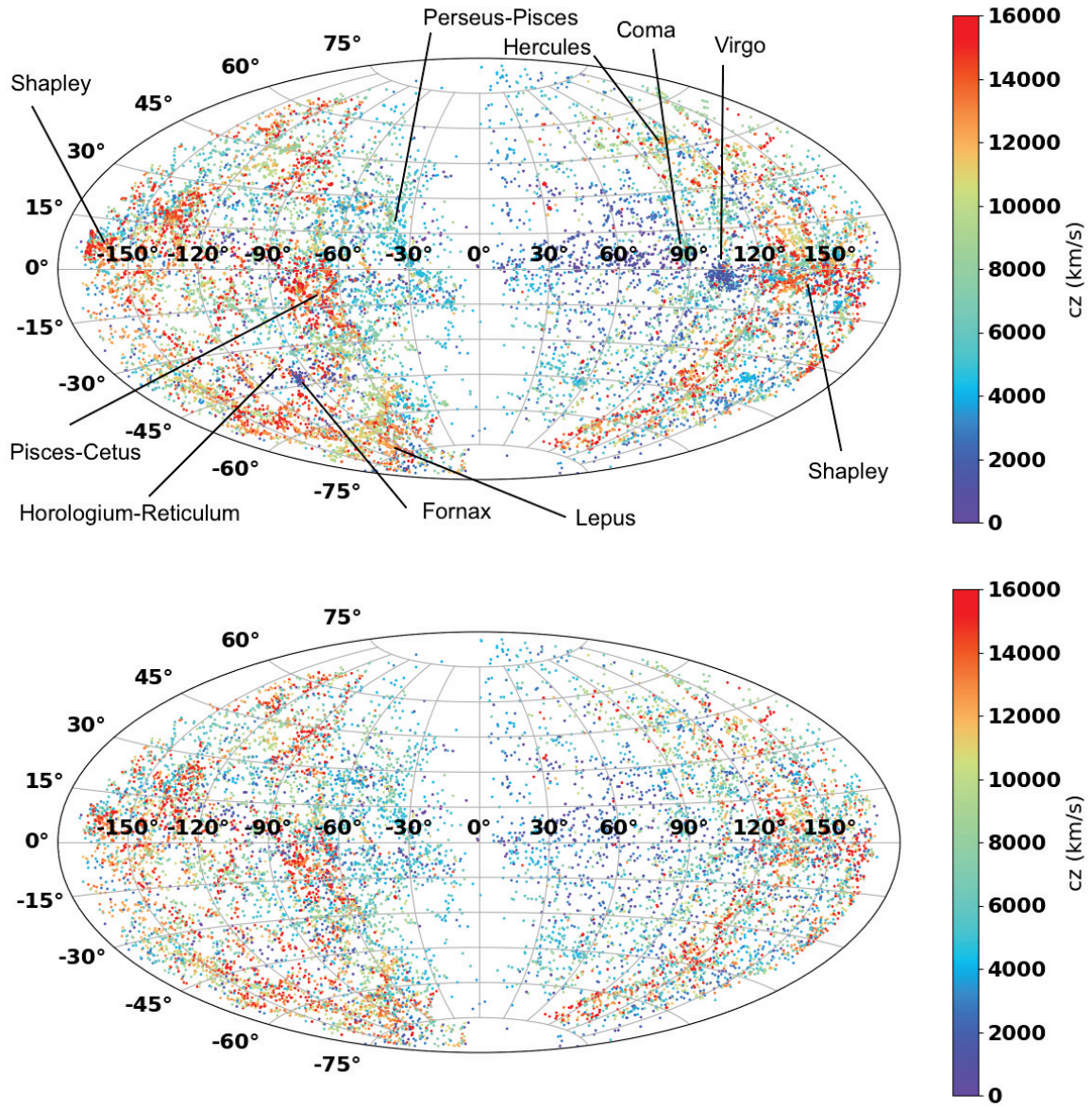


FIGURE 2.1 – Distribution des galaxies (haut) et des groupes de galaxies (bas) du catalogue *Cosmicflows-3* en coordonnées supergalactiques. La couleur correspond à l'éloignement des galaxies (ou des groupes) en redshift. Certains superamas sont identifiés parmi les galaxies, tels que Virgo, Fornax, Persée-Poisson et Shapley, etc.

La figure 2.2 gauche montre une tranche SGX-SGY centrée en $SGZ = 0$ km/s et de largeur 3,000 km/s. Toutes les galaxies sont positionnées à leur redshift. Les galaxies provenant de CF1 sont représentées en vert, CF2 en bleu et CF3 en rouge. Les cercles représentent la limite approximative des données, les différents rayons sont 3,000 km/s, 8,000 km/s et 16,000 km/s pour CF1, CF2 et CF3 respectivement. Par ailleurs, la figure 2.2 droite présente les histogrammes des distances issues de CF1, CF2 et CF3 en vert, bleu et rouge respectivement. On remarque que CF1 s'étend jusqu'à 30 Mpc, CF2 jusqu'à environ 100 Mpc et CF3 jusqu'à 150 Mpc.

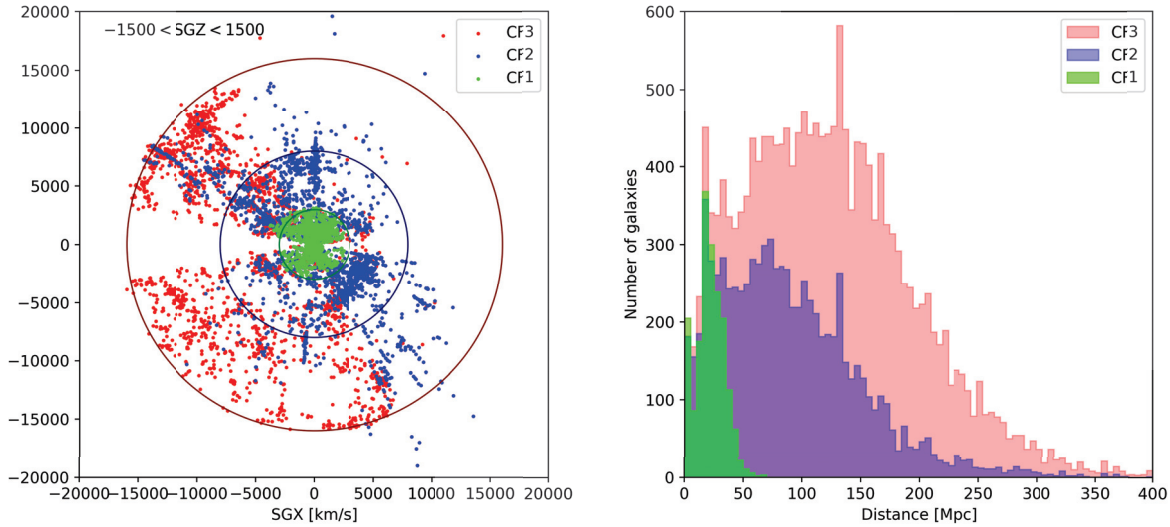


FIGURE 2.2 – Évolution des catalogues *Cosmicflows*. À gauche est représentée une tranche SGX-SGY centrée en $SGZ = 0$ km/s et de largeur 3,000 km/s. Toutes les galaxies sont positionnées à leur redshift. Les galaxies provenant de CF1 sont en vert, CF2 en bleu et CF3 en rouge. Les cercles représentent la limite approximative des données en redshift. Les différents rayons sont 3,000 km/s, 8,000 km/s et 16,000 km/s pour CF1, CF2 et CF3 respectivement. La figure de droite présente les histogrammes des distances issues de CF1, CF2 et CF3 en vert, bleu et rouge respectivement.

La figure 2.2 permet ainsi de visualiser l'évolution sur 10 ans des catalogues de distance *Cosmicflows*, allant de plus en plus loin et couvrant ainsi petit à petit un plus grand volume de notre univers local. En effet, CF1 a permis d'analyser le groupe local, tandis que CF2 a permis d'identifier de nouvelles structures telles que Coma, Persée-Poisson, Laniakea et légèrement Shapley. De plus, la collaboration *Cosmicflows* a pu étudier jusqu'aux orientations du champ de vitesse des galaxies naines (Libeskind et al., 2015). Enfin, le catalogue CF3 apporte maintenant davantage de détails sur le superamas Shapley, ainsi que de nouvelles structures dans l'hémisphère Sud qui seront identifiées dans le chapitre 4. Malheureusement, il n'y a rien de nouveau dans l'hémisphère Nord en raison d'une asymétrie entre les hémisphères terrestres Nord et Sud qui sera détaillée ci-dessous.

La table 2.2 présente les superamas qui ont été identifiés dans les catalogues *Cosmicflows*. Certains d'entre eux seront évoqués plus tard dans le chapitre 4. Les noms des amas sont donnés dans la colonne (1), et leur redshift est donné dans la colonne (2). Les coordonnées équatoriales : l'ascension droite RA et la déclinaison δ ; des amas sont listées dans les

colonnes (3) et (4) en heures et en degrés respectivement. Les coordonnées galactiques (longitude l et latitude b) sont données dans les colonnes (5) et (6), en degrés. Enfin, les coordonnées supergalactiques SGX,SGY et SGZ sont listées dans les colonnes (7), (8) et (9) respectivement et sont exprimées en Mpc/ h .

TABLE 2.2 – Superamas identifiés dans les catalogues *Cosmicflows*. Les noms des amas sont donnés dans la colonne (1), et leur redshift est donné dans la colonne (2). Les coordonnées équatoriales : l’ascension droite RA et la déclinaison δ ; des amas sont listées dans les colonnes (3) et (4). Les coordonnées galactiques (longitude l et latitude b) sont données dans les colonnes (5) et (6). Les coordonnées supergalactiques SGX,SGY et SGZ sont listées dans les colonnes (7), (8) et (9) respectivement.

	cz	RA	δ	l	b	SGL	SGB	SGX	SGY	SGZ
	km/s	h	deg	deg	deg	deg	deg	Mpc/ h	Mpc/ h	Mpc/ h
Virgo	1079	13	12	284	74	103	-2	-2	11	-0
Centaurus	3418	13	-41	302	22	156	-11	-31	13	-7
Hydra	3777	11	-28	270	26	139	-38	-23	19	-23
Perseus-Pisces	4900	3	42	143	-17	343	-7	47	-14	-6
Pavo-Indus	5180	22	-6	52	-43	274	41	3	-39	34
Coma	6895	12	24	239	83	91	-0	-2	69	-1
Phoenix	8095	1	-48	301	-70	250	-12	-27	-75	-16
Hercules	10000	16	18	30	49	106	45	-19	68	71
Leo	9615	11	20	224	64	88	-19	2	91	-32
Columba	10020	6	-36	243	-22	228	-68	-25	-28	-93
Shapley	12891	13	-33	306	30	149	-6	-110	66	-14
Norma	4707	16	-61	325	-7	188	7	-46	-7	6
Southern-Wall	5000	1	10	129	-53	307	0	30	-40	0
Ursa-Major	19007	10	56	155	51	53	-6	114	151	-20
Pisces	17988	1	8	131	-55	305	-2	103	-147	-5
Great-Arch	20125	17	26	47	33	90	63	0	91	179
Arch-knot	6083	19	12	47	-3	270	81	-0	-10	60
Great-Wall	9165	14	18	15	66	104	26	-20	80	40
Lepus	14142	6	-28	233	-24	252	-71	-14	-43	-134
Vela	18000	9	-50	273	0	174	-45	-128	14	-126
Antlia	16031	10	-36	273	18	151	-39	-109	60	-101
Aquarius	17388	23	-20	44	-65	268	19	-4	-165	55
Pisces-Cetus	16279	1	-20	96	-83	276	-1	16	-162	-4
Great-Attractor	4472	11	-46	279	10	162	-36	-35	11	-26
Apus	17000	17	-74	319	-20	200	-1	-160	-58	-3
Horologium-Reticulum	20000	20	-25	15	-22	223	48	-99	-92	148
Telescopium	20000	19	-44	354	-24	212	30	-148	-92	99
Sculptor Wall	9000	0	-25	47	-84	270	0	-0	-90	0
Ophiuchus	10000	17	-23	0	9	173	43	-72	9	69

Vitesses particulières

Afin de calculer les vitesses particulières à partir des mesures de distance, il faut choisir une valeur de H_0 pour ensuite utiliser les équations 2.6, 2.7, 2.8 ou 2.9 permettant de calculer la partie radiale des vitesses particulières. Le choix de la valeur de ce paramètre est discuté ci-après.

Tully et al. (2016) montre que la valeur de la constante de Hubble la plus adéquate pour estimer la composante radiale des vitesses particulières des galaxies est $H_0 = 75 \text{ km s}^{-1} \text{ Mpc}^{-1}$. En effet, cette valeur permet de minimiser le monopole à partir des distances provenant de CF3, et d'obtenir un flot entrant ou sortant radialement pas trop important, comme le montre la figure 2.3, extraite de Tully et al. (2016). Les points avec barres d'erreurs correspondent à la moyenne des vitesses particulières, déduites des distances CF3, distribuées dans des bins de redshift cz . Les galaxies trop locales pour lesquelles $cz < 3,000 \text{ km/s}$ ne sont pas prises en compte. Les vitesses particulières ont été calculées pour différentes valeurs de H_0 , de 68 km/s/Mpc à 77 km/s/Mpc . Les bins de vitesse obtenus pour la même valeur de H_0 sont reliés par des lignes en trait plein ou pointillés. La ligne horizontale en pointillé noir signale un flot entrant ou sortant nul. La zone au dessus de cette ligne indique un flot sortant plus important, tandis que la zone en dessous révèle un flot entrant considérable. Une valeur de H_0 trop grande résulte en moyenne de vitesses particulières trop faibles, tandis qu'une valeur de H_0 trop petite donne l'inverse. La valeur de H_0 idéale associée aux distances CF3 est ainsi 75 km/s/Mpc . Ce choix est représenté en rouge sur la figure 2.3.

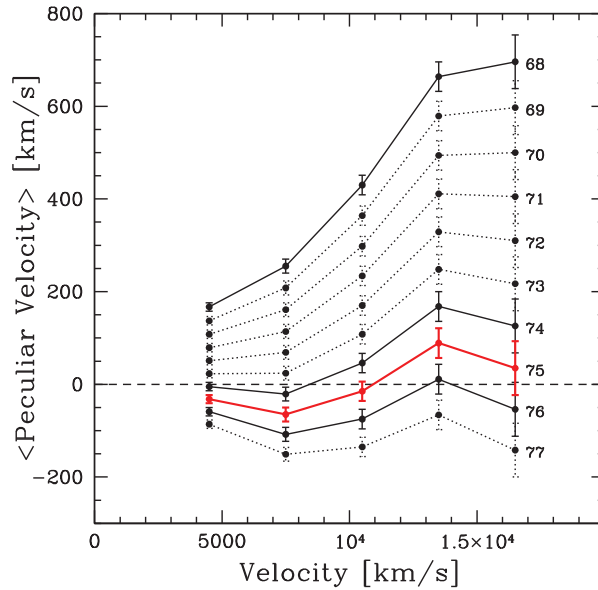


FIGURE 2.3 – Vitesses particulières radiales selon la valeur de H_0 . Les vitesses particulières obtenues pour différentes valeurs de H_0 (de 68 à 77) sont moyennées dans des bins de cz . Figure extraite de Tully et al. (2016).

Pour ces raisons, la valeur $H_0 = 75 \text{ km s}^{-1} \text{ Mpc}^{-1}$ sera utilisée tout au long de cette thèse avec les distances provenant de *Cosmicflows-3*. On écrira donc Mpc/h_{75} lorsque cette valeur sera utilisée. De plus, l'équation 2.9 sera utilisée pour estimer les vitesses particulières à partir des distances.

La figure 2.4 présente deux tranches en coordonnées supergalactiques cartésiennes SGX-SGY, centrées en $SGZ = 0$ km/s et de largeur 3,000 km/s. Les points correspondent aux galaxies (gauche) et aux groupes (droite) de CF3, positionnés à leur redshift. Les flèches bleues et rouges représentent la composante radiale des vitesses particulières, se dirigeant vers l'observateur et s'éloignant de l'observateur respectivement. On peut déjà observer sur cette figure les zones dans lesquelles les grandes structures sont localisées. En effet, un attracteur est nettement visible lorsqu'un flot gravitationnel montre une chute avant et une chute arrière. Le super-amas de Persée-Poisson en $X, Y = (3,000, -3,000)$ km/s est un excellent exemple. La figure 2.4 permet aussi de mettre en évidence une asymétrie entre les hémisphères terrestres Nord et Sud. En effet, seulement 30 % (5372 galaxies) des distances de galaxies fournies par CF3 sont dans le Nord, contre 70 % (12,276 galaxies) dans le Sud. En ce qui concerne les groupes, 3,394 distances d'ensembles de galaxies sont dans le Nord contre 8,542 dans le Sud.

La figure 2.5 montre les histogrammes des vitesses particulières des catalogues CF3 non groupé (en rouge) et groupé (en bleu). On peut observer que les deux histogrammes sont centrés en 0. Par ailleurs, on remarque que les queues de distribution tendent vers des valeurs aberrantes pour le régime linéaire ($> 4,000$ km/s). Ceci est dû aux grandes erreurs de mesure sur les distances.

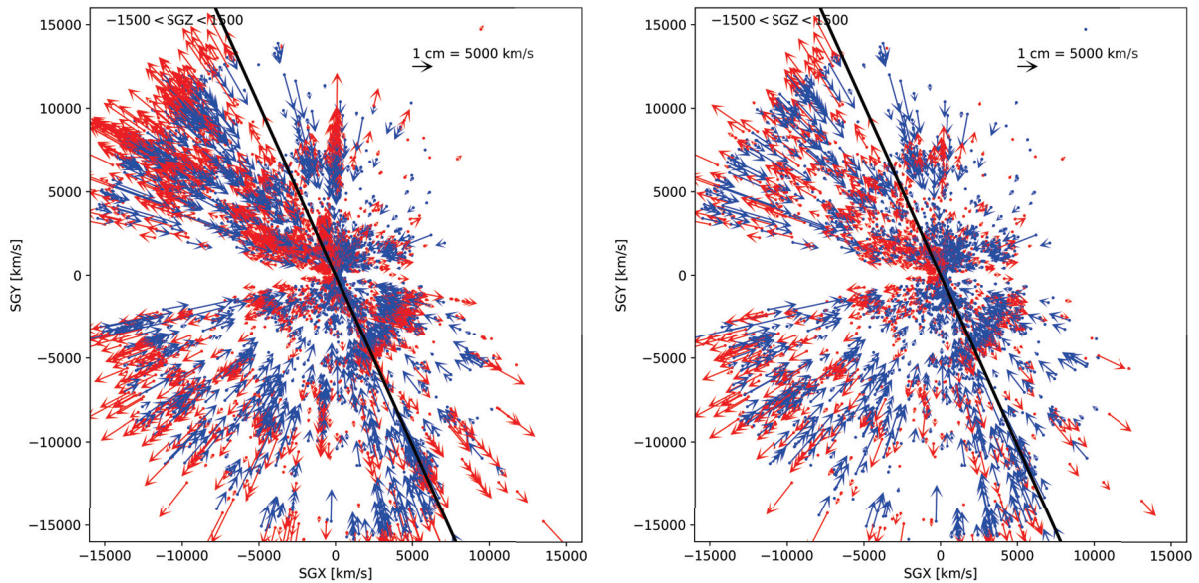


FIGURE 2.4 – Tranches en coordonnées supergalactiques cartésiennes SGX-SGY, centrées en $SGZ = 0$ km/s et de largeur 3,000 km/s. Les points correspondent aux galaxies (gauche) et aux groupes (droite) de CF3, positionnés à leur redshift. Les flèches bleues et rouges représentent la composante radiale des vitesses particulières, se dirigeant vers l'observateur et s'éloignant de l'observateur respectivement. Les zones de convergence chute avant - chute arrière mettent en évidence la présence d'un attracteur. Voir par exemple le super-amas de Persée-Poisson en $X, Y = (3,000, -3,000)$ km/s.

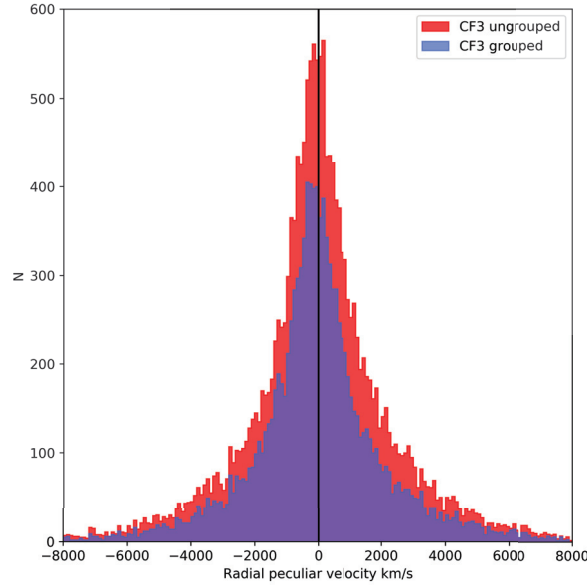


FIGURE 2.5 — Histogrammes des vitesses particulières des catalogues CF3 non groupé (en rouge) et groupé (en bleu). On peut observer que les deux histogrammes sont centrés en 0. Cependant, on remarque que les queues de distribution tendent vers des valeurs aberrantes pour le régime linéaire ($> 4,000$ km/s). Ceci est dû aux grandes erreurs de mesure sur les distances.

2.2.3 *Altius* — Cosmicflows-4 : encore plus haut ... vers le Nord

Motivations

Notre galaxie la Voie Lactée se déplace à une vitesse de 630 km/s par rapport au CMB. Les raisons sont multiples. D'une part, le superamas de Shapley et le Grand Attracteur jouent le rôle d'attracteur et attirent notre galaxie dans leur direction. D'autre part, les données CF2 ont permis de mettre en évidence une région vide avec très peu de galaxies dans la direction opposée à Shapley. Cette région a été nommée le Repousseur du Dipole (*Dipole Repeller* — DR, Hoffman et al., 2017) et est située dans l'hémisphère Nord (qui est une région dans laquelle les données manquent et où des mesures provenant des radiotélescopes de Green Bank et Arecibo vont être rajoutées). Les effets à la fois d'évacuation de la part du DR et d'attraction de la part de Shapley et du Grand Attracteur sont comparables à l'amplitude de la vitesse de la Voie Lactée. Par ailleurs, CF3 a permis de mettre en évidence l'existence d'une autre grande structure dans la direction de la région la plus froide du CMB : le Repoussoir du Point Froid (*Cold Spot Repeller* — CSR, Courtois et al., 2017), qui reste cependant à confirmer. Par conséquent, l'existence d'attracteurs et de repousseurs semble ainsi importante pour expliquer le mouvement de notre galaxie dans l'univers. Il est cependant impossible de confirmer l'existence de ces structures avec les données actuelles, car ces régions sont beaucoup trop lointaines.

Le prochain catalogue du projet Cosmicflows, *Cosmicflows-4* (CF4) a pour but d'atteindre ces régions encore peu connues, afin de mener une étude encore plus approfondie.

CF4 — Nord

CF4 donnera accès à environ 25,000 distances, soit 7,000 distances de plus par rapport à CF3, et s'étendra jusqu'à environ 10,000 km/s dans le Nord. En effet, le but de CF4 est de corriger l'asymétrie Nord/Sud de CF3. Par conséquent, les nouvelles distances de galaxies proviendront d'observations menées aux radiotélescopes situés dans l'hémisphère Nord. De plus, ces distances seront obtenues par la relation de Tully-Fisher. La partie radio (largeur de raie HI) sera soit tirée de la littérature, les observations ayant été effectuées par d'autres équipes (données ALFALFA à Arecibo, données KLUN17 et NIBLES à Nancay), soit obtenue à partir d'observations effectuées par la collaboration *Cosmicflows* au radiotélescope de Green Bank. Cette partie sera discutée en détail dans le chapitre 3. La photométrie sera extraite des résultats PanSTARRS, SDSS et WISE.

La figure 2.6 présente une visualisation à trois dimensions de la forme prévisionnelle de CF4. Les galaxies CF3 sont en gris, tandis que les galaxies complémentaires dans le Nord sont en couleur : données Nancay (NIBLES et KLUN17) en jaune, données provenant de la collaboration ALFALFA (Arecibo) en bleu et galaxies qui seront potentiellement observées à Green Bank en violet. On peut remarquer sur cette figure que ces nouvelles galaxies permettront de corriger en partie l'asymétrie Nord / SUD de CF3, tout en allant plus loin pour atteindre le DR dans le Nord. Cependant, ces structures sont encore trop lointaines pour les radiotélescopes situés dans l'hémisphères Nord, leurs données ne permettant pas d'atteindre le DR suffisamment.

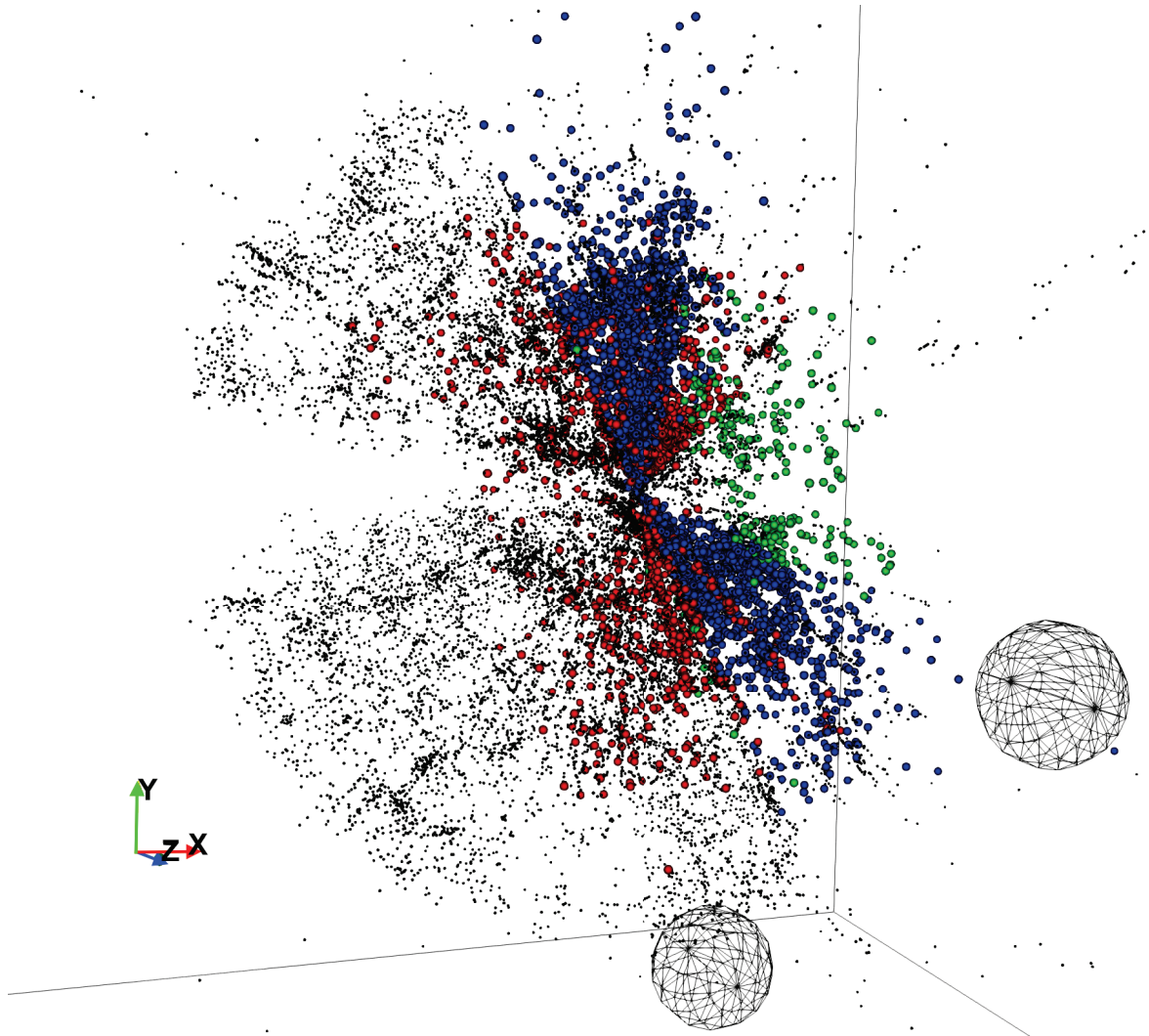


FIGURE 2.6 – Visualisation à trois dimensions de la forme prévisionnelle du prochain catalogue *Cosmicflows-4*. Les galaxies CF3 sont représentées par de petits points noirs, tandis que les nouvelles galaxies qui seront ajoutées à CF3 sont illustrées par des sphères de couleur : les données Nancay (provenant des sondages HI NIBLES et KLUN17) en rouge, données provenant de la collaboration ALFALFA en bleu et galaxies détectées au radiotélescope de Green Bank en vert. Les deux sphères en *fil de fer* correspondent aux deux repousseurs : le Repousseur du Dipole (DR) et le Cold Spot Repeller (CSR).

Chapitre 3

Vers Cosmicflows-4 : observations HI

Le chapitre 2 a montré que le catalogue actuel de distances de galaxies *Cosmicflows-3* est très asymétrique : 70% des galaxies sont dans l'hémisphère terrestre Sud contre seulement 30% dans le Nord. L'objectif du futur catalogue *Cosmicflows-4* est de corriger cette asymétrie en ajoutant de nouvelles galaxies dans l'hémisphère Nord obtenues par la relation de Tully-Fisher. Cette méthode de mesure de distances nécessite à la fois les données de télescopes optiques (photométrie) et radio (spectres HI). Les chapitres 4 et 5 expliqueront comment les distances des galaxies peuvent être utilisées pour développer de nouvelles méthodes permettant de caractériser et d'identifier les grandes structures de l'univers local et mesurer leur compacité.

Ce chapitre introduit tout d'abord brièvement la relation de Tully-Fisher. Puis sont décrites les deux types de données nécessaires pour mesurer une distance : tout d'abord la photométrie, puis les observations à 21 cm pour mesurer la largeur de raie des spectres HI des galaxies, qui constituent l'essentiel de ce chapitre. Les parties suivantes présentent l'analyse de divers échantillons HI par différentes méthodes, selon le type de données disponibles, et qui vont être ajoutées au catalogue *Cosmicflows-4*.

Sommaire

3.1	La relation de Tully-Fisher	52
3.1.1	Radioastronomie : observations à 21 cm	54
3.2	Données provenant du radiotélescope Nançay	61
3.2.1	Kinematic in the Local UNiverse	62
3.2.2	Nançay Interstellar Baryons Legacy Extragalactic Survey	67
3.3	Données provenant du radiotélescope Arecibo	72
3.3.1	Arecibo Legacy Fast ALFA : description	72
3.3.2	Intégration des données ALFALFA à Cosmicflows-4	73
3.4	Projet GBT18A-021 : Cosmic flows in the Local Universe	80
3.4.1	Sélection des sources HI	80
3.4.2	Observations à 21 cm au GBT	85
3.4.3	Observations au GBT : réduction des données	90
3.4.4	Premiers résultats	94
3.5	Conclusion : données HI pour Cosmicflows-4	97

3.1 La relation de Tully-Fisher

L'étude des courbes de rotation des galaxies spirales ayant une distance connue a permis de mettre en évidence une relation empirique entre la luminosité L de ces galaxies et leur largeur de raie à 21 cm de l'hydrogène neutre W qui correspond à deux fois leur vitesse de rotation. En effet, les galaxies plus lumineuses, contenant plus de gaz et d'étoiles, sont donc plus massives et ont par conséquent une vitesse de rotation plus élevée pour maintenir un état d'équilibre. C'est la relation de Tully-Fisher (TF, [Tully and Fisher, 1977](#)) :

$$L = k \left(\frac{W}{2 \sin(i)} \right)^\alpha, \quad (3.1)$$

où k et α sont des constantes et i l'inclinaison de la galaxie. On peut aussi écrire cette équation sous la forme :

$$M = a + b \log_{10} V_{\max}, \quad (3.2)$$

où M est la magnitude absolue et V_{\max} la vitesse de rotation maximale d'une galaxie. Les paramètres a et b sont des constantes.

Pour déterminer la distance d'une galaxie à l'aide de la relation TF, deux types d'observations conduites sur deux différents types de télescopes sont nécessaires :

- avec un télescope optique : photométrie de surface pour mesurer la magnitude apparente m ;
- avec un radiotélescope : mesure de la largeur de raie HI W pour estimer la vitesse de rotation V_{\max} puis la magnitude absolue M .

Les observations donnent donc accès à m et M , et il est ainsi possible de déduire la distance d :

$$m - M = \mu = 5 \log_{10} (d) + 25. \quad (3.3)$$

Cette relation est tout d'abord calibrée par des galaxies proches dont la distance est connue (obtenue par exemple par la relation période-luminosité des Céphéides). Elle est ensuite appliquée aux galaxies dont on veut mesurer la distance.

Photométrie

La photométrie permet de mesurer la magnitude apparente des galaxies. Elle peut se faire dans différentes bandes spectrales, listées dans la table 3.1 de l’ultraviolet jusqu’à l’infrarouge moyen.

Un grand nombre de données observationnelles sont accessibles dans les bandes optiques jusqu’au proche infrarouge : *Sloan Digital Sky Survey* (SDSS, bandes U, G, R, I, Z , Abazajian et al., 2003), *Two Micron All Sky Survey* (2MASS, bandes J, H, K , Huchra et al., 2012).

Par ailleurs, les observations menées au télescope spatial Spitzer (Werner et al., 2004; Fazio et al., 2004) par la collaboration *Cosmicflows* ont apporté une grande contribution à la préparation du catalogue *Cosmicflows-3* (Sorce et al., 2012, 2014; Tully et al., 2016). Les observations en infrarouge moyen conduites sur des télescopes terrestres sont difficiles et demandent un temps de pose important car contaminées par l’atmosphère de la Terre, mais les observations conduites depuis l’espace règlent ce problème et les magnitudes totales apparentes des galaxies peuvent être obtenues en quelques minutes seulement.

Pour la construction des prochains catalogues *Cosmicflows-4* (ou plus), les données photométriques proviendront de la mission *Wide-field Infrared Survey Explorer* (WISE, Liu et al., 2008), et du sondage Pan-STARRS (bandes g, r, i, z, y Kaiser et al., 2002).

Dans ce chapitre, les images optiques des galaxies observées à 21 cm seront données dans les bandes g, r, i, z, y et proviendront du sondage Pan-STARRS.

TABLE 3.1 – Bandes spectrales photométriques. Les lettres associées aux différentes bandes sont listées dans les colonnes (1) et (4). Les longueurs d’onde médianes approximatives de chaque bande sont données dans les colonnes (2) et (5) en nm. Les colonnes (3) et (6) donnent des exemples de filtres appartenant aux différentes bandes. Table tirée de Sorce (2014).

Ultraviolet			Proche infrarouge		
U	~ 300	u	I	~ 800	I _C , i
			Z	~ 900	z
	Visible		Y	~ 1020	y
B (bleu)	~ 400		J	~ 1200	
V (violet)	~ 500		H	~ 1600	
G (vert)	~ 600	g	K	~ 2200	K _s , K'
R (rouge)	~ 700	R _C , r	Au-delà : infrarouge moyen		

3.1.1 Radioastronomie : observations à 21 cm

Raie à 21 cm de l'atome d'Hydrogène neutre

Dans l'atome d'Hydrogène neutre (HI), un électron tourne autour d'un proton. Par ailleurs, ces deux particules possèdent chacune un spin qui peut être positif ou négatif. Il existe donc deux états : dans l'un, les spins des deux électrons sont parallèles, et dans l'autre les spins sont antiparallèles. Lors du passage de l'état parallèle (d'énergie E_1) à l'état antiparallèle (d'énergie E_2), on observe un rayonnement d'énergie sous forme d'un photon de fréquence $\nu_0 = 1420$ MHz, donnée par la relation de Planck $E_1 - E_2 = h\nu_0$, ou h est la constante de Planck. Cette fréquence est équivalente à une longueur d'onde $\lambda = 21$ cm.

La raie spectrale à 21 cm est une raie interdite, c'est-à-dire que la transition associée n'est normalement pas autorisée selon les règles de la mécanique quantique, et elle est peu probable. Il faut en effet 11 millions d'années pour qu'un atome d'hydrogène subisse cette transition. Cependant, les galaxies contiennent une quantité importante d'hydrogène neutre, c'est pourquoi il est possible d'observer ces transitions. Par exemple, notre galaxie, la Voie Lactée, contient environ 10^{67} atomes d'hydrogène, et 10^{52} atomes par seconde émettent un photon de fréquence 1420 MHz. L'hydrogène joue un rôle dans la formation des étoiles dans les galaxies. En présence de gaz d'hydrogène, une galaxie peut former des étoiles. Or, si le gaz présent dans la galaxie a déjà été consommé en formation stellaire, il ne reste alors que des étoiles vieilles.

À cause de la vitesse radiale d'une galaxie v (somme de la vitesse due à l'expansion et de la composante radiale de la vitesse particulière), la raie spectrale à 21 cm est décalée vers une fréquence ν (ou longueur d'onde λ) par effet Doppler. Il existe deux conventions, optique et radio, pour déterminer la vitesse v à partir de la fréquence ou longueur d'onde :

$$\text{Convention optique} \quad v_{\text{opt}} = cz = c \frac{\lambda - \lambda_0}{\lambda_0} = c \frac{\nu_0 - \nu}{\nu}; \quad (3.4)$$

$$\text{Convention radio} \quad v_{\text{radio}} = c \frac{\nu_0 - \nu}{\nu_0} = c \frac{\lambda - \lambda_0}{\lambda} \neq cz. \quad (3.5)$$

Spectres à 21 cm des galaxies spirales

Le spectre HI d'une galaxie spirale inclinée (vue sur la tranche), a une forme typique : une raie divisée en deux pics. Ceci est dû à la rotation de la galaxie (et donc du nuage d'hydrogène) sur elle-même. D'un côté de la galaxie, une partie du nuage de gaz se rapproche de l'observateur, tandis que du côté opposé, une partie du nuage de gaz s'éloigne de l'observateur. Par effet Doppler, la raie est donc élargie, et deux pics sont visibles. Dans le cas d'une galaxie vue de face, il n'y a aucun mouvement radial par rapport à l'observateur, le spectre HI est donc constitué d'un seul pic.

Dans la collaboration *Cosmicflows*, afin de mesurer les largeurs de raies HI, les spectres HI sont inspectés un par un visuellement et la détection de la raie HI se fait à l'aide d'un logiciel IDL *maison*.

Pourquoi la mesure de la largeur de raie n'est-elle pas automatique ? Les spectres sont

tous différents. Il existe beaucoup de cas particuliers et de cas difficiles. Un cas idéal, dont la mesure de largeur de raie est quasi-automatique, est un spectre avec un bon signal sur bruit. Il existe cependant des cas plus difficiles avec un faible signal sur bruit, où il est alors très difficile de détecter la raie de façon totalement automatique. On a donc besoin d'inspecter tous les spectres un par un.

Mesure de la largeur de raie à 21 cm sur un spectre HI

La mesure de la largeur de la raie dans la collaboration Cosmicflows se fait de la façon suivante, tirée de [Courtois et al. \(2009\)](#).

Tout d'abord, les bords de la raie sont identifiés : la où elle commence (0% de la raie) et la où elle se termine (100% de la raie). Ils sont facilement identifiables dans le spectre intégré (i.e la somme accumulée du flux). Afin de minimiser les erreurs dues aux bords des raies parfois mal définis (souvent à cause du bruit), les premiers et derniers 5% sont ignorés pour la suite : on considère ainsi seulement l'intervalle [5%, 95%] du flux intégré.

Si le logiciel ne trouve pas les limites de la raie HI automatiquement (signal sur bruit trop faible), l'utilisateur peut re-définir de nouvelles limites [0%, 100%] en cliquant sur le spectre.

La figure 3.1 gauche présente le spectre intégré de la galaxie PGC 156999 extrait des données KLUN17 décrites dans la section 3.2.1. Les traits verticaux rouges correspondent aux limites de l'intervalle [0%, 100%] du flux intégré, définies automatiquement par l'algorithme ou redéfinies par l'utilisateur. Les traits verticaux bleus correspondent aux limites de l'intervalle [5%, 95%] du flux intégré. On remarque ainsi que l'intervalle [5%, 95%] définit mieux la raie HI que l'intervalle [0%, 100%].

Après avoir repéré les bords de la raie HI (0-100% et 5-95%), la largeur de la raie W_{50} est mesurée. Le paramètre W_{50} correspond à la largeur de raie contenant 50% du flux intégré dans l'intervalle [5%, 95%].

La figure 3.1 droite montre le spectre final de la même galaxie PGC 156999 après la mesure de la largeur de raie HI. Les traits verticaux bleus correspondent à nouveau aux limites de l'intervalle [5%, 95%] du flux intégré. Le trait vert horizontal représente la hauteur à 50% du flux intégré pour la mesure de W_{50} . En effet, le paramètre W_{50} est obtenu à partir de l'intersection entre les traits verticaux bleus et le trait horizontal vert.

Il existe cependant des cas difficiles, par exemples lorsque les bords de la raie HI sont mal définis. Le spectre HI peut alors être lissé afin de supprimer le bruit pour que les bords de la raie soient plus facilement identifiables. Cela a cependant un inconvénient : le lissage augmente la résolution spectrale et élargit la raie, et l'erreur de mesure sur W_{50} augmente elle aussi.

Corrections des effets instrumentaux et cosmologiques

Le paramètre W_{50} mesuré sur les données observationnelles doit être corrigé de différents effets instrumentaux ou cosmologiques. Pour corriger l'élargissement dû à la résolution

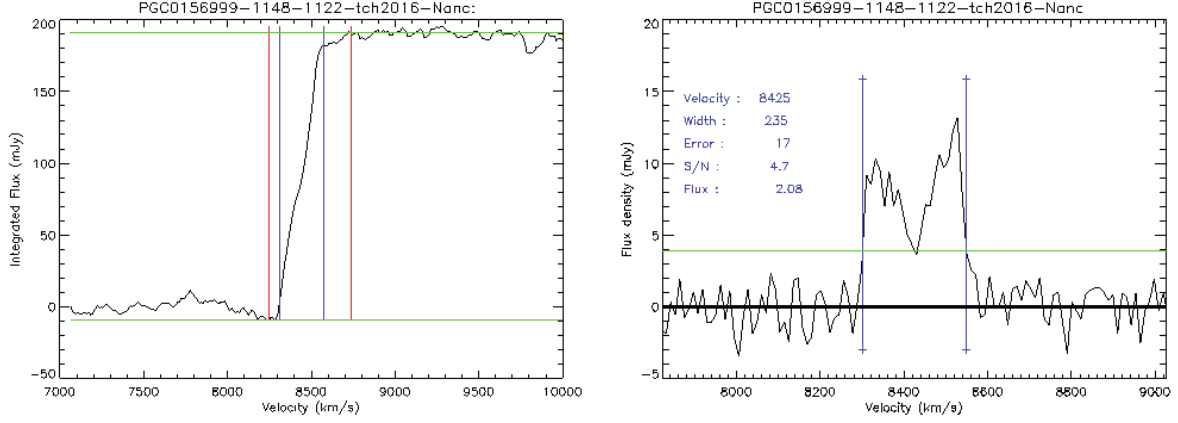


FIGURE 3.1 – Spectre à 21 cm de la galaxie *PGC 156999*, extrait du sondage HI KLUN17 détaillé dans la partie suivante. La figure de gauche correspond au spectre intégré sur lequel sont détectés les bords de la raie HI. Les traits verticaux rouges correspondent aux limites de l'intervalle [0%, 100%] du flux intégré, et les traits verticaux bleus correspondent aux limites de l'intervalle [5%, 95%] du flux intégré. La figure de droite représente le même spectre à 21 cm de la galaxie. Les traits verticaux bleus correspondent à nouveau aux limites de l'intervalle [5%, 95%] du flux intégré. Le trait vert horizontal représente la hauteur à 50% du flux intégré pour la mesure de W_{50} .

spectrale et au redshift, la largeur de raie W_{50} mesurée sur le profil HI est ajustée avec l'équation suivante :

$$W_{50}^c = \frac{W_{50}}{1+z} - 2\Delta v \lambda \quad (3.6)$$

où Δv est la résolution spectrale (après un éventuel smoothing) et $\lambda = 0.25$ une constante dont la valeur est déterminée empiriquement à partir de tests sur différents profils HI (Courtois et al., 2009).

On peut aussi ajuster le paramètre W_{50} pour qu'il ait une signification physique plus importante, i.e pour qu'il corresponde à deux fois la vitesse de rotation V_{\max} . Tully and Fouque (1985) ont introduit une équation permettant d'obtenir une telle quantité :

$$W_{mx}^2 = W_{50}^2 + W_{t50}^2 \left(1 - 2e^{-\left(\frac{W_{50}}{W_{c50}}\right)^2} \right) - 2W_{50}W_{t50} \left(1 - e^{-\left(\frac{W_{50}}{W_{c50}}\right)^2} \right), \quad (3.7)$$

avec $W_{c50} = 100$ km/s et $W_{t50} = 9$ km/s. Cette équation corrige le paramètre W_{50} des effets de turbulence et de dispersion des vitesses internes des galaxies spirales, causant un élargissement de la raie.

Attention à l'inclinaison !

Cependant, le paramètre W_{mx} d'un spectre HI d'une galaxie spirale définit dans l'équation 3.7 n'est toujours pas équivalent à sa vitesse de rotation V_{\max} . Il faut considérer un autre paramètre afin de déterminer de façon correcte V_{\max} : l'inclinaison de la galaxie. En effet, le disque de gaz d'hydrogène d'une galaxie étant en rotation, l'amplitude de rotation observée dépend de l'orientation du disque par rapport à l'observateur, i.e son inclinaison. Par conséquent, après dé-projection de l'inclinaison de disque i , on peut estimer la vitesse de rotation d'une galaxie à partir de la largeur de son profil HI corrigée W_{mx} :

$$2V_{\max} = \frac{W_{mx}}{\sin(i)}. \quad (3.8)$$

L'inclinaison i d'une galaxie par rapport à la ligne de visée d'un observateur est donnée par (Holmberg, 1958) :

$$\cos(i) = \sqrt{\frac{\left(\frac{b}{a}\right)^2 - q_0^2}{1 - q_0^2}}, \quad (3.9)$$

où $\frac{b}{a}$ est le ratio petit axe - grand axe de la galaxie, et q_0 son *épaisseur*, i.e le ratio petit axe - grand axe de la galaxie si vue sur la tranche. Donc par définition, l'inclinaison d'une galaxie vue parfaitement de face est de 0° , et l'inclinaison d'une galaxie vue sur la tranche est de 90° .

La valeur de référence pour le paramètre q_0 est de 0.20 dans la collaboration *Cosmicflows*. Des publications proposent d'autres valeurs sur ce paramètre dépendant du type morphologique des galaxies. Giovanelli et al. (1997) propose $q_0 = 0.13$ pour les galaxies de type Sc. Cependant, et heureusement, le choix de q_0 a un effet négligeable sur les mesures de distances. Par exemple, pour $b/a = 0.20$, on a :

$$\text{si } q_0 = 0.20 \quad \text{alors } i = 90^\circ \text{ et } 1/\sin(i) = 1, \quad (3.10)$$

$$\text{si } q_0 = 0.13 \quad \text{alors } i = 81^\circ \text{ et } 1/\sin(i) = 1.012. \quad (3.11)$$

Cela résulte en une différence 1.2% pour le facteur $1/\sin(i)$ présent dans l'équation 3.8 pour le calcul de $2V_{\max}$. Pour un ratio petit axe-grand axe plus grand tel que $b/a = 0.90$:

$$\text{si } q_0 = 0.20 \quad \text{alors } i = 26.42^\circ \text{ et } 1/\sin(i) = 2.275, \quad (3.12)$$

$$\text{si } q_0 = 0.13 \quad \text{alors } i = 26.08^\circ \text{ et } 1/\sin(i) = 2.248, \quad (3.13)$$

ce qui donne dans ce cas une différence de 2.7% pour $1/\sin(i)$. L'incertitude sur la correction apportée au paramètre W_{mx} pour qu'il soit équivalent à deux fois la vitesse de rotation devient alors plus importante pour les galaxies vues de face. Pour cette raison,

seules les galaxies pour lesquelles $i > 45^\circ$ sont considérées pour l'estimation des distances par la relation de Tully-Fisher.

L'incertitude sur l'inclinaison des galaxies est une part importante de l'erreur sur leur distance, et une mesure précise de i est souvent difficile à obtenir. La figure 3.2 (a) compare une image optique (droite) et une carte HI (gauche) de la même galaxie *NGC6946*, provenant du Digitized Sky Survey, extraites de Boomsma et al. (2008). Pour faciliter la comparaison, les deux images sont à la même échelle. Cette figure permet de constater que le disque de gaz de la galaxie, mis en évidence à gauche, est beaucoup plus étendu que son disque stellaire.

Sur les planches (b) et (c) de la figure 3.2 sont comparées les images composites radio-optique de deux galaxies, *NGC5055* et *M33* respectivement, extraites de Wijnholds et al. (2010) et <http://images.nrao.edu/398>. Le disque d'hydrogène neutre est superposé sur le disque stellaire (toujours la même échelle) en bleu sur la planche (b) et violet sur la planche (c). Les disques de gaz sont à nouveau plus étendu que le disque stellaire. De plus, l'inclinaison des deux disques semble différer, notamment sur la planche (b). Il faut se rappeler que la largeur de raie W_{mx} doit être dé-projetée de l'inclinaison du disque de gaz, et non du disque stellaire. Or, les inclinaisons sont la plupart du temps estimées à partir d'images optiques, et donc sur les disques stellaires, induisant probablement des erreurs qui ne sont pas actuellement prises en compte. L'obtention de disques HI en 3D par les réseaux de radiotélescopes multi-antennes comme *ASKAP* (*Australian SKA Pathfinder*) et *MeerKAT* permettra de résoudre définitivement le problème de cette erreur sur la distance due à la mauvaise estimation de l'inclinaison du disque de gaz.

Erreur sur la largeur de raie à 21 cm

L'attribution d'une erreur e_W sur la mesure de la largeur de raie W_{50} se fait selon le signal sur bruit. Les règles utilisées sont celles définies dans Courtois et al. (2009).

Le ratio signal sur bruit SNR est défini par $SNR = S/N$, où S est le signal, i.e le flux moyen dans la raie HI, et N est le bruit (*noise*), i.e flux moyen en dehors de la raie HI.

Dans le meilleur des cas, i.e quand le signal est au moins 17 fois plus grand que le bruit, une erreur e_W de 8 km/s (qui correspond à la meilleure résolution spectrale) est accordée sur la mesure de la largeur de raie. Sinon, l'erreur attribuée augmente inversement par rapport au SNR. Les règles d'assignation des erreurs e_W selon le signal sur bruit SNR sont les suivantes :

$$\text{Si } SNR \geq 17, \quad e_W = 8 \text{ km/s}; \quad (3.14)$$

$$\text{Si } 2 < SNR < 17, \quad e_W = 21.6 - 0.8 SNR \text{ km/s}; \quad (3.15)$$

$$\text{Si } SNR < 2, \quad e_W = 70 - 25 SNR \text{ km/s}. \quad (3.16)$$

L'erreur ne peut tout de même pas être inférieure à la résolution spectrale. Si c'est le cas, on remplace l'erreur trop petite par l'erreur minimale, i.e la résolution spectrale. Par ailleurs, il existe une convention pour les cas particuliers dans le catalogue ADHI de la

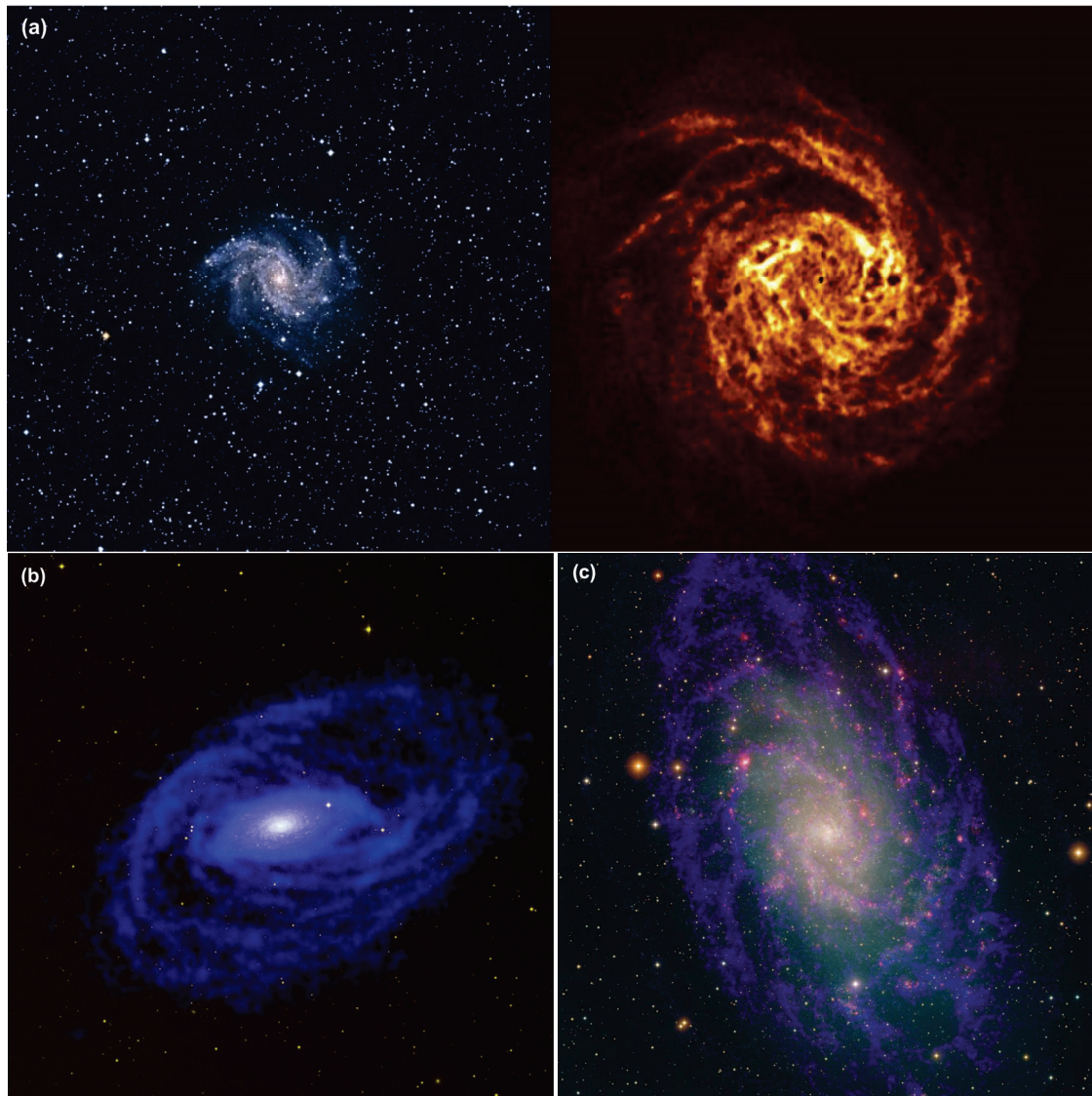


FIGURE 3.2 – Haut : image optique (gauche), provenant du Digitized Sky Survey, et une carte HI (droite) de la même galaxie *NGC6946*, extraites de [Boomsma et al. \(2008\)](#). Les deux images sont à la même échelle. Bas : images composites radio-optique de deux galaxies, *NGC5055* et *M33* respectivement, extraites de [Wijnholds et al. \(2010\)](#) et <http://images.nrao.edu/398>. Le disque d’hydrogène neutre est superposé sur le disque stellaire (toujours la même échelle) en bleu sur la planche (b) et violet sur la planche (c).

base de données EDD (*Extragalactic Distance Database*) afin de repérer les cas de spectres particuliers. Ces cas particuliers sont les suivants :

- une erreur $e_W = 100$ km/s est attribuée dans les cas de confusion (i.e plusieurs galaxies voisines en HI dans le radiolobe) ;
- une erreur $e_W = 500$ km/s est assignée aux cas très incertains ou sans détection.

Les profils HI adéquats pour estimer des distances avec la relation Tully-Fisher sont les profils ayant une erreur sur la mesure de la largeur de raie e_W inférieure à 20 km/s. Les profils dont e_W se situent au-delà de ce seuil ne sont pas considérés pour estimer des distances à intégrer aux catalogues *Cosmicflows* par TF (Tully and Pierce, 2000).

Autres paramètres

À partir de la vitesse héliocentrique V_{hel} et du flux intégré dans la raie HI exprimé en Jy.km/s que l'on notera FI, la masse HI M_{HI} est donnée par Courtois et al. (2015) :

$$M_{HI} = 2.36 \times 10^5 D^2 \text{FI}, \quad (3.17)$$

ou $D = V_{hel}/H_0$ est la distance en Mpc. Il est aussi possible d'estimer la magnitude apparente d'une galaxie à 21 cm, définie par :

$$m_{21} = -2.5 \log_{10}(\text{FI}) + 17.4. \quad (3.18)$$

Sondages à 21 cm

Il existe deux types de sondages en HI extragalactique : les sondages à l'aveugle et les sondages avec cibles pointées.

Dans les sondages à l'aveugle, le ciel est tout d'abord entièrement scanné. Il faut ensuite extraire les sources HI qui ont été détectées. Cette méthode possède à la fois des avantages et des inconvénients. En effet, cette méthode est non-biaisée, c'est-à-dire qu'elle permet de détecter des galaxies encore inconnues, qui n'ont jamais été observées auparavant et qui sont donc non listées dans d'autres catalogues ou bases de données. Cependant, il n'y a pas des galaxies partout dans le ciel, et donc pas de détections HI pour une grande partie du scan. Cela entraîne une perte de temps, i.e une majeure partie du temps a été passé à scanner le ciel et peu de temps à poser sur des galaxies. De plus, le temps d'intégration est constant sur tout le ciel, et donc sur toutes les sources HI. Or, les galaxies spirales n'ont pas toutes le même flux apparent (de part leur luminosité intrinsèque et de part leur éloignement), et nécessitent ainsi des temps d'observation différents.

Les sondages avec cibles pointées consistent à pointer le télescope sur un grand nombre de galaxies, une par une. Contrairement aux sondages à l'aveugle, ce type d'observations permettent un gain de temps. En effet, moins de temps est perdu à regarder le ciel puisque chaque scan garantit une détection (à condition d'avoir sélectionner les sources HI à

observer correctement auparavant !). Cependant, cette méthode est biaisée : on ne pointe que sur des galaxies que l'on connaît déjà (i.e des galaxies déjà observées et présentes dans les bases de données). Cela limite les chances d'obtenir de nouvelles découvertes.

Voici une liste de sondages à 21 cm récents :

- **HIPASS** (*HI Parkes All Sky Survey*) : sondage à l'aveugle conduit au radiotélescope Parkes et réunissant les spectres HI de 5,317 galaxies (Barnes et al., 2001; Meyer et al., 2004; Wong et al., 2006). Ces données sont déjà incluses dans CF3 ;
- **ALFALFA** (*Arecibo Legacy Fast ALFA*) : sondage à l'aveugle conduit au radiotélescope Arecibo Giovanelli et al. (2005); Haynes et al. (2011). Un total d'environ 30,000 sources HI ont été détectées ;
- **AGES** (*AGN and Galaxy Evolution Survey*) : sondage à l'aveugle conduit à Arecibo (Auld et al., 2006; Cortese et al., 2008; Minchin et al., 2010; Taylor et al., 2012, 2013, 2014). Un total de 874 sources HI ont été détectées ;
- **GASS** (*Galactic All Sky Survey*) : sondage à cibles pointées conduit au télescope Arecibo durant lequel environ 800 sources ont été détectées (Catinella et al., 2013, 2012, 2010). Ces données sont déjà incluses dans CF3 ;
- **EBHIS** (*Effelsberg-Bonn HI survey*) : sondage à l'aveugle conduit au radiotélescope Effelsberg (Winkel et al., 2010). Les données ne sont pas encore publiques ;
- **2MTF** (*2MASS Tully-Fisher*) : conduit aux radiotélescopes GBT et Parkes (Masters et al., 2007; Hong et al., 2013; Masters et al., 2014). Un total de 2,062 galaxies spirales ont été détectées. Ces données sont déjà incluses dans CF3 ;
- **NIBLES** (*Nançay Interstellar Baryons Legacy Extragalactic Survey*) : sondage à cibles pointées conduit au radiotélescope Nançay (van Driel et al., 2008, 2016b), Butcher et al. (2016). Un total de 2,839 spectres HI ont été réunis ;
- **KLUN** (*Kinematics of the Local Universe*) : sondage à cibles pointées conduit à Nançay durant lequel 4,876 galaxies ont été détectées (Theureau et al., 1998, 2005, 2017).

Les données obtenues par les sondages en gras sont celles qui sont analysées dans ce chapitre.

3.2 Données provenant du radiotélescope Nançay

Durant cette thèse, les largeurs des raies à 21 cm ont été mesurées sur les données des spectres provenant du radiotélescope Nançay, situé dans le centre de la France, dans la commune de Nançay.

Pour cela, deux ensembles de spectres à 21 cm ont été considérés ; l'un a été publié par la collaboration KLUN en 2017, et l'autre correspond aux observations effectuées par la collaboration NIBLES. Dans les deux cas, le récepteur FORT a été utilisé lors des sessions d'observation. Les spectres ont été obtenus dans deux polarisations (horizontale et verticale), pour 4096 canaux. Les galaxies ont été observées avec une largeur de bande de 50 MHz, ce qui permet d'obtenir une très fine résolution spectrale de 2.6 km/s à 21 cm.

Les données KLUN17 et NIBLES ont déjà été réduites par les collaborations ayant effectué les observations. Cependant, il a fallu re-mesurer les largeurs de raie à 21 cm des ces deux échantillons avec la méthode utilisée par la collaboration *Cosmicflows* afin de les intégrer à la compilation *All Digital HI* (ADHI) et ainsi au futur catalogue *Cosmicflows-4*. Le catalogue ADHI est une compilation de toutes les largeurs de raie HI mesurées par la collaboration *Cosmicflows*.

3.2.1 Kinematic in the Local UNiverse

Cette section commence par la description de l'échantillon de spectres HI publié dans Theureau et al. (2017). Cet article est la toute dernière publication de la série de sondages HI nommés EDS (*Extragalactic Distance Scale*, Bottinelli et al., 1992, 1993) et KLUN (*Kinematic in the Local UNiverse*, di Nella et al., 1996; Theureau et al., 1998; Paturel et al., 2003b; Theureau et al., 2005) dont les observations à 21 cm ont été conduites au radiotélescope de Nançay.

Sélection des sources HI

Les galaxies observées ont été sélectionnées parmi une compilation de 2.7 millions de galaxies, extraites des données DSS (Paturel et al., 2003a), DENIS (Paturel et al., 2005) et 2MASS (dans l'infrarouge, Jarrett et al., 2000a). Cette compilation a été préparée à l'aide de la base de données extragalactiques HYPERLEDA (Makarov et al., 2014). Toutes les galaxies de cette compilation ont été inspectées à l'oeil, une par une, afin de sélectionner uniquement les galaxies spirales, i.e de type morphologique allant de **Sa** à **Sm**.

Les critères de sélection considérés afin de distinguer les galaxies adéquates pour des observations à 21 cm sont les suivants :

- les galaxies doivent avoir une déclinaison telle que $\delta > -40^\circ$ (déclinaison limite de Nançay) ;
- la collaboration KLUN s'est concentrée surtout sur les galaxies ayant une vitesse de récession $4000 < cz < 10000$ km/s. Cet intervalle correspond à la région où l'on constate un manque de mesures TF dans la littérature. Quelques galaxies plus proches ont aussi été considérées.

Stratégie d'observation

Le temps de pose par galaxie est compris entre 15 minutes et 6 heures, et avec un temps de pose moyen de 2 heures. La majorité des spectres obtenus ayant un flux apparent assez faible, un temps de pose assez long était nécessaire afin d'obtenir des paramètres avec une erreur acceptable pour estimer une distance.

Après chaque session, les spectres obtenus pour chaque galaxie sont inspectés. Si le signal sur bruit est supérieur à 5 (ou si $\text{RMS} < 2$ mJy, ce qui correspond à un seuil raisonnable

pour une détection à Nançay), la galaxie est classée comme *complétée* et ne sera pas observée davantage.

Le spectre final (pour publication) est obtenu avec la somme de toutes les intégrations dans les deux polarisations. De plus, un smoothing est aussi appliqué afin d'éliminer un maximum d'interférences radio (RFI). La résolution spectrale typique après smoothing est de 10 km/s à 21 cm, ce qui est moyen.

Toutes publications incluses, la collaboration KLUN a réuni les spectres HI de 4876 galaxies spirales en 6 ans, obtenus entre 2000 et 2006, dont 3039 ont déjà été analysés par la collaboration *Cosmicflows* (Courtois et al., 2009). Seuls les derniers spectres de la publication Theureau et al. (2017) restaient à re-analyser. Cet échantillon rassemblant 828 spectres HI sera nommé par la suite KLUN17. Il apporte un cinquième de nouvelles données Nançay.

Analyse des spectres HI pour intégration à Cosmicflows-4

Au final, après avoir mesuré la largeur de raie HI pour les 828 spectres cités précédemment, il apparait que seulement 500 detections ont été obtenues avec la méthode de détection de raie de la collaboration *Cosmicflows*, soit un taux de détection de 60%. Concernant les nouvelles additions à ADHI, 393 nouvelles galaxies ont été ajoutées dont 325 avec une erreur $e_W \leq 20$ adéquate pour estimer des distances de galaxies, soit un taux de succès de 40% (2 sur 5 spectres adéquats pour appliquer la relation de TF).

La figure 3.3 gauche présente une carte du ciel en projection aïtoff en coordonnées équatoriales. Les points correspondent aux galaxies KLUN17 dont les spectres ont été analysés durant cette thèse. Les galaxies associées à tous les spectres considérés (soit 828) sont représentées par des points rouges. Les galaxies détectées et qui étaient déjà listées dans ADHI correspondent aux carrés gris, tandis que les détections qui n'étaient pas présentes dans ADHI correspondent aux étoiles oranges. On peut remarquer que la distribution des galaxies est quasi-homogène dans la zone couverte par KLUN17.

La figure 3.3 montre les histogrammes des vitesses héliocentriques des données KLUN17. Le code couleur est le même que ci-dessus. On peut remarquer que les données comportent des galaxies lointaines allant jusqu'à 10,000 km/s, y compris les nouveaux ajouts à ADHI. Les distances obtenues à partir des largeurs de raies mesurées sur les données KLUN17 seront donc un élément essentiel de CF4.

La table 3.2 présente les paramètres HI obtenus à partir des données KLUN17 pour 10 galaxies. Les colonnes sont les suivantes : (1) numéro PGC, (2) coordonnées J2000, (3) vitesse héliocentrique en km/s, (4) largeur de raie W_{50} en km/s, (5) erreur sur la mesure de la largeur de raie e_W en km/s, (6) masse HI M_{HI} en masses solaires, et (7) magnitude à 21 cm m_{21} . La planche 3.3 présente les spectres HI après la mesure de W_{50} et les images multi-bandes extraites de PanSTARRS pour les mêmes galaxies. Les galaxies sont triées par largeur de raie croissante, i.e de faible à grande vitesse de rotation.

Les largeurs de raie obtenues, notées $W_{50}^{\text{Cosmicflows}}$ sont comparées, sur la figure 3.6, à celles publiées par Theureau et al. (2017) et notées W_{50}^{KLUN} . Le panel de gauche compare directement ces deux paramètres, tandis que le panel de droite met en évidence leur différence.

Les droites de référence sont tracées en rouge et correspondent à $W_{50}^{\text{Cosmicflows}} = W_{50}^{\text{KLUN}}$. Les largeurs de raie obtenues à l'aide de la méthode de la collaboration *Cosmicflows* sont cohérentes avec celles obtenues par [Theureau et al. \(2017\)](#). Une légère différence est visible pour les petites et grandes largeurs de raie ($W_{50} < 200$ km/s et $W_{50} > 400$ km/s).

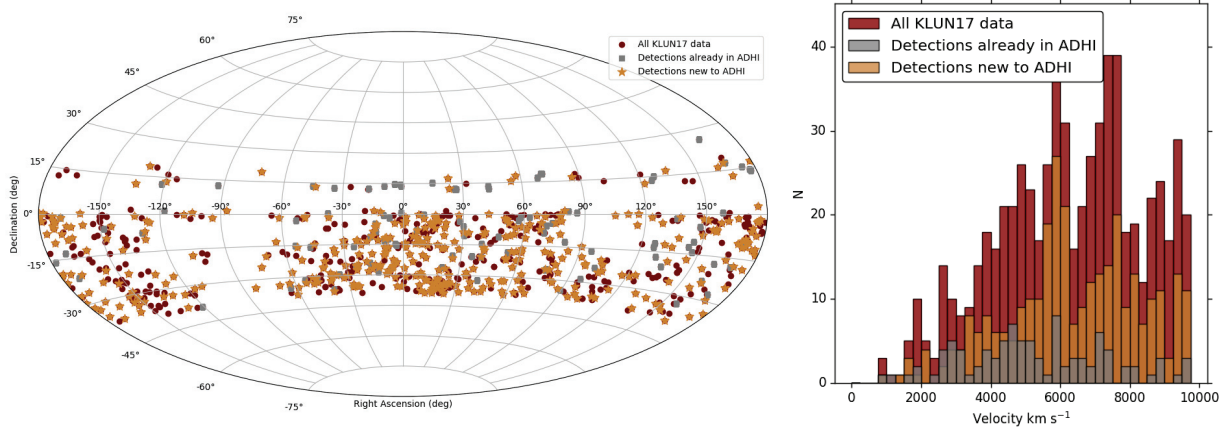


FIGURE 3.3 – Gauche : carte du ciel en projection aitoff en coordonnées équatoriales. Les points correspondent aux galaxies KLUN17 dont les spectres ont été analysés pendant cette thèse. Droite : histogrammes des vitesses héliocentriques de ces mêmes galaxies. Les galaxies associées à tous les spectres considérés (soit 828) sont représentées en rouge. Les galaxies détectées appartenant déjà à ADHI sont en gris, tandis que les détections qui n'étaient pas présentes dans ADHI auparavant sont en orange.

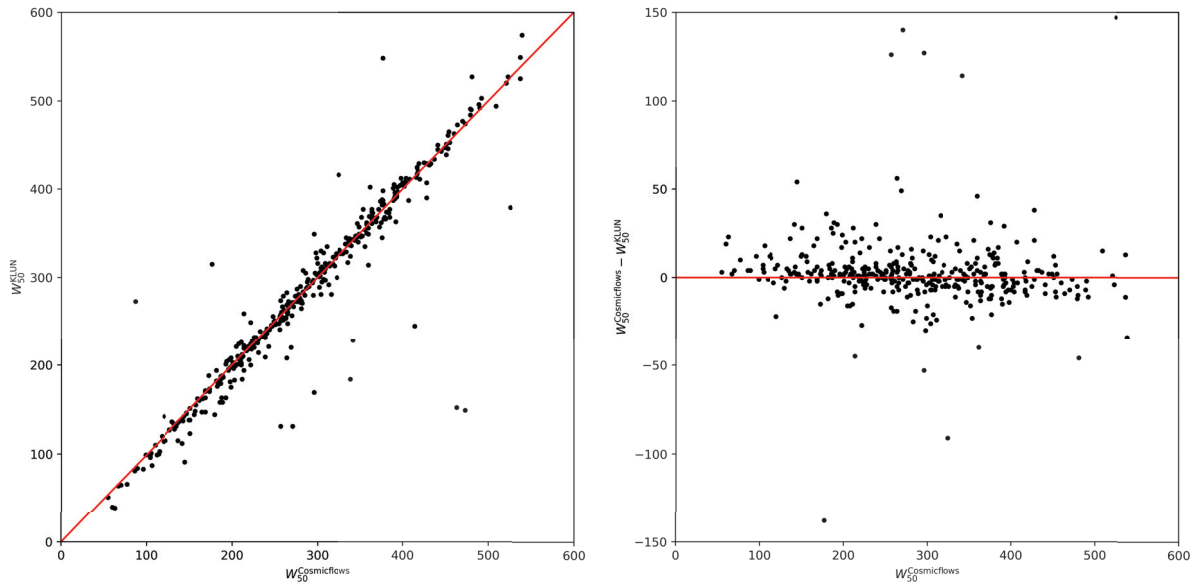
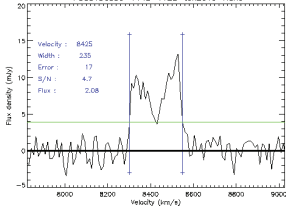

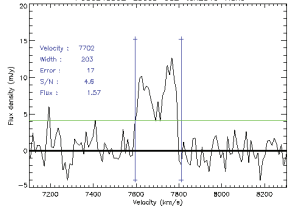
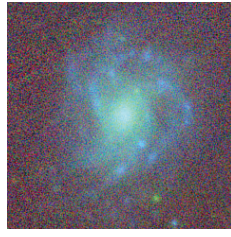
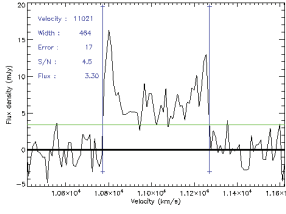

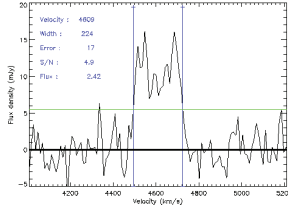

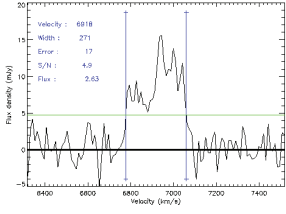

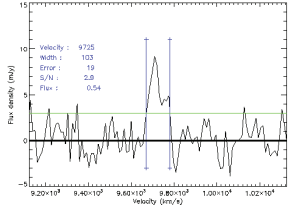

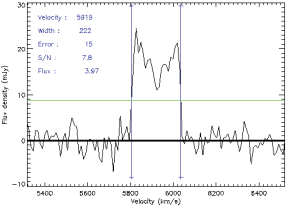

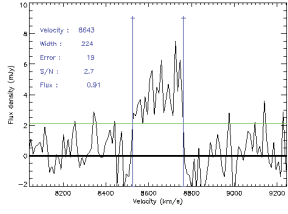

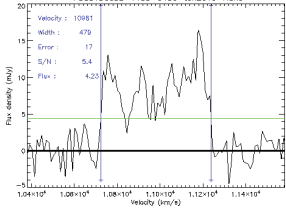

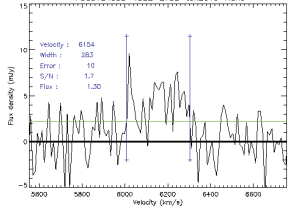



FIGURE 3.4 – Comparaison des largeurs de raie $W_{50}^{\text{Cosmicflows}}$ obtenues à l'aide de la méthode de la collaboration *Cosmicflows* à celles obtenues par [Theureau et al. \(2017\)](#), notées W_{50}^{KLUN} . Le panel de gauche compare directement ces deux paramètres, tandis que le panel de droite met en évidence leur différence. Les droites de référence telles que $W_{50}^{\text{Cosmicflows}} = W_{50}^{\text{KLUN}}$ sont tracées en rouge.

TABLE 3.2 – Exemples de paramètres HI extraits des spectres à 21 cm provenant du sondage HI KLUN17. Les colonnes sont les suivantes : (1) numéro PGC de la galaxie ; (2) coordonnées J2000 ascension droite (en h :m :s) — déclinaison (en d :m :s) ; (3) vitesse héliocentrique (en km/s) ; (4) largeur de raie à 21 cm (non corrigée, en km/s) ; (5) erreur sur la largeur de raie à 21 cm (en km/s) ; (6) logarithme décimal de la masse HI (divisée par la masse solaire) ; (7) magnitude à 21 cm.

PGC	J2000 RA-DEC	V_{hel} km s ⁻¹	W_{50} km s ⁻¹	e_W km s ⁻¹	$\log_{10}(M_{HI})$ M_{\odot}	m_{21}
156999	J114815.5-1122	8425	235	17	9.79	16.60
845592	J232053.8-2009	7702	203	17	9.59	16.91
10190	J024117.8-0955	11021	464	17	10.23	16.10
20014	J070121.5-2434	4609	224	17	9.33	16.44
71447	J232641.6-1137	6918	271	17	9.72	16.35
135364	J225629.7-1929	9725	103	19	9.33	18.07
5187	J012429.5-0433	5919	222	15	9.77	15.90
16889	J051131.2-0035	8643	224	19	9.46	17.50
156695	J113338.5-0456	10981	479	17	10.33	15.83
184993	J152200.2-2752	6154	283	10	9.32	17.12

TABLE 3.3 – Quelques spectres HI après la mesure de W_{50} et images multi-bandes extraites de PanSTARRS pour 10 galaxies extraites de l'échantillon KLUN17. Les galaxies sont triées par largeur de raie croissante, i.e de faible à grande vitesse de rotation. Les colonnes sont les suivantes : (1) et (3) numéro PGC, vitesse héliocentrique (en km/s), et spectre HI; (2) et (4) largeur de raie HI (en km/s), logarithme décimal de la masse HI (divisée par la masse solaire), magnitude à 21 cm, et image PanSTARRS de 1 arcmin de coté.

PGC / V_{hel} Profile	W_{50} / $\log M_{HI}$ / m_{21} PanSTARRS	PGC / V_{hel} Profile	W_{50} / $\log M_{HI}$ / m_{21} PanSTARRS
PGC156999 / 8425	235 / 9.79 / 16.60	PGC845592 / 7702	203 / 9.59 / 16.91
			
PGC10190 / 11021	464 / 10.23 / 16.10	PGC20014 / 4609	224 / 9.33 / 16.44
			
PGC71447 / 6918	271 / 9.72 / 16.35	PGC135364 / 9725	103 / 9.33 / 18.07
			
PGC5187 / 5919	222 / 9.77 / 15.90	PGC16889 / 8643	224 / 9.46 / 17.50
			
PGC156695 / 10981	479 / 10.33 / 15.83	PGC184993 / 6154	283 / 9.32 / 17.12
			

3.2.2 Nançay Interstellar Baryons Legacy Extragalactic Survey

L'échantillon NIBLES (*Nançay Interstellar Baryons Legacy Extragalactic Survey*) rassemble les spectres à 21 cm obtenus par observations de HI extragalactique à Nançay. Ces galaxies ont été sélectionnées depuis la DR5 de SDSS.

Sélection des sources HI

Tout comme les données KLUN, les galaxies qui ont été observées par la collaboration NIBLES ont été extraites d'un catalogue de départ. Dans le cas des données NIBLES, l'échantillon de départ contient environ 3000 galaxies qui ont été extraites de SDSS DR5. Ces galaxies ont ensuite été sélectionnées selon les critères suivants :

- pour une galaxie donnée, on doit avoir accès à la fois à sa magnitude absolue (derivée du redshift) dans la bande z et à son spectre optique dans la base de données SDSS (DR5) ;
- on considère seulement les galaxies situées dans l'univers local, c'est-à-dire celles ayant un redshift situé dans l'intervalle $900 < cz < 12000 \text{ km/s}$;
- afin d'obtenir une distribution uniforme en magnitude absolue, la collaboration NIBLES considère des bins en magnitude absolue M_z de 0.5 mag de largeur, où chaque bin contient environ le même nombre de galaxies (environ 150). Pour les bins les moins peuplés, ils considèrent tout les objets disponibles dans les données SDSS ;
- les galaxies les plus proches sont prioritaires. En effet, [van Driel et al. \(2016b\)](#) ont décidé de se concentrer d'abord sur les galaxies les plus proches, car la probabilité de les détecter est plus élevée que pour des galaxies plus lointaines. De plus, un volume d'univers similaire a été observé dans chaque bin de magnitude ;
- il n'y a aucun critère de selection concernant la couleur ;
- enfin, les galaxies faisant partie de l'amas Virgo et les galaxies ayant une déclinaison $0^\circ < \delta < 38^\circ$ n'ont pas été prises en compte.

Stratégie d'observation

Les spectres NIBLES ont été obtenus par *position-switch* : cycles de 10-10 paires de scans ON-OFF (sur et à côté de la source HI afin d'extraire le ciel). Chaque paire ON-OFF dure 4-4 secondes, soit un total de 8 secondes par paire. Les observations ont été conduites de Janvier 2007 à Décembre 2010. Un temps de télescope total de 3450 heures a été utilisé.

Le spectre final (pour publication) est obtenu avec la somme de toutes les intégrations dans les deux polarisations, avec correction d'éventuelles interférences radio. De plus, la ligne de base a été soustraite afin d'obtenir un spectre plat en dehors de la raie HI. Un smoothing a aussi été appliqué. La résolution spectrale finale est plutôt basse avec 18 km/s par pixel à 21 cm : cela correspond presque à la limite absolue de qualité spectrale nécessaire pour mesurer une distance par la relation TF (20 km/s, erreur instrumentale de 10% sur la largeur de raie).

Un total de 2839 spectres a été rassemblé par la collaboration NIBLES. Pour la suite, seulement l'échantillon de spectres étiquetés en tant que *détection* par NIBLES sera considéré, soit une compilation de 1864 profils HI. Cette compilation est disponible sur la base de données Vizier [van Driel et al. \(2016a\)](#).

Analyse des spectres HI pour intégration à *Cosmicflows-4*

Parmi les 2839 spectres obtenus par les observations à Nançay décrites ci-dessus, un total de 1864 spectres ont été analysés (ceux qui sont disponibles sur Vizier). Cela a conduit à 1443 détections en HI, ce qui donne un taux de détection de 51%. Cela a permis d'ajouter 808 nouvelles galaxies au catalogue ADHI, dont 564 avec erreur $e_W \leq 20$, i.e adéquates pour appliquer la TF et ainsi mesurer la distance, ce qui donne un taux de succès de seulement 20%, ce qui est très faible par rapport au nombre d'heure passées au télescope (3500h).

La figure 3.5 gauche présente une carte du ciel en projection aitoiff en coordonnées supergalactiques. Les points représentent les galaxies NIBLES dont les spectres ont été analysés durant cette thèse. Les galaxies associées à tous les spectres considérés (soit environ 1900) sont représentées par des points bleus. Les galaxies qui ont été détectées mais qui étaient déjà listées dans ADHI correspondent aux carrés gris, tandis que les détections correspondant aux galaxies qui n'étaient pas dans ADHI correspondent aux étoiles turquoises. Tout comme KLUN17, on peut voir que la distribution des galaxies est quasi-homogène dans la zone couverte par NIBLES.

La figure 3.5 montre les histogrammes des vitesses héliocentriques des données NIBLES considérées dans cette section. Le code couleur est le même que ci-dessus. On peut remarquer que les données comportent une grande majorité de galaxies proches, et peu de galaxies lointaines situées aux distances les plus intéressantes pour CF4 ($6,000 < cz < 10,000$ km/s).

La table 3.5 montre les paramètres HI mesurés sur les données KLUN17 pour 10 galaxies. Les colonnes sont les suivantes : (1) numéro PGC, (2) coordonnées J2000, (3) vitesse héliocentrique en km/s, (4) largeur de raie W_{50} en km/s, (5) erreur sur la mesure de la largeur de raie e_W en km/s, (6) masse HI M_{HI} en masses solaires, et (7) magnitude à 21 cm m_{21} . La planche 3.4 présente les spectres HI après la mesure de W_{50} et les images multibandes extraites de PanSTARRS pour les mêmes galaxies. Les galaxies sont triées par largeur de raie croissante, i.e de faible à grande vitesse de rotation définie par $v_{\text{rot}} = \frac{W_{50}}{2 \sin(i)}$ où i est l'inclinaison de la galaxie. Par ailleurs, comme on peut le voir sur la planche 3.4, certaines galaxies sont beaucoup trop de face pour être de bonnes candidates pour l'utilisation de la TF (par exemple PGC 2688).

Les largeurs de raie obtenues, notées $W_{50}^{\text{Cosmicflows}}$ sont comparées, sur la figure 3.6, à celles publiées par [van Driel et al. \(2016b\)](#) et notées W_{50}^{NIBLES} . Le panel de gauche compare directement ces deux paramètres, tandis que le panel de droite met en évidence leur différence. Les droites de référence sont tracées en rouge et correspondent à $W_{50}^{\text{Cosmicflows}} = W_{50}^{\text{NIBLES}}$. Un décalage important entre les deux méthodes de mesure de W_{50} est mis en évidence. Pour exemple, le spectre de la galaxie PGC 2220, pour lequel la différence entre $W_{50}^{\text{Cosmicflows}}$ et W_{50}^{NIBLES} est égale à 64 km/s, est montré sur la figure 3.7. Le panel de droite présente la mesure de la largeur de raie par la collaboration *Cosmicflows*, délimitée

par les traits verticaux bleus. Celui de gauche présente la mesure faite par la collaboration NIBLES (extraite de [van Driel et al., 2016b](#)), schématisée par la flèche bleue en dessous du spectre. Contrairement à *Cosmicflows*, la collaboration NIBLES n'a pas inclus le pic vers 5,150 km/s, ce qui explique la différence entre les deux mesures.

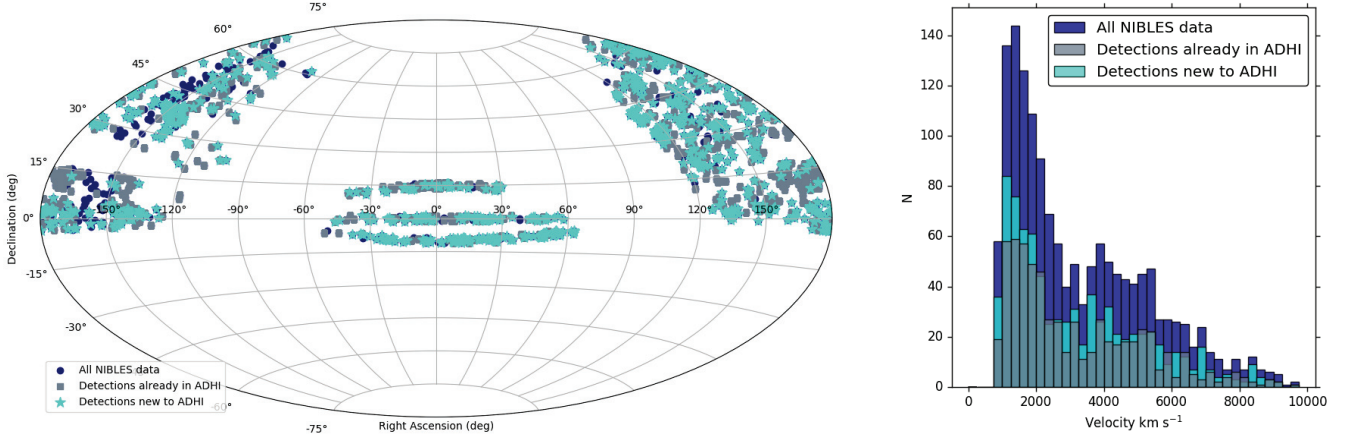


FIGURE 3.5 – Gauche : carte du ciel en projection aitoff en coordonnées équatoriales. Les points correspondent aux galaxies NIBLES dont les spectres ont été analysés pendant cette thèse. Droite : histogrammes des vitesses héliocentriques de ces mêmes galaxies. Les galaxies associées à tous les spectres considérés sont représentées en bleu foncé. Les galaxies détectées appartenant déjà à ADHI sont en gris, tandis que les détections qui n'étaient pas présentes dans ADHI auparavant sont en bleu clair.

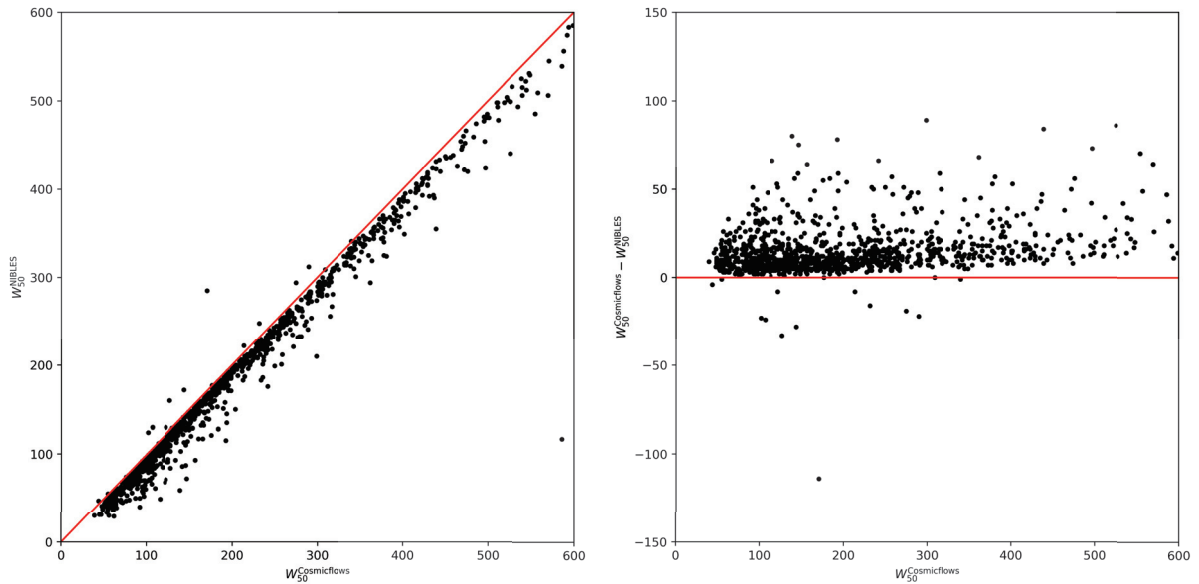


FIGURE 3.6 – Comparaison des largeurs de raie $W_{50}^{\text{Cosmicflows}}$ obtenues à l'aide de la méthode de la collaboration *Cosmicflows* à celles obtenues par [van Driel et al. \(2016b\)](#), notées W_{50}^{NIBLES} . Le panel de gauche compare directement ces deux paramètres, tandis que le panel de droite met en évidence leur différence. Les droites de référence telles que $W_{50}^{\text{Cosmicflows}} = W_{50}^{\text{NIBLES}}$ sont tracées en rouge.

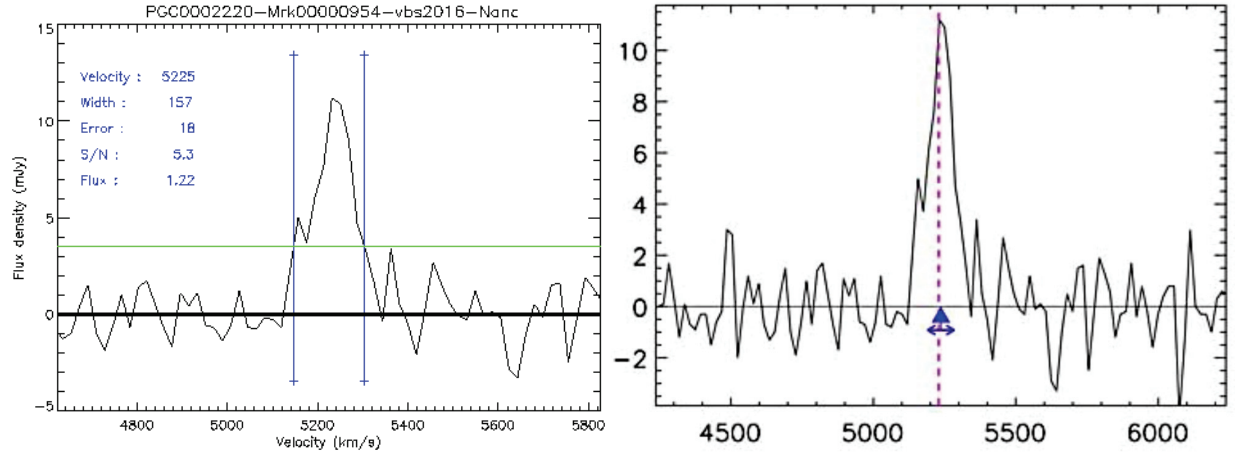
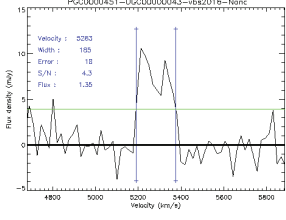

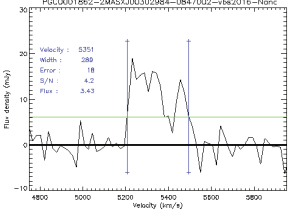

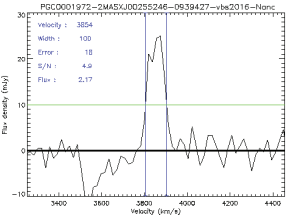
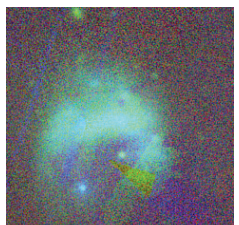
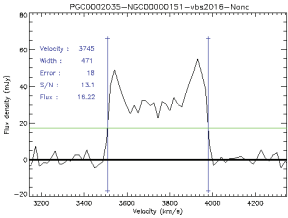
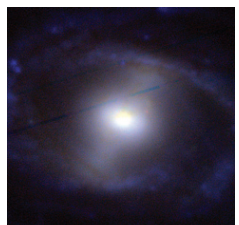
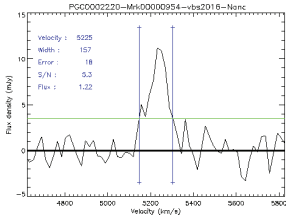

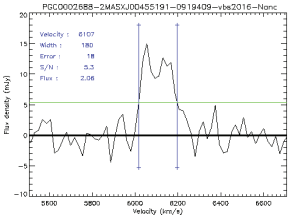

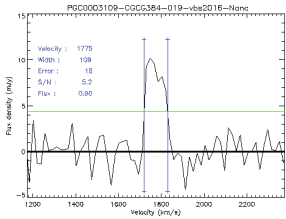
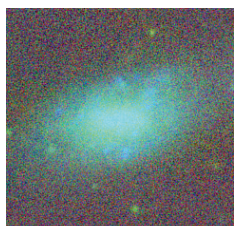
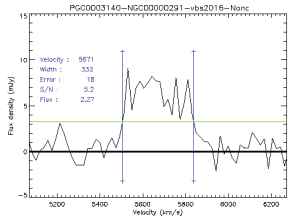
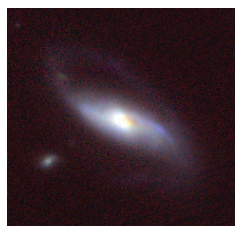
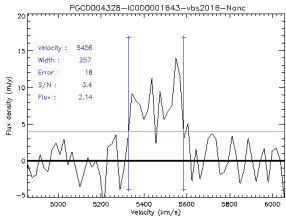

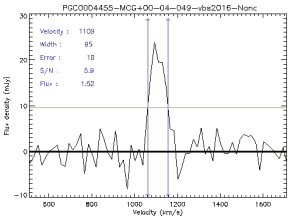
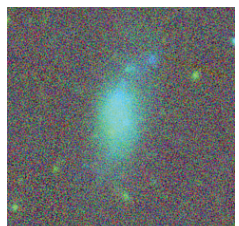


FIGURE 3.7 – Spectre de la galaxie PGC 2220, pour lequel la différence entre $W_{50}^{\text{Cosmicflows}}$ et W_{50}^{NIBLES} est égale à 64 km/s. Droite : mesure de W_{50} faite par la collaboration NIBLES (traits verticaux bleus). Gauche : mesure de W_{50} faite par la collaboration NIBLES (flèche bleue, [van Driel et al., 2016b](#))

TABLE 3.4 – Paramètres HI extraits des spectres à 21 cm provenant du sondage HI NIBLES. Les colonnes sont les suivantes : (1) numéro PGC de la galaxie ; (2) coordonnées J2000 ascension droite (en h :m :s) — déclinaison (en d :m :s) ; (3) vitesse héliocentrique (en km/s) ; (4) largeur de raie à 21 cm (non corrigée, en km/s) ; (5) erreur sur la largeur de raie à 21 cm (en km/s) ; (6) logarithme décimal de la masse HI (divisée par la masse solaire) ; (7) magnitude à 21 cm.

PGC	J2000 RA-DEC	V_{hel} km s ⁻¹	W_{50} km s ⁻¹	e_W km s ⁻¹	$\log_{10}(M_{\text{HI}})$ M_{\odot}	m_{21}
156999	J114815.5-1122	8425	235	17	9.79	16.60
845592	J232053.8-2009	7702	203	17	9.59	16.91
10190	J024117.8-0955	11021	464	17	10.23	16.10
20014	J070121.5-2434	4609	224	17	9.33	16.44
71447	J232641.6-1137	6918	271	17	9.72	16.35
135364	J225629.7-1929	9725	103	19	9.33	18.07
5187	J012429.5-0433	5919	222	15	9.77	15.90
16889	J051131.2-0035	8643	224	19	9.46	17.50
156695	J113338.5-0456	10981	479	17	10.33	15.83
184993	J152200.2-2752	6154	283	10	9.32	17.12

TABLE 3.5 – Spectres HI après la mesure de W_{50} et images multi-bandes extraites de PanSTARRS pour 10 galaxies extraites de l'échantillon NIBLES. Les galaxies sont triées par largeur de raie croissante, i.e de faible à grande vitesse de rotation. Les colonnes sont les suivantes : (1) et (3) numéro PGC, vitesse héliocentrique (en km/s), et spectre HI ; (2) et (4) largeur de raie HI (en km/s), logarithme décimal de la masse HI (divisée par la masse solaire), magnitude à 21 cm, et image PanSTARRS de 1 arcmin de coté.

PGC / V_{hel} Profile	W_{50} / $\log M_{HI}$ / m_{21} PanSTARRS	PGC / V_{hel} Profile	W_{50} / $\log M_{HI}$ / m_{21} PanSTARRS
PGC451 / 5283	185 / 9.20 / 17.07	PGC1862 / 5351	289 / 9.61 / 16.06
			
PGC1972 / 3854	100 / 9.13 / 16.56	PGC2035 / 3745	471 / 9.98 / 14.37
			
PGC2220 / 5225	157 / 9.15 / 17.18	PGC2688 / 6107	180 / 9.51 / 16.62
			
PGC3109 / 1775	109 / 8.08 / 17.51	PGC3140 / 5671	332 / 9.49 / 16.51
			
PGC4328 / 5456	257 / 9.43 / 16.57	PGC4455 / 1109	95 / 7.89 / 16.95
			

3.3 Données provenant du radiotélescope Arecibo

3.3.1 Arecibo Legacy Fast ALFA : description

Tous les détails sur le sondage ALFALFA discutés dans cette partie sont extraits de Giovanelli et al. (2005) et Haynes et al. (2011).

Arecibo Legacy Fast ALFA (ALFALFA) est un sondage de HI extragalactique utilisant le radiotélescope Arecibo. La prise de données a débuté en 2005 et s'est terminée en 2012. Plus de 30,000 sources HI ont été détectées jusqu'à $z \approx 0.06$. Le but de ALFALFA est de sonder l'univers local en HI sur un volume cosmologique significatif. En effet, les résultats ALFALFA fournissent un catalogue dont les données atteignent les superamas Persée-Poisson et *Great Wall*, grande structure connectant les superamas Coma et Hercules.

Le radiotélescope Arecibo est le radiotélescope le plus sensible au monde jusqu'à maintenant. Il sera éventuellement dépassé par le radiotélescope FAST localisé en Chine. Les observations à Arecibo sont possibles sur la bande L à 1.4 GHz. La précision de pointage de Arecibo est supérieure à 20 arcmin. Cela permet d'identifier immédiatement l'équivalent optique le plus probable pour chaque détection HI (qui est identifié avec SDSS dans la collaboration ALFALFA).

ALFA (*Arecibo L Band Feed Array*) est un ensemble de sept cornets d'alimentation qui peuvent recevoir deux polarisations (horizontale et verticale), dans des fréquences situées entre 1225 MHz et 1525 MHz. La taille du radiolobe du radiotélescope Arecibo est de 3.3×3.8 arcmin (azimut x zenith), ce qui est très pratique par rapport à Nançay et au GBT par exemple, car cela évite les sources possible de confusion entre galaxies au même redshift. Ainsi, cela permet de réaliser des mesures de galaxies d'amas, ce qui est impossible avec des télescopes ayant de grands radiolobes. Par ailleurs, six cornets d'alimentation sont arrangés aux coins d'un hexagone régulier, tandis que le septième cornet est situé au centre de l'hexagone.

La résolution spectrale des spectres à 21 cm obtenus par ALFALFA est de 5.2 km/s. Cette grande résolution permet d'obtenir des mesures de largeur de raie plus précises, car il sera plus facile de détecter les limites [0% et 100%] des raies à 21 cm qu'avec une résolution plus faible.

En tout, les données ALFALFA couvrent 7000 deg² de ciel. Les ascensions droites α et déclinaisons δ considérées dans le sondage ALFALFA sont $7h30 < \alpha < 16h30$ (region de printemps), $22h < \alpha < 3h$ (region d'automne) et $0^\circ < \delta < +38^\circ$.

La stratégie d'observation de la collaboration ALFALFA est la suivante. Tout d'abord, toutes les observations sont effectuées durant la nuit afin de diminuer la contamination des RFI. De plus, le ciel est scanné en continu (*drift scan*, scan en mode balayage) : le radiolobe central suit une trajectoire telle que la déclinaison reste fixe (avec de légers déplacements du télescope). Le bras de collection de la lumière (*feed*) est positionné de telle sorte que les sept *feed cornets* du récepteur ALFA observent des déclinaisons équidistantes. Par ailleurs, une session d'observation est dédiée à une seule déclinaison. Les données obtenues pour un balayage du ciel pour une déclinaison unique ont deux dimensions : le temps et la fréquence.

Pour des scans balayant le ciel, la méthode permettant de soustraire le ciel est différente de celle pour les observations pointées (paire de scans ON-OFF). En effet, durant un scan, les données sont recueillies continuellement sans (presque) déplacer le télescope, il est donc impossible d'effectuer des paires de scans ON-OFF.

Le taux d'échantillonnage des scans dans le cadre du sondage ALFALFA est de 1 Hz, ce qui signifie qu'un spectre de 4096 chaînes est enregistré chaque seconde pour chaque polarisation et chaque cornet d'alimentation. Ainsi, les caractéristiques de la bande passante varient avec le temps, et il est donc possible de modéliser ces variations (en dehors des sources HI) avec des polynômes d'ordre faible afin de soustraire le ciel. En effet, tout les 600 spectres (10 min de scan en balayage), on obtient une carte en deux dimensions de la bande passante qui peut être soustraite à chaque spectre.

Les données ont été réduites par la collaboration ALFALFA. Chaque spectre (après soustraction du ciel et dans chaque polarisation) est inspecté afin de traiter les RFI.

3.3.2 Intégration des données ALFALFA à Cosmicflows-4

Le catalogue $\alpha 100$ contenant les résultats HI de toutes les sources extragalactiques mesurées par la collaboration ALFALFA est public¹ depuis Mars 2018. Cependant les spectres à 21 cm ne sont pas encore accessibles. Il a donc fallu analyser les paramètres HI obtenus par ALFALFA et comparer leurs paramètres à ceux correspondant à 40% du survey ($\alpha 40$) que la collaboration *Cosmicflows* avait déjà mesuré sur spectres avec la méthode *maison* pour W_{50} , afin d'intégrer ces nouveaux résultats au catalogue de la base de données EDD : *All Digital HI* (ADHI).

La figure 3.8 montre une projection aitoft en coordonnées supergalactiques des 32,612 sources HI issues du catalogue $\alpha 100$. Les galaxies dont l'inclinaison est inférieure à 45° sont représentées en rouge, et les galaxies dont l'inclinaison est supérieure à 45° sont en vert, i.e les candidates potentielles pour la TF. Dans les deux cas, une distribution uniforme dans les zones du ciel couvertes par ALFALFA est constatée.

Par ailleurs, la figure 3.9 présente un histogramme des vitesses héliocentriques des sources $\alpha 100$ à gauche et un histogramme de leur inclinaison à droite. La majorité des galaxies ont une inclinaison supérieure à 45° . De plus les distances obtenues par TF à partir de ces données HI seront une bonne extension de CF3 vers des distances plus lointaines car en effet, la couverture en redshift va jusqu'à 17,000 km/s.

1. Le catalogue $\alpha 100$ est disponible sur <http://egg.astro.cornell.edu/alfalfa/data/index.php>.

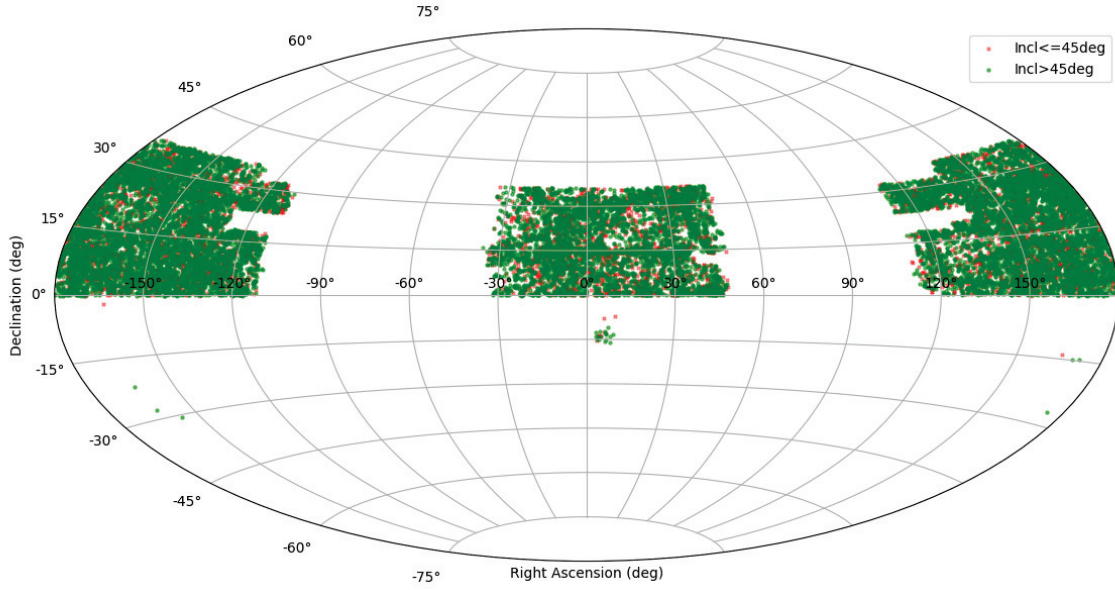


FIGURE 3.8 – Projection aïtoff en coordonnées équatoriales des 32,612 sources HI issues du catalogue $\alpha 100$. Les galaxies dont l’inclinaison est inférieure à 45° sont en rouge, et les galaxies dont l’inclinaison est supérieure à 45° en vert (i.e les candidates potentielles pour la TF).

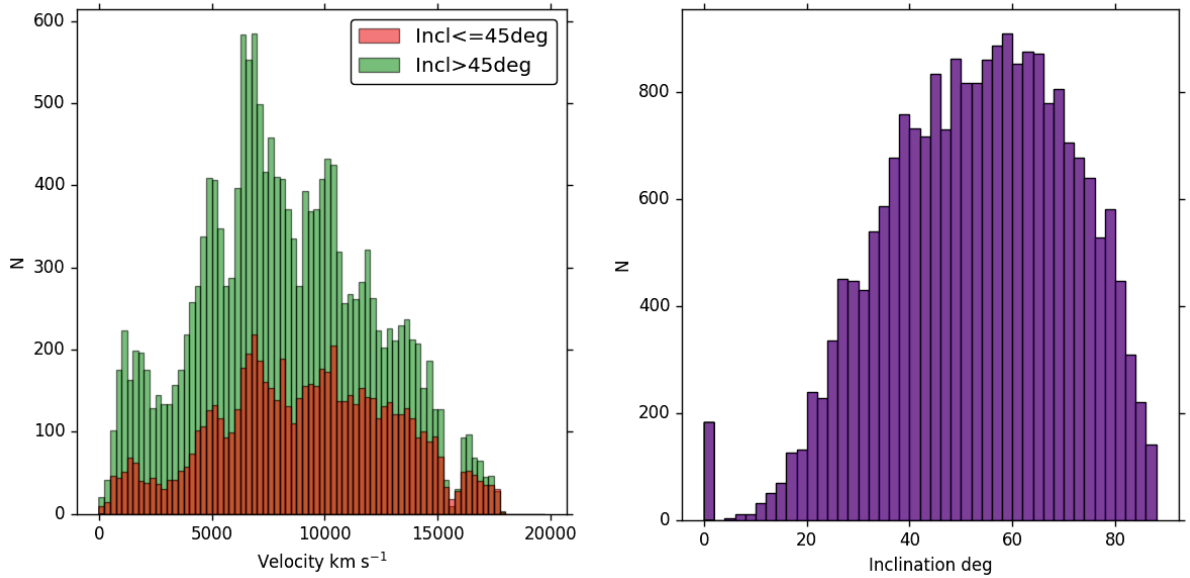


FIGURE 3.9 – Gauche : histogramme des vitesses héliocentriques des sources $\alpha 100$. Droite : histogramme de l’inclinaison des galaxies du catalogue $\alpha 100$. La majorité des galaxies ayant une inclinaison supérieure à 45° et leur redshift allant jusqu’à environ 17,000 km/s, les distances obtenues par TF à partir de ces données HI seront un bon complément à CF3.

Comparaison des données $\alpha 40$ aux paramètres ADHI

Le but est d'identifier les spectres HI de qualité suffisante afin de déduire une distance par la relation de Tully-Fisher, i.e les mesures de largeur de raie avec une erreur inférieure à 20 km/s obtenues avec la méthode *maison* de la collaboration.

Haynes et al. (2011) ont publié le catalogue $\alpha 40$ contenant 15,000 sources HI extragalactiques dont les largeurs de raie ont été mesurées par la collaboration ALFALFA. Ce catalogue représente 40% du volume total couvert par le sondage ALFALFA. Les spectres du catalogue $\alpha 40$ ont été ré-analysés par la collaboration *Cosmicflows*, et les largeurs de raies ont été ainsi re-mesurées. À partir des 15,000 sources HI, un total de 4,393 largeurs de raie ont été obtenues dont 3,898 avec une erreur inférieure à 20 km/s (i.e adéquates pour mesurer une distance). Il a donc été possible de comparer les paramètres HI obtenus pour ces 3,898 galaxies par les deux collaborations, ALFALFA et *Cosmicflows*, afin de pouvoir par la suite sélectionner les sources HI du catalogue $\alpha 100$ adéquates, dont les spectres ne sont pas accessibles, pour mesurer une distance, et éventuellement estimer les valeurs qu'auraient divers paramètres HI (dont W_{50} et e_W) des spectres HI ALFALFA si ils avaient été analysés par la collaboration *Cosmicflows*. Par souci de clarté, les différents paramètres HI qui seront considérés dans la suite de cette sous-partie sont définis de la manière suivante : W_{50}^i , e_W^i et SNR^i pour la largeur de raie et son erreur, et le ratio signal sur bruit respectivement, où l'indice i représente la collaboration considérée. La notation $i = \text{CF}$ sera utilisée pour la collaboration *Cosmicflows* (i.e ADHI), et $i = \alpha$ pour la collaboration ALFALFA.

Les erreurs sur les largeurs de raie e_W^α listées dans les catalogues $\alpha 40$ et $\alpha 100$ sont définies de la manière suivante : elles correspondent à la racine de la somme des carrés des deux composantes. La première composante est une erreur statistique qui dépend principalement du ratio signal sur bruit SNR^α . La seconde est une erreur systématique associée aux erreurs faites par l'utilisateur lorsqu'il estime les limites de la raie HI. Dans la majorité des cas, l'erreur systématique est significativement plus petite que l'erreur statistique. Ainsi, on ignore la première au profit de la seconde. On a donc en majorité une relation directe entre e_W^α et SNR^α . La définition de e_W^α est ainsi similaire à celle de la collaboration *Cosmicflows*, cependant la relation exacte n'est pas donnée par la collaboration ALFALFA. Sur la figure 3.10 gauche, les erreurs e_W^α issues du catalogue $\alpha 40$ sont comparées à celles obtenues par la collaboration *Cosmicflows* sur les mêmes spectres HI et notées e_W^{CF} . Il apparaît que $e_W^\alpha \neq e_W^{\text{CF}}$.

La figure 3.10 milieu compare les ratios signal sur bruit SNR^α et SNR^{CF} . Seules les galaxies pour lesquelles $e_W^{\text{CF}} \leq 20$ sont considérées. On peut à nouveau remarquer que les paramètres obtenus par les deux collaborations sont très différents : $\text{SNR}^\alpha \neq \text{SNR}^{\text{ADHI}}$. En modélisant la relation entre SNR^α et SNR^{ADHI} par une fonction affine, on obtient (en vert sur la figure 3.10 milieu) :

$$\text{SNR}^\alpha = 1.7 + 2.9 \text{SNR}^{\text{CF}}. \quad (3.19)$$

L'équation 3.19 relie les valeurs des ratios signals sur bruits mesurées par les deux collaborations, *Cosmicflows* et ALFALFA, malgré leur définition différente. En utilisant cette équation et les équations 3.14, 3.15 et 3.16, il est ainsi possible de passer d'une définition à l'autre, et ainsi estimer les erreurs sur les largeurs de raie e_W^{CF} à partir du signal sur bruit

SNR^α . Ces erreurs estimées à partir des données ALFALFA sont comparées à celles qui ont été obtenues auparavant par la collaboration *Cosmicflows* sur la figure 3.10 droite, qui montre un histogramme de la différence entre ces deux erreurs. Les erreurs estimées sont notées $e_W^{\text{CF},\text{est}}$ et les erreurs mesurées sont notées $e_W^{\text{CF},\text{mes}}$. Cet histogramme est modélisé par une distribution normale centrée $\mu = 0.4$ et de largeur $\sigma = 1$, représentée en vert. Les erreurs mesurées par la collaboration *Cosmicflows* sont retrouvées avec succès à partir des données ALFALFA, avec une incertitude maximale de 5 km/s. Cela permet ainsi d'identifier les sources ALFALFA avec une erreur sur largeur de raie $e_W^{\text{CF},\text{est}} \leq 20$ km/s.

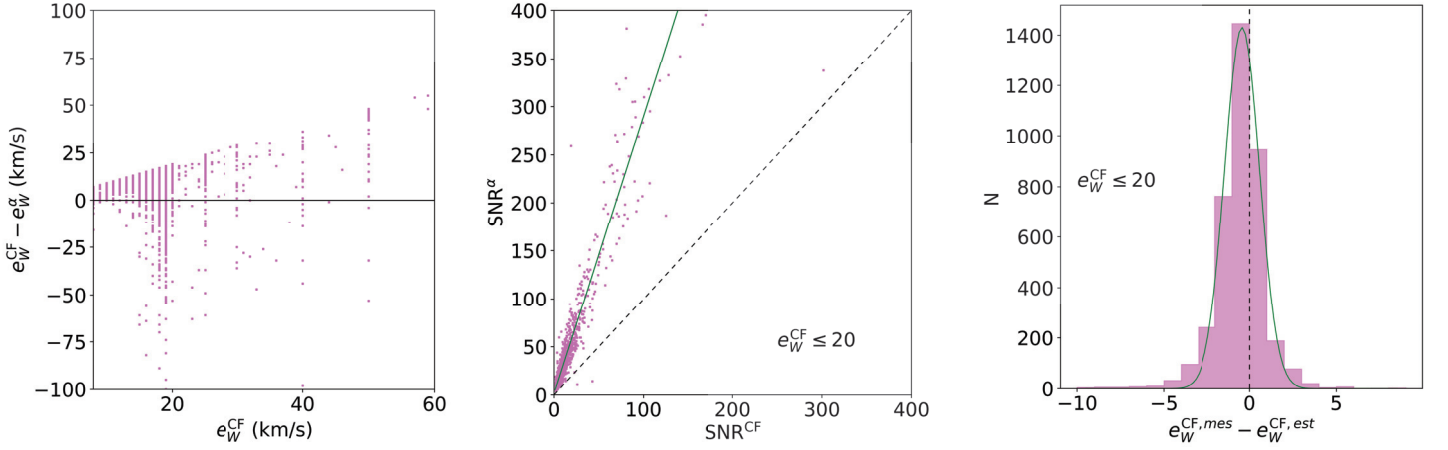


FIGURE 3.10 – Gauche : comparaison des erreurs sur les largeurs de raie à 21 cm e_W^{CF} et e_W^α . Milieu : comparaison des ratios signal sur bruit SNR^α et SNR^{CF} . La droite en trait plein vert représente la relation entre les deux paramètres modélisée par une fonction affine (voir équation 3.19). Droite : histogramme de la différence entre les vraies erreurs sur les largeurs de raie *Cosmicflows* $e_W^{\text{CF},\text{mes}}$ et les erreurs ALFALFA dans la définition *Cosmicflows* $e_W^{\text{CF},\text{est}}$. Cet histogramme est modélisé par une distribution normale centrée $\mu = 0.4$ et de largeur $\sigma = 1$, représentée en vert.

Après avoir identifié les spectres HI avec une erreur adéquate pour appliquer la TF, il faut comparer les largeurs de raie W_{50} que l'on a mesuré avec la méthode *Cosmicflows* W_{50}^{CF} , à celles obtenues par la collaboration ALFALFA W_{50}^α , afin de corriger un éventuel biais (qui peut-être dû par exemple à une méthode de mesure différente ou à une résolution spectrale différente).

La largeur de raie W_{50c}^α est définie dans Haynes et al. (2011) de la manière suivante. Cette quantité est mesurée à 50% du flux de chaque coté de la raie. Les effets d'élargissement dus à la résolution spectrale Δv et au redshift z sont corrigés de la manière suivante :

$$W_{50c}^\alpha = \frac{\sqrt{(W_{50}^\alpha)^2 - (\Delta v)^2}}{1 + z}. \quad (3.20)$$

Cependant aucune correction due aux effets de turbulence ou à l'inclinaison de la source n'est effectuée. Les largeurs de raie W_{50}^{CF} n'étant pas corrigées des effets de résolution spectrale et effets cosmologiques, elles seront comparées aux largeurs de raie ALFALFA non corrigées W_{50}^α afin de déduire une éventuelle relation entre ces deux paramètres.

La figure 3.11 gauche montre la différence $W_{50}^{\text{CF}} - W_{50}^{\alpha}$ en fonction du signal sur bruit SNR^{α} . Une dispersion importante est observée pour des ratios signal sur bruit faibles, i.e environ $\text{SNR}^{\alpha} < 15$. Cela montre que les largeurs de raie à 21 cm obtenues par les deux collaborations *Cosmicflows* et ALFALFA sont très différentes lorsque le ratio signal sur bruit est trop faible. Un histogramme de la différence $W_{50}^{\text{CF}} - W_{50}^{\alpha}$ pour quatre seuils sur le SNR est présenté sur la figure 3.11 milieu : sans seuil en bleu, $\text{SNR} > 10$ en rouge, $\text{SNR} > 15$ en vert et $\text{SNR} > 20$ en jaune. Seules les galaxies pour lesquelles $e_W^{\text{CF}} \leq 20$ km/s sont prises en compte. Cette figure montre que la différence $W_{50}^{\text{CF}} - W_{50}^{\alpha}$ est réduite pour $\text{SNR} > 15$ et $\text{SNR} > 20$.

Pour estimer les distances des galaxies par l'application de la relation de Tully-Fisher à l'aide des largeurs de raie provenant du catalogue ALFALFA, ce sont seulement les galaxies avec $e_W^{\text{CF,est}} \leq 20$ km/s (voir détails un peu plus haut) et $\text{SNR}^{\alpha} > 15$ qui seront prises en compte afin de minimiser les erreurs sur les distances.

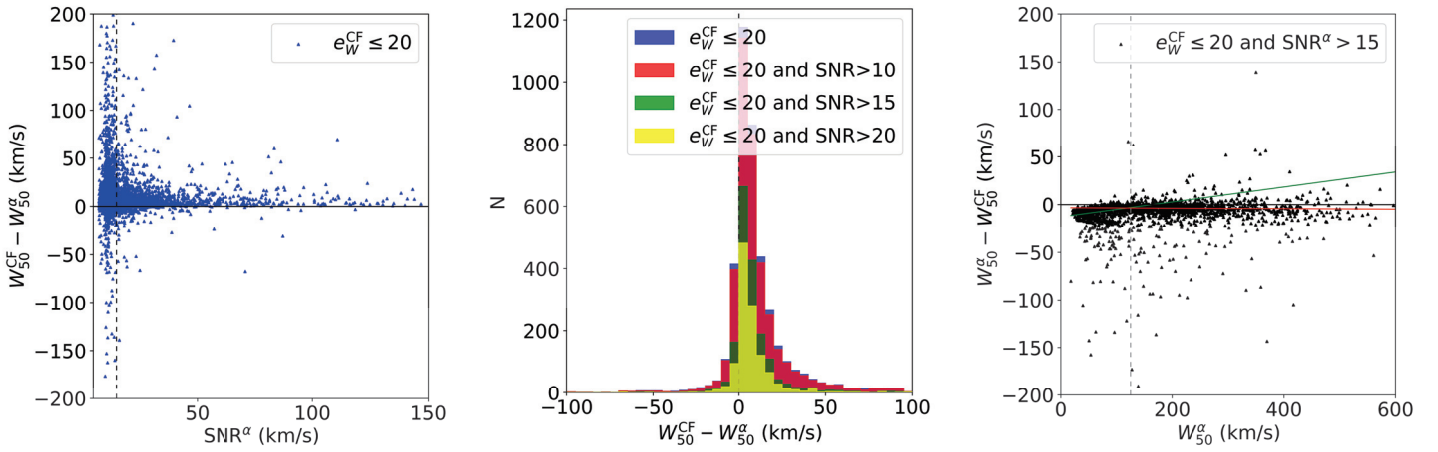


FIGURE 3.11 – Gauche : différence $W_{50}^{\text{CF}} - W_{50}^{\alpha}$ en fonction du signal sur bruit SNR^{α} . Milieu : histogramme de la différence $W_{50}^{\text{CF}} - W_{50}^{\alpha}$ pour quatre seuils (sans seuil en bleu, $\text{SNR} > 10$ en rouge, $\text{SNR} > 15$ en vert et $\text{SNR} > 20$ en jaune). Droite : comparaison des largeurs de raies mesurées par les deux collaborations pour les galaxies avec $e_W^{\text{CF,est}} \leq 20$ km/s et $\text{SNR}^{\alpha} > 15$. Les droites verte et rouge modélisent la distribution des points par une ligne brisée (voir équations 3.21 et 3.22).

Par ailleurs, les figures 3.11 gauche et milieu mettent en évidence dans la distribution de $W_{50}^{\text{CF}} - W_{50}^{\alpha}$ un léger décalage par rapport à zéro. Ce décalage est aussi présent sur la figure 3.11 droite, qui compare les largeurs de raies mesurées par les deux collaborations pour les galaxies avec $e_W^{\text{CF,est}} \leq 20$ km/s et $\text{SNR}^{\alpha} > 15$ en représentant la différence entre ces deux paramètres $W_{50}^{\text{CF}} - W_{50}^{\alpha}$ en fonction de W_{50}^{α} . Une cassure vers $W_{50}^{\alpha} \approx 150$ est observée. Afin de déterminer une relation liant W_{50}^{CF} à W_{50}^{α} , la distribution des points est modélisée par une ligne brisée définie par :

$$W_{50}^{\text{CF}} = 0.92W_{50}^{\alpha} + 12.9 \quad \text{si } W_{50}^{\alpha} \leq 150 \text{ km/s}, \quad (3.21)$$

$$W_{50}^{\text{CF}} = W_{50}^{\alpha} + 3.3 \quad \text{si } W_{50}^{\alpha} > 150 \text{ km/s}. \quad (3.22)$$

$$(3.23)$$

En utilisant le même principe que pour les erreurs sur les largeurs de raie HI, ces deux relations seront utilisées afin de déduire les valeurs des largeurs de raie ALFALFA dans la définition de la collaboration *Cosmicflows*.

Application au catalogue $\alpha 100$

Les équations 3.21 et 3.22 sont donc utilisées afin de déduire les valeurs des largeurs de raies du catalogue $\alpha 100$, si elles avaient été mesurées par la collaboration *Cosmicflows*. De même pour les erreurs sur les largeurs de raie à l'aide des équations 3.14, 3.15, 3.16 et 3.19. Pour le reste de cette section, les différents paramètres HI sont déduits de ces valeurs de largeurs de raie.

Pour la suite, les galaxies non adéquates pour l'application de la relation de Tully-Fisher sont exclues, i.e les galaxies pour lesquelles $e_W > 20$ km/s, $\text{SNR} < 15$ et $i < 45^\circ$. Au final, un échantillon contenant 3,121 galaxies ALFALFA adaptées pour mesurer leur distance par le relation TF est obtenu. Les galaxies du catalogue $\alpha 40$ sont exclues, ayant leur largeur de raie ayant déjà été mesurée par la collaboration *Cosmicflows*.

La figure 3.12 gauche montre la distribution en redshift cz des sources ALFALFA adéquates pour la mesure d'une distance, i.e les galaxies avec $e_W \leq 20$ km/s, $\text{SNR} > 15$ et $i > 45^\circ$ tel qu'il est décrit ci-dessus. Trois pics sont observés dans la distribution, vers 2,000 km/s, 5,000 km/s et 8,000 km/s. L'échantillon ALFALFA va permettre d'ajouter des galaxies lointaines à *Cosmicflows-4*, allant jusqu'à environ 17,000 km/s : 559 galaxies sont à une distance supérieure à 80,000 km/s. La figure 3.12 droite représente l'histogramme des largeurs de raie ALFALFA (définition ADHI) pour ces mêmes galaxies. Tous les types de galaxies sont représentés dans cet échantillon, allant des petites galaxies avec une petite vitesse de rotation (i.e W_{50} petit), aux galaxies plus massives tournant plus vite (grand W_{50} jusqu'à plus de 500 km/s).

La figure 3.13 gauche présente le logarithme du flux intégré en fonction du logarithme de la largeur de raie pour les galaxies telles que $e_W \leq 20$ km/s, $\text{SNR} > 15$ et $i > 45^\circ$. Cette figure montre qu'un sondage à l'aveugle tel que ALFALFA détecte majoritairement les galaxies à grandes largeurs de raie et flux important. Peu de petites galaxies (flux intégrés faibles et petites largeurs de raie) sont détectées. Ceci est dû au fait que le même temps de pose est utilisé pour tout le ciel, et donc pour toutes les galaxies, et n'est donc pas adapté pour obtenir des profils HI adéquats pour chaque galaxie. Par ailleurs, cette figure permet de mettre en évidence que les sondages à l'aveugle ne détectent pas les galaxies à forte rotation et faible flux.

La figure 3.13 droite montre la distribution de la masse en HI en fonction de la distance pour les galaxies adéquates pour l'application de la relation de Tully-Fisher. Les galaxies avec une masse HI faible sont observées pour les petites distances seulement. Aux grandes distances, seules les galaxies massives sont détectées.

Le catalogue ADHI disponible sur la base de données EDD comptait jusqu'à maintenant 4,427 spectres provenant des données ALFALFA adéquats pour la mesure d'une distance par la relation TF. L'analyse du catalogue complet des données ALFALFA $\alpha 100$ permet ainsi d'ajouter à ADHI 3,121 nouvelles largeurs de raie dont l'erreur est assez faible pour la mesure d'une distance.

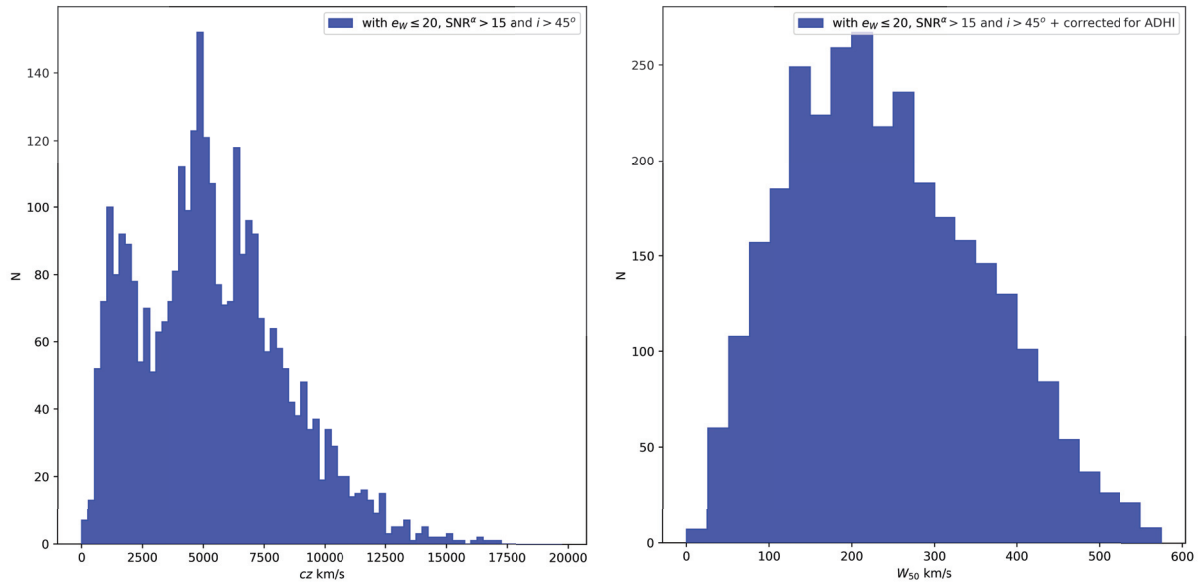


FIGURE 3.12 – Gauche : distribution en redshift cz des 3,121 sources ALFALFA adéquates pour la mesure d’une distance, i.e les galaxies avec $e_W \leq 20$ km/s, $\text{SNR} > 15$ et $i > 45^\circ$. Droite : histogramme des largeurs de raie ALFALFA (définition *Cosmicflows*) pour ces mêmes galaxies.

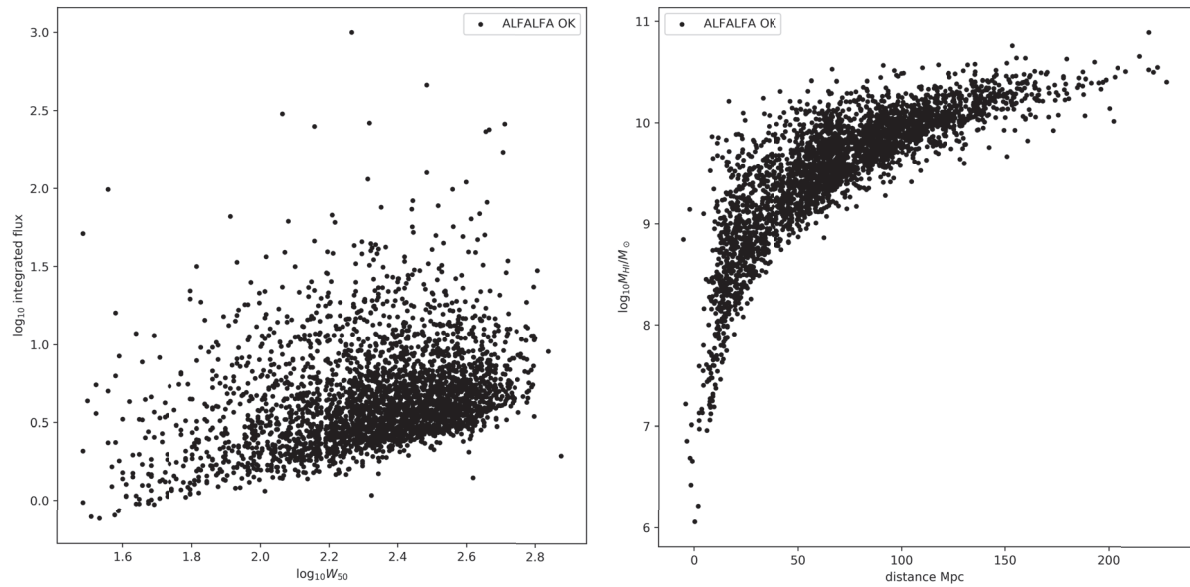


FIGURE 3.13 – Gauche : logarithme du flux intégré en fonction du logarithme de la largeur de raie pour les galaxies ALFALFA telles que $e_W \leq 20$ km/s, $\text{SNR} > 15$ et $i > 45^\circ$. Droite : distribution de la masse en HI en fonction de la distance pour ces mêmes galaxies.

3.4 Projet GBT18A-021 : Cosmic flows in the Local Universe

La collaboration a fait une demande de temps de télescope au GBT fin 2017. Cette demande a été acceptée et 518 heures d’observations ont été attribuées pour le semestre 18A, de février à juillet 2018. Un objectif de 500 galaxies a été fixé (environ 1 heure de pose par galaxie en moyenne). Les observations ont commencé dès le mois de décembre. Cela a donc apporté jusqu’en février de nombreuses heures supplémentaires d’observations (25 sessions avant le 1er février, 67 heures de télescope).

Il est obligatoire pour les nouveaux observateurs, dont je faisais partie, de participer à l’école d’observations du GBT, qui a eu lieu en septembre 2017, durant une semaine, à Green Bank, USA. Les participants ont appris à utiliser un radiotélescope, notamment le GBT : cours sur la radio astronomie, sur l’utilisation du GBT, etc.

Les différents modes d’observation possibles ont été présentés :

- le mode *position-switch* : un scan sur la source suivi d’un scan légèrement décalé, sur le ciel, à la même fréquence sont effectués ;
- le mode *frequency switch* : au lieu de faire varier la position ciblée par le télescope, on peut faire varier la fréquence. Un scan à la fréquence correspondant à la vitesse de la galaxie suivi d’un scan à une fréquence plus ou moins élevée (où la galaxie n’apparaît pas). Ce mode d’observation est plus efficace la nuit (les observations conduites le jour étant contaminées par le Soleil).

Durant les premières observations effectuées au GBT dans le cadre du projet GBT18A-021, les deux modes d’observation ont été testés. Cependant, le mode *frequency switch* s’est révélé inefficace.

3.4.1 Sélection des sources HI

Préparation du catalogue de sources HI

Les cibles HI ont été extraites d’une vaste compilation de galaxies pouvant être potentiellement observées au GBT. Au départ, cette compilation est composée de deux échantillons :

- échantillon de galaxies dites *plates*, en réalité vues par la tranche, provenant du *Revised Flat Galaxy Catalogue* (RFGC, [Karachentsev et al., 1999](#)) couvrant tout le ciel. Cet échantillon a été coupé à $\delta > 36^\circ$ afin de ne pas couvrir la même zone que le sondage ALFALFA avec Arecibo ;
- échantillon en direction du Repousseur du Dipole. Les galaxies de cet échantillon ont été sélectionnées avec les critères suivants : $20h < \alpha < 24h$ pour l’ascension droite, $16^\circ < \delta < 65^\circ$ pour la déclinaison et $3,000 < cz < 12,000$ km/s en ce qui concerne l’éloignement. De plus, seules les galaxies spirales avec une inclinaison supérieure à 50° ont été considérées.

Afin de ne pas ré-observer des galaxies dont les mesures de largeur de raie sont déjà utilisables pour appliquer la TF, les galaxies ayant une erreur $e_W \leq 20$ km/s dans ADHI ne sont pas prises en compte. Par ailleurs, sont aussi laissées de côté les galaxies avec les codes d'erreur $e_W = 500$ km/s (pas détectées) et $e_W = 100$ km/s (cas confus). Toutefois, le catalogue dans lequel les galaxies à observer sont piochées est beaucoup trop réduit. Il reste trop peu de galaxies après inspection (dont les critères seront discutés un peu plus bas) : 208 galaxies extraites de l'échantillon de galaxies plates, et seulement 7 galaxies tirées de l'échantillon du DR.

De manière à avoir plus de galaxies adéquates aux observations HI, le catalogue de départ a été agrandi en ajoutant deux nouveaux échantillons couvrant tout le ciel plus uniformément en dehors de la zone couverte par ALFALFA :

- premier échantillon sélectionné depuis LEDA. Il contient des galaxies spirales sélectionnées avec les critères suivants : $\delta > 36^\circ$, $3,000 < cz < 9,000$ km/s, inclinaison supérieure à 55° . Tout comme les échantillons précédents, les galaxies ayant déjà un spectre adéquat sur ADHI, ainsi que les galaxies non détectées et les cas confus ont été enlevés. Cette nouvelle sélection a apporté 2608 nouvelles galaxies au catalogue de départ ;
- deuxième compilation extraite de LEDA. Les contraintes sont cette fois relâchées pour obtenir davantage de galaxies. En effet, les galaxies doivent maintenant remplir les conditions suivantes : $\delta > 38^\circ$, $3,000 \leq cz \leq 10,000$ km/s, $40^\circ \leq i \leq 85^\circ$ où i est l'inclinaison. À l'égal des autres échantillons, les galaxies ne doivent pas avoir de mesure de largeur de raie avec une erreur $e_W \leq 20$ km/s, $e_W = 100$ km/s ou $e_W = 500$ km/s. Cet échantillon enrichit lui aussi le catalogue de départ de 2595 nouvelles galaxies ;
- les galaxies qui ont des spectres GB300 ou Nançay presque satisfaisants mais qui sont à refaire pour pouvoir utiliser la relation de TF sont aussi ajoutées. Ces galaxies sont au nombre de 4.

Au final, le catalogue de cibles HI potentielles contient 5420 galaxies. Les deux derniers échantillons ont fait grossir considérablement le catalogue. Cela demande plus de temps pour inspecter les galaxies une par une, afin de vérifier si elles sont des cibles adéquates pour les observations HI et pour la mesure de distance avec la TF.

Cependant, nous verrons par la suite qu'il y a peu de galaxies adéquates pour les observations à 21 cm. Il faut donc partir d'un nombre important de galaxies afin d'avoir au final un nombre suffisant de galaxies à observer et adéquates pour la mesure d'une distance.

Inspection des galaxies avant observation

Les galaxies sont en effet inspectées une par une par les membres de la collaboration. Une galaxie qui peut être observée au GBT dans le cadre du projet *Cosmicflows* doit remplir **toutes** les conditions décrites ci-dessous :

1. La cible doit tout d'abord être une galaxie spirale contenant un disque de gaz HI. Les galaxies elliptiques et les galaxies spirales susceptibles de ne pas avoir de disque de gaz ne sont pas observées. Le disque d'hydrogène n'étant pas visible à l'oeil nu,

il faut deviner la présence de gaz à partir des images PanSTARRS dans différentes bandes, g , r , i , z et y . En effet, les galaxies plus brillantes dans le bleu sont preuves de galaxies comportant des étoiles jeunes, et donc de la présence très probable d'un disque d'hydrogène. Alors que les galaxies brillantes dans le rouge (orange) contiennent seulement de vieilles étoiles : il n'y a plus de formation stellaire, qui est le signe qu'il n'y a probablement plus d'hydrogène disponible.

Deux exemples sont donnés sur la figure 3.14. La figure 3.14 (a) présente des images optiques provenant de Pan-STARRS de la galaxie PGC 137148. L'image la plus à gauche est une composite des bandes g , i et y . La galaxie y apparaît orange. Les bandes des autres images sont, de gauche à droite, g , r , i , z et y , soit de bleu à rouge. Cette galaxie est très brillante dans la bande y , i.e dans le rouge, et de moins en moins brillante vers le bleu. Cela signifie qu'elle est composée d'une vieille population stellaire et qu'elle contient peu de jeunes étoiles. Cette galaxie ne doit pas être observée. Le cas inverse est montré sur la figure 3.14 (b) pour la galaxie PGC 55355. L'organisation de cette planche est identique à la planche (a). Dans ce cas, cette galaxie apparaît bleue dans l'image composite. De plus, elle est peu brillante dans le rouge (bande y) où seul son bulbe, composé de vieilles étoiles, est visible. Le disque de cette galaxie est cependant de plus en plus brillant et détaillé vers le bleu : il est donc composé majoritairement de jeunes étoiles. Il est alors probable que cette galaxie contient du gaz HI. Cette galaxie peut donc être observée.

Il est aussi possible de regarder les spectres optiques des sondages SDSS (DR12 Alam et al., 2015) ou LAMOST (Cui et al., 2012), si disponibles sur LEDA² (Makarov et al., 2014). La figure 3.15 présente deux spectres provenant des données SDSS DR12. À gauche, il y a un spectre de la galaxie PGC 55355 qui a été détectée en HI au GBT durant les sessions d'observations. Le spectre SDSS de cette galaxie indique une raie d'émission $H\alpha$, qui est la trace de la présence d'étoiles jeunes dans cette galaxie. Cela indique qu'il est probable que cette galaxie contienne un disque de gaz d'hydrogène neutre. Le spectre de droite est le spectre associé à la galaxie PGC 12596, qui n'a pas été sélectionnée comme source HI pour observations au GBT car elle ne remplissait pas toutes les conditions. Ce spectre correspond en effet au spectre typique d'une galaxie elliptique : avec un continuum en pente et des raies d'absorption telles que Na, Mg et Ca qui mettent en évidence une population stellaire vieille. C'est une galaxie qui a donc trop peu de gaz d'hydrogène.

Enfin, si la galaxie a déjà été observée et détectée à 21 cm par une autre collaboration, cela signifie qu'elle contient en effet un disque de gaz d'hydrogène. On peut donc la ré-observer pour intégrer une nouvelle mesure de distance à ADHI. On peut aussi s'aider des mesures de magnitude à 21 cm (disponibles sur LEDA pour quelques galaxies observées en HI) : on observe les galaxies pour lesquelles $m_{21} < 17$.

2. par ailleurs aucune collision de galaxies n'est observée, ou des galaxies en interaction perturbées par des galaxies voisines. La figure 3.16 montre une image en lumière optique (gauche) et radio (droite) du triplet de galaxies M81, M82 et NGC 3077. On observe dans l'image radio des filaments d'hydrogène reliant les trois galaxies, mettant en évidence des transferts de gaz d'une galaxie à l'autre. Ces *boucles de marée* d'hydrogène, baptisées *Arp-Loop*, résultent d'interactions entre ces trois galaxies, et rendent difficile l'estimation de la vitesse de rotation des galaxies à partir de leur spectre HI. Ces filaments sont cependant invisibles dans l'image optique, où les trois galaxies ne semblent pas perturbées. À droite de la figure 3.16 sont présentées les images composites Pan-STARRS de trois galaxies du catalogue de sources HI poten-

2. <http://leda.univ-lyon1.fr/>

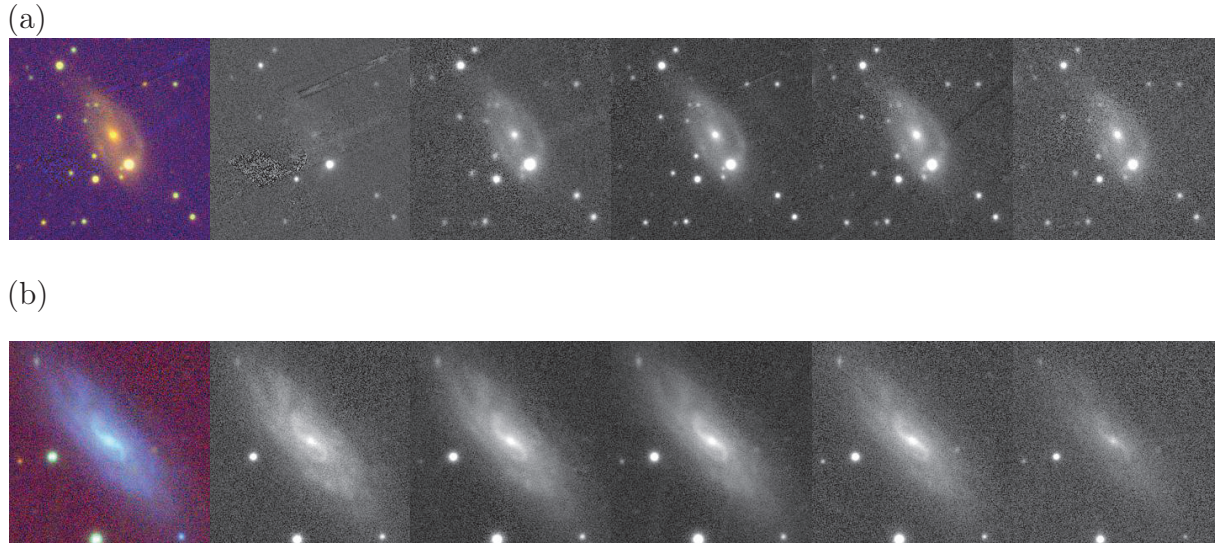


FIGURE 3.14 – Images optiques provenant de Pan-STARRS des galaxies (a) PGC 137148 et (b) PGC 55355. L'image la plus à gauche est une composite des bandes g, i et y. Les bandes des autres images sont, de gauche à droite, g, r, i, z et y, soit de bleu à rouge.

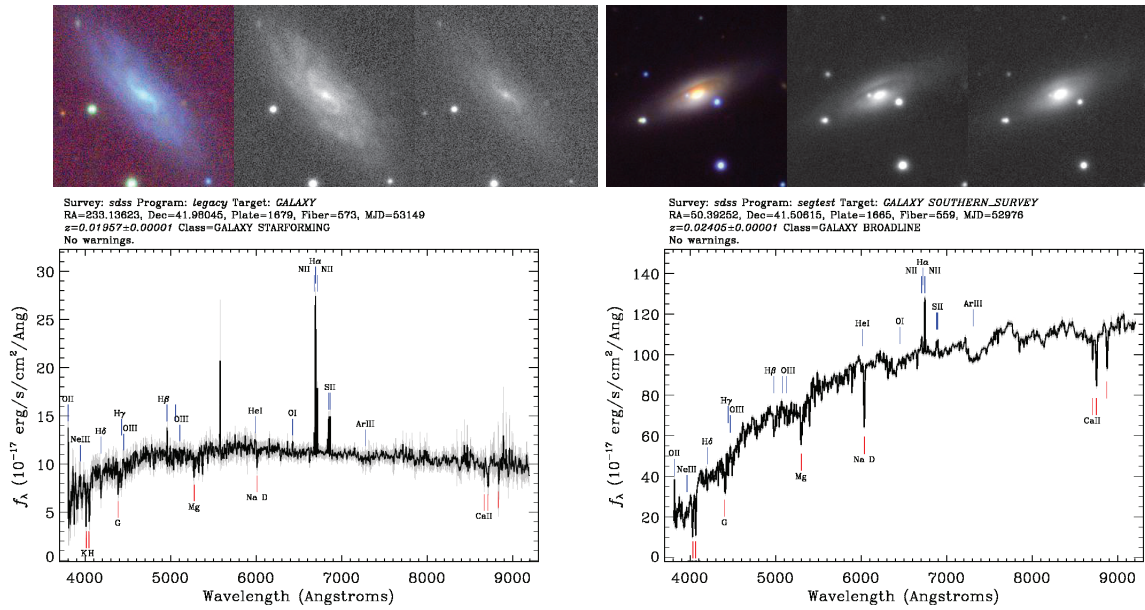


FIGURE 3.15 – Spectres optiques provenant des données SDSS DR12 des galaxies (a) PGC 55355 et (b) PGC 12596. La composite des bandes g, i et y, et les images dans les bandes g et y, provenant de Pan-STARRS, pour ces deux galaxies sont données au dessus de chaque spectre.

tielles pour observation au GBT. Ces galaxies étant en collision ou en interaction, elles n'ont pas été observées ;

- il ne doit pas y avoir de galaxies voisines ayant une vitesse de récession similaire à celle de la galaxie cible dans le radiolobe du GBT dont le diamètre est de 9 arcmin.

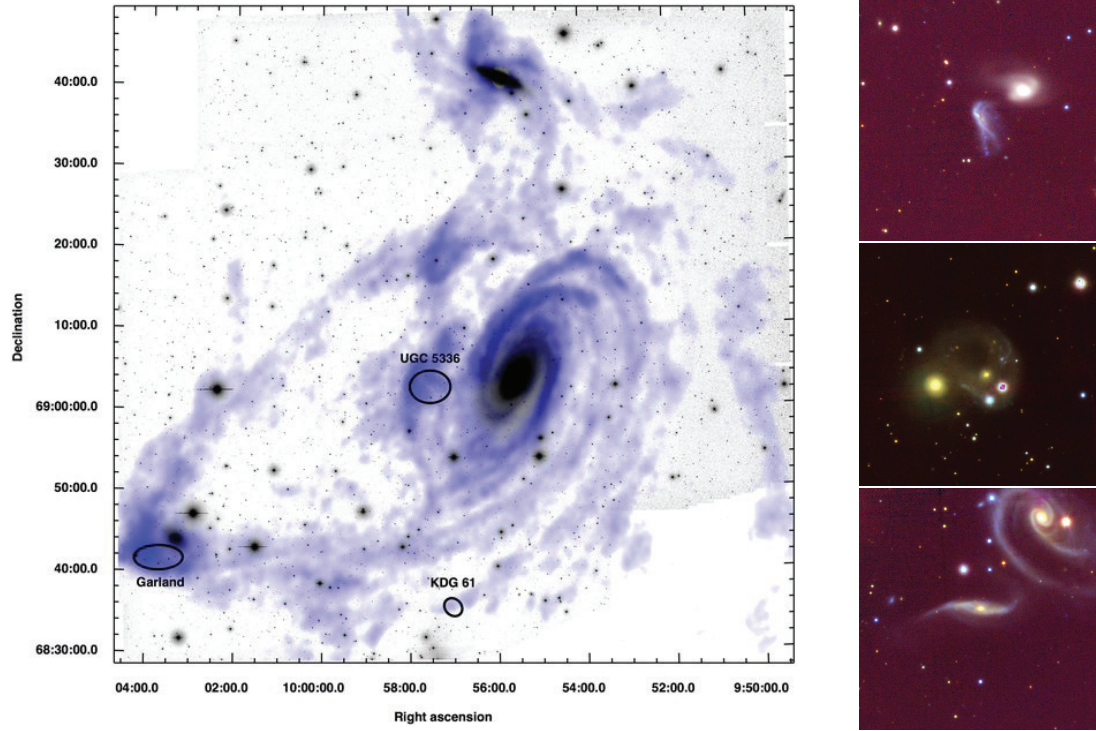


FIGURE 3.16 – Image HI du groupe de galaxies *M81*, *M82* et *NGC 3077* superposée en bleu sur une image optique en bande *i*. Figure extraite de Yun et al. (1994); Croxall et al. (2009). À droite de la figure 3.16 sont présentées les images composites Pan-STARRS de trois galaxies du catalogue de sources HI potentielles pour observation au GBT : de haut en bas, PGC 93089, PGC 9060 et PGC 8970.

Par ailleurs, il faut prendre en compte les lobes secondaires du radiotélescope, qui peuvent détecter du signal dans un diamètre de 15 arcmin environ. Cela entraîne une confusion du spectre HI : les raies HI de galaxies voisines situées à des redshifts similaires sont très proches et souvent confondues. Il devient donc impossible d'identifier deux raies HI et de les associer aux galaxies respectives. Afin d'éviter toute confusion, les galaxies situées dans le radiolobe doivent avoir une différence en redshift d'au moins 1000 km/s.

La figure 3.17 montre une image PanSTARRS dans la bande *g* de la galaxie spirale *PGC 922* (ou *IC 1535*), qui a une vitesse de récession de 5231 km/s. La taille de cette image est environ de 10 arcmin de côté (625 arcsec). Les traits en pointillés jaunes correspondent aux coordonnées de la galaxie, soit la position de pointage du GBT. Le cercle jaune représente la taille du radiolobe du télescope, et son diamètre est ainsi de 9 arcmin. Deux autres galaxies sont situées dans le radiolobe : *IC 1534* située à 2 arcmin ayant une vitesse de 5224 km/s, et *IC 1536* située à 4 arcmin et avec une vitesse de 5106 km/s. Si ces trois galaxies contiennent chacune un disque d'hydrogène neutre, ces deux galaxies ayant un redshift similaire à la galaxie cible, les trois spectres HI des ces galaxies seront donc confondus, et la galaxie *PGC 922* ne doit pas être observée. Cependant, si les galaxies voisines ne contiennent pas de disque de gaz (par exemple si ce sont des galaxies elliptiques ou oranges, voir détails ci-dessus), elles ne donneront rien à 21 cm et la galaxie *PGC 922* peut donc être observée. Il faut ainsi bien inspecter les galaxies voisines afin de décider si une

galaxie peut être observée.

Il faut toutefois faire attention aux galaxies situées en dehors du radiolobe. En effet, une quatrième galaxie est située juste à la limite du radiolobe : NGC 0048. Cette galaxie peut tout de même être détectée par les lobes secondaires du radiotélescope. La vitesse de cette galaxie étant de 1776 km/s, elle ne pose pas de problèmes puisqu'elle est beaucoup plus proche de l'observateur et il n'y a pas de risque de confusion. Il faut ainsi vérifier qu'il n'y ait pas de galaxies voisines situées dans un diamètre de 15 arcmin pour toutes les galaxies du catalogue afin de prendre en compte la taille du radiolobe ainsi que les lobes secondaires ;

4. il ne faut pas qu'il y ait d'étoiles gênantes pour la photométrie (par exemple sur un bras) qui pourrait rendre l'estimation de la magnitude apparente de la galaxie difficile ;
5. la galaxie ne doit pas être vue de face afin de limiter les erreurs sur la mesure de distance (voir section 3.1.1) ;
6. enfin, toutes les galaxies avec une forme ou structure anormale, et celles dont le type ou l'inclinaison sont difficiles à déterminer, sont rejetées, afin de limiter les erreurs, les non-détections et ainsi les pertes de temps.

Si toutes les conditions sont remplies, la galaxie peut être observée. Des erreurs sont tout de même inévitables, et cela donne donc lieu à des abandons : cas de non-détections et confusion. Certaines erreurs peuvent être évitées en faisant plus attention. Cependant, il y a des pièges ! Certaines galaxies remplissent toutes ces conditions mais ne sont pas ou peu détectées au GBT.

3.4.2 Observations à 21 cm au GBT

Le radiotélescope de Green Bank, ou *Robert C. Byrd Green Bank Telescope* (GBT) est un radiotélescope à antenne unique de 100 mètres de diamètre. Il est situé dans la ville de Green Bank, dans l'état de Virginie de l'Ouest aux États-Unis. Sa couverture en déclinaison va de -46.5° à 90° . Le rayon du radiolobe dans le ciel est de 9 arcmin.

La configuration du télescope pour les observations en HI est la suivante :

- le récepteur en bande L noté *Rcvr1_2* a été utilisé. Ce récepteur peut recueillir des fréquences allant de 1.15 GHz à 1.73 GHz ;
- le spectromètre VEGAS a été utilisé avec 4096 canaux. Il correspond au système d'acquisition de données associé aux raies spectrales au GBT ;
- les galaxies ont été observées dans une largeur de bande de 16.9 MHz, ce qui permet d'atteindre une superbement fine résolution spectrale de 0.9 km/s. La résolution finale (après lissage) est de 3.6 km/s (voir plus loin) ;
- le temps d'intégration est fixé à 30 secondes. De plus, afin de soustraire le ciel, il faut avoir recours aux observations en mode *position-switch* : à chaque session d'observation, deux scans pour chaque source HI sont effectués, i.e un scan ON (sur la source) et un scan OFF (sur le ciel). La stratégie d'observation sera discutée plus en détail ci-dessous ;
- enfin, les spectres ont été enregistrés en utilisant la convention optique pour convertir la fréquence en vitesse.

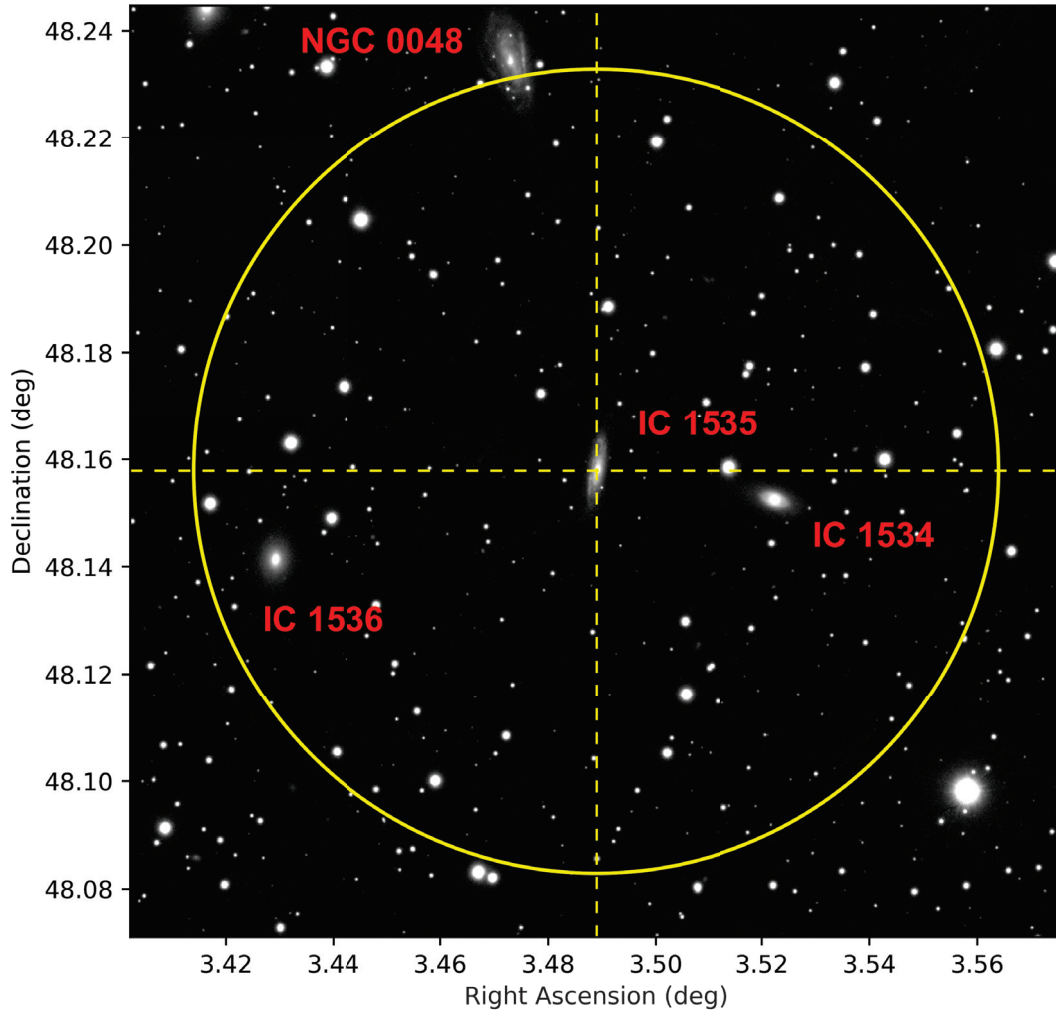


FIGURE 3.17 – Image PanSTARRS dans la bande g de la galaxie spirale *PGC 922* (ou *IC 1535*). La taille de cette image est environ de 10 arcmin de côté (625 arcsec). Les traits en pointillés jaunes correspondent aux coordonnées de la galaxie, soit la position de pointage du GBT. Le cercle jaune représente la taille du radiolobe du télescope, et son diamètre est ainsi de 9 arcmin.

Stratégie d’observation

Le mode d’observation appelé *position-switch* a été utilisé de façon à soustraire le ciel à chaque scan sur une source HI. En effet à chaque session d’observation, deux scans pour chaque source sont effectués. Le premier scan est un scan ON et est effectué en pointant sur la galaxie. Il faut ensuite procéder à un scan OFF, sur le ciel, en décalant le télescope de 1° en déclinaison environ. Le temps de pose pour chaque scan (ON ou OFF) est habituellement de 5 minutes. Chaque scan contient donc 10 intégrations d’une durée

de 30 secondes.

Afin de gagner du temps, un nouveau mode d'observation a été élaboré : le mode ON-OFF-ON. Tout comme un scan ON-OFF, il s'agit d'abord d'effectuer un scan ON sur la source, suivi d'un scan OFF (décalé de 1° environ). Le scan OFF est ensuite suivi d'un second scan ON sur la même source HI. Le temps de pose de chaque scan est fixé à 5 minutes. Le temps de télescope utilisé pour une même source HI est donc 15 minutes dont 10 minutes passées sur la source et 5 minutes perdues sur le ciel. Ce gain de 5 minutes par source HI par session est considérable.

Les galaxies qui sont à observer, ont été observées, etc... sont identifiées par mots-clé dans le catalogue :

- au départ, toutes les galaxies ont le mot-clé **verify** qui leur est attribué. Les galaxies doivent être inspectées par les membres de la collaboration afin de décider si elle peut être observée ;
- le mot-clé **tbo** (*to be observed*) est assigné aux galaxies qui ont été inspectées et qui sont adéquates pour être observées en HI. Dans le cas contraire, pour les galaxies qui ne sont pas adaptées pour les observations à 21 cm, le mot-clé **no** est donné ;
- les galaxies qui ont été observées et dont le spectre est utilisable pour appliquer la TF ($e_W \leq 20\text{km/s}$) ont comme mot-clé **OK**. Celles pour qui il y a une détection mais la raie n'est pas assez marquée pour mesurer la largeur de raie se voient attribuer le mot-clé **continue**. Enfin pour les galaxies pour lesquelles il n'y a pas de détection, ou bien pour les cas confus, on donne le mot-clé **abandon** ;
- on donne le mot-clé **reduce** aux galaxies dont les données d'une session n'ont pas encore été réduites avant la session suivante afin de ne pas ré-observer des galaxies qui n'ont pas besoin davantage de temps de pose.

Avant une session d'observation, il faut préparer une liste de galaxies à observer. Il faut observer les sources qui sont levées pendant la session, donc selon le temps sidéral local (*Local sidereal Time*, LST) au moment de la session. Pour cela il y a deux solutions. On peut choisir les galaxies à la main : on prépare une liste des sources observables selon le LST et l'ascension droite des sources, que l'on intègre dans le code d'observation avant chaque session.

Il est aussi possible de le faire de façon automatique : en adaptant un script d'observation d'un autre projet GBT, un algorithme qui sélectionne automatiquement les galaxies à observer dans le catalogue de départ pendant une session d'observation a été développé. Cet algorithme dépend à la fois de la position du Soleil dans le ciel, de la position pointée par le télescope et des statuts mots-clé qui ont été attribués aux galaxies du catalogue de départ et détaillés ci-dessus. L'algorithme liste tout d'abord les galaxies qui peuvent être observées à un moment donné, et qui remplissent ainsi les conditions suivantes : une déclinaison comprise entre $0^\circ < \delta < 85^\circ$, une séparation angulaire par rapport au Soleil supérieure à 10° (les scans trop près du Soleil sont de mauvaise qualité), le statut de la source dans le catalogue doit correspondre à **tbo** ou **continue**. Ensuite, est choisie la galaxie remplissant ces conditions pour laquelle le temps de déplacement du GBT est le plus court en prenant en compte la vitesse de déplacement du télescope. La galaxie choisie est observée par scans ON-OFF-ON d'une durée totale habituelle de 15 minutes. Enfin, l'algorithme attribue temporairement le statut **reduce** aux galaxies observées pendant

une session afin qu'elles se soit pas ré-observées pendant cette même session. Le catalogue original n'est pas modifié.

Les étapes détaillées ci-dessus sont répétées tout au long de la session. À la fin du temps alloué, le script est interrompu (après le scan OFF pour éviter de perdre un scan ON).

Le temps de pose dépend de la source. On pose jusqu'à ce que le spectre soit utilisable pour estimer une distance avec la TF (i.e $e_W \leq 20\text{km/s}$). Si c'est le cas, le statut `tbo` ou `continue` est modifié en `OK`. Si on voit clairement qu'il n'y a pas de détection (ou autre, par exemple cas confus), on attribue le statut `abandon`. La qualité des spectres obtenus est inspectée après chaque session lors de la réduction des données expliquée un peu plus loin.

Planification dynamique des observations

Les sessions d'observations sont programmées tout au long du semestre 2018A. Le planning *dynamique* des observations est préparé par le GBT.

Les sessions d'observations au GBT ne sont pas planifiées dès le début pour tout le semestre. En effet, le planning d'une journée est publié 24 à 48 heures à l'avance. Les projets sont choisis en fonction de leur priorité, de la météo, et de la disponibilité des différents récepteurs. Une fois que le planning est prêt, les observateurs des différents projets dont une session d'observation a été attribuée sont avertis. Il leur est donc nécessaire d'être toujours prêts à observer dans les 48h. Il est possible d'indiquer des périodes de *blackout*, i.e d'indisponibilité. Dans ce cas, pendant ces périodes d'indisponibilité, le projet ne sera pas pris en compte pour planifier les sessions.

La figure 3.18 présente la planification des sessions d'observation entre le 29 avril 2018 au 9 juin 2018 pour le projet *Comisflows* au GBT. Les encadrés bleus désignent les sessions d'observations, et l'heure de début de chaque session est donné en gras. Les sessions d'observations du projet *Cosmicflows* ne sont pas planifiées chaque jour.

La figure 3.19 gauche montre la distribution du temps sidéral local (*Local Sidereal Time*, LST, à Greenbank) au début des sessions d'observation du projet *Cosmicflows* au GBT. La distribution n'est pas homogène. Certains temps sidéraux n'ont pas été attribués : entre 7h et 17h.

La figure 3.19 milieu présente un histogramme de la durée des sessions d'observation en heures. Les sessions d'observations sont en majorité assez courtes, entre une heure et trois heures. Cela ne permet d'observer qu'un petit nombre de galaxies, une dizaine de galaxies au maximum si l'on prend en compte le temps que prend le télescope pour changer de position de pointage. Cependant, quelques sessions longues, pouvant durer jusqu'à huit heures, ont été attribuées au projet *Cosmicflows*, autorisant l'observation d'un plus grand nombre de galaxies (entre 20 et 30 galaxies).

La figure 3.19 droite présente deux histogrammes superposés : un histogramme du nombre de sessions en vert, et un histogramme du nombre total d'heures d'observation en bleu, selon les mois. Le trait rouge indique le début du semestre 2018A. Au total, 104 sessions ont été planifiées et 316.5 heures d'observation ont été données au projet *Cosmicflows*, soit 61% du temps total attribué au projet (518h). Cependant, les heures *facturées* (soit 245.25h,

i.e sessions à partir de février) ne représentent que 47% du temps total attribué au projet. La figure 3.19 droite met en évidence le fait que la majorité des sessions d'observation pour le projet *Cosmicflows* ont été planifiées en Janvier, Mai, Juin et Juillet. Durant le semestre 2018A (Février-Juillet), 80% du temps d'observation a été planifié entre Mai et Juillet.

Par ailleurs, du fait de la météo à Greenbank durant les deux hivers précédents, peu d'observations à haute-fréquence ont été programmées. Par conséquent, en anticipation de l'hiver à venir, il a été proposé au programme *Cosmicflows* de continuer les observations au prochain semestre (2018B), en mode *filler* (ou *bouche-trou*). Ainsi, les observations supposées s'arrêter le 31 Juillet 2018 ont redémarré en Août 2018. Le temps d'observation restant au programme *Cosmicflows*, soit 272.75h, pourra donc être utilisé jusqu'à Février 2019.

May 2018

Sun	Mon	Tue	Wed	Thu	Fri	Sat
29	30	1	2	3	4	5
	17:30 Observing GBT18A-021 - 05	16:45 Observing GBT18A-021 - 05	18:15 Observing GBT18A-021 - 05	22:30 Observing GBT18A-021 - 03	21:30 Observing GBT18A-021 - 05	
6	7	8	9	10	11	12
		17:30 Observing GBT18A-021 - 05				
13	14	15	16	17	18	19
		12:45 Observing GBT18A-021 - 05	12:45 Observing GBT18A-021 - 04	14:15 Observing GBT18A-021 - 04		11:30 Observing GBT18A-021 - 05
20	21	22	23	24	25	26
18:30 Observing GBT18A-021 - 03			12:00 Observing GBT18A-021 - 05	12:30 Observing GBT18A-021 - 05		
27	28	29	30	31	1	2
8:00 Observing GBT18A-021 - 05		18:45 Observing GBT18A-021 - 03	11:45 Observing GBT18A-021 - 05 18:45 Observing GBT18A-021 - 03	9:45 Observing GBT18A-021 - 04 18:15 Observing GBT18A-021 - 03		
3	4	5	6	7	8	9
12:00 Observing GBT18A-021 - 04 18:00 Observing GBT18A-021 - 05	8:30 Observing GBT18A-021 - 05				12:00 Observing GBT18A-021 - 05	7:45 Observing GBT18A-021 - 05 17:30 Observing GBT18A-021 - 03

FIGURE 3.18 – Planification des sessions d'observation entre le 29 avril 2018 au 9 juin 2018 pour le projet *Comisflows* au GBT. Les encadrés bleus désignent les sessions d'observations, et l'heure de début de chaque session est donné en gras.

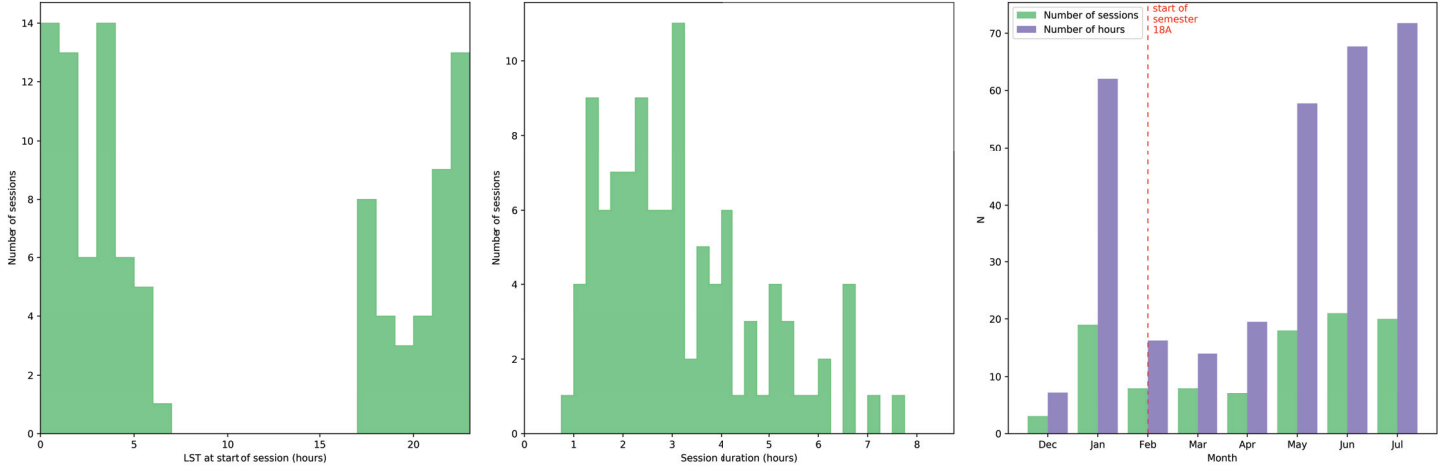


FIGURE 3.19 – Gauche : distribution du temps sidéral local (*Local Sidereal Time*, LST, à Greenbank) au début des sessions d’observation du projet *Cosmicflows* au GBT. Milieu : histogramme de la durée des sessions d’observation en heures. Droite : histogrammes du nombre de sessions (vert) et du nombre total de temps d’observation (en heures, bleu) selon les mois. Le trait pointillé rouge indique le début du semestre 2018A.

3.4.3 Observations au GBT : réduction des données

De l’observation d’une galaxie à la mesure de sa largeur de raie à 21 cm

Après chaque session, les données sont réduites afin de mettre à jour le catalogue d’observation pour la prochaine session et éventuellement modifier les statuts des galaxies observées. La réduction des données se fait avec GBTIDL (<http://gbtidl.nrao.edu/>) qui est le logiciel de réduction des données fourni par le GBT. Il est écrit en IDL et contient des fonctions diverses permettant d’analyser les données provenant du GBT. La méthode de réduction est détaillée ci-dessous.

Pour chaque session, paire de scans ON-OFF, polarisation et intégration, la procédure `getsigref` utilisée calibre le spectre total à partir de deux spectres *signal* (ON) et *reference* (OFF) avec l’équation suivante :

$$S = \frac{\text{signal} - \text{reference}}{\text{reference}} \times T_{\text{sys}} \times G \times e^{\tau}, \quad (3.24)$$

où $T_{\text{sys}} \approx 20$ K correspond à la température du système et est calculé par GBTIDL à partir du spectre *reference*, $G = 2$ K/Jy est le gain du GBT et $\tau = 0.008$ le paramètre de correction de l’opacité de l’atmosphère. Le spectre obtenu est ensuite inspecté afin de supprimer d’éventuelles interférences radio (RFI). Les RFI rencontrées au cours des observations ainsi que les méthodes utilisées pour tenter de supprimer les interférences sont décrites dans la section suivante 3.4.3.

Par la suite, les spectres calibrés pour toutes les intégrations, polarisations, paires ON-OFF et sessions d’une même galaxie sont sommés et moyennés puis lissés avec une fenêtre de Hann (procédure `hanning`) afin d’obtenir une résolution de 3.6 km/s. Enfin, la baseline

est modélisée par un polynôme de degré 3 (ou plus si besoin) puis soustraite au spectre afin d'obtenir une ligne de base plate et un niveau de flux à zero en dehors de la raie HI.

Après chaque réduction, le spectre final de chaque galaxie observée est inspecté. Si la raie HI est clairement visible et ses bords clairement définis, sa largeur W_{50} est mesurée avec la méthode décrite dans la section 3.1.1. Si l'on obtient une erreur sur W_{50} adéquate pour estimer une distance avec la relation de Tully-Fisher (i.e ≤ 20 km/s), on attribue le statut OK à la galaxie : l'observation de cette galaxie est terminée, on a suffisamment posé dessus. Si les bords de la raie HI sont mal définis, ou bien le ratio signal sur bruit n'est pas assez grand pour donner une erreur sur W_{50} inférieure ou égale à 20 km/s, il faut encore poser sur la galaxie, on lui attribue ainsi le statut **continue** dans le catalogue afin de l'observer dans une session prochaine. Enfin, si on constate que la raie HI est peu ou pas visible malgré un temps de pose considérable, on décide de ne plus poser sur la galaxie afin de ne pas perdre davantage de temps de télescope en lui attribuant le mot-clé **abandon**.

Interférences radio

Tous les spectres sont inspectés afin de supprimer les interférences radio (RFI). Deux types de RFI contaminant les spectres à 21 cm ont été observés.

D'une part, on peut observer de petites interférences radio correspondant à des pics très fins. Celles-ci sont remplacées par du bruit avec GBTIDL en interpolant le bruit avant et après le pic, afin d'éviter de perdre du temps d'intégration. Ce type d'interférences radio ne sont pas gênantes si elles sont en dehors de la raie HI. La figure 3.20 gauche illustre ce type d'interférence observé dans le spectre HI de la galaxie PGC 2481777.

D'autre part, une autre interférence est apparue, qui est cette fois beaucoup plus gênante. Elle est située à 1.381 MHz (soit 8,500 km/s). Elle est causée par un système GPS utilisé pour les tests nucléaires. Aucun filtre, configuration GBT, ou fonction GBTIDL n'est disponible pour supprimer cette interférence radio. Ce radar ruine la plupart des spectres HI des galaxies lointaines vers 8000–9000 km/s, qui sont celles qui sont les plus intéressantes pour améliorer CF3. Cependant, il peut apparaître dans les scans ON ou OFF, mais n'apparaît pas toujours dans toutes les intégrations. Dans l'attente d'une solution qui permettrait de sauver les spectres, les intégrations dans lesquelles le radar est présent sont exclues lors de la réduction des données. La figure 3.20 droite présente le spectre de la galaxie PGC 6177 située à une distance de 8,252 km/s. Le radar est présent autour de 8,500 km/s, et la raie HI n'est pas visible, masquée par le bruit causé par le radar.

La figure 3.21 montre les dix intégrations de 30 secondes qui sont sommées pour obtenir le spectre HI de PGC 6177 représenté sur la figure 3.20. Le radar est présent sur les intégrations 0, 1, 4, 5, 6, 7, 8, 9 ; soit 8 intégrations sur 10. Par conséquent, seulement deux intégrations sont utilisables. On perd donc dans ce cas 4 minutes de temps de pose, auxquelles il faut ajouter les 5 minutes perdues en posant sur le ciel lors du scan OFF.

Une tentative de suppression de cette RFI sur les intégrations concernées a été faite en utilisant la méthode présentée par [Hunt et al. \(2016\)](#), qui permet aux auteurs de supprimer les interférences à bande large (*wideband*) dans les intégrations contaminées par les RFI.

Tout d'abord, si la différence entre le flux maximal et le flux minimal du spectre d'une in-

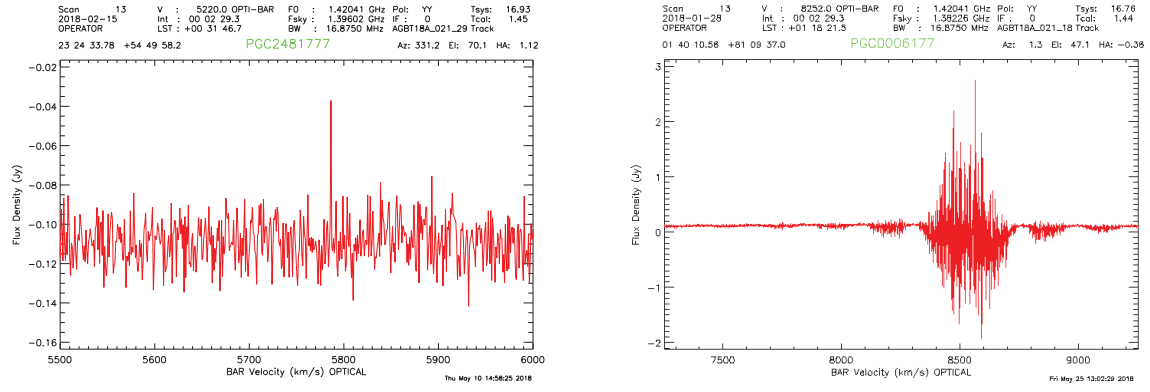


FIGURE 3.20 – Les deux types principaux d’interférences radio au GBT : à gauche une pic fin (*spike*), et à droite une émission large d’un radar ou d’un GPS.

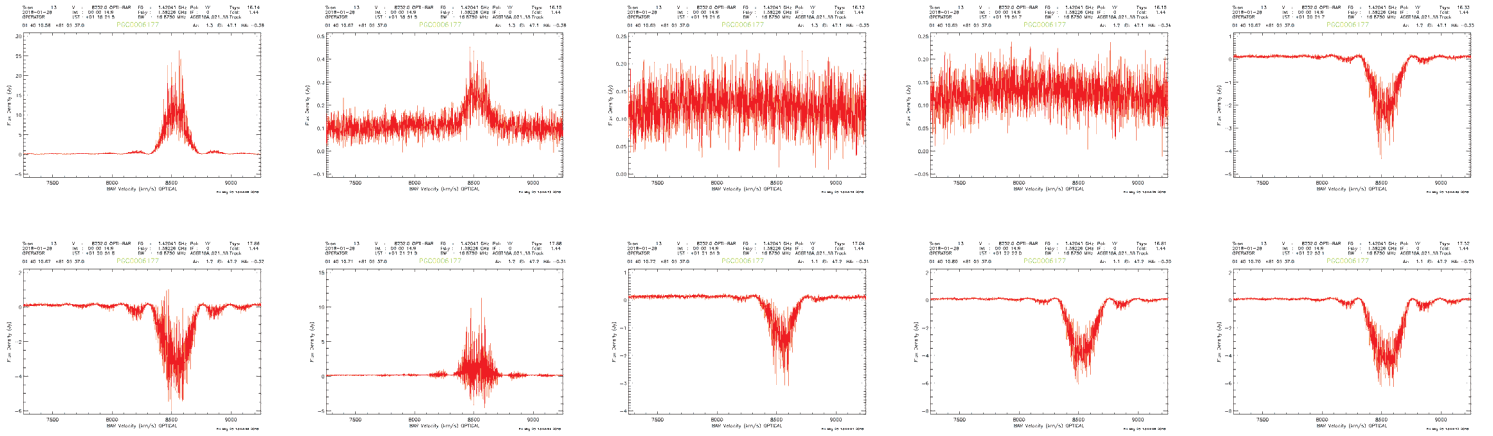


FIGURE 3.21 – Scan de 5 minutes composé de 10 intégrations de 30 secondes centré sur la galaxie PGC 6177.

tégration donnée dépasse un certain seuil, cela signifie que cette intégration est contaminée par le radar situé à 8500 km/s. Cela permet de détecter automatiquement les intégrations dans lesquelles le radar est présent. La méthode décrite ci-dessous est ensuite appliquée afin de supprimer cette RFI.

Ensuite, si le spectre est en effet contaminé par le radar, sa transformée de Fourier est tout d’abord calculée. Dans l’espace de Fourier, le radar se manifeste par de grands pics dans les premiers et derniers canaux. La transformée de Fourier d’un spectre non-contaminé par cette RFI ne contient pas de tels pics. La méthodologie Sigma Clipping est utilisée, afin de fixer les canaux dont le flux dépasse 4.3σ à zéro et de faire de même pour les huit canaux voisins, où σ correspond à la déviation standard typique d’un spectre non-contaminé par le radar. Ces deux valeurs ont été fixées par [Hunt et al. \(2016\)](#) après plusieurs tests sur différents spectres où les RFI à bande large sont présentes afin de déterminer la combinaison de valeurs permettant de les supprimer totalement. Le Sigma Clipping est effectué sur les parties réelles et imaginaires de la transformée de Fourier du

spectre considéré. Enfin, la transformée de Fourier inverse est calculée afin de revenir dans l'espace des fréquences où le radar est à présent absent.

[Hunt et al. \(2016\)](#) ont validé leur méthode avec des tests sur des observations GBT auxquelles ils ont rajouté des interférences à bande large artificielles, ont appliqué leur méthode de Sigma Clipping dans l'espace de Fourier et ont supprimé ces interférences. Cependant, les auteurs ont montré que cette méthode est efficace seulement lorsque la raie HI n'est pas visible dans une intégration et mélangée au bruit (mais devient visible lorsque l'on accumule plusieurs intégrations). En effet, une raie HI visible dans un spectre se traduit elle aussi par des pics dans l'espace de Fourier, ce qui peut induire une perte du signal *intéressant* lorsque l'on applique le Sigma Clipping. De plus, cette méthode est efficace seulement quand moins de 10% des intégrations sont contaminées par les RFI, ce qui n'est malheureusement, dans la plupart du temps, pas le cas : il faudrait que seulement 1 intégration sur 10 soit contaminée.

La figure 3.22 montre le spectre de la galaxie PGC 6177 obtenu en sommant les intégrations présentées en figure 3.21. À gauche, seules les intégrations dans lesquelles le radar n'apparaît pas ont été prises en compte. À droite, les intégrations dans lesquelles le radar apparaît ont été corrigées à l'aide de la méthode de [Hunt et al. \(2016\)](#) décrite auparavant. Cette figure montre que la méthode de Sigma Clipping dans l'espace de Fourier n'est pas efficace sur les observations du projet *Cosmicflows* : la raie HI est légèrement visible à gauche, mais elle n'est plus visible à droite, ce qui signifie malheureusement que le signal est perdu lors de l'application de la méthode de Sigma Clipping.

De telles interférences ont été observées par la collaboration NIBLES lors d'observations conduites au radiotélescope de Nançay. Selon [van Driel et al. \(2016a\)](#), des signaux de radars perturbent les observations vers 8,300 km/s. Leur amplitude est beaucoup plus importante que les signaux correspondant à la raie HI, ce qui rend la suppression de ces radars très difficile. La solution choisie par [van Driel et al. \(2016a\)](#) est elle aussi d'ignorer les intégrations contaminées par ces interférences radio.

De telles interférences représentent ainsi un problème touchant tous les télescopes, faisant perdre un temps précieux aux observateurs. Une étude plus approfondie du radar touchant le GBT est en cours, dans le but d'envoyer un rapport à ce sujet à l'observatoire de Greenbank.

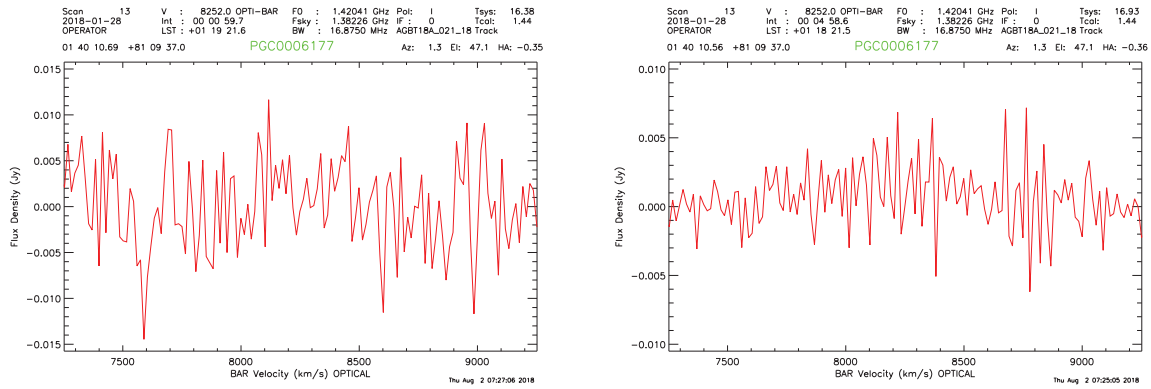


FIGURE 3.22 – Spectre de la galaxie PGC 6177 obtenu en sommant les intégrations présentées en figure 3.21. Gauche : seules les intégrations dans lesquelles le radar n’apparaît pas ont été prises en compte. Droite : les intégrations dans lesquelles le radar apparaît ont été corrigées à l’aide de la méthode de [Hunt et al. \(2016\)](#).

3.4.4 Premiers résultats

À ce jour, parmi le catalogue de départ contenant 5460 galaxies, il y a :

- 5 galaxies n’ont pas encore été inspectées (statut **verify**) ;
- 4,804 qui ne remplissaient pas toutes les conditions pour une observation en HI après inspection (statut **no**). Ainsi, les galaxies non observables en HI (statut **no**) représentent près de 90% du catalogue d’observation ;
- 194 galaxies prêtes à être observées lors d’une prochaine session d’observation (statut **tbo**) ;
- 259 sources HI qui ont été détectées (statut **OK**) ;
- 187 abandons (statut **abandon**) ;
- et enfin 11 galaxies qui requièrent du temps de pose additionnel afin d’obtenir un profil HI adéquat pour appliquer la TF (statut **continue**).

Toutes ces galaxies sont représentées sur une projection aitoff en coordonnées supergalactiques en Figure 3.23. Les galaxies refusées après inspection sont symbolisées par de petites croix grises. Les galaxies prêtes à être passées au GBT sont matérialisées par des carrés bleus, et les galaxies déjà passées au GBT mais exigeant davantage de temps d’intégration par des triangles jaunes. Enfin les galaxies dont l’observation HI est terminée sont représentées par des points rouges pour les cas non-détectés ou confus, et par des étoiles vertes pour les détections satisfaisantes.

À ce jour, les largeurs de raies ont été extraites de 209 spectres, dont 119 cas satisfaisants (**OK**) et 90 abandons (**abandons**). Les cas satisfaisants ont donné lieu à 109 largeurs de raies, soit 92 %, avec une erreur inférieure à 20 km/s, i.e adéquates pour la mesure d’une distance par la méthode TF. Parmi les abandons, 16 cas de confusion ont été identifiés et 72 sources HI n’ont pas été détectées.

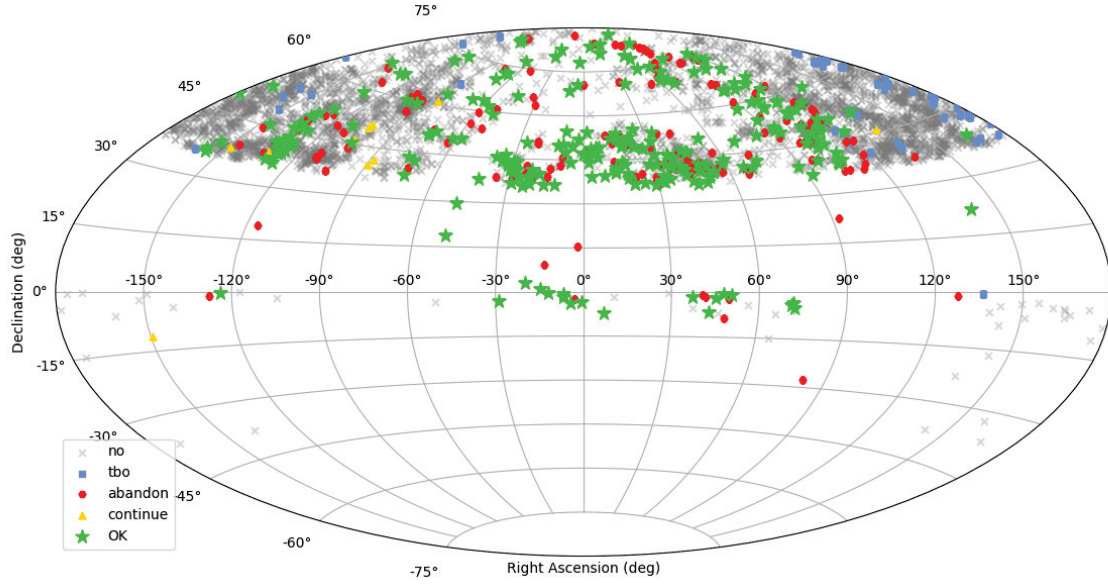


FIGURE 3.23 – map statuts catalogue GBT

Les résultats obtenus à partir de l'analyse de ces 209 spectres sont donnés dans la figure 3.24 sous forme d'histogrammes de la vitesse héliocentrique des sources HI, des largeurs de raie obtenues, du ratio signal sur bruit et du flux intégré dans la raie HI. Les cas satisfaisants sont représentés en vert, et les abandons en rouge.

Par ailleurs, la figure 3.25 gauche présente le logarithme du flux intégré en fonction du logarithme de la largeur de raie, et la figure 3.25 droite représente le logarithme de la masse HI en fonction de la distance en Mpc. Les 109 cas satisfaisants et adéquats pour la mesure d'une distance sont représentés en rouge. Pour comparaison, les points noirs représentent toutes les largeurs de raie HI provenant d'observations au GBT et disponibles dans le catalogue ADHI. Le panel de droite montre qu'à petite distance, les galaxies les moins massives aux plus massives sont détectées. Cependant, à grande distance, seules les galaxies les plus massives ne peuvent être détectées par le GBT.

Les paramètres HI mesurés sur ces 209 galaxies détectées au GBT (statut OK) sont listés dans la table A.1. Les colonnes correspondent à : (1) numéro PGC de la source HI ; (2) ascension droite J2000 exprimée en heures ; (3) déclinaison J2000 en degrés ; (4) vitesse radiale héliocentrique en km/s ; (5) flux intégré dans la raie HI en Jy.km/s ; (6) et (7) largeur de raie W_{50} et son erreur associée e_W en km/s ; (8) signal sur bruit ; (9) masse HI ; (10) magnitude 21. Par ailleurs, la planche B.1 présente pour chaque détection (soit 119) le profil HI et l'image multi-bande associée (y , i et g) extraite de PanSTARRS et de 1 arcmin de côté. Enfin, les spectres des 90 non-détectés ou cas confus sont donnés dans la planche C.1.

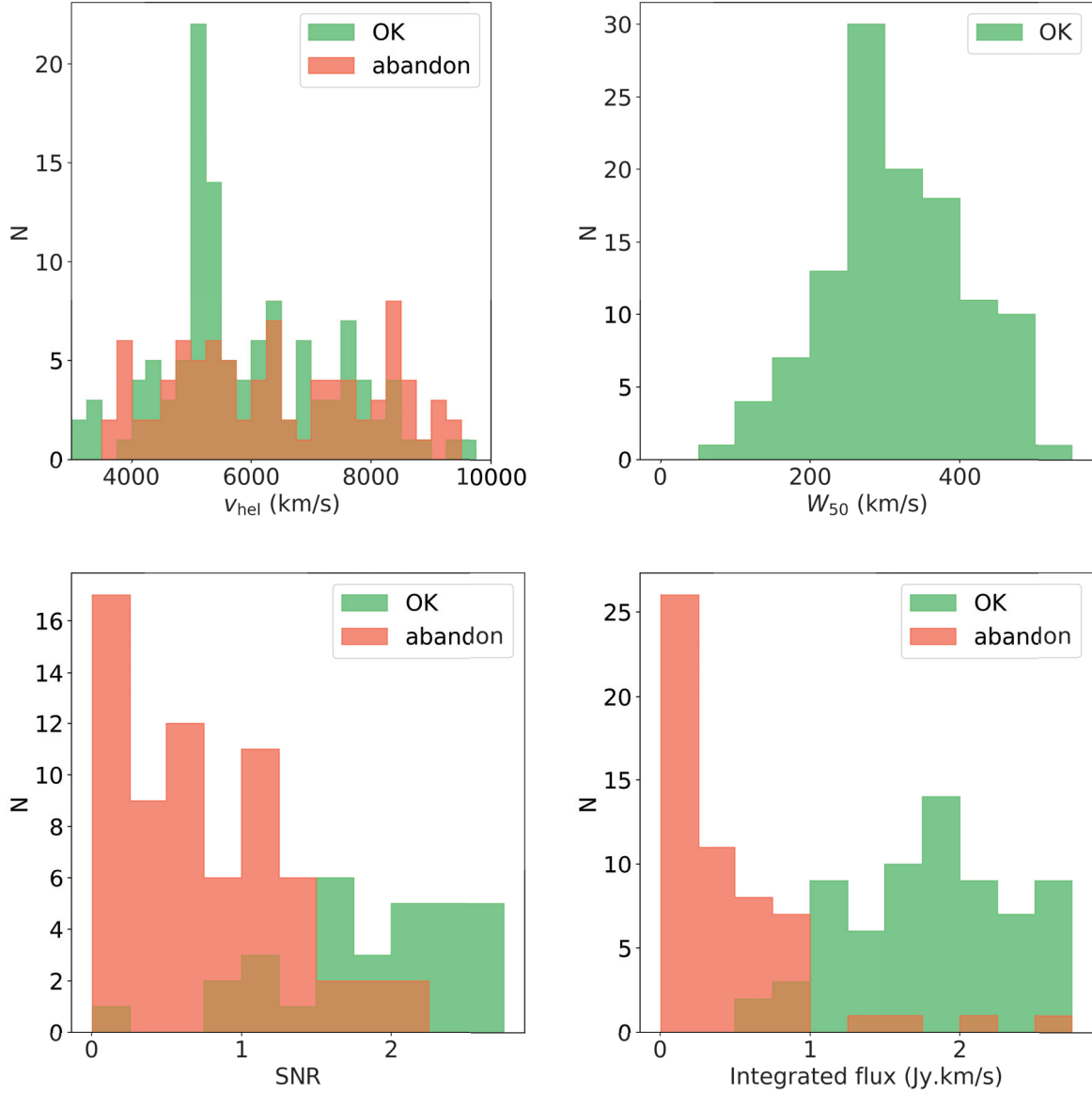


FIGURE 3.24 – De gauche à droite et de haut en bas : histogrammes de la vitesse héliocentrique des sources HI, des largeurs de raie obtenues, du ratio signal sur bruit et du flux intégré dans la raie HI. Les cas satisfaisants sont représentés en vert, et les abandons en rouge.

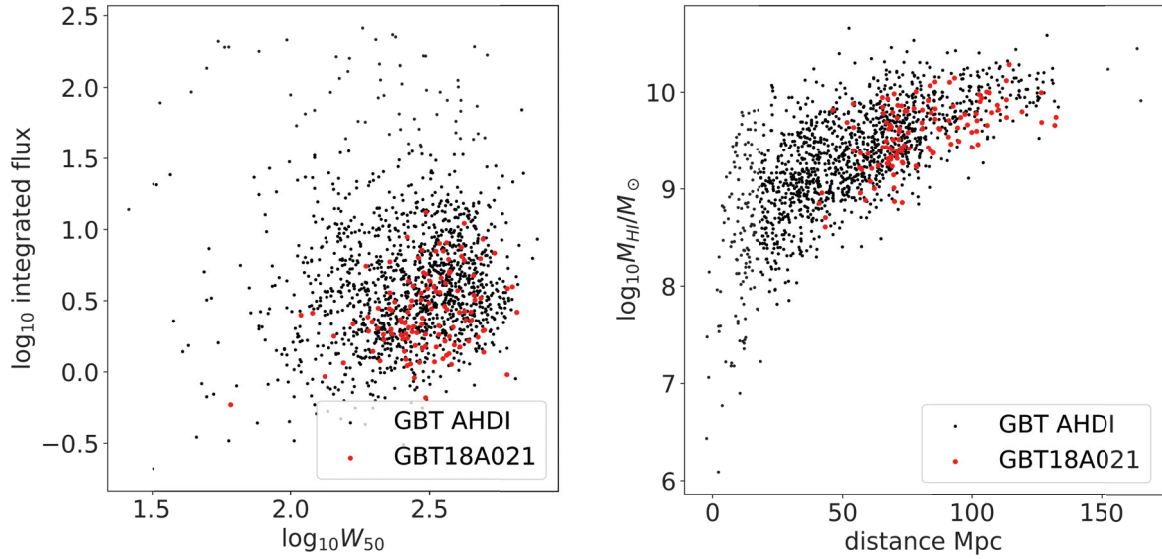


FIGURE 3.25 – Gauche : logarithme du flux intégré en fonction du logarithme de la largeur de raie. Droite : logarithme de la masse HI en fonction de la distance en Mpc. Les 109 cas satisfaisants et adéquats pour la mesure d’une distance sont représentés en rouge. Pour comparaison, les points noirs représentent toutes les largeurs de raie HI provenant d’observations au GBT et disponibles dans le catalogue ADHI.

3.5 Conclusion : données HI pour Cosmicflows-4

Les différentes analyses des données HI provenant de divers radiotélescopes ont permis d’ajouter au total 4,269 nouvelles galaxies, dont la largeur de raie est adéquate pour la mesure d’une distance, au catalogue ADHI disponible sur la base de données EDD : 889 sources HI détectées à Nançay (données KLUN17 et NIBLES), 3,121 galaxies extraites des données ALFALFA (si l’on exclu les sources HI provenant du catalogue α 40, étant déjà disponibles dans ADHI), et enfin, à ce jour, 259 détections satisfaisantes au GBT. Ces nouvelles compilations sont représentées sur une projection aitoff en coordonnées équatoriales sur la figure 3.26. Les nouvelles données provenant de Nançay sont données par des carrés rouges, celles provenant du télescope Arecibo (ALFALFA) par des points bleus et les sources HI détectées au GBT par des étoiles vertes. Ces ajouts couvrent en majorité l’hémisphère terrestre Nord, ce qui permet d’atteindre l’objectif qui a été fixé dans le chapitre 2 pour la construction du prochain catalogue *Cosmicflows-4*.

La figure 3.27 présente deux histogrammes du redshift des galaxies qui seront ajoutées à CF4 : à gauche sont comprises toutes les galaxies, tandis qu’à droite ne sont prises en compte que les galaxies situées dans l’hémisphère Nord. Le catalogue CF3 est représenté en vert, et la future compilation CF4 en vert. Les ajouts décrits ci-dessus sont représentés en bleu. Par ailleurs, la figure 3.28 présente une tranche SGX-SGY de 30 Mpc/h de largeur en SGZ, sur laquelle sont représentés les trois catalogues CF1, CF2 et CF3 en vert, bleu et rouge respectivement. Les nouvelles galaxies qui permettront de construire le catalogue CF4 sont présentées en jaune. Les cercles vides noirs correspondent aux sources HI qui vont être observées au GBT durant le prochain semestre 2018B. Les figures 3.27 et 3.28 montrent qu’un grand nombre de galaxies seront ajoutées dans l’hémisphère Nord,

permettant de combler en partie le manque de données dans cette région. Cependant, il existe toujours un manque de données lointaines dans le Nord, les nouvelles données ajoutées allant en majorité jusqu'à environ 8,000 km/s.

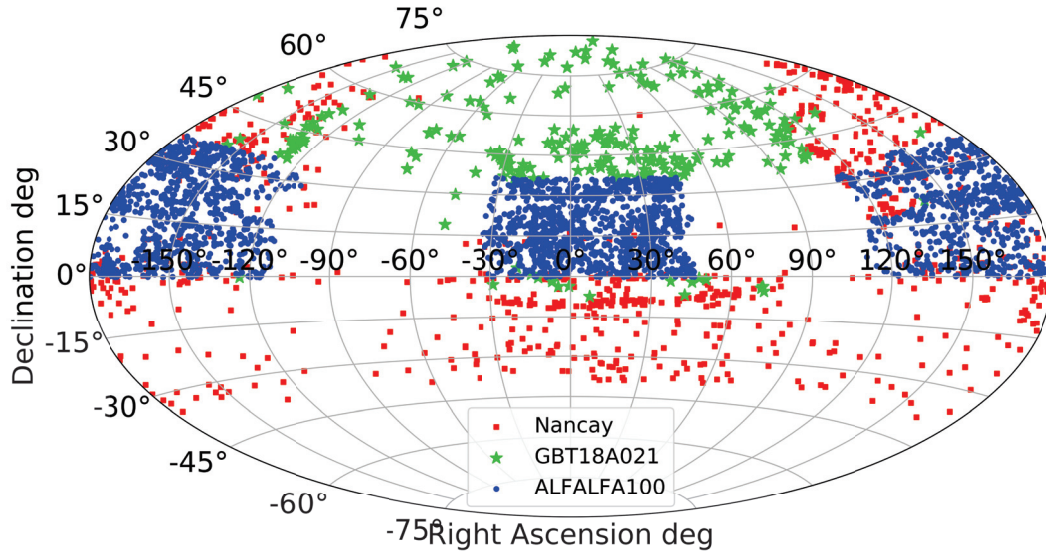


FIGURE 3.26 – Projection aitoff en coordonnées équatoriales sur laquelle sont représentées les nouvelles données ajoutées par les différentes analyses de données HI décrites dans ce chapitre : les données provenant de Nancay sont données par des carrés rouges, celles provenant du télescope Arecibo (ALFALFA) par des points bleus et les sources HI détectées au GBT par des étoiles vertes.

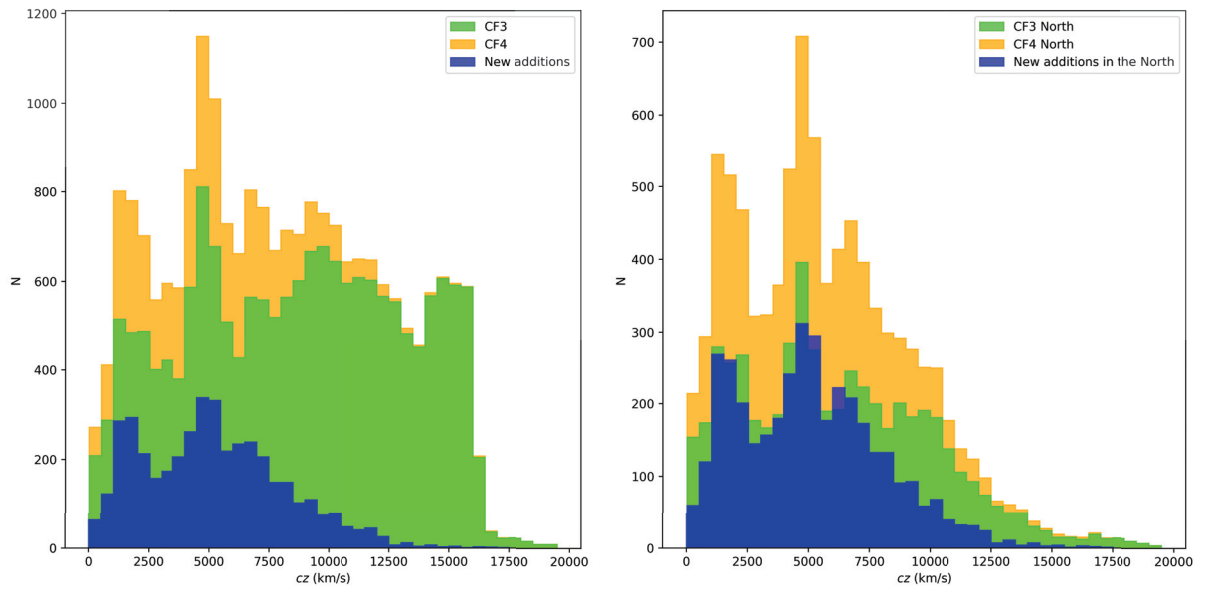


FIGURE 3.27 – Histogrammes du redshift cz des galaxies qui seront ajoutées à CF4. À gauche : toutes les galaxies. À droite : galaxies situées dans l’hémisphère Nord seulement. Le catalogue CF3 est représenté en vert, et la future compilation CF4 en vert. Les ajouts à CF4 décrit dans ce chapitre sont représentés en bleu.

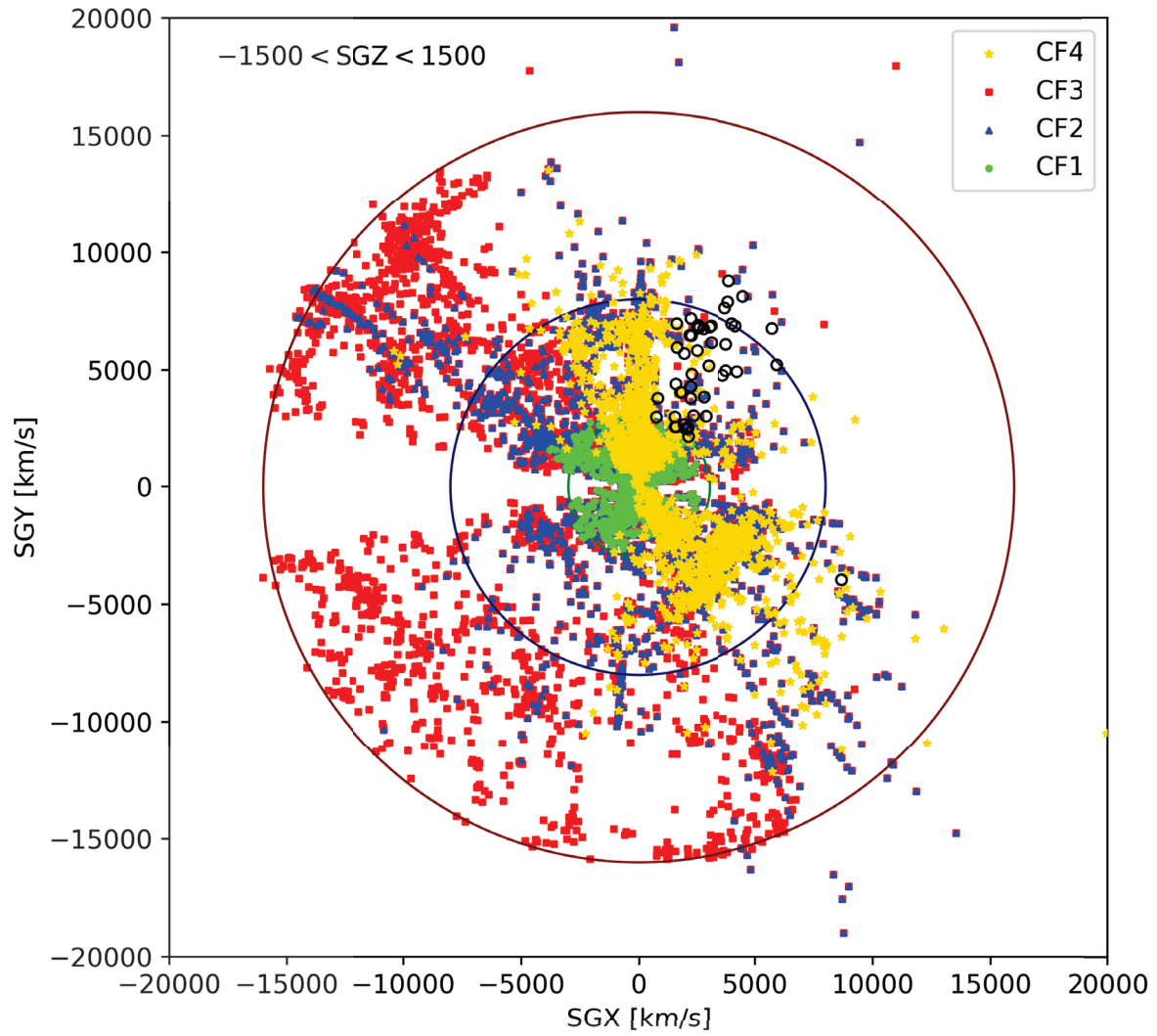


FIGURE 3.28 – Tranche SGX-SGY de 30 Mpc/ h de largeur en SGZ. Les trois catalogues CF1, CF2 et CF3 sont représentés en vert, bleu et rouge respectivement. Les nouvelles galaxies qui permettront de construire le catalogue CF4 sont présentées en jaune. Les cercles vides noirs correspondent aux sources HI qui vont être observées au GBT durant le prochain semestre 2018B.

Chapitre 4

Identification, caractérisation et visualisation des grandes structures

Les mouvements des galaxies dans l'univers sont responsables de la formation de grandes structures. Les catalogues de redshift montrent que les galaxies sont regroupées dans des amas et filaments, et qu'il existe des régions contenant très peu de galaxies : des vides. Ces régions sont tout de même difficiles à délimiter. Cependant, pour la première fois, [Tully et al. \(2014\)](#) ont défini et identifié, à l'aide des vitesses particulières des galaxies du catalogue *Cosmicflows-2*, le volume d'un superamas : celui dans lequel nous sommes situés, Laniakea ; sous forme de bassin gravitationnel.

Ce chapitre est organisé de la façon suivante. Pour commencer, les deux méthodes de reconstruction des champs de vitesse particulière tri-dimensionnels seront décrites. Ensuite, les algorithmes utilisés pour exploiter de tels champs de vitesse et identifier les bassins et vallées gravitationnelles seront détaillés. Ces algorithmes seront par la suite appliqués à une simulation Λ CDM, puis au champ de vitesse reconstruit à partir des données *Cosmicflows-3*.

Sommaire

4.1	Champs de vitesse particulière tri-dimensionnels	102
4.1.1	Technique du filtre de Wiener et des réalisations contraintes . .	102
4.1.2	Reconstruction par échantillonnage de Gibbs	105
4.2	Exploitation des champs de vitesse particulière	106
4.2.1	Lignes de flux — <i>streamlines</i>	107
4.2.2	Segmentation des bassins gravitationnels	113
4.2.3	Vallées gravitationnelles	117
4.2.4	Implémentation en C++	118
4.3	Simulation	120
4.4	Application aux données <i>Cosmicflows</i>	129
4.4.1	Champs de surdensité et de vitesse <i>Cosmicflows</i>	129
4.4.2	<i>Cosmicflows-2</i> : notre superamas Laniakea	132
4.4.3	<i>Cosmicflows-3</i> : résultats préliminaires	136

4.1 Champs de vitesse particulière tri-dimensionnels

Il est possible, à l'aide des données observationnelles telles que les catalogues *Cosmicflows*, de reconstruire les champs de densité et de vitesse particulière à trois dimensions. Deux méthodes sont utilisées au sein de la collaboration *Cosmicflows* : la technique du Filtre de Wiener et des réalisations contraintes, et la méthode de reconstruction par échantillonnage de Gibbs. Les caractéristiques principales de ces deux techniques de reconstruction sont détaillées ci-après.

4.1.1 Technique du filtre de Wiener et des réalisations contraintes

Filtre de Wiener

Le filtre de Wiener a été élaboré par [Wiener, N. \(1964\)](#) dans le cadre du traitement du signal. Il consiste à réduire la quantité de bruit présent dans un signal donné en reconstruisant une estimation du signal sous-jacent réel (i.e sans bruit) à l'aide d'un modèle statistique. La technique du filtre de Wiener a par la suite fait son chemin jusqu'au domaine de l'astrophysique ([Rybicki and Press, 1992](#)), puis de la cosmologie : [Zaroubi et al. \(1995, 1999\)](#) ont développé un modèle théorique permettant de reconstruire les champs de densité et vitesse particulière à trois dimensions sous-jacents à partir des données observationnelles (catalogues de redshift ou de vitesses particulières radiales).

Tout d'abord, supposons que nous voulons reconstruire le champ de surdensité sous-jacent δ , et considérons un ensemble de M contraintes c_i , situées aux positions \vec{x}_i et contaminées par des erreurs statistiques ϵ_i ($i \in [1, M]$). Les contraintes c_i correspondent aux données observationnelles, et sont, dans le cadre de cette thèse, les vitesses particulières radiales des galaxies déduites des distances issues des catalogues *Cosmicflows*. Les erreurs ϵ_i correspondent alors aux erreurs sur les vitesses particulières radiales, obtenues en propageant les erreurs sur les distances observées. Les contraintes échantillonnent le champ de surdensité sous-jacent δ_j aux positions \vec{x}_j et tel que $j \in [1, N]$. On peut alors écrire :

$$c_i = \sum_{j=1}^N R_{ij} \delta_j + \epsilon_i, \quad (4.1)$$

où R_{ij} est la matrice de réponse et relie le champ de surdensité sous-jacent réel δ aux données observationnelles. Afin de reconstruire le champ de surdensité, on cherche un filtre F tel que :

$$\delta_i^F = \sum_{j=1}^M F_{ij} c_j, \quad (4.2)$$

où F est une matrice de dimensions $N \times M$. Le filtre optimal est celui minimisant la différence entre le champ de surdensité réel δ et le champ de surdensité estimé δ^F , i.e

minimisant la variance du résidu $D_i = \delta_i - \delta_i^F$ donnée par $\langle D_i D_j \rangle = \langle (\delta_i - \delta_i^F)(\delta_j - \delta_j^F) \rangle$. [Hoffman and Ribak \(1991\)](#) montrent que F correspond au filtre de Wiener, noté WF :

$$WF_{ij} = \sum_{k=1}^M \langle \delta_i c_j \rangle \langle c_j c_k \rangle^{-1}, \quad (4.3)$$

où $\langle \delta_i c_j \rangle$ correspond à la matrice de corrélation entre le champ de surdensité et les contraintes, et $\langle c_i c_j \rangle$ à la matrice d'autocorrélation des contraintes. Ces deux quantités dépendent de la matrice de réponse R_{ij} et de la matrice de covariance du champ de surdensité $\langle \delta_i \delta_j \rangle$, donnée par la transformée de Fourier du spectre de puissance $P(k)$, à condition que les erreurs ϵ_i ne soient pas corrélées aux contraintes et au champ de surdensité. Le champ de surdensité reconstruit par le filtre de Wiener est alors donné par :

$$\delta_i^{WF} = \sum_{j=1}^M \sum_{k=1}^M \langle \delta_i c_j \rangle \langle c_j c_k \rangle^{-1} c_k. \quad (4.4)$$

Réalisations Contraintes

La technique des réalisations contraintes a été mise en place par [Hoffman and Ribak \(1991, 1992\)](#). Nous avons vu ci-dessus que le champ de surdensité réel δ correspond à la somme du champ de surdensité estimé δ^{WF} et du résidu D . Ce dernier n'est pas connu mais il est possible d'en estimer une réalisation avec une distribution statistique identique au résidu réel. Cela permet ainsi de générer une réalisation contrainte (CR pour *Constrained Realization*) du champ de surdensité telle que $\delta^{CR}(\vec{x}) = \delta^{WF}(\vec{x}) + D(\vec{x})$.

En injectant l'équation 4.4 dans la variance du résidu $\langle D_i D_j \rangle$ donnée un peu plus haut, on peut écrire ([Hoffman and Ribak, 1991, 1992](#)) :

$$\langle D_i D_j \rangle = \langle \delta_i \delta_j \rangle - \sum_{k=1}^M \sum_{l=1}^M \langle \delta_i c_k \rangle \langle c_k c_l \rangle^{-1} \langle \delta_j c_l \rangle. \quad (4.5)$$

Cette équation montre que $\langle D_i D_j \rangle$ ne dépend pas explicitement des valeurs des contraintes c_i . En effet, l'équation 4.5 fait intervenir les contraintes seulement à travers les matrices d'autocorrélation $\langle c_i c_j \rangle$ et les matrices de corrélation contraintes-champ $\langle \delta_i c_j \rangle$. Il est ainsi possible de générer une réalisation contrainte du champ δ^{CR} de la façon décrite ci-après. Tout d'abord, une réalisation aléatoire (RR, *Random Realization*) du champ, notée δ^{RR} , est générée à partir de la matrice de covariance du champ $\langle \delta_i \delta_j \rangle$. Puis un ensemble de contraintes synthétiques est construit en extrayant les valeurs de δ^{RR} aux positions des données observationnelles. La réalisation contrainte du champ est ainsi obtenue par :

$$\delta^{CR}(\vec{x}) = \delta^{RR}(\vec{x}) + \sum_{i=1}^M \sum_{j=1}^M \langle \delta(\vec{x}) c_i \rangle \langle c_i c_j \rangle^{-1} (c_i - c_j^{RR}). \quad (4.6)$$

Dans les zones non-couvertes par les données, on a $\delta^{CR} = \delta^{RR}$. Dans le cas contraire, i.e aux positions \vec{x}_i des contraintes, on a $\delta^{CR} = \delta^{WF}$. De plus, comme $\langle \delta^{RR} \rangle = 0$, on a $\langle \delta^{CR} \rangle = \delta^{WF}$.

Champ de vitesse particulière

De manière complètement analogue, il est possible de reconstruire le champ de vitesse particulière à partir des données observationnelles à l'aide du filtre de Wiener. L'équation 4.4 devient :

$$v_{\alpha}^{WF}(\vec{x}) = \sum_{i=1}^M \sum_{j=1}^M \langle v_{\alpha}(\vec{x}) c_i \rangle \langle c_i c_j \rangle^{-1} c_j, \quad (4.7)$$

où v_{α}^{WF} représente la composante cartésienne α du champ de vitesse. Enfin, les réalisations contraintes du champ de vitesse sont données par :

$$v_{\alpha}^{CR}(\vec{x}) = v_{\alpha}^{RR}(\vec{x}) + \sum_{i=1}^M \sum_{j=1}^M \langle v_{\alpha}(\vec{x}) c_i \rangle \langle c_i c_j \rangle^{-1} (c_i - c_j^{RR}). \quad (4.8)$$

La méthode de reconstruction par filtre de Wiener a été appliquée aux données *Cosmicflows-2*. Les champs de surdensité et de vitesse particulière reconstruits seront présentés dans la section 4.4.1. Cependant, il a été montré par [Graziani et al. \(2018\)](#) que cette méthode n'est pas adaptée pour l'application aux données *Cosmicflows-3* et que cela nécessite la mise en place d'une nouvelle méthode de reconstruction.

4.1.2 Reconstruction par échantillonnage de Gibbs

L'application de la technique du filtre de Wiener et des réalisations contraintes sur les données *Cosmicflows-3* résulte en une reconstruction irréaliste (avec par exemple un inflow global artificiel en direction de l'observateur de plusieurs milliers de km/s). En effet, une modélisation incorrecte des données observationnelles, i.e des mesures des modules de distances, peut introduire divers biais. Afin de minimiser de tels biais, l'erreur sur la mesure du module de distance doit être propagée sur la mesure des vitesses particulières observées et sur les positions des galaxies. Ceci a ainsi motivé la mise en place d'une nouvelle méthode, basée sur l'idée suivante : les distances des galaxies sont des paramètres libres, qui doivent être ajustés afin de minimiser la différence entre les vraies vitesses particulières des galaxies et celles estimées par la reconstruction. Le modèle développé par [Graziani et al. \(2018\)](#) pour la mise en oeuvre de cette méthode est une adaptation du modèle bayésien proposé par [Lavaux \(2016\)](#).

Fonction de vraisemblance

Le but est donc de reconstruire le champ de vitesse dans une grille à partir des observations provenant des catalogues *Cosmicflows*, i.e les modules de distance μ et les redshifts z . Pour ce faire, les observations sont modélisées à partir des paramètres du problème, i.e les distances d , le champ de vitesse linéaire à trois dimensions \vec{v}_g reconstruit sur une grille de dimension M^3 , la constante de Hubble H_0 (qui sera représentée par h_{eff} et la dispersion due aux non-linéarités σ_{NL} .

Il faut tout d'abord construire la fonction de vraisemblance, i.e la probabilité conditionnelle des observations (z, μ) sachant les paramètres du modèle $(d, \vec{v}_g, h_{eff}, \sigma_{NL})$. Celle-ci s'obtient simplement en combinant les probabilités conditionnelles sur μ et z puisqu'elles sont indépendantes, puis en factorisant sur toutes les galaxies du catalogue observationnel considéré. Ainsi, la fonction de vraisemblance est donnée par :

$$\mathcal{L}(\mathcal{M}, \mathcal{Z}|\mathcal{D}, h_{eff}, \sigma_{NL}, \vec{v}_g) = \prod_i \left\{ \frac{1}{2\pi\sigma_{\mu,i}} \exp \left[-\frac{1}{2} \frac{\left(\mu_i - 5 \log_{10} \frac{h_{eff} d_i}{10 \text{ pc}} \right)^2}{\sigma_{\mu,i}^2} \right] \right. \\ \left. \times \frac{1}{\sqrt{c\sigma_z^2(1+z_{\text{mod},i})^{-2} + \sigma_{NL}^2}} \exp \left[-\frac{1}{2} \frac{(u(z_i, d_i) - \vec{v}_g(\vec{r}_i) \cdot \hat{\vec{r}}_i)^2}{c\sigma_z^2(1+z_{\text{mod},i})^{-2} + \sigma_{NL}^2} \right] \right\}. \quad (4.9)$$

Le premier terme correspond à la probabilité conditionnelle sur le module de distance, tandis que le second terme correspond à la probabilité conditionnelle sur le redshift. On suppose ici que les erreurs sur μ et z sont gaussiennes et notées σ_μ et σ_z respectivement.

Par ailleurs, le paramètre h_{eff} est introduit afin de modéliser l'incompatibilité potentielle du point zéro de la calibration de l'indicateur de distance considéré avec la valeur de la constante de Hubble H_0 choisie. En effet, la relation module de distance – distance d'une galaxie n'est connue qu'à une constante près, i.e le point zéro, qui peut être estimé.

Celui-ci est cependant extrêmement corrélé à la valeur de H_0 . Afin de mesurer le module de distance d’une galaxie, il faut supposer la valeur de l’un de ces deux paramètres puis en déduire la valeur de l’autre. [Graziani et al. \(2018\)](#) ont choisi de fixer la valeur de la constante de Hubble et de laisser varier la valeur du point zéro. La non-compatibilité de ces deux paramètres est exprimée à travers le paramètre h_{eff} .

Enfin, la quantité $u(z, d)$ représente la vitesse particulière radiale calculée à partir de la distance observée d’une galaxie avec l’équation 2.7.

Échantillonnage

À partir de cette fonction de vraisemblance, il est possible de déduire la probabilité postérieure des paramètres du modèle donnée par $\mathcal{P}(\mathcal{D}, h_{eff}, \sigma_{NL}, \vec{v}_g | \mathcal{M}, \mathcal{Z})$ à l’aide du théorème de Bayes :

$$\mathcal{P}(\mathcal{D}, h_{eff}, \sigma_{NL}, \vec{v}_g | \mathcal{M}, \mathcal{Z}) \propto \mathcal{L}(\mathcal{M}, \mathcal{Z} | \mathcal{D}, h_{eff}, \sigma_{NL}, \vec{v}_g) \times \mathcal{P}(\mathcal{D}) \mathcal{P}(h_{eff}) \mathcal{P}(\sigma_{NL}) \mathcal{P}(\vec{v}_g). \quad (4.10)$$

Le but est ainsi de reconstruire le champ de vitesse particulière à partir des observations *Cosmicflows*, i.e trouver la *solution* qui maximise la probabilité $\mathcal{P}(\mathcal{D}, h_{eff}, \sigma_{NL}, \vec{v}_g | \mathcal{M}, \mathcal{Z})$. Ceci peut être effectué à l’aide de l’algorithme proposé par [Lavaux \(2016\)](#), qui échantillonne le modèle par la méthode de Monte-Carlo par chaînes de Markov. Cette technique d’échantillonnage est un échantillonnage dit de Gibbs ([Metropolis et al., 1953](#); [Gelfand and Smith, 1990](#); [Geman and Geman, 1984](#); [Hastings, 1970](#)).

4.2 Exploitation des champs de vitesse particulière

Un champ de vitesse particulière à trois dimensions reconstruit à partir des observations peut être exploité afin d’extraire diverses informations et caractéristiques sur les grandes structures de l’univers local. Pour cela, durant cette thèse, deux algorithmes permettant d’identifier et caractériser les grandes structures se trouvant dans le champ de vitesse local ont été développés dans le cadre d’une collaboration réunissant des informaticiens (LIRIS) et astrophysiciens (IPNL) de Lyon. Ces deux algorithmes sont tous les deux basés sur les lignes de flux. Le premier est un algorithme de segmentation permettant de localiser les attracteurs et repoussoirs et d’identifier les bassins gravitationnels associés (bassins d’attraction et de répulsion). Le second est une méthode innovatrice permettant de visualiser et identifier des structures qui sont présentes dans le champ de vitesse, même dans des zones peu ou pas couvertes par les données observationnelles. Après avoir défini les lignes de flux, les deux algorithmes, utilisés tout au long de ce chapitre pour identifier les grandes structures de l’univers local, seront décrits, et validés sur un champ de vitesse synthétique construit selon la méthode suivante :

- tout d’abord, le nombre de structures (attracteurs ou repoussoirs) que l’on souhaite placer dans le champ de vitesse est fixé (ou choisi aléatoirement) et noté N_{struct} ;

- les positions dans le champ de vitesse de ces N_{struct} structures, notées $\vec{r}_{\text{struct}}^i$ (où $i = 1, \dots, N_{\text{struct}}$), sont choisies aléatoirement ;
- un facteur F_{struct}^i , représentant la masse de chaque structure, est aussi fixé aléatoirement. Son signe détermine le type de la structure : il peut être négatif (attracteurs) ou positif (repousseurs) ;
- enfin, un champ de vitesse synthétique \vec{v}_{mock} contenant ces structures est construit. Sa valeur à la position \vec{r} est donnée par :

$$\vec{v}_{\text{mock}}(\vec{r}) = \sum_{i=1}^{N_{\text{struct}}} \frac{F_{\text{struct}}^i}{|\vec{r}_{\text{struct}}^i - \vec{r}|^2} (\vec{r}_{\text{struct}}^i - \vec{r}). \quad (4.11)$$

La validation des algorithmes est effectuée sur un champ de vitesse généré avec l'équation 4.11. Le champ de vitesse synthétique considéré par la suite contient 10 structures (attracteurs ou repousseurs selon le signe de F_{struct}) situées à des coordonnées choisies aléatoirement, qui seront comparées aux coordonnées des attracteurs et repousseurs détectés automatiquement par notre algorithme. Le champ de vitesse est construit sur une grille de 500 Mpc/h de côté et de dimensions $128 \times 128 \times 128$, ce qui donne une résolution de 7.8125 Mpc/h. La taille de cette boîte est similaire à celles dans lesquelles les champs de vitesse reconstruits à partir des données observationnelles ont été calculés, et qui seront considérées dans ce chapitre (320 Mpc/h pour CF2 et 600 Mpc/h pour CF3).

4.2.1 Lignes de flux — *streamlines*

Un champ de vitesse peut être visualisé de différentes façons. Ces techniques de visualisation reposent sur l'utilisation de particules sans masse en tant que *traceurs* afin de visualiser leur trajectoire (Jobard, Bruno, 2000) :

- *streamline* : c'est une courbe tangente, en tout point, au champ de vitesse considéré. Dans le cas d'un champ de vitesse stationnaire (indépendant du temps), une streamline correspond à la trajectoire qu'aurait une particule sans masse si elle était placée dans le champ de vitesse ;
- *pathlines* : trajectoire qu'une particule sans masse suivrait si elle était relâchée dans un champ de vitesse non-stationnaire. Cette trajectoire représente les positions successives de la particule dans le temps ;
- *streaklines* : courbe reliant les positions à un instant donné de plusieurs particules relâchées à partir d'un même point de départ fixe dans un champ de vitesse non-stationnaire.

Pour un champ de vitesse indépendant du temps, tel que les champs de vitesse qui seront considérés tout au long de ce chapitre, les courbes de l'espace représentant les streamlines, pathlines et streaklines sont identiques. Le terme de streamline, ou ligne de flux, sera par la suite utilisé afin de décrire une telle trajectoire dans les champs de vitesse considérés.

Définition

Dans un champ de vitesse stationnaire noté \vec{v} , une ligne de flux $\vec{s}(\tau)$ paramétrisée par la variable τ (représentant la "distance" dans la ligne de flux), est en tout point tangente au champ de vitesse \vec{v} . Le vecteur tangent à $\vec{s}(\tau)$ étant donné par $\frac{d\vec{s}}{d\tau}$, une ligne de flux est obtenue en résolvant l'équation différentielle suivante :

$$\frac{d\vec{s}}{d\tau} = \vec{v}(\vec{s}), \quad (4.12)$$

avec $\vec{s}(\tau = 0) = \vec{s}_0$ comme condition initiale correspondant au point de départ de la ligne de flux, que l'on appelle aussi point *seed*. La position à travers laquelle passe la ligne de flux après un pas $\Delta\tau$ est donnée par :

$$\vec{s}(\tau + \Delta\tau) = \vec{s}(\tau) + \int_{\tau}^{\tau + \Delta\tau} \vec{v}(\vec{s}(\tau)) d\tau \quad (4.13)$$

La solution de cette équation, i.e la ligne de flux, est obtenue en intégrant les composantes du champ de vitesse depuis le point *seed*. En calculant l'équation 4.13 pour N pas $\Delta\tau$ consécutifs, un ensemble de points $\{\vec{s}_i\}_{i \in [0, N]}$, appelés points de contrôle, est généré. Cet ensemble de point représente la ligne de flux sous forme discrétisée, la courbe de l'espace passant par tous ces points étant la ligne de flux. Ceci est schématisé sur la figure 4.1. Les flèches rouges représentent les vecteurs vitesse. La ligne de flux est construite à partir du point *seed* \vec{s}_0 représenté en bleu. Les point de contrôle $\vec{s}_1, \vec{s}_2, \vec{s}_3, \vec{s}_N$ générés en intégrant le vecteur vitesse pour plusieurs pas consécutifs sont représentés en vert. La courbe de l'espace passant par tous ces points et représentant la ligne de flux est tracée en noir. Par souci de clarté pour la suite de ce chapitre, le point de contrôle \vec{s}_N , auquel le calcul de la ligne de flux a été arrêté, sera noté comme \vec{s}_{stop} et appelé point d'arrêt.

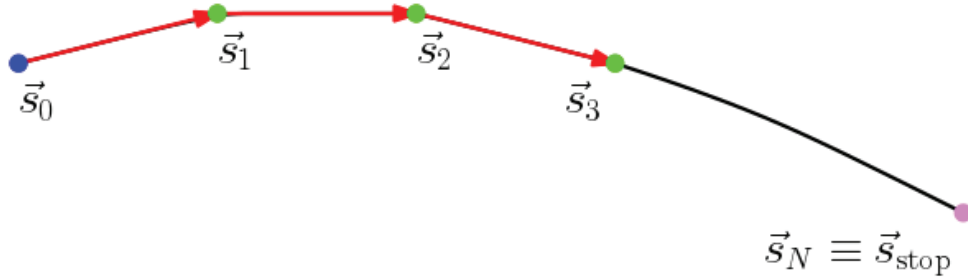


FIGURE 4.1 – Exemple de la définition d'une ligne de flux.

Les lignes de flux peuvent être construites selon trois points de vue permettant de visualiser le champ de vitesse de trois façons différentes :

- point de vue *avant*, ou *forward* : la ligne de flux est construite à partir d'un point *seed* donné en intégrant le champ de vitesse \vec{v} ;

- point de vue *arrière*, ou *backward* : la ligne de flux est construite à partir d'un point seed donné en intégrant l'opposé du champ de vitesse, i.e $-1 \times \vec{v}$;
- point de vue *double*, ou *both* : ce point de vue correspond à la réunion des deux précédents, c'est-à-dire que la ligne de flux est construite à partir d'un point seed donné en intégrant à la fois \vec{v} et $-1 \times \vec{v}$.

Les lignes de flux convergent vers les points critiques du champ de vitesse considéré, qui sont les points où le vecteur vitesse est nul. Dans un champ de vitesse particulière, de tels points correspondent aux attracteurs ou repoussoirs. La visualisation d'un champ de vitesse à l'aide des lignes de flux permettent par conséquent de représenter la topologie de ce champ. Les points de convergence des lignes de flux construites dans le point du vue avant correspondent aux attracteurs, zones pointées par le champ de vitesse. Inversement, la convergence des lignes de flux générées dans le point de vue *arrière* représentent les repoussoirs. Ainsi, les points d'arrêt \vec{s}_{stop} des lignes de flux sont très utiles quand à l'étude de leur convergence pour détecter les attracteurs, repoussoirs, et les bassins gravitationnels qui leur sont associés. Par ailleurs, le point de vue *double* permet de visualiser ces deux types de structures à la fois, cependant cette combinaison perd un sens physique et ne sera pas considérée par la suite.

Les lignes de flux générées à partir du champ de vitesse synthétique considéré dans cette section peuvent être visualisées sur la figure 4.2. Cette figure présente une visualisation tri-dimensionnelle du champ de vitesse synthétique. Les sphères rouges et bleues sont placées aux positions des attracteurs et repoussoirs situés dans le champ de vitesse. A gauche, les lignes de flux sont construites dans le point de vue *avant* (*forward*), et convergent en effet vers les attracteurs. Inversement, à droite, les lignes de flux sont calculées dans le point de vue *arrière* (*backward*) et permettent ainsi de visualiser les repoussoirs.

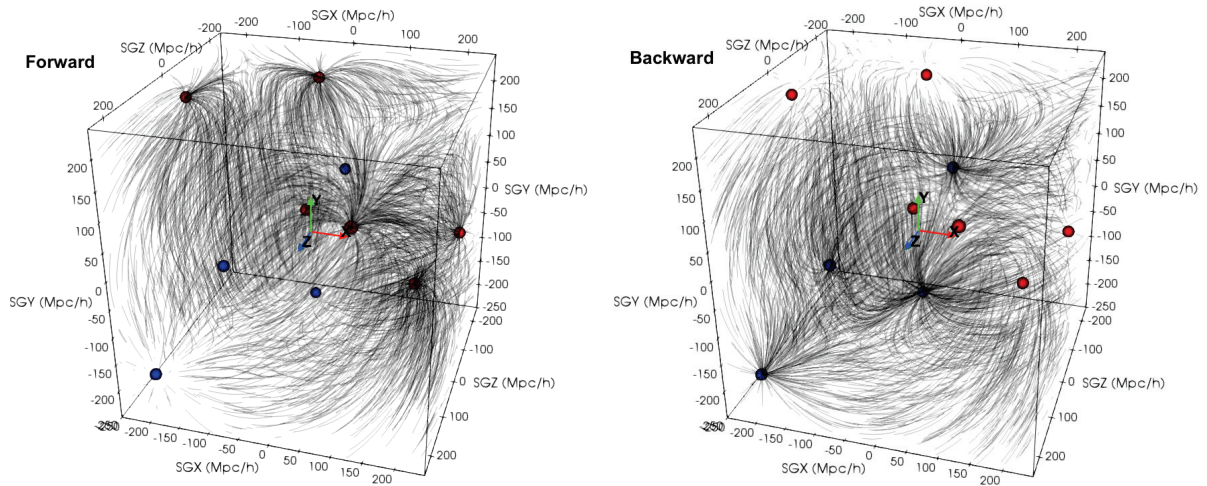


FIGURE 4.2 – Visualisation tri-dimensionnelle du champ de vitesse synthétique considéré dans cette section à l'aide de lignes de flux calculées dans le point de vue *avant* (gauche) et *arrière* (droite). Les sphères rouges et bleues sont placées aux positions des attracteurs et repoussoirs situés dans le champ de vitesse.

Intégration numérique

En pratique, la résolution de l'équation 4.13 se fait par intégration numérique. De nombreuses méthodes d'intégration numérique existent, dont l'intégrateur d'Euler et l'intégrateur Runge-Kutta.

L'intégrateur d'Euler (aussi appelé intégrateur Runge-Kutta d'ordre 1) permet d'obtenir la suite de points de contrôle $\{\vec{s}_i\}$ par la formule :

$$\vec{s}_{i+1} = \vec{s}_i + \Delta\tau \vec{v}(\vec{s}_i). \quad (4.14)$$

Ainsi, la ligne de flux est construite depuis le point seed \vec{s}_0 :

$$\vec{s}_i = \vec{s}_0 + \sum_{u=0}^i \Delta\tau \vec{v}(\vec{s}_u). \quad (4.15)$$

Cet intégrateur est simple mais peu précis. Afin d'améliorer la précision de cet intégrateur, il est possible d'utiliser un point supplémentaire \vec{k} pour chaque pas $\Delta\tau$. Ceci correspond à la méthode d'intégration dite de Runge-Kutta d'ordre 2 :

$$\vec{k} = \Delta\tau \vec{v}\left(\vec{s}_i + \frac{\Delta\tau \vec{v}(\vec{s}_i)}{2}\right), \quad (4.16)$$

$$\vec{s}_{i+1} = \vec{s}_i + \vec{k}. \quad (4.17)$$

Il est possible de faire encore mieux en utilisant l'intégration numérique Runge-Kutta d'ordre 4, qui utilise plus de points intermédiaires. Le point de contrôle \vec{s}_{i+1} de la ligne de flux est donné en fonction du point précédent \vec{s}_i par la formule :

$$\vec{s}_{i+1} = \vec{s}_i + \left(\frac{1}{6}\vec{k}_1 + \frac{1}{3}\vec{k}_2 + \frac{1}{3}\vec{k}_3 + \frac{1}{6}\vec{k}_4\right), \quad (4.18)$$

où les vecteurs \vec{k}_1 , \vec{k}_2 , \vec{k}_3 et \vec{k}_4 correspondent à des approximations de la tangente au champ de vitesse \vec{v} entre les étapes τ_i et τ_{i+1} et sont donnés par :

$$\vec{k}_1 = \vec{v}(\vec{s}_i), \quad (4.19)$$

$$\vec{k}_2 = \vec{v}\left(\vec{s}_i + \vec{k}_1 \frac{\Delta\tau}{2}\right), \quad (4.20)$$

$$\vec{k}_3 = \vec{v}\left(\vec{s}_i + \vec{k}_2 \frac{\Delta\tau}{2}\right), \quad (4.21)$$

$$\vec{k}_4 = \vec{v}(\vec{s}_i + \vec{k}_3 \Delta\tau). \quad (4.22)$$

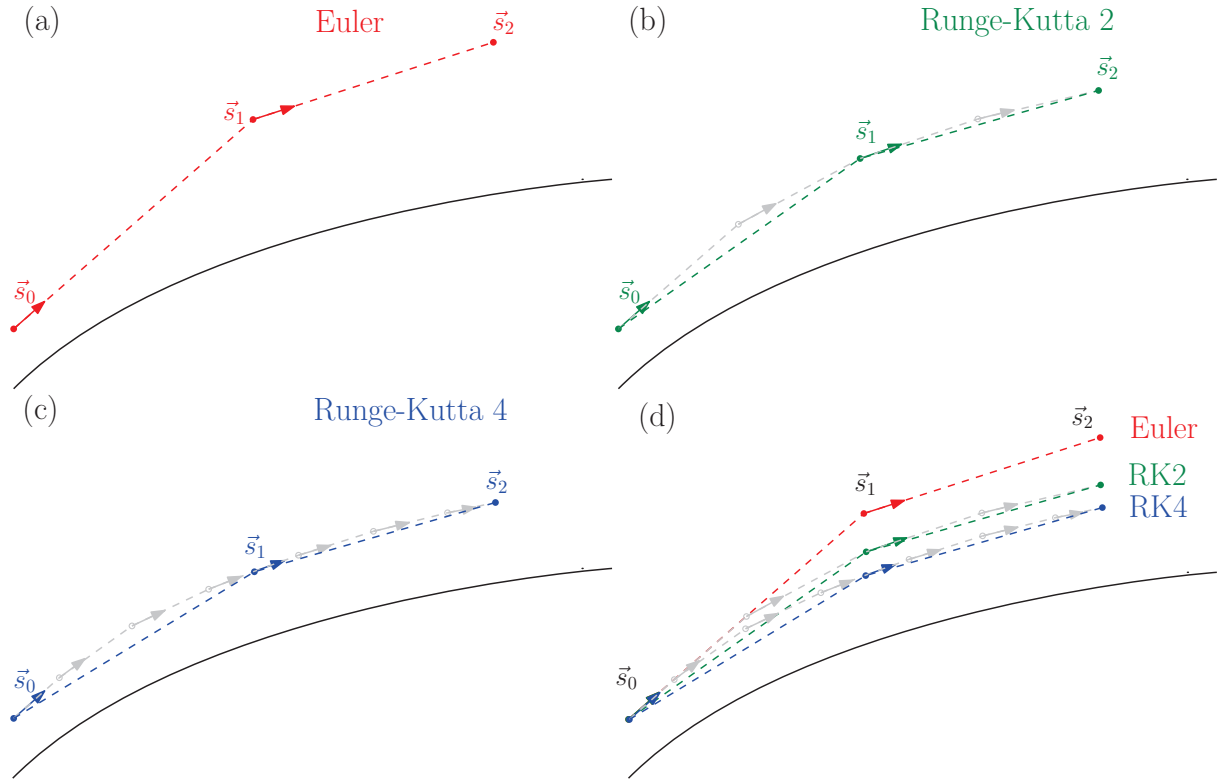


FIGURE 4.3 – Schéma de différentes méthodes d'intégration numérique : (a) Euler, (b) Runge-Kutta d'ordre 2 (RK2), (c) Runge-Kutta d'ordre 4 (RK4), (d) comparaison des solutions numériques obtenues à l'aide des trois intégrateurs précédents.

Les intégrateurs numériques d'Euler, de Runge-Kutta d'ordre 2 et de Runge-Kutta d'ordre 4 sont schématisés sur les figures 4.3 (a), (b) et (c) respectivement. Ces trois figures sont réunies dans la figure 4.3 (d) pour pouvoir comparer plus facilement les solutions numériques obtenues à l'aide des trois méthodes d'intégration numérique. La solution numérique approximative est tracée en traits pointillés en couleur : rouge pour la méthode d'Euler, vert pour Runge-Kutta d'ordre 2 et bleu pour Runge-Kutta d'ordre 4. La courbe noire représente la vraie solution $\vec{s}(\tau)$. Les flèches représentent les dérivées de $\vec{s}(\tau)$, i.e le champ de vitesse, utilisées pour positionner les points de contrôle à partir du point seed et construire la ligne de flux pas par pas. Les points et dérivées intermédiaires utilisés par les méthodes Runge-Kutta d'ordres 2 et 4 pour obtenir un point de contrôle à partir du précédent sont tracés en gris. Les intégrateurs Runge-Kutta d'ordres 2 et 4 sont plus précis que la méthode d'Euler. De plus, l'intégration numérique par la méthode Runge-Kutta d'ordre 4 a une précision légèrement plus haute que celle d'ordre 2.

Dans cette thèse, les lignes de flux sont calculées à l'aide de la classe `vtkStreamTracer` implémentée dans la librairie VTK (voir section 4.2.3) et utilisant l'intégrateur numérique de Runge-Kutta d'ordre 4 pour générer des lignes de flux. Celles-ci sont générées à partir de chaque voxel de la grille contenant le champ de vitesse particulière de dimensions $M \times M \times M$ (soit M^3 points seed). Le pas $\Delta\tau$ est choisi de façon à ce que les points de contrôle des lignes de flux soient espacés d'un voxel : $\Delta\tau$ correspond alors à la résolution de la grille considérée. Enfin, le nombre maximal d'itérations est choisi assez haut afin d'avoir un nombre suffisant de lignes de flux convergeant vers les points attracteurs ou

repousseurs du champ de vitesse considéré.

La figure 4.4 montre l'évolution des lignes de flux en fonction du nombre maximal d'itérations N choisi, pour un pas fixe (égal à un voxel) : (a) $N = 5$, (b) $N = 25$, (c) $N = 50$ et (d) $N = 100$; dans une tranche SGX-SGY du champ de vitesse synthétique de largeur 50 Mpc/h. La longueur des lignes de flux augmente avec le nombre maximal d'itérations. Lorsque N est trop faible (voir par exemple $N = 5$), il est difficile de visualiser le champ de vitesse et ainsi identifier les points attracteurs. Ceux-ci sont de plus en plus visibles quand le nombre d'itérations augmente. Nous utiliserons par la suite $N = 100$ pour calculer les lignes de flux sur les champs de vitesse, car nous verrons que cette valeur est largement suffisante, y compris avec les données observationnelles.

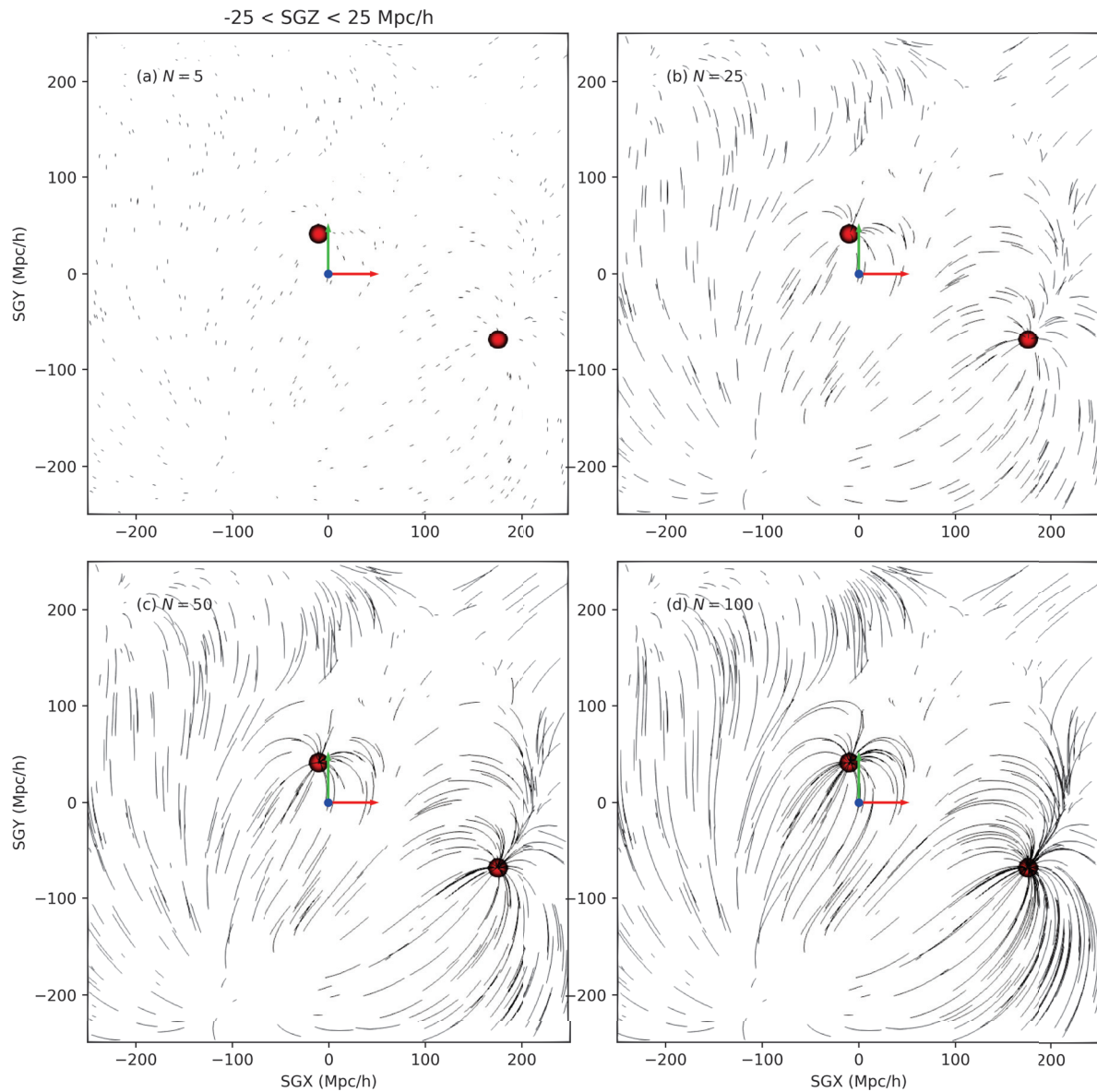


FIGURE 4.4 – Évolution des lignes de flux en fonction du nombre maximal d'itérations N , pour un pas fixe (égal à un voxel), dans une tranche SGX-SGY du champ de vitesse synthétique de largeur 50 Mpc/h : (a) $N = 5$, (b) $N = 25$, (c) $N = 50$ et (d) $N = 100$.

4.2.2 Segmentation des bassins gravitationnels

L'algorithme de segmentation développé durant cette thèse permet d'identifier les grandes structures en leur attribuant des bassins d'attraction et de répulsion. Il faut tout d'abord identifier les centres des bassins d'attraction (ou de répulsion), i.e les points de convergence des lignes de flux, dans le champ de vitesse considéré. Celui-ci est par la suite divisé en régions, représentant les bassins versants, en faisant croître les bassins gravitationnels autour des points attracteurs ou repousseurs.

Attracteurs et repousseurs

Dans un champ de vitesse, les attracteurs et les repousseurs correspondent aux positions où les lignes de flux convergent. Selon la direction d'intégration choisie (ou point de vue), on identifie ainsi des attracteurs (point de vue *avant*) ou des repousseurs (point de vue *arrière*).

Ces régions sont de fait les positions correspondant à la destination finale du plus grand nombre de lignes de flux. Cette définition est par conséquent utilisée afin de trouver automatiquement les positions des attracteurs ou des repousseurs dans un cube de données contenant le champ de vitesse de dimensions $M \times M \times M$ et de côté L .

Pour ce faire, les lignes de flux sont calculées au départ de chaque voxel du champ de vitesse, générant ainsi un ensemble de points de contrôle $\{\vec{s}_i\}$ pour chaque voxel. Une matrice M_{stop} , elle aussi de dimensions $M \times M \times M$ et de côté L , est ensuite construite. Dans cette matrice, chaque voxel contient une valeur qui est égale au nombre de lignes de flux qui ont ce voxel en tant que destination finale \vec{s}_{stop} .

La figure 4.5 montre une visualisation tri-dimensionnelle de la matrice M_{stop} obtenue à partir du champ de vitesse synthétique considéré dans cette section. Les points de vue avant et arrière, calculés séparément, sont superposés. Les points noirs correspondent à des surdensités dans la matrice, le seuil utilisé est $M_{\text{stop}} > 100$. Les sphères rouges et bleues sont placées aux positions des attracteurs et repousseurs respectivement. Cette figure montre ainsi que des surdensités dans la matrice M_{stop} sont observées aux positions des attracteurs et repousseurs. Cela confirme en effet qu'un grand nombre de lignes de flux ont comme destination finale ces points attracteurs ou repousseurs, selon le point de vue considéré.

Les attracteurs ou les repousseurs correspondent alors aux maxima locaux de M_{stop} , selon la direction d'intégration choisie. Ainsi en identifiant les maxima locaux de cette matrice dont les positions seront notées \vec{x}_{max} , un ensemble de points $\{\vec{x}_{\text{max},i}\}$ est obtenu, où $i \in [1, N_{\text{max}}]$ et N_{max} est le nombre de maxima locaux détectés.

Les positions des structures (attracteurs et repousseurs) dans le champ de vitesse synthétique étant connues, il est possible de vérifier si les maxima locaux de M_{stop} sont en effet positionnés aux coordonnées des attracteurs et des repousseurs.

La figure 4.6 présente deux visualisations à trois dimensions du champ de vitesse synthétique $\vec{v}_{\text{mock}}(\vec{r})$, dans le point de vue *avant* (gauche) et *arrière* (droite). Les sphères rouges et bleues sont positionnées respectivement aux vraies coordonnées des attracteurs

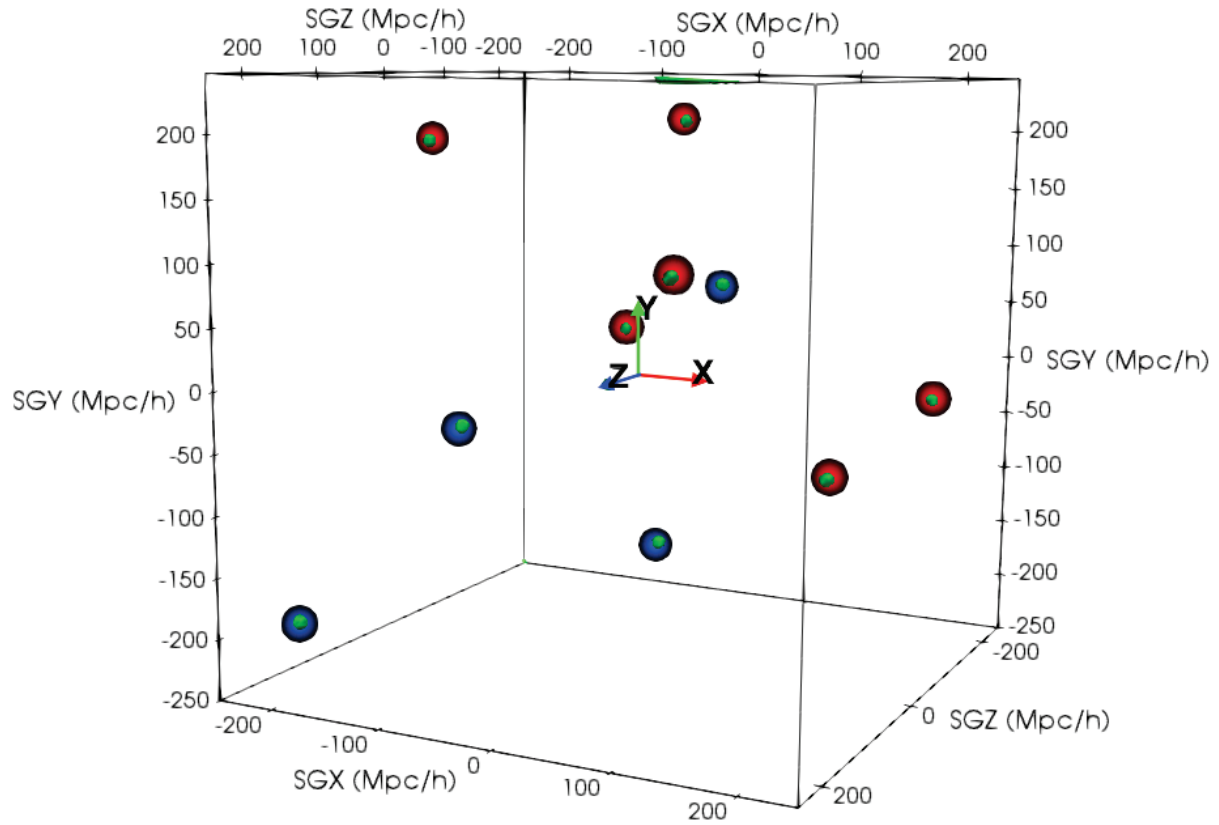


FIGURE 4.5 — Visualisation tri-dimensionnelle de la matrice M_{stop} obtenue à partir du champ de vitesse synthétique. Les points de vue avant et arrière, calculés séparément, sont superposés. Les points noirs correspondent à des surdensités dans la matrice ($M_{\text{stop}} > 100$). Les sphères rouges et bleues sont placées aux positions des attracteurs et repoussoirs respectivement.

et repoussoirs positionnés dans le champ de vitesse synthétique. Les triangles verts sont placés aux coordonnées des maxima locaux détectés dans M_{stop} obtenue à partir du champ synthétique. Les triangles verts étant superposés aux sphères rouges et bleues, cela prouve que tous les attracteurs et tous les repoussoirs sont détectés avec succès.

De plus, les coordonnées des maxima locaux détectés dans la matrice M_{stop} sont listées dans la table 4.1 (colonne (1)) et comparées aux vraies positions des attracteurs et repoussoirs (colonne (2)). L'algorithme retrouve presque parfaitement les coordonnées des différentes structures (attracteurs ou repoussoirs), avec une erreur maximale d'un voxel, i.e la résolution de la grille dans laquelle le champ de vitesse a été construit.

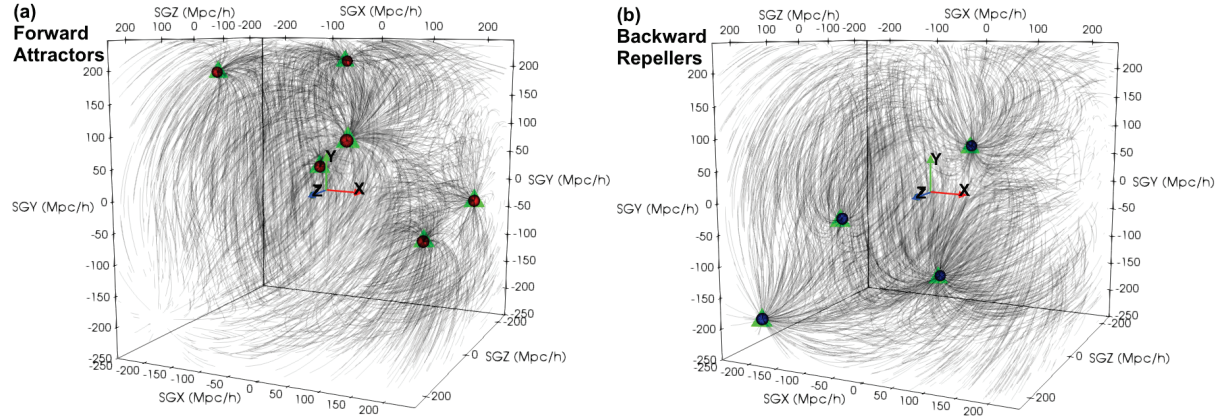


FIGURE 4.6 – Visualisations à trois dimensions du champ de vitesse synthétique $\vec{v}_{\text{mock}}(\vec{r})$, dans le point de vue *avant* (gauche) et *arrière* (droite). Les sphères rouges et bleues sont positionnées respectivement aux vraies coordonnées des attracteurs et repoussoirs positionnés, tandis que les triangles verts sont placés aux coordonnées des maxima locaux détectés dans M_{stop} .

TABLE 4.1 – Test sur un champ de vitesse synthétique. Les coordonnées des structures (attracteurs et repoussoirs) placées dans le champ de vitesse sont connues et listées dans la colonne (1). Les coordonnées des attracteurs et repoussoirs détectés automatiquement par l’algorithme décrit en section 4.2.2 sont données dans la colonne (2). Dans les deux colonnes, les coordonnées sont données dans le système de coordonnées supergalactiques cartésiennes (SGX,SGY,SGZ), en Mpc/h. Les étiquettes associées aux bassins gravitationnels sont listées en colonne (3).

Attracteurs	Vraies coordonnées	Détection automatique	Étiquette
	(−35.16, 207.03, −167.97)	(−39.06, 207.03, −167.97)	1
	(226.56, −15.62, −93.75)	(226.56, −15.62, −93.75)	2
	(−242.19, 187.50, −62.50)	(−242.19, 187.50, −66.41)	3
	(171.88, −70.31, 3.91)	(171.88, −70.31, 0.00)	4
	(−11.72, 39.06, 0.00)	(−11.72, 39.06, 0.00)	5
	(109.38, 105.47, 171.88)	(109.38, 105.47, 167.97)	6
Repoussoirs			
	(23.44, 66.41, −121.09)	(19.53, 62.50, −125.00)	1
	(−19.53, −164.06, −85.94)	(−23.44, −167.97, −85.94)	2
	(−128.91, −42.97, 70.31)	(−132.81, −46.88, 70.31)	3
	(−207.03, −191.41, 195.31)	(−210.94, −195.31, 191.41)	4

Segmentation du champ de vitesse en bassins gravitationnels

À l'aide des points d'arrêt des lignes de flux, calculés pour obtenir la matrice M_{stop} , et des coordonnées des attracteurs (ou repousseurs) détectés précédemment, il est possible de construire des bassins d'attraction (ou de répulsion) avec la méthode de segmentation décrite ci-après.

Pour un attracteur ou un repousseur donné, un bassin d'attraction (*basin of attraction*, BoA), ou de répulsion (*basin of repulsion*, BoR), est défini comme la région contenant tous les points de l'espace dont les lignes de flux convergent vers la position de l'attracteur ou du repousseur respectif.

Le but de la méthode de segmentation est de construire une matrice B , toujours de dimensions $M \times M \times M$ et de côté L , dans laquelle les voxels sont distribués dans différents bassins gravitationnels. Pour y parvenir, chaque maximum local $\vec{x}_{\text{max},i}$ (attracteur ou repoussoir, voir plus haut) est associé à un bassin gravitationnel. Des *étiquettes* (*labels*) $m \in [1, N_{\text{max}}]$ sont attribuées aux différents bassins afin de pouvoir les identifier facilement dans la matrice B . Par exemple, dans le cas du champ de vitesse synthétique, six attracteurs et quatre repousseurs ont été détectés, on a donc six bassins d'attraction et quatre bassins de répulsion. Les étiquettes qui leur sont attribuées sont données dans la colonne (3) de la table 4.1.

La méthode de segmentation utilisée est la suivante :

- chaque bassin est initialisé à un unique voxel qui est celui situé à la position du maximum local $\vec{x}_{\text{max},i}$ associé. On attribue donc à ce voxel l'étiquette m respective ;
- pour chaque voxel du champ de vitesse, nous calculons la ligne de flux partant de ce voxel. Si la destination finale \vec{s}_{stop} d'une ligne de flux correspond à un voxel appartenant à un bassin gravitationnel, alors le point *seed* \vec{s}_0 de cette ligne de flux appartient lui aussi à ce bassin gravitationnel. L'étiquette associée à ce bassin est ainsi attribuée au voxel correspondant à la position \vec{s}_0 ;
- l'étape précédente est répétée jusqu'à convergence, pour un nombre d'itération donné.

De cette manière, les voxels du champ de vitesse sont distribués petit à petit dans les différents bassins associés aux attracteurs ou repousseurs localisés dans le champ de vitesse. Ceci fait ainsi croître, à chaque itération, les bassins autour des voxels correspondant aux positions des attracteurs ou repousseurs. Après un certain nombre d'itération, les bassins se stabilisent. Au final, la grille est divisée en bassins gravitationnels (BoA et BoR). Chaque voxel appartient à un seul bassin d'attraction et à un seul bassin de répulsion. Cependant, les voxels dont les streamlines sortent de la boîte de calcul et n'ont donc aucun point de convergence n'appartiennent à aucun bassin : dans ce cas, on leur attribue l'étiquette 0.

La figure 4.7 montre les bassins d'attraction (gauche) et de répulsion (droite), obtenus à partir du champ de vitesse synthétique, représentés respectivement par des isosurfaces colorées dans une tranche SGX-SGX de largeur 50 Mpc/h et centrée en SGZ = 0Mpc/h (gauche) et SGZ = -105Mpc/h (droite). Une couleur différente est attribuée à chaque bassin. De plus, le champ de vitesse synthétique est représenté par des lignes de flux

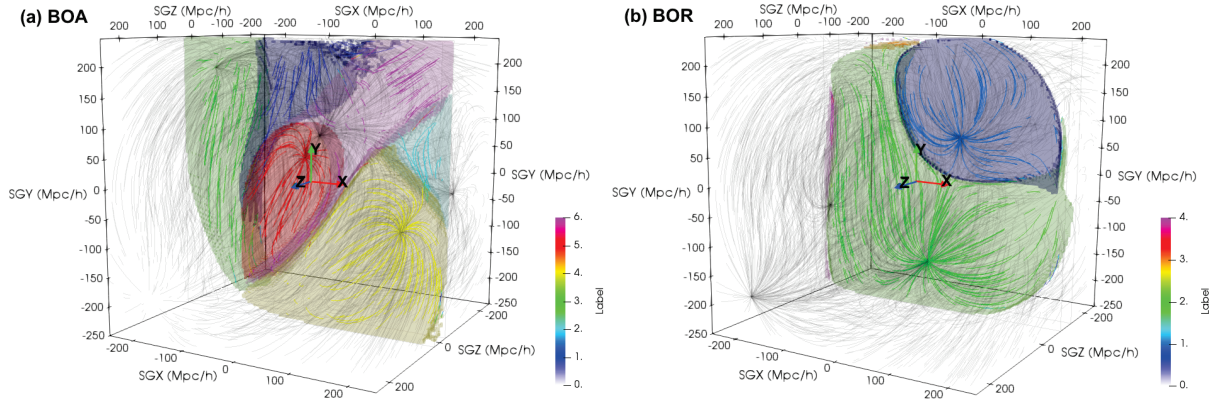


FIGURE 4.7 – Bassins d’attraction (gauche) et de répulsion (droite), obtenus à partir du champ de vitesse synthétique, représentés par des isosurfaces colorées dans une tranche SGX-SGX de largeur 50 Mpc/h et centrée en SGZ = 0Mpc/h (gauche) et SGZ = −105Mpc/h (droite). Les correspondances couleur / bassin sont données sur la palette de couleur, sur laquelle les bassins sont identifiés par l’étiquette qui leur est associée (voir table 4.1). Le champ de vitesse synthétique est représenté par des lignes de flux (point de vue *avant* à gauche, *arrière* à droite).

(point de vue *avant* à gauche, *arrière* à droite). Les lignes de flux situées dans les tranches considérées sont tracées avec des traits plus opaques et colorés. Leur couleur varie selon le bassin auquel elles appartiennent. Les correspondances couleur / bassin sont données sur la palette de couleur, sur laquelle les bassins sont identifiés par l’étiquette qui leur est associée (voir table 4.1). La correspondance lignes de flux / bassin gravitationnels est excellente : les bassins contiennent bien les lignes de flux convergeant vers l’attracteur ou le repoussoir associé. Cependant, sur la figure 4.7, un seul bassin d’attraction apparaît fermé. Les bassins des structures se situant vers les bords de la boîte s’étendent jusqu’aux bords car il n’y a pas de structures plus lointaines qui permettraient de fermer les bassins versants.

4.2.3 Vallées gravitationnelles

Dans cette section est détaillé un outil permettant de visualiser des vallées gravitationnelles. Cet outil exploite lui aussi les lignes de flux, et sa construction est très similaire à celle de la matrice des points d’arrêt M_{stop} détaillée dans la section 4.2.2. Il permet de visionner, indirectement, les lignes de flux générées à partir d’un champ de vitesse, et de visualiser leurs zones de convergence qui ne correspondent pas toujours à des points attracteurs ou repousseurs. Cet outil est obtenu en construisant un *histogramme des lignes de flux*.

Ce nouvel outil représente ainsi des *vallées gravitationnelles*, dans lesquelles passent les *rivières gravitationnelles* : les gradients du champ potentiel. L’histogramme montre en effet les régions où les lignes de flux se regroupent, i.e où les *rivières* passent.

Algorithme

Pour y parvenir, les lignes de flux sont calculées au départ de chaque voxel du champ de vitesse considéré, générant ainsi un ensemble de points de contrôle $\{\vec{s}_i\}$ pour chaque voxel. Ensuite, un *histogramme des points de contrôle* des lignes de flux est construit dans une matrice notée S , de dimensions $M \times M \times M$ et de côté L (tout comme le champ de vitesse). Chaque voxel de cette matrice est tout d'abord initialisé à zéro. Puis la valeur du voxel correspondant à la position de chaque point de contrôle de chaque ligne de flux, calculée au départ de chaque voxel du champ de vitesse, est incrémentée de un. Ce procédé est ainsi proche de la construction d'un histogramme des points de contrôle des lignes de flux. Chaque voxel de cette matrice S contient alors une valeur qui est égale au nombre de ligne de flux passant à travers ce même voxel. Cela permet par conséquent de cartographier les lignes de flux générées à partir d'un champ de vitesse. Cet histogramme contient alors des valeurs allant de 1 (une seule ligne de flux passe par ce voxel, i.e la ligne de flux démarrant de ce voxel), à une très grande valeur (attracteurs ou repoussoirs, i.e voxels où un très grand nombre de lignes de flux convergent).

La figure 4.8 présente à gauche une visualisation tri-dimensionnelle du champ de vitesse synthétique qui a été considéré jusqu'à maintenant (lignes de flux rouges : point de vue *avant*; lignes de flux bleues : point de vue *arrière*). Les sphères rouges et bleues sont positionnées aux coordonnées des attracteurs et repoussoirs respectivement. Les isosurfaces grises représentent des surdensités dans la matrice S telles que $S > 700$ (les deux points de vues *avant* et *arrière* sont superposés). Une surdensité est observée autour de chaque attracteur. Ces surdensités forment des plans représentant les zones privilégiées par lesquelles passent les lignes de flux. À droite sont représentés deux plans SGX-SGY et SGX-SGZ de l'histogramme des lignes de flux S (point de vue *avant*), centrés respectivement en $SGZ = 0 \text{ Mpc}/h$ et $SGY = 0 \text{ Mpc}/h$. L'échelle des couleurs est logarithmique afin de faciliter la visualisation, la plupart des voxels de S contenant une valeur faible et peu de voxels possédant une valeur élevée. Les lignes de flux (point de vue *avant*) sont aussi représentées en bleu dans deux tranches d'épaisseur $50 \text{ Mpc}/h$ centrées sur les deux plans. Cette figure permet de mettre en évidence les lignes de flux passant par les surdensités de S .

4.2.4 Implémentation en C++

Les algorithmes permettant de calculer les différentes matrices et d'obtenir les résultats présentés précédemment ont été implémentés sous la forme d'un programme en C++ reposant sur la librairie VTK (*Visualization ToolKit*). Cette bibliothèque rassemble des classes et des fonctions permettant la visualisation et l'analyse de données à trois dimensions.

La construction des matrices M_{stop} et S nécessite de calculer les lignes de flux au départ de chaque voxel du champ de vitesse, soit M^3 fois pour un champ de vitesse de dimensions $M \times M \times M$. Pour la matrice M_{stop} , le temps de calcul est environ de 90 minutes pour $M = 128$, et plus de 12 heures pour $M = 256$. Il est d'autant plus long pour la construction de la matrice S qui nécessite, en plus du calcul des lignes de flux pour chaque voxel, une itération sur tous les points de contrôle de chaque ligne de flux. Ainsi, le temps d'exécution est environ de 120 minutes pour $M = 128$, et plus de 15 heures pour $M = 256$.

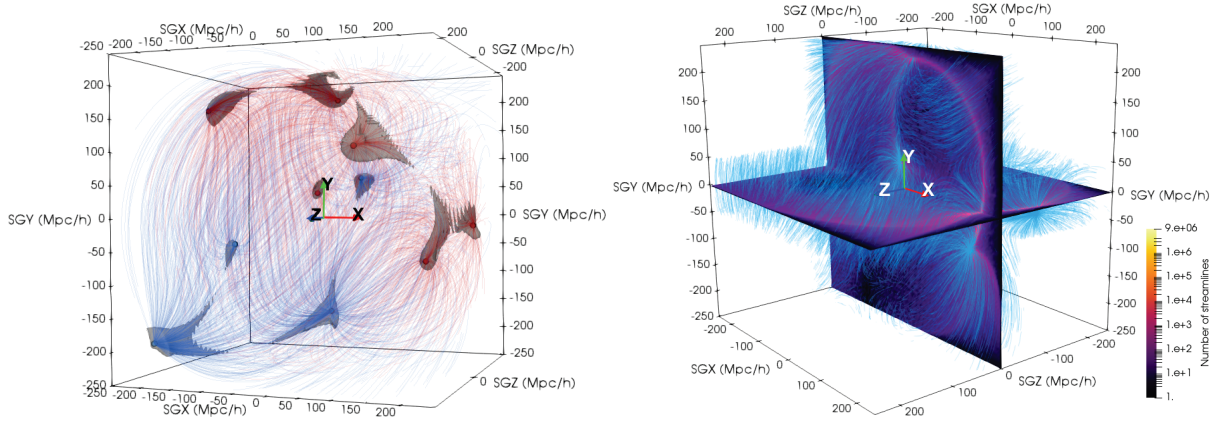


FIGURE 4.8 – Gauche : visualisation tri-dimensionnelle du champ de vitesse synthétique (lignes de flux rouges : point de vue *avant*; lignes de flux bleues : point de vue *arrière*). Les sphères rouges et bleues sont positionnées aux coordonnées des attracteurs et repoussoirs respectivement. Les isosurfaces grises représentent des surdensités dans la matrice S telles que $S > 700$ (points de vues *avant* et *arrière* superposés). Droite : plans SGX-SGY et SGX-SGZ de l'histogramme des lignes de flux S (point de vue *avant*), centrés respectivement en $SGZ = 0$ Mpc/h et $SGY = 0$ Mpc/h. L'échelle des couleurs est logarithmique afin de faciliter la visualisation. Les lignes de flux (point de vue *avant*) sont représentées en bleu dans deux tranches d'épaisseur 50 Mpc/h centrées sur les deux plans.

Les temps d'exécution sont ainsi beaucoup trop longs. L'algorithme nécessite alors d'être parallélisé afin d'accélérer le calcul des lignes de flux et de réduire le temps de calcul. Le programme en C++ a été parallélisé à l'aide de la librairie MPI de la façon suivante. La grille représentant le champ de vitesse est divisée en sous-parties qui sont par la suite distribuées à différents cœurs. Les cœurs calculent ainsi simultanément les lignes de flux au départ de chaque voxels du sous-cube de données qui leur est associé.

Les temps de calcul obtenus après parallélisation pour le calcul des matrices M_{stop} et S sont listés dans la table 4.2 pour différentes tailles de grilles (128 et 256) et différents nombres de cœurs (1 ie sans parallélisation, 8 et 16). Cette table permet de constater que la parallélisation réduit considérablement le temps d'exécution : il est environ six fois plus court pour 8 cœurs, et dix fois plus faible pour 16 cœurs par rapport au temps de calcul sans parallélisation. Cependant, le temps d'exécution reste toujours important pour une grille de dimensions 256^3 et plus.

TABLE 4.2 – Temps de calcul obtenus avant et après parallélisation pour le calcul des matrices M_{stop} et S , pour différentes tailles de grilles M (128 et 256) et différents nombres de cœurs.

Nombre de cœurs	$M_{\text{stop}}, M = 128$	$M_{\text{stop}}, M = 256$	$S, M = 128$	$S, M = 256$
1	≈ 90	> 720	≈ 120	> 900
8	≈ 15	≈ 90	≈ 20	≈ 180
16	< 10	≈ 60	≈ 15	≈ 120

4.3 Simulation

Présentation

La recherche des bassins gravitationnels et la construction d'un histogramme des lignes de flux ont été appliquées au champ de vitesse d'une simulation *Cosmicflows-2*. Cette simulation a été préparée dans le cadre des collaborations *Cosmicflows* et *Constrained Local UniversE Simulations* (CLUES, Yepes et al., 2009; Gottloeber et al., 2010; Courtois and Tully, 2012; Sorce et al., 2016). Ces deux collaborations ont pour objectifs d'étudier la structure de l'univers local à l'aide des vitesses particulières des galaxies. La collaboration CLUES utilise les catalogues de vitesses particulières fournis par la collaboration *Cosmicflows* pour simuler l'univers local et étudier l'histoire de sa formation, et son évolution.

La simulation considérée dans cette section est une simulation de matière noire (non hydrodynamique) fournissant le champ de densité de matière ρ_{simu} et les trois composantes cartésiennes du champ de vitesse particulière \vec{v}_{simu} , calculés chacun dans une boîte de dimensions 256^3 et de largeur 500 Mpc/h, en assumant une cosmologie Λ CDM et en utilisant les données CF2 comme conditions initiales.

La figure 4.9 présente quatre plans SGX-SGY en $\text{SGZ} = 0$ Mpc/h. Les planches (a) et (b) correspondent toutes les deux à la même tranche de densité de la simulation. Une structure filamentaire est observée dans tout le plan. Il est cependant difficile d'identifier visuellement des bassins gravitationnels. Sur la planche (a) sont superposés deux bassins d'attraction, notés A et B et obtenus par la méthode de segmentation décrite en section 4.2.2, alors que deux bassins de repulsion, notés C et D, sont tracés sur la planche (b). Les formes des bassins gravitationnels sont difficiles à comprendre à l'aide du champ de densité seulement. Il faut visualiser le champ de vitesse, et pour cela, l'histogramme des lignes de flux (voir section 4.2.3) obtenu à partir de la simulation est utilisé.

Les planches (c) et (d) de la figure 4.9 représentent deux tranches de l'histogramme des lignes de flux S obtenu en intégrant le champ de vitesse de la simulation en avant et en arrière respectivement. Les bassins d'attraction et de repulsion A, B, C et D sont à nouveau tracés. Les structures des bassins gravitationnels sont ici beaucoup plus visibles et il est possible de deviner à l'oeil nu la forme des différents bassins. Par ailleurs, il est possible d'identifier des régions de haute densité dans la structure filamentaire de ρ correspondant à des surdensités dans l'histogramme des lignes de flux.

Analyse

L'algorithme de segmentation détecte 87 bassins d'attraction, et 76 bassins de répulsion. Il est ainsi possible de remarquer qu'un nombre similaire de bassins d'attraction et de répulsion sont identifiés dans la simulation Λ CDM.

Sur la figure 4.10 est tracé le volume des bassins gravitationnels en fonction de la densité totale enclose à l'intérieur des bassins. Les bassins d'attraction sont représentés par des carrés rouges et les bassins de répulsion par des triangles bleus. Ces deux paramètres sont liés. La distribution des points peut être modélisée par une loi de proportionnalité : $y = 7.5x$ pour les bassins d'attraction, et $y = 7x$ pour les bassins de répulsion, où x et

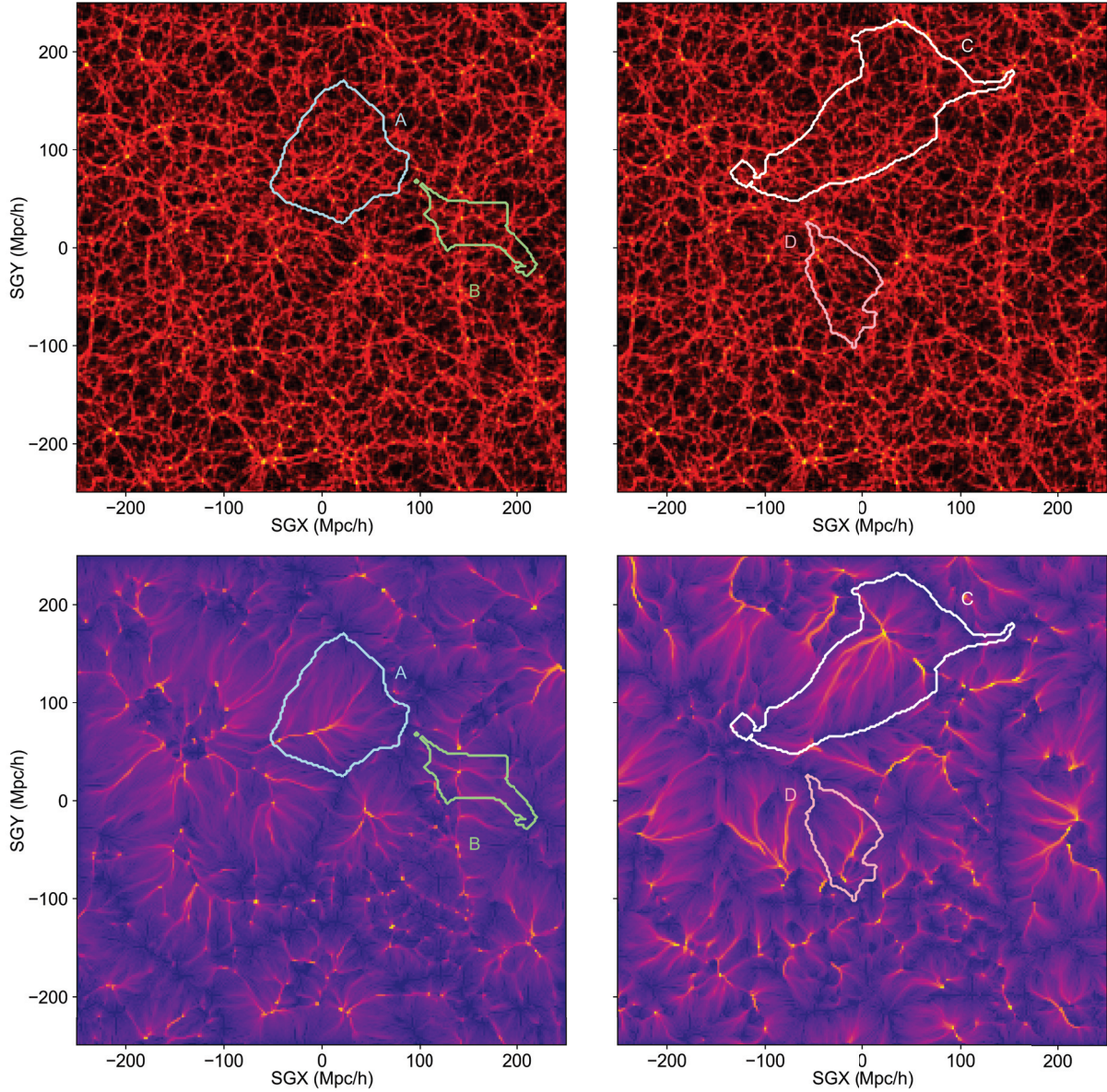


FIGURE 4.9 — Quatre plans SGX-SGY en $SGZ = 0$ Mpc/h d’une simulation CF2. Planches (a) et (b) : tranche de densité sur laquelle sont superposés respectivement deux bassins d’attraction segmentés notés A et B, et deux bassins de repulsion notés C et D. Planches (c) et (d) : deux tranches de l’histogramme des lignes de flux S obtenu en intégrant le champ de vitesse de la simulation en avant et en arrière respectivement. Les bassins d’attraction et de repulsion A, B, C et D sont à nouveau tracés.

y correspondent respectivement aux axes de la figure 4.10, soit la somme de la densité contenue dans un bassin et le volume du bassin. Les bassins gravitationnels A, B et D apparaissant sur la figure 4.9 sont identifiés. Le bassin C n’apparaît pas car il sort du cadre, étant trop gros et trop massif.

La figure 4.11 présente les histogrammes des volumes des bassins d’attraction (gauche) et des bassins de repulsion (droite) segmentés. Le volume typique d’un bassin d’attraction est ainsi situé environ entre 250,000 et 500,000 (Mpc/h)³. Les bassins de repulsions sont en général légèrement plus gros, leur volume moyen étant entre 250,000 et 700,000 (Mpc/h)³.

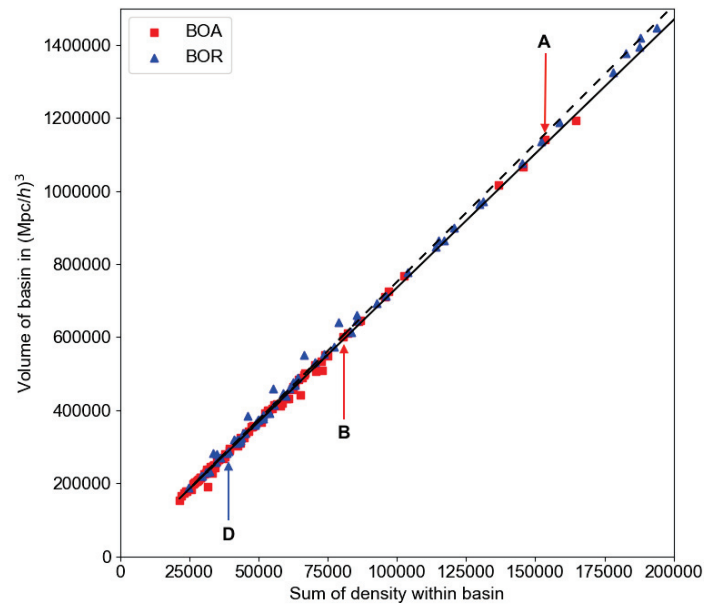


FIGURE 4.10 – Volume des bassins gravitationnels en fonction de la densité totale enclose à l'intérieur des bassins. Les bassins d'attraction sont représentés par des carrés rouges et les bassins de répulsion par des triangles bleus. Les droites en trait plein et traits pointillés noirs modélisent la distribution des points pour les bassins d'attraction et de répulsion respectivement.

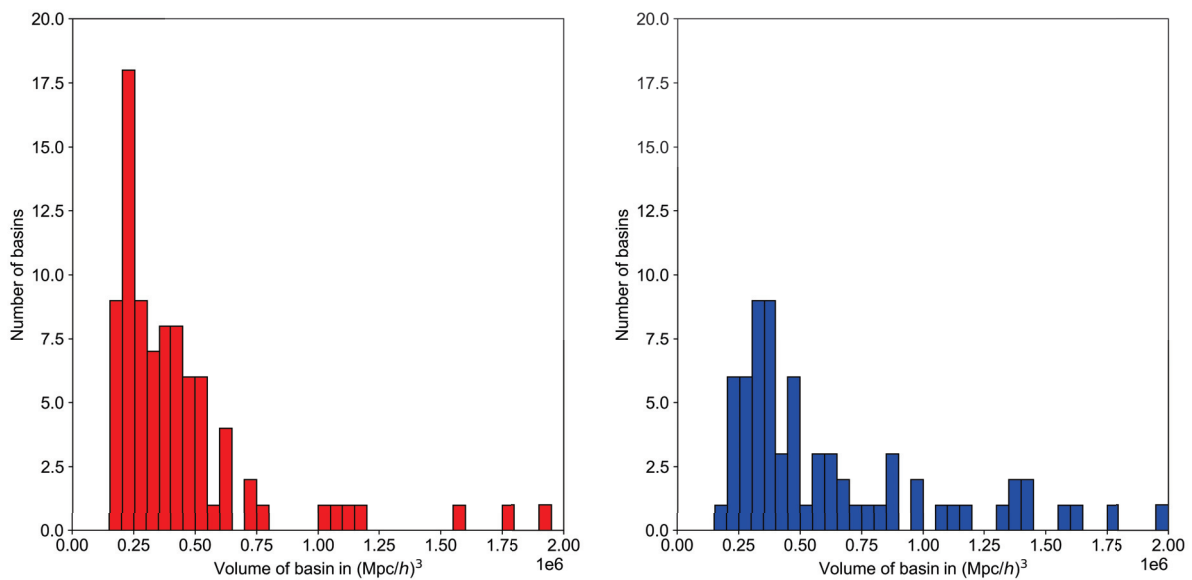


FIGURE 4.11 – Histogramme des volumes des bassins d'attraction (gauche) et des bassins de repulsion (droite) segmentés.

La figure 4.12 présente les profils de densité des bassins d’attraction (gauche) et de répulsion (droite). Les profils de densité des vides ont déjà été étudiés par [de Fromont and Alimi \(2018\)](#); [Diemer and Kravtsov \(2014\)](#). Cependant, c’est la première fois que des profils de densité sont tracés pour des bassins gravitationnels.

Dans les quatre panels, les couleurs représentent le volume des bassins gravitationnels, du plus petit en bleu, au plus grand en rouge. Sur les deux panels du haut, la somme de la densité dans une sphère de rayon r et centrée sur le point de convergence des lignes de flux, est tracée en fonction de r , pour chaque bassin gravitationnel. Un comportement similaire est observé pour les deux types de bassins (attraction et répulsion). Cette figure permet de remarquer que les petits bassins ont des gradients moins forts que les bassins plus gros. Par ailleurs, les plus petits bassins ont un rayon d’environ 50 Mpc/h, tandis que les plus gros bassins sont trois fois plus grand : leur rayon peut aller jusqu’à 150 Mpc/h. Dans les deux panels du bas, la densité est moyennée au lieu d’être sommée. Dans le cas des bassins d’attraction, toutes les courbes atteignent un plateau vers 50 Mpc/h. Tandis que pour les bassins de répulsion, le plateau n’est pas atteint avant 150 Mpc/h.

Les figures 4.13 et 4.14 présentent les profils de l’histogramme des lignes de flux S , calculé en point de vue avant et arrière respectivement. L’organisation de ces figures est la même que la figure 4.12, mais dans ce cas, on considère l’histogramme des lignes de flux S (en point de vue avant ou arrière) au lieu de la densité ρ . Dans le cas des bassins d’attraction, un comportement similaire à la densité ρ (voir figure 4.12) est observé.

La figure 4.15 présente une tranche SGX-SGY d’épaisseur 50 Mpc/h en SGZ. Sur cette tranche sont tracées des isosurfaces rouges de densité telles que $\rho > 0.5$, sur lesquelles sont superposées des isosurfaces jaunes de l’histogramme des lignes de flux construit en point de vue avant et telles que $S > 100$ lignes de flux. Les deux quantités coïncident parfaitement : l’histogramme des lignes de flux forme des *filaments*, qui relient les surdensités de ρ les unes aux autres. L’histogramme est alors relié à une des quantités accessibles par les observations : la densité ρ .

La figure 4.16 correspond à la figure 4.15 sur laquelle sont tracées les bassins A et D (voir figure 4.9) sous forme d’isosurfaces grillagées rouge et bleue respectivement. L’histogramme des lignes de flux coïncide parfaitement avec les deux bassins versants. Ce nouvel outil est donc aussi relié aux bassins gravitationnels.

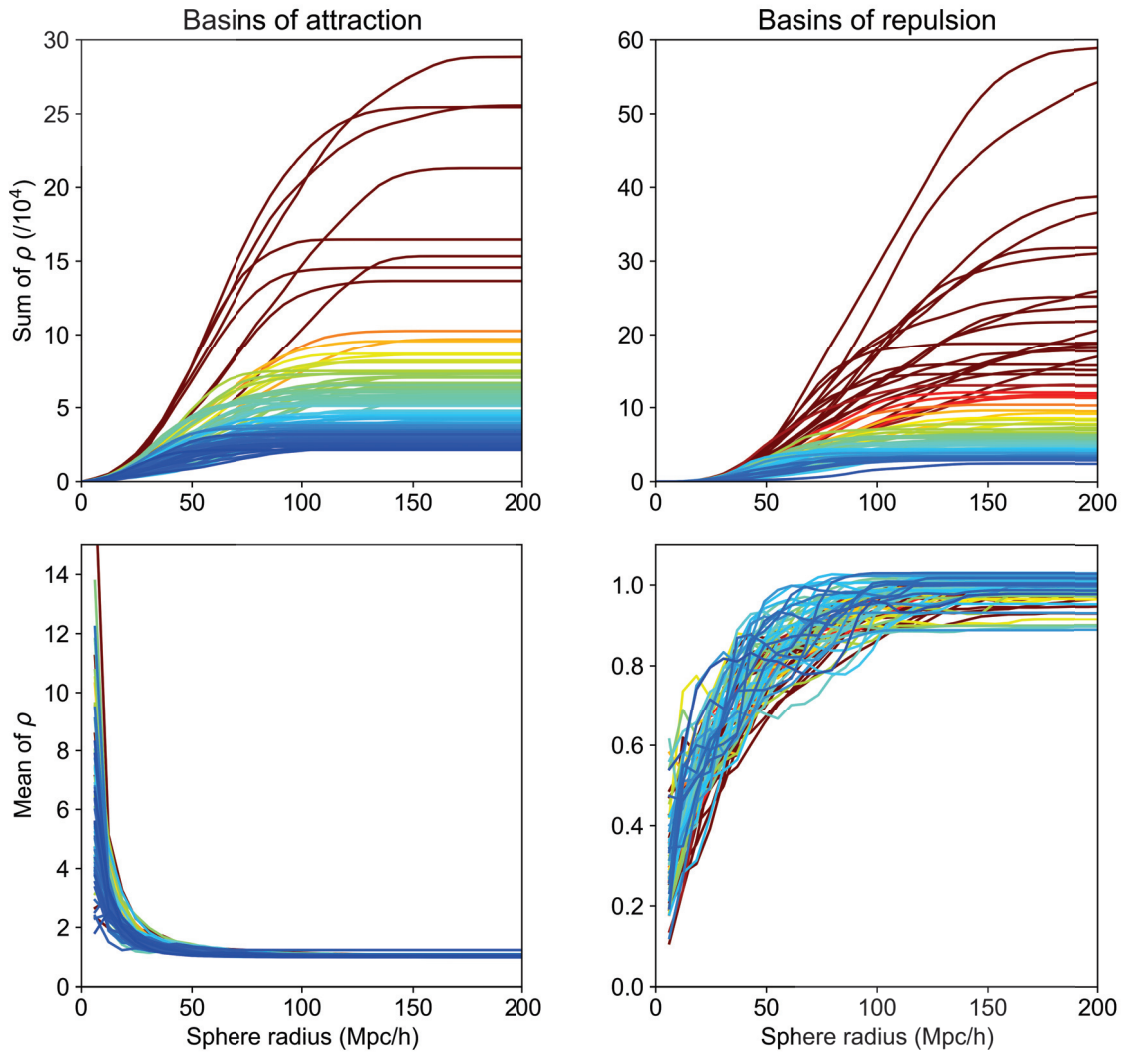


FIGURE 4.12 – Profils de densité des bassins d’attraction (gauche) et de répulsion (droite). La palette de couleurs représente le volume des bassins gravitationnels, du plus petit en bleu, au plus grand en rouge. Haut : somme de la densité dans une sphère de rayon r et centrée sur le point de convergence des lignes de flux, tracée en fonction de r , pour chaque bassin gravitationnel. Bas : ici la densité est moyennée au lieu d’être sommée.

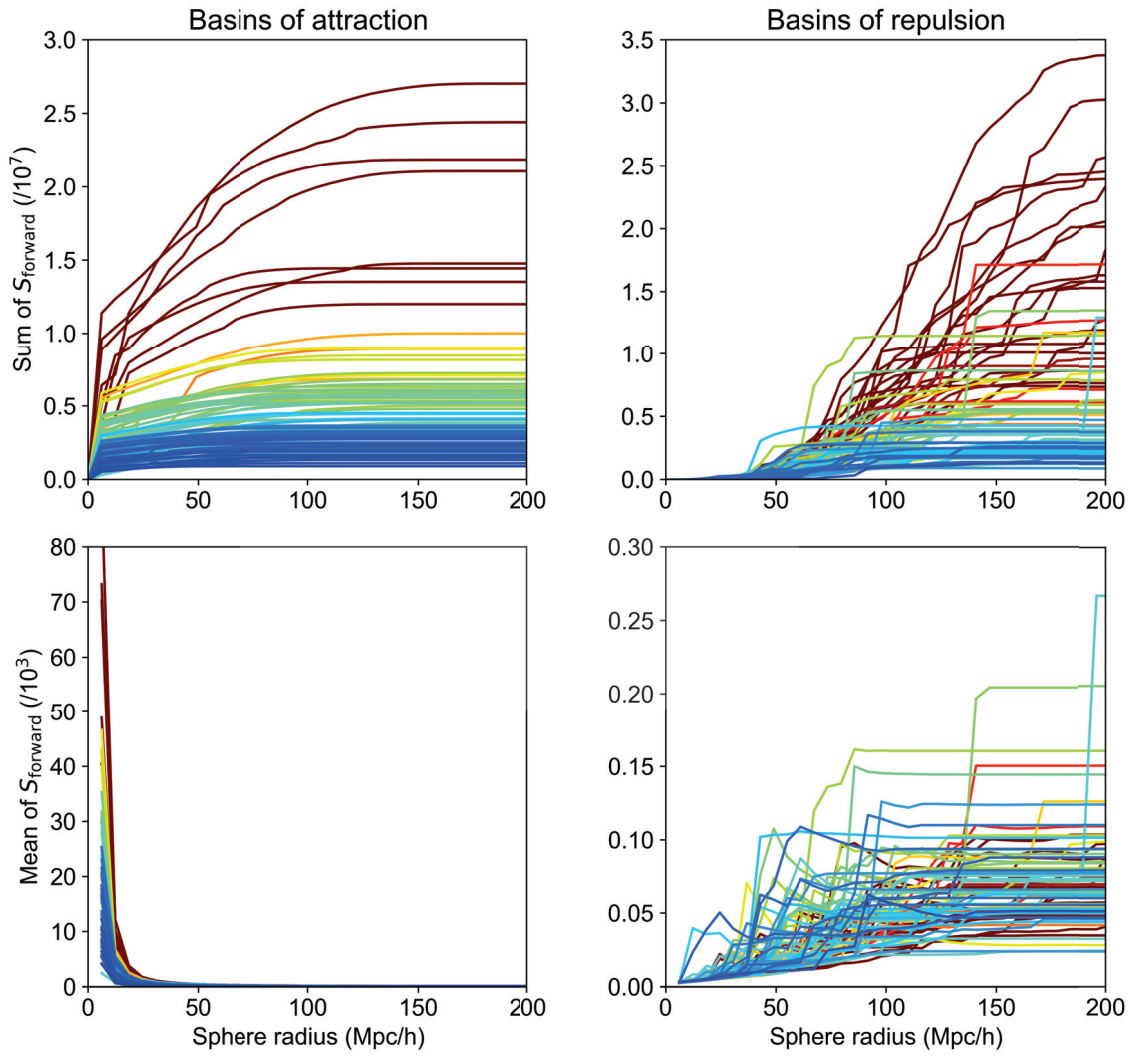


FIGURE 4.13 – Profils de l’histogramme des lignes de flux S , calculé en point de vue avant. Voir figure 4.12 pour la description de l’organisation de la figure.

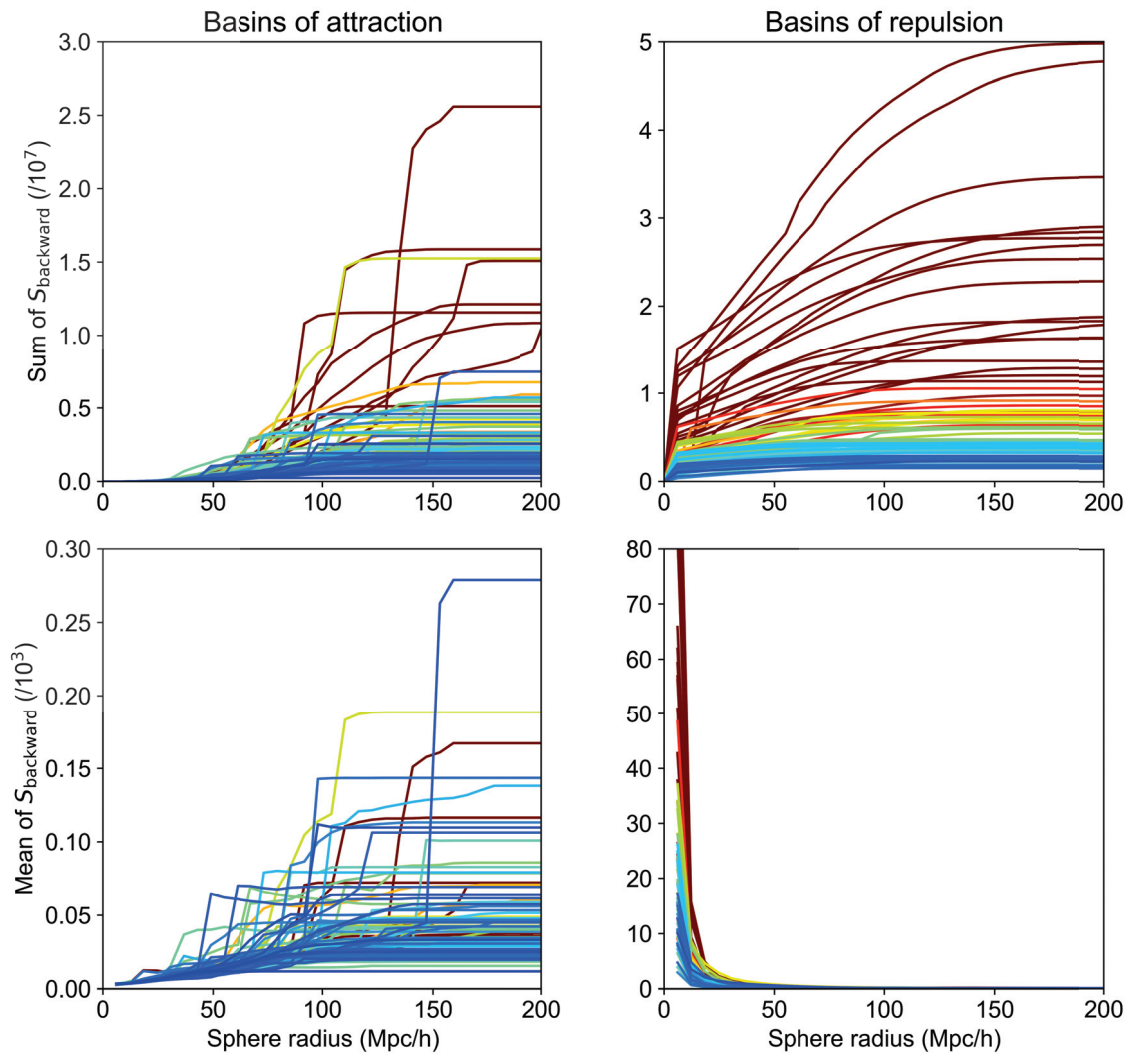


FIGURE 4.14 – Profils de l’histogramme des lignes de flux S , calculé en point de vue arrière. Voir figure 4.12 pour la description de l’organisation de la figure.

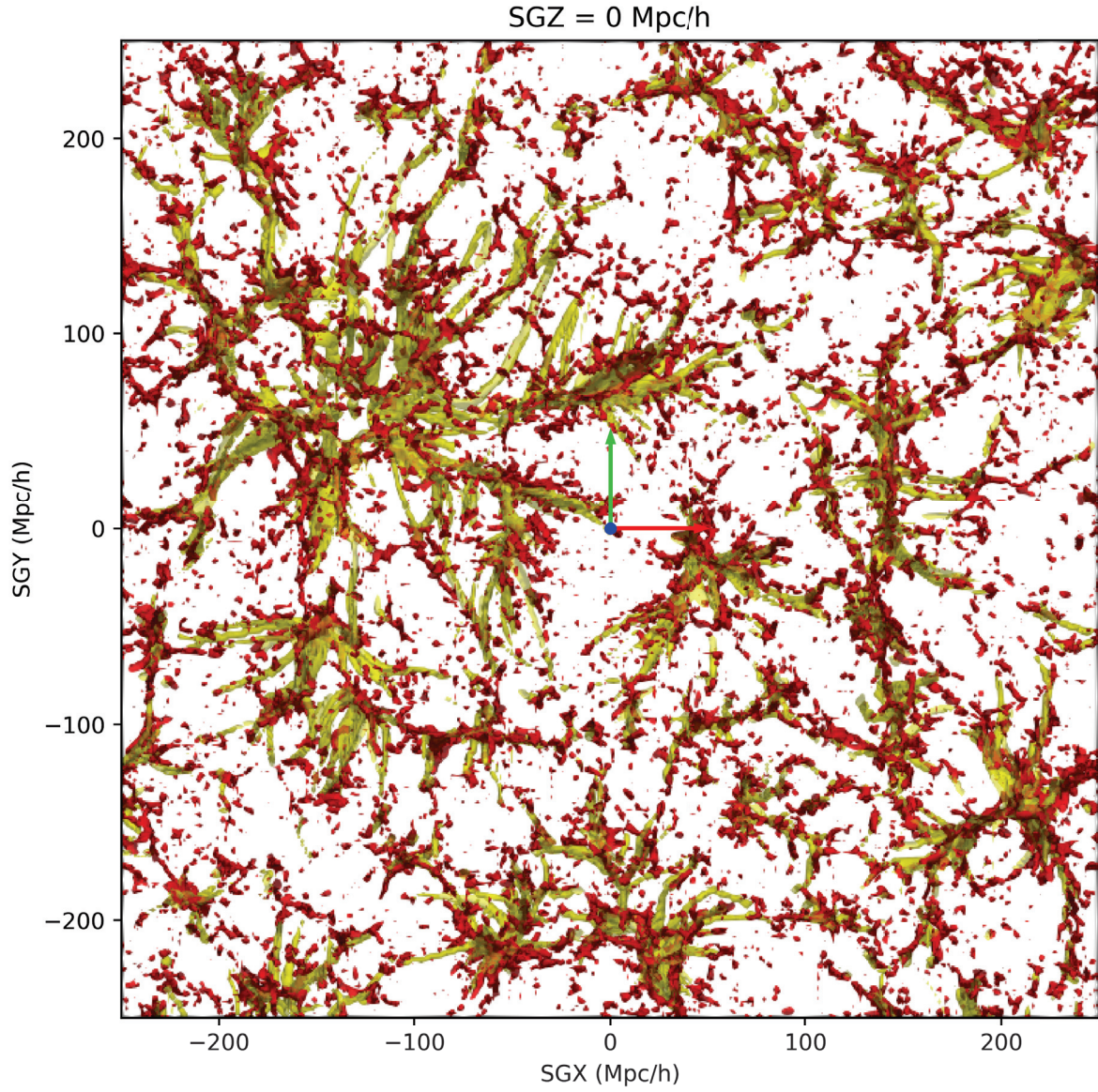


FIGURE 4.15 – Tranche SGX-SGY d'épaisseur $50 \text{ Mpc}/h$ en SGZ. Les isosurfaces rouges sont des isosurfaces de la densité ρ et sont tracées telles que $\rho > 0.5$. Les isosurfaces jaunes sont elles des isosurfaces de l'histogramme des lignes de flux, construit en point de vue avant, et sont telles que $S > 100$ lignes de flux.

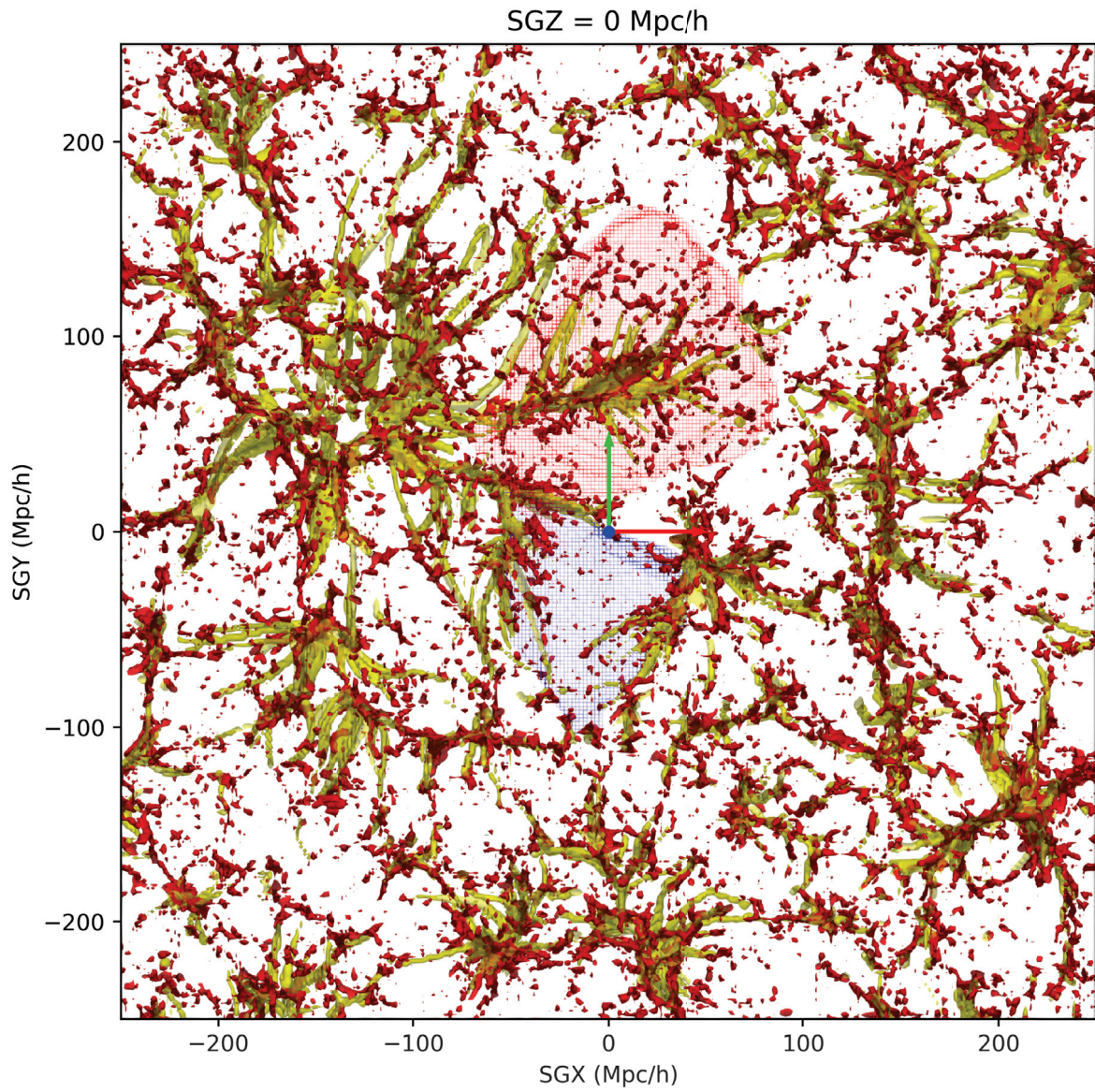


FIGURE 4.16 – Figure 4.15 sur laquelle sont superposés deux bassins gravitationnels sous forme d'isosurfaces grillagées : le bassin d'attraction A (en rouge) et le bassin de répulsion D (en bleu), apparaissant sur la figure 4.9.

4.4 Application aux données *Cosmicflows*

Après validation sur des données synthétiques, les algorithmes ont été appliqués sur les champs de vitesse reconstruits à partir des données observationnelles *Cosmicflows*. Nous avons tout d'abord commencé par effectuer un test sur le champ de vitesse *Cosmicflows-2*. Ensuite, les méthodes ont été mises en pratique sur le tout dernier champ de vitesse reconstruit à partir des données *Cosmicflows-3*.

4.4.1 Champs de surdensité et de vitesse *Cosmicflows*

Dans cette partie sont présentés et comparés les champs de surdensité et de vitesse particulière à trois dimensions reconstruits à partir des données groupées CF2 et CF3, avec deux méthodes différentes (filtre de Wiener pour CF2 et échantillonnage de Gibbs pour CF3). Les champs de vitesses CF2 et CF3 sont préparés respectivement par [Hoffman et al. \(2015\)](#), avec la méthode du filtre de Wiener, et [Graziani et al. \(2018\)](#), à l'aide de l'échantillonnage de Gibbs (voir section 4.1). Il sera vu dans le chapitre que la reconstruction CF3 n'est pas finalisée, et qu'il existe différentes versions, selon la constante de Hubble choisie, la prise en compte ou non des outliers, etc. Les figures et résultats qui suivent ont été obtenus à partir de la dernière version disponible à ce jour.

La figure 4.17 montre différentes tranches des champs de surdensité reconstruits à partir des données CF2 (gauche) et CF3 (droite). De haut en bas sont présentées une tranche SGX-SGY centrée en $SGZ = 0 \text{ Mpc}/h$, une tranche SGX-SGZ centrée en $SGY = -70 \text{ Mpc}/h$ et une tranche SGY-SGZ centrée en $SGX = 0 \text{ Mpc}/h$. Les zones de surdensité (amas) sont représentées en rouge tandis que les zones de sous-densité (vides) sont représentées en bleu. Les noms des différents amas associés aux surdensités sont affichés sur les différentes tranches. Le champ de vitesse est tracé par des flèches grises qui pointent vers les surdensités. Ces tranches permettent de constater l'évolution entre les données CF2 et CF3. La différence entre les deux reconstructions est clairement vue sur la tranche SGX-SGZ, sur laquelle la surdensité CF2 est très faible, alors qu'un grand nombre d'amas apparaissent sur la reconstruction CF3. De nouveau amas et des structures lointaines pour lesquels la surdensité CF2 était très faible sont clairement mis en évidence par les données CF3 : Sculptor Wall, Pisces-Cetus, Lepus, Hercules, Apus, Telescopium, Horologium-Reticulum...

La figure 4.18 présente une tranche SGX-SGY centrée en $-25 < SGZ < 25 \text{ Mpc}/h$ de la surdensité de la reconstruction CF3. Les galaxies provenant du sondage de redshift 2MRS (*2MASS Redshift Survey*, [Huchra et al., 2012](#)) sont superposées sous forme de points noirs. La concordance surdensité - galaxies est parfaite au niveau de chaque amas.

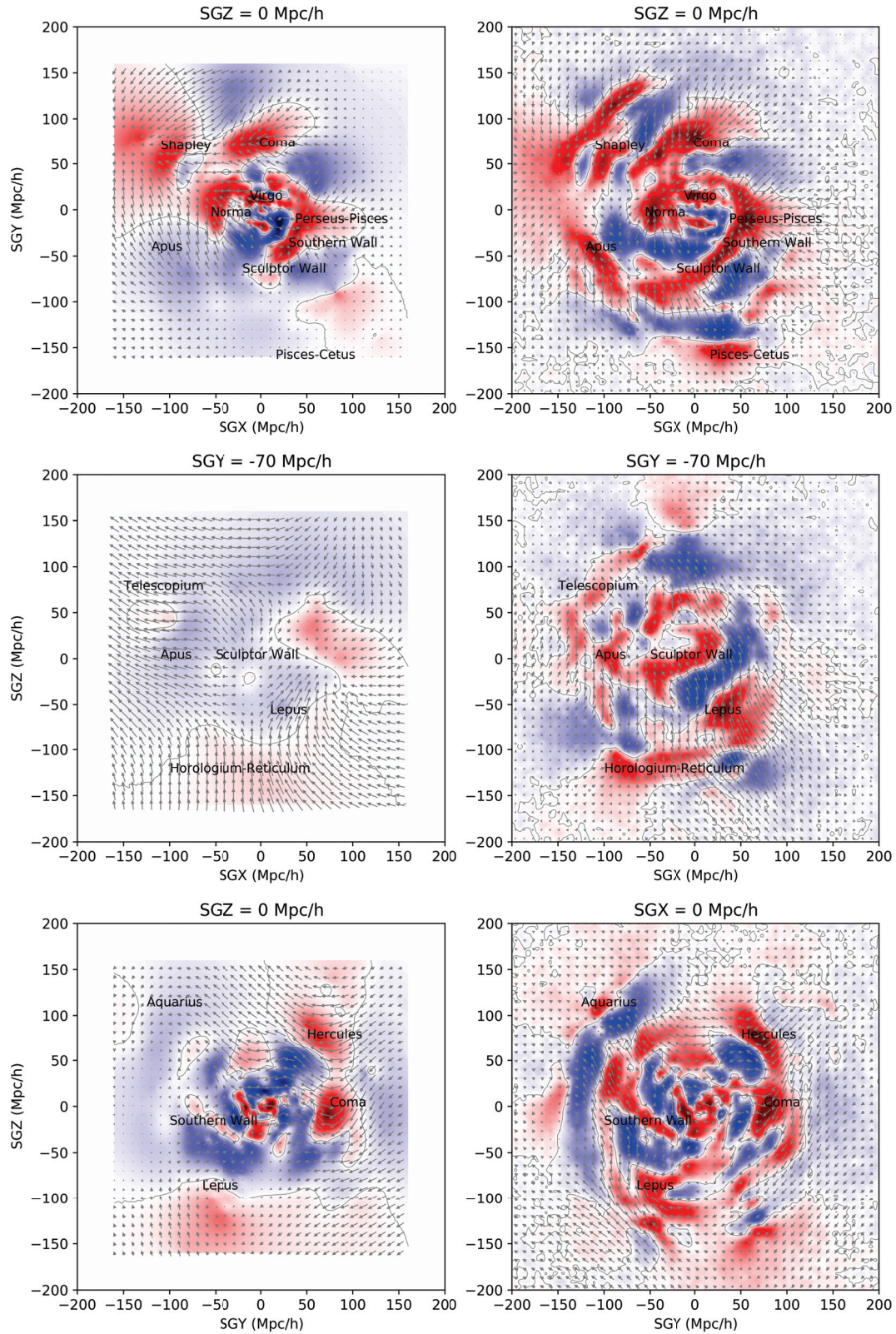


FIGURE 4.17 – Tranches des champs de surdensité reconstruits à partir des données CF2 (gauche) et CF3 (droite). De haut en bas : tranche SGX-SGY centrée en $SGZ = 0$ Mpc/h, tranche SGX-SGZ centrée en $SGY = -70$ Mpc/h et tranche SGY-SGZ centrée en $SGX = 0$ Mpc/h. Le champ de vitesse est tracé par des flèches grises. Les noms des différents amas associés aux surdensités sont affichés sur les différentes tranches. Voir la figure 4.18 pour la palette de couleurs.

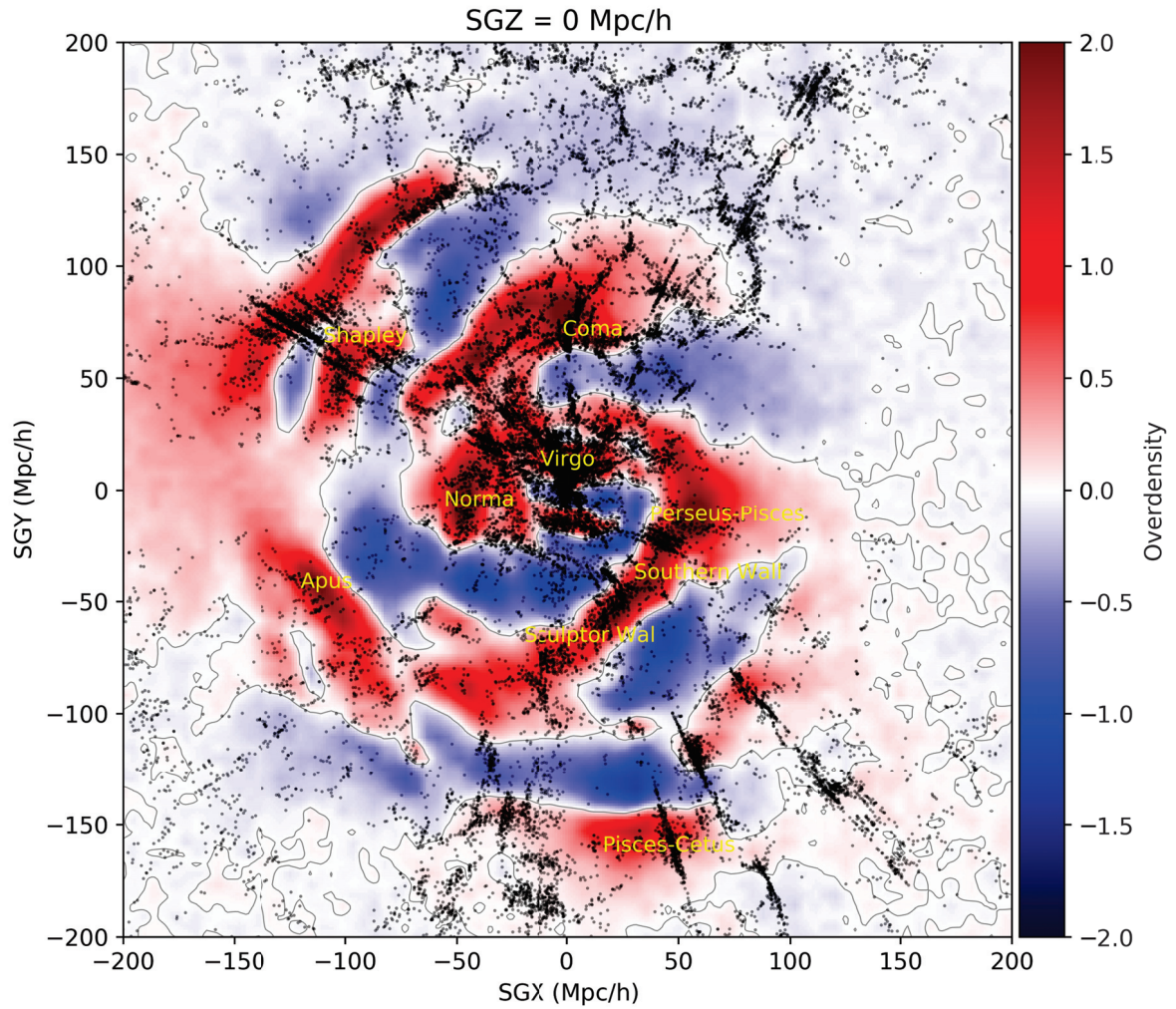


FIGURE 4.18 – Tranche SGX-SGY centrée en $-25 < SGZ < 25$ Mpc/h de la surdensité de la reconstruction CF3. Les galaxies provenant du sondage de redshift 2MRS sont superposées sous forme de points noirs.

4.4.2 Cosmicflows-2 : notre superamas Laniakea

Cette section présente les résultats obtenus lors de l'application de l'algorithme de segmentation au champ de vitesse reconstruit à partir du catalogue observationnel CF2. Le but ici est de vérifier le résultat de Tully et al. (2014) en retrouvant automatiquement le bassin d'attraction de Laniakea, le superamas dont nous faisons partie. Ceci permettrait ainsi de valider notre algorithme sur des données observationnelles. Le bassin d'attraction de Laniakea a été obtenu par Tully et al. (2014) en cherchant la *ligne de partage des eaux*, i.e la surface à laquelle le champ de vitesse diverge (Pomarède et al., 2017).

Champ de vitesse local autour de Norma

Pour y parvenir, le champ de vitesse qui a été calculé pour Tully et al. (2014) pour mettre en évidence le bassin d'attraction de Laniakea a été utilisé. Ce champ de vitesse correspond au champ de vitesse local autour de Norma (centre approximatif de Laniakea).

En effet, le bassin de Laniakea n'est pas visible dans le champ de vitesse total reconstruit à partir des données CF2 avec la méthodologie du filtre de Wiener. Ceci est dû à la composante de marée (à longue portée) qui est provoquée par les grandes structures voisines. Afin d'obtenir le bassin d'attraction de Laniakea, il faut décomposer le champ de vitesse total en deux composantes : un champ de vitesse local et une composante due aux surdensités extérieures. Ceci se fait de la manière suivante. Considérons seulement le champ de surdensité δ_{DIV} issu du filtre de Wiener et contenu dans une sphère de rayon R_{DIV} et de centre C_{DIV} . Il est possible de déduire le champ de vitesse local \vec{v}_{DIV} associé à ce champ de surdensité *isolé*. Pour ce faire, il suffit de résoudre l'équation de Poisson suivante, reliant le champ de surdensité au champ de vitesse :

$$\nabla \cdot \vec{v}_{\text{DIV}} = -H_0 f \delta_{\text{DIV}}. \quad (4.23)$$

Le champ de vitesse local considéré dans Tully et al. (2014) a été reconstruit à partir du champ de surdensité contenu dans une sphère de rayon 100 Mpc/h et centrée en $(-51, 4, 5)$ Mpc/h. Les coordonnées et valeurs choisies pour le centre et le rayon de la sphère sont celles permettant d'isoler au mieux le bassin d'attraction de Laniakea des structures voisines telles que Coma, Persée-Poissons et Shapley.

Resultats

Une fois appliqué au champ de vitesse local autour de Norma, l'algorithme de segmentation détecte un attracteur à la position $(\text{SGX}, \text{SGY}, \text{SGZ}) = (-63.75, 20, 13.75)$ Mpc/h, soit $(\text{SGL}, \text{SGB}) = (163^\circ, 12^\circ)$ et $cz = 6821$ km/s. Ceci coïncide avec la direction du Grand Attracteur, qui correspond à l'attracteur associé au bassin versant du superamas Laniakea.

Ainsi, le bassin d'attraction de Laniakea est segmenté automatiquement sans intervention particulière. Celui-ci est représenté sur la figure 4.19 haut, qui montre une tranche SGX-SGY en SGZ = 0 Mpc/h du champ de surdensité reconstruit à partir des données CF2.

Le champ de vitesse local autour de Norma est représenté par des lignes de flux noires. Le bassin d'attraction segmenté de Laniakea est tracé en trait plein vert. Il est comparé à celui obtenu par Tully et al. (2014) qui est lui représenté en jaune. Les lignes de flux appartenant au bassin gravitationnel de Laniakea segmenté sont mises en valeur et tracées en vert. Les deux bassins (celui reconstruit automatiquement par segmentation et celui de Tully et al. (2014)) sont presque identiques. Ils se chevauchent parfaitement vers les différents attracteurs connus (Persée-Poissons, Coma, Shapley), et diffèrent seulement dans une zone peu contrainte par les données, vers $(SGX, SGY) = (-40, -75)$ Mpc/h.

La différence entre les deux bassins est mise en évidence sur la figure 4.19 bas. Celle-ci présente une visualisation tri-dimensionnelle des deux bassins d'attraction de Laniakea : celui obtenu par la méthode de segmentation est représenté par une isosurface verte, et celui de Tully et al. (2014) en jaune. Les lignes de flux appartenant au bassin d'attraction de Laniakea sont tracées en noir. Le bassin d'attraction fourni par Tully et al. (2014) est coupé et n'inclut pas toutes les lignes de flux convergeant vers le Grand Attracteur, contrairement au bassin reconstruit automatiquement par l'algorithme de segmentation. Le bassin d'attraction segmenté est donc plus correct que celui obtenu par les auteurs.

Par ailleurs, le bassin d'attraction de Laniakea segmenté automatiquement (en vert sur la figure 4.19) enclot un volume de 2.3×10^6 (Mpc/h)³, tandis que celui fourni par Tully et al. (2014) enclot un volume de 1.7×10^6 (Mpc/h)³.

Par conséquent, l'algorithme de segmentation reconstruit le bassin d'attraction de Laniakea avec succès.

Incertitudes sur les bassins versants

En appliquant l'algorithme de segmentation à des réalisations contraintes reconstruites à partir des données observationnelles, il est possible d'obtenir des barres d'erreurs sur les positions des attracteurs et les bassins gravitationnels associés.

Afin d'obtenir des barres d'erreur sur le bassin d'attraction de Laniakea, un ensemble de 20 réalisations contraintes reconstruites avec le catalogue CF2 a été utilisé. Afin d'obtenir des réalisations ressemblant aux champs de surdensité et de vitesse reconstruits par la méthode du filtre de Wiener à partir des données CF2, un lissage Gaussien est appliqué aux grilles contenant les champs de surdensité des 20 CR, avec un facteur de lissage dépendant de l'éloignement par rapport au centre : 5 Mpc/h jusqu'à 10 Mpc/h (Virgo), 10 Mpc/h jusqu'à 80 Mpc/h (Coma), 35 Mpc/h jusqu'à 130 Mpc/h (Shapley) et 50 Mpc/h jusqu'à 160 Mpc/h (Repoussoir du Dipole, Hoffman et al., 2017). Ce lissage permet d'obtenir un champ de vitesse ressemblant aux données observationnelles : plus la distance par rapport au centre de la grille est grande, moins il y a de données disponibles, donc plus le facteur de lissage est grand. Une méthode de lissage plus appropriée pourrait être développée dans le cadre des données CF3, lorsque les réalisations contraintes de CF3 seront disponibles.

Vingt champs de vitesse sont ensuite déduits des champs de surdensité lissés à l'aide de l'équation de Poisson reliant le champ de vitesse au champ de surdensité. C'est sur ces 20 champs de vitesse qu'est appliqué l'algorithme de segmentation. Un ensemble d'attracteurs associés à des bassins d'attraction est obtenu. Afin d'identifier les bassins d'attraction correspondant à Laniakea (et les attracteurs associés), les bassins dont les attracteurs

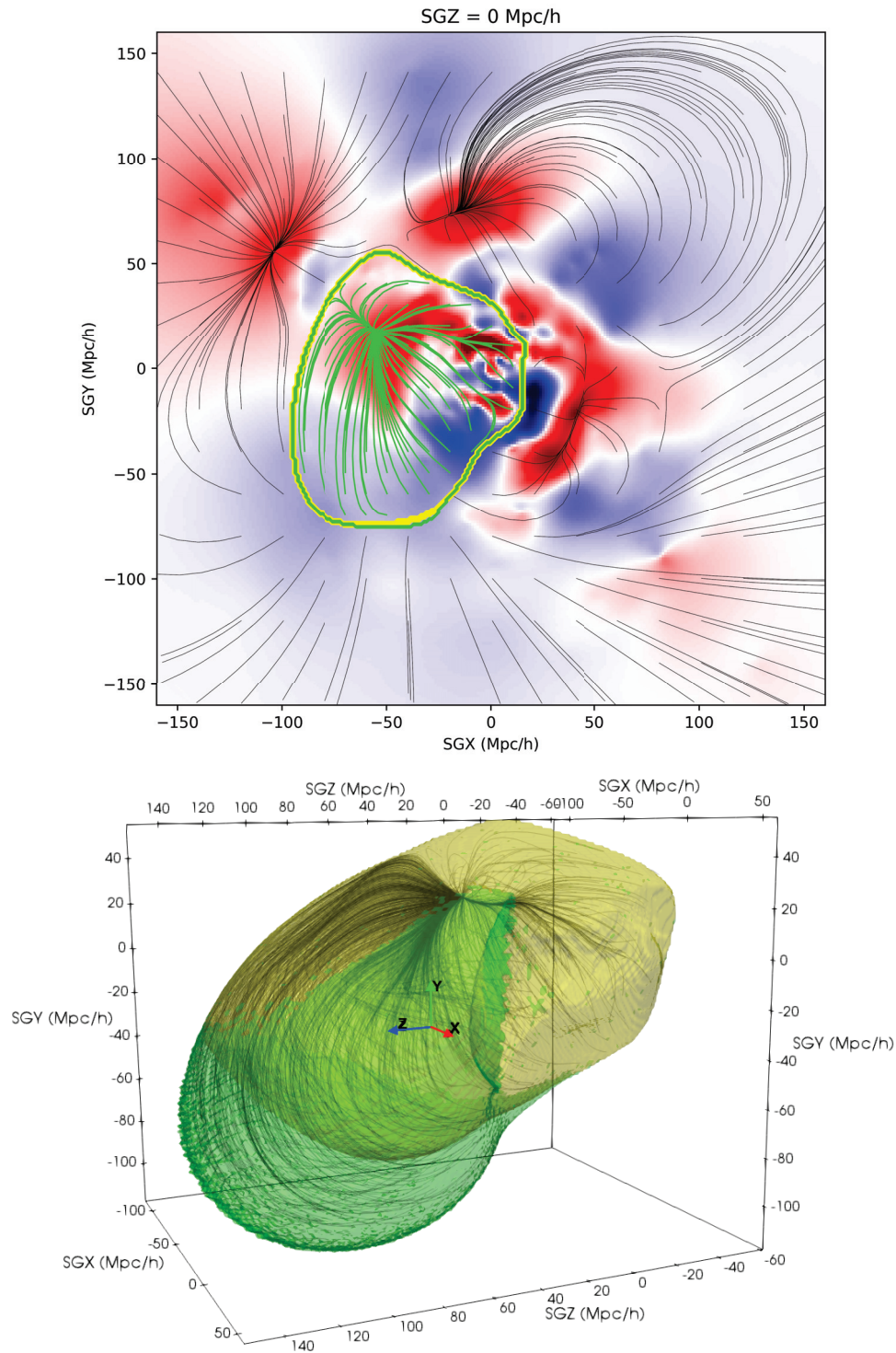


FIGURE 4.19 – Haut : tranche SGX-SGY en $SGZ = 0 \text{ Mpc/h}$ du champ de surdensité CF2. Le champ de vitesse local autour de Norma est représenté par des lignes de flux noires. Le bassin d’attraction segmenté de Laniakea est tracé en trait plein vert, et celui obtenu par Tully et al. (2014) en jaune. Les lignes de flux appartenant au bassin gravitationnel de Laniakea segmenté sont mises en valeur et tracées en vert. Bas : visualisation tri-dimensionnelle des deux bassins d’attraction de Laniakea : celui obtenu par la méthode de segmentation est représenté par une isosurface verte, et celui de Tully et al. (2014) en jaune. Les lignes de flux appartenant au bassin d’attraction de Laniakea sont tracées en noir.

associés sont situés dans une sphère centrée autour de la position du Grand Attracteur (détectée sur le champ de vitesse CF2, voir section 4.4.2) sont sélectionnés et identifiés comme bassins d'attraction de Laniakea.

Les résultats obtenus sont présentés sur la figure 4.20. Cette figure montre une tranche SGX-SGY en $SGZ = 0$ Mpc/h du champ de surdensité reconstruit par la méthode du filtre de Wiener à l'aide des données CF2. Les bassins d'attraction de Laniakea obtenus sur ces mêmes données par Tully et al. (2014) et par la méthode de segmentation sont tracés en trait plein jaune et vert respectivement. Les croix noires sont positionnées aux coordonnées des attracteurs associés aux bassins d'attraction détectés dans les 20 CR et identifiés comme bassins d'attraction de Laniakea. Ces bassins d'attraction sont eux représentés par des contours noirs.

Il est ainsi possible, à l'aide des réalisations contraintes reconstruites à l'aide de données observationnelles, d'extraire des barres d'erreurs *visuelles* sur les bassins gravitationnels, ce qui n'a pas été fait dans Tully et al. (2014).

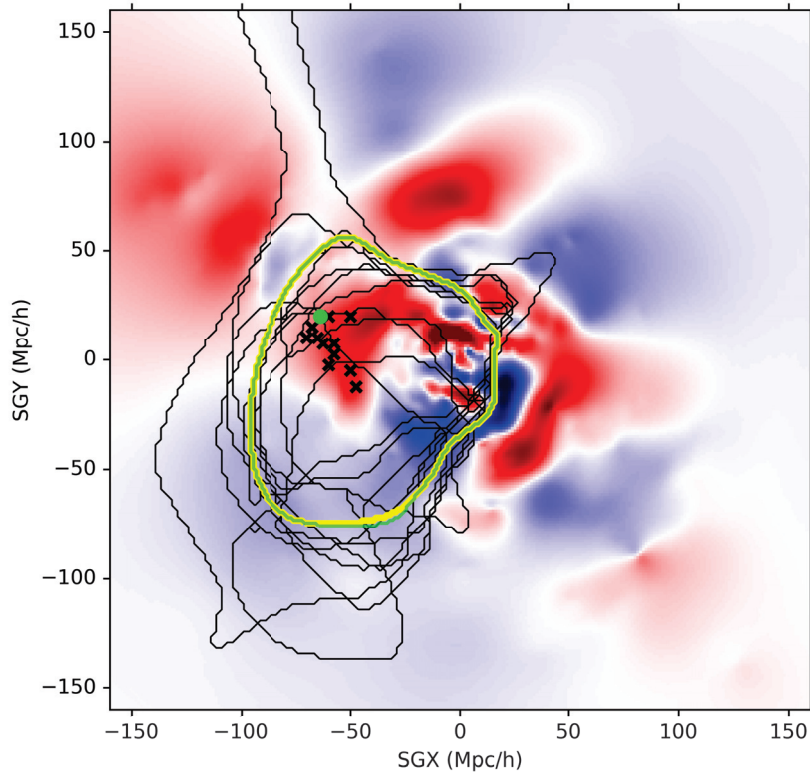


FIGURE 4.20 – Tranche SGX-SGY en $SGZ = 0$ Mpc/h du champ de surdensité reconstruit par la méthode du filtre de Wiener à l'aide des données CF2. Les bassins d'attraction de Laniakea obtenus sur ces mêmes données par Tully et al. (2014) et par la méthode de segmentation sont tracés en trait plein jaune et vert respectivement. Les croix noires sont positionnées aux coordonnées des attracteurs associés aux bassins d'attraction détectés dans les 20 CR et identifiés comme bassins d'attraction de Laniakea. Ces bassins d'attraction sont eux représentés par des contours noirs.

4.4.3 Cosmicflows-3 : résultats préliminaires

Cette dernière section de ce chapitre est dédiée aux résultats obtenus à partir des données *Cosmicflows-3*. Ces résultats sont cependant préliminaires, le champ de vitesse reconstruit à partir du catalogue CF3 n'étant pas finalisé.

La figure 4.21 présente l'évolution des lignes de flux en fonction du nombre d'itérations. Les lignes de flux sont intégrées en avant sur le champ de vitesse CF3. Les trois panels correspondent à trois nombres d'itérations différents, soit, de gauche à droite : 40, 70 et 100 itérations. Les lignes de flux s'étendent au fur et à mesure, et les points de convergence (ici les attracteurs) sont de plus en plus visibles. Par la suite, les lignes de flux sont calculées avec 100 itérations, pour que les lignes de flux démarrant des voxels les plus éloignés puissent rejoindre leur destination finale. Il serait cependant intéressant d'ajouter une condition d'arrêt supplémentaire, par exemple sur la vitesse, pour que les lignes de flux soient automatiquement stoppées lorsqu'elles approchent un attracteur ou un repoussoir.

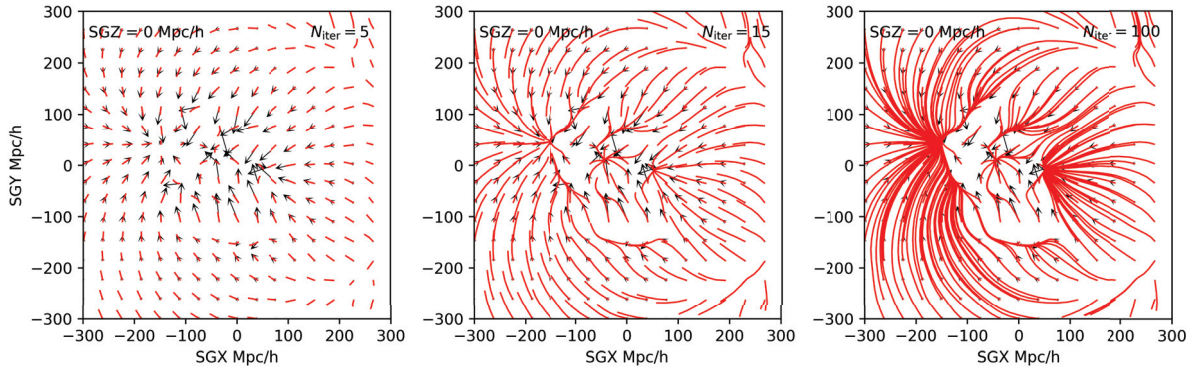


FIGURE 4.21 – Évolution des lignes de flux, intégrées en avant sur le champ de vitesse CF3, en fonction du nombre d'itérations. De gauche à droite : 40, 70 et 100 itérations.

La figure 4.21 droite est présentée plus en détails sur la figure 4.22 haut. Celle-ci montre une tranche SGX-SGY, d'épaisseur $-25 < \text{SGZ} < 25$ Mpc/h, sur laquelle sont tracées les lignes de flux de la figure 4.21 droite, et des isosurfaces de surdensité telles que $\delta > 0.8$ (gris) et $\delta > 1.5$ (rouge). La palette de couleur, utilisée pour colorer les lignes de flux, représente la valeur de la vitesse. Cela permet de remarquer que la vitesse diminue à l'approche d'un point attracteur ou repoussoir. La figure 4.22 bas est identique au panel du haut, mais les lignes de flux sont cette fois-ci intégrées en arrière, afin de visualiser les repoussoirs. Dans les deux cas, la correspondance entre les lignes de flux et les surdensités est parfaite.

Les attracteurs détectés par l'algorithme de segmentation, appliqué sur le champ de vitesse CF3, sont listés dans la table 4.3. Leur nom est donnée en colonne (1), les coordonnées supergalactiques cartésiennes SGX, SGY, SGZ sont données dans les colonnes (2), (3) et (4) respectivement, et les volumes des bassins d'attraction segmentés sont listés en colonne (5). Par ailleurs, les bassins d'attraction obtenus sont présentés dans une tranche SGX-SGY d'épaisseur $-25 < \text{SGZ} < 25$ Mpc/h en figure 4.23. Les isosurfaces noires correspondent aux surdensités telles que $\delta > 1.5$, et les lignes de flux sont tracées en point de vue avant. Sur cette figure, cinq bassins gravitationnels sont visibles : les bassins de

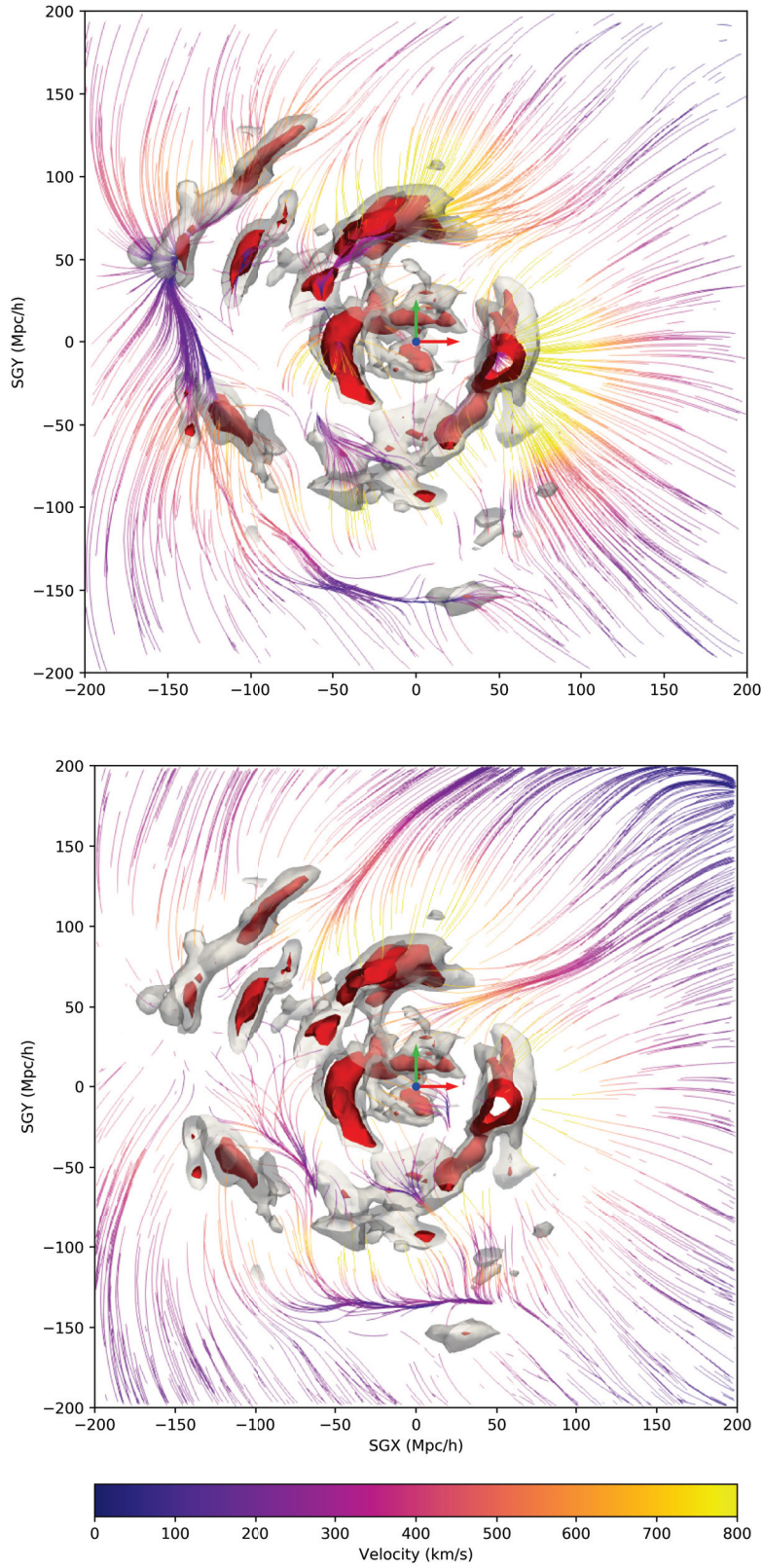


FIGURE 4.22 — Tranche SGX-SGY telle que $-25 < SGZ < 25$ Mpc/h. Les lignes de flux et des isosurfaces de surdensité telles que $\delta > 0.8$ (gris) et $\delta > 1.5$ (rouge), obtenues à partir de la reconstruction CF3, sont tracées. La palette de couleur utilisée pour colorer les lignes de flux représente la valeur de la vitesse. Haut : lignes de flux obtenues en intégrant en avant. Bas : lignes de flux obtenues en intégrant en arrière.

TABLE 4.3 – Attracteurs détectés par l’algorithme de segmentation, appliqué sur le champ de vitesse CF3.

Nom	SGX	SGY	SGZ	V_{BOA}
superamas	Mpc/h	Mpc/h	Mpc/h	(Mpc/h) ³
Hercules	-46.88	-18.75	-98.44	2.8×10^7
Persée-Poissons	42.19	-9.38	-32.81	1.6×10^7
Shapley I (loin)	-150.00	46.88	-14.06	1.1×10^8
Shapley II (proche)	-107.81	42.19	4.69	4.4×10^6
Laniakea	-51.56	4.69	9.38	3.5×10^6
Coma	-60.94	18.75	23.44	3.2×10^7

Shapley I en bleu, Shapley II en rouge, Laniakea en vert, Coma en rose, et Persée-Poisson en jaune. Les petites sphères bleues sont positionnées aux coordonnées des attracteurs détectés automatiquement par l’algorithme de segmentation et listées dans la table 4.3.

Le superamas Shapley est divisé en deux bassins : un bassin plus proche et plus petit, et un immense bassin situé juste derrière. Le bassin gravitationnel de notre superamas Laniakea inclut une autre structure vers $(X, Y) = (0, -100)$ Mpc/h approximativement. De plus, les bassins d’attraction ne sont pas *fermés*, et incluent des voxels très éloignés, jusqu’aux bords de la boîte. Pour cette raison, les volumes listés en table 4.3 ne sont pas fiables, faussés par ces effets de bord.

À la recherche de Vela ?

Vela est une surdensité située dans la zone d’obscurité (*Zone of Avoidance*, ZOA) causée par la Voie Lactée, dans l’hémisphère terrestre Sud. Sa position la rend par conséquent très difficile à être identifiée directement à partir des données observationnelles, plus particulièrement à l’aide de vitesses particulière (les étoiles de notre galaxie rendant la photométrie, et ainsi la mesure d’une distance par la relation de Tully-Fisher impossible). Kraan-Korteweg et al. (2017) ont cependant mis en évidence un superamas dans la direction de Vela, à l’aide de mesures de décalages spectraux de galaxies situées dans la direction de Vela ($245^\circ \lesssim l \lesssim 295^\circ$ et $|b| \lesssim 10^\circ$).

La figure 4.24 présente quatre tranches SGX-SGY d’épaisseur $-25 < \text{SGZ} < 25$ Mpc/h. Sur ces tranches sont superposées les galaxies d’une compilation de redshifts (jusqu’à $cz < 30,000$ km/s, extraite de la base de données LEDA (<http://leda.univ-lyon1.fr/>)) sous forme de petits points noirs ; et les galaxies du sondage de redshift de Kraan-Korteweg et al. (2017) sous forme de points noirs légèrement plus gros, regroupés de chaque côté de la ZOA. Par souci de visibilité, les points sont tracés en blanc dans les deux panels de droite.

À gauche, la palette de couleur représente le champ de surdensité δ . Une très légère surdensité est visible entre les deux groupes de galaxies du catalogue de Kraan-Korteweg et al. (2017), en direction de la ZOA. En effet, Vela se situe à la fois dans une zone cachée et à la limite des données CF3. Il y a donc très peu de données disponibles dans cette

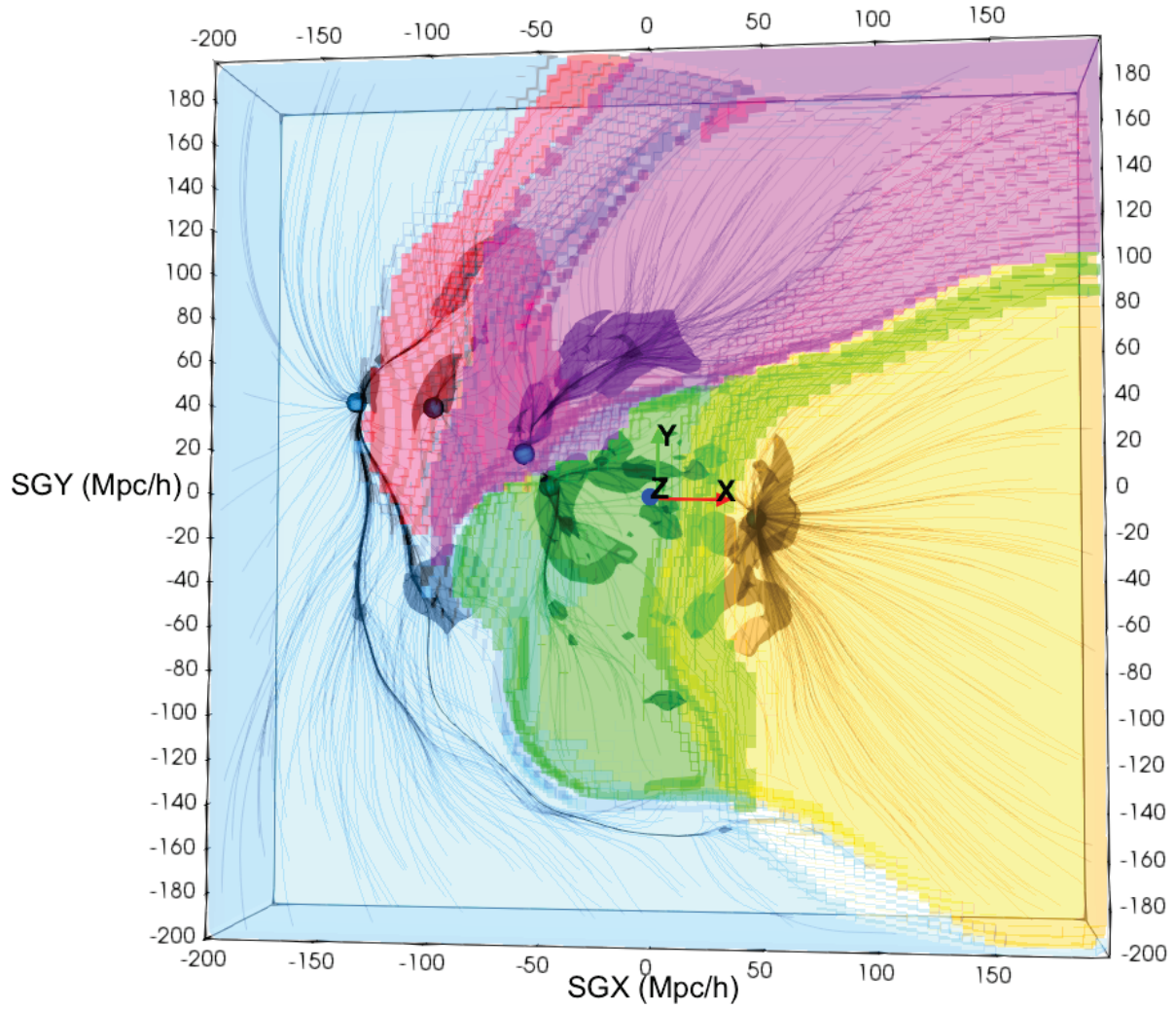


FIGURE 4.23 – Tranche SGX-SGY d’épaisseur $-25 < \text{SGZ} < 25$ Mpc/h. Les isosurfaces noires correspondent aux surdensités telles que $\delta > 1.5$, et les lignes de flux sont tracées en point de vue avant. Les isosurfaces colorées correspondent aux bassins d’attraction segmentés des structures listées en table 4.3 : Shapley I en bleu, Shapley II en rouge, Laniakea en vert, Coma en rose, et Persée-Poisson en jaune.

région.

À droite est cartographié l'histogramme des lignes de flux obtenus à partir du champ de vitesse CF3 (point de vue avant). Les zones oranges et jaunes représentent des vallées et rivières gravitationnelles, i.e régions que beaucoup de lignes de flux traversent. Une surdensité est observée dans la région de Vela. Cela met ainsi en évidence le passage d'une rivière gravitationnelle au niveau de Vela, mais ne confirme pas l'existence de ce superamas. Comme il a été vu sur une simulation, les isocontours de l'histogramme des lignes de flux relient les surdensités de δ entre elles, et ne correspondent donc pas forcément à une surdensité.

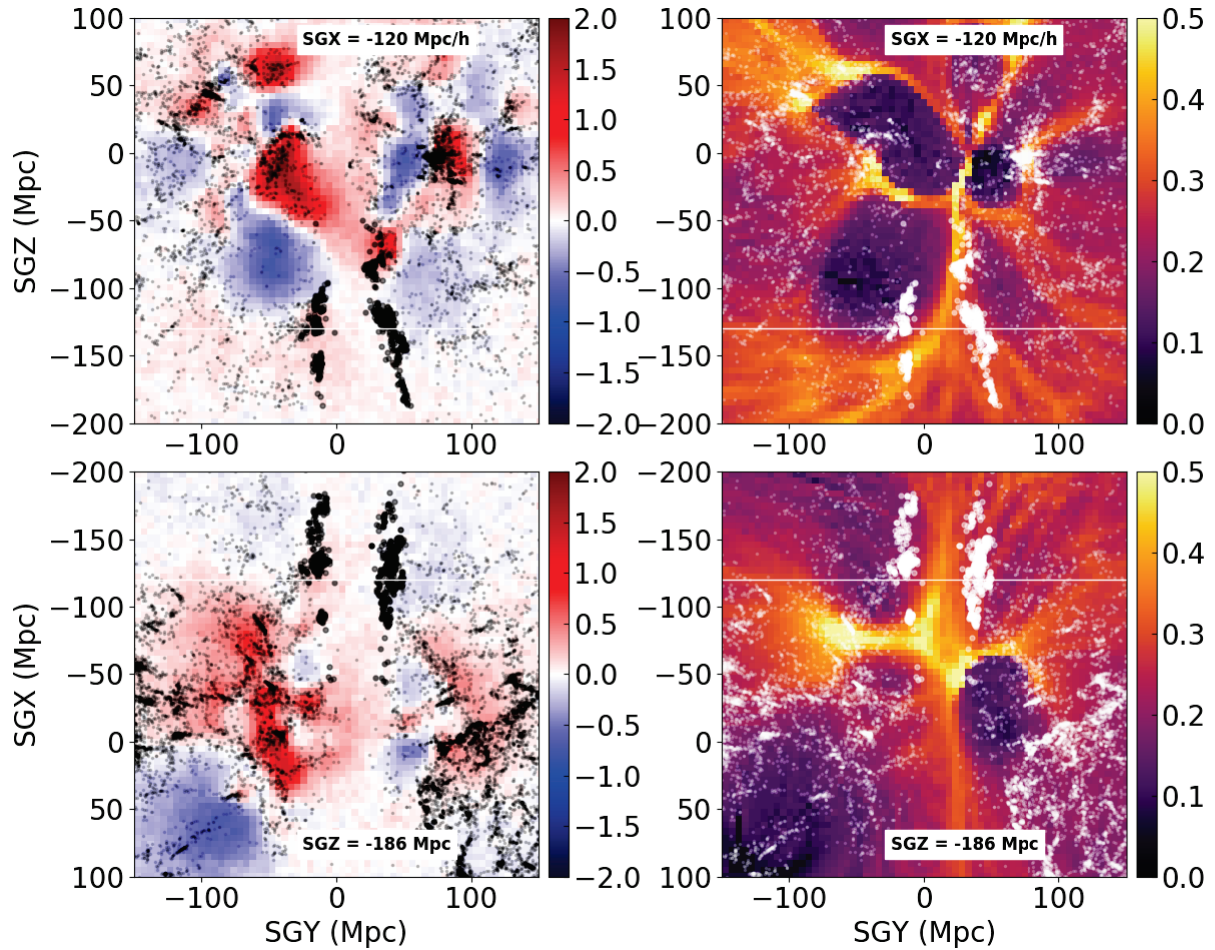


FIGURE 4.24 – Tranches SGX-SGY d'épaisseur $-25 < SGZ < 25 \text{ Mpc/h}$. Les petits points noirs correspondent aux galaxies d'une compilation de redshifts extraite de LEDA. Les plus gros points noirs correspondent aux galaxies du sondage de redshift de Kraan-Korteweg et al. (2017). À gauche, la palette de couleur représente le champ de surdensité δ CF3, tandis qu'à droite, elle représente l'histogramme des lignes de flux obtenu à partir du champ de vitesse CF3 (point de vue avant).

Chapitre 5

Taux de croissance local

Les vitesses particulières des galaxies sont responsables de la formation des structures de l'univers. Elles peuvent donc servir de sondes cosmologiques, et leur analyse permet de mesurer le taux de croissance actuel de l'univers, paramètre représentant le *tir à la corde* en cours entre l'expansion et la gravitation. Ce chapitre présente l'estimation de ce paramètre à l'aide des données observationnelles *Cosmicflows-3* décrites en section 2.2.2.

Sont présentées tout d'abord différentes méthodes existantes permettant d'estimer le taux de croissance, et qui varient selon les données utilisées : les distortions dans l'espace des redshifts, les oscillations acoustiques des baryons et les vitesses particulières des galaxies. La méthode utilisée dans cette thèse exploitant les statistiques à deux points des vitesses particulières sera ensuite expliquée plus en détails : les trois estimateurs considérés v_{12} , ψ_1 et ψ_2 , ainsi que l'estimation des erreurs sur ces statistiques, seront décrites. Puis, cette méthode sera validée sur des catalogues fictifs. Ceux-ci permettront d'estimer les erreurs dues à la variance cosmique, qui sont significatives du fait de la taille du catalogue de vitesses particulières *Cosmicflows-3*. Le choix de la valeur de la constante de Hubble sera ensuite discuté. Enfin, le taux de croissance local sera mesuré sur les données du catalogue observationnel *Cosmicflows-3*, ce qui clôturera ce chapitre et le contenu de cette thèse.

Sommaire

5.1	Bibliographie	142
5.1.1	Distorsions dans l'espace des redshifts	142
5.1.2	Oscillations acoustiques des baryons	145
5.1.3	Vitesses particulières	147
5.2	Statistiques à deux points des vitesses particulières	148
5.2.1	Vitesse relative de paires	149
5.2.2	Corrélations à deux points des vitesses particulières	151
5.2.3	Estimation des erreurs sur les statistiques à deux points	154
5.3	Taux de croissance sur les données Cosmicflows-3	155
5.3.1	Validation sur des catalogues fictifs	156
5.3.2	Variance cosmique	159
5.3.3	Constante de Hubble	161
5.3.4	Taux de croissance local avec les données Cosmicflows-3	163

5.1 Bibliographie

Jusqu'à présent, diverses estimations du taux de croissance f des grandes structures de l'univers ont été publiées. Les méthodes utilisées pour la détermination de ce paramètre varient selon les données, i.e les sondes cosmologiques considérées. Dans cette section, les principales caractéristiques des méthodes utilisées pour trois sondes cosmologiques différentes sont rapidement décrites : les distortions dans l'espace des redshifts, les oscillations acoustiques des baryons, et les vitesses particulières.

5.1.1 Distorsions dans l'espace des redshifts

Les mesures de redshift évaluent à la fois la contribution de l'expansion de l'univers (flot de Hubble), ainsi que l'apport de la gravitation à travers les vitesses particulières. Cela crée des déformations lorsque les galaxies sont positionnées à leur redshift observé z_{obs} au lieu d'être placées à leur distance d . Ces déformations sont appelées distortions dans l'espace des redshifts, ou *Redshift Space Distortions* (RSD). Les vitesses particulières v_p des galaxies sont à l'origine des ces déformations ([Sargent and Turner, 1977](#)) :

$$z_{\text{obs}} = z_{\text{true}} + \frac{v_p}{c}, \quad (5.1)$$

où z_{true} correspond au vrai redshift des galaxies.

Les distortions dans l'espace des redshifts se manifestent à différentes échelles, de deux façons distinctes. Celles-ci sont schématisées sur la figure 5.1 gauche, qui compare les formes apparentes d'un amas (représenté en jaune) et d'un superamas (représenté en bleu) dans l'espace réel (i.e si les galaxies sont positionnées à leur distance) et dans l'espace des redshifts (i.e si les galaxies sont positionnées à leur redshift). Les flèches correspondent aux vitesses particulières des galaxies. L'observateur est représenté par un oeil en bas de la figure.

Aux petites échelles (< 1 Mpc/h), les galaxies au centre des amas ont des mouvements relatifs les unes par rapport aux autres très aléatoires. Cette forte dispersion des vitesses crée des vitesses particulières sur la ligne de visée de l'observateur, et les amas paraissent donc étirés radialement, comme pointant vers l'observateur. Cet effet est appelé *effet des doigts de Dieu* (*finger of God* [Jackson, 1972](#)). Les galaxies encore aux bords des amas tombent vers le centre de ces derniers. Dans ce cas, les mouvements des galaxies sont cohérents, et les structures n'apparaissent pas étirées mais aplaties. Cet effet est appelé *effet Kaiser* ([Kaiser, 1987](#)).

L'outil statistique permettant de mettre en évidence les distortions dans l'espace des redshifts est la fonction de corrélation à deux points dans l'espace des redshifts, notée $\xi(\sigma, \pi)$. Cet outil mesure la probabilité de trouver une paire de galaxies ayant une séparation transverse σ et une séparation radiale π par rapport à la ligne de visée de l'observateur. Dans un univers isotrope, ξ est indépendant de la direction. Cependant, ce n'est pas le cas dans l'espace des redshifts à cause des distortions dues aux vitesses particulières.

La figure 5.2 droite, extraite de [Peacock et al. \(2001\)](#), montre la fonction de corrélation

à deux points dans l'espace des redshifts $\xi(\sigma, \pi)$ obtenue avec les galaxies du sondage de redshift 2dFGRS (*2dF Galaxy Redshift Survey*, [Peacock et al., 2001](#)). Sur cette figure, $\xi(\sigma, \pi)$ est tracée en fonction de σ sur l'axe X et π sur l'axe Y. La palette de couleurs va du noir au jaune, i.e d'une faible probabilité à une forte probabilité de trouver une paire de galaxies ayant une séparation transverse σ et une séparation radiale π . La fonction est obtenue pour un quart de la figure, puis est répétée dans les trois autres afin de bien mettre en évidence les déformations, les déviations par rapport à une symétrie circulaire (cas d'un univers isotrope). Les distortions dans l'espace des redshifts sont en effet clairement visibles et sont mises en valeur par des flèches blanches. À petite échelle, une élongation de type *doigts de Dieu* est observée, tandis qu'à grande échelle, un aplatissement dû à l'effet Kaiser est visible.

[Kaiser \(1987\)](#) a montré qu'il est possible d'utiliser les distorsions dans l'espace des redshifts pour extraire des paramètres cosmologiques. En effet, le regroupement des galaxies (*galaxy clustering*), i.e la forme de $\xi(\sigma, \pi)$ est sensible à l'amplitude de $f\sigma_8$. Ce paramètre correspond au produit du taux de croissance des structures f et de l'amplitude des fluctuations de matière σ_8 et est plus couramment appelé taux de croissance normalisé des grandes structures de l'univers. Dans la figure 5.2 droite, les contours noirs représentent la forme de $\xi(\sigma, \pi)$ prédite par le modèle Λ CDM. Les seuils des contours représentés sont $\xi = 10, 5, 2, 1, 0.5, 0.2, 0.1$.

[Peacock et al. \(2001\)](#); [Hawkins et al. \(2003\)](#); [Beutler et al. \(2012\)](#); [Samushia et al. \(2012\)](#); [de la Torre et al. \(2013\)](#) ont utilisé les distorsions dans l'espace des redshifts afin de mesurer $f\sigma_8$. Les résultats obtenus par ces publications sont donnés dans la table 1.1.

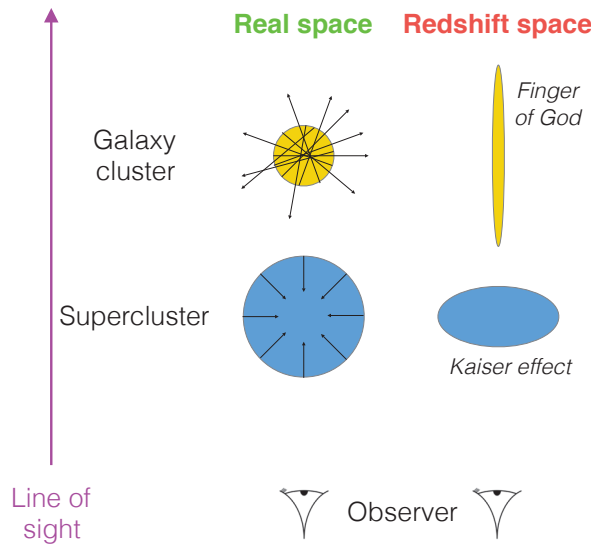


FIGURE 5.1 – Distortions dans l'espace des redshifts à différentes échelles. Les formes apparentes d'un amas (représenté en jaune) et d'un superamas (représenté en bleu) sont comparées dans l'espace réel et dans l'espace des redshifts. Les flèches correspondent aux vitesses particulières des galaxies. L'observateur est représenté par un oeil en bas de la figure.

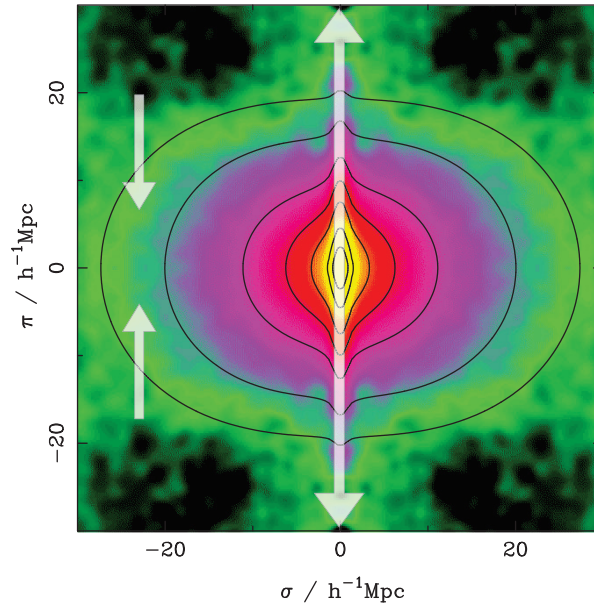


FIGURE 5.2 – Fonction de corrélation à deux points dans l’espace des redshifts $\xi(\sigma, \pi)$, obtenue avec les galaxies de 2dFGRS, et tracée en fonction de σ sur l’axe X et π sur l’axe Y. La palette de couleurs va du noir au jaune, i.e d’une faible probabilité à une forte probabilité de trouver une paire de galaxie ayant une séparation transverse σ et une séparation radiale π . Les contours noirs représentent la forme de $\xi(\sigma, \pi)$ prédite par le modèle Λ CDM. Les seuils des contours représentés sont $\xi = 10, 5, 2, 1, 0.5, 0.2, 0.1$. Extraite de [Peacock et al. \(2001\)](#).

5.1.2 Oscillations acoustiques des baryons

Que sont les oscillations acoustiques des baryons ? Considérons le plasma primordial, c'est à dire l'univers très jeune et très chaud, dans lequel il y a des fluctuations de densité (baryons, photons, matière noire, ...). Dans une fluctuation de densité, les collisions entre les baryons et les photons créent une onde sonore sphérique qui se propage autour de cette fluctuation. À la recombinaison ($z = 10000$), les baryons et les photons deviennent indépendants. Il n'y a donc plus de plasma, et la vitesse de l'onde acoustique créée par la collision diminue puis devient nulle à une distance de 150 Mpc de la fluctuation où a eu lieu la collision. On retrouve donc, en plus de la fluctuation originale, une surdensité de baryons dans une coquille située à 150 Mpc de la fluctuation initiale. Les deux pics de surdensité sont soumis à la gravitation et attirent donc les baryons et la matière noire se trouvant à proximité. Ils sont ainsi à l'origine de la formation des grandes structures que l'on connaît aujourd'hui : les galaxies, amas et superamas de galaxies. Ceci est valable pour toutes les fluctuations de densité présentes dans le plasma primordial : pour chaque fluctuation, on a une coquille de surdensité de matière dans un rayon de 150 Mpc de la surdensité initiale.

Dans les données observationnelles, cela se traduit par un pic dans la fonction de corrélation à deux points dans l'espace des redshifts $\xi(r)$ à une séparation de $r = 150$ Mpc. Il est ainsi plus probable de trouver des galaxies séparées de 150 Mpc que 120 ou 180 Mpc. Ce pic est aussi présent sous forme d'oscillations dans le spectre de puissance de matière linéaire, qui correspond à la transformée de Fourier de la fonction de corrélation.

Les oscillations sont ainsi prédites par le modèle standard de la cosmologie. Cependant, la première détection des oscillations acoustiques des baryons (*Baryons Acoustic Oscillations*, BAO) n'a eu lieu qu'en 2005 (Eisenstein et al., 2005; Cole et al., 2005). Les sondages de redshifts précédents n'avaient pas une couverture en redshift assez grande pour pouvoir détecter les BAO.

Les BAO constituent donc une sonde cosmologique permettant d'étudier la croissance des structures. Il est possible de contraindre des paramètres cosmologiques dont le taux de croissance des grandes structures en comparant le pic BAO dans la fonction de corrélation (ou spectre de puissance) obtenu avec les données de redshift surveys aux fonctions prédites par différents modèles cosmologiques.

La figure 5.3 représente la fonction de corrélation à deux points dans l'espace des redshifts $\xi(r)$ obtenue à partir du sondage de redshift SDSS-LRG (*Sloan Digital Sky Survey - Luminous Red Galaxies*, extraite de Eisenstein et al., 2005). Les résultats obtenus par les auteurs sont représentés par des points noirs. Un pic vers $r = 110 \text{ Mpc}/h = 150 \text{ Mpc}$ correspondant au pic BAO est observé.

Les traits pleins en vert, rouge et bleu correspondent aux modèles Λ CDM calculés avec $\Omega_m h^2 = 0.12, 0.13$ et 0.14 respectivement. Le paramètre $\Omega_b h^2 = 0.024$ est le même pour les trois cas. Le pic BAO est à nouveau clairement mis en évidence. Par ailleurs, le trait rose représente un modèle sans constante cosmologique. Dans ce cas, aucun pic n'est visible vers 150 Mpc.

La forme du pic BAO dépend ainsi du modèle et des valeurs des paramètres cosmologiques considérés. Ceux-ci peuvent ainsi être mesurés en comparant $\xi(r)$ mesuré sur des sondages

de redshift à sa forme prédite par un modèle cosmologique.

Blake et al. (2011); Reid et al. (2012) ont utilisé les oscillations acoustiques des baryons pour estimer $f\sigma_8$. Les valeurs obtenus sont listées dans la table 1.1.

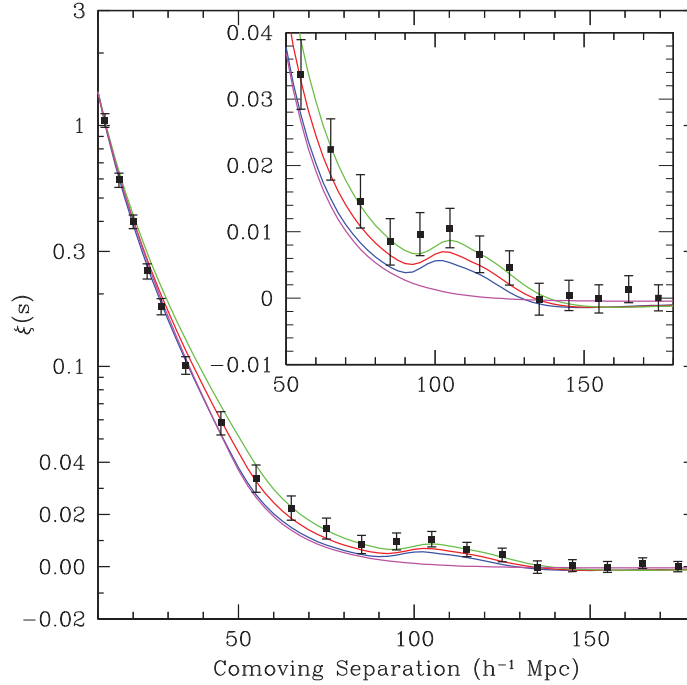


FIGURE 5.3 — Fonction de corrélation à deux points dans l’espace des redshifts $\xi(r)$ obtenue à partir du sondage de redshift SDSS-LRG (*Sloan Digital Sky Survey - Luminous Red Galaxies*, extraite de Eisenstein et al., 2005). Les résultats obtenus par les auteurs sont représentés par des points noirs. Les traits pleins en vert, rouge et bleu correspondent aux modèles Λ CDM calculés avec $\Omega_m h^2 = 0.12, 0.13$ et 0.14 respectivement. Le paramètre $\Omega_b h^2 = 0.024$ est le même pour les trois cas. Par ailleurs, le trait rose représente un modèle sans constante cosmologique.

5.1.3 Vitesses particulières

Les effets des vitesses particulières, i.e la gravitation responsable de la formation des structures, peuvent être étudiés en utilisant les vitesses particulières de galaxies estimées directement à partir des distances observées (voir section 2.1.2). Celles-ci sont issues de catalogues observationnels tels que les catalogues *Cosmicflows*. L'utilisation des vitesses particulières en tant que sondes cosmologiques est un moyen prometteur de mesure du taux de croissance des grandes structures de l'univers.

Dans le régime linéaire, le champ de vitesse particulière $\vec{v}(\vec{r})$ à une position \vec{r} est relié au champ de surdensité $\delta(\vec{r})$ par la formule :

$$\vec{v}(\vec{r}) = \frac{H_0 f}{4\pi} \int_0^\infty d^3\vec{x} \delta(\vec{x}) \frac{\vec{x} - \vec{r}}{|\vec{x} - \vec{r}|^3}, \quad (5.2)$$

où f est le taux de croissance des grandes structures. À partir d'un sondage de redshift, il est possible d'obtenir le champ de surdensité de galaxies noté $\delta_g(\vec{r})$, puis d'estimer le champ de vitesse \vec{v}_{pred} prédit par le champ de surdensité :

$$\vec{v}(\vec{r}) = \frac{H_0 \beta}{4\pi} \int_0^\infty d^3\vec{r}' \delta_g(\vec{r}') \frac{\vec{r}' - \vec{r}}{|\vec{r}' - \vec{r}|^3}, \quad (5.3)$$

où $\beta = f/b$ et b est le biais entre la densité de galaxies et la densité de masse. La comparaison des vitesses particulières observées, issues de catalogues observationnels, à celles prédites permet alors d'extraire le paramètre β . Par ailleurs, les sondages de redshift donnent accès au paramètre $\sigma_{8,gal}$ correspondant aux fluctuations de densité de galaxies dans une sphère top-hat de rayon 8 Mpc/ h . Il est ainsi possible d'en déduire le taux de croissance normalisé $f\sigma_8 = \beta\sigma_{8,gal}$.

Cette méthode est celle utilisée par [Turnbull et al. \(2012\)](#). Les auteurs ont obtenu $f\sigma_8 = 0.4 \pm 0.07$ à $z = 0.03$ en analysant les vitesses particulière des galaxies hôtes des SNIa de la compilation *First Amendment* (A1).

[Johnson et al. \(2014\)](#) utilisent les vitesses particulières des galaxies du catalogue *6-degree Field Galaxy Survey* (6dFGS) pour mesurer le taux de croissance des grandes structures $f\sigma_8$. Les auteurs obtiennent, en analysant les statistiques à deux points des vitesses particulières (spectre de puissance des vitesses), $f\sigma_8 = 0.428^{+0.079}_{-0.068}$ à un décalage spectral $z = 0.05$.

La méthode utilisée dans ce chapitre pour estimer le taux de croissance des grandes structures de l'univers utilise les vitesses particulières des galaxies, issues du catalogue *Cosmicflows-3*. Cette méthode, basée sur l'analyse des statistiques à deux points des vitesses particulières, est décrite ci-après dans la section 5.2.

5.2 Statistiques à deux points des vitesses particulières

Les catalogues préparés par la collaboration *Cosmicflows* donnent accès aux vitesses particulières des galaxies. Ce chapitre présente les résultats en cours de publication par [Dupuy et al. \(2018\)](#). Dans cette publication, le taux de croissance des grandes structures de l'univers local est estimé à l'aide des vitesses particulières du catalogue *Cosmicflows-3*. La méthode utilisée pour y parvenir consiste à analyser les statistiques à deux points des vitesses particulières des galaxies. Les statistiques considérées sont détaillées dans cette section et sont au nombre de trois : la vitesse relative de paires v_{12} , et les fonctions de corrélation des vitesses particulières ψ_1 et ψ_2 .

Pour la suite est considérée une paire de galaxies qui sont situées, par rapport à l'observateur, aux positions \vec{r}_A et \vec{r}_B respectivement. Leur vecteur de séparation est donné par $\vec{r} = \vec{r}_A - \vec{r}_B$. Ces deux galaxies ont pour vitesse particulière \vec{v}_A et \vec{v}_B respectivement. Les composantes radiales de ces vitesses particulières sont données par $\vec{u}_A = u_A \hat{r}_A = (\vec{v}_A \cdot \hat{r}_A) \hat{r}_A$ et $\vec{u}_B = u_B \hat{r}_B = (\vec{v}_B \cdot \hat{r}_B) \hat{r}_B$, où $\hat{r}_{A,B}$ sont les vecteurs unitaires pointant vers les galaxies A et B respectivement. Les cosinus des angles entre les différentes directions sont donnés par $\cos \theta_A = \hat{r}_A \cdot \hat{r}$, $\cos \theta_B = \hat{r}_B \cdot \hat{r}$ et $\cos \theta_{AB} = \hat{r}_A \cdot \hat{r}_B$. La figure 5.4 schématise cette paire de galaxies avec tous les paramètres définis ci-dessus.

Pour éviter en partie le biais de Malmquist, les galaxies sont positionnées à leur redshift. Placer les galaxies à leur distance observée introduit un large biais de Malmquist, associé à des erreurs significatives ([Strauss and Willick, 1995](#); [Lynden-Bell et al., 1988](#)). Le biais de Malmquist est aussi présent dans l'espace des redshifts, mais est beaucoup moins important, particulièrement pour les galaxies lointaines.

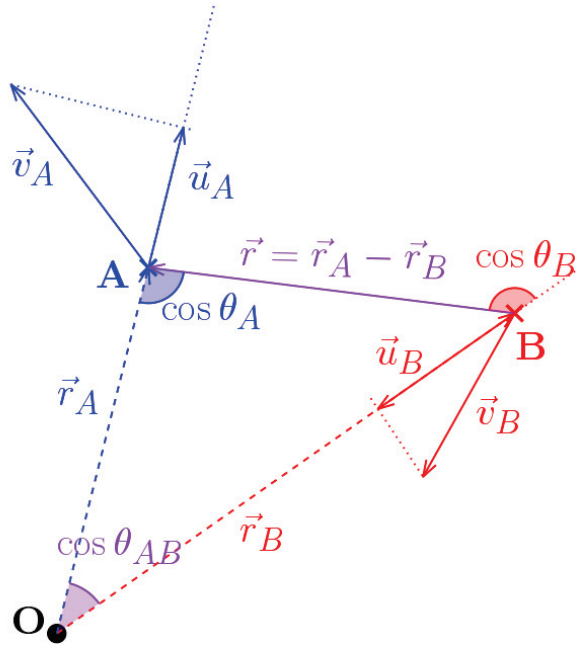


FIGURE 5.4 – Paire de galaxies considérée dans ce chapitre.

5.2.1 Vitesse relative de paires

Modèle analytique

La vitesse relative de paires v_{12} a été introduite pour la première fois dans le cadre de la théorie Bogolyubov-Born-Green-Kirkwood-Yvon (BBGKY, [Yvon, 1935](#); [N. N. Bogoliubov, 1946](#); [Kirkwood, 1946, 1947](#); [Born and Green, 1946](#)). Dans cette théorie, l'équation de conservation des paires relie la fonction de corrélation à deux points $\xi(r)$ au taux de croissance f et à la vitesse particulière relative de paires de galaxies $v_{12}\vec{r}/r$, où \vec{r} est le vecteur de séparation comobile d'une paire de galaxies ([Davis and Peebles, 1977](#); [Peebles, 1980](#)). Pour une paire de galaxies séparées par une distance r , la vitesse relative de paires est donnée par ([Juszkiewicz et al., 1998](#)) :

$$v_{12}(r) = \left\langle (\vec{u}_A - \vec{u}_B) \cdot \hat{\vec{r}} \right\rangle_\rho = \frac{\langle (\vec{u}_A - \vec{u}_B) (1 + \delta_A)(1 + \delta_B) \rangle}{1 + \xi(r)}, \quad (5.4)$$

où \vec{u}_i et δ_i sont la vitesse particulière et le contraste de densité à la position $i = A, B$ et $\langle \dots \rangle_\rho$ correspond à une moyenne pondérée sur les paires où le poids est $(1 + \delta_A)(1 + \delta_B)(1 + \xi(r))^{-1}$.

Dans le régime non-linéaire, c'est-à-dire pour les paires très proches, on a $\xi(r) \gg 1$ et la solution de l'équation de conservation des paires est $v_{12}(r) = -H_0 r$: la vitesse relative de paires compense le flot de Hubble. Dans le régime linéaire, c'est-à-dire pour les grandes distances de séparation, on a $\xi(r) \ll 1$ et la solution de l'équation de conservation pour v_{12} est donnée par l'expansion perturbative de $\xi(r)$.

Afin d'estimer des paramètres cosmologiques tels que le paramètre de densité de matière Ω_m , [Juszkiewicz et al. \(1999\)](#) introduit une solution pour v_{12} valide dans les deux régimes en interpolant les solution linéaires et non-linéaires :

$$v_{12}(r) \approx -\frac{2}{3}Hr f \bar{\xi}(r) \left[1 + \alpha \bar{\xi}(r) \right], \quad (5.5)$$

où $\bar{\xi}(r) = \bar{\xi}(r)/[1 + \xi(r)]$ et $\bar{\xi}(r) = 3r^{-3} \int_0^r \xi(x)x^2 dx$ est la fonction de corrélation à deux points moyennée dans une sphère de rayon r . La constante α est définie par :

$$\alpha = 1.2 - 0.65\gamma, \quad (5.6)$$

où γ est la pente en échelle logarithmique de la fonction de corrélation à deux points à la frontière entre les regimes non-linéaire et linéaire, c'est-à-dire lorsque $\xi(r) = 1$:

$$\gamma = - \left. \frac{d \ln \xi(r)}{d \ln r} \right|_{\xi=1}. \quad (5.7)$$

L'équation 5.6 est valide pour des valeurs de γ comprises entre 0 et 3 (0 et 3 exclus).

À partir de la solution approximative de $v_{12}(r)$ (equation 5.5), on retrouve la solution linéaire pour $\xi \rightarrow 0$, et la solution valide dans le regime non-linéaire si $x \rightarrow 0$. En effet, l'équation 5.5 a été testée et validée par Juszkiewicz et al. (1999) sur des simulations N-body pour $0.1 < \xi(r) < 1000$.

L'équation 5.5 montre que la vitesse particulière relative v_{12} dépend du taux de croissance f . Donc en analysant les vitesses relatives observées et en les comparant à la théorie, il est possible d'estimer f .

La figure 5.5 montre la vitesse relative de paires v_{12} en fonction de la distance de séparation r , calculée avec l'équation 5.5 pour diverses valeurs du taux de croissance f . La valeur correspondant au modèle Λ CDM, i.e $f \approx 0.5$, est représentée en trait plein, tandis que les autres valeurs sont représentées par des traits pointillés. La correspondance des couleur – f est donnée par la palette de couleurs. L'amplitude de v_{12} dépend de f , permettant d'utiliser cette statistique pour estimer le taux de croissance à partir de données observationnelles.

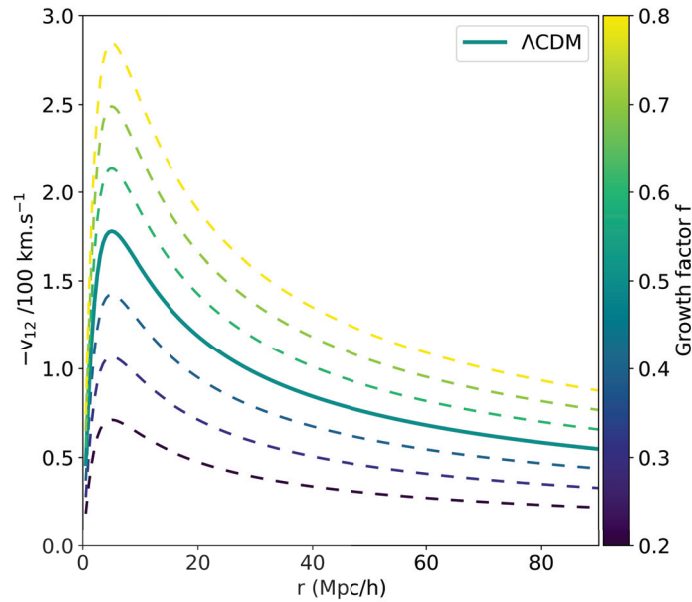


FIGURE 5.5 – Vitesse relative de paires v_{12} en fonction de la distance de séparation r , calculée avec l'équation 5.5 pour diverses valeurs du taux de croissance f . La valeur correspondant au modèle Λ CDM, i.e $f \approx 0.5$, est mise en valeur et représentée en trait plein.

Estimateur observationnel

Cependant, les observations ne donnent accès qu'à la composante radiale des vitesses particulières des galaxies. Il est donc impossible d'utiliser l'équation 5.5 directement à partir de données observationnelles. Il faut donc passer par un estimateur déterminant $v_{12}(r)$ à partir des vitesses particulières radiales (observées) seulement.

Ferreira et al. (1999) introduit un tel estimateur permettant d'estimer $v_{12}(r)$ directement à partir des catalogues observationnels de vitesses particulières de galaxies. L'estimateur

radial de $v_{12}(r)$ est donné par :

$$v_{12}(r) = \frac{2 \sum (u_A - u_B)(\cos \theta_A + \cos \theta_B)}{\sum (\cos \theta_A + \cos \theta_B)^2}. \quad (5.8)$$

Les sommes sont calculées pour toutes les paires séparées par une distance r .

Il est donc ainsi possible d'extraire le taux de croissance des grandes structures de l'univers en comparant la vitesse relative de paires observée obtenue avec l'équation 5.8 à celle prédite et dépendant de f donnée par l'équation 5.5.

Des estimations du taux de croissance à partir de la mesure de la vitesse relative de paires ont été présentées dans diverses publications. En combinant les résultats avec les données provenant des sondages de redshift, il est possible de déduire les paramètres Ω_m et σ_8 . Juskiewicz et al. (2000) a mesuré $\Omega_m = 0.35 \pm 0.15$ en étudiant la vitesse relative de paires obtenue à partir des vitesses particulières de galaxies du catalogue *Mark III*. Feldman et al. (2003) a obtenu $\Omega_m = 0.30^{+0.17}_{-0.07}$ et $\sigma_8 = 1.13^{+0.22}_{-0.23}$ par l'analyse des vitesses relatives de paires à l'aide de divers catalogues de vitesses particulières (*Mark III*, *Spiral Field I-Band*, *Nearby Early-type Galaxies Survey* and the *Revised Flat Galaxy Catalog*). Ma et al. (2015) obtient $\Omega_m^{0.6} h = 0.102^{+0.384}_{-0.044}$ et $\sigma_8 = 0.39^{+0.73}_{-0.1}$ en analysant la vitesse relative de paires du catalogue *Cosmicflows-2*.

5.2.2 Corrélations à deux points des vitesses particulières

Les deux autres statistiques considérées dans ce chapitre sont des estimateurs de la fonction de corrélation des vitesses particulières.

Fonction de corrélation des vitesses particulières

Dans le cadre du régime linéaire, le tenseur de corrélation des vitesses à deux points pour un champ de vitesse homogène $\vec{v}(\vec{r})$ est défini par Monin and Yaglom (1975) :

$$\Psi_{ij}(r) \equiv \langle v_i(\vec{r}_A) v_j(\vec{r}_B) \rangle = [\Psi_{\parallel}(r) - \Psi_{\perp}(r)] \hat{r}_i \hat{r}_j + \Psi_{\perp}(r) \delta_{ij}, \quad (5.9)$$

où i et j correspondent aux coordonnées cartésiennes (x, y, z) , $r = |\vec{r}|$ est le vecteur de separation d'une paire de galaxies A et B , et $\hat{r} = \vec{r}/r$. Les quantités $\Psi_{\parallel}(r)$ et $\Psi_{\perp}(r)$ sont les fonctions de corrélations des vitesses radiales et transversales (par rapport à \vec{r}). Ces quantités sont données par Gorski (1988) :

$$\Psi_{\parallel}(r) = \frac{H_0^2 (f \sigma_8)^2}{2\pi^2} \int P_0(k) \left[j_0(kr) - 2 \frac{j_1(kr)}{kr} \right] dk, \quad (5.10)$$

et

$$\Psi_{\perp}(r) = \frac{H_0^2 (f\sigma_8)^2}{2\pi^2} \int P_0(k) \frac{j_1(kr)}{kr} dk, \quad (5.11)$$

où $j_0(x)$ et $j_1(x)$ sont les fonctions de Bessel sphériques de première espèce :

$$j_0(x) = \frac{\sin x}{x}, \quad j_1(x) = \frac{\sin x}{x^2} - \frac{\cos x}{x}, \quad (5.12)$$

et $P(k)$ est le spectre de puissance linéaire non normalisé. Dans ce chapitre, $P(k)$ est calculé à l'aide du logiciel **CAMB Python** (Lewis and Bridle, 2002; Lewis and Challinor, 2011) avec les paramètres cosmologiques donnés par Planck 2015 (Planck Collaboration et al., 2016).

Les quantités $\Psi_{\parallel}(r)$ et $\Psi_{\perp}(r)$ dépendent toutes les deux du paramètre $f\sigma_8$ correspondant au taux de croissance normalisé. Ces fonctions de corrélation peuvent donc être utilisées pour contraindre ce paramètre.

Estimateurs observationnels

Cependant, comme les catalogues de vitesses particulières de galaxies fournissent seulement la partie radiale, il n'est pas possible d'appliquer les équations 5.10 et 5.11 sur des données observationnelles.

Pour cette raison, Gorski et al. (1989) introduit deux statistiques des vitesses particulières, notées ψ_1 et ψ_2 , dépendant seulement des composantes radiales des vitesses particulières.

La première statistique à deux points $\psi_1(r)$ est définie par :

$$\psi_1(r) = \frac{\sum \vec{u}_A \cdot \vec{u}_B}{\sum (\hat{\vec{r}}_A \cdot \hat{\vec{r}}_B)^2} = \frac{\sum u_A u_B \cos \theta_{AB}}{\sum \cos^2 \theta_{AB}}, \quad (5.13)$$

et $\psi_2(r)$ est donnée par :

$$\psi_2(r) = \frac{\sum (\vec{u}_A \cdot \hat{\vec{r}}) (\vec{u}_B \cdot \hat{\vec{r}})}{\sum (\hat{\vec{r}}_A \cdot \hat{\vec{r}}_B) (\hat{\vec{r}}_A \cdot \hat{\vec{r}}) (\hat{\vec{r}}_B \cdot \hat{\vec{r}})} = \frac{\sum u_A u_B \cos \theta_A \cos \theta_B}{\sum \cos \theta_{AB} \cos \theta_A \cos \theta_B}. \quad (5.14)$$

Tout comme l'équation 5.8, les sommes sont calculées pour toutes les paires de galaxies séparées par une distance r .

Modèles

Il faut cependant relier ces deux estimateurs ψ_1 et ψ_2 aux fonctions de corrélation radiales et transverse Ψ_{\parallel} et Ψ_{\perp} afin d'obtenir une dépendance en f . La fonction de corrélation des composantes radiales des vitesses particulières des galaxies est donnée par :

$$\langle u_A^m u_B^n \rangle = \hat{r}_A^m \hat{r}_B^n \Psi_{ij}(r) \hat{r}_A^i \hat{r}_B^j, \quad (5.15)$$

où i, j, m et n sont les composantes cartésiennes x, y, z des vecteurs. En insérant l'équation 5.15 dans les équations 5.13 et 5.14, on peut écrire les quantités $\psi_1(r)$ et $\psi_2(r)$ en fonction des fonctions de corrélation des vitesses radiales et transversales [Gorski et al. \(1989\)](#) :

$$\psi_{1,2}(r) = \mathcal{A}_{1,2}(r) \Psi_{\parallel}(r) + [1 - \mathcal{A}_{1,2}(r)] \Psi_{\perp}(r), \quad (5.16)$$

où

$$\mathcal{A}_1(r) = \frac{\sum \cos \theta_A \cos \theta_B \cos \theta_{AB}}{\sum \cos^2 \theta_{AB}}, \quad (5.17)$$

et

$$\mathcal{A}_2(r) = \frac{\sum \cos^2 \theta_A \cos^2 \theta_B}{\sum \cos \theta_A \cos \theta_B \cos \theta_{AB}}. \quad (5.18)$$

Les fonctions $\mathcal{A}_1(r)$ et $\mathcal{A}_2(r)$ contiennent des informations sur la géométrie du catalogue de vitesses particulières considéré. De plus, elles mesurent les contributions de $\Psi_{\parallel}(r)$ et $\Psi_{\perp}(r)$ aux statistiques $\psi_1(r)$ et $\psi_2(r)$.

En comparant les modèles définis dans l'équation 5.16 aux estimateurs donnés dans les équations 5.13 et 5.14, on peut extraire la valeur sous-jacente du taux de croissance normalisé $f\sigma_8$ à partir de données observationnelles.

La figure 5.6 montre les estimateurs ψ_1 (gauche) et ψ_2 (droite) en fonction de la distance de séparation r , calculés avec l'équation 5.16 pour diverses valeurs du taux de croissance normalisé $f\sigma_8$. Les quantités \mathcal{A}_1 et \mathcal{A}_2 ont été calculées à partir des galaxies du catalogue *Cosmicflows-3*. La valeur de $f\sigma_8$ correspondant au modèle Λ CDM, i.e $f \approx 0.5$, est représentée en trait plein, tandis que les autres valeurs sont représentées par des traits pointillés. La correspondance des couleur – f est donnée par la palette de couleurs. Tout comme v_{12} , l'amplitude de ces statistiques dépend de f , permettant de les utiliser pour estimer le taux de croissance à partir de données observationnelles. Cependant, ψ_1 semble plus adéquat, les courbes correspondant aux différents $f\sigma_8$ étant plus espacées, rendant l'estimation de ce paramètre plus précise. Ceci n'est pas le cas pour ψ_2 , pour lequel les courbes se resserrent et se croisent vers 80 Mpc/h.

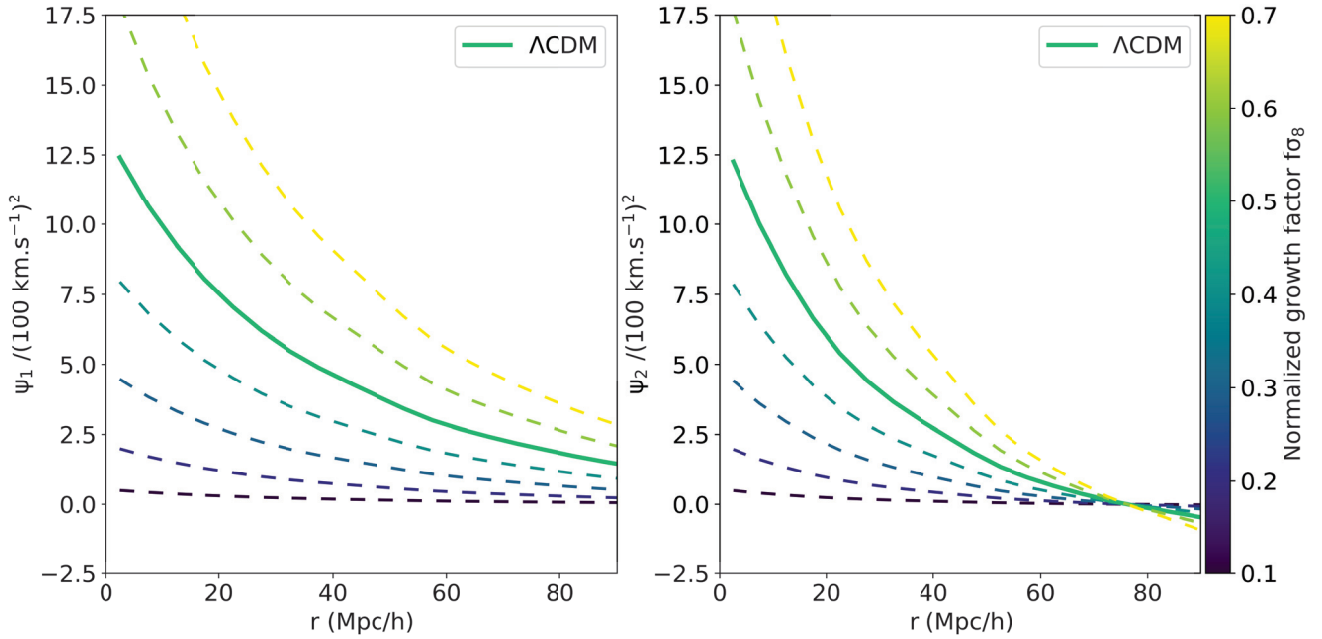


FIGURE 5.6 – Estimateurs ψ_1 (gauche) et ψ_2 (droite) en fonction de la distance de séparation r , calculés avec l'équation 5.16 pour diverses valeurs du taux de croissance normalisé $f\sigma_8$. Les quantités \mathcal{A}_1 et \mathcal{A}_2 ont été calculées à partir des galaxies du catalogue *Cosmicflows-3*. La valeur de $f\sigma_8$ correspondant au modèle Λ CDM, i.e $f \approx 0.5$, est représentée en trait plein, tandis que les autres valeurs sont représentées par des traits pointillés. La correspondance des couleurs selon les valeurs de f est donnée par la barre verticale.

5.2.3 Estimation des erreurs sur les statistiques à deux points

Deux types d'incertitudes sur les statistiques de vitesses particulières doivent être prises en compte lors de l'estimation du taux de croissance à partir de données observationnelles : les erreurs observationnelles, dues à aux erreurs de mesure des distances ; et les erreurs dues à la variance cosmique.

Erreurs observationnelles

Les erreurs observationnelles sur les mesures de distances des galaxies doivent être prises en compte lors de l'estimation du taux de croissance. Les erreurs observationnelles sur les statistiques des vitesses particulières v_{12} , ψ_1 et ψ_2 , i.e les erreurs dues à l'incertitude sur la mesure de la distance, sont estimées par des réalisations Monte-Carlo.

Ces réalisations sont construites en ajoutant une erreur aléatoire à la vitesse particulière radiale. Cette erreur est extraite d'une distribution normale avec une déviation standard égale à l'erreur sur la vitesse particulière déduite de l'incertitude sur le module de la distance. Un total de 100 réalisations ont été calculées pour chaque échantillon considéré

dans la section 5.3 (données fictives et observationnelles). Les statistiques à deux points sont ensuite calculées pour chaque réalisation (avec les équations 5.8, 5.13 ou 5.14), puis l'erreur observationnelle correspond à la déviation standard des valeurs obtenues.

Matrice de covariance et chi-square

Une même galaxie pouvant appartenir à différents bins de séparation, les erreurs entre les bins sont corrélées. Par conséquent, la matrice de covariance doit être prise en compte lorsque l'on compare les statistiques obtenues à partir de mocks ou données observationnelles aux statistiques prédites par un modèle cosmologique afin d'extraire le taux de croissance normalisé $f\sigma_8$.

À partir des mocks décrits ci-dessus, il est possible de calculer la matrice de covariance C entre différents bins de distance de séparation. La covariance entre les bins de séparation r_i et r_j est donnée par :

$$C_{mn} = \frac{1}{N_k - 1} \sum_{k=1}^{N_k} (\psi_k(r_m) - \overline{\psi_m}) (\psi_k(r_n) - \overline{\psi_n}), \quad (5.19)$$

où $\overline{\psi_{m,n}} = \frac{1}{N_k} \sum_{k=1}^{N_k} \psi_k(r_{m,n})$ correspond à la moyenne des statistiques obtenues pour tous les mocks dans les bins r_m et r_n respectivement et $N_k = 100$ est le nombre de mocks considérés.

Le taux de croissance normalisé $f\sigma_8$ est estimé en comparant le modèle théorique d'une statistique à deux points ψ_{mod} (équations 5.5 ou 5.16), à l'estimateur observationnel de la même statistique ψ_{meas} (équations 5.8, 5.13 et 5.14) mesuré à partir des échantillons de vitesses particulières. La mesure de $f\sigma_8$ est obtenue en minimisant le χ^2 défini par une somme sur tous les bins de distance de séparation des paires :

$$\chi^2(f\sigma_8) = \sum_{i,j=0}^{N_{\text{bins}}} \left[\left(\psi_{\text{meas}}(r_i) - \psi_{\text{mod}}(r_i; f\sigma_8) \right) C_{ij}^{-1} \left(\psi_{\text{meas}}(r_j) - \psi_{\text{mod}}(r_j; f\sigma_8) \right) \right]. \quad (5.20)$$

En pratique, dans cette thèse, la minimisation est effectuée à l'aide de l'interface `python iminuit` du logiciel C++ MINUIT (James and Roos, 1975; iminuit team). L'erreur sur le taux de croissance est aussi calculée par ce même logiciel à l'aide de la dérivée seconde du χ^2 .

5.3 Taux de croissance sur les données Cosmicflows-3

Dans cette section, les trois estimateurs des trois statistiques décrites dans la section 5.2 sont tout d'abord appliquées à des données fictives, ressemblant aux données observationnelles, afin de valider les estimateurs.

Pour le reste de ce chapitre, les statistiques v_{12} , ψ_1 et ψ_2 sont calculées avec les équations 5.8, 5.13 et 5.14 respectivement, dans 20 bins de distance de séparation des paires r , allant de 0 Mpc/h à 95 Mpc/h. Ces bins ont ainsi une largeur de 5 Mpc/h. Dans les figures montrant les statistiques, les points sont situés au milieu des bins.

5.3.1 Validation sur des catalogues fictifs

Catalogues fictifs

Les catalogues fictifs utilisés afin de valider les estimateurs des statistiques à deux points sont construits avec la méthode suivante. Les vitesses particulières à trois dimensions \vec{v}_p prédites par le modèle Λ CDM sont extraites d'une Réalisation Contrainte de CF3 (voir section 4.1.1, Graziani et al., 2018; Hoffman and Ribak, 1991, 1992) aux positions des galaxies d'un catalogue donné. La composante radiale u , i.e observable, des vitesses particulières est ensuite déduite : $u = \vec{v}_p \cdot \hat{r}$.

On considère trois mocks construits avec la méthode décrite ci-dessus à partir de trois catalogues différents : 2MASS Redshift Survey (Huchra et al., 2012, 2MRS) et *Cosmicflows-3* non-groupé et groupé. Toute la dérivation montrée dans la section 5.2 n'est valable que pour des paires de vraies galaxies, et non pour des paires de groupes de galaxies. Cependant, le catalogue CF3 groupé sera tout de même considéré par la suite comme un test pour ceux qui utiliseraient ces données. Un certain nombre de publications utilisent en effet les versions groupées des catalogues *Cosmicflows* pour estimer $f\sigma_8$ (ou $\beta = f/b$), tel que Ma et al. (2015); Nusser (2017).

Vitesse relative de paires

La figure 5.7 montre la vitesse relative de paires $v_{12}(r)$ en fonction de la distance de séparation r . La courbe en trait plein noir correspond au modèle Λ CDM, calculé à l'aide de la fonction de corrélation à deux points $\xi(r)$ de 2MRS extraite de Lavaux (2010). Les résultats obtenus à l'aide de l'estimateur radial de v_{12} à partir des vitesses particulières radiales extraites des catalogues fictifs 2MRS et CF3 (non-groupé et groupé) sont représentés respectivement par des cercles noirs, des triangles rouges et des carrés bleus. Les barres d'erreur correspondent aux barres d'erreur observationnelles et sont obtenues à l'aide de la méthode décrite dans la section 5.2.3.

Les résultats obtenus pour les trois mocks ne suivent pas le modèle Λ CDM, et aucune barre d'erreur n'inclut la courbe noire. Cet estimateur ne permet donc pas de retrouver le modèle Λ CDM attendu. Cela signifie que cet estimateur ne peut pas être utilisé avec les données *Cosmicflows* pour estimer le taux de croissance local. Ceci explique pourquoi Ma et al. (2015) trouvent des valeurs totalement incohérentes de Ω_m et σ_8 en utilisant v_{12} sur CF2.

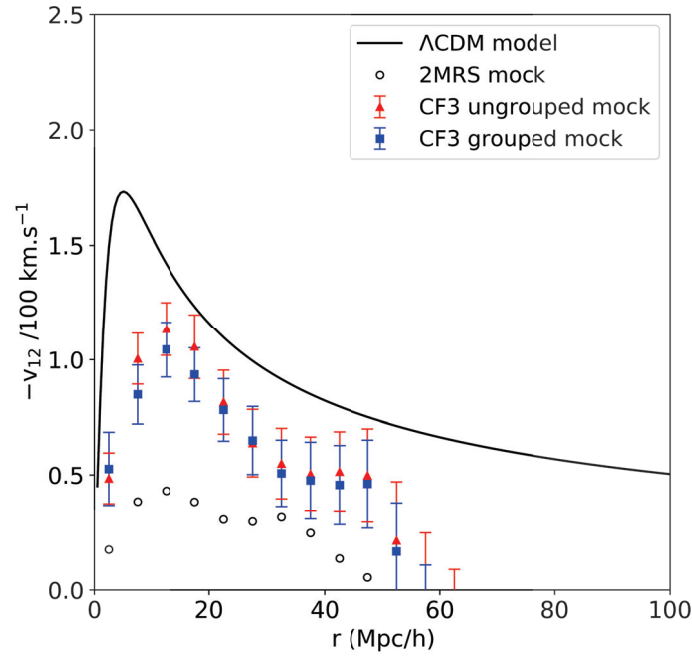


FIGURE 5.7 – Vitesse relative de paires $v_{12}(r)$, prédite par une réalisation contrainte Λ CDM des données CF3, en fonction de la distance de séparation r . Le modèle Λ CDM est représenté par la courbe en trait plein noir. Les cercles noirs, des triangles rouges et des carrés bleus représentent respectivement les résultats obtenus à l’aide de l’estimateur radial de v_{12} à partir des vitesses particulières radiales extraites des catalogues fictifs 2MRS et CF3 (non-groupé et groupé). Les barres d’erreur correspondent aux barres d’erreur observationnelles obtenues à l’aide de la méthode décrite dans la section 5.2.3.

Fonctions de corrélation des vitesses particulières

La figure 5.8 montre les statistiques à deux points ψ_1 et ψ_2 en fonction de la distance de séparation r . Les modèles Λ CDM ($\Omega_m = 0.3$, $\gamma = 0.55$, $\sigma_8 = 0.83$) calculés pour les galaxies de 2MRS, et les galaxies et les groupes de CF3 sont représentés par des courbes noire, rouge et bleue respectivement. Les résultats obtenus à l’aide des estimateurs et à partir des vitesses radiales extraites des mocks 2MRS et CF3 (non groupé et groupé) comme détaillé ci-dessus sont représentés par des cercles noirs, triangles rouges et carrés bleus respectivement.

Pour l’estimateur ψ_1 , le mock 2MRS suit le modèle jusqu’à 80 Mpc/h. Par ailleurs, l’amplitude des mocks CF3 est légèrement plus faible mais les barres d’erreur incluent le modèle Λ CDM jusqu’à 60 Mpc/h. Cela montre que la géométrie et la distribution spatiale des galaxies ou groupes du catalogue *Cosmicflows-3* ne permet pas de contraindre le taux de croissance des grandes structures pour des distances de séparation supérieures à 60 Mpc/h. Pour l’estimateur ψ_2 , la situation est similaire à ψ_1 . En effet, cet estimateur ne peut pas être utilisé pour estimer le taux de croissance au-delà de 60 Mpc/h pour les deux échantillons CF3.

De plus, la figure 5.8 montre que, en globalité pour les deux statistiques à deux points,

la différence entre les échantillons CF3 groupé et non groupé est très faible. Cela montre que les effets non-linéaires n'ont pas d'impact sur les estimateurs ψ_1 et ψ_2 , excepté pour des classes de séparations faibles qui contiennent des galaxies proches les unes des autres (i.e à l'intérieur d'amas). Lorsque l'on extrait $f\sigma_8$ à partir de ces statistiques, les effets de non-linéarité ne sont pas pris en compte : en effet, les bins correspondant à des séparations inférieures à 20 Mpc/h seront omis.

Par ailleurs, le mock 2MRS recouvre presque parfaitement pour chaque bins le modèle Λ CDM utilisé pour construire le champ de vitesse des CR en dépit de la variance cosmique. Celle-ci sera discuté sur des catalogues fictifs du type CF3 dans la section 5.3.2.

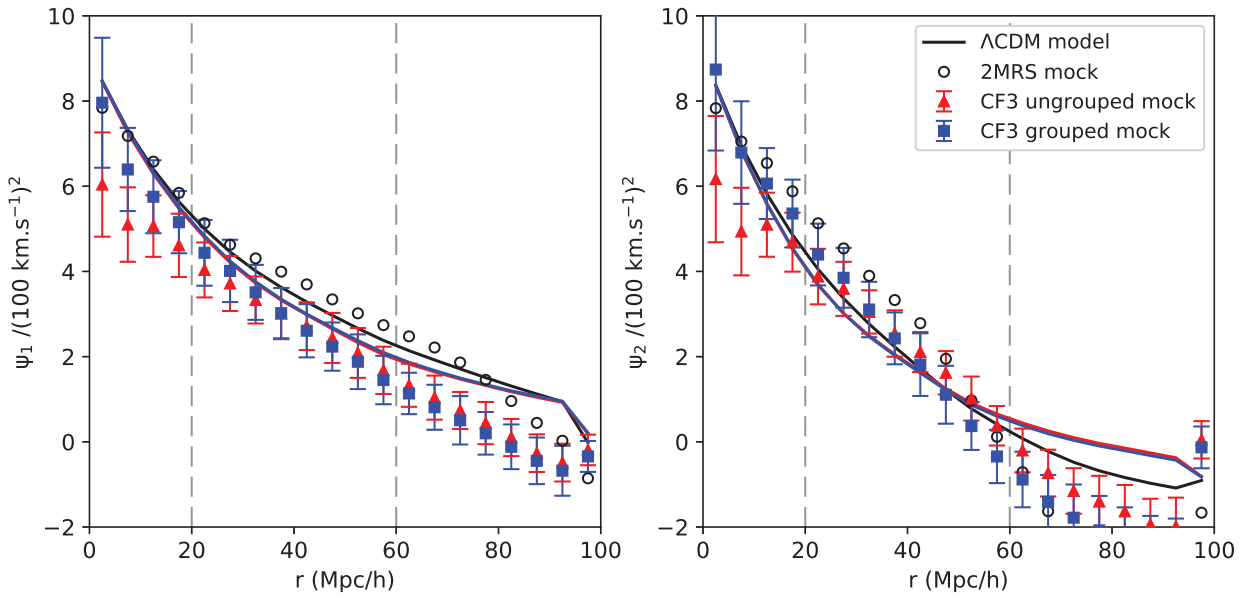


FIGURE 5.8 – Statistiques à deux points des vitesses particulières des galaxies ψ_1 (gauche) et ψ_2 (droite), prédite par une réalisation contrainte Λ CDM des données CF3, en fonction de la distance de séparation r . Le modèle Λ CDM est représenté par des courbes en traits pleins noir (2MRS), rouge (galaxies CF3) et bleu (groupes CF3). Les résultats obtenus à l'aide des estimateurs des statistiques sont représentés par des cercles noir (2MRS), triangles rouges (galaxies CF3) et carrés bleus (groupes CF3). Les lignes verticales en trait pointillés délimitent la zone où $f\sigma_8$ peut être estimé.

Pour la suite, seules les statistiques ψ_1 et ψ_2 sont considérées, la statistique v_{12} n'étant pas adaptée pour application aux données *Cosmicflows*.

5.3.2 Variance cosmique

Les tests effectués sur des données fictives illustrés sur la figure 5.8 montrent que la valeur sous-jacente de $f\sigma_8$ de la CR est retrouvée par les estimateurs ψ_1 et ψ_2 . Cependant, en considérant la taille de CF3 et de la CR utilisée, la valeur mesurée du taux de croissance à partir de ces données est sa valeur locale. À cause de la variance cosmique, il est possible que cette valeur locale ne représente pas la valeur globale du taux de croissance des grandes structures de l'univers.

Afin d'analyser l'impact de la variance cosmique sur les corrélations des vitesses et la mesure du taux de croissance, il faut réaliser des tests sur les halos de matière noire extraits de la simulation MultiDark Planck 2 (MDPL2, Prada et al., 2012) de taille 1 Gpc/h. La cosmologie sous-jacente de cette simulation est la cosmologie de Planck ($H_0 = 67.77$ km s⁻¹ Mpc⁻¹, $\Omega_\Lambda = 0.69$, $\Omega_m =$, $\sigma_8 = 0.82$). Pour construire les mocks détaillés ci-dessous, seuls les halos de masse typiquement galactique comprise entre $10^{11} M_\odot$ et $10^{12} M_\odot$ sont pris en compte.

Échantillons indépendants

Le catalogue CF3 non groupé peut être contenu (quasiment) en sa totalité dans une sphère de rayon 250 Mpc/h. Dans un cube de côté 1 Gpc/h, il est possible de placer huit de ces sphères indépendantes. Un total de 8 échantillons CF3 ont donc été générés à partir de chaque octant du cube de la simulation. L'observateur de chaque échantillon est placé au centre de son sous-cube associé. Ensuite, les composantes radiales des vitesses particulières sont extraites aux positions des galaxies CF3 centrées sur ce nouvel observateur. Les échantillons sont ainsi complètement indépendants, chaque halo de la simulation est associé au maximum à un seul échantillon.

Les statiques ψ_1 et ψ_2 obtenues pour ces huit échantillons sont présentées sur la figure 5.9 à gauche et à droite respectivement. La courbe en trait plein noir correspond au modèle Λ CDM pour le catalogue non-groupé CF3. Les triangles rouges avec barres d'erreur représentent la moyenne et la déviation standard des résultats obtenus à l'aide des estimateurs associés aux statistiques pour les huit échantillons. La cosmologie sous-jacente de cette simulation est bien retrouvée par les deux estimateurs $\psi_1(r)$ et $\psi_2(r)$ car toutes les barres d'erreurs incluent le modèle Λ CDM.

Par ailleurs, cette figure montre que la variance cosmique a un effet plus important sur l'estimateur ψ_2 , les barres d'erreur étant plus grandes et l'amplitude étant plus élevée que la valeur attendue prédite par CDM, y compris dans l'intervalle de robustesse 20 – 60 Mpc/h dans laquelle $f\sigma_8$ peut être mesuré (voir section 5.3.1). Par conséquent, l'estimateur ψ_2 étant moins robuste, il ne sera pas considéré pour la suite de ce chapitre.

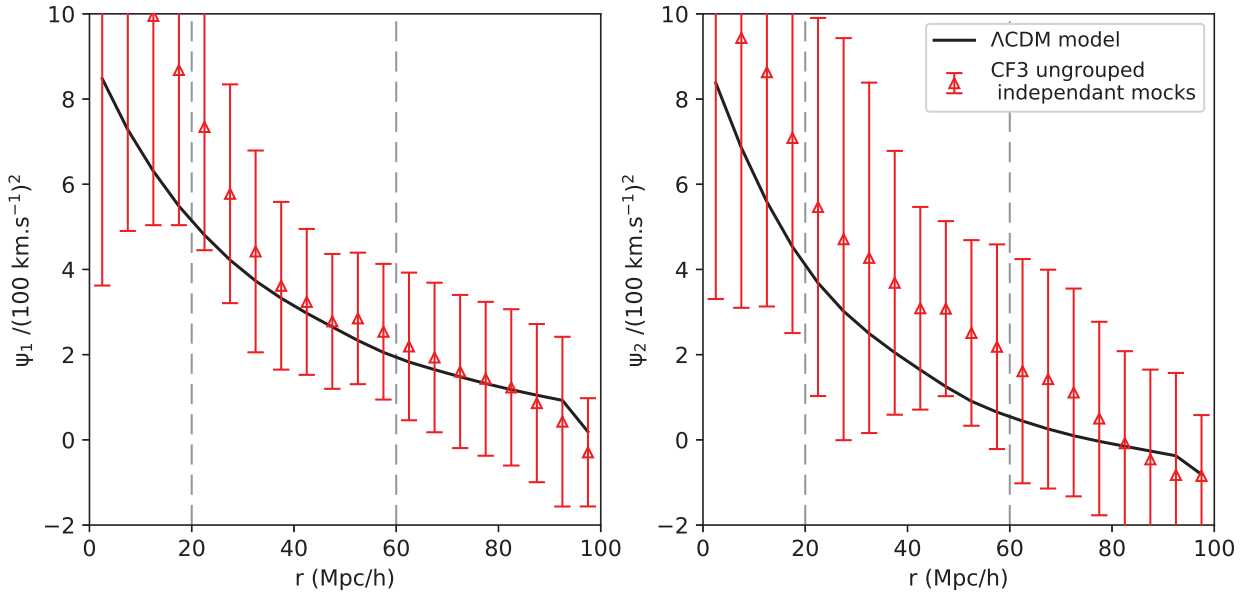


FIGURE 5.9 – Statistiques ψ_1 (gauche) et ψ_2 (droite) obtenues pour les huit échantillons indépendants du type CF3 non groupé. La courbe en trait plein noir correspond au modèle Λ CDM pour le catalogue non-groupé CF3. Les triangles rouges avec barres d'erreur représentent la moyenne et la déviation standard des résultats obtenus à l'aide des estimateurs associés aux statistiques pour les huit échantillons.

Observateurs aléatoires

Cependant, il est difficile de conclure avec seulement 8 échantillons. Il est nécessaire de considérer un plus grand nombre d'échantillon afin d'avoir une estimation plus correcte des erreurs dues à la variance cosmique sur l'estimateur ψ_1 .

Pour cette raison, un total de 100 nouveaux échantillons ont été générés, à partir de la même grande simulation, avec la méthode suivante. Au lieu de placer huit observateurs de façon à ce que les échantillons n'aient pas de halos en commun, les observateurs sont placés à des positions choisies aléatoirement dans la grille. Ensuite, les composantes radiales des vitesses particulières sont extraites aux positions des galaxies CF3 centrées sur chaque nouvel observateur. Il y a cependant un inconvénient : un même halo peut maintenant appartenir à plusieurs échantillons, et les résultats sont donc hautement corrélés.

La figure 5.10 montre la statistique ψ_1 obtenue avec ces 100 échantillons. La courbe en trait plein noir correspond au modèle Λ CDM pour le catalogue non-groupé CF3. Les triangles rouges avec barres d'erreur représentent la moyenne et la déviation standard des résultats obtenus à l'aide de l'estimateur observationnel de ψ_1 pour les cent échantillons. La cosmologie sous-jacente de cette simulation est bien retrouvée par l'estimateur $\psi_1(r)$. Cela montre qu'en dépit de leur taille, les échantillons du type CF3 pourraient permettre une mesure du taux de croissance des grandes structures. Cependant les erreurs dues à la variance cosmique sont considérables et ne doivent pas être ignorées lorsque l'on contraint

$f\sigma_8$ avec les données CF3.

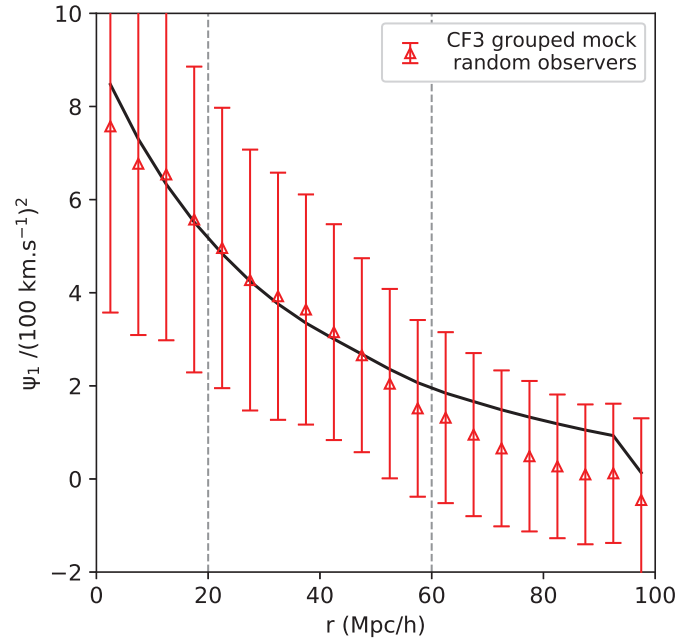


FIGURE 5.10 – Statistique ψ_1 obtenue pour les cent échantillons, du type CF3 non groupé, dont les observateurs ont été placés aléatoirement. La courbe en trait plein noir correspond au modèle Λ CDM pour le catalogue non-groupé CF3. Les triangles rouges avec barres d'erreur représentent la moyenne et la déviation standard des résultats obtenus à l'aide de l'estimateur observationnel de ψ_1 pour les cent échantillons.

5.3.3 Constante de Hubble

Le calcul des vitesses particulières des galaxies à partir de leur distance observée se fait par l'équation 2.9, qui dépend de la constante de Hubble H_0 . Par conséquent, le calcul des vitesses particulières nécessite de supposer une valeur de H_0 . La dépendance des statistiques des vitesses particulières à deux points nécessite donc d'être étudiée.

La figure 5.11 montre ψ_1 en fonction de la distance de séparation r . La courbe noire correspond au modèle Λ CDM. Celui-ci ne dépend de H_0 qu'à travers le spectre de puissance, mais sa forme et son amplitude ne sont pas modifiées. L'estimateur ψ_1 a été calculé à partir des vitesses particulières des groupes CF3 pour différentes valeurs de H_0 : 66, 68, 70, 72, 73, 74, 75, 76, 77 et 79 km/s/Mpc. Cette figure montre que les valeurs qui permettraient aux données d'être en accord avec le modèle Λ CDM sont $H_0 = 70, 71, 72$ km/s/Mpc. Les autres valeurs donnent des résultats non-raisonnables. Cependant, choisir $H_0 = 70, 71, 72$ km/s/Mpc pour la suite, i.e l'estimation du taux de croissance sur les données CF3, n'est pas raisonnable et biaiserait les résultats.

Néanmoins, cette figure montre que ψ_1 pourrait permettre de contraindre à la fois le taux de croissance $f\sigma_8$, et la constante de H_0 , à partir de catalogues observationnels de vitesses particulières.

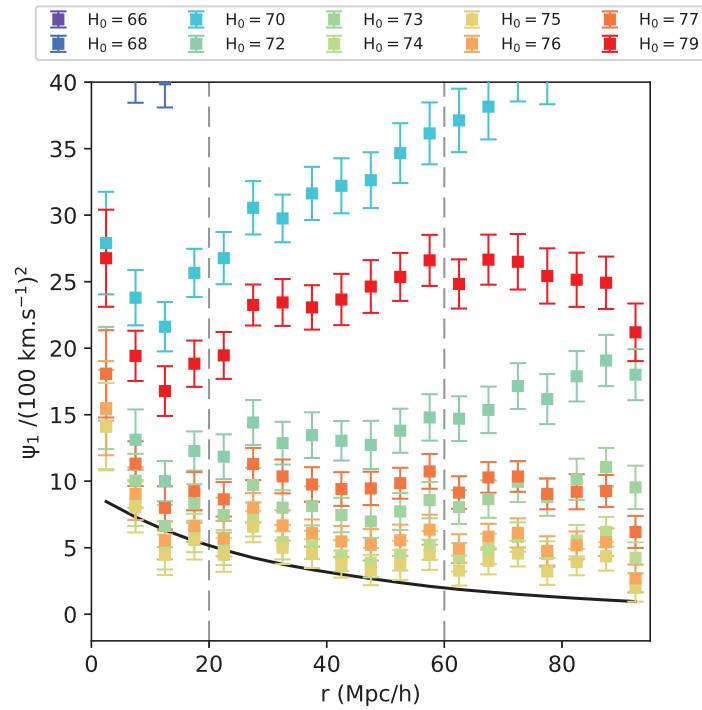


FIGURE 5.11 – Statistique ψ_1 en fonction de la distance de séparation r . La courbe noire correspond au modèle Λ CDM. L'estimateur ψ_1 a été calculé à partir des vitesses particulières des groupes CF3 pour différentes valeurs de H_0 : 66, 68, 70, 72, 73, 74, 75, 76, 77 et 79 km/s/Mpc.

5.3.4 Taux de croissance local avec les données *Cosmicflows-3*

Pour terminer, cette dernière section présente les résultats obtenus à partir de la statistique à deux points des vitesses particulières ψ_1 , appliquée aux vitesses particulières observées provenant du catalogue *Cosmicflows-3*.

Statistiques mesurées sur le catalogue *Cosmicflows-3*

La figure 5.12 montre les résultats du calcul de ψ_1 sur les vitesses particulières observées du catalogue *Cosmicflows-3*. Les échantillons non-groupés et groupés sont représentés respectivement par des triangles rouges et des carrés bleus.

Les barres d'erreurs incluent à la fois les incertitudes observationnelles (estimées en propageant l'erreur sur le module de distance, voir section 5.2.3) et celles dues à la variance cosmique. Ces dernières correspondent aux erreurs obtenues sur les deux statistiques sur les cent échantillons préparés en plaçant les observateurs aléatoirement dans une grande simulation (voir section 5.3.2). Les statistiques prédites par différents modèles cosmologiques sont illustrées par des lignes pleines (Λ CDM, $f\sigma_8 = 0.4$) et en pointillés ($f\sigma_8 = 0.2, 0.3, 0.5, 0.6$).

Cet estimateur, auquel les vitesses particulières des galaxies du catalogue CF3 ont été appliquées, et à l'intérieur de son intervalle de robustesse (indiqué par des traits pointillés verticaux), est cohérent avec le modèle Λ CDM, comme prédit par les tests sur les données fictives effectués en section 5.3.1. Par ailleurs, les résultats obtenus à partir des vitesses particulières des groupes de CF3 ont une amplitude légèrement plus élevée.

Estimation du taux de croissance local

Les données sur la figure 5.12 sont comparées aux modèles décrits dans l'équation 5.16 en minimisant le χ^2 défini dans l'équation 5.20 afin d'extraire le taux de croissance $f\sigma_8$. Les comparaisons sont faites à l'intérieur de l'intervalle $[20, 60]$ Mpc h^{-1} de distances de séparations.

Les résultats sont présentés dans la Table 5.1 qui montre les valeurs de $f\sigma_8$ (colonne (2)) extraites directement à partir des données observationnelles *Cosmicflows-3*. L'échantillon (galaxies ou groupes) et la statistique (ψ_1 ou ψ_2) utilisés pour mesurer le taux de croissance sont indiqués dans la colonne (1). Les incertitudes de mesure observationnelles et celles dues à la variance cosmique sur le taux de croissance sont toutes les deux présentées, dans les colonnes (3) et (4) respectivement.

La statistique ψ_1 mène à une valeur $f\sigma_8(z = 0.05) = 0.43 (\pm 0.03)_{\text{obs}} (\pm 0.11)_{\text{cosmic}}$ pour les données non-groupées, la couverture en redshift des données *Cosmicflows-3* s'étendant jusqu'à $z = 0.05$. Les erreurs indiquées par les indices "obs" et "cosmic" correspondent aux erreur observationnelles et de variance cosmique respectivement. Cela permet de confirmer localement le modèle Λ CDM, avec une erreur de 25%.

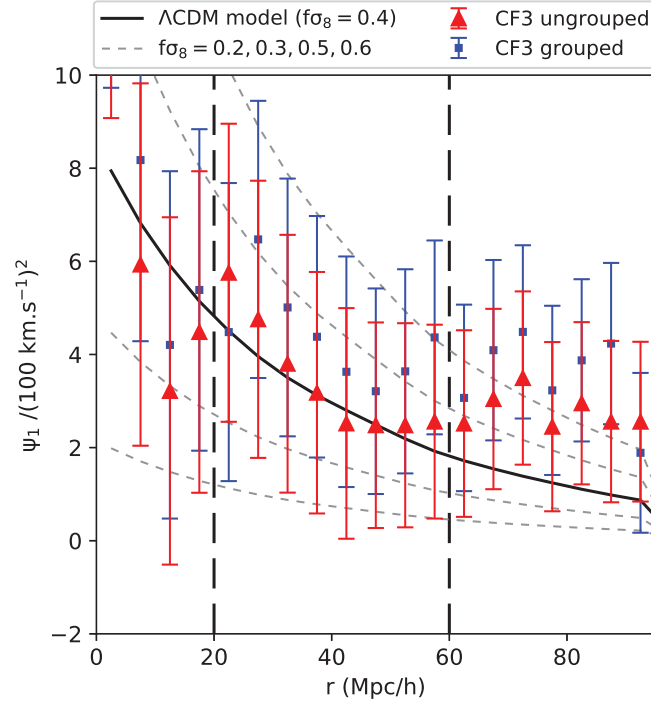


FIGURE 5.12 – Statistique ψ_1 en fonction de la distance de séparation r . Le modèle Λ CDM, i.e $f\sigma_8 \approx 0.4$, prédit pour les galaxies CF3 est représenté par une courbe en trait plein noir. Les statistiques prédites par d’autres modèles cosmologiques : $f\sigma_8 = 0.2, 0.3, 0.5, 0.6$; sont représentés par des courbes en traits pointillés noirs. Les points disséminés avec des barres d’erreurs correspondent au résultats obtenus à partir des vitesses particulières observées issues du catalogue *Cosmicflows-3*. L’échantillon non-groupé est représenté par des triangles rouges, tandis que l’échantillon groupé est représenté par des carrés bleus. Les lignes verticales en trait pointillé délimitent la zone dans laquelle le taux de croissance peut être estimé.

TABLE 5.1 – Mesures du taux de croissance normalisé, obtenues à partir des statistiques ψ_1 and ψ_2 calculées sur les données observationnelles *Cosmicflows-3*. Les colonnes (3) et (4) correspondent aux erreurs observationnelles $\Delta f\sigma_8^{\text{obs}}$ et de variance cosmique $\Delta f\sigma_8^{\text{cosmic}}$ respectivement.

Échantillon CF3 — statistique	$f\sigma_8$		$\Delta f\sigma_8^{\text{obs}}$		$\Delta f\sigma_8^{\text{cosmic}}$
CF3 non-groupé — ψ_1	0.43	\pm	0.03	\pm	0.11
CF3 non-groupé — ψ_2	0.51	\pm	0.03	\pm	0.12
CF3 groupé — ψ_1	0.45	\pm	0.04	\pm	0.11
CF3 groupé — ψ_2	0.52	\pm	0.04	\pm	0.12

La Figure 5.13 montre le taux de croissance normalisé $f\sigma_8$ en fonction de redshift z . Cette figure est celle qui est montrée dans le chapitre 1 (figure 1.4) à laquelle le résultat obtenu durant cette thèse a été ajouté.

Pour rappel, la courbe en trait plein noir correspond au modèle Λ CDM ($\gamma = 0.55$). D'autres modèles cosmologiques (gravité modifiée avec $\gamma \neq 0.55$) sont représentés par des lignes noires en pointillé et en tiret. Les points disséminés avec des barres d'erreurs montrent diverses contraintes du taux de croissance extraites de la littérature qui sont présentés dans la table 1.1. Les couleurs dépendent de la sonde cosmologique utilisée pour contraindre le taux d'expansion : SN Ia en jaune (Turnbull et al., 2012), les vitesses particulières des galaxies en bleu (Johnson et al., 2014), les oscillations acoustiques des baryons en vert (Blake et al., 2011; Reid et al., 2012), et les distortions dans l'espace des redshift en rose (Hawkins et al., 2003; Samushia et al., 2012; Beutler et al., 2012; de la Torre et al., 2013). Les prédictions des mesures de taux de croissance pour de futurs sondages cosmologiques sont montrées en rouge, avec une barre d'erreur prédite de 10% (McConnachie et al., 2016; Amendola et al., 2016; da Cunha et al., 2017).

La mesure locale du taux de croissance, obtenue dans cette thèse à partir des vitesses radiales particulières des galaxies du catalogue CF3, est représentée dans la figure 5.13 par un point bleu avec des barres d'erreur. La méthodologie utilisée pour mesurer le taux de croissance n'était pas applicable aux paires de groupes de galaxies, le résultat obtenu pour CF3 groupé n'est pas tracé sur cette figure. Les barres d'erreur incluent seulement les erreurs observationnelles, par cohérence avec tous les autres points de mesure présentés sur cette figure. Les erreurs dues à la variance cosmique ne sont pas tracées. Le résultat obtenu à partir des données CF3 conforte le modèle Λ CDM avec Λ CDM puisqu'il recouvre en effet la courbe $\gamma = 0.55$ avec $z = 0.05$.

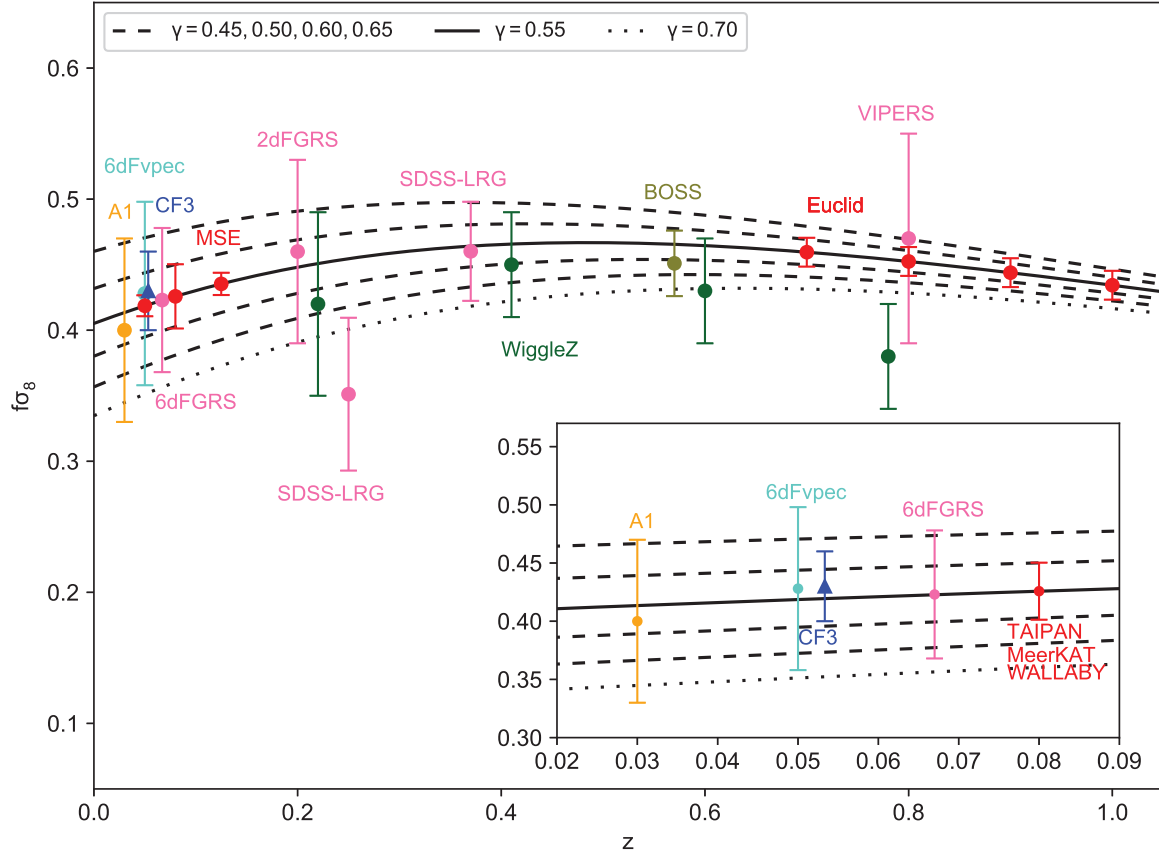


FIGURE 5.13 – Taux de croissance normalisé $f\sigma_8$ en fonction du redshift z . Le modèle Λ CDM ($\gamma = 0.55$) est représenté par une courbe en trait plein noir. Les lignes noires en pointillé représentent d'autres modèles cosmologiques (gravité modifiée avec $\gamma \neq 0.55$). Les points disséminés avec barres d'erreur correspondent à diverses contraintes de $f\sigma_8$ issues de la littérature et listées dans la table 1.1. Les couleurs dépendent de la sonde cosmologique utilisée pour estimer le taux de croissance : les SN Ia en jaune, les vitesses particulières en bleu, les oscillations acoustiques des baryons en vert et les distortions dans l'espace des redshifts en rose. Les prévisions pour de futurs relevés cosmologiques sont représentées en rouge. Le résultat obtenu durant cette thèse à partir des vitesses radiales particulières des galaxies du catalogue CF3, est représenté par un point bleu avec des barres d'erreur. Ces dernières incluent seulement les erreurs observationnelles, par cohérence avec tous les autres points de mesure présentés sur cette figure.

Chapitre 6

Conclusion et perspectives

Afin de conclure cette thèse basée sur les techniques d'analyse des données observationnelles telles que les catalogues *Cosmicflows*, quelques perspectives et ouvertures qui pourraient être engagées par la suite sont proposées dans ce dernier chapitre.

Catalogues Cosmicflows

Cette thèse a exposé l'analyse de diverses compilations de données HI, provenant des radiotélescopes Nançay (KLUN et NIBLES) et Arecibo (ALFALFA). Sont aussi présentés les premiers résultats obtenus lors d'observations conduites au radiotélescope de Greenbank de Décembre 2017 à Juillet 2018. Les largeurs de raies obtenues à partir de ces études et observations seront utilisées pour ajouter plus de 4,000 nouvelles distances de galaxies au catalogue *Cosmicflows-4*. Ainsi, le prochain catalogue *Cosmicflows-4* contiendrait, à ce jour, 42% de galaxies dans l'hémisphère terrestre Nord, et 58% dans l'hémisphère Sud. Cela représente une belle amélioration : l'asymétrie Nord - Sud de *Cosmicflows-3* a été corrigée avec succès¹. Et ce n'est pas terminé ! Le programme d'observation *Cosmicflows* au télescope de Greenbank est toujours en cours, et ne se termine que fin Janvier 2019. Cela permettra d'ajouter encore un grand nombre de galaxies à *Cosmicflows-4*. Cependant, les observations au télescope de Greenbank sont gênées par l'apparition d'un signal GPS vers 8,000 km/s. Cela ruine les spectres des galaxies lointaines, les plus intéressantes. De telles interférences sont vues à Nançay ([van Driel et al., 2016b](#)) et à Greenbank ([Hunt et al., 2016](#)), et il est très difficile de les supprimer tout en gardant le signal HI. Il est prévu d'étudier plus en détail ce signal GPS afin de sauver les spectres des galaxies concernées.

Par ailleurs, l'arrivée de nouveaux télescopes et sondages cosmologiques permettra aux compilations *Cosmicflows* d'aller encore plus loin. Le sondage de galaxies TAIPAN, conduit au *1.2m UK Schmidt Telescope* installé en Australie, a démarré en 2018 et prévoit d'obtenir 40,000 mesures de distances de galaxies elliptiques d'ici 2020. De plus, le sondage HI WALLABY, installé au réseau de 36 radiotélescopes *Australian Square Kilometre Array Pathfinder* (ASKAP), prévoit, lui aussi, la mesure de 40,000 distances de galaxies spirales. Il est cependant possible d'obtenir davantage de mesures : il se peut que WALLABY obtienne jusqu'à 80,000 distances, ce type de sondages étant assez rapide. L'arrivée

1. Pour rappel, seulement 30% des galaxies du catalogue *Cosmicflows-3* sont situées dans l'hémisphère terrestre Nord (voir chapitre 2).

d'un autre radiotélescope, MeerKAT situé en Afrique du Sud, permettra l'obtention de distances de galaxies spirales supplémentaires dans l'hémisphère Sud.

Tous ces télescopes et sondages sont situés dans l'hémisphère Sud. Ils seront complétés dans l'hémisphère Nord par le sondage de supernovae de type Ia *Zwicky Transient Facility* (ZTF), qui prévoit d'observer 100,000 SNIa en 10 ans jusqu'à 90,000 km/s, et par un sondage, prévu avec l'installation d'un nouvel instrument au télescope de 3.5m à Calar Alto en Espagne, et qui apporterait 20,000 distances de galaxies en deux ans. Enfin, l'arrivée du radiotélescope *Five hundred meter Aperture Spherical Telescope* (FAST) en Chine pourrait permettre de compléter WALLABY et MeerKAT, mais ses performances à 21 cm ne sont pas encore connues.

Concernant les sondages HI en cours EBHIS et ALFALFA, les spectres HI ne sont pas encore publics. Dès qu'ils le seront, ils seront re-analysés afin de mesurer la largeur de raie de chaque spectre avec la méthode *Cosmicflows*. L'analyse des données obtenues par d'autres collaborations, NIBLES et ALFALFA, ont montré que les mesures de largeurs de raie par la collaboration *Cosmicflows* sont en moyenne plus grandes que celles des autres collaboration. Il est donc important pour la collaboration *Cosmicflows* de re-calculer ces largeurs de raie avant de les insérer dans le catalogue ADHI.

Algorithme de recherche de bassins et vallées gravitationnelles

Cette thèse a présenté des méthodes d'exploitation des champs de vitesse tri-dimensionnels reconstruits à l'aide de données observationnelles telles que les catalogues *Cosmicflows*. Ces techniques permettent d'identifier et caractériser des bassins et vallées gravitationnelles dans l'univers local. Elles ont d'ailleurs parmi, pour la première fois, de tracer les profils de densité de bassins gravitationnels.

Jusqu'à présent, les grandes structures de l'univers local ont été caractérisées par diverses techniques telles que la méthode DisPerSE (Sousbie, 2011), recherchant le *squelette* de l'univers local, et la méthode BISOUS (Tempel et al., 2014) qui extrait la structure filamentaire de l'univers. Toutes ces techniques exploitent la distribution des galaxies dans des sondages de décalages spectraux, i.e le champ de surdensité. Très peu de méthodes utilisent les vitesses particulières des galaxies et le champ de vitesse associé, et encore moins les vitesses particulières observées. À partir du champ de vitesse estimé à l'aide des galaxies sondage SDSS, Leclercq et al. (2017) identifie des amas, filaments et vides, mais cette méthode n'a pas été appliqué à des vitesses particulières observationnelles.

La technique d'identification des bassins et vallées gravitationnelles présentée dans cette thèse représente ainsi une des premières méthodes d'analyse de champs de vitesse particulière observée. Cette méthode a été appliquée à la version courante du champ de vitesse reconstruit à partir des vitesses particulières du catalogue *Cosmicflows-3*. Il est cependant difficile, à ce jour, d'en tirer des résultats et conclusion, le champ de vitesse CF3 étant toujours en préparation et non finalisé. Différentes versions du champ de vitesse ont jusqu'à maintenant été préparées, avec des valeurs de la constante de Hubble différentes, la prise en compte ou non des outliers, etc (voir Graziani et al., 2018). La reconstruction des champs de vitesse particulière est un art délicat. Bien que les données CF3 aient été rendues publiques en 2016, aucun consortium n'a encore réussi à publier un champ de vitesse tridimensionnel car ces études sont très complexes et très longues, même dans la

collaboration *Cosmicflows*, qui a une meilleure compréhension du catalogue CF3.

Les résultats obtenus varient d'une version à l'autre, comme le montrent les figures 6.1 et 6.2. Ces deux figures sont les mêmes que les figures 4.23 et 4.24 respectivement, mais sont obtenues à partir d'une version précédente du champ de vitesse CF3. Un plus grand nombre de bassins d'attraction sont obtenus : Shapley I, Shapley II, Laniakea/Coma, Persée-Poisson, Pisces-Cetus. Dans cette version, les bassins de Lepus et Hercule sont aussi identifiés mais ne sont pas visibles dans la figure 6.1. Cependant, sur la figure 6.2, les rivières gravitationnelles autour du superamas Vela sont beaucoup moins marquées que sur la version courante. Il faut donc continuer de travailler sur les méthodes de reconstruction tridimensionnelle du champ de vitesse particulière avant de tirer des conclusions sur les bassins versants, et sur le superamas Vela.

Taux de croissance des grandes structures

Les vitesses particulières peuvent être utilisées comme sondes cosmologiques pour tester le modèle Λ CDM. Cette thèse a proposé une méthode permettant d'estimer le taux de croissance des grandes structures de l'univers à partir des vitesses particulières observées (radiales) seulement. Cette technique a été appliquée aux données non-groupées *Cosmicflows-3* et a ainsi permis de mesurer le taux de croissance de l'univers local : $f\sigma_8(z = 0.05) = 0.43 (\pm 0.03)_{\text{obs}} (\pm 0.11)_{\text{cosmic}}$ (erreurs observationnelles et variance cosmique). Cela permet de confirmer localement le modèle Λ CDM, avec une erreur de 25%. Il a aussi été vu que cette méthode n'est pas adaptée pour être appliquée sur les catalogues groupés *Cosmicflows*.

L'arrivée des nouveaux sondages discutés un peu plus haut va permettre l'obtention de mesures beaucoup plus précises, notamment du côté de la variance cosmique. Le catalogue *Cosmicflows-3* ne permet pour l'instant d'obtenir une grande précision sur la mesure de $f\sigma_8$ à cause de la taille de ce catalogue et de l'impact de la variance cosmique. Les futurs catalogues auront une couverture en redshift beaucoup plus grande et permettront ainsi d'atteindre une précision plus haute, et de diminuer l'erreur due à la variance cosmique.

Une ouverture possible au sujet du taux de croissance des grandes structures serait d'utiliser les algorithmes d'identification des bassins et vallées gravitationnelles pour mesurer ce paramètre, en mesurant la *compacité* des bassins versants segmentés.

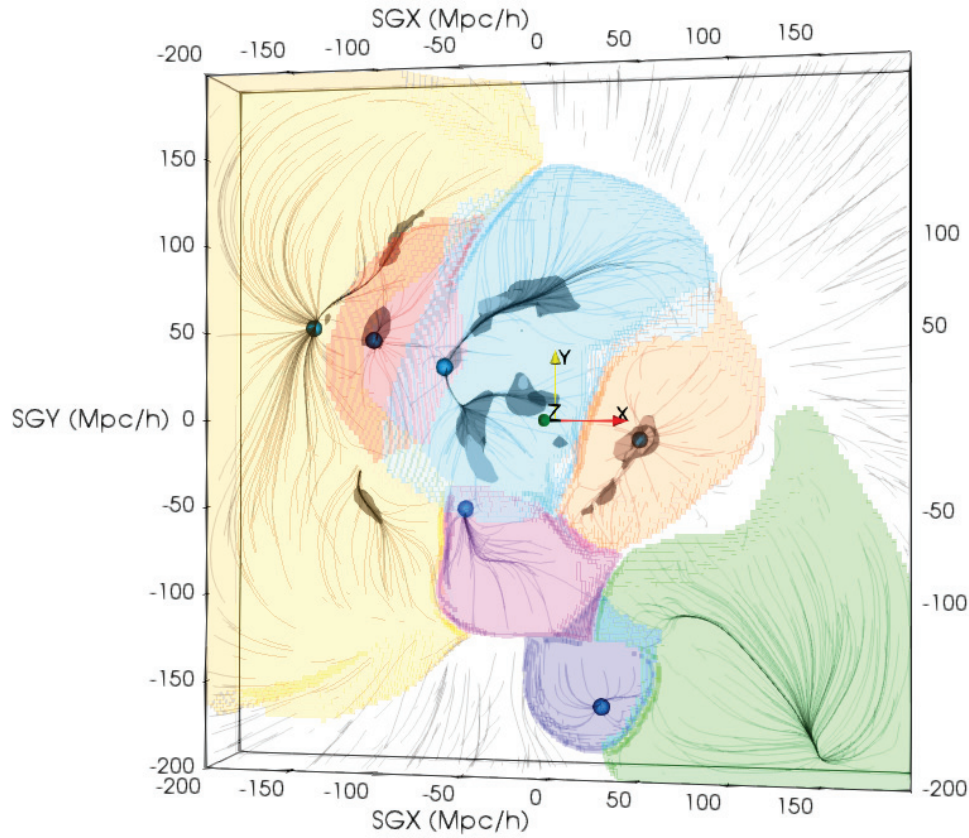


FIGURE 6.1 — Figure 4.23 pour une version légèrement différente (constante de Hubble, outliers) du champ de vitesse obtenu à l'aide des données *Cosmicflows-3*. — Tranche SGX - SGY d'épaisseur $-25 < SGZ < 25$ Mpc/h. Les isosurfaces noires correspondent aux surdensités telles que $\delta > 1.5$, et les lignes de flux sont tracées en point de vue avant. Les isosurfaces colorées correspondent aux bassins d'attraction segmentés.

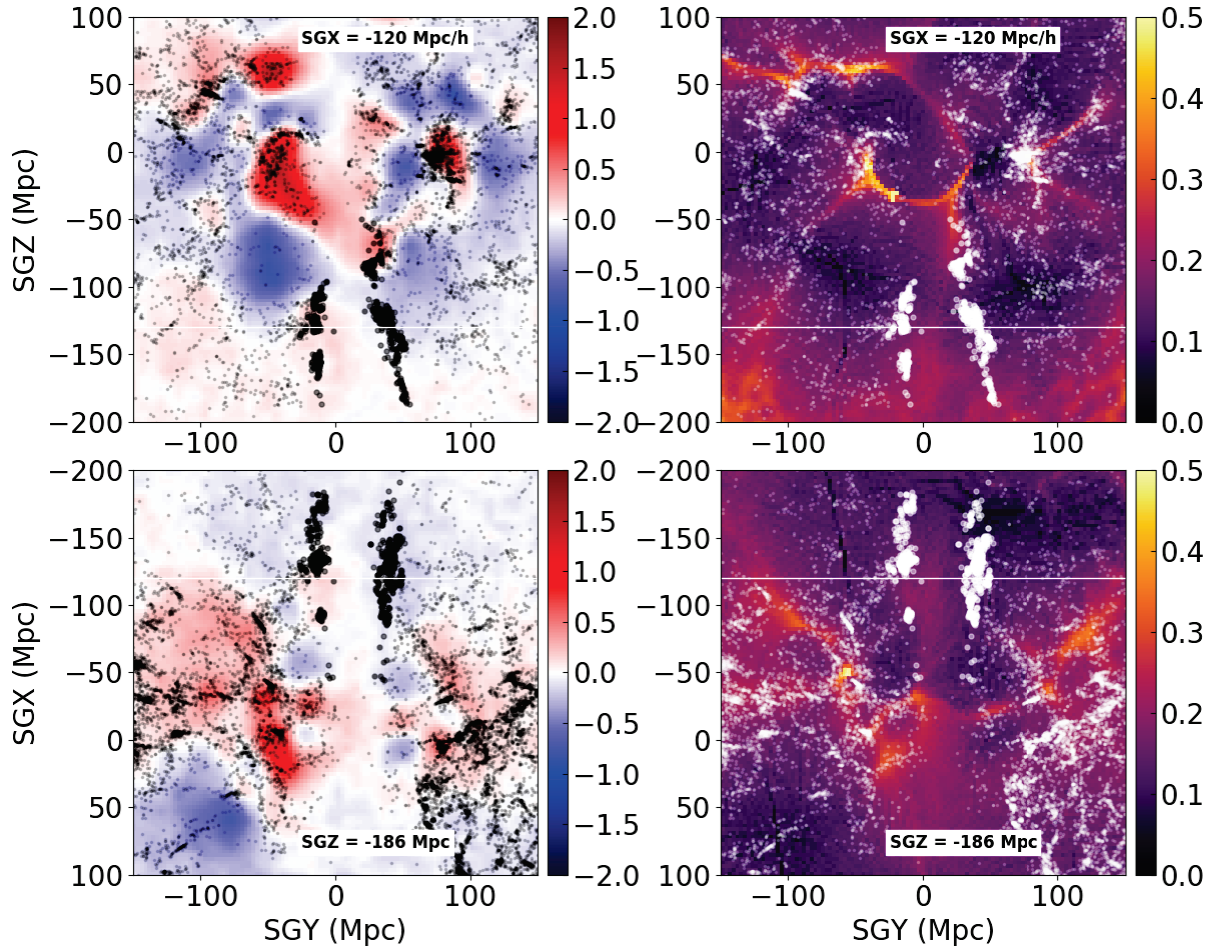


FIGURE 6.2 — Figure 4.24 pour une version légèrement différente (constante de Hubble, outliers) du champ de vitesse obtenu à l’aide des données *Cosmicflows-3*. — Tranches SGX - SGY d’épaisseur $-25 < SGZ < 25 \text{ Mpc/h}$. Les petits points noirs correspondent aux galaxies d’une compilation de redshifts extraite de *LEDA*. Les plus gros points noirs correspondent aux galaxies du sondage de redshift de *Kraan-Korteweg et al. (2017)*. À gauche, la palette de couleur représente le champ de surdensité δ_{CF3} , tandis qu’à droite, elle représente l’histogramme des lignes de flux obtenu à partir du champ de vitesse *CF3* (point de vue avant).

Appendices

Annexe A

Observations au GBT : paramètres HI des galaxies adéquates pour Cosmicflows-4

TABLE A.1 – Paramètres HI extraits des spectres à 21 cm obtenus à partir des observations conduites par la collaboration *Cosmicflows* au radiotélescope de Greenbank. Les colonnes sont les suivantes : (1) numéro PGC de la galaxie ; (2) et (3) coordonnées J2000 ascension droite (en heures) et déclinaison (en degrés) ; (4) vitesse héliocentrique (en km/s) ; (5) flux intégré dans la raie HI ; (6) largeur de raie à 21 cm (non corrigée, en km/s) ; (7) erreur sur la largeur de raie à 21 cm (en km/s) ; (8) ratio signal sur bruit ; (9) logarithme décimal de la masse HI (divisée par la masse solaire) ; (10) magnitude à 21 cm.

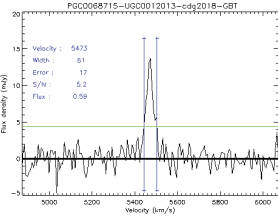
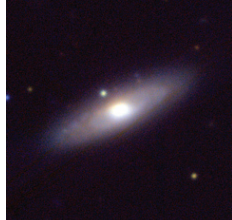
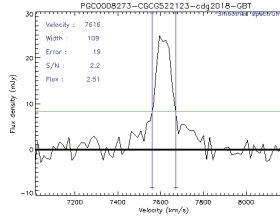

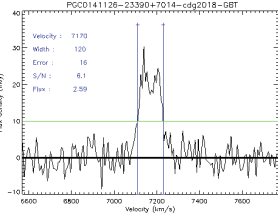
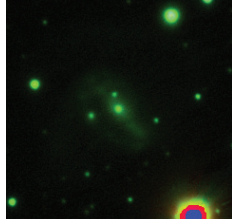
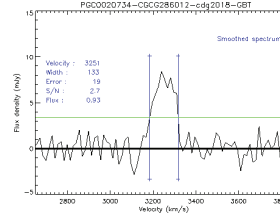

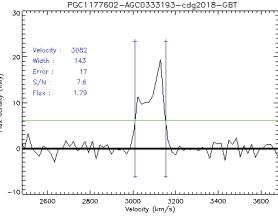

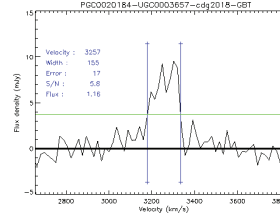

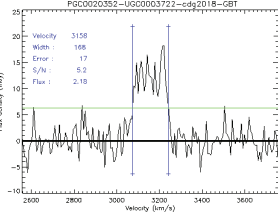

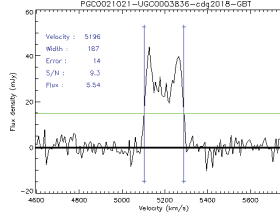
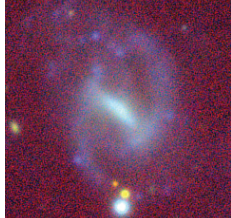
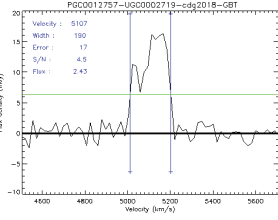

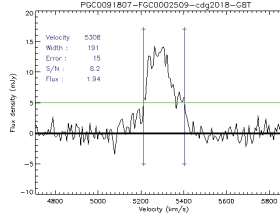

PGC	RA h	DEC deg	V_{hel} km s ⁻¹	Int. Fl. Jy.km s ⁻¹	W_{50} km s ⁻¹	e_W km s ⁻¹	SNR	$\log_{10}(M_{HI})$ M_{\odot}	m_{21}
PGC0068715	22.38	36.05	5473	0.59	61	17	5.20	8.86	17.97
PGC0141126	23.69	70.51	7170	2.59	120	16	6.10	9.74	16.37
PGC1177602	23.02	0.86	3082	1.79	143	17	7.60	8.85	16.77
PGC0020352	7.19	54.38	3158	2.18	168	17	5.20	8.96	16.55
PGC0012757	3.42	38.94	5107	2.43	190	17	4.50	9.42	16.44
PGC0019998	7.01	38.63	5186	2.25	197	17	4.90	9.40	16.52
PGC0090426	23.47	47.40	7368	2.04	206	15	7.20	9.66	16.63
PGC0090769	7.12	44.66	5873	1.20	210	18	4.20	9.24	17.20
PGC0165319	1.38	50.92	7493	1.68	217	27	1.70	9.59	16.84
PGC0069135	22.56	38.94	5451	5.94	228	18	14.60	9.86	15.47
PGC0004401	1.23	49.61	5450	2.00	230	15	7.70	9.39	16.65
PGC0014215	3.98	43.35	4594	1.85	232	18	3.30	9.21	16.73
PGC0019447	6.69	64.00	6566	2.33	249	18	4.10	9.62	16.48
PGC0022527	8.03	75.94	5527	2.25	251	17	5.10	9.45	16.52
PGC0068316	22.22	38.83	6355	1.40	257	18	3.50	9.37	17.03
PGC0001414	0.37	49.13	5121	1.86	260	17	5.60	9.31	16.73
PGC0004184	1.17	43.29	4897	8.85	263	17	20.90	9.95	15.03
PGC0068871	22.44	40.17	4669	1.11	263	34	1.40	9.00	17.29
PGC0008844	2.32	42.73	6056	4.35	266	15	7.10	9.82	15.80
PGC0009480	2.49	42.25	5641	2.90	270	18	3.30	9.59	16.24
PGC0011926	3.20	40.24	4267	3.17	273	18	4.50	9.37	16.15
PGC0000676	0.16	47.36	5158	4.02	276	16	6.40	9.64	15.89
PGC0070716	23.22	36.46	5229	0.91	279	20	1.60	9.01	17.50
PGC0071231	23.37	50.16	8937	1.89	285	18	3.60	9.80	16.71
PGC0166736	22.07	36.90	5884	1.17	291	18	3.30	9.23	17.23
PGC0018389	6.10	82.72	5393	3.57	295	17	5.10	9.63	16.02
PGC0004390	1.22	38.89	6276	7.17	297	15	7.70	10.07	15.26
PGC0012812	3.43	42.06	6535	2.39	298	18	3.50	9.63	16.45
PGC0069182	22.57	41.32	5180	1.52	305	19	2.80	9.23	16.95
PGC0090374	20.83	29.61	4775	1.85	306	19	2.00	9.24	16.73
PGC0165921	23.34	68.50	3465	13.16	307	9	15.60	9.82	14.60
PGC0018128	5.95	76.51	5005	3.06	315	17	4.60	9.50	16.19
PGC0060434	17.52	59.49	8486	2.78	321	19	3.00	9.92	16.29
PGC0021013	7.44	58.33	7670	1.18	329	19	2.30	9.46	17.22
PGC2412511	1.39	52.34	5873	7.05	333	18	7.10	10.01	15.28
PGC0001751	0.47	-7.39	7640	2.12	341	19	2.80	9.71	16.58
PGC0021712	7.75	70.03	3882	7.98	343	15	8.00	9.69	15.14
PGC0020380	7.20	41.78	5326	4.59	348	17	5.30	9.74	15.75

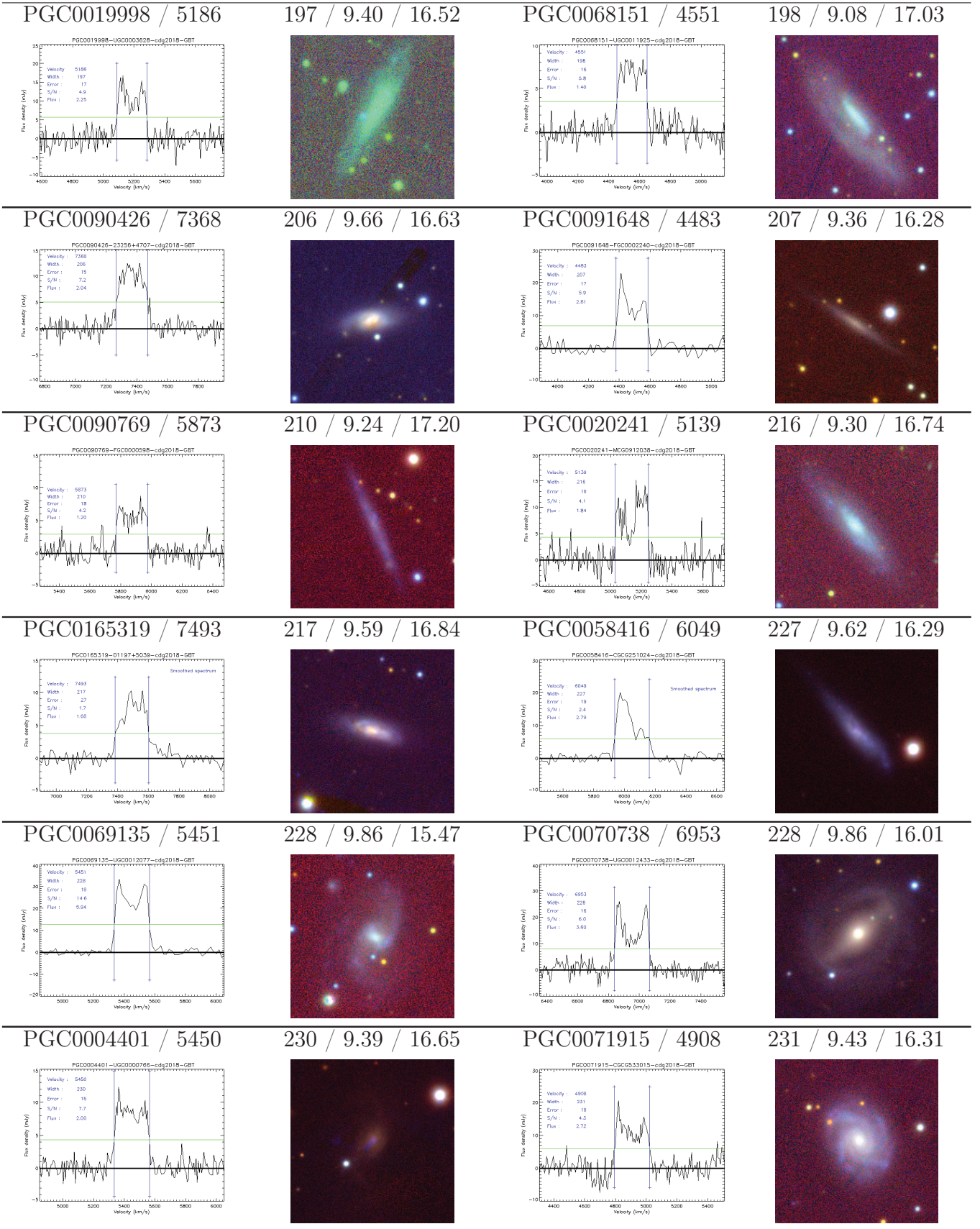
PGC0022301	7.96	66.44	4911	7.07	354	17	5.40	9.85	15.28
PGC0090761	6.69	45.62	6225	1.65	359	19	3.20	9.43	16.86
PGC0090722	4.74	-4.04	9503	1.31	364	27	1.70	9.69	17.11
PGC0063347	19.54	72.47	5372	2.75	365	19	2.90	9.51	16.30
PGC0090616	2.49	-1.62	9919	1.35	371	43	1.10	9.74	17.07
PGC0019794	6.89	47.17	6228	5.08	373	16	5.90	9.92	15.64
PGC0019794	6.89	47.17	6227	5.05	375	17	5.70	9.91	15.64
PGC0166421	2.27	53.89	7081	2.07	384	50	0.80	9.64	16.61
PGC0090760	6.65	57.38	7988	3.73	399	18	8.20	10.00	15.97
PGC0091839	23.96	-3.60	7118	1.50	406	65	0.20	9.50	16.96
PGC0016272	4.90	42.31	6847	6.49	412	18	5.10	10.10	15.37
PGC0018709	6.25	71.14	4061	11.06	421	17	9.60	9.88	14.79
PGC0006177	1.67	81.16	8259	2.31	427	18	3.80	9.82	16.49
PGC0068000	22.08	44.06	9493	2.64	446	19	3.20	10.00	16.35
PGC0000002	0.00	47.27	5004	4.76	455	16	6.30	9.69	15.71
PGC0090748	6.17	50.78	7729	3.21	460	19	2.20	9.91	16.13
PGC0166743	22.22	50.50	8565	6.25	482	17	4.60	10.28	15.41
PGC0066791	21.46	45.48	5238	8.58	492	100	9.50	9.98	15.07
PGC0020088	7.07	50.68	6378	1.38	495	19	2.40	9.37	17.05
PGC0017326	5.47	49.96	6991	6.84	542	18	4.60	10.14	15.31
PGC0011447	3.03	36.10	7921	3.87	602	18	3.60	10.00	15.93
PGC0066281	21.19	37.88	8223	2.63	652	29	1.60	9.87	16.35

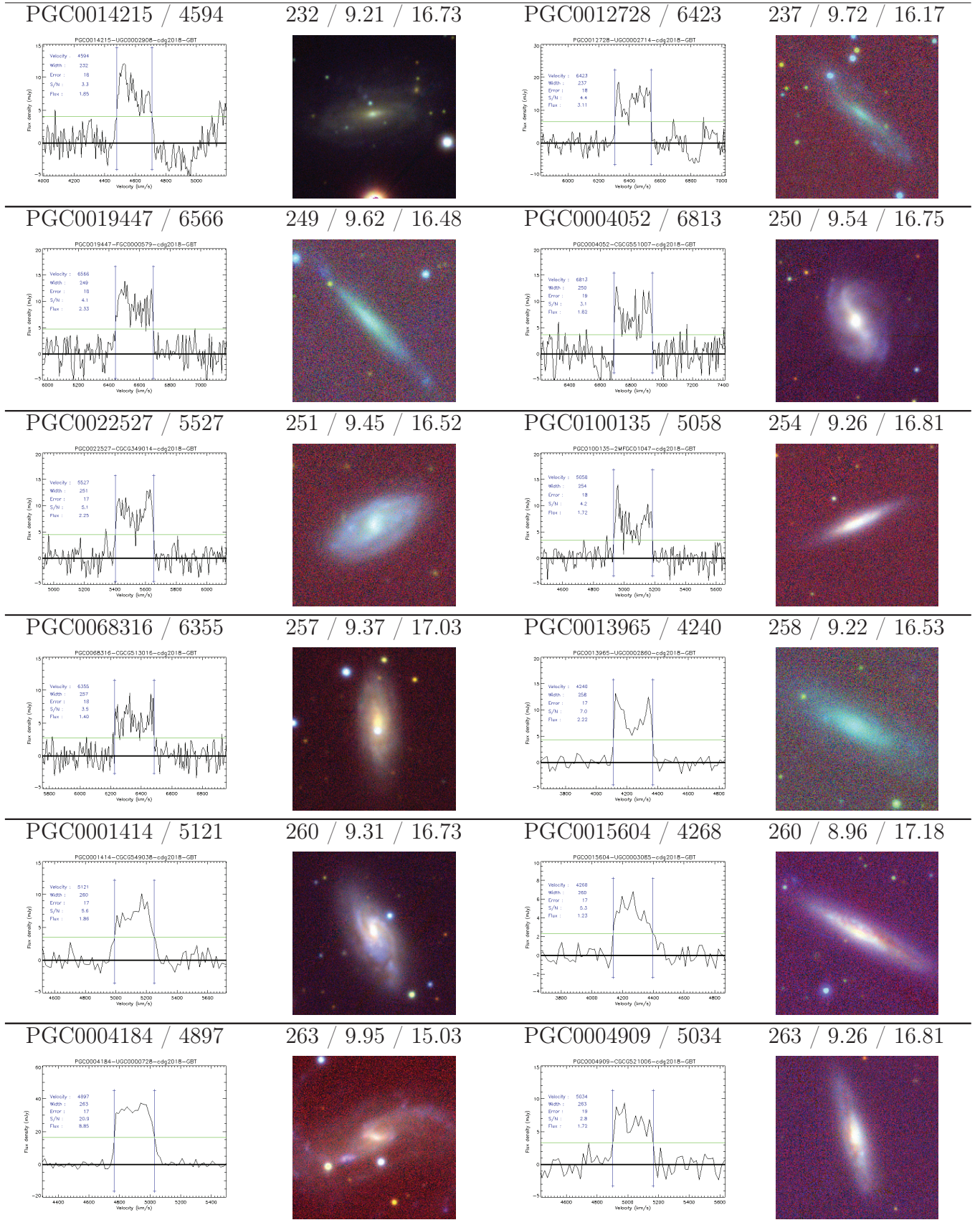
Annexe B

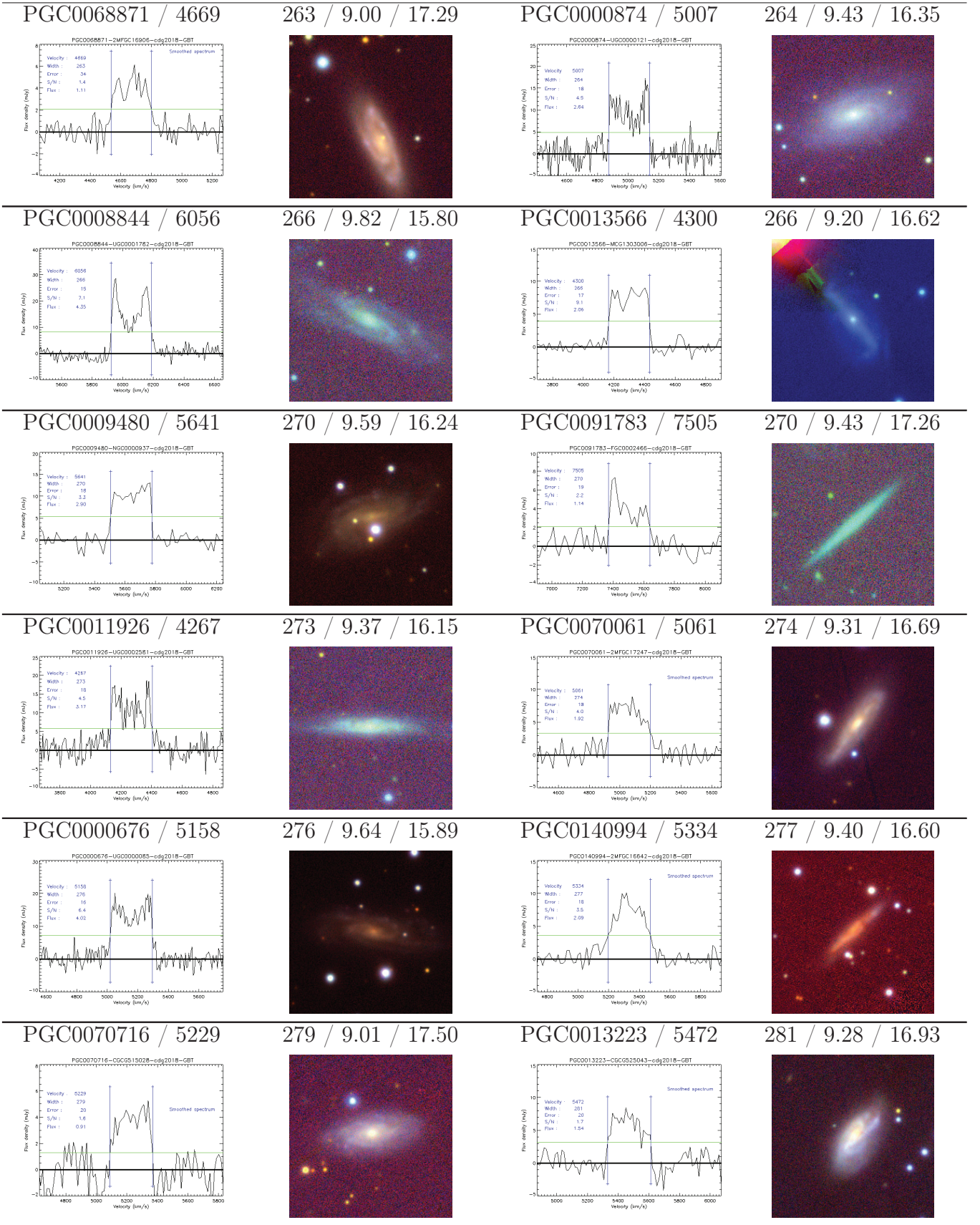
Observations au GBT : spectres HI adéquats pour Cosmicflows-4

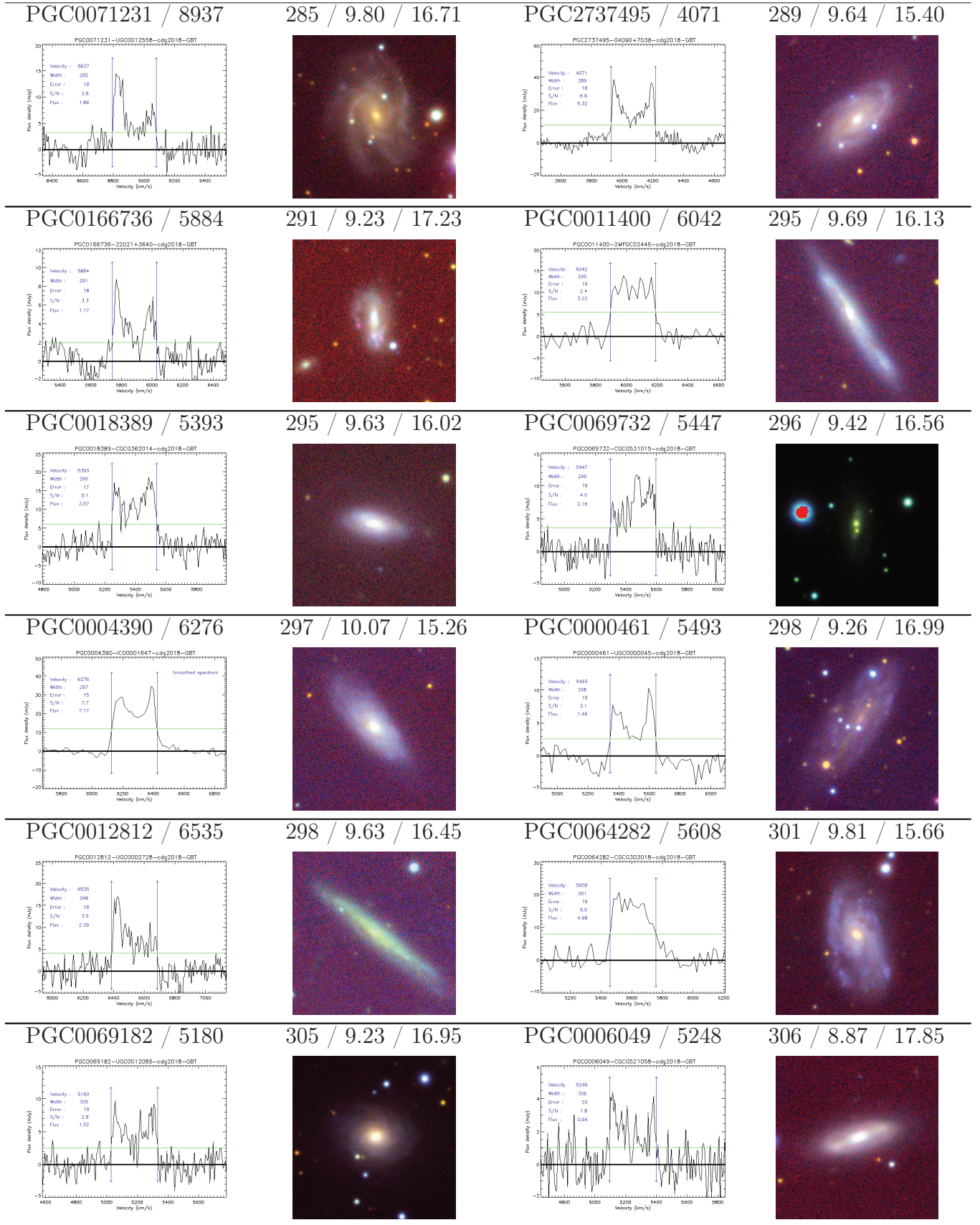
TABLE B.1 – Spectres HI après la mesure de W_{50} et images multi-bandes extraites de PanSTARRS pour les galaxies observées et détectées par la collaboration *Cosmicflows* au radiotélescope de Greenbank. Les galaxies sont triées par largeur de raie croissante, i.e de faible à grande vitesse de rotation. Les colonnes sont les suivantes : (1) et (3) numéro PGC, vitesse héliocentrique (en km/s), et spectre HI; (2) et (4) largeur de raie HI (en km/s), logarithme décimal de la masse HI (divisée par la masse solaire), magnitude à 21 cm, et image PanSTARRS de 1 arcmin de coté.

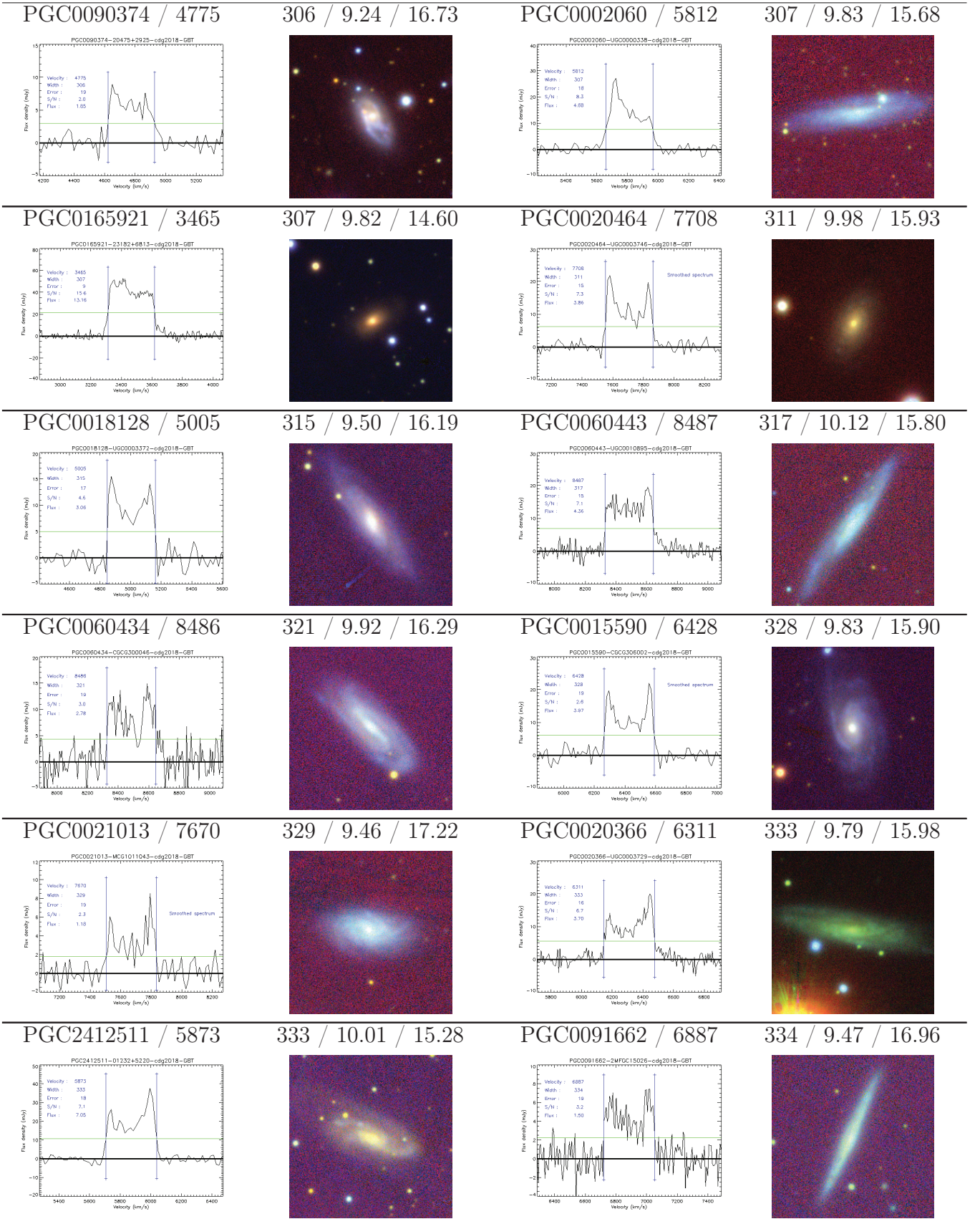
PGC / V_{hel} in km s ⁻¹ Profile	W_{50} / $\log_{10}(M_{HI})$ / m_{21} PanSTARRS	PGC / V_{hel} Profile	W_{50} / $\log M_{HI}$ / m_{21} PanSTARRS
PGC0068715 / 5473	61 / 8.86 / 17.97	PGC0008273 / 7616	109 / 9.78 / 16.40
			
PGC0141126 / 7170	120 / 9.74 / 16.37	PGC0020734 / 3251	133 / 8.61 / 17.48
			
PGC1177602 / 3082	143 / 8.85 / 16.77	PGC0020184 / 3257	155 / 8.70 / 17.24
			
PGC0020352 / 3158	168 / 8.96 / 16.55	PGC0021021 / 5196	187 / 9.79 / 15.54
			
PGC0012757 / 5107	190 / 9.42 / 16.44	PGC0091807 / 5308	191 / 9.35 / 16.68
			

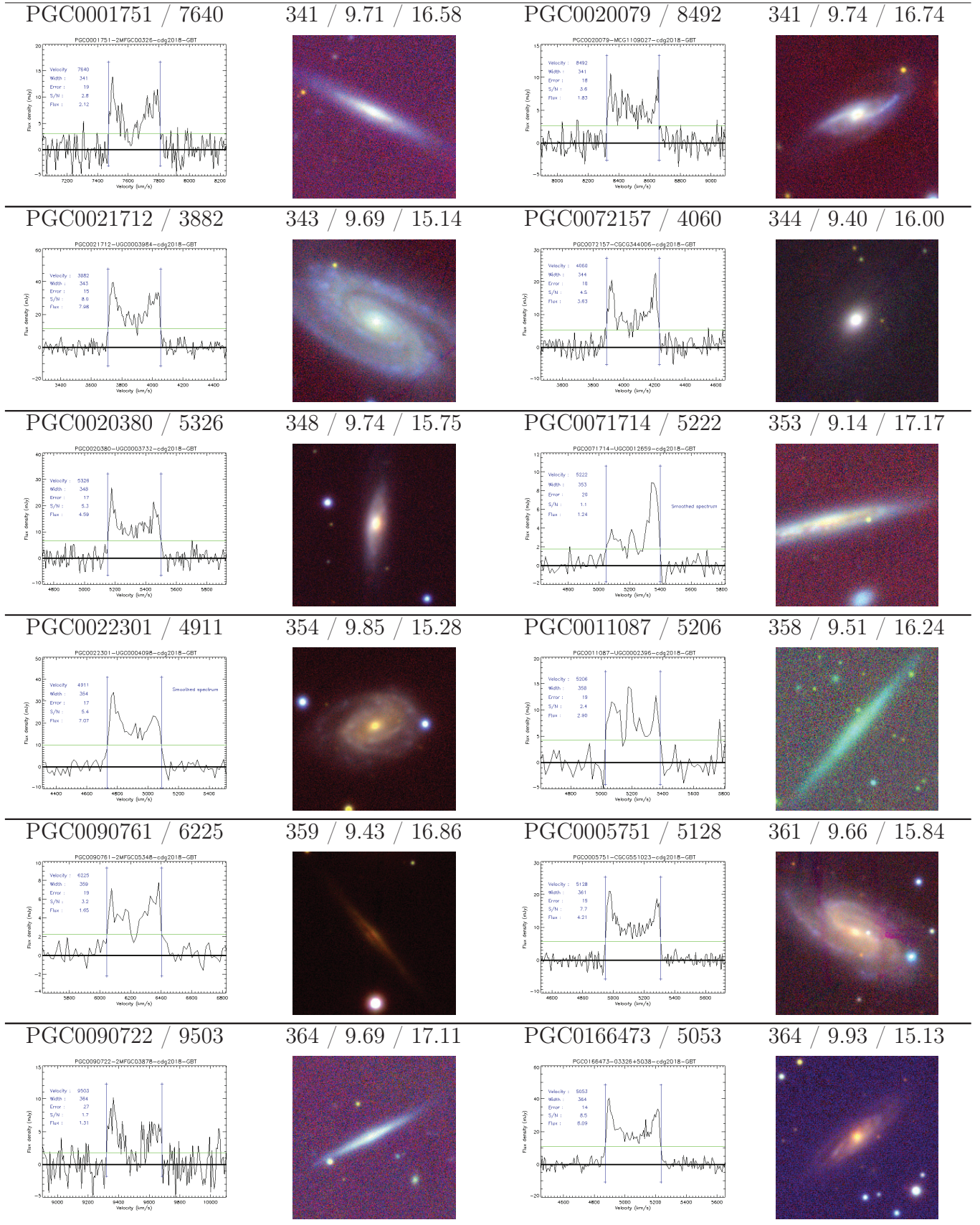


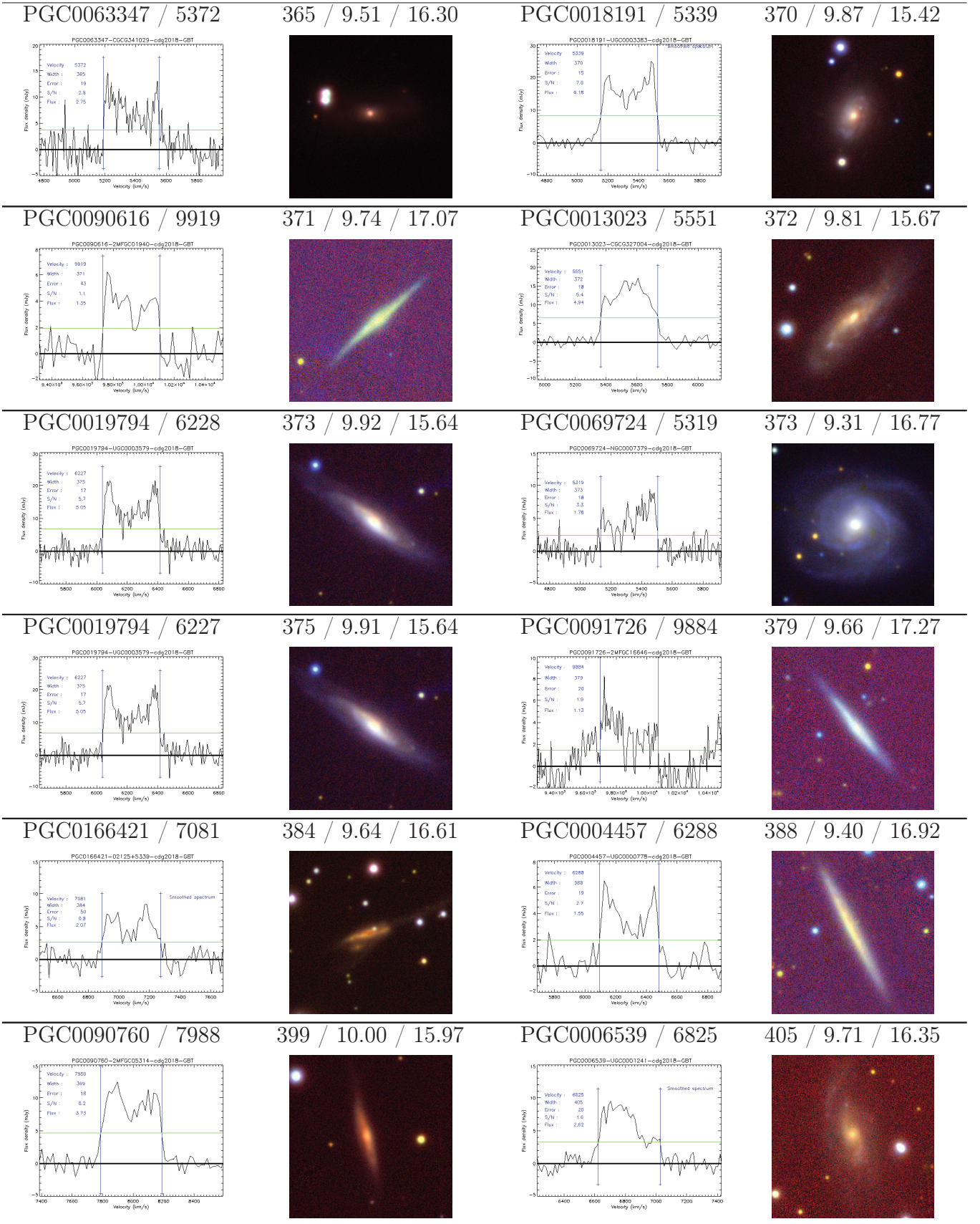


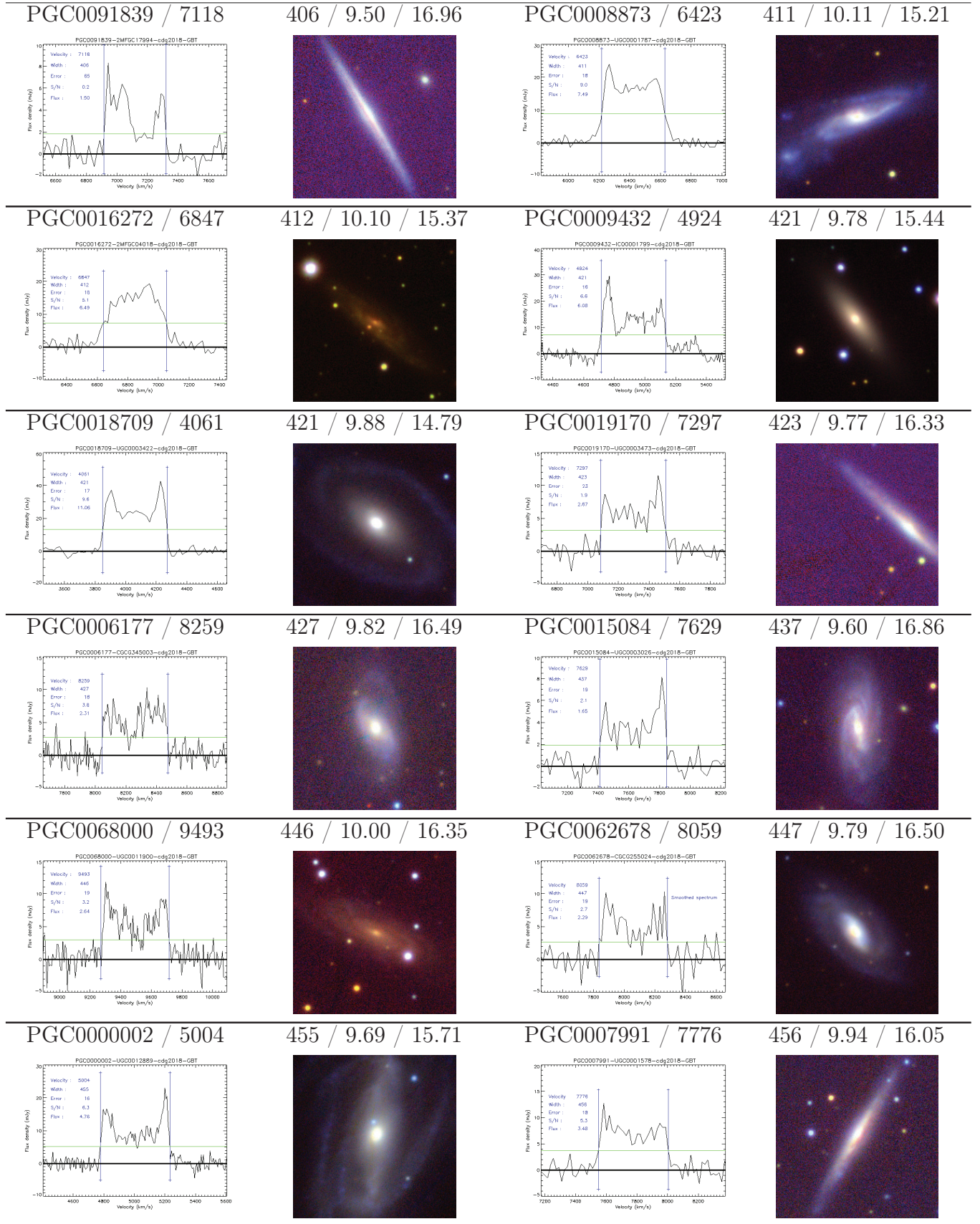


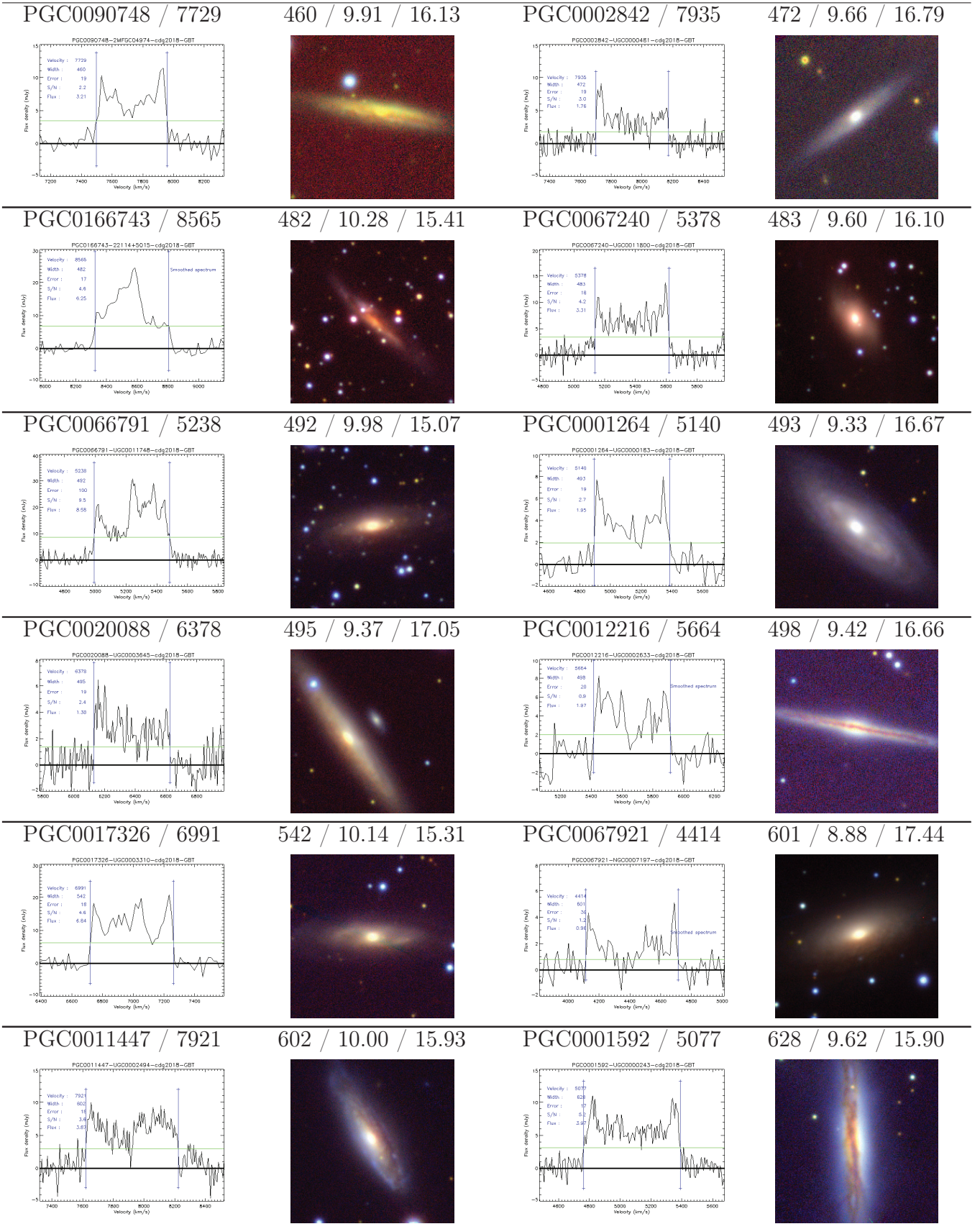






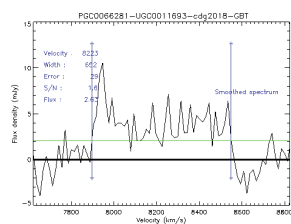






PGC0066281 / 8223

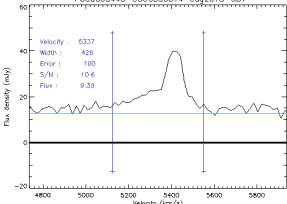

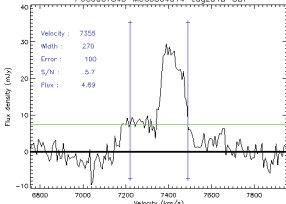
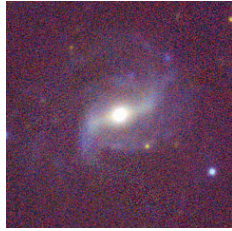
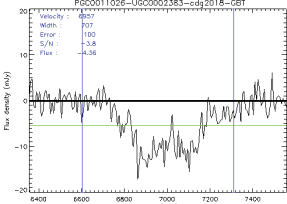

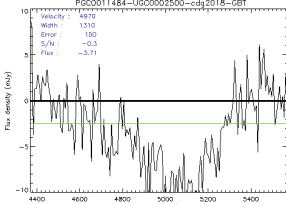
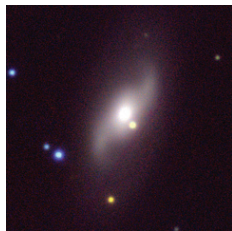
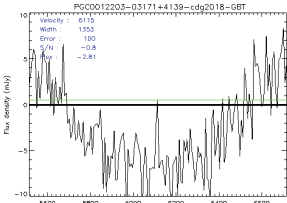
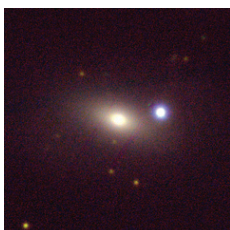
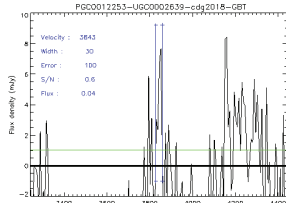
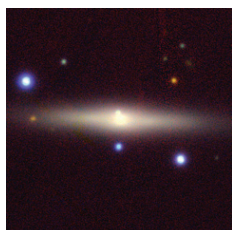
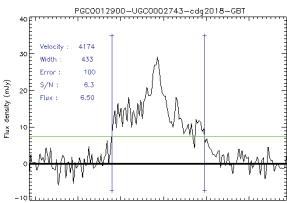
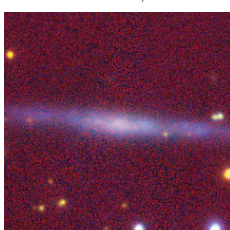
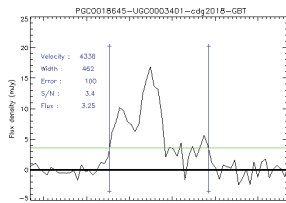
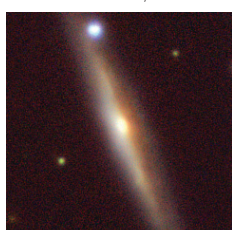
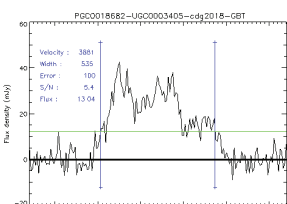
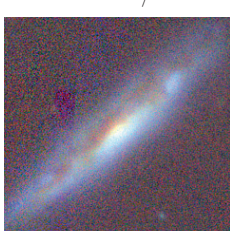
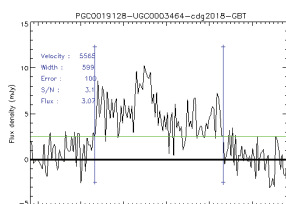
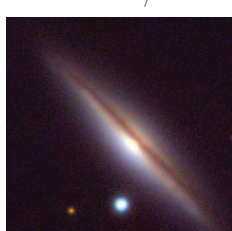
652 / 9.87 / 16.35

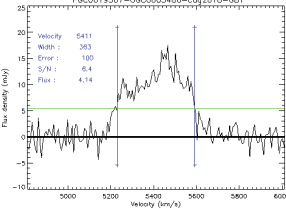
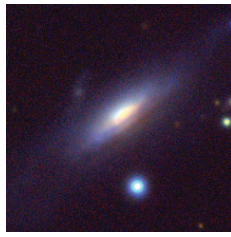
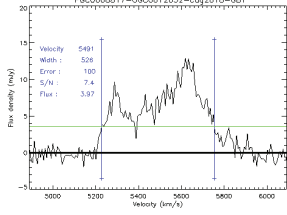

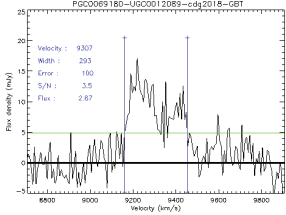

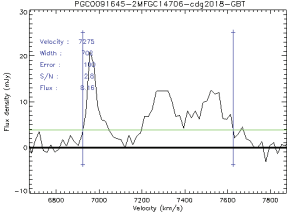

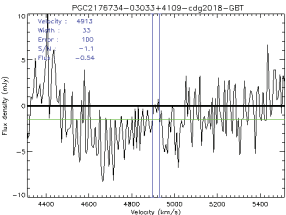
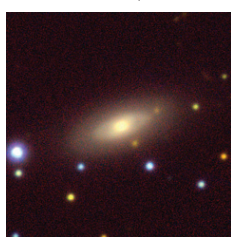
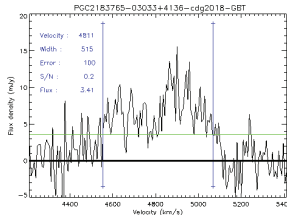
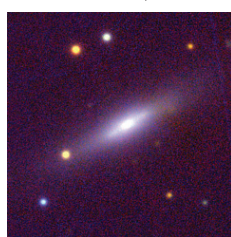
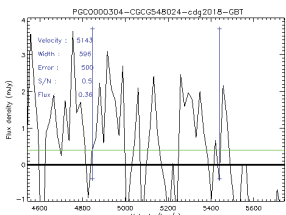

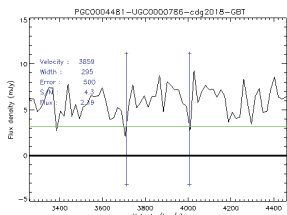

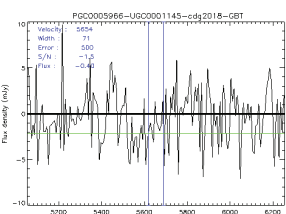

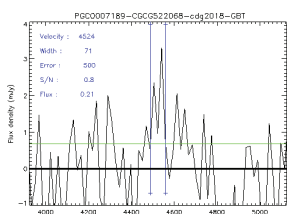

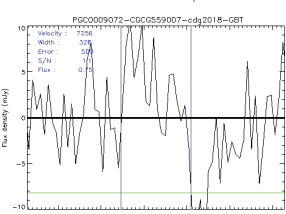

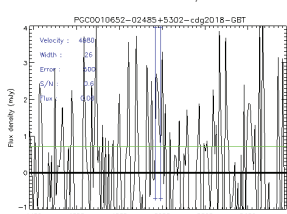



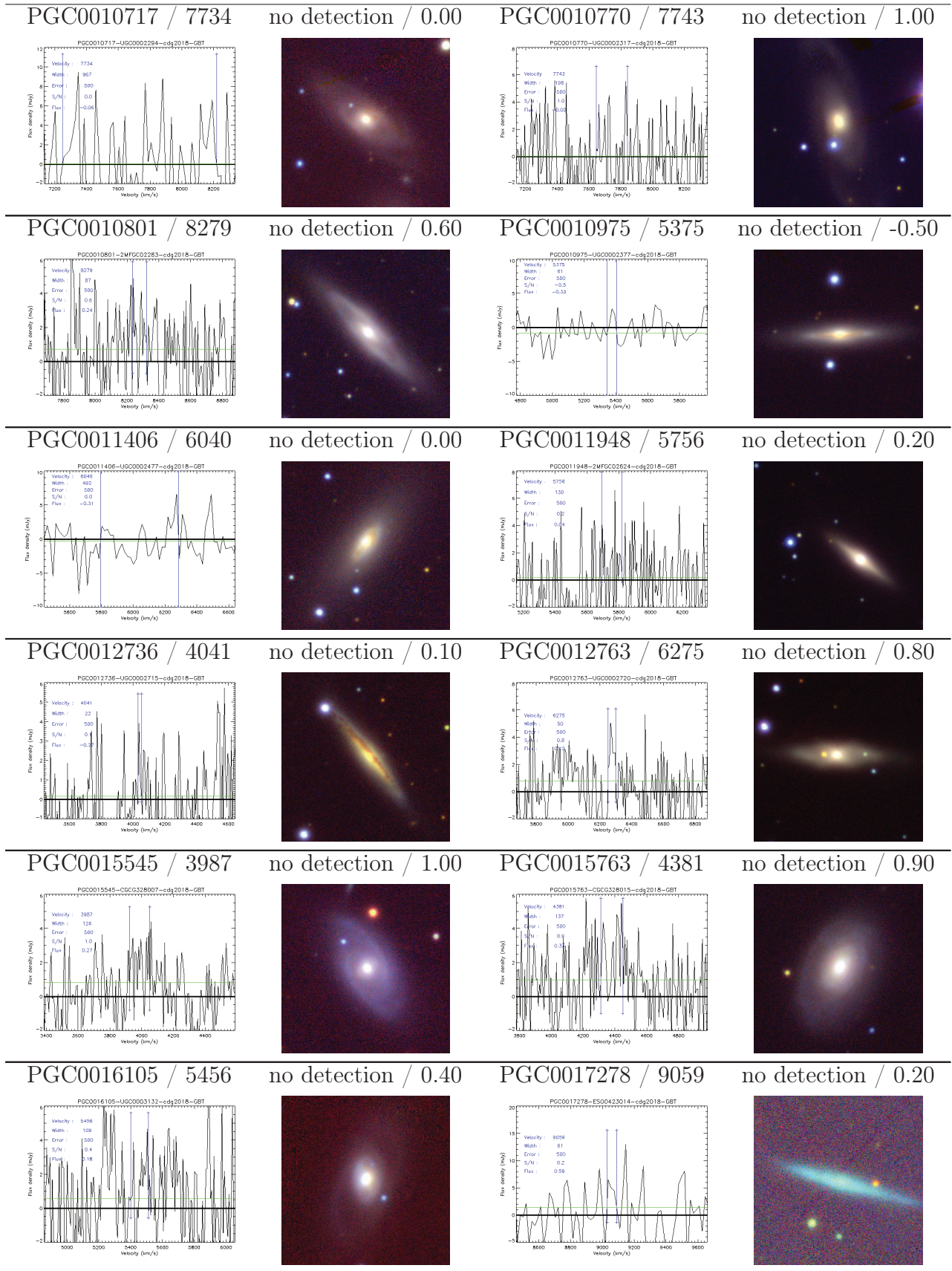
Annexe C

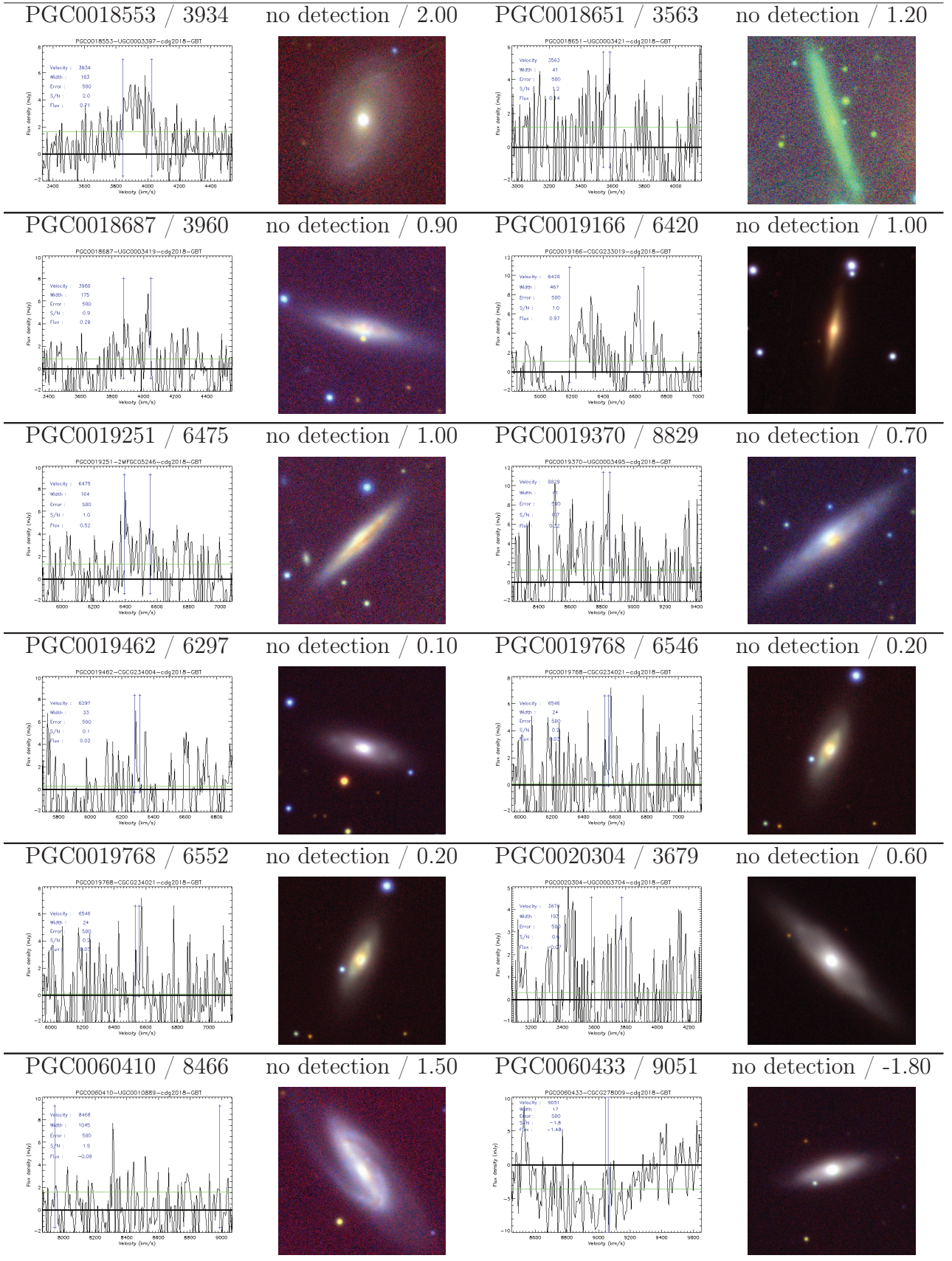
Observations au GBT : spectres HI des sources confuses ou non détectées

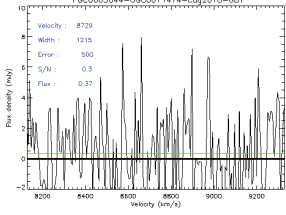
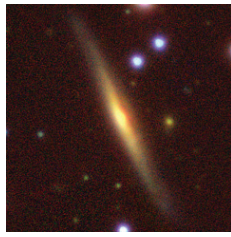
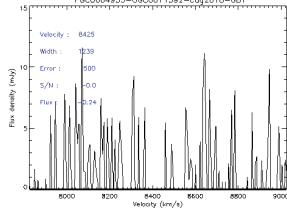

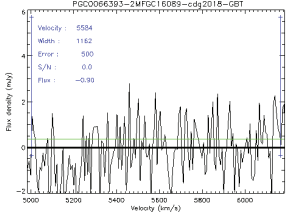

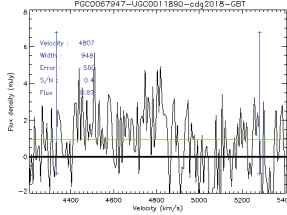
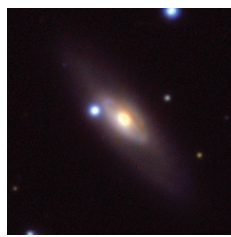
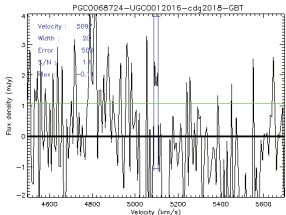

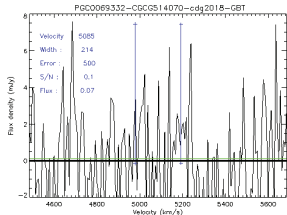
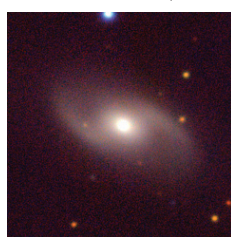
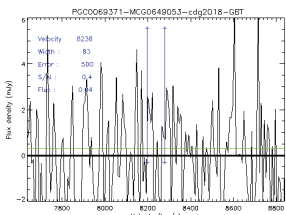
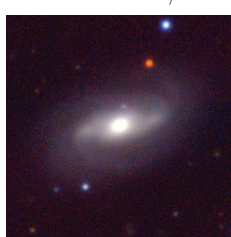
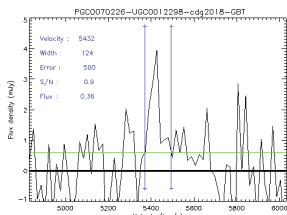
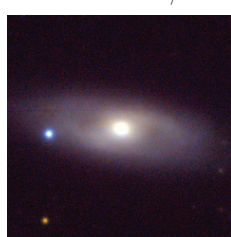
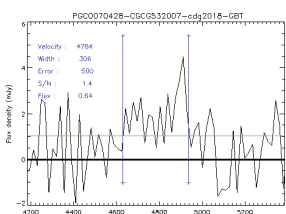

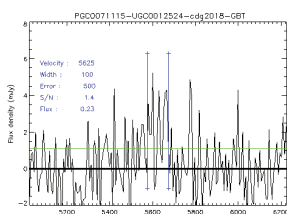
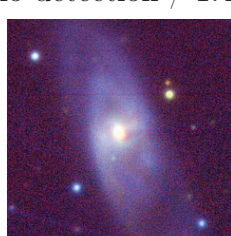
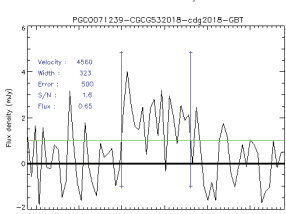
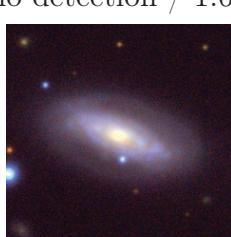
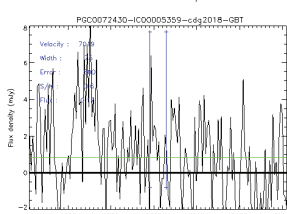

TABLE C.1 – Spectres HI après la mesure de W_{50} et images multi-bandes extraites de PanSTARRS pour les galaxies observées mais non détectées (ou cas de confusion) par la collaboration *Cosmicflows* au radiotélescope de Greenbank. Les galaxies sont triées par largeur de raie croissante, i.e de faible à grande vitesse de rotation. Les colonnes sont les suivantes : (1) et (3) numéro PGC, vitesse héliocentrique (en km/s), et spectre HI ; (2) et (4) largeur de raie HI (en km/s), logarithme décimal de la masse HI (divisée par la masse solaire), magnitude à 21 cm, et image PanSTARRS de 1 arcmin de coté.

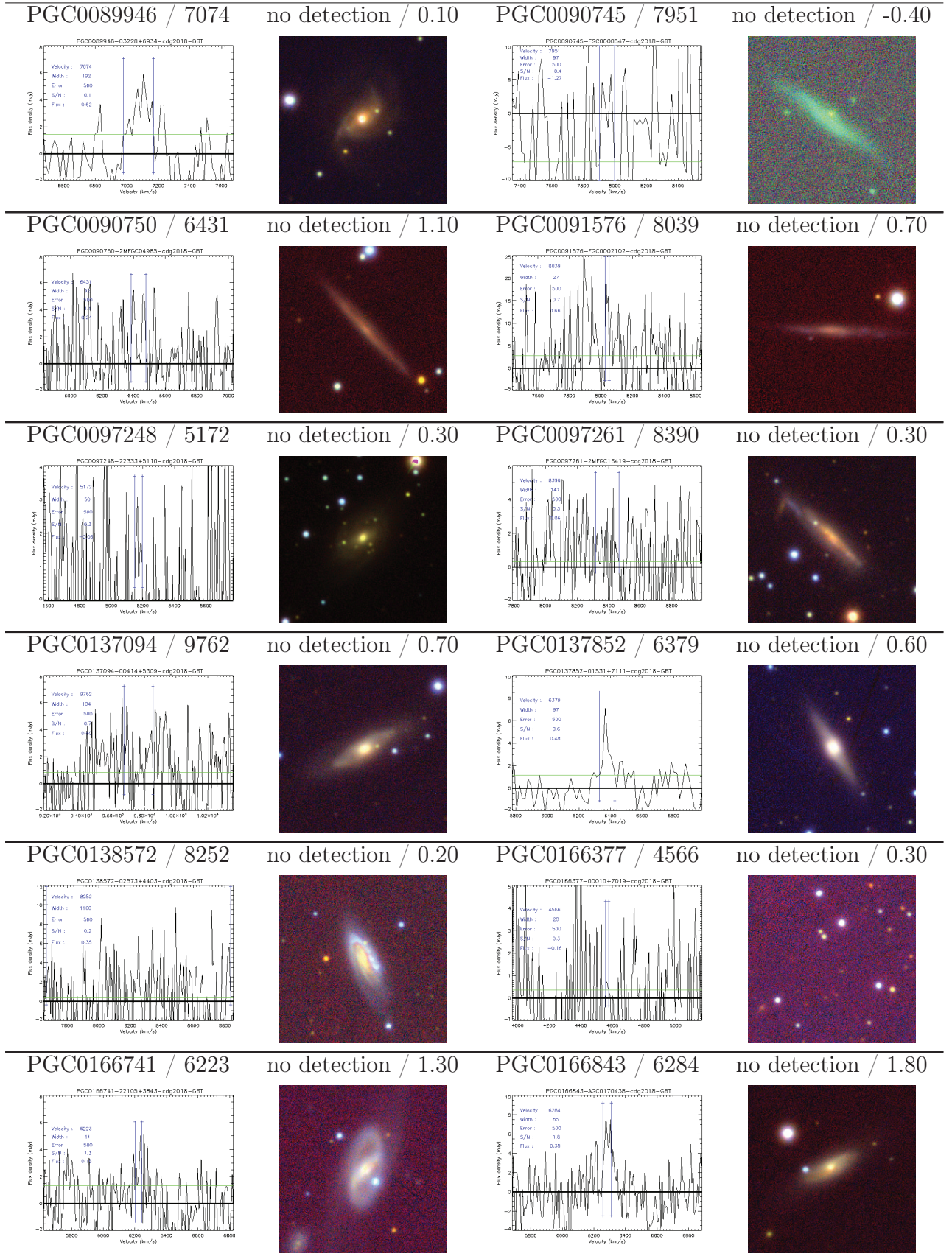
PGC / V_{hel} in km s ⁻¹ Profile	Comment / SNR PanSTARRS	PGC / V_{hel} Profile	Comment / SNR PanSTARRS
PGC0003448 / 5337	confused / 10.60	PGC0007845 / 7355	confused / 5.70
			
PGC0011026 / 6957	confused / -3.80	PGC0011484 / 4970	confused / -0.30
			
PGC0012203 / 6115	confused / -0.80	PGC0012253 / 3843	confused / 0.60
			
PGC0012900 / 4174	confused / 6.30	PGC0018645 / 4338	confused / 3.40
			
PGC0018682 / 3881	confused / 5.40	PGC0019128 / 5565	confused / 3.10
			

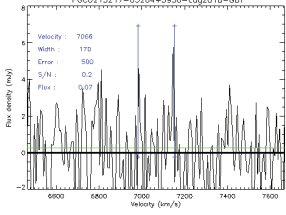
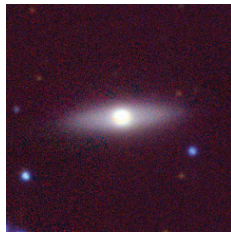
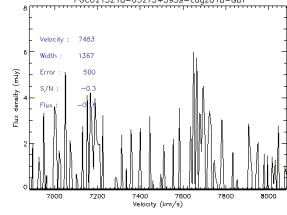
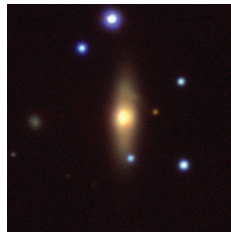
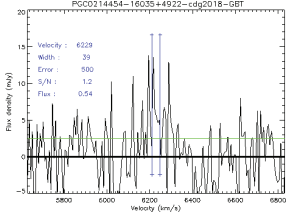

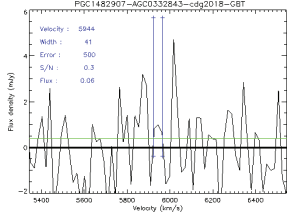

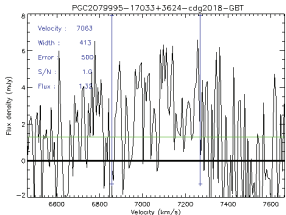
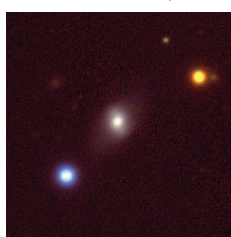
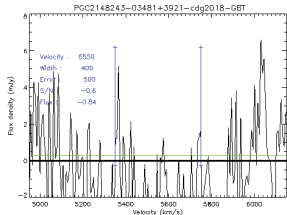
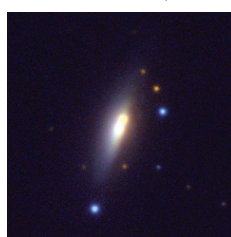
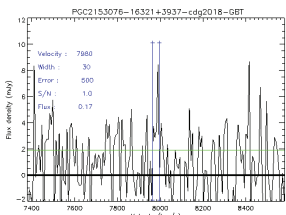

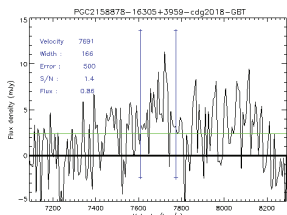

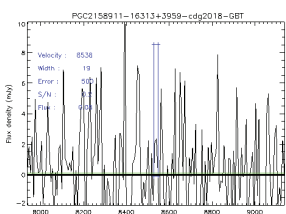

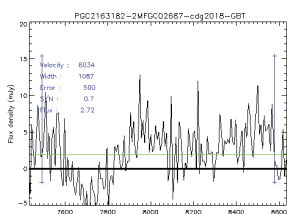

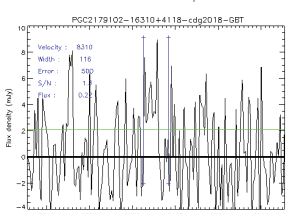

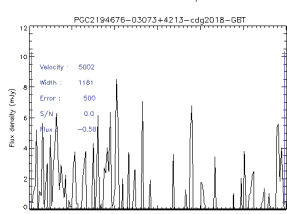
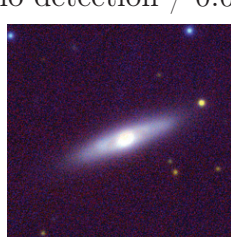
PGC0019307 / 5411	confused / 6.40	PGC0068817 / 5491	confused / 7.40
			
PGC0069180 / 9307	confused / 3.50	PGC0091645 / 7275	confused / 2.80
			
PGC2176734 / 4913	confused / -1.10	PGC2183765 / 4811	confused / 0.20
			
PGC0000304 / 5143	no detection / 0.50	PGC0004481 / 3859	no detection / 4.30
			
PGC0005966 / 5654	no detection / -1.50	PGC0007189 / 4524	no detection / 0.80
			
PGC0009072 / 7256	no detection / 1.10	PGC0010652 / 4980	no detection / 0.60
			





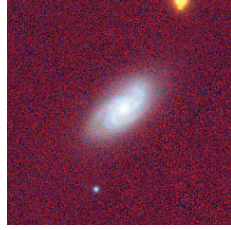
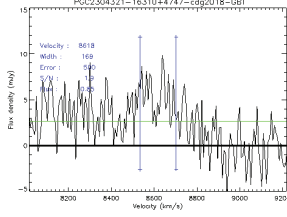
PGC0063644 / 8729	no detection / 0.30	PGC0064933 / 8425	no detection / -0.00
			
PGC0066393 / 5584	no detection / 0.00	PGC0067947 / 4807	no detection / 0.40
			
PGC0068724 / 5097	no detection / 1.10	PGC0069332 / 5085	no detection / 0.10
			
PGC0069371 / 8238	no detection / 0.40	PGC0070226 / 5432	no detection / 0.90
			
PGC0070428 / 4784	no detection / 1.40	PGC0071115 / 5625	no detection / 1.40
			
PGC0071239 / 4560	no detection / 1.60	PGC0072430 / 7019	no detection / 0.60
			



PGC0213217 / 7066	no detection / 0.20	PGC0213218 / 7483	no detection / -0.30
			
PGC0214454 / 6229	no detection / 1.20	PGC1482907 / 5944	no detection / 0.30
			
PGC2079995 / 7063	no detection / 1.00	PGC2148243 / 5550	no detection / -0.60
			
PGC2153076 / 7980	no detection / 1.00	PGC2158878 / 7691	no detection / 1.40
			
PGC2158911 / 8538	no detection / 0.20	PGC2163182 / 8034	no detection / 0.70
			
PGC2179102 / 8310	no detection / 1.30	PGC2194676 / 5002	no detection / 0.00
			

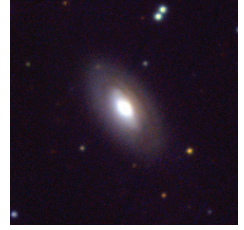
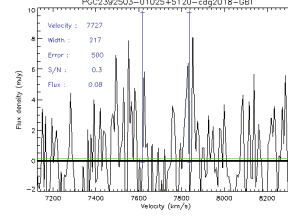
PGC2304321 / 8618

no detection / 1.90



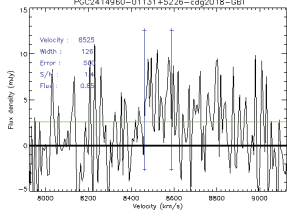
PGC2392503 / 7727

no detection / 0.30



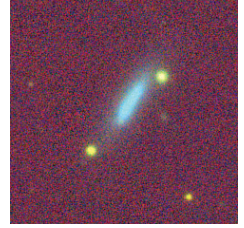
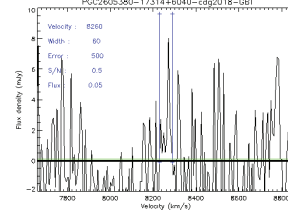
PGC2414960 / 8525

no detection / 1.40



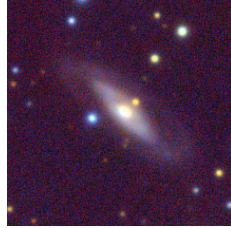
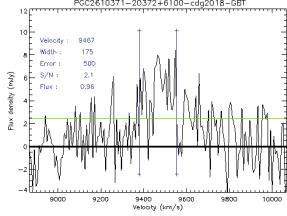
PGC2605380 / 8260

no detection / 0.50



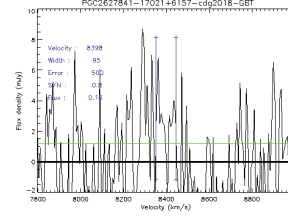
PGC2610371 / 9467

no detection / 2.10



PGC2627841 / 8398

no detection / 0.80



Bibliographie

- Abazajian, K., Adelman-McCarthy, J. K., Agüeros, M. A., Allam, S. S., Anderson, S. F., Annis, J., Bahcall, N. A., Baldry, I. K., Bastian, S., Berlind, A. et al. *The First Data Release of the Sloan Digital Sky Survey*. AJ, 126 :2081–2086, 2003.
- Alam, S., Albareti, F. D., Allende Prieto, C., Anders, F., Anderson, S. F., Anderton, T., Andrews, B. H., Armengaud, E., Aubourg, É., Bailey, S. et al. *The Eleventh and Twelfth Data Releases of the Sloan Digital Sky Survey : Final Data from SDSS-III*. ApJS, 219 :12, 2015.
- Amanullah, R., Lidman, C., Rubin, D., Aldering, G., Astier, P., Barbary, K., Burns, M. S., Conley, A., Dawson, K. S., Deustua, S. E. et al. *Spectra and Hubble Space Telescope Light Curves of Six Type Ia Supernovae at $0.511 < z < 1.12$ and the Union2 Compilation*. ApJ, 716 :712–738, 2010.
- Amendola, L., Appleby, S., Avgoustidis, A., Bacon, D., Baker, T., Baldi, M., Bartolo, N., Blanchard, A., Bonvin, C., Borgani, S. et al. *Cosmology and Fundamental Physics with the Euclid Satellite*. ArXiv e-prints, 2016.
- Auld, R., Minchin, R. F., Davies, J. I., Catinella, B., van Driel, W., Henning, P. A., Linder, S., Momjian, E., Muller, E., O’Neil, K. et al. *The Arecibo Galaxy Environment Survey : precursor observations of the NGC 628 group*. MNRAS, 371 :1617–1640, 2006.
- Barnes, D. G., Staveley-Smith, L., de Blok, W. J. G., Oosterloo, T., Stewart, I. M., Wright, A. E., Banks, G. D., Bhathal, R., Boyce, P. J., Calabretta, M. R. et al. *The HI Parkes All Sky Survey : southern observations, calibration and robust imaging*. MNRAS, 322 : 486–498, 2001.
- Beutler, F., Blake, C., Colless, M., Jones, D. H., Staveley-Smith, L., Poole, G. B., Campbell, L., Parker, Q., Saunders, W. and Watson, F. *The 6dF Galaxy Survey : $z \approx 0$ measurements of the growth rate and σ_8* . MNRAS, 423 :3430–3444, 2012.
- Blake, C., Kazin, E. A., Beutler, F., Davis, T. M., Parkinson, D., Brough, S., Colless, M., Contreras, C., Couch, W., Croom, S. et al. *The WiggleZ Dark Energy Survey : mapping the distance-redshift relation with baryon acoustic oscillations*. MNRAS, 418 : 1707–1724, 2011.
- Boomsma, R., Oosterloo, T. A., Fraternali, F., van der Hulst, J. M. and Sancisi, R. *HI holes and high-velocity clouds in the spiral galaxy NGC 6946*. A&A, 490 :555–570, 2008.
- Born, M. and Green, H. S. *A General Kinetic Theory of Liquids. I. The Molecular Distribution Functions*. Proceedings of the Royal Society of London Series A, 188 : 10–18, 1946.

- Bottinelli, L., Durand, N., Fouque, P., Garnier, R., Gouguenheim, L., Paturel, G. and Teerikorpi, P. *Observational data for the kinematics of the local universe. I - Radial velocity measurements.* A&AS, 93 :173–188, 1992.
- Bottinelli, L., Durand, N., Fouque, P., Garnier, R., Gouguenheim, L., Loulergue, M., Paturel, G., Petit, C. and Teerikorpi, P. *Observational Data for the Kinematics of the Local Universe - Part Two - Second Set of Radial Velocity Measurements.* A&AS, 102 : 57, 1993.
- Brownstein, J. R. and Moffat, J. W. *The Bullet Cluster 1E0657-558 evidence shows modified gravity in the absence of dark matter.* MNRAS, 382 :29–47, 2007.
- Butcher, Z., Schneider, S., van Driel, W., Lehnert, M. D. and Minchin, R. *NIBLES - an HI census of stellar mass selected SDSS galaxies. II. Arecibo follow-up HI observations.* A&A, 596 :A60, 2016.
- Catinella, B., Schiminovich, D., Kauffmann, G., Fabello, S., Wang, J., Hummels, C., Lemonias, J., Moran, S. M., Wu, R., Giovanelli, R. et al. *The GALEX Arecibo SDSS Survey - I. Gas fraction scaling relations of massive galaxies and first data release.* MNRAS, 403 :683–708, 2010.
- Catinella, B., Schiminovich, D., Kauffmann, G., Fabello, S., Hummels, C., Lemonias, J., Moran, S. M., Wu, R., Cooper, A. and Wang, J. *The GALEX Arecibo SDSS Survey. VI. Second data release and updated gas fraction scaling relations.* A&A, 544 :A65, 2012.
- Catinella, B., Schiminovich, D., Cortese, L., Fabello, S., Hummels, C. B., Moran, S. M., Lemonias, J. J., Cooper, A. P., Wu, R., Heckman, T. M. et al. *The GALEX Arecibo SDSS Survey - VIII. Final data release. The effect of group environment on the gas content of massive galaxies.* MNRAS, 436 :34–70, 2013.
- Cole, S., Percival, W. J., Peacock, J. A., Norberg, P., Baugh, C. M., Frenk, C. S., Baldry, I., Bland-Hawthorn, J., Bridges, T., Cannon, R. et al. *The 2dF Galaxy Redshift Survey : power-spectrum analysis of the final data set and cosmological implications.* MNRAS, 362 :505–534, 2005.
- Cortese, L., Minchin, R. F., Auld, R. R., Davies, J. I., Catinella, B., Momjian, E., Rosenberg, J. L., Taylor, R., Gavazzi, G., O’Neil, K. et al. *The Arecibo Galaxy Environment Survey - II. A HI view of the Abell cluster 1367 and its outskirts.* MNRAS, 383 : 1519–1537, 2008.
- Courtois, H. M. and Tully, R. B. *Cosmic Flows surveys and CLUES simulations.* Astronomische Nachrichten, 333 :436, 2012.
- Courtois, H. M., Tully, R. B., Fisher, J. R., Bonhomme, N., Zavodny, M. and Barnes, A. *The Extragalactic Distance Database : All Digital H I Profile Catalog.* AJ, 138 : 1938–1956, 2009.
- Courtois, H. M., Tully, R. B. and Héraudeau, P. *Cosmic flows : University of Hawaii 2.2-m I-band photometry.* MNRAS, 415 :1935–1942, 2011a.
- Courtois, H. M., Tully, R. B., Makarov, D. I., Mitronova, S., Koribalski, B., Karachentsev, I. D. and Fisher, J. R. *Cosmic Flows : Green Bank Telescope and Parkes H I observations.* MNRAS, 414 :2005–2016, 2011b.

- Courtois, H. M., Zaritsky, D., Sorce, J. G. and Pomarède, D. *Giant disc galaxies : where environment trumps mass in galaxy evolution*. MNRAS, 448 :1767–1778, 2015.
- Courtois, H. M., Tully, R. B., Hoffman, Y., Pomarède, D., Graziani, R. and Dupuy, A. *Cosmicflows-3 : Cold Spot Repeller ?* ApJ, 847 :L6, 2017.
- Croxall, K. V., van Zee, L., Lee, H., Skillman, E. D., Lee, J. C., Côté, S., Jr. Kennicutt, R. C. and Miller, B. W. *Chemical Abundances of Seven Irregular and Three Tidal Dwarf Galaxies in the M81 Group*. ApJ, 705 :723–738, 2009.
- Cui, X.-Q., Zhao, Y.-H., Chu, Y.-Q., Li, G.-P., Li, Q., Zhang, L.-P., Su, H.-J., Yao, Z.-Q., Wang, Y.-N., Xing, X.-Z. et al. *The Large Sky Area Multi-Object Fiber Spectroscopic Telescope (LAMOST)*. Research in Astronomy and Astrophysics, 12 :1197–1242, 2012.
- da Cunha, E., Hopkins, A. M., Colless, M., Taylor, E. N., Blake, C., Howlett, C., Magoulas, C., Lucey, J. R., Lagos, C., Kuehn, K. et al. *The Taipan Galaxy Survey : Scientific Goals and Observing Strategy*. Publ. Astron. Soc. Australia, 34 :e047, 2017.
- Davis, M. and Peebles, P. J. E. *On the integration of the BBGKY equations for the development of strongly nonlinear clustering in an expanding universe*. ApJS, 34 :425–450, 1977.
- Davis, T. M. and Scrimgeour, M. I. *Deriving accurate peculiar velocities (even at high redshift)*. MNRAS, 442 :1117–1122, 2014.
- de Fromont, P. and Alimi, J.-M. *Reconstructing matter profiles of spherically compensated cosmic regions in Λ CDM cosmology*. MNRAS, 473 :5177–5194, 2018.
- de la Torre, S., Guzzo, L., Peacock, J. A., Branchini, E., Iovino, A., Granett, B. R., Abbas, U., Adami, C., Arnouts, S., Bel, J. et al. *The VIMOS Public Extragalactic Redshift Survey (VIPERS) . Galaxy clustering and redshift-space distortions at $z \simeq 0.8$ in the first data release*. A&A, 557 :A54, 2013.
- di Nella, H., Paturel, G., Walsh, A. J., Bottinelli, L., Gouguenheim, L. and Theureau, G. *Kinematics of the local universe. III. Neutral hydrogen observations of southern galaxies*. A&AS, 118 :311–321, 1996.
- Diemer, B. and Kravtsov, A. V. *Dependence of the Outer Density Profiles of Halos on Their Mass Accretion Rate*. ApJ, 789 :1, 2014.
- Dupuy, A., Courtois, H. M. and Graziani, R. *Local growth rate from Cosmicflows peculiar velocities*. 2018.
- Einasto, J., Kaasik, A. and Saar, E. *Dynamic evidence on massive coronas of galaxies*. Nature, 250 :309–310, 1974.
- Eisenstein, D. J., Zehavi, I., Hogg, D. W., Scoccimarro, R., Blanton, M. R., Nichol, R. C., Scranton, R., Seo, H.-J., Tegmark, M., Zheng, Z. et al. *Detection of the Baryon Acoustic Peak in the Large-Scale Correlation Function of SDSS Luminous Red Galaxies*. ApJ, 633 :560–574, 2005.
- Fazio, G. G., Hora, J. L., Allen, L. E., Ashby, M. L. N., Barmby, P., Deutsch, L. K., Huang, J.-S., Kleiner, S., Marengo, M., Megeath, S. T. et al. *The Infrared Array Camera (IRAC) for the Spitzer Space Telescope*. ApJS, 154 :10–17, 2004.

- Feldman, H., Juszkievicz, R., Ferreira, P., Davis, M., Gaztañaga, E., Fry, J., Jaffe, A., Chambers, S., da Costa, L., Bernardi, M. et al. *An Estimate of Ω_m without Conventional Priors*. ApJ, 596 :L131–L134, 2003.
- Ferreira, P. G., Juszkievicz, R., Feldman, H. A., Davis, M. and Jaffe, A. H. *Streaming Velocities as a Dynamical Estimator of Ω* . ApJ, 515 :L1–L4, 1999.
- Freedman, W. L., Madore, B. F., Gibson, B. K., Ferrarese, L., Kelson, D. D., Sakai, S., Mould, J. R., Jr. Kennicutt, R. C., Ford, H. C., Graham, J. A. et al. *Final Results from the Hubble Space Telescope Key Project to Measure the Hubble Constant*. ApJ, 553 : 47–72, 2001.
- Gelfand, Alan E. and Smith, Adrian F. M. *Sampling-Based Approaches to Calculating Marginal Densities*. Journal of the American Statistical Association, 85(410) :398–409, 1990.
- Geman, Stuart and Geman, Donald. *Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images*. IEEE Trans. Pattern Anal. Mach. Intell., 6(6) :721–741, 1984.
- Giovanelli, R., Haynes, M. P., Herter, T., Vogt, N. P., Wegner, G., Salzer, J. J., da Costa, L. N. and Freudling, W. *The I band Tully-Fisher relation for cluster galaxies : data presentation*. AJ, 113 :22–52, 1997.
- Giovanelli, R., Haynes, M. P., Kent, B. R., Perillat, P., Saintonge, A., Brosch, N., Catinella, B., Hoffman, G. L., Stierwalt, S., Spekkens, K. et al. *The Arecibo Legacy Fast ALFA Survey. I. Science Goals, Survey Design, and Strategy*. AJ, 130 :2598–2612, 2005.
- Gorski, K. *On the pattern of perturbations of the Hubble flow*. ApJ, 332 :L7–L11, 1988.
- Gorski, K. M., Davis, M., Strauss, M. A., White, S. D. M. and Yahil, A. *Cosmological velocity correlations - Observations and model predictions*. ApJ, 344 :1–19, 1989.
- Gottloeber, S., Hoffman, Y. and Yepes, G. *Constrained Local Universe Simulations (CLUES)*. ArXiv e-prints, 2010.
- Graziani, R., Courtois, H. M., Lavaux, G., Hoffman, Y., Tully, R. B. and Copin, Y. *The peculiar velocity field up to $z \approx 0.054$ from Cosmicflows-3*. 2018.
- Hall, M., Courteau, S., Dutton, A. A., McDonald, M. and Zhu, Y. *An investigation of Sloan Digital Sky Survey imaging data and multiband scaling relations of spiral galaxies*. MNRAS, 425 :2741–2765, 2012.
- Hastings, W. K. *Monte Carlo Sampling Methods Using Markov Chains and Their Applications*. Biometrika, 57(1) :97–109, 1970.
- Hawkins, E., Maddox, S., Cole, S., Lahav, O., Madgwick, D. S., Norberg, P., Peacock, J. A., Baldry, I. K., Baugh, C. M., Bland-Hawthorn, J. et al. *The 2dF Galaxy Redshift Survey : correlation functions, peculiar velocities and the matter density of the Universe*. MNRAS, 346 :78–96, 2003.
- Haynes, M. P., Giovanelli, R., Martin, A. M., Hess, K. M., Saintonge, A., Adams, E. A. K., Hallenbeck, G., Hoffman, G. L., Huang, S., Kent, B. R. et al. *The Arecibo Legacy Fast ALFA Survey : The $\alpha.40$ H I Source Catalog, Its Characteristics and Their Impact on the Derivation of the H I Mass Function*. AJ, 142 :170, 2011.

- Hoffman, Y. and Ribak, E. *Constrained realizations of Gaussian fields - A simple algorithm*. ApJ, 380 :L5–L8, 1991.
- Hoffman, Y. and Ribak, E. *Primordial Gaussian perturbation fields - Constrained realizations*. ApJ, 384 :448–452, 1992.
- Hoffman, Y., Courtois, H. M. and Tully, R. B. *Cosmic bulk flow and the local motion from Cosmicflows-2*. MNRAS, 449 :4494–4505, 2015.
- Hoffman, Y., Pomarède, D., Tully, R. B. and Courtois, H. M. *The dipole repeller*. Nature Astronomy, 1 :0036, 2017.
- Holmberg, E. *A photographic photometry of extragalactic nebulae*. Meddelanden fran Lunds Astronomiska Observatorium Serie II, 136 :1, 1958.
- Hong, T., Staveley-Smith, L., Masters, K. L., Springob, C. M., Macri, L. M., Koribalski, B. S., Jones, D. H., Jarrett, T. H. and Crook, A. C. *2MTF - II. New Parkes 21-cm observations of 303 southern galaxies*. MNRAS, 432 :1178–1188, 2013.
- Hubble, E. *A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae*. Proceedings of the National Academy of Science, 15 :168–173, 1929.
- Hubble, E. and Humason, M. L. *The Velocity-Distance Relation among Extra-Galactic Nebulae*. ApJ, 74 :43, 1931.
- Huchra, J. P., de Lapparent, V., Geller, M. J., Kurtz, M. J., Horine, E., Peters, J. and Tokarz, S. *A Slice of the Universe*. In *Bulletin of the American Astronomical Society*, volume 17 of BAAS, page 857, 1985.
- Huchra, J. P., Macri, L. M., Masters, K. L., Jarrett, T. H., Berlind, P., Calkins, M., Crook, A. C., Cutri, R., Erdoğan, P., Falco, E. et al. *The 2MASS Redshift Survey — Description and Data Release*. ApJS, 199 :26, 2012.
- Hunt, L. R., Pisano, D. J. and Edel, S. *The Search for HI Emission at $z \approx 0.4$ in Gravitationally Lensed Galaxies with the Green Bank Telescope*. AJ, 152 :30, 2016.
- iminuit team, . *iminuit – A Python interface to Minuit*. <https://github.com/iminuit/iminuit>. Accessed : 2018-03-05.
- Jackson, J. C. *A critique of Rees’s theory of primordial gravitational radiation*. MNRAS, 156 :1P, 1972.
- Jacobs, B. A., Rizzi, L., Tully, R. B., Shaya, E. J., Makarov, D. I. and Makarova, L. *The Extragalactic Distance Database : Color-Magnitude Diagrams*. AJ, 138 :332–337, 2009.
- James, F. and Roos, M. *Minuit – a system for function minimization and analysis of the parameter errors and correlations*. Computer Physics Communications, 10 :343–367, 1975.
- Jarrett, T.-H., Chester, T., Cutri, R., Schneider, S., Rosenberg, J., Huchra, J. P. and Mader, J. *2MASS Extended Sources in the Zone of Avoidance*. AJ, 120 :298–313, 2000a.
- Jarrett, T. H., Chester, T., Cutri, R., Schneider, S., Skrutskie, M. and Huchra, J. P. *2MASS Extended Source Catalog : Overview and Algorithms*. AJ, 119 :2498–2531, 2000b.

- Jobard, Bruno, . *Visualisation de champs de vecteurs bidimensionnels à base de stream-lines*. PhD thesis, Université du Littoral-Côte d'Opale, 2000.
- Johnson, A., Blake, C., Koda, J., Ma, Y.-Z., Colless, M., Crocce, M., Davis, T. M., Jones, H., Magoulas, C., Lucey, J. R. et al. *The 6dF Galaxy Survey : cosmological constraints from the velocity power spectrum*. MNRAS, 444 :3926–3947, 2014.
- Jones, D. H., Read, M. A., Saunders, W., Colless, M., Jarrett, T., Parker, Q. A., Fairall, A. P., Mauch, T., Sadler, E. M., Watson, F. G. et al. *The 6dF Galaxy Survey : final redshift release (DR3) and southern large-scale structures*. MNRAS, 399 :683–698, 2009.
- Juszkiewicz, R., Fisher, K. B. and Szapudi, I. *Skewed Exponential Pairwise Velocities from Gaussian Initial Conditions*. ApJ, 504 :L1–L4, 1998.
- Juszkiewicz, R., Springel, V. and Durrer, R. *Dynamics of Pairwise Motions*. ApJ, 518 : L25–L28, 1999.
- Juszkiewicz, R., Ferreira, P. G., Feldman, H. A., Jaffe, A. H. and Davis, M. *Evidence for a Low-Density Universe from the Relative Velocities of Galaxies*. Science, 287 :109–112, 2000.
- Kaiser, N. *Clustering in real space and in redshift space*. MNRAS, 227 :1–21, 1987.
- Kaiser, N., Aussel, H., Burke, B. E., Boesgaard, H., Chambers, K., Chun, M. R., Heasley, J. N., Hodapp, K.-W., Hunt, B., Jedicke, R. et al. *Pan-STARRS : A Large Synoptic Survey Telescope Array*. In Tyson, J. A. and Wolff, S., editors, *Survey and Other Telescope Technologies and Discoveries*, volume 4836 of *Proc. SPIE*, pages 154–164, 2002. doi: 10.1117/12.457365.
- Karachentsev, I. D., Karachentseva, V. E., Kudrya, Y. N., Sharina, M. E. and Parnovskij, S. L. *The revised Flat Galaxy Catalogue*. Bulletin of the Special Astrophysics Observatory, 47, 1999.
- Karachentsev, I. D., Mitronova, S. N., Karachentseva, V. E., Kudrya, Y. N. and Jarrett, T. H. *The 2MASS Tully-Fisher relation for flat edge-on galaxies*. A&A, 396 :431–438, 2002.
- Karachentsev, I. D., Karachentseva, V. E., Huchtmeier, W. K. and Makarov, D. I. *A Catalog of Neighboring Galaxies*. AJ, 127 :2031–2068, 2004.
- Karachentsev, I. D., Dolphin, A., Tully, R. B., Sharina, M., Makarova, L., Makarov, D., Karachentseva, V., Sakai, S. and Shaya, E. J. *Advanced Camera for Surveys Imaging of 25 Galaxies in Nearby Groups and in the Field*. AJ, 131 :1361–1376, 2006.
- Kirkwood, J. G. *The Statistical Mechanical Theory of Transport Processes I. General Theory*. J. Chem. Phys., 14 :180–201, 1946.
- Kirkwood, J. G. *The Statistical Mechanical Theory of Transport Processes II. Transport in Gases*. J. Chem. Phys., 15 :72–76, 1947.
- Kraan-Korteweg, R. C., Cluver, M. E., Bilicki, M., Jarrett, T. H., Colless, M., Elagali, A., Böhringer, H. and Chon, G. *Discovery of a supercluster in the Zone of Avoidance in Vela*. MNRAS, 466 :L29–L33, 2017.

- Lahav, O., Lilje, P. B., Primack, J. R. and Rees, M. J. *Dynamical effects of the cosmological constant*. MNRAS, 251 :128–136, 1991.
- Lavaux, G. *Precision constrained simulation of the local Universe*. MNRAS, 406 :1007–1013, 2010.
- Lavaux, G. *Bayesian 3D velocity field reconstruction with VIRBIUS*. MNRAS, 457 :172–197, 2016.
- Leclercq, F., Jasche, J., Lavaux, G., Wandelt, B. and Percival, W. *The phase-space structure of nearby dark matter as constrained by the SDSS*. J. Cosmology Astropart. Phys., 6 :049, 2017.
- Lewis, A. and Bridle, S. *Cosmological parameters from CMB and other data : A Monte Carlo approach*. Phys. Rev. D, 66(10) :103511, 2002.
- Lewis, A. and Challinor, A. *CAMB : Code for Anisotropies in the Microwave Background*. Astrophysics Source Code Library, 2011.
- Libeskind, N. I., Hoffman, Y., Tully, R. B., Courtois, H. M., Pomarède, D., Gottlöber, S. and Steinmetz, M. *Planes of satellite galaxies and the cosmic web*. MNRAS, 452 :1052–1059, 2015.
- Linder, E. V. *Cosmic growth history and expansion history*. Phys. Rev. D, 72(4) :043529, 2005.
- Liu, F., Cutri, R., Greanias, G., Duval, V., Eisenhardt, P., Elwell, J., Heinrichsen, I., Howard, J., Irace, W., Mainzer, A. et al. *Development of the Wide-field Infrared Survey Explorer (WISE) mission*. In *Modeling, Systems Engineering, and Project Management for Astronomy III*, volume 7017 of *Proc. SPIE*, page 70170M, 2008. doi: 10.1117/12.790087.
- Lynden-Bell, D., Faber, S. M., Burstein, D., Davies, R. L., Dressler, A., Terlevich, R. J. and Wegner, G. *Spectroscopy and photometry of elliptical galaxies. V - Galaxy streaming toward the new supergalactic center*. ApJ, 326 :19–49, 1988.
- Ma, Y.-Z., Li, M. and He, P. *Constraining cosmology with pairwise velocity estimator*. A&A, 583 :A52, 2015.
- Makarov, D., Prugniel, P., Terekhova, N., Courtois, H. and Vauglin, I. *HyperLEDA. III. The catalogue of extragalactic distances*. A&A, 570 :A13, 2014.
- Masters, K. L., Huchra, J. P., Crook, A. C., Macri, L. M., Jarrett, T., Koribalski, B., Staveley-Smith, L., Jones, H. and Springob, C. M. *2MTF - The 2MASS Tully-Fisher Survey*. In *American Astronomical Society Meeting Abstracts #210*, volume 39 of *Bulletin of the American Astronomical Society*, page 193, 2007.
- Masters, K. L., Crook, A., Hong, T., Jarrett, T. H., Koribalski, B. S., Macri, L., Springob, C. M. and Staveley-Smith, L. *2MTF III. H I 21 cm observations of 1194 spiral galaxies with the Green Bank Telescope*. MNRAS, 443 :1044–1056, 2014.
- McConnachie, A., Babusiaux, C., Balogh, M., Driver, S., Côté, P., Courtois, H., Davies, L., Ferrarese, L., Gallagher, S., Ibata, R. et al. *The Detailed Science Case for the Maunakea Spectroscopic Explorer : the Composition and Dynamics of the Faint Universe*. ArXiv e-prints, 2016.

- Mei, S., Blakeslee, J. P., Côté, P., Tonry, J. L., West, M. J., Ferrarese, L., Jordán, A., Peng, E. W., Anthony, A. and Merritt, D. *The ACS Virgo Cluster Survey. XIII. SBF Distance Catalog and the Three-dimensional Structure of the Virgo Cluster*. ApJ, 655 : 144–162, 2007.
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. and Teller, E. *Equation of State Calculations by Fast Computing Machines*. J. Chem. Phys., 21 :1087–1092, 1953.
- Meyer, M. J., Zwaan, M. A., Webster, R. L., Staveley-Smith, L., Ryan-Weber, E., Drinkwater, M. J., Barnes, D. G., Howlett, M., Kilborn, V. A., Stevens, J. et al. *The HIPASS catalogue - I. Data presentation*. MNRAS, 350 :1195–1209, 2004.
- Minchin, R. F., Momjian, E., Auld, R., Davies, J. I., Valls-Gabaud, D., Karachentsev, I. D., Henning, P. A., O’Neil, K. L., Schneider, S., Smith, M. W. L. et al. *The Arecibo Galaxy Environment Survey. III. Observations Toward the Galaxy Pair NGC 7332/7339 and the Isolated Galaxy NGC 1156*. AJ, 140 :1093–1118, 2010.
- Monin, A. S. and Yaglom, A. M. *Statistical fluid mechanics*. Cambridge : MIT Press, 1975.
- N. N. Bogoliubov, . *Kinetic Equations*. Journal of Experimental and Theoretical Physics, 16 :691–702, 1946.
- Nusser, A. *Velocity-density correlations from the cosmicflows-3 distance catalogue and the 2MASS Redshift Survey*. MNRAS, 470 :445–454, 2017.
- Paturel, G., Petit, C., Prugniel, P., Theureau, G., Rousseau, J., Brouty, M., Dubois, P. and Cambrésy, L. *HYPERLEDA. I. Identification and designation of galaxies*. A&A, 412 :45–55, 2003a.
- Paturel, G., Theureau, G., Bottinelli, L., Gouguenheim, L., Coudreau-Durand, N., Hallet, N. and Petit, C. *HYPERLEDA. II. The homogenized HI data*. A&A, 412 :57–67, 2003b.
- Paturel, G., Vauglin, I., Petit, C., Borsenberger, J., Epchtein, N., Fouqué, P. and Mamon, G. *A catalog of LEDA galaxies with DENIS measurements*. A&A, 430 :751–759, 2005.
- Peacock, J. A. *Cosmological Physics*. 1999.
- Peacock, J. A., Cole, S., Norberg, P., Baugh, C. M., Bland-Hawthorn, J., Bridges, T., Cannon, R. D., Colless, M., Collins, C., Couch, W. et al. *A measurement of the cosmological mass density from clustering in the 2dF Galaxy Redshift Survey*. Nature, 410 : 169–173, 2001.
- Peebles, P. J. E. *The large-scale structure of the universe*. 1980.
- Perlmutter, S., Aldering, G., Goldhaber, G., Knop, R. A., Nugent, P., Castro, P. G., Deustua, S., Fabbro, S., Goobar, A., Groom, D. E. et al. *Measurements of Ω and Λ from 42 High-Redshift Supernovae*. ApJ, 517 :565–586, 1999.
- Planck Collaboration, , Ade, P. A. R., Aghanim, N., Arnaud, M., Ashdown, M., Aumont, J., Baccigalupi, C., Banday, A. J., Barreiro, R. B., Bartlett, J. G. et al. *Planck 2015 results. XIII. Cosmological parameters*. A&A, 594 :A13, 2016.

- Pomarède, D., Courtois, H. M., Hoffman, Y. and Tully, R. B. *Cosmography and Data Visualization*. PASP, 129(5) :058002, 2017.
- Prada, F., Klypin, A. A., Cuesta, A. J., Betancort-Rijo, J. E. and Primack, J. *Halo concentrations in the standard Λ cold dark matter cosmology*. MNRAS, 423 :3018–3030, 2012.
- Reid, B. A., Samushia, L., White, M., Percival, W. J., Manera, M., Padmanabhan, N., Ross, A. J., Sánchez, A. G., Bailey, S., Bizyaev, D. et al. *The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey : measurements of the growth of structure and expansion rate at $z = 0.57$ from anisotropic clustering*. MNRAS, 426 : 2719–2737, 2012.
- Rest, A., Scolnic, D., Foley, R. J., Huber, M. E., Chornock, R., Narayan, G., Tonry, J. L., Berger, E., Soderberg, A. M., Stubbs, C. W. et al. *Cosmological Constraints from Measurements of Type Ia Supernovae Discovered during the First 1.5 yr of the Pan-STARRS1 Survey*. ApJ, 795 :44, 2014.
- Riess, A. G., Filippenko, A. V., Challis, P., Clocchiatti, A., Diercks, A., Garnavich, P. M., Gilliland, R. L., Hogan, C. J., Jha, S., Kirshner, R. P. et al. *Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant*. AJ, 116 : 1009–1038, 1998.
- Roberts, M. S. and Rots, A. H. *Comparison of Rotation Curves of Different Galaxy Types*. A&A, 26 :483–485, 1973.
- Rubin, V. C. and Jr.Ford, W. K. *Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions*. ApJ, 159 :379, 1970.
- Rubin, V. C., Jr.Ford, W. K. and Thonnard, N. *Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605 $/R = 4\text{kpc}/$ to UGC 2885 $/R = 122\text{kpc}/$* . ApJ, 238 :471–487, 1980.
- Rybicki, G. B. and Press, W. H. *Interpolation, realization, and reconstruction of noisy, irregularly sampled data*. ApJ, 398 :169–176, 1992.
- Samushia, L., Percival, W. J. and Raccañelli, A. *Interpreting large-scale redshift-space distortion measurements*. MNRAS, 420 :2102–2119, 2012.
- Sargent, W. L. W. and Turner, E. L. *A statistical method for determining the cosmological density parameter from the redshifts of a complete sample of galaxies*. ApJ, 212 :L3–L7, 1977.
- Sheth, K., Regan, M., Hinz, J. L., Gil de Paz, A., Menéndez-Delmestre, K., Muñoz-Mateos, J.-C., Seibert, M., Kim, T., Laurikainen, E., Salo, H. et al. *The Spitzer Survey of Stellar Structure in Galaxies (S4G)*. PASP, 122 :1397, 2010.
- Sorce, J. *From Spitzer Mid-Infrared Observations and Measurements of Peculiar Velocities to Constrained Simulations of the Local Universe*. Theses, Université Claude Bernard Lyon 1, 2014. URL <http://hal.in2p3.fr/tel-01138701>.
- Sorce, J. G., Courtois, H. M. and Tully, R. B. *The Mid-infrared Tully-Fisher Relation : Spitzer Surface Photometry*. AJ, 144 :133, 2012.

- Sorce, J. G., Tully, R. B., Courtois, H. M., Jarrett, T. H., Neill, J. D. and Shaya, E. J. *From Spitzer Galaxy photometry to Tully-Fisher distances*. MNRAS, 444 :527–541, 2014.
- Sorce, J. G., Gottlöber, S., Yepes, G., Hoffman, Y., Courtois, H. M., Steinmetz, M., Tully, R. B., Pomarède, D. and Carlesi, E. *Cosmicflows Constrained Local Universe Simulations*. MNRAS, 455 :2078–2090, 2016.
- Sousbie, T. *The persistent cosmic web and its filamentary structure - I. Theory and implementation*. MNRAS, 414 :350–383, 2011.
- Springob, C. M., Masters, K. L., Haynes, M. P., Giovanelli, R. and Marinoni, C. *SFI++ . II. A New I-Band Tully-Fisher Catalog, Derivation of Peculiar Velocities, and Data Set Properties*. ApJS, 172 :599–614, 2007.
- Springob, C. M., Magoulas, C., Colless, M., Mould, J., Erdoğan, P., Jones, D. H., Lucey, J. R., Campbell, L. and Fluke, C. J. *The 6dF Galaxy Survey : peculiar velocity field and cosmography*. MNRAS, 445 :2677–2697, 2014.
- Strauss, M. A. and Willick, J. A. *The density and peculiar velocity fields of nearby galaxies*. Phys. Rep., 261 :271–431, 1995.
- Taylor, R., Davies, J. I., Auld, R. and Minchin, R. F. *The Arecibo Galaxy Environment Survey - V. The Virgo cluster (I)*. MNRAS, 423 :787–810, 2012.
- Taylor, R., Davies, J. I., Auld, R., Minchin, R. F. and Smith, R. *The Arecibo Galaxy Environment Survey - VI. The Virgo cluster (II)*. MNRAS, 428 :459–469, 2013.
- Taylor, R., Minchin, R. F., Herbst, H., Davies, J. I., Rodriguez, R. and Vazquez, C. *The Arecibo Galaxy Environment Survey - VII. A dense filament with extremely long H I streams*. MNRAS, 443 :2634–2649, 2014.
- Tempel, E., Kipper, R., Saar, E., Bussov, M., Hektor, A. and Pelt, J. *Galaxy filaments as pearl necklaces*. A&A, 572 :A8, 2014.
- Theureau, G., Bottinelli, L., Coudreau-Durand, N., Gouguenheim, L., Hallet, N., Loulergue, M., Paturel, G. and Teerikorpi, P. *Kinematics of the local universe. VII. New 21-cm line measurements of 2112 galaxies*. A&AS, 130 :333–339, 1998.
- Theureau, G., Coudreau, N., Hallet, N., Hanski, M., Alsac, L., Bottinelli, L., Gouguenheim, L., Martin, J.-M. and Paturel, G. *Kinematics of the local universe . XII. 21-cm line measurements of 586 galaxies with the new Nançay receiver*. A&A, 430 :373–383, 2005.
- Theureau, G., Coudreau, N., Hallet, N., Hanski, M. O. and Poulain, M. *Kinematics of the Local Universe. XIV. Measurements from the 21 cm line and the HI mass function from a homogeneous catalog gathered with the Nançay radio telescope*. A&A, 599 :A104, 2017.
- Tonry, J. and Schneider, D. P. *A new technique for measuring extragalactic distances*. AJ, 96 :807–815, 1988.
- Tonry, J. L., Blakeslee, J. P., Ajhar, E. A. and Dressler, A. *The SBF Survey of Galaxy Distances. I. Sample Selection, Photometric Calibration, and the Hubble Constant*. ApJ, 475 :399–413, 1997.

- Tonry, J. L., Dressler, A., Blakeslee, J. P., Ajhar, E. A., Fletcher, A. B., Luppino, G. A., Metzger, M. R. and Moore, C. B. *The SBF Survey of Galaxy Distances. IV. SBF Magnitudes, Colors, and Distances.* ApJ, 546 :681–693, 2001.
- Tully, R. B. *Nearby galaxies catalog.* 1988.
- Tully, R. B. *Galaxy Groups : A 2MASS Catalog.* AJ, 149 :171, 2015.
- Tully, R. B. and Fisher, J. R. *A new method of determining distances to galaxies.* A&A, 54 :661–673, 1977.
- Tully, R. B. and Fouque, P. *The extragalactic distance scale. I - Corrections to fundamental observables.* ApJS, 58 :67–80, 1985.
- Tully, R. B. and Pierce, M. J. *Distances to Galaxies from the Correlation between Luminosities and Line Widths. III. Cluster Template and Global Measurement of H_0 .* ApJ, 533 :744–780, 2000.
- Tully, R. B., Shaya, E. J., Karachentsev, I. D., Courtois, H. M., Kocevski, D. D., Rizzi, L. and Peel, A. *Our Peculiar Motion Away from the Local Void.* ApJ, 676 :184–205, 2008.
- Tully, R. B., Courtois, H. M., Dolphin, A. E., Fisher, J. R., Héraudeau, P., Jacobs, B. A., Karachentsev, I. D., Makarov, D., Makarova, L., Mitronova, S. et al. *Cosmicflows-2 : The Data.* AJ, 146 :86, 2013.
- Tully, R. B., Courtois, H., Hoffman, Y. and Pomarède, D. *The Laniakea supercluster of galaxies.* Nature, 513 :71–73, 2014.
- Tully, R. B., Courtois, H. M. and Sorce, J. G. *Cosmicflows-3.* AJ, 152 :50, 2016.
- Turnbull, S. J., Hudson, M. J., Feldman, H. A., Hicken, M., Kirshner, R. P. and Watkins, R. *Cosmic flows in the nearby universe from Type Ia supernovae.* MNRAS, 420 : 447–454, 2012.
- van Driel, M., Butcher, Z., Schneider, S., Lehnert, M., Minchin, R., Blyth, S.-L., Chemin, L., Hallet, N., Joseph, T., Kotze, P. et al. *VizieR Online Data Catalog : NIBLES. I. The Nançay HI survey (van Driel+, 2016).* VizieR Online Data Catalog, 359, 2016a.
- van Driel, W., Schneider, S. E., Lehnert, M. and Minchin, R. *NIBLES : an H I census of local SDSS galaxies.* In Minchin, R. and Momjian, E., editors, *The Evolution of Galaxies Through the Neutral Hydrogen Window*, volume 1035 of *American Institute of Physics Conference Series*, pages 256–258, 2008. doi: 10.1063/1.2973597.
- van Driel, W., Butcher, Z., Schneider, S., Lehnert, M. D., Minchin, R., Blyth, S.-L., Chemin, L., Hallet, N., Joseph, T., Kotze, P. et al. *NIBLES : an H I census of stellar mass selected SDSS galaxies. I. The Nançay H I survey.* A&A, 595 :A118, 2016b.
- Walker, E. S., Baltay, C., Campillay, A., Citrenbaum, C., Contreras, C., Ellman, N., Feindt, U., González, C., Graham, M. L., Hadjiyska, E. et al. *First Results from the La Silla-QUEST Supernova Survey and the Carnegie Supernova Project.* ApJS, 219 : 13, 2015.
- Watkins, R. and Feldman, H. A. *An unbiased estimator of peculiar velocity with Gaussian distributed errors for precision cosmology.* MNRAS, 450 :1868–1873, 2015.

- Werner, M. W., Roellig, T. L., Low, F. J., Rieke, G. H., Rieke, M., Hoffmann, W. F., Young, E., Houck, J. R., Brandl, B., Fazio, G. G. et al. *The Spitzer Space Telescope Mission*. ApJS, 154 :1–9, 2004.
- Wiener, N., . *Extrapolation, interpolation, and smoothing of stationary time series : with engineering applications*. Technology press books in science and engineering. Technology Press of the Massachusetts Institute of Technology, 1964.
- Wijnholds, S., van der Tol, S., Nijboer, R. and van der Veen, A.-J. *Calibration challenges for future radio telescopes*. IEEE Signal Processing Magazine, 27 :30–42, 2010.
- Winkel, B., Kalberla, P. M. W., Kerp, J. and Flöer, L. *The Effelsberg-Bonn H I Survey : Data Reduction*. ApJS, 188 :488–499, 2010.
- Wong, O. I., Ryan-Weber, E. V., Garcia-Appadoo, D. A., Webster, R. L., Staveley-Smith, L., Zwaan, M. A., Meyer, M. J., Barnes, D. G., Kilborn, V. A., Bhathal, R. et al. *The Northern HIPASS catalogue - data presentation, completeness and reliability measures*. MNRAS, 371 :1855–1864, 2006.
- Yepes, G., Martínez-Vaquero, L. A., Gottlöber, S. and Hoffman, Y. *The CLUES project : Constrained Local UniversE Simulations*. In Balazs, C. and Wang, F., editors, *American Institute of Physics Conference Series*, volume 1178 of *American Institute of Physics Conference Series*, pages 64–75, 2009. doi: 10.1063/1.3264558.
- Yun, M. S., Ho, P. T. P. and Lo, K. Y. *A high-resolution image of atomic hydrogen in the M81 group of galaxies*. Nature, 372 :530–532, 1994.
- Yvon, J. *La théorie statistique des fluides et l'équation d'état*. Actualités scientifiques et industrielles. Hermann & cie, 1935. URL https://books.google.fr/books?id=SLI_AQAAIAAJ.
- Zaroubi, S., Hoffman, Y., Fisher, K. B. and Lahav, O. *Wiener Reconstruction of the Large-Scale Structure*. ApJ, 449 :446, 1995.
- Zaroubi, S., Hoffman, Y. and Dekel, A. *Wiener Reconstruction of Large-Scale Structure from Peculiar Velocities*. ApJ, 520 :413–425, 1999.
- Zwicky, F. *Die Rotverschiebung von extragalaktischen Nebeln*. Helvetica Physica Acta, 6 :110–127, 1933.