

ÉCOLE DOCTORALE 182

Observatoire astronomique de Strasbourg

THÈSE

présentée par :

Maxime BEURET

soutenue le **21 Septembre 2016**

pour obtenir le grade de : **Docteur de l'université de Strasbourg**

Discipline/ Spécialité : Astrophysique

Formation stellaire dans la galaxie et interaction avec le milieu interstellaire.

THÈSE dirigée par :

Mr. CAMBRÉSY Laurent
Mr. BIENAYMÉ Olivier

Astronome HDR, Observatoire astronomique de Strasbourg
Astronome HDR, Observatoire astronomique de Strasbourg

RAPPORTEURS :

Mr. Braine Jonathan
Mme Russeil Delphine

Astronome HDR, Laboratoire d'Astrophysique de Bordeaux
Maître de conférences HDR, Laboratoire d'Astrophysique
de Marseille

EXAMINATEUR :

Mme Lançon Ariane

Professeur HDR, Observatoire astronomique de Strasbourg

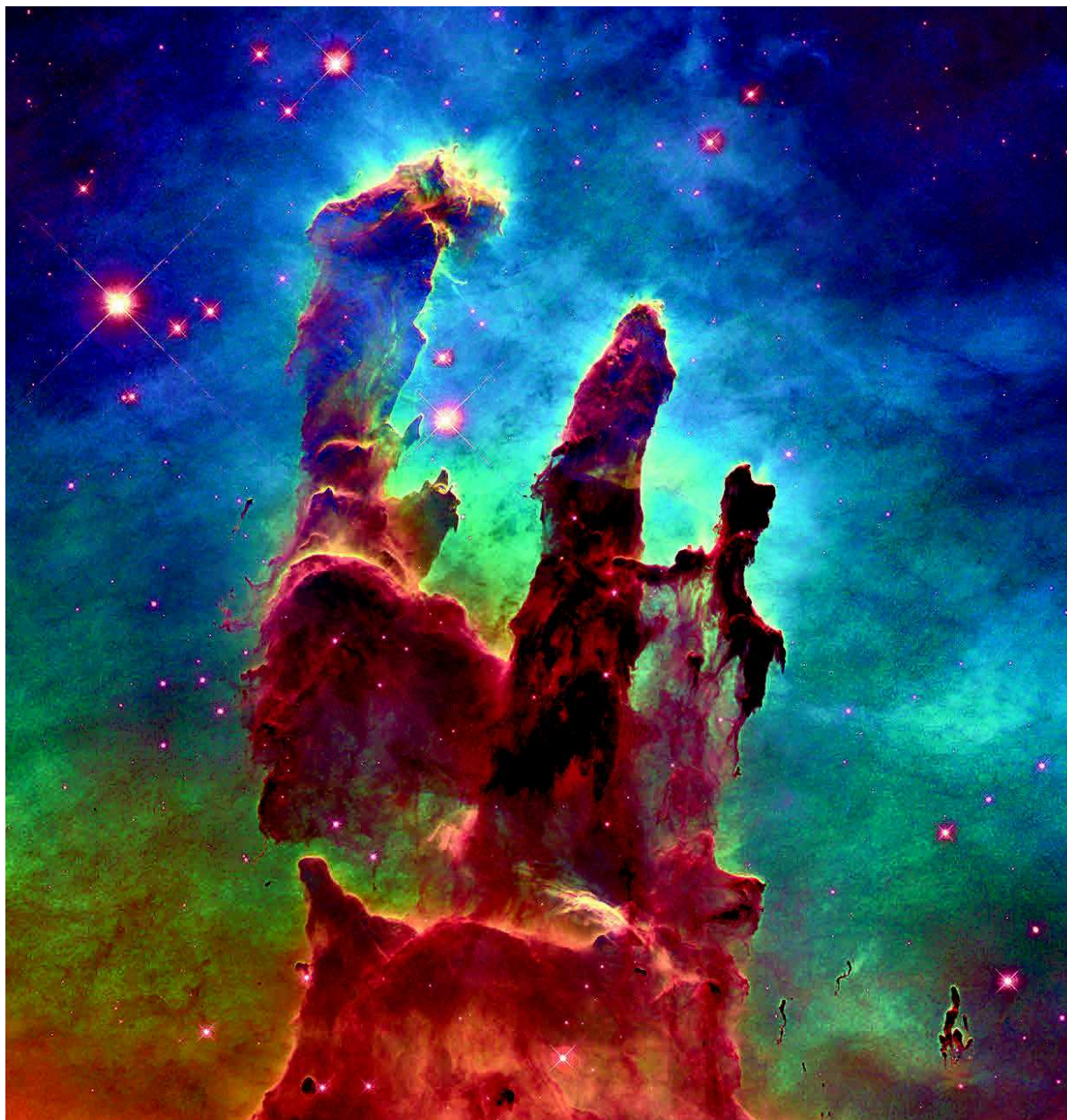


FIGURE 1: Pilliers de la création dans M16 vu par Hubble. Crédit image : NASA, ESA et The Hubble Heritage Team.

*"Les gens ont des étoiles qui ne sont pas les mêmes.
Pour les uns, qui voyagent, les étoiles sont des guides.
Pour d'autres elles ne sont rien que de petites lumières.
Pour d'autres qui sont savants elles sont des
problèmes..."*

Antoine de Saint-Exupéry (Le Petit Prince, 1943)

*"Ce sont les étoiles, les étoiles tout là-haut qui
gouvernent notre existence."*

William Shakespeare (Le Roi Lear, 1603-1606)

Remerciements

La thèse est un travail de longue haleine qui permet d'accroître à la fois ses connaissances dans le domaine d'étude choisi, mais aussi dans son développement personnel à travers les échanges, les questionnements et les difficultés rencontrées. Je tiens ainsi à remercier mon directeur de thèse, Laurent Cambrésy, pour sa disponibilité, ses conseils et son temps investi tout au long de ces trois années. Je remercie aussi mon co-directeur de thèse Olivier Bienaymé pour m'avoir permis de mener à bien cette thèse. Je remercie Nicolas Billot, post-doc à l'IRAM, pour l'apport tant scientifique que personnel qu'il m'a apporté pendant ce doctorat. Je tiens tout particulièrement à remercier Roberta Paladini, chercheuse à l'IPAC/Caltech, pour son aide précieuse durant la deuxième moitié de ma thèse et pour m'avoir accueilli à Caltech d'Avril à Juillet 2015.

Je remercie aussi le directeur de l'Observatoire astronomique de Strasbourg, Hervé Wozniak, pour m'avoir permis d'effectuer ce doctorat au sein de cet établissement. Je remercie toute l'équipe du CDS de m'avoir accueilli. Je remercie notamment Bernd Vollmer, Caroline Bot et Françoise Génova, mais aussi les chercheurs d'autres équipes comme Nicolas Martin, Laurence Boirin, Ariane Lançon ou encore Christian Boily pour leurs discussions sur des domaines variés.

Je remercie ensuite les rapporteurs de ma thèse d'avoir accepté cette tâche, Jonathan Braine, astronome au Laboratoire d'Astrophysique de Bordeaux et Delphine Russeil, maître de conférences au Laboratoire d'Astrophysique de Marseille. Je remercie aussi Ariane Lançon, professeur à l'Observatoire astronomique de Strasbourg, pour avoir accepté d'être examinateur de ma thèse.

Je remercie aussi tous les doctorants de l'observatoire pour leur pouvoir de rendre l'observatoire encore plus chaleureux et leur joie de vivre, notamment les deux Nicolas, Julien, JB, François, Mathieu, Guillaume, Enmanuelle, Jérôme, Jeremy, Khyati. Dans le même cadre, je remercie mes amis de m'avoir accompagné en m'encourageant et en me permettant de penser à autres choses dans les étapes plus difficiles. Flo, NicoLi, NicoBi, Lex, Dany, Gillou, merci les gars !

Je termine en remerciant mes parents pour leurs soutiens et pour m'avoir laissé l'opportunité de faire les études que je voulais afin d'arriver jusqu'ici. Enfin, je remercie mille fois Anaïs pour m'avoir supporté pendant ces trois ans (et bien plus encore) ainsi que pour son soutien et son réconfort.

TABLE DES MATIÈRES

Sommaire	ix
Liste des tables	xiii
Liste des figures	xv
Introduction	xxiii
 I Contexte astrophysique	 1
1 Milieu Interstellaire et formation stellaire	3
1.1 Histoire du Milieu Interstellaire	4
1.2 Constitution du MIS	4
1.2.1 Rayons cosmiques	6
1.2.2 Champ magnétique	6
1.2.3 Abondances	6
1.2.4 Le gaz	7
1.2.4.1 Le gaz neutre	7
1.2.4.2 Le gaz moléculaire	8
1.2.4.3 Le gaz ionisé	9
1.2.5 La poussière	10
1.2.5.1 Rôles des poussières	10
1.2.5.2 Types de poussières	12
1.3 Cycle de vie et structure du Milieu Interstellaire	13
1.3.1 Cycle de vie	13
1.3.2 Des grandes structures aux étoiles	14
1.3.3 Régions associées à la formation stellaire	15
1.4 Les différents stades évolutifs des proto-étoiles.	17

II	Travaux et résultats	23
2	Catalogue multi-longueurs d'onde	25
2.1	Observations dans les domaines infrarouge et sub-millimétrique	26
2.1.1	Télescopes antérieurs à Herschel	26
2.1.2	Herschel	28
2.1.3	Relevé Hi-GAL	29
2.2	Les données et outils utilisés	30
2.2.1	Catalogues single-band Hi-GAL	30
2.2.2	Specfind	31
2.3	Catalogue multi-longueurs d'onde	32
2.3.1	Conditions d'ajustement	32
2.3.2	Résultats et catalogue final	33
2.4	Catalogue de l'équipe Hi-GAL	37
2.4.1	Contenu	37
2.4.2	Notion de <i>clumps</i>	37
2.4.3	Comparaison et Discussion	38
2.4.4	Remarques	42
3	Le catalogue d'amas	45
3.1	La méthode	46
3.1.1	Les méthodes de sélection de sur-densités	46
3.1.1.1	Carte de densité	46
3.1.1.2	Méthode des k plus proches voisins	46
3.1.1.3	Cellules de Voronoï	47
3.1.1.4	Minimum Spanning Tree	48
3.1.2	Choix de la méthode MST	49
3.2	Construction du catalogue d'amas	50
3.2.1	Stades évolutifs des clumps et confusion	50
3.2.2	Contraintes imposées par la structure galactique	53
3.2.3	Distances héliocentriques et catalogue d'amas fiables	59
3.3	Caractéristiques du catalogue d'amas	61
3.4	Discussion	64
3.5	Description des colonnes	66
3.5.1	Catalogue d'amas fiables	66
3.5.2	Catalogue archive	68
3.6	Conclusion	69
4	Analyse du catalogue d'amas	71
4.1	Distribution spatiale des amas	72
4.1.1	Validité des amas	72
4.1.2	Distribution spatiale	74
4.1.2.1	Comparaison avec les bras spiraux	74
4.1.2.2	Hauteur des amas	76
4.2	Distribution spatiale des clumps	80
4.2.1	Concentration et fractale	81
4.2.2	Distribution spatiale	83
4.3	Distribution des masses des clumps	85
4.3.1	Ségrégation de masse	85
4.3.2	Fonction de masse des clumps	88

4.3.2.1	IMF et CMF	88
4.3.2.2	Application aux données	90
4.3.2.3	Multiplicité et efficacité stellaire	94
4.4	Association avec des objets de la littérature	95
4.4.1	Régions H II	96
4.4.2	IRDCs	97
4.4.3	Comparaison avec des catalogues d'amas enfouis pré-existants . .	98
4.5	Discussions et biais observationnels	99
4.6	Conclusion	103
III	Conclusion et perspectives	105
5	Conclusion et perspectives	107
5.1	Conclusion	108
5.1.1	Catalogue de clumps	108
5.1.2	Catalogue d'amas	108
5.1.3	Analyse des amas	109
5.2	Perspectives	110
IV	Annexes	113
Annexes		115
A	Article "Sources clustering in the Hi-GAL survey using a Minimum Spanning Tree method."	115
B	Catalogue d'amas	139
C	Constantes et variables	155
D	Glossaire	157
D.1	Acronymes	158
D.2	Instruments dans l'infrarouge et le sub-millimétrique	159
Bibliographie		161

LISTE DES TABLEAUX

1.1	Ordre de grandeur de la densité volumique, température et masse totale des différentes composantes du milieu interstellaire.	5
1.2	Abondance des principaux éléments du MIS. Les abondances solaires sont exprimées comme $12 + \log(X/H)$, où X est le nombre d'atomes de l'élément et H représente l'hydrogène. Pour ρ Ophiuchus, les valeurs représentent la différence avec les abondances solaires, $[X/H] = \log(X/H) - \log(X/H)_{\odot}$. Données de Savage & Sembach (1996)	7
1.3	Grandeurs caractéristiques des composantes du milieu interstellaire froid.	16
2.1	Resolutions de divers télescopes terrestres et spatiaux dans l'infrarouge et le sub-millimétrique	29
2.2	Nombre de sources dans les catalogues single-band.	31
2.3	Caractéristiques des catalogues pour chaque condition de départ. Les deux conditions se rapportent à l'ajustement par un corps noir modifié et concernent la prise en compte du catalogue à $70 \mu m$ et la possibilité de laisser β comme paramètre libre.	34
2.4	Taille des résolutions angulaires à différentes distances héliocentriques.	38
4.1	Résultats de l'ajustement de la CMF des clumps en amas et des clumps isolés pour l'ensemble de la partie interne du plan galactique ainsi que pour les premier et quatrième quadrants.	92
B.1	Complete list of cluster candidates.	140
C.1	Constantes utiles	156
C.2	Conversions d'unité utiles	156

TABLE DES FIGURES

1	Pilliers de la création dans M16 vu par Hubble. Crédit image : NASA, ESA et The Hubble Heritage Team.	iii
1.1	Distribution de la brillance du ciel aux latitudes galactiques élevées, de l'infrarouge proche au millimétrique. Les lignes pointillées (associées aux termes OH, O 2 et Airglow) correspondent aux émissions dans la haute atmosphère. L'oxygène atomique (OI) et la raie Lyman α ($Ly\alpha$) proviennent de la magnétosphère terrestre. Les étoiles (faint stars), la lumière zodiacale et le CMB participe tous au spectre observé de manière non négligeable. Les cirrus sont des nuages diffus, observés aux hautes latitudes. Enfin, la ligne en tiret correspond à l'émission des différents types de poussières. Extrait de Leinert et al. (1998)	5
1.2	Carte CO de Dame et al. (2001) permettant de tracer la molécule H 2. . .	9
1.3	Représentation d'un modèle d'agglomération des grains de poussières. (credit : http://www.astro.ucla.edu/~wright/dust/)	10
1.4	Extinction par atome d'hydrogène calculée pour le modèle de Draine & Lee (1984) en comparaison avec les observations de l'extinction galactique moyenne. Extrait de Draine & Lee (1984)	11
1.5	Modèle d'émission des grains de poussières composés de plusieurs espèces, confronté aux observations de cirrus. Extrait de Desert et al. (1990)	12
1.6	Cycle de vie du MIS (crédit Kaiser, 2002)	14
1.7	Vue d'artiste de la Voie Lactée et de ses bras spiraux vu par un observateur extérieur. Credit : NASA/Adler/U. Chicago/Wesleyan/JPL-Caltech	15
1.8	relation entre les tailles des différentes structures, allant des bras spiraux aux étoiles, et leur temps de formation. Extrait de (Efremov & Elmegreen, 1998).	16
1.9	Exemple de régions H II avec des PDRs visibles et quasi-sphériques, provenant de Deharveng et al. (2005) . Composition de couleurs d'images du télescope MSX. La couleur turquoise correspond à l'émission des PAHs dans la bande A ($8.3 \mu m$). La couleur rouge correspond à l'émission des VSGs dans la bande E ($21.3 \mu m$).	18

1.10	Image composite du filament <i>snake</i> observé avec Spitzer, à $4.5\ \mu m$ (bleu), $8\ \mu m$ (vert) et $24\ \mu m$ (rouge). Les deux flèches localisent des sources infrarouges identifiées en proto-étoiles. Extrait de Wang et al. (2014)	19
1.11	Haut : Filaments de la partie sud de la region Serpens (André et al., 2010). Bas : Filaments dans IC 5146 (Arzoumanian et al., 2013, 2011).	20
1.12	Principales régions de formation stellaire dans le plan galactique. L'image de fond correspond à la distribution du gaz CO (Dame et al., 2001).	21
1.13	SED des proto-étoiles à leurs différents stades évolutifs. De gauche à droite : Classe 0, Classe I, Classe II, Classe III et la ZAMS (crédit Schulz, 2012) .	21
1.14	SED des proto-étoiles à leurs différents stades évolutifs. Haut-gauche : Classe 0, Haut-droite : Classe I, bas-gauche : Classe II, bas-droite : Classe III. (Crédit : http://newt.phys.unsw.edu.au/jacara/pilotscience.php)	22
2.1	Transmission de l'atmosphère terrestre en fonction des domaines du spectre électromagnétique. La transmission est quasi nulle dans l'infrarouge et le sub-millimétrique, à l'exception de quelques bandes dans l'infrarouge proche. (Crédit : https://fr.wikipedia.org/wiki/Atmosph%C3%A8re_terrestre)	26
2.2	Intervalles de longueur d'onde des principaux télescopes spatiaux actuels et passés. (Crédit : https://fr.wikipedia.org/wiki/T%C3%A9lescope_spatial)	27
2.3	Composition RGB des images à 70 , 160 et $250\ \mu m$ entre 348 et 354 degrés. 30	
2.4	Exemple de confusion due à la résolution. L'objet est en fait composé d'au moins deux sous-objets à $70\ \mu m$	32
2.5	Histogramme des rapports entre l'erreur sur le flux et le flux à longueur d'onde. Le pic des histogrammes se situe aux alentours de 10% du flux. .	33
2.6	Loi du corps noir modifiée comme développée dans l'équation 2.1, pour différentes températures avec $\beta = 2$. Les lignes pointillées localisent les 5 longueurs d'onde utilisées dans le cadre du projet Hi-GAL. On s'aperçoit qu'à $100\ K$ nous sommes dans le régime Rayleigh-Jeans à toutes les longueurs d'onde.	35
2.7	Corrélation apparente entre la température du corps noir modifiée et β en prenant en compte toutes les longueurs d'onde disponibles (partie gauche) et sans $70\ \mu m$ (partie droite). Les températures minimales et maximales ont été fixées à $5\ K$ et $100\ K$ comme expliqué dans le texte plus haut. . . .	36
2.8	Histogramme des distances héliocentriques des clumps. Il semble y avoir deux zones privilégiées de localisation des clumps, vers $2-3\ kpc$ et dans l'intervalle $8-12\ kpc$	40
2.9	Gauche : Histogramme des distances héliocentriques des clumps. Droite : Histogramme des masses des clumps (log).	40
2.10	Gauche : Histogramme des températures des clumps pré-stellaires (rouge) et proto-stellaires (bleu). Droite : Même histogramme mais pour le catalogue <i>product</i>	41
2.11	Différence observée entre les flux des catalogues single-band et du catalogue bandmerged. La redéfinition des flux au-delà de $250\ \mu m$ entraîne un écart important avec les flux originaux.	41
2.12	Gauche : Comparaison de la distribution des masses entre la dernière version du catalogue <i>product</i> et celle utilisée. Droite : Même comparaison mais pour les températures.	43

3.1	Diagramme de Voronoï et triangulation de Delaunay. Gauche : Construction de la triangulation de Delaunay. Aucune source ne doit se trouver à l'intérieur des cercles circonscrits. Centre : Lien entre la triangulation de Delaunay (noir) et le diagramme de Voronoï (rouge). Les points rouges sont à la fois les sommets du diagramme de Voronoï et le centre des cercles circonscrits. Droite : Diagramme de Voronoï final d'un échantillon de points (rouge).	48
3.2	Exemple d'un graphe et du MST associé. Le poids attribué est proportionnel aux distances entre les points.	49
3.3	Distribution des clumps dans un champ de 2 degrés carrés, centré à $l = 330$ degrés.	51
3.4	Distribution des tailles de branches (en gris) correspondant au champ de la figure 3.3. La distribution cumulative est représentée en noir. Les deux segments ajustés à cette distribution cumulative permettent d'obtenir la valeur de Λ_{cut} (en pointillé).	51
3.5	Résultat du MST en suivant la méthode de Koenig et al. (2008) et Gutermuth et al. (2009) , appliquée sur l'échantillon de la figure figure 3.3. Les points et lignes gris correspondent aux clumps rejetés ($\Lambda > \Lambda_{cut}$) alors que les lignes et points bleus correspondent à $\Lambda < \Lambda_{cut}$. Les lignes rouges encadrent les sur-densités avec $N \geq 10$	52
3.6	Carte de densité des clumps unbound pour une région localisée entre $26 \leq l \leq 31$ degrés. La même distribution est observée dans l'échantillon de données total.	52
3.7	Même représentation que la figure 3.6 pour les clumps pré-stellaires. . . .	53
3.8	Même représentation que la figure 3.6 pour les clumps proto-stellaires. . .	53
3.9	Relation entre Λ_{cut} et la densité moyenne de la zone considérée. La ligne rouge correspond à la fonction de corrélation $\Lambda_{cut} \propto d_{mean}^{-0.24}$	54
3.10	Exemple de reconstruction d'un catalogue de sur-densités. La taille de fenêtre correspond à 10 degrés. Les points gris correspondent aux clumps pré-stellaires et proto-stellaires. Les polygones gris sont les sur-densités possédant aux moins 10 clumps. Les polygones rouges sont les sur-densités divisées en deux par les fenêtres originelles. Les rectangles noirs (lignes pointillées et pleines) sont les régions utilisées pour reconstruire le catalogue. La fenêtre du dessus est la fenêtre de chevauchement de cette région. Les deux fenêtres du milieu sont les fenêtres originelles et celle en dessous est la fenêtre finale.	56
3.11	Distribution de la fraction moyenne des clumps dans les sur-densités en fonction de la longitude. La ligne noire correspond à la fraction moyenne des clumps dans les sur-densités pour chaque catalogue (un pour chaque taille de fenêtre). La taille de bin est définie à 5 degrés. La barre d'erreur correspond à l'erreur à 1σ de la fraction de clumps dans les sur-densités. Les deux lignes pointillées correspondent à la même distribution mais pour une variation de valeur de $\pm 5\%$ de Λ_{cut}	57
3.12	Fraction de clumps dans les sur-densités en fonction de la taille des fenêtres, pour un champ de 2 degrés carrés centré sur $l = 335$ degrés (gauche) et $l = 20$ degrés (droite).	57
3.13	Distribution de Λ_{cut} en fonction de la longitude pour quatre tailles de fenêtre, 2, 10, 20 et 30 degrés. La longitude correspond au centre de chaque fenêtre.	58

3.14	Distribution du nombre de clumps dans les sur-densités, à partir de deux clumps.	59
3.15	Distribution des distances héliocentriques des amas les plus fiables. Haut : Distribution des distances calculées à partir du pic de la distribution des HDEs des clumps. Bas : Correlation entre les distances calculées à partir du pic de la distribution des HDEs des clumps et à partir de leur moyenne. Le coefficient de corrélation vaut 0.99925 ce qui signifie que les distributions de HDEs sont significativement piquées.	62
3.16	Haut : Distribution des rayons équivalents des amas calculés à partir des demi-petit axe et demi-grand axe de l'ellipse. Bas : Rapport entre le petit axe et le grand axe de l'ellipse.	63
3.17	Distribution des masses volumiques des amas calculées à partir du total des masses des clumps de chaque amas.	64
3.18	haut : Distribution du nombre d'amas en fonction de Λ_{cut} pour l'ensemble des fenêtres de 10 degrés. Bas : Distribution des Λ_{cut} de notre étude (en rouge) suivant la longitude en comparaison avec celle des Λ_{cut} (en noir) de la méthode de Battinelli (1991). Les longitudes des points correspondent au centre des fenêtres.	65
3.19	Distribution spatiale des sur-densités et de leur clumps dans une zone dense du plan galactique. Haut : Amas fiables. Bas : Ensemble des sur-densités trouvées pour cette zone. La figure de fond correspond à une composition colorée à 160, 250 et 350 μ m.	67
3.20	Exemple d'une sur-densité rejetée pour cause de confusion sur la ligne de visée. Les distances héliocentriques sont comprises entre 2.8 kpc et 12.3 kpc et il est difficile d'extraire une sur-densité fiable même dans les zones les plus lumineuses. La figure de fond correspond à une composition colorée à 160, 250 et 350 μ m.	68
3.21	Exemple d'une région de formation stellaire caractérisée par la présence de deux amas du catalogue fiable mais qui ne comprend en réalité qu'un seul amas enfoui. La figure de fond correspond à une composition colorée à 160, 250 et 350 μ m.	69
4.1	Distribution des amas dans un diagramme <i>clustering parameter-magnitude</i>	73
4.2	Exemple d'amas composés de filaments. L'image de fond est une image Hi-GAL à 250 μ m.	74
4.3	Carte de densité des amas. Les trois zones montrées permettent de comparer les premier et quatrième quadrants. La différence la plus significative se situe respectivement entre $l=46$ à $l=30$ deg et $l=314$ et $l=330$ deg.	75
4.4	Distribution des amas dans la galaxie en comparaison des bras spiraux provenant de Englmaier et al. (2011).	76
4.5	Distribution des amas dans la galaxie en comparaison des bras spiraux provenant de Taylor & Cordes (1993). Quelques régions de formation stellaire connues ont été ajoutées. Les amas semblent mieux ajuster les bras comparé à ceux de la figure 4.4.	77
4.6	Histogramme de la distance galactocentrique des amas, R. Le pic se situe aux alentours de 6 kpc.	78
4.7	Histogramme de la hauteur des amas, pour tous les amas (figure de gauche) et en séparant les premier et quatrième quadrants (figure de droite).	79

4.8	Distribution de la valeur moyenne de la hauteur des amas dans des bins de taille $0.1R_0$, en fonction de la distance galactocentrique. La partie inférieure de la figure représente la dispersion des valeurs σ_z	80
4.9	Carte en 2D de la fonction a_{mod} correspondante à l'équation 4.5. Les bras spiraux proviennent de Taylor & Cordes (1993)	81
4.10	Exemple d'un amas possédant une distribution centrée, provenant de simulation N-corps (Küpper et al., 2011).	82
4.11	Plusieurs exemples de fractale. <i>Gauche</i> : Ensemble de Mandelbrot. <i>Milieu</i> : Construction du flocon de Koch. <i>Droite</i> : Représentation d'un flocon de Koch.	82
4.12	Relation entre \bar{m} et \bar{s} de la fonction Q (équation 4.6), caractérisant le profil de distribution des clumps d'un amas. La droite correspond à la valeur $Q = 0.8$ séparant une distribution centrée d'une distribution fractale.	84
4.13	Exemple de distribution spatiale des clumps d'un amas avec l'ajout des valeurs des masses. Le fond correspond à l'image de l'amas vu à $70 \mu m$	86
4.14	Figure de Kirk et al. (2016) montrant la ségrégation de masse dans Orion A, Orion B et des YSOs en amas d'une étude précédente (Kirk & Myers, 2011).	87
4.15	Méthode de Kirk et al. (2016) pour étudier la ségrégation de masse au sein des amas. Le graphique de gauche montre le lien entre la plus grande masse et le rapport d'offset. Le graphique de droite représente la même chose mais pour la deuxième plus grande masse.	87
4.16	CMF des clumps en amas. Les lignes rouges correspondent aux segments ajustés et nous fournis les pentes α_1 , α_2 and α_3 ainsi que les masses de transition m_1 and m_2 . <i>Gauche</i> : clumps avec HDEs. <i>milieu</i> : Ensemble des clumps des amas. <i>Droite</i> : Clumps avec recalcule des HDEs. Les trois considérations n'altèrent pas le profil de la CMF.	91
4.17	CMF des clumps en amas (gauche) et des clumps isolés (droite). Les lignes rouges correspondent aux mêmes segments que dans la figure 4.16.	91
4.18	Comparaison des valeurs des trois pentes de la CMF provenant de la table 4.1 entre les clumps isolés ($\alpha_{isolated}$) et les clumps en amas ($\alpha_{clustered}$). α_1 est en rouge, α_2 est en bleu et α_3 est en vert. Les valeurs des pentes de l'IMF de Kroupa sont aussi montrées en noir. La ligne noire représente la zone de valeurs égales entre les clumps isolés et les clumps en amas.	93
4.19	Distributions de la masse des clumps dans W3 pour une zone de forte densité (triangles) et une zone de faible densité (cercles). Les deux parties de la figure correspondent à deux calculs différents des masses. Dans toutes les distributions, nous pouvons observer un <i>saut</i> suivie d'une <i>bosse</i> entre $\log_{10}(m) = 1.5 (M_\odot)$ et $\log_{10}(m) = 2 (M_\odot)$. Cette figure est extraite de (Moore et al., 2007).	94
4.20	Distribution des masses des clumps en amas. Rouge : Amas possédant une ou plusieurs régions H II. Bleu : Amas sans association avec une région H II.	97
4.21	Distribution des diamètres des amas de Maj2013 et Mor2013 (en noir) comparé aux diamètres des amas de notre étude (en rouge).	99
4.22	Distribution de la distance des clumps en amas. La ligne verte représente la taille de la résolution à $70 \mu m$. La ligne orange correspond à la même taille mais à $500 \mu m$. La ligne grise correspond à la longueur de Jeans ($0.1 pc$).	100

4.23	Logarithme de la densité surfacique des amas (pc^{-2}) en fonction du logarithme de la masse des clumps. Le panel de couleur correspond aux distance héliocentriques de amas en kpc (barre de couleurs).	101
4.24	Logarithme des distances moyennes entre les clumps, évaluées à partir des tailles des branches λ du MST, en fonction du logarithme de la masse des clumps. Le panel de couleur correspond aux distance héliocentriques de amas en kpc (barre de couleurs).	101
4.25	Logarithme de l'aire des amas (pc^2) en fonction du logarithme de la masse des clumps. Le panel de couleur correspond aux distance héliocentriques de amas en kpc (barre de couleurs).	102
4.26	Partie (d) : Logarithme du nombre de clumps par amas en fonction du logarithme de la masse des clumps. Le panel de couleur correspond aux distance héliocentriques de amas en kpc (barre de couleurs).	102
5.1	Régions H II proche de W51.	110
5.2	Amas enfouis de notre étude associés à la région H II G5198+054 situé à 10 kpc. Une structure filamentaire se dessine sur la partie haute-droite. . .	111

INTRODUCTION

L'Univers observable est constitué d'une multitude d'objets observables à différentes échelles qui ont tendance à se regrouper ensemble sous l'effet de la gravitation. En remontant dans le temps, quelques centaines de millions d'années après le Big Bang, on s'aperçoit que les premiers objets massifs à s'être formés sont les quasars, c'est-à-dire des galaxies primordiales possédant un noyau actif (très lumineux). Dans le même temps les premières étoiles, dites de population III, se seraient formées c'est-à-dire constituées uniquement d'Hydrogène et d'Hélium. Ensuite, de plus grandes structures se seraient formées, les galaxies, que l'on observe actuellement à l'aide des grands observatoires terrestres et spatiaux. Notre galaxie, la Voie Lactée, s'est donc formée il y a très longtemps comme les étoiles qui la composent. Ces étoiles de population II et I, c'est-à-dire respectivement pauvres et riches en éléments lourds, peuvent être encore visibles dans des amas globulaires pour les moins massives d'entre elles. Cependant les étoiles de notre galaxie vivent et meurent afin de laisser place à de nouvelles générations d'étoiles et des éléments chimiques de plus en plus complexes et abondants, notamment nécessaires à la formation de planètes et de vies si les conditions sont réunies.

Pour comprendre l'origine de notre système solaire et de la vie sur Terre, il est prépondérant de répondre à la question : comment les étoiles se forment-elles et dans quel environnement ?

Notre existence est étroitement liée aux étoiles et notamment aux premières étapes de la formation stellaire. Nous savons depuis quelques dizaines d'années que l'espace entre les étoiles n'est pas vide. Au début du ^{xx}^e siècle siècle ([Hartmann, 1904](#)) fit la découverte de ce qu'on appelle aujourd'hui le Milieu Interstellaire. Celui-ci est composé de gaz et de poussières de différentes natures. Des travaux plus récents, c'est-à-dire dans la deuxième moitié du ^{xx}^e siècle siècle, nous ont montré que cette poussière était aussi présente dans notre système solaire et est responsable de ce qu'on appelle communément la lumière zodiacale. Brian May, ancien guitariste du groupe de rock anglais Queen, en a d'ailleurs étudié les propriétés dans sa thèse ([May, 2008](#)). Des études scientifiques importantes ont vu le jour afin de comprendre le lien entre cette poussière, la formation des étoiles, notre système solaire et la présence de vie dans l'Univers. Ces travaux étant d'une ampleur considérable, plusieurs domaines ont vu le jour afin de compartimenter les réflexions.

Mon sujet de thèse est ainsi à la rencontre de deux domaines : la formation stellaire et le milieu interstellaire. Notre galaxie est un vaste laboratoire d'études de cette formation et je me suis donc intéressé aux premières étapes de la formation des étoiles, allant du

nuage moléculaire à la proto-étoile. Nous savons depuis quelques années que les étoiles ont une tendance à se regrouper en amas, comme la plupart des objets astrophysiques évoqués en début d'introduction. C'est ainsi que ma thèse s'inscrit dans cette problématique à l'aide de nouvelles données, afin d'étudier les propriétés des régions de formation stellaire. En effet, j'ai principalement utilisé des données provenant du télescope Herschel qui nous fournit des images ainsi que des données dans l'infrarouge lointain et le domaine sub-millimétrique à une résolution angulaire inégalée, à des longueurs d'onde encore peu ou pas exploitées. Plus particulièrement, le pilier de ma thèse est l'utilisation de données provenant du projet Hi-GAL (Herschel infrared Galactic Plane Survey). Hi-GAL est un relevé de l'ensemble du plan galactique et un projet majeur du télescope Herschel. De nombreuses études ont utilisé et utilisent toujours ce relevé avec à la clé quelques découvertes majeures, dans un champ d'étude encore peu exploré. Il est ainsi intéressant de continuer dans cette lignée et d'avoir l'opportunité d'utiliser des données encore peu exploitées, les premières études datant de 2010 et seulement sur des régions réduites du plan galactique.

Je vais présenter mes travaux de thèse à travers ce manuscrit en commençant par une introduction sur le milieu interstellaire, ces principaux composants et les observations et découvertes majeures faites jusqu'à présent (chapitre 1). Je continuerai par une introduction sur la formation stellaire en me concentrant sur les toutes premières phase, avant que l'on classe ces objets en tant qu'étoile (chapitre 2). Les chapitres suivants seront consacrés à mes résultats. J'ai commencé mes travaux de recherche par l'élaboration d'un catalogue de sources jeunes à partir des catalogues Hi-GAL composés de sources extraites à partir des images à chaque longueur d'onde. La méthode ainsi que les résultats sont accessibles dans le chapitre 3. Ce catalogue m'a ensuite permis d'élaborer un catalogue d'amas de sources jeunes dans toute la partie interne de notre galaxie (chapitre 4). Par la suite, j'ai étudié les propriétés physiques et spatiales de ces amas et de leurs sources dans le chapitre 5.

Première partie

Contexte astrophysique

CHAPITRE

1

MILIEU INTERSTELLAIRE ET FORMATION STELLAIRE

Sommaire

1.1	Histoire du Milieu Interstellaire	4
1.2	Constitution du MIS	4
1.2.1	Rayons cosmiques	6
1.2.2	Champ magnétique	6
1.2.3	Abondances	6
1.2.4	Le gaz	7
1.2.4.1	Le gaz neutre	7
1.2.4.2	Le gaz moléculaire	8
1.2.4.3	Le gaz ionisé	9
1.2.5	La poussière	10
1.2.5.1	Rôles des poussières	10
1.2.5.2	Types de poussières	12
1.3	Cycle de vie et structure du Milieu Interstellaire	13
1.3.1	Cycle de vie	13
1.3.2	Des grandes structures aux étoiles	14
1.3.3	Régions associées à la formation stellaire	15
1.4	Les différents stades évolutifs des proto-étoiles.	17

Le but de ce chapitre est d'introduire et de décrire le Milieu Interstellaire. Je commence par faire un bref rappel de l'histoire de ce champ d'étude en rappelant les découvertes principales. Je continue en analysant le contenu du Milieu Interstellaire en différenciant les catégories d'éléments le composant, en me concentrant sur le gaz et la poussière. Je termine enfin par décrire sa structure des grandes échelles (Galaxie) aux échelles plus petites que sont les régions de formation stellaire.

1.1 Histoire du Milieu Interstellaire

Le milieu interstellaire, plus connu sous l'acronyme MIS ou ISM en anglais (pour *InterStellar Medium*) est un champ d'étude très vaste et pluridisciplinaire de l'astrophysique contemporaine.

La première observation de particules de matière entre les étoiles fut énoncé par [Hartmann \(1904\)](#) lors de l'observation des raies spectrales de δ *Orionis*. Il constata des raies fixes du Ca II qui ne pouvait être indéniablement du à l'étoile en mouvement. La seule proposition plausible était la présence de matière sur la ligne de visée entre la Terre et l'étoile.

1.2 Constitution du MIS

Ce bref historique nous amène à nous intéresser aux constituants principaux du MIS. Le MIS se regroupe en trois catégories : la matière ordinaire, les rayons cosmiques et le champ magnétique. De plus celui-ci baigne dans un champ de rayonnement composé de photons à toutes les énergies du spectre électromagnétique.

Pour rappel, la Galaxie est formée d'un disque en rotation, d'un bulbe et d'un halo. Le disque s'étend sur plus de 20 kpc. Le bulbe est visible dans la partie centrale de la Galaxie et est visible sur 2-3 kpc. Enfin le halo est très peu lumineux et suit le bulbe dans les parties externes de la Galaxie. La majeure partie du milieu interstellaire est observable dans le disque. L'ensemble de la masse de la Galaxie varie selon les considérations que l'on fait. Ainsi la masse est estimée à $M_G \approx 1.7 \cdot 10^{11} M_\odot$ dans un rayon de 20 kpc mais peu dépasser $10^{12} M_\odot$ en prenant en compte un rayon plus important ([McMillan, 2011](#)). Le milieu interstellaire possède une masse d'environ 5% de la masse totale des étoiles et 0.5% de la masse totale de la Galaxie. 70% de sa masse est due à l'hydrogène, 28% à l'hélium et les 2% restants sont constitués de divers éléments. Le milieu interstellaire peut être divisé en plusieurs composantes suivant leur densité et leur température. On notera notamment la distinction entre les milieux atomique, moléculaire et ionisé. Chacun de ces milieux possède des composantes à différentes températures. Ces composantes ne sont pas spatialement dissociées et se décomposent suivant les processus physiques en cours et le champ de rayonnement ambiant. La table 1.1 résume les diverses grandeurs de la densité volumique, température et masse de ces composantes. Il est à noter que ces valeurs ne représentent uniquement qu'un ordre de grandeur. Certaines masses sont d'ailleurs incertaines, comme par exemple pour le milieu moléculaire.

L'émission spectrale du MIS couvre tout les domaines de longueur d'onde. Il couvre en particulier les domaines de l'infrarouge, sub-millimétrique et radio comme observé sur la figure 1.1. D'autres composantes interviennent et perturbent la distribution spectrale, comme le fond diffus cosmologique (CMB), la lumière zodiacale, les étoiles ainsi que des

TABLE 1.1: Ordre de grandeur de la densité volumique, température et masse totale des différentes composantes du milieu interstellaire.

Milieu		Densité volumique (cm^{-3})	température (K)	Masse (M_{\odot})
atomique	Froid	20	≈ 100	$1.5 \cdot 10^9$
	Tiède	0.2	≈ 8000	$1.5 \cdot 10^9$
moléculaire		≥ 1000	≤ 100	10^9
Ionisé	Régions H II	$1 - 10^4$	≈ 10000	$5 \cdot 10^7$
Ionisé	Diffus	0.03	≈ 8000	$< 10^9$
Ionisé	Chaud	$6 \cdot 10^{-3}$	$\approx 5 \cdot 10^5$	10^8

atomes et molécules dans la haute atmosphère lorsque les observations se font depuis le sol. Cependant, dans l'infrarouge moyen, lointain et le sub-millimétrique, on observe les poussières du MIS que nous traiterons dans la section 1.2.5.

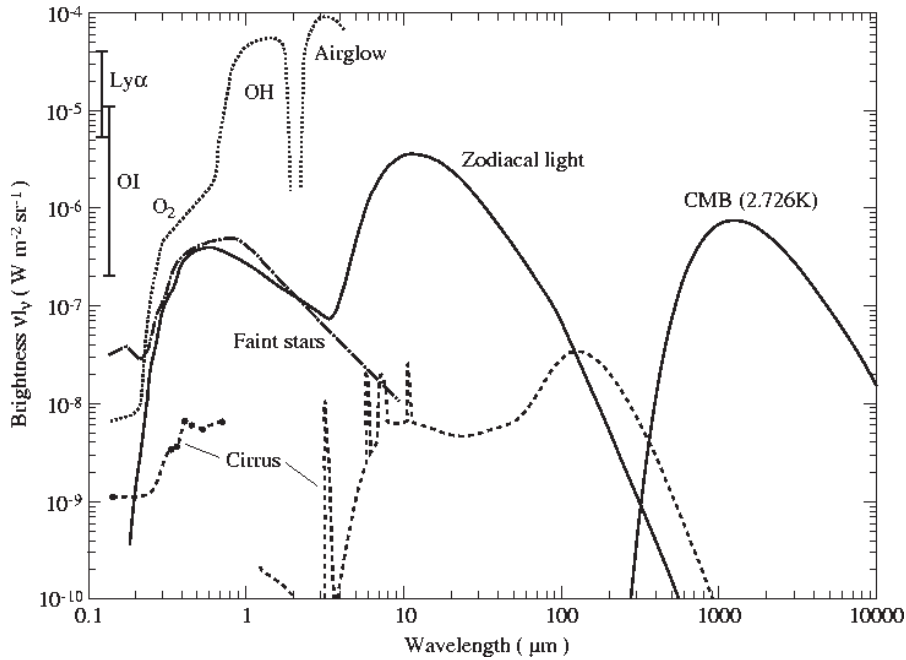


FIGURE 1.1: Distribution de la brillance du ciel aux latitudes galactiques élevées, de l'infrarouge proche au millimétrique. Les lignes pointillées (associées aux termes OH, O₂ et Airglow) correspondent aux émissions dans la haute atmosphère. L'oxygène atomique (OI) et la raie Lyman α ($\text{Ly}\alpha$) proviennent de la magnétosphère terrestre. Les étoiles (faint stars), la lumière zodiacale et le CMB participe tous au spectre observé de manière non négligeable. Les cirrus sont des nuages diffus, observés aux hautes latitudes. Enfin, la ligne en tiret correspond à l'émission des différents types de poussières. Extrait de [Leinert et al. \(1998\)](#).

1.2.1 Rayons cosmiques

Les rayons cosmiques sont composés de noyaux d'atomes, principalement des protons et des noyaux d'héliums, ainsi que d'électrons. Ceux-ci sont très énergétiques, pouvant atteindre des valeurs de 10^{20} eV, en raison de leur vitesse relativiste. Les rayons cosmiques parviennent jusqu'à la Terre et sont capables de créer des avalanches de particules jusqu'au sol, détectées sur plusieurs milliers de m^2 . Leur origine dépend de leur énergie. Ainsi, les rayons les plus énergétiques peuvent provenir du milieu intergalactique ainsi que de galaxies proches comme le Petit Nuage de Magellan. Cependant en dessous d'une certaine énergie, 10 GeV selon [Sreekumar et al. \(1993\)](#), ces rayons semblent être d'origine galactique.

Les rayons cosmiques constituent une source d'énergie importante pour le milieu interstellaire. Les interactions entre ces rayons et les noyaux du gaz interstellaire produisent des rayonnement gamma d'une énergie de l'ordre de la centaine de MeV. Ce rayonnement est produit par trois mécanismes : les interactions nucléaires, le rayonnement de freinage et l'effet Compton inverse. Ces rayons gamma permettent notamment de déterminer la masse du milieu interstellaire dans la Galaxie et notamment la masse moléculaire.

1.2.2 Champ magnétique

Comme pour les rayons cosmiques, le champ magnétique est une source d'énergie importante. Il est d'ailleurs intéressant de remarquer que la densité d'énergie est la même pour ces deux constituants du milieu interstellaire, avec une valeur de 1 eV cm^{-3} . C'est aussi la valeur que l'on retrouve pour le champ de rayonnement au voisinage solaire. Une étude approfondie de la dynamique du milieu interstellaire nécessite ainsi la prise en compte de ces phénomènes.

Le champ magnétique du milieu interstellaire est de l'ordre de $5 \mu G$ en moyenne mais augmente dans les régions denses, à environ $30 \mu G$, ainsi que pour la formation des étoiles. Il participe à l'effondrement et à la formation stellaire, des nuages moléculaires jusqu'aux étoiles de la PMS. À grande échelle, il s'oppose à l'attraction gravitationnelle de la matière du disque avec l'aide de la pression cinétique non thermique et des particules chargées.

Différentes mesures permettent de déterminer le champ magnétique, par effet Zeeman, rotation Faraday et à partir du rayonnement synchrotron galactique. L'analyse de la polarisation de la lumière permet de déterminer la direction du champ magnétique. [Heiles & Crutcher \(2005\)](#) ont ainsi déterminé la direction du champ magnétique dans la galaxie et ont montré qu'il était perpendiculaire au plan.

1.2.3 Abondances

Le Milieu Interstellaire est composé de gaz sous différentes formes, compositions et états. Comme nous l'avons vu plus haut, la majeure partie du gaz est sous forme atomique et neutre. Les éléments les plus abondants sont l'hydrogène (H), l'hélium (He), suivis d'éléments lourds comme le C, N, O et Fe à hauteur de 2%, pour l'ensemble des éléments lourds. Les éléments les plus abondants au voisinage solaire et ρ Ophiuchus sont réunis dans la table 1.2. ρ Ophiuchus est aussi choisi comme référence car il caractérise une zone dense.

TABLE 1.2: Abondance des principaux éléments du MIS. Les abondances solaires sont exprimées comme $12 + \log(X/H)$, où X est le nombre d'atomes de l'élément et H représente l'hydrogène. Pour ρ Ophiuchus, les valeurs représentent la différence avec les abondances solaires, $[X/H] = \log(X/H) - \log(X/H)_{\odot}$. Données de [Savage & Sembach \(1996\)](#).

Éléments	Système solaire	ρ Ophiuchus froid
H	12,00	—
He	10,99	—
C	8,55	-0.41
N	7,97	-0.07
O	8,87	-0.39
Fe	7,50	-2.27

Les abondances proviennent de l'analyse des raies d'absorption des éléments dans l'UV ainsi que, pour une partie d'entre eux, dans le spectre visible des étoiles. Contrairement à la phase gazeuse, les abondances des molécules sont moins bien connues car leur analyse est réalisée à partir des raies d'émission qui dépendent elles mêmes de l'excitation des molécules. Ces abondances peuvent néanmoins être calculées en comparant les valeurs d'abondance totales estimées (gaz et poussières) aux abondances du gaz. Les éléments lourds font à la fois partie du gaz et de la poussière. Plus particulièrement, dans des zones denses comme les nuages moléculaires, ces éléments s'assemblent pour former des molécules de tailles différentes (de la taille de quelques atomes jusqu'à des grains de $0.1 \mu m$).

1.2.4 Le gaz

Dans certains endroits, les éléments restent sous forme de gaz. Nous pouvons dissocier le gaz en trois catégories : le gaz neutre atomique, le gaz moléculaire et le gaz ionisé comme montré dans la table 1.1.

1.2.4.1 Le gaz neutre

Le gaz atomique est la composante la plus importante en masse du MIS. Il existe trois raies typiques permettant de caractériser ce gaz atomique : la raie à 21 cm de l'hydrogène atomique, les raies de structure fine et les raies d'absorption interstellaires.

Les raies de structure fine sont primordiales dans le MIS car elles permettent de refroidir le milieu. La structure fine est en fait le dédoublement d'un niveau d'énergie, généralement le niveau fondamental, en raison de l'interaction entre le moment orbital total des électrons et leur spin total. Les raies du C II à $157 \mu m$ et du O I à $63 \mu m$ sont d'ailleurs les deux principales raies permettant le refroidissement dans le MIS.

La raie à 21 cm est une des raies les plus connues de l'histoire de l'astronomie. Elle

permet notamment d'étudier la distribution et la masse de l'élément le plus abondant de la Galaxie avec son homologue moléculaire H_2 . Cette raie correspond à la transition hyperfine de l'état fondamental de l'atome d'hydrogène induit par l'inversion du spin de l'électron par rapport à celui du proton. Il est à noter que cette émission n'est pas fréquente. Cependant l'hydrogène atomique étant très abondant dans la Galaxie, il n'est pas difficile d'obtenir un signal suffisamment élevé pour les radiotélescopes.

L'hydrogène atomique est largement utilisée aujourd'hui pour tracer, notamment, le *warp* et le *Flare* de la Galaxie (Burton & Hartmann, 1988) mais aussi les structures à l'intérieur des bras spiraux. Le gaz HI est visible dans tout le disque galactique, même dans la partie externe de la galaxie située à plus de 15 kpc ce qui permet l'étude de la distorsion du plan galactique. Il se trouve aussi que plus de 60% de l'hydrogène atomique se retrouve dans les nuages chauds avec des températures supérieures à 500 K. Dans ces zones, la plupart de l'hydrogène reste sous forme atomique. L'explication est que l'hydrogène moléculaire formé par collision d'hydrogène atomique se désagrège par photo-dissociation causée par les photons UV du milieu chaud. Nous pouvons ainsi observer une balance entre l'hydrogène moléculaire et atomique.

Les troisièmes raies, d'absorption, trahissent la présence d'atomes lors de l'observation du spectre d'une étoile. Ces raies restent fixes dans le spectre de l'étoile contrairement aux raies de la photosphère de l'étoile en raison de sa vitesse radiale et ainsi de l'effet Doppler. Ces deux types de raies ne peuvent pas être confondues. Ces raies aident à comprendre les conditions physico-chimiques qui règnent dans le milieu diffus. Les éléments principalement observés sont Na, K, Ca, des ions Ca^+ , Ti^+ et des molécules CN et CH.

1.2.4.2 Le gaz moléculaire

Comme pour le gaz atomique, il existe plusieurs façons d'observer le gaz moléculaire. Nous allons simplement les énumérer ici mais en s'arrêtant sur les deux molécules les plus abondantes, le CO et H_2 . Les transitions électroniques sont l'une des particularités des molécules. Elles sont équivalentes aux transitions atomiques. La plupart des émissions de ces raies se retrouvent dans l'UV lointain. Il existe ensuite les transitions vibrationnelles résultant de modes de vibration de la molécule comme son nom l'indique. Comme les raies d'absorption interstellaires, les molécules possèdent des raies d'absorption, appelées bandes interstellaires diffuses. Celles-ci sont très faibles et se trouvent majoritairement dans le domaine visible.

Enfin, il existe une dernière sorte de raies, les transitions rotationnelles. Plusieurs molécules ont été observées possédant ces raies. Nous pouvons notamment citer les molécules CO, H_2 , NH_3 , H_2CO et H_2O . Nous observons ces raies dans les domaines sub-millimétrique, millimétrique et centimétrique.

Le CO est en fait la molécule la plus observée. La raison est que la molécule H_2 ne possède qu'une très faible signature spectroscopique. Ceci est due à son manque de moment dipolaire et ainsi aucunes transitions rotationnelles ne sont permises. Il existe quelques raies de H_2 dans l'UV, cependant l'extinction dans ces longueurs d'onde est importante. Il y a aussi des transitions rotationnelles et vibrationnelles dans l'infrarouge lorsque l'on regarde la composante chaude de H_2 , qui est localisée vers les régions de formation stellaire. Cependant pour évaluer le reste de H_2 , il est préférable d'utiliser le CO. Ainsi, en connaissant le rapport d'abondance entre le CO et H_2 , environ 10^{-4} , il est possible de

mesurer indirectement cette dernière molécule. Deux isotopes sont utilisés, ^{12}CO ($J=1-0$) et ^{13}CO ($J=1-0$).

Contrairement au HI, la molécule H_2 n'est localisée que vers le centre galactique. Il devient d'ailleurs difficile de l'observer après 12 kpc. La figure 1.2 montre une carte CO permettant de tracer la molécule H_2 (Dame et al., 2001). On remarque qu'au-delà de 90 degrés du centre galactique, la densité chute brutalement. Il en est de même pour H_2 .

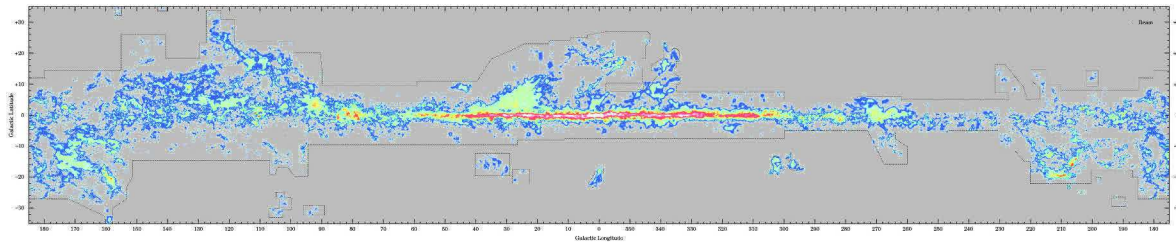


FIGURE 1.2: Carte CO de Dame et al. (2001) permettant de tracer la molécule H_2 .

Enfin, la poussière permet d'obtenir le facteur de conversion entre l'intensité du CO et la densité de colonne de H_2 . Ce rapport est important pour déterminer le rapport entre la poussière et le gaz. C'est actuellement la méthode la plus utilisée pour évaluer les différences de rapport au sein des différentes structures de la Galaxie, du fait que le gaz et la poussière n'ont pas été distribués de manière homogène.

Il a été montré que l'excès de couleur $E(B-V)$ est corrélé à la colonne de densité d'hydrogène N_H par la relation $E(B-V) \approx 1,7 \cdot 10^{-22} N_H \text{ mag.cm}^2$ (Rachford et al., 2002). En utilisant les relations de Kramers-kronig (Purcell, 1969), il est possible de lier cette corrélation au rapport poussière sur gaz.

1.2.4.3 Le gaz ionisé

Le gaz ionisé se retrouve partout où l'on observe des rayonnements UV, X, des particules chargées et des chocs. Il se retrouve notamment dans les régions H II, le milieu interstellaire chaud et le milieu ionisé diffus.

Aux alentours des régions H II, le gaz produit une émission continue en raison de plusieurs mécanismes. Tout d'abord l'émission libre-libre qui caractérise le freinage des électrons libres passant à proximité des ions. L'émission libre-lié est une autre émission quasi-continue, produite par la recombinaison des électrons et des ions. Celle-ci entraîne tout de même des discontinuités dans le spectre connue sous différents noms, Balmer (raies de recombinaison sur le niveau $n=2$), Paschen (raies de recombinaison sur le niveau $n=3$), etc...

Toujours autour des régions H II, il existe un autre type de raies, appelées raies de recombinaison. Ces raies correspondent aux désexcitations progressives qui suivent la recombinaison sur les niveaux élevés d'un atome, en général H, He, C, O, etc...

Nous avons vu que les régions H II ne forment qu'un milieu parmi d'autres zones ionisées. Cependant je ne détaille pas ces régions ici car ce qui nous intéresse dans cette thèse sont les zones de formation stellaire.

1.2.5 La poussière

1.2.5.1 Rôles des poussières

La poussière est un assemblage d'atomes allant de quelques nanomètres jusqu'à quelques micromètres. L'abondance des poussières est beaucoup moins importante que le gaz et elles ne représentent qu'à peu près un centième de la masse du milieu interstellaire compte tenu du rapport gaz sur poussière mentionné plus haut. Cependant elles jouent un rôle important dans la physique et la chimie du MIS. Les poussières jouent par exemple le rôle de catalyseur. Certains atomes et molécules, en se collant à la surface des grains, forment un manteau de glace ou se produisent des réactions chimiques. C'est dans ce contexte que se forme la molécule H_2 . Les grains se détruisent dans le MIS ce qui enrichit le milieu environnant. D'autres grains coagulent. Un modèle de coagulation est d'ailleurs présenté dans la figure 1.3.

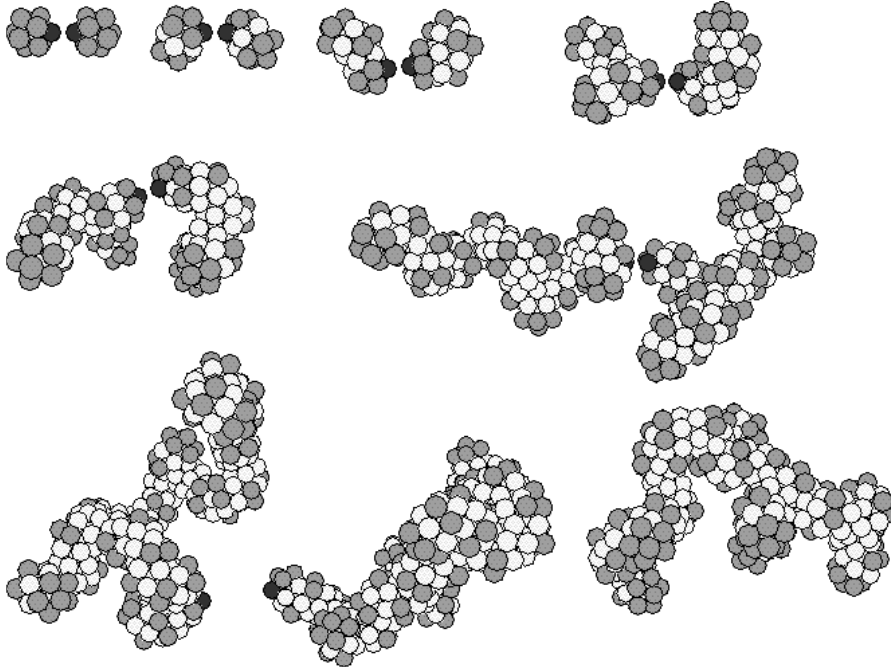


FIGURE 1.3: Représentation d'un modèle d'agglomération des grains de poussières. (credit : <http://www.astro.ucla.edu/~wright/dust/>)

Les poussières absorbent et diffusent aussi la lumière des étoiles. Ces deux phénomènes caractérisent l'extinction causée par les poussières. La lumière diffusée tend vers le bleu alors que celle qui passe à travers le nuage est plus rouge. L'extinction suit une loi particulière et dépend du modèle de grain considéré. Ainsi pour ajuster les observations, certains chercheurs ont développé des modèles de grains possédant une distribution de taille mais aussi de composition différente. Ainsi les grains peuvent être des grains silicates amorphes, carbonés ainsi que des grains mixtes composé d'un pourcentage de chacune des espèces. Mathis et al. (1977) utilise une distribution de taille :

$$dn_i = A_i n_H a^{-3.5} da \quad (1.1)$$

Où A_i est une constante de normalisation, n_H est la densité de noyaux d'hydrogène et dn_i est le nombre de grains d'espèce i dans un intervalle de tailles a et $a + da$. l'épaisseur

optique s'écrit ainsi :

$$\tau_\lambda = \sum_i \int_{a_{min}}^{a_{max}} Q_e(a_i, \lambda) \pi a_i^2 N_i(a_i) da_i \quad (1.2)$$

Où $N_i(a_i)$ est la densité de colonne des grains d'espèce i et de rayon a_i , $Q_e(a_i, \lambda)$ est l'efficacité d'extinction. Tant que les grains restent petits devant λ , τ_λ est proportionnelle à leur volume. On peut voir sur la figure 1.4, montrant l'extinction pour un modèle de [Draine & Lee \(1984\)](#), que l'extinction diffère suivant les composants mais que le modèle ajuste bien l'extinction galactique moyenne. Cependant les processus physiques aux petites échelles peuvent induire des différences de formation de grains, d'espèces et d'abondances de certains éléments. Ainsi le modèle doit être ajusté à chaque zone, avec des tailles de grains, des densités de colonne et des espèces différentes.

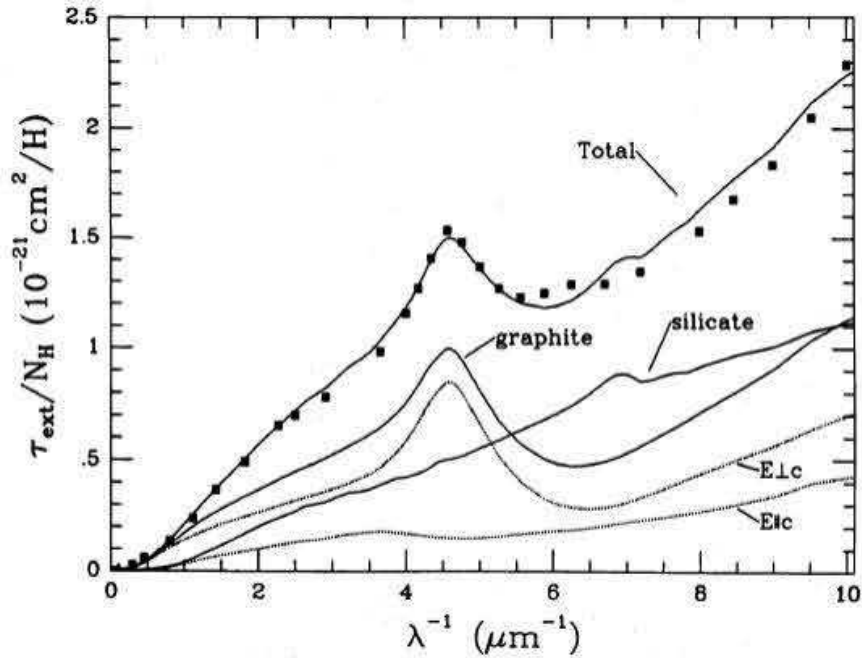


FIGURE 1.4: Extinction par atome d'hydrogène calculée pour le modèle de [Draine & Lee \(1984\)](#) en comparaison avec les observations de l'extinction galactique moyenne. Extrait de [Draine & Lee \(1984\)](#).

Enfin, l'énergie absorbée par les poussières, issue des rayonnements UV, visible et infrarouge proche, est réémise dans des longueurs d'onde plus grandes, allant de l'infrarouge moyen au sub-millimétrique. À peu près 50% de l'énergie émise par les étoiles est absorbée tandis que l'autre moitié est diffusée ou alors n'est pas affectée.

L'émission dépend aussi de plusieurs facteurs :

$$E_{em} = \int_0^\infty 4\pi a^2 Q_a(\nu) \pi B_\nu(T) d\nu \quad (1.3)$$

Où B_ν est la fonction du corps noir. $4\pi a^2$ est la surface d'émission, $Q_a(\nu)$ est l'efficacité d'émission et $\pi B_\nu(T)$ est la puissance émise par unité de fréquence dans le demi-espace par un corps noir de température T .

L'émission se caractérise par plusieurs profils différents dans le spectre de puissance, notamment entre 1 et 1000 μm comme nous pouvons le voir sur la figure 1.5, traduisant l'émission des poussières suivant un modèle de grains. On remarque que le spectre est composé de trois composantes que nous allons détailler dans la section suivante.

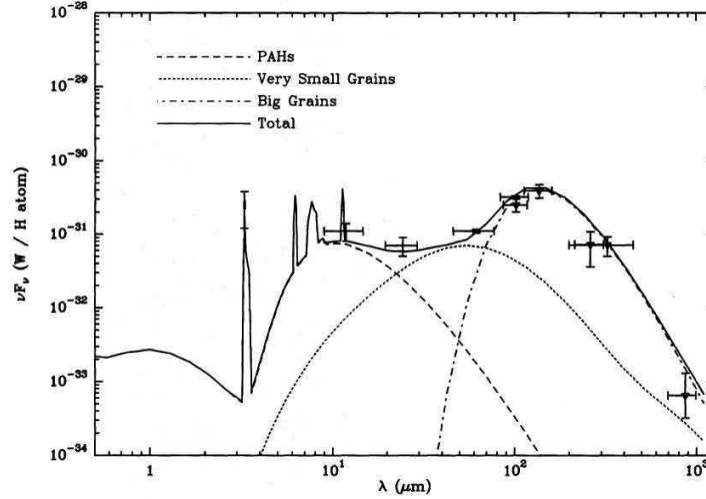


FIGURE 1.5: Modèle d'émission des grains de poussières composés de plusieurs espèces, confronté aux observations de cirrus. Extrait de [Desert et al. \(1990\)](#).

1.2.5.2 Types de poussières

L'émission des poussières est caractérisée par l'équation 1.5. Celle-ci, ainsi que les modèles, nous indique que les grains doivent suivre une certaine distribution de taille mais aussi de composition afin d'ajuster les données.

Actuellement les grains se différencient en 4 catégories principales : les PAHs (Polycyclic Aromatic Hydrocarbon), les SGs (Small Grains), les VSGs (Very Small Grains) et les BGs (Big Grains). Les SGs ne sont pas présent dans le modèle de [Desert et al. \(1990\)](#) comme le montre la figure 1.5.

Les petits grains (SGs) sont des grains hors équilibre thermique. Ces grains subissent des variations importantes de leur température dues à l'absorption d'un photon dans l'UV ou le visible et une réémission aussi rapide et donc un refroidissement quasi-instantané. Ce genre d'émission est appelé chauffage stochastique. Le spectre émis par ce genre de grain se calcule en connaissant la probabilité $P(a, T)$ qu'un grain de rayon a ait une température comprise entre T et $T + dT$ ([Draine & Anderson, 1985](#)) :

$$I_\nu(a) = \int_0^\infty Q_a(a, \nu) B_\nu(T) P(a, T) dT \quad (1.4)$$

La distribution ne dépend pas du champ de rayonnement externe mais de la température atteinte après l'absorption d'un photon. Cependant elle dépend de la nature des grains et de l'énergie du photon absorbé.

Les PAHs caractérisent l'émission des poussières à certaines bandes spectrales, 3,3, 6,2, 7,7, 8,6, 11,3 et 12,7 μm . Leur origine est restée longtemps inconnue, depuis leur découverte dans la nébuleuse NGC 7027 ([Gillett et al., 1973](#)), plusieurs noms leur furent

attribuées comme UIB (Unidentified Infrared Bands), IEF (Infrared Emission Features) ou encore AIB (Aromatic Infrared Bands). C'est [Allamandola et al. \(1989\)](#) qui proposa que ces PAHs sont formés par des très petits grains formés de cycles aromatiques ou de grosses molécules chauffées par l'absorption de photons UV.

Les gros grains sont responsables de l'émission au-delà de $60 \mu m$ comme nous pouvons le voir sur la figure 1.5. Ces grains à l'équilibre ont une émission caractérisée par un corps noir modifié :

$$\lambda I_\lambda = k_{ref} \left(\frac{\lambda}{\lambda_{ref}} \right)^\beta \lambda B_\nu(T_{eq}) \quad (1.5)$$

Où β représente l'indice spectral caractérisant une modification de la fonction du corps noir. La loi d'opacité k_ν suit une loi de puissance $k_\nu = k_{ref}(\lambda/\lambda_{ref})^\beta$.

Ces grains contiennent une grande partie de la matière solide. L'interaction entre ces grains et des plus petits grains rend compte du bilan énergétique du MIS.

Enfin, les très petits grains possèdent un statut ambigu car ils font la liaison entre les PAHs et les BGs, c'est à dire dans l'infrarouge moyen. Ces grains sont dans un régime intermédiaire entre un chauffage stochastique et un équilibre.

1.3 Cycle de vie et structure du Milieu Interstellaire

1.3.1 Cycle de vie

Les différents types de gaz et de poussières ne sont pas dispersés uniformément dans le Milieu Interstellaire. Plus précisément, certains types de gaz et poussières sont associés à des régions spécifiques de ce milieu. Afin d'introduire ces différents types nous allons commencer par introduire le cycle de vie du Milieu Interstellaire. La figure 1.6 montre les principales étapes de ce cycle. Tout commence par une nuage HI géant et diffus qui commence à se fragmenter pour former des structures de plus en plus petites afin d'arriver jusqu'à la formation d'étoiles. À partir de là deux scénarios sont possibles. Premièrement les étoiles de faibles masses évoluent lentement jusqu'à la séquence principale, en plusieurs millions d'années. Ces étoiles ressemblent au Soleil et vont rester plusieurs milliards d'années sur la séquence principale avant de céder une grande partie de leur enveloppe au changement de fusions nucléaires. A la fin il ne restera qu'une naine blanche qui s'éteindra petit à petit. Cependant une nébuleuse planétaire y sera associé et viendra enrichir le MIS. Ce procédé est beaucoup plus rapide et efficace lorsque l'on parle d'étoiles massives. En effet, dans ce cas les étoiles ne restent que quelques centaines de millions d'années sur la séquence principale avant de former des éléments de plus en plus lourds, avant d'exploser en supernovae et souffler le gaz et la poussière autour d'elle. C'est dans ce genre d'environnement que se forme la poussière, ou la densité est suffisamment importante pour favoriser la coagulation des atomes.

Au final tous ces éléments se retrouvent mélangés dans le MIS pour un nouveau cycle.

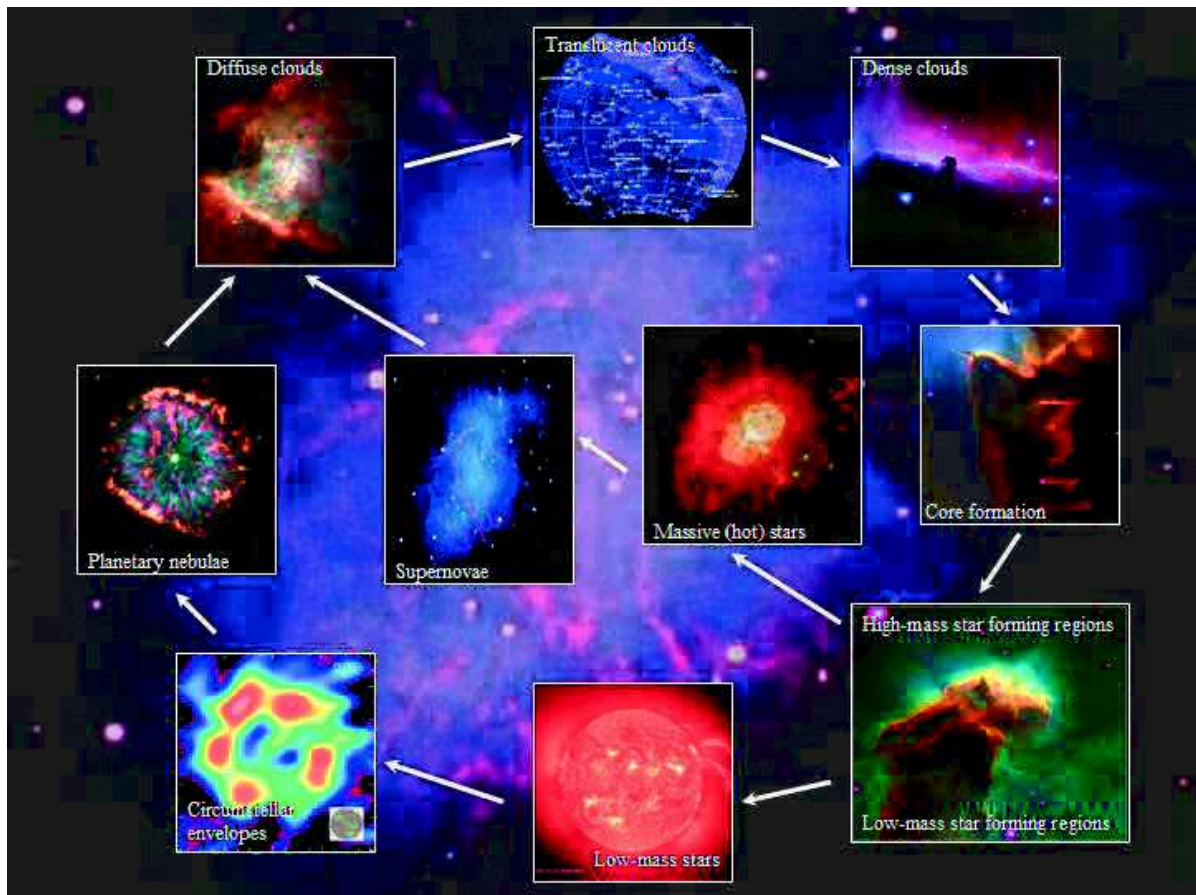


FIGURE 1.6: Cycle de vie du MIS (crédit Kaiser, 2002)

1.3.2 Des grandes structures aux étoiles

Le cycle décrit les étapes de l'enrichissement du MIS en éléments lourds. Cependant, ces étapes ont lieu à des échelles différentes qui possèdent chacune leur particularité. Ce processus qui permet de passer des grandes structures aux étoiles s'appelle la *fragmentation*. Cette fragmentation résulte d'une instabilité gravitationnelle et est liée à d'autres facteurs comme le champ magnétique, la turbulence ou encore la pression thermique.

Tout commence à l'échelle galactique avec l'apparition des bras spiraux comme représenté dans la figure 1.7. L'origine et le maintien de ces bras fait office de nombreux travaux et ne sont toujours pas élucidés. C'est principalement à l'intérieur de ces bras que prend place la formation stellaire car ces bras sont en fait des zones de sur-densité de gaz et/ou d'étoiles. Il existe toutefois des régions de formation stellaire dans le milieu interbras. Cependant les temps de fragmentation, en vue d'obtenir des étoiles à partir de grandes structures, sont beaucoup plus lents en raison de la densité plus faible. A l'intérieur de ces bras des structures de plus en plus fines se forment. Cela commence par des nuages d'une taille de l'ordre du kpc, appelé super nuage HI. Ceux-ci s'effondrent en un laps de temps relativement long comparé à la formation des structures suivantes. En fait, plus les tailles diminuent et plus les temps de formation deviennent petits. Efremov & Elmegreen (1998) présentent une étude de ces différentes structures et montrent leur temps de formation, visible dans la figure 1.8 extraite de cette étude. La suite de la fragmentation se passe à l'échelle de la centaine de parsecs avec l'apparition des nuages moléculaires géants, appelés GMC, sur une échelle de temps d'une dizaine de millions

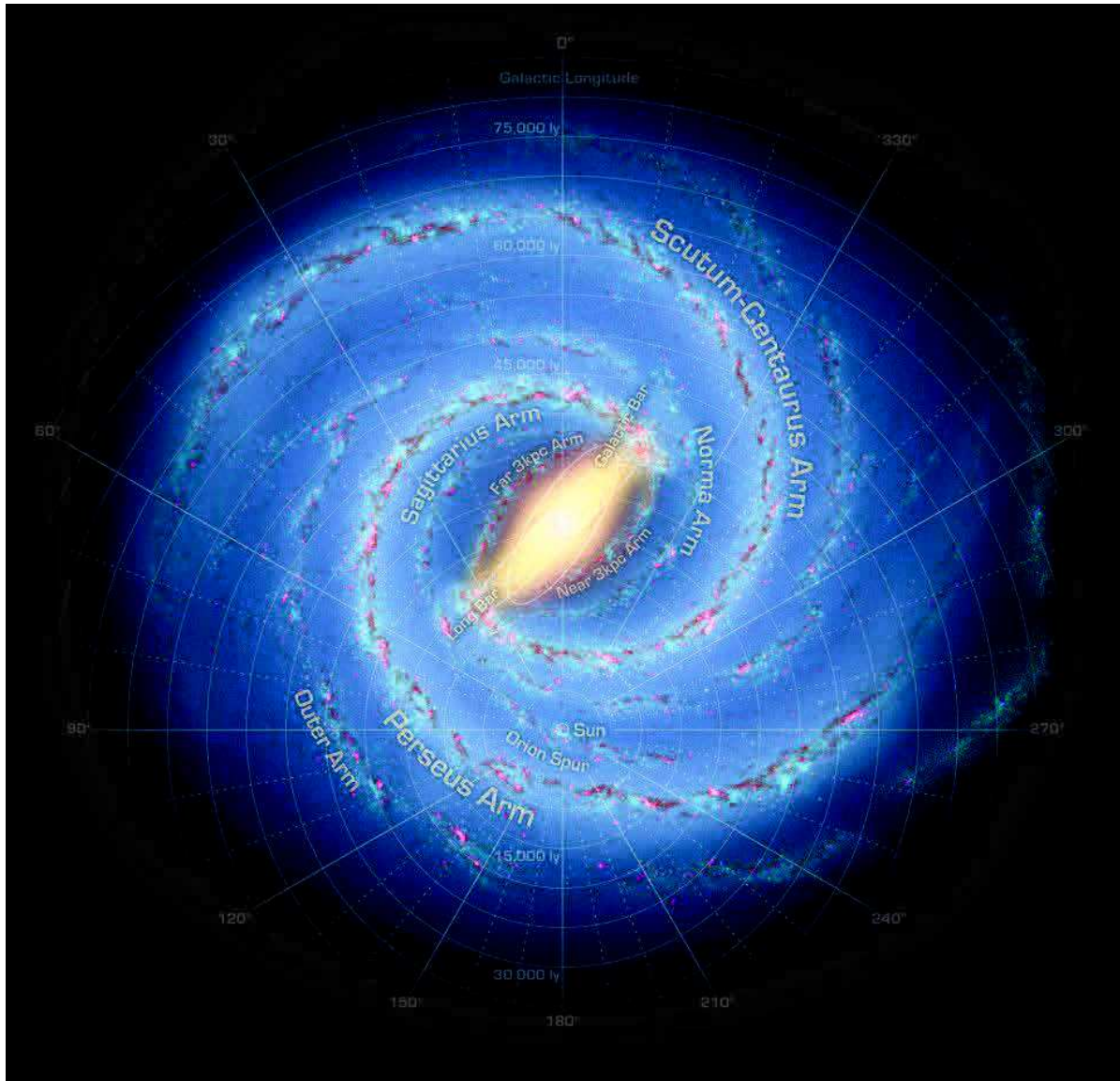


FIGURE 1.7: Vue d'artiste de la Voie Lactée et de ses bras spiraux vu par un observateur extérieur. Credit : NASA/Adler/U. Chicago/Wesleyan/JPL-Caltech

d'années. C'est à cette échelle que commence à se former les premières molécules. Les densités sont par exemple suffisante pour former de l'hydrogène moléculaire, à partir des grains de poussières. Les GMC donnent ensuite naissance à des nuages moléculaires de la taille du parsec voir de la dizaine de parsecs. Ces derniers sont supposés être stable en raison des forces internes qui agitent les particules et du champ magnétique. Cependant une perturbation peut se produire et ces nuages moléculaires se fragmentent pour laisser place à des coeurs moléculaires plus denses et plus froids. La table 1.3 comprend les densités et températures typiques de ces régions. Enfin les coeurs moléculaires s'effondrent et forment les proto-étoiles qui commencent à chauffer le milieu environnant.

1.3.3 Régions associées à la formation stellaire

Les nuages moléculaires proviennent de la fragmentation de super amas HI comme nous l'avons vu précédemment. Néanmoins, ils peuvent aussi provenir de la collision

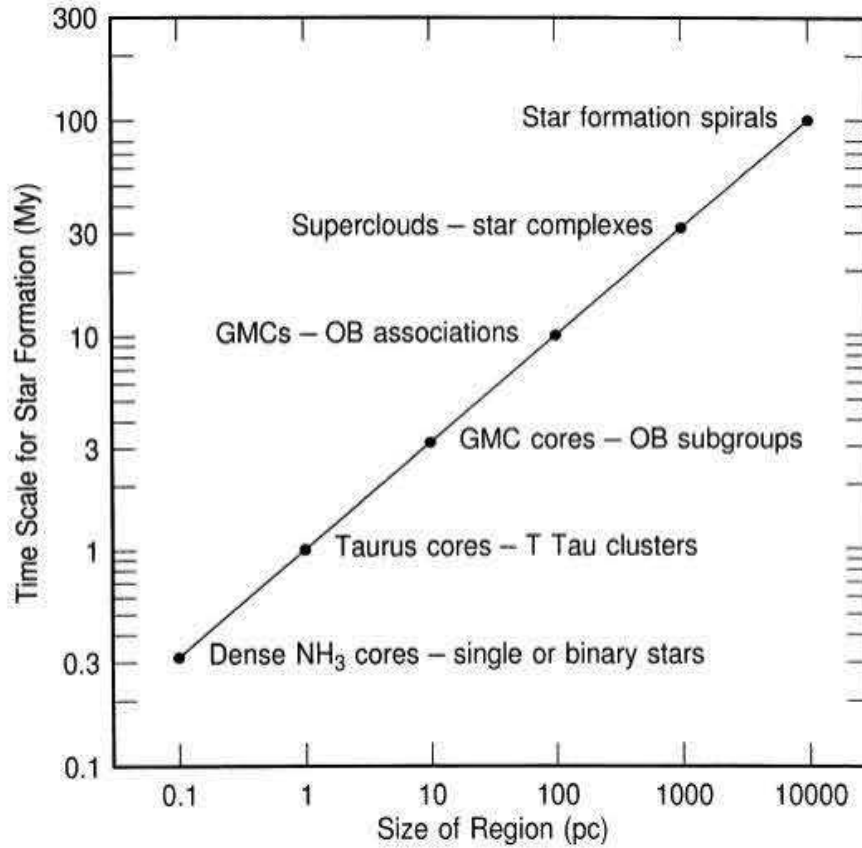


FIGURE 1.8: relation entre les tailles des différentes structures, allant des bras spiraux aux étoiles, et leur temps de formation. Extrait de (Efremov & Elmegreen, 1998).

de ces super amas. On observe généralement ces nuages dans l'infrarouge, le sub-millimétrique et le domaine radio. Leur temps de formation caractéristique est d'environ 10^7 Yr et leur masse peut aller de quelques centaines de masses solaires à plusieurs dizaines de milliers pour les nuages moléculaires géants (GMC), pour des températures inférieures à 100 K et des tailles inférieures à la dizaine de parsecs. Il existe un lien entre ces nuages moléculaires et les régions H II que l'on évoque dans la section suivante.

Les étoiles massives et chaudes, de type OB, évoluent rapidement et ne nécessitent que quelques millions d'années pour atteindre la fin de la séquence principale. Ces étoiles émettent un rayonnement très intense depuis leur plus jeune âge. Le rayonnement ayant une énergie supérieure à 13.6 eV, correspondant à une émission dans l'UV, ionise les atomes d'hydrogène présents aux alentours de l'étoile. Ainsi, les régions H II se forment à partir d'une première génération d'étoile.

TABLE 1.3: Grandeurs caractéristiques des composantes du milieu interstellaire froid.

	Densité volumique (cm^{-3})	température (K)	Taille (pc)
Milieu atomique froid	30	100	10 – 100
Nuages moléculaires	200	40	1 – 10
Coeurs moléculaires	10^4	10	0.1

Les régions H II sont souvent associées avec les nuages moléculaires. Non seulement elles ionisent leur environnement mais en plus elles érodent les structures autour d'elles à travers des phénomènes de photo-évaporation et de photo-ionisation. Ainsi une région H II peut désagréger entièrement un nuage moléculaire et ainsi stopper les processus de formation stellaire. Cependant cela n'affecte quand majeure partie les nuages moléculaires les plus massifs, avec des masses supérieures à $10^5 M_{\odot}$. En dessous de cette masse, il est peu probable que ces nuages forment des étoiles suffisamment massives pour ioniser leur environnement (Williams & McKee, 1997).

Dans d'autres cas, l'expansion des régions H II s'accompagne de chocs qui perturbent les régions encore sous forme atomique. Ces zones, encore sous forme atomique, sont appelées régions de photo-dissociation (PDRs). Elles se manifestent à la rencontre d'un milieu chaud et ionisé avec un milieu plus dense, froid et sous forme atomique et moléculaire. La figure 1.9 montre plusieurs régions H II avec leurs PDRs (Deharveng et al., 2005). Le chauffage et la chimie de ces zones sont régulés par les photons UV.

Les IRDCs, pour Infrared Dark Cloud, caractérisent des régions denses et froides du MIS. Les premières observations ont été faites avec les instruments ISO (Infrared Space Observatory) et MSX (Midcourse Space Experiment) (Egan et al., 1998). La figure 1.10 montre l'IRDC du serpent (Wang et al., 2014). Des études récentes ont montré, utilisant les données Spitzer à $24 \mu m$, que certains IRDCs sont associés avec des sources brillantes à cette longueur d'onde.

Ces IRDCs sont en général associés avec des régions de formation d'étoiles massives.

L'arrivée d'Herschel a permis l'obtention de nouvelles données. L'une des observations majeures est l'identification de filaments (André et al., 2010; Arzoumanian et al., 2013, 2011) comme le montre la figure 1.11. Ces régions sont proches, $\approx 400 pc$ pour la région Serpens, et n'appartiennent pas au plan galactique ce qui rend l'identification de ces structures plus aisée. André et al. (2010) a montré que ces filaments avaient une taille caractéristique de $0.1 pc$ et étaient associés avec les premières étapes de la formation d'étoiles. Molinari et al. (2010a) montrent que les sources Herschel sont associées avec un grand nombre de filaments.

Enfin la figure 1.12 représente les régions de formation stellaire les plus connues du plan galactique, superposées à la carte CO de Dame et al. (2001). Ces régions possèdent des sources évoluées comme des sources jeunes. Elles permettent ainsi d'explorer un panel complet d'âges et de masses.

1.4 Les différents stades évolutifs des proto-étoiles.

Les proto-étoiles, appelé YSOs (Young Stellar Objects), caractérisent le début de la formation d'un objet central compact au sein d'une sur-densité de gaz et de poussières. Ces proto-étoiles poursuivent le stade des coeurs pré-stellaires en continuant l'effondrement gravitationnel et l'accrétion de matière. Cette phase peut durer plusieurs millions d'années dans le cas d'un objet de faible masse (avoisinant la masse du Soleil) mais peut passer inaperçu dans le cadre d'étoiles massives. C'est durant cette phase que le coeur commence à chauffer la matière l'entourant. Cette phase est en fait subdivisée en plusieurs stades évolutifs, que l'on nomme classes ici, qui peuvent être différenciées par la

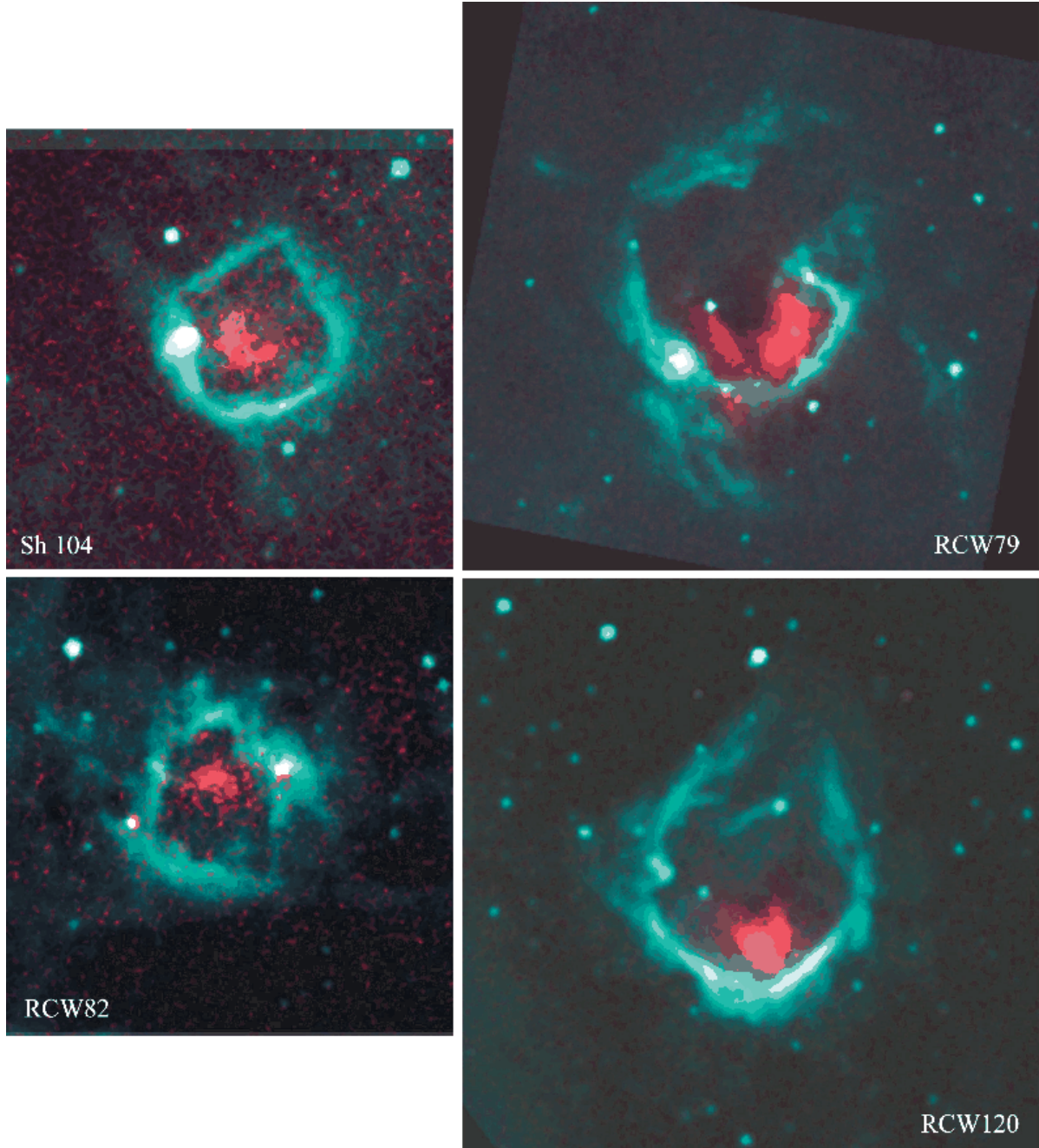


FIGURE 1.9: Exemple de régions H II avec des PDRs visibles et quasi-sphériques, provenant de [Deharveng et al. \(2005\)](#). Composition de couleurs d'images du télescope MSX. La couleur turquoise correspond à l'émission des PAHs dans la bande A ($8.3 \mu m$). La couleur rouge correspond à l'émission des VSGs dans la bande E ($21.3 \mu m$).

SED de la source.

Les YSOs sont divisés en quatre classes pour les étoiles de faibles masses : les classes 0, les classes I, les classes II et les classes III comme nous pouvons le voir dans la table 1.13 issue de [Schulz \(2012\)](#). La classe 0 se caractérise par une émission dans l'infrarouge moyen et lointain. Il n'y a encore aucune émission de la source centrale dans les domaines de l'infrarouge proche car on observe seulement l'enveloppe de la source. L'émission est entièrement déterminée par les gros grains. L'enveloppe a une masse de 0.14 à $2.8 M_{\odot}$.

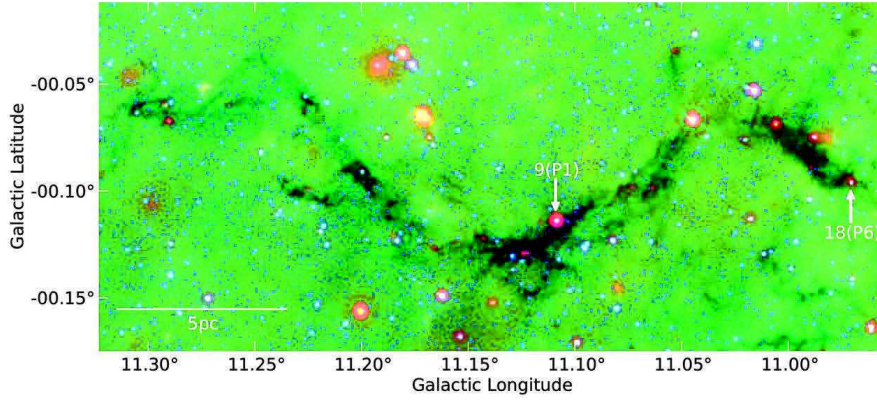


FIGURE 1.10: Image composite du filament *snake* observé avec Spitzer, à $4.5 \mu m$ (bleu), $8 \mu m$ (vert) et $24 \mu m$ (rouge). Les deux flèches localisent des sources infrarouges identifiées en proto-étoiles. Extrait de Wang et al. (2014)

La classe I voit son enveloppe devenir plus fine et commence à laisser place à des émissions de la source centrale. Cependant l'émission est encore dominée par les gros grains. L'enveloppe a une masse de 0.015 à $0.15 M_{\odot}$. Un champ magnétique de plus en plus fort s'installe.

Ensuite, lors des deux dernières étapes, classes II et III, le flux de la source devient de plus en plus important et l'objet se rapproche de plus en plus d'une étoile possédant un disque protoplanétaire. L'émission n'est plus dominée par les gros grains comme nous pouvons le voir sur la figure 1.14. Ces étoiles de faibles masses finissent en étoile T-Tauri dans leur classes II et III avant de passer à la PMS (Pre-Main Sequence) et ensuite la séquence principale.

Dans le cas d'objets de masses intermédiaires, les étoiles deviennent des étoiles de Herbig Ae/Be, pour une masse comprise entre 2 et $8 M_{\odot}$. Pour les étoiles encore plus massives, l'accrétion et l'évolution se fait très rapidement et les étapes citées précédemment n'ont pas lieu d'être.

Ces descriptions m'amènent vers l'étude menée lors de mon doctorat. Celle-ci a débuté avec la construction d'un catalogue de sources jeunes comme nous allons le voir de le chapitre suivant.

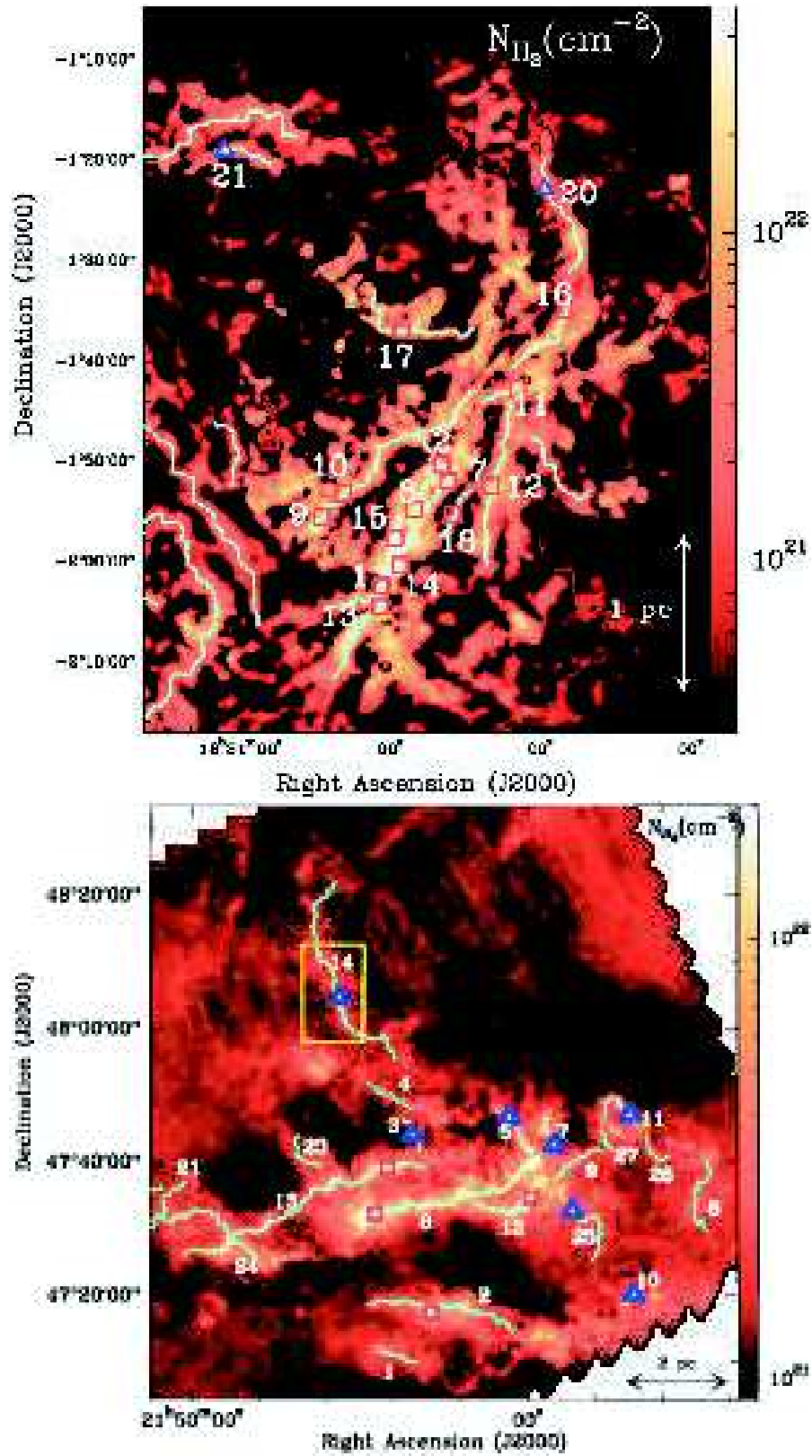


FIGURE 1.11: Haut : Filaments de la partie sud de la region Serpens (André et al., 2010).
Bas : Filaments dans IC 5146 (Arzoumanian et al., 2013, 2011).

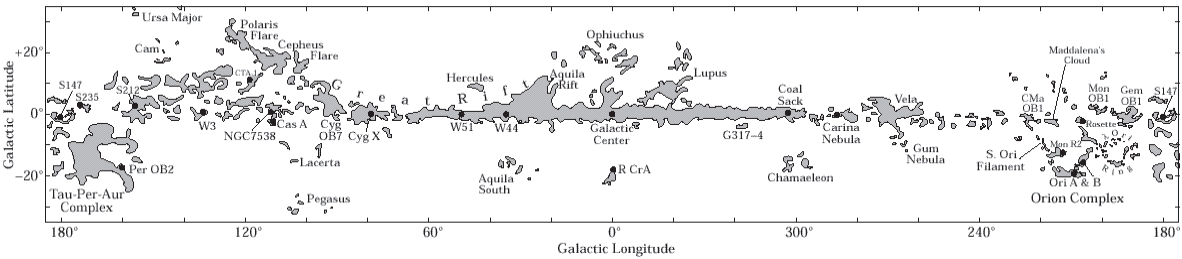


FIGURE 1.12: Principales régions de formation stellaire dans le plan galactique. L'image de fond correspond à la distribution du gaz CO (Dame et al., 2001).

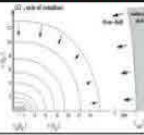
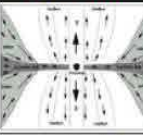
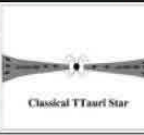

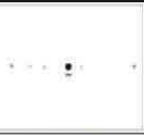
	Infalling protostar	Accreting protostar	Contracting PMS star		MS star
YSO properties					
Phase	adiabatic (A,B,C)	accretion (D) deuterium burning onset of convection	convective radiative onset of nuclear burning		convective radiative full nuclear burning
Matter flows	mostly infall disk & outflows form	some infall mostly accretion outflows, jets	low accretion	?	—
Envelope/ disk size	< 10000 AU	< 1000 AU	< 400 AU	~ 100 AU	—
Infall/ accretion rate	10^{-4}	10^{-5}	$10^{-6} \text{ -- } 10^{-7}$?	—
Age	$10^4 \text{ -- } 10^5 \text{ yr}$	10^5 yr	$10^6 \text{ -- } 10^7 \text{ yr}$	$10^6 \text{ -- } 10^7 \text{ yr}$	—
Emission bands (except IR)	thermal radio X-ray?	radio X-ray	radio optical strong X-ray	non-therm. radio optical strong X-ray	non-therm. radio optical X-ray
Classes	Class 0	Class I	Class II	Class III	ZAMS

FIGURE 1.13: SED des proto-étoiles à leurs différents stades évolutifs. De gauche à droite : Classe 0, Classe I, Classe II, Classe III et la ZAMS (crédit Schulz, 2012)

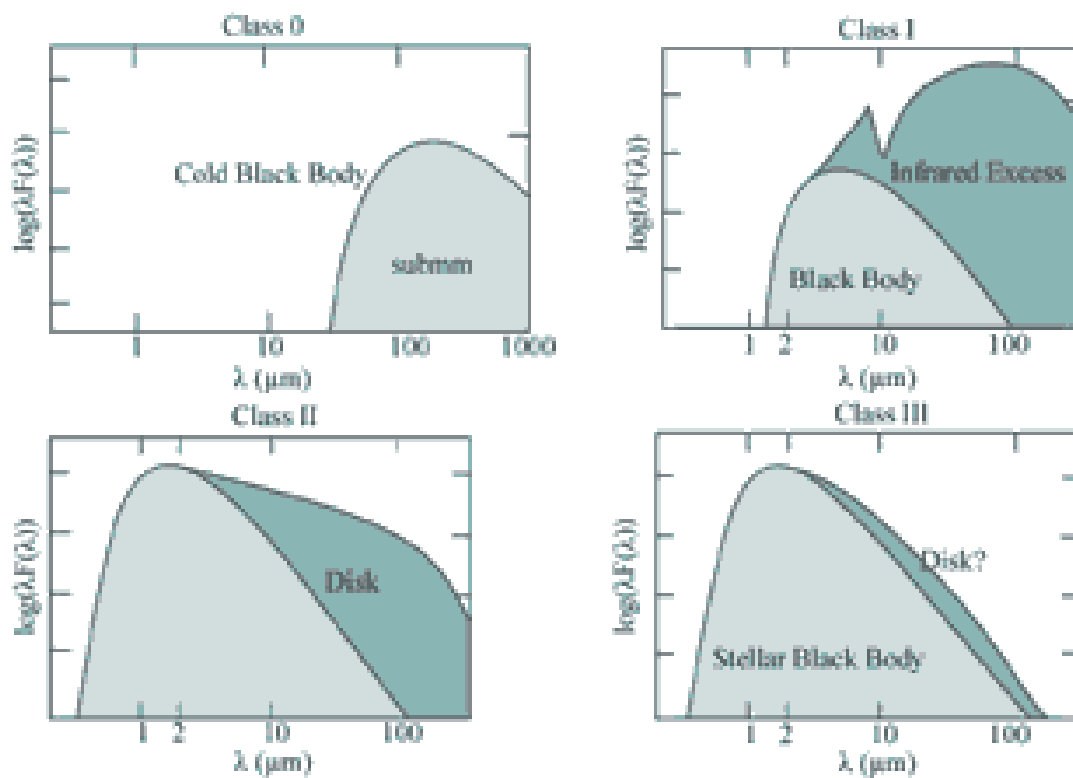


FIGURE 1.14: SED des proto-étoiles à leurs différents stades évolutifs. Haut-gauche : Classe 0, Haut-droite : Classe I, bas-gauche : Classe II, bas-droite : Classe III. (Crédit : <http://newt.phys.unsw.edu.au/jacara/pilotscience.php>)

Deuxième partie

Travaux et résultats

CHAPITRE

2

CATALOGUE MULTI-LONGUEURS D'ONDE

Sommaire

2.1	Observations dans les domaines infrarouge et sub-millimétrique	26
2.1.1	Télescopes antérieurs à Herschel	26
2.1.2	Herschel	28
2.1.3	Relevé Hi-GAL	29
2.2	Les données et outils utilisés	30
2.2.1	Catalogues single-band Hi-GAL	30
2.2.2	Specfind	31
2.3	Catalogue multi-longueurs d'onde	32
2.3.1	Conditions d'ajustement	32
2.3.2	Résultats et catalogue final	33
2.4	Catalogue de l'équipe Hi-GAL	37
2.4.1	Contenu	37
2.4.2	Notion de <i>clumps</i>	37
2.4.3	Comparaison et Discussion	38
2.4.4	Remarques	42

J'ai commencé mon doctorat en élaborant un catalogue d'objets stellaires jeunes à partir des catalogues *single-bands* provenant de l'équipe du relevé Hi-GAL. Je commencerai par des généralités sur les télescopes précédents avant d'introduire le télescope spatial Herschel et le relevé Hi-GAL. Je continuerai par introduire les données utilisées pour ensuite évoquer mon travail en commençant par la réalisation du catalogue. Je continuerai par une discussion sur les particularités de ce catalogue, pour finir par une confrontation avec un autre catalogue provenant de l'équipe Hi-GAL et le lien avec la suite de mon projet de thèse.

2.1 Observations dans les domaines infrarouge et sub-millimétrique

2.1.1 Télescopes antérieurs à Herschel

Comme nous l'avons évoqué dans le chapitre précédent, les objets stellaires jeunes sont principalement observables dans le domaine infrarouge et sub-millimétrique. En particulier l'infrarouge lointain et le domaine sub-millimétrique sont les longueurs d'ondes prédisposées à l'étude des sources les plus jeunes, c'est-à-dire des coeurs pré-stellaires aux proto-étoiles de classes 0/I.

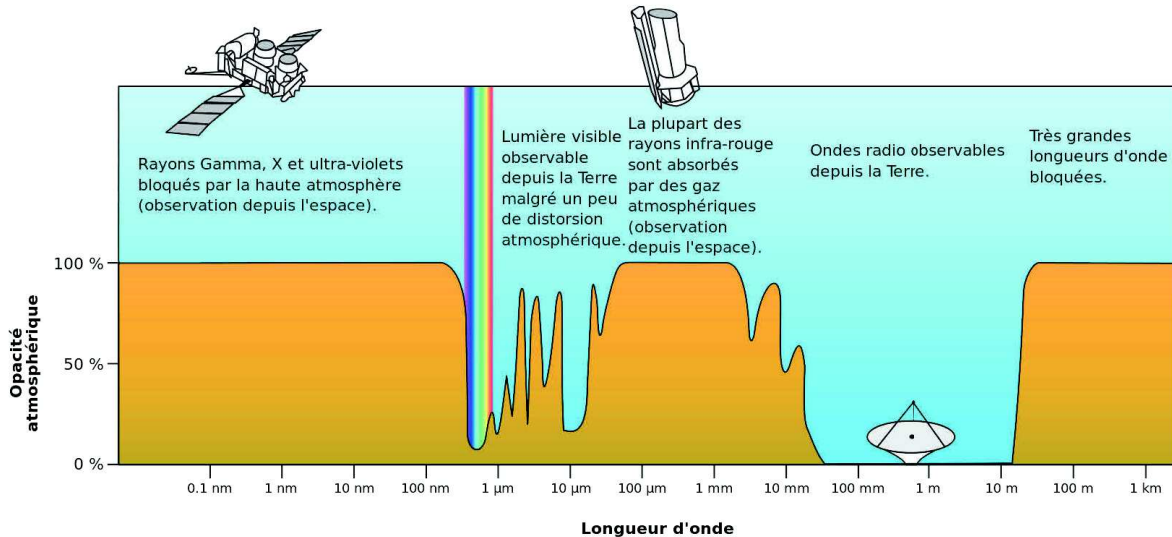


FIGURE 2.1: Transmission de l'atmosphère terrestre en fonction des domaines du spectre électromagnétique. La transmission est quasi nulle dans l'infrarouge et le sub-millimétrique, à l'exception de quelques bandes dans l'infrarouge proche. (Crédit : https://fr.wikipedia.org/wiki/Atmosph%C3%A8re_terrestre)

Il se trouve que dans ces longueurs d'onde l'atmosphère terrestre absorbe la quasi-totalité de la lumière reçue comme le montre la figure 2.1. Il faut ainsi utiliser principalement des télescopes spatiaux mis en orbite autour de la Terre. De multiples télescopes spatiaux possèdent des intervalles de longueur d'onde couvrant l'infrarouge proche et lointain mais peu couvrent à la fois l'infrarouge lointain et le sub-millimétrique. On peut

notamment citer COBE (Cosmic Background Explorer) et IRTS (Infrared Telescope in Space) comme indiqué sur la figure 2.2.

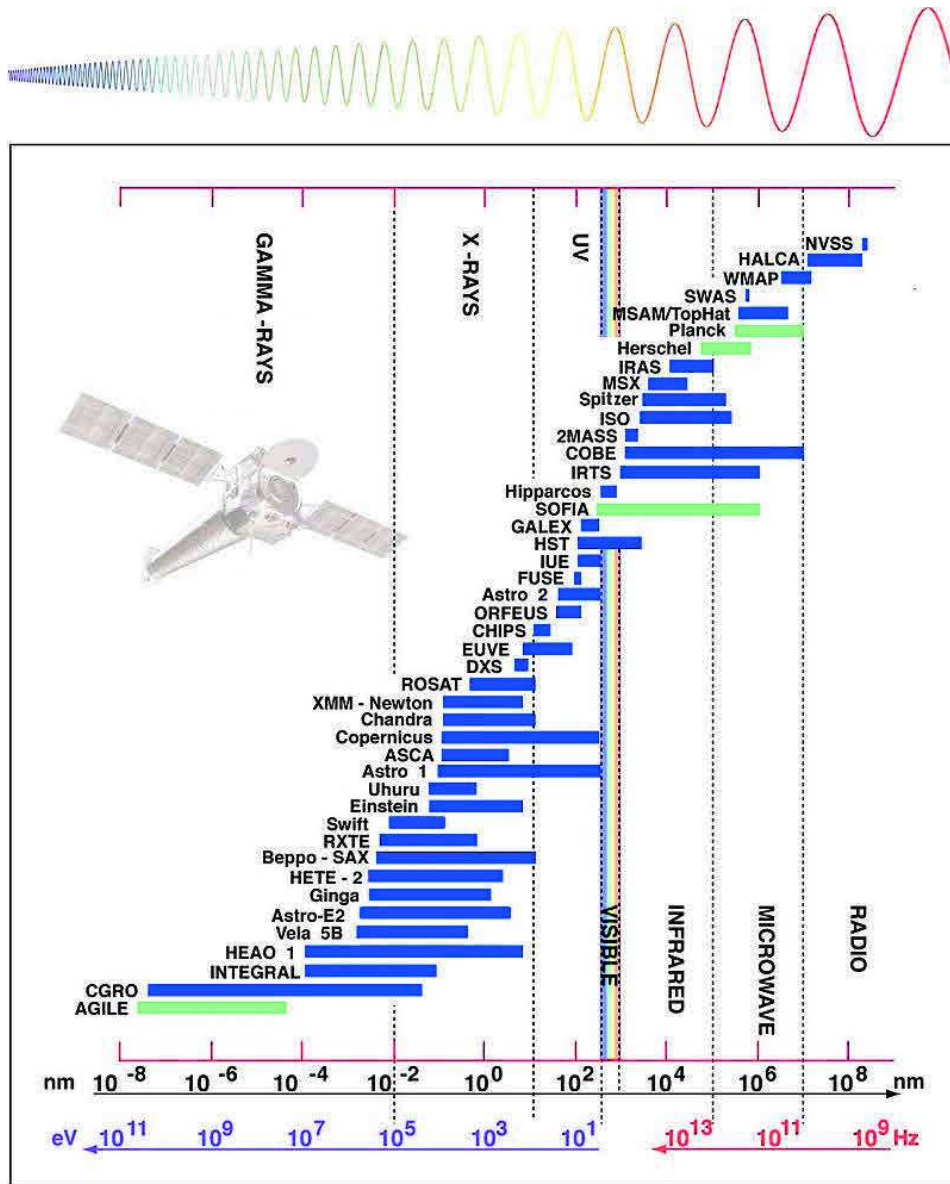


FIGURE 2.2: Intervalles de longueur d'onde des principaux télescopes spatiaux actuels et passés. (Crédit : https://fr.wikipedia.org/wiki/T%C3%A9lescope_spatial)

Nous pouvons aussi citer SOFIA¹ (Stratospheric Observatory for Infrared Astronomy) qui est un télescope particulier, puisqu'embarqué dans un avion. Ce n'est donc pas un télescope spatial même s'il observe à partir d'une haute altitude (12 km minimum), limitant ainsi l'absorption des rayonnements infrarouges moyens et lointains ainsi que sub-millimétriques par l'atmosphère. D'autres expériences ont été menées à l'aide de ballons stratosphériques comme par exemple avec PRONAOS². COBE³ est un télescope dédié à l'étude du fond diffus cosmologique, cependant le détecteur DIRBE⁴ était dédié, entre autre, à l'étude de l'émission des poussières. Enfin le télescope japonais IRTS⁵ ne releva

1. Temi et al. (2014)
2. Duran & Buisson (1990)
3. Boggess et al. (1992); Mather (1982)
4. Silverberg et al. (1993)
5. Murakami et al. (1994)

que 7% du ciel durant sa mission de 28 jours. De plus, tous ces télescopes, datant des années 80-90 ne disposaient pas de résolution et de sensibilité suffisantes pour étudier les jeunes objets stellaires ainsi que les régions de formation d'étoiles dans leur ensemble. D'autres télescopes ont vu le jour dans les années 1990-2000, à l'instar du télescope Spitzer⁶, WISE⁷, AKARI⁸, MSX⁹, ISO¹⁰ et IRAS¹¹ dans les différents domaines de longueurs d'onde que couvre l'infrarouge, et Planck¹² de l'infrarouge lointain aux ondes centimétriques. Il existe aussi quelques fenêtres atmosphériques dans le domaine sub-millimétrique qui permettent l'étude du milieu interstellaire froid, les principaux étant le télescope JCMT avec l'instrument SCUBA¹³ (450 et 850 μm), APEX avec l'instrument LABOCA¹⁴ (870 μm) et CSO avec l'instrument Bolocam¹⁵ (1.1, 1.4 et 2.1 mm). Malgré un intervalle large de longueurs d'onde, combiné grâce à Spitzer, Bolocam, LABOCA et SCUBA, il manquait toujours des données homogènes à hautes résolutions dans l'infrarouge lointain et le sub-millimétrique, entre 100 et 500 μm , Planck n'ayant pas une résolution suffisante pour l'étude de sources quasi-ponctuelles. L'arrivée du télescope Herschel¹⁶, lancé en même temps que Planck, a donc permis de combler un manque de données cruciales pour l'étude des objets stellaires jeunes. Le télescope possède, à l'heure actuelle, la meilleure résolution dans le domaine infrarouge lointain et sub-millimétriques. La table 2.1 montre les résolutions des différents instruments des télescopes évoqués plus haut en comparaison avec celles des instruments PACS et SPIRE d'Herschel. Les résolutions se rapprochant le plus de celles d'Herschel sont celles de l'instrument MIPS de Spitzer. Cependant un facteur trois est applicable aux résolutions de MIPS afin d'obtenir les résolutions d'Herschel.

2.1.2 Herschel

Le télescope Herschel est un télescope de l'ESA (Agence Spatiale Européenne), lancé le 14 mai 2009 par la fusée Ariane 5 au point de Lagrange L2 et en service du 14 juin 2009 jusqu'à fin avril 2013. Il s'agit, à l'heure actuelle, du télescope spatial disposant du plus grand miroir jamais construit dans le domaine de l'infrarouge lointain et sub-millimétrique et possédant ainsi la meilleure résolution angulaire. Trois instruments sont embarqués à son bord et permettent l'acquisition de nombreuses données nécessaires à l'étude du milieu interstellaire, de la formation stellaire dans notre galaxie mais aussi dans d'autres galaxies. Deux caméras constituées de bolomètres, PACS et SPIRE, permettent l'acquisition d'images de la poussière interstellaire. PACS cartographie la poussière dans l'infrarouge lointain, dans trois filtres centrés sur les longueurs d'onde 70 μm , 100 μm et 160 μm . SPIRE effectue la même opération mais dans le sub-millimétrique et aussi à l'aide de trois filtres centrés sur les longueurs d'onde 250 μm , 350 μm et 500 μm . Enfin le troisième instrument, HIFI, se compose d'un spectromètre à haute résolution approprié pour

-
6. [Werner et al. \(2004\)](#)
 7. [Wright et al. \(2010\)](#)
 8. [Murakami et al. \(2007\)](#)
 9. [Egan et al. \(2003\)](#)
 10. [Kessler et al. \(1996\)](#)
 11. [Neugebauer et al. \(1984\)](#)
 12. [Tauber et al. \(2010\)](#)
 13. [Holland et al. \(1999, 2013\)](#)
 14. [Siringo et al. \(2009\)](#)
 15. [Glenn et al. \(1998\)](#)
 16. [Pilbratt et al. \(2010\)](#)

TABLE 2.1: Résolutions de divers télescopes terrestres et spatiaux dans l'infrarouge et le sub-millimétrique

Télescopes	Instruments	Longueurs d'onde (μm)	Résolutions
APEX	LABOCA	870	18.6''
COBE	DIRBE	1.25-240	42'
CSO	Bolocam	1100-2100	34-65''
Herschel	PACS/SPIRE	70-500	5-37''
IRAS	Survey Array	8-120	30-120''
ISO	SWS/LWS/ISOCAM/ISOPHOT	2.5-240	1.5-90''
JCMT	SCUBA	450-850	14''
Planck	LFI/HFI	330-10 000	10-30'
Spitzer	IRAC/MIPS	3.6-160	2-38''
WISE		3.4-22	6.1-12''

une étude chimique des composants de l'univers proche. Plus d'informations sur le télescope Herschel ainsi que sur les différents instruments sont donnés dans les articles référents (mission Herschel : [Pilbratt et al. \(2010\)](#), PACS : [Poglitsch et al. \(2010\)](#), SPIRE : [Griffin et al. \(2010\)](#) et HIFI : [de Graauw et al. \(2010\)](#)).

Herschel a pour but de répondre à plusieurs problématiques actuelles, dont deux majeures : Comment les galaxies se forment et évoluent dans l'Univers jeune ? Comment les étoiles se forment, évoluent et interagissent avec le milieu interstellaire ? D'autres recherches sont venues se rajouter à ces questions, comme l'analyse de la chimie de l'Univers et des compositions chimiques des atmosphères et surfaces des comètes, des planètes et de leurs satellites dans le système solaire.

Herschel couvra une grande partie du ciel par le biais d'observations dédiées et donna lieu à de nombreux projets et collaborations tant dans notre galaxie que dans les domaines extra-galactiques. Il a notamment couvert l'ensemble du plan galactique qui est la structure qui nous intéresse tout particulièrement ici car elle comprend la majorité des régions de formation stellaire comme nous l'avons vu dans le chapitre 1. Celui-ci a donné lieu à un programme-clé, Hi-GAL¹⁷ (Herschel infrared Galactic Plane Survey) que j'évoque dans la section suivante.

2.1.3 Relevé Hi-GAL

Hi-GAL constitue le premier projet d'étude de l'ensemble du plan galactique dans l'infrarouge lointain et sub-millimétrique ([Molinari et al. \(2010b\)](#)) grâce aux détecteurs PACS et SPIRE, aux longueurs d'onde 70 μm , 160 μm , 250 μm , 350 μm et 500 μm . C'est aussi l'un des plus grands programmes d'observation du télescope Herschel.

17. <https://hi-gal.ifs-i-roma.inaf.it/higal/home.html>

Outre les objectifs d'Herschel, déjà mentionnés plus haut, Hi-GAL tente de répondre à d'autres questions au sein de la Galaxie. Il cherche notamment à connaître la contribution de la matière froide à la masse de la Galaxie, les variations du rapport gaz sur poussières ainsi que les variations d'efficacité de formation stellaire en fonction des conditions environnementales.

Comme évoqué précédemment, il permet l'étude d'objets jeunes et enfouis dans le gaz et la poussière de manière individuelle mais aussi de manière statistique pour la première fois dans ces longueurs d'onde, avec ces résolutions et avec des données homogènes. Ainsi il permet l'étude multi-échelle du plan galactique à partir des coeurs pré-stellaires jusqu'au bras spiraux.

L'obtention des images s'est faite grâce aux détecteurs PACS et SPIRE en mode parallèle ce qui permet l'observation de grandes zones du ciel en même temps, d'où l'homogénéité des données. Ainsi l'ensemble du plan fut observé en une fois avec des champs de 2 degrés carrés qui se superposent légèrement. La figure 2.3 est une image couleur composée des images à $70\mu m$ (bleu), $160\mu m$ (vert) et $250\mu m$ (rouge), montrant 4 de ces champs entre 348 et 354 degrés.

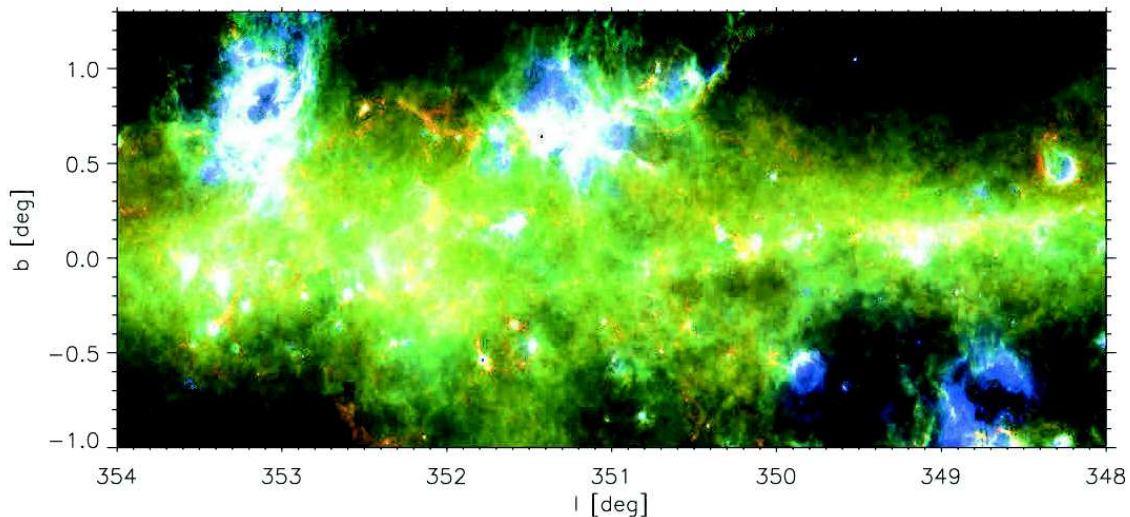


FIGURE 2.3: Composition RGB des images à 70, 160 et $250\mu m$ entre 348 et 354 degrés.

Une grande partie de mon travail de thèse repose sur ce relevé, comme par exemple l'élaboration d'un catalogue multi-longueurs d'onde de sources compactes jeunes que j'évoque dans les sections suivantes.

2.2 Les données et outils utilisés

2.2.1 Catalogues single-band Hi-GAL

Dans le but d'étudier des régions de formation stellaire et plus particulièrement les premiers stades évolutifs des étoiles, il a été nécessaire de créer un catalogue de sources réunissant les données à chaque longueur d'onde du relevé Hi-GAL. Pour ce faire, le consortium a mis à ma disposition les catalogues de sources compactes correspondant chacun à une longueur d'onde. La première étape de l'équipe Hi-GAL était de produire

TABLE 2.2: Nombre de sources dans les catalogues single-band.

	Version finale	Version utilisée
Longueur d'onde (μm)	Nombre de sources	Nombre de sources
70	123 210	122 971
160	308 509	292 051
250	280 685	280 258
350	160 972	161 855
500	85 460	85 880

des catalogues pour la partie interne de la Galaxie, qui concentre la majeure partie des régions de formation stellaire de part la profondeur du champ sondé, entre 20 et 25 kpc. La partie externe étant toujours en cours de réduction, mon travail a porté sur cette partie interne qui constitue un champ de 140 degrés en longitude, entre -70 et 70 degrés.

Les sources ont été extraites à partir des images photométriques Hi-GAL produites grâce au pipeline ROMAGAL ([Traficante et al. \(2011\)](#)). L'extraction en elle même fut réalisée grâce à l'algorithme CuTEX qui fut, contrairement à d'autres algorithmes d'extraction de sources comme getsources ou SExtractor, développé principalement dans l'idée de l'utiliser sur les données Hi-GAL. En effet ce code permet d'extraire des sources compactes dans divers environnements plus ou moins structurés à l'instar du plan galactique (voir les détails dans [Molinari et al. \(2011\)](#)).

[Molinari et al. \(2016a\)](#) présente une analyse complète de l'extraction et de la création des catalogues de sources compactes. Cette version finale des catalogues single-band n'était pas encore disponible au début de mes travaux. J'ai ainsi utilisé une version antérieure se rapprochant fortement de la version finale. La table 2.2 résume le nombre de sources trouvées à chaque longueur d'onde pour les deux versions. Comme expliqué dans [Molinari et al. \(2016a\)](#), la complétude varie selon la longitude en raison des différences de densité le long du plan galactique. Ainsi une complétude différente est trouvée pour chaque champ de 2 degrés carrés ce qui se traduit par des différences de nombre de sources extraites en plus des différences déjà présentes en raison des différentes structures et donc densités dans notre galaxie, surtout dans sa partie interne. Cependant plusieurs dizaines de milliers de sources ont été extraites dans chaque bande ce qui est, à l'heure actuelle, le plus grand échantillon de sources jeunes et qui nous permettra des études statistiques inégales.

2.2.2 Specfind

Afin de construire un catalogue de clumps j'ai utilisé SPECFIND, un algorithme d'identification croisée de catalogues développé par Bernd Vollmer ([Vollmer et al., 2005](#)). L'objectif initial de SPECFIND était d'identifier des spectres radio continus dans de grands échantillons de données. Pour cela une première identification par position est effectuée qui donne lieu à une hiérarchisation des sources de chaque catalogue. L'identification se poursuit par une recherche à l'aide des flux de chaque source hiérarchisée. L'avantage de cette algorithme réside dans l'établissement d'une hiérarchie en fonction des tailles des sources et des résolutions à chaque longueur d'onde. D'ailleurs le code fonctionne aussi

avec des catalogues aux mêmes longueur d'onde mais avec des résolutions différentes. Il se trouve que cet algorithme est bien adapté à notre étude puisque les sources à $500\ \mu\text{m}$ peuvent englober plusieurs sources à des longueurs d'onde inférieures comme par exemple à $70\ \mu\text{m}$ comme nous pouvons le voir sur la figure 2.4.

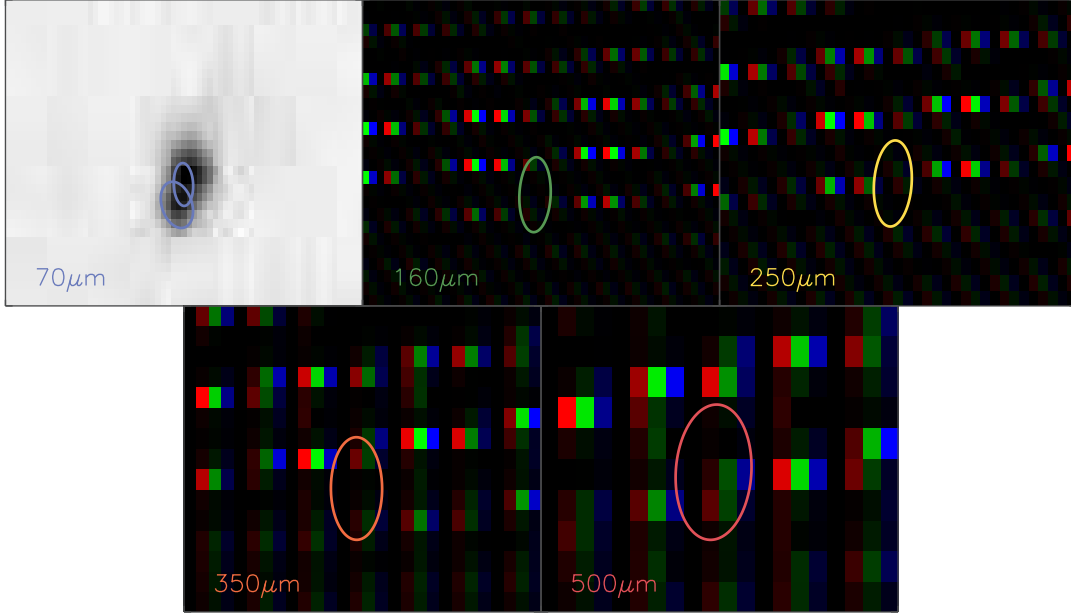


FIGURE 2.4: Exemple de confusion due à la résolution. L'objet est en fait composé d'au moins deux sous-objets à $70\ \mu\text{m}$.

Il a cependant fallu adapter SPECFIND afin qu'il s'adapte à l'étude de sources de type stellaire. En effet, à l'origine le code comprenait une identification par flux à l'aide d'une loi de puissance caractérisant l'émission free-free dans le domaine radio alors que, dans notre étude, les sources sont représentées par un corps noir modifié (équation 2.1) du à l'émission des poussières. Ainsi nous considérons l'ajustement de l'équation suivante :

$$F_\lambda = A\lambda^{-\beta} B(T, \lambda) \quad (2.1)$$

$B(T, \lambda)$ représente la loi de Planck caractérisant la luminance du corps de noir, β l'indice caractérisant l'émissivité des poussières et A la constante ajustable comprenant différentes caractéristiques de la source et globale que nous évoquerons dans la section 4.3. Nous allons donc ajuster cette loi et étudier les différents résultats, suivant différents critères, dans la section suivante.

2.3 Catalogue multi-longueurs d'onde

2.3.1 Conditions d'ajustement

Nous allons ainsi combiner les outils introduits dans la section précédente afin de construire un catalogue *bandmerged* de sources jeunes. Nous nous aidons des positions, des tailles des sources et des flux fournis par les 5 catalogues (un pour chaque longueur d'onde) ainsi que les résolutions aux différentes longueurs d'onde. Les erreurs sur le flux sont un critère important pour la bonne évaluation des identifications. Dans toutes les bandes le pic des erreurs correspond à approximativement 10% du flux comme

nous pouvons le voir sur la figure 2.5. 71%, 84%, 83%, 80% et 76% des objets possèdent, respectivement, des erreurs inférieures à 20% du flux à $70\ \mu m$, $160\ \mu m$, $250\ \mu m$, $350\ \mu m$ et $500\ \mu m$.

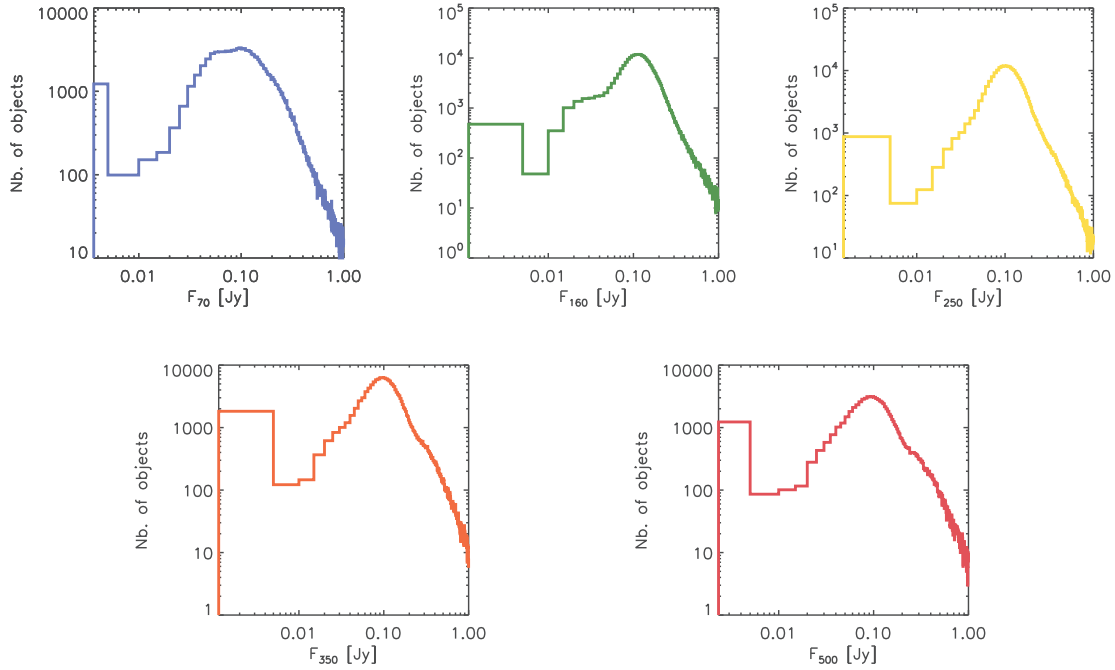


FIGURE 2.5: Histogramme des rapports entre l'erreur sur le flux et le flux à longueur d'onde. Le pic des histogrammes se situe aux alentours de 10% du flux.

Dans tous les résultats qui suivent, les valeurs des erreurs sur le flux inférieures à 20% du flux sont réajustées à ce pourcentage afin d'obtenir des erreurs homogènes pour l'ensemble des données. Cela permet aussi d'obtenir une valeur des erreurs plus réaliste, prenant en compte les erreurs dues à la non symétrie et l'aspect non gaussien des sources, ces causes étant prises en compte pour l'estimation du flux intégré mais pas pour l'estimation des erreurs. Enfin, l'estimation de l'erreur se faisant dans une fenêtre comprenant plusieurs sources, donc moyennée sur toute cette fenêtre, une hausse des valeurs des erreurs permet de prendre en compte un bruit plus important au niveau de la source où le flux est estimé.

Nous voulons aussi éviter les sources possédant peu de points et donc ayant des paramètres peu contraints. C'est pourquoi seules les sources détectées dans au moins trois longueurs d'onde sont gardées.

2.3.2 Résultats et catalogue final

La première étape est d'éliminer les sources ne possédant pas au moins trois points consécutifs. Il est d'une importance cruciale d'obtenir un nombre suffisant de points à ajuster afin d'obtenir la meilleure fiabilité possible des sources, en évitant par exemple les cas limites où le nombre de paramètres à ajuster est égal aux nombres de variables. Cependant, il est aussi important que ces points soient consécutifs ce qui démontre une fiabilité supplémentaire de ces sources. Les sources ne présentant pas cette caractéristique sont ainsi rejetées du catalogue.

TABLE 2.3: Caractéristiques des catalogues pour chaque condition de départ. Les deux conditions se rapportent à l'ajustement par un corps noir modifié et concernent la prise en compte du catalogue à $70 \mu m$ et la possibilité de laisser β comme paramètre libre.

Conditions initiales	Nb. total de sources	Nb. rejetées	Nb. d'AGBs	Nb. SED ≥ 4 points	Nb. avec $T > 100$ K	NB. final
$70 \mu m + \beta$	89 252	35 506	865	6 979	1 557	53 746
β	88 384	32 643	3 046	29 987	1 597	55 741
$\beta = 2, 70 \mu m$	58 576	22 031	25	727	70	36 545
$\beta = 2$	60 526	20 194	2 411	9 622	85	40 332

L'opération suivante à effectuer afin d'analyser correctement les données sorties de SPEC-FIND est d'éliminer les contaminations causées par des objets n'appartenant pas aux objets stellaires jeunes. Les sources possédant au moins une identification à 70 , 160 , 250 et $350 \mu m$ peuvent notamment servir à l'analyse de la contamination des catalogues par les AGBs. Martinavarro et al. (en préparation) propose une identification des AGBs grâce à un diagramme couleur-couleur basé sur ces quatre longueurs d'onde. Les critères de sélection sont décrits dans [Veneziani et al. \(2013b\)](#) et se résument aux intervalles de valeurs suivants :

$$0.2 < \frac{\log \frac{F_{70}}{F_{160}}}{\log \frac{F_{70}}{F_{250}}} < 0.9 \quad \text{et} \quad 1.2 < \frac{\log \frac{F_{70}}{F_{350}}}{\log \frac{F_{160}}{F_{350}}} < 3.3 \quad (2.2)$$

Où F_{70} , F_{160} , F_{250} et F_{350} représentent les flux aux différentes longueurs d'onde. Toutes les sources qui tombent dans ces deux intervalles sont ainsi supprimées du catalogue. La table 2.3 indique le nombre de source correspondant à des AGBs. Le nombre est à comparer avec le nombre de source ayant au moins les 4 points demandés par les critères de suppression d'AGBs, aussi indiqué dans la table 2.3.

Après avoir fait le tri des sources les plus fiables, nous pouvons maintenant étudier le catalogue. En fait, nous avons construit quatre catalogues différents correspondant à deux conditions de départ. La première condition était de laisser l'indice caractérisant l'émission des petits grains β libre ce qui sous-entend d'avoir trois paramètres libres avec A , T comme exprimés dans l'équation 2.1. la deuxième condition est de construire un catalogue sans l'aide du catalogue à $70 \mu m$. Comme nous avons pu le voir dans le chapitre 1, il existe différents contaminants du spectre d'énergie des jeunes objets stellaires, à savoir les PAHs, VSGs et BGs. À $70 \mu m$, les VSGs jouent un rôle important ce qui peut augmenter la valeur du flux à cette longueur d'onde et peut ainsi rendre difficile l'ajustement par un corps noir modifié. Nous l'avons quand même utilisé pour observer les différences occasionnées dans les deux cas. Ceci revient à dire que dans un cas nous avons ajusté l'ensemble de la SED (5 points théoriques) et que dans l'autre, la présence d'une source à $70 \mu m$ n'est utilisée que pour une identification par position. La table 2.3 résume les caractéristiques des catalogues pour chaque condition. Il est à noter que lorsque le flux à $70 \mu m$ n'est pas pris en compte dans l'ajustement, le nombre d'AGBs et le nombre de sources avec 4 points consécutifs ($70 \mu m$ inclu) donnés dans la table 2.3 sont purement indicatifs.

Il se trouve que dans les quatre catalogues des valeurs de température aberrantes

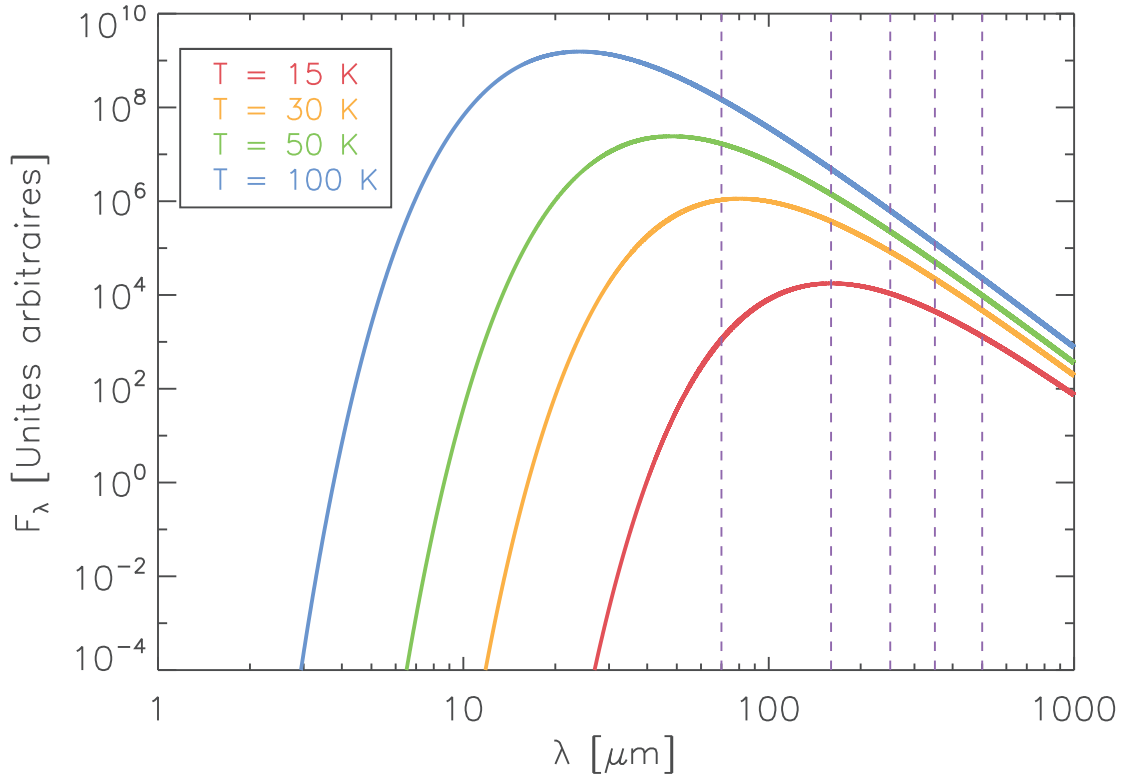


FIGURE 2.6: Loi du corps noir modifiée comme développée dans l'équation 2.1, pour différentes températures avec $\beta = 2$. Les lignes pointillées localisent les 5 longueurs d'onde utilisées dans le cadre du projet Hi-GAL. On s'aperçoit qu'à 100 K nous sommes dans le régime Rayleigh-Jeans à toutes les longueurs d'onde.

apparaissent. La température dépend fortement des contraintes appliquées par les flux à chaque longueur d'onde. La figure 2.6 donne une idée de la contribution des flux en respectant l'équation 2.1, pour différentes températures avec β fixé à 2. Les lignes pointillées localisent les 5 longueurs d'onde utilisées dans le cadre du projet Hi-GAL. On peut remarquer qu'au delà de 50 K, les flux à $\lambda \geq 160 \mu m$ peuvent être ajustés par une droite. À partir de 100 K cela s'étend à toutes les longueurs d'onde. Au-delà de cette température la source est en fait complètement dans le régime Rayleigh-Jeans qui devient prédominant lorsque $\lambda \gg hc/k_B T$ et s'écrit :

$$B(\lambda, T) = \frac{2ck_B T}{\lambda^{4+\beta}} \quad (2.3)$$

Dans ce régime la température joue le rôle de constante, en plus de celle déjà mentionnée dans l'équation 2.1. Une légère fluctuation du flux ou des erreurs larges peuvent entraîner de grandes valeurs de température. Nous éliminons ainsi les sources ayant une température supérieure à 100 K. Cette température représente plutôt une limite supérieure au delà de laquelle les ajustements sont considérés comme erronés plutôt qu'une valeur physique. Il se trouve d'ailleurs que les sources dans ce domaine de longueurs d'onde ont des températures inférieures à 50 K, comme nous pouvons le voir dans [Elia et al. \(2013\)](#) et [Veneziani et al. \(2013b\)](#). Nous avons aussi éliminé les sources ayant une température inférieure à 5 K, causé par une corrélation entre β et T comme nous allons le voir ci-dessous. Le nombre total de sources éliminées selon le critère de température est indiqué dans la table 2.3. Au total nous avons éliminé quasiment 40% des associations.

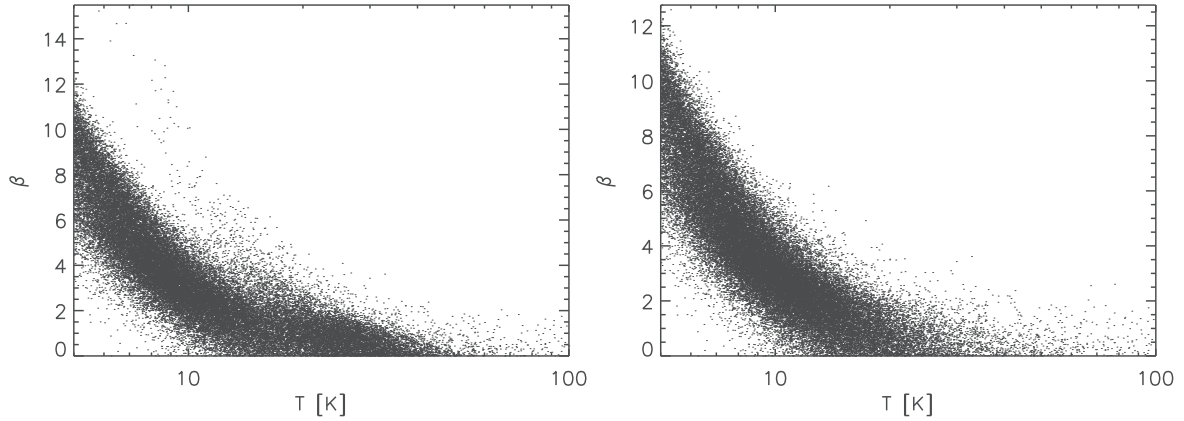


FIGURE 2.7: Corrélation apparente entre la température du corps noir modifié et β en prenant en compte toutes les longueurs d'onde disponibles (partie gauche) et sans $70 \mu m$ (partie droite). Les températures minimales et maximales ont été fixées à 5 K et 100 K comme expliqué dans le texte plus haut.

Les sources restantes composent ainsi le catalogue final. Il reste cependant à déterminer les conditions initiales d'ajustement les plus adaptées à l'étude de ces sources jeunes. Comme nous l'avons expliqué plus haut, les flux à $70 \mu m$ sont contaminés par les VSGs. Ainsi le catalogue final ne comprendra qu'une évaluation par position de la composante à $70 \mu m$.

En ce qui concerne β , il se trouve qu'il existe une dégénérescence forte avec la température comme nous pouvons le voir sur les deux parties de la figure 2.7. La figure montre la corrélation apparente température- β lors de l'ajustement avec et sans l'utilisation du flux à $70 \mu m$. Les valeurs de β sont anormalement élevées et devraient se situer entre 0 et 3, 0 correspondant au corps noir. Il est tout de fois à noter que 90% des valeurs de β se situent en dessous de 10. Lorsque nous ne considérons que les sources ayant au moins 4 points, le pourcentage monte à 99%. Autrement dit nous obtenons quelques valeurs aberrantes dues à une association litigieuse en plus de la dégénérescence avec la température. Plusieurs recherches ont été effectuées afin d'étudier et d'éliminer cette dégénérescence. Nous pouvons notamment citer [Kelly et al. \(2012\)](#), [Veneziani et al. \(2013a\)](#) et [Elia & Pezzuto \(2016\)](#). Les deux premières recherches utilisent une méthode bayésienne qui permet de réduire l'intervalle de valeur de β . [Veneziani et al. \(2013a\)](#) trouve notamment une valeur moyenne $\beta = 1.6$ proche de celle trouvée grâce au satellite Planck ([Planck Collaboration et al., 2011](#)). [Elia & Pezzuto \(2016\)](#) ont développé deux formules analytiques pour calculer de manière indépendante la température et β . Appliquées à des sources réelles, ils trouvent des valeurs de β comprises entre 1 et 2.

Afin de contourner ce problème, nous avons choisi la valeur de β à 2. Cette valeur est largement utilisée dans la littérature, cependant β varie selon les environnements et notamment suivant les différentes abondances de grains ([Arab et al., 2012](#)).

Au final le catalogue retenu est celui avec un β fixé et un ajustement sans la composante à $70 \mu m$. Nous allons maintenant nous intéresser au catalogue produit par le consortium Hi-GAL et le comparer avec notre catalogue en étudiant certaines caractéristiques physiques des sources.

2.4 Catalogue de l'équipe Hi-GAL

Un des objectifs de l'équipe Hi-GAL était de fournir un catalogue de sources jeunes à partir des catalogues *singlebands*. Ce projet est encore en cours actuellement avec des versions de plus en plus abouties de ce catalogue. J'ai pu accéder à une des premières versions fiables quelques mois après le début de ma thèse, ce qui m'a permis de le comparer avec le catalogue que j'ai créé. La version du catalogue utilisée ici comprend 99181 objets. Nous allons décrire le catalogue dans cette partie et le comparer à notre catalogue. Cependant, avant de présenter cette partie de mon travail, il est important de définir le terme et notion de *clumps* qui reviendra souvent en fin de ce chapitre et dans les chapitres suivants.

2.4.1 Contenu

Le catalogue (Elia et al., in prep) se présente en fait sous la forme de deux fichiers, appelés respectivement *filtered catalog* et *physical properties catalog* que nous appellerons par la suite catalogue *product*. Le premier comprend, entre autre, les flux et les tailles des sources aux différentes longueurs d'onde. Il est à noter que le catalogue comprend aussi les associations faites, seulement par position, avec d'autres catalogues, à savoir MSX, Bolocam, LABOCA (APEX), WISE et MIPS GAL¹⁸ (relevé du plan galactique avec l'instrument MIPS de Spitzer). Comme pour les catalogues à chaque longueur d'onde, les données sont comprises dans la partie interne de la galaxie, entre -70 et 70 deg en longitude. 99 083 clumps composent ce catalogue.

Le deuxième catalogue comprend des propriétés physiques des sources calculées à partir des ajustements de SED. Comme dans notre étude, l'ajustement ne prend pas en compte la composante à 70 μm du flux de la source. De plus l'indice spectral d'émissivité β est aussi fixé à 2. L'avantage principal de ce catalogue réside dans l'utilisation des distances héliocentriques pour dériver certaines propriétés physiques des sources comme la masse par exemple. Je reviendrai sur la notion de distance héliocentrique dans le chapitre suivant concernant l'élaboration d'un catalogue d'amas de sources jeunes.

D'autres conditions d'ajustement, qui diffèrent légèrement de celles de notre étude, sont évoquées dans la section 4.3 afin de discuter de la comparaison des deux catalogues de sources.

2.4.2 Notion de *clumps*

En reprenant les tailles des objets aux différentes longueurs d'onde, on remarque que malgré une bonne résolution dans ces longueurs d'onde ($\leq 37''$), nous n'avons pas exclusivement affaire à des proto-étoiles uniques. En effet, nous pouvons calculer une taille réelle moyenne des sources à partir du pic de distribution des tailles angulaires des sources aux différentes longueurs d'onde ainsi qu'en utilisant différentes distances héliocentriques calquées sur celles fournies dans le catalogue. La table 2.4 nous renseigne sur ces tailles réelles.

Nous pouvons voir qu'au delà d'une certaine distance, les sources ne peuvent plus être considérées comme unique. Nous arrivons facilement à la taille du système solaire

18. Carey et al. (2009)

TABLE 2.4: Taille des résolutions angulaires à différentes distances héliocentriques.

Longueur d'onde (μm)	résolution	Tailles (pc)			
		1 kpc	5 kpc	10 kpc	15 kpc
70	5''	0.024	0.121	0.242	0.364
160	12''	0.058	0.291	0.582	0.873
250	18''	0.087	0.436	0.873	1.309
350	25''	0.121	0.606	1.212	1.818
500	37''	0.179	0.897	1.794	2.691

à 1 kpc, ce qui constitue une limite possible pour un objet unique, et à la distance Soleil-Proxima Centauri (1.3 pc) au delà de 10 kpc. Comme les étoiles se forment plus proche les unes des autres puis se dispersent, il est ainsi aisé de penser que nous avons affaire à des systèmes multiples. Nous parlerons ainsi de **clumps** dans toute la suite de ce manuscrit lorsque nous parlerons des sources Herschel.

2.4.3 Comparaison et Discussion

Premièrement nous pouvons identifier le nombre de clumps en commun dans les deux catalogues à l'aide d'une identification croisée par position. Les coordonnées du catalogue *product* correspondent aux coordonnées à la plus petite longueur d'onde de chaque source. Il se trouve que 73% des clumps présents dans le catalogue de notre étude sont aussi présents dans le catalogue Hi-GAL.

Il est ensuite possible de comparer certaines propriétés physiques des clumps des deux catalogues, celui du consortium Hi-GAL et de notre étude, en prenant en compte certaines considérations ainsi qu'en discutant des hypothèses faites lors de l'ajustement et décrites dans [Elia et al. \(2013\)](#).

Tout d'abord nous pouvons nous intéresser aux masses des clumps. La constante A utilisée dans l'ajustement à travers l'équation 2.1 renferme plusieurs composantes. Pour comprendre d'où proviennent et quelles sont ces composantes, nous devons revenir à la définition commune de l'équation du corps noir modifié à ajuster, déjà présentée dans le chapitre 1 et rappelée dans l'équation 2.4.

$$F_\lambda = \tau \Omega B(T, \lambda) \quad (2.4)$$

τ représente l'opacité et Ω l'angle solide de la source. On notera ici que l'on remplace l'expression originelle de la profondeur optique $1 - e^{-\tau}$ par τ car nous observons ici l'enveloppe de la source, considérée comme optiquement mince ($\tau \ll 1$). τ peut s'écrire $\tau = (\lambda_0/\lambda)^\beta$ ou λ_0 correspond à la longueur d'onde à partir de laquelle la source devient optiquement épaisse. Nous pouvons aussi écrire τ sous la forme suivante :

$$\tau = k_{ref} \left(\frac{\lambda_{ref}}{\lambda} \right)^\beta \frac{M}{A_{ref}} \quad (2.5)$$

Où $k_{ref} = k_{250}$ représente l'opacité à λ_{ref} . [Elia et al. \(2013\)](#) utilisent comme longueur d'onde de référence $\lambda_{ref} = 250\mu m$ et $k_{250} = 0.1cm^2 g^{-1}$ ([Hildebrand, 1983](#)). M représente la masse de l'enveloppe et A_{re} la surface du clumps. En posant $\Omega = A/D^2$ avec D la distance héliocentrique, nous pouvons réécrire l'équation 2.4 :

$$F_{\lambda} = k_{ref} \left(\frac{\lambda_{ref}}{\lambda} \right)^{\beta} \frac{M}{D^2} B(T, \lambda) \quad (2.6)$$

Ainsi en utilisant cette forme avec l'équation 2.1, nous pouvons extraire une expression de la constante A :

$$A = k_{ref} \lambda_{ref}^{\beta} \frac{M}{D^2} \quad (2.7)$$

En inversant l'expression nous pouvons calculer la masse des clumps à condition de connaître la distance héliocentrique. Nous ne possédons pas cette information au contraire du catalogue *product*. Cependant, en admettant un certain intervalle de valeurs des distances, calqué sur les valeurs présentes dans le catalogue *product*, il est possible d'obtenir un intervalle de masse des clumps. La figure 2.8 montre la distribution des distances héliocentriques provenant du catalogue *product*. On choisit trois valeurs de distances caractéristiques afin de calculer les masses, 3 kpc, 10 kpc et 15 kpc. 3 kpc et 10 kpc correspondent à deux pics de l'histogramme des distances tandis que 15 kpc représente une valeur limite supérieure. La figure 2.9 compare la distribution des masses provenant du catalogue *product* et à celles de notre étude pour les trois valeurs de distances choisies. Les intervalles sont significativement différents suivant les valeurs utilisées dans le calcul. Cependant les intervalles sont en parfait accord avec ceux du catalogue *product*. En considérant un mixte de valeur de distances, nous aurions probablement un pic situé entre 10^2 et 10^3 masse solaire. On peut constater que l'intervalle de masse est important, allant de quelques masses solaires à plusieurs dizaines de milliers. Il est possible que cet écart de masses nous indique la présence d'objets différents. Ce point sera étudié dans le chapitre 5.

Après s'être intéressé aux masses des clumps, nous pouvons nous intéresser aux stades évolutifs. Dans le catalogue *product*, trois stades sont différenciés, les clumps *unbound*, pré-stellaires et proto-stellaires. Sur les 99 083 clumps du catalogue il y a 17 424 unbound, 56 773 pré-stellaires et 24 886 proto-stellaires. Les clumps pré-stellaires sont différenciés des clumps *unbound* en utilisant la relation masse-rayon de Larson connue aussi sous le nom *troisième loi de Larson*, $M(r) > 460 M_{\odot} (r/pc)^{1.9}$, avec r le rayon de la source ([Larson, 1981](#)). Ensuite, les flux à $70 \mu m$ sont utilisés dans le but de discriminer les clumps pré-stellaires des proto-stellaires. [Dunham et al. \(2008\)](#) montrent que la présence d'un flux à $70 \mu m$ est lié à une proto-étoile et peut être difficilement associé au chauffage seul du MIS. Nous procédons de la même façon afin de comparer les températures et la nature des sources de notre catalogue et celui du consortium Hi-GAL. Cependant, comme nous n'avons pas accès à des masses fiables, nous ne pouvons distinguer que les clumps pré-stellaires des proto-stellaires. La figure 2.10 illustre nos propos, en comparant les distributions de température des deux catégories de clumps pour le catalogue *product* et notre étude. On remarque d'abord que les valeurs sont similaires dans les deux catalogues. De plus, les clumps proto-stellaires (en bleu sur les figures) semblent être plus chaud en moyenne que les clumps pré-stellaires et les clumps unbound (en rouge sur les figures). Ceci est aussi cohérent avec les études de [Bontemps et al. \(2010\)](#) et [Gianini et al. \(2012\)](#) qui montrent que le flux à $70 \mu m$ est un critère de distinction entre les

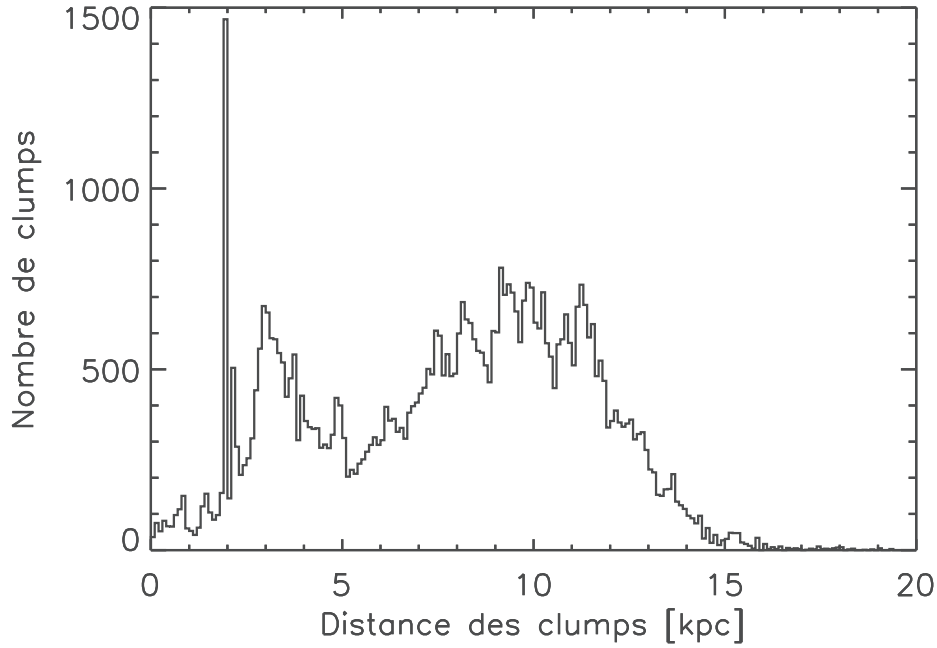


FIGURE 2.8: Histogramme des distances héliocentriques des clumps. Il semble y avoir deux zones privilégiées de localisation des clumps, vers 2-3 kpc et dans l'intervalle 8-12 kpc.

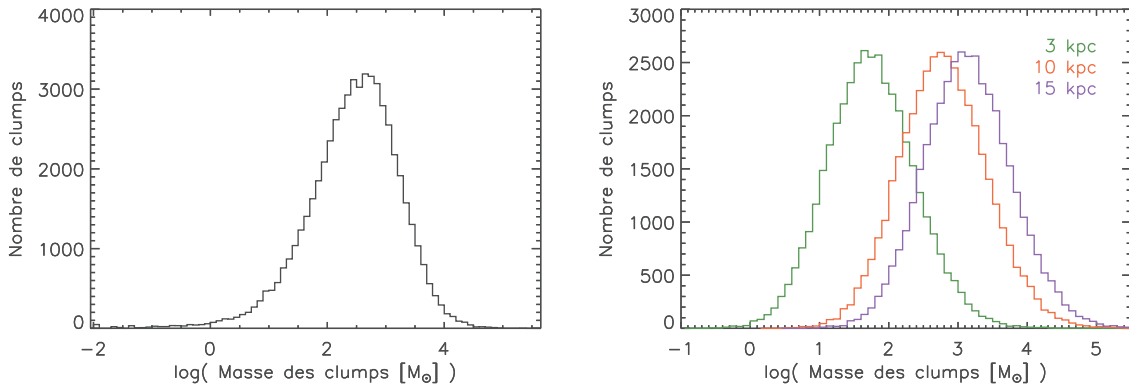


FIGURE 2.9: Gauche : Histogramme des distances héliocentriques des clumps. Droite : Histogramme des masses des clumps (log).

sources possédant un chauffage interne, donc plus évoluées et chaudes, et les sources pré-stellaires.

Ensuite, une autre différence entre le catalogue de notre étude et le catalogue *product* réside dans les paramètres d'ajustement. Il se trouve que λ_0 , provenant de l'expression de τ , fait partie des paramètres ajustés et présents dans le catalogue *product*. Ils distinguent les sources optiquement fines des sources optiquement épaisses suivant si l'ajout du paramètre libre λ_0 ajuste mieux les SED des clumps.

Enfin, les flux au-delà de $250 \mu m$ sont recalculés à partir de la taille déconvoluée des objets et selon la formule suivante (Elia et al., 2013) :

$$F_{\lambda, new} = F_{\lambda, old} \left(\frac{\theta_{250}}{\theta_{\lambda}} \right) \text{ pour } \lambda > 250 \mu m \quad (2.8)$$

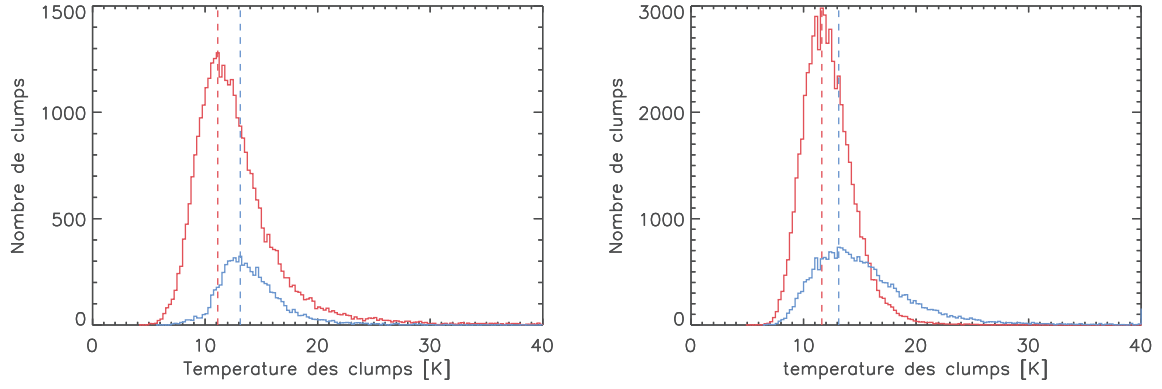


FIGURE 2.10: Gauche : Histogramme des températures des clumps pré-stellaires (rouge) et proto-stellaires (bleu). Droite : Même histogramme mais pour le catalogue *product*.

En général les sources ont une taille plus importante aux plus grandes longueurs d'onde non pas seulement en raison de la résolution mais aussi à cause du volume de poussières froides sondé. Comme développé dans [Elia et al. \(2013\)](#), ils ont ainsi préféré réajuster les tailles des sources par soucis de cohérence entre longueurs d'onde. La figure 2.11 montre l'écart entre les flux originaux et recalculés pour une source brillante et une source faible. Il est clair que ce recalcul change les valeurs des paramètres libres. Il est aussi probable que cela impact directement la fiabilité des ajustements et ainsi le nombre d'associations trouvées. En ajoutant à ça la différence occasionnée par le processus d'ajustement et les paramètres laissés libres, il est possible d'observer un tel écart de nombre de sources entre le catalogue de notre étude et celui du consortium Hi-GAL (plus du double dans le catalogue *product*).

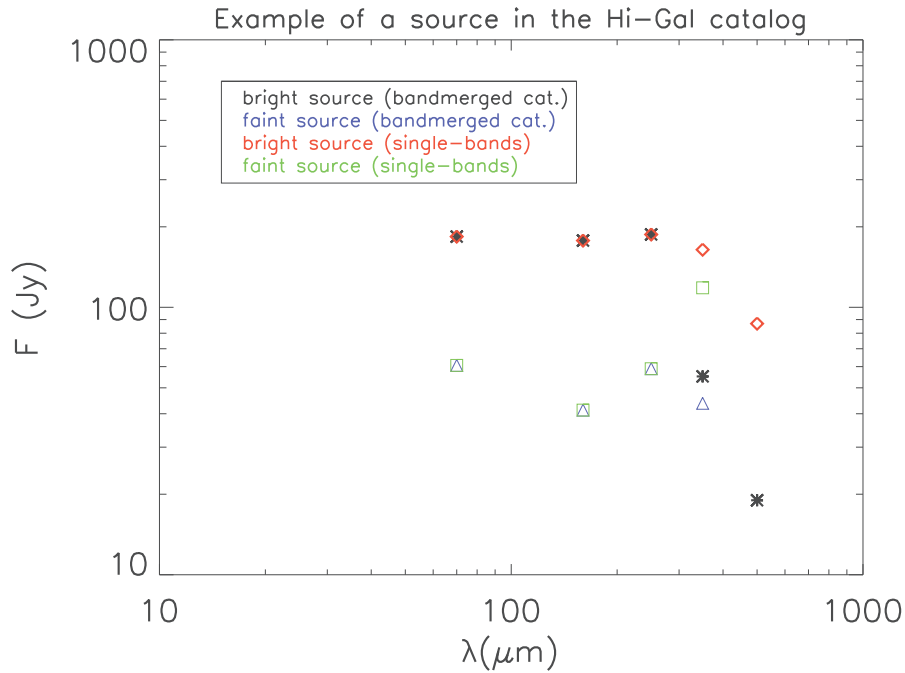


FIGURE 2.11: Différence observée entre les flux des catalogues single-band et du catalogue bandmerged. La redéfinition des flux au-delà de $250\mu m$ entraîne un écart important avec les flux originaux.

Afin d'être rigoureux dans l'élaboration d'un tel catalogue, il serait utile d'appliquer une

correction de couleur à chaque longueur d'onde. Cela constituerait un travail obligatoire en vue d'une utilisation et d'une publication de ce catalogue. Cependant cette étape n'a pas été terminée en raison de l'arrivée du catalogue de l'équipe Hi-GAL. En effet, ce catalogue est plus abouti en raison du temps investi (plus de deux ans) et des données supplémentaires mises à leur disposition comme par exemple les distances héliocentriques. Il s'ensuit que j'ai utilisé le catalogue *product* afin de poursuivre plus rapidement la suite de mes travaux de recherche.

2.4.4 Remarques

A la fin de mon doctorat, une nouvelle version du catalogue de propriétés physiques des clumps a été produit. Cette nouvelle version a été élaborée à partir des nouvelles versions des catalogues single-band, évoquées dans la section 2.1. Quelques modifications ont été apportées par rapport aux versions précédentes. La plus importante est la distinction entre les sources fiables et celles rejetées, basé sur le critère du nombre de composantes de la SED (> 3 points). Je n'ai délibérément pas pris en compte ces modifications lors de mon étude car le temps investi est beaucoup trop important au vu des différences de résultats escomptées. Je rappelle que nous faisons face à une étude statistique qui n'a de sens que dans le cas de grands échantillons de données. Les modifications apportés à ce catalogue étant minimes, il est fortement probable qu'elles conduisent aux mêmes résultats.

Pour s'en convaincre, voici les différences observées avec celui utilisé pour ma thèse, pour quelques propriétés physiques des sources. La dernière version possède 100 922 clumps tandis que la version utilisée en possède 99 083. La figure 2.12 compare l'histogramme des masses et des températures des deux catalogues. On remarque qu'il n'y a que très peu de différence. Le résultat est le même pour tous les paramètres.

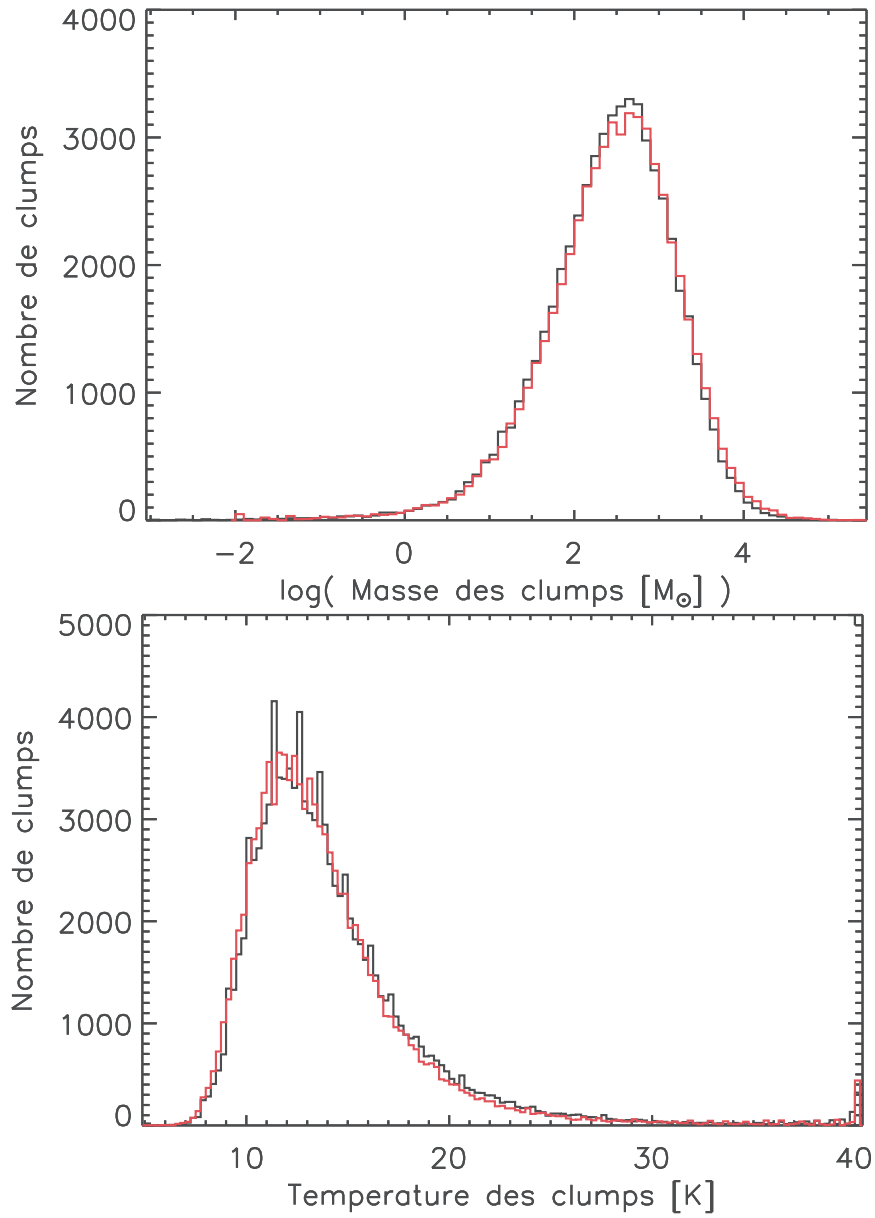


FIGURE 2.12: Gauche : Comparaison de la distribution des masses entre la dernière version du catalogue *product* et celle utilisée. Droite : Même comparaison mais pour les températures.

CHAPITRE

3

LE CATALOGUE D'AMAS

Sommaire

3.1	La méthode	46
3.1.1	Les méthodes de sélection de sur-densités	46
3.1.1.1	Carte de densité	46
3.1.1.2	Méthode des k plus proches voisins	46
3.1.1.3	Cellules de Voronoï	47
3.1.1.4	Minimum Spanning Tree	48
3.1.2	Choix de la méthode MST	49
3.2	Construction du catalogue d'amas	50
3.2.1	Stades évolutifs des clumps et confusion	50
3.2.2	Contraintes imposées par la structure galactique	53
3.2.3	Distances héliocentriques et catalogue d'amas fiables	59
3.3	Caractéristiques du catalogue d'amas	61
3.4	Discussion	64
3.5	Description des colonnes	66
3.5.1	Catalogue d'amas fiables	66
3.5.2	Catalogue archive	68
3.6	Conclusion	69

Dans cette section j'évoque le premier grand travail de ma thèse, à savoir la construction d'un catalogue d'amas d'objets jeunes à l'aide du catalogue de clumps élaboré par une partie de l'équipe Hi-GAL (catalogue *product* utilisé dans le chapitre précédent). Je commence par évoquer la méthode utilisée pour poursuivre ensuite par des tests de cette méthode afin d'obtenir une estimation de l'efficacité de celle-ci. Je finirai par la construction du catalogue et j'évoquerai les particularités de celui-ci.

3.1 La méthode

Le but de cette section est d'introduire les différentes méthodes d'identification de sur-densités et de converger vers la méthode utilisée dans notre étude.

3.1.1 Les méthodes de sélection de sur-densités

3.1.1.1 Carte de densité

Une carte de densité est un outil simple et efficace pour trouver une majeure partie des sur-densités. La technique est simple car il suffit de diviser la région sondée en bins d'une taille donnée et de compter le nombre d'étoiles associées à chacun d'entre eux. Les sur-densités sont ensuite choisies suivant l'écart de densité à la moyenne de l'ensemble des bins. Il est d'usage d'utiliser un seuil entre 2σ et 5σ .

Il est cependant nécessaire de choisir une taille de bin appropriée suivant l'échelle sondée. En effet, cette méthode s'adapte à toutes les échelles, de l'extra galactique aux amas de quelques étoiles. Elle permet aussi de trouver des sous-structures dans les amas, comme par exemple dans la PDR d'une région H II. Une taille de bin trop grande ne permettra pas de reconnaître les différentes structures des amas et une taille trop petite sera trop sensible aux fluctuations statistiques de la densité de sources. La méthode peut être améliorée en utilisant plusieurs tailles de bins pour un même échantillon de données (Kirsanova et al., 2008). Cela permet d'explorer différentes échelles mais aussi d'étudier de grands jeux de données où les densités et les tailles de structures varient significativement.

Cependant cette technique ne permet pas de déterminer avec précision les contours d'un amas ainsi que des paramètres élémentaires comme la position centrale de celui-ci lorsque la densité n'est pas piquée à un seul endroit. Nous allons donc explorer d'autres méthodes plus élaborées dans la suite.

3.1.1.2 Méthode des k plus proches voisins

L'approche ressemble à la méthode précédente. Il s'agit en effet de construire une carte de densité mais cette fois-ci à partir des sources présentes à proximité de chacune des sources. Cette méthode est utilisée dans de nombreux domaines et problèmes scientifiques comme par exemple dans des problèmes de classification en intelligence artificielle. La première utilisation de cette méthode en astrophysique semble provenir de

Casertano & Hut (1985) en s'appuyant sur des travaux de von Hoerner (1963). L'estimateur de la densité locale est décrit comme suit :

$$\rho_k = \frac{k-1}{\zeta(r)} m \quad (3.1)$$

Avec k l'indice du k -ème plus proche voisin, m la masse de chaque source et ζ le volume où l'aire de la sphère de rayon r , suivant si les sources sont étudiées dans un espace tridimensionnel ou projetées sur un plan. $m = 1$ lorsque l'on veut calculer une densité à partir du nombre de sources.

L'avantage de cette méthode réside dans le fait qu'il n'y ait pas de taille de bin à choisir comme dans la première méthode évoquée. Cependant cette méthode est aussi sensible aux variations statistiques de densité et il faut ainsi choisir avec prudence le nombre de voisins pour déterminer la densité locale ainsi qu'un seuil de densité (typiquement 3σ au dessus de la densité du fond). Casertano & Hut (1985) proposent d'utiliser une valeur $k \geq 6$. Il est aussi évident qu'une valeur trop grande de k ne permettra pas d'observer les sous-structures les plus fines. Il est cependant possible de faire varier k sur le même jeu de données afin d'étudier les sur-densités sur plusieurs échelles de taille.

Un autre avantage de cette méthode est la possibilité de définir d'autres quantités des sur-densités comme la position centrale de celles-ci :

$$x_{d,k} = \frac{\sum_{i=1}^k x_i \rho_k^i}{\sum_{i=1}^k \rho_k^i} \quad (3.2)$$

x_i est la position de la i -ème source de la sur-densité et ρ_k^i la densité comprenant le k -ème plus proche voisin autour de cette source. Il est aussi possible d'obtenir le rayon de densité qui est défini comme étant la valeur moyenne de la densité pondérée par la distance de chaque source au centre de la sur-densité $x_{d,k}$:

$$r_{d,k} = \frac{\sum_{i=1}^k |x_i - x_{d,k}| \rho_k^i}{\sum_{i=1}^k \rho_k^i} \quad (3.3)$$

Plusieurs études sont consacrées à l'étude d'amas d'étoiles avec l'aide de cette méthode (Cambrésy et al., 2013; Gutermuth et al., 2005; Román-Zúñiga et al., 2008).

Nous allons voir dans la suite de cette section que d'autres méthodes suivent des développements plus complexes.

3.1.1.3 Cellules de Voronoï

Les cellules de Voronoï ou mosaïque de Voronoï, diagramme de Voronoï, "Voronoi tessellation" en anglais, est une autre méthode non paramétrique pour déterminer la densité locale dans un échantillon de données. Cette méthode fait partie d'un ensemble d'algorithmes développés dans le cadre de la théorie des graphes. Cette théorie mathématique, développée par le mathématicien suisse Leonhard Euler, vise à résoudre des problèmes

complexes dans de nombreux domaines liés à la notion de réseau mais aussi dans des domaines annexes comme par exemple en génétique.

La mosaïque de Voronoï est liée à la triangulation de Delaunay ([Shamos & Hoey, 1975](#); [Toussaint, 1980](#)). Cette dernière consiste à relier tous les points d'un échantillon à l'aide de triangles tel qu'aucun point, autre que ceux des sommets du triangle, ne se retrouve à l'intérieur de leur cercle circonscrit. La partie gauche de la figure 3.1 montre un exemple de triangulation de Delaunay. On peut décrire simplement le diagramme de Voronoï à partir de la triangulation de Delaunay. En fait le diagramme de Voronoï représente le graphe dual de la triangulation de Delaunay c'est-à-dire qu'il est défini à partir des éléments de cette dernière méthode. Les sommets du diagramme de Voronoï correspondent au centre des cercles circonscrits des triangles de la triangulation de Delaunay. La partie centrale de la figure 3.1 illustre ce propos. Enfin la partie droite de la figure 3.1 montre un diagramme de Voronoï. Il existe plusieurs méthodes de construction de ce diagramme dont l'une des plus connues provenant de [Shamos & Hoey \(1975\)](#).

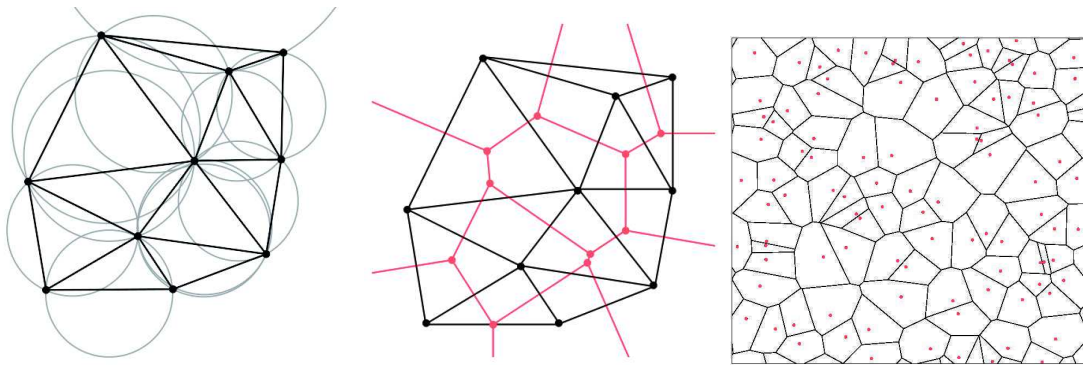


FIGURE 3.1: Diagramme de Voronoï et triangulation de Delaunay. Gauche : Construction de la triangulation de Delaunay. Aucune source ne doit se trouver à l'intérieur des cercles circonscrits. Centre : Lien entre la triangulation de Delaunay (noir) et le diagramme de Voronoï (rouge). Les points rouges sont à la fois les sommets du diagramme de Voronoï et le centre des cercles circonscrits. Droite : Diagramme de Voronoï final d'un échantillon de points (rouge).

Cette méthode permet d'entourer chaque source de l'échantillon par un polygone. Plus l'aire du polygone est petite et plus la densité locale est importante. On peut ainsi remonter à la densité locale en inversant l'aire du polygone. En appliquant un certain seuil, comme pour les deux méthodes mentionnées dans les sous-sections précédentes, il est possible de définir des sur-densités. Cependant la méthode est sensible aux petites variations de densité. [González & Padilla \(2009\)](#) lisent la densité estimée en prenant en compte les cellules voisines dans le but d'obtenir une meilleure mesure de la densité locale.

Cette méthode a été principalement utilisée dans le domaine de l'astrophysique pour trouver des amas de galaxies. Quelques études d'amas d'étoiles ont été menées à l'aide de cette méthode ([Espinoza et al., 2009](#); [Kuhn et al., 2015](#)).

3.1.1.4 Minimum Spanning Tree

Enfin, la méthode des arbres couvrants de poids minimal, Minimum Spanning Tree en anglais, abrégé MST, est une méthode faisant partie intégrante de la théorie des graphes

comme la méthode des cellules de Voronoï. MST suit le même principe que cette dernière mais en cherchant à relier chacun des points de la distribution au lieu de chercher à les isoler par des cellules. La base de cette méthode est d'utiliser la triangulation de Delaunay afin de connecter chaque source, appelé sommet comme pour le diagramme de Voronoï, avec des segments, appelés ici branches, sans créer de boucle et en minimisant le poids total des branches. Dans le cas où toutes les sources ont un poids égal, la méthode cherche à minimiser la taille totale des branches. La figure 3.2 montre un exemple de résultat du MST, superposé à un graphe de type de la triangulation de Delaunay. Dans le cas particulier où toutes les sources possèdent le même poids, il existe plusieurs MST pour le même échantillon. Cependant dans un cas réel comme notre étude, où nous utilisons les positions des sources, la solution est unique.

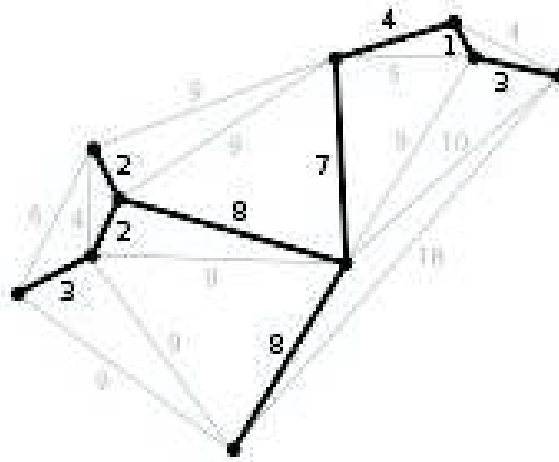


FIGURE 3.2: Exemple d'un graphe et du MST associé. Le poids attribué est proportionnel aux distances entre les points.

Le premier algorithme fut décrit dans [Borůvka \(1926a,b\)](#) dont une traduction peut être trouvée dans [Nešetřil & Nešetřilová \(2012\)](#). Plusieurs algorithmes ont ensuite été développés soit à une même époque par des chercheurs indépendants, soit à des dates ultérieures afin d'augmenter la rapidité d'exécution des premiers algorithmes. [Graham & Hell \(1985\)](#) décrit plusieurs algorithmes dont celui utilisé dans notre étude, l'algorithme de Prim ([Prim, 1957](#)).

Les premières utilisations dans le domaine de l'astrophysique ont été faites dans le cadre d'études des amas de galaxies et notamment dans le contexte des filaments aux grandes échelles ([Barrow et al., 1985](#); [Colberg, 2007](#)). Plus tard, cette méthode fut appliquée aux amas stellaires comme dans les études de [Koenig et al. \(2008\)](#), [Gutermuth et al. \(2009\)](#), [Billot et al. \(2011\)](#) et [SaraI et al. \(2015\)](#). Elle fut aussi récemment utilisée pour observer la ségrégation de masse dans les amas d'étoiles comme dans les études de [Allison et al. \(2009\)](#), [Maschberger & Clarke \(2011\)](#) et [Parker et al. \(2012\)](#), où encore pour étudier le profil de distribution des étoiles au sein d'amas ([Cartwright & Whitworth, 2004](#)).

3.1.2 Choix de la méthode MST

[Schmeja \(2011\)](#) a comparé les quatre méthodes citées précédemment et montre que chacune des méthodes détectent différents types d'amas et dépend donc de la nature

des données. Il montre notamment que la méthode MST est adaptée pour les amas ayant une densité largement supérieure à la densité du fond.

Comme nous l'avons vu dans le chapitre 1, l'analyse des images Herschel montre clairement des structures filamenteuses. Certaines sources extraites de ces images ont tendance à suivre le profil de ces filaments (Molinari et al., 2010a). Lorsqu'on compare les méthodes entre elles, on remarque que la méthode la plus adaptée pour extraire des sur-densités constituées de filaments est la méthode MST que nous utiliserons ainsi dans la suite de notre étude.

Il existe cependant plusieurs méthodes afin d'extraire les sur-densités. Leur point commun est de définir une longueur de branche de coupure, Λ_{cut} , pour laquelle toutes les branches ayant $\Lambda \leq \Lambda_{cut}$ sont considérées comme faisant partie d'une sur-densité (Barrow et al., 1985). Λ_{cut} est défini différemment suivant les études. Ainsi Barrow et al. (1985) et Campana et al. (2008) utilisent la valeur moyenne des branches, $\Lambda_{cut} = \langle \Lambda \rangle$, ou un de ces multiples. Battinelli (1991) utilise plusieurs valeurs de Λ_{cut} et décrit la valeur adéquate comme étant celle permettant l'extraction du plus grand nombre de sur-densités. Maschberger et al. (2010) utilise une procédure quasiment similaire, à la différence que les sur-densités choisies sont celles présentant des propriétés similaires à celles sélectionnées visuellement. Enfin, Koenig et al. (2008) et Gutermuth et al. (2009) procédèrent à une autre méthode. Celle-ci consiste à construire un histogramme et sa distribution cumulative de toutes les tailles de branches. Il s'agit ensuite d'ajuster un segment aux petites tailles de branches ainsi qu'aux grandes tailles. Enfin l'intersection de ces deux segments permet d'obtenir la taille de coupure. C'est cette méthode que nous avons retenu pour notre étude car la valeur de Λ_{cut} est obtenue automatiquement. Un exemple du fonctionnement de cette méthode, appliquée aux données de notre étude, est donné dans les figure 3.3, 3.4 et 3.5. la figure 3.3 représente la distribution des clumps Hi-GAL pour un champ de 2 degrés carrés, centré à $l = 330$ degrés. La figure 3.4 montre ensuite la méthode d'évaluation de Λ_{cut} . La distribution cumulative (en noir) est calculée à partir de la distribution des tailles des branches (en gris). L'intersection des deux segments ajustés correspond ici à une valeur $\Lambda_{cut} = 160''$. Nous verrons dans la section suivante que notre étude se base sur une distribution en deux dimensions, c'est-à-dire que nous utilisons les distances angulaires entre les clumps. Le résultat de cet exemple est montré au travers de la figure 3.5. Une condition nécessaire pour que le résultat ne soit pas sensible aux fluctuations statistiques de la densité est de contraindre le nombre minimum de clumps, N , présents dans les sur-densités. Gutermuth et al. (2009) montre que N doit être au moins supérieur à 6 mais une valeur minimum de 10 est préférable. Ainsi nous suivons cette recommandation dans la suite de notre étude.

3.2 Construction du catalogue d'amas

Nous allons maintenant décrire les différentes étapes de construction du catalogue final d'amas de clumps. Il a d'abord fallu évaluer les contraintes imposées par la nature des données et la structure du plan galactique. La deuxième partie consiste à extraire les amas fiables en utilisant des critères de distance héliocentrique.

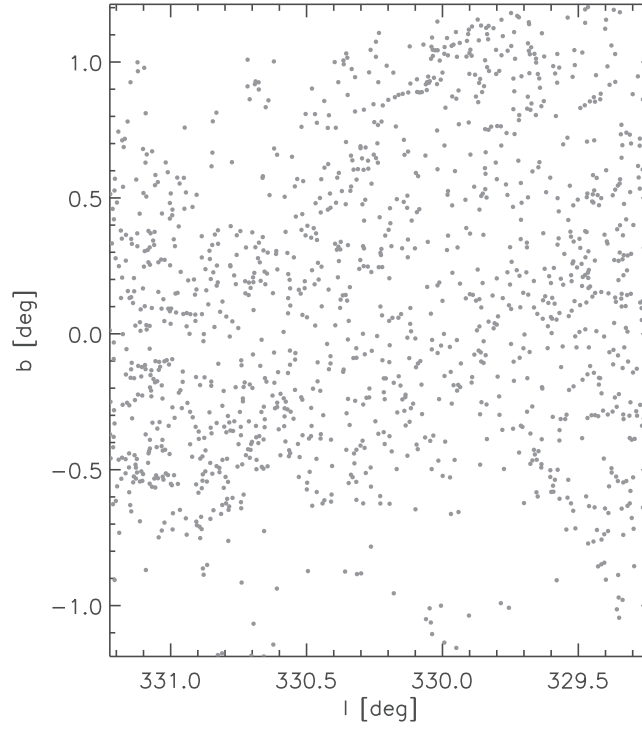


FIGURE 3.3: Distribution des clumps dans un champ de 2 degrés carrés, centré à $l = 330$ degrés.

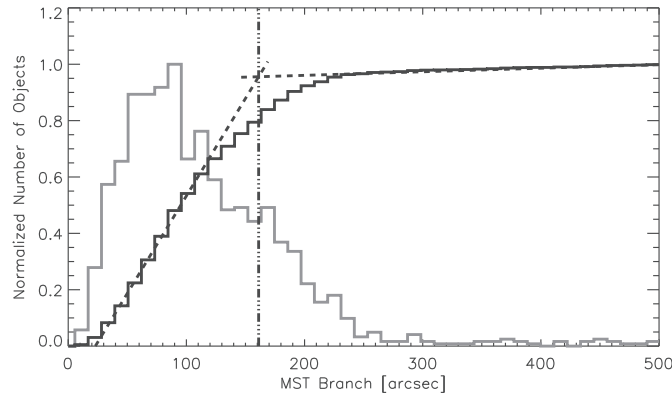


FIGURE 3.4: Distribution des tailles de branches (en gris) correspondant au champ de la figure 3.3. La distribution cumulative est représentée en noir. Les deux segments ajustés à cette distribution cumulative permettent d'obtenir la valeur de Λ_{cut} (en pointillé).

3.2.1 Stades évolutifs des clumps et confusion

La première étape de la construction d'un catalogue d'amas est de savoir si les données ont une distribution spatiale suffisamment hétérogène pour extraire des sur-densités. Pour rappel, nous avons vu dans le chapitre précédent que les clumps sont divisés en trois stades évolutifs dans le catalogue *product*, unbound, pré-stellaire et proto-stellaire. En construisant une carte de densité pour chaque type de clumps, nous pouvons observer leurs distributions spatiales et les comparer entre elles. Les figures 3.6, 3.7 et 3.8 montrent un exemple de cartes de densité pour chaque type de clumps, calculées pour la région localisée entre $26 \leq l \leq 31$ degrés. On peut observer que les clumps unbound ne sont que très peu présents au centre du plan galactique ($-0.5 \leq b \leq 0.5$ degrés). On observe la distribution inverse pour les clumps proto-stellaires se qui laisse présumer

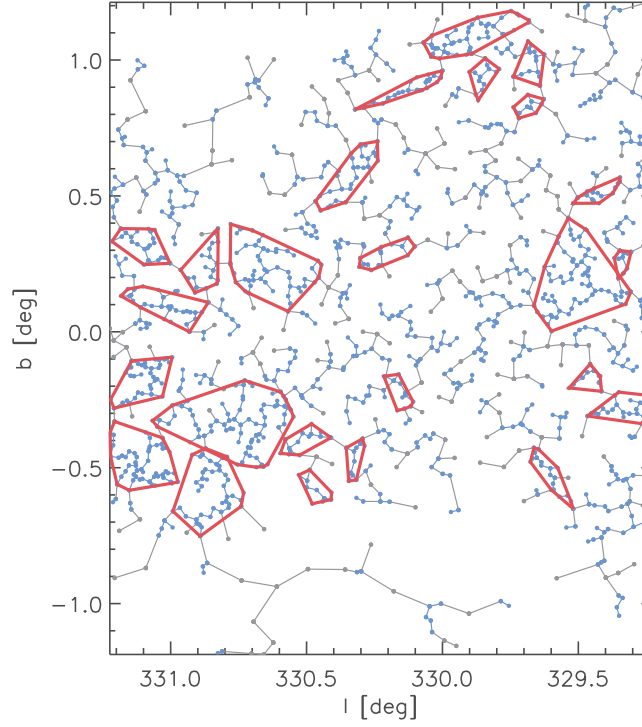


FIGURE 3.5: Résultat du MST en suivant la méthode de [Koenig et al. \(2008\)](#) et [Gutermuth et al. \(2009\)](#), appliquée sur l'échantillon de la figure figure 3.3. Les points et lignes gris correspondent aux clumps rejetés ($\Lambda > \Lambda_{cut}$) alors que les lignes et points bleus correspondent à $\Lambda < \Lambda_{cut}$. Les lignes rouges encadrent les sur-densités avec $N \geq 10$.

que les deux distributions sont liées. En effet, les clumps proto-stellaires sont des objets très brillants en raison du chauffage de leur partie interne. Il se trouve ainsi que la luminosité du fond dans leur environnement proche est plus élevée que dans des régions sans clumps proto-stellaires ce qui a pour effet d'augmenter la complétude en flux dans ces zones. Il en résulte que les clumps unbound, étant froids et peu lumineux, ne sont pas aussi facilement extraits que les clumps proto-stellaires. En plus de cette limitation, la complétude augmente aussi avec la ligne de visée. Plus les sources sont proches du centre du plan galactique et plus le nombre d'objets sur la ligne de visée est important. Cet effet est d'autant plus important que les sources sont froides et faibles.

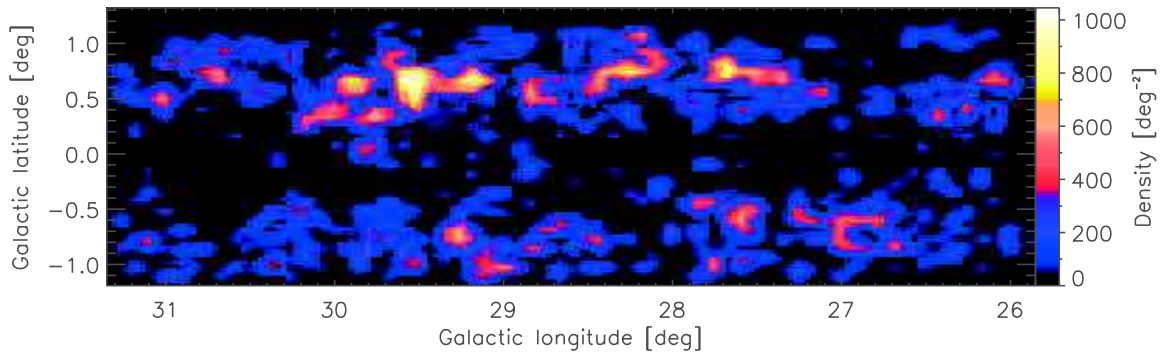


FIGURE 3.6: Carte de densité des clumps unbound pour une région localisée entre $26 \leq l \leq 31$ degrés. La même distribution est observée dans l'échantillon de données total.

Les clumps pré-stellaires ne semblent pas être affectés par ces effets. Leur distribu-

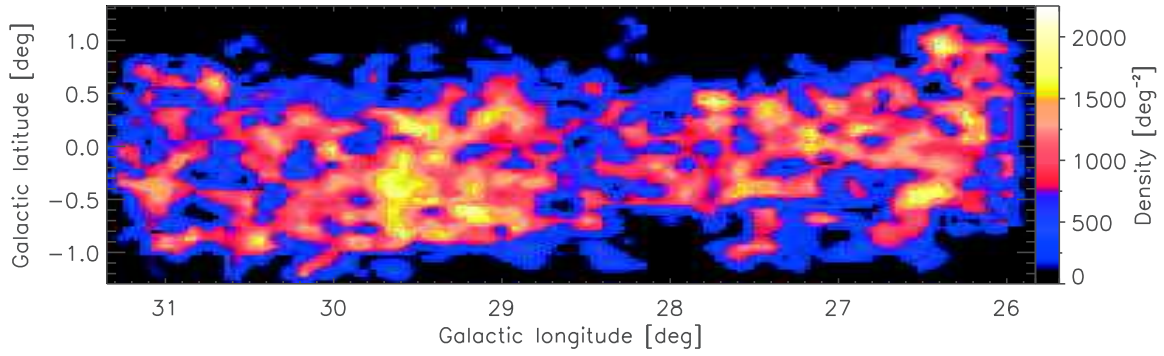


FIGURE 3.7: Même représentation que la figure 3.6 pour les clumps pré-stellaires.

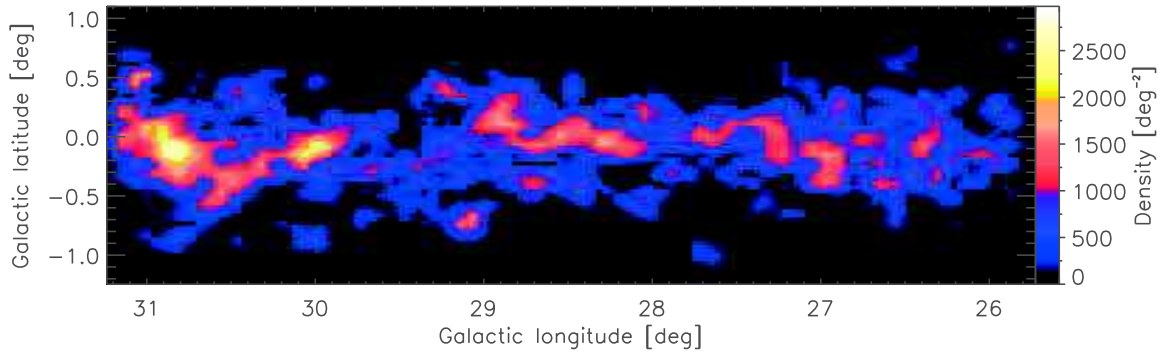


FIGURE 3.8: Même représentation que la figure 3.6 pour les clumps proto-stellaires.

tion est la plus homogène de toutes même si des sur-densités se dessinent à "vu d'oeil". La distribution la plus hétérogène est celle des clumps pro-stellaires. Nous pourrions lier l'évolution des clumps à leur degré de concentration en supposant que ces deux distributions ne sont pas affectées par les effets de confusion. Cependant le catalogue comprend quasiment deux fois plus de clumps pré-stellaires que de clumps proto-stellaires ce qui semble être élevé compte tenu de la durée de chaque stade. En effet, le stade pré-stellaire ne dure qu'environ 100 000 ans alors que le stade proto-stellaire peut durer quelques millions d'années (un à deux millions d'années pour les classes 0/I). Il est ainsi probable que certains clumps pré-stellaires soient en réalité des clumps proto-stellaires. Pour cette raison nous ne séparerons pas ces deux types de clumps dans tout le reste de l'étude. Enfin, la construction du catalogue ne se fera qu'avec les clumps pre et proto-stellaires puisque la distribution spatiale des clumps unbound semble être perturbée par les autres types de clumps. Sur les 99 083 clumps, on compte 81 659 clumps pré-stellaires et proto-stellaires.

3.2.2 Contraintes imposées par la structure galactique

Toutes les méthodes d'extraction de sur-densités sont sensibles à la densité de sources. Dans le cas de la méthode MST utilisée avec les conditions énoncées dans la section 1.2, la détermination de Λ_{cut} est associée à la densité moyenne de la région sondée. Plus la densité moyenne est élevée et plus la valeur de Λ_{cut} est petite. La figure 3.9 montre la relation entre Λ_{cut} et la densité moyenne dans le cadre de notre étude. Les points représentent différentes solutions de la méthode MST suivant la taille de la fenêtre, c'est-à-dire l'aire correspond au sous échantillon considéré, ainsi que suivant différentes longitudes. L'ensemble des points parcourt l'ensemble de la distribution spatiale des clumps. On peut observer une corrélation très marquée entre ces deux quantités, bien décrite par une loi

de puissance, $\Lambda_{cut} \propto d_{mean}^{-\beta}$ avec une valeur de pente $\beta \approx 0.24 \pm 0.01$. On constate aussi une dispersion des points. Λ_{cut} dépend aussi du contraste entre la densité du fond et les sur-densités. Ainsi pour une même densité moyenne, le contraste peut varier ce qui induit une variation de la distribution des tailles de branches et ainsi des valeurs différentes de Λ_{cut} .

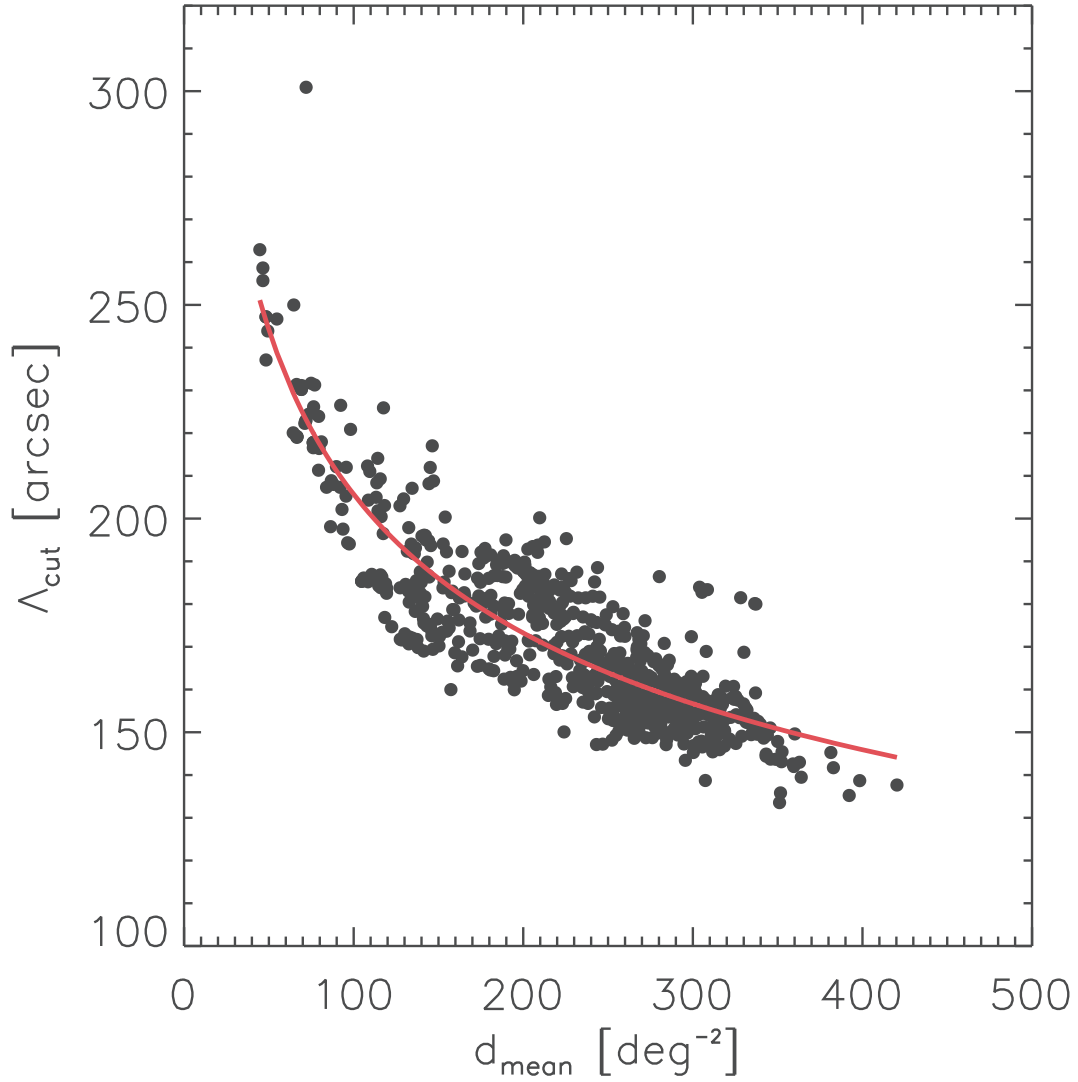


FIGURE 3.9: Relation entre Λ_{cut} et la densité moyenne de la zone considérée. La ligne rouge correspond à la fonction de corrélation $\Lambda_{cut} \propto d_{mean}^{-0.24}$.

Cette corrélation nous indique que la structure galactique a un impact sur la détermination de Λ_{cut} et ainsi sur la détectabilité des sur-densités. Nous savons que la densité varie le long du plan galactique, principalement en raison de la présence de bras spiraux et du centre galactique qui ont tendance à augmenter la densité locale. Comme nous nous intéressons à la distribution dans le plan galactique, entre les latitudes $-2 \leq b \leq 2$ degrés, les variations en latitude sont négligeables.

Cette variation de densité, que l'on pourrait qualifier de physique, nous oblige à diviser le catalogue en plusieurs fenêtres suivant les longitudes. L'avantage de cette technique est d'obtenir une valeur de Λ_{cut} pour chaque zone particulière de la partie du plan galactique sondée. Cependant cela nous oblige aussi à faire chevaucher les différentes fenêtres afin d'éviter de rejeter des clumps appartenant à des sur-densités, voir même de diviser des

sur-densités. Pour éviter ce problème, de nouvelles fenêtres sont créées en décalant les fenêtres originelles par la moitié de leur taille en longitude. En posant N_f le nombre de fenêtres originelles, nous devons obtenir $N_f - 1$ fenêtres chevauchantes avec une valeur supplémentaire de Λ_{cut} pour chacune d'entre elles. Cette technique pose un problème puisque maintenant pratiquement toutes les régions sondées possèdent deux valeurs de Λ_{cut} . Seul les deux moitiés de fenêtres à chaque extrémité ne sont pas affectées par cette technique. Pour éliminer ce problème, nous considérons que la valeur de Λ_{cut} à privilégier pour une sur-densité donnée est celle où cette dernière est la plus éloignée des extrémités de la fenêtre à laquelle elle appartient. En d'autre terme nous prenons seulement en compte la moitié centrée de chaque fenêtre (hormis pour les extrémités). Cela permet de considérer les sur-densités lorsqu'elles sont au milieu de leur environnement et non pas celui d'une structure voisine.

La figure 3.10 permet de mieux comprendre cette technique de sélection de la fenêtre adéquate pour chaque sur-densité. Cette figure montre un cas réel d'application sur les données de notre étude. Tout d'abord la fenêtre du dessus correspond à la fenêtre de chevauchement. Les deux fenêtres du milieu, accolées l'une à l'autre, sont les fenêtres originelles. Enfin, la grande fenêtre de dessous correspond à la fenêtre finale, c'est-à-dire le catalogue reconstruit à ces longitudes. Dans l'ensemble des fenêtres les polygones gris correspondent aux sur-densités avec au moins 10 clumps. Les polygones rouges sont les sur-densités divisées en deux à cause du partitionnement en longitude. Les rectangles noirs (lignes pointillées et pleines) sont les régions utilisées pour reconstruire le catalogue.

Le partitionnement était un premier problème pour la construction d'un catalogue d'amas. Le deuxième problème consiste à choisir une taille de fenêtre idéale correspondant aux échelles de variations de densité des données. La méthode retenue dans notre étude consiste à comparer différentes tailles de fenêtre, de 2 degrés à 30 degrés par pas de un degré. La taille minimum, de 2 degrés, a été choisie pour correspondre à la taille des champs de démonstration du relevé Hi-GAL (SDPFs, [Veneziani et al., 2013b](#)). Elle est aussi plus grande que la taille typique des sur-densités de notre étude, comme montré dans la figure 3.4, ce qui évite d'avoir deux fenêtres de chevauchement pour une sur-densité. De plus cette taille permet de garder suffisamment de clumps dans une même fenêtre afin d'obtenir une bonne estimation de Λ_{cut} . Enfin, la taille maximum, de 30 degrés, a été choisie afin d'obtenir une fenêtre aussi large que les plus grandes structures apparentes, c'est-à-dire les tangentes des bras spiraux.

Il est possible, à partir de ces différentes tailles de fenêtre, d'évaluer si les clumps ne sont pas distribués aléatoirement. [Campana et al. \(2008\)](#) montrent que la taille totale des branches du MST est proportionnelle à $\sqrt{AN_{tot}}$ où A est l'aire de la fenêtre et N_{tot} le nombre total de clumps dans celle-ci. Ils montrent aussi que la taille moyenne des branches est proportionnelle à $\sqrt{A/N_{tot}}$. Afin de savoir si la distribution n'est pas aléatoire, il faut la comparer avec une distribution aléatoire simulée. Pour la taille totale et la taille moyenne des branches, [Campana et al. \(2008\)](#) trouvent un coefficient de ≈ 0.65 , réalisé à partir de simulation Monte-Carlo, ce qui constitue une valeur limite supérieure. En prenant en compte toutes les tailles de fenêtres et toutes les longitudes, nous trouvons des coefficients respectifs de $\approx 0.51 \pm 0.01$ et $\approx 0.40 \pm 0.01$ pour la taille totale et moyenne des branches. Ces valeurs sont significativement plus basses que celles attendues pour un champ aléatoire ce qui démontre que les clumps ne sont pas distribués aléatoirement. Comme nous sommes en présence de véritables sur-densités (non aléatoires) indépendamment de la taille des fenêtres, il est possible de toutes les comparer entre elles.

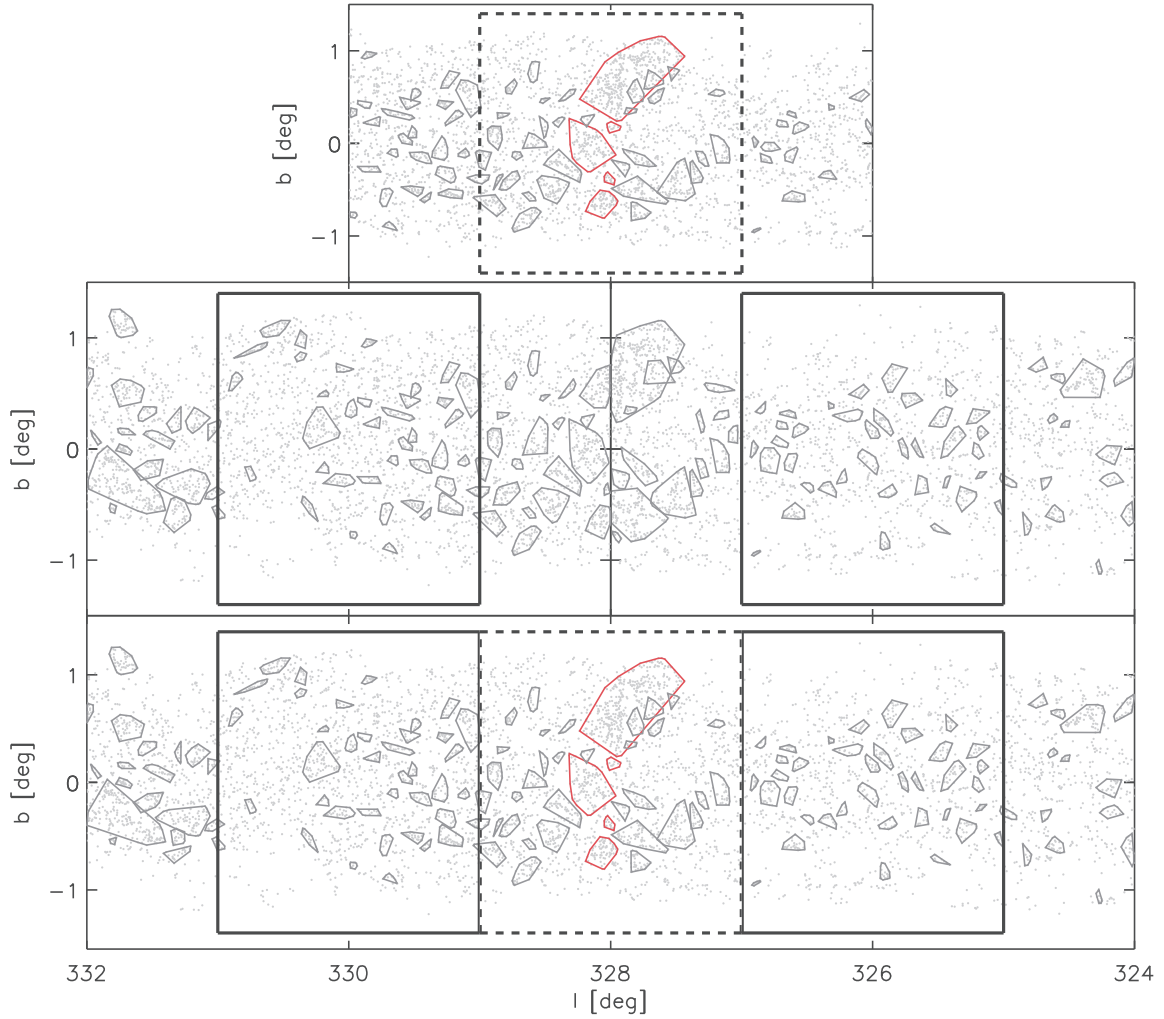


FIGURE 3.10: Exemple de reconstruction d'un catalogue de sur-densités. La taille de fenêtre correspond à 10 degrés. Les points gris correspondent aux clumps pré-stellaires et proto-stellaires. Les polygones gris sont les sur-densités possédant aux moins 10 clumps. Les polygones rouges sont les sur-densités divisées en deux par les fenêtres originelles. Les rectangles noirs (lignes pointillées et pleines) sont les régions utilisées pour reconstruire le catalogue. La fenêtre du dessus est la fenêtre de chevauchement de cette région. Les deux fenêtres du milieu sont les fenêtres originelles et celle en dessous est la fenêtre finale.

La figure 3.11 montre un test de l'effet de la taille de fenêtre en fonction de la longitude. La courbe représente la valeur moyenne de la fraction de clumps dans les sur-densités calculée à partir des 29 catalogues construits, un pour chaque taille de fenêtre, en fonction de la longitude et en utilisant un bin de 5 degrés. La barre d'erreur correspond à l'erreur à 1σ de la fraction de clumps dans les sur-densités. Les deux lignes pointillées correspondent à la même distribution mais pour une variation de valeur de $\pm 5\%$ de Λ_{cut} . Cette variation correspond approximativement à $10''$ et représente une variation de 10 à 15% (suivant les longitudes) du nombre de clumps dans les sur-densités. La figure montre qu'il n'y a donc pas de variations importantes causées par l'utilisation de différentes tailles de fenêtre, sauf aux extrémités de la distribution ce qui correspond à un effet de bord.

Afin de choisir une taille de fenêtre et comme la technique précédente ne semble pas fonctionner, nous pouvons étudier des régions particulières sans moyenner la fraction de

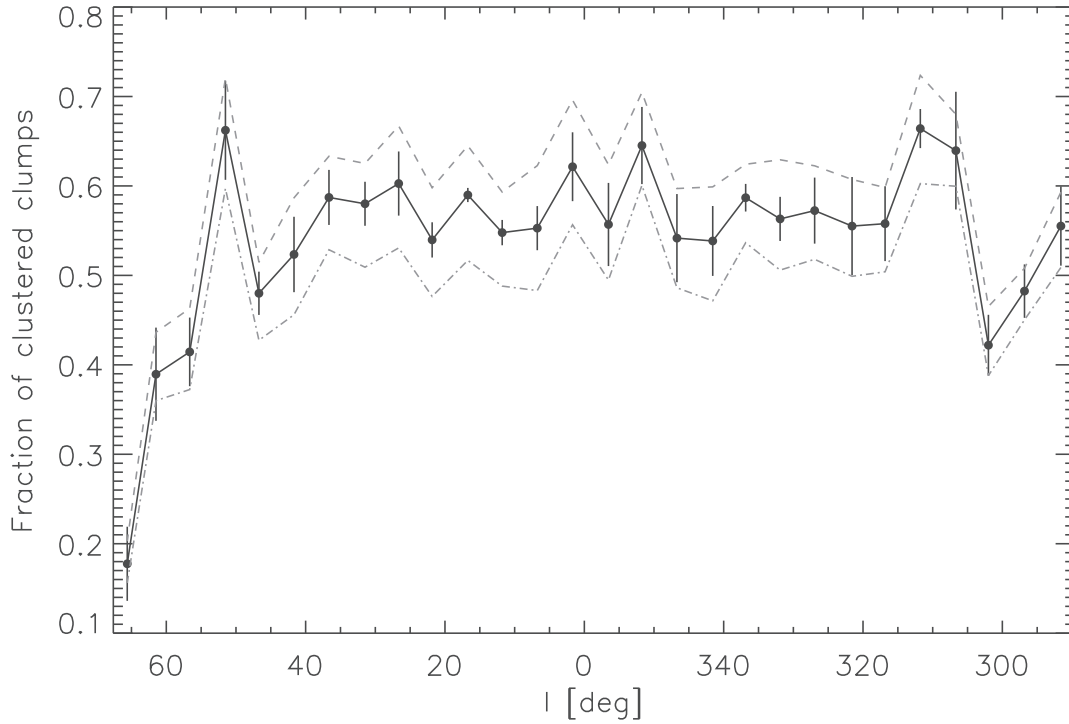


FIGURE 3.11: Distribution de la fraction moyenne des clumps dans les sur-densités en fonction de la longitude. La ligne noire correspond à la fraction moyenne des clumps dans les sur-densités pour chaque catalogue (un pour chaque taille de fenêtre). La taille de bin est définie à 5 degrés. La barre d'erreur correspond à l'erreur à 1σ de la fraction de clumps dans les sur-densités. Les deux lignes pointillées correspondent à la même distribution mais pour une variation de valeur de $\pm 5\%$ de Λ_{cut} .

clumps appartenant à des sur-densités. Pour rappel, seul les sur-densités avec plus de 10 clumps sont considérées. La figure 3.12 montre la fraction de clumps dans des sur-densités en fonction de la taille des fenêtres pour deux régions en particulier de 2 degrés carrés, l'une centrée à $l = 335$ degrés (partie gauche) et l'autre à $l = 20$ degrés (partie droite). On remarque que la distribution est différente entre les deux régions, celle à $l = 335$ degrés possédant un plus grand nombre de clumps dans les sur-densités que pour l'autre région. Dans les deux cas il existe un intervalle de tailles où la fraction ne varie pas. Cependant, pour la région à $l = 335$ degrés, une taille de fenêtre supérieure à 20 degrés provoque un saut significatif dans la valeur de la fraction. Cette fraction est influencée par Λ_{cut} ainsi que le nombre de clumps et leur distribution.

Il est aussi utile d'éliminer les variations statistiques observées dans les trop petites tailles de fenêtre. Ainsi la fenêtre de 2 degrés correspond à ce critère et n'est donc pas retenue pour la suite de l'étude. Afin de terminer la sélection de la taille de fenêtre adéquate, nous pouvons comparer la distribution de Λ_{cut} suivant la longitude et pour toutes les tailles de fenêtre. La figure 3.13 montre quatre tailles de fenêtre, 2, 10, 20 et 30 degrés. On remarque qu'une grande taille de fenêtre, typiquement ≥ 20 degrés, induit une distribution trop lissée de Λ_{cut} . Il est alors impossible d'observer les structures galactiques.

Au final, en comparant les figures 3.12 et 3.13, nous pouvons définir la meilleur taille de fenêtre comme étant comprise entre 10 et 15 degrés. Nous avons décidé de choisir 10 degrés dans la suite de notre étude. La figure 3.14 montre la distribution du nombre de sur-densités ayant au moins deux clumps, c'est-à-dire à partir du plus petit nombre

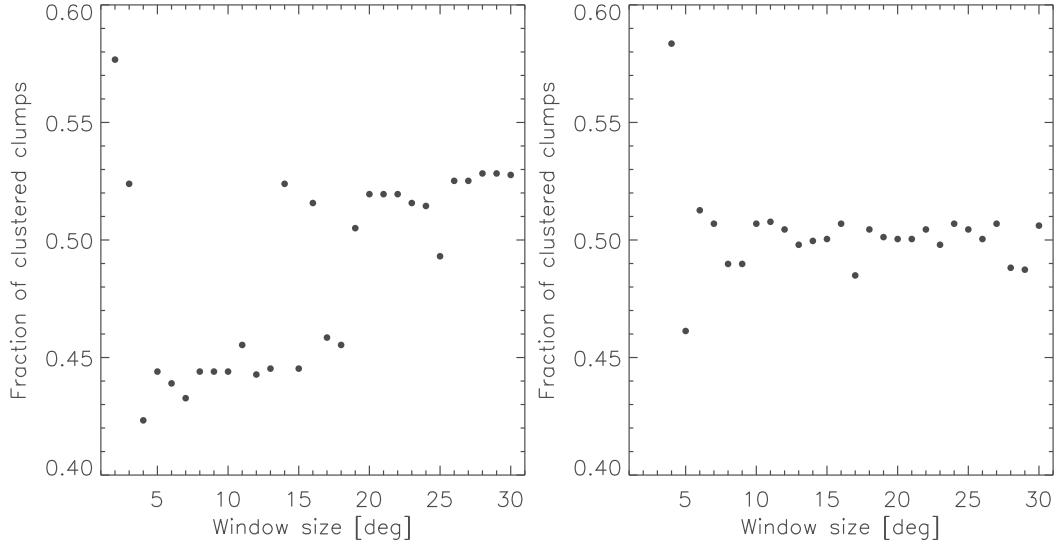


FIGURE 3.12: Fraction de clumps dans les sur-densités en fonction de la taille des fenêtres, pour un champ de 2 degrés carrés centré sur $l = 335$ degrés (gauche) et $l = 20$ degrés (droite).

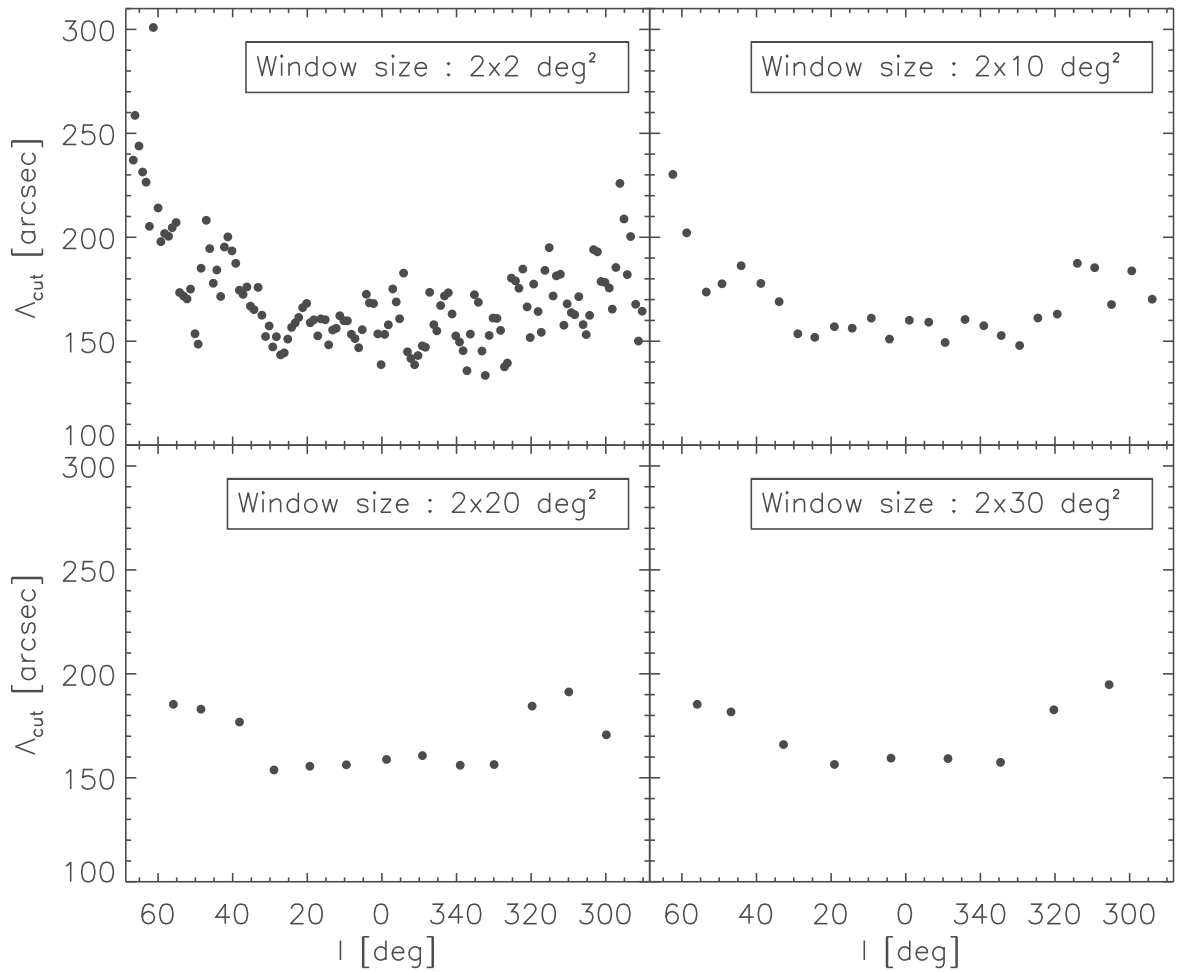


FIGURE 3.13: Distribution de Λ_{cut} en fonction de la longitude pour quatre tailles de fenêtre, 2, 10, 20 et 30 degrés. La longitude correspond au centre de chaque fenêtre.

d'associations possibles. Beaucoup de sur-densités possèdent entre 8 et 10 clumps. Cependant nous allons voir dans la sous-section suivante que le nombre minimum de 10

clumps a été choisi afin de satisfaire certains critères.

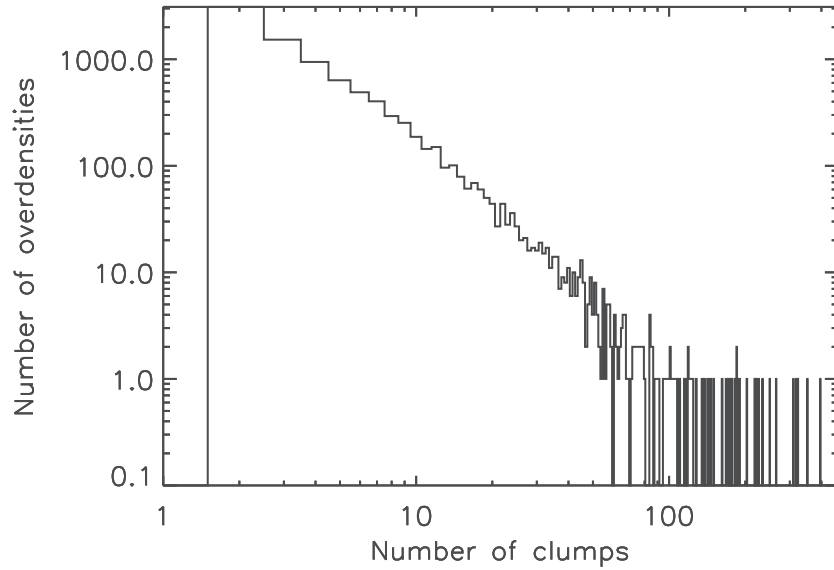


FIGURE 3.14: Distribution du nombre de clumps dans les sur-densités, à partir de deux clumps.

Le catalogue final se compose de 1705 sur-densités possédant au moins 10 clumps et correspond à 45 323 clumps (soit 55% du nombre total de clumps). Nous allons maintenant évoquer la dernière étape de la construction du catalogue, qui nous permet de différencier les sur-densités des amas, à savoir l'utilisation des distances héliocentriques.

3.2.3 Distances héliocentriques et catalogue d'amas fiables

Les 1705 sur-densités représentent des distributions de clumps en deux dimensions. Cependant, comme la galaxie est une structure en trois dimensions, une troisième coordonnée serait préférable. Elle permettrait entre autre d'éliminer les sur-densités avec une confusion importante sur la ligne de visée. Cette sélection peut justement être faite en utilisant les estimations de distances héliocentriques, abrégées HDEs dans la suite, du catalogue *product* de clumps. Ce sont des distances cinématiques ce qui induit 2 valeurs différentes dans la partie interne de la galaxie. Afin de choisir la distance proche ou lointaine, représentative de la région sondée, plusieurs autres données sont utilisées en complément, incluant une carte d'extinction, les vitesses du CO, les vitesses du NH 3, une courbe de rotation de la galaxie, des données HI ainsi que des images dans le domaine optique et infrarouge. De plus, lorsque des clumps sont associés avec une région dont la distance héliocentrique est connue, celle-ci leur est attribuée. [Russeil et al. \(2011\)](#) développent un exemple de calcul de distances cinématiques dans les SDPFs du relevé Hi-GAL.

Sur les 99 083 clumps contenus dans le catalogue *product*, 57% possèdent des HDEs. Sur les 83 851 clumps pré-stellaires et proto-stellaires, 59% ont une HDE. Cependant les HDEs n'ont pas été obtenues pour les sources entre $l = -10$ degrés et $l = 14$ degrés. En effet, dans les régions proches du centre galactique, les courbes de rotation nécessaires au calcul des distances ne sont pas fiables et les vitesses obtenues sont sujettes à une confusion sur la ligne de visée due respectivement à des perturbations gravitationnelles

importantes et à une forte densité de sources dans cette zone. De plus, une petite variation de vitesse radiale entraîne une grande variation de distances dans cette zone. En évitant cete zone, il reste 74 085 clumps et on constate que 67% (49 373) des clumps pré-stellaires et proto-stellaires ont une HDE.

La première possibilité pour construire un catalogue fiable d'amas à partir de ces distances est d'utiliser la méthode MST en trois dimensions. Cependant les incertitudes sur les distances sont relativement importantes comme montré par [Russeil et al. \(2011\)](#) dans les SDPFs, avec des erreurs respectives de 0.6 kpc et 0.9 kpc à $l = 30$ degrés et $l = 59$ degrés. Cette raison empêche une détection fiable des sur-densités en trois dimensions car l'algorithme cherche à relier les sources les plus proches les unes des autres et ainsi il risque de relier des sources d'une source-densité potentielle à celles d'autres, toutes situées dans la même zone tri-dimensionnelle. Il existe aussi une deuxième limite à cette méthode relative aux clumps sans HDEs. En effet, les clumps ne possédant pas de distances constituent aussi une part importante des sur-densités et peuvent aussi renfermer des informations requises, notamment, pour l'étude plus approfondie des amas. Il faut cependant définir des critères de sélection des sur-densités fiables, à partir des HDEs, afin de réduire le risque de confusion lié aux clumps sans HDEs.

Ceci nous amène ainsi à l'étape de différenciation des sur-densités fiables et non fiables. Les HDEs sont utilisées en tant que troisième dimension afin d'évaluer la compacité des sur-densités, c'est-à-dire leur capacité à former une région dense localisée dans l'espace. Celle-ci est calculée en utilisant l'histogramme des HDEs pour chaque sur-densité. On assume qu'un clump appartient à une sur-densité si son HDE est compris dans un intervalle de distance de ± 1 kpc à partir du pic de l'histogramme. Cette valeur correspond aux incertitudes sur les distances cinématiques mentionnées plus haut. Nous utiliserons l'acronyme CDR, pour *Compact Distance Range*, afin de désigner cet intervalle de distance dans la suite de cette section.

Deux rapports sont calculés pour évaluer la compacité des sur-densités. Le premier correspond au pourcentage de clumps possédant des HDEs dans les sur-densités. Le deuxième est le rapport entre le nombre de clumps dans le CDR et le nombre total de clumps avec des HDEs. Nous voulons maximiser le nombre de clumps avec HDEs sans pour autant restreindre le résultat aux sur-densités ne possédant que des clumps avec des HDEs. Pour cela nous considérons les sur-densités comme éligible au rang de sur-densité réel si au moins 50% des clumps possèdent une HDE et si 75% d'entre eux sont dans le CDR. Ces critères expliquent pourquoi nous avons choisi un nombre minimum de 10 clumps appartenant à une sur-densité. En dessous de 10 clumps, il devient difficile d'utiliser ces critères car le nombre de clumps avec des HDEs devient en général trop petit pour construire un histogramme de distances.

Sur les 1705 sur-densités, 607 satisfont ces critères. Cependant, seulement 219 d'entre ces dernières possèdent des clumps à la même distance héliocentrique, en accord avec le CDR. Pourtant, ces sur-densités sont considérées comme les plus fiables. Afin d'augmenter leur nombre, nous avons décidé d'éliminer les quelques clumps qui ne semblent pas faire partie de ces 607 sur-densités, en accord avec la définition du CDR, et de recalculer une solution à partir de la méthode MST. Outre la possibilité d'éliminer une part de la confusion créée par la projection du plan galactique en deux dimensions, cette procédure nous permet aussi de tester la méthode MST dans le but d'observer le nombre de sur-densités retrouvées avec un échantillon légèrement différent.

Cette procédure enlève 1531 clumps sur les 74 085 clumps de l'échantillon total possé-

dant des HDEs. À partir des 82 320 clumps restants (avec et sans HDEs), nous recalculons donc une solution de la méthode MST à partir de tailles de fenêtre de 10 degrés. Comme développé dans la sous-section précédente, la nouvelle analyse peut toujours être considérée comme une distribution non-aléatoire puisque la taille moyenne des branches est proportionnelle à $\sqrt{A/N_{tot}}$ avec un coefficient de proportionnalité de 0.36 ± 0.01 contre 0.65 trouvé par [Campana et al. \(2008\)](#) dans le cas d'une distribution aléatoire. 1633 sur-densités ont été trouvées avec cette nouvelle compilation. 449 d'entre elles n'ont pas de HDEs et sont principalement localisées vers le centre galactique où aucunes HDEs ne sont fournies. 648 ont une distribution de HDEs confuse c'est-à-dire qu'il est difficile d'établir une valeur unique de distances soit parce qu'il existe plusieurs pics dans la même distribution situés à des distances différentes (écart supérieur au CDR), soit parce qu'il est impossible de définir un pic dans la distribution. Enfin 536 satisfont les critères de compacité mentionnés plus haut. Le résultat est encore meilleur puisque sur les 536 sur-densités, 496 d'entre elles possèdent tous leurs clumps dans le CDR. De plus, ces 536 sur-densités correspondent aux sur-densités trouvées lors de la première compilation de la méthode MST.

Ce sont ces 1633 sur-densités qui définissent l'échantillon final et qui nous permettent de définir deux catalogues de sur-densités que nous allons définir dans la section suivante.

3.3 Caractéristiques du catalogue d'amas

Comme résultat du travail de la section 3.2, nous avons créé deux catalogues de sur-densités. Le premier comprend les sur-densités les plus fiables, c'est-à-dire les 496 des 536 qui satisfont les critères de compacité. Les 40 sur-densités restantes sont associées aux 648 sur-densités qui possèdent une distribution de HDEs confuse. Ces 688 sur-densités ainsi que les 449 sans HDEs composent le deuxième catalogue.

Afin de distinguer les deux catalogues, nous appelons le premier catalogue : catalogue d'amas fiables. Le deuxième catalogue est nommé : catalogue archive. Nous utilisons désormais le terme d'amas pour les sur-densités les plus fiables car elles ont une véritable forme en trois dimensions. Ces amas sont même qualifiés d'amas enfouis car ils comprennent des sources très jeunes encore associées avec le milieu interstellaire environnant. En d'autre terme les sources de ces amas n'ont pas encore eu le temps d'effacer les traces de leur formation et sont toujours en train d'accréter du matériel venant du milieu interstellaire. Il est d'ailleurs à noter que la plus grande partie de la masse de ce type d'amas est encore sous forme de gaz et poussières répartis entre les clumps. La définition d'amas enfouis correspond à celle donnée, par exemple, par [Lada & Lada \(2003\)](#) qui étudient l'évolution de ces amas à partir de leur formation dans les GMCs.

Comme le catalogue d'amas fiables est celui permettant d'utiliser les HDEs comme distances héliocentriques des amas, il est possible de définir certaines caractéristiques des amas sans passer par des mesures angulaires, c'est-à-dire projetées dans le plan du ciel. Nous allons ainsi développer ces caractéristiques.

Tout d'abord, 70% des amas ont entre 10 et 20 clumps et seulement 4 amas possèdent plus de 70 clumps. La somme des 496 amas correspond à 9608 clumps ce qui correspond à 12% du total de 82 320 clumps. 8248 (86%) de ces clumps possèdent des HDEs.

La figure 3.15 représente la distribution des distances héliocentriques des amas fiables. Il existe deux façons de les calculer, soit à partir du pic de la distribution des HDEs des clumps pour chaque amas, soit à partir de la moyenne de cette même distribution. La partie haute de la figure 3.15 représente le pic de la distribution des HDEs des clumps. On remarque que la distribution ressemble à celle des clumps du catalogue *product* d'Hi-GAL (figure 2.8). La partie basse représente la corrélation entre les distances héliocentriques (moyenne et pic). Le coefficient de corrélation est égal à 0.99925 ce qui signifie que les distributions des HDEs sont significativement piquées. Dans toute la suite de cette étude (chapitre suivant compris) la distance héliocentrique d'un amas correspondra à la distance du pic de la distribution des HDEs des clumps.

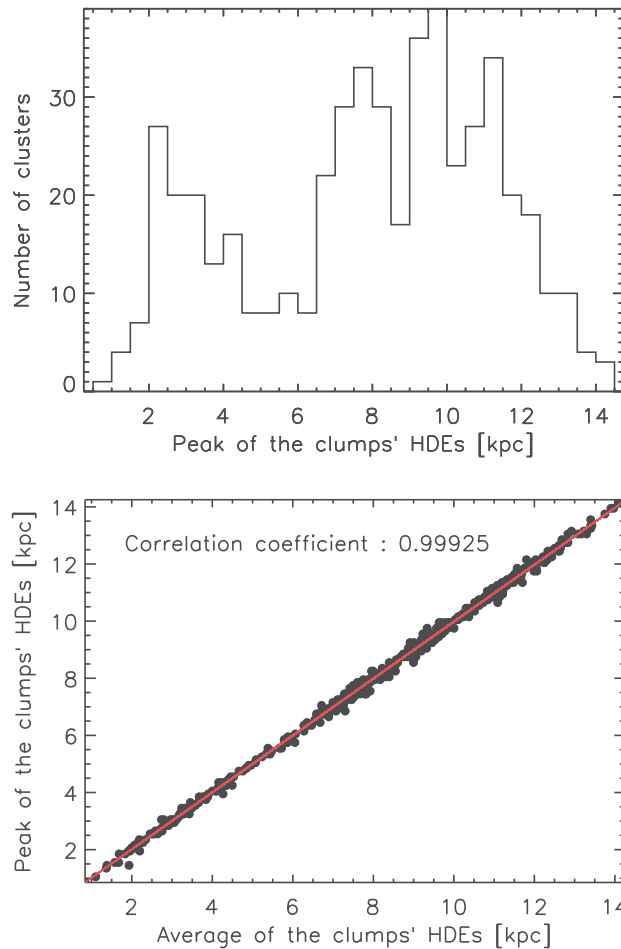


FIGURE 3.15: Distribution des distances héliocentriques des amas les plus fiables. Haut : Distribution des distances calculées à partir du pic de la distribution des HDEs des clumps. Bas : Corrélation entre les distances calculées à partir du pic de la distribution des HDEs des clumps et à partir de leur moyenne. Le coefficient de corrélation vaut 0.99925 ce qui signifie que les distributions de HDEs sont significativement piquées.

À partir de ces distances héliocentriques il est possible de calculer la taille des amas. Pour cela, nous ajustons une ellipse sur chacun des amas. Lorsque l'ajustement échoue nous considérons un cercle ayant pour diamètre la distance entre les deux clumps les plus éloignés l'un de l'autre dans l'amas. La partie haute de la figure 3.16 montre le rayon de l'amas, calculé à l'aide de la formule $r_{eq} = \sqrt{a_{maj}a_{min}}$ lorsque l'amas est ajusté par une ellipse, où a_{maj} représente le grand axe et a_{min} le petit axe. On s'aperçoit que la majorité des amas ont un rayon équivalent inférieur à 10 pc, avec un pic aux alentours de

5 pc. Pour information, une grande région de formation stellaire comme W51, possédant quelques milliers de YSOs (et non clumps), a un demi-grand axe d'environ 25 pc dans l'infrarouge lointain. La partie basse de la figure 3.16 montre le rapport entre le petit axe et le grand axe permettant d'évaluer l'élongation des amas. Sur les 496 amas, 118 ($\approx 24\%$) sont ajustés par un cercle. Le facteur d'élongation le plus important, avec une valeur de 5, caractérise en général les amas composés d'une structure filamentaire.

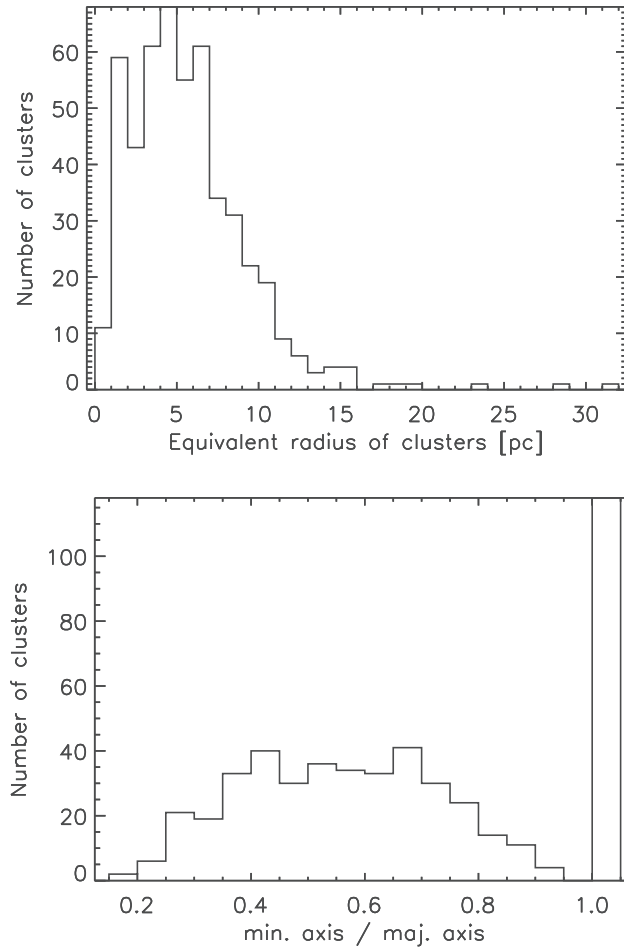


FIGURE 3.16: Haut : Distribution des rayons équivalents des amas calculés à partir des demi-petit axe et demi-grand axe de l'ellipse. Bas : Rapport entre le petit axe et le grand axe de l'ellipse.

Il est maintenant possible d'évaluer la masse volumique de l'amas en utilisant les masses des clumps en gardant à l'esprit que ceux-ci continuent d'accréter du gaz. Cependant les clumps ne possèdent pas tous des HDEs, même dans les amas. Le catalogue produit d'Hi-GAL fournit tout de même les masses de ces clumps avec une distance fixée à 1 kpc. Il suffit ainsi de recalculer les masses en s'aidant de l'équation 2.6 et en l'inversant :

$$M = \frac{F_\nu D^2}{\Omega k_\nu B(T, \nu)} \quad (3.4)$$

Comme F_ν , Ω , k_ν et $B(T, \nu)$ sont considérés comme des constantes pour la même

source, la masse peut se recalculer de la façon suivante :

$$M_{new} = \left(\frac{D_{new}}{D} \right)^2 M_{old} \quad (3.5)$$

M_{new} correspond à la nouvelle masse après normalisation de la précédente valeur M_{old} par la distance D de la source et la nouvelle distance D_{new} , dont la valeur correspond à la distance héliocentrique de l'amas mentionnée plus haut. Une fois la masse totale de l'amas calculée, il suffit de la diviser par le volume de l'amas. Cependant nous n'avons accès qu'à la surface projetée de ces amas. En admettant que le rayon équivalent de l'amas corresponde au troisième axe de l'ellipsoïde nous pouvons calculer la masse volumique des amas. La figure 3.17 montre la distribution des valeurs de masse volumique pour l'ensemble des amas. La plus petite valeur est de $0.9 M_{\odot}.pc^{-3}$ et le pic se situe autour de $10 M_{\odot}.pc^{-3}$. En suivant les analyses de [Bok \(1934\)](#) et [Spitzer \(1958\)](#), [Lada & Lada \(2003\)](#) soutiennent que le terme d'amas est employé lorsque les sources qui le composent sont suffisamment proches et massives pour contrer la force de marée de la galaxie. Ceci induit une valeur typique de la masse volumique de $0.1 M_{\odot}.pc^{-3}$. Une valeur supérieure à $1 M_{\odot}.pc^{-3}$ est requise afin qu'un amas d'une taille typique de celle des amas de notre étude ne se désagrège pas rapidement, le temps de rupture étant d'environ 200 millions d'années (une rotation complète de la galaxie). En comparant ces valeurs types aux valeurs de masse volumique des amas de notre étude, on se rend compte que les clumps devraient rester grouper même après plusieurs rotations de notre galaxie et ainsi avoir le temps d'évoluer, au moins, jusqu'au stade de la séquence principale. Nous devrions ainsi être amené à observer encore des groupements d'étoiles à ces endroits dans un futur lointain. Cependant, d'autres facteurs participent à l'isolement des sources comme par exemple des perturbations gravitationnelles internes qui insufflent de l'énergie aux sources et les éloignent les unes des autres ([Spitzer, 1940](#)). Un modèle dynamique comprenant toutes ces perturbations ainsi que les caractéristiques des amas de notre étude permettrait de mieux étudier l'évolution de ceux-ci.

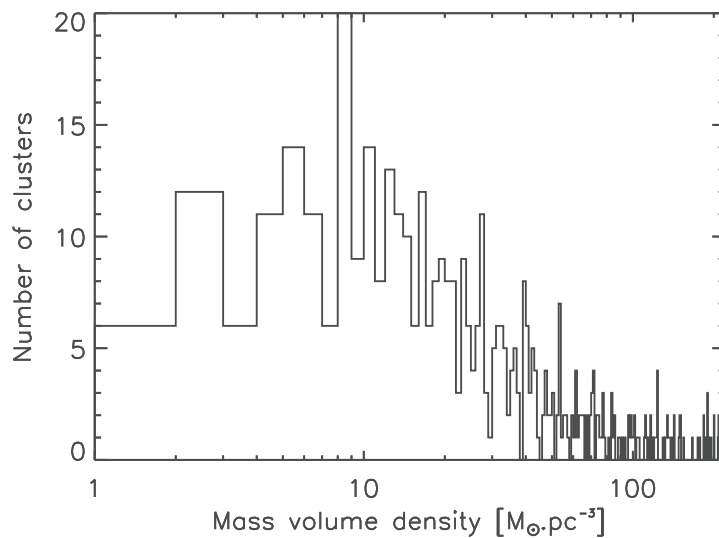


FIGURE 3.17: Distribution des masses volumiques des amas calculées à partir du total des masses des clumps de chaque amas.

3.4 Discussion

Nous avons évoqué, dans la section 3.1.2, plusieurs procédures pour déterminer Λ_{cut} . L'une d'entre elle consiste à évaluer Λ_{cut} comme étant la taille de branche qui permet d'extraire le plus de sur-densités (Battinelli, 1991). Nous avons appliqué cette procédure afin de comparer les valeurs Λ_{cut} avec celles de la méthode retenue pour la construction du catalogue d'amas. Pour cela nous avons calculé différentes solutions de la méthode MST à l'aide d'une grille de Λ_{cut} , entre 10 et 1450'' par pas de 1, calquée sur les valeurs de longueur de branches des différentes solutions du MST de notre étude (Exemple avec la figure 3.4). Nous nous sommes seulement intéressé aux solutions du MST pour une taille de fenêtre de 10 degrés, correspondant à la taille retenue dans notre étude. La partie haute de la figure 3.18 montre la distribution du nombre d'amas en fonction de Λ_{cut} pour l'ensemble des fenêtres de 10 degrés (27 au total en comptant les fenêtres chevauchantes). La partie basse de la figure montre la distribution des Λ_{cut} de notre étude (en rouge) suivant la longitude en comparaison avec celle des Λ_{cut} (en noir) de la méthode de Battinelli (1991). Les longitudes des points correspondent au centre des fenêtres. La plupart des différences de valeur entre Λ_{cut} des deux procédures ne dépassent pas 20''. Nous avons déjà vu dans la section 3.2.1 qu'une différence de 10-15'' ne changeait pas significativement les résultats, il en est de même ici. Les zones où l'on constate les plus grandes différences sont situées aux extrémités de l'échantillon, c'est-à-dire dans les zones de plus faibles densités. Cependant peu d'amas sont associés à ces zones ce qui signifie qu'une variation de Λ_{cut} n'entraîne pas une variation importante du nombre d'amas. Au final nous aurions aussi pu choisir cette procédure pour notre étude sans modifier significativement le résultat.

Le nombre de sur-densités fiables (496) comparé au nombre total (1633) est relativement faible. Ceci nous donne une idée de l'effet de projection sur la ligne de visée qui est d'autant plus important dans la partie interne de la galaxie. Afin d'augmenter la fiabilité et le nombre d'amas, il faudrait obtenir des HDEs plus précises ainsi qu'en plus grand nombre. Le catalogue *product* de clumps est une version provisoire et devrait à terme, comporter plus d'HDEs. La figure 3.19 montre la même région située entre $l=22$ degrés et $l=25$ degrés mais en ne superposant soit que les amas (partie haute), soit l'ensemble des sur-densités, c'est-à-dire les deux catalogues réunis (partie basse). Cette zone est relativement dense et illustre le problème d'identification de certains amas. Les zones les plus lumineuses souffrent d'une confusion importante et ne permettent pas de considérer les zones de formation stellaire comme fiables. Cette confusion est à la fois due à une difficulté de calcul des HDEs mais aussi à une difficulté d'extraire les clumps des images. En effet, le fond est aussi très lumineux dans cette zone et seul les clumps lumineux et donc très massifs peuvent être détectés. Un exemple de confusion est montré dans la figure 3.20. Bien que cette sur-densité semble réunir plusieurs régions lumineuses, il est difficile d'extraire chacune d'entre elles en raison de l'intervalle et du nombre important de valeurs d'HDEs, de 2.8 kpc à 12.3 kpc avec plusieurs pics entre 2 et 4 kpc, 4 et 8 kpc et au delà.

Cependant ces sur-densités rejetées peuvent être utiles pour étudier des régions particulières à l'aide d'autres données, comme par exemple les catalogues à plus courtes longueurs d'onde (WISE, MIPS GAL, GLIMPSE, etc...), ainsi qu'en recalculant les distances héliocentriques des sources.

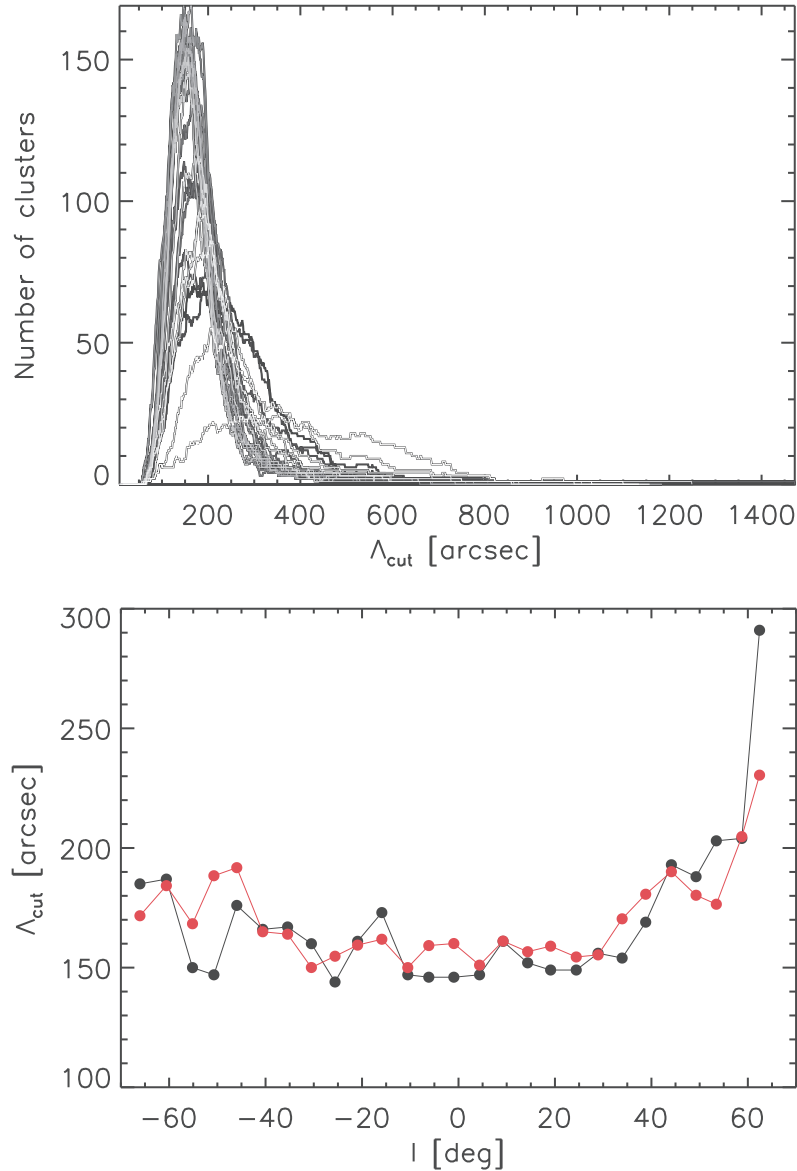


FIGURE 3.18: haut : Distribution du nombre d'amas en fonction de Λ_{cut} pour l'ensemble des fenêtres de 10 degrés. Bas : Distribution des Λ_{cut} de notre étude (en rouge) suivant la longitude en comparaison avec celle des Λ_{cut} (en noir) de la méthode de Battinelli (1991). Les longitudes des points correspondent au centre des fenêtres.

Enfin, plusieurs amas semblent appartenir à la même région de formation stellaire. La figure 3.21 montre une région H II, RCW 120 (Deharveng et al., 2009), associée à deux amas enfouis. Elle confirme l'importance d'évaluer l'ensemble des amas et des surdensités dans une certaine zone afin de l'étudier en profondeur.

3.5 Description des colonnes

Les deux catalogues seront disponibles prochainement dans la base de données VizieR¹ (Ochsenbein et al., 2000) du CDS. Cependant, une version du catalogue d'amas fiables est disponible dans l'annexe B. Les colonnes des deux catalogues sont décrites

1. <http://vizier.u-strasbg.fr/viz-bin/VizieR>

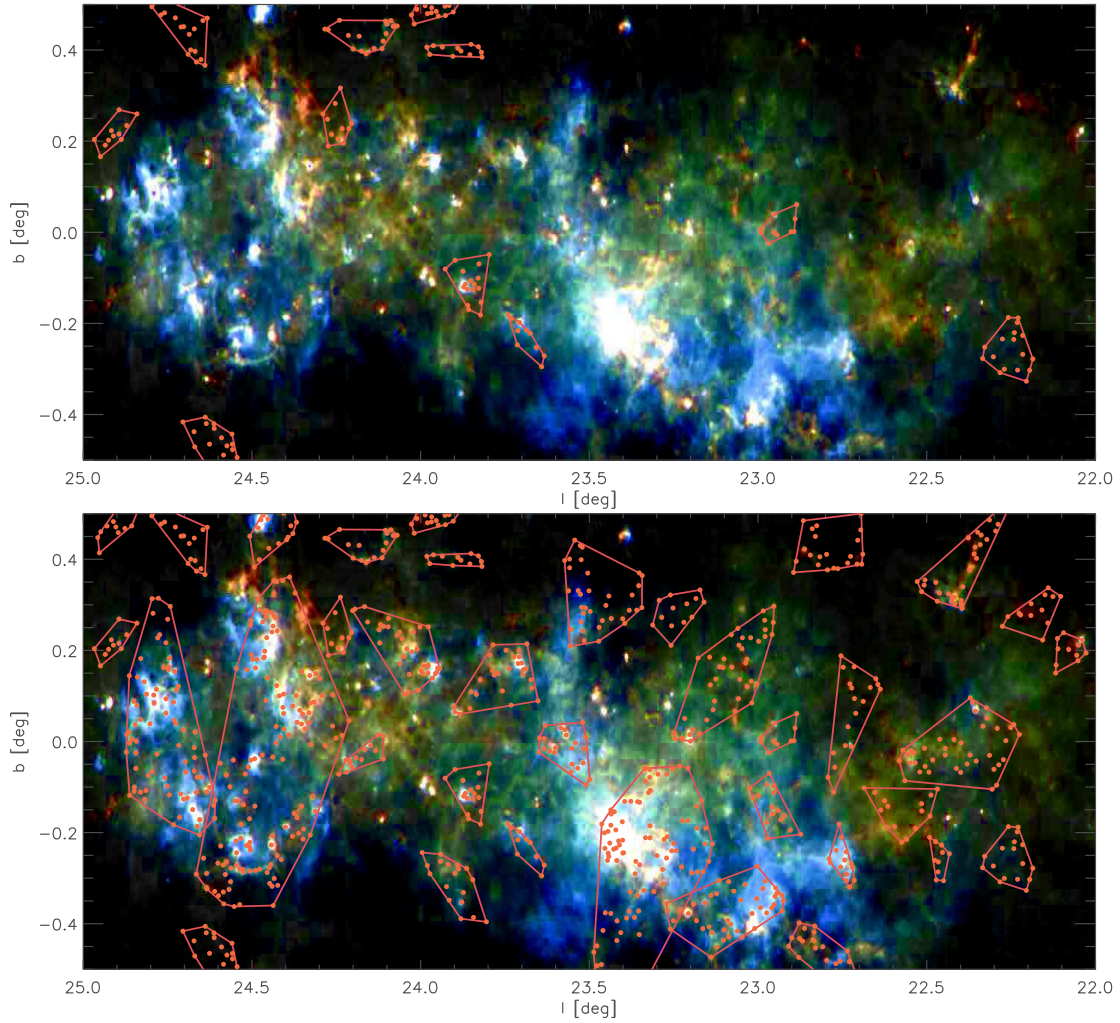


FIGURE 3.19: Distribution spatiale des sur-densités et de leur clumps dans une zone dense du plan galactique. Haut : Amas fiables. Bas : Ensemble des sur-densités trouvées pour cette zone. La figure de fond correspond à une composition colorée à 160, 250 et 350 μ m.

dans cette section.

3.5.1 Catalogue d'amas fiables

- Column [1] Nom de l'amas. La forme est $GLLL.ll+BB.bb$ où $LLL.ll$ et $BB.bb$ correspondent à la position moyenne en longitude et latitude de l'ensemble des clumps.
- Columns [2-3] correspondent aux coordonnées galactiques (l, b) de l'amas calculées comme décrit dans la colonne [1].
- Columns [4-5] correspondent à la taille angulaire de l'amas en arc-minute. La taille est obtenue en ajustant une ellipse sur chaque amas. La colonne [4] et la colonne [5] correspondent respectivement au petit axe et grand axe.
- Column [6] est l'angle de position (PA) de l'ellipse en degré. Il est mesuré dans le sens des aiguilles d'une montre à partir de la direction positive de l'axe de la longi-

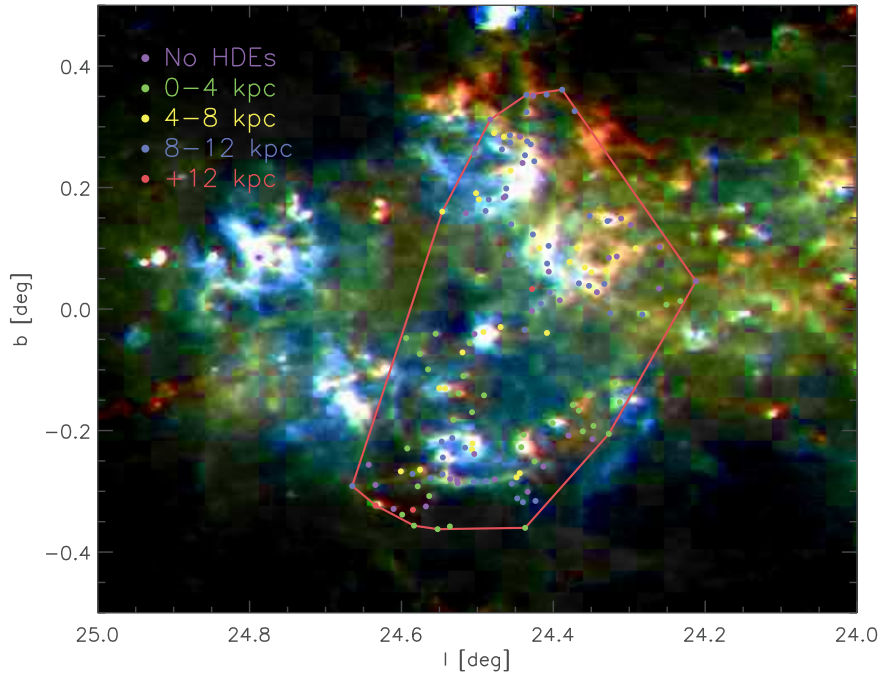


FIGURE 3.20: Exemple d'une sur-densité rejetée pour cause de confusion sur la ligne de visée. Les distances héliocentriques sont comprises entre 2.8 kpc et 12.3 kpc et il est difficile d'extraire une sur-densité fiable même dans les zones les plus lumineuses. La figure de fond correspond à une composition colorée à 160, 250 et 350 μ m.

tude.

- Column [7] correspond à la distance héliocentrique de l'amas, D_{peak} , estimée à partir du pic de l'histogramme des distances comme décrit dans la section 2.2.
- Column [8] correspond à la distance héliocentrique moyenne de l'amas, \bar{D} comme décrit dans la section 2.2.
- Column [9] est la distance galactocentrique de l'amas, R , calculée à partir de l'équation 4.3.
- Column [10] est la hauteur de l'amas, Z , par rapport à la latitude 0 du plan galactique et calculée à partir de l'équation 4.4.
- Column [11-12] correspondent au petit axe et grand axe de l'amas en parsec, calculés à partir de leur équivalent angulaire (colonne [4] et colonne [5]) et de D_{peak} .
- Column [13] est la densité de clumps dans l'amas, en pc^{-2} .
- Column [14] correspond au nombre total de clumps dans l'amas.
- Column [15] correspond au rapport entre le nombre de clumps pré-stellaires et proto-stellaires. Un rapport de 1 signifie que l'amas n'est composé que de clumps pré-stellaires alors qu'une valeur de 0 indique la présence de seulement les clumps proto-stellaires.

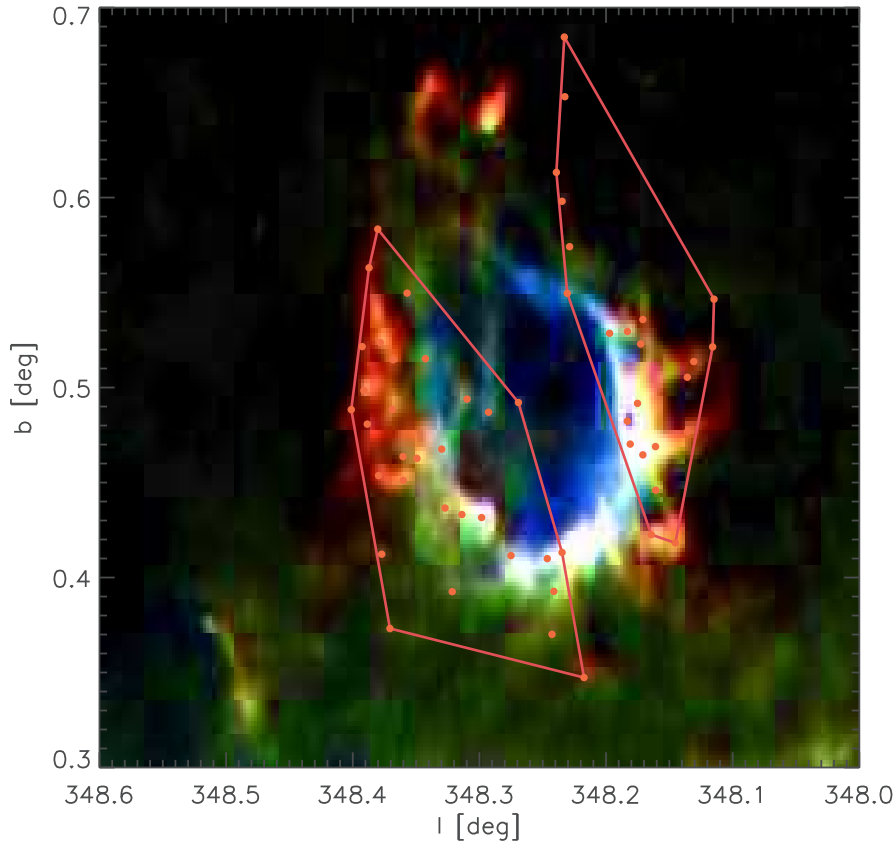


FIGURE 3.21: Exemple d'une région de formation stellaire caractérisée par la présence de deux amas du catalogue fiable mais qui ne comprend en réalité qu'un seul amas enfoui. La figure de fond correspond à une composition colorée à 160 , 250 et $350\mu\text{m}$.

- Column [16] correspond à la région H II la plus proche du centre de l'amas.

Les colonnes 9, 10 et 16 sont expliquées dans le chapitre suivant.

3.5.2 Catalogue archive

- Columns [1-6] sont les mêmes colonnes que celles de la section 3.1.
- Column [7] est la densité de clumps dans la sur-densité, en arcmin^{-2} .
- Columns [8-9] correspondent à la description des colonnes [14-15] dans la section 3.1.
- Column [10] est un *flag* qui détermine la classification d'une sur-densité rejetée. NDA, pour *No Distance Available*, classe les sur-densités sans HDEs. SD, pour *Spread Distance*, classe les sur-densités avec peu de clumps avec des HDEs et/ou avec aucun pic significatif dans la distribution des HDEs. Enfin, SCC, *Satisfying Compactness Criteria* qui sont des sur-densités satisfaisant les premiers critères de sélection d'amas mais ayant des clumps n'appartenant pas à l'amas. Ces clumps entraînent une confusion qui ne permettent pas de savoir si l'amas est réel.

3.6 Conclusion

MST est la méthode la mieux adaptée aux données de notre étude puisqu'elle permet d'extraire des sur-densités dans différents environnements allant de zones ténues à très denses. Elle permet notamment l'extraction de structures filamenteuses, difficile à obtenir avec les autres méthodes classiques. L'application de cette méthode ainsi que l'utilisation des estimations de distances héliocentriques, fournies par le catalogue *product* de clumps, nous a permis de construire un catalogue de sur-densités fiables, que l'on nomme amas enfouis de sources jeunes, ainsi qu'un catalogue archive comprenant les sur-densités les plus ambiguës mais utile à l'exploration de régions particulières de formation stellaire. Le catalogue fiable constitue le premier catalogue d'amas enfouis, dans la partie interne de la galaxie, composés essentiellement de coeurs pré-stellaires et de proto-étoiles de classes 0/I. Un autre catalogue d'amas enfouis à l'échelle galactique a été construit par [Camargo et al. \(2015\)](#) et mis à jour par [Camargo et al. \(2016\)](#). Cependant ce catalogue correspond à la partie externe de la galaxie, entre $l = 145$ degrés et $l = 290$ degrés. De plus les amas enfouis sont extraits à partir des données WISE dont les longueurs d'onde traduisent de la présence de sources plus évoluées (proto-étoiles de classes II/III) comme nous l'avons vu dans le chapitre 1.

Nous allons maintenant étudier les propriétés physiques et spatiales des amas et de leur clumps dans le chapitre suivant afin de comprendre leur rôle joué dans la formation stellaire.

CHAPITRE

4

ANALYSE DU CATALOGUE D'AMAS

Sommaire

4.1	Distribution spatiale des amas	72
4.1.1	Validité des amas	72
4.1.2	Distribution spatiale	74
4.1.2.1	Comparaison avec les bras spiraux	74
4.1.2.2	Hauteur des amas	76
4.2	Distribution spatiale des clumps	80
4.2.1	Concentration et fractale	81
4.2.2	Distribution spatiale	83
4.3	Distribution des masses des clumps	85
4.3.1	Ségrégation de masse	85
4.3.2	Fonction de masse des clumps	88
4.3.2.1	IMF et CMF	88
4.3.2.2	Application aux données	90
4.3.2.3	Multiplicité et efficacité stellaire	94
4.4	Association avec des objets de la littérature	95
4.4.1	Régions H II	96
4.4.2	IRDCs	97
4.4.3	Comparaison avec des catalogues d'amas enfouis pré-existants .	98
4.5	Discussions et biais observationnels	99
4.6	Conclusion	103

Après avoir construit les catalogues d'amas, je me suis intéressé à l'étude du catalogue d'amas fiables afin d'étudier statistiquement les régions de formation d'étoiles jeunes et leur lien avec leur milieu interstellaire environnant. Plusieurs analyses ont été faites, tant sur la distribution spatiale des amas que l'analyse des propriétés physiques des sources jeunes.

4.1 Distribution spatiale des amas

Le catalogue d'amas a été formé de la manière la plus fiable possible à l'aide de plusieurs critères. Cependant nous ne connaissons toujours pas l'organisation spatiale de ces amas. Nous pouvons nous attendre à observer des structures galactiques au sein de cette distribution, principalement les bras spiraux, ce que nous allons voir dans les sous-sections suivantes.

4.1.1 Validité des amas

Les amas sont caractérisés par une densité supérieure à celle du milieu environnant. Afin d'évaluer le contraste entre ces amas et le fond constitué de l'ensemble des sources sur la ligne de visée et ne faisant pas partie des amas, j'ai appliqué deux fonctions similaires provenant de [Campana et al. \(2008\)](#) et [Massaro et al. \(2009\)](#) :

$$g_k = \frac{\bar{\Lambda}}{\bar{\Lambda}_k} \quad (4.1)$$

g_k est appelé *clustering parameter* et correspond au rapport entre la longueur de branche moyenne $\bar{\Lambda}$ et la longueur de branche moyenne du k-ème amas $\bar{\Lambda}_k$. Les deux longueurs de branche sont calculées à partir des branches reliant chaque clumps provenant de la solution de la méthode MST. $\bar{\Lambda}$ correspond donc à la longueur de branche moyenne dans le cadre d'une fenêtre de 10 deg en longitude puisque cette taille de fenêtre fut sélectionnée parmi plusieurs autres tailles (plus de détails dans le chapitre 4).

$$M_k = N_{k,cl} g_k \quad (4.2)$$

M_k est appelé *magnitude* et correspond au produit du *clustering parameter* et du nombre de clumps $N_{k,cl}$ présents dans le k-ème amas. Ainsi nous pouvons comparer un amas avec un faible *clustering parameter* mais possédant un nombre important de clumps avec un amas ayant un meilleur contraste avec le fond mais avec peu de clumps.

La figure 4.1 montre les valeurs de g_k (histogramme du haut) et M_k (histogramme de droite) pour les 496 amas du catalogue ainsi que la relation entre les deux, c'est-à-dire $N_{k,cl}$, représentée par des lignes grises. Les pics des distributions de g_k et M_k sont localisés entre 1.4 et 1.5 ainsi que 14 et 16, respectivement. Ce sont les valeurs attendues puisque la distribution du nombre de clumps dans les amas est piquée autour de 10. Les valeurs sont similaires à celles trouvées par [Campana et al. \(2008\)](#) et [Massaro et al. \(2009\)](#). Dans ces deux articles, ils utilisent ces méthodes pour différencier les amas appartenant à une distribution aléatoire des amas fiables. Pour cela ils simulent une distribution aléatoire et compare les valeurs de M_k des amas sortant de cette distribution à leurs régions observées. L'écart entre les deux distributions est le seuil à partir duquel ils

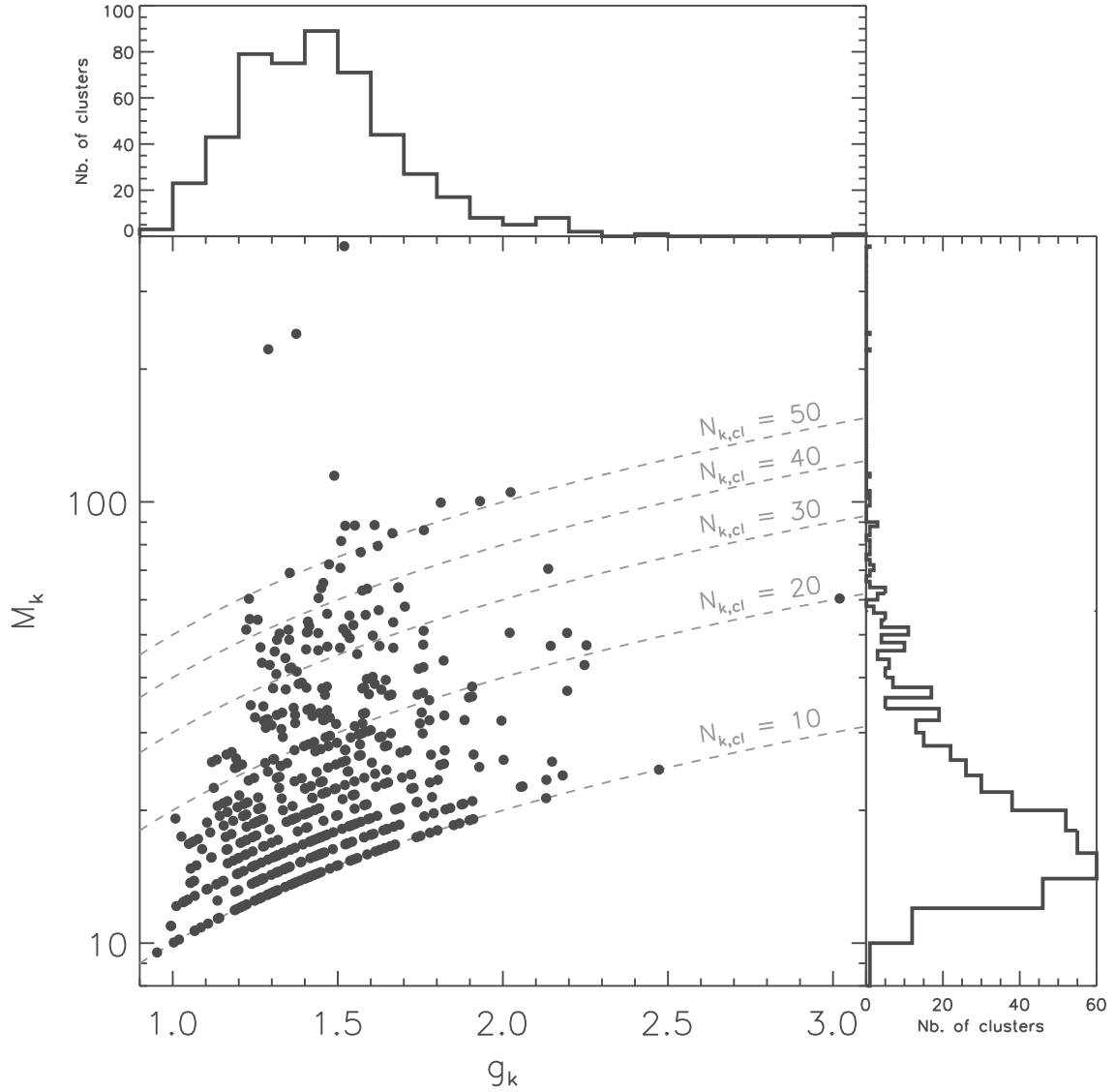


FIGURE 4.1: Distribution des amas dans un diagramme *clustering parameter-magnitude*.

considèrent les amas comme étant fiables, même si la distribution est toujours contaminée par quelques amas ambigus au dessus de cette limite.

Nous n'utilisons pas ces méthodes comme critères de sélection des amas fiables mais plutôt comme une analyse de la capacité de la méthode MST à évaluer des amas enfouis dans un fond relativement dense. Même si certains amas possèdent une petite valeur de g_k et M_k (< 1.5 ou < 15), ces amas n'en restent pas moins fiables en raison du nombre de clumps présents aux mêmes HDEs. Il est aussi utile de rappeler que certains amas sont formés essentiellement de filaments ce qui se traduit par une forte densité linéaire (le long du filament) sans pour autant observer cette même densité pour l'ensemble de l'amas. Nous pouvons ainsi confronter les amas possédant une faible *magnitude* avec les images Herschel comme c'est le cas avec la figure 4.2 où nous pouvons observer une structure filamentaire. Ainsi M_k ou g_k peuvent avoir une valeur faible ce qui n'est pas le cas d'un amas à symétrie sphérique, centralement concentré par exemple.

Il se trouve aussi que les études menées par [Campana et al. \(2008\)](#) et [Massaro et al. \(2009\)](#) concernent des sources gamma. Il est ainsi difficile de comparer ce domaine avec le notre puisque les sources ne s'organisent probablement pas de la même façon au sein

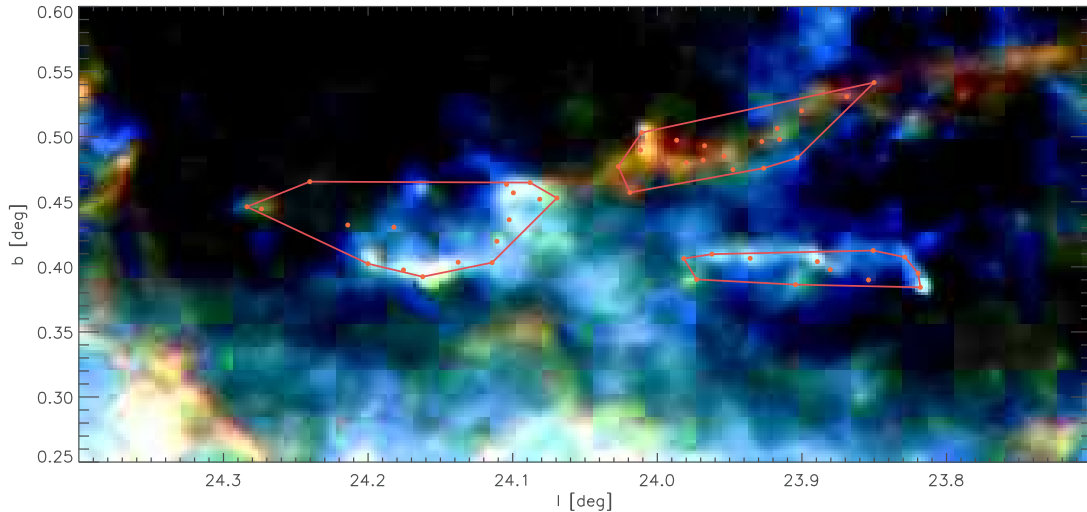


FIGURE 4.2: Exemple d'amas composés de filaments. L'image de fond est une image Hi-GAL à $250 \mu m$.

d'amas. De plus les auteurs ne semblent pas être en présence de structures filamenteuses ce qui s'ajoute à la différence de nos études respectives.

Enfin une analyse en 3 dimensions serait probablement plus appropriée dans notre cas, réduisant fortement l'impact du fond sur l'analyse. Pour cela, il faudrait obtenir des distances plus précises comme évoqué précédemment dans le chapitre 4, concernant l'élaboration d'un catalogue d'amas en 3 dimensions.

Nous avons montré que la plupart des amas de notre étude possèdent un contraste marqué avec les clumps du fond.

Nous allons maintenant nous intéresser à la distribution spatiale des amas.

4.1.2 Distribution spatiale

4.1.2.1 Comparaison avec les bras spiraux

Après avoir discuté de la véracité de ces amas, nous nous intéressons maintenant à la distribution spatiale de ceux-ci. Plusieurs études ont été menées sur les différentes structures de notre galaxie, notamment sur les bras spiraux, ce qui nous amène à comparer la distribution spatiale de nos amas de différentes façons. La première particularité est le nombre d'amas dans chaque quadrant. Le plan galactique est divisé en quatre quadrants à partir du centre de la galaxie. Les amas se répartissent entre le premier et le quatrième quadrant. 176 amas sont localisés dans le premier quadrant et donc les 320 autres amas sont localisés dans le quatrième quadrant dans le même intervalle de longitude. La distribution est donc plutôt asymétrique comme nous pouvons aussi le voir sur la figure 4.3.

Cette figure montre les cartes de densité aux différentes longitudes en comparant le premier quadrant avec son équivalent dans le quatrième. La différence la plus flagrante se situe entre $l=314$ et $l=330$ deg, comparé avec l'intervalle de $l=46$ à $l=30$ deg. Afin de comprendre cette différence de densité, nous pouvons observer la distribution des amas en la confrontant avec les bras spiraux. Nous ne connaissons pas exactement la position des bras spiraux, l'étendu de ceux-ci en longitude ainsi que leur nombre. Cependant nous pou-

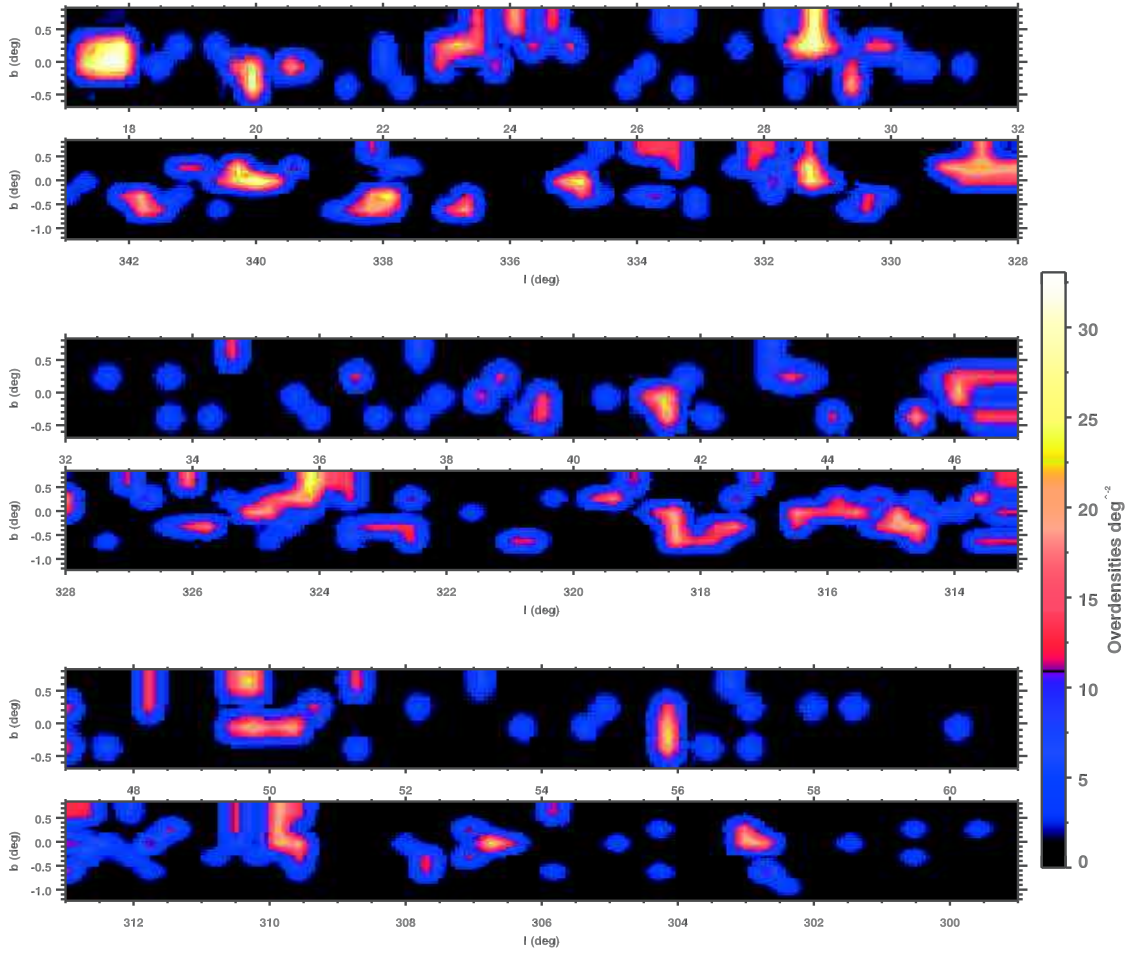


FIGURE 4.3: Carte de densité des amas. Les trois zones montrées permettent de comparer les premier et quatrième quadrants. La différence la plus significative se situe respectivement entre $l=46$ à $l=30$ deg et $l=314$ et $l=330$ deg.

vons nous baser sur les recherches précédentes afin d'observer le lien entre les amas de sources jeunes et ces bras spiraux. Englmaier et al. (2011) ont reconstruit la distribution 3D du gaz moléculaire (CO) à partir d'un modèle dynamique adapté à notre galaxie. Ils déduisent de cette distribution quatre bras spiraux comme montré dans la figure 4.4. On peut observer un décalage systématique entre les bras et les amas, ces derniers étant situés en avant des bras. Ce décalage disparaît lorsque l'on prend en compte d'autres formes de bras comme ceux de Georgelin & Georgelin (1976). Une mise à jour a été faite par Taylor & Cordes (1993), cette version étant montrée dans la figure 4.5. Ces bras ont été définis à partir de régions H II ce qui explique la différence avec les bras de Englmaier et al. (2011). En effet, les régions H II sont des objets plus localisés que le CO mais elles tracent aussi les premières étapes de la formation stellaire. En ce sens, il est normal que les amas composés de sources jeunes ajustent mieux ces bras spiraux plutôt que ceux traçant le gaz moléculaire. Dans ces deux figures nous observons quand même que les amas décrivent un motif particulier ressemblant à celui des bras spiraux. La plupart de ces amas semblent être localisés à proximité des bras nommés Perseus, Carina-Sagittarius et Scutum-Centaurus.

On s'aperçoit que l'asymétrie des bras provoque une différence de distribution spatiale des amas entre le premier et quatrième quadrant. Il est aussi probable qu'une différence d'efficacité de formation stellaire entre les différents bras, voir même à l'intérieur de cha-

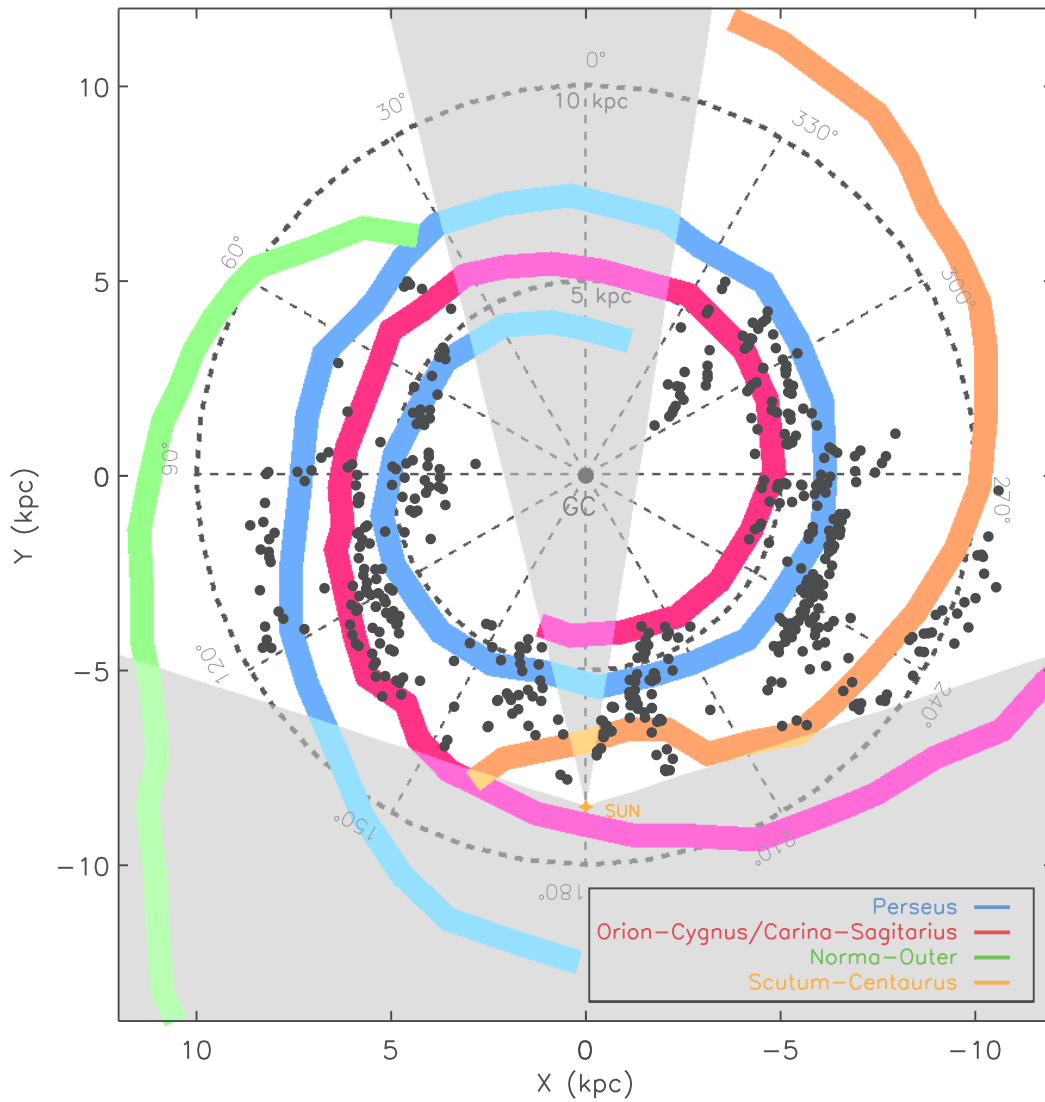


FIGURE 4.4: Distribution des amas dans la galaxie en comparaison des bras spiraux provenant de [Englmaier et al. \(2011\)](#).

cun des bras, peut être à l'origine de cette différence entre le premier et quatrième quadrants. Cependant [Moore et al. \(2012\)](#), [Eden et al. \(2013\)](#) et [Eden et al. \(2015\)](#) montrent qu'il n'y a pas de variations de la CFE, *Clump Formation Efficiency* dans les bras spiraux à l'échelle du kpc. Les différences observées sont plutôt à des échelles de l'ordre de la dizaine de parsecs ([Eden et al., 2012](#); [Vutisalchavakul et al., 2014](#)). Une autre explication de cette asymétrie pourrait venir de la densité de gaz et de poussières, différentes d'une région à l'autre. En effet, les amas semblent se regrouper dans des zones particulières qui pourraient être la signe d'une sur-densité de gaz et de poussières.

4.1.2.2 Hauteur des amas

Après avoir comparé les amas avec les bras spiraux, nous nous intéressons à la hauteur de ceux-ci par rapport au plan galactique. Afin d'observer une quelconque forme par-

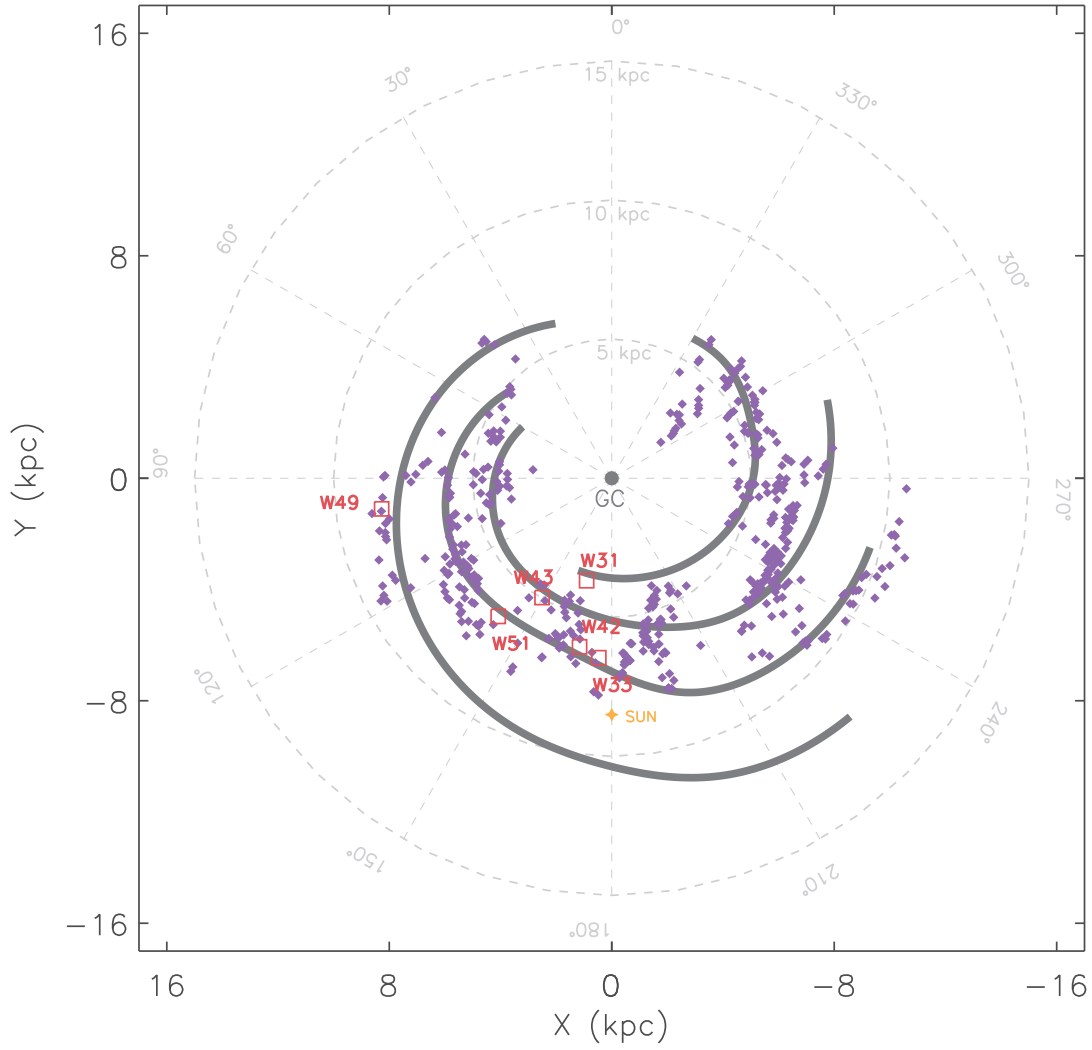


FIGURE 4.5: Distribution des amas dans la galaxie en comparaison des bras spiraux provenant de [Taylor & Cordes \(1993\)](#). Quelques régions de formation stellaire connues ont été ajoutées. Les amas semblent mieux ajuster les bras comparé à ceux de la figure 4.4.

ticulière dans le plan, il est préférable de partir du centre galactique. Il faut ainsi calculer la distance galactocentrique des amas, à l'aide de l'équation 4.3 :

$$R = \sqrt{R_0^2 + D_{pic}^2 - 2R_0D_{pic} \cos(l)} \quad (4.3)$$

Où R représente la distance galactocentrique des amas. R_0 correspond à la distance galactocentrique du Soleil et est fixée à 8.5 Kpc. Enfin D_{pic} et l représentent respectivement la distance héliocentrique des amas comme calculée dans le chapitre 4 et la longitude des amas.

La figure 4.6 montre la distribution de R avec l'ajustement d'une gaussienne afin d'étudier la position du pic. Celui-ci est situé aux alentours de 6 kpc. Peu d'amas sont localisés en dessous de 4 kpc car cette zone englobe la barre galactique et les premiers bras spiraux. Nous avons vu dans le chapitre 4 que l'estimation de distance est difficile dans cette zone.

Nous poursuivons notre étude en calculant la hauteur des amas par rapport au plan galactique à l'aide de l'équation 4.4 :

$$Z = D_{pic} \sin(b) \quad (4.4)$$

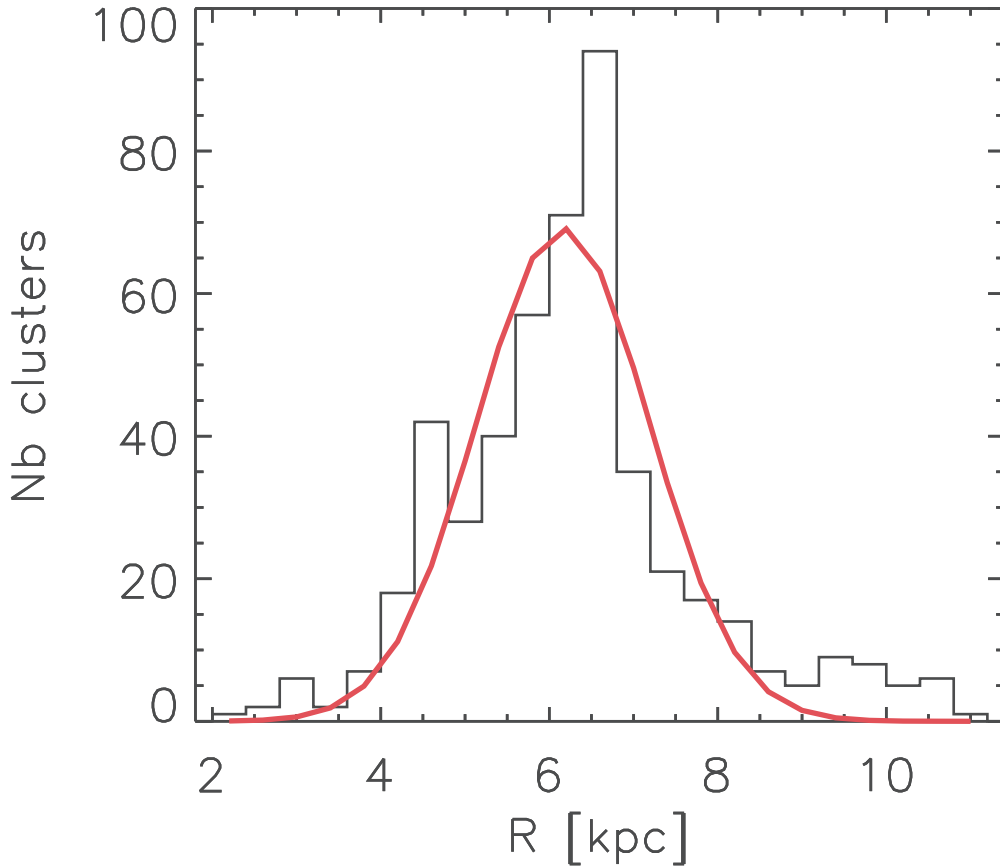


FIGURE 4.6: Histogramme de la distance galactocentrique des amas, R . Le pic se situe aux alentours de 6 kpc.

Où b correspond à la latitude des amas. La figure 4.7 représente l'histogramme de la hauteur des amas. Tous les amas semblent être compris dans un disque, que nous évoquerons plus bas, entre -250 et 135 pc. La valeur moyenne des hauteurs est $\langle Z \rangle \approx -6.2$ pc. Cette hauteur moyenne peut-être expliquée par la hauteur du Soleil par rapport au plan. En effet, [Brand & Blitz \(1993\)](#), [Reed \(1997\)](#) et [Reed \(2006\)](#) ont montré que le Soleil était localisé à $Z_0 = 10 - 20$ pc au dessus du plan galactique, ce qui est en accord avec notre résultat.

En séparant les hauteurs en fonction des quadrants, on remarque que seul les amas dans le quatrième quadrant sont localisés en dessous de 200 pc par rapport au plan.

Afin de connaître le profil de ce disque qui englobe les amas, nous allons comparer la hauteur avec la distance galactocentrique. La figure 4.8 nous montre cette distribution ainsi que la dispersion des valeurs σ_z . En fait, la distribution montre la valeur moyenne des hauteurs dans une taille de bin donnée, fixée ici à $0.1R_0$. Cette taille de bin permet d'éviter des fluctuations d'ordre statistique ainsi que de trop lisser la distribution auquel cas nous ne pourrions plus observer de profil particulier.

Il se trouve que dans le premier quadrant, la distribution fluctue avec une pente négative pour des distances inférieures à 6 kpc. au-delà de 6-7 kpc, la distribution monte avant de redescendre après 8-9 kpc. Dans le quatrième quadrant, au contraire, la distribution ne fait que suivre une pente négative après 6-7 kpc. Les amas situés à $Z \leq 200$ pc sont majoritairement situés dans cette zone.

au-delà de 6-7 kpc, nos résultats sont en accord avec ceux de [Paladini et al. \(2004\)](#) qui étudient le même profil pour les régions H II. Cependant nous ne pouvons pas comparer

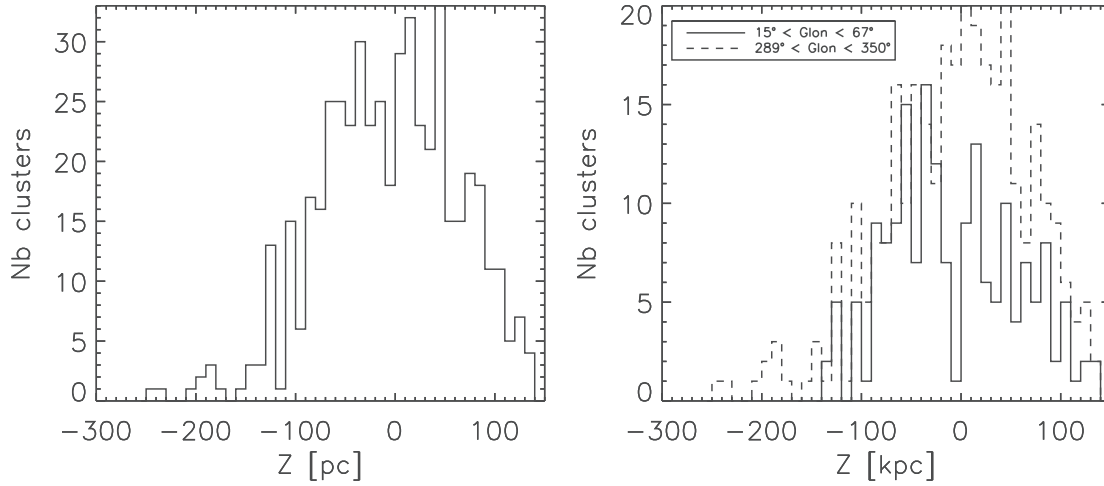


FIGURE 4.7: Histogramme de la hauteur des amas, pour tous les amas (figure de gauche) et en séparant les premier et quatrième quadrants (figure de droite).

l'ensemble de leur distribution puisque nous n'avons accès qu'à très peu d'amas au-delà de 10 kpc et seulement dans le quatrième quadrant. Nous devrions donc nous attendre à une pente positive dans le premier quadrant, au-delà de 10-11 kpc.

Ce profil similaire s'explique par le fait que les régions H II sont étroitement liées aux amas jeunes. Nous pouvons ainsi en déduire que cette distribution est liée au *warp* du plan galactique. Plusieurs études ont été menées sur ce *warp*, observé à l'origine dans le disque H I (Burton & Hartmann, 1988; Henderson et al., 1982). D'autres composantes semblent aussi faire partie de cette déformation du plan, comme les nuages moléculaires, les étoiles OB et 2MASS. On peut retrouver une liste exhaustive des composants principaux de cette déformation dans Reylé et al. (2008). Il faut cependant être prudent avec ce résultat car le *warp* est principalement visible au-delà de 10 kpc. Une valeur de 8.4 ± 0.2 kpc est calculé par Derriere & Robin (2001) comme distance de départ du *warp*, à partir de données DENIS et du modèle de Besançon. Ce que nous voyons n'est donc que le début de cette déformation. L'ambiguïté dans le premier quadrant peut aussi être interprétée par le fait que la déformation du disque dans le cas de la composante étoiles est plus fine c'est-à-dire que la pente observée est plus plate (Voir Marshall et al., 2006; Reylé et al., 2008). Le plan galactique est aussi évasé à ses extrémités. Paladini et al. (2004) l'observe dans leurs données, cependant nous ne pouvons pas car ce *flare* devient significatif à plus de 15 kpc du centre galactique. Nous pouvons d'ailleurs l'observer dans la distribution de la dispersion des hauteurs dans la figure 4.8, celle-ci étant quasiment plate dans les deux quadrants. Ceci est en accord avec Derriere & Robin (2001) qui montre que seule la partie externe du disque stellaire est évasée avec une variation en longitude de la distance de départ du *flare*, entre 8.8 et 10.5 kpc.

La distribution des hauteurs moyennes dans le premier quadrant peut aussi être vu comme une distorsion du plan galactique. Molinari et al. (2016b) démontre, à l'aide des données Herschel, que le plan est modulé en fonction de la longitude. Nous pouvons, à l'instar de cette étude, observer cette modulation à l'aide d'une fonction a_{mod} :

$$a_{mod} = \frac{C_{pos} - C_{neg}}{C_{pos} + C_{neg}} \quad (4.5)$$

Où C_{pos} et C_{neg} représentent respectivement le nombre d'amas au dessus et en dessous de $Z = 0$, dans une taille de bin fixée à $0.1R_0$. Ceci nous permet de créer une carte en 2D

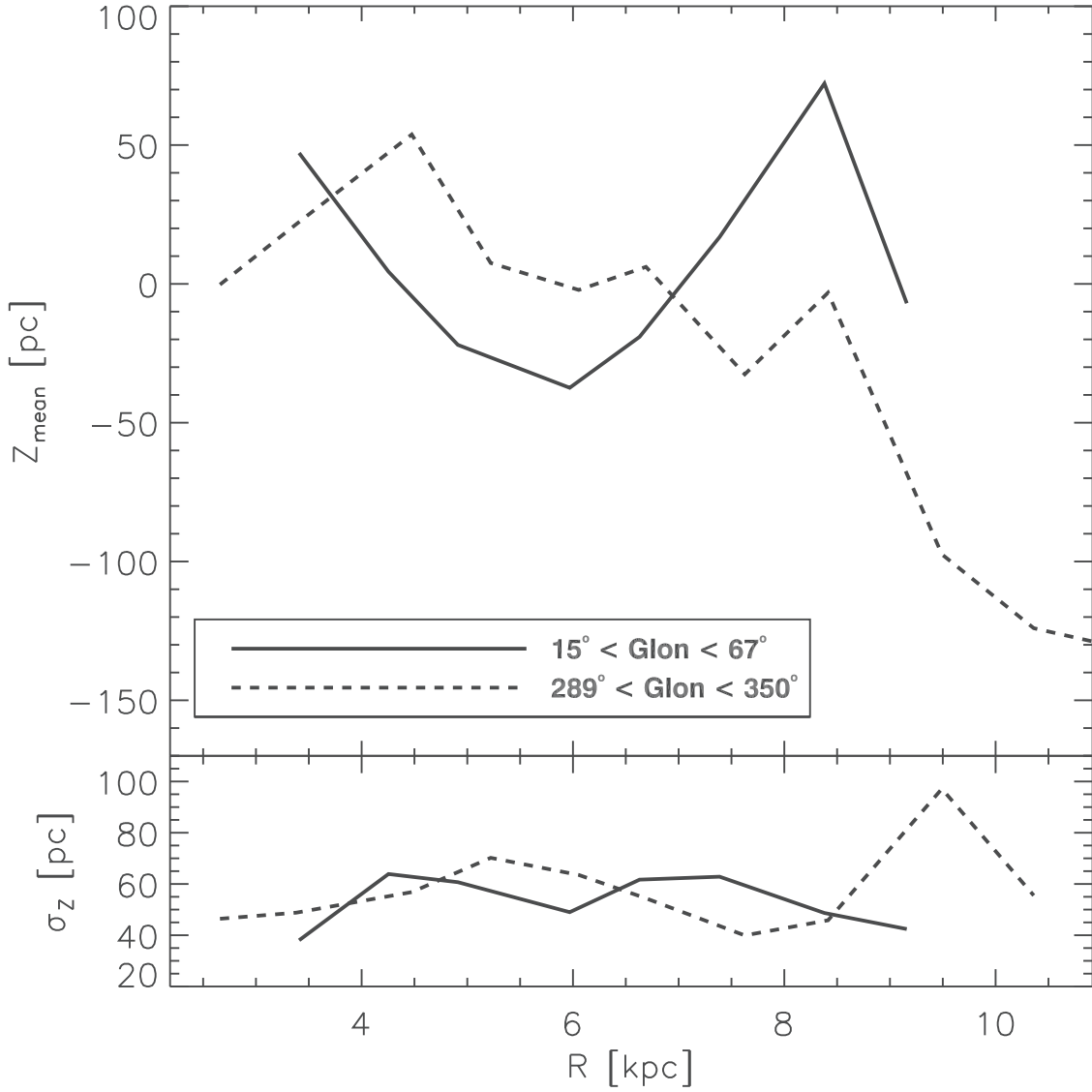


FIGURE 4.8: Distribution de la valeur moyenne de la hauteur des amas dans des bins de taille $0.1R_0$, en fonction de la distance galactocentrique. La partie inférieure de la figure représente la dispersion des valeurs σ_z .

de la fraction d'amas au dessus et en dessous du plan en fonction de différentes zones de notre galaxie. La figure 4.9 montre cette carte. Nous avons aussi ajouté les bras spiraux provenant de [Taylor & Cordes \(1993\)](#) en comparaison. Nous observons effectivement une distorsion, celle-ci ne semble pas être simplement liée à la longitude mais aussi aux bras spiraux, surtout dans le quatrième quadrant.

4.2 Distribution spatiale des clumps

Après avoir étudié la distribution spatiale des amas dans notre galaxie, nous pouvons nous intéresser à la distribution spatiale des clumps qui les composent. Nous allons voir que la méthode MST permet aussi de caractériser la distribution spatiale des clumps en amas. Le but de cette étude est de connaître le profil spatial des clumps et d'observer une tendance à cette échelle.

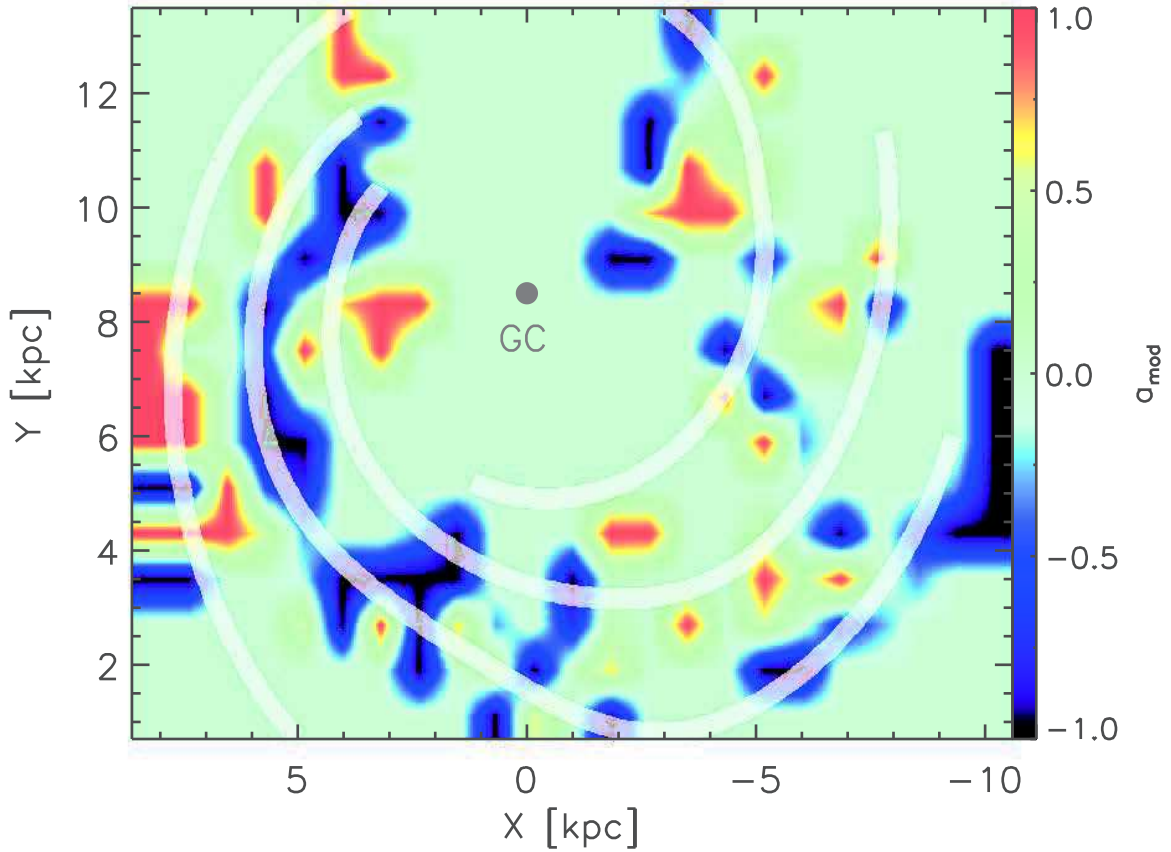


FIGURE 4.9: Carte en 2D de la fonction a_{mod} correspondante à l'équation 4.5. Les bras spiraux proviennent de [Taylor & Cordes \(1993\)](#).

4.2.1 Concentration et fractale

Lorsque nous évoquons le terme de distribution spatiale des amas, deux profils particuliers se distinguent en plus d'une distribution aléatoire. Le premier correspond à une concentration centrée que l'on peut renommer comme distribution centrée ou distribution centralement concentrée. C'est un profil simple à observer et à construire qui traduit un gradient de densité au sein de la distribution des sources. On peut observer ce profil dans des amas globulaires ou des amas ouverts. Un exemple résultant de simulation à N-corps ([Küpper et al., 2011](#)) est présenté dans la figure 4.10.

Le deuxième profil correspond au monde des fractales et est très utile pour comprendre le monde nous entourant. En effet, on retrouve ce profil un peu partout dans la nature et à différentes échelles, d'où le principe de fractale, mais qui ne sont pas forcément liées entre elles. Les meilleurs exemples dans la nature sont les flocons de neige, les arbres, l'ADN, le cerveau humain ou encore le brocoli. Le type de fractale le plus connu est probablement l'ensemble de Mandelbrot dont un exemple est donné dans la figure 4.11. Il existe des dizaines de méthode de construction des fractales mais en général elles correspondent à un polynôme d'ordre 2. Un autre exemple de représentation de fractale est donné dans la figure 4.11, représentant le flocon et la courbe de Koch (avec les étapes de construction).

Il est possible d'attribuer une dimension à toutes ces fractales, prénommée dimension

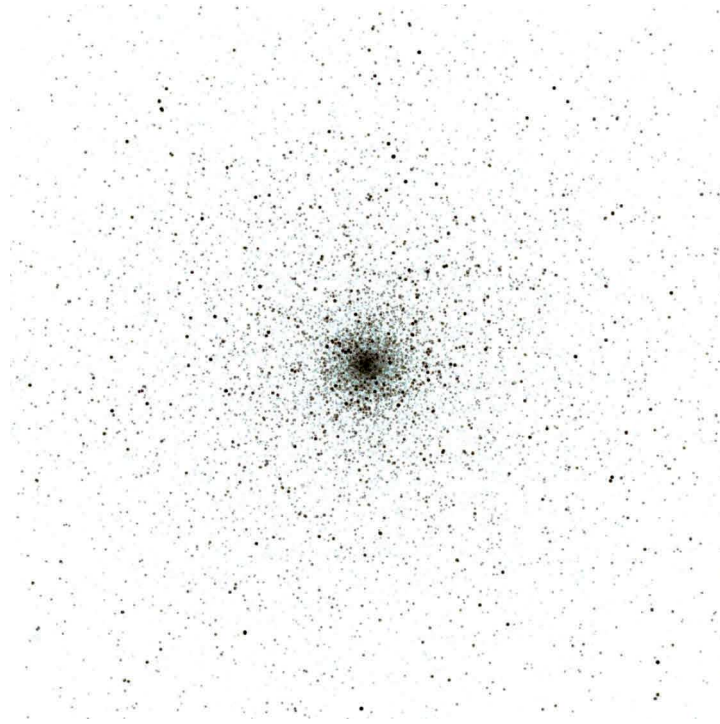


FIGURE 4.10: Exemple d'un amas possédant une distribution centrée, provenant de simulation N-corps (Küpper et al., 2011).

d'Hausdorff, afin de comparer le degré de sous-structures. Une figure *simple* en trois dimensions, du type d'une sphère, correspondra à une dimension fractale de 3 alors qu'une fractale complexe possédant un nombre accru de sous-structures se verra attribuer une dimension fractale plus petite (pouvant être plus petite que 1 pour certaines fractales).

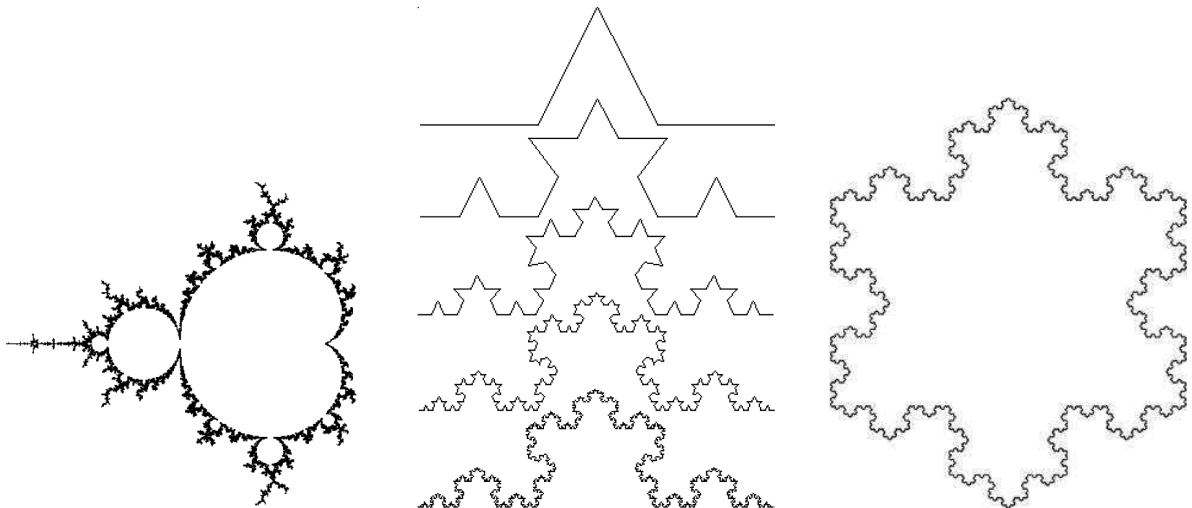


FIGURE 4.11: Plusieurs exemples de fractale. *Gauche* : Ensemble de Mandelbrot. *Milieu* : Construction du flocon de Koch. *Droite* : Représentation d'un flocon de Koch.

Nous allons ainsi analyser le profil de distribution des clumps dans la sous-section suivante afin de tenter d'observer l'un de ces deux profils particuliers.

4.2.2 Distribution spatiale

[Cartwright & Whitworth \(2004\)](#) ont développé une fonction permettant de séparer des profils différents à l'aide des distances entre clumps (équation 4.6). Trois profils sont considérés, à savoir une distribution centralement concentrée (avec différentes concentrations), aléatoire et fractale (avec différentes dimensions d'Hausdorff). Cette fonction est d'autant plus adaptée à notre étude puisqu'elle s'appuie en partie sur la méthode MST.

$$Q = \frac{\bar{m}}{\bar{s}} \quad (4.6)$$

\bar{m} correspond à la taille de branches moyenne normalisée de l'amas. Pour rappel, les branches sont les distances angulaires entre les sources calculées par l'algorithme de la méthode MST. La normalisation s'exprime $(N_{\text{clumps}} A)^{0.5} / (N_{\text{clumps}} - 1)$ avec N_{clumps} le nombre total de clumps dans l'amas et A l'aire de l'amas en degré. La deuxième partie de cette fonction Q , \bar{s} , correspond à la distance angulaire moyenne entre les clumps normalisée par le rayon angulaire de l'amas. [Cartwright & Whitworth \(2004\)](#) ont procédé à leur étude en utilisant principalement des amas simulés en deux et trois dimensions. Ils ont trouvé qu'un seuil $Q = 0.8$ permettait de distinguer les amas possédant une distribution centrée ($Q > 0.8$) de ceux ayant une dimension fractale ($Q < 0.8$). Ils ont aussi remarqué que plus le degré de concentration des sources au centre de l'amas était élevé et plus la valeur de Q augmentait. À l'inverse, plus la dimension fractale diminue, donc le nombre de sous-structure augmente, et plus Q diminuait. [Cartwright & Whitworth \(2004\)](#) ont aussi comparé leurs amas simulés avec des amas réels, à savoir IC2391, Chamaeleon, Taurus, ρ Ophiuchus et IC348. Les trois premiers posséderaient une distribution fractale tandis que les deux derniers auraient une distribution centrée. Les valeurs les plus extrêmes reviennent à l'amas du Taurus ($Q = 0.47$) et IC348 ($Q = 0.98$).

La figure 4.12 montre la distribution des valeurs de \bar{m} en fonction de \bar{s} . La droite caractérise la limite entre les différents profils, $Q = 0.8$. On remarque que pratiquement tous les amas, 489 sur les 496 du catalogue ($\approx 98,6\%$), se situent en dessous de ce seuil et donc correspondent à une distribution fractale. Les valeurs de Q se situent entre 0.4 et 0.84, en accord avec les valeurs trouvées pour les amas simulés et réels dans [Cartwright & Whitworth \(2004\)](#). Des travaux plus récents, présentés dans [Cartwright \(2009\)](#), montrent que la limite $Q = 0.8$ n'est pas une limite franche et ainsi que les amas situés entre $Q = 0.75$ et $Q = 0.85$ peuvent être considérés comme ayant l'une des trois distributions citées plus haut. Dans le même cadre, [Cartwright & Whitworth \(2009\)](#) suggèrent que la valeur de Q est sensible à l'élongation de l'amas, c'est-à-dire le rapport entre le demi-grand axe et le demi-petit axe. Plus particulièrement, ils notent dans leur étude qu'au-delà d'un rapport d'élongation de 3 l'impact sur Q devient significatif et nécessite une correction à appliquer en fonction de la corrélation entre ces deux valeurs. Dans notre étude, seulement 40 amas ($\approx 8\%$) ont une élongation supérieure à 3 et aucune corrélation entre cette valeur et Q ne peut être trouvée, comme celle observée dans [Cartwright & Whitworth \(2009\)](#). Nous avons alors décidé de ne pas considérer cette correction. L'explication à cette absence de corrélation dans notre étude peut s'expliquer par la présence de filaments. En effet, la principale différence entre notre étude et celle de [Cartwright & Whitworth \(2004\)](#) est qu'ils considèrent des amas d'étoiles plus âgés que ceux de notre étude, qui ont eu le temps d'effacer leurs empreintes initiales dans le nuage moléculaire les englobants. Dans notre étude, les clumps sont enfouis dans le gaz et la poussière de leur géniteur et sont, en plus, associés à des filaments. Cela signifie aussi qu'une distribution considérée comme aléatoire dans notre étude ($Q = 0.8$) peut aussi correspondre à une structure filamentaire.

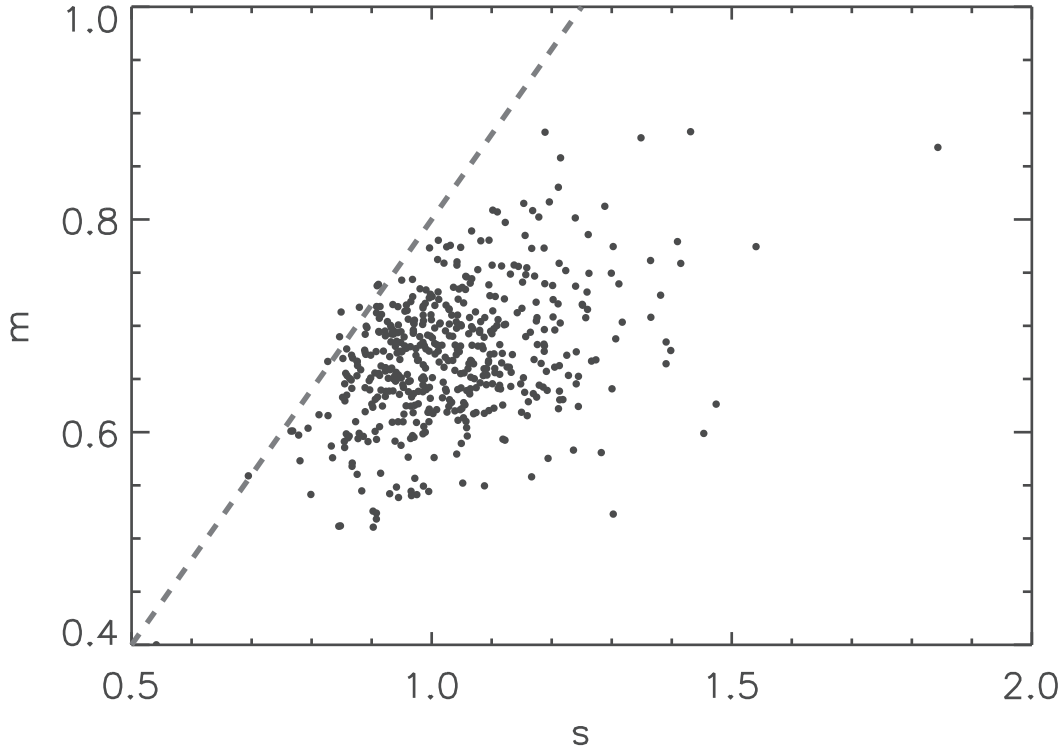


FIGURE 4.12: Relation entre \bar{m} et \bar{s} de la fonction Q (équation 4.6), caractérisant le profil de distribution des clumps d'un amas. La droite correspond à la valeur $Q = 0.8$ séparant une distribution centrée d'une distribution fractale.

En considérant la zone ambiguë de distribution spatiale des clumps mentionnée plus haut, $0.75 \leq Q \leq 0.85$, 90% des amas possèdent une distribution fractale avec $Q \leq 0.75$. Il s'en suit que la plupart des amas possèdent une distribution fractale, avec plusieurs dimensions fractales différentes allant de l'observation de sous-structures marquées (comme pour l'amas du Taurus) à des observations plus ambiguës (comme pour l'amas du Chamaeleon). En accord avec [Cartwright & Whitworth \(2004\)](#), les dimensions d'Hausdorff vont de 1.5 à 3 avec un pic aux alentours de 2.

[Parker & Dale \(2015\)](#) ont étudié la distribution spatiale des amas d'étoiles jeunes ($\leq 3 \text{ Myr}$) à travers des simulations hydrodynamiques, à partir de l'évolution de GMCs. Ils ont montré que ces amas suivent aussi une distribution fractale avec des valeurs similaires de Q comparées à notre étude. Cependant, [Cartwright & Whitworth \(2004\)](#) montrent que certains amas réels ont une distribution centrée. [Schmeja & Klessen \(2006\)](#) ont analysé les mêmes amas en différenciant les sources selon leurs stades évolutifs, de classe 0 à classe 3. Outre le fait que la distance entre les étoiles devient de plus en plus grande au fur et à mesure du temps, ils trouvent que le profil de distribution spatiale est lié à l'évolution temporelle des objets. Ainsi les amas forment tout d'abord des sous-structures aux mêmes stades évolutifs que les clumps de notre étude, puis ces sous-structures disparaissent pour former des amas avec une concentration centrée. On peut ainsi s'attendre à ce que les amas de notre étude finissent par évoluer dans le même sens.

Dans notre étude et dans le cadre de sous-structures marquées, il est difficile d'étudier le profil spatial de chacune des sous-structures en raison du peu de clumps présents, la majorité des amas possédant moins de 20 clumps. Nous ne pouvons, par exemple, pas connaître le nombre de magnitudes d'échelle pour lequel nous retrouvons les mêmes structures.

Combes (2000) montrent que le MIS possède aussi une dimension fractale, sur 6 magnitudes d'échelle, avec une dimension d'Hausdorff comprise entre 1.6 et 2 c'est-à-dire équivalente à nos résultats. Cet article montre aussi que les regroupements de galaxies possèdent la même forme sur 4 magnitudes d'échelle avec le même intervalle de valeurs des dimensions d'Hausdorff.

Le résultat de notre étude poursuit directement les résultats de cet article et montre à quel point la notion de fractale est importante dans la nature, sur des échelles et domaines différents. Cela signifie aussi que les mêmes processus physiques sont à l'oeuvre entre la formation d'amas de galaxies et la formation d'amas d'étoiles au sein de ces mêmes galaxies.

4.3 Distribution des masses des clumps

Après avoir étudié la distribution spatiale des clumps appartenant aux amas, nous pouvons continuer notre analyse du catalogue d'amas en nous concentrant sur les propriétés physiques de clumps. Le catalogue *product* nous renseigne sur certaines propriétés physiques des clumps et nous allons particulièrement nous intéresser aux masses dans cette section.

4.3.1 Ségrégation de masse

Nous avons vu précédemment que la distribution des clumps au sein des amas était de nature fractale. Ceci sous-entend que ces amas possèdent des sous-structures et donc que les objets les plus massifs soient dispersés dans tout l'amas. Cependant, du fait de la nature particulière des amas composés en grande partie de filaments, nous pouvons aussi imaginer que les objets les plus massifs se réunissent à l'embranchement de certains filaments ou simplement dans une zone plus dense. La figure 4.13 montre l'exemple d'une distribution spatiale des clumps d'un amas avec les valeurs des masses de chacun. Nous allons donc étudier la distribution spatiale des masses à l'aide de deux méthodes différentes. Il est à noter que les clumps ne possèdent pas tous des estimations de masse et qu'il est nécessaire de les calculer de la même façon que dans la section 3.3.

La première méthode, développée par Allison et al. (2009), utilise la méthode MST afin d'étudier la ségrégation de masse au sein d'amas simulés. Il s'agit en fait de recalculer une solution de MST seulement pour un certain nombre de sources de l'amas, N_{MST} . Dans notre cas N_{MST} est compris entre 5 et 10 ce qui est petit comparé à l'étude d'amas de plusieurs milliers de sources. Cette méthode se caractérise par la définition d'un rapport appelé *Mass Segregation Ratio* et est symbolisé par Λ_{MSR} . L'équation 4.7 correspond à la définition analytique de ce rapport.

$$\Lambda_{MSR} = \frac{\langle \Lambda_{rand} \rangle}{\Lambda_{massive}} \pm \frac{\sigma_{rand}}{\Lambda_{massive}} \quad (4.7)$$

$\langle \Lambda_{rand} \rangle$ correspond à la moyenne des longueurs totales des branches de MST pour un sous-échantillon de N_{MST} clumps choisis aléatoirement. C'est une valeur moyenne car nous utilisons 500 sous-échantillons de N_{MST} clumps afin d'obtenir une bonne estimation de l'erreur, caractérisée ici par la déviation standard σ_{rand} . Les valeurs de $\langle \Lambda_{rand} \rangle$ sont

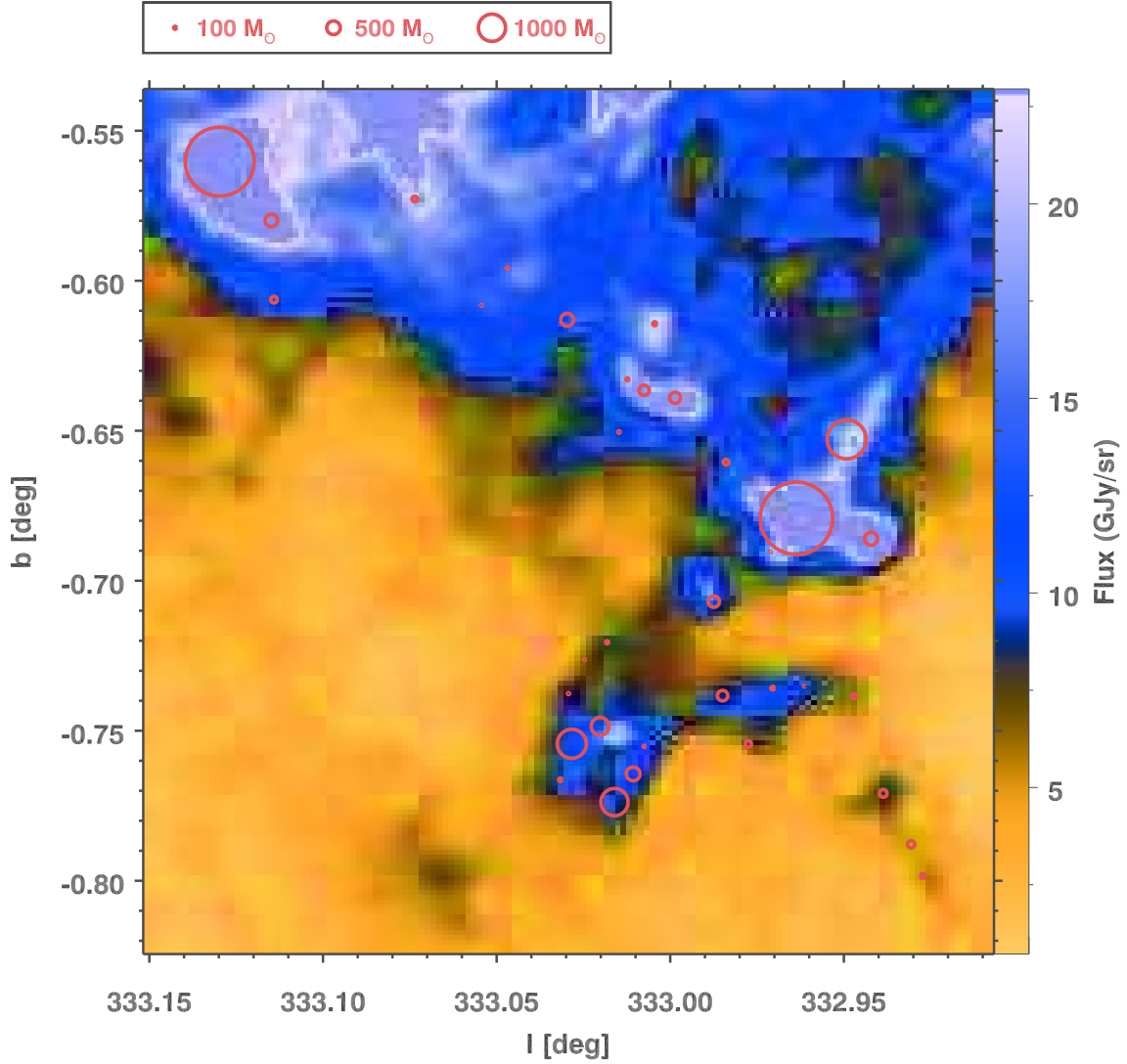


FIGURE 4.13: Exemple de distribution spatiale des clumps d'un amas avec l'ajout des valeurs des masses. Le fond correspond à l'image de l'amas vu à $70 \mu m$.

comprises entre 0.088 et 1.37 degrés. $\Lambda_{massive}$ est la longueur totale des branches de MST pour les N_{MST} clumps les plus massifs de l'amas. $\Lambda_{MSR} = 1$ indique une absence de ségrégation de masse car les sources les plus massives sont distribuées de la même façon que les sources choisies aléatoirement. $\Lambda_{MSR} < 1$ indique une ségrégation inverse alors que $\Lambda_{MSR} > 1$ est le cas que nous recherchons, c'est-à-dire avec une densité plus importante des objets massifs comparée à celle des sources choisies aléatoirement.

Cette méthode permet d'obtenir une valeur limite de N_{MST} en dessous de laquelle les clumps massifs sont ségrégués. Dans notre étude, Λ_{MSR} est compris entre 0.65 et 4.5 pour tous les amas et toutes les valeurs de N_{MST} , avec une erreur maximale $\sigma_{rand} = 0.2$. Il se trouve aussi que 0.65 correspond à la valeur la plus grande de certains amas. 32.6% des amas ont $\Lambda_{MSR} < 1$ mais seulement 6% ont une valeur de Λ_{MSR} inférieure à 0.8. Au total, 56.6% des amas n'ont pas de ségrégation de masse apparente et 37.4% ont une faible ségrégation de masse avec une valeur maximale de Λ_{MSR} ne dépassant que rarement 2. En comparaison avec notre étude, [Allison et al. \(2009\)](#) trouve des valeurs de Λ_{MSR} supérieures à 8 en appliquant leur méthode à la nébuleuse d'Orion.

La deuxième méthode a été développée par [Kirk et al. \(2016\)](#) et consiste à comparer

la distance qui sépare le centre de l'amas du clump le plus massif et la distance moyenne séparant les clumps du centre de l'amas. Comme pour [Allison et al. \(2009\)](#), cette méthode définit un rapport de ces deux distances que l'on nomme *rapport d'offset*. La figure 4.15 montre la distribution de la plus grande masse et la deuxième plus grande masse de chaque amas, normalisée à la masse de l'amas, en fonction de ce rapport d'offset. La figure 4.14 provenant de [Kirk et al. \(2016\)](#) (figure 5 de l'article) montre un exemple de ségrégation de masse dans les sous-structures d'Orion A et B (aussi dans Lane et al., en préparation) en comparaison avec les YSOs en amas provenant d'une étude précédente ([Kirk & Myers, 2011](#)). Plus l'objet le plus massif se rapproche du centre de la sous-structure et plus la masse est importante en comparaison avec la masse totale de cette sous-structure. Nous n'observons pas cette distribution particulière dans notre étude que ce soit pour le clumps le plus massif ou le deuxième le plus massif. Même en séparant les amas à l'aide de leurs distances héliocentriques, il n'est pas possible d'obtenir le résultat de [Kirk et al. \(2016\)](#).

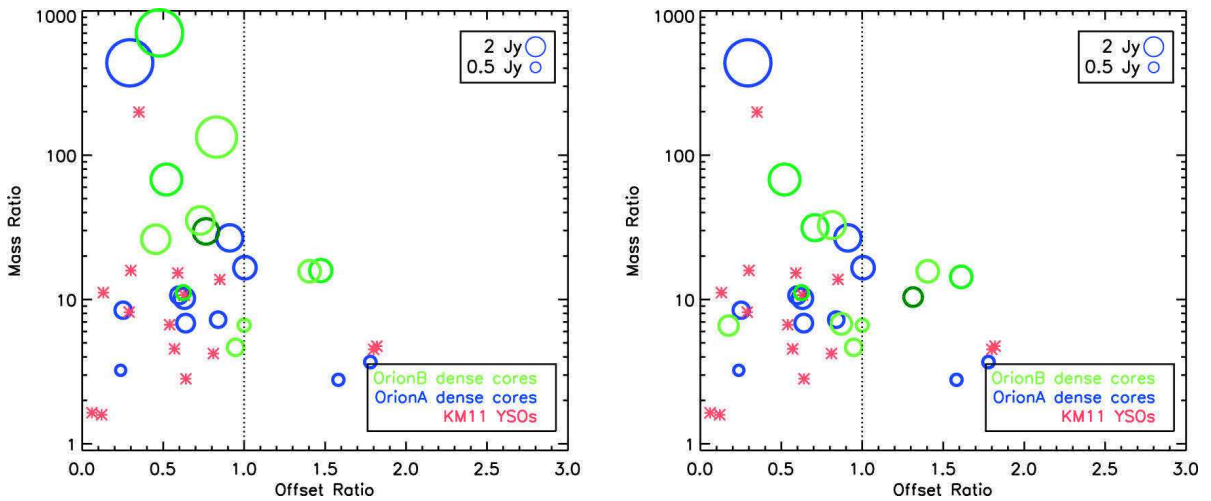


FIGURE 4.14: Figure de [Kirk et al. \(2016\)](#) montrant la ségrégation de masse dans Orion A, Orion B et des YSOs en amas d'une étude précédente ([Kirk & Myers, 2011](#)).

Une troisième méthode nommée $m-\Sigma$, consiste à comparer la masse totale comprise dans une certaine aire ou volume avec la densité de sources locale ([Maschberger & Clarke, 2011](#)). Nous avons fait le choix de ne pas utiliser cette méthode dans notre étude car le nombre de clumps par amas est trop peu élevé pour calculer une densité de sources locale fiable. Les résultats issus de cette méthode seraient ainsi bien plus discutables que ceux des deux méthodes présentées précédemment.

Au final, les deux méthodes appliquées ne suggèrent pas une ségrégation de masse ce qui rejoint notre analyse du profil spatial des clumps. Nous n'observons pas d'amas formant des clumps massifs à un seul endroit mais plutôt dans l'ensemble des sous-structures de ceux-ci. Ceci signifie que la ségrégation s'effectue probablement à une échelle plus fine. En effet, les différentes études menées jusqu'à présent traitent d'objets individuels dans le voisinage solaire, dans des amas de quelques parsecs de diamètre. Ces régions correspondent ainsi à la zone occupée par un voir deux clumps de notre étude.

Malgré une limitation de cette étude de ségrégation de masse due à la résolution, il

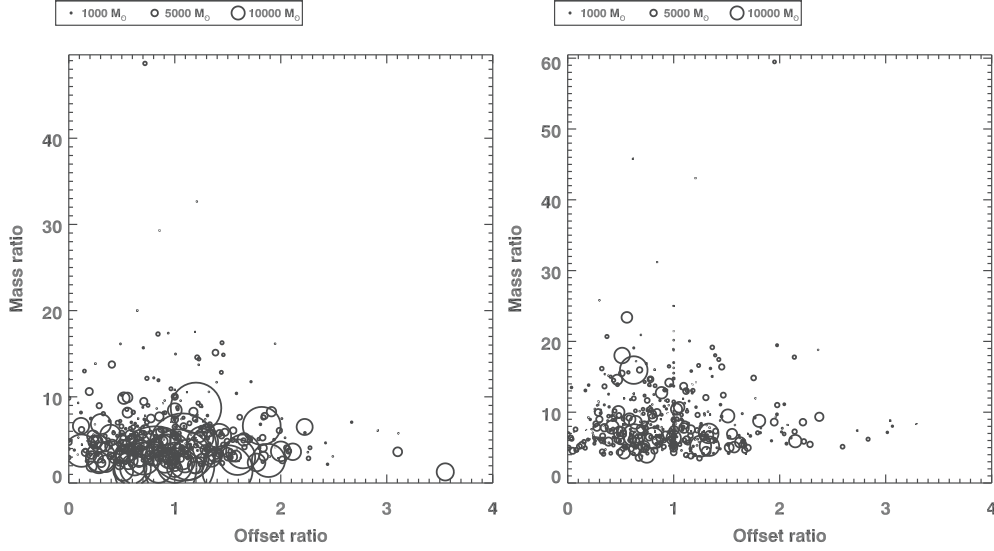


FIGURE 4.15: Méthode de [Kirk et al. \(2016\)](#) pour étudier la ségrégation de masse au sein des amas. Le graphique de gauche montre le lien entre la plus grande masse et le rapport d'offset. Le graphique de droite représente la même chose mais pour la deuxième plus grande masse.

est possible d'analyser la fonction de masse des clumps, dans la section suivante.

4.3.2 Fonction de masse des clumps

4.3.2.1 IMF et CMF

La distribution de masses est largement utilisée dans divers domaines de l'astrophysique, allant des ensembles de galaxies aux groupements d'étoiles. Pour ce dernier cas, il existe une fonction particulière que l'on nomme fonction de masse initiale où *Initial Mass Function* (IMF). Elle caractérise la distribution de masse d'objets stellaires, typiquement des étoiles en formation, dans un certain volume d'espace (généralement en pc^3). La première estimation de l'IMF provient de [Salpeter \(1955\)](#) pour des étoiles situées dans le voisinage solaire. L'expression originelle de cette fonction est la suivante :

$$\Phi_{Salpeter}(M) = \frac{dN}{d(\log_{10}m)} \frac{dt}{T_0} \quad (4.8)$$

Où dN est le nombre d'étoiles dans l'intervalle de masse dm créé pendant l'intervalle de temps dt . T_0 est une normalisation correspondant à l'âge de la galaxie, fixant ainsi l'âge maximum des étoiles dans cette galaxie. [Salpeter \(1955\)](#) étudie principalement la fonction de luminosité qui est liée à l'IMF par l'équation suivante :

$$\kappa(M_v) = -\Phi(M) \frac{d(\log_{10}m)}{dM_v} \quad (4.9)$$

Où M_v représente la magnitude visuelle absolue. Cela suppose de connaître la relation masse-luminosité (MLR) des étoiles et objets jeunes. Cependant de nombreuses incertitudes, tant sur les observations de la MLR que sur les théories d'évolution et de structure des étoiles, entraîne des incertitudes sur la relation $\kappa(M_v) = \Phi(m)$. La MLR dépend notamment de l'âge des sources, de leur métallicité, de leur rotation ainsi que leur transfert

de matière avec leur environnement proche (Massey et al., 2002).

L'IMF fut très étudiée à partir de l'étude Salpeter (1955) et demeure à l'heure actuelle une fonction incontournable de l'étude de la formation stellaire. Salpeter (1955) montra que l'IMF était proportionnelle à $m^{-\Gamma}$ avec $\Gamma = 1.35$ dans la partie des étoiles massives. Plusieurs études suivantes ont montré que l'ensemble de l'IMF peut être approximée par trois ou quatre pentes (Kroupa, 2001, 2002; Kroupa et al., 1993) avec un changement de pente aux très grandes masses visibles dans certaines données.

Plusieurs équipes de chercheurs se sont penchées sur la signification de cette fonction de masse ainsi que son étude dans plusieurs environnements différents. L'IMF est intéressante au sens où elle permet d'établir une connexion entre les étoiles de faibles masses, intermédiaires et les étoiles massives, allant de $\approx 0.01 M_{\odot}$ (étoiles naines) à $\approx 120 M_{\odot}$ (Massey, 1998). Plusieurs processus physiques sont à l'oeuvre pour former ces différents types d'étoiles et des simulations ont ainsi vu le jour afin de faire correspondre les théories et les observations. On apprend notamment que les simulations montrent que l'IMF devrait varier selon les conditions, les environnements et les processus physiques de la formation stellaire (Kroupa, 2001). Cependant les observations montrent une invariabilité dans cette distribution de masse, contraire aux théories en vigueur et aux simulations (Kroupa, 2002).

Nous allons ainsi étudier cette distribution de masse dans notre étude. Les clumps de notre étude sont jeunes (1-2 millions d'années au maximum) comparé aux étoiles étudiées dans la plupart des papiers traitant de l'IMF (plusieurs dizaines de millions d'années au minimum). De plus les clumps ne sont pas des objets stellaires individuels mais un ensemble de coeurs pré-stellaires et/ou de proto-étoiles. Ainsi, nous ne pouvons pas parler d'IMF dans notre cas mais de CMF pour *Clump Mass Function*, à ne pas confondre avec *Core Mass Function*, que nous nommons CoMF ici, qui correspond à la distribution de masse des coeurs pré-stellaires et proto-étoiles même si les deux distributions sont liées.

Comme les clumps possèdent approximativement tous le même ordre de grandeur d'âge, la CMF s'écrit plus simplement en s'affranchissant de la partie temporelle dt/T_0 . La fonction établie par Pavel Kroupa (Kroupa et al., 1993) qui correspond à l'ajustement de la CMF par trois segments, à trois intervalles de masses différents, s'écrit comme suit :

$$\Phi_{Kroupa} = \frac{dn}{dm} \propto m^{-\alpha} \text{ with } \alpha = \begin{cases} \alpha_1, & m_0 < m < m_1 \\ \alpha_2, & m_1 < m < m_2 \\ \alpha_3, & m_2 < m \end{cases} \quad (4.10)$$

Où m_0 , m_1 et m_2 sont les trois masses à laquelle la distribution change de valeur de pente. Dans le cas de l'IMF, $m_0 = 0.01 M_{\odot}$, $m_1 = 0.1 M_{\odot}$ et $m_2 = 0.5 M_{\odot}$. De même les pentes dans ces intervalles valent : $\alpha_1 = 0.3 \pm 0.7$, $\alpha_2 = 1.3 \pm 0.5$ and $\alpha_3 = 2.3 \pm 0.3$ (Kroupa et al., 1993). La dernière valeur peut être égale à 2.7 suivant si l'on prend en compte les sources les plus massives ($\approx 10 M_{\odot}$ pour les étoiles évoluées) de la distribution et qui traduit le peu d'étoiles massives présentes dans les amas (statistique faible). Il est à noter que Φ_{Kroupa} utilise une variation de masse dm au lieu d'une variation logarithmique comme pour $\Phi_{Salpeter}$. Les deux relations sont liées par la formule suivante :

$$\Phi_{Kroupa} = \frac{1}{m \ln(10)} \Phi_{Salpeter} \quad (4.11)$$

Ainsi, $\alpha = \Gamma + 1$ et les valeurs de Γ et α_3 sont cohérentes entre elles. D'autres fonctions

ont été présentées afin d'ajuster au mieux l'IMF. [Maschberger \(2013\)](#) fournit un résumé de toutes ces fonctions. Cependant ces fonctions tendent à représenter la même notion (l'IMF) avec des paramètres différents. Nous allons nous concentrer sur la fonction présentée par Pavel Kroupa ([Kroupa et al., 1993](#)) dans la suite de notre étude.

4.3.2.2 Application aux données

Nous avons vu dans le chapitre précédent que les clumps ne possèdent pas tous des estimations de distance ce qui diminue leur nombre et ainsi les statistiques dans l'étude de la CMF. Cependant, nous avons vu qu'il était possible de calculer les masses des clumps ne possédant pas d'HDEs dans la section 4.3 (équation 3.5). De plus cette formule nous permet de recalculer la masse des clumps afin qu'elle corresponde à la distance héliocentrique de l'amas en question (le pic de la distribution des HDEs). La figure 4.16 montre la forme de la CMF ainsi que les trois segments ajustés pour les masses recalculées des clumps possédant des HDEs (partie gauche), l'ensemble des clumps avec recalcule des masses (partie centrale) et seulement les clumps possédant des HDEs sans recalcule de masses (partie droite). La distribution mélange l'ensemble des clumps des amas car nous n'avons pas assez de statistiques pour étudier les amas un par un. La taille du bin correspond à une règle statistique, $W = 2(IQR)/\sqrt[3]{N}$, où W est la taille du bin, IQR l'écart interquartile et N le nombre de sources de la distribution ([Freedman & Diaconis, 2011](#)). Cependant le choix de la taille de bin n'affecte pas significativement les valeurs de α car le nombre de sources est important (entre 8000 et 9000 suivant si on considère les clumps avec ou sans HDEs). [Rosolowsky \(2005\)](#) montre qu'au-delà de $N > 500$, la taille du bin n'affecte plus significativement les résultats. Les barres d'erreur correspondent à la valeur à 1σ des erreurs de Poisson, $\sigma = \sqrt{N}/\Delta m$, relatives aux comptages des sources. Nous n'avons pas considéré l'erreur sur le calcul des masses, provenant principalement de l'erreur sur les distances héliocentriques, car les erreurs de Poisson constituent la principale source d'erreur sur la CMF. Concernant la technique d'ajustement, les 3 segments sont ajustés en même temps et nous cherchons à minimiser le χ^2 total. Les pentes de ces 3 segments sont ensuite comparées aux valeurs des pentes de l'IMF de Kroupa mentionnées plus haut. Afin d'établir un lien entre la CMF et l'IMF, nous imposons une valeur fixe de $m_0 = 10 M_{\odot}$. Pour les trois CMF de la figure 4.16, nous obtenons les mêmes valeurs de masses de transition m_1 et m_2 ainsi que les mêmes valeurs de pentes, α_1 , α_2 and α_3 . Nous décidons ainsi de ne choisir que les clumps avec HDEs et sans recalculer les masses dans la suite de cette étude.

Concernant la valeur des pentes, on remarque qu'elles sont toutes les trois identiques aux valeurs des pentes de l'IMF de Kroupa. Pour aller plus loin, nous pouvons aussi étudier l'impact de l'environnement sur le profil de la CMF. En effet, MST permet d'extraire des sur-densités d'un échantillon de données mais il permet aussi d'observer les sources isolées. Celles-ci sont définies à partir de la distance avec leur plus proches voisines. Si une source n'est reliée à aucune autre source avec une distance inférieure ou égale à Λ_{cut} , alors nous la considérons comme isolée. Nous obtenons ainsi environ 4700 clumps isolés que nous comparons avec les 8200 clumps en amas possédant des HDEs. La figure 4.17 compare la CMF des clumps en amas et celle des clumps isolés.

Enfin, avant de passer aux résultats, nous avons séparé le ciel entre le premier et le quatrième quadrant afin d'observer une éventuelle asymétrie dans la distribution de

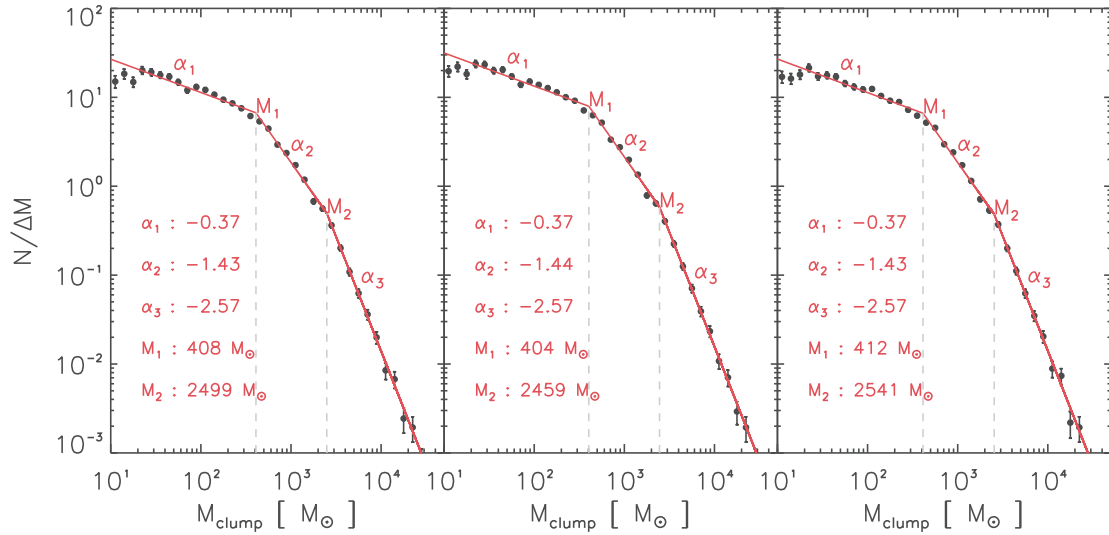


FIGURE 4.16: CMF des clumps en amas. Les lignes rouges correspondent aux segments ajustés et nous fournis les pentes α_1 , α_2 and α_3 ainsi que les masses de transition m_1 and m_2 . Gauche : clumps avec HDEs. milieu : Ensemble des clumps des amas. Droite : Clumps avec recalcul des HDEs. Les trois considérations n'altèrent pas le profil de la CMF.

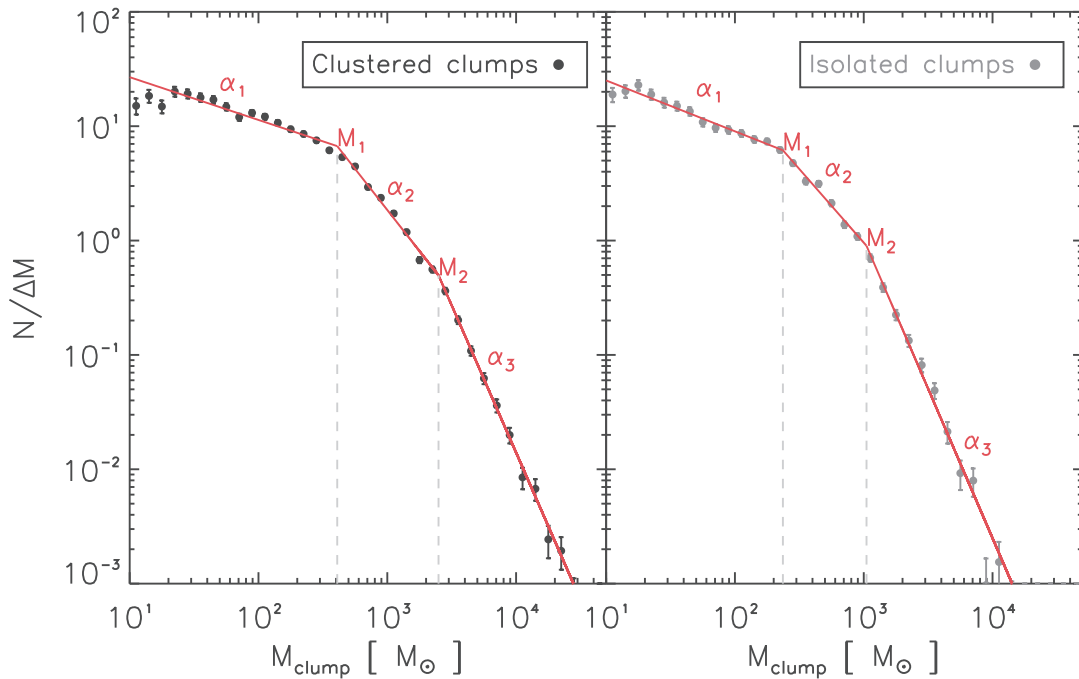


FIGURE 4.17: CMF des clumps en amas (gauche) et des clumps isolés (droite). Les lignes rouges correspondent aux mêmes segments que dans la figure 4.16.

la CMF à l'échelle galactique. Une autre possibilité serait de séparer le ciel suivant les distances héliocentriques des amas. En effet, la figure 3.15 montre que les amas sont majoritairement distribués en dessous de 5 kpc ou au dessus de 6 kpc. Cependant beaucoup plus d'amas se situent au-delà de 6 kpc, trois fois plus en réalité, ce qui entraîne une statistique réduite en dessous de 5 kpc et ainsi des fluctuations et erreurs importantes sur la distribution de la CMF. Nous aurions aussi pu séparer les données suivant leur stade évolutif. Encore une fois le nombre de clumps pré-stellaires est beaucoup plus important

que le nombre de clumps proto-stellaires ce qui sous-entend d'être en la présence d'un biais d'estimation du stade évolutif des clumps, comme évoqué dans la section 3.2.1, mais aussi de réduire la statistique des clumps proto-stellaires. Au final, dans les deux possibilités évoquées, nous ne pourrions pas conclure sur une variabilité ou sur une stabilité du profil de la CMF.

La table 4.1 montre les valeurs des pentes et des masses de transition de la CMF des clumps en amas et des clumps isolés pour l'ensemble de la partie interne du plan galactique ainsi que pour les premiers et quatrièmes quadrants. De plus les valeurs des trois pentes sont comparées avec les valeurs des pentes de l'IMF de Kroupa dans la figure 4.18. Cette comparaison montre que les pentes ne varient pas significativement entre les clumps isolés et les clumps en amas. De plus, on remarque que les pentes sont cohérentes avec les valeurs de l'IMF en considérant les erreurs à 3σ , ajoutées sur la figure.

TABLE 4.1: Résultats de l'ajustement de la CMF des clumps en amas et des clumps isolés pour l'ensemble de la partie interne du plan galactique ainsi que pour les premier et quatrième quadrants.

Clumps	α_1	α_2	α_3	$m_1 (M_\odot)$	$m_2 (M_\odot)$
En amas	0.37 ± 0.02	1.43 ± 0.03	2.57 ± 0.07	408	2 499
Isolés	0.44 ± 0.02	1.29 ± 0.05	2.47 ± 0.08	235	981
Premier quadrant					
En amas	0.15 ± 0.05	1.29 ± 0.06	2.60 ± 0.14	324	2 081
Isolés	0.29 ± 0.05	1.37 ± 0.10	2.52 ± 0.13	213	969
Quatrième quadrant					
En amas	0.46 ± 0.02	1.55 ± 0.04	2.56 ± 0.11	483	2 880
Isolés	0.53 ± 0.03	1.24 ± 0.08	2.35 ± 0.10	251	960

Cependant, nous devons faire attention à l'intervalle de masses et les types d'objets considérés ici. En effet, au-delà de $1000 M_\odot$, les clumps peuvent être considérés comme des nuages moléculaires. Considérant cette éventualité, la fonction de masse des nuages moléculaires (MCMF), qui correspond à la partie la plus massive de la CMF, est en accord avec la pente définie par Kroupa. La deuxième pente α_2 est en accord avec la pente trouvée par [Klessen & Burkert \(2000\)](#) et [Elmegreen & Clemens \(1985\)](#), $dN/dm \approx m^{-1.5}$ qui utilisèrent différentes simulations pour enquêter sur la fragmentation des nuages moléculaires. Cependant ils trouvent une valeur plus petite de la pente α_3 . [Tsuboi & Miyazaki \(2012\)](#) et [Tsuboi et al. \(2015\)](#) montrent que la pente α_3 aux grandes masses varie vers le centre galactique et correspond à la valeur de notre étude, au dessus de $10^4 M_\odot$ c'est-à-dire dans la partie la plus massive de la CMF de notre étude.

Dans le même cadre, nous verrons dans la section 4.5 que les masses des clumps sont corrélées aux distances héliocentriques jusqu'à 6 kpc. En mélangeant des clumps de masses différentes, entre $10 M_\odot$ et $10\,000 M_\odot$, nous pourrions supposer que nous mélangeons des types différents de clumps. Cependant, nous devrions observer un *saut* et/ou

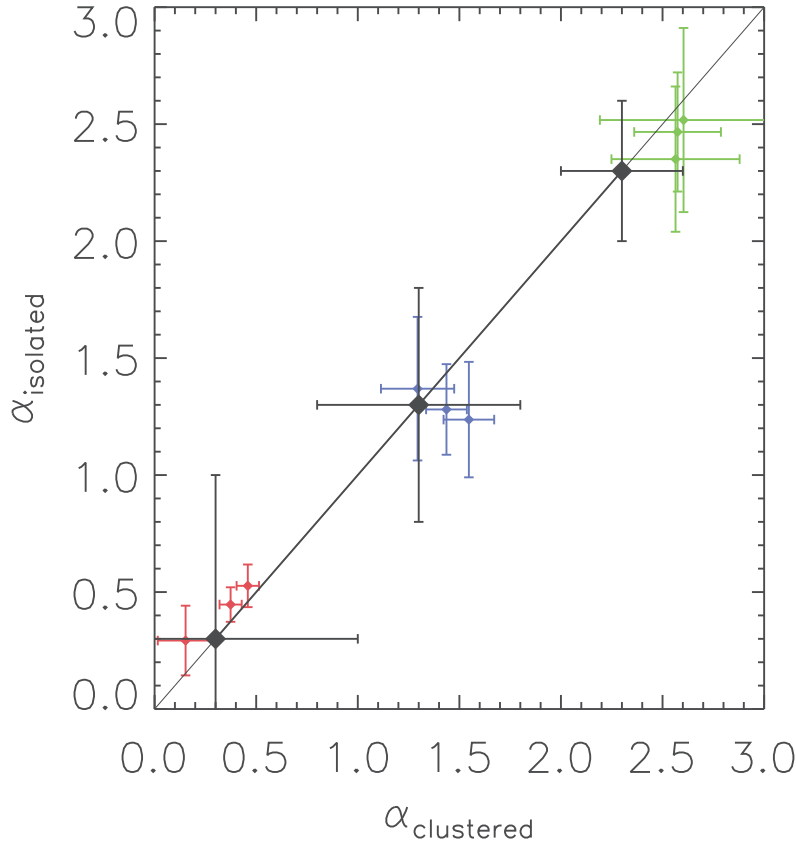


FIGURE 4.18: Comparaison des valeurs des trois pentes de la CMF provenant de la table 4.1 entre les clumps isolés ($\alpha_{isolated}$) et les clumps en amas ($\alpha_{clustered}$). α_1 est en rouge, α_2 est en bleu et α_3 est en vert. Les valeurs des pentes de l'IMF de Kroupa sont aussi montrées en noir. La ligne noire représente la zone de valeurs égales entre les clumps isolés et les clumps en amas.

une *bosse* dans le profil de la CMF entre les clumps de faibles masses et les clumps massifs (Moore et al., 2007; Reid et al., 2010), comme montré dans la figure 4.19. Le fait de ne pas observer ce profil particulier nous indique que les clumps font partie de la même distribution. De plus, la distribution est dominée par les clumps au dessus de 6 kpc car leur nombre est beaucoup plus important que les clumps situés en dessous de 5-6 kpc.

Michel et al. (2011) montrent que la CoMF ne suit pas nécessairement le profil de l'IMF dans les groupes de coeurs les plus denses, surtout pour les étoiles massives. De même Hatchell & Fuller (2008), en étudiant les coeurs pré-stellaires et proto-stellaires dans Perseus, montre que la CoMF devrait être plus pentue que l'IMF aux grandes masses. Cependant ils montrent aussi qu'il est possible d'obtenir une valeur de pente cohérente avec celle de l'IMF si les coeurs pré-stellaires se fragmentent et forment des systèmes multiples de proto-étoiles.

D'autres travaux montrent un lien fort entre la CMF et l'IMF et une invariabilité selon l'environnement. Nous pouvons notamment citer Wienen et al. (2015) qui, en étudiant les clumps provenant d'ATLASGAL (Schuller et al., 2009), montrent une invariabilité de la CoMF dans plusieurs amas enfouis. Simpson et al. (2008) montrent que la CoMF correspond à l'IMF dans le nuage moléculaire Ophiuchus. Alves et al. (2007) suggèrent que l'IMF trouve son origine dans la CoMF. Bastian et al. (2010) suggèrent aussi que l'IMF est invariant suivant les environnements. Si la CMF suit le profil de l'IMF, une différence

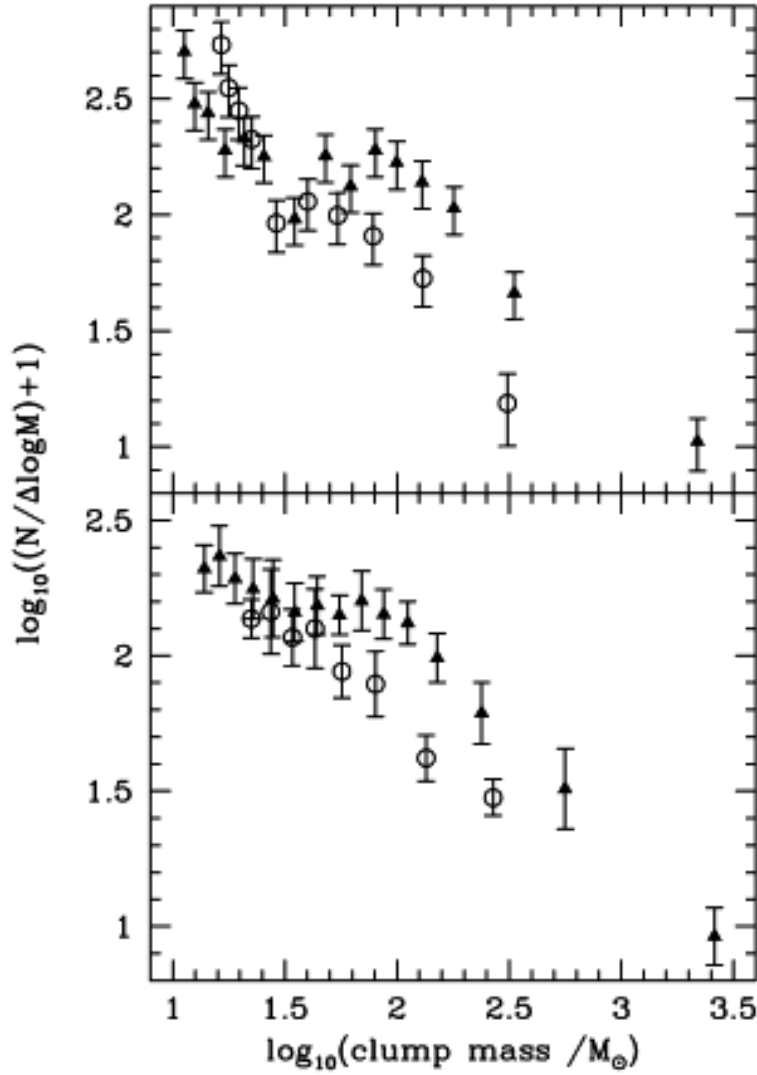


FIGURE 4.19: Distributions de la masse des clumps dans W3 pour une zone de forte densité (triangles) et une zone de faible densité (cercles). Les deux parties de la figure correspondent à deux calculs différents des masses. Dans toutes les distributions, nous pouvons observer un *saut* suivie d'une *bosse* entre $\log_{10}(m) = 1.5$ (M_{\odot}) et $\log_{10}(m) = 2$ (M_{\odot}). Cette figure est extraite de (Moore et al., 2007).

observée dans le profil de la CMF devrait entraîner une différence dans le profil de l'IMF.

Au final, malgré les remarques faites sur les types de clumps et l'intervalle de masse de la CMF, notre étude tend à montrer que le profil de la CMF suit le profil de la CoMF et de l'IMF. En d'autre terme, notre étude suggère qu'il existe un lien fort entre la CMF, la CoMF et ainsi l'IMF et que cette dernière prend son origine dans les toutes premières étapes de la formation stellaire, au moins à partir du moment où le nuage moléculaire devient instable gravitationnellement et que les premiers grumeaux pré-stellaires se forment.

Nous remarquons aussi que les valeurs des masses de transition sont plus importantes pour les clumps en amas que pour les clumps isolés. Nous allons étudier cette différence dans la section suivante.

4.3.2.3 Multiplicité et efficacité stellaire

En considérant que les clumps sont des objets similaires, indépendamment de la distance et de leur masse, et en supposant le lien entre la CMF et l'IMF montré précédemment, il existe un coefficient de proportionnalité entre la masse des clumps et celui des étoiles plus évoluées. Ce coefficient comprend lui même deux notions, lorsque les objets sont suffisamment massifs et de tailles importantes, que sont l'efficacité de formation stellaire (SFE) et le facteur de multiplicité, n_* , qui caractérise le nombre d'étoiles potentiellement formées ou formables dans un clumps. Ainsi, la relation entre les valeurs de masse dans l'IMF et celles dans la CMF s'écrit :

$$m_{IMF} = \frac{\epsilon}{n_*} m_{CMF} \quad (4.12)$$

Où ϵ représente la SFE, m_{IMF} est la masse de l'étoile correspondant à la courbe de l'IMF et m_{CMF} est la masse du clump correspondant à la courbe de la CMF. En analysant la CMF de la section précédente, on peut s'apercevoir que les masses de transition sont plus importantes pour les clumps en amas que pour ceux isolés. Cela signifie une différence dans la distribution de n_* ou une variation de la SFE.

En supposant une valeur moyenne de SFE de $\epsilon = 0.3 \pm 0.1$ (Alves et al., 2007) et une fonction de distribution de masses, il est possible d'évaluer la distribution de n_* . Une étude similaire a été faite par Holman et al. (2013) sans utiliser une valeur fixée de la SFE et en cherchant une valeur moyenne de n_* , \bar{n}_* . Cependant la dégénérescence entre la SFE et \bar{n}_* oblige à contraindre l'un de ces paramètres. Hatchell & Fuller (2008) et Goodwin et al. (2008) utilisèrent un modèle multiple, *fully multiple model* qui permet de fixer la valeur de \bar{n}_* afin d'évaluer celle de la SFE. Cependant leur modèle se limite à l'étude de systèmes composés de 2 ou 3 coeurs et ne peut pas s'appliquer à notre étude.

Afin d'évaluer rapidement les valeurs possibles de n_* , nous utilisons sa valeur moyenne \bar{n}_* . Ce qui nous intéresse aussi ici, c'est de comparer les clumps en amas et les isolés. Pour cela il suffit d'obtenir le rapport entre les masses des clumps en amas et celles de clumps isolés, à partir des masses de transition et en considérant une SFE fixée à 0.3, afin d'obtenir le rapport entre le facteur de multiplicité moyen $\bar{n}_{*, isolated}$ et $\bar{n}_{*, clustered}$:

$$R_n = \frac{m_{CMF, isolated}}{m_{CMF, clustered}} = \frac{\bar{n}_{*, isolated}}{\bar{n}_{*, clustered}} \quad (4.13)$$

Les masses de transition m_1 and m_2 sont bien adaptées pour ce calcul. Pour l'ensemble des trois CMFs nous trouvons une valeur de R_n comprise entre 0.3 et 0.65 ce qui se signifie que $\bar{n}_{*, clustered}$ est en moyenne deux fois supérieur à $\bar{n}_{*, isolated}$. Les clumps en amas, connus pour être des zones particulièrement denses, forment soit plus de proto-étoiles, soit des étoiles plus massives que les clumps isolés. Cependant cette dernière supposition serait en contradiction avec le profil universel de l'IMF. Une étude approfondie, à une meilleur résolution, d'un clump en amas et d'un clump isolé permettrait probablement d'en apprendre plus sur le mécanisme de formation stellaire et de confronter les différents scénarios de formation d'étoiles massives (Krumholz et al., 2005).

4.4 Association avec des objets de la littérature

Nous nous sommes intéressés à la distribution spatiales des amas, des clumps les composant ainsi que de la masse de ceux-ci. Cependant nous n'avons pas encore évoqué le lien entre les amas et les composantes du MIS. Nous allons voir ici si les amas peuvent être associés à deux de ces composantes, les régions H II et les IRDCs. Enfin, nous terminerons par une comparaison avec des amas et régions de formation stellaire provenant de la littérature.

4.4.1 Régions H II

Comme nous l'avons vu dans le chapitre 1, les régions H II sont liées à la formation stellaire par la PDR qui les entoure. Ces régions sont particulièrement liées à la formation d'étoiles massives (Urquhart et al., 2013) puisqu'elles abritent, ou ont abrité, dans leur centre une étoile massive de type OB. De plus c'est typiquement dans ce genre de région que l'on s'attend à observer une formation induite d'étoiles (Deharveng et al., 2005; Kendrew et al., 2012; Thompson et al., 2012), scénario en faveur de la formation d'étoiles massives.

Dans cette section, nous utilisons les catalogues de régions H II provenant de Anderson et al. (2014)¹ (And2014) et Paladini et al. (2003) (Pal2003) afin d'effectuer une identification croisée par position avec les amas de notre étude. Le catalogue de And2014 regroupe 8400 régions H II évaluées à l'aide des données de WISE. Une grande partie de ces régions sont connues tandis qu'une autre majorité ne possèdent pas d'émissions dans le domaine radio contrairement au catalogue de Pal2003, qui regroupe 1142 régions H II. 7589 des régions H II de l'ensemble des ces deux catalogues sont localisées dans le même intervalle de coordonnées galactiques que les amas, avec respectivement 6561 et 1008 pour les catalogues And2014 et Pal2003.

Une région H II n'est associé avec un amas que si au moins un clumps est localisé dans le rayon de cette région H II. Ceci permet de garantir que l'amas est au moins localisé à l'endroit de la PDR de la région H II. Au total 274 amas (55%) sont chacun associés avec au moins une région H II. Dans le sens contraire, 443 régions H II sont chacune localisées avec au moins un amas. Il est évident que le nombre de régions H II associées serait bien plus important si les sur-densités du catalogue archive était suffisamment fiables pour être prises en compte dans cette étude.

Pour aller plus loin, nous pouvons extraire les régions H II aussi grandes que les amas, c'est-à-dire de l'ordre de la dizaine de parsecs. Ceci permet d'éliminer les régions H II localisées dans le coeur de l'amas et qui ne seraient à associer qu'avec certains clumps. Nous trouvons que 178 amas sont associés avec 177 régions H II. 168 de ces régions proviennent du catalogue de And2014 et seulement 9 du catalogue de Pal2003. La principale raison de ce faible nombre d'associations avec le catalogue de Pal2003 est que les régions H II sont plus petites que les amas lorsque l'on parle de taille angulaire, c'est-à-dire sans utilisation des distances héliocentriques. Ce dernier point caractérise la limite principale avec les catalogues de notre étude. En effet, seul le catalogue de And2014 fournit des distances héliocentriques. Sur les 8300 régions H II, 1400 possèdent une estimation de distance. Sur les 496 amas, seulement 54 sont associés avec ces régions H II (58 d'entre elles). Si on associe les distances avec la taille des régions H II, seulement 24 amas sont associés avec 19 régions H II de même taille. Cependant la distance n'est

1. Les régions H II sont visibles à cette adresse : <http://astro.phys.wvu.edu/wise/>

pas un critère décisif pour l'identification croisée car en analysant les images Herschel, on remarque que la majeure partie des clumps des amas (les 168 possédant une région H II de taille comparable) sont associés avec la PDR de ces régions H II.

Comme nous l'avons évoqué précédemment, la principale limite à cette identification est la taille des amas ainsi que le nombre de sur-densités rejetées dans le catalogue archive. Il est difficile de trouver des régions H II de 10-20 pc, la plupart étant des régions ultra-compactes (UCH II) ou compactes.

Enfin, nous pouvons comparer la distribution de masse des clumps appartenant à des amas avec ou sans régions H II. Nous confrontons cette distribution comprenant les clumps des 274 amas associés avec au moins une région H II avec les 222 amas restants. La figure 4.20 montre cette distribution avec en rouge les amas associés à au moins une région H II et en bleu les autres amas. Aucune différence ne peut être constatée entre ces deux distributions c'est-à-dire entre deux environnements différents. Ceci rejoint ainsi le résultat de la section 4.3.2.2.

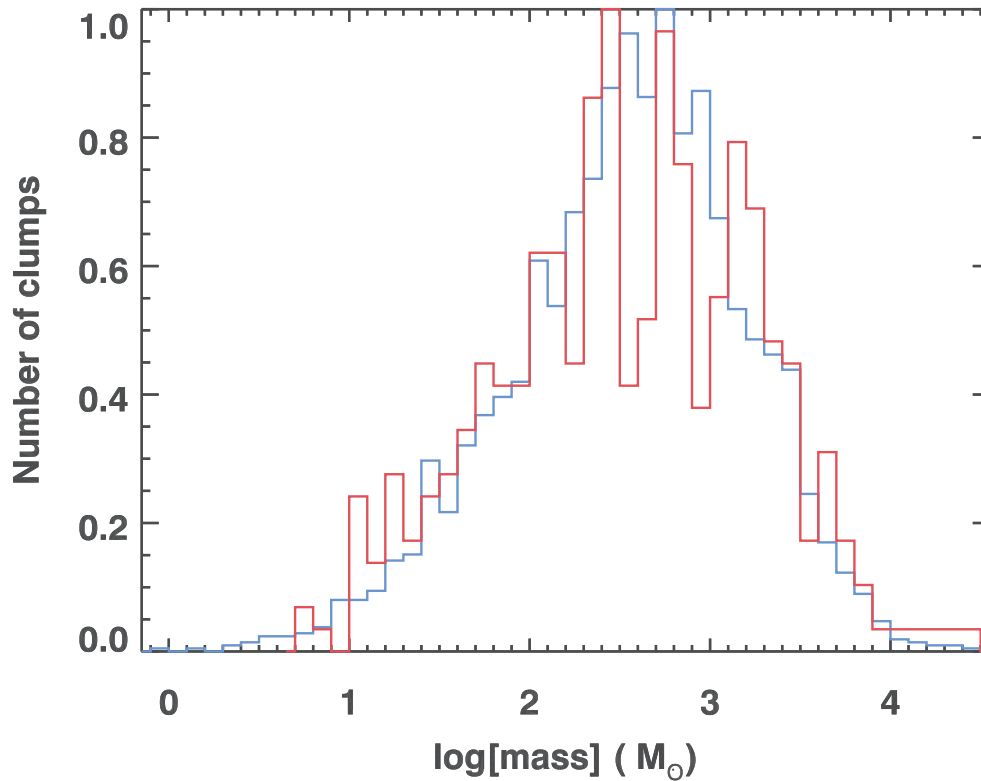


FIGURE 4.20: Distribution des masses des clumps en amas. Rouge : Amas possédant une ou plusieurs régions H II. Bleu : Amas sans association avec une région H II.

Les régions H II associées avec les amas sont inscrits dans la dernière colonne de la table de l'annexe B montrant les différentes colonnes du catalogue d'amas.

4.4.2 IRDCs

Les régions H II ne sont pas les seuls traceurs de la formation d'étoiles massives. Les IRDCs font aussi partie de cette catégorie. Dans notre étude nous effectuons une iden-

tification croisée avec deux catalogues d'IRDCs, celui de [Simon et al. \(2006\)](#) (Sim2006) et celui de [Peretto & Fuller \(2009\)](#) (Per2009). Nous utilisons deux catalogues d'IRDCs, alors qu'il existe d'autres catalogues, car les deux présentés ici chevauchent ces autres catalogues. Le catalogue de Sim2006 utilise la bande à $8.3\ \mu\text{m}$ du télescope MSX et trouve 10 931 IRDCs localisés majoritairement dans les premier et quatrième quadrants c'est-à-dire dans la zone de notre étude. De même les IRDCs du catalogue de Per2009, au nombre de 11 303, sont localisés dans principalement dans cette zone. Ce deuxième catalogue utilise les données issues des relevés GLIMPSE et MIPS GAL. Ces deux catalogues sont complémentaires avec un total de 17 109 IRDCs uniques trouvés à partir de leur combinaison. Un chevauchement de seulement 20% est observé entre les deux ([Peretto & Fuller, 2009](#)).

Les IRDCs sont en général de la taille des clumps de notre étude, ainsi la comparaison ne peut se faire avec un amas entier. L'association se fait seulement par position car les distances ne sont pas fournies dans les catalogues d'IRDCs. Seulement 38% des clumps (sur l'ensemble des clumps du catalogue *product*) tombent dans l'ellipse d'un IRDC, calculée à partir du demi-petit axe, demi-grand axe et angle de position fournit par les deux catalogues. Ce nombre est similaire à celui trouvé par [Billot et al. \(2011\)](#) qui étudiaient la distribution des clumps dans les SDPFs d'HIGAL. Cependant ils montrent que 97% des associations sont faites avec des clumps en amas. Dans notre étude, nous trouvons que ce nombre descend à 10%. La différence s'explique par l'identification des clumps ainsi que les critères de sélections des amas. En effet, nous n'utilisons que le catalogue d'amas alors qu'une partie non négligeable des sur-densités se retrouvent dans le catalogue archive. En utilisant ainsi l'ensemble des deux catalogues, d'amas et de sur-densités, le nombre d'association avec les clumps en amas passe de 10% à 92%. Cette fois, ce chiffre correspond à celui de l'étude de [Billot et al. \(2011\)](#). Ce résultat montre que les amas ou les sur-densités, si elles sont réelles, sont propices à former des étoiles plus massives en raison de leur environnement dense.

De plus, 333 amas (67%) possèdent soit des clumps associés avec des IRDCs, soit les IRDCs sont dans l'ellipse formée par l'amas. Cela signifie qu'à l'intérieur de ces amas, des scénarios de formation stellaire différents ont lieu à des localisations différentes mais aussi que des variations de densité importantes peuvent y être observées.

En règle générale nous devrions être capable d'obtenir plus d'associations. La raison principale est que les IRDCs ne sont observés que dans des régions de fort contraste avec la luminosité du fond.

4.4.3 Comparaison avec des catalogues d'amas enfouis pré-existants

Il existe différents catalogues d'amas stellaires dans la littérature. Nous allons utiliser ces catalogues afin d'identifier par position les amas enfouis de notre étude connus auparavant. Pour cela nous allons utiliser deux catalogues provenant de [Majaess \(2013\)](#) (Maj2013) et [Morales et al. \(2013\)](#) (Mor2013). Maj2013 identifiaient des YSOs à partir des données WISE et 2MASS et ont extrait des amas enfouis à partir d'une observation visuelle de leur distribution spatiale. Le catalogue se compose de 230 amas enfouis au total, mais seulement 41 sont localisés dans le même intervalle de coordonnées galactiques que les amas de notre étude. Le catalogue de Mor2013 se compose d'amas ouverts et d'amas enfouis de la littérature (la liste des catalogues utilisés peut être trou-

vée dans le papier de cette étude) , ainsi que 75 nouveaux amas extraits à partir des données GLIMPSE, tous dans l'intervalle de coordonnées galactiques $-60 \leq l \leq 60$ et $-1.5 \leq b \leq 1.5$. Au total, le catalogue est formé par 695 amas dont 531 dans le même intervalle de coordonnées galactiques que les amas de notre étude. Cependant le catalogue comprend deux types principaux d'amas : les amas enfouis et les amas ouverts. Il est peu probable d'obtenir une association significative avec les amas ouverts car les étoiles les composant sont beaucoup plus âgées et évoluées que les clumps de notre étude. En effet, ces étoiles ont eu le temps de migrer de leur site de formation à cause de leur propre vitesse de déplacement et des influences gravitationnelles extérieures à ces étoiles, voir même à l'amas. Une association avec les amas de notre étude ne serait que fortuite. Sur les 531 amas, 263 sont des amas enfouis.

Les deux catalogues, Maj2013 et Mor2013, ne se recouvrent que très peu car seulement 5 amas enfouis de Mor2013 sont associés avec ceux de Maj2013. En éliminant ces 5 amas, nous pouvons effectué une identification croisée par position et à partir des diamètres des amas.

Sur les 496 amas de notre étude, seulement 30 sont associés avec des amas de Maj2013 et Mor2013. 4 d'entre eux sont associés avec plusieurs amas de la littérature. Cela s'explique par la différence de taille entre les amas de notre étude et ceux de Maj2013 et Mor2013 comme nous pouvons le voir sur la figure 4.21.

Le faible nombre d'associations tant à montrer que plus de 90% des amas de notre étude sont de nouveaux amas, encore inconnus et inexplorés.

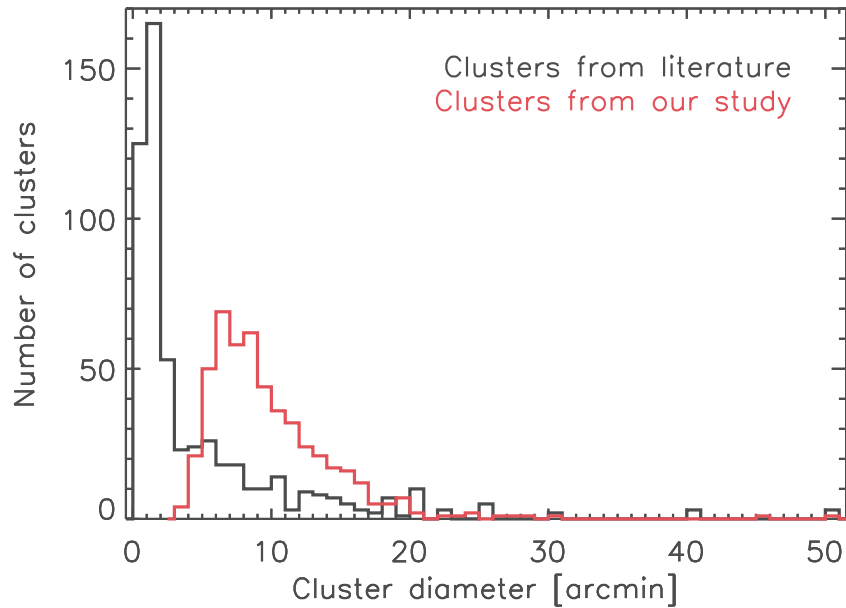


FIGURE 4.21: Distribution des diamètres des amas de Maj2013 et Mor2013 (en noir) comparé aux diamètres des amas de notre étude (en rouge).

4.5 Discussions et biais observationnels

La distribution fractale des amas implique qu'une distribution non aléatoire des clumps peut être observée mais aussi que certains clumps sont plus espacés que dans le cadre

d'une distribution centrée. On peut étudier la distance séparant les clumps en amas et un possible biais en les comparant avec la longueur de Jeans des nuages moléculaires (Equation 4.14), qui caractérise l'habilité d'une sphère à s'effondrer, ainsi que la résolution aux longueurs d'onde $70 \mu m$ et $500 \mu m$.

$$\lambda_J = \sqrt{\frac{15k_B T}{4\pi G \mu n m_p}} \quad (4.14)$$

k_B est la constante de Boltzmann, T est la température du nuage, n est la densité volumique moyenne du nuage, m_p la masse du proton et μ est le poids moléculaire moyen. On considère ici trois températures de 10, 20 et 30 K. La densité volumique moyenne est fixée à 10^5 cm^{-3} et la masse par particule moyenne est fixée à $\mu m_p = 4 \times 10^{-27} \text{ kg}$ avec $\mu \approx 2.5$ en supposant un nuage composé à 80% d'Hydrogène et 20% d'Hélium. Ces considérations induisent des valeurs de la longueur de Jeans entre 0.05 pc et 0.1 pc. Cette longueur serait plus grande d'un facteur deux si on considérait un nuage avec un champ magnétique et sous une pression externe au nuage. La figure 4.22 montre la relation entre les distances entre les clumps en parsec (projetée sur le plan du ciel) calculées à partir de la taille angulaire des branches Λ et des distances héliocentriques des l'amas. La figure montre aussi la taille en parsec de la résolution à $70 \mu m$ et à $500 \mu m$ dans l'intervalle de distances que couvre notre échantillon. La longueur de Jeans est aussi montrée (0.1 pc). Toutes les distances entre clumps sont plus grandes que la résolution à $70 \mu m$, les longueurs de Jeans et plus de 90% sont plus grandes que la résolution à $500 \mu m$. De plus la distance entre les clumps est une distance projetée. Cela signifie que les clumps sont séparés par une distance plus grande que leur empreinte laissée dans le nuage moléculaire, caractérisant l'espace créé par une fragmentation. Ainsi les résolutions des instruments PACS et SPIRE introduisent un biais en ne permettant pas la détection de clumps espacés d'une distance inférieure à la longueur de Jeans.

Les figures 4.23, 4.24, 4.25 et 4.26 montrent respectivement différentes propriétés des amas, la densité surfacique de l'amas (pc^{-2}), la distance moyenne séparant les clumps, l'aire des amas, le nombre de clumps, en fonction de la masse moyenne des clumps. La barre de couleurs correspond aux distances héliocentriques des amas en kpc. On remarque une corrélation significative entre ces propriétés, la masse moyenne des clumps et la distance héliocentrique, sauf pour le nombre de clumps. La figure 4.23 montre que plus la valeur de la densité surfacique de l'amas est petite et plus la masse moyenne des clumps est petite. Il existe aussi une relation entre la masse moyenne des clumps et la distance séparant chacun d'entre eux comme montré dans la figure 4.24. On observe aussi que l'aire des amas augmente sans augmenter le nombre de clumps (figures 4.25 et 4.26).

Comme mentionné plus haut, les résolutions des instruments PACS et SPIRE affectent la distance entre les clumps. Elles affectent aussi les masses car plus la distance héliocentrique devient importante et plus les clumps réuniront de matières (gaz et poussières) à l'intérieur de l'aire formée par la résolution. Nos données sont probablement affectées par un biais sur les masses qui induit un biais sur la distance entre les clumps. Il est possible qu'à partir d'une certaine distance, il ne soit plus possible de détecter certains clumps en raison de la sensibilité des instruments PACS et SPIRE et de la confusion sur la ligne de visée. Nous pouvons estimer la distance critique à partir de laquelle un clump d'une certaine masse n'est plus observable en utilisant la limite de complétude à 90% en flux pour chacune des longueurs d'onde du relevé Hi-GAL. [Molinari et al. \(2016a\)](#) nous

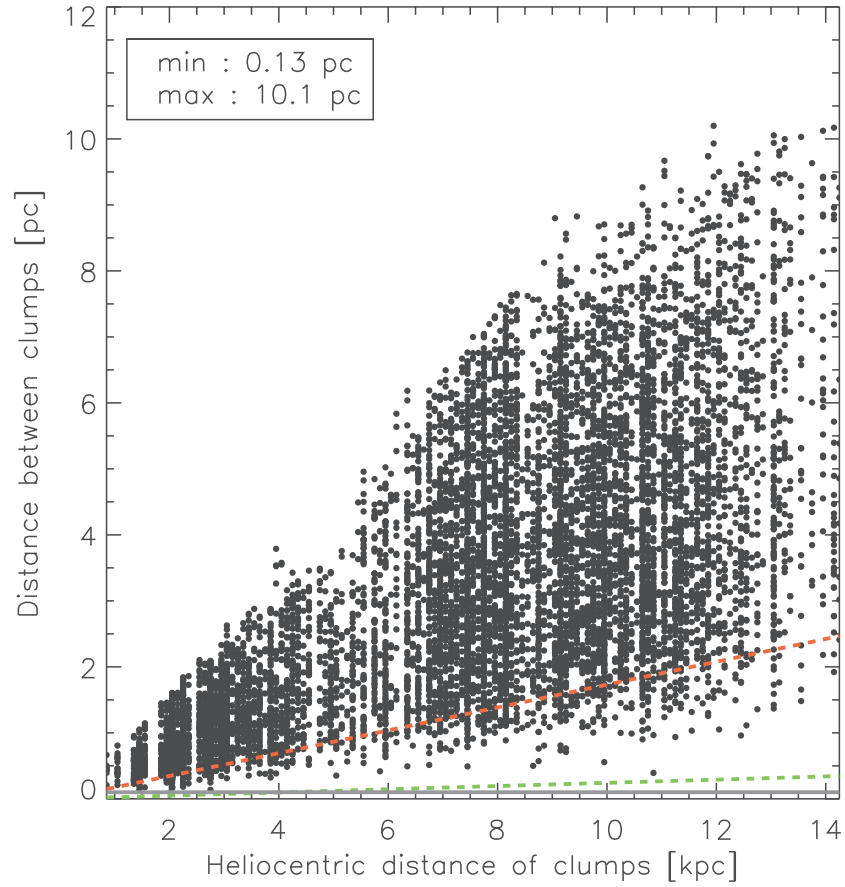


FIGURE 4.22: Distribution de la distance des clumps en amas. La ligne verte représente la taille de la résolution à $70 \mu m$. La ligne orange correspond à la même taille mais à $500 \mu m$. La ligne grise correspond à la longueur de Jeans (0.1 pc).

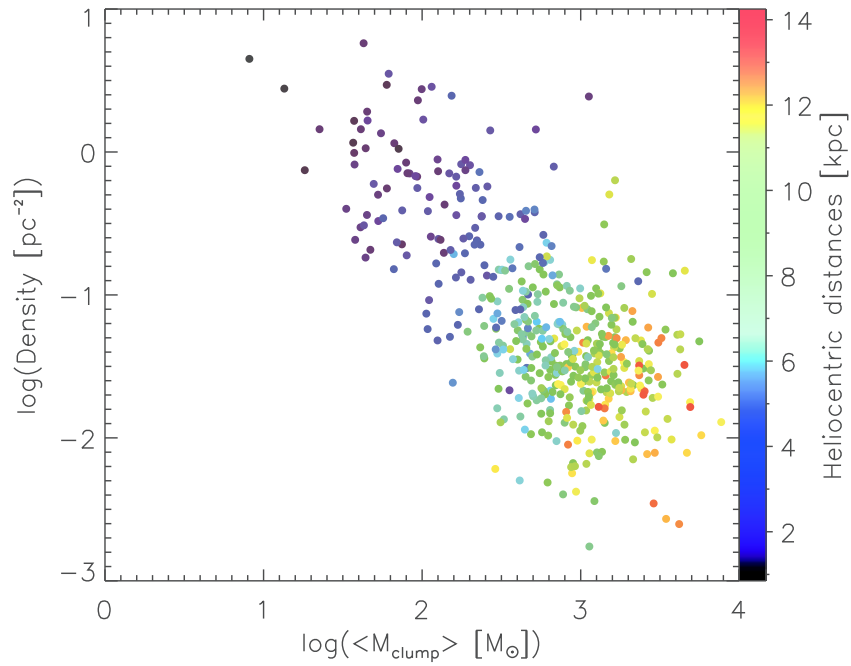


FIGURE 4.23: Logarithme de la densité surfacique des amas (pc^{-2}) en fonction du logarithme de la masse des clumps. Le panel de couleur correspond aux distance héliocentriques de amas en kpc (barre de couleurs).

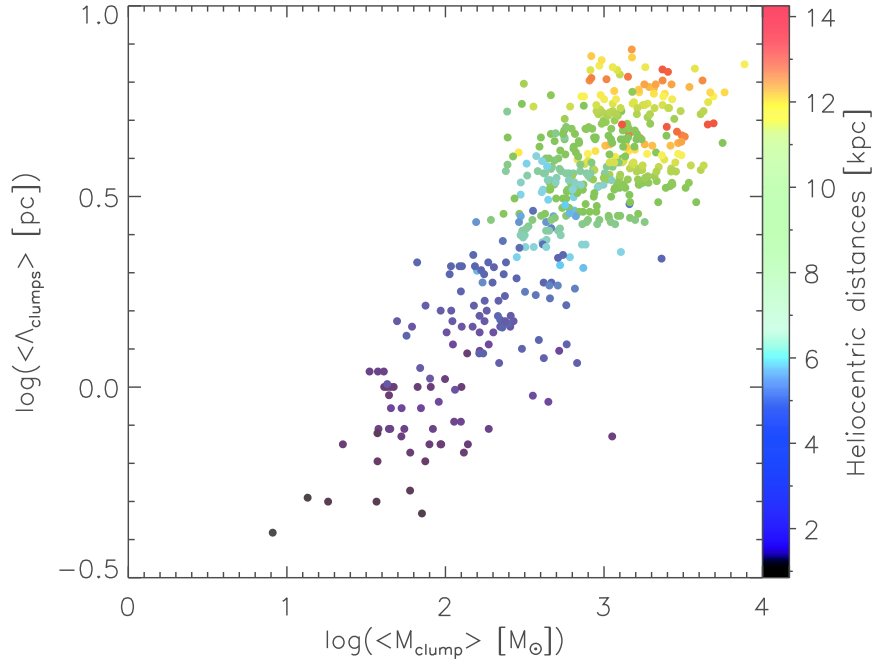


FIGURE 4.24: Logarithme des distances moyennes entre les clumps, évaluées à partir des tailles des branches λ du MST, en fonction du logarithme de la masse des clumps. Le panel de couleur correspond aux distance héliocentriques de amas en kpc (barre de couleurs).

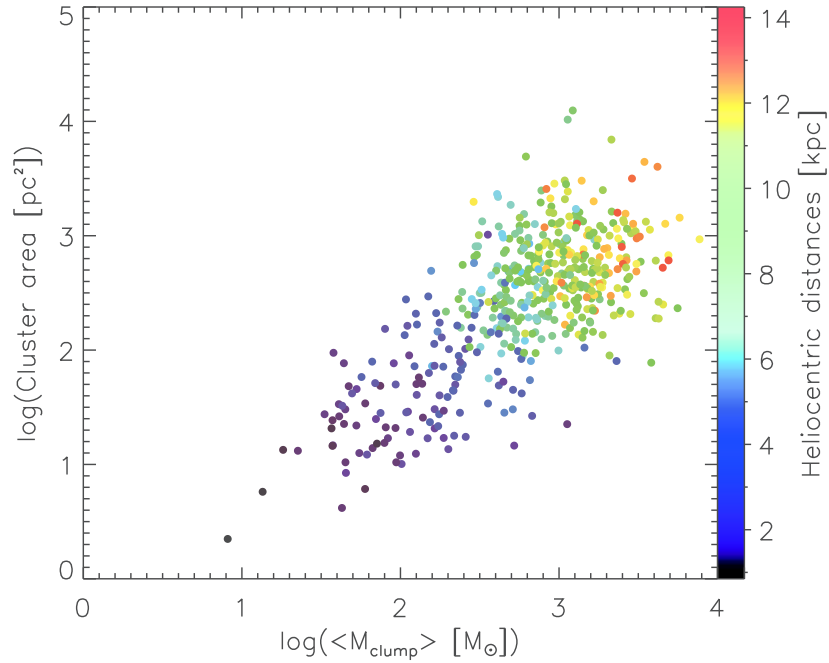


FIGURE 4.25: Logarithme de l'aire des amas (pc^2) en fonction du logarithme de la masse des clumps. Le panel de couleur correspond aux distance héliocentriques de amas en kpc (barre de couleurs).

renseigne sur ces limites de complétude. Il est à noter que ces limites dépendent de la zone sondée et notamment de la longitude. Notre calcul ne sert ici qu'à donner un ordre d'idée de l'influence de la sensibilité du télescope Herschel sur la détectabilité des clumps.

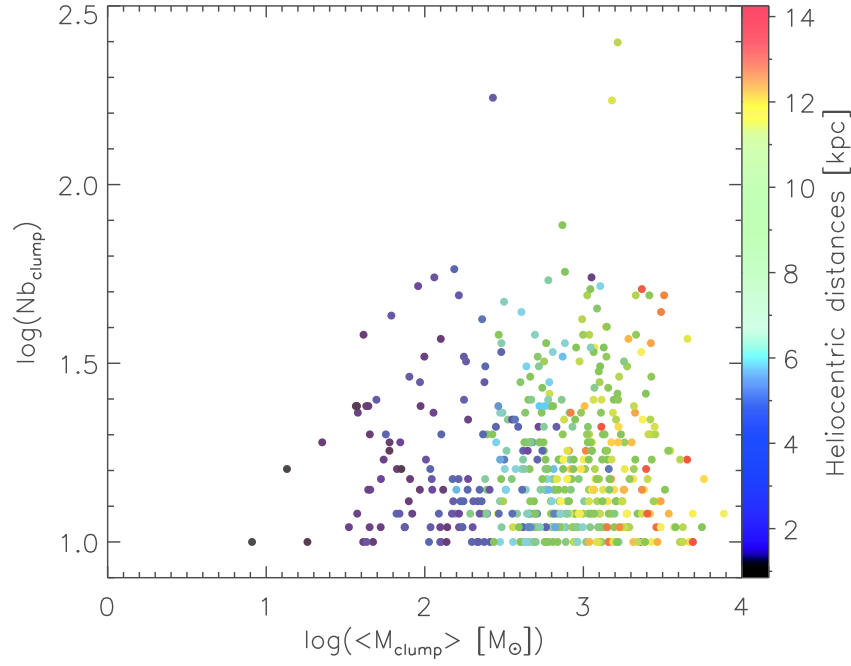


FIGURE 4.26: Partie (d) : Logarithme du nombre de clumps par amas en fonction du logarithme de la masse des clumps. Le panel de couleur correspond aux distance héliocentriques de amas en kpc (barre de couleurs).

En inversant l'équation 2.6 on obtient :

$$D = \sqrt{\frac{M k_{ref} \left(\frac{\nu}{\nu_{ref}} \right)^2 B(T, \nu)}{F_\nu}} \quad (4.15)$$

La distance héliocentrique est calculée en reprenant les valeurs déjà mentionnées dans le chapitre 3, $k_{ref} = 0.1 \text{ cm}^2 \text{ g}^{-1}$ à $\lambda_{ref} = 250 \mu\text{m}$ (Hildebrand, 1983), ainsi qu'en considérant une loi de corps noir à $T = 15 \text{ K}$. Les limites de complétudes sont estimées respectivement à 0.7, 1.5, 2.0, 2.0, 3.0 Jy pour 70, 160, 250, 350 et 500 μm (Molinari et al., 2016a). Le résultat est qu'au-delà de 6-7 kpc il devient difficile de détecter des clumps avec une masse inférieure à $100 M_\odot$. C'est ce que nous observons dans notre étude, la plupart des clumps de masses inférieures à $100 M_\odot$ ne sont observés qu'en dessous de 5 kpc. Si ces clumps ne sont pas observés dans les amas les plus éloignés alors la distance entre les clumps augmentent et correspond à ce qu'on observe dans la figure 4.24.

4.6 Conclusion

Nous nous sommes concentrés sur les propriétés physiques des amas enfouis dans ce chapitre. Le premier travail fut d'étudier la distribution spatiale des amas, en utilisant les distances héliocentriques estimées dans le chapitre précédent. En analysant la distribution en deux dimensions dans le plan galactique, nous observons que les amas suivent le profil des bras spiraux. Il semble d'ailleurs y avoir un décalage suivant les modèles de bras spiraux, les amas étant préférentiellement associés aux bras spiraux issues de l'analyse des régions H II plutôt que du CO. Ceci semble traduire le fait que les amas enfouis de notre étude commencent légèrement à se dissocier du gaz. Il peut toutefois s'agir

d'un décalage dû à la fois aux erreurs sur les distance cinématiques des clumps composant les amas et les erreurs sur l'estimation du profil des bras. Cependant ce décalage est systématique avec les bras issues de carte CO (Englmaier et al., 2011). Nous nous sommes ensuite aperçus, en calculant la hauteur des amas par rapport au plan galactique, que ceux-ci semblaient correspondre au *warp* de la galaxie généralement associé avec le disque HI.

Nous avons ensuite étudié la distribution spatiale des clumps au sein de chaque amas en utilisant une relation basée sur la méthode MST. On en déduit que la distribution qui gouverne la grande majorité des amas est une distribution fractale. Cela signifie que les amas sont formés de sous structures. Cela peut se comprendre car une partie des clumps sont associés avec des structures filamenteuses se regroupant et formant des noeuds. L'analyse de la dimension d'Hausdorff, caractérisant le degré de sous-structure au sein d'une distribution, nous apprend que les amas ont une tendance significative à posséder la même dimension fractale que le MIS à l'échelle galactique.

Cette étude s'est poursuivie en analysant les masses des clumps. Nous avons tenté d'observer une ségrégation de masse au sein des amas, cependant la nature fractale de ceux-ci ainsi que le nombre réduit de clumps dans chacun des amas nous empêche d'observer une ségrégation de masse centrée. Nous sommes donc rapidement passé à une partie plus intéressante en raison de la possibilité d'obtenir un résultat avec ce type de données, nous parlons ici du profil de la distribution de masse des clumps (CMF). Nous avons confronté ce profil à celui de l'IMF et plus particulièrement au profil composé de trois segments de l'IMF originellement introduit par Pavel Kroupa. On trouve que le profil de masse des clumps suit particulièrement bien l'IMF de Kroupa, des petites aux grandes masses (10 à 10 000 M_{\odot}). Ce résultat nous indique que l'IMF prend son origine dans les toutes premières étapes de la formation stellaire.

Nous avons ensuite comparé la distribution de masse des clumps en amas avec celle des clumps isolés, c'est à dire dont MST n'a pu trouver aucune association ne serait-ce qu'avec un seul autre clump. Il en résulte que les CMF des clumps en amas et des clumps isolés suivent le même profil. Ceci nous indique que l'environnement n'aurait pas d'influence sur la CMF et ainsi renforce l'hypothèse d'un profil invariant. Nous trouvons cependant que les masses de transition, entre les différents segments de la CMF, sont quasiment deux fois plus élevées pour les clumps en amas que pour les clumps isolés. En fixant une efficacité de formation stellaire, cela signifierait qu'un plus grand nombre de proto-étoiles se formerait à partir des clumps en amas.

Enfin, nous finissons l'étude des amas en analysant les associations avec deux traceurs de formation d'étoiles massives, les régions H II et les IRDCs. 55% des amas sont associés avec au moins une région H II et 67% possèdent au moins un IRDC. Le profil de la distribution de masse des amas possédant au moins une région H II et ceux n'en possédant pas est la même ce qui rejoint l'hypothèse de l'universalité de la CMF et de l'IMF par la même occasion.

Troisième partie

Conclusion et perspectives

CHAPITRE

5

CONCLUSION ET PERSPECTIVES

Sommaire

5.1	Conclusion	108
5.1.1	Catalogue de clumps	108
5.1.2	Catalogue d'amas	108
5.1.3	Analyse des amas	109
5.2	Perspectives	110

5.1 Conclusion

Nous savons que les processus de formation stellaire obéissent à certaines lois, tant à l'échelle locale des objets stellaires, qu'à des échelles bien plus importantes. Nous savons notamment que les étoiles jeunes se regroupent dans des régions denses, appelées régions de formation stellaire, associées à des nuages moléculaires. Cependant les mécanismes de formation d'étoiles sont encore mal compris. Il est ainsi nécessaire d'étudier l'ensemble des différents stades évolutifs des étoiles afin de comprendre cette formation en créant des scénarios dynamiques. L'arrivée du télescope Herschel a permis d'en apprendre un peu plus en étudiant les toutes premières phases de la formation stellaire.

La problématique générale de cette thèse s'inscrit dans ce cadre. Il s'agit de comprendre comment les régions de formation stellaire jeunes interagissent avec leur environnement et quelles sont leurs spécificités.

Nous nous sommes tout particulièrement intéressé aux régions de formation stellaire très jeunes, de quelques millions d'années au maximum, en utilisant les données provenant du relevé Hi-GAL.

5.1.1 Catalogue de clumps

La première étape de mon doctorat fut de créer un catalogue de sources jeunes. Nous avons couplé les catalogues de points à chacune des 5 longueurs d'onde du relevé Hi-GAL avec un algorithme d'identification croisée, SPEC-FIND, capable d'identifier des sources à partir de leur position et de leur distribution spectrale d'énergie. Après avoir adapté l'algorithme aux données infrarouges et sub-millimétriques, nous avons obtenu un catalogue de pratiquement 30 000 sources jeunes. En utilisant divers critères de distinction des stades évolutifs de sources, nous avons trouvé que la plupart des sources jeunes sont des sources pré-stellaires, c'est à dire des sources de quelques dizaines ou centaines de milliers d'années.

Ces sources jeunes sont en fait appelées clumps en raison de leur taille et masse. En effet leur taille est typiquement de l'ordre du parsec et leur masse peut dépasser les $1000 M_{\odot}$, suivant les distances héliocentriques que nous avons imposé, ce qui signifie que plusieurs coeurs pré-stellaires ou proto-stellaires sont en cours de formation au sein de chacun d'entre eux.

Au moment de la réalisation de ce catalogue, un autre catalogue a été construit par une partie de l'équipe Hi-GAL. Nous avons comparé ce catalogue à celui construit dans notre étude et nous remarquons que les sources possèdent les mêmes propriétés, notamment en température. Dans les deux catalogues ces clumps possèdent une température moyenne de 10 K pour les clumps pré-stellaires et de 15 K pour les clumps proto-stellaires, avec une distribution plus piquée dans le cas des clumps pré-stellaires.

5.1.2 Catalogue d'amas

Nous avons ensuite construit un catalogue d'amas à partir du catalogue de clumps de l'équipe Hi-GAL. Le choix a été fait d'utiliser ce catalogue plutôt que celui construit dans notre étude car il présentait des distances héliocentriques importantes pour l'estimation des propriétés physiques des amas.

La construction du catalogue d'amas s'est faite en deux étapes. La première étape consistait à utiliser l'algorithme de détection de sur-densités, MST (Minimum Spanning Tree), afin d'évaluer les effets de fenêtre. En effet les structures galactiques telle que les bras spiraux ou le centre galactique modifient la densité sur la ligne de visée. La méthode MST ne réagira ainsi pas de la même façon. D'ailleurs nous avons vu qu'il existe une corrélation significative entre la longueur de branche de coupure, seuil automatique pour définir une sur-densité, et la densité de l'échantillon. Après plusieurs tests et vérifications nous avons gardé une taille de fenêtre correspondante à la nature des données.

Nous avons ensuite utilisé les distances héliocentriques afin d'extraire des sur-densités fiables. Au total 496 sur-densités ont été requalifiées en amas. Les autres sur-densités, 1137 au total, sont classées dans un catalogue archive. De plus ces amas sont qualifiés d'amas enfouis en raison des clumps les composants. Leur masse volumique nous indique aussi que ces amas sont liés, en tout cas à ce stade.

Ce catalogue sera bientôt accessible par la communauté scientifique à travers le CDS.

5.1.3 Analyse des amas

Nous avons vu dans ce chapitre que les 496 amas, issus du catalogue fiable, étaient distribués dans les régions denses du plan de notre galaxie, à savoir les bras spiraux. Ils se distribuent d'ailleurs en-dessous et au-dessus du plan galactique en laissant apparaître une distorsion bien connue appelée *warp* de la Galaxie.

La distribution des clumps est aussi intéressante puisque nous observons qu'à cette échelle, c'est à dire de quelques parsecs à quelques dizaines de parsecs, les clumps en amas suivent une distribution fractale avec une dimension moyenne similaire avec les précédentes études du MIS à l'échelle galactique. Cela renforce l'idée de sous-structures et de fragmentation selon les mêmes processus physiques, des grandes échelles galactiques aux plus petites, dans des régions denses des bras spiraux.

J'ai poursuivi mon étude en analysant la distribution des masses des clumps. Nous avons comparé le profil de distribution avec l'IMF, la fonction de masse initiale. L'IMF fait l'objet de nombreux débats, concernant son origine mais aussi son universalité. Les modèles ne s'accordent d'ailleurs pas avec l'ensemble des observations. En analysant la CMF, la fonction de masse des clumps, nous voulions évaluer les différences possibles entre les clumps isolés et les clumps en amas. Les clumps isolés sont supposés se former dans des régions moins denses ce qui implique un changement d'environnement. Nous n'avons observé aucune différence entre les deux distributions. De plus la CMF correspond à l'IMF, que ce soit aux petites masses qu'aux grandes masses. Nous pouvons en tirer deux conclusions. La première est qu'il n'y a pas de différences de profil suivant ces deux environnements. La seule différence s'inscrit dans les masses. Les clumps isolés sont en moyenne moins massifs que les clumps en amas. Nous pourrions ainsi nous attendre à une différence du nombre d'étoiles formés à partir de ces deux types de clumps (un facteur deux), si l'on considère une efficacité de formation stellaire invariante. Ceci ne semble pas être le cas d'après des études récentes sur des amas d'étoiles connus. Cependant nous faisons une étude à l'échelle galactique en moyennant les nombres d'étoiles potentiellement formées. Dans ce cadre, une valeur fixe de la SFE est adaptée. Mais surtout, l'IMF semble prendre racine dans les tous premiers stades de la formation stellaire. Il semble ainsi y avoir un lien fort entre l'IMF, la CMF et la CoMF (core mass function). Tout ceci s'inscrit dans une universalité de l'IMF, dans l'espace et dans le temps puisque que plusieurs millions d'années séparent la distribution des étoiles et des clumps.

L'identification avec des régions de formation d'étoiles massives comme les régions H II et les IRDCs, nous indique le potentiel de cette étude sur cette question. En effet, plus de 50% des amas sont associés avec des IRDCs et/ou des régions H II. C'est dans ce genre de zones que la question de la formation d'étoiles massives trouvera sa réponse. Enfin, l'analyse des amas enfouis de la littérature nous renseigne sur la nouveauté de notre catalogue. Seulement 30 amas de notre étude, sur les 496 au total, sont associés à des amas enfouis connus.

5.2 Perspectives

Une version finale du catalogue de clumps de l'équipe d'Hi-GAL pourrait nous permettre d'obtenir plus d'amas et ainsi plus de statistiques, nous permettant d'étudier l'interaction entre ces amas et le milieu interstellaire à une échelle plus fine.

A la fin de ma thèse, j'ai commencé à étudier des régions particulières de formation stellaire, comme sur les figures 5.1 et 5.2, associées avec une ou plusieurs régions H II. L'étude consiste à utiliser d'autres catalogues de sources dans l'infrarouge moyen (MIPSGAL, GLIMPSE, WISE, 2MASS) ainsi que des traceurs du milieu interstellaire (CO, H I) afin d'évaluer l'âge des sources (clumps, classe II et III) et des régions H II. Un profil des âges et une comparaison avec l'âge de régions H II permettrait d'argumenter envers une formation induite ou non. Nous pourrions aussi utiliser ces données afin d'établir si un seuil d'extinction ou de densité est requis pour la formation stellaire. Enfin, une carte de colonne de densité nous aiderait à extraire des possibles filaments et de comparer leur structure avec les étoiles en formation.

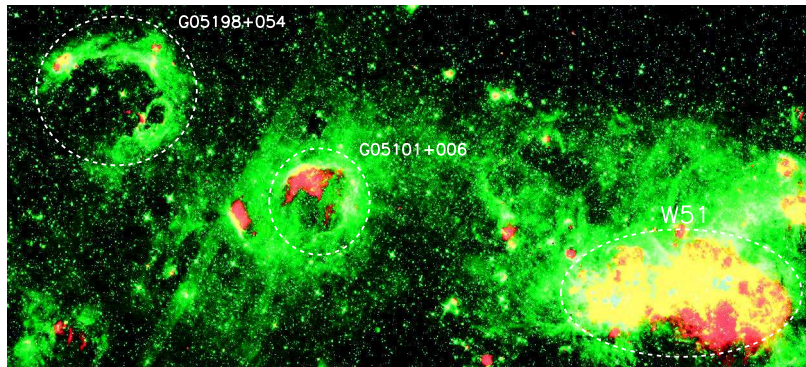


FIGURE 5.1: Régions H II proche de W51.

Comme évoqué dans le chapitre 1 de la première partie, l'arrivée du télescope Herschel a permis la découverte d'un nouveau type d'objet appartenant au milieu interstellaire, les filaments. À partir du catalogue d'amas, il serait intéressant d'en explorer certains ayant déjà des structures filamentaires observées visuellement. En utilisant les images Herschel afin de calculer les cartes de densités ainsi que des traceurs tel que le CO et le HI, nous pourrions évaluer le lien entre ces filaments et les clumps. L'une des pistes de recherche serait d'essayer de contraindre la formation d'étoiles massives en analysant les densités de colonne d'hydrogène ainsi que l'extinction provenant des cartes CO et HI.

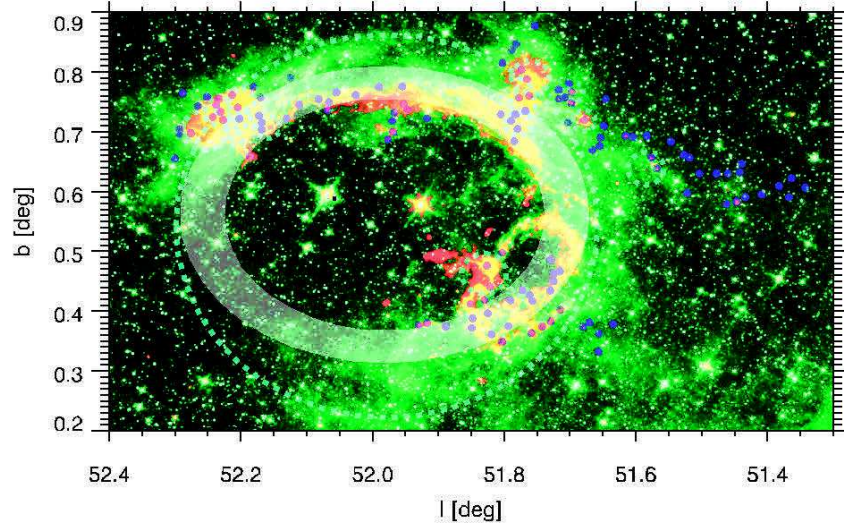


FIGURE 5.2: Amas enfouis de notre étude associés à la région H II G5198+054 situé à 10 kpc. Une structure filamentaire se dessine sur la partie haute-droite.

Il est difficile d'estimer si ces amas perdureront dans le temps et il faudrait utiliser des modèles dynamiques d'amas plongés dans un potentiel galactique pour étudier les effets de la force de marée, des perturbations causées par les objets multiples et tout autre phénomène causant des échanges d'énergie importants. Ceci a déjà été commencé par [Dorval et al. \(2016\)](#) pour l'aspect théorique.

Enfin, dans un futur proche, les observatoires NOEMA (radiotélescope millimétrique) et ALMA (interféromètre millimétrique) permettront d'en apprendre plus sur la formation d'étoiles massives, notamment leur lien avec les filaments d'Herschel, en complétant les données Herschel à plus grandes longueurs d'onde. Il s'agira notamment d'évaluer si la formation stellaire suit les mêmes lois de l'échelle des filaments aux grandes structures galactiques.

Quatrième partie

Annexes

ANNEXE

A

ARTICLE "SOURCES CLUSTERING
IN THE HI-GAL SURVEY USING A
MINIMUM SPANNING TREE
METHOD."

Sources clustering in the Hi-GAL[★] survey using a Minimum Spanning Tree method

M. Beuret¹, N. Billot², L. Cambresy¹, D. J. Eden³, D. Elia⁴, S. Molinari⁴, S. Pezzuto⁴, and E. Schisano⁴

¹ Observatoire astronomique de Strasbourg, Université de Strasbourg, CNRS, UMR 7550, 11 rue de l'Université, F-67000 Strasbourg, France

e-mail: maxime.beuret@astro.unistra.fr

² Instituto de Radio Astronomía Milimétrica, Avenida Divina Pastora, 7, Local 20, 18012 Granada, Spain

³ Astrophysics Research Institute, Liverpool John Moores University, Ic2 Liverpool Science Park, 146 Brownlow Hill Liverpool L3 5RF, UK

⁴ INAF – Istituto di Astrofisica e Planetologia Spaziali, via Fosso del Cavaliere 100, I-00133 Roma, Italy

ABSTRACT

Aims. We aim at studying the clustering properties of the far-infrared sources from the Herschel infrared Galactic Plane Survey (Hi-GAL). The spatial distribution of sources inside star forming regions is an imprint of the original condition in the molecular cloud. We also delivered a catalog of over-densities.

Methods. We used a Minimum Spanning Tree method to find over-densities in the Hi-GAL survey in the inner part of the galaxy. We used a *product* catalog from the Hi-GAL consortium which give us a list of far-infrared clumps in a longitude range $289 \leq l \leq 360$ deg and $0 \leq b \leq 67$ deg and also in a latitude range $-2 \leq b \leq 2$ deg. We built a catalog of over-densities that allow us to explore the clustering properties of the young clumps. Using the heliocentric distance estimates of clumps we built a catalog of more reliable over-densities which was the pillar of the clustering properties study. We then linked these over-densities with environmental conditions as H II regions and IRDCs.

Results. We found 1633 over-densities with more than 10 members. Using the heliocentric distance estimations and compactness criteria, 496 are considered as reliable and define as cluster candidates. We built a catalog of cluster candidates and another one composed by the 1137 remaining. We found a difference of the spatial distribution of the cluster candidates between the first and fourth quadrants that following the inner spiral structure of the Milky Way. We demonstrate that our cluster candidates have a fractal shape according to the ISM shape. We also study the distribution of masses of clustered clumps and isolated clumps. We show that the CMF tend to be same than the Kroupa's IMF with a 3 segment fitting. To complete our study we analyze the correlation between the environmental condition and these cluster candidates. We found that 50% of them have at least one H II region and 60% have at least one IRDC.

Key words. Stars:formation – Stars:mass function – Galaxy:structure – Infrared:Stars – Catalogs

1. Introduction

In recent years, the invention of Galactic Plane surveys have allowed for the statistical properties of the star-formation process to be examined, as they encompass many different Galactic environments. These environments can cover all length scales, from studying interactions on kpc scales in the spiral arms, to the sub-10-pc scales within individual star-forming regions. Studies have started to show that, on the scale of spiral arms, there are no major changes in the star-formation efficiency caused by the spiral arms (Moore et al. 2012; Eden et al. 2013, 2015) but that the sub-10-pc scales may be the most important, with the most significant variations found on the smaller scales (Eden et al. 2012; Vutisalchavakul et al. 2014). These results point towards the smaller scales as the most important, as a result, local triggering may be vital in the star-formation process (e.g. Deharveng et al. 2005; Thompson et al. 2012; Kendrew et al. 2012).

Billot et al. (2011) used the Herschel infrared Galactic Plane Survey (Hi-GAL; Molinari et al. 2010b, 2016) Science Demonstration Phase (SDP) fields, to study the clustering of star-forming clumps ("clump" is preferred to "core" because most of sources are not individual sources), within two, 2x2 deg fields. In this paper, we extend this to cover 1/3 of the Galactic Plane, in the Galactic longitude range of -71 deg to 67 deg, spanning the first and fourth quadrants.

As with the Billot et al. (2011) study, we use the Minimum Spanning Tree (MST) method to investigate how the band-merged (Elia et al., in prep), far-infrared traced (70, 160, 250, 350 and 500 μ m) star-forming clumps cluster with each other. A byproduct of identifying the clumps that fall into these clusters, is that field clumps will also be identified, allowing for a comparison between these two environments. Measurements of the stellar initial mass function (IMF) in both the Galaxy and in extragalactic structures, find that the IMF is invariant, with no significant differences found (Bastian et al. 2010). If the clump mass function (CMF) matches the slope of the IMF, as well as the turnover mass, with a constant mass offset applied, as pre-

[★] Hi-GAL is a key-project of the Herschel Space Observatory survey (Pilbratt et al. 2010) and uses the PACS (Poglitsch et al. 2010) and SPIRE (Griffin et al. 2010) cameras in parallel mode.

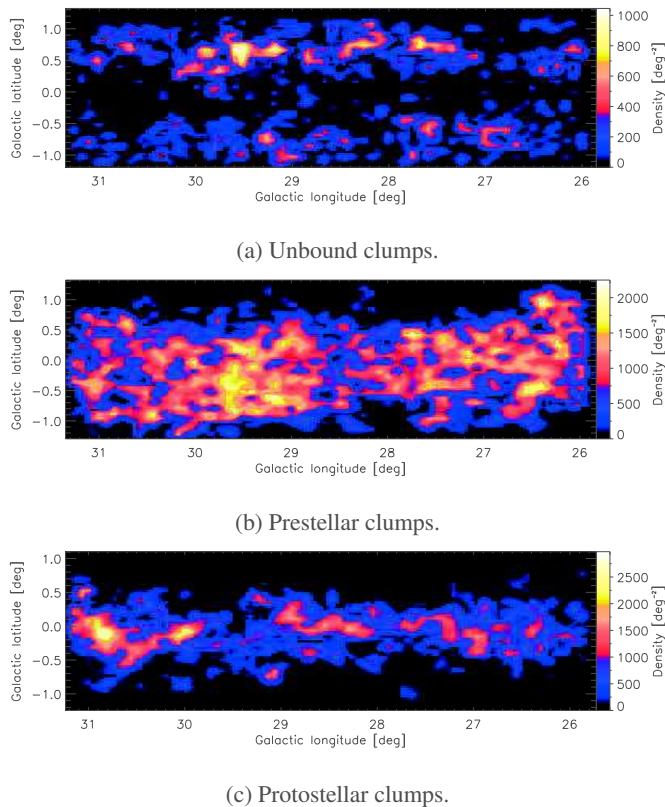


Fig. 1: Density map of a sub-sample of the three different types of clumps defined in the Hi-GAL product catalog, located in a longitude range $26 \leq l \leq 31$ deg. The same distribution is observed for the whole sample.

vious observers have found (e.g. Beltrán et al. 2006; Simpson et al. 2008), then any changes detected in the CMFs of different environments may hint at a change in the IMF.

The band-merged Hi-GAL product catalog (Elia et al., in prep), with a similar description of how this was built found in Elia et al. (2013), provides SED fit parameters to the individual clumps, allowing for the evolutionary stage of the clumps to be identified. Three stages are defined: unbound objects, pre-stellar cores and protostars. Pre- and proto-stellar stages are identified from unbound objects by calculating the Bonnor-Ebert mass, $M_{BE} \approx 2.4 R_{BE}^2 / G$ (Giannini et al. 2012). The pre- and proto-stellar stages are distinguished from each other by the presence of a $70 \mu\text{m}$ source in a proto-stellar clump (e.g. Dunham et al. 2008; Veneziani et al. 2013). The distribution of the three evolutionary stages is shown for part of the Galactic Plane in Fig. 1. Each sub-figure represents an evolutionary stage in the longitude range $26 \leq l \leq 31$ deg. The pre-stellar clumps are more extended in Galactic latitude, compared to proto-stellar clumps. The unbound clump distribution is hard to characterise due to their obscuration by the proto-stellar clumps in the mid-plane, therefore we only consider clustered over-densities composed of pre- and proto-stellar clumps. It is important to notice that this distribution is the same for the whole sample i.e. for the longitude range $289 \leq l \leq 360$ deg and $0 \leq l \leq 67$ deg.

* $a = \sqrt{k_b T / \mu}$ where k_b is the Boltzmann constant, T the clump temperature and μ the mean molecular weight. R_{BE} is the Bonnor-Ebert radius which can be approximated by the radius of the clump. G is the gravitational constant.

Further evolutionary markers, such as Hii regions, a marker of high-mass star formation (Urquhart et al. 2013) and infrared dark clouds (IRDC) can be associated with the clustered clumps, giving further insight into the environmental conditions.

This paper will be arranged as follows. Section 2 will introduce the Minimum Spanning Tree method, allowing the clustered clumps to be identified, with Section 3 describing the process of applying the method to the Hi-GAL data. In Section 4, we present the results of the study, with section 5 comprising of the discussion, whilst section 6 presents the catalogue of over densities. Finally, Section 7 contains the summary and conclusions.

2. Minimum Spanning Tree method

Clumps in these wavelengths are mostly associated with filaments (eg. Molinari et al. 2010a). For this reason we decided to use a Minimum Spanning Tree (MST) method which is well suited to find sources along filamentary structures as explained in Schmeja (2011). This method was first mentioned in Borůvka (1926a,b). An English translation of this paper can be found in Nešetřil & Nešetřilová (2012). Since this date, several algorithms have been described. Graham & Hell (1985) provides a detailed historical evolution of the MST algorithm. For our study we use the Prim's algorithm which have been developed by Prim (1957). The MST method belong to the graph theory. Especially it is a sub-graph of the Delaunay triangulation (Shamos & Hoey 1975; Toussaint 1980). It aims at connecting all points, called vertices, with branches, called also edges, without creating closed loops while minimizing the total length of the branches. The solution of the MST analysis is unique. In our case the vertices correspond to the clumps and the edges, Λ , correspond to the angular distance separating the clumps regarding to the solution found with the MST method. More details about the Prim's algorithm can also be found in Schmeja (2011). This algorithm was originally used in the astrophysical field for galaxy clusters. MST belong to a group of algorithms particularly useful for large scale statistics as explain in March et al. 2012. It is more recently used for star clusters as in Koenig et al. (2008), Gutermuth et al. (2009), Billot et al. (2011) and Saral et al. (2015). Especially this method was also recently used to observe the mass segregation into star clusters as in Allison et al. (2009), Maschberger & Clarke (2011), Parker et al. (2012). To analyse the tree formed by the MST algorithm several methods can be applied. For our study we follow the recommendations of Koenig et al. (2008) and Gutermuth et al. (2009). We perform the histogram of branches distribution. It yields the estimation of a branch cut-off length Λ_{cut} by fitting segments on the cumulative distribution. This procedure provides us with a cut-off criterion determined by the algorithm itself. It means we do not have to insert a hand-made threshold whereas it is the case for the other methods.

Figure 2 shows the distribution of Hi-GAL protostellar and prestellar clumps for a $2 \times 2 \text{ deg}^2$ region of the sky centered on $l=330$ deg. From this sub-sample we can determine the branch cut-off length by using the branches distribution as shown in the Fig. 3. It shows the cumulative distribution (in grey) from the histogram of branches (in black) of this sub-sample. We can fit two segments on the whole cumulative distribution. The first one will fit the small branches whereas the other one will fit the large branches. The intersection of these two segments provide us with the branch cut-off length Λ_{cut} which corresponds to $160''$ for this sub-sample. We can form some sub-trees from

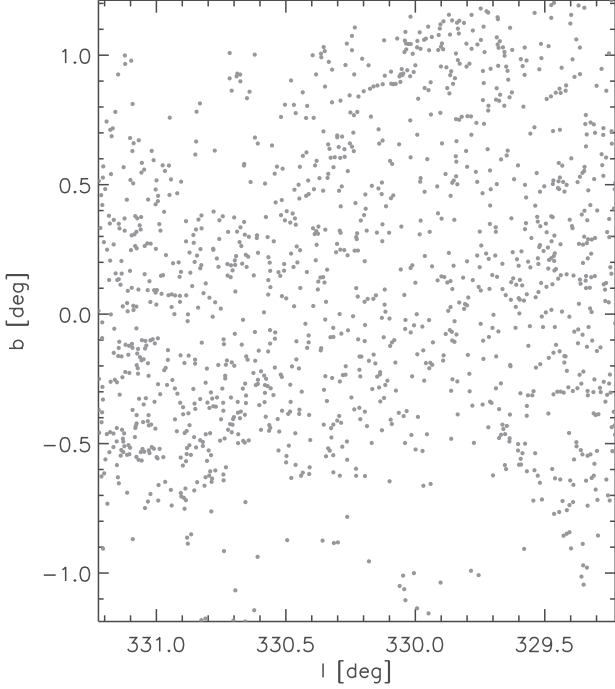


Fig. 2: Distribution of prestellar and protostellar clumps in a 2×2 deg field.

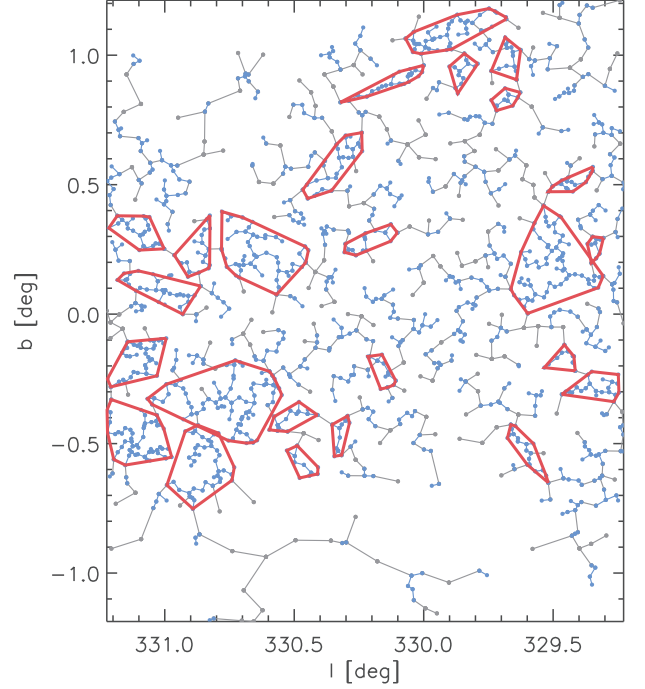


Fig. 4: Result of the MST method applied on a small region (Fig. 2). Grey dots and lines correspond to $\Lambda > \Lambda_{\text{cut}}$ whereas blue dots and lines correspond to $\Lambda < \Lambda_{\text{cut}}$. Red lines surround groups with $N \geq 10$.

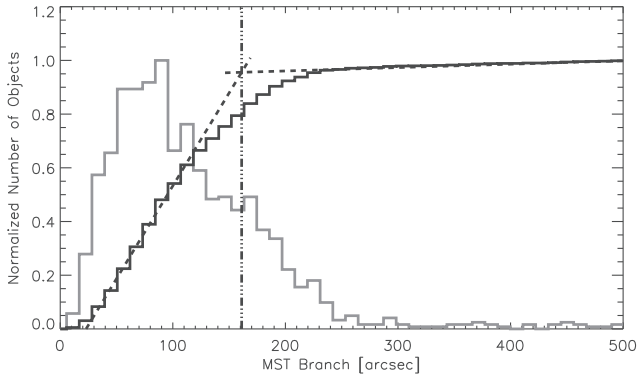


Fig. 3: Histogram of MST branch length. The grey line corresponds to the branch length histogram. The black line is the cumulative distribution. The two dashed black lines are the fitted segments. The dashed-dot line shows the cut-off branch length.

the total tree form by the MST algorithm by considering only branches smaller than Λ_{cut} . The sub-tree are considered as the over-densities formed by N members and $N-1$ branches. The result is presented in Fig. 4. We also use a threshold in term of number of clumps to form an over-density. We follow the recommendation of Gutermuth et al. (2009) by using a threshold of $N = 10$ clumps which allow us to avoid statistical over-densities. We can see that most of the *visual* over-densities in Fig. 2 can be found by the MST method.

3. Catalog of over densities

3.1. Effect of the window size

As shown in Fig. 3, Λ_{cut} depends on the branch distribution resulting from the MST analysis. The determination of Λ_{cut} can be associated with the mean density of the sample. Higher mean densities involve smaller Λ_{cut} . Figure 5 shows the Λ_{cut} distribution along the mean density. Each point corresponds to a solution of MST for different regions probed and different ranges of longitude (see below for more explanation). It seems to have a correlation between Λ_{cut} and the mean density which is well represented by a power law $\Lambda_{\text{cut}} \propto D_{\text{mean}}^{-\beta}$ with a slope $\beta \approx 0.24 \pm 0.01$ (we only consider the error from the fit here). The slope can fluctuate because of the density contrast between the over-densities and the background as well as the number of over-densities which both define the clustering rate of the region probed. It means that the same mean density can imply different values of Λ_{cut} depending of its clustering rate. Remember these points, this correlation allow us to find over-densities in every environments. As explained in Schmeja (2011), MST allow us to find over-densities with a high density contrast compared to the background.

We wanted to split the Hi-GAL data into several windows along the galactic longitude because we observed several different regions of the Milky Way into the galactic plane with different densities. In fact, the density through the galactic plane mainly fluctuate along the longitude due to the center of the galaxy and the spiral arms which increase the number of sources. We do not split the data along the galactic latitude because we considered that the latitude range $-1 \leq b \leq 1$ deg is small enough to not observe strong spatial variation due to the galactic struc-

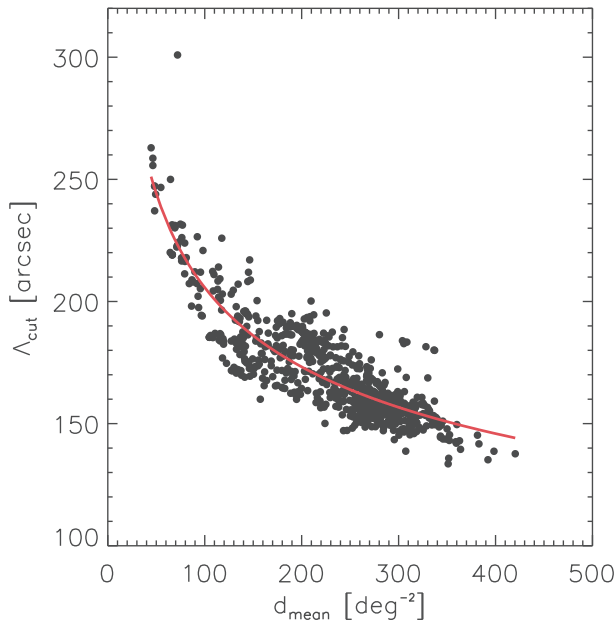


Fig. 5: Relation between Λ_{cut} and the mean density of the window considered. The red line corresponds to the correlation function $\Lambda_{\text{cut}} \propto d_{\text{mean}}^{-0.24}$.

ture.

We also need to overlap all adjacent windows to avoid to reject certain clumps that potentially belong to a same over-density. We created a new set of windows shifted by half the size of the original ones in term of galactic longitude. These windows allow us to keep over-densities potentially split in two in the originally windows. For each window we built a branch histogram as shown in Fig. 3. It means that we have two different Λ_{cut} for a same over-density and sometimes two different shapes because of the overlap. It implies to consider the most representative Λ_{cut} for each over-density which is the one from the window (original one or shifted ones) where the over-density is the closest from the center to avoid to be affected by other regions of the sky, i.e. the window that gather the whole region where is located the over-density.

Figure 6 illustrates the process to re-build the whole catalog. The window on the top of the Fig. corresponds to the overlap window whereas the two windows on the middle are the original ones. The window on the bottom of Fig. 6 is the final catalog for this range of galactic longitude. The grey hulls correspond to the over-densities found for each windows with at least 10 members. The black hulls correspond to over-densities cut in two in the original windows. The boxes correspond to the limits of the area considered to build the final catalog. For this example some over-densities are clearly split in two but we gather them in the final version.

We performed several window sizes to observe the effect of splitting the data set on the result. We also want to select a window size that corresponds to the spatial distribution of the data. We decide to explore a range from 2 to 30 deg by step of 1 window along all data i.e. from -70 to 67 deg in galactic longitude. The minimum size of windows has been chosen to corresponds to the Hi-GAL SDP fields. This size allow to avoid small window with a small number of sources. This will

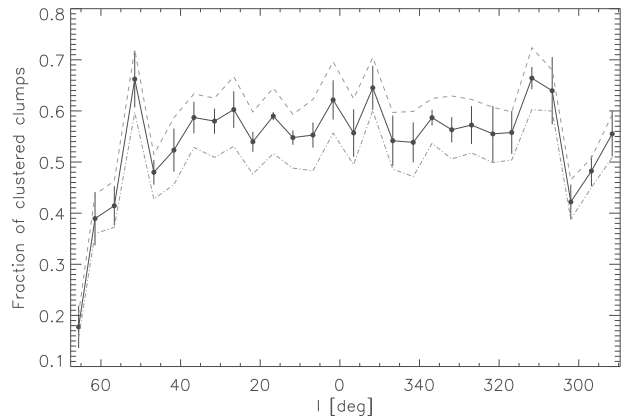


Fig. 7: Spatial distribution of the fraction of clustered clumps along the galactic longitude. The black line corresponds to the mean fraction of clustered clumps of each master catalog (one for each window size) for each bin (5 deg). The black dots correspond to this mean and are located at the average longitude of each bin. We also compute the 1σ error which corresponds to the standard deviation of the fraction of clustered clumps for each bin. The grey dashed and dot-dashed lines correspond to a variation of $\pm 5\%$ of Λ_{cut} .

allow us to have a good branch distribution (as Fig. 3) for each window size. We also want a minimal window size larger than the typical size of over-densities. As over-densities are typically smaller than 0.5 deg, the 2 deg are enough to conserve most of them. The maximum value of window along the sky has been chosen to 30 deg which is supposed to be larger than the different tangents of the spiral arms and correspond to half the angular size of the galactic center.

We can evaluate the possibility to have a non-random clusterization i.e. the over-densities are real and not due to a statistical effect. We use the definition developed by Campana et al. (2008). As explain by the author the total length, Λ , of a MST analysis is proportional to $\sqrt{AN_{\text{tot}}}$ where A is the window area and N_{tot} is the total number of clumps inside this window. They found a coefficient of ≈ 0.65 from their Monte Carlo simulations on random-field MST. Considering all our window sizes we found a coefficient of $\approx 0.51 \pm 0.01$ but the most important part is the relation between the mean length of a MST and $\sqrt{A/N_{\text{tot}}}$. We found a coefficient of $\approx 0.40 \pm 0.01$ that we compare with their result which is the same than for the total length i.e. ≈ 0.65 for a random-field MST. For the same area and the same number of sources, the mean length of a MST with our data is much smaller than the one for a random-field MST. It means that there is a peculiar distribution in our data i.e. it presents a non-random clusterization. This result allows us to compare all window sizes because over-densities should appear in the different windows. Figure 7 shows results on the effect of the window size on the over-densities with more than 10 members. We use all the window sizes (from 2 to 30 deg) i.e. we build 29 master catalogs of over-densities (one for each window size). Each of these master catalogs contain over-densities over the whole sky considered on our study. We compute the fraction of clustered clumps in a 5 deg bin along the longitude. The plotted dots correspond to the average longitude of each bin. Then we compute the mean of the fraction of clustered clumps of each master catalog in each bin. We also compute the standard deviation of the fraction of

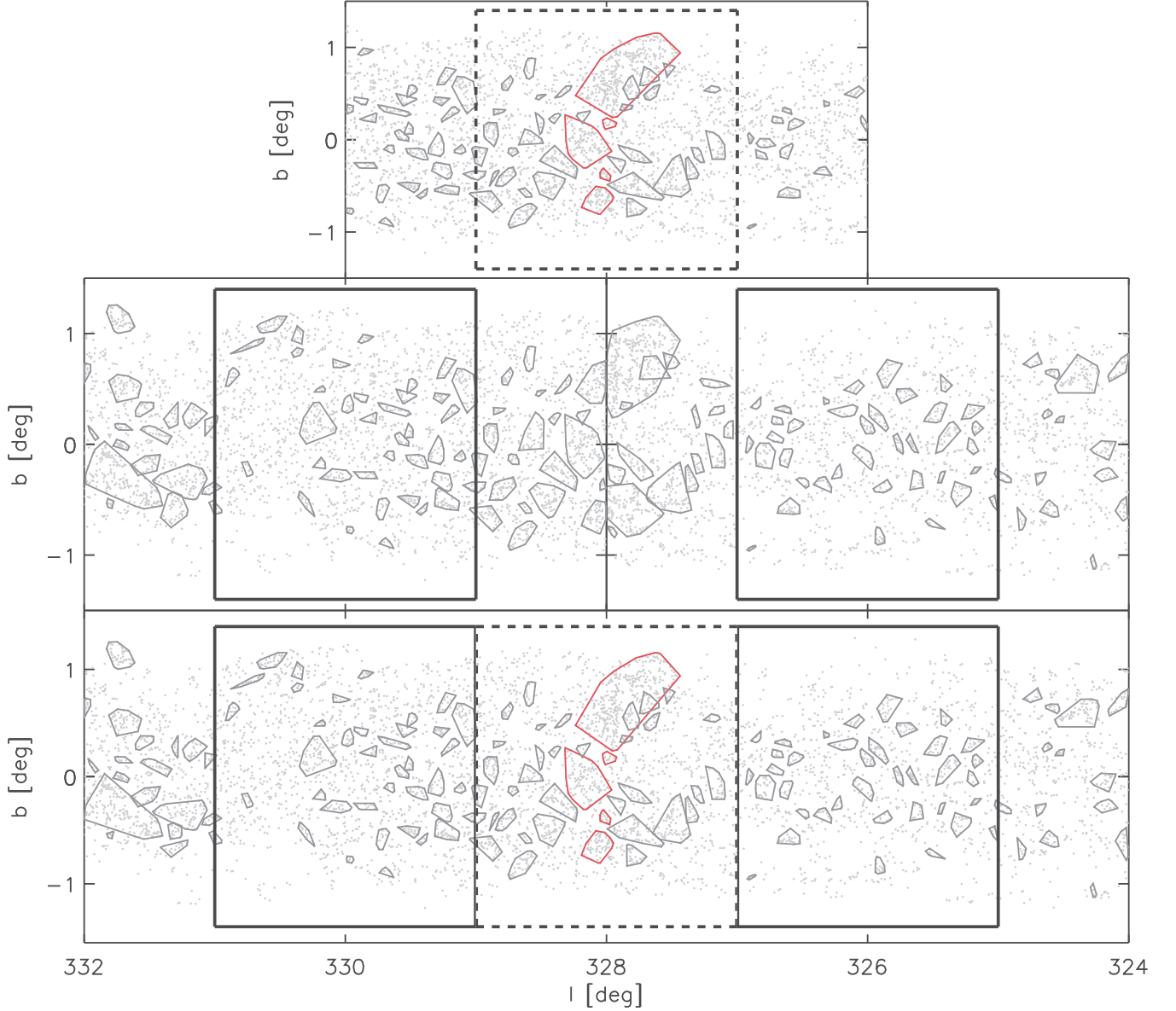


Fig. 6: Example of the overlapping method to re-build the whole catalog of over-densities found by the MST method. This example field is 10 deg large. Grey dots are the prestellar and protostellar clumps. The grey hulls correspond to over-densities with at least 10 members. The red hulls are the over-densities splitted in two in the original windows. The black dashed and solid lines mark the separation limits for each window. The top window is the overlap window. The middle windows are the two original windows of this field. The bottom window is the result of the overlapping method.

clustered clumps in the same bin. We compare this standard deviation to the fraction of clustered clumps for a variation of $\pm 5\%$ of the branch cut-off length (also the average of each master catalog). The difference created by the variation of the window sizes is usually smaller than a variation of $\pm 5\%$ of the branch cut-off length which typically corresponds to $10''$. It means that the variation created by the window sizes is small but significant enough as we can see at $l = 305$ deg.

For this reason we need to complete our study by analysing the variation of the fraction of clustered clumps in particular regions and for every window sizes. We only consider clumps into over-densities with at least 10 members to avoid statistical

over-densities as explain in Sect. 2. Figures 8a and 8b shows to particular regions. Both of them are a 2 deg field located in the fourth quadrant, at 295 deg (Fig. 8a) and 20 deg (Fig. 8b). We can observe that there is a common area where the variation is less important. The jump around 22-23 deg is due to a lower source density in the window. When we take into account a larger window, we gather different regions of the sky as explain above. Sometimes we can gather a region with a small number of clumps. This region change the branch distribution and also the number of clustered clumps. This effect does not appear in Fig. 8b because the distribution of clumps is different compared to the region of Fig. 8a. We also want a window size that remove

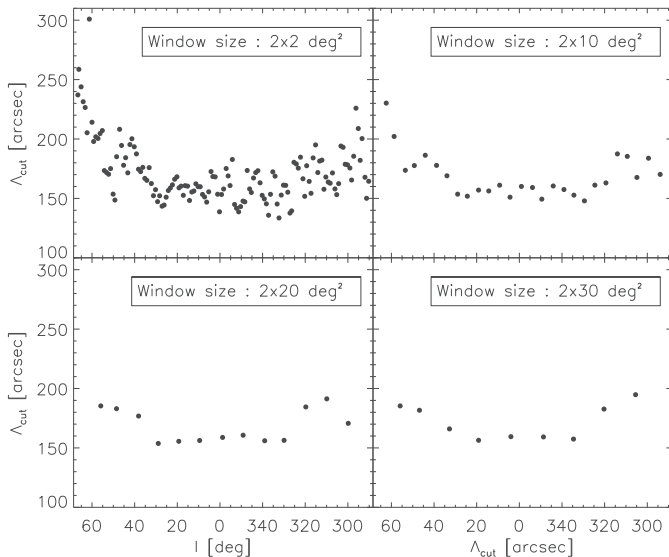


Fig. 9: Profile of Λ_{cut} along the galactic longitude. l corresponds to the center of each window.

the variation at small scale (typically below 2 deg) due to the statistical fluctuations of density over the sky. Those fluctuations are mainly due to the different completeness along the longitude. Figure 9 shows the variation of Λ_{cut} along the galactic longitude for 4 different window sizes, 2, 10, 20 and 30 deg. A larger window size implies a smoother distribution of Λ_{cut} . Beyond 20 deg the distribution is too smooth and we can not observe the variation of densities in the spiral arms or in the center of the galaxy due to the clustering. Only the increasing part of Λ_{cut} remains beyond ± 40 deg. The best range of window sizes seems to be between 10 and 15 deg. Comparing all 2 deg field and for all the reasons mentioned above, we decide to keep a window size of 10 deg.

We found 1705 over-densities with at least 10 members.

3.2. Heliocentric distance estimations

We characterized an over-density as a specific distribution of clumps in 2 dimensions. We would like to improve the characterization of the over-density and build a catalog of cluster candidates using the heliocentric distance estimations (HDEs) from the Hi-GAL product catalog. These HDEs correspond to the kinematic distance of the source determined by using a rotational curve of the galaxy and also some other information to resolve the distance ambiguity (as the extinction map, velocity of CO, NH₃ and HI data, infrared images and optical images). A case study can be found in Russeil et al. (2011) who investigates these kinematic distances in the Hi-GAL science demonstration phase fields.

About 57% of the 99 083 sources (unbound, prestellar and protostellar clumps) of the Hi-GAL product catalog have HDEs. In term of protostellar clumps, almost 64% have HDEs. These HDEs have not been obtained for sources between -10 and +14 deg of galactic longitude due to the small values of radial velocity which are comparable to the peculiar motion of the sources in this range. Excluding this range almost 80 percents of protostellar clumps have HDEs.

A first approach consists in using these HDEs to compute MST in 3D dimensions. However, the uncertainties on HDEs are too large (≈ 0.6 kpc and ≈ 0.9 kpc for the SDP $l=30$ deg and $l=59$ deg;

respectively. See Russeil et al. 2011) to consider this approach as a good one (See Billot et al. 2011). Another approach could be to reprocess the MST analysis for prestellar and protostellar clumps with only HDEs. However, we would remove all clumps without HDEs which means we would have rejected a part of information from the catalog.

For this reason, the HDEs information is using after computing MST, as a third dimension to evaluate the compactness of each over-density. We build a histogram of HDEs for each over-density. We also need to define a range of HDEs for each over-density to determine their reliability. According to the uncertainty mentioned above (From Russeil et al. 2011) we consider a range of 2kpc, called RD for *Reliable Distance*, which corresponds to an uncertainty of ± 1 kpc from the peak of the HDE histogram.

The compactness can be defined as a comparison between the ratio of the number of sources inside the RD over the total number of sources with HDEs and the ratio of the number of sources with HDEs over the total number of sources. We overlay the over-densities on the Hi-GAL images at the different wavelengths for different ratios. We decide to only keep over-densities with at least 70% of clumps with HDEs and 40% of clumps inside the RD. We also consider a reliable over-density as an over-density with at least 5 clumps with HDEs. With less than 5 sources it becomes easier to satisfy the compactness criteria without to be reliable. In this way we found that 607 over-densities satisfy the compactness criteria over the 1705 over-densities found above. 1243 over-densities have at least 1 clump with HDEs.

Only 219 (36%) of these over-densities have all their clumps inside the RD i.e. we do not have any confusion on the line of sight excepted from the clumps without HDEs. We can improve this fraction by removing clumps that do not belong to the RD and only for the over-densities that already satisfying the compactness criteria define above (i.e. the 388 over-densities remaining). We do not remove clumps from the other over-densities because it could completely change the distribution of clumps and the determination of the cut-off branch length by removing a large number of clumps.

We removed 1531 clumps and recompute MST with the 80128 clumps remaining. We found a total of 1633 over-densities. 449 of them have no clumps with HDEs. Over the 1184 over-densities remaining, 536 satisfying the criteria. It is less than the first MST computation but 496 (93%) of these over-densities have all their clumps inside the RD. A majority of these over-densities are the same than in the first MST computation but without the clumps outside the RD. The 40 over-densities remaining have still clustered clumps outside the RD and could change the spatial association of the other clumps, that belong to the same over-density, if we removed them. For this reason we decided to remove these 40 over-densities. The final list is composed by 496 over-densities. 364 of them have between 10 and 20 clumps ($\approx 70\%$) whereas only 4 over-densities have at least 70 members. This final list is composed by a total number of clumps of 9608 i.e. $\approx 12\%$ of the total sample (80128 clumps). 8248 of these clumps have HDEs ($\approx 86\%$ of clustered clumps in the final list).

Recomputing MST with a different sample could also change the evaluation of a non-random clusterization as used in Sect. 3.1. We found a coefficient of $\approx 0.36 \pm 0.01$ between the mean MST length and $\sqrt{A/N_{\text{tot}}}$ which is still under the upper limit of ≈ 0.65 found by Campana et al. (2008) to quantify a non-random clusterization.

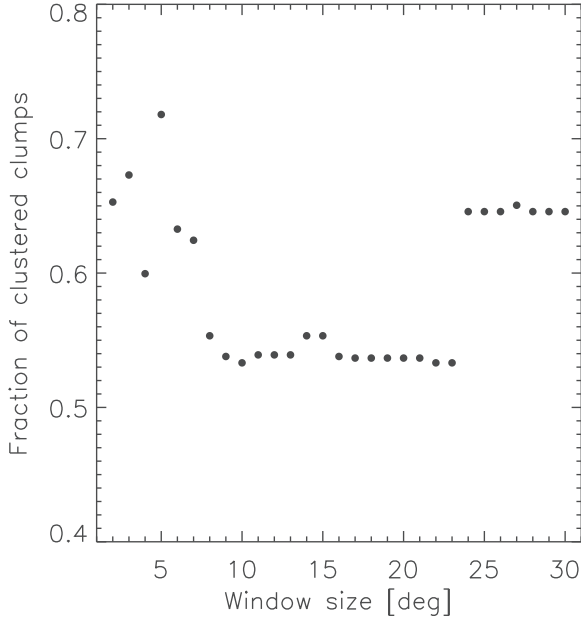
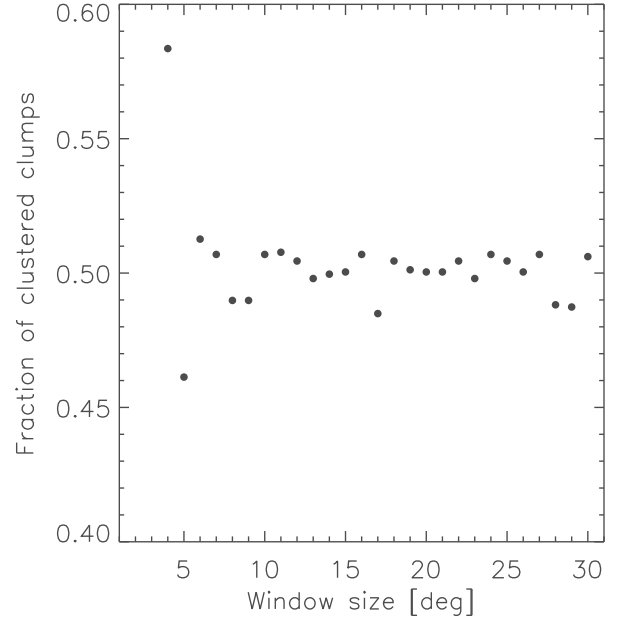

(a) 2×2 deg field located at $l=295$ deg.

(b) 2×2 deg field located at $l=20$ deg.

Fig. 8: Variation of the fraction of clustered clumps into 2×2 deg fields, for all window sizes.

4. Description of the catalogs

As described in Sect. 3.2 we split the over-densities in two lists. These two lists formed two catalogs. The first one is composed by all over-densities rejected for our study with more than 10 members, called *potential clusters catalog*. The other catalog is the one which gather the most reliable over-densities, called cluster candidates catalog, used in Sect. 5 and 6 to analyze the properties of the cluster and it is composed of 496 cluster candidates. A part of these catalogs is available in tables B.1, B.2 and ?? after the reference. The two catalogs are described in the following sub-section.

4.1. Cluster candidates catalog

- Column [1] corresponds to the name of the cluster candidates. The form is GLLL.ll+BB.bb where LLL.ll and BB.bb correspond to the average position of all clustered clumps.
- Columns [2-3] correspond to the galactic longitude and latitude (l,b) of the cluster candidates determined as explain in Col. [1].
- Columns [4-5] correspond to their size in arc-minute. To build this column we fit an ellipse on each over-density. $\approx 75\%$ of cluster candidates are not spherical. This is probably due to the fact that a part of these clumps are associated with filaments (eg. Molinari et al. 2010a). Column [4] corresponds to the semi-major axis whereas Col. [5] is the semi-minor axis.
- Column [6] gives the position angle (PA) of the ellipse in degree. This angle represents the tilt between the positive direction of the longitude axis and the semi-major axis,

rotated clockwise.

- Column [7] corresponds to the heliocentric distance D_{peak} of the cluster candidates estimated from the clumps HDEs. These distances correspond to the HDEs peak values of clumps inside the thickness use in Sect. 3.2.
- Column [8] is the same than Col. [7] but for mean HDEs values \bar{D} .
- Column [9] is the galactocentric distance computed using Eq. 2.
- Column [10] is the height Z from the galactic plane computed using the Eq. 1.
- Column [11-12] provide the cluster candidates size in parsec. We use the angular distances provided by the semi-major and semi-minor axis of columns [4-5] as well as the peak heliocentric distances of the cluster candidates (described in Col. [7] above).
- Column [13] corresponds to the average clumps surface density (in pc^{-2}). We use the area of the ellipse fitted on cluster candidates which is πAB where A is the semi-major axis and B is the semi-minor axis from Col. [11-12].
- Column [14] corresponds to the total number of clumps (both prestellar and protostellar clumps).
- Column [15] corresponds to the ratio of number of prestellar clumps over protostellar clumps. A ratio of 1 corresponds to an over-density composed by only prestellar clumps. Reciprocally a value of 0 means that no prestellar clumps belong to the over-density.

- Column [16] corresponds to the closest H II region from the center of the cluster candidates. This column is explained in Sect. 4.3.

4.2. Potential cluster candidates catalog

- Column [1] corresponds to the name of the cluster candidates. The form is GLLL.ll+BB.bb where LLL.ll and BB.bb correspond to the average position of all clustered clumps.
- Columns [2-3] correspond to the galactic longitude and latitude (l,b) of the cluster candidates determined as explain in Col. [1].
- Columns [4-5] correspond to their size in arc-minute as in Col. [4-5] of the cluster candidates catalog. Column [4] corresponds to the semi-major axis whereas Col. [5] is the semi-minor axis.
- Column [6] gives the position angle (PA) of the ellipse in degree. The reference is the same than for the cluster candidates catalog.
- Column [7] corresponds to the average clumps surface density (in arcmin^{-2}). We use the area of the ellipse fitted on cluster candidates which is πAB where A is the semi-major axis and B is the semi-minor axis from Col. [4-5].
- Column [8] corresponds to the total number of clumps (both prestellar and protostellar clumps).
- Column [9] corresponds to the ratio of number of prestellar clumps over protostellar clumps as in Col. [15] of the cluster candidates catalog.
- Column [10] corresponds to a flag that differentiate over-densities. NDA, for No Distance Available, corresponds to over-densities with clumps without HDEs. SD, for Spread Distances, corresponds to over-densities without a peculiar range of HDEs. SCC corresponds to cluster candidates but with clumps outside the thickness define in Sect. 3.2 that modify the clumps distribution.

5. Results

Section 3 allowed us to split the over-densities in two lists according to the HDEs. The first list corresponds to potential cluster candidates i.e. the 1137 over-densities which have been rejected in Sect. 3.2. These over-densities do not allow us to perform a study of physical properties of star forming regions but it is composed by potential clusters which could be used in other studies as for instance the star formation processes and high mass star formation. The second one is formed by the remaining 496 over-densities. At this stage we prefer to rename these over-densities as cluster candidates. In this section we perform a study of the physical properties and the spatial distribution of these 496 cluster candidates.

5.1. Spatial distribution and clusters characterization

5.1.1. Cluster candidates spatial distribution

Among the 496 cluster candidates, only 176 ($\approx 35\%$) are located in the first quadrant. We split the first and the fourth quadrant

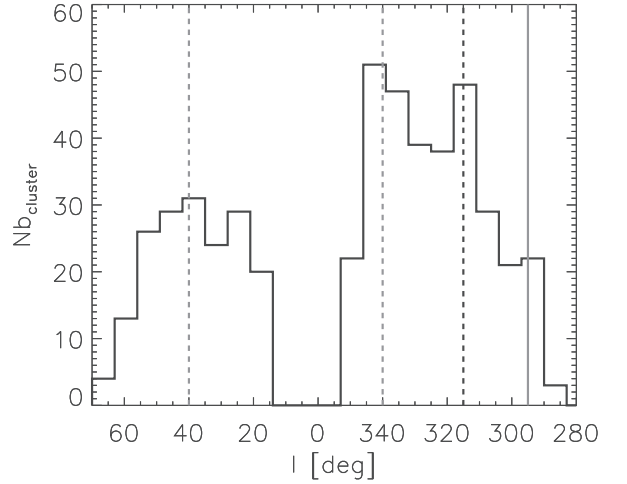


Fig. 10: Longitude distribution of the 496 cluster candidates. The visible tangents of the spiral arms are located at $l=315$ deg (black dashed line). The tangent at $l=295$ deg is marked by a grey line whereas the peak at $l=30$ deg and $l=340$ deg are marked by a grey dashed line.

in two longitude range $[15, 40]$, $[40, 60]$ and $[320, 345]$, $[300, 320]$, respectively. These ranges correspond to a visual separation made from the main peaks in Fig. 10 which shows the longitude distribution of the cluster candidates. In the range $l=[15, 40]$ we have 97 cluster candidates whereas in the range $l=[320, 345]$ we have 151 of them. For the other ranges ($[40, 60]$ and $[300, 320]$) we have 72 cluster candidates for the range in the first quadrant and 101 in the fourth quadrant. There is a difference between the two quadrants which is observed along all longitude range. This difference can be explained by the difference of the spiral structure between the first and fourth quadrant. As shown on Fig. 11, the cluster candidates seems to follow the spiral arms considering the error on the HDEs and the small shift. This figure is obtain using the spiral arms from Englmaier et al. (2011) who study the spiral structure of the Milky Way from the molecular gas. The cluster candidate locations agree with these spiral arms because we observe very young clumps which did not have enough time to be dispersed. Most of the cluster candidates are located beyond 5 kpc from the sun, on the tangential parts of the spiral arms. The star formation seems to be more important in the fourth quadrant than in the first one. The difference in the range $40 \leq l \leq 60$ deg and $300 \leq l \leq 320$ deg can be explained by a low density corresponding to a higher Λ_{cut} value as seen in the 2×2 deg fields of Fig. 9. We can clearly see the tangent at $l=315$ deg on Fig. 10. Other tangents not appear significantly from the background noise as the one at $l=40$ deg because cluster candidates seems to be spread instead of to have a peak around this longitude as shown on Fig. 11. The tangent at $l=295$ deg also appears insignificantly due to the lack of cluster candidates at $l=295$ deg (end of the sample). We also observe a significant peak at $l=340$ deg and a slightly one at $l=30$ deg but those peaks are due to cluster candidates much closer to the Sun than the ones located at other longitudes.

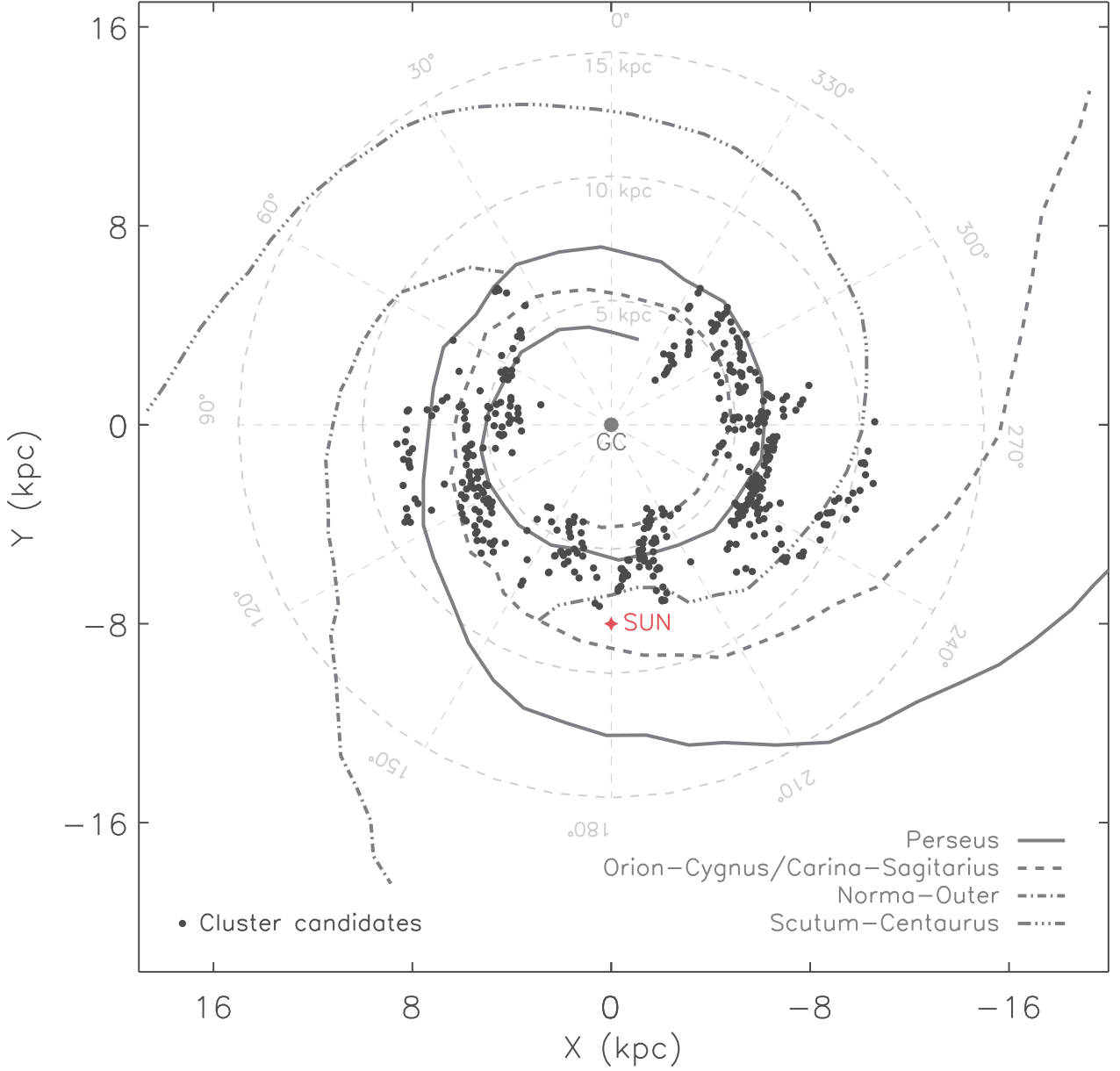


Fig. 11: Distribution of cluster candidates on the spiral arms from their HDEs. The star symbol corresponds to the sun. We consider 4 spiral arms associated with gas from Englmaier et al. (2011). The lines mark the location of these spiral arms. Light grey dashed lines corresponds to galactocentric distances. The dots mark the position of each cluster candidates.

We also analyze the height Z of the cluster candidates from the galactic plane using the expression :

$$Z = D_{\text{peak}} \sin(b) \quad (1)$$

Where D_{peak} is the heliocentric distance of the cluster candidates. These distances correspond to the HDEs peak value of clumps. The galactic coordinate corresponds to the central position of the cluster candidate. We would like to emphasize the link between Z and the galactocentric distance R defines as follow :

$$R = \sqrt{R_0^2 + D_{\text{peak}}^2 - 2R_0D_{\text{peak}} \cos(l)} \quad (2)$$

Where $R_0 \approx 8.5$ kpc is the galactocentric distance of the sun. We define a bin size along R which corresponds to $\approx 0.1R_0$. In each bin we compute the mean of Z , Z_{mean} . Figure 12 shows the profile of the Z_{mean} along R for the first and fourth quadrants. We compare our result with Paladini et al. (2004) who used the distribution of H II regions in the galactic plane but for all quadrants. We observe the same shape for the fourth quadrant in our study than in the third and fourth quadrants of Paladini et al. (2004). Especially we observe the same break around 7 kpc. For cluster candidates in the range $15 \leq l \leq 67$ deg the Z_{mean} decrease until 6 kpc. After 6 kpc Z_{mean} increase as indicated in Paladini et al. (2004) before to decrease again after 8.5 kpc. However, the

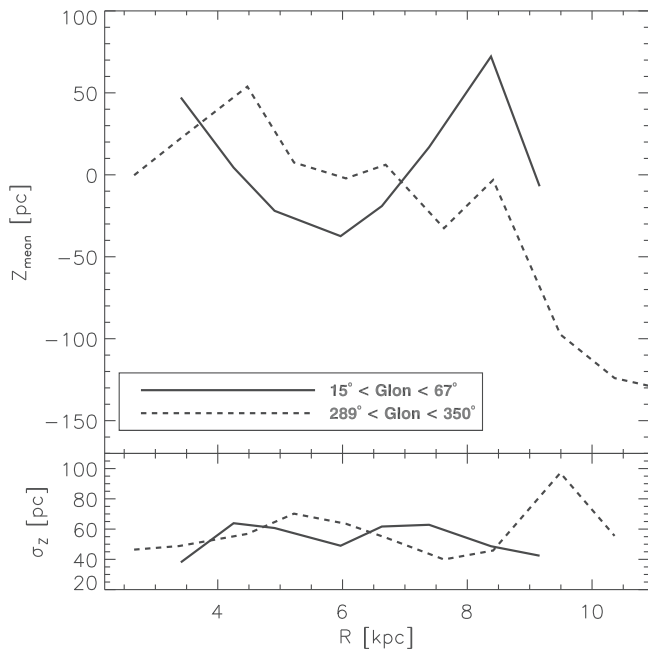


Fig. 12: *Top* : Distribution of the height Z_{mean} versus the galactocentric distance of cluster candidates R . The black line corresponds to the cluster candidates in the first quadrant whereas the grey dashed line corresponds to the fourth quadrant. *Bottom* : Distribution of σ_z along the bins of R .

range of R is small compared to the fourth quadrant and we do not have access to the part after 10 kpc where it supposed to increase quickly as in Paladini et al. (2004). According to this paper and our study the greatest difference between the two quadrants is observed after 7 kpc. This peculiar profile is probably due to the warp of the galactic plane ((See Burton & Hartmann 1988)). This warp is well known for HI disks (See Henderson et al. 1982) but also for other components of the Milky Way as the molecular clouds, OB stars and 2MASS stars. An exhaustive list of the main components of the warped disks can be found in Reyl   et al. (2008). The warp mostly take places after 10 kpc so our results have to be taken with caution. However, it seems that all cluster candidates lie in a thin, asymmetric and warped disk in a range of Z from -248 to 134 pc. The asymmetry in the range of Z is mainly due to the galactic warp, considering a warp counterpart before 10 kpc, but also to the lack of data in the first quadrant beyond 10 kpc. It is also due to the location of the Sun which is located above the galactic plane, about $Z_{\odot} = 10 - 20$ pc according to Reed (1997) and Reed (2006) (See also Brand & Blitz 1993). The sun location can also explain the mean value of Z , $\langle Z \rangle \approx -6.2$ pc.

As explain in Marshall et al. (2006); Reyl   et al. (2008), the stars warp is more ambiguous than the gas warp and it could explain the ambiguity in our result. We also do not observe any flare as described in Paladini et al. (2004), as shown in the bottom part of Fig. 12 which presents the standard deviation σ_z for each bin along R . It agrees with previous works showing that the flares mostly belong to the outer stellar disk Derriere & Robin (See 2001).

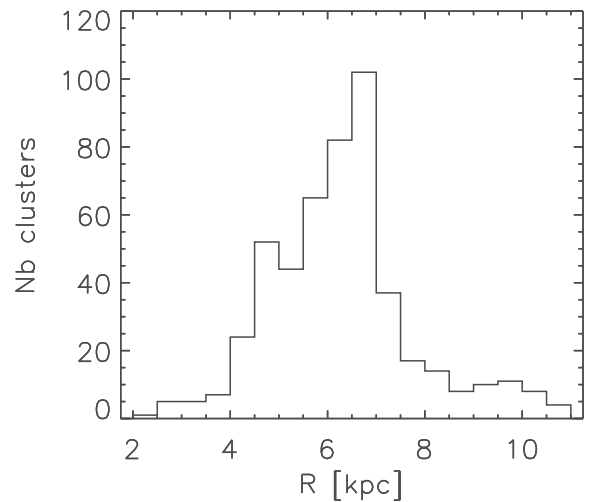
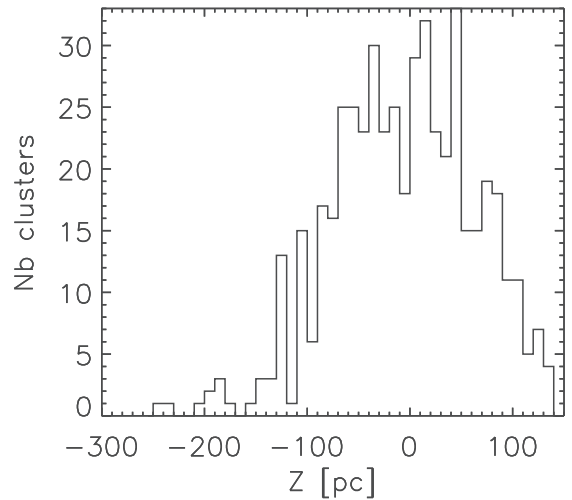


Fig. 13: *Top* : Distribution of cluster candidates along Z . The mean value of Z is $\langle Z \rangle \approx -6.2$ pc. *Bottom* : Distribution of cluster candidates along R . The peak about 6 kpc is also observed in Paladini et al. (2004).

5.1.2. Clumps spatial distribution in clusters

We complete our study by analyzing the clumps spatial distribution. We used the parameter Q introduced by Cartwright & Whitworth (2004) to distinguish centrally concentrated cluster from fractal distribution. This parameter is define as follow :

$$Q = \frac{\bar{m}}{\bar{s}} \quad (3)$$

\bar{m} is the normalized mean branch length of the sub-tree corresponding to the cluster candidates. Mentioned earlier, the sub-tree is a part of the tree found by the MST method for a window size of 10 deg. For other window sizes the sub-tree could be slightly different i.e. the path between clumps could have a different form. Here the difference does not change the result. The mean branch length found for each cluster candidate

is normalized by $(N_{\text{clumps}}A)^{0.5}/(N_{\text{clumps}} - 1)$ where N_{clumps} is the total number of clumps inside each cluster candidate and A is the area of the cluster. \bar{s} corresponds to the mean distance between clumps divided by the radius of the cluster candidates. Cartwright & Whitworth (2004) used different simulated clusters in 2 and 3 dimensions but projected in 2 dimensions. They also applied the Q parameter on real cluster in 2 dimensions. They were able to distinguish a fractal cluster from a centrally concentrated cluster using the Q parameter. Fractal seems to have $Q \leq 0.8$ whereas $Q > 0.8$ corresponds to concentrated clusters. For Cartwright & Whitworth (2004) $Q = 0.8$ corresponds to unclustered clumps i.e. clumps with a random distribution. However, as shown in Cartwright (2009), clusters with $0.75 \leq Q \leq 0.85$ could also follow a concentrated or fractal distribution. Cartwright & Whitworth (2009) shows also that the elongation, the ratio between the semi-major axis and the minor-axis, could impact the value of Q for high value (above 3). Here, only 40 cluster candidates have an elongation above 3 and no correlation between these two values was observed. We decided to not apply any correction to Q . Using this parameter we found a range from $Q = 0.4$ to $Q = 0.84$ with 98.6% of cluster candidates below $Q = 0.8$ and 90% below $Q = 0.75$. It means that our cluster candidates mainly follow a fractal distribution as the ISM (See Combes 2000). The range of Q is similar to the one found by Parker & Dale (2015) in hydrodynamical simulations. Schmeja et al. (2008) found that embedded clusters formed by Class 0/I objects seems to have a hierarchical distribution while older clusters formed with Class II/III objects are more centrally concentrated. This result joins our study because the cluster candidates from our study are formed by very young clumps, typically Class 0/I and younger. However, this parameter do not seems to make the difference between filaments and real fractal clusters. Cluster candidates with $Q = 0.8$ do not mean a random spatial distribution because some of them are composed by filaments.

The fractal distribution suggests that clumps are not gravitationally bound in their host cluster. It can be verified comparing the distance between each clumps (in pc) with the Jeans length (Eq. 4) which characterize the stability for a spherical distribution.

$$\lambda_j = \sqrt{\frac{15k_B T}{4\pi G \mu n m_p}} \quad (4)$$

Where k_B corresponds to the Boltzmann constant, T is the temperature of the cloud, n is the mean volume density of the cloud, m_p the mass of the proton and μ is the mass per particle. We consider a typical temperature of the molecular cloud of 10, 20 and 30 K and a volume density n of 10^5 particles per cm^3 and a mass per particle μ of 4×10^{-27} kg for a cloud composed only by 80% of hydrogen and 20% of helium. This consideration gives a Jeans length from 0.05 pc to 0.1 pc. This length would be different if we consider a cloud with magnetic field and under external pressure. Figure 14 shows the distribution of 2D distances (in pc) between clustered clumps. The smallest value is comparable to the Jeans length but most of them are larger than 1 pc. We also compare a projected distance tot a real 3D distance (Jeans length). Using accurate HDEs lead to increase the distance between clumps. This analysis confirm that cluster candidates are not bound and the spatial distribution could not be centrally concentrated.

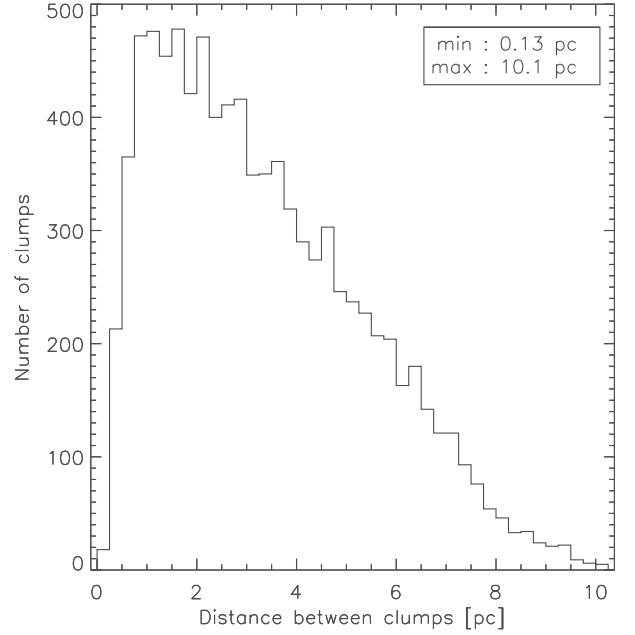


Fig. 14: Distribution of the distance between clustered clumps in cluster candidates.

We also use the criterion of Campana et al. (2008) to determine if our cluster candidates are significant i.e. if their shapes do not correspond to a random distribution. It is directly linked to the criterion used in Sect. 3 to analyze a non-random clusterization. This criterion is described as follow :

$$g_k = \frac{\bar{\Lambda}}{\bar{\Lambda}_k} \quad (5)$$

Where $\bar{\Lambda}$ is the mean branch length of the total tree for window sizes of 10 deg. $\bar{\Lambda}_k$ is the mean branch length of the k -th cluster. We found a range of g_k from 0.95 to 3.02 with a peak about 1.5. We can also use the parameter of Massaro et al. (2009) which used the relation between the number of clumps inside each cluster and g_k . This criterion is described as follow :

$$M_k = N_{\text{clumps}} g_k \quad (6)$$

Where N_{clumps} is the total number of clumps in the k -th cluster. This criteria called *magnitude* allow to compared clusters with a small number of sources and larger ones. We found a range of M_k from 9.5 to 380 with a peak value about 15 which corresponds to $g_k = 1.5$ associated with 10 clumps (the majority of number of clumps inside cluster candidates). 54% have $M_k > 20$ which is a possible limit to determine reliable clusters. However, cluster candidates which have small M_k values (≤ 15) have still a large number of clumps at the same HDEs. The low value of M_k and g_k is the result of low $\bar{\Lambda}$. It happens when the number of clustered clumps (not only cluster candidates) is high and dominates the background in a window. The 2D distribution limits also our analysis because we do not have access to the real distribution inside each cluster candidate which could change the g_k values and probably increase them.

5.2. Distribution of mass

We compare the masses of clustered clumps of the 496 cluster candidates to those which are considered as isolated. As a reminder, in the MST analysis all clumps are connected by branch lengths. If a clump is connected to the other ones with branch lengths greater than the cut-off branch length, this clump is considered as isolated. We count 4752 isolated clumps which represent almost half the total number of clustered clumps (8248). We recompute the mass of all clustered clumps using Eq. 7 and Eq. 8. We use the distance of each cluster candidate as the distance of normalization D .

$$M = \frac{F_\nu D^2}{\pi k_\nu B(T, \nu)} \quad (7)$$

$$M_{\text{peak}} = \frac{D_{\text{peak}}^2}{D^2} M \quad (8)$$

The mass distribution help us to compute the Clumps Mass Function (CMF) slopes for both clustered and isolated clumps. We assumed the relation described in Eq. 9 which is the expression used in (Kroupa et al. 1993) for a multi-segment power law fitting on the IMF. α is linked to the Salpeter slope (Salpeter 1955) by the relation $\alpha = \Gamma + 1$.

$$\Phi = \frac{dn}{dm} \propto m^{-\alpha} \text{ with } \alpha = \begin{cases} \alpha_1, M_0 < M < M_1 \\ \alpha_2, M_1 < M < M_2 \\ \alpha_3, M_2 < M \end{cases} \quad (9)$$

In order to find the CMF slopes, we need to define a bin of mass. In our case the choice of the bin does not affect significantly the result because of the number of clumps used to build the histogram of mass ($N > 500$), as explain in Rosolowsky (2005). We chose a logarithmic bin of mass of 0.1 for all histogram. This value corresponds to an approximation of the value found by using the Freedman and Diaconis' rule (Freedman & Diaconis 2011), $W = 2(\text{IQR})/\sqrt[3]{N}$, where W is the bin width, IQR is the interquartile range and N the number of sources used to make the histogram.

Figure 15 shows the CMF for clustered clumps without HDEs (black dots) and for isolated clumps (grey dots). The error bars correspond to the 1-sigma errors of the Poisson statistic, $\sigma = \sqrt{N}/\Delta M$. We do not use the error on the masses that mainly come from the distance uncertainties. We fit a 3-segment power law on each CMF, as described in Eq. 9, to compare with the 3 Kroupa's IMF slopes (See Kroupa 2001, 2002). We took a constant value for $M_0 = 10 M_\odot$ which determine the starting point of the fit. The fit provides us with the 3 slopes α_1 , α_2 and α_3 as well as the masses of the two breaks M_1 and M_2 . Table 1 shows the fitting results for both clustered and isolated clumps. The different slopes seems to agree between clustered and isolated clumps. These slopes also agree with the Kroupa's IMF slopes. We also compared the CMF shape between the first and fourth quadrants. The CMF shape in the first quadrant seems to agree with Kroupa's IMF slopes. It is the same result in the fourth quadrant but only for the clustered clumps. The 3 slopes of the CMF of the isolated clumps are greater than the IMF slope expectations, especially α_3 which corresponds to high masses. Table 1 shows also the fitting results for the first and fourth quadrants. Figures of the CMF in the first and fourth quadrant are available in appendix A.

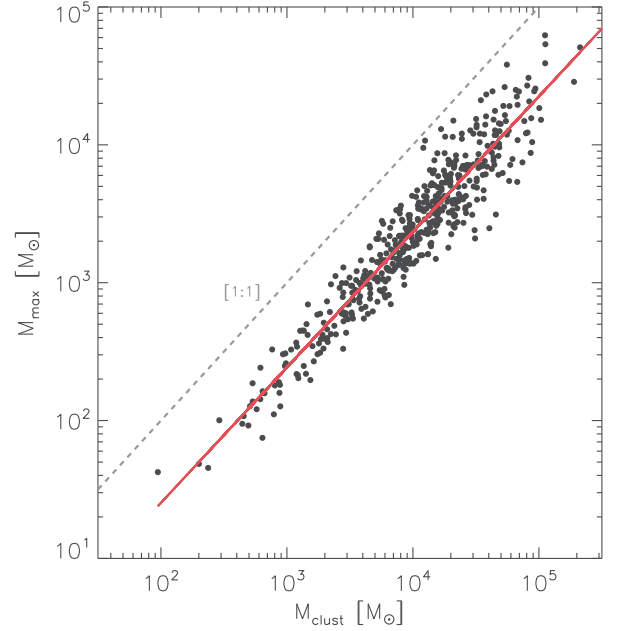


Fig. 16: Biggest mass of each cluster candidates versus cluster masses. All cluster candidates means also to gather all heliocentric distances. The grey dashed line corresponds to the upper limit of the maximum mass which is the cluster mass. The red line corresponds to the linear fit on the distribution.

In all case, excepted for the isolated clumps in the fourth quadrant, the break masses M_1 and M_2 are much smaller for isolated clumps than for clustered clumps.

We also have a correlation between the biggest mass of each cluster candidates and the cluster masses as shown in Fig 16. We used all clustered clumps, even those without HDEs, to compute the total mass of the cluster candidates. The dashed line corresponds to the upper limit of the maximum mass which is the cluster mass. The red line corresponds to the linear fit on the distribution that provides us with the correlation between the maximum mass and the cluster mass, $M_{\text{cloud,max}} = \zeta M_{\text{cluster}}^\eta$ with $\zeta = 0.27 \pm 0.04$ and $\eta = 0.98 \pm 0.02$. This figure gather all cluster candidates, at all heliocentric distances. The correlation remains the same if we split the sample along the heliocentric distance. Kirk & Myers (2012) observed the same correlation for YSO groups. We can reasonably think that the mass of the clustered clumps is linked to the mass of their group from small scale to larger scale.

6. Discussion

6.1. Catalogs completeness

MST is a powerful method to find over-densities in a sample. It is especially appropriate to find over-densities composed by filamentary structures as explained in Sect. 2. The cut-off branch length Λ_{cut} can be determined by different methods. In this paper we used the method described in Koenig et al. (2008) and Gutermuth et al. (2009). Other methods could also be applied to select Λ_{cut} . Maschberger et al. (2010) chose the cut-off branch length such that properties of the over-densities found corresponds to the ones selected by eye. Battinelli (1991) defined Λ_{cut} as the branch length which allow us to find the

Table 1: CMF slopes for isolated clumps and clustered clumps only with HDEs.

Clump types	α_1	α_2	α_3	$M_1 (M_\odot)$	$M_2 (M_\odot)$
Clustered clumps	0.37 ± 0.02	1.43 ± 0.03	2.57 ± 0.07	408.44	2498.98
Isolated clumps	0.44 ± 0.02	1.29 ± 0.05	2.61 ± 0.04	235.78	1050.64
First quadrant					
Clustered clumps	0.15 ± 0.05	1.29 ± 0.06	2.60 ± 0.14	323.61	2081.14
Isolated clumps	0.31 ± 0.06	1.35 ± 0.07	2.63 ± 0.09	215.81	1009.34
Fourth quadrant					
Clustered clumps	0.46 ± 0.02	1.55 ± 0.04	2.56 ± 0.11	483.33	2879.93
Isolated clumps	0.61 ± 0.02	1.99 ± 0.04	4.35 ± 0.27	494.03	3805.42

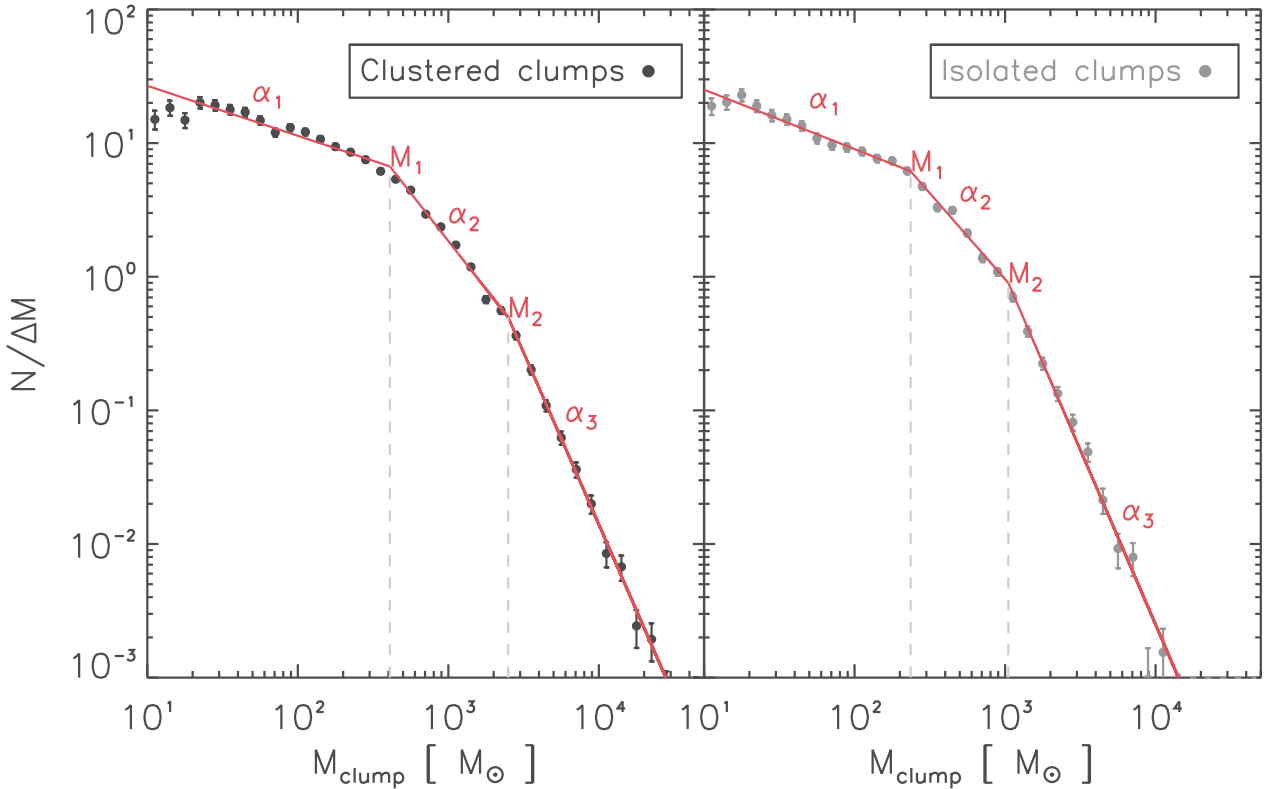


Fig. 15: Mass distribution of clustered clumps (left panel) and isolated clumps (right panel). The mass of clustered clumps are recomputed considering Eq. 8 and the distance of the cluster candidates. We only used clumps with HDEs for this figure. The two plots correspond to the function Φ (Eq. 9) versus mass. The red lines correspond to the fitting segments on the CMF which provide us the slopes α_1 , α_2 and α_3 as well as the break masses M_1 and M_2 .

maximum number of over-densities in a sample.

The low number of cluster candidates compared to the total number of over-densities found after recomputing MST (1633) is due to the HDEs. We removed 464 over-densities with no HDEs. 668 have a large range of HDEs resulting to a strong confusion on the line of sight. The improvement of the accuracy on the estimation of the heliocentric distances is an important topic which could have an impact on the number of cluster candidates found. We would also be able to compute MST in 3D which allow us to have a better spatial characterization of the cluster candidates as well as a possible investigation of the

mass segregation. The heliocentric distances is also a big source of uncertainty in the computation of the mass. It implies a certain uncertainty on the mass distribution and the CMF slope. A greater number of heliocentric distances should also help us to reduce the poisson error which remains the main source of uncertainty in the CMF slope computation.

Figure 17 shows different characteristics of cluster candidates as a function of the mean clump mass. We compare the cluster density (Fig. 17a), the mean length $\langle \Lambda_{\text{clumps}} \rangle$ between clustered clumps (Fig. 17b) which corresponds to the angular length $\bar{\Lambda}_k$ of Eq. 5 but in parsec, the cluster area (Fig. 17c) and

the number of clumps (Fig. 17d) in each cluster candidate to the mean clump mass. The color bar corresponds to the heliocentric distances. There is a strong correlation with the heliocentric distances excepted for the number of clumps. The cluster density (Fig. 17a) clearly shows a relation with the distances. Lower is the density, higher is the mean clump mass and the distance. The same analyze can be applied for mean clump length (Fig. 17b). Farer cluster candidates imply a larger angular distance between clumps but also a larger cluster area (Fig. 17c). Fig. 17b shows also a break about 6 kpc. It probably means that, beyond 6 kpc, the resolution does not change the material gathered in each clumps. Only the length between clumps increase which also means that the sensitivity of the Herschel telescope probably limit our study at high heliocentric distances. The confusion also increase with the heliocentric distance because of the foreground on the line of sight and because we only observe clumps with a high background contrast. All these constraints affect the density of clusters and the mass of clumps which imply the distribution in Fig. 17a, Fig. 17b and Fig. 17c.

6.2. Discussion about the CMF and relation with IMF

The Hi-GAL physical catalog provides also a mass for clumps without HDEs using a distance of 1kpc. We also use Eq. 8 to recompute the mass of these clumps with $D = 1$ kpc and with the help of the distances of the cluster candidates as described in column 7 of Sect. 4.1. The CMF slopes and break masses do not change by using the whole sample of clustered clumps. The results remain also the same if we do not recompute the masses. Comparison of these 3 methods to build the CMF can be found in appendix A.

We do not split the clumps according to their evolutionary stages. The determination of the prestellar and protostellar stages, as explained in the introduction, induced a huge number of prestellar clumps. Because of the possibility to have a misidentifying criterion we prefer to not include this analyze in this paper.

As explain in the previous section, farer clumps seems to gather more materials because of the resolution. It implies that the high mass part of the CMF is dominated by clumps which are located beyond 5 kpc. Reversely, the low and middle mass parts are dominated by clumps located below 5 kpc. However, all these clumps seem to be similar objects in term of evolutionary stages and the CMF slopes remain the same below or beyond 5 kpc.

The results mentioned in Sect. 4.2 about the distributions of mass suggest that the CMF follow the Kroupa's IMF. Most of results agree with the slopes in Kroupa (2001), which are $\alpha_1 = 0.3$, $\alpha_2 = 1.3$ and $\alpha_3 = 2.3$ according to Eq. 9 and the 3σ error bars (Table 1 only shows 1σ error). As explain in Sect.5.2., we neglect errors of the mass measurement ($\geq 20\%$) which are mainly due to the distance estimation. Some other sources of uncertainty could raise the total error on the mass measurement to $\approx 50\%$, as listed in Rosolowsky (2005). According to our result, the relation between CMF and IMF could mean that the future IMF that will take place with later stages could already appear in the early stages of the star formation, independently of the environment. Wienen et al. (2015) found a similar result by analyzing the mass distribution of several embedded clusters, in various environments, with ATLASGAL (See Schuller et al. 2009). Simpson et al. (2008)

investigates on the CMF in the Ophiuchus cloud L1688 and suggests that the IMF is already determined by the prestellar stage. Bastian et al. (2010) suggests that the IMF do not tend to vary significantly with the environment which could confirm the link between CMF and IMF. However, Reid et al. (2010) suggests that the CMF tends to follow the IMF even if perturbations like noise, source blending and spatial filtering modify the mass distribution but it does not mean that the CMF observed is the real CMF. Michel et al. (2011) also suggests that the CMF does not especially tend to follow the IMF in groups of cores, especially in the densest groups.

This issue on the CMF does not explain the steeper slopes for the isolated clumps in the fourth quadrant (see table 1). In this case, it seems that the steepest slope corresponds to high masses. Hatchell & Fuller (2008) shows that prestellar core mass function could be steeper than the IMF if the most massive prestellar cores will fragment to form several less massive protostellar cores. Even if they study cores instead of clumps (more massive and gather more sources than cores), we could extrapolate to our CMF.

Another possibility is that this CMF would be better fitted by 2 segments instead of 3 by removing the effect of the highest masses in the fit (see Fig. A.3). In this case, we found $\alpha_0 = -0.65 \pm 0.02$ and $\alpha_1 = -2.38 \pm 0.03$ for a break mass of $M_0 = 638.3 M_\odot$. The new α_1 slope agrees with the Kroupa's α_3 slope but α_0 is still steeper than excepted for the IMF.

If we still assume this relation between the CMF and the IMF, i.e. trusting the computed slopes and break masses, we can reasonably supposed that it exists a mass coefficient between clumps and stars. This coefficient depends on a star formation efficiency (SFE) ϵ as well as a multiplicity factor which characterize the number of star formed in each clumps n_* . The relation between the IMF masses and CMF masses would be $M_{\text{IMF}} = (\epsilon/n_*)M_{\text{CMF}}$. Assuming a SFE = 0.3 ± 0.1 (Alves et al. 2007) and a probability function for the distribution of mass, it would be possible to find the distribution of n_* . A similar work have been done by Holman et al. (2013) without using a fixed value of the SFE and with the mean value of n_* , \bar{n}_* . However, the degeneracy between the SFE and \bar{n}_* force us to constrain one of those parameters. Hatchell & Fuller (2008) and Goodwin et al. (2008) used a *fully multiple model* which allows to fix \bar{n}_* in order to find the SFE for cores. However, it can not work for clumps because they are much more massives and so, they gather much more sources. There is also a possibility to have substructures inside the clumps which do not allow us to extrapolate model used for the cores. In this case, we need a model for each clumps to find a distribution of n_* which is a really hard task is not the purpose of this paper.

Even if the environment does not change the slopes of the CMF, it seems to affect the break masses and tend to move it to higher values for the clustered clumps. It means that we should observe a difference in the histogram of n_* between clustered and isolated clumps or a variation of the SFE.

A quick computation can provides us a SFE ratio between clustered and isolated clumps, $\epsilon_R = \epsilon_{\text{isolated}}/\epsilon_{\text{clustered}}$, by fixing the multiplicity factor n_* at the same mean value, $\bar{n}_{*,\text{clustered}} = \bar{n}_{*,\text{isolated}}$. Using $M_{\text{IMF}} = (\epsilon/n_*)M_{\text{CMF}}$ we define the relation between ϵ_R and the clustered and isolated masses, $\epsilon_R = M_{\text{CMF,clustered}}/M_{\text{CMF,isolated}}$. The break masses M_1 and M_2 are appropriate to compute ϵ_R . We found $\epsilon_R \in [1.7 : 2.4]$ which means that $\epsilon_{\text{isolated}} \approx 2 \times \epsilon_{\text{clustered}}$. This the reverse result for the isolated clumps in the fourth quadrant where $\epsilon_R \in [0.60 : 0.75]$ i.e. $\epsilon_{\text{clustered}} \approx 1.5 \times \epsilon_{\text{isolated}}$.

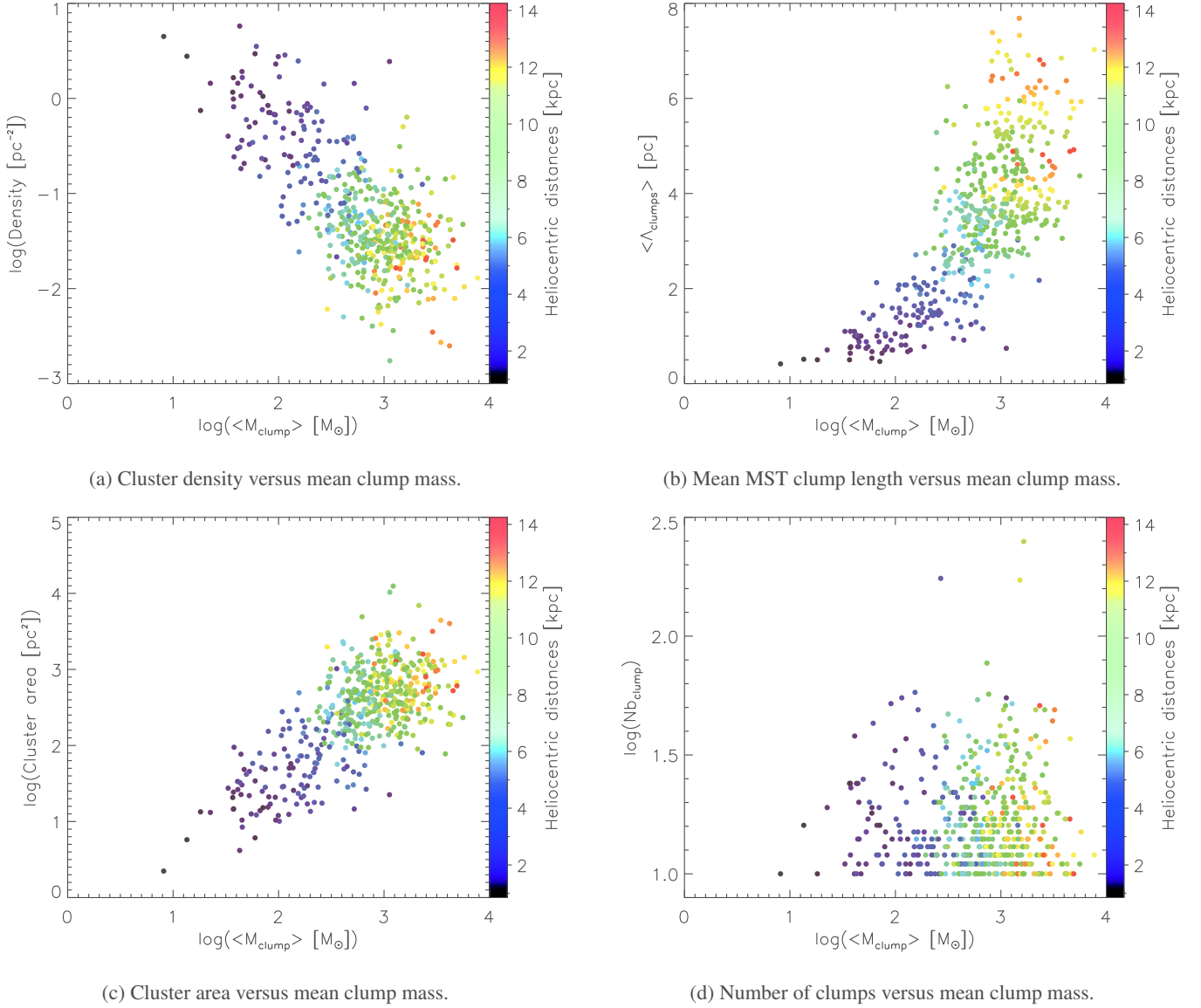


Fig. 17: Cluster characteristics versus the mean clump mass for each cluster candidates. The color bar correspond to the heliocentric distances.

6.3. Comparison with environmental conditions

6.3.1. H II regions

Young star forming regions, as the cluster candidates of our study, are part of clouds embedded in the ISM. Several events occur in these clouds and lead to the formation of stars. The formation of H II regions is one of them. In this section, we would like to investigate on the relation between the cluster candidates and these H II regions. We used the catalog of Anderson et al. (2014) (And2014) which gather 8400 H II regions. We also complete this catalog by using the catalog of Paladini et al. (2003) (Pal2003) which contains 1142 H II regions. 7569 of the total of 9542 H II regions are located in the Hi-GAL product catalog galactic range which correspond to 6561 from And2014 and 1008 from Pal2003. We only associate H II regions and cluster candidates if at least one of the clustered clumps is located inside the limit radius of the H II regions. A total of 274 ($\approx 55\%$) of cluster candidates are associated with at least one H II region. A total of 443 H II regions are

associated with at least one cluster candidate. 380 of them are from And2014 whereas 63 are from Pal2003. These H II regions are firstly associated with cluster candidates independently of their sizes. A probably better association is to differentiate H II regions associated with only few clumps and those associated with the whole cluster candidates. We consider that the H II regions have to be as large as the cluster candidates to be potentially associated with all their clumps. By circularizing cluster candidates we can compare them to the radius of the H II regions. 178 cluster candidates are associated with at least one H II region whereas 177 H II regions are associated with at least one cluster candidates. 168 of the H II regions are from And2014 and 9 from Pal2003.

And2014 provides us also with heliocentric distances. We are able to associate cluster candidates and H II regions using these distances by considering the same error as introduced in Sect. 3.2 which corresponds to ± 1 kpc. Only 58 H II regions from the 380 H II regions firstly associated, from And2014, are still associated with cluster candidates and lead to 54 cluster

candidates associations. By considering only large H II regions, 19 of the 380 below are still associated which lead to 24 cluster candidates associations. However, by analyzing some Herschel images at the different wavelengths, we observe that most of the H II regions without distance estimations are associated with the clumps of the cluster candidates. The small number of association with large H II regions could also be explained because of the size of the cluster candidates. Most of them are larger than 20 pc which implies to have a large and mainly old H II region to host clusters so large. This case remains rare and explains why the association correspond to smaller H II regions which are gathered inside the cluster candidates.

6.3.2. IRDCs

IRDCs have been recently associated with the early phase of the star formation. We investigate their relations with cluster candidates using 2 catalogs of IRDCs. Simon et al. (2006) (Sim2006) used the Midcourse Space Experiment data (Mill et al. 1994; Egan et al. 1998) at $8.3 \mu\text{m}$ to obtain a catalog of 10931 IRDC candidates located in the first and fourth quadrant of the galactic plane. Peretto & Fuller (2009) (Per2009) used the Spitzer surveys GLIMPSE (Benjamin et al. 2003) and MIPS GAL (Carey et al. 2009) catalogs to obtain a list of 11303 IRDCs located almost in the same coordinates range than the cluster candidates. As explain in Per2009, these two catalogs are complementary because only 20 % of the Per2009 IRDCs overlap the IRDC candidates spatial distribution of Sim2006. A total number of 17109 IRDCs (77 %) fall into the longitude and latitude ranges of the cluster candidates.

38 % of clumps fall into the two dimensional ellipse computed using the semi-major axis, semi-minor axis and position provided by these catalogs. Only 10 % are clustered clumps and ≈ 7 % are isolated clumps. Other clumps belong to the potential cluster candidate catalog or to over-densities with less than 10 members. The results are equivalent in the Hi-GAL SDP fields. Billot et al. (2011) found a comparable result for the Hi-GAL SDP fields and for the total number of clumps associated with IRDCs. However, he found that almost 97 % of clumps are located in a cluster. This difference could be explained by the difference of selection of the clusters in the two studies. The Hi-GAL sub-catalog on this two field could also be different. The Heliocentric distances have been probably recomputed and the merge between the different wavelengths is also different from Billot et al. (2011).

However, 333 (67 %) cluster candidates are associated with IRDCs. These IRDCs are not necessarily associated with a clustered clumps but are located inside the two dimensional ellipse formed by using the parameter in the cluster candidates catalog.

In general terms, we should be able to obtain more association between clustered clumps or cluster candidates and IRDCs. The main reason is that IRDCs are only observed in region with a high contrast with the background luminosity. It means that not only dense and cold IRDCs are observed which are supposed to form massive stars.

Table 2: Catalogs of stellar clusters and star forming regions from the literature.^(a)Number of objects in the whole catalog.^(b)Number of objects in the galactic range of the Hi-GAL product catalog.

Reference	Nb _{clust} ^(a)	Nb _{clust} ^(b)
Avedisova (2002)	3235	1519
Sridharan et al. (2002); Beuther et al. (2002)	69	48
Lee et al. (2012)	280	214
Majaess (2013)	230	59
Morales et al. (2013)	695	682
Solin et al. (2014)	160	145

6.4. Comparison with other cluster catalogs

We compared the cluster candidates catalog with other catalogs from the literature. Table 2 lists the reference of the catalogs used, the number of clusters, their types and the number of clusters in the Hi-GAL product catalog range. We have a list of 2667 stellar clusters and star forming regions in the range of the cluster candidates.

As for the H II regions and IRDCs, we have to define an association radius. We decided to keep associated clusters which are located in $3'$ and $9'$ of the central position of the clusters candidates. $3'$ corresponds to the peak size of the distribution of estimated radius, from semi-minor and semi-major axis, of the cluster candidates. 95% of clusters have an estimated radius smaller than $9'$. For $3'$, 102 (≈ 20 %) cluster candidates have an association with at least one of the 2667 clusters. A total of 148 clusters from the literature have been associated with the cluster candidates sample. For $9'$, 242 cluster candidates have an association and 445 clusters from the literature are associated. Compared to the total sample of 2667 clusters, only few of them are associated with the cluster candidates. One reason is because some regions are much older than our sample as some star forming regions and open clusters. These regions are more spread and it becomes more difficult to associate them with young clusters.

7. Summary and perspectives

We have analyzed the sources clustering of the Hi-GAL physical catalog. We used the MST (Minimum Spanning Tree) method to find over-densities in the range allowed by the catalog i.e. $0 \leq l \leq 67\text{deg}$ and $290 \leq l \leq 360\text{deg}$. We found a total of 1705 over-densities with at least 10 members. We used the clump HDEs to differentiate cluster candidates from potential cluster candidates. After recomputing MST to improve our final list we found 1633 over-densities. 496 of them have all clump HDEs inside a thickness $\pm 1\text{kpc}$ and are considered as cluster candidates. The 1137 other over-densities are considered as potential cluster candidates.

We present the largest catalog of embedded clusters of nascent stars. This could help identifying star forming regions for next studies about star formation processes, associated with higher resolution infrared data and spanning a wide range of wavelength (from near-infrared to far-infrared).

We complete our study by analyzing the spatial distribution of the cluster candidates. We found that almost all cluster candidates follow the same shape than spiral arms and are

located in front of the spiral arms from gas of Englmaier et al. (2011) as more evolved stars. We also analyze the Z distribution and we probably observe a part of the well known warp of the galactic plane.

The spatial distribution of clumps shows that the cluster candidates follow a fractal distribution as the ISM. A small part of these cluster candidates seems to have clumps randomly distributed. By visually inspecting these cluster candidates, we see that almost all of them correspond to a cluster composed by several filamentary structures.

Then we analyze the mass distribution which leads to the CMF. The CMF slopes are significantly the same for isolated and clustered clumps which tend toward a universal mass distribution. These values are also consistent with the Kroupa's slopes which means that the IMF already takes place in the early stages of the star formation.

Finally we compare with different environmental components as H II regions, IRDCs and other cluster catalogs from the literature. We found that 55% of cluster candidates are associated with H II regions and 68% with IRDCs. A small number are associated with clusters from the literature because these clusters are usually older than those from our study.

A further work will be to explore the outer part of the Milky Way as well as investigate on the star forming regions containing these young clumps.

Acknowledgements. M. Beuret is grateful the Hi-GAL team as well as Nicolas Billot which provided us with the Hi-GAL product catalog. He also thanks Laurent Cambresy and David J. Eden for the discussion on the results as well as Arnaud Siebert for provided me the file with the shape of the spiral arms. This research has made use of the VizieR catalogue access tool, CDS, Strasbourg, France. The original description of the VizieR service was published in A&AS 143, 23. This research has also made use of NASA's Astrophysics Data System Bibliographic Services.

References

- Allison, R. J., Goodwin, S. P., Parker, R. J., et al. 2009, *MNRAS*, 395, 1449
- Alves, J., Lombardi, M., & Lada, C. J. 2007, *A&A*, 462, L17
- Anderson, L. D., Bania, T. M., Balser, D. S., et al. 2014, *ApJS*, 212, 1
- Avedisova, V. S. 2002, *Astronomy Reports*, 46, 193
- Bastian, N., Covey, K. R., & Meyer, M. R. 2010, *ARA&A*, 48, 339
- Battinelli, P. 1991, *A&A*, 244, 69
- Beltrán, M. T., Brand, J., Cesaroni, R., et al. 2006, *A&A*, 447, 221
- Benjamin, R. A., Churchwell, E., Babler, B. L., et al. 2003, *PASP*, 115, 953
- Beuther, H., Schilke, P., Menten, K. M., et al. 2002, *ApJ*, 566, 945
- Billot, N., Schisano, E., Pestalozzi, M., et al. 2011, *ApJ*, 735, 28
- Borůvka, O. 1926a, *Práce morav. přírodověd. spol. v Brně*, 3, 37, 58
- Borůvka, O. 1926b, *Elektronický obzor*, 15, 153, 154
- Brand, J. & Blitz, L. 1993, *A&A*, 275, 67
- Burton, W. B. & Hartmann, D. 1988, *Structure and dynamics of the galactic system. Triennial report 1984 - 1987*, IAU Commission 33.
- Campana, R., Massaro, E., Gasparini, D., Cutini, S., & Tramacere, A. 2008, *MNRAS*, 383, 1166
- Carey, S. J., Noriega-Crespo, A., Mizuno, D. R., et al. 2009, *PASP*, 121, 76
- Cartwright, A. 2009, *MNRAS*, 400, 1427
- Cartwright, A. & Whitworth, A. P. 2004, *MNRAS*, 348, 589
- Cartwright, A. & Whitworth, A. P. 2009, *MNRAS*, 392, 341
- Combes, F. 2000, *Advanced Series in Astrophysics and Cosmology*, 10, 143
- Deharveng, L., Zavagno, A., & Caplan, J. 2005, *A&A*, 433, 565
- Derriere, S. & Robin, A. C. 2001, in *Astronomical Society of the Pacific Conference Series*, Vol. 232, *The New Era of Wide Field Astronomy*, ed. R. Clowes, A. Adamson, & G. Bromage, 229
- Dunham, M. M., Crapsi, A., Evans, II, N. J., et al. 2008, *ApJS*, 179, 249
- Eden, D. J., Moore, T. J. T., Morgan, L. K., Thompson, M. A., & Urquhart, J. S. 2013, *MNRAS*, 431, 1587
- Eden, D. J., Moore, T. J. T., Plume, R., & Morgan, L. K. 2012, *MNRAS*, 422, 3178
- Eden, D. J., Moore, T. J. T., Urquhart, J. S., et al. 2015, *MNRAS*, 452, 289
- Egan, M. P., Shipman, R. F., Price, S. D., et al. 1998, *ApJ*, 494, L199
- Elia, D., Molinari, S., Fukui, Y., et al. 2013, *ApJ*, 772, 45
- Englmaier, P., Pohl, M., & Bissantz, N. 2011, *Memorie della Societa Astronomica Italiana Supplementi*, 18, 199
- Freedman, D. & Diaconis, P. 2011, *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 57, 453
- Giannini, T., Elia, D., Lorenzetti, D., et al. 2012, *A&A*, 539, A156
- Goodwin, S. P., Nutter, D., Kroupa, P., Ward-Thompson, D., & Whitworth, A. P. 2008, *A&A*, 477, 823
- Graham, R. L. & Hell, P. 1985, *Annals of the History of Computing*, 7, 43
- Griffin, M. J., Abergel, A., Abreu, A., et al. 2010, *A&A*, 518, L3
- Gutermuth, R. A., Megeath, S. T., Myers, P. C., et al. 2009, *ApJS*, 184, 18
- Hatchell, J. & Fuller, G. A. 2008, *A&A*, 482, 855
- Henderson, A. P., Jackson, P. D., & Kerr, F. J. 1982, *ApJ*, 263, 116
- Holman, K., Walch, S. K., Goodwin, S. P., & Whitworth, A. P. 2013, *MNRAS*, 432, 3534
- Kendrew, S., Simpson, R., Bressert, E., et al. 2012, *ApJ*, 755, 71
- Kirk, H. & Myers, P. C. 2012, *ApJ*, 745, 131
- Koenig, X. P., Allen, L. E., Gutermuth, R. A., et al. 2008, *ApJ*, 688, 1142
- Kroupa, P. 2001, *Mon. Not. Roy. Astron. Soc.*, 322, 231
- Kroupa, P. 2002, *ASP Conf. Ser.*, 285, 86
- Kroupa, P., Tout, C. A., & Gilmore, G. 1993, *MNRAS*, 262, 545
- Lee, E. J., Murray, N., & Rahman, M. 2012, *ApJ*, 752, 146
- Majaess, D. 2013, *Ap&SS*, 344, 175
- Marshall, D. J., Robin, A. C., Reylé, C., Schultheis, M., & Picaud, S. 2006, *A&A*, 453, 635
- Maschberger, T. & Clarke, C. J. 2011, *MNRAS*, 416, 541
- Maschberger, T., Clarke, C. J., Bonnell, I. A., & Kroupa, P. 2010, *MNRAS*, 404, 1061
- Massaro, E., Tinebra, F., Campana, R., & Tosti, G. 2009, *ArXiv e-prints*
- Michel, M., Kirk, H., & Myers, P. C. 2011, *ApJ*, 735, 51
- Mill, J. D., O'Neil, R. R., Price, S., et al. 1994, *Journal of Spacecraft and Rockets*, 31, 900
- Molinari, S., Schisano, E., Elia, D., et al. 2016, *ArXiv e-prints*
- Molinari, S., Swinyard, B., Bally, J., et al. 2010a, *A&A*, 518, L100
- Molinari, S., Swinyard, B., Bally, J., et al. 2010b, *PASP*, 122, 314
- Moore, T. J. T., Urquhart, J. S., Morgan, L. K., & Thompson, M. A. 2012, *MNRAS*, 426, 701
- Morales, E. F. E., Wyrowski, F., Schuller, F., & Menten, K. M. 2013, *A&A*, 560, A76
- Nešetřil, J. & Nešetřilová, H. 2012, *Documenta Mathematica*, Vol Extra Vol: Optimization Stories, 127
- Paladini, R., Burigana, C., Davies, R. D., et al. 2003, *A&A*, 397, 213
- Paladini, R., Davies, R. D., & De Zotti, G. 2004, *MNRAS*, 347, 237
- Parker, R. J. & Dale, J. E. 2015, *MNRAS*, 451, 3664
- Parker, R. J., Maschberger, T., & Alves de Oliveira, C. 2012, *MNRAS*, 426, 3079
- Peretto, N. & Fuller, G. A. 2009, *A&A*, 505, 405
- Pilbratt, G. L., Riedinger, J. R., Passvogel, T., et al. 2010, *A&A*, 518, L1
- Poglitsch, A., Waelkens, C., Geis, N., et al. 2010, *A&A*, 518, L2
- Reed, B. C. 1997, *PASP*, 109, 1145
- Reed, B. C. 2006, *JRASC*, 100, 146
- Reid, M. A., Wadsley, J., Petitclerc, N., & Sills, A. 2010, *ApJ*, 719, 561
- Reylé, C., Marshall, D. J., Robin, A. C., & Schultheis, M. 2008, in *Astronomical Society of the Pacific Conference Series*, Vol. 396, *Formation and Evolution of Galaxy Disks*, ed. J. G. Funes & E. M. Corsini, 215
- Rosolowsky, E. 2005, *PASP*, 117, 1403
- Rusell, D., Pestalozzi, M., Mottram, J. C., et al. 2011, *A&A*, 526, A151
- Salpeter, E. E. 1955, *ApJ*, 121, 161
- Saral, G., Hora, J. L., Willis, S. E., et al. 2015, *ApJ*, 813, 25
- Schmeja, S. 2011, *Astronomische Nachrichten*, 332, 172
- Schmeja, S., Kumar, M. S. N., & Ferreira, B. 2008, *MNRAS*, 389, 1209
- Schuller, F., Menten, K. M., Contreras, Y., et al. 2009, *A&A*, 504, 415
- Simon, R., Jackson, J. M., Rathborne, J. M., & Chambers, E. T. 2006, *ApJ*, 639, 227
- Simpson, R. J., Nutter, D., & Ward-Thompson, D. 2008, *MNRAS*, 391, 205
- Solin, O., Haikala, L., & Ukkonen, E. 2014, *A&A*, 562, A115
- Sridharan, T. K., Beuther, H., Schilke, P., Menten, K. M., & Wyrowski, F. 2002, *ApJ*, 566, 931
- Thompson, M. A., Urquhart, J. S., Moore, T. J. T., & Morgan, L. K. 2012, *MNRAS*, 421, 408
- Urquhart, J. S., Thompson, M. A., Moore, T. J. T., et al. 2013, *MNRAS*, 435, 400
- Veneziani, M., Elia, D., Noriega-Crespo, A., et al. 2013, *A&A*, 549, A130
- Vutisalchavakul, N., Evans, II, N. J., & Battersby, C. 2014, *ApJ*, 797, 77
- Wienen, M., Wyrowski, F., Menten, K. M., et al. 2015, *A&A*, 579, A91

Appendix A: Distribution of mass

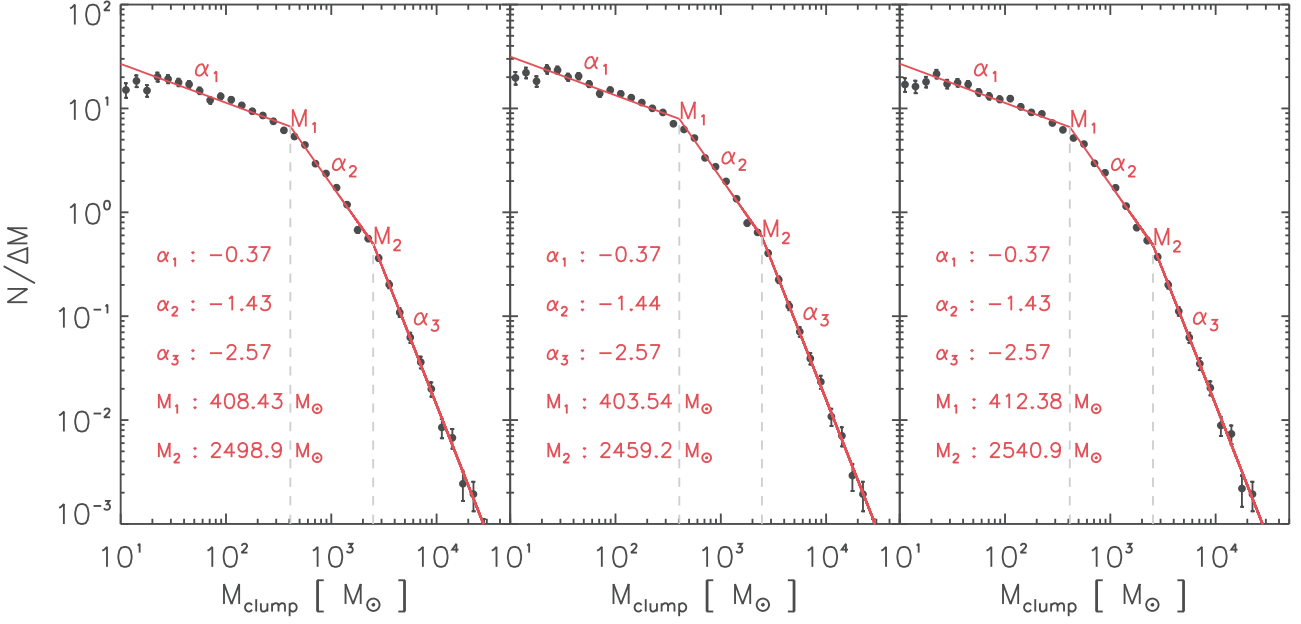


Fig. A.1: Mass distribution of clustered clumps with HDEs (left panel), with and without HDEs (middle panel) and without recomputing masses (right panel). For masses which are recomputed we consider Eq. 8 and the distance of the cluster candidates. The two plots correspond to the function Φ (Eq. 9) versus mass. The red lines correspond to the fitting segments on the CMF which provide us the slopes α_1 , α_2 and α_3 as well as the break masses M_1 and M_2 . We notice that the 3 considerations do not change significantly the result of slopes and break masses.

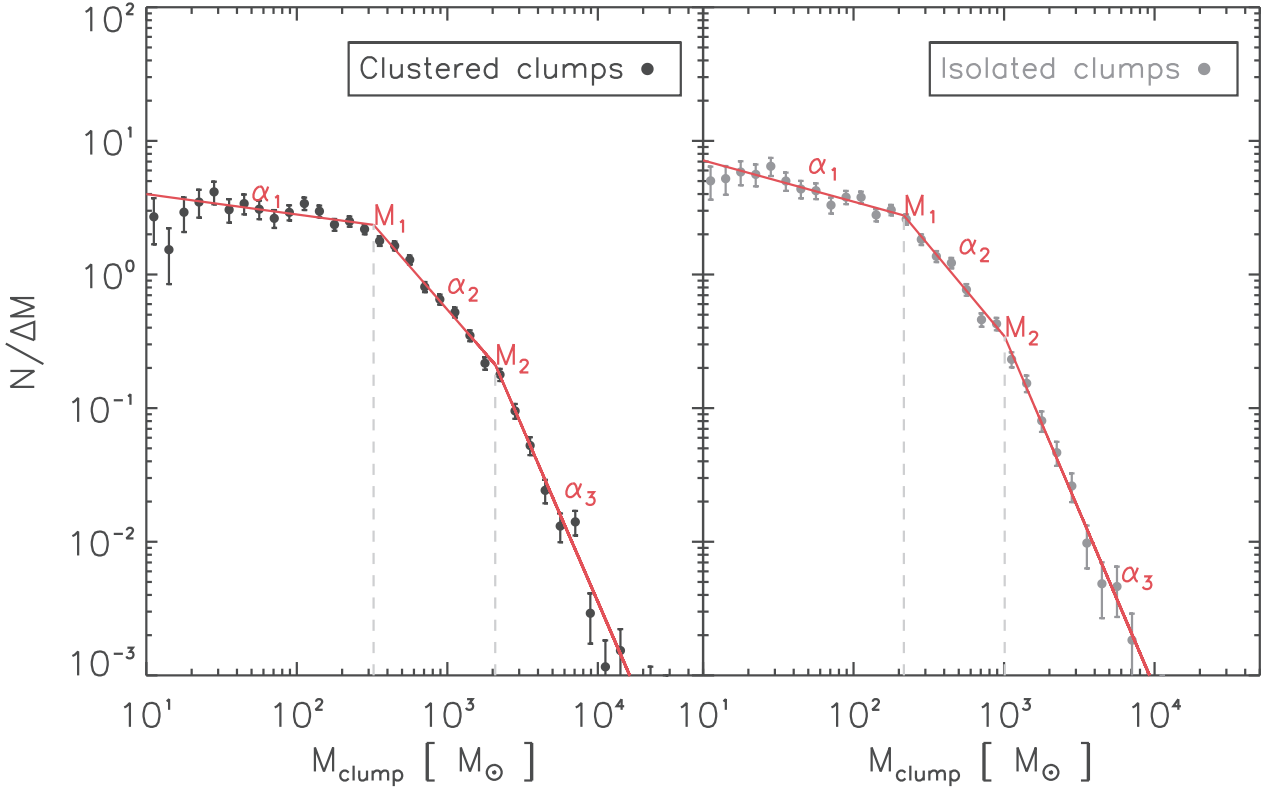


Fig. A.2: Mass distribution of clustered clumps (left panel) and isolated clumps (right panel) in the first quadrant. The mass of clustered clumps are recomputed considering Eq. 8 and the distance of the cluster candidates. We only used clumps with HDEs for this figure. The two plots correspond to the function Φ (Eq. 9) versus mass. The red lines correspond to the fitting segments on the CMF which provide us the slopes α_1 , α_2 and α_3 as well as the break masses M_1 and M_2 .

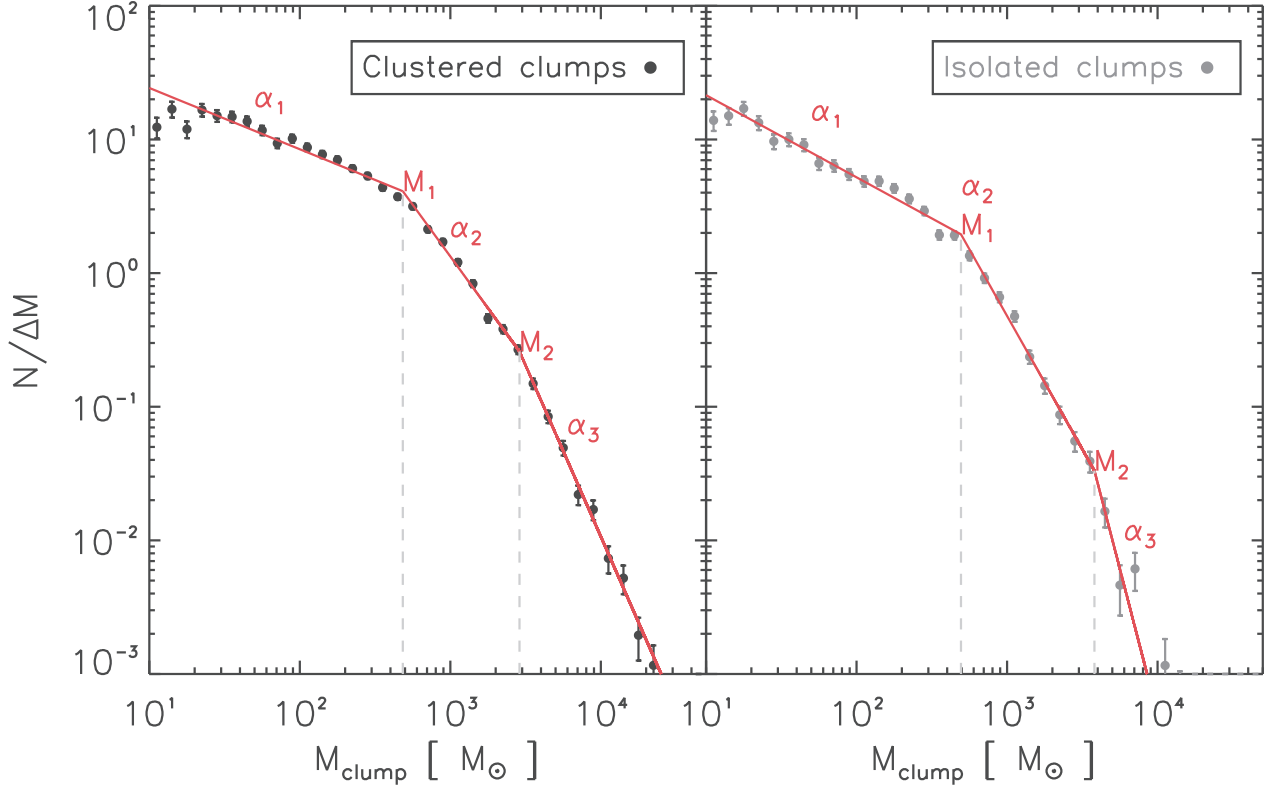


Fig. A.3: Mass distribution of clustered clumps (left panel) and isolated clumps (right panel) in the fourth quadrant. The mass of clustered clumps are recomputed considering Eq. 8 and the distance of the cluster candidates. We only used clumps with HDEs for this figure. The two plots correspond to the function Φ (Eq. 9) versus mass. The red lines correspond to the fitting segments on the CMF which provide us the slopes α_1 , α_2 and α_3 as well as the break masses M_1 and M_2 .

Appendix B: Catalog tables

Table B.1: Partial list of cluster candidates. Column [1-8].

[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Cluster candidates name	l_{center} (deg)	b_{center} (deg)	a-axis (arcmin)	b-axis (arcmin)	PA (deg)	D_{peak} (kpc)	\bar{D} (kpc)
G289.85-0.764	289.87	-0.76	7.36	2.85	16.4	8.05	8.05
G289.57-0.656	289.58	-0.65	3.56	2.28	43.5	7.45	7.45
G290.52-0.771	290.59	-0.78	9.14	7.54	162.	7.55	7.55
G291.84-0.684	291.84	-0.68	4.21	4.21	0.00	9.05	9.05
G292.27-0.015	292.27	-0.01	4.72	1.83	50.0	5.85	5.85
G292.39-0.437	292.39	-0.43	4.83	3.76	4.96	5.45	5.45
G290.30-0.014	290.30	-0.01	6.16	6.16	0.00	6.05	6.05
G290.56-1.412	290.59	-1.42	6.23	2.68	23.5	7.35	7.35
G290.77-1.371	290.79	-1.38	4.28	3.75	117.	7.45	7.45
G290.46-0.338	290.45	-0.31	6.51	5.54	86.3	8.25	8.25
G293.89-0.824	293.89	-0.82	11.8	7.42	47.2	10.3	10.3
G293.60-1.616	293.58	-1.52	9.93	6.17	56.9	2.35	2.35
G294.61-1.201	294.60	-1.20	4.89	3.07	178.	9.55	9.55
G295.14-0.577	295.12	-0.55	4.87	3.20	50.2	10.8	10.8
G295.15-1.293	295.16	-1.27	6.53	3.11	102.	2.25	2.25
G295.14-1.589	295.12	-1.56	9.06	7.75	153.	2.25	2.25
G294.91-1.667	294.91	-1.66	4.72	4.72	0.00	2.25	2.25
G294.23-0.477	294.23	-0.47	6.09	6.09	0.00	7.25	7.25
G294.08-1.614	294.08	-1.61	6.53	1.82	146.	2.25	2.25
G294.50-1.621	294.49	-1.64	3.89	2.22	87.8	2.25	2.25
G289.48+0.112	289.48	0.11	3.70	3.70	0.00	8.15	8.15
G294.97+0.103	294.97	0.10	3.66	3.66	0.00	9.45	9.45
G295.07+0.538	295.09	0.53	3.56	2.83	0.00	7.55	7.55
G296.79-1.120	296.74	-1.10	18.0	10.3	62.1	9.95	9.95
G296.85-1.457	296.89	-1.44	8.18	4.36	16.5	9.85	9.85
G297.14-1.357	297.14	-1.36	8.53	4.09	6.62	9.75	9.75
G297.50-0.822	297.49	-0.79	13.9	7.14	53.0	10.7	10.7
G297.32-0.269	297.28	-0.27	5.88	3.00	170.	9.35	9.35
G297.47-0.008	297.46	0.01	4.64	3.84	0.00	9.45	9.45
G298.18-0.321	298.17	-0.30	11.1	6.34	80.7	11.3	11.3
G298.74-0.019	298.74	-0.01	7.60	5.03	129.	10.2	10.2
G298.85+0.170	298.85	0.17	5.55	5.55	0.00	10.4	10.4
G298.38-0.069	298.38	-0.06	6.39	6.39	0.00	11.0	11.0
G298.15-0.617	298.15	-0.61	7.17	7.17	0.00	11.9	11.9
G299.52+0.119	299.52	0.11	3.37	3.37	0.00	7.75	7.75
G298.91+0.461	298.90	0.46	6.17	4.21	82.4	2.55	2.55
G300.34-0.281	300.34	-0.28	5.58	5.58	0.00	11.8	11.8
G300.39+0.303	300.38	0.31	9.67	6.20	84.3	7.15	7.15
G301.55-0.259	301.55	-0.25	6.56	4.40	1.43	6.75	6.75
G302.44+0.025	302.42	0.03	11.9	7.57	40.1	5.95	5.95
G302.72+0.190	302.67	0.13	9.27	6.52	49.3	5.55	5.55
G302.74-0.078	302.74	-0.07	4.84	4.84	0.00	5.55	5.55
G302.19-0.856	302.19	-0.85	7.38	7.38	0.00	12.0	12.0
G302.50-0.728	302.50	-0.72	4.47	4.47	0.00	12.0	12.0
G302.67-0.680	302.70	-0.70	6.40	5.19	62.8	11.9	11.9
G303.88-0.792	303.88	-0.79	4.21	1.87	10.4	12.4	12.4
G304.52-0.076	304.52	-0.07	9.06	9.06	0.00	7.45	7.45
G304.91+0.018	304.91	0.01	4.71	4.71	0.00	7.25	7.25
G305.65+0.522	305.65	0.52	3.97	3.97	0.00	6.85	6.85
G306.28-0.033	306.29	-0.03	4.72	2.58	1.65	2.25	2.25
G305.38-0.257	305.38	-0.25	4.08	0.93	178.	2.55	2.55
G306.62+0.008	306.62	0.00	3.07	3.07	0.00	7.75	7.75
G304.01+0.381	304.01	0.38	9.99	5.58	98.0	5.75	5.75
G305.50-0.707	305.49	-0.70	9.37	3.97	152.	8.35	8.35

Table B.2: Column [9-16]. ⁽¹⁾H II regions from Anderson et al. (2014), ⁽²⁾H II regions from Paladini et al. (2003)

[9]	[10]	[11]	[12]	[13]	[14]	[15]	[16]
R (kpc)	Z (pc)	a-axis (pc)	b-axis (pc)	Density Clumps.pc ⁻²	Nb _{clumps}	Nb _{proto} /Nb _{pre}	H II regions
8.01	-5.5	17.1	6.80	0.065	24	1.00	G289.799-00.839 ⁽¹⁾
9.72	-4.5	7.71	5.14	0.088	11	0.57	G289.582-00.636 ⁽¹⁾
2.46	-5.3	20.0	16.4	0.023	24	1.00	290.5-00.8 ⁽²⁾
8.08	-5.7	11.2	11.2	0.035	14	3.66	G291.614-00.525 ⁽¹⁾
9.39	-0.0	8.07	2.85	0.137	10	0.25	
9.82	-2.3	7.75	5.95	0.082	12	0.71	
4.78	-0.0	10.8	10.8	0.043	16	0.23	
2.54	-7.2	13.4	5.67	0.050	12	0.20	
1.26	-7.3	9.27	8.13	0.063	15	0.25	
3.56	-2.5	15.6	13.3	0.030	20	0.42	
18.8	-7.6	35.7	22.3	0.019	49	1.45	G293.898-00.826 ⁽¹⁾
10.6	-2.3	6.78	4.21	0.311	28	1.33	G293.613-01.291 ⁽¹⁾
17.3	-8.9	13.5	8.72	0.042	16	0.33	G294.793-01.330 ⁽¹⁾
16.5	-5.7	15.4	9.91	0.039	19	1.11	G295.163-00.675 ⁽¹⁾
9.65	-2.1	4.25	2.05	0.545	15	0.15	G294.793-01.330 ⁽¹⁾
9.72	-2.2	5.91	5.09	0.422	40	0.90	G294.793-01.330 ⁽¹⁾
10.0	-2.2	3.10	3.10	0.461	14	0.75	G294.793-01.330 ⁽¹⁾
15.6	-3.3	12.7	12.7	0.046	24	0.71	G294.206-00.464 ⁽¹⁾
10.7	-2.2	4.25	1.09	1.089	16	0.33	
10.5	-2.2	2.57	1.55	0.795	10	1.50	G294.517-01.626 ⁽¹⁾
10.7	0.91	8.89	8.89	0.040	10	1.00	G289.505+00.127 ⁽¹⁾
15.9	0.97	10.3	10.3	0.029	10	0.42	G295.048+00.060 ⁽¹⁾
13.8	3.82	7.82	6.38	0.063	10	0.00	
4.10	-8.8	52.3	29.9	0.010	51	0.96	G296.660-00.925 ⁽¹⁾
2.82	-9.7	23.5	12.7	0.019	18	0.20	
1.26	-9.5	24.2	11.6	0.032	29	0.93	G297.089-01.343 ⁽¹⁾
3.85	-7.6	43.4	22.2	0.018	55	1.20	G297.497-00.758 ⁽¹⁾
1.35	-2.5	16.1	7.90	0.034	14	0.55	G297.312-00.295 ⁽¹⁾
2.80	0.17	12.6	10.3	0.036	15	0.66	G297.570-00.036 ⁽¹⁾
9.86	-3.4	36.7	21.1	0.021	52	2.46	G298.224-00.334 ⁽¹⁾
13.3	-0.2	22.6	15.0	0.022	24	0.50	G298.529-00.251 ⁽¹⁾
14.2	1.77	16.9	16.9	0.020	18	2.00	G298.756+00.059 ⁽¹⁾
11.3	-0.7	20.5	20.5	0.013	18	1.57	G298.529-00.251 ⁽¹⁾
10.1	-6.9	25.0	25.0	0.007	14	0.40	
15.0	0.92	7.56	7.56	0.061	11	0.83	
9.27	1.14	4.57	3.17	0.350	16	0.23	G298.924+00.473 ⁽¹⁾
20.3	-3.2	19.1	19.1	0.021	25	1.27	G300.084-00.485 ⁽¹⁾
15.6	2.24	20.0	12.8	0.040	33	0.17	G300.566+00.169 ⁽¹⁾
12.4	-1.7	12.9	8.72	0.050	18	0.50	G301.615-00.340 ⁽¹⁾
7.43	0.22	20.6	13.1	0.067	58	0.61	G302.270+00.073 ⁽¹⁾
5.96	0.72	14.9	10.4	0.062	31	1.81	G302.631+00.030 ⁽¹⁾
5.57	-0.4	7.90	7.90	0.066	13	0.18	G302.582-00.083 ⁽¹⁾
12.4	-9.1	25.9	25.9	0.011	24	0.71	
9.92	-8.0	15.5	15.5	0.024	19	1.11	G302.503-00.762 ⁽¹⁾
8.12	-7.7	22.2	17.9	0.019	24	2.00	G302.614-00.756 ⁽¹⁾
5.93	-8.8	15.4	6.07	0.043	13	1.60	G303.872-00.792 ⁽¹⁾
8.19	-0.5	19.5	19.5	0.030	37	0.54	G304.465-00.023 ⁽¹⁾
10.5	0.13	10.0	10.0	0.031	10	0.11	
13.7	3.41	7.84	7.84	0.072	14	0.40	G305.637+00.515 ⁽¹⁾
10.6	-0.0	3.10	1.73	0.648	11	0.37	
9.69	-0.6	3.04	0.88	1.185	10	1.50	G305.322-00.255 ⁽¹⁾
16.2	0.06	7.07	7.07	0.069	11	0.22	
4.78	2.13	16.7	9.31	0.077	38	1.71	G303.896+00.407 ⁽¹⁾
14.3	-5.4	22.7	9.56	0.030	21	0.50	

Table B.3: Partial list of potential cluster candidates.

[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
Overdensities name	l_{center} (deg)	b_{center} (deg)	a-axis (arcmin)	b-axis (arcmin)	PA (deg)	Density (Clumps/arcmin ²)	Nb _{clumps}	Nb _{proto} /Nb _{pre}	Over-density type
G289.78-1.314	289.78	-1.20	23.3	13.4	76.1	0.15	128	2.55	SD
G289.97-0.911	289.98	-0.93	6.19	3.94	106.	0.11	15	0.87	SD
G289.82-0.619	289.80	-0.62	3.60	1.88	14.0	0.12	10	0.25	SD
G289.88-1.157	289.88	-1.15	5.50	2.51	126.	0.24	11	2.66	SD
G290.03-1.244	290.03	-1.24	6.69	6.69	0.00	0.20	21	1.33	SD
G290.14-1.400	290.15	-1.41	7.33	4.06	89.9	0.28	20	0.05	SD
G290.25-0.823	290.25	-0.82	5.12	2.37	44.7	0.39	20	1.85	SD
G291.26-0.745	291.26	-0.74	15.6	15.6	0.00	0.19	122	1.34	SD
G291.21-0.259	291.19	-0.29	13.1	5.26	154.	0.24	50	2.33	SD
G291.76-1.037	291.74	-1.06	5.77	3.07	95.7	0.26	14	0.55	NDA
G291.97-0.972	291.94	-0.98	9.40	3.13	157.	0.16	22	2.14	NDA
G292.07-1.091	292.10	-1.11	9.15	3.66	177.	0.18	22	0.29	SD
G291.59-0.450	291.57	-0.47	15.0	11.2	102.	0.22	101	4.31	SCC
G291.79-0.412	291.79	-0.41	6.46	6.46	0.00	0.24	20	Inf	SD
G291.97-0.167	291.97	-0.16	9.24	9.24	0.00	0.17	25	1.08	SD
G291.90-0.013	291.90	-0.01	3.40	2.15	36.7	0.16	10	0.25	SD
G290.25-1.002	290.24	-1.04	8.92	5.14	18.8	0.21	19	0.58	SD
G292.26-0.679	292.25	-0.64	5.63	3.93	69.8	0.12	14	2.50	SD
G290.66+0.332	290.62	0.31	6.55	6.09	119.	0.29	31	6.75	NDA
G291.48-1.621	291.48	-1.62	2.42	2.42	0.00	0.45	10	0.66	NDA
G293.05-0.956	293.05	-0.95	8.01	8.01	0.00	0.18	31	2.87	NDA
G293.47-0.968	293.53	-0.94	14.0	4.99	136.	0.17	40	0.53	SD
G293.61-0.946	293.58	-0.97	4.61	4.32	0.00	0.34	10	1.00	SD
G293.66-1.223	293.62	-1.27	10.3	8.80	156.	0.15	37	0.37	NDA
G294.22-0.814	294.22	-0.81	3.98	2.19	50.6	0.22	10	0.66	NDA
G295.16-0.720	295.16	-0.74	12.3	4.32	14.0	0.19	35	2.88	SD
G294.47-0.482	294.47	-0.48	7.11	7.11	0.00	0.24	22	0.69	SD
G294.36-1.515	294.36	-1.51	4.68	4.68	0.00	0.13	13	1.16	SD
G295.86-0.143	295.85	-0.14	4.24	3.83	168.	0.20	20	0.81	SD
G297.26-0.969	297.26	-0.96	7.75	7.75	0.00	0.19	18	0.05	NDA
G298.74-0.345	298.70	-0.38	22.9	16.6	12.8	0.15	177	2.00	SD
G298.70+0.095	298.68	0.11	3.62	2.85	177.	0.22	13	0.62	SD
G299.06+0.188	299.06	0.18	6.74	6.74	0.00	0.26	21	0.61	SD
G297.96-0.515	297.96	-0.51	6.86	6.86	0.00	0.23	10	0.11	SCC
G299.34-0.297	299.33	-0.30	9.84	4.93	98.9	0.17	28	0.75	SD
G299.80+0.021	299.78	0.04	8.86	3.82	4.05	0.11	14	0.55	SD
G300.16-0.110	300.14	-0.10	5.16	3.33	0.00	0.34	16	1.66	SD
G300.09-0.384	300.08	-0.38	3.13	2.87	89.9	0.35	10	0.25	NDA
G300.20-0.353	300.21	-0.38	5.53	2.22	67.9	0.23	14	1.00	SD

ANNEXE

B

CATALOGUE D'AMAS

TABLE B.1: Complete list of cluster candidates.

[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]	[14]	[15]	[16]
Cluster candidates name	l_{center} (deg)	b_{center} (deg)	a-axis (arcmin)	b-axis (arcmin)	PA (deg)	D_{peak} (kpc)	\bar{D} (kpc)	R (kpc)	Z (pc)	a-axis (pc)	b-axis (pc)	Density Clumps.pc ⁻²	Nb _{clumps}	Nb _{proto} /Nb _{pre}	H II regions
G289.85-0.764	289.87	-0.76	7.36	2.85	16.4	8.05	8.05	8.01	-5.5	17.1	6.80	0.065	24	1.00	G289.799-00.839 ⁽¹⁾
G289.57-0.656	289.58	-0.65	3.56	2.28	43.5	7.45	7.45	9.72	-4.5	7.71	5.14	0.088	11	0.57	G289.582-00.636 ⁽¹⁾
G290.52-0.771	290.59	-0.78	9.14	7.54	162.	7.55	7.55	2.46	-5.3	20.0	16.4	0.023	24	1.00	290.5-00.8 ⁽²⁾
G291.84-0.684	291.84	-0.68	4.21	4.21	0.00	9.05	9.05	8.08	-5.7	11.2	11.2	0.035	14	3.66	G291.614-00.525 ⁽¹⁾
G292.27-0.015	292.27	-0.01	4.72	1.83	50.0	5.85	5.85	9.39	-0.0	8.07	2.85	0.137	10	0.25	
G292.39-0.437	292.39	-0.43	4.83	3.76	4.96	5.45	5.45	9.82	-2.3	7.75	5.95	0.082	12	0.71	
G290.30-0.014	290.30	-0.01	6.16	6.16	0.00	6.05	6.05	4.78	-0.0	10.8	10.8	0.043	16	0.23	
G290.56-1.412	290.59	-1.42	6.23	2.68	23.5	7.35	7.35	2.54	-7.2	13.4	5.67	0.050	12	0.20	
G290.77-1.371	290.79	-1.38	4.28	3.75	117.	7.45	7.45	1.26	-7.3	9.27	8.13	0.063	15	0.25	
G290.46-0.338	290.45	-0.31	6.51	5.54	86.3	8.25	8.25	3.56	-2.5	15.6	13.3	0.030	20	0.42	
G293.89-0.824	293.89	-0.82	11.8	7.42	47.2	10.3	10.3	18.8	-7.6	35.7	22.3	0.019	49	1.45	G293.898-00.826 ⁽¹⁾
G293.60-1.616	293.58	-1.52	9.93	6.17	56.9	2.35	2.35	10.6	-2.3	6.78	4.21	0.311	28	1.33	G293.613-01.291 ⁽¹⁾
G294.61-1.201	294.60	-1.20	4.89	3.07	178.	9.55	9.55	17.3	-8.9	13.5	8.72	0.042	16	0.33	G294.793-01.330 ⁽¹⁾
G295.14-0.577	295.12	-0.55	4.87	3.20	50.2	10.8	10.8	16.5	-5.7	15.4	9.91	0.039	19	1.11	G295.163-00.675 ⁽¹⁾
G295.15-1.293	295.16	-1.27	6.53	3.11	102.	2.25	2.25	9.65	-2.1	4.25	2.05	0.545	15	0.15	G294.793-01.330 ⁽¹⁾
G295.14-1.589	295.12	-1.56	9.06	7.75	153.	2.25	2.25	9.72	-2.2	5.91	5.09	0.422	40	0.90	G294.793-01.330 ⁽¹⁾
G294.91-1.667	294.91	-1.66	4.72	4.72	0.00	2.25	2.25	10.0	-2.2	3.10	3.10	0.461	14	0.75	G294.793-01.330 ⁽¹⁾
G294.23-0.477	294.23	-0.47	6.09	6.09	0.00	7.25	7.25	15.6	-3.3	12.7	12.7	0.046	24	0.71	G294.206-00.464 ⁽¹⁾
G294.08-1.614	294.08	-1.61	6.53	1.82	146.	2.25	2.25	10.7	-2.2	4.25	1.09	1.089	16	0.33	
G294.50-1.621	294.49	-1.64	3.89	2.22	87.8	2.25	2.25	10.5	-2.2	2.57	1.55	0.795	10	1.50	G294.517-01.626 ⁽¹⁾
G289.48+0.112	289.48	0.11	3.70	3.70	0.00	8.15	8.15	10.7	0.91	8.89	8.89	0.040	10	1.00	G289.505+00.127 ⁽¹⁾
G294.97+0.103	294.97	0.10	3.66	3.66	0.00	9.45	9.45	15.9	0.97	10.3	10.3	0.029	10	0.42	G295.048+00.060 ⁽¹⁾
G295.07+0.538	295.09	0.53	3.56	2.83	0.00	7.55	7.55	13.8	3.82	7.82	6.38	0.063	10	0.00	
G296.79-1.120	296.74	-1.10	18.0	10.3	62.1	9.95	9.95	4.10	-8.8	52.3	29.9	0.010	51	0.96	G296.660-00.925 ⁽¹⁾
G296.85-1.457	296.89	-1.44	8.18	4.36	16.5	9.85	9.85	2.82	-9.7	23.5	12.7	0.019	18	0.20	
G297.14-1.357	297.14	-1.36	8.53	4.09	6.62	9.75	9.75	1.26	-9.5	24.2	11.6	0.032	29	0.93	G297.089-01.343 ⁽¹⁾
G297.50-0.822	297.49	-0.79	13.9	7.14	53.0	10.7	10.7	3.85	-7.6	43.4	22.2	0.018	55	1.20	G297.497-00.758 ⁽¹⁾
G297.32-0.269	297.28	-0.27	5.88	3.00	170.	9.35	9.35	1.35	-2.5	16.1	7.90	0.034	14	0.55	G297.312-00.295 ⁽¹⁾
G297.47-0.008	297.46	0.01	4.64	3.84	0.00	9.45	9.45	2.80	0.17	12.6	10.3	0.036	15	0.66	G297.570-00.036 ⁽¹⁾
G298.18-0.321	298.17	-0.30	11.1	6.34	80.7	11.3	11.3	9.86	-3.4	36.7	21.1	0.021	52	2.46	G298.224-00.334 ⁽¹⁾
G298.74-0.019	298.74	-0.01	7.60	5.03	129.	10.2	10.2	13.3	-0.2	22.6	15.0	0.022	24	0.50	G298.529-00.251 ⁽¹⁾
G298.85+0.170	298.85	0.17	5.55	5.55	0.00	10.4	10.4	14.2	1.77	16.9	16.9	0.020	18	2.00	G298.756+00.059 ⁽¹⁾
G298.38-0.069	298.38	-0.06	6.39	6.39	0.00	11.0	11.0	11.3	-0.7	20.5	20.5	0.013	18	1.57	G298.529-00.251 ⁽¹⁾
G298.15-0.617	298.15	-0.61	7.17	7.17	0.00	11.9	11.9	10.1	-6.9	25.0	25.0	0.007	14	0.40	
G299.52+0.119	299.52	0.11	3.37	3.37	0.00	7.75	7.75	15.0	0.92	7.56	7.56	0.061	11	0.83	
...

TABLE B.1: continued.

[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]	[14]	[15]	[16]
Cluster candidates name	l_{center} (deg)	b_{center} (deg)	a-axis (arcmin)	b-axis (arcmin)	PA (deg)	D_{peak} (kpc)	\bar{D} (kpc)	R (kpc)	Z (pc)	a-axis (pc)	b-axis (pc)	Density Clumps.pc ⁻²	Nb _{clumps}	Nb _{proto} /Nb _{pre}	H II regions
G298.91+0.461	298.90	0.46	6.17	4.21	82.4	2.55	2.55	9.27	1.14	4.57	3.17	0.350	16	0.23	G298.924+00.473 ⁽¹⁾
G300.34-0.281	300.34	-0.28	5.58	5.58	0.00	11.8	11.8	20.3	-3.2	19.1	19.1	0.021	25	1.27	G300.084-00.485 ⁽¹⁾
G300.39+0.303	300.38	0.31	9.67	6.20	84.3	7.15	7.15	15.6	2.24	20.0	12.8	0.040	33	0.17	G300.566+00.169 ⁽¹⁾
G301.55-0.259	301.55	-0.25	6.56	4.40	1.43	6.75	6.75	12.4	-1.7	12.9	8.72	0.050	18	0.50	G301.615-00.340 ⁽¹⁾
G302.44+0.025	302.42	0.03	11.9	7.57	40.1	5.95	5.95	7.43	0.22	20.6	13.1	0.067	58	0.61	G302.270+00.073 ⁽¹⁾
G302.72+0.190	302.67	0.13	9.27	6.52	49.3	5.55	5.55	5.96	0.72	14.9	10.4	0.062	31	1.81	G302.631+00.030 ⁽¹⁾
G302.74-0.078	302.74	-0.07	4.84	4.84	0.00	5.55	5.55	5.57	-0.4	7.90	7.90	0.066	13	0.18	G302.582-00.083 ⁽¹⁾
G302.19-0.856	302.19	-0.85	7.38	7.38	0.00	12.0	12.0	12.4	-9.1	25.9	25.9	0.011	24	0.71	
G302.50-0.728	302.50	-0.72	4.47	4.47	0.00	12.0	12.0	9.92	-8.0	15.5	15.5	0.024	19	1.11	G302.503-00.762 ⁽¹⁾
G302.67-0.680	302.70	-0.70	6.40	5.19	62.8	11.9	11.9	8.12	-7.7	22.2	17.9	0.019	24	2.00	G302.614-00.756 ⁽¹⁾
G303.88-0.792	303.88	-0.79	4.21	1.87	10.4	12.4	12.4	5.93	-8.8	15.4	6.07	0.043	13	1.60	G303.872-00.792 ⁽¹⁾
G304.52-0.076	304.52	-0.07	9.06	9.06	0.00	7.45	7.45	8.19	-0.5	19.5	19.5	0.030	37	0.54	G304.465-00.023 ⁽¹⁾
G304.91+0.018	304.91	0.01	4.71	4.71	0.00	7.25	7.25	10.5	0.13	10.0	10.0	0.031	10	0.11	
G305.65+0.522	305.65	0.52	3.97	3.97	0.00	6.85	6.85	13.7	3.41	7.84	7.84	0.072	14	0.40	G305.637+00.515 ⁽¹⁾
G306.28-0.033	306.29	-0.03	4.72	2.58	1.65	2.25	2.25	10.6	-0.0	3.10	1.73	0.648	11	0.37	
G305.38-0.257	305.38	-0.25	4.08	0.93	178.	2.55	2.55	9.69	-0.6	3.04	0.88	1.185	10	1.50	G305.322-00.255 ⁽¹⁾
G306.62+0.008	306.62	0.00	3.07	3.07	0.00	7.75	7.75	16.2	0.06	7.07	7.07	0.069	11	0.22	
G304.01+0.381	304.01	0.38	9.99	5.58	98.0	5.75	5.75	4.78	2.13	16.7	9.31	0.077	38	1.71	G303.896+00.407 ⁽¹⁾
G305.50-0.707	305.49	-0.70	9.37	3.97	152.	8.35	8.35	14.3	-5.4	22.7	9.56	0.030	21	0.50	
G306.82+0.237	306.82	0.23	9.67	9.67	0.00	7.35	7.35	15.7	1.72	20.6	20.6	0.027	37	0.27	G306.740+00.264 ⁽¹⁾
G306.85-0.228	306.85	-0.20	3.96	2.04	80.1	7.95	7.95	16.3	-1.6	9.10	4.75	0.073	10	0.11	
G307.63-0.257	307.60	-0.21	11.9	5.50	70.8	7.45	7.45	13.9	-1.6	25.8	11.7	0.045	44	1.58	G307.585-00.249 ⁽¹⁾
G307.92+0.020	307.92	0.02	5.50	5.50	0.00	4.05	4.05	10.2	0.08	6.55	6.55	0.103	14	0.55	G307.862+00.043 ⁽¹⁾
G307.59-0.586	307.58	-0.58	4.59	2.08	41.9	7.55	7.55	14.1	-4.1	10.0	4.51	0.083	12	1.00	G307.559-00.585 ⁽¹⁾
G307.46-0.837	307.46	-0.82	3.83	1.98	138.	13.3	13.3	19.9	-9.7	14.5	7.98	0.030	11	2.66	G307.412-00.795 ⁽¹⁾
G309.52-0.094	309.50	-0.08	6.25	2.47	59.3	6.85	6.85	2.42	-0.6	12.5	4.73	0.075	14	0.55	
G309.57-0.738	309.55	-0.75	3.24	2.81	47.4	7.55	7.55	1.74	-5.1	6.89	6.38	0.072	10	0.66	G309.540-00.720 ⁽¹⁾
G308.74+0.607	308.70	0.66	13.3	6.66	93.4	6.35	6.35	7.56	3.89	24.7	12.4	0.050	49	1.45	G308.571+00.500 ⁽¹⁾
G310.20-0.144	310.23	-0.16	8.12	2.43	17.2	7.05	7.05	4.08	-1.1	16.6	4.86	0.078	20	1.50	G310.168-00.132 ⁽¹⁾
G310.52-0.207	310.52	-0.20	3.29	3.29	0.00	6.35	6.35	6.02	-1.3	6.20	6.20	0.107	13	1.16	G310.519-00.220 ⁽¹⁾
G311.46+0.137	311.46	0.14	5.52	2.93	172.	7.15	7.15	11.9	1.01	11.5	6.04	0.068	15	0.50	G311.471+00.106 ⁽¹⁾
G311.11+0.978	311.12	0.99	4.85	2.91	79.0	7.25	7.25	10.1	6.06	10.3	6.13	0.065	13	0.00	
G310.49+0.761	310.49	0.76	6.87	6.87	0.00	6.85	6.85	5.88	4.72	13.7	13.7	0.025	15	0.15	
G311.22-0.501	311.23	-0.47	13.3	4.33	98.0	7.45	7.45	10.9	-3.4	28.9	9.27	0.042	36	0.44	
G310.88-0.381	310.88	-0.38	13.4	13.4	0.00	6.75	6.75	8.42	-2.5	26.3	26.3	0.022	49	1.33	G310.808-00.399 ⁽¹⁾
...

TABLE B.1: continued.

[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]	[14]	[15]	[16]
Cluster candidates name	l_{center} (deg)	b_{center} (deg)	a-axis (arcmin)	b-axis (arcmin)	PA (deg)	D_{peak} (kpc)	\bar{D} (kpc)	R (kpc)	Z (pc)	a-axis (pc)	b-axis (pc)	Density Clumps.pc ⁻²	Nb _{clumps}	Nb _{proto} /Nb _{pre}	H II regions
G310.30+0.576	310.30	0.57	3.58	2.02	3.90	6.75	6.75	4.62	3.67	6.99	4.03	0.112	10	0.42	
G311.66-0.118	311.66	-0.12	9.88	4.24	8.18	6.95	6.95	12.7	-0.8	19.9	8.65	0.042	23	1.87	
G310.60-0.434	310.60	-0.43	6.84	6.84	0.00	6.85	6.85	6.65	-2.8	13.5	13.5	0.031	18	2.00	G310.630-00.421 ⁽¹⁾
G311.38+0.707	311.38	0.70	7.61	7.61	0.00	7.25	7.25	11.6	4.71	16.0	16.0	0.034	28	0.64	G311.396+00.583 ⁽¹⁾
G309.73-0.026	309.72	-0.04	4.79	2.61	173.	7.65	7.65	0.85	-0.3	10.5	5.90	0.071	14	0.07	
G309.71+0.306	309.71	0.30	6.77	6.77	0.00	6.45	6.45	2.05	1.94	12.7	12.7	0.027	14	0.16	
G309.61+0.505	309.62	0.49	7.85	1.83	122.	6.65	6.65	2.02	3.17	15.2	3.24	0.096	15	0.36	
G309.45+0.496	309.45	0.49	4.02	2.67	129.	7.45	7.45	2.45	3.54	8.53	5.75	0.077	12	0.33	G309.362+00.622 ⁽¹⁾
G310.46+0.309	310.46	0.30	3.67	3.67	0.00	7.85	7.85	5.89	2.39	8.57	8.57	0.056	13	0.85	G310.465+00.198 ⁽¹⁾
G310.82+1.074	310.80	1.09	4.28	1.91	136.	6.75	6.75	7.91	5.99	8.40	4.03	0.093	10	0.00	
G311.88+0.445	311.90	0.44	5.43	2.44	157.	7.05	7.05	13.7	3.03	11.1	4.86	0.087	15	0.36	
G312.50+0.039	312.42	0.07	26.5	24.4	148.	7.75	7.75	15.8	0.55	59.8	55.0	0.016	172	0.91	
G312.80-0.194	312.80	-0.19	3.84	3.84	0.00	7.65	7.65	16.1	-1.4	8.35	8.35	0.054	12	0.50	
G311.81-0.749	311.77	-0.73	5.59	5.03	89.9	8.15	8.15	14.1	-5.4	13.1	11.9	0.028	14	0.55	
G312.28-0.324	312.26	-0.30	7.03	3.74	157.	7.45	7.45	15.2	-2.2	15.2	8.13	0.038	15	0.50	G312.299-00.327 ⁽¹⁾
G312.39-0.442	312.41	-0.44	6.21	4.71	168.	7.75	7.75	15.8	-3.3	13.9	10.7	0.034	16	0.33	G312.385-00.416 ⁽¹⁾
G312.13-0.460	312.13	-0.46	6.12	2.17	163.	7.55	7.55	14.9	-3.3	13.5	4.51	0.062	12	0.09	
G312.67-0.433	312.65	-0.42	4.43	2.85	64.8	7.95	7.95	16.3	-3.2	10.2	6.72	0.046	10	0.11	
G313.12-0.646	313.12	-0.64	8.09	3.58	139.	8.35	8.35	16.7	-5.0	19.5	8.64	0.039	21	0.23	
G313.59+0.085	313.58	0.05	30.5	23.0	6.51	8.15	8.15	15.6	0.41	72.4	54.7	0.014	175	0.57	G313.411+00.127 ⁽¹⁾
G313.23+0.136	313.23	0.13	6.16	6.16	0.00	8.15	8.15	16.3	1.10	14.6	14.6	0.017	12	1.40	G313.411+00.127 ⁽¹⁾
G313.94+0.221	313.94	0.22	9.25	9.25	0.00	8.15	8.15	14.3	1.79	21.9	21.9	0.012	19	0.35	
G313.69-0.499	313.70	-0.47	12.5	5.14	172.	8.65	8.65	15.6	-3.9	31.4	13.0	0.031	40	0.81	G313.696-00.550 ⁽¹⁾
G314.08-0.111	314.08	-0.10	9.01	3.07	162.	8.25	8.25	13.8	-0.8	21.6	7.53	0.031	16	0.14	G313.949-00.037 ⁽¹⁾
G314.25+0.111	314.25	0.11	6.35	2.60	40.4	7.85	7.85	12.6	0.87	14.5	6.06	0.068	19	2.80	G314.217-00.009 ⁽¹⁾
G314.23+0.416	314.24	0.41	11.0	6.53	63.6	6.95	6.95	12.0	2.81	22.3	13.1	0.059	55	1.20	G314.078+00.483 ⁽¹⁾
G313.16+0.452	313.15	0.42	13.6	7.23	169.	6.95	6.95	15.3	2.88	27.5	14.5	0.029	37	0.27	G313.088+00.362 ⁽¹⁾
G314.30+0.938	314.30	0.93	7.79	7.79	0.00	7.95	7.95	12.4	6.41	17.9	17.9	0.010	11	0.00	
G313.37-0.627	313.37	-0.62	3.59	2.40	94.9	8.25	8.25	16.2	-4.8	8.54	5.69	0.071	11	0.00	
G314.39-0.403	314.39	-0.40	7.02	7.02	0.00	8.35	8.35	12.2	-3.2	17.0	17.0	0.018	17	0.06	
G314.76-0.131	314.76	-0.13	12.1	12.1	0.00	7.95	7.95	9.66	-1.0	28.1	28.1	0.016	42	0.20	
G315.03-0.335	315.04	-0.36	4.60	2.42	90.4	8.15	8.15	7.80	-2.9	10.8	5.62	0.051	10	0.11	
G315.04+0.205	315.04	0.20	3.53	3.53	0.00	7.85	7.85	7.68	1.59	8.13	8.13	0.048	10	0.25	
G314.42-0.685	314.42	-0.68	7.86	7.86	0.00	8.85	8.85	12.4	-5.6	20.2	20.2	0.008	11	0.10	G314.409-00.708 ⁽¹⁾
G315.69+0.003	315.67	0.03	4.27	2.70	140.	8.15	8.15	2.82	0.29	10.1	6.29	0.049	10	0.00	
G315.68+0.170	315.68	0.17	9.87	9.87	0.00	8.35	8.35	2.78	1.41	23.9	23.9	0.008	16	0.23	
...

TABLE B.1: continued.

[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]	[14]	[15]	[16]
Cluster candidates name	l_{center} (deg)	b_{center} (deg)	a-axis (arcmin)	b-axis (arcmin)	PA (deg)	D_{peak} (kpc)	\bar{D} (kpc)	R (kpc)	Z (pc)	a-axis (pc)	b-axis (pc)	Density Clumps.pc ⁻²	Nb _{clumps}	Nb _{proto} /Nb _{pre}	H II regions
G316.03+0.049	316.01	0.05	4.01	3.07	133.	8.85	8.85	0.35	0.48	10.1	8.08	0.042	11	0.83	
G316.34-0.019	316.35	-0.01	7.91	2.19	28.1	9.05	9.05	2.99	-0.1	20.7	5.41	0.042	15	0.87	
G316.23-0.292	316.25	-0.28	5.75	1.70	177.	9.15	9.15	2.18	-2.6	15.4	4.46	0.046	10	1.00	
G316.42-0.490	316.42	-0.49	9.31	6.28	71.8	9.25	9.25	3.66	-4.3	25.1	16.8	0.025	34	1.26	G316.369-00.368 ⁽¹⁾
G315.35-0.001	315.35	-0.00	8.83	3.23	162.	8.05	8.05	5.40	-0.0	20.6	7.35	0.025	12	0.00	
G315.27+0.115	315.25	0.12	3.46	2.50	144.	7.85	7.85	6.13	0.97	8.13	5.42	0.072	10	0.00	
G312.63+0.606	312.64	0.59	5.73	1.76	158.	7.65	7.65	16.0	4.28	12.6	3.73	0.073	11	0.37	
G311.86+0.857	311.84	0.83	5.61	3.71	102.	7.55	7.55	13.9	5.57	12.2	8.24	0.044	14	1.00	
G317.09+0.265	317.10	0.27	5.40	2.67	30.4	9.15	9.15	9.12	2.44	14.4	7.06	0.071	23	3.60	G317.030+00.028 ⁽¹⁾
G317.27+0.300	317.30	0.31	6.81	3.23	154.	9.55	9.55	10.8	2.91	18.9	8.72	0.032	17	4.66	G317.030+00.028 ⁽¹⁾
G317.41+0.231	317.41	0.24	3.75	2.79	11.7	9.65	9.65	11.6	2.34	10.5	8.16	0.040	11	1.75	G317.030+00.028 ⁽¹⁾
G317.42-0.272	317.42	-0.27	4.74	2.19	138.	9.65	9.65	11.7	-2.5	13.3	5.77	0.053	13	0.30	G317.300-00.211 ⁽¹⁾
G317.93-0.544	317.92	-0.52	5.26	2.90	9.87	8.25	8.25	13.6	-4.1	12.7	6.97	0.046	13	0.30	G317.979-00.672 ⁽¹⁾
G318.20-0.188	318.19	-0.18	4.52	2.27	124.	9.85	9.85	16.2	-1.8	13.1	6.80	0.035	10	0.42	
G317.85+0.162	317.85	0.16	4.92	4.92	0.00	9.65	9.65	14.4	1.56	13.7	13.7	0.028	17	1.42	G317.817+00.217 ⁽¹⁾
G317.66+0.453	317.67	0.44	2.78	1.98	30.7	9.35	9.35	13.1	4.04	7.90	5.59	0.079	11	0.10	
G317.51-0.377	317.51	-0.38	9.17	6.01	147.	9.65	9.65	12.3	-3.6	25.8	16.9	0.019	27	0.92	G317.466-00.350 ⁽¹⁾
G318.45-0.513	318.45	-0.51	7.75	7.75	0.00	10.0	10.0	17.4	-4.9	22.7	22.7	0.011	19	0.46	
G318.26-0.578	318.26	-0.57	7.14	2.76	166.	9.75	9.75	16.4	-5.3	20.1	7.52	0.043	21	0.61	G318.141-00.697 ⁽¹⁾
G318.53-0.187	318.53	-0.18	4.46	4.46	0.00	9.85	9.85	17.4	-1.8	12.7	12.7	0.021	11	0.22	
G316.96+0.535	316.97	0.51	9.18	3.46	131.	9.55	9.55	8.34	4.73	25.5	9.89	0.027	22	0.37	G317.030+00.028 ⁽¹⁾
G319.15-0.384	319.14	-0.36	4.30	2.72	63.5	11.3	11.3	19.8	-4.0	14.1	8.76	0.035	14	3.66	G319.164-00.421 ⁽¹⁾
G319.26-0.318	319.25	-0.32	6.60	5.18	47.2	11.6	11.6	20.1	-3.7	22.3	17.5	0.021	26	3.33	G319.188-00.329 ⁽¹⁾
G319.44-0.402	319.44	-0.39	3.46	2.74	169.	11.8	11.8	20.1	-4.5	12.2	9.14	0.031	11	0.57	G319.293-00.426 ⁽¹⁾
G319.26+0.234	319.27	0.23	3.16	1.44	36.2	9.75	9.75	18.2	2.28	8.90	3.36	0.106	10	1.50	G319.229+00.225 ⁽¹⁾
G320.35+0.113	320.32	0.10	7.92	4.71	142.	12.4	12.4	17.6	1.25	28.8	17.1	0.021	34	1.12	G320.257+00.173 ⁽¹⁾
G320.45-0.805	320.45	-0.80	11.1	4.84	76.2	9.15	9.15	14.0	-6.5	29.8	13.0	0.026	32	0.68	G320.286-00.890 ⁽¹⁾
G320.29-0.783	320.29	-0.80	7.48	3.66	74.8	8.95	8.95	14.6	-6.4	19.5	9.77	0.025	15	0.25	G320.286-00.890 ⁽¹⁾
G320.59+0.048	320.59	0.04	10.8	6.99	156.	8.15	8.15	12.5	0.39	25.6	16.6	0.031	42	0.75	G320.590+00.190 ⁽¹⁾
G320.74-0.699	320.74	-0.69	6.95	1.32	74.7	9.05	9.05	12.3	-5.8	18.2	3.12	0.067	12	0.09	
G320.00+0.919	320.00	0.91	3.94	1.20	0.50	2.85	2.85	10.6	2.26	3.26	0.98	1.090	11	0.00	
G319.87+0.798	319.89	0.83	6.09	3.57	2.12	2.95	2.95	10.8	2.18	5.19	3.05	0.401	20	0.53	G319.884+00.793 ⁽¹⁾
G318.61+0.086	318.62	0.05	3.74	2.99	135.	9.45	9.45	17.3	0.54	10.3	7.99	0.046	12	0.09	
G320.23+0.433	320.23	0.43	6.21	3.36	87.6	10.8	10.8	16.6	4.55	19.4	10.5	0.032	21	1.33	G320.248+00.442 ⁽¹⁾
G320.96-0.803	320.96	-0.80	7.77	3.24	14.0	9.25	9.25	11.0	-6.6	20.9	8.44	0.037	21	Inf	G321.061-00.763 ⁽¹⁾
...

TABLE B.1: continued.

[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]	[14]	[15]	[16]
Cluster candidates name	l_{center} (deg)	b_{center} (deg)	a-axis (arcmin)	b-axis (arcmin)	PA (deg)	D_{peak} (kpc)	\bar{D} (kpc)	R (kpc)	Z (pc)	a-axis (pc)	b-axis (pc)	Density Clumps.pc ⁻²	Nb _{clumps}	Nb _{proto} /Nb _{pre}	H II regions
G318.51-0.761	318.49	-0.79	4.54	3.19	119.	2.55	2.55	10.6	-1.8	3.40	2.32	0.480	12	0.20	
G319.33+0.876	319.33	0.87	8.52	5.20	12.7	2.95	2.95	11.4	2.26	7.34	4.43	0.322	33	0.65	
G319.04+0.477	319.06	0.47	3.83	2.65	50.3	9.95	9.95	18.4	4.58	10.8	7.68	0.041	11	0.37	G319.089+00.458 ⁽¹⁾
G321.26+0.349	321.26	0.34	5.29	2.38	142.	11.0	11.0	9.91	3.78	17.0	7.63	0.031	13	0.18	
G318.73+0.972	318.73	0.97	4.40	1.35	23.6	2.95	2.95	11.2	2.43	3.81	1.01	0.820	10	0.00	
G322.49+0.284	322.49	0.28	2.29	1.71	135.	11.3	11.3	3.44	3.18	7.83	5.54	0.080	11	0.37	G322.490+00.218 ⁽¹⁾
G322.39+0.209	322.38	0.21	3.39	1.84	32.4	11.5	11.5	3.15	2.48	11.2	5.63	0.065	13	1.16	G322.396+00.201 ⁽¹⁾
G322.46+0.643	322.45	0.63	4.23	3.22	2.40	9.75	9.75	1.90	5.75	12.1	8.90	0.044	15	0.25	
G323.23+0.309	323.23	0.30	7.51	3.94	18.8	10.3	10.3	8.63	3.10	22.6	11.8	0.019	16	0.06	G323.101+00.268 ⁽¹⁾
G323.80+0.007	323.80	0.00	5.73	2.05	121.	9.95	9.95	12.6	0.07	16.4	5.95	0.058	18	1.25	G323.806+00.020 ⁽¹⁾
G323.84+0.477	323.85	0.45	6.45	2.81	169.	10.2	10.2	13.2	4.54	19.3	8.66	0.037	20	0.05	
G324.04+0.635	324.03	0.64	8.34	5.77	64.1	10.1	10.1	14.2	6.06	24.5	17.1	0.019	26	0.08	G323.968+00.649 ⁽¹⁾
G324.33+0.505	324.35	0.51	5.76	3.86	0.00	10.1	10.1	15.9	5.01	17.1	11.6	0.027	17	0.13	
G323.58+0.278	323.58	0.27	8.26	3.53	46.6	10.1	10.1	11.2	2.77	24.2	10.5	0.024	20	0.25	G323.349+00.205 ⁽¹⁾
G324.63+0.387	324.63	0.38	2.17	2.17	0.00	11.1	11.1	18.1	4.20	6.66	6.66	0.071	10	0.66	
G325.34+0.007	325.35	0.01	3.42	1.91	95.7	11.1	11.1	19.6	0.18	10.8	6.66	0.056	13	0.62	G325.354-00.036 ⁽¹⁾
G325.26-0.165	325.26	-0.16	3.26	3.26	0.00	8.25	8.25	16.6	-1.3	8.05	8.05	0.053	11	2.66	G325.240-00.120 ⁽¹⁾
G325.92-0.366	325.92	-0.36	5.53	3.49	96.3	9.45	9.45	17.4	-3.3	15.3	9.78	0.046	22	0.83	
G326.04-0.511	326.04	-0.50	10.3	4.55	134.	9.95	9.95	17.6	-4.8	29.9	13.3	0.026	33	1.06	G326.036-00.497 ⁽¹⁾
G326.27-0.512	326.24	-0.46	7.41	5.78	105.	9.95	9.95	17.0	-4.4	21.4	16.8	0.022	25	0.78	G326.474-00.292 ⁽¹⁾
G326.55+0.837	326.46	0.77	27.5	18.4	43.7	2.75	2.75	10.2	1.91	22.0	14.7	0.244	250	0.47	G326.270+00.783 ⁽¹⁾
G326.31+0.483	326.29	0.49	5.10	3.08	49.9	3.05	3.05	10.7	1.46	4.59	2.78	0.547	22	0.57	G326.270+00.783 ⁽¹⁾
G326.30+0.765	326.29	0.75	4.76	3.61	144.	3.05	3.05	10.7	2.09	4.21	3.15	0.334	14	0.40	G326.270+00.783 ⁽¹⁾
G325.64+0.301	325.66	0.30	5.32	2.39	20.0	10.8	10.8	19.2	3.24	16.7	7.49	0.027	11	0.22	
G325.86+0.248	325.87	0.25	6.39	2.34	98.5	11.1	11.1	19.2	2.76	20.7	7.69	0.023	12	0.20	G325.775+00.100 ⁽¹⁾
G324.20-0.153	324.23	-0.12	6.24	3.01	114.	10.2	10.2	15.4	-1.3	18.7	8.66	0.025	13	0.44	
G324.69-0.329	324.68	-0.32	4.01	1.22	30.7	10.4	10.4	17.6	-3.3	11.9	3.60	0.081	11	0.83	G324.642-00.321 ⁽¹⁾
G324.96-0.604	324.96	-0.60	6.66	2.53	1.43	9.75	9.75	17.7	-5.5	19.0	7.52	0.051	23	0.91	G324.924-00.569 ⁽¹⁾
G323.48+0.530	323.48	0.53	3.39	3.39	0.00	10.5	10.5	10.7	5.34	10.3	10.3	0.032	11	0.00	G323.349+00.205 ⁽¹⁾
G322.60+0.729	322.64	0.75	5.37	3.98	0.00	10.1	10.1	3.61	6.98	15.6	11.6	0.022	13	0.00	
G328.12-0.107	328.13	-0.10	3.74	2.43	75.1	9.75	9.75	4.21	-1.0	10.6	6.73	0.053	12	1.40	
G328.23-0.367	328.23	-0.36	3.73	3.73	0.00	2.85	2.85	5.90	-1.0	3.11	3.11	0.361	11	0.10	
G327.52-0.830	327.54	-0.82	9.97	4.53	49.8	4.15	4.15	7.34	-3.0	12.0	5.54	0.180	38	0.65	G327.454-00.633 ⁽¹⁾
G328.28+0.683	328.26	0.67	6.53	3.08	49.1	8.55	8.55	2.70	5.32	16.1	7.81	0.035	14	0.27	
G328.31+0.443	328.30	0.41	4.68	2.87	0.00	8.45	8.45	2.37	3.41	11.6	7.14	0.076	20	1.00	G328.286+00.464 ⁽¹⁾
...

TABLE B.1: continued.

[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]	[14]	[15]	[16]
Cluster candidates name	l_{center} (deg)	b_{center} (deg)	a-axis (arcmin)	b-axis (arcmin)	PA (deg)	D_{peak} (kpc)	\bar{D} (kpc)	R (kpc)	Z (pc)	a-axis (pc)	b-axis (pc)	Density Clumps.pc ⁻²	Nb _{clumps}	Nb _{proto} /Nb _{pre}	H II regions
G328.58+0.551	328.59	0.53	4.28	2.80	61.8	8.05	8.05	0.45	4.06	10.0	6.80	0.074	16	0.33	
G328.66+0.423	328.63	0.42	5.74	3.71	177.	8.35	8.35	0.46	3.45	13.8	9.11	0.030	12	1.00	
G328.96+0.219	328.96	0.21	3.51	3.51	0.00	9.15	9.15	3.37	1.99	9.47	9.47	0.042	12	3.00	G328.961+00.248 ⁽¹⁾
G328.78-0.097	328.78	-0.09	3.15	3.15	0.00	11.4	11.4	3.55	-1.1	10.4	10.4	0.034	12	1.40	G328.807-00.078 ⁽¹⁾
G329.04+0.401	329.06	0.42	4.71	2.50	162.	9.55	9.55	4.44	3.93	13.1	6.59	0.040	11	0.83	
G328.48-0.134	328.49	-0.15	4.80	2.98	139.	9.75	9.75	1.49	-1.5	13.4	8.24	0.034	12	0.71	
G328.37-0.107	328.35	-0.11	4.69	4.40	30.5	9.75	9.75	2.43	-1.1	13.4	12.5	0.033	18	0.63	
G329.20+0.340	329.20	0.35	4.65	3.59	82.4	9.65	9.65	5.62	3.32	12.9	9.99	0.034	14	0.40	
G329.68+0.828	329.68	0.83	3.76	2.25	32.7	3.35	3.35	7.59	2.47	3.65	2.31	0.376	10	0.25	G329.804+00.917 ⁽¹⁾
G329.93+1.064	329.91	1.07	7.84	3.15	32.7	5.35	5.35	8.92	4.70	12.2	4.88	0.143	27	0.42	G329.804+00.917 ⁽¹⁾
G327.47+0.761	327.45	0.77	7.13	3.64	79.4	9.25	9.25	9.52	6.48	19.1	9.58	0.029	17	0.30	G327.537+00.627 ⁽¹⁾
G330.84+0.275	330.84	0.27	5.64	2.09	102.	4.35	4.35	11.7	1.18	7.20	2.60	0.203	12	0.33	
G331.10-0.156	331.11	-0.17	8.59	4.32	35.3	9.35	9.35	17.0	-1.5	23.2	11.6	0.042	36	2.27	G331.020-00.143 ⁽¹⁾
G330.86-0.509	330.86	-0.50	4.90	3.63	81.1	10.6	10.6	17.4	-5.1	15.1	11.0	0.038	20	1.22	G330.775-00.398 ⁽¹⁾
G330.84-0.647	330.84	-0.64	7.23	7.23	0.00	10.7	10.7	17.4	-6.4	22.5	22.5	0.012	20	0.17	G331.628-00.926 ⁽¹⁾
G331.37-0.009	331.37	-0.02	4.83	2.84	3.78	9.55	9.55	17.7	-0.2	13.5	8.07	0.055	19	18.0	G331.300-00.071 ⁽¹⁾
G330.52-0.389	330.49	-0.39	6.68	2.75	25.0	10.9	10.9	15.9	-4.1	21.3	8.45	0.024	14	0.27	
G330.14-0.232	330.14	-0.23	4.45	4.45	0.00	11.6	11.6	14.3	-2.6	15.0	15.0	0.015	11	0.37	
G330.47-0.582	330.47	-0.57	4.62	3.04	130.	10.7	10.7	15.6	-5.8	14.3	9.82	0.022	10	0.00	
G331.36+0.527	331.36	0.50	10.8	7.31	7.38	11.6	11.6	19.8	5.59	36.6	24.7	0.015	43	0.79	G331.365+00.521 ⁽¹⁾
G331.38+0.152	331.34	0.15	5.41	2.00	167.	3.75	3.75	12.0	0.56	5.93	2.24	0.334	14	0.75	G331.371+00.113 ⁽¹⁾
G331.10-0.524	331.09	-0.51	7.07	4.83	173.	10.6	10.6	18.2	-5.2	22.0	15.1	0.033	35	1.18	G330.986-00.433 ⁽¹⁾
G331.37+1.109	331.38	1.12	9.49	6.33	137.	4.95	4.95	13.2	4.46	13.6	9.04	0.126	49	1.33	G331.344+01.055 ⁽¹⁾
G332.13-0.450	332.13	-0.45	5.18	1.41	95.4	11.2	11.2	19.3	-4.8	16.9	3.88	0.082	17	1.83	G331.628-00.926 ⁽¹⁾
G332.65-0.629	332.62	-0.65	13.9	4.89	10.0	3.85	3.85	11.3	-2.3	15.5	5.48	0.167	45	1.36	G332.657-00.622 ⁽¹⁾
G333.51-0.264	333.52	-0.29	4.62	3.94	62.5	11.8	11.8	12.9	-3.4	15.8	13.5	0.031	21	2.00	G333.497-00.275 ⁽¹⁾
G334.41+0.062	334.39	0.06	6.01	1.29	169.	10.0	10.0	4.58	0.68	17.6	3.46	0.062	12	0.20	
G333.00-0.706	333.02	-0.68	8.97	4.01	132.	3.85	3.85	10.2	-2.4	10.0	4.40	0.251	35	0.75	G332.990-00.619 ⁽¹⁾
G334.12+0.593	334.12	0.59	6.72	4.11	141.	11.2	11.2	7.63	6.29	21.9	13.4	0.027	26	0.23	
G333.25-0.754	333.33	-0.79	9.05	6.54	4.45	3.65	3.65	9.12	-2.6	9.59	6.90	0.153	32	0.39	
G333.99-0.392	333.98	-0.39	3.76	2.45	78.7	11.8	11.8	9.17	-4.5	12.9	8.18	0.030	10	0.42	
G332.82+0.201	332.80	0.20	3.82	2.72	51.9	11.3	11.3	17.0	2.28	12.3	8.76	0.035	12	0.20	
G334.26+0.451	334.27	0.45	5.04	2.11	17.0	11.2	11.2	6.37	4.97	16.4	6.72	0.037	13	0.18	
G334.42-0.239	334.45	-0.24	4.87	3.42	148.	12.0	12.0	5.49	-2.9	17.1	11.7	0.029	19	1.71	G334.412-00.230 ⁽¹⁾
G334.99-0.278	334.99	-0.27	4.01	1.64	108.	12.8	12.8	4.55	-3.5	14.7	6.27	0.034	10	1.50	
...

TABLE B.1: continued.

[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]	[14]	[15]	[16]
Cluster candidates name	l_{center} (deg)	b_{center} (deg)	a-axis (arcmin)	b-axis (arcmin)	PA (deg)	D_{peak} (kpc)	\bar{D} (kpc)	R (kpc)	Z (pc)	a-axis (pc)	b-axis (pc)	Density Clumps.pc ⁻²	Nb _{clumps}	Nb _{proto} /Nb _{pre}	H II regions
G333.83-0.598	333.81	-0.61	6.14	1.83	151.	12.0	12.0	10.7	-6.9	21.6	5.88	0.025	10	0.00	
G334.60+1.005	334.59	0.99	4.59	1.38	152.	4.85	4.85	4.04	4.07	6.48	1.67	0.322	11	0.22	
G332.45+0.566	332.45	0.56	4.30	2.36	109.	11.3	11.3	18.5	6.08	14.1	7.83	0.034	12	0.09	G332.500+00.476 ⁽¹⁾
G333.74+0.363	333.70	0.36	4.19	2.05	27.6	2.55	2.55	7.84	0.91	3.04	1.52	0.821	12	0.50	G333.725+00.364 ⁽¹⁾
G335.28-0.285	335.28	-0.29	12.1	4.06	161.	3.05	3.05	5.83	-0.8	10.7	3.64	0.266	33	2.00	
G335.30-0.382	335.30	-0.37	3.52	2.07	127.	2.95	2.95	5.95	-1.0	3.05	1.76	0.708	12	0.33	
G335.23+0.034	335.23	0.03	5.02	5.02	0.00	12.2	12.2	5.28	0.42	17.9	17.9	0.010	11	0.37	
G335.36-0.023	335.36	-0.02	8.64	8.64	0.00	12.3	12.3	6.32	-0.2	31.0	31.0	0.005	18	0.38	G335.499-00.142 ⁽¹⁾
G335.49+0.005	335.49	0.00	3.64	3.64	0.00	12.3	12.3	7.43	0.06	12.7	12.7	0.021	11	1.75	G335.469+00.027 ⁽¹⁾
G335.57+0.214	335.57	0.21	5.33	3.51	129.	12.4	12.4	8.11	2.65	19.2	12.8	0.021	17	0.70	G335.794+00.153 ⁽¹⁾
G336.27+0.042	336.28	0.03	3.41	2.66	140.	2.85	2.85	8.52	0.09	2.78	2.20	0.623	12	1.00	
G336.22+0.207	336.22	0.20	4.23	4.23	0.00	12.5	12.5	13.5	2.58	15.6	15.6	0.014	11	4.50	G336.202+00.199 ⁽¹⁾
G335.32+0.405	335.32	0.40	5.66	5.66	0.00	11.4	11.4	5.39	4.51	18.9	18.9	0.019	22	0.37	G335.326+00.409 ⁽¹⁾
G334.92+0.491	334.90	0.49	4.92	2.90	89.9	11.2	11.2	2.77	5.37	16.0	9.51	0.029	14	0.00	
G332.33+0.230	332.33	0.23	3.64	3.64	0.00	3.65	3.65	11.6	0.83	3.78	3.78	0.267	12	0.71	
G334.51+0.818	334.51	0.81	3.29	2.70	80.9	4.95	4.95	4.20	3.61	4.83	3.82	0.275	16	4.33	G334.521+00.818 ⁽¹⁾
G335.56+0.665	335.55	0.65	9.87	5.28	16.9	10.6	10.6	6.73	6.50	30.5	16.4	0.019	30	0.00	
G331.92+0.394	331.92	0.39	7.65	3.32	170.	10.8	10.8	19.2	4.17	24.2	10.5	0.021	17	0.13	
G332.17+0.617	332.16	0.62	2.76	2.34	146.	2.85	2.85	11.1	1.65	2.20	1.96	0.882	12	0.20	
G333.25+0.819	333.22	0.80	6.59	2.63	26.9	2.95	2.95	9.20	2.13	5.67	2.27	0.369	15	0.15	
G336.24+1.064	336.24	1.06	4.01	4.01	0.00	5.05	5.05	9.00	4.41	5.78	5.78	0.133	14	0.16	
G336.95-0.183	336.93	-0.19	6.90	5.07	159.	10.8	10.8	16.6	-2.0	21.8	15.8	0.022	25	2.57	G336.919-00.188 ⁽¹⁾
G337.98-0.146	337.98	-0.14	6.00	6.00	0.00	12.1	12.1	20.6	-1.7	21.3	21.3	0.013	19	1.71	G337.986-00.144 ⁽¹⁾
G338.39+0.108	338.37	0.14	14.8	5.97	62.2	13.0	13.0	21.2	1.87	56.6	22.5	0.019	77	2.85	G338.350+00.221 ⁽¹⁾
G338.94-0.667	338.92	-0.67	8.04	3.27	12.9	3.65	3.65	11.1	-2.2	8.54	3.56	0.188	18	0.63	G339.169-00.698 ⁽¹⁾
G338.09+0.056	338.09	0.05	4.97	4.97	0.00	11.7	11.7	20.2	0.65	17.2	17.2	0.016	15	1.50	G338.075+00.016 ⁽¹⁾
G338.01+0.154	338.01	0.15	5.69	5.69	0.00	12.1	12.1	20.6	1.86	20.1	20.1	0.011	15	2.75	
G339.67-0.332	339.70	-0.35	4.52	3.23	89.9	13.3	13.3	14.9	-4.5	17.8	12.1	0.014	10	0.25	
G338.61+0.574	338.61	0.57	6.32	3.36	145.	11.5	11.5	19.1	6.24	21.1	11.2	0.022	17	0.21	G338.706+00.645 ⁽¹⁾
G338.78+0.508	338.76	0.50	4.26	1.46	20.1	11.4	11.4	18.5	5.56	14.2	5.59	0.039	10	1.00	G338.861+00.597 ⁽¹⁾
G338.46+0.619	338.47	0.62	5.08	1.84	179.	11.6	11.6	19.6	6.85	17.0	5.68	0.039	12	0.00	
G339.97+0.054	339.95	0.05	6.03	3.40	170.	13.2	13.2	12.8	0.75	23.3	12.9	0.014	14	0.27	G339.952+00.052 ⁽¹⁾
G340.43+0.100	340.43	0.10	4.21	2.18	36.0	12.7	12.7	8.41	1.27	15.8	7.62	0.028	11	0.83	
G340.62+0.088	340.62	0.08	10.2	10.2	0.00	12.5	12.5	6.73	1.11	37.5	37.5	0.005	23	0.27	
G339.84+0.299	339.84	0.29	8.06	8.06	0.00	13.5	13.5	13.9	3.99	31.7	31.7	0.008	28	1.00	G339.781+00.243 ⁽¹⁾
G340.16+0.104	340.18	0.10	4.69	2.35	31.9	13.0	13.0	10.8	1.32	18.0	9.01	0.019	10	0.42	
...

TABLE B.1: continued.

[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]	[14]	[15]	[16]
Cluster candidates name	l_{center} (deg)	b_{center} (deg)	a-axis (arcmin)	b-axis (arcmin)	PA (deg)	D_{peak} (kpc)	\bar{D} (kpc)	R (kpc)	Z (pc)	a-axis (pc)	b-axis (pc)	Density Clumps.pc ⁻²	Nb _{clumps}	Nb _{proto} /Nb _{pre}	H II regions
G340.70-0.260	340.71	-0.24	10.2	2.94	60.8	12.6	12.6	6.09	-3.0	37.8	10.6	0.018	23	0.09	G341.090-00.017 ⁽¹⁾
G340.97-0.033	340.97	-0.01	8.02	2.56	73.7	13.1	13.1	5.00	-0.1	30.7	10.1	0.017	17	0.30	
G341.44+0.268	341.44	0.26	5.55	1.76	138.	2.05	2.05	6.56	0.54	3.31	1.00	1.245	13	0.44	
G341.57-0.208	341.57	-0.20	3.28	3.28	0.00	12.7	12.7	6.09	-2.6	12.4	12.4	0.022	11	0.83	
G341.30+0.334	341.31	0.34	4.64	4.10	148.	2.05	2.05	6.48	0.68	2.74	2.45	0.710	15	1.14	G338.407-00.918 ⁽¹⁾
G338.47-0.707	338.48	-0.71	4.39	3.16	0.00	3.45	3.45	11.6	-2.2	4.45	3.15	0.362	16	0.23	
G338.28-0.761	338.31	-0.74	10.1	6.45	29.3	3.05	3.05	11.4	-2.0	8.99	5.76	0.220	36	0.09	
G337.50-0.544	337.50	-0.54	3.74	3.74	0.00	12.7	12.7	20.5	-6.6	13.9	13.9	0.021	13	0.85	
G339.26-0.775	339.25	-0.76	7.04	3.90	60.0	3.15	3.15	9.95	-2.1	6.43	3.60	0.288	21	0.40	G337.487-00.534 ⁽¹⁾
G340.86+0.182	340.86	0.18	3.92	2.84	11.2	12.4	12.4	4.90	2.25	14.2	10.5	0.023	11	1.75	G339.169-00.698 ⁽¹⁾
G340.66+0.210	340.66	0.21	2.92	2.39	177.	12.7	12.7	6.57	2.66	10.7	8.80	0.033	10	0.42	G340.910+00.167 ⁽¹⁾
G340.94-0.708	340.94	-0.70	3.85	3.85	0.00	3.45	3.45	5.17	-2.2	3.95	3.95	0.224	11	0.00	G340.600+00.342 ⁽¹⁾
G337.25-0.715	337.23	-0.72	4.75	3.79	129.	3.25	3.25	11.0	-2.1	4.48	3.54	0.279	14	0.07	
G340.61+0.352	340.61	0.35	4.00	4.00	0.00	12.5	12.5	6.81	4.32	14.3	14.3	0.024	16	1.00	
G340.04+0.929	340.05	0.91	4.38	3.20	144.	4.75	4.75	7.63	3.75	6.13	4.33	0.167	14	0.40	
G340.30+1.037	340.30	1.03	7.35	7.35	0.00	4.55	4.55	6.44	3.91	9.68	9.68	0.061	18	0.05	G340.070+00.927 ⁽¹⁾
G338.90+0.860	338.90	0.85	5.16	3.02	90.7	4.25	4.25	11.6	3.19	6.39	3.59	0.166	12	0.00	
G339.12+0.787	339.06	0.77	7.76	4.66	168.	4.15	4.15	11.1	2.89	9.39	5.54	0.170	28	0.16	
G338.90+0.564	338.91	0.58	4.23	3.57	23.1	4.45	4.45	11.7	2.44	5.53	4.60	0.236	19	1.37	
G339.52+0.967	339.52	0.96	4.08	1.22	126.	4.25	4.25	9.72	3.50	5.08	1.46	0.426	10	0.00	G338.861+00.597 ⁽¹⁾
G341.23+0.963	341.22	0.95	4.61	2.80	16.9	4.35	4.35	4.17	3.54	5.81	3.67	0.178	12	0.00	
G340.89+0.699	340.87	0.70	3.82	1.87	22.2	4.05	4.05	4.73	2.62	4.42	1.97	0.436	12	0.09	
G342.32+0.572	342.32	0.57	3.18	1.33	155.	2.05	2.05	7.93	1.11	1.87	0.70	2.641	11	2.66	
G342.08-0.398	342.08	-0.41	4.44	2.61	49.8	2.05	2.05	7.45	-0.8	2.64	1.58	0.759	10	0.66	G342.957+00.014 ⁽¹⁾
G342.43-0.325	342.43	-0.33	5.57	1.62	158.	2.05	2.05	8.15	-0.6	3.31	1.00	1.149	12	0.50	
G343.09-0.152	343.11	-0.14	4.72	2.79	166.	2.05	2.05	9.47	-0.2	2.83	1.73	1.102	17	0.30	
G343.39-0.340	343.42	-0.36	10.5	4.24	93.5	2.05	2.05	9.95	-0.7	6.29	2.55	0.456	23	0.76	
G343.72-0.196	343.73	-0.19	9.55	4.77	18.8	2.05	2.05	10.2	-0.3	5.70	2.83	0.571	29	1.90	G343.344-00.330 ⁽¹⁾
G343.91-0.375	343.94	-0.37	5.90	3.05	166.	2.05	2.05	10.4	-0.7	3.53	1.87	0.720	15	0.00	G343.636-00.150 ⁽¹⁾
G341.93-0.678	341.92	-0.69	5.13	4.51	164.	2.05	2.05	7.18	-1.3	3.08	2.64	0.934	24	0.14	G346.259+00.288 ⁽¹⁾
G344.16+0.213	344.15	0.21	3.57	1.85	89.9	11.4	11.4	19.9	2.42	11.8	5.59	0.048	10	0.66	
G344.58+0.078	344.58	0.07	5.79	2.66	144.	11.7	11.7	20.0	0.91	19.8	9.07	0.026	15	0.07	
G342.66-0.261	342.66	-0.27	6.24	3.68	42.9	2.05	2.05	8.63	-0.5	3.74	2.23	0.759	20	0.33	
G344.86+0.445	344.84	0.42	3.79	2.03	118.	12.1	12.1	19.8	5.03	13.2	7.26	0.039	12	0.09	G346.259+00.288 ⁽¹⁾
G345.12-0.002	345.13	-0.03	6.28	4.84	67.3	13.9	13.9	20.5	-0.5	25.4	19.8	0.008	14	1.33	
G346.25+0.285	346.25	0.26	5.72	4.54	175.	13.0	13.0	12.5	3.47	21.6	17.4	0.016	19	0.35	
...

TABLE B.1: continued.

[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]	[14]	[15]	[16]
Cluster candidates name	l_{center} (deg)	b_{center} (deg)	a-axis (arcmin)	b-axis (arcmin)	PA (deg)	D_{peak} (kpc)	\bar{D} (kpc)	R (kpc)	Z (pc)	a-axis (pc)	b-axis (pc)	Density Clumps.pc ⁻²	Nb _{clumps}	Nb _{proto} /Nb _{pre}	H II regions
G346.53-0.137	346.51	-0.16	4.89	3.43	161.	10.7	10.7	8.76	-1.7	15.3	10.4	0.023	12	0.20	
G345.49+0.387	345.49	0.38	4.20	4.20	0.00	2.15	2.15	9.48	0.81	2.67	2.67	0.622	14	1.80	G345.486+00.399 ⁽¹⁾
G345.20-0.746	345.20	-0.74	4.00	4.00	0.00	2.65	2.65	10.3	-1.7	3.03	3.03	0.622	18	0.50	G345.094-00.779 ⁽¹⁾
G345.43-0.718	345.43	-0.71	4.41	2.13	62.3	2.65	2.65	9.89	-1.7	3.42	1.58	0.704	12	0.00	
G344.57-0.658	344.59	-0.67	8.41	4.94	176.	2.65	2.65	11.0	-1.6	6.46	3.77	0.326	25	0.25	G344.472-00.557 ⁽¹⁾
G346.40+0.642	346.40	0.64	5.45	4.13	55.9	1.95	1.95	7.67	1.17	3.08	2.33	0.663	15	0.07	
G346.11-0.346	346.11	-0.34	4.98	1.83	12.8	13.1	13.1	13.7	-4.4	19.2	6.42	0.033	13	0.62	G346.095-00.361 ⁽¹⁾
G344.22-0.646	344.21	-0.65	6.18	4.56	52.9	2.65	2.65	11.1	-1.6	4.75	3.54	0.434	23	0.64	344.2-00.6 ⁽²⁾
G342.90+0.974	342.89	0.97	6.15	1.47	69.2	4.85	4.85	10.4	4.01	8.70	2.36	0.169	11	0.10	
G346.42+0.948	346.42	0.94	5.28	3.31	112.	2.15	2.15	7.57	1.74	3.31	2.09	0.730	16	0.14	
G344.31+0.476	344.31	0.47	4.83	2.15	21.4	11.5	11.5	20.0	5.29	16.4	6.90	0.033	12	0.09	
G344.10+0.485	344.10	0.48	4.45	4.45	0.00	11.2	11.2	19.6	5.24	14.5	14.5	0.015	10	0.00	
G345.92+0.589	345.93	0.57	4.79	2.07	99.7	2.15	2.15	8.62	1.17	2.96	1.28	1.000	12	0.20	
G345.86+0.791	345.87	0.78	5.72	1.56	35.2	1.55	1.55	8.61	1.09	2.56	0.75	2.130	13	0.18	
G346.04+0.870	346.04	0.87	5.22	5.22	0.00	1.85	1.85	8.37	1.41	2.78	2.78	0.492	12	0.00	
G345.58+0.758	345.58	0.75	7.31	7.31	0.00	1.85	1.85	9.17	1.27	3.93	3.93	0.492	24	0.20	
G345.51+1.074	345.51	1.09	6.40	3.52	45.9	1.55	1.55	9.14	1.37	2.88	1.60	1.169	17	0.00	G345.235+01.408 ⁽¹⁾
G345.25+1.040	345.25	1.03	10.4	5.39	160.	1.95	1.95	9.74	1.67	5.94	3.08	0.815	47	1.13	G345.202+01.027 ⁽¹⁾
G344.96+1.111	344.96	1.11	7.57	5.71	100.	2.05	2.05	10.1	1.83	4.53	3.39	0.703	34	0.47	G345.235+01.408 ⁽¹⁾
G345.27-0.557	345.27	-0.55	3.40	3.40	0.00	13.7	13.7	19.7	-7.2	13.4	13.4	0.021	12	0.33	G345.247-00.597 ⁽¹⁾
G342.25-0.531	342.23	-0.52	4.19	2.80	150.	2.05	2.05	7.75	-1.0	2.55	1.73	1.007	14	0.00	
G347.11+0.256	347.11	0.25	3.92	3.92	0.00	10.8	10.8	3.84	2.75	12.4	12.4	0.024	12	1.40	G347.087+00.091 ⁽¹⁾
G347.21-0.215	347.21	-0.22	5.16	1.65	112.	11.4	11.4	3.67	-2.4	17.2	5.59	0.036	11	0.10	
G347.25+0.036	347.29	0.01	7.41	5.48	173.	11.0	11.0	2.87	0.19	23.8	17.4	0.014	19	0.46	G347.087+00.091 ⁽¹⁾
G347.38+0.250	347.38	0.25	4.40	4.40	0.00	10.5	10.5	2.10	2.61	13.6	13.6	0.018	11	10.0	G347.495+00.171 ⁽¹⁾
G347.57+0.049	347.57	0.04	4.64	4.64	0.00	10.4	10.4	2.37	0.51	13.9	13.9	0.022	14	2.50	G347.495+00.171 ⁽¹⁾
G347.62+0.132	347.63	0.11	2.57	1.42	89.9	10.3	10.3	2.63	1.21	7.99	3.57	0.122	11	2.66	G347.495+00.171 ⁽¹⁾
G347.93-0.431	347.93	-0.43	6.62	6.62	0.00	10.2	10.2	4.98	-4.2	19.7	19.7	0.019	24	0.84	G348.000-00.496 ⁽¹⁾
G348.99+0.148	348.99	0.16	8.37	4.96	47.5	11.0	11.0	13.9	1.85	26.9	16.1	0.026	36	0.80	G348.987+00.168 ⁽¹⁾
G349.78-0.002	349.78	-0.00	2.84	1.76	1.97	9.95	9.95	17.0	-0.0	8.41	4.85	0.085	11	1.20	
G349.85-0.555	349.86	-0.61	10.4	5.43	53.4	2.25	2.25	10.3	-1.2	6.86	3.55	0.495	38	2.80	G349.418-00.649 ⁽¹⁾
G348.34+0.463	348.34	0.46	8.60	4.26	110.	1.45	1.45	7.70	0.64	3.64	1.80	1.499	31	1.21	G348.261+00.485 ⁽¹⁾
G348.17+0.521	348.17	0.57	9.35	3.43	99.8	1.35	1.35	7.56	0.73	3.67	1.31	1.513	23	0.76	G348.261+00.485 ⁽¹⁾
G348.51-0.610	348.52	-0.60	6.38	3.78	81.1	1.45	1.45	7.93	-0.8	2.69	1.58	1.864	25	4.00	G348.691-00.826 ⁽¹⁾
G348.79-0.575	348.79	-0.57	3.25	1.38	109.	12.5	12.5	13.5	-6.8	12.2	4.33	0.059	10	1.50	G348.691-00.826 ⁽¹⁾
...

TABLE B.1: continued.

[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]	[14]	[15]	[16]
Cluster candidates name	l_{center} (deg)	b_{center} (deg)	a-axis (arcmin)	b-axis (arcmin)	PA (deg)	D_{peak} (kpc)	\bar{D} (kpc)	R (kpc)	Z (pc)	a-axis (pc)	b-axis (pc)	Density $Clumps.pc^{-2}$	Nb _{clumps}	Nb _{proto} /Nb _{pre}	H II regions
G347.77+0.237	347.77	0.23	4.33	1.81	123.	11.0	11.0	4.13	2.60	13.7	5.39	0.047	11	0.57	
G348.91-0.793	348.91	-0.79	7.03	3.56	122.	2.25	2.25	8.58	-1.6	4.59	2.33	0.683	23	2.83	G348.691-00.826 ⁽¹⁾
G015.45+0.194	15.45	0.19	4.00	4.00	0.00	4.25	4.25	12.6	0.81	4.86	4.86	0.161	12	1.00	015.4+00.2 ⁽²⁾
G015.15+0.186	15.15	0.18	7.36	7.36	0.00	13.2	13.2	20.9	2.46	28.5	28.5	0.008	23	1.55	G015.212+00.167 ⁽¹⁾
G019.64-0.120	19.66	-0.11	6.80	4.54	11.8	11.7	11.7	8.55	-1.3	23.3	15.7	0.029	34	3.25	G019.629-00.095 ⁽¹⁾
G018.06+0.004	18.05	0.02	4.42	2.04	131.	12.1	12.1	8.63	0.26	15.6	7.26	0.039	14	1.33	G018.068-00.013 ⁽¹⁾
G018.77+0.263	18.77	0.26	3.40	3.40	0.00	14.2	14.2	5.81	3.70	13.9	13.9	0.016	10	2.33	G018.741+00.250 ⁽¹⁾
G018.61+0.318	18.61	0.31	3.13	3.13	0.00	14.1	14.1	6.19	4.42	12.9	12.9	0.032	17	0.88	G018.594+00.321 ⁽¹⁾
G018.97-0.775	18.98	-0.77	7.52	2.32	179.	4.55	4.55	4.04	-3.1	9.93	3.14	0.183	18	0.20	
G019.25+0.375	19.25	0.37	4.86	4.86	0.00	14.1	14.1	7.18	5.18	20.1	20.1	0.016	21	0.90	
G019.79-0.432	19.79	-0.43	7.05	4.80	72.9	11.2	11.2	9.32	-4.7	22.9	15.5	0.020	23	0.27	G019.783-00.628 ⁽¹⁾
G019.88-0.807	19.88	-0.80	7.37	7.37	0.00	3.45	3.45	7.37	-2.4	7.43	7.43	0.143	25	0.38	G019.783-00.628 ⁽¹⁾
G020.01-0.986	20.01	-0.99	4.30	2.48	0.00	3.25	3.25	7.80	-2.7	4.04	2.24	0.385	11	0.00	
G017.82+0.075	17.83	0.08	4.33	2.93	0.00	9.25	9.25	8.66	0.79	11.5	7.82	0.035	10	1.50	G017.818+00.067 ⁽¹⁾
G017.56+0.072	17.56	0.07	4.57	3.42	23.2	13.9	13.9	14.1	1.00	18.6	13.6	0.020	16	1.66	
G017.27-0.140	17.27	-0.14	6.98	6.98	0.00	12.3	12.3	15.0	-1.7	25.2	25.2	0.009	19	1.37	
G017.44-0.229	17.41	-0.25	6.05	2.79	132.	12.0	12.0	13.7	-3.0	21.2	10.1	0.022	15	0.36	
G017.23+0.144	17.23	0.17	5.43	3.10	96.3	12.2	12.2	15.2	2.14	19.3	11.1	0.020	14	0.27	
G018.33-0.833	18.33	-0.83	6.87	6.87	0.00	3.85	3.85	5.48	-2.8	7.75	7.75	0.100	19	0.46	
G017.65+0.215	17.65	0.21	2.87	2.87	0.00	2.35	2.35	7.94	0.50	1.98	1.98	0.886	11	1.75	
G017.41+0.743	17.41	0.74	5.77	5.77	0.00	1.95	1.95	8.46	1.31	3.29	3.29	0.555	19	1.71	
G021.45-0.657	21.45	-0.64	2.66	2.40	0.00	10.9	10.9	18.7	-6.5	8.45	7.56	0.049	10	0.11	G021.424-00.661 ⁽¹⁾
G020.50-0.848	20.49	-0.86	4.85	2.91	6.75	4.35	4.35	9.84	-3.3	6.19	3.67	0.195	14	0.00	
G021.87-0.197	21.89	-0.17	8.75	7.78	129.	10.7	10.7	19.2	-1.8	27.2	24.3	0.018	38	1.11	G021.870-00.162 ⁽¹⁾
G022.24-0.267	22.24	-0.26	4.49	4.49	0.00	11.6	11.6	19.9	-3.0	15.0	15.0	0.023	17	0.88	G022.090-00.117 ⁽¹⁾
G023.41+0.509	23.42	0.53	5.95	3.35	12.2	9.45	9.45	13.5	4.80	16.3	9.22	0.027	13	0.00	
G023.88+0.404	23.89	0.39	5.48	0.97	178.	9.25	9.25	10.3	3.59	14.6	3.19	0.081	12	0.50	G023.689+00.377 ⁽¹⁾
G023.85-0.114	23.85	-0.11	4.36	4.36	0.00	10.7	10.7	11.6	-1.2	13.8	13.8	0.028	17	1.42	G023.847-00.128 ⁽¹⁾
G024.68+0.450	24.68	0.45	6.11	2.95	151.	9.05	9.05	3.92	3.94	16.2	7.65	0.043	17	0.13	G024.919+00.294 ⁽¹⁾
G024.91+0.211	24.91	0.21	4.29	1.72	39.4	9.65	9.65	2.31	2.02	12.0	4.71	0.061	11	4.50	G024.897+00.211 ⁽¹⁾
G024.24+0.227	24.24	0.25	4.07	2.62	76.3	8.95	8.95	7.48	2.21	10.7	6.90	0.047	11	0.37	G024.185+00.211 ⁽¹⁾
G025.11-0.456	25.11	-0.45	3.21	3.21	0.00	3.85	3.85	4.65	-1.6	3.51	3.51	0.257	10	0.11	
G024.13+0.436	24.16	0.43	6.78	2.40	179.	9.85	9.85	8.58	4.16	19.5	6.80	0.040	17	0.54	G024.139+00.432 ⁽¹⁾
G023.95+0.493	23.93	0.50	5.82	2.11	23.9	9.95	9.95	10.4	4.82	16.8	5.95	0.057	18	0.38	
G026.54+0.476	26.55	0.47	4.36	3.18	151.	9.95	9.95	12.1	4.55	12.8	9.08	0.040	15	0.36	
...

TABLE B.1: continued.

[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]	[14]	[15]	[16]
Cluster candidates name	l_{center} (deg)	b_{center} (deg)	a-axis (arcmin)	b-axis (arcmin)	PA (deg)	D_{peak} (kpc)	\bar{D} (kpc)	R (kpc)	Z (pc)	a-axis (pc)	b-axis (pc)	Density Clumps.pc ⁻²	Nb _{clumps}	Nb _{proto} /Nb _{pre}	H II regions
G022.95+0.001	22.95	0.00	3.69	1.85	25.5	10.6	10.6	17.0	0.01	11.6	5.20	0.052	10	0.66	G023.689-00.205 ⁽¹⁾
G023.70-0.216	23.69	-0.23	4.88	1.33	138.	11.0	11.0	13.0	-2.6	15.7	3.81	0.053	10	1.00	
G022.61-0.647	22.60	-0.66	3.74	2.86	0.79	11.0	11.0	18.6	-6.7	12.0	9.34	0.033	12	0.09	
G026.27-0.065	26.25	-0.10	6.77	4.59	103.	9.35	9.35	9.53	-0.9	18.5	12.5	0.026	19	1.11	G026.090-00.079 ⁽¹⁾
G026.10-0.064	26.08	-0.07	3.58	3.22	37.4	9.15	9.15	8.14	-0.6	9.47	8.35	0.048	12	2.00	
G024.59-0.465	24.61	-0.47	6.53	3.24	130.	4.05	4.05	5.37	-1.8	7.65	3.69	0.224	20	0.17	
G023.03-0.504	23.01	-0.49	3.55	2.49	20.7	10.8	10.8	16.9	-5.1	11.2	7.49	0.041	11	0.10	G022.585-01.081 ⁽¹⁾
G026.35+0.816	26.35	0.81	4.64	2.17	63.9	3.35	3.35	8.00	2.43	4.47	2.00	0.390	11	0.10	
G024.97+0.603	24.96	0.59	4.07	1.29	14.2	8.55	8.55	1.44	4.78	10.2	2.95	0.105	10	0.25	
G025.74+1.063	25.74	1.06	3.93	1.63	63.1	3.15	3.15	6.18	2.75	3.60	1.53	0.573	10	0.42	G029.065+00.387 ⁽¹⁾
G023.71-0.804	23.71	-0.80	7.59	7.59	0.00	4.25	4.25	8.90	-3.0	9.39	9.39	0.057	16	0.06	
G022.86+0.831	22.86	0.83	3.75	3.75	0.00	2.95	2.95	10.6	2.18	3.22	3.22	0.306	10	0.66	
G022.43-0.906	22.43	-0.90	2.87	2.87	0.00	4.45	4.45	12.6	-3.5	3.76	3.76	0.224	10	0.66	G028.590-00.352 ⁽¹⁾
G027.60+0.077	27.60	0.07	6.58	6.58	0.00	9.95	9.95	17.4	0.77	19.1	19.1	0.025	29	1.23	
G029.02+0.303	29.02	0.30	5.48	3.49	162.	9.05	9.05	16.3	2.70	14.3	9.37	0.030	13	0.30	
G029.38+0.150	29.36	0.15	5.55	2.51	154.	9.95	9.95	15.7	1.54	16.1	6.87	0.034	12	0.00	G030.067+00.097 ⁽¹⁾
G028.85+0.340	28.84	0.34	3.70	1.74	77.6	9.65	9.65	17.4	3.27	10.5	4.71	0.076	12	2.00	
G028.59-0.359	28.57	-0.36	2.40	1.85	89.9	5.35	5.35	13.7	-1.8	3.69	2.61	0.395	12	3.00	
G029.63-0.221	29.62	-0.22	5.87	4.27	58.5	8.75	8.75	13.4	-1.9	14.8	10.8	0.047	24	0.71	G030.546-00.178 ⁽¹⁾
G029.39-0.537	29.39	-0.53	4.09	4.09	0.00	10.7	10.7	16.3	-5.5	12.8	12.8	0.021	11	0.37	
G029.61-0.597	29.61	-0.58	4.71	3.48	124.	10.0	10.0	14.5	-5.5	13.8	10.4	0.033	15	0.15	
G029.15+0.456	29.16	0.45	2.96	1.78	137.	13.0	13.0	19.5	5.68	11.0	6.37	0.049	11	4.50	G031.016-00.039 ⁽¹⁾
G030.09+0.089	30.09	0.08	5.27	2.09	15.4	7.95	7.95	10.0	0.70	12.2	4.75	0.081	15	0.66	
G030.29+0.087	30.29	0.08	7.00	7.00	0.00	8.75	8.75	9.16	0.76	17.8	17.8	0.024	25	0.78	
G030.62-0.124	30.63	-0.13	3.13	1.36	27.6	8.85	8.85	6.61	-1.2	8.08	3.05	0.141	11	2.66	G031.047+00.355 ⁽¹⁾
G031.10+0.038	31.10	0.03	5.03	5.03	0.00	11.8	11.8	4.60	0.45	17.3	17.3	0.023	22	2.14	
G031.04+0.355	31.05	0.37	3.32	2.64	22.7	9.65	9.65	3.45	3.53	9.42	7.45	0.045	10	2.33	
G030.62+0.629	30.62	0.63	5.28	3.96	167.	9.25	9.25	6.88	5.48	14.2	10.5	0.033	16	0.06	G033.306-00.150 ⁽¹⁾
G030.36-0.833	30.36	-0.83	5.76	5.76	0.00	4.75	4.75	7.41	-3.5	8.03	8.03	0.059	12	0.09	
G026.96+0.441	26.96	0.43	4.02	1.15	154.	9.05	9.05	13.9	3.79	10.8	3.12	0.112	12	0.33	
G030.07-1.088	30.07	-1.07	5.90	2.72	39.2	4.35	4.35	8.60	-3.8	7.50	3.35	0.151	12	0.00	G034.130-00.174 ⁽¹⁾
G027.45-0.866	27.45	-0.86	3.14	2.33	29.3	5.25	5.25	12.6	-3.9	4.79	3.62	0.219	12	0.09	
G031.06+0.767	31.04	0.75	2.80	1.88	31.8	3.35	3.35	5.51	2.28	2.83	2.00	0.560	10	0.42	
G033.45-0.104	33.45	-0.10	5.92	5.92	0.00	8.45	8.45	14.4	-0.8	14.5	14.5	0.020	14	0.40	G035.040-00.500 ⁽¹⁾
G034.30-0.189	34.30	-0.18	8.21	1.91	1.86	10.6	10.6	18.9	-1.9	25.4	6.36	0.031	16	0.23	
G035.04-0.470	35.05	-0.48	5.96	3.41	45.2	10.6	10.6	18.5	-4.9	18.3	10.4	0.036	22	0.46	
...

TABLE B.1: continued.

[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]	[14]	[15]	[16]
Cluster candidates name	l_{center} (deg)	b_{center} (deg)	a-axis (arcmin)	b-axis (arcmin)	PA (deg)	D_{peak} (kpc)	\bar{D} (kpc)	R (kpc)	Z (pc)	a-axis (pc)	b-axis (pc)	Density Clumps.pc ⁻²	Nb _{clumps}	Nb _{proto} /Nb _{pre}	H II regions
G034.56-0.410	34.56	-0.41	4.86	3.20	144.	10.3	10.3	18.8	-4.2	14.7	9.45	0.029	13	0.85	
G035.98+0.076	35.96	0.06	5.00	1.67	171.	11.2	11.2	15.1	0.77	16.4	5.49	0.045	13	0.30	G036.106+00.145 ⁽¹⁾
G035.00+0.342	35.00	0.34	4.52	4.52	0.00	3.35	3.35	11.6	1.12	4.47	4.47	0.206	13	1.16	G034.997+00.330 ⁽¹⁾
G034.85+0.447	34.86	0.44	4.11	1.99	67.6	3.55	3.55	11.9	1.51	4.24	2.12	0.353	10	0.25	
G034.72+0.514	34.75	0.49	6.39	4.97	36.7	3.45	3.45	11.9	1.63	6.41	5.05	0.127	13	0.18	
G034.73-0.734	34.71	-0.72	8.50	3.70	80.3	3.05	3.05	11.5	-2.0	7.52	3.33	0.355	28	0.86	G034.757-00.669 ⁽¹⁾
G036.71-0.624	36.69	-0.64	7.46	3.90	161.	9.75	9.75	8.88	-5.8	21.2	11.1	0.017	13	0.08	G036.620-00.557 ⁽¹⁾
G036.05-0.412	36.09	-0.41	6.81	6.03	125.	9.85	9.85	13.2	-3.9	19.5	17.3	0.023	25	0.13	
G036.10+0.574	36.10	0.57	5.98	2.45	66.3	8.75	8.75	12.3	4.75	15.1	6.04	0.045	13	0.44	
G034.33-0.682	34.33	-0.68	4.85	2.27	123.	0.85	0.85	9.33	-0.5	1.21	0.58	4.481	10	0.25	
G035.60-0.546	35.58	-0.56	4.48	1.30	137.	10.1	10.1	16.2	-5.4	13.1	3.50	0.069	10	4.00	G035.559-00.824 ⁽¹⁾
G035.32-0.917	35.35	-0.94	4.02	2.86	39.3	2.55	2.55	10.4	-2.0	3.04	2.15	0.580	12	0.50	G035.559-00.824 ⁽¹⁾
G035.13-0.753	35.13	-0.75	5.51	3.90	94.4	2.25	2.25	10.4	-1.5	3.64	2.57	0.745	22	1.20	G035.126-00.755 ⁽¹⁾
G036.41-0.762	36.41	-0.76	9.12	2.65	178.	9.85	9.85	11.0	-6.8	26.1	7.60	0.032	20	0.11	G036.262-00.721 ⁽¹⁾
G035.33+1.054	35.33	1.05	6.07	6.07	0.00	5.15	5.15	12.6	4.47	9.06	9.06	0.065	17	0.70	G035.187+00.892 ⁽¹⁾
G037.16-0.417	37.15	-0.40	3.18	2.60	66.0	11.2	11.2	5.93	-4.4	10.2	8.68	0.035	10	1.00	G037.175-00.440 ⁽¹⁾
G036.90+0.144	36.89	0.17	8.68	4.14	142.	9.75	9.75	7.24	1.71	24.7	11.6	0.020	19	0.26	
G038.27-0.209	38.27	-0.20	11.2	11.2	0.00	9.15	9.15	5.07	-1.9	29.9	29.9	0.018	51	0.41	G038.365-00.062 ⁽¹⁾
G037.87-0.391	37.85	-0.41	9.74	7.38	17.6	9.15	9.15	1.55	-3.6	25.8	19.7	0.021	35	1.33	G037.872-00.399 ⁽¹⁾
G039.41-0.361	39.38	-0.35	16.4	5.86	5.61	8.75	8.75	12.9	-3.0	41.9	14.8	0.029	57	1.19	G039.294-00.311 ⁽¹⁾
G039.88-0.374	39.91	-0.41	7.31	5.99	89.9	8.95	8.95	15.6	-3.5	19.0	15.4	0.019	18	0.63	G039.883-00.346 ⁽¹⁾
G039.90-0.176	39.92	-0.14	8.52	6.11	133.	9.15	9.15	15.8	-1.3	22.7	16.4	0.026	31	0.82	G039.856-00.184 ⁽¹⁾
G039.61-0.196	39.58	-0.16	7.47	4.85	33.1	9.15	9.15	14.2	-1.5	19.9	13.0	0.024	20	1.50	G039.434-00.035 ⁽¹⁾
G040.80-0.198	40.80	-0.19	7.49	4.10	11.4	11.0	11.0	19.5	-2.1	24.1	13.2	0.023	24	0.71	G040.795-00.216 ⁽¹⁾
G040.99-0.557	40.99	-0.55	5.64	5.64	0.00	8.85	8.85	17.2	-4.6	14.6	14.6	0.022	15	0.15	G040.965-00.621 ⁽¹⁾
G041.51-0.178	41.55	-0.21	7.99	5.08	33.1	8.15	8.15	15.5	-1.7	18.8	11.9	0.026	19	0.58	G041.516-00.142 ⁽¹⁾
G040.75+0.063	40.72	0.06	5.92	4.28	90.0	11.3	11.3	19.8	0.79	19.5	14.1	0.013	12	0.33	
G039.19+0.247	39.19	0.24	4.77	1.75	103.	11.1	11.1	13.5	2.72	15.3	5.44	0.041	11	0.83	G039.196+00.224 ⁽¹⁾
G040.70-0.648	40.63	-0.62	10.0	4.26	22.0	8.05	8.05	16.4	-4.6	23.4	10.0	0.021	16	0.33	G040.620-00.651 ⁽¹⁾
G038.25+0.822	38.25	0.82	7.89	7.89	0.00	3.05	3.05	6.12	2.23	6.98	6.98	0.189	29	0.07	G037.989+00.871 ⁽¹⁾
G039.92-0.901	39.94	-0.90	5.50	1.09	13.7	3.55	3.55	11.0	-2.7	5.61	1.22	0.508	11	0.37	
G039.45-0.976	39.45	-0.97	5.78	5.78	0.00	3.55	3.55	9.80	-2.9	6.00	6.00	0.132	15	0.36	
G038.35-0.910	38.35	-0.90	9.55	1.94	87.1	1.05	1.05	7.69	-0.8	2.92	0.62	2.775	16	0.60	
G041.77-0.246	41.78	-0.25	5.43	3.66	114.	7.45	7.45	14.2	-1.8	11.7	8.13	0.053	16	0.45	
G041.95-0.411	41.95	-0.41	8.58	8.58	0.00	7.55	7.55	13.6	-3.0	18.7	18.7	0.018	21	0.16	G042.006-00.500 ⁽¹⁾
...

TABLE B.1: continued.

[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]	[14]	[15]	[16]
Cluster candidates name	l_{center} (deg)	b_{center} (deg)	a-axis (arcmin)	b-axis (arcmin)	PA (deg)	D_{peak} (kpc)	\bar{D} (kpc)	R (kpc)	Z (pc)	a-axis (pc)	b-axis (pc)	Density Clumps.pc ⁻²	Nb _{clumps}	Nb _{proto} /Nb _{pre}	H II regions
G042.12-0.612	42.12	-0.61	4.52	4.52	0.00	5.05	5.05	11.0	-2.9	6.75	6.75	0.153	22	1.20	G042.006-00.500 ⁽¹⁾
G042.61-0.373	42.60	-0.36	6.13	1.68	11.5	7.95	7.95	10.4	-2.8	14.2	3.88	0.063	11	0.22	
G044.10+0.119	44.11	0.12	5.56	3.71	104.	7.45	7.45	1.48	0.94	12.0	8.13	0.055	17	0.54	G044.224+00.085 ⁽¹⁾
G044.21-0.193	44.19	-0.20	8.73	2.37	109.	7.65	7.65	1.91	-1.5	19.4	5.28	0.037	12	0.20	
G044.58-0.308	44.54	-0.31	5.24	4.23	0.00	7.05	7.05	4.56	-2.1	10.8	8.77	0.036	11	0.10	G044.552-00.239 ⁽¹⁾
G044.35-0.219	44.35	-0.21	4.43	4.43	0.00	6.85	6.85	3.26	-1.4	8.84	8.84	0.040	10	0.00	G044.379-00.327 ⁽¹⁾
G044.39-0.403	44.39	-0.40	5.26	5.26	0.00	7.35	7.35	3.41	-2.8	11.3	11.3	0.029	12	0.00	G044.379-00.327 ⁽¹⁾
G044.29+0.041	44.29	0.04	4.60	4.60	0.00	8.05	8.05	2.57	0.33	10.7	10.7	0.038	14	1.00	G044.224+00.085 ⁽¹⁾
G045.24-0.581	45.22	-0.58	5.11	3.36	17.5	7.95	7.95	9.58	-4.3	11.9	7.76	0.065	19	0.18	G045.191-00.486 ⁽¹⁾
G045.39-0.563	45.39	-0.55	3.15	2.76	0.00	7.75	7.75	10.5	-4.0	7.07	5.98	0.082	11	0.83	
G045.46-0.689	45.46	-0.68	11.8	11.8	0.00	7.05	7.05	10.5	-4.4	24.3	24.3	0.020	38	0.40	G045.391-00.725 ⁽¹⁾
G045.76-0.592	45.74	-0.63	9.01	4.06	113.	6.75	6.75	11.8	-4.0	17.7	8.07	0.035	16	0.33	
G045.94-0.530	45.94	-0.54	4.68	3.10	152.	6.95	6.95	12.8	-3.5	9.59	6.34	0.062	12	0.71	
G043.01-0.573	42.98	-0.56	7.29	3.29	111.	8.35	8.35	8.02	-4.4	17.7	8.15	0.037	17	0.13	
G043.45+0.541	43.48	0.53	7.28	2.58	27.3	11.8	11.8	5.97	6.07	25.2	9.14	0.023	17	0.30	G043.432+00.516 ⁽¹⁾
G045.52-0.297	45.52	-0.29	5.04	5.04	0.00	7.85	7.85	11.3	-2.2	11.4	11.4	0.031	13	0.44	
G046.36-0.217	46.33	-0.22	16.3	6.39	156.	7.45	7.45	14.7	-1.6	35.5	13.8	0.034	54	0.58	G046.495-00.241 ⁽¹⁾
G045.82-0.312	45.83	-0.29	10.6	6.56	23.1	7.15	7.15	12.5	-2.1	22.0	13.7	0.036	35	0.59	G045.825-00.291 ⁽¹⁾
G043.84+0.408	43.85	0.42	6.18	2.42	157.	11.8	11.8	3.59	4.83	21.2	8.18	0.023	13	1.16	G043.818+00.395 ⁽¹⁾
G046.51+0.012	46.51	0.01	3.71	3.71	0.00	8.25	8.25	15.9	0.10	9.00	9.00	0.039	10	0.42	
G046.11+0.340	46.11	0.34	5.26	4.03	59.0	6.75	6.75	13.3	2.28	10.4	8.07	0.049	13	0.85	
G046.39-0.444	46.39	-0.44	3.67	2.06	141.	7.45	7.45	14.9	-3.2	8.13	4.45	0.105	12	0.20	
G046.56+0.307	46.56	0.30	4.83	4.83	0.00	11.3	11.3	19.0	3.43	16.1	16.1	0.021	18	0.20	G046.792+00.284 ⁽¹⁾
G042.41-0.708	42.41	-0.70	5.81	5.81	0.00	8.65	8.65	12.1	-5.6	14.6	14.6	0.014	10	0.00	
G043.86-0.865	43.86	-0.86	5.07	4.16	127.	8.25	8.25	1.02	-6.2	12.0	9.86	0.053	20	0.25	G043.935-00.832 ⁽¹⁾
G047.09+0.309	47.12	0.39	17.5	8.57	153.	6.55	6.55	15.0	2.50	33.4	16.3	0.030	52	0.44	G046.792+00.284 ⁽¹⁾
G047.03-0.373	47.05	-0.39	5.56	3.70	44.5	7.95	7.95	16.4	-3.0	12.8	8.68	0.039	14	0.40	G047.046-00.400 ⁽¹⁾
G048.63+0.246	48.65	0.28	5.55	3.09	105.	10.6	10.6	13.9	2.96	17.2	9.72	0.041	22	1.00	G048.630+00.230 ⁽¹⁾
G049.72+0.036	49.73	0.03	3.72	2.19	97.1	6.95	6.95	4.28	0.24	7.58	4.15	0.111	11	0.10	G049.694-00.042 ⁽¹⁾
G050.00-0.124	49.99	-0.13	10.0	5.65	157.	7.15	7.15	2.48	-0.9	20.9	11.8	0.032	25	0.47	G049.728+00.129 ⁽¹⁾
G050.03+0.557	50.01	0.52	7.19	6.07	10.5	11.2	11.2	3.65	5.65	23.6	19.8	0.017	25	0.56	G049.728+00.129 ⁽¹⁾
G049.82+0.453	49.79	0.42	3.71	3.00	89.9	10.6	10.6	4.89	4.39	11.6	9.00	0.030	10	0.66	G049.728+00.129 ⁽¹⁾
G048.58+0.468	48.55	0.45	5.08	3.49	48.1	11.0	11.0	14.8	4.82	16.1	11.4	0.022	13	0.62	G048.542+00.464 ⁽¹⁾
G050.35-0.433	50.34	-0.44	8.89	5.16	155.	3.25	3.25	5.26	-1.4	8.39	4.89	0.248	32	1.00	G050.317-00.421 ⁽¹⁾
G050.99+0.212	51.02	0.17	7.32	3.59	30.6	6.35	6.35	5.86	1.13	13.5	6.57	0.068	19	3.75	G051.010+00.060 ⁽¹⁾
...

TABLE B.1: continued.

[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]	[14]	[15]	[16]
Cluster candidates name	l_{center} (deg)	b_{center} (deg)	a-axis (arcmin)	b-axis (arcmin)	PA (deg)	D_{peak} (kpc)	\bar{D} (kpc)	R (kpc)	Z (pc)	a-axis (pc)	b-axis (pc)	Density Clumps.pc ⁻²	Nb _{clumps}	Nb _{proto} /Nb _{pre}	H II regions
G051.58-0.364	51.57	-0.37	3.97	1.10	163.	6.75	6.75	9.38	-2.4	7.72	2.33	0.176	10	1.50	G051.457-00.286 ⁽¹⁾
G051.21-0.777	51.20	-0.79	7.16	4.26	11.0	6.45	6.45	7.03	-4.5	13.3	8.02	0.041	14	0.40	G051.203-00.739 ⁽¹⁾
G050.92-0.579	50.92	-0.57	4.24	4.24	0.00	3.25	3.25	6.26	-1.7	4.04	4.04	0.233	12	0.33	
G051.22+0.617	51.19	0.55	8.45	4.18	80.6	10.4	10.4	8.70	5.48	25.7	12.4	0.012	13	0.44	
G051.65+0.693	51.65	0.69	15.1	15.1	0.00	10.6	10.6	12.3	6.80	46.9	46.9	0.007	49	0.53	G051.678+00.718 ⁽¹⁾
G049.50-0.897	49.51	-0.90	6.44	4.45	160.	7.75	7.75	6.01	-6.1	14.4	10.0	0.059	27	0.42	G049.048-00.886 ⁽¹⁾
G049.28-0.971	49.27	-0.99	3.52	2.83	51.7	7.95	7.95	7.83	-6.6	8.23	6.72	0.068	12	0.09	G049.048-00.886 ⁽¹⁾
G049.03-0.956	49.03	-0.95	4.79	3.50	34.0	7.75	7.75	9.42	-6.3	10.7	8.02	0.037	10	0.00	G049.048-00.886 ⁽¹⁾
G050.09+0.852	50.09	0.85	4.34	2.61	99.5	8.75	8.75	1.46	6.58	10.8	6.75	0.047	11	0.37	
G051.77+0.396	51.78	0.41	10.5	4.75	173.	10.3	10.3	13.0	4.14	31.9	14.2	0.022	32	1.46	G051.831+00.462 ⁽¹⁾
G052.30-0.575	52.30	-0.57	7.21	4.95	88.6	7.65	7.65	13.7	-4.1	16.0	10.8	0.056	31	0.29	G052.174-00.567 ⁽¹⁾
G052.77-0.554	52.77	-0.55	3.07	1.28	58.6	5.35	5.35	13.1	-2.8	4.88	1.84	0.387	11	0.22	G052.799-00.534 ⁽¹⁾
G052.17+0.737	52.17	0.73	11.6	3.87	4.85	10.2	10.2	15.3	6.88	34.6	11.7	0.029	38	0.58	G051.978+00.542 ⁽¹⁾
G052.75+0.200	52.74	0.24	8.13	4.24	135.	4.25	4.25	12.1	1.02	10.0	5.29	0.119	20	0.53	
G053.62+0.408	53.60	0.45	4.98	3.40	68.2	7.45	7.45	15.8	3.26	10.9	7.27	0.056	14	0.40	G053.935+00.228 ⁽¹⁾
G054.69+0.964	54.69	0.96	5.35	2.26	144.	7.15	7.15	12.5	5.87	11.0	4.93	0.070	12	0.33	
G054.61+0.805	54.63	0.82	4.23	2.01	158.	7.25	7.25	12.8	5.30	9.02	4.33	0.089	11	0.00	
G055.43+0.160	55.49	0.05	16.5	12.2	45.0	6.55	6.55	7.75	0.38	31.4	23.3	0.019	44	0.25	
G054.73-0.424	54.73	-0.42	4.41	1.89	151.	5.75	5.75	11.3	-2.3	7.42	3.43	0.137	11	0.22	G054.913-00.339 ⁽¹⁾
G056.37+0.332	56.37	0.29	7.29	3.49	114.	5.75	5.75	2.99	1.69	12.2	5.95	0.074	17	0.41	
G056.38-0.283	56.35	-0.29	6.10	4.34	156.	5.85	5.85	2.98	-1.7	10.2	7.28	0.059	14	0.40	
G056.40-0.012	56.40	-0.01	5.73	5.73	0.00	10.0	10.0	2.05	-0.1	16.6	16.6	0.026	23	0.27	G056.413-00.028 ⁽¹⁾
G054.37-0.476	54.36	-0.48	3.80	2.21	97.8	6.75	6.75	13.5	-3.1	7.36	4.03	0.106	10	0.25	
G056.36-0.587	56.36	-0.58	4.70	4.70	0.00	6.55	6.55	2.36	-3.6	9.04	9.04	0.038	10	0.11	
G056.94-0.252	56.91	-0.26	12.2	3.97	21.0	6.15	6.15	3.53	-1.6	21.9	7.04	0.049	24	0.50	
G057.56-0.286	57.56	-0.28	5.46	5.46	0.00	9.35	9.35	8.70	-2.6	14.7	14.7	0.024	17	0.88	G057.541-00.279 ⁽¹⁾
G057.73+0.080	57.73	0.08	7.60	3.42	4.36	8.55	8.55	9.51	0.68	18.9	8.34	0.038	19	0.46	G057.715+00.173 ⁽¹⁾
G058.11+0.192	58.13	0.17	9.03	4.02	174.	9.15	9.15	12.5	1.61	24.0	10.4	0.020	16	0.14	
G058.20+0.778	58.21	0.79	13.5	7.07	84.5	6.35	6.35	11.0	4.53	25.0	13.1	0.029	31	0.34	
G058.69+0.659	58.69	0.65	8.38	8.38	0.00	5.55	5.55	12.4	3.40	13.5	13.5	0.041	24	0.33	G058.602+00.639 ⁽¹⁾
G060.67+0.007	60.67	0.00	5.92	3.72	36.1	9.45	9.45	15.8	0.06	16.3	10.3	0.028	15	0.15	G060.659-00.031 ⁽¹⁾
G061.50+0.109	61.53	0.13	9.12	5.91	167.	5.95	5.95	8.94	0.82	15.7	10.2	0.064	33	0.22	G061.467+00.380 ⁽¹⁾
G062.37+0.297	62.36	0.30	4.17	2.46	114.	9.25	9.25	4.17	2.75	11.0	6.38	0.045	10	0.25	
G063.21-0.442	63.21	-0.44	4.10	1.08	31.3	9.05	9.05	3.35	-3.8	10.8	3.12	0.094	10	0.66	G063.093-00.401 ⁽¹⁾
G063.77-0.087	63.77	-0.07	7.58	3.91	179.	9.25	9.25	8.11	-0.7	20.4	10.5	0.022	15	0.25	
...

TABLE B.1: continued.

[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]	[14]	[15]	[16]
Cluster candidates name	l_{center} (deg)	b_{center} (deg)	a-axis (arcmin)	b-axis (arcmin)	PA (deg)	D_{peak} (kpc)	\bar{D} (kpc)	R (kpc)	Z (pc)	a-axis (pc)	b-axis (pc)	Density $Clumps.pc^{-2}$	Nb _{clumps}	Nb _{proto} /Nb _{pre}	H II regions
G064.26-0.428	64.24	-0.40	10.0	3.81	148.	3.95	3.95	8.78	-1.5	11.5	4.31	0.133	21	0.50	G064.130-00.475 ⁽¹⁾
G066.82+0.874	66.81	0.88	6.85	1.94	157.	3.95	3.95	11.5	3.04	7.83	2.36	0.343	20	3.00	G066.746+00.850 ⁽¹⁾
End of the list															

ANNEXE

C

CONSTANTES ET VARIABLES

TABLE C.1: Constantes utiles

Vitesse de la lumière	$c = 2,99792453 \cdot 10^8 \text{ m.s}^{-1}$
Constante de Planck	$h = 6,6260755 \cdot 10^{-34} \text{ m}^2.\text{kg}.\text{s}^{-1}$
Constante de Boltzmann	$k_B = 1,3806485 \cdot 10^{-23} \text{ m}^2.\text{kg}.\text{s}^{-2}.\text{K}^{-1}$
Masse du proton	$m_p = 1,672623 \cdot 10^{-27} \text{ kg}$
Unité astronomique	$UA = 1,5 \cdot 10^{11} \text{ m}$
Année lumière	$al = 9,46 \cdot 10^{15} \text{ m}$
parsec	$pc = 3,086 \cdot 10^{16} \text{ m}$
Masse du Soleil	$M_{\odot} = 1,989 \cdot 10^{30} \text{ kg}$

TABLE C.2: Conversions d'unité utiles

Longueur	$1 \text{ pc} = 3,26 \text{ al}$
Température	$1 \text{ K} = -273,15 \text{ }^{\circ}\text{C}$
Masse	$1 \text{ kg} = 1000 \text{ g}$
Densité de flux	$1 \text{ Jy} = 10^{-26} \text{ W.m}^{-2}.\text{Hz}^{-1}$
	$1 \text{ Jy} = 10^{-23} \text{ erg.s}^{-1}.\text{cm}^{-2}.\text{Hz}^{-1}$
Puissance	$1 \text{ W} = 10^7 \text{ erg.s}^{-1} = \text{J.s}^{-1}$

ANNEXE

D

GLOSSAIRE

D.1 Acronymes

AGB : *Asymptotic Giant Branch*. Lorsque les étoiles peu massives ($\leq 8 M_{\odot}$) ont épuisé leur hydrogène et hélium en leur cœur elles subissent des changements morphologiques radicaux et appartiennent désormais à la branche des géantes asymptotiques.

BG : *Big Grain*, ces grains émettent dans l'infrarouge moyen/lointain et dans le sub-millimétrique. C'est la plus grosse taille de grain et ils comprennent la plupart de la matière solide.

CMF : *Clump Mass Function*, fonction de masse des clumps.

CoMF : *Core Mass Function*, fonction de masse des cœurs.

CDR : *Compact Distance Range*. Il définit l'intervalle de valeur des estimations de distances héliocentriques pour lequel les clumps sont considérés comme associés l'un à l'autre dans une sur-densité et pour un repère tri-dimensionnel (section 2.3, chapitre 3).

IMF : *Initial Mass Function*, fonction de masse initiale.

IRDC : *InfraRed Dark Cloud*. Ce sont des régions denses et froides qui ne sont visibles que par contraste avec leur environnement. Les dernières études montrent une relation significative entre les IRDCs et la formation d'étoiles massives.

GMC : *Giant Molecular Cloud*, nuage moléculaire géant. Ce sont de grands nuages de gaz et poussières, ayant une taille de plusieurs dizaines pc, généralement associés aux bras spiraux et permettant la formation de plusieurs régions de formation stellaire.

HDE : *Heliocentric Distance Estimate*, distance héliocentrique des sources. Elles sont notamment utilisées pour délimiter les amas.

MIS : *Milieu InterStellaire*. Ce terme regroupe l'ensemble des constituants comblant le vide entre les étoiles, d'où son nom explicite. L'étude de ce milieu est tellement large car il interagit avec d'autres domaines de l'astrophysique, comme la formation stellaire, qu'il est considéré comme un domaine à part entière lui aussi.

MCMF : *Molecular Cloud Mass Function*, fonction de masse des nuages moléculaires.

PAH : *Polycyclic Aromatic Hydrocarbon*, ce sont des molécules responsables de pics d'émission dans l'infrarouge moyen.

PDR : *Photo-Dissociation Region*, partie externe d'une région H II délimitant la partie ionisée de la partie neutre et moléculaire dense.

PMS : *Pre Main Sequence*, zone caractérisant l'évolution des étoiles avant qu'elles n'appartiennent à la séquence principale. Ces étoiles sont pratiquement entièrement formées et les réactions de fusion de l'hydrogène sont sur le point de commencer.

SED : *Spectral Energy Distribution*, profil en flux de la source aux différentes longueurs d'onde.

SFE : *Star Formation Efficiency*, efficacité de formation stellaire. Elle caractérise le pourcentage de masse disponible dans un amas de gaz et poussière pour former une étoile.

SDPF : *Science Demonstration Phase Field*. En général c'est le premier champ obtenue d'un projet d'un télescope et permettant d'effectuer les premières études apportant des résultats scientifiques. Exemple : Hi-GAL SDP.

UCH II : *Ultra-compact H II region*. Ce sont des régions H II caractérisant la formation récente d'une étoile massive encore enfouie dans le nuage moléculaire.

VSG : *Very Small Grain*, ces grains ont été introduits afin de combler l'émission entre les PAHs et les BGs.

YSO : *Young Stellar Object*, nom donné aux objets stellaires durant leur formation et pendant les premières étapes de leur évolution.

D.2 Instruments dans l'infrarouge et le sub-millimétrique

AKARI : Originellement connu sous le nom ASTRO-F, AKARI est un télescope spatial développé par l'agence spatiale japonaise (JAXA).

APEX : APEX est un télescope terrestre de 12 mètres de diamètre, développé par l'ESA et se situant dans la région de l'Atacama au Chili.

COBE : Le Cosmic Background Explorer est un télescope spatial de la NASA, lancé le 18 novembre 1989. Il est l'un des prédecesseurs du satellite Planck.

CSO : Le Caltech Submillimeter Observatory est un télescope terrestre dans le domaine sub-millimétrique, maintenu en fonctionnement jusqu'en 2015 par l'université californienne Caltech et localisé près du sommet du Mauna Kea à Hawaï.

Herschel : Le télescope Herschel est un télescope spatial de l'ESA (Agence Spatiale Européenne) dans le domaine infrarouge et sub-millimétrique, lancé le 14 mai 2009 par la fusée Ariane 5 au point de Lagrange L2 et en service du 14 juin 2009 jusqu'à fin avril 2013.

IRAS : L'Infrared Astronomical Satellite est un télescope spatial développé conjointement par la NASA, l'agence néerlandaise (NIVR) et l'agence de la recherche britannique (SERC). C'est un télescope dans le domaine de l'infrarouge proche et lointain, lancé le 25 janvier 1983 et dont la mission dura 2 mois.

IRTS : IRTS est un télescope embarqué à bord du Space Flier Unit (SFU), lancé en 1995. Ce télescope balaie un spectre large de longueurs d'onde, allant de l'infrarouge proche au domaine millimétrique.

ISO : L'Infrared Space Observatory est un télescope spatial développé par l'ESA et lancé le 17 novembre 1995. Son domaine de longueurs d'onde va de l'infrarouge proche à lointain.

JCMT : Le James Clerk Maxwell Telescope est un télescope terrestre dans le domaine sub-millimétrique possédant un diamètre de 15m. Tout comme le télescope CSO, il est situé près du Mauna Kea à Hawaï.

MSX : Ce télescope est un télescope militaire américain, lancé le 24 avril 1996, développé dans le but d'étudier les émissions infrarouges depuis l'espace. Il était à l'origine utilisé pour surveiller les mouvements de sources infrarouges, notamment les missiles balistiques, mais il a pris part à la recherche scientifique ensuite.

Planck : Planck est un télescope spatial de l'ESA, lancé avec Herschel le 14 mai 2009, successeur du satellite COBE et étudiant le fond diffus cosmologique. Son domaine de longueurs d'onde va du sub-millimétrique jusqu'au radio.

PRONAOS : C'est un ballon français de 2 mètres de diamètre observant dans la haute atmosphère et dans le domaine sub-millimétrique, dédié à l'observation du milieu interstellaire et du fond diffus cosmologique. Trois vols ont été effectués en 1994, 1996 et 1999.

SOFIA : Le Stratospheric Observatory for Infrared Astronomy est un télescope embarqué à bord d'un avion, se composant de plusieurs instruments, et observant dans le visible et l'infrarouge. Le projet a été développé par la NASA et l'agence spatiale allemande. Le premier vol fut effectué en 2007.

Spitzer : Le télescope Spitzer, connu originellement sous le nom de SIRTf (Space Infrared Telescope Facility), a été développé par la NASA. Il observe dans le domaine infrarouge, moyen et lointain. Il a été lancé le 25 août 2003. En mai 2009 l'épuisement du système de refroidissement stoppa la mission *froide*.

WISE : Wide-Field Infrared Survey Explorer. Ce télescope développé par la NASA étudie le ciel dans l'infrarouge moyen. Il a été lancé le 14 décembre 2009 et mis en hiber-

nation en 2011 avant d'être réactivé en 2013 afin de rechercher de nouveaux astéroïdes.

BIBLIOGRAPHIE

- Allamandola, L. J., Tielens, A. G. G. M., & Barker, J. R. 1989, *ApJS*, 71, 733 13
- Allison, R. J., Goodwin, S. P., Parker, R. J., et al. 2009, *MNRAS*, 395, 1449
- Alves, J., Lombardi, M., & Lada, C. J. 2007, *A&A*, 462, L17
- Anderson, L. D., Bania, T. M., Balser, D. S., et al. 2014, *ApJS*, 212, 1
- André, P., Men'shchikov, A., Bontemps, S., et al. 2010, *A&A*, 518, L102 xvi, 17, 20
- Arab, H., Abergel, A., Habart, E., et al. 2012, *A&A*, 541, A19 36
- Arzoumanian, D., André, P., Peretto, N., & Könyves, V. 2013, *A&A*, 553, A119 xvi, 17, 20
- Arzoumanian, D., André, P., Didelon, P., et al. 2011, *A&A*, 529, L6 xvi, 17, 20
- Barrow, J. D., Bhavsar, S. P., & Sonoda, D. H. 1985, *MNRAS*, 216, 17
- Bastian, N., Covey, K. R., & Meyer, M. R. 2010, *ARA&A*, 48, 339
- Battinelli, P. 1991, *A&A*, 244, 69
- Billot, N., Schisano, E., Pestalozzi, M., et al. 2011, *ApJ*, 735, 28
- Boggess, N. W., Mather, J. C., Weiss, R., et al. 1992, *ApJ*, 397, 420 27
- Bok, B. J. 1934, *Harvard College Observatory Circular*, 384, 1
- Bontemps, S., André, P., Könyves, V., et al. 2010, *A&A*, 518, L85 39
- Borůvka, O. 1926a, *Práce morav. přírodověd. spol. v Brně*, 3, 37, 58
- . 1926b, *Elektronický obzor*, 15, 153, 154
- Brand, J., & Blitz, L. 1993, *A&A*, 275, 67
- Burton, W. B., & Hartmann, D. 1988, *Structure and dynamics of the galactic system. Triennial report 1984 - 1987*, *IAU Commission 33*. 8, 79

- Camargo, D., Bica, E., & Bonatto, C. 2015, *New A*, 34, 84
- . 2016, *MNRAS*, 455, 3126
- Cambr sy, L., Marton, G., Feher, O., T th, L. V., & Schneider, N. 2013, *A&A*, 557, A29
- Campana, R., Massaro, E., Gasparrini, D., Cutini, S., & Tramacere, A. 2008, *MNRAS*, 383, 1166
- Carey, S. J., Noriega-Crespo, A., Mizuno, D. R., et al. 2009, *PASP*, 121, 76 37
- Cartwright, A. 2009, *MNRAS*, 400, 1427
- Cartwright, A., & Whitworth, A. P. 2004, *MNRAS*, 348, 589
- . 2009, *MNRAS*, 392, 341
- Casertano, S., & Hut, P. 1985, *ApJ*, 298, 80
- Colberg, J. M. 2007, *MNRAS*, 375, 337
- Combes, F. 2000, *Advanced Series in Astrophysics and Cosmology*, 10, 143
- Dame, T. M., Hartmann, D., & Thaddeus, P. 2001, *ApJ*, 547, 792 xv, xvi, 9, 17, 21
- de Graauw, T., Helmich, F. P., Phillips, T. G., et al. 2010, *A&A*, 518, L6 29
- Deharveng, L., Zavagno, A., & Caplan, J. 2005, *A&A*, 433, 565 xv, 17, 18, 96
- Deharveng, L., Zavagno, A., Schuller, F., et al. 2009, *A&A*, 496, 177
- Derriere, S., & Robin, A. C. 2001, in *Astronomical Society of the Pacific Conference Series*, Vol. 232, *The New Era of Wide Field Astronomy*, ed. R. Clowes, A. Adamson, & G. Bromage, 229
- Desert, F.-X., Boulanger, F., & Puget, J. L. 1990, *A&A*, 237, 215 xv, 12
- Dorval, J., Boily, C. M., Moraux, E., Maschberger, T., & Becker, C. 2016, *MNRAS*, 459, 1213 111
- Draine, B. T., & Anderson, N. 1985, *ApJ*, 292, 494 12
- Draine, B. T., & Lee, H. M. 1984, *ApJ*, 285, 89 xv, 11
- Dunham, M. M., Crapsi, A., Evans, II, N. J., et al. 2008, *ApJS*, 179, 249 39
- Duran, M., & Buisson, F., eds. 1990, *The PRONAOS telescope* 27
- Eden, D. J., Moore, T. J. T., Morgan, L. K., Thompson, M. A., & Urquhart, J. S. 2013, *MNRAS*, 431, 1587
- Eden, D. J., Moore, T. J. T., Plume, R., & Morgan, L. K. 2012, *MNRAS*, 422, 3178
- Eden, D. J., Moore, T. J. T., Urquhart, J. S., et al. 2015, *MNRAS*, 452, 289
- Efremov, Y. N., & Elmegreen, B. G. 1998, *MNRAS*, 299, 588 xv, 14, 16

- Egan, M. P., Price, S. D., & Kraemer, K. E. 2003, in *Bulletin of the American Astronomical Society*, Vol. 35, American Astronomical Society Meeting Abstracts, 1301–28
- Egan, M. P., Shipman, R. F., Price, S. D., et al. 1998, *ApJ*, 494, L199–17
- Elia, D., & Pezzuto, S. 2016, *ArXiv e-prints*, arXiv :1606.02496–36
- Elia, D., Molinari, S., Fukui, Y., et al. 2013, *ApJ*, 772, 45–35, 38, 39, 40, 41
- Elmegreen, B. G., & Clemens, C. 1985, *ApJ*, 294, 523
- Englmaier, P., Pohl, M., & Bissantz, N. 2011, *Memorie della Societa Astronomica Italiana Supplementi*, 18, 199
- Espinoza, P., Selman, F. J., & Melnick, J. 2009, *A&A*, 501, 563
- Freedman, D., & Diaconis, P. 2011, *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 57, 453
- Georgelin, Y. M., & Georgelin, Y. P. 1976, *A&A*, 49, 57
- Giannini, T., Elia, D., Lorenzetti, D., et al. 2012, *A&A*, 539, A156–39
- Gillett, F. C., Forrest, W. J., & Merrill, K. M. 1973, *ApJ*, 183, 87–12
- Glenn, J., Bock, J. J., Chattopadhyay, G., et al. 1998, in *Proc. SPIE*, Vol. 3357, *Advanced Technology MMW, Radio, and Terahertz Telescopes*, ed. T. G. Phillips, 326–334–28
- González, R. E., & Padilla, N. D. 2009, *MNRAS*, 397, 1498
- Goodwin, S. P., Nutter, D., Kroupa, P., Ward-Thompson, D., & Whitworth, A. P. 2008, *A&A*, 477, 823
- Graham, R. L., & Hell, P. 1985, *Annals of the History of Computing*, 7, 43
- Griffin, M. J., Abergel, A., Abreu, A., et al. 2010, *A&A*, 518, L3–29
- Gutermuth, R. A., Megeath, S. T., Myers, P. C., et al. 2009, *ApJS*, 184, 18
- Gutermuth, R. A., Megeath, S. T., Pipher, J. L., et al. 2005, *ApJ*, 632, 397
- Hartmann, J. 1904, *ApJ*, 19, doi :10.1086/141112–xxiii, 4
- Hatchell, J., & Fuller, G. A. 2008, *A&A*, 482, 855
- Heiles, C., & Crutcher, R. 2005, in *Lecture Notes in Physics*, Berlin Springer Verlag, Vol. 664, *Cosmic Magnetic Fields*, ed. R. Wielebinski & R. Beck, 137–6
- Henderson, A. P., Jackson, P. D., & Kerr, F. J. 1982, *ApJ*, 263, 116
- Hildebrand, R. H. 1983, *QJRAS*, 24, 267–39, 103
- Holland, W. S., Robson, E. I., Gear, W. K., et al. 1999, *MNRAS*, 303, 659–28
- Holland, W. S., Bintley, D., Chapin, E. L., et al. 2013, *MNRAS*, 430, 2513–28
- Holman, K., Walch, S. K., Goodwin, S. P., & Whitworth, A. P. 2013, *MNRAS*, 432, 3534

- Kaiser, R. I. 2002, *Chemical Reviews*, 102, 1309 xv, 14
- Kelly, B. C., Shetty, R., Stutz, A. M., et al. 2012, *ApJ*, 752, 55 36
- Kendrew, S., Simpson, R., Bressert, E., et al. 2012, *ApJ*, 755, 71
- Kessler, M. F., Steinz, J. A., Anderegg, M. E., et al. 1996, *A&A*, 315, L27 28
- Kirk, H., & Myers, P. C. 2011, *ApJ*, 727, 64
- Kirk, H., Johnstone, D., Di Francesco, J., et al. 2016, *ApJ*, 821, 98
- Kirsanova, M. S., Sobolev, A. M., Thomasson, M., et al. 2008, *MNRAS*, 388, 729
- Klessen, R. S., & Burkert, A. 2000, *ApJS*, 128, 287
- Koenig, X. P., Allen, L. E., Gutermuth, R. A., et al. 2008, *ApJ*, 688, 1142
- Kroupa, P. 2001, *MNRAS*, 322, 231
- . 2002, *Science*, 295, 82
- Kroupa, P., Tout, C. A., & Gilmore, G. 1993, *MNRAS*, 262, 545
- Krumholz, M. R., McKee, C. F., & Klein, R. I. 2005, *Nature*, 438, 332
- Kuhn, M. A., Getman, K. V., & Feigelson, E. D. 2015, *ApJ*, 802, 60
- Küpper, A. H. W., Maschberger, T., Kroupa, P., & Baumgardt, H. 2011, *MNRAS*, 417, 2300
- Lada, C. J., & Lada, E. A. 2003, *ARA&A*, 41, 57
- Larson, R. B. 1981, *MNRAS*, 194, 809 39
- Leinert, C., Bowyer, S., Haikala, L. K., et al. 1998, *A&AS*, 127, 1 xv, 5
- Majaess, D. 2013, *Ap&SS*, 344, 175
- Marshall, D. J., Robin, A. C., Reyl  , C., Schultheis, M., & Picaud, S. 2006, *A&A*, 453, 635
- Maschberger, T. 2013, *MNRAS*, 429, 1725
- Maschberger, T., & Clarke, C. J. 2011, *MNRAS*, 416, 541
- Maschberger, T., Clarke, C. J., Bonnell, I. A., & Kroupa, P. 2010, *MNRAS*, 404, 1061
- Massaro, E., Tinebra, F., Campana, R., & Tosti, G. 2009, *ArXiv e-prints*, arXiv :0912.3843
- Massey, P. 1998, in *Astronomical Society of the Pacific Conference Series*, Vol. 142, *The Stellar Initial Mass Function (38th Herstmonceux Conference)*, ed. G. Gilmore & D. Howell, 17
- Massey, P., Penny, L. R., & Vukovich, J. 2002, *ApJ*, 565, 982
- Mather, J. C. 1982, *Optical Engineering*, 21, 769 27
- Mathis, J. S., Rumpl, W., & Nordsieck, K. H. 1977, *ApJ*, 217, 425 10
- May, B. 2008, *A Survey of Radial Velocities in the Zodiacal Dust Cloud* xxiii

- McMillan, P. J. 2011, MNRAS, 414, 2446 4
- Michel, M., Kirk, H., & Myers, P. C. 2011, ApJ, 735, 51
- Molinari, S., Schisano, E., Faustini, F., et al. 2011, A&A, 530, A133 31
- Molinari, S., Swinyard, B., Bally, J., et al. 2010a, A&A, 518, L100 17, 49
- . 2010b, PASP, 122, 314 29
- Molinari, S., Schisano, E., Elia, D., et al. 2016a, ArXiv e-prints, arXiv :1604.05911 31, 103
- Molinari, S., Noriega-Crespo, A., Bally, J., et al. 2016b, A&A, 588, A75
- Moore, T. J. T., Bretherton, D. E., Fujiyoshi, T., et al. 2007, MNRAS, 379, 663
- Moore, T. J. T., Urquhart, J. S., Morgan, L. K., & Thompson, M. A. 2012, MNRAS, 426, 701
- Morales, E. F. E., Wyrowski, F., Schuller, F., & Menten, K. M. 2013, A&A, 560, A76
- Murakami, H., Bock, J., Freund, M. M., et al. 1994, ApJ, 428, 354 27
- Murakami, H., Baba, H., Barthel, P., et al. 2007, PASJ, 59, S369 28
- Neugebauer, G., Habing, H. J., van Duinen, R., et al. 1984, ApJ, 278, L1 28
- Nešetřil, J., & Nešetřilová, H. 2012, Documenta Mathematica, Vol Extra Vol : Optimization Stories, 127
- Ochsenbein, F., Bauer, P., & Marcout, J. 2000, A&AS, 143, 23
- Paladini, R., Burigana, C., Davies, R. D., et al. 2003, A&A, 397, 213
- Paladini, R., Davies, R. D., & De Zotti, G. 2004, MNRAS, 347, 237
- Parker, R. J., & Dale, J. E. 2015, MNRAS, 451, 3664
- Parker, R. J., Maschberger, T., & Alves de Oliveira, C. 2012, MNRAS, 426, 3079
- Peretto, N., & Fuller, G. A. 2009, A&A, 505, 405
- Pilbratt, G. L., Riedinger, J. R., Passvogel, T., et al. 2010, A&A, 518, L1 28, 29
- Planck Collaboration, Abergel, A., Ade, P. A. R., et al. 2011, A&A, 536, A25 36
- Poglitsch, A., Waelkens, C., Geis, N., et al. 2010, A&A, 518, L2 29
- Prim, R. C. 1957, Bell System Technical Journal, 36, 1389
- Purcell, E. M. 1969, ApJ, 158, 433 9
- Rachford, B. L., Snow, T. P., Tumlinson, J., et al. 2002, ApJ, 577, 221 9
- Reed, B. C. 1997, PASP, 109, 1145
- . 2006, JRASC, 100, 146
- Reid, M. A., Wadsley, J., Petitclerc, N., & Sills, A. 2010, ApJ, 719, 561

- Reyl , C., Marshall, D. J., Robin, A. C., & Schultheis, M. 2008, in *Astronomical Society of the Pacific Conference Series*, Vol. 396, *Formation and Evolution of Galaxy Disks*, ed. J. G. Funes & E. M. Corsini, 215
- Rom n-Z niga, C. G., Elston, R., Ferreira, B., & Lada, E. A. 2008, *ApJ*, 672, 861
- Rosolowsky, E. 2005, *PASP*, 117, 1403
- Russeil, D., Pestalozzi, M., Mottram, J. C., et al. 2011, *A&A*, 526, A151
- Salpeter, E. E. 1955, *ApJ*, 121, 161
- Saral, G., Hora, J. L., Willis, S. E., et al. 2015, *ApJ*, 813, 25
- Savage, B. D., & Sembach, K. R. 1996, *ARA&A*, 34, 279 xiii, 7
- Schmeja, S. 2011, *Astronomische Nachrichten*, 332, 172
- Schmeja, S., & Klessen, R. S. 2006, *A&A*, 449, 151
- Schuller, F., Menten, K. M., Contreras, Y., et al. 2009, *A&A*, 504, 415
- Schulz, N. S. 2012, *The Formation and Early Evolution of Stars*, doi :10.1007/978-3-642-23926-7 xvi, 18, 21
- Shamos, M. I., & Hoey, D. 1975, in *Proceedings of the 16th Annual Symposium on Foundations of Computer Science, SFCS '75* (Washington, DC, USA : IEEE Computer Society), 151–162
- Silverberg, R. F., Hauser, M. G., Boggess, N. W., et al. 1993, in *Proc. SPIE*, Vol. 2019, *Infrared Spaceborne Remote Sensing*, ed. M. S. Scholl, 180–189 27
- Simon, R., Jackson, J. M., Rathborne, J. M., & Chambers, E. T. 2006, *ApJ*, 639, 227
- Simpson, R. J., Nutter, D., & Ward-Thompson, D. 2008, *MNRAS*, 391, 205
- Siringo, G., Kreysa, E., Kov cs, A., et al. 2009, *A&A*, 497, 945 28
- Spitzer, Jr., L. 1940, *MNRAS*, 100, 396
- . 1958, *ApJ*, 127, 17
- Sreekumar, P., Bertsch, D. L., Dingus, B. L., et al. 1993, *Physical Review Letters*, 70, 127 6
- Tauber, J. A., Mandolesi, N., Puget, J.-L., et al. 2010, *A&A*, 520, A1 28
- Taylor, J. H., & Cordes, J. M. 1993, *ApJ*, 411, 674
- Temi, P., Marcum, P. M., Young, E., et al. 2014, *ApJS*, 212, 24 27
- Thompson, M. A., Urquhart, J. S., Moore, T. J. T., & Morgan, L. K. 2012, *MNRAS*, 421, 408
- Toussaint, G. T. 1980, *Pattern Recognition*, 12, 261
- Traficante, A., Calzoletti, L., Veneziani, M., et al. 2011, *MNRAS*, 416, 2932 31

- Tsuboi, M., & Miyazaki, A. 2012, PASJ, 64, arXiv :1204.5535
- Tsuboi, M., Miyazaki, A., & Uehara, K. 2015, PASJ, 67, 109
- Urquhart, J. S., Thompson, M. A., Moore, T. J. T., et al. 2013, MNRAS, 435, 400
- Veneziani, M., Piacentini, F., Noriega-Crespo, A., et al. 2013a, ApJ, 772, 56–36
- Veneziani, M., Elia, D., Noriega-Crespo, A., et al. 2013b, A&A, 549, A130–34, 35, 55
- Vollmer, B., Davoust, E., Dubois, P., et al. 2005, A&A, 431, 1177–31
- von Hoerner, S. 1963, ZAp, 57
- Vutisalchavakul, N., Evans, II, N. J., & Battersby, C. 2014, ApJ, 797, 77
- Wang, K., Zhang, Q., Testi, L., et al. 2014, MNRAS, 439, 3275–xvi, 17, 19
- Werner, M. W., Roellig, T. L., Low, F. J., et al. 2004, ApJS, 154, 1–28
- Wienen, M., Wyrowski, F., Menten, K. M., et al. 2015, A&A, 579, A91
- Williams, J. P., & McKee, C. F. 1997, ApJ, 476, 166–17
- Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., et al. 2010, AJ, 140, 1868–28

Formation stellaire dans la galaxie et interaction avec le milieu interstellaire.

Résumé

Les étoiles jouent un rôle important dans l'évolution des galaxies ainsi que dans l'évolution des systèmes planétaires et l'apparition d'une vie complexe tel que la notre. Il est ainsi important de répondre à cette question : Comment les étoiles se forment-elles ? Cette vaste question fait appel à des connaissances dans plusieurs domaines dont deux majeurs, la Formation Stellaire et le Milieu Interstellaire. C'est dans ce cadre générale que s'inscrit ma thèse. Notre galaxie est un vaste laboratoire d'études de cette formation et je me suis donc intéressé aux premières étapes de la formation des étoiles, allant du nuage moléculaire à la proto-étoile. J'ai principalement utilisé des données provenant du télescope Herschel qui nous fournit des images et des données dans l'infrarouge lointain et le domaine sub-millimétrique à une résolution inégalée. J'ai d'abord construit un catalogue de sources à l'aide d'un algorithme d'identification croisée, SPECFIND, puis appliqué un algorithme de clustering, MST, sur près de 100 000 sources afin de construire le premier catalogue d'amas d'objets stellaires jeunes à l'échelle galactique. Ceci m'a conduit à étudier les propriétés de ces amas et des sources les constituant.

Mots clés Astrophysique, Catalogue, Amas enfouis, Formation d'étoiles, Structure galactique, Infrarouge, Fonction de masse, Milieu Interstellaire, Clustering

Abstract

Stars play an important role for the evolution of galaxies as well as for the evolution of planetary systems and the emergence of a complex life such as ours. Therefore, it is important to answer this question : How stars form ? This broad question uses knowledges in several areas, including two majors, the Star Formation and the Interstellar Medium. My thesis is a part of this overall framework. Our galaxy is a laboratory complex for the study of this formation. I became interested in the first stages of the star formations, from Molecular Clouds to protostars. I mainly used data from the Herschel telescope which provides us with images and data in the far infrared and sub-millimeter at an unparalleled resolution. First of all, I built a catalogue of young clumps using SPECFIND, an algorithm of cross-identification. Then I applied an algorithm of clustering, MST, over 100 000 young clumps to find over-densities in order to release the first catalogue of young stellar clusters in a galactic scale. Finally, I studied the physical properties of these clusters and their young clumps.

Keywords Astrophysics, Catalogue, Embedded clusters, Star formation, Galactic Structure, Infrared, Mass function, Interstellar Medium, Clustering