THE DYNAMICS OF INFRARED DARK HUBS AND THE FORMATION OF MASSIVE STARS

Michael Pryderi Hughes Anderson

A THESIS SUBMITTED TO
CARDIFF UNIVERSITY
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

30TH JUNE 2023

Acknowledgements

Well, this was a long haul, wasn't it? I've been at Cardiff for what feels like a very long time, and although the past five years have been exceptionally challenging, I am extremely fortunate to have worked with some incredible people in that time. Although it is a massive cliché, I don't think that I can really do justice to my gratitude in writing, but here we go anyway...

Firstly my supervisory team, Nicolas Peretto, Sarah Ragan and Andy Rigby. Thank you so much for your endless encouragement, counsel, patience, honesty, and also dragging me out of countless rabbit holes by the scruff of my neck. Thank you for all of the guidance, opportunities and experiences you have provided. As tough as it has been over the past five years, I consider myself very lucky to have worked with you, and thank you for managing to make it fun along the way too!

To my examiners, Tim Davis and Steven Longmore, thank you for putting me at ease and, despite my extreme nerves, making the viva a both an enjoyable and fulfilling experience.

I would also like to thank the Science & Technology Facilities Council and Cardiff University for funding my PhD.

Liz Watkins, how lucky I was to have such a beyond-cromulent mentor! I have learnt so much from you since the start of the PhD. Your passion for research is formidable, and I was frequently impressed/horrified by your drive to add yet another dimension to an already dense plot. Your rants were always spectacular, your puns painful, and your choice of adjectives to describe food hilarious.

Gerwyn Jones, you were a fleeting but integral part of the office. Thank you for bringing cheer and a bit of silliness whenever you were in, and I really appreciate all the candid chats after many a long day. You always made reaching a high *Peñaloza index* look effortless.

Thank you to Felix Priestley, for helping me understand so many convoluted theoretical papers. I don't think anyone has made me laugh as much in talks with your consistently blunt delivery (both when presenting and asking questions). Roger Wesson, thanks for keeping me caffeinated and recommending some relaxing music

to write my thesis to. Holly Davies, thanks for nerding out with me over various *Star Wars* shows. Zeinab Khorrami, thanks for lightening the mood and setting off unexpected chains of laughter in the office.

Thank you Virginia D'Emilio, Charlotte Braithwaite, Vassilis Skliris, for bringing some normality to the office during some very strange and difficult times.

Thank you to Terri Pearce, for always being real and telling it like it is. Your music recommendations were always on point. Jíxuán Zhōu, thank you for offering me all those snacks as a pick-me-up during the day, your positivity, and for the long list of places to visit in China. Ilaria Ruffa, thank you for showing me the breadth of Italian profanity that can be used to berate *CASA*. Hearing it from the office over always cracked me up. Benjamin Beringue, thank you for being such a kind officemate. Your crêpes were incredible!

Eva Durán Camacho, the *Cuarenta y Tres*, all of the *queso* smuggled over from Spain, and introducing me to *What If...?* were almost certainly critical to my research! I can't thank you enough for all the laughs, long conversations, and always checking in on me no matter how stressed out I looked. You are an exceptional colleague and friend to have, and wherever you end up people will be lucky to have you around.

Thank you Ana Duarte Cabral, for giving me my first taste of astronomy research, introducing me to the group at Cardiff, and for all your motivational chats and advice over the years. Paul Clark, thanks for all of our discussions (sensible or otherwise), the frequent reminders to never take astronomy too seriously and, of course, for inducting me to the *Flute & Tankard*. Ant Whitworth, thanks for your abundant thought provoking questions and enthusiasm. Paul Roche and Richard Lewis, thanks for sharing your wealth of experience (and the nuances of navigating University bureaucracy).

Thank you to all the support staff, who without the School would collapse, especially Jess Phillips, Neil Sigllechner, Gareth Keene. Thank you to the druid of computing, Richard Frewin, for ensuring that I was never hampered by technical issues, for answering my plentiful support requests, and keeping *Reblochon* in working order despite the constant strain I put it under.

Thank you to Paul Horrocks and Alex Muhl-Richardson for talking *utter nonsense* late into the nights. It was always a welcome distraction from the daily grind.

Gayathri Eknath, I will forever admire your drive to change things for the better, and your determination when the going gets tough. Thank you for all the help, letting me talk through my problems at length, and for putting up with my rambling. Thanks also for all the trips to numerous cheap eats to find some much needed *dosa*, *vada pav* and other comfort food. Escaping out into nature with you to find some serenity amongst all the chaos of the PhD got me through a lot of stressful moments. Here's to many more trips, but maybe with less of the chaos and stress!

To my father, thank you for always encouraging me to ask questions, engage in discussion on a vast range of topics from a young age, and for taking me out to look at the night sky.

Finally, to my mother, thank you for your unending empathy, calming influence, pragmatism, and providing such a safe environment for me throughout my life. There is no question that this PhD would not have been possible without your support over all these years.

Abstract

Hub-filament systems (HFSs) are networks of converging interstellar filaments, often with active star formation at their centres, that may play an important role in high-mass star formation. In this thesis, I have studied a sample of six HFSs observed with ALMA 2.9 mm continuum and $N_2H^+(J=1\text{–}0)$ molecular line emission to investigate their physical and dynamical properties.

The first part of this research project was to investigate the how much of their mass is contained within cores, and with a broader sample of clumps, whether their morphology plays a part in what fraction of their mass is locked up in their most massive core. I showed that this fraction is significantly higher in IR-dark hubs than IR-bright clumps, suggesting that the most-massive cores form early on.

In the second project, I investigated the kinematic and dynamical properties of these six HFSs using the dense gas tracer N_2H^+. The data show intricate emission structures, and in total contain around 180,000 spectra, with a high fraction of them appearing very complex. At the time there was no programme that could automatically fit such complex spectra en masse. As a consequence I led the development of `mwydyn`, a fully-automated, multiple velocity component, hyperfine line-fitting code.

Results from this code revealed that the emission invariably consists of quiescent individual filaments in the outskirts that converge towards dynamically active hub centres, where a systematic increase in velocity dispersion and number of velocity components is observed, which I suggest could be the result of an accretion shock at the filament-hub intersection, or complex unresolved gas flows within the hub. Also, despite the masses of the HFSs spanning over an order of magnitude, the distributions of their velocity centroid are remarkably similar. I propose that this could be a sign that these HFSs are sheet-like structures seen mostly face-on.

However, in order to draw robust conclusions it is clear that further investigation is required with a much larger, unbiased sample of clumps covering a broad range of masses, evolutionary stage, and morphology.

Publications

First author publications

- Anderson et al. (2021)[1]

Co-author publications

- Rigby et al. (2021)
- Faustino Vieira et al. (2023a)
- Faustino Vieira et al. (2023b)
- Rigby et al. (2024)[2]

---

[1] Part of the work presented in this thesis is published in this paper.

[2] The code `mwydyn`, presented in this thesis, is published in this paper.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 The interstellar medium

Looking up at the night sky with the naked eye it would be easy to believe that space is an unchanging, vast, empty void peppered with stars. This couldn't be further from the truth. The interstellar medium (ISM) is the term all of the material between stars. Comprised predominantly of hydrogen ($\sim 70\%$ by mass), with some helium ($\sim 28\%$) and trace quantities ($\sim 2\%$) of all other elements which astronomers (confusingly) refer to as "metals" out of convenience. The overwhelming majority of elements are in the gas phase ($\sim 99\%$) that is either molecular, atomic (neutral), or ionised, with the remaining $\sim 1\%$ locked up in interstellar dust grains (Ferrière 2001), which are small solid particles comprised of silicates and carbon compounds, typically less than $1\,\mu\mathrm{m}$ in size (Draine 2011).

It is a diverse, dynamic and complex medium that covers a broad range of temperatures and densities. Astronomers tend to break up the ISM into phases, where the three-phase model was first proposed by McKee & Ostriker (1977), we now more commonly separate the ISM into four (Ward-Thompson & Whitworth 2011), or sometimes as many as seven phases (Draine 2011). Here, we will discuss five of the more distinct phases.

Starting from the largest scales, the Hot Ionised Medium (HIM) surrounds galaxies and is extremely hot ($10^6\,\mathrm{K}$ to $10^7\,\mathrm{K}$) and tenuous, with number densities less than $10^{-2}\,\mathrm{cm}^{-3}$. Regions recently subjected to a supernova explosion also have a significant amount of material in this phase.

Next, there is the Warm and Cold Neutral Medium (WNM and CNM, respectively), where the gas is largely in the form of atomic hydrogen. The WNM has densities $< 1\,\mathrm{cm}^{-3}$ and temperatures in the range of thousands to $10^4\,\mathrm{K}$, whereas the the CNM typically spans densities of around $1\,\mathrm{cm}^{-3}$ to $10^2\,\mathrm{cm}^{-3}$ and temperatures ranging between $30\,\mathrm{K}$ to $\sim 10^2\,\mathrm{K}$.

Figure 1.1: Part of the *Spitzer* GLIMPSE survey of the Galactic Plane, showing molecular emission in red, bright white emission from stellar nuseries, and dark dust clouds in extinction. Credit: NASA/JPL-Caltech/University of Wisconsin.

The coldest and densest phase is the Cold Molecular Medium (CMM), more commonly referred to by the structures that are in this phase – Molecular Clouds (MCs). The gas in this phase is largely in the molecular phase (i.e. primarily H_2) at densities from $\sim 10^2\,\mathrm{cm}^{-3}$ to upward of $10^6\,\mathrm{cm}^{-3}$, and temperatures $< 30\,\mathrm{K}$ to as low as a few Kelvin.

Finally, the presence of embedded high-mass (O- and B-type stars) within these molecular clouds ionise their surroundings (leading to temperatures around $10^4\,\mathrm{K}$, and densities around $0.2\,\mathrm{cm}^{-3}$ to $10^4\,\mathrm{cm}^{-3}$) giving us the fifth phase of the ISM – the warm ionised medium (WIM). Matter in this phase is also found around evolved solar- and intermediate-mass stars, where material shed from their outer layers gets ionised by the exposed hot stellar core.

Whether these phases are called molecular, atomic or ionised is governed by which state the bulk of the hydrogen is in that medium, given that it is the most abundant element.

It is worth noting that gas does not remain in these phases indefinitely, but in fact constantly cycles between each of these phases in the ISM, with gas being heating by extreme events such as supernova explosions on one end, and collisions and shocks leading to density perturbations, that eventually lead to runaway cooling and the formation of the densest structures in the ISM.

We can express the net cooling rate (per unit volume) of the gas with the equation:

$$\begin{aligned}\Lambda_{\rm net} &= \Lambda(n,T) - \Gamma(n,T) \\ &= n^2\Lambda(T) - n\Gamma(T)\end{aligned} \tag{1.1}$$

where n is the number density of the gas in the ISM, Λ and Γ are the cooling and heating rates per unit volume. The second form moves the density dependent parts of the rates out of Λ and Γ. Cooling is largely dominated by collisions between particles, and hence is dependent on n^2, whereas heating is dominated by emission, and so is dependent on n.

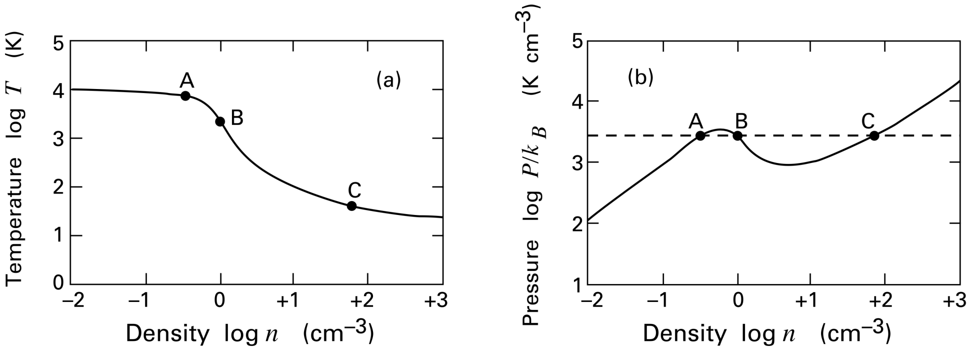

Figure 1.2: *(a)* Plot showing the theoretical prediction for the equilibrium temperature T of the ISM as a function of number density n. *(b)* Equilibrium pressure ($P/k_{\rm B} = nT$) as a function of number density. The dashed line represents the mean empirical value of $nT = 3 \times 10^3\,{\rm K\,cm^{-3}}$ for the ISM, and intersections with the equilibrium curve marked by A, B and C. Taken from Stahler & Palla (2004).

The thermal energy content of the gas comes from the balance of these heating and cooling processes (Stahler & Palla 2004). This is illustrated in Figure 1.2, where the solid black line denotes the equilibrium temperature and pressure from theoretical models when $\Lambda_{\rm net} = 0$. Gas with a pressure above this line will have $\Lambda_{\rm net} > 0$ (i.e. net cooling), and gas below this line will have $\Lambda_{\rm net} < 0$.

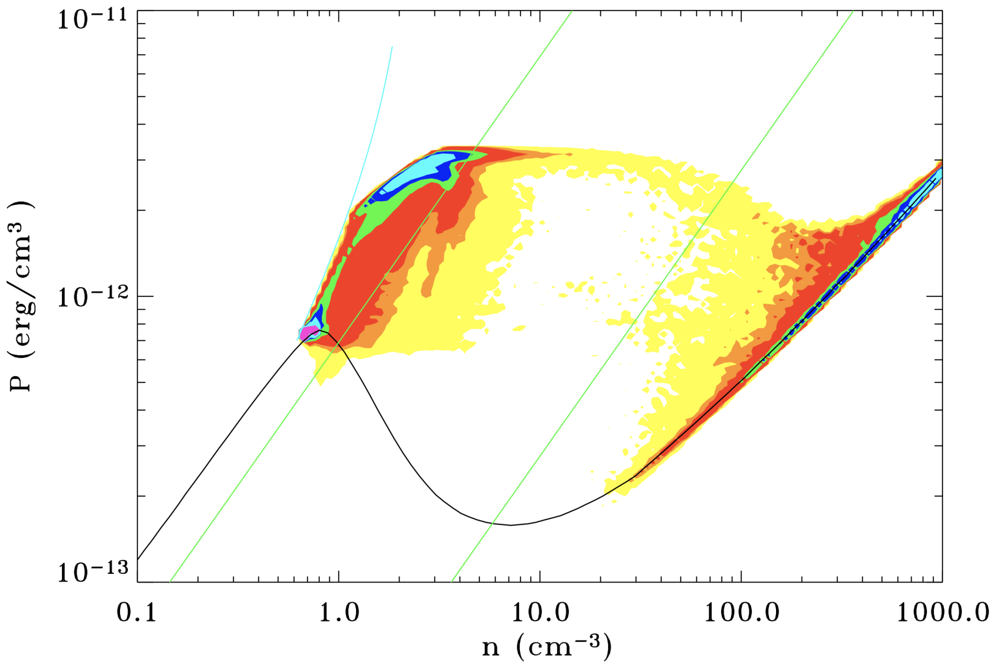

Figure 1.3: Pressure-density diagram from Audit & Hennebelle (2005). The colours represent the fraction of the gas with a given density and pressure in their simulation, with yellow being the lowest, and cyan/magenta being the highest. The black line is the thermal equilibrium curve, and the cyan line represents the initial conditions of the gas.

The mean measured value of nT for observed HI is denoted by the dashed line, which we see intersects at points A, B and C. The points A and C represent stable solutions, representing the WNM and the CNM, respectively. Gas located around point B are at a thermally unstable point, where if it is compressed slightly it will cool until reaching point C. However if the gas at point B expands slightly, it will heat up until reaching point A. As a result of this modelling it is thought that the atomic gas in the ISM is divided into two phases, with different n and T, in a pressure equilibrium. This is nicely illustrated by the gas mass distributions produced in simulations of turbulent converging flows by Audit & Hennebelle (2005), shown in Figure 1.3, which largely occupy two density regimes. Most of the cold, denser gas is closer to the equilibrium curve (i.e. it has cooled faster due to density enhancements from shocks), whereas the warmer, less dense gas is largely somewhere between the initial conditions and the equilibrium curve.

The transition from atomic to the molecular phase is a little different. Although the direct conversion of H to H_2 can occur via two hydrogen atoms combining, and releasing a photon, the rate of this transition is extremely slow. Instead, the bulk of the H_2 in the ISM is primarily by the conversion of H (from the CNM) into H_2 on the surfaces of dust grains (Hollenbach et al. 1971). Hydrogen atoms are adsorbed onto dust grains, and react with other hydrogen to form H_2. The formation of this bond releases some energy, which is partly transfered to the dust grain, and the rest kicks the newly formed H_2 molecule back into the gas phase. Molecular hydrogen is relatively easily destroyed by photodissociation by UV photons, so to maintain a fairly stable bulk amount of H_2 there needs to be a high enough density of it for it to self-shield from the incoming ionising radiation.

All stars in the Milky Way form in molecular clouds, and therefore they set the initial conditions for star formation (Lada & Lada 2003; Heyer & Dame 2015). The work in this thesis will focus on the CMM, and specifically structures that are found within molecular clouds.

1.2 Molecular clouds and their substructures

Numerous processes lead to the formation structures within the ISM, and although transient with fuzzy boundaries they generally are classified by their morphology and physical properties into a nested hierarchy of structures.

1.2.1 Molecular clouds

The largest structures within the CMM are molecular clouds, comprised of cold, dense molecular (and some atomic) gas. Most molecular clouds have masses of around $10^3\,M_\odot$ to $10^5\,M_\odot$, with the most extreme massive clouds, known as Giant Molecular Clouds (GMCs), containing as much as $10^7\,M_\odot$ (Miville-Deschênes et al.

2017)[1]. These cloud structures can range anywhere from a few to a few hundred parsecs across. Around half of the molecular gas in the Milky Way is contained within the ~ 460 clouds with $M > 8.4 \times 10^5\,\mathrm{M}_\odot$, derived from CO observations (Miville-Deschênes et al. 2017), with around half of all ionising radiation coming from the 18 most luminous *WMAP* sources, largely attributed to high-mass star clusters (Murray & Rahman 2010). Similarly, the 30 most massive structures identified ($\sim 16\%$ of the sample) in the ATLASGAL 870 µm continuum survey of the inner ($-80\,\mathrm{deg} < l < 60\,\mathrm{deg}$) Milky Way account for around 36% of the mass of clouds, and 52% of the total bolometric luminosity Urquhart et al. (2018). The high density of MCs shield the bulk of the material within them from being heated, ionised and dispersed by the surrounding radiation field. Dust emission and metal line emission facilitate cooling of the gas within them, leading to temperatures of around 10 K for the bulk of the gas in MCs, and sometimes as low as 5 K (Glover & Clark 2012). This temperature floor is largely set by heating due to cosmic rays and the CMB.

There are a few processes which are believed to form MCs, but generally all of them involve the supersonic compression of gas. Such sources of compression include the sweeping up of material by stellar winds, HII regions, supernova events, and turbulence. Larger molecular clouds are thought to be assembled from smaller clouds at the bottom of the potential well generated by stars, as they move through spiral arms (Duarte-Cabral & Dobbs 2017).

Despite H_2 being the most abundant element in the CMM phase, most of the information gathered about the physical properties of molecular clouds have been obtained through observations of carbon monoxide, CO, and its isotopologues. This is because the temperature of most of the gas in the CMM is too cold to excite H_2, a symmetric molecule requiring temperatures upward of ~ 500 K to exite its rovibrational transitions (Roueff et al. 2019), means it is near impossible to observe directly. Many surveys have been conducted to fully map the Galactic Plane to study the molecular clouds that reside within them (e.g. Dame et al. 2001; Jackson

---

[1] This value is rather extreme, which may in part be due to the method used to assign distances to the clouds in this study. A more reasonable maximum cloud mass is likely an order of magnitude lower.

et al. 2006; Dempsey et al. 2013; Rigby et al. 2016; Schuller et al. 2017). The properties of the molecular clouds identified within these surveys are typically inferred using a conversion factor that converts the brightness in CO to a column density of H_2. There are some uncertainties involved with this procedure, as this assumes that there is a constant abundance ratio of CO to H_2, and often temperature affects are not accounted for. Synthetic observations of molecular clouds has shown that the estimated cloud masses can be uncertain by a factor of two (Szűcs et al. 2016). Observationally, the column density does not follow a linear relationship with integrated intensity, especially when considering optical depth effects (Barnes et al. 2015), and on small scales it has been shown to be a poor tracer of the total column density (Lewis et al. 2022).

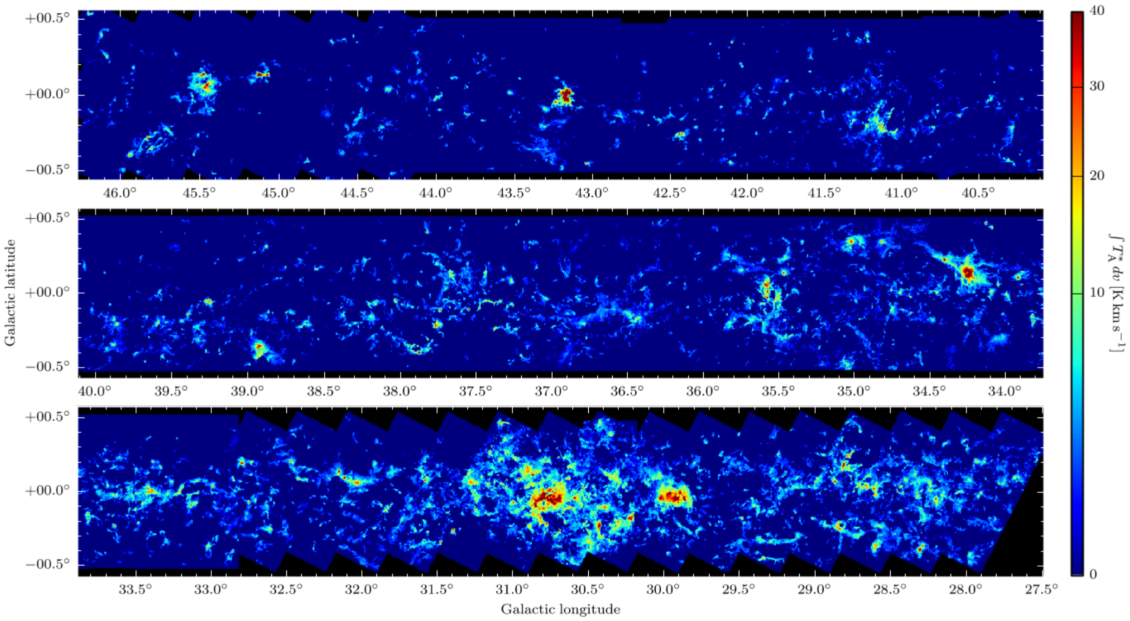

Figure 1.4: ^{13}CO(J=3–2) integrated intensity map of the Galactic Plane from the CHIMPS survey (Rigby et al. 2016).

Using a range of different tracers, Larson (1981) found a tight correlation between the velocity dispersion σ and a cloud's size L, with $\sigma \propto L^{0.38}$. This power-law index of $\gamma = 0.38$ is very similar to Kolmogorov-type turbulence which scales as $\sigma \propto L^{1/3}$, leading to the interpretation that molecular clouds are turbulent structures. This has been further investigated, with a study of ~ 270 clouds by Solomon et al. (1987) finding $\sigma \propto L^{1/2}$, steeper than the $\gamma = 0.38$ found by Larson (1981). Heyer et al.

(2009) re-examined the Solomon et al. (1987) sample with improved observations, but found no Larson relationship, and found that that the ratio of $\sigma/L^{1/2}$ varies systematically with a cloud's mass surface density as $\sim \Sigma^{1/2}$, compatible with the clouds being self-gravitating structures.

The virial parameter, which is the balance of kinetic to gravitational potential energy, is a measure that is often used to discuss whether molecular clouds are gravitationally bound structures, with $\alpha_{\mathrm{vir}} \gtrsim$ 2–3 defining the boundary where clouds are unbound (depending on the density profile). However, using simulations Dobbs et al. (2011) found that in fact most of the molecular gas is gravitationally unbound, however substructures within the denser parts of the clouds are bound and able to form stars. More recent analysis of the Dame et al. (2001) ^{12}CO(J=1–0) survey of the Galactic Plane by Miville-Deschênes et al. (2017) seem to agree with this picture, finding that finding that only 15% of the $\sim$ 8100 identified molecular clouds (containing 40% of the mass) have $\alpha_{\mathrm{vir}} \leq 3$. Analysis of the much higher resolution SEDIGISM survey of the inner Galactic Plane, observing the optically thinner tracer ^{13}CO(J=2–1), catalogued over 10,000, and found most clouds to have $\sigma/L^{0.52}$ and $\alpha_{\mathrm{vir}} \leq 2$. The differences between these results is likely driven by a few factors, firstly the latter study uses a slightly less extended tracer, and hence is probably tracing more gravitationally bound material, secondly the SEDIGISM survey is much higher resolution ($28''$) than the Dame et al. (2001) ^{12}CO(J=1–0) survey ($8.5'$), and lastly, the method of cloud of cloud extraction and analysis differ between the studies, which will also impact their results.

A topic interest has been whether molecular cloud properties vary with Galactic environment. Ragan et al. (2018) found that while the Milky Way's spiral arms host some of tha Galaxy's most massive clusters, there seems to be very little enhancements to the overall star formation fraction associated with spiral arm structure. Duarte-Cabral et al. (2020) also only saw marginal enhancements in some of the properties (mass, velocity dispersion, high-mass star-formation) of the most extreme clouds with environment. However, there is some evidence that clouds in the Central Molecular Zone (CMZ) are different from the rest of the Galaxy. Longmore et al. (2013) showed that the star formation rate is an order of magnitude lower than predicted, even though there is a large reservoir of dense molecular gas available, possibly due to the increased turbulence measured in CMZ clouds relative the rest of the MW. This could be supported by the shape of the column density PDF observed in G0.253+0.016, a molecular cloud located in the CMZ. Rathborne et al. (2014) found that the column density PDF shows a log-normal distribution at the top end, a signature that the cloud is dominated by turbulence, whereas star-forming clouds/clouds dominated by gravity have column density PDFs with power-law like wings (Kainulainen et al. 2009).

The cold temperatures within their centres, and hence lack of thermal support, leads to the creation of overdensities through gravitational instabilities. These overdensities lead to local fragmentation and collapse into structures such as clumps (parsec-scale dense molecular cloud structures), filaments, and in turn prestellar cores.

1.2.2 Molecular cloud substructures

Within GMCs and MCs, overdensities exists that are a few parsecs in size, are commonly known as clumps. They span a broad range of masses ($\sim 10^2\,\mathrm{M}_\odot$ to $10^4\,\mathrm{M}_\odot$), and typically have density profiles that follow $\rho(r) \propto r^{-2}$ (Williams et al. 2000). They are formed through a combination of both supersonic turbulence and self-gravity, and are thought to be the precursors of star clusters (Williams et al. 2000; Krumholz et al. 2019).

One particular class of clumps are Infrared dark clouds (IRDCs)[2]. IRDCs are cold ($\sim$10 K to 20 K), high column density ($\sim 10^{22}\,\mathrm{cm}^{-2}$) molecular clouds seen as a silhouette against the diffuse IR ($\sim$7 µm to 25 µm) background emission of the Galactic Plane (Perault et al. 1996; Egan et al. 1999; Simon et al. 2006; Ragan et al. 2009; Peretto & Fuller 2009, 2010). Their low temperatures and lack of bright infrared dust emission, suggest that they harbour systems (protoclusters) in the earliest stages of star formation. They cover a broad range of sizes, masses and morphologies, and likely lead to a diverse range of different star-forming regions.

---

[2] IRDCs are not exclusively clump-sized objects (i.e. many molecular clouds are IR-dark), but the work presented in this thesis largely focuses on clump- and smaller-scale objects.

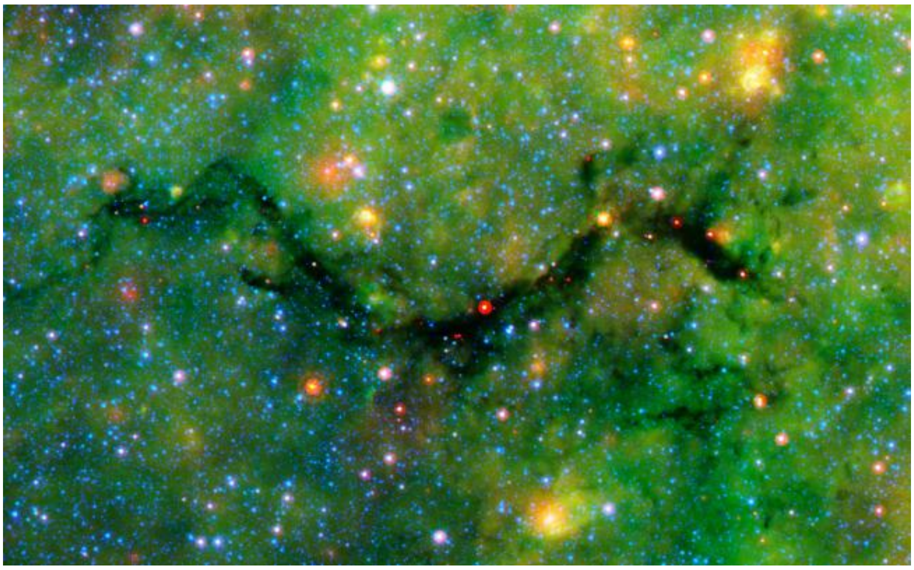

Figure 1.5: *Spitzer* 3-colour image of the IRDC called "the Snake", showing the high density filamentary features in extinction. Credit: NASA JPL-Caltech/S. Carey.

Filaments are high-aspect ratio structures comprised of gas and dust (Schneider & Elmegreen 1979; Goldsmith et al. 2008). Filamentary structures are seen on all scales throughout the Universe, but here we will focus on the filaments found within our galaxy, specifically ones found within the CMM. A significant fraction (around 50% to 70%) of the mass in the form of dense gas within MCs appear to be found in filaments (Könyves et al. 2015). These structures can be as long as $\sim 100\,\mathrm{pc}$, such as the "Nessie" filament (Jackson et al. 2010), or the Giant Molecular Filaments presented by Ragan et al. (2014) containing as much as $10^4\,\mathrm{M}_\odot$ to $10^5\,\mathrm{M}_\odot$ located within the inter-arm regions of the Milky Way.

Observations from *Herschel* (at 70 µm to 500 µm) of molecular clouds in the Galactic Plane revealed countless filaments on the scale of a few parsec. Analysis of observations from *Herschel* have shown that filaments represent a key stage in the star formation process, and are ubiquitous with star forming clouds, with the vast majority ($\sim 75\%$) of pre- and protostellar cores being embedded within dense filaments (Arzoumanian et al. 2011; Polychroni et al. 2013; Könyves et al. 2015). These parsec-scale filaments typically have sub-parsec widths, around 0.1 pc, which somewhat controversially has been proposed as a "characteristic width" (Arzoumanian et al. 2011; André et al. 2016), whereas others have found the distribution of filament widths to be much broader, and suggest that this measured characteristic width is an artefact of the methodology or tracers used (Panopoulou et al. 2017). Some have shown that these parsec-scale filaments are in fact composed of bundles of smaller, trans-sonic, filamentary structures. Termed “fibers”, with widths as low as $\sim 0.02\,\mathrm{pc}$, it has been suggested that the concentration of these fibers may determine the masses of the stars that form within them (e.g. Hacar et al. 2017, 2018). What is clear from studies of filaments, is that they set the initial conditions for the formation of cores.

1.3 Star formation

1.3.1 Low-mass star formation

Cores, which are the smallest substructure in the star formation process, are gravitationally bound fragments of dense gas that are between 0.01 pc to 0.1 pc in size, and have masses from around $0.1\,\mathrm{M}_\odot$ to tens of solar masses. They are denser ($n > 10^5\,\mathrm{cm}^{-3}$) than the structures that they are embedded within (Andre et al. 1993; Ward-Thompson et al. 1994), and hence are well shielded from external sources of radiation and hence have extremely cold interiors ($\sim 10\,\mathrm{K}$) in the earliest stages of their lifetimes. Whereas clumps are the precursors of stellar clusters, cores are the precursors of individual stellar systems (Andre et al. 2000).

In broad terms core can be split into two types: pre-stellar and protostellar. A prestellar core is a gravitationally bound overdensity within a clump/filament which does not yet contain a protostellar object in hydrostatic equilibrium (i.e. internal pressure balances gravitational collapse) within it (Andre et al. 2000; Ward-Thompson & Whitworth 2011). Pre-stellar cores are very cold, and so they emit almost all of their radiation at far-IR and longer wavelengths. Observationally, the boundary that defines whether a core is pre-stellar or protostellar is whether a core is emitting a lot of radiation in the mid-IR. For example, *Herschel* 70 µm emission is often used to classify more evolved cores.

However, the lack of this emission is not a guarantee that a core is in its pre-stellar phase, as observations are often limited in their sensitivity and resolution, and many of these cores are embedded in obscured, high opacity regions. In fact high-resolution ALMA observations of so-called "starless" cores (typically pre-stellar due to lack of shorter wavelength emission) have shown the presence of outflow signatures, an indication that these sources may be much more evolved (i.e. protostellar) than they appear in lower resolution dust continuum surveys (Svoboda et al. 2019).

A quantity that is used to decide whether a parcel of gas is bound is the Jeans mass (Jeans 1928). In the case of a spherically symmetric parcel of gas of mean density ρ, it must have a mass that exceeds the Jeans mass in order to collapse (in the absence of any support against gravity). This is given by

$$M_{\mathrm{J}} = \frac{4\pi}{3}\frac{c_{\mathrm{s}}^3}{G^{3/2}\rho^{1/2}} \tag{1.2}$$

where G is the gravitational constant, and c_{s} is the isothermal sound speed of the gas at temperature T, which is given by:

$$c_{\mathrm{s}} = \sqrt{\frac{k_{\mathrm{B}}T}{\mu m_{\mathrm{H}}}} \tag{1.3}$$

where m_{H} is the mass of a hydrogen atom, μ is the mean molecular weight (per hydrogen atom), and k_{B} is the Boltzmann constant.

The Jeans mass at $n = 10^4\,\mathrm{cm}^{-3}$ and $T = 10\,\mathrm{K}$ is around $1\,\mathrm{M}_{\odot}$, which is comparable to the typical core mass within nearby star-forming clouds (André et al. 2014). The fact that most cores are found to lie on top of dense filaments strongly suggests that their formation is related to the gravitational instability running along those filaments (Polychroni et al. 2013).

Assuming that there is no support against gravity, this parcel of gas with $M > M_{\mathrm{J}}$ will collapse in a free-fall time, which is the time it takes for a completely static, spherically symmetric parcel of gas to collapse into a single point. The free-fall time is often used as a reference timescale in star formation research. This is given by:

$$t_{\mathrm{ff}} = \sqrt{\frac{3\pi}{32G\rho}} \tag{1.4}$$

For a parcel of gas with a density of $n = 10^4\,\mathrm{cm}^{-3}$, its corresponding t_{ff} is around $3 \times 10^5\,\mathrm{yr}$.

Understanding what physical processes determine the mass of stars is an active area of astrophysics research. The similarity between the shape of the mass distribution of prestellar cores identified in nearby star-forming regions and that of the initial mass function of stars suggests that the latter may be inherited from the former, with a one to one correlation between core and stellar masses, and a uniform core to star formation efficiency across all core masses of $\sim 30\%$ (e.g. Motte et al. 1998; Johnstone et al. 2001; Nutter & Ward-Thompson 2007; Könyves et al. 2010, 2015). The determination of stellar masses via core accretion is often referred to as *core-fed* accretion (e.g. Wang et al. 2010). As a result of the analysis of *Herschel* observations of Gould belt star-forming regions (André et al. 2010), it has been proposed that the mass of cores is, in turn, determined by the fragmentation of gravitationally unstable filaments whose local Jeans mass is $\sim 1\,\mathrm{M}_\odot$, i.e. the peak of the core mass function in these regions (André et al. 2010, 2014, 2019; Roy et al. 2015).

1.3.2 High-mass star formation

High mass stars are defined as having around $8\,\mathrm{M}_\odot$ or more (i.e. they are the progenitors of supernovae). Although accounting for a small fraction of all stars by number (around 1%), the luminosity of a galaxy is dominated by the emission from high-mass stars, as the stellar luminosity is given by

$$L_* \propto M_*^{3.5} \tag{1.5}$$

where M_* is the stellar mass. The sheer amount of emission from these stars inject vast amount of energy into the ISM over their short lifetimes, which is inversely proportional to the star's mass:

$$\tau_{\mathrm{lifetime}} \propto M_*^{-2.5} \tag{1.6}$$

At the end of their lives they explode, altering the large scale structure of galaxies and disperse heavier elements across the ISM. Despite their importance and far-reaching influence on their environment, the formation of high-mass stars is not nearly as well understood as the formation of their low-mass counterparts. This is

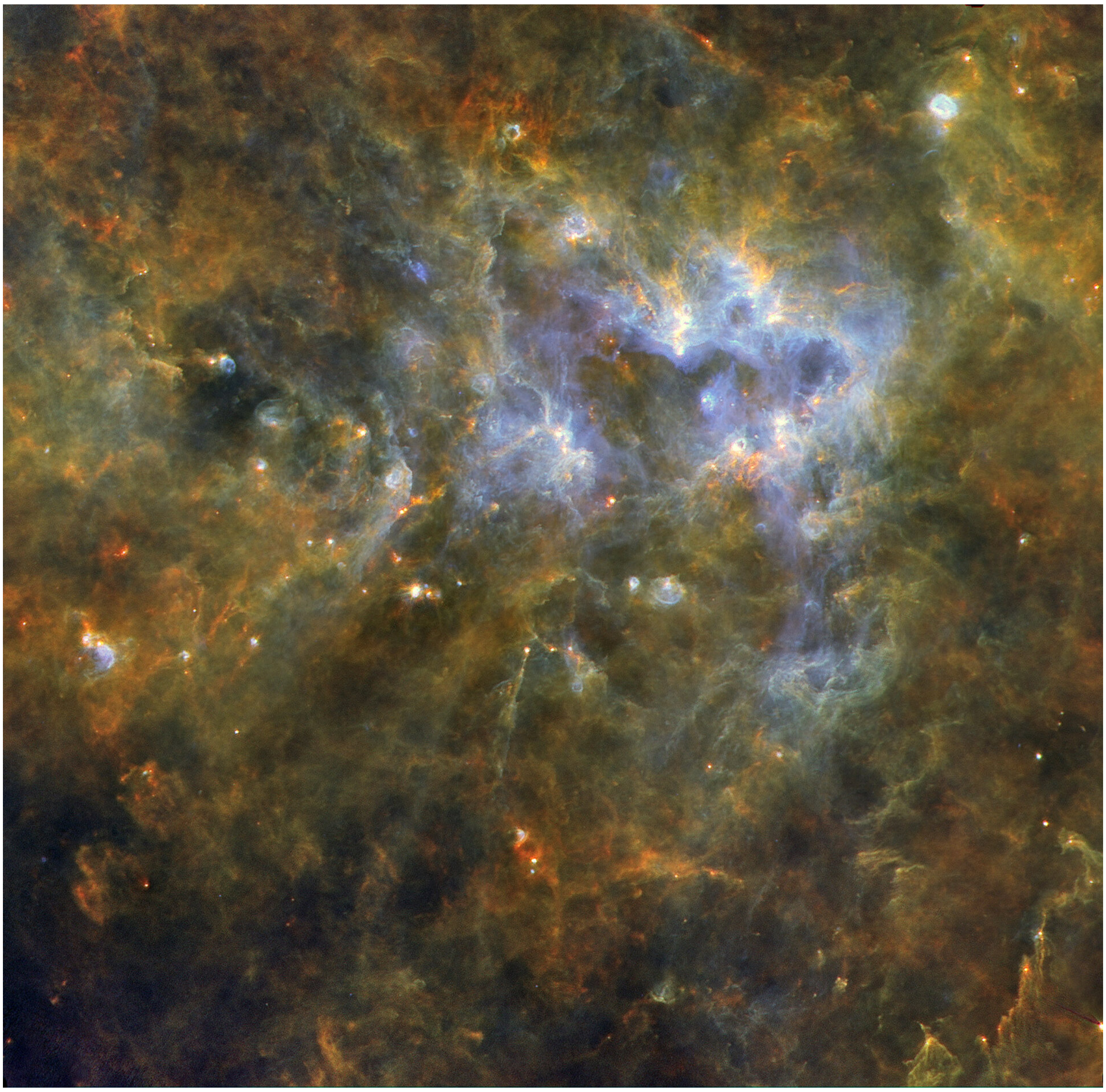

Figure 1.6: *Herschel* 3-colour image of the G305 complex from the Hi-GAL survey. Red shows the cooler, dense gas observed at 250 µm, the centre of the complex harbours the high-mass star WR48a and the clusters Danks 1 and 2. Their strong stellar winds push away and heat up the gas in the neighbourhood, which is seen in blue 70 µm emission. Credit: ESA/Herschel/Hi-GAL.

partly driven by their rarity, their short lifetimes and evolutionary timescales, and the fact that they are embedded within high dust extinction, crowded environments during the early phases of their formation. Nearby star-forming regions also seem to mostly host the formation of low-mass stars, meaning that the majority of observations of these regions are inherently limited in their spatial/physical resolution.

Understanding how high-mass stars are formed is crucial to our general understanding of star (and planet) formation, the structure and evolution of galaxies, and also the history and future of the Universe as a whole.

While the scenario described in Section 1.3.1 might be relevant for determining the masses of low-mass cores and stars, it seems rather inappropriate when it comes to the formation of the most massive stars ($M_\star > 8\,\mathrm{M}_\odot$). The most massive prestellar cores identified in the far-infrared and sub-millimetre surveys of Gould Belt regions are typically about $10\,\mathrm{M}_\odot$ (Könyves et al. 2015, 2020), implying a stellar mass of about $3\,\mathrm{M}_\odot$ when accounting for the core to star formation efficiency derived by the same authors. Much more massive prestellar cores, typically $30\,\mathrm{M}_\odot$ and above, would need to be found in order to form massive stars in a core-fed-type scenario.

Searches for such massive cores have now failed to find a significant population (e.g. Motte et al. 2007; Svoboda et al. 2019; Sanhueza et al. 2019), and as of today only a few exceptional cases are known (e.g. Cyganowski et al. 2014; Nony et al. 2018), despite an ever increasing database of high-angular resolution observations of cold and compact sources. In observations of high-mass star forming regions, the core mass functions (CMFs) constructed from extracted cores have been shown to be largely top-heavy (Zhang et al. 2015; Motte et al. 2018a; Pouteau et al. 2022; Nony et al. 2023; Pouteau et al. 2023), with the bulk of the higher mass cores being protostellar as opposed to pre-stellar, further implying that a core-fed scenario is unlikely for high-mass star formation.

Recently, Peretto et al. (2020) have used (sub-)millimetre dust continuum observations of Galactic plane star-forming regions to show that the evolution of massive compact sources ($m_{\mathrm{gas}} > 30\,\mathrm{M}_\odot$) in mass vs. temperature diagrams is better explained by an accretion scenario in which cores gain mass while simultaneously collapsing to form protostars. In a similar manner, (Rigby et al. 2021) find evidence for the mass growth of clumps, suggesting that same accretion processes may occur over a wider range of scales. The mass growth of the core is believed to be the result of the collapse of the surrounding parsec-scale mass reservoir called *clump*, hence the accretion scenario described above is referred to as *clump-fed* (Wang et al. 2010).

The results from Rigby et al. (2021) suggest that there must be a link between the properties of a clump and the stars that form within it. Such a link has been searched for in the past. For instance, Palau et al. (2014, 2021) found a correlation between the fragmentation level within massive 0.1 pc-size cores and their average volume density, as expected from Jeans instability. On larger scales, Barnes et al. (2021) found a similar result, larger parsec-size clouds having larger number of cores embedded within them. They also find a correlation between the cloud mass and the mass of its most massive core. The existence of such a relation has also been explored by Lin et al. (2019), who found a tight correlation between the mass of a sub-sample of massive ATLASGAL clumps and the mass of the most massive fragment they identify on SABOCA 350 µm continuum images. However, the small difference in angular resolution between LABOCA (18″) and SABOCA (8.5″) might play a significant part in driving the observed correlation. On the other hand, Urquhart et al. (2014) argued that clumps with signposts of active massive star formation are more spherical than those which do not have such associated tracers, while Rigby et al. (2018) suggested that more spherical clumps are more efficient at concentrating their mass within their most massive core. These studies suggest that a combination of clump mass and morphology might be important parameters for the formation of massive stars.

1.3.3 Hub-filament systems

Here, we focus on a specific morphological category of MCs: hub-filamentary systems (HFS) (Myers 2009). Hubs are small networks of converging interstellar filaments, at the centre of which active star formation is often observed (e.g. Liu et al. 2012; Kirk et al. 2013; Peretto et al. 2013, 2014; Treviño-Morales et al. 2019). They are found in all types of region, from low-mass star-forming clouds (e.g. Myers 2009; Kirk et al. 2013), to high-mass star-forming regions (e.g. Peretto et al. 2013; Schwörer et al. 2019), with the most luminous ($L \geq 10^5\,\mathrm{L}_\odot$) young stellar objects within the Galaxy are systematically embedded within HFSs (Kumar et al. 2020; Peretto et al. 2022). They have even been observed in our closest neighbouring galaxy (Fukui et al. 2019; Tokuda et al. 2019). They are therefore a structure of global interest when it comes to star formation theories.

The formation mechanism of such hubs are not yet fully understood. However, in the past 15 years or so, a large number of different scenarios have been proposed to explain their formation. Most of them start with a compression phase whereby a sheet of gas is formed first, soon followed by filament formation (e.g Nakamura & Li 2008; Myers 2009; Wang et al. 2010; Gómez & Vázquez-Semadeni 2014; Balfour et al. 2015; Inoue et al. 2018; Padoan et al. 2020). The origin and associated physical scale of the compression differ according to the models, along with the physics that lead to the formation of filaments. The convergence of those filaments is either linked to the formation of the sheet itself (Myers 2009), the result of gravity that drags filaments towards the bottom of the potential well (e.g. Li & Nakamura 2006; Vázquez-Semadeni et al. 2019), or the consequence of multiple turbulence-driven sheet-sheet (e.g. Inutsuka et al. 2015; Federrath 2016; Padoan et al. 2016) or filament-filament (Kumar et al. 2020) interactions. In those models, star formation naturally occurs at the convergence point of the filaments, i.e. the hub, since gas density is the largest there. One of the key differences between those models is the physical connection between the filaments and central hub, with, in particular, one fundamental question that still needs answering: Do filaments regulate the mass growth of their central hub and the cores within it?

In parallel to the theoretical effort mentioned above, numerous observational studies have investigated the density structure and gas kinematics of hub-filament systems (e.g. Liu et al. 2012; Peretto et al. 2013; Kirk et al. 2013; Busquet et al. 2013; Peretto et al. 2014; Dewangan et al. 2017; Williams et al. 2018; Treviño-Morales et al. 2019; Arzoumanian et al. 2021; Zhou et al. 2022; Liu et al. 2023). Interestingly, the vast majority of those studies interpret their data in the context of collapse scenarios, whereby filaments feed the central hub via gravity-driven mass inflows. Evidence for such scenarios is based on the presence of velocity gradients along filaments, and the mass segregation of cores within the hub. Williams et al. (2018) argued that the centres of hub filament systems, where the filaments converge, are privileged locations of massive core formation as they correspond to the locations of maximum gradient of gravitational acceleration, as opposed to individual uniform density filaments where these are located at their ends (e.g. Hartmann & Burkert

2007; Clarke & Whitworth 2015).

However, only very few studies have the necessary angular resolution, spatial coverage, and sensitivity to follow the gas flows along individual filaments and down to the hub. For instance, with more than 100 clumps observed with ALMA at $\sim 2''$ resolution, the ATOMS project (Liu et al. 2020) statistically investigated the kinematics of HFSs (Zhou et al. 2022; Xu et al. 2023). However, by focusing on the hub, the connection with the surrounding filaments is only partially addressed. Most of the other studies have mapped entire HFSs at similar angular resolution, but targeting only single sources that are often selected for their ability to form high-mass stars (e.g. Peretto et al. 2013; Beltrán et al. 2022), preventing them from drawing general conclusions on the physical link between filaments and hubs.

The key questions that the work in this thesis attempts to answer regard the properties of HFSs and the cores that are formed within them. Firstly, are HFSs distinct from the Galactic clump population regarding core formation, and are HFSs preferable sites for the formation of high-mass cores? Secondly, what are the kinematic properties of HFSs, how do the kinematic properties of of the filaments vary along their lengths, and is there a clear transition from filament to hub? Finally, how do the environmental conditions vary around the cores found in HFSs, and do the physical properties of the cores vary/correlate with changes in the local environment?

1.4 Thesis structure

In Chapter 2, I outline the fundamentals of single-dish observations at millimetre/radio wavelengths, and also describe the principles behind radio interferometry and data reduction. In Chapter 3 I present a sample of six HFSs, and the ALMA 2.9 mm continuum and N_2H^+(J=1–0) emission line data that was acquired to study them as part of the work presented in this thesis. In Chapter 4 I present a study of cores within this sample of six HFSs using the continuum part of the aforementioned ALMA data, and compare them with a broader sample of cores/clumps from literature. In Chapter 5, I present a study of the six HFSs using N_2H^+(J=1–0) to trace the kinematic properties filamentary and dense gas structures within them. This study required the development of a brand new, multiple velocity component, hyperfine line-fitting code, which is also described in this chapter. Finally, in Chapter 6, I summarise the main findings and conclusions of this thesis, present some techniques that will be used as part of a deeper study of the filamentary structures within these HFSs, and outline future work that, using observations of a larger sample of clumps and synthetic observations, will hopefully shed more light on the nature of HFSs and their role in Galactic star formation.

Chapter 2

Radio interferometry

2.1 Radio telescope fundamentals

The Earth's atmosphere acts as a barrier to observing the universe, with emission in many different wavelength regimes being absorbed before reaching the surface. To observe at these wavelengths, it is often necessary to build space observatories. However, there are two privileged wavelength windows where most of the light is not absorbed by the atmosphere, the optical/near-IR, and the radio regime ($200\,\mu\mathrm{m} \lesssim \lambda \lesssim 30\,\mathrm{m}$)).

Radio telescopes that operate at the shorter end of the wavelength range ($\lambda \lesssim 30\,\mathrm{cm}$) are usually in the form of some parabolic dish comprised of mesh or solid metal panels. This parabolic shape means that the telescope is highly directional, and reflects incoming light onto a receiver, where the signal is amplified and can then be processed. The receivers of radio telescopes are often cryogenically cooled to reduce thermal noise and enabling the detection of incredibly weak incoming signals.

For a telescope pointed in a given direction, the spectral flux density S_ν is defined as the amount of incident energy over a given area, per second, per unit frequency (i.e. with SI units of $[\mathrm{W\,m^{-2}\,Hz^{-1}}]$). As radio astronomy deals with incredibly faint signals, the spectral flux density is more commonly measured in units of Janskys (Jy), where $1\,\mathrm{Jy} = 10^{-26}\,\mathrm{W\,m^{-2}\,Hz^{-1}}$. Given that astronomical sources have a finite size, their incoming spectral flux density is usually measured per unit solid angle, called the specific intensity $I_\nu(\theta, \phi)$ (sometimes referred to as a brightness), which is related to S_ν by:

$$S_\nu = \int_{\Omega_\mathrm{s}} I_\nu(\theta, \phi)\mathrm{d}\Omega \tag{2.1}$$

where Ω_s is some solid angle over which the spectral flux density is being measured, typically the solid angle subtended by a source. The units of I_ν are often shown in [$\mathrm{MJy\,sr^{-1}}$], or in the case of radio interferometry in [$\mathrm{Jy\,beam^{-1}}$], which will be discussed in the subsequent section.

For a region that is in thermal equilibrium at temperature T, the specific intensity as a function of frequency ν (or wavelength λ) of the radiation field is described by the Planck function:

$$\begin{aligned} B_\nu(T) &= \frac{2h\nu^3}{c^2}\frac{1}{\exp(\frac{h\nu}{k_\mathrm{B}T}) - 1} \\ B_\lambda(T) &= \frac{2hc^2}{\lambda^5}\frac{1}{\exp(\frac{hc}{\lambda k_\mathrm{B}T}) - 1} \end{aligned} \tag{2.2}$$

where h is the Planck constant, c is the speed of light in a vacuum, and k_B is the Boltzmann constant. For low frequency emission (i.e. long wavelength), the Planck function can be approximated by a simpler expression. When $h\nu \ll k_\mathrm{B}T$ (or $hc/\lambda \ll k_\mathrm{B}T$), an expansion of the exponential term yields

$$\begin{aligned} \exp\left(\frac{h\nu}{k_\mathrm{B}T}\right) &\approx 1 + \frac{h\nu}{k_\mathrm{B}T} + ... \\ \exp\left(\frac{hc}{\lambda k_\mathrm{B}T}\right) &\approx 1 + \frac{hc}{\lambda k_\mathrm{B}T} + ... \end{aligned} \tag{2.3}$$

which can then be substituted into equation 2.2 to yield the Rayleigh–Jeans (R–J) approximation:

$$\begin{aligned} B_\nu^\mathrm{R-J}(T) &= \frac{2\nu^2}{c^2}k_\mathrm{B}T \\ B_\lambda^\mathrm{R-J}(T) &= \frac{2c}{\lambda^4}k_\mathrm{B}T \end{aligned} \tag{2.4}$$

given that radio telescopes operate within this long wavelength regime, it is often convenient to use this approximation. See Figure 2.1 for an illustration of the accuracy of this approximation as a function of wavelength.

The R–J approximation also serves to define a frequently used quantity known as the brightness temperature T_B (in units of [K]), which is related to the specific intensity by the equation:

$$I_\nu = \frac{2\nu^2}{c^2}k_\mathrm{B}T_\mathrm{B} \tag{2.5}$$

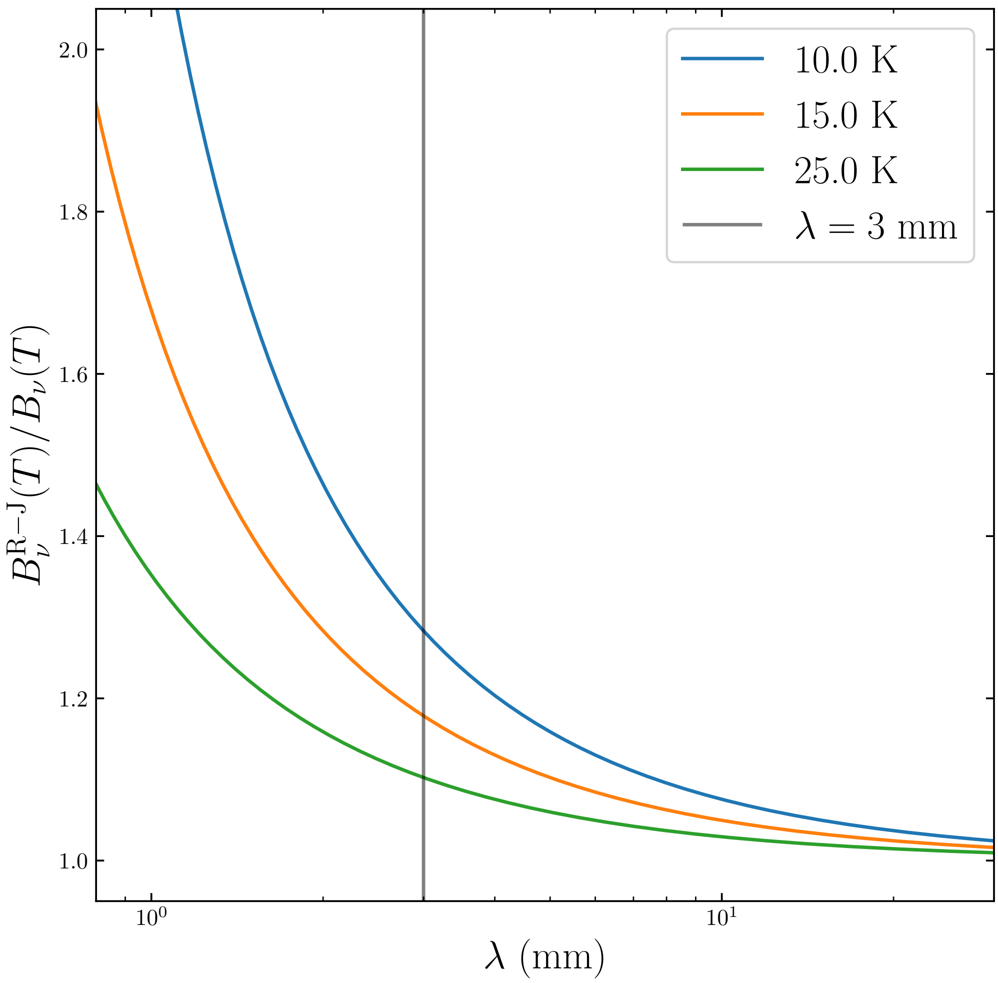

Figure 2.1: Plot showing how much the Rayleigh–Jeans approximation overestimates the spectral flux density of a blackbody as a function wavelength, for a set of temperatures. A vertical line is shown marking $\lambda = 3\,\mathrm{mm}$, where $B_\nu^{\mathrm{R-J}}$ overestimates B_ν by around 18% (for $T = 15\,\mathrm{K}$).

which effectively corresponds to the specific intensity emitted by a blackbody (in thermodynamic equilibrium) at temperature T_B.

An idealised, circular aperture telescope has and angular resolution θ (in [rad]) given by the equation:

$$\theta \approx 1.220\frac{\lambda}{D} \tag{2.6}$$

where λ corresponds to the observed wavelength, and D is the diameter of the aperture (both in [m]). This resolution element, θ, corresponds to the FWHM of the main beam of the antenna response, which is illustrated in Figure 2.2.

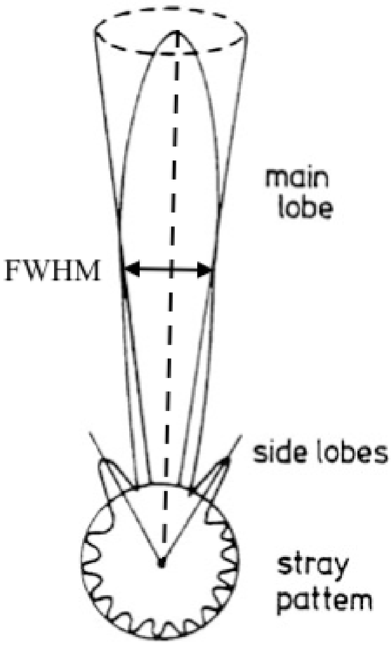

Figure 2.2: An illustration of the antenna response power pattern, with the majority of the sensitivity directed toward the source direction in the "main lobe". Adapted from Wilson et al. (2013).

2.2 Radio interferometry

While it is possible to improve the angular resolution by observing at shorter wavelengths, often this will lead to the emission probing different structures, or different astronomical phenomena altogether. Therefore, it is usually necessary to increase the diameter of the telescope aperture to attain higher resolution observations. This, however, presents its own challenges. As a telescopes mirror/dish increases in size, the harder it is to maintain its structural integrity and maintain a high degree of surface accuracy. Even if one could construct an arbitrarily large telescope dish, it would become increasingly complex and expensive to engineer a method to steer/point the telescope to a particular point in the sky with sufficient accuracy over the course of an exposure.

Fortunately, interferometry is a technique that makes it possible reach significantly higher angular resolutions without the need manufacture single, gigantic telescopes. An array of many precisely positioned single-dish telescopes spread over a large area work in unison to observe the same target, with the data from each telescope carefully combined to produce a higher resolution image than any of the individual telescopes in the array. Each pair of telescopes are separated from each other by a distance b, called a baseline. The angular resolution of an intereformeter can be approximated by the equation:

$$\theta \approx k\frac{\lambda}{b_{\mathrm{max}}} \tag{2.7}$$

where b_{max} is the maximum physical separation between telescopes in the array, also known as the maximum baseline, and k is a factor of order unity that is dependent on how measurements from the telescope are weighted during data reduction. The consequence of this is that it is possible to create an astronomical intereformeter of almost any size, so long as the telescopes positions are stable relative to each other, and measurements from them can be accurately synchronised together.

Unfortunately, this astonishing increase in resolution comes with some limitations. Firstly, an interferometer simply has less collecting area than a filled aperture telescope with a diameter equal to the longest baseline. Secondly, because an interferometer cannot sample spatial frequencies smaller than a baseline roughly equal to an antenna diameter, it will be insensitive to large scale variations (e.g. extended emission) of a target source. This is termed the zero-spacing problem. However, this can be somewhat accounted for by combining interferometric data with single-dish data using a variety of techniques. The maximum recoverable scale (MRS), is approximately given by:

$$\theta_{\mathrm{MRS}} \approx 0.6\frac{\lambda}{b_{\mathrm{min}}} \tag{2.8}$$

where b_{min} is the minimum baseline length in the antenna array.

The field of view of an in interferometer is a circle with a diameter given by the FWHM of the primary beam of one of the constituent antenna in the array. So to observe a larger patch of the sky “mosaicking” is required. The response of the primary beam will decrease towards the edges (i.e. will be less sensitive), so when creating a mosaic it is important to ensure sufficient overlap between these single pointings, and hence maximise the area of high-sensitivity in the field. The optimal method it to place the centres of each field on the vertices of a triangular grid spaced

$\lambda/(\sqrt{3}D)$ apart, where D is the antenna diameter.

2.2.1 Two-element interferometer

Here we will describe a simple two-element intereformeter, comprised of two single-dish telescopes x and y (see Figure 2.3). They are both fixed on a level plane, separated by a distance (or baseline) $\mathbf{b}$. Both telescopes are pointed at a distant, monochromatic source emitting a plane electromagnetic wave, which is in the direction denoted by the unit vector $\hat{\mathbf{s}}$, which is at an angle θ_s relative to the plane[1]. The signal from an observed sources arrives at antenna x first at time t_x, and later arrives at antenna y after a brief delay at time $t_y = t_x + \tau_{\text{geo}}$, where τ_{geo} is called the

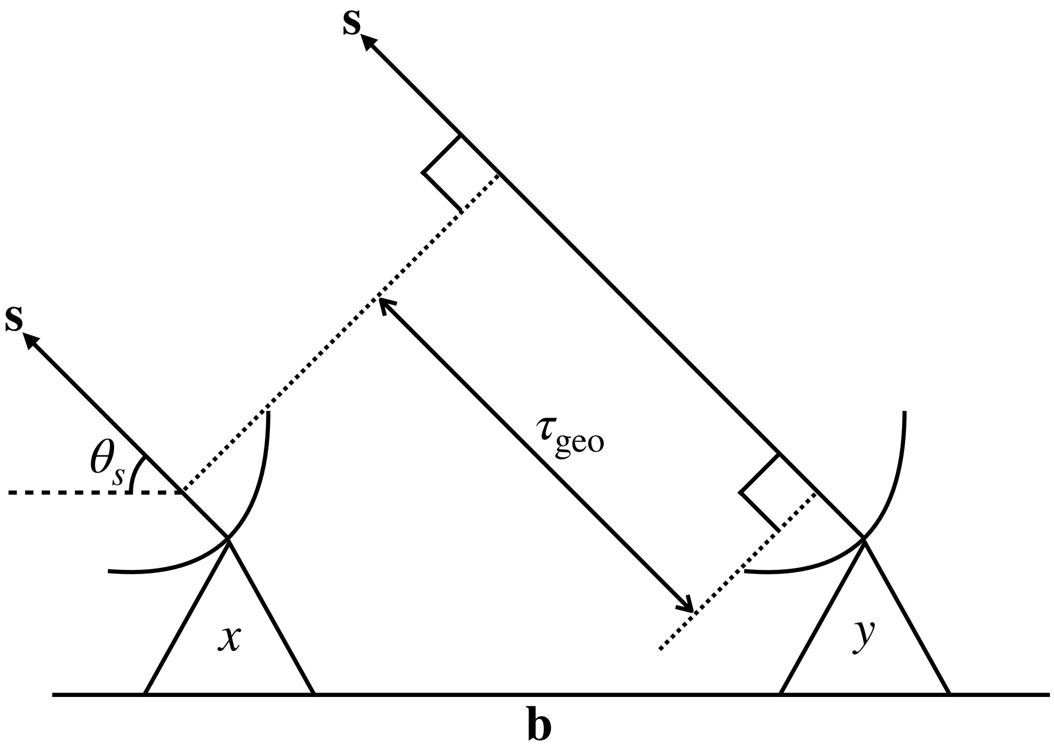

Figure 2.3: Schematic of a two-element interferometer.

---

[1] Each telescope are actually pointed in slightly different directions $\hat{\mathbf{s}}_x$ and $\hat{\mathbf{s}}_y$ (at angles $\theta_{s,x}$ and $\theta_{s,y}$ above the plane), respectively, but given that the astronomical source is significantly further away than the separation between the two telescopes ($b \ll s_x; b \ll s_y$), then $\hat{\mathbf{s}}_x \cdot \hat{\mathbf{s}}_y \to 1$, and therefore $\hat{\mathbf{s}}_x \sim \hat{\mathbf{s}}_y \equiv \hat{\mathbf{s}}$.

geometric time delay. This is given by:

$$\tau_{\text{geo}} = \frac{\mathbf{b} \cdot \hat{\mathbf{s}}}{c} \tag{2.9}$$

Often the baseline length is measured in terms of the number of wavelengths (of the observed emission), such that $\mathbf{b} = \lambda \mathbf{b}_\lambda$, which can be substituted into equation equation 2.9 to yield the expression:

$$\tau_{\text{geo}} = \frac{\lambda \mathbf{b}_\lambda \cdot \hat{\mathbf{s}}}{c} = \frac{b_\lambda \cos(\theta_s)}{\nu_0} \tag{2.10}$$

where ν is the frequency of the emission[2]. Before passing on the signal received by the telescopes to the correlator, the geometrical time delay needs to be compensated for to ensure that the signals arrive at the same phase, which is called the instrument delay τ_{inst}. We can therefore define the difference between the geometrical and instrumental delays as $\tau = \tau_{\text{geo}} - \tau_{\text{inst}}$. This instrumental delay τ_{inst} is a constant value, such that when the telescope is pointed exactly in the direction of $\hat{\mathbf{s}}$ the expression $\tau_{\text{geo}} - \tau_{\text{inst}} = 0$.

The signals from the antennae are then passed to a receiver, where the voltages (as a function of time) of the signals received are proportional to:

$$\begin{aligned} V_x &\propto E e^{i\omega t} \\ V_y &\propto E e^{i\omega(t+\tau_{\text{geo}})} \end{aligned} \tag{2.11}$$

where V_x and V_y are the output voltages of the signal received at each telescope, $\omega = 2\pi\nu$ is the angular frequency, and E is the amplitude of the electromagnetic wave from the distance source (Wilson et al. 2013). These two signals[3] are then be combined by a correlator, which computes the time-averaged product of the two voltages.

$$\begin{aligned} R(\tau) &\propto \frac{E^2}{T} \int_0^T e^{i\omega t} e^{-i\omega(t-\tau)} \mathrm{d}t \\ R(\tau) &\propto \frac{E^2}{T} \int_0^T e^{i\omega\tau} \mathrm{d}t \end{aligned} \tag{2.12}$$

---

[2] The final expression comes from the fact that $\mathbf{b}_\lambda \cdot \hat{\mathbf{s}} = b_\lambda s \cos(\theta_s)$ and given that $\hat{\mathbf{s}}$ is a unit vector, $s = 1$.

[3] In fact, these signals are first mixed with a similar frequency generated by a local oscillator (LO), resulting in only the beat frequency (i.e. lower than the initial frequency) of each signal being passed to the correlator. Correlating these much lower frequency signals is a much simpler task than using the raw signal.

Where T is the integration time. This then becomes:

$$R(\tau) \propto E^2 e^{i\omega\tau} \tag{2.13}$$

In terms of measurable quantities, we can express $R(\tau)$, the output from the correlator, as

$$R(\tau)_{xy} = A(\hat{\mathbf{s}}) I_\nu(\hat{\mathbf{s}}) e^{i\omega\tau} \mathrm{d}\Omega \mathrm{d}\nu \tag{2.14}$$

where $I_\nu(\hat{\mathbf{s}})$ is the brightness distribution in the direction $\hat{\mathbf{s}}$, $A(\hat{\mathbf{s}})$ is the effective collecting area of each antenna (assumed to be identical). The term $A(\hat{\mathbf{s}}) I_\nu(\hat{\mathbf{s}}) \mathrm{d}\Omega \mathrm{d}\nu$ is the power received per bandwidth $\mathrm{d}\nu$ over solid angle element $\mathrm{d}\Omega$.

Using these definitions we can define the visibility function for the baseline $\mathbf{b}$, by integrating over the solid angle subtended by the source (Ω_S):

$$R(\mathbf{b}) = \iint\limits_{\Omega_\mathrm{S}} A(\hat{\mathbf{s}}) I_\nu(\hat{\mathbf{s}}) \exp\left[i2\pi\nu \left(\frac{\mathbf{b} \cdot \hat{\mathbf{s}}}{c} - \tau_\mathrm{inst} \right) \right] \mathrm{d}\Omega \mathrm{d}\nu \tag{2.15}$$

The brightness distribution of the observed source, $I_\nu(\hat{\mathbf{s}})$, is embedded within the above visibility function (what is measured by an intereformeter). So, it is necessary to try and extract this quantity from $R(\mathbf{b})$, which can be challenging in practice.

2.2.2 The u,v-plane

Solving equation equation 2.15 for I_ν can be made simpler by the introduction of a new coordinate system. First, we can define a unit vector that points to the centre of an imaged field as $\hat{\mathbf{s}}_0$, where the phase of the received signals is set to zero (this is also referred to as the phase centre). We can express the vector $\mathbf{s}$ in relation to this as

$$\mathbf{s} = \hat{\mathbf{s}}_0 + \boldsymbol{\sigma} \tag{2.16}$$

where $\boldsymbol{\sigma}$ is some (small) deviation from the phase centre, referred to as the sky offset vector. We can substitute this into equation equation 2.15, which after some manipulation becomes:

$$R(\mathbf{b}) = \exp\left[i2\pi\nu \left(\frac{\mathbf{b} \cdot \hat{\mathbf{s}}_0}{c} - \tau_\mathrm{inst} \right) \right] \mathrm{d}\nu \iint\limits_{\Omega_\mathrm{S}} A(\boldsymbol{\sigma}) I_\nu(\boldsymbol{\sigma}) \exp\left[i2\pi\nu \left(\frac{\mathbf{b} \cdot \boldsymbol{\sigma}}{c} \right) \right] \mathrm{d}\Omega \tag{2.17}$$

The term outside the integral describes the phase of $R(\mathbf{b})$ at the phase centre, whereas the integral describes the visibility:

$$V(\mathbf{b}) = \iint\limits_{\Omega_{\mathrm{S}}} A(\boldsymbol{\sigma}) I_\nu(\boldsymbol{\sigma}) \exp\left[i 2\pi\nu \left(\frac{\mathbf{b}\cdot\boldsymbol{\sigma}}{c}\right)\right] \mathrm{d}\Omega \tag{2.18}$$

or equivalently, in terms of the baseline vector measured in wavelengths ($\mathbf{b}_\lambda$):

$$V(\mathbf{b}_\lambda) = \iint\limits_{\Omega_{\mathrm{S}}} A(\boldsymbol{\sigma}) I_\nu(\boldsymbol{\sigma}) \exp\left(i 2\pi \mathbf{b}_\lambda \cdot \boldsymbol{\sigma}\right) \mathrm{d}\Omega \tag{2.19}$$

It is useful to define a convenient coordinate system for which we can relate the vectors $\boldsymbol{\sigma}$ and $\mathbf{b}_\lambda$, and hence measure $V(\mathbf{b}_\lambda)$. The three orthogonal components of the baseline vector $\mathbf{b}_\lambda$ are (u, v, w), where the unit vector $\hat{\mathbf{s}}_0$ defines the direction of the w axis as pointing towards the phase center of the observed field

$$\mathbf{b}_\lambda = \begin{pmatrix} u \\ v \\ w \end{pmatrix}, \text{ where } \hat{\mathbf{s}}_0 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \tag{2.20}$$

The axes u and v are projected in the celestial easterly (i.e. Right Ascension) and northerly (Declination) directions, respectively. This plane, which is perpendicular to the direction of the source (or more accurately, the phase center) is more commonly known as the u,v-plane (Burke et al. 2019). These axes are usually measured in wavelengths. An illustration of this coordinate system is show in Figure 2.4.

The sky offset vector $\boldsymbol{\sigma}$ is expressed in terms of the direction cosines (l, m, n), where

$$\begin{pmatrix} l \\ m \\ n \end{pmatrix} = \begin{pmatrix} \cos\alpha \\ \cos\beta \\ \cos\gamma \end{pmatrix} \tag{2.21}$$

where α and β are the angles between the vector $\mathbf{s}$ and the u and v axes, respectively. The plane defined by (l, m) is parallel to the u,v-plane, and n is parallel to both w and $\hat{\mathbf{s}}_0$. This is illustrated in Figure 2.5.

For a unit vector, $l^2 + m^2 + n^2 = 1$, and hence we can express n as

$$n = \sqrt{1 - l^2 - m^2} \tag{2.22}$$

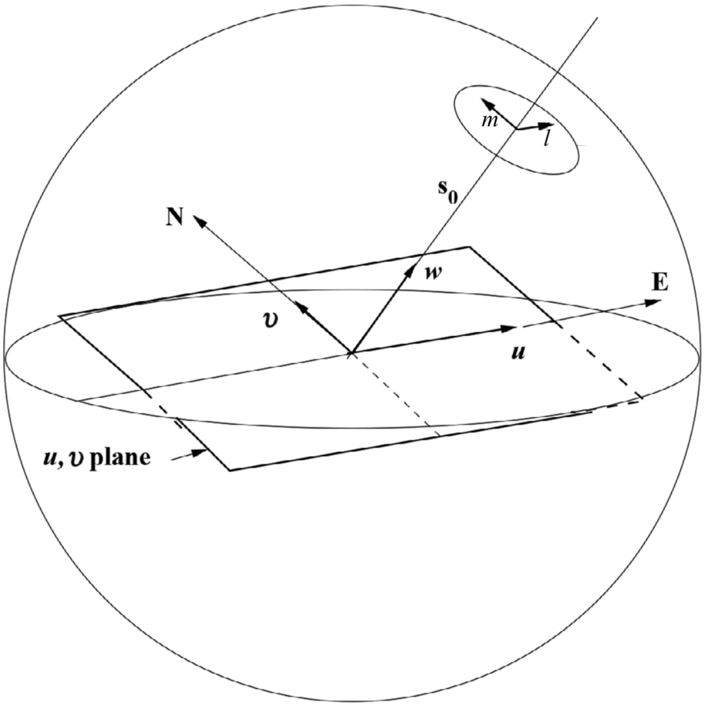

Figure 2.4: Illustration of the u,v-plane and it's relationship to the celestial sphere. Adapted from Burke et al. (2019).

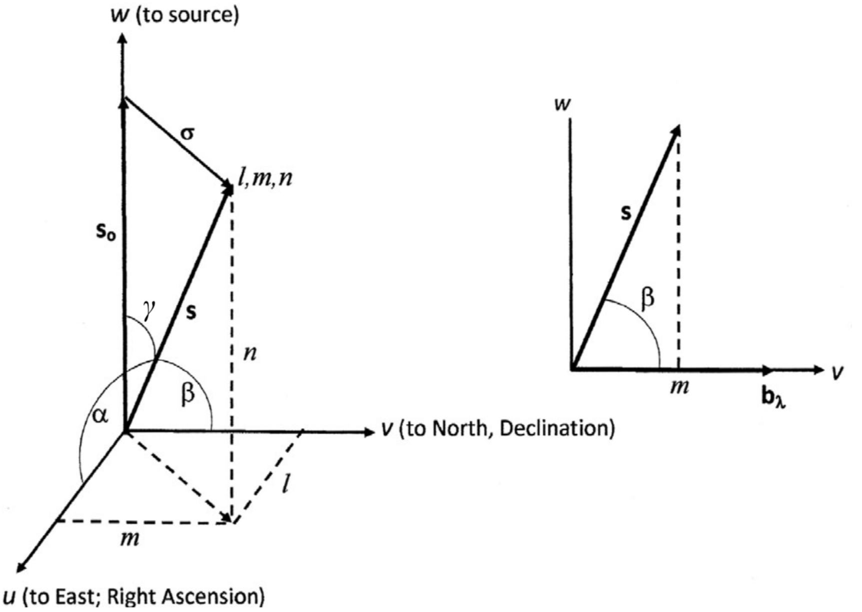

Figure 2.5: Illustration of the coordinates of $\boldsymbol{\sigma}$ in relation to the u,v-plane. The right hand set of axes shows a 2D projection along the u-axis. Adapted from Burke et al. (2019).

Note that (l, m, n) axes are measured in the spatial domain, whereas the (u, v, w) axes are measured in terms of spatial frequency. Since $\boldsymbol{\sigma} = \mathbf{s} - \hat{\mathbf{s}}_0$, and $\hat{\mathbf{s}}_0$ is a unit vector:

$$\boldsymbol{\sigma} = \begin{pmatrix} l \\ m \\ \sqrt{1 - l^2 - m^2} - 1 \end{pmatrix} \tag{2.23}$$

Using these newly defined coordinates we can substitute these into the exponential term of equation 2.19:

$$2\pi \mathbf{b}_\lambda \cdot \boldsymbol{\sigma} = 2\pi[ul + vm + w(\sqrt{1 - l^2 - m^2} - 1)] \tag{2.24}$$

and hence, we can now rewrite equation 2.19 in terms of these coordinates:

$$\begin{aligned} V(u, v, w) = &\iint_{\Omega_\mathrm{S}} A(l, m) I_\nu(l, m) \\ &\cdot \exp\left(i2\pi[ul + vm + w(\sqrt{1 - l^2 - m^2} - 1)]\right) \frac{\mathrm{d}l\mathrm{d}m}{\sqrt{1 - l^2 - m^2}} \end{aligned} \tag{2.25}$$

where

$$\mathrm{d}\Omega = \frac{\mathrm{d}l\mathrm{d}m}{\sqrt{1 - l^2 - m^2}} \tag{2.26}$$

When mapping only a small region of the sky, the values of both l and m are very small, and hence the term $\sqrt{1 - l^2 - m^2} \approx 1$. This simplifies the equation, leading to:

$$V(u, v, w) = e^{i2\pi w} \iint_{\Omega_\mathrm{S}} A(l, m) I_\nu(l, m) e^{i2\pi(ul + vm)} \mathrm{d}l\mathrm{d}m \tag{2.27}$$

Finally, if we restrict the baseline to occupy only the u,v-plane, $w = 0$, and we are left with the expression:

$$V(u, v) = \iint_{\Omega_\mathrm{S}} A(l, m) I_\nu(l, m) e^{i2\pi(ul + vm)} \mathrm{d}l\mathrm{d}m \tag{2.28}$$

Giving out final expression for the visibility function. We can then recover the brightness distribution from this measurement of $V(u, v)$ by performing an inverse Fourier transform, yielding:

$$A(l, m) I_\nu(l, m) = \iint_{\Omega_\mathrm{S}} V(u, v) e^{-i2\pi(ul + vm)} \mathrm{d}u\mathrm{d}v \tag{2.29}$$

This relationship between the complex visibility and the brightness distribution is stated by the van Cittert-Zernike theorem (van Cittert 1934; Zernike 1938).

2.2.3 Synthesis imaging

For each baseline and integration time, an intereformeter samples one Fourier component (visibility) in the u,v-plane. While it is possible to increase the number of samples by performing multiple integrations at different times by relying on the rotation of the Earth, hence changing the orientation of the baseline relative to the source, to sample different parts of the u,v-plane. This only increases the number of samples by the the number of integration times, and this alone is not a very efficient way of thoroughly sampling the u,v-plane.

However, increasing the number of antennas in the array not only adds more collecting area, but also adds more baselines to the intereformeter, and hence more visibilities. The number of baselines, and hence visibilities sampled per integration time in the u,v-plane, is given by

$$N_{\text{baseline}} = \frac{N(N-1)}{2} \tag{2.30}$$

where N is the number of antennas in the array. This dramatically increases the number of samples over the course of an observing period, and by carefully distributing the positions of the antennae in the array it is possible to more thoroughly and efficiently sample the u,v-plane.

For small fields of view, the van Cittert-Zernike theorem states that the measured interference pattern (or complex visibility) of a distant, incoherent source is the 2D Fourier transform of the source's brightness distribution on the sky. As we increase the number of antennae and integrations, we construct a more "complete" set of visibilities, which in turn will yield a "truer" measurement of the source's brightness distribution and therefore a better image of the source.

Given the practical limitations of sampling an infinite number of visibilities, the resulting brightness distribution will contain artefacts as a result of the incomplete u,v coverage.

With a set of visibilities obtained, one must calibrate the measurements prior to imaging. In short, this is comprised of a few components. Flux calibration involves observing an astronomical source with a well characterised brightness model (often a planet or moon in the solar system), and using this measurement to set the correct flux scale for the scientific target. As the phase of the incoming radiation is critical to interferometry, phase calibration needs to be performed. This usually involves observing (at regular intervals during an observing run) a bright point source near the target field to track and correct for atmospheric variations. And finally, bandpass and gain calibration are required to monitor the variations of phase and amplitude with frequency and time, respectively. These may drift due to the source elevation varying, or changes in weather, over the observing run. Very bright point sources, typically quasars near the target, are chosen to track these variations.

All of these calibrators will be monitored and checked regularly by the observatory to ensure accurate calibration.

Once the data are calibrated, the Fourier transform of the set of visibilities yields an image known as a dirty map, I_{D}, which can be expressed as a convolution[4]:

$$I_{\mathrm{D}}(l,m) = P_{\mathrm{D}}(l,m) * I(l,m) \tag{2.31}$$

where $I(l,m)$ is the true source brightness distribution, and $P_{\mathrm{D}}(l,m)$ is the dirty beam, which is effectively is the point spread function (PSF) of the intereformeter. This is in fact a poor-quality PSF that contains artefacts (such as sidelobes) that will add spurious emission to an image if not accounted for.

This PSF can be well characterised, as it is a function of the array element positions, and the location of the source of the sky, both of which can be measured with a high degree of accuracy. It is clear then, that in order to recover the source brightness distribution I[5], the dirty map I_{D} needs to be deconvolved with the dirty beam P_{D}.

---

[4] The convolution theorem states that $\mathcal{F}\{f * g\} = \mathcal{F}\{f\} \cdot \mathcal{F}\{g\}$, where $\mathcal{F}$ denotes the Fourier transform operator.

[5] In practice I is not possible to recover exactly as this would require perfect and complete sampling of the u,v-plane, so strictly speaking some similar brightness distribution I' is being measured. However, as the u,v sampling becomes more complete, this approaches the true distribution, and when considering the inclusion of instrumental noise, the final image is close enough to the true source brightness distribution.

2.2.4 The CLEAN algorithm

Högbom (1974) developed the CLEAN algorithm, a brute force approach to performing deconvolution. CLEAN first locates the point of brightest emission in I_{D} and then subtracts the PSF (P_{D}) from that location in I_{D}. The algorithm takes note of the position and emission subtracted, and then proceeds to the next brightest point in I_{D}, and continues to iterate through each subsequent brightest point in the dirty map until PSF-subtracted, or residual map, resembles noise.

The subtracted fluxes are then added into an empty map, which is then convolved with a restoring beam (clean beam), which is typically some idealised elliptical Gaussian (based on the size of the main beam of the dirty beam), yielding a model brightness map. This can then be added to the field of noise, producing a final "cleaned map", which is a close approximation of the source brightness distribution I_{D}.

The CLEAN algorithm has been implemented in the Common Astronomy Software Applications (CASA), which is the main data processing software suite for the Atacama Large Millimeter/submillimeter Array (ALMA) and Karl G. Jansky Very Large Array (VLA), and is often used for reducing data from other radio telescopes (McMullin et al. 2007; The Casa Team et al. 2022). This implementation is based on the Cotton-Schwab CLEAN algorithm (Schwab 1984), which is comprised of two layers of iterative outer loops (major cycles), and inner loops (minor cycles). The major cycle is responsible for transforming the data between the Fourier and image domains, and the minor cycles are purely based in the image domain. These cycles run an iterative, weighted χ^2 minimisation process on a set of linear equations of the same form as equation equation 2.31. A schematic of this process is shown in Figure 2.6, taken from CASA's documentation[6].

Iterations start with some initial guess of the model brightness distribution. The major cycle creates a model visibility distribution and subtracts it from the data visibilities, creating residual visibilities. These residual visibilities are then gridded, and subsequently inverse fast Fourier transformed (iFFT) into a residual image. This residual image is then passed onto the minor cycle, where deconvolution

---

[6] `https://casadocs.readthedocs.io/`

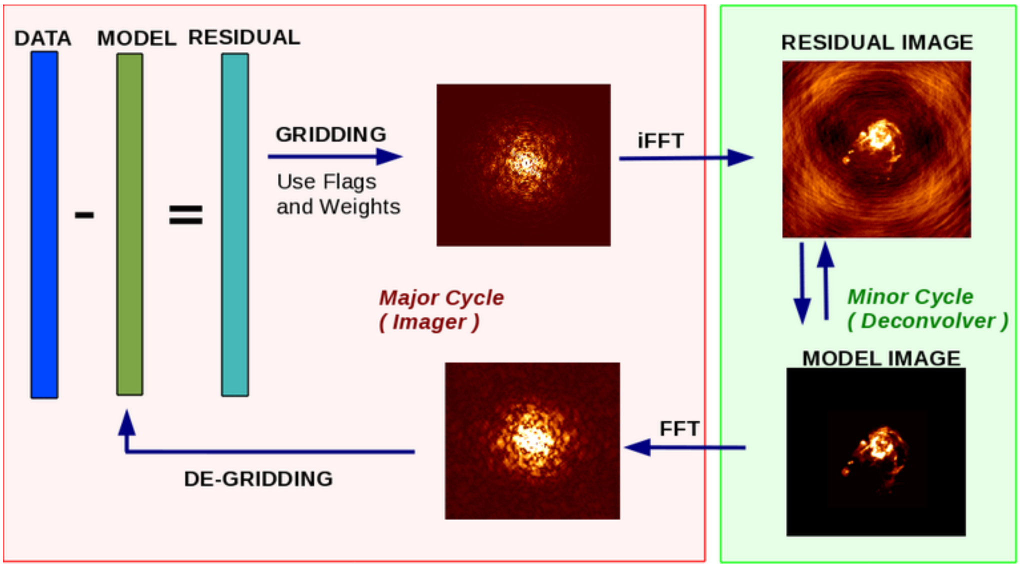

Figure 2.6: Schematic of CASA's implementation of the Cotton-Schwab CLEAN algorithm.

with the dirty beam P_{D} is performed iteratively, effectively building up a model of the brightness distribution by separating it from the PSF, until some minor cycle stopping criterion is reached and a model image is created. This model image is then passed back to the major cycle and fast Fourier transformed (FFT) into model visibilities, which then begins the next iteration of the major cycle. This process continues until some major cycle or global stopping criterion (e.g. number of iterations, or if the RMS of the residual reaches some desired noise level), and a final model image is created. This final model image is then added to the residuals and convolved with the restoring beam, yielding a final reconstructed image of the source.

The data are not uniformly distributed in the u,v-plane and are usually centrally concentrated, due to the increased number of short baselines in the telescope array. Within the imaging process, the visibilities can be weighted to alter the telescope response, and hence optimise various properties, such as dynamic range, synthesised beam size, and sensitivity. Although there are numerous weighting schemes implemented in CASA, here we will focus on three of the more commonly used schemes. The first is "natural weighting", which does not modify the intrinsic weighting of each visibility, and so the weighting parameter is given by:

$$w_i = \frac{1}{\sigma_i^2} \tag{2.32}$$

where σ_i^2 is the noise variance of the i-th visibility. Because of the central concentration of sampling, larger scale features will be optimised, degrading resolution slightly. This leads to higher point-source sensitivity, at the expense of the inclusion of some undesirable structure within the PSF.

On the other extreme, "uniform weighting" tries to compensate for the uneven sampling density within the u,v-plane, giving equal weight to each measured spatial frequency regardless of sample density. The weight is given by

$$w_i = \frac{1}{N_{\mathrm{S}}(k)} \frac{1}{\sigma_i^2} \tag{2.33}$$

where $N_{\mathrm{S}}(k)$ is the number of visibilities within some symmetric region (cell) k of the u,v-plane. The PSF produced by uniform weighting has the smallest possible angular resolution (∼70% the size of PSF from natural weighting). However, the sensitivity is notably reduced (∼55% relative to natural weighting), and artefacts can crop into the PSF due to sparsely sampled parts of the u,v-plane being given artificially high weights.

Finally, "Briggs weighting" (also known as "robust weighting"), is a weighting scheme that can be altered depending on the desired beam properties to be optimised (Briggs 1995). It is based around the SNR of the measurements and a tuneable "robust" parameter R that defines some noise threshold. High SNR samples are weighted by the sample density, like in uniform weighting, to optimise for a smaller PSF, and low SNR samples are naturally weighted to optimise for sensitivity. This robust parameter is effectively a sliding scale between uniform and natural weighting, where a value of $R = -2$ is closest to uniform weighting, and $R = 2$ is closest to natural weighting. A value of $R = 0.5$ is commonly used, as the sensitivity is almost as good as natural weighting (∼95%), with a significantly smaller beam (∼80%), although the exact amounts will vary depending on the density and distribution of u,v samples (Briggs 1995; Briggs et al. 1999).

More elaborate versions of the CLEAN algorithm have been developed to overcome some of the shortcomings of the original Högbom (1974) algorithm, which assumes that the source's brightness distribution is solely comprised of point sources, with little to no other emission in the field-of-view. One particular extension of the CLEAN algorithm, suitable for sources with extended emission, is the Multi-scale CLEAN (MS-CLEAN) (Cornwell 2008). MS-CLEAN constructs a model of the brightness distribution using both delta functions for point sources (as in the original CLEAN algorithm), but also Gaussian brightness components at a set of user-defined scales. This better accounts for emission on a broad range of scales, which in turn helps produce a better fitting set of model visibilities in the major cycles. The set of scales are usually set to representative sizes of dominant emission within the field, typically beam-sized, and around 3–5 times the beam size (in addition to the point source scale), although care should be taken not to choose size scales that correspond to unmeasured scales caused by lack of short-spacings in the interferometer.

For more careful image reconstruction, it is often beneficial to restrict which regions of a field the CLEAN algorithm operates within, for example to avoid adding brightness components to regions where there is no signal. This can be specified by a binary mask. For simple, point-like sources, a mask could be as simple as a single elliptical mask covering the region of emission. For more complex or extended sources, a more sophisticated mask may be required for optimal model construction. For example, this could be defined as a set of user-defined polygons that are pre-defined and used for the entire imaging procedure. Another option is to alter the mask after every major cycle, which is often hand-drawn by the user to cover regions they deem to contain "real" emission.

This presents a couple of issues. Determining what emission in a dirty image is "real" is rather subjective, and so two different users could end up with a different final image of a source depending on their judgment. Another issue is that for large fields, with highly complex emission, it can be extremely time-consuming to draw appropriate masks that encompass the observed emission. For cube imaging, a mask will need to be defined for every channel for which there is emission. In addition to this, if many major cycles are required to CLEAN the data to a required RMS noise level or stopping criterion, the mask may need to change over time, requiring further user intervention that will likely get less accurate, due to user fatigue, with subsequent cycles.

To overcome these issues, an automated masking algorithm called “auto-multithresh” has been developed by Kepley et al. (2020), and implemented in CASA primarily for imaging ALMA data, but can also be used for imaging data from other radio interferometers. The algorithm uses multiple noise and sidelobe level thresholds in the residuals to find regions of emission to mask. The mask is updated at the start of every minor cycle, based on the most recently generated residual image. Small areas of the mask can also be “pruned” automatically, as regions that are small enough are unlikely to be real emission. Low SNR emission regions adjacent to the created mask are then incorporated (by binary dilation) down to some low SNR threshold. Finally, the mask is convolved with a Gaussian with a size that is a multiple of the synthesised beam, and then only regions with emission that are above some fraction of the peak brightness are retained in the final mask for that iteration of deconvolution.

All of the thresholds and parameters for each of these operations can be altered by the user, however CASA has some recommended values for each parameter that depends on various array configuration and whether continuum or line emission is being imaged.

For imaging line emission, the data are gridded into a spectral cube, which can be represented in position-position-frequency space, or with a set reference frequency and doppler corrections, transformed into some preferred velocity reference frame (e.g. the Kinematic Local Standard of Rest, or LSRK) producing a position-position-velocity (PPV) cube.

For imaging continuum emission, it is common to measure a source’s brightness over a very broad range of frequencies, and combine all of the frequency channels during imaging. This is called multi-frequency synthesis (MFS). MFS was initially developed by Conway et al. (1990) as means to increase the quality of radio images from sparsely sampled sets of visibilities, such as for interferometers with a small number of baselines. By effectively increasing u,v sampling, it is possible to greatly increase image fidelity and point source sensitivity. One requirement is that a source's brightness should not vary greatly over the frequency range of the band. When performing MFS, the resulting image will be 2D, in position-position (PP).

2.3 The Atacama Large Millimeter/submillimeter Array (ALMA)

This thesis makes use of both 2.9 mm continuum and $N_2H^+(J{=}1\text{–}0)$ line emission data from the Atacama Large Millimeter/submillimeter Array (ALMA). ALMA, one of the largest and most expensive ground-based telescope projects, is a set of two interferometers comprised of an extended fifty 12 m antenna array, and the Atacama Compact Array (ACA), a compact twelve 7 m antenna array, supplemented by four 12 m single-dish telescopes (TP) operating in the millimeter/submillimeter wavelength regime, from wavelengths of around 0.32 mm to 8.5 mm.

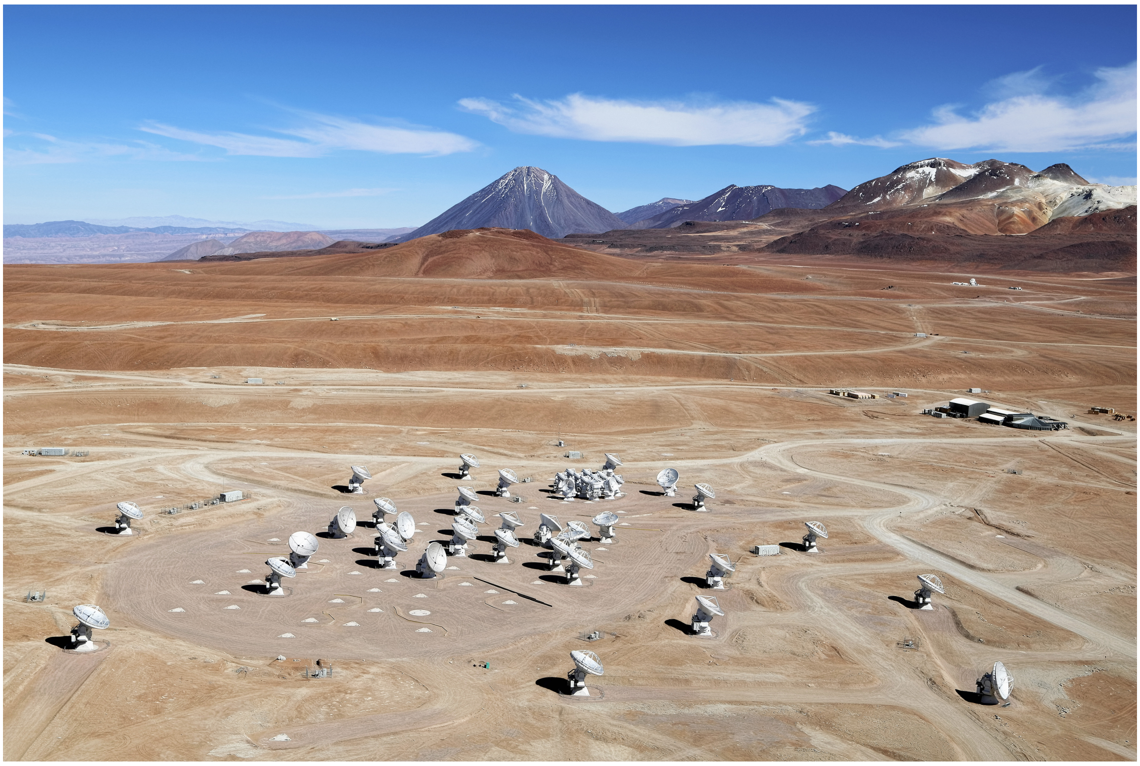

Figure 2.7: Aerial view of the Chajnantor plateau showing the larger 12 m antennae, along with the ACA towards the centre of the image. Credit: Clem & Adri Bacri-Normier/ESO.

Located in the Chajnantor plateau high up in the Chilean Andes, at an elevation of 4576 m to 5044 m, the observatory is an international partnership between Chile, Europe, the United States, Canada, Japan, South Korea, and Taiwan. Light in the millimeter and submillimeter regime can help us explore some of the coldest molecular cloud structures in space (the birthplace of stars), and observe emission from the earliest, most distant galaxies in the Universe. These faint signals are however strongly absorbed by water vapour in the Earth's atmosphere, and so the Chajnantor plateau is the perfect location for such an observatory, being such a dry location, coupled with the high elevation meaning that light has to pass through significantly less of the atmosphere before reaching the telescopes.

The antenna in the 12 m array can be moved into multiple different locations, with baselines ranging from 14.6 m, to anywhere between 160.7 m to 16.2 km for the most compact and extended configurations, respectively. The ACA has baselines from 8.7 m to 45.0 m. With the large number of antennae, and hence baselines (1225 baselines for the 12 m array alone), ALMA can very efficiently sample the u,v-plane over a vast range of spatial scales (and when combined with the TP telescopes, all spatial scales), and observe at resolutions ranging from $31.5''$ (7 m array, at $\nu = 40\,\mathrm{GHz}$) to $0.0048''$ (12 m array, most extended configuration, at $\nu = 870\,\mathrm{GHz}$).

Figure 2.8: A view of several ALMA antennae, with a the Galactic Centre of the Milky Way visible in the background. Credit: B. Tafreshi/ESO.

Chapter 3

Observations of a sample of six hub-filament systems

3.1 Sample selection

For the purpose of this study, 6 infrared dark clouds we selected, all part of the Peretto & Fuller (2009) catalogue. One of these, SDC335, was already examined by my collaborators in a series of studies (Peretto et al. 2013; Avison et al. 2015, 2021). The Peretto & Fuller (2009) catalogue contains over 11,000 IRDCs with a column density above $10^{22}\,\mathrm{cm}^{-2}$ identified within the *Spitzer* GLIMPSE survey (Churchwell et al. 2009) of the Galactic Plane between $10^\circ < |l| < 65^\circ$ and $|b| < 1^\circ$. The IRDCs in this catalogue cover a mass range from around $10\,\mathrm{M}_\odot$ to a few $10^4\,\mathrm{M}_\odot$, and a distance range of roughly 2 kpc to 8 kpc (Peretto & Fuller 2010)

These 6 clouds were selected as they exhibit a well defined hub-filament morphology seen in extinction at 8 µm, with an easily identified filament convergence point[1]. They all have high extinction contrast against a relatively uniform mid-infrared background. They have been selected so that their distances lie within a narrow range, i.e. from 2 kpc to 3.2 kpc, so that their properties can easily be compared to each other. Finally, they have been chosen so that they cover a large range

---

[1] It is not yet known what fraction of IRDCs/clumps in the Milky Way exhibit a hub-filament morphology, but preliminary results from within the group suggest that around 5% to 10% of clumps are HFSs.

of masses, from a few hundred to a few thousand solar masses, to try to evaluate the impact of the hub morphology on core formation independently of the clump mass.

3.2 ALMA observations

Five HFSs (see Table 3.1) were observed with the Atacama Large Millimeter/submillimeter Array (ALMA) 12 m array between 20th–23rd January 2016 with a total of 41–46 antennas (C36-1 configuration), and with the Atacama Compact Array (ACA) between 17th April and 25th July 2016 (during Cycle 3) with 11 antennas (Project ID: 2015.1.01014.S; PI: Peretto). The number of 12 m (7 m) pointings was 61 (23), with a total on-source observing duration of 3.81 h (10.76 h) (for all five fields).

An additional HFS, SDC335, was observed with the ALMA 12 m array between 27th September–19th November 2011 with the 16 available antennas during Cycle 0 in the compact configuration (Project ID: 2011.0.00474.S; PI: Peretto). A complete description of the observations are presented in Peretto et al. (2013). Follow-up observations of SDC335 were performed with the ACA between 6th–8th November 2016 (during Cycle 4) with 10 antennas (Project ID: 2016.1.00810.S; PI: Peretto). The total number of 12 m (7 m) pointings was 11 (6), with a total on-source observing duration of 4.11 h (1.33 h).

The data were reduced and calibrated using the same CASA[2] (McMullin et al. 2007) versions as used by the ALMA pipeline, using the standard pipeline scripts. The quasars J1531-5108, J1604-4441, J1617-5848, J1706-4600, J1650-5044 were used for phase calibration. Flux and bandpass calibration were performed using Mars, Ganymede, Neptune, J1427-4206, J1617-5848, J1733-1304, J1924-2914. The uncertainty in absolute flux calibration is $\sim 5\%$ in Band 3, according to the ALMA Cycle 3 Technical Handbook[3]. The weights of the 12 m SDC335 data were adjusted using `statwt()` on line-free channels prior to combination with the Cycle 4 ACA data.

The calibrated ACA and 12 m visibilities were then concatenated and imaged using CASA version 5.5.0, utilising its implementation of the Multi-Scale CLEAN

---

[2] `https://casa.nrao.edu`

[3] `https://almascience.eso.org/documents-and-tools/cycle3/alma-technical-handbook`

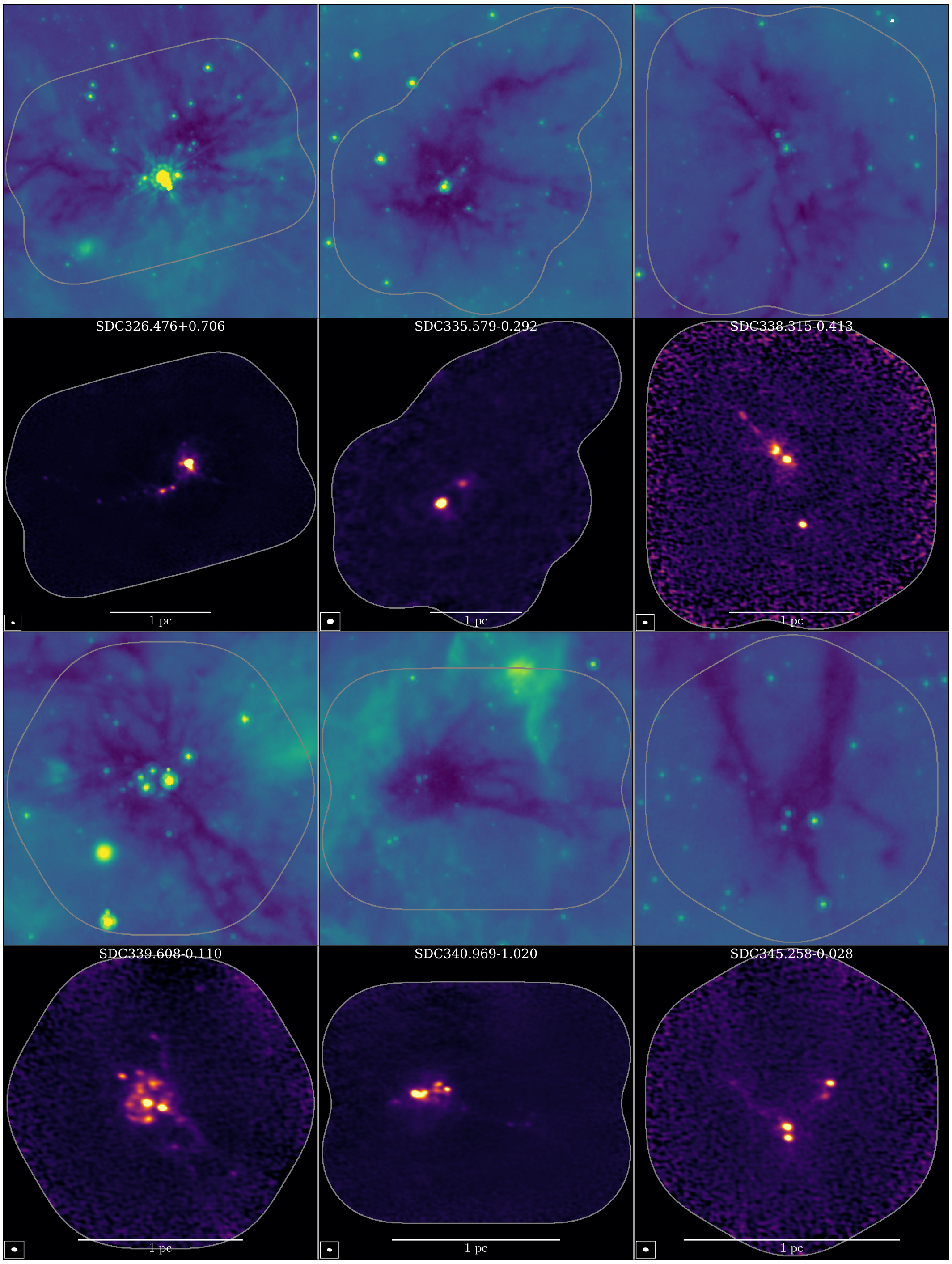

Figure 3.1: *(1st and 3rd row) Spitzer* 8 µm images of the six HFSs observed with ALMA. Below each *Spitzer* image is the corresponding ALMA combined 7 m+12 m continuum images at 2.9 mm of each HFS. The synthesised beam size of each image is shown in the lower left hand corner, the grey contour shows the extent of our ALMA fields.

(Cornwell 2008) within the `tclean()` task. This was to better recover extended emission features that are larger than the beam. The data were imaged with Briggs weighting (Briggs 1995) with a robust parameter of 0.5. All of the images were primary beam corrected. Automatic masking was used during imaging.

Continuum data

Table 3.1: The 6 HFSs we observed with ALMA, their estimated distances, and a summary of observational properties of the combined 7 m+12 m 0 2.9 mm continuum observations. The HFSs will hereafter be referred to by their shorthand names highlighted in bold. The linear resolution x_d corresponds to the physical size of the synthesised beam major axis at the distance d. The mass sensitivities were calculated assuming a source temperature of 12 K.

SDC Name (Peretto & Fuller 2009)	d (pc)	Beam ($'' \times ''$)	PA (°)	RMS noise (µJy/beam)	x_d (pc)	M_{min} ($\mathrm{M}_\odot$)
SDC326.476+0.706	2610	2.80 × 2.16	69.12	78.68	0.035	0.27
SDC335.579–0.292	3230	4.69 × 3.63	−79.00	379.72	0.073	1.98
SDC338.315–0.413	2940	2.91 × 1.92	80.19	49.00	0.041	0.21
SDC339.608–0.113	2740	2.88 × 1.90	80.42	66.06	0.038	0.24
SDC340.969–1.020	2210	2.88 × 1.92	79.92	100.10	0.031	0.24
SDC345.258–0.028	2090	2.84 × 1.87	80.65	50.88	0.029	0.12

For the five hubs observed in Cycle 3 the central frequency was 103.6 GHz, and the total continuum bandwidth used was 3.735 GHz. For SDC335 only 0.2 GHz of continuum bandwidth was used for imaging (split over two bands 104.0 GHz to 104.1 GHz, 105.0 GHz to 105.1 GHz), as this was the frequency coverage of the original Cycle 0 dataset. The central frequency was 104.55 GHz. Table 3.1 contains a summary of the observational parameters for the six fields.

An angular resolution of $\sim 2.8''$ to $4.7''$ is achieved, which at the distance of the targets corresponds to a linear resolution range of 0.029 pc to 0.073 pc. This is at least a factor of two smaller than the Jeans length (which ranges between 0.10 pc to 0.21 pc) computed from the clump's average density, assuming a sound speed of

$0.2\,\mathrm{km\,s^{-1}}$.

N_2H^+(J=1–0) data

Table 3.2: Summary of observational properties of of our combined TP+7 m+12 m observations of N_2H^+(J=1–0). The RMS noise (in line-free channels) varies between spectra in each field, so here it represents the mean RMS for the whole cube.

Name	d (pc)	Beam ($'' \times ''$)	PA ($^\circ$)	RMS (K)	x_d (pc)	V_{sys} ($\mathrm{km\,s^{-1}}$)
SDC326	2610	3.26×2.49	67.42	0.15	0.041	-39.4
SDC335	3230	5.21×4.16	-74.50	0.08	0.082	-46.5
SDC338	2940	3.38×2.25	80.47	0.16	0.048	-38.0
SDC339	2740	3.36×2.21	81.95	0.18	0.045	-32.9
SDC340	2210	3.36×2.23	80.73	0.17	0.040	-22.7
SDC345	2090	3.29×2.17	81.43	0.18	0.033	-16.4

In addition to the continuum band setup mentioned above, spectral line bands were observed in parallel. The bands were centred around the frequency of the brightest hyperfine transition of N_2H^+(J=1–0) at 93 173.7643 MHz, with a spectral resolution of $0.2\,\mathrm{km\,s^{-1}}$. Whereas the continuum bands were chosen to measure the thermal dust emission from cores within the HFSs, the spectral line setup were chosen to trace the kinematics of the dense gas in the HFSs.

In addition to the 12 m and 7 m array observations, single-dish observations were taken of all fields (in parallel with the ACA) using the Total Power (TP). These were taken in order to complement the combined 7 m+12 m spectral line setups, by providing the zero-spacing information, and hence recover the large scale emission in the fields. The total amount of TP observations for all six fields was ~ 36.3 h.

Once the combined 7 m+12 m N_2H^+(J=1–0) datacubes were cleaned, they needed to be combined with the TP data. This was performed using a technique called “feathering”, which is implemented in CASA task `feather()`. Feathering involves Fourier transforming both datasets, and then applying a high-pass filter to the high resolution (i.e. 7 m+12 m) data, and a low-pass filter to the low resolution (TP) data. The filter weights are derived from the beam sizes of the input data. The two filtered cubes are then summed, and then inverse Fourier transformed. The angular resolution of the final N_2H^+(J=1–0) datacubes are $\sim 3.3''$ to $5.2''$, with a mean RMS noise (in line-free channels) between 0.08 K to 0.18 K. A summary of the observational parameters are shown in Table 3.2.

3.3 Ancillary *Spitzer*, *WISE* and *Herschel* data

This thesis makes use of publicly available *Spitzer* GLIMPSE 8 µm data[4] (Churchwell et al. 2009) and *WISE* 12 µm data[5] (Wright et al. 2010), at an angular resolution of $\sim 2.4''$ and $\sim 6.5''$, respectively. Temperature and column density maps are used, first presented in Peretto et al. (2016), at a resolution of $\sim 18''$, which were constructed from 160 µm and 250 µm data from the *Herschel* Hi-GAL survey (Molinari et al. 2010). Finally, the Molinari et al. (2016) 70 µm compact source catalogue was also used in this work. Figure 3.1 shows the *Spitzer* 8 µm fields for all six HFSs, along with the final imaged ALMA 2.9 mm continuum fields.

---

[4] `https://irsa.ipac.caltech.edu/data/SPITZER/GLIMPSE`

[5] `https://irsa.ipac.caltech.edu/applications/wise/`

Chapter 4

An ALMA study of hub-filament systems: I. On the clump mass concentration within the most massive cores

The following material was published in Anderson et al. (2021).

In this paper we aimed to constraining the efficiency of hubs at concentrating their mass into their most massive core, and this for a large range of clump masses. The end goal is to disentangle the effects of clump mass to those related to clump morphology. We do this by analysing new ALMA 2.9 mm observations of a sample of six HFSs.

4.1 ALMA 2.9 mm continuum images

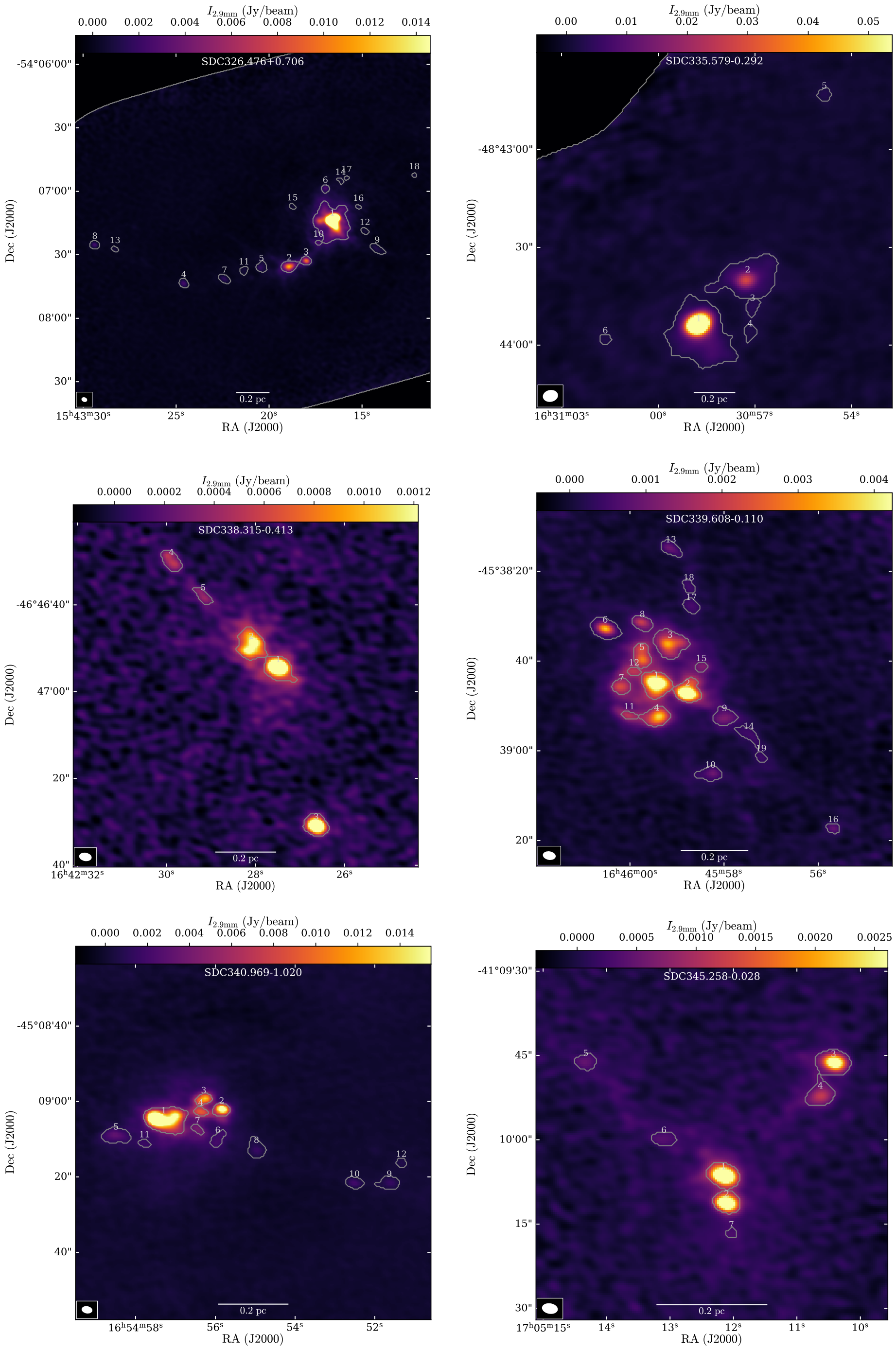

Figure 4.1: Close up view of ALMA 2.9 mm continuum images of our sample of 6 hub-filament systems, showing all of our extracted “cores”. Each extracted core is labeled with their MM#, with the grey contours showing each core’s dendrogram structure footprint.

4.2 Mass fragmentation

4.2.1 Core extraction

To extract the structures from our ALMA continuum images we use a dendrogram-based method using `astrodendro`, a Python package based on the Rosolowsky et al. (2008) implementation of dendrograms to analyse astronomical data. Our extraction method required that the minimum structure size $n_{\mathrm{pix,min}}$ must be greater than half the synthesised beam size (typically 18 pixels), the faintest included emission structure to be at a specific intensity of $I_{\mathrm{min}} = 5 \times \sigma_{\mathrm{global}}$, and minimum structure significance $\Delta I_{\mathrm{min}} = 1 \times \sigma_{\mathrm{global}}$, where σ_{global} is the RMS specific intensity calculated from the residual image of the field produced after imaging (see Table 3.1). This ensured that all of the extracted emission structures were at least detected five times above the global RMS in an image, with a peak at least 6 times the RMS.

For our analysis, we consider the leaves of the dendrogram (i.e. emission structures that do not have any detectable substructure) to be "cores". We are aware that these sources may well be sub-fragmented at higher resolution. We impose the constraint that only cores that are contained within the contour at >50% of the primary beam power are included in the dendrogram. This is to avoid noise peaks that appear brighter and occur more frequently toward the edges of the fields, caused by the non-uniform response of the primary beam. After applying this constraint, the constructed dendrograms contained 71 candidate cores across the six fields.

We produced error maps by performing a windowed RMS calculation on our residuals, with a window size of 4x4 beam major axis lengths. We then construct signal-to-noise (SNR) maps to better estimate the strength of the detections, given that the noise varies across the fields, and to help remove any spurious detections. Core candidates with at least $n_{\mathrm{pix,min}}$ pixels with a SNR ≥ 3 are classed as detections. Extracted structures that do not satisfy this condition are discarded. After applying this criteria we obtain a set of 67 cores. Figure 4.1 shows a zoomed in view of these 67 cores, along with their extent as defined by the dendrogram.

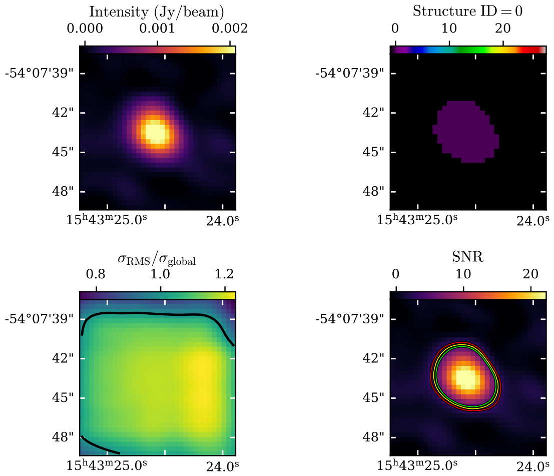

Figure 4.2: Example of core checking plot for Structure ID 0 in SDC326 (G326.4745+0.7027-MM4), which was flagged as a detection. *(top left)* Continuum image. *(top right)* Dendrogram leaf of structure extracted from image. *(bottom left)* RMS noise map divided by global RMS value used for dendrogram construction. The black contour represents where $\sigma_{\mathrm{RMS}}/\sigma_{\mathrm{global}} = 1$. *(bottom right)* SNR map, with contours for SNR levels of 3, 4 and 5 shown in red, yellow and green, respectively.

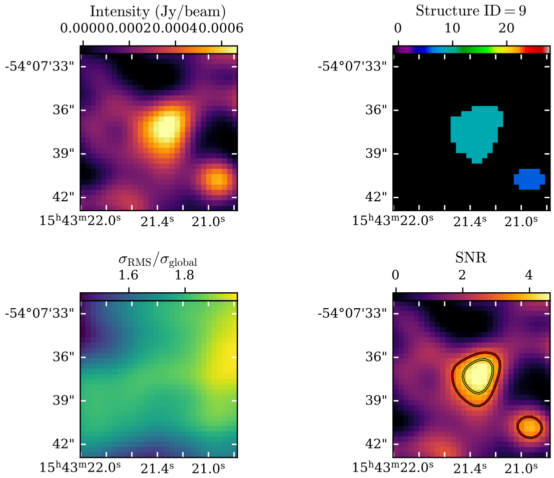

Figure 4.3: Another example of a core checking plot (as above) for Structure ID 9 in SDC326 (G326.4745+0.7027-MM11), which was also flagged as a detection. Structure ID 6 is shown to the lower right, which was discarded due to its small size and low signal-to-noise.

4.2.2 Core sizes and masses

Assuming that the cores are in local thermodynamic equilibrium (LTE) and that the dust emission is optically thin, the core masses can then be calculated using

$$M_{\mathrm{core}} = \frac{d^2 S_\nu}{\kappa_\nu B_\nu(T)} \tag{4.1}$$

where d is the distance to the source, S_ν is the integrated flux density of the source, κ_ν is the specific dust opacity, and $B_\nu(T)$ is the Planck function at a given dust temperature T (Kauffmann et al. 2008). We assume the same specific dust opacity relation as Marsh et al. (2015, 2017):

$$\kappa_\nu(\lambda) = 0.1\,\mathrm{cm}^2/\mathrm{g}\left(\frac{\lambda}{300\,\mu\mathrm{m}}\right)^{-\beta} \tag{4.2}$$

with a dust opacity index $\beta = 2$, for a given wavelength λ, and accounting for a gas-to-dust mass ratio of 100. The uncertainty in the dust opacity is around ±50% (Ossenkopf & Henning 1994; Roy et al. 2013, 2015).

To estimate core temperatures we use a combination of two methods. Our primary method is to use dust temperature maps derived from *Herschel* 160 µm/250 µm ratio maps as presented in Peretto et al. (2016). We simply take the temperature (T_{col}) at the position of each core's intensity-weighted centroid. These maps cover a temperature range of around 12–30 K for our set of fields. Note that because we assume a unique temperature along the line of sight and that the typical background temperature of the Galactic Plane is ~ 18 K, we may overestimate the temperatures of dense clumps colder than this background value (Peretto et al. 2010; Battersby et al. 2011; Marsh et al. 2015).

For warmer sources (such as massive protostellar cores), this may be significantly underestimating their temperature, and hence overestimating their mass. To try and counter this effect, we use the Hi-GAL 70 µm Compact Source Catalogue (Molinari et al. 2016) to see which cores in our sample have an associated 70 µm source, as the 70 µm flux density is known to be a good tracer of the luminosity of embedded sources (Dunham et al. 2008; Ragan et al. 2012). If a 70 µm source is present within the equivalent radius R_{eq} of a core, which is the radius of a circle with equal area to the core's corresponding dendrogram mask, we say they are associated. We then convert the 70 µm flux densities to bolometric (internal) luminosities using the following relation (Elia et al. 2017):

$$L_{\rm int} = 25.6 \left(\frac{S_{70\,\mu\rm m}}{10\,\rm Jy}\right) \left(\frac{d}{1\,\rm kpc}\right)^2 \rm L_\odot \tag{4.3}$$

Where $S_{70\,\mu\rm m}$ is the integrated 70 µm flux density of the source, and d is the distance to the clump. Assuming that the dust emission from a protostellar core is optically thin and is predominantly in the far-infrared, we calculate the mean mass-weighted temperature of the core, $\overline{T_{\rm d}}$ (Emerson 1988; Terebey et al. 1993):

$$\overline{T_{\rm d}} = \frac{3}{2} T_0 \left(\frac{L_{\rm int}}{L_0}\right)^{1/6} \left(\frac{r}{r_0}\right)^{-1/3} \tag{4.4}$$

Where $L_{\rm int}$ is the source's internal luminosity, r is the core's radius, and reference values $T_0 = 25\,\rm K$, $L_0 = 520\,\rm L_\odot$, $r_0 = 0.032\,\rm pc$. This form of the equation assumes $\beta = 2$, and that the density profile of the core follows $\rho(r) \propto r^{-2}$, (as used by e.g. Bontemps et al. 2010; Svoboda et al. 2019).

We use equation 4.4 to calculate the mean temperature within core equivalent radius $r = R_{\rm eq}$. For our set of sources, $\overline{T_{\rm d}}$ ranges between 18–76 K.

If a 70 µm flux density derived temperature can be obtained for a core, we assign the core $T_{\rm core} = \overline{T_{\rm d}}$, and otherwise assign $T_{\rm core} = T_{\rm col}$. We assume that the temperature of the gas and dust are coupled as the cores have a density at least $\sim 10^6\,\rm cm^{-3}$, the threshold at which Goldsmith (2001) states that the dust and gas temperatures become essentially equal.

We use the Revised Kinematic Distance Calculator[1] (Reid et al. 2009, 2014) to estimate the distances to the IRDCs, using the LSR velocities for each clump. We assume that the IRDCs are located at the near distance as they are IR-dark at 8 µm, but do not assume whether the clump is located within a spiral arm or in an inter-arm region. The typical distance uncertainty is between 10–20%.

The integrated flux density of the cores comes from our dendrogram extraction, following the "clipped" paradigm (see Rosolowsky et al. 2008). Since we care about the cores as being overdensities, by using a clipped method we minimise the contribution from the background on the mass estimates, which could be particularly

---

[1] `http://bessel.vlbi-astrometry.org/revised_kd_2014`

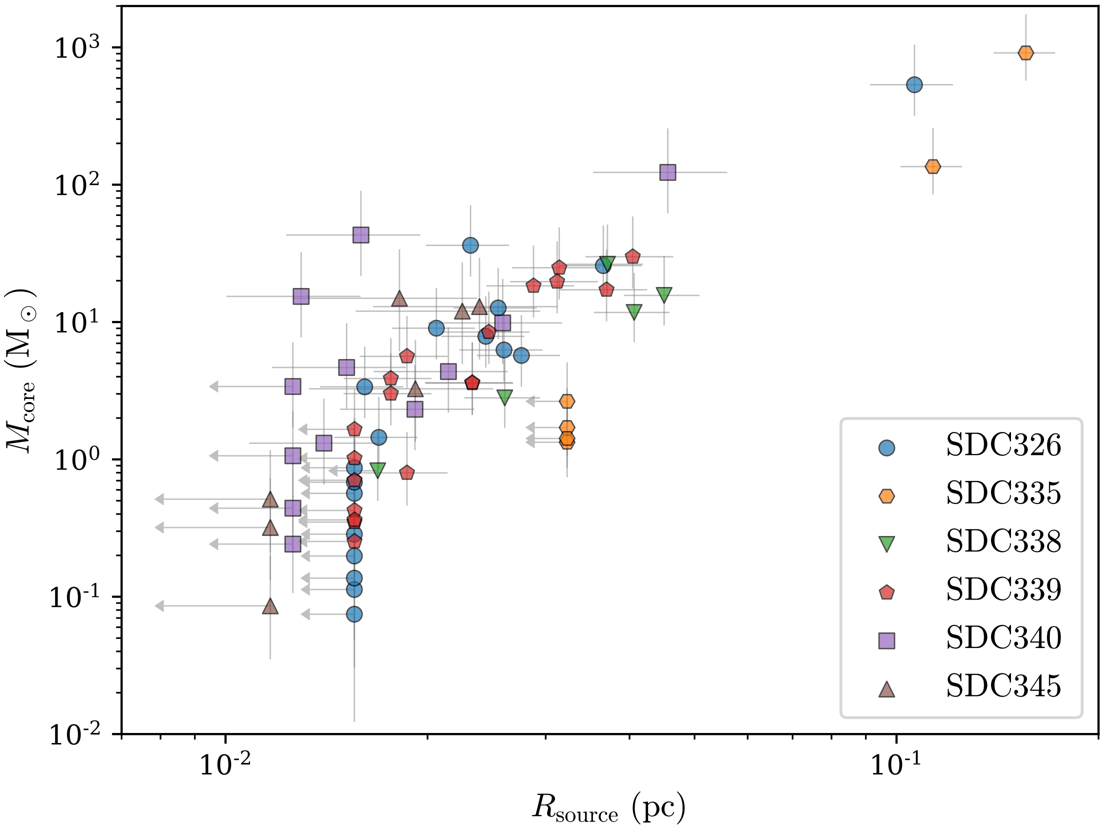

Figure 4.4: Core mass against deconvolved radius ($R_{\rm source}$) plot for all extracted cores from all 6 IRDCs. Upper limits for the radii of unresolved sources are indicated with arrows pointing towards the left.

large for the crowded areas at the centre of the hub-filamentary systems. This way, we are being conservative in the mass estimates, and are possibly underestimating the mass of some of these cores at the centre of the hubs. The error in integrated flux calculated from the quadrature sum within the core mask of our error maps, multiplied by pixel area.

By substituting all of these values into equation 4.1, we obtain masses for all cores. The error in the core masses was calculated using Monte-Carlo methods, by randomly sampling over each variable in equation 4.1, assuming Gaussian errors. We also calculate a deconvolved source radius, $R_{\rm source}$, which is given by

$$R_{\rm source} = \sqrt{R_{\rm eq}^2 - \frac{\theta_{\rm maj}\theta_{\rm min}}{4}} \tag{4.5}$$

where $\theta_{\rm maj}$ and $\theta_{\rm min}$ are the major and minor beam axes, respectively. A table of derived core properties is supplied as online supplementary material, and a plot of core mass against deconvolved radius is shown in Figure 4.4.

We see that our sample of extracted cores cover a broad mass range of 0.07–911 $\mathrm{M}_{\odot}$, with a mean mass of 32 $\mathrm{M}_{\odot}$. The core masses we present here follow the "clipped" paradigm, which subtracts all of the flux below the core's contour in the dendrogram (similar to a background subtraction). Our mass sensitivity ranges between 0.12–0.27 $\mathrm{M}_{\odot}$, depending on the field, with the exception of SDC335 for which the mass sensitivity is 1.98 $\mathrm{M}_{\odot}$. Note that these mass sensitivities were calculating assuming a source temperature of 12 K, an hence when cores are assigned a temperature warmer than 12 K they can have a lower calculated mass than our listed sensitivity. Also note that these mass sensitivities correspond to a clipped mass sensitivity, whereas often in literature the "bijective" (i.e. no background subtraction) mass sensitivity is quoted. Given the dendrogram parameters we have used for our extraction, a corresponding bijective mass sensitivity would be ~ 6 times higher (i.e. less sensitive) than the sensitivities quoted above.

Contrary to Csengeri et al. (2017), we do find intermediate-mass cores in the sample, likely due to not using a single $T_{\mathrm{core}} = 25\,\mathrm{K}$ for all cores, the assumption made in their core mass calculations. Two cores (SDC335-MM1 and SDC326-MM1) are exceptionally high mass, at 911 $\mathrm{M}_{\odot}$ and 534 $\mathrm{M}_{\odot}$ contained within a deconvolved radius of 0.156 pc and 0.106 pc, respectively. They also correspond to the two of the largest sources identified. They are therefore excellent candidates for the formation of very high-mass stars.

Four of the HFS contain at least one core with $M_{\mathrm{core}} \geq 30\,\mathrm{M}_{\odot}$ and so, assuming a core to star formation efficiency of 30%, could form at least one high-mass star with $M_{\star} > 8\,\mathrm{M}_{\odot}$.

4.2.3 Core formation efficiencies

As discussed in the introduction, the ability of a clump to concentrate its mass within cores is a fundamental, but poorly understood characteristic of star-forming regions. In this paper we will refer to parsec-scale dense molecular cloud structures as "clumps", within which stellar clusters and large systems can form (Eden et al.

Table 4.1: Core properties of the most-massive cores within each IRDC in our sample, ordered by clump mass. R_{source} is the deconvolved equivalent radius of the core, and R_{clump} is the equivalent radius of the clump. The full table can be found online[2], along with a table of properties for all of the extracted cores.

Core ID	R_{source} (pc)	T_{core} (K)	M_{core} ($M_\odot$)	R_{clump} (pc)	M_{clump} ($M_\odot$)	f_{MMC} (%)	CFE (%)
SDC345-MM2	0.018	14.6	15^{+19}_{-9}	0.27	135	11.0	32.7
SDC338-MM3	0.037	15.9	26^{+25}_{-10}	0.42	213	12.4	27.0
SDC339-MM3	0.040	15.9	30^{+29}_{-12}	0.54	942	3.2	15.3
SDC340-MM1	0.046	46.3	123^{+134}_{-61}	0.53	1768	7.0	11.8
SDC326-MM1	0.106	41.5	534^{+512}_{-216}	0.80	2399	22.2	26.9
SDC335-MM1	0.156	41.2	911^{+835}_{-338}	0.95	3739	24.4	28.2

2012; Motte et al. 2018b). Here, we calculate the core formation efficiency (CFE),

$$\text{CFE} = \frac{\sum_i M_{\text{core},i}}{M_{\text{clump}}} \tag{4.6}$$

which is the sum of core masses in a given clump, divided by the clump's mass. This tells us how much of a clump's mass is contained within compact sources. The clump masses are obtained from *Herschel* column density maps (Peretto et al. 2016), where the clump boundary is defined by the H_2 column density contour at $N_{H_2} = 3 \times 10^{22}\,\text{cm}^{-2}$.

As far as massive star formation is concerned, another quantity of interest is the fraction of the clump mass contained within its most-massive core (MMC),

$$f_{\text{MMC}} = \frac{M_{\text{MMC}}}{M_{\text{clump}}}. \tag{4.7}$$

Table 4.1 shows a summary of the properties of the MMCs for each IRDC in our sample, and the CFE for each clump. We see that the CFE varies between 11–33%, while f_{MMC} ranges between 3–24%. Note that the CFE calculated here does not take into account the variation in sensitivity between each field, and hence are not directly comparable.

4.3 The relationship between clump and core masses

4.3.1 Broader sample of clumps and cores

In order to get a sense of how f_{MMC} values from our hub sample compare to those from a less biased Galactic plane population of clumps, we use the Csengeri et al. (2017) sample of high-mass ATLASGAL sources observed with ALMA (Project ID: 2013.1.00960.S; PI: Csengeri). This sample contains 35 clumps that have been observed with ALMA ACA at 878 µm (Band 7). These ACA data have similar angular resolution as ours, with a mean beam size of $3.8''$. Also, the distance of these clumps span a very similar range ($1.3\,\mathrm{kpc} < d < 4.2\,\mathrm{kpc}$) to our set of sources. Note that as our 7 m+12 m observations are Band 3, and hence the dust emission we are comparing between datasets may arise from slightly different layers of the cores.

For consistency we use the same procedure for source extraction as described in Section 4.2.1. However, note that the Csengeri et al. (2017) observations are single-pointing only, and are somewhat less sensitive. We therefore cannot compare the CFE values from both samples, and instead focus on comparing f_{MMC}. Core temperatures and clump masses for the Csengeri et al. (2017) sample are estimated in the same way as for our sample of clumps (see Section 4.2).

Three of the clumps overlap between our samples, so we preferentially choose extracted fluxes from our data due to greater coverage, sensitivity and resolution. In the two instances where two clumps share the same $N_{\mathrm{H_2}}$ contour, we merge the clumps and assign it the name of the "original" clump containing the brightest source. Given that our method to measure clump mass is dependent on *Herschel* coverage, one source from the Csengeri et al. (2017) sample has been discarded. The joined sample therefore contains 35 clumps in total, and within those clumps we detect 129 cores. *Spitzer* 8 µm cutout images of each clump are shown in Figure A.1, with the *Herschel* column density contours (that define our clump boundaries) overlaid.

4.3.2 Clump classification

In this paper, we use two distinct clump classification schemes, one that qualitatively identifies the amount of star formation activity within it, and another that determines whether or not a most massive core is at the centre of a hub filament system. Despite both schemes having their own limitations (see below), they can still provide insight into the time evolution of the clumps for the former, and the filamentary environment of the most massive cores for the latter.

We first classify the clumps based on the mid-infrared brightness within the $N_{\mathrm{H}_2} = 3 \times 10^{22}\,\mathrm{cm}^{-2}$ contour used to define the clump boundaries (see Figure A.1). Infrared brightness has recently been shown to be a reliable time evolution tracer (Rigby et al. (2021); Watkins et al. in prep). We classify clumps into three infrared brightness category, from the less evolved to the more evolved: “IR-dark”, no 8 µm extended emission within clump, prominent extinction features; “IR-bright”, significant 8 µm extended emission within the clump, without prominent extinction features; or “Intermediate”, having both clear extinction and emission features within the clump. This classification is made by eye, and is therefore subject to some subjectivity, especially for borderline cases. However, it still provides a reasonable classification of the inner star formation activity of a clump. Out of the the 35 clumps, we classify 13 as IR-dark, 16 as Intermediate, and 6 as IR-bright.

Clumps are then further classified as either HFS or non-HFS according to the location of the most massive core with respect to its local network of filaments. For that purpose we utilise a Hessian-based method, similar to Schisano et al. (2014); Orkisz et al. (2019), to extract filamentary structures from *Herschel* 250 µm images of the clumps. We then classify a clump as a HFS if there are at least three filaments pointing towards the location of the most massive core. One caveat of this method is the relatively low angular resolution of the *Herschel* 250 µm image compared to the ALMA data ($\sim 18''$ vs. $\sim 3''$) which prevents us from making a robust association between filaments and cores. Also, for the same reason, a lot of the filamentary structures within the clumps will not be resolved or even identified. We therefore use the *Spitzer* 8 µm images in conjunction with our extracted filaments to inform our final classification, by checking each one of the clumps for filamentary structures

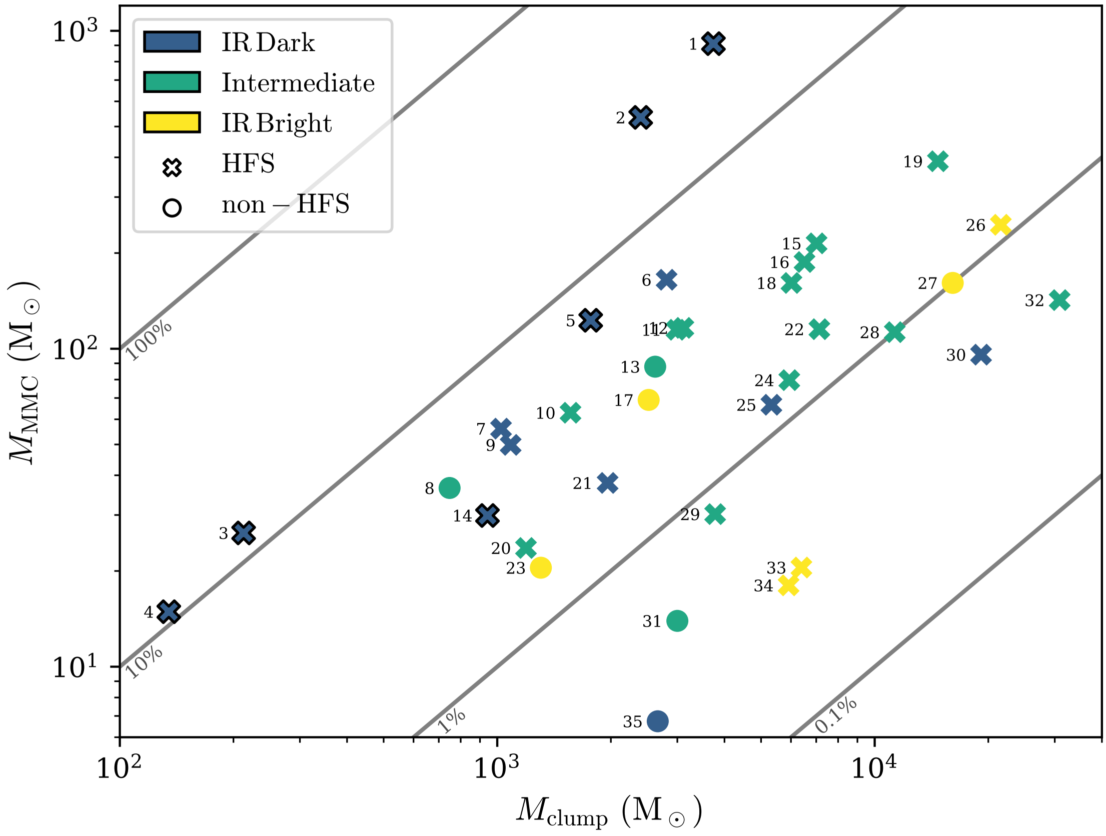

Figure 4.5: Mass of clumps against the mass of the most-massive core (MMC) within that clump. The crosses represent clumps that have been classified as HFS, and circular points are non-HFS clumps. The point fill colours represent the three IR-brightness classes. Points with a black outline are sources observed at 2.9 mm, and points without outlines are sources observed at 878 µm. Clump 20 (G339.6802-1.2090) had no *Spitzer* 8 µm coverage, so *WISE* 12 µm was used for IR-brightness classification. The diagonal grey lines represent lines of constant $f_{\rm MMC}$.

seen in extinction at 8 µm. Instances where clumps were classified by *Spitzer* 8 µm are noted in our table of MMC properties as part of the online supplementary materials associated with this paper.

Out of the 35 clumps, 28 are classified as hubs and 7 as non-hubs, making our sample hub-dominated. This is likely to be a consequence of how the sample has been built: the merging of 6 infrared dark hubs with a sample of 29 massive clumps, which are known to often be associated to hubs (Kumar et al. 2020).

4.3.3 Mass concentration within most massive cores

One argument is that a clump's ability to form high-mass stars is directly linked to the amount of material within that clump (Beuther et al. 2013). Therefore we first investigate the relation between the clump mass (M_{clump}) and the mass of their most massive cores (M_{MMC}). Figure 4.5 shows that, when considering the entire clump sample, there is only a fairly moderate correlation between these two quantities, with a Spearman's rank correlation coefficient $r_{\mathrm{S}} = 0.535$, and a p-value=0.0009. It is possible that this correlation may be influenced by the sparse sampling of the parameter space below a clump mass of $< 1000\,\mathrm{M}_{\odot}$. Above a clump mass of $1000\,\mathrm{M}_{\odot}$, the distribution of the most massive core mass is fairly uniform between $10\,\mathrm{M}_{\odot}$ and $1000\,\mathrm{M}_{\odot}$, suggesting a wide range of f_{MMC} values. If we exclude all datapoints (4 clumps) with $M_{\mathrm{clump}} < 1000\,\mathrm{M}_{\odot}$, then we obtain a correlation coefficient of $r_{\mathrm{S}} = 0.447$ (p-value=0.01), which is moderately weaker than for the full sample. However, if we now only consider the 6 new infrared-dark hubs we observed, we notice that the correlation, even though less statistically significant, is much stronger, with a correlation coefficient of 1 (p-value=0). We speculate that this could point towards a time-dependent correlation between clump and core mass. We will discuss that point further below.

A tight correlation between clump mass and mass of the most massive core was found by Lin et al. (2019) for a sample of ATLASGAL clumps covering a large range of evolutionary stages as traced by their luminosity to mass ratio. This is at odds with the results discussed above for the full sample. It is likely that the tight correlation observed by Lin et al. (2019) is artificially driven by the small range

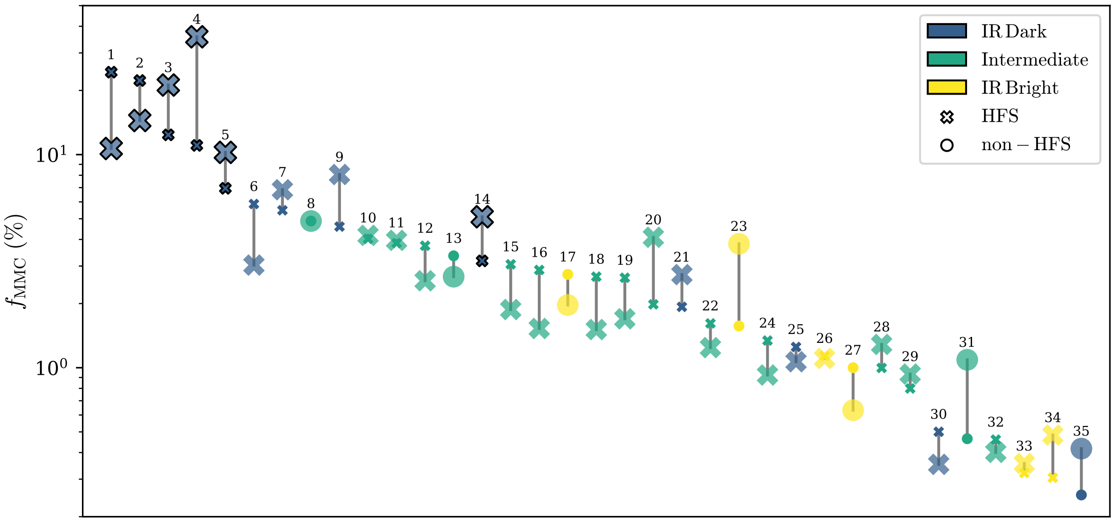

Figure 4.6: Fraction of each clump's total mass contained within its most-massive core (f_{MMC}). The larger points are the f_{MMC} values multiplied by the median core radius of 0.07 pc over the radius of that MMC. The number above each pair of points represents the clump ID number.

of scales they probe, typically 0.3 pc for what they call cores and 0.7 pc for their clumps (a factor of ∼2.3). In our study, the range of scales we probe between the median core size (0.08 pc) and the median clump size (1.5 pc) is a lot larger, a factor of ∼18.8, therefore probing clearly distinct structures.

The smaller set of symbols in Figure 4.6 shows the same information as presented in Figure 4.5 but in the form of f_{MMC} values, with each clump marked by their unique ID number. The points use the same colour scheme as used in Figure 4.5. What is apparent is that some of the clump categories, such as IR-dark clumps, have on average larger f_{MMC} values than others. However, one possible bias that may affect such comparison is the difference in core radii, with some cores being more massive simply by being much larger. In order to remove that bias, the larger set of symbols in Figure 4.6 shows the same quantity as the small set of symbols but rescaled by the median R_{eq} of the MMCs (0.07 pc) over the the core's R_{eq}. By doing this rescaling we effectively compare f_{MMC} at the same core radius, assuming that the density profiles of these cores scale as $\rho(r) \propto r^{-2}$ (Bontemps et al. 2010; Svoboda et al. 2019). We now see that, even though there has been a bit of reshuffling, the

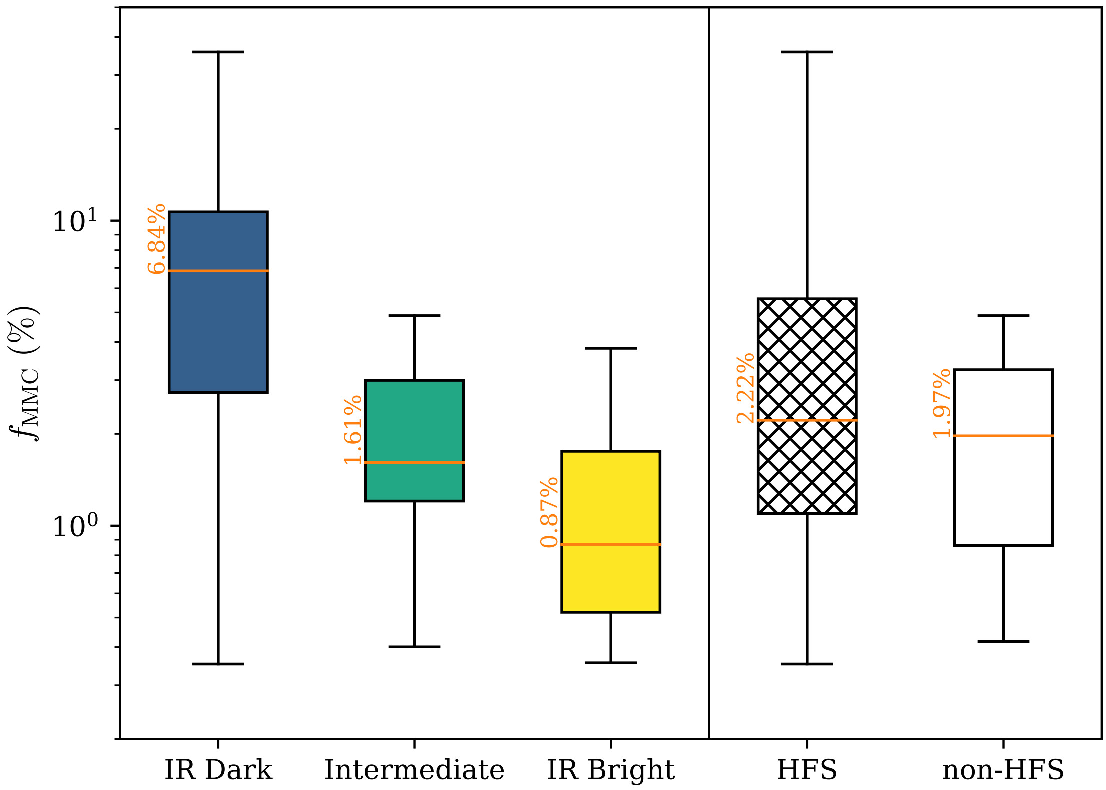

Figure 4.7: Distribution of rescaled $f_{\rm MMC}$ values for each of our clump categories. The orange lines represent the median (with the values also in orange), the boxes represent the interquartile range (IQR), and the "whiskers" represent the full extent (i.e. the 0th and 100th percentile) of the data. From left to right, there are 13, 16, 6, 28, and 7 clumps in each category.

individual $f_{\rm MMC}$ have not drastically changed.

Figure 4.7 shows the distribution of rescaled $f_{\rm MMC}$ values for each category of clump, where the orange line represents the median, and the "whiskers" of the box-plot represent the full extent of the data. We can see that IR-dark clumps have a median $f_{\rm MMC}$ around 7.9 times higher than IR-bright clumps, while Intermediate clumps have a median value 2.4 times higher than their IR-bright counterparts. Even more striking, is our sample of 6 IR-dark hubs (see Table 4.1) that have a median value of 12.6%, which is 14.5 times higher than IR-bright clumps. In contrast, the median rescaled $f_{\rm MMC}$ values for HFS and non-HFS are only separated by a factor of 1.1. Although the median values for these two clump categories are close, the distributions shown in Figure 4.7 appear to be different. To test this we perform a two-sample Kolmogorov–Smirnov (K-S) test to check whether the two samples (HFS and non-HFS) come from the same distribution. We find a p-value of 0.705 for the test, and therefore we cannot reject the null hypothesis that the two samples were drawn from the same distribution (at a significance level of 5%). Although this test is inconclusive, it is likely biased by the very small sample size of non-HFS clumps.

4.4 Conclusions of this chapter

As shown in Section 4.3, although the distributions of $f_{\rm MMC}$ values of the hub filament and non-hub systems appear different, this apparent difference is not statistically significant. Whether this is because there is a common mass concentration efficiency between the two types of clumps or due to the small size of the non-HFS sample and a bias in the sample construction is unclear. Distinguishing between these two possibilities requires observations of a larger, well selected sample of non-HFS sources. The source selection bias is such that we are, by construction, focusing on high-mass star-forming clumps. These have been shown to be preferentially associated to hubs (Kumar et al. 2020). As a result, we may be missing out on a large population of non-hub clumps that have much lower $f_{\rm MMC}$ values. The relatively low resolution of the data used to derive filament skeletons (compared to the ALMA data used for core characterisation) may cause us to mis-classify a large fraction of clumps altogether, in either direction, which would lead to averaging out $f_{\rm MMC}$ values for both hub and non-hub clumps. As it is, we believe that we cannot provide any robust conclusions on the ability of hubs to concentrate more mass within their most massive cores compared to non-hub clumps.

Interestingly, Figure 4.7 shows a clear trend of $f_{\rm MMC}$ values with our infrared brightness classification, $f_{\rm MMC}$ decreases by more than one order of magnitude when going from IR-dark to IR-bright clumps. If one takes this infrared brightness classification as a rough proxy for time evolution, then our results suggest that the clump efficiency in concentrating mass within their most massive cores decreases with time. When inspecting in details the origin of this decrease, we realise that this trend is due to an increase of median clump masses (IR-dark clump: 1961 $\mathrm{M_\odot}$; Intermediate clump: 4859 $\mathrm{M_\odot}$; IR-bright clump: 6155 $\mathrm{M_\odot}$) and not due to a decrease of median core masses (IR-dark clumps: 56 $\mathrm{M_\odot}$; Intermediate clumps: 114 $\mathrm{M_\odot}$; IR-bright clumps: 45 $\mathrm{M_\odot}$). Note as well that the sub-sample of 6 infrared dark hubs we observed displays the highest median f_{MMC} value (12.6%) of all categories. While it is not completely clear what bias in the way we selected these 6 sources is responsible for driving such high f_{MMC} values, in the context of the trend discussed above, these sources represent some of the earliest stages of clump evolution (with a median clump mass of 1355 $\mathrm{M_\odot}$ and a median MMC mass of 76 $\mathrm{M_\odot}$). It is possible that we overestimate the temperatures of our clumps, and hence underestimate their mass leading to a potentially artificially higher f_{MMC} for IR-dark clumps in particular. However Figure 3 in Peretto et al. (2016) shows that for a clump with a mean temperature of 12 K and $N_{\mathrm{H_2}} \geq 3 \times 10^{22}\,\mathrm{cm^{-2}}$ the column density (and therefore mass) is at worst underestimated by ~30%, far from the factor of 3 required to bring the median mass of IR-dark and IR-bright clumps in line.

In light of these results, we propose a scenario in which HFS are formed very early on during the time evolution of a clump, efficiently funnelling mass into its most massive core. The early global collapse of the clump is likely to be driving force behind the early formation of these cores (Peretto et al. 2013, 2014; Williams et al. 2018). During these early stages of clump evolution the mass of the MMC most likely correlates with the mass of the clump itself. As time goes on, clump mass grows (Peretto et al. 2020; Rigby et al. 2021), accreting matter from its surrounding environment (without increasing the mass of its MMC), resulting in a decreasing f_{MMC} over time.

There are a couple of consequences to this scenario. First, the core mass function (CMF) at early stages is likely to be top-heavy, as observed by Zhang et al. (2015); Motte et al. (2018a). Second, despite the subsequent mass growth of the clump, the most-massive cores that are formed within are those that are formed at early stages. This could be explained by radiative feedback disrupting the hubs after the first few massive stars have formed (Geen et al. 2017), or by mechanisms such as fragmentation induced starvation (Peters et al. 2010). By using observations of the optically thick HCO$^+$(1–0) line, Jackson et al. (2019) measured the level of the blue asymmetry — which signifies the presence of gravitational collapse — in a sample of $\sim$ 1000 MALT90 clumps. The significance of the asymmetry feature was found to decrease as a function of evolutionary stage in these clumps, which would seem to support our proposed reduction in the efficiency of mass concentration over time.

This scenario needs to be further tested by enlarging the sample to cover a wider range of masses, and selected to be representative of the population of Galactic clumps. Mapping the kinematics of these HFS would allow us to look for signatures of clump collapse and accretion, infer whether the filaments in these hubs are really converging, and investigate any link with various core properties. We will address the latter in a following study.

Chapter 5

An ALMA study of hub-filament systems: II. Quiescent filaments converging towards highly dynamic hubs

Hub-filament systems are networks of converging interstellar filaments, often with active star formation at their centres, that may play an important role in high-mass star formation. In the previous chapter (which contained work published in Anderson et al. (2021), hereafter Paper I) I presented an analysis of the core population of a sample of 6 infrared dark hub-filament systems mapped with ALMA. It was shown that a higher fraction of a clump's mass ends up within the most massive core (MMC) of IR-dark hubs compared to a sample of IR-bright clumps taken from Csengeri et al. (2017). Such early massive core formation requires large inflow rates and dynamically active IR-dark clumps. To trace the kinematics of the dense gas, six IR-dark hubs were mapped in $N_2H^+(J=1\text{–}0)$ at $\sim 3''$ resolution with ALMA 12m+ACA. The data show intricate emission structures and complex spectra. In this chapter, I present this data, and derived kinematic properties obtained using a newly developed multiple velocity component, hyperfine line-fitting code.

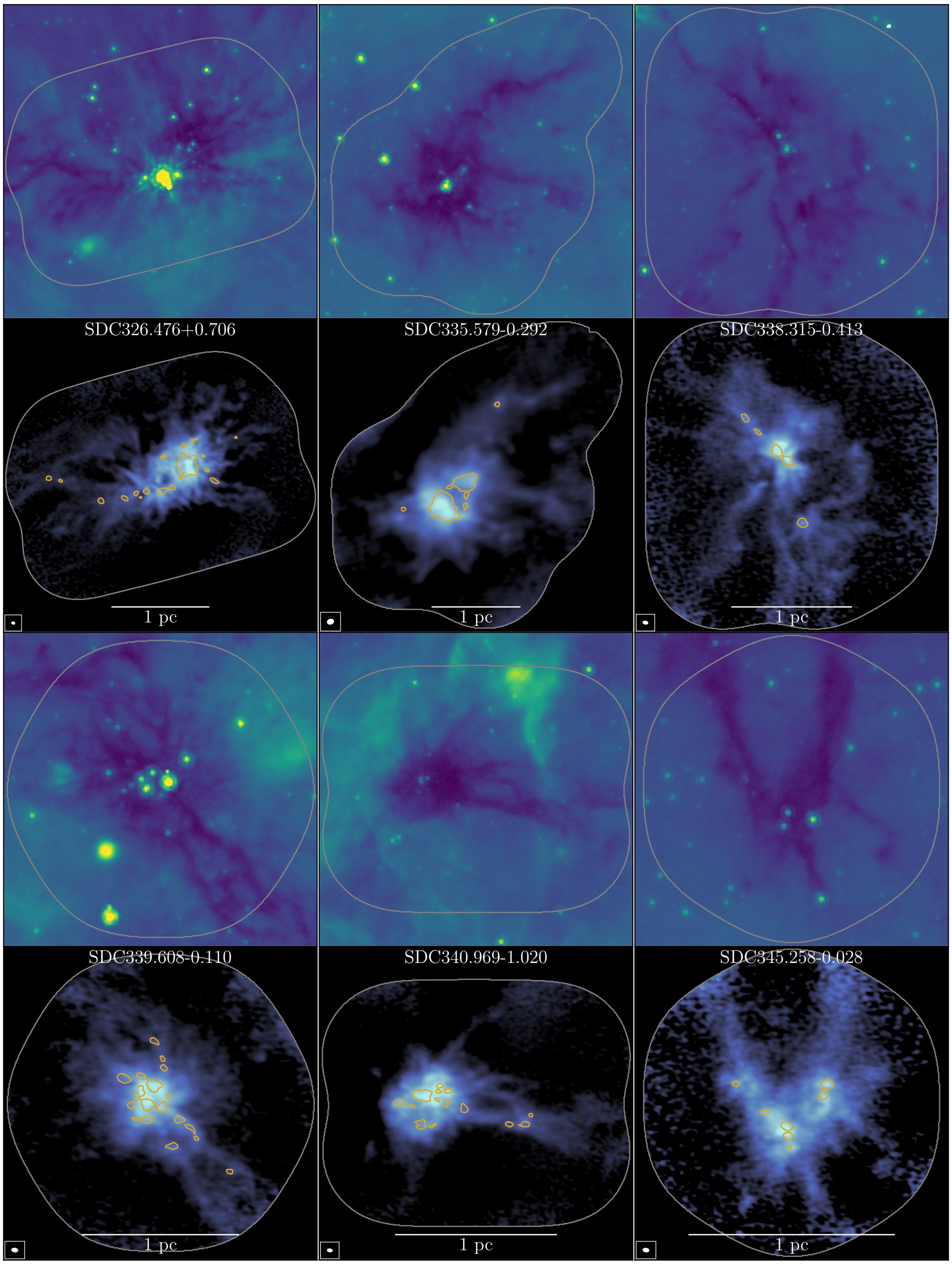

Figure 5.1: *First and third row*: *Spitzer* 8 µm images of the six HFSs we observed with ALMA, showing prominent extinction features in a hub-filament system configuration. Below each *Spitzer* image is the corresponding ALMA combined 12 m+7 m+TP N_2H^+(J=1–0) integrated intensity images. The synthesised beam size of each image is shown in the lower left hand corner, the grey contour shows the extent of our ALMA fields. The orange contours denote the footprints of the 2.9 mm continuum cores (Anderson et al. 2021).

5.1 N_2H^+(J=1–0) integrated intensity maps and column densities

In this paper, we focus on the analysis of the N_2H^+(J=1–0) emission line. With a critical density of $\sim 3 \times 10^4\,\mathrm{cm}^{-3}$ to $6 \times 10^4\,\mathrm{cm}^{-3}$ (Shirley 2015) and a formation path that prevails in low CO abundance environments (i.e. where CO is depleted), N_2H^+ preferentially traces cold, dense gas (Caselli et al. 1995, 2002; Bergin et al. 2002). N_2H^+(J=1–0) integrated intensities have been shown to correlate extremely well with H_2 column densities estimated from dust emission for the $\sim 10^{22}\,\mathrm{cm}^{-2}$ to $10^{23}\,\mathrm{cm}^{-2}$ regime (Tafalla 2001; André et al. 2007; Hacar et al. 2018; Barnes et al. 2020; Peretto et al. 2023). Towards higher column density cores, the agreement between dust emission and N_2H^+(J=1–0) worsen (Peretto et al. 2013), either because of opacity or the destruction of N_2H^+ through the heating of dust grains and the release of CO (Sternberg & Dalgarno 1995; Bergin et al. 2002; Bergin & Tafalla 2007). Another interesting aspect of the N_2H^+(J=1–0) line is that it has a hyperfine structure that can (if the velocity dispersion of the gas is low enough) be resolved into 7 components, some of which blend together in higher velocity dispersion environments into 3 groups of lines, one of which is isolated. As a result of the existence of those hyperfine components at fixed spectral separation, line fitting provides accurate measurements of the gas systemic velocity and velocity dispersion.

Figure 3.1 shows the *Spitzer* 8 µm images of the 6 infrared dark HFSs along with the corresponding N_2H^+(J=1–0) integrated intensity maps. The footprints of the cores identified in Paper I are displayed as orange contours. The visual comparison between the filamentary extinction features and the integrated emission clearly shows that N_2H^+(J=1–0) is an excellent tracer of the infrared dark regions. We also notice that the match between cores is not as good, even though peaks of N_2H^+(J=1–0) integrated intensities are present towards most core footprints.

In order to better quantify the correlation between H_2 column densities and the N_2H^+(J=1–0) integrated intensities, we constructed a pixel-by-pixel density plot of those two quantities. For that purpose, we computed H_2 column density maps from *Herschel* Hi-GAL data (Molinari et al. 2016) for each of the 6 sources using the

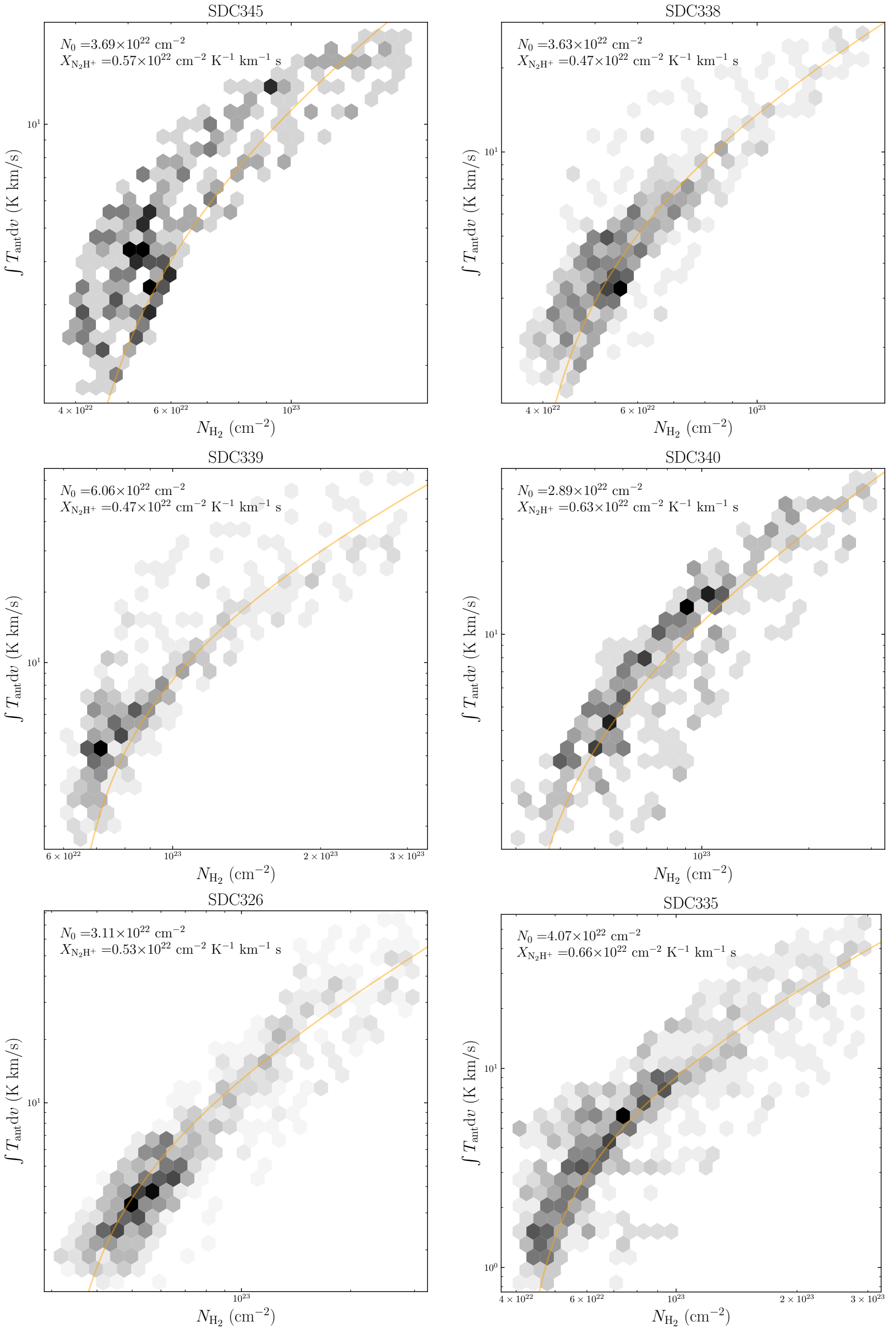

Figure 5.2: Histogram of the ALMA integrated intensity of N_2H^+(J=1–0) against column density derived from *Herschel* for each cloud. The integrated intensity maps were convolved and regridded to match the *Herschel* column density map resolution. The column density offset, $N_{H_2,0}$, is the column density at which we do not detect any N_2H^+ emission in our maps. The orange line represents our best linear fit of the form $\int T_{ant}dv = 1/X_{N_2H^+} \cdot (N_{H_2} - N_{H_2,0})$.

method presented in Peretto et al. (2016). The angular resolution of those maps is 18″. We then convolved our N_2H^+(J=1–0) integrated intensity maps to the same angular resolution and match their astrometry and pixel sizes to the column density maps. The corresponding density plots of each source are shown in Figure 5.2. One can see that, despite the presence of some scatter, both quantities are reasonably well correlated across the board. We further quantify those correlations by performing a linear fit using the following model:

$$N_{\mathrm{H_2}} = N_{\mathrm{H_2},0} + X_{\mathrm{N_2H^+}} \int T_{\mathrm{ant}} \mathrm{d}v \tag{5.1}$$

where $N_{\mathrm{H_2}}$ is the *Herschel* column density, $N_{\mathrm{H_2},0}$ is the median column density where the SNR ~ 1 in our integrated integrated maps, $X_{\mathrm{N_2H^+}}$ is the intensity to column density conversion factor, and $\int T_{\mathrm{ant}} \mathrm{d}v$ is the N_2H^+(J=1–0) integrated intensity. Here, the only free parameter is the conversion factor $X_{\mathrm{N_2H^+}}$. The best fit and associated best fit parameters for each cloud are displayed in Figure 5.2. We notice that the conversion factor $X_{\mathrm{N_2H^+}}$ has very similar values for each of the 6 fields, all within $X_{\mathrm{N_2H^+}} = (0.55 \pm 0.12) \times 10^{22} \mathrm{cm^{-2}\,K^{-1}\,km^{-1}\,s}$. This value is also very similar to the value obtained by Hacar et al. (2018) for the Integral Shape Filament (ISF) in Orion, for a temperature of $\sim 12\,\mathrm{K}$.

New clump masses were derived using these column density conversion factors, which are shown in Table 5.1 alongside the masses derived in Paper I.

5.2 N_2H^+(J=1–0) model fitting

Across the six N_2H^+(J=1-0) sources, over 180,000 spectra with SNR > 10 are detected, amongst which a significant fraction displays evidence of multiple velocity components. In order to extract the gas kinematics out of those spectra, one needs to fit them in an automated way with little to no supervision (i.e. provide the expected number of velocity components and initial guesses of fit parameters).

There are now plenty of fully automated multi Gaussian component fitting algorithm available such as, e.g., SCOUSE (Henshaw et al. 2016), GaussPy+ (Riener et al. 2019), BTS (Clarke et al. 2018). However, when it comes to hyperfine line

Table 5.1: List of sources and their properties from Anderson et al. (2021) (with superscript A+21), and updated properties in this work.

SDC Name (Peretto & Fuller 2009)	d (pc)	$M_{\rm clump}^{A+21}$ (M$_\odot$)	$f_{\rm MMC}^{A+21}$ (%)	$M_{\rm clump}$ (M$_\odot$)	$f_{\rm MMC}$ (%)
SDC326.476+0.706	2610	2399	22.2	3307	16.1
SDC335.579–0.292	3230	3739	24.4	4926	18.5
SDC338.315–0.413	2940	213	12.4	672	3.9
SDC339.608–0.113	2740	942	3.2	1150	2.6
SDC340.969–1.020	2210	1768	7.0	1622	7.6
SDC345.258–0.028	2090	135	11.0	285	8.1

fitting codes, only a few exist, all requiring some level of supervision: CLASS (Pety 2018), PySpecKit (Ginsburg et al. 2022), HfS (Estalella 2017).

At the time of performing this analysis, no fully automated, multiple velocity component, hyperfine line fitting code was available. Hence I have, in collaboration with A. J. Rigby and N. Peretto, developed `mwydyn` (*Welsh*: worm; *pronounced*: muy-din; IPA: [ˈmʊi̯dɪn])[1]. A detailed description of how the code works can be found in Section 5.2.1, and examples of fits are shown in Figure 5.3.

To illustrate the necessity to fit the full hyperfine structure of N_2H^+, Figure 5.4 shows the distributions of FWHM values obtained from calculating the 2nd moment of the isolated component, and fitted FWHM obtained using `mwydyn`. It is immediately apparent that one would obtain a completely different picture of the kinematics when using the moment method, especially in objects similar to those in this sample that contain multiple velocity components, with complex blending.

---

[1] `https://github.com/mphanderson/mwydyn`

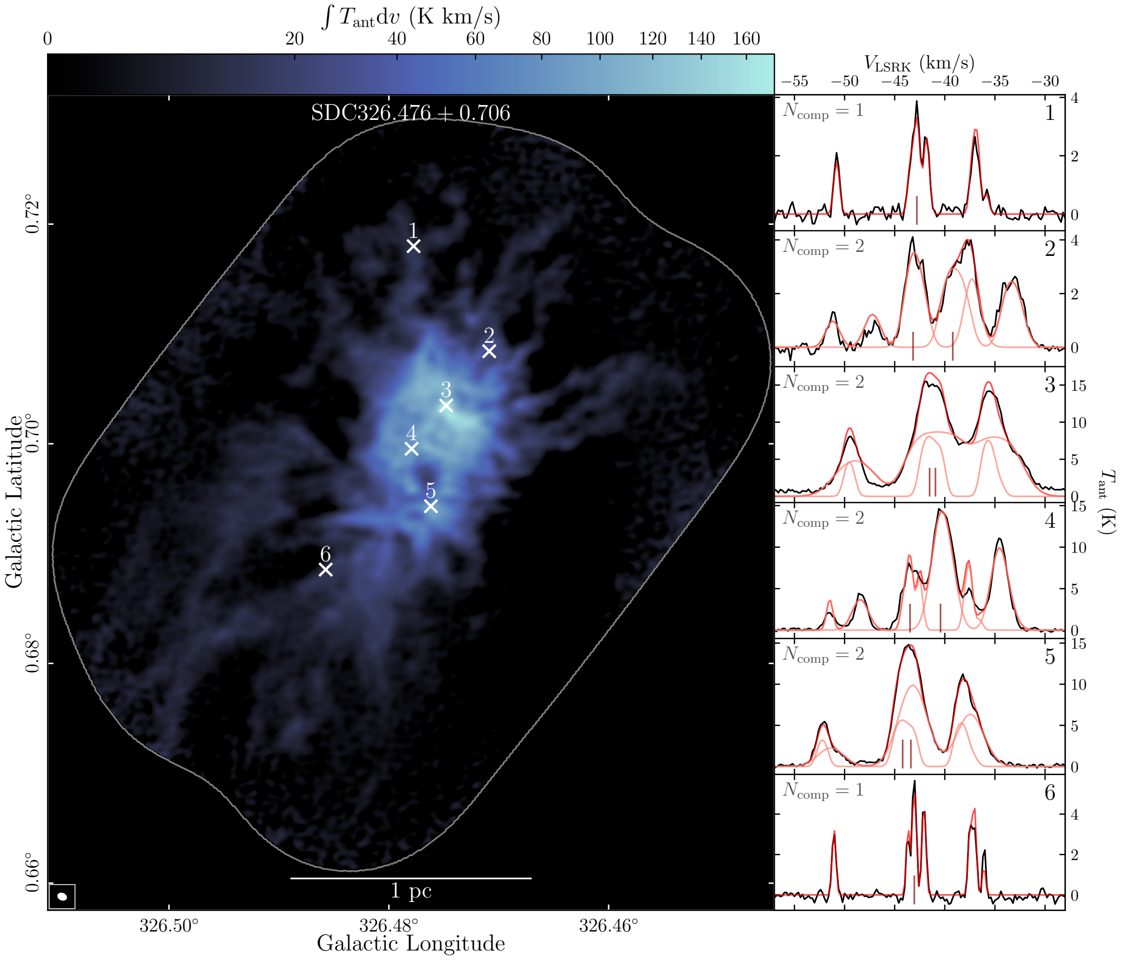

Figure 5.3: Integrated intensity N_2H^+ image of SDC326, with example spectra in black (numbered 1–6, with locations marked on the map), and their corresponding best fitting models produced by `mwydyn` in red. For models comprised of multiple velocity components, their constituent sub-models are shown in light red. The locations of the best fit velocity centroids are marked with dark red vertical bars.

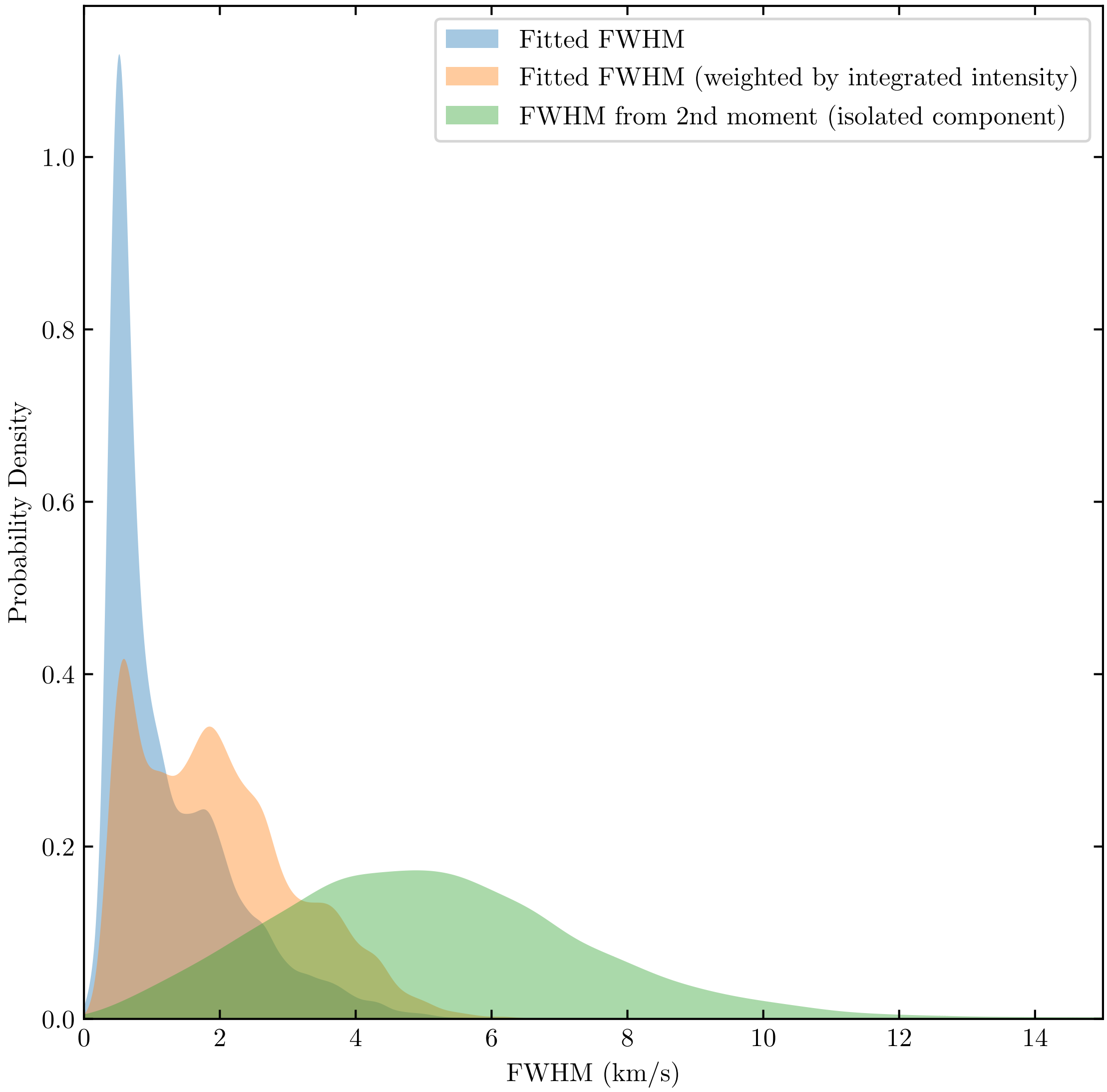

Figure 5.4: Comparison of the distributions of FWHM values for SDC326 obtained by fitting with `mwydyn` (both raw values and weighted by the integrated intensity), and by calculating the 2^{nd} moment of the isolated component.

5.2.1 Description of `mwydyn`

At its core, our fitting programme is based upon the procedure and assumptions of the Hyperfine Structure (HFS) fitting method from CLASS, a set of continuum and line emission analysis tools that is in turn part of the GILDAS[2] software package (Pety 2018). The HFS method fits a molecular line emission spectrum with an individual hyperfine multiplet with a model comprised of four free parameters: $T_{\mathrm{ant}} \cdot \tau$, v_{cen} (centroid velocity of the reference hyperfine component), Δv (FWHM linewidth), and τ_{main}. We also refer to these parameters as p_1, p_2, p_3, and p_4, respectively. The method depends on five assumptions:

1. Each component of the hyperfine multiplet has the same excitation temperature
2. The opacities of each hyperfine component have a Gaussian profile as a function of frequency
3. Each hyperfine component have the same linewidth
4. The multiplet components do not overlap
5. The main beam temperature is well suited to the source (i.e. the source is well resolved)

The total opacity (as a function of frequency/velocity) of the multiplet is expressed as:

$$\tau(v) = p_4 \sum_{i=1}^{N} r_i \cdot \exp\left[-4\ln 2\left(\frac{v - \delta v_i - p_2}{p_3}\right)\right] \tag{5.2}$$

where p_2 is the velocity of the reference component, δv_i is the velocity offset is the i^{th} component of the multiplet (relative to the reference component), p_3 is the FWHM linewidth, r_i is the relative strength of each hyperfine component (where $\sum_i^N r_i = 1$), N is the number of hyperfine components in the multiplet, and p_4 is the sum of the opacities at the component line-centres.

---

[2] `https://www.iram.fr/IRAMFR/GILDAS`

Table 5.2: Parameters of the N_2H^+(J=1–0) multiplet used in our fitting model. The velocities are with respect to the reference component, shown in bold, for which our fitted centroid velocities correspond to.

Transition ($JF_1F \rightarrow J'F_1'F'$)	ν_i (MHz)	r_i	δv_i ($\mathrm{km\,s^{-1}}$)
1 1 0 → 0 1 1	93171.6086	1/27	6.9360
1 1 2 → 0 1 2	93171.9054	5/27	5.9841
1 1 1 → 0 1 0	93172.0403	3/27	5.5452
1 2 2 → 0 1 1	93173.4675	5/27	0.9560
1 2 3 → 0 1 2	**93173.7643**	**7/27**	**0.0000**
1 2 1 → 0 1 1	93173.9546	3/27	−0.6109
1 0 1 → 0 1 2	93176.2527	3/27	−8.0064

The total line profile (for which we use to fit our spectra) is then expressed as:

$$T_{\mathrm{ant}}(v) = \frac{p_1}{p_4}(1 - e^{-\tau(v)}) \tag{5.3}$$

where p_1 is effectively the peak intensity of the line, scaled by the opacity.

For this study, we use the same model of the multiplet parameters for the N_2H^+(J=1–0) transition as used by CLASS, which are shown in Table 5.2.1. We chose the brightest hyperfine component as the reference component for our model fitting.

Our fitting programme `mwydyn` extends the CLASS hyperfine fitting procedure to be able to fit a superposition of multiple hyperfine multiplets to whole datacube of spectra, as well as automating initial guess parameters and determining how many multiplets are required to produce a “good fit”. The algorithm also attempts to reduce the discontinuities between fit parameters between adjacent spectra. Here, we will outline all of the steps the programme runs though over the course of fitting an input data cube.

1. A user specified configuration file is read by the programme that contains various options from the chosen hyperfine line model, fitting bounds and options,

parallel processing settings, and whether to save the data products/produce summary figures.

2. The input data cube is converted into units of brightness temperature (K), from surface brightness (Jy beam^{-1}).

3. The RMS of the first and last 25 channels are used to measure the noise level in all spectra in the cube, generating an RMS map. This is used to ensure that the algorithm only fits spectra that have a peak signal-to-noise ratio (SNR) greater than 10 (for the brightest hyperfine component). When $\tau \ll 1$, this corresponds to an SNR ~ 4.3 for the isolated component.

4. All spectra that satisfy this SNR condition ($\mathrm{SNR_{lim}} = 10$) are cycled through in succesion. Initially one N_2H^+(J=1–0) hyperfine multiplet (hereafter velocity component) is fit to the spectrum. We use LMFIT (Newville et al. 2014), a fitting package based around the Levenberg-Marquardt algorithm for our fitting. LMFIT extends many of the methods from SciPy's `optimize` module and provides a high-level interface for creating custom models, setting fit parameters guesses and bounds, and also adds the ability to algebraically define relationships and constrains between model parameters.

   We set out our initial guesses and bounds on the fit parameters as follows. We locate the brightest channel in the spectrum, and use its amplitude as the the initial guess for p_1 and its velocity coordinate as the guess for p_2. The centroid velocity is allowed to vary ±20 km s^{-1} around this initial guess. The FWHM linewidth p_3 is initialised at 0.5 km s^{-1}, which is roughly the isothermal sound speed at 10 K, and is allowed to vary between 0.1 km s^{-1} to 10 km s^{-1}. As N_2H^+(J=1–0) is generally considered to be optically thin, we set our initial guess of the total opacity p_4 as 0.2, with bounds between 0.1–30, consistent with the range adopted by CLASS.

   In addition to the previously mentioned parameter bounds, we add a constraint that requires that the peak intensity of the total hyperfine multiplet is greater than $\mathrm{SNR_{lim}}$. This is less important for single velocity component fits, however when multiple velocity component models are fit this helps prevent adding numerous additional, very low intensity sub-models that would not be independently detected.

5. After fitting a single velocity component model to the spectrum, the algorithm tries to fit a 2- and 3-component model to the spectrum. In principle the algorithm can fit almost arbitrarily many velocity components, however with each additional submodel another four free parameters are added, the parameter space becomes more complex and slower to explore, and hence the fitting takes significantly longer to converge. Our limit on $N_{\rm comp}^{\rm max} = 3$ is partly motivated by this dramatic increase in computation time, and also by the difficultly in interpreting models that contain more than three velocity components.

   For both the 2- and 3-component fits the initial guess for one of the submodels are set in the same manner as the 1-component model initial guesses. For the second and third set of four parameters (i.e. for the second and third velocity components/submodels), we copy the guesses and bounds for parameters p_1, p_3, and p_4, but adjust our guess for p_2. We estimate the velocity range of detected emission by finding the first and last channels where the emission is greater than $\mathrm{SNR_{lim}}/2 = 5$, meaning that there is less than a 1 in $\sim 1.7 \times 10^6$ chance of a spurious noise spike in one channel registering as a false positive detection of emission. Once we have this range of detected emission, we compare this range to the range of velocities expected from a single hyperfine multiplet ($\sim 15\,\mathrm{km\,s^{-1}}$), i.e. if the range of detected emission is $\sim 18\,\mathrm{km\,s^{-1}}$ then we initialise the guesses for p_2 at $\pm 3\,\mathrm{km\,s^{-1}}$.

6. Once we have the set of 3 models for the current spectrum, the algorithm needs to determine the best fitting model to the data. We use the Bayesian Information Criterion (BIC) for evaluating the quality of the three models. Essentially, the BIC is the log-likelihood with an additional term that penalises the use of models with a larger number of free parameters, thus helping reduce the chance of overfitting. The model with the lowest BIC is chosen as the best model, but when additional velocity components are added (i.e. 2- and 3-

component models) we require that there should be a significant change in the BIC, in this study ΔBIC $= 20$, to adopt a more complex/higher-order model.

In all prior steps, each individual spectrum were fit independently of the surrounding spectra. Although the fitting algorithm explores the parameter space well and generally provides reasonable solutions, there may be some discontinuities in the fit parameters of the adjacent models. Given that spectra that lie within the beam area are correlated, we do not expect strong discontinuities between the models in adjacent spectra. Therefore, we perform several iterations of identifying and refitting spectra whose models strongly deviate from their adjacent counterparts.

7. We cycle through all of our initial best fits for the spectra and identify spectra (one-by-one) within a radius of 2 pixels which have a lower BIC than the spectrum in question. If a spectrum within search radius had a lower BIC value, we refit the current spectrum by taking the best fitting parameters from lower-BIC model spectrum and use those parameters as initial guesses for the refit. If the new model spectrum has a BIC value that is significantly better (i.e. ΔBIC $= 20$) than the previous model, we accept the new model as the best fitting model for the spectrum.

8. We repeat this process 10 times (5 times in each direction) across the cube to ensure that there is ample opportunity for the models to converge on a reasonable, locally consistent solution.

9. With the fitting procedure now complete we write all of the results to a FITS table containing a list of all model/submodel parameters for all fitted spectra. We also write FITS cubes of the model spectra and residuals, and maps such as the the RMS, number of fitted components ($N_{\rm comp}$) in each spectrum, and the final model BIC values.

Note that, as with CLASS, `mwydyn` works for any hyperfine molecular line. We support user generated hyperfine models in the same format as our N_2H^+(J=1–0) model, but note that we have only extensively tested the programme with N_2H^+(J=1–

0). Also, all of the fitting and refitting `mwydyn` is fully parallelised, significantly shortening the runtime on multi-processor machines. The user can specify how many processors they wish to utilise for the programme.

5.3 Fit results

`mwydyn` was run on all six of the N_2H^+ fields, the following subsections present the results.

5.3.1 Global distributions

First, we will present the global distributions of fit properties for the six hub-filament systems in our sample. The two fit properties we will focus on are the fitted velocity centroids ($V_{\rm LSR}$) and the linewidths (FWHM). The global distributions of both of these quantities are shown for each hub-filament system in Figure 5.5, where the distributions have been ordered by clump mass. The two upper panels are the raw kernel density estimations (KDEs) of fitted properties, and the two lower panels have been weighted by the integrated intensity of each submodel. We present these weighted distributions as they will be more representative of the amount of mass in the system that has a given value of a quantity, assuming that the integrated intensity correlates well with the column density/mass. For example there are considerably more spectra in the outer filamentary regions of the clumps compared to the hub region, but there is much less intensity/likely much less mass than in the much brighter central region, so the unweighted distributions are biased by the relative area of the different parts of the clump.

The $V_{\rm LSR}$ distributions are, in broad terms, similar between hubs. All of the distributions cover a similar range in velocity of $\sim 3\,{\rm km\,s^{-1}}$ to $4\,{\rm km\,s^{-1}}$, and there does not seem to be any correlation with the mass of each clump. This appears to be the case in both the raw and intensity-weighted distributions. All of the distributions display small peaks that are separated by roughly $1\,{\rm km\,s^{-1}}$, which may in part be due to the filamentary substructures we see in the data. These peaks somewhat less apparent in the intensity-weighted distributions. It is clear that based on the global

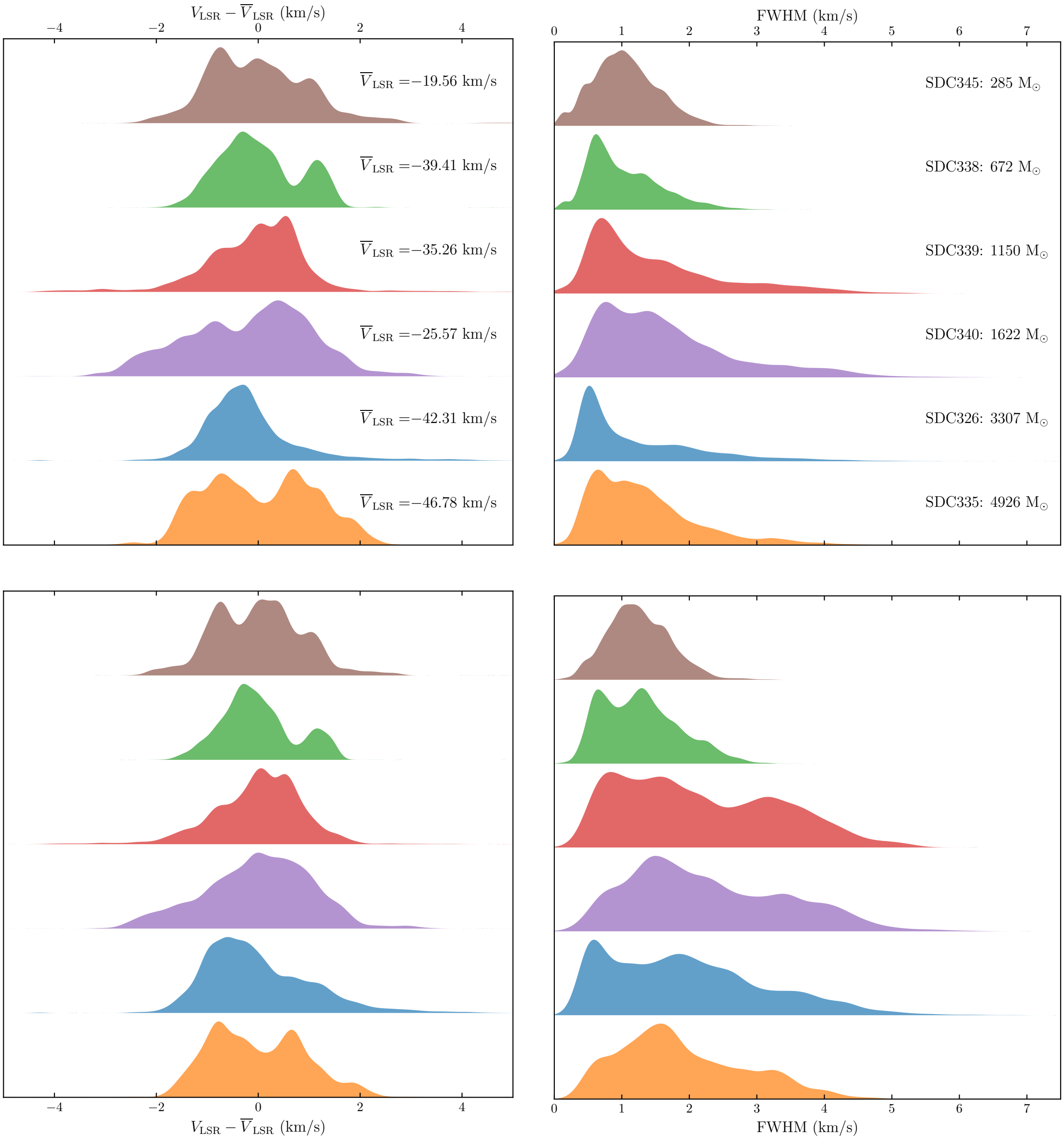

Figure 5.5: *(top row)* Distributions of fitted velocity centroids (left) and fitted FWHM (right) produced by `mwydyn`. *(bottom row)* The same as above, but with each value weighted by the integrated intensity of their respective submodel. The integrated intensity-weighted mean centroid velocity ($\overline{V}_{\rm LSR}$) of each cloud has been subtracted from their respective velocity distributions to aid comparison. The distributions have been ordered in ascending order of clump mass derived from N_2H^+.

distribution of velocities alone we would not be able to tell which clump is the most massive, contains the highest mass core, or has the highest f_{MMC} (see Table 5.1).

The FWHM distributions, shown in the right column of Figure 5.5, are somewhat more varied. Beginning with the unweighted distributions we see that the higher FWHM tail of the distributions is far less extended when looking at the two low mass hubs compared to the higher mass hubs in the sample. One similarity in all of the distributions (except SDC345) is a large peak at a FWHM value of ~ 0.6km/s, accompanied by a strong shoulder at around 1–2km/s. This implies that the vast majority of spectra (in terms of number) are dominated by narrow linewidth, low dispersion gas.

When weighting the FWHM distributions by the integrated intensity of each fitted submodel, the bimodality/multimodality of the distributions become far more apparent, with what was a shoulder in the unweighted distribution appearing more like an additional peak in the distribution. There is also significantly more power in the FWHM wings, indicating that the brighter regions of the cloud generally have a higher linewidth. Overall the location of the mean shifts towards higher FWHM with increasing clump mass, and in particular, the total power above a given FWHM seems to increase with clump mass. This occurs up to around $10^3\,\mathrm{M}_{\odot}$, and then decreases slightly. However, this sample size far too small to make any conclusions regarding this trend.

The lowest mass clump (SDC345) has a peak at a FWHM of around $1\,\mathrm{km\,s^{-1}}$. This is somewhat unexpected given how IR-dark it is, and the relative low-mass of the filaments that comprise the system. That being said the mapped area of SDC345 is limited to the densest central (hub) region, with very little coverage of the outer filamentary structures seen in the *Spitzer* data. This, combined with the overall lower SNR of the $\mathrm{N_2H^+}$ data for this field (likely due to the lower column density of the object) means that only a small, central patch (SNR > 10) of the mapped area was fit by `mwydyn`. We speculate that the lack of a low linewidth peak may be due to the limited mapping of the outermost filamentary structures, or that the energy involved in the formation of the hub has not been dissipated yet.

Throughout the remainder of the paper our analysis will focus on the integrated intensity-weighted fit properties unless otherwise stated. We believe the weighted quantities are more representative of the clump properties as the intensity roughly correlates with column density/mass.

5.3.2 Spatial distributions

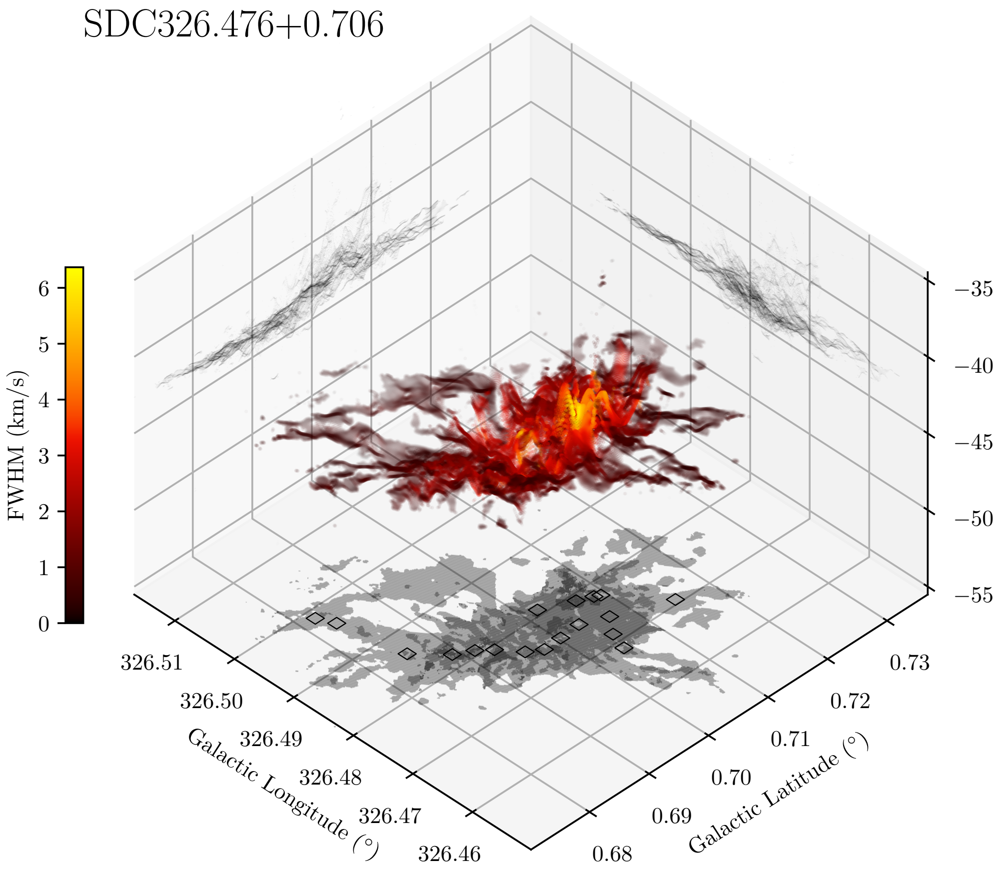

Figure 5.6: 3D PPV Galactic longitude-latitude-V_{LSR} (l,b,v) visualisation of the fit results produced by `mwydyn` for SDC326. The points are coloured by the best fit FWHM value. On the lower box surface is a 2D projection (l,b) showing the number of fitted velocity components along the line of sight, with darker grey indicating more components. The positions of the "cores" presented in Anderson et al. (2021) are marked with diamonds. The left and right surfaces show 2D projections in both (b,v) and (l,v), respectively.

Here we examine how the properties of the gas vary spatially across each of the hub-filament systems. In Figure 5.6 we show a 3D visualisation of our fit results for SDC326 in PPV, i.e. Galactic longitude-latitude-V_{LSR} (l,b,v), with the points coloured by the fitted FWHM. On each of the three "walls"/panels in the plots are 2D projections along the axis perpendicular to the given wall. The bottom surface represents the number of fitted components along the line of sight, with the darkest grey representing 3-component fits, and the lightest grey representing 1-component fits. On the two side panels are position-velocity projections of the fit results (Galactic latitude-velocity and longitude-velocity on the left and right, respectively). The darker grey/black represents a high number of points along that projected axis. The visualisations for all of the hubs are shown in Figure 5.7.

Despite the varied appearance and mass of the six hubs, we see many consistent attributes in their PPV visualisations. All of the structures appear to be coherent in velocity, with filamentary features seen in all hubs. We see that the gas has typically low velocity dispersion in the outskirts, which gradually increases towards the hub centres. The number of fitted components also increases towards the hub centres, with some increases elsewhere that may be due to overlapping filamentary structures or the presence of cores.

In the PV projections, it also appears that the dispersion in the centroid velocities increases toward the centres. Note that the points plotted here are the raw values only, and have not had their appearance modified/weighted by the integrated intensity of their respective submodel. Some of the clumps (SDC339, SDC340) exhibit quite clear clump-scale velocity gradients, which only slight gradients present in SDC335 and SDC345. No clear velocity gradients are seen for SDC326 and SDC338 (at the clump scale).

Small, semi-sinusoidal features are seen in position-velocity in all of the objects. Similar features have been seen in other work (e.g. Henshaw et al. 2016, 2020). It is important to note that these features are not an artefact of the fitting process, as these are clearly visible within position-velocity diagrams shown in Figure 5.12. Although a full characterisation of the filament population will be the topic of a follow-up paper, we can already comment on the fact that a large fraction of those velocity oscillations are not associated with density fluctuations, e.g. cores[3]. It

---

[3] At least, cores detected within our continuum data.

is possible that these oscillatory features originate from some magneto-turbulent fluctuations.

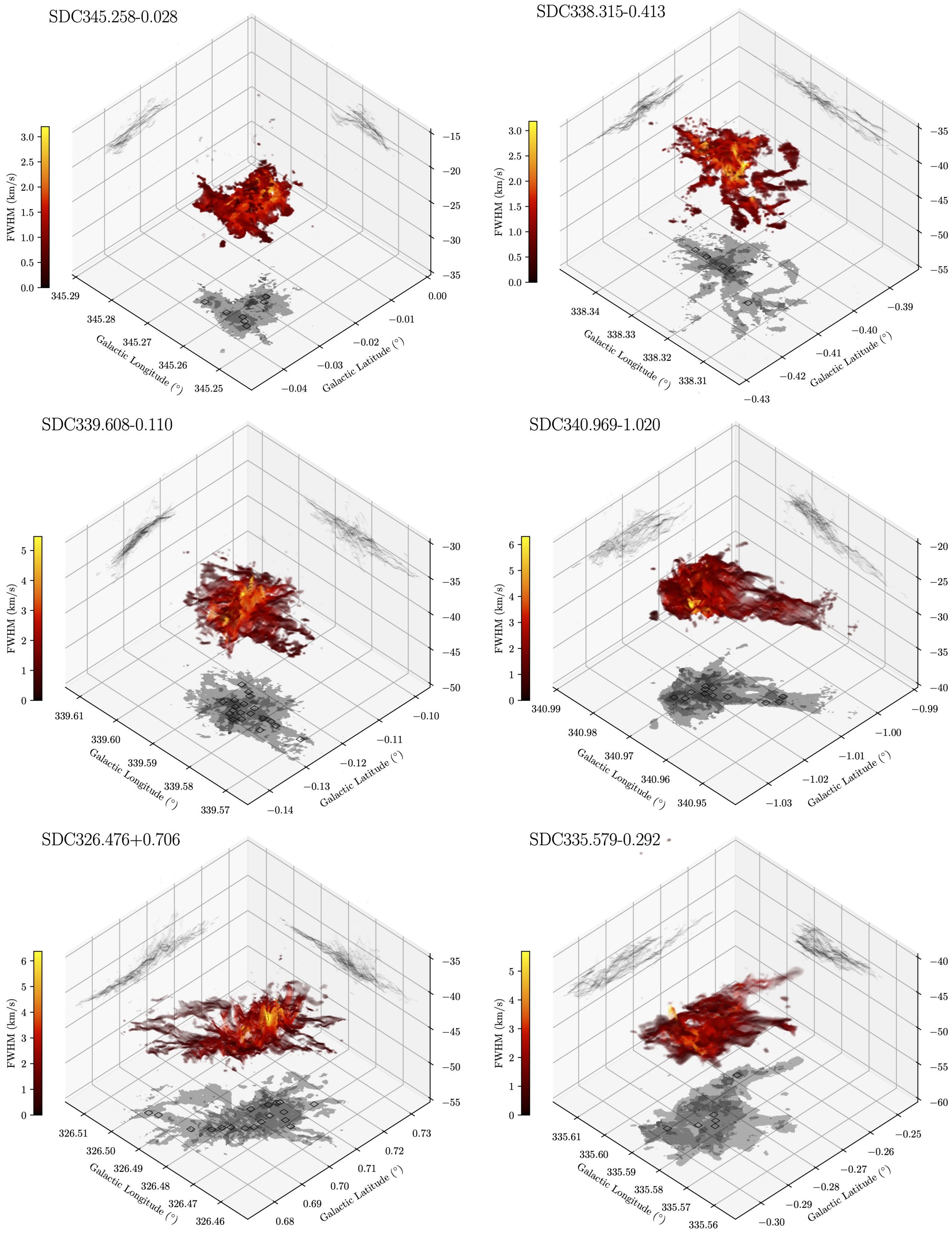

Figure 5.7: 3D PPV (l,b,v) visualisations of the fit results produced by `mwydyn`. The points are coloured by the best fit FWHM value. On the lower box surface is a 2D projection (l,b) showing the number of fitted velocity components along the line of sight, with darker grey indicating more components. The positions of the "cores" presented in Anderson et al. (2021) are marked with diamonds. The left and right surfaces show 2D projections in both (b,v) and (l,v), respectively.

5.3.3 Radial trends

In the previous section we explored some of the spatial features and trends in a qualitative manner. Here, we focus on investigating how the gas properties behave as a function of radius more quantitatively. In Figure 5.8 and Figure 5.9 we divide our hub-filament systems into approximately radial bins, and examine how the distributions of fit properties vary within each annulus. The leftmost panels shows the ALMA integrated intensity N_2H^+(J=1–0) image of the clouds with the contour levels used to generate the radial bins superimposed. The contours have been generated using *Herschel* column density maps (Peretto et al. 2016) of the clouds using a square-root scaling of contour levels. This was used in part because the column density drops off as $\sim 1/r^2$, and hence it was relatively easy to generate contours with equivalent radii ($R_{\rm eq}$) that varied at a similar rate (i.e. the physical scales of each contour are roughly similar) across clouds. Points that do not lie within the outermost contour level are grouped together into one distribution, and hence don't have an $R_{\rm eq}$ value assigned to them, i.e. the $R_{\rm eq}$ of the outermost contour acts as a lower limit on their radius value. The contour levels are colour-coded to match the mean number of fitted components (right column of Figure 5.8), and the distributions of fitted centroid velocities, fitted FWHM, and the calculated total dispersion values (in Figure 5.9).

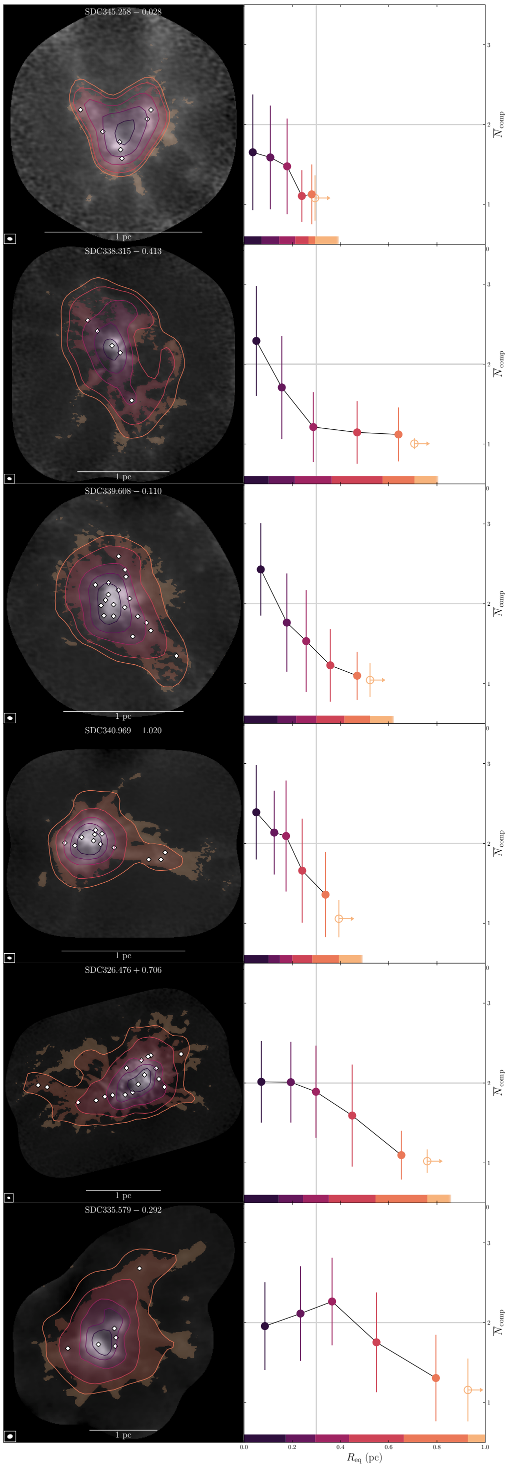

Figure 5.8: *Left*: N_2H^+(J=1–0) integrated intensity images for each cloud, with *Herschel* column density contours at similar physical scales. The positions of the cores presented in Anderson et al. (2021) are marked with diamonds. *Right*: Mean number of fitted components within each contour level as a function of R_{eq}. The vertical lines represent the standard deviation of N_{comp}. The unfilled circles represent the outermost points for which we do not have a reliable R_{eq} value, and hence their radii are lower limits.

First we will consider the mean number of fitted components, i.e. $\overline{N}_{\mathrm{comp}}$ (see Figure 5.8). As seen in the PPV visualisations in Figure 5.7 we see that the number of fitted components increases with decreasing radius. This is largely an expected behaviour given the increased complexity of the spectra that we see towards the centre (e.g. as in Figure 5.3) that require a more complex model (i.e. more components). When comparing these plots between HFSs, we see that the increase in N_{comp} occurs earlier (i.e. at larger radii) in the higher mass systems. Within $R_{\mathrm{eq}} < 0.3\,\mathrm{pc}$, N_{comp} increases steadily from SDC345 (the lowest mass) to SDC335 (the highest mass), with the region of multiple-component fits extending in radius. For the two most massive HFSs, this increase appears to flatten off at around 0.1 pc to 0.3 pc. For SDC335 we see a decrease at smallest radii. This is driven by two factors: i) a ring of spectra around the third contour that show prominent self absorption, which `mwydyn` attempts to fit with higher-complexity models; ii) The spectra in the centremost region are actually rather simple, with one peak, a very broad linewidth and wings, hence generally needing fewer components to model. See Figure 5.10 for a set of example spectra and models for SDC335.

In all cases, Figure 5.9 shows that the FWHM distributions peak at low values (i.e. $< 1\,\mathrm{km\,s^{-1}}$) at large radii, and then increases steadily towards the central contour level. For SDC345, the distributions are fairly similar across all radii, which again may be due to the limited mapping as described in Section 5.3.1. The variations/number of peaks within the distributions also vary steadily with increasing mass until SDC326, with a more subtle change in the distributions for SDC335. Note that the spatial resolution is ~ 1.6 times worse for SDC335 that the rest of the sample, so we are likely resolving less substructure within it.

The velocity centroid distributions overall show a smaller degree of variation (in terms of breadth) with radius, and the variations with radius are less consistent across clouds. In some cases (e.g. SDC335) have trends that are the inverse of trends in the FWHM distributions, i.e. the distributions get simpler and narrower with decreasing radius. These distributions are complex, and it is difficult to extract more meaningful conclusions from them as opposed to the FWHM distributions, especially as the shapes of these distributions will be heavily influenced by the 3D

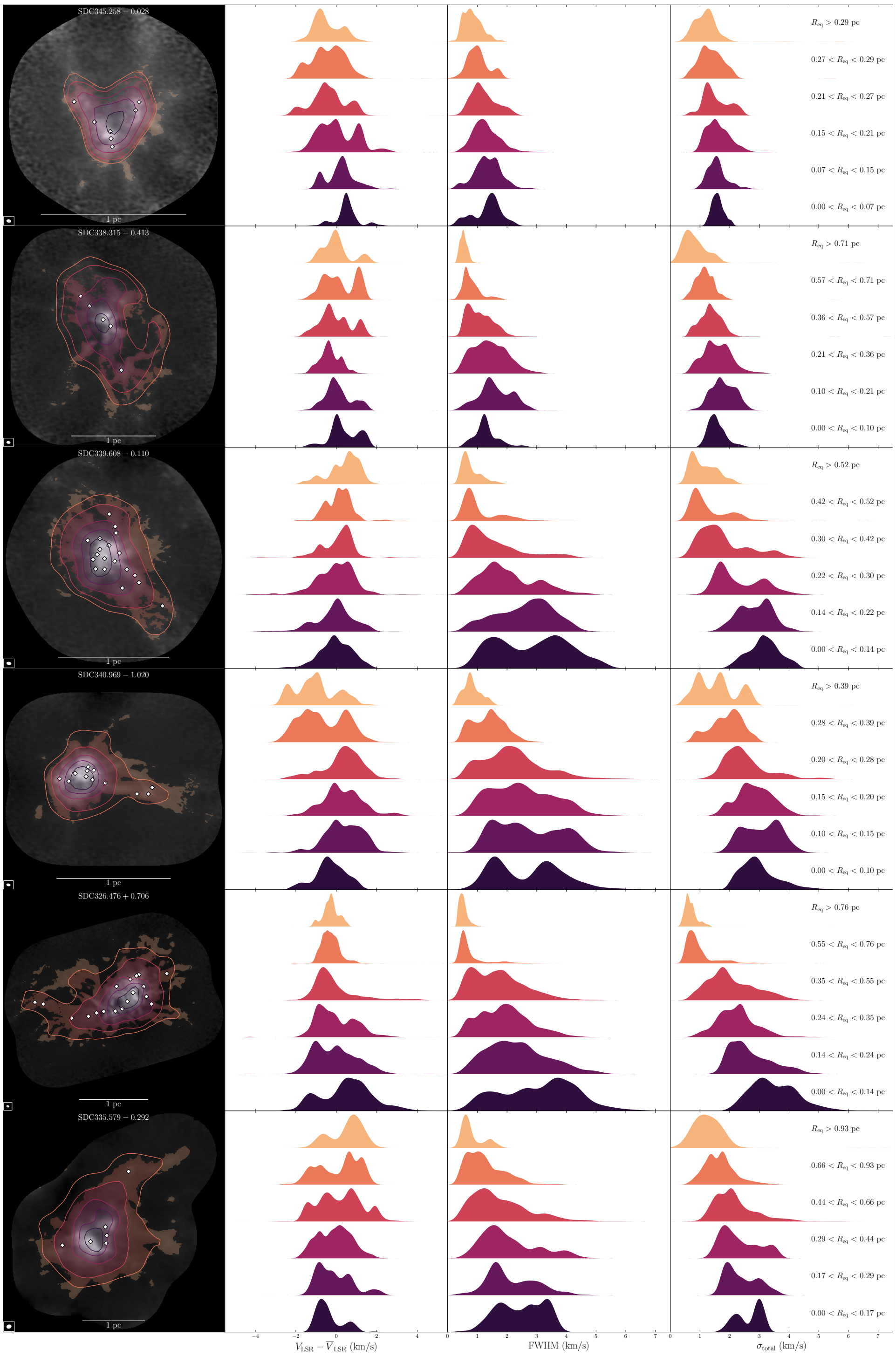

Figure 5.9: *From left to right*: N_2H^+(J=1–0) integrated intensity images for each cloud, with *Herschel* column density contours (same as Figure 5.8), and then the KDEs of the fitted velocity centroids, FWHM values, and the total velocity dispersions colour-coded by their respective contours in the image on the left. The points that do not fall within the outermost contour level are assigned to the topmost distribution. All of these distributions are integrated intensity weighted.

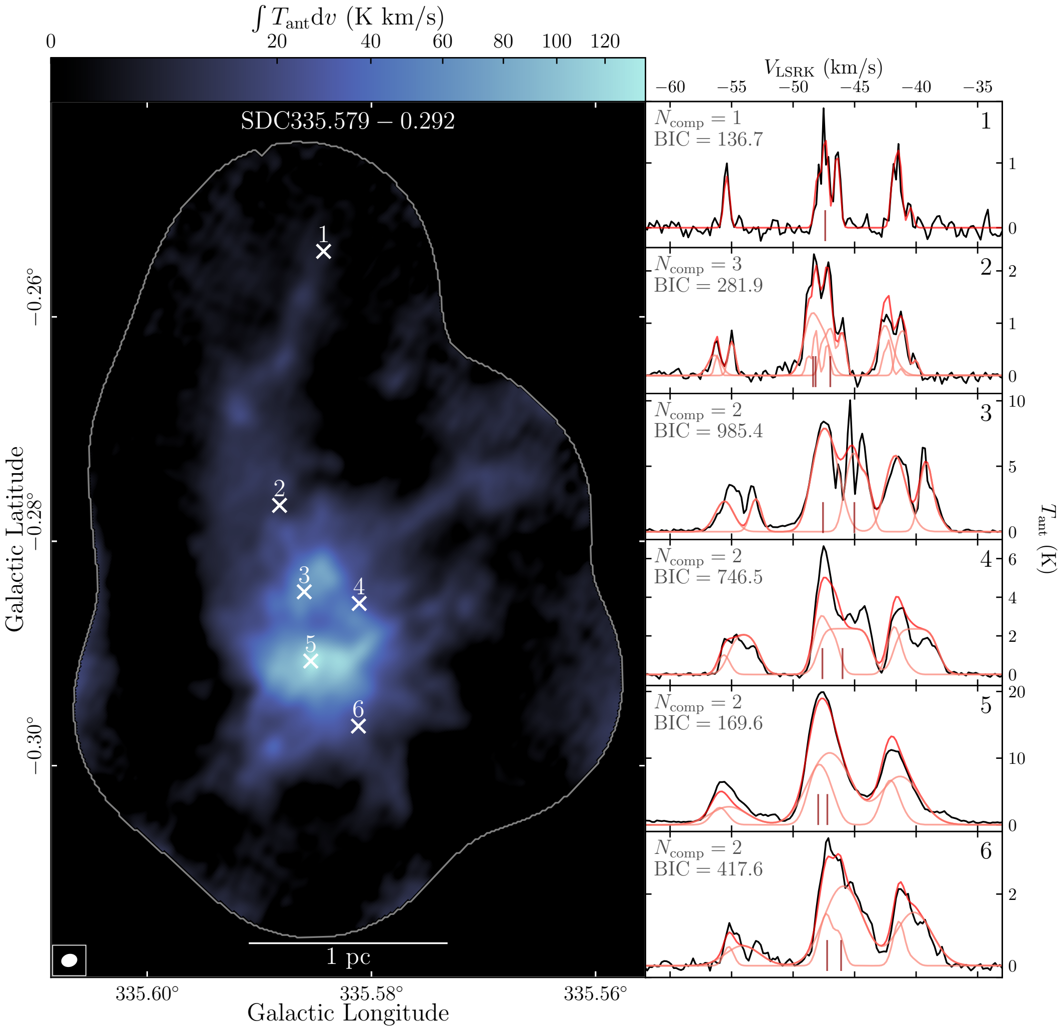

Figure 5.10: Integrated intensity N_2H^+(J=1–0) image of SDC335, with example spectra in black (numbered 1–6, with locations marked on the map), and their corresponding best fitting models produced by `mwydyn` in red.

morphologies of the HFSs and also their 2D projections on the sky.

What is clear from those distributions of velocity dispersions and centroid velocities is that they both carry important information regarding the intrinsic kinematics, and energetics, of the HFSs. In order to obtain a global picture, we decided to combine them in the form of a single total velocity dispersion measurement $\sigma_{\rm total}$ defined as:

$$\sigma_{\rm total} = \sqrt{\frac{\sum_i^{N_{\rm comp}} w_i[(v_i - \bar{v})^2 + \sigma_i^2]}{\sum_i^{N_{\rm comp}} w_i}} \tag{5.4}$$

where the sum is over the number of fitted components $N_{\rm comp}$ in a given spectrum,

v_i and σ_i are the fitted velocity centroid and velocity dispersions of the i-th submodel, w_i is a weight, which in our calculation is the integrated intensity of the i-th submodel, and $\bar{v}$ is the intensity-weighted mean of all fitted centroid velocities in a given cloud. Note that we do not fit for σ_i in `mwydyn`, so this is simply converted from the FWHM by: $\sigma_i = \mathrm{FWHM}_i/2\sqrt{2\ln 2}$.

By comparing σ and σ_{total} one is able to estimate the relative contribution of the velocity dispersion versus the contribution of the inter-component dispersion in the velocity centroids to the kinematics. Figure 5.11 displays such a comparison as a function of R_{eq} (the full distributions of σ_{total} as a function of radius are shown in Figure 5.9). We can see that, at large radii, where the filaments lie, the total velocity dispersion is dominated by the inter-component kinematics, with σ_{total} is around 3 to 4 times larger than σ. At the smallest radii, this ratio drops to around 2 to 3. We also notice that the two clumps with the smallest masses (SDC345 and SDC338) display flatter velocity dispersion profiles that do not peak nearly as much as the other four HFSs and a much more similar relative contribution of σ to σ_{total} at all radii.

5.4 Discussion

5.4.1 The 3D morphology of infrared-dark hub-filament systems

As discussed in the introduction, many models for the formation of HFSs start with the formation of a compressed sheet of gas. However, finding observational evidence of the existence of those sheets is extremely difficult, mostly because of the observations' inability to probe the depth of cloud along the line-of-sight. Using observations of different transitions of the same molecule one can estimate its average volume density which, when combined with a H_2 column density map, can provide the line-of-sight depth of a cloud. In the few cases that this has been done, the filamentary nature of the clouds have been confirmed (Li & Goldsmith 2012; Bonne et al. 2020). However in the case of Musca, a 6 pc long filamentary cloud located

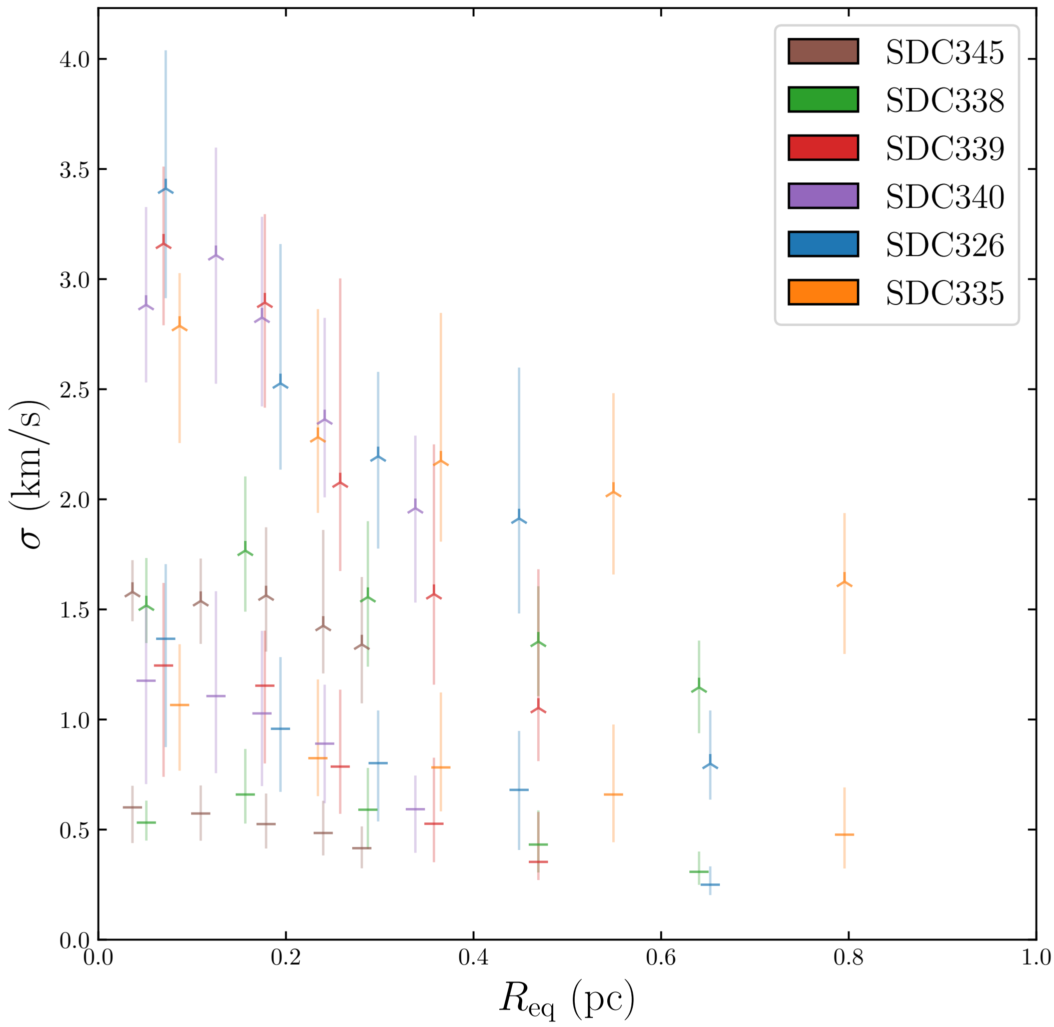

Figure 5.11: Comparison of the median velocity dispersions (horizontal lines), and total velocity dispersions (three-pointed stars) as a function of radius for each cloud. The errorbars represent the interquartile range of the distributions shown in Figure 5.9.

around 140 pc from the Sun, it has been argued that its global morphology is rather that of a sheet-like cloud seen edge-on (Tritsis & Tassis 2018; Tritsis et al. 2022). Here, we discuss the 3D morphology of HFSs through the lens of the their dense gas kinematics.

There is an ever growing body of evidence that parsec-scale star-forming clumps are in a state of global collapse, whether those clumps are hub-filament systems (Peretto et al. 2013; Kirk et al. 2013), or not (Peretto et al. 2006; Schneider et al. 2010; Ragan et al. 2015). In fact, (Peretto et al. 2023) showed that all infrared dark clumps are dynamically decoupled from their parent molecular clouds as the result of clump collapse. It is therefore most likely that the six hub-filament systems targeted within our study are no exception, and are also collapsing. One of the sources in this sample (SDC335), was studied by Peretto et al. (2013), has evidence for being in a state of global collapse.

When considering the velocity distributions in Figure 5.5 (bottom left), the strong similarity of the centroid velocities between HFSs is challenging to explain if one considers a collapsing, fragmented, spheroidal clump. Indeed, gravitational acceleration being larger for the more massive/denser clumps, SDC335 ($M \sim 5000\,\mathrm{M}_{\odot}$) should show evidence of a more dynamic gas kinematics than SDC345 that is ~ 20 times less massive. If the filaments that connect to the hub were to be randomly distributed within a sphere, then differences in gravitational acceleration would appear via broader centroid velocity distributions for the most massive HFSs. This feature is not apparent in Figure 5.5. It is therefore more likely that the six infrared dark hub-filament systems we selected have a sheet-like morphology. However, if those sheets were to be randomly oriented with respect to the plane-of-the-sky we would also see differences in the centroid velocity distributions. Therefore, we propose that our sources are sheet-like clumps seen mostly face-on.

While at first it could seem unlikely to have selected, by chance, only face-on sheet-like clumps, our source selection is heavily biased. Indeed, by design, we selected sources that exhibit clear hub-filament morphologies in mid-infrared extinction, and in order to see such clear morphologies in sheet-like clumps, they are likely be close to be face-on, similar to the Monoceros R2 HFS presented in

Treviño-Morales et al. (2019).

The distributions shown in Figure 5.5 (bottom right) show that, unlike the centroid velocities, the FWHM (and hence velocity dispersion) does correlate with mass, i.e. broader distributions for higher mass clumps. However, those large velocity dispersions are not randomly spread over the clumps, but are almost exclusively located within the central hub regions, as opposed to the filament regions (see Figure 5.11 and Figure 5.9). This strongly suggests that the dynamical interaction of the gas flows channelled by the multi-directional filaments (but confined to a sheet) produces a randomisation of the flow directions in the hub region, leading to larger line-of-sight velocity dispersions. Figure 5.12 shows position-velocity diagram going across the hub-filament systems in order to illustrate the clear change in the gas kinematic properties when moving from the filament region to the hub.

5.4.2 The filament-hub transition

The picture depicted above is based on the assumption that parsec-scale hub-filament systems are globally collapsing. However, as a result of the sheet-like morphology of the HFSs and their near plane-of-sky orientation, direct evidence of collapse via the observed line-of-sight velocity field is marginal. For instance, on the clump scale, Figure 5.7 shows clear velocity gradients only for SDC339 and SDC340. Since velocity gradients are usually the main evidence upon which claims of collapse are made, it is not directly obvious that the six HFSs studied here are globally collapsing.

Another example is provided by Lee & Hennebelle (2016a), where a flattened protocluster is forming at the centre of a globally collapsing cloud. The protocluster itself is stabilised as the result of its rotational and turbulent energy while accreting mass from its surrounding cloud. Although not directly addressed in that paper, such a scenario might lead to a similar accretion shock at the outer boundary of the protocluster.

This shock may be caused by the inflowing gas, driven by the global collapse of the parent clump, hitting the outer boundary of the growing hub. How such shocks would appear in N_2H^+(J=1–0) observations is not clear. However the relatively

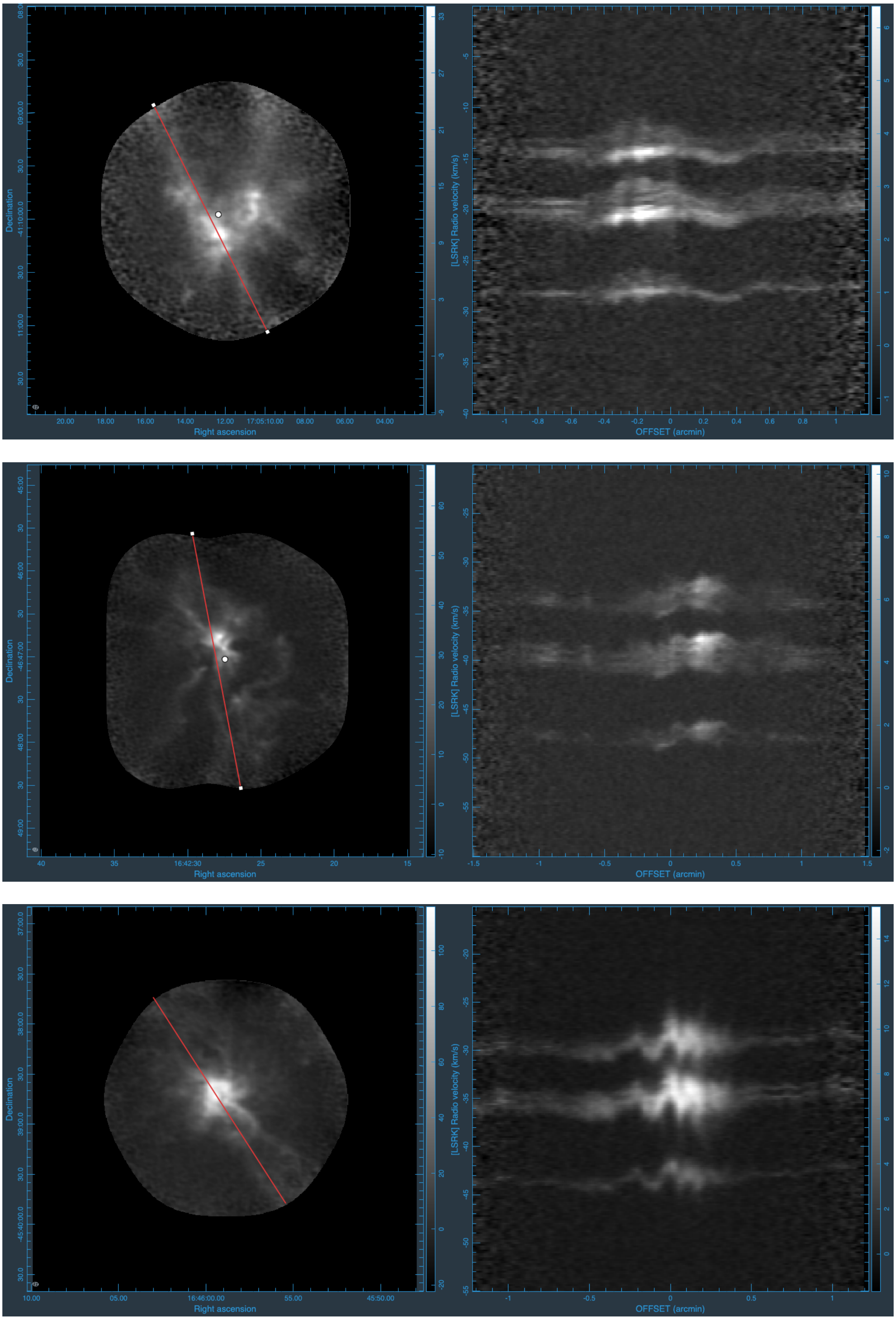

Figure 5.12: The left panels show the integrated N_2H^+(J=1–0) intensity data for each hub, with the red line indicating the axis for which the PV-diagram was generated, which are show in the right panels.

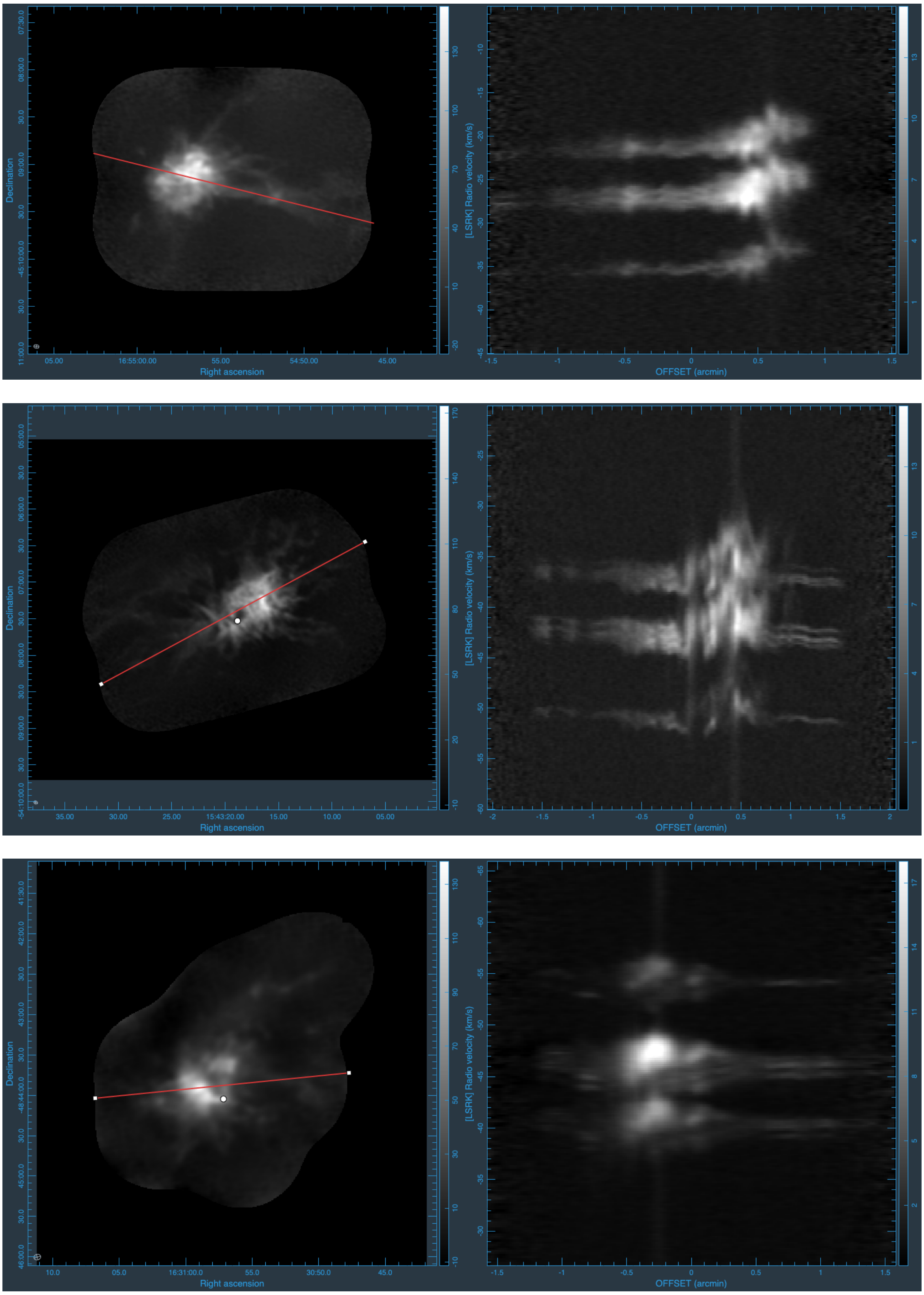

Figure 5.12: *(continued)*

sharp transitions that appear in at least some of the PPV visualisations and PV diagrams (where the filaments meet the hub centre), might be evidence of such a process.

Note that the implication is that hubs are dynamically decoupled from the surrounding collapsing gas, and whether they are in a state of collapse or supported by a combination of rotational, magnetic and turbulent energy remains to be shown.

5.4.3 The energy budget of hub-filament systems

In order to determine the energy budget of the six HFSs we constructed their virial ratio profiles, i.e. the ratio of the kinetic energy over gravitational energy. To do this, we first need to derive the mass profiles of each HFS. This is done by first obtaining the N_2H^+(J=1–0) integrated intensity within the same contours as those used in Figure 5.8 and then convert them into mass by using the $X_{N_2H^+}$ factors derived in Section 5.1. Note that the total masses derived this way differ from the values obtained in Paper I by a factor of up to 3 in the case of SDC345 (see Table 5.1). These differences comes from the different methods that have been used to derive masses. In Paper I clump masses correspond to the bijective masses (Rosolowsky et al. 2008) enclosed within a column density contour of $3 \times 10^{22}\,\mathrm{cm}^{-2}$ in the Peretto et al. (2016) column density maps whose larges-scale structures had been removed with a $10'$ median filter. This contour was chosen as a compromise so that we had reasonable boundaries for all 35 clumps studied within that paper, however in reality a lot of the cloud structures seen in *Spitzer* 8 μm extinction and N_2H^+ emission are not contained within said contours, hence likely underestimating their masses. In the present study, masses are also bijective masses but estimated through the means of a non-filtered column density map and and the removal of a local HFS background $N_{H_2,0}$. i.e. the column density at which we do not detect N_2H^+ emission towards that clump (see Section 5.1). These differences, combined with the imperfect correlation between H_2 column density and N_2H^+(J=1–0) integrated intensity, most likely account for most of the differences.

Figure 5.13 presents the mass profiles of the six HFSs. We first notice that the profile shapes are very similar, with a slight deviation in the case of SDC345. This is

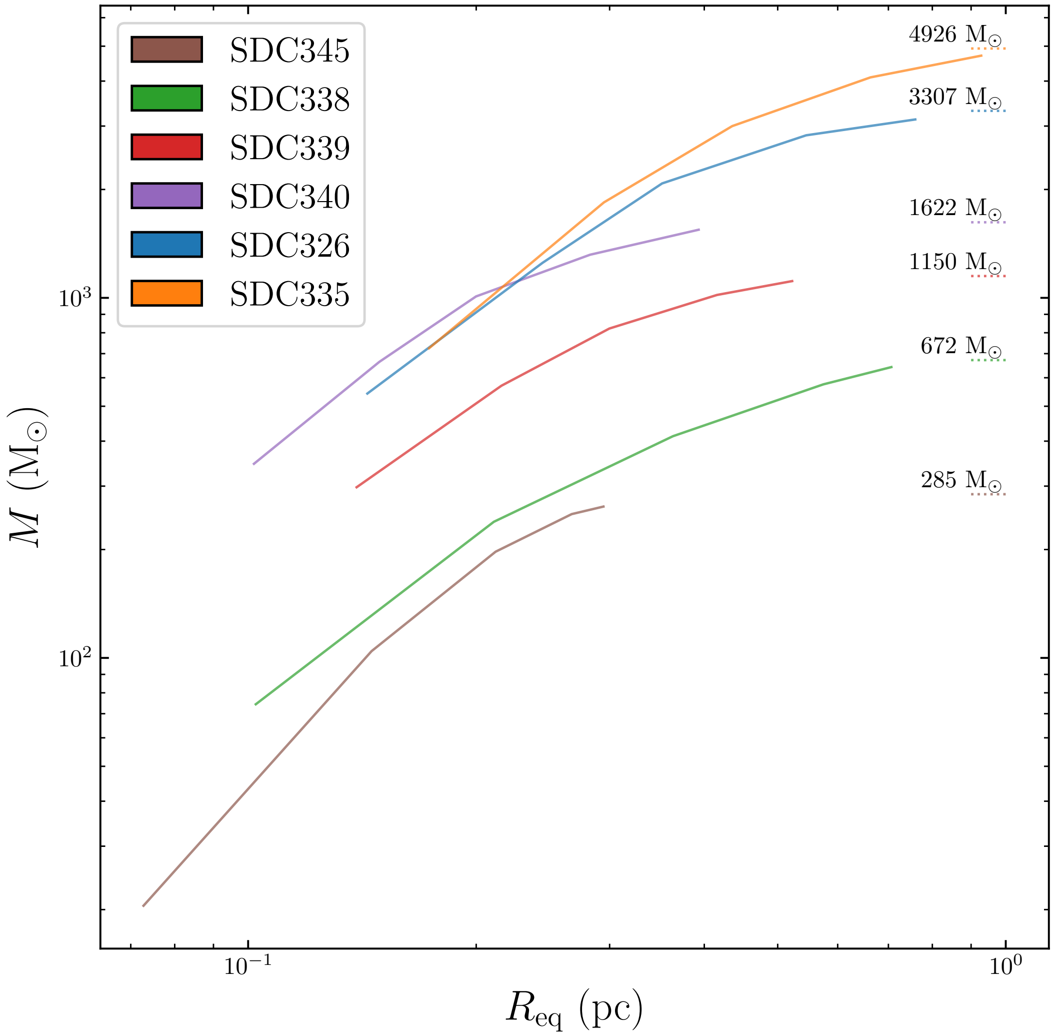

Figure 5.13: Enclosed mass within each contour level (see Figure 5.8), as a function of the equivalent radius ($R_{\rm eq}$) of that contour level. The dotted line represents the total mass, including the points which lie outside of the outermost contour, and hence have an undefined $R_{\rm eq}$.

is likely the consequence of the mismatch between the *Herschel* column density peak used to define the contours and the ALMA integrated intensity within the smallest contour (see Figure 5.8). We also notice that the shape of the profiles, while on a log-log plot, are not straight lines. For spherical clouds, it has been shown that even for power-law mass profiles of the form $M \propto R_{\rm eq}^{\alpha}$, curved profiles are expected as a result of projection effects (Peretto et al. 2023). However, for face-on sheets, those projection effects are minimised and the change in the slope of the mass profiles directly links to a change of the mass surface density profile slope. The profiles here are broadly consistent with $M \propto R_{\rm eq}$ in the filament region and $M \propto R_{\rm eq}^2$ in the hub region, which, in terms of radial dependence of the mass surface density Σ translates into $\Sigma \propto R_{\rm eq}^{-1}$ and $\Sigma \propto R_{\rm eq}^{0}$ (i.e. constant), respectively. However, we have to stress that these N_2H^+-based mass profiles are very uncertain, in particular within the hub region where systematics (i.e. line optical depth, excitation temperature variations, abundance variations) could affect the shape of those profiles. For instance, a constant mass surface density profile within the central $\sim 0.8\,\mathrm{pc}$ region of SDC335, as suggested by Figure 5.13, should be spatially resolved in the *Herschel* column density map as a central flat plateau, something that we do not observe (see column density contours in Figure 5.8).

Combining the estimated mass profiles with the velocity dispersion profiles presented in Figure 5.11, we can derive their virial ratio profiles. The virial ratio $\alpha_{\rm vir}$ is often used as a tool to discuss whether a cloud is bound or unbound, but here we merely use it as a measure of the balance between kinetic and gravitational energy. The virial ratio is defined as:

$$\alpha_{\rm vir} = \frac{2E_{\rm kin}}{|E_{\rm grav}|} \tag{5.5}$$

where the kinetic energy $E_{\rm kin}$ of an enclosed mass M with 1D velocity dispersion $\sigma_{\rm 1D}$ is defined as:

$$E_{\rm kin} = \frac{3}{2} M \sigma_{\rm 1D}^2 \tag{5.6}$$

and the gravitational potential energy $E_{\rm grav}$ of an enclosed mass M with radius R is given by:

$$E_{\rm grav} = -\frac{(3-k_\rho)}{(5-2k_\rho)} \frac{GM^2}{R} \tag{5.7}$$

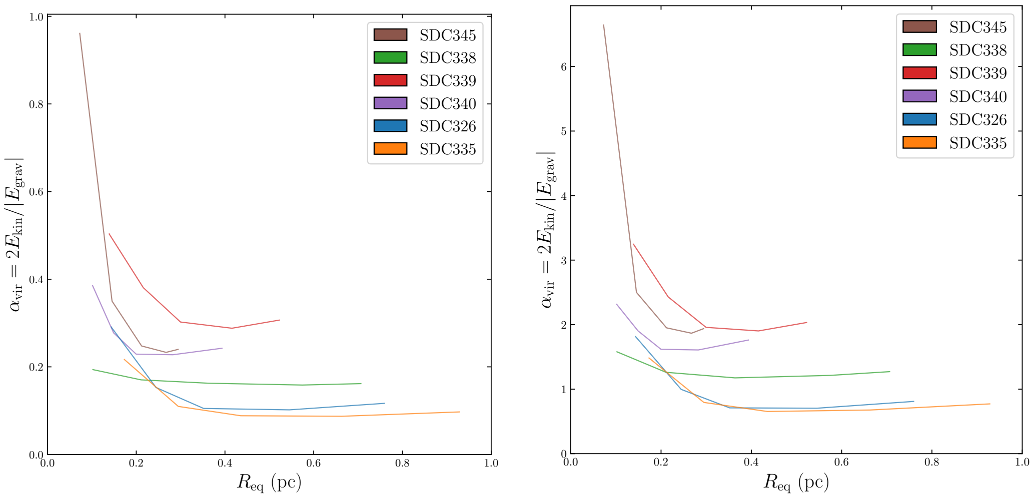

Figure 5.14: Virial parameter (of each contour level), as a function of the equivalent radius ($R_{\rm eq}$) of the outer boundary of that contour level. The left panel shows the virial parameter calculated using the velocity dispersion σ, with the right panel using $\sigma_{\rm total}$.

where G is the gravitational constant, and k_ρ is the power law index of the density profile of the clump.

For an oblate clump, the total (i.e. estimated at the outermost radius) gravitational energy is very similar to the spherical case (Bertoldi & McKee 1992). However, the profiles of the gravitational potential of oblate clumps are different to those of spherical clumps (e.g. Lee & Hennebelle 2016b), and so the mass located outside a given radius $R_{\rm eq}$ does impact the calculation of the gravitational potential energy. However, some recent numerical work have shown that the interaction between the mass within the clump and the gravitational potential external to it does not contribute much to its total gravitational energy (Ganguly et al. 2022), even though it might artificially lead to systematics (i.e. higher virial ratios) towards cores (Ballesteros-Paredes et al. 2018).

Figure 5.14 shows the virial ratio profiles obtained with $k_\rho = 2$ (Williams et al. 2000; Peretto et al. 2023), for both velocity dispersion profiles (i.e. $\sigma(R_{\rm eq})$ and $\sigma_{\rm total}(R_{\rm eq})$). While the exact values of the virial ratios are uncertain (see above), all HFSs show a similar behaviour whereby the virial ratio increases from values typical of gravitationally bound structures (i.e. $\alpha_{\text{vir}} \lesssim 2$) within the filament region to values typical unbound structures (i.e. $\alpha_{\text{vir}} \gtrsim 2$) towards the hub region. The central increase of the virial ratio is the direct consequence of the large velocity dispersions observed towards the hubs. Though, it is important to note that active, sometimes massive, star-forming cores are located in all of the hubs, which makes it very unlikely that the gas lying within the central hubs is unbound. In fact, Ballesteros-Paredes et al. (2018) argue that higher virial ratios towards cores can be attributed to ignoring the surrounding mass in the gravitational energy calculation. Combined with the likely under-estimation of the hub masses, unbound virial ratio values in the densest regions of HFSs are to be expected.

5.4.4 Determining factors for the mass of the most massive cores

In Paper I we showed that, on average, the fraction of mass in a hub-filament system that locked up its most massive core f_{MMC} is larger for infrared dark HFSs compared to infrared-bright clumps. These fractions could reach values as high as $\sim 24\%$ for SDC335. Here, we want to understand what physical characteristics of HFSs might set those high values f_{MMC}, as they are central to our understanding of high-mass star formation.

First, as mentioned above, the masses of the HFSs have been revised. We have therefore used those new masses to calculate new f_{MMC} estimates (see Table 5.1). The changes are most significant for the two lower-mass HFSs, SDC345 and SDC338. In Figure 5.15, we show a scatter plots of f_{MMC} versus clump mass and virial ratio. One can see that, while there is a correlation with clump mass, there is a tentative anti-correlation with α_{vir}.

Granted, this is a very small sample, but this could be a suggestion that the global dynamical state of the clump is important in determining the mass of the most massive cores within clumps. This is reminiscent of the findings in Bonnell et al. (2011), who showed that only clumps with low virial ratios will develop to form high-mass stars (and a complete stellar Initial Mass Function). It is interesting to

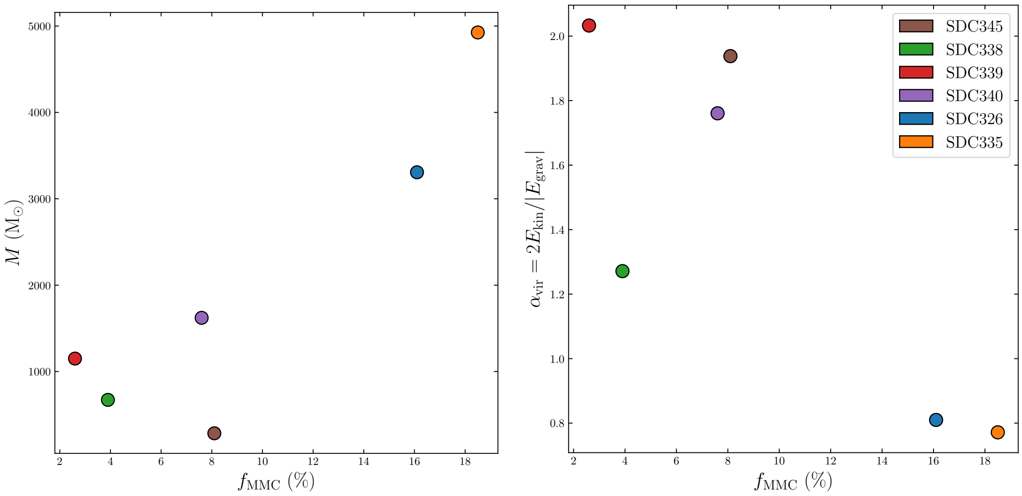

Figure 5.15: Clump mass and virial parameter at the outermost clump contour against the $f_{\rm MMC}$ value for each hub.

note that the clump with the highest global $\alpha_{\rm vir}$ value, SDC339, has both a large number of cores (19), yet no core with a mass greater than $M_{\rm core} \sim 30\,{\rm M}_\odot$, and that the most massive core in the sample, with $M_{\rm core} \sim 900\,{\rm M}_\odot$, belongs to the hub with the lowest global virial ratio (SDC335).

The properties of the initial compression of the gas into a sheet is probably responsible for defining the sheet's main characteristics, including how bound it is. For instance, in the case of cloud-cloud collision, low-velocity shocks lead to a set of more convergent filaments and more massive cores (e.g. Balfour et al. 2017). This seems to be consistent with our observations of infrared dark hub-filament systems.

Chapter 6

Summary

6.1 Conclusions

In this thesis, I presented new ALMA observations of a sample of six infrared dark hub-filament systems. In Chapter 4 I presented an analysis of the core population, and showed that a higher fraction of a clump's mass ends up within the most massive core (MMC) of IR-dark hubs compared to a sample of IR-bright clumps. In Chapter 5, I analysed the dense gas kinematics of the HFS through the fitting of the N_2H^+(J=1–0) emission line using the new fully-automated, multiple velocity component, hyperfine line-fitting code `mwydyn`.

This analysis shows that the hub and filament regions of hub-filaments systems are dynamically distinct, i.e. filaments are made of quiescent gas with low-velocity dispersion, while hubs present extremely complex kinematics alongside very large velocity dispersions. Also, the distribution of velocity centroids across this sample HFSs are extremely similar between clumps, despite spanning a factor of ~ 20 in mass, suggesting that hub-filament systems are compressed layers seen largely face-on. In some HFSs it seems that the velocity profile across the filament-hub transition is relatively sharp, a feature that might be consistent with accretion shock at the outer boundary of the hub. It may also be the result of complex gas flows within the hub centre that are largely unresolved in our observations. What is clear is the need for synthetic observations of N_2H^+(J=1–0) to further investigate the complex gas dynamics within the centres of HFSs. Finally, there is a tentative anti-correlation between f_{MMC} and the global α value for the six HFSs, suggesting that the HFS global virial ratio might set what fraction of its mass will end in the most massive core.

It appears that the physical origin of the initial compression of the gas into a layer is likely to be important in determining the evolution and star formation ability of hub-filament systems. It is not completely clear, though, what this origin might be, whether it is external (e.g. cloud-cloud collision) or internal (e.g. gravity). This will be an important future avenue to explore using observations of a much larger sample of objects across a broad range of evolutionary states, or through the analysis HFSs and non-HFSs formed in simulations.

Another question is to what degree (if at all) these hubs are dynamically decoupled from their surrounding molecular cloud. If indeed clump dynamical decoupling is universal, then the implication is that clouds hosting HFSs have, at least, three dynamically distinct regions: i. A semi-stable/slowly collapsing molecular cloud on large scales; ii. A globally collapsing clump on the scale of a few parsecs; iii. A relatively "static" but growing central hub on a few tenths of parsecs. Only large clump statistics across mass, evolution, and morphology can settle those questions.

6.2 Future work

The data upon which the research presented in this thesis is based is rich with information that is yet to be exploited in full. One obvious avenue to explore is to perform a deep analysis on the filamentary structures that we see. I believe it is necessary to extract the filamentary structures that we see to work towards quantifying the strength of the velocity gradients present, and also to investigate the kinematic properties of the filaments and how they change when they merge.

6.2.1 Tracing the kinematics along the filaments to the hub

Although not completed in time to present in full within this thesis, I have laid the foundations for this work, and have tested some methods for extracting the filamentary structures we see in PPV, which will appear in a forthcoming paper.

Here I will outline the two methods tested so far, both of which are based around calculating the Hessian matrix for an input image or datacube. The Hessian matrix has been used numerous times in astronomy literature[1] for identifying filamentary structures in 2D (position-position) images (e.g. Schisano et al. 2014; Orkisz et al. 2019), spiral arm features in position-veloctiy maps (Durán-Camacho et al. 2024), and as mentioned earlier, was a method employed in Paper I to help classify clumps into hubs and non-hubs.

The Hessian matrix is not restricted to 2D fields, and the general expression for n dimensions is given by:

$$\mathbf{H}_f = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix} \tag{6.1}$$

where f is some function that maps an input vector $\mathbf{x}$ (with n dimensions) to a scalar array. The Hessian of this array, is a square matrix with rows and columns indexed by i and j, respectively. Each element of $\mathbf{H}_f$ can be expressed using the shorthand f_{ij}, i.e.

$$(\mathbf{H}_f)_{i,j} = \frac{\partial^2 f}{\partial x_i \partial x_j} = f_{ij} \tag{6.2}$$

One can then diagonalise the Hessian matrix to obtain its eigenvalue matrix, Λ, which is a matrix containing n eigenvalues λ_n.

The eigenvalues of the Hessian matrix describe the magnitude of the curvature along the direction of their respective eigenvector. For a 2D image, if both eigenvalues λ_1 and λ_2 are positive, this represents a point in the input array which is concave (a bowl). If both λ_1 and λ_2 are negative, this point is convex (i.e. a peak). If one of the eigenvalues are positive and the other is negative, this corresponds to a saddle point (or ridge). A filamentary structure in a 2D image is similar to a ridge-like structure, where along the axis perpendicular to the filament (i.e. the shortest axis) it is strictly concave, and along the filament it is less curved. We can define the criteria for a filament in 2D as $\lambda_2 < \lambda_1$ and $\lambda_2 < 0$. If we search for regions in the

---

[1] Also in many other fields involving image processing.

map where these criteria are met, and set some minimum curvature threshold, we can obtain a boolean mask in which a filamentary structure exists[2].

One can apply this method of finding filaments in a PPV datacube by taking each velocity channel, and one-by-one constructing a mask for each velocity plane, and then merging each of these 2D masks into one 3D mask of the PPV filaments. An example of one plane of these masks is shown in the lower left panel of Figure 6.1. These masks can then be used to obtain the fitted kinematic properties, intensities, and by extension column densities along these filaments.

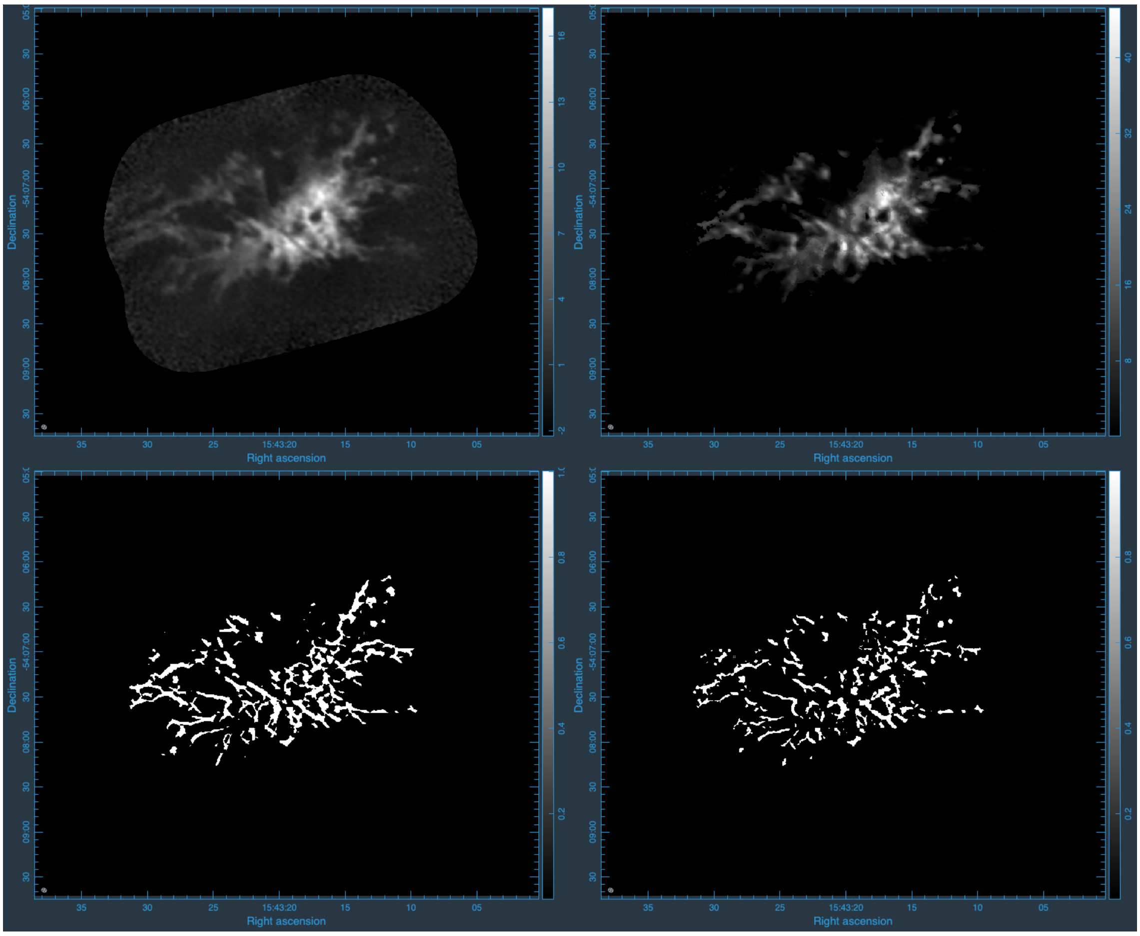

Figure 6.1: *(Top left)* One plane of the N_2H^+(J=1–0) cube for SDC326. *(Top right)* The same plane of a “Gaussian” model cube generated by `mwydyn`. *(Bottom left)* Filament mask obtained by per-plane construction. *(Bottom right)* Filament mask (with $\mathcal{S}_{\rm fil} = 0.02$) obtained using a method adapted from the NEXUS algorithm.

---

[2] Some fine-tuning is required with real images which contain noise to get reasonable masks, but measuring the typical curvature of the noise field is a good starting point.

The method described above treats each velocity plane independently, and so doesn't directly take into account the correlations between velocity channels. Is important information being lost using this method? Is it possible to extract filaments directly from a 3D cube? As the Hessian can easily be generated for a 3D input array (i.e. a PPV datacube), it is possible to measure the curvature of the structures seen in emission as in the 2D case. However, extending the criteria above to PPV is challenging, especially as the velocity axis does not vary over similar scales as the position axes, and attempts to do so yielded "unnatural" looking filamentary structure with many artefacts.

Cautun et al. (2013) developed an algorithm called NEXUS to identify and analyse structures in the cosmic web, such as "clusters", "filaments", "walls" and "voids". Although primarily developed for analysis of PPP data (from cosmological simulations), it can be applied to any 3D field. This algorithm also utilises the eigenvalues of the Hessian matrix to characterise the morphology of structures in input data. In an attempt to improve on my previous attempts of PPV filament extraction, I adopted a much simplified version of their algorithm.

First, the eigenvalues λ_1, λ_2 and λ_3 need to be obtained from the Hessian of the input datacube. The eigenvalues are sorted, such that $\lambda_1 < \lambda_2 < \lambda_3$. In Cautun et al. (2013), they define a so-called "signature function" $\mathcal{S}$, which is used to identify the morphological classes ("clusters", "filaments" and "walls") in their input density fields. This function maps a 3D field into another 3D field with varying values of $\mathcal{S}$, where a high value of indicates a strong signature of some morphological signature. We are primarily interested in the filament component of this function, which is expressed as

$$\mathcal{S}_{\text{fil}} = \mathcal{I}_{\text{fil}} \times |\lambda_2| \theta(-\lambda_1) \theta(-\lambda_2) \tag{6.3}$$

where $\mathcal{I}_{\text{fil}}$ is the "shape strength", given by

$$\mathcal{I}_{\text{fil}} = \left| \frac{\lambda_2}{\lambda_1} \right| \Theta \left(1 - \left| \frac{\lambda_3}{\lambda_1} \right| \right) \tag{6.4}$$

where $\Theta(x) = x\theta(x)$, and

$$\theta(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases} \tag{6.5}$$

Using these definitions, along with the constraints $|\lambda_1| \simeq |\lambda_2| \gg |\lambda_3|$ and $\lambda_1 < 0; \lambda_2 < 0$, a filamentary structure is defined as being anywhere $\mathcal{S}_{\rm fil} > 0$. We can therefore create a PPV filament mask by choosing some value of $\mathcal{S}_{\rm fil}$, an example of which can be seen in the lower right panel of Figure 6.1. Some 3D visualisations of the isocontour $\mathcal{S}_{\rm fil} = 0.02$ for SDC326 are shown in Figure 6.2.

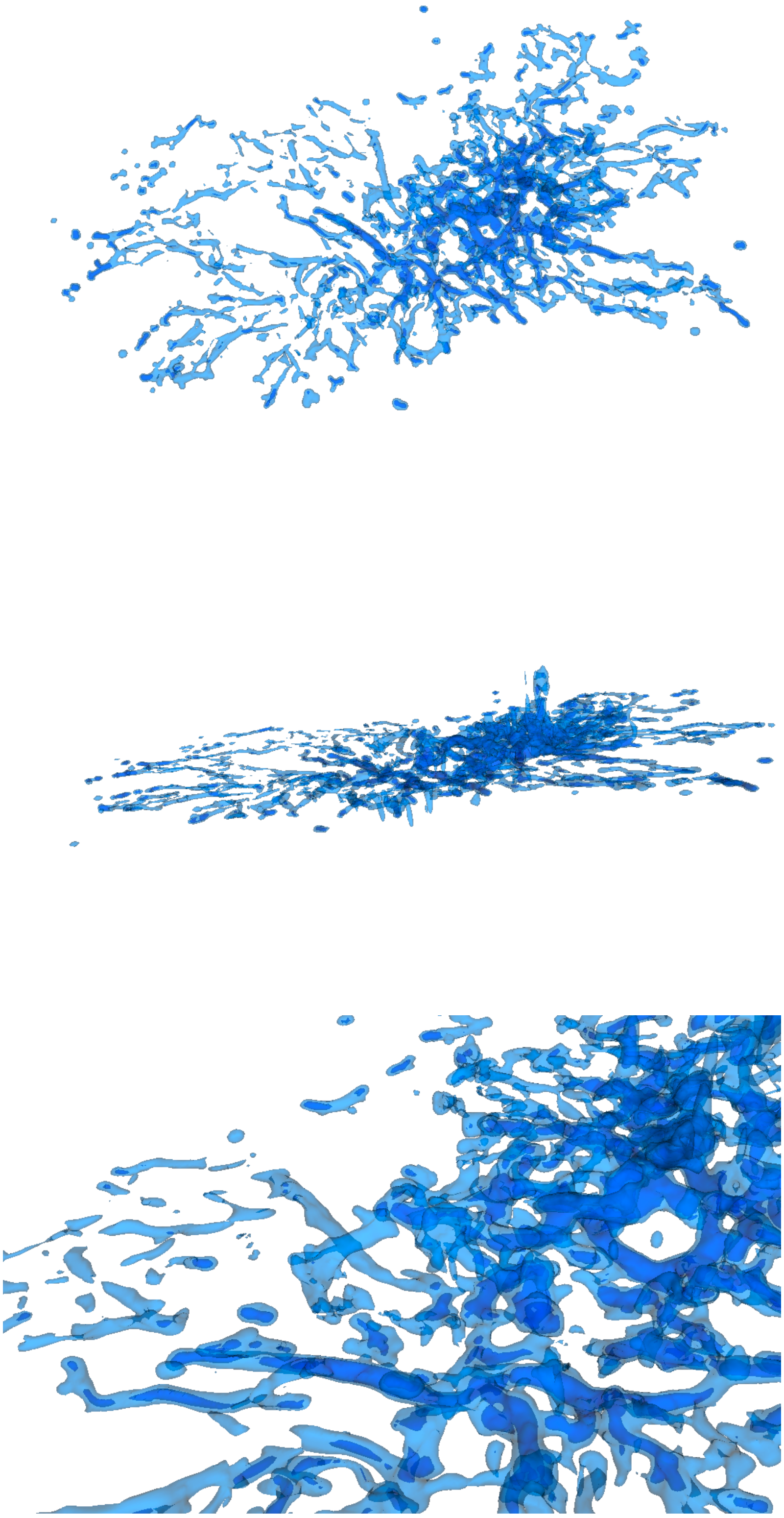

Figure 6.2: 3D visualisations at different viewing angles of the filamentary structures found in SDC326 using equation 6.3, and $\mathcal{S}_{\mathrm{fil}} = 0.02$.

Subjectively, this seems to match very well the filamentary structures seen in the original data. However, at present it is unclear which of the two methods meaningfully obtains the “best” filamentary structures, or exactly how to interpret filaments found in PPV-space. This will require further investigation, for instance by generating synthetic filaments and using some distance measure to see how well the extracted filaments match the original.

Once a suitable method has been devised and tested for extracting these filamentary structures, it will make it much easier to investigate any correlations between the core properties found in Paper I, and the kinematic information obtained using `mwydyn`.

6.2.2 Constraining the morphology of HFSs

Some of the suggestions in Chapter 5 could also be tested. If there are indeed shocks present where the filaments meet the central hub, then we should not be able to see a continuation of the incoming filamentary structure beyond this front, or see any significant filamentary structure within the hub. It may be possible to check whether the velocity gradients (in combination with estimated column density values) are driven by the acceleration caused gravitational potential of the central hub, and also calculate mass accretion rates (albeit with a large degree of uncertainty). If these hubs are largely sheet-like structures seen largely face-on, it might also be possible to put some contraints on the inclination angle using a combination of the measured gravitational potential and the measured gradients along the filaments.

As mentioned in Chapter 5, it is possible to investigate the 3D structure of molecular clouds using by using multiple transitions of the same molecule (e.g. Li & Goldsmith 2012; Bonne et al. 2020). In Bonne et al. (2020), the J=2–1 and J=3–2 transitions of ^{13}CO were observed observed with APEX. Using the ratio of these two transitions, radiative transfer was performed using RADEX to construct a density profile for the Musca filament, in combination with temperature maps derived from *Herschel* observations. Using *Herschel* column density maps and estimated density profiles, they showed that Musca typically had a line-of-sight thickness of $\sim 0.2\,\mathrm{pc}$ to $0.5\,\mathrm{pc}$. In principle it would be possible to perform this technique for the sample of hubs presented in this thesis, if observations of a higher transition of N_2H^+ (perhaps J=2–1 or J=3–2) were obtained with a more compact 12 m configuration and/or the ACA. This could help determine, if these hubs are sheets oriented on the plane of the sky, how thick they are. Although, given the higher morphological and kinematic complexity of these hubs compared to Musca, this might be a challenging task.

6.2.3 Overcoming small number statistics

For any significant conclusions to be drawn regarding the nature of hub-filament systems, and how they compare to other clumps, we need a much larger set of observations of an unbiased sample of clumps. Some progress is being made on this front already, with an ALMA Cycle 9 programme (for which I am a co-investigator) that is in the process of being observed ($\sim$ 70% complete). The work presented in this thesis has effectively acted as a proof of concept for some of the avenues of investigation in the proposal.

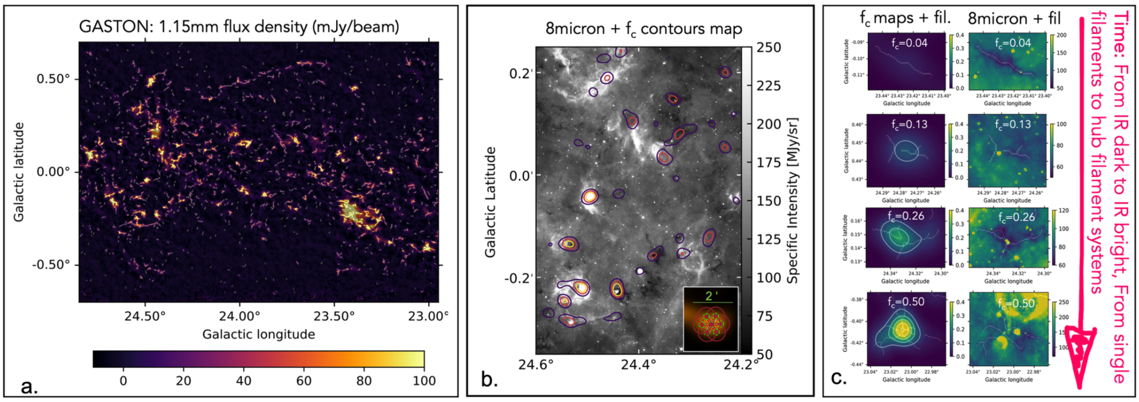

Figure 6.3: *(a):* GASTON 1.15 mm image, with identified filamentary structures overlain. *(b):* Sub-region of the left panel in *Spitzer* 8 µm, with contours representing values of f_c starting at $f_c = 0.1$, in 0.1 increments. *(c):* A selection of clumps showing both the f_c and *Spitzer* 8 µm maps, illustrating the correlation of f_c with $f_{\rm IRB}$.

Results from Rigby et al. (2021) suggest that the IR-brightness fraction $f_{\rm IRB}$, is a tracer of clump evolution, with clumps gathering mass from their environment over time. Using the same set of observations, Peretto et al. (2022) devised a new technique that systematically quantifies how convergent a set of filaments are. Using this filament convergence parameter, f_c, they showed that although HFSs represent only a small fraction of filamentary structures, they host more massive and luminous compact sources than non-hubs. They also showed that hubs appear to be more evolved than non-hubs. taken together, the results from both of these studies suggest an evolutionary sequence whereby filamentary structures evolve from IR-dark, individual filaments, toward IR-bright HFSs, and that during this process compact sources within them grown in mass.

This proposal aims to unveil how filamentary structures in the ISM evolve towards the formation of HFSs by surveying 136 clumps (mosaicked) with ALMA in N_2H^+(J=1–0) and 3 mm dust continuum, with a very similar observational setup to the six fields presented here. By observing a more diverse, much larger sample of clumps it will be possible to test the hypothesis that clumps gather mass over time, and in doing so increase their morphological complexity using f_{IRB} as an evolutionary tracer, and f_c as a measure of complexity. The kinematics of more than 100 clumps and ~ 1000 filaments will be able to be investigated by applying `mwydyn` to the N_2H^+(J=1–0) observations, hopefully giving us more clues as to how HFSs are formed, and whether the results seen for the sample in this thesis are representative.

6.2.4 Testing observed kinematics using synthetic observations

A significant source of uncertainty for observational astronomers is how well the properties we derive from our observations match the true physical properties that belong to those observed objects. While synthetic observations of simulations are becoming more common, there is still a lot more work to be done in this area regarding molecular line emission and kinematics. Clarke et al. (2018) performed simulations of the growth of a filament in a turbulent medium, and then generated synthetic $C^{18}O$ observations of those same filaments to investigate whether structures seen in PPV (obtained through fitting the observations with BTS, a multiple

Gaussian line-fitting code) were coherent with features in PPP. They found that filamentary structures identified in PPV do not correspond to filamentary structures that appear in PPP, and hence cautioning that care should be taken when using velocity-coherent features to constrain substructure within filaments.

An important test would to be investigate how closely the recovered velocity field from fitting molecular line emission is to velocity field of the observed object. For example, one could perform AREPO simulations of molecular clouds (e.g. Clark et al. 2019; Smith et al. 2020), and use a chemical network, e.g. UCLCHEM (Holdship et al. 2017), to obtain the abundances of molecular species (e.g. N_2H^+). One could then perform radiative transfer modelling, using LIME (Brinch & Hogerheijde 2010), of the full hyperfine structure of N_2H^+(J=1–0), to obtain a brightness distribution of the simulation at multiple viewing angles. This brightness distribution could then be fed through a pipeline that generates synthetic observations, for example the ALMA simulator, to generate data that can then be treated as observations of molecular clouds. The synthetically observed spectra could be fit using `mwydyn`, yielding some kinematic properties that could then be compared to the actual velocity fields in the input simulations. Although this would be a computationally expensive, laborious and time-consuming process, it may well be able to help us undertand what causes some of the complex line shapes, and challenge our interpretations of what we observe in the ISM.

Appendix A

Spitzer 8 µm Clump Cutouts

Spitzer 8 µm images of each clump in our sample. The images are in Galactic longitude and latitude (l, b), with the axes marked in the lower right. The red contours represent the "boundary" of our clumps defined by the $N_{\mathrm{H}_2} = 3 \times 10^{22}\,\mathrm{cm}^{-2}$ level in our *Herschel* column density maps. The orange diamond marks the location of the most-massive core (MMC) in each clump, and the grey contours show the ALMA field of view coverage for each object. The filamentary structures identified in the *Herschel* 250 µm maps are overlaid in white. The symbol to the upper left represents how we classified each clump, along with the clump ID number. Crosses represent clumps that have been classified as HFS, and circular points are non-HFS clumps. The point fill colours represent the three IR-brightness classes. Points with a black outline are sources observed at 2.9 mm, and points without outlines are sources observed at 878 µm.

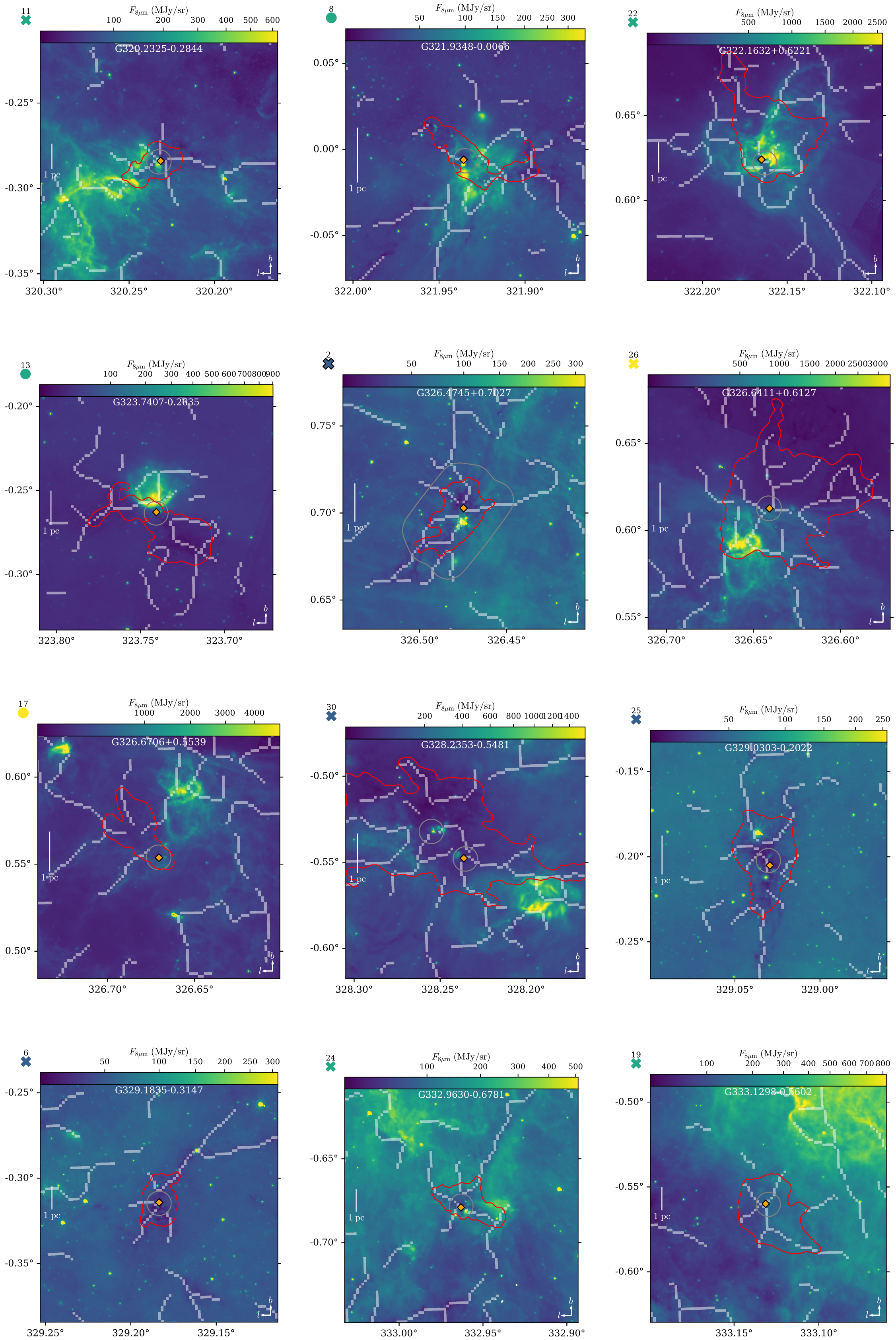

Figure A.1: *Spitzer* 8 µm images of the clump in the extended sample, showing our cloud boundaries and classifications.

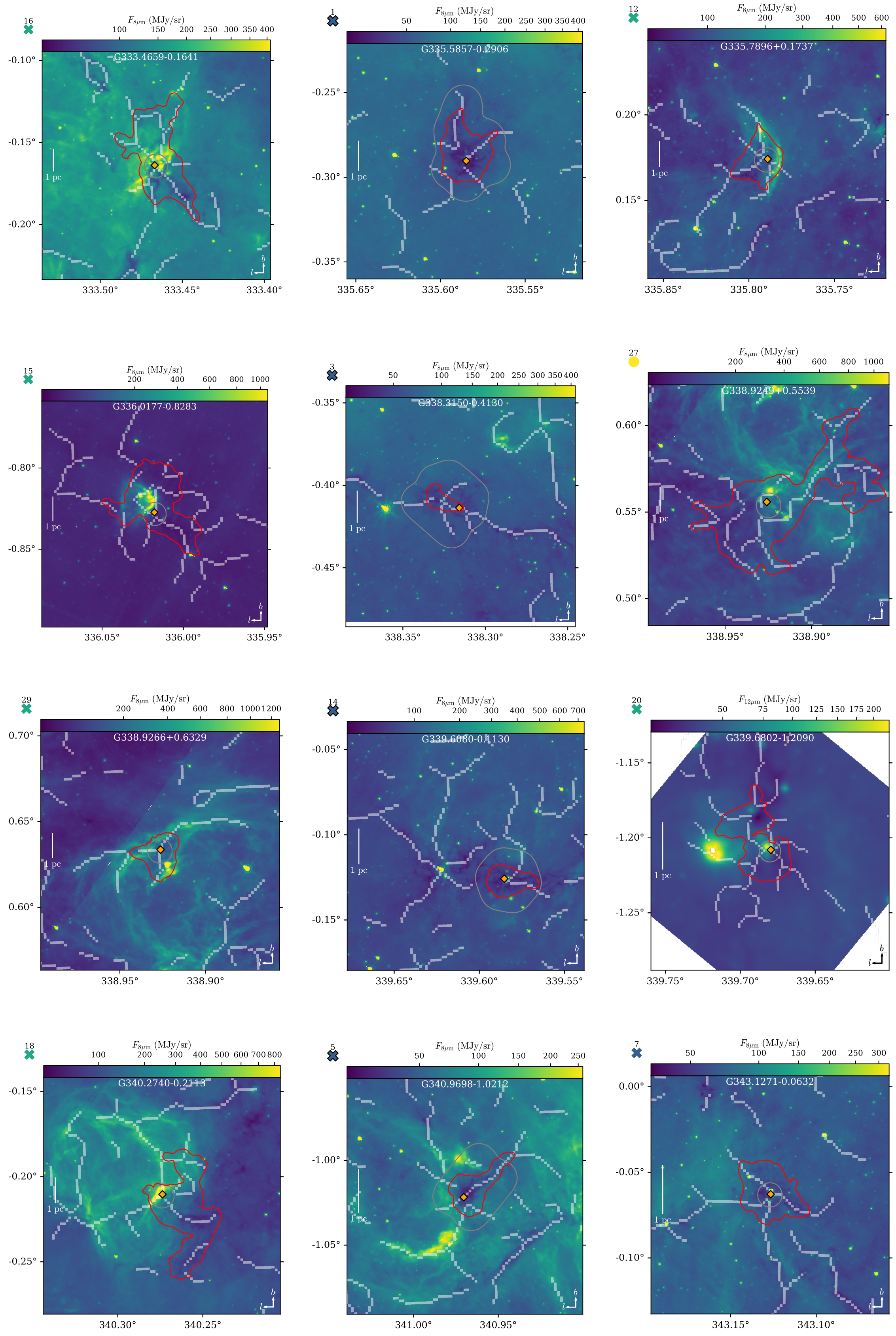

Figure A.1: *(continued)*

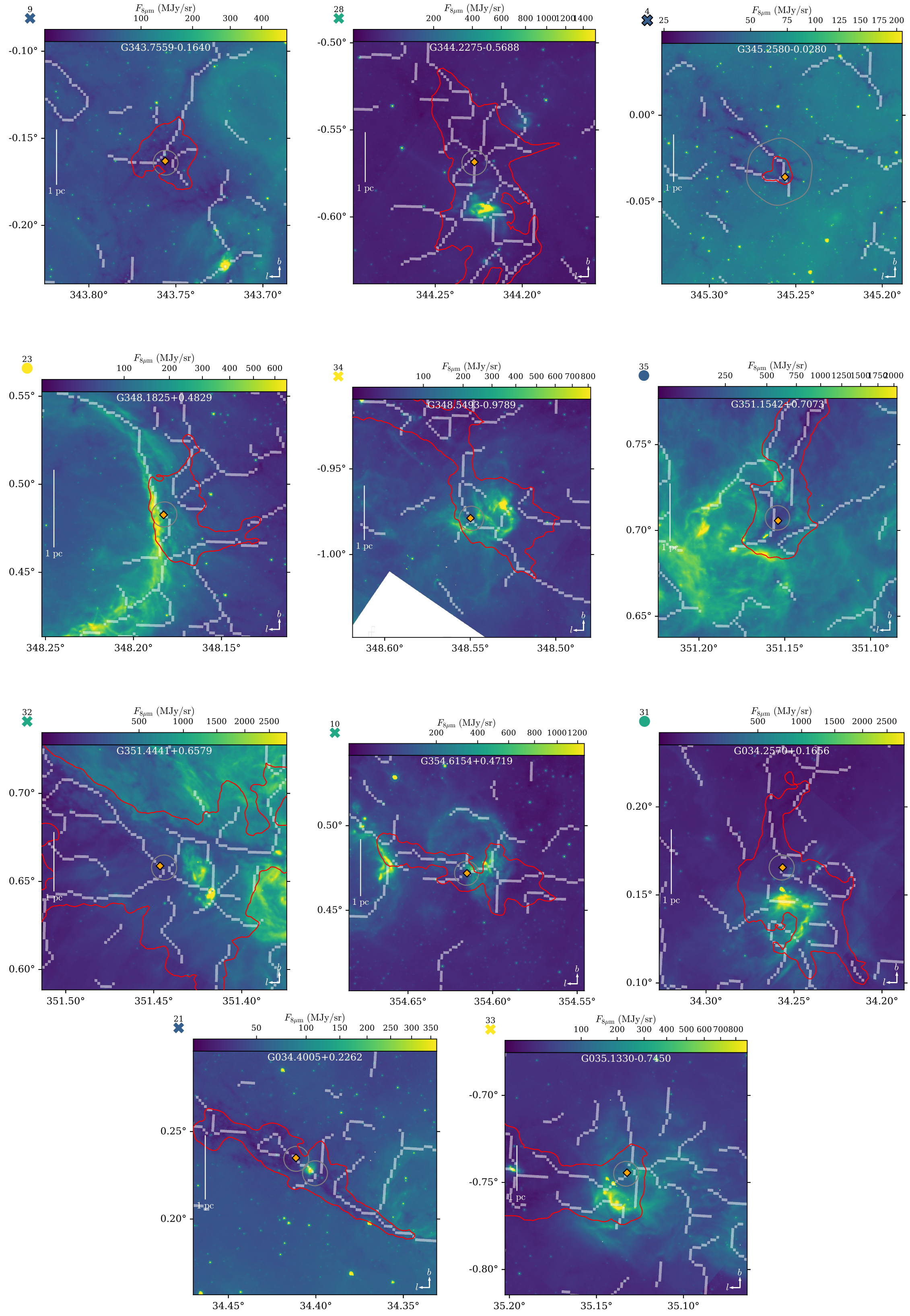

Figure A.1: *(continued)*

Appendix B

Software used

This research made use of the Python packages Astropy[1] (Astropy Collaboration et al. 2013, 2018, 2022), astrodendro[2], IPython[3] (Pérez & Granger 2007), LMFIT[4] (Newville et al. 2014, 2022), Matplotlib[5] (Hunter 2007), NumPy[6] (Harris et al. 2020), SciPy[7] (Virtanen et al. 2020), and scikit-image[8] (van der Walt et al. 2014).

This research also made use of NASA's Astrophysics Data System Bibliographic Services, CARTA[9] (Comrie et al. 2021), CASA[10] (The Casa Team et al. 2022), TOPCAT[11] (Taylor 2005), and SAOImageDS9[12] (Joye & Mandel 2003).

This thesis was typeset using LaTeX[13][14].

---

[1] `https://astropy.org`

[2] `http://dendrograms.org`

[3] `https://ipython.org`

[4] `https://lmfit.github.io/lmfit-py`

[5] `https://matplotlib.org`

[6] `https://numpy.org`

[7] `https://scipy.org`

[8] `https://scikit-image.org`

[9] `https://cartavis.org`

[10] `https://casa.nrao.edu`

[11] `http://www.star.bris.ac.uk/~mbt/topcat`

[12] `http://ds9.si.edu`

[13] `https://www.latex-project.org`

[14] `https://tug.org`

Bibliography

Anderson M., et al., 2021, MNRAS, 508, 2964

Andre P., Ward-Thompson D., Barsony M., 1993, ApJ, 406, 122

Andre P., Ward-Thompson D., Barsony M., 2000, in Mannings V., Boss A. P., Russell S. S., eds, Protostars & Planets IV. p. 59, doi:10.48550/arXiv.astro-ph/9903284

André Ph., Belloche A., Motte F., Peretto N., 2007, A&A, 472, 519

André Ph., et al., 2010, A&A, 518, L102

André P., Di Francesco J., Ward-Thompson D., Inutsuka S.-I., Pudritz R. E., Pineda J., 2014, in , Protostars & Planets VI. University of Arizona Press, Tucson, AZ, doi:10.2458/azu_uapress_9780816531240-ch002

André Ph., et al., 2016, A&A, 592, A54

André Ph., Arzoumanian D., Könyves V., Shimajiri Y., Palmeirim P., 2019, A&A, 629, L4

Arzoumanian D., et al., 2011, A&A, 529, L6

Arzoumanian D., et al., 2021, A&A, 647, A78

Astropy Collaboration et al., 2013, A&A, 558, A33

Astropy Collaboration et al., 2018, AJ, 156, 123

Astropy Collaboration et al., 2022, ApJ, 935, 167

Audit E., Hennebelle P., 2005, A&A, 433, 1

Avison A., Peretto N., Fuller G. A., Duarte-Cabral A., Traficante A., Pineda J. E., 2015, A&A, 577, A30

Avison A., et al., 2021, A&A, 645, A142

Balfour S. K., Whitworth A. P., Hubber D. A., Jaffa S. E., 2015, MNRAS, 453, 2471

Balfour S. K., Whitworth A. P., Hubber D. A., 2017, MNRAS, 465, 3483

Ballesteros-Paredes J., Vázquez-Semadeni E., Palau A., Klessen R. S., 2018, MNRAS

Barnes P. J., Muller E., Indermuehle B., O'Dougherty S. N., Lowe V., Cunningham M., Hernandez A. K., Fuller G. A., 2015, ApJ, 812, 6

Barnes A. T., et al., 2020, MNRAS, 497, 1972

Barnes A. T., et al., 2021, MNRAS, 503, 4601

Battersby C., et al., 2011, A&A, 535, A128

Beltrán M. T., Rivilla V. M., Kumar M. S. N., Cesaroni R., Galli D., 2022, A&A, 660, L4

Bergin E. A., Tafalla M., 2007, Annu. Rev. Astron. Astrophys., 45, 339

Bergin E. A., Alves J., Huard T., Lada C. J., 2002, ApJ, 570, L101

Bertoldi F., McKee C. F., 1992, ApJ, 395, 140

Beuther H., et al., 2013, A&A, 553, A115

Bonne L., et al., 2020, A&A, 644, A27

Bonnell I. A., Smith R. J., Clark P. C., Bate M. R., 2011, MNRAS, 410, 2339

Bontemps S., Motte F., Csengeri T., Schneider N., 2010, A&A, 524, A18

Briggs D. S., 1995, PhD thesis, The New Mexico Institue of Mining and Technology, Socorro, New Mexico, `http://www.aoc.nrao.edu/dissertations/dbriggs/`

Briggs D. S., Schwab F. R., Sramek R. A., 1999, in Taylor G. B., Carilli C. L., Perley R. A., eds, Astronomical Society of the Pacific Conference Series Vol. 180, Synthesis Imaging in Radio Astronomy II. p. 127

Brinch C., Hogerheijde M. R., 2010, A&A, 523, A25

Burke B. F., Graham-Smith F., Wilkinson P. N., 2019, An Introduction to Radio Astronomy, 4th Edition, 4th edn. Cambridge University Press

Busquet G., et al., 2013, ApJ, 764, L26

Caselli P., Myers P. C., Thaddeus P., 1995, ApJ, 455

Caselli P., Benson P. J., Myers P. C., Tafalla M., 2002, ApJ, 572, 238

Cautun M., Van De Weygaert R., Jones B. J. T., 2013, MNRAS, 429, 1286

Churchwell E., et al., 2009, Publ. Astron. Soc. Pac., 121, 213

Clark P. C., Glover S. C. O., Ragan S. E., Duarte-Cabral A., 2019, MNRAS, 486, 4622

Clarke S. D., Whitworth A. P., 2015, MNRAS, 449, 1819

Clarke S. D., Whitworth A. P., Spowage R. L., Duarte-Cabral A., Suri S. T., Jaffa S. E., Walch S., Clark P. C., 2018, MNRAS, 479, 1722

Comrie A., et al., 2021, CARTA: The Cube Analysis and Rendering Tool for Astronomy, Zenodo, doi:10.5281/ZENODO.3377984

Conway J. E., Cornwell T. J., Wilkinson P. N., 1990, MNRAS, 246, 490

Cornwell T. J., 2008, MNRAS, 2, 793

Csengeri T., et al., 2017, A&A, 600, L10

Cyganowski C. J., et al., 2014, ApJ, 796, L2

Dame T. M., Hartmann D., Thaddeus P., 2001, ApJ, 547, 792

Dempsey J. T., Thomas H. S., Currie M. J., 2013, ApJS, 209, 8

Dewangan L. K., Ojha D. K., Baug T., 2017, ApJ, 844, 15

Dobbs C. L., Burkert A., Pringle J. E., 2011, MNRAS, 413, 2935

Draine B. T., 2011, Physics of the Interstellar and Intergalactic Medium. Princeton University Press, doi:10.1515/9781400839087

Duarte-Cabral A., Dobbs C. L., 2017, MNRAS, 470, 4261

Duarte-Cabral A., et al., 2020, MNRAS, 500, 3027

Dunham M. M., Crapsi A., Evans II N. J., Bourke T. L., Huard T. L., Myers P. C., Kauffmann J., 2008, ApJS, 179, 249

Durán-Camacho E., et al., 2024, Self-Consistent Modelling of the Milky Way Structure Using Live Potentials (arxiv:2405.09503), `http://arxiv.org/abs/2405.09503`

Eden D. J., Moore T. J. T., Plume R., Morgan L. K., 2012, MNRAS, 422, 3178

Egan M. P., Carey S. J., Price S. D., Shipman R. F., Feldman P., Redman R., 1999, in Cox P., Kessler M., eds, ESA Special Publication Vol. 427, The Universe as Seen by ISO. p. 671

Elia D., et al., 2017, MNRAS, 471, 100

Emerson J. P., 1988, in Dupree A. K., Lago M. T. V. T., eds, , Formation and Evolution of Low Mass Stars. Springer Netherlands, Dordrecht, pp 21–44, doi:10.1007/978-94-009-3037-7_2

Estalella R., 2017, PASP, 129, 025003

Faustino Vieira H., Duarte-Cabral A., Davis T. A., Peretto N., Smith M. W. L., Querejeta M., Colombo D., Anderson M., 2023a, MNRAS, p. stad1876

Faustino Vieira H., Duarte-Cabral A., Davis T. A., Peretto N., Smith M. W. L., Querejeta M., Colombo D., Anderson M., 2023b, Monthly Notices of the Royal Astronomical Society, 527, 3639

Federrath C., 2016, MNRAS, 457, 375

Ferrière K. M., 2001, Rev. Mod. Phys., 73, 1031

Fukui Y., et al., 2019, ApJ, 886, 14

Ganguly S., Walch S., Clarke S. D., Seifried D., 2022, arXiv

Geen S., Soler J. D., Hennebelle P., 2017, MNRAS, 471, 4844

Ginsburg A., Sokolov V., de Val-Borro M., Rosolowsky E., Pineda J. E., Sipőcz B. M., Henshaw J. D., 2022, AJ, 163, 291

Glover S. C. O., Clark P. C., 2012, Monthly Notices of the Royal Astronomical Society, pp no–no

Goldsmith P. F., 2001, ApJ, 557, 736

Goldsmith P. F., Heyer M., Narayanan G., Snell R., Li D., Brunt C., 2008, ApJ, 680, 428

Gómez G. C., Vázquez-Semadeni E., 2014, ApJ, 791, 124

Hacar A., Tafalla M., Alves J., 2017, A&A, 606, A123

Hacar A., Tafalla M., Forbrich J., Alves J., Meingast S., Grossschedl J., Teixeira P. S., 2018, A&A, 610, A77

Harris C. R., et al., 2020, Nature, 585, 357

Hartmann L., Burkert A., 2007, ApJ, 654, 988

Henshaw J. D., et al., 2016, MNRAS, 457, 2675

Henshaw J. D., et al., 2020, Nat Astron, 4, 1064

Heyer M., Dame T., 2015, Annu. Rev. Astron. Astrophys., 53, 583

Heyer M., Krawczyk C., Duval J., Jackson J. M., 2009, ApJ, 699, 1092

Högbom J. A., 1974, A&AS, 15, 417

Holdship J., Viti S., Jiménez-Serra I., Makrymallis A., Priestley F., 2017, AJ, 154, 38

Hollenbach D. J., Werner M. W., Salpeter E. E., 1971, ApJ, 163, 165

Hunter J. D., 2007, Comput. Sci. Eng., 9, 90

Inoue T., Hennebelle P., Fukui Y., Matsumoto T., Iwasaki K., Inutsuka S.-i., 2018, Publ. Astron. Soc. Jpn., 70

Inutsuka S.-i., Inoue T., Iwasaki K., Hosokawa T., 2015, A&A, 580, A49

Jackson J. M., et al., 2006, ApJS, 163, 145

Jackson J. M., Finn S. C., Chambers E. T., Rathborne J. M., Simon R., 2010, ApJ, 719, L185

Jackson J. M., et al., 2019, ApJ, 870, 5

Jeans J. H., 1928, Astronomy and Cosmogony. Cambridge University Press, `https://ui.adsabs.harvard.edu/abs/1928asco.book.....J`

Johnstone D., Fich M., Mitchell G. F., Moriarty-Schieven G., 2001, ApJ, 559, 307

Joye W. A., Mandel E., 2003, in Payne H. E., Jedrzejewski R. I., Hook R. N., eds, Astronomical Society of the Pacific Conference Series Vol. 295, Astronomical Data Analysis Software and Systems XII. p. 489

Kainulainen J., Beuther H., Henning T., Plume R., 2009, A&A, 508, L35

Kauffmann J., Bertoldi F., Bourke T. L., Evans N. J., Lee C. W., 2008, A&A, 487, 993

Kepley A. A., Tsutsumi T., Brogan C. L., Indebetouw R., Yoon I., Mason B., Meyer J. D., 2020, PASP, 132, 024505

Kirk H., Myers P. C., Bourke T. L., Gutermuth R. A., Hedden A., Wilson G. W., 2013, ApJ, 766, 115

Könyves V., et al., 2010, A&A, 518, L106

Könyves V., et al., 2015, A&A, 584, A91

Könyves V., et al., 2020, A&A, 635, A34

Krumholz M. R., McKee C. F., Bland-Hawthorn J., 2019, Annu. Rev. Astron. Astrophys., 57, 227

Kumar M. S. N., Palmeirim P., Arzoumanian D., Inutsuka S. I., 2020, A&A, 642, A87

Lada C. J., Lada E. A., 2003, Annu. Rev. Astron. Astrophys., 41, 57

Larson R. B., 1981, MNRAS, 194, 809

Lee Y.-N., Hennebelle P., 2016a, A&A, 591, A30

Lee Y.-N., Hennebelle P., 2016b, A&A, 591, A31

Lewis J. A., Lada C. J., Dame T. M., 2022, ApJ, 931, 9

Li D., Goldsmith P. F., 2012, ApJ, 756, 12

Li Z.-Y., Nakamura F., 2006, ApJ, 640, L187

Lin Y., Csengeri T., Wyrowski F., Urquhart J. S., Schuller F., Weiss A., Menten K. M., 2019, A&A, 631, A72

Liu H. B., Jiménez-Serra I., Ho P. T. P., Chen H.-R., Zhang Q., Li Z.-Y., 2012, ApJ, 756, 10

Liu T., et al., 2020, MNRAS, 496, 2790

Liu H.-L., et al., 2023, MNRAS, 522, 3719

Longmore S. N., et al., 2013, MNRAS, 429, 987

Marsh K. A., Whitworth A. P., Lomax O., 2015, MNRAS, 454, 4282

Marsh K. A., et al., 2017, MNRAS, 471, 2730

McKee C. F., Ostriker J. P., 1977, ApJ, 218, 148

McMullin J. P., Waters B., Schiebel D., Young W., Golap K., 2007, in Shaw R. A., Hill F., Bell D. J., eds, Astronomical Society of the Pacific Conference Series Vol. 376, Astronomical Data Analysis Software and Systems XVI. p. 127

Miville-Deschênes M.-A., Murray N., Lee E. J., 2017, ApJ, 834, 57

Molinari S., et al., 2010, A&A, 518, L100

Molinari S., et al., 2016, A&A, 591, A149

Motte F., Andre P., Neri R., 1998, A&A, 336, 150

Motte F., Bontemps S., Schilke P., Schneider N., Menten K. M., Broguière D., 2007, A&A, 476, 1243

Motte F., et al., 2018a, Nat Astron, 2, 478

Motte F., Bontemps S., Louvet F., 2018b, Annu. Rev. Astron. Astrophys., 56, 41

Murray N., Rahman M., 2010, ApJ, 709, 424

Myers P. C., 2009, ApJ, 700, 1609

Nakamura F., Li Z.-Y., 2008, ApJ, 687, 354

Newville M., Stensitzki T., Allen D. B., Ingargiola A., 2014, LMFIT: Non-Linear Least-Square Minimization and Curve-Fitting for Python, Zenodo, doi:10.5281/ZENODO.11813

Newville M., et al., 2022, Lmfit/Lmfit-Py: 1.1.0, Zenodo, doi:10.5281/ZENODO.598352

Nony T., et al., 2018, A&A, 618, L5

Nony T., et al., 2023, A&A, 674, A75

Nutter D., Ward-Thompson D., 2007, MNRAS, 374, 1413

Orkisz J. H., et al., 2019, A&A, 624, A113

Ossenkopf V., Henning T., 1994, A&A, 291, 943

Padoan P., Pan L., Haugbølle T., Nordlund Å., 2016, ApJ, 822, 11

Padoan P., Pan L., Juvela M., Haugbølle T., Nordlund Å., 2020, ApJ, 900, 82

Palau A., et al., 2014, ApJ, 785, 42

Palau A., et al., 2021, ApJ, 912, 159

Panopoulou G. V., Psaradaki I., Skalidis R., Tassis K., Andrews J. J., 2017, MNRAS, 466, 2529

Perault M., et al., 1996, A&A, 315, L165

Peretto N., Fuller G. A., 2009, A&A, 505, 405

Peretto N., Fuller G. A., 2010, ApJ, 723, 555

Peretto N., André Ph., Belloche A., 2006, A&A, 445, 979

Peretto N., et al., 2010, A&A, 518, L98

Peretto N., et al., 2013, A&A, 555, A112

Peretto N., et al., 2014, A&A, 561, A83

Peretto N., Lenfestey C., Fuller G. A., Traficante A., Molinari S., Thompson M. A., Ward-Thompson D., 2016, A&A, 590, A72

Peretto N., et al., 2020, MNRAS, 496, 3482

Peretto N., et al., 2022, EPJ Web Conf., 257, 00037

Peretto N., Rigby A. J., Louvet F., Fuller G. A., Traficante A., Gaudel M., 2023, MNRAS, 525, 2935

Pérez F., Granger B. E., 2007, Comput. Sci. Eng., 9, 21

Peters T., Klessen R. S., Low M.-M. M., Banerjee R., 2010, ApJ, 725, 134

Pety J., 2018, Gildas And Class, Zenodo, doi:10.5281/ZENODO.1205423

Polychroni D., et al., 2013, ApJ, 777, L33

Pouteau Y., et al., 2022, A&A, 664, A26

Pouteau Y., et al., 2023, A&A, 674, A76

Ragan S. E., Bergin E. A., Gutermuth R. A., 2009, ApJ, 698, 324

Ragan S. E., Heitsch F., Bergin E. A., Wilner D., 2012, ApJ, 746, 174

Ragan S. E., Henning Th., Tackenberg J., Beuther H., Johnston K. G., Kainulainen J., Linz H., 2014, A&A, 568, A73

Ragan S. E., Henning T., Beuther H., Linz H., Zahorecz S., 2015, A&A, 573, A119

Ragan S. E., Moore T. J. T., Eden D. J., Hoare M. G., Urquhart J. S., Elia D., Molinari S., 2018, MNRAS, 479, 2361

Rathborne J. M., et al., 2014, ApJ, 795, L25

Reid M. J., et al., 2009, ApJ, 700, 137

Reid M. J., et al., 2014, ApJ, 783, 130

Riener M., Kainulainen J., Henshaw J. D., Orkisz J. H., Murray C. E., Beuther H., 2019, A&A, 628, A78

Rigby A. J., et al., 2016, MNRAS, 456, 2885

Rigby A. J., et al., 2018, A&A, 615, A18

Rigby A. J., et al., 2021, MNRAS, p. stab200

Rigby A. J., et al., 2024, Monthly Notices of the Royal Astronomical Society, p. stae030

Rosolowsky E. W., Pineda J. E., Kauffmann J., Goodman A. A., 2008, ApJ, 679, 1338

Roueff E., Abgrall H., Czachorowski P., Pachucki K., Puchalski M., Komasa J., 2019, A&A, 630, A58

Roy A., et al., 2013, ApJ, 763, 55

Roy A., et al., 2015, A&A, 584, A111

Sanhueza P., et al., 2019, ApJ, 886, 102

Schisano E., et al., 2014, ApJ, 791, 27

Schneider S., Elmegreen B. G., 1979, ApJS, 41, 87

Schneider N., Csengeri T., Bontemps S., Motte F., Simon R., Hennebelle P., Federrath C., Klessen R., 2010, A&A, 520, A49

Schuller F., et al., 2017, A&A, 601, A124

Schwab F. R., 1984, AJ, 89, 1076

Schwörer A., et al., 2019, A&A, 628, A6

Shirley Y. L., 2015, PASP, 127, 299

Simon R., Jackson J. M., Rathborne J. M., Chambers E. T., 2006, ApJ, 639, 227

Smith R. J., et al., 2020, MNRAS, 492, 1594

Solomon P. M., Rivolo A. R., Barrett J., Yahil A., 1987, ApJ, 319, 730

Stahler S. W., Palla F., 2004, The Formation of Stars. Wiley-VCH, Weinheim

Sternberg A., Dalgarno A., 1995, ApJS, 99, 565

Svoboda B. E., et al., 2019, ApJ, 886, 36

Szűcs L., Glover S. C. O., Klessen R. S., 2016, MNRAS, 460, 82

Tafalla M., 2001, Astronomical Society of the Pacific, 243, 103

Taylor M. B., 2005, in Shopbell P., Britton M., Ebert R., eds, Astronomical Society of the Pacific Conference Series Vol. 347, Astronomical Data Analysis Software and Systems XIV. p. 29

Terebey S., Chandler C. J., Andre P., 1993, ApJ, 414, 759

The Casa Team et al., 2022, PASP, 134, 114501

Tokuda K., et al., 2019, ApJ, 886, 15

Treviño-Morales S. P., et al., 2019, A&A, 629, A81

Tritsis A., Tassis K., 2018, Science, 360, 635

Tritsis A., Bouzelou F., Skalidis R., Tassis K., Enßlin T., Edenhofer G., 2022, MNRAS, 514, 3593

Urquhart J. S., Figura C. C., Moore T. J. T., Hoare M. G., Lumsden S. L., Mottram J. C., Thompson M. A., Oudmaijer R. D., 2014, MNRAS, 437, 1791

Urquhart J. S., et al., 2018, MNRAS, 473, 1059

Vázquez-Semadeni E., Palau A., Ballesteros-Paredes J., Gómez G. C., Zamora-Avilés M., 2019, MNRAS, 490, 3061

Virtanen P., et al., 2020, Nat Methods, 17, 261

Wang P., Li Z.-Y., Abel T., Nakamura F., 2010, ApJ, 709, 27

Ward-Thompson D., Whitworth A. P., 2011, An Introduction to Star Formation. Cambridge University Press, Cambridge, doi:10.1017/CBO9780511974021

Ward-Thompson D., Scott P. F., Hills R. E., Andre P., 1994, MNRAS, 268, 276

Williams J. P., Blitz L., McKee C. F., 2000, in Mannings V., Boss A. P., Russell S. S., eds, Protostars & Planets IV. p. 97, doi:10.48550/arXiv.astro-ph/9902246

Williams G. M., Peretto N., Avison A., Duarte-Cabral A., Fuller G. A., 2018, A&A, 613, A11

Wilson T. L., Rohlfs K., Hüttemeister S., 2013, Tools of Radio Astronomy. Astronomy and Astrophysics Library, Springer Berlin Heidelberg, Berlin, Heidelberg, doi:10.1007/978-3-642-39950-3

Wright E. L., et al., 2010, AJ, 140, 1868

Xu F.-W., et al., 2023, MNRAS, 520, 3259

Zernike F., 1938, Physica, 5, 785

Zhang Q., Wang K., Lu X., Jiménez-Serra I., 2015, ApJ, 804, 141

Zhou J.-W., et al., 2022, MNRAS, 514, 6038

van Cittert P., 1934, Physica, 1, 201

van der Walt S., Schönberger J. L., Nunez-Iglesias J., Boulogne F., Warner J. D., Yager N., Gouillart E., Yu T., 2014, PeerJ, 2, e453