PRESTELLAR CORES IN PERSEUS

PRESTELLAR CORES IN PERSEUS

by DAMIEN ROBERTSON

A Thesis Submitted to the School of Graduate Studies in Partial Fulfilment
of the Requirements for the Degree Master of Science

McMaster University MASTER OF SCIENCE (2014) Hamilton, Ontario (Physics)

TITLE: Prestellar Cores in Perseus

AUTHOR: Damien Robertson, B.Sc. (Saint Mary's University)

SUPERVISOR: Dr. C.D. Wilson

NUMBER OF PAGES: xii, 76

Abstract

Star formation is a complex hierarchical process that witnesses the transfer of mass among a range of scales from large diffuse molecular clouds to crowded clumps and finally down to prestellar cores. The final stage of this process has prestellar cores actively accreting matter while undergoing gravitational collapse on their way to becoming main sequence stars. This thesis presents multi wavelength submillimeter observations of the Perseus molecular cloud using 160 μm, 250 μm, 350 μm, and 500 μm maps of thermal dust emission from the *Herschel* space observatory. Additionally $C^{18}O$ $J = 3 \rightarrow 2$ spectral line emission is observed in four star forming clumps within Perseus using the James Clerk Maxwell Telescope. Spectral line emission allows for the separation of material along the line of sight. Prestellar core mass is derived from observational maps using various source finding algorithms. The mass is overestimated when compared to prestellar core mass found from spectral line data. This overestimation can be mitigated with careful selection of source finding algorithm and background removal. Further, the prestellar core mass derived from spectral line data was the closest match to the initial stellar mass function over dust maps. However, both the spectral line masses and dust map masses do not agree with the IMF confirming a star forming efficiency factor in the evolutionary step between prestellar core and main sequence star. Lastly, a filamentary analysis finds that high mass stars preferentially form in crowded regions close to, or contained within, filament structure.

Acknowledgements

The emotional support for this work was solely provided by my family. The birth of my son Beckett will forever be linked to my research here. Thank you Jill for your infinite patience, understanding and help with achieving this goal.

Thank you Kaz, Tara, Max and Jon for being great officemates and friends. Your help at all hours of the day and night was appreciated. My supervisor, Dr. Christine Wilson, deserves much credit for allowing me to travel off the beaten path to investigate areas of interest on my random walk to the final results - it was very much appreciated.

Contents

List of Figures

List of Tables

List of Abbreviations and Symbols

α_{high} High mass power law slope from double power law CMF

α_{low} Low mass power law slope from double power law CMF

β Spectral index

μ Mean mass from lognormal CMF

ν_0 Reference frequency

σ Standard deviation from lognormal CMF

σ_{rms} Root Mean Square

M_{break} Break mass from double power law CMF

2PL Double Power Law CMF (Equation 5.1)

CMF Cumulative Mass Function

DEC Declination

DMF Discrete Mass Function

FWHM Full-Width Half Maximum

GMC Giant Molecular Cloud

HARP Heterodyne Array Receiver Program

HIPE Herschel Interactive Processing Environment

IMF Interstellar Mass Function

JCMT James Clerk Maxwell Telescope

K-S Kolmogorov-Smirnov

LOGN Lognormal CMF (Equation 5.2)

$M_\odot$ Solar Mass (1 $M_\odot$ = 1.989 $\times$ 10^{30} kg)

MC Monte Carlo

PACS Photodetector Array Camera and Spectrometer

pc parsec (1 pc = 3.0857 $\times$ 10^{16} m)

PDF Probability Distribution Function

PP Position-Position

PPP Position-Position-Position

PPV Position-Position-Velocity

RA Right Ascension

S/N Signal to Noise Ratio

SED Spectral Energy Distribution

SFE Star Formation Efficiency

SPIRE Spectral and Photometric Imaging Receiver

YSO Young Stellar Object

Declaration of Academic Achievement

The work presented here is the result of research performed by myself during the years 2012–2014. Results which have substantial contributions from other authors are clearly prefaced and the contributions of those authors are indicated.

Chapter 1

Introduction

In a general sense star formation is a hierarchical process: a giant molecular cloud (GMC) collapses and turbulence fragments the cloud into large clump-like substructure (Larson, 1981; Lada & Lada, 2003). As turbulence dissipates, filamentary structure emerges and prestellar cores take shape within gravitational instabilities (Arzoumanian et al., 2011). The molecular clouds which host star formation can be quite large, often tens of parsecs in size with 10^3 - 10^5 solar masses ($M_\odot$) of gas and dust (Tielens, 2005). These molecular clouds are also quite dense, usually $\sim$ 10 - 100 cm^{-3} but can reach densities of $>10^3$ cm^{-3} in star forming structures (Tielens, 2005; Pineda et al., 2010). Due to the density found in these GMCs they are intrinsically self-shielding. Ionizing radiation or interstellar radiation fields are blocked from penetrating deep into the cloud (Tielens, 2005). As a result, these opaque clouds tend to have very low internal temperatures, $<$ 50 K and can be as low as $\sim$ 10 K in denser star forming regions (Tielens, 2005). The various gaseous molecules can atomically de-excite which radiates energy away at infrared and radio wavelengths (Tielens, 2005). Observing GMCs at these wavelengths provides some of the only clues to the processes which govern star formation.

1.1 Young Stellar Objects

The distinction between a prestellar core and an unbound clump is generally size and density. Cores tend to be $\lesssim 0.1$ pc in diameter and have densities $\gtrsim 10^4$ cm^{-3} (Myers & Benson, 1983). In order to categorize the evolutionary stage of observed prestellar cores, or young stellar objects (YSO), a classification scheme developed by Charles Lada is generally accepted (Lada, 1987). The classification scheme is based on spectral index, a, which measures the slope of the spectral energy distribution (SED) between $1 \leq \lambda \leq 10$ μm (Lada, 1987). In this early work Lada first proposes three YSO classes (class I - III). Class I YSO are defined with spectral index $0 < a \leq 3$ and have broader than blackbody SEDs with rising flux at longer wavelengths (top right panel in Figure 1.1) (Lada, 1987). Class II YSO are defined with spectral index $-2 < a \leq 0$ and have broader than blackbody SEDs with flat or falling flux at longer wavelengths (bottom left panel in Figure 1.1) (Lada, 1987). Class III YSO are defined with spectral index $-3 < a \leq -2$ and the SED is well represented with a reddened black body (bottom right panel in Figure 1.1) (Lada, 1987). Figure 1.1 shows example SEDs of class I - III YSO with the emerging protostellar SED.

Later, Lada along with Adams et al. (1987) revised the spectral index a to encompass wavelengths from $1 \leq \lambda \leq 10$ μm and attempted to delineate class I and II YSO (Adams et al. 1987). The authors suggested class I_1 YSO are in a pure mass infall phase whereas class I_2 YSO have collimated outflow and show sharply increasing SEDs (Adams et al., 1987). The authors also suggest class II_1 YSO which have a double hump or wide, flat SED and a class II_2 which show only a single hump in the SED (Adams et al., 1987). In 1993 André et al. discovered evidence of a YSO with no detected flux at $\lambda < 10$ μm. This seemingly unique YSO was designated as class 0 and represents the earliest

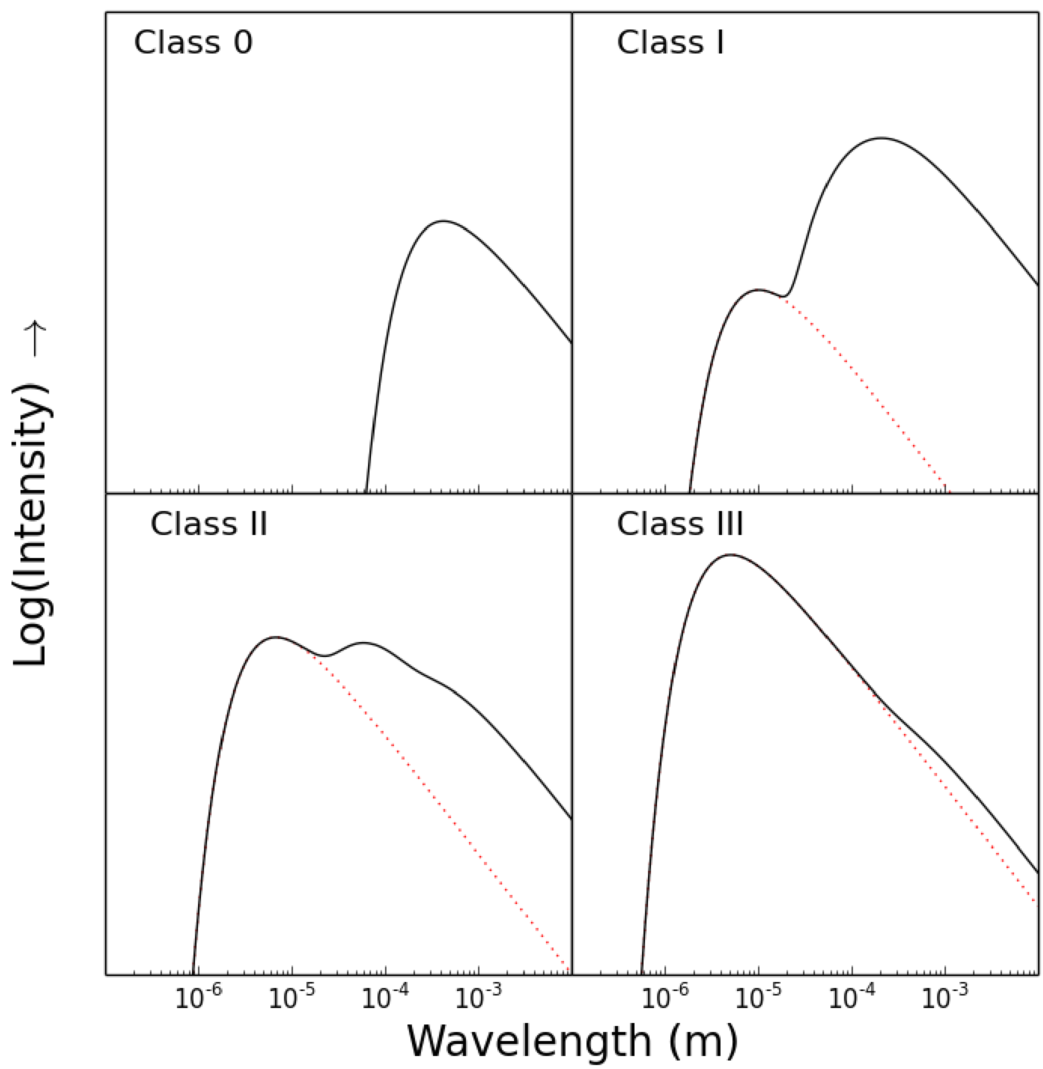

Figure 1.1: Classification scheme for the spectral energy distributions of embedded young stellar objects. Solid line is a combination of emerging protostellar SED (red dotted line) and dust SED of envelope.

stage in protostar evolution in which no blackbody component of the protostar is observed (top left panel in Figure 1.1) (Andre et al., 1993). Class 0 and starless cores, can resemble a cold dusty black body with a characteristic temperature of < 30 K (Hatchell et al., 2007b). The YSO classes ranging from 0 - III are thought to mimic an evolutionary track for cores as they condense, form protostars and shed their dust and gas envelopes on the way to the main sequence (Andre et al., 1993). Despite the work of Adams et al. (1987) to delineate class I and II YSO it seems that the standard class 0 - III of Lada (1987) and Andre et al. (1993) are widely used. Starless cores are noted as having a similar SED profile to class 0 YSO, however; they do not have a central protostar yet. Starless cores may be distingushed from class 0 YSOs using the central density profile or looking for emission from the protostar at radio wavelengths and probing for inflows or outflows. Work completed by Hatchell et al. (2007b) found 103 YSOs in Perseus, 47 of those are starless cores, 34 are class 0 and 22 are class I.

1.2 Mass Evolution

It has been a challenge to reconcile the star formation processes that span large scale molecular clouds to the observed initial stellar mass function (IMF). It is thought that turbulence plays a part in not only breaking up molecular clouds into clumps but also fragmenting clumps into cores within clouds (Larson, 1981). The initial conditions of the clouds are thought to set the stage for clump properties which in turn dictate core fragmentation (Padoan & Nordlund, 2002; Hennebelle & Chabrier, 2008). However, the connection from core properties to the observed IMF is perhaps the most difficult to reconcile. Recent theoretical papers attempt to accomplish this by incorporating gravoturbulent fragmentation, competitive accretion among cores and varying star forming efficiencies (Veltchev et al., 2011; Donkov et al., 2012). Ultimately, it may be a combination of many factors which settle core properties into the IMF that occur as a stochastic process in the late stages of star formation.

A convenient way to quantify a population of prestellar core masses is by creating mass distributions, or simply, a mass function. Edwin Salpeter (1955) was the first to record a census of stars and create a initial stellar mass function. Salpeter found that the initial mass function of stars generally follows a power law,

$$\frac{dN}{d\mathrm{Log}(M)} = AM^{-x}, \tag{1.1}$$

where A is a scaling factor and x is the power law slope (Salpeter, 1955; Chabrier, 2003). This power law relation holds for the mass function when complete sampling is achieved and a cut-off mass defines the lower mass limit of this relation. Many authors attempt to fit multiple power law relations to the same mass function and define a break mass which separates power law regimes (Chabrier, 2003). Often, it is convenient to express the power law slope as $\alpha = x + 1$ due to variations in constructing the mass function, with Salpeter's canonical slope as $x = 1.35$, $\alpha = 2.35$ (Salpeter, 1955). Equation 1.1 is also called the discrete mass function (DMF) because it divides the masses into bins that are $\mathrm{Log}(M)$ in width (top panel in Figure 1.2). This mass binning is often arbitrary and populations with a low sample number ($\lesssim 100$) can be difficult to interpret (Reid & Wilson, 2006). A cumulative mass function (CMF) does not suffer from the arbitrariness of mass binning data or the difficulty with low sample numbers. The counterpart to Equation 1.1 for the CMF is

$$N(>M) = -\frac{A}{\alpha - 1} M^{-(\alpha-1)}, \tag{1.2}$$

where A is a scaling factor and α is now the power law slope (Reid & Wilson, 2006). It is worth noting that the power law slopes of Equation 1.1 and 1.2 are equivalent when using the relation $\alpha = x + 1$. Figure 1.2 shows the same mass population interpreted using the DMF, top panel, and the CMF, bottom panel. The mass population shown in Figure 1.2 is that of the IMF from Chabrier (2005),

$$\xi = \begin{cases} 0.093 \exp\left[\frac{-(\mathrm{Log}(M) - \mathrm{Log}(0.2))^2}{2(0.55)^2}\right], & M < 1.0M_\odot \\ 0.043M^{-1.35}, & M \geq 1.0M_\odot, \end{cases} \tag{1.3}$$

with power law slopes shown as dotted red lines that have $\alpha = 2.35$ fit to $M > 1.0M_\odot$.

Salpeter discovered that among larger stars ($\gtrsim 1$ M$_\odot$) for which the census was complete the mass distribution had a well defined power law slope of $\alpha = -2.35$ (Salpeter, 1955). However, when looking at the mass function of clumps within the GMCs, before a stellar population has been established, the power law slope is generally shallower (Pekruhl et al., 2013). For example, Pekruhl et al. (2013) determines a power law slope of $\alpha \sim -1.9$ for a population of prestellar candidates in the Carina Nebula. Results like this imply that evolutionary processes like star formation efficiency, core fragmentation and competitive accretion shape the distribution of masses among cores before finally arriving at the IMF (Offner et al., 2013). An example of this is seen in Figure 1.3 as the dotted line that has a power law slope of $\alpha = 1.9$ above $\sim 1.0M_\odot$. A shallow power law slope, like that seen in Figure 1.3, implies that there are more high mass objects than if the slope was steeper. In order for the dotted CMF to evolve into the solid CMF the high mass cores need to lose mass to match the IMF power law slope. The break mass between two power law mass regions can imply a star formation efficiency (SFE) (Offner et al., 2013). For example,

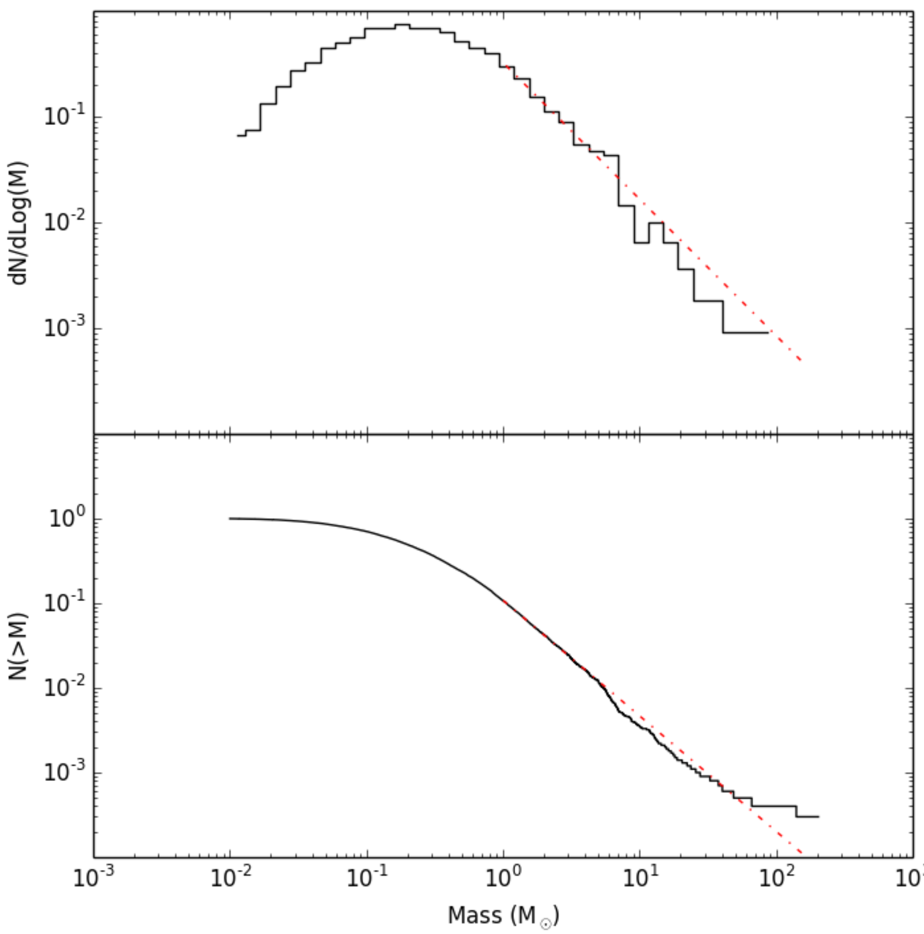

Figure 1.2: *Top:* DMF of stellar IMF, *bottom:* CMF of stellar IMF. Dotted red line in both panels is $\alpha = 2.35$.

in Figure 1.3 both the solid and dotted CMFs have a break mass of $\sim 1.0M_{\odot}$; however the dashed CMF has a break mass of $\sim 7.0M_{\odot}$ with $\alpha = 2.35$. If the dashed CMF were to evolve into the solid CMF the population as a whole would need to lose mass. In the example listed above the dashed CMF would need to uniformly shed about 86 % of core mass to match the solid CMF implying a SFE of approximately 14 %. Low SFE would suggest that much of the mass of the protostar is lost before becoming a bona fide star (Offner et al., 2013). Apparently, anywhere from 50% - 80% of protostellar mass is lost, recycled back into the GMC before formation stops (Offner et al., 2013). However, discussion of mass functions and star formation efficiency requires accurate mass estimation.

When considering the crowded environments in which stars form it is ill-advised to comb through observations looking for prestellar core candidates by hand. Instead, most astronomers employ various source-finding algorithms which provide unbiased candidate lists using standardized and predictable exploration. However, depending on observation or environmental specific factors and choice of algorithm, prestellar core mass estimation may have an intrinsic bias often towards overestimation. Recently work by Ward et al. (2012) suggests that prestellar core mass may be overestimated by a factor of three. This overestimation comes from the fact that any two-dimensional map is a flattened projection of an inherently three-dimensional cloud along the line of sight (Ward et al., 2012). Material along the line of sight or the combination of more than one core can cause an overestimation in mass. The impact of overestimating prestellar core mass has implications for all of stellar evolution from initial cloud collapse to final stellar population.

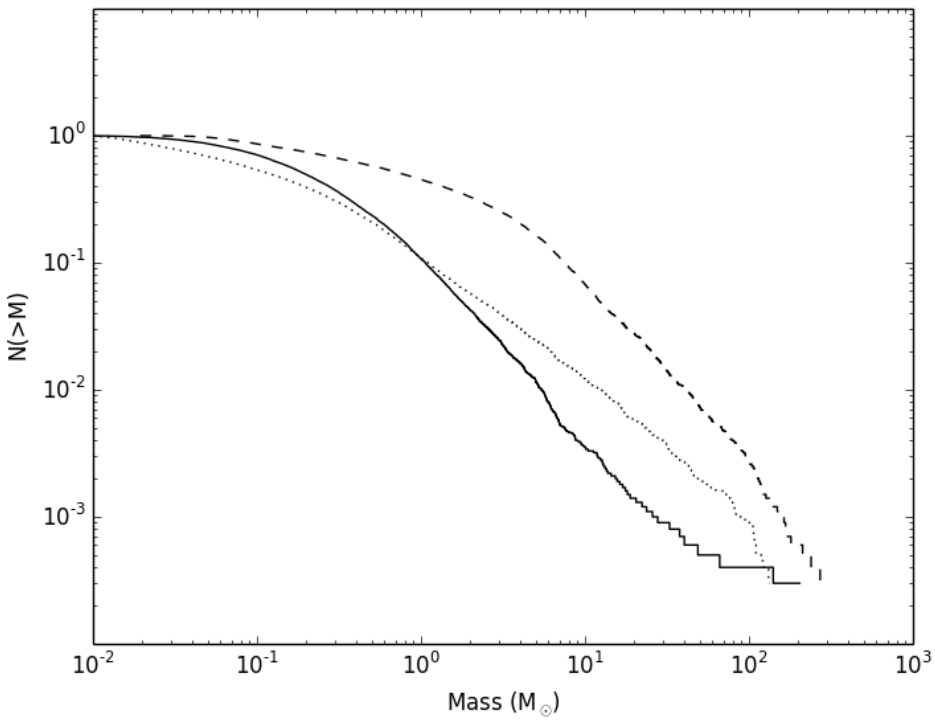

Figure 1.3: Solid line is a CMF with $\alpha = 2.35$ for $M{>}1.0M_\odot$, dotted line is a CMF with $\alpha = 1.9$ for $M{>}1.0M_\odot$ and dashed line is a CMF with $\alpha = 2.35$ for $M{>}7.0M_\odot$

1.3 Perseus Molecular Cloud

The Perseus molecular cloud has been extensively studied in the last decade (see Kirk et al. (2006, 2007); Hatchell et al. (2005, 2007a); Lombardi et al. (2010)). The popularity of Perseus as a subject of study is mainly due to the close proximity to Earth and active star formation. The Perseus molecular cloud is located approximately $\sim$ 250 pc away from Earth in the constellation of Perseus (RA 3:34:00 and DEC +31:00:00) (Kirk et al., 2006). Recent work using extinction maps suggests the GMC is tilted away from Earth from West (212 $\pm$ 18 pc) to East (260 $\pm$ 18 pc) (Lombardi et al., 2010). The Perseus molecular cloud has been noted as a site of active star formation with observed YSOs in a range of evolutionary stages from class 0 - I in several star forming clumps, most notably NGC 1333 and IC 348 (see Figure 2.1) (Hatchell et al., 2007b; Curtis et al., 2010a). Clumps in the Perseus GMC are areas of higher than average densities with clusters of protostars forming within, see Figure 2.1 (Curtis et al., 2010a). The Perseus GMC is associated with Per OB2, the second closest OB association from the Sun, which could also be influencing the evolution of star formation within the GMC as evident in IC 348 and triggering star formation (Curtis et al., 2010a; Sadavoy, 2013). In total the Perseus molecular cloud covers an angular extent of 1.5×5.0 degrees with an estimated mass of $\sim\ 1.7 \times 10^4$ $M_\odot$ (Bachiller & Cernicharo, 1986).

Most recent observations of the Perseus GMC have observed the thermal emission by dust via 450 and 850 μm and the de-excitation of N_2H^+ and CO isotopologues using radio spectral line data (Kirk et al., 2006; Hatchell et al., 2007a; Curtis et al., 2010a). Observing the thermal emission of dust can offer wide field maps and very fine detail (Sadavoy et al., 2012). However, thermal dust emission yields a two dimensional position-position (PP) map which is a flattened projection of a three dimensional object along the line of sight. Observing the de-excitation of molecules offers not just the position of the emission in the sky but also the radial velocity of the radiating material (Kirk et al., 2006). These spectral line data sets provide three dimensional position-position-velocity (PPV) data. It is worth making a distinction that PPV data cubes are not the same as true position-position-position (PPP) representations. With PPV data the third dimension is *velocity*, not physical space as in PPP data and emission is still flattened along the line of sight. The advantage of PPV data is that distinct cores that happen to be combined along a line of sight in PP data may be separated in velocity space. Recent work by Ward et al. (2012) using synthetic observations suggests that core masses derived from PPV data match true PPP mass more accurately than PP derived masses.

There are a few recent attempts to estimate a protostellar mass function of the Perseus GMC. Using submillimeter continuum maps at 850 μm and extinction data, Kirk et al. (2006) find a high mass ($\gtrsim 1$ $M_{\odot}$) power law slope of ~ -2.5 to -3.0 of a sample of <60 protostellar cores, far steeper than the Salpeter IMF value of -2.35. These results seem contradictory to the work of Curtis & Richer (2010) who also use a survey of 850 μm emission and find a high mass ($\gtrsim 6$ $M_{\odot}$) power law slope of -2.0 ± 0.1 and -3.15 ± 0.8 depending on the choice of source finding algorithms and deduce a SFE of 10 % - 20 %. These results straddle the Salpeter IMF and highlight that estimated mass is dependent on how one chooses to determine the extent of YSOs (Curtis & Richer, 2010). Additional work by Enoch et al. (2008) using 1.1 mm dust emission of Perseus, Serpens, and Ophiuchus estimates a high mass (>0.8 $M_{\odot}$) slope of -2.3 ± 0.4 along with a SFE of 25 %.

The Perseus molecular cloud is an excellent region to observe active star formation where YSOs exist as both starless cores and protostars. This work will compare mass estimates from new multi-band submillimeter dust emission as well as optically thin $C^{18}O$ spectral line data. The goal of the mass comparison is to test various source findings algorithms and discover any intrisic biases. Chapter 2 will discuss the data products used in this analysis, Chapter 3 discusses source finding algorithms and Chapter 4 addresses data analysis and mass determination. Chapter 5 will explore various mass comparisons between algorithms both indirectly on core population and directly on a core-by-core basis. Chapter 5 also presents unique core-to-filament distance analysis using submillimeter dust emission.

Chapter 2

Instrumentation and Data

The cold temperatures (<50 K) in molecular clouds imply that thermal emission from dust peaks at submillimeter wavelengths. Typical temperatures found in denser star forming regions can be as low as 10 - 30 K which correspond to black body emission peaking 97 - 290 μm (André et al., 2010). Unfortunately the Earth's atmosphere is opaque to these wavelengths. Additionally, the low energy rotational transition, $J = 3 \rightarrow 2$, of the Hydrogen tracer CO emits at millimeter and submillimeter wavelengths (Curtis et al., 2010a). The Earth's atmosphere does has several transmission windows where CO can be observed. However, to mitigate for losses from atmospheric seeing and water vapour one needs to locate a telescope high above the bulk of the atmosphere.

2.1 *Herschel*

The *Herschel* Space Observatory was a submillimeter telescope capable of observing low temperature thermal dust emission. Commissioned by the European Space Agency and launched on May 14, 2009, *Herschel* solves the issue of atmospheric opacity by rising above it and observing from space. Equipped with a 3.5 m diameter Cassegrain telescope *Herschel* can observe far infrared wavelengths from 55-671 μm (Pilbratt et al., 2010). There are three main scientific instruments on board *Herschel*: two cameras and a heterodyne spectrometer (Pilbratt et al., 2010). This work uses the two main cameras on board *Herschel* which cover a spectral range of 70 - 500 μm.

The first of the two cameras is the Photodetector Array Camera and Spectrometer (PACS). PACS consists of two parts, an imaging photometer and a integral field spectrometer. This work uses data from the imaging photometer. Light coming into the PACS photometer is split into three bands based on wavelength. The red band is the longest wavelength (130 - 210 μm) and is directed to a 32 $\times$ 16 pixel bolometer array (Poglitsch et al., 2010). Either the green or blue band (60 - 85 μm and 85 - 130 μm respectively) is directed to a 64 $\times$ 32 pixel bolometer array (Poglitsch et al., 2010). The field of view provided in all three bands is $3.5' \times 1.75'$ (Poglitsch et al., 2010). In order to decrease thermal noise, the bolometer arrays are kept at an operating temperature of 0.3 K using liquid He (Poglitsch et al., 2010). PACS is able to observe in the red and either the blue or green bands simultaneously.

The second camera is the Spectral and Photometric Imaging REceiver (SPIRE). Similar to PACS, SPIRE contains an imaging photometer and a spectrometer; again this work only uses data from the imaging photometer. The SPIRE imaging photometer observes in three spectral bands centered on 250, 350 and 500 μm (Griffin et al., 2010). The pixel count for each band is 139, 88 and 43 respectively and pixels are arranged in 15 $\times$ 9, 13 $\times$ 7 and 9 $\times$ 5 close packed grid (Griffin et al., 2010). The field of view for each band is $4' \times 8'$ which is offset by $11'$ from the center of the *Herschel* telescope (Griffin et al., 2010). Similar to PACS, the bolometer arrays are kept at an operating temperature of 0.3 K using liquid He. SPIRE is able to observe in all three bands simultaneously.

Herschel can observe using both the PACS and SPIRE cameras simultaneously (Pilbratt et al., 2010). When *Herschel* observes large areas of the sky it does so by scanning an area along rows and then along columns (Poglitsch et al., 2010). This pattern allows *Herschel* to 'paint' the sky allowing the detectors to fully sample the observation area. The rate at which *Herschel* scans when observing with both PACS/SPIRE cameras, called parallel mode, is set at $20''$ and $60''$ per second, (slow and fast) (Poglitsch et al., 2010).

2.1.1 *Herschel* Observations and Data Reduction

The Perseus molecular cloud was observed in February 2010 to map the Western half of the cloud and again in February 2011 to map the Eastern half. The complete *Herschel* Perseus molecular cloud map covers an area of approximately 12.3 deg^2. Both *Herschel* observations were taken at scan speeds of $60''$/s in parallel mode yielding simultaneous observations at 70, 160, 250, 350 and 500 μm wavelengths.

The *Herschel* observations were reduced by Sarah Sadavoy (UVic); full data reduction details can be found in her Ph.D. thesis[1]. Here I provide a summary of the processes that were performed on the data. The PACS/SPIRE raw data were reduced using the Herschel Interactive Processing Environment (HIPE) (Sadavoy, 2013). The first step performed was to convert the raw data, voltages from the bolometers, into a digital numerical quantity while filtering bad pixels and ascribing telemetry information (Sadavoy, 2013). The second step was to then convert the numerical quantities into flux units using up-to-date calibration information (Sadavoy, 2013). In this step any problems with data, e.g. spurious or erratic readings, are corrected for (Sadavoy, 2013). The last

[1] https://dspace.library.uvic.ca:8443//handle/1828/4725

step is to create a sky-map with the final data products using a separate program called *scanamorphos* (Roussel, 2012). *Scanamorphos* was used for map creation over the HIPE map routines due to the superior performance when correcting for large scale stripes from the scan process and negative artefacts around bright sources (Sadavoy, 2013). Finally, the maps have had a zero-point offset corrected based on Planck HFI and IRAS data. Refer to Table 2.1 for information regarding the final data products. Due to the fast scan rate the 70 μm waveband was undersampled and not used in data analysis (Sadavoy, 2013).

PACS/SPIRE Band	70 μm	160 μm	250 μm	350 μm	500 μm
Pixel Size	3.2″	4.5″	6.0″	10.0″	14.0″
BEAM FWHM	8.4″	13.5″	18.2″	24.9″	36.3″

Table 2.1: *Herschel* PACS/SPIRE Map Details

2.2 James Clerk Maxwell Telescope

The James Clerk Maxwell Telescope (JCMT) is a 15 m diameter reflecting telescope located on Mauna Kea, Hawaii designed for submillimeter wavelengths, primarily 850 μm (353 GHz) (Dent et al., 2000). JCMT is also equipped with a Heterodyne Array, formally known as HARP. HARP is capable of observing spectral lines found at 325 - 375 GHz (800 - 923μm) with up to 31 KHz (0.03 km s^{-1}) resolution. Spectral lines observed in this frequency range are predominately the $J = 3 \rightarrow 2$ rotational transition of the three isotopologues of CO, ^{12}CO, ^{13}CO and C^{18}O found at 345.8, 330.6 and 329.3 GHz, respectively (Buckle et al., 2009; Curtis et al., 2010a). HARP is comprised of 16

Superconductor-Insulator-Superconductor detectors on a 4 × 4 grid with 16″ separation between detectors (Buckle et al., 2009). The under-sampled field of view is 104″ × 104″. There are many observation modes for HARP; however, when observing large regions a scanning or on-the-fly mapping/rastering technique must be used (Buckle et al., 2009). Since the HARP beam is under-sampled the array must be rotated 14.48 deg with respect to the scan direction to ensure full sky coverage on a 7.3″ grid (Buckle et al., 2009). Similar to *Herschel* scanning is performed by sweeping first along rows, then along columns (Buckle et al., 2009).

2.2.1 HARP Observations and Data Reduction

HARP observed four regions within Perseus for the $J = 3 \rightarrow 2$ transition of $C^{18}O$ (329.3 GHz): IC 348 in the Eastern half and NGC 1333, L1448 and L1455 in the Western half. These sub-regions were chosen due to their high extinction, which indicates areas of active star formation(Curtis et al., 2010a), see Figure 2.1. Spectral line observations were carried out over nine nights from December 17, 2007 to January 12, 2008 with NGC 1333 mapped on July 27, 2007 (Curtis et al., 2010a).

HARP observations of $C^{18}O$ were reduced by Emily Curtis and published in Curtis et al. (2010a,b); Curtis & Richer (2011). Here I provide a summary of the processes that were performed on the data; for full data reduction description see Curtis et al. (2010a). The four regions observed in Perseus were mapped using a scanning technique with the following provisions: spacing between scan rows was set to half the width of the detector array; regions would be scanned twice in a 'basket weave' pattern with a small offset made on the repeat scans (Curtis et al., 2010a). This method ensures that more than one detector would observe any patch of the region and in total 4 scans were made

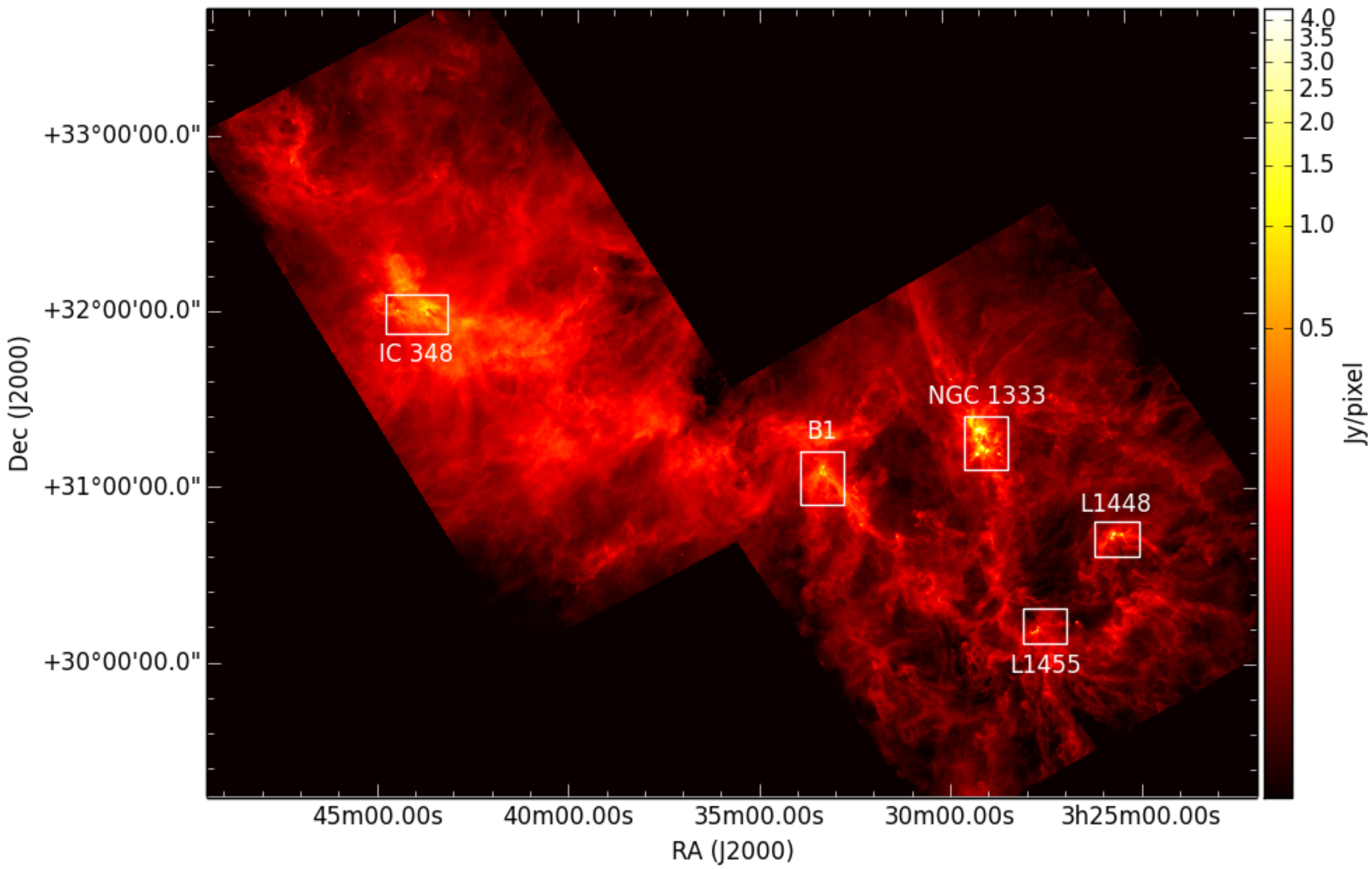

Figure 2.1: Perseus molecular cloud at 250 μm from *Herschel* SPIRE camera. Only the subregions NGC 1333, IC 348, L1148, and L1455 are observed in $C^{18}O$ with HARP.

of each region (Curtis et al., 2010a). The raw data from HARP was processed with the *Starlink*[2] software collection. Any bad pixel or noisy data was flagged and removed via the software and the remaining spectra were sampled on a $3''$ grid with an equivalent beam full-width half maximum (FWHM) of $17.7''$ (Curtis et al., 2010a). The spectra were re-binned to provide a resolution of 0.15 km s^{-1} (Curtis et al., 2010a). After gridding a baseline was subtracted using line-free (off-source) samples of the data cube (Curtis et al., 2010a).

[2]http://starlink.jach.hawaii.edu

Chapter 3

Source Finding Algorithms

Since it is not practical or advisable for an individual to scour data sets manually looking for sources it is important to thoroughly understand how source finding algorithms function. Source finding algorithms are advantageous to use because they are predictable, repeatable and allow minute adjustments which offer time savings over manually searching. Presented here are the source finding algorithms used in this work. The summaries of the algorithms are presented in order of increasing complexity and one should consult the original work for full details on performance.

3.1 `Clumpfind`

The source finding algorithm `clumpfind` was developed in 1994 to mimic how an observer might locate and measure sources in spectral line PPV data. The algorithm adopts the simple notion that all sources found in a data set have a local maximum that can be traced down to base level noise (Williams et al., 1994). `Clumpfind` accomplishes this by considering the data as a discrete set of linear evenly spaced contours with the lowest contour characterizing the noise of the data (Williams et al., 1994). The algorithm begins at the highest contour level, looks for local maxima and follows connected pixels down to lower contour levels (Williams et al., 1994). An important issue to consider when source finding in crowded regions is blending. Source blending occurs if two or more sources are individually unresolved but instead resemble a hybrid, non-circular lump of emission with resolved local maximums; see left panel of Figure 3.1 for example. `Clumpfind` does not deblend sources, instead shared pixels are partitioned off to the source whose contour level is closer; see right panel of Figure 3.1 (Williams et al., 1994).

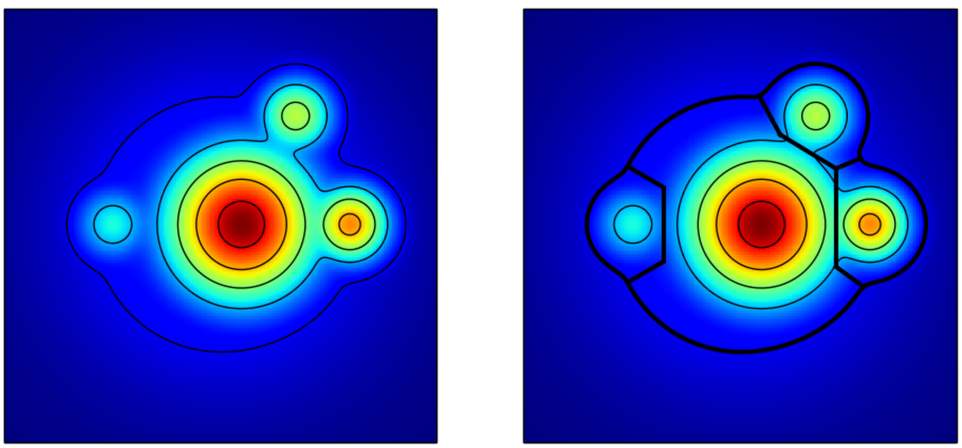

Figure 3.1: Simulated contoured sources with shared pixels at lower contours, darker lines in right panel show source boundaries from `clumpfind`.

It is worth noting that any given pixel may only belong to one source; `clumpfind` does not divide emission from a pixel based on blending (Williams et al., 1994). Additionally, `clumpfind` does not subtract a background and reports all pixels above the lowest threshold, partitioned to a source, as total emission; see top panel of Figure 3.2. The `clumpfind` algorithm was later modified to accommodate not just three dimensional data but also two dimensional data. The suffix '2D' or '3D' will be used to differentiate when `clumpfind` is used on two or three dimensional data sets.

3.2 SExtractor

Developed primarily for optical and near infrared extragalactic data `SExtractor` functions similarly to `clumpfind` with a few additions. Before source detection `SExtractor` attempts to identify and remove a user defined local background via a combination of sigma-clipping and a median filter using mode estimation (Bertin & Arnouts, 1996). A background map is created and smoothed using bilinear interpolation and subtracted from the original map (Bertin & Arnouts, 1996). Sources are detected using the background subtracted map and contour thresholding, a method similar to what is employed in `clumpfind`. `SExtractor` also employs a deblending routine. Sources are deblended based on a dendrogram method of identifying child sources from a parent source; see middle panel of Figure 3.2 (Bertin & Arnouts, 1996).

Blended sources are re-thresholded at 30 exponentially spaced contours where source peaks are traced down to a common contour level (Bertin & Arnouts, 1996). Shared pixels are then divided based on expected contribution from a bivariate Gaussian fit to the child source (Bertin & Arnouts, 1996). An exmaple of `SExtractor` deblending can be seen in the middle panel of Figure 3.2. The light gray parent contour will have emission divided among the dark and mid gray child sources. The lowest contour level is estimated via the background map and a user defined S/N is specified for a source to be reported. Total emission per deblended source is reported above the lowest estimated contour. However, unlike `clumpfind`, `SExtractor` only works on two dimensional PP maps.

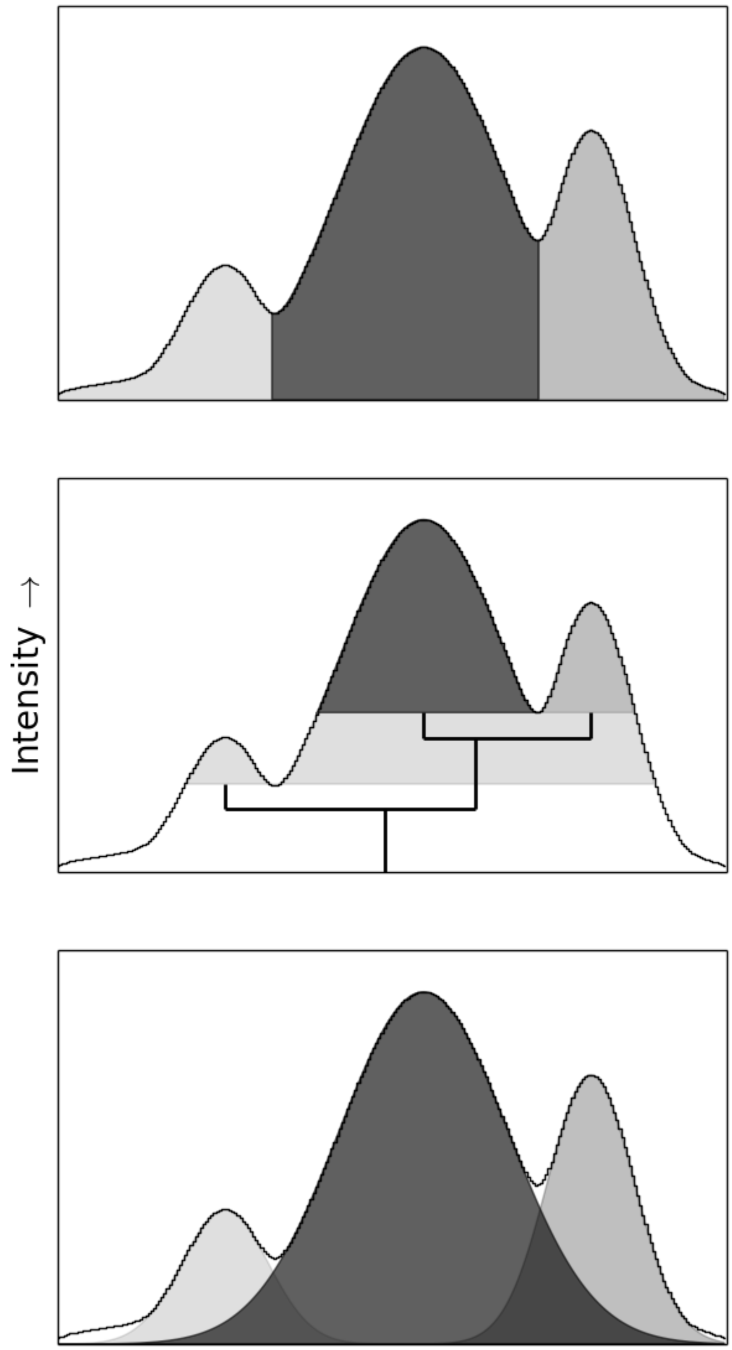

Figure 3.2: Simulated deblending of sources, *top:* `clumpfind` attempts no deblending, *middle:* `SExtractor` divides parent source to child sources, *bottom:* `getsources` Deblending based on source FWHM.

3.3 Getsources

With the advent of *Herschel*, submillimeter maps of thermal dust emission have become incredibly detailed and can have a dynamic range that spans 10^3 or more. The source finding algorithm `getsources` was designed specifically for highly detailed maps and functions in a fundamentally different way than `clumpfind` and `SExtractor` (Men'shchikov et al., 2012). This change in analysis ideology is based on the fact that sources may take on any number of sizes and shapes based on resolution of data, noise, detailed anisotropic background and multiple source blending (Men'shchikov et al., 2012). `Getsources` functions by detecting sources in maps that have been filtered according to spatial scales (Men'shchikov et al., 2012). This spatial scale filtering can be imagined by examining spatial frequencies of a Fourier transformed map; see Figure 3.3.

The spatial scales explored are defined at the lowest end by the FWHM resolution of the observation and at the highest end by a user defined parameter (Men'shchikov et al., 2012). The extent of a source is identified once a maximum response is discovered at a given spatial scale; see panel 3 in Figure 3.3. Examining sources using spatial filtering intrinsically removes any background contamination and noise simply because they do not exist at those spatial scales (Men'shchikov et al., 2012). Once sources have been detected they are measured on unfiltered maps because the scale filtering process does not conserve flux (Men'shchikov et al., 2012). Any blended sources are deblended according to the source shape found using spatial filtering and individual peak source emission, while dividing shared emission (Men'shchikov et al., 2012). A local background is subtracted from sources based on the type of structure the source may be residing in using a separate algorithm called `getfilaments` which is incorporated into `getsources` (Men'shchikov, 2013). `Getfilaments`

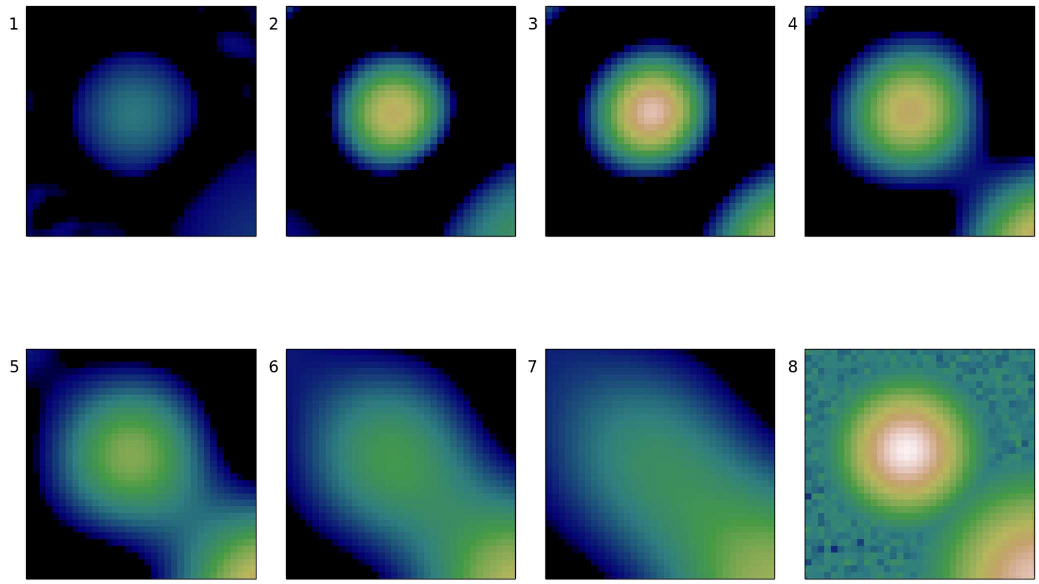

Figure 3.3: Simulated spatial filtering of increasing scale size (panels 1-7) for a Gaussian source (panel 8), spatial scale in panel 3 shows a maximum response among scales; colour scale is normalized among all panels.

improves on large scale background box-type filters by discriminating against not just background noise but also the added, often higher intensity, filamentary structure (Men'shchikov et al., 2012; Men'shchikov, 2013).

Since `getsources` was designed for *Herschel* data it is intrinsically able to handle multiple wavelength observations with changing angular resolution while tracking individual sources through wavebands. The user may specify in how many wavebands a source must be found above a threshold to be considered for reporting (Men'shchikov et al., 2012). Each waveband is filtered separately and then combined by `getsources` as a final step. The finished `getsources` source list reports sources and gives waveband specific parameters such as flux, flux error, major and minor axis length and S/N ratio. `Getsources` currently only works on two dimensional PP maps.

Chapter 4

Data Analysis

4.1 Mass Estimation

There are a few methods available to estimate mass from the PP and PPV data sets of Perseus. Each method requires the total emission contained within the source aperture, or simply the flux S, as well as an estimate of the temperature of the medium.

4.1.1 Dust Emission

The high resolution multi wavelength observations from *Herschel* allow for a Spectral Energy Distribution (SED) to be created of the emitting dust. This process of assembling an SED allows for a direct estimate of mass and temperature.

In order to estimate mass and temperature from an SED one has to manipulate the canonical source function from radiative transfer,

$$S_\nu = \Omega B_\nu(T)[1 - e^{-\tau_\nu}], \tag{4.1}$$

where S_ν is the monochromatic source flux, Ω is the solid angle of the source, $B_\nu(T)$ is the Planck function at a frequency ν and τ_ν is the optical depth (Rybicki & Lightman, 1979; Olmi et al., 2013). Equation 4.1 assumes that the dust emission is not illuminated from behind and that all emission is generated from the dust itself - a fair assumption given that a zero point offset has been removed in the final stages of data reduction. To remove the frequency dependence on τ_ν, a scaling factor is introduced, $(\nu/\nu_0)^\beta$ (Planck Collaboration et al., 2013). The scaling factor reflects the opacity dependence on frequency with a reference frequency ν_0 and a spectral index β (Planck Collaboration et al., 2013). Under the assumption of optically thin emission, $1 \gg \tau$, Equation 4.1 can be simplified by taking the first term of a series expansion of $e^{-\tau_\nu}$,

$$S_\nu = \Omega \tau_{\nu_0} B_\nu(T) \left(\frac{\nu}{\nu_0}\right)^\beta, \tag{4.2}$$

where τ_{ν_0} is the optical depth at reference frequency ν_0. Equation 4.2 is known as an optically thin modified black body.

A fair question to ask is how much difference might assuming optically thin emission make with regards to an SED. Figure 4.1 shows the fractional difference in flux when using an optically thin modified black body (Equation 4.2) over a standard modified black body (Equation 4.1). The dotted horizontal line in Figure 4.1 marks a 10% deviation in flux from a standard black body which is approximately the relative uncertainty in calibration from the *Herschel* PACS/SPIRE data, see Section 4.4. The 70 μm waveband is the most susceptible to problems from assuming optically thin transmission with effects becoming less severe at longer wavelengths. Optical depths of $\tau_{\nu_0} \gtrsim 0.01$ at

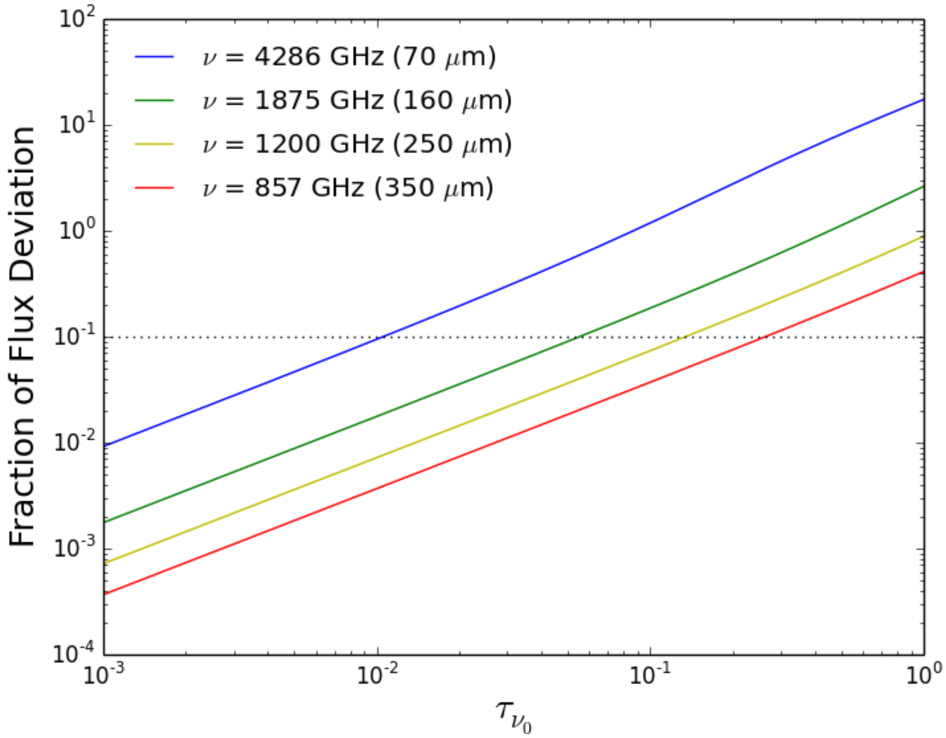

Figure 4.1: Fractional flux difference between an optically thin and standard modified black body; horizontal line marks a 10% change, assuming $\beta = 2.0$, $\nu_0 = 1000$ GHz (300 μm).

70 μm experience deviations of more than 10% whereas the 160 μm waveband experiences these deviations at $\tau_{\nu_0} \gtrsim 0.06$ and 250 - 500 μm wavebands at $\tau_{\nu_0} \gtrsim 0.1$.

Equation 4.2 allows one to fit multi wavelength data to an optically thin modified black body with the free parameters being temperature, T, and optical depth, τ_{ν_0}. In order to include the spectral index β as a free parameter of the fit one needs to have an additional data point in the Rayleigh-Jeans tail ($\gtrsim$ 300 μm) of the SED otherwise a temperature/spectral index degeneracy can inflate the uncertainty of the respective fit parameters (Sadavoy, 2013). Optical depth is related to mass via

$$\tau_{\nu_0} = \frac{M\kappa_{\nu_0}}{\Omega d^2}, \tag{4.3}$$

where κ_{ν_0} is the reference dust emissivity, M is mass, and d is the distance to the source (Olmi et al., 2013). To be consistent with the current paradigm of Gould Belt Survey papers we adopt the values $\kappa_{\nu_0} = 0.1$, $\beta = 2.0$ for reference dust emissivity and spectral index at a reference frequency of $\nu = 1000$ GHz (300 μm) (André et al., 2010). An important fact to note is that a gas-to-dust ratio of 100:1 is assumed and folded into the values for κ_{ν_0} (André et al., 2010). Thus, the mass estimated is the total mass, gas and dust, not just the mass of the dust. Equations 4.2 and 4.3 may be combined to fit mass and temperature directly,

$$S_\nu = \frac{MB_\nu(T)\kappa_{\nu_0}}{d^2}\left(\frac{\nu}{\nu_0}\right)^\beta. \tag{4.4}$$

4.1.2 C^{18}O Emission

Mass estimation using C^{18}O spectral line data requires a different approach to what is used for the *Herschel* data. The procedure can be thought of as a series of steps: first, determine the column density of an optically thin line,

$$\mathrm{N_u} = \frac{4\pi I_{\mathrm{ul}}}{A_{\mathrm{ul}}h\nu_{\mathrm{ul}}}, \tag{4.5}$$

where u and l are the upper and lower transitions (3 and 2), $\mathrm{N_u}$ is column density for the upper level, A_{ul} is the Einstein coefficient of the transition, $h\nu_{\mathrm{ul}}$ is the energy level separation and I_{ul} is the intensity of the transition (Tielens, 2005). A convenient way to write Equation 4.5 with integrated intensity of the spectral line is

$$\mathrm{N}_3 = \frac{8k\pi\nu^2}{hc^3}\frac{1}{A_{3\rightarrow 2}}\int T_{mb}\ dv, \tag{4.6}$$

where T_{mb} is the main beam temperature (Klaassen, 2008). Incorporating a partition function converting $\mathrm{N}_3 \rightarrow \mathrm{N}(\mathrm{C}^{18}\mathrm{O})$ along with the proper Einstein coefficient for the $\mathrm{C}^{18}\mathrm{O}$ $J = 3 \rightarrow 2$ transition from spectral line emission,

$$\mathrm{N}(\mathrm{C}^{18}\mathrm{O}) = 5.14 \times 10^{12}\frac{T_{\mathrm{ex}}}{e^{-(T_{\mathrm{trans}}/T_{\mathrm{ex}})}}\int \frac{T_A^*}{\eta_{mb}}\ dv, \tag{4.7}$$

where $T_{\mathrm{trans}} = 31.6$ K, T_{ex} is the excitation temperature, $T_{mb} = T_A^*/\eta_{mb}$ and T_A^* is antenna temperature from observational data products, η_{mb} is the main beam efficiency and the integral is a summation through velocity channels (Minchin et al., 1993). Next, one needs to use a $\mathrm{C}^{18}\mathrm{O}$ to H_2 abundance ratio to convert $\mathrm{N}(\mathrm{C}^{18}\mathrm{O})$ to $\mathrm{N}(\mathrm{H}_2)$ since $\mathrm{C}^{18}\mathrm{O}$ is only a tracer of the primary constituent of gas, H_2. Recent work in the Taurus Molecular Cloud found an abundance ratio for CO, $\mathrm{CO}/\mathrm{H}_2 = 1.1 \times 10^{-4}$ (Pineda et al., 2010). Using ratios for $\mathrm{CO}/^{13}\mathrm{CO} = 77$ (see Curtis et al. (2010a) for ratio details), along with region specific $^{13}\mathrm{CO}/\mathrm{C}^{18}\mathrm{O}$ ratios found for NGC 1333 (3.7 ± 1.3), IC 348 (4.1 ± 1.8), L1448 (2.8 ± 0.7) and L1455 (3.6 ± 0.9) a $\mathrm{C}^{18}\mathrm{O}$ to H_2 ratio can determined by bootstrapping off of the Pineda et al. (2010) CO/H_2 ratio (Curtis et al., 2010a). Equation 4.7 can be expressed as

$$\mathrm{N}(\mathrm{H}_2) = 3.6 \times 10^{18}\frac{R_{\mathrm{C}^{18}\mathrm{O}}T_{\mathrm{ex}}}{e^{-(31.6\ \mathrm{K}/T_{\mathrm{ex}})}}\int \frac{T_A^*}{\eta_{mb}}\ dv\ [\mathrm{cm}^{-2}], \tag{4.8}$$

where $R_{\mathrm{C}^{18}\mathrm{O}}$ is the region specific $^{13}\mathrm{CO}/\mathrm{C}^{18}\mathrm{O}$ ratio found by Curtis et al. (2010a). The last step to determine mass is to convert the column density using

$$M = \mathrm{N}(\mathrm{H}_2)Am_{\mathrm{H}_2} \tag{4.9}$$

where A is the pixel area (cm^2) and $m_{\mathrm{H_2}}$ is the mass of an H_2 molecule (Minchin et al., 1993).

4.2 SED Pixel Maps

In order to estimate mass from spectral line data one needs a reliable estimate of the excitation temperature, see Equation 4.8. Reliable estimates of environmental dust temperature can be found by fitting the entire Perseus region with an optically thin modified black body. Each waveband specific *Herschel* map has a different intrinsic resolution and pixel size. In order to compute a pixel-by-pixel SED fit each map needs to have the same resolution and pixel size. Ultimately, the final product will have the same resolution as the lowest resolution data, in this case the 500 μm data (see Table 2.1). In order to put all maps to a common resolution one must convolve the maps with a convolution kernel. This work adopts the convolution kernels of Aniano et al. (2011) which consider the point spread function of different instruments along with Gaussian convolution FWHM size. Next, the program `SWarp`[1] is used to re-grid, match pixel size and align maps to the 500 μm map (Bertin et al., 2002). Thus, the 160 μm - 500 μm maps now have a common resolution, FWHM of 36.3$''$ with a pixel size of 14.0$''$ and are ready for pixel-by-pixel SED fitting.

Each pixel in the four (160 μm - 500 μm) maps was fit with Equation 4.2 using a Levenberg - Marquardt algorithm (damped least squares minimization method) solving for two free parameters; temperature and optical depth, (Figures 4.2 and 4.3). Uncertainties on mass and temperature are taken from the square root of the covariance matrix from the Levenberg - Marquardt algorithm. The SED fits assumed a spectral index, $\beta = 2.0$ at reference frequency,

[1]http://www.astromatic.net/software/swarp

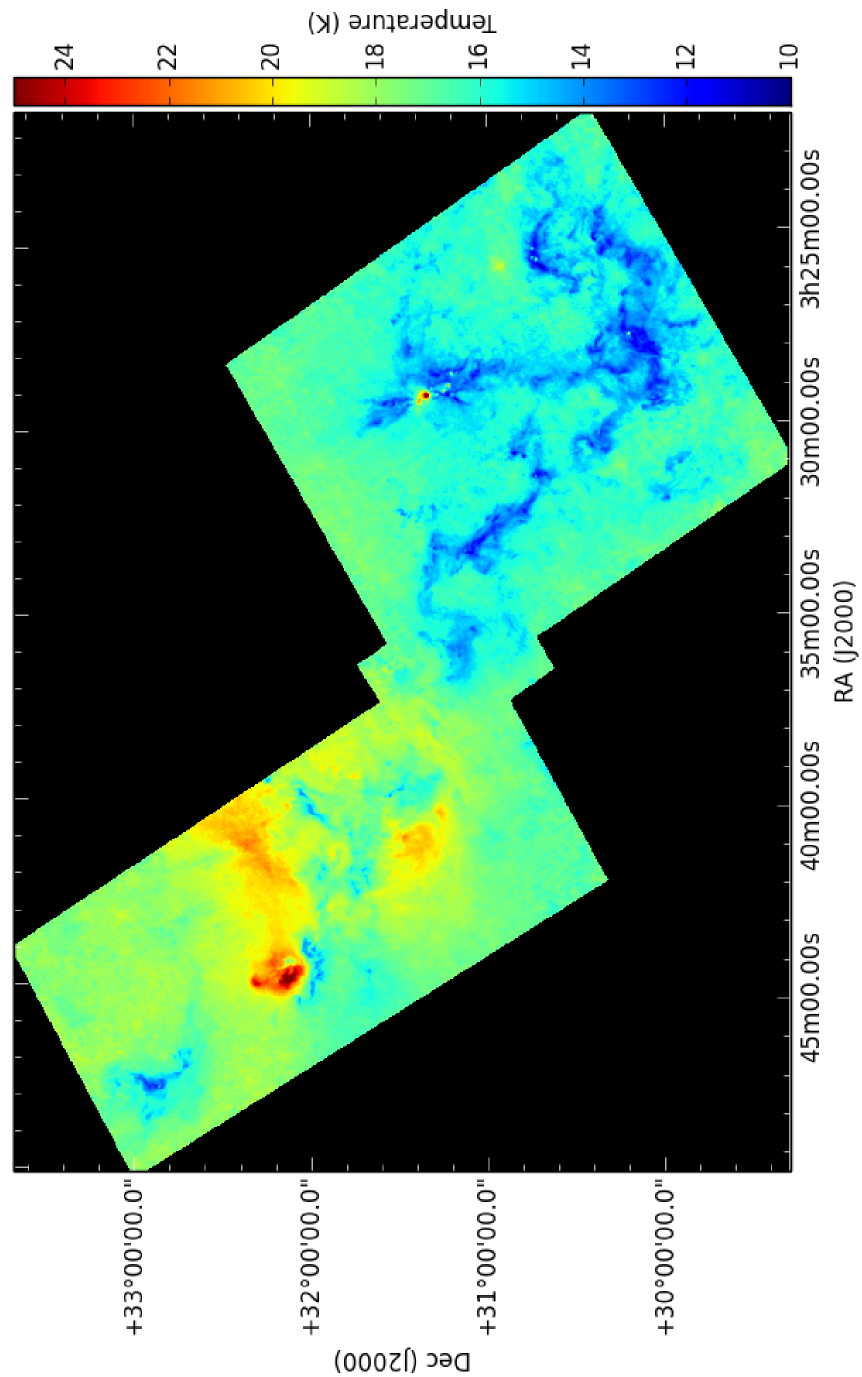

Figure 4.2: 160 μm - 500 μm pixel-by-pixel SED fit temperature map; assumed $\beta = 2.0$, $\nu_0 = 1000$ GHz (300 μm)

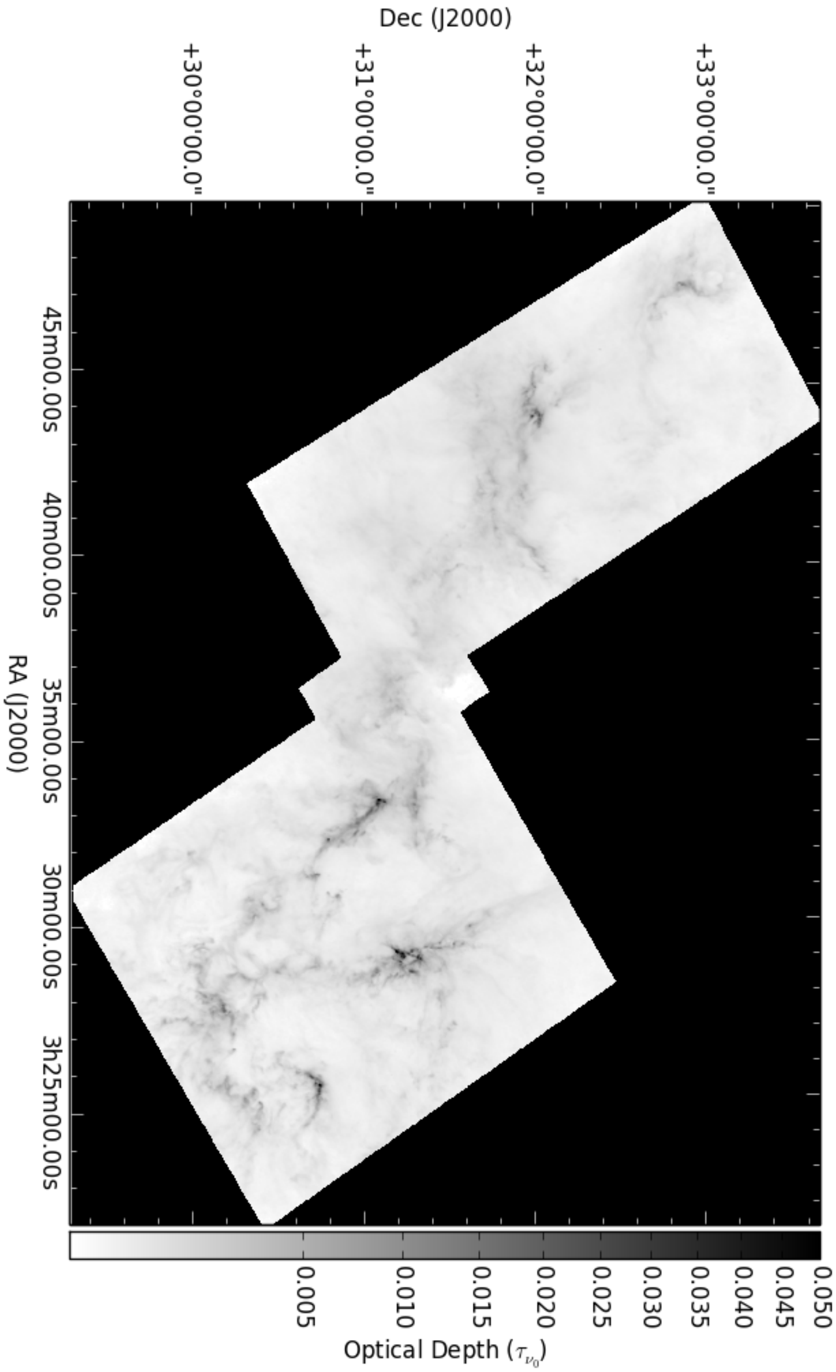

Figure 4.3: 160 μm - 500 μm pixel-by-pixel SED fit optical depth map; assumed $\beta = 2.0$, $\nu_0 = 1000$ GHz (300 μm)

$\nu_0 = 1000$ GHz for consistency with André et al. (2010). The temperature map may be used along with spectral line emission to determine mass for $C^{18}O$ data cubes. It is worth noting that in Figure 4.3 the estimated optical depths do not exceed the values identified in Figure 4.1 indicating that the optically thin modified black body is reliable.

4.3 Source Catalogs

The following section describes the procedures and parameters used with the chosen source finding algorithms. The sources in the produced catalogs are then subject to various selection criteria before they are designated potential cores.

4.3.1 Clumpfind 3D

The `clumpfind 3D` algorithm requires two inputs from the user, the lowest contour level and contour interval, which are suggested to be chosen as integer factors of the root mean square (σ_{rms}) noise of the data cube (Williams et al., 1994). To determine an estimate of the root mean sqaured of the data cube an iterative mean sigma clipping was performed until the $i+1$ mean agreed with the i'th mean to within 0.1%. The iterative sigma clipping procedure is as follows: determine the statistical mean of emission for every pixel in the map, keep pixels that are within $\pm 2\sigma$ of the mean, repeat until convergence. In practice this procedure produced similar σ_{rms} values as one would find simply dropping an aperture at arbitrary spots of low emission in the data. When choosing the lowest contour level the algorithm authors recommend using two or three times the root mean squared, with a contour interval every 2 times the root mean squared (Williams et al., 1994). This method provides a standard template for choosing contour levels that are not so fine as to detect spurious sources but not so coarse that they may miss sources (Williams et al., 1994). For this work the lowest contour level is defined as $3\sigma_{\rm rms}$ with contour interval every $2\sigma_{\rm rms}$(see Table 4.1).

Region	NGC 1333	IC 348	L1448	L1455
$\sigma_{\rm rms}$ (K km s^{-1})	0.19	0.13	0.12	0.12
Lowest contour ($3\sigma_{\rm rms}$) [K km s^{-1}]	0.57	0.39	0.36	0.36
Contour interval ($2\sigma_{\rm rms}$) [K km s^{-1}]	0.38	0.26	0.24	0.24
No. Sources	1710	1410	398	423
No. Filtered Sources	181	173	70	1

Table 4.1: `Clumpfind 3D` $C^{18}O$ data cube parameters.

After performing `clumpfind 3D` on the four regions within Perseus a raw source catalog of 3941 potential sources is created and a three dimensional Gaussian function is fit to each source to characterize spatial and velocity scales only. Most of these potential sources are not cores. The raw catalog is filtered to accept sources which satisfy the following selection criteria:

1. A source must not have emission above the lowest contour level that runs off any of the six sides of the data cube.

2. A source must have at least one pixel above the $7\sigma_{\rm rms}$ contour level.

3. A source FWHM in either spatial dimension must be larger than the beam FWHM, 17.7″ (0.02 pc).

4. A source FWHM in either spatial dimension must be smaller than 263.3″ (0.3 pc).

5. A source FWHM must span a minimum velocity dispersion of 0.12 km s^{-1}.

6. A source must have more than 47 contributing pixels above the lowest contour level.

Condition 1 ensures no flux from the source is missing and thus giving an underestimate of the emission. Condition 2 ensures the source has a statistically significant peak in emission (Hatchell et al., 2007b). Condition 3 and 4 ensure that a source is of appropriate size and not overly elongated in one direction. Condition 5 is a consequence of the contributing factors to velocity dispersion in a source,

$$\sigma^2_{C^{18}O} = \sigma^2_{\mathrm{T}} + \sigma^2_{\mathrm{NT}} \tag{4.10}$$

where $\sigma^2_{C^{18}O}$ is the observed velocity dispersion, σ^2_{T} is the velocity dispersion from thermal motion and σ^2_{NT} is non thermal velocity dispersion (Curtis & Richer, 2011). The velocity dispersion from non thermal motion is expected to be non zero since the progenitor dense knots that collapse into cores are observed to exist in environments with supersonic turbulence (Larson, 1981). The velocity dispersion from thermal motion is,

$$\sigma^2_{\mathrm{T}} = \frac{kT}{m_{\mathrm{C^{18}O}} m_p}, \tag{4.11}$$

where k is the Boltzmann constant, T is temperature, $m_{\mathrm{C^{18}O}}$, m_p is the mass of a proton and $m_{\mathrm{C^{18}O}}$ is the relative molecular mass ($m_{\mathrm{C^{18}O}} = 30$) (Curtis & Richer, 2011). Assuming a lower estimate for temperature in dense cores of 10 K, the minimum thermal velocity dispersion would be $\sigma^2_{\mathrm{T}} = 2.75 \times 10^{-3}$ km^2 s^{-2} or simply $\mathrm{FWHM_T} = 0.12$ km s^{-1}. Condition 6 is a combination of conditions 3 and 5 and represents the minimum number of pixels a source must have to satisfy the above conditions. The selection criteria allow for reliable minimum detection of 0.02 $M_\odot$ based on an assumed lower estimate of excitation temperature of 10 K and a $^{13}CO/C^{18}O$ ratio of 2.8. A total of 425 sources passed the selection criteria with condition 3 filtering out the most candidate sources. Table 4.1 shows the number of catalog sources before and after filtering for each region.

4.3.2 Clumpfind 2D

Similar to `clumpfind 3D`, the `clumpfind 2D` algorithm needs two inputs from the user, the lowest contour and contour interval (Williams et al., 1994). The method to determine the map $\sigma_{\rm rms}$ is the same used for `clumpfind 3D` algorithm, iterative mean sigma clipping. Lastly, $3\sigma_{\rm rms}$ is used for the lowest contour level with contour intervals every $2\sigma_{\rm rms}$, see Table 4.2. Unlike the pixel-by-pixel SED maps (Figures 4.2 and 4.3) `clumpfind 2D` is performed on the original *Herschel* maps.

<table>
<tr><th></th><th>160 μm</th><th>250 μm</th><th>350 μm</th><th>500 μm</th></tr>
<tr><td>$\sigma_{\rm rms}$ (Jy)</td><td>0.025</td><td>0.043</td><td>0.062</td><td>0.055</td></tr>
<tr><td>Lowest contour ($3\sigma_{\rm rms}$) [Jy]</td><td>0.075</td><td>0.129</td><td>0.186</td><td>0.165</td></tr>
<tr><td>Contour interval ($2\sigma_{\rm rms}$) [Jy]</td><td>0.050</td><td>0.086</td><td>0.124</td><td>0.110</td></tr>
<tr><td>No. Sources</td><td>614</td><td>388</td><td>298</td><td>227</td></tr>
<tr><td>No. Filtered Sources</td><td colspan="4">54</td></tr>
</table>

Table 4.2: `Clumpfind 2D` *Herschel* map parameters.

After performing `clumpfind 2D` on the four *Herschel* wavebands, 160 μm - 500 μm, for the entire Perseus map a raw source catalog of 1527 potential sources is created. The raw catalog is filtered to accept sources which satisfy selection criteria:

1. A source must not have emission above the lowest contour level that runs off any of the four sides of the map.

2. A source must be present in all four wavebands with at least one pixel above the $5\sigma_{\rm rms}$ contour level in each waveband.

3. A source must have at least 11 contributing pixels above the lowest contour level in the 160, 250 μm waveband and at least 8 contributing pixels above the lowest contour level in the 350, 500 μm waveband.

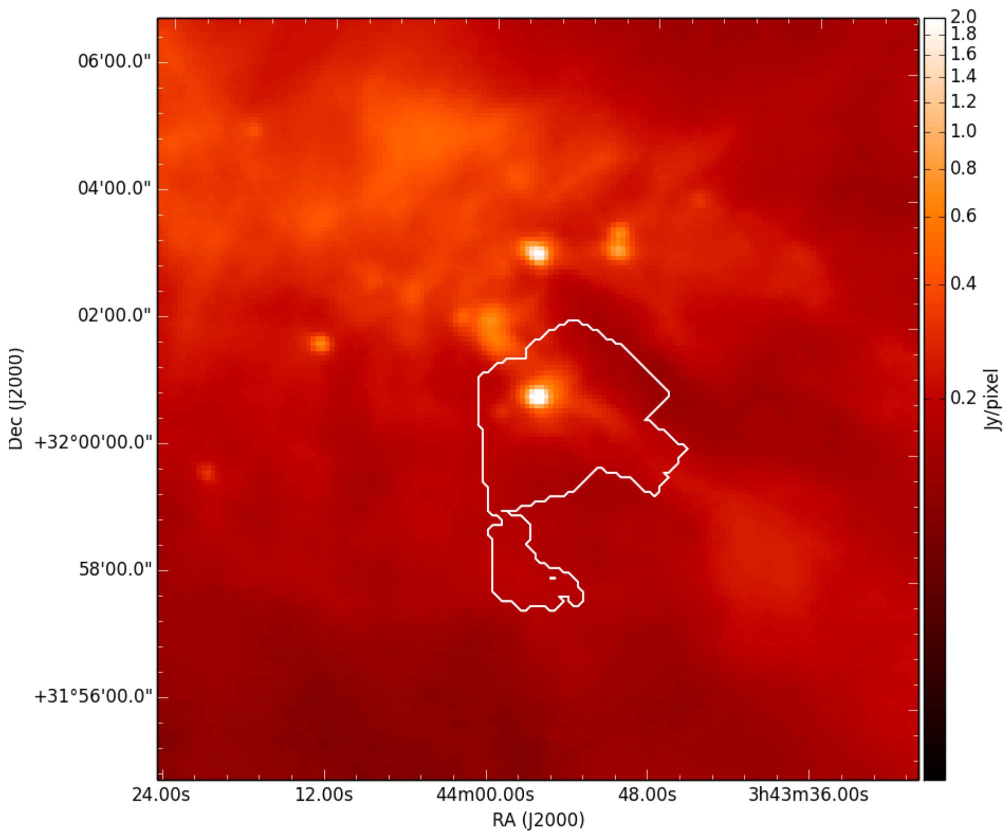

Figure 4.4: Zoomed portion of IC 348 (160 μm *Herschel* map) within Perseus; the white contour shows all pixels partitioned to the central source using `clumpfind 2D`.

Condition 1 ensures no flux from the source is missing thus giving an underestimate of the emission. Condition 2 ensures the source has a statistically significant peak in emission in all wavebands (Hatchell et al., 2007b; Curtis & Richer, 2010). The $5\sigma_{\rm rms}$ significance was chosen instead of the $7\sigma_{\rm rms}$ level used in `clumpfind 3D` because of the requirement that a source be present in all wavebands. Condition 3 ensures that any source reported has the minimum number of pixels required to fill the beam FWHM, a minimum resolution element. Unlike `clumpfind 3D` there is no Gaussian fit to each source in the `clumpfind 2D` catalog. The `clumpfind 2D` algorithm tends to partition low contour level pixels to a source, especially if the source lies on a elongated filamentary structure (see Figure 4.4). This behavior also prevents a source check of maximum size which is normally performed. The selection criteria allow for reliable minimum detection of 0.4 $M_\odot$ based on a lower temperature estimate of 10 K. A total of 54 sources passed the selection criteria with condition 2 filtering out the most candidate sources (see Figure 4.5).

4.3.3 SExtractor

The source finding algorithm `SExtractor` requires a few inputs from the user. Since the algorithm subtracts a background the user must specify the background mesh size. The mesh size is simply how much area to consider when computing a background from the data. In all four *Herschel* wavebands a background mesh size of $263.3'' \times 263.3''$ (0.3×0.3 pc) was chosen since that is the upper limit on source size. The second and third parameters that must be supplied are both related to detection. A detection threshold, similar to the lowest contour level in `clumpfind`, is specified as a factor multiplied by the derived $\sigma_{\rm rms}$. As well, a minimum number of pixels above detection threshold must be supplied for a source to be considered (see Table 4.3). The detection

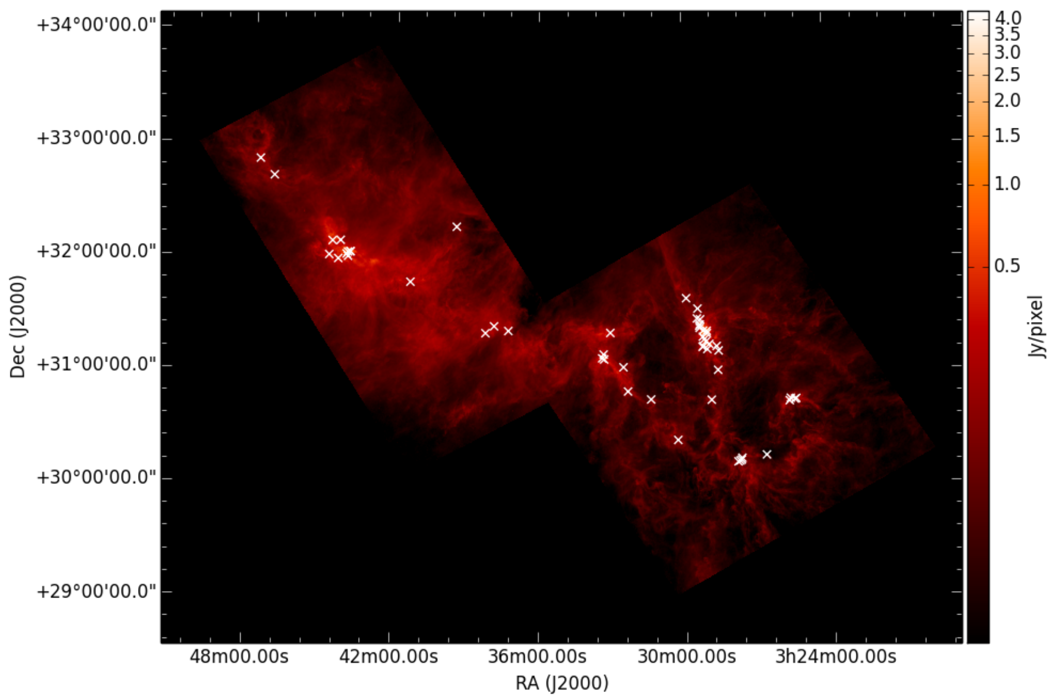

Figure 4.5: Locations of the 54 cores that pass the selection criteria using `clumpfind 2D`.

	160 μm	250 μm	350 μm	500 μm
Background Mesh (pixel×pixel)	57×57	43×43	25×25	19×19
Detection Threshold (σ_{rms})	3.0			
Detection Min. Area (pixels)	11	11	8	8
No. Sources	1125	998	752	511
No. Filtered Sources	88			

Table 4.3: `SExtractor` *Herschel* map parameters.

parameters chosen are consistent with what is chosen for the `clumpfind 2D` algorithm. Like `clumpfind 2D`, `SExtractor` is performed on original *Herschel* maps.

After performing `SExtractor` on the four *Herschel* wavebands, 160 μm - 500 μm, for the entire Perseus map a raw source catalog of 3386 potential sources is created. The raw catalog is filtered to accept sources which satisfy selection criteria:

1. A source must not have emission above the lowest contour level that runs off any of the four sides of the map.

2. A source must be present in all four wavebands with at least one pixel above the $5\sigma_{\text{rms}}$ contour level in each waveband.

3. A source must have no more than: 3879 contributing pixels in the 160 μm band; 2182 contributing pixels in the 250 μm band; 786 contributing pixels in the 350 μm band; and 401 contributing pixels in the 500 μm band.

Condition 1 ensures no flux from the source is missing thus giving an underestimate of the emission. Condition 2 ensures the source has a statistically significant peak in emission in all wavebands (Hatchell et al., 2007b; Curtis & Richer, 2010). There is no condition for minimum size because that is pro-

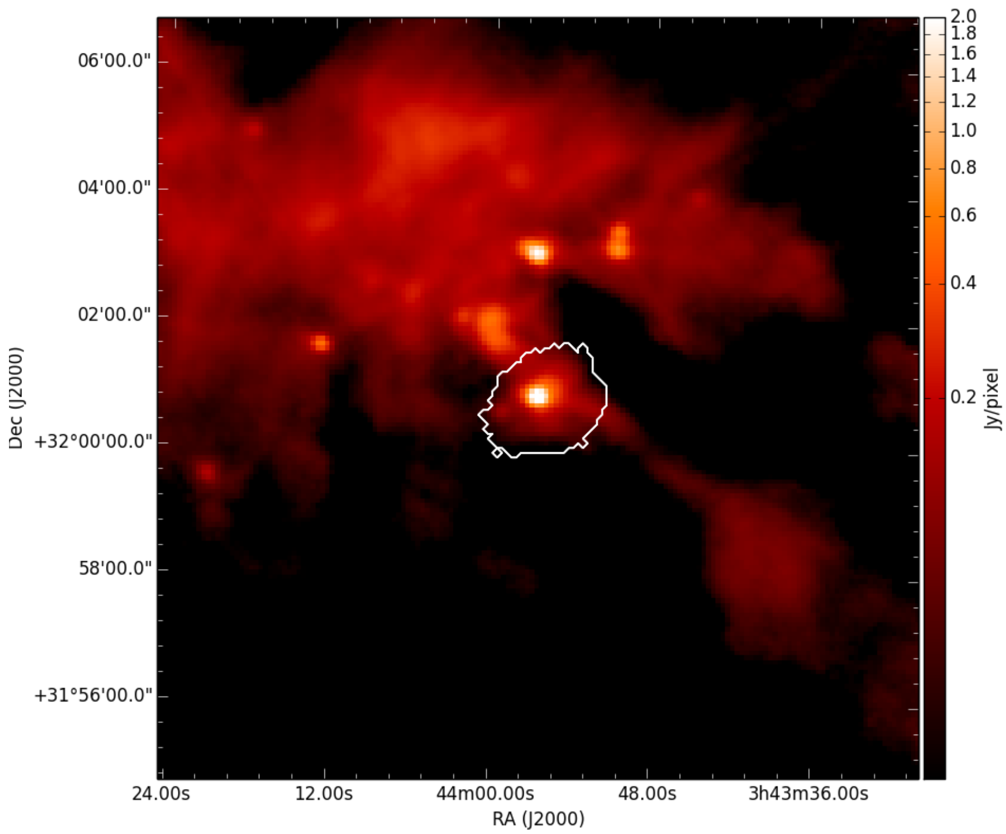

Figure 4.6: Zoomed portion of IC 348 (160 μm *Herschel* map) within Perseus; the white contour shows all pixels partitioned to the central source using `SExtractor`.

vided as a user parameter prior to running `SExtractor`. However, condition 3 ensures that no source has an area greater than 263.3″ × 263.3″ (0.3 × 0.3 pc). In practice the background subtraction in `SExtractor` was more efficient at eliminating lower contour run-off than `clumpfind 2D` (compare Figure 4.4 and 4.6). The selection criteria allow for reliable minimum detection of 0.02 $M_\odot$ based on a lower temperature estimate of 10 K. A total of 88 sources passed the selection criteria with condition 2 filtering out the most candidate sources (see Figure 4.7).

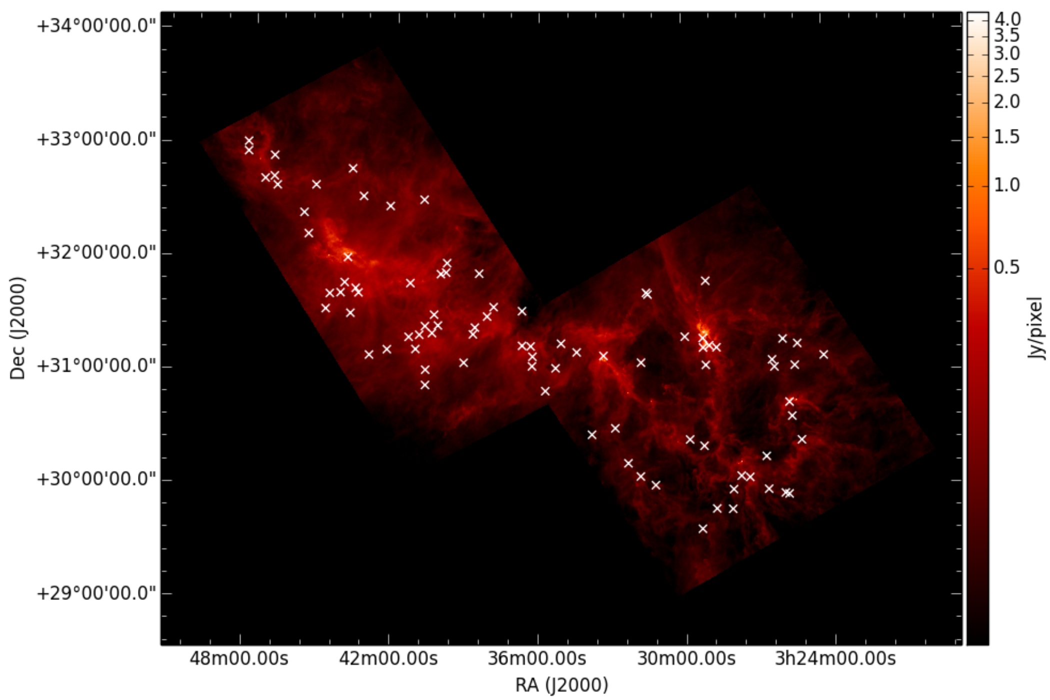

Figure 4.7: Locations of the 88 cores that pass the selection criteria using `SExtractor`.

4.3.4 Getsources

Since `getsources` was designed for *Herschel* data products the process of source detection is highly automated. The user needs to supply the appropriate data; *Herschel* 160 μm - 500 μm maps, the default configuration is optimal for detection. `Getsources` produces a source catalog where sources are already spatially correlated among wavebands and have at least a minimum size governed by the beam FWHM in each waveband. The raw `getsources` catalog contained 1705 potential sources. The raw catalog is filtered to accept sources which satisfy selection criteria:

1. A source must not have emission that runs off any of the four sides of the map.
2. A source must be present in all four wavebands with a S/N of at least 7.0 in each waveband.
3. A source FWHM in one of the two spatial dimensions must be smaller than 263.3$''$ (0.3 pc).
4. A source must not be sub-structured by a smaller source.

Condition 1 ensures no flux from the source is missing thus giving an underestimate of the emission. Condition 2 ensures that all sources have reliable detection. The `getsources` catalog identifies sources with a S/N of 7.0 as reliable; sources with a S/N <7.0 are deemed tentative and are not included. Condition 3 ensures that no source is larger than the maximum size or overly elongated. Condition 4 is unique to the `getsources` algorithm as it prevents substructured detections that may be misidentified as sources but are actually larger plateau background emission that hosts a smaller core or cores within its borders. Due to the construction of `clumpfind` and `SExtractor` misidentifying sources in this manner is not a problem. Additionally, `getsources` identifies sources confined to a much smaller area than `clumpfind` and `SExtractor` (compare Figures 4.8 to 4.4 and 4.6 for example). The selection criteria allow for reliable minimum detection of 0.05 $M_\odot$ based on a lower temperature estimate of 10 K. A total of 284 sources passed the selection criteria with condition

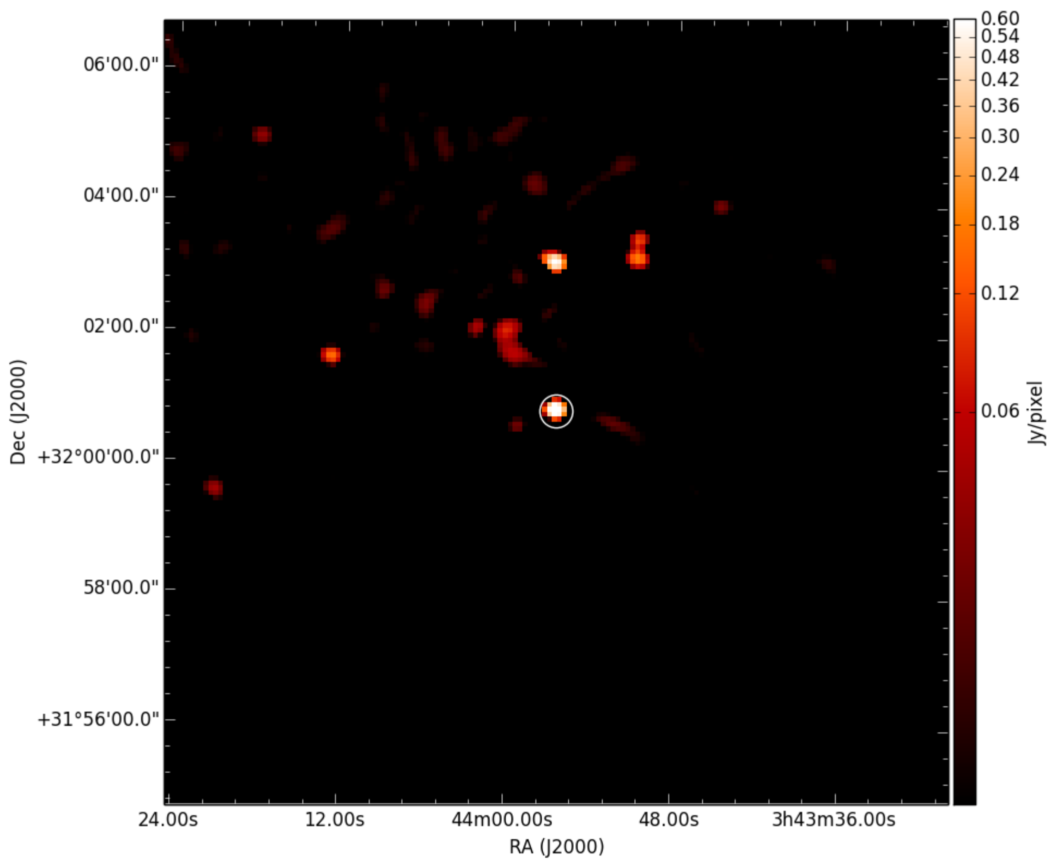

Figure 4.8: Zoomed portion of IC 348 (160 μm *Herschel* map) within Perseus; the white contour shows a 2σ region of the central source using a detection image from `getsources`.

2 filtering out the most candidate sources (see Figure 4.9).

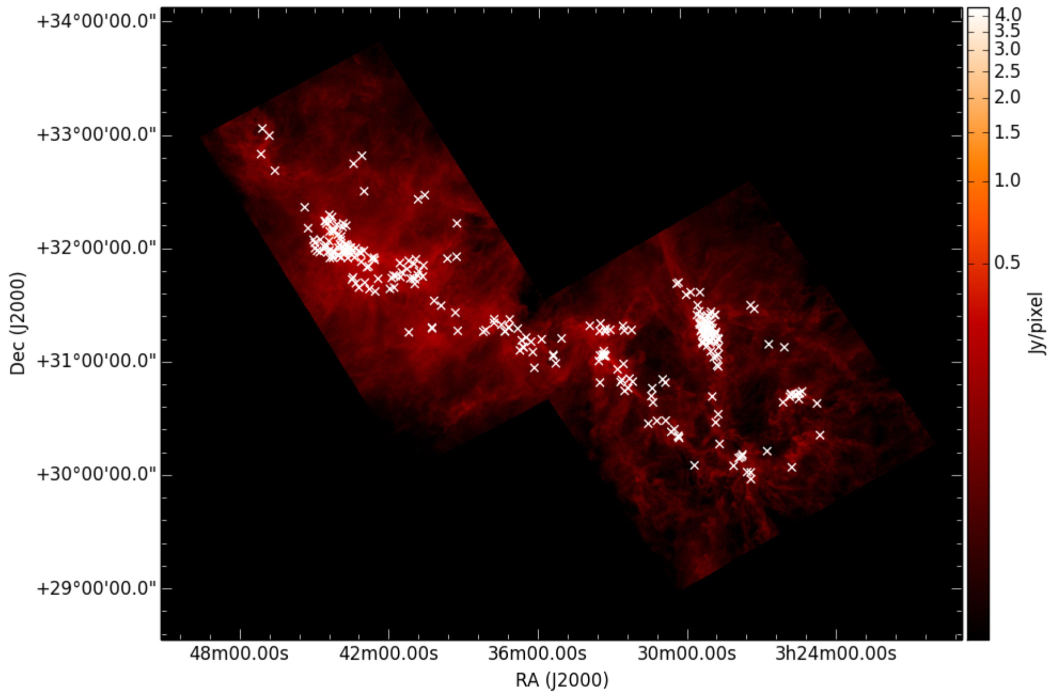

Figure 4.9: Locations of the 284 cores that pass the selection criteria using `getsources`.

4.4 Source Mass

Every source in each catalog that passed the selection criteria discussed in the previous section had a mass and temperature estimated for it. Masses for sources in the `clumpfind 3D` catalog were estimated using Equations 4.8 and 4.9. Excitation temperature, $T_{\rm ex}$, was estimated from the pixel-by-pixel SED fit map of Perseus (Figure 4.2) using the location of the source on that map averaged from a beam sized aperture. The uncertainty on the estimate of mass was calculated using the uncertainty of the antenna temperature, T_A^*, as the single velocity channel root mean squared noise, the uncertainty on the excitation temperature found from the pixel-by-pixel SED fit map and the uncertainty in the $R_{\mathrm{C^{18}O}}$ ratio provided by Curtis et al. (2010a). The cumulative mass function of the `clumpfind 3D` source masses can be found in the top left panel of Figure 4.10; the estimated temperatures from the pixel-by-pixel SED fit maps are shown in Figure 4.11.

	Colour Correction Factor ($\times$ S_ν)	Sensitivity Error (Jy)	Calibration Error (% $\times$ S_ν)
160 μm	1.01	0.0685	5.0
250 μm	0.986	0.0017	7.0
350 μm	0.989	0.0014	7.0
500 μm	1.001	0.002	7.0

Table 4.4: *Herschel* multi band colour correction and error estimation (Poglitsch et al., 2010; Griffin et al., 2010)

Sources in the three *Herschel* catalogs were fit with an optically thin modified black body, Equation 4.4, which estimated mass and temperature as parameters of the fit. Each flux measurement in the four wavebands was subject to a colour correction and the error for each flux measurement was summed in quadrature from Table 4.4. In addition to the two sources of error listed in Table 4.4 a third source of error, measurement error, was incorporated. The measurement error used for the `clumpfind 2D` catalog was the root mean squared of the *Herschel* maps (Table 4.2), while the measurement errors for the `SExtractor` and `getsources` catalog were estimated with flux measurements by the algorithm since σ_{rms} varied source-by-source (Bertin & Arnouts, 1996; Men'shchikov et al., 2012). Each source was fit using a Levenberg - Marquardt algorithm (damped least squares minimization method) solving for mass and temperature. The cumulative mass functions for the `clumpfind 2D`, `SExtractor` and `getsources` can be found in Figure 4.10, along with temperatures in Figure 4.11.

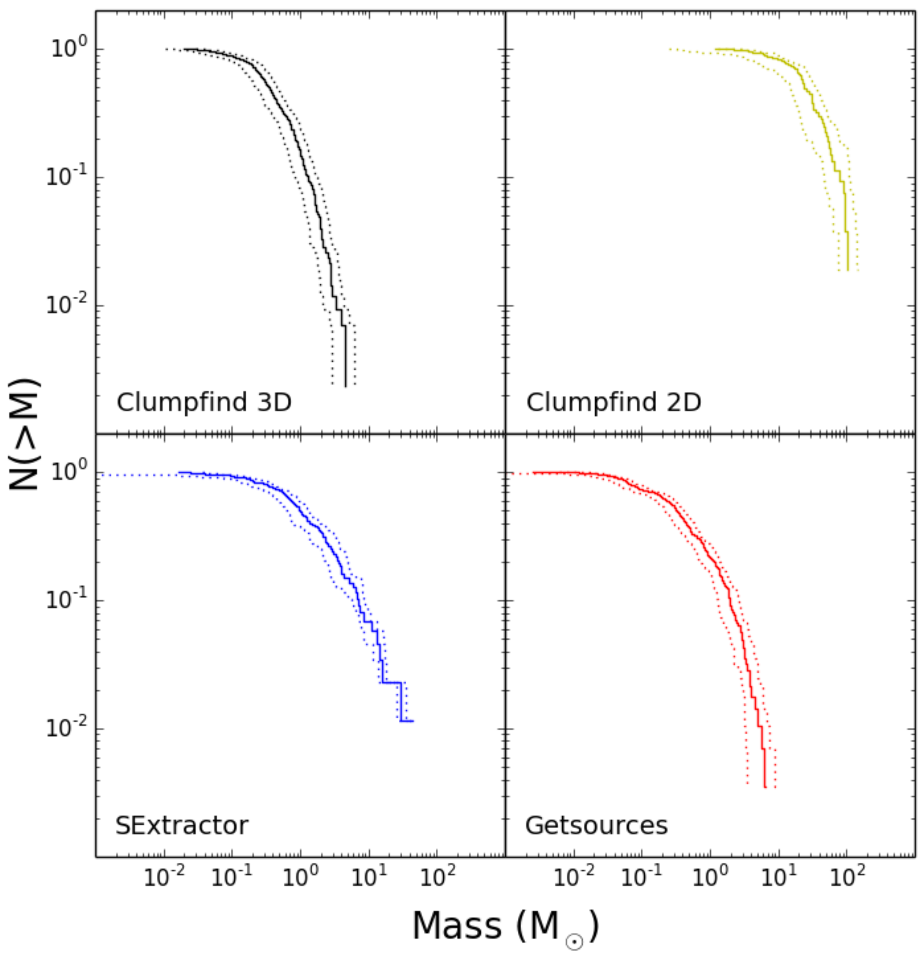

Figure 4.10: Cumulative mass function (solid) with $\pm 1\sigma$ (dotted) for; *top left* `clumpfind 3D`, *top right* `clumpfind 2D`, *bottom left* `SExtractor` and *bottom right* `getsources`.

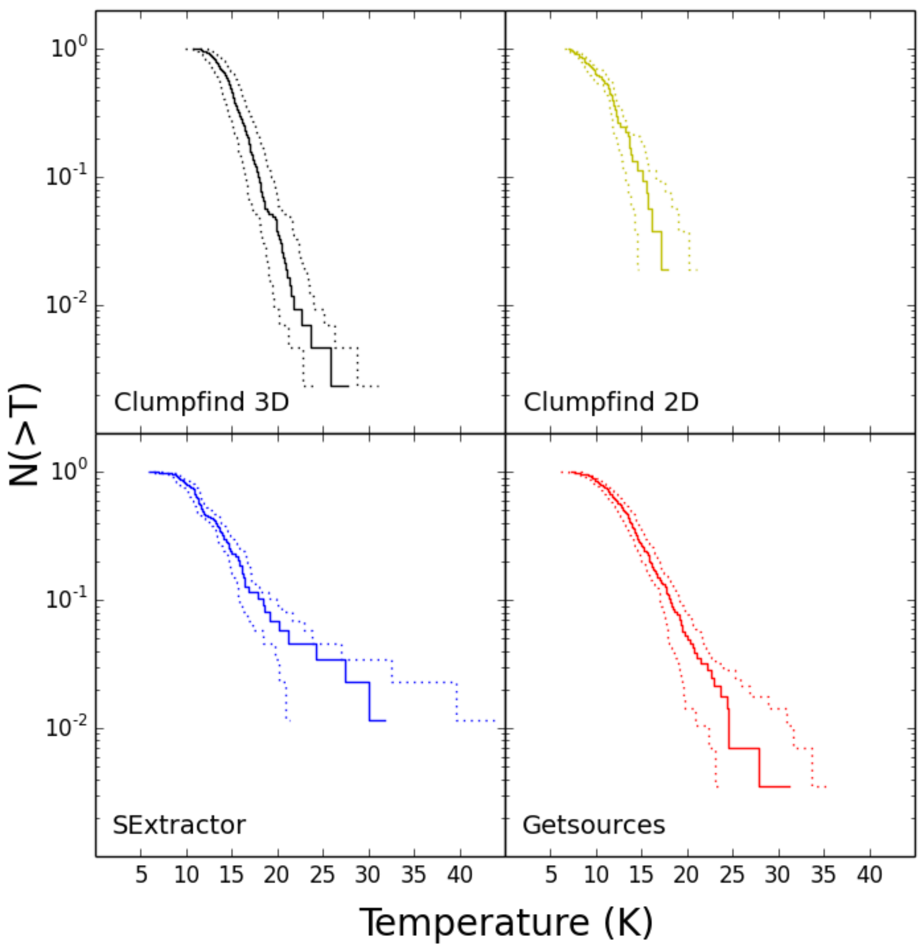

Figure 4.11: Cumulative temperature function (solid) with $\pm 1\sigma$ (dotted) for; *top left* `clumpfind 3D`, *top right* `clumpfind 2D`, *bottom left* `SExtractor` and *bottom right* `getsources`.

Chapter 5

Results

5.1 Mass Comparison

5.1.1 Core Mass Distributions

The CMFs for each catalog are fit with two functional forms. The first functional form is a double power law fit,

$$N(>M) = \begin{cases} AM_{\text{break}}^{(x_{\text{low}}-x_{\text{high}})} M^{-(x_{\text{low}}+1)}, & M < M_{\text{break}} \\ AM^{-(x_{\text{high}}+1)}, & M \geq M_{\text{break}} \end{cases} \tag{5.1}$$

where x_{low} and x_{high} are the low and high mass power law slopes respectively (Reid & Wilson, 2006). The second functional form is a lognormal fit,

$$N(>M) = \frac{1}{2}\left[1 - \text{erf}\left(\frac{\ln(M) - \ln(\mu)}{\sqrt{2}\sigma}\right)\right], \tag{5.2}$$

where μ and σ are the mean and standard deviation of the fit (Reid & Wilson, 2006). The mean and standard deviation of the lognormal fit can be thought as complimentary to break mass and high mass power law slope. As break mass increases so does the mean of a lognormal fit. Likewise, shallower slopes are indicative of larger standard deviations in a lognormal fit. While a direct comparison between parameters of different functional forms is impossible they may be used for relative comparisons between catalogs using the same functional form. These two functional forms are chosen based on the work by Reid & Wilson (2006) who compared a collection of functional fits to twelve observed CMFs and found Equations 5.1 and 5.2 to be the best performing based on χ^2 and p-value tests. A benefit for using Equation 5.1 is that the estimated high mass power law slope can be compared to contemporary studies of Perseus.

Each mass distribution was fit using a Levenberg-Marquardt algorithm. Additionally, the stellar IMF is randomly sampled from the probability distribution function (PDF) given by Chabrier (2005) using rejection sampling,

$$\xi = \begin{cases} 0.093 \exp\left[\frac{-(\mathrm{Log}(M)-\mathrm{Log}(0.2))^2}{2(0.55)^2}\right], & 0.08M_\odot \leq M < 1.0M_\odot \\ 0.043M^{-1.35}, & M \geq 1.0M_\odot. \end{cases} \tag{5.3}$$

Rejection sampling is a Monte Carlo (MC) method of populating a distribution by uniformly sampling a parameter space that fully encompasses a custom PDF, (see Press et al. (2002)). A randomly drawn number is checked to see if it lies within the defined PDF. If it fails the number is rejected, if it passes it is kept (Press et al., 2002). The process repeats until the accepted sample size is reached. By construction, a CMF has uncertainty on the independent axis, M, instead of the dependent axis, $N(>M)$, which makes uncertainty analysis difficult. Instead of using the covariance matrix from the Levenberg-Marquardt algorithm to estimate parameter uncertainty a Monte Carlo (MC) sampling approach is used. The IMF (Equation 5.3) has no uncertainty on mass so

	`Clumpfind 3D`	`Clumpfind 2D`	`SExtractor`	`Getsources`	IMF
N	425	54	88	284	
2PL					
α_{low}	-1.23 ± 0.03	-1.16 ± 0.05	-1.13 ± 0.03	-1.16 ± 0.03	-1.49 ± 0.04
α_{high}	-2.21 ± 0.07	-2.4 ± 0.2	-1.77 ± 0.06	-1.97 ± 0.06	-2.2 ± 0.1
M_{break} $(M_\odot)$	0.31 ± 0.02	22.1 ± 1.2	0.7 ± 0.1	0.33 ± 0.03	0.33 ± 0.04
χ^2	2.0	0.3	0.4	1.6	0.4
LOGN					
μ $(M_\odot)$	0.33 ± 0.01	23.0 ± 2.0	1.00 ± 0.06	0.29 ± 0.01	0.31 ± 0.02
σ	1.09 ± 0.03	1.00 ± 0.09	1.52 ± 0.07	1.57 ± 0.03	0.95 ± 0.04
χ^2	0.2	0.2	0.09	0.6	1.1

Table 5.1: CMF fit parameters for double power law (2PL) and lognormal (LOGN) functional forms.

each MC iteration samples new points from the PDF. Each core had a mass sampled from a Gaussian distribution centered on the original mass estimate with a standard deviation taken from the mass estimate uncertainty. In total 10^3 fits were made with each functional form that produced a distribution of fit parameters. Final fit parameters were estimated from the mean and standard deviation of the respective randomly drawn distribution. Table 5.1 contains the estimated parameters along with uncertainties and Figure 5.1 shows the graphical fits for both functional forms.

In all four catalogs the lognormal functional form fit had a lower calculated χ^2 than the double power law fit. However, the IMF had a lower χ^2 with a double power law fit which is unsurprising given the fact that the PDF itself is partly a power law (Equation 5.3). The high mass power law slope, $\alpha_{\rm high}$, of both `clumpfind` algorithms captures $\alpha_{\rm high}$ of the IMF within uncertainty. However, $\alpha_{\rm high}$ derived from `getsources` is shallower than the IMF and the slope of `SExtractor` is shallower still. Break mass determined from `clumpfind 3D` and `getsources` are equal to each other, and to the IMF, within uncertainty. Break mass for `clumpfind 2D` greatly overshoots the IMF and `clumpfind 3D` break mass. Lastly, `SExtractor` obtains a break mass not much higher than the IMF, `clumpfind 3D` and `getsources`. Contemporary studies generally find high mass power law slopes greater than -2.35 with break mass 0.8 - 1.0 $M_\odot$ (Kirk et al., 2006; Curtis & Richer, 2010). An interesting effect to note regarding the double power law is the fact that the break mass is lower than 1$M_\odot$ for the IMF despite being explicitly defined in Equation 5.3. The behaviour may be explained by the need to fit the high and low mass power law slope together. Regarding the lognormal functional form, the mean of the `getsources` catalog captures the IMF mean with the `clumpfind 3D` mean equal within uncertainty. Similar to break mass, the means of `SExtractor` and `clumpfind 2D` miss the mark by an increasing margin. Additionally, like

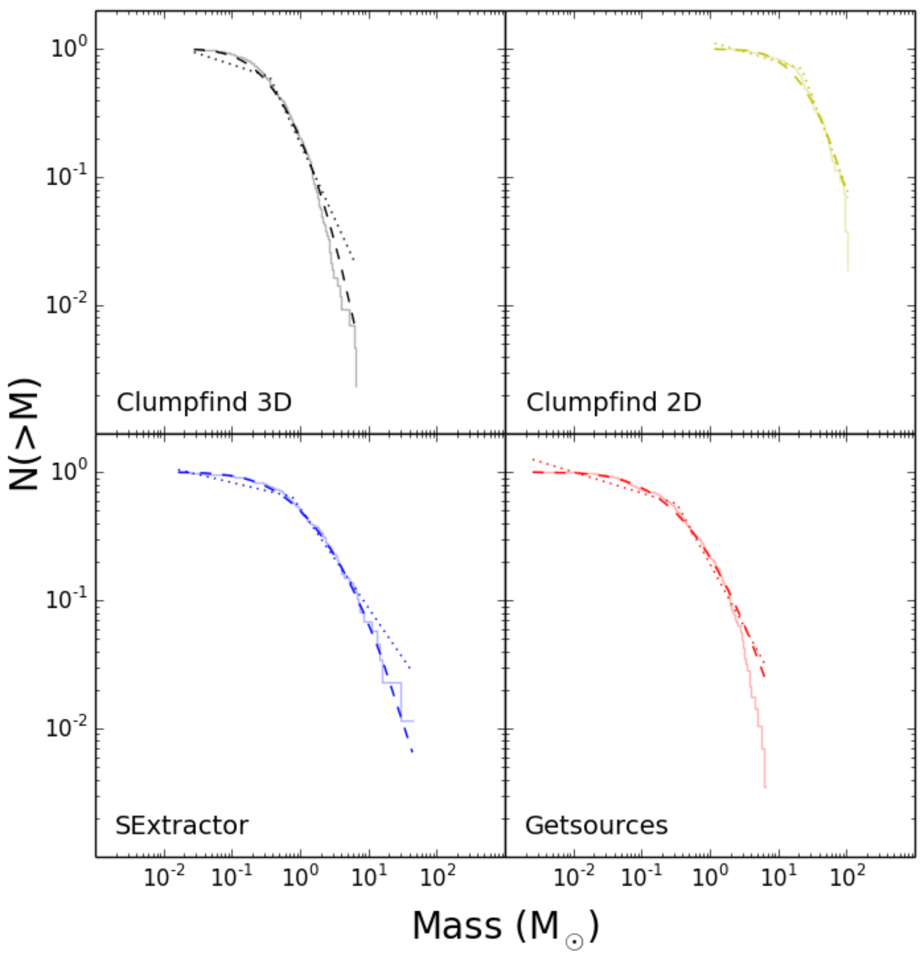

Figure 5.1: Cumulative mass function (faint solid) with double power law (dotted) and lognormal (dashed) fits; *top left:* `clumpfind 3D`, *top right:* `clumpfind 2D`, *bottom left:* `SExtractor` and *bottom right:* `getsources`.

α_{high} the standard deviation of the IMF and `clumpfind 2D` are equal within uncertainty with `clumpfind 3D` just missing out. The shallower high mass power law slopes of `SExtractor` and `getsources` are reflected in larger values of standard deviation.

Based on the results of the functional form fits the cores from the PPV data, found using `clumpfind 3D`, are the closest among the four catalogs to match the IMF. This result was predicted by Ward et al. (2012) who concluded from synthetic observations that core masses determined from spectral line data more closely resemble true mass than masses found from continuum maps which tend to overestimate. Of the algorithms run on the *Herschel* continuum maps `getsources` matches `clumpfind 3D` and the IMF closer than `SExtractor` and `clumpfind 2D`. Interestingly, despite the relatively small sample size of the `clumpfind 2D` catalog it consistently matched the width parameter (α_{high} and σ) of the IMF. However, the entire `clumpfind 2D` distribution was pushed to much higher masses as evident in the break mass and mean and is plainly visible in Figure 5.1.

If the algorithms performed on the continuum maps were to be ranked based on location parameter (M_{break} and μ) they would also correspond to how much background is subtracted and how tightly the source region is defined.

Figure 5.2 shows a cross section at constant DEC of a core candidate in *Herschel* continuum maps found using `clumpfind 2D` (yellow line) and `SExtractor` (blue line) with the black dashed line marking the $3\sigma_{\mathrm{rms}}$ level. This particular candidate is in the B1 region of Perseus, see Pezzuto et al. (2012); Sadavoy et al. (2012) for specific studies on the B1 region. The `clumpfind 2D` algorithm does not subtract a background and low level emission bleeds off the source cross section above the $3\sigma_{\mathrm{rms}}$ level. Modest background subtraction using a median/mode hybrid filter is employed in `SExtractor` which reduces

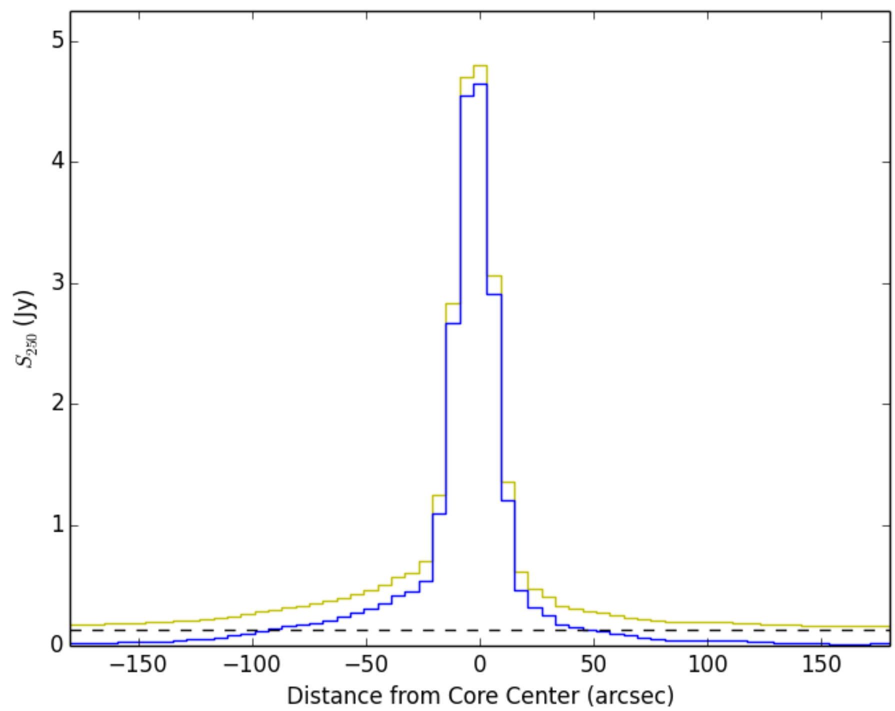

Figure 5.2: Cross section at constant DEC of core (RA: 3:33:17.7, DEC: +31:09:32.1) using `clumpfind 2D` (yellow) and `SExtractor` (blue) with dashed line showing $3\sigma_{\rm rms}$.

the low level emission but some emission runoff is still present. This low level runoff may be understood if the central core is immersed in an envelope of material as suggested by Pezzuto et al. (2012) who analyzed a handful of cores in the B1 region using the same *Herschel* data. The spatial scale analysis performed by `getsources` suggests a FWHM of 18.2″for this source and seems to cut the low level emission off completely capturing just the prestellar core and none of the envelope. The estimated mass for this core using SED fitting is $48.2 \pm 5.2\ M_{\odot}$ for `clumpfind 2D`, $15.7 \pm 1.7\ M_{\odot}$ for `SExtractor` and $1.6 \pm 0.5\ M_{\odot}$ for `getsources`. The difference in estimated masses is significant, a factor of ~ 30 for `clumpfind 2D` over `getsources` and ~ 10 for `SExtractor` over `getsources`.

Kolmogorov-Smirnov Test

A third independent test can be performed to investigate if the algorithms are characterising cores similarly. The Kolmogorov-Smirnov (K-S) test is a non-parametric test to investigate if two discrete populations are drawn from the same underlying distribution (Press et al., 2002). The K-S test accomplishes this by creating a normalized cumulative distribution (not to be confused with the CMF) from catalog masses and calculates the K-S statistic, D, which is the point at which the two distributions are at their most different,

$$D = \max \mid S_1(x) - S_2(x) \mid, \ -\infty < x < \infty, \tag{5.4}$$

where S_1 and S_2 are the normalized cumulative distributions (Press et al., 2002). The K-S statistic alone does not provide a probability that the two populations are drawn from the same underlying distribution, but only the fractional distance between cumulative populations. A two-sided p-value, p_v, can provide an estimate of the level of rejection of the hypothesis that both populations are drawn from the same distribution,

$$p_v(z) = 2\sum_{j=1}^{\infty}(-1)^{j-1}e^{-2j^2z^2},$$

where z is found by,

$$z = D\left(N_e + 0.12 + \frac{0.11}{N_e}\right), \; N_e = \sqrt{\frac{N_1N_2}{N_1+N_2}},$$

and where N_1 and N_2 are the number of samples in the respective populations (Press et al., 2002). Simply, the p-value states the probability of randomly obtaining the K-S statistic having drawn two populations from the same underlying distribution.

Data Set 1	Data Set 2	D
IMF	`Clumpfind 3D`	0.108
`Clumpfind 3D`	`Getsources`	0.157
IMF	`Getsources`	0.225
`Clumpfind 3D`	`SExtractor`	0.382
IMF	`SExtractor`	0.440
IMF	`Clumpfind 2D`	0.953
`Clumpfind 3D`	`Clumpfind 2D`	0.970

Table 5.2: K-S test results on estimated mass catalogs and IMF, note $p_v = 0$ for all cases.

Table 5.2 shows the results for catalog comparison between the IMF/`clumpfind 3D` and other catalogs. Additionally, Figure 5.3 shows the K-S tests with the four lowest K-S statistic. While no two mass catalogs were shown to be drawn from the same underlying distribution the K-S statistic does support

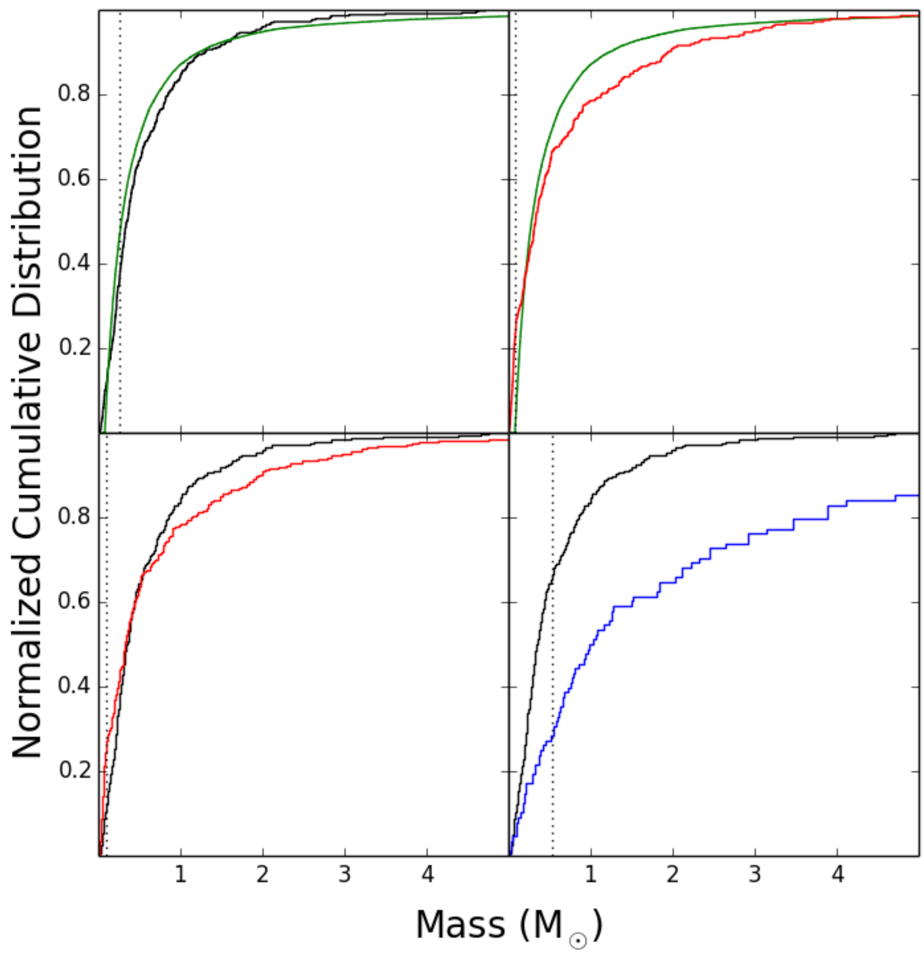

Figure 5.3: K-S test results, dotted line shows location of K-S statistic, D; *top left:* IMF (green) and `clumpfind 3D` (black), *top right:* IMF (green) and `getsources` (red), *bottom left:* `clumpfind 3D` (black) and `getsources` (red) and *bottom right:* `clumpfind 3D` (black) and `SExtractor` (blue).

the functional CMF fits in the previous section. The mass catalogs produced by `clumpfind 3D` and `getsources` are *close* to the IMF. The top two panels in Figure 5.3 ultimately show that the estimated masses lie to the right of the IMF at a given fraction for `clumpfind 3D` and `getsources`. This may suggest that the mass distributions of prestellar cores is fundamentally different than the IMF and can also be explained by a SFE between the CMF and IMF. Likewise, the bottom two panels of Figure 5.3 show a rightward shift in mass derived from continuum maps. A clear result is emerging that is illustrated well in the bottom left panel of Figure 5.3. If the spectral line data does trace the underlying true mass function as Ward et al. (2012) suggest, then `getsources` is the most successful source finding algorithm among the three performed on the continuum maps based on core populations.

5.1.2 Individual Core Comparison

Mass distributions compare entire populations; the results of the previous section suggest that using `getsources` on continuum maps best captures the mass distribution of the IMF and CMF from spectral line data. However, how do the masses compare between algorithms for individual cores? Cores in each of the continuum catalogs were spatially correlated with cores in the spectral line `clumpfind 3D` catalog. Two cores are deemed spatially correlated if the position of peak flux is within a separation, δ,

$$\delta = \sqrt{\mathrm{FWHM}^2_{\mathrm{C^{18}O}} + \mathrm{FWHM}^2_{500}} \tag{5.5}$$

where $\mathrm{FWHM}_{\mathrm{C^{18}O}}$ is the beam FWHM for the spectral line data and FWHM_{500} is the beam FWHM of the 500 μm waveband. Between `clumpfind 3D` and `clumpfind 2D` there are 23 spatially correlated cores, `clumpfind 3D` and `SExtractor`

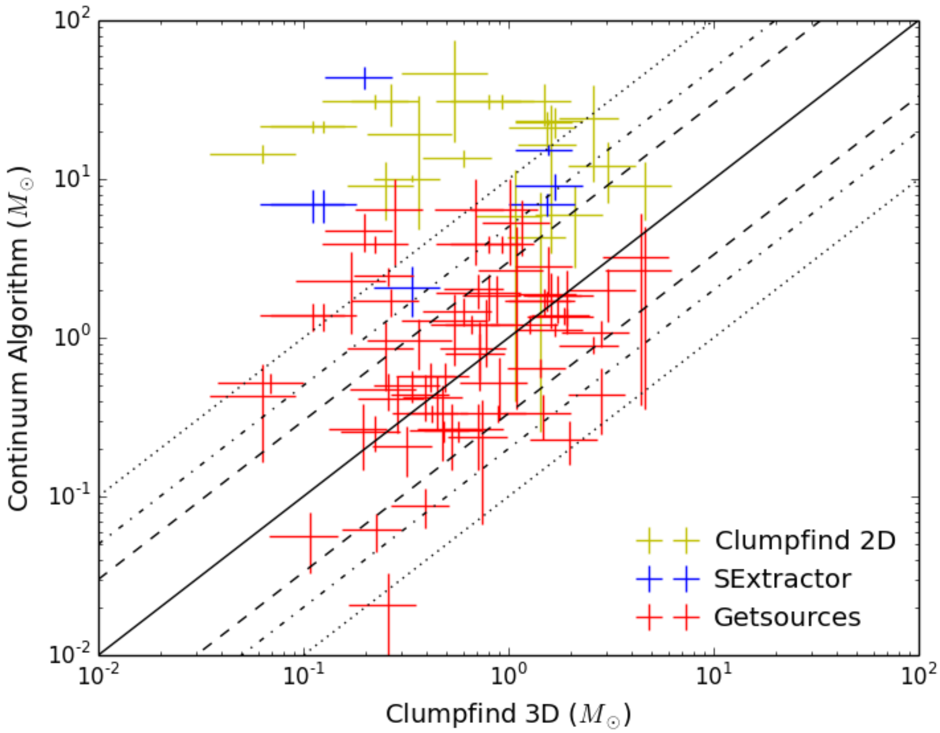

Figure 5.4: Individual core comparison between `clumpfind 3D` and; `clumpfind 2D` (yellow), `SExtractor` (blue) and `getsources` (red). Solid line shows equal mass, dashed, dot-dashed and dotted lines show mass difference by factors 3, 5 and 10 respectively.

there are 7 spatially correlated cores, and `clumpfind 3D` and `getsources` there are 69 spatially correlated cores. The number of spatially correlated cores between `clumpfind 3D` and `SExtractor` is surprisingly low given that `SExtractor` found more cores than `clumpfind 2D`. The answer may lie in Figures 4.5, 4.7 and 4.9. Both `clumpfind 2D` and `getsources` tend to find sources along the main filamentary superstructure in Perseus whereas `SExtractor` is finding sources all over the map. This is important because the four star forming clumps mapped with HARP are found in this superstructure. Figure 5.4 shows individual core comparisons between `clumpfind 3D` and the three algorithms performed on the continuum maps. The solid black line in Figure 5.4 shows equal mass while the dashed, dot-dashed and dotted black lines show mass deviation by factors 3, 5 and 10 respectively. Cores that are plotted above the solid black line in Figure 5.4 indicate continuum mass overestimating spectral line cores while cores plotted below the solid black line indicate the opposite. A single core found in continuum maps may be associated with multiple cores identified in spectral line data. This is a consequence of spectral line cores being separated in velocity but otherwise aligned in position space. An example of this can be seen in a core found at ~ 0.13 $M_\odot$ in the `clumpfind 3D` catalog in Figure 5.4. Here two spectral line cores are associated with one continuum core at ~ 20 $M_\odot$ for `clumpfind 2D`, ~ 7 $M_\odot$ for `SExtractor` and ~ 1.2 $M_\odot$ for `getsources`.

All of the spatially correlated cores from `clumpfind 2D` and `SExtractor` overestimate their `clumpfind 3D` counterparts. Of the spatially correlated cores in the `clumpfind 2D` catalog 17 of 23 (73.9%) overestimate the `clumpfind 3D` mass by over a factor of 10. This result contrasts to 3 of 7 (42.9%) `SExtractor` cores and only 6 of 69 (8.7%) `getsources` cores overestimating spectral lines mass by more than a factor of 10. Just over half of the `getsources` cores, 40 of 69, overestimate cores found in spectra line data. However, the `getsources` cores fare better that the other two continuum catalogs with 46 of 69 (66.7%) cores are within a factor of 3; 54 of 69 (78.3%) cores within a factor of 5 to the spectra line mass.

5.2 Core Environments

Results from the previous section suggest that with accurate background subtraction one can mitigate the amount of overestimation when determining core mass from detailed continuum maps. However, both `SExtractor` and `getsources` subtract a background so why is `getsources` more successful at matching spectral line mass? Both algorithms are different in how they treat background subtraction. `SExtractor` applies a median/mode hybrid filter to each pixel whereas `getsources` identifies structure like filaments and subtracts an appropriate local background based on that structure. The key difference is that a broad filter works well for near isotropic or low dynamic range background whereas filamentary structure is inherently anisotropic.

Many authors have suggested that prestellar cores preferentially form in high density environments like filamentary structure (see Arzoumanian et al. (2011); Pezzuto et al. (2012) for example). In order to determine how much of a role anisotropic background plays in source finding and mass determination the filamentary structure of Perseus was traced out and core distances from these filaments are determined. In order to trace out the filamentary structure of Perseus the optical depth map created from pixel-by-pixel SED fitting (Figure 4.3) was adaptively thresholded using a 0.3 pc (263.3″) box to create a mask where pixels above the median value of the box are included. The top two panels of Figure 5.5 show a section of NGC 1333 (top left) and the initial threshold mask (top right). The initial thresholding flushes out prominent structure that

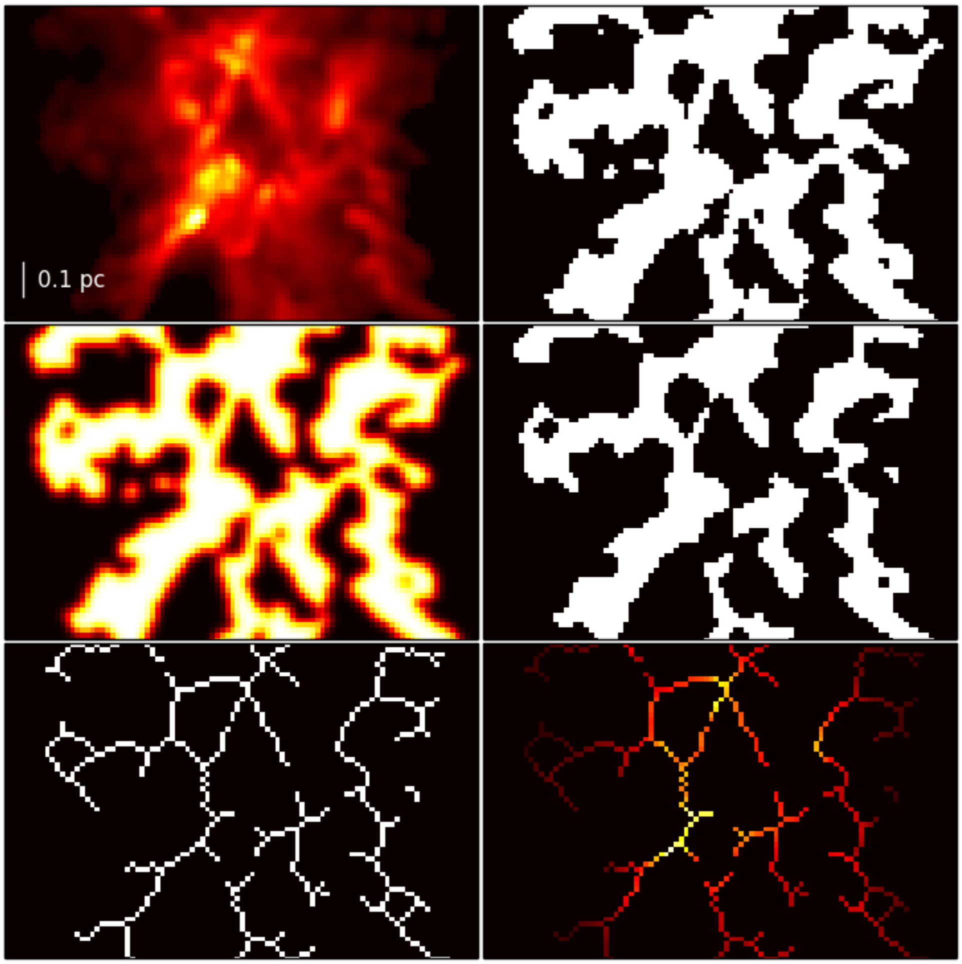

Figure 5.5: Filament tracing process; *top left:* Optical depth ($> 3\sigma_\tau$) map of NGC 1333, *top right:* adaptive median threshold to 0.3 pc box, *middle left:* Gaussian filter to remove small scale noise, *middle right:* global threshold to isolate true structure, *bottom left:* medial axis transformation and *bottom right:* weighted medial axis transformation based on background emission

has a optical depth of at least $3\sigma_\tau$ (root mean squared value: $\sigma_\tau = 0.0005$). The next step filters out small structure noise by applying a Gaussian filter with a standard deviation that matches the standard deviation of the 500 μm beam (middle left panel of Figure 5.5). The map is then globally thresholded taking pixels ≥ 0.85 to create a mask that only contains prominent structure while rejecting spurious small, faint structure and isolated circular noise (middle right panel of Figure 5.5) note in particular the difference between the top right and middle right panels. A medial axis transformation is then applied to the second global mask to determine the skeleton structure. The medial axis of a binary shape is all internal pixels that have 2 or more border pixels that are the shortest and equal distance, often called a topological skeleton (Blum, 1967). The result is a single pixel skeleton that traces through the structure and branches organically (bottom left panel of Figure 5.5). Lastly, the binary filament is weighted corresponding to the factor of optical depth above the noise (σ_τ) at each point in the skeleton (bottom right panel in Figure 5.5).

Each core in the `getsources` catalog had the distance to the nearest weighted skeleton branch calculated. The determined distance is then weighted by the strength of the skeleton by dividing the distance by the σ_τ factor above the noise in the optical depth map. Figure 5.6 shows the results of the weighted core to filament distance against core mass; the vertical dash-dotted line indicates the mean mass found from the lognormal CMF fit (μ, Table 5.1) and the horizontal dotted lines mark the median weighted distance to filament above and below the mean mass μ. Figure 5.6 confirms that high mass cores are preferentially found close to high optical depth (density) structure. As mass decreases cores are predominantly found further away from filaments; however the spread in weighted distances does increase towards lower optical depth. The median weighted distance for the high mass cores (right of the

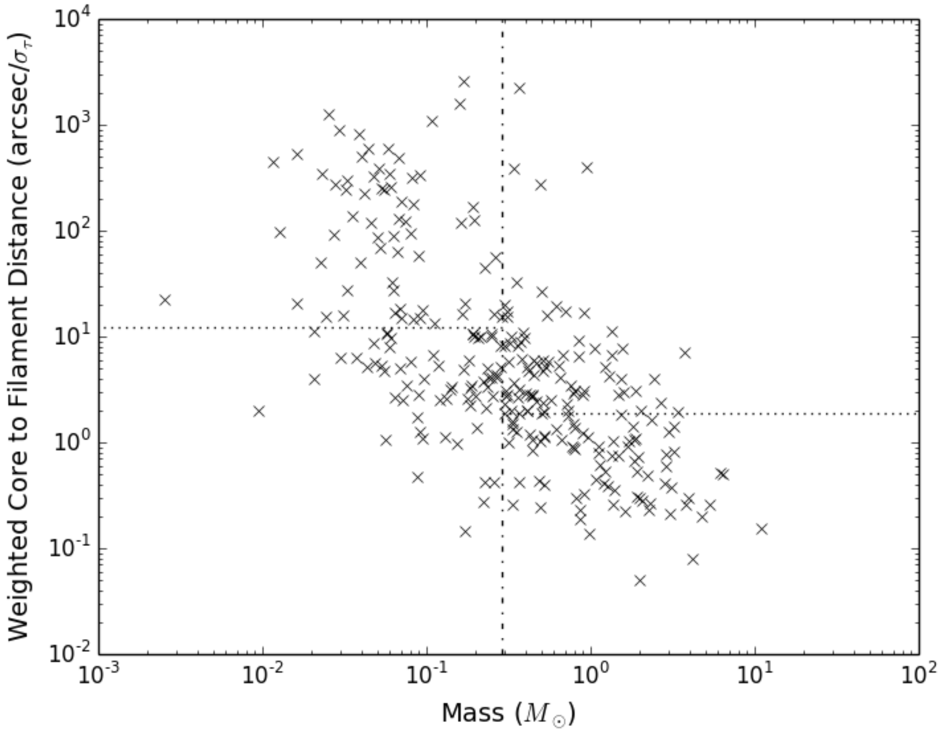

Figure 5.6: Core mass to nearest filament comparison, vertical dash-dotted line indicates lognormal CMF mean (μ), horizontal dotted lines show median weighted distance to filament for core mass above and below CMF mean. Typical mass error is $\sim 30\%$.

dash-dotted line in Figure 5.6) is $1.9''/\sigma_\tau$ while median weighted distance for low mass cores is $12.2''/\sigma_\tau$. A second qualitative interpretation may be that there are predominantly two main groups, a group of higher mass cores found close to filament structure and another group of lower mass cores dispersed away from the structure (Figures 5.6 and 5.7). This core to structure analysis supports previous work (see Arzoumanian et al. (2011); Pezzuto et al. (2012)) suggesting that filaments play a key role in high mass ($> 0.33\mathrm{M}_\odot$) prestellar core formation and evolution.

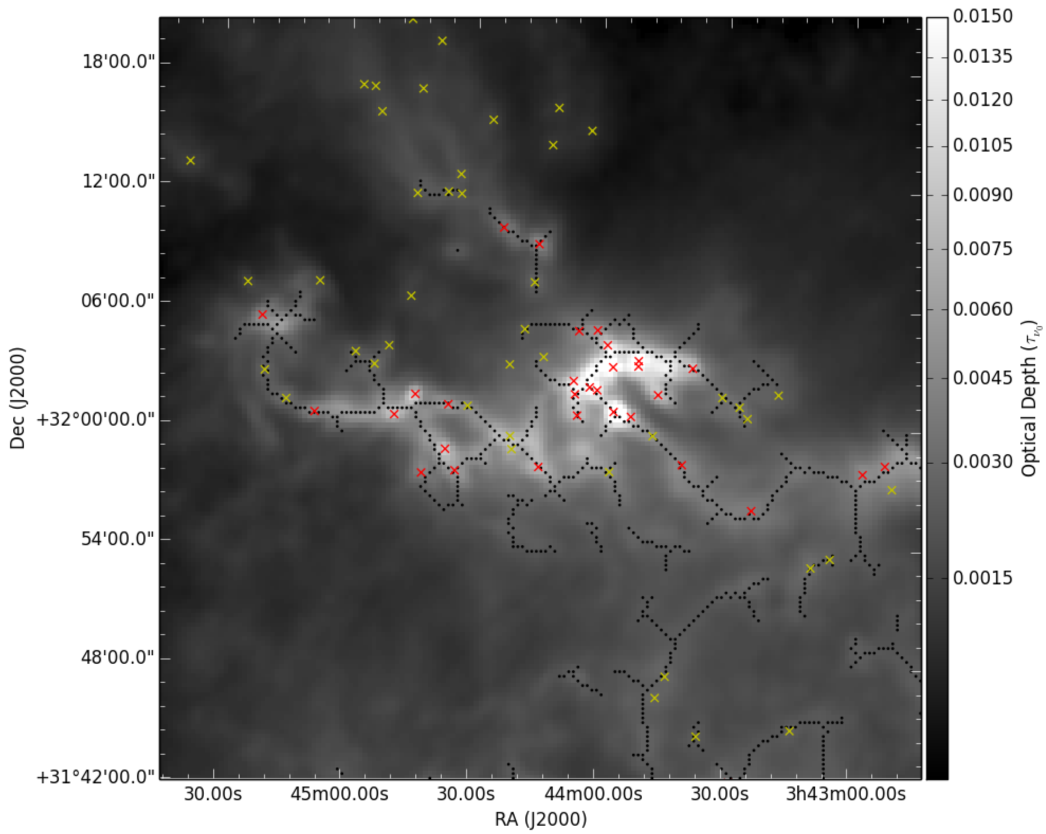

Figure 5.7: Optical depth map of IC 348 showing filament structure as black points, high mass cores (red x) and low mass cores (yellow x).

Chapter 6

Summary

This thesis presents original research involving prestellar core mass comparison from thermal dust emission and spectral line data. In this work I have extracted robust core catalogs using a variety of source finding algorithms. Core masses are determined and populations are compared against each other using parametric and non-parametric methods. Additionally, individual cores are spatially correlated between continuum and spectral line data allowing for direct mass comparison. Lastly, a unique core-to-filament distance is computed for each core found in the `getsources` core catalog to probe the role that environment plays in core mass determination.

The factors that bias prestellar core selection and mass measurement are beginning to take shape. The unambiguous results are that mass estimation from continuum maps does intrinsically overestimate both IMF and mass derived from spectral line data. This mass overestimation can have an impact on an interpreted SFE. The `clumpfind 2D` example may be a perfect illustration of this effect. The estimated mass from this catalog along with CMF functional form fits suggest that the width parameter (α_{high} or σ) accurately captures the IMF. However, the location parameter (M_{break} or μ) greatly overestimates the IMF location suggesting a SFE < 10 %. Despite limited sample size, the core-by-core comparison between `clumpfind 2D` and spectral line mass (Figure 5.4) supports the overestimate claims above. However, if spectral line data was not available one might incorrectly conclude a severely underestimated SFE based on the results of the CMF functional form fitting alone. Even a modest background subtraction employed in `SExtractor` can help to mitigate issues of incorrect SFE. Results from the CMF fitting of the `SExtractor` catalog suggests a global SFE that tends to decrease at high masses because of the larger width parameter (small α_{high}). The trade off here is that incorrectly implemented background subtraction causes many cores in crowded regions to be missed especially if they are of low mass. These missed cores in the `SExtractor` catalog are clearly seen in the dearth of found cores in the crowded main star forming clumps. Exploring spatial scales excels at identifying the core alone without low level background, as `getsources` shows by matching the location parameters of the IMF and spectral line masses. Like the `SExtractor` catalog the width parameter found by `getsources` does suggest a SFE that tends to decrease towards higher core masses. They both simply find too many high mass cores which draws the conclusion that some of these massive cores fragment before becoming main sequence stars. Alternatively, these high mass cores could be the direct result of `SExtractor` and `getsources` inadvertently finding too much mass because of line of sight effects as predicted by (Ward et al., 2012). Given the results of the CMF fitting to the spectral line masses the former explanation seems less likely than the latter especially considering that there is direct evidence that high mass cores evolve in crowded filamentary regions. The core-by-core comparison suggests that `getsources` does a fair job of matching `clumpfind 3D` when considering the cores that were spatially correlated, neither systematically over or under estimating spectral line mass. There is an intrinsic spread in compared masses that comes from the various factors and assumptions that contribute to uncertainties when estimating mass in both PP and PPV data. Considering the results of the three algorithms performed on continuum maps there is a systematic overestimation present when estimating mass from continuum maps that is common to all 2D algorithms. A key finding of this work is that choosing the correct algorithm with thorough background subtraction can mitigate much of the overestimation but in crowded regions it is unavoidable.

The issue of SFE from prestellar cores to the IMF ultimately seems to be unsettled. There does seem to be a mass discrepancy between spectral line and continuum derived masses suggesting line of sight effects are relevant and should not be ignored especially for more massive cores. Additionally, mass estimates from both spectral line and continuum data disagree with the IMF hinting at some global SFE during the last stage in stellar evolution. However, now we can frame contemporary studies which find and estimate core mass in continuum maps appropriately while being mindful of the tremendous bias that is potentially introduced by source finding algorithms.

6.1 Future Work

This work can easily be expanded to included more star forming regions in new GMCs and by comparing more algorithms specifically in 3D data sets. The B1 region has been recently found to contain a crowded star forming clump in the early stages of collapse (see Pezzuto et al. (2012); Sadavoy et al. (2012)) that has yet to be mapped in $C^{18}O$. Obtaining an additional estimate of core flux at ≥ 300 μm (the Rayleigh-Jeans tail) will allow for SED fitting of β since the spectral index has been shown to vary by as much as ± 0.3 in higher density regions (Sadavoy, 2013; Planck Collaboration et al., 2013). As more data becomes publicly available *Herschel* observed GMCs such as Orion, Taurus and others in the Gould Belt Survey will be open for this type of comparison. Additionally, work into quantifying the filament/core relationship has just begun. This work probed the weighted distance but much more effort could be put into categorizing the effect the host clump has on filaments and further down to cores. The *Herschel* and HARP data used in this work provide an excellent amount of data to examine filament properties in both PP and PPV data.

Bibliography

Adams, F. C., Lada, C. J., & Shu, F. H. 1987, ApJ, 312, 788

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

André, P., Men'shchikov, A., Bontemps, S., et al. 2010, A&A, 518, L102

Aniano, G., Draine, B. T., Gordon, K. D., & Sandstrom, K. 2011, PASP, 123, 1218

Arzoumanian, D., André, P., Didelon, P., et al. 2011, A&A, 529, L6

Bachiller, R., & Cernicharo, J. 1986, A&A, 166, 283

Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393

Bertin, E., Mellier, Y., Radovich, M., et al. 2002, in Astronomical Society of the Pacific Conference Series, Vol. 281, Astronomical Data Analysis Software and Systems XI, ed. D. A. Bohlender, D. Durand, & T. H. Handley, 228

Blum, H. 1967, A transformation for extracting new descriptors of form

Buckle, J. V., Hills, R. E., Smith, H., et al. 2009, MNRAS, 399, 1026

Chabrier, G. 2003, PASP, 115, 763

Chabrier, G. 2005, in Astrophysics and Space Science Library, Vol. 327, The Initial Mass Function 50 Years Later, ed. E. Corbelli, F. Palla, & H. Zinnecker, 41

Curtis, E. I., & Richer, J. S. 2010, MNRAS, 402, 603

—. 2011, MNRAS, 410, 75

Curtis, E. I., Richer, J. S., & Buckle, J. V. 2010a, MNRAS, 401, 455

Curtis, E. I., Richer, J. S., Swift, J. J., & Williams, J. P. 2010b, MNRAS, 408, 1516

Dent, W., Duncan, W., Ellis, M., et al. 2000, in Astronomical Society of the Pacific Conference Series, Vol. 217, Imaging at Radio through Submillimeter Wavelengths, ed. J. G. Mangum & S. J. E. Radford, 33

Donkov, S., Veltchev, T. V., & Klessen, R. S. 2012, MNRAS, 423, 889

Enoch, M. L., Evans, II, N. J., Sargent, A. I., et al. 2008, ApJ, 684, 1240

Griffin, M. J., Abergel, A., Abreu, A., et al. 2010, A&A, 518, L3

Hatchell, J., Fuller, G. A., & Richer, J. S. 2007a, A&A, 472, 187

Hatchell, J., Fuller, G. A., Richer, J. S., Harries, T. J., & Ladd, E. F. 2007b, A&A, 468, 1009

Hatchell, J., Richer, J. S., Fuller, G. A., et al. 2005, A&A, 440, 151

Hennebelle, P., & Chabrier, G. 2008, ApJ, 684, 395

Kirk, H., Johnstone, D., & Di Francesco, J. 2006, ApJ, 646, 1009

Kirk, H., Johnstone, D., & Tafalla, M. 2007, ApJ, 668, 1042

Klaassen, P. 2008, PhD thesis, McMaster University

Lada, C. J. 1987, in IAU Symposium, Vol. 115, Star Forming Regions, ed. M. Peimbert & J. Jugaku, 1–17

Lada, C. J., & Lada, E. A. 2003, ARA&A, 41, 57

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

Lombardi, M., Lada, C. J., & Alves, J. 2010, A&A, 512, A67

Men'shchikov, A. 2013, A&A, 560, A63

Men'shchikov, A., André, P., Didelon, P., et al. 2012, A&A, 542, A81

Minchin, N. R., White, G. J., & Padman, R. 1993, A&A, 277, 595

Myers, P. C., & Benson, P. J. 1983, ApJ, 266, 309

Offner, S. S. R., Clark, P. C., Hennebelle, P., et al. 2013, ArXiv e-prints, arXiv:1312.5326

Olmi, L., Anglés-Alcázar, D., Elia, D., et al. 2013, A&A, 551, A111

Padoan, P., & Nordlund, Å. 2002, ApJ, 576, 870

Pekruhl, S., Preibisch, T., Schuller, F., & Menten, K. 2013, A&A, 550, A29

Pezzuto, S., Elia, D., Schisano, E., et al. 2012, A&A, 547, A54

Pilbratt, G. L., Riedinger, J. R., Passvogel, T., et al. 2010, A&A, 518, L1

Pineda, J. L., Goldsmith, P. F., Chapman, N., et al. 2010, ApJ, 721, 686

Planck Collaboration, Abergel, A., Ade, P. A. R., et al. 2013, ArXiv e-prints, arXiv:1312.1300

Poglitsch, A., Waelkens, C., Geis, N., et al. 2010, A&A, 518, L2

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 2002, Numerical recipes in C++ : the art of scientific computing

Reid, M. A., & Wilson, C. D. 2006, ApJ, 650, 970

Roussel, H. 2012, Scanamorphos: Maps from scan observations made with bolometer arrays, astrophysics Source Code Library, ascl:1209.012

Rybicki, G. B., & Lightman, A. P. 1979, Radiative processes in astrophysics

Sadavoy, S. 2013, PhD thesis, University of Victoria

Sadavoy, S. I., di Francesco, J., André, P., et al. 2012, A&A, 540, A10

Salpeter, E. E. 1955, ApJ, 121, 161

Tielens, A. G. G. M. 2005, The Physics and Chemistry of the Interstellar Medium

Veltchev, T. V., Klessen, R. S., & Clark, P. C. 2011, MNRAS, 411, 301

Ward, R. L., Wadsley, J., Sills, A., & Petitclerc, N. 2012, ApJ, 756, 119

Williams, J. P., de Geus, E. J., & Blitz, L. 1994, ApJ, 428, 693