Università degli Studi di Roma "Tor Vergata"
Dottorato di Ricerca in Astronomia – XXII ciclo

Multiwave analysis of High-Mass star forming regions.

by

Fabiana Faustini

November 2009

Program Coordinator
Prof. Pasquale Mazzotta

Thesis Advisors
Dr. Sergio Molinari
Dr. Francesco Berrilli

Ringraziamenti

Questo cammino è stato lungo, e non sempre facile... ma accanto a me ho avuto persone che mi hanno incoraggiato ricordandomi che non ero mai sola ad affrontarlo. Mio marito, Fabio, è stato la mia forza nei momenti più difficili quando pensavo che in fondo non ne valeva la pena e che era meglio lasciar perdere,lui mi ha spinto a portare a termine questo percorso credendo fortemente nelle mie capacità... I miei genitori, mi hanno dato la possibilità di intraprendere questi studi spianando la mia strada e mai hanno dubitato di me... Tutti i miei parenti più vicini e i miei amici mi hanno sempre supportato, rallegrando i momenti più bui... Vi ringrazio tutti per quello che avete fatto per me, ed oggi divido con voi la gioia di aver tagliato questo traguardo...

grazie a tutti di cuore,
Fabiana

Abstract

There have been considerable efforts to understand how stars form from both a theoretical and an observational point of view. We have reached a good understanding of how isolated lowmass stars form (Klein et al. 2006). The widely accepted scenario is that low-mass stars form by the gravitational collapse of a prestellar core followed at later stages by disk accretion. Extending this theory to high-mass stars is not trivial. Highmass (proto-)stars reach the zero age main sequence while still accreting. When the central protostar reaches a mass of about 10 $M_{\odot}$ hydrogen fusion ignites in the core and the star's radiation pressure and wind should prevent further accretion. Several theories are today proposed, we discuss about them in the introduction, and we try to discriminate between these theoretical models through the re-building of the Star-Formation Hystori of clusters formed in high-mass star formation regions.

The presentation of this work is divided into three section.

- The first part presents the analysis of our sample and the discussion of our scientific results, it is divided in three chapter. In the chapter 2 we presents the results of our analysis in the Near-IR banbs to characterized the properties of low mass cluster in our sample, while in the chapter 3 we shown the SEDs building for intermediate and high-mass objects and the fits with theoretical models. In the chapter4 we take again our results on all the examinated wavelengths to extrapolate the information about the clusters star formation history.

- In the second section the structure and the performances of our data analysis algorithm is presented.

- The last section recapitulates the results obtain in all this work

Contents

Chapter 1

Introduction to the Star Formation

Stars are the "Atoms" at the basis of the universe, the problems of how stars form is one of the central themes of the contemporary astrophysic. By transforming gas and interstellar medium into stars, the formation process change the medium and determines the structure and the evolution of galaxies. A lot of questions are opened on processes that drive the evolution of a molecular cloud (situated in a bigger and homogenuos cloud) into a clusters of cores and these cores into protostars.

1.1 The importance of high-mass stars

There have been considerable efforts to understand how stars form from both a theoretical and an observational point of view. We have reached a good understanding of how isolated lowmass stars form (Klein et al. 2006). The widely accepted scenario is that low-mass stars form by the gravitational collapse of a prestellar core followed at later stages by disk accretion.

Extending this theory to high-mass stars is not trivial. Highmass (proto-)stars reach the zero age main sequence while still accreting. When the central protostar reaches a mass of about 10 $M_{\odot}$ hydrogen fusion ignites in the core and the star's radiation pressure and wind should prevent further accretion. This is obviously a paradox given that yet more massive stars do form. Several theories have been put forward to solve this dilemma (Zinnecker & Yorke 2007), such as accretion rates of up to three orders of magnitude higher than in the case of lowmass stars (Cesaroni 2005), and non-spherical accretion geometries (Nakano 1989; Yorke 2002; Keto 2003), or coalescence in dense (proto-)stellar clusters (Bonnel et al. 1998). All of these theories have predictions that can, in principle, be tested observationally. Significant effort has been made to detect massive accretion disks (Cesaroni et al. 2006), powerful outflows (Beuther et al. 2002; Cesaroni et al. 2005), and dense protostellar clusters (Testi et al. 1999; de Wit et al. 2005), all of which are predicted by one or other formation theory. None of these efforts have provided conclusive arguments in favour or against any of the theories.

Why the comprehension of high-mass star forming process is so important? High-mass sources are fundamental in the evolution of the galaxy. With their short but bright life they are the greater sources of energy for the interstellar medium. With their birth and their death (that it happens passing through the supernovae phase, it's the opposite of what happens for the low mass stars) they give an high radiation quantity to the galactic atmosphere. They

inject, during their explosion, an high amount of heavy elements that enrich and change the chimical composition of galaxy. This renewed medium is the site of birth of evolved stars with an higher contents of heavy elements; in this way the stellar generation changes. We must to comprise their distribution, spatial and mass distribution, and their evolutive phase to comprise the physical state, the chimical content and the morphology of our, and external, galaxy.

1.2 Star forming sites

Today is accepted that the star formation sites are the *giant molecolar cluods* (GMCs) where it's found the molecolar gas out of which stars form. They occupy a small fraction of volume of the inter stellar medium (ISM) but comprise a significant fraction of the mass. GMCs have masses that excess of 10^4 $M_\odot$ and dimensions of the order of 1 Kpc. They are sorrounded by a layer of atomic gas that shields the molecules from the interstellar UV radiation (that would dissociate them), for example a column density of about $N_H \simeq 2 \times 10^{20} cm^{-2}$ for the atomic layer is required to form the CO. This molecule is strongly associated to star forming regions, indeed its presence is critical to reach the gas temperature of about 10 K with the cooling process that is required to permits the Jeans fragmentation, that assures the formation of the GMCs sub-structures, described below.

Larson (1981) summarized some of the key of dynamical features of GMCs into three laws. GMCs :

- are supersonically turbolent with velocity dispersions that increase as a power oh the size (line width-size relation) in conformity with the relation

$$\sigma \propto R_{pc}^{1/2} \tag{1.1}$$

  where $R_{pc} \equiv R/(1pc)$.

- are gravitationally bound ($\alpha_{vir} \simeq 1$ where α_{vir} is the virial parameter).

- all have similar column densities

Solomon et al. (1987), in their study of GMCs in the first Galactic quadrant (that contain the GMCs of the solar circle), confirm the validity of these laws finding a line width-size relation of $\sigma = (0.72 \pm 0.07) R_{pc}^{0.5\pm0.05} kms^{-1}$, an $\alpha_{vir} = 1.1$ and a mean column density of $\bar{N}_H = (1.4 \pm 0.3) \times 10^{22} R_{pc}^{0.0\pm0.1} cm^{-2}$

As Larson pointed out, these laws are not indipendent and any two of them imply the third. If we express the line width-size relation as $\sigma \equiv \sigma_{pc} R_{pc}^{1/2}$, the equation of the virial teorem becomes

$$\alpha_{vir} = \left(\frac{5}{\pi pc}\right) \frac{\sigma_{pc}^2}{G\Sigma} = 3.7 \left(\frac{\sigma_{pc}}{1 kms^{-1}}\right)^2 \left(\frac{100 M_\odot pc^{-2}}{\Sigma}\right) \tag{1.2}$$

and relates the three scaling laws.

There are actually two main theories for interpreting the GMC properties. The first is that GMCs are dynamic, transient entities in which the turbolence is driven by large-scale colliding gas flows that create the clouds (Vázquez-Semadeni et al. 2006; Ballesteros-Paredes et al.

2007). The second theory foresees that GMCs are formed by large-scale self-graviting instabilities, and the turbolence that they contain is due to a combination of natural turbolence of diffuse ISM, conversion of gravitational energy to turbolence ones during the structure contraction and the energy injection from forming stars. The contribution of the different elements changes during the time (McKee 1999; McKee & Holliman 1999).

GMCs structure is no uniform but presents a hierarchical structure that extend from the scale of the cluod down to the thermal Jeans mass[1] for the case of gravitationally bound clouds, and down to smaller masses for unbound clouds (Langer et al. 1995). Overdense regions within GMCs are termed *clumps*, massive clumps are the sites where the star clusters forms and they are generally gravitationally bound. Clumps present further fragmentation, denser regions where the single stars (or multiple sistems like binaries) form, these are the *cores* that are necessarily gravitationally bound. Most ^{13}CO clumps are unbound, and therefore do not obey Larson's laws; the mass distribution of such clumps can extend in an unbroken power law from several tens of solar masses down to Jupiter masses (Heithausen et al., 1998). On the other hand, Bertoldi & McKee (1992) found that most of the mass in the clouds is concentrated in the most massive clumps, and these appear to be gravitationally bound.

1.3 Star Formation

The stars formation theories are traditionally divided into two parts: low and high mass. The distinction between these two group is fixed at a mass of 8 $M_\odot$. Protostars that will forms stars with masses below this limit have luminosities dominates by accretion. While protostars above this mass value have luminosities dominated by nuclear burning unless the accretion rate is very high. Low-mass stars undergo extensive pre-main sequence (pre-MS) evolution in the Hertzsprung-Russell (HR) diagram from the birthline[2] , and they become observable before their arrive on the MS. While high-mass protostars have a brief pre-MS phase and they arrive on the zero-age main-sequence (ZAMS) still obscured by their dense envelope. These features of high-mass star formation process make difficult the observations of these objects.

1.3.1 From Core to Star

Low-mass stars appear to form from gravitationally bound cores. At the outset of theoretical studies of star formation, it was realized that isothermal cores undergoing gravitational collapse become very centrally concentrated, with a density profile that becomes approximately $\rho \propto r^{-2}$ (Bodenheimer & Sweigart, 1968; Larson, 1969). The initial configuration is the singular isothermal sphere (SIS), which is an unstable hydrostatic equilibrium. The collapse starts at the center, and the point at which the gas begins to fall inward propagates

[1] The Jeans mass is the mass value over that the considered object becomes gravitationally unstable. This value is correlated to the density and the temperature of the objects through the law

$$M_{Jeans} \approx 3.3 \cdot 10^{22} \left(\frac{T^3}{\rho} \right)^{1/2}$$

[2] The birthline rapresents, on the HR diagram, the ... from a purely convective objects to a radiative ones (Palla & Stahler 1992, 1993).

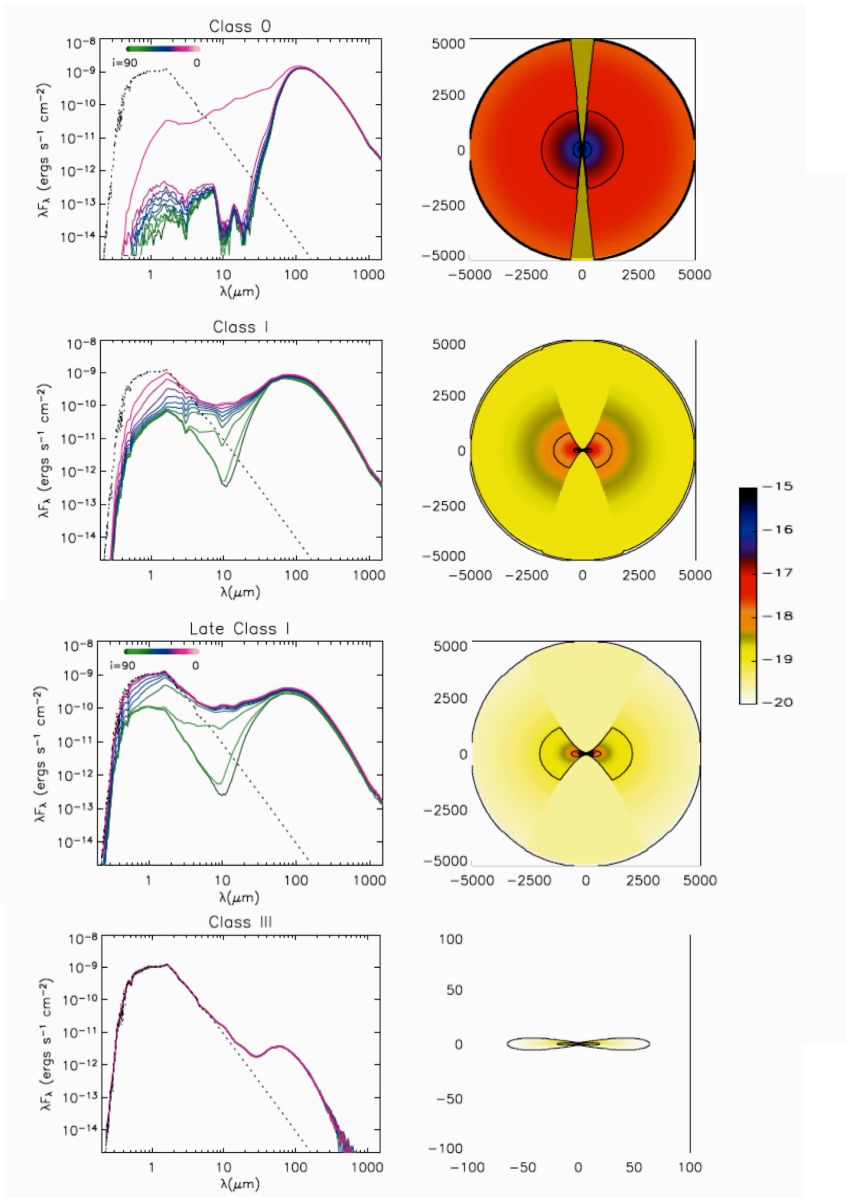


Figure 1.1. Spectral energy distribution (on the left side) are presented for different evolutive phases, from Class 0 (top) to Class III (bottom) for several observation angles. On the right side is shown the gas distribution and its temparature gradient (color-scale) for the same classes.

outward at the sound speed (the "expansion wave"). This solution is therefore termed an "inside-out" collapse.

When the central region collapses to generate the protostar its dimensions decreases and, for the angular momentum conservation, the angular velocity increases. This rotation aims to flat the gas and dust distribution in the nearest region of the collapsing objects

The growth of protostars can be inferred through the characterization of the spectral energy distribution (SED) of the continuum, see Fig. 1.1. Protostellar SEDs are conventionally divided into four classes, which are believed to represent an evolutionary progression (Myers et al., 1987 divided sources into two classes; Lada, 1987 introduced Classes I-III; and Andre, Ward- Thompson, & Barsony, 1993 introduced Class 0). The classification is summarized in the following scheme:

- ***Class 0***: sources with a central protostar that are extremely faint in the optical and near IR (i.e., undetectable at $\lambda < 10$ μm with the technology of the 1990's) and that have a significant submillimeter luminosity, L_{smm}/L_{bol} 0.5%. Sources with these properties have $M_{envelope} \gtrsim M_{\star}$Protostars are believed to acquire a significant fraction, if not most, of their mass in this embedded phase.
- ***Class I***: sources with $\alpha_{IR} > 0$, where $\alpha_{IR} \equiv \frac{dlog(\lambda F_{\lambda})}{dlog\lambda}$ is the slope of the SED over the wavelength range between 2.2 μm and 10-25 μm. Such sources are believed to be relatively evolved protostars with both circumstellar disks and envelopes.
- ***Class II***: sources with $-1.5 < \alpha_{IR} < 0$ are believed to be pre-main sequence stars with significant circumstellar disks (classical T Tauri stars).
- ***Class III***: sources with $\alpha_{IR} < -1.5$ are pre-main sequence stars that are no longer accreting significant amounts of matter (weak-lined T Tauri stars).

Figure 1.1 shows how the SEDs, and the distribution of material (gas and dust) in the components of protostars, change during the evolution of the protostars. In all the SED plots the dashed curve rapresents the protostars photosphere emission. Class 0 show a dominant emission from the mid-IR to mm related to the luminosity produced by the central objects and reprocessed by the dense envelope. This contribution decreases during the evolution of the source, when the envelope becomes less dense, and become visible the direct emission of the protostar. Class III SED shows only the emission of the source and a little contribution of the disk in the mid-IR and far-IR wavelenghts.

1.3.2 The High Mass Protostars

High mass star formation is a process less clear respect to the low mass one. Indeed the high-mass objects have very short pre-MS phases and arrive at the ZAMS still embedded in their dense envelope when their main accretion phase is still going. The observation of these very young objects is not trivial and only today with powerfull technical means we can start to find and characterize high-mass protostars.

Two main categories of theories are today proposed:

- *Accretion theories*, are based on the extrapolation, at higher mass, of the low-mass theories. When the central protostar reaches a mass of about 10 $M_{\odot}$ hydrogen fusion ignites in the core and the star's radiation pressure and wind should prevent

further accretion. This is obviously a paradox given that yet more massive stars do form. Several theories have been put forward to solve this dilemma (Zinnecker & Yorke 2007), such as accretion rates of up to three orders of magnitude higher than in the case of low-mass stars (Cesaroni 2005), and non-spherical accretion geometries (Nakano 1989; Yorke 2002; Keto 2003) or competitive accretion (Bonnell et al. 2001).

- *Coalescence theories* (Bonnell et al. 1998) in dense (proto-)stellar clusters. This theory foresees that high-mass form not from a single dense core, as low mass protostar, but from a merging of some core that are located in the inner and denser region of the protocluster.

Some evidence has been collected that accretion has a fundamental rule also in the formation of massive star, such as observation of disks and outflows (Churchwell 1997; Zhang et al. 2001; Beuther et al. 2002) directly collegated to the accretion fenomena. This scenario prefers that models that foresee the principle mechanism at the basis of massive star formation is the accretion respect to coalescence; indeed this last theory foresee that disks and outflows must be destroyed during the merging between the different sources.

These observational evidences prefer accretion theories, but this branch includes several theories that foresee different scenario, so other discriminate factor are required to understand the real process at the basis of high-mass cluster formation.

1.4 Observing High Mass star forming region from Near-IR to millimeter.

High-mass stars form in populated clusters, where low mass objects are numerically dominant (Lada & Lada 2003). Starting from this assumption, verified with a lot of observations, we have made a multiwave analisys of high-mass star forming regions to comprehend the history of the formation of these objects.

Each wavelength trace a different source population, as it's shown in Fig.1.2:

- Near-IR bands: J, H and K mostly reveal emission from low mass sources. Younger massive objects don't emit in this bands yet.
- Mid-IR bands: Only the emission of a small number of low mass sources survives at these wavelengths. Most intermediate mass objects dominate, and the most massive object starts to be visible.
- Far-IR and Sub-mm bands: The massive object becomes the brightest source in the Far-IR and completely dominates the emission of cold dust in the sub-millimeter.

The SED changes during the evolution of the source, as previously said for low-mass classification, so a good characterization of the spectral shape can be used to understand the evolutive phase of our sources. Besides, the simultaneous study of low and high mass populations allows us to cast light on some aspects of the star formation process that are not evident when the two different populations are considered separately, and allow us to rebuild the Star Formation History (SFH) of our young cluster that is a powerful tool to discriminate between several theories of star formation in clusters. Indeed different theories foresee different time for the formation of a star clusters:

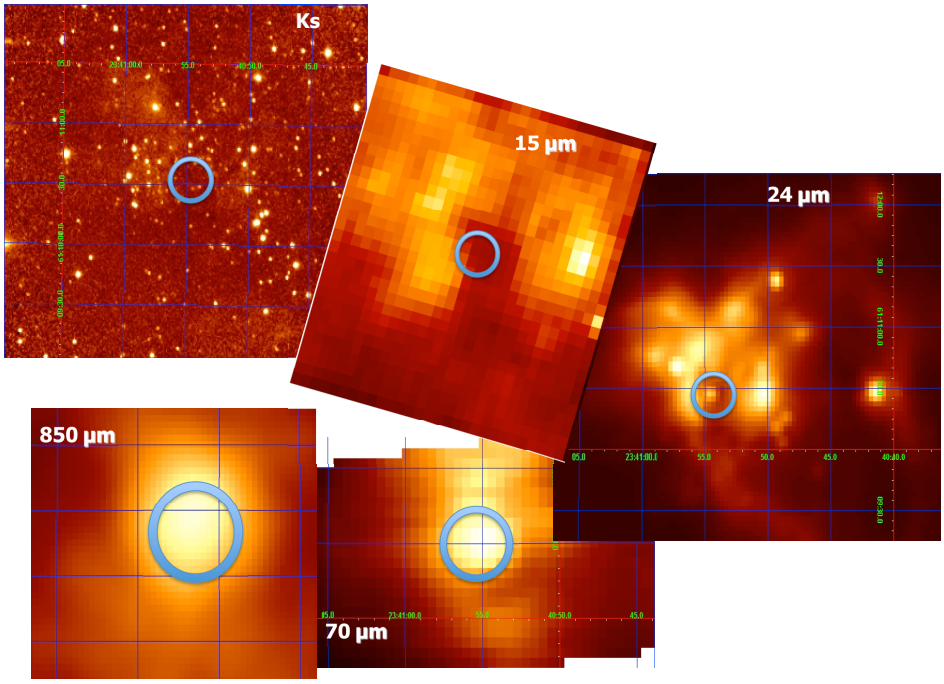


Figure 1.2. A star forming region (Mol160, see table 1.1) is shown in several wavelenghts in ascending order from the left-top (k_S) to the left-bottom (850μm).

- competitive accretion (Bonnell et al. 2001) foresee a clump with a weak turbolence and the formation of all the stars simultaneously in about $2t_{ff}$
- models based prevalently on gravitational collapse (McKee & Tan 2003; Krumholz, McKee, & Klein 2006) foresee a medium with a turbolence comparable with the gravitational energy that implies a longer star forming time (few t_{ff})

Our aim is to obtain information about the time of formation of our clusters to discriminate between these theories. In the chapter 2 we presents the results of our analysis in the Near-IR banbs to characterized the properties of low mass cluster in our sample, while in the chapter 3 we shown the SEDs building for intermediate and high-mass objects and the fits with theoretical models. In the chapter 4 we take again our results on all the examinated wavelengths to extrapolate the information about the clusters star formation history. The properties and the performance of our algorithm, developed for the analysis of these regions, are presented in the chapter 5. Finally the conclusion of this work are summerized in the last chapter, 6.

1.5 The source sample

Our sample was selected from a larger sample of candidate high-mass protostars selected and analyzed by Molinari et al. (1996, 1998, 2000, 2002); Brand et al. (2001).The samples contain sources grouped into High and Low sources mainly according to their [25-12] $\gtrless$ 0.57 IRAS color value. The adopted threshold was recommended by Wood & Churchwell

(1989) to select sources with a high probability of association with UCHII regions, and these are the High sources in our terminology. A robust case has been built in the past to show that the two groups of sources exhibit different properties; the Low sources appear to be less strongly characterised than High but they show convincing evidence of being on average relatively younger. Program fields are listed in Table 1.1. All the sources have been observed in the Near-IR with two telescopes from earth, as describe in chapter 2. Some field are included in two Spitzer[3] Legacy Programs, GLIMPSE[4] (Mid-IR bands) and MIPSGAL[5] (Mid and Far -IR bands), described in chapter 3. Finally for a filed subsample we have some millimetric observation, as explain in chapter 3.

The possibility to analyze these field in a multi-wave perspective has gave us some information about the global evolution of all the members of the cluster, from low to high mass, as shown in the following sections.

[3]NASA's Spitzer Space Telescope is a space-based infrared observatory, part of NASA's Great Observatories program,and consisting of a 0.85 meter telescope and three cryogenically-cooled science instruments. During its mission, Spitzer will obtain images and spectra by detecting the infrared energy, or heat, radiated by objects in space between wavelengths of 3 and 180 μm.

[4]Galactic Legacy Mid-Plane Survey Extraordinaire

[5]MIPS/Spitzer Survey of the Galactic Plane

Table 1.1. List of Sources

Source Mol[a]	IRAS Name	α(J2000)	δ(J2000)	Tel. NIR	GLIMPSE	MIPSGAL	mm
3	00420+5530	00:44:57.6	+55:46:52	Pal	N	N	Y
8	05137+3919	05:17:13.3	+39:22:14	Pal	N	N	Y
9	05168+3634	05:20:16.2	+36:37:21	Pal	N	N	N
11	05345+3157	05:37:47.8	+31:59:24	Pal	N	N	Y
12	05373+2349	05:40:24.4	+23:50:54	NTT	N	N	Y
15	06056+2131	06:08:41.0	+21:31:01	Pal	N	N	N
28	06584−0852	07:00:51.0	−08:56:29	Pal	N	N	N
30	17450−1742	17:48:09.3	−27:43:21	NTT	Y	Y[c]	N
38	18024−2119	18:06:18.0	−21:42:00	NTT	Y	Y	Y
45	18144−1723	18:17:24.2	−17:22:13	NTT	Y	Y	Y
50	18162−1612	18:19:07.5	−16:11:21	NTT	Y	Y	N
59	18278−1009	18:30:35.2	−10:07:12	Pal	Y	Y	Y
75	18511+0146	18:53:38.1	+01:50:27	Pal	Y	Y	Y
82	18565+0349	18:59:03.4	+03:53:22	NTT	Y	Y	N
84	18567+0700	18:59:13.6	+07:04:47	NTT	N	N	N
98	19092+0841	19:11:37.4	+08:46:30.0	NTT	Y	Y	Y
99	19094+0944	19:11:52.0	+09:49:46	Pal	Y	Y	N
103	19213+1723	19:23:37.0	+17:28:59	NTT	Y	Y[c]	Y
109	19374+2352	19:39:33.2	+23:59:55	NTT[b]	Y	Y	Y
110	19388+2357	19:40:59.4	+24:04:39	NTT[b]	Y	Y	Y
136	21307+5049	21:32:31.5	+51:02:22.0	Pal	N	N	Y
139	21519+5613	21:53:38.8	+56:27:53.0	Pal	N	N	Y
143	22172+5549	22:19:09.0	+56:04:45.0	Pal	N	N	Y
148	22305+5803	22:32:24.3	+58:18:58.2	Pal	N	N	Y
151	22506+5944	22:52:38.6	+60:00:56.0	Pal	N	N	Y
160	23385+6053	23:40:53.3	+61:10:19.1	Pal	N	Y[d]	Y

[a] Source running number from Molinari et al. (1996).

[b] Imaged only in K_s.

[c] Imaged only in 24 μm.

[d] Observed within a Garanteed Time Programs.

I Part

Scientific Analysis

Chapter 2

Near InfraRed: low mass clusters

2.1 Observations and data analysis

Program fields are listed in Table 1.1 and were imaged in J, H, and K_s bands. A total of 15 fields were observed in three nights in November 1998 at the Palomar 60-inch telescope equipped with a 256×256 NICMOS-3 array of pixel scale 0.″62/pix and total FOV 2.′6×2.′6. The remaining 11 fields were observed in 3 nights in August 2000 at the ESO-NTT using the 1024×1024 SOFI camera with a pixel scale of 0.″29/pix and a total FOV of 4.′9×4.′9. Standard dithering techniques were used to minimize the impact of bad pixels and optimize flat-fielding, allowing us to achieve for each field a total of 5min integration time per band (in the central portion of the observed field) within an area of 3.′5×3.′5 of Palomar observations, and 20min (10min for the band K_s) at NTT of a total covered area of 6.′5×6.′5. Suitable calibration sources from the list of Hunt at al. (1988) were observed regularly during the observations to track atmospheric variations for different airmasses. Standard stars and target fields were observed at airmasses no greater than 1.7 at NTT, and 1.3 at Palomar; we determined average zero-point magnitudes for each night and used them to calibrate our photometry. For each field, the images in the three bands were registered and astrometric solutions were determined using a few bright optically visible sources.

The K_s images for all observed fields, with superimposed submillimeter continuum emission distribution when available (Molinari et al. 2008a) are available online[1]

2.1.1 Point source extraction and photometry

The extraction and photometry of point sources for all images were completed using the IRAF package. The r.m.s of the background signal and the FWHM of point sources were measured throughout the images to characterize the image noise and PSF properties; these parameters were fed to the DAOFIND task for source extraction, where a detection threshold of 3σ was used for all images. Sources with saturated pixels were excluded from the analysis; the linearity of the system response was checked *a posteriori* comparing, both for the Palomar and the NTT data, the magnitudes obtained to those from 2MASS using a few stars with magnitudes reaching up to the maximum values found in our photometry files; the relations between the 2MASS magnitudes and ours in the three bands were found to be linear over the entire magnitude range of the detected sources. There were clearly brighter

[1] at http://galatea.ifsi-roma.inaf.it/faustini/maps/

objects in the various fields, but their peaks were already flagged as saturated and excluded from the detection process.

The photometry of sources is difficult to determine in very dense stellar fields such as the inner Galactic plane, where all our target fields are located and the crowding is such that more than one source can enter either any plausible aperture chosen or any annulus used for background estimation. This problem is of course more extreme in the clustered environments close to detected sites of massive star formation (see Sect2.2.1 below).

The first alternative approach that we tried to follow was PSF-fitting photometry that should be less affected by these problems. We chose a subsample of test fields with different levels of stellar crowding. In this procedure, an important aspect was the modeling of the PSF. To test this, we completed several trials selecting a variable number of point-like sources (from 3 to 30) of different brightness levels and different positions in the field. We found that the resulting PSF model was not particularly sensitive to the choice of numbers and/or brightnesses of the stars. However the results, were quite dependent on the mean stellar density of the field. The photometry was carried out using the ALLSTAR task, which was particularly suited to crowded fields. However, we also tested the other two tasks (PEAK and NSTAR) and obtained comparable results for most of the sources. We note, however, that in the most crowded areas in particular, the subtraction of the PSF-fitted sources from the image introduced two spurious effects: an unacceptably high level of residuals with brightness levels well above the detection threshold used and a significant number of negative holes, indicating that the PSF-fit included some background in the source flux estimation and therefore overestimated its value. Both effects are caused by both the limited accuracy of the PSF model that can be obtained for very crowded fields, where faint neighboring stars can enter the area where the PSF model is estimated, and the presence of a significant and variable background, which is quite common and expected for the Galactic plane. A similar conclusion was reached by Hillenbrand & Carpenter (2000) in their study of the inner Orion Nebula Cluster.

The second approach that we followed was standard aperture photometry. The choice of radii for both the aperture and the background annuli was of course extremely important. The optimum aperture should be neither too large to include nearby sources nor too small to truncate significantly the PSF and underestimate the flux significantly. We completed several attempts for one of the most crowded fields (Mol30, observed at NTT) with three different aperture radii equal to the PSF FWHM (typically $0\farcs7$ at NTT and $1\farcs4$ at Palomar in K_s), and twice and thrice this value. For each photometry run, we analyzed the source flux distribution and, as expected, the median flux was found to increase with increasing aperture radius. Increasing the aperture from one to two PSF FWHMs increased the median source flux by an amount compatible with the inclusion of the first ring of an Airy diffraction pattern. In contrast, when the aperture radius was increased to a factor of three higher than the PSF FWHM, the flux increase was far higher than could be attributed to the additional fraction of the Airy profile entering the aperture, and must therefore have been caused by the inclusion of nearby sources. We adopted an aperture radius equal to the PSF FWHM to minimize neighbor contamination, and then applied an aperture correction factor to the fraction of the PSF removed by the aperture; this was estimated by multi-aperture photometry (starting from a size of 1 FWHM) on relatively isolated stars in the target fields.

Given the crowding of our fields, a further effect to be corrected for is the possible contamination by the tails of the brightness profiles of neighbouring stars. To quantify this contamination, we created a grid of simulations with two symmetric Gaussians with a wide

variety of peak contrasts and different reciprocal distances. We computed the fraction of the Gaussian profile of the neighbouring source within the photometry aperture centered on the main source, and hence generated a matrix of photometry corrections for different source distances and peak contrasts. We then processed the magnitude file produced by the aperture photometry task and for each source we applied a magnitude correction depending on the presence, distance, and contrast ratios with respect to other neighbouring stars.

In spite of the various issues discussed above, the photometric data obtained with the two methods were in good agreement with each other, apart from at faint magnitudes. For these faint objects, we consistently found that the PSF photometry tends to produce brighter magnitudes (and hence stronger sources) than the aperture photometry; this effect can be easily understood from our finding (see above) that the subtraction of PSF-fitted sources always leaves negative holes in the residual image, and this effect is far more important for faint stars. We thus decided to adopt the magnitudes determined from aperture photometry.

For each target field, we estimated the limiting magnitude (LM) using artificial star experiments. The fields were populated using the IRAF task ADDSTAR with 400 fake stars with magnitudes distributed in bins of 0.25 mag between values of 15 and 21; the percentage of recovered stars as a function of magnitude provides an estimate of the completeness level of our photometry. The star recovery percentage was not found to decrease monotonically with increasing magnitude because fake stars can also be placed very close to bright real stars and then go undetected by the finding algorithm. However, we find that the limit of 85-90% recovery fraction is reached on average at around J=18.7, H=17.7, and K_s=17.4 for NTT images, and J=18.0, H=17.3, and K_s=16.6 for Palomar images. We found that the typical photometric uncertainty is below 0.1mag close to the limiting magnitude.

To verify the integrity of our photometry, we compared our magnitudes with those extracted from 2MASS point source catalog for all the fields in our sample. Considering the differences in spatial resolutions between 2MASS and the telescopes used for our observations, this comparison was limited to 2MASS point-like sources associated with a single source in the Palomar or NTT images. The median differences with respect to 2MASS for the various fields are of the order of -0.1, -0.2 and -0.3 mag for J, H, and K_s bands, respectively. Within each field, the scatter around these median values is ~ 0.1 mag in all three bands, confirming the internal consistency of our photometry. Noticeable departures ($\sim$0.5 mag) of the median difference with 2MASS from the above values are observed for the field of source Mol11 (Palomar), and for sources Mol103, Mol109 and Mol110 (NTT). However, the latter sources were observed on the same night, observations for which our log registered as not good due to sky variations that were not tracked by night-averaged zero points. We emphasize again, however, that these are systematic differences with respect to 2MASS in this limited number of cases; the r.m.s. scatter about these median differences are ~ 0.1 mag in all bands and this should provide confidence that the internal consistency of the photometry in each field is preserved. We then decided to rescale our photometry to the 2MASS photometric system to remove these systematic effects. The (J-H) and (H-K) color differences between 2MASS and our photometry are not correlated with the magnitude, so that no magnitude-dependent color effect is introduced in this rescaling.

2.2 Results

2.2.1 Cluster Identification

The identification of a cluster results from the analysis of stellar density in the field. Since our target fields are sites of massive star formation associated with local peaks of dust column densities and hence of visual extinction, the K_s images are clearly more suited for this type of analysis.

Stellar density maps were compiled for each field by counting stars in a running boxcar of size equal to 20″. The box size was determined empirically to enhance the statistical significance of local stellar density peaks and to maximize the ability to detect the clusters. Larger boxes tend to smear the cluster into the background stellar density field decreasing the statistical significance of the peak, which may lead to non-detection of a clearly evident cluster, particularly in the rich inner Galaxy fields. Smaller boxes produce noisy density maps where the number of sources in each bin starts to be comparable to the fluctuations in the background density field caused either by intrinsic variations in the field star density or to variable extinction from diffuse foreground ISM in the Galactic Plane (where all of our sources are located). For most of our objects in the outer Galaxy, this analysis is used to locate the position of the peak stellar density, since the clusters are obvious already from visual inspection. For the remaining fields, the density maps are used to ascertain the presence of a cluster; toward the inner Galaxy in particular, the density maps tend to show more than one peak at comparable levels. It is important to remember, however, that this is a search for stellar clusters toward regions where indications of active star formation are already available, and this information can be used. In particular, the coincidence of these peaks with cold dust clumps traced by intense submillimeter and millimeter emission (Beltrán et al. 2006; Molinari et al. 2008a) is critical before we can consider the density peak to be a true feature associated with the star formation region. Casual association is excluded by the high number of positive associations (see Table2.1).

As further confirmation of the positive detection of a cluster we compiled radial stellar density profiles where stars were counted inside annuli of increasing internal radius and constant width and then divided by the area of the annuli (Testi et al. 1998); uncertainties were assigned assuming Poisson statistics for the number of stars in each annulus. We then assigned a positive cluster identification if the radial profile exhibited at least two annuli that had values above the background. To refine the location of the density peak, we repeated the radial density profile analysis starting from several locations within 10″ of the peak derived from the density maps; the location that maximizes the overall statistical significance of the annuli was then assigned to the cluster center. Figure 2.1 shows the typical footprint of a cluster, where the stellar density is plotted as a function of distance r from the start location; the density has a maximum at r=0 and decreases until it reaches a constant value, which is the average background/foreground stellar density.

There were two exceptions in this analysis. The first was for source Mol160. The K_s-band image shows clear stellar density enhancement in a semi-circular annulus surrounding the northern side of the dense millimeter core, which appears devoid of stars. This stellar density enhancement is coincident with the emission patterns visible in the mid-IR (Molinari et al. 2008b), so is clearly a stellar population associated with the star-forming region. Since the millimeter peak is at the center of symmetry of the semi-circular stellar distribution, we consider this tobe the center of the cluster. This is only for completeness,

since we cannot say whether the low density of stars at the millimeter peak is an effect of extreme visual extinction or reflects an intrinsic paucity of NIR-visible forming stars, as the proposed extreme youth of the massive YSO accreting in its depth would seem to suggest (Molinari et al. 2008b).

The second exception was for source Mol8. The stellar density analysis shows two peaks that are coincident with two distinct dust cores; we therefore assumed the presence of two distinct clusters, rather than a subclustering feature within the same cluster. The radial density profile analysis could not be used here, so we fit elliptical Gaussians to the peaks in the density maps, allowing for an underlying constant level representing the background stellar density. The resulting cluster richness was obtained by integrating the fitted Gaussian, and the cluster radius was taken to be equal to the fitted FWHM (the fitted Gaussians were nearly circular).

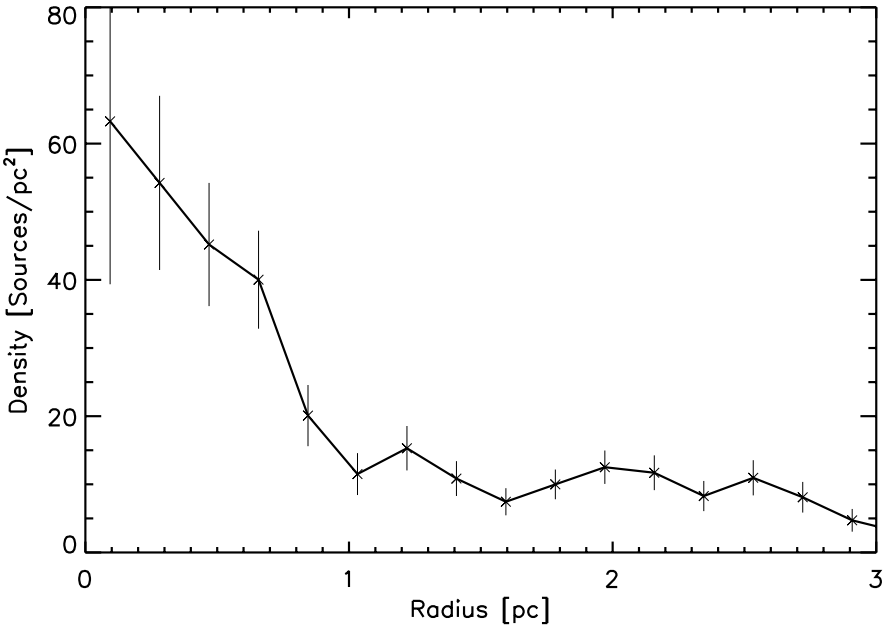


Figure 2.1. Stellar density (in stars/pc^2), for Mol28, as a function of the radial distance (in parsecs) from cluster center. Error-bars are computed as the Poissonian fluctuations of source counts in each bin.

Always following Testi et al. (1998), we determined the richness indicator of the cluster I_c by integrating the background-subtracted density profile; the cluster radius was taken to be the radial distance from the start location where the density profile reaches a constant value. This richness indicator is a very convenient figure to use when no detailed information is available for each single star in the region and the membership of the cluster cannot be established for each single star. These values are reported in Col. 3 of Table2.1 for all fields where a cluster has been clearly revealed. Column 1 gives the target name (cf. Table 1.1); its kinematic distance is listed in Col. 2. The parameter N_{obs} (Col. 4) is

the number of cluster members derived (see Sect. 2.3.1 below) from the integration of the background-subtracted K_s luminosity function (hereafter KLF, see Sect. 2.3.1). Also reported in Col. 8 is the mass of the hosting molecular clump; this was derived from the cold dust emission as reported in Molinari et al. (2008a, 2000), integrated over the entire spatial extent of the cluster; conversion into masses was achieved based on the optically thin assumption and by assuming T=30 K, $\beta = 1.5$ (Molinari et al. 2008a), and a mass opacity $\kappa_{230GHz} = 0.005 cm^2 g^{-1}$ which corresponds to a gas/dust weight ratio of 100 (Preibisch et al. 1993). The IRAS source bolometric luminosity, Col. 9, is taken from Molinari et al. (1996, 2000, 2002, 2008a); in Col. 10 we list the A_V at the peak cluster position estimated from submm observations (Molinari et al. 2008a, 2000). In Cols. 11 and 12 the coordinates of the centers of the identified clusters are reported. Columns 6 and 7 contain parameters that are described later in the text (see Sect. 2.2.2).

Table 2.1. Results for cluster detection

Sou.	d[a]	I_c	N_{obs}	R_{clu}	Pre-MS		M_{gas}	L_{bol}	$A_{V\ peak}$	Cluster Center	
Mol	kpc			pc	*CC*(%)	*CM*(%)	$M_\odot$	$10^3 L_\odot$	mag	α(J2000)	δ(J2000)
3	2.17	78	78	1.7	34	99	910	12.4	18	00:44:57.4	+55:47:20.0
8A	11.5	25	30	1.3	37	27	1650	57.0	18	05:17:13.8	+39:22:29.7
8B	11.5	27	24	1.3	9	4	1780	5.5	8	05:17:12.0	+39:21:51.8
9	6.2	7	7	0.6	16	–[g]	–	24	–	05:20:16.9	+36:37:22.0
11	2.1	51	48	0.5	12	41	360	4.6	35	05:37:47.7	+31:59:24.0
12	1.6	12	13	0.3	0	30	72	1.6	46	05:40:24.4	+23:51:54.8
15	1.5	64	61	0.3	34	–[g]	–	5.8	–	06:08:41.0	+21:31:00.0
28	4.5	75	75	1.0	58	95	220	9.1	4	07:00:51.5	−08:56:18.2
30	0.3	200[e]		no cluster detected[b]			-	0.14	–	–	–
38	0.5	−5		no cluster detected[c]			-[h]	0.19	40	–	–
45	4.3	28	27	0.5	68	19	1340	21.2	77	18:17:24.1	−17:22:12.3
50	4.9	46	43	0.6	36	41	80	17.3	25	18:19:07.6	−16:11:21.0
59	5.7	0		no cluster detected[c]			-[h]	11	29	–	–
75	3.9	8	7	0.7	37	24	1310	13.3	38	18:53:38.1	+01:50:26.5
82	6.8	8	10	0.3	0	52	590	15.4	52	18:59:03.2	+03:53:16.7
84	2.2	21	21	0.3	0	54	28	4.3	15	18:59:14.3	+07:04:52.3
98	4.5	6		no cluster detected[f]			-[h]	9.2	68	–	–
99	6.1	45	38	0.9	49	–[g]	–	37.3	–	19:11:51.4	+09:49:35.4
103	4.1	105	107	0.7	46	87	510	28.2	42	19:23:36.2	+17:28:58.1
109	4.3	19	17	0.5	[d]	[d]	1030	26.7	90	19:39:33.0	+24:00:21.3
110	4.3	23	20	0.5	[d]	[d]	400	14.8	55	19:40:58.5	+24:04:36.3
136	3.6	21	19	0.6	18	52	230	4	21	21:32:31.4	+51:02:23.1
139	7.3	25	24	1.2	10	31	1870	1.35	20	21:53:39.2	+56:27:50.7
143	5.0	25	22	0.8	10	76	630	7.8	26	22:19:09.0	+56:04:58.7
148	5.1	43	41	0.9	43	64	22	7.8	13	22:32:23.4	+58:19:01.3
151	5.4	15	14	0.9	12	30	2020	25	40	22:52:38.3	+60:00:44.6
160	5.0	36	34	1.3	30	76	1830	16	32	23:40:53.1	+61:10:21.0

[a] Kinetic distance using the rotation curve from Brand & Blitz (1993).

[b] Stellar density analysis inconclusive due to extreme crowdedness of this field.

[c] Stellar density reveales no peaks close to the IRAS position or the submm peak.

[d] Only observed in K_s.

[e] Detection refused due to extreme field complexity (see text).

[f] Detection refused because only 1 annulus in the radial density profile is above background (see text).

[g] No extinction estimate is available due to lack of submm information to evaluate de-reddening correction.

[h] Extinction estimate is available from single-pointing submillimeter data (Molinari et al. 2000) but not from maps, so that a reliable clump mass estimate is not possible.

Following the procedure described, a cluster was detected within 1′ of the IRAS position for 22 out of the 26 observed fields (85% detection rate). In two cases (Mol38 and Mol59), the stellar density map does not show a clear peak above the fluctuations of the field stellar density. For Mol 98, the radial density profile only shows one annulus above the background, and therefore fails the criterion that the stellar density enhancement should be resolved significantly above the background in two annuli. In one case (Mol30), several stellar density peaks were found in proximity to the IRAS source, but the lack of information about the submillimeter/millimeter continuum prevents us from drawing any firm conclusion.

Figure 2.2 shows I_c as a function of the peak A_V and suggests that with higher dust extinction, we may find it more difficult, or it becomes less likely, to detect a cluster at 2.2 μm.

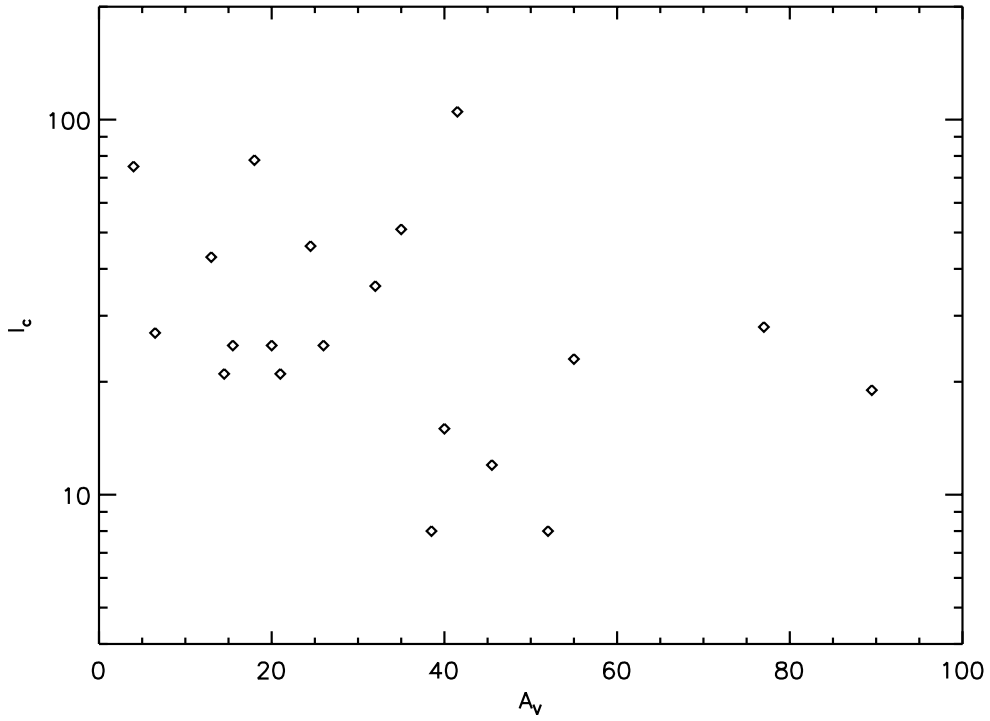


Figure 2.2. Cluster richness indicator I_c as a function of A_V at the cluster center; for a few detected clusters, we do not have an estimate of A_V).

Our detection rate is quite high and this implies that young stellar clusters in sites of intermediate and massive star formation are ubiquitous. While this was established for relatively old Pre-MS systems such as Herbig Ae/Be stars (Testi et al. 1999), we hereby verify that this is also true in much younger systems, where the most massive stars may even be in a pre-Hot Core stage (Molinari et al. 2008a).

Our detection rate is higher compared to other similar searches of stellar clusters toward high-mass YSOs. For example Kumar et al. (2006) used the 2MASS archive and reported a rate of 25% (rising to 60% when neglecting the inner Galaxy regions) toward a larger sample, which also includes the sources of this work; in particular, we detect all clusters

also detected by Kumar et al. and in addition we reveal clusters toward 13 objects for which Kumar et al. report no detection. The reason for this discrepancy may be because we obtained dedicated observations, while Kumar et al. used data from the 2MASS archive; the diffraction-limited spatial resolution of our data is between a factor of 4 and a factor of 10 better with respect to 2MASS, and this certainly facilitates cluster detection especially in particularly crowded areas such as the inner Galactic plane. To test this hypothesis, we degraded the NTT K_s image of Mol103, also considered in Kumar et al., to the 2MASS resolution; extraction and photometry were performed as outlined above but the search for a cluster based on the stellar radial density profiles revealed no cluster. The estimated number of members (corrected for the contribution of fore/background stars) for 7 out of the 10 clusters detected both by us and by Kumar et al. was at least a factor of two less in the latter study.

Kumar & Grave (2008) conducted a similar study on a large sample of high-mass YSOs, that included some of our sources, using data from the GLIMPSE survey (Benjamin et al. 2003). They detect no significant cluster around any targets in a sample of 509 objects. As the authors say in their paper, however, GLIMPSE data are sensitive to 2-4 $M_\odot$ pre-main sequence stars at the distance of 3 kpc. Based on color-magnitude analysis (see later below), our mass sensitivity is of the order of 1 $M_\odot$ at a distance of 3.6 kpc and ~0.6 $M_\odot$ at a distance of 2.1 Kpc. Probing longer wavelengths, GLIMPSE is likely to be more sensitive to younger sources compared to the classical J, H, K range, which also samples relatively older pre-MS objects. The combination of sampling higher-mass (and hence rarer stars because of the shape of the IMF) and relatively younger stars (which, as indeed our analysis finds, may not be the majority in a young cluster) may plausibly be the reason for the negative cluster detection results of Kumar & Grave.

The distribution of the radii of the detected clusters indicated by a full line in Fig.2.3; the median value is 0.7 pc. The dashed histogram (which refers to the upper X-axis) shows the distribution of the cluster richness indicator I_c, with a median number of stars of 27. We note that the value of I_c for many of our clusters is less than the limit of 35 suggested by Lada & Lada (2003) to be a *bona fide* cluster. This definition stems from the argument that a less rich agglomerate may not survive the formation process as an entity. Our interest, however, is to investigate the spatial properties of the young stellar population in a star-forming region at the time of active formation, without worrying about its possible persistence as a cluster at the end of the formation phase. However, we prefer not to introduce a new term to identify the structures that we see and still use the term cluster, although in a milder way than Lada & Lada.

2.2.2 Properties of identified clusters

We first derive qualitative measurements related to the nature of the identified clusters using simple diagnostic tools such as color-color and color-magnitude diagrams. These diagrams have been drawn for all detected clusters and are available in electronic form; we illustrate here the particular case for Mol28.

Color-color analysis

Figure 2.4 shows the [J-H] *versus* [H-K_s] diagram for all sources detected within a distance equal to R_{cl} centered on the stellar density peak. The full circles represent all sources

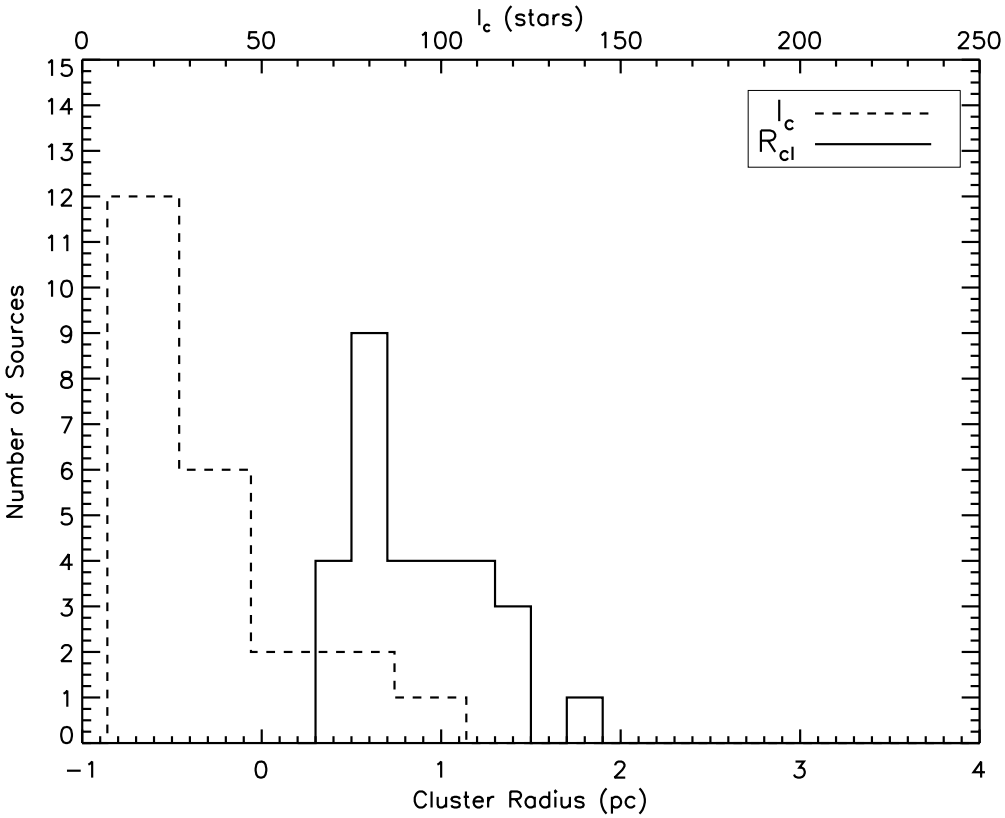


Figure 2.3. Distribution of the cluster radii in parsecs (full line) and the cluster richness indicator I_c in number of stars (dashed lines); the median values for the two distributions are 0.6 pc and 37 stars, respectively

detected in all three bands, the arrows representing sources with lower limits (to their magnitude) in the J band. The plot shows more stars than the I_c value reported in Table 2.1 because we also include the fore/background stars that cannot be individually distinguished from the true cluster members. A significant fraction of the sources have colors compatible with main-sequence stars that have a variable amount of extinction reddening (computed by adopting the Rieke & Lebofsky (1985) extinction curve), but many sources have colors that are typical of young pre-MS objects with an intrinsic IR excess produced by warm circumstellar dust distributed in disks (Lada & Adams 1992). The set of dotted curves represents the locus of two-component black bodies with temperatures as indicated at the start and end of each dotted line; along each curve, the relative contribution of the the two black bodies is varied. These curves mimic the effect of a temperature stratification in the dusty circumstellar envelopes, and the presence of sources in the area covered by these curves is an indication of the presence of warm circumstellar dust.

A straightforward indication of the youth of the cluster may be provided by the fraction of sources that are not compatible with being reddened MS stars, i.e., those with IR excess. The number of stars with an IR excess is normalized to the total number of stars detected in the cluster area, corrected for the expected number of fore/background stars estimated from the areas surrounding the cluster (but still in the same imaged field). To be conservative we extend the region of the MS by 0.2 magnitudes to the right corresponding to about a 2σ uncertainty in measured magnitudes. This ratio is reported as a percentage value in Col. 6

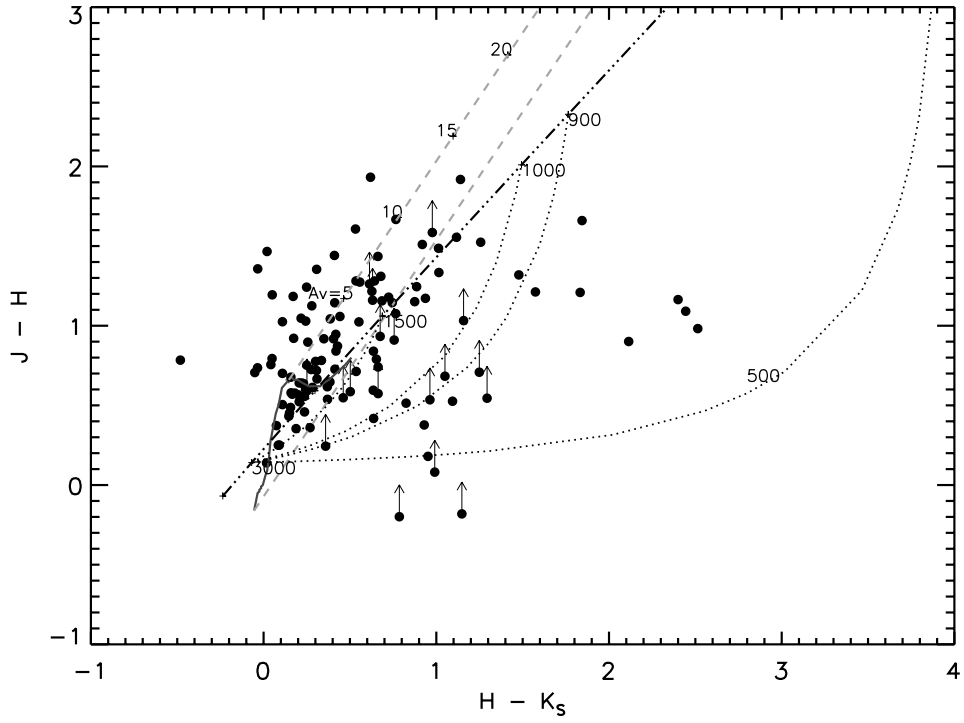


Figure 2.4. [J-H] *vs* [H-K_s] diagram for Mol28. Upward pointing arrows are the sources not detected in J. The continuous curve at the bottom-left represents the main sequence, while the dashed grey lines represent the effect of reddening (Rieke & Lebofsky 1985) for variable amounts of extinction as indicated along the lines. The dashed-dotted black line is the black-body curve, and the dotted curves are two-component black-body curves with varying relative contribution (respectively, from the inner to the outer curve, 3000-1500K, 3000-1000K, 3000-900K and 3000-500K).

of Table 2.1.

Color-magnitude analysis

Additional evolutionary indications of the detected clusters may be derived from the K_s-[H-K_s] diagram, reported for Mol28 in Fig. 2.5. Compared to the main sequence (the leftmost almost vertical curve in the figure) a significant fraction of the sources are on its right, where the evolutionary tracks for Pre-MS sources (Palla & Stahler 1999) can also be found, and could therefore be interpreted as very young pre-MS objects. The distribution of sources in the diagram spans a much larger region than that covered by the Pre-MS isochrones, because of the combined effect of extinction reddening and IR excess. The extinction effects can be seen from the dotted lines originating in the main sequence and extending toward the bottom-right for increasing values of A_V. On the other hand, the presence of a warm dusty circumstellar envelope implies an increase in both absolute emission and SED steepness, which would shift a pure photosphere toward the top-right of the diagram (as shown by the arrow labeled 'IREX' in Fig. 2.5). In a similar way to the color-color analysis, it is impossible to estimate the age of individual stellar sources based on their location on the pre-MS isochrones, because we do not know the amount of A_V by which

we should de redden each object. We follow a conservative approach by dereddening each object using half of the exctinction estimated for each location from millimeter maps; this corresponds to placing each object midway through the clump.

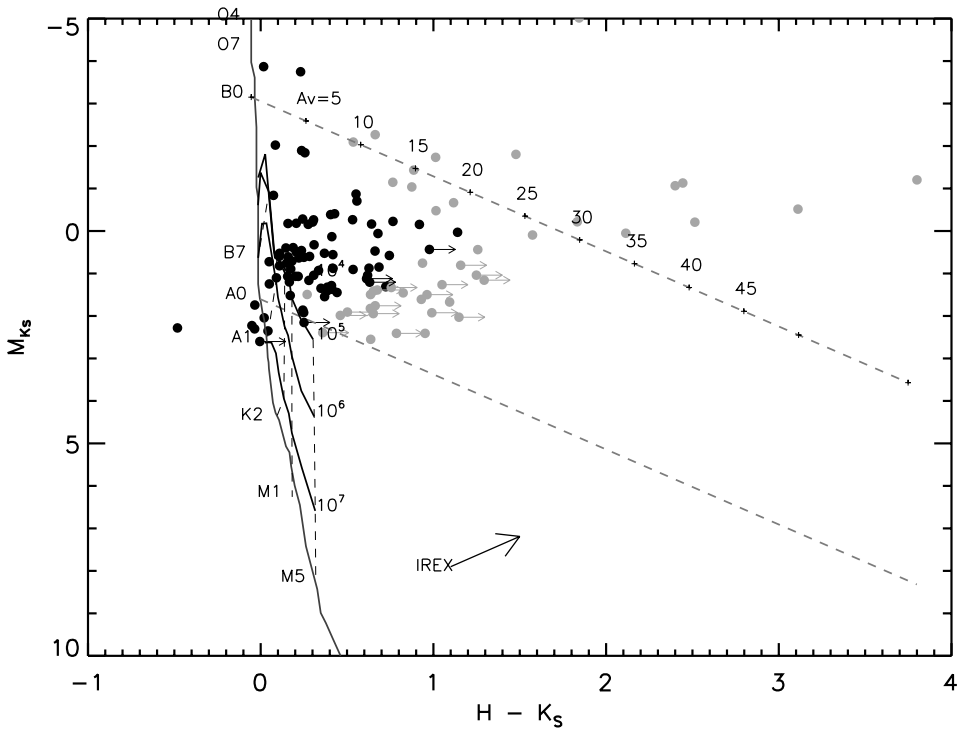


Figure 2.5. K_s *vs* [H-K_s] diagram for Mol28. The leftmost curve represents the main sequence, while the dashed lines represent the effect of reddening for variable amounts of extinction. Isochrones from Palla & Stahler (1999) are also indicated with full lines for different Pre-MS ages. The arrow labeled IREX indicates the direction of change due to IR-excess (see Sect. 2.3.2). Rightward pointing arrows represent those sources not detected in H. Symbols in grey color indicate sources with IR excess as determined from the color-color diagrams (see Fig. 2.4).

A further correction is to remove the IR excess for those sources, which is apparent in the color-color diagram (of fig. 2.4), estimated using the formulation suggested by Hillenbrand & Carpenter (2000), and used later in this work (see Sect. 2.3.2). The ratio of pre-MS stars to the total in each cluster area will remain contaminated by fore/background stars; to estimate this contamination, we choose an off-cluster area in the same imaged field and simply compute the ratio of sources with pre-MS colors to the total (in these off-cluster regions in which there is no significant reddening to correct for). for each cluster, Col. 7 of Table 2.1 reports the fraction of stars (detected in the cluster area in all three bands) situated more than 0.2 mag to the right of the MS after the various corrections have been applied.

2.3 Initial mass functions and star formation histories

As is apparent from the qualitative analysis presented in the previous paragraphs, the diagnostic power of our observations is limited because we do not know which objects in the

cluster area are true cluster members nor the precise amount of dust extinction (originating within the hosting clump) and IR excess (originating in the immediate circumstellar environment) pertaining to each source. Without this detailed knowledge of individual stars in the clusters, fundamental quantities such as the initial mass function (IMF) and the star formation history (SFH) cannot be derived directly from, e.g., the K_s luminosity function (KLF). We are compelled to obtain these using statistical simulations of clusters based on different input parameters and performing a statistical comparison between synthetic and observed KLFs and HKCFs.

We first derive the observed KLFs from the observations. We then illustrate in detail the model used for the cluster simulations, exploring the sensitivity of the results to a wide range of input parameters finally, modeled and observed KLFs are compared to infer statistically the IMF and SFH for our clusters.

2.3.1 Observed K_s luminosity functions

The KLF of each cluster is obtained by simply counting all detected sources within the cluster area as identified from the cluster density profile (see Sect2.2.1). In a similar way to the other diagnistic tools (Sects2.2.2 and 2.2.2), the KLF is contaminated by field stars that cannot be individually identified. To account for the field star contamination in a statistical way we subtract from the KLF that was compiled for the cluster area, the KLF for a region outside the cluster area but still in the same imaged field, after normalising the different areas. The regions in which the field star KLF is compiled have a lower extinction with respect to the cluster KLF, so the background contribution to the cluster KLF is likely to be overestimated. Field-subtracted KLFs for all clusters are available online[2].

The integral of the KLF provides an independent estimate of the number of cluster members, and these values are reported as N_{obs} in Table 2.1. Their agreement with the richness indicator I_c confirms the consistency of our analysis. All KLFs show a dominant peak that is always close to the completeness limit, showing that our observations are insufficiently sensitive to the low-mass stellar component of our clusters. Many of the KLFs present a separate small peak at low magnitudes (one or two sources at most, on average). Could this be caused by confusion because of source crowding and insufficient spatial resolution ? For each cluster, we studied the distribution of distances of each star from its nearest neighbour and found that there are two types of distributions, reported in Fig.2.6. In the first type (full line in figure), the distribution has a peak corresponding to an inter-star distance significantly higher than the value corresponding to half the PSF FWHM (the full vertical line); in this case, the suggestion is that all cluster members have been resolved from their neighbour. In the second type (dashed line in the figure), the distribution has its peak very close to half the PSF's FWHM (the dashed vertical line), indicating that source blending should certainly be considered possible. We verified that all clusters with a distance distribution of the second type do exhibit a second faint peak at high brightness in their KLFs, therefore confirming that this feature is an artifact of the relatively low spatial resolution, which in some cases is insufficient to resolve all cluster members.

[2] at http://galatea.ifsi-roma.inaf.it/faustini/KLF/

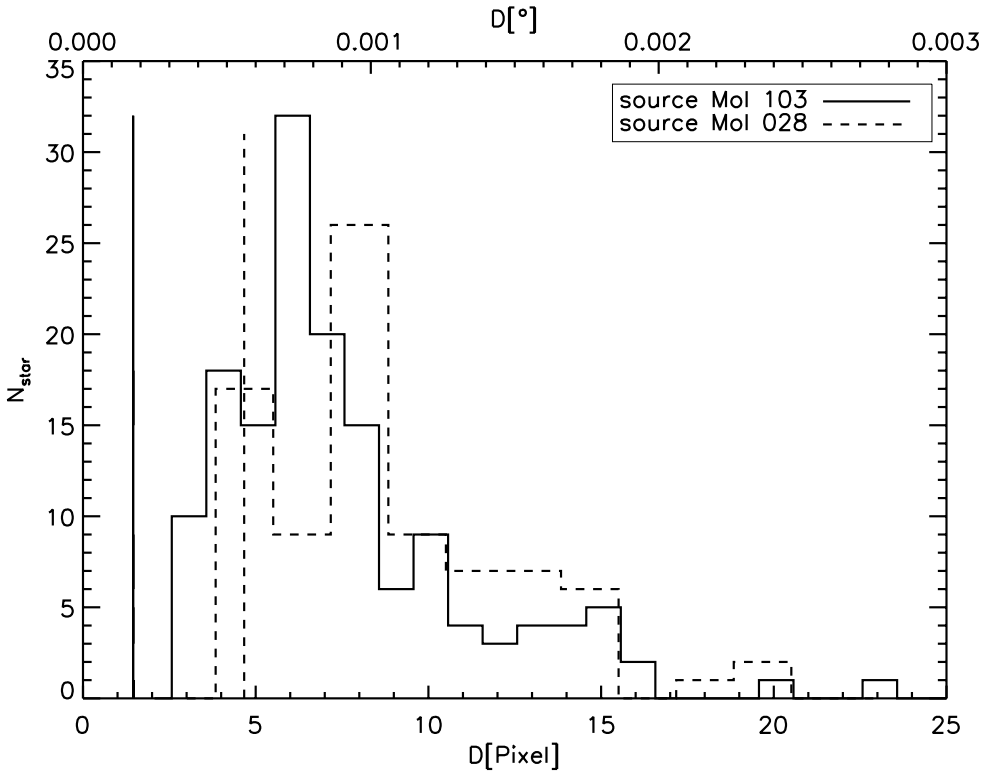


Figure 2.6. Distribution of identified sources as a function of nearest-neighbor distance (D) for two of our examined fields (Mol28 dashed line and Mol103 full line).

2.3.2 Synthetic KLF. Synthetic cluster generator: a near-IR cluster simulator

As already mentioned, we cannot derive masses and ages from our data alone. We thus developed a model to create statistically significant cluster simulations obtained for different assumptions of IMF and SFH (source ages and their distribution), and compare the synthetic KLFs with the observed field-subtracted KLFs. This model we called the synthetic cluster generator (SCG).

SCG: model description

A cluster is created by adding stars whose masses and ages are assigned via a Monte-Carlo extraction according to the chosen IMF and SFH; the pre-MS evolutionary tracks of Palla & Stahler (1999) are then used to convert them into J, H, and K_s magnitudes. The 3D distribution of stars is obtained by randomly choosing for each star a set of x,y,z coordinates using the observed stellar density profile (see Sect. 2.2.1), approximated to be a radially symmetric Gaussian, as weight-function; using submm continuum images, this is needed to assign the proper column of cold dust "required" to extinguish the near-IR radiation. Other analytical functions could have been used, e.g., a King profile, but the statistics of our clusters are insufficiently high to explore the effect of different radial profile assumptions.

To convert the submm flux into dust column density, we used the dust temperature and emissivity exponent β determined in Molinari et al. (2000); mean values from the latter

work were adopted for those fields not covered by our work.

To properly simulate the pre-MS stars, we also need to include the effect of an IR excess caused by warm dust in the circumstellar envelopes and disks. We used the distribution (modeled as a Gaussian) of $[H\text{-}K_s]_{ex}$ color excesses as measured for a sample of Pre-MS stars in Taurus, as used by Hillenbrand & Carpenter (2000), as a weight-function to randomly assign a $[H\text{-}K_s]_{ex}$ to each simulated star in our model; the K_s *vs* $[H\text{-}K_s]_{ex}$ relationship adopted in the above mentioned work was then used to derive the H and K_s excess-corrected magnitudes. The K_s magnitude of the synthetic star was then compared with the limiting magnitude typical of the cluster being simulated to determine whether the star could have been detected in our observations. This procedure is repeated until the number of synthetic detectable stars equals the value of I_c determined for our observations; at this point, the cluster generation process is complete.

Since the simulation is based on Monte Carlo extraction of stellar mass, age, and position in the cluster, each independent run for a fixed set of input parameters can in principle result in very different outputs in terms of cluster luminosity, total stellar mass, maximum stellar mass, and synthetic KLF. To determine the statistical significance, the model is run 200 times for any given set of input parameters, and the median KLF is later adopted for comparison with the observed one. Clearly, the predictive power of this simulation model resides in its capability to characterize the cluster properties of any given parameter set. In other words, the distribution of the resulting quantities should not be uniform but peaked around characteristic values. We return to this point in Sect.2.3.2

SCG: input assumptions

We tested three different assumptions about the star formation histories in our cluster simulations. The first was to assume that stars in the cluster formed in a single burst-like event (hereafter SB) some t_1 years ago. The explored range in the simulations is $10^3 \leq t_1 \leq 10^8$ yrs. The second was that the formation of stars proceeds at a constant rate (hereafter CR) from a time t_1 years ago to a time t_2 years ago. The ranges explored in the simulations are $10^4 \leq t_1 \leq 10^8$ yrs and $10^3 \leq t_2 \leq 10^7$ yrs, where we always assume that $t_1 > t_2$. The third possibility that we explored was a variation in the previous assumption, where the star formation rate is not constant but varies with time as a Gaussian function (hereafter GR). Within the boundaries of the start and end of the star formation process, t_1 and t_2 that were varied as above, we also varied both the time t_c of the Gaussian peak in the range $10^{3.7} \leq t_c \leq 10^{7.7}$ and $\mathrm{Log}_{10}(\sigma)$ of the Gaussian-like SFH, which was allowed to have one of two values 0.1 and 0.5.

We allowed three different choices of IMFs, i.e.,Kroupa (1993); Scalo (1998); Salpeter (1955), with the latter modified by introducing a different slope for $M<1\ M_\odot$ coinciding with that of the Scalo (1998) IMF; the three IMFs were labeled IMF1, IMF2, and IMF3, respectively. The IMF from Kroupa et al. provides a more accurate description of the low-mass end of the distribution, while the classical Salpeter IMF is flatter at low mass but heavier at intermediate and high masses (above 1 $M_\odot$). The properties of the Scalo IMF is in-between the other two, resembling Salpeter's one below 1 $M_\odot$ and above 10 $M_\odot$, and Kroupa's for $1\ M_\odot<M<10\ M_\odot$.

SCG: predictive power

To verify our model's predictive power, we completed 200 simulations for a cluster with a Salpeter IMF and a constant star formation rate with $t_1=10^6$ yrs and $t_2=10^4$ yrs. Figure 2.7 shows the distribution of the predicted number of stars and the total luminosity for the 200 simulations. The number of cluster members shows very little variation, as expected since the number of detectable stars is the parameter that we use to stop the simulation; on the other hand, the distribution of the total luminosity is not particularly peaked, as the central 3 bins containing about 60% of the simulations span almost two decades in luminosity.

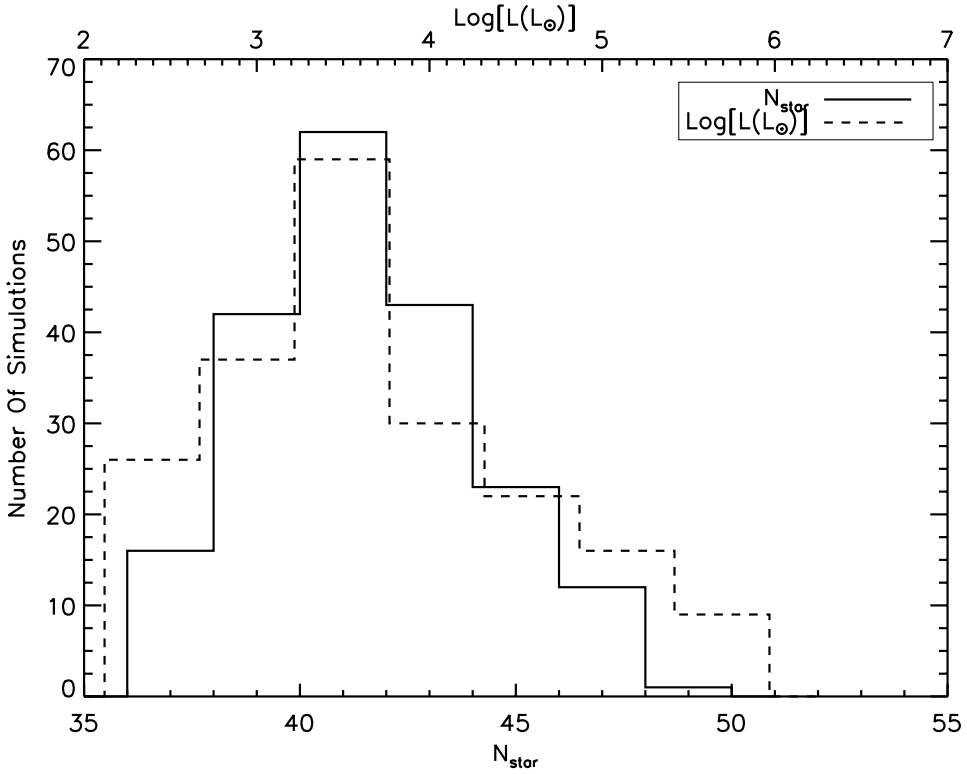


Figure 2.7. Distribution of the predicted number of cluster members (full line) and total luminosity (dashed line) for 200 SCG runs for Mol160 with a Salpeter IMF and a constant star formation rate with $t_1=10^6$ yrs and $t_2=10^4$ yrs.

On the other hand, the distributions for the total cluster stellar mass, and for the mass of the most massive member (see Fig. 2.8) are rather peaked and highlight a relatively higher predictive power of the model for these two quantities. It is to be noted that the distributions are rather skewed, suggesting that neither the mean nor the median are particularly suited to characterize the peak of the distribution. We indeed found that these quantities assume at their distributions peak a more representative value of mass and use them in the following discussion.

Concerning the reproducibility of the KLF, for each of the 200 runs the resulting KLF was fitted with a Gaussian function and the center, peak and σ were determined. Figure2.9 reports the distribution of these three parameters for the 200 runs and shows that all of them are remarkably peaked and symmetric. The formal r.m.s. spread for the three quantities, estimated via a Gaussian fit to the distributions in the figure, is $\approx$ 0.3 mag for the KLF

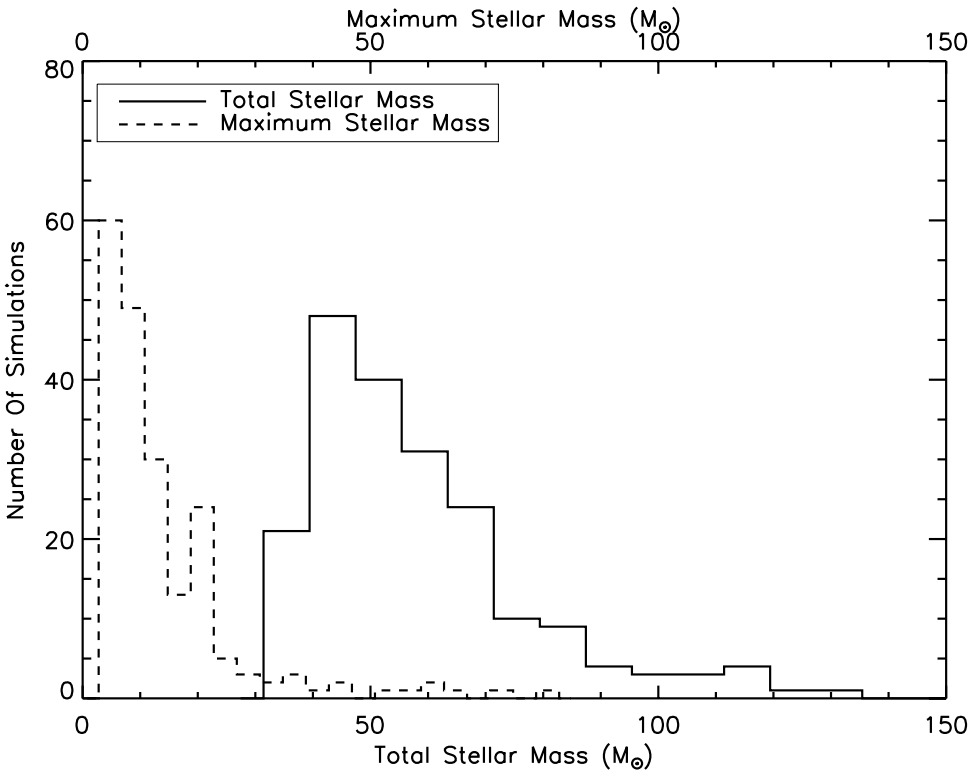


Figure 2.8. Distribution of the predicted total stellar mass (full line) and mass for the most massive star (dashed line) in a cluster for 200 SCG runs for Mol160 (same inputs as in Fig. 2.7).

center, ≈ 12% for the KLF peak (about 1.2 sources out of a mean KLF peak of 10), and ≈ 0.25 mag for the KLF FWHM.

We completed a similar analysis for HKCF (H-K_s color function; see Sect. 2.2.2). Figure 2.10 shows the distribution of Gaussian function centers, peaks and σ's for HKCFs obtained for the same 200 runs used previously for the KLFs. Gaussian fits to the three distributions in the figure infer an r.m.s. that is ≈ 0.15 mag for the HKCF center and ≈ 0.14 mag for the HKCF FWHM, while the "peak" distribution is flatter and has an r.m.s. value of ≈ 21% for the HKCF peak (about 3.2 sources out of a mean HKCF peak of 15). It is worthwhile to stress that since the position that is assigned to each simulated star in the cluster is different in each of the 200 runs of the model (for any given set of input parameters), the scatter in the properties of the synthetic KLFs and HKCFs also statistically tends to account for the effects of extinction variations in the cluster's hosting clump, which may in principle be relevant in such heavily embedded systems (see Table2.1).

For a given set of input parameters, we conclude that the model results, have a good reproducibility, except concerning the total luminosity. The model therefore has a strong predictive power concerning the median properties of a synthetic cluster. The spread in KLF center magnitudes is indeed, less than the bin amplitude used in compiling the KLFs for the simulations (and is used in the remainder of the work); the median synthetic KLF therefore provides a good representation of the cluster luminosity distribution.

In conclusion, 200 simulation runs for each combination of input parameters (IMF and SFH) can provide a robust assessment of the statistical significance of the synthetic observ-

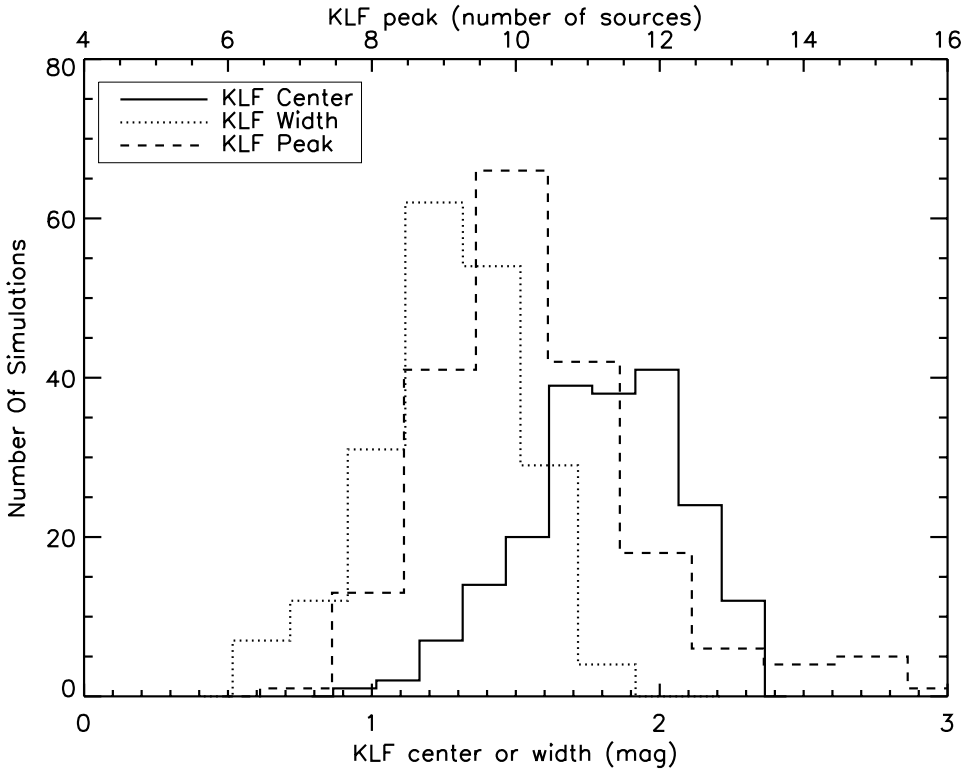


Figure 2.9. Distribution of the predicted center magnitude (full line - bottom X-axis scale), width (dotted line - bottom X-axis scale) and peak value (dashed line - top X-axis scale) of the predicted Gaussian-fitted KLFs for 200 SCG runs for Mol160 (same inputs as in Fig. 2.7).

able properties (KLFs and HKCFs). Although the distributions for the KLFs' (HKCFs') parameters seem rather symmetrical, we adopt the median KLF (HKCF) of the 200 runs as a more reliable characterization for that particular parameters' set. The use of the mean KLF (HKCF) for the comparison does not significantly alter the results.

Exploring the SCG parameter space: cluster parameters

After verifying the robustness of model results in independent runs for the same input parameters, we now measure the sensitivity of the model results to changes in these parameters. We first concentrate on simulated cluster physical parameters (number of cluster members, total luminosity, stellar mass distribution), and in the next paragraph we examine how the KLFs and the appearance of the color-magnitude diagrams, which are the main observables used in our analysis, behave in this respect.

Number of stars N_{stars} - As a general rule, the older the cluster is allowed to be, irrespective of the detailed SFH adopted, the higher is the number of produced stars. This is easily understood since the SCG cluster formation is stops when the number of the K_s-detectable stars equals the number of observed objects; if a cluster is old, the stars will be intrinsically fainter due to the shape of Pre-MS tracks and statistically less likely to extract stars bright enough to be detectable. As long as $t_l \leq 10^6$ yrs, N_{stars} does not depend significantly on

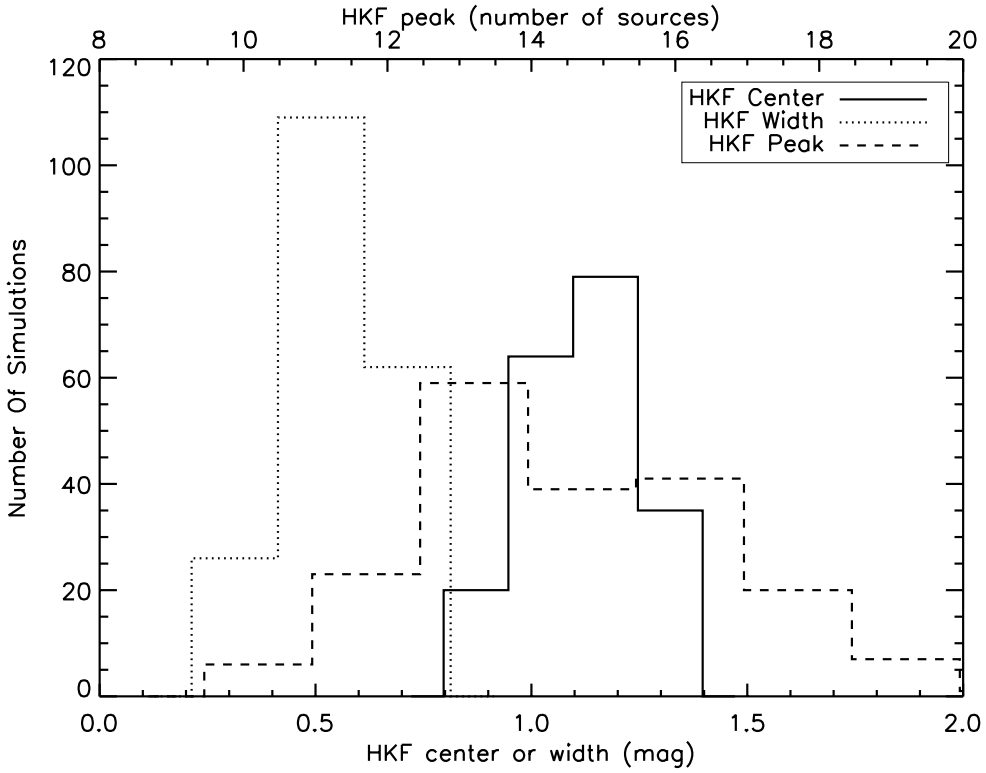


Figure 2.10. Distribution of the predicted center color (full line - bottom X-axis scale), width (dotted line - bottom X-axis scale) and peak value (dashed line - top X-axis scale) of the predicted Gaussian-fitted HKCFs for 200 SCG runs for Mol160.

the IMF choice, while for older systems, IMF1 (Kroupa 1993) produces nearly twice as many stars as IMF3 (Salpeter 1955) with IMF2 (Scalo 1998) in-between.

Stellar Masses - Likewise, the total stellar mass and the mass of the most massive star will be higher the older the cluster is allowed to be. If an IMF1 cluster is a very old SB or a CR with $t_1=10^8$ yrs and $t_2=10^7$ yrs for example, $M_{\star Tot}$ and $M_{\star Max}$ will be respectively a factor of 5 and 2 higher than clusters that are younger and/or are allowed to form stars until recent times (i.e., allowing a CR with $t_2 = 10^4$ yrs). The explanation follows directly from the argument made for the N_{stars} behavior above; matching the number of K_s-detected stars in a relatively old cluster with intrinsically fainter stars will require that stars will have to be on average more massive objects, and this will clearly result also in a higher total stellar mass.

Going from IMF1 to IMF3, both $M_{\star Tot}$ and $M_{\star Max}$ significantly increase, as expected. The trend of $M_{\star Tot}$ with cluster age is less pronounced because with IMF2 and IMF3 it is statistically more likely to produce relatively more massive (and hence more easily detectable in K_s) stars requiring a lower number of star extractions and hence a lower relative total mass at the end of the simulation. The age-trend of $M_{\star Max}$ is instead the same (only shifted toward higher masses) because the probability of extracting a massive star is the same for all ages and is only a function of the chosen IMF.

Total stellar luminosities and massive object luminosities - The total stellar luminosity, like the luminosity of the most massive star ($L_{\star Max}$), exhibits the same behaviour as $M_{\star Max}$. This is easily understood given the steep power-law dependence of the stellar luminosity on mass, and confirms that the total luminosity (L_{Tot}) is largely dominated by the most massive stellar object in the cluster: $L_{Tot} \propto L_{\star Max}$. Of great interest is the ratio of $L_{\star Max}$ to L_{Tot}; for the vast majority of clusters, its value varies between 0.6 and 0.8. This is further confirmation that global properties of our clusters are dominated by the most massive source. This ratio does not present any particular dependence on the value of $M_{\star Max}$, or of N_{stars}; only for the most populated clusters (clusters with of the order of a hundred members, such as Mol103, where the contribution of a great number of low-mass sources becomes important do we find a lower value for this ratio.

Exploring the SCG parameter space: KLF variations

We now briefly analyze the diagnostic power of the KLF and the HKCF against changes in IMF and SFH choices. Figure 2.11 shows the KLFs predicted for source Mol3 adopting the same SFH parameters (as indicated in the figure) and using the three different IMF choices.

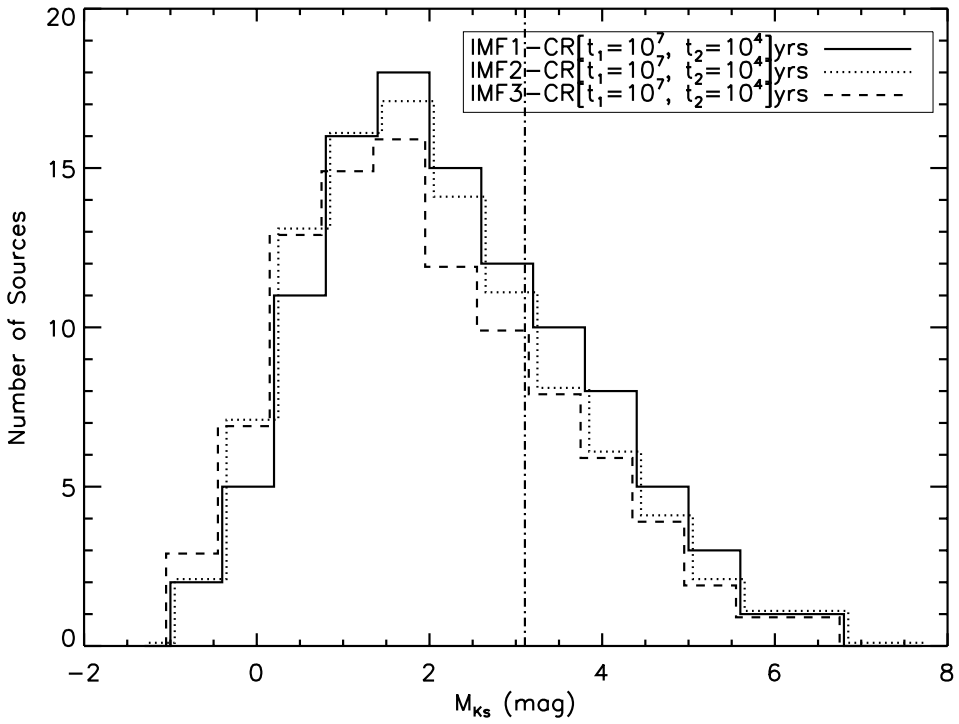


Figure 2.11. KLF (using the absolute K_s magnitudes) for Mol3 predicted by SCG for a CR cluster with $t_1=10^7$ yrs and $t_2=10^4$ yrs, for three different choices of the IMF (line styles as indicated in the figure). The dash-dotted line represents the completeness limit for this source given the K_s limiting magnitude.

The shape of the resulting KLF changes throughout the M_K range; going from Kroupa et al.'s IMF1 to Salpeter's IMF3, the distribution becoming more skewed toward lower magnitudes; this was expected since IMF1 produces more lower mass stars than IMF3. One

can certainly argue that the change is not dramatic, but on the other hand the modification does not affect one or two bins but the entire KLF consistently. The change is more apparent in the region between the peak and the completeness limit than at the bright end of the KLF, and for this reason the ability of the model to discriminate between different IMFs is higher for those sources, as Mol3 in the figure, where the KLF's peak is clearly detected above the completeness limit.

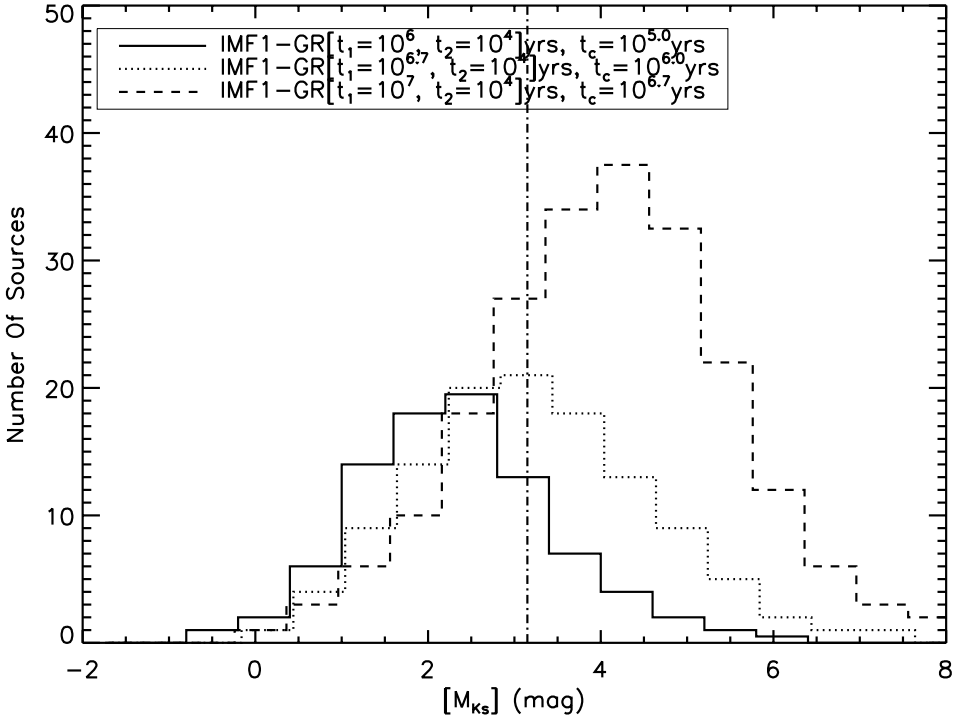


Figure 2.12. KLF (using the absolute K_s magnitudes) for Mol3 predicted by SCG for an IMF3 cluster with different ages as indicated in the plot. Older clusters produce a KLF peaked tower lower magnitudes. The dash-dotted line represents the completeness limit of this source given the K_s limiting magnitude.

The difference in predicted KLFs is much more dramatic if different age ranges are assumed, while keeping fixed the shape of the SFH and the IMF, as it is apparent in Fig.2.12. The peak of the KLF shifts considerably toward higher magnitudes as the median stellar ages (t_c) increase. A similar trend is observed by comparing SB models with different ages, although SB models always produce KLFs that are considerably narrower than CR or GR models. Older cluster ages would shift the peak of the KLF beyond the completeness limit; in other words, our analysis is insensitive to ages for the majority of stars in excess of $\sim 5\$\times 10^6 \div 10^7$ yrs; these old cluster ages would be hard to justify given that they are still heavily embedded in dense clumps.

Finally, we briefly show how the KLFs change for different choices of the SFHs. Figure 2.13 shows the KLFs obtained for a SB with $t_1=10^6$ yrs, compared with a CR with $t_1=10^7$ yrs and $t_2=10^4$ yrs and a GR with the same start and end star formation period, and with a peak times for star formation rate of $t_c=10^6$ yrs. The KLFs are clearly different, with

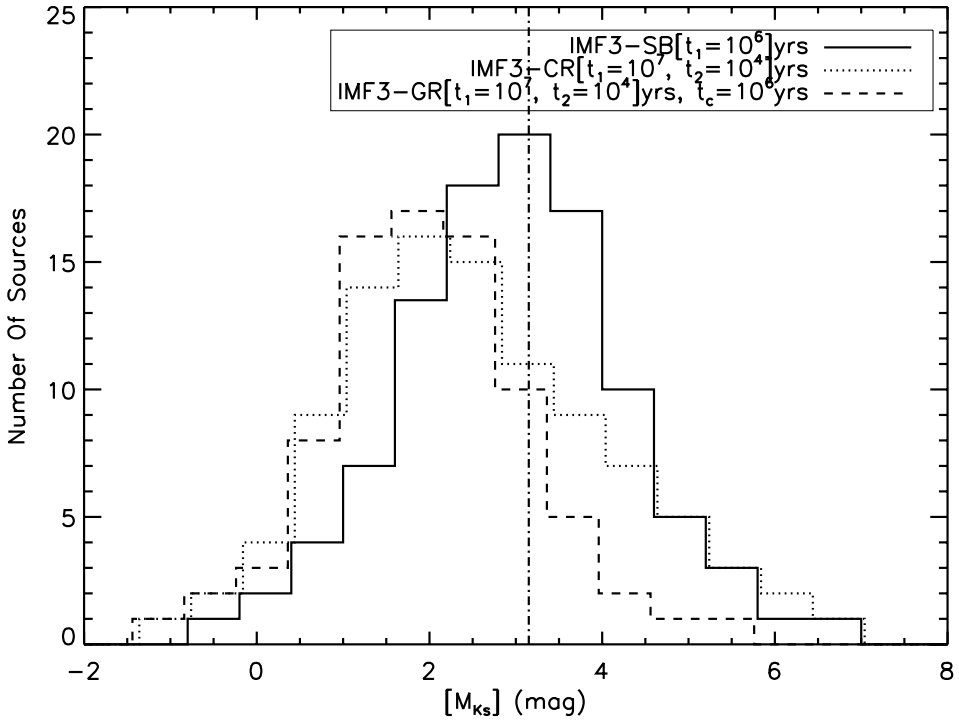


Figure 2.13. KLF (using the absolute K_s magnitudes) for Mol3 predicted by SCG for an IMF3 SB cluster with $t_1=10^6$ yrs (full line), CR with $t_1=10^7$ yrs and $t_2=10^4$ yrs (dotted line), and GR with same t_1 and t_2 and $t_c=10^6$ yrs (dashed line). The dash-dotted line represents the completeness limit for this source given the K_s limiting magnitude.

a peak magnitude which is quite sensitive to the formation rate typology and the peak time for star production.

The [H-K_s] color functions are found to be insensitive to the choice of IMF. As for the KLFs, the main differences between synthetic HKCFs are more evident for different age ranges especially in the number of detectable stars.

2.3.3 Comparing observed and synthetic KLFs and HKCFs

The detailed comparison of the model KLFs and HKCFs functions to those observed was carried out only for those sources where the number of detected stars was sufficient ($I_c \geq 15$) to allow a statistically significant comparison (see Col. 3 of Table 2.1), and where submm information was available to allow meaningful estimates of extinction (this excludes Mol15 and Mol99). The number of clusters fulfilling these criteria were 16 out of 23 detected clusters. The comparison of the observed KLFs and HKCFs (KLF_{Obs}, $HKCF_{Obs}$), with the synthetic ones produced by SCG (KLF_{Syn} $HKCF_{Syn}$), for the full set of input parameters (IMF, SFH and age parameters) was carried out automatically; to ease the process, the observed and synthetic functions were computed on the same M_K and H-K grid.

The comparison procedure between synthetic and observed KLFs is described below, but it is the same for HKCFs. The KLFs are first compared bin by bin (the comparison being limited to those bins brighter than the completeness limit) identifying with i each

bin of the observed KLF, starting from i=1 for the lower-M_K non-zero bin to N for the bin where the completeness limit for that source is reached (the number N differs clearly for each cluster). In the case of HKCFs, we exclude objects with H and K magnitudes brighter than the observed limiting magnitudes for these bands. A matching flag m_i is set to be 1 for those bins where the number of sources coincide within the 1σ Poissonian error bar of the observed KLF, i.e.,

$$|N_{\star Syn_i} - N_{\star Obs_i}| \leq \sqrt{N_{\star Obs_i}} \tag{2.1}$$

The total number of bins where a match is found is divided by the total number of bins useful to the comparison to obtain a KLF compatibility figure (in %) of

$$C = 100 \times \frac{\sum_{i=1}^{N} m_i}{N} \tag{2.2}$$

The higher is C, the closer is the overall match between KLF_{Obs} and KLF_{Syn}.

However, the same value of C may result from bins concentrated at the low-M_K end of the KLF, where there are few sources, or in the region around the peak and in the proximity of the completeness limit, where instead there are more sources and hence higher statistical significance. A further figure of merit is then introduced,

$$W = \sum_{i=1}^{N} \frac{m_i \cdot N_{\star Obs_i}}{N_{\star Obs_{tot}}}, \tag{2.3}$$

where $N_{\star Obs_{tot}}$ is the total number of sources present in all the bins used for comparison. This parameter weights each matching bin by its relative richness, favoring the bins closer to the completeness limit and the KLF peak over those in the bright tail of the KLF, and favoring for HKCF the bin closer to the peak of the distribution. This choice is because the KLF (HKCF) peak region is the most sensitive to changes in the SCG input parameters.

In this automatic procedure, we select models for which the parameters C and W (at the same time for KLFs and HKCFs) are maximum. For Mol8B, Mol45, and Mol84, the observed KLF has a very irregular and multiple-peaked shape that cannot be matched by any model, and are therefore discarded from further considerations. We are then left with 14 clusters for which a series of models can be found with at least 75% of the bins matching the observations. We find that the best values of C and W are never found for one single set of parameters, but rather we identify ranges of parameter values that produce the best match; in other words, there is a level of degeneracy that the models cannot remove, and this varies from source to source. In 4 clusters (Mol109, 110, 136 and 151), this degeneracy is essentially complete and the model is unable to make any prediction; in one case (Mol148), the comparison selects models with very old stellar ages but with total stellar luminosities by far in excess of the measured bolometric luminosity obtained by integrating the observed luminosities for this region from the mid-IR to the millimeter (see Table 2.1). In the 9 remaining cases, some degeneracy persists especially in the IMF, confirming (Sect2.3.2) that our models are weakly sensitive to the IMFs, but there are clear indications concerning the SFH and ages.

Table 2.2 reports a summary of the results.The IMF of matching classes of synthetic cluster models is shown in Col. 2. Columns 3, 4, and 5 contain the times for the formation of 10%, 50% and 90% of the total number of cluster members; these values are the median of the values that these times have in all models that match the observations. Column 6 is the

Table 2.2. Results for SCG runs on detected clusters

Source Mol	IMF	$t_{10\%}$	t_{med} yrs	$t_{90\%}$	$N_\star$	$M_{\star Max}$ $M_\odot$	$M_{\star Tot}$ $M_\odot$
3	1	$10^{4.98}$	$10^{5.40}$	$10^{5.81}$	91	3.7	61
8A	2-3	$10^{4.71}$	$10^{5.38}$	$10^{5.83}$	31	3.8	30
11	1-2-3	$10^{6.06}$	$10^{6.41}$	$10^{6.82}$	48	3.8	47
28	3	$10^{6.03}$	$10^{6.12}$	$10^{6.24}$	77	9.9	105
50	1	$10^{6.16}$	$10^{6.48}$	$10^{6.66}$	53	3.5	36
103	1-2	$10^{6.57}$	$10^{6.70}$	$10^{6.83}$	115	4.0	80
139	3	$10^{5.34}$	$10^{5.96}$	$10^{6.46}$	25	2.9	16
143	2	$10^{6.57}$	$10^{6.70}$	$10^{6.82}$	27	3.1	21
160	1-2-3	$10^{6.57}$	$10^{6.7}$	$10^{6.83}$	89	4.3	63

number of cluster members $N_\star$, Col. 7 shows the mass of the most massive object $M_{\star Max}$, and Col. 8 reports the total stellar mass of the cluster $M_{\star Tot}$. We emphasize again that the analysis selects classes of models other than single models; the values reported in Table 2.2 are the median value of the parameters for each class of matching models. The table shows that for some fields multiple IMFs are compatible with the data and, in general, SFHs with constant (CR) or Gaussian (GR) star-formation rates provide acceptable solutions for certain age ranges (as reported in the Table). Simulations of SFHs with a single burst are, in general, not accepted. Our modeling is insensitive to bulk stellar ages in a cluster in excess of $5\$\times10^6 \div 10^7$ years (Sect. 2.3.2).

2.4 Discussion

2.4.1 Cluster ages and star formation histories

Perhaps the most important result of this work is that in all clusters where the comparison of observed KLFs with the ones predicted by the SCG model is possible (see previous paragraph), the observations are consistent with a star formation that continues over time intervals that in most cases have a duration of between about few 10^5 and few 10^6 yrs , and with a median cluster age of a few 10^6 yrs. In most cases, we cannot discriminate clearly between a constant or variable SFR but we are confident that we can exclude the possibility that on average the stars in our clusters are coeval and originate in a single burst of formation. Detailed studies toward the Orion Nebula Cluster indicate that stars have been forming for at least 10 t_{dyn}, or 20 t_{ff} (Palla & Stahler 1999; Hoogerwerf et al. 2001; Hillenbrand 1997), and our results would seem to generalize this on a larger sample of intermediate and high-mass star-forming regions.

In principle it can be argued that our analysis is incomplete since we did not consider longer wavelength data, which could identify heavily extincted objects that are barely visible, or not visible at all, in the near infrared. However, this does not modify appreciably our conclusions about the age spread within the clusters. Vig et al. (2007) indeed, applied a different analysis to the specific region Mol075, a field not included in our final analysis (Table 2.2) because the background-subtracted KLF is populated by too few objects for a statistically significant model comparison. Vig et al. also considered *Spitzer* IRAC and

MIPS data, looking for the brighter and redder objects in the area covered by submillimeter emission. In this way, they could identify the younger and more massive objects in the field with an estimated age of the order of 10^6 years or less. This approach, however, is insensitive to low mass and relatively older pre-MS objects, for which our method is ideally designed. While for this particular field, for the reasons explained above, we cannot perform a direct comparison to our approach, it is clear that the inclusion of longer wavelength data in the analysis might have identified a different, younger, population of objects, rather than increasing the observed age spread deduced for the clusters.

Models of cluster formation by competitive accretion appear to produce an IMF close to those observed because of thanks to the spread in the accretion rates of the competitive accretion mechanism. However, the prediction that all stars are formed in about 5×10^5 yrs (Bonnel et al. 2004) for typical conditions in young clusters, corresponding to a dynamical time or so, seems to disagree with our results. We instead favour scenarios (Tan & McKee 2002) in which stars continue to forming over several free-fall times thus providing the required age spread. The finding that the most massive object in the fields considered in this work are still being formed or have just finished a phase of intense accretion (Molinari et al. 2008a) is a further indication that star formation seems to be a long-duration process in the life of a molecular clump.

How do our clusters compare to more evolved systems, such as the sample of Herbig Ae/Be stars observed by Testi et al. (1997, 1998) ? Figure 2.14 shows the relationship between the mass of the most massive source and the total number of cluster stars as provided by the SCG simulations for our modeled clusters. The filled symbols represent the clusters that we could model (Table 2.2); the empty symbols instead represent the clusters that could not be modeled for a variety of reasons (see Sect2.3.3), while Testi et al.'s clusters are reported as asterisks (see figure caption for detailed explanation of the symbols). The figure suggests that the clusters presented in this study are richer than those surrounding Herbig Ae/Be stars for any given value of the most massive star in each cluster. The trend persists if we use similar indicators (e.g,. I_c, the full triangles in the figure). Furthermore, we note that while the limiting magnitudes of our observations and those of Testi et al. (1998) are similar, higher A_V values toward our sources and the typically greater distance from the Sun would justify the non-detection of the fainter cluster members predicted by the SCG models. It is thus likely that the values of I_c derived from our observations tend to underestimate the cluster memberships.

This evidence is clear for values of the most massive star in the cluster below $\sim 10 M_\odot$, where the 9 clusters for which we could compare observations with SCG predictions lie (the diamonds). In the clusters for which we haven't simulation results (the empty triangles), the mass for the highest-mass star was estimated by assuming that a fraction of between 30 and 100% of the bolometric luminosity originates in a single ZAMS star. In this case the trend toward richer clusters than Herbig stars (i.e., the asterisks) would become marginal. These estimates, however, place the latter clusters systematically to the right in the plot, compared to the 9 modeled clusters; indeed, if we were to estimate in the same way a maximum stellar mass also for the 9 modeled clusters, we would obtain values in excess (between a factor two and three) of those provided by the detailed SCG modeling. In other words, the evidence that our clusters are richer than those around Herbig stars is marginal at worst (i.e., using the most conservative approach of estimating the mass of the highest-mass star).

The plausibility of this interpretation is strengthened by the results of Baumgardt & Kroupa

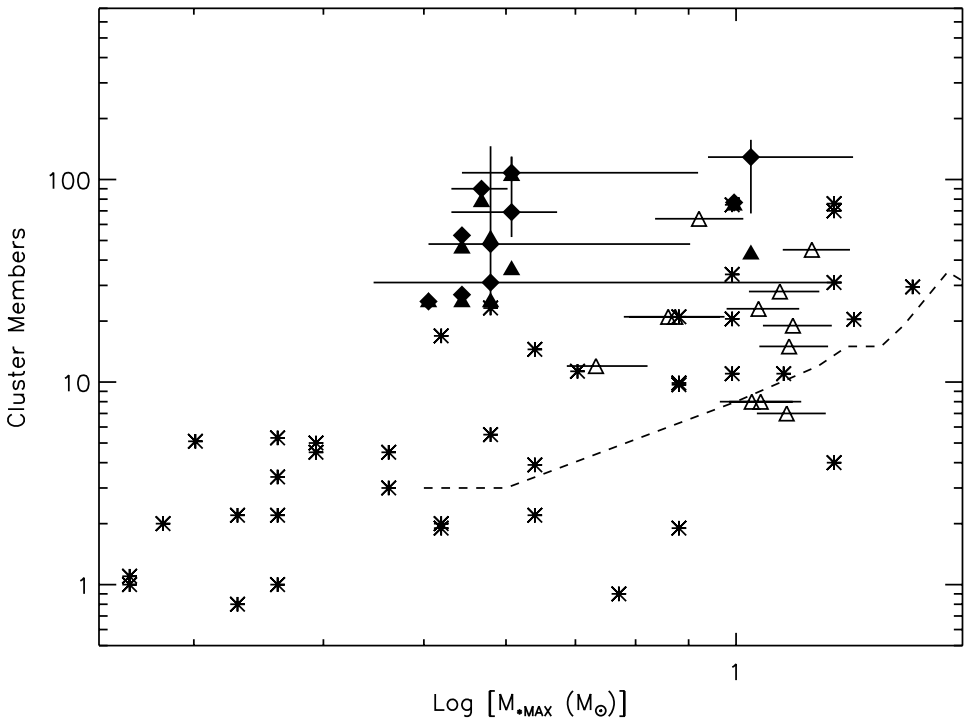


Figure 2.14. Number of cluster members as a function of mass of the highest mass star. Asterisks are for Testi et al. (1997, 1998) Herbig Ae/Be sample. Other points are for our source sample, where the full diamonds are the $N_\star$ obtained from our cluster simulation analysis; and the full lines passing through the full diamonds represent the total spread in the parameters from all SCG models that match the observed KLF and HKCF. Full triangles are the same clusters with the observationally derived I_c instead of the model-provided $N_\star$. Empty triangles are those clusters that were not analyzed with SCG, or exhibit complete model degeneracy; in these cases I_c was used for the cluster membership (from Table 2.1), while the maximum stellar mass has been derived by assuming that half of the bolometric luminosity is generated by a single ZAMS star (the horizontal lines through the empty triangles represent the mass spread assuming that a fraction of between 30% and 100% of L_{bol} comes from a single star). The dashed line delimits the region where 25% of the clusters randomly extracted from the IMF would follow in the statistical models (Testi et al. 2001).

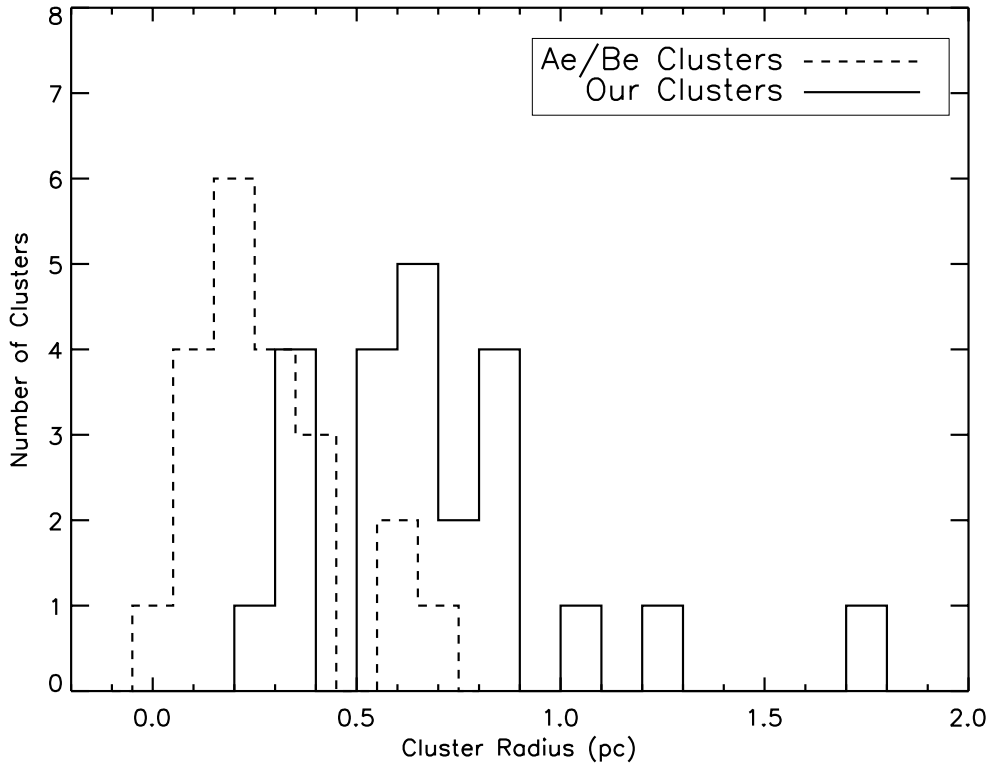


Figure 2.15. Distribution of radii for our clusters (full line) and those associated with Herbig Ae/Be stars (from Testi et al. 2001, dashed line).

(2007) who completed extensive numerical simulations of the evolution of stellar clusters as a function of, among other parameters, the cluster gas content. They show that for a wide range of initial conditions and star formation efficiencies, the dispersal of the gas with age causes the cluster to expand overall and disperse a fraction of the stars originally belonging to the cluster. As the cluster expands, its decreasing stellar density ensures that the low-density outer regions of the cluster become ever more difficult to detect against the field stars (especially in the Galactic plane, where all these objects lie), mimicking a smaller cluster from an observational viewpoint. Figure 2.15 shows the distribution of radii for our clusters (full line) and for those surrounding Ae/Be stars; the radii were derived by the same analysis in the two samples and the histogram clearly shows that our clusters are indeed larger in size, confirming the prediction of Baumgardt Baumgardt & Kroupa (2007).

This age effect on cluster size is also revealed within our clusters sample. Figure 2.16 presents the relationship between the cluster radii derived from the observations and their ages derived from the modeling. The reported correlation has a Spearman rank correlation coefficient of ~ -0.6, indicating a good correlation with a significance of 92% (between 2 and 3σ). Ongoing gas dispersal is certainly plausible in our clusters, given the common detection in these systems of molecular outflows (Zhang et al. 2001, 2005), which are highly effective in transferring material away from the forming objects and possibly out of the star-forming region; parsec-scale flows are also commonly observed from low-mass YSOs.

The final stage of gas dispersal, eventually leading to optically revealed clusters such as those around Ae/Be stars, might be triggered by the powerful winds and radiation fields

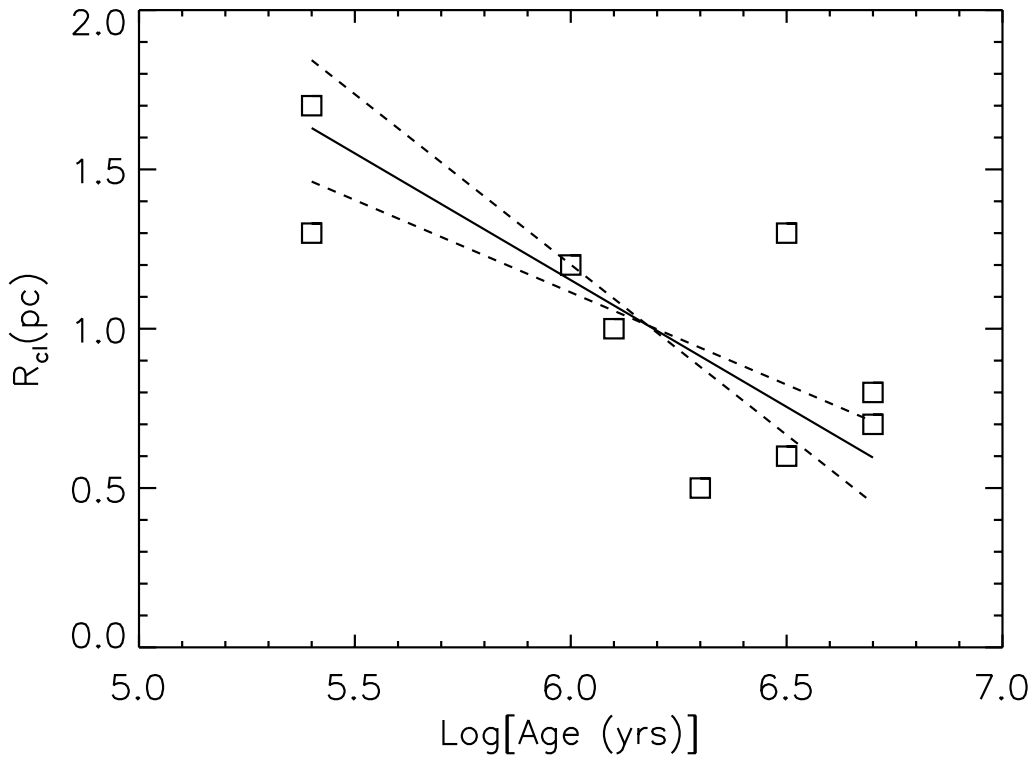


Figure 2.16. Cluster radii (from Tab. 2.1) as a function of the cluster median ages (from Tab. 2.2). Dashed lines represent the linear fits obtained fitting in turn one plotted variable as a function of the other; the full line is the bisector of the two dashed lines and represents the fit which minimizes the quadratic geometric (i.e. not along the X or Y axis alone) distance of the data from the fit. Following Isobe at al. (1990), this is the adequate approach when the nature of the data scatter around the linear fit is not known (and it is not due to classical measurement uncertainties); the slope of the full line is -0.8 ± 0.2 and the 1σ spread is within the area enclosed by the two dashed lines.

from newborn O and B stars. The indications are (Molinari et al. 2008a,b) that the most massive objects forming in the densest regions of the clumps hosting our clusters may not yet have reached the ZAMS, or are just starting to develop their HII regions. It is quite likely that this event will mark the moment of maximum efficiency of gas dispersal and further evolution of our clusters toward the Ae/Be's ones.

2.4.2 Physical *vs.* statistical models for cluster formation

Figure 2.14 can also be used as a diagnostic to discriminate between different classes of models for the origin of clusters. Testi et al. (2001) called *physical* the class of models that implies a physical relationship between the most massive star that forms and other environmental properties such as the cluster richness or the mass of the gaseous clump where the stars originate from; examples are the "turbulent core" (McKee & Tan 2003), the "coalescence" (Bonnel et al. 1998), or the competitive accretion models (Bonnel & Bate 2006). In a second class of models, called *statistical*, the relationship between the most massive star in a cluster and its richness originates in the higher probability of finding the rare massive stars in rich clusters rather than in isolation (Bonnel & Clarke 1999). If

clusters are populated by randomly picking stars from the field stars' initial mass function, and considering a cluster membership-size distribution in the form of an appropriate power law, then the observations of Testi et al. (1999) can be naturally explained. Nevertheless, this model predicts that a significant fraction of high-mass stars are still associated with relatively poorly populated clusters, in other words that massive stars can be found both in high-N clusters and, to a lesser extent, in low-N clusters.

The dashed line in Fig. 2.14 is the upper boundary of the region that should contain 25% of the cluster realizations obtained by randomly extracting stars from the IMF (Testi et al. 2001). If we consider our measurements of I_c for our entire sample (i.e., the full and empty triangles), a fraction of about 15% of the clusters is found marginally below the dashed line. However, we note that our modeling was possible only for clusters above a membership threshold derived with I_c, which is thus a biased subsample toward rich clusters. Based the extreme assumption that the fields with no detected cluster are cases of systems consisting of a single heavily extincted star, and thus would fall below the dashed line in Fig. 2.14, this fraction would approach the 25% level. This, however, is an extreme case because, as we have already discussed, the high value of the extinction derived from submillimeter maps may be the reason for not detecting clusters in at least a number of observed fields.

As an additional means of differentiating between physical and statistical cluster models, Weidner & Kroupa (2006) argued that a non-trivial correlation exists between the highest-mass star in a cluster, $M_{\star Max}$, and the total stellar mass of the cluster itself, $M_{\star Tot}$ (Fig. 2.17). Numerical simulations show that the relationship obtained by pure random sampling of the IMF with an imposed physical limit of 150 $M_\odot$ for the maximum stellar mass (the dashed-dot line in Fig. 2.17) clearly does not represent our data.

A substantially different result (the dashed line) is obtained if cluster members are selected in ascending order and constrained to total cluster masses distributed according to a cluster total mass function (Weidner & Kroupa 2006). Basically, this second option infers that drawing 10 clusters of 100 stars will not deliver the same $M_{\star Max}$=f($M_{\star Tot}$) as drawing 1 cluster of 1000 stars. This trend closely resembles a semi-analytical approximation of $M_{\star Max}$=f($M_{\star Tot}$) obtained by Weidner & Kroupa (2004), again assuming that total cluster stellar masses are distributed according to a power-law mass function. Weidner & Kroupa (2006) suggested that this *sorted sampling* way of populating a cluster can be physically understood in terms of a pre-stellar clump where initial low-amplitude perturbations start low-mass star formation; as further perturbations with larger amplitude grow, higher mass stars will start to form until the feedback from the latter will begin to disrupting the cloud. This scenario of organised star formation where low-mass stars are the first to form, is the same as we favor (see Sect.2.4.1) considering the age spread that we find in our clusters in which, based on independent considerations (Molinari et al. 2008a), the most massive star may not have formed. By the way, this latter possibility does not change the substance of the agreement between our data and the *physical* cluster models in Fig.2.17, since the late addition of the highest mass star currently not yet visible in the near-IR would shift the points toward the top-right of the plot.

We verified *a posteriori* that the range of $M_{\star Max}$ and $M_{\star Tot}$ parameters values explored by our simulations is much wider than the area spanned by the bars attached to the single points, and also includes regions compatible with the random sampling cluster model. We then conclude that the result of Fig. 2.17 is not likely to be produced by a biased sampling of the clusters' physical parameters explored in our models.

The predictions of the *sorted sampling* descibed above, with which our data points best

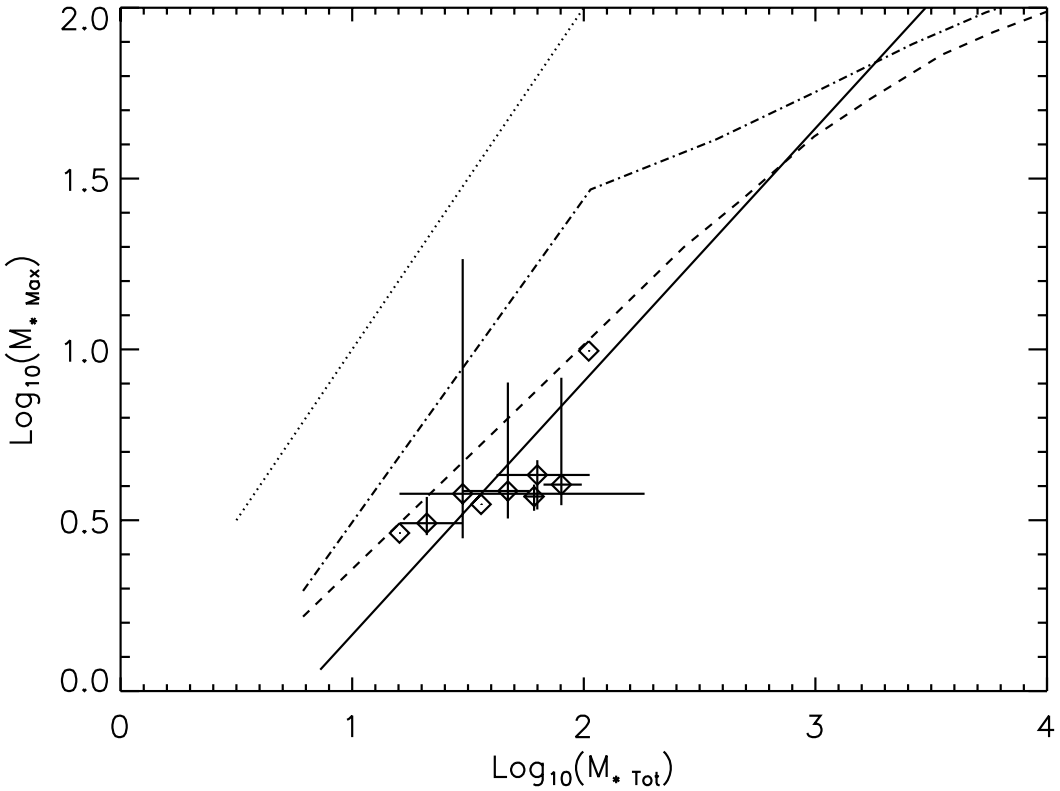


Figure 2.17. Mass for the highest mass star as a function of the total stellar mass for the 9 modeled clusters (see Table 2.2); the bars associated to each cluster (the diamond symbols) represent the total span of the parameter values for the classes of models selected by our analysis (the diamond marks the median values). The dash-dotted, and dashed lines represents the relationship obtained for numerical simulations of clusters drawn from pure *random sampling* of the IMF, and using a so-called *sorted sampling*, following Weidner & Kroupa (2006). The full line is a semi-analytical approximation of this relationship obtained by Weidner & Kroupa (2004). The dotted line is the limit where a cluster is made of just one star.

agree, are also in good agreement with the results from simulations of clusters forming in a competitive accretion scenario (Bonnel et al. 2004). This model, however, seems to be excluded by the observed ages and age spreads in our clusters which are in clear disagreement with the predicted cluster formation timescales of 2-3 free-fall times.

2.4.3 Influence of binarity on the interpretation of age spread

Weidner et al. (2008) carried out extensive numerical simulations to determine how the presence of unresolved binary/multiple stars can affect the observational properties of a young cluster in a massive star-forming region. Assuming 100% binarity in a cluster of coeval sources, they find that unresolved binaries may lead an observer to conclude that a significant age spread is instead present in the cluster; the full line in Fig.2.18 shows the *a posteriori* age determination assuming that all binaries are unresolved. We see that the measured age spread for the large majority of stars simulated in Weidner et al.'s simulation is comparable to a log(age) Gaussian distribution with σ=0.1, which is one of the possible choices of star formation histories in our cluster models. However, the comparison of our

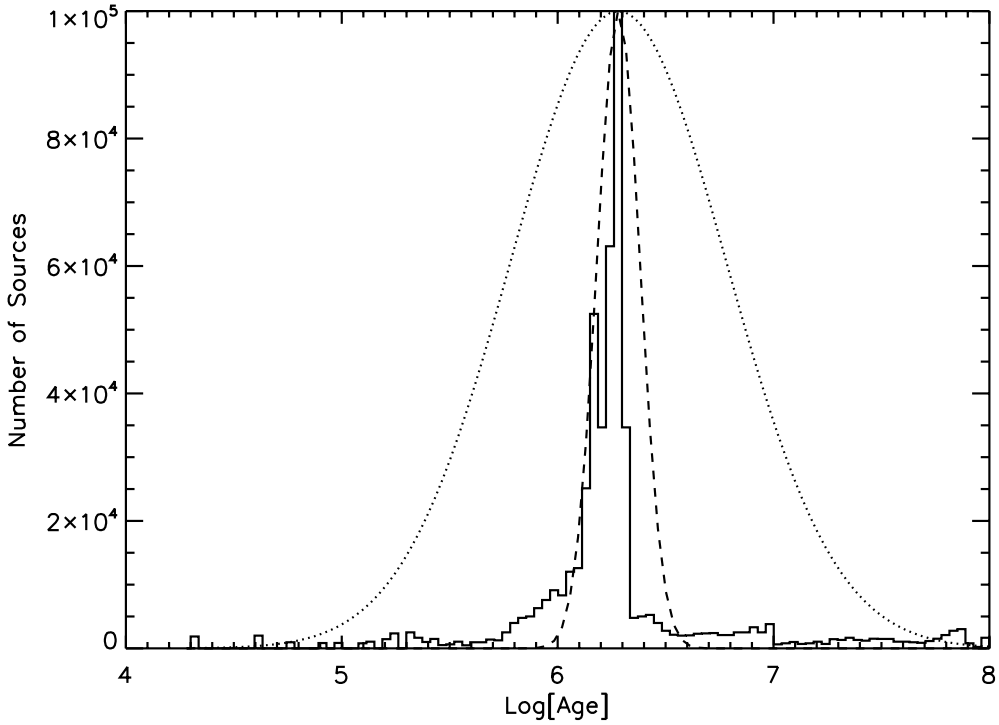


Figure 2.18. Full line represents the age spread resulting from the simulations of Weidner et al. (2008) for a cluster with an input age of 2×10^6 yrs and 100% binarity fraction. The two gaussians are the age weight functions used in our simulations of Gaussian Star Formation Histories, with σ=0.1 (dashed line) and 0.5 (dotted line) respectively. In case of constant SFH models, the adopted age spread is comparable to the σ=0.5 distribution in the figure.

observed KLFs with the SCG models (Table 2.2) suggests age spreads much larger than this, and more comparable to a log(age) distribution with σ=0.5 (the dotted line in Fig. 2.18). We then conclude that unresolved binaries in our clusters cannot account for the observed age spread.

Chapter 3

From Mid InfraRed to Millimetric: High mass objects

The empirical classification of the SED of low mass YSOs has been used as a powerful tool to constrain theoretical models. In particular, is was shown that the peak of the SED shifted from wavelengths longward of 100 μm for Class 0 objects toward shorter wavelengths the more the YSO approaches the Main Sequence (MS). An additional tool that proved to have good diagnostic power is the relationship between the bolometric luminosity of the YSO and the envelope mass out of which it forms. Objects known to be in different stages according to the SED shape classification also occupy distinct regions in the L_{bol}-M_{env} diagram, and the entire Class 0-II path can easily be followed even with very simple models (Saraceno et al. 1996).

Establishing a similar scenario for intermediate and highmass YSOs is much more difficult due to the clustered environments in which they are born and to the large (d≥1 kpc) characteristic distances. Qualitative schemes have been proposed (Evans et al. 2002); a more quantitative assessment especially in terms of lifetimes of the duration of the different phases is still elusive. IRAS data have been mainly used to select samples of massive YSOs using color criteria, but such investigations can only go as far as providing statistical indications about the evolutionary stage that different groups of sources exhibiting specific global SEDs properties may be in. The complexity of the situation clearly emerges when these bright IRAS sources are observed at higher resolution in the mid-IR, and are often resolved in clusters of several members with a diffuse emission component local to the star forming region. The same region observed in the sub-millimeter even with the limited spatial resolution attainable with single-dish facilities may or may not offer a scenario consistent with the mid-IR picture. In the following sections we presents our methodology to analize our multi-wavelenghts data. First we build the SEDs from the observation and then we fit them with theoretical models, described in Sec. 3.3. All the steps are presented in this chapter.

3.1 SED Building

We build the SEDs for all the sample fields for that we have the image in all the wavelenghts between mid-IR and mm. All the sources are identified in each wavelenght separately and then the fluxes at the different wavelenghts are associated through their coordinates.

The analysis of each image was made with an algorithm developed by our group to analize the star formation regions. Source detection is based on the analysis of the brightness profile usig the second derivative of the image. The derivative is a powerfull means to find the local maxima and to identify the changes in curvature, so we can de-blend the close sources even in presence of a spatially variable background. The derivative is besides used to describe the source shape through their perimeter identified finding the first minima around the local peaks. The source position and their perimeters are used as constrains to a bi-dimensional gaussian fit algorithms, that extimates the source flux making a simultaneus analysis of the source profile and the plateau where the source is located; this is a fondamental step in the analysis of a star forming region where all the source are inside a variable background. The detail of the problematics finded in the analysis of a star forming regions and the solution obtained are presented in the last chapter (see chapter5), where this technique is described in detail.

In this way each source is characterize with its spectral form and this is used to obtain the evolutive information on the objects making a comparison with modelled SED, following introduced. Bolometric luminosity (L_{bol}) of these source was been obtain integrating the SED, this information is usable to plot our source in an evolutive diagram, described in the last paragraph 3.5, with another protostar property, the envelope mass that is an physical quantity important for the evolutive history of the protostars.

3.2 Data Analisys

Program fields are the same of NIR analysis, but we have not observations at all the wavelengths for all of them. Fields included in GLIMPSE and MIPSGAL survey, which will be described later, are indicated with the Y letters in col. 4 and 5 of Table 1.1, while those regions that have a millimetric observation are cheked with an Y in col. 6. In the following subsection we will outline the general method to building an observative SED, and in detail the variuos step. We want identify YSOs in our sample and estimate their fluxes from the mid-IR to the millimeter. To achieve our aims, we use the "NOMEPRO" algorithm, developed by us and described in the chapter 5, for sources extraction and photometry. This software is able to identify and to make photometry of the compat core associated with massive forming objects only, separating the contribution of the core from what is more likely from the diffuse clump medium.

3.2.1 MiD-Far Infrared: from MSX and IRAS to SPITZER surveys

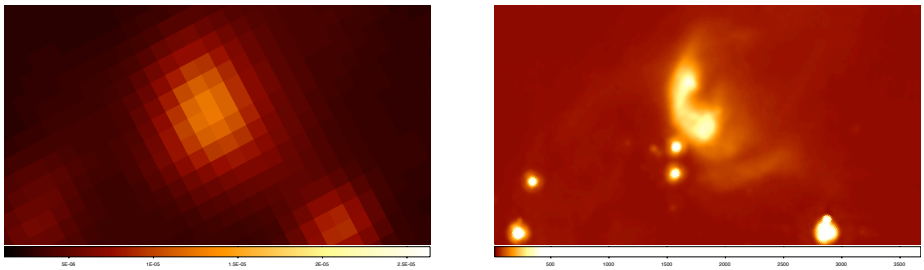
A previous SED study on these fields was been do in the past using MSX[1] data for mid-IR and IRAS[2] data for far-IR (Molinari et al. 2008a). In Fig. 3.1(a) and 3.2(a) we show two MSX image at 8.3 and 21.4 μm for a filed of our sample, Mol98. A compact core is clearly visible in these images, but increasing the spatial resolution of the observation this single object is, in reality, compound by a set of diffuse emission and compact cores, as shown in Fig. 3.1(b) and 3.2(b) where are presented two images obtained with SPITZER, precisely from IRAC at 8 μm the first and from MIPS at 24 μm the other. These images are acquired

[1]Midcourse Space Experiment (MSX) is an IRSA project with the SPIRIT III infrared telescope (33 cm aperture) that cover an area between 0°<l<360°, |b| <5° in four Mid-Ir bands (8.3, 12.1, 14.5, 21.4 μm).

[2]The Infrared Astronomical Satellite (IRAS) was a joint project of the US, UK and the Netherlands. The IRAS mission performed an unbiased, sensitive all sky survey at 12, 25, 60 and 100 μm.

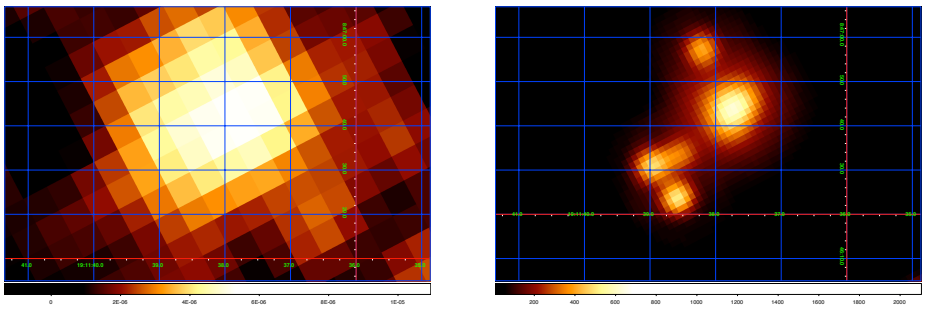
by the satellite within two great surveys GLIMPSE and MIPSGAL. The great number of our field presents a multi-sources structure in the mid-IR bands but the situation is less complex at 70 μm because due to the lower spatial resolution, so the various sources in the area can be blended in a large patch of emission. In spite of this fact, MIPSGAL 70 μm observations have a better spatial resolution respect to the IRAS data, so the improvment is high within trying to isolate as much as possible the contribution of YSOs out of the total that may due to other nearby objects or extended components.

We describe now the properties of our data.



(a) MSX 8.3 μm (b) IRAC 8 μm

Figure 3.1.

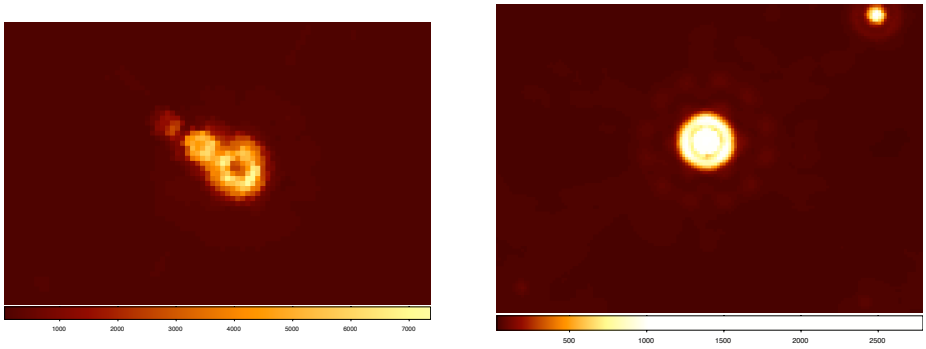


(a) MSX 21 μm (b) MIPS 24 μm

Figure 3.2.

The GLIMPSE project observed the Galactic plane ($10°\leq|l|\leq65°$ and $|b|\leq1°$) in the four IRAC bands (centered at 3.6, 4.5, 5.8, 8 μm) in 400 hours. The MIPSGAL survey

coverage are the regions $10° \leq |l| \leq 60°$ and $300° \leq |l| \leq 336°$, in the IR bands of 24 and 70 μm with MIPS photometer. The mosaics will span the survey region in 1.1 by 1.1 degree tiles. Our fields are candidate protostars of high mass, so they are characterized by an high luminosity. Therefore there is a probability that our sources are saturated at almost all the wavelenghts (only at 70 μm we reach not the saturation limit). Indeed at the GLIMPSE bands and at 24 μm images we have some cases of saturation. Spitzer team provides a routine (BCDsaturation.py, it is a Python-interactive IRAC/MIPS BCD saturation recovery from T. Jarrett) to correct the source brightness profile. The saturation produces differents features in the image obtained with MIPS or IRAC instrument. Saturated sources at 24 μm present the classical "*donut*" shape, see Fig. 3.3(b), but in the IRAC bands there are a further problem. Indeed the saturation produce the *donut* shape for the bright source, but introduce a sequence of "*ghost*" figures of the real image, as shown in Fig. 3.3(a). These *ghost* images are located along one direction that depart from the real source, they have a decresing flux going faraway the center. They are ever three. The Jarrett routine resolves very well, with an interpolation, the problems at 24 μm but it has some difficulties at the IRAC bands, and it is not able to remove the gost image.



(a) GLIMPSE saturated source (b) MIPS saturated source

Figure 3.3.

So in those cases where the saturation problem is present at GLIMPSE bands we eximate a lower limit for that objects integrating the *donut* of real sources by hand with costum procedure using GILDAS package. Source detection and photometry was doing with "NOMEPRO" algorithm. We find the optimal threshold for detection for each wavelenght testing our routine of a subsample of little subfields (one for each band), with this procedure we run the algorithm in an automatic way on all the sample. The results of data analysis are labelled in table 3.1.

The association of the source at the differents wavelenghts was made by hand, considering the coordinates of the peak of the sources but even their shape, so a source can correspond to more than one objects in a different wavelenght. The IR nomenclature, shown in Col.2 of table 3.1, are choose on the basis of source distribution in the 24 μm images.

Table 3.1. Results of analysis of mid and far infrared data

Source	IR	Fluxs [Jy]					
Mol		3.6[μm]	4.5[μm]	5.8[μm]	8.0[μm]	24[μm]	70[μm]
38	A	0.0027	0.0266	0.0752	0.3317	0.9300	–
	B	0.0043	0.0050	0.0086	0.0094	1.0726	58.19
	C	0.0017	0.1732	0.0019	0.0007	0.0766	54.51
	C1	0.0262	0.0074	0.0428	0.0110	–	–
	D	-	-	0.0047	0.0300	0.0203	3.73 [a]
	E	0.0681	0.2780	0.2927	0.0469	0.0182	2.20 [a]
	F	0.0556	0.1365	0.0558	0.1052	0.0295	2.99 [a]
	G	0.0086	0.1099	0.1828	0.2187	0.0295	1.66 [a]
45	A	0.1738	0.1450	0.9126	1.2569	72.80	167.47
	A1	-	-	0.1436	0.5901	–	–
	B	0.0104	0.0131	0.0067	0.0037	0.0665	4.79 [a]
	C	0.0041	0.0113	0.0037	0.0088	0.0124	2.55 [a]
59	A	0.7035 [b]	1.5367 [b]	3.3839 [b]	3.6180 [b]	38.807	46.40
75	A	1.0170 [b]	1.3830 [b]	11.760 [b]	9.1800 [b]	150.61	128.49
	B	0.1898	1.4113	1.0923	0.2757	2.3122	–
	C	0.3115	0.2773	2.2568	0.9773	2.8847	–
	D	-	0.0033	0.0074	0.0408	0.9701	–
	E	0.0060	0.0040	0.0607	0.0408	0.9701	–
	F	-	-	-	0.1709	0.5321	–
	G	0.0108	0.0072	0.0110	0.0103	0.1660	4.69 [a]
	H	[c]	[c]	[c]	[d]	0.3789	19.31
	70I	-	-	-	-	-	10.38
98	A	0.0003	0.0017	0.0048	0.0056	0.2284	74.54
	A1	-	0.0029	0.0044	0.0028	-	–
	B	0.0017	0.0018	0.0852	0.0697	0.2859	–
	C	0.0143	0.0309	0.0408	0.1167	2.0287	9.29 [a]
	D	0.0052	0.0016	0.0032	0.0550	2.7474	5.30 [a]
	D1	0.0007	0.0024	0.0054	0.0968	–	–
	D2	0.0020	0.0015	0.0031	0.1143	–	–
	E	0.0009	0.0036	0.0237	0.1788	0.3656	6.51 [a]
109	A	0.0165	0.0213	0.0316	0.057	79.83	122.19
	A1	-	0.0106	0.150	0.0086	-	-
	B	0.0048	0.0132	0.0056	0.3347	1.6809	13.59 [a]
	C	0.0029	0.0026	0.0079	0.0328	1.1579	16.59 [a]
	D	0.0070	0.0036	0.0181	0.1161	5.7339	9.63 [a]
	E	0.0007	0.0002	0.0087	0.1448	0.1361	4.33 [a]
	F	0.00003	0.0014	0.0066	0.0246	0.3128	1.91 [a]
	G	[d]	[d]	[d]	[d]	1.6526	7.09 [a]
	G1	0.0008	0.0007	0.0113	0.0558	–	–
	H	0.0025	0.0015	0.0386	0.0552	0.4395	5.18 [a]
110	A	0.0034	0.0018	0.0656	0.0649	5.5510	86.91
	A1	0.0046	0.0199	0.0087	0.0400	–	–
	A2	0.0065	0.0140	0.0014	0.0037	–	–
	A3	0.0055	0.0081	0.0190	0.0232	–	–
	B	0.0011	0.0003	0.0032	0.0344	0.0726	3.88 [a]
	C	0.0015	0.0012	0.0025	0.0134	0.2769	3.49 [a]
	D	0.0037	0.0004	0.0124	0.0307	0.0527	2.29 [a]
	E	0.0009	0.0010	0.0017	0.0200	0.0221	1.31 [a]
	F	0.0222	0.0356	0.0178	0.0189	0.0057	0.482 [a]

[a] Upper limit
[b] Lower limit
[c] Out of image
[d] Diffuse emission

Table 3.2. Results of analysis of millimetric data

Source Mol	MM	Fluxs [Jy] 450[μm]	850[μm]	1200[μm]	IR-ass[a]
38	1	38.24	3.99	[b]	B
	2	[c]	1.36	[b]	C
45	1	64.77	6.17	[b]	A
59	1	16.50	1.25	0.71	A
75	1	46.51	5.88	[b]	A
	2	31.45	2.45	[b]	B
98	1	54.61	4.47	2.12	A
109	1	[b]	[b]	1.86	A
110	1	[b]	[b]	1.32	A

[a] Association with a IR source, the nomenclature is that assigned in Col. 2 of table3.1
[b] No observed in this wavelenght
[c] Out of image

3.2.2 Millimeter

Sub-mm/mm continuum maps are available for all sources in table 1.1 where there is the Y letters in col. 6. Large maps covering several square arcminutes at 1.2 mm (HPBW=24″) were obtained at the SEST telescope with SIMBA by Beltrán et al. (2006) for some fields with 08h$\leq\alpha_{J2000}\leq$18h. For all other fields small (FOV$\sim$2′^{2}) 850 μm (and occasionally 450 μm) jiggle-maps were acquired in service observing at the JCMT telescope using the SCUBA instrument (HPBW=15″@850 μm). Our data analysis code make a good detection and extraction of the core, separing its contribution from which of diffuse material. For some of the fields previously analized (see Molinari et al. (2008a)), we compare our results with those obtained with a different technique finding a good correspondence.

3.3 The SED Models

Two codes are publicly available to generate synthetic SEDs suited for star formation: DUSTY (Ivezić & Elitzur 1997) and the 3D Radiation Transfer Code by Whitney et al. (2003a,b). We adopted the latter due to its 3D nature which allows us to model the disc around the central star, to take into account the polar cavities due to outflows, and to derive, for the same input parameters, how the SED changes when seen through different lines of sight. Besides this code contemplates how the SED change along an evolutionary sequence.

3.3.1 Embedded Protostar Models

Whitney's models attempt to give a realistic desription of a protostar considering th 3D structure of this objects and development a radiative transfer method. These models are very accurate and take into account the several aspects necessary to obtain a good rapresentation of a protostellar object.

- ***Dust properties***. The dominant opacity source in the radiative transfer models is dust grains. Whitney models have the capability to include different grains in different regions.

- ***Effects of geometry***. The structure of embedded protostars is not trivial and several components may be considered.

  - *Disk*. Its presence appears to be ubiquitous in star-forming regions.
  - *Bipolar Cavities*. Embedded protostars require bipolar cavities to fit the observations (Whitney et al. 1997), they include them in these models. They choose two shapes to describe the cavities since both appear to be evident in observations (Padgett 1999; Reipurth et al. 2000). These are a streamline, which is conical on large scales, and a curved cavity. The streamline cavity shape might occur if precessing jets carve out a conical shape while infalling material outside of the cavity continues to fall in along streamlines.
  - *Envelope*. They choose for the envelope geometry the infall solution that includes rotation (Terebey et al. 1984). On larger scales, magnetic fields (Li 1998; Ciolek & Königl 1998) and initial motions prior to dynamical collapse (Foster & Chevalier 1993) may affect the structure, but within the infalling region, the effects of rotation likely dominate over magnetic effects. This region contributes most to the SED and images.
  - *Protostar*. Most of the source luminosity comes from the central star, with a amount (that varies due to the different evolutive phase) of accretion luminosity from the envelope, in the first phase of formation, and prevantly from the disk, in the following phases.

  The models properly account for scattering and inclination effects. The largest variations with inclination appear in the mid-IR, with edge-on sources showing a broad dip near 10 μm in the Class I–II sources. The 9.8 μm amorphous silicate feature is much narrower than this. The broad dip is due to two effects: the large extinction in the disk midplane, which blocks thermal radiation from the inner disk+envelope, and the low albedo, which prevents radiation from scattering out the polar regions.

- ***Effects of evolution***. Protostars have been grouped into an evolutionary sequence, Class 0–Class III, on the basis of their spectral energy distributions (Lada & Wilking 1984; Strom et al. 1989). This evolutionary sequence is characterized by a decreasing envelope infall rate, increasing disk radius, increasing bipolar cavity opening angle, and decreasing cavity density. Each evolutive phase is characterized by several differerences from the others. The evolution of principle features are explain following.

  - *Class 0*. Class 0 sources are in the early protostellar collapse stage, with spectral energy distributions (SEDs) that resemble blackbodies with T$\leqq$30 (André et al. 1993). The majority of the source mass resides in the infalling envelope. Even at this early stage of collapse (of order a few times 10^4 yr), these sources exhibit powerful, bipolar molecular flows.
  - *Class I*. The Class I stage is a later stage of protostellar collapse that lasts a few times 10^5 yr. These sources display very broad SEDs that peak near 100 μm. Their envelope masses are similar to the mass of the central pre–main-sequence core. They have well-developed accretion disks, and their envelopes have bipolar cavities excavated by outflows (Terebey et al. 1993; Tamura et al. 1991).

- *Class II*. The Class II stage is characterized by the presence of excess infrared emission above that expected for a stellar photosphere, with the SED emission peak occurring in the nearinfrared. This is the signature of a classical T Tauri star surrounded by an accretion disk. Infall from the cloud has ceased owing to dispersal of the remnant infall envelope by the combined effects of infall and outflow. Class II sources have ages of about 10^6–10^7 yr (Strom et al. 1989).
- *Class III*. The Class III sources, also known as "weak-lined" T Tauri stars, have SEDs that resemble a stellar photosphere. The Class III stage can last for about 10^7 yr, ending in a zero-age main-sequence star (Palla 1999).

The SED models, for several conditions, are presented in Fig. 3.4, 3.5 and 3.6. Each figure shows the SEDs for all the different evolutive phases at a fixed angle of the line of sight. The dashed line show the stellar photosphere emission as it would be if there is not the extinction of the dust. The disks and cavities have effects on the resulting SEDs in the mid-IR, while the envelope emission is important in the far-IR and mm range. During the protostar evolution the contribution of the several components change, producing analogue changes in the SED; the envelope contribution in the far-IR becomes less dominant while becomes visible the emission of photosphere and of the disk. The object SED, for the same protostar class, changes modifing the angle of the observation. Indeed the protostar photosphere emission becomes more visible decreasing the angle from the edge-on position (87° shown in Fig. 3.4) to the pole-on (17° shown in Fig3.6), this is due to the bipolar cavities "excavated" from the polar jets. These jets remove some material from the accreting envelope along the rotating axis (othervise the polar direction) opening a window on the internal objects.

The big accurancy of these models help us to retrieve important information about the evolution and the physical properties of our sources.

3.4 Observed SED

In this section we present our selection criteria and, for all the sources for which we can build the observed SED, the fitting of real SED to the model previously described. The results of the SED fitting are labeled in table 3.3.

In the following discussion we indicate with *main source* the object (or objects, indeed if we can resolve more than one source in the Far-IR and in the sub-millimeter images this implies a possibility to have two or three or more main sources) associated with the Far-IR and millimetric emission, while the secondary sources are all the others.

3.4.1 Model Selection Criteria

Whitney gives us about 200000 synthetic SEDs builded with their model to compare them with our observed SED, Robitaille at al. 2007. Several parameters vary between one SED and another, such as objects mass, temperature, radius, envelope, mass accretion rate, evolutive phase, inclination angle (this has a fondamental rule in how the SED appears). All the models are build for a source that is at a distance of 1.0 Kpc.

The first step of the comparison between our observed SED and the models is a *preselection* of synthetic SED that can be reproduced our observation. This is based on the bolometric luminosity of each source, only models that have luminosities comparable with

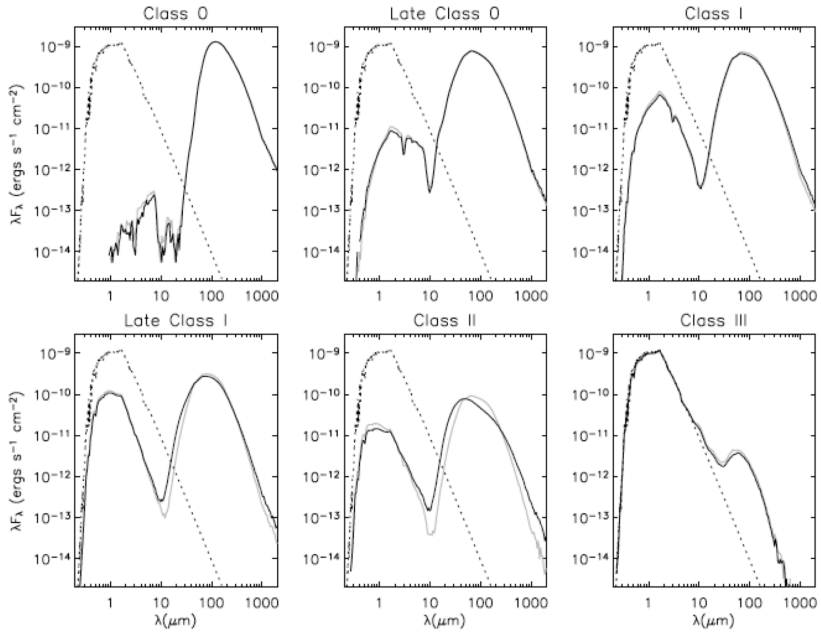


Figure 3.4. The graphics show how the SED change along the evolutive way of a source. Six differents phase are presented, from Class0 (top-left) to ClassIII (bottom-right). These SED are obtain for an object observed with an edge-on line of sight angle, $\theta = 87$, which coincides with the maximum extinction.

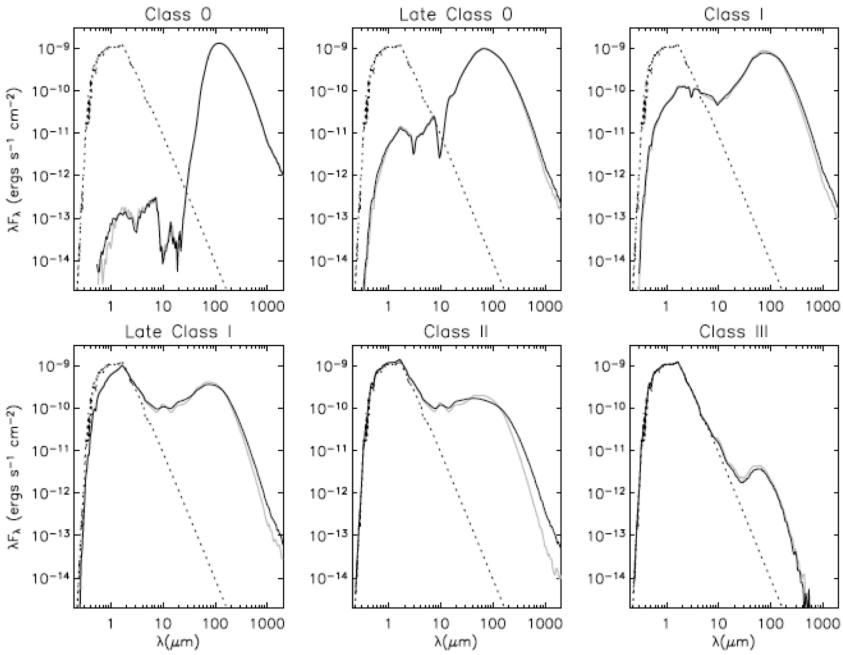


Figure 3.5. The same of Fig. 3.4 but with an line of sight angle of about $\theta = 56$.

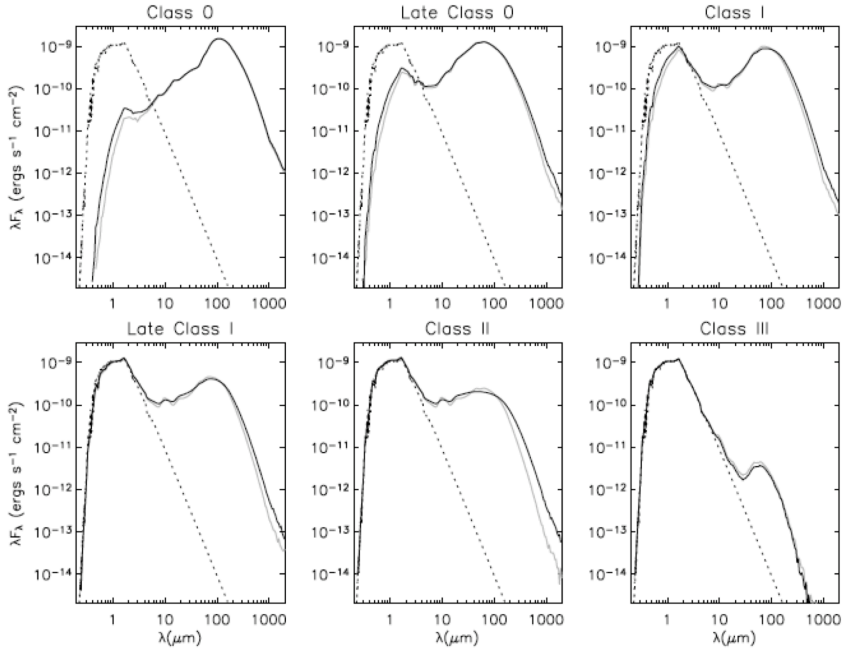


Figure 3.6. The same of Fig. 3.4 but with an line of sight angle of about $\theta = 17$, polar on, which coincides with the minimum extinction.

the observed one can be associated with our sources. This is used to reduce the number of comparison to do and to speed up the running schedule, removing a sub-sample of incompatible models. The observative bolometric luminosity is calculated integrating the SED. Our sources have several distances (between few hundred to any thousands of parsec) from the Earth, while all the modelled sources are at a distance of 1 Kpc, so all the observed luminosities are "*rescaled*" to a distance of 1.0 Kpc to carry out the comparison.

The effective comparison is made on this models subsample. We present now the general criteria for the choice of the best models; in some cases these are modified to take in account the different conditions and situations that we meet with some objects.

The best fits of models with the observed SED are selected with a weigthed least-sqares algorithm in which the weigth are related to the goodness and the reliability of the data in the several wavelenghts rather than to the errorbars on the fluxes. Indeed a flux value has a different quality, and therefore a weight, due to the spatial resolution and the sensibility that we can obtain at a given wavelenght. Our SEDs are split into three regions with associated a weight ($\frac{1}{\zeta^2_{[i,j,k]}}$), these values for the regions are calculated so that the weights sum would be 1 and that the reciprocal rapport between MIR:FIR:MM would be 3:1:3 to give more importance, in the fitting, to those regions that are characterized from a good resolution and hence better resolved sources.

The quantity that we minimize, in the fit, to the best model selection is

$$\xi = \sum_{i=1}^{N_{MIR}} \frac{(F_i - M_i)^2}{\zeta_i^2} + \sum_{j=1}^{N_{FIR}} \frac{(F_j - M_j)^2}{\zeta_j^2} + \sum_{k=1}^{N_{mm}} \frac{(F_k - M_k)^2}{\zeta_k^2} \tag{3.1}$$

where $F_{i,j,k}$ are the fluxes of each band (excluding the upper and the lower limits) and the $M_{i,j,k}$ are the models fluxes value at the same wavelenghts. Whitney gives us the models with the SED builded in several filters, we make the comparison, obviously, using only those filters that coincide with the our observed bands. Our regions, and the corresponding weights, are:

- ***MIR*** with $3.5\mu m \leq \lambda \leq 25\mu m$. This region include the IRAC bands and the MIPS 24 μm, so these images have a very good resolution and we can assume that all the sources presented in the fields are detected and resolved (with the exception of those objects so blended to be indistinguishable). For all parts in this region we assign a weight of $\frac{1}{\zeta_i^2} = \frac{3}{7}$

- ***FIR*** with $\lambda = 70\mu m$. The FIR region include only the MIPS 70 μm images; these are the data at the lower resolution and we expect that improving this resolution we can resolve some of the actual "blob" in several components. So we associate a low weight to this region, equal to $\frac{1}{\zeta_j^2} = \frac{1}{7}$

- ***MM*** with $\lambda \geq 300\mu m$. The millimeter bands, obtained from the Earth with tow different telescope (see Section 3.2.2) have good resolution and sensibility so we assume that all the objects that emit in these bands are reveal. Such as for the MIR region, even the MM one has an high weight about $\frac{1}{\zeta_k^2} = \frac{3}{7}$

For the SED models we choose the large aperture (5000 AU) that foresees, see section 3.3, the inclusion of the envelope contribution in the SED. We verify that the dimension of

our fitted objects is similar to this value, so we can affirm that even the observed cores are the sum of a contribution of an embedded source and an envelope.

3.4.2 Fitting Results

For a complete discussion we compare our results with those obtained using MSX data for Mid-IR bands and IRAS data for Mid-IR and Far-IR ones (see Molinari et al. (2008a)), to underline the improvement in the SED building and in the source identification due to the improvement of spatial resolution and sensibility. We want to anticipate that the Higal data will give a further improvement both in the source association between the different bands (the spatial resolution at 70 μm will be like the actual Spitzer 24 μm, so the big "blob" visible in the Far-IR can be resolved in several sources, if there are) and in the SED building (covering the most important region of the SED, the peak).

Mol038

The image at 24 μm, see Fig. 3.7(b), shows a group of several source, while in the Far-IR only two cores are visible. These two objects are associated with two sources that are present at Mid-IR wavelenghts. A complete SED can be build for six sources that present different evolutive state. Three source (E, F, G) are fitted by ZAMS stars SED of intermediate mass (about 6-7 $M_{\odot}$) and and age about 2-4$\times10^6$ years. All the characteristics of the selected fits for these objects underline the main-sequence phase of them. Indeed they have a residual envelope and a $\dot{M}$ practically null, and a $M_{\star}$ and $R_{\star}$ tipically of a MS stars.

The other three sources are younger. Two of them (B and C, the main sources) posses a consistent envelope again (thousands of $M_{\odot}$) and have high values of $\dot{M}$, these facts underline their pre-main sequence status. They are high mass sources (13-15 $M_{\odot}$), but they have not reach their zero-age main-sequence mass values. These are the two sources that have their contourpart in the Far-IR and mm bands. The last objects (D) correspond to a fit of a young source (age is about 2×10^5) but with a poor residual envelope and a low mass accretion rate. So we classify this objects as pre-main sequence because it reachs not its main-sequence value, besides it has an high radius to be already collapsed sources.

This global images of this star forming region shows a no trivial situation, where star with different evolutive phases exist at the same time.

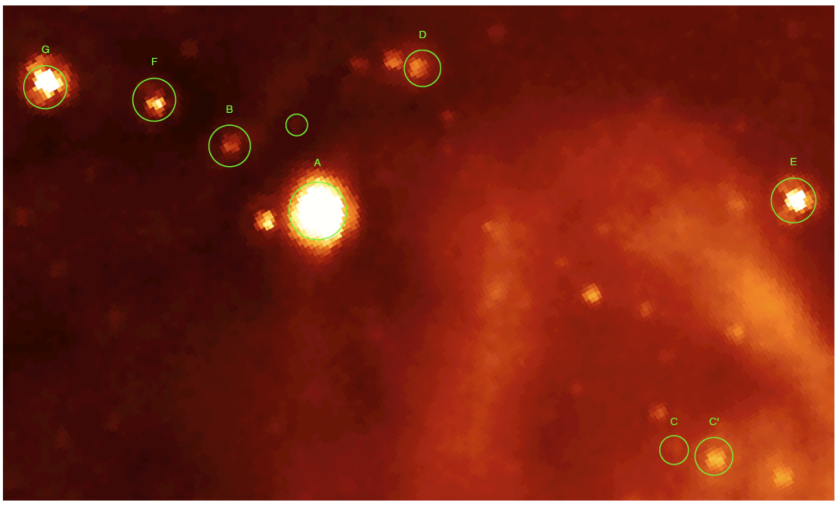
The comparing with the previous analysis with MSX+IRAF data shows that all the sources visible at MSX 8 μm are detected at MIPS 24 μm, while they were desappear at MSX 21 μm where a single weak source was present (G source that has not corrispondence with sources in the others band), our B source. This G source was fitted with a grey-body and it results as a millimetric source, while our B source has a complite SED so with an high-resolution analysis we have identified the Mid-IR counterpart that correspondes to the Far-IR/mm source. This objects is well fitted with a very young accreting object, so we confirm its protostars condition. This is the unique source fitted in the previous analysis. In the actual analysis we find another main source (C) that is fitted as pre main-sequence objects.

Table 3.3. Best fits selected for the sample sources.

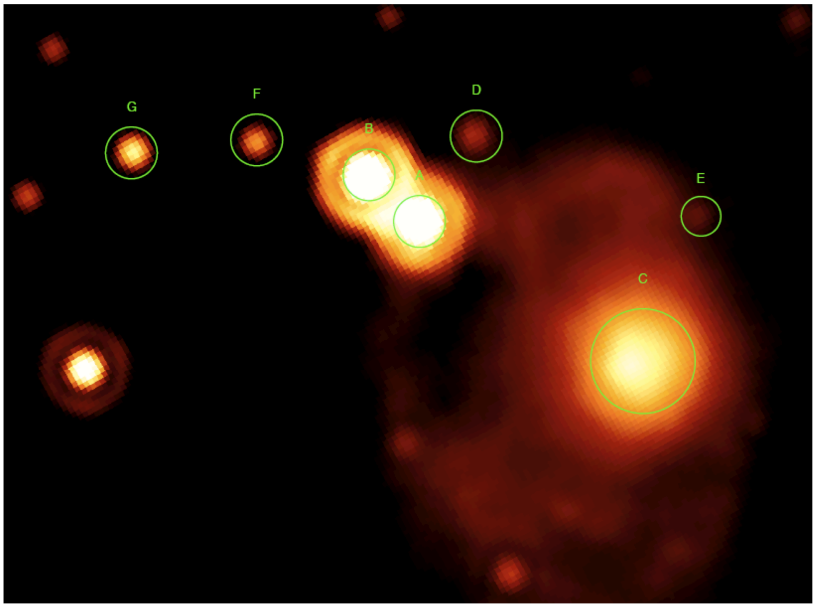
Source Mol	IR	Angle	L_{tot}	Time	M_{env}	$\dot{M}$	$T_\star$	$M_\star$	$R_\star$	Main[a]
38	B	31	1.3E+04	1.1E+04	3.0E+03	5.3E-03	7.8E+03	15	58	Y
	C	87	5.3E+03	1.5E+04	1.5E+03	2.9E-03	6.5E+03	12.9	60	Y
	D	75	6.3E+01	1.9E+05	2.2E+00	3.7E-05	4.9E+03	4.5	15	N
	E	21	2.1E+03	2.3E+06	2.5E-07	0	2.1E+04	7.5	3.3	N
	F	18	2.1E+03	4.5E+06	6.6E-08	0	2.1E+04	7.2	3.3	N
	G	41	9.4E+02	4.5E+06	1.0E-07	0	2.0E+04	6.6	3.1	N
45	A	18	9.0E+03	9.6E+03	3.1E+03	5.1E-03	4.8E+03	12.2	104	Y
	B	63	2.5E+01	1.8E+05	5.6E-02	1.0E-05	4.6E+03	3	9.8	N
	C	69	6.4E+00	5.4E+06	7.4E-04	0	1.1E+04	2.7	1.92	N
59	A	18	1.5E+04	3.4E+03	6.7E+01	1.0E-04	4.4E+03	17	158	Y
75	A	31	6.1E+03	1.6E+03	6.7E+01	1.0E-04	4.2E+03	13.6	109	Y
	B	41	2.3E+03	6.3E+04	1.1E+03	1.9E-03	8.2E+03	8.8	22	Y
	C	81	9.9E+02	9.7E+05	4.3E+00	1.7E-07	2.0E+04	7.5	4.4	N
	D	87	9.0E+02	6.4E+05	1.9E+01	2.8E-08	2.0E+04	6.7	3.3	N
	E	18	1.3E+02	6.0E+03	1.5E+00	2.9E-05	4.2E+03	3.6	21.8	N
	G	18	3.4E+01	1.6E+05	2.0E-01	9.7E-06	4.4E+03	2.2	9.6	N
98	A	49	2.0E+03	7.9E+03	2.1E+03	2.9E-03	4.5E+03	10.8	76.4	Y
	B	81	7.2E+01	5.0E+06	3.2E-06	0	1.3E+04	3.7	2.5	N
	C	81	4.3E+01	9.5E+04	2.6E+00	2.0E-05	4.7E+03	6.5	29.7	N
	D	87	4.5E+01	5.6E+06	6.0E-01	5.6E-06	1.8E+04	6.1	3.4	N
	E	63	6.2E+01	3.4E+03	4.0E-01	1.8E-05	4.1E+03	1.6	14	N
109	A	69	3.4E+03	1.5E+03	1.3E+02	3.9E-04	4.2E+03	9.5	102	Y
	B	41	2.4E+02	6.0E+03	2.5E+00	4.0E-05	4.2E+03	4.9	31.6	N
	C	87	1.4E+02	4.7E+04	5.0E-01	2.4E-05	4.5E+03	5.9	23.9	N
	D	87	8.9E+02	4.9E+05	4.2E+00	4.8E-06	1.7E+04	6.6	5.3	N
	E	87	2.0E+01	4.9E+06	4.9E-05	0	1.33E+04	3.6	2.4	N
	F	49	9.4E+01	2.5E+04	5.0E-01	8.5E-05	4.4E+03	3.16	16.5	N
	H	18	7.0E+01	9.7E+03	7.0E-01	2.4E-05	4.3E+03	2.4	16.5	N
110	A3	41	3.7E+03	2.7E+03	3.4E+02	5.2E-04	4.3E+03	13.2	110	Y
	B	87	1.2E+01	2.6E+06	2.0E-02	5.0E-06	8.0E+03	3.0	5.2	N
	C	56	4.0E+01	4.5E+03	1.0E+00	1.1E-05	4.2E+03	1.8	13.2	N
	D	81	3.2E+01	6.7E+06	4.0E-08	0	1.3E+04	3.5	2.2	N
	E	87	6.2E+01	9.4E+06	1.3E-05	0	1.1E+04	2.5	1.8	N
	F	56	6.6E+02	3.0E+06	1.2E-07	0	1.8E+04	5.8	2.9	N

[a] This keyword is yes (Y) for the main source.

[b] This value is zero when the star is already an the zero-age main-sequence.

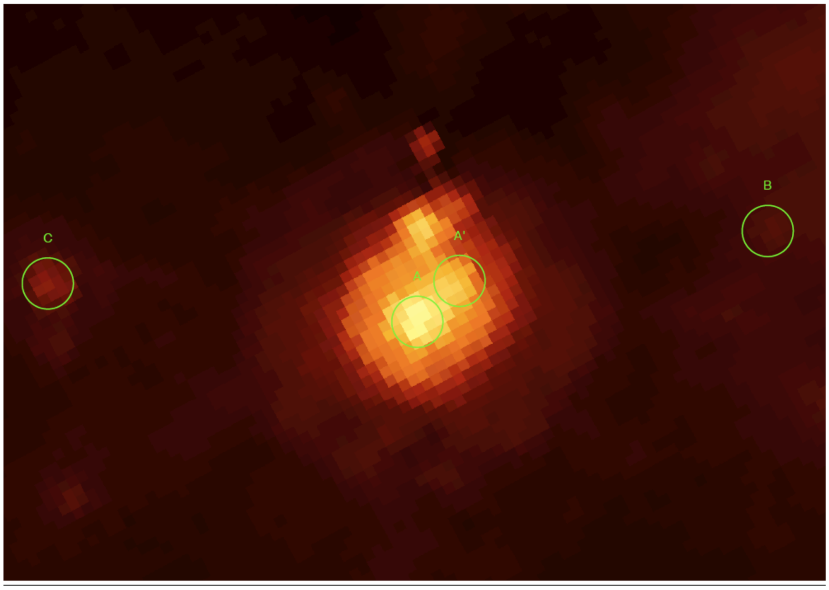


(a) Field image at 8 μm

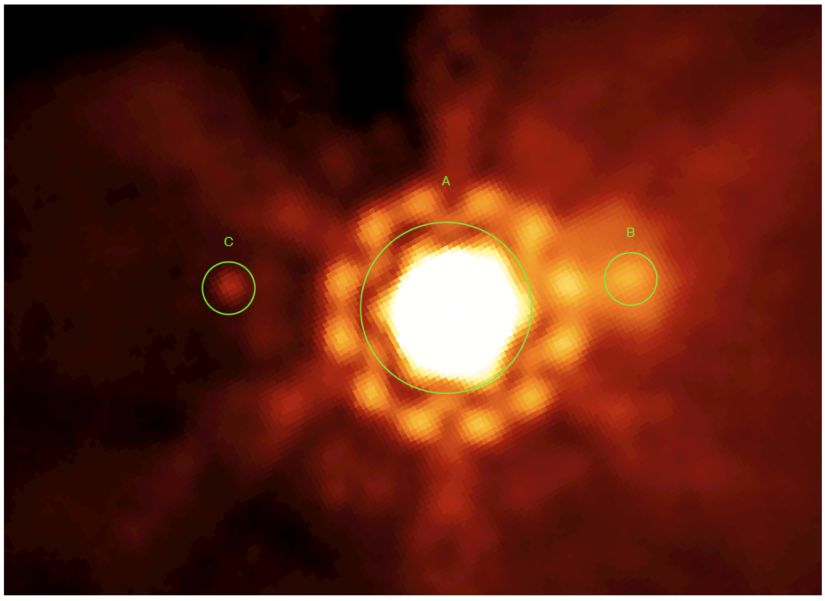


(b) Field image at 24 μm

Figure 3.7. We present two image of the field Mol038 where are identified with an empty green circle the detected objects with associated its source name(green letters). a) a mid-IR image at 8 μm , b) an image at 24 μm.



(a) Field image at 8 μm



(b) Field image at 24 μm

Figure 3.8. The same of Fig. 3.7.

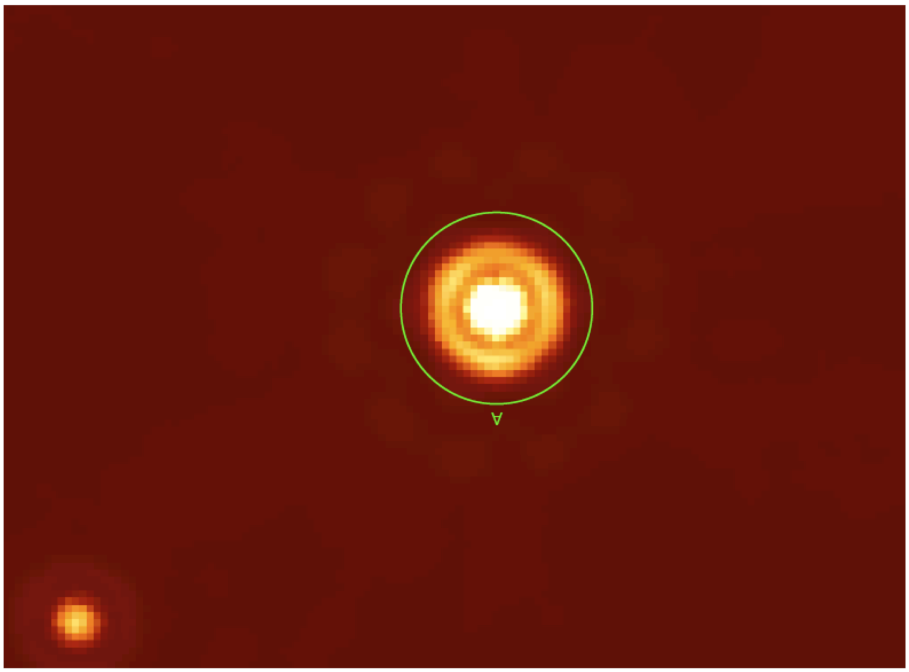


Figure 3.9. The same of Fig. 3.7.

Mol045

Three objects can be associated to the star forming region in the 24 μm, one much bright (A) while the other two are less intense (B and C). Source A corrisponds at the peak at 70 μm and the millimeter, it result very young (9×10^3) from the fitting with SED model, with a great reserve of envelope and a big mass accretion rate, about 5.1×10^{-3} $\dot{M}$. Its mass is high (12.24 $M_\odot$).. Even the source B is a pre-main sequence object with a less mass (3 $M_\odot$), its envelope is *quasi* exhausted and its $\dot{M}$ is no high, it's compatible with its mass.

The last source is a main-sequence object, with an ended accretion (envelope and $\dot{M}$ null). Its final mass is about 2.7 $M_\odot$.

Our three source coincide with the source A in the MSX images.

Mol059

A single source is visible in this region at 24 μm, it is associated to the far-IR and millimetric object. The Mid-IR fluxes in the IRAC bands are lower limits because this source is too bright and it is saturated in GLIMPSE images and in MIPS 24 μm, just that at 24 μm the saturation was been resolve using interpolation routine of Jarret, while in the other bands the problem persists.

This objects its very massive and young (about 17 $M_\odot$ with an age of 3×10^3) but its residual envelope is poor but the mass accretion rate is not so low (1×10^{-4}). The source is not a MS source because it has even a very big radius tipical of a very young collapsing source.

Mol075

Six sources have a complete observative SED of the several objects that are present at 24 μm. Two of these have already reach the main-sequence (C and D) while the others are younger.

The main object (A source) that is associated to 70 μm core is in its accretion phase but it is in proximity of the zero-age main-sequence. Its mass is high (13.60 $M_\odot$).

The other three are pre-main-sequence objects, two of intermediate masses (E and G with rispective 3.6 and 2. $M_\odot$now and 9 and 6 $M_\odot$when they arrive on the main-sequence) and the last (source B) is an high mass stars (about 9 $M_\odot$today but 43 $M_\odot$when it starts to burn Hydrogen).

Even in this region the evolutive situation is complicated and there are several sources in different conditions.

Mol098

In the 24 μm image there are five bright sources and for all of them we can build an observative SED.

Sources B and D are main-sequence stars with intermediate masses. While the other three are young pre main-sequence stars. Source A is the main source, it is the more massive stars in the full of accreting phase (about 10 $M_\odot$and an $\dot{M} \sim 2.9\times10^{-3}$). The last two are protostars one (C) is in proximity of main sequence while the other (E) is the younger source of this star forming region, its actual mass is about 1.6 $M_\odot$so its mass accretion rate is not so high as that of A source.

Mol109

Seven sources are fitted with the SED models of rich group that appears at 24 μm. Two objects are main-sequence stars (D and E), they have an intermediate mass (6.6 and 3.6 $M_\odot$) and ages of about 5×10^5 and 5×10^6 years respectively. Five stars are pre-main sequence objects, they have ages included between 1.5×10^3 and 4.7×10^4.

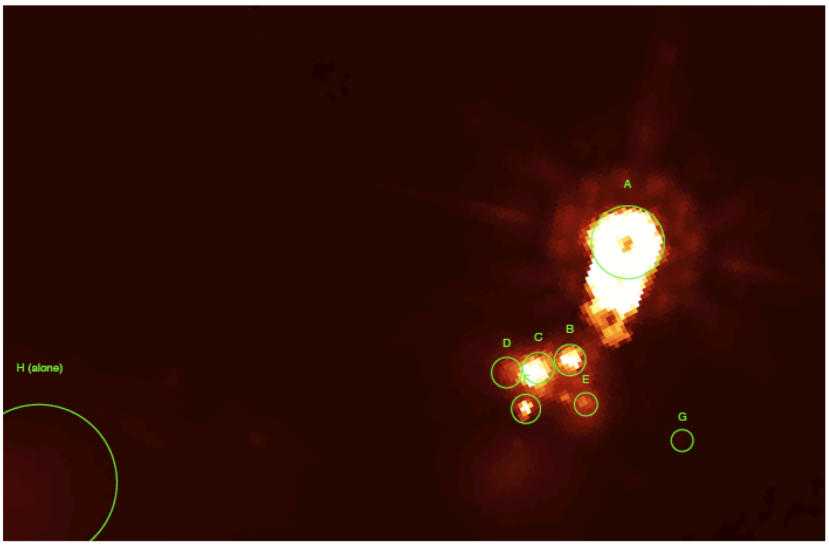
The younger source is the more massive (A the main objects), it has a residual envelope and its accretion phase is active (with an accretion rate about 3.9×10^{-4}).

The other sources are pre-main sequence objects with an intermediate-high mass but the accretion rate is lower about 10^{-5} and the residual envelope is too poor to permit them to reach their zero-age pre main-sequence mass.

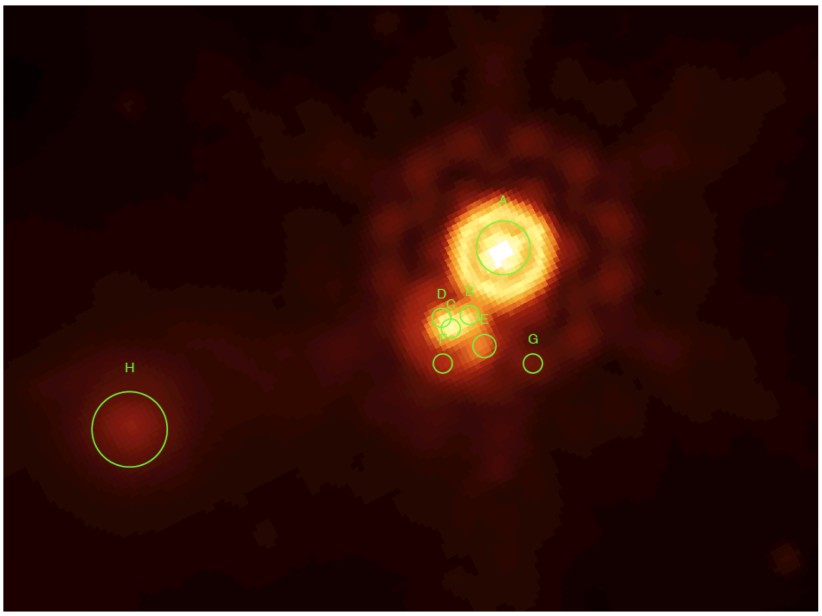
Mol110

Mol110 at 24 μm present a complex situation with compact sources and diffuse emission. A group of several sources is visible in this band, but only for six of them we can build a complete observative SED.

Three sources (D, E and F) are evolved sources that have finished their accretion phase, while source B that have not reach its main-sequence mass but its acretion rate is low and the residual envelope poor, indeed its SED appear as a zero-age main-sequence source and has an age compatible with its arrive on the main-sequence.

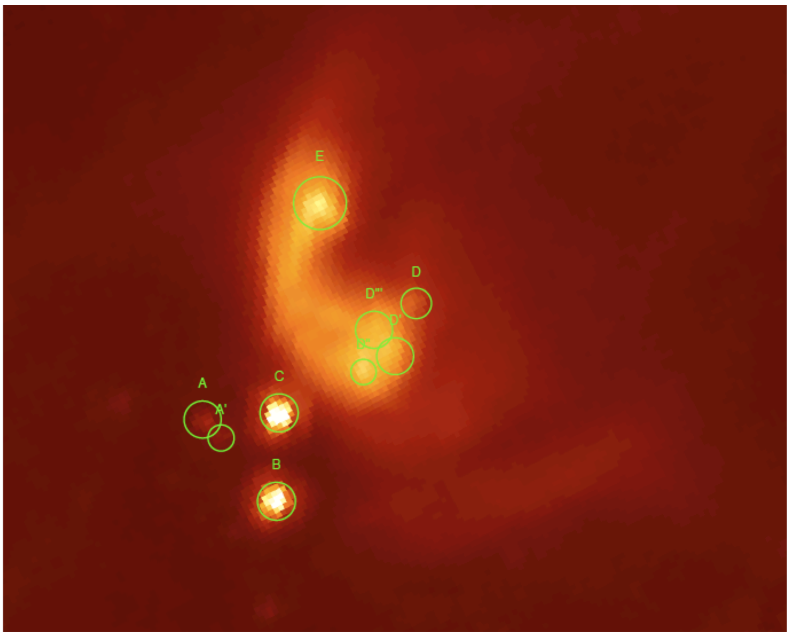


(a) Field image at 8 μm

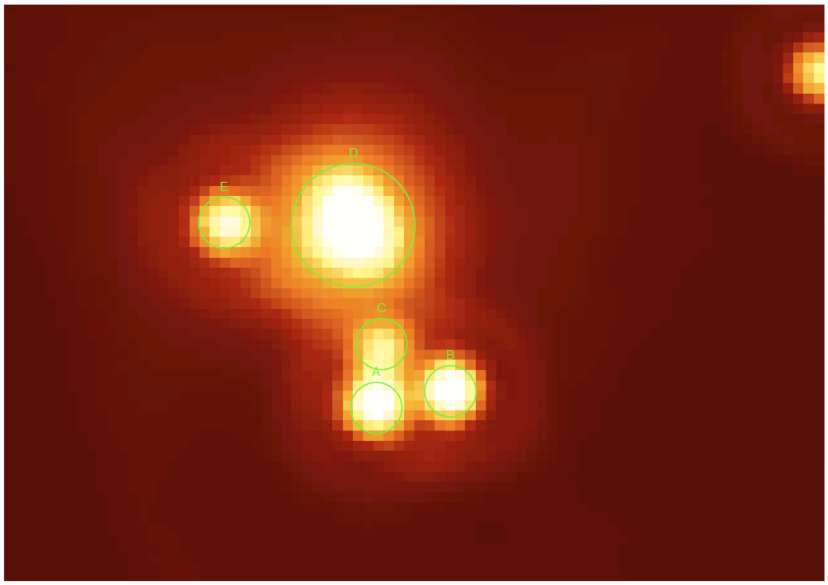


(b) Field image at 24 μm

Figure 3.10. The same of Fig. 3.7.

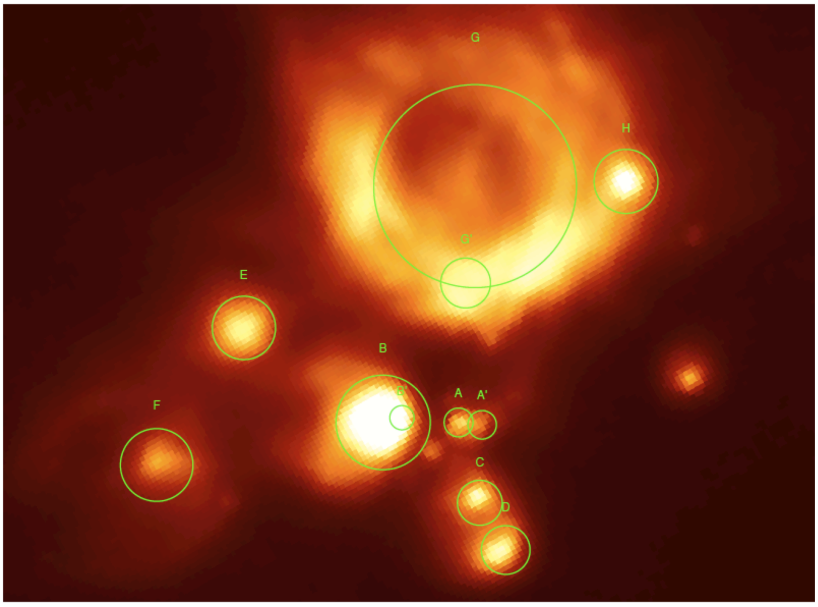


(a) Field image at 8 μm

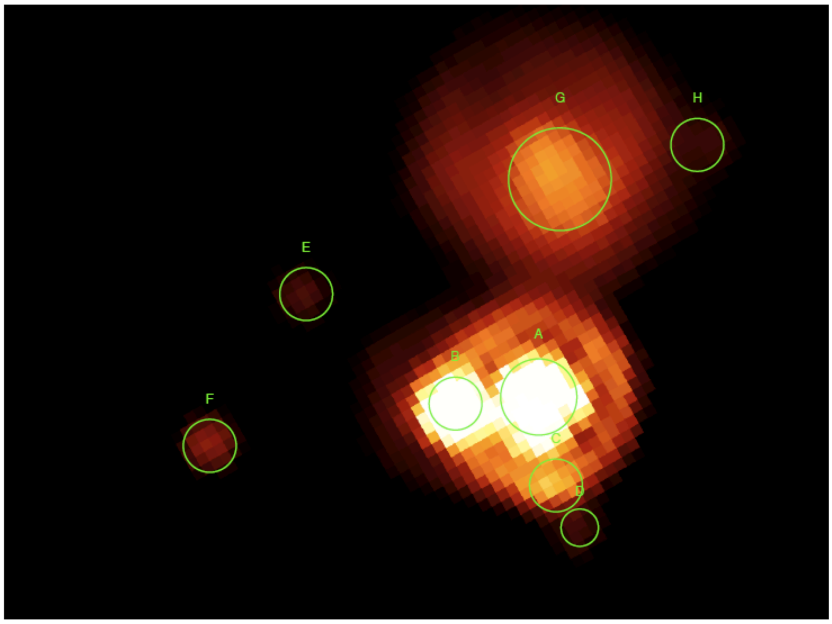


(b) Field image at 24 μm

Figure 3.11. The same of Fig. 3.7.

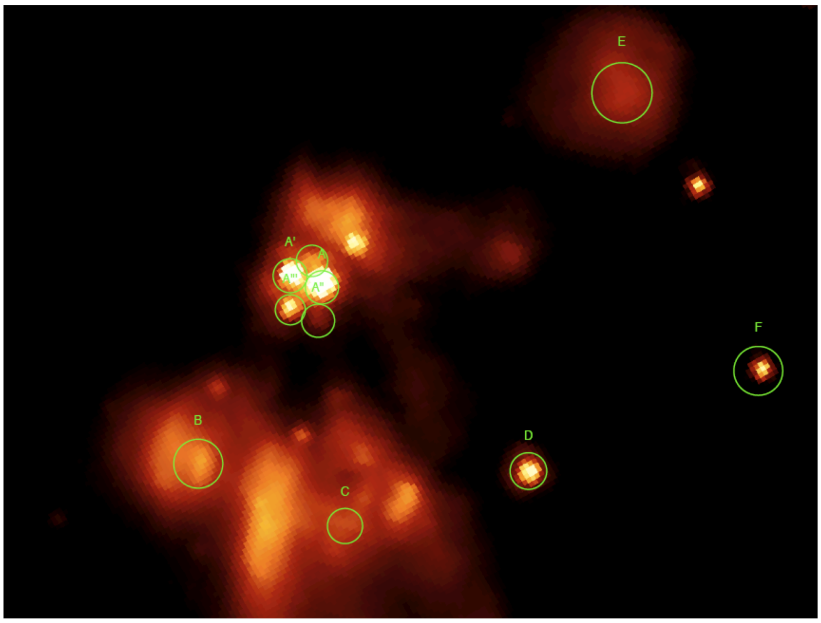


(a) Field image at 8 μm

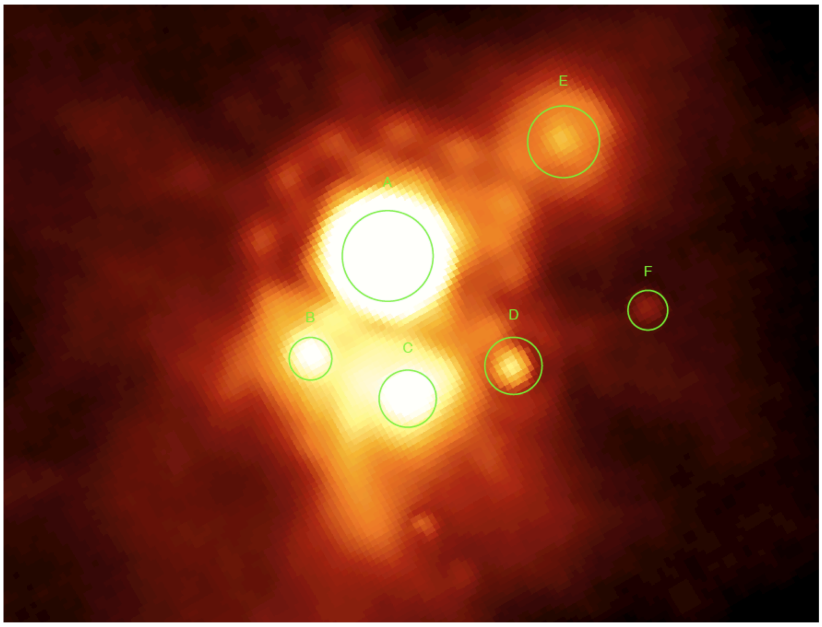


(b) Field image at 24 μm

Figure 3.12. The same of Fig. 3.7.



(a) Field image at 8 μm



(b) Field image at 24 μm

Figure 3.13. The same of Fig. 3.7.

Only the objects A and C are pre main-sequence sources, the first is the main object and has an high mass value with ana a still intense $\dot{M}$ value. Even source C is young and its accretion is active, but its mass is smaller respect to that of principle source.

3.5 Discussion

3.5.1 Young Stellar Object evolution: observative diagram

The diagram in Fig. 3.14 is used in the previous SED analysis (see Molinari et al. (2008a)) and was introduced by Saraceno et al. (1996) in their study of low mass young objects. As previously said in the introduction, for the high mass sources the HR diagram is no a good instrument to understand their evolutive phase in an intuitive and simple way because high-mass protostars arrive on the ZAMS still complitely obscured from their dust envelope, so there is not possible estimate their T_{eff} from the photosphere emission. This diagram L_{bol}-M_{env} can be used as an evolutive diagram, as the HR, but it is usefull even for the high-mass objects.

In the work of Molinari et al. (2008a) we identified two tipologies of observative SED for the source sample, one fitted with an embedded ZAMS models ("IR" source) and another fitted with two components: a greybody model ("MM" source) and a moderately embedded ZAMS ("IR-S"). In detail, sources for which successful fits could be found are labeled “IR-P” (filled black circles in Fig. 3.14). If a fit cannot be found we adopt a two component approach where the warm Mid-IR portion of the SED was successfully fit with a moderately embedded ZAMS (the “IR-S”, asterisks and plus in Fig. 3.14), while the cold Far-IR/mm part was fit with a single temperature grey-body (the “MM-P”, empty circles in the diagram). Positionated on the diagram, these three tipologies of sources occupy different region of the figure that coincide with extrapolation at high mass of the position of low mass Class II (==> IR-S), Class I (==> IR) and the Class 0 (==> MM).

Where are located our sources in this diagram? We find that a lot of the main objects are in the plain region where are localized the MM sources (red diamonds in Fig. 3.14), only two objects that are fitted with an embedded ZAMS are in the middle of IR and IR-S regions (yellow diamonds). Two of the red diamonds (the more thick) are no main source, but are fitted with more young pre main-sequence objects. Blue diamonds are the secondary objects classify as pre main-sequence sources, they have actually intermediate masses included between 1.6 and 6.0 $M_{\odot}$. They are positionated in the Class II region of the diagram as we aspect seing their phisycal conditions.

These three regions (MM, IR-P, IR-S) are populated by objects with several chracteristic that correspond really to different evolutive phases, indeed sources that belong to these region have analogous differences to those present between the different phases of low mass objects (Class 0, Class I, Class II).

We present a brief summarizing scheme of the distribution of our sources in the diagram, to underline the changes in the physical properties (such as luminosity, mass accretion rate, temperature, radius and obviously age) that can correspond to different evolutionary state:

- **MM**: sources that are located in this region present some characteristic of young accreting objects, indeed they are fitted as very young protostars with a mean age of few 10^3 years. They have a very high accretion rate $\dot{M}$ in agreement with the

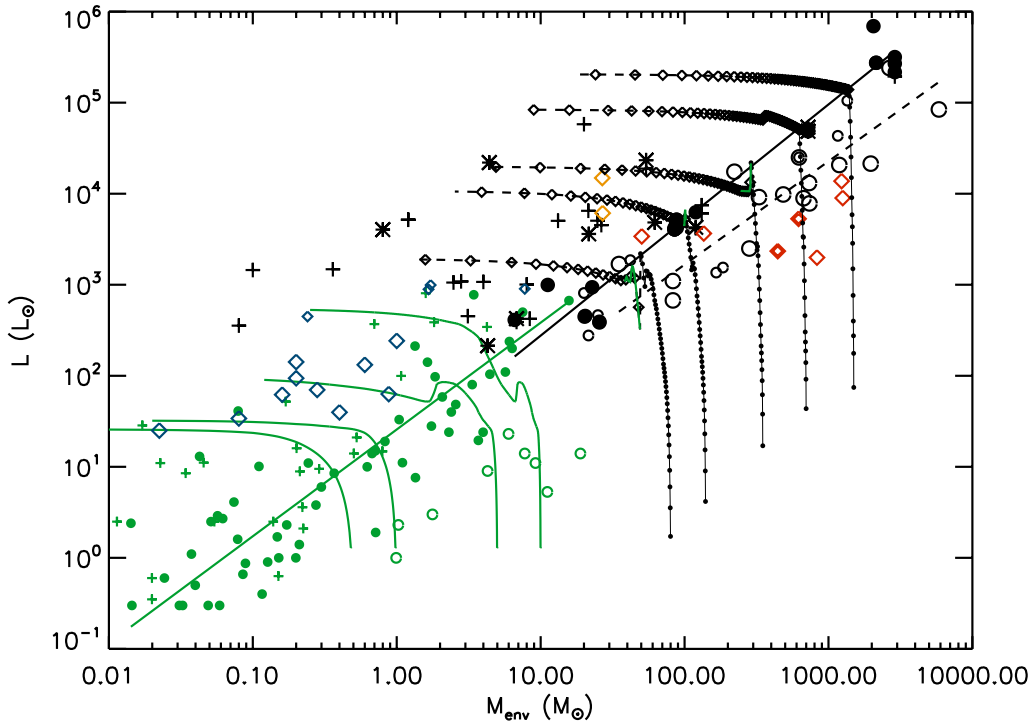


Figure 3.14. L_{bol}-M_{env} diagram for sources presented in Table 6. Lines and symbols in green are from Saraceno et al. (1996) and depict the situation for the low-mass regime (Class 0, I and II sources are represented respectively by open circles, filled circles and crosses, with the full thin line representing the log-log linear fit to the Class I source distribution). The heavy black lines and symbols represent the sources in the work of Molinari et al. (2008a). MM sources are represented by open circles (smaller circles for MM-S objects) while IR-P, IR-S1 and IR-S2 are indicated by filled circles, asterisks and crosses, respectively. Lines represent the models as described in the text; limited to the high-mass regime, the dot-marked full thick lines represent the accelerating accretion phase (the dots mark 10^4yr intervals), the grey line is the portion of the pre-MS evolution, while the dashed diamond-marked portion (at 10^5yr intervals) represents the final envelope clean-up phase. The different models (left to right) are for different initial envelope masses of 80, 140, 350, 700 and 2000 $M_\odot$. The full and dashed black lines are the best log-log linear fit to the IR-P and MM-P sources, respectively. Empty diamonds are the objects analized in these work. Red diamonds are younger objects, the two symbols more thick are young sources but are not the main objects of the field. The two yellow diamonds are main objects but they are ZAMS objects. The blue diamonds are the secondary objects fitted as pre main-sequence sources.

theorized value for accreting high mass protostars (few 10^{-3}), and an high luminosity (about few 10^4 $L_\odot$) typical of high-mass objects. They have big values of $R_\star$ and of residual envelope masses that underlined their condition of pre main-sequence stars. All the objects of our sample that are situated in this diagram zone are main sources.

- **IR-P**: Two of the our main sources are located in the IR-P region. They present an high luminosity and a big $R_\star$ (about 100 $R_\odot$) but a mass accretion rate lower than the $\dot{M}$ of MM objects (even if they have an high $\dot{M}$ about few 10^{-4}), and a lower residual envelope. This characteristics underline their a step forward in their evolutive way reapect to the MM sources.

- **IR-S**: The last two main source are situated in the region of IR-S, they have already reach the mass that must to have on the main-sequence and are associated to a still lower $\dot{M}$ and envelope, but they are young and with a radius and a luminonosity much big tipycal of no ZAMS objects. They are problably reaching the main-sequence.

- **Class II**: About the half of our secondary objects are more or less evolved protostars of intermediate masses. They are older than the main sources, with a very low residual envelope and $\dot{M}$ but a large radius, they have not reach their ZAMS mass so they are pre main-sequence objects. Their properties are analogous to IR-S ones, scaled down due to the lower mass values.

So these sources show different physical properties that can corrispond to several evolutive phase and that locate them in a different region of L_{bol}-M_{env} diagram. Is there a theory that can justify all that we see in the observations and the changing position between a diagram region to another?

In Fig. 3.14 we use the **?** to predict the behaviour of luminosity and circumstellar mass as a function of time, dot-marked full thick lines, as explain in the next paragraph. These lines seem to describe the passage between the different region of diagram in a very good way supporting an evolutionary scenario.

3.5.2 Young Stellar Object evolution: model

A lot of the current high-mass star formation theories seem to converge on a scenario where accretion is a time-dipendent process and its rate increasing with time during the formation. In turbulent core models this is due to the expanding inside-out collapse reaching out to material characterised by supersonic speeds and where higher accretion rates are possible; in competitive accretion models (Bonnel et al. 2003) the object doomed to be the higher mass in the region is the one which happens to be closer to the potential well where higher is the density and the probability where fresh material not initially bound to the collapsing core is drawn from the outer clump regions.

(**?**) offer a convenient analytical prescription for the accretion rate as a function of time. In particular, for a standard free-fall envelope density profile the accretion rate is expressed as

$$\dot{m}_\star = 4.6 \times 10^{-4} \left(\frac{m_{\star f}}{30 M_\odot} \right)^{3/4} \Sigma_{cl}^{3/4} \left(\frac{m_\star}{m_{\star f}} \right)^{0.5} M_\odot / yr \qquad (3.2)$$

where $m_\star$ and $m_{\star f}$ are the instantaneous and the final stellar mass, respectively; Σ_{cl} is the core surface density in $g\,cm^{-2}$. Once $m_{\star f}$ and Σ_{cl} are fixed it is easy to compute

the evolution of $m_\star$ as a function of time; in the L_{bol}-M_{env} diagram we will actually follow the envelope mass which, from a given initial value M_{env_i} will decrease in time both because of the material accreting onto the star at a rate $\dot{m}_\star$, and because of the material expelled by the powerful molecular outflows which characterize the earliest phases of massive YSOs formation. The fraction of accreting material that is expelled in the outflow is deemed to be of the order of f_{out} ~0.5 (**?**). An additional effect contributing to the clump envelope dispersal is the fact that massive YSOs form in clusters, and so there will be other less massive YSOs in the same cluster which will draine material from the molecular clump. Faustini et al. (2009) have detected young stellar clusters toward many of the objects included in the present work, and a relationship is found between the mass of the most massive star in the cluster and the total stellar mass

$$Log[m_{\star Tot}] = 0.55 + 1.41\, Log[m_{\star Max}] \tag{3.3}$$

Assuming $m_{\star Max}$=$m_{\star f}$, for each choice of $m_{\star f}$ we derive an estimate of the total mass bound in stars; adopting a mean stellar age of $t_{cl} \sim 3 \times 10^6$ years for the clusters (Faustini et al. 2009), we can define a time-averaged rate for clump dispersal onto cluster stars (except the most massive one), as

$$\dot{m}_{cl} \sim \frac{m_{\star Tot} - m_{\star f}}{t_{cl}} \tag{3.4}$$

The evolution of the clump mass will then be

$$M_{env}(t) \sim M_{env_i} - [t\ (\dot{m}_\star \cdot f_{out} + \dot{m}_{cl})] \tag{3.5}$$

As concerns the total luminosity emitted by the star **?** plot in their fig. 6 the luminosity as a function of stellar mass, taking into account the contributions from accretion, deuterium burning and core contraction; we extracted a point-by-point L*vs* $m_\star$ relationship and used it in our toy-model. The relevant free parameters in modelling the evolution of the massive YSO in the L_{bol}-M_{env} diagram are the initial mass of the core envelope and the final mass of the star $m_{\star f}$ where we assume that the main accretion stops; this part corresponds to the black dotted full line section of the evolutive tracks reported in fig.3.14 for different initial envelope masses: M_{env_i}=[80, 140, 350, 700, 2000] $M_\odot$.

For any choice M_{env_i} the tracks will all start at very low luminosity; as the stellar mass grows the tracks cross the area occupied by the MM sources and proceed almost vertically toward the IR-P sources. It is important at this point the mass $m_{\star f}$ for which we choose to terminate the main process of accelerating accretion according to eq.3.2: for the tracks in fig. 3.14 we adopted $m_{\star f}$=[5.5, 8, 13.5, 18, 35] $M_\odot$ to simulate a situation where the stellar formation is ended in correspondence of the locus of the IR-P sources[3]. Looking at the dot frequency along the tracks (each dot represents a span of 10^4 years) we see that

[3]This is clearly arbitrary but the alternatives are unlikelies. We can choose to stop the accelerating accretion in the MM region. With these scenario the evolution from MM to IR-P objects would not only be due to an increase of stellar mass and luminosity caused by the residual accretion, but also to the envelope clearing process. Since after $t = t_{\star f}$ the accretion runs at a much slower rate, the duration of this process would be of the order of few 10^6 years, much longer than the few times $10^4 \div 10^5$ years duration. Besides, the possibility that the main accretion can be extended to later times with respect to the standard scenario seems excluded by the fact that the MM and IR-P sources are distributed in a rather well confined strip of the L_{bol}-M_{env} diagram. Since the accelerating accretion runs much faster than the subsequent residual accretion phase, it is quite unlikely that the accelerating accretion can be prolonged beyond the IR-P region.

evolution speeds up as we go toward the rightmost track not only because of the higher $m_{\star f}$ which justifies a higher accretion rate according to eq. 3.2, but also because of higher clump surface densities Σ_{cl}. Indeed we assign Σ_{cl} using a relationship which we find in our MM sources between the core mass and the column density, determined dividing the masses in table 6 of Molinari et al. (2008a) by the core extensions derived from the fits to the millimeter cores:

$$Log(\Sigma_{cl}) = -2.0 + 0.7\, Log(M_{env_i}) \tag{3.6}$$

Based on our choice of the initial envelope masses we derived, following eq. 3.6, Σ_{cl}=[0.25, 0.36, 0.7, 1.15, 2.4] g cm^{-3}, in good agreement with the interval measured for Galactic cores as reported by **?**. The times by which the final masses $m_{\star f}$ are achived in the simulations of fig. 3.14 are $t_{\star f}$=[4.8, 4.0, 2.8, 2.0, 1.4]$\times 10^5$ years, and at that time the accretion rates are $\dot{m}_\star$ ~[0.46, 0.8, 1.9, 3.4, 10]$\times 10^{-4}$ M$_\odot$ yr^{-1}. These rates have significant impact on the mass at the arrival onto the ZAMS. Palla & Stahler (1992) showed that the accretion flow has the important effect of delaying the time of arrival onto the Main Sequence; a YSO that joins the MS at 8M$_\odot$ for $\dot{M}_{acc} = 10^{-5}$M$_\odot$ yr^{-1} would see its ZAMS mass increased to 15M$_\odot$ if $\dot{M}$ was 10 times higher. Rates even one order of magnitude higher, as our simulations would seem to suggest, are usually postulated based on the inferred mass loss rates measured for molecular outflows in high-mass YSOs (Zhang et al. 2001, 2005); under these conditions the arrival onto the ZAMS can be deferred to higher and higher masses provided the required ever higher accretion rates can be maintained.

When $m_\star = m_{\star f}$ the accretion that has been driving the protostar assembly according to eq. 3.2 will stop; our choice of $m_{\star f}$ is clearly arbitrary at the moment, and in turbulent core models this basically corresponds to setting the core radius and density structure. At this time the YSO is already on the ZAMS, or very close to it, if sufficiently massive (the rightmost three tracks in fig. 3.14), or it will approach it following standard PMS tracks (Palla & Stahler 1993; Behrend & Meader 2001)) if not. This early ZAMS phase correspond to the stages observationally identified with Hot Cores and UCHII regions. The very high accretion rate that we find at the end of this phase is compatible with the outflows rate of observed hot-cores and HCHII regions, and it is suitable to confine the expanding HII regions, as it's necessary to justify the lifetime of Hot-cores and HCHIIs region, as we explain in the introduction. After the end of accelerating accretion, the clump mass decrease because of material expelled in the molecular outflows (at rates = 0.5 $\dot{m}_\star$) and drained by accretion onto other objects forming in the same regions. The luminosity radiated in this second phase is expressed as $L = L_\star + L_{acc}$ where $L_\star$ is the luminosity radiated during the approach to the ZAMS aong standard pre main-sequence tracks (if the object is not already on the ZAMS, $m_{\star f} \lesssim 18$M$_\odot$), or ZAMS luminosity. While the accretion luminosity is computed using eq.(74) in **?**. $L_\star$ and L_{acc} will vary with time because both depend on the stellar mass that will increase with time due to the persisting residual accretion. These second phase corresponds to the dashed section of the evolutive tracks in fig. 3.14, and spans the region between IR-P and IR-S. The simulation is stopped when the clump disappears.

These simple model does not the ambition to give a detailed description of the formation process but it is sufficient to give a qualitative descrption of the evolution phases and to correlate the different region of the L$_{bol}$-M$_{env}$ along an evolutive way.

3.5.3 Sources ages

The complex evolutive situation appears clearly from the scenario presented in the last paragraph and in the previous summarizing scheme. Each single high-mass star forming region presents a no trivial star formation history, as we can explain in deteail the next Chapter combining the results of the Near-IR analysis with the SED one. This situation is deduced even from the SED analysis alone, when you can resolve the several cluster members present in a region. This fact support a scenario of a process of formation that carries on in a continuous way for a long time.

Chapter 4

Star Formation History of High-Mass Star Forming Regions

Observational evidences show that theories of high-mass star formation based on accretion are preferred. Indeed the presence of a disk and powerfull outflows appear to be ubiquitous in star-forming regions, these features are tight related to accretion process and they are a discriminative factor to reject or accept some theories.

Coalescence foresees that the formation of an high mass object takes place through the merging of low-intermediate mass protostars in the central region of very dense and crowded cluster. This process foresee the distruction of the the disk and the jet that are present in the original protostars, so the evidence of the presence of these objects associated to high-mass protostars need to tend to other desciption of high-mass objects formation, such as process based on the accretion models.

4.1 Accretion models

This branch of the star-formation model include several theories rapresented from two principal models:

- *Competitive Accretion*. Accretion in clusters is a dynamic phenomenon as both the stars and the gas respond to the same gravitational potential. Accretion rate are highly non-uniform with the stars nearer the cluster center that accreting more than the others.

- *Gravitational Collapse in Turbolent Medium*: Massive stars form in regions supersonically turbulent, where the energy due to this turbulence is comparable or greater than the energy due to the gravitational collapse.

In these theories the dominant paradigm for star formation is gravitationally collapse but they are based on a different scenario of the cluster formation and foresee different evolutive time. In both theories, a star initially forms when a gravitationally bound gas core collapses. The crucial distinction between them is their prediction for what happens subsequently. In gravitational collapse, after a protostar has consumed or expelled all the gas in its initial core, it may continue accreting from its parent clump. However, it will not accrete enough to change its mass substantially. In contrast, competitive accretion requires

that the amount accreted after the initial core is consumed be substantially larger than the protostellar mass.

The competive accretion model is ineffective if the medium has turbolent energy comparable to the gravitational energy, in a approximately virialized clumps, whereas it is effective if the turbolence is sufficently weak. So Bonnel and his collaborators (Bonnel and Bate 2006) argue that the gas throughout star-forming clumps has a very low turbulent velocity so that protostars in clusters can accrete efficently. On the other hand, Krumholz, McKee and Klein (2005) argue that stellar feedback and the cascade of turbolence are sufficiently turbulent, at all the scales, to be virialized.

The timescale for the formation of star clusters is an important discriminant between these models. Star clusters form in about 1-2 t_{ff} if turbulence is allowed to decay (bonnel bate e vine 2003), whereas it can be significantly longer if turbolence is maintained, for the formation of massive objects from turbulent cores Mckee & Tan foresee several t_{ff}.

This controversy can be resolved through a more detailed decsription of the star formation history of the high-mass clusters.

4.2 High-Mass Clusters SFH

Can we understand something about the story of the clusters formation? Can we use it to discriminate between these formation models?

High-mass protostars are actually "unknown" objects for several complication that we have explain in the introduction of this work (see Chapter 1). For those reasons our understanding of the formation process of these objects is actually not as good as for the low-mass protostars.

Our multiwave analysis has pointed out some aspects of the high-mass star formation regions that have not emerged in the past. This was due to the capability to analize, in the same region, different stellar population in several evolutive phase. Indeed, we can find interesting results from the Near-IR clusters analysis and from the SED building from Mid-IR to millimetric bands separately, but the most usefully results is found where the two studies are considered togheter.

First, clustering seems almost ubiquitous in young High-mass star forming regions, where low mass objects are numerically dominant. In our Near-IR study we find young clusters (mean ages about 10^6 years) detected around the position of our sample of candidate high-mass protostars. The cluster members age shows a spread, within each clusters, between a few 10^5 to few 10^6 years. We have verified that this is a real spread in the age distribution of low-mass objects, rejecting the possibility that it is related to another phenomenon, such as the presence of a lot of unresolved binaries, or multiple systems, in our clusters.

On the contrary, the highest mass objects appear to be only 1×10^4-1×10^5 years old at most (depending on mass). The high-resolution of our data permits us to resolved some subclusters of intermediate-high mass in the Mid infrared bands. The analysis of these sources reveals that an age spread in present even within of these last group, where the massive object is, in all the examinated cases, the younger objects. These objects are prevalently younger than the Near-IR low-mass cluster, and a lot of them are pre-MS sources. The most massive source of each field is associated with the first phases of evolution of a protostars, they are in the main accretion phase where its evolution is dominated from the accreting

process.

These global situation show a non-trivial distribution of the evolutive phases within the each star forming region and give us the possibility to try some conclusions about the star formation history in the high-mass clusters. We find that low-mass evolved protostars, that are dominant in number within the star forming region, cohesist with a smaller group of intermediate-high mass objects that are younger pre-MS stars, but the emission in the Far-IR and millimetric band is dominated by the emission of a one, or at most two, high-mass objects that are in the first phase of their formation. Rather than a rapid burst process, star formation appears to be a relatively long-lasting and continuous event with low-mass stars forming first and the highest-mass object appearing only at relatively later times.

Sumarized this result we can affirm that:

"The star formation is a slow and continuus process in the high-mass regions, it proceeds for several thousands and even millions od years. The age of each members within each cluster is probably related to the its mass, the low mass sources start their formation firstly, while the high mass objects are the last."

This suggests a way to resolve the doubt about different high-mass cluster evolution models. Massive clumps, where these clusters form, have free-fall time of about $t_{ff} \sim 1\times10^5$ years Tan et al. (2005); Bonnel et al. (2003). Indeed, our cluster have mean ages, eximated from the low-mass population ages that are the oldest objects in our fields, of about 10^6 and few 10^6 years that correspond to ten and few tens times the free-fall time. This is a fundamental result that can be used to discriminate between star formation theories. A so long formation time prefers those models that foresee a higly turbolent medium where the gravitational energy is balanced by the turbolence ones, such as the model of gravitational collapse of Tan & McKee (2003); Krumholz McKee & Klein (2006). The turbolence supports the cores against the gravity and do longer the collapsing time. In this scenario greater is the turbolence intensity greater is the time of formation, the turbulence can support even the more massive core, more over the Jeans mass, permits the formation of the cores without form the high mass source.

Two different process, one hand the increase of turbulence and the other the increase of temperature, can justify the other result that we find: *The more massive object is the last source to form*. Indeed when the low-mass source evolve they ignite the polar outflows, these process add turbolence to the clump where the cluster is forming. An high value of turbolence energy permit the formation of bound cores that are more massives (Tan & McKee 2003; Krumholz McKee & Klein 2006), in this way the most massive core is the last to form and so it's the last to star its collapse. Potentially this scenario leads to a formation of even more massive stars, but when a very high-mass form and ignites the hydrogen combustion it destroys the circumstellar medium and stops the formation process.

On the other hand, the radiation produced from the low-mass objects warms the clumps where the cluster is forming (Krumholz & Tan 2007). The temperature increases stabilizes the clump against a further fragmentation and permits the massive core build-up and then the formation of a very high-mass protostar.

These two process are not incompatible, and they can act simultaneously to permit the massive cores formation and to slow down their collapse.

II Part

Software Development

Chapter 5

"DERBIGA" : an algorithm for star forming regions analysis

In the last decades the use of two-dimensional images is largely diffused in astrophysics; high-resolution and sensibility images are produced at all the wavelenghts from telescopes, ballons and satellites. The great technological development has permitted the building of more careful and powerful instruments opening new horizons of modern scientific research. Actually great spatial missions can help the astrophysics to take a step forward, Spitzer will have finished while Planck and Hershel has been launched and soon we will receive their data.
Contemporaneously computers developments allows complex calculations, speeds up analysis, and permits the use of complicated softwares.

Today the scientific community has powerful means in its hands, but even they must to face up to new problems. The improvement in the observed images had often reveals some new aspects of an objects and new problematics that are "invisible" in the previous data.

A region observed at different waveleghts presents changing properties, an improvement in sensibility and resolution show distinct characteristic. The astronomers must deal with different problems correlated with the specific scientific objective that they want to reach. A different astrophysics target implies a different wavelenghts, energies, physical conditions, ecc... afterwards another world. This variedness involves a variety of different approaches to face the analysis of these data; actually there are a lot of programs to do source extraction and photometry in astrophysicsal images, and they are based on differents tecniques.

5.1 Star forming regions: data analysis problems.

The photometry of a well isolated star is, today, a resolved problem but the situation is different for a crowding region and even more for a star forming region. The photometric difficulties become much worse when we try to study these two cases which contain much of the most interesting astrophysics goals.

The star forming process can be studied in a wide intervall of wavelenghts between Near Infrared and Millimetric to investigate the different aspects of this fenomena. In the most wavelenghts, images of these regions presents some problematics, due to the contributions of quite a few objects to the emission (more or less evolved stars, gas and dust at

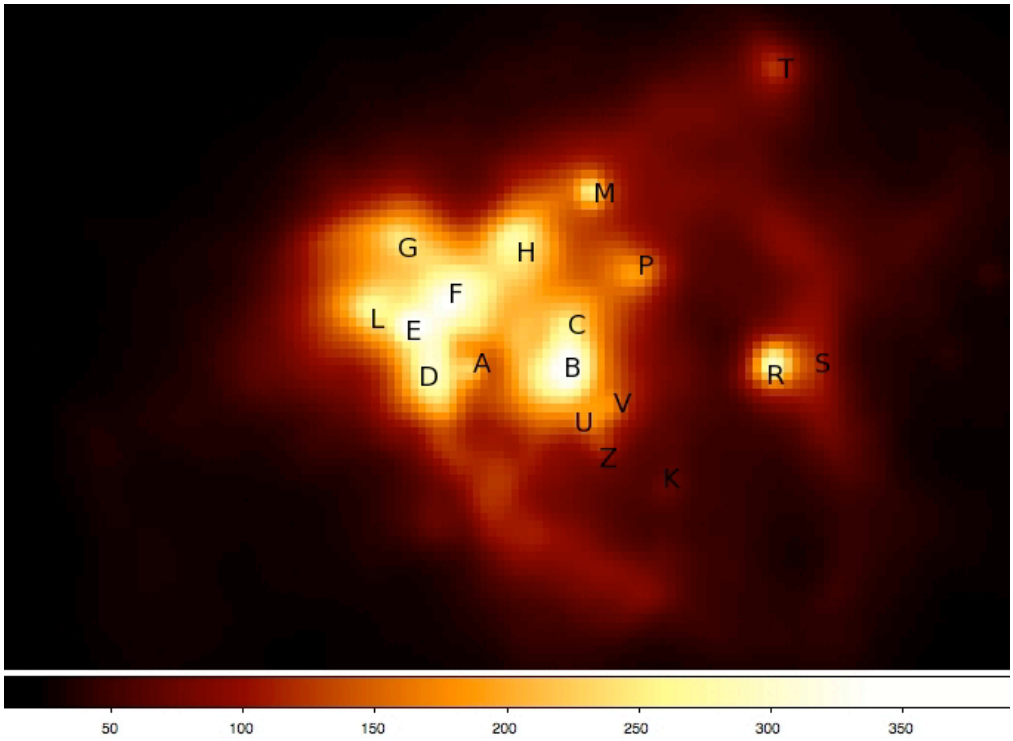


Figure 5.1. Image at 24 μm of mol160. The principal sources are indicated with a letter between A and T.

several temperatures, densities and distributions):

- *Crowding:* we can say, on the basis of the actual scientific knowledge, that all the star form in clusters (Lada & Lada 2003; Faustini et al. 2009), and their richness is proportional to the mass of the highest mass star (Testi et al. 1997, 1998).
- *Highly spatially variable background:* all star born in molecular clouds, so they are deep in gas and dust that are non-homogeneously distributed.
- *No-psf profile for all the sources:* the protostars are embedeed in their envelope during their accreting phase which no necessarily have a gaussian or sferical-like density distribution. The source profile can changing depending on the stage of the accretion phase, becoming sigar-like from a sferical distribution.

Figure 5.1 shows the Mol160/IRAS23385+6053 field at 24 μm, a classical example of crowded intermediate-high mass star forming region. We have analized its nature in the recent past (Molinari et al. 2008b). The extraction and photometry of point sources was attempted using several publicly avaible packages for automatic source detection and extraction, such as DAOPHOT (Stetson 1987), SEXTRACTOR (Bertin & Arnouts 1996), or MOPEX (Makovoz & Marleau 2005). These algorithms have excellent results in the analysis crowded fields but have some problems with regions like Mol160/IRAS23385+6053. As is apparent from the image, this field presents the problematics listed before. Indeed the various peaks that are visible sit on a complex patch of extended emission which is strongly

variable on small and large scales. This poses severe problems for reliable source detection and photometry since the various packages we tried assume either a constant background, or estimate one by using a box of fixed size across the image. Sources like A in the figure, for example, are undetected with SExtractor and also with MOPEX if the box for estimating the background is too large. Using a small box would solve this, but on the other hand leads one to significantly understimate the fluxes for sources like D or F which instead lie on a less variable and broader patch of extended emission. For this reason we decided to visually identify compact peaks in the map, and to manually perform the photometry at 24 μm with custom procedures using the GILDAS package, optimising by hand the area for the background estimate.

This way to attempt this problem is acceptable in the analysis of a single region but it's absurd to think to do an analysis by hand on a survey data. These obstacles cannot be completely exceeded with the classical methods and a new approach is necessary. We must solve this problem because we have the prospect to analize the data of a big Hershel sky survey (HiGal[1]).

5.2 DERBIGA: a new data analysis algorithm

In the previous sections we have mentionated the actual situation of software for imaging analysis, we discussed about their potentialities in the study of complex situation, like crowded fields, but the star forming regions present other problematics that are no to be resolved in an easy way.

In this section we present the reasoning and the mathematic that are on the basis of our algorithm, DERBIGA[2]. The software is written in IDL language and its use is very simple. Indeed no preliminary analysis of the image must to do, instead the other software require some informations about the input image as FWHM, rms value, ecc...
As explain following, the only input parameter that must to choose is the sigma, see subsection 5.3.3. The other parameters are automatically extimated within the algorithm or they are uselesses for the intrinsic characteristic of our procedure.

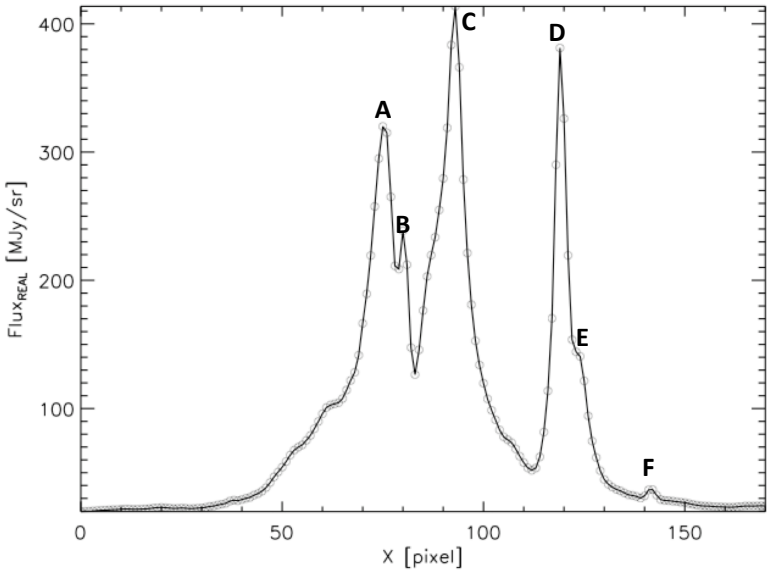
I develope the detection routine, the idea at the basis of the algorithms and its realization are mine, while the photometry one is realizes by Dr. Molinari.

5.3 Detection Software

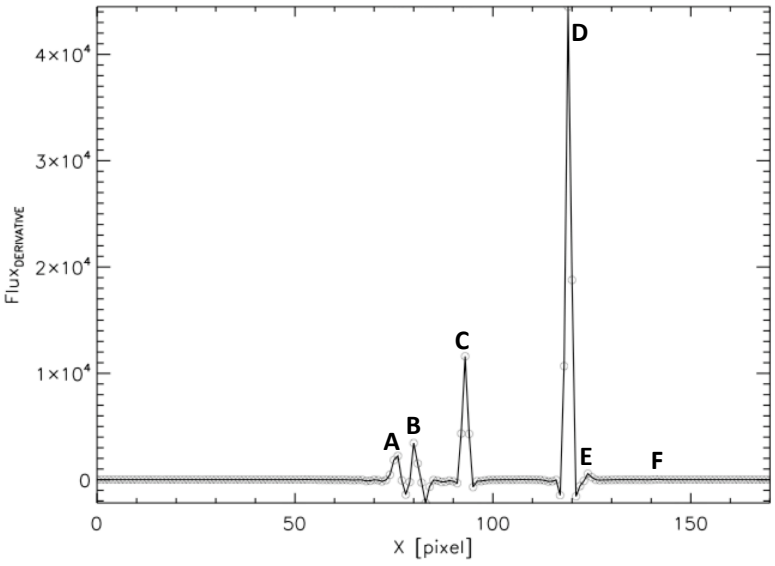
We present a detailed description of the methods applied to carry out the source detection. In simple function analysis the use of derivatives (first and second order) is basic to point out the presence of local maxima and flexa. The application of derivative techniques in a two-dimensional image isn't trivial, because it's a discrete set and the classical methods for derivative cannot be used. We use the Lagrangian methods for numerical differentiation: the 5-points formula (a description of these interpolation and differentiation techniques can be find in Hildebrand (1956)). A brief introduction to the Lagrangian methods is presented in the following section.

[1] Hi-GAL is a big international project (PI Sergio Molinari) that will cover the inner 120° (in longitude) of the Galactic Plane. We will collect data at 70, 170, 250, 350 and 500 μm simultaneously.

[2] Derivative and Bi-dimensional Gaussian fit



(a) Input image section



(b) Derivative image section

Figure 5.2. We present the result of our detection software run on a star forming region, in abscissa are plotted the pixel coordinate and in ordinate the Flux/pixel in the case of real image and the Derivate of the flux for pixel in the (b) images. (a) A vertical section of an image of a star forming region, (b) the same vertical section but in the derivative image.

5.3.1 The Lagrangian methods

The Lagrangian formula (for numerical interpolation, integration and differentiation) are expressed explicitly in terms of the ordinates involved rather then in term of their differences; this permits a more direct consideration of the effect on the result of a change in one or more of the ordinates. We have a discrete function that can be describe by the Lagrangian interpolation polynomial of degree n

$$f(x) = \sum_{k=0}^{n} l_k(x) f(x_k) + E(x) \tag{5.1}$$

where $l_i(x)$ are the *Lagrangian Coefficient function* given from the expression

$$l_i(x) = C_i[(x - x_0)....(x - x_{i-1})(x - x_{i+1}).....(x - x_n)] \tag{5.2}$$

where C_i is a constant derived by the requirement that the result of replacing f(x) with its Lagrangian approximation be an identity when is an arbitrary polynomial of degree n or less, this condition is equivalent to the request that $l_i(x_j)$ can be described with a so-called *Kronecker delta*, so

$$l_i(x_j) = \delta_{ij} \qquad with \qquad \delta_{ij} = \begin{cases} 0 & if \quad i \neq j \\ 1 & if \quad i = j \end{cases} \tag{5.3}$$

For the request that $l_i(x_i) = 1$ it is possible to determine the constant value

$$C_i = \frac{1}{(x_i - x_0)....(x_i - x_{i-1})(x_i - x_{i+1}).....(x_i - x_n)}. \tag{5.4}$$

In order to put this results in a somewhat more compact forms, we introduce the notation

$$\pi(x) = (x - x_0)(x - x_1).....(x - x_n) \tag{5.5}$$

where its derivative is related to the constant value, indeed

$$\pi'(x_i) = (x_i - x_0).....(x_i - x_n) = \frac{1}{C_i}. \tag{5.6}$$

With this notation the Lagrangian Coefficients are

$$l_i(x) = \frac{\pi(x)}{(x - x_i)\pi'(x_i)} = \frac{(x - x_0)....(x - x_{i-1})(x - x_{i+1}).....(x - x_n)}{(x_i - x_0)....(x_i - x_{i-1})(x_i - x_{i+1}).....(x_i - x_n)}. \tag{5.7}$$

The first expression is useful in theoretical consideration, the second in the actual calculation of the function. With this notation the formula for the error associated to the Lagrangian approximation is

$$E(x) = \pi(x) f[x_0, x_1, ..., x_n, x] = \pi(x) \frac{f^{(n+1)}(\xi)}{(n+1)!}. \tag{5.8}$$

where ξ is some number in the intervall I limited by the largest and smallest of the numbers $x_0, x_1, ..., x_n$ and x. By differentiating 5.1 r times, one obtains the approximation and its error

$$f^{(r)}(x) = \sum_{k=0}^{n} l_k^{(r)}(x) f(x_k) \tag{5.9}$$

$$E^{(r)}(x) = \frac{1}{(n+1)!}\frac{d^r}{dx^r}[\pi(x) f^{(n+1)}(\xi)] = \frac{d^r}{dx^r}\{\pi(x) f[x_0, x_1, ..., x_n, x]\} \tag{5.10}$$

Appling the *Leibnitz' formula* for the rth derivative of a product, the error expression becomes

$$E^{(r)}(x) = \sum_{i=0}^{r} \begin{pmatrix} r \\ i \end{pmatrix} \pi^i(x) \frac{d^{r-i}}{dx^{r-i}} f[x_0, x_1, ..., x_n, x] = \sum_{i=0}^{r} \frac{r!}{(n+r-i+1)!i!} \pi^{(i)}(x) f^{(n+r-i+1)}(\xi_i) \tag{5.11}$$

For the first order of derivative, r=1, we formula 5.11 becomes

$$E'(x) = \pi'(x)\frac{f^{(n+1)}(\xi_1)}{(n+1)!} + \pi(x)\frac{f^{(n+2)}(\xi_0)}{(n+2)!}. \tag{5.12}$$

In particular, for numerical differentation at a tabular point, there follows

$$f'(x_i) = \sum_{k=0}^{n} l'_k(x_i) f(x_k) + \pi'(x_i)\frac{f^{(n+1)}(\xi_1)}{(n+1)!} \tag{5.13}$$

since $\pi(x)$ vanishes when x=x_i, where the factor $\pi'(x_i)$ has the simple form shows in equation 5.6.

In the cases when the abscissas are uniformly spaced, the Lagrangian coefficient functions have been tabulated rather extensively for various values of n. Formulas that involving an odd number of ordinates are most often used, and, if that number is n+1=2m+1, the abscissa are then conventionally renumbered as $x_{-m}, ..., x_{-1}, x_0, x_1, ..., x_m$. If the uniform spacing $x_{k+1} - x_k$ is denoted by h, and if s is measured from the central point, so that

$$x = x_0 + hs \qquad x_i = x_0 + hs_i. \tag{5.14}$$

We present now the differentiating five-point Lagrangian interpolation formula that we have used to derivate our images.

$$f'_{-2} = \frac{1}{12h}(-25f_{-2} + 48f_{-1} - 36f_0 + 16f_1 - 3f_2) + \frac{h^4}{5}f^v(\xi) \tag{5.15}$$

$$f'_{-1} = \frac{1}{12h}(-3f_{-2} - 10f_{-1} + 18f_0 - 6f_1 + f_2) + \frac{h^4}{20}f^v(\xi) \tag{5.16}$$

$$f'_0 = \frac{1}{12h}(f_{-2} - 8f_{-1} + 8f_1 - f_2) + \frac{h^4}{30}f^v(\xi) \tag{5.17}$$

$$f'_1 = \frac{1}{12h}(-f_{-2} + 6f_{-1} - 18f_0 + 10f_1 + 3f_2) + \frac{h^4}{20}f^v(\xi) \tag{5.18}$$

$$f'_2 = \frac{1}{12h}(3f_{-2} - 16f_{-1} + 36f_0 - 48f_1 + 25f_2) + \frac{h^4}{5}f^v(\xi) \tag{5.19}$$

These equations, as clearly explain in the following section, are used to calculate the derivative along the horizontal and vertical direction using h=1, as the distance between

the centers of the pixel is equal to the pixel unit, while to build the derivatives along the diagonal directions we use h= $\sqrt{2}$, as the distance between two pixel centers is equal to the diagonal of the square that has for side the pixel dimension.

5.3.2 Four direction derivatives

The input image is acquired into the algorithm and and second order derivatives are calculated, point to point, along four directions (x, y, and the two bisectors at 45 degrees) producing 4 arrays with the same dimensions of the original image. The derivatives along horizontal and vertical direction are simple and we use our routine of 5-point differentiation that works row to row and column to column producing the first two image. Derivatives along the diagonals presents some difficulties. To apply a rotation to the image introduces changes in the images due to the convolutions applied to rotate the array. So we developes a specific routine to select the pixel that belong to the same diagonal (starting from the left-bottom diagonal to the right-top for the first derivative and from the right-bottom to the left-top for the second one) and apply the derivation routine as it would be a row (or a column). The shorter diagonal must to have at least 5 point to use our differentiation formula. In this way we obtain the other two derivative images. Obviosly, we must consider that the derivatives along horizontal and vertical directions have a different step respect to the other two directions. Indeed in the firsts the distance between the center of one pixel and the following is equal to 1 (pixels unit), while the derivative step along the two diagonal directions is $\sqrt{2}$ (pixels unit), as we moving along the 45 degrees bisector where the distance between the centers of two pixels in succession is really $\sqrt{2}$ pixels.

The four derivatives arrays are considered together, simultaneously, so that the result can be indipendent by the differentiation direction. We choose to use four direction to improve the software capacity to find the sources, especially the nearby objects. The use of only two differentiation directions (i.e. along x and y) implies the introduction of a preferential direction of detection; indeed the slope changes due to closest source, especially due to a weak source, would be detected only along x and y. While the adoption of four directions garantee a source extraction indipendent by the differentiation direction, indeed a single pixel has another eight pixels that surround its position. These points are complitely covered by derivative operation along the four direction (2 pixel for each direction) so a change in the brightness profile trend will be detected in any position it gets on. Why cannot we use more than four derivatives? The informations obtained with other directions would be redundant respect those finded with the first four. In the second crown of pixel around the selected pixel there are some points covered not by the four directions, but this is not really true. Indeed the derivatives images are built analysing the input image point to point, so the skipped pixel in the second crown are covered when we change the point to apply derivatives, as clearly shown in Fig.5.3. Moving the central point, the pixel where is applied the derivative operation, each nearby pixel is covered with all the derivative directions, in this way all the surface of the brightness profile is examined.

The four derivative images are inverted in their signs, this operation has not a basis in a mathematical reasoning but has a pratical rule, indeed is more intuitiveto have a positive source location instead of a negative one (indeed the real sign of second order derivative is negative in corrispondence of a candidate source). Afterwards the images are coadded point to point to obtained a mean image to help us in a rapid control of the results of derivative operation.

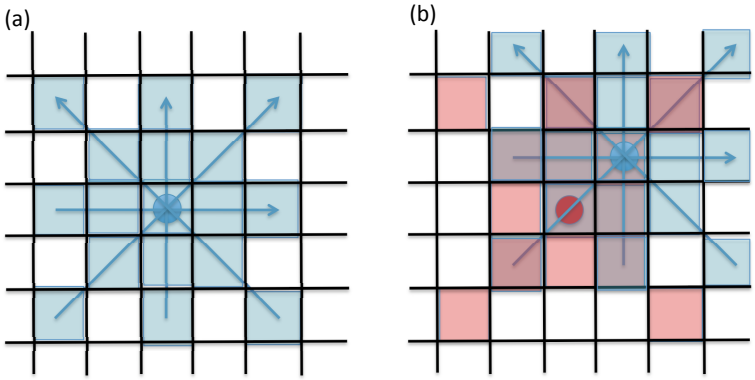


Figure 5.3. These image rapresents the four direction of derivative, and the covering of pixels during this operation. (a) show the region covered in the analysis of a single pixel along the four direction, while (b) show the changing in the pixel covered moving the central pixel, point to point along all the image.

During the analysis of the four derivatives, their signs has a fondamental rule. Indeed the sign gives the information about the changes of the slope of brilliance profile and then about the presence of a source. In our derivative arrays the sign of second derivative are inverted to obtain an image where the sources location is identified by a positive peaks to make simple and intuitive a visual inspection of these results.

From the simultaneous analysis of the four derivative we can characterized several slope changes correlated to structures at different scale, for exampe deep changes in the slope that correspond to high values of derivative can be associated to dense core and compact sources, while a more flat distribution coincides to low value in the derivative and then to bigger structure. Very big scale of emission are completely flattened from the derivative operation that, for its construction its able to exalt the small scale, slightly larger of 5 pixels (the lenght of derivative filter). We chose a threshold on the second derivative values suitable to select only the compact objects. A candidate source location is selected when it passes the threshold verify at all the derivatives directions (i.e. the derivative values is grater than the threshold value in the four derivative images).

We describe now the characteristic of the output images and how these properties are important in the analysis of our star forming regions.

This technique is "*a priori*" independent to the source shape (symmetric or not, psf-like or elliptical) as the trend in the derivative is similar and has the same peculiarities.

Figure 5.2 show the same vertical cut from an image and form its second order derivative. The input image section (see Fig 5.2(a)) presents a typical profile where several peaks (more or less broad) are superimposed over background (spatially variable background) from the interstellar medium belonging to the star forming region. As clearly seen (comparing Fig. 5.2(a) and 5.2(b)), the variability of background is completely removed from derivative operation, this is a fundamental step to distinguish sources from diffused material of native cloud.

Another basic element is the separation of the blended source. With the use of this tecnique, as of the greater part of psf methods, it is easy to distinguish two sources even when they are very close if their brilliance profile peaks are both visible. A problem arises when one of the two peaks is included in the profile of the other source and is visible like a flexa. The closest sources are no detected with the greater part of actual algorirthms, instead derivative change rapidly sign in coincidence with a change of slope due to two close objects, producing a clearly identified double peak in correspondence of the two sources. In Fig.5.2(a) is shown an example of two blended sources where the peak of the second is hidden in the profile of the first one (sources D and E). These objects are clearly separed in derivative image (sources D and E in Fig5.2(b)).
Source B is in particular conditions, indeed it is partially included in the brilliance profile of the source A and besides it's located in a region where the background emission is highly descending, its detection is simple with this algorithm.

The derivative image is clearly interpreted by sight but we want an algorithm able to do this interpretation in an automatic way. The detection software calculates internally the derivative mean value and the associated rms. The derivative mean value is eximated calculating the mode of the pixel values distribution in all the derivative image. This value, as expected, is around zero, indeed constant sky background and the broad emissions are "flattened" around this value. A different address can be do for rms value. Noise in the derivatives image is computed with a *sigma clipping* technique to avoid the contribution of the sources and "negative rings" around them due to the derivative changing sign in proximity of real image local minima. Obviously, this derivative mean value and rms are concerning to the derivative image and there are not related to the background and rms values in the observed image.

Source location propability is identified through this threshold, but it's not sufficient to identified the source peaks but only the regions where they are located. These selected region where automatically analized and only those point that are local maxima of second order derivatives image are choosen as candidate sources. This selection is used to compile a candidate sources list. Figure 5.4 shows a contuors map of the analized star forming region with the overlap of derivative contours.

5.3.3 Input Parameters

As previously said, the use of this software is very simple and it requires not a preliminary analysis of input image. The input command line contain some information commands, a sigma and some keyword. These commands are listed follow:

- The *sigma (σ)* is utilized to select the source over the background emission. This value is used in analogous way to the signal to noise ratio (S/N) within the formula, calculated pixel to pixel,

$$Th(x, y) = Bkg(x, y) + \sigma \times RMS \tag{5.20}$$

  where *Th* is the threshold to be used, in the coordinates *(x,y)* in pixel, to select or discard a candidate source, *Bkg* is the mean value of the derivatives at that point and *RMS* is the derivative image rms value. The σ required in input change a lot with the characteristic of the observed image from few to hundreds units, it is important underlined that this value is absolutely no correlated with the real S/N.

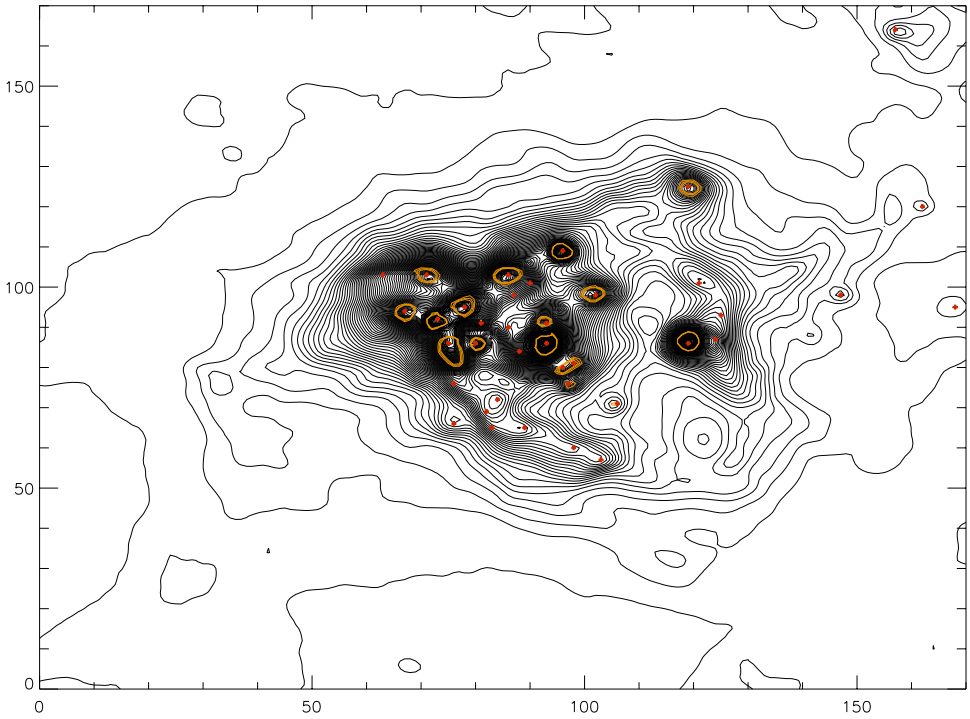


Figure 5.4. Black line rapresent countour map of a star forming region (mol160), with the overlap of derivative contours (yellow line). The red point are the detected sources.

- *General informations:*
  - Input/Output directories
  - Input/Output file names
- *Keywords:*
  - *smoothing*, if the input image is too pixelized can be useful to do a smoothing to simplify the derivative operations.

5.3.4 Source Perimeters

We find, from second derivative images, the perimeter of the sources identifying the first minima located around the detected peaks in the four directions. The first minima characterize the point in that change the slope, better the concavity, of the brightness profile, so it rapresents the point at which the source startsto merge in the plateau emission. this information is used as constrain to the gaussian fit, following described, to obtain a better fit of the sources.

5.3.5 Source detection tests

We have tested our algorithm on real and simulated field at different wavelenghts. We present, in the following paragraphs, the results of the application of our software to an

Hershel simulation and to a Spitzer observation. For the first one we have the truth-table to compare the position of extracted sources to the real coordinates, while we dont know the real position of the objects in the observed image, so to comprehend the entity of our detection results we compare those with the extraction of two of the better data analysis software: Daofind and Mopex.

We want to underline a fundamental characteristic of our method, all the presented results are obtained with a single run of our software. This have a great importance if you want use the algorithm in an automatic way on the data of a survey.

Test on simulated fields

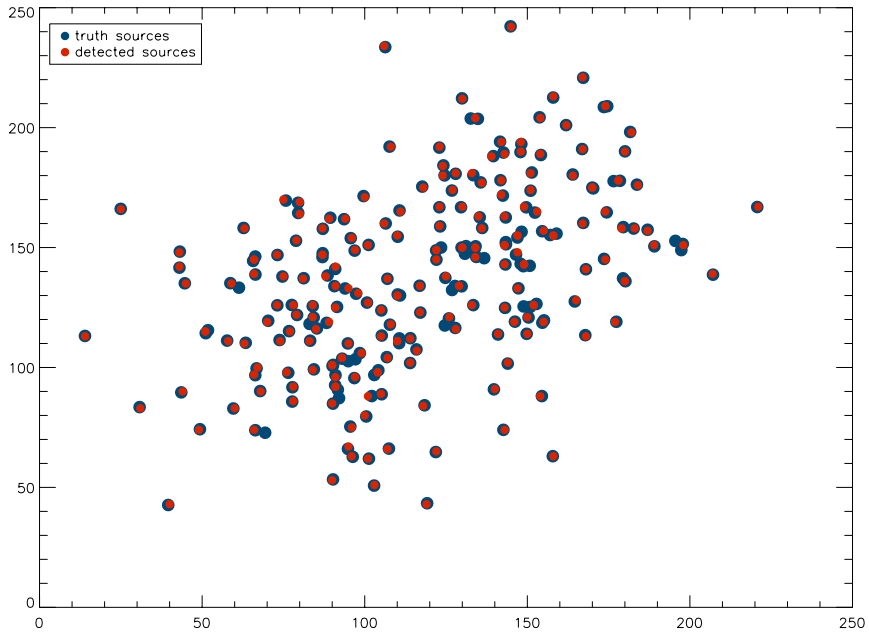
To show the application of this method, we present the results of source extraction for a PACS[3] simulated fields at 70 μm with different source shapes (circular and elliptical gaussian profiles) with and without noise.

Figure 5.5 shows the results of source extraction in the field with and without noise, the number of extracted source is similar in the two cases and is the equivalent to the detection percentage about 93-95%. The greater number of undetected sources are hidden in the brightness profile of a near objects, as it is shown in Fig 5.6 where the central source in fact is the result of a blending between two objects. This percentage of detected source is a very good results, considering the difficulties present in these simulations.

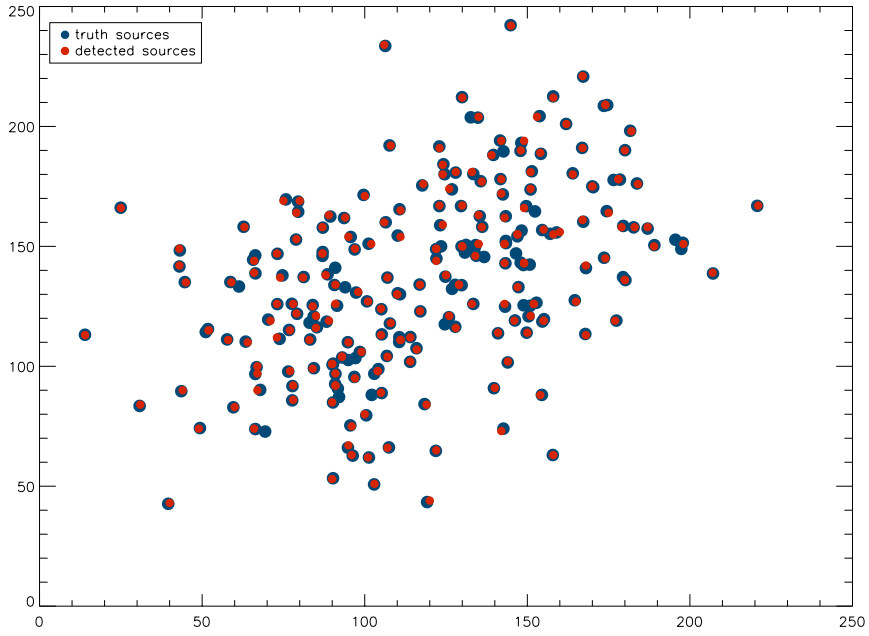
Test on observed field

We present in Fig.5.7 the detection results on a field observed with Spitzer, mol160. This observation at 24 μm, previously introduced, was been the crucial point to decide to find a new data analysis approach to resolve our problems in forecast to the future surveys. As previously said, we have not a table with the source position of this real observation, so we test our algorithm comparing our detection with those obtain with two of the most used data analysis package (DAOFIND and MOPEX). The increase in the source detection is evident from images presented in Figure 5.7. Another important thing is that the new source extracted are not false positive but are real objects as can be underline from a surface analysis of the image flux trend (see Figure 5.8). In the first image (see 5.7(a)) are shown the DAOFIND results only a few number of objects are found. Blended sources are no detected and even those objects that are located on a rapidly variable background. The detection number increase using MOPEX, but the presented results are obtained with multi-run of the software, changing the dimension of the box used to eximate the background and to fit the shape of the source (green crosses with black contour in Figure 5.7(b)). Our results are obtained with a unique run of our algorithm and show an higher number of sources, as can be seen in Figure 5.7(c). Further the blended sources, resolved in a very good way from oun routine, we find objects even in region when the diffuse emission have a variable structure, as the five sources located on the bottom branch of cloud that there are never detected by the other routines but that are clearly visible in the surface plot (objects indicated by the blue arrows in Figure 5.8. This no-detection with the other tasks is due to the choise of background value for the extraction of the source, indeed to detect the brightest sources located on the diffuse emission we must use an high value for background

[3] Photodetector Array Camera & Spectrometer for Herschel



(a) Source detection results for a PACS simulation at 70μm without noise



(b) Source detection results for a PACS simulation at 70μm with noise

Figure 5.5. The detection results for two simulated PACS field. Blue filled circles rapresent the position of real position of the sources while the red ones are the detected objects.

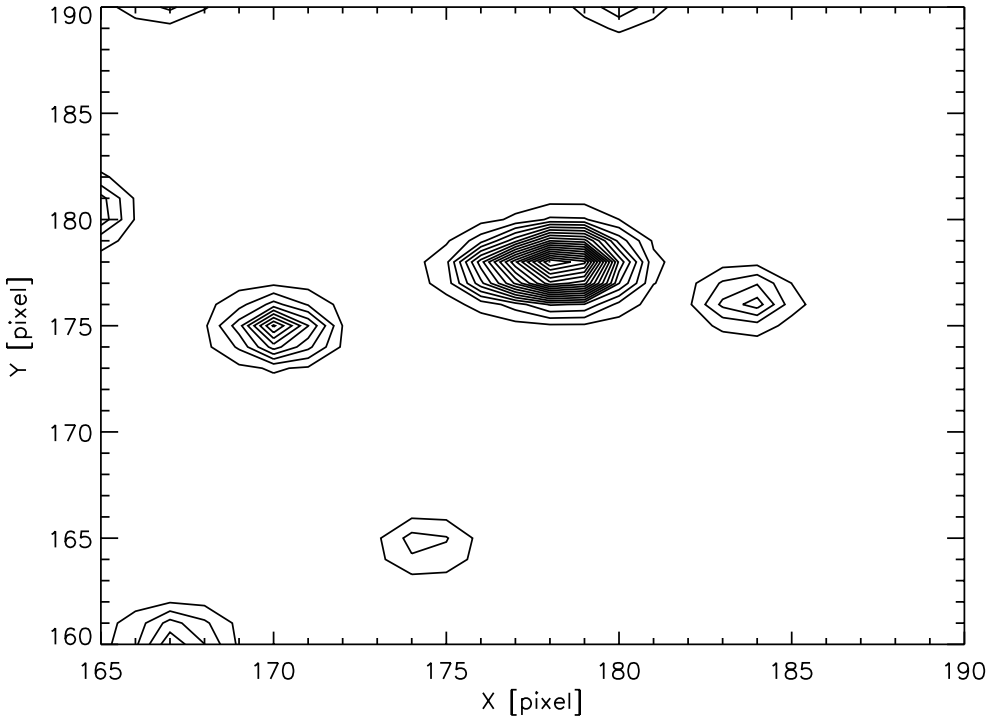


Figure 5.6. Contours maps for two blended sources

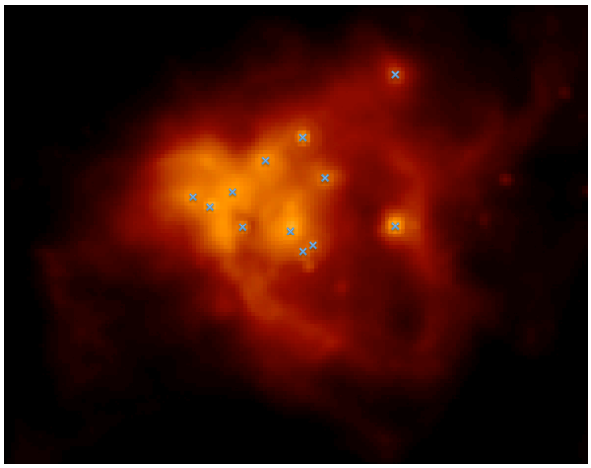
for DAOFIND and a too small box for MOPEX (even this choice leads to an overextimation of the background value).

5.4 Bi-dimensional gaussian photometry

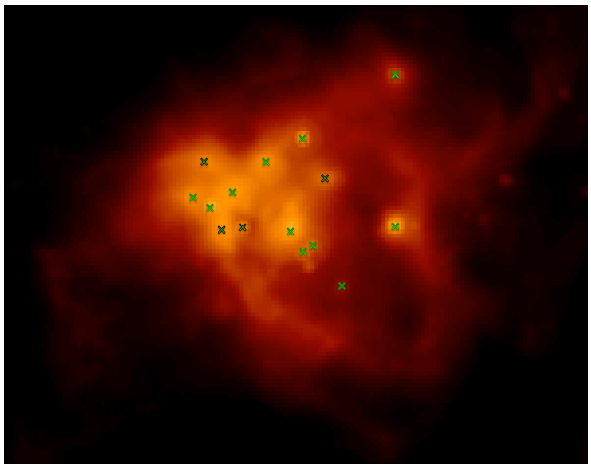
The following description is no detailed like the previous of detection software, because the development of this routine is no mine but it's of Dr. Molinari, so the complete description is no included in this work. So a brief introduction is necessary because it is used for the data analysis described previously.

The sources list produced by detection software is used as input to the photometry routine. In these second analysis all the peak position where fit with a 2D Gaussians+plateau sistem. All would be simple with only isolated objects, but sources tend to appear clustered. The first step of photometry routine is to attempt to group sources by their relative distance, in this way we select several list of objects dividing isolate source from group of two, three, four, etc... each group was fitted in a different way. The routine analizes at first the single objects, for each peak it fit a 2D gaussian profile for the source and a variable value for the plateau due to the diffuse emission of the compact cloud where the core is. The fit of 2D gaussian is made allowing parameter variations as

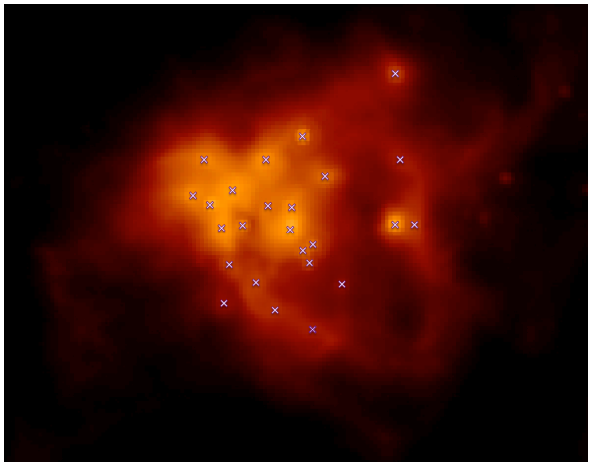
- position
- the ratio between minor and major axis



(a) Daofind results



(b) Mopex results



(c) Our Detection results

Figure 5.7. Comparison between detection results. (a) Blue crosses are the Daofind results. (b) Green crosses are Mopex results. Crosses with black contour are detected with software run with different input condition. (c) White crosses with red contours are our results.

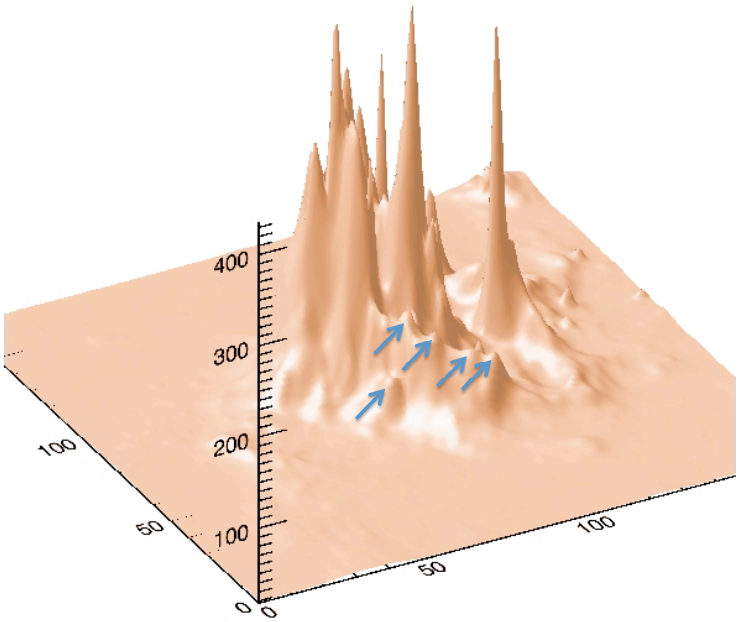


Figure 5.8. A surface plot for mol160 observed at 24 μm. The blue arrows indicate five sources located on a filamentary emission of the cloud. these objects are detected in a simple way with our algorithm.

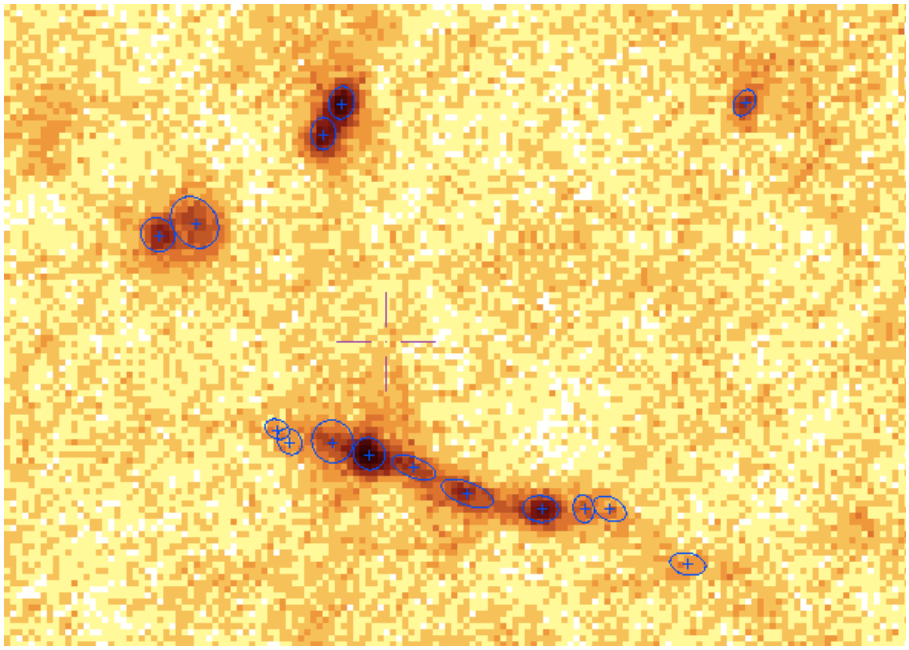


Figure 5.9. Gaussian fit results. The backgronud image is an ATLASGAL image while the blue contours are the fitted gaussians.

- the angle

so, there are a lot of free parameters that make the fit more adaptable to the differents shape of the esamined objects. This algorithm contain and internal subroutine, based on the find software, so it has the same functions of DAOFIND package. This can be used if you have not an input list of sources. We explain now the input command line.

5.4.1 Input Parameters

The use of this routine is simple, as said for the detection one. The input parameters include some information commands, keyword and some limits for the variables that are used for the fitting.

- *General informations:*
  - Input/Output directories
  - Input/Output file names
- *Keywords:*
  - *instrument* there is a pre-impostated list of observative instruments.
  - *smoothing*, if the input image is too pixelized can be useful to do a smoothing to simplify the fitting of sources profile.

  - *Input source list*, this keyword is used to specify if the input file comes to our detection algorithm or to the find routine

- *Variables for the fit:*

  - *margin*, the fit of 2D gaussian has problems with a source profile truncated by the edge of the image, so you can set a margin value to esclude the edges from the fitting.
  - *posback*, this value can set equal to 0 or 1. Posback keyword is used to keep the fitting plateau non-negative (1) or free (0).
  - *possource*, if the input coordinates are fixed the possource must set equal to 0, while if it is set equal to 1 the peak position is optimized around the input position.

- *Threshold* for find routine. If there is not an input source list, this value is used as threshold for internal detection routine.

This command line contains all the information for the use of the program.

5.5 Benchmarking of data analysis softwares for Hershel

Whitin of the Hershel team for data analysis, there was a benchmark between the different algorithms for sources detection and photometry. We have partecipate to this with good result. This test was made in three step with several type of simulated images.

Simulation for the first step was made to test the detection of sources at different distances and with several dimension, they have not particullary phisical properties (as wavelenghts, or psf) and correspond not real telescope observations. In spite of all these problems, we obtain good results in detection and photometry, as shown in Figure 5.10, where the blue point are the simulated source while the red ones are our detection. It's clearly visible how our detection results are very good, about 98%.

A second step of simulations was made to reproduce the observations in the six Hershel bands (three for PACS photometer centered at 75, 110 and 170 μm and three for SPIRE photometer at 250, 350 and 500 μm). In Figures 5.11 we present our results of detection and photometry for the simulated image at 110. Our results are perfect for the first showed wavelenght (110μm) where all the sources have compact dimension. Indeed in this image we detect and fit in a simple way all the compact sources even those much blended. In other wavelenghts, such as at 250 μm, there are several cases where the dimensions of the source are so big that the profile of the objects is complitely flattened by the derivative operation and the source results no detected. Where there is a large objects with compact source place on the top of its, we detect without problems the objects on the big source (diffuse emission), only if this large source has a big flux we find it. The truth-table of this second step of simulation are good and we have made the comparison between the fitted and the simulated fluxes. In Figure 5.12 we present our results for the peak and the integrated flux. As it's clearly shown in these figures both the peak flux and the integrated one are extracted with good value that are similar to the simulated ones. Indeed the point are located on the bisector.

To resolve this noise problem and make realistic simulations, we made new image using PACS and SPIRE simulators to make a pre-analysis of the characteristics and the problems

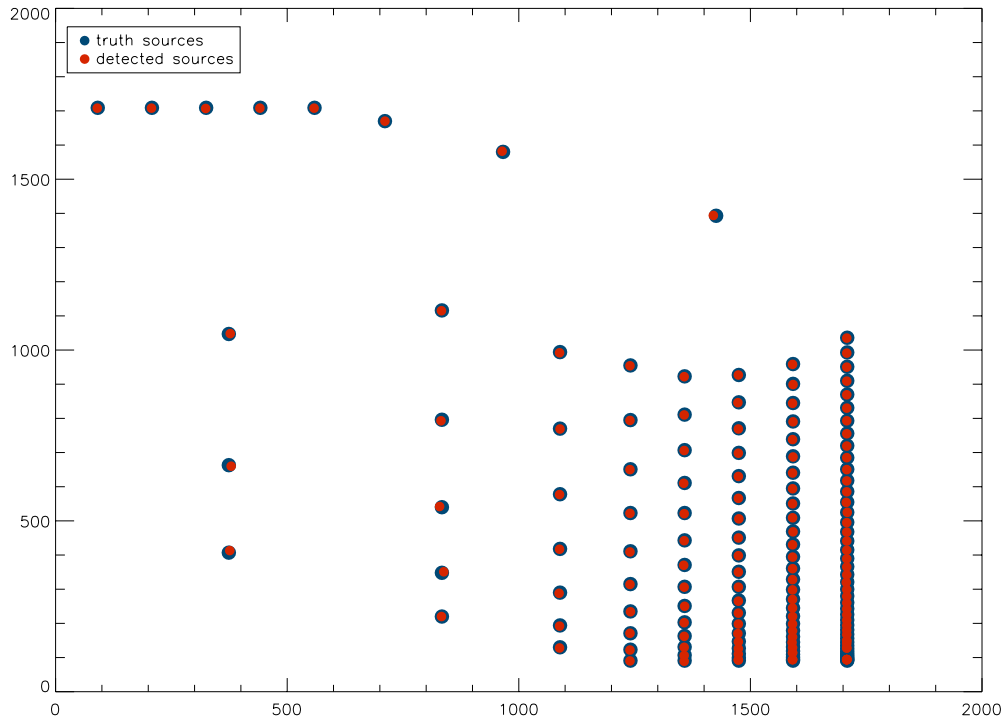
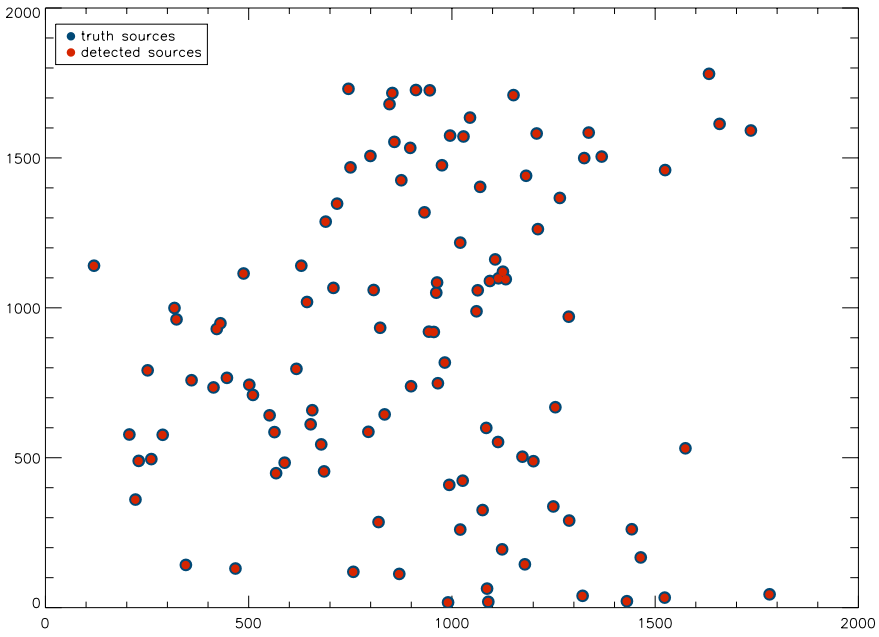
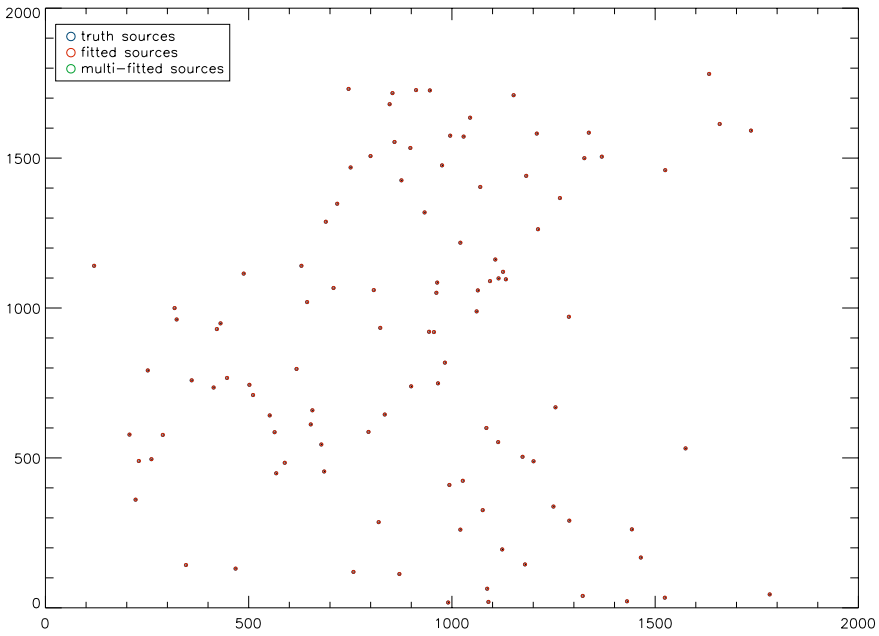


Figure 5.10. Detection results for a simulated field. Filled blue circles are the positions of the simulated sources while the red ones are the detected objects.

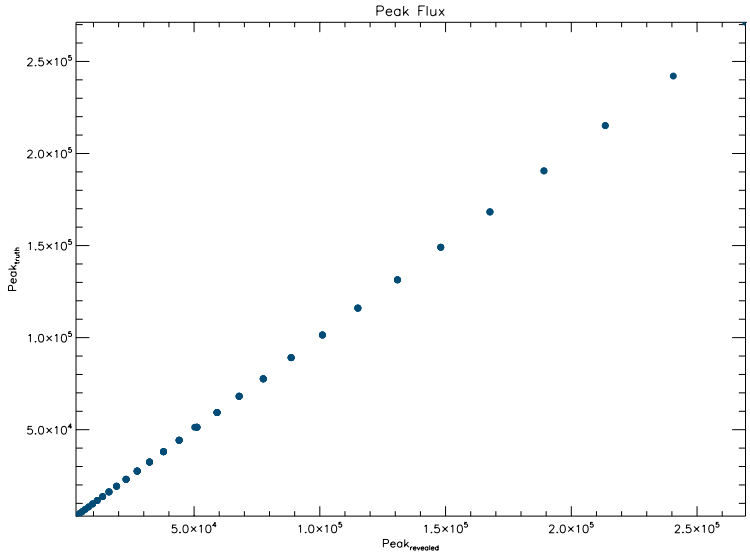


(a) Detection results

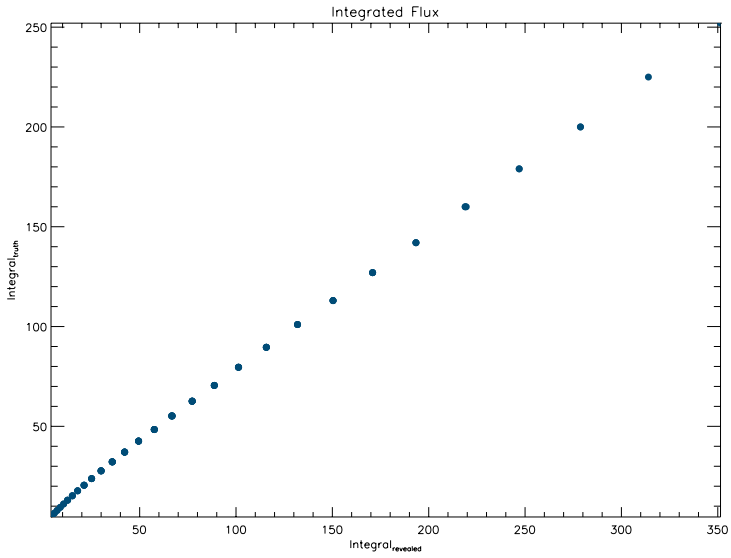


(b) Photometry results

Figure 5.11. Results of our algorithm on a simulated PACS image at 110 μm. a) detection results: filled blue circle are rapresent the position of silulated sources and filled red circles are our detection. b) photometry results: empty blue circle rapresent the shapes of simulated objects while the red ones are our fits.



(a) Peak Fluxes



(b) Integrated Fluxes

Figure 5.12. Photometry results. Values find with our algorithm are plotted on the abscissa while the real simulated values are on the ordinate. Figure a) shows the results for the peak fluxes while on the b) graphic is presented the results for the integrated fluxes.

that will have the real images that soon arrive from Hershel. These simulators, developed by Sibthorpe et al. (2009), are based on the properties of detectors and on their reply to input signal. The produced simulations contain a realistic treatment of the photonic noise and help us to confront the problematics of the real observations.

5.6 Discussion

Our algorithm make a good job in critical condition like those present in a star forming region, where crowding, spatially variable background and no psf-like source are summed to buil a very complicated situation. We show how the actual free distributed data analysis softwares are no able to work very well in these conditions.

As previously said, our aim is not to develop a software able to resolve all the problems that can be find in the astrophysic data analysis, but to build an algorithm that can be overcome our difficulties, in the prospect to use its in the HiGal survey.

The capability to remove the contribution of diffuse emission is fundamental to detect all the objects in a star forming region. This is one of the best results of our detection method that make it in a simple way exploiting the intrinsic properties of derivative operation.

Built source list, a very important step is the capacity to fit source profiles in an adaptable way. Our photometry algorithm fit no-psf profile (much common in star forming regions) and can do it in a crowded group separating the contribution of different sources and, this is one of the fundamental results, separating the contribution of these sources from the emission of diffuse emission (gave by a costant platue contribution).

Another important thing is that "DERBIGA" is able to detect and extract all the tipology of objects present in a region (point-like, elliptical, blended) in a single run of the software, and, found the optimal threshlod, it can work automatically over a big data volume without our intervention.

These were the principle aims that we have wanted when we decide to do a new instrument for data analysis. Obviously this routine works very well in less complicated situation like open clusters or single objects, but for its characteristic is no adapt to analize, for example, the structures of diffuse emissions.

III Part

Conclusion

Chapter 6

Conclusion

6.1 Scientific results

The main results of this work are the following:

- *Near-IR clusters results:*
  - In the J, H, and K_s NIR bands, we have imaged 26 intermediate and high-mass star-forming regions selected from a larger sample of sources spanning a range in luminosities and presumed youth. We have identified the presence of 23 young stellar clusters in 22 fields.
  - The detected clusters have richness indicator values of between ten and several tens of objects and have median radii of 0.7 pc. Compared to clusters around Herbig Ae/Be stars, our clusters seem richer and larger for any given mass for the most massive star in each cluster. Color-color diagrams show that these clusters are young: many sources exhibit colors typical of young pre-MS objects with an intrinsic IR excess originating in warm circumstellar dust. This is confirmed by the analysis of color-magnitude diagrams, where a significant fraction of stars in each cluster are found to be related to the Pre-MS evolutionary tracks, even after conservative de-reddening is applied.
  - We have been unable to perform a direct inversion of stellar luminosities (and colors) into masses and ages; we use a synthetic cluster generator (SCG) model to create statistically-significant cluster simulations for different initial parameters (IMF, SFH, source ages, and their distribution), and compare the synthetic KLFs and HKCFs with the observed (field star-subtracted) ones. For the fraction of clusters for which this comparison selects a clearly defined region of the parameter space, we conclude that star formation in these regions *cannot* be represented by a single burst, but is a process that is spread out in time. Clusters have mean ages of a few 10^6 yrs; the ages of most of the clusters members are spread, within each cluster, between a few 10^5 yrs and a few 10^6 yrs. Together with the independent evidence that the most massive stars in these systems are very young, or not even yet on the ZAMS, this result is difficult to reconcile with any model predicting cluster formation in a crossing time.
  - The cluster radii seem to be inversely proportional to their age, as also confirmed by the comparison of cluster parameters with those typical of Ae/Be

systems, which are smaller and less rich. As suggested by numerical simulations in the literature, dispersal of intra-cluster gas (by, e.g., molecular outflows or radiation fields from massive stars) may lead to the loss of a fraction of cluster stellar population, thus indeed leading to smaller and less rich clusters. Our results seem in line with this prediction.

– The relation between the mass of the most massive star in a cluster and the cluster's richness indicator suggests that a *physical* rather than *statistical* nature of the cluster origin is more likely.

- *SED Results:*

  – We verify that, similarly to the low-mass regime, the L_{bol}-M_{env} diagram is a valuable tool to follow the pre-MS evolution of massive YSOs in their embedded phase. The distribution of massive YSOs in the diagram clearly suggests that MM, IR-P and IR-S objects are the high-mass analogues of Class 0, I and II sources in the classical classification established for low-mass YSOs.

  – Using high-resolution images of Spitzer (GLIMPSE survey for the Mid-IR bands and MIPSGAL survey for Mid/Far-IR ones) we resolve several core associated at each high-mass star forming region. Each region contain one, or more, main sources that have a Far-IR/mm counterpart. The main sources are the more massive and younger objects in the sample fields ($M_\star$ about few tens of $M_\odot$ and age about few 10^3 yr) and a great number of them is in the accelerating accretion phase (equivalent of Class 0 for low-mass).

  – Using a simple toy-model based on the prescriptions from the turbulent core model (McKee & Tan McKee & Tan (2003)), we show that MM sources preceeds IR-P sources in the evolution toward the ZAMS. The distribution of the different classes of objects with respect to these evolutionary tracks suggest that the IR-P stage may correspond to the arrival of the YSO on the ZAMS. The MM objects would represent an earlier stage.

  – The analysis of the age and the evolutive phase of the sources, identified in each region, shows that several evolutive phases coesist in a single cluster.

- *Global Results:*

  – Clustering seems almost ubiquitous in young High-mass star forming regions. The ages of most of the clusters low-mass population members are spread, within each cluster, between a few 10^5 to few 10^6 years. The highest mass objects, on the contrary, appear to be only 1-4$\times10^5$ years old at most (depending on mass).

  – Rather than a rapid burst process, star formation appears to be a relatively long-lasting and continuous event with low-mass stars forming first and the highest-mass object appearing only at relatively later times.

  – a long time for star formation seems to prefer those theoretical models that are based an accretion on o single massive core (Tan & McKee 2003; Krumholz McKee & Klein 2006).

6.2 Software development results

We have developed an algorithm aale to work in star forming region, where differents problematics are present to complicate our job (see chapter 5). The principle aims are:

- *General results:*
  - No preliminary image analysis is required
  - A single run to make detection and photometry of all the tipology of the sources present in an image.
- *Detection results:*
  - Source detection is not affected by the background variations (big-scale variations are flattened by the derivatives)
  - Source detection is not affected by the shape of the sources
  - Second derivatives are very sensitive to slope changes so we can de-blend the closest sources
  - There are not a lot of parameters to be set... Only one: the desidered threshold
- *Photometry results:*
  - Separation of the flux of a source from the contribution of the clump where it is located
  - Fitting of different source profiles (circular gaussian, elliptical gaussian) with an adaptable set of parameters
  - Multi-fitting for crowded regions

6.3 Future: Data Analysys for Hi-GAL

Hi-GAL is a big international project (PI Sergio Molinari) that will cover the inner 120° (in longitude) of the Galactic Plane. We will collect data at 70, 170, 250, 350 and 500 μm simultaneously. Hi-GAL embodies the optimum combination of Herschel wavelength coverage, sensitivity, mapping strategy and speed to deliver, in a single and homogeneous dataset of extraordinary legacy value, the ultimate census, temperature, luminosity, mass and Spectral Energy Distribution of star forming regions and cold ISM structures in all the environments of the Galactic ecosystem, at unprecedented resolutions, and at all scales from massive objects in protoclusters to the full spiral arm.

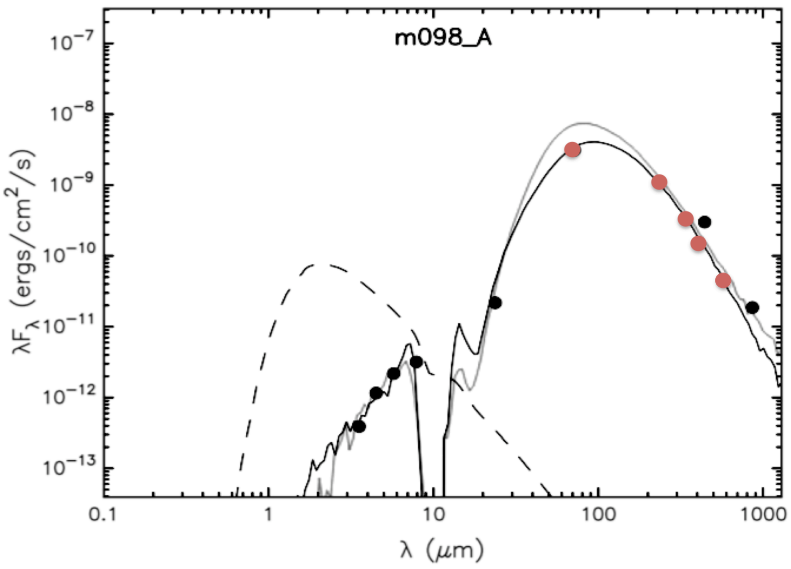

m098_A
10^{-7}
10^{-8}
10^{-9}
10^{-10}
10^{-11}
10^{-12}
10^{-13}
λF_λ (ergs/cm^2/s)
0.1
1
10
100
1000
λ (μm)

Figure 6.1.

Bibliography

Aitken, D. A., Efstathiou, A., McCall, A., & Hough, J. H. 2002, MNRAS,329, 647

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

Bachiller, R. & Tafalla, M. 1999, in The Origin of Stars and Planetary Systems, ed. C. J. Lada & N.D. Kylafis (Dordrecht: Kluwer), 227

Baumgardt, H., & Kroupa, P. 2007, MNRAS 380, 1589

Ballesteros-Paredes, J., Klessen, R. S., Mac Low, M.-M., & Vázquez-Semadeni, E., 2007, Protostars and Planets V, 63

Behrend, R. & Meader, A. 2001, A&A 373, 190

Beltrán, M. T., Cesaroni, R., Neri, R., Codella, C., Furuya, R. S., Testi, L., & Olmi, L. 2004, ApJ, 601, L187

Beltrán, M.T., Brand, J., Cesaroni, R., Fontani, F., Pezzuto, S., Testi, L., Molinari, S. 2006, A&A 447, 221

Benjamin, R. A., Churchwell, E. B., Babler, B., et al. 2003, PASP 115, 953

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

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

Bessel, M. S., 1979, PASP, 91, 377

Bessel, M. S. 1991, AJ, 102, 303

Bessel, M. S., & Brett, J.M. 1988, PASP, 100, 1134

Beuther, H., Schilke, P., Sridharan, T. K., Menten, K. M., Walmsley, C. M., & Wyrowski, F. 2002, A&A, 383, 892

Binney, S. J., Tremaine,S. 1987, Galactic Dynamics, Princeton Univ. Press, Princeton NJ

Bjorkman, J. E. 1997, *in Stellar Atmospheres: Theory and Observations*, ed. J. P. De Greve, R. Blomme, & H. Hensberge (New York: Springer), 239

Bjorkman, J. E. & Wood, K. 2001 ApJ 554,615

Bonnell, I. A., Bate, M.R., & Zinnecher, H. 1998, MNRAS, 298, 93

Bonnell, I. A., & Davies, M.B. 1998, MNRAS, 295, 691

Bonnell, I. A., & Clarke, C. 1999, MNRAS, 309, 461

Bonnell, I. A., & Bate, M.R. 2002, MNRAS, 336, 659

Bonnell, I. A., & Bate, M.R. 2006, MNRAS, 370, 488

Bonnell, I. A., Bate, M.R., Clarke, C., & Pringle, J.E. 2001, MNRAS, 323, 785

Bonnell, I. A., Vine, S.G., & Bate, M.R. 2004, MNRAS, 349, 735

Bonnell, I. A., Bate, M.R., & Vine, S.G. 2003, MNRAS, 343, 413

Brand, J., & Blitz, L. 1993, A&A, 275, 67

Brand, J., Cesaroni, R., Palla, F., Molinari, S. 2001, A&A, 370, 230

Burrows, C. J., et al. 1996, ApJ, 473, 437

Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, ApJ, 345, 245

Cesaroni, R., Neri, R., Olmi, L., Testi, L., Walmsley, C. M., & Hofner, P. 2005, A&A, 434, 1039

Cesaroni, R. 2005, Ap&SS, 295, 5

Cesaroni, R., Galli, D., Lodato, G., Walmsley, C.M., & Zhang, Q. 2006, Protostars & Planet V, *Disks Around Young O-B (Proto)Stars: Observation and Theory*

Chini, R., Hoffmeister, V., Kimeswenger, S., Nielbock, M., Nürnberger, D., Schmidtobreick, L., & Sterzik, M. 2004, Nature, 429, 155

Churchwell, E. 1997, ApJL, 479, L59

Ciolek, G. E., & Königl, A. 1998, ApJ, 504, 257

Clarke, C., & Delgrado-Donate, E. 2003, *Galactic Star Formation Across the Stellar Mass Spectrum*. *ASP Conference Series*, p.287

Cohen, M. 1993, AJ, 105, 1860

Cotera, A., et al. 2001, ApJ, 556, 958

de Wit, W.-J., Testi, L., Palla, F., Zinnecker, H. 2005, A&A, 437, 247

di Francesco, J., Evans, N. J., II, Caselli, P., Myers, P. C., Shirley, Y., Aikawa, Y., & Tafalla, M., 2007, Protostars and Planets V, 17

Dullemond & Natta, A. 2003, A&A, 405, 597

Faustini, F., Molinari, S., Testi, L., & Brand, J. 2009, A&A, 503, 801

Foster, P., & Chevalier, R. A. 1993, ApJ, 416, 303

Hillenbrand, L. A., & Carpenter, J. M. 2000, ApJ, 540, 236

Hoogerwerf, R., de Bruijne, J. H. J. & de Zeeuw, P. T. 2001, A&A, 365, 49

Hartmann, L. 1998, Accretion Processes in Star Formation (Cambridge: Cambridge Univ. Press)

Hartmann, L. W., Jones, B. F., Stauffer, J. R. & Kenion, S. J. 1991, AJ, 101, 1050

Hunt, L. K., Mannucci, F., Testi, L., Migliorini, S., Stanga, R. M., Baffa, C., Lisi, F., Vanzi, L., 1998, AJ, 115, 259

Hillenbrand, L. A. 1997, AJ, 113, 1733

Hildebrand, F. B. 1956 "*Introduction to numerical analysis*", 64

Isobe, , Feigelson, E., Akritas, , Babu , 1990, ApJ 364, 104

Ivezić, Z., Elitzur, M., 1997, MNRAS 287, 799

Klein, R. I., Inutsuka, S., Padoan, P., Tomisaka, & K. 2006, Protostars & Planet V, *Current Advances in the Methodology and Computational Simulation of the Formation of Low-Mass Stars*

Kenyon, S., & Hartmann, L. 1995, ApJS, 101, 117

Kenyon, S. J., & Hartmann, L. 1987, ApJ, 323, 714

Keto, E. 2003, ApJ, 599, 1196

Kim, S.-H.,Martin, P. G., & Hendry, P. D. 1994, ApJ, 422, 164

Koornneef, J. 1983, A&A, 128, 84

Kumar, M.S.N., Keto, E., Clerkin, E. 2006, A& A 449, 1033

Kumar, M.S.N., & Grave, J.M.C. 2008, ASPC 387, 323

Kroupa, P., Tout, C. A., & Gilmore, G. 1993, MNRAS, 262, 545

Krumholz, M. R., McKee, C. F., & Klein, R. I. 2006, ApJ 638, 369

Krumholz, M. R., & Tan, J. C. 2007, ApJ 654, 304

Lada, C., & Adams, F. 1992, ApJ 393, 278

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

Lada, C. J., & Wilking, B. A. 1984, ApJ, 287, 610

Langer, W. D., Velusamy, T., Kuiper, T. B. H., Levin, S., Olsen, E., & Migesen, V., 1995, ApJ, 453, 293

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

Li, Z.-Y. 1998, ApJ, 497, 850

Makovoz, D. & Marleau, F. R. 2005 PASP 117, 1113

McKee C.F. 1999, NATO ASIC Proc.540, 29

McKee C.F., & Holliman, J. H., II 1999, ApJ 522, 313

McKee C.F., & Tan, J. C. 2003, ApJ 585, 850

Molinari, S., Brand, J., Cesaroni, R., & Palla, F. 1996, A&A, 308, 573

Molinari, S., Brand, J., Cesaroni, R., Palla, F., & Palumbo, G. G. C. 1998, A&A, 336, 339

Molinari, S., Brand, J., Cesaroni, R., & Palla, F. 2000, A&A, 355, 617

Molinari, S., Testi, L., Rodrguez, L.F., & Zhang, Q. 2002, ApJ, 570, 758

Molinari, S., Pezzuto, S., Cesaroni, R., Brand, J., Faustini, F., Testi, L. 2008a, A&A, 481, 345

Molinari, S., Faustini, F., Testi, L., Pezzuto, S., Cesaroni, R., Brand, J. 2008b, A&A, 487, 1119

Moraux, E., Lawson, W. A. & Clarke, C. 2007, A&A, 473, 163

Nakano, T. 1989, ApJ, 345, 464

Padgett, D. L., Brandner, W., Stapelfeldt, K. R., Strom, S. E., Terebey, S., & Koerner, D. 1999, AJ, 117, 1490

Palla, F. 1999, in The Origin of Stars and Planetary Systems, ed. C. J. Lada & N.D. Kylafis (NATO ASI Ser. C, 540:Dordrecht: Kluwer), 375

Palla, F., & Stahler, S. W. 1992, ApJ 392, 667

Palla, F., & Stahler, S. W. 1993, ApJ 418, 414

Palla, F., & Stahler, S. W. 1999, ApJ, 525, 722

Preibisch, Th., Ossenkopf, V., Yorke, H. W., Henning, Th. 1993, A&A, 279, 577

Reipurth, B., & Raga, A. C. 1999, in The Origin of Stars and Planetary Systems, ed. C. J. Lada & N.D. Kylafis (Dordrecht: Kluwer), 267

Reipurth, B., Yu, K. C., Heathcote, S., Bally, J., & Rodrìguez, L. F. 2000, AJ, 120, 1449

Rieke, G. H., & Lebofsky, M. J. 1985, ApJ, 288, 618

Saito, M., Kawabe, R., Ishiguro, M.,Miyama, S. M., & Hayashi,M. 1995, ApJ, 453, 384

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

Saraceno P., Andrè P., Ceccarelli C., Griffin M., Molinari S., 1996, A&A 309, 827

Scalo, J. 1998, *The Stellar Initial Mass Function, ASP Conference Proceedings* Vol. 142, p. 201 (eds. G. Gilmore, D. Howell)

Sibthorpe, B., Chanial, P., & Griffin, M. J., 2009, A&A 503, 625

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

Stetson, Peter B. 1987, PASP 99, 191

Strom, K. M., Strom, S. E., Edwards, S., Cabrit, S., & Skrutskie, M. F. 1989, AJ, 97, 111

Tamura, M., Gatley, I.,Waller, W., & Werner, M. W. 1991, ApJ, 374, L25

Tan, J.C., Krumholz, M.R., McKee, C.F. 2005, ApJ 632, 65

Tan, J.C., Krumholz, M.R., McKee, C.F. 2006, ApJ 641, 121

Terebey, S., Shu, F. H., & Cassen, P. 1984, ApJ, 286, 529

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

Testi, L., Palla, F., Prusti, T., Natta, A., & Maltagliati, S. 1997, A&A, 320, 159

Testi, L., Palla, F., & Natta, A., 1998, A&AS, 133, 81

Testi, L., Palla, F., & Natta, A., 1999, A&A 342, 515

Testi, L., Palla, F., & Natta, A., 2001, *From Darkness to Light, ASP Conference Series* Vol. 243, p. 377 (eds. T. Montmerle, Ph. André)

Tan, J. C., 2005, *Cores to Clusters: Star Formation with Next Generation Telescopes Astrophysics and Space Science Library*, Vol. 324, p. 87

Tan, J. C., & McKee C.F. 2002, *Hot Star Workshop III: The earliest Stages of Massive Star Birth. ASP Conference Proceedings*, 267,p. 267

Tan, J. C., & McKee C.F. 2003, IAU Symposium 221, 274P

Vázquez-Semadeni, E., Ryu, D., Passot, T., Gonzáles, R. F., & Gazol, A. 2006, ApJ 643, 245

Vig. S., Testi, L., Walmsley, M., Molinari, S., Carey, S., Noriega-Crespo, A. 2007, A&A 470, 977

Wainscoat, R. J., Cohen, M., Volk, K., Walker, H. J., & Schwartz, D. E. 1992, ApJS, 83, 111

Ward-Thompson, D., Andre, P., Crutcher, R., Johnstone, D., Onishi, T., & Wilson, C. 2007, Protostars and Planets V, 33

Weidner, C., Kroupa, P. 2004, MNRAS 348, 187

Weidner, C., Kroupa, P. 2006, MNRAS 365, 1333

Weidner, C., Kroupa, P., & Maschberger, T. 2008 arXiv0811.3730W

Whittet, D. C. B., Gerakines, P. A., Hough, J. H., & Shenoy, S. S. 2001, ApJ, 547, 872

Whitney, B. A., Kenyon, S. J., & Gómez, M. 1997, ApJ, 485, 703

Whitney, B.A., & Wolff, M.J. 2002 ApJ 574, 205

Whitney, B.A.,Wood, K., Bjorkman, J.E.,Wolff, M.J. 2003 ApJ 591, 1049

Whitney, B.A.,Wood, K., Bjorkman, J.E.,Wolff, M.J. 2003 ApJ 598, 1079

Wood D.O.S., & Churchwell, E. 1989, ApJ 340,265

Wood, K., Wolff, M. J., Bjorkman, J. E., & Whitney, B. 2002, ApJ, 564, 887

Yorke, H. W. 2002, *Hot Star Workshop III: The earliest Stages of Massive Star Birth. ASP Conference Proceedings*, 267,p. 165

Zhang, Q., Hunter, T. R., Brand, J., Sridharan, T. K., Molinari, S., Kramer, M. A., Cesaroni, R. 2001, ApJ 552, L167

Zhang, Q., Hunter, T. R., Brand, J., Sridharan, T. K., Cesaroni, R., Molinari, S., Wang, J., Kramer, M. A. 2005, ApJ 625, 864

Zinnecker, H. & Yorke, H. W. 2007 ARA&A 45, 481