Astronomy
&
Astrophysics

Evolving structures of star-forming clusters⋆

S. Schmeja and R. S. Klessen

Astrophysikalisches Institut Potsdam, An der Sternwarte 16, 14482 Potsdam, Germany
e-mail: sschmeja@aip.de;rklessen@aip.de

ABSTRACT

Context. Understanding the formation and evolution of young star clusters requires quantitative statistical measures of their structure.
Aims. We investigate the structures of observed and modelled star-forming clusters. By considering the different evolutionary classes in the observations and the temporal evolution in models of gravoturbulent fragmentation, we study the temporal evolution of the cluster structures.
Methods. We apply different statistical methods, in particular the normalised mean correlation length and the minimum spanning tree technique. We refine the normalisation of the clustering parameters by defining the area using the normalised convex hull of the objects and investigate the effect of two-dimensional projection of three-dimensional clusters. We introduce a new measure ξ for the elongation of a cluster. It is defined as the ratio of the cluster radius determined by an enclosing circle to the cluster radius derived from the normalised convex hull.
Results. The mean separation of young stars increases with the evolutionary class, reflecting the expansion of the cluster. The clustering parameters of the model clusters correspond in many cases well to those from observed ones, especially when the ξ values are similar. No correlation of the clustering parameters with the turbulent environment of the molecular cloud is found, indicating that possible influences of the environment on the clustering behaviour are quickly smoothed out by the stellar velocity dispersion. The temporal evolution of the clustering parameters shows that the star cluster builds up from several subclusters and evolves to a more centrally concentrated cluster, while the cluster expands slower than new stars are formed.

1. Introduction

Almost all stars form in clusters. Embedded clusters contain various types of young stars, making them ideally suited to study the early stages of star formation as they provide a large and genetically homogeneous sample (see Lada & Lada 2003, for a review). Understanding the formation and evolution of young stellar clusters requires quantitative statistical measures of their structure, which may give important clues to the formation process. While some clusters are centrally concentrated with a smooth radial density gradient, others show filaments and signs of fractal subclustering. If and how different structures are connected to the environmental conditions of the molecular clouds and how they depend on the evolutionary stage of the cluster is not yet clear.

Different methods have been used to describe the clustering properties of star clusters, e.g., the mean surface density of companions (Larson 1995) or spanning trees. Cartwright & Whitworth (2004, hereafter CW04) presented a review of various statistical methods for analysing the structures of star clusters and a detailed investigation of both observed and artificially created clusters. We apply the methods discussed there to

clusters created by numerical simulations of gravoturbulent star formation and investigate the clustering behaviour with time. We extend the analysis of observed clusters by considering different evolutionary classes.

Four classes of young stellar objects (YSOs) are distinguished according to the properties of their spectral energy distributions (SEDs) (e.g. André et al. 2000): Class 0 sources are deeply embedded protostars with a large sub-mm to bolometric luminosity ratio ($L_{smm}/L_{bol} > 0.005$). The Class 0 stage is the main accretion phase and lasts only a few 10^4 yr. Class 1 objects are relatively evolved protostars, which are surrounded by an accretion disc and a circumstellar envelope. Pre-main-sequence stars in Class 2 and 3 correspond to classical and weak line T Tauri stars, respectively. They are characterised by a circumstellar disc (optically thick in Class 2, optically thin in Class 3) and the lack of a dense circumstellar envelope. These four classes are usually interpreted as an evolutionary sequence from Class 0 to 3. The progenitors of these forming stars are prestellar cores (starless cores, prestellar condensations). These are gravitationally bound, dense molecular cloud cores with typical stellar masses that may already be in a state of collapse, but have not formed a central protostellar object yet.

Section 2 explains the statistical methods used. In Sects. 3 and 4 we describe the observations and the models,

---

⋆ Appendices are only available in electronic form at
http://www.edpsciences.org

respectively, and in Sect. 5 the application of the methods to the data. The results are analysed and discussed in Sect. 6, while we present our conclusions in Sect. 7. The Appendices give the details about the normalisation of the clustering parameters.

2. Statistical methods

2.1. Normalised correlation length

A wide range of statistical methods has been developed to analyse the structure of star clusters (see CW04 for a review). A simple approach is to study the distribution of source separations, as it has been done e.g. by Kaas et al. (2004) for the Serpens cloud core. Larson (1995), extending the analysis by Gomez et al. (1993), introduced the mean surface density of companions $\Sigma(\theta)$, a tool that since then has often been used to study star-forming clusters (e.g. Bate et al. 1998; Gladwin et al. 1999; Klessen & Kroupa 2001). The mean surface density of companions (MSDC) specifies the average number of neighbours per square degree on the sky at an angular separation θ for each cluster star. Knowing the distance to the cluster, θ can be converted to an absolute distance r to determine $\Sigma(r)$. The physical interpretation of the MSDC can be difficult, and CW04 have shown that the normalised correlation length is a better indicator for the clustering behaviour. The normalised correlation length $\bar{s}$ is the mean separation s between stars in the cluster, normalised by dividing by the radius of the cluster, R_{cluster}. This radius is defined via the normalised convex hull of the objects (see Appendix A). The $\bar{s}$ values are independent of the number of stars in the cluster (CW04).

2.2. Minimum spanning trees

The minimum spanning tree (MST), a construct from graph theory, is the unique set of straight lines ("edges") connecting a given set of points ("vortices") without closed loops, such that the sum of the edge lengths is a minimum (Kruskal 1956; Prim 1957; Gower & Ross 1969). In astrophysics, minimum spanning trees have so far mainly been used to analyse the structure of galaxy clusters (e.g. Barrow et al. 1985; Adami & Mazure 1999; Doroshkevich et al. 2004). From the MST the normalised mean edge length $\bar{m}$ is derived. Unlike the mean separation length s, the mean edge length m depends on the number of stars in the cluster, therefore it has to be normalised by the factor $(A/n)^{1/2}$ (Marcelpoil 1993), where n is the total number of stars and A the two-dimensional area of the cluster. In the three-dimensional case the normalisation factor is $(V/n)^{1/3}$, where V is the volume of the cluster. Area and volume are defined by the normalised convex hull of the objects (see Appendix A). The normalisation factors are discussed in detail in Appendix B.

An additional reducing operation, called *separating*, can be used to isolate subclusters (Barrow et al. 1985). Separating means removing all edges of the MST whose lengths exceed a certain limit. When removing edges from a MST, each remaining subgraph is again a MST of its vortices (Robins et al. 2000).

2.3. The combined measure Q

The values $\bar{s}$ and $\bar{m}$, on their own, can quantify, but cannot distinguish between, a smooth large-scale radial density gradient and multiscale fractal subclustering. Dussert et al. (1986) combine the mean edge length m of a MST and its standard deviation σ_{m} and use the (m, σ_{m})-plane to separate different degrees of order in various systems. However, CW04 show that this is not sufficient to differentiate between a smooth large-scale radial density gradient and fractal subclustering. Therefore, CW04 introduced the parameter $Q = \bar{m}/\bar{s}$, which can provide this distinction. Large Q values ($Q \geq 0.8$) indicate centrally concentrated clusters having a volume density $n \propto r^{-\alpha}$, where Q increases with increasing α (i.e. with increasing degree of central concentration). Small Q values ($Q \leq 0.8$) describe clusters with fractal substructure, where Q decreases with increasing degree of subclustering.

2.4. The elongation of a cluster

We also investigate the elongation of the clusters. We define the elongation ξ of a cluster as the ratio of the cluster radius defined by an enclosing circle to the cluster radius derived from the normalised convex hull of the objects. A value of $\xi \approx 1$ describes a spherical cluster, while a value of $\xi \approx 3$ corresponds to an elongated elliptical cluster with an axis ratio of $a/b \approx 10$. See Appendix C for details.

3. Observations

Our observational data are based on the sample of YSOs in embedded clusters compiled from various sources as discussed in detail by Schmeja et al. (2005, hereafter SKF05). However, in the current analysis only the regions ρ Ophiuchi, Taurus and Serpens will be studied in detail. These are clusters, where sufficient information on the evolutionary classes as well as on the positions is given in the literature. IC 348 and Chamaeleon I are used for determining additional clustering parameters. The data of ρ Ophiuchi are taken from Bontemps et al. (2001) (YSOs) and Stanke et al. (2005) (prestellar cores), the Serpens data are from Kaas et al. (2004) (Class 1/2), Hurt & Barsony (1996), and Froebrich (2005) (Class 0), the Taurus data are taken from Hartmann (2002) (YSOs) and Lee & Myers (1999) (prestellar cores). The data of IC 348 are taken from Luhman et al. (2003) and those of Chamaeleon I from Cámbresy et al. (1998). For further details on the compilation of the original sample see SKF05. The numbers of objects are given in Table 1, the positions of all YSOs are plotted in the upper panel of Fig. 1. The adopted distances used to determine the absolute values of s and m are 140 pc for ρ Oph (Bontemps et al. 2001), 260 pc for Serpens (Kaas et al. 2004), 140 pc for Taurus (Hartmann 2002), 315 pc for IC 348 (Luhman et al. 2003), and 150 pc for Cha I (Haikala et al. 2005).

4. The models

We perform numerical simulations of the fragmentation and collapse of turbulent, self-gravitating gas clouds and the

Table 1. Clustering measures and numbers for the observed star-forming regions for prestellar cores (p), Class 1, 2, 3 sources, and all YSOs. (Note that the total number of YSOs is larger than the sum of objects in the individual classes, because it contains Class 0 sources and objects with unclear classification not considered in the analysis of the individual classes.) For details see the discussion in Sect. 6.1.

Region	Class	n	s [']	s [pc]	m [']	m [pc]	$\bar{s}$	$\bar{m}$	Q	$\bar{s}^*$	$\bar{m}^*$	Q^*	ξ
ρ Ophiuchi	p	98	20.45	0.83	2.71	0.11	0.67	0.55	0.82	0.49	0.40	0.81	1.36
	1	15	11.82	0.48	3.15	0.13	1.13	0.68	0.60	0.77	0.43	0.56	1.47
	2	111	15.86	0.65	2.25	0.09	0.69	0.58	0.83	0.54	0.63	1.15	1.27
	3	77	16.25	0.66	2.53	0.10	0.82	0.63	0.76	0.70	0.66	0.93	1.17
	all	205	15.63	0.64	1.67	0.07	0.67	0.57	0.85	0.53	0.45	0.85	1.28
Serpens	1	19	3.43	0.26	1.12	0.09	0.99	0.82	0.83	0.27	0.22	0.79	3.61
	2	18	7.85	0.59	1.74	0.13	0.74	0.61	0.82	0.52	0.42	0.80	1.42
	all	80	5.98	0.45	1.06	0.08	0.60	0.54	0.90	0.37	0.32	0.89	1.63
Taurus	p	52	281.0	11.47	45.88	1.87	0.77	0.48	0.62	0.59	0.36	0.61	1.29
	1	25	265.6	10.84	73.18	2.98	0.83	0.60	0.73	0.59	0.41	0.70	1.39
	2	108	333.5	13.62	34.30	1.40	0.69	0.41	0.59	0.43	0.54	1.26	1.61
	3	72	414.4	16.96	52.99	2.16	0.84	0.51	0.60	0.77	0.63	0.82	1.09
	all	197	356.8	14.58	25.98	1.06	0.67	0.39	0.58	0.46	0.27	0.58	1.45
IC 348	all	288	6.86	0.63	0.66	0.06	0.56	0.52	0.93	0.49	0.45	0.92	1.16
Chamaeleon I	all	180	41.51	1.81	3.74	0.16	0.55	0.37	0.68	0.50	0.34	0.68	1.32

resulting formation and evolution of protostars as described in Schmeja & Klessen (2004, hereafter SK04). We use a code based on smoothed particle hydrodynamics (SPH; Monaghan 1992) in order to resolve large density contrasts and to follow the evolution over a long timescale. The code includes periodic boundary conditions (Klessen 1997) and sink particles (Bate et al. 1995) that replace high-density cores while keeping track of mass and linear and angular momentum. We determine the resolution limit of our SPH calculations using the Bate & Burkert (1997) criterion. This is sufficient for the highly non-linear density fluctuations created by supersonic turbulence as confirmed by convergence studies with up to 10^7 SPH particles (Jappsen et al. 2005; Li et al. 2005).

Our simulations consist of two globally unstable models that contract from Gaussian initial conditions without turbulence and of 22 models where turbulence is maintained with constant rms Mach numbers $\mathcal{M}$, in the range $0.1 \leq \mathcal{M} \leq 10$. We distinguish between turbulence that carries its energy mostly on large scales, at wavenumbers $1 \leq k \leq 2$, on intermediate scales, i.e. $3 \leq k \leq 4$, and on small scales with $7 \leq k \leq 8$. The naming of the models, G1 and G2 for the Gaussian runs, and M$\mathcal{M}$kk (with rms Mach number $\mathcal{M}$ and wavenumber k) for the turbulent models, follows SK04. Details of the individual models are given in their Table 1.

The dynamical behaviour of isothermal self-gravitating gas is scale free and depends only on the ratio α between internal energy and potential energy: $\alpha = E_{\text{int}}/|E_{\text{pot}}|$. Since we are only interested in the positions of the young stars, and since all the clustering parameters are normalised by the cluster radius, the physical scaling is irrelevant for the present study.

The YSO classes are determined as follows (see SKF05 and Froebrich et al. 2005 for a detailed discussion): The beginning of Class 0 is identified with the formation of the first hydrostatic core, when the central object has a mass of about 0.01 $M_\odot$ (Larson 2003). The transition from Class 0 to Class 1 is reached when the envelope mass is equal to the mass of the central protostar (André et al. 2000). We determine the

transition from Class 1 to Class 2 when the optical depth of the remaining envelope becomes unity at 2.2 μm (K-band). The end of Class 0 and Class 1 corresponds to a mass of about 0.43 and 0.85 times the final mass, respectively (SKF05). Lacking a feasible criterion to distinguish Class 2 from Class 3 objects, we consider both classes combined. Prestellar cores are identified by a clump-finding algorithm described in Klessen & Burkert (2000; see also SKF05).

We only consider models with a numerical resolution of at least 200 000 particles. Furthermore, in order to get reasonable numbers of protostars for the statistics, we select only those models where more than 37 protostars with $M_{\text{end}} \geq 0.1 \, M_\odot$ (roughly corresponds to the detection limits of the observations) are formed. This reduces our set of models to 16. Again, see Table 1 of SK04 for further details.

5. Application to the data

We construct the MST following Prim's (1957) algorithm, as also described by Gower & Ross (1969). For both the observations and the models, the parameters $\bar{s}$, $\bar{m}$, Q, and ξ are computed for the different evolutionary classes independently (provided, there is a sufficiently large number of YSOs of that class) as well as for the entire cluster. In no region is the number of Class 0 sources large enough to be included. In addition, these parameters are determined at frequent timesteps of the simulations to obtain the temporal evolution of the parameters. In the case of $\bar{m}$, this requires to construct the MST anew at every chosen timestep. As an example, Fig. 2 shows the MST of one model at different evolutionary stages. While the parameters $\bar{s}$, $\bar{m}$ and Q are calculated using the normalisation factor given above, $\bar{s}^*$, $\bar{m}^*$ and Q^* are determined with the normalisation factor and cluster radius of CW04 (see also the discussion in the Appendix).

As shown by CW04, the effect of binary stars on the clustering parameters is not negligible. Since binaries create very short edge lengths, a large fraction of binaries will significantly
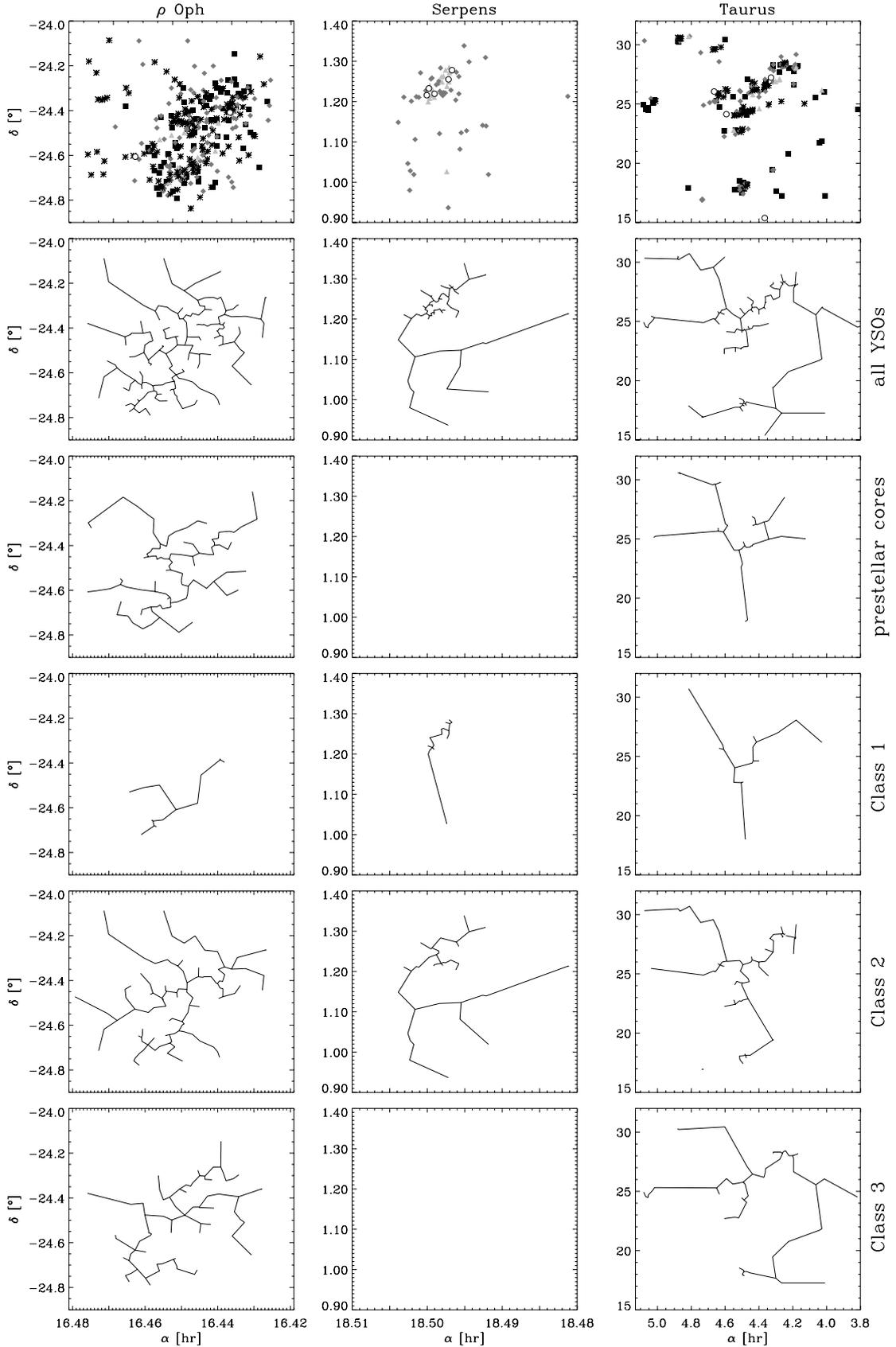
Fig. 1. *Upper panel*: observational data for the star-forming clusters ρ Ophiuchi, Serpens, and Taurus. Circles: Class 0, triangles: Class 1, diamonds: Class 2, squares: Class 3, asterisks: prestellar cores. The sources of the observational data are given in the text. *Other panels*: the MSTs of the same regions for all YSOs (second row), prestellar cores (third row), and Class 1, 2, 3 objects (fourth, fifth, and sixth row). For Serpens, no data on prestellar cores and Class 3 objects are available.
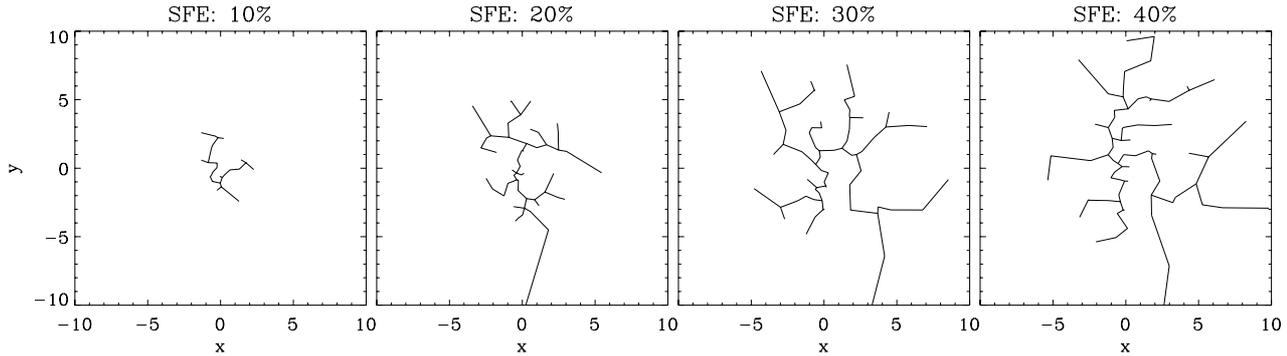
Fig. 2. The 2D minimum spanning tree for the YSOs of model M6k4a projected into the xy-plane at a star formation efficiency of 10%, 20%, 30% and 40% (*from left to right*).

reduce the mean edge length $\bar{m}$ (and as a consequence also change Q). As the binaries are not part of the clustering regime, their influence on the clustering parameters has to be minimized. While for most of the clusters it is not relevant, Taurus is known to have a large binary population. Thus we removed the known binary companions from our sample of YSOs in Taurus.

We calculate the clustering parameters $\bar{s}$, $\bar{m}$, Q, and ξ for the model cluster in three dimensions, and, in order to compare them with the apparent two-dimensional observational data, projected into the xy-, xz-, and yz-plane. The positions of the stars are corrected for the periodic boundary conditions. Only objects inside a box of ten times the side length of the original computational box are considered, if there are any objects outside this volume, we assume that they have left the cluster and are not relevant for the clustering process.

In the beginning stages of the cluster as a whole, and of the individual evolutionary classes, the number of objects is rather low. In such cases the analysis may not be statistically significant and has to be taken with a grain of salt.

6. Discussion

6.1. Observations

Table 1 lists the clustering measures introduced in Sect. 2 for the observations. Columns 4 to 7 list the (non-normalised) mean separation of objects and mean MST edge lengths in arcminutes and parsec. Columns 8 to 10 give the parameters $\bar{s}$, $\bar{m}$ and Q, while Cols. 11 to 13 list the same values calculated using the normalisation of CW04 ($\bar{s}^*$, $\bar{m}^*$, Q^*). The last column gives the elongation ξ. The values of $\bar{s}^*$, $\bar{m}^*$ and Q^* agree with the values given by CW04, except for $\bar{s}^*$ and $\bar{m}^*$ of Taurus (interestingly, however, our Q^* is the same as theirs). We attribute small differences to the slightly different underlying samples, and the discrepancy for Taurus to the different treatment of binaries. Due to the different definition of the radius/area of the cluster, $\bar{s}^*$ and $\bar{m}^*$ differ significantly from $\bar{s}$ and $\bar{m}$, while Q and Q^* are roughly the same, in particular for large samples (see the discussion in Appendix B). Taurus and Chamaeleon I have substructure, while ρ Oph and IC 348 are centrally concentrated clusters (see the discussion in CW04). Serpens (not discussed by CW04) has $Q = 0.90$, corresponding to a central

concentration with a radial density exponent $\alpha \approx 1.9$, similar to IC 348.

The linear distances s and m are in the same range for ρ Oph, Serpens, and IC 348, significantly larger for Chamaeleon I, and about an order of magnitude larger in the case of Taurus, confirming the notion that Taurus represents a somewhat less clustered mode of star formation. However, when considering only the central part of the Taurus region, the values decrease significantly to $s = 3.05$ pc and $m = 0.49$ pc. The latter is comparable to the 0.3 pc estimated as the average distance to the nearest stellar neighbour in the central region of Taurus by Gomez et al. (1993) and Hartmann (2002).

The mean separation s increases with the evolutionary class for all three clusters investigated in this regard, reflecting the expansion of the cluster, which can also be seen in Fig. 1. While Class 0 protostars are formed in the high-density central regions, more evolved YSOs already had time to move to more remote regions (see also Kaas et al. 2004). Note that the prestellar cores do not fit into this sequence, as they are distributed over an area roughly as large as the entire cluster. Thus we speculate that not all objects identified as prestellar cores will eventually form stars. Only in the central parts of the cluster may the density be high enough to make the cores collapse. This is consistent with the findings of Vázquez-Semadeni et al. (2005) that a significant number of "failed cores" should exist, which may redisperse and which may correspond to the observed starless cores.

The elongation of the clusters ranges from $\xi = 1.16$ for the almost perfectly spherical cluster IC 348 to $\xi = 1.63$ for Serpens. The ξ value can differ significantly for subclusters, e.g. the filamentary central part of the Taurus region has an elongation of $\xi = 1.83$.

6.2. Models

We compute the clustering parameters $\bar{s}$, $\bar{m}$, Q, and ξ for all models and compare them with each other as well as with those from the observations. Figure 3 shows the clustering parameters for all models, sorted by the Mach number at a star formation efficiency (SFE) of 10% and 40%. The clustering parameters of the observed clusters (from Table 1) are shown as horizontal lines at their Mach number (taken from SKF05). The $\bar{s}$ and $\bar{m}$ values of the models are in general significantly larger
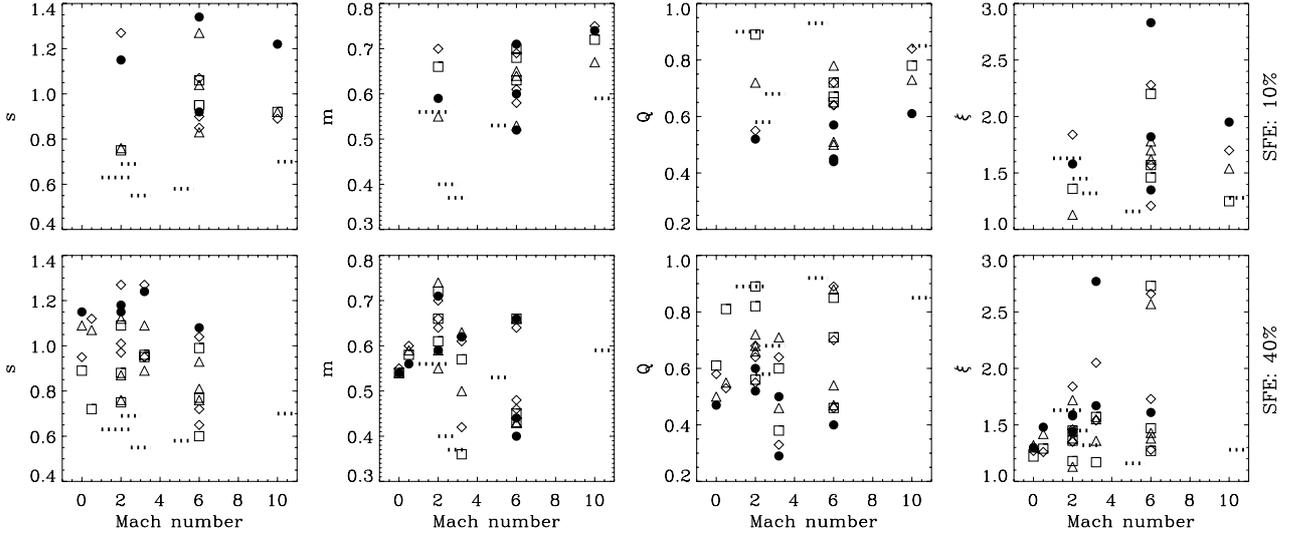
Fig. 3. The clustering parameters $\bar{s}$, $\bar{m}$, Q, and ξ for all models at a SFE of 10% (upper panel) and 40% (lower panel), plotted versus the Mach number. Shown are the values for the projection into the xy-, xz-, and yz-plane (diamonds, triangles and squares, respectively), and for the 3D analysis (filled circles). The horizontal lines show the corresponding values from the observational data for Serpens, Taurus, Chamaeleon I, IC 348, and ρ Oph (*from left to right*).
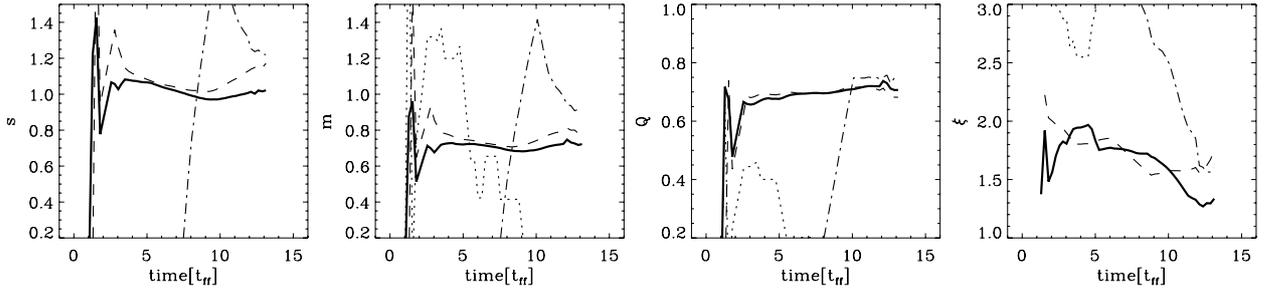

Fig. 4. The temporal evolution of the 3D clustering parameters $\bar{s}$, $\bar{m}$, Q, and ξ for one model (M6k4a), dotted: Class 0, dashed: Class 1, dash-dotted: Class 2/3, solid line: entire cluster.

than those from the observations. A particularly large discrepancy is noted when the model cluster is strongly elongated. A large ξ value reduces the normalisation factor and increases $\bar{s}$ and $\bar{m}$. While the observed clusters show rather moderate elongations in the range $1.1 < \xi < 1.6$, in some models the stars form in a single filament with an elongation $\xi > 3$. Model clusters with elongations in the range of the observations show also a good agreement in the other parameters. For example, the fairly spherical cluster of model M6k4a (shown in Fig. 2) has (at 40% SFE and projected into the xy-plane) an elongation of $\xi = 1.28$ and the clustering parameters $\bar{s} = 0.72$, $\bar{m} = 0.64$, and $Q = 0.89$. These values are almost identical to those of ρ Ophiuchi ($\xi = 1.28$, $\bar{s} = 0.70$, $\bar{m} = 0.59$, $Q = 0.85$). For the other projections the values differ by less than 8%. Most of the Q values, which are independent of the area, lie in the same range as those from the observations. We find no correlation of the clustering parameters with the properties of the turbulent driving (Mach number or wave number k) of the models. Neither do the values from the observations show any correlation with the Mach number. Thus we conclude that if there is any systematic influence of the turbulent environment on the

clustering behaviour, it is only existent in the earliest phase of cluster formation, before it is smoothed out by the motions of the individual protostars (see also Bate et al. 1998).

Figure 2 shows the MST of the model cluster M6k4a at different evolutionary stages and reveals the expansion of the cluster. We analyse the temporal evolution of the clustering parameters: $\bar{s}(t)$, $\bar{m}(t)$, $Q(t)$, $\xi(t)$ in all models. As an example, Fig. 4 shows this sequence for the said model M6k4a. The general behaviour is the same for all models: $\bar{s}$ and $\bar{m}$ decline slightly with time, while Q increases slowly or stays at a roughly constant value. This evolution is shown by the entire cluster as well as by the individual classes, although in later stages the latter values might fall to zero as the number of objects in a particular class becomes zero. Decreasing $\bar{s}$ and $\bar{m}$ values indicate that star formation sets in in different, rather dispersed regions of the cloud. The cluster becomes denser as more and more gas is turned into protostars. New stars are formed faster than the cluster expands. According to the Q values, the cluster evolves slowly from fractal subclustering to a more centrally concentrated cluster although in no model does Q_{2D} rise significantly above the "divding value" of 0.8. This again shows that the
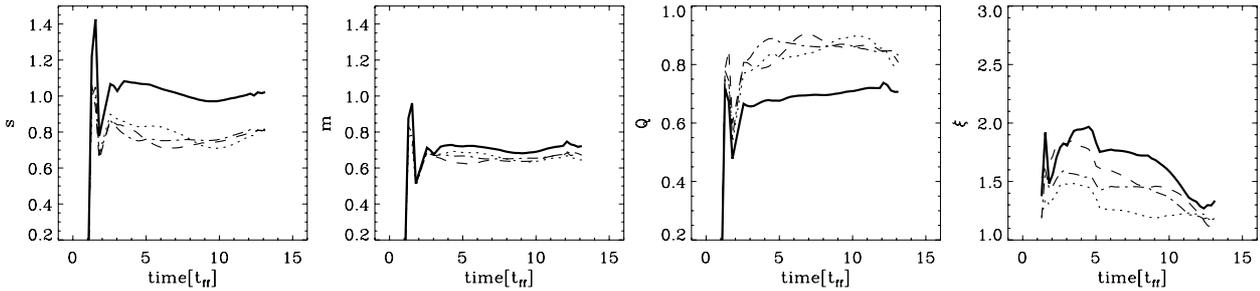
Fig. 5. The temporal evolution of the clustering parameters $\bar{s}$, $\bar{m}$, Q, and ξ of the model M6k4a for the 3D analysis (solid line) and for the projection into the *xy* (dotted), *xz* (dashed), and *yz* plane (dash-dotted).
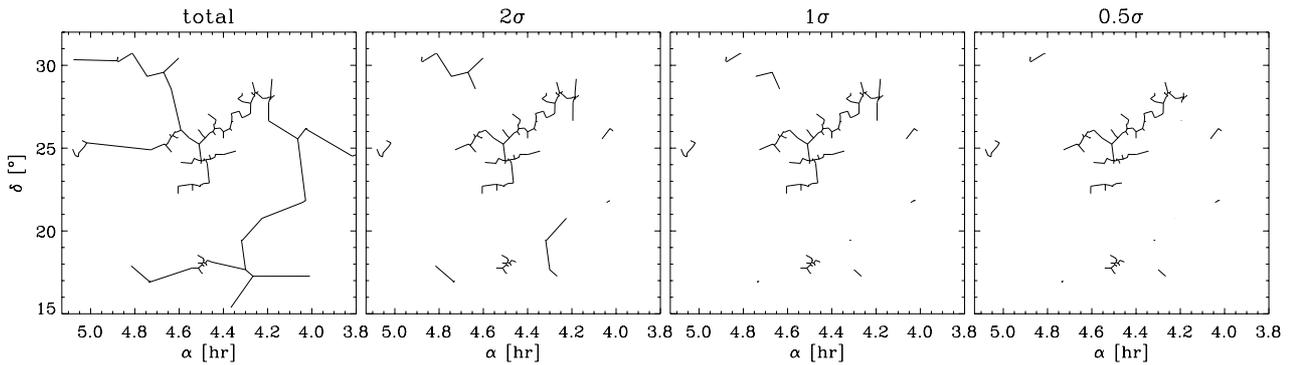

Fig. 6. The MST of the Taurus cluster as a whole (*left panel*) and separated at $l + 2\sigma_{\rm m}$, $l + \sigma_{\rm m}$, and $l + 0.5\sigma_{\rm m}$ (*from left to right*).

cluster builds up from separate groups which will grow into a single, more centrally concentrated cluster, as also found by Bonnell et al. (2003) and Clark et al. (2005).

6.3. The effect of projection

Looking at the two-dimensional projections of the 3D model clusters does not significantly change the picture as a whole. The individual $\bar{s}$, $\bar{m}$ and Q values can indeed differ for the projection into the *xy*, *xz*, and *yz* plane, but the qualitative behaviour of the evolution is more or less the same, independent of the projection (Fig. 5). While the $\bar{s}_{\rm 3D}$ and $\bar{m}_{\rm 3D}$ values usually are higher than the values of the projections, $Q_{\rm 3D}$ tends to be lower than the values of the projections. The investigation of several hundred randomly created clusters shows that $\bar{s}_{\rm 3D}$ is always expected to be larger than the value for the projections ($\bar{s}_{\rm 3D}/\bar{s}_{\rm 2D} \approx 1.2$), while the $\bar{m}$ values can be in the range $0.95 \lesssim \bar{m}_{\rm 3D}/\bar{m}_{\rm 2D} \lesssim 1.4$ with a mean value of 1.1. In the extreme case, $Q_{\rm 3D}$ can differ by up to 30% from the 2D value. Note that the physical interpretation of Q as given by CW04 is based on the two-dimensional analysis. Therefore, for an interpretation of the numerical values (and not only the trend) of Q, the projected values have to be used.

The elongation measure ξ, on the other hand, depends strongly on the projection. Obviously, an elongated, filament-like structure seen from the side will look spherical when observed along its major axis. In the case of the models the three-dimensional value of ξ can be used to describe the true shape of the cluster. A large scatter in the $\xi_{\rm 2D}$ values means a high $\xi_{\rm 3D}$ value and vice versa.

6.4. The effect of separating

We separate the observed clusters by succesively removing all MST edges with lengths l larger than $l + 3\sigma_{\rm m}$, $l + 2\sigma_{\rm m}$, $l + \sigma_{\rm m}$, and $l + 0.5\sigma_{\rm m}$, where $\sigma_{\rm m}$ denotes the standard deviation of the mean edge length of the MST. This process is demonstrated for the Taurus star-forming region in Fig. 6. When we compare all three regions with sufficient data, we see that for Serpens and Taurus only solitary stars in the outskirts of the cluster are removed, while the central part of the cluster stays connected. However, the more homogeneous ρ Oph cluster breaks down in two roughly equally large clusters at the last step. In Fig. 7 we show the clustering parameters of the largest remaining subcluster (i.e., the one containing the largest number of stars after each level of separation). The parameter Q is less affected by the separating procedure than $\bar{s}$ and $\bar{m}$. Significant changes are only seen when large parts of the cluster are excluded (as in the step from 1 to 0.5σ for ρ Oph, when the cluster breaks up into two large subclusters). The elongation measure ξ varies significantly with the level of separation. For example, the remaining subcluster in Taurus shows a filamentary structure and thus a larger ξ ($\xi = 1.83$) than the cluster as a whole.

7. Summary and conclusions

We show that the normalised mean correlation length $\bar{s}$ and the mean edge length of the minimum spanning tree $\bar{m}$, and in particular the combination of both parameters, Q, as proposed by CW04, are very useful tools to study the structures of star clusters, both observed ones and those from numerical simulations. We refine the definiton of the cluster area by using the normalised convex hull rather than a circular or rectangular area
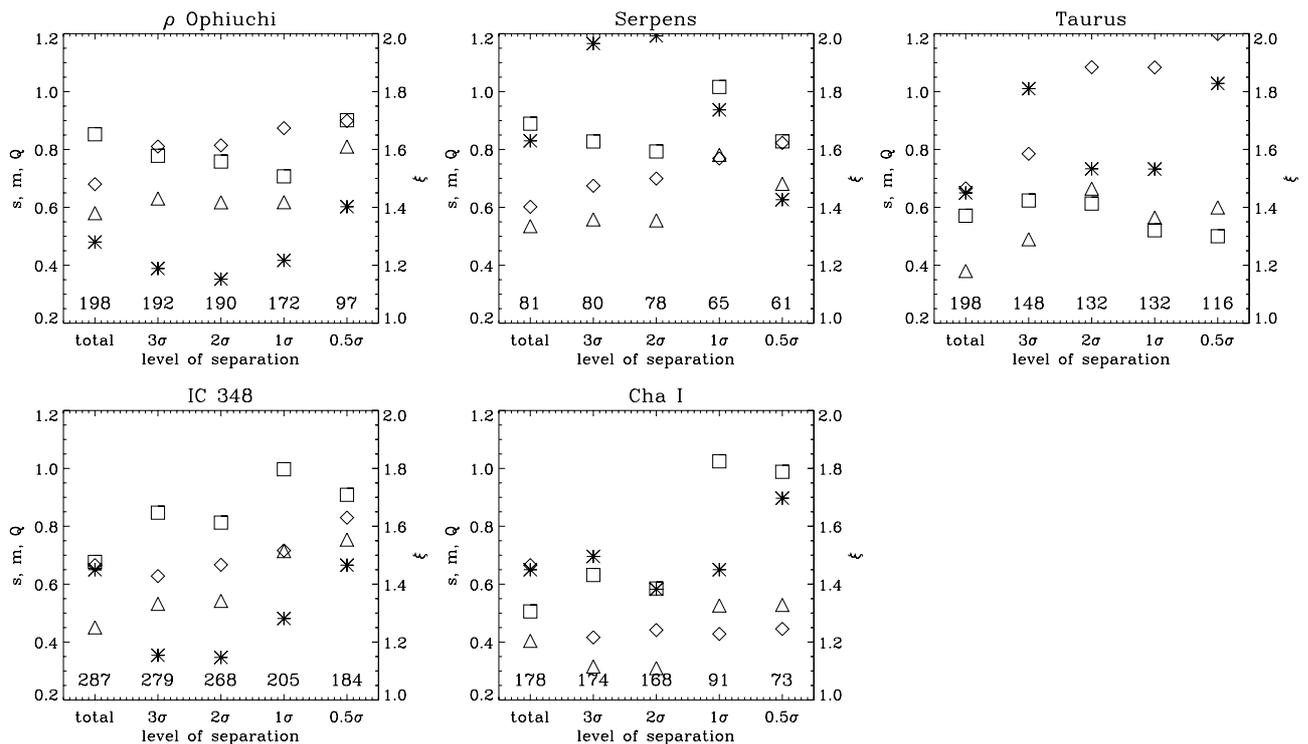
Fig. 7. The clustering parameters $\bar{s}$ (diamonds), $\bar{m}$ (triangles), Q (squares), and ξ (asterisks; scale on the right-hand axis) of the largest remaining subcluster at different levels of separation. The numbers along the abscissa give the number of YSOs contained in the particular subcluster.

around the objects. Unlike $\bar{s}$ and $\bar{m}$ the parameter Q is independent of the definiton of the area of the cluster. In addition, it is less affected when removing stars from the cluster by separating the MST.

We introduce a new measure ξ for the elongation of a cluster. It is defined as the ratio of the cluster radius determined by an enclosing circle to the cluster radius derived from the normalised convex hull. This is a stable statistical measure not influenced by fractal substructre, which could also be applied to the filamentary structure of molecular clouds.

The mean separation s increases with the evolutionary class, reflecting the expansion of the cluster. The prestellar cores do not follow that sequence, leading us to the speculation that not all objects classified as prestellar cores will eventually form stars. The clustering values of the models lie roughly in the same range as those from observed clusters. A particularly good agreement is reached when the clusters have similar elongation values ξ. No correlation with the Mach number or the wave number of the turbulence is found. We conclude that possible influences of the turbulent environment on the clustering behaviour are quickly smoothed out by the velocity dispersion of the young stars.

The temporal evolution of the clustering parameters shows that the star cluster builds up from several subclusters and evolves to a more centrally concentrated state. New stars are formed faster than the cluster expands. The projection of the 3D models into a 2D plane changes the clustering parameters, but not the general behaviour with time.

References

Adami, C., & Mazure, A. 1999, A&AS, 134, 393

Adams, F. C., & Wiseman, J. J. 1994, ApJ, 435, 693

André, P., Ward-Thompson, D., & Barsony, M. 2000, in Protostars and Planets IV, ed. V. Mannings, A. P. Boss, & S. S. Russell (Tucson: University of Arizona Press), 59

Barber, C. B., Dobkin, D. P., & Huhdanpaa, H. T. 1996, ACM Transactions on Mathematical Software, 22, 469, http://www.qhull.org

Barrow, J. D., Bhavsar, S. P., & Sonoda, D. H. 1985, MNRAS, 216, 17

Bate, M. R., & Burkert, A. 1997, MNRAS, 288, 1060

Bate, M. R., Bonnell, I. A., & Price, N. M. 1995, MNRAS, 277, 362

Bate, M. R., Clarke, C. J., & McCaughrean, M. J. 1998, MNRAS, 297, 1163

Beardwood, J., Halton, J. H., & Hammersley, J. M. 1959, Proc. Cambridge Philos. Soc., 55, 299

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

Bontemps, S., André, P., Kaas, A. A., et al. 2001, A&A, 372, 173

Cambrésy, L., Copet, E., Epchtein, N., et al. 1998, A&A, 338, 977

Cartwright, A., & Whitworth, A. P. 2004, MNRAS, 348, 589 (CW04)

Clark, P. C., Bonnell, I. A., Zinnecker, H., & Bate, M. R. 2005, MNRAS, 359, 809

Doroshkevich, A., Tucker, D. L., Allam, S., & Way, M. J. 2004, A&A, 418, 7

Dussert, C., Rasigni, G., Rasigni, M., Palmari, J., & Llebaria, A. 1986, Phys. Rev. B, 34, 3528

Froebrich, D. 2005, ApJS, 156, 169

Froebrich, D., Schmeja, S., Smith, M. D., & Klessen, R. S. 2005, MNRAS, submitted

Gladwin, P. P., Kitsionas, S., Boffin, H. M. J., & Whitworth, A. P. 1999, MNRAS, 302, 305

Gomez, M., Hartmann, L., Kenyon, S. J., & Hewett, R. 1993, AJ, 105, 1927

Gower, J. C., & Ross, G. J. S. 1969, Appl. Stat., 18, 54

Haikala, L. K., Harju, J., Mattila, K., & Toriseva, M. 2005, A&A, 431, 149

Hartmann, L. 2002, ApJ, 578, 914

Hoffman, R., & Jain, A. K. 1983, Pattern Recognition Letters, 1, 175

Hurt, R. L., & Barsony, M. 1996, ApJ, 460, L45

Jappsen, A.-K., Klessen, R. S., Larson, R. B., Li, Y., & Mac Low, M.-M. 2005, A&A, 435, 611

Kaas, A. A., Olofsson, G., Bontemps, S., et al. 2004, A&A, 421, 623

Klessen, R. S. 1997, MNRAS, 292, 11

Klessen, R. S., & Burkert, A. 2000, ApJS, 128, 287

Klessen, R. S., & Kroupa, P. 2001, A&A, 372, 105

Kruskal, J. B., Jr. 1956, Proc. Amer. Math. Soc., 7, 48

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

Larson, R. B. 1995, MNRAS, 272, 213

Larson, R. B. 2003, Rep. Prog. Phys., 66, 1651

Lee, C. W., & Myers, P. C. 1999, ApJS, 123, 233

Li, Y., Mac Low, M.-M., & Klessen, R. S. 2005, ApJ, 626, 823

Luhman, K. L., Stauffer, J. R., Muench, A. A., et al. 2003, ApJ, 593, 109

Marcelpoil, R. 1993, Analytical Cellular Pathology, 5, 177

Monaghan, J. J. 1992, ARA&A, 30, 543

Prim, R. C. 1957, Bell Systems Tech. J., 36, 1389

Ripley, B. D., & Rasson, J.-P. 1977, J. Appl. Prob., 14, 483

Robins, V., Meiss, J. D., & Bradley, E. 2000, Physica D, 139, 276

Schmeja, S., & Klessen, R. S. 2004, A&A, 419, 405 (SK04)

Schmeja, S., Klessen, R. S., & Froebrich, D. 2005, A&A, 437, 911 (SKF05)

Stanke, T., Smith, M. D., Gredel, R., & Khanzadyan, T. 2005, A&A, in press [arXiv:astro-ph/0511093]

Vázquez-Semadeni, E., Kim, J., Shadmehri, M., & Ballesteros-Paredes, J. 2005, ApJ, 618, 344
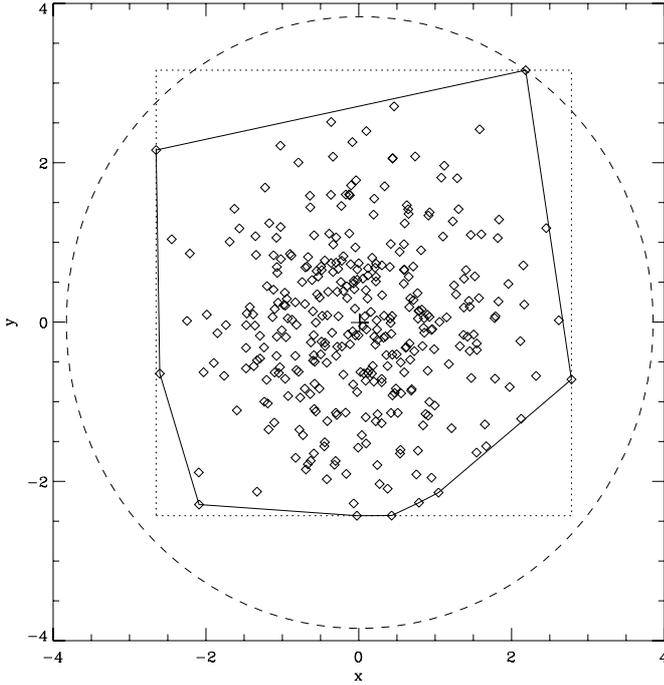
Online Material

Fig. A.1. Randomly distributed data points and the area they span according to different definitions (circle, rectangle, convex hull).

Appendix A: Radius and area of a cluster

For the normalisation of the parameters $\bar{s}$ and $\bar{m}$ the radius and the area (or volume) of the cluster are needed (see Appendix B). Since a star cluster has no well-defined natural boundary, different approaches to determine the cluster area have been used. CW04 define the cluster radius as the distance between the mean position of all cluster members and the most distant star and the area as a circle with the cluster radius: $A = \pi R^2_{\text{cluster}}$. Adami & Mazure (1999) define the area used for the normalisation of the MST edge lengths as the maximum rectangle of the point set. However, both methods tend to significantly overestimate the area of the cluster, in particular if the cluster is elongated or irregularly shaped rather than spherical. Therefore, we estimate the area A of the cluster using the convex hull of the data points, normalised by an additional geometrical factor taking into account the ratio of the number of objects inside and on the convex hull:

$$A = m(H)/[1 - (v_{\text{H}}/n)] \qquad (A.1)$$

(Hoffman & Jain 1983; Ripley & Rasson 1977), where H is the convex hull, $m(H)$ its area, v_{H} the number of vertices on H and n the number of objects. The correction factor is used since the convex hull by itself tends to be smaller than the true sampling window. To be consistent, we define the cluster radius as the radius of a circle with the same area A. In the three-dimensional case the volume V and the radius are defined analogously. The convex hull is computed using the programme *Qhull* (Barber et al. 1996).

Figure A.1 demonstrates that the definition of the cluster area is crucial, since it can differ by a factor of two or more. In the given example of randomly distributed data points, the area of the circle is 46.33, that of the rectangle is 30.40, the area of

the convex hull is 24.19, and the normalised area (Eq. (A.1)) is 24.83. For the calculation of the parameter $Q = \bar{m}/\bar{s}$ the size of the radius/area is irrelevant (as long as the area used for the normalisation of $\bar{m}$ depends on the radius used for the normalisation of $\bar{s}$ or vice versa). The radius is cancelled since it is used to normalise both, $\bar{m}$ and $\bar{s}$.

Appendix B: Normalisation of the minimum spanning tree

The edge lengths of a minimum spanning tree depend on the number of points and on the area, therefore the mean edge length m has to be normalised, in order to compare the results from samples with different numbers of objects and/or different areas. Beardwood et al. (1959) show that the length of the shortest closed path through n points $l(P^n)$ in a plane region of area A is asymptotically proportional to $\sqrt{nA}$ for large n. Since the number of edges in a MST is $(n-1)$, Hoffman & Jain (1983) and CW04 claim that the expected length of a randomly selected edge of a MST is asymptotically proportional to

$$\frac{\sqrt{nA}}{n-1}. \qquad (B.1)$$

They use this factor to normalise the mean edge lengths of different data sets. Beardwood et al. (1959), however, deal with the *closed path* through n points (i.e., the travelling salesman problem), not a MST, as cited incorrectly by Hoffman & Jain (1983) and CW04. Marcelpoil (1993), on the other hand, deduces the normalisation factor

$$\sqrt{A/n} \qquad (B.2)$$

for comparing the mean edge lengths of different samples. Since

$$\lim_{n\to\infty} \frac{\sqrt{nA}/(n-1)}{\sqrt{A/n}} = \lim_{n\to\infty} \frac{n}{n-1} = 1, \qquad (B.3)$$

both methods lead to the same result for large n, e.g. for $n > 100$ the difference is $\lesssim 1\%$ (see Fig. B.1). In this work we use the normalisation factor of Marcelpoil (1993), which seems more plausible to us.

In the three-dimensional case the general result for k dimensions of Beardwood et al. (1959) (their Eq. (2)) can be written as $l(P^n) \propto \sqrt[3]{Vn^2}$, leading to the factor

$$\frac{\sqrt[3]{Vn^2}}{n-1} \qquad (B.4)$$

to normalise the mean edge lengths of the 3D MST. However, we follow the reasoning of Marcelpoil (1993) and use the normalisation factor

$$\sqrt[3]{V/n} \qquad (B.5)$$

in the 3D case. The volume V is the volume of the convex hull as discussed above. Again, the ratio of both normalisation factors asymptotically approaches 1, making them interchangeable for large n.

Note that we use different definitions of the radius and the area from CW04 (Appendix A), resulting in an additional difference between our and their parameters $\bar{m}$ and $\bar{s}$. However,
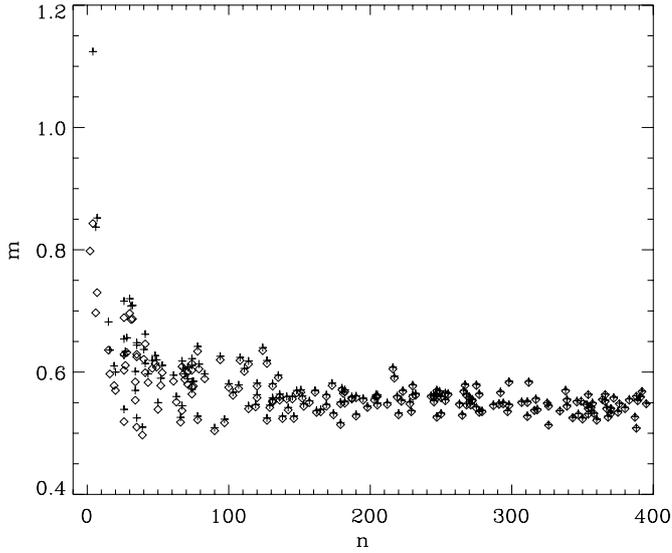
Fig. B.1. Mean edge lengths $\bar{m}$ for 200 randomly created 2D sets of points ($3 \leq n \leq 400$), normalised with the factor B.2 (crosses) and B.1 (diamonds), respectively, plotted versus the number of points.
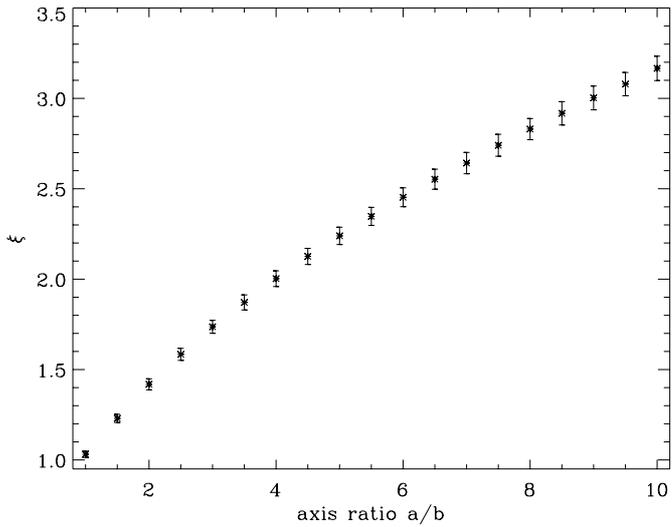

Fig. C.1. The relation between the axis ratio a/b of an elliptical area and the elongation ξ, with the 1σ error also indicated.

in the calculation of Q the radius contained in the normalisation factor is cancelled out, so the difference in the Q parameters boils down to Eq. (B.3). Unlike $\bar{m}$ and $\bar{s}$, the more relevant parameter Q is the same for a large number of objects, independent of the normalisation method. To allow comparison, we list the parameters computed using the normalisation factor and the cluster radius according to CW04 in Table 1 as well.

Appendix C: Elongation of a cluster

We notice that some of the model clusters are strongly elongated, causing a large difference in the cluster area depending on whether it is defined by the enclosing circle or by the normalised convex hull. We use this fact to propose a new, statis-

tically stable characterisation of the elongation of a cluster. We define the elongation measure ξ as the ratio

$$\xi = \frac{R_{\text{cluster}}^{\text{circle}}}{R_{\text{cluster}}^{\text{conv. hull}}}. \tag{C.1}$$

This is an approach similar to the "filament index" $\mathcal{F}$ introduced by Adams & Wiseman (1994). However, while Adams & Wiseman (1994) use the actual area of a cloud obtained from column density maps, we have to estimate the area of the cluster of point sources by the convex hull. This makes ξ a reliable measure for the elongation, since it excludes the possibility that fractal substructure in an otherwise spherical cluster leads to a large ratio of the two radii. The measure ξ could also be used to quantify the filamentary structure in molecular clouds.

In order to test if ξ is indeed a good measure for the elongation we place 350 points randomly on an elliptical area with increasing axis ratio ($1 \leq a/b \leq 10$). To minimise the statistical scatter we perform 500 different realisations for each a/b and determine the mean ξ and its standard deviation for each step. Figure C.1 shows that ξ increases with the axis ratio a/b. For $a/b = 1$ (i.e., a circle), $\xi \approx 1$, as expected.