# Structure and mass segregation in Galactic stellar clusters

Sami Dib,[1,2]★ Stefan Schmeja[3,4] and Richard J. Parker[5]†

[1]*Instituto de Astrofísica y Ciencias Planetarias, Universidad de Atacama, Copayapu 485, Copiapó, Chile*
[2]*Niels Bohr International Academy, Niels Bohr Institute, Blegdamsvej 17, DK-2100 Copenhagen, Denmark*
[3]*Technische Informationsbibliothek, Welfengarten 1b, D-30167 Hannover, Germany*
[4]*Astronomisches Rechen-Institut, Zentrum für Astronomie der Universität Heidelberg, Mönchhofstraße 12-14, D-69120 Heidelberg, Germany*
[5]*Department of Physics and Astronomy, The University of Sheffield, Hicks Building, Hounsfield Road, Sheffield S3 7RH, UK*

## ABSTRACT

We quantify the structure of a very large number of Galactic open clusters and look for evidence of mass segregation for the most massive stars in the clusters. We characterize the structure and mass segregation ratios of 1276 clusters in the Milky Way Stellar Cluster (MWSC) catalogue containing each at least 40 stars and that are located at a distance of up to $\approx 2$ kpc from the Sun. We use an approach based on the calculation of the minimum spanning tree of the clusters, and for each one of them, we calculate the structure parameter $\mathcal{Q}$ and the mass segregation ratio $\Lambda_{\mathrm{MSR}}$. Our findings indicate that most clusters possess a $\mathcal{Q}$ parameter that falls in the range 0.7–0.8 and are thus neither strongly concentrated nor do they show significant substructure. Only 27 per cent can be considered centrally concentrated with $\mathcal{Q}$ values $>0.8$. Of the 1276 clusters, only 14 per cent show indication of significant mass segregation ($\Lambda_{\mathrm{MSR}} > 1.5$). Furthermore, no correlation is found between the structure of the clusters or the degree of mass segregation with their position in the Galaxy. A comparison of the measured $\mathcal{Q}$ values for the young open clusters in the MWSC to *N*-body numerical simulations that follow the evolution of the $\mathcal{Q}$ parameter over the first 10 Myr of the clusters life suggests that the young clusters found in the MWSC catalogue initially possessed local mean volume densities of $\rho_*$ $\approx 10$–$100\,\mathrm{M}_\odot\,\mathrm{pc}^{-3}$.

**Key words:** stars: formation – Galaxy: general – open clusters and associations: general – open clusters and associations: individual – Galaxy: stellar content.

## 1 INTRODUCTION

Star clusters are fundamental building blocks of galactic discs and most stars, if not all, form in clusters (e.g. Carpenter 2000; Lada & Lada 2003; Dib 2011; Dib et al. 2011a,b; Dib 2014; Mallick et al. 2014; Hony et al. 2015). The dynamics of stars in the clusters as well as the structure of clusters measured as a function of cluster age hold important clues on the processes of star formation and stellar evolution. As clusters age, the expulsion of gas by stellar feedback as well as dynamical interactions between stars and binary systems in the cluster soften its gravitational potential, leading to their expansion and to their partial or total dissolution into the field of their host galaxy (e.g. Spitzer & Harm 1958; Dib 2011; Dib et al. 2011a,b; Parker & Meyer 2012; Dib et al. 2013; Parker & Dale 2013; Pfalzner & Kaczmarek 2013; Brinkmann et al. 2017). Clusters can also be disrupted by close encounters with giant molecular clouds as they

orbit the Galactic Centre (e.g. Gieles et al. 2006) or by being subjected to strong tidal fields (e.g. Dalessandro et al. 2015 ; Martinez-Medina et al. 2017 ; Zhai et al. 2017).

The initial spatial distribution of stars in young clusters may reflect the structure of the parental protostellar clump/cloud (e.g. Dib et al. 2010a; Lomax, Whitworth & Cartwright 2011; Gouliermis, Hony & Klessen 2014; Hony et al. 2015). However, as the clusters evolve, their structure is shaped by the gravitational interactions between member stars and by tidal effects, and the structure of the clusters will reflect their dynamical evolution. Numerical simulations of star cluster formation show that clusters can build up in a hierarchical way from several sub-clusters which evolve dynamically and merge into a single, centrally concentrated cluster (e.g. Bonnell & Bate 2006; Schmeja & Klessen 2006; Allison et al. 2010; Moeckel & Bate 2010; Padoan et al. 2014; Parker et al. 2014; Fujii 2015) or from the direct collapse of a single gravitationally bound clump (e.g. Banerjee & Kroupa 2015). The latter scenario is likely to be required in order to reproduce the high star formation efficiencies and short age spreads observed in massive clusters (Dib et al. 2013). Substructure in a fractal cluster may be erased

★ E-mail: sami.dib@gmail.com
† Royal Society Dorothy Hodgkin Fellow

rapidly or preserved for a longer time, depending on the stellar velocity dispersions. Results from *N*-body simulations (e.g. Goodwin & Whitworth 2004) indicate that in clusters with low initial stellar velocity dispersions, the resulting collapse of the cluster tends to erase substructure to a large extent. In clusters with virial ratios[1] of 0.5 or higher, however, initial substructure survives for several crossing times. Spatial substructure has been observed in clusters as old as ≈100 Myr (e.g. Sánchez & Alfaro 2009). However, the structure of open clusters may also be a result of later dynamical evolution.

In many star clusters, the brightest, most massive stars are concentrated towards the centre of the cluster, which is usually attributed to mass segregation (e.g. Dib et al. 2010a; Hasan & Hasan 2011; Haghi et al. 2015; Sheikhi et al. 2016). Whether mass segregation occurs due to an evolutionary effect or is of primordial origin is not yet entirely clear. In the first case, massive stars formed elsewhere in the cluster eventually sink to the cluster centre through the effects of two-body relaxation (e.g. McMillan, Vesperini & Portegies Zwart 2007; Allison et al. 2009a). This is corroborated by numerical simulations in which mass segregation occurs on time-scales that are of the order of the clusters ages (e.g. Allison et al. 2010; Parker et al. 2014). In the second scenario, massive stars form preferentially in the central region of the cluster either by efficiently accreting gas due to their location at the bottom of the cluster potential well (e.g. Dib et al. 2010a) or by a coalescence process of less massive stars (Dib, Kim & Shadmehri 2007a; Dib et al. 2008a). The fact that mass segregation is also observed in young clusters (e.g. Littlefair et al. 2003; Gouliermis et al. 2004; Stolte et al. 2006; Chen, de Grijs & Zhao 2007; Sharma et al. 2007; Gennaro et al. 2011; Feigelson et al. 2013; Habibi et al. 2013; Pang et al. 2013) might suggest that the second scenario is more likely, but the question is still under intense debate.

In the Milky Way, several studies have explored the dependence (or lack of it) of some of the properties of open clusters with their age and surface density, such as their members' richness (e.g. Tadross et al. 2002), size (Schilbach et al. 2006; Tadross 2014), galactic scaleheight (Buckner & Froebrich 2014), mass segregation (Bukowiecki et al. 2012) or structure (Gregorio-Hetem et al. 2015), as well as the dependence of the cluster structure and metallicity on their position in the Galactic disk (e.g. Friel 1995; Froebrich et al. 2010; Bukowiecki et al. 2011; Tadross 2014). The Milky Way Star Cluster (MWSC) catalogue (Kharchenko et al. 2012,2013; Schmeja et al. 2014; Scholz et al. 2015; Dib, Schmeja & Hony 2017) offers the largest homogeneous sample of Galactic open clusters, allowing us to study the spatial structure and mass segregation in a large number of clusters over a wide range of cluster ages ranging from young clusters with ages ≈1 Myr to older clusters with ages of about 5 Gyr. In Section 2, we briefly recount some of the characteristics of the MWSC, while in Section 3, we describe the methods we use to describe the structure and mass segregation of clusters in the catalogue. The results on the existence of correlations (or lack of it) between the structure and mass segregation levels in the clusters

versus cluster properties are presented and discussed in Section 4, and in Section 5, we conclude.

## 2 DATA

The MWSC catalogue (Kharchenko et al. 2012, 2013) with its latest addition of predominantly old open clusters (Schmeja et al. 2014) contains 3145 confirmed Galactic open clusters, globular clusters and compact associations. They have been analysed in a homogeneous way using 2MASS and PPMXL (Röser, Demleitner & Schilbach 2010), resulting in uniform structural, kinematic and astrophysical data, such as radii, tidal radii, distances, ages, as well as the membership probability of stars in the cluster, among several other properties. Scholz et al. (2015) added 63 additional clusters to the catalogue, raising the total number to 3208. Apart from about 60 old (ages ≳ 1 Gyr) open clusters missing within 1 kpc of the Sun (Schmeja et al. 2014), the MWSC is complete to a distance of about 1.8 kpc. We also refer the reader to Dib, Schmeja & Hony (2017) for further discussion on the completeness of the MWSC catalogue with respect to various implied initial cluster mass functions in the Galaxy. The MWSC contains spatial, kinematic and photometric membership probabilities $P_s$, $P_{kin}$, $P_{JH}$ and $P_{JK}$ for each star within the cluster area. For more details on the determination of these probabilities, see Kharchenko et al. (2012). The combined membership probability is thus defined as:

$$P = P_s \cdot \min(P_{kin}, P_{JH}, P_{JK}). \tag{1}$$

In this work, following Kharchenko et al. (2013), we consider a star to be a cluster member when it has a membership probability of $P \geq 61$ per cent or higher. In addition, we require that the 2MASS quality flag $Q\_flg = A$ (corresponding to a signal-to-noise ratio $S/N > 10$) in each photometric band for stars fainter than $K_s = 7$ (Kharchenko et al. 2012).

## 3 METHODS

In order to study the clusters structure and mass segregation, we use two methods based on a minimum spanning tree (MST) which is the unique set of straight lines ('edges') connecting a given set of points without closed loops, such that the sum of all edge lengths is a minimum (Borůvka 1926; Kruskal 1956; Prim 1957; Gower & Ross 1969). These methods are detailed below.

### 3.1 Structure parameter

A commonly used quantity to characterize the structure of clusters is the $\mathcal{Q}$ parameter (Cartwright & Whitworth 2004, 2009), which is given by

$$\mathcal{Q} = \frac{\bar{\ell}_{MST}}{\bar{s}}. \tag{2}$$

The parameter combines the normalized correlation length $\bar{s}$, i.e. the mean distance between all stars, and the normalized mean edge length $\bar{\ell}_{MST}$ derived from the MST. The $\mathcal{Q}$ parameter is used to quantify the structure of a cluster and to distinguish between clusters with a central density concentration and hierarchical clusters with a fractal substructure. Large $\mathcal{Q}$ values ($\mathcal{Q} > 0.8$) are associated with centrally condensed clusters with radial density profiles $\rho(r) \propto r^{-\alpha}$, while small $\mathcal{Q}$ values ($\mathcal{Q} < 0.8$) indicate clusters with a fractal substructure. $\mathcal{Q}$ is correlated with $\alpha$ for $\mathcal{Q} > 0.8$ and anticorrelated with the fractal dimension $D$ for $\mathcal{Q} < 0.8$ (Cartwright & Whitworth 2004, in particular see fig. 5 in their paper). An interesting aspect of the $\mathcal{Q}$

[1] Our definition of the virial ratio is $\alpha_{vir} = E_k/E_{grav}$, where $E_k$ and $E_{grav}$ are the total kinetic and potential energy, respectively. In the *N*-body models we compare our observations in Section 4.3, an initial value of $\alpha_{vir} < 0.3$ refers to subvirial initial conditions, whereas $\alpha_{vir} = 0.5$ refers to a virial case. However, since the models considered in Section 4.3 have spatial and velocity substructure, a value of $\alpha_{vir} = 0.5$ does not necessarily imply virial equilibrium. For more discussion on this point, we refer the reader to Parker et al. (2014).

parameter is that it measures the level of substructure present in a cluster independent of the cluster density. A detailed description of the method, and in particular its implementation used in this study, is given in Schmeja & Klessen (2006).

### 3.2 Mass segregation ratio

Allison et al. (2009a) introduced the mass segregation ratio ($\Lambda_{MSR}$) as a measure to identify and quantify mass segregation in clusters. The method is based on a calculation of the length of the MST, $l_{MST}$, which measures the compactness of a given sample of vertices in the MST. The mass segregation of a cluster is measured by comparing the value of $l_{MST}$ of the $n_{MST}$ most massive stars, $l_{MST}^{mp}$, with the average $l_{MST}$ of k sets of n random stars, $\langle l_{MST}^{rand} \rangle$. The value of $\Lambda_{MSR}$ is then given by

$$\Lambda_{MSR} = \frac{\langle l_{MST}^{rand} \rangle}{l_{MST}^{mp}}. \tag{3}$$

The error on $\Lambda_{MSR}$ is given by

$$\Delta \Lambda_{MSR} = \frac{\Delta l_{MST}^{rand}}{l_{MST}^{mp}}, \tag{4}$$

where $\Delta l_{MST}^{rand}$ is the standard deviation from the random sets. The method has been modified by Olczak, Spurzem & Henning (2011) by using the geometric mean rather than the arithmetic mean in order to minimize the influence of outliers. This method works by constructing the MST for the $n_{MST}$ most massive stars and determining the mean edge length $\gamma_{mp}$. Then, we construct the MST of the same number of randomly selected stars from the entire sample and determine the mean edge length $\gamma_{rand}$. The value of the MSR following Olczak, Spurzem & Henning (2011), $\Gamma_{MSR}$, is then given by:

$$\Gamma_{MSR} = \frac{\langle \gamma_{MST}^{rand} \rangle}{\gamma_{MST}^{mp}}, \tag{5}$$

and the associated standard deviation of $\Gamma_{MSR}$ is given by:

$$\Delta \Gamma_{MSR} = \Delta \gamma_{MST}^{rand}. \tag{6}$$

In this work, we compute both $\Lambda_{MSR}$ using the arithmetic mean as in Allison et al. (2009a), and $\Gamma_{MSR}$ using the geometric mean following Olczak, Spurzem & Henning (2011). In each case, this is done 100 times in order to obtain the quantities $\langle l_{MST}^{rand} \rangle$ and $\langle \gamma_{MST}^{rand} \rangle$. A value of $\Lambda_{MSR} \approx 1$ (respectively, $\Gamma_{MSR} \approx 1$) implies that both samples of stars (i.e. the most massive and the randomly selected) are distributed in a similar manner, whereas $\Lambda_{MSR} \gg 1$ (respectively, $\Gamma_{MSR} \gg 1$) indicates mass segregation and $\Lambda_{MSR} \ll 1$ (respectively, $\Gamma_{MSR} \ll 1$) points to inverse mass segregation, i.e. the massive stars are more spread outwards than the rest.

Since the vast majority of the clusters in the sample have ages that are much larger than a few million years and are therefore unaffected by extinction effects, we use the magnitudes of stars as a proxy for the mass. This also has the advantage of avoiding to introduce additional uncertainties when converting the observed luminosities into masses. Fig. 1 displays the dependence of $\Lambda_{MSR}$ (top row) and $\Gamma_{MSR}$ (bottom row) on $n_{MST}$ for the different bands available in the MWSC, namely the $B, V, J, H$ and $K_s$ bands for two selected clusters. The figure displays the case of a mass segregated cluster (MWSC 2202, left column) and of non-mass segregated cluster (MWSC 0141, right column). We observe that the level of mass segregation is insensitive to the choice of wavelength. For the remaining clusters in our sample (1276 in total, see Section 3.3 below), we calculate $\Lambda_{MSR}$ and $\Gamma_{MSR}$ using the $J$ band observations. Although the number
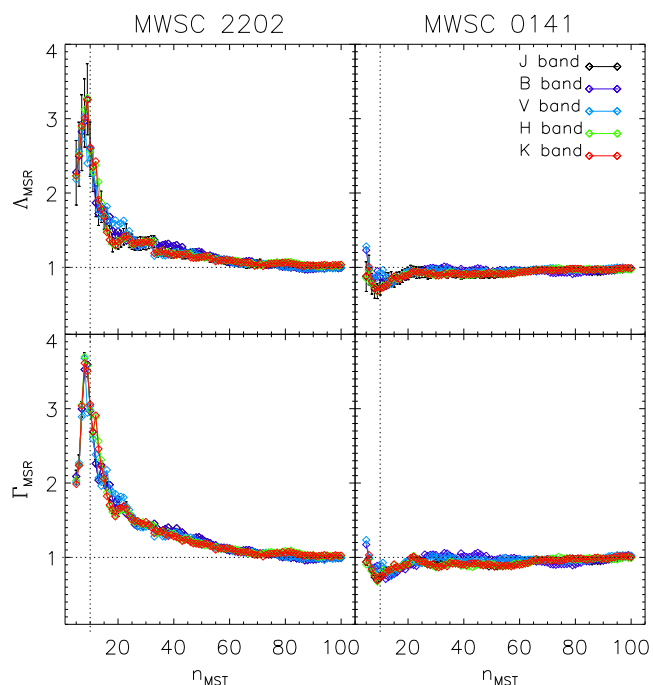


**Figure 1.** Examples of the mass segregation ratios $\Lambda_{MSR}$ (Allison et al. 2009a, top row) and $\Gamma_{MSR}$ (Olczak, Spurzem & Henning 2011, bottom row) as a function of the number of stars used for computing them $n_{MST}$, using the different bands available in the MWSC catalogue for two clusters. For the sake of clarity, error bars are only shown for the $J$-band measurements. The left-hand panels display the case of a clearly mass-segregated cluster (MWSC 2202 = NGC 5460) and the right-hand panels a cluster without any signs of mass segregation (MWSC 141 = ASCC 6, lower panel). The horizontal dotted line shows the division between non-mass-segregated and mass-segregated clusters at $\Lambda_{MSR} = 1$ (or $\Gamma_{MSR} = 1$). The vertical line is placed at $n_{MST} = 10$, the value used for comparing the mass segregation ratios in the remaining of the text.

of stars in each cluster varies greatly in the sample (between a few tens to more than 4000), all clusters show a similar behaviour. We have verified that if a cluster shows evidence of mass segregation, this is usually seen only for $n_{MST} \lesssim 20$, regardless of the total number of cluster members. Therefore, $n_{MST} = 10$ is a well-justified choice for comparing different clusters. Hereafter, we will refer to the $\Lambda_{MSR}$ and $\Gamma_{MSR}$ parameters as $\Lambda_{10}^J$ and $\Gamma_{10}^J$.

### 3.3 Application to the data

In order to minimize biases and selection effects, we only consider clusters closer than 2 kpc from the Sun. With a decreasing number of stars in a cluster, the error on $\mathcal{Q}$ increases, and the $\mathcal{Q}$ values become less reliable (e.g. Gouliermis et al. 2012). Also the MSR analysis requires a minimum number of objects to give meaningful results. Therefore, we consider only those clusters with 40 or more members (where $\sigma_{\mathcal{Q}} \lesssim 10$ per cent). Applying the restrictions ($d < 2$ kpc; a minimum number of stars of 40 with $P > 61$ per cent in a cluster) leaves 1276 clusters that are used in this study.

## 4 RESULTS AND DISCUSSION

We calculated $\mathcal{Q}$ and the MSR ($\Lambda_{10}^J$ and $\Gamma_{10}^J$) for all of the 1276 clusters in the sample. Fig. 2 displays the distribution of the $\mathcal{Q}$ parameter for the entire sample (left-hand panel). The values of $\mathcal{Q}$ lie in the range $0.67 < \mathcal{Q} < 0.97$ with an arithmetic mean value and standard deviation of $\bar{\mathcal{Q}} = 0.78 \pm 0.04$. Only 344 clusters (26.95 per cent)

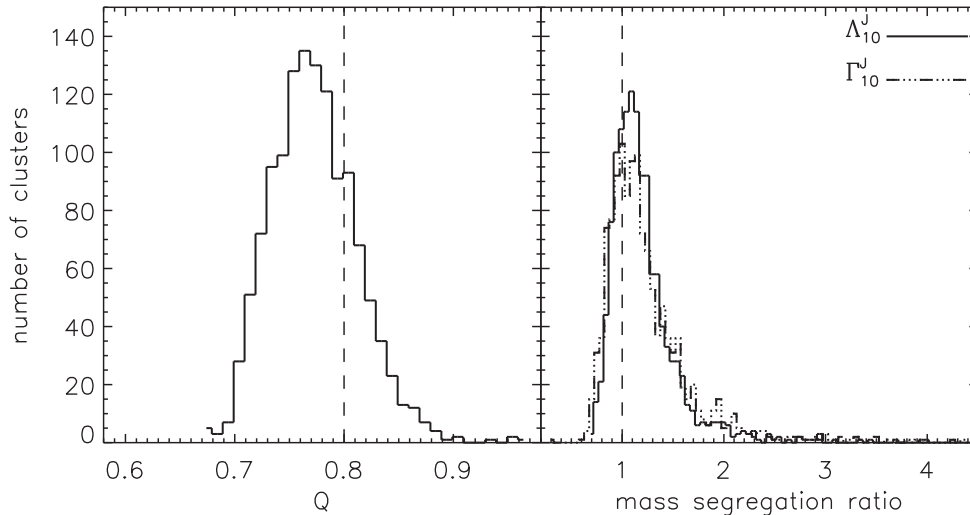**Figure 2.** Distribution of the structure parameter $\mathcal{Q}$ (left-hand panel) and of the mass segregation ratios $\Lambda_{10}^J$ and $\Gamma_{10}^J$ (right-hand panel). The vertical dotted lines show the division between hierarchical and centrally concentrated clusters at $\mathcal{Q} = 0.8$ and the division between non-mass-segregated and mass-segregated clusters at $\Lambda_{10}^J, \Gamma_{10}^J = 1$, respectively.

**Table 1.** Parameters of the selected clusters from the MWSC catalogue (1) ID in the MWSC catalogue (2) cluster name (3) R. A. (2000) (4) declination (2000) (5) Galactic longitude (6) Galactic latitude (7) distance from the Sun (8) log (age) (9) core radius (in pc) (10) tidal radius (in pc) (11) number of stars (12) distance to the galactic centre (13) distance to the galactic plane (14) $Q$ parameter (15) $\Lambda_{10}^J$ (Allison et al. 2009a) (16) $\Gamma_{10}^J$ (Olczak, Spurzem & Henning 2011). The complete list for the 1276 clusters is available in the online version of the paper.

| ID | Name | $\alpha$ (deg) | $\delta$ (deg) | $l$ (deg) | $b$ (deg) | $D$ (pc) | $\log(\tau_{cl})$ (yr) | $r_c$ (pc) | $r_t$ (pc) | $N_*$ | $d_{GC}$ (pc) | $z$ (pc) | $Q$ | $\Lambda_{10}^J$ | $\Gamma_{10}^J$ |
|----|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| 2 | NGC 7801 | 0.082 | 50.727 | 114.717 | −11.331 | 1953 | 9.255 | 0.61 | 9.93 | 64 | 9301 | −384 | 0.747 | 1.360 | 1.517 |
| 5 | Berkeley 59 | 0.559 | 67.425 | 118.219 | 5.001 | 1000 | 6.100 | 0.55 | 6.51 | 88 | 8971 | 87 | 0.834 | 2.167 | 2.201 |
| 6 | Cep OB4 | 0.735 | 67.500 | 118.299 | 5.062 | 850 | 6.100 | 9.540 | 13.56 | 653 | 8901 | 75 | 0.758 | 1.308 | 1.835 |

possess a value of $\mathcal{Q} > 0.8$, i.e. are centrally concentrated. The majority of clusters (72.1 per cent) lie in the range $0.7 < \mathcal{Q} < 0.8$, showing neither central concentration nor significant substructure. This is also the range expected for a random distribution of stars. Fig. 2 (right-hand panel) also displays the distribution of the MSRs $\Lambda_{10}^J$ and $\Gamma_{10}^J$ for the sample. The distributions of $\Lambda_{10}^J$ and $\Gamma_{10}^J$ are nearly identical with the distribution of $\Gamma_{10}^J$ being slightly broader. Given this result, we use $\Lambda_{10}^J$ as a description of the MSR in the remaining sections of the paper. The values of $\Lambda_{10}^J$ lie in the range $0.69 < \Lambda_{10}^J < 4.65$ with an arithmetic mean value and standard deviation of $\bar{\Lambda}_{10}^J = 1.23 \pm 0.37$. Only 180 clusters (14.1 per cent) have values of $\Lambda_{10}^J > 1.5$ and can be considered as being significantly mass segregated. Table 1 lists the parameters of the selected 1276 clusters along with their respective values of $\mathcal{Q}$, $\Lambda_{10}^J$ and $\Gamma_{10}^J$ derived in this work.

### 4.1 Correlation of cluster structure and mass segregation with cluster age

The structure of clusters changes, from the onset of their formation and thorough their subsequent dynamical evolution (e.g. Schmeja & Klessen 2006; Schmeja et al. 2008a; Parker 2014). In gravitationally bound clusters, self-gravity leads to a centrally condensed configuration, while gravitationally unbound clusters will approach nearly homogeneous distributions (with $\mathcal{Q} \approx 0.8$). It may take several crossing times to reach an equilibrium state (Goodwin & Whitworth 2004). Simulations (Schmeja & Klessen 2006; Moeckel & Bate 2010; Parker & Alves de Oliveira 2017) and ob-

servations (Schmeja, Kumar & Ferreira 2008b) indicate an increase of $\mathcal{Q}$ during the first few Myr of a forming cluster. In their simulations, Parker et al. (2014) found this behaviour of $\mathcal{Q}$ only for subvirial star-forming regions, while in supervirial regions $\mathcal{Q}$ stays at a constant low level. Fig. 3 (top panel) displays the $\mathcal{Q}$ values of our sample as a function of cluster age, $\tau_{cl}$. No correlation between $\mathcal{Q}$ and $\tau_{cl}$ is observed. Sánchez & Alfaro (2009) studied a small sample of 16 open clusters spanning a wide range of ages $(6.9 < \log(\tau_{cl}) < 9.6)$, and determined their $\mathcal{Q}$ values (red triangles in Fig. 3). The latter authors argued for a weak correlation of $\mathcal{Q}$ with age, but their conclusion is not substantiated by our findings. They also argued that a correlation exists between $\mathcal{Q}$ and the ratio of the cluster age divided by the tidal radius which is proportional to the age of the cluster expressed in units of the crossing time. They find the relation $\mathcal{Q} = (0.07 \pm 0.03)\log(\tau_{cl}/r_t) + (0.35 \pm 0.21)$, where $r_t$ is the tidal radius. We seek the same correlations between $\mathcal{Q}$ and $\tau_{cl}/r_t$ in our sample. The result is displayed in Fig. 3 (bottom panel) along with the data points of Sánchez & Alfaro (2009). Our data do not suggest the existence of a correlation between $\mathcal{Q}$ and $(\tau_{cl}/r_t)$. In fact, most of the $\mathcal{Q}$ values for the clusters of the MWSC lie far below the correlation suggested by Sánchez & Alfaro (2009). We attribute this discrepancy to the different samples and to the small number of clusters studied by Sánchez & Alfaro (2009). We also note a large difference in the $\mathcal{Q}$ parameter (up to $\Delta\mathcal{Q} \approx 0.25$) for a few clusters (e.g. for the cluster MWSC 3008, we find a value of $\mathcal{Q} = 0.77$, whereas Sánchez & Alfaro found $\mathcal{Q} = 1.02$. A possible interpretation of the absence of a correlation between $\mathcal{Q}$ and cluster age (or between $\mathcal{Q}$ and $\tau_{cl}/r_t$) implies that even though if it is
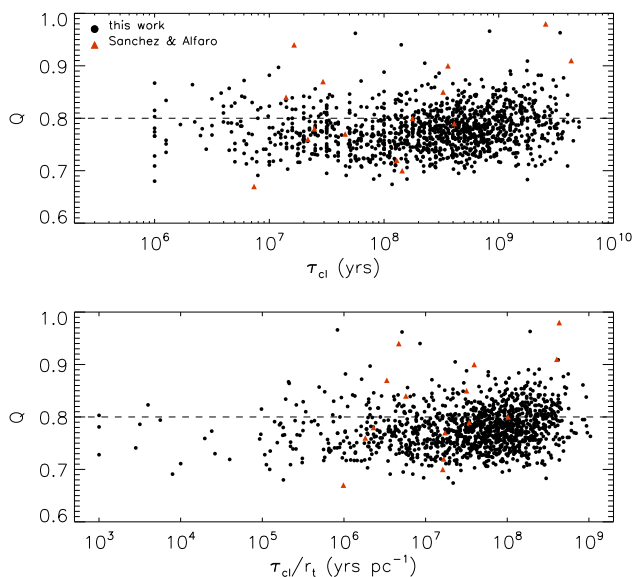
**Figure 3.** The top panel displays $\mathcal{Q}$ as a function of cluster age, $\tau_{cl}$. The horizontal dotted line at $\mathcal{Q} = 0.8$ indicates the division between hierarchical and centrally concentrated clusters. The lower panel displays $\mathcal{Q}$ plotted as a function of the ratio of the cluster age to its tidal radius. The purple symbols show the values found by Sánchez & Alfaro (2009).

likely that $\mathcal{Q}$ increases with time for individual clusters, at least in the early formation period, clusters start from having different $\mathcal{Q}$ values and follow a distinct individual evolution, such that a general correlation for all clusters is not to be expected. We should also point out that if we were missing stars that are located in the outskirts of the clusters (i.e. outliers), this would have the effect of artificially decreasing the $\mathcal{Q}$ parameter. The effect of missing outliers is difficult to quantify, because obviously this effect may depend on their numbers and spatial distributions. An example of this effect for the young clusters IC 348 and NGC 1333 is demonstrated in Parker & Alves de Oliveira (2017). The $\mathcal{Q}$ parameter for these two clusters decreases by 0.1–0.15 when the outer regions are omitted in the calculations.

Tadross (2014) found a weak correlation between the age of a cluster and its diameter and Schilbach et al. (2006) found a dependence of cluster radius on age. However, they attributed this dependence to the effects of mass segregation which are ubiquitous for clusters older than 30 Myr in their sample. Fig. 4 displays the core radius, $r_c$ (top panel) and the tidal radius, $r_t$, (bottom panel) plotted as a function of the cluster age, $\tau_{cl}$. There are no visible correlations between $r_c$ or $r_t$ with the cluster age.

A correlation of mass segregation levels with age may be expected from the dynamical evolution of clusters. However, this effect may be overshadowed by the existence of different levels of primordial mass segregation in the clusters (Dib, Kim & Shadmehri 2007a; Dib et al. 2010a). From a dynamical point of view, mass segregation can occur on short time-scales of a few Myr, leading to a rapid rise of $\Lambda_{MSR}$ (e.g. Allison et al. 2009b). A similar trend for $\Lambda_{MSR}$ has been observed by Parker et al. (2014). In their simulations, supervirial regions show no sign of mass segregation, i.e. $\Lambda_{MSR}$ stays at unity for the entire time of the simulation. On the other hand, subvirial regions show a wide variety in the evolution of $\Lambda_{MSR}$. Usually, $\Lambda_{MSR}$ increases over the first few Myr due to dynamical mass segregation with values up to $\Lambda_{MSR} \approx 10$, after which, $\Lambda_{MSR}$ can evolve in many different ways. In some models, $\Lambda_{MSR}$ remains at high values, while in others it drops again to $\Lambda_{MSR} \approx 1$. On the observational
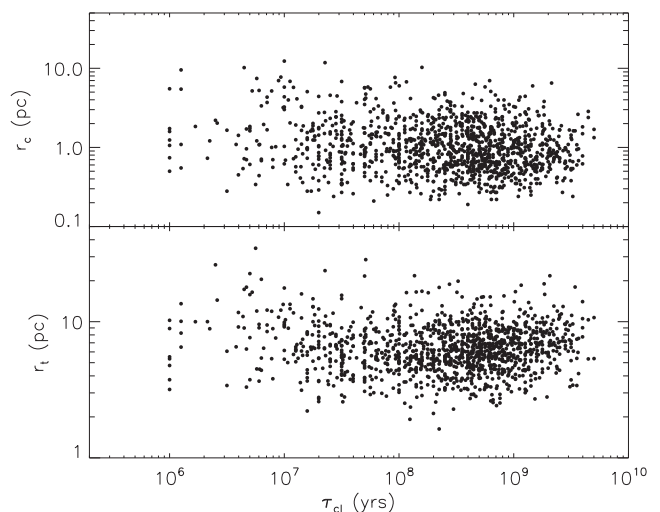


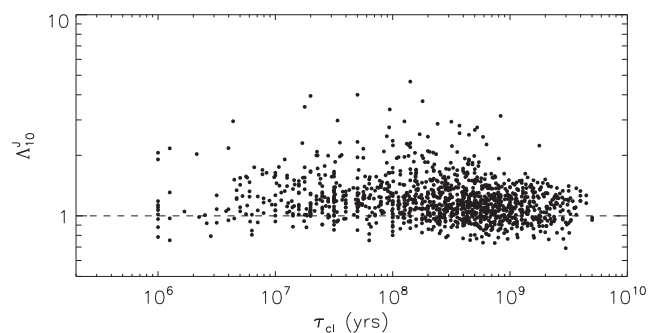**Figure 4.** Core radius (top panel) and tidal radius (bottom panel) as a function of cluster age.



**Figure 5.** Mass segregation ratio $\Lambda_{10}^{J}$ as a function of cluster age.

side, Bukowiecki et al. (2012) used a sample of 599 open clusters selected from the *Two Micron All Sky Survey* (2MASS) and argued that there is a tendency of mass segregation to increase with age. Fig. 5 displays the values of $\Lambda_{10}^{J}$ plotted as a function of cluster age for the sample of 1276 clusters used in this study. No correlation is visible between $\Lambda_{10}^{J}$ and age.

### 4.2 Correlation of cluster structure and mass segregation with galactic position

As the environment may have an influence on the structure of the clusters, we investigate the existence of potential correlations of the clusters parameters with their position in the Galaxy, characterized by the Galactocentric distance $d_{GC}$, the distance from the Galactic plane $|z|$, and the location of the clusters in or outside of the spiral arms. Tadross (2014) found a slight correlation between Galactocentric radius and $|z|$ with the size of the clusters. Froebrich et al. (2010) noted that more extended clusters are found more often at large Galactocentric distances as well as at larger $|z|$. Schilbach et al. (2006) found a systematic increase of cluster size with $|z|$, which becomes significant for clusters older than $\log(\tau_{cl}) = 8.35$.

In our sample, $\mathcal{Q}$ does not show any correlation with $d_{GC}$ or $|z|$ (see Figs 6 and 7). Neither do we find a correlation of the core or tidal radii with $d_{GC}$ (Fig. 8). While no correlation is observed between $r_c$ and $|z|$ (Fig. 9, top panel), we do however, find a correlation of $r_t$ with $|z|$ (Fig. 9, bottom panel). The correlation is given by $\log(r_t) = 0.076(\pm 0.008)\log|z| + 0.89(\pm 0.01)$ (with a
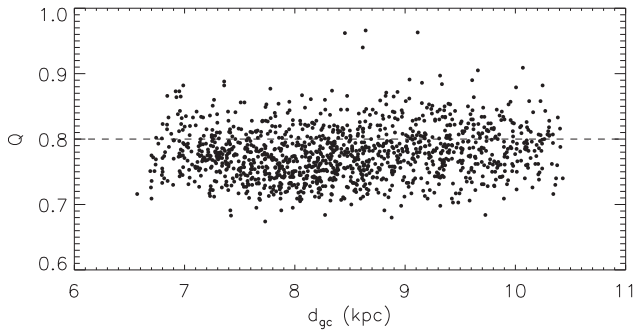
**Figure 6.** $\mathcal{Q}$ as a function of Galactocentric distance. The Sun is assumed to lie at $d_{\mathrm{GC}_\odot} = 8.5\,\mathrm{kpc}$.
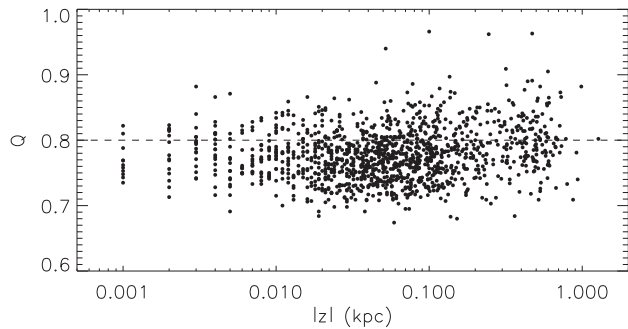


**Figure 7.** $\mathcal{Q}$ as a function of the distance $|z|$ from the Galactic plane.
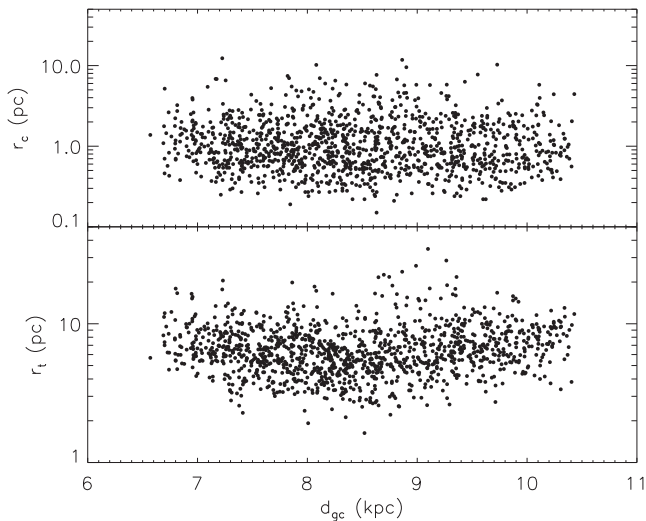


**Figure 8.** Core radius (top panel) and tidal radius (bottom panel) as a function of Galactocentric distance $d_{\mathrm{GC}}$.
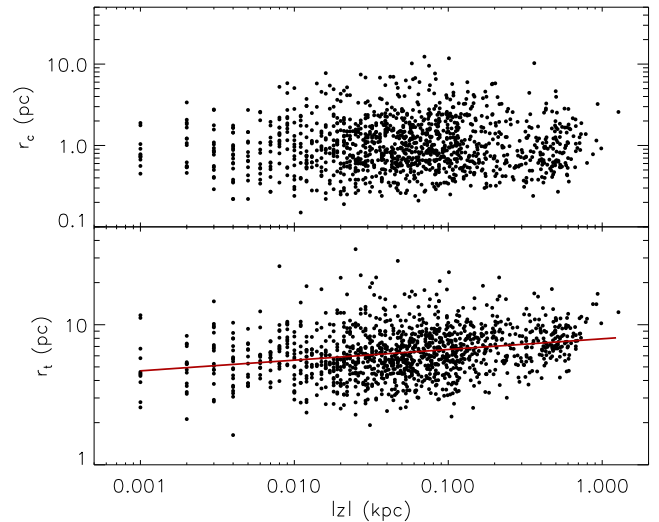


**Figure 9.** Core radius (top panel) and tidal radius (bottom panel) as a function of $|z|$. The red line shows a linear fit to the data. The parameters of the fit are reported in the text.
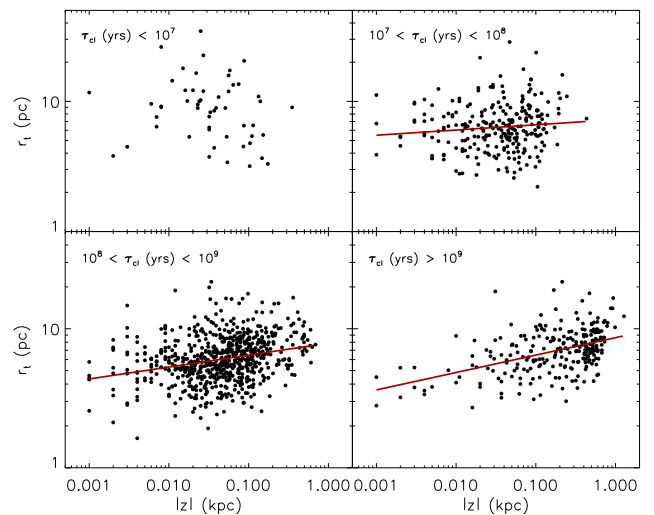


**Figure 10.** $|z|$ versus tidal radius in four age bins. The red lines show linear fits to the data. The parameters of the fits are reported in the text.

Pearson correlation coefficient of $\approx 0.25$). This is in agreement with the findings of Schilbach et al. (2006), Froebrich et al. (2010), Bukowiecki et al. (2011) and Tadross (2014). We analysed the same relation for different age bins (Fig. 10). While there is no obvious correlation in the age bin $\log{(\tau_{\mathrm{cl}})}(\mathrm{yr}) < 7$, we find a correlation between $|z|$ and the tidal radius for ages $7 < \log{(\tau_{\mathrm{cl}})}(\mathrm{yr}) < 8$ ($\log(r_{\mathrm{t}}) = 0.040(\pm 0.023)\log|z| + 0.86(\pm 0.03)$), $8 < \log{(\tau_{\mathrm{cl}})}(\mathrm{yr}) < 9$ ($\log(r_{\mathrm{t}}) = 0.085(\pm 0.011)\log|z| + 0.89 \pm (0.01)$), and $\log{(\tau_{\mathrm{cl}})}(\mathrm{yr}) \geq 9$ ($\log(r_{\mathrm{t}}) = 0.12(\pm 0.015)\log|z| + 0.93(\pm 0.01)$). The Pearson correlation coefficients are 0.10, 0.27 and 0.46 for the age bins $[10^7$–$10^8]$ yr, $[10^8, 10^9]$ yr and $>10^9$ yr, respectively,

indicating an increase in the correlation between $r_{\mathrm{t}}$ and $|z|$ with increasing age.

$\Lambda_{10}^J$ shows no correlation with $d_{\mathrm{GC}}$ (Fig. 11). Higher values of $\Lambda_{10}^J$ ($\gtrsim 2$) are only found for cluster at smaller $|z|$ ($|z| \lesssim 250$ pc (Fig. 12). This is likely to be only a statistical effect, since there are many more clusters close to the Galactic plane than at high $|z|$. However, the mean value of $\Lambda_{10}^J$ for consecutive bins containing each 50 clusters does not change significantly with $|z|$ (not shown). We also test whether the cluster parameters show any dependence with respect to their location inside/outside of the Galactic spiral arms (Fig. 13). The positions of the spiral arms (Perseus and Sagittarius arms) are taken from Vallée (2014). The average $\mathcal{Q}$ parameter is exactly the same, $\bar{\mathcal{Q}} = 0.78 \pm 0.04$, for the clusters inside and outside the arms. The $\Lambda_{10}^J$ values are on average slightly higher outside the spiral arms ($\bar{\Lambda}_{10}^J = 1.25 \pm 0.38$) than inside ($\bar{\Lambda}_{10}^J = 1.18 \pm 0.33$), but
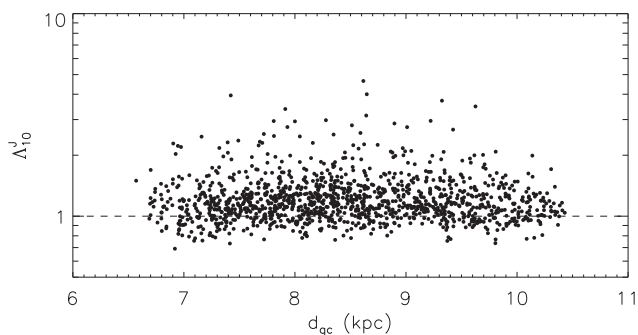
**Figure 11.** MSR parameter ($\Lambda_{10}^J$) versus Galactocentric distance. The Sun is assumed to lie at $d_{GC_\odot} = 8.5$ kpc).
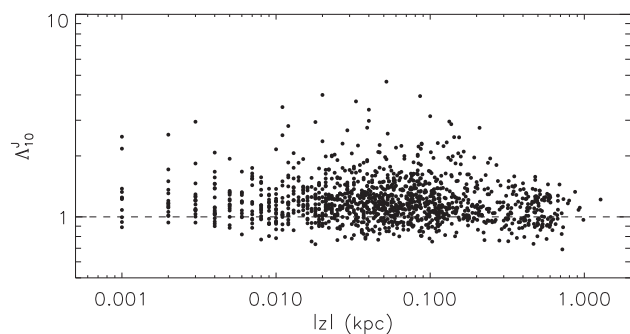


**Figure 12.** MSR parameter ($\Lambda_{10}^J$) versus distance $|z|$ from the Galactic plane.

these values are compatible within the $1\sigma$ uncertainty. Of the clusters outside the arms, $\approx$16 per cent of them show significant mass segregation ($\Lambda_{10}^J > 1.5$), while this is only the case for $\approx$7 per cent of all the clusters in the spiral arms. The average tidal radii also do not show a significant difference inside ($\bar{r}_t = 7.36 \pm 2.81$ pc) and outside ($\bar{r}_t = 6.73 \pm 3.23$ pc) the spiral arms.

### 4.3 Implications for star formation in the local volume

It is possible to gain insight into the physical conditions prevalent at the time the young clusters in the MWSC formed by comparing their structure and mass segregation levels with those derived from numerical simulations. Magnetohydrodynamical simulations of star cluster formation have the advantage of taking into account the effects of the gas on the structure of the nascent clusters and can follow the evolution of the cluster properties during its build up. However, star cluster formation simulations are computationally expensive and can therefore sample only a limited subset of the the initial conditions of the parent protocluster clumps (e.g. Dib et al. 2007b; Dib et al. 2008b; Dib et al. 2010b; Padoan et al. 2014). An appealing alternative is to use *N*-body simulations which can follow the evolution of a cluster over an extended period of time. These simulations can start either from a gas free cluster and can use as initial conditions of the stars positions and kinematics the input of star formation models or be constructed with more controlled and idealized initial conditions. Parker (2014), Parker et al. (2014), Parker, Dale & Ercolano (2015) and Parker & Alves de Oliveira (2017) presented a number of such simulations. Parker et al. (2014) simulated the dynamical evolution of initially hierarchically struc-
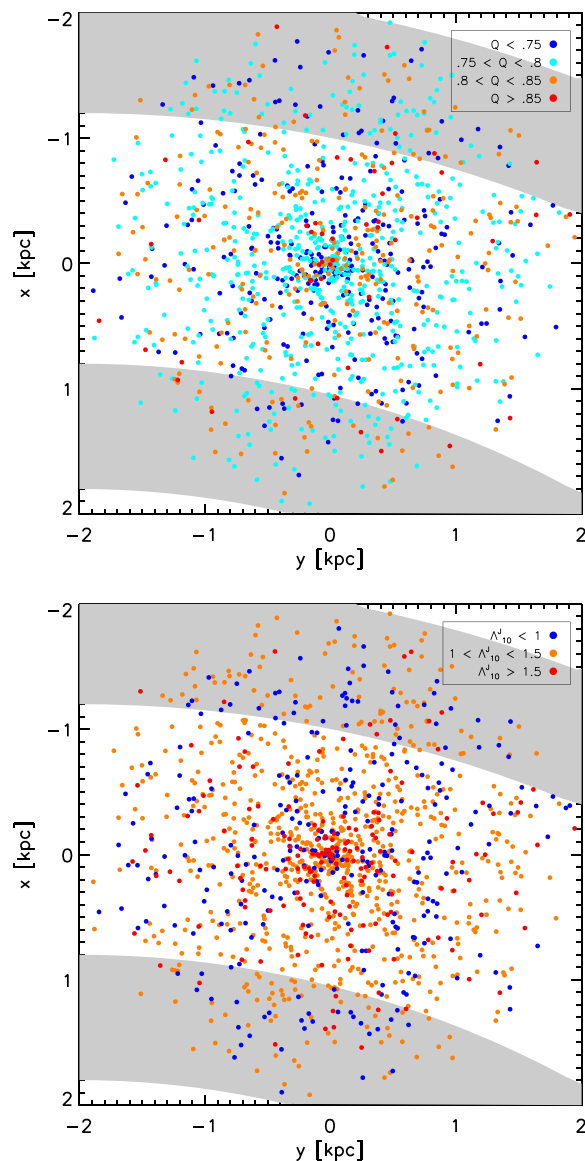


**Figure 13.** Spatial distribution of the clusters centred around the Galactic position of the Sun. The grey areas are the inner (Sagittarius) and outer (Perseus) spiral arms around the Sun. The clusters are coded by the value of their structure parameter $\mathcal{Q}$ (top panel), and their mass segregation ratio $\Lambda_{10}^J$ (bottom panel). There are no obvious correlations between the $\mathcal{Q}$ or $\Lambda_{10}^J$ values of the clusters with their position in the arms or in the interarm regions.

tured clusters over the first 10 Myr starting from different initial conditions, and followed the evolution of $\mathcal{Q}$ and $\Lambda_{MSR}$. These *N*-body simulations explored the effect of starting from subvirial or supervirial conditions and the effect of a different initial fractal dimension of the clusters. Parker et al. (2014) found that in subvirial regions ($\alpha_{vir}$ in the range [0.3–0.5]), substructure is erased rapidly and $\mathcal{Q}$ rises to values >1 within 1 Myr. On the other hand, in supervirial regions ($\alpha_{vir} = 1.5$), substructure is preserved and a constant low $\mathcal{Q}$ characteristic of cluster with substructure in maintained. Cases of clusters that have $\alpha_{vir} \approx 0.5$–1.5 are not yet fully explored and could display an intermediate behaviour, i.e. a moderate rise in the value of $\mathcal{Q}$ followed by a saturation at that level.
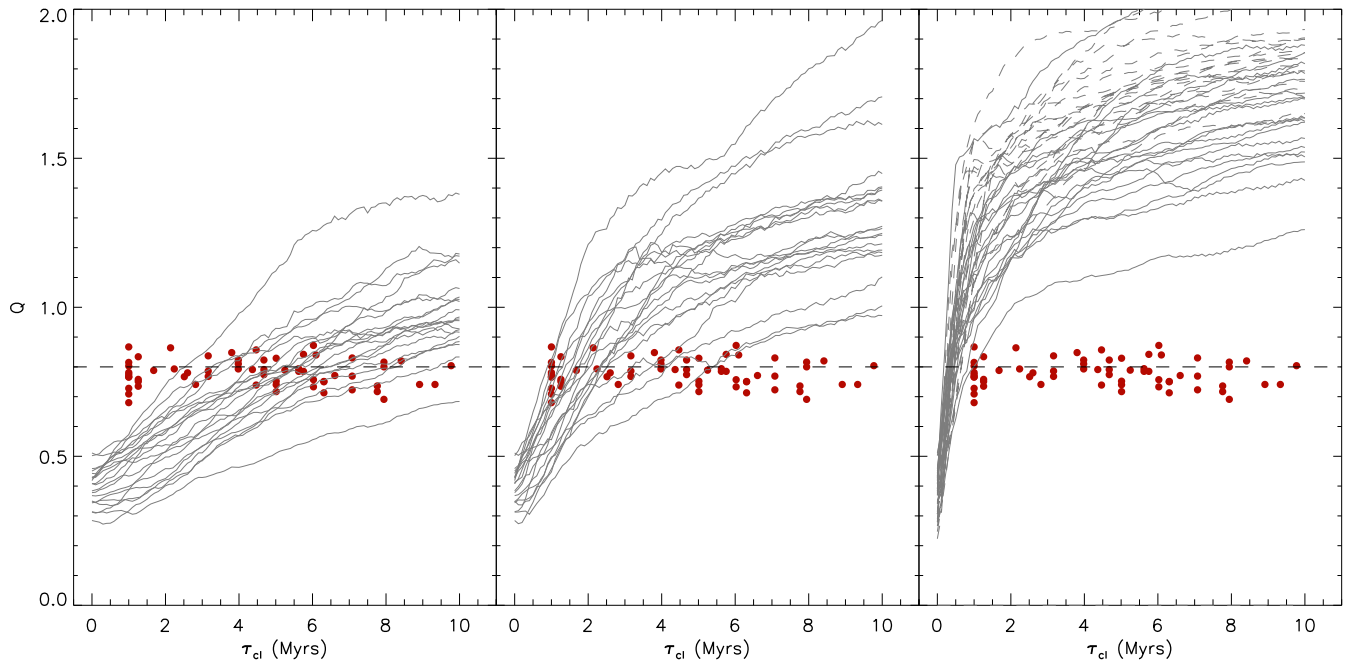
**Figure 14.** Comparison of the $\mathcal{Q}$ parameter for the young clusters in the MWSC to the evolution of $\mathcal{Q}$ in *N*-body models of stellar clusters. The *N*-body models include ones for clusters with $N_* = 425$ in which the initial local volume densities in the clusters falls in the range $10–60$ $M_{\odot}$ pc$^{-3}$ (full line, left-hand panel), $100–500$ $M_{\odot}$ pc$^{-3}$ (full line, middle panel) and $5\,000–10\,000$ $M_{\odot}$ pc$^{-3}$ (full line, right-hand panel) and with $N_* = 1500$ with initial local volume densities that are in the range $2 \times 10^3–2 \times 10^4$ $M_{\odot}$ pc$^{-3}$ (dashed line, right-hand panel).

Parker ([2014](#)) explored the effect of changing the initial cluster density[2] ($\rho_*$) of the star-forming region on the evolution of the $\mathcal{Q}$ parameter (fig. 3 in his paper). High-density regions ($\rho_* \approx 10^4$ $M_{\odot}$ pc$^{-3}$) lose substructure within 1 Myr and reach values $\mathcal{Q} > 1$ after 10 Myr (i.e. at the end of the simulations), medium-density regions ($\rho_* \approx 10^2$ $M_{\odot}$ pc$^{-3}$) lose substructure within 3–5 Myr and end up with $0.7 \lesssim \mathcal{Q} \lesssim 1.2$, whereas low-density regions ($\rho_* \approx 10$ $M_{\odot}$ pc$^{-3}$) retain substructure for the entire time and stay at values $0.4 \lesssim \mathcal{Q} \lesssim 0.7$.

In this work, we compare the $\mathcal{Q}$ values derived for the population of young clusters in the MWSC to a set of *N*-body models presented in Parker & Alves de Oliveira ([2017](#)). These models follow the time evolution of the $\mathcal{Q}$ parameter in clusters with $N_* = 425$ stars[3] and have an initial virial ratio of $\alpha_{vir} = 0.3$. The clusters are initially substructured and have a fractal dimension $D = 1.6$. The models include cases with initial cluster radii of 0.5, 1.5 and 3 pc. The stellar masses of the 425 stars are drawn from the observed IMF of IC 348 (Luhman, Esplin & Loutrel [2016](#))[4] and the initial positions of the stars within the cluster are randomly assigned and

no correlation between the masses of the stars and their positions within the cluster is imposed. For the three chosen values of the cluster initial radii, this leads to local volume densities in the clusters that fall in the range $\rho_* \approx 10–60$ $M_{\odot}$ pc$^{-3}$ when the radius is 3 pc, $100–500$ $M_{\odot}$ pc$^{-3}$ for a cluster radius of 1.5 pc and $5\,000–10\,000$ $M_{\odot}$ pc$^{-3}$ for a cluster radius of 0.5 pc. We also include a set of simulations which have a higher number of stars $N_* = 1500$, whose masses are randomly drawn from a Galactic field like IMF (Maschberger [2013](#))[5] and whose spatial positions are randomly assigned within the fractal structure. This additional set of simulations are performed with $\alpha_{vir} = 0.3$, $D = 1.6$ and an initial cluster radius of 1 pc, corresponding to initial local volume densities that fall in the range $\rho_* \approx 2 \times 10^3–10^4$ $M_{\odot}$ pc$^{-3}$. The simulations do not have stellar evolution switched on, nor do they feature an external Galactic tidal field. Simulations used in Parker et al. ([2016](#)) find no appreciable difference between the long-term dynamical evolution of clusters with or without stellar evolution.

In Fig. [14](#), we compare the $\mathcal{Q}$ values in the observations and the *N*-body models. The $\mathcal{Q}$ values for the young MWSC clusters which fall in the range of $0.7 \lesssim \mathcal{Q} \lesssim 0.9$ are better matched with the models that start with initial stellar surface densities of $\rho_* = 10–100$ $M_{\odot}$ pc$^{-3}$ (left-hand panel and middle panel in Fig. [14](#)). Models that possess higher stellar surface densities (right-hand panel) fail to reproduce this set of observations. This implies that the young clusters in the MWSC were formed with roughly similar initial stellar volume densities and probably out of protoclusters clumps with similar structural and dynamical properties.

---

[2] This is a working assumption as stars will not form simultaneously. Instead this concept of initial stellar density could be understood as an initial peak stellar density.

[3] As the $\mathcal{Q}$ parameter displays a slight dependence on the number of stars (e.g. Lomax, Whitworth & Cartwright [2011](#)), ideally, each cluster in the observational sample should be compared to simulations that are performed using the same number of stars. However, in practice, the scatter between *N*-body simulations constructed with the same set of parameters but with different random seed number for the spatial and kinematic distributions of the stars is larger than the effect of the number of stars.

[4] The Luhman, Esplin & Loutrel ([2016](#)) paper lists the photometric data of the 425 stars in IC 348. The stellar masses have been derived in Parker & Alves de Oliveira ([2017](#)) using the Luhman, Esplin & Loutrel ([2016](#)) data following a procedure described in detail in section 2 of their paper.

[5] The functional form for the IMF proposed by Maschberger ([2013](#)) is an order-3 Logistic function which is described by three parameters, namely the slope in the low-mass regime, the slope in the intermediate to high mass regime and a parameter that ensure the continuity across these two mass regimes.

Gregorio-Hetem et al. ([2015](#)) performed a similar comparison for their sample of 25 stellar associations with earlier *N*-body models performed by Parker & Dale ([2013](#)). They found that their data points are better reproduced with models that have initial volume densities of $\approx 5$ $M_\odot$ pc$^{-3}$. The presence of an external Galactic tidal field can expedite the dissolution of star clusters. However, the absence of a tidal field in our simulations will not affect our interpretation that the initial stellar densities were no higher than 10–100 $M_\odot$ pc$^{-3}$ for two reasons. First, initially high stellar densities ($>1000$ $M_\odot$ pc$^{-3}$) would produce high values of $\mathcal{Q}$ and $\Lambda_{\rm MSR}$, which we do not see in our sample of observed clusters. These high densities would lead to significant dynamical interactions during the early stages of the clusters' lives, where two-body relaxation would dominate over the effects of the tidal field. Secondly, Parker et al. ([2016](#)) have recently shown that clusters do not approach energy equipartition, where the lowest mass objects would be ejected to the outskirts of the cluster. Therefore, we would not expect the influence of an external tidal field to preferentially remove low-mass objects from the cluster and bias the measurement of $\Lambda_{\rm MSR}$. We also note that even in this unlikely scenario, the $\mathcal{Q}$ parameter would be unaffected, as it is independent of stellar mass.

### 4.4 Discussion

As discussed above, numerical simulations show that the $\mathcal{Q}$ parameter can rise quickly in a star-forming cluster (e.g. Schmeja & Klessen [2006](#); Moeckel & Bate [2010](#), Parker et al. [2014](#); Parker [2014](#)). In order to better understand the transition from the embedded phase into the gas-free phase in terms of the clusters structure, we analyse an additional sample of embedded clusters, taken from a study performed with the *Spitzer Space Telescope* located within 1 kpc of the Sun (Gutermuth et al. [2009](#)). Computing the $\mathcal{Q}$ values for this sample gives a mean value of $\mathcal{Q} = 0.86 \pm 0.08$ for the 20 clusters with 40 or more stars. This value is higher than the mean value for our open clusters ($\mathcal{Q} = 0.78 \pm 0.04$). A similar value of $\mathcal{Q} = 0.87 \pm 0.07$ was found by (Jaehnig, Da Rio & Tan [2015](#)) for 22 young (ages $\approx 1$–3 Myr) clusters in Galactic star-forming regions. However, considering only the MWSC clusters with ages $<5$ Myr (the maximum time for clusters expected to be embedded) results in a mean value of $\mathcal{Q} = 0.78 \pm 0.04$, exactly the same as for the entire sample. The discrepancy may be attributed to the different small samples, as well as to the problematic definition of embedded clusters (e.g. Kroupa [2011](#)). So the notion that young clusters may have, on average, higher $\mathcal{Q}$ values, as noted for a different sample in (Schmeja et al. [2008a](#)), may not hold.

We check whether possible biases are induced by the cluster sample which is affected by incompleteness (Dib, Schmeja & Hony [2017](#)). We perform simple comparisons with the sample used in previous sections by selecting clusters that are either at distances $d < 1$ kpc (323 clusters) or $d < 0.5$ kpc (78 clusters) from the Sun, or clusters with more than 500 members (78 clusters). The results of these tests show that this selection criteria do not change any of our results. The chosen membership probability is also not critical to the results. When varying the required membership probability between $P > 30$ per cent and $P > 75$ per cent, $\mathcal{Q}$ changes on average by less than 5 per cent and $\Lambda_{10}^J$ by 17 per cent, while the number of stars changes on average by $\approx 30$–40 per cent. Likewise using a different value of (i.e. $n_{\rm MST} = 5$ or $n_{\rm MST} = 15$) or different filters for the MSR comparison (see Fig. [1](#) for two examples) does not result in a different behaviour (for more details on the effects of changing these quantities, see the Appendix).

## 5 SUMMARY

We analysed 1276 Galactic open clusters with uniform astrophysical data from the MWSC catalogue and computed their structure parameter $\mathcal{Q}$ and their mass segregation ratio $\Lambda_{\rm MSR}$. Our main findings can be summarized as follows:

(i) Most clusters possess values of the $\mathcal{Q}$ parameter that fall in the range $0.7 < \mathcal{Q} < 0.8$, indicating neither central concentration nor significant substructure. Only $\approx 27$ per cent can be considered centrally concentrated ($\mathcal{Q} > 0.8$).

(ii) Most clusters show mass segregation values around $\Lambda_{\rm MSR} \approx 1$, indicating a similar distribution of massive and low-mass stars. The distribution function of $\Lambda_{\rm MSR}$ is positively skewed and $\approx 14$ per cent of the clusters show signs of significant mass segregation ($\Lambda_{\rm MSR} > 1.5$).

(iii) No correlation is found between $\mathcal{Q}$, $\Lambda_{\rm MSR}$, or the cluster radius with the cluster age. Some of the correlations claimed by other authors using much smaller cluster samples could not be confirmed.

(iv) No significant correlation is found between $\mathcal{Q}$, $\Lambda_{\rm MSR}$ or cluster radius with the cluster position in the Galaxy i.e. the distance from the Galactic Centre $d_{\rm GC}$, the distance from the Galactic plane $|z|$ or the position in the arm/interam regions. There is a tendency for clusters at larger distances from the Galactic plane (i.e. large $|z|$) to have larger tidal radii, which holds in particular for older clusters.

(v) Embedded and open clusters show on average the same $\mathcal{Q}$ values.

(vi) A comparison of the observed $\mathcal{Q}$ values of the young clusters in the MWSC to a suite of *N*-body numerical simulations of the early evolution of stellar clusters suggests that the clusters found in the MWSC catalogue were formed from sub-virial/virial conditions and with mean local volume densities of $\rho_* \approx 10$–100 $M_\odot$ pc$^{-3}$.

## REFERENCES

Allison R. J., Goodwin S. P., Parker R. J., Portegies Zwart S. F., de Grijs R., Kouwenhoven M. B. N., 2009a, MNRAS, 395, 1449
Allison R. J., Goodwin S. P., Parker R. J., de Grijs R., Portegies Zwart S. F., Kouwenhoven M. B. N., 2009b, ApJ, 700, L99
Allison R. J., Goodwin S. P., Parker R. J., Portegies Zwart S. F., de Grijs R., 2010, MNRAS, 407, 1098
Banerjee S., Kroupa P., 2015, MNRAS, 447, 728
Bonnell I. A., Bate M. R., 2006, MNRAS, 370, 488
Borůvka O., 1926, Práce, moravské přírodovědecké společnosti, 3, 37
Brinkmann N., Banerjee S., Mowtani B., Kroupa P., 2017, A&A, 600, 49
Buckner A. S. M., Froebrich D., 2014, MNRAS, 444, 290

Bukowiecki Ł., Maciejewski G., Konorski P., Strobel A., 2011, Acta Astron. 61, 231

Bukowiecki Ł., Maciejewski G., Konorski P., Niedzielski A., 2012, Acta Astron. 62, 281

Carpenter J. M., 2000, MNRAS, 120, 3139

Cartwright A., Whitworth A. P., 2004, MNRAS, 348, 589

Cartwright A., Whitworth A. P., 2009, MNRAs, 392, 341

Chen L., de Grijs R., Zhao J. L., 2007, AJ, 134, 1368

Dalessandro E., Miocchi P., Carraro G., Jílková L., Moitinho A., 2015, MNRAS, 449, 1811

Dib S., 2011, in Alfaro Navarro E. J., Gallego Calvente A. T., Zapatero Osorio M. R., eds, Stellar Clusters and Associations: A RIA Workshop on Gaia, p. 30

Dib S., 2007, JKAS, 40, 157

Dib S., 2014, MNRAS, 444, 1957

Dib S., Kim J., Shadmehri M., 2007a, MNRAS, 381, L40

Dib S., Kim J., Vázquez-Semadeni E., Burkert A., Shadmehri M., 2007b, ApJ, 661, 262

Dib S., Shadmehri M., Gopinathan M., Kim J., Henning T., 2008a, in Beuther H., Linz H., Henning T., eds, ASP Conf. Ser. Vol 387, Massive Star Formation: Observations Confront Theory. Astron. Soc. Pac., San Francisco, p. 282

Dib S., Brandenburg A., Kim J., Gopinathan M., André P., 2008b, ApJ, 678, 105

Dib S., Shadmehri M., Padoan P., Maheswar G., Ojha D. K., Khajenabi F., 2010a, MNRAS, 405, 401

Dib S., Hennebelle P., Pineda J. E., Csengeri T., Bontemps S., Audit E., Goodman A. A., 2010b, ApJ, 721, 425

Dib S., Piau L., Mohanty S., Braine J., 2011a, MNRAS, 415, 3439

Dib S., Piau L., Mohanty S., Braine J., 2011b, in Alecian G., Belkacem K., Samadi R., Valls-Gabaud D., eds, SF2A-2011, Proc. Annu. Meeting French Soc. Astron. Astrophys., p. 275

Dib S., Gutkin J., Brandner W., Basu S., 2013, MNRAS, 436, 3727

Dib. S., Schmeja S., Hony S., 2017, MNRAS, 464, 1738

Feigelson E. D. et al., 2013, ApJS, 209, 26

Friel E. D., 1995, ARA&A, 33, 381

Froebrich D., Schmeja S., Samuel D., Lucas P. W., 2010, MNRAS, 409, 1281

Fujii M., 2015, PASJ, 67, 59

Gennaro M., Brandner W., Stolte A., Henning, Th., 2011, MNRAS, 412, 2469

Gieles M., Portegies Zwart S. F., Baumgardt H., Athanassoula E., Lamers H. J. G. L. M., Sipior M., Leenarts J., 2006, MNRAS, 371, 793

Goodwin S. P., Whitworth A. P., 2004, A&A, 413, 929

Gouliermis D., Keller S. C., Kontizas M., Kontizas E., Bellas-Velidis I., 2004, A&A, 416, 137

Gouliermis D., Schmeja S., Dolphin A. E., Gennaro M., Tognelli E., Prada Moroni P. G., 2012, ApJ, 748, 64

Gouliermis D., Hony S., Klessen R. S., 2014, MNRAS, 439, 3775

Gower J. C., Ross G. J. S., 1969, J. Roy. Stat. Soc. Ser. C (Appl. Stat.), 18, 54

Gregorio-Hetem J., Hetem A., Santos-Silva T., Fernandes B., 2015, MNRAS, 448, 2504

Guthermuth R. A., Megeath S. T., Myers P. C., Allen L. E., Pipher J. L., Fazio G. G., 2009, ApJS, 184, 18

Habibi M., Stolte A., Brandner W., Hußman B., 2013, A&A, 556, 26

Haghi H., Zonoozi A. H., Kroupa P., Banerjee S., Baumgardt H., 2015, MNRAS, 454, 3872

Hasan P., Hasan S. N., 2011, MNRAS, 413, 2345

Hony S. et al., 2015, MNRAS, 448, 184

Jaehnig K. O., Da Rio N., Tan J. C., 2015, ApJ, 798, 126

Kharchenko N. V., Piskunov A. E., Schilbach E., Röser S., Scholz R.-D., 2012, A&A, 543, 156

Kharchenko N. V., Piskunov A. E., Schilbach E., Röser S., Scholz R.-D., 2013, A&A, 558, 53

Kroupa P., 2011, in Alfaro Navarro E. J., Gallego Calvente A. T., Zapatero Osorio M. R., eds, Stellar Clusters & Associations: A RIA Workshop on Gaia. Granada, Spain, p. 17

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

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

Littlefair S. P., Naylor T., Jeffries R. D., Devey C. R., Vine S., 2003, MNRAS, 345, 1205

Lomax O., Whitworth A. P., Cartwright A., 2011, MNRAS, 412, 627

Luhman K. L., Esplin T. L., Loutrel N. P., 2016, ApJ, 827, 52

Mallick K. K. et al., 2014, MNRAS, 443, 3218

Martinez-Medina L. A., Pichardo B., Peimbert A., Moreno E., 2017, ApJ, 834, 58

Maschberger T., 2013, MNRAS, 429, 1725

McMillan S. L. W., Vesperini E., Portegies Zwart S. F., 2007, ApJ, 655, 45

Moeckel N., Bate M. R., 2010, MNRAS, 404, 721

Olczak C., Spurzem R., Henning T., 2011, A&A, 532, 119

Padoan P., Haugbølle T., Nordlund, Å., 2014, ApJ, 797, 32

Pang X., Grebel E. K., Allison R. J. et al., 2013, 764, 73

Parker R. J., 2014, MNRAS, 445, 4037

Parker R. J., Alves de Oliveira C., 2017, MNRAS, 468, 4340

Parker R. J., Dale J. E., 2013, MNRAS, 432, 986

Parker R. J., Meyer M. R., 2012, MNRAS, 427, 637

Parker R. J., Wright N. J., Goodwin S. P., Meyer M. R., 2014, MNRAS, 438, 620

Parker R. J., Dale J. E., Ercolano B., 2015, 446, 4278

Parker R. J., Goodwin S. P., Wright N. J., Meyer M. R., Quanz S. P., 2016, MNRAS, 459, 119

Pfalzner S., Kaczmarec T., 2013, A&A, 559, 38

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

Röser S., Demleitner M., Schilbach E., 2010, AJ, 139, 2440

Sánchez N., Alfaro E., 2009, ApJ, 696, 2086

Schilbach E., Kharchenko N. V., Piskunov A. E., Röser S., Scholz R.-D., 2006, A&A, 456, 523

Schmeja S., Klessen R. S., 2006, MNRAS, 449, 151

Schmeja S., Kumar M. S. N., Froebrich D., Klessen R. S., 2008a, in Vesperini E., Giersz M., Sills A., eds, Proc. IAU Symp. 246, Dynamical Evolution of Dense Stellar Systems. Cambridge Univ. Press, Cambridge, p. 50

Schmeja S., Kumar M. S. N., Ferreira B., 2008b, MNRAS, 389, 1209

Schmeja S., Kharchenko N. V., Piskunov A. E., Röser S., Schilbach E., Froebrich D., Scholz R.-D., 2014, A&A, 568, 51

Scholz R.-D., Kharchenko N. V., Piskunov A. E., Röser S., Schilbach E., 2015, A&A, 581, 39

Sharma S., Pandey A. K., Ojha D. K., Chen W. P., Ghosh S. K., Bhatt B. C., Maheswar G., Sagar R., 2007, MNRAS, 380, 1141

Sheikhi N., Hasheminia M., Khalaj P., Haghi H., Zonoozi A. H., Baumgardt H., 2016, MNRAS, 457, 1028

Spitzer L., Jr, Harm R., 1958, ApJ, 127, 544

Stolte A., Brandner W., Brandl B., Zinnecker H., 2006, ApJ, 132, 253

Tadross A. L., 2014, NRIAG J. Astron. Geophys., 3, 88

Tadross A. L., Werner P., Osman A., Marie M., 2002, New A., 7, 553

Vallée J. P., 2014, ApJS, 215, 1

Zhai M., Abt H., Zhao G., Li C., 2017, AJ, 153, 57

## SUPPORTING INFORMATION

Supplementary data are available at *MNRAS* online.

**Table1.dat**

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

## APPENDIX: SENSITIVITY OF THE RESULTS TO THE CHOICE OF $n_{MST}$ AND CLUSTER MEMBERSHIP PROBABILITY

Here, we compare the distributions of $\mathcal{Q}$ and $\Lambda_{10}^{J}$ obtained with different value of the number of most massive stars considered for mass segregation, $n_{MST}$, and of the stellar membership probability,
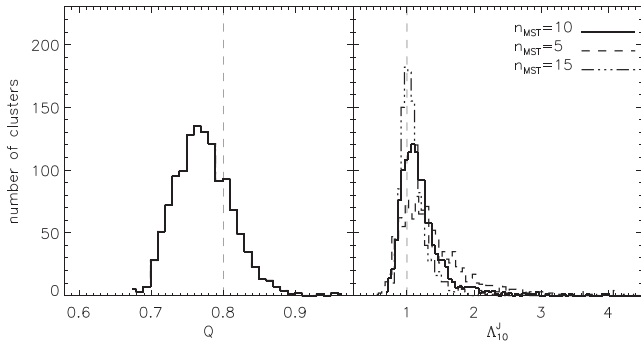
**Figure A1.** The dependence of the distribution of the structure parameter $\mathcal{Q}$ (left-hand panel) and of the mass segregation ratios $\Lambda_{10}^J$ (right-hand panel) on the choice of $n_{\mathrm{MST}}$. The cases shown are for $n_{\mathrm{MST}} = 5$, $n_{\mathrm{MST}} = 10$ (fiducial case) and $n_{\mathrm{MST}} = 15$. The vertical dotted lines show the division between hierarchical and centrally concentrated clusters at $\mathcal{Q} = 0.8$ and the division between non-mass-segregated and mass-segregated clusters at $\Lambda_{10}^J = 1$, respectively.
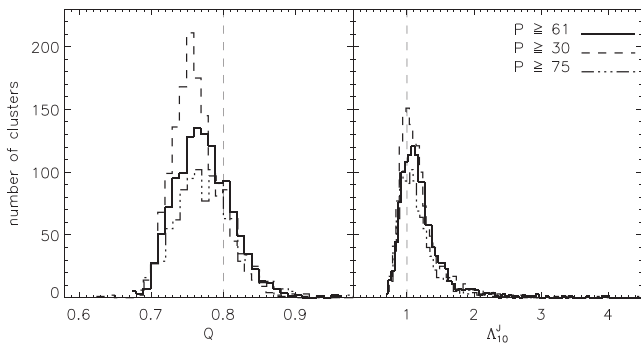


**Figure A2.** The dependence of the distribution of the structure parameter $\mathcal{Q}$ (left-hand panel) and of the mass segregation ratios $\Lambda_{10}^J$ (right-hand panel) on the choice of the stellar membership probability, $P$. The fiducial case corresponds to the case with $P \geq 61$ per cent (full line, 1276 clusters). A higher/smaller value of the threshold probability reduces/increases the number of clusters in the sample. For $P \geq 30$ per cent and $P \geq 75$ per cent, the number of clusters is 1464 and 998, respectively. The vertical dotted lines show the division between hierarchical and centrally concentrated clusters at $\mathcal{Q} = 0.8$ and the division between non-mass-segregated and mass-segregated clusters at $\Lambda_{10}^J = 1$, respectively.

$P$, to those of the fiducial case where $n_{\mathrm{MST}} = 10$ and $P \geq 61$ per cent. Fig. A1 displays the distribution of $\mathcal{Q}$ (left-hand panel) and of $\Lambda_{10}^J$ for $n_{\mathrm{MST}} = 5$, 10 and 15, with $P$ being fixed at $P \geq 61$ per cent. The total number of clusters is the same (i.e. 1276 clusters). The distributions of $\mathcal{Q}$ are identical. For higher $n_{\mathrm{MST}}$ values, the peak of the distribution of $\Lambda_{10}^J$ is shifted towards $\approx 1$. This is not too surprising since for higher value of $n_{\mathrm{MST}}$, the distribution of the $n_{\mathrm{MST}}$ most massive stars becomes more similar to the one of the total stellar population in the cluster. The arithmetic mean values and standard deviations are $[0.778 \pm 0.039, 0.778 \pm 0.039, 0.778 \pm 0.039]$ and $[1.46 \pm 0.73, 1.23 \pm 0.37, 1.14 \pm 0.25]$, whereas the median values of the $\mathcal{Q}$ and $\Lambda_{10}^J$ distributions are $[0.776, 0.776, 0.776]$ and $[1.27, 1.15, 1.08]$ for $n_{\mathrm{MST}} = 5$, 10 and 15, respectively.

For the cases with different membership probability $P$, a higher value of the threshold probability reduces the number of clusters that fulfil our selection criterion of $N_* \geq 40$, and the reverse is true for smaller $P$ values. Fig. A2 displays the distributions of $\mathcal{Q}$ (left-hand panel) and of $\Lambda_{10}^J$ (right-hand panel) for cases with $P \geq 30$ per cent (1464 clusters), $P \geq 61$ per cent (fiducial, 1276 clusters) and $P \geq 75$ per cent (998 clusters). In all three cases here, $n_{\mathrm{MST}} = 10$. The effect of low $P$ is to contaminate the clusters with mostly low-mass field stars. In turn, this shifts the peak of the $\lambda_{10}^J$ distribution towards lower values, and it also causes the structure of the clusters to be less centrally condensed (i.e. smaller $\mathcal{Q}$ values). The arithmetic mean values and standard deviations are $[0.771 \pm 0.078, 0.778 \pm 0.039, 0.780 \pm 0.042]$ and $[1.20 \pm 0.34, 1.23 \pm 0.37, 1.18 \pm 0.29]$, whereas the median values of the $\mathcal{Q}$ and $\Lambda_{10}^J$ distributions are $[0.765, 0.776, 0.777]$ and $[1.12, 1.15, 1.11]$ for $P \geq 30$ per cent, $\geq 61$ per cent and $\geq 75$ per cent, respectively.

This paper has been typeset from a TeX/LaTeX file prepared by the author.