eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Evolution of spatio-kinematic structures in star-forming regions: are Friends of Friends worth knowing?

Richard J. Parker[1]★† and Nicholas J. Wright[2]

[1]*Department of Physics and Astronomy, The University of Sheffield, Hicks Building, Hounsfield Road, Sheffield S3 7RH, UK*
[2]*Astrophysics Group, Keele University, Keele, Staffordshire ST5 5BG, UK*

ABSTRACT

The Friends of Friends algorithm identifies groups of objects with similar spatial and kinematic properties, and has recently been used extensively to quantify the distributions of gas and stars in young star-forming regions. We apply the Friends of Friends algorithm to N-body simulations of the dynamical evolution of subvirial (collapsing) and supervirial (expanding) star-forming regions. We find that the algorithm picks out a wide range of groups (1–25) for statistically identical initial conditions, and cannot distinguish between subvirial and supervirial regions in that we obtain similar mode and median values for the number of groups it identifies. We find no correlation between the number of groups identified initially and either the initial or subsequent spatial and kinematic tracers of the regions' evolution, such as the amount of spatial substructure, dynamical mass segregation, or velocity dispersion. We therefore urge caution in using the Friends of Friends algorithm to quantify the initial conditions of star formation.

Key words: methods: numerical – stars: formation – open clusters and associations: general – kinematics and dynamics.

1 INTRODUCTION

The majority of stars form in regions where the median stellar density exceeds the density in the Galactic field by several orders of magnitude (e.g. Lada & Lada 2003; Bressert et al. 2010). These young stars often follow a hierarchical or substructured spatial distribution (Cartwright & Whitworth 2004; Gutermuth et al. 2009; Gouliermis, Hony & Klessen 2014; Kuhn et al. 2014; Wright et al. 2014; Jaehnig, Da Rio & Tan 2015) and also display correlated velocities (i.e. low-velocity dispersions) and kinematic substructure on local scales (e.g. Larson 1981; Jeffries et al. 2014; Foster et al. 2015; Hacar et al. 2016; Da Rio et al. 2017).

There is also mounting evidence that the gas from which stars form also exhibits significant spatial (Cartwright, Whitworth & Nutter 2006; Henshaw et al. 2016b; Williams et al. 2018) and kinematic (Peretto, André & Belloche 2006; Hacar et al. 2013; Hacar, Tafalla & Alves 2017) substructure, although analysis of simulations that directly follow the conversion of gas to stars suggests that the link between their respective spatial and kinematic properties is highly non-trivial (Parker & Dale 2015; Vázquez-Semadeni, González-Samaniego & Colín 2017; Kuznetsova, Hart-

mann & Ballesteros-Paredes 2018) and may not be a direct one-to-one mapping.

Quantifying how these spatial and kinematic structures form and evolve is crucial in order to understand the typical environment where most stars form, and the implications for planet formation and stability (e.g. Scally & Clarke 2001; Adams et al. 2004; Parker & Quanz 2012; Vincke, Breslau & Pfalzner 2015; Portegies Zwart 2016; Cai et al. 2017) as well as their collective evolution in the context of Galaxy-scale astrophysical processes (e.g. Kereš et al. 2009).

A significant amount of effort has been invested in quantifying both the spatial distributions in young star-forming regions (Larson 1995; Cartwright & Whitworth 2004; Kuhn et al. 2014; Jaffa et al. 2018) and the kinematic distributions (Foster et al. 2015; Wright et al. 2016; Wright & Mamajek 2018). Recently, the Friends of Friends algorithm, originally used to quantify clusters of galaxies (Huchra & Geller 1982), has been used to quantify the initial stages of star formation by simultaneously incorporating both the spatial *and* kinematic information. This is usually realized by using the x- and y-position and the radial velocity measurement of either gas parcels, or individual stars.

The most notable result from these Friends of Friends analyses has been the discovery of 'bundles of fibres' within filaments in star-forming regions (Hacar et al. 2013; Hacar et al. 2017, 2018) as well as significant substructure in both the distribution of gas (Henshaw et al. 2016a) and of the protostars (Hacar et al. 2016; Da

★ E-mail: R.Parker@sheffield.ac.uk
† Royal Society Dorothy Hodgkin Fellow.

Rio et al. [2017]). These results potentially indicate a universality in the spatial and kinematic substructure in star-forming regions, but crucially, the Friends of Friends method has not been extensively tested on either synthetic data (Cartwright & Whitworth [2004]; Lomax, Whitworth & Cartwright [2011]; Parker & Goodwin [2015]; Jaffa et al. [2018]), though a notable recent exception is the work by Clarke et al. ([2018]), or on simulation data.

An exception to the latter is the work by Kuznetsova et al. ([2018]), who show that the kinematic structures are likely influenced by the accretion histories of the groups of stars. However, even star-forming regions with a moderate stellar density ($\sim 100\,\mathrm{M_\odot\,pc^{-3}}$) evolve significantly in the first few Myr after star formation, and an obvious avenue of investigation is the longevity of groups identified by the Friends of Friends algorithm during their subsequent dynamical evolution.

Furthermore, no studies have tested the Friends of Friends algorithm on multiple random realizations of the same distribution. Tests of spatial distribution algorithms (such as the $\mathcal{Q}$-parameter or mass segregation algorithms; Cartwright [2009]; Parker & Goodwin [2015]) are essential to understand the significance of any single observed (or simulated) result.

In this paper we test the Friends of Friends algorithm on *N*-body simulations of the dynamical evolution of star-forming regions. The paper is organized as follows. In Section 2 we describe the Friends of Friends algorithm, and the *N*-body simulations we utilize. We present our results in Section 3 and provide a discussion in Section 4. We conclude in Section 5.

2 METHOD

In this section we describe the set-up of our *N*-body simulations and the method used to define groups based on the Friends of Friends method.

2.1 *N*-body simulations

Observations and simulations of the early stages of star formation both suggest that stars form with a spatially and kinematically substructured distribution. (In part, this has inspired the proliferation of the use of Friends of Friends algorithms in star formation studies (Hacar et al. [2013]; Da Rio et al. [2017]), as star-forming regions are clearly inhomogeneous in terms of their spatial and kinematic properties.) To mimic this substructure, our simulated star-forming regions are set up using the box fractal method described in Goodwin & Whitworth ([2004]) and Cartwright & Whitworth ([2004]). We direct the reader to Goodwin & Whitworth ([2004]), Allison et al. ([2010]), and Parker et al. ([2014]) for full details of this method, but we briefly summarize it here.

A box fractal is constructed by defining a 'parent' in the centre of a cube of side N_{div} (we adopt $N_{\mathrm{div}} = 2$), which spawns N_{div}^3 subcubes. Each of these subcubes contains a first generation 'child' at its centre. The construction of the fractal distribution proceeds by determining which of the children then go on to become parents. The probability that a child becomes a parent is given by N_{div}^{D-3}, where D is the fractal dimension. This process is repeated recursively and the final generation of children becomes stars, which are positioned randomly within the fractal distribution. When the fractal dimension is low, fewer children spawn their own offspring and the resultant fractal distribution contains more substructure.

In two sets of simulations we present in this paper, the fractal dimension is $D = 1.6$, which gives the highest degree of substructure in a three-dimensional distribution. This is in order to facilitate the

detection of multiple groups of stars by the Friends of Friends algorithm. A higher fractal dimension, e.g. $D = 2.0$ or 2.6, would lead to fewer distinct groups of stars in the resultant distribution. However, as we discuss below, the fractal dimension also governs the velocity structure in our box fractal method and may influence the way spatio-kinematic groups disperse, so we ran a further two sets of simulations with the fractal dimension set to $D = 2.0$ and 2.6, respectively.

The velocities of the parents in the fractal are drawn from a Gaussian with a mean zero, and the children inherit the velocities of their parents plus an extra random component, the size of which scales as N_{div}^{D-3} (i.e. in a similar fashion to the spatial distribution) and decreases through each successive generation. This results in a kinematic distribution where stars on local scales have very similar velocities, but on larger scales the velocities can be quite different.

In the box fractal method, on local scales of size L the velocities scale as $v(L) \propto L^{3-D}$. We expect the time-scale for the erasure of substructure to be of the order of $t(L) \sim L/v(L)$, so for a fractal with $D = 1.6$ the time-scale for the erasure of structure is $t(L) \propto L^{-0.4}$, for $D = 2.0$ $t(L) \propto L$, and for $D = 2.6$ $t(L) \propto L^{0.6}$. This implies that structure is erased faster on large scales in the case of $D = 1.6$, but erased faster on small scales in the cases of $D = 2.0$ and 2.6.

Interestingly, the $D = 2.0$ and 2.6 fractals are more consistent with the Larson ([1981]) observed line-width relations,[1] where $v(L) \propto L^{0.38}$ [and therefore $t(L) \propto L^{0.6}$], so we might expect that spatio-kinematic structure in observed star-forming regions would be erased faster on smaller scales.

We scale the velocities of the stars to a virial ratio $\alpha_{\mathrm{vir}} = T/|\Omega|$, where T and $|\Omega|$ are the total kinetic and potential energies, respectively. In one set of simulations $\alpha_{\mathrm{vir}} = 0.3$, i.e. subvirial, where the global motion of the stars is to fall towards the centre of the potential. In the second set of simulations $\alpha_{\mathrm{vir}} = 1.5$, i.e. supervirial, where the global motion is for the star-forming region to expand. However, due to the (local) correlated velocities in the fractal distributions, a significant degree of violent relaxation (Lynden-Bell [1967]) occurs (see Allison et al. [2010]; Parker et al. [2014]; Parker & Wright [2016], for examples of the dynamical evolution of these types of systems). In our simulations, violent relaxation occurs within the substructure, while the global bulk motion of the simulation is either to collapse (in the $\alpha_{\mathrm{vir}} = 0.3$, subvirial case) or rapidly expand (in the $\alpha_{\mathrm{vir}} = 1.5$, supervirial case).

Each simulation contains $N = 1500$ single stars, drawn from the Maschberger ([2013]) formulation of the initial mass function (IMF) with a probability distribution

$$p(m) \propto \left(\frac{m}{\mu}\right)^{-\alpha} \left(1 + \left(\frac{m}{\mu}\right)^{1-\alpha}\right)^{-\beta}, \tag{1}$$

where $\mu = 0.2\,\mathrm{M_\odot}$ is the average stellar mass, $\alpha = 2.3$ is the Salpeter ([1955]) power-law exponent for higher mass stars, and $\beta = 1.4$ describes the slope of the IMF for low-mass objects (which also deviates from the log-normal form; Bastian, Covey & Meyer [2010]). We sample this distribution in the mass range 0.01–50 $\mathrm{M_\odot}$.

The radii of our fractal star-forming regions are set to 1 pc. This radius, and the degree of spatial substructure as set by the fractal dimension, gives high to moderate *local* stellar densities ($\tilde{\rho} \sim 10^4\,\mathrm{M_\odot\,pc^{-3}}$ for $D = 1.6$, $\tilde{\rho} \sim 10^3\,\mathrm{M_\odot\,pc^{-3}}$ for $D = 2.0$, and $\tilde{\rho} \sim 10^2\,\mathrm{M_\odot\,pc^{-3}}$ for $D = 2.6$), which means that the initial

---

[1]Note that recent work by Traficante et al. ([2018]) has shown that massive star-forming clumps deviate quite strongly from the Larson ([1981]) relations.

Table 1. Summary of the variables in the initial conditions of our *N*-body simulations. The columns are the fractal dimension, *D*, the virial ratio, $\alpha_{\rm vir}$, and the local density $\tilde{\rho}$.

Fractal dimension	Virial ratio	Local density
$D = 1.6$	$\alpha_{\rm vir} = 0.3$	$\tilde{\rho} \sim 10^4 \, {\rm M}_\odot \, {\rm pc}^{-3}$
$D = 1.6$	$\alpha_{\rm vir} = 1.5$	$\tilde{\rho} \sim 10^4 \, {\rm M}_\odot \, {\rm pc}^{-3}$
$D = 2.0$	$\alpha_{\rm vir} = 1.5$	$\tilde{\rho} \sim 10^3 \, {\rm M}_\odot \, {\rm pc}^{-3}$
$D = 2.6$	$\alpha_{\rm vir} = 1.5$	$\tilde{\rho} \sim 10^2 \, {\rm M}_\odot \, {\rm pc}^{-3}$

substructure will dynamically evolve. It is unclear if there is a typical initial density for star formation (and if there is one, what it is; Marks & Kroupa 2012; Parker 2014; Parker & Alves de Oliveira 2017), but $\tilde{\rho} \sim 10^3 - 10^4 \, {\rm M}_\odot \, {\rm pc}^{-3}$ is consistent with models of the formation and evolution of the Orion Nebula Cluster (Allison et al. 2010; Allison & Goodwin 2011).

In summary, we evolve four sets of *N*-body simulations, each with the same number of stars and initial radius, but we vary the initial degree of substructure, the initial local density, and the initial virial ratio $\alpha_{\rm vir}$. We use the kira integrator in the Starlab environment (Portegies Zwart et al. 1999, 2001) to evolve the star-forming regions for 10 Myr. We do not include stellar evolution. A summary of the simulation initial conditions is given in Table 1.

2.2 Friends of Friends group detection

As with other recent work (Hacar et al. 2013, 2016; Da Rio et al. 2017; Kuznetsova et al. 2018), we base our Friends of Friends detection algorithm on the original method for classifying clusters of galaxies in redshift space by Huchra & Geller (1982). We perform our analysis using the full six-dimensional information available (*x*, *y*, and *z*, as well as the corresponding velocity components v_x, v_y, and v_z). Most observational studies are done in three dimensions (usually *x*, *y*, and v_z), but we have verified that our results do not significantly differ if done in fewer dimensions. However, by using the full 6D information we would expect the Friends of Friends algorithm to identify real group structures without being hampered by projection effects.

The algorithm proceeds as follows. If a star is not yet assigned a group, we search for companions to that star that are less than a distance d_L and a velocity difference less than v_L. If the nearby star fulfils both criteria then we start a new group and add companions that are less than d_L and v_L from any star in the group. If no further stars fulfil the criteria another unassigned star is chosen randomly and we repeat the process.

Arguably, the most challenging aspect of the Friends of Friends analysis is to define the distance and velocity thresholds, d_L and v_L, above which stars are assigned into different groups. Because we are analysing multiple *N*-body simulations, each containing multiple snapshots of data, we have automated the process of defining d_L and v_L. We create an ordered list of all possible separations between stars, and all possible differences in velocity. We set d_L to be the median separation divided by three ($\tilde{\Delta}r/3$) and v_L to be the median velocity difference ($\tilde{\Delta}v$). There is no real physical basis behind these choices, other than they divide the star-forming region into a reasonable number of groups (i.e. 1–20). These thresholds vary from simulation to simulation (and over time), but the initial values are typically $\tilde{\Delta}r/3 \sim 0.3$ pc and $\tilde{\Delta}v \sim 1$ km s^{-1}, which are similar to the threshold lengths adopted in observational studies (Hacar et al. 2016; Da Rio et al. 2017). However, we

note that dividing a hierarchical fractal distribution into constituent groups is somewhat artificial, and we discuss this issue further in Section 4.

Finally, in order to mimic observational studies, and to avoid the potentially artificial imposition of boundaries, we set an automatic stellar density threshold where any star that resides in the lowest quartile of an ordered list of stellar densities is not assigned to a group. This is intended to reduce the prospect of 'bridges' of only one or two stars between groups causing the algorithm to merge two otherwise distinct groups. However, as we will see, this conservative threshold does not alleviate confusion in the Friends of Friends group detection.

2.3 Other kinematic and spatial measures

In Section 3 we will also look for a dependence of the number of groups identified by the Friends of Friends algorithm on other kinematic and spatial measures. These techniques have been presented in previous papers so we direct the interested reader to the relevant literature, however we briefly summarize them here.

We determine the $\mathcal{Q}$-parameter (Cartwright & Whitworth 2004; Cartwright 2009; Lomax et al. 2011; Jaffa, Whitworth & Lomax 2017), which compares the average length of the edges on the minimum spanning tree of all the stars in a region, $\bar{m}$, to the average separation between stars, $\bar{s}$:

$$Q = \frac{\bar{m}}{\bar{s}}, \tag{2}$$

where $\mathcal{Q} < 0.7$ indicates a substructured distribution and $\mathcal{Q} > 0.9$ indicates a smooth, centrally concentrated distribution. In Section 3 we will plot $\mathcal{Q}$ against three other measures.

First, we will take the ratio of the statistical radial velocity dispersion (the velocity measured along the *z*-axis), σ, to the interquartile range (IQR) of the radial velocities (σ/IQR). Parker & Wright (2016) show that this ratio exceeds unity for clusters that have formed via violent relaxation and merging of substructure. Secondly, we determine the relative surface density of the most massive stars, $\Sigma_{\rm LDR}$, by comparing the median surface density of the 10 most massive stars $\tilde{\Sigma}_{10}$ to the median surface density of all stars, $\tilde{\Sigma}_{\rm all}$ (Küpper et al. 2011; Maschberger & Clarke 2011; Parker et al. 2014):

$$\Sigma_{\rm LDR} = \frac{\tilde{\Sigma}_{10}}{\tilde{\Sigma}_{\rm all}}, \tag{3}$$

where $\Sigma \gg 1$ indicates that the most massive stars are in areas of significantly higher than average stellar density. Finally, we will follow the evolution of the mass segregation ratio, $\Lambda_{\rm MSR}$, which compares the length of the minimum spanning tree of the $N_{\rm MST}$ most massive stars, $l_{\rm subset}$, to the average length $\langle l_{\rm average} \rangle$ of $N_{\rm MST}$ randomly chosen stars (Allison et al. 2009). There is a dispersion associated with the average length of random MSTs, which is roughly Gaussian and can be quantified as the standard deviation of the lengths $\langle l_{\rm average} \rangle \pm \sigma_{\rm average}$. Instead of using $\sigma_{\rm average}$, we conservatively estimate the lower (upper) uncertainty as the MST length which lies 1/6 (5/6) of the way through an ordered list of all the random lengths (corresponding to a 66 per cent deviation from the median value, $\langle l_{\rm average} \rangle$). This determination prevents a single outlying object from heavily influencing the uncertainty. The mass segregation ratio is then

$$\Lambda_{\rm MSR} = \frac{\langle l_{\rm average} \rangle}{l_{\rm subset}} {}^{+\sigma_{5/6}/l_{\rm subset}}_{-\sigma_{1/6}/l_{\rm subset}}, \tag{4}$$
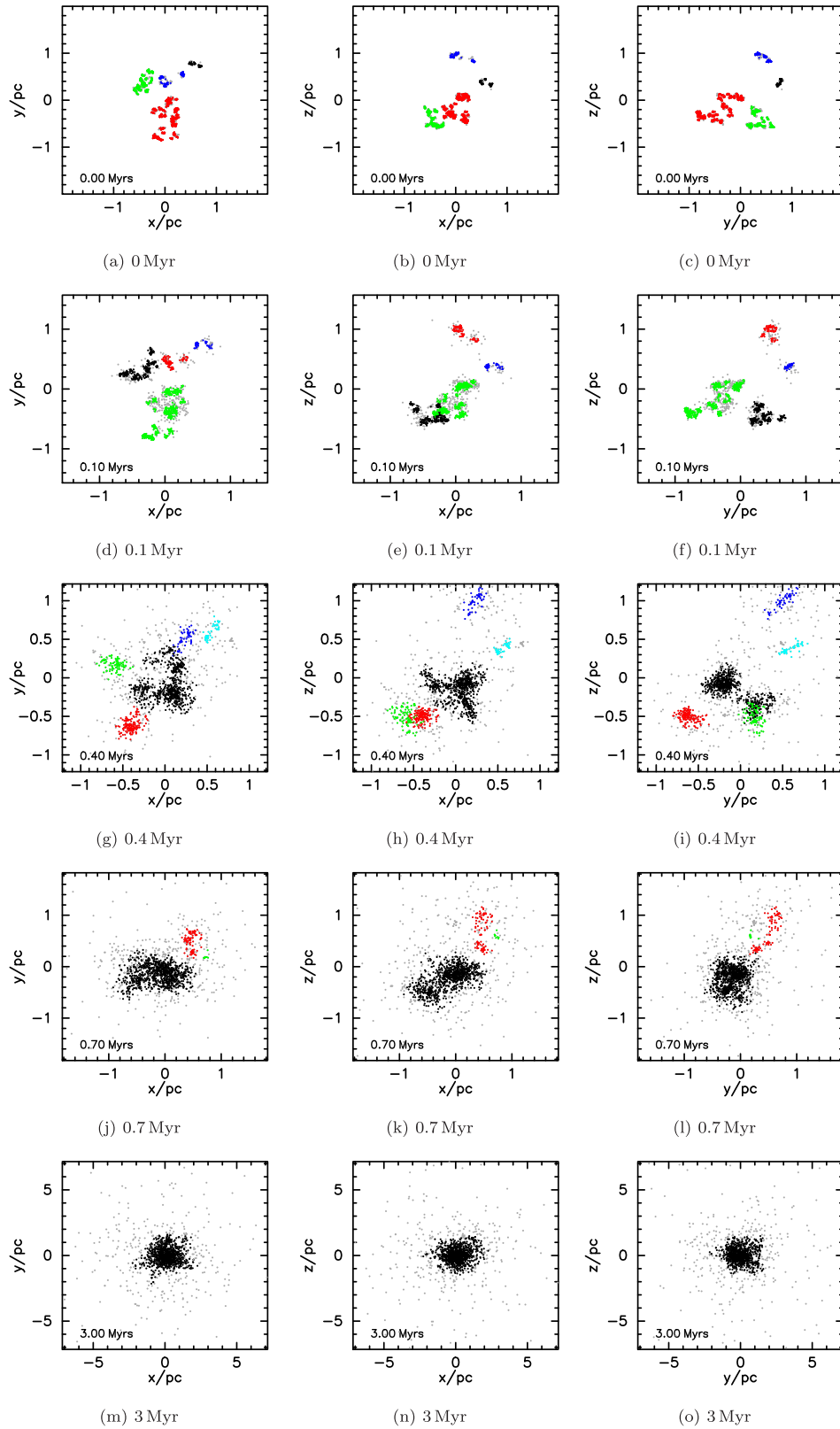
Figure 1. Evolution of groups defined by the Friends of Friends algorithm in a simulated subvirial ($\alpha_{vir} = 0.3$) star-forming region with initial fractal dimension $D = 1.6$. Stars that have a local stellar density below the first quartile in the distribution are not assigned to a group and are coloured grey. The colours in the subsequent snapshots do not correspond to the colours in the first snapshot ($t = 0\,\mathrm{Myr}$).
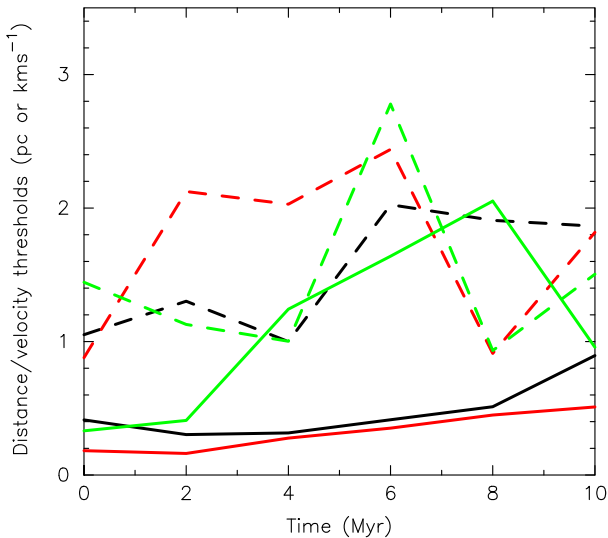
Figure 2. The evolution of distance and velocity threshold lengths automatically calculated at every snapshot in the simulation. The solid lines are the distance thresholds, $\tilde{\Delta}r/3$ and dashed lines are the velocity thresholds $\tilde{\Delta}v$ for three simulations with 4 (the mode – black lines), 8 (median – green lines), and 25 (extreme – red lines) groups initially identified by the Friends of Friends algorithm.

where $\Lambda_{MSR} \gg 1$ indicates that a star-forming region is significantly mass segregated.

3 RESULTS

3.1 Subvirial (collapsing) star-forming regions

We show five snapshots of the evolution of a typical subvirial simulation in Fig. 1, with three different viewing angles (along the z-, y-, and x-axes, respectively) for each snapshot. Each group detected by the Friends of Friends algorithm is shown by a different colour, with stars sitting below the density threshold shown by the grey points. We emphasize that due to the disappearance (or sometimes formation) of groups, long-lived groups are not necessarily shown by the same colour in different snapshots. The Friends of Friends algorithm has detected four distinct groups at $t = 0$ Myr in this particular simulation (panels a–c). As this simulation has a subvirial bulk motion, we would expect the individual groups to evolve and merge as the star-forming region coalesces to a cluster, which occurs during the first 1 Myr. Interestingly, however, the number of groups briefly increases to 5 by 0.4 Myr (panels g–i) before reducing to one main group after 0.7 Myr (panels j–o).

All of the subvirial simulations lose their substructure within the first 1 Myr and form a bound cluster (Parker & Meyer 2012; Parker et al. 2014). Interestingly, the number of distinct groups that the Friends of Friends algorithm identifies at $t = 0$ Myr varies significantly. In our suite of 20 simulations, identical apart from the random number seed used to set the positions, velocities and stellar masses, the number of groups identified varies between 1 and 25, where the mode is 4 and the median number of groups is 8.

To check whether this is due to the automatically calculated threshold lengths, in Fig. 2 we plot the evolution of both the distance length, $\tilde{\Delta}r/3$ (solid lines) and the velocity length, $\tilde{\Delta}v$ (dashed lines) for simulations where the Friends of Friends algorithm identifies 4 groups (black lines), 8 groups (green lines), and 25 groups (red lines). While the distance threshold is smaller when more groups are

identified, this trend is not fulfilled for the velocity thresholds. We also note that the differences in both thresholds between the three simulations are very small (all are less than 0.4 pc and ~ 1 km s^{-1}).

Fig. 3 shows a different subvirial simulation, with 13 groups identified initially, which evolves to a similar-looking single cluster after 1 Myr, which is indistinguishable from the cluster shown in Fig. 1.

To investigate whether the long-term evolution of the cluster depends on the initial number of groups identified by the Friends of Friends algorithm, in Fig. 4 we plot the $\mathcal{Q}$-parameter against the kinematic and spatial diagnostics of cluster evolution, σ/IQR, Σ_{LDR}, and Λ_{MSR}. (We refer the interested reader to Parker & Wright (2016) and Parker et al. (2014) for detailed descriptions of how these diagnostics evolve over time due to dynamical relaxation.) For this paper, we colour code these plots according to the number of groups the Friends of Friends algorithm picks out. The green coloured points indicate simulations where the Friends of Friends algorithm picks out 5 groups or fewer at $t = 0$ Myr; the blue symbols indicate between 5 and 15 groups, and the magenta symbols indicate that the Friends of Friends algorithm has picked out more than 15 groups. It is clear that the initial number of groups picked out by the algorithm is not related to the magnitude of the σ/IQR ratio, or the spatial structure, $\mathcal{Q}$, the relative local surface density, Σ_{LDR}, or the occurrence and amount of mass segregation as measured by Λ_{MSR}.

## 3.2 Supervirial (expanding) star-forming regions

We show a typical example of a supervirial star-forming region in Fig. 5. As in Fig. 1, we show five snapshots in time and three viewing angles for each snapshot. Our first result is that the numbers of groups identified by the Friends of Friends algorithm at $t = 0$ Myr in the suite of 20 simulations are almost identical to the numbers identified in the subvirial simulations, with a range between 1 and 20, a mode of 4, and a median of 8. However, the virial ratio in these supervirial simulations is $\alpha_{vir} = 1.5$, meaning that the global velocity dispersion is significantly higher than in the subvirial simulations ($\alpha_{vir} = 0.3$) presented above. This difference in velocity scaling between the two sets of simulations leads to very different dynamical evolution (expansion and preservation of some substructure versus collapse and erasing of all substructure), yet the number of groups identified by the Friends of Friends algorithm does not betray the future evolution of an individual star-forming region. This is because the threshold lengths are automatically calculated, and so any linear scaling of the velocities (such as changing the virial ratio) will not give statistically different results.

In these simulations a significant degree of substructure is retained, and so we might expect the Friends of Friends algorithm to identify multiple distinct groups of stars. While this is the case, the number of groups identified after 0.1–0.5 Myr of dynamical evolution is typically only 1–3, despite the region displaying rather obvious visual spatial substructure (e.g. panels j–l in Fig. 5). At earlier stages of this simulation (0.4 Myr – panels g–i), bridges of stars are apparent between the (visual) groups of stars which would explain why the Friends of Friends algorithm classifies them as one large group (the red points in panels g–i). However, at later stages these bridges are not as apparent, and the physical distances between the visual groups are larger.

The reason the Friends of Friends algorithm does not detect many distinct groups in this simulation (and others) appears to be due to the large amount of spatial and kinematical mixing that occurs throughout the dynamical evolution of the star-forming region. In Fig. 6 we show the x–y plane of the snapshot at 3 Myr
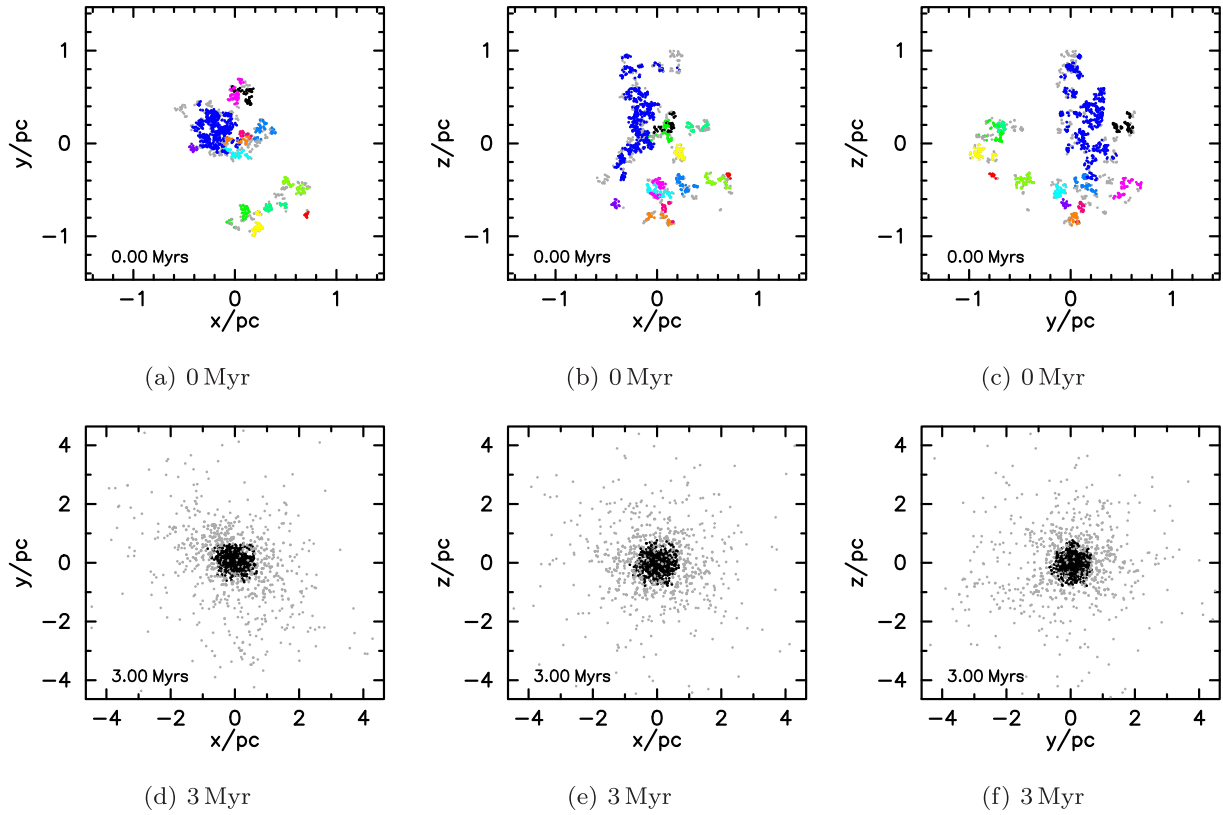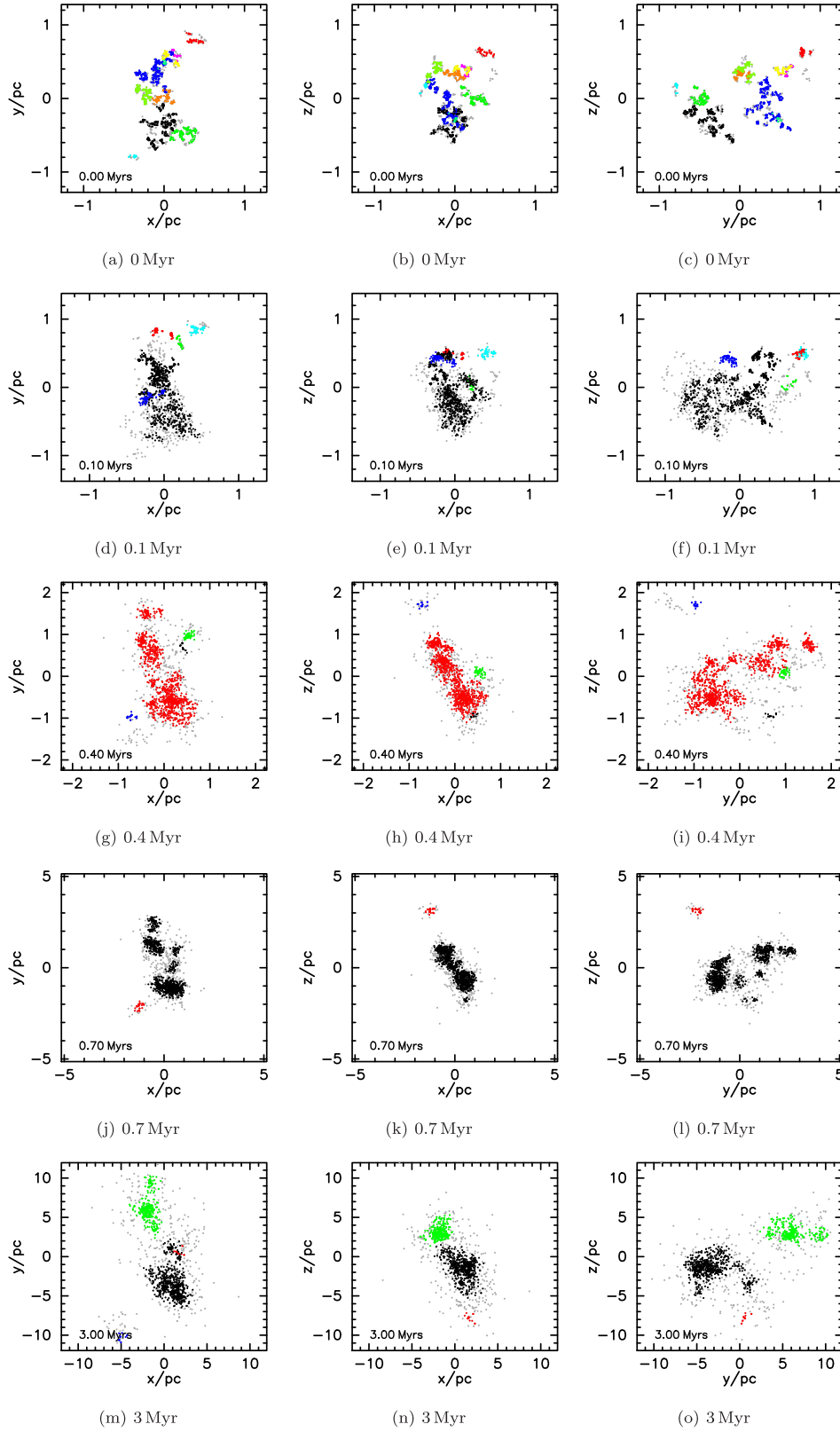
Figure 3. As Fig. 1, but showing the evolution of groups defined by the Friends of Friends algorithm for a different realization of a simulated subvirial ($\alpha_{vir} = 0.3$) star-forming region with $D = 1.6$. This simulation is identical to that in Fig. 1, save for the random number seed used to assign positions, velocities, and masses to the stars. In this simulation, 13 distinct groups are found by the Friends of Friends algorithm, compared to 4 in the simulation shown in Fig. 1. Stars that have a local stellar density below the first quartile in the distribution are not assigned to a group and are coloured grey. The colours in the subsequent snapshots do not correspond to the colours in the first snapshot ($t = 0\,\mathrm{Myr}$).
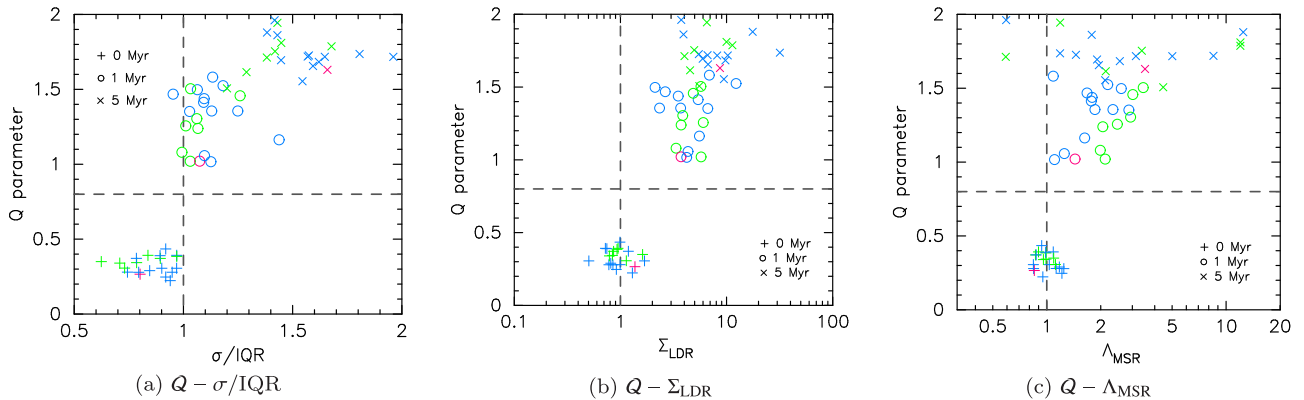
Figure 4. Different measures of spatial and kinematic evolution in our subvirial ($\alpha_{vir} = 0.3$) simulations. In panel (a) we show the Cartwright & Whitworth (2004) $\mathcal{Q}$-parameter against the radial velocity dispersion divided by the interquartile range of radial velocities (Parker & Wright 2016). In panel (b) we show the $\mathcal{Q}$-parameter against the relative local surface density ratio of the 10 most massive stars, Σ_{LDR} (Küpper et al. 2011; Parker et al. 2014). In panel (c) we show the $\mathcal{Q}$-parameter against the mass segregation ratio Λ_{MSR} (Allison et al. 2009). The boundary between hierarchically substructured and centrally concentrated distributions is shown by the horizontal dashed lines, and the vertical dashed lines correspond to unity for the other measures, indicating no special configuration for the massive stars. In all panels, the green symbols indicate simulations where the Friends of Friends algorithm picks out 5 groups or fewer at $t = 0\,\mathrm{Myr}$; the blue symbols indicate between 5 and 15 groups, and the magenta symbols indicate that the Friends of Friends algorithm has picked out more than 15 groups. There is no strong dependence of the dynamical evolution on the initial number of groups.
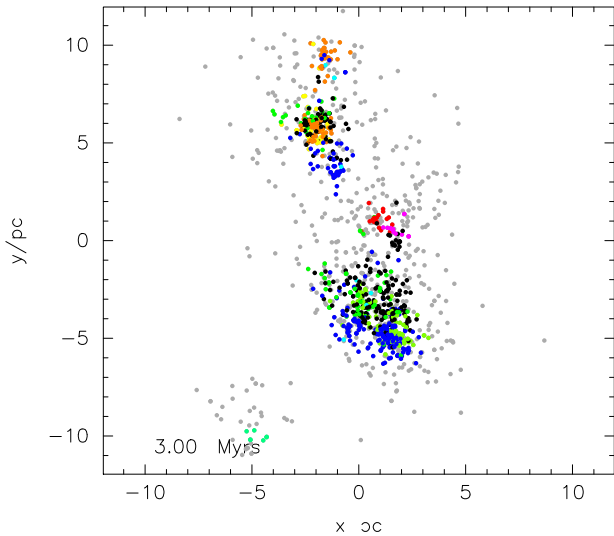
from Fig. 5(m), but the stars are colour coded according to their original groups at $t = 0\,\mathrm{Myr}$ (Fig. 5a). Clearly, stars have migrated significantly, yet the star-forming region has preserved some spatial substructure. Interestingly, Arnold et al. (2017) find a similar result

when examining the formation of binary star clusters – two clusters orbiting a common centre of mass – from initially supervirial, substructured star-forming regions like those in this paper. Arnold et al. (2017) find that the stars that constitute the components of the

(a) 0 Myr   (b) 0 Myr   (c) 0 Myr

(d) 0.1 Myr   (e) 0.1 Myr   (f) 0.1 Myr

(g) 0.4 Myr   (h) 0.4 Myr   (i) 0.4 Myr

(j) 0.7 Myr   (k) 0.7 Myr   (l) 0.7 Myr

(m) 3 Myr   (n) 3 Myr   (o) 3 Myr

Figure 5. Evolution of groups defined by the Friends of Friends algorithm in a simulated supervirial ($\alpha_{\rm vir} = 1.5$) star-forming region with $D = 1.6$. Stars that have a local stellar density below the first quartile in the distribution are not assigned to a group and are coloured grey. The colours in the subsequent snapshots do not correspond to the colours in the first snapshot ($t = 0$ Myr).

Figure 6. Same as Fig. 5(m), but showing the stars colour coded according to their original groups in Fig. 5(a). Significant mixing of the groups has occurred, despite the region retaining spatial and kinematic substructure.

binary clusters do not originate in the same location as their fellow constituents, making it impossible to predict where a star will end up during the evolution of a supervirial star-forming region.

We now investigate whether the long-term evolution of the supervirial star-forming regions depends on the initial number of groups identified by the Friends of Friends algorithm. In Fig. 7 we plot the $\mathcal{Q}$-parameter against the kinematic and spatial diagnostics of dynamical evolution, σ/IQR, Σ_{LDR}, and Λ_{MSR} and we colour code these plots according to the number of groups the Friends of Friends algorithm picks out. The green points indicate simulations where the Friends of Friends algorithm picks out 5 groups or fewer at $t = 0$ Myr; the blue symbols indicate between 5 and 15 groups, and the magenta symbols indicate that the Friends of Friends algorithm has picked out more than 15 groups. As with the subvirial simulations, the initial number of groups picked out by the algorithm is not related to the subsequent magnitude of the σ/IQR ratio, or the spatial structure, $\mathcal{Q}$, the relative local surface density, Σ_{LDR}, or the occurrence and amount of mass segregation, Λ_{MSR} (though the amount of mass segregation in supervirial star-forming regions is minimal because the massive stars rarely interact with each other).

3.3 Varying the initial degree of substructure

As discussed in Section 2.1, we might expect the evolution of spatio-kinematic groups identified by the Friends of Friends algorithm to be correlated with the initial fractal dimension of the simulation. So far, we have presented the results for the cases where $D = 1.6$, where we expect the time-scale for structure erasure to be $t(L) \propto L^{-0.4}$, which means that structure is erased more quickly on larger scales. Conversely, when the fractal dimension is higher ($D = 2.0$ or 2.6, then structure is erased more quickly on smaller scales [where $t(L) \propto L$ for $D = 2.0$ and for $D = 2.6$ $t(L) \propto L^{0.6}$].

We show a typical example of a simulation where the initial fractal dimension is $D = 2.0$ in Fig. 8, with stars above the density threshold of the first quartile assigned to groups. Again, the colours of groups identified at later snapshots are not correlated with the colours of the groups identified at $t = 0$ Myr. These initial conditions ($D = 2.0$ and $\alpha_{vir} = 1.5$) often lead to the formation of a binary

cluster (Arnold et al. 2017), and this happens in more than 50 per cent of the realizations of this set of initial conditions.

We do not see any clear evidence for the time-scale for substructure erasure to be different for the $D = 2.0$ simulations compared to the simulations with $D = 1.6$ (with similar results for the $D = 2.6$ simulations that we do not show here for the sake of brevity). This may be because the groups identified by the Friends of Friends algorithm are not representative of the physical scales in our box fractal, or simply that the Friends of Friends algorithm cannot distinguish between multiple groups once dynamical evolution takes place.

As in the corresponding supervirial simulations with $D = 1.6$, the simulations with less substructure also expose the same problem with the Friends of Friends algorithm, namely that distinct clumps of stars are assigned to the same group due to bridging stars, and some of these groups are transient between snapshots (such as the cyan coloured group in Figs 8j–l, which disappears in snapshots either side of 0.7 Myr).

We identify fewer groups initially than in the $D = 1.6$ case (a mode of 2 and a median of 5), but with similarly large range in group number (1–15). We therefore conclude that the issues identified with the Friends of Friends technique are not unique to a particular set of initial conditions.

4 DISCUSSION

Our analysis of *N*-body simulations using the Friends of Friends algorithm to identify spatio-kinematic groups exposes several issues with using this technique to quantify the initial conditions of star formation.

First, we have performed our analysis on fractal distributions. By definition, fractals are hierarchical and self-similar, with the only boundary conditions being the size scale of the distribution and the velocity dispersion. Therefore, any groups that are identified are necessarily artificial and arbitrary (see also Parker & Goodwin 2015), without any physical meaning. In an observed star-forming region, it is also usually extremely unclear whether the region can be broken down into constituent parts. This is especially relevant if star formation is inherently hierarchical or self-similar (e.g. Elmegreen 2018), which implies that there is no scale length for star formation.

Secondly, we find a wide range in the number of groups identified by the Friends of Friends algorithm. Our simulations that are set up to be in subvirial collapse and have a high degree of initial spatial and kinematic substructure contain between 1 and 25 groups, with a mode of 4 and a median of 8, for statistically similar fractal distributions, identical apart from the random number seed used to initialize the positions and velocities. Worryingly, we find very similar numbers of groups when the fractals are scaled to be supervirial (i.e. expanding), despite the regions having very different initial velocity dispersions. Therefore, the number of groups identified in a star-forming region (or filaments – cf. Hacar et al. 2013, 2017) does not betray any information about the physical initial conditions and may be subject to line-of-sight confusion (Clarke et al. 2018).

Thirdly, during the subsequent dynamical evolution of our star-forming regions, stars move between groups, often creating bridges between groups so that two distinct groups become one larger group. Observational studies (Hacar et al. 2016; Da Rio et al. 2017) often attempt to mitigate for this by introducing a density threshold, so that lone stars (or gas parcels) do not unduly influence the number of groups (or filaments) identified. We also applied a density threshold to our *N*-body simulations, but find that bridges of stars still occur, and that the Friends of Friends algorithm cannot separate larger groups. In part, this is due to the fact that stars migrate sig-

(a) $Q - \sigma/\mathrm{IQR}$ (b) $Q - \Sigma_{\mathrm{LDR}}$ (c) $Q - \Lambda_{\mathrm{MSR}}$
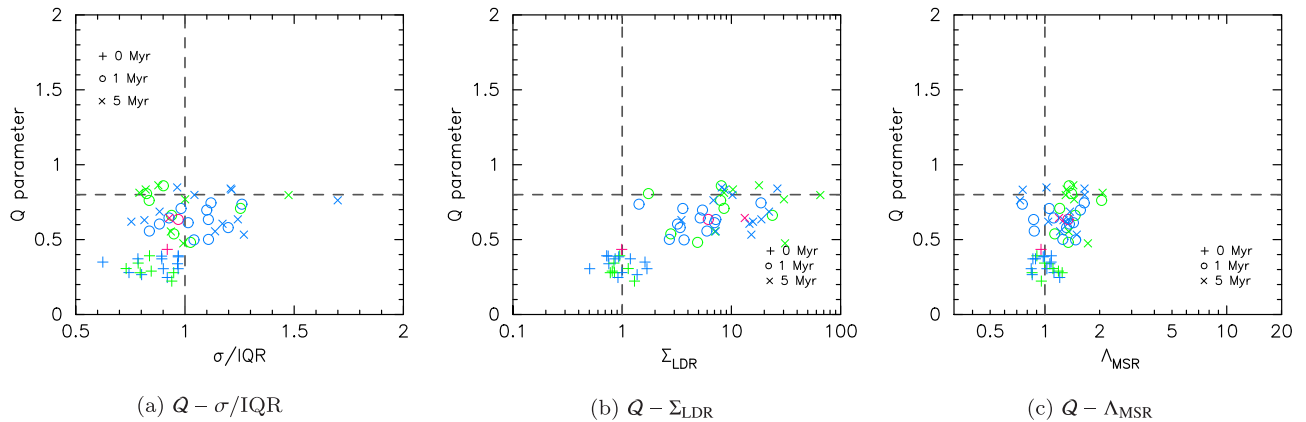
Figure 7. Different measures of spatial and kinematic evolution in our supervirial ($\alpha_{\mathrm{vir}} = 1.5$) simulations. In panel (a) we show the Cartwright & Whitworth (2004) $\mathcal{Q}$-parameter against the radial velocity dispersion divided by the interquartile range of radial velocities (Parker & Wright 2016). In panel (b) we show the $\mathcal{Q}$-parameter against the relative local surface density ratio of the 10 most massive stars, Σ_{LDR} (Küpper et al. 2011; Parker et al. 2014). In panel (c) we show the $\mathcal{Q}$-parameter against the mass segregation ratio Λ_{MSR} (Allison et al. 2009). The symbols and lines are the same as in Fig. 4.

nificant distances during the dynamical evolution of the simulations (Fig. 6) and swap between groups (see also Arnold et al. 2017). The problem of 'bridging' is likely ignored in observational studies if the thresholds for group definition are tuned to each specific region at a given time, rather than being automated as we have done here.

Fourthly, we have found no dependence on the later evolutionary state of the star-forming regions on the number of groups identified initially by the Friends of Friends algorithm. We measured the amount of spatial substructure, mass segregation, velocity dispersion, and relative surface densities, and find no dependence on the number of initial groups.

We note that some of our simulations have initially very high stellar densities ($\tilde{\rho} \sim 10^4$ stars pc^{-3}). This facilitates rapid dynamical evolution, but does not affect the number of groups identified initially. We ran a set of low-density simulations ($\tilde{\rho} \sim 10$ stars pc^{-3}) and found similar behaviour, albeit on longer dynamical time-scales. Similarly, changing the initial degree of spatial and kinematic substructure (which can invert the time-scales on which we would expect structure to be erased) does not affect our conclusions.

We have performed our analysis using the full six-dimensional information (as the position and velocity vector of every star contain three components). We also performed it in three dimensions (x, y, and v_z) to mimic the information available in (most) observational studies and found similar results. Indeed, one could argue that the confusion present when all of the spatio-kinematic information is available should preclude any further use of the technique in fewer dimensions.

5 CONCLUSIONS

We analyse *N*-body simulations of the dynamical evolution of subvirial (collapsing) and supervirial (expanding) star-forming regions and apply an automated Friends of Friends algorithm to pick out groups or stars that have similar spatial and kinematic properties. Our conclusions are the following:

(i) The Friends of Friends technique picks out wide-ranging numbers of groups in statistically identical spatio-kinematic fractal distributions, despite the threshold lengths for distance and velocity

varying little between individual simulations. The mode is 4 and the median is 8 but the number of groups identified across 20 identical simulations ranges from 1 to 25.

(ii) We do not see any difference in the number of groups identified in subvirial and supervirial simulations. The mode and median numbers of groups are identical, and the range is almost identical. This is because although the initial velocity scalings are very different (0.3 km s^{-1} for the subvirial simulations versus 1.5 km s^{-1} for the supervirial simulations), the scaling is linear and our Friends of Friends algorithm automatically calculates distance and velocity thresholds. In practice, this means that any automated analysis would not be able to distinguish between very different initial star formation conditions.

(iii) The dynamical evolution of the star-forming regions causes the groups to merge together in Friends of Friends space, even if (as in the case of the supervirial simulations) there are still distinct spatial substructures. This occurs because stars migrate between groups as the simulation progresses, but the groups do not dynamically mix with each other. Therefore, at a given point in the evolution of a star-forming region, a spatio-kinematic group is not retaining any information on the initial properties of that group.

(iv) Furthermore, there is no dependence of the later spatial and kinematic evolution of the star-forming regions on the number of groups identified by the Friends of Friends algorithm. The amount of mass segregation, overall spatial structure, and velocity dispersion that develops with time are unrelated to the initial number of groups, implying that the global dynamical evolution of the star-forming region cannot be related to any group structure defined by the Friends of Friends method.

Taken together, our results suggest that the Friends of Friends algorithm may not be particularly useful for quantifying the initial conditions of star-forming regions, and we urge users to include tests on synthetic data sets in any future analyses (see also Clarke et al. 2018).
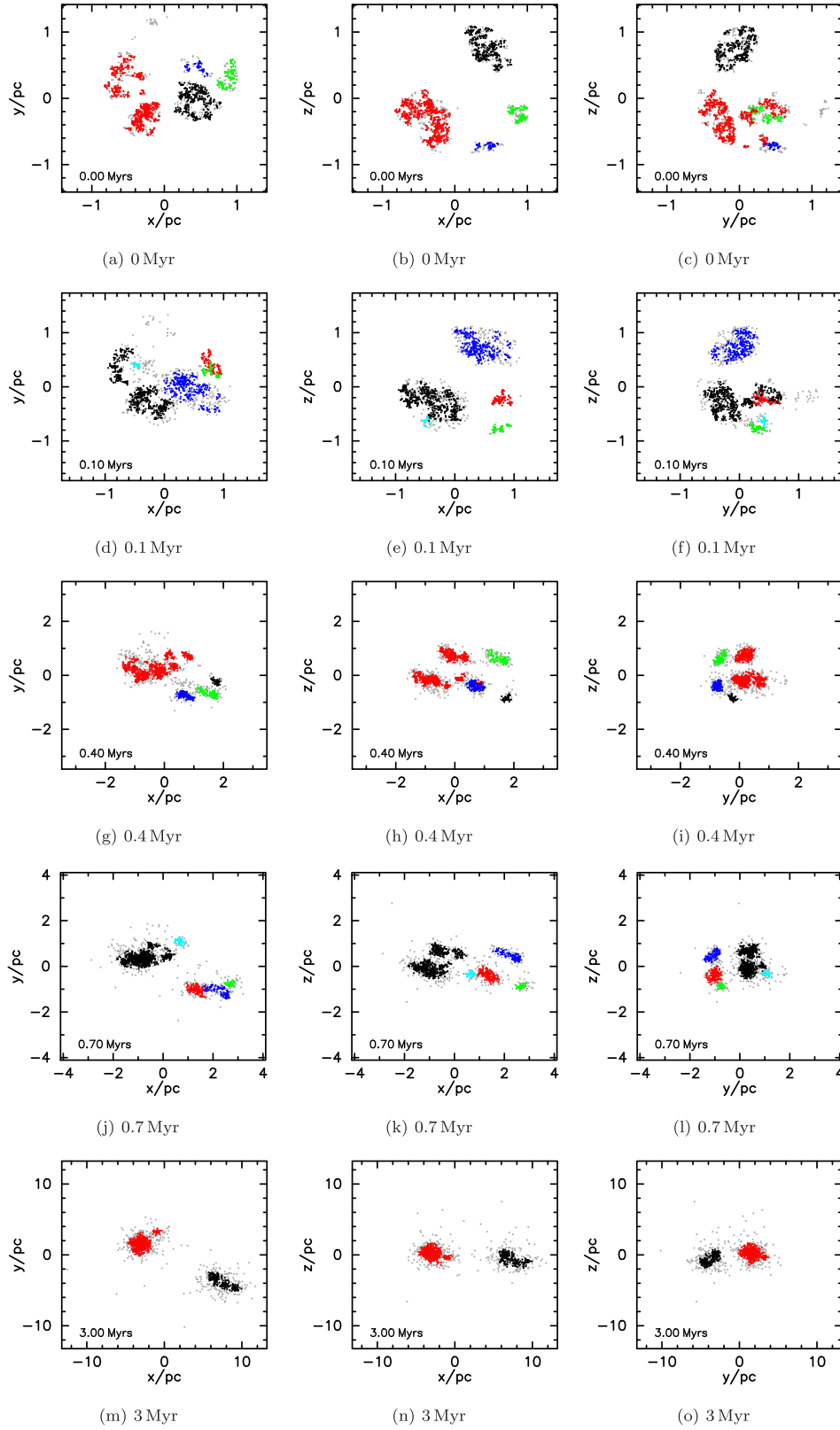
Figure 8. Evolution of groups defined by the Friends of Friends algorithm in a simulated supervirial ($\alpha_{\rm vir} = 1.5$) star-forming region with moderate levels of initial spatial and kinematic substructure (fractal dimension $D = 2.0$). Stars that have a local stellar density below the first quartile in the distribution are not assigned to a group and are coloured grey. The colours in the subsequent snapshots do not correspond to the colours in the first snapshot ($t = 0\,{\rm Myr}$).

REFERENCES

Adams F. C., Hollenbach D., Laughlin G., Gorti U., 2004, ApJ, 611, 360
Allison R. J., Goodwin S. P., 2011, MNRAS, 415, 1967
Allison R. J., Goodwin S. P., Parker R. J., Portegies Zwart S. F., de Grijs R., Kouwenhoven M. B. N., 2009, MNRAS, 395, 1449
Allison R. J., Goodwin S. P., Parker R. J., Portegies Zwart S. F., de Grijs R., 2010, MNRAS, 407, 1098
Arnold B., Goodwin S. P., Griffiths D. W., Parker R. J., 2017, MNRAS, 471, 2498
Bastian N., Covey K. R., Meyer M. R., 2010, ARA&A, 48, 339
Bressert E. et al., 2010, MNRAS, 409, L54
Cai M. X., Kouwenhoven M. B. N., Portegies Zwart S. F., Spurzem R., 2017, MNRAS, 470, 4337
Cartwright A., 2009, MNRAS, 400, 1427
Cartwright A., Whitworth A. P., 2004, MNRAS, 348, 589
Cartwright A., Whitworth A. P., Nutter D., 2006, MNRAS, 369, 1411
Clarke S. D., Whitworth A. P., Spowage R. L., Duarte-Cabral A., Suri S. T., Jaffa S. E., Walch S., Clark P. C., 2018, MNRAS, 479, 1722
Da Rio N. et al., 2017, ApJ, 845, 105
Elmegreen B. G., 2018, ApJ, 853, 88
Foster J. B. et al., 2015, ApJ, 799, 136
Goodwin S. P., Whitworth A. P., 2004, A&A, 413, 929
Gouliermis D. A., Hony S., Klessen R. S., 2014, MNRAS, 439, 3775
Gutermuth R. A., Megeath S. T., Myers P. C., Allen L. E., Pipher J. L., Fazio G. G., 2009, ApJS, 184, 18
Hacar A., Tafalla M., Kauffmann J., Kovács A., 2013, A&A, 554, A55
Hacar A., Alves J., Forbrich J., Meingast S., Kubiak K., Großschedl J., 2016, A&A, 589, A80
Hacar A., Tafalla M., Alves J., 2017, A&A, 606, A123
Hacar A., Tafalla M., Forbrich J., Alves J., Meingast S., Grossschedl J., Teixeira P. S., 2018, A&A, 610, A77
Henshaw J. D. et al., 2016a, MNRAS, 457, 2675
Henshaw J. D. et al., 2016b, MNRAS, 463, 146
Huchra J. P., Geller M. J., 1982, ApJ, 257, 423
Jaehnig K. O., Da Rio N., Tan J. C., 2015, ApJ, 798, 126
Jaffa S. E., Whitworth A. P., Lomax O., 2017, MNRAS, 466, 1082
Jaffa S. E., Whitworth A. P., Clarke S. D., Howard A. D. P., 2018, MNRAS, 477, 1940
Jeffries R. D. et al., 2014, A&A, 563, A94

Kereš D., Katz N., Davé R., Fardal M., Weinberg D. H., 2009, MNRAS, 396, 2332
Kuhn M. A. et al., 2014, ApJ, 787, 107
Küpper A. H. W., Maschberger T., Kroupa P., Baumgardt H., 2011, MNRAS, 417, 2300
Kuznetsova A., Hartmann L., Ballesteros-Paredes J., 2018, MNRAS, 473, 2372
Lada C. J., Lada E. A., 2003, ARA&A, 41, 57
Larson R. B., 1981, MNRAS, 194, 809
Larson R. B., 1995, MNRAS, 272, 213
Lomax O., Whitworth A. P., Cartwright A., 2011, MNRAS, 412, 627
Lynden-Bell D., 1967, MNRAS, 136, 101
Marks M., Kroupa P., 2012, A&A, 543, A8
Maschberger T., 2013, MNRAS, 429, 1725
Maschberger T., Clarke C. J., 2011, MNRAS, 416, 541
Parker R. J., 2014, MNRAS, 445, 4037
Parker R. J., Alves de Oliveira C., 2017, MNRAS, 468, 4340
Parker R. J., Dale J. E., 2015, MNRAS, 451, 3664
Parker R. J., Goodwin S. P., 2015, MNRAS, 449, 3381
Parker R. J., Meyer M. R., 2012, MNRAS, 427, 637
Parker R. J., Quanz S. P., 2012, MNRAS, 419, 2448
Parker R. J., Wright N. J., 2016, MNRAS, 457, 3430
Parker R. J., Wright N. J., Goodwin S. P., Meyer M. R., 2014, MNRAS, 438, 620
Peretto N., André P., Belloche A., 2006, A&A, 445, 979
Portegies Zwart S. F., 2016, MNRAS, 457, 313
Portegies Zwart S. F., Makino J., McMillan S. L. W., Hut P., 1999, A&A, 348, 117
Portegies Zwart S. F., McMillan S. L. W., Hut P., Makino J., 2001, MNRAS, 321, 199
Salpeter E. E., 1955, ApJ, 121, 161
Scally A., Clarke C., 2001, MNRAS, 325, 449
Traficante A. et al., 2018, MNRAS, 477, 2220
Vázquez-Semadeni E., González-Samaniego A., Colín P., 2017, MNRAS, 467, 1313
Vincke K., Breslau A., Pfalzner S., 2015, A&A, 577, A115
Williams G. M., Peretto N., Avison A., Duarte-Cabral A., Fuller G. A., 2018, A&A, 613, A11
Wright N. J., Mamajek E. E., 2018, MNRAS, 476, 381
Wright N. J., Parker R. J., Goodwin S. P., Drake J. J., 2014, MNRAS, 438, 639
Wright N. J., Bouy H., Drew J. E., Sarro L. M., Bertin E., Cuillandre J.-C., Barrado D., 2016, MNRAS, 460, 2593

This paper has been typeset from a TEX/LATEX file prepared by the author.