



Kinematics in Young Star Clusters and Associations with *Gaia* DR2

Michael A. Kuhn¹ , Lynne A. Hillenbrand¹, Alison Sills² , Eric D. Feigelson³ , and Konstantin V. Getman³ 

¹Department of Astronomy, California Institute of Technology, Pasadena, CA 91125, USA; mkuhn@astro.caltech.edu

²Department of Physics & Astronomy, McMaster University, 1280 Main Street West, Hamilton, ON L8S 4M1, Canada

³Department of Astronomy & Astrophysics, Pennsylvania State University, 525 Davey Laboratory, University Park, PA 16802, USA

Received 2018 July 5; revised 2018 November 4; accepted 2018 November 5; published 2019 January 2

Abstract

The *Gaia* mission has opened a new window into the internal kinematics of young star clusters at the sub-km s⁻¹ level, with implications for our understanding of how star clusters form and evolve. We use a sample of 28 clusters and associations with ages from ~1–5 Myr, where lists of members are available from previous X-ray, optical, and infrared studies. Proper motions from *Gaia* DR2 reveal that at least 75% of these systems are expanding; however, rotation is only detected in one system. Typical expansion velocities are on the order of ~0.5 km s⁻¹, and in several systems, there is a positive radial gradient in expansion velocity. Systems that are still embedded in molecular clouds are less likely to be expanding than those that are partially or fully revealed. One-dimensional velocity dispersions, which range from $\sigma_{1D} = 1$ to 3 km s⁻¹, imply that most of the stellar systems in our sample are supervirial and that some are unbound. In star-forming regions that contain multiple clusters or subclusters, we find no evidence that these groups are coalescing, implying that hierarchical cluster assembly, if it occurs, must happen rapidly during the embedded stage.

Key words: astrometry – open clusters and associations: general – stars: formation – stars: kinematics and dynamics

Supporting material: figure sets, data behind figure

1. Introduction

While star formation in the Galaxy occurs in large star-forming complexes, most of the stars formed in these complexes quickly disperse into the field without remaining bound to one another as members of open clusters (Adams 2010; Gouliermis 2018). Star formation takes place in turbulent, clumpy giant molecular clouds, and yet the star clusters that do remain bound tend to have smooth stellar density distributions. The processes of cluster assembly, equilibration, and dissolution have remained poorly constrained by observation for two reasons: the difficulty of obtaining reliable samples of cluster members in nebulous regions with many field-star contaminants and the absence of kinematic information for faint stars. The nearby Orion Nebula Cluster (ONC) has been extensively characterized, but these difficulties have been overcome only more recently for other massive star-forming complexes. In these regions, samples of members emerge from multiwavelength surveys (Feigelson 2018), while the *Gaia* mission (Gaia Collaboration et al. 2016) is expected to revolutionize the study of their internal kinematics (Allison 2012; Clarke et al. 2015; Moraux 2016).

The kinematics of young stars should reflect the processes of star cluster formation. Two principal paths from theoretical studies are “monolithic” cluster formation, in which a young star cluster is born in a single molecular cloud core, and “hierarchical” cluster formation, in which larger clusters are built via the accumulation of smaller subclusters (Elmegreen 2000; Bonnell et al. 2003; Banerjee & Kroupa 2015). Young stellar objects (YSOs) that are embedded or partially embedded in star-forming clouds typically have clumpy distributions, where groups are known as subclusters (e.g., Aarseth & Hills 1972; Lada & Lada 1995; Kuhn et al. 2014; Megeath et al. 2016; Sills et al. 2018). It has been unclear, however, whether the subclusters will disperse once the molecular gas disperses, or if they will merge into larger, possibly bound clusters that survive gas expulsion. Examination

of the motions of these subclusters can help constrain cluster formation scenarios.

The evolution of a young cluster depends on its dynamical state (e.g., Parker et al. 2014; Sills et al. 2018), as well as on changes in the gravitational potential due to the dispersal of the molecular cloud (e.g., Tutukov 1978; Hills 1980) and tidal interactions with clouds and clusters (e.g., Kruijssen 2012). Much attention has been focused on the role of gas expulsion, which will cause a cluster to expand and can cause bound embedded clusters to become unbound (Elmegreen 1983; Mathieu 1983; Adams 2000; Kroupa et al. 2001; Goodwin & Bastian 2006; Pelupessy & Portegies Zwart 2012; Brinkmann et al. 2017). However, recent simulations have suggested that gas expulsion may play a less important role in cluster disruption if there is spatial decoupling between stars and gas (Dale et al. 2015).

The empirical relationship between the size and density of young star clusters and associations provides indirect evidence for their expansion (Pfalzner 2009, 2011; Pfalzner et al. 2014). For subclusters, size has been found to be negatively correlated with density and positively correlated with age, as would be expected if they were expanding (Kuhn et al. 2015a; Getman et al. 2018). Furthermore, Marks & Kroupa (2012) argued that the low binary fractions observed in young star clusters imply these clusters were more compact upon formation.

Direct evidence of cluster expansion via proper-motion measurements has been difficult to obtain. Recent studies of OB associations have yielded either no evidence of expansion (Ward & Kruijssen 2018) or no evidence of expansion from a single compact system (Wright et al. 2016; Wright & Mamajek 2018). For the ONC, there has long been debate about whether the system is expanding or contracting (Muench et al. 2008). In a recent work, Dzib et al. (2017) found no evidence of expansion or contraction of the ONC using proper motions from radio observations, while Da Rio et al. (2017) reported a correlation between radial velocity (RV) and extinction that can be explained

by expansion, and Kounkel et al. (2018) reported a preference for expansion among stars around the outer edges of Orion A.

In this paper, we use the superb astrometry of *Gaia*'s Second Data Release (DR2; Gaia Collaboration et al. 2018a) to elucidate the formation (evidence for or against hierarchical assembly) and dynamical state (expansion, contraction, or equilibrium) of young star clusters. We use a sample of young star clusters and associations, ranging from 0.3 to 3.7 kpc, that were the focus of a series of studies combining NASA's *Chandra* X-ray, *Spitzer* mid-infrared, and ground-based optical and near-infrared images to provided reliable samples of tens of thousands of YSOs (Feigelson et al. 2013; Getman et al. 2017; Kuhn et al. 2017b).

Section 2 describes the data, and Section 3 derives the basic cluster properties that are used in the analysis. Section 4 addresses the question of whether clusters are in equilibrium, expanding, or contracting, and whether they are rotating. Section 5 examines the velocity dispersions in clusters. Section 6 addresses the question of whether young clusters form by merging of subclusters. Section 7 relates internal cluster kinematics to other physical properties. A discussion of the observational results, along with a comparison to a cluster formation simulation, is provided in Section 8. Section 9 is the conclusion.

2. Data Sets

2.1. YSOs in Clusters and Associations

This *Gaia* study is based on samples of YSOs, with typical ages of 1–5 Myr, from the Massive Young Star-forming Complex Study in Infrared and X-ray survey (MYStIX; Broos et al. 2013; Feigelson et al. 2013), the Star Formation in Nearby Clouds study (SFINC; Getman et al. 2017), and a similar study of NGC 6231 (Kuhn et al. 2017b). In each of these studies, X-ray emission was used to classify probable cluster members based on the higher expected X-ray luminosities of pre-main-sequence stars compared to main-sequence field stars. MYStIX and SFINC also include sources selected by infrared excess, which indicates YSOs with disks and envelopes that may or may not be X-ray emitters. The infrared-selected samples come from Povich et al. (2013) and Getman et al. (2017), with careful attention to reduce contamination by post-main-sequence dusty red giants. Reducing field-star contamination is particularly important for these massive star-forming regions that lie in the Galactic Plane.

In this study, we include only the regions that contain rich clusters visible in the optical. We omit the regions with few stars and those for which *Gaia* is limited by high optical extinction (e.g., Serpens South or DR 21). To be included in our study, a region must contain at least 20 cluster members with reliable *Gaia* proper motions. These criteria yield a sample of 28 star clusters and associations that reside in 21 star-forming regions (Table 1).

In this paper, we use the term “stellar system” to indicate a major group of spatially associated stars, which includes embedded clusters, bound open clusters, and compact associations of unbound stars (Kuhn et al. 2014).⁴ Some of the regions contain multiple stellar systems that are analyzed individually.

⁴ Historically, it has been difficult to distinguish between bound and unbound young stellar systems from observations. Studies of cluster density have shown that there is no density threshold that divides “clustered” and “distributed” star formation (e.g., Bressert et al. 2010; Gieles et al. 2012; Kuhn et al. 2015b), and uncertainties in measurements of system masses and kinematics limit dynamical modeling. The *Gaia* mission will undoubtedly improve the situation for the last two issues.

Notable examples include NGC 2264 (containing a loose association around S Mon and embedded clusters around IRS 1 and IRS 2 adjacent to the Cone Nebula), NGC 6357 (containing Pismis 24, G353.1+0.6, and G353.2+0.7), the Carina Nebula (including Tr 14, Tr 15, and Tr 16), and the Cep OB3b association (containing a group to the east adjacent to the Cep B cloud and a group to the west around V454 Cep). For the Rosette Nebula region, which includes stars both in the cluster NGC 2244 and in the Rosette Molecular Clouds, we use only NGC 2244 for expansion analysis. In the Orion star-forming region, we focus only on the ONC. The sample of 28 stellar systems is given in Table 1.

2.2. Cross-matching to *Gaia* DR2

Our study is primarily based on astrometric measurements from the *Gaia* DR2 catalog (Gaia Collaboration et al. 2018a; Lindegren et al. 2018). We use the astrometric notation α and δ for R.A. and decl., ϖ for parallax in units of milliarcseconds (mas), and μ_{α^*} and μ_{δ} for proper motions in units of mas yr^{-1} , where $\mu_{\alpha^*} \equiv \mu_{\alpha} \cos \delta$.

We cross-matched 30,839 objects from the YSO catalogs to sources in the *Gaia* catalog. Significant effort has already been devoted to identifying the best match between X-ray and optical/infrared sources in the MYStIX, SFINC, and NGC 6231 catalogs (e.g., Naylor et al. 2013; Getman et al. 2017). The optical or infrared source coordinates are often more precise than the X-ray positions, and in such cases, we use those coordinates for cross-matching with *Gaia*. The match radius for matching to *Gaia* sources was set to 1.2 arcsec, and we select the nearest *Gaia* source within that match radius.

The cross-matching led to 20,716 matches with the *Gaia* DR2 catalog, 17,509 of which have the five-parameter “astrometric global iterative solution” (AGIS) involving position, parallax, and proper motion (Lindegren et al. 2018). The median magnitude of these sources is $G = 18.1$ mag (interquartile range: 16.6–19.1 mag), and the median proper-motion precision is $\sigma_{\mu} = 0.4 \text{ mas yr}^{-1}$ (interquartile range: 0.2–0.8 mas yr^{-1}). The *Gaia* catalog includes statistical uncertainties on the astrometric properties calculated from the astrometric model. For AGIS models that do not converge, solutions are provided with relaxed criteria, with up to 20 mas of astrometric excess noise as defined by Lindegren et al. (2012). We only accept sources with `astrometric_excess_noise` < 1 mas as providing reliable kinematics. We also omit likely non-member contaminants (Section 3.3) and sources with statistical uncertainties >3 km s^{-1} on tangential velocity (Section 3.4). The final sample contains 6507 objects.

In Figure 1, the left panel shows a near-infrared color-magnitude diagram for both the full catalog of YSOs in NGC 6530 and the subsample used for the *Gaia* analysis. In this region, at a distance of $d = 1.34$ kpc, the *Gaia* sources include stars down to $\sim 0.5 M_{\odot}$, and tend to be the least absorbed sources. The right panel shows proper motions, and their uncertainties, as a function of *Gaia*'s G magnitude for these sources. The mass range of the *Gaia* sample is different in different regions, depending on distance and extinction. However, in nearly all regions, stars down to 0.5–1 M_{\odot} are included. Several nearby regions have *Gaia* data that probe to even lower mass stars, notably NGC 1333, IC 348, the ONC, and NGC 2264. Conversely, our sample for the distant region NGC 1893 contains only stars with masses >2 M_{\odot} .

Table 1
Cluster/Association Sample

Region	α_0 (J2000)	δ_0 (J2000)	n_{samp} (stars)	$\mu_{\alpha^*,0}$ (mas yr $^{-1}$)	$\mu_{\delta,0}$ (mas yr $^{-1}$)	ϖ_0 (mas)	Distance (pc)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Berkeley 59	0 02 14.91	+67 25 07.6	225	-1.61 ± 0.10	-1.92 ± 0.09	0.91 ± 0.04	1100_{-50}^{+50}
NGC 1333	3 29 07.96	+31 20 39.3	47	7.24 ± 0.28	-9.68 ± 0.13	3.36 ± 0.06	296_{-5}^{+5}
IC 348	3 44 33.88	+32 09 31.8	180	4.63 ± 0.14	-6.45 ± 0.13	3.09 ± 0.05	324_{-5}^{+5}
LkH α 101	4 30 10.17	+35 16 04.9	65	2.16 ± 0.20	-5.18 ± 0.42	1.77 ± 0.05	564_{-14}^{+15}
NGC 1893	5 22 53.74	+33 26 54.5	88	-0.24 ± 0.08	-1.40 ± 0.08	0.26 ± 0.04	3790_{-510}^{+700}
ONC	5 35 15.68	-05 23 40.1	378	1.51 ± 0.11	0.50 ± 0.12	2.48 ± 0.04	403_{-6}^{+7}
Mon R2	6 07 47.58	-06 22 42.6	97	-2.91 ± 0.11	1.05 ± 0.18	1.06 ± 0.04	948_{-38}^{+42}
Rosette	6 32 26.76	+04 47 37.1	468	-1.63 ± 0.07	0.15 ± 0.07	0.64 ± 0.04	1560_{-90}^{+110}
— NGC 2244	6 31 55.77	+04 55 7.8	272	-1.70 ± 0.07	0.20 ± 0.07	0.65 ± 0.04	1550_{-90}^{+100}
NGC 2264	6 40 57.93	+09 40 49.0	519	-1.76 ± 0.08	-3.72 ± 0.07	1.35 ± 0.04	738_{-21}^{+23}
— S Mon	6 40 49.83	+09 51 03.3	242	-1.62 ± 0.08	-3.71 ± 0.07	1.36 ± 0.04	738_{-21}^{+23}
— NGC 2264 IRS 2	6 41 00.64	+09 35 55.7	151	-2.29 ± 0.14	-3.61 ± 0.08	1.34 ± 0.04	748_{-23}^{+24}
— NGC 2264 IRS 1	6 41 07.03	+09 28 09.3	126	-2.05 ± 0.18	-3.90 ± 0.09	1.36 ± 0.04	736_{-22}^{+23}
NGC 2362	7 18 42.90	-24 57 44.0	246	-2.83 ± 0.07	2.95 ± 0.08	0.75 ± 0.04	1332_{-68}^{+75}
Carina	10 45 02.23	-59 41 59.8	1285	-6.55 ± 0.07	2.17 ± 0.07	0.38 ± 0.04	2620_{-250}^{+310}
— Tr 14	10 44 01.68	-59 32 48.1	401	-6.54 ± 0.07	2.06 ± 0.07	0.38 ± 0.04	2640_{-250}^{+310}
— Tr 15	10 44 42.08	-59 22 40.6	194	-6.25 ± 0.08	2.06 ± 0.08	0.38 ± 0.04	2630_{-250}^{+310}
— Tr 16	10 44 53.49	-59 43 10.1	268	-6.90 ± 0.07	2.63 ± 0.08	0.38 ± 0.04	2610_{-250}^{+310}
NGC 6231	16 54 15.90	-41 49 59.0	615	-0.55 ± 0.07	-2.17 ± 0.07	0.59 ± 0.04	1710_{-110}^{+13}
RCW 120	17 12 23.88	-38 29 15.7	29	-0.82 ± 0.12	-2.38 ± 0.19	0.59 ± 0.04	1680_{-110}^{+130}
NGC 6357	17 25 18.73	-34 17 24.8	178	-0.90 ± 0.08	-2.29 ± 0.10	0.56 ± 0.04	1770_{-120}^{+140}
— Pismis 24	17 24 44.06	-34 13 20.3	75	-0.83 ± 0.08	-2.08 ± 0.10	0.56 ± 0.04	1790_{-130}^{+150}
— G353.1+0.6	17 25 34.09	-34 24 49.5	53	-0.98 ± 0.09	-2.34 ± 0.11	0.56 ± 0.04	1780_{-130}^{+150}
— G353.2+0.7	17 25 59.77	-34 16 04.4	47	-1.09 ± 0.09	-2.66 ± 0.12	0.56 ± 0.04	1780_{-130}^{+150}
M20	18 02 23.10	-23 01 50.0	116	0.41 ± 0.12	-1.69 ± 0.09	0.79 ± 0.05	1264_{-68}^{+76}
NGC 6530	18 04 14.84	-24 21 45.9	669	1.32 ± 0.08	-2.07 ± 0.08	0.75 ± 0.04	1336_{-68}^{+76}
NGC 6611	18 18 42.20	-13 47 03.0	356	0.21 ± 0.08	-1.56 ± 0.08	0.57 ± 0.04	1740_{-120}^{+130}
M17	18 20 28.50	-16 10 58.0	82	-0.04 ± 0.17	-1.40 ± 0.13	0.60 ± 0.04	1680_{-110}^{+130}
IC 5146	21 53 31.47	+47 15 54.5	115	-2.87 ± 0.12	-2.52 ± 0.18	1.28 ± 0.04	783_{-25}^{+26}
NGC 7160	21 53 46.36	+62 35 07.6	71	-3.53 ± 0.10	-1.43 ± 0.13	1.04 ± 0.04	961_{-38}^{+41}
Cep OB3b	22 55 31.19	+62 38 54.1	678	-0.75 ± 0.09	-2.31 ± 0.08	1.16 ± 0.04	863_{-29}^{+31}
— Cep B	22 56 40.29	+62 42 06.2	482	-0.61 ± 0.10	-2.28 ± 0.09	1.15 ± 0.04	868_{-30}^{+32}
— V454 Cep	22 53 47.06	+62 35 46.6	196	-1.26 ± 0.11	-2.61 ± 0.11	1.18 ± 0.04	847_{-29}^{+31}

Note. Column 1: region name. Systems that are part of larger star-forming complexes are indented. Columns 2–3: coordinates of the system center. Column 4: the sizes of the sample of stars used in the analysis. Column 5–6: proper motion of the system center. Column 7: parallax of the system center.

The presence of a visual binary or the acceleration of a source can cause the *Gaia* astrometric solution to be rejected (Lindegren et al. 2018). Nevertheless, astrometric binaries remain a possible contributor to scatter in proper-motion distributions. The velocity dispersion induced by binaries depends on the individual stellar masses, binary separations, eccentricity of orbits, and inclinations of the systems. Binary orbital motions are unlikely to have a preferred orientation, so they should not bias observed bulk shifts in velocity, but they can contribute a high-velocity tail to velocity distributions.

The effect of binaries on velocity dispersions can be partially mitigated by filtering out sources with high astrometric excess noise. We use the ONC as a testbed to examine the link between binarity and astrometric excess noise. Duchêne et al. (2018) provided a list of ONC stars with and without companions at separations of 10–60 au, based on *HST* imaging. For the visual binaries, 50% have `astrometric_excess_noise` > 1 mas, while only 7% of the non-visual binary stars in their sample exceed this threshold. This result supports our decision to only use sources

below a 1 mas threshold for measuring median properties of stellar kinematics. We also note that the well-known O-star system, θ^1 Ori C, has `astrometric_excess_noise` > 1 mas and thus is not included in our sample. For applications that require accurate estimates of measurement uncertainty, such as measuring velocity dispersions, we limit the sample to data with `astrometric_excess_noise` = 0 mas.

2.3. Subclusters

Some of the systems we analyze as a unit in Sections 4–5 can be further decomposed into multiple subclusters. These include the clumpy distributions of stars that make up systems like NGC 6530 or M17 or smaller groups of stars surrounding a main cluster in systems like NGC 6611 and M20. In general, these subclusters contain too few stars to investigate their internal kinematics with DR2 data. Instead, we examine the relative velocities of different subclusters to test for signs of merger or dispersal (Section 6).

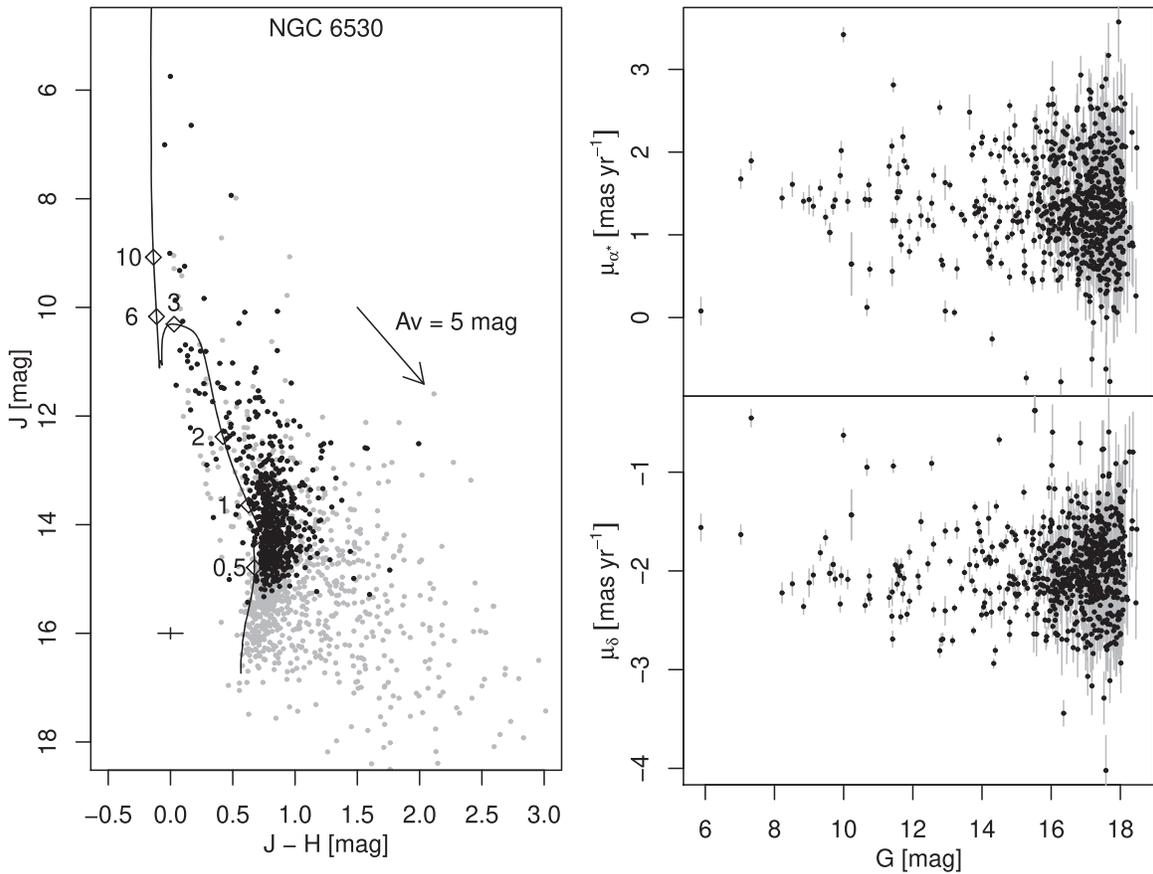


Figure 1. Left: J vs. $J - H$ color-magnitude diagram for sources in NGC 6530. *Gaia* sources used in the analysis are shown in black, while other MYStIX probable members are shown in gray. Only sources with uncertainties in J and $J - H$ less than 0.1 mag are plotted. A reddening vector (Rieke & Lebofsky 1985), a 2 Myr isochrone (Bressan et al. 2012) with several masses indicated, and a cross showing maximum allowed 1σ uncertainties are shown for guidance. Right: proper-motion measurements and uncertainties vs. G -band magnitude for YSOs in NGC 6530. Only sources included in the study are plotted.

(The complete figure set (21 images) is available.)

Subclusters in the star-forming regions studied here were cataloged by Kuhn et al. (2014) for MYStIX, Getman et al. (2018) for SFINCs, and Kuhn et al. (2017a) for NGC 6231. Subcluster identification in these papers was based on mixture models, a statistical cluster analysis method that is well adapted to cases where the size and density of clusters can vary and the number of clusters is uncertain (McLachlan & Peel 2000; Kuhn & Feigelson 2017). In some star-forming regions, subclusters can be found outside the main cluster, while in others, subclusters are clumps of stars that make up the main cluster.

3. Basic Properties of Stellar Systems

3.1. Parallaxes

Distances to stellar systems can be estimated assuming that their members span a small range of distances (see Bailer-Jones et al. 2018, their Section 3.5). This assumption is approximately true for our sample—the nearest region, NGC 1333, has a diameter only $\sim 0.4\%$ of the distance to the system. *Gaia* parallax measurements can then be considered to be random variables drawn from a distribution where the mean is ϖ_0 , the parallax of the system center; and the standard deviation is the measurement uncertainties given in the *Gaia* catalog. Using multiple stars to estimate the parallax of the center of a system, ϖ_0 , will yield a measurement with uncertainties smaller than

those for individual stars. However, gains in precision from pooling stars are limited due to correlated uncertainties of up to ~ 0.04 mas that are noticeable on spatial scales smaller than 1° (Lindgren et al. 2016).

We estimate the system parallax using the weighted median of individual stellar parallax measurements. This method is robust against contaminants while taking into account measurement uncertainties. For this analysis, we use the conventional $1/\text{error}^2$ weights and the weighted median function from the CRAN package *spatstat* (Baddeley et al. 2015) within the *R* statistical software environment (R Core Team 2018). Uncertainties on the weighted-median parallax are calculated using bootstrap analysis, with random sampling with replacement from the set of measurements (with added random errors) and weights. Finally, we add in quadrature the systematic uncertainty of 0.04 mas described by Lindgren et al. (2016), which provides a noise floor.⁵

Table 1 provides the new parallax estimates that we use in this study. Parallaxes are calculated independently for each

⁵ Several papers have proposed correction factors for systematic errors in *Gaia* astrometry (e.g., Kounkel et al. 2018; Muraveva et al. 2018; Stassun & Torres 2018), all of which are consistent with the systematic uncertainties reported by Lindgren et al. (2016). The parallax and proper-motion values reported in Table 1 are based on *Gaia* astronomy with no correction applied.

stellar system in a star-forming complex. In all cases, the uncertainties on median parallaxes are dominated by the systematic uncertainties in *Gaia* astrometry and not by statistical dispersion.

Several highlights from the revised distances in Table 1 are mentioned here.

1. Our analysis of the *Gaia* data places the ONC at a distance of 403^{+7}_{-6} pc, which is slightly closer than the estimate of 414 ± 7 pc from Menten et al. (2007), but farther than the 388 ± 5 pc found by Kounkel et al. (2017), both based on Very Long Baseline Array (VLBA) data. In Appendix A, we compare our distance estimate to 386 ± 8 pc from Kounkel et al. (2018) and discuss effects of the three-dimensional structure of Orion A.
2. Our distance estimate for NGC 1333 of 296 ± 5 pc is $\sim 25\%$ larger than the distance from Hirota et al. (2008) of 235 ± 18 pc. A larger distance to NGC 1333 will move stellar age estimates downward, which would resolve some anomalies in age estimates noted by Luhman et al. (2016) and Richert et al. (2018).
3. The *Gaia* distance estimate to M20 is 1260 ± 70 pc—a factor of ~ 2 nearer than the distance estimate of 2700 pc from Cambr esy et al. (2011). The revised distance places M20 at a similar distance as NGC 6530, from which it is separated by only 1.4° on the sky.
4. In star-forming regions made up of multiple stellar systems, the nearly identical parallax measurements for the different components provide confidence in the accuracy of the measurements and reassurance that these systems are physically associated. For example, we have confirmed that Tr 14, 15, and 16 are all at a distance of ~ 2600 pc, putting to rest claims in older literature that Tr 15 is in the background (see Walborn 1995).

3.2. Proper Motions

The proper motions of the system centers, $\mu_{\alpha^*,0}$ and $\mu_{\delta,0}$, are estimated using the same weighted-median strategy as above. For systematic uncertainties due to correlated errors, we use ± 0.07 mas yr⁻¹ (Lindgren et al. 2016). These proper motions are reported in Table 1. Given that the systems each subtend $< 1^\circ$ on the sky, we ignore spherical geometry effects in determining the median—the small corrections calculated below for individual stars do not affect the proper motion of the system center.

The effects of correlated errors in *Gaia* DR2 on the relative velocities of stars within systems was investigated in simulations by Bianchini et al. (2018). They found that the effects were $< 0.5 \mu\text{as yr}^{-1}$ on length scales of 10 arcmin, which suggest that these effects will be negligible in our study of internal kinematics.

3.3. Refined Sample

Parallax and proper-motion measurements allow us to identify likely field-star contaminants in the MYStIX/SFiNCs/NGC 6231 catalogs. Stars are removed from the sample if their parallax measurements are inconsistent with the median parallax by more than three standard deviations, taking into account uncertainty on the measurement and on the median. Outliers in proper motion by more than > 5 standard

deviations (using the maximum absolute deviation as a robust estimator of standard deviation) are also filtered out as likely contaminants.

Estimation of median parallax and proper motion and refinement of membership are performed iteratively until convergence. For example, in NGC 6231 with 1760 *Gaia* counterparts, 121 (7% of the total) are removed as likely contaminants, leaving 1639 bona fide system members in *Gaia*. (Only 615 of these have sufficient astrometric precision for inclusion in the analysis.) Overall, contamination rates were about 13%, with contamination rates for individual systems mostly falling into the range 7%–15%. Several systems with much higher contamination among the *Gaia* sources include M17 (23%), NGC 2362 (25%), NGC 7160 (28%), and RCW 120 (38%). Table 1 gives the number of stars in the final, refined sample.

3.4. Stellar Kinematics

We are interested in obtaining two components of stellar velocities (from a three-dimensional Cartesian coordinate system) based on the proper motions μ_{α^*} and μ_{δ} . Since the systems in our sample are sufficiently distant, compact, and slow-moving, the relative proper motions of their stars are dominated by their physical velocities, so only small correction factors, calculated below, are necessary.

The observed proper motions of stars in stellar systems can be affected by perspective and the motion of the center of the system (van Leeuwen 2009). In particular, radial motion of a system can produce an effect known as “perspective expansion,” which is seen in *Gaia* measurements of globular clusters (Gaia Collaboration et al. 2018), many of which have RVs of hundreds of kilometers per second. A first-order approximation to the perspective expansion can be obtained from Equation (13) in van Leeuwen (2009). Here, α_0 , δ_0 , ϖ_0 , $\mu_{\alpha^*,0}$, and $\mu_{\delta,0}$ are the astrometric parameters of the system center, and $\Delta\alpha_i$ and $\Delta\delta_i$ are the difference in R.A. and decl. between an individual star and the system center. The equations for the additional shift in the proper motion of a star are

$$\Delta\mu_{\alpha^*,\text{per}} \approx \Delta\alpha_i \left(\mu_{\delta,0} \sin \delta_0 - \frac{v_r \varpi_0}{\kappa} \cos \delta_0 \right), \quad (1)$$

$$\Delta\mu_{\delta,\text{per}} \approx -\Delta\alpha_i \mu_{\alpha^*,0} \sin \delta_0 - \Delta\delta_i \frac{v_r \varpi_0}{\kappa}, \quad (2)$$

where $\kappa = 4.74$ is the conversion factor from mas yr⁻¹ to km s⁻¹ at a distance of 1 kpc. The first term in each equation relates to motion in a spherical coordinate system while the second term relates to the apparent expansion/contraction of an object as it gets farther/nearer.

These equations, calculated using RVs in Appendix B, contribute small shifts to apparent stellar proper motions. Following the strategy recommended by Brown et al. (1997), these contributions can be subtracted from the observed proper motions (relative to the system center) to isolate the effects of internal kinematics. The region with the largest corrections is the Orion A cloud with corrections on the order of ~ 0.07 mas yr⁻¹. Most other regions have corrections with magnitudes < 0.02 mas yr⁻¹. Similarly, for regions spanning $< 1^\circ$, the projection effects of spherical geometry are small ($< 0.02\%$) and are not expected to affect science results.

The correction factors computed above allow us to approximate a two-dimensional velocity of each star, $\mathbf{v} = (v_x, v_y)$,

relative to the rest frame of the center of the system. First, we calculate the components of the corrected velocities parallel to lines of constant R.A. and decl.,

$$v_{\alpha} \approx -\kappa \left(\frac{\Delta\mu_{\alpha^*}^{\text{obs}} - \Delta\mu_{\alpha^*}^{\text{per}}}{\varpi_0} \right), \quad (3)$$

$$v_{\delta} \approx \kappa \left(\frac{\Delta\mu_{\delta}^{\text{obs}} - \Delta\mu_{\delta}^{\text{per}}}{\varpi_0} \right). \quad (4)$$

Then, these velocities are transformed to velocities in a Cartesian coordinate system, v_x and v_y , using the orthographic projection (Gaia Collaboration et al. 2018b). Uncertainties on v can be described by a covariance matrix $\Sigma_{\text{err},i}$. This is obtained from the covariance matrix for the astrometric solution, reconstructed from the DR2 catalog using Equation B.3 in Lindegren et al. (2018), multiplied by a correction factor of 1.1^2 (Section 5.2 in Lindegren et al. 2018), and scaled by $(\kappa/\varpi)^2$ to convert angular motions into velocities. This covariance matrix does not include the systematic effects of uncertainties in system parallax or RV.

Velocities can also be expressed in different coordinate systems. For example, the outward and azimuthal components of the projected velocity with respect to the center of the system are

$$v_{\text{out}} = \mathbf{v} \cdot \hat{\mathbf{r}}, \quad (5)$$

$$v_{\text{az}} = \mathbf{v} \cdot \hat{\boldsymbol{\phi}}, \quad (6)$$

where $\hat{\mathbf{r}}$ is a unit vector pointing away from the system center and $\hat{\boldsymbol{\phi}}$ is a unit vector pointing in the azimuthal direction relative to the system center. The uncertainty on a velocity component in the $\hat{\mathbf{u}}$ direction can be obtained from the covariance matrix

$$\sigma_{\text{err},i} = (\hat{\mathbf{u}}^T \Sigma_{\text{err},i} \hat{\mathbf{u}})^{1/2}. \quad (7)$$

3.5. Visualizations of Kinematics

Several approaches to the visualization of the observed four-dimensional kinematic structure of stellar systems are shown in Figures 2–4. The ONC and NGC 6530 are shown as examples, but plots for the other 26 stellar systems are available in the figure sets. In each case, we depict velocities in the rest frames of their system centers.

The kinematics of stellar systems can be displayed using arrows to show the velocity vectors for each star, as in Figure 2. In regions with large numbers of stars, these produce crowded plots that are difficult to interpret. The figure highlights sources with the highest quality proper-motion measurements.

Figure 3 shows the direction of motion (no amplitude) for individual stars. The color of the arrowhead is also determined by the direction of motion (as indicated by the compass wheel) for clearer visualization. In an area of the diagram where arrowheads are mostly one color, the stars are mostly moving in one direction. The color saturation of the symbol is related to the statistical weighting of the data point. Only points with uncertainties $<3 \text{ km s}^{-1}$ are shown, with the darkest points representing points with uncertainties $<0.5 \text{ km s}^{-1}$. The X's mark the system centers used for measuring outward velocities v_{out} .

Figure 4 shows the relationship between star positions and velocities. For each system, the four scatter plots show velocities perpendicular and parallel to each coordinate axis.

A simple radial contraction or expansion will produce a gradient in the $v_x - \text{R.A.}$ and $v_y - \text{decl.}$ diagrams (upper left and lower right) but will not affect the other diagrams (upper right and lower left). The different patterns of stellar motion within different systems are analyzed in the next section.

4. Bulk Stellar Motions

4.1. Contraction versus Expansion Velocities

The expansion or contraction of a stellar system would produce a bulk outward or inward motion of its stars. In studies of stellar systems in star-forming regions, positive correlations between cluster (or subcluster) size and age suggest that these systems expand with velocities of $\sim 0.25\text{--}2 \text{ km s}^{-1}$ (Pfalzner 2009; Kuhn et al. 2015a; Getman et al. 2018), where the slope of the relation is interpreted as an expansion velocity. Expansion velocities in this range would be detectable with *Gaia* in the star-forming regions in our sample.

Velocities that are preferentially oriented outward can be seen in some stellar systems and not in others. Expansion would show up as arrows pointing radially outward in Figure 2, coherent patterns in arrow direction and hue in Figure 3, and correlations in positions and velocity in Figure 4.

NGC 6530 exhibits these characteristics. Stellar velocity vectors are primarily oriented away from the center, making the arrows in Figure 2 point preferentially outward and showing up as a gradient in the color of arrowheads in Figure 3. There are also correlations between v_x and α , and v_y and δ (Figure 4) that are consistent with the expectations for an expanding system. In contrast, for the ONC, these figures show that stars with differently oriented velocity vectors are more mixed up, and there is no visually obvious correlation between position and velocity of stars.

We quantify the system expansion (or contraction) using the weighted-median value of v_{out} for each system using all members of the “refined sample” of stars. Uncertainties on the median are calculated by bootstrap resampling, as earlier. The expansion velocities are provided in Table 2.

The distribution of v_{out} values that go into calculating the median are illustrated in Figure 5, with weighted kernel-density estimates (KDE) of v_{out} obtained with the function density in R . Shifts of the distribution toward positive values indicate expansion while shifts to negative values indicate contraction.

In our sample of systems, median v_{out} values are more likely to be directed outward than inward. Figure 6 shows a histogram of these expansion/contraction velocities and a histogram of the signal-to-noise ratio, calculated as the ratio of the median v_{out} to the estimated uncertainty on the median v_{out} . Although most systems show effect sizes $<3\sigma$, the distribution is clearly shifted to the right from what would be expected if systems had zero expansion, and all non-zero measurements were due to measurement uncertainty. There are six systems with statistically significant outward velocities at the $>3\sigma$ level, while no individual systems show statistically significant inward velocities at this level. The velocity shifts range from -2.0 to $+2.0 \text{ km s}^{-1}$, but are predominantly positive with a mean value of 0.5 km s^{-1} .

Systems where expansion is detected at the $>3\sigma$ level include NGC 1893, NGC 2244, Tr 16, NGC 6530, NGC 6611, and Cep B. In these regions, the expansion pattern is often visually apparent on the plots in Figures 2–4. These cases happen to be the ones with high expansion velocities greater than 0.9 km s^{-1} .

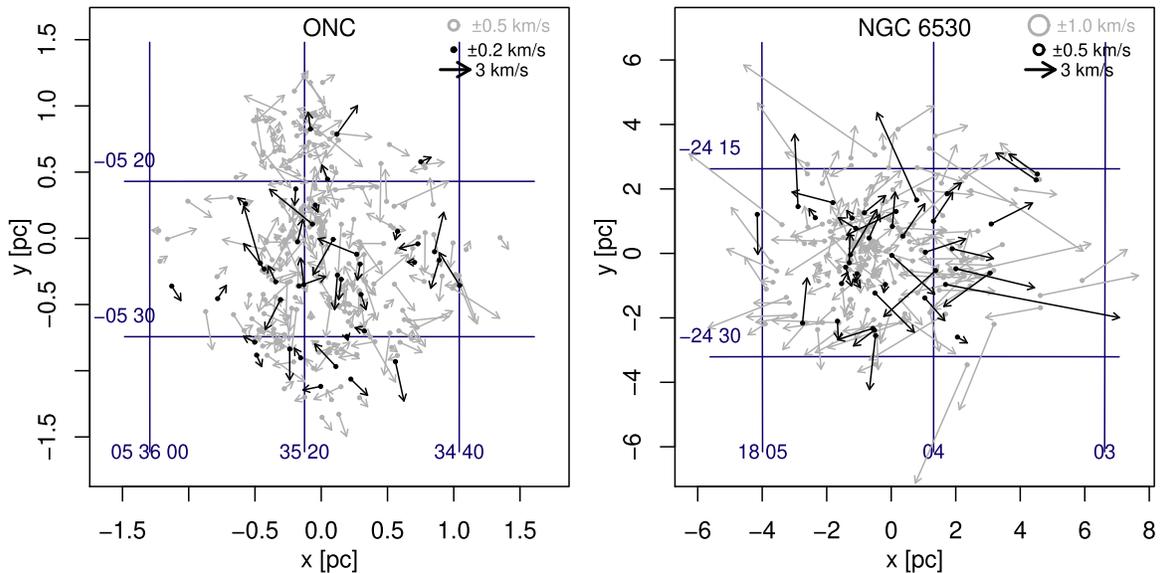


Figure 2. Maps of the velocity vectors (magnitude and direction) for the best quality *Gaia* data in the ONC (left) and NGC 6530 (right). The quality of the velocity measurements is indicated by the arrow color (black for most precise and gray for less precise), with the maximum allowed uncertainties for each group shown by the circles in the legend.

(The complete figure set (28 images) is available.)

However, a number of other systems have measured expansion velocities on the order of $\sim 0.4 \text{ km s}^{-1}$, but the magnitude of the expansion velocity is not sufficient to reach the $>3\sigma$ level given typical uncertainties on an expansion velocity of $\sim 0.2 \text{ km s}^{-1}$. This group includes the ONC, S Mon, Tr 14, Tr 15, Pismis 24, G353.1+0.6, IC 5146, and V454 Cep. Overall, it is likely that most of the systems in this second group have mild expansion (Section 4.3), but the evidence for expansion of any individual system is not definitive.

Not all systems are likely to be in a state of expansion. The systems NGC 1333, IC 348, NGC 2264 IRS 2, NGC 2362, and NGC 7160 show no evidence for either expansion or contraction, and M17 and NGC 6231 show signs of contraction, albeit not at a high significance level. M17 has a fairly fast contraction rate (-2 km s^{-1}), but the *Gaia* data are limited by high extinction in this region, so the result is significant only at the 2σ level. NGC 6231 has a fairly slow contraction velocity of -0.2 km s^{-1} . Contraction of this system would be interesting because NGC 6231 is physically larger than most systems, giving it the appearance of having expanded in the past. Finally, the systems Berkeley 59, LkH α 101, Mon R2, NGC 2264 IRS 1, RCW 120, G353.2+0.7, and M20 have large-enough uncertainties relative to the expansion/contraction velocity that results are ambiguous.

4.2. Radial Dependence of Expansion Velocity

Models for the expansion of OB associations suggest linear relationships between expansion velocity and distance of stars from a system center (Blaauw 1964; Brown et al. 1997). This occurs because, in an unbound system, faster stars will travel farther from their point of origin, causing the stars to spatially sort themselves by velocity.

Figure 7 shows the expansion velocity as a function of radius. For this plot, stars are binned by radial distance using bin sizes of ~ 60 stars, and expansion velocities are estimated using the same method as above. We fit the points with a line of

the form $y = ax + b$ using weighted least-squares regression, where weights are proportional to the reciprocal of the uncertainty squared. The slopes of the regression lines have units of $\text{km s}^{-1} \text{ pc}^{-1}$ ($\approx \text{Myr}^{-1}$) and intercepts have units of km s^{-1} .

For the ONC, which shows evidence for only mild expansion, all points are consistent with the mean value. Both slope (0.0 ± 0.4) and intercept (0.4 ± 0.3) are consistent with there being either no or mild expansion. For NGC 6530, most of the points follow a positive linear relationship between radius and expansion velocity, and the two points that deviate from this relationship have large uncertainties. The slope of 0.6 ± 0.2 is statistically significant, while the intercept 0.0 ± 0.4 is not statistically significant. The regression analysis, using the `lm` function in *R*, gives a p -value of 0.002, providing strong evidence that NGC 6530 has a radially dependent expansion velocity.

Several other systems fall into each class. Examples that illustrate possible linear relationships between radius and expansion rate include Cep B ($p = 0.02$), NGC 2244 ($p = 0.003$), and Tr 16 ($p = 0.08$). For these systems, the slopes of the relationships range from ~ 0.5 to $1 \text{ km s}^{-1} \text{ pc}^{-1}$. NGC 6611 also shows the fastest expansion in the outer regions, consistent with this pattern, but the slope of the relation is not statistically significant ($p = 0.06$). Others systems, like Berkeley 59, NGC 2362, NGC 6231, S Mon, Tr 14, and Tr 15, show flat relationships—the latter group includes both systems with and without evidence for net expansion.

4.3. Fraction of Systems that are Expanding

Overall, 21 out of 28 systems (i.e., 75%) have positive values of median v_{out} . However, calculating the fraction of systems that are expanding is complicated by measurement errors, which can change the sign of median v_{out} for systems with small or zero expansion velocity. To evaluate this effect, we include uncertainties on the median v_{out} in a statistical

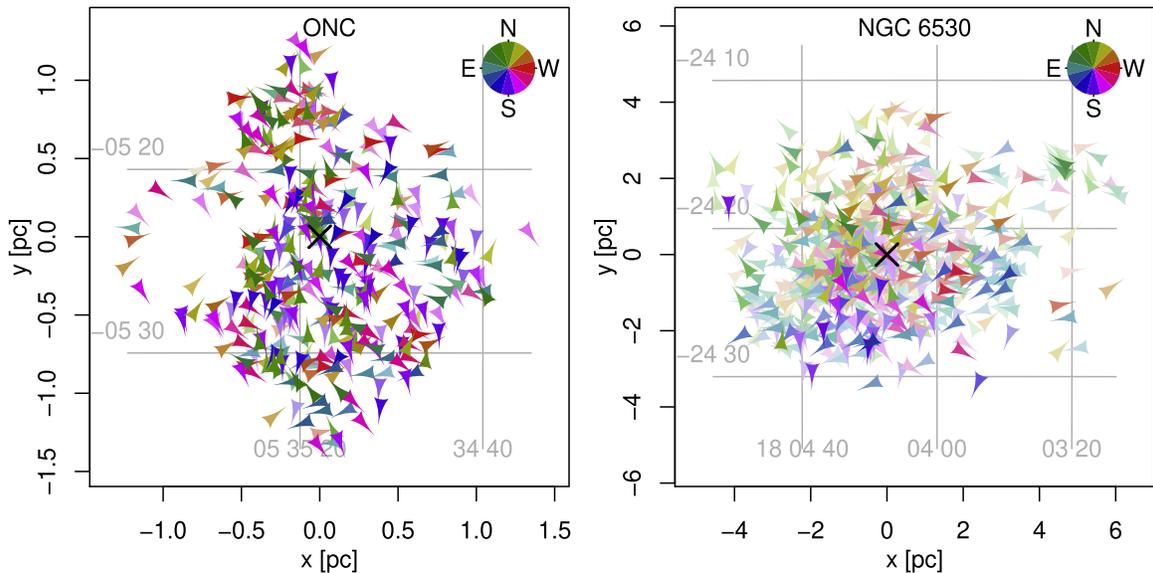


Figure 3. Direction of motion of individual stars in the rest frame of the system. Diagrams are shown for the ONC (left) and NGC 6530 (right). The orientations of the arrows and their hues indicate their direction, while the saturation indicates weighting based on statistical uncertainty. In the ONC, stars with different velocities are mixed together, while in NGC 6530, many stars have directions of motions away from the system center (as indicated by the outward pointing arrows and color segregation by azimuth).

(The complete figure set (28 images) is available.)

model akin to the Extreme Deconvolution method of Bovy et al. (2011). The construction of such a model is described more thoroughly in Section 5, where it is used for a different application.

The observed distribution of expansion velocities (Figure 6) can be modeled as an intrinsic distribution convolved with measurement uncertainties. We examine several models for the intrinsic distribution, which are fit to the data using the Monte Carlo Markov Chain (MCMC) method. We first try modeling the distribution of the median v_{out} as a single Gaussian, which yields an $f = 0.86$ for systems in expansion (0.65–0.98, 95% credible interval). We next use a mixture of two Gaussians, noting that Gaussian mixture models can be used as flexible models for unknown probability density functions (e.g., Kelly et al. 2008). This model yields a fraction of $f = 0.88$ (0.79–0.94, 95% credible interval). For MCMC analysis, we use “non-informative” priors for mixture models: for the standard deviation of the Gaussians, we use a uniform distribution between 0 and 6, for the mixing parameters we use a Dirichlet distribution with $\alpha = 1$, and for the mean we use a uniform distribution between -3 and 3 . We use the Metropolis–Hastings algorithm to sample from the posterior. Through experimentation, we find that variations in the functional form of the priors have relatively minor effects on the results.

Our conclusion from this analysis is that our assumptions on the distribution of expansion velocities for systems have little effect on the fraction that are expanding ($\sim 85\%$) and that the effect of measurement errors is to slightly decrease the fraction of systems observed undergoing expansion. Thus, we can reliably claim that $\gtrsim 75\%$ of systems in our sample are expanding.

4.4. Cluster Rotation

The stellar velocity measurements can also be used to look for evidence of bulk rotation. The angular momentum of a star

with mass m and velocity \mathbf{v} at a position \mathbf{r} relative to the center of the system is

$$\mathbf{L} = \mathbf{r} \times m \mathbf{v}, \quad (8)$$

so the component of the angular momentum along the line of sight is

$$L \cos i = -m R v_{\text{az}}, \quad (9)$$

where i is the inclination of the angular momentum vector and R is the projected distance of the star from the center of the system. Thus, a non-zero median value of v_{az} can indicate bulk rotation of a system. For v_{az} , we use the same method to calculate the median and error on the median as what we used for v_{out} . These values are provided in Table 2.

For the 28 systems in the sample, the values of the median v_{az} are distributed around zero with an average of $0.06 \pm 0.10 \text{ km s}^{-1}$. For most systems, the value of the median v_{az} is within 2σ of $v_{\text{az}} = 0 \text{ km s}^{-1}$. A few cases have more statistically significant values, including LkH α 101, the ONC, Tr 14, and G353.2+0.7 at 2σ significance, and Tr 15 (median $v_{\text{az}} = 1.7 \text{ km s}^{-1}$) at $>3\sigma$ significance. Typical uncertainties on median azimuthal velocities are $<0.4 \text{ km s}^{-1}$, so rotational velocities less than this value may not be detectable.

Under the assumption that the median values of v_{az} are all results of measurement uncertainty, the ratios of these values to their uncertainties should be drawn from a normal distribution with a mean 0 and standard deviation 1. Figure 8 shows a histogram of the ratio of observed rotation to measurement uncertainty, which is compared to the normal distribution expected given the null hypothesis. The distributions are remarkably similar, suggesting that no real rotation is detected in most of the stellar systems. However, Tr 15 is difficult to explain using the null hypothesis, so rotation may be real in this individual case.

Although bulk rotation has been reliably measured in globular clusters (e.g., Bianchini et al. 2018; Kamann et al. 2018), very

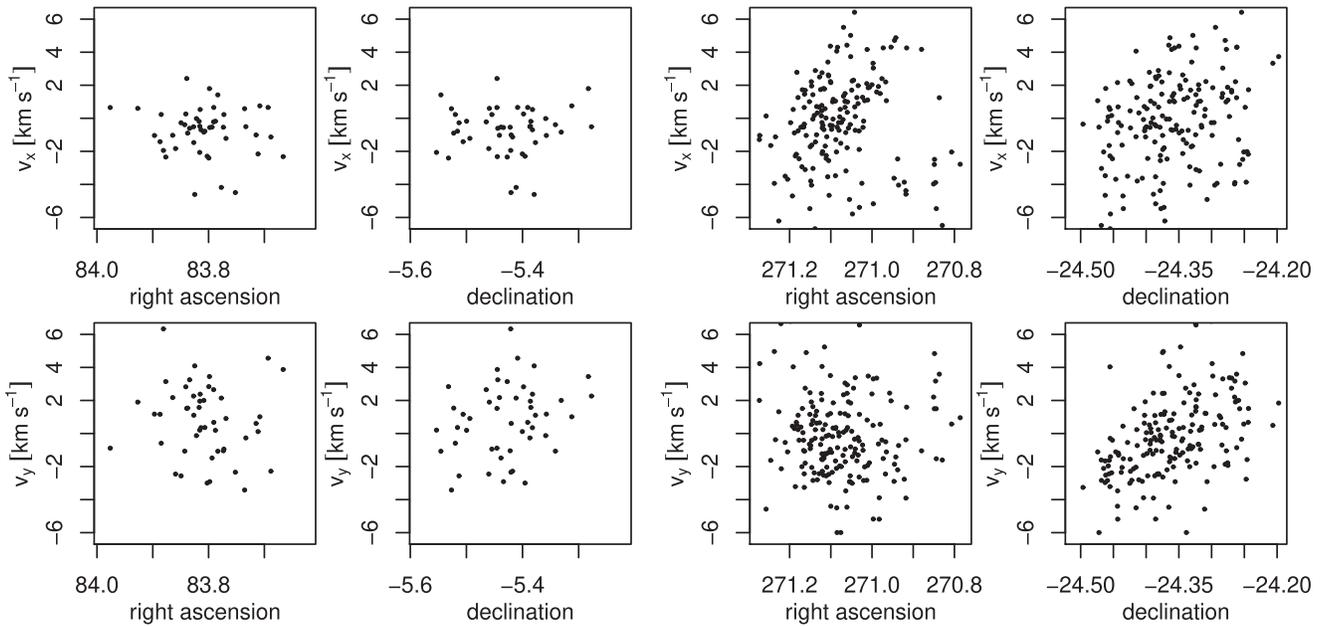


Figure 4. Dispersions of stars in position and velocity for the ONC (left) and NGC 6530 (right). For each system, the four panels show all combinations of R.A., decl., v_x , and v_y . (Recall that v_x is motion east to west and v_y is motion south to north). All panels show the same velocity range to facilitate comparison of velocity dispersions. Expansion may show up as a correlation between v_x and R.A. or between v_y and decl. Only stars with no astrometric excess noise are included to reduce the presence of velocity outliers.

(The complete figure set (24 images) is available.)

few attempted measurements exist in the literature for open clusters. A recent study of *Gaia* DR1 data by Reino et al. (2018) concluded that there was no evidence for bulk rotation in the ~ 600 Myr old Hyades cluster, based on measurement of azimuthal motion at only the 2σ level. Our results are consistent with cluster rotation being rare, but we are less sensitive to rotational velocities below several tenths of a km s^{-1} .

5. Velocity Dispersions

Calculation of the velocity dispersion is particularly sensitive to measurement errors, which can broaden the observed distribution. The astrometric measurement uncertainties reported in the *Gaia* catalog are heteroscedastic and comparable to the velocity dispersion, so their effect must be carefully modeled. For *Gaia* sources with `astrometric_excess_noise` > 0 , the statistical uncertainties derived from the *Gaia* DR2 AGIS model may not represent the real errors in relative proper motions (Lindegren et al. 2012). Thus, we use only sources with `astrometric_excess_noise` = 0 when modeling velocity distributions.

We use the maximum likelihood to estimate the intrinsic velocity dispersion in the presence of measurement error (see Walker et al. 2006). We model the observed velocity of a star i as the sum of its intrinsic velocity \mathbf{v}_i and an error term ϵ_i ,

$$\mathbf{v}_{\text{obs},i} = \mathbf{v}_i + \epsilon_i. \quad (10)$$

Both \mathbf{v}_i and ϵ_i are assumed to be multivariate normally distributed, with

$$\mathbf{v}_i \sim N(\boldsymbol{\mu}_v, \Sigma_v), \quad (11)$$

$$\epsilon_i \sim N(0, \Sigma_{\text{err},i}), \quad (12)$$

where $\boldsymbol{\mu}_v$ and Σ_v are the mean and covariance matrix of the intrinsic velocity distribution, and $\Sigma_{\text{err},i}$ is the covariance matrix for the measurement uncertainty on the i th star. Then the log-likelihood is

$$\mathcal{L}(\boldsymbol{\mu}_v, \Sigma_v | \mathbf{v}_{\text{obs},i}, \Sigma_{\text{err},i}) = \sum_{i=1}^N \log \phi(\mathbf{v}_{\text{obs},i} | \boldsymbol{\mu}_v, \Sigma_v + \Sigma_{\text{err},i}), \quad (13)$$

where ϕ is the probability density of the normal distribution. The maximum likelihood $\boldsymbol{\mu}_v$ and Σ_v parameters can be found by numerical optimization. We used the BFGS algorithm (Fletcher & Reeves 1964) implemented in the *R* function `optim`. Examination of the likelihood function shows it to be approximately normal, so we use `optim` to numerically calculate the Hessian matrix at the maximum (also called the Fisher Information Matrix) and invert it to estimate the covariance matrix.

5.1. Velocity Phase Space

The total velocity dispersions of the individual stellar systems will incorporate both the bulk expansion velocity (if non-zero) and a velocity spread due to the orbital motions of individual stars. Figure 9 shows stars plotted in velocity phase space with coordinates (v_x, v_y) , where the ONC and NGC 6530 are used as examples.

The observed velocity distributions are not entirely isotropic. Stars in the ONC are preferentially moving north or south, rather than east or west, while stars in NGC 6530 are preferentially moving east or west, rather than north or south. However, the referencing of stellar motions to a frame that is based on the equatorial coordinate system is arbitrary.

The Gaussian model of the velocity distribution provides two velocity dispersion components. These are the two

Table 2
Bulk Expansion and Rotation Velocities

Region	n_{samp} (stars)	Median v_{out} (km s^{-1})	Median v_{az} (km s^{-1})
(1)	(2)	(3)	(4)
Berkeley 59	225	0.36 ± 0.24	-0.39 ± 0.31
NGC 1333	47	0.23 ± 0.28	-0.45 ± 0.21
IC 348	180	0.16 ± 0.18	-0.27 ± 0.23
LkH α 101	65	0.97 ± 0.68	0.72 ± 0.33
NGC 1893	88	1.34 ± 0.32	0.19 ± 0.53
ONC	378	0.43 ± 0.20	-0.30 ± 0.14
Mon R2	97	-0.12 ± 0.53	0.58 ± 0.30
NGC 2244	272	1.23 ± 0.17	-0.10 ± 0.22
S Mon	242	0.39 ± 0.15	-0.03 ± 0.07
NGC 2264 IRS 2	151	-0.27 ± 0.28	-0.13 ± 0.21
NGC 2264 IRS 1	126	0.36 ± 0.40	-0.25 ± 0.32
NGC 2362	246	0.02 ± 0.28	-0.10 ± 0.15
Tr 14	401	0.39 ± 0.34	0.52 ± 0.19
Tr 15	194	0.64 ± 0.38	1.72 ± 0.47
Tr 16	268	0.84 ± 0.22	-0.06 ± 0.43
NGC 6231	615	-0.23 ± 0.14	-0.09 ± 0.17
RCW 120	29	-0.28 ± 1.45	-0.12 ± 0.50
Pismis 24	75	0.91 ± 0.44	0.54 ± 0.44
G353.1+0.6	53	2.07 ± 1.10	0.07 ± 0.52
G353.2+0.7	47	-0.24 ± 0.65	-0.95 ± 0.48
M20	116	0.33 ± 0.37	-0.21 ± 0.35
NGC 6530	669	0.99 ± 0.19	-0.29 ± 0.15
NGC 6611	356	0.90 ± 0.23	0.29 ± 0.18
M17	82	-2.06 ± 1.00	-0.06 ± 1.09
IC 5146	115	0.48 ± 0.25	0.05 ± 0.49
NGC 7160	71	-0.20 ± 0.30	0.21 ± 0.18
Cep B	482	0.95 ± 0.29	0.39 ± 0.14
V454 Cep	196	0.55 ± 0.34	-0.14 ± 0.35

Note. Column 1: system name. Column 2: number of stars used to calculate median velocities. Column 3: median v_{out} —measure of expansion or contraction. Column 4: median v_{az} —measure of rotation.

principal components of the velocity dispersion, where the first component is defined to be the one with the largest variance. Thus, σ_{pc1} is the semimajor axis of the ellipses in Figure 9, while σ_{pc2} is the semiminor axis. The use of Equation (13) allows for the heteroscedastic uncertainties to be taken into account in the principal component analysis. Table 3 provides σ_{pc1} and σ_{pc2} , as well as their ratio, the position angle of the first principal component, and statistical uncertainties.

The ratios of σ_{pc1} to σ_{pc2} show that most systems do not have statistically significant velocity anisotropy (i.e., values that are significantly greater than 1). However, the ONC has $\sigma_{\text{pc1}}/\sigma_{\text{pc2}} = 1.5$ and NGC 6530 has $\sigma_{\text{pc1}}/\sigma_{\text{pc2}} = 1.5$, both significant at $\sim 2.5\sigma$. In both cases, the orientation of the velocity anisotropy is approximately aligned with the spatial elongation of the system.

5.2. Shape of the Velocity Distribution

We can examine the shape of the velocity distributions by comparing to bivariate normal distributions. Henze & Zirkler (1990) provided a test of multivariate normality, implemented in the *R* package MVN (Korkmaz et al. 2014). The systems where velocity distributions are consistent with multivariate normal distributions include Berkeley 59, the ONC, IRS 1, and IC 5146 ($p > 0.05$). The systems Mon R2, NGC 2244, NGC 7160, and Cep B show moderate statistically significant deviation from normality ($0.05 < p < 0.001$), while NGC 1893, S Mon, IRS 2,

NGC 2362, Tr 14/15/16, NGC 6231, M20, NGC 6530, and NGC 6611 show large statistically significant deviations from normality ($p < 0.001$). We note that this hypothesis test is more sensitive to deviations when sample sizes are larger, and it does not indicate how the distribution deviates from normality.

Deviations from normality can be visualized using plots of observed data quantiles versus theoretical quantiles (the $Q-Q$ plot). The data quantiles are the differences between each measurement and the mean value, normalized by the standard deviation,

$$Q_{\text{data},i} = \frac{v_{\text{obs},i} - \mu}{\sqrt{\sigma_v^2 + \sigma_{\text{err},i}^2}}, \quad (14)$$

while the theoretical quantiles are

$$Q_{\text{theo},i} = \sqrt{2} \operatorname{erf}^{-1}(2(r_i - 0.5)/n - 1), \quad (15)$$

the quantile function of the Gaussian distribution, where n is the number of velocity measurements, r_i is the rank of the i th measurement, and erf^{-1} is the inverse of the error function. We calculate a test envelope (95%) for the null hypothesis that the data are correctly described by our model using Monte Carlo simulations.

$Q-Q$ plots are produced for each velocity component (Figure 10). The ONC velocity distribution is remarkably well fit by a normal distribution. This result is astrophysically interesting in itself (see Section 8), but also implies that standard deviations of the distribution calculated using Equation (13) will be reliable.

For NGC 6530, the distribution closely follows a normal distribution out to $\pm 2\sigma$, but beyond this threshold, there is a significant excess of stars with higher velocities than would be expected from a normal distribution. Even a small number of outliers can have a large effect on estimates of the standard deviation, so the observed deviations from a normal distribution suggest that a standard deviation estimated from Equation (13) will overestimate the width of the distribution.

For nearly half the systems, the shape of the observed velocity distribution indicates the presence of outliers. The nature of the outliers is uncertain. While these could represent a population of higher velocity stars that are escaping the system at a faster rate, they could also represent points with large errors not represented well by the AGIS uncertainties or field stars contaminating the “refined sample.” In Table 3, the more distant systems tend to be the ones where outliers are detected. While the astrometric excess noise parameter is effective at flagging known binary stars in the ONC, it is not likely to be as effective for more distant regions. Overall, the outliers do not seem to have a preferential spatial location within the systems.

A possible strategy to cope with outliers is to use a mixture model to represent the main population and the outliers as two distinct Gaussian components. In our model, these components will have the same means but different covariance matrices (e.g., Bravi et al. 2018). The complete data likelihood for this mixture model is

$$\begin{aligned} \mathcal{L}(\boldsymbol{\mu}_v, \Sigma_{v,1}, \Sigma_{v,2}, \eta | \mathbf{v}_{\text{obs},i}, \Sigma_{\text{err},i}) \\ = \sum_{i=1}^N \log((1 - \eta) \phi(\boldsymbol{\mu}_v, \Sigma_{v,1} + \Sigma_{\text{err},i}) \\ + \eta \phi(\boldsymbol{\mu}_v, \Sigma_{v,2} + \Sigma_{\text{err},i})), \end{aligned} \quad (16)$$

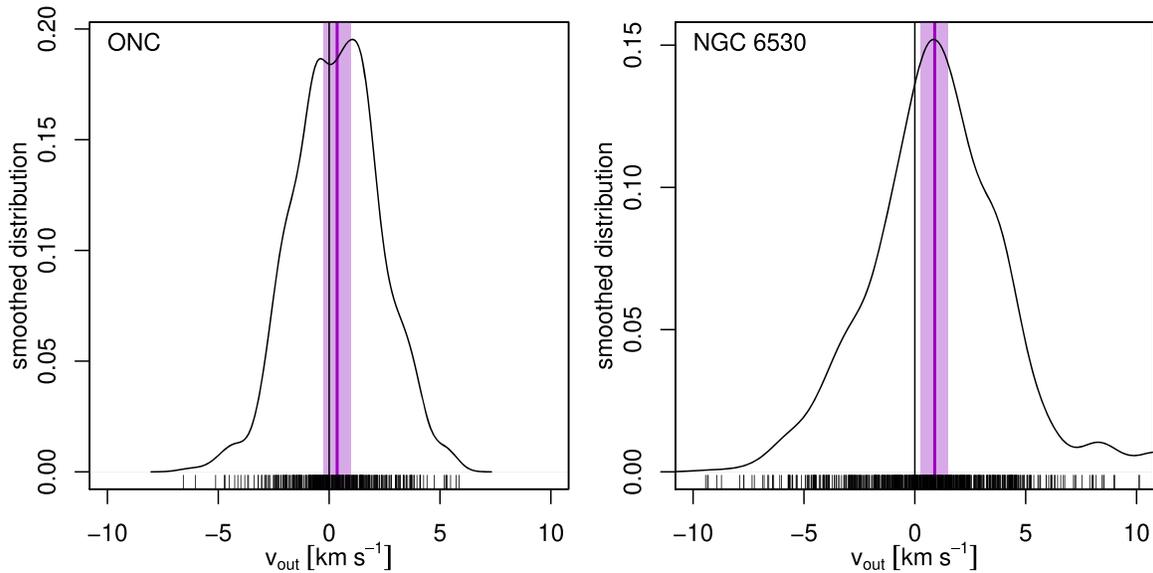


Figure 5. Distributions of v_{out} shown as KDE plots for stars in the ONC (left) and NGC 6530 (right). The black lines indicate $v_{\text{out}} = 0 \text{ km s}^{-1}$, the solid magenta lines indicate the median of the distribution, and the shaded magenta regions indicate the 3σ uncertainty on the median. Medians greater than 0 indicate bulk expansion while medians less than zero indicate bulk contraction.

(The complete figure set (28 images) is available.)

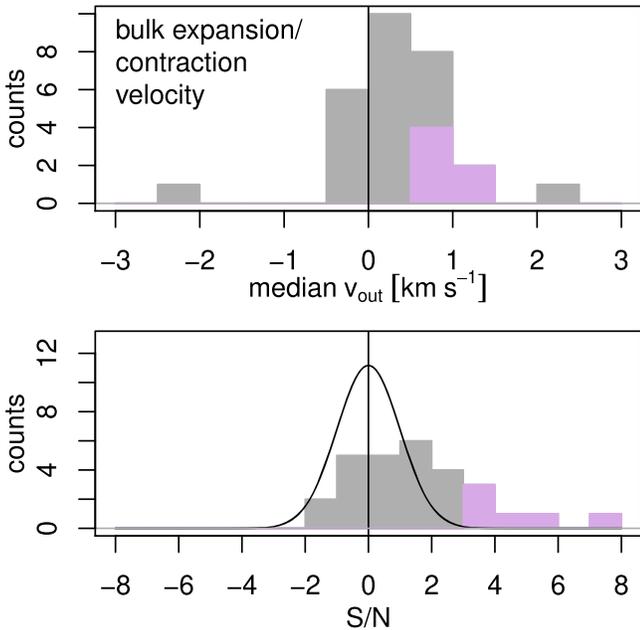


Figure 6. Histograms of the expansion velocities (top) and signal-to-noise ratio of the expansion velocities (bottom) for the stellar systems in our sample. Expansion significant at the $>3\sigma$ level is marked in magenta. The Gaussian curve in the bottom plot shows the distribution that would be expected under the null hypothesis that all measurements of expansion and contraction were products of measurement uncertainty.

where $\Sigma_{v,1}$ and $\Sigma_{v,2}$ are the covariance matrices describing the velocity distribution and $0 < \eta < 1$ is the mixing parameter. For the second component, we use a radially symmetric normal distribution (i.e., $\Sigma_{v,2} = a I_n$) and require the dispersion to be larger than that for the first component. If the second component is significantly wider than the first but has a much

smaller fraction of the stars, then it can be considered a model for the “outliers.”

For stellar systems where outliers can be seen on the $Q-Q$ plot, we use Equation (16) to estimate the velocity dispersion of the main component, the fraction of sources that are outliers (the mixing parameter η in the model), and a velocity dispersion for the outliers. For both the single component and the mixture model, we calculate the change in the Bayesian Information Criterion (BIC), which is a penalized likelihood used for selecting between models with different numbers of parameters where the model with the lowest BIC is preferred (Schwarz 1978). For cases where the model appears to correctly identify outliers (i.e., η is small and the velocity dispersion is large for the second component), we report values from the mixture model method in Table 3. In eight out of nine cases where this model was applied, the two-component model produced a significant improvement in the BIC ($\Delta\text{BIC} < -6$). We note that in all cases where the outlier model improves the fit, Henze–Zirkler’s test showed strong deviation from normality.

5.3. One-dimensional Velocity Dispersions

Formulas for stellar dynamics are often given in terms of a one-dimensional velocity dispersion, σ_{1D} , because this quantity can be obtained from RV measurements alone. A characteristic one-dimensional velocity dispersion can also be obtained from multidimensional velocity dispersions by taking the mean variance

$$\sigma_{1D}^2 = \frac{\sigma_{\text{pc1}}^2 + \sigma_{\text{pc2}}^2}{2}. \quad (17)$$

These values are also recorded in Table 3.

The one-dimensional velocity dispersions found for clusters in our sample range from 0.8 to 2.8 km s^{-1} , with a mean of 1.8 km s^{-1} . For the ONC, the *Gaia*-based velocity dispersion is

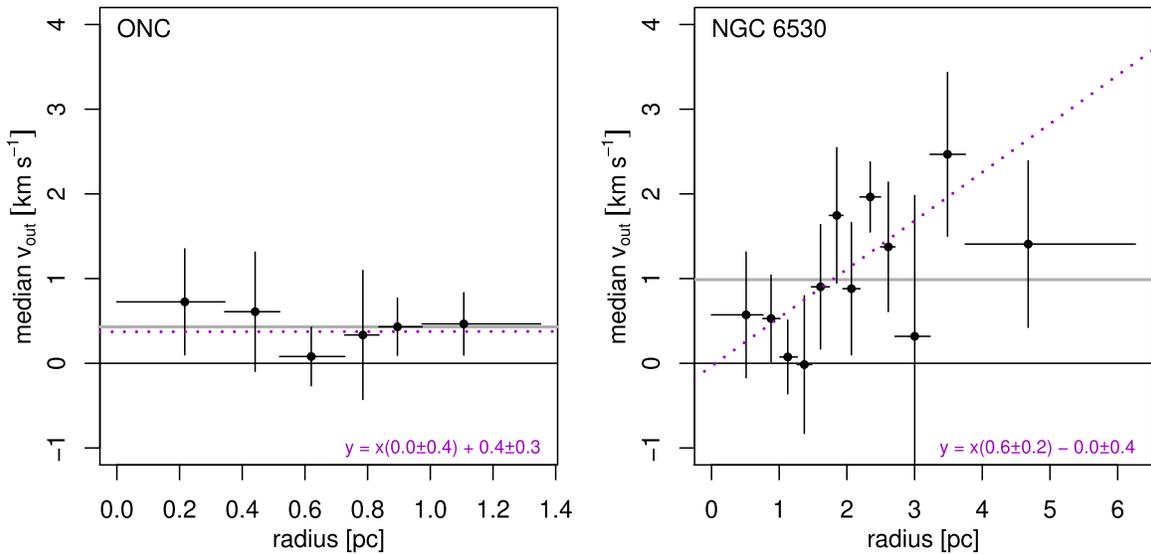


Figure 7. Expansion velocity as a function of distance from the system center. Stars are binned by radial distance in groups of 60. In each panel, the black line indicates the division between contraction and expansion, the gray line is the median expansion velocity for all stars, the points are the binned data, and the magenta line is a linear regression to the data. The equation of the regression line is included in the plot.

(The complete figure set (15 images) is available.)

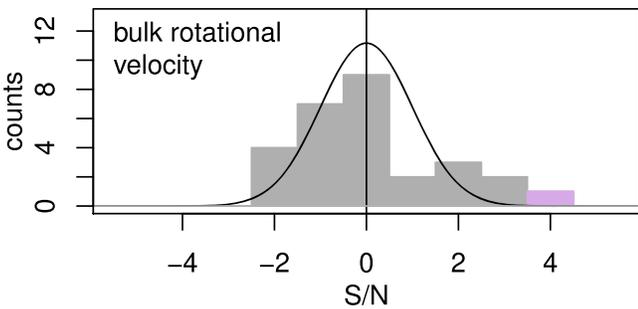


Figure 8. Histogram of the signal-to-noise ratio in rotational velocity (median v_{az} divided by its uncertainty). The observed distribution is compared to the expected Gaussian distribution (black curve) for the null hypothesis that there is no rotation, and non-zero values are effects of measurement uncertainty. Tr 15, significant at $>3\sigma$, is highlighted in magenta.

$\sigma_{1D} = 1.8 \pm 0.1 \text{ km s}^{-1}$. Estimates from earlier studies include $\sim 2.3 \text{ km s}^{-1}$ from a proper-motion study by Jones & Walker (1988), 3.1 km s^{-1} from an RV study by Fűrész et al. (2008), $\sim 2.3 \text{ km s}^{-1}$ from an RV study by Tobin et al. (2009), $\sim 2.5 \text{ km s}^{-1}$ from a RV study by Kounkel et al. (2016), $\sim 2.3 \text{ km s}^{-1}$ from a radio proper-motion study by Dzib et al. (2017), and 1.7 km s^{-1} from an RV study by Da Rio et al. (2017) after corrections to take into account spatial variations in the mean velocity dispersion. Our value is smaller than most of these estimates, but approximately equal to the estimate by Da Rio et al. (2017). We note that our estimate is based only on the central cluster, rather than on larger areas that were the focuses of studies by Fűrész et al. (2008) and Kounkel et al. (2016), for which the total velocity dispersions will include broadening effects from the velocity gradients identified by Da Rio et al. (2017).

There is a positive correlation between σ_{1D} and expansion velocity (Figure 11); Kendall’s rank correlation test shows the dependence to be marginally statistically significant ($p = 0.03$). The one-dimensional velocity dispersions are significantly higher than the expansion velocities, mostly exceeding the bulk

expansion velocities by a factor of 2–3. For cases where the expansion rate varies with radius (e.g., Figure 7), this effect will contribute to velocity dispersions. For example, in a toy model with pure expansion from a central point, we would expect $\sigma_{1D} \approx 1.5 \bar{v}_{out}$. However, velocity dispersions that are much larger than this imply that not all stars are moving outward.

In this initial *Gaia* study, the comparison of velocity dispersions in different regions comes with the caveat that the mass ranges of stars in the samples for different regions are not identical (e.g., Figure 1). If there is a relationship between a star’s mass and its velocity, then the selection of stars with good *Gaia* astrometry, which tend to be the brightest stars in a region, can affect the observed velocity dispersion. The systems for which velocity dispersions are derived in Table 3 typically have samples that include stars with masses down to $0.5 M_{\odot}$.

In order to determine how sample selection may affect estimated velocity dispersions, we show plots of velocity versus absolute magnitude in the *J* band, where the absolute magnitude is

$$M_J = J + 5 \log \varpi_0 [\text{mas}] - 10. \quad (18)$$

For the ONC and NGC 6530, we show the plots of both v_x and v_y versus M_J in Figure 12. In both cases, for stars with $1 < M_J < 4$ mag (approximately a mass range of $0.5\text{--}2.5 M_{\odot}$; constituting the bulk of the sample), velocity dispersion stays relatively constant with magnitude.

There has been much interest, both observational and theoretical, in the effect of stellar mass on velocity dispersions in open clusters and globular clusters (e.g., Bianchini et al. 2016; Parker et al. 2016; Spera et al. 2016; Webb & Vesperini 2017; Bravi et al. 2018). Overall, studies suggest that open clusters do not achieve energy equipartition (see Spitzer 1969). We leave the full investigation of the kinematic implications of stellar mass and mass segregation in the *Gaia* DR2 data to a future study.

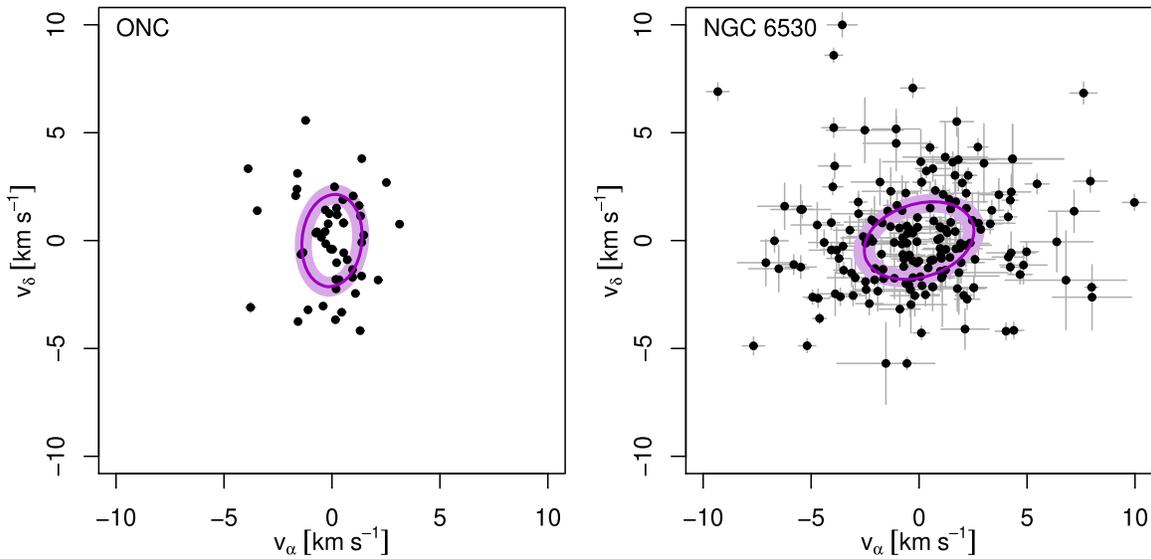


Figure 9. Velocity distributions for the ONC (left) and NGC 6530 (right). The magenta ellipses show the velocity dispersion from the best-fit bivariate normal distribution—the ellipse is the isodensity contour for this normal distribution at the 1σ level. The shaded regions show the uncertainty resulting from two times the standard error on the velocity dispersions. Only sources with no astrometric excess noise are used.

(The complete figure set (19 images) is available.)

5.4. Velocity Dispersion as a Function of Radius

The radial variation in velocity dispersion of a stellar system will reflect its dynamical state. Figure 13 shows the velocity dispersion as a function of radius for several stellar systems for which there are a sufficient number of stars to measure velocity dispersions in radial bins. For this analysis, we only use stars with `astrometric_excess_noise` = 0 that are not classified as “outliers.”

The top row in Figure 13 shows velocity dispersion as a function of radius in two systems, NGC 6530 and Cep B, that are clearly expanding (Section 4). In these two systems, velocity dispersion increases with distance from the center. Given that both of them were found to have increasing expansion velocity as a function of radius (Section 4.2), the trend in velocity dispersion supports our earlier result.

The bottom row in Figure 13 shows three systems, ONC, NGC 6231, and NGC 2362, with mild or no expansion. In the ONC, velocity is approximately constant with radius, and for NGC 6231 and NGC 2362, velocity decreases with radius. In the plots for these three systems, we use the cluster core radius r_c as a length scale in order to better compare with theoretical models for gravitationally bound clusters. The core radii were measured by Hillenbrand & Hartmann (1998) and Kuhn et al. (2017a, 2014) for these three clusters, who updated them with the new distance estimates (Section 8.4). For the ONC and NGC 2362, velocity is nearly constant out to a radius of eight times r_c , while for NGC 6231 it decreases steeply with radius—by a factor of ~ 2 at a distance of $4r_c$ from the center (Figure 13).

Given the lack of strong expansion of the ONC, NGC 6231, and NGC 2362, these may be some of the best candidates in our sample for gravitationally bound clusters, and it may be useful to compare them to commonly used cluster distribution functions. For the isothermal sphere model, velocity dispersion is independent of position throughout the cluster. On the other hand, for the Plummer sphere and the lowered isothermal sphere models (also known as King models), velocity dispersions decrease monotonically with radius (Binney & Tremaine 2008). In Plummer spheres, the velocity dispersion at

a point at radius r is proportional to $(1 + r/r_c)^{-1/4}$. The family of King models is characterized by a parameter W_0 describing the central potential (King 1966), and curves of one-dimensional velocity dispersion as a function of projected radius are shown by Binney & Tremaine (2008, their Figure 4.11) for several values of W_0 . For the ONC and NGC 2362, their velocity profiles would be consistent with isothermal spheres or King models with $W_0 \geq 9$. However, the Plummer model is rejected. For NGC 6231, the velocity dispersion is consistent with the Plummer model or King models with $3 < W_0 < 6$.

6. Subcluster Motions

Many of the stellar complexes from the MYStIX and SFINC studies contain subclusters that have been delineated by Kuhn et al. (2014) and Getman et al. (2017). The clearest examples that have good *Gaia* data include NGC 2264, Cep OB3b, NGC 6530, the Rosette Nebula, NGC 6357, NGC 6611, and the Carina OB1 association. For this subset, we now examine the kinematics of the substructures relative to each other. Subclusters with insufficient *Gaia* data (< 10 stars) are omitted.

Table 4 gives the properties of subclusters, including subcluster centers from earlier studies⁶ and the kinematics properties derived from *Gaia*. We find the bulk motions of subclusters by calculating the weighted-median velocities of their stellar members. In a few cases, subclusters have been combined when overlapping subclusters represent core–halo structures (Kuhn et al. 2017a)—these are indicated in Table 4. The projected subcluster velocities, relative to the association rest frame, range up to $\sim 8 \text{ km s}^{-1}$, with a median value of $\sim 2 \text{ km s}^{-1}$ and an interquartile range of $0.9\text{--}3.5 \text{ km s}^{-1}$. The velocity dispersions of subclusters can be quite different in different regions. The clusters in the Carina OB1 association have relative velocities of $5\text{--}8 \text{ km s}^{-1}$, while the various

⁶ In some cases, the subcluster centers may not be perfectly centered among the stars seen by *Gaia*. This is an effect of differential absorption across a subcluster, limiting *Gaia*’s sensitivity in the high-extinction part.

Table 3
Velocity Dispersions in Select Clusters

Region	n_{samp} (stars)	σ_{1D} (km s^{-1})	σ_{pc1} (km s^{-1})	σ_{pc2} (km s^{-1})	$\sigma_{\text{pc1}}/\sigma_{\text{pc2}}$	PA (deg)	η_{outliers} (%)	σ_{outliers} (km s^{-1})	ΔBIC
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Berkeley 59	18	1.2 ± 0.2	1.5 ± 0.3	0.8 ± 0.2	1.9 ± 0.6	178			
NGC 1893	54	2.6 ± 0.5	3.3 ± 0.7	1.7 ± 0.7	1.9 ± 0.9	43	15 ± 10	13	-40
ONC	48	1.8 ± 0.1	2.2 ± 0.2	1.4 ± 0.1	1.5 ± 0.2	5			
Mon R2	13	1.6 ± 0.4	2.0 ± 0.5	1.1 ± 0.5	1.7 ± 0.9	172			
NGC 2244	89	2.0 ± 0.1	2.2 ± 0.2	1.8 ± 0.2	1.2 ± 0.2	31			
S Mon	67	1.1 ± 0.1	1.3 ± 0.1	0.9 ± 0.1	1.4 ± 0.2	89			
NGC 2264 IRS 2	29	1.8 ± 0.2	2.3 ± 0.4	1.2 ± 0.2	1.9 ± 0.5	97			
NGC 2264 IRS 1	30	1.5 ± 0.2	1.9 ± 0.3	1.0 ± 0.2	1.9 ± 0.4	76			
NGC 2362	98	0.8 ± 0.1	0.9 ± 0.2	0.7 ± 0.1	1.2 ± 0.3	153	14 ± 17	2.0	-6.9
Tr 14	145	2.5 ± 0.2	2.8 ± 0.2	2.1 ± 0.3	1.3 ± 0.2	46	19 ± 7	7.3	-42
Tr 15	105	2.4 ± 0.2	2.5 ± 0.2	2.2 ± 0.2	1.1 ± 0.2	114	11 ± 5	9.3	-34
Tr 16	121	2.8 ± 0.2	3.4 ± 0.4	1.9 ± 0.2	1.8 ± 0.3	126	9	5.5	5.4
NGC 6231	278	1.6 ± 0.1	1.7 ± 0.1	1.5 ± 0.1	1.2 ± 0.1	132	10 ± 4	5.4	-106
M20	36	1.6 ± 0.2	1.9 ± 0.3	1.2 ± 0.3	1.6 ± 0.4	73	11 ± 10	6.1	-18
NGC 6530	185	2.3 ± 0.1	2.7 ± 0.2	1.8 ± 0.2	1.5 ± 0.2	106	13 ± 5	7.3	-73
NGC 6611	94	1.8 ± 0.2	2.2 ± 0.3	1.4 ± 0.2	1.6 ± 0.4	22	23 ± 9	5.8	-37
IC 5146	11	0.9 ± 0.2	1.1 ± 0.3	0.7 ± 0.2	1.5 ± 0.5	179			
NGC 7160	25	1.5 ± 0.2	1.8 ± 0.2	1.1 ± 0.2	1.7 ± 0.4	51			
Cep B	27	1.9 ± 0.2	2.1 ± 0.3	1.6 ± 0.3	1.3 ± 0.3	177			

Note. Column 1: system name. Column 2: number of stars used to calculate velocity dispersions. Column 3: characteristic one-dimensional velocity dispersion. Column 4–5: velocity dispersion in the first and second principal components for the two-dimensional velocity model. Column 6: ratio of velocity dispersions in the first and second velocity principal components—a measure of overall velocity anisotropy. Column 7: position angle (east from north) of the semimajor axis of the velocity dispersion. Column 8: fraction of stars belonging to the second component of the mixture model. Column 9: velocity dispersion for the second component of the mixture model. Column 10: change in BIC when the second component was added. The last three columns were only used when a multiple component velocity model was required.

subclusters in NGC 2264 have relative velocities of $0\text{--}2 \text{ km s}^{-1}$.

Figure 14 shows spatial maps of the stars assigned to each subcluster along with the velocity vectors of each group. In general, the subcluster motions are not convergent, but appear either randomly oriented or divergent. This pattern is seen in almost every star-forming region investigated, ranging from well-delineated clusters in regions like NGC 6357 to clumpy distributions of stars in regions like NGC 2264. The main exception is in Carina, where the motions of Tr 14 and Tr 16 are directed toward each other, but the apparent convergence of these clusters could be a chance alignment.

The contrast between internal cluster velocities and global kinematics of a complex can be clearly seen in the Carina OB1 association, which contains several clusters, including Tr 14/15/16 (included in this study) as well as Bochum 11 and the Treasure Chest to the south. Figure 15 (right) shows star positions in decl. plotted against their v_δ velocity component. In this complex, the individual clusters are shifted with respect to one another in velocity, but internal velocity dispersions within the individual clusters can also be seen. In the south of Carina OB1, there is a velocity gradient stretching from Bochum 11 to the Treasure Chest to Tr 16, while Tr 14 and Tr 15, to the north, have significantly different motions from the clusters to the south. Within Tr 14 and Tr 16, the positive correlation between decl. and v_δ characteristic of an expanding cluster can be seen. The total velocity difference between different clusters is significantly larger than the velocity dispersions within the clusters. In contrast, in the lower mass NGC 2264 region (left panel in Figure 15), small differences can be seen in the velocities of the different subgroups, but the

magnitudes of these differences and the total velocity dispersions are much smaller.

We find no evidence for the hierarchical assembly of rich clusters from subclusters in our sample. Evidence for this process would be converging motions of subclusters. This failure is expected, as *Gaia* does not provide access to the youngest embedded subclusters, but restricts our analysis to older clusters where the molecular cloud is at least partially dispersed. Thus, hierarchical cluster assembly, if it occurs, must occur promptly when a cluster is still embedded and must involve subclusters separated by $<5 \text{ pc}$ and ages $<1 \text{ Myr}$.

Subcluster motions are linked to the large-scale kinematics in molecular clouds, which may include effects of supersonic turbulence (Larson 1981; Padoan & Nordlund 2002; Mac Low & Klessen 2004; Krumholz & McKee 2005), free-fall velocities of collapsing clouds (Vázquez-Semadeni et al. 2017), and/or accretion of material onto molecular clouds (Fukui et al. 2014; Ibáñez-Mejía et al. 2017). The systems observed today at ages 1–5 Myr were formed in different dense molecular cores at widely separated portions in giant molecular clouds. It is therefore natural that they inherit the motion of their natal cloud cores and exhibit spatially correlated but essential random motions with respect to each other.

7. Relationships between Kinematics and Other Properties of Stellar Systems

Physical properties of stellar systems, including their masses, sizes, and ages, may be linked to their internal kinematics. These links may arise from stellar dynamics or through the star formation process. For example, the relationship between cloud size and cloud velocity dispersion (Larson 1981) could yield a

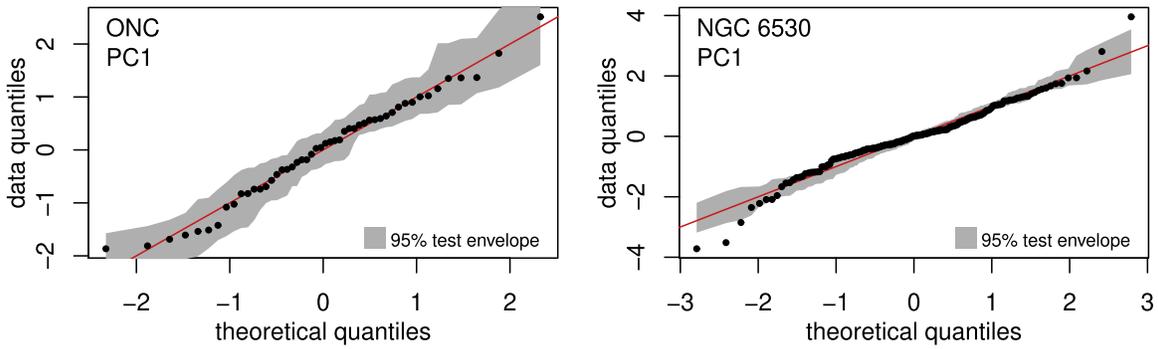


Figure 10. $Q-Q$ plot to evaluate the Gaussianity of stellar velocity distributions. The plots show theoretical quantiles vs. data quantiles, assuming that the data are fit by a normal distribution. Lower values on the left and higher values on the right show an excess number of outliers with higher than expected velocities. The red line shows the expected value for a normal distribution, and the gray shaded region shows a 95% test envelope. Only sources with no astrometric excess noise are used. (The complete figure set (38 images) is available.)

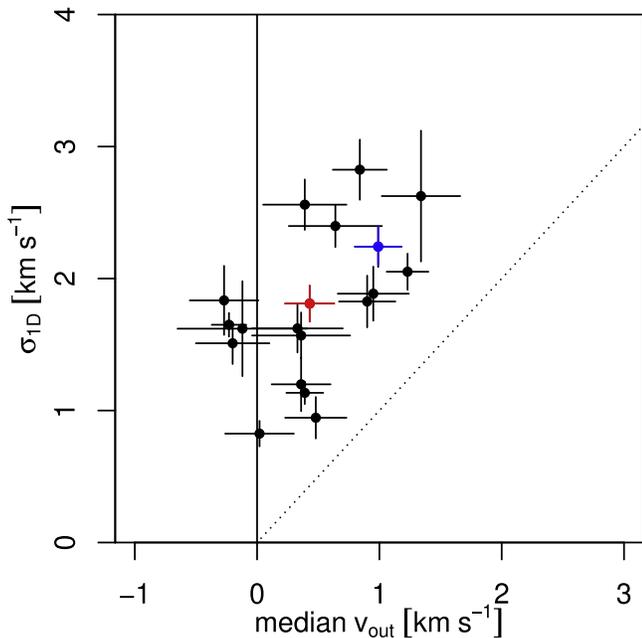


Figure 11. One-dimensional velocity dispersion vs. expansion velocity. The solid line indicates the division between contracting and expanding systems, and the dashed line indicates where both quantities are equal. The data points for the ONC (red) and NGC 6530 (blue) are marked.

relationship between the mass of the resulting stellar system and its velocity dispersion.

We estimate characteristic masses, sizes, and ages for objects in our sample using the data and methods from the previous MYSIX/SFINCS/NGC 6231 studies, but updating the values with new distance estimates. Estimates of the system mass, M_{cl} (corrected for incompleteness), are calculated using the methods from Kuhn et al. (2015b)—the X-ray luminosity function for pre-main-sequence stars is used to extrapolate the completeness fraction, and star counts are converted to masses using a mean mass of $\bar{m} = 0.61 M_{\odot}$ per star based on the Maschberger (2013) initial mass function (IMF). Kuhn et al. (2015b) estimated typical uncertainties of 0.25 dex on masses estimated with this method. Median ages for systems are calculated using the Age_{JX} method from Getman et al. (2014). Half-mass radii, r_{hm} , are calculated by taking the median

distance of stars in our samples from the center of the system. However, in the case of the ONC where our sample truncates the outer region of the cluster, we adopt the half-mass radius of 0.9 pc from Da Rio et al. (2014).

These estimates may be subject to a variety of systematic uncertainties, the effect of which is difficult to determine. For example, mass and age estimates can be affected by systematic errors in inference of stellar properties and choice of model isochrones (Richert et al. 2018). Furthermore, M_{cl} and r_{hm} could be underestimated due to difficulties determining the outer boundaries of clusters and the finite fields of view in the MYSIX/SFINCS/NGC 6231 studies.⁷ The derived physical properties of the systems are shown in Figure 16, and the values used here are available from the “data behind the figure.”

The gravitational effects of the natal molecular clouds are also likely to affect young stellar systems. The systems in our sample are in various stages of gas dispersal. Systems are considered to be embedded when the stars are reddened and projected on the molecular cloud and revealed when there is little extinction from the cloud. Partially embedded clusters are an intermediate stage and often represent systems at the edge of a cloud or within a bubble. The geometry of the system and projection effects may affect how systems are classified in ambiguous cases.

Figure 16 shows the relationships between kinematic properties (σ_{1D} and median v_{out}) and M_{cl} , r_{hm} , and age. Points are color-coded by degree of “embeddedness.” Unsurprisingly, embedded systems tend to be younger while revealed systems tend to be older. There is no statistically significant relation between σ_{1D} and embedded state. However, there is a statistical difference ($p_{\text{KS}} < 0.01$) between the median v_{out} of embedded systems, which tend to have no expansion, and that of partially embedded/revealed systems, which tend to be expanding.

Statistically significant correlations are found between the velocity dispersion and both M_{cl} and r_{hm} using Kendall’s rank correlation test ($p < 0.01$). However, no statistical correlation is found between the expansion velocity and the measured physical properties. As mass and radius are themselves related to each other, multivariate regression analysis is needed to treat interdependencies. We use the R package relaimpo (Grömping 2006) to evaluate the importance of $\log M_{\text{cl}}$, $\log r_{\text{hm}}$, and

⁷ See Kuhn et al. (2017a) for a detailed discussion of challenges involved in obtaining M_{cl} and r_{hm} .

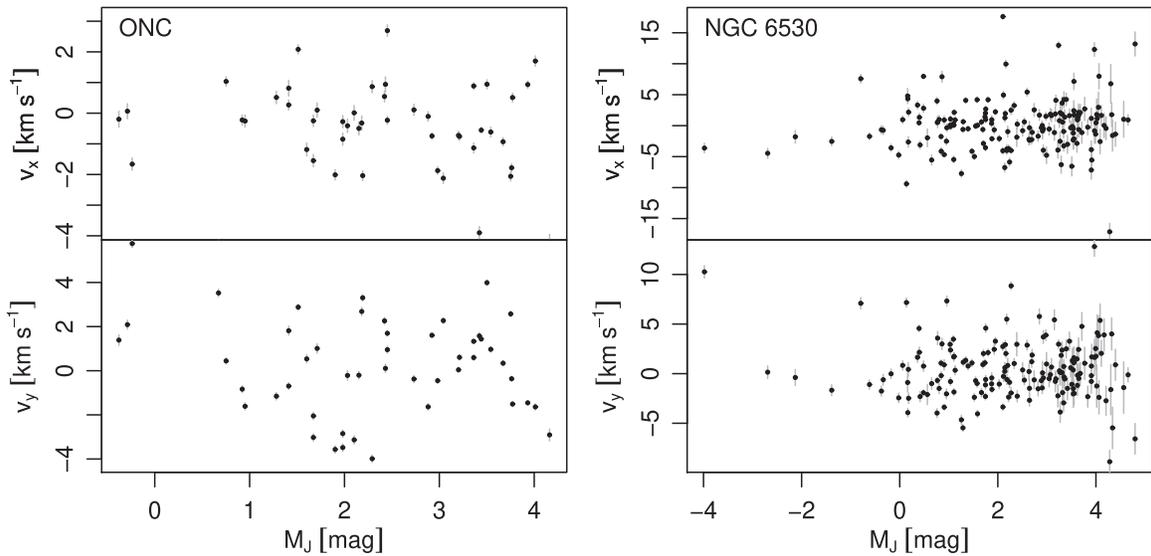


Figure 12. Velocity components v_x and v_y vs. J -band absolute magnitude. Statistical 1σ uncertainties on velocities are shown by the gray error bars. Only sources with no astrometric excess noise are used.

log age to predict $\log \sigma_{1D}$ and the median v_{out} in a linear regression. This analysis identified M_{cl} as the only important predictor of velocity dispersion and did not identify any of these variables as a statistically significant predictor of expansion. Although the relationship between system mass and velocity dispersion was expected, the result is interesting because it provides empirical evidence for the relationship based on independent estimates of system masses and velocity dispersions.

The crossing time, defined as $t_{cross} = 2 r_{hm} / \sigma_{1D}$, and the ratio of age to crossing time are two additional quantities that are important to stellar dynamics (Binney & Tremaine 2008). Crossing times range from 0.4 to 3 Myr in our sample, while the ratios of system age to crossing time range from 0.6 to 3. This indicates that the systems are all dynamically very young and would not have had time to relax through two-body interactions. Expansion can also drive the ratio of age to crossing time toward a small value, if a system expands at a rate proportional to its velocity dispersion. Surprisingly, there is no statistically significant correlation between either of these quantities and either velocity dispersion or expansion rate for the objects in our sample (Figure 16).

7.1. Virial State

The observed velocity dispersion of a system can be compared to the velocity dispersion needed for virial equilibrium σ_{virial} to estimate whether it is subvirial, virial, supervirial, or unbound. If $\sigma_{1D} > \sqrt{2} \sigma_{virial}$, the total energy of the system would be positive, and the system would be unbound. Given a mass M_{cl} and half-mass radius r_{hm} , the velocity dispersion of a virialized cluster is given by the equation

$$\sigma_{virial} = \left(\frac{G M_{cl}}{\eta r_{hm}} \right)^{1/2}, \quad (19)$$

where G is the gravitational constant and η is a constant that depends on the mass profile of a cluster. A Plummer model yields $\eta \approx 10$. Many young stellar clusters have $\eta < 10$ due to

relatively broad density profiles (e.g., Portegies Zwart et al. 2010; Grudić et al. 2018).

Figure 17 shows σ_{1D} versus σ_{virial} . There is a clear positive correlation between these two quantities ($p_{Kendall} < 0.001$). For the assumption that $\eta = 10$ (shown in the plot), all systems lie above the solid line showing the relationship for virial equilibrium, and most lie above the dashed line indicating zero total energy, which suggests that most of them are unbound. If we were to assume $\eta = 5$ instead, the ONC would be in approximate virial equilibrium, and several other systems, including Berkeley 59, NGC 2362, Tr 14, and NGC 6611, would have negative total energy, suggesting that they are likely bound. In addition, uncertainties on M_{cl} and r_{hm} could affect whether systems are in the “bound” or “unbound” regime of this plot. Although this plot is useful for demonstrating a statistical correlation between σ_{1D} versus σ_{virial} , in most cases, whether or not a particular system is gravitationally bound remains ambiguous due to systematic uncertainties.

Two non-expanding systems (NGC 6231, and NGC 2362), one mildly expanding system (ONC), and two systems with clear expansion (NGC 6530 and Cep B) are labeled on the figure. The mildly/non-expanding systems lie along the bottom of the distribution, while the expanding systems lie toward the top, as would be expected if expansion is driven by systems being unbound. The ONC is within the region considered gravitationally bound, and NGC 6231 and NGC 2362 could be within this region given uncertainties in measurements or assumptions. At the other extreme, NGC 6530 and Cep B are sufficiently far from being in virial equilibrium that they can be classified as being unbound even in the presence of systematic uncertainties.

Given the large dynamical range in M_{cl} , which spans a factor of ~ 100 in mass, it is notable that the relationship between σ_{1D} and σ_{virial} is relatively tight with only a factor of ~ 1.5 scatter.

7.2. Dynamical State of Expanding Systems

The two examples with the most statistically significant expansion are NGC 6530 and Cep B, both of which have “Hubble flow” like expansion patterns (Figure 7) which hint at free expansion. They have one-dimensional velocity dispersions

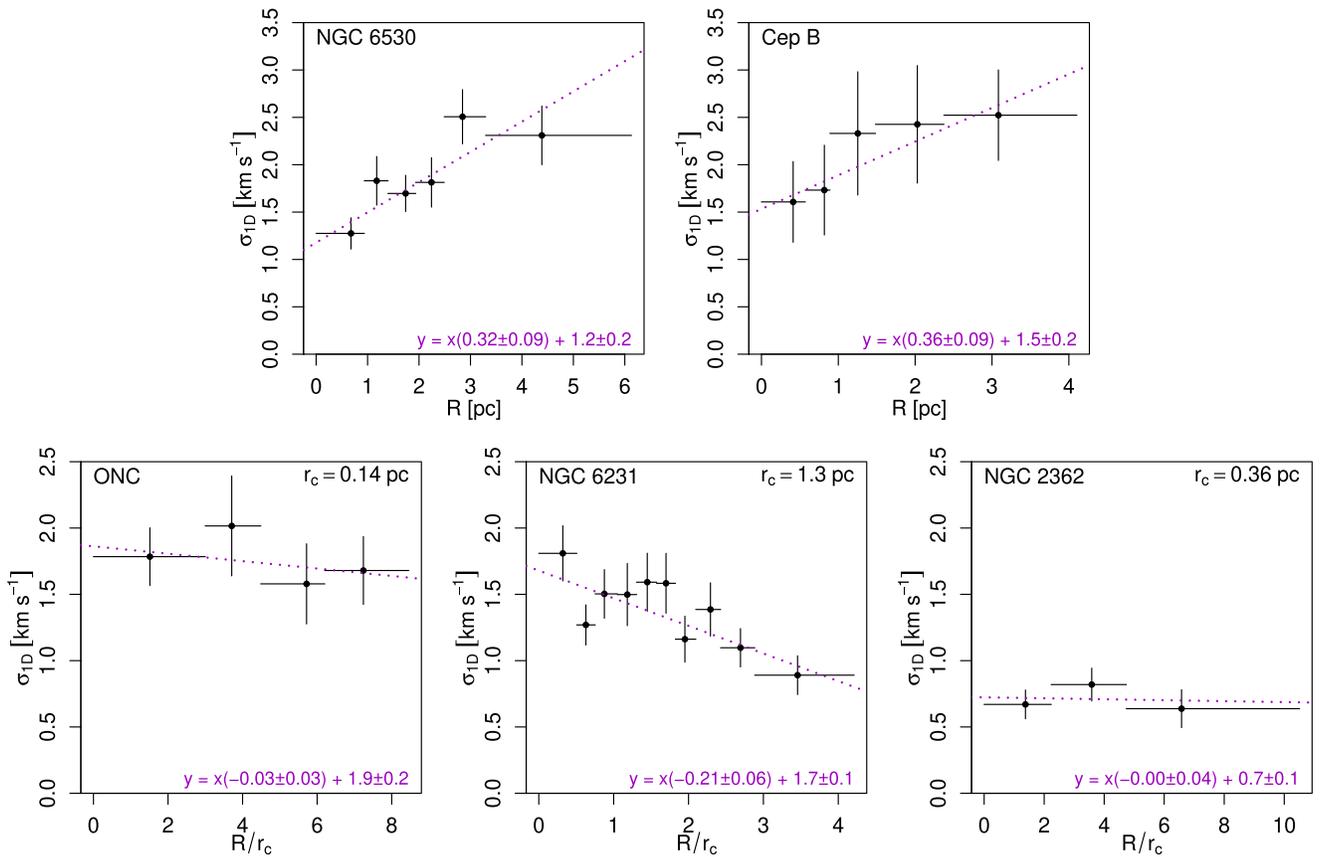


Figure 13. Velocity dispersion as a function of distance from the cluster center. In each panel, the points show the velocity dispersion in each radial bin, and the error bars show 1σ uncertainties. The weighted least-squares regression line for y as a function of x is shown for each plot (dotted magenta line), and the equation is given in the lower right corner. Top row: plots for the expanding systems NGC 6530 and Cep B. Bottom row: plots for systems with mild (ONC) or no (NGC 6231 and NGC 2362) expansion. For the bottom row, the x -axis is normalized by the cluster core radius r_c (Section 7) to facilitate comparison with cluster models.

of $2.2 \pm 0.2 \text{ km s}^{-1}$ and $1.9 \pm 0.2 \text{ km s}^{-1}$ and approximate half-mass radii of $\sim 2 \text{ pc}$ and $\sim 1.5 \text{ pc}$, respectively. For these values and the assumptions above, NGC 6530’s virial mass would be $\sim 20,000 M_\odot$ and Cep B’s virial mass would be $\sim 10,000 M_\odot$. However, the estimated system masses are only $\sim 4000 M_\odot$ and $\sim 1000 M_\odot$ (Kuhn et al. 2015b). Thus, both the gradient in expansion velocity and the inferred stellar populations indicate that these two systems are not gravitationally bound. This places these expanding associations near the top of the distribution on Figure 17.

For a freely expanding association, stars will sort themselves by velocity as they move away from the center of the system, effectively decreasing the local velocity dispersion. This can be tested in NGC 6530 by plotting the ratio of expansion velocity to the dispersion in v_{out} as a function of radius (Figure 18). Values of this ratio >1 suggest that nearly all stars are moving outward, while values <1 suggest some stars are moving inward even as the system expands overall. In this case, it turns out that the expansion velocity is always less than or equal to the local velocity dispersion.

Many of the expanding systems show internal substructure (Kuhn et al. 2014). We have not analyzed the kinematics within individual subclusters due to insufficient numbers of stars in our sample. In principal, it would be possible for subgroups of stars to be locally bound, even if the total energy of a region as a whole is positive. Future *Gaia* data releases are likely to provide

more information about these groups due to larger samples of stars with good astrometry and higher overall precision.

7.3. Dynamical State of Non-expanding Systems

The three objects ONC, NGC 6231, and NGC 2362 may be the best candidates in our sample of 28 systems for being gravitationally bound clusters because the systems have little to no expansion, and their surface density profiles are close to what would be expected for a system in virial equilibrium.

The core radius r_c —the radius where the apparent surface density of stars drops by a factor of ~ 2 —can serve as a length scale with which to connect the spatial and kinematic cluster properties. For a dynamically relaxed cluster, with a surface density profile given by a King profile or an isothermal sphere, the density at the center of the clusters, ρ_0 , is related to the cluster core radius, r_c , and the velocity dispersion, σ_{1D} , by the equation

$$\rho_{0,\text{virial}} = \frac{9\sigma_{1D}^2}{4\pi G r_c^2}, \quad (20)$$

where G is the gravitational constant (Binney & Tremaine 2008). For clusters in which the potential is dominated by the stars (versus needing to account for $M_{\text{stars}} + M_{\text{gas}}$), the mass density implied by the cluster dynamics can be compared to the observed cluster number density n_0 to infer the mean mass per star. This, in turn, can be compared to the average mass per star

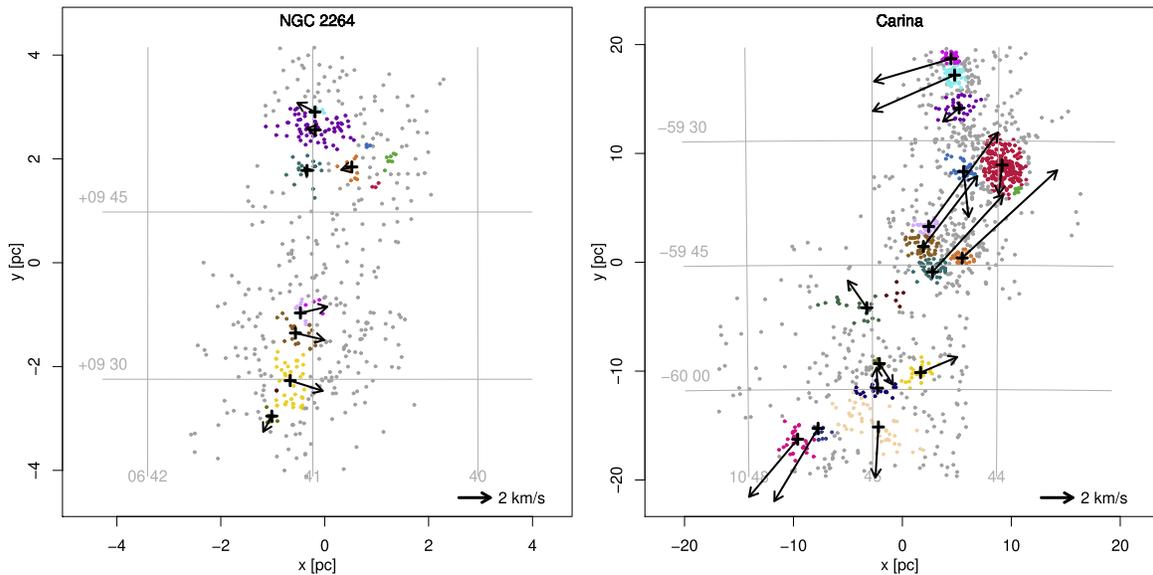


Figure 14. Kinematics of subclusters in NGC 2264 (left) and the Carina OB1 association (right). Stars included in the study are marked with a color symbol indicating the subcluster to which they were assigned in Kuhn et al. (2014). The crosses mark subcluster centers, and the vectors indicate velocities of the subclusters, as indicated by the velocity scale. Subcluster velocities in Carina tend to be much larger than in the smaller, nearby NGC 2264 region. In both NGC 2264 and Carina, nearby groups of stars tend to move in similar directions, but there is no overall sign of subcluster mergers.

(The complete figure set (7 images) is available.)

predicted by a standard IMF. Adopting the Maschberger (2013) form of the IMF over the mass range $0.08\text{--}150 M_{\odot}$ yields $\bar{m} = 0.61 M_{\odot}$ for single stars or $\bar{m} = 0.78 M_{\odot}$, assuming a population including unresolved binary systems.

The ONC is a smooth, centrally concentrated distribution of stars, which was modeled by Hillenbrand & Hartmann (1998) using a King profile (King 1966; Binney & Tremaine 2008) with a core radius of $\sim 0.15\text{--}0.20$ pc.⁸ Although the cluster is located within the Orion A Cloud, the central region of the cluster is dominated by stars, not dense gas; an ionization front propagates back into the molecular cloud, located approximately 0.2 pc behind the cluster center, and a neutral gas “lid” is located >1 pc in the foreground (O’dell 2001). Hillenbrand & Hartmann (1998) used velocity measurements from Jones & Walker (1988) to compare the virial central density of the ONC to the observed number density of stars. We repeat this experiment, instead using the $\sigma_{1D} = 1.8$ km s⁻¹ from *Gaia* measurements and scaling their core radius to $r_c = 0.14$ pc and their observed density of stars to $n_0 = 2.7 \times 10^4$ stars pc⁻³ based on differences in distance assumptions. This gives $\rho_0 = 2.8 \times 10^4 M_{\odot} \text{ pc}^{-3} = 1.0 M_{\odot} n_0$. A value of $1.0 M_{\odot}$ per star is slightly higher than the expected mean stellar mass, but some of the virial mass may be made up by accounting for some amount of remaining gas in the outer parts of the cluster. Thus, the ONC is close to virial equilibrium, although a slight discrepancy could account for the mild expansion.

NGC 6231 and NGC 2362 are both larger, older, and less dense than the ONC. The large sizes suggest that they may

have already expanded, but may have reached a turnaround radius where outward expansion has halted.

For NGC 6231, values of $r_c = 1.3$ pc and $n_0 = 180$ stars pc⁻³ were found by Kuhn et al. (2017a), scaled for difference in assumed distances.⁹ A velocity dispersion of $\sigma = 1.6$ km s⁻¹ yields $\rho_{0,\text{virial}} = 260 M_{\odot} \text{ pc}^{-3} = 1.4 M_{\odot} n_0$, a slightly larger ratio than that for the ONC. In this case, the cluster is completely devoid of molecular cloud, so gas mass cannot contribute significantly to the total mass. Assuming the average mass per star (or binary) is $\sim 0.6\text{--}0.8 M_{\odot}$, this would place NGC 6231 near the threshold for being unbound, but whether the cluster is bound or unbound depends on the total mass of the cluster.

For NGC 2362, values of $r_c = 0.36$ and $n_0 = 600$ stars pc⁻³ were found by Kuhn et al. (2015a), scaled for difference in distance.¹⁰ A velocity dispersion of $\sigma = 0.8$ km s⁻¹ yields $\rho_{0,\text{virial}} = 870 M_{\odot} \text{ pc}^{-3} = 1.5 M_{\odot} n_0$. NGC 2362 is also a system from which gas has been expelled, so, again, the molecular cloud cannot provide the additional mass. However, uncertainty in mass estimation could make the difference between a bound cluster and unbound system.

8. Discussion

The *Gaia* data reveal considerable diversity in kinematic properties in our sample of nearby young clusters, with examples of both expanding and non-expanding systems (Figures 2–6). The expected pattern of increased expansion

⁸ The structure of the ONC core is more complex than accounted for by a single King profile. For example, there are small concentrations of stars with densities much greater than accounted for by the smooth King model (e.g., Henney & Arthur 1998; Rivilla et al. 2013; Kuhn et al. 2014). We use the simpler model form from Hillenbrand & Hartmann (1998) because it is easier to interpret dynamically.

⁹ The estimation of the number of stars in the cluster (on which n_0 depends linearly) is based on extrapolation of the IMF to account for incompleteness in the observation and thus, depends on assumptions of stellar age, pre-main-sequence isochrones, cluster distance, and the shape of the IMF. Kuhn et al. (2017a) estimated a total of 5700 stars down to $0.08 M_{\odot}$ projected within the field of view, while Damiani et al. (2016) estimated a similar cluster mass using different assumptions.

¹⁰ As for NGC 6231, the estimated density of stars at the center of NGC 2362 depends on model-based corrections for incompleteness in the observed sample.

Table 4
Relative Subcluster Kinematics

Region	Subcluster	α_0 (J2000)	δ_0 (J2000)	n_{samp} (stars)	$\langle v_x \rangle$ (km s ⁻¹)	$\langle v_y \rangle$ (km s ⁻¹)
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Rosette	A	6 30 57.1	+04 57 57	26	0.5 ± 0.5	-1.2 ± 0.7
...	C	6 31 32.0	+04 50 58	12	1.3 ± 1.2	0.0 ± 1.0
...	D+E	6 31 59.3	+04 54 50	241	0.5 ± 0.3	0.5 ± 0.3
...	H	6 33 07.2	+04 46 57	17	0.3 ± 2.0	-0.8 ± 0.5
...	L	6 34 10.7	+04 25 06	50	-0.5 ± 0.4	-0.6 ± 1.0
NGC 2264	D	6 40 45.8	+09 49 03	11	-0.7 ± 0.2	-0.2 ± 0.6
...	E	6 40 59.1	+09 52 22	67	-0.5 ± 0.4	0.09 ± 0.1
...	F	6 40 59.2	+09 53 59	10	-1.1 ± 0.5	0.6 ± 0.3
...	H	6 41 02.1	+09 48 44	16	-0.2 ± 1.0	0.2 ± 0.6
...	I	6 41 04.5	+09 35 57	13	1.7 ± 0.7	0.4 ± 0.2
...	J	6 41 06.3	+09 34 09	20	1.8 ± 0.2	-0.4 ± 1.0
...	K	6 41 08.2	+09 29 53	43	2.1 ± 0.6	-0.7 ± 0.4
...	M	6 41 14.9	+09 26 42	11	-0.5 ± 0.8	-1.0 ± 0.4
Carina	B+C	10 43 56.4	-59 32 54	262	-0.2 ± 0.4	-1.9 ± 0.2
...	D	10 44 32.9	-59 33 42	36	0.3 ± 0.9	-2.9 ± 0.9
...	E	10 44 34.0	-59 44 08	31	6.0 ± 0.7	5.6 ± 1.0
...	F	10 44 37.4	-59 26 03	34	-1.0 ± 0.9	-0.9 ± 0.8
...	H	10 44 41.8	-59 22 05	59	-5.2 ± 0.7	-2.3 ± 0.6
...	I	10 44 45.3	-59 20 07	24	-4.9 ± 0.4	-1.4 ± 1.4
...	J	10 45 02.4	-59 45 50	41	4.5 ± 1.1	4.9 ± 0.7
...	K	10 45 06.2	-59 40 21	30	4.4 ± 0.5	6.0 ± 0.5
...	L	10 45 11.1	-59 42 46	51	3.4 ± 0.3	4.5 ± 0.6
...	M	10 45 13.7	-59 57 58	23	2.3 ± 1.3	1.0 ± 1.0
...	O	10 45 53.4	-59 56 53	11	0.9 ± 1	-1.3 ± 0.7
...	P	10 45 54.4	-60 04 32	47	-0.2 ± 1.6	-3.2 ± 1.3
...	Q	10 45 55.2	-59 59 51	23	-0.1 ± 1.0	1.3 ± 1.1
...	R	10 46 05.4	-59 50 09	19	-1 ± 1	1.7 ± 0.9
...	S	10 46 52.7	-60 04 40	13	-2.8 ± 0.7	-4.6 ± 1.5
...	T	10 47 12.5	-60 05 58	33	-3.1 ± 0.6	-3.7 ± 0.8
NGC 6357	A	17 24 43.7	-34 12 07	34	-1.1 ± 0.5	3.1 ± 0.5
...	B	17 24 46.7	-34 15 23	23	-1.6 ± 0.9	1.3 ± 0.7
...	C+D	17 25 34.3	-34 23 10	27	1.0 ± 0.6	2.4 ± 1.2
...	E	17 25 47.9	-34 27 12	11	-4.1 ± 2.1	-2.7 ± 0.8
...	F	17 26 02.2	-34 16 42	21	1.8 ± 1.1	-1.0 ± 0.8
NGC 6530	A	18 03 23.8	-24 15 19	15	-3.4 ± 1.2	1.9 ± 0.8
...	C	18 03 46.3	-24 22 01	14	2.2 ± 3.3	-1.9 ± 0.9
...	D	18 03 51.3	-24 21 08	11	4.1 ± 2.1	-1.1 ± 0.8
...	E	18 04 07.6	-24 25 53	55	-0.9 ± 1.1	-1.8 ± 0.5
...	F	18 04 13.3	-24 18 27	173	1.5 ± 0.4	1.1 ± 0.3
...	G	18 04 20.1	-24 22 51	14	0.6 ± 1.0	0.6 ± 1.0
...	H	18 04 23.3	-24 21 13	65	0.03 ± 0.4	0.8 ± 0.2
...	I	18 04 28.3	-24 22 46	110	0.3 ± 0.2	-0.5 ± 0.5
...	J	18 04 39.6	-24 23 20	31	-0.7 ± 0.5	0.04 ± 0.3
...	K	18 04 50.5	-24 26 19	45	-2.3 ± 0.9	-0.1 ± 0.8
NGC 6611	A+B	18 18 42.2	-13 47 03	213	0.5 ± 0.3	-0.1 ± 0.3
...	D	18 18 57.3	-13 45 23	71	-1.7 ± 0.4	0.9 ± 0.7
Cep OB3b	A	22 53 47.1	+62 35 47	195	2.2 ± 0.4	-0.9 ± 0.4
...	B	22 54 58.4	+62 34 09	23	-0.1 ± 0.5	-0.5 ± 0.4
...	C	22 56 40.3	+62 42 06	401	-0.6 ± 0.3	0.2 ± 0.3

Note. Column 1: the names of the star-forming regions. Column 2: the names of the subclusters defined by Kuhn et al. (2014) and Getman et al. (2018). Columns 3–4: coordinates of the subcluster centers. Column 5: the number of *Gaia* sources in each subcluster. Column 6–7: subcluster velocity projected in the plane of the sky relative to the center-of-mass rest frame of the entire association.

velocity with distance from the cluster center is seen (Figure 7). Velocity dispersions range from 1 to 3 km s⁻¹, exhibit Gaussian distributions, and often exceed the expectations of virial equilibrium (Figures 9–11, 17).

On larger scales, we examine the relative motions of subclusters in star-forming complexes. Subcluster trajectories are typically divergent, reflecting their origins in turbulent

clouds, rather than convergent as expected if clusters are currently assembling from smaller components (Figure 14).

8.1. Expectations from Simulations

Theoretical models of star cluster formation, informed by hydrodynamical and *N*-body codes, have led to predictions about the stellar dynamics of very young clusters—even before

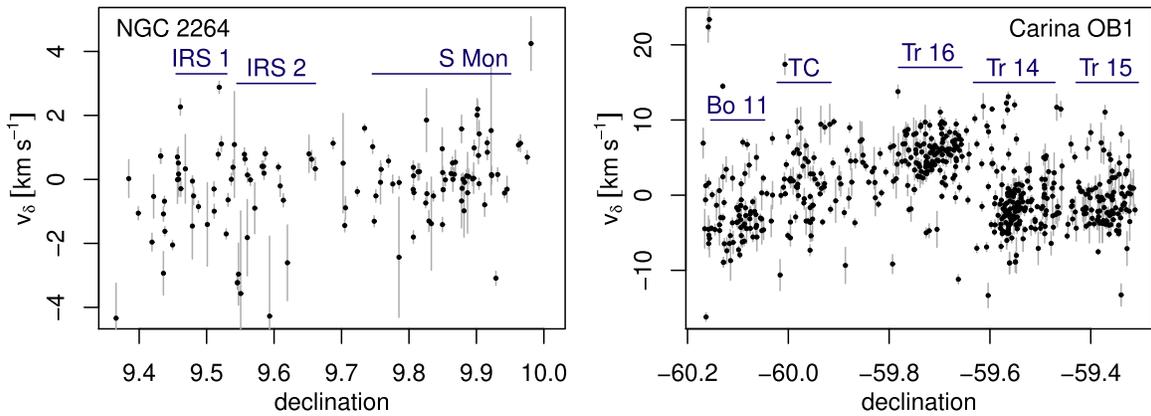


Figure 15. Velocity v_8 vs. decl. for stars in NGC 2264 (left) and the Carina OB1 association (right). The decl. range of several constituent clusters are marked, including S Mon, IRS 1, and IRS 2 in NGC 2264; and Tr 14, 15, 16, the Treasure Chest (TC) and Bochum 11 (Bo 11) in Carina. The plot reveals the different mean velocities of the different clusters as well as cluster expansion within some individual clusters. Note that the velocity range for NGC 2264 is much smaller than that for Carina. Only sources with no astrometric excess noise are used.

sufficiently precise kinematic data were available to test these models.

Cluster simulations typically show an initial collapse during the first crossing time where the global contraction is accompanied by increasing velocity dispersion. This brief phase is followed by a re-expansion with outward velocity accompanied by a slight decrease in velocity dispersion (e.g., Proszkow et al. 2009, their Figure 1). Departure from spherical symmetry, such as elongated clusters, will produce an observed kinematic structure dependent on viewing angle. These projection effects can produce apparent velocity gradients and can influence the velocity dispersion by about a factor of 2, mostly during the expansion phase (Proszkow et al. 2009).

In gas-rich environments, subclusters are likely to merge into more massive young star clusters (Fellhauer et al. 2009; Maschberger et al. 2010). Kuznetsova et al. (2015) found that the gas potential dominates the free-fall time on timescales <0.75 – 1 , encompassing the initial infall and star formation phases. By the time the stellar potential begins to dominate, subclusters have already merged. Kuznetsova et al. (2015) reported stellar velocity dispersions of 3 – 4 km s^{-1} compared to gas velocity dispersions of 1 – 2 km s^{-1} for an isothermal (cold, subvirial) model of an ONC-like cluster. Other simulations (Banerjee & Kroupa 2018) implied that some observed massive young star clusters are too compact to be produced through hierarchical formation, thus suggesting the monolithic formation route.

At the more advanced age at which young clusters are revealed in optical wavelengths, their initial kinematic state will have been altered by loss of residual gas, causing the system to expand and/or disperse (e.g., Adams 2000; Kroupa et al. 2001). Early studies focused on the role of star formation efficiency in determining the final state of the system. These studies showed that for simple models of gas loss, there may be a threshold of 20%–33% star formation efficiency that governed whether a cluster will survive (e.g., Mathieu 1983; Goodwin & Bastian 2006) and that star formation efficiency will affect the fraction of stars stripped from a surviving cluster (e.g., Kroupa et al. 2001; Bastian & Goodwin 2006). However, cluster survival is also strongly influenced by the initial dynamical state, with subvirial initial states leading to higher survival probabilities and supervirial initial states leading to lower survival probabilities (Goodwin 2009). The situation becomes even more complicated due to effects of cloud and star cluster structure. Simulations indicate that spatial

decoupling between gas and stars (Dale et al. 2015) and highly substructured spatial distributions of stars (Farias et al. 2018) will attenuate the effect of gas expulsion on a stellar system. Other factors such as dynamical ejection of massive stars from very dense clusters have also been identified as possible contributors to cluster mass loss leading to cluster expansion (Pfalzner & Kaczmarek 2013). Energy injected into clusters by the hardening or formation of binary systems can also contribute to cluster expansion (Banerjee & Kroupa 2017).

The expansion of a cluster or association can follow different tracks depending on its mass and virial state. Even systems that have a positive total energy may leave behind a bound core of stars (e.g., Goodwin & Bastian 2006, their Figure 4). Escaping stars may lead to an excess number of stars at larger distances from a system center (e.g., Bastian & Goodwin 2006), and this could contribute to a positive radial expansion gradient. The timescale on which massive clusters may expand and/or evaporate in response to gas expulsion is on the order of 10 Myr (Pfalzner et al. 2014 and references therein). Long-term cluster survival (10 Myr to 1 Gyr) is positively correlated with cluster mass, but in clusters that arose from low star formation efficiency environments, the correlation may be weaker (Shukirgaliyev et al. 2018). The virial state of systems will also affect their structure as they evolve—substructure will be erased in a few dynamical timescales from systems that are initially virial or subvirial, while supervirial systems may retain substructure as they expand (Parker & Wright 2016).

In the following discussion (Sections 8.2–8.3), we use a simulation of cluster assembly from Sills et al. (2018) as a benchmark to compare with the observed clusters in our study. The model used for comparison is the “DR 21 Fiducial Model” designed to understand the current gas and stellar distribution in DR 21, a massive filamentary star-forming region in Cygnus X. A full description of the model’s initial conditions and assumptions is given in Sills et al. (2018). Briefly, this is a simulation of a $3000 M_{\odot}$ star cluster in which stars originate in a chain of subclusters, spatially dispersed following the pattern of observed stars in DR 21. Each subcluster is modeled with an elongated Plummer sphere and a velocity dispersion based on the virial parameter. The subclusters have no initial bulk motion relative to each other in this simulation. Gas was modeled without star formation or stellar feedback and moves gently out of the cluster center during the simulation. The model simulates the first 10 Myr of dynamical evolution,

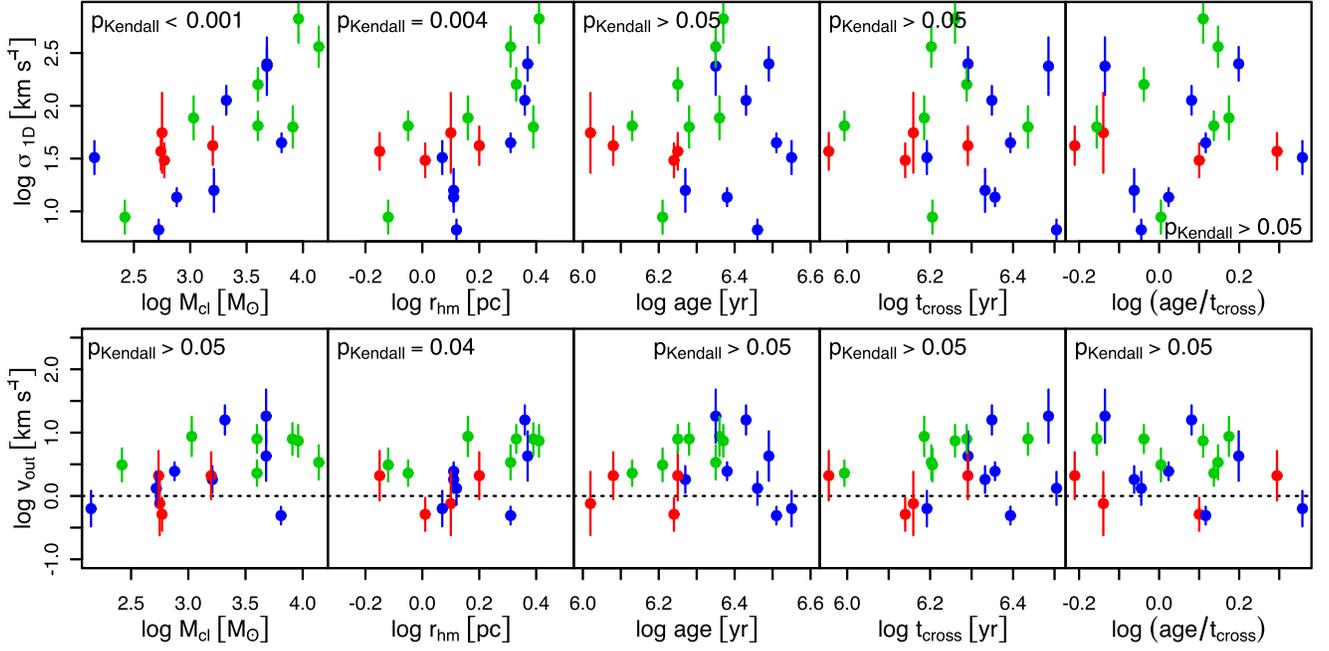


Figure 16. Scatter plots showing relationships between cluster kinematics and other physical cluster properties for the subset of stellar systems included in Table 3. Kinematic quantities include the velocity dispersion (σ_{1D}) and expansion velocity (median v_{out}), and other physical properties include the system mass (M_{cl}), size (r_{hm}), age, crossing time (t_{cross}), and ratio of age to crossing time. Systems are color-coded based on properties of the natal cloud: embedded (red), partially embedded (green), and revealed (blue). Statistical significance of the correlation is assessed using the Kendall rank correlation test (Hollander & Wolfe 1973), and the p -values are shown on the plot. Positive correlations exist between velocity dispersion, mass, and size, but no statistically significant correlations can be found for other properties. The data used to create this figure are available.

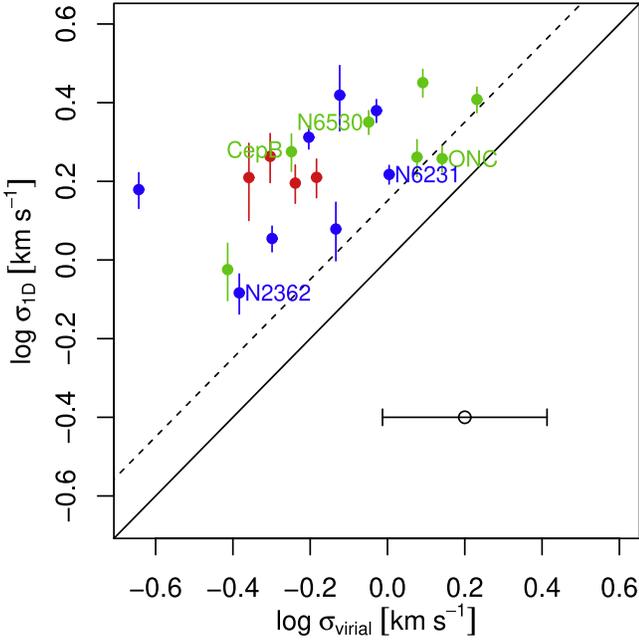


Figure 17. Virial velocity dispersion calculated using observed M_{cl} and r_{hm} vs. observed velocity dispersion. Two non-expanding systems (NGC 6231 and NGC 2362), one mildly expanding system (ONC), and two expanding systems (NGC 6530 and Cep B) are labeled. Colors of points indicate the degree of “embeddedness” as in Figure 16. The solid line shows the expected relationship for a cluster in virial equilibrium, while the dashed line shows the limit for a bound cluster. The values of σ_{virial} are calculated assuming $\eta = 10$ in Equation (19). The large error bar shows how a factor of 2 systematic uncertainty on mass or radius would affect σ_{virial} .

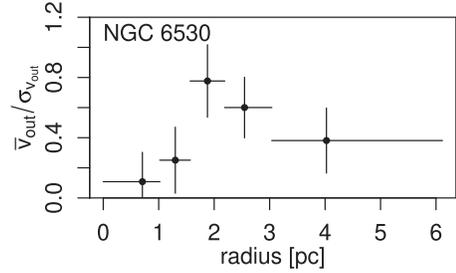


Figure 18. Ratio of expansion velocity to dispersion in v_{out} as a function of radius. Larger values indicate that velocity vectors are mostly directed outward, while smaller values indicate significant fractions of stars moving both outward and inward.

showing the progression from a clumpy assembly of subclusters to a single, virialized massive young star cluster.

This simulation results in a bound system—a different outcome than observed in many, but not all, of the systems in our sample—and thus, the model is unlikely to be representative for the largely unbound systems in this study (Section 8.3). However, the simulation is useful for revealing what kinematic effects may be observed in bound clusters (Section 8.4) and therefore allow us to evaluate which phenomena can be used as evidence for a system being unbound versus bound.

8.2. Assembly of Clusters and Associations

The stellar systems included in our *Gaia*-based study are typically several million years old, and the clouds from which stars are forming have been partially or completely dispersed,

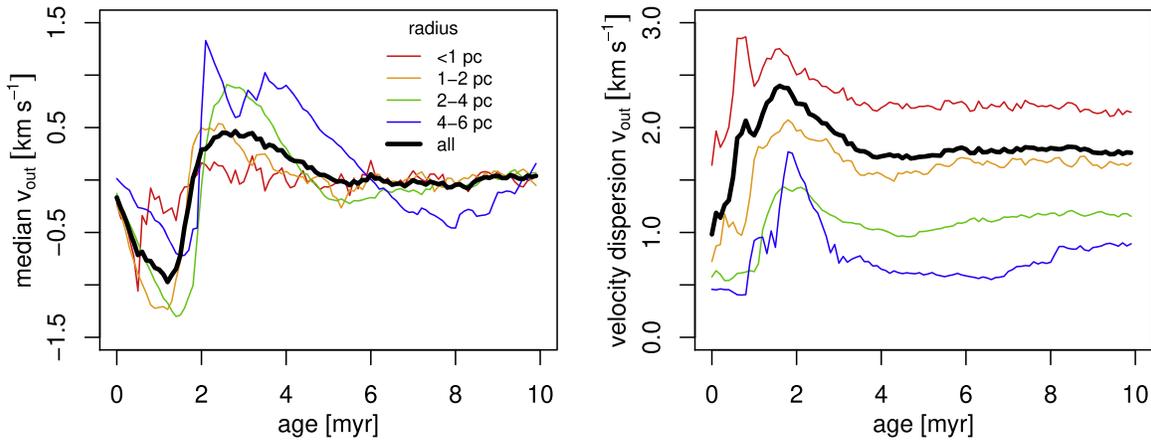


Figure 19. Characteristics of the velocity distribution for the stars in the Sills et al. (2018) simulation as a function of time. The left panel shows the median v_{out} while the right panel shows the standard deviation of v_{out} . These statistics are calculated for stars in various radial bins as indicated by the legend. During the simulation, stars stream into the center of the cluster, yielding bulk inward velocities, before rebounding with a bulk outward velocity and settling into a “quasi-static” state.

so the epoch of cluster assembly will have mostly finished. However, kinematic properties such as rotation and the motions of subclusters can provide constraints on how systems were assembled.

In the simulation from Sills et al. (2018), starting at an earlier stage of evolution, subclusters rapidly converge to form a single cluster in ~ 1 Myr. The evolution of the median v_{out} and velocity dispersion, calculated in the same way as above for the real clusters, is shown for the first 10 Myr in Figure 19. The infall of subclusters produces an initial negative median v_{out} for the central cluster for the first 2 Myr. The velocity distribution rebounds, yielding cluster expansion, before settling into a quasi-static state at ~ 6 Myr. The infall velocity reaches -0.75 km s^{-1} , while the rebound outward velocity peaks at $\sim 0.5 \text{ km s}^{-1}$. The velocity dispersion increases during infall, peaking at the point of rebound, then settling down to just under 2 km s^{-1} .

The only system in our study that may be caught in a phase of initial collapse is M17, which appears to have a large (though uncertain) contraction velocity (median $v_{\text{out}} = -2.1 \pm 1.0 \text{ km s}^{-1}$). The spatial structure of stars in M17 is also distinctive, being one of the few systems in the MYStIX survey with a “clumpy” structure, suggesting that the merging of many subclusters is still underway (Kuhn et al. 2014). However, the other embedded stellar systems do not appear to be undergoing such rapid collapse. The systems NGC 1333 and NGC 2264 IRS 1/2, much less massive than M17, are also gas-rich and contain protostars, but neither of them have statistically significant expansion or contraction.

For stellar systems in more evolved star-forming regions, where molecular gas has been partially or fully dispersed, the motions of subclusters are inconsistent with mergers. This supports the theoretical prediction that hierarchical cluster assembly, if it occurs, must happen promptly before gas is expelled by stellar feedback.

Rotational properties of young stellar systems can also be used to test the hierarchical assembly scenario. Simulations of star formation by Mapelli (2017) indicated that bulk rotation should be common for clusters at ages of 1–2 Myr due to large-scale torques imparted on the gas as cluster assembly occurs. Lee & Hennebelle (2016) also highlighted the prominence of cluster rotation, which in their simulations results naturally from conservation of angular momentum during the global

collapse. Here, rotation accounts for approximately one-third of the total kinetic energy, and together with turbulence acts to counteract gravity to keep the cluster globally virialized. In both the Mapelli and Lee & Hennebelle models, rotation signatures should be stronger than expansion signatures. We do not find this result here, presenting an important constraint on the physical processes of cluster assembly.

8.3. Expanding Associations: Unbound or Bound

The clear signs of expansion in many of the systems in our sample fit with the view that most star formation in massive star-forming complexes yields unbound systems, as described by Gouliermis (2018 and references therein). This view that many of these systems are not gravitationally bound is corroborated by the comparison of virial velocity dispersion to observed velocity dispersion (Figure 17), which indicates that many of the systems are highly supervirial. In a few cases, specifically NGC 6530 and Cep B, the difference between the observed system mass and the mass needed to bind the system is sufficiently large (Section 7.2) to definitively show that the systems are dispersing.

The expansion of stellar systems is correlated with gas loss in our sample of star-forming regions (Section 7). This correlation is consistent with a picture of system expansion as a reaction to changes in the gravitational potential due to the dispersal of the molecular cloud. However, it is also plausible that cloud dispersal and system expansion are both evolutionary processes that happen around the same time in star-forming regions but are independent of each other.

Observing the expansion of a stellar system on the order of $\sim 0.5 \text{ km s}^{-1}$ is insufficient evidence to demonstrate that a system is unbound. In the simulated cluster of Sills et al. (2018), the median v_{out} (Figure 19) increases after initial contraction, reaching an expansion speed of $\sim 0.5 \text{ km s}^{-1}$, similar to the typical value observed for our expanding systems (Figures 6 and 11). Velocity dispersion is also near maximum at this time, with a value of 2.5 km s^{-1} . We also note that at the point of maximum expansion velocity in the simulation at ~ 2.5 Myr, the expansion velocity increases with radial distance from the system center (Figure 19, left), similar to the trends seen in Figure 7 for NGC 6530 and Cep B. However, a difference between the cluster expansion in the simulation and the pattern of expansion seen in NGC 6530 and Cep B is that,

for the expanding associations, velocity dispersion increases with radius (Figure 13) while velocity dispersion decreases with radius in the simulation of the bound cluster (Figure 19, right).

Apart from clear cases like NGC 6530 and Cep B, other expanding systems still lack sufficient information to definitively classify them as bound or unbound. To clearly distinguish between “bound clusters” and “unbound associations,” we will require better constraints on the total mass of stars and gas in star-forming regions. In addition, some systems may have escaping stars, while a cluster core remains bound—NGC 6611 in the Eagle Nebula is discussed below (Section 8.4) as a possible example of this.

Considering slightly older systems, the conditions of formation of large OB associations like Scorpius–Centaurus (Sco-Cen) have been uncertain, with debate about whether these were produced by the expansion of an association or widely distributed star formation events (Wright et al. 2016; Ward & Kruijssen 2018; Wright & Mamajek 2018). The systems in our sample are currently fairly compact, with sizes of several parsecs, but their kinematic properties mean that they will inevitably grow to sizes larger than 100 pc across.

NGC 6530 may be a good analog of a precursor to a massive OB association. The bulk expansion velocity of NGC 6530 is 0.9 km s^{-1} , but its velocity dispersion is 2.2 km s^{-1} . The mass of the system compared to its virial mass is sufficiently low that self-gravity will have little effect on the evolution of stellar velocities, and the association will expand ballistically. In 10 Myr, a star traveling at 2.2 km s^{-1} will travel 22 pc while a star with a velocity of 4.5 km s^{-1} (2σ from the mean) would travel 45 pc, so that $\sim 95\%$ of the stars would be found within a region with a length of ~ 90 pc. This size is quite similar to the size of Upper Scorpius in the Sco-Cen association, which is ~ 75 – 100 pc long along its longest axis (Galli et al. 2018) and has an age of ~ 10 Myr (Pecaut & Mamajek 2016). The estimated mass for NGC 6530 of $\sim 4000 M_{\odot}$ is somewhat higher than $1400 M_{\odot}$ for Upper Scorpius (Preibisch & Mamajek 2008).

Several of the expanding systems exhibit significant substructure in the spatial distribution of their stars (see Kuhn et al. 2014, their Figure 5). In NGC 6530, stars are not smoothly distributed but instead are clumped. In NGC 1893, the system is made up of a chain of subclusters, which likely traces the shape of the molecular filament from before gas was expelled from the system. If the expansion of unbound stellar systems is truly “Hubble-like,” this substructure would get expanded to larger sizes, as seen in large OB associations like Upper Scorpius. On the other hand, if spatial substructure is not reflected in the kinematics, then spatial distributions of stars would tend to get smoother as stars drift farther from their point of origin. The preservation of substructure during the dynamical evolution of a young stellar system is a result that has been predicted for supervirial systems (Section 8.1).

8.4. Bound Clusters

Our sample includes a few systems that are likely to survive (temporarily) as bound open clusters (Section 8.4). These systems are not currently expanding at a significant rate and have total energies near the division between bound and unbound systems. The ONC is still quite compact, while NGC 6231 and NGC 2362 have probably already expanded. Despite the young ages of these systems, which are all less than their

virialization timescales for two-body interactions, they all show properties that would be expected for virialized systems. They are relatively well fit by cluster models, and the cluster density, cluster core radius, and velocity distributions are consistent with theoretical expectations given the uncertainties. Velocity dispersions in these clusters are either constant with radius (expected for an isothermal structure) or decreasing.

In our sample, the ONC provides the best case for modeling velocity dispersions. For an isothermal sphere with equal-mass stars, the distribution of stellar velocities will be Maxwellian (i.e., a Gaussian distribution). Due to a number of assumptions that are violated in real young stellar systems, there is no reason to expect a priori that this distribution should accurately describe stellar velocities. Nevertheless, it turns out that the velocity distribution in the ONC is consistent with Gaussianity (Figure 10), but the distribution is elongated north–south (Figure 9). The elongation is parallel to the orientation of the Orion A cloud, so it could result from perturbations of the gravitational potential from asymmetry in the star-forming complex, or it could be a residual effect of the cluster formation from a molecular filament (e.g., Proszkow & Adams 2009; Kuznetsova et al. 2015).

The properties of probable bound clusters like the ONC, NGC 6231, and NGC 2362 can be compared to the simulated cluster from Sills et al. (2018) to test the accuracy of its predictions. For several time points from the simulation, Figure 20 shows stellar velocity v_{out} as a function of projected radius R . For each evolutionary stage shown, velocity dispersions are higher in the cluster center and lower in the cluster periphery. This is also clearly seen in Figure 19, where velocity dispersions are strongly affected by distance from the center of the cluster. This pattern is similar to the radial dependence of velocity dispersion found in NGC 6231 (Figure 13), but is different from the isothermal distributions in the ONC and NGC 2362. Future simulations can be used to test which initial conditions can produce the different types of the observed velocity profile.

Bound systems seem to be in the minority in our sample based on the evidence from Figure 17. Nevertheless, as discussed above, it is possible for a system with positive total energy to leave behind a bound core after most of the stars are lost. Given that many expanding systems show a radial gradient in expansion velocity, stars near the center might be less likely to be escaping from the system. A possible example could be NGC 6611, which is composed of a dense central cluster surrounded by lower-density groups of stars. Although the median expansion velocity of the system is $\sim 1 \text{ km s}^{-1}$, the expansion velocity as measured within the inner 2 pc is lower at $\sim 0.5 \text{ km s}^{-1}$ (Figure 7). Simulations of these systems may be helpful for determining whether a bound cluster is likely to remain in such cases.

9. Conclusions

The superb astrometric measurements of the *Gaia* spacecraft provided in DR2 have allowed us to examine the kinematics of nearby young star clusters and associations including many of the most massive star-forming complexes in our neighborhood of the Galaxy. The study also makes use of large samples of YSOs in these regions from the MYStIX and SFiNCs projects. Results from our sample of 28 systems likely represent the environments in which most star formation takes place in mature spiral galaxies.

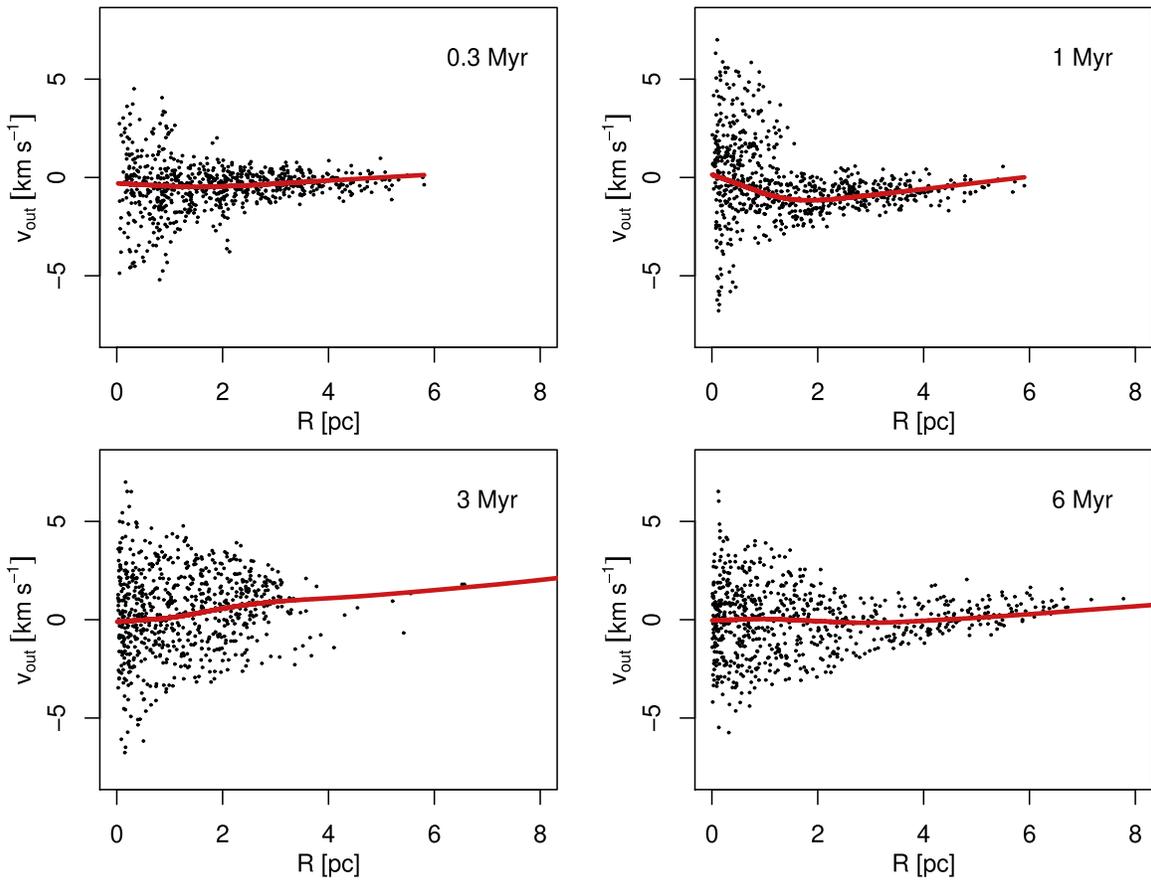


Figure 20. Snapshots of the outward velocity v_{out} as a function of distance R from the cluster center in the Sills et al. (2018) simulation at 0.3, 1.0, 3.0, and 6.0 Myr. Only stars with $M > 1 M_{\odot}$ are shown for ease of plotting.

The main scientific results of this study are as follows:

1. Bulk expansion is commonly seen for young stellar systems during the first few million years. At least 75% of the systems in the sample have positive expansion velocities (likely 85%–90%), and expansion velocities range up to 2 km s^{-1} , with a median value of $\sim 0.5 \text{ km s}^{-1}$. Significant expansion around 1 km s^{-1} is measured for NGC 6530 in the Lagoon Nebula, Cep B in the Cep OB3b association, Tr 16 in the Carina Nebula, NGC 2244 in the Rosette Nebula, NGC 6611 in the Eagle Nebula, and NGC 1893.
2. The most rapidly expanding systems have positive total energies, so their expansion can be explained as cluster dispersal. Velocity dispersions are sufficiently large for these associations to expand to the size of well-known OB associations like Sco-Cen in $\sim 10 \text{ Myr}$.
3. A positive radial gradient in expansion velocity is seen in some expanding systems like NGC 6530, Cep B, NGC 2244, and Tr 16. Radial gradients in velocity can result from faster stars traveling larger distances, which yields positional sorting of stars by velocity. This phenomenon would be most prominent in systems where stellar trajectories are affected little by the gravitation of the cluster, so detection of this gradient provides evidence that these regions are unbound associations. In contrast, in a simulation of expansion in a gravitationally bound cluster, no radial dependence is seen.
4. Several systems appear likely to be gravitationally bound. Notable in this group is the still-compact ONC and the larger clusters NGC 6231 and NGC 2362 that may have expanded in the past. A mild expansion velocity is measured for the ONC, and there is no evidence for the expansion of NGC 6231 or NGC 2362. All three have velocity dispersions consistent with virial equilibrium; the ONC and NGC 2362 have isothermal velocity distributions while velocity dispersion decreases with radius in NGC 6231. The velocity dispersion in the ONC is approximately Gaussian. These findings are consistent with predictions of bound cluster models, and we expect that these clusters will emerge as gravitationally bound, main-sequence, open clusters.
5. Stellar systems that are no longer embedded in their natal molecular clouds or only partially embedded are statistically more likely to be in a state of expansion than systems that are still embedded ($p < 0.001$). This result is consistent with expansion as a consequence of gas expulsion, but it is also possible that cloud dispersal and expansion of stellar systems occur simultaneously but independently.
6. Among the more embedded clusters, NGC 1333 and IC 348 in the Perseus cloud do not show signs of expansion or contraction. The embedded cluster M17 is unique in showing evidence (2σ significance) of contraction with velocity -2 km s^{-1} . This is consistent with its clumpy

stellar structure and suggests it is still in the process of assembly.

7. There is no evidence for rotation in all but one case, Tr 15. Theoretical simulations indicate that rotation is expected to be present in clusters that inherit angular momentum from the merging of large subclusters, so its absence provides constraints on cluster formation. More sensitive measurements are required to further constrain the rotation of young clusters.
8. In our sample of young stellar systems, one-dimensional velocity dispersions, σ_{1D} , range from 1 to 3 km s⁻¹. The velocity dispersion for the ONC (1.8 ± 0.1 km s⁻¹) is slightly lower than most previous estimates based on RV studies. In the full cluster sample, velocity dispersions are typically greater than bulk expansion velocities by a factor of $\sim 2-3$.
9. The relative motions of clusters within massive star-forming regions generally show random motions, likely inherited from the parent molecular clouds. They do not generally have convergent motions expected from hierarchical assembly, indicating that any cluster merging occurred during an embedded phase before the clusters were observed.

While indirect evidence of cluster dispersal had been claimed previously (e.g., Pfalzner 2009), the *Gaia* observations provide direct evidence that stars produced in compact massive star-forming regions are more likely to immediately disperse after gas expulsion. Only a few pre-main-sequence systems are likely to produce gravitationally bound, main-sequence, open clusters. Future *Gaia* data releases are anticipated to dramatically improve astrometry for fainter sources, which will allow kinematic studies of stellar populations with higher absorption than the ones included in this study. This will improve constraints on processes of cluster assembly and help distinguish which star formation environments produce stellar systems of different dynamical fates.

We thank Alberto Krone-Martins and Lennart Lindegren for expert advice on *Gaia*, Floor van Leeuwen for advice about astrometry, and Fred Adams, Michael Grudić, and Anushree

Sengupta for useful feedback on the article. We would also like to thank the referee for helpful comments. E.D.F. and K.V.G. are supported by NASA grant NNX15AF42G and *Chandra* grants AR7-18002X, GO7-18001X, and G06-17015x, and the *Chandra* ACIS Team contract SV474018 (G. Garmire & L. Townsley, principal investigators) issued by the *Chandra* X-ray Center under contract NAS8-03060.

Facility: *Gaia*.

Software: astro (Kelvin 2014), celestial (Robotham 2016), MVN (Korkmaz et al. 2014), R (R Core Team 2018), relaimpo (Grömping 2006), spatstat (Baddeley & Turner 2005), TOPCAT (Taylor 2005).

Appendix A Distance to Orion

The Orion region is sufficiently close that distance measurements differ slightly depending on the location in the cloud. Based on our weighted-median method described in the body of the paper, we find a median parallax of $\varpi = 2.482 \pm 0.041$ for the X-ray-selected probable cluster members in the ONC *Chandra* field, corresponding to a distance of 403 pc. This field includes the dense central cluster, also known as the Trapezium Cluster, but excludes stars associated with Orion A that lie more than 1.5 pc from the center. Kounkel et al. (2018) reported a *Gaia*-derived distance of 386 ± 3 pc for the ONC, which is based on a much larger region spanning $\delta = -7^\circ$ to -4° in decl. Their analysis also differs in that they apply a correction to the *Gaia* parallaxes to account for systematic differences between astrometry from *Gaia* and the VLBA based on Kounkel et al. (2017). This correction shifts objects at the distance of the ONC nearer by ~ 10 pc. In our work, we have not applied this correction.

Figure 21 shows *Gaia* parallax measurements as a function of decl. in the Orion A cloud. We fit a non-parametric loess curve (Cleveland 1979) and compute 95% confidence intervals, which reveals some variation in mean parallax for stars in different parts of the Orion A complex. The ONC itself is recessed (403 pc), while the stars to the north and south (including populations of stars both embedded in the filament and outside it) are nearer (395 pc). Given that this study is

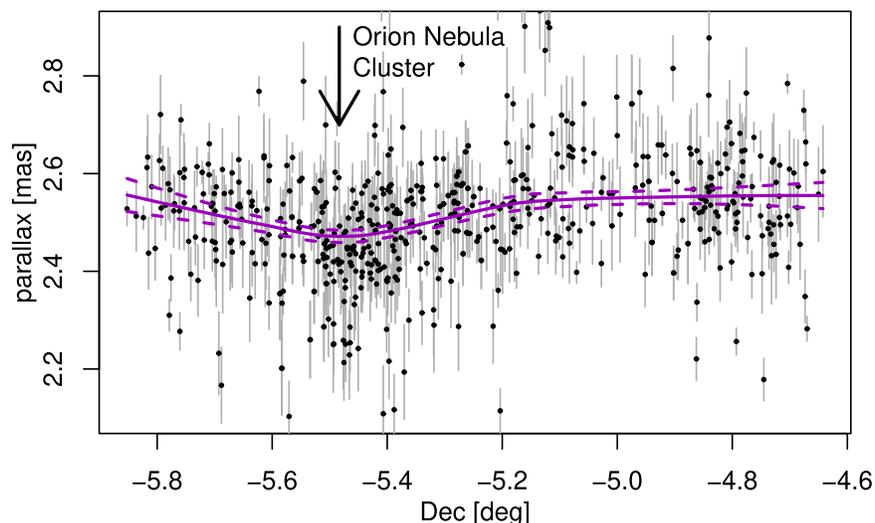


Figure 21. Parallax measurements for probable cluster members in the Orion A clouds, including the ONC. Only points with parallax uncertainties < 0.1 mas are included. The magenta lines show a non-parametric regression line and 95% confidence interval found using the loess regression in the R programming language with options `span = 0.6`, `degree = 1`, `family = "symmetric,"` `iterations = 4`, and `surface = "direct."`

focused on the kinematics of the main cluster, we use the 403 pc distance. However, the discrepancy between our value and the distance from Kounkel et al. (2018) can be explained by a combination of the different sizes of the region analyzed and whether additional correction factors are applied to *Gaia* astrometry.

Appendix B Proper-motion Corrections due to RVs

Calculating the contribution of perspective expansion to proper motions in Equations (1) and (2) requires measurements of the RVs of clusters. We have compiled a list of RVs for the clusters (Table 5) based on both previously published RV measurements and RV measurements made by the *Gaia* spacecraft (Cropper et al. 2018). There are a few cases of discrepancies in RVs from the literature and/or median RVs from *Gaia*. In order to assign RV measurements to all clusters so that proper-motion corrections can be computed, we favor (1) the median *Gaia* RV when at least five stars are available, (2) the most recent literature RV, and finally (3) the *Gaia* RVs

Table 5
RVs of Clusters

Region	Sample	RV _{<i>Gaia</i>} (km s ⁻¹)	RV _{lit} (km s ⁻¹)
Berkeley 59	0	...	-14.5 (a), -6.50 (b)
NGC 1333	0	...	8.02 (c)
IC 348	8	19.7 ± 2.2	14.00 (b), 15.37 (d)
LkHα 101	2	5.74 ± 3	11-13 (e)*
NGC 1893	0	...	-2.25 (b)
Orion A Cloud			
— ONC	12	21.8 ± 6.6	26 (f), 28.94 (b)
— OMC 2-3	9	22.6 ± 5.5	
— Orion Flank S	5	25.2 ± 4.8	
— Orion Flank N	7	27.4 ± 3.3	
Mon R2	1	35.6 ± 3	28 (g)
NGC 2244	1	19.1 ± 0.96	26.16 (b)
NGC 2264	7	15.8 ± 2.9	22 (h), 17.68 (b)
NGC 2362	1	28.9 ± 5.7	25.33 (b)
Carina			
— Tr 14	0	...	-6.7 (i), -15.0 (b)
— Tr 15	0
— Tr 16	0	...	-25.00 (b)
NGC 6231	1	-19.1 ± 1	-27.28 (b)
RCW 120	1	13 ± 4.1	-16 (j) *
NGC 6357	0	...	-12 (k)*
M20	0	...	-2.13 (b)
NGC 6530	0	...	-13.32 (b)
NGC 6611	3	16:	18.00 (b), 10 (l)
M17	0	...	-17.00 (b)
IC 5146	1	4.59 ± 10	-5.5 (m), ~5 (n)
NGC 7160	1	-28 ± 8.1	-23.8 (m), -26.28 (b)
Cep OB3b	4	-28.7 ± 4.4	-22 (o), -15.12 (b)

Note. Column 1: region name. Column 2: number of members with *Gaia* RV measurements. Column 3: median RV (barycentric). Column 4: RV measurements from the literature (barycentric). Measurements for cloud gas are marked with an asterisk.

RV References. (a) Liu et al. (1989), (b) Kharchenko et al. (2005), (c) Foster et al. (2015), (d) Cottaar et al. (2015), (e) Dewdney & Roger (1982), (f) Fűrész et al. (2008), (g) Kim et al. (2004), (h) Fűrész et al. (2006), (i) Dias et al. (2002), (j) Blitz et al. (1982), (k) Caswell & Haynes (1987), (l) Evans et al. (2005), (m) Liu et al. (1989), (n) Liu et al. (1991), (o) Pozzo et al. (2003).

based on one to four stars. For Tr 15, which lacks its own independent RV measurement, we assign it an RV of -20 km s⁻¹ based on the motion of the Carina Nebula as a whole.

ORCID iDs

Michael A. Kuhn  <https://orcid.org/0000-0002-0631-7514>
 Alison Sills  <https://orcid.org/0000-0003-3551-5090>
 Eric D. Feigelson  <https://orcid.org/0000-0002-5077-6734>
 Konstantin V. Getman  <https://orcid.org/0000-0002-6137-8280>

References

- Aarseth, S. J., & Hills, J. G. 1972, *A&A*, **21**, 255
 Adams, F. C. 2000, *ApJ*, **542**, 964
 Adams, F. C. 2010, *ARA&A*, **48**, 47
 Allison, R. J. 2012, *MNRAS*, **421**, 3338
 Baddeley, A., Rubak, E., & Turner, R. 2015, *Spatial Point Patterns: Methodology and Applications with R* (Boca Raton, FL: CRC Press)
 Baddeley, A., & Turner, R. 2005, *Journal of Statistical Software*, **12**, 1
 Bailer-Jones, C. A. L., Rybizki, J., Fousneau, M., Mantelet, G., & Andrae, R. 2018, *AJ*, **156**, 58
 Banerjee, S., & Kroupa, P. 2015, *MNRAS*, **447**, 728
 Banerjee, S., & Kroupa, P. 2017, *A&A*, **597**, A28
 Banerjee, S., & Kroupa, P. 2018, in *The Birth of Star Clusters*, Vol. 424, ed. S. Stahler (Cham: Springer), 143
 Bastian, N., & Goodwin, S. P. 2006, *MNRAS*, **369**, L9
 Bianchini, P., van de Ven, G., Norris, M. A., Schinnerer, E., & Varri, A. L. 2016, *MNRAS*, **458**, 3644
 Bianchini, P., van der Marel, R. P., del Pino, A., et al. 2018, *MNRAS*, **481**, 2125
 Binney, J., & Tremaine, S. 2008, *Galactic Dynamics* (2nd ed.; Princeton, NJ: Princeton Univ. Press)
 Blaauw, A. 1964, *ARA&A*, **2**, 213
 Blitz, L., Fich, M., & Stark, A. A. 1982, *ApJS*, **49**, 183
 Bonnell, I. A., Bate, M. R., & Vine, S. G. 2003, *MNRAS*, **343**, 413
 Bovy, J., Hogg, D. W., & Roweis, S. T. 2011, *AnApS*, **5**, 1657
 Bravi, L., Zari, E., Sacco, G. G., et al. 2018, *A&A*, **615**, A37
 Bressan, A., Marigo, P., Girardi, L., et al. 2012, *MNRAS*, **427**, 127
 Bressert, E., Bastian, N., Gutermuth, R., et al. 2010, *MNRAS*, **409**, L54
 Brinkmann, N., Banerjee, S., Motwani, B., & Kroupa, P. 2017, *A&A*, **600**, A49
 Broos, P. S., Getman, K. V., Povich, M. S., et al. 2013, *ApJS*, **209**, 32
 Brown, A. G. A., Dekker, G., & de Zeeuw, P. T. 1997, *MNRAS*, **285**, 479
 Cambrésy, L., Rho, J., Marshall, D. J., & Reach, W. T. 2011, *A&A*, **527**, A141
 Caswell, J. L., & Haynes, R. F. 1987, *A&A*, **171**, 261
 Clarke, C. J., Mathieu, R. D., & Reid, I. N. 2015, in *Dynamics of Young Star Clusters and Associations*, ed. L. Eyer & C. P. Bell (1st ed.; Berlin: Springer)
 Cleveland, W. S. 1979, *J. Am. Stat. Assoc.*, **74**, 829
 Cottaar, M., Covey, K. R., Foster, J. B., et al. 2015, *ApJ*, **807**, 27
 Cropper, M., Katz, D., Sartoretti, P., et al. 2018, *A&A*, **616**, A5
 Da Rio, N., Tan, J. C., Covey, K. R., et al. 2017, *ApJ*, **845**, 105
 Da Rio, N., Tan, J. C., & Jaehnig, K. 2014, *ApJ*, **795**, 55
 Dale, J. E., Ercolano, B., & Bonnell, I. A. 2015, *MNRAS*, **451**, 987
 Damiani, F., Micela, G., & Sciortino, S. 2016, *A&A*, **596**, A82
 Dewdney, P. E., & Roger, R. S. 1982, *ApJ*, **255**, 564
 Dias, W. S., Alessi, B. S., Moitinho, A., & Lépine, J. R. D. 2002, *A&A*, **389**, 871
 Duchêne, G., Lacour, S., Moraux, E., Goodwin, S., & Bouvier, J. 2018, *MNRAS*, **478**, 1825
 Dzib, S. A., Loinard, L., Rodríguez, L. F., et al. 2017, *ApJ*, **834**, 139
 Elmegreen, B. G. 1983, *MNRAS*, **203**, 1011
 Elmegreen, B. G. 2000, *ApJ*, **530**, 277
 Evans, C. J., Smartt, S. J., Lee, J.-K., et al. 2005, *A&A*, **437**, 467
 Fariás, J. P., Fellhauer, M., Smith, R., Domínguez, R., & Dabringhausen, J. 2018, *MNRAS*, **476**, 5341
 Feigelson, E. D. 2018, in *The Birth of Star Clusters*, Vol. 424, ed. S. Stahler (Cham: Springer), 119
 Feigelson, E. D., Townsley, L. K., Broos, P. S., et al. 2013, *ApJS*, **209**, 26
 Fellhauer, M., Wilkinson, M. I., & Kroupa, P. 2009, *MNRAS*, **397**, 954

- Fletcher, R., & Reeves, C. M. 1964, *CompJ*, 7, 149
- Foster, J. B., Cottaar, M., Covey, K. R., et al. 2015, *ApJ*, 799, 136
- Fukui, Y., Ohama, A., Hanaoka, N., et al. 2014, *ApJ*, 780, 36
- Fűrész, G., Hartmann, L. W., Megeath, S. T., Szentgyorgyi, A. H., & Hamden, E. T. 2008, *ApJ*, 676, 1109
- Fűrész, G., Hartmann, L. W., Szentgyorgyi, A. H., et al. 2006, *ApJ*, 648, 1090
- Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018a, *A&A*, 616, A1
- Gaia Collaboration, Helmi, A., van Leeuwen, F., et al. 2018b, *A&A*, 616, A12
- Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., et al. 2016, *A&A*, 595, A1
- Galli, P. A. B., Joncour, I., & Moraux, E. 2018, *MNRAS*, 477, L50
- Getman, K. V., Broos, P. S., Kuhn, M. A., et al. 2017, *ApJS*, 229, 28
- Getman, K. V., Feigelson, E. D., Kuhn, M. A., et al. 2014, *ApJ*, 787, 108
- Getman, K. V., Kuhn, M. A., Feigelson, E. D., et al. 2018, *MNRAS*, 477, 298
- Gieles, M., Moeckel, N., & Clarke, C. J. 2012, *MNRAS*, 426, L11
- Goodwin, S. P. 2009, *Ap&SS*, 324, 259
- Goodwin, S. P., & Bastian, N. 2006, *MNRAS*, 373, 752
- Gouliermis, D. A. 2018, *PASP*, 130, 072001
- Grömping, U. 2006, *Journal of Statistical Software*, 17, 1
- Grudić, M. Y., Hopkins, P. F., Faucher-Giguère, C.-A., et al. 2018, *MNRAS*, 475, 3511
- Henney, W. J., & Arthur, S. J. 1998, *AJ*, 116, 322
- Henze, N., & Zirkler, B. 1990, *Communications in Statistics-Theory and Methods*, 19, 3595
- Hillenbrand, L. A., & Hartmann, L. W. 1998, *ApJ*, 492, 540
- Hills, J. G. 1980, *ApJ*, 235, 986
- Hirota, T., Bushimata, T., Choi, Y. K., et al. 2008, *PASJ*, 60, 37
- Hollander, M., & Wolfe, D. 1973, *Nonparametric Statistical Methods* (New York: Wiley)
- Ibáñez-Mejía, J. C., Mac Low, M.-M., Klessen, R. S., & Baczynski, C. 2017, *ApJ*, 850, 62
- Jones, B. F., & Walker, M. F. 1988, *AJ*, 95, 1755
- Kamann, S., Husser, T. O., Dreizler, S., et al. 2018, *MNRAS*, 473, 5591
- Kelly, B. C., Fan, X., & Vestergaard, M. 2008, *ApJ*, 682, 874
- Kelvin, L. 2014, astro: Astronomy Functions, Tools and Routines, 1.2, CRAN, <https://CRAN.R-project.org/package=astro>
- Kharchenko, N. V., Piskunov, A. E., Röser, S., Schilbach, E., & Scholz, R.-D. 2005, *A&A*, 438, 1163
- Kim, B. G., Kawamura, A., Yonekura, Y., & Fukui, Y. 2004, *PASJ*, 56, 313
- King, I. R. 1966, *AJ*, 71, 64
- Korkmaz, S., Goksuluk, D., & Zararsiz, G. 2014, *The R Journal*, 6, 151
- Kounkel, M., Covey, K., Suárez, G., et al. 2018, *AJ*, 156, 84
- Kounkel, M., Hartmann, L., Loinard, L., et al. 2017, *ApJ*, 834, 142
- Kounkel, M., Hartmann, L., Tobin, J. J., et al. 2016, *ApJ*, 821, 8
- Kroupa, P., Aarseth, S., & Hurley, J. 2001, *MNRAS*, 321, 699
- Kruijssen, J. M. D. 2012, *MNRAS*, 426, 3008
- Krumholz, M. R., & McKee, C. F. 2005, *ApJ*, 630, 250
- Kuhn, M. A., & Feigelson, E. D. 2017, arXiv:1711.11101
- Kuhn, M. A., Feigelson, E. D., Getman, K. V., et al. 2014, *ApJ*, 787, 107
- Kuhn, M. A., Feigelson, E. D., Getman, K. V., et al. 2015a, *ApJ*, 812, 131
- Kuhn, M. A., Getman, K. V., & Feigelson, E. D. 2015b, *ApJ*, 802, 60
- Kuhn, M. A., Getman, K. V., Feigelson, E. D., et al. 2017a, *AJ*, 154, 214
- Kuhn, M. A., Medina, N., Getman, K. V., et al. 2017b, *AJ*, 154, 87
- Kuznetsova, A., Hartmann, L., & Ballesteros-Paredes, J. 2015, *ApJ*, 815, 27
- Lada, E. A., & Lada, C. J. 1995, *AJ*, 109, 1682
- Larson, R. B. 1981, *MNRAS*, 194, 809
- Lee, Y.-N., & Hennebelle, P. 2016, *A&A*, 591, A30
- Lindgren, L., Hernández, J., Bombrun, A., et al. 2018, *A&A*, 616, A2
- Lindgren, L., Lammers, U., Bastian, U., et al. 2016, *A&A*, 595, A4
- Lindgren, L., Lammers, U., Hobbs, D., et al. 2012, *A&A*, 538, A78
- Liu, T., Janes, K. A., & Bania, T. M. 1989, *AJ*, 98, 626
- Liu, T., Janes, K. A., & Bania, T. M. 1991, *AJ*, 102, 1103
- Luhman, K. L., Esplin, T. L., & Loutrel, N. P. 2016, *ApJ*, 827, 52
- Mac Low, M.-M., & Klessen, R. S. 2004, *RvMP*, 76, 125
- Mapelli, M. 2017, *MNRAS*, 467, 3255
- Marks, M., & Kroupa, P. 2012, *A&A*, 543, A8
- Maschberger, T. 2013, *MNRAS*, 429, 1725
- Maschberger, T., Clarke, C. J., Bonnell, I. A., & Kroupa, P. 2010, *MNRAS*, 404, 1061
- Mathieu, R. D. 1983, *ApJL*, 267, L97
- McLachlan, G., & Peel, D. 2000, *Finite Mixture Models* (New York: Wiley)
- Megeath, S. T., Gutermuth, R., Muzerolle, J., et al. 2016, *AJ*, 151, 5
- Menten, K. M., Reid, M. J., Forbrich, J., & Brunthaler, A. 2007, *A&A*, 474, 515
- Moraux, E. 2016, in *EAS Publ. Ser. 80, Stellar Clusters: Benchmarks of Stellar Physics and Galactic Evolution*, ed. E. Moraux, Y. Lebreton, & C. Charbonnel (Les Ulis: EDP Sciences), 73
- Muench, A., Getman, K., Hillenbrand, L., & Preibisch, T. 2008, in *Handbook of Star Forming Regions, Volume I: The Northern Sky, Vol. 4*, ed. B. Reipurth (San Francisco, CA: ASP), 483
- Muraveva, T., Delgado, H. E., Clementini, G., Sarro, L. M., & Garofalo, A. 2018, *MNRAS*, 481, 1195
- Naylor, T., Broos, P. S., & Feigelson, E. D. 2013, *ApJS*, 209, 30
- O'dell, C. R. 2001, *ARA&A*, 39, 99
- Padoan, P., & Nordlund, Å. 2002, *ApJ*, 576, 870
- Parker, R. J., Goodwin, S. P., Wright, N. J., Meyer, M. R., & Quanz, S. P. 2016, *MNRAS*, 459, L119
- 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
- Pecaut, M. J., & Mamajek, E. E. 2016, *MNRAS*, 461, 794
- Pelupessy, F. I., & Portegies Zwart, S. 2012, *MNRAS*, 420, 1503
- Pfalzner, S. 2009, *A&A*, 498, L37
- Pfalzner, S. 2011, *A&A*, 536, A90
- Pfalzner, S., & Kaczmarek, T. 2013, *A&A*, 559, A38
- Pfalzner, S., Parmentier, G., Steinhausen, M., Vincke, K., & Menten, K. 2014, *ApJ*, 794, 147
- Portegies Zwart, S. F., McMillan, S. L. W., & Gieles, M. 2010, *ARA&A*, 48, 431
- Povich, M. S., Kuhn, M. A., Getman, K. V., et al. 2013, *ApJS*, 209, 31
- Pozzo, M., Naylor, T., Jeffries, R. D., & Drew, J. E. 2003, *MNRAS*, 341, 805
- Preibisch, T., & Mamajek, E. 2008, in *Handbook of Star Forming Regions, Volume II: The Southern Sky, Vol. 5*, ed. B. Reipurth (San Francisco, CA: ASP), 235
- Proszkow, E.-M., & Adams, F. C. 2009, *ApJS*, 185, 486
- Proszkow, E.-M., Adams, F. C., Hartmann, L. W., & Tobin, J. J. 2009, *ApJ*, 697, 1020
- R Core Team 2018, *R: A Language and Environment for Statistical Computing* (Vienna: R Foundation for Statistical Computing)
- Reino, S., de Bruijne, J., Zari, E., d'Antona, F., & Ventura, P. 2018, *MNRAS*, 477, 3197
- Richert, A. J. W., Getman, K. V., Feigelson, E. D., et al. 2018, *MNRAS*, 477, 5191
- Rieke, G. H., & Lebofsky, M. J. 1985, *ApJ*, 288, 618
- Rivilla, V. M., Martín-Pintado, J., Jiménez-Serra, I., & Rodríguez-Franco, A. 2013, *A&A*, 554, A48
- Robotham, A. S. G. 2016, *Celestial: Common astronomical conversion routines and functions*, Astrophysics Source Code Library, ascl:1602.011
- Schwarz, G. 1978, *AnSta*, 6, 461
- Shukirgaliyev, B., Parmentier, G., Just, A., & Berczik, P. 2018, *ApJ*, 863, 171
- Sills, A., Rieder, S., Scora, J., McCloskey, J., & Jaffa, S. 2018, *MNRAS*, 477, 1903
- Spera, M., Mapelli, M., & Jeffries, R. D. 2016, *MNRAS*, 460, 317
- Spitzer, L., Jr. 1969, *ApJL*, 158, L139
- Stassun, K. G., & Torres, G. 2018, *ApJ*, 862, 61
- Taylor, M. B. 2005, in *ASP Conf. Ser. 347, Astronomical Data Analysis Software and Systems XIV*, ed. P. Shopbell, M. Britton, & R. Ebert (San Francisco, CA: ASP), 29
- Tobin, J. J., Hartmann, L., Fűrész, G., Mateo, M., & Megeath, S. T. 2009, *ApJ*, 697, 1103
- Tutukov, A. V. 1978, *A&A*, 70, 57
- van Leeuwen, F. 2009, *A&A*, 497, 209
- Vázquez-Semadeni, E., González-Samaniego, A., & Colín, P. 2017, *MNRAS*, 467, 1313
- Walborn, N. R. 1995, *RMxAC*, 2, 51
- Walker, M. G., Mateo, M., Olszewski, E. W., et al. 2006, *AJ*, 131, 2114
- Ward, J. L., & Kruijssen, J. M. D. 2018, *MNRAS*, 475, 5659
- Webb, J. J., & Vesperini, E. 2017, *MNRAS*, 464, 1977
- Wright, N. J., Bouy, H., Drew, J. E., et al. 2016, *MNRAS*, 460, 2593
- Wright, N. J., & Mamajek, E. E. 2018, *MNRAS*, 476, 381